

Wilhelm Kleppmann

Praxisreihe

Qualität



Versuchs- planung

Produkte und
Prozesse optimieren



10., überarbeitete Auflage

HANSER

Kleppmann
Versuchsplanung



Bleiben Sie auf dem Laufenden!

Hanser Newsletter informieren Sie regelmäßig über neue Bücher und Termine aus den verschiedenen Bereichen der Technik. Profitieren Sie auch von Gewinnspielen und exklusiven Leseproben. Gleich anmelden unter

www.hanser-fachbuch.de/newsletter

Herausgeber der Praxisreihe Qualität (vormals Praxisreihe Qualitätswissen):
von 1991 (Gründungsjahr) bis 2016 Franz J. Brunner; seit 2016 Kurt Matyas.

In der Praxisreihe Qualität sind bereits erschienen:

Jörg Brenner

Lean Production

Praktische Umsetzung zur Erhöhung der Wertschöpfung
3., überarbeitete Auflage
ISBN 978-3-446-45664-8

Jörg Brenner

Lean Administration

Verschwendungen erkennen, analysieren, beseitigen
ISBN 978-3-446-45472-9

Franz J. Brunner

Japanische Erfolgskonzepte

Kaizen, KVP, Lean Production Management, Total Productive Maintenance, Shopfloor Management, Toyota Production Management, GD3 – Lean Development
4., überarbeitete Auflage
ISBN 978-3-446-45428-6

Franz J. Brunner

Qualität im Service

Wege zur besseren Dienstleistung
ISBN 978-3-446-42241-4

Franz J. Brunner, Karl W. Wagner

Mitarbeit: Peter H. Osanna, Kurt Matyas, Peter Kuhlang

Qualitätsmanagement

Leitfaden für Studium und Praxis
6., überarbeitete Auflage
ISBN 978-3-446-44712-7

Werner Friedrichs

Das Fitnessprogramm für KMU

Methoden für mehr Effizienz im Automobil-, Anlagen- und Sondermaschinenbau
ISBN 978-3-446-45341-8

Werner Friedrichs

Ressourcenmanagement in KMU

ISBN 978-3-446-45766-9

Menderes Güneş, Marwan Hamdan, Mirko Klug

Gewährleistungsmanagement

ISBN 978-3-446-44795-0

Marco Einhaus, Florian Lugauer, Christina Häußinger

Arbeitsschutz und Sicherheitstechnik

Der Schnelleinstieg für (angehende) Führungskräfte:
Basiswissen, Haftung, Gefährdungen, Rechtslage
ISBN 978-3-446-45474-3

René Kiem

Qualität 4.0

QM, MES und CAQ in digitalen Geschäftsprozessen der Industrie 4.0
ISBN 978-3-446-44736-3

Jörg Brenner

Shopfloor Management und seine digitale Transformation

Die besten Werkzeuge in 45 Beispielen
ISBN 978-3-446-46000-3

Veit Kohnhauser, Markus Pollhamer

Entwicklungsqualität

ISBN 978-3-446-42796-9

Karl Koltze, Valeri Souchkov

Systematische Innovation

TRIZ-Anwendung in der Produkt- und Prozessentwicklung
2., überarbeitete Auflage
ISBN 978-3-446-45127-8

Kurt Matyas

Instandhaltungslogistik

Qualität und Produktivität steigern
7., erweiterte Auflage
ISBN 978-3-446-45762-1

Arno Meyna, Bernhard Pauli

Zuverlässigkeitstechnik

Quantitative Bewertungsverfahren
2., überarbeitete und erweiterte Auflage
ISBN 978-3-446-41966-7

Markus Schneider

Lean und Industrie 4.0

Eine Digitalisierungsstrategie mit der Wertstrommethode und Information Flow Design
ISBN: 978-3-446-45917-5

Wilfried Sihl, Alexander Sunk, Tanja Nemeth,
Peter Kuhlang, Kurt Matyas

Produktion und Qualität

Organisation, Management, Prozesse
ISBN 978-3-446-44735-6

Friedrich Peschke, Carsten Eckhardt

Flexible Produktion durch Digitalisierung

Entwicklung von Use Cases
ISBN 978-3-446-45746-1

Konrad Wälde, Olga Wälde

Statistische Methoden der Qualitäts sicherung

Praktische Anwendung mit MINITAB und JMP
ISBN 978-3-446-43217-8

Johann Wappis

Null-Fehler-Management

Umsetzung von Six Sigma
6., aktualisierte Auflage
ISBN 978-3-446-45875-8

Wilhelm Kleppmann

Versuchsplanung

Produkte und Prozesse optimieren

10., überarbeitete Auflage

Praxisreihe Qualität
Herausgegeben von Kurt Matyas

HANSER

Der Autor:

Prof. Dr. Wilhelm Kleppmann, Hochschule Aalen, Beethovenstraße 1, 73430 Aalen
wilhelm.kleppmann@hs-aalen.de

Zu diesem Buch gehört ein Download. Unter dem Reiter Extras können Sie sich hier die Dateien zum Buch herunterladen:

<https://www.hanser-fachbuch.de/buch/Versuchsplanung/9783446461468>



Alle in diesem Buch enthaltenen Informationen wurden nach bestem Wissen zusammengestellt und mit Sorgfalt geprüft und getestet. Dennoch sind Fehler nicht ganz auszuschließen. Aus diesem Grund sind die im vorliegenden Buch enthaltenen Informationen mit keiner Verpflichtung oder Garantie irgendeiner Art verbunden. Herausgeber, Autor und Verlag übernehmen infolgedessen keine Verantwortung und werden keine daraus folgende oder sonstige Haftung übernehmen, die auf irgendeine Art aus der Benutzung dieser Informationen – oder Teilen davon – entsteht.

Ebenso wenig übernehmen Herausgeber, Autor und Verlag die Gewähr dafür, dass beschriebene Verfahren usw. frei von Schutzrechten Dritter sind. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürfen.

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Buches, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder ein anderes Verfahren) – auch nicht für Zwecke der Unterrichtsgestaltung – reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

© 2020 Carl Hanser Verlag München

www.hanser-fachbuch.de

Lektorat: Dipl.-Ing. Volker Herzberg

Herstellung: le-tex publishing services GmbH, Leipzig

Coverkonzept: Marc Müller-Bremer, www.rebranding.de, München

Titelgrafik: © istockphoto.com/matdesign24

Coverrealisation: Max Kostopoulos

Satz: Kösel Media GmbH, Krugzell

Druck und Bindung: Hubert & Co. KG Buchpartner, Göttingen

Printed in Germany

Print-ISBN: 978-3-446-46146-8

E-Book-ISBN: 978-3-446-46397-4

Vorwort

Total Quality Management (TQM), Prozessorientierung in der DIN ISO 9001, SixSigma-Programme, Kontinuierliche Verbesserungsprogramme (KVP), Kaizen, ... – uns allen ist die Notwendigkeit der ständigen Verbesserung bewusst. Versuchsplanung ist eine Sammlung von Ideen und Verfahren, dabei systematisch vorzugehen, um mit möglichst geringem Aufwand möglichst viel zu lernen.

Im Rahmen einer SixSigma-Strategie ist Versuchsplanung das Werkzeug zur eigentlichen Verbesserung und nimmt damit eine zentrale Stellung ein. Dadurch hat Versuchsplanung in den letzten Jahren wesentlich an Bedeutung und Verbreitung gewonnen, und so ist nun schon die 10. Auflage dieses Buches erforderlich. Diese Chance habe ich wieder zur Aktualisierung und Ergänzung genutzt.

Ziel ist es, Praktikern in Entwicklung, Konstruktion und Fertigung, sowie Studenten einen anwendungsorientierten Einstieg und Überblick zu geben. Die Methoden der klassischen Statistischen Versuchsplanung werden mit Ideen von Shainin, Taguchi u.a. zu einer neuen Kombination verbunden.

SixSigma und Versuchsplanung sind Teamarbeit. Jedes Teammitglied muss über Ziele, Möglichkeiten und die prinzipielle Vorgehensweise Bescheid wissen. Aber nicht jedes Teammitglied muss alle Einzelheiten kennen.

- Kapitel 1 bis 5 geben einen allgemeinen Überblick über die Versuchsplanung und behandeln einfache Verfahren, die bei der Vorbereitung weiterer Versuche nützlich sind. Sie sind für alle Teammitglieder gedacht.
- Kapitel 6 bis 12 behandeln die statistischen Grundlagen und die wichtigsten Versuchspläne und ihre Auswertung. Sie wenden sich an das Teammitglied, das die Versuche plant und die Ergebnisse dann auswertet. Abschnitte, die mit einem Stern * gekennzeichnet sind und Ergänzungen in Fußnoten sind für das Verständnis der folgenden Kapitel nicht erforderlich und können zunächst ausgelassen werden.
- Kapitel 13 bis 20 behandeln verschiedene weiterführende Themen. Sie können bei Bedarf und unabhängig voneinander gelesen werden.

Im Download-Bereich des Hanser-Verlags zu diesem Buch finden Sie Visualisierungen in JavaScript, die Ihnen dabei helfen sollen, statistische Ideen besser zu begreifen. Erleben Sie selbst, wie Versuchsergebnisse streuen, was ein Vertrauensbereich ist, wie lineare Regression funktioniert und wie ein Fertigungsprozess immer besser wird.

Um das Verständnis für die Bedeutung statistischer Aussagen zu fördern, werden die meisten Beispiele ausführlich vorgerechnet. Eingestraute Übungsaufgaben verdeutlichen

und vertiefen die jeweiligen Inhalte. Nutzen Sie diese Übungsmöglichkeit – die folgende Lösung dient der Selbstkontrolle.

Obwohl aus didaktischen Gründen die Beispiele und Aufgaben hier von Hand vorgerechnet werden, empfehle ich ab Kapitel 6 parallel den Einsatz einer Software. Sie vereinfacht die Auswertung wesentlich und erlaubt vielfältige grafische Darstellungen.

Die Beschreibung der Versuchsplanung in diesem Buch ist unabhängig von einer speziellen Software. Viele gute Programme sind erhältlich. Kapitel 19 gibt Entscheidungshilfen zur Auswahl und einen Überblick über neun dieser Programme. Im Download-Bereich des Hanser-Verlags zu diesem Buch finden Sie Dateien mit Beispielen aus diesem Buch in den Formaten dieser Programme. Somit können Sie die Programme anhand bekannter Beispiele testen, direkt vergleichen und das Programm auswählen, das Ihnen am besten gefällt. Wenn Sie dann die Beispiele mit der gewählten Software nachvollziehen, werden Sie feststellen:

- Das Aufstellen von Versuchsplänen und die Auswertung der Versuchsergebnisse sind nicht schwer.
- Die Darstellung der Ergebnisse unterscheidet sich etwas von der Darstellung in diesem Buch. Jede Software ist anders, anhand der durchgerechneten Beispiele sollte es jedoch kein Problem sein, die Bedeutung der Ausgaben zu verstehen.
- Mit etwas Übung erscheint dann alles plötzlich ganz einfach. Aber auch darin liegt ein gewisses Risiko. Vergewissern Sie sich immer, dass die Daten und die Ergebnisse sinnvoll sind. Verwenden Sie Ihren gesunden Menschenverstand. Versuchsplanung ist ein sehr wertvolles Hilfsmittel. Aber es soll den gesunden Menschenverstand nicht ersetzen, sondern schärfen.

Als erste eigene Anwendung empfehle ich ein überschaubares Problem mit nur wenigen Faktoren und klar definierten Zielen. Bitte achten Sie auf die sorgfältige Vorbereitung Ihrer Versuche – sie ist entscheidend für den Erfolg.

Ich möchte darauf hinweisen, dass wesentliche Teile dieses Buches (insbesondere in den Kapiteln 7 bis 12) ursprünglich den ebenfalls von mir erstellten Schulungsunterlagen der Deutschen Gesellschaft für Qualität e. V. (DGQ), Frankfurt am Main, entnommen sind. Der Lehrgang „Statistische Versuchsplanung“ wird durch dieses Buch vertieft und ergänzt. Daher kann das Buch als begleitende oder weiterführende Literatur verwendet werden. Umgekehrt bietet der Lehrgang eine gute Einführung bzw. Ergänzung zu diesem Buch. Interessierte Leser können sich unter www.dgq.de über das Weiterbildungsangebot der DGQ informieren.

Zum Schluss möchte ich allen danken, die zu diesem Buch beigetragen haben, insbesondere der DGQ für die Genehmigung, Teile aus ihren Lehrgangsunterlagen zu verwenden, B. Schäfer und anderen Lesern für ihre hilfreichen inhaltlichen Anregungen und den Projektarbeitern der Hochschule Aalen für die Visualisierungen in JavaScript.

Allen Lesern bin ich dankbar für konstruktive Anregungen und Kritik. Ich wünsche Ihnen viel Erfolg bei der Anwendung der Versuchsplanung.

Aalen, im September 2019

Wilhelm Kleppmann

Inhalt

| | |
|--|----|
| Vorwort | v |
| 1 Einführung | 1 |
| 1.1 Warum Versuche? | 1 |
| 1.2 Warum Statistik? | 1 |
| 1.3 Warum Versuchsplanung? | 4 |
| 1.4 Welche Art von Ergebnissen kann man erwarten? | 6 |
| 1.5 Versuche oder systematische Beobachtung? | 7 |
| 1.6 Versuchsplanung und Six-Sigma-Strategie | 8 |
| 2 Ausgewählte Begriffe | 11 |
| 2.1 Zielgrößen | 12 |
| 2.2 Einflussgrößen | 12 |
| 2.3 Steuergrößen | 13 |
| 2.4 Störgrößen | 13 |
| 2.5 Faktoren | 14 |
| 2.6 Faktorstufen | 14 |
| 2.7 Quantitative und qualitative Faktoren | 15 |
| 3 Vorgehensweise im Überblick | 17 |
| 3.1 Ausgangssituation beschreiben | 17 |
| 3.2 Untersuchungsziel festlegen | 19 |
| 3.2.1 Optimale Lage des Mittelwerts | 19 |
| 3.2.2 Reduzierung der Streuung/Robustheit | 20 |
| 3.2.3 Erkennen der wichtigsten Störgrößen in der Fertigung | 21 |
| 3.2.4 Gleichzeitig fertigen und lernen | 22 |
| 3.2.5 Funktion und Zuverlässigkeit nachweisen | 22 |
| 3.3 Zielgrößen und Faktoren festlegen | 22 |
| 3.3.1 Auswahl der Zielgrößen | 22 |
| 3.3.2 Sammlung der Einflussgrößen | 24 |
| 3.3.3 Auswahl der Faktoren | 25 |
| 3.3.4 Festlegung der Faktorstufen | 26 |
| 3.3.5 Einflussgrößen, die nicht untersucht werden | 28 |
| 3.4 Versuchsplan aufstellen | 29 |
| 3.4.1 Festlegung der Faktorstufenkombinationen | 29 |

| | | |
|----------|---|----|
| 3.4.2 | Anzahl der Realisierungen | 30 |
| 3.4.3 | Blockbildung | 31 |
| 3.4.4 | Randomisierung | 31 |
| 3.4.5 | Aufwandsabschätzung | 33 |
| 3.5 | Versuche durchführen | 35 |
| 3.5.1 | Vorbereitung | 35 |
| 3.5.2 | Durchführung | 36 |
| 3.6 | Versuchsergebnisse auswerten | 37 |
| 3.7 | Ergebnisse interpretieren und Maßnahmen ableiten | 40 |
| 3.7.1 | Interpretation | 40 |
| 3.7.2 | Maßnahmen | 41 |
| 3.8 | Absicherung, Dokumentation, weiteres Vorgehen | 42 |
| 3.8.1 | Absicherung der Verbesserungen | 42 |
| 3.8.2 | Dokumentation | 42 |
| 3.8.3 | Weiteres Vorgehen | 43 |
| 4 | Systematische Beobachtung | 45 |
| 4.1 | Multi-Vari-Bild | 45 |
| 4.2 | Darstellung der örtlichen Verteilung von Fehlern | 49 |
| 4.3 | Prozessvergleich | 52 |
| 4.4 | Paarweiser Vergleich von Produkten | 54 |
| 5 | Einfache Versuche | 57 |
| 5.1 | Variablenvergleich zur Prozessverbesserung | 57 |
| 5.2 | Komponententausch zur Produktverbesserung | 61 |
| 5.3 | Überblick über die Methoden nach D. Shainin | 64 |
| 6 | Statistische Grundlagen | 65 |
| 6.1 | Verteilung | 65 |
| 6.1.1 | Häufigkeitsverteilung von Versuchsergebnissen | 65 |
| 6.1.2 | Verteilungsdichte und Verteilungsfunktion | 68 |
| 6.1.3 | Normalverteilung | 70 |
| 6.2 | Auswertung einer Stichprobe | 71 |
| 6.2.1 | Repräsentative Stichprobe | 71 |
| 6.2.2 | Eintragung ins Wahrscheinlichkeitsnetz | 73 |
| 6.2.3 | Schätzwerte für Mittelwert μ und Varianz σ^2 | 76 |
| 6.2.4 | Vertrauensbereiche | 77 |
| 6.3 | Vergleich von zwei Mittelwerten | 82 |
| 6.3.1 | Auswertung von Versuchsergebnissen | 83 |
| 6.3.2 | Festlegung des Stichproben- bzw. Versuchsumfangs | 89 |
| 6.3.3 | Voraussetzungen | 91 |
| 6.4 | Transformation von Messwerten | 95 |
| 6.4.1 | Logarithmische Normalverteilung | 95 |
| 6.4.2 | Poisson-Verteilung | 96 |
| 6.4.3 | Box-Cox-Transformation | 98 |

| | |
|---|-----|
| 7 Vollständige faktorielle Versuchspläne | 101 |
| 7.1 Zwei Faktoren auf je zwei Stufen | 101 |
| 7.1.1 Versuchsplan und Effekte | 101 |
| 7.1.2 Auswerteformalismus und Beurteilung der Signifikanz | 104 |
| 7.1.3 Interpretation von Wechselwirkungen | 106 |
| 7.1.4 Randomisierung und Blockbildung | 108 |
| 7.2 k Faktoren auf je zwei Stufen | 113 |
| 7.2.1 Versuchsplan | 113 |
| 7.2.2 Auswertung | 115 |
| 7.2.3 Versuchsumfang | 120 |
| 7.3 Auswertung von Versuchsplänen mit $n = 1$ | 121 |
| 7.3.1 Wahrscheinlichkeitsdarstellung der Effekte | 121 |
| 7.3.2 Schätzung der Zufallsstreuung durch „Pooling“ | 125 |
| 7.3.3 Risiken | 127 |
| 8 Screening-Versuchspläne | 129 |
| 8.1 Hintergrund | 129 |
| 8.2 Fraktionelle faktorielle Versuchspläne | 130 |
| 8.2.1 Der fraktionelle faktorielle 2^{4-1} -Plan als Beispiel | 130 |
| 8.2.2 Anwendung des 2^{4-1} -Plans zur Blockbildung | 134 |
| 8.2.3 Fraktioneller faktorieller 2^{k-p} -Plan | 137 |
| 8.2.4 Was bedeutet Vermengung? | 140 |
| 8.2.5 Auflösung | 143 |
| 8.2.6 Überblick über 2^{k-p} -Pläne | 144 |
| 8.2.7 Praxisbeispiel Reflowlöten | 147 |
| 8.3 Plackett-Burman-Versuchspläne* | 156 |
| 8.3.1 Plackett-Burman-Versuchspläne der Auflösung III | 157 |
| 8.3.2 Plackett-Burman-Versuchspläne der Auflösung IV | 159 |
| 8.3.3 Übersättigte Pläne | 160 |
| 8.4 Funktionstest* | 160 |
| 8.5 Definitive Screening Pläne* | 162 |
| 8.6 Einsatzempfehlungen | 164 |
| 9 Robuste Produkte/Prozesse | 167 |
| 9.1 Ziel und Strategie von G. Taguchi | 168 |
| 9.1.1 Qualitätsziel: Streuung minimieren | 168 |
| 9.1.2 Entwicklungsstrategie: Robuste Produkte/Prozesse | 168 |
| 9.2 Taguchis Versuchspläne und ihre Auswertung | 170 |
| 9.3 Alternative Ansätze | 178 |
| 9.3.1 Aus der Differenz von Messwerten abgeleitete Zielgrößen | 178 |
| 9.3.2 Wechselwirkung zwischen Steuer- und Rauschfaktoren | 179 |
| 9.4 Anmerkungen zu den „Orthogonalen Feldern“ u. ä.* | 180 |
| 9.4.1 Orthogonale Felder | 180 |
| 9.4.2 Lineare Graphen und Dreieckstabellen | 181 |
| 9.4.3 Dummy Levels, Pseudo Factor Designs, Idle Columns | 182 |

| | |
|--|-----|
| 10 Regressionsanalyse | 185 |
| 10.1 Einfache lineare Regression | 186 |
| 10.1.1 Methode der kleinsten Quadrate | 186 |
| 10.1.2 Bestimmtheitsmaß und Korrelationskoeffizient | 188 |
| 10.1.3 Grafische Beurteilung der Residuen | 191 |
| 10.1.4 Vertrauensbereiche und Signifikanz | 194 |
| 10.1.5 Zusammenhang lineare Regression – Mittelwertvergleich | 199 |
| 10.1.6 Quasilineare Regression | 200 |
| 10.2 Mehrfache Regression | 200 |
| 10.2.1 Zweifache lineare Regression | 201 |
| 10.2.2 Transformierte Einflussgrößen | 204 |
| 10.2.3 Prinzip der schrittweisen Regression | 207 |
| 10.2.4 Beurteilung des Regressionsmodells | 208 |
| 11 Versuchspläne für nichtlineare Zusammenhänge | 211 |
| 11.1 Zentral zusammengesetzte Versuchspläne | 211 |
| 11.1.1 Orthogonaler Versuchsplan | 213 |
| 11.1.2 Technisch bedingte Abweichungen vom Versuchsplan | 214 |
| 11.1.3 Bekannte nichtlineare Abhängigkeiten | 215 |
| 11.1.4 Varianten von zentral zusammengesetzten Plänen | 215 |
| 11.1.5 Praxisbeispiel Laserschneiden | 218 |
| 11.2 Alternative Pläne* | 226 |
| 11.2.1 3^k - und 3^{k-p} -Pläne | 226 |
| 11.2.2 Box-Behnken-Pläne | 226 |
| 11.2.3 Kleine zusammengesetzte Pläne | 227 |
| 11.2.4 Optimale Pläne | 229 |
| 11.3 Grenzen des quadratischen Modells | 230 |
| 11.3.1 Beispiele und Lösungsmöglichkeiten | 231 |
| 11.3.2 Allgemeine Lösungsansätze | 233 |
| 11.4 Einsatzempfehlungen | 234 |
| 12 Varianzanalyse | 237 |
| 12.1 Einfache balancierte Varianzanalyse | 238 |
| 12.2 Mehrfache Varianzanalyse | 243 |
| 12.3 Feste und zufällige Effekte* | 246 |
| 12.4 Nicht vollständige Randomisierung* | 248 |
| 12.4.1 Alle Realisierungen einer Kombination gemeinsam | 248 |
| 12.4.2 Split-Plot Versuche | 249 |
| 13 Screening für mehrstufige Faktoren* | 251 |
| 13.1 Versuchspläne | 251 |
| 13.2 Auswertung | 252 |
| 13.3 Einsatzempfehlungen | 254 |

| | |
|--|-----|
| 14 Versuchspläne für Mischungen* | 255 |
| 14.1 Mischungspläne ohne Begrenzungen | 256 |
| 14.2 Auswertung von Mischungsplänen | 258 |
| 14.3 Mischungspläne mit Begrenzungen | 258 |
| 14.4 Kombinierte Versuchspläne | 259 |
| 15 Spezielle Zielgrößen* | 261 |
| 15.1 Gut-Schlecht-Ergebnisse | 261 |
| 15.1.1 Möglichkeiten zur Vermeidung | 261 |
| 15.1.2 Auswertung | 263 |
| 15.2 Anzahl Fehler | 267 |
| 15.3 Mehrere Zielgrößen | 268 |
| 16 Erweiterung von Versuchsplänen* | 277 |
| 16.1 Trennung vermengter Wechselwirkungen | 277 |
| 16.2 Zentrumspunkt | 279 |
| 16.3 Zuordnung quadratischer Effekte | 281 |
| 16.4 Nicht realisierbare Faktorstufenkombinationen | 283 |
| 17 Alternative Modellierungsansätze* | 287 |
| 17.1 Andere Verteilungen: Maximum Likelihood Prinzip | 287 |
| 17.2 Robuste Regression | 288 |
| 17.3 Andere Modellansätze | 288 |
| 17.3.1 Überanpassung und Unteranpassung | 288 |
| 17.3.2 Lokale Modelle | 290 |
| 17.3.3 Neuronale Netze | 291 |
| 17.3.4 Gaussian Process Models | 291 |
| 17.3.5 Einsatzempfehlungen | 292 |
| 18 Sequentielle Optimierungsverfahren* | 295 |
| 18.1 Evolutionary Operations (EVOP) | 296 |
| 18.2 Methode des steilsten Anstiegs | 298 |
| 18.3 Simplexverfahren | 299 |
| 18.4 Neuere Entwicklungen | 301 |
| 19 Software* | 303 |
| 19.1 Allgemeine Hinweise | 303 |
| 19.2 Beschreibung ausgewählter Programme | 304 |
| 19.3 Spezielle Anwendungsgebiete | 311 |
| 20 Beispiele* | 315 |
| 20.1 Beispiel Motoroptimierung | 315 |
| 20.2 Literaturbeispiele | 320 |

| | |
|--|------------|
| 20.3 Übungsbeispiele | 321 |
| 20.3.1 Papier-Rotor | 321 |
| 20.3.2 Nürnberger Trichter | 324 |
| Anhang | 327 |
| Anhang A – Abkürzungen und Formelzeichen | 327 |
| Anhang B – Statistische Tabellen | 328 |
| Anhang C – Wegweiser durch die Verfahren | 330 |
| Anhang D – Ablauf einer Versuchsplanung | 331 |
| Anhang E – Ablauf einer Datenauswertung | 332 |
| Anhang F – Glossar deutsch/englisch | 333 |
| Anhang G – JavaScript-Visualisierungen zum Herunterladen | 337 |
| Anhang H – Beispiele zum Herunterladen | 341 |
| Anhang J – Software/Demos im Internet | 341 |
| Index | 345 |

* Für das Verständnis der folgenden Kapitel nicht erforderlich

1

Einführung

Dieses Kapitel beschreibt Prinzip und Hintergrund der Versuchsplanung. Unser Ziel ist es, Versuche so zu planen, dass wir die gewünschte Information mit einem Minimum an Zeit und Kosten erhalten.

■ 1.1 Warum Versuche?

Unternehmen müssen sich im Wettbewerb am Markt behaupten. Dazu müssen ihre Produkte und Fertigungsprozesse ständig verbessert werden:

- Der Funktionsumfang der Produkte muss erhöht werden. Die Anforderungen der Kunden müssen immer besser erfüllt werden.
- Die Kosten müssen gesenkt werden, z. B. durch geringere Materialkosten oder höhere Ausbeute.
- Die Entwicklungszeit neuer Produkte und ihre Durchlaufzeit in der Fertigung müssen immer weiter verkürzt werden.

Diese Verbesserungen können nicht allein durch Analyse von Daten aus der Fertigung und kritisches Nachdenken erreicht werden. Dazu sind die Zusammenhänge in Entwicklung, Fertigung und Qualitätsmanagement zu kompliziert und vielschichtig. Um den Einfluss von Designänderungen auf die Eigenschaften eines neuen Produktes oder den Einfluss von Änderungen von Prozessparametern auf das Prozessergebnis zu bestimmen, sind gezielte Versuche notwendig.

Jede Neu- oder Weiterentwicklung durchläuft daher in einem „globalen Versuchsplan“ eine Vielzahl von Versuchen (Pilotversuche, Prototypenversuche, Baumusterprüfungen, Dauerläufe, Zuverlässigkeitstests, Vorserien-Großversuche, Produktionsversuchsserien u. a.). Versuchsplanung (auch DOE = Design of Experiments genannt) hilft bei jedem dieser Schritte, gezielt zu besseren, wiederholbaren Ergebnissen zu gelangen.

■ 1.2 Warum Statistik?

Trotz aller Sorgfalt erhält man bei der Wiederholung eines einzelnen Versuchs meist nicht genau den gleichen Zahlenwert als Ergebnis. Zufällige Unterschiede, z. B. bei Ausgangs-

material, Umgebungsbedingungen und Messung, führen zu Unterschieden – die Versuchsergebnisse streuen.

Bild 1-1 zeigt ein Beispiel: Auf 30 Teilen wurde unter nominell gleichen Bedingungen galvanisch eine Schicht abgeschieden. Die gemessene Schichtdicke ist bei jedem Teil etwas anders, z.T. aufgrund echter Unterschiede in der Schicht, z.T. aufgrund von Unterschieden bei der Messung. Der Mittelwert der Schichtdicke beträgt in diesem Beispiel 30 µm – er beschreibt die Lage der Werte. Die Einzelwerte streuen um den Mittelwert. Die Standardabweichung der Schichtdicke beträgt 2 µm – sie ist ein Maß für die Breite des Bereichs, über den die Werte streuen. Ca. $\frac{2}{3}$ der Einzelwerte liegen im Bereich Mittelwert \pm Standardabweichung (hier 28 bis 32 µm).¹

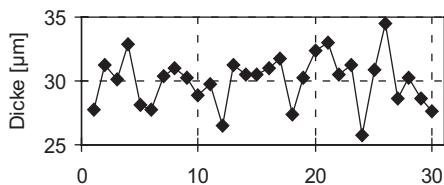


BILD 1-1

Dicke von 30 unter gleichen Bedingungen abgeschiedenen Schichten

Die Streuung der Werte ist kein Problem, wenn sie wesentlich kleiner ist als Unterschiede, die erkannt werden sollen. Bild 1-2 zeigt als Beispiel Messwerte von je 30 Teilen, die auf zwei verschiedenen Anlagen gefertigt wurden. Die Mittelwerte der Dicken betragen 30 µm bzw. 40 µm, die Standardabweichungen jeweils 2 µm. Auch ohne Einsatz statistischer Verfahren erkennt man in diesem Fall, dass die Teile von der einen Anlage eine größere Schichtdicke haben als die Teile von der anderen Anlage. Man kann sich eine Linie zwischen den beiden Messreihen denken, so dass alle Werte oberhalb der Linie zur einen Gruppe und alle Werte unterhalb zur anderen Gruppe gehören.

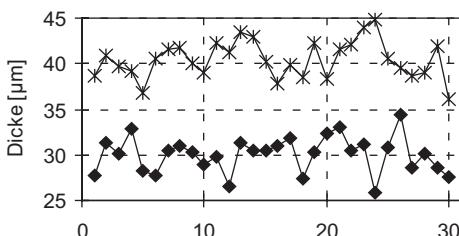


BILD 1-2

Zwei Messreihen mit Mittelwerten von 30 µm bzw. 40 µm und Standardabweichung von 2 µm

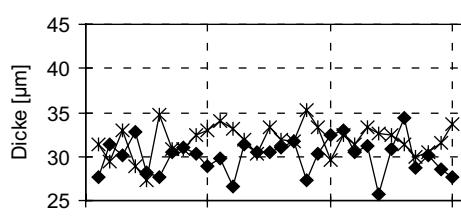


BILD 1-3

Zwei Messreihen mit Mittelwerten von 30 µm bzw. 32 µm und Standardabweichung von 2 µm

Häufig liegt jedoch eine Situation wie in Bild 1-3 vor. Es zeigt zwei Messreihen mit Mittelwerten von 30 µm bzw. 32 µm und Standardabweichungen von jeweils 2 µm (d.h. die Mittelwerte unterscheiden sich gerade um eine Standardabweichung). In Bild 1-3 ist der Unterschied zwischen den Anlagen nicht sofort erkennbar. Dies ist kein Problem, wenn der

¹ Hier soll nur das Prinzip erläutert werden, Einzelheiten siehe Kapitel 6.

kleine Unterschied keine Auswirkungen auf die Produkteigenschaften hat. Ist der Unterschied von 2 μm jedoch relevant, so kommt man ohne den Einsatz statistischer Methoden leicht zum falschen Ergebnis.

Statistische Verfahren basieren auf zwei Prinzipien:

1. Man sichert sich gegen Fehlentscheidungen ab. Zwei Versuchsergebnisse werden nur dann als unterschiedlich akzeptiert, wenn der beobachtete Unterschied so groß ist, dass er nur mit ausreichend kleiner Wahrscheinlichkeit zufällig auftritt, obwohl in Wirklichkeit kein Unterschied besteht. So ermöglicht Statistik rationale Entscheidungen trotz Zufallsstreuung.
2. Der Mittelwert von mehreren Messungen streut weniger als die Einzelwerte, weil zufällige Abweichungen sich teilweise kompensieren. Je größer die Anzahl der Einzelwerte ist, desto kleiner ist die Streuung ihres Mittelwerts. Bild 1-4 zeigt als Beispiel das Verhalten der Mittelwerte von je 30 Werten, die einzeln streuen wie in Bild 1-3. Bei den Mittelwerten ist der Unterschied zwischen den Anlagen deutlich erkennbar.

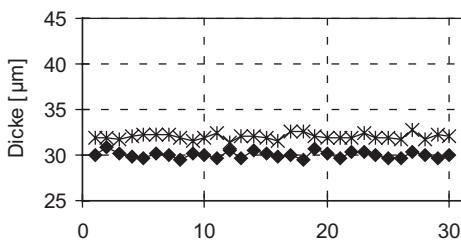


BILD 1-4

Mittelwerte von je 30 Einzelwerten, sonst wie Bild 1-3

Statistische Verfahren sind erforderlich,

- wenn bei den Versuchen Unterschiede zwischen Prozess- bzw. Produktvarianten erkannt werden sollen, die kleiner sind als ca. das Fünffache der Standardabweichung der Zufallsstreuung (d. h. kleiner als in Bild 1-2), oder
- wenn Unterschiede quantitativ ermittelt werden sollen, d. h. wenn man z. B. in Bild 1-2 angeben möchte, in welchem Bereich der wahre Unterschied zwischen den Schichtdicken mit einer bestimmten Wahrscheinlichkeit liegt.

Je kleinere Unterschiede man noch erkennen möchte bzw. je genauer man diese Unterschiede bestimmen möchte, desto größer ist der erforderliche Versuchsumfang. Als Dauermenregel kann gelten:

Will man einen Unterschied von einer Standardabweichung (d. h. von 2 μm im obigen Beispiel) erkennen, so benötigt man ca. 30 Einzelwerte je Prozess- bzw. Produktvariante (vgl. Bild 1-4).

Begnügt man sich jedoch damit, einen Unterschied von zwei Standardabweichungen (d. h. von 4 μm im obigen Beispiel) zu erkennen, so benötigt man um den Faktor vier weniger Einzelversuche (d. h. ca. 8 je Prozess- bzw. Produktvariante).²

² Hier soll nur ein Gefühl für die Größenordnung des erforderlichen Versuchsumfangs vermittelt werden, Einzelheiten siehe Kapitel 6.

■ 1.3 Warum Versuchsplanung?

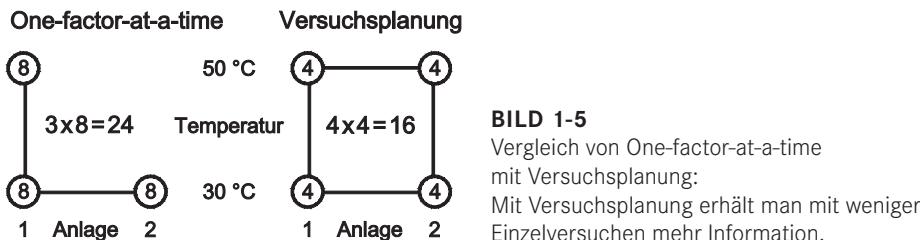
Versuche kosten Zeit und Geld, daher soll die Anzahl der Einzelversuche (= Versuchsumfang) möglichst klein sein. Ist der Versuchsumfang jedoch zu klein, so erkennt man relevante Unterschiede oft nicht. Daher bedeutet Versuchsplanung zunächst – wie in Abschnitt 1.2 –, den richtigen Versuchsumfang festzulegen.

So sind z. B. 8 Einzelversuche je Variante erforderlich, wenn man einen Unterschied von zwei Standardabweichungen erkennen möchte. Sollen also z. B. zwei Anlagen mit dieser Genauigkeit verglichen werden, so benötigt man je 8 Einzelversuche bei Anlage 1 und bei Anlage 2 – insgesamt also 16 Einzelversuche.

Oft soll zusätzlich der Einfluss anderer Größen ermittelt werden. Man kann z. B. daran interessiert sein, ob eine Erhöhung der Temperatur von 30 °C auf 50 °C die Schichtdicke verändert.

In solchen Fällen wird häufig der Rat gegeben, immer nur eine der Größen Anlage und Temperatur (man nennt sie Faktoren³) zu verändern (One-factor-at-a-time, links in Bild 1-5). So kann die Ursache für eine Veränderung des Ergebnisses leicht zugeordnet werden. Diese Vorgehensweise hat jedoch eine Reihe von gravierenden Nachteilen:

- Für die gewünschte Genauigkeit benötigt man 8 Einzelversuche je Variante, also insgesamt $3 \times 8 = 24$ Einzelversuche. Mit jedem zusätzlichen Faktor nimmt diese Anzahl um 8 zu. Der Aufwand wird schnell sehr groß.
- Trotz dieses großen Aufwands erhält man z. B. keine Information über das Verhalten von Anlage 2 bei der Temperatur 50 °C. Hat die Erhöhung der Temperatur bei Anlage 2 den gleichen Effekt wie bei Anlage 1 oder nicht?
- Die Kombination Anlage 1/Temperatur 30 °C hat eine größere Bedeutung als die anderen beiden Kombinationen.



Die Versuchsplanung dagegen empfiehlt, alle möglichen Kombinationen der Faktoren Anlage und Temperatur zu untersuchen (Bild 1-5 rechts), d. h. zusätzlich die Kombination Anlage 2/Temperatur 50 °C. Auf den ersten Blick sieht das nach mehr Aufwand aus. Und außerdem wurden z. T. beide Größen gleichzeitig verändert. Wie kann man da die Einflüsse trennen?

³ Definition der Begriffe in Kapitel 2.

Eine genauere Betrachtung zeigt jedoch, dass diese Vorgehensweise nur Vorteile hat:

- Insgesamt enthält der Versuchsplan nur $4 \times 4 = 16$ Einzelversuche, im Vergleich zu $3 \times 8 = 24$ bei One-factor-at-a-time. Trotzdem hat man 8 Wertepaare zur Berechnung des Einflusses der Anlage (man spricht vom „Effekt“ der Anlage), nämlich 4 bei 30 °C und 4 bei 50 °C. Und man hat 8 Wertepaare zur Berechnung des Einflusses der Temperatur, nämlich 4 bei Anlage 1 und 4 bei Anlage 2. Die statistische Absicherung ist daher ebenso gut wie bei den 24 Einzelversuchen bei One-factor-at-a-time.
- Man hat auch Versuchsergebnisse bei Anlage 2 und 50 °C und kann so überprüfen, ob der Einfluss der Temperatur von der Anlage abhängt oder nicht.
- Der Versuchsplan ist ausgewogen, keine Kombination hat eine größere Bedeutung als die anderen.

Auf diese Weise erhält man mit weniger Einzelversuchen mehr Information. Das Geheimnis des Erfolgs liegt in der Ausgewogenheit des Versuchsplans rechts in Bild 1-5. Sie erlaubt es, jedes Versuchsergebnis für die Berechnung des Effekts der Anlage **und** des Effekts der Temperatur zu benutzen. Jedes Ergebnis wird somit mehrfach genutzt. Daraus resultiert dann die Einsparung.

Bei mehr als zwei Faktoren ist die Einsparung durch Versuchsplanung noch größer. Bei 4 Faktoren genügen z.B. immer noch nur 16 Einzelversuche, während bei One-factor-at-a-time $(4 + 1) \times 8 = 40$ Einzelversuche erforderlich sind.

Beim praktischen Einsatz der Versuchsplanung ergeben sich noch weitere Vorteile:

- Weil die Faktoren gleichzeitig verändert werden, muss bereits am Anfang festgelegt werden, was untersucht werden soll. Dies erzwingt eine systematische Vorgehensweise, ausgehend vom Untersuchungsziel. „Einfach einmal probieren“ wird vermieden.
- Da der Aufwand von Anfang an besser abgeschätzt werden kann, ist eine verbesserte Kosten-Nutzen-Analyse bereits vor der Versuchsdurchführung möglich. Es wird vermieden, dass immer wieder „nur noch ein Versuch“ durchgeführt wird und damit letztlich der Kosten- und Zeitrahmen überschritten wird.
- Die systematische Darstellung der Ergebnisse erlaubt eine bessere Ableitung von Maßnahmen aus den Ergebnissen – und nur Verbesserungsmaßnahmen rechtfertigen den Aufwand für eine Untersuchung.
- Die systematische Darstellung erleichtert die Dokumentation der Ergebnisse und damit die Übertragung der Erfahrungen auf zukünftige Entwicklungen und andere Benutzer (z.B. Urlaubsvertretung, Nachfolger).

Typische Anwender berichten (vgl. z.B. [1]) eine

- Verkürzung von Projektlaufzeiten um 40 – 75 % und
- Senkung der Versuchskosten um 40 – 75 %.

■ 1.4 Welche Art von Ergebnissen kann man erwarten?

Als Ergebnis der Versuchsplanung und der anschließenden Auswertung erhält man ein empirisches Modell, das den Zusammenhang zwischen den untersuchten Faktoren (z.B. Prozessparametern) und den Zielgrößen (z.B. Schichtdicke, Ausbeute, Messwerten für Produktmerkmale) quantitativ beschreibt.

Da es sich um ein empirisches Modell handelt, dessen mathematische Form vorgegeben werden muss, beschreibt es den experimentell ermittelten Zusammenhang

- nur im untersuchten Bereich, d. h. eine Extrapolation ist nicht zulässig,
- nur im Rahmen der Möglichkeiten der vorgegebenen Form und
- nur unter Berücksichtigung der Zufallsstreuung.

Trotz dieser Einschränkungen sind die Ergebnisse sehr wertvoll für die Produkt- bzw. Prozessoptimierung. Bild 1-6 zeigt als Beispiel ein Teilergebnis einer Untersuchung zum Laserschneiden von Aluminiumblech. In Kapitel 11 wird dieses Beispiel ausführlich behandelt.

Durch den Laser wird das Aluminium geschmolzen. Druckgas entfernt das geschmolzene Material und führt so zur Trennung. Leider erstarrt ein Teil dieses Materials bereits an der Schnittkante wieder und bildet dort einen Grat oder Bart (Verletzungsrisiko). Ziel ist ein bartfreier Schnitt und gleichzeitig eine möglichst geringe Oberflächenrauheit. Dazu wurden u.a. die Prozessparameter (= Faktoren) Laserleistung und Schneidgeschwindigkeit im Rahmen eines Versuchsplans verändert. Für jede Kombination von Prozessparametern im Versuchsplan wurden die mittlere Barthöhe und Oberflächenrauheit gemessen. An die Ergebnisse wurde ein quadratisches Modell angepasst.

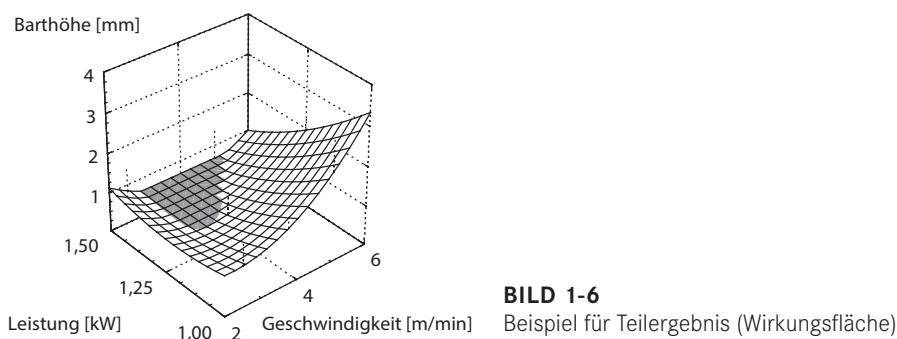


Bild 1-6 zeigt die Abhängigkeit der Barthöhe von Laserleistung und Schneidgeschwindigkeit in diesem Modell, wenn die anderen Prozessparameter festgehalten werden. Man erkennt sofort, dass eine hohe Laserleistung (1,5 kW ist die Maximalleistung des verwendeten Lasers) und eine mittlere Schneidgeschwindigkeit bezüglich der Barthöhe günstig sind. Im dunkler markierten Bereich tritt kein Bart auf. Mit abnehmender Laserleistung nimmt die optimale Schneidgeschwindigkeit ab.

Ähnliche Darstellungen erhält man natürlich auch für die Abhängigkeit der Barthöhe von den anderen Prozessparametern und für die Abhängigkeit der Rauheit von den Prozessparametern. Aus solchen Darstellungen kann man erkennen,

- welche Werte der Prozessparameter besonders günstig für die verschiedenen Zielgrößen sind,
- welche Kompromisse zwischen evtl. widersprüchlichen Anforderungen aus den verschiedenen Zielgrößen nötig bzw. möglich sind, und
- in welche Richtung evtl. noch weitere Verbesserungen möglich sind (in diesem Beispiel könnte man sich überlegen, ob es sich lohnt, einen Laser mit höherer Leistung einzusetzen).

Im Beispiel von Bild 1-6 war das Hauptziel, die Abhängigkeit der Zielgrößen Barthöhe und Rauheit von den Prozessparametern quantitativ (empirisch) zu erfassen. Aus dieser quantitativen Beschreibung der Abhängigkeit lassen sich dann Verbesserungsmaßnahmen für den Prozess ableiten. Für diese Art von Fragestellung sind die sogenannten **klassischen Methoden** der Versuchsplanung besonders geeignet.

Häufig ist das Hauptziel eine möglichst geringe Abhängigkeit von Störgrößen. Im Fall eines Fertigungsprozesses kann dies z.B. bedeuten, dass das Prozessergebnis möglichst wenig von Schwankungen der Prozessparameter, der Umgebungsbedingungen oder des Ausgangsmaterials beeinflusst wird. Im Fall eines Produktes kann dies auch eine möglichst geringe Abhängigkeit von den Einsatzbedingungen (z.B. Versorgungsspannung und Umgebungstemperatur bei einer elektronischen Schaltung) sein. Man spricht dann von robusten Prozessen bzw. Produkten. Für diese Fragestellung sind die **Methoden nach G. Taguchi** besonders geeignet.

In wieder anderen Fällen möchte man mit möglichst einfachen Mitteln und geringem Aufwand die wichtigsten Größen identifizieren, die die Streuung verursachen. Für diese Fragestellung sind (zumindest als Ansatzpunkt) die **Methoden nach D. Shainin** besonders geeignet.

Insgesamt betrachtet ist Versuchsplanung eine Sammlung von Ideen und Verfahren zur systematischen Verbesserung. Versuchsplanung ist ein Werkzeugkasten – einzelne Werkzeuge sind nicht besser oder schlechter als andere, sondern mehr oder weniger gut für die Lösung bestimmter Aufgaben geeignet. Ziel dieses Buches ist es, die Werkzeuge zu beschreiben und bei der Auswahl des jeweils geeigneten Werkzeugs zu helfen. Häufig ist es sinnvoll, verschiedene Werkzeuge nacheinander zu verwenden.

■ 1.5 Versuche oder systematische Beobachtung?

Geplante Versuche sind immer mit Aufwand verbunden. Daher stellt sich manchmal die Frage, ob es nicht ausreicht, die Fertigung eines Produktes systematisch zu beobachten und Ursache-Wirkungs-Zusammenhänge aus der gemeinsamen Veränderung von Prozessparametern und Ergebnissen (Korrelation) zu erkennen.

Nachteil dieser Vorgehensweise ist, dass eine solche Korrelation viele Ursachen haben kann und kein Beweis für einen Ursache-Wirkungs-Zusammenhang ist.



Beispiel

Bei einer chemischen Reaktion ist ein Ausgangsstoff manchmal verunreinigt. Die Verunreinigung führt einerseits zu einer Verringerung der Ausbeute, andererseits zu Schaumbildung. Um die Schaumbildung zu unterdrücken, wird dann der Druck erhöht.

Trägt man nun die Ausbeute gegen den Prozessparameter Druck auf, so ist bei hohem Druck immer die Ausbeute schlecht – man beobachtet eine deutliche Korrelation zwischen Druck und Ausbeute. Es besteht jedoch kein Ursache-Wirkungs-Zusammenhang. Vielmehr ist die Verunreinigung die gemeinsame Ursache für den hohen Druck und die schlechte Ausbeute.

Bei geplanten Versuchen werden die Prozessparameter gezielt verändert. Wenn die Änderung der Ergebnisse (Zielgrößen) durch die Änderung der Prozessparameter an- und wieder abgeschaltet werden kann, ist der Ursache-Wirkungs-Zusammenhang dadurch nachgewiesen. Allerdings kann nur der Effekt derjenigen Prozessparameter (Faktoren) erkannt werden, die im Versuchsplänen enthalten sind.

Daraus resultiert folgende Arbeitsteilung:

- Systematische Beobachtung gibt – neben Expertenwissen, Erfahrung, u.ä. – Hinweise auf vermutlich wichtige Einflussgrößen (vgl. Kapitel 4 und 10),
- die dann als Faktoren in geplante Versuche aufgenommen werden (vgl. z.B. Kapitel 3, 5, 7).

■ 1.6 Versuchsplanung und Six-Sigma-Strategie

In den letzten Jahren gewann die Six-Sigma-Strategie weite Verbreitung als einheitlicher Ansatz zur Beurteilung und Verbesserung aller Prozesse in einem Unternehmen (z.B. [2 – 4]). Eine Prozessverbesserung nach der Six-Sigma-Strategie besteht aus den fünf Phasen DMAIC:

- Definieren (**Define**):
Kunden und seine Anforderungen identifizieren, den zu verbesserten Prozess beschreiben und das Verbesserungsziel festlegen
- Messen (**Measure**):
Messgrößen zur Beurteilung des Prozessergebnisses (Zielgrößen) festlegen, Messmittelhaftigkeit und Prozessfähigkeit (σ -Niveau) bestimmen
- Analysieren (**Analyse**):
Systematische Beobachtung des Prozesses, Datensammlung und Suche nach Zusammenhängen zwischen Prozessparametern (Inputs) und Zielgrößen (Outputs)
- Verbessern (**Improve**):
Versuchsplanung und -auswertung, um Ursache-Wirkungs-Zusammenhänge zwischen Prozessparametern und Zielgrößen quantitativ zu bestimmen und daraus Verbesserungen abzuleiten

- Regeln (Control):

Prozessregelung, um die erreichte Verbesserung auf Dauer beizubehalten.

Dieses Buch über Versuchsplanung behandelt Methoden für die Schritte „Analyse“ (Kapitel 4 und 10) und „Improve“ (ab Kapitel 5), die Vorbereitungsphasen „Define“ und „Measure“ werden in Kapitel 3 angesprochen, „Control“ wird hier nicht behandelt.

Während der Schwerpunkt der Six-Sigma-Literatur auf den organisatorischen und strategischen Aspekten liegt, wird hier die praktische Anwendung der Versuchsplanung als Weg zur Prozessverbesserung beschrieben.

Literatur

- [1] *Barker, T.B.*: Quality by Experimental Design. Marcel Dekker, New York 1985
- [2] *Magnusson, K./Kroslid, D./Bergman, B.*: Six Sigma umsetzen. Hanser Verlag, München 2. Auflage 2003
- [3] *Melzer, A.*: Six Sigma – kompakt und praxisnah. Springer Gabler, Wiesbaden 2. Auflage 2019
- [4] *Wappis, J./Jung, B.*: Null-Fehler-Management: Umsetzung von Six Sigma. Hanser Verlag, München 6. Auflage 2019

2

Ausgewählte Begriffe

Dieses Kapitel erläutert einige Begriffe und ihre Anwendung in der Versuchsplanung.

Um die Begriffe möglichst anschaulich erläutern zu können, wird zunächst je ein Beispiel aus der Produktentwicklung und aus der Prozessentwicklung bzw. Fertigung beschrieben.



1. Optimierung einer Pumpenkonstruktion

Ausgehend von einer vorhandenen Pumpenkonstruktion soll eine ähnliche Pumpe mit um mindestens 2 % verbessertem Wirkungsgrad und 10% niedrigeren Herstellungskosten entwickelt werden. Das Saugvermögen soll möglichst nahe bei einem vorgegebenen Wert liegen.

Um den Einfluss verschiedener Konstruktionsparameter wie Gehäuse-, Steuerscheiben- und Schaufelradgeometrie zu untersuchen, werden diese in einem Versuch systematisch verändert. Für jede untersuchte Kombination von Konstruktionsparametern werden Saugvermögen und Wirkungsgrad der Pumpe gemessen. Die erwarteten Herstellungskosten werden von der Fertigungsabteilung geschätzt.

2. Optimierung eines chemischen Abscheideprozesses

Bei der Fertigung von integrierten Bauelementen (ICs) werden viele Schichten chemisch abgeschieden. Dazu werden z. B. 100 hintereinander aufgestellte Halbleiterscheiben gleichzeitig in einem Rohrofen auf hohe Temperatur erhitzt, durch den ein Gemisch aus Trägergas und verschiedenen Reaktionsgasen gepumpt wird. Bei der hohen Temperatur im Rohrofen zersetzen sich die Reaktionsgase. Zersetzungsprodukte scheiden sich auf den Scheiben ab, und eine Schicht wächst auf.

Ziel ist es, bei möglichst hoher Abscheiderate eine qualitativ hochwertige Schicht zu erhalten, deren Dicke nur geringfügig von der Position auf der Scheibe (Mitte oder Rand) und der Position der Scheibe im Ofen (vorne – mitten – hinten, in Strömungsrichtung der Reaktionsgase gesehen) abhängt.

Bei den Versuchen werden die Temperatur im Ofen (in verschiedenen Heizzonen), der Druck und die Durchflussmengen verschiedener Reaktionsgase verändert. Gemessen wird jeweils die Dielektrizitätskonstante der abgeschiedenen Schicht (sie gibt Aufschluss über die Qualität der Schicht) und die Dicke der Schicht in der Mitte und am Rand auf mehreren Scheiben, die an definierten Stellen im Ofen standen.

■ 2.1 Zielgrößen

Zielgrößen beschreiben das Ergebnis eines Versuchs. Zielgrößen können Messwerte sein, aber auch Größen, die aus einem oder mehreren Messwerten errechnet werden. Bei einem Versuch können mehrere Zielgrößen bestimmt werden.



1. Zielgrößen im Pumpenbeispiel

Wirkungsgrad, Saugvermögen und Herstellungskosten.

Das Saugvermögen wird direkt gemessen. Der Wirkungsgrad wird aus Messwerten wie Leistungsaufnahme, Förderhöhe u.ä., die Herstellungskosten werden aus Teilekosten und Montageaufwand berechnet.

2. Zielgrößen im Abscheidebeispiel

Dielektrizitätskonstante (als Indikator für die Qualität), mittlere Abscheiderate und Streuung der Abscheiderate (z.B. die Differenz zwischen Mitte und Rand auf einer Scheibe, die Differenz zwischen Scheiben von verschiedenen Positionen im Ofen).

Die Dielektrizitätskonstante wird direkt gemessen. Die Abscheiderate wird aus der Schichtdicke und der Abscheidezeit berechnet. Mittlere Abscheiderate und Streuung der Abscheiderate werden jeweils aus mehreren Messwerten berechnet.

■ 2.2 Einflussgrößen

Einflussgrößen sind Größen, die die Versuchsergebnisse (Zielgrößen) möglicherweise beeinflussen.

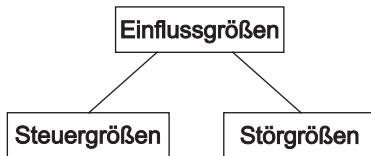


1. Einflussgrößen im Pumpenbeispiel

Die verschiedenen Konstruktionsparameter wie Gehäuse-, Steuerscheiben- und Schaufelradgeometrie, die im Versuch verändert werden, aber auch weitere Konstruktionsparameter, die das Ergebnis ebenfalls beeinflussen würden. Außerdem können z. B. auch Einsatzbedingungen oder Montagebedingungen Einflussgrößen sein.

2. Einflussgrößen im Abscheidebeispiel

Die Temperatur im Ofen (in verschiedenen Heizzonen), der Druck und die Durchflussmengen verschiedener Reaktionsgase, aber auch die Position auf der Scheibe, die Position der Scheibe im Ofen, Umgebungsbedingungen, die chemische Zusammensetzung der Reaktionsgase, das Trägergas, die Hersteller der Gase und der Scheiben usw.

**BILD 2-1**

Einflussgrößen werden von G. Taguchi in Steuergrößen und Störgrößen unterteilt

■ 2.3 Steuergrößen

Steuergrößen sind Einflussgrößen, deren Wert für das Produkt bzw. den Fertigungsprozess auf einen bestimmten Wert eingestellt und dort (in gewissen Grenzen) gehalten werden kann (Konstruktions- bzw. Prozessparameter, Bild 2-1).



1. Steuergrößen im Pumpenbeispiel

Die verschiedenen Konstruktionsparameter wie Gehäuse-, Steuerscheiben- und Schaufelradgeometrie, die im Versuch verändert werden, aber auch weitere Konstruktionsparameter, die das Ergebnis ebenfalls beeinflussen würden. Aufgabe der Entwicklung ist es, geeignete Werte für die Steuergrößen festzulegen.

2. Steuergrößen im Abscheidebeispiel

Prozessparameter, wie die Temperatur im Ofen (in verschiedenen Heizzonen), der Druck und die Durchflussmengen verschiedener Reaktionsgase, die chemische Zusammensetzung der Reaktionsgase, das Trägergas usw. Für die Fertigung werden geeignete Werte für die Steuergrößen festgelegt.

■ 2.4 Störgrößen

Störgrößen sind Einflussgrößen, deren Wert für das Produkt bzw. den Fertigungsprozess nicht vorgegeben werden kann (oder z.B. aus Kostengründen nicht vorgegeben werden soll, Bild 2-1).



1. Störgrößen im Pumpenbeispiel

Einsatzbedingungen beim Kunden (z.B. Umgebungstemperatur, Führung der Saug- und Druckleitung) oder Montagebedingungen, die nicht vorgegeben werden können; aber auch die zufällige Abweichung eines Konstruktionsparameters von seinem Sollwert (innerhalb der Spezifikation) wirkt wie eine Störgröße.

2. Störgrößen im Abscheidebeispiel

Die Position auf der Scheibe, die Position der Scheibe im Ofen, Umgebungsbedingungen; aber auch die zufällige Abweichung eines Prozessparameters wie der Temperatur von seinem Sollwert wirkt wie eine Störgröße.

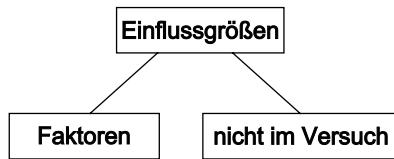


BILD 2-2

Einflussgrößen, die für den Versuch ausgewählt werden, heißen Faktoren

■ 2.5 Faktoren

Aus der Vielzahl der Einflussgrößen werden für den Versuch die vermuteten wesentlichen Einflussgrößen ausgewählt. Diese für den Versuch ausgewählten Einflussgrößen heißen Faktoren (Bild 2-2).



1. Faktoren im Pumpenbeispiel

Die verschiedenen Konstruktionsparameter wie Gehäuse-, Steuerscheiben- und Schaufelradgeometrie, die im Versuch verändert werden.

2. Faktoren im Abscheidebeispiel

Die Temperatur im Ofen (in verschiedenen Heizzonen), der Druck und die Durchflussmengen verschiedener Reaktionsgase. Die Position auf der Scheibe und die Position der Scheibe im Ofen können je nach Auswertung ebenfalls als Faktoren betrachtet werden.

Hinweis: Für den Versuch ausgewählte Steuergrößen heißen Steuerfaktoren. In einem Versuch können auch Störgrößen als Faktoren verändert werden, sie heißen Rauschfaktoren. Ziel des Versuchs ist es dann, Einstellungen der Steuerfaktoren zu finden, bei denen sich die Rauschfaktoren möglichst wenig auswirken.

■ 2.6 Faktorstufen

Nach der Auswahl der Faktoren muss festgelegt werden, welche Werte die Faktoren im Versuch annehmen sollen. Diese ausgewählten Werte werden als Stufen bezeichnet.



1. Faktorstufen im Pumpenbeispiel

Das Gehäuse kann im Versuch z. B. an einer bestimmten Stelle eine Nut haben oder nicht, die Stufen für den Faktor „Gehäusenut“ sind dann „mit Nut“ und „ohne Nut“.

Ähnlich müssen für alle Faktoren mindestens zwei Stufen festgelegt werden.

2. Faktorstufen im Abscheidebeispiel

Die Temperatur kann z. B. die Stufen 700 °C und 720 °C haben, d.h. ein Teil (normalerweise die Hälfte) der Versuche wird bei 700 °C, der Rest bei 720 °C durchgeführt.

Der Druck kann z. B. die Stufen 130 mbar und 160 mbar haben usw.

■ 2.7 Quantitative und qualitative Faktoren

Je nach Art der Faktorstufen wird zwischen quantitativen (= numerischen) und qualitativen (= kategorischen) Faktoren unterschieden. Bei quantitativen Faktoren werden die Stufen mit Zahlenwerten auf einer Messskala beschrieben, bei qualitativen Faktoren gibt es keine Skala, nur Namen, Beschreibungen oder Bezeichnungen. Bei quantitativen Faktoren gibt es – im Gegensatz zu qualitativen Faktoren – zumindest im Prinzip auch Zwischenwerte und es ist sinnvoll, Ergebnisse zu interpolieren.



1. Pumpenbeispiel

„Gehäusenut“ mit den Stufen „mit Nut“ und „ohne Nut“ ist ein Beispiel für einen qualitativen Faktor. Der Abstand zwischen Gehäuse und Schaufelrad und andere geometrische Dimensionen sind quantitative Faktoren.

2. Abscheidebeispiel

Temperatur und Druck sind Beispiele für quantitative Faktoren. Die Position auf der Scheibe (Mitte – Rand) oder der Scheibe im Ofen (vorne – mitte – hinten) sind qualitative Faktoren, wenn sie so angegeben werden, als Abstand von der Mitte oder als Nummer der Scheibe im Ofen sind sie quantitativ. Werden mehrere Öfen verwendet, so ist der Ofen ein qualitativer Faktor mit den Stufen Ofen 1, Ofen 2, Ofen 3 usw. (im Gegensatz zur Nummer der Scheibe im Ofen enthält die Nummer des Ofens keine inhaltliche, sondern nur organisatorische Information – es ist nicht sinnvoll, zwischen Ofen 1 und Ofen 3 zu interpolieren).

3

Vorgehensweise im Überblick

In diesem Kapitel werden die Einzelschritte beschrieben, aus denen ein geplanter Versuch besteht (siehe auch [1 - 4]). Der Schwerpunkt liegt hier auf der Vorbereitung und Nachbereitung. Die eigentliche Versuchsplanung und die Auswertung der Ergebnisse werden später im Detail beschrieben.

Ein wesentlicher Aspekt der Versuchsplanung ist, dass bereits in der Planungsphase alle Betroffenen mit eingebunden werden. Soll z.B. die Entwicklungsabteilung eine Untersuchung zu einem neuen Produkt durchführen, so müssen die Anforderungen (Untersuchungsziele) mit dem Marketing abgestimmt sein, damit das Produkt die Wünsche der Kunden erfüllt, und mit der Fertigung, damit das Produkt kostengünstig gefertigt werden kann. Die beteiligten Mitarbeiter aus den betroffenen Bereichen müssen zwar keine Versuchsplanung im Detail durchführen können, sollen jedoch einen Überblick über Vorgehensweise und Möglichkeiten der Versuchsplanung haben. Dann können sie besser zur Definition der Untersuchungsziele beitragen.

Ziel dieses Kapitels ist es, den so von der Untersuchung Betroffenen genügend Information zu geben, dass sie gezielt zur Vorbereitung, insbesondere zur Definition der Untersuchungsziele, beitragen können.

■ 3.1 Ausgangssituation beschreiben

Zur Vorbereitung einer Untersuchung gehört, dass man sich zunächst Rechenschaft ablegt über das Umfeld. Dazu gehören folgende Fragen:

- Wer ist der Kunde?

Für wen wird die Untersuchung gemacht? Was stört ihn? Was braucht er? Was ist ihm eine Verbesserung wert? Die Kundenorientierung hilft bei der Formulierung der Ziele und beim Setzen von Prioritäten. Der Kunde kann extern oder firmenintern sein.

- Was ist die langfristige Zielsetzung?

Jede Untersuchung kostet Zeit und Geld. Sie ist daher nur zu verantworten, wenn sie einen entsprechenden Nutzen bringt. Um den Nutzen einer Untersuchung beurteilen zu können, muss sie in eine Gesamtstrategie eingeordnet sein.

- Welches (Teil-)Problem soll durch die jetzt geplante Untersuchung gelöst werden?

Insbesondere bei komplexen Problemen ist es sinnvoll, sie in überschaubare Teile zu zerlegen und in mehreren Schritten vorzugehen. So kann das Ergebnis des einen Schrit-

tes bei der Planung des nächsten berücksichtigt werden und jeder Einzelschritt bleibt einfach.

- Wie viel Zeit und Geld stehen maximal zur Verfügung?

Bei der Verbesserung von Fertigungsprozessen hängt dies von der erzielbaren Einsparung und Wertsteigerung ab. Einsparungen können z.B. durch geringeren Ausschuss, niedrigere Materialkosten oder kürzere Bearbeitungszeiten erzielt werden. Eine Wertsteigerung ergibt sich z.B. bei einer leichteren Verarbeitbarkeit in Folgeprozessen.

Bei der Produktentwicklung hängt dies von der erwarteten Wertsteigerung aus Sicht des Kunden oder von der Kosteneinsparung ab.

Normalerweise besteht eine Optimierung aus mehreren Einzelschritten. Der Aufwand für einen einzelnen Versuchsplan sollte daher ca. ein Drittel des maximalen Aufwands nicht übersteigen.

Für die spätere Dokumentation der Ergebnisse wird empfohlen, die Ausgangssituation quantitativ zu erfassen. So wird eine Kosten-Nutzen-Analyse möglich.

- Wer ist von der geplanten Untersuchung betroffen – und: Sind alle eingebunden?

Um eine falsche oder unvollständige Zielsetzung zu vermeiden, müssen alle Betroffenen eingebunden sein. Die Entwicklungsabteilung kann z.B. Untersuchungsziele nicht ohne Berücksichtigung der Kundenwünsche (Marketing) und der Möglichkeiten der Fertigung und Zulieferer festlegen.

Für Versuche in der Fertigung ist der wichtigste Teilnehmer der Mann vor Ort, wie z.B. der Anlagenbediener. Er kennt das Problem am besten. Außerdem ist psychologisch wichtig, dass er aktiv mitwirkt und weiß, worum es geht. Ihm soll geholfen werden. Er muss diese Hilfe aber auch akzeptieren.

- Wer ist für das Projektmanagement verantwortlich?

Zielsetzung, Randbedingungen u.ä. werden im Team festgelegt. Aber einer muss für das Gesamtprojekt verantwortlich sein. Er kümmert sich um die Einhaltung des Termin- und Kostenplans. Er überwacht das Gesamtprojekt („Kümmerer“, bei SixSigma: Black Belt).

- Was ist über das zu untersuchende Problem bereits bekannt?

Das Rad muss nicht noch einmal erfunden werden. Es ist wichtig, dass Erfahrungen mit ähnlichen Problemen aus der Vergangenheit berücksichtigt werden. Oft finden sich auch in der Fachliteratur nützliche Hinweise.

Eine Darstellung des Prozessablaufs (z.B. als Flussdiagramm) hilft, die Diskussion im Team zu fokussieren. Bereits bekannte Zusammenhänge und noch offene Fragen können übersichtlich eingetragen werden.

Systematische Beobachtung (Kapitel 4) kann wertvolle Hinweise für die Planung aktiver Versuche liefern. Bereits vorliegende Daten sollten daraufhin analysiert werden, ob sie Hinweise auf die Problemursachen geben (Vorsicht, auf Scheinkorrelationen achten!).

■ 3.2 Untersuchungsziel festlegen

Vor Beginn der detaillierten Planung müssen sich alle Betroffenen über das Untersuchungsziel einig sein. Ohne eine solche Einigung werden die Ergebnisse unbefriedigend bleiben.

Dieser Schritt ist eine große Kunst. Er erfordert sowohl Sachkenntnis als auch Fingerspitzengefühl im Umgang mit den Betroffenen. Oft müssen gegensätzliche Abteilungsinteressen miteinander vereinbart werden.

Da es verschiedene Methoden der Versuchsplanung gibt, die jeweils für unterschiedliche Fragestellungen optimal sind, werden hier als Hilfe bei der Festlegung des Untersuchungsziels die verschiedenen grundsätzlichen Möglichkeiten kurz vorgestellt:

3.2.1 Optimale Lage des Mittelwerts

Das Prozessergebnis oder ein Produktparameter soll einen bestimmten Wert annehmen. In diesem Fall ist das Prozessergebnis bzw. der Produktparameter selbst eine geeignete Zielgröße. Aus statistischen Gründen wird dann meist der Mittelwert betrachtet.



Beispiele

- Der Wirkungsgrad einer Pumpe soll möglichst hoch sein.
- Der Mittelwert für das Saugvermögen einer Pumpe soll möglichst nahe an einem vorgegebenen Wert sein.
- Die Abscheiderate soll möglichst hoch sein.
- Die Ausbeute einer chemischen Reaktion soll möglichst hoch sein.

In der Untersuchung möchte man ermitteln, wie der Mittelwert von den verschiedenen Faktoren abhängt. Aus dieser Abhängigkeit ergeben sich dann optimale Einstellungen für die Faktoren.

Häufig sollen mehrere Zielgrößen gleichzeitig optimiert werden. Faktoreinstellungen, die für eine Zielgröße optimal sind, sind dann für andere Zielgrößen meist nicht gleichzeitig optimal. Aus der quantitativen Kenntnis der Abhängigkeiten aller Zielgrößen von den Faktoren kann man bei Zielkonflikten optimale Kompromisse finden.



Beispiele

- Bei der Entwicklung einer neuen Pumpe sollen eine vorgegebene Förderhöhe und Fördermenge bei möglichst geringer Leistungsaufnahme und möglichst geringen Fertigungskosten erreicht werden. Dazu wird die Abhängigkeit der Zielgrößen Förderhöhe, Fördermenge, Leistungsaufnahme und Fertigungskosten von verschiedenen Konstruktionsparametern (Faktoren) untersucht.
- Bei einem Fertigungsprozess treten verschiedene Fehlerarten auf. Das Problem ist, dass Einstellungen der Prozessparameter, die für eine Fehlerart günstig sind, für andere Fehlerarten u.U.

ungünstig sind. Gesucht wird eine Kompromisseinstellung, die die gesamten Fehlerkosten minimiert. Dazu wird die Abhängigkeit der verschiedenen Fehlerarten von den Prozessparametern untersucht.

- Für eine neue Infrarotlötanlage soll die Abhängigkeit der Temperatur an verschiedenen Stellen auf der Leiterplatte von der Transportgeschwindigkeit und der Temperatur der Infrarotstrahler an verschiedenen Stellen in der Anlage bestimmt werden. Ziel ist ein bestimmter zeitlicher Verlauf der mittleren Temperatur auf der Leiterplatte und die Minimierung der Temperaturdifferenz auf der Leiterplatte.
-

Für die quantitative Bestimmung der Abhängigkeit der Mittelwerte von den Faktoren bieten die **klassischen Methoden** der Versuchsplanung die besten und vielseitigsten Möglichkeiten. Diese werden vor allem in den Kapiteln 6 bis 8 und 10 bis 15 ausführlich behandelt.

3.2.2 Reduzierung der Streuung/Robustheit

In der Fertigung ist oft weniger die Lage des Mittelwerts des Prozessergebnisses problematisch als dessen Streuung. Die Streuung bestimmt z.B. die Prozessfähigkeit. Soll die Streuung reduziert werden, so ist die Standardabweichung von mehreren Versuchsrealisierungen eine geeignete Zielgröße.



Beispiele

- Das Saugvermögen einer Pumpe ist spezifiziert mit $1700 - 1750 \text{ m}^3/\text{h}$. Aufgrund kleiner, zufälliger Unterschiede der Einzelkomponenten und der Montagebedingungen hat nicht jede gefertigte Pumpe genau das gleiche Saugvermögen. Zur Erhöhung der Prozessfähigkeit soll die Zufallsstreuung des Saugvermögens reduziert werden.
 - Der Durchmesser von Wellen nach dem Schleifen streut zu stark, die Streuung soll reduziert werden.
-

In diesen Beispielen ist man an der Streuung des Ergebnisses interessiert. Die Ursachen für die Streuung sind unbekannt, es handelt sich um Unterschiede zwischen Teilen, die unter gleichen Bedingungen gefertigt wurden.

G. Taguchi hat diese Idee noch weiter entwickelt. Häufig sind einige wenige Störgrößen für einen großen Teil der Streuung verantwortlich. Dann kann man gezielt nach Einstellungen von Steuergrößen suchen, bei denen sich eine Veränderung dieser Störgrößen im üblichen Bereich möglichst wenig auswirkt. Taguchi verwendet dafür den Begriff Robustheit.

Ein robuster Prozess ist ein Prozess, dessen Ergebnis möglichst wenig von bestimmten Störgrößen abhängt. Ein robustes Produkt ist ein Produkt, dessen Eigenschaften möglichst wenig von bestimmten Störgrößen (d.h. bestimmten Fertigungs- oder Einsatzbedingungen) abhängen.



Beispiele

- Die Streuung der Schichtdicke auf einer Halbleiterscheibe soll minimiert werden (es gibt systematische Unterschiede zwischen Mitte und Rand der Scheibe, die Unterschiede sind nicht zufällig).
- Der Einfluss der Umgebungstemperatur und der Eingangsspannung auf die Ausgangsspannung einer elektronischen Schaltung soll minimiert werden.

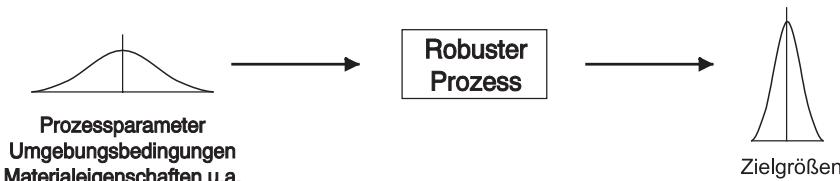


BILD 3-1 Untersuchungsziel nach G. Taguchi

Bild 3-1 erläutert die Zielsetzung von G. Taguchi: Die Steuergrößen (z.B. Prozessparameter) sind so festzulegen, dass trotz zufälliger Streuung dieser Größen um ihren Sollwert und/oder trotz der Streuung von Störgrößen (z.B. Umgebungsbedingungen, Materialeigenschaften) die Werte der Zielgrößen nur wenig streuen. Dadurch kann der Aufwand zur Begrenzung der Streuung reduziert (Kosteneinsparung) und trotzdem gute Qualität an den Kunden geliefert werden.

Die Behandlung der Streuung als Zielgröße und die Besonderheiten der Vorgehensweise nach G. Taguchi werden in Kapitel 9 beschrieben. Als Grundlage dienen die in den Kapiteln 7 und 8 behandelten Versuchspläne.

3.2.3 Erkennen der wichtigsten Störgrößen in der Fertigung

Manchmal werden Probleme erst in der Fertigung erkannt. Die Fertigungsstreuung ist zu groß oder sporadische Fertigungseinbrüche treten auf. Dann möchte man mit möglichst einfachen Mitteln erkennen, welche Störgrößen dafür vor allem verantwortlich sind. Diese Störgrößen können dann gezielt überwacht (enger toleriert) werden. Dies ist zwar mit Kosten verbunden, sie werden aber in Kauf genommen, um den Versuchsaufwand gering zu halten.

Ziel der Methoden von **D. Shainin** ist es, durch eine systematische Beobachtung der Fertigung und einfache Versuche herauszufinden, welche Störgrößen besonders wichtig sind. Diese Vorgehensweise unterscheidet sich wesentlich von der bei den anderen Methoden. Der Rest dieses Kapitels ist daher nicht direkt anwendbar. Einem Anwender, der sich zunächst auf gezielte Beobachtung seiner Fertigung und einfache Versuche beschränken möchte, wird daher empfohlen, zunächst Kapitel 4 und 5 zu lesen. Dort werden die Besonderheiten der Vorgehensweise nach D. Shainin beschrieben.

In günstigen Fällen führt diese Vorgehensweise sehr rasch und mit geringen Kosten zum Erkennen und Beseitigen der Probleme. In weniger günstigen Fällen erleichtern die Ergebnisse wenigstens die Festlegung der Zielgrößen und Faktoren in Abschnitt 3.3.

3.2.4 Gleichzeitig fertigen und lernen

Neue Fertigungsverfahren werden oft in einer Pilotlinie entwickelt und optimiert. Einstellungen der Prozessparameter, die in der Pilotlinie optimal waren, müssen in der großtechnischen Fertigungslinie nicht optimal sein. Durch systematische Veränderung der Prozessparameter innerhalb ihrer Spezifikation in der laufenden Fertigung kann man Verbesserungsmöglichkeiten erkennen. Die Fertigung liefert so gleichzeitig das Produkt und Information über Verbesserungsmöglichkeiten. Diese Optimierungsstrategie wird mit Evolutionary Operations (**EVOP**) bezeichnet und in Abschnitt 18.1 behandelt. Grundlage sind Versuchspläne aus Kapitel 7.

3.2.5 Funktion und Zuverlässigkeit nachweisen

Bevor ein neues Produkt freigegeben werden kann, ist nachzuweisen, dass es unter allen spezifizierten Einsatzbedingungen bestimmungsgemäß funktioniert und eine ausreichende Zuverlässigkeit erreicht.

Das Problem besteht nun darin, dass die Einsatzbedingungen durch viele verschiedene Größen beschrieben werden, wie z.B. Umgebungstemperatur, Luftfeuchtigkeit, Versorgungsspannung, Last, Details der Systemkonfiguration. Für all diese Größen sind Bereiche spezifiziert und das Produkt muss für beliebige Konfigurationen funktionieren – aber häufig ist nicht klar, welche Kombination von Einsatzbedingungen die ungünstigste ist.

Abschnitt 8.4 zeigt, wie mit relativ wenigen Einzelversuchen (fast) alle Extremkombinationen von beliebigen drei oder vier dieser Größen erfasst werden können.

■ 3.3 Zielgrößen und Faktoren festlegen

Im Folgenden werden Hinweise zur Auswahl geeigneter Zielgrößen, Faktoren und Faktorstufen gegeben. Die geeignete Auswahl dieser Größen ist entscheidend für den Erfolg einer Versuchsplanung. Die Auswahl sollte daher mit besonderer Sorgfalt erfolgen.

3.3.1 Auswahl der Zielgrößen

Bei der Auswahl der Zielgrößen ist auf folgende Aspekte zu achten:

- Kundenorientierung/Relevanz

Zielgrößen müssen die Probleme des Kunden abbilden und in möglichst engem Zusammenhang zu den Untersuchungszielen stehen.

- Quantifizierung

Zielgrößen sollten möglichst quantitative Größen sein. Messwerte (oder daraus berechnete quantitative Größen) enthalten wesentlich mehr Information als Gut/Schlecht-Aussagen. Daraus ergibt sich eine kleinere Zufallsstreuung und damit ein kleinerer Ver-

suchsumfang. Bei nicht direkt messbaren Größen kann man sich mit Noten behelfen. Eine Abstufung von 1 bis 6 enthält mehr Information als nur bestanden/nicht bestanden.

■ Vollständigkeit

Alle wesentlichen Prozessergebnisse bzw. Produkteigenschaften müssen als Zielgrößen erfasst werden. Es ist wichtig, dass nicht nur die momentan problematischen Größen berücksichtigt werden, denn bei der Optimierung bezüglich dieser Größen könnte sich ansonsten eine andere Größe unbemerkt verschlechtern. Die meisten Anwendungen haben daher mehrere Zielgrößen. Trotzdem wird in den Übungsbeispielen wegen der leichteren Überschaubarkeit normalerweise nur eine Zielgröße behandelt. Bei mehreren Zielgrößen werden die Auswertungen für jede Zielgröße getrennt durchgeführt.

■ Verschiedenheit

Die Interpretation der Ergebnisse wird erleichtert, wenn die Anzahl der Zielgrößen möglichst klein ist und jede Zielgröße einen anderen, möglichst grundlegenden Zusammenhang erfasst.

Manchmal muss zwischen Relevanz für den Kunden und Aufwand der Messung einer Zielgröße abgewogen werden. So ist z.B. die Ermittlung der Lebensdauer eines Produktes sehr aufwändig – daher wird oft nach Indikatorgrößen gesucht, die leichter zu messen sind. Dies setzt jedoch voraus, dass die verwendete Messgröße wirklich ein Indikator für die Lebensdauer ist (dies muss schon vorher bekannt sein).

Die Messstreuung sollte klein im Vergleich zu den tatsächlichen Unterschieden sein. Dies kann mit einer Messmittelfähigkeitsuntersuchung überprüft werden. Ist die Messstreuung nicht klein genug, sollte das Messverfahren verbessert werden. Notfalls kann die Messung auch mehrfach wiederholt und der Mittelwert verwendet werden.

Als Zielgröße kann eine Messgröße selbst oder eine aus mehreren Messgrößen berechnete Größe verwendet werden (z.B. Differenz der Abscheiderate an Mitte und Rand) – wichtig ist, dass die Untersuchungsziele erfasst werden.



Hinweis

Bei einem geregelten System ist manchmal wählbar, was Einflussgröße und was Zielgröße ist. Bei einem Einspritzmotor besteht z.B. ein Zusammenhang zwischen Drehzahl, Drehmoment und Einspritzmenge (Verbrauch). Je größer die Einspritzmenge, desto größer werden Drehzahl und Drehmoment. Auf einem Prüfstand kann man jedoch auch Drehzahl und Drehmoment vorgeben und die Einspritzmenge dann so regeln, dass die vorgegebenen Werte eingehalten werden. Die nötige Einspritzmenge (= Verbrauch) wird damit zu einer Zielgröße. Dies ist günstig, wenn der Einfluss von Drehzahl, Drehmoment und verschiedenen Einspritzparametern auf den Verbrauch bestimmt werden soll. Hier bestimmt die geplante Anwendung, ob eine Größe Einflussgröße oder Zielgröße ist. Vgl. dazu auch das Beispiel in Abschnitt 20.1.

3.3.2 Sammlung der Einflussgrößen

Die Bedeutung einer Einflussgröße, die nicht als Faktor in der Untersuchung enthalten ist, kann natürlich nicht erkannt werden. Und verändert sich eine wichtige Einflussgröße unkontrolliert während des Versuchs, so kann sie die Ergebnisse verfälschen. Daher dürfen bei der Planung keine wichtigen Einflussgrößen vergessen werden.

Es hat sich bewährt, die Auswahl der Faktoren in zwei Schritte zu zerlegen: Zunächst werden möglichst viele Einflussgrößen gesammelt. Erst danach wird aus der Vielzahl der Einflussgrößen eine handhabbare Anzahl von Faktoren für die weitere Untersuchung ausgewählt.

Zur Sammlung der Einflussgrößen ist das Brainstorming geeignet. Dies ist eine Technik zur Ideenfindung, bei der das kreative Denkvermögen einer Gruppe angeregt wird.

Beim Brainstorming ist wichtig, dass alle Teammitglieder zu Wort kommen. Dies kann durch eine mündliche Befragung reihum oder eine schriftliche Abfrage von Ideen auf Kärtchen (Metaplanteknik) erfolgen. Bei der mündlichen Befragung können Teilnehmer sich gegenseitig inspirieren, aber auch beeinflussen. Mit der schriftlichen Abfrage wird die Beeinflussung vermieden; die Dominanz einzelner Teammitglieder wird damit reduziert. In jedem Fall ist entscheidend, dass während der Ideenfindung keine Idee kritisiert oder gar als untauglich abgewertet wird.

Grafische Darstellungen helfen, die Diskussion im Team zu strukturieren und möglichst nichts zu vergessen. Besonders bewährt haben sich dazu folgende Hilfsmittel:

- **Prozessablaufdiagramm** (Flussbild)

Der (Fertigungs-)Prozess wird in Einzelschritte zerlegt und für jeden Schritt werden Vorgehensweise, sowie Eingangs- und Ausgangsgrößen beschrieben.

- **Ursache-Wirkungs-Diagramm** (auch Ishikawa- oder Fischgrätdiagramm)

Ursache-Wirkungs-Zusammenhänge werden grafisch immer weiter unterteilt, meist angefangen mit Mensch, Maschine, Material, Methode, Mitwelt und Messung.

- **Einflussgrößen-Zielgrößen-Matrix**

Das im Team bereits vorhandene Wissen über die Wirkzusammenhänge zwischen Einflussgrößen und Zielgrößen wird in Form einer Matrix zusammengestellt – jede Einflussgröße ist eine Zeile, jede Zielgröße eine Spalte, das Wissen wird Feld um Feld zusammengetragen [1, 3].

1. Schritt: Für jede Zielgröße eine Spalte anlegen.

2. Schritt: In der 1. Spalte alle Einflussgrößen sammeln, die irgendeine Zielgröße beeinflussen könnten (Brainstorming, Erfahrung).

3. Schritt: Einflussgrößen in Steuergrößen und Störgrößen unterteilen.

4. Schritt: Für jede Einflussgröße das vorhandene Wissen über Größe und Art ihres Einflusses auf jede Zielgröße sammeln (z.B. starker, schwacher, kein Einfluss, unbekannt, linear, nicht linear, usw.).



Beispiel aus der chemischen Industrie (Ausgangspunkt)

Im Rest dieses Kapitels werden die Einzelschritte an einem Beispiel aus der chemischen Industrie erläutert:

Zur Fertigung einer bestimmten Chemikalie werden mehrere Ausgangsstoffe einschließlich Katalysator in einem Reaktionsgefäß vermischt. Die Mischung wird anschließend über längere Zeit unter Röhren erhitzt, dabei erfolgt die Reaktion. Dann wird das Reaktionsprodukt abgetrennt. Ziel ist eine Erhöhung der Ausbeute bei möglichst geringen Kosten.

Tabelle 3.1 zeigt einen Auszug aus der Liste der Einflussgrößen, mit den Einschätzungen des Teams zur Größe ihres Einflusses auf die beiden Zielgrößen. Die Spalte „Wert“ wird erst später ausgefüllt. Bei Einflussgrößen, die als Faktoren in den Versuch aufgenommen werden, enthält sie dann den untersuchten Wertebereich, bei den anderen Einflussgrößen den verwendeten festen Wert – so können später leicht verschiedene Versuche verglichen werden, da alle wichtigen Größen dokumentiert sind.

| Einflussgröße | Art | Wert | Ausbeute | Kosten |
|------------------|-------------|------|----------|--------|
| Temperatur | Steuergröße | | stark | gering |
| Reaktionszeit | Steuergröße | | stark | stark |
| Katalysatormenge | Steuergröße | | stark | gering |
| Rührrate | Steuergröße | | gering | kein |
| Materialcharge | Störgröße | | gering | kein |
| Bediener | Störgröße | | gering | kein |
| Verunreinigung | Störgröße | | stark | gering |
| usw | | | | |

TABELLE 3.1

Einflussgrößen-Zielgrößen-Matrix für das Beispiel (Auszug), mit der im Team erwarteten Größe des jeweiligen Einflusses

Temperatur, Reaktionszeit, Katalysatormenge und Rührrate können an der Anlage eingestellt werden und sind damit Steuergrößen. Die anderen Einflussgrößen können in der späteren Produktion nicht vorgegeben werden und sind damit Störgrößen. Das Ausgangsmaterial kann zu einer Steuergröße werden, wenn z. B. verschiedene Hersteller miteinander verglichen werden und der beste ausgewählt werden soll. Verunreinigung kann zu einer Steuergröße werden, wenn z. B. verschiedene Reinigungsverfahren zur Wahl stehen.

3.3.3 Auswahl der Faktoren

Erst nach Abschluss der Ideenfindung werden die Ideen bewertet und gewichtet. In dieser Phase wird aus der Vielzahl der gefundenen (möglichen) Einflussgrößen eine handhabbare Anzahl von Faktoren für die weitere Untersuchung ausgewählt (typisch 3–6, nur selten mehr als 10).

Das wichtigste Auswahlkriterium ist die vermutete Bedeutung der Einflussgrößen für die Lösung des zu untersuchenden Problems. Zusätzlich kann die Genauigkeit und Reproduzierbarkeit, mit der verschiedene Stufen der Einflussgrößen eingestellt werden können, und der Aufwand für die Einstellung und Veränderung der Stufen berücksichtigt werden.

Zur Vorbereitung von Versuchen in der Fertigung wird empfohlen, zunächst die Fertigung systematisch zu beobachten und die Ergebnisse dieser Beobachtung bei der Auswahl der

Faktoren zu berücksichtigen. Dazu sind die Methoden von D. Shainin sehr wertvoll (Kapitel 4 und 5).

Wichtig ist, dass die Auswahl im Konsens erfolgt. Gehören die Teammitglieder unterschiedlichen Hierarchieebenen an, so ist besonders darauf zu achten, dass nur fachliche Aspekte berücksichtigt werden. Prozess- bzw. Produktkenntnisse sind hier wichtiger als die Stellung eines Teammitglieds im Unternehmen.

Die ausgewählten Faktoren müssen unabhängig voneinander veränderbar sein. Manchmal gibt es mehrere Möglichkeiten, solche unabhängigen Faktoren festzulegen. Man sollte dann die Festlegung verwenden, für die man (z.B. aus technischen Gründen) die direkteste und einfachste Abhängigkeit von den Zielgrößen erwartet.



Geometrisches Beispiel (vgl. auch Kapitel 20)

Bei einem Rechteck können die Länge L und die Breite B unabhängig voneinander verändert werden. Die Fläche $F = L \cdot B$ und das Verhältnis $V = L/B$ ergeben sich dann und können nicht mehr unabhängig festgelegt werden. Als Faktoren können beliebige zwei dieser vier Größen verwendet werden, d.h. man hat z.B. folgende Möglichkeiten:

- Faktoren L und B, wenn man vermutet, dass die Zielgrößen physikalisch oder technisch bedingt direkt von diesen Größen abhängen, oder
- Faktoren F und V, wenn man vermutet, dass die Zielgrößen direkt von diesen Größen abhängen (direkter als von L und B getrennt), usw.

Für die ausgewählten Faktoren werden Stufen festgelegt (s.u.). Die Werte für die anderen beiden Größen werden dann mit den obigen Formeln errechnet.

Wenn man vermutet, dass die Zielgrößen nur von einer dieser Größen abhängen, so genügt es, diese eine Größe als Faktor zu verwenden. Eine andere der Größen wird dann festgehalten, die anderen beiden werden errechnet.

3.3.4 Festlegung der Faktorstufen

Anzahl der Faktorstufen

Enthält ein Versuchsplan viele Faktoren (mehr als 5 bis 6), so werden zur Begrenzung des Aufwands normalerweise nur zwei Stufen für jeden Faktor untersucht. Mit einem solchen Versuchsplan kann man feststellen,

- welche der Faktoren wichtig sind,
- wie groß ihr linearer Effekt auf jede der Zielgrößen ist, und
- welche Stufe eines jeden Faktors für jede der Zielgrößen günstig ist (erlaubt auch Kompromisse bei Zielkonflikten).

Enthält ein Versuchsplan nur wenige Faktoren, so können auch mehr als zwei Stufen für jeden Faktor untersucht werden. Damit kann man auch das Maximum/Minimum einer Zielgröße finden.

Bei der Festlegung der Anzahl der Stufen spielt auch die Art des Faktors eine Rolle:

- Quantitative Faktoren wie Temperatur, Druck o.ä. können meist beliebige Werte (in einem sinnvollen Bereich) annehmen. Soll als erste Näherung der lineare Effekt von vielen Faktoren bestimmt werden oder erwartet man nur eine lineare Abhängigkeit, so verwendet man normalerweise auch nur zwei Stufen. Erwartet man dagegen eine nicht-lineare Abhängigkeit und ist die Anzahl der Faktoren ausreichend klein, so verwendet man 3 bis 5 Stufen.
- Qualitative Faktoren wie Anlage, Hersteller o.ä. können meist nur bestimmte Werte annehmen (z.B., weil zwei Anlagen oder drei Hersteller miteinander verglichen werden sollen). Dann ergibt sich die Anzahl der Stufen oft aus der Problemstellung. Soll der Effekt vieler Faktoren bestimmt werden, beschränkt man sich zunächst auf die zwei wichtigsten Stufen (z.B. die beiden wichtigsten Hersteller).

Werte der Faktorstufen (quantitative Faktoren)

- Ausgangspunkt ist normalerweise der bisher beste Fertigungsprozess bzw. das bisher beste Produkt. Ohne weitere Information untersucht man einen Bereich, der symmetrisch dazu liegt. Vermutet man eine Verbesserung in eine bestimmte Richtung, nimmt man als eine Stufe den bisher besten Zustand, als zweite Stufe einen Wert in die vermutlich bessere Richtung.
- Je kleiner der Abstand zwischen den Stufen ist, desto kleiner ist auch der Unterschied zwischen den Ergebnissen und desto größer ist der Versuchsumfang, der erforderlich ist, um trotzdem einen Unterschied zwischen den Ergebnissen zu erkennen.
- Kann ein Faktor nicht genau gemessen werden, so sollte der Abstand der Faktorstufen mindestens 6σ sein, wenn σ die Standardabweichung für die Messung des Faktors ist (kann z.B. die Temperatur bei einer chemischen Reaktion nur mit einer Standardabweichung von 5°C gemessen werden, so liegen Stufenwerte von 60°C und 70°C zu dicht beieinander, 50°C und 80°C liegen dagegen ausreichend weit auseinander).
- Je größer der Abstand zwischen den Stufen ist, desto größer kann die Abweichung der Ergebnisse von der Linearität sein. Mit zwei Stufen können diese Abweichungen jedoch nicht erkannt werden, sie verfälschen daher das Ergebnis. Insbesondere darf zwischen den Stufen kein neues physikalisches oder technisches Phänomen auftreten (z.B. Wasser gefrieren).
- Eine Extrapolation der Ergebnisse über den untersuchten Bereich hinaus ist nicht zulässig, daher sollte die Untersuchung den interessanten Bereich möglichst enthalten.
- Bei EVOP werden als Faktorstufen meist die Spezifikationsgrenzen für Prozessparameter verwendet.
- Bei der Behandlung von Störgrößen als Faktoren und beim Funktionstest werden realistische Werte verwendet (d.h. Werte, die auch in Wirklichkeit auftreten, wie z.B. untere und obere Spezifikationsgrenze für die Einsatzbedingungen eines Produktes).

Bild 3-2 zeigt qualitativ, dass die Faktorstufen nicht zu nah beieinander, aber auch nicht zu weit auseinander liegen sollten.

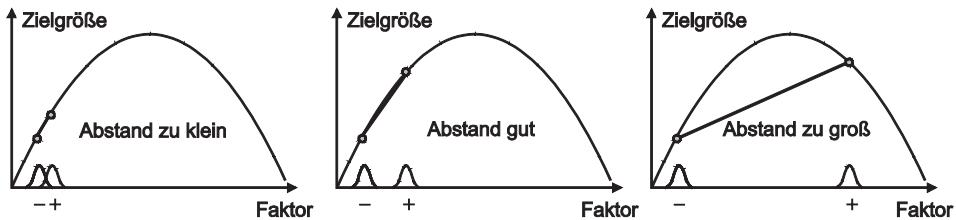


BILD 3-2 Liegen die Stufenwerte eines Faktors zu nah beieinander, wird sein Effekt von Zufallsstreuung verdeckt; liegen die Stufenwerte zu weit auseinander, so wird der Zusammenhang schlecht beschrieben (zweistufiger Faktor)



Beispiel aus der chemischen Industrie (Fortsetzung)

Aus den vielen Einflussgrößen in Tabelle 3.1 wurden folgende als vermutlich wichtigste Faktoren ausgewählt (ein starker Einfluss auf die Ausbeute wird erwartet):

- Temperatur der Mischung:
Bisheriger Wert 120 °C. Bringt eine Erhöhung der Temperatur auf den Maximalwert der Anlage von 140 °C eine Verbesserung der Ausbeute?
- Reaktionszeit:
Bisheriger Wert 2 h. Bringt eine Verlängerung auf 4 h eine Verbesserung?
- Menge des Katalysators:
Bisheriger Wert 0,1%. Bringt eine Erhöhung auf 0,5% eine Verbesserung?

Jeder der drei quantitativen Faktoren wird auf zwei Stufen untersucht. Ausgangspunkt ist jeweils der bisherige Wert. Da bekannt ist, in welche Richtung eine Verbesserung zu erwarten ist, verwendet man als eine Stufe den bisherigen Wert, als zweite Stufe einen Wert in diese Richtung.

Bei der Temperatur ist die Obergrenze durch die Anlage vorgegeben, bei den anderen Faktoren wird als zweite Stufe ein deutlich höherer Wert als der bisherige verwendet. Alle Faktoren können genau gemessen werden, der Unterschied zwischen den Stufen ist jeweils wesentlich größer als 6σ . Kein neues Phänomen ist zu erwarten. Die Abhängigkeit der Ausbeute von der Zeit dürfte nichtlinear sein (Sättigung), hier interessiert aber im Wesentlichen die Größe der Abhängigkeit. Um den Aufwand zu begrenzen, begnügt man sich daher zunächst mit zwei Stufen.

Hinweis: Zur Überprüfung der Linearität ist es bei quantitativen Faktoren sinnvoll, zusätzlich Einzelversuche im Zentrum des untersuchten Wertebereichs durchzuführen (vgl. Abschnitt 16.2). Zeigt sich dabei, dass eine wesentliche Abweichung von der Linearität besteht, so kann der Versuchsplan nachträglich erweitert werden (vgl. Abschnitt 16.3).

3.3.5 Einflussgrößen, die nicht untersucht werden

Nicht untersuchte Einflussgrößen (vor allem Störgrößen) werden möglichst konstant gehalten, um so die Zufallsstreuung zu minimieren. Je kleiner die Zufallsstreuung ist, desto kleinere Effekte der Faktoren können noch erkannt werden, bzw. desto weniger Einzelversuche sind erforderlich. Gelingt es z.B. die Zufallsstreuung zu halbieren, so genügt bereits ein Viertel der Einzelversuche.



Beispiel aus der chemischen Industrie (Fortsetzung)

- Als Rührrate wird der übliche Wert 5/min festgelegt.
- Ausgangsmaterial aus einer Charge wird verwendet (Chargennummer notieren).
- Derselbe Bediener führt alle Einzelversuche in derselben Anlage durch.
- Die Anlage wird vor jedem Einzelversuch sorgfältig gereinigt.
- Die Umgebungsbedingungen werden möglichst konstant gehalten.
- Alle Einzelversuche werden in möglichst kurzem zeitlichem Abstand durchgeführt.
- Derselbe Bediener trennt das Reaktionsprodukt ab und führt alle Messungen mit jeweils demselben Instrument durch.

Die Festlegungen werden z.B. in der Einflussgrößen-/Zielgrößenmatrix dokumentiert. Festgehaltene Werte werden notiert, damit die Versuchsbedingungen später möglichst genau reproduziert werden können. Werte von Störgrößen, die nicht konstant gehalten werden können, werden erfasst und aufgezeichnet – so kann ihr Einfluss nachträglich erkannt werden.

■ 3.4 Versuchsplan aufstellen

Der optimale Versuchsplan hängt vom Untersuchungsziel, von der Anzahl der Faktoren, von der Anzahl der Stufen je Faktor, von der gewünschten Genauigkeit der Ergebnisse und von der Größe der Zufallsstreuung ab. Hier werden die wichtigsten Aspekte schrittweise am Beispiel aus der chemischen Industrie erläutert. Einzelheiten werden ab Kapitel 6 behandelt.

3.4.1 Festlegung der Faktorstufenkombinationen

Zunächst wird festgelegt, welche Kombinationen der Faktorstufen untersucht werden sollen. Dazu gibt es Standardpläne, die vor allem in Kapitel 7, 8, 11, 13 und 14 beschrieben werden.



Beispiel aus der chemischen Industrie (Fortsetzung)

Drei Faktoren sollen auf je zwei Stufen untersucht werden. Ein sogenannter „vollständiger faktorieller Versuchsplan“ (Kapitel 7) besteht aus allen acht möglichen Faktorstufenkombinationen (siehe Tabelle 3.2).¹

¹ Ergänzende Versuche im Zentrum des untersuchten Bereichs sind sinnvoll und werden in Abschnitt 16.2 behandelt.

| syst. Nr. | Temperatur [°C] | Zeit [h] | Katalysator [%] |
|-----------|-----------------|----------|-----------------|
| 1 | 120 | 2 | 0,1 |
| 2 | 140 | 2 | 0,1 |
| 3 | 120 | 4 | 0,1 |
| 4 | 140 | 4 | 0,1 |
| 5 | 120 | 2 | 0,5 |
| 6 | 140 | 2 | 0,5 |
| 7 | 120 | 4 | 0,5 |
| 8 | 140 | 4 | 0,5 |

TABELLE 3.2

Liste der Faktorstufenkombinationen für das Beispiel

3.4.2 Anzahl der Realisierungen

Je genauer man den Effekt der Faktoren bestimmen möchte und je stärker die Versuchsergebnisse aufgrund zufälliger Unterschiede streuen (Maßzahl ist die Standardabweichung σ , vgl. Absatz 6.1.3), desto mehr Einzelversuche sind erforderlich. Für Versuche mit zweistufigen Faktoren kann als Daumenregel gelten, dass man insgesamt

$$N \approx 60 \cdot \left(\frac{\sigma}{\Delta\mu} \right)^2 \quad (3.1)$$

Einzelversuche durchführen sollte, wenn ein Effekt von $\Delta\mu$ technologisch relevant ist und mit hoher Wahrscheinlichkeit erkannt werden soll, falls er existiert.

Besteht der Versuch aus m Faktorstufenkombinationen, so wird jede dieser Kombinationen

$$n \approx \frac{N}{m} \quad (3.2)$$

mal durchgeführt (man spricht von n -maliger Realisierung). Einzelheiten werden in Kapitel 6 und 7 behandelt.



Beispiel aus der chemischen Industrie (Fortsetzung)

Aus der Fertigungsüberwachung ist bekannt, dass die Standardabweichung der Ausbeute aufgrund der Zufallsstreuung $\sigma = 1\%$ beträgt. Wenn ein Ausbeuteunterschied von 2% relevant ist und im Versuch erkannt werden soll, so benötigt man

$$N \approx 60 \cdot \left(\frac{1}{2} \right)^2 = 15 \text{ Einzelversuche.}$$

Jede der acht Faktorstufenkombinationen sollte daher $15/8 \approx$ zweimal realisiert werden (d. h. für jede Kombination werden zwei Einzelversuche durchgeführt).

Falls σ nicht bekannt ist, wird empfohlen, in zwei Schritten vorzugehen. Zunächst wird mit einem kleinen Wert für n (2 oder 3) begonnen. Falls dies bereits genügt, um ausreichend abgesicherte Ergebnisse zu erhalten, ist der Versuch damit abgeschlossen. Andernfalls kann aus den Ergebnissen ein Schätzwert für σ und damit ein ausreichender Wert für n berechnet werden.

3.4.3 Blockbildung

Ziel der Blockbildung ist es, die Zufallsstreuung zu minimieren, um so mit möglichst wenigen Einzelversuchen abgesicherte Ergebnissen zu erhalten.

Unter Blockbildung versteht man die Einteilung der Einzelversuche bzw. Versuchsobjekte in Gruppen, so dass innerhalb jeder Gruppe (Block genannt)

- einerseits die zufälligen Unterschiede möglichst klein sind und
- andererseits jede Faktorstufenkombination möglichst gleich häufig auftritt.

Eventuell vorhandene Unterschiede zwischen den Blöcken können aufgrund der Ausgewogenheit innerhalb der Blöcke erkannt und aus der Rechnung eliminiert werden. So kann die Zufallsstreuung reduziert werden.



Beispiele

- Reicht eine Charge des Ausgangsmaterials nicht für alle geplanten Einzelversuche, so ist es sinnvoll, aus den Einzelversuchen mit Ausgangsmaterial einer Charge jeweils einen Block zu bilden. Eventuelle Unterschiede zwischen den Chargen können dann aus der Zufallsstreuung herausgerechnet werden.
- Kann man nicht alle Einzelversuche in einer Woche durchführen, so ist es sinnvoll, aus den Einzelversuchen einer Woche jeweils einen Block zu bilden. Eventuelle Unterschiede zwischen den Wochen (z.B. aufgrund des Abschaltens am Wochenende) können dann aus der Zufallsstreuung herausgerechnet werden.
- Auch ohne erkennbare Einschnitte können Versuchsergebnisse sich allmählich verändern (Trend). So kann z.B. das Reaktionsgefäß verschmutzen, die Temperaturanzeige driften, das Material altern oder der Bediener immer geübter werden. Dann sind die Unterschiede zwischen zeitlich aufeinanderfolgenden Einzelversuchen kleiner als zwischen Versuchen in großem zeitlichem Abstand. Durch eine Unterteilung der Versuche in Blöcke kann man die Unterschiede innerhalb eines Blocks reduzieren.

Werden die Faktorstufenkombinationen n -mal realisiert, so kann man die Einzelversuche einer Realisierung jeweils als einen Block betrachten (ergibt n Blöcke). In Kapitel 8 werden weitere Möglichkeiten der Blockbildung behandelt.

3.4.4 Randomisierung

Um zu verhindern, dass ein Trend oder eine andere unerkannte Änderung der Ergebnisse die Schätzung der Effekte der Faktoren verfälscht, werden die Einzelversuche in jedem Block in zufälliger Reihenfolge durchgeführt. Man spricht von Randomisierung.

Achtung: „Zufällige Reihenfolge“ bedeutet nicht „beliebige Reihenfolge“. Die Reihenfolge wird vor der Versuchsdurchführung mit Hilfe von Zufallszahlen festgelegt. Diese Reihenfolge muss dann eingehalten werden. Eine spätere Änderung der Reihenfolge (z.B. um weniger oft zwischen den Stufen eines bestimmten Faktors wechseln zu müssen) birgt das Risiko einer Verfälschung der Ergebnisse.

Werden für die Durchführung der Versuche Teile benötigt, so werden diese Teile den Faktorstufenkombinationen zufällig zugeordnet, um eine Verfälschung der Ergebnisse durch evtl. vorhandene Unterschiede zwischen den Teilen zu vermeiden. Dies verursacht keinen Mehraufwand, bringt aber mehr Sicherheit.

Durch Blockbildung und Randomisierung zusammen erreicht man eine optimale Absicherung gegen Veränderungen der Versuchsbedingungen. Dabei gilt das Prinzip:

- Zunächst durch Blockbildung bekannte, kontrollierbare Veränderungen aus der Zufallsstreuung heraus halten
- und dann innerhalb der Blöcke randomisieren, um eine Verfälschung der Effekte durch verbliebene unbekannte, nicht kontrollierbare Veränderungen zu vermeiden.



Beispiel aus der chemischen Industrie (Fortsetzung)

Tabelle 3.3 zeigt eine mögliche Reihenfolge der Einzelversuche, die die Ideen von Blockbildung und Randomisierung berücksichtigt. Jede Faktorstufenkombination wird einmal in zufälliger Reihenfolge realisiert, dann noch einmal in anderer Reihenfolge.

TABELLE 3.3 Liste der Einzelversuche im Beispiel (syst. Nr. aus Tabelle 3.2)

| Vers. Nr. | syst. Nr. | Realisierung (Block) | Temperatur [°C] | Zeit [h] | Katalysator [%] | Ausbeute [%] |
|-----------|-----------|-------------------------|--------------------|-------------|--------------------|-----------------|
| 1 | 8 | 1 | 140 | 4 | 0,5 | |
| 2 | 3 | 1 | 120 | 4 | 0,1 | |
| 3 | 1 | 1 | 120 | 2 | 0,1 | |
| 4 | 7 | 1 | 120 | 4 | 0,5 | |
| 5 | 6 | 1 | 140 | 2 | 0,5 | |
| 6 | 2 | 1 | 140 | 2 | 0,1 | |
| 7 | 4 | 1 | 140 | 4 | 0,1 | |
| 8 | 5 | 1 | 120 | 2 | 0,5 | |
| 9 | 4 | 2 | 140 | 4 | 0,1 | |
| 10 | 3 | 2 | 120 | 4 | 0,1 | |
| 11 | 8 | 2 | 140 | 4 | 0,5 | |
| 12 | 5 | 2 | 120 | 2 | 0,5 | |
| 13 | 1 | 2 | 120 | 2 | 0,1 | |
| 14 | 6 | 2 | 140 | 2 | 0,5 | |
| 15 | 7 | 2 | 120 | 4 | 0,5 | |
| 16 | 2 | 2 | 140 | 2 | 0,1 | |

3.4.5 Aufwandsabschätzung

Wenn die Einzelversuche und ihre Reihenfolge wie z. B. in Tabelle 3.3 feststehen, kann der Versuchsaufwand abgeschätzt werden. Er muss in einem angemessenen Verhältnis zum Wert der erhofften Verbesserung bzw. Einsparung stehen.



Beispiel aus der chemischen Industrie (Fortsetzung)

Eine Kostenanalyse ergibt folgende Werte:

- Grundkosten für Material, Anlage, Reinigung des Produkts und Personal bei den momentanen Einstellungen der Fertigung (Temperatur 120 °C, Zeit 2 h und Katalysator 0,1%): 5000 € pro Füllung des Reaktionsgefäßes
- Wert (Verkaufserlös) des Produkts: 100 € pro % Ausbeute und Füllung (d. h. Kostendeckung bei 50 % Ausbeute)
- Anzahl Füllungen pro Jahr: ca. 500 (mit derselben Anlage werden auch andere Produkte gefertigt)
- Erhoffte Ausbeuteerhöhung: 5 % (das bedeutet $5 \times 100 \times 500 = 250\,000$ € zusätzlicher Gewinn pro Jahr)
- Zusätzliche Energiekosten zur Erhöhung der Temperatur von 120 auf 140 °C: 50 €/Füllung
- Zusätzliche Anlagenkosten zur Verlängerung der Zeit von 2 h auf 4 h: 1000 €/Füllung
- Zusätzliche Materialkosten zur Erhöhung des Katalysators von 0,1 % auf 0,5 %: 50 €/Füllung
- Ingenieurkosten für Vorbereitung, Durchführung und Auswertung des geplanten Versuchs (bestehend aus 16 Einzelversuchen): 10 000 €
- Der Versuch hat keinen Einfluss auf das Produkt nach der Reinigung, es kann verkauft werden. Da der Versuch in der Fertigung durchgeführt wird, fallen die Grundkosten sowieso an. Jede Füllung wird erst im Reaktionsgefäß gemischt und erhitzt, daher entstehen keine Zusatzkosten durch die Änderung der Versuchsbedingungen (Randomisierung). Als Kosten für die Versuche fallen daher nur die Ingenieurkosten und die zusätzlichen Energie-, Anlagen- und Materialkosten bei jeweils der Hälfte der Einzelversuche an:

$$10\,000 + 8 \times 50 + 8 \times 1000 + 8 \times 50 \approx 20\,000 \text{ €}$$
- Die Versuchskosten von 20 000 € stehen in einem sehr günstigen Verhältnis zu den erhofften zusätzlichen Einnahmen von 250 000 € pro Jahr. Der Versuch wird daher wie geplant durchgeführt.

Blockbildung und Randomisierung führen dazu, dass für jeden Faktor die Stufenwerte häufig verändert werden müssen. Je nach Aufwand für die Veränderung der Faktorstufen kann dies den Versuchsaufwand u. U. deutlich erhöhen. Die damit verbundenen Zusatzkosten und der zusätzliche Zeitbedarf dürfen bei der Aufwandsabschätzung nicht übersehen werden.

In begründeten Einzelfällen kann es aus Aufwandsgründen erforderlich sein, von Blockbildung und/oder voller Randomisierung abzusehen. Dies gilt vor allem, wenn

- alle Einzelversuche in relativ kurzer Zeit und unter konstanten Versuchsbedingungen durchgeführt werden können,
- die Faktorstufen genau und reproduzierbar eingestellt werden können, und
- der Zeitaufwand oder die Kosten für die Änderung der Faktorstufen (evtl. auch nur für bestimmte Faktoren) sehr hoch sind.

In solchen Fällen nimmt man bewusst in Kauf, dass die im Versuch ermittelte Zufallsstreuung evtl. kleiner ist als die Zufallsstreuung bei voller Randomisierung und ein evtl. vorhandener Trend die Schätzung für einen nur selten geänderten Faktor verfälschen kann. Also Vorsicht bei der Interpretation der Ergebnisse!



Beispiel

Beim Wellenlöten strebt man an, dass die meisten Leiterplatten fehlerfrei sind. Auch vor der Optimierung ist die Anzahl der Lötfehler pro Leiterplatte normalerweise schon relativ niedrig, sie unterliegt daher einer großen Zufallsstreuung. Um Unterschiede zwischen verschiedenen Einstellungen der Lötanlage zu erkennen, müssen somit viele Leiterplatten gelötet werden.

Bei unveränderter Einstellung der Lötanlage ist der Aufwand für das Löten zusätzlicher Leiterplatten nur sehr klein (wenige Sekunden). Andererseits ist der Aufwand für eine Veränderung der Temperatur des Lots sehr groß (Stunden). Daher sieht man meist von einer vollen Randomisierung ab, um den Aufwand zu begrenzen. Mit einer Einstellung der Lötanlage (Faktorstufenkombination) werden mehrere Leiterplatten nacheinander gelötet (d. h. keine Blockbildung wie im Beispiel aus der chemischen Industrie). Und bei der Randomisierung der Faktorstufenkombinationen achtet man meist darauf, dass die Temperatur des Lots nur selten (z. B. viermal) geändert werden muss.

Eine andere Art der Blockbildung ist beim Wellenlöten jedoch sehr einfach zu realisieren und daher empfehlenswert: Sollen später z. B. Leiterplatten verschiedener Hersteller oder unterschiedliche Typen mit derselben Einstellung der Anlage gelötet werden, so sollte man bei jeder Einstellung der Anlage hintereinander je eine (oder mehrere) Leiterplatte der verschiedenen Hersteller bzw. Typen löten.

Steht der Aufwand in einem angemessenen Verhältnis zum erhofften Ergebnis, so wird der Versuchsplan freigegeben und der Versuch kann durchgeführt werden. Erscheint der Aufwand zu groß, so ist zu untersuchen, ob

- durch Verzicht auf Faktoren oder Faktorstufen,
- durch Verzicht auf volle Blockbildung und/oder Randomisierung,
- durch eine kleinere Anzahl von Realisierungen

der Aufwand evtl. reduziert werden kann, ohne das Untersuchungsziel zu gefährden, oder ob ein weniger ehrgeiziges Ziel wesentlich kostengünstiger erreicht werden kann. Dann wird der Versuchsplan entsprechend modifiziert.

Manchmal ist das erhoffte Ergebnis auch so gering, dass man auf die Untersuchung völlig verzichten kann. Dann gibt es meist wichtigere Probleme, auf deren Lösung man sich konzentrieren sollte.

Diese Aufwandsabschätzung am Anfang ist ein wesentlicher Vorteil geplanter Versuche. Dadurch wird vermieden, dass durch ungeplantes Vorgehen nach dem Motto „Nur noch ein Versuch“ der Zeit- und/oder Kostenrahmen überschritten wird.

■ 3.5 Versuche durchführen

Statistik ist kein Ersatz für fehlende Sorgfalt. Im Gegenteil: Jeder einzelne Messwert wird zur Berechnung aller Effekte verwendet. Er beeinflusst daher auch alle Schlussfolgerungen. Ein unentdeckter Fehler kann alle Ergebnisse verfälschen. Die Einzelversuche im Versuchsplan müssen daher mit größtmöglicher Sorgfalt vorbereitet und durchgeführt werden.

3.5.1 Vorbereitung

Eine sorgfältige Planung und Vorbereitung gewährleistet einen reibungslosen und fehlerfreien Ablauf der Versuche. Dazu gehören u. a. folgende Schritte:

- Planung der erforderlichen Ressourcen:
Welche Anlagen, Messmittel, Rohmaterialien und Teile werden wann benötigt? Wer führt wann welche Arbeiten durch?
- Überprüfung der Messgeräte und Messverfahren:
Sind alle Messgeräte kalibriert und justiert? Ist die Messstreuung bekannt und ausreichend klein? Im Zweifel ist eine Messmittelanalyse eine wichtige und nützliche Vorbereitung, da der Einfluss von Probenaufbereitung und Messung oft deutlich unterschätzt wird.
- Auswahl, Kennzeichnung und Zuordnung der Teile:
Sind die Teile repräsentativ? Sind die Teile einheitlich bzw. sind vorhandene Unterschiede als Blockfaktor im Versuchsplan berücksichtigt? Sind die Teile gekennzeichnet, damit eine eindeutige Zuordnung zwischen Teil und Versuchsnummer gewährleistet ist? Sind die Teile den Versuchsnummern zufällig zugeordnet?
- Festlegung des Versuchs- und Messablaufs:
Ist der Versuchs- und Messablauf eindeutig festgelegt? Ist dadurch gewährleistet, dass die Umgebungsbedingungen und sonstigen Randbedingungen für alle Einzelversuche so einheitlich wie möglich sind? Werden die wichtigsten Umgebungsbedingungen erfasst und dokumentiert? Wie werden Zuordnungs- und Übertragungsfehler vermieden?
- Zuordnung und Einweisung des Personals:
Ist eindeutig festgelegt, wer welche Aufgaben hat (insbesondere eine Aufgabe = eine Person)? Ist allen Beteiligten klar, dass die vorgegebene Versuchsreihenfolge einzuhalten ist und dass jedes Einzelergebnis zur Berechnung aller Effekte benutzt wird und daher Sorgfalt und einheitliche Vorgehensweise extrem wichtig sind?
- Durchführung eines Pilotversuchs:
Wurde ein Pilotversuch (z. B. zwei extreme Einzelversuche) durchgeführt, um die Realisierbarkeit der geplanten Faktorstufenkombinationen zu überprüfen und den Versuchsablauf zu testen und einzubüben?

Falls in dieser Phase Probleme festgestellt werden, müssen sie vor der eigentlichen Versuchsdurchführung behoben werden. Dazu kann es erforderlich sein, den Versuchsplan anzupassen.

**Beispiel aus der chemischen Industrie (Fortsetzung)**

Da es sich um einen normalen Fertigungsprozess handelt, sind die Abläufe festgelegt und allgemein bekannt. Es muss nur überprüft werden, dass ausreichend Ausgangsmaterial einer Charge für 16 Füllungen des Reaktionsgefäßes vorhanden ist.

Ein Pilotversuch zeigt die Realisierbarkeit der extrem vom momentanen Zustand der Fertigung abweichenden Faktorstufenkombination Temperatur 140 °C, Zeit 4 h und Katalysator 0,5 %.

Der Anlagenbediener wird über Ziel und Bedeutung der Versuche informiert. Die weiteren Maßnahmen sind in Absatz 3.3.5 beschrieben.

3.5.2 Durchführung

Bei Versuchsdurchführung, Messung und Übertragung der Ergebnisse ist auf besondere Sorgfalt zu achten. Der Verantwortliche sollte immer dabei bleiben. Trotz aller Sorgfalt können sich bei der Versuchsdurchführung unvorhergesehene Zwischenfälle ereignen, die evtl. Einfluss auf die Ergebnisse haben. Es ist deshalb erforderlich, parallel zum Ablauf des Versuchs alle Besonderheiten und Abweichungen vom Plan zu dokumentieren. Bei der Auswertung werden die Ist-Werte der Faktoren berücksichtigt, nicht die Soll-Werte.

Wenn Umgebungsbedingungen oder Anlagenzustand nicht konstant gehalten werden können, ist es sinnvoll, Veränderungen aufzuzeichnen. Dann kann ihr Einfluss auf die Ergebnisse ggf. nachträglich erkannt werden.

**Beispiel aus der chemischen Industrie (Fortsetzung)**

Tabelle 3.4 zeigt den Versuchsplan aus Tabelle 3.3 mit den Ergebnissen. Bei der Versuchsdurchführung traten keine Besonderheiten auf.

TABELLE 3.4 Einzelversuche aus Tabelle 3.3 mit den Versuchsergebnissen

| Vers. Nr. | syst. Nr. | Realisierung (Block) | Temperatur [°C] | Zeit [h] | Katalysator [%] | Ausbeute [%] |
|-----------|-----------|-------------------------|--------------------|-------------|--------------------|-----------------|
| 1 | 8 | 1 | 140 | 4 | 0,5 | 68,5 |
| 2 | 3 | 1 | 120 | 4 | 0,1 | 56,7 |
| 3 | 1 | 1 | 120 | 2 | 0,1 | 52,8 |
| 4 | 7 | 1 | 120 | 4 | 0,5 | 56,5 |
| 5 | 6 | 1 | 120 | 2 | 0,5 | 62,2 |
| 6 | 2 | 1 | 140 | 2 | 0,1 | 61,5 |
| 7 | 4 | 1 | 140 | 4 | 0,1 | 67,9 |
| 8 | 5 | 1 | 120 | 2 | 0,5 | 53,6 |
| 9 | 4 | 2 | 140 | 4 | 0,1 | 70,2 |
| 10 | 3 | 2 | 120 | 4 | 0,1 | 55,2 |
| 11 | 8 | 2 | 140 | 4 | 0,5 | 67,2 |

| Vers. Nr. | syst. Nr. | Realisierung (Block) | Temperatur [°C] | Zeit [h] | Katalysator [%] | Ausbeute [%] |
|-----------|-----------|----------------------|-----------------|----------|-----------------|--------------|
| 12 | 5 | 2 | 120 | 2 | 0,5 | 54,1 |
| 13 | 1 | 2 | 120 | 2 | 0,1 | 54,1 |
| 14 | 6 | 2 | 140 | 2 | 0,5 | 62,9 |
| 15 | 7 | 2 | 120 | 4 | 0,5 | 54,6 |
| 16 | 2 | 2 | 140 | 2 | 0,1 | 61,8 |

■ 3.6 Versuchsergebnisse auswerten

Die statistische Auswertung der Versuchsergebnisse wird ab Kapitel 6 ausführlich beschrieben. Hier werden nur ausgewählte Aspekte kurz erläutert.

Bei vollständigen faktoriellen Versuchsplänen wird jede Faktorstufenkombination gleich oft realisiert. Für jede Faktorstufenkombination können die Mittelwerte der Versuchsergebnisse berechnet werden.

Durch Vergleich der Einzelergebnisse bei derselben Faktorstufenkombination kann man Ausreißer in den Ergebnissen erkennen. Ausreißer sind Ergebnisse, die offensichtlich nicht zu den anderen Ergebnissen passen. Sie können

- durch einen falsch eingestellten Faktor,
- durch falsches Ablesen eines Messgeräts,
- durch einen Übertragungsfehler o.ä.

verursacht werden, d. h. durch Fehler beim Versuch. Da Ausreißer das Ergebnis der Auswertung verfälschen würden, dürfen sie in der weiteren Analyse nicht berücksichtigt werden. Wenn möglich, sollte die Ursache für die Abweichung gesucht und das Ergebnis korrigiert oder der betroffene Einzelversuch wiederholt werden.

Ist eine Wiederholung des Einzelversuchs nicht möglich, wird dieser Wert gestrichen. Dabei muss man jedoch sehr vorsichtig sein. Insbesondere wenn die Faktorstufenkombination nur einmal realisiert wurde, kann ein Ausreißer die berechneten Effekte unbemerkt verfälschen. Umgekehrt kann ein scheinbarer Ausreißer auch auf bisher unbekannte, wichtige Einflüsse hinweisen.



Beispiel aus der chemischen Industrie (Fortsetzung)

Durch Umsortieren der Versuchsergebnisse von Tabelle 3.4 in die systematische Reihenfolge von Tabelle 3.2 erhält man Tabelle 3.5.

TABELLE 3.5 Ergebnisse für die Ausbeute in der systematischen Reihenfolge

| Nr. | Temperatur [°C] | Zeit [h] | Katalysator [%] | Einzelergebnisse [%] | | Mittel \bar{y}_i [%] |
|-----|-----------------|----------|-----------------|----------------------|------|------------------------|
| 1 | 120 | 2 | 0,1 | 52,8 | 54,1 | 53,45 |
| 2 | 140 | 2 | 0,1 | 61,5 | 61,8 | 61,65 |
| 3 | 120 | 4 | 0,1 | 56,7 | 55,2 | 55,95 |
| 4 | 140 | 4 | 0,1 | 67,9 | 70,2 | 69,05 |
| 5 | 120 | 2 | 0,5 | 53,6 | 54,1 | 53,85 |
| 6 | 140 | 2 | 0,5 | 62,2 | 62,9 | 62,55 |
| 7 | 120 | 4 | 0,5 | 56,5 | 54,6 | 55,55 |
| 8 | 140 | 4 | 0,5 | 68,5 | 67,2 | 67,85 |

Die Unterschiede zwischen den beiden Ergebnissen bei einer Faktorstufenkombination in Tabelle 3.5 sind alle vergleichbar groß. Es gibt daher keinen Hinweis auf Ausreißer.

Bei einem vollständigen faktoriellen Versuch gibt es je zwei Faktorstufenkombinationen, die sich nur in einem bestimmten Faktor unterscheiden. So unterscheiden sich in Tabelle 3.5 z. B. Nr. 1 und 2 bzw. Nr. 3 und 4, Nr. 5 und 6 und Nr. 7 und 8 jeweils nur in der Temperatur. Der Mittelwert dieser Unterschiede ist ein Maß für den Einfluss dieses Faktors auf das Versuchsergebnis und heißt Effekt oder auch Haupteffekt dieses Faktors (z. B. Effekt der Temperatur). Für jeden Faktor kann so der Effekt berechnet werden.



Beispiel aus der chemischen Industrie (Fortsetzung)

$$\begin{aligned}\text{Effekt der Temperatur} &= \frac{1}{4}((\bar{y}_2 - \bar{y}_1) + (\bar{y}_4 - \bar{y}_3) + (\bar{y}_6 - \bar{y}_5) + (\bar{y}_8 - \bar{y}_7)) = \\ &= \frac{1}{4}((61,65 - 53,45) + (69,05 - 55,95) + \dots) = 10,575\end{aligned}$$

Dieses Ergebnis bedeutet:

Bei 140 °C ist die Ausbeute im Mittel um 10,575 % höher als bei 120 °C.

$$\begin{aligned}\text{Effekt der Zeit} &= \frac{1}{4}((\bar{y}_3 - \bar{y}_1) + (\bar{y}_4 - \bar{y}_2) + (\bar{y}_7 - \bar{y}_5) + (\bar{y}_8 - \bar{y}_6)) = \\ &= \frac{1}{4}((55,95 - 53,45) + (69,05 - 61,65) + \dots) = 4,225\end{aligned}$$

Nach 4 h ist die Ausbeute im Mittel um 4,225 % höher als nach 2 h.

$$\begin{aligned}\text{Effekt des Katalysators} &= \frac{1}{4}((\bar{y}_5 - \bar{y}_1) + (\bar{y}_6 - \bar{y}_2) + (\bar{y}_7 - \bar{y}_3) + (\bar{y}_8 - \bar{y}_4)) = \\ &= \frac{1}{4}((53,85 - 53,45) + (62,55 - 61,65) + \dots) = -0,075\end{aligned}$$

Mit 0,5 % Katalysator ist die Ausbeute im Mittel um 0,075 % niedriger als mit 0,1 %.

Der Effekt eines Faktors auf das Versuchsergebnis hängt oft davon ab, welchen Wert (welche Stufe) ein anderer Faktor hat. Man spricht dann von einer Wechselwirkung der Faktoren.



Beispiel aus der chemischen Industrie (Fortsetzung)

Bei 4 h Reaktionszeit beträgt der Effekt der Temperatur:

$$= \frac{1}{2}((\bar{y}_4 - \bar{y}_3) + (\bar{y}_8 - \bar{y}_7)) = \frac{1}{2}((69,05 - 55,95) + (67,85 - 55,55)) = 12,7$$

Bei 2 h Reaktionszeit beträgt der Effekt der Temperatur dagegen nur:

$$= \frac{1}{2}((\bar{y}_2 - \bar{y}_1) + (\bar{y}_6 - \bar{y}_5)) = \frac{1}{2}((61,65 - 53,45) + (62,55 - 53,85)) = 8,45$$

Der Effekt der Temperatur ist der Mittelwert (siehe oben):

$$= \frac{1}{4}((\bar{y}_4 - \bar{y}_3) + (\bar{y}_8 - \bar{y}_7)) + ((\bar{y}_2 - \bar{y}_1) + (\bar{y}_6 - \bar{y}_5)) = \frac{1}{2}(12,7 + 8,45) = 10,575$$

Der Effekt der Wechselwirkung zwischen Temperatur und Zeit ist die Hälfte der Differenz:

$$= \frac{1}{4}((\bar{y}_4 - \bar{y}_3) + (\bar{y}_8 - \bar{y}_7)) - ((\bar{y}_2 - \bar{y}_1) + (\bar{y}_6 - \bar{y}_5)) = \frac{1}{2}(12,7 - 8,45) = 2,125$$

Dieses Ergebnis bedeutet:

Bei 4 h Reaktionszeit ist der Effekt der Temperatur um 2,125 % größer als im Mittel und um $2 \cdot 2,125\% = 4,25\%$ größer als bei 2 h Reaktionszeit.

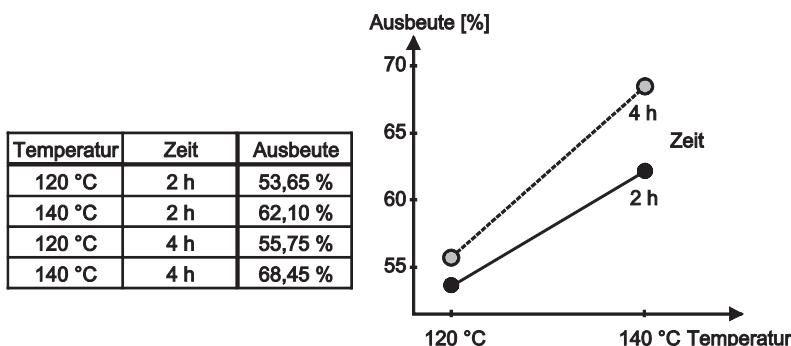


BILD 3-3 Mittelwerte der Versuchsergebnisse bei verschiedenen Temperaturen und Reaktionszeiten

Durch Umsortieren der Werte kann man auch zeigen:

Bei einer Temperatur von 140 °C ist der Effekt der Reaktionszeit um $2 \cdot 2,125\% = 4,25\%$ größer als bei 120 °C.

Bild 3-3 zeigt die Mittelwerte der Versuchsergebnisse bei den Faktorstufenkombinationen Temperatur 120 °C und Zeit 2 h, Temperatur 140 °C und Zeit 2 h usw. als Zahlenwerte und grafisch. In der grafischen Darstellung ist die Abweichung der beiden Linien von der Parallelität Ausdruck der Wechselwirkung zwischen Temperatur und Zeit.

Trotz aller Sorgfalt streuen Versuchsergebnisse aufgrund von zufälligen Einflüssen. Da die Effekte aus diesen Versuchsergebnissen berechnet wurden, streuen auch diese. Die aus den Versuchsergebnissen berechneten (beobachteten) Effekte weichen von den unbekannten wahren Werten zufällig ab. Allerdings treten große Abweichungen nur selten auf. Dies nutzt man, um die Signifikanz von Effekten zu beurteilen: Ist der Effekt größer als seine Zufallsstreuung (formal: die Breite seines Vertrauensbereichs), so nennt man ihn „signifikant“ – man ist überzeugt, dass der Effekt echt ist. Ist er kleiner als seine Zufallsstreuung, so nennt man ihn „nicht signifikant“ – das heißt nicht, dass er 0 ist, aber er könnte auch 0 sein (Einzelheiten in Kapitel 6 und 7).



Beispiel aus der chemischen Industrie (Fortsetzung)

Die Effekte der Temperatur (A), der Zeit (B) und der Wechselwirkung zwischen Temperatur und Zeit (AB) sind signifikant, der Effekt des Katalysators und der Wechselwirkungen zwischen Katalysator und Temperatur bzw. Zeit sind nicht signifikant.

■ 3.7 Ergebnisse interpretieren und Maßnahmen ableiten

Das Ergebnis der rein statistischen Analyse in Abschnitt 3.6 sind Zahlenwerte für die Größe der Effekte, die Breite der Vertrauensbereiche und daraus abgeleitet Aussagen über die Signifikanz der Effekte. Dieses Ergebnis muss jetzt technisch verstanden, interpretiert und in Verbesserungsmaßnahmen umgesetzt werden.

3.7.1 Interpretation

Ist die Wechselwirkung zwischen zwei Faktoren signifikant, so bedeutet dies, dass der Effekt des einen Faktors davon abhängt, welchen Wert (Stufe) der andere Faktor hat. Die beiden Faktoren müssen gemeinsam betrachtet werden. Da Wechselwirkungen als Zahlenwerte nur schwer zu interpretieren sind, wird empfohlen, Mittelwerte für alle Faktorstufenkombinationen dieser beiden Faktoren wie in Bild 3-3 zu berechnen und Konsequenzen aus diesen Mittelwerten abzuleiten.

Wichtig ist, dass das Ergebnis technisch plausibel ist. Die besten Ergebnisse sind diejenigen, bei denen im Nachhinein alle Beteiligten denken: Das ist doch eigentlich klar. Warum sind wir da nicht schon längst draufgekommen?



Beispiel aus der chemischen Industrie (Fortsetzung)

In Bild 3-3 sind alle signifikanten Effekte grafisch dargestellt:

- Der Effekt der Temperatur von 10,575 % bedeutet, dass der Mittelwert der beiden Ausbeuten bei 140 °C (rechts) um 10,575 % höher ist als bei 120 °C.

- Der Effekt der Reaktionszeit von 4,225 % bedeutet, dass der Mittelwert der beiden Ausbeuten bei 4 h (oben) um 4,225 % höher ist als bei 2 h.
- Der Effekt der Wechselwirkung ist die Abweichung von der Parallelität der beiden Linien (große Wechselwirkung heißt große Abweichung von der Parallelität). Da der Zahlenwert selbst aber ziemlich unanschaulich ist, wird bei signifikanten Wechselwirkungen immer eine Darstellung wie in Bild 3-3 empfohlen.

Es ist technisch plausibel, dass bei einer chemischen Reaktion die Ausbeute mit der Temperatur und der Reaktionszeit zunimmt. Es ist auch plausibel, dass der Effekt der Temperatur bei längerer Zeit größer ist als bei kürzerer Zeit (bei noch höheren Temperaturen bzw. Zeiten würde jedoch irgendwann eine Sättigung eintreten, eine Extrapolation der Ergebnisse ist nicht zulässig).

Die Katalysatormenge hat – anders als ursprünglich erwartet – keinen signifikanten Effekt. Das bedeutet, dass auch bei 0,1 % Katalysator bereits eine Sättigung erreicht ist. Da der Katalysator bei der Reaktion unverändert bleibt, ist auch dieses Ergebnis plausibel.

3.7.2 Maßnahmen

Plausibilität bezieht sich auf die qualitative Abhängigkeit der einzelnen Zielgrößen von den Faktoren. Dabei ist es normal, dass bei mehreren Zielgrößen Zielkonflikte auftreten. Was für eine Zielgröße günstig ist, ist manchmal für eine andere ungünstig. In dieser Situation hilft die quantitative Kenntnis der Abhängigkeit. Damit kann man gezielt Kompromisse finden und alle (widersprüchlichen) Ziele soweit möglich miteinander verbinden. Auch und gerade Kostenminimierung und Prozessvereinfachung sind wesentliche Ziele. Aus diesen quantitativen und technischen Betrachtungen resultiert schließlich ein vorläufiger Maßnahmenkatalog für konkrete Prozess- bzw. Produktverbesserungen. Werden aus den Ergebnissen keine Maßnahmen abgeleitet, so war der Aufwand verschwendet. Dabei kann natürlich im Einzelfall auch die Bestätigung der bisherigen Vorgehensweise ein wichtiges Ergebnis sein.



Beispiel aus der chemischen Industrie (Fortsetzung)

Die Erhöhung der Temperatur von 120 °C auf 140 °C erbringt im Mittel eine Erhöhung der Ausbeute von über 10 % und damit pro Füllung 1000 € Mehrerlös. Auch bei 2 h Reaktionszeit beträgt die Erhöhung der Ausbeute noch über 8 %. Die Mehrkosten betragen nur 50 €. Daher wird die Prozesstemperatur (zunächst probeweise) auf 140 °C erhöht.

Die Erhöhung der Katalysatormenge erbringt keine erkennbare Erhöhung der Ausbeute, würde jedoch zusätzliche Kosten verursachen. Daher wird die Katalysatormenge bei 0,1 % belassen.

Die Verlängerung der Reaktionszeit von 2 h auf 4 h erbringt im Mittel eine Erhöhung der Ausbeute von ca. 4 %, bei 140 °C sogar um 6 % und damit pro Füllung 600 € Mehrerlös. Die Mehrkosten betragen jedoch 1000 € pro Füllung. Daher wird die Reaktionszeit bei 2 h belassen.

■ 3.8 Absicherung, Dokumentation, weiteres Vorgehen

Auch nach dem Ableiten der Maßnahmen ist die Arbeit noch nicht abgeschlossen. Es muss gewährleistet werden, dass die Verbesserungen wirklich erreicht werden und auf Dauer erhalten bleiben.

3.8.1 Absicherung der Verbesserungen

Aus den Versuchsergebnissen wurden Maßnahmen zur Prozess- oder Produktverbesserung abgeleitet. Bevor diese Maßnahmen endgültig umgesetzt werden, ist ein Bestätigungsversuch bzw. eine länger andauernde Probeumstellung der Fertigung empfehlenswert. Damit werden folgende Aspekte überprüft:

- Bei Kompromissen zwischen widersprüchlichen Zielen und bei fraktionellen Plänen ist die empfohlene Faktorstufenkombination im Versuch evtl. nicht enthalten. Dann muss zunächst überprüft werden, ob die errechnete Verbesserung auch tatsächlich eintritt.
- Bei Versuchen wird besonders sorgfältig gearbeitet. Nicht immer bleibt die errechnete Verbesserung auch unter Fertigungsbedingungen bestehen. Evtl. vorhandene Unterschiede können frühzeitig erkannt und beseitigt werden.

3.8.2 Dokumentation

Zum Nachweis des Erfolgs und als Ausgangspunkt für weitere Verbesserungen muss der Versuch ausreichend dokumentiert werden. Dazu gehören:

- Darstellung der Ausgangslage
- durchgeführte Versuche und Ergebnisse
- abgeleitete Maßnahmen
- Darstellung des erreichten Zustands
- Kosten des Versuchs und Einsparung durch den Versuch.

Die systematische Vorgehensweise bei der Versuchsplanung erleichtert die Dokumentation wesentlich.



Beispiel aus der chemischen Industrie (Fortsetzung)

Ausgangslage:

- Ausbeute ca. 54 % ⇒ Erlös pro Füllung ca. 5400 €
- Kosten pro Füllung 5000 € ⇒ Gewinn pro Füllung ca. 400 €

Durchgeführte Versuche, Ergebnisse und abgeleitete Maßnahmen: siehe oben.

Erreichter Zustand:

- Ausbeute ca. 62 % \Rightarrow Erlös pro Füllung ca. 6200 €
- Kosten pro Füllung 5050 € \Rightarrow Gewinn pro Füllung ca. 1150 €

Kosten des Versuchs ca. 20 000 €

Zusätzlicher Gewinn bei unveränderter Anzahl Füllungen (etwas weniger bei unveränderter Produktmenge):

$$750 \text{ €} \cdot 500 = 375\,000 \text{ €/Jahr.}$$

3.8.3 Weiteres Vorgehen

In diesem Kapitel wurde der Durchlauf eines Verbesserungszyklus beschrieben. Danach kann, aufbauend auf den Ergebnissen des gerade abgeschlossenen Versuchs, ein neuer Zyklus durchlaufen werden, mit z.B.

- anderen Stufen für die Faktoren
- anderen Faktoren
- anderem Versuchsumfang
- manchmal auch weiteren Zielgrößen.

Oft ist das Problem aber auch schon nach einem Durchlauf gelöst und man kann sich anderen Aufgaben widmen. Auch diese Aufgaben können wieder mit Versuchsplanung gelöst werden und so sammelt man Schritt um Schritt Wissen und Erfahrung mit der Versuchsplanung.



Beispiel aus der chemischen Industrie (Ende)

Mit dem Wissen aus diesem Versuch kann man z.B. überlegen, ob es sich lohnt,

- den Katalysator auf 0,05 % zu reduzieren
- mit einer Reaktionszeit zwischen 2 und 4 h einen besseren Kompromiss zwischen Ausbeute und Anlagenkosten zu suchen (falls die Ausbeute nichtlinear von der Zeit abhängt)
- eine neue Anlage mit höherer Maximaltemperatur zu entwickeln
- den Einfluss des Drucks, des Mischungsverhältnisses der Ausgangsmaterialien oder anderer Einflussgrößen zu untersuchen, usw.

In Abschnitt 16.3 wird als Beispiel für eine vertiefende Untersuchung gezeigt, wie nachträglich die nichtlineare Abhängigkeit der Ausbeute von der Zeit bestimmt werden kann.

Literatur

- [1] *Flamm, R.J.*: Entwicklung eines Systemkonzeptes zur wissensbasierten, systemtechnisch unterstützten Versuchsmethodik. FQS Schrift, Beuth Verlag, Berlin 1995
- [2] *Holst, G.*: Systematisierung der Planungsphase der Statistischen Versuchsmethodik für die industrielle Anwendung. Verlag Shaker, Aachen 1995
- [3] *Gundlach, C.*: Entwicklung eines ganzheitlichen Vorgehensmodells zur problemorientierten Anwendung der statistischen Versuchsplanung. Kassel University Press, 2004
- [4] *Dean, A./Voss, D./Draguljic, D.*: Design and Analysis of Experiments. Springer Verlag, New York 2. Auflage 2017

4

Systematische Beobachtung

Die richtige Auswahl der Faktoren hat entscheidenden Einfluss auf den Erfolg einer Untersuchung. Die Bedeutung einer Einflussgröße, die im Versuch nicht verändert wird, kann auch nicht erkannt werden.

In Abschnitt 3.3 wurde gezeigt, wie mit Brainstorming wichtige Einflussgrößen als Faktoren ausgewählt werden können. Dabei ist man auf die Erfahrung und Objektivität der Beteiligten angewiesen. In diesem Kapitel werden vier Verfahren behandelt, die durch systematische Beobachtung nützliche Hinweise auf wichtige Einflussgrößen liefern. Grundlage ist das Motto von D. Shainin [1]: „Lasst nicht die Ingenieure raten, lasst die Teile sprechen.“

- Das Multi-Vari-Bild, die Darstellung der örtlichen Verteilung von Fehlern und der Prozessvergleich (Abschnitte 4.1 bis 4.3) helfen bei der Verbesserung von Fertigungsprozessen, insbesondere bei der Eingrenzung von Ursachen für Fertigungsprobleme.
- Der paarweise Vergleich (Abschnitt 4.4) hilft vor allem bei der Verbesserung von Produkten.

Voraussetzung für den Einsatz dieser Verfahren ist, dass bereits etwas Beobachtbares vorliegt. Daher liegt ihr Anwendungsschwerpunkt in der Fertigung und in späten Entwicklungsstadien. Aber auch für neue Produkte bzw. Prozesse kann die Beobachtung bereits bestehender ähnlicher Produkte bzw. Prozesse nützlich sein.

Die folgende Beschreibung soll Ideen für eine systematische Sammlung und Darstellung von Daten geben. Diese Ideen können mit gesundem Menschenverstand sicher noch erweitert werden.

■ 4.1 Multi-Vari-Bild

Das Multi-Vari-Bild wurde von L. Seder [2] entwickelt und von D. Shainin in seine Methodensammlung übernommen [3].

Ausgangspunkt ist eine Zielgröße, deren Streuung zu groß ist. Ziel des Multi-Vari-Bildes ist es, durch systematische Beobachtung der Fertigung (d.h. ohne Eingriff und Veränderung) die möglichen Streuursachen einzugrenzen. Diese ergeben dann die Faktoren für den anschließenden geplanten Versuch.

Dazu werden Messwerte für die Zielgröße von

- verschiedenen Stellen an einem Teil
- verschiedenen, nacheinander gefertigten Teilen
- zu verschiedenen Zeiten gefertigten Teilen

systematisch gesammelt und gegeneinander grafisch dargestellt.

Man beobachtet, ob der größte Unterschied z.B.

- zwischen verschiedenen Stellen an einem Teil
- zwischen Teilen von verschiedenen Positionen in einer Anlage (z.B. bei Mehrfachformen)
- zwischen Teilen von verschiedenen Anlagen (bei mehreren parallel eingesetzten Anlagen)
- zwischen aufeinanderfolgenden Teilen in einer Charge
- zwischen Fertigungschargen
- zwischen Losen des Ausgangsmaterials
- von Stunde zu Stunde, Schicht zu Schicht usw.

auftritt. Die Datensammlung wird fortgesetzt, bis mindestens 80 % der üblichen Streuung erfasst wurden.

Dann überlegt man, welche der vielen möglichen Ursachen (Einflussgrößen) mit dem beobachteten Verhalten konsistent sind. Nur diese können verantwortlich sein. Normalerweise kommen nur noch einige wenige Ursachen in Frage.



Beispiel aus der Metallbearbeitung (nach Bhote [3])

Der Durchmesser von Ankerwellen für Elektromotoren streute zu stark. Der Solldurchmesser betrug $0,250 \pm 0,001$ Zoll. Für ein Multi-Vari-Bild wurden folgende Daten erfasst:

- an einer Stelle: maximaler und minimaler Durchmesser (Unrundheit)
- an einem Teil (einer Welle): Durchmesser links und rechts
- zwischen Teilen zu einer Zeit: drei nacheinander gefertigte Teile
- zu verschiedenen Zeiten: fünf Stichproben im Abstand von einer Stunde.

Für jedes Teil wurden vier Messwerte erfasst (min-max, links-rechts). Bild 4-1 zeigt die Messwerte grafisch, wobei zusammengehörige Messwerte auch nebeneinander dargestellt sind (Multi-Vari-Bild).

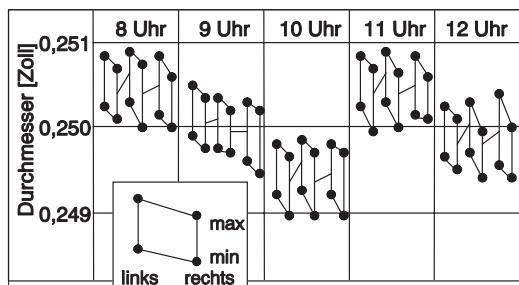


BILD 4-1
Multi-Vari-Bild:
Darstellung der vier Durchmesserwerte an
einem Teil an jeweils drei nacheinander ge-
fertigten Teilen, zu fünf verschiedenen Zeiten
entnommen.

Folgende Unterschiede fallen in Bild 4-1 auf:

1. Die Wellendurchmesser nehmen von 8 bis 10 Uhr kontinuierlich ab, nehmen dann sprunghaft zu, dann wieder ab. Von 10 Uhr 30 bis 11 Uhr ist Frühstückspause. D.h. nach dem Anschalten der Anlage sind die Werte hoch und nehmen dann im Betrieb ab. Dies kann auf Erwärmung der Anlage zurückzuführen sein. Eine Überprüfung ergab einen zu niedrigen Kühlmittelstand.
2. Die Wellendurchmesser rechts sind immer etwas kleiner als links. Eine Überprüfung ergab, dass das Werkzeug nicht parallel zur Achse ausgerichtet war.
3. Der Unterschied zwischen Minimal- und Maximaldurchmesser an einer Stelle ist sehr groß, d.h. die Welle ist unrund. Eine Überprüfung ergab ein ausgeschlagenes Lager.

Im vorliegenden Beispiel konnten die Ursachen identifiziert und mit geringem Aufwand beseitigt werden. Dies führte zu wesentlichen Kosteneinsparungen und zu einer Reduktion der Streuung um den Faktor 6.

Normalerweise führt das Multi-Vari-Bild nicht direkt zur Identifizierung der Ursachen selbst. Aber ein bestimmtes Muster im Multi-Vari-Bild ist nur mit manchen Ursachen konsistent und nicht mit anderen. Dadurch können viele Ursachen ausgeschlossen werden. Die anschließende Suche unter den verbliebenen Ursachen ist damit wesentlich direkter und zielführender.

Bei einer Versuchsplanung brauchen dann nur die Einflussgrößen als Faktoren berücksichtigt zu werden, die mit dem Multi-Vari-Bild konsistent sind. So erleichtert das Multi-Vari-Bild die Auswahl der Faktoren (vgl. Abschnitt 3.3).

Ziel der folgenden Aufgabe ist es, diese Einschränkung der möglichen Einflussgrößen zu verdeutlichen. Die Aufgabe sollte auch ohne Kenntnisse aus der Gießerei lösbar sein. Bitte, denken Sie über die vier beschriebenen Situationen nach, bevor Sie die Musterlösung lesen. Es kommt weniger auf die Einzelheiten als auf die Erkenntnis an, dass je nach Muster im Multi-Vari-Bild immer nur manche Ursachen in Frage kommen.



Aufgabe

In einer Gießerei werden bestimmte Teile gefertigt. Ihre Dichte ist eine wesentliche Zielgröße, da eine niedrige Dichte auf Einschlüsse, Poren und Lunker hinweist. Daher wird bisher die Dichte aller Teile vor der Weiterverarbeitung gemessen. Ist sie an einem Teil zu niedrig, so ist dieses Teil schlecht und wird wieder eingeschmolzen.

Der Anteil schlechter Teile ist zu hoch und soll gesenkt werden. Momentan erfolgt die Messung getrennt von der Fertigung. Daher sind viele Ursachen für die schlechten Teile denkbar. Um die Ursachen einzuschränken, wird die Dichte daher in vier Schichten (je 8 h) fertigungsbezogen erfasst und in einem Multi-Vari-Bild dargestellt.

Folgendes sei über den Fertigungsprozess bekannt:

- Die Teile werden in einer Zweifach-Form gefertigt, d.h. mit einer Füllung der Form werden gleichzeitig zwei Teile gegossen.
- Das Material stammt aus einem Vorratsbehälter, dessen Füllung jeweils für ca. 4 h reicht. Danach wird eine neue Füllung erschmolzen.

Für das Multi-Vari-Bild wird viermal pro Schicht (d.h. alle 2 h) die Dichte der jeweils 2 Teile von 3 aufeinanderfolgenden Füllungen der Form bestimmt. Dabei wird festgehalten, in welcher Form (1 oder 2) und bei welcher Füllung das Teil gegossen wurde.

Bild 4-2 zeigt als a, b, c und d vier verschiedene denkbare Ergebnisse. Analysieren Sie für jede der vier Möglichkeiten das Muster und überlegen Sie dann, welche Ursachen jeweils verantwortlich sein können.

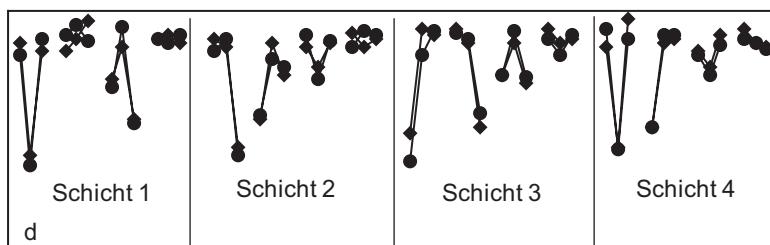
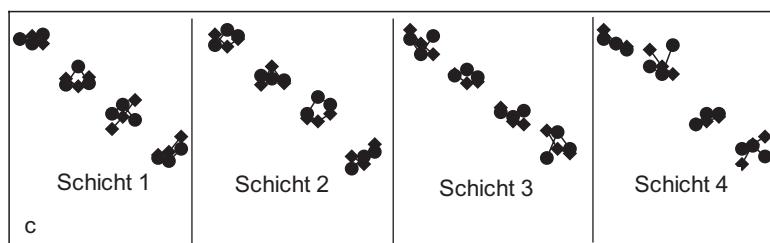
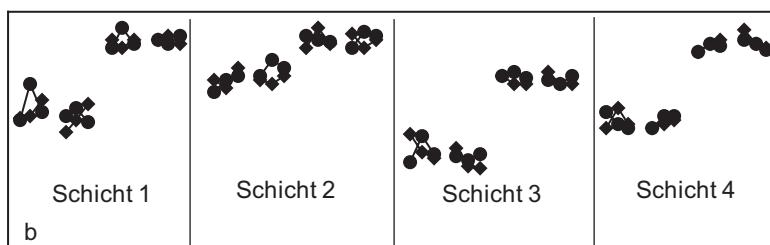
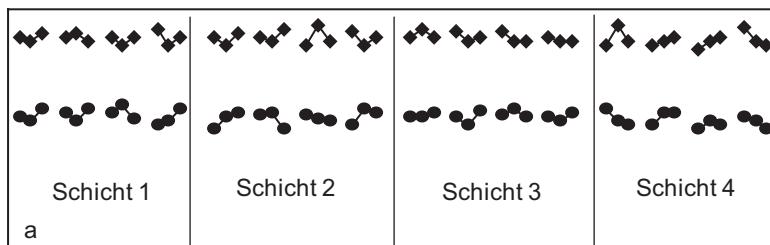


BILD 4-2 Vier mögliche Ergebnisse für das Multi-Vari-Bild:

- ◆ Form 1
- Form 2

Lösung

- a) Es besteht ein großer Unterschied zwischen den beiden Formen. Für jede Form einzeln streuen die Ergebnisse sehr wenig. Die Ursache muss in der Form liegen, z. B. in der Auslegung der Form, der Dynamik der Formfüllung o.ä.
- b) Es besteht kein Unterschied zwischen den beiden Formen. Immer zwei Stichproben sind ähnlich. Das ist gerade die Zeit, für die eine Füllung des Vorratsbehälters reicht. Die Ursache für die Streuung muss daher in Unterschieden zwischen den Füllungen der Vorratsbehälter liegen, z. B. Materialzusammensetzung, Aufbereitung der Füllung o.ä.
- c) Innerhalb einer Schicht fallen die Messwerte systematisch. Innerhalb der Gruppen von drei aufeinander folgenden Formfüllungen und zwischen den Formen gibt es keine Unterschiede. Die Ursache muss sich also innerhalb einer Schicht systematisch ändern, z. B. Ermüdung der Mannschaft, Erwärmung der Umgebung o.ä.
- d) Es besteht kein Unterschied zwischen den beiden Formen. Aber gelegentlich erhält man für einzelne Formfüllungen wesentlich niedrigere Werte als normal. Dies könnte z. B. an einer nicht richtig geschlossenen Form o.ä. liegen.

Zweck dieser Aufgabe war zu verdeutlichen, dass immer nur manche Ursachen mit dem Muster des Multi-Vari-Bildes konsistent sind. Die möglichen Ursachen werden eingegrenzt.

■ 4.2 Darstellung der örtlichen Verteilung von Fehlern

Werden in einem Fertigungsschritt viele Teile gleichzeitig nebeneinander hergestellt, so erhält man zusätzliche Information aus der örtlichen Verteilung von Fehlern bzw. Problemen. Ähnliches gilt, wenn ein ausgedehntes Teil mehrere Fehler enthalten kann.

Die Darstellung der örtlichen Verteilung von Fehlern ist eine Verfeinerung der Betrachtung der Streuung innerhalb eines Teiles, wie sie auch im Multi-Vari-Bild dargestellt wird.

Auch das Ziel ist dasselbe: Eine bestimmte Verteilung bzw. Änderung der Verteilung ist nur mit manchen Ursachen vereinbar. Aus der beobachteten Verteilung erhält man daher eine Eingrenzung der möglichen Ursachen. Die Auswahl der Faktoren für den folgenden Versuch wird erleichtert.



Beispiel aus der Leiterplattenfertigung

20 kleine Leiterplatten eines bestimmten Typs werden zum Löten auf einem Nutzen zusammengefasst, sie sind in fünf Reihen mit je vier Stück angeordnet. Dieser Nutzen wird mit der Unterseite in ein Lotbad eingetaucht, alle 20 Leiterplatten werden so gleichzeitig gelötet.

Manchmal werden einzelne Kontakte nicht gelötet, manchmal treten Brücken (Kurzschlüsse) zwischen Kontakten auf. Die Anzahl der Fehler insgesamt ist zu hoch und soll reduziert werden.

Zur Eingrenzung der Ursachen wird die Anzahl der Fehler für jede Leiterplatte einzeln erfasst. Bild 4-3 zeigt die Anzahl der Fehler je Leiterplatte für jeweils drei unmittelbar nacheinander gelötete Nutzen (untereinander), die an vier verschiedenen Tagen gelötet wurden.

| Tag | 1 | 2 | 3 | 4 |
|-----|---|---|---|---|
| Nr. | 1 0 0 0 3 0 0 1 2 0 0 0 1 0 0 0 1 0 0 0 | 1 1 2 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 0 0 2 0 1 0 1 0 0 0 2 0 0 0 4 0 0 0 1 | 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 |
| 1 | 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 | 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 | 1 0 0 0 3 0 0 0 5 0 0 0 2 0 0 0 2 0 0 0 |
| 2 | 1 3 4 3 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 | 0 0 0 1 0 0 0 4 0 0 0 2 0 0 0 1 0 0 0 0 | 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 | 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 |
| 3 | 0 | 0 | 0 | 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 |

BILD 4-3 Anzahl der Lötfehler je Leiterplatte an je drei nacheinander gelöteten 4 × 5 Nutzen (in Spalten übereinander), von vier verschiedenen Tagen.

Auffallend in Bild 4-3 ist, dass die Fehler überwiegend am Rand der Nutzen liegen, allerdings nicht immer am gleichen Rand (wenn überhaupt Fehler in nennenswertem Umfang auftreten). Dabei kann bei aufeinanderfolgenden Nutzen das Erscheinungsbild durchaus unterschiedlich sein. Zwischen den vier Tagen sind keine wesentlichen Unterschiede erkennbar.

Die Fehlerursache muss sich von Nutzen zu Nutzen ändern können und jeweils einen bestimmten Rand betreffen.

Als Ursache wurde schiefes Eintauchen des Nutzens in das Lotbad identifiziert. Wie die fehlerfreien Nutzen zeigen, kann das Problem mit guter Justierung vermieden werden.



Beispiel aus der Halbleiterfertigung

Auf einer Silizium-Einkristall-Scheibe werden gleichzeitig viele ICs (Integrierte Schaltungen) gefertigt, für einen bestimmten Typ z.B. ca. 100. Bereits ein kleiner Fehler im Bereich eines ICs führt zum Ausfall des betroffenen ICs. Fehler, die zum Ausfall führen, können Versetzungen im Einkristall, Defekte in aufgewachsenen oder aufgetragenen Schichten, die Unterbrechung einer Leiterbahn, Staubpartikel und vieles mehr sein.

Ziel ist es, die Ausbeute ständig schrittweise zu erhöhen, indem systematische Ausfallursachen identifiziert und beseitigt werden. Da der Fertigungsprozess aus mehreren hundert Einzelschritten besteht, ist die Identifizierung der Ausfallursachen ein langwieriger Prozess, der sich für eine neue IC-Generation jeweils über Jahre erstreckt. Dabei werden die verschiedensten Techniken eingesetzt, und die Ausbeute wird von anfangs niedrigen Werten schrittweise gesteigert. Hier soll einer dieser Schritte beschrieben werden.

Betrachtet man einzelne Scheiben, so sind die Ausfälle sehr unterschiedlich über die Scheiben verteilt. Aufgrund der großen Zufallsstreuung und der vielen beteiligten Mechanismen ist keine

Systematik erkennbar. Um einen Überblick über die Verteilung der Ausbeute über die Scheiben zu erhalten, wurde für jede Position auf der Scheibe die mittlere Ausbeute von ca. 1000 Scheiben berechnet (Wafermap). Dabei ergab sich grob die in Bild 4-4 dargestellte Verteilung. Dass die Ausbeute am Rand niedriger als in der Mitte war, entsprach den Erwartungen. Bei mehreren der Einzelprozesse ist die Partikeldichte am Rand höher als in der Mitte. Überrascht hat jedoch die noch niedrigere Ausbeute am Flat (das ist eine Abflachung der Scheibe an einer Seite, die der Ausrichtung der Scheiben bei manchen Prozessschritten dient).

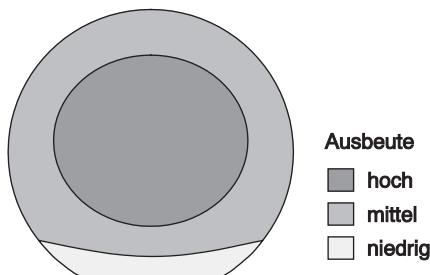


BILD 4-4 Schematische Darstellung der Ausbeuteverteilung über eine Scheibe

Was ist am Flat anders als am übrigen Rand? Als Ursache kamen nur Prozessschritte in Frage, bei denen die Scheibe ausgerichtet ist oder bei denen sich die Störung der Symmetrie am Flat bemerkbar macht.

Durch solche Überlegungen fiel der Verdacht auf einige Ofenprozesse, bei denen die Scheiben schnell erhitzt werden. Durch die Erhitzung entstehen thermische Spannungen, die Versetzungen im Einkristall auslösen können. Durch die Störung der Symmetrie am Flat sind die Spannungen dort besonders hoch. Besonders viele Versetzungen entstehen und sie reduzieren die Ausbeute am Flat.

Diese Überlegungen reichten natürlich nicht zur Identifizierung der genauen Ursache aus. Aber sie waren Anlass für geplante Versuche, bei denen als Faktoren die Aufheizraten bei mehreren kritischen Ofenprozessen verändert wurden (Stufe 1 jeweils: momentane Rate; Stufe 2: reduzierte Rate). Dabei zeigte sich, dass einer dieser Prozesse verantwortlich war. Die Aufheizrate für diesen Prozess wurde reduziert. Dadurch wurde die Ausbeute am Flat deutlich erhöht. Auch im übrigen Randbereich wurde eine (weniger deutliche) Erhöhung der Ausbeute erreicht. Insgesamt wurde die Ausbeute um ca. 5 % erhöht.

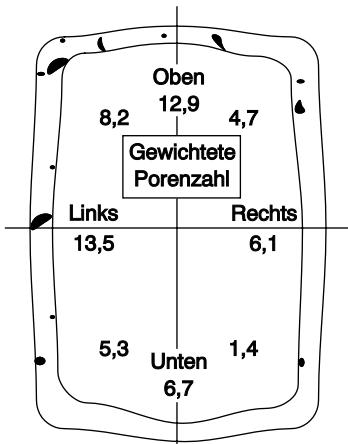


Beispiel aus einer Gießerei (nach Traver [4])

In einer Gießerei wurden Dichtringe gegossen. Aus unbekannter Ursache verschlechterte sich die Qualität. Ein hoher Anteil der Ringe musste wegen zu großer Poren wieder eingeschmolzen werden.

Zur Quantifizierung wurde ein Bewertungsmaßstab für die Porengröße eingeführt (von „1: sehr klein“ bis „5: riesiges Loch“). Die mit der Porengröße multiplizierte (gewichtete) Porenzahl wurde im Stile eines Multi-Vari-Bildes jeweils an mehreren hintereinander gegossenen Ringen zu verschiedenen Zeiten erfasst. Die mittlere gewichtete Porenzahl pro Ring betrug 19,7 und änderte sich nicht wesentlich von Ring zu Ring und Zeit zu Zeit.

Bild 4-5 zeigt die mittlere gewichtete Porenzahl je Quadrant und Hälfte und die Verteilung der Poren für einen typischen Ring. Traver schlägt vor, die Defekte rot zu kennzeichnen und nennt das Ergebnis dann „Masern-Diagramm“ [4]. Entscheidend ist die detaillierte Betrachtung der örtlichen Verteilung der Defekte innerhalb eines Teils. Aus ihr erkennt man klar die Häufung der Poren oben links.

**BILD 4-5**

Mittlere gewichtete Porenzahl je Quadrant und Ringhälfte und Anordnung der Poren in einem typischen Beispiel. Links oben ist die mittlere Porenzahl am höchsten (8,2), rechts unten am niedrigsten (1,4).

Was war oben links anders als in den übrigen Bereichen? Ein Ventilator blies von oben links Kühlluft in den Arbeitsbereich, um die Arbeitstemperatur erträglich zu halten und die Formen schneller abzukühlen. Die Verteilung der Poren gab damit einen Hinweis darauf, dass die Kühlung entscheidenden Einfluss auf die Porenbildung hat. In geplanten Versuchen wurde anschließend die Kühlung optimiert und das Problem beseitigt. Die Beobachtung der Fertigung ergab die entscheidenden Hinweise zur Eingrenzung der Ursachen. Die wichtigen Faktoren konnten ausgewählt werden.

■ 4.3 Prozessvergleich

Der Prozessvergleich [5] dient zur Fehleranalyse bei Prozessen mit sporadisch auftretenden Problemen. Man vergleicht die Fertigungsbedingungen und Ergebnisse von Inline-Kontrollen während der Fertigung bei guten und schlechten Produkten. Wiederkehrende Unterschiede können auf Problemursachen hinweisen.

Wichtig ist, dass die richtigen Größen in der Fertigung über einen ausreichend langen Zeitraum erfasst werden. Als Daumenregel kann gelten, dass das Problem im Erfassungszeitraum ca. achtmal auftritt, wobei die Einzelereignisse unabhängig voneinander sein müssen. Bei einem Chargenprozess bedeutet dies, dass acht Chargen betroffen sein sollten.



Beispiel Vakuumverpackung (nach Bhote [5])

Blattsalat wurde in Plastiktüten abgefüllt, evakuiert und versiegelt. Bei einem Teil der Packungen war die Versiegelung nicht dicht, was zum Verderb der Ware führte. Zur Prozessverbesserung sollte geklärt werden, welche Prozessparameter für die schlechte Abdichtung verantwortlich sind.

Durch Befragung der Prozessexperten wurde zunächst folgende Liste potentiell wichtiger Einflussgrößen zusammengestellt:

- Lage der Absaugdüse relativ zur Füllung der Tüten
- Absaugdauer

- Haltezeit nach der Absaugung
- Ausfahrgeschwindigkeit der Absaugdüse
- Versiegelungstemperatur
- Haltezeit der Temperatur
- Anpressdruck
- Abkühlzeit
- Auswurfverzögerung

Diese Größen wurden durch Fertigungsbeobachtung über einen ausreichend langen Zeitraum erfasst, in dem mehrfach Probleme mit undichten Packungen auftraten. Ein Vergleich der Werte der Einflussgrößen bei dichten und undichten Packungen ergab, dass bei den undichten Packungen die Absaugdüse mehr oder weniger tief in die Füllung eintauchte. Packungen, bei denen sich die Absaugdüse oberhalb der Füllung befand, waren dicht.

Wenn die Absaugdüse eintauchte, wurden Blätter in den Bereich der Versiegelung verschleppt, die zu Undichtigkeiten führten. Die Ursache war somit identifiziert und konnte beseitigt werden.



Beispiel aus der Halbleiterfertigung (2)

Bei der Halbleiterfertigung durchläuft jedes Fertigungslos über einen Zeitraum von mehreren Wochen mehrere hundert genau festgelegte Einzelprozessschritte. Für jeden Prozessschritt stehen parallel gleichartige Anlagen zur Verfügung, die je nach momentaner Belegung eingesetzt werden.

In einer Datenbank wird für jedes Los hinterlegt, wann jeder Prozessschritt an welcher Anlage durchgeführt wurde und was die zugeordneten Werte der kritischen Anlagenparameter waren. Ergänzend werden Prozessergebnisse wie Schichtdicke, Reflektivität u.ä. an mitlaufenden Teststrukturen erfasst und abgespeichert.

Erst am Ende der Scheibenfertigung kann die Funktion der Bauelemente überprüft werden. Findet man nun Lose mit niedrigerer Ausbeute als normal oder mit anderen erkennbaren Problemen, so wird sofort versucht, mit Hilfe des Fehlerbildes bei elektrischen und mikroskopischen Untersuchungen die Ursachen einzuziehen.

Parallel dazu findet ein Prozessvergleich mit Hilfe der Datenbank statt. Gibt es mehrere Lose mit demselben Problem, so wird gesucht, für welche Prozessschritte diese Lose über dieselbe Anlage liefen, insbesondere innerhalb kurzer Zeit. Außerdem werden die den schlechten Losen zugeordneten Messwerte für Prozessergebnisse wie Schichtdicke usw. mit denen der guten Lose verglichen.

Die Analyse des Fehlerbildes kombiniert mit dem Prozessvergleich führt normalerweise sehr schnell zum Erkennen und Abstellen der Problemursache.

In diesem Beispiel kann aufgrund der langen Durchlaufzeit der Lose mit der Datenerfassung nicht bis zum Auftreten eines Problems gewartet werden. Daher werden alle für eine spätere Analyse evtl. wichtigen Daten vorbeugend erfasst, um bei Auftreten eines Problems den Prozessvergleich sofort durchführen zu können.

■ 4.4 Paarweiser Vergleich von Produkten

Der Paarweise Vergleich [3] dient zur Fehleranalyse bei Produkten. Man geht folgendermaßen vor:

- Mehrere Paare aus je einem guten und einem schlechten Teil mit möglichst ähnlicher Vergangenheit werden gesammelt und paarweise auf Unterschiede untersucht.
- Sich wiederholende Unterschiede könnten die Ursache für das Problem sein.

In günstigen Fällen führt der Paarweise Vergleich direkt zur Identifizierung der Problemursache, sonst ist er eine Hilfe bei der Auswahl der Faktoren für einen anschließenden Versuch.

Beispiel Elektronikmodul (nach Bhote [3])

Eine bestimmte Diode im Elektronikmodul eines Autos fiel zu häufig aus. Mehrere ausgefallene Dioden wurden von den Werkstätten zurückgeliefert, zusammen mit nicht ausgefallenen Dioden des selben Typs, möglichst aus demselben Modul. Gute und schlechte Einheiten wurden paarweise einer REM-Prüfung (Raster-Elektronen-Mikroskop) unterzogen. Tabelle 4.1 zeigt die festgestellten Unterschiede.

| Paar Nr. | Gute Diode | Schlechte Diode |
|----------|------------|--|
| 1 | fehlerfrei | beschädigtes Plättchen, Oxidschäden, Kupfermigration |
| 2 | fehlerfrei | Legierungsunregelmäßigkeiten, Oxidschäden |
| 3 | fehlerfrei | Oxidschäden, Verunreinigung |
| 4 | fehlerfrei | Oxidschäden, beschädigtes Plättchen |

TABELLE 4.1

Paarweiser Vergleich von Dioden: Oxidschäden sind wahrscheinlich die Ursache für die Ausfälle

Beispiel Stoßdämpfer (nach Quentin [6])

Bei einem bestimmten Stoßdämpfertyp häuften sich die Reklamationen. Vergleiche der defekten Stoßdämpfer mit der Spezifikation zeigten keine Abweichungen. Daher wurde ein Paarweiser Vergleich durchgeführt.

Um möglichst ähnliche Vergleichsteile zur Verfügung zu haben, wurden für eine begrenzte Zeit bei einer Reklamation immer beide Stoßdämpfer einer Achse ausgebaut. Anschließend wurden der gute und der schlechte Stoßdämpfer jeweils paarweise miteinander verglichen. Sie hatten die gleiche Belastung im Einsatz gesehen, evtl. Unterschiede sollten daher Aufschluss über die Ursache der Reklamationen geben.

Neun Stoßdämpferpaare wurden untersucht. Dabei zeigte es sich, dass bei den schlechten Stoßdämpfern sieben undicht waren, bei den guten dagegen nur zwei. Alle anderen Unterschiede waren wesentlich seltener.

Eine genauere Untersuchung der undichten Stoßdämpfer zeigte, dass in den meisten Fällen die Dichtung beschädigt und die Oberflächenrauheit der Kolbenstange relativ groß war.

Die Spezifikation der Rauheit betrug $R_a < 6 \mu\text{m}$. Alle untersuchten Stoßdämpfer erfüllten diese Spezifikation. Von den insgesamt neun undichten Stoßdämpfern hatten jedoch sieben eine Rauheit im

Bereich $3 \mu\text{m} < R_a < 6 \mu\text{m}$, nur zwei hatten eine geringere Rauheit. Von den dichten Stoßdämpfern hatten alle eine Rauheit $R_a < 3 \mu\text{m}$.

Als wirkliche Ursache für die Reklamationen wurde so die zu große Rauheit der Kolbenstange identifiziert. Die Produktspezifikation wurde auf $R_a < 3 \mu\text{m}$ geändert, und das Problem war beseitigt.

Literatur

- [1] *Shainin, D./Shainin, P.*: „Better than Taguchi Orthogonal Tables“, in: Quality and Reliability Engineering International 4 (1988), 143 – 149
- [2] *Seder, L. A.*: „Diagnosis with Diagrams I bzw. II“, in: Industrial Quality Control (Januar bzw. März 1950)
- [3] *Bhote, K. R.*: Qualität – Der Weg zur Weltspitze. IQM, Großbottwar 1990
- [4] *Traver, R. W.*: Manufacturing Solutions for Consistent Quality and Reliability. AMACOM, New York 1995
- [5] *Bhote, K. R., Bhote, A. K.*: World Class Quality. AMACOM, New York 2. Auflage 2000
- [6] *Quentin, H.*: Versuchsmethoden im Qualitäts-Engineering. Vieweg, Braunschweig 1994

5

Einfache Versuche

D. Shainin hat eine einfache Versuchsstrategie beschrieben. Sie erlaubt es, mit geringem Aufwand die wichtigsten Faktoren zu identifizieren, deren Veränderung für die Streuung eines Fertigungsprozesses bzw. für den Ausfall eines Produktes verantwortlich ist [1 - 3].

Wichtigste Voraussetzung für einen Erfolg der Strategie ist, dass die Zufallsstreuung sehr viel kleiner als der Effekt der Faktoren ist. Nur dann kann auf mehrmalige Realisierung und statistische Auswertung verzichtet werden.

Je nach Anwendungsgebiet (und damit der Art der Faktoren) verwendet Shainin verschiedene Bezeichnungen:

- Prozessverbesserung: Variablenvergleich (Abschnitt 5.1)
- Produktverbesserung: Komponententausch (Abschnitt 5.2).

Shainin nennt den wichtigsten Faktor „Rotes X“. Ziel ist es, dieses „Rote X“ zu erkennen.

■ 5.1 Variablenvergleich zur Prozessverbesserung

Ziel des Variablenvergleichs ist, unter einer begrenzten Anzahl von Faktoren diejenigen zu erkennen, deren Veränderung den Hauptbeitrag zur Streuung der Zielgröße erbringt. Als Faktoren werden Größen betrachtet, die sich im Verlauf der Fertigung innerhalb gewisser Grenzen verändern, wie z.B.

- Prozessparameter (innerhalb ihrer Spezifikation)
- Umgebungsbedingungen (innerhalb ihrer natürlichen Grenzen)
- Unterschiede im Ausgangsmaterial, z. B. Charge, Alter, Lieferant

Ausgangspunkt:

- Die (bis ca. 20) Faktoren A, B, C, D, E, ... stehen im Verdacht, die Streuung der Zielgröße zu verursachen.
- Für jeden Faktor gibt es einen vermutlich „guten“ Wert (für A sei das Ag) und einen vermutlich „schlechten“ Wert (As). Beide Werte (= Stufen) müssen realistisch sein, d.h. in der Fertigung auch wirklich auftreten.

In Bild 5-1 wird der Variablenvergleich schematisch dargestellt. Als erster Schritt werden Vorversuche durchgeführt, um zu erkennen, ob die ausgewählten Faktoren wirklich fast die gesamte Streuung verursachen:

- zwei Einzelversuche mit allen Faktoren auf dem „guten“ Wert (Ergebnisse G₁, G₂)
- zwei Einzelversuche mit allen Faktoren auf dem „schlechten“ Wert (Ergebnisse S₁, S₂).

Die Differenz D zwischen den Mittelwerten der beiden „guten“ und der beiden „schlechten“ Ergebnisse ist ein Maß für den Effekt der Faktoren:

$$D = \left| \frac{G_1 + G_2}{2} - \frac{S_1 + S_2}{2} \right| \quad (5.1)$$

Der Mittelwert d der Unterschiede zwischen den beiden „guten“ bzw. „schlechten“ Ergebnissen ist ein Maß für die Zufallsstreuung:

$$d = \frac{|G_1 - G_2| + |S_1 - S_2|}{2} \quad (5.2)$$

Die Zufallsstreuung ist so klein (im Vergleich zum Effekt der Faktoren), dass keine mehrmalige Realisierung und keine formale statistische Analyse erforderlich ist, wenn

$$D : d \geq 5 : 1 \quad (5.3)$$

Wird diese Bedingung nicht erfüllt, so bestehen folgende Möglichkeiten:

- Der wichtigste Faktor (das „Rote X“) ist nicht im Versuch enthalten. Manchmal hilft eine bessere Auswahl der Faktoren.
- Die Zuordnung „gut“ und „schlecht“ ist nicht richtig. Manchmal hilft das Vertauschen oder die Änderung von einzelnen Werten.
- Die Zufallsstreuung durch andere Ursachen ist zu groß. Dann ist die Vereinfachung von D. Shainin zu grob, und die in den folgenden Kapiteln beschriebene (klassische) Versuchsplanung muss eingesetzt werden. Da der Aufwand für die vier Vorversuche klein im Vergleich zur möglichen Einsparung bei einem Erfolg ist, lohnen sie sich trotzdem.

Wenn die Bedingung D : d $\geq 5 : 1$ erfüllt ist, kann der eigentliche Variablenvertausch beginnen. Für die Faktoren A, B, C, D, ... werden nun einzeln die Werte von „gut“ und „schlecht“ vertauscht.

Zunächst wird ein Einzelversuch durchgeführt, bei dem alle Faktoren außer A auf „gut“ gehalten werden, nur A hat den Wert „schlecht“ (Bezeichnung in Bild 5-1 AsRg = A schlecht, Rest gut). Und ein Einzelversuch wird durchgeführt, bei dem alle Faktoren außer A auf „schlecht“ gehalten werden, nur A hat den Wert „gut“ (Bezeichnung AgRs = A gut, Rest schlecht).

Für alle Faktoren werden nacheinander diese beiden Einzelversuche durchgeführt (ähnlich zu One-factor-at-a-time). Führt das Vertauschen von guter und schlechter Stufe zu keiner wesentlichen Veränderung der Zielgröße, so ist dieser Faktor unwichtig. In Bild 5-1 sind A bis D unwichtig. Der Faktor E dagegen vertauscht gut und schlecht. Der Faktor E ist dominant („Rotes X“). Der Faktor H liefert einen wesentlichen Beitrag, ist aber nicht alleinige Ursache („Rosa X“).

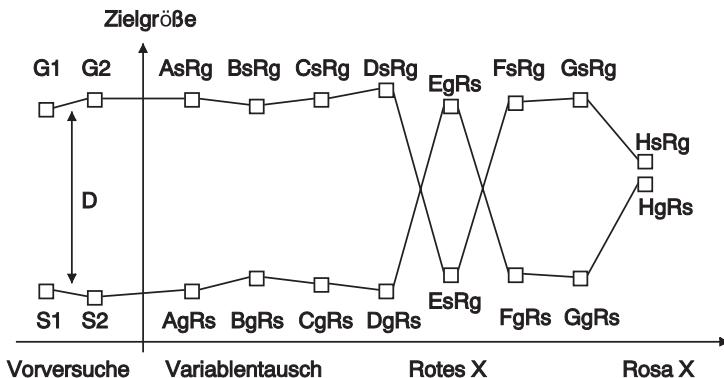


BILD 5-1 Grafische Darstellung der Ergebnisse der vier Vorversuche und des Variablentauschs

Mit dem Variablenvergleich können auch Wechselwirkungen erkannt werden (im Gegensatz zu einer reinen One-factor-at-a-time-Vorgehensweise).

Ändert sich bei einem Vertauschen nur eines der Ergebnisse der Zielgröße (gut oder schlecht, D bzw. G in Bild 5-2), so weist dies darauf hin, dass dieser Faktor eine wichtige Wechselwirkung mit einem anderen Faktor hat.

D. Shainin schlägt vor, nach Abschluss des Variablentauschs alle Faktoren, die als „Rotes X“, „Rosa X“ und Wechselwirkungen identifiziert wurden, in einem faktoriellen Versuch zu untersuchen (siehe Kapitel 7).

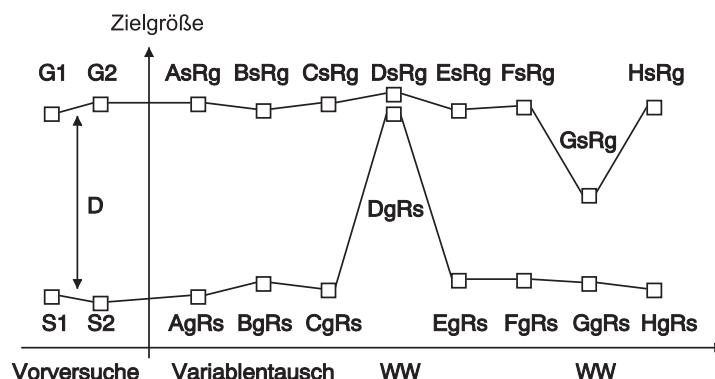


BILD 5-2 Grafische Darstellung der Ergebnisse bei Wechselwirkungen



Beispiel Metallpresse (nach Bhote [2])

In einer Presse werden Metallteile umgeformt. Die Toleranz für ein bestimmtes kritisches Maß beträgt $\pm 0,005$ Zoll. Manchmal streuen die Teile jedoch wesentlich mehr. In einem Brainstorming wurden die Faktoren in Tabelle 5.1 als mögliche Ursachen gesammelt.

Die Streuung des kritischen Maßes ist die Zielgröße. Um sie zu ermitteln, werden jeweils fünf Teile gefertigt, und die Differenz zwischen dem größten und kleinsten der fünf Werte (Spannweite) wird verwendet.

| Faktor | gut | schlecht |
|-------------------------|--------------|--------------------|
| A: Ausrichtung der Form | ausgerichtet | nicht ausgerichtet |
| B: Metalldicke | dick | dünn |
| C: Metallhärte | hart | weich |
| D: Metallbiegung | flach | gebogen |
| E: Stößelaufnahme | kalibriert | mit Spiel |
| F: Halten des Materials | waagrecht | nicht waagrecht |

TABELLE 5.1

Faktoren und vermutete gute und schlechte Stufen im Beispiel Metallpresse

Ergebnisse des Vorversuchs (in 0,001 Zoll)

alle Faktoren gut 4 4 Spannweite

alle Faktoren schlecht 47 61 Spannweite

Voraussetzung für weiteres Vorgehen überprüfen:

$$D = \left| \frac{4+4}{2} - \frac{47+61}{2} \right| = 50 \quad \text{und}$$

$$d = \frac{|4-4|}{2} + \frac{|47-61|}{2} = 7$$

$$\Rightarrow D : d = 50 : 7 > 5 : 1.$$

Der Unterschied zwischen den Ergebnissen bei „alle Faktoren gut“ und „alle Faktoren schlecht“ ist so groß, dass der eigentliche Variablenvergleich ohne Wiederholungen durchgeführt werden kann. Signifikante Effekte sind auch ohne formale Signifikanztests („mit bloßem Auge“) erkennbar. Tabelle 5.2 zeigt die Ergebnisse des eigentlichen Variablenvergleichs.

| Versuch | Kombination der Stufen | Spannweite [0,001 Zoll] | Schlussfolgerung |
|---------|------------------------|-------------------------|------------------|
| 1 | AsRg | 3 | A unwichtig |
| 2 | AgRs | 102 | |
| 3 | BsRg | 5 | B unwichtig |
| 4 | BgRs | 47 | |
| 5 | CsRg | 7 | C unwichtig |
| 6 | CgRs | 72 | |
| 7 | DsRg | 23 | Rosa X |
| 8 | DgRs | 30 | |
| 9 | EsRg | 7 | unklar |
| 10 | EgRs | 20 | |
| 11 | FsRg | 73 | Rotes X |
| 12 | FgRs | 18 | |
| Test | DsFsRg DgFgRs | 70 4 | völlige Umkehr |

TABELLE 5.2

Ergebnisse des Variablenvergleichs im Beispiel Metallpresse

Ergebnisse für die praktische Umsetzung:

- Die Materialbiegung und das Halten des Materials sind kritisch.
Eine Vorrichtung wurde gebaut, damit der Maschinenbediener das Material immer waagrecht hält; dadurch wurden bedienerbedingte Variationen beseitigt.
- Die Materialdicke und -härte sind nicht wichtig, daher konnten Toleranzen erweitert werden.

Mit nur einer Versuchsreihe wurde die Streuung um den Faktor 5 reduziert.

Der Variablenvergleich ist ein effizientes Verfahren zur einfachen Verbesserung von Fertigungsprozessen. Er ist bei der (Weiter-)Entwicklung der Prozesse und in der Fertigung einsetzbar, solange die Zufallsstreuung ausreichend klein ist.

■ 5.2 Komponententausch zur Produktverbesserung

Der Komponententausch ist im Grunde ein Variablenvergleich, angewendet auf Produkte, die zerlegt und wieder zusammengebaut werden können. Der Ausgangspunkt ist ein gutes und ein schlechtes Produkt. „Schlecht“ bedeutet, der Messwert für eine bestimmte Zielgröße liegt außerhalb der Toleranz. Ziel ist es, die Komponente zu identifizieren, die für den Ausfall des schlechten Produktes verantwortlich ist.

Die Komponenten entsprechen den Faktoren, die Stufe „gut“ sind die Komponenten des guten Produkts, die Stufe „schlecht“ sind die Komponenten des schlechten Produkts, und jeder Einzelversuch besteht darin, das Produkt zu zerlegen und wieder neu zusammenzubauen.

Daraus ergibt sich folgende Vorgehensweise:

Vorversuche durchführen:

- Zielgröße am „guten“ Produkt messen, dann Produkt zerlegen, wieder zusammenbauen und Zielgröße wieder messen (Ergebnisse G_1, G_2).
- Zielgröße am „schlechten“ Produkt messen, dann Produkt zerlegen, wieder zusammenbauen und Zielgröße wieder messen (S_1, S_2).

Voraussetzung für weiteres Vorgehen überprüfen:

Wie beim Variablenvergleich muss auch beim Komponententausch der Unterschied zwischen dem „guten“ und dem „schlechten“ Produkt ausreichend groß sein:

- Differenz zwischen „gut“ und „schlecht“ berechnen:

$$D = \left| \frac{G_1 + G_2}{2} - \frac{S_1 + S_2}{2} \right| \quad (5.1)$$

- Maß für die Streuung berechnen:

$$d = \frac{|G_1 - G_2| + |S_1 - S_2|}{2} \quad (5.2)$$

- Die Zufallsstreuung ist so klein, dass keine statistische Analyse nötig ist, wenn

$$D:d \geq 5:1. \quad (5.3)$$

Wird dies nicht erreicht, so kann auch die Montage für den Unterschied verantwortlich sein.

Wenn $D:d \geq 5:1$ erfüllt ist, werden für die Komponenten A, B, C, D, ... einzeln ausgetauscht. Nach dem Austausch wird an beiden Produkten die Zielgröße gemessen, dann wird wieder zurückgetauscht. Die Bewertung erfolgt wie in Bild 5-1 und 5-2 beim Variablenvergleich.



Beispiel Zeitzähler (nach Bhote [2])

Ein Zeitzähler soll bis zu einer Temperatur von -40°C störungsfrei arbeiten. Alle Produkte arbeiten zwar bei 0°C störungsfrei, viele fallen jedoch bei Temperaturen um -5°C bereits aus. Als Zielgröße wird daher die Temperatur festgelegt, bei der ein Produkt gerade noch funktioniert.

Der Zeitzähler besteht aus einer Elektronik, die Zählimpulse erzeugt, und mehreren mechanischen Teilen, die für die Anzeige benötigt werden.

Ergebnisse des Vorversuchs (in $^{\circ}\text{C}$)

gutes Produkt vor: -40 nach Zerlegen und Zusammenbau: -35

schlechtes Produkt vor: 0 nach Zerlegen und Zusammenbau: -5

Voraussetzung für weiteres Vorgehen überprüfen:

$$D = \left| \frac{(-40) + (-35)}{2} - \frac{0 + (-5)}{2} \right| = 35 \quad \text{und}$$

$$d = \frac{|(-40) - (-35)|}{2} + \frac{|0 - (-5)|}{2} = 5$$

$$\Rightarrow D:d = 35:5 = 7:1 > 5:1.$$

Tabelle 5.3 zeigt die wichtigsten Einzelkomponenten und Tabelle 5.4 die Ergebnisse des eigentlichen Komponententauschs (grafische Darstellung in Bild 5-3).

| Kennbuchstabe | Komponente |
|---------------|------------------------------|
| A | Zylinderspule mit Magnetkern |
| B | Zwischenradwelle |
| C | Ziffernwelle |
| D | Gehäuse |
| E | Kniehebel |
| F | Zwischenräder |
| G | Ziffernscheiben |
| H | Elektronik |

TABELLE 5.3

Liste der wichtigsten Komponenten des Zeitzählers

| Versuch Nr. | Kombination der Komponenten | Ergebnis [°C] | Schlussfolgerung |
|-------------|-----------------------------|---------------|------------------|
| 1 | AsRg | - 40 | A unwichtig |
| 2 | AgRs | - 5 | |
| 3 | BsRg | - 35 | B unwichtig |
| 4 | BgRs | 0 | |
| 5 | CsRg | - 35 | C unwichtig |
| 6 | CgRs | - 5 | |
| 7 | DsRg | - 20 | Wechselwirkung |
| 8 | DgRs | - 5 | |
| 9 | EsRg | - 40 | E unwichtig |
| 10 | EgRs | 0 | |
| 11 | FsRg | - 40 | F unwichtig |
| 12 | FgRs | - 5 | |
| 13 | GsRg | - 20 | Wechselwirkung |
| 14 | GgRs | - 5 | |
| 15 | HsRg | - 35 | H unwichtig |
| 16 | HgRs | 0 | |
| Test | DsGsRg DgGgRs | 0 - 40 | völlige Umkehr |

TABELLE 5.4
Ergebnisse des Komponententauschs am Beispiel Zeitzähler

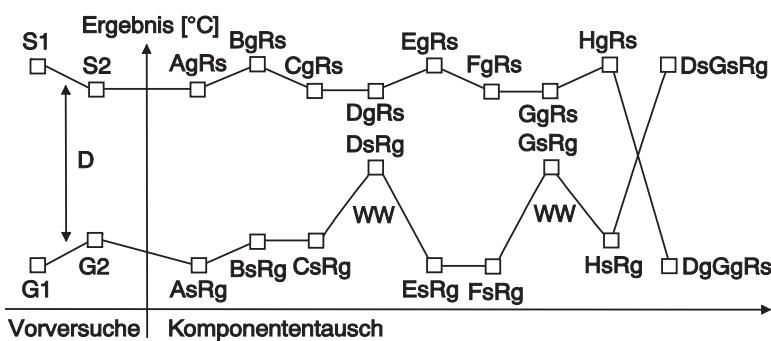


BILD 5-3 Grafische Darstellung der Versuchsergebnisse von Tabelle 5.4

D und G verschlechtern zwar das „Gut“-Ergebnis, verbessern aber nicht das „Schlecht“-Ergebnis, d. h. sie wirken nicht allein, sondern nur zusammen (Wechselwirkung). Dies zeigt sich auch im Testlauf, bei dem beide gleichzeitig vertauscht wurden.

Eine anschließende, genauere Analyse des Problems ergab, dass manche Gehäuse (D) einen Schwund aufweisen und manche Ziffernscheiben (G) nicht zentrisch sind. Wenn beide Abweichungen zusammentreffen (Wechselwirkung), klemmt die Mechanik bei tiefen Temperaturen. Nach einer Änderung der Ziffernscheibe war das Problem völlig beseitigt.

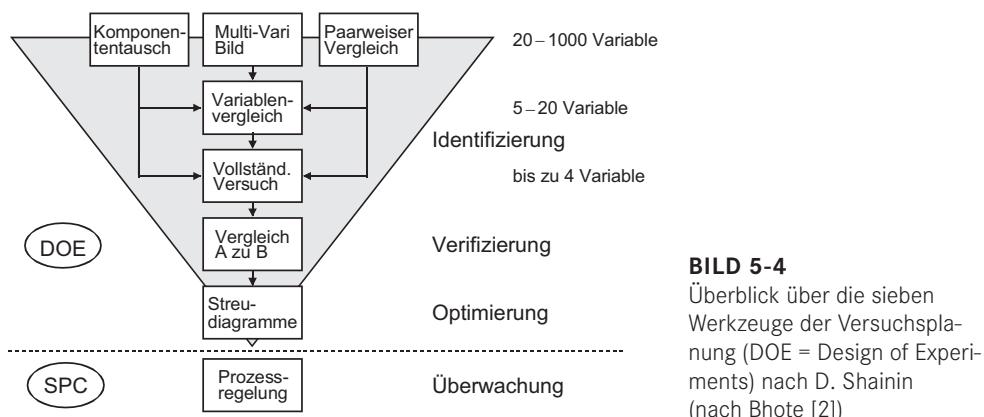
■ 5.3 Überblick über die Methoden nach D. Shainin

In den Kapiteln 4 und 5 wurden im Wesentlichen die Methoden nach D. Shainin behandelt [1, 2]. Shainins Strategie ist:

- mit Beobachtung und einfachen Versuchen die wichtigsten Faktoren erkennen,
- deren Toleranz einengen (das verursacht zwar Kosten, bringt aber große Verbesserung) und
- die Toleranz für unwichtige Faktoren aufweiten (das bringt eine Einsparung).

Hintergrund dieser Strategie ist das **Pareto-Prinzip**: Normalerweise dominieren einige wenige Ursachen (Faktoren). Den wichtigsten Faktor nennt D. Shainin „**Rotes X**“ (red X), den nächstwichtigsten „**Rosa X**“ (pink X).

Diese dominanten Faktoren gilt es zu finden und zu beherrschen. Dazu bietet D. Shainin sieben einfache Werkzeuge an. Bild 5-4 zeigt diese Werkzeuge im Überblick [2]. In der ersten Ebene (Multi-Vari-Bild, Paarweiser Vergleich, Komponententausch) sind keine Vorkenntnisse erforderlich. In den unteren Ebenen wird zunehmend Vorwissen erforderlich.



Der vollständige faktorielle Versuch wird in Kapitel 7 behandelt. Der „Vergleich A zu B“ ist im Wesentlichen der in Kapitel 6 behandelte Vergleich von zwei Mittelwerten. Ein Streudiagramm ist die Darstellung von zwei Messgrößen gegeneinander, wie sie z. B. in Kapitel 10 behandelt wird. Auf diese Werkzeuge wird daher hier nicht weiter eingegangen.

Literatur

- [1] Shainin, D./Shainin, P.: „Better than Taguchi Orthogonal Tables“, in: Quality and Reliability Engineering International 4 (1988), 143 – 149
- [2] Bhote, K. R.: Qualität – Der Weg zur Weltspitze. IQM, Großbottwar 1990
- [3] Klein, B.: Versuchsplanung – DoE. Oldenbourg Verlag, München 4. Auflage 2014

6

Statistische Grundlagen

Bisher wurde auf die statistische Auswertung von Versuchsergebnissen verzichtet. Dies ist möglich, wenn die Zufallsstreuung sehr viel kleiner ist als technologisch relevante Unterschiede. Leider ist diese Bedingung in vielen praktischen Anwendungen nicht erfüllt. Daher werden in diesem Kapitel ausgewählte statistische Grundlagen vermittelt. Im Mittelpunkt steht dabei die Darstellung der Ideen, für Einzelheiten wird auf die Standardliteratur verwiesen, z. B. die Einführungen [1 – 4], das Nachschlagewerk [5] oder das Engineering Statistics Handbook [6]. Die beiliegende DVD enthält Visualisierungen und Simulationen, die beim Verständnis der Ideen helfen sollen – Einzelheiten beschreibt Anhang G. Zunächst werden ausschließlich zufällige Unterschiede behandelt. Werden z. B. mehrere Teile unter nominell gleichen Bedingungen hergestellt, erhält man trotz aller Anstrengungen nicht genau dieselbe Dicke, Masse, Spannung, usw. In Abschnitt 6.1 wird gezeigt, wie man die Verteilung der Werte beschreiben kann, und in Abschnitt 6.2 wird gezeigt, was man aus Messwerten (sie werden Stichprobe genannt) über die zugrunde liegende „Wahrheit“ (sie wird Grundgesamtheit genannt) lernen kann.

Anschließend wird in Abschnitt 6.3 gezeigt, wie man die Ergebnisse bei zwei verschiedenen Versuchsbedingungen miteinander vergleichen kann. Es geht z. B. um die Frage, ob die Dicke von auf einer Anlage hergestellten Teilen wirklich größer ist als die Dicke von auf einer anderen Anlage herstellten. Und es wird gezeigt, wie man schon vor der Durchführung eines solchen Vergleichs einen sinnvollen Versuchsumfang festlegen kann. Genügt es, je drei Teile miteinander zu vergleichen, oder sollte man besser je 300 Teile vergleichen?

■ 6.1 Verteilung

Der Zusammenhang zwischen der Häufigkeitsverteilung von Versuchsergebnissen und der Verteilungsfunktion der Grundgesamtheit wird erläutert. Häufig ist die Normalverteilung eine gute Näherung für die Verteilung von Messwerten.

6.1.1 Häufigkeitsverteilung von Versuchsergebnissen

Trotz aller Sorgfalt erhält man bei der Wiederholung eines Einzelversuchs unter nominell gleichen Versuchsbedingungen nicht genau den gleichen Zahlenwert für das Ergebnis.

Die Versuchsergebnisse streuen aufgrund von zufälligen Unterschieden z.B. beim Ausgangsmaterial, bei den Umgebungsbedingungen oder bei der Messung.

Die Breite des Bereichs, in dem die Werte streuen, hängt z.B. von der Komplexität des Einzelversuchs, von der Breite der Streuung des Ausgangsmaterials und von der Genauigkeit der Messung ab. Wenn man das Ergebnis auf genügend viele Nachkommastellen angibt, wird man immer Unterschiede finden. Die Frage, ob diese Unterschiede technologisch relevant sind, wird erst später betrachtet. Momentan geht es nur um die phänomenologische Beschreibung der Streuung.



Beispiel aus der Galvanik (Erweiterung des Beispiels aus Abschnitt 1.2)

Auf 50 Teilen wurde unter nominell gleichen Bedingungen galvanisch eine Schicht abgeschieden. Tabelle 6.1 zeigt die Messwerte für die Schichtdicke in der Reihenfolge der Abscheidung und Messung, Bild 6-1 zeigt diese Messwerte grafisch.

TABELLE 6.1 Gemessene Schichtdicke von 50 unter nominell gleichen Bedingungen abgeschiedenen Schichten (alle Werte in µm)

| Nr. | Dicke |
|-----|-------|-----|-------|-----|-------|-----|-------|-----|-------|
| 1 | 27,7 | 9 | 30,3 | 16 | 31,0 | 23 | 31,2 | 30 | 27,6 |
| 2 | 31,3 | 10 | 27,7 | 17 | 31,8 | 24 | 25,9 | 31 | 29,4 |
| 3 | 30,1 | 11 | 28,7 | 18 | 27,4 | 25 | 30,8 | 32 | 27,4 |
| 4 | 32,9 | 12 | 26,3 | 19 | 30,2 | 26 | 34,4 | 33 | 30,9 |
| 5 | 28,2 | 13 | 31,3 | 20 | 32,4 | 27 | 28,8 | 34 | 26,8 |
| 6 | 31,0 | 14 | 30,5 | 21 | 33,0 | 28 | 30,3 | 35 | 29,3 |
| 7 | 30,4 | 15 | 30,5 | 22 | 30,5 | 29 | 28,6 | 36 | 32,8 |
| 8 | 29,8 | | | | | | | 37 | 28,8 |
| | | | | | | | | 38 | 28,8 |
| | | | | | | | | 39 | 30,4 |
| | | | | | | | | 40 | 31,0 |
| | | | | | | | | 41 | 32,0 |
| | | | | | | | | 42 | 29,1 |
| | | | | | | | | 43 | 29,8 |
| | | | | | | | | 44 | 28,4 |
| | | | | | | | | 45 | 32,3 |
| | | | | | | | | 46 | 29,8 |
| | | | | | | | | 47 | 29,8 |
| | | | | | | | | 48 | 33,3 |
| | | | | | | | | 49 | 29,3 |
| | | | | | | | | 50 | 27,5 |

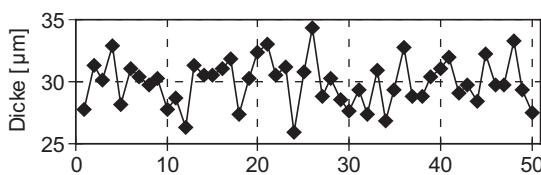


BILD 6-1

Gemessene Schichtdicke von 50 unter nominell gleichen Bedingungen abgeschiedenen Schichten

Die Versuchsergebnisse liegen in einem bestimmten Wertebereich. Dabei sind Werte in der Mitte des Bereichs (normalerweise) häufiger als Werte am Rande des Bereichs. Um eine Vorstellung von der Verteilung der Werte zu erhalten, ist es sinnvoll, den gesamten Bereich in Teilbereiche zu unterteilen und abzuzählen, wie viele Werte in jedem Teilbereich liegen. Diese Teilbereiche heißen Klassen.

Die Anzahl der Werte in jeder Klasse (in jedem Teilbereich) heißt (absolute) Häufigkeit.

Für den Vergleich von Messreihen mit unterschiedlich vielen Werten ist es nützlich, die Anzahl der Werte in jeder Klasse durch die Gesamtzahl der Werte zu teilen. So erhält man die relative Häufigkeit.

Die grafische Darstellung der (absoluten oder relativen) Häufigkeit gegen die Klassenmitten heißt Histogramm. Auch ein Histogramm unterliegt dem Zufall in dem Sinne, dass man bei einer Wiederholung der gesamten Versuchsreihe andere Einzelwerte und damit auch eine andere Häufigkeitsverteilung erhält.

Ist man an der Anzahl von Werten interessiert, die unterhalb einer bestimmten Grenze liegen, so kann man die Häufigkeit in den Klassen unterhalb dieser Grenze aufaddieren. Man erhält die kumulierte Häufigkeit.



Beispiel aus der Galvanik (Fortsetzung)

Teilt man die Werte aus Tabelle 6.1 in Klassen ein, z. B.

$$25,5 < \text{Dicke} \leq 26,5 \text{ (alle Werte in } \mu\text{m)}$$

$$26,5 < \text{Dicke} \leq 27,5 \text{ usw.}$$

und zählt dann ab, wie viele Werte in jeder dieser Klassen liegen, so erhält man die Spalte „absolute Häufigkeit“ in Tabelle 6.2. Die „2“ in der ersten Zeile bedeutet z.B., dass genau 2 der Werte in Tabelle 6.1 zwischen den Klassengrenzen 25,5 und 26,5 liegen (einschließlich 26,5), nämlich die Werte Nr. 12 (26,3) und Nr. 24 (25,9).

Die Spalte „kumulierte Häufigkeit“ in Tabelle 6.2 enthält die Anzahl der Werte kleiner oder gleich der jeweiligen Klassenobergrenze. Die „6“ in der zweiten Zeile bedeutet z. B., dass genau 6 der Werte in $25,5 < \text{Dicke} \leq 27,5$ sind, nämlich 2 Werte mit $25,5 < \text{Dicke} \leq 26,5$ und 4 Werte mit $26,5 < \text{Dicke} \leq 27,5$ usw. Die „kumulierte Häufigkeit“ ist die Summe der „absoluten Häufigkeiten“ bis zur jeweiligen Zeile. Sie wird auch Summenhäufigkeit genannt.

Die „relative Häufigkeit“ erhält man, indem man die „absolute Häufigkeit“ durch die Gesamtzahl der Werte (hier 50) teilt, die „relative kumulierte Häufigkeit“ erhält man entsprechend aus der „kumulierten Häufigkeit“.

Bild 6-2 zeigt die grafische Darstellung der absoluten Häufigkeit als Histogramm. Das Histogramm gibt eine unmittelbare Vorstellung von der Verteilung der Werte.

TABELLE 6.2 Klassierung der Werte aus Tabelle 6.1

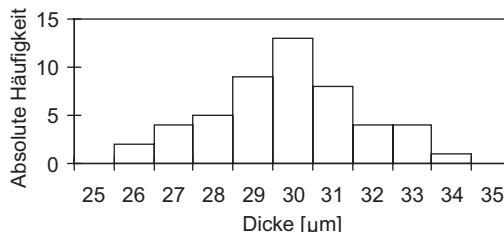
absolute Häufigkeit = Anzahl der Werte in der Klasse

kumulierte Häufigkeit = summierte Anzahl = Anzahl bis zur Klassenobergrenze

relative Häufigkeit = absolute Häufigkeit/Gesamtzahl der Werte

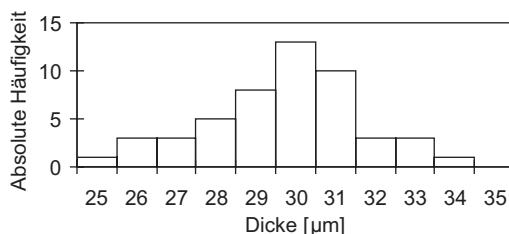
rel. kum. Häufigkeit = kumulierte Häufigkeit/Gesamtzahl der Werte

| Klasse | absolute Häufigkeit | kumulierte Häufigkeit | relative Häufigkeit | rel. kum. Häufigkeit |
|---------------------------------|---------------------|-----------------------|---------------------|----------------------|
| $25,5 < \text{Dicke} \leq 26,5$ | 2 | 2 | 0,04 | 0,04 |
| $26,5 < \text{Dicke} \leq 27,5$ | 4 | 6 | 0,08 | 0,12 |
| $27,5 < \text{Dicke} \leq 28,5$ | 5 | 11 | 0,10 | 0,22 |
| $28,5 < \text{Dicke} \leq 29,5$ | 9 | 20 | 0,18 | 0,40 |
| $29,5 < \text{Dicke} \leq 30,5$ | 13 | 33 | 0,26 | 0,66 |
| $30,5 < \text{Dicke} \leq 31,5$ | 8 | 41 | 0,16 | 0,82 |
| $31,5 < \text{Dicke} \leq 32,5$ | 4 | 45 | 0,08 | 0,90 |
| $32,5 < \text{Dicke} \leq 33,5$ | 4 | 49 | 0,08 | 0,98 |
| $33,5 < \text{Dicke} \leq 34,5$ | 1 | 50 | 0,02 | 1,00 |

**BILD 6-2**

Grafische Darstellung der absoluten Häufigkeit als Histogramm

Führt man die 50 Einzelversuche unter nominell gleichen Bedingungen ein weiteres Mal durch, so erhält man andere Versuchsergebnisse und damit ein anderes Histogramm. Bild 6-3 zeigt das Ergebnis einer Wiederholung der Versuchsreihe.¹

**BILD 6-3**

Grafische Darstellung einer Wiederholung der Versuchsreihe

6.1.2 Verteilungsdichte und Verteilungsfunktion

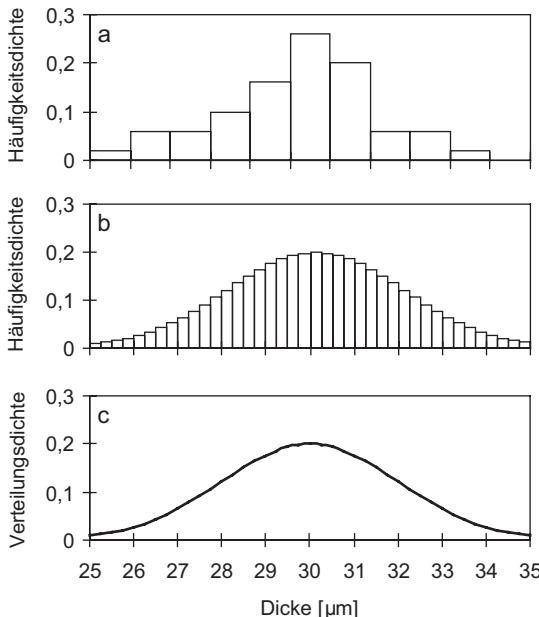
Je größer die Anzahl der Messwerte wird, desto weniger unterscheiden sich Histogramme der relativen Häufigkeit voneinander.

Unterteilt man nun die Klassen, so werden die Stufen kleiner. Allerdings verkleinern sich auch alle relativen Häufigkeiten. Dies erschwert den Vergleich von Histogrammen mit unterschiedlichen Klassenbreiten. Teilt man die relative Häufigkeit aber durch die Klassenbreite, so sind die Werte (abgesehen von der Stufung) unabhängig von der Klassenbreite – man erhält die sogenannte Häufigkeitsdichte (Bild 6-4 b).

Stellt man sich nun vor, die Anzahl der Messwerte wird beliebig groß und die Klassen werden immer feiner unterteilt, so geht die Häufigkeitsdichte in eine stetige Funktion über, die die tatsächliche Verteilung der Versuchsergebnisse für die gewählten Versuchsbedingungen beschreibt (Bild 6-4 c). Sie heißt Verteilungsdichte der Grundgesamtheit. Die relative kumulierte Häufigkeit geht bei immer größerer Anzahl von Werten und immer feinerer Unterteilung in die Verteilungsfunktion der Grundgesamtheit über (Bild 6-5).

Mit dem Begriff Grundgesamtheit bezeichnet man in der Statistik die „Wahrheit“, über die Aussagen gemacht werden sollen, z.B. die Durchmesser aller Teile einer Liefercharge, oder sämtliche Versuchsergebnisse, die man erhielt, wenn man einen bestimmten Einzelversuch unter genau definierten Bedingungen beliebig oft durchführen könnte. Sie ist für eine bestimmte Situation fest, aber nicht bekannt.

¹ Die Visualisierung „Normalverteilte Zufallszahlen“ erzeugt Zufallszahlen und stellt sie ähnlich zu Bild 6-1 und 6-2 dar. Zusätzlich werden Mittelwerte, Varianz und Standardabweichung dargestellt – vgl. Anhang G.

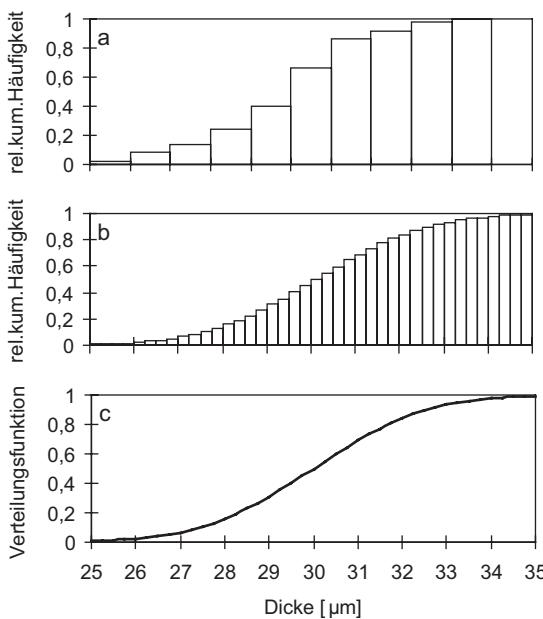
**BILD 6-4**

Mit zunehmender Anzahl der Messwerte n und immer feinerer Unterteilung in Klassen geht die Häufigkeitsdichte über in die Verteilungsdichte

a) $n = 50$, Breite $1\mu\text{m}$ (= Bild 6-3)

b) $n = 10^5$, Breite $0,25\mu\text{m}$

c) $n \rightarrow \infty$, Breite $\rightarrow 0$

**BILD 6-5**

Mit zunehmender Anzahl der Messwerte n und immer feinerer Unterteilung in Klassen geht die relative kumulierte Häufigkeit über in die Verteilungsfunktion

In praktischen Anwendungen hat man nur einige (meist wenige) Versuchsergebnisse vorliegen. Diese Versuchsergebnisse sind eine Stichprobe aus der Grundgesamtheit. Stichprobenergebnisse unterliegen dem Zufall. Bei zwei Stichproben aus derselben Grundgesamtheit erhält man nicht dasselbe Ergebnis.

6.1.3 Normalverteilung

Wenn die Zufallsstreuung die Summe vieler Einflüsse (wie Umgebungsbedingungen, Materialeigenschaften usw.) ist, werden die Messwerte häufig in guter Näherung durch die sogenannte Normalverteilung beschrieben.

Die Dichte der Normalverteilung ist gegeben durch

$$g(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6.1)$$

wobei

μ = Mittelwert und

σ = Standardabweichung (bzw. σ^2 = Varianz) der Normalverteilung

zwei Parameter sind, die Lage bzw. Breite der Verteilung beschreiben.

Bild 6-6 zeigt die Dichte der Normalverteilung für verschiedene Mittelwerte μ , Bild 6-7 für verschiedene Standardabweichungen σ .

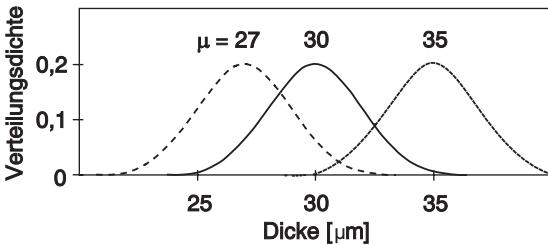


BILD 6-6

Dichte der Normalverteilungen mit Mittelwerten $\mu = 27, 30$ und $35 \mu\text{m}$, Standardabweichung $\sigma = 2 \mu\text{m}$

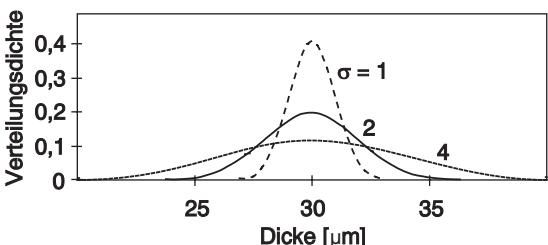


BILD 6-7

Dichte der Normalverteilungen mit Mittelwert $\mu = 30 \mu\text{m}$ und Standardabweichungen $\sigma = 1, 2$ und $4 \mu\text{m}$

- Die Dichte der Normalverteilung ist symmetrisch zum Mittelwert μ .
- An den Stellen $\mu \pm \sigma$ ist die Dichte am steilsten.
- 68,3 % der Gesamtfläche liegt im Bereich $\mu - \sigma \leq x \leq \mu + \sigma$.
- 95,5 % der Gesamtfläche liegt im Bereich $\mu - 2\sigma \leq x \leq \mu + 2\sigma$.
- 99,73 % der Gesamtfläche liegt im Bereich $\mu - 3\sigma \leq x \leq \mu + 3\sigma$.

Durch die lineare Transformation

$$u = \frac{x - \mu}{\sigma} \quad (6.2)$$

geht die Normalverteilung in Gleichung (6.1) in die sogenannte standardisierte Normalverteilung

$$g(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}} \quad (6.3)$$

mit Mittelwert 0 und Standardabweichung 1 über. Bild 6-8 zeigt diesen Übergang grafisch. Vorteil der standardisierten Normalverteilung ist, dass sie nur von der Größe u abhängt und daher leicht berechnet und tabelliert werden kann.

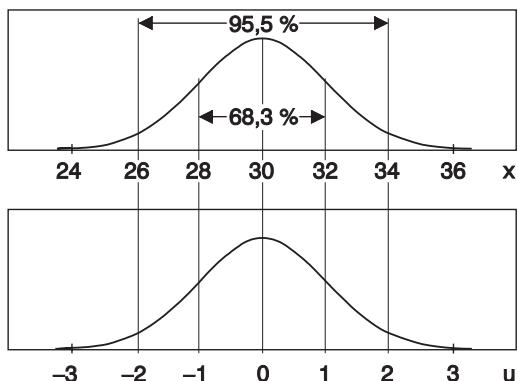


BILD 6-8

Verteilungsdichte der Normalverteilung mit Mittelwert $\mu = 30$ und Standardabweichung $\sigma = 2$, verglichen mit der standardisierten Normalverteilung.

■ 6.2 Auswertung einer Stichprobe

In praktischen Anwendungen ist die Verteilung der Grundgesamtheit (die Wahrheit) nicht bekannt. Man hat nur (mehr oder weniger viele) Messwerte vorliegen. Diese Messwerte stellen eine Stichprobe aus der Grundgesamtheit dar. Die Anzahl der Messwerte wird als Stichprobenumfang n bezeichnet.

6.2.1 Repräsentative Stichprobe

Aus der Stichprobe möchte man etwas über die Grundgesamtheit lernen. Dazu muss die Stichprobe repräsentativ für die Grundgesamtheit sein. Dies erreicht man z.B. durch zufällige Entnahme. Jedes Element der Grundgesamtheit hat dann die gleiche Wahrscheinlichkeit, in die Stichprobe zu kommen (vgl. Randomisierung und Blockbildung in den Absätzen 3.4.3 und 3.4.4).

In praktischen Anwendungen wird diese Randbedingung leider häufig missachtet. Dann sind Enttäuschungen vorprogrammiert, etwa folgender Art: „Jetzt habe ich statistische Methoden angewendet, und trotzdem bin ich zum falschen Ergebnis gekommen.“



Beispiel

Eine Schicht wird galvanisch abgeschieden. Die Schichtdicke an einer definierten Stelle des Teils ist das gesuchte Versuchsergebnis. Zehn Teile befinden sich gleichzeitig in der Anlage. Von diesen zehn Teilen einer Anlagenfüllung werden fünf zufällig ausgewählt und gemessen.

Die statistische Auswertung dieser fünf Messwerte macht nur eine Aussage über die Verteilung der Schichtdicken bei dieser einen Füllung. Das Ergebnis ist auf andere Füllungen nur dann übertragbar, wenn es keinerlei Unterschiede zwischen den Füllungen gibt.

Da dies nicht gewährleistet werden kann, ist es besser, Teile aus verschiedenen Füllungen zu messen. Zufällige Unterschiede zwischen den Füllungen sind dann in der Stichprobe enthalten.

Aber auch bei einer Stichprobe aus mehreren Füllungen sind noch falsche Schlussfolgerungen möglich, z.B. weil Langzeitänderungen auftreten, die von der Stichprobe nicht erfasst wurden.²



Aufgabe

Die Zufallsstreuung eines bestimmten Längenmessgeräts soll ermittelt werden. Dazu nimmt ein Bediener ein repräsentatives Teil, legt es ein und liest den Messwert zehnmal hintereinander ab. Was ist von dieser Vorgehensweise zu halten? Was könnte er besser machen? (Bitte denken Sie über den Ablauf einer Messung und die beschriebene Vorgehensweise nach, bevor Sie die Lösung lesen. Ziel der Aufgabe ist, dass Sie sich das Problem bewusst machen; es wird nicht erwartet, dass Sie die vollständige Lösung selbst finden.)

Lösung

Mit der beschriebenen Vorgehensweise wird der Bediener eine sehr kleine Streuung erhalten, evtl. erhält er sogar zehnmal den gleichen Messwert. Die so ermittelte Streuung sagt jedoch nichts über die Zufallsstreuung des Messgeräts aus. Die zehn Messwerte sind keine repräsentative Stichprobe.

Folgende Verbesserungen sind möglich (je nach Fragestellung):

- Die Mindestmaßnahme ist, dass das Teil nach jeder Messung entnommen und wieder neu eingelegt wird. So wird die Zufallsstreuung durch unterschiedliches Einlegen mit erfasst.
- Besteht der Verdacht, dass die Zufallsstreuung vom zu messenden Teil abhängt, so müssen Messungen an mehreren zufällig entnommenen Teilen durchgeführt werden (die wahren Unterschiede zwischen den Teilen müssen dann allerdings herausgerechnet werden).
- Besteht der Verdacht, dass das Messgerät driftet oder dass der Messwert von den Umgebungsbedingungen abhängt, so müssen die Messungen über einen längeren Zeitraum verteilt werden, wenn diese Streuursache mit erfasst werden soll.
- Wird die Messung in der Anwendung von verschiedenen Bedienern durchgeführt, so sind auch Unterschiede zwischen den Bedienern mit zu erfassen.
- Werden in der Anwendung mehrere baugleiche Messgeräte wahlweise eingesetzt, so sind auch Unterschiede zwischen den Messgeräten zu erfassen.

Im Einzelfall können weitere Streuursachen von Bedeutung sein. Sie müssen dann auch bei der Untersuchung berücksichtigt werden.

Die Erfassung der verschiedenen Streuursachen kann dadurch erfolgen, dass in der Stichprobe zufällig verschiedene Teile, zu verschiedenen Zeiten, von verschiedenen Bedienern, auf verschiedenen Messgeräten usw. enthalten sind. Dann wird die Gesamtstreuung durch all diese Ursachen erfasst. Eine Zuordnung von Einzelbeiträgen zu den Ursachen ist jedoch nicht möglich.

Bei Messgerätefähigkeitsuntersuchungen behandelt man Messgerät, Bediener usw. als Faktoren in einem geplanten Versuch und verändert sie systematisch. Dann kann man zwischen den einzelnen Streubeiträgen unterscheiden.

² Es gibt statistische Methoden, mit denen man solche Langzeitänderungen erkennen kann (z.B. Regelkarten).

6.2.2 Eintragung ins Wahrscheinlichkeitsnetz

Die Normalverteilung beschreibt Messwerte häufig in guter Näherung. Es können jedoch auch Abweichungen auftreten. Da die folgenden Auswertungsverfahren die Normalverteilung voraussetzen, sollte zunächst überprüft werden, ob die Messwerte damit konsistent sind.

Die Eintragung der Messwerte ins Wahrscheinlichkeitsnetz (genau: Wahrscheinlichkeitsnetz für normalverteilte Merkmalswerte) ist ein einfaches grafisches Verfahren dafür. Das Wahrscheinlichkeitsnetz beruht auf folgender Idee:

Aufgrund der S-Form der Verteilungsfunktion (Bild 6-5 bzw. 6-9) ist es schwer, Messwerte direkt mit der Normalverteilung zu vergleichen. Durch eine geeignete Umskalierung der %-Achse kann man jedoch erreichen, dass die Verteilungsfunktion eine Gerade darstellt. Messwerte aus einer normalverteilten Grundgesamtheit liegen dann (abgesehen von zufälligen Abweichungen) auf einer Geraden.

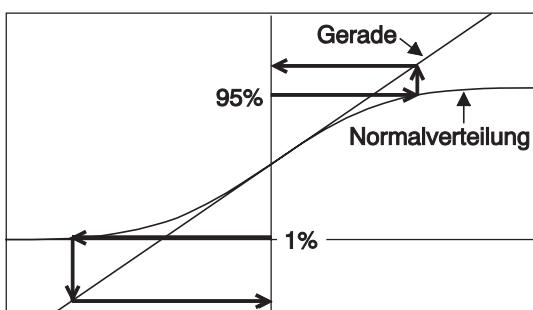


BILD 6-9

Entstehung des Wahrscheinlichkeitsnetzes:
Durch eine geeignete Skalierung der y bzw.
%-Achse (%-Werte nahe 0 und 100% wer-
den entsprechend den Pfeilen nach außen
verschoben) wird die S-förmige Vertei-
lungsfunktion auf eine Gerade abgebildet.

Das Wahrscheinlichkeitsnetz kann man nutzen, um auf einfache grafische Art zu überprüfen, ob n Messwerte aus einer Normalverteilung stammen können.

Dazu ordnet man zunächst die n Werte nach ihrer Größe und nummeriert sie in dieser Reihenfolge von 1 bis n . Dem Wert Nr. i ordnet man dann einen Prozentwert zu, der

$$\text{ungefähr } \frac{i - 0,5}{n} \cdot 100\% \quad (6.4)$$

beträgt. Dies ist die relative kumulierte Häufigkeit (abgesehen von einer kleinen Korrektur, die bewirkt, dass der kleinste Wert ($i = 1$) den gleichen Abstand von 0 % wie der größte Wert ($i = n$) von 100 % hat).³ Schließlich trägt man den Messwert auf der x-Achse und den Prozentwert auf der y-Achse im Wahrscheinlichkeitsnetz auf.

Stammen die Werte aus einer normalverteilten Grundgesamtheit, so liegen sie näherungsweise auf einer Geraden, wie in Bild 6-10 ($n = 8$ Werte) bzw. 6-11 ($n = 50$ Werte).

Liegt der niedrigste oder höchste Wert (evtl. auch mehr als ein Wert) wie in Bild 6-12 weitab von der Geraden durch die restlichen Werte, so handelt es sich vermutlich um

³ In der Literatur finden sich auch andere, ähnliche Näherungsformeln. Hier geht es nur darum, das Prinzip der Eintragung zu erläutern. Software verwendet meist die exakten Werte.

einen Ausreißer, er passt nicht zu den anderen Werten. Man sollte überprüfen, ob bei diesem Wert evtl. ein Mess- oder Übertragungsfehler vorliegt.

Liegen die Werte insgesamt auf einer deutlich gekrümmten Linie wie in Bild 6-13, so stammen die Werte vermutlich aus einer nicht normalverteilten Grundgesamtheit. Man sollte nach einer geeigneten Transformation der Werte suchen und z.B. mit dem Logarithmus der Messwerte rechnen, statt mit den Werten selbst. Einzelheiten werden in Abschnitt 6.4 behandelt.



Beispiel aus der Galvanik

Die Dicke von $n = 8$ unter nominell gleichen Bedingungen abgeschiedenen Schichten beträgt (die ersten Werte aus Tabelle 6.1):

27,7 31,3 30,1 32,9 28,2 31,0 30,4 29,8

Ordnet man diese Messwerte der Größe nach, so erhält man Tabelle 6.3.

| i | Messwert Dicke [μm] | $\frac{i - 0,5}{n} \cdot 100\%$ |
|---|-------------------------------------|---------------------------------|
| 1 | 27,7 | 6,3 |
| 2 | 28,2 | 18,8 |
| 3 | 29,8 | 31,3 |
| 4 | 30,1 | 43,8 |
| 5 | 30,4 | 56,2 |
| 6 | 31,0 | 68,7 |
| 7 | 31,3 | 81,2 |
| 8 | 32,9 | 93,7 |

TABELLE 6.3

Nach der Größe geordnete Messwerte und aus der Ordnungsnummer i berechnete %-Werte für die Eintragung ins Wahrscheinlichkeitsnetz

Die Eintragung dieser Werte in ein Wahrscheinlichkeitsnetz ergibt Bild 6-10. Trägt man alle 50 Werte aus Tabelle 6.1 in ein Wahrscheinlichkeitsnetz ein, so erhält man das Bild 6-11.

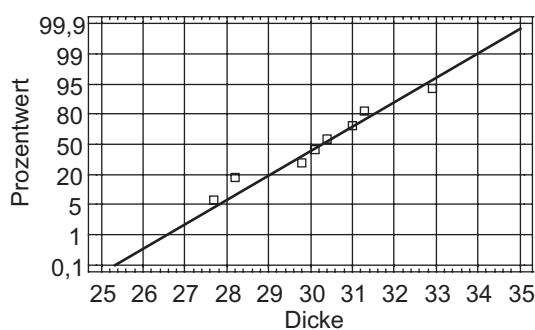
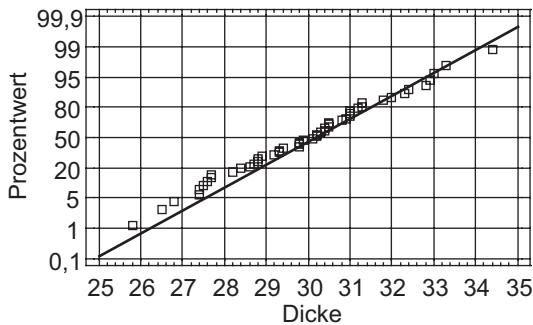
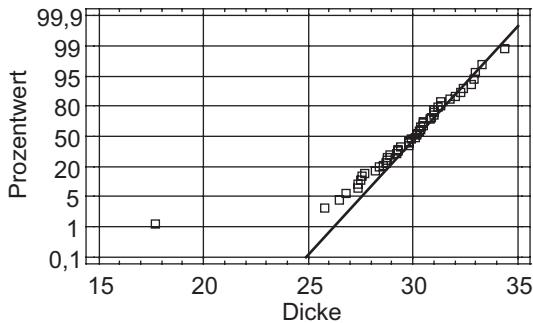


BILD 6-10

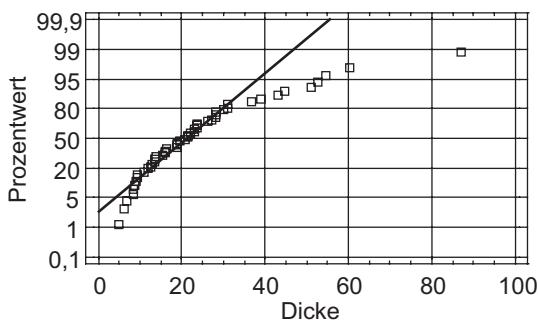
Eintragung der acht Werte aus Tabelle 6.3 ins Wahrscheinlichkeitsnetz

**BILD 6-11**

Eintragung der 50 Werte aus Tabelle 6.1 ins Wahrscheinlichkeitsnetz

**BILD 6-12**

Eintragung von 50 Werten mit einem Ausreißer (der einzelne Wert bei ca. 18 µm) ins Wahrscheinlichkeitsnetz

**BILD 6-13**

Eintragung von 50 nicht normalverteilten Werten ins Wahrscheinlichkeitsnetz – eine deutliche Krümmung ist erkennbar

Jede gute Statistiksoftware enthält die Darstellung von Daten im Wahrscheinlichkeitsnetz (meist englisch „probability plot“). Während die Eintragung der Werte von Hand mühsam ist, ist sie mit Software auch für eine große Anzahl von Werten sehr einfach.⁴

Als Entscheidungshilfen beim Erkennen von Abweichungen von der Normalverteilung werden in Softwarepaketen zusätzlich auch rechnerische Tests, wie z. B. der Kolmogoroff-Smirnoff- oder der χ^2 -Test (für großes n) und der Shapiro-Wilk-Test (auch für kleines n [1 – 6] angeboten.

⁴ Q-Q-Plot und P-P-Plot haben dasselbe Ziel. Bei diesen Plots wird in Bild 6-9 die x-Achse (statt der y-Achse) umskaliert, um eine Gerade zu erhalten. Dadurch erscheinen die Punkte im Plot gleichmäßiger verteilt.

6.2.3 Schätzwerte für Mittelwert μ und Varianz σ^2

Messwerte werden hier und im Folgenden grundsätzlich mit y bezeichnet, da in der Versuchsplanung meist die Zielgröße gemessen wird, in Abhängigkeit von Faktorstufen x .

Der Mittelwert \bar{y} von n Messwerten y_i

$$\text{Mittelwert } \bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i \quad (6.5)$$

ist ein Schätzwert für den Mittelwert μ der Verteilung⁵ und die

$$\text{Varianz } s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (y_i - \bar{y})^2 \quad (6.6)$$

ist ein Schätzwert für die Varianz σ^2 der Verteilung.

Da die Einzelwerte y_i dem Zufall unterliegen, unterliegen auch die daraus berechneten Schätzwerte dem Zufall. Aus einer anderen Stichprobe erhält man andere Schätzwerte, ohne dass deswegen der eine richtig und der andere falsch ist. Sie unterscheiden sich zufällig.

Die so berechneten Mittelwerte und Varianzen haben zwei angenehme Eigenschaften:

1. Je größer der Stichprobenumfang n wird, desto weniger streuen die Schätzwerte. Sie nähern sich immer mehr den wahren Werten, weil sich zufällige Unterschiede zwischen den Einzelwerten immer besser kompensieren.
2. Der Mittelwert vieler Schätzwerte nähert sich ebenfalls immer mehr den wahren Werten (man nennt diese Eigenschaft „erwartungstreu“).

$s = \sqrt{s^2}$ ist ein Schätzwert für die wahre Standardabweichung σ . Dieser Schätzwert ist aber nicht erwartungstreu. Daher werden im Folgenden bei der Zusammenfassung mehrerer Stichproben immer Mittelwerte der Varianzen und nicht der Standardabweichungen verwendet. Die Visualisierung „Normalverteilte Zufallszahlen“ erlaubt es, diese Zusammenhänge spielerisch zu erleben (vgl. Anhang G).



Beispiel 1

Vier Teile werden galvanisch beschichtet. Als Schichtdicken erhält man (in μm):

27,7 31,3 30,1 32,9

Der Mittelwert beträgt:

$$\bar{y} = \frac{1}{4} \cdot (27,7 + 31,3 + 30,1 + 32,9) = 30,5 \mu\text{m}$$

⁵ Hier und im Folgenden bezeichnen griechische Buchstaben immer Parameter der Grundgesamtheit (d. h. die „Wahrheit“), lateinische Buchstaben bezeichnen Schätzwerte für diese Parameter, die aus einer Stichprobe berechnet wurden.

Die Varianz beträgt:

$$s^2 = \frac{1}{4-1} \cdot \left((27,7 - 30,5)^2 + (31,3 - 30,5)^2 + (30,1 - 30,5)^2 + (32,9 - 30,5)^2 \right) = 4,8$$

Die Standardabweichung beträgt:

$$s = \sqrt{4,8} = 2,19 \text{ } \mu\text{m}$$

Die meisten Taschenrechner besitzen eine Statistikfunktion, die aus den Einzelwerten direkt Mittelwert und Standardabweichung berechnet.

Beispiel 2

Vier weitere Teile werden unter identischen Bedingungen galvanisch beschichtet. Als Schichtdicken erhält man (in μm):

28,2 31,0 30,4 29,8

Der Mittelwert beträgt $\bar{y} = 29,85 \text{ } \mu\text{m}$.

Die Standardabweichung beträgt $s = \sqrt{1,45} = 1,20 \text{ } \mu\text{m}$.

Der Unterschied zwischen Beispiel 1 und Beispiel 2 ist zufallsbedingt. Man kann nicht sagen, das eine Ergebnis ist richtig, das andere falsch. Es ist auch nicht so, dass im Beispiel 2 genauer gemessen wurde, weil die Streuung kleiner ist. Auch dieser Unterschied ist zufällig.

Übrigens: Beide „Versuche“ sind Zufallszahlen aus einer Normalverteilung mit Mittelwert $\mu = 30 \text{ } \mu\text{m}$ und Standardabweichung $\sigma = 2 \text{ } \mu\text{m}$. Aber diese Information hat man bei einem echten Versuch natürlich nicht. Dann kennt man nur das Ergebnis und versucht, daraus etwas über den wahren Wert von μ und σ zu lernen.



Aufgabe

Sie messen die maximale Leistung von 5 Motoren und erhalten folgende Ergebnisse (in kW):

52,1 49,8 50,3 51,4 50,9

Berechnen Sie Mittelwert, Varianz und Standardabweichung der Stichprobe.

Lösung

Mittelwert $\bar{y} = 50,9 \text{ kW}$

Varianz $s^2 = 0,815 \text{ (kW)}^2$

Standardabweichung $s = 0,903 \text{ kW}$

6.2.4 Vertrauensbereiche

Vertrauensbereich für den Mittelwert μ

Der Mittelwert \bar{y} einer Stichprobe unterliegt dem Zufall. Bei jeder Stichprobe erhält man einen anderen Mittelwert. Alle diese Mittelwerte liegen in der Nähe des wahren, aber unbekannten Mittelwerts μ der Grundgesamtheit.

Die Streuung der Stichprobenmittelwerte ist umso kleiner,

- je kleiner die Standardabweichung ist und
- je größer der Stichprobenumfang n ist.⁶



Beispiel

Berechnet man in Tabelle 6.1 die Mittelwerte von jeweils zwei aufeinanderfolgenden Werten, so erhält man:

$$(27,7 + 31,3)/2 = 29,5$$

$$(30,1 + 32,9)/2 = 31,5$$

$$(28,2 + 31,0)/2 = 29,6 \text{ usw.}$$

Die zufälligen Abweichungen vom wahren, aber unbekannten Mittelwert μ kompensieren sich teilweise. Diese Kompensation der zufälligen Abweichungen ist umso besser, aus je mehr Werten der Stichprobenmittelwert berechnet wird, d.h. je größer n ist. Die Zufallsstreuung der Mittelwerte ist damit kleiner als die der Einzelwerte.

Ganz allgemein kann man zeigen, dass die Varianz des Mittelwerts \bar{y} von n Einzelwerten gegeben ist durch

$$\text{Varianz des Mittelwerts } \sigma_{\bar{y}}^2 = \frac{\sigma^2}{n} \quad (6.7)$$

d.h. die Varianz des Mittelwerts ist $1/n$ der Varianz der Einzelwerte (weil zufällige Abweichungen sich kompensieren).

Wenn man nun bei einer Stichprobe einen bestimmten Mittelwert \bar{y} beobachtet hat, erwartet man daher umgekehrt, dass der wahre, aber unbekannte Mittelwert μ der Grundgesamtheit in der Nähe liegt. Und zwar umso näher, je kleiner die Standardabweichung s und je größer der Umfang n der Stichprobe ist. Gemäß (6.7) ist

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} \quad (6.8)$$

ein Schätzwert für die Standardabweichung des Mittelwerts. Man kann zeigen, dass der Bereich

$$\bar{y} - t \cdot s_{\bar{y}} \leq \mu \leq \bar{y} + t \cdot s_{\bar{y}} \quad (6.9)$$

den wahren, aber unbekannten Mittelwert μ mit einer Wahrscheinlichkeit von 95 %, 99 % bzw. 99,9 % enthält, wenn man für t die Werte aus Tabelle 6.4 verwendet. Dieser Bereich heißt zweiseitiger Vertrauensbereich für den Mittelwert, die Wahrscheinlichkeit (z.B. 95 %) heißt Vertrauensniveau. Man spricht auch kurz vom 95 %-Vertrauensbereich usw.

⁶ Die Visualisierung „Vertrauensbereich für den Mittelwert“ zieht Stichproben beliebigen Umfangs aus einer Normalverteilung und lässt die Veränderung von Mittelwert und Vertrauensbereich von Stichprobe zu Stichprobe beobachten – vgl. Anhang G.

| Freiheitsgrad f | t-Werte für Vertrauensniveau | | |
|-----------------|------------------------------|-------|--------|
| | 95 % | 99 % | 99,9 % |
| 1 | 12,71 | 63,66 | 636,62 |
| 2 | 4,303 | 9,925 | 31,60 |
| 3 | 3,182 | 5,841 | 12,92 |
| 4 | 2,776 | 4,604 | 8,610 |
| 5 | 2,571 | 4,032 | 6,869 |
| 6 | 2,447 | 3,707 | 5,959 |
| 7 | 2,365 | 3,499 | 5,408 |
| 8 | 2,306 | 3,355 | 5,041 |
| 9 | 2,262 | 3,250 | 4,781 |
| 10 | 2,228 | 3,169 | 4,587 |
| 12 | 2,179 | 3,055 | 4,318 |
| 15 | 2,131 | 2,947 | 4,073 |
| 20 | 2,086 | 2,845 | 3,850 |
| 30 | 2,042 | 2,750 | 3,646 |
| 40 | 2,021 | 2,704 | 3,551 |
| 50 | 2,009 | 2,678 | 3,496 |
| 70 | 1,994 | 2,648 | 3,435 |
| 100 | 1,984 | 2,626 | 3,390 |
| 1000 | 1,962 | 2,581 | 3,300 |
| ∞ | 1,960 | 2,576 | 3,291 |

TABELLE 6.4t-Werte zur Berechnung zweiseitiger Vertrauensbereiche⁷

t hängt vom gewünschten Vertrauensniveau ab: Je sicherer man sein möchte, dass der Vertrauensbereich den wahren Wert für μ enthält, desto breiter ist der Bereich, d. h. desto größer ist t.

Der Schätzwert s in (6.8) wird aus der Stichprobe berechnet. Er unterliegt daher selbst dem Zufall. Je weniger Werte zur Berechnung von s verwendet werden, desto größer ist seine Zufallsstreuung. Dies wird mit einem größeren Wert von t berücksichtigt. Nach Abzug des Mittelwerts verbleiben von den n Werten in der Stichprobe nur noch $n - 1$ Werte zur Berechnung von s. Diese Anzahl wird in Tabelle 6.4 als Freiheitsgrad f bezeichnet. Es gilt:

$$\text{Freiheitsgrad } f = n - 1 \quad (6.10)$$

⁷ Für statistisch Interessierte:

Die Zahlenwerte in Tabelle 6.4 sind die kritischen Werte $t_{f, 1-\alpha/2}$ der t- oder Student-Verteilung (z. B. [5]). Um die Rechnung zu erleichtern, werden hier direkt die Faktoren zur Berechnung der zweiseitigen Vertrauensbereiche angegeben.

**Beispiel 1 (Fortsetzung)**

Vier Teile werden galvanisch beschichtet. Als Schichtdicken erhält man (in μm):

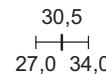
27,7 31,3 30,1 32,9

Der Mittelwert beträgt $\bar{y} = 30,5 \mu\text{m}$, die Standardabweichung $s = 2,19 \mu\text{m}$.

Für den Vertrauensbereich erhält man mit $f = 4 - 1 = 3$:

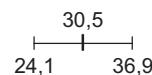
95 %-Vertrauensniveau: $t = 3,182$

$$30,5 - \frac{3,182 \cdot 2,19}{\sqrt{4}} = 27,0 \leq \mu \leq 34,0 = 30,5 + \frac{3,182 \cdot 2,19}{\sqrt{4}}$$



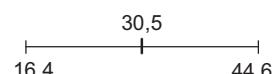
99 %-Vertrauensniveau: $t = 5,841$

$$30,5 - \frac{5,841 \cdot 2,19}{\sqrt{4}} = 24,1 \leq \mu \leq 36,9 = 30,5 + \frac{5,841 \cdot 2,19}{\sqrt{4}}$$



99,9 %-Vertrauensniveau: $t = 12,92$

$$30,5 - \frac{12,92 \cdot 2,19}{\sqrt{4}} = 16,4 \leq \mu \leq 44,6 = 30,5 + \frac{12,92 \cdot 2,19}{\sqrt{4}}$$



Je höher das Vertrauensniveau, desto breiter ist der Vertrauensbereich.

Beispiel 2 (Fortsetzung)

Vier weitere Teile werden unter identischen Bedingungen galvanisch beschichtet. Als Schichtdicken erhält man (in μm):

28,2 31,0 30,4 29,8

Der Mittelwert beträgt $\bar{y} = 29,85 \mu\text{m}$, die Standardabweichung $s = 1,20 \mu\text{m}$.

95 %-Vertrauensniveau: $29,85 - \frac{3,182 \cdot 1,20}{\sqrt{4}} = 27,9 \leq \mu \leq 31,8 = 29,85 + \frac{3,182 \cdot 1,20}{\sqrt{4}}$

99 %-Vertrauensniveau: $29,85 - \frac{5,841 \cdot 1,20}{\sqrt{4}} = 26,3 \leq \mu \leq 33,4 = 29,85 + \frac{5,841 \cdot 1,20}{\sqrt{4}}$

99,9 %-Vertrauensniveau: $29,85 - \frac{12,92 \cdot 1,20}{\sqrt{4}} = 22,1 \leq \mu \leq 37,6 = 29,85 + \frac{12,92 \cdot 1,20}{\sqrt{4}}$

Der Unterschied zwischen Beispiel 1 und Beispiel 2 ist zufallsbedingt. Man kann nicht sagen, dass eine Ergebnis ist richtig, das andere falsch. Beide Vertrauensbereiche sind „richtig“ in dem Sinn, dass sie den in diesem simulierten Beispiel bekannten Mittelwert $\mu = 30 \mu\text{m}$ enthalten.

Bei jeder Stichprobe erhält man andere Einzelwerte und damit auch andere Vertrauensbereiche. Im Mittel über viele Stichproben enthalten 95 % (bzw. 99 % bzw. 99,9 %) der mit der Berechnungsformel (6.9) berechneten Vertrauensbereiche den normalerweise unbekannten wahren Wert. Bei einem einzelnen Vertrauensbereich kann man ohne Zusatzinformation nicht feststellen, ob er den wahren Wert enthält.



Aufgabe (Fortsetzung)

Sie messen die maximale Leistung von 5 Motoren und erhalten folgende Ergebnisse (in kW):

52,1 49,8 50,3 51,4 50,9

Berechnen Sie den Vertrauensbereich für den Mittelwert zum Vertrauensniveau 95 %, 99 % und 99,9 %.

Lösung

$n = 5; f = 4$

95 %-Vertrauensbereich: $49,8 \text{ kW} \leq \mu \leq 52,0 \text{ kW}$

99 %-Vertrauensbereich: $49,0 \text{ kW} \leq \mu \leq 52,8 \text{ kW}$

99,9 %-Vertrauensbereich: $47,4 \text{ kW} \leq \mu \leq 54,4 \text{ kW}$.

Hinweise zum Vertrauensbereich für den Mittelwert

- Wiederholt man einen Versuch unter identischen Bedingungen, so erhält man aufgrund der Zufallsstreuung unterschiedliche Versuchsergebnisse. Aus jeder Stichprobe vom Umfang n berechnet man einen anderen Mittelwert \bar{y} und eine andere Standardabweichung s . Daraus errechnet man unterschiedliche Vertrauensbereiche für den Mittelwert (vgl. Beispiel 1 und 2), d.h. der Vertrauensbereich unterliegt dem Zufall. Im Mittel über viele Versuche enthält z.B. ein Anteil von 95 % der mit (6.9) berechneten 95 %-Vertrauensbereiche den wahren, aber unbekannten Mittelwert der Verteilung. Bild 6-14 zeigt diese Situation grafisch.⁸
- Der Vertrauensbereich für den Mittelwert überdeckt den wahren **Mittelwert** μ mit einer vorgegebenen Wahrscheinlichkeit. Der Vertrauensbereich darf nicht verwechselt werden mit dem Bereich, in dem einzelne Versuchsergebnisse liegen. Einzelergebnisse streuen über einen wesentlich breiteren Bereich.

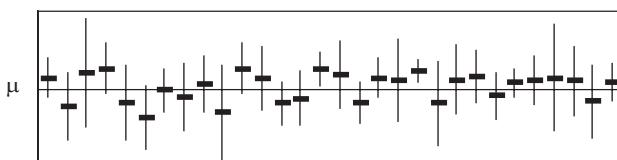


BILD 6-14 Grafische Darstellung der 95 %-Vertrauensbereiche für den Mittelwert μ , die aus 30 verschiedenen Stichproben (Versuchen) des Umfangs $n = 4$ berechnet wurden:

Aus jedem Versuch erhält man einen anderen Stichprobenmittelwert \bar{y} (kurze waagrechte Linie) und eine andere Stichprobenstandardabweichung s (bestimmt die Breite des Vertrauensbereichs). 28 der 30 Vertrauensbereiche überdecken den wahren, aber normalerweise unbekannten Wert μ , d.h. ca. 95 %.

⁸ Die Visualisierung „Vertrauensbereich für den Mittelwert“ zeigt ähnliche Darstellungen für Stichproben mit beliebigem Umfang aus einer beliebigen Normalverteilung – vgl. Anhang G.

Vertrauensbereiche für andere Größen

Wie hier für den Mittelwert gezeigt, kann man Vertrauensbereiche auch für andere Größen berechnen. Der folgende Vergleich von zwei Mittelwerten und die Regression im Kapitel 10 benutzen ebenfalls Vertrauensbereiche. Die Interpretation ist immer dieselbe:

- Der Vertrauensbereich wird aus einem Stichprobenergebnis berechnet und unterliegt daher dem Zufall.
- Der Vertrauensbereich überdeckt den wahren, festen, aber unbekannten Wert der gesuchten Größe mit einer Wahrscheinlichkeit, die mit dem Vertrauensniveau vorgegeben wird.
- Für einen einzelnen Vertrauensbereich weiß man nicht, ob er „richtig“ oder „falsch“ ist. D.h. man kann nicht sagen, ob gerade dieser Vertrauensbereich den wahren Wert überdeckt oder nicht.
- Wird der gesamte Versuch wiederholt, so erhält man einen anderen Vertrauensbereich. Dies ist kein Fehler, sondern eine Folge der unvermeidlichen Zufallsstreuung.

Der Vertrauensbereich ist umso schmäler, je kleiner die Zufallsstreuung und je größer der Stichprobenumfang n ist.

■ 6.3 Vergleich von zwei Mittelwerten

Häufig sollen zwei Verfahren, Prozess- oder Produktvarianten, Anlagen o.ä. miteinander verglichen werden. Man möchte wissen, ob sich die wahren, aber unbekannten Mittelwerte der beiden Gruppen (statistisch: der beiden Grundgesamtheiten) unterscheiden oder nicht.



Beispiele

- Teile werden in zwei parallelen Fertigungslinien hergestellt. Gibt es einen Unterschied zwischen den Längen, Massen, Dichten, Farben, Drücken usw. der Teile von den beiden Linien?
- Ein Qualitätsmerkmal wird mit zwei verschiedenen Verfahren gemessen. Unterscheiden sich die Ergebnisse?
- An einer Anlage wurde eine Wartungsarbeit durchgeführt. Sind die Ergebnisse nachher wieder so wie vorher?
- Die Entwicklungsabteilung hat ein neues Verfahren entwickelt. Sind die Ergebnisse wirklich besser als mit dem alten Verfahren?
- Eine neue Materialcharge wird eingesetzt. Ist die Dichte der damit hergestellten Teile gleich geblieben?
- Hat die Temperatur einen Einfluss auf die Ausbeute einer Reaktion? Bisher wurde z. B. eine Temperatur von 120 °C verwendet, aber sind 140 °C vielleicht besser?

In diesem Abschnitt wird gezeigt:

1. wie man Versuchsergebnisse auswerten kann, wenn sie bereits vorliegen,
2. wie viele Einzelversuche man durchführen sollte (d.h. Festlegung des Stichprobenumfangs n), wenn man einen bestimmten Unterschied zwischen den Gruppen mit hoher Wahrscheinlichkeit erkennen möchte und
3. welche Voraussetzungen erfüllt sein müssen, und wie man sicherstellen bzw. überprüfen kann, dass sie erfüllt sind.

6.3.1 Auswertung von Versuchsergebnissen

Zunächst gehen wir davon aus, dass die Versuchsergebnisse bereits vorliegen:

- n Messwerte aus Gruppe 1, mit Mittelwert \bar{y}_1 und Varianz s_1^2 und
- n Messwerte⁹ aus Gruppe 2, mit Mittelwert \bar{y}_2 und Varianz s_2^2 .

Beim Vergleich von zwei Gruppen beträgt die Gesamtzahl N der Messwerte somit

$$\text{Gesamtzahl } N = 2 \cdot n. \quad (6.11)$$

n heißt Stichprobenumfang, N Versuchsumfang.

Unter der Voraussetzung, dass

- die Einzelwerte repräsentativ für die beiden Gruppen sind,
- die Einzelwerte für jede Gruppe einzeln normalverteilt sind und
- die Standardabweichung für beide Gruppen gleich ist,

kann man folgendermaßen vorgehen:

1. Man berechnet die Differenz zwischen den beiden Stichprobenmittelwerten (in der Versuchsplanung heißt diese Differenz „Effekt“)

$$\bar{d} = \bar{y}_2 - \bar{y}_1. \quad (6.12)$$

\bar{d} ist ein Schätzwert für die wahre Differenz δ der Mittelwerte der beiden Verfahren.

2. Man berechnet den Vertrauensbereich¹⁰ für die Differenz δ

$$\bar{d} - t \cdot s_{\bar{d}} \leq \delta \leq \bar{d} + t \cdot s_{\bar{d}} \quad (6.13)$$

⁹ Die Vorgehensweise lässt sich ohne Probleme auch auf ungleiche Stichprobenumfänge erweitern, allerdings werden die Formeln dann etwas komplizierter (vgl. z.B. [5]). Der Vergleich ist jedoch am effizientesten, wenn die Stichprobenumfänge n gleich sind. Bei geplanten Versuchen achtet man daher auf gleiche n . Gute Software berücksichtigt auch ungleiche Stichprobenumfänge, ohne dass sich der Anwender darum kümmern muss.

¹⁰ (6.13) entspricht offensichtlich (6.9). Verglichen mit (6.8) enthält (6.14) einen zusätzlichen Faktor $\sqrt{2}$, da die Varianz der Differenz von zwei Mittelwerten doppelt so groß ist wie die Varianz eines Mittelwerts. (6.15) enthält als Varianz den Mittelwert über die beiden Gruppen. Der Freiheitsgrad in (6.16) ist die Summe der Freiheitsgrade der beiden Gruppen – oder andersherum betrachtet: Bei der Berechnung der Streuung wurden von den insgesamt N Einzelwerten 2 Gruppenmittelwerte abgezogen.

wobei

$$s_d = \sqrt{\frac{2}{n} \cdot s^2} = \sqrt{\frac{4}{N} \cdot s^2} = \text{Standardabweichung des Effekts (standard error)} \quad (6.14)$$

$$s^2 = \frac{s_1^2 + s_2^2}{2} = \text{Varianz der Einzelwerte (mittel)} \quad (6.15)$$

$$f = 2 \cdot (n - 1) = N - 2 = \text{Freiheitsgrad.} \quad (6.16)$$

3. Enthält der Vertrauensbereich den Wert 0, so kann die wahre Differenz $\delta = 0$ sein. Die Daten sind konsistent damit, dass kein Unterschied zwischen den beiden Gruppen besteht. Die beobachtete Differenz kann auch zufällig aufgetreten sein. Man sagt, der Effekt ist nicht signifikant.¹¹

Enthält der Vertrauensbereich den Wert 0 nicht, so geht man davon aus, dass wirklich ein Unterschied besteht. Man sagt, der Effekt ist signifikant.

Dabei ist es üblich, die 95 %-, 99 %- und 99,9 %-Vertrauensbereiche zu berechnen, und das Ergebnis wie folgt zu beurteilen (vgl. z.B. DGO-Lehrgänge):

Wenn bereits der 95 %-Vertrauensbereich den Wert 0 enthält, wird das Ergebnis mit – bewertet. Wenn nicht einmal der 99,9 %-Vertrauensbereich den Wert 0 enthält, wird das Ergebnis mit *** bewertet. Dazwischen erhält man folgende Abstufung:

- 95 % * 99 % ** 99,9 % ***

Dabei bedeutet:

- kein Hinweis auf Unterschied, wenn der Versuchsumfang groß genug war
- * indifferent, möglichst mehr Daten sammeln
- ** signifikanter Unterschied
- *** hochsignifikanter Unterschied



Beispiel

Eine Firma bezieht eine bestimmte Stahlsorte von zwei verschiedenen Herstellern. Sie möchte wissen, ob sich die Hersteller bezüglich der Streckgrenze unterscheiden. Dazu entnimmt sie verteilt über einen längeren Zeitraum je 11 Proben und bestimmt deren Streckgrenzen. Aus den Einzelwerten wurde berechnet:

$$\text{Hersteller 1: } \bar{y}_1 = 312 \text{ N/mm}^2 \quad s_1 = 21 \text{ N/mm}^2$$

$$\text{Hersteller 2: } \bar{y}_2 = 345 \text{ N/mm}^2 \quad s_2 = 25 \text{ N/mm}^2$$

Besteht ein Unterschied zwischen den Streckgrenzen der beiden Hersteller?

¹¹ Für statistisch Interessierte:

In der Literatur (z.B. [5]) ist es üblich, den Mittelwertvergleich als Testproblem zu formulieren (t-Test):

Nullhypothese: Effekt = 0

Alternativhypothese: Effekt $\neq 0$

Die Prüfgröße $\frac{|\bar{y}_2 - \bar{y}_1|}{s_d}$ wird mit dem t-Wert aus Tabelle 6.4 verglichen,

je nach Ergebnis wird die Nullhypothese verworfen oder nicht verworfen. Diese Vorgehensweise führt zum selben Ergebnis wie die hier vorgestellte Variante, ist m. E. jedoch weniger anschaulich.

$$\bar{d} = 345 - 312 = 33 \text{ N/mm}^2 = \text{Effekt}$$

$$s = \sqrt{\frac{21^2 + 25^2}{2}} = \sqrt{533} = 23,1 = \text{Standardabweichung der Einzelwerte (mittel)}$$

$$N = 2 \cdot 11 = 22 \Rightarrow s_d = \sqrt{\frac{4}{22}} \cdot 23,1 = 9,84 \text{ N/mm}^2 = \text{Standardabweichung des Effekts}$$

$$f = 22 - 2 = 20 = \text{Freiheitsgrad}$$

Mit Tabelle 6.4 erhält man:

$$95\%: \quad t = 2,086 \Rightarrow t \cdot s_d = 2,086 \cdot 9,84 = 20,5 \text{ N/mm}^2 \Rightarrow \\ (33 - 20,5) = 12,5 \text{ N/mm}^2 \leq \delta \leq 53,5 \text{ N/mm}^2 \quad (= 33 + 20,5)$$

$$99\%: \quad t = 2,845 \Rightarrow t \cdot s_d = 2,845 \cdot 9,84 = 28,0 \text{ N/mm}^2 \Rightarrow \\ (33 - 28,0) = 5,0 \text{ N/mm}^2 \leq \delta \leq 61,0 \text{ N/mm}^2 \quad (= 33 + 28,0)$$

$$99,9\%: \quad t = 3,850 \Rightarrow t \cdot s_d = 3,850 \cdot 9,84 = 37,9 \text{ N/mm}^2 \Rightarrow \\ (33 - 37,9) = -4,9 \text{ N/mm}^2 \leq \delta \leq 70,9 \text{ N/mm}^2 \quad (= 33 + 37,9)$$

Der 99 %-Vertrauensbereich enthält den Wert 0 nicht, aber der 99,9 %-Vertrauensbereich enthält den Wert 0.

\Rightarrow ** Es besteht ein signifikanter Unterschied zwischen den beiden Herstellern.

Zum selben Ergebnis kommt man auch, wenn man den Betrag (= die Größe) des Effekts, hier 33, direkt mit der Breite der Vertrauensbereiche vergleicht, hier 20,5 – 28,0 und 37,9.



Aufgabe 1

Sie wollen wissen, ob der Nachdruck beim Druckgießen eines bestimmten Teils Einfluss auf die Dichte des Teils hat (je größer die Dichte, desto weniger Lunker und Fehlstellen enthält ein Teil). Dazu gießen Sie je 3 Teile bei einem Nachdruck von 300 bar und 400 bar (in zufälliger Reihenfolge) und messen die Dichte (in g/cm^3). Sie erhalten folgende Ergebnisse:

300 bar: 2,635 2,639 2,637

400 bar: 2,641 2,640 2,645

Skizzieren Sie diese Werte, und überlegen Sie zunächst intuitiv, ob Sie glauben, dass der Druck einen Einfluss auf die Dichte hat.

Berechnen Sie anschließend den Effekt des Drucks und die 95 %-, 99 %-, und 99,9 %-Vertrauensbereiche für den Effekt, und beurteilen Sie so die Signifikanz des Effekts.

Lösung

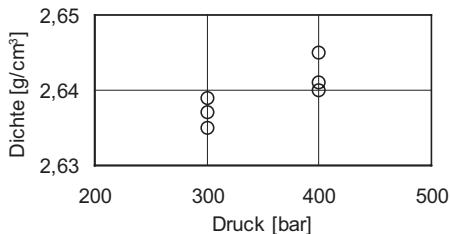


BILD 6-15

Grafische Darstellung der je drei Messwerte für die Dichte bei Druck 300 und 400 bar

$$300 \text{ bar: } \bar{y}_1 = 2,637 \text{ g/cm}^3 \quad s_1 = 0,002 \text{ g/cm}^3$$

$$400 \text{ bar: } \bar{y}_2 = 2,642 \text{ g/cm}^3 \quad s_2 = 0,00265 \text{ g/cm}^3$$

$$\bar{d} = 2,642 - 2,637 = 0,005 \text{ g/cm}^3 = \text{Effekt}$$

$$s = \sqrt{\frac{0,002^2 + 0,00265^2}{2}} = 0,00235 \text{ g/cm}^3 = \text{Standardabweichung der Einzelwerte}$$

$$N = 6 \Rightarrow s_{\bar{d}} = \sqrt{\frac{4}{6}} \cdot 0,00235 = 0,00191 \text{ g/cm}^3 = \text{Standardabweichung des Effekts}$$

$$f = 6 - 2 = 4 = \text{Freiheitsgrad}$$

Mit Tabelle 6.4 erhält man:

$$95\%: \quad t = 2,776 \quad \Rightarrow t \cdot s_{\bar{d}} = 2,776 \cdot 0,00191 = 0,0053 \text{ g/cm}^3 \Rightarrow \\ -0,0003 \text{ g/cm}^3 \leq \delta \leq 0,0103 \text{ g/cm}^3$$

Bereits der 95 %-Vertrauensbereich enthält den Wert 0.

\Rightarrow – Die beobachteten Unterschiede sind kein ausreichender Hinweis darauf, dass der Druck die Dichte beeinflusst.

Intuitiv vermutet man aufgrund der Messwerte in Bild 6-15, dass der Druck einen Einfluss auf die Dichte hat. Dazu sind zwei Bemerkungen angebracht:

1. Liegen nur wenige Versuchsergebnisse vor, so überschätzt man intuitiv die Bedeutung der Ergebnisse. Insbesondere wenn man bereits vorher an eine Abhängigkeit glaubt, neigt man dazu, ein Ergebnis wie das obige als Beweis zu betrachten. Umgekehrt unterschätzt man bei großem Versuchsumfang die Bedeutung von Unterschieden. Statistik schützt vor überreilten Schlüssen.
2. Die Aussage, dass der beobachtete Unterschied nicht signifikant ist, heißt nicht, dass es keinen Unterschied gibt. Sie bedeutet nur, dass der beobachtete Unterschied mit einer Wahrscheinlichkeit von >5 % auch zufällig auftritt. Untersucht man mehr Teile, so kann sich durchaus zeigen, dass ein signifikanter Unterschied besteht. Der Versuchsumfang ist zu klein, um zu einer sinnvollen Entscheidung zu kommen.

Aufgabe 2

Wie Aufgabe 1, nur dass je 11 Teile gegossen wurden. Ergebnisse (nur Mittelwerte und Standardabweichungen werden angegeben):

$$300 \text{ bar: } \bar{y}_1 = 2,637 \text{ g/cm}^3 \quad s_1 = 0,002 \text{ g/cm}^3$$

$$400 \text{ bar: } \bar{y}_2 = 2,642 \text{ g/cm}^3 \quad s_2 = 0,00265 \text{ g/cm}^3 \quad (\text{wie in Aufgabe 1})$$

Berechnen Sie den Effekt des Drucks und die 95 %-, 99 %-, und 99,9 %-Vertrauensbereiche für den Effekt, und beurteilen Sie so die Signifikanz des Effekts. Vergleichen Sie das Ergebnis mit Aufgabe 1.

Lösung

$$\bar{d} = 2,642 - 2,637 = 0,005 \text{ g/cm}^3 = \text{Effekt (wie in Aufgabe 1)}$$

$$s = \sqrt{\frac{0,002^2 + 0,00265^2}{2}} = 0,00235 \text{ g/cm}^3 = \text{Standardabweichung der Einzelwerte}$$

$$N = 22 \Rightarrow s_{\bar{d}} = \sqrt{\frac{4}{22}} \cdot 0,00235 = 0,0010 \text{ g/cm}^3 = \text{Standardabweichung des Effekts}$$

$$f = 22 - 2 = 20 = \text{Freiheitsgrad}$$

Mit Tabelle 6.4 erhält man:

$$95\%: t = 2,086 \Rightarrow t \cdot s_d = 2,086 \cdot 0,0010 = 0,0021 \text{ g/cm}^3 \Rightarrow \\ 0,0029 \text{ g/cm}^3 \leq \delta \leq 0,0071 \text{ g/cm}^3$$

$$99\%: t = 2,845 \Rightarrow t \cdot s_d = 2,845 \cdot 0,0010 = 0,0028 \text{ g/cm}^3 \Rightarrow \\ 0,0022 \text{ g/cm}^3 \leq \delta \leq 0,0078 \text{ g/cm}^3$$

$$99,9\%: t = 3,850 \Rightarrow t \cdot s_d = 3,850 \cdot 0,0010 = 0,0038 \text{ g/cm}^3 \Rightarrow \\ 0,0012 \text{ g/cm}^3 \leq \delta \leq 0,0088 \text{ g/cm}^3$$

Auch der 99,9 %-Vertrauensbereich enthält den Wert 0 nicht.

$\Rightarrow \star\star\star$ Der Druck hat einen hochsignifikanten Einfluss auf die Dichte.

Je nach Stichprobenumfang kann derselbe Effekt bei derselben Standardabweichung als zufällig oder als hochsignifikant eingestuft werden. Der Grund für diese unterschiedliche Bewertung liegt darin, dass mit zunehmendem Stichprobenumfang der Vertrauensbereich schmäler wird. Bei größerem Stichprobenumfang überdeckt der Vertrauensbereich den Wert 0 nicht mehr.

Grafische Darstellung

Bilder 6-16 und 6-17 zeigen zwei alternative Möglichkeiten, die Signifikanz von Effekten grafisch darzustellen.

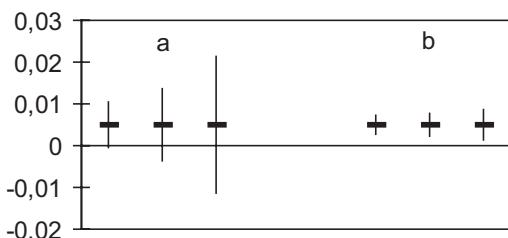


BILD 6-16
95-, 99%- und 99,9 %-Vertrauensbereiche
für den Druck im Vergleich zum Wert 0
a je drei Messwerte
b je elf Messwerte

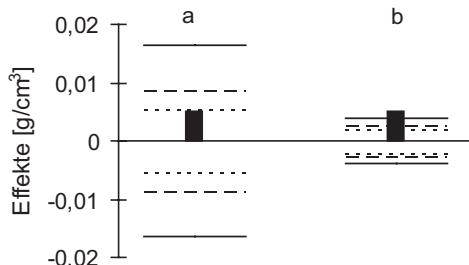


BILD 6-17
Effekt des Drucks und Breite des 95%-
(Punkte), 99%- (Striche) und des
99,9 %-Vertrauensbereichs (Linien)
a je drei Messwerte
b je elf Messwerte

In Bild 6-16 wird der Vertrauensbereich (senkrechte Linie) wie oben beschrieben direkt mit dem Wert 0 verglichen. Der Vertrauensbereich liegt symmetrisch zum beobachteten Effekt (waagrechte Linie). Der Vertrauensbereich überdeckt mit der vorgegebenen Wahrscheinlichkeit den wahren Wert des Effekts. Enthält der Vertrauensbereich den Wert 0 nicht, so ist der Effekt auf dem zugehörigen Niveau signifikant.

Bild 6-16 a zeigt das Ergebnis von Aufgabe 1:

Alle drei Vertrauensbereiche enthalten den Wert 0. Der beobachtete Effekt des Drucks kann auch zufällig auftreten, ohne dass wirklich ein Unterschied besteht (Bewertung -).

Bild 6-16 b zeigt das Ergebnis von Aufgabe 2:

Keiner der drei Vertrauensbereiche enthält den Wert 0 (Bewertung ***: hochsignifikant). Nach einem solchen Ergebnis ist man überzeugt davon, dass die Dichte wirklich vom Druck abhängt. Dabei hat sich die Größe des Effekts im Vergleich zu Bild 6-16 a nicht verändert – aufgrund des größeren Stichprobenumfangs ist der Vertrauensbereich jedoch schmäler.

In Bild 6-17 wird der Effekt mit der Breite der 95 %-, 99 %- und 99,9 %-Vertrauensbereiche verglichen. Die waagrechten Linien liegen bei $\pm t \cdot s_{\bar{d}}$. Ragt ein Effekt (als Balken dargestellt) aus diesem Band heraus, so ist er auf dem zugehörigen Niveau signifikant.

In Bild 6-17 a liegt der Balken (knapp) innerhalb des 95 %-Vertrauensbereichs (Bewertung –), in Bild 6-17 b ragt der Balken sogar über den 99,9 %-Vertrauensbereich hinaus, der Effekt ist hochsignifikant (***).

Der Effekt ist unverändert, grafisch zeigt sich das in der unveränderten Länge des Balkens. Aber die Breite der Vertrauensbereiche in b ist aufgrund des größeren Versuchsumfangs kleiner.

Die Darstellungen in Bild 6-16 und 6-17 sind völlig gleichwertig. Im Folgenden wird die Darstellung von Bild 6-17 verwendet, da in dieser Darstellung alle drei Vertrauensniveaus in einem Bild zusammengefasst sind und daher das gesamte Testergebnis mit einem Blick erfasst werden kann.

Anmerkung zum Mittelwertvergleich mit Software

Mit den meisten Statistikpaketen können Vertrauensbereiche für die Differenz der Mittelwerte (den Effekt) zu vorgegebenen Vertrauensniveaus berechnet werden, wie hier beschrieben. Durch Vergleich der 95 %-, 99 %- und 99,9 %-Vertrauensbereiche mit 0 erhält man die oben beschriebene Bewertung.

Ergänzend wird von den meisten Statistikpaketen als Entscheidungshilfe eine „Wahrscheinlichkeit“ (p , probability level, $prob>t$) oder ein „Signifikanzniveau“ (significance -level) berechnet. Dies ist die Wahrscheinlichkeit dafür, dass der beobachtete Effekt oder ein noch größerer Effekt nur zufällig auftritt, wenn der wahre Effekt 0 ist. Je kleiner diese Wahrscheinlichkeit ist, desto mehr wird man davon überzeugt sein, dass wirklich ein Unterschied zwischen den beiden Gruppen besteht. Somit erhält man folgende Beurteilung des Ergebnisses:

| | | |
|---------------------------------------|-----|-----------------------------------|
| Wahrscheinlichkeit > 5 % (bzw. 0,05): | – | kein Hinweis auf Unterschied |
| 5 % > Wahrscheinlichkeit > 1 %: | * | indifferent, möglichst mehr Daten |
| 1 % > Wahrscheinlichkeit > 0,1 %: | ** | signifikanter Unterschied |
| 0,1 % > Wahrscheinlichkeit: | *** | hochsignifikanter Unterschied |

Die Angabe der Wahrscheinlichkeit erlaubt die Entscheidung ohne Tabelle 6.4 und stellt somit eine wesentliche Erleichterung dar.

6.3.2 Festlegung des Stichproben- bzw. Versuchsumfangs

Je größer der Stichprobenumfang ist, desto schmäler wird der Vertrauensbereich für den Effekt, d.h. desto schmäler wird der Bereich, der den wahren Effekt mit vorgegebener Wahrscheinlichkeit enthält. Will man daher bereits einen kleinen wahren Effekt erkennen, so benötigt man einen großen Stichprobenumfang.



Beispiel

Bild 6-18 zeigt die 99 %-Vertrauensbereiche für den Effekt, die bei 30 „simulierten Versuchen“ mit Stichprobenumfang $n = 3$ (Versuchsumfang $N = 6$) ermittelt wurden. Bei der Simulation wurden ein „wahrer Effekt“ von 1 und eine Standardabweichung von 1 zugrunde gelegt. Die meisten Vertrauensbereiche enthalten den Wert 0. Der Versuchsumfang ist zu klein, um den tatsächlich vorhandenen Unterschied zu erkennen.¹²

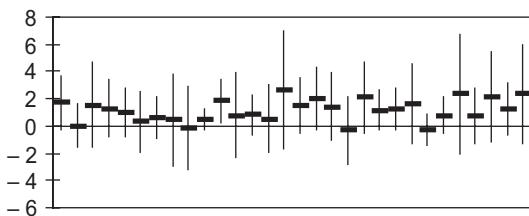


BILD 6-18

99 %-Vertrauensbereiche für die Differenz von zwei Gruppen (Effekt), 30 Zufallsstichproben, Versuchsumfang je $N = 6$, $\Delta\mu = 1$, $\sigma = 1$

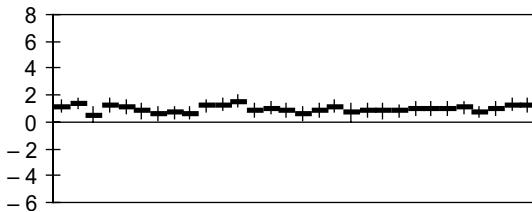


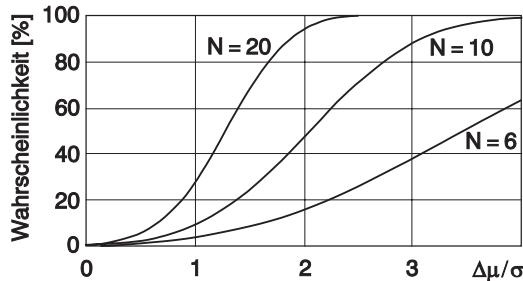
BILD 6-19

Wie Bild 6-18, aber Versuchsumfang je $N = 60$

Bild 6-19 zeigt die Vertrauensbereiche, die unter ansonsten gleichen Bedingungen wie in Bild 6-18 mit Stichprobenumfang $n = 30$ (Versuchsumfang $N = 60$) ermittelt wurden. Die Vertrauensbereiche sind deutlich schmäler, die meisten Vertrauensbereiche enthalten den Wert 0 nicht mehr, d.h. bei diesem größeren Stichprobenumfang erkennt man in den meisten Fällen, dass ein Unterschied zwischen den beiden Gruppen besteht.

Die Wahrscheinlichkeit, mit der ein tatsächlich vorhandener Unterschied $\Delta\mu$ erkannt wird (d.h. die Wahrscheinlichkeit dafür, dass der 99 %-Vertrauensbereich den Wert 0 nicht enthält), hängt vom Versuchsumfang N und vom Verhältnis $\Delta\mu/\sigma$ ab. Bild 6-20 zeigt diese Abhängigkeit für ausgewählte Werte von N .

¹² Die Visualisierung „Vergleich von zwei Mittelwerten“ zeigt, wie sich der wahre Unterschied der Mittelwerte $\Delta\mu$, die Standardabweichung σ und der Stichprobenumfang n auf den Mittelwertvergleich auswirken, sowohl anhand der Verteilungen als auch mit simulierten Stichprobenergebnissen – vgl. Anhang G.

**BILD 6-20**

Wahrscheinlichkeit, mit der ein tatsächlich vorhandener Unterschied $\Delta\mu$ erkannt wird, in Abhängigkeit vom Verhältnis $\Delta\mu/\sigma$, für Versuchsumfang $N = 6, 10$ und 20

Je größer der Versuchsumfang N und das Verhältnis $\Delta\mu/\sigma$ sind, mit desto größerer Wahrscheinlichkeit wird der Unterschied gefunden.



Beispiel

Aus Bild 6-20 kann man ablesen, dass ein Unterschied von $\Delta\mu = 1\sigma$ bei einem Versuchsumfang von $N = 6$ ($n = 3$) mit einer Wahrscheinlichkeit von nur 4% gefunden wird, bei $N = 10$ mit 9% und bei $N = 20$ mit 28%. Ein Unterschied von $\Delta\mu = 2\sigma$ wird dagegen bei einem Versuchsumfang von $N = 6$ mit einer Wahrscheinlichkeit von 15% gefunden, bei $N = 10$ mit 47% und bei $N = 20$ mit 94%.

Bei der Versuchsplanung stellt sich meist die umgekehrte Frage: Welcher Versuchsumfang ist erforderlich, wenn ein bestimmter Unterschied mit hoher Wahrscheinlichkeit (meist verwendet man 90 %) gefunden werden soll, falls er existiert.

In Bild 6-20 bedeutet dies: Für welchen Wert von N erhält man für ein vorgegebenes Verhältnis $\Delta\mu/\sigma$ eine Wahrscheinlichkeit von 90 % (0,90)?

Man kann zeigen, dass bei einem Stichprobenumfang n bzw. einem Versuchsumfang $N = 2 \cdot n$ (= Gesamtzahl der Messwerte) von

$$\text{Versuchsumfang } N = 2 \cdot n = 60 \cdot \left(\frac{\sigma}{\Delta\mu} \right)^2 \quad (6.17)^{13}$$

ein wahrer Unterschied zwischen den Gruppen (Effekt) $\Delta\mu$ mit hoher Wahrscheinlichkeit entdeckt wird, wenn die Standardabweichung der Einzelwerte σ beträgt. Genauer gesagt, im Mittel über viele Versuche enthalten 90 % der 99 %-Vertrauensbereiche für den Effekt nicht den Wert 0, wenn der wahre Wert $\Delta\mu$ beträgt.¹⁴

¹³ (6.17) ist eine Näherung und gilt für große Werte von N . Für kleine Werte ($N < \text{ca. } 40$) sollte man einen etwas größeren Versuchsumfang verwenden.

¹⁴ Für statistisch Interessierte:

In der Testtheorie werden folgende Begriffe und Bezeichnungen verwendet:

Der Vertrauensbereich enthält nicht den Wert 0, obwohl kein Unterschied besteht: Fehler 1. Art.

Ein tatsächlich vorhandener Unterschied wird nicht erkannt:

Fehler 2. Art.

Wahrscheinlichkeit für den Fehler 1. Art: Signifikanzniveau α ,

hier ist $\alpha = 1\%$.

Wahrscheinlichkeit für den Fehler 2. Art: β ,

hier ist $\beta = 10\%$.

Für andere α und β wird der Faktor 60 ersetzt durch $4 \cdot (u_{1-\alpha/2} + u_{1-\beta})^2$



Beispiel

Eine Firma bezieht eine bestimmte Stahlsorte von zwei verschiedenen Herstellern. Sie möchte wissen, ob sich die Hersteller bezüglich der Streckgrenze unterscheiden. Aus Erfahrung ist bekannt, dass die Standardabweichung der Streckgrenze ca. $\sigma = 20 \text{ N/mm}^2$ beträgt. Ein Unterschied von $\Delta\mu = 30 \text{ N/mm}^2$ ist technologisch relevant und soll mit hoher Wahrscheinlichkeit erkannt werden, falls er existiert.

Aus (6.17) erhält man für den erforderlichen Versuchsumfang

$$N = 60 \cdot \left(\frac{20}{30} \right)^2 = 26,7 \approx 28$$

d. h. von jedem Hersteller sollten ca. 14 Proben untersucht werden.

Ist bereits ein Unterschied von $\Delta\mu = 10 \text{ N/mm}^2$ technologisch relevant, so erhält man aus (6.17)

$$N = 60 \cdot \left(\frac{20}{10} \right)^2 = 240$$

d. h. von jedem Hersteller sollten ca. 120 Proben untersucht werden. Je kleiner der Unterschied ist, desto erkannt werden soll, desto größer ist der erforderliche Versuchsumfang.



Aufgabe

Beim Druckgießen beträgt die Standardabweichung der Dichte erfahrungsgemäß ca. $0,0025 \text{ g/cm}^3$. Es soll untersucht werden, ob sich die Dichte von Teilen, die mit einem Nachdruck von 400 bar gegossen wurden von der unterscheidet, die mit 300 bar gegossen wurden. Ein Unterschied von $0,002 \text{ g/cm}^3$ soll mit hoher Wahrscheinlichkeit entdeckt werden, falls er existiert. Welchen Versuchsumfang würden Sie verwenden?

Lösung

$$N = 60 \cdot \left(\frac{0,0025}{0,002} \right)^2 = 93,75 \approx 94$$

Es ist sinnvoll, je 47 (oder 50) Teile bei 300 bar und bei 400 bar zu gießen.

Achtung: Man darf jedoch auf keinen Fall erst alle 50 Teile bei einem Druck und dann alle 50 Teile beim anderen Druck gießen, obwohl dies von der Durchführung her natürlich am einfachsten wäre. Randomisieren oder wenigstens mehrmals wechseln!

6.3.3 Voraussetzungen

In Abschnitt 6.3.1 wurde darauf hingewiesen, dass (6.13) nur gilt, wenn

- die Einzelwerte repräsentativ für die beiden Gruppen sind,
- die Einzelwerte für jede Gruppe einzeln normalverteilt sind und
- die Standardabweichung für beide Gruppen gleich ist.

Die erste dieser Bedingungen ist die wichtigste und die am häufigsten missachtete. Auf ihre Einhaltung ist bei der Planung der Versuche zu achten. Die anderen beiden Bedin-

gungen werden im Nachhinein überprüft. Für verteilungsunabhängige Vergleiche (z.B. Mann-Whitney-Wilcoxon) und Vergleiche von Zähldaten (z.B. u- oder χ^2 -Test) wird auf die Literatur verwiesen [1 – 6].

Repräsentative Stichproben

Eine Stichprobe sagt immer nur etwas aus über die Grundgesamtheit, aus der sie entnommen wurde. Für den Vergleich von zwei Verfahren bedeutet dies z.B.: Die beiden Stichproben müssen repräsentativ für die beiden Verfahren sein. Dabei helfen Blockbildung und Randomisierung (vgl. Absätze 3.4.3 und 3.4.4). Bei der Planung eines Versuchs muss man sich genau überlegen, was man erkennen will, und wie man daher vorgehen muss, damit das Ergebnis wirklich repräsentativ ist. Folgende Beispiele sollen auf Probleme und Möglichkeiten hinweisen.



Beispiel 1: Reihenfolge der Einzelversuche

Der Versuchsaufwand ist natürlich am geringsten, wenn zunächst alle Einzelversuche mit dem einen Verfahren und anschließend alle Einzelversuche mit dem anderen Verfahren durchgeführt werden. Dann kann jedoch ein Trend die Ergebnisse verfälschen, und zufällige Unterschiede zwischen Wiederholungen derselben Einstellung werden nicht erfasst. Besser ist es, die beiden Verfahren abwechselnd einzusetzen (Blockbildung).

Aber auch mit dieser Vorsichtsmaßnahme vergleicht man die beiden Verfahren unter den momentan herrschenden Bedingungen. Wenn sich an diesen Bedingungen etwas Grundlegendes ändert, kann dies natürlich den Vergleich beeinflussen.

Beispiel 2: Ausgangsmaterial

Es ist nicht zulässig, für ein Verfahren Ausgangsmaterial aus einer Charge zu verwenden und für das andere Verfahren Ausgangsmaterial aus einer anderen Charge. Dann könnte nämlich nicht unterschieden werden, ob ein evtl. gefundener signifikanter Unterschied durch die unterschiedlichen Verfahren oder durch die unterschiedlichen Materialchargen verursacht wird. Folgende Möglichkeiten bestehen:

- Alle Versuche für beide Verfahren werden mit Ausgangsmaterial derselben Charge durchgeführt. Dadurch wird die Zufallsstreuung reduziert, weil Unterschiede im Ausgangsmaterial nicht in die Zufallsstreuung eingehen. Wenn ein Verfahren jedoch für Material einer Charge besonders günstig ist, bedeutet das nicht notwendigerweise, dass es auch für Material anderer Chargen günstig ist. Dazu sind zumindest ergänzende technische Überlegungen erforderlich, Statistik kann dazu keine Aussagen machen.
- Ausgangsmaterial verschiedener Chargen wird zufällig in beiden Verfahren verwendet (Randomisierung). Dann gehen Unterschiede zwischen den Chargen in die Zufallsstreuung ein.
- Das Ausgangsmaterial wird als Blockfaktor verwendet. Das bedeutet hier, dass Ausgangsmaterial verschiedener Chargen verwendet wird – aber aus jeder Charge dieselbe Anzahl Teile für jedes Verfahren. Die Charge-Nummer des Ausgangsmaterials wird dann als zusätzlicher Faktor behandelt. Diese Vorgehensweise führt zur besten Absicherung der Ergebnisse.

Normalverteilung

Durch Eintragung der Messwerte ins Wahrscheinlichkeitsnetz kann überprüft werden, ob sie konsistent sind mit einer Normalverteilung (Abschnitt 6.2).

Für Messwerte aus zwei Gruppen kann man dies natürlich für jede der beiden Gruppen getrennt durchführen. Um bereits kleinere Abweichungen von der Normalverteilung zu erkennen, ist es wünschenswert, möglichst alle Messwerte zusammenzufassen. Da sich die Mittelwerte der beiden Gruppen jedoch unterscheiden können, sind die zusammengenfassten Messwerte von zwei Gruppen selbst meist nicht normalverteilt.

Daher zieht man von den Messwerten jeder Gruppe den jeweiligen Stichprobenmittelwert ab. Der Rest wird als Residuum bezeichnet. Diese Residuen kann man dann zusammen in einem Wahrscheinlichkeitsnetz auftragen. Liegen die Residuen näherungsweise auf einer Geraden, so sind die Werte konsistent damit, dass sie aus normalverteilten Grundgesamtheiten stammen.

Gruppe 1: Einzelwerte: $y_{11}, y_{12}, y_{13}, \dots, y_{1n}$

Mittelwert: \bar{y}_1

Residuen: $y_{11} - \bar{y}_1, y_{12} - \bar{y}_1, y_{13} - \bar{y}_1, \dots, y_{1n} - \bar{y}_1$

Gruppe 2: Einzelwerte: $y_{21}, y_{22}, y_{23}, \dots, y_{2n}$

Mittelwert: \bar{y}_2

Residuen: $y_{21} - \bar{y}_2, y_{22} - \bar{y}_2, y_{23} - \bar{y}_2, \dots, y_{2n} - \bar{y}_2$

(6.18)



Beispiel

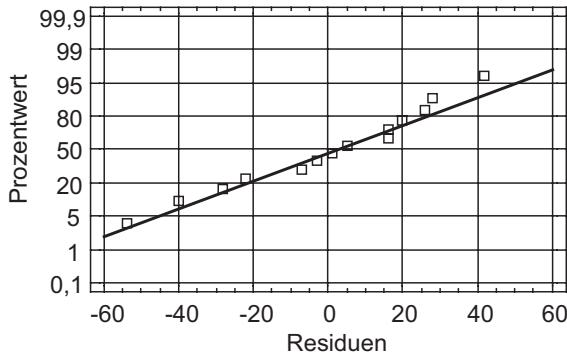
In einer Halbleiterfertigung werden für die Abscheidung einer SiO_2 -Schicht wahlweise zwei baugleiche Anlagen eingesetzt. Es soll untersucht werden, ob die Schichtdicke unterschiedlich ist. Tabelle 6.5 zeigt die Messwerte und die daraus berechneten Residuen.

TABELLE 6.5 Messwerte für die Schichtdicke (in nm) von je 7 zufällig entnommenen Scheiben von zwei Anlagen mit den daraus berechneten Residuen (Wert – Mittelwert)

| Nr. | Anlage 1 | Residuen 1 | Anlage 2 | Residuen 2 |
|----------------------|----------|------------|----------|------------|
| 1 | 440 | -54 | 494 | -28 |
| 2 | 472 | -22 | 548 | 26 |
| 3 | 499 | 5 | 523 | 1 |
| 4 | 514 | 20 | 538 | 16 |
| 5 | 510 | 16 | 519 | -3 |
| 6 | 487 | -7 | 550 | 28 |
| 7 | 536 | 42 | 482 | -40 |
| Mittelwert \bar{y} | 494 | | 522 | |

Mit Hilfe der Residuen können die insgesamt 14 Werte zusammen in einem Wahrscheinlichkeitsnetz dargestellt werden. Würde man die Messwerte selbst in einem Netz eintragen, wären sie nicht normalverteilt. In Bild 6-21 sind die Residuen aus Tabelle 6.5 eingetragen. Nach ihrer Größe geordnet sind dies:

| | | | | | | | | | | | | | | |
|------|-----|-----|-----|-----|----|----|---|---|----|----|----|----|----|----|
| Nr. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Wert | -54 | -40 | -28 | -22 | -7 | -3 | 1 | 5 | 16 | 16 | 20 | 26 | 28 | 42 |

**BILD 6-21**

Residuen (Abweichungen der Einzelwerte von den Gruppenmittelwerten) erlauben es, mehrere Gruppen in einem Wahrscheinlichkeitsnetz zusammenzufassen



Aufgabe

Sie wollen wissen, ob der Nachdruck beim Druckgießen eines bestimmten Teils Einfluss auf die Dichte hat. Dazu gießen Sie je drei Teile bei 300 bar und bei 400 bar Nachdruck. Sie erhalten folgende Ergebnisse (in g/cm^3):

300 bar: 2,635 2,639 2,637

400 bar: 2,641 2,640 2,645

Berechnen Sie die Residuen.

Lösung

300 bar: Mittelwert: 2,637 \Rightarrow Residuen -0,002 0,002 0,000

400 bar: Mittelwert: 2,642 \Rightarrow Residuen -0,001 -0,002 0,003

Residuen nach der Größe geordnet für die Eintragung ins Wahrscheinlichkeitsnetz:

| Nr. | 1 | 2 | 3 | 4 | 5 | 6 |
|------|--------|--------|--------|-------|-------|-------|
| Wert | -0,002 | -0,002 | -0,001 | 0,000 | 0,002 | 0,003 |

Gleiche Standardabweichung

Wenn sich die Standardabweichungen der beiden Gruppen sehr deutlich unterscheiden, so kann dies auch bei der Eintragung der Residuen ins Wahrscheinlichkeitsnetz erkannt werden.

Ergänzend gibt es den sogenannten F-Test, der darauf basiert, dass bei gleicher Standardabweichung das Verhältnis der Varianzen der beiden Gruppen in der Nähe von 1 liegt. Ein Verhältnis, das sehr stark von 1 abweicht, ist nicht konsistent mit gleichen Standardabweichungen. Diese Grundidee des F-Tests ist auch die Basis der Varianzanalyse in Kapitel 12. Der F-Test ist in den meisten Statistikpaketen implementiert.



Hinweis

Der Vergleich von zwei Mittelwerten ist relativ unempfindlich gegenüber moderaten Abweichungen von der Normalverteilung und ungleichen Standardabweichungen. Große Abweichungen können durch Transformationen reduziert werden (Einzelheiten in Abschnitt 6.4).

■ 6.4 Transformation von Messwerten

Voraussetzung bei der Berechnung der Vertrauensbereiche in Abschnitt 6.2 und beim Vergleich von Mittelwerten in Abschnitt 6.3 ist die Normalverteilung der Messwerte (in Abschnitt 6.3 sind zusätzlich gleiche Standardabweichungen vorausgesetzt).

Große Abweichungen von der Normalverteilung können mit Hilfe einer Transformation der Messwerte reduziert werden. Häufig führt diese Transformation auch zu einer Reduktion der Unterschiede in der Standardabweichung.

Transformation der Messwerte bedeutet, dass statt der Messwerte selbst eine Größe verwendet wird, die sich durch eine einfache Umrechnung aus den Messwerten ergibt. Für alle weiteren Analysen werden dann diese umgerechneten Werte verwendet.

Im Folgenden werden einige besonders nützliche Transformationen behandelt. Abschnitt 15.1 behandelt weitere Transformationen. Eine Alternative zur Transformation von Messwerten sind die in Kapitel 17 beschriebenen Verfahren.

6.4.1 Logarithmische Normalverteilung

Die logarithmische Normalverteilung tritt auf, wenn viele Einflüsse sich multiplikativ auf den Messwert auswirken (Normalverteilung bei additiven Einflüssen, vgl. Abschnitt 6.1).

Die logarithmische Normalverteilung ist häufig bei der Beschreibung von Lebensdauerdaten in Zuverlässigkeitstests¹⁵ nützlich. Die Messwerte in Bild 6-13 stammen z.B. aus einer logarithmischen Normalverteilung. Charakteristisch für die logarithmische Normalverteilung ist, dass

- keine negativen Messwerte auftreten und
- die Verteilung der Messwerte einen langen „Schwanz“ bei hohen Werten hat.

Verwendet man statt der Messwerte selbst die Logarithmen der Messwerte, so sind diese normalverteilt. Alle Verfahren aus Abschnitt 6.2 und 6.3, aber auch aus den folgenden Kapiteln, sind für die Logarithmen dann anwendbar.



Beispiel

In einer Halbleiterfertigung werden Aluminium-Leiterbahnen durch Sputtern erzeugt. Die Entwicklungsabteilung hat zwei verschiedene Prozessvarianten entwickelt, die bzgl. der Lebensdauer verglichen werden sollen. Dazu werden mit beiden Varianten je 10 Teststrukturen erzeugt und einem beschleunigten Lebensdauertest unterzogen (erhöhte Temperatur und Stromdichte). Zu folgenden Zeitpunkten (in h) fielen die Strukturen aus:

| | | | | | | | | | | |
|-------------|----|----|----|----|-----|-----|-----|-----|-----|-----|
| Variante 1: | 27 | 49 | 67 | 90 | 134 | 148 | 164 | 245 | 330 | 735 |
| Variante 2: | 5 | 11 | 16 | 22 | 27 | 45 | 60 | 81 | 90 | 200 |

Man erkennt sofort, dass bei großen Zeiten die Ausfallzeitpunkte weiter auseinander liegen als bei kleinen Zeiten. Dies ist eine Folge der logarithmischen Normalverteilung der Werte. Die Transformation besteht darin, statt der Ausfallzeitpunkte ihre Logarithmen zu verwenden:

¹⁵ Alternativ wird die Weibullverteilung verwendet.

| | | | | | | | | | | |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| In (Variante 1): | 3,3 | 3,9 | 4,2 | 4,5 | 4,9 | 5,0 | 5,1 | 5,5 | 5,8 | 6,6 |
| In (Variante 2): | 1,6 | 2,4 | 2,8 | 3,1 | 3,3 | 3,8 | 4,1 | 4,4 | 4,5 | 5,3 |

Die weitere Analyse erfolgt wie in (6.12) bis (6.16), aber mit diesen Logarithmen:

$$\text{In (Variante 1): } \bar{y}_1 = 4,88 \quad s_1 = 0,962$$

$$\text{In (Variante 2): } \bar{y}_2 = 3,53 \quad s_2 = 1,106$$

$$d = \bar{y}_2 - \bar{y}_1 = 3,53 - 4,88 = -1,35$$

$$s_d = \sqrt{\frac{s_1^2 + s_2^2}{2}} = \sqrt{\frac{0,962^2 + 1,106^2}{2}} = 1,036$$

$$s_d = \sqrt{\frac{4}{N} \cdot s} = \sqrt{\frac{4}{20}} \cdot 1,036 = 0,463$$

$$f = N - 2 = 20 - 2 = 18$$

Mit Tabelle 6.4 erhält man (mit Interpolation):

$$95\%: t \cdot s_d = 2,10 \cdot 0,463 = 0,97$$

$$99\%: t \cdot s_d = 2,88 \cdot 0,463 = 1,33$$

$$99,9\%: t \cdot s_d = 3,92 \cdot 0,463 = 1,81$$

$$\Rightarrow 1,33 < |d| = 1,35 < 1,81 \quad \Rightarrow **$$

Es besteht ein signifikanter Unterschied zwischen den Varianten. Da das Ziel eine möglichst lange Lebensdauer ist, ist es besser, Variante 1 zu verwenden.

Die logarithmische Transformation der Messwerte hat noch eine weitere Auswirkung, die häufig erwünscht ist:

Bei logarithmisch normalverteilten Daten nimmt die Standardabweichung meist proportional zum Mittelwert zu. Dies ist auch im obigen Beispiel der Fall: Die Standardabweichung der Lebensdauer von Variante 1 ist deutlich größer als die Standardabweichung der Lebensdauer von Variante 2. Die Standardabweichungen der Logarithmen unterscheiden sich nur aufgrund der Zufallsstreuung. Durch die logarithmische Transformation wurden also die beiden statistischen Bedingungen für den Mittelwertvergleich gleichzeitig erfüllt:

- Normalverteilung und
- gleiche Standardabweichung (vgl. auch Absatz 10.1.3).

6.4.2 Poisson-Verteilung

Häufig ist der Messwert ein Zählergebnis, z.B. die Anzahl Lötfehler auf einer Leiterplatte, die Anzahl Astlöcher in einem Brett, die Anzahl Kratzer auf einer Glasplatte, die Anzahl Käufer eines bestimmten Produktes pro Tag.

Wenn die Einzelereignisse (z.B. Lötfehler) unabhängig voneinander sind, sind solche Zählergebnisse poissonverteilt. Wenn die Anzahl ausreichend groß ist, kann die Poisson-verteilung durch eine Normalverteilung angenähert werden, deren Standardabweichung aber mit dem Mittelwert zunimmt (wie $\sqrt{\mu}$).

Anschaulich äußert sich dies z.B. darin, dass der Unterschied zwischen 0 und 1 Fehlern bedeutender ist als der Unterschied zwischen 100 und 101 Fehlern. In der Analyse kann man dies dadurch berücksichtigen, dass man als Zielgröße nicht die Anzahl selbst, sondern $\sqrt{\text{Anzahl}}$ verwendet. Für unabhängige Einzelereignisse ist die Standardabweichung dieser transformierten Zielgröße unabhängig vom Mittelwert.

Aber auch wenn die Einzelereignisse nicht völlig unabhängig sind (wie z.B. die Lötfehler auf einer Leiterplatte, wo ein verrutschtes Bauelement mehrere Lötfehler verursachen kann), führt die Wurzeltransformation zu einer Verbesserung. Die Standardabweichung von $\sqrt{\text{Anzahl}}$ hängt dann zwar immer noch vom Mittelwert ab, aber weniger als bei der Anzahl selbst.



Beispiel

In einer Fertigungslinie werden zwei Lötanlagen parallel eingesetzt. Diese beiden Anlagen sollen bezüglich der Anzahl Lötfehler miteinander verglichen werden.¹⁶ Da die Anzahl Lötfehler insgesamt sehr klein ist, werden als Einheit jeweils 100 Leiterplatten betrachtet. Die folgenden Zahlen stellen die Gesamtzahl Lötfehler von jeweils 100 Leiterplatten gleichen Typs dar, die auf den Anlagen jeweils an einem Tag gelötet wurden.

| | | | | | | | | |
|------------------|----|---|----|----|----|----|----|---|
| Anzahl Anlage 1: | 10 | 3 | 8 | 1 | 18 | 4 | 6 | 9 |
| Anzahl Anlage 2: | 19 | 9 | 25 | 33 | 15 | 26 | 50 | 7 |

Diese Zahlen streuen stärker, als bei poissonverteilten Fehlerzahlen zu erwarten wäre. Es gibt offensichtlich gewisse Unterschiede im Anlagenzustand von Tag zu Tag oder eine Abhängigkeit zwischen den Fehlern. Die Transformation besteht darin, statt der Anzahl Fehler die Wurzel aus der Anzahl zu verwenden:

$$\begin{array}{llllllll} \sqrt{\text{Anzahl1}} : & 3,2 & 1,7 & 2,8 & 1,0 & 4,2 & 2,0 & 2,4 & 3,0 \\ \sqrt{\text{Anzahl2}} : & 4,4 & 3,0 & 5,0 & 5,7 & 3,9 & 5,1 & 7,1 & 2,6 \end{array}$$

Die weitere Analyse erfolgt wie in (6.12) bis (6.16), aber mit diesen Wurzeln:

$$\sqrt{\text{Anzahl1}} : \quad \bar{y}_1 = 2,54 \quad s_1 = 0,99$$

$$\sqrt{\text{Anzahl2}} : \quad \bar{y}_2 = 4,60 \quad s_2 = 1,46$$

$$\bar{d} = \bar{y}_2 - \bar{y}_1 = 4,60 - 2,54 = 2,06$$

$$s = \sqrt{\frac{s_1^2 + s_2^2}{2}} = \sqrt{\frac{0,99^2 + 1,46^2}{2}} = 1,25$$

$$s_d = \sqrt{\frac{4}{N}} \cdot s = \sqrt{\frac{4}{16}} \cdot 1,25 = 0,625$$

$$f = N - 2 = 16 - 2 = 14$$

¹⁶ Für den hier behandelten Vergleich von zwei Gruppen gibt es auch Verfahren, die direkt auf der Poissonverteilung basieren (vgl. z.B. [5]). Vorteil der hier behandelten Transformation ist, dass sie es erlaubt, Zählergebnisse mit demselben Formalismus wie Messwerte zu behandeln. Außerdem ist sie als Näherung auch dann noch einsetzbar, wenn die Fehler nicht unabhängig sind.

Mit Tabelle 6.4 erhält man (mit Interpolation):

$$95\%: t \cdot s_d = 2,14 \cdot 0,625 = 1,33$$

$$99\%: t \cdot s_d = 2,98 \cdot 0,625 = 1,86$$

$$99,9\%: t \cdot s_d = 4,14 \cdot 0,625 = 2,59$$

$$\Rightarrow 1,86 < |\bar{d}| = 2,06 < 2,59 \Rightarrow **$$

Es besteht ein signifikanter Unterschied zwischen den Anlagen.

6.4.3 Box-Cox-Transformation

Wenn die Standardabweichung von Messwerten y mit ihrem Mittelwert μ zunimmt wie

$$\sigma_y \propto \mu^{1-\lambda},$$

dann hat die transformierte Variable

$$Y = y^\lambda \quad (6.19)$$

eine Standardabweichung, die unabhängig vom Mittelwert ist (man spricht daher auch von einer varianzstabilisierenden Transformation) [7].

Für $\lambda = 0$ erhält man (als Grenzwert) die logarithmische Transformation aus Absatz 6.4.1, für $\lambda = 1/2$ die Wurzeltransformation aus Absatz 6.4.2. Diese Transformationen können als Sonderfälle der allgemeineren Transformation (6.19) betrachtet werden. Box und Cox haben ein Kriterium zur Auswahl des „besten“ Werts von λ vorgeschlagen. Durch eine geeignete Normierung werden Residuen für verschiedene Werte von λ vergleichbar gemacht und dann minimiert [7] (die sogenannte Box-Cox-Transformation).

Als Entscheidungshilfe bietet Statistik-Software eine Darstellung der normierten Residuen gegen λ , etwa wie in Bild 6-22. Das Minimum gibt den besten Wert von λ , hier sehr nahe an $\lambda = 0$. Als zusätzliche Hilfe wird der 95 %-Vertrauensbereich für λ gezeigt. Enthält dieser Vertrauensbereich nicht den Wert $\lambda = 1$ (keine Transformation), so sollten die Daten mit einem geeigneten glatten Wert aus dem Vertrauensbereich transformiert werden.

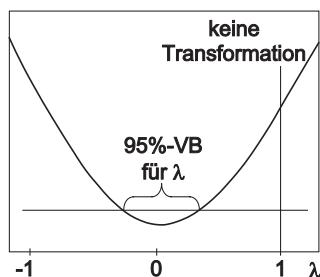


BILD 6-22

Box-Cox-Transformation am Beispiel der Daten aus Abschnitt 6.4.1:

$\lambda = 0$ entspricht der logarithmischen Transformation, sie ist signifikant besser als keine Transformation ($\lambda = 1$)



Hinweis

Die Notwendigkeit von Transformationen wurde hier statistisch begründet. Eine Transformation kann jedoch durchaus auch technisch sinnvoll sein. So kann man als Maß für den Durchsatz einer Anlage z.B. die Zeit pro Teil verwenden, aber genauso gut die Anzahl Teile pro Zeiteinheit. Mit $\lambda = -1$ in (6.19) werden diese beiden möglichen Zielgrößen ineinander umgerechnet ($1/y$). Welche der beiden Zielgrößen besser ist, ist nicht offensichtlich.

Literatur

- [1] Fahrmeir, L./Heumann, C./Künstler, R./Pigeot, I./Tutz, G.: Statistik – Der Weg zur Datenanalyse. Springer Verlag, Berlin, 8. Auflage 2016
- [2] Stahel, W.: Statistische Datenanalyse. Vieweg Verlag, Wiesbaden, 5. Auflage 2008
- [3] Ross, S. M.: Statistik für Ingenieure und Naturwissenschaftler, Spektrum Verlag, Heidelberg, 3. Auflage 2006
- [4] Köhler, W./Schachtel, G./Voleske, P.: Biostatistik, Springer Spektrum, Berlin, 5. Auflage 2012
- [5] Graf, U./Henning, H.J./Stange, K./Wilrich, P.-T.: Formeln und Tabellen der angewandten mathematischen Statistik. Springer Verlag, Berlin, 3. Auflage 1998
- [6] NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>
- [7] Box, G. E. P./Hunter, W. G./Hunter, J. S.: Statistics for Experimenters. John Wiley, New York, 2. Auflage 2005

7

Vollständige faktorielle Versuchspläne

Grundlage der Versuchsplanung

Dieses Kapitel behandelt Aufbau und Auswertung von vollständigen faktoriellen Versuchsplänen mit k Faktoren auf je zwei Stufen.

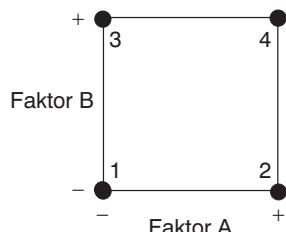
Beim Mittelwertvergleich in Kapitel 6 wurde ein Faktor auf zwei Stufen betrachtet. Vollständige faktorielle Versuchspläne sind eine Verallgemeinerung des Mittelwertvergleichs. Sie sind ein wichtiges „Arbeitspferd“ der Versuchsplanung und bilden die Grundlage für viele Weiterentwicklungen [1 – 5].

■ 7.1 Zwei Faktoren auf je zwei Stufen

In diesem Abschnitt wird der Mittelwertvergleich zunächst auf zwei Faktoren erweitert. Dies ist der Beginn der eigentlichen Versuchsplanung.

7.1.1 Versuchsplan und Effekte

Bild 7-1 zeigt einen vollständigen faktoriellen Versuchsplan mit zwei Faktoren A und B auf je zwei Stufen, die mit – und + bezeichnet werden (z. B. den Faktor Druck auf den Stufen 450 mTorr und 600 mTorr und den Faktor Temperatur auf den Stufen 710 °C und 720 °C). Der Versuchsplan besteht aus $2^2 = 4$ Faktorstufenkombinationen, daher heißt er auch 2^2 -Plan.



| syst. Nr. | Faktoren | |
|--------------|----------|---|
| | A | B |
| 1 | - | - |
| 2 | + | - |
| 3 | - | + |
| 4 | + | + |

BILD 7-1 Ein vollständiger faktorieller 2^2 -Versuchsplan besteht aus den vier Faktorstufenkombinationen der zwei Faktoren A und B auf den Stufen – und +. Die Nummern 1–4 im Bild links entsprechen den Nummern in der Tabelle rechts (systematische Reihenfolge).

Bezeichnet man die Ergebnisse für die Zielgröße y bei den vier Faktorstufenkombinationen mit y_1, y_2, y_3 und y_4 , so ist z. B.

$$\frac{y_2 + y_4}{2} \text{ der Mittelwert der Versuchsergebnisse mit A auf der Stufe +}$$

$$\frac{y_1 + y_3}{2} \text{ der Mittelwert der Versuchsergebnisse mit A auf der Stufe -.}$$

Die Differenz der beiden Mittelwerte ist ein Maß für die Wirkung des Faktors A. Man nennt sie Effekt des Faktors A.

$$\text{Effekt A} = \frac{y_2 + y_4}{2} - \frac{y_1 + y_3}{2} \quad (7.1)$$

Analog ist der Effekt von B ein Maß für die Wirkung des Faktors B.

$$\text{Effekt B} = \frac{y_3 + y_4}{2} - \frac{y_1 + y_2}{2} \quad (7.2)$$

Von einer Wechselwirkung AB spricht man, wenn der Effekt von A auf die Zielgröße davon abhängt, welchen Wert B hat (oder umgekehrt).

$y_4 - y_3$ ist der Effekt von A, wenn B auf der Stufe + steht.

$y_2 - y_1$ ist der Effekt von A, wenn B auf der Stufe - steht.

Die Differenz dieser beiden Effekte ist ein Maß dafür, wie stark die Wirkung des Faktors A vom Wert von B abhängt. Die Hälfte der Differenz nennt man Effekt der Wechselwirkung AB.

$$\text{Effekt AB} = \frac{y_4 - y_3}{2} - \frac{y_2 - y_1}{2} = \frac{y_1 - y_2 - y_3 + y_4}{2} \quad (7.3)$$

Zum selben Ergebnis gelangt man, wenn man berechnet, wie stark die Wirkung des Faktors B vom Wert von A abhängt.

Aus den Ergebnissen bei den vier Faktorstufenkombinationen werden somit drei Effekte berechnet. Als vierte Größe kann der Mittelwert aller Ergebnisse berechnet werden. Er beschreibt die Lage der Versuchsergebnisse insgesamt, macht aber keine Aussage über die Wirkung der Faktoren.



Beispiel

Durch thermische Zersetzung in einem Rohrofen wird SiO_2 auf Si-Scheiben abgeschieden. Der Druck und die Temperatur im Ofen werden als Faktoren verändert, Zielgröße ist die Abscheiderate. Für den Faktor Druck werden als Stufen 450 mTorr und 600 mTorr, für den Faktor Temperatur 710 °C und 720 °C verwendet. Tabelle 7.1 zeigt die Versuchsergebnisse für die Abscheiderate (Mittelwerte von mehreren Realisierungen).

| syst. Nr. | Druck [mTorr] | Temperatur [°C] | Abscheiderate [nm/min] |
|-----------|---------------|-----------------|------------------------|
| 1 | 450 (-) | 710 (-) | 6,0 |
| 2 | 600 (+) | 710 (-) | 7,5 |
| 3 | 450 (-) | 720 (+) | 6,6 |
| 4 | 600 (+) | 720 (+) | 10,3 |

TABELLE 7.1
Versuchsplan und Ver-
suchsergebnisse

Bild 7-2 zeigt links die Abscheiderate für die vier Faktorstufenkombinationen. Der Mittelwert der Versuchsergebnisse beträgt

$$\text{bei } 600 \text{ mTorr: } (7,5 + 10,3)/2 = 8,9 \text{ nm/min,}$$

$$\text{bei } 450 \text{ mTorr: } (6,0 + 6,6)/2 = 6,3 \text{ nm/min.}$$

Die Differenz der beiden Mittelwerte beträgt $8,9 - 6,3 = 2,6 \text{ nm/min}$. Dies ist der Effekt des Drucks. Bild 7-2 zeigt rechts die beiden Mittelwerte und den Effekt des Drucks. Er ist ein Maß für die Wirkung des Drucks auf die Abscheiderate.

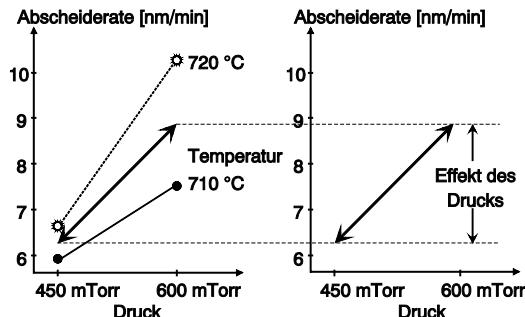


BILD 7-2

Grafische Darstellung der Abscheiderate für jede Faktorstufenkombination (links). Effekt des Drucks = 2,6 bedeutet:
Der Mittelwert der Abscheiderate beim Druck 600 mTorr ist um 2,6 nm/min höher als beim Druck 450 mTorr.

Bild 7-3 zeigt links die Abscheiderate für die vier Faktorstufenkombinationen wie in Bild 7-2. Der Mittelwert der Versuchsergebnisse beträgt

$$\text{bei } 720^\circ\text{C: } (6,6 + 10,3)/2 = 8,45 \text{ nm/min,}$$

$$\text{bei } 710^\circ\text{C: } (6,0 + 7,5)/2 = 6,75 \text{ nm/min.}$$

Die Differenz der beiden Mittelwerte beträgt $8,45 - 6,75 = 1,7 \text{ nm/min}$. Dies ist der Effekt der Temperatur. Bild 7-3 zeigt rechts die beiden Mittelwerte und den Effekt der Temperatur.

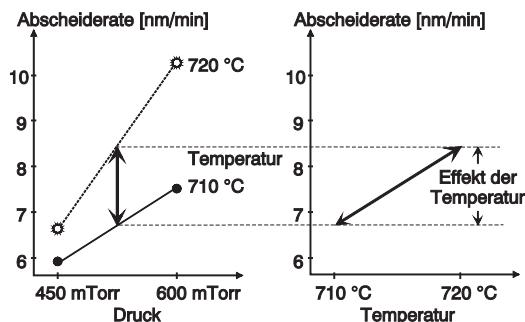


BILD 7-3

Grafische Darstellung der Abscheiderate, vgl. Bild 7-2. Effekt der Temperatur = 1,7 bedeutet:
Der Mittelwert der Abscheiderate bei Temperatur 720 °C ist um 1,7 nm/min höher als bei Temperatur 710 °C.

Bild 7-4 zeigt links die Abscheiderate für die vier Faktorstufenkombinationen wie in Bild 7-2. Der Effekt der Temperatur beträgt

$$\text{bei } 600 \text{ mTorr: } 10,3 - 7,5 = 2,8 \text{ nm/min,}$$

$$\text{bei } 450 \text{ mTorr: } 6,6 - 6,0 = 0,6 \text{ nm/min.}$$

Die Differenz der beiden Effekte beträgt $2,8 - 0,6 = 2,2 \text{ nm/min}$. Der Effekt der Wechselwirkung ist die Hälfte dieser Differenz, also $1,1 \text{ nm/min}$. Die Wechselwirkung ist somit ein Maß für die Abweichung von der Parallelität der Linien.

Bild 7-4 zeigt rechts eine alternative Darstellung der Wechselwirkung: Auf der x-Achse ist nun die Temperatur aufgetragen. Der Effekt des Drucks beträgt

$$\text{bei } 720^\circ\text{C: } 10,3 - 6,6 = 3,7 \text{ nm/min,}$$

$$\text{bei } 710^\circ\text{C: } 7,5 - 6,0 = 1,5 \text{ nm/min.}$$

Die Differenz der beiden Effekte beträgt $3,7 - 1,5 = 2,2 \text{ nm/min}$, wie oben.

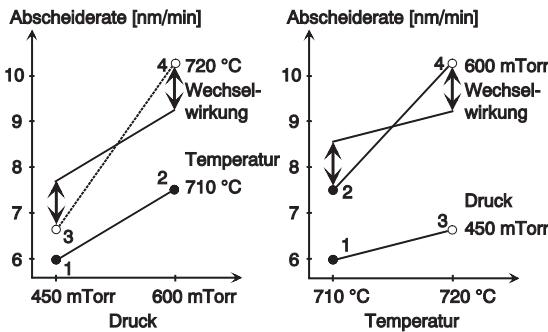


BILD 7-4

Grafische Darstellung der Abscheiderate, vgl. Bild 7-2. Der Effekt der Wechselwirkung zwischen Druck und Temperatur ist die Abweichung der Linien von der Parallelität (Pfeile links oder rechts).

7.1.2 Auswertereformalismus und Beurteilung der Signifikanz

Die Vorzeichen in den Gleichungen (7.1) und (7.2) zur Berechnung der Effekte von A und B sind identisch zu den Stufen der Faktoren A und B im Versuchsplan von Tabelle 7.1. Die Vorzeichen in Gleichung (7.3) zur Berechnung des Effekts der Wechselwirkung AB sind die Produkte der Vorzeichen von A und B. Aus dieser Beobachtung erhält man einen einfachen Auswertereformalismus (siehe Tabelle 7.2).

Die Signifikanz der Effekte wird wie in Abschnitt 6.3 beurteilt:

Wenn jede Faktorstufenkombination n-mal realisiert wurde und die Standardabweichung für alle Faktorstufenkombinationen gleich ist ($\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = \sigma$), dann ist analog zu (6.15)

$$s^2 = \frac{s_1^2 + s_2^2 + s_3^2 + s_4^2}{4} \quad (7.4)$$

ein Schätzwert für die Varianz der Einzelwerte mit Freiheitsgrad

$$f = 4 \cdot (n - 1) = N - 4 \quad (7.5)$$

Jeder Effekt ist die Differenz zweier Mittelwerte von jeweils N/2 Einzelwerten. Daher erhält man wie in (6.14) als Schätzwert für die Standardabweichung eines Effekts

$$s_{\bar{d}} = \sqrt{\frac{4}{N} \cdot s^2} \quad (7.6)$$

Ein Vergleich von (7.5) und (7.6) mit (6.14) und (6.16) zeigt eine große Ähnlichkeit, wenn man den Versuchsumfang N als Vergleichsbasis verwendet. Daher ist in der Versuchsplanung der Versuchsumfang N nützlicher als die Anzahl n der Realisierungen jeder Faktorstufenkombination (der Stichprobenumfang n). Im Folgenden wird in Formeln daher immer der Versuchsumfang N verwendet.

Um die Signifikanz der Effekte zu beurteilen, vergleicht man sie wie in Bild 6-17 mit der Breite des 95%--, 99%- und 99,9%-Vertrauensbereichs, d. h. man vergleicht jeden Effekt mit $\pm t \cdot s_{\bar{d}}$.



Beispiel (Fortsetzung)

Die Versuchsergebnisse in Tabelle 7.1 sind Mittelwerte von jeweils n = 4 Realisierungen, die in randomisierter Reihenfolge durchgeführt wurden (vgl. Tabelle 7.3). Tabelle 7.2 zeigt die Ergebnisse in systematischer Reihenfolge (syst. Nr. aus Bild 7-1).

TABELLE 7.2 Formalismus zur Auswertung von Versuchsergebnissen:

Spalten A und B: Versuchsplan gemäß Bild 7-1 für die Faktoren A und B

Spalte AB: Produkt der Vorzeichen in Spalten A und B

Einzelwerte: Ergebnisse der Einzelversuche

\bar{y}_i : Mittelwert der Einzelwerte einer Faktorstufenkombination (Zeile)

s_i^2 : Varianz der Einzelwerte einer Faktorstufenkombination (Zeile)

Zeile „Σ“ (Summe):

Spalten A, B, AB: Summe der Spalte \bar{y}_i mit Vorzeichen der jeweiligen Spalte
(z. B. in Spalte A: 5,2 = - 6,0 + 7,5 - 6,6 + 10,3)

Spalte s_i^2 : Summe der Varianzen (2,38 = 0,247 + 1,333 + 0,500 + 0,300)

Zeile „Effekt“: Summe/(Anzahl der Wertepaare) (hier 2; z. B. 2,6 = 5,2/2)

s^2 : (Summe der Varianzen)/(Anzahl der Werte) (0,595 = 2,38/4)

| syst. Nr. | Druck A | Temperatur B | WW AB | Einzelwerte Abscheiderate [nm/min] | | | | \bar{y}_i | s_i^2 |
|--------------|------------|-----------------|----------|---------------------------------------|------|------|------|-------------|---------|
| 1 | - | - | + | 6,1 | 5,9 | 5,4 | 6,6 | 6,0 | 0,247 |
| 2 | + | - | - | 6,1 | 7,7 | 8,9 | 7,3 | 7,5 | 1,333 |
| 3 | - | + | - | 5,8 | 6,4 | 7,5 | 6,7 | 6,6 | 0,500 |
| 4 | + | + | + | 9,7 | 11,0 | 10,4 | 10,1 | 10,3 | 0,300 |
| Σ | 5,2 | 3,4 | 2,2 | | | | | | 2,38 |
| Effekt | 2,6 | 1,7 | 1,1 | | | | | $s^2 =$ | 0,595 |

Beurteilung der Signifikanz der Effekte:

$$s^2 = \frac{0,247 + 1,333 + 0,500 + 0,300}{4} = 0,595 = \text{Mittelwert der Einzelvarianzen}$$

$$s_d = \sqrt{\frac{4}{16} \cdot 0,595} = 0,3857 = \text{Standardabweichung der Effekte}$$

$$f = 16 - 4 = 12 = \text{Freiheitsgrad.}$$

Mit Tabelle 6.4 erhält man:

$$95\%: t \cdot s_d = 2,179 \cdot 0,3857 = 0,840$$

$$99\%: t \cdot s_d = 3,055 \cdot 0,3857 = 1,178$$

$$99,9\%: t \cdot s_d = 4,318 \cdot 0,3857 = 1,665$$

Der Vergleich der Breiten der Vertrauensbereiche mit den Effekten zeigt:

$$\text{Effekt des Drucks} = 2,6 \quad ***$$

$$\text{Effekt der Temperatur} = 1,7 \quad ***$$

$$\text{Effekt der Wechselwirkung} = 1,1 \quad *.$$

Analog zu Bild 6-17 zeigt Bild 7-5 den Vergleich der Effekte mit der Breite der Vertrauensbereiche grafisch. *** bedeutet, dass der Effekt aus dem durchgezogenen 99,9%-Niveau herausragt. * bedeutet, dass der Effekt zwischen dem gestrichelten 99% und dem punktierten 95%-Niveau liegt.

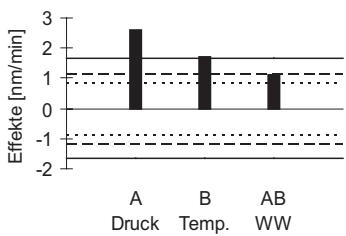


BILD 7-5

Grafische Darstellung des Signifikanztests:

Die Länge der Balken stellt die Effekte dar, die waagrechten Linien die Breite des 95%- 99%- und 99,9%-Vertrauensbereichs

7.1.3 Interpretation von Wechselwirkungen

Eine Wechselwirkung zwischen zwei Faktoren (Zwei-Faktor-Wechselwirkung, kurz 2FWW genannt) bedeutet, dass der Effekt eines Faktors davon abhängt, welchen Wert der andere Faktor hat. Erfahrungsgemäß ist es schwierig, sich eine Wechselwirkung als eigenständige Größe vorzustellen. Daher wird empfohlen, bei der Interpretation von Versuchsergebnissen wie folgt vorzugehen:

- Die Effekte aller Faktoren und Wechselwirkungen werden formal berechnet und auf Signifikanz überprüft.
- Sind nur die Effekte der Faktoren selbst signifikant, so können die Faktoren getrennt betrachtet werden. Zur Auswahl der günstigeren Stufe für jeden (signifikanten) Faktor vergleicht man wie in Bild 7-2 oder 7-3 rechts die Mittelwerte der Versuchsergebnisse von Faktorstufe + und - miteinander (ihre Differenz ist der Effekt).
- Ist der Effekt der Wechselwirkung zwischen zwei Faktoren signifikant, so müssen diese beiden Faktoren immer gemeinsam betrachtet werden. Man berechnet die vier Mittelwerte für die verschiedenen Faktorstufenkombinationen getrennt und wählt die güns-

tigste Kombination aus. Was günstig ist, hängt von der Zielsetzung ab. Mit dem Ergebnis von Bild 7-4 kann man z.B. zu folgenden Ergebnissen gelangen:

- wenn das Ziel ein hoher Wert der Abscheiderate ist, ist die Kombination Temperatur 720 °C und Druck 600 mTorr günstig,
- wenn das Ziel eine geringe Abhängigkeit von der Temperatur ist, ist der Druck 450 mTorr günstig,
- wenn das Ziel eine geringe Abhängigkeit vom Druck ist, ist die Temperatur 710 °C günstig.

Wechselwirkungen sind weit verbreitet. Es ist normal, dass der Effekt eines Faktors davon abhängt, welchen Wert ein anderer Faktor hat. Einige typische Situationen, in denen Wechselwirkungen wichtig sind, sollen dies verdeutlichen:

- Die Lage des Optimums für einen Faktor hängt vom Wert des anderen Faktors ab (a).
- Wenn zwei ungünstige Faktorstufen zusammentreffen, treten Probleme auf (b).
- Zwei Faktoren wirken sich multiplikativ auf die Zielgröße aus (c).



Beispiele

1. In einer Fertigung werden Teile von zwei verschiedenen Lieferanten A und B auf zwei parallelen Anlagen 1 und 2 verarbeitet. Die Teile von Lieferant A lassen sich besser auf Anlage 1, die Teile von Lieferant B besser auf Anlage 2 verarbeiten (a).
2. Bei der Optimierung der Einstellung einer Belichtungsanlage erhält man gute Ergebnisse bei großer Blende und kurzer Belichtungszeit und bei kleiner Blende und großer Belichtungszeit (d.h. bei mittlerer Lichtmenge) (a).
3. In einer Fertigung werden zugelieferte Teile von zwei verschiedenen Lieferanten A und B auf zwei parallelen Anlagen 1 und 2 weiterverarbeitet. Teile von Lieferant A lassen sich nur an Anlage 1 gut verarbeiten, bei Lieferant B ist das Ergebnis unabhängig von der Anlage (b).
4. In einer Fertigung werden zwei verschiedene Lotpasten eingesetzt. Eine dieser Lotpasten macht Probleme, wenn die Luftfeuchtigkeit hoch ist (b).
5. Beim Setzen von Kleberpunkten mit einer Spritze führt eine bestimmte Vergrößerung des Kanülen Durchmessers zu einer Verdoppelung des Punktvolumens, und eine bestimmte Verlängerung der Zeit führt ebenfalls zu einer Verdoppelung. Dann erwartet man bei großem Durchmesser und langer Zeit eine Vervierfachung. Bei kleiner Kanüle führt die Verlängerung zu einer Vergrößerung des Volumens von 1 auf 2, also um 1. Bei großer Kanüle führt die Verlängerung zu einer Vergrößerung des Volumens von 2 auf 4, also um 2 (c).

Manchmal kann man durch eine geeignete Wahl der Faktoren und/oder Zielgrößen vorhersehbare Wechselwirkungen vermeiden. Im obigen Belichtungsbeispiel ist es günstiger, die Lichtmenge als Faktor zu verwenden, im Beispiel Kleberpunkte vermeidet man die Wechselwirkung, wenn man den Logarithmus des Volumens als Zielgröße verwendet. Man kann jedoch nicht davon ausgehen, dass dies immer gelingt. Daher ist es sicherer, Wechselwirkungen mit zu berücksichtigen.

7.1.4 Randomisierung und Blockbildung

In Absatz 3.4.3 und 3.4.4 wurde beschrieben, wie durch Blockbildung und Randomisierung innerhalb der Blöcke eine Verfälschung der Versuchsergebnisse durch einen eventuell vorhandenen Trend oder andere systematische Unterschiede weitestgehend vermieden werden kann. Hier soll die Umsetzung von Blockbildung und Randomisierung am Beispiel von Absatz 7.1.1 gezeigt werden.

TABELLE 7.3 Vollständiger faktorieller 2^2 -Versuch mit 4 Realisierungen mit Blockbildung und randomisierter Reihenfolge innerhalb der Blöcke:

Jede Zeile stellt einen Einzelversuch dar.

1. Spalte: Reihenfolge, in der Einzelversuche durchgeführt werden
2. Spalte: Nummer der Faktorstufenkombination in Tabelle 7.1 bzw. 7.2
3. Spalte: Nummer der Realisierung (d. h. Block)
4. und 5. Spalte: Stufen der Faktoren bei jedem Einzelversuch
6. Spalte: Versuchsergebnisse aus Tabelle 7.2
7. Spalte: Postulierte Verfälschung durch einen Trend in den Ergebnissen
8. Spalte: Versuchsergebnisse mit Trend (Summe von 6. und 7. Spalte)

| Vers. Nr. | syst. Nr. | Realisierung (Block) | Faktor A | Faktor B | Rate ohne Trend | Trend | Rate mit Trend |
|-----------|-----------|----------------------|----------|----------|-----------------|-------|----------------|
| 1 | 2 | 1 | + | - | 6,1 | 0,1 | 6,2 |
| 2 | 3 | 1 | - | + | 5,8 | 0,2 | 6,0 |
| 3 | 1 | 1 | - | - | 6,1 | 0,3 | 6,4 |
| 4 | 4 | 1 | + | + | 9,7 | 0,4 | 10,1 |
| 5 | 4 | 2 | + | + | 11,0 | 0,5 | 11,5 |
| 6 | 2 | 2 | + | - | 7,7 | 0,6 | 8,3 |
| 7 | 3 | 2 | - | + | 6,4 | 0,7 | 7,1 |
| 8 | 1 | 2 | - | - | 5,9 | 0,8 | 6,7 |
| 9 | 1 | 3 | - | - | 5,4 | 0,9 | 6,3 |
| 10 | 3 | 3 | - | + | 7,5 | 1,0 | 8,5 |
| 11 | 2 | 3 | + | - | 8,9 | 1,1 | 10,0 |
| 12 | 4 | 3 | + | + | 10,4 | 1,2 | 11,6 |
| 13 | 3 | 4 | - | + | 6,7 | 1,3 | 8,0 |
| 14 | 1 | 4 | - | - | 6,6 | 1,4 | 8,0 |
| 15 | 4 | 4 | + | + | 10,1 | 1,5 | 11,6 |
| 16 | 2 | 4 | + | - | 7,3 | 1,6 | 8,9 |

Die vier Faktorstufenkombinationen einer Realisierung bilden jeweils einen Block. Sie werden nacheinander durchgeführt. Randomisierung bedeutet, dass die Reihenfolge der vier Einzelversuche in jeder Realisierung zufällig (und unterschiedlich) ist. So ergibt sich z. B. die in Tabelle 7.3 dargestellte Reihenfolge der Einzelversuche. Um die Wirkung von Blockbildung und Randomisierung zu demonstrieren, wird in Tabelle 7.3 ein Trend unterstellt, der dazu führt, dass sich das Ergebnis von Einzelversuch zu Einzelversuch jeweils um 0,1 erhöht.

TABELLE 7.4 Auswertung der Versuchsergebnisse mit Trend aus Tabelle 7.3 (Bezeichnungen wie in Tabelle 7.2, ohne Korrektur für den Trend)

| syst. Nr. | A | B | AB | Einzelwerte Rate [nm/min] | | | | \bar{y}_i | s_i^2 |
|-----------|------|-----|------|---------------------------|------|------|------|-------------|---------|
| 1 | - | - | + | 6,4 | 6,7 | 6,3 | 8,0 | 6,85 | 0,617 |
| 2 | + | - | - | 6,2 | 8,3 | 10,0 | 8,9 | 8,35 | 2,550 |
| 3 | - | + | - | 6,0 | 7,1 | 8,5 | 8,0 | 7,40 | 1,207 |
| 4 | + | + | + | 10,1 | 11,5 | 11,6 | 11,6 | 11,2 | 0,540 |
| Σ | 5,3 | 3,4 | 2,3 | | | | | | 4,914 |
| Effekt | 2,65 | 1,7 | 1,15 | | | | | $s^2 =$ | 1,228 |

Tabelle 7.4 zeigt die Ergebnisse aus Tabelle 7.3 in der systematischen Reihenfolge (wie in Tabelle 7.2). Die Auswertung in Tabelle 7.4 nutzt nur die Randomisierung aus. Die Blockstruktur ist nicht berücksichtigt.

Durch Randomisierung wurde vermieden, dass der Trend die Schätzwerte für die Effekte mehr als zufällig verfälscht. Allerdings hat sich der Schätzwert für die Varianz erhöht, weil die Einzelwerte in Tabelle 7.4 aufgrund des Trends von Spalte zu Spalte zunehmen. Dadurch verändert sich die Signifikanz der Effekte.

TABELLE 7.5 Auswertung der Versuchsergebnisse mit Trend aus Tabelle 7.3 (mit Korrektur für den Trend – von jedem Block [= jeder Spalte] wurde der jeweilige Mittelwert abgezogen)

| syst. Nr. | A | B | AB | korrigierte Einzelwerte | | | | \bar{y}_i | s_i^2 |
|-----------|------|-----|------|-------------------------|------|------|--------|-------------|---------|
| 1 | - | - | + | -0,775 | -1,7 | -2,8 | -1,125 | -1,6 | 0,785 |
| 2 | + | - | - | -0,975 | -0,1 | 0,9 | -0,225 | -0,1 | 0,594 |
| 3 | - | + | - | -1,175 | -1,3 | -0,6 | -1,125 | -1,05 | 0,095 |
| 4 | + | + | + | 2,925 | 3,1 | 2,5 | 2,475 | 2,75 | 0,097 |
| Σ | 5,3 | 3,4 | 2,3 | | | | | | 1,571 |
| Effekt | 2,65 | 1,7 | 1,15 | | | | | $s^2 =$ | 0,393 |

In Tabelle 7.5 wird zusätzlich die Blockbildung ausgenutzt. In jedem Block (= Spalte der Einzelwerte) tritt jede Faktorstufenkombination genau einmal auf. Daher sollten sich die Mittelwerte der Blöcke nur zufällig unterscheiden. Von den Versuchsergebnissen für jeden Block wird daher der Mittelwert dieses Blocks abgezogen, d.h. in Tabelle 7.5 sind von den Spalten der Einzelwerte in Tabelle 7.4 jeweils die Mittelwerte der Spalten (7,175; 8,40; 9,10 und 9,125) abgezogen.

Das Abziehen der Mittelwerte hat keinen Einfluss auf die Effekte, man erhält jedoch eine deutlich reduzierte Streuung, weil Unterschiede zwischen den Blöcken nicht mehr eingehen.

Durch das Abziehen der Blockmittelwerte hat man jedoch nicht nur die systematischen Unterschiede zwischen den Blöcken eliminiert, sondern auch die zufälligen Unterschiede (in Tabelle 7.2 sind die Mittelwerte der Spalten auch nicht gleich, obwohl es keinen Trend gibt). Dadurch unterschätzt man die Streuung jetzt, und eine Korrektur ist erforderlich. Bei b Blöcken reduziert sich durch das Abziehen der b Blockmittelwerte der Freiheits-

grad f um $b - 1$ (der Gesamtmittelwert geht auch ohne Berücksichtigung der Blöcke nicht in die Rechnung ein) und man erhält statt (7.4) und (7.5):¹

$$f_{\text{korr}} = f - b + 1 \quad (7.7)$$

$$s_{\text{korr}}^2 = \frac{f}{f_{\text{korr}}} \cdot s^2 \quad (7.8)$$

Für Tabelle 7.5 erhält man ($b = 4$ Blöcke):

$$f_{\text{korr}} = f - 4 + 1 = 12 - 3 = 9$$

$$s_{\text{korr}}^2 = \frac{12}{9} \cdot 0,393 = 0,524$$

Diese korrigierte Varianz unterscheidet sich nur zufällig vom Ergebnis ohne Trend im Anschluss an Tabelle 7.2. Durch Randomisierung konnte vermieden werden, dass der Trend die Effekte verfälscht. Aufgrund der Blockbildung konnte der Unterschied zwischen den Blöcken aus der Zufallsstreuung herausgerechnet werden (vgl. Tabelle 7.5 mit 7.4).

Tabelle 7.3 stellt somit eine ideale Reihenfolge für die Durchführung der Versuche dar. Für diese Reihenfolge müssen die Faktorstufen von Einzelversuch zu Einzelversuch geändert werden. Dies kann zu einem hohen Versuchsaufwand führen.

In der Praxis wird daher häufig der Wunsch geäußert, alle Realisierungen derselben Faktorstufenkombination (mit derselben systematischen Nummer) hintereinander durchzuführen. Dies sollte nur gemacht werden, wenn die Einstellung aller Faktorstufen genau reproduzierbar ist und wenn keine systematische Veränderung der Ergebnisse mit der Zeit (d.h. kein Trend) zu befürchten ist.

Ist die Änderung eines der Faktoren sehr aufwändig, so wird häufig der Wunsch geäußert, zunächst alle Versuche mit einer Stufe dieses Faktors durchzuführen und erst dann alle Versuche mit der anderen Stufe. Dies sollte nur im Extremfall geschehen, da dann ein Trend den Schätzwert für den Effekt dieses Faktors verfälschen kann (vgl. dazu auch Absatz 3.4.5).



Aufgabe

Die Ausbeute einer chemischen Reaktion soll erhöht werden. In einem Brainstorming wurden als vermutlich wichtigste Faktoren die Temperatur und der Druck festgelegt. Momentan sind bei der Anlage eine Temperatur von 100 °C und ein Druck von 2 bar eingestellt. Es soll untersucht werden, ob eine Erhöhung der Temperatur auf 120 °C und/oder eine Erhöhung des Drucks auf 3 bar zu einer Verbesserung der Ausbeute führt.

¹ Software führt diese Korrekturen durch, ohne dass sich der Anwender darum kümmern muss. Ziel dieser Darstellung ist vor allem, durch eine einfache Rechnung den Hintergrund zu erläutern. Software nutzt zur Darstellung der Ergebnisse häufig Bezeichnungen aus der Varianzanalyse (siehe Kapitel 12), für das Verständnis der Vorgehensweise ist Varianzanalyse jedoch nicht erforderlich.

- a) Stellen Sie einen Versuchsplan auf. In welcher Reihenfolge würden Sie die Versuche durchführen, wenn jede Faktorstufenkombination zweimal realisiert werden soll? Diskutieren Sie Vor- und Nachteile verschiedener Alternativen.
- b) Bei der Durchführung der Versuche wurden folgende Ausbeuten gemessen. Vervollständigen Sie die Tabelle, und berechnen Sie die Effekte von Temperatur, Druck und Wechselwirkung. Welche Effekte sind signifikant? Welche Faktorstufenkombination werden Sie in Zukunft verwenden?

| syst. Nr. | Temperatur | Druck | WW | Einzelausbeute [%] | | Mittel \bar{y} | Varianz s^2 |
|-----------|------------|-------|----|--------------------|------|------------------|---------------|
| 1 | - | - | | 70,3 | 69,2 | 69,75 | 0,605 |
| 2 | + | - | | 64,5 | 65,0 | 64,75 | 0,125 |
| 3 | - | + | | 58,0 | 59,9 | 58,95 | 1,805 |
| 4 | + | + | | 72,6 | 71,9 | | |

Lösung

c) Vollständiger faktorieller 2^2 -Versuchsplan in systematischer Reihenfolge:

| syst. Nr. | Temperatur | Druck |
|-----------|------------|-------|
| 1 | 100 °C | 2 bar |
| 2 | 120 °C | 2 bar |
| 3 | 100 °C | 3 bar |
| 4 | 120 °C | 3 bar |

Mögliche Reihenfolgen der Einzelversuche (Auswahl – die Beispiele sollen nur das Prinzip zeigen):

1. Randomisierte Reihenfolge der Einzelversuche mit $n = 2$

| Vers. Nr. | syst. Nr. | Block | Temperatur | Druck | Ausbeute |
|-----------|-----------|-------|------------|-------|----------|
| 1 | 2 | 1 | 120 °C | 2 bar | |
| 2 | 4 | 1 | 120 °C | 3 bar | |
| 3 | 1 | 1 | 100 °C | 2 bar | |
| 4 | 3 | 1 | 100 °C | 3 bar | |
| 5 | 4 | 2 | 120 °C | 3 bar | |
| 6 | 1 | 2 | 100 °C | 2 bar | |
| 7 | 3 | 2 | 100 °C | 3 bar | |
| 8 | 2 | 2 | 120 °C | 2 bar | |

Vorteile: Wegen Randomisierung keine Verfälschung der Effekte durch Trend.

Zufallsstreuung wird voll erfasst.

Unterschied zwischen Blöcken kann erkannt und eliminiert werden.

Nachteil: Häufige Änderung der Stufen, dadurch u.U. hoher Versuchsaufwand.

2. Randomisierte Reihenfolge der Faktorstufenkombinationen

| Vers. Nr. | syst. Nr. | Temperatur | Druck | Ausbeute |
|-----------|-----------|------------|-------|----------|
| 1 | 2 | 120 °C | 2 bar | |
| 2 | 2 | 120 °C | 2 bar | |
| 3 | 3 | 100 °C | 3 bar | |
| 4 | 3 | 100 °C | 3 bar | |
| 5 | 4 | 120 °C | 3 bar | |
| 6 | 4 | 120 °C | 3 bar | |
| 7 | 1 | 100 °C | 2 bar | |
| 8 | 1 | 100 °C | 2 bar | |

Vorteile: Wegen Randomisierung keine Verfälschung der Effekte durch Trend (zumindest bei Versuchen mit mehr Faktoren).

Geringerer Aufwand als bei Randomisierung der Einzelversuche.

Nachteile: Zufallsstreuung durch Ungenauigkeit der Einstellung der Faktorstufen wird nicht erfasst, Zufallsstreuung daher evtl. unterschätzt.

Trend kann nicht als Unterschied zwischen den Blöcken erkannt und aus Zufallsstreuung eliminiert werden.

3. Geordnete Reihenfolge der Faktorstufen für den Faktor Temperatur

(Annahme: Änderung der Temperatur ist sehr aufwändig)

| Vers. Nr. | syst. Nr. | Temperatur | Druck | Ausbeute |
|-----------|-----------|------------|-------|----------|
| 1 | 2 | 120 °C | 2 bar | |
| 2 | 2 | 120 °C | 2 bar | |
| 3 | 4 | 120 °C | 3 bar | |
| 4 | 4 | 120 °C | 3 bar | |
| 5 | 3 | 100 °C | 3 bar | |
| 6 | 3 | 100 °C | 3 bar | |
| 7 | 1 | 100 °C | 2 bar | |
| 8 | 1 | 100 °C | 2 bar | |

Vorteil: Geringer Aufwand – die Temperatur muss nur einmal verändert werden.

Nachteile: Zufallsstreuung durch Ungenauigkeit der Einstellung der Faktorstufen wird nicht erfasst, Zufallsstreuung daher evtl. unterschätzt.

Trend kann Effekt der Temperatur verfälschen.

Anmerkung:

Je nach Problemstellung sind auch andere Reihenfolgen denkbar. Vom statistischen Standpunkt aus betrachtet ist die 1. Reihenfolge ideal, da sie Blockbildung und Randomisierung berücksichtigt. Jede andere Reihenfolge birgt Risiken der Verfälschung, diese werden jedoch u.U. zur Reduzierung des Aufwandes bewusst in Kauf genommen.

b)

| syst. Nr. | Temperatur | Druck | WW | Einzelwerte [%] | \bar{y} | s^2 |
|-------------|------------|-------|------|-----------------|-----------|-------|
| 1 | - | - | + | 70,3 | 69,2 | 69,75 |
| 2 | + | - | - | 64,5 | 65,0 | 64,75 |
| 3 | - | + | - | 58,0 | 59,9 | 58,95 |
| 4 | + | + | + | 72,6 | 71,9 | 72,25 |
| Σ | 8,3 | -3,3 | 18,3 | | | 2,780 |
| Effekt | 4,15 | -1,65 | 9,15 | | | 0,695 |
| Signifikanz | ** | * | *** | | | |

$$s_{\bar{y}} = \sqrt{\frac{4}{N} \cdot s^2} = \sqrt{\frac{4}{8} \cdot 0,695} = 0,59$$

$$f = N - 4 = 8 - 4 = 4.$$

Mit Tabelle 6.4 erhält man:

$$95\%: t \cdot s_{\bar{y}} = 2,776 \cdot 0,59 = 1,64$$

$$99\%: t \cdot s_{\bar{y}} = 4,604 \cdot 0,59 = 2,71$$

$$99,9\%: t \cdot s_{\bar{y}} = 8,610 \cdot 0,59 = 5,08$$

Da der Effekt der Wechselwirkung AB (hoch-)signifikant ist, müssen die Faktoren A und B gemeinsam betrachtet werden. Die Mittelwerte für die Faktorstufenkombinationen aus obiger Tabelle zeigen, dass die Kombination Temperatur 120 °C mit Druck 3 bar die höchste Ausbeute liefert.

■ 7.2 k Faktoren auf je zwei Stufen

In diesem Abschnitt wird die Erweiterung auf (im Prinzip) beliebig viele Faktoren behandelt.

7.2.1 Versuchsplan

Erweitert man den Versuchsplan in Bild 7-1, so verdoppelt sich mit jedem neuen Faktor die Anzahl der Faktorstufenkombinationen. Für k Faktoren auf 2 Stufen erhält man

$$m = 2^k \text{ Faktorstufenkombinationen} \quad (7.9)$$

Tabelle 7.6 zeigt diese Faktorstufenkombinationen (in der allgemeinen Bezeichnung der Stufen mit + und -). Da alle möglichen Faktorstufenkombinationen enthalten sind, heißt dieser Plan vollständiger faktorieller 2^k -Versuchsplan (oder kurz: 2^k -Plan).

In Bild 7-6 sind die $2^3 = 8$ Faktorstufenkombinationen eines Versuchsplans für drei Faktoren grafisch dargestellt: Sie bilden die Eckpunkte eines Würfels. Mehr als drei Faktoren lassen sich nicht mehr einfach grafisch darstellen, die Kombinationen bilden jedoch immer die Eckpunkte eines k -dimensionalen Würfels.

TABELLE 7.6 Faktorstufenkombinationen vollständiger faktorieller 2^k -Pläne für k Faktoren auf 2 Stufen: Für jeden zusätzlichen Faktor verdoppelt sich die Anzahl der Kombinationen.

| syst. Nr. | Faktor A | Faktor B | Faktor C | Faktor D | ... |
|-----------|----------|----------|----------|----------|-----|
| 1 | - | - | - | - | - |
| 2 | + | - | - | - | - |
| 3 | - | + | - | - | - |
| 4 | + | + | - | - | - |
| 5 | - | - | + | - | - |
| 6 | + | - | + | - | - |
| 7 | - | + | + | - | - |
| 8 | + | + | + | - | - |
| 9 | - | - | - | + | - |
| 10 | + | - | - | + | - |
| 11 | - | + | - | + | - |
| 12 | + | + | - | + | - |
| 13 | - | - | + | + | - |
| 14 | + | - | + | + | - |
| 15 | - | + | + | + | - |
| 16 | + | + | + | + | - |

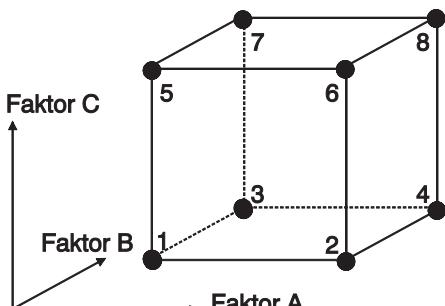


BILD 7-6

Grafische Darstellung der Faktorstufenkombinationen eines vollständigen faktoriellen 2^3 -Plans

Hinweis:

Ergänzende Versuche im Zentrum des Würfels sind sinnvoll und werden in Abschnitt 16.2 behandelt



Beispiel

Siehe Abschnitt 3.4: Das Beispiel aus der chemischen Industrie ist ein vollständiger faktorieller 2^3 -Plan

7.2.2 Auswertung

Die Auswertung verläuft analog zur Auswertung des 2^2 -Plans. Für einen 2^k -Plan kann man insgesamt $2^k - 1$ Effekte berechnen (und zusätzlich den Gesamtmittelwert der Daten). Tabelle 7.7 zeigt die $2^3 - 1 = 8 - 1 = 7$ Effekte und ihre Vorzeichenspalten im Beispiel 2^3 . Die Vorzeichenspalten für die drei Faktoren A, B und C ergeben sich direkt aus dem Versuchsplan. Die Vorzeichenspalten der Wechselwirkungen ergeben sich als Produkte der Vorzeichenspalten der beteiligten Faktoren, also

$$\text{z.B. } AB = A \cdot B \text{ und } ABC = A \cdot B \cdot C \quad (7.10)$$

Es gibt drei Zwei-Faktor-Wechselwirkungen (2FWW) AB, AC und BC und eine Drei-Faktor-Wechselwirkung (3FWW) ABC.

TABELLE 7.7 Vorzeichenspalten für die Berechnung der Effekte der drei Faktoren (A, B und C), der drei 2FWW (AB, AC und BC) und der einen 3FWW (ABC)

| syst. Nr. | A | B | C | AB | AC | BC | ABC |
|-----------|---|---|---|----|----|----|-----|
| 1 | - | - | - | + | + | + | - |
| 2 | + | - | - | - | - | + | + |
| 3 | - | + | - | - | + | - | + |
| 4 | + | + | - | + | - | - | - |
| 5 | - | - | + | + | - | - | + |
| 6 | + | - | + | - | + | - | - |
| 7 | - | + | + | - | - | + | - |
| 8 | + | + | + | + | + | + | + |

Die Effekte erhält man, indem man die (Mittelwerte der) Versuchsergebnisse mit den jeweiligen Vorzeichenspalten multipliziert und dann addiert. Das Ergebnis wird durch die Anzahl Paare dividiert. Bei 3 Faktoren (mit $m = 8$ Faktorstufenkombinationen) dividiert man also durch 4, allgemein durch $m/2$.

$$\text{Effekt} = \frac{2}{m} \cdot \sum_{i=1}^m (\text{Vorzeichen} \cdot \bar{y}_i) \quad (7.11)$$

Der Mittelwert der Varianzen innerhalb der Faktorstufenkombinationen ist ein Schätzwert für die Varianz.

$$s^2 = \frac{1}{m} \cdot \sum_{i=1}^m s_i^2 \quad (7.12)$$

Die Standardabweichung des Effekts ist

$$s_{\bar{d}} = \sqrt{\frac{4}{N} \cdot s^2} \quad (7.13)$$

Der Freiheitsgrad ist (Gesamtzahl der Einzelversuche) – (Anzahl Faktorstufenkombinationen), also

$$f = N - m \quad (7.14)$$

Um die Signifikanz der Effekte zu beurteilen, vergleicht man sie wie in Bild 6-17 bzw. 7-5 mit der Breite der 95%--, 99%- und 99,9%-Vertrauensbereiche, d. h. man vergleicht jeden Effekt mit $\pm t \cdot s_{\bar{d}}$.



Beispiel

Siehe Abschnitt 3.4 – Beispiel aus der chemischen Industrie.

Ergänzend zu Abschnitt 3.4 wird hier gezeigt, wie die Effekte mit (7.11) berechnet und ihre Signifikanz mit (7.12) bis (7.14) beurteilt werden können.

Faktoren und Stufen:

Faktor A Temperatur – = 120 °C + = 140 °C

Faktor B Zeit – = 2 h + = 4 h

Faktor C Katalysator – = 0,1% + = 0,5%

Durch Umcodierung der Stufen und Einfügen der Vorzeichenspalten der Wechselwirkungen erhält man aus Tabelle 3.5 die Tabelle 7.8.

TABELLE 7.8 Tabelle zur Berechnung der Effekte der Faktoren und Wechselwirkungen

| syst. Nr. | Temp. A | Zeit B | Kat. C | AB | AC | BC | ABC | Einzelergebnisse [%] | | \bar{y}_i | s_i^2 |
|-----------|---------|--------|--------|-------|--------|--------|--------|----------------------|------|-------------|---------|
| 1 | – | – | – | + | + | + | – | 52,8 | 54,1 | 53,45 | 0,845 |
| 2 | + | – | – | – | – | + | + | 61,5 | 61,8 | 61,65 | 0,045 |
| 3 | – | + | – | – | + | – | + | 56,7 | 55,2 | 55,95 | 1,125 |
| 4 | + | + | – | + | – | – | – | 67,9 | 70,2 | 69,05 | 2,645 |
| 5 | – | – | + | + | – | – | + | 53,6 | 54,1 | 53,85 | 0,125 |
| 6 | + | – | + | – | + | – | – | 62,2 | 62,9 | 62,55 | 0,245 |
| 7 | – | + | + | – | – | + | – | 56,5 | 54,6 | 55,55 | 1,805 |
| 8 | + | + | + | + | + | + | + | 68,5 | 67,2 | 67,85 | 0,845 |
| Σ | 42,3 | 16,9 | -0,3 | 8,5 | -0,3 | -2,9 | -1,3 | | | | 7,680 |
| Eff. | 10,575 | 4,225 | -0,075 | 2,125 | -0,075 | -0,725 | -0,325 | | | | 0,960 |
| Sig. | *** | *** | – | ** | – | – | – | | | | |

Die Zeile „Σ“ jeder Effektspalte in Tabelle 7.8 gibt die Summe der Mittelwerte \bar{y}_i , multipliziert mit den Vorzeichen der entsprechenden Spalte. Schätzwerte für die Effekte „Eff.“ erhält man, indem man Σ durch $m/2$ dividiert, hier 4.

Der Mittelwert der Varianzen aller Zeilen ist ein Schätzwert für die Varianz der Zufallsstreuung s^2 der Einzelwerte. In diesem Beispiel erhält man

$$s^2 = \frac{1}{8} \sum_{i=1}^8 s_i^2 = \frac{7,68}{8} = 0,96 \text{ (letzte Spalte)}$$

$$s_d = \sqrt{\frac{4}{N} \cdot s^2} = \sqrt{\frac{4}{16} \cdot 0,96} \approx 0,49$$

$$f = N - 8 = 16 - 8 = 8$$

Mit Tabelle 6.4 erhält man:

$$95\%: t \cdot s_d = 2,306 \cdot 0,49 = 1,13$$

$$99\%: t \cdot s_d = 3,355 \cdot 0,49 = 1,64$$

$$99,9\%: t \cdot s_d = 5,041 \cdot 0,49 = 2,47$$

Der Vergleich der Breiten der Vertrauensbereiche mit den Effekten ergibt die Beurteilung in der letzten Zeile von Tabelle 7.8. Bild 7-7 stellt die Effekte im Vergleich zur Breite des 95%--, 99%- und 99,9%-Vertrauensbereichs grafisch dar. Die Effekte der Temperatur (A), der Zeit (B) und der Wechselwirkung zwischen Temperatur und Zeit (AB) sind signifikant, der Effekt des Katalysators und der Wechselwirkungen zwischen Katalysator und Temperatur bzw. Zeit sind nicht signifikant.

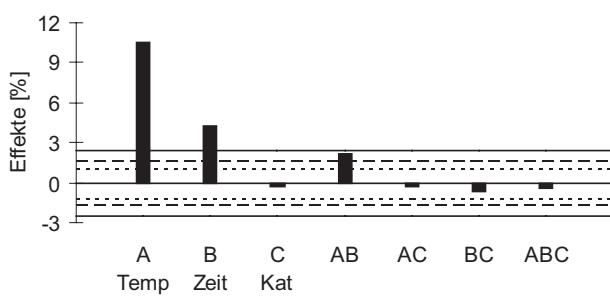


BILD 7-7

Effekte verglichen mit der Breite der Vertrauensbereiche (waagrechte Linien)

Da die Wechselwirkung AB signifikant ist, müssen die Faktoren A und B gemeinsam betrachtet werden. Die Mittelwerte der Faktorstufenkombinationen sind:

| Temperatur | Zeit | Mittelwert Ausbeute |
|------------|------|-------------------------------------|
| 120°C | 2 h | $\frac{53,45 + 53,85}{2} = 53,65\%$ |
| 140°C | 2 h | $\frac{61,65 + 62,55}{2} = 62,10\%$ |
| 120°C | 4 h | $\frac{55,95 + 55,55}{2} = 55,75\%$ |
| 140°C | 4 h | $\frac{69,05 + 67,85}{2} = 68,45\%$ |

Bild 3-3 zeigt diese Mittelwerte der vier Faktorstufenkombinationen grafisch. Die Abweichung von der Parallelität ist Ausdruck der 2FWW zwischen Temperatur und Zeit. Interpretation und Konsequenzen aus den Ergebnissen wurden in Kapitel 3 behandelt.

Für jede signifikante Zwei-Faktor-Wechselwirkung (2FWW) betrachtet man die Mittelwerte der vier Faktorstufenkombinationen für das betroffene Faktorenpaar (siehe oben).

Sollte eine Drei-Faktor-Wechselwirkung (3FWW) signifikant sein, so bedeutet dies, dass der Effekt eines Faktors davon abhängt, welche Faktorstufenkombination der beiden anderen gerade betrachtet wird. Die drei Faktoren müssen gemeinsam behandelt werden. Dazu betrachtet man die Mittelwerte an allen acht Faktorstufenkombinationen (d.h. an allen acht Würfelecken in Bild 7-6).

3FWW sind nur selten von praktischer Bedeutung in dem Sinne, dass sie die Schlussfolgerungen aus einem Versuch wirklich verändern.



Aufgabe (Gewinderollen)

Durch Anpressen von geeigneten Werkzeugen auf einen sich drehenden, runden Rohling werden Gewinde gerollt. Ziel ist es, möglichst reproduzierbar eine Gewindetiefe von 4,3 mm zu erhalten. Dazu werden Rollversuche mit zwei verschiedenen Werkzeugformen (1 und 2) und zwei verschiedenen Anpressdrücken (60 und 90 bar) durchgeführt. Um den Einfluss des Ausgangsmaterials abzuschätzen, werden Rohlinge aus zwei verschiedenen Lieferchargen α und β (weiches und hartes Material) verwendet. Für jede der acht Faktorstufenkombinationen werden drei Gewinde gerollt. Die Tiefe der Gewinde wird gemessen. Umsortiert in die systematische Reihenfolge erhält man folgende Ergebnisse:

| syst. Nr. | Material A | Druck B | Werkzeug C | Einzelergebnisse [mm] | | | \bar{y}_i | s_i^2 |
|--------------|---------------|------------|---------------|--------------------------|------|------|-------------|---------|
| 1 | α | 60 bar | 1 | 4,12 | 4,21 | 4,21 | 4,18 | 0,0027 |
| 2 | β | 60 bar | 1 | 4,17 | 4,03 | 4,10 | 4,10 | 0,0049 |
| 3 | α | 90 bar | 1 | 4,28 | 4,30 | 4,32 | 4,30 | 0,0004 |
| 4 | β | 90 bar | 1 | 4,31 | 4,32 | 4,30 | 4,31 | 0,0001 |
| 5 | α | 60 bar | 2 | 4,26 | 4,31 | 4,33 | 4,30 | 0,0013 |
| 6 | β | 60 bar | 2 | 4,23 | 4,19 | 4,21 | 4,21 | 0,0004 |
| 7 | α | 90 bar | 2 | 4,42 | 4,39 | 4,42 | 4,41 | 0,0003 |
| 8 | β | 90 bar | 2 | 4,46 | 4,43 | 4,40 | 4,43 | 0,0009 |

Berechnen Sie die Effekte aller Faktoren und Wechselwirkungen, und beurteilen Sie ihre Signifikanz. Stellen Sie die 2FWW AB grafisch dar. Welchen Druck und welches Werkzeug würden Sie aufgrund dieser Ergebnisse verwenden? Warum?

Lösung

| syst. Nr. | Material A | Druck B | Werkzeug C | AB | AC | BC | ABC | \bar{y}_i | s_i^2 |
|--------------|---------------|------------|---------------|-------|-------|-------|-------|-------------|----------|
| 1 | - | - | - | + | + | + | - | 4,18 | 0,0027 |
| 2 | + | - | - | - | - | + | + | 4,10 | 0,0049 |
| 3 | - | + | - | - | + | - | + | 4,30 | 0,0004 |
| 4 | + | + | - | + | - | - | - | 4,31 | 0,0001 |
| 5 | - | - | + | + | - | - | + | 4,30 | 0,0013 |
| 6 | + | - | + | - | + | - | - | 4,21 | 0,0004 |
| 7 | - | + | + | - | - | + | - | 4,41 | 0,0003 |
| 8 | + | + | + | + | + | + | + | 4,43 | 0,0009 |
| Σ | -0,14 | 0,66 | 0,46 | 0,20 | 0,00 | 0,00 | 0,02 | | 0,0110 |
| Eff. | -0,035 | 0,165 | 0,115 | 0,050 | 0,000 | 0,000 | 0,005 | $s^2 =$ | 0,001375 |
| Sig. | * | *** | *** | ** | - | - | - | | |

$$s_d = \sqrt{\frac{4}{N} \cdot s^2} = \sqrt{\frac{4}{24} \cdot 0,001375} = \sqrt{0,0002292} = 0,01514$$

$$f = N - 8 = 24 - 8 = 16$$

Mit Tabelle 6.4 erhält man (Interpolation):

$$95\%: t \cdot s_d = 2,12 \cdot 0,01514 = 0,032$$

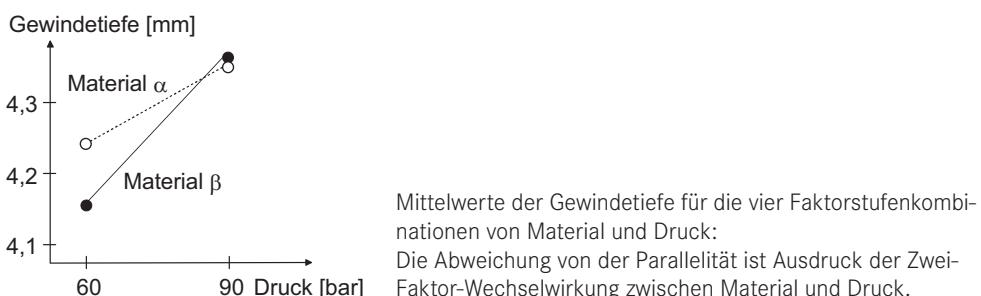
$$99\%: t \cdot s_d = 2,92 \cdot 0,01514 = 0,044$$

$$99,9\%: t \cdot s_d = 4,02 \cdot 0,01514 = 0,061$$

Da die Wechselwirkung AB signifikant ist, müssen diese beiden Faktoren gemeinsam betrachtet werden. Die Mittelwerte für die Faktorstufenkombinationen sind:

| Material | Druck | Mittlere Gewindetiefe |
|----------|--------|--|
| α | 60 bar | $\frac{4,18 + 4,30}{2} = 4,240 \text{ mm}$ |
| β | 60 bar | $\frac{4,10 + 4,21}{2} = 4,155 \text{ mm}$ |
| α | 90 bar | $\frac{4,30 + 4,41}{2} = 4,355 \text{ mm}$ |
| β | 90 bar | $\frac{4,31 + 4,43}{2} = 4,370 \text{ mm}$ |

Bei niedrigem Druck gibt es wesentliche Unterschiede zwischen den Materialien, bei hohem Druck ist die Gewindetiefe kaum vom Material abhängig. Dieses Ergebnis ist plausibel. Der niedrige Druck reicht nicht aus, um das Werkzeugprofil vollständig auf den Rohling zu übertragen. Je nach Härte des Röhrlings erhält man bei geringem Druck unterschiedliche Gewindetiefen. Der hohe Druck (90 bar) reduziert die Abhängigkeit vom Material und ist daher besser geeignet.



Das Werkzeug hat einen signifikanten Einfluss auf die Gewindetiefe. Um das günstigere Werkzeug auszuwählen, betrachtet man die Gewindetiefe beim besseren Druck von 90 bar. Man erhält folgende Mittelwerte:

$$\text{Werkzeug 1: } \frac{4,30 + 4,31}{2} = 4,305 \text{ mm}$$

$$\text{Werkzeug 2: } \frac{4,41 + 4,43}{2} = 4,420 \text{ mm}$$

Das Werkzeug 1 ist günstiger, weil die Gewindetiefe näher am Zielwert von 4,3 mm liegt.

7.2.3 Versuchsumfang

Jeder Effekt ist die Differenz der Mittelwerte von jeweils $N/2$ Werten. Daher ist der erforderliche Versuchsumfang N (= Gesamtzahl der Einzelversuche) in erster Näherung unabhängig von der Anzahl der Faktoren, und (6.17) bleibt (näherungsweise) gültig. Da die Anzahl der Faktorstufenkombinationen $m = 2^k$ beträgt, nimmt die Anzahl der Realisierungen n ab, da

$$N = n \cdot m = 60 \cdot \left(\frac{\sigma}{\Delta\mu} \right)^2 \quad (7.15)$$

Bei Einsatz der faktoriellen Versuchspläne können mehr Faktoren gleichzeitig untersucht werden, (fast) ohne Zunahme des Versuchsumfangs.



Aufgabe

Ein Versuch zur Ausbeuteoptimierung einer chemischen Reaktion soll geplant werden. Aus Erfahrung ist bekannt, dass die Ausbeute aufgrund von Zufallseinflüssen mit einer Standardabweichung von etwa 2 % streut. Tatsächliche Unterschiede von 2 % sind technologisch relevant und sollen mit hoher Wahrscheinlichkeit erkannt werden, falls sie existieren.

Bestimmen Sie den Versuchsumfang N und die Anzahl Realisierungen n jeder Faktorstufenkombination, wenn der Einfluss von $k = 1, 2, 3, 4$ bzw. 5 Faktoren mit je zwei Stufen auf die Ausbeute untersucht werden soll.

Überlegen Sie, was der Versuchsumfang bei 6 und mehr Faktoren sein könnte. Welche neuen Aspekte/Probleme ergeben sich?

Lösung

$$\frac{\sigma}{\Delta\mu} = \frac{2\%}{2\%} = 1$$

Aus (7.15) erhält man $N = n \cdot m = 60$ und damit für

- a) $k = 1$: $m = 2^1 = 2$, für die Anzahl n der Realisierungen jeder Faktorstufenkombination (Stichprobenumfang) erhält man daher $n = 30$
- b) $k = 2$: $m = 2^2 = 4$, daher $n = 15$
- c) $k = 3$: $m = 2^3 = 8$, daher $n = 8$ (aufgerundet von 7,5)
- d) $k = 4$: $m = 2^4 = 16$, daher $n = 4$
- e) $k = 5$: $m = 2^5 = 32$, daher $n = 2$

D. h. die Gesamtzahl der Versuche N ist (fast) unabhängig von der Anzahl der Faktoren. Durch die systematische Anordnung der Faktorstufenkombinationen in der Versuchsplanung erhält man mehr Information ohne zusätzlichen Aufwand.

Solange $n \geq 2$ ist, nimmt die erforderliche Gesamtzahl der Versuche nicht wesentlich zu (mit zunehmender Anzahl Faktoren k nimmt die Anzahl der Faktorstufenkombinationen zu, der Freiheitsgrad f nimmt daher ab, der Unterschied im t-Wert ist aber relativ klein).

Bei sechs Faktoren ($k = 6$) erhält man mit $m = 64$ nur $n = 1$. Die bisher behandelten Auswertungsverfahren versagen. Folgende Möglichkeiten bestehen:

- der Versuchsumfang wird auf $N = 128$ erhöht
- die Varianz s^2 ist bereits vorher bekannt oder wird durch mehrmalige Realisierung einzelner Faktorstufenkombinationen getrennt ermittelt

- die Varianz s^2 wird durch „Pooling zufälliger Effekte“ geschätzt (Abschnitt 7.3)
- nicht alle 2^k Faktorstufenkombinationen werden untersucht (Kapitel 8).

Ähnliches gilt bei sieben und mehr Faktoren.

■ 7.3 Auswertung von Versuchsplänen mit n = 1

Bei Versuchsplänen mit nur einer Realisierung $n = 1$ kann man die Varianz s^2 nicht aus der Streuung zwischen Wiederholungen abschätzen. Stattdessen verwendet man „zufällige Effekte“ zur Schätzung der Varianz.

Vollständige faktorielle Versuchspläne mit $n = 1$ lassen sich meist relativ gut auswerten, wenn die Anzahl der Faktoren $k \geq 5$ ist (d.h. die Anzahl der Faktorstufenkombinationen $m \geq 32$). Bei $k = 4$ oder gar $k = 3$ ist die hier beschriebene Auswertung häufig problematisch und subjektiv.

7.3.1 Wahrscheinlichkeitsdarstellung der Effekte

Wenn alle Effekte in Wirklichkeit 0 sind, so sind die aus Versuchsergebnissen berechneten Effekte aufgrund von zufälligen Unterschieden normalverteilt. Dies kann man nutzen, um zwischen nur zufällig von 0 abweichenden Effekten und wahren Effekten zu unterscheiden.

Wie im Abschnitt 7.2 berechnet man zunächst die Effekte für alle Faktoren und Wechselwirkungen. Diese Effekte werden ins Wahrscheinlichkeitsnetz wie z.B. in Bild 6-11 einge tragen.

Die meisten Effekte sind nahe 0 und liegen näherungsweise auf einer Geraden. Diese Effekte kann man zur Schätzung der Zufallsstreuung verwenden. „Ausreißer“ wie in Bild 6-12 sind Effekte, deren Absolutwert deutlich größer ist, als aufgrund der Zufallsstreuung erwartet – sie sind die vermuteten wahren Effekte. Positive Effekte sind Ausreißer nach oben, negative Effekte nach unten.



Beispiel

Ausbeute einer Halbleiterfertigung (C. D. Montgomery, IBM Schulungsunterlagen von G. Schuster-Kreuzer, nach [5]).

Zielgröße: Ausbeute (linear transformierte Werte)

Faktoren (Parameter eines Phototechnikschriftes und anschließender Ätzung):

| | | |
|--------------------|------|-----------|
| A Blende | 1 | 2 |
| B Belichtungszeit | -20 | + 20 % |
| C Entwicklungszeit | 30 | 45 sec |
| D Masken-Dimension | 1 | 2 |
| E Ätzzeit | 14,5 | 15,5 min. |

TABELLE 7.9 Versuchsplan und Ergebnisse eines vollständigen faktoriellen 2⁵-Versuchs mit n = 1 (umsortiert in die systematische Reihenfolge)

| syst. Nr. | Blende A | Belichtungszeit B | Entwicklungszeit C | Maske D | Ätzzeit E | Ausbeute |
|--------------|-------------|----------------------|-----------------------|------------|--------------|----------|
| 1 | 1 | - 20 % | 30 sec | 1 | 14,5 min | 7 |
| 2 | 2 | - 20 % | 30 sec | 1 | 14,5 min | 9 |
| 3 | 1 | + 20 % | 30 sec | 1 | 14,5 min | 34 |
| 4 | 2 | + 20 % | 30 sec | 1 | 14,5 min | 55 |
| 5 | 1 | - 20 % | 45 sec | 1 | 14,5 min | 16 |
| 6 | 2 | - 20 % | 45 sec | 1 | 14,5 min | 20 |
| 7 | 1 | + 20 % | 45 sec | 1 | 14,5 min | 40 |
| 8 | 2 | + 20 % | 45 sec | 1 | 14,5 min | 60 |
| 9 | 1 | - 20 % | 30 sec | 2 | 14,5 min | 8 |
| 10 | 2 | - 20 % | 30 sec | 2 | 14,5 min | 10 |
| 11 | 1 | + 20 % | 30 sec | 2 | 14,5 min | 32 |
| 12 | 2 | + 20 % | 30 sec | 2 | 14,5 min | 50 |
| 13 | 1 | - 20 % | 45 sec | 2 | 14,5 min | 18 |
| 14 | 2 | - 20 % | 45 sec | 2 | 14,5 min | 21 |
| 15 | 1 | + 20 % | 45 sec | 2 | 14,5 min | 44 |
| 16 | 2 | + 20 % | 45 sec | 2 | 14,5 min | 61 |
| 17 | 1 | - 20 % | 30 sec | 1 | 15,5 min | 8 |
| 18 | 2 | - 20 % | 30 sec | 1 | 15,5 min | 12 |
| 19 | 1 | + 20 % | 30 sec | 1 | 15,5 min | 35 |
| 20 | 2 | + 20 % | 30 sec | 1 | 15,5 min | 52 |
| 21 | 1 | - 20 % | 45 sec | 1 | 15,5 min | 15 |
| 22 | 2 | - 20 % | 45 sec | 1 | 15,5 min | 22 |
| 23 | 1 | + 20 % | 45 sec | 1 | 15,5 min | 45 |
| 24 | 2 | + 20 % | 45 sec | 1 | 15,5 min | 65 |
| 25 | 1 | - 20 % | 30 sec | 2 | 15,5 min | 6 |
| 26 | 2 | - 20 % | 30 sec | 2 | 15,5 min | 10 |
| 27 | 1 | + 20 % | 30 sec | 2 | 15,5 min | 30 |
| 28 | 2 | + 20 % | 30 sec | 2 | 15,5 min | 53 |
| 29 | 1 | - 20 % | 45 sec | 2 | 15,5 min | 15 |
| 30 | 2 | - 20 % | 45 sec | 2 | 15,5 min | 20 |
| 31 | 1 | + 20 % | 45 sec | 2 | 15,5 min | 41 |
| 32 | 2 | + 20 % | 45 sec | 2 | 15,5 min | 63 |

Gemäß (7.11) werden Summen und Effekte mit Hilfe der Vorzeichenspalten für die Faktoren und Wechselwirkungen berechnet. Tabelle 7.10 zeigt eine Auswahl.

TABELLE 7.10 Auswahl aus den insgesamt 31 Vorzeichenpalten und Summen

| Nr. | A | B | C | D | E | AB | AC | AD | ... | ABC | ABD | ... | ABCD | ... | Ausbeute |
|----------|-----|-----|-----|-----|---|-----|----|----|-----|-----|-----|-----|------|-----|----------|
| 1 | - | - | - | - | - | + | + | + | | - | - | | + | | 7 |
| 2 | + | - | - | - | - | - | - | - | | + | + | | - | | 9 |
| 3 | - | + | - | - | - | - | + | + | | + | + | | - | | 34 |
| 4 | + | + | - | - | - | + | - | - | | - | - | | + | | 55 |
| 5 | - | - | + | - | - | + | - | + | | + | - | | - | | 16 |
| 6 | + | - | + | - | - | - | + | - | | - | + | | + | | 20 |
| 7 | - | + | + | - | - | - | - | + | | - | + | | + | | 40 |
| 8 | + | + | + | - | - | + | + | - | | + | - | | - | | 60 |
| 9 | - | - | - | + | - | + | + | - | | - | + | | - | | 8 |
| 10 | + | - | - | + | - | - | - | + | | + | - | | + | | 10 |
| 11 | - | + | - | + | - | - | + | - | | + | - | | + | | 32 |
| 12 | + | + | - | + | - | + | - | + | | - | + | | - | | 50 |
| 13 | - | - | + | + | - | + | - | - | | + | + | | + | | 18 |
| 14 | + | - | + | + | - | - | + | + | | - | - | | - | | 21 |
| 15 | - | + | + | + | - | - | - | - | | - | - | | - | | 44 |
| 16 | + | + | + | + | - | + | + | + | | + | + | | + | | 61 |
| 17 | - | - | - | - | + | + | + | + | | - | - | | + | | 8 |
| 18 | + | - | - | - | + | - | - | - | | + | + | | - | | 12 |
| 19 | - | + | - | - | + | - | + | + | | + | + | | - | | 35 |
| 20 | + | + | - | - | + | + | - | - | | - | - | | + | | 52 |
| 21 | - | - | + | - | + | + | - | + | | + | - | | - | | 15 |
| 22 | + | - | + | - | + | - | + | - | | - | + | | + | | 22 |
| 23 | - | + | + | - | + | - | - | + | | - | + | | + | | 45 |
| 24 | + | + | + | - | + | + | + | - | | + | - | | - | | 65 |
| 25 | - | - | - | + | + | + | + | - | | - | + | | - | | 6 |
| 26 | + | - | - | + | + | - | - | + | | + | - | | + | | 10 |
| 27 | - | + | - | + | + | - | + | - | | + | - | | + | | 30 |
| 28 | + | + | - | + | + | + | - | + | | - | + | | - | | 53 |
| 29 | - | - | + | + | + | + | - | - | | + | + | | + | | 15 |
| 30 | + | - | + | + | + | - | + | + | | - | - | | - | | 20 |
| 31 | - | + | + | + | + | - | - | - | | - | - | | - | | 41 |
| 32 | + | + | + | + | + | + | + | + | | + | + | | + | | 63 |
| Σ | 189 | 543 | 155 | -13 | 7 | 127 | 7 | -1 | | -7 | 5 | | -1 | | |

Die Anzahl der Faktorstufenkombinationen beträgt $m = 32$. Daher gilt:

$$\text{Effekt} = \Sigma/16.$$

Tabelle 7.11 zeigt alle 31 Summen und die daraus berechneten Effekte. In der Spalte „Nr. i“ sind die Effekte der Größe nach nummeriert.

TABELLE 7.11 Alle 31 Summen, die daraus berechneten Effekte ($=\Sigma/16$) und Nummer i in nach der Größe geordneter Reihenfolge. Die vier „Ausreißer“ in Bild 7-8 sind fett markiert.

| | Σ | Effekt | Nr. i |
|-----|----------|--------------|-------|
| A | 189 | 11,81 | 30 |
| B | 543 | 33,94 | 31 |
| C | 155 | 9,69 | 29 |
| D | -13 | -0,81 | 3 |
| E | 7 | 0,44 | 19 |
| AB | 127 | 7,94 | 28 |
| AC | 7 | 0,44 | 20 |
| AD | -1 | -0,06 | 11 |
| AE | 15 | 0,94 | 25 |
| BC | 1 | 0,06 | 13 |
| BD | -11 | -0,69 | 5 |
| BE | 9 | 0,56 | 22 |
| CD | 13 | 0,81 | 23 |
| CE | 5 | 0,31 | 16 |
| DE | -19 | -1,19 | 1 |
| ABC | -7 | -0,44 | 6 |

| | Σ | Effekt | Nr. i |
|-------|----------|--------|-------|
| ABD | 5 | 0,31 | 17 |
| ABE | -3 | -0,19 | 9 |
| ACD | -7 | -0,44 | 7 |
| ACE | 5 | 0,31 | 18 |
| ADE | 13 | 0,81 | 24 |
| BCD | 7 | 0,44 | 21 |
| BCE | 15 | 0,94 | 26 |
| BDE | 3 | 0,19 | 14 |
| CDE | -13 | -0,81 | 4 |
| ABCD | -1 | -0,06 | 12 |
| ABCE | 3 | 0,19 | 15 |
| ABDE | 15 | 0,94 | 27 |
| ACDE | -5 | -0,31 | 8 |
| BCDE | -15 | -0,94 | 2 |
| ABCDE | -3 | -0,19 | 10 |

Um zwischen wahren und nur zufälligen Effekten zu unterscheiden, trägt man die Effekte wie in Bild 6-11 ins Wahrscheinlichkeitsnetz ein, wobei sich der Prozentwert wie in (6.4) aus der Ordnungsnummer i ergibt. Bild 7-8 zeigt das Ergebnis.

Die meisten Effekte sind nahe 0 und liegen näherungsweise auf einer Geraden. Diese Effekte kann man als zufällig betrachten. Einige wenige Effekte an den Extremen oben oder unten (hier A, B, C und AB alle oben) liegen deutlich von dieser Geraden entfernt (bei der Auftragung von Messwerten würde man sie als Ausreißer betrachten). Diese „Ausreißer“ sind die vermuteten wahren Effekte.

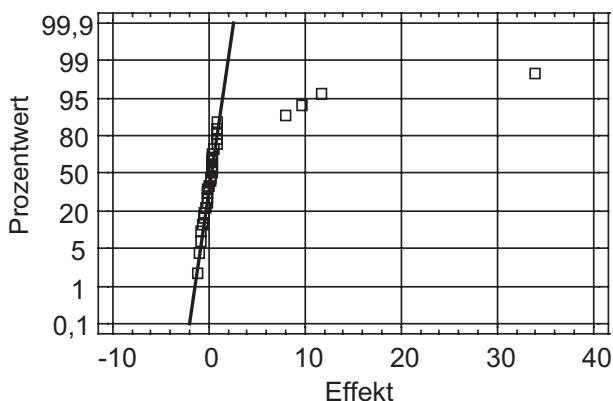


BILD 7-8

Auftragung der Effekte im Wahrscheinlichkeitsnetz

7.3.2 Schätzung der Zufallsstreuung durch „Pooling“

Zur Überprüfung der Signifikanz einzelner Effekte benötigt man ihre Zufallsstreuung. Einen Schätzwert erhält man aus den (vermutlich) zufälligen Effekten – sie liegen auf der Geraden in Bild 7-8. Der Freiheitsgrad f ist die Anzahl dieser Effekte.

$$\text{Varianz } s_d^2 = \frac{1}{\text{Anzahl zufällige Effekte}} \cdot \sum_{\text{zufällige Effekte}} (\text{Effekte})^2 \quad (7.16)$$

$$\text{Freiheitsgrad } f = \text{Anzahl zufällige Effekte} \quad (7.17)$$

Diese Zusammenfassung zufälliger Effekte zu einer Schätzung der Zufallsstreuung wird als „Pooling“ bezeichnet. Eine gewisse Subjektivität bei der Entscheidung darüber, welche Effekte gepoolt werden sollen, ist dabei unvermeidlich. Sie ist jedoch umso weniger kritisch, je größer die Anzahl untersuchter Faktorstufenkombinationen ist. Als Entscheidungshilfen können folgende Kriterien verwendet werden:

- Effekte liegen im Wahrscheinlichkeitsnetz auf einer Geraden
- es handelt sich um Wechselwirkungen vieler Faktoren.



Beispiel (Fortsetzung)

Wendet man diese Kriterien auf das Beispiel zur Ausbeute einer Halbleiterfertigung an, so kann man z. B. entscheiden, dass in Tabelle 7.11 alle 2FWW außer AB und alle 3FWW, 4FWW und 5FWW gepoolt werden sollen. Dann erhält man:

$$s_d^2 = \frac{1}{9+10+5+1} \cdot (0,44^2 + (-0,06)^2 + 0,94^2 + 0,06^2 + \dots + (-0,19)^2) = 0,36$$

$$s_d = \sqrt{0,36} = 0,60$$

$$f = 9 + 10 + 5 + 1 = 25 .$$

Mit Tabelle 6.4 erhält man (Interpolation):

$$95\%: t \cdot s_d = 2,06 \cdot 0,60 = 1,24$$

$$99\%: t \cdot s_d = 2,79 \cdot 0,60 = 1,67$$

$$99,9\%: t \cdot s_d = 3,72 \cdot 0,60 = 2,23 .$$

Der Vergleich der Breiten der Vertrauensbereiche mit den Effekten zeigt:

$$\text{Effekt der Blende (A)} = 11,81 \text{ ***}$$

$$\text{Effekt der Belichtungszeit (B)} = 33,94 \text{ ***}$$

$$\text{Effekt der Entwicklungszeit (C)} = 9,69 \text{ ***}$$

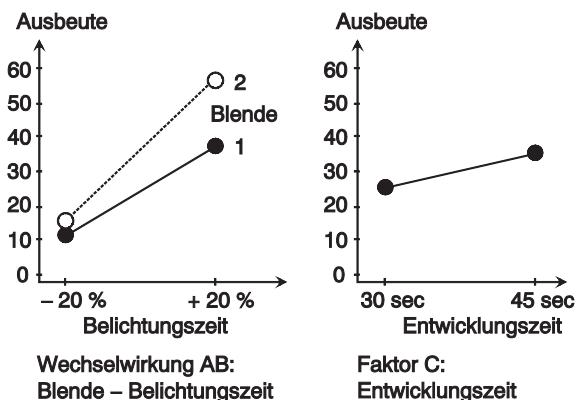
$$\text{Effekt der Maske (D)} = -0,81 \text{ -}$$

$$\text{Effekt der Ätzzeit (E)} = 0,44 \text{ -}$$

$$\text{Effekt der Wechselwirkung AB} = 7,94 \text{ ***.}$$

Die Signifikanz der Wechselwirkung AB bedeutet, dass die Faktoren Blende (A) und Belichtungszeit (B) gemeinsam betrachtet werden müssen. Bild 7-9 zeigt den Effekt der Wechselwirkung AB und des

Faktors C grafisch. Jeder Punkt in der Darstellung der Wechselwirkung AB ist der Mittelwert der jeweils 8 Ausbeuten für die betrachtete Kombination von Blende und Belichtungszeit.

**BILD 7-9**

Grafische Darstellung der signifikanten Effekte:
2FWW AB und Faktor C.

Praktische Konsequenzen aus den Ergebnissen

Eine möglichst hohe Ausbeute erreicht man mit der Faktorstufenkombination:

Blende: 2

Belichtungszeit: + 20 %

Entwicklungszeit: 45 sec.

Maske und Ätzzeit haben keinen signifikanten Einfluss auf die Ausbeute und können zur Optimierung anderer Zielgrößen verwendet werden.



Aufgabe

Wiederholen Sie die Rechnung, indem Sie nur die zehn 3FWW, fünf 4FWW und die eine 5FWW poinen. Vergleichen Sie mit den obigen Ergebnissen.

Lösung

$$s_{\bar{d}}^2 = \frac{1}{10+5+1} \cdot \left((-0,44)^2 + 0,31^2 + (-0,19)^2 + (-0,44)^2 + \dots + (-0,19)^2 \right) = 0,31$$

$$s_{\bar{d}} = \sqrt{0,31} = 0,56$$

$$f = 10 + 5 + 1 = 16.$$

Mit Tabelle 6.4 erhält man (Interpolation):

$$95\%: t \cdot s_{\bar{d}} = 2,12 \cdot 0,56 = 1,19$$

$$99\%: t \cdot s_{\bar{d}} = 2,92 \cdot 0,56 = 1,64$$

$$99,9\%: t \cdot s_{\bar{d}} = 4,02 \cdot 0,56 = 2,25$$

Die Bewertung bleibt unverändert.

7.3.3 Risiken

Bei Versuchen mit nur einer Realisierung $n = 1$ muss die Zufallsstreuung wie gezeigt durch Pooling ermittelt werden. Welche Effekte dabei gepoolt werden, ist zu einem gewissen Grad subjektiv. Streng genommen verwendet man für die Auswertung bereits das (vermutete) Ergebnis.

Die Darstellung im Wahrscheinlichkeitsnetz bietet eine Entscheidungshilfe². Sie ist umso besser, je mehr Effekte eingetragen werden können, d.h. je mehr Faktorstufenkombinationen untersucht wurden.

Die Zufallsstreuung wird dann aus den vermutlich zufälligen Effekten geschätzt. Wird ein Effekt fälschlicherweise zur Zufallsstreuung gerechnet, so wirkt sich dies umso weniger aus, je mehr Effekte insgesamt vorhanden sind.

Beides zusammen bewirkt, dass Versuchspläne mit $n = 1$ umso besser ausgewertet werden können, je mehr Faktorstufenkombinationen der Versuchsplan enthält.

Ein vollständiger faktorieller 2^5 -Versuchsplan enthält 32 Faktorstufenkombinationen. Dies ist meist unkritisch (vgl. Beispiel). Bei wesentlich weniger Faktorstufenkombinationen führt die Darstellung der Effekte im Wahrscheinlichkeitsnetz häufig nicht zu eindeutigen Entscheidungen. Außerdem ist der Aufwand für die Wiederholung der Faktorstufenkombinationen relativ klein.

Daher wird empfohlen, bei Versuchen mit 16 oder weniger Faktorstufenkombinationen normalerweise mindestens $n = 2$ zu verwenden.

Literatur

- [1] Box, G. E. P./Hunter, W. G./Hunter, J. S.: Statistics for Experimenters. John Wiley, New York, 2. Auflage 2005
- [2] Montgomery, D. C.: Design and Analysis of Experiments. John Wiley, New York, 8. Auflage 2012
- [3] Scheffler, E.: Statistische Versuchsplanung und Auswertung; DVG, 3. Auflage Stuttgart 1997
- [4] Petersen, H.: Grundlagen der Statistik und der statistischen Versuchsplanung. Ecomed, Landsberg/Lech 1991
- [5] Qualitätsverbesserung durch Versuchsmethodik, Schulungsunterlagen der DGQ (Deutsche Gesellschaft für Qualität), Frankfurt am Main 1997
- [6] Wu, C. F. J./Hamada, M.: Experiments: Planning, Analysis, and Optimization. John Wiley, New York, 2. Auflage 2009

² In Literatur und Software finden sich auch alternative Entscheidungshilfen:

- Half-Normal-Plot: Darstellung der Beträge der Effekte in einem halben Wahrscheinlichkeitsnetz, dadurch liegen die Ausreißer alle am oberen Ende und sind etwas leichter erkennbar.
- Lenth margin of error: Aus den Medianwerten kleiner Effekte wird ein quantitatives Kriterium zur Unterscheidung zwischen wahren und zufälligen Effekten abgeleitet [6].

Diese helfen bei der Entscheidung, welche Effekte wahr und welche zufällig sind. Sie nehmen dem Anwender die Entscheidung jedoch nicht ab. Es bleibt eine subjektive Entscheidung.

8

Screening- Versuchspläne

Bei vollständigen faktoriellen Versuchen nimmt die Anzahl der Faktorstufenkombinationen mit der Anzahl der Faktoren k sehr schnell zu. In diesem Kapitel wird gezeigt, wie die Anzahl von Faktorstufenkombinationen (und damit der Versuchsaufwand) wesentlich reduziert werden kann (z.B. [1 – 2]).

Die hier behandelten Versuchspläne sind besonders geeignet, wenn erst wenig über eine Problemstellung bekannt ist. Dann sollen meist viele Faktoren untersucht werden und man möchte vor allem wissen, welche dieser vielen Faktoren wichtig sind und in welche Richtung sie die Zielgröße(n) beeinflussen. Aufgrund dieser relativ grobmaschigen Zielsetzung heißen diese Versuchspläne auch „Screening-Versuchspläne“ (screening = aussieben).

Bei unkritischer Anwendung bergen die Screening-Versuchspläne jedoch auch erhebliche Risiken. Daher wird gezeigt, wie diese Risiken begrenzt werden können.

■ 8.1 Hintergrund

Bei vollständigen faktoriellen Versuchen nimmt die Anzahl der Faktorstufenkombinationen m mit der Anzahl der Faktoren k wie $m = 2^k$ zu. Mit acht Faktoren erhält man z.B. bereits 256 Kombinationen. Der Versuchsaufwand wird zu groß, die Gefahr einer Vertauschung von Versuchseinheiten nimmt zu, und es ist praktisch unmöglich, alle Einzelversuche unter identischen Randbedingungen durchzuführen.

Aus den $m = 2^k$ Faktorstufenkombinationen kann man $m - 1$ Effekte berechnen. Mit zunehmender Anzahl Faktoren nimmt dabei vor allem die Anzahl der Wechselwirkungen von mehr als 2 Faktoren zu, wie in Tabelle 8.1 dargestellt.

Dies bedeutet, dass man bei großen k in einem vollständigen faktoriellen Versuch hauptsächlich Wechselwirkungen von mehr als 2 Faktoren untersucht. Wenn man davon ausgeht, dass diese Wechselwirkungen häufig vernachlässigbar sind, erhebt sich die Frage, ob man anstelle dieser „höheren“ Wechselwirkungen weitere Faktoren untersuchen kann.

TABELLE 8.1 Anzahl der Effekte der Faktoren, Zwei-Faktor-Wechselwirkungen (2FWW) und Wechselwirkungen von mehr als 2 Faktoren in Abhängigkeit von der Anzahl der Faktoren

| Anzahl Faktoren k | der Faktoren | Anzahl der Effekte | |
|----------------------|--------------|--------------------|-----------------|
| | | der 2FWW | 3FWW, 4FWW, ... |
| 1 | 1 | - | - |
| 2 | 2 | 1 | - |
| 3 | 3 | 3 | 1 |
| 4 | 4 | 6 | 5 |
| 5 | 5 | 10 | 16 |
| 6 | 6 | 15 | 42 |
| 7 | 7 | 21 | 99 |
| 8 | 8 | 28 | 219 |

■ 8.2 Fraktionelle faktorielle Versuchspläne

Bei den fraktionellen faktoriellen Versuchsplänen werden anstelle von „höheren“ Wechselwirkungen weitere Faktoren untersucht. Dadurch bleibt die Anzahl der Faktorstufenkombinationen unverändert, während die Anzahl der Faktoren erhöht wird.

Diese Vorgehensweise wird zunächst an einem Beispiel demonstriert und anschließend verallgemeinert.

8.2.1 Der fraktionelle faktorielle 2^{4-1} -Plan als Beispiel

Ein vollständiger faktorieller Versuchsplan für die drei Faktoren A, B und C besteht aus 2^3 Faktorstufenkombinationen (oben in Tabelle 8.2). Aus den Ergebnissen kann man die Effekte der Faktoren A, B und C, der 2FWW AB, AC und BC und der 3FWW ABC berechnen.

Kann man nun davon ausgehen, dass die 3FWW ABC vernachlässigbar ist, so ist es naheliegend, anstelle dieser 3FWW einen zusätzlichen Faktor D zu untersuchen. Dies führt zum neuen Versuchsplan in der Mitte von Tabelle 8.2. Die Stufen (= Vorzeichen) in der Spalte für D sind identisch zu denen der Spalte für ABC im oberen Teil.

Der vollständige faktorielle Versuchsplan für die vier Faktoren A, B, C und D (im unteren Teil von Tabelle 8.2) besteht aus $2^4 = 16$ Kombinationen. Die acht Faktorstufenkombinationen des neuen Versuchsplans sind in den 16 Kombinationen des 2^4 -Plans enthalten (die Pfeile markieren die Kombinationen, die im neuen Plan enthalten sind, und geben ihre Nummern).

TABELLE 8.2 Durch die Zuordnung eines vierten Faktors D zur Spalte der 3FWW ABC im 2^3 -Plan entsteht ein Versuchsplan, bei dem vier Faktoren mit acht Faktorstufenkombinationen untersucht werden. Dieser neue 2^{4-1} -Plan (in der Mitte) enthält genau die Hälfte der Faktorstufenkombinationen des 2^4 -Plans im unteren Teil.

Ausgangspunkt ist der vollständige faktorielle 2^3 -Plan;

| syst. Nr. | Faktoren/Effekte | | | | | | |
|--------------|------------------|---|---|----|----|----|-----|
| | A | B | C | AB | AC | BC | ABC |
| 1 | - | - | - | + | + | + | - |
| 2 | + | - | - | - | - | + | + |
| 3 | - | + | - | - | + | - | + |
| 4 | + | + | - | + | - | - | - |
| 5 | - | - | + | + | - | - | + |
| 6 | + | - | + | - | + | - | - |
| 7 | - | + | + | - | - | + | - |
| 8 | + | + | + | + | + | + | + |



mit der Zuordnung
 $ABC \Rightarrow D$ entsteht
daraus der neue
 2^{4-1} -Plan

| syst. Nr. | Faktoren/Effekte | | | |
|--------------|------------------|---|---|---|
| | A | B | C | D |
| 1 | - | - | - | - |
| 2 | + | - | - | + |
| 3 | - | + | - | + |
| 4 | + | + | - | - |
| 5 | - | - | + | + |
| 6 | + | - | + | - |
| 7 | - | + | + | - |
| 8 | + | + | + | + |

als Teil des vollständigen faktoriellen 2^4 -Plans

(die Zahlen in Klammern geben die Nummern aus dem 2^{4-1} -Plan an)

| syst. Nr. | Faktoren/Effekte | | | |
|--------------|------------------|---|---|---|
| | A | B | C | D |
| 1 | - | - | - | - |
| 2 | + | - | - | - |
| 3 | - | + | - | - |
| 4 | + | + | - | - |
| 5 | - | - | + | - |
| 6 | + | - | + | - |
| 7 | - | + | + | - |
| 8 | + | + | + | - |
| 9 | - | - | - | + |
| 10 | + | - | - | + |
| 11 | - | + | - | + |
| 12 | + | + | - | + |
| 13 | - | - | + | + |
| 14 | + | - | + | + |
| 15 | - | + | + | + |
| 16 | + | + | + | + |

↔ (1)

↔ (4)

↔ (6)

↔ (7)

↔ (2)

↔ (3)

↔ (5)

↔ (8)

Da der neue Versuchsplan nur einen Teil der 16 Faktorstufenkombinationen des 2^4 -Plans enthält, nennt man ihn fraktionellen faktoriellen Versuchsplan. In diesem Fall handelt es sich um die Hälfte eines vollständigen faktoriellen 2^4 -Versuchsplans, daher wird dieser fraktionelle faktorielle Versuchsplan üblicherweise mit

$$\frac{1}{2} \cdot 2^4 = 2^{4-1}$$

bezeichnet.

Um die Eigenschaften des neuen Versuchsplans in Tabelle 8.2 weiter zu untersuchen, werden wie in Tabelle 7.7 die Vorzeichenspalten für sämtliche Wechselwirkungen berechnet. Man erhält Tabelle 8.3.

TABELLE 8.3 Vorzeichenspalten für die Berechnung der Effekte der vier Faktoren A, B, C und D, der sechs 2FWW AB, AC, AD, BC, BD und CD, der vier 3FWW ABC, ABD, ACD und BCD und der 4FWW ABCD.

| Nr. | A | B | C | D | AB | AC | AD | BC | BD | CD | ABC | ABD | ACD | BCD | ABCD |
|-----|---|---|---|---|----|----|----|----|----|----|-----|-----|-----|-----|------|
| 1 | - | - | - | - | + | + | + | + | + | + | - | - | - | - | + |
| 2 | + | - | - | + | - | - | + | + | - | - | + | - | - | + | + |
| 3 | - | + | - | + | - | + | - | - | + | - | + | - | + | - | + |
| 4 | + | + | - | - | + | - | - | - | - | + | - | - | + | + | + |
| 5 | - | - | + | + | + | - | - | - | - | + | + | + | - | - | + |
| 6 | + | - | + | - | - | + | - | - | + | - | - | + | - | + | + |
| 7 | - | + | + | - | - | - | + | + | - | - | - | + | + | - | + |
| 8 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |

Die Vorzeichenspalten der Wechselwirkungen ergeben sich als Produkte der Vorzeichenspalten der beteiligten Faktoren, also

$$\text{z.B. } AC = A \cdot C \text{ und } ABD = A \cdot B \cdot D \quad (= 7.10)$$

In Tabelle 8.3 sind jeweils zwei Spalten miteinander identisch (z.B. A und BCD). Bei der Auswertung kann man nicht zwischen diesen beiden Effekten unterscheiden. Berechnet man den Effekt von A (mit Hilfe seiner Vorzeichenspalte), so erhält man in Wirklichkeit die Summe A + BCD. Die Effekte von A und BCD können nicht getrennt werden; man sagt, sie sind miteinander vermenkt. Man verwendet auch den Begriff „Alias“. Für den fraktionellen faktoriellen 2^{4-1} -Plan kann man nur folgende Summeneffekte berechnen:

- 1: A + BCD
- 2: B + ACD
- 3: C + ABD
- 4: D + ABC
- 5: AB + CD

6: AC + BD

7: AD + BC

(8: ABCD + Mittelwert).

Aus den $2^4 = 16$ Faktorstufenkombinationen des vollständigen faktoriellen Versuchsplans kann man 15 Effekte und den Mittelwert (also 16 Größen) berechnen. Aus den $2^{4-1} = 2^3 = 8$ Faktorstufenkombinationen des fraktionellen faktoriellen Versuchsplans kann man wie beim 2^3 -Plan nur 7 Effekte und den Mittelwert (also 8 Größen) berechnen. Jeweils zwei der 15 Effekte sind miteinander vermenkt. Bei der 4FWW ABCD ist für alle Faktorstufenkombinationen $ABCD = +1$, sie ist daher mit dem Mittelwert vermenkt – im Folgenden wird dafür auch die Bezeichnung $ABCD = I$ verwendet (I steht für Identität).

Von den 16 Faktorstufenkombinationen des vollständigen faktoriellen 2^4 -Plans gilt für die eine Hälfte $ABCD = +1$ und für die andere $ABCD = -1$. Das Produkt ABCD bestimmt, welche Faktorstufenkombinationen aus dem vollständigen 2^4 -Plan im 2^{4-1} -Plan verwendet werden. ABCD heißt daher auch „Generator“ des fraktionellen Plans.

Mit Hilfe des Generators und einiger einfacher Rechenregeln kann man leicht die miteinander vermenkten Effekte bestimmen. Aus den Beziehungen:

$$(-1) \cdot (-1) = (+1) \cdot (+1) = +1 \text{ und } (+1) \cdot (-1) = (-1) \cdot (+1) = -1$$

ergeben sich folgende Rechenregeln (jeder Buchstabe steht für eine ganze Vorzeichen-Spalte; I steht für eine Spalte, die nur + enthält):

$$AA = BB = \dots = I \dots AA = BB = \dots = I \quad (8.1)$$

$$AI = IA = A \text{ usw. für alle Faktoren.} \quad (8.2)$$

Durch Einsetzen z.B. in Tabelle 8.3 kann man sich von diesen Rechenregeln leicht überzeugen. $AA = I$ bedeutet nämlich nur, dass man in jeder Zeile + erhält, wenn man die Spalte A mit sich selbst multipliziert.

Mit diesen Rechenregeln erhält man aus der

$$\text{Zuordnung } D = ABC \quad (8.3)$$

durch Multiplikation beider Seiten mit D den

$$\text{Generator } I = ABCD. \quad (8.4)$$

Die miteinander vermengten Effekte erhält man durch Multiplikation mit dem Generator:

- A ist vermengt mit AABCD = BCD
- B ist vermengt mit BABCD = ACD
- C ist vermengt mit CABCD = ABD
- D ist vermengt mit DABCD = ABC
- AB ist vermengt mit ABABCD = CD
- AC ist vermengt mit ACABCD = BD
- AD ist vermengt mit ADABCD = BC
- I ist vermengt mit ABCD, d.h. ABCD ist mit dem Mittelwert vermengt.

Anmerkungen

- Durch die Zuordnung $D = ABC$ werden aus den 16 Faktorstufenkombinationen des Plans 2^4 in Tabelle 8.2 die gekennzeichneten 8 ausgewählt. Für diese Zuordnung ist z.B. A mit BCD vermengt, d.h. man bestimmt $A + BCD$. Alternativ kann man die Zuordnung $D = -ABC$ verwenden. Dann wählt man die anderen 8 Versuche aus. Der Generator ist $I = -ABCD$ und die berechneten Effekte sind z.B. $A - BCD$ usw.
- Fraktionelle faktorielle Versuchspläne können bei Bedarf zu vollständigen faktoriellen Versuchen ergänzt werden, indem man die weggelassenen Kombinationen nachholt.
- Dadurch, dass die fraktionellen faktoriellen Versuchspläne gegenüber den vollständigen faktoriellen Plänen zusätzliche Faktoren enthalten, sind sie sehr rationell. Der Preis, den man dafür bezahlen muss, ist das Vermengen der Effekte. Die Versuchspläne müssen daher sorgfältig anhand der vorhandenen Informationen oder Annahmen über mögliche Wechselwirkungen ausgewählt werden.
- Der fraktionelle faktorielle 2^{4-1} -Plan in Tabelle 8.2 beispielsweise eignet sich immer dann, wenn
 - die 3FWW vernachlässigt werden können und
 - einer der vier Faktoren A, B, C, D keine 2FWW mit den anderen Faktoren aufweist.
 Diese Bedingungen sind z.B. erfüllt, wenn der Faktor D ein Blockfaktor ist.

8.2.2 Anwendung des 2^{4-1} -Plans zur Blockbildung

Wenn ein Versuchsplan aus einer größeren Anzahl von Einzelversuchen besteht, ist es zweckmäßig, den Versuch in mehreren Blöcken durchzuführen. So kann man leichter sicherstellen, dass die Einzelversuche innerhalb eines Blocks unter gleichen Randbedingungen durchgeführt werden.

Bei einem faktoriellen 2^3 -Plan kann man z.B. D = ABC als Blockfaktor verwenden. Alle Einzelversuche mit „-“ in der Spalte ABC bilden Block 1, alle Einzelversuche mit „+“ in der Spalte ABC bilden Block 2 (Tabelle 8.4).

| Block D = ABC | syst. Nr. | Faktoren | | |
|------------------|-----------|----------|---|---|
| | | A | B | C |
| 1 (-) | 1 | - | - | - |
| 1 (-) | 4 | + | + | - |
| 1 (-) | 6 | + | - | + |
| 1 (-) | 7 | - | + | + |
| 2 (+) | 2 | + | - | - |
| 2 (+) | 3 | - | + | - |
| 2 (+) | 5 | - | - | + |
| 2 (+) | 8 | + | + | + |

TABELLE 8.4

Blockbildung mit Hilfe des Blockfaktors D = ABC:
 In Block 1 ist ABC -, in Block 2 ist ABC +. Der Versuchsplan entsteht durch Umsortieren des üblichen 2³-Plans (syst. Nr. gibt die Nummer im 2³-Plan, z.B. oben in Tabelle 8.2).

Blockbildung ist z. B. sinnvoll, wenn Material aus zwei verschiedenen Chargen verwendet werden muss oder wenn sich ein Versuch über mehrere Tage erstreckt. Dann verhalten sich die beiden Chargen meist nicht genau gleich bzw. die Randbedingungen an verschiedenen Tagen sind nicht absolut gleich. Durch Blockbildung erreicht man, dass hierdurch die Effekte nicht verfälscht werden (und die Streuung nicht vergrößert wird).

Block 1 sind dann die Einzelversuche mit Material aus Charge 1 oder die Versuche, die an Tag 1 durchgeführt werden. Man darf davon ausgehen, dass der Blockfaktor, in diesen Beispielen die Charge bzw. der Tag, keine Wechselwirkung mit den übrigen Faktoren des Versuchs aufweist.

Die hier beschriebene Blockbildung ist eine Weiterführung der in Kapitel 7 beschriebenen Blockbildung. In Kapitel 7 wurden Blöcke aus m Einzelversuchen gebildet. Das hier beschriebene Verfahren führt zu Blöcken aus m/2 Einzelversuchen, d.h. eine feinere Unterteilung. Bei einer Abweichung von der vollen Randomisierung kann man die hier beschriebene Blockbildung für eine grobe Absicherung verwenden.



Beispiel

Eine Metallschicht wird galvanisch auf ein unregelmäßig geformtes Teil abgeschieden. Aufgrund der Geometrie des Teils erhält man nicht überall die gleiche Schichtdicke. Aus Erfahrung ist bekannt, wo man eine besonders geringe bzw. besonders hohe Schichtdicke erhält. Ziel ist eine möglichst kleine Schichtdickendifferenz. Der Absolutwert kann über die Abscheidezeit leicht verändert werden.

Der Einfluss folgender Faktoren soll untersucht werden:

- A Strom 500 600 A
- B Temperatur 50 70 °C
- C Zusatzstoff ohne mit

Die Veränderung dieser Faktoren ist sehr aufwändig und dauert jeweils mehr als eine Stunde. Die Stufenwerte können genau und reproduzierbar eingestellt werden. Zwischen verschiedenen (auch nacheinander gefertigten) Teilen gibt es jedoch zufällige Unterschiede. Daher entscheidet man sich für folgende Vorgehensweise:

- Für jede Faktorstufenkombination werden nacheinander drei Teile gefertigt und die Schichtdicke an der Stelle mit der erfahrungsgemäß kleinsten und größten Schichtdicke wird gemessen (n = 3).
- Wegen des Zeitaufwands für die Umstellung der Faktorstufen muss der Versuch über zwei aufeinanderfolgende Tage verteilt werden. Da sich die Eigenschaften des Galvanikbades über Nacht verändern können, soll der Tag als Blockfaktor berücksichtigt werden.

- Innerhalb jeden Blocks sollen Temperatur und Zusatzstoff möglichst wenig verändert werden, da sie besonders aufwändig zu verändern sind.
- Da die mittlere Schichtdicke vom Produkt Strom \times Zeit (= Ladung) abhängt, wird die Abscheidezeit so verändert, dass das Produkt Strom \times Zeit konstant bleibt. Bei Strom 500 A beträgt die Abscheidezeit 60 sec, bei 600 A nur 50 sec.

Tabelle 8.5 zeigt die Ergebnisse in der Reihenfolge der Versuchsdurchführung (alle Schichtdicken in μm).

Als Zielgröße für die Analyse eignet sich in diesem Beispiel die Differenz der Schichtdicke an der Position mit erfahrungsgemäß minimaler und maximaler Dicke an jedem Teil. D.h. bei der ersten Faktorstufenkombination erhält man bei Teil 1 einen Wert von $8,3 - 6,6 = 1,7$. Im Folgenden wird nur mit diesen Werten weitergerechnet. In Tabelle 8.6 werden Mittelwerte und Varianzen dieser Differenzen für die drei Teile jeder Faktorstufenkombination angegeben. Die Einzelversuche wurden in die Standardreihenfolge (syst. Nr.) sortiert. Die Stufenwerte wurden durch – bzw. + ersetzt.

TABELLE 8.5 Schichtdicke (in μm) an Position mit erfahrungsgemäß minimaler und maximaler Dicke, gemessen an je drei Teilen, in der Reihenfolge der Versuchsdurchführung

| Vers. Nr. | syst. Nr. | Tag | Strom A | Temp. B | Zusatz C | Teil 1 min | max | Teil 2 min | max | Teil 3 min | max |
|-----------|-----------|-----|---------|---------|----------|------------|-----|------------|-----|------------|-----|
| 1 | 1 | 1 | 500 A | 50 °C | ohne | 6,6 | 8,3 | 6,4 | 8,2 | 6,6 | 8,5 |
| 2 | 4 | 1 | 600 A | 70 °C | ohne | 5,7 | 8,0 | 5,7 | 8,0 | 6,0 | 8,0 |
| 3 | 7 | 1 | 500 A | 70 °C | mit | 5,8 | 7,4 | 6,0 | 7,3 | 6,3 | 7,3 |
| 4 | 6 | 1 | 600 A | 50 °C | mit | 6,2 | 8,2 | 6,1 | 8,0 | 6,3 | 8,1 |
| 5 | 5 | 2 | 500 A | 50 °C | mit | 6,1 | 8,1 | 6,4 | 8,1 | 6,5 | 8,2 |
| 6 | 8 | 2 | 600 A | 70 °C | mit | 5,7 | 7,6 | 5,6 | 7,9 | 5,7 | 7,8 |
| 7 | 3 | 2 | 500 A | 70 °C | ohne | 6,1 | 8,1 | 6,2 | 8,1 | 6,2 | 8,3 |
| 8 | 2 | 2 | 600 A | 50 °C | ohne | 6,0 | 8,9 | 6,4 | 8,6 | 6,6 | 8,7 |

TABELLE 8.6 Berechnung der Effekte der Faktoren und Wechselwirkungen (Einzelversuche sortiert nach syst. Nr., der Standardreihenfolge für 2^3 -Pläne)

| Vers. Nr. | syst. Nr. | Tag+ ABC | A | B | C | AB | AC | BC | Teil 1 Diff. | Teil 2 Diff. | Teil 3 Diff. | \bar{y}_i | s_i^2 |
|-----------|-----------|--------------------|---------------------|---------------------|----------------------|-------------------|-------------------|---------------------|--------------|--------------|--------------|-------------|--------------------------|
| 1 | 1 | – | – | – | – | + | + | + | 1,7 | 1,8 | 1,9 | 1,8 | 0,01 |
| 8 | 2 | + | + | – | – | – | – | + | 2,9 | 2,2 | 2,1 | 2,4 | 0,19 |
| 7 | 3 | + | – | + | – | – | + | – | 2,0 | 1,9 | 2,1 | 2,0 | 0,01 |
| 2 | 4 | – | + | + | – | + | – | – | 2,3 | 2,3 | 2,0 | 2,2 | 0,03 |
| 5 | 5 | + | – | – | + | + | – | – | 2,0 | 1,7 | 1,7 | 1,8 | 0,03 |
| 4 | 6 | – | + | – | + | – | + | – | 2,0 | 1,9 | 1,8 | 1,9 | 0,01 |
| 3 | 7 | – | – | + | + | – | – | + | 1,6 | 1,3 | 1,0 | 1,3 | 0,09 |
| 6 | 8 | + | + | + | + | + | + | + | 1,9 | 2,3 | 2,1 | 2,1 | 0,04 |
| Σ | | 1,1 0,275 ** | 1,7 0,425 *** | -0,3 -0,075 – | -1,3 -0,325 ** | 0,3 0,075 – | 0,1 0,025 – | -0,3 -0,075 – | | | | | 0,41 0,051 $s^2 =$ |
| Eff. | | | | | | | | | | | | | |
| Sig. | | | | | | | | | | | | | |

Die Berechnung der Effekte und die Beurteilung ihrer Signifikanz verlaufen wie bei vollständigen faktoriellen Versuchsplänen gemäß (7.11ff):

$$s_d = \sqrt{\frac{4}{N} \cdot s^2} = \sqrt{\frac{4}{24} \cdot 0,051} = \sqrt{0,0085} = 0,092$$

$$f = N - 8 = 24 - 8 = 16.$$

Mit Tabelle 6.4 erhält man für die Breite der Vertrauensbereiche:

$$95\%: t \cdot s_d = 2,12 \cdot 0,092 = \pm 0,20$$

$$99\%: t \cdot s_d = 2,92 \cdot 0,092 = \pm 0,27$$

$$99,9\%: t \cdot s_d = 4,02 \cdot 0,092 = \pm 0,37$$

Der Vergleich mit den Effekten ergibt die Beurteilung ihrer Signifikanz in der letzten Zeile von Tabelle 8.6. Im Gegensatz zu vollständigen faktoriellen Plänen sind hier aber immer zwei Effekte miteinander vermischt. So ist z. B. A mit BCD vermischt. Der Effekt BCD gibt an, wie stark sich die 2FWW BC von einem Tag zum anderen verändert hat. Man geht davon aus, dass diese Änderung vernachlässigbar ist und man interpretiert das Ergebnis als A (analog für C). D.h. nur die Effekte der Faktoren Strom und Zusatz auf die Differenz sind signifikant. Da das Ziel eine möglichst kleine Differenz ist, wählt man:

Strom 500 A

Zusatz mit.

Dadurch erniedrigt sich die Schichtdickendifferenz um $0,425 + 0,325 = 0,75$ gegenüber der ungünstigsten Faktorstufenkombination.

Die Temperatur hat keinen signifikanten Einfluss. Sie kann daher aufgrund anderer Überlegungen festgelegt werden (z. B. Kosten oder Einfluss auf andere Zielgrößen).

Der signifikante Blockfaktor bedeutet, dass sich von einem Tag zum anderen etwas verändert hat, das ebenfalls die Differenz beeinflusst. Eine genauere Analyse der Veränderungen zwischen den beiden Tagen kann evtl. Hinweise auf die Ursache ergeben. Hätte man die Versuche ohne Blockbildung durchgeführt, hätte dieser Unterschied die Effekte (und damit die Schlussfolgerungen) verfälschen können.

8.2.3 Fraktionaler faktorieller 2^{k-p} -Plan

Analog zum Übergang vom vollständigen faktoriellen 2^3 -Plan zum fraktionellen faktoriellen 2^{4-1} -Plan kann man

- von beliebigen vollständigen faktoriellen Versuchsplänen für $k-p$ Faktoren ausgehen und
- p Faktoren durch p Zuordnungen zusätzlich aufnehmen.

Der so entstandene Versuchsplan enthält dann insgesamt k Faktoren und 2^{k-p} Faktorstufenkombinationen. D.h. es wird nur ein

Anteil $\frac{1}{2^p}$ der insgesamt 2^k Faktorstufenkombinationen des vollständigen faktoriellen

Versuchsplans, also $\frac{1}{2^p} \cdot 2^k = 2^{k-p}$

untersucht. Dadurch wird eine wesentliche Reduktion im Versuchsaufwand erreicht. Der Preis dieser Reduktion ist, dass auch nur ein

Anteil $\frac{1}{2^p}$ der insgesamt $2^k - 1$ Effekte des vollständigen Plans

berechnet werden kann. Das bedeutet, dass jeder Effekt des fraktionellen faktoriellen 2^{k-p} -Versuchsplans die Summe von 2^p Effekten ist (Vermengung). Die Vermengungsstruktur ergibt sich aus den p Zuordnungen mit Hilfe der Rechenregeln (8.1)–(8.4).



Beispiel 1

Ausgehend vom vollständigen faktoriellen 2^4 -Plan mit den 4 Faktoren A, B, C und D wird durch die Zuordnung E = ABCD ein fraktioneller faktorieller 2^{5-1} -Plan mit den 5 Faktoren A, B, C, D und E erzeugt. Dieser Plan hat 16 Faktorstufenkombinationen. Der Generator des Plans ist I = ABCDE. Jeder Effekt des 2^{5-1} -Plans ist die Summe von $2^1 = 2$ Effekten, die sich direkt aus dem Generator ergeben.

A ist vermengt mit AABCDE = BCDE, usw. D. h. man kann nur folgende Effekte berechnen:

| | | | |
|----------|----------|----------|-----------|
| A + BCDE | B + ACDE | C + ABDE | D + ABCE |
| E + ABCD | AB + CDE | AC + BDE | AD + BCE |
| AE + BCD | BC + ADE | BD + ACE | BE + ACD |
| CD + ABE | CE + ABD | DE + ABC | (I+ABCDE) |

Effekte der Faktoren sind mit den Effekten von Vier-Faktor-Wechselwirkungen (4FWW) vermischt, und 2FWW sind mit 3FWW vermischt. I + ABCDE bedeutet, dass die 5FWW ABCDE den Mittelwert aller Versuchsergebnisse verfälschen kann. Da der Mittelwert eigentlich kein Effekt ist, wurde dieser Term eingeklammert, er ist jedoch für das Verständnis der Systematik nützlich.

Beispiel 2

Ausgehend vom vollständigen faktoriellen 2^3 -Plan mit den 3 Faktoren A, B und C wird durch die Zuordnungen D = ABC und E = AB ein fraktioneller faktorieller 2^{5-2} -Plan mit den 5 Faktoren A, B, C, D und E erzeugt. Dieser Plan hat 8 Faktorstufenkombinationen. Zwei Generatoren I = ABCD und I = ABE des Plans erhält man direkt aus den beiden Zuordnungen. Jeder Effekt des 2^{5-2} -Plans ist die Summe von $2^2 = 4$ Effekten. Es fehlt also noch ein Generator. Dieser ergibt sich aus dem Produkt der beiden anderen Generatoren I · I = I = ABCD · ABE = CDE.

Insgesamt erhält man: I = ABCD = ABE = CDE.

A ist daher vermischt mit AABCD = BCD, AABE = BE und ACDE usw.

Man kann nur folgende Effekte berechnen:

| | |
|---------------------|------------------------|
| A + BCD + BE + ACDE | B + ACD + AE + BCDE |
| C + ABD + ABCE + DE | D + ABC + ABDE + CE |
| E + ABCDE + AB + CD | AC + BD + BCE + ADE |
| AD + BC + BDE + ACE | (I + ABCD + ABE + CDE) |

Effekte der Faktoren sind mit den Effekten von 2FWW, 3FWW und 4FWW vermischt.

Beispiel 3

Ausgehend vom vollständigen faktoriellen 2^3 -Plan mit den 3 Faktoren A, B und C wird durch die Zuordnungen D = ABC, E = AB und F = AC ein fraktioneller faktorieller 2^{6-3} -Plan mit den 6 Faktoren A, B, C, D, E und F erzeugt. Dieser Plan hat 8 Faktorstufenkombinationen. Als Generatoren des Plans erhält man direkt aus den Zuordnungen I = ABCD, I = ABE und I = ACF. Jeder Effekt des 2^{6-3} -Plans ist die Summe von $2^3 = 8$ Effekten. Es fehlen also noch vier Generatoren. Diese ergeben sich aus den Produkten der drei anderen Generatoren

$$I \cdot I = I = ABCD \cdot ABE = CDE$$

$$I \cdot I = I = ABCD \cdot ACF = BDF$$

$$I \cdot I = I = ABE \cdot ACF = BCEF$$

$$I \cdot I \cdot I = I = ABCD \cdot ABE \cdot ACF = ADEF.$$

Insgesamt erhält man: I = ABCD = ABE = ACF = CDE = BDF = BCEF = ADEF.

Statt A berechnet man daher wegen der Vermengung:

$$A + BCD + BE + CF + ACDE + ABDF + ABCEF + DEF \text{ usw.}$$



Aufgabe 1

Ausgehend vom vollständigen faktoriellen 2^2 -Plan mit den 2 Faktoren A und B wird durch die Zuordnung C = AB ein fraktioneller faktorieller 2^{3-1} -Plan mit den 3 Faktoren A, B und C erzeugt. Bestimmen Sie Generator und Vermengungen dieses Plans. Wie lautet der Versuchsplan?

Lösung 1

Generator I = ABC

Vermengungen:

$$A + BC \quad B + AC \quad C + AB \quad (I + ABC)$$

Versuchsplan (Faktorstufenkombinationen):

| syst. Nr. | Faktoren | | |
|--------------|----------|---|---|
| | A | B | C |
| 1 | – | – | + |
| 2 | + | – | – |
| 3 | – | + | – |
| 4 | + | + | + |

Aufgabe 2

Ausgehend vom vollständigen faktoriellen 2^2 -Plan mit den 2 Faktoren A und B wird durch die Zuordnung C = -AB ein fraktioneller faktorieller 2^{3-1} -Plan mit den 3 Faktoren A, B und C erzeugt. Bestimmen Sie Generator und Vermengungen dieses Plans. Wie lautet der Versuchsplan? Welchen Plan erhält man, wenn man die Versuchspläne von Aufgabe 1 und 2 zusammenfügt?

Lösung 2

Generator I = -ABC

Vermengungen:

$$A - BC \quad B - AC \quad C - AB \quad (I - ABC)$$

Versuchsplan (Faktorstufenkombinationen):

| syst. Nr. | Faktoren | | |
|--------------|----------|---|---|
| | A | B | C |
| 1 | - | - | - |
| 2 | + | - | + |
| 3 | - | + | + |
| 4 | + | + | - |

Zusammen ergeben die Versuchspläne von Aufgabe 1 und 2 den vollständigen faktoriellen 2^3 -Plan.

Geometrische Darstellung

Der fraktionelle faktorielle 2^{k-p} -Plan enthält nur einen Anteil $1/2^p$ der Faktorstufenkombinationen des vollständigen faktoriellen 2^k -Plans. Die Auswahl der untersuchten Faktorstufenkombinationen (= Versuchspunkte) erfolgt dabei so, dass der Versuchsräum durch diese Punkte möglichst gleichmäßig überdeckt wird. Bild 8-1 zeigt als Beispiel den vollständigen faktoriellen 2^3 -Plan (alle acht Eckpunkte des Würfels) und die durch die Zuordnung $C = -AB$ daraus ausgewählten vier Eckpunkte für den 2^{3-1} -Plan (wie in Aufgabe 2, mit Kreisen markiert). Mit der Zuordnung $C = AB$ wählt man die vier nicht markierten Eckpunkte aus (wie in Aufgabe 1). Aus Bild 8-1 erkennt man:

- Die vier Versuchspunkte füllen den Würfel so gleichmäßig aus, wie das mit vier Punkten möglich ist.
- Stellt sich nachträglich heraus, dass einer der Faktoren keinen Einfluss hat, so erhält man einen vollständigen Versuchsplan in den verbliebenen zwei Faktoren (eine der gestrichelt dargestellten Projektionen).
- Der fraktionelle Versuchsplan kann zum vollständigen Versuchsplan ergänzt werden, indem die fehlenden Versuchspunkte nachgeholt werden.

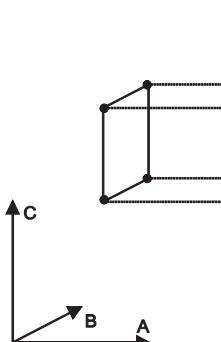


BILD 8-1

Vollständiger faktorieller Versuchsplan = alle Eckpunkte des Würfels

Fraktioneller faktorieller Versuchsplan = nur markierte (oder nur nicht markierte) Eckpunkte

8.2.4 Was bedeutet Vermengung?

Wenn zwei oder mehr Effekte miteinander vermengt sind, dann bedeutet das, dass bei der Auswertung nicht zwischen diesen Effekten unterschieden werden kann.

Im fraktionellen Versuchsplan wurde nur ein Teil der Faktorstufenkombinationen des vollständigen Plans realisiert, und bei diesen Faktorstufenkombinationen unterscheiden sich die vermengten Effekte nicht.

Statt des gewünschten Effekts allein erhält man immer die Summe der miteinander vermengten Effekte.

Wenn man durch technische Überlegungen sagen kann, welcher der miteinander vermengten Effekte das Ergebnis bestimmt, dann hat man im Versuch (zumindest näherungsweise) diesen Effekt bestimmt. Man hat Einzelversuche eingespart.

Sind die technischen Überlegungen falsch, oder ignoriert man die potentielle Verfälschung einfach, so kann man zu einem falschen Ergebnis kommen.

Hat man nach dem Versuchsergebnis Zweifel an der richtigen Interpretation, so kann man den fraktionellen Plan gezielt erweitern, um zwischen zwei Effekten zu unterscheiden.

Bei der Entscheidung, welcher der miteinander vermengten Effekte wichtig ist und ob ggf. weitere Einzelversuche zur eindeutigen Unterscheidung notwendig sind, **helfen nur technische Überlegungen, nicht Statistik.**



Beispiel zur Auswirkung von Vermengungen

Wenn zwei Effekte miteinander vermengt sind, so bedeutet das, dass man nur die Summe der Effekte bestimmen kann und nicht die Einzeleffekte. Um dies zu verdeutlichen, gehen wir von einem vollständigen faktoriellen 2^3 -Versuch aus, für den die Effekte aller Faktoren und Wechselwirkungen bestimmt werden können. Anschließend betrachten wir die beiden fraktionellen Pläne, die man durch Auswahl der Hälfte der Faktorstufenkombinationen aus dem vollständigen Versuch erhält.

Um zwischen Zufallsstreuung und Vermengung zu trennen, betrachten wir einen hypothetischen Versuch ohne Zufallsstreuung, mit den Mittelwerten aus der Aufgabe zum Gewinderollen (Absatz 7.2.2) als Ergebnisse.

In diesem Beispiel sind die Effekte von 2FWW mit denen von Faktoren vermengt. Dadurch können 2FWW die Effekte von Faktoren verfälschen. Wie dieses Beispiel zeigt, sind solche extremen Pläne mit erheblichen Risiken verbunden und sollten möglichst vermieden werden.

In den folgenden Absätzen wird dann gezeigt, wie man die Vorteile von fraktionellen Plänen nutzen und gleichzeitig die hier aufgezeigten Risiken vermeiden kann. Zunächst soll jedoch **Problemwusstsein** geschaffen werden.

Beginnen wir mit dem vollständigen faktoriellen 2^3 -Versuch. Tabelle 8.7 zeigt die Ergebnisse.

| Nr. | Material A | Druck B | Werkzeug C | Ergebnisse [mm] |
|-----|------------|---------|------------|-----------------|
| 1 | α | 60 bar | 1 | 4,18 |
| 2 | β | 60 bar | 1 | 4,10 |
| 3 | α | 90 bar | 1 | 4,30 |
| 4 | β | 90 bar | 1 | 4,31 |
| 5 | α | 60 bar | 2 | 4,30 |
| 6 | β | 60 bar | 2 | 4,21 |
| 7 | α | 90 bar | 2 | 4,41 |
| 8 | β | 90 bar | 2 | 4,43 |

TABELLE 8.7

Ergebnisse des vollständigen faktoriellen 2^3 -Versuchs (ohne Zufallsstreuung)

Daraus ergeben sich die Effekte gemäß Tabelle 8.8, wobei Effekt = $\Sigma / 4$, da die Anzahl der Faktorstufenkombinationen (Zeilen) $m = 8$ beträgt.

TABELLE 8.8 Berechnung der Effekte aus den Ergebnissen in Tabelle 8.7

| Nr. | Material A | Druck B | Werkzeug C | AB | AC | BC | ABC | y_i |
|----------|------------|---------|------------|-------|-------|-------|-------|-------|
| 1 | - | - | - | + | + | + | - | 4,18 |
| 2 | + | - | - | - | - | + | + | 4,10 |
| 3 | - | + | - | - | + | - | + | 4,30 |
| 4 | + | + | - | + | - | - | - | 4,31 |
| 5 | - | - | + | + | - | - | + | 4,30 |
| 6 | + | - | + | - | + | - | - | 4,21 |
| 7 | - | + | + | - | - | + | - | 4,41 |
| 8 | + | + | + | + | + | + | + | 4,43 |
| Σ | -0,14 | 0,66 | 0,46 | 0,20 | 0,00 | 0,00 | 0,02 | |
| Eff. | -0,035 | 0,165 | 0,115 | 0,050 | 0,000 | 0,000 | 0,005 | |

Der fraktionelle faktorielle 2^{3-1} -Plan, der mit der Zuordnung $C = AB$ entsteht, enthält nur die in Tabelle 8.9 dargestellten Faktorstufenkombinationen (dies sind die vier Zeilen aus Tabelle 8.7, für die $ABC = +1$ ist). Daraus berechnen wir die Effekte in Tabelle 8.10 (Effekt = $\Sigma/2$, da $m = 4$).

| Nr. | Material A | Druck B | Werkzeug C | Ergebnisse [mm] |
|-----|------------|---------|------------|-----------------|
| 2 | B | 60 bar | 1 | 4,10 |
| 3 | A | 90 bar | 1 | 4,30 |
| 5 | A | 60 bar | 2 | 4,30 |
| 8 | B | 90 bar | 2 | 4,43 |

TABELLE 8.9

Ergebnisse des fraktionellen faktoriellen 2^{3-1} -Versuchs, der durch die Zuordnung $C = AB$ entsteht (die Hälfte der Zeilen von Tabelle 8.7)

TABELLE 8.10 Berechnung der Effekte aus den Ergebnissen in Tabelle 8.9 (nur zur Demonstration mit allen Spalten aus Tabelle 8.8)

| Nr. | Material A | Druck B | Werkzeug C | AB | AC | BC | ABC | y_i |
|----------|------------|---------|------------|-------|-------|--------|-------|-------|
| 2 | + | - | - | - | - | + | + | 4,10 |
| 3 | - | + | - | - | + | - | + | 4,30 |
| 5 | - | - | + | + | - | - | + | 4,30 |
| 8 | + | + | + | + | + | + | + | 4,43 |
| Σ | -0,07 | 0,33 | 0,33 | 0,33 | 0,33 | -0,07 | 17,13 | |
| Eff. | -0,035 | 0,165 | 0,165 | 0,165 | 0,165 | -0,035 | | |

Dadurch, dass im Vergleich zu Tabelle 8.8 die Hälfte der Einzelversuche fehlt, kann z.B. nicht mehr zwischen den Effekten von A und BC unterschieden werden; bei beiden erhält man $A + BC$. In diesem Beispiel ist die Wechselwirkung $BC = 0$, daher wird A nicht verändert. Für BC erhält man jedoch statt 0 ebenfalls das Ergebnis -0,035.

Statt des Effekts von C erhält man $C + AB$, also $0,115 + 0,050 = 0,165$. Hat man nur dieses Ergebnis vorliegen, kann man nicht feststellen, wie groß C und AB einzeln sind. Stellt man sich auf den Standpunkt, dass die Wechselwirkung vernachlässigbar ist, so meint man $C = 0,165$.

Die 3FWW ABC kann aus Tabelle 8.10 nicht bestimmt werden. Der Mittelwert der Einzelergebnisse in Tabelle 8.10 beträgt 4,2825, der Mittelwert der Einzelergebnisse in Tabelle 8.8 beträgt 4,28. Beim 2^{3-1} -Plan ist der Mittelwert also um $ABC/2$ verschoben, man berechnet $I + ABC$ (Faktor 2, da bei der Berechnung des Mittelwerts durch m , bei der Berechnung des Effekts nur durch $m/2$ dividiert wird).

Hätten wir den fraktionellen faktoriellen 2^{3-1} -Plan mit der Zuordnung $C = -AB$ erzeugt, so hätten wir gerade die anderen vier Faktorstufenkombinationen in Tabelle 8.7 durchgeführt (für die $ABC = -1$ ist) und daraus die Effekte in Tabelle 8.11 berechnet.

TABELLE 8.11 Ergebnisse des fraktionellen faktoriellen 2^{3-1} -Versuchs, der durch die Zuordnung $C = -AB$ entsteht (die andere Hälfte von Tabelle 8.7) und die Berechnung der Effekte wie in Tabelle 8.10

| Nr. | Material A | Druck B | Werkzeug C | AB | AC | BC | ABC | y_i |
|----------|------------|---------|------------|--------|--------|-------|--------|-------|
| 1 | - | - | - | + | + | + | - | 4,18 |
| 4 | + | + | - | + | - | - | - | 4,31 |
| 6 | + | - | + | - | + | - | - | 4,21 |
| 7 | - | + | + | - | - | + | - | 4,41 |
| Σ | -0,07 | 0,33 | 0,13 | -0,13 | -0,33 | 0,07 | -17,11 | |
| Eff. | -0,035 | 0,165 | 0,065 | -0,065 | -0,165 | 0,035 | | |

Wie in Tabelle 8.10 kann nicht mehr zwischen dem Effekt von A und BC unterschieden werden, bei beiden erhält man $A - BC$. In diesem Beispiel ist die Wechselwirkung $BC = 0$, daher wird A nicht verändert. Für BC erhält man jedoch 0,035 statt 0.

Statt des Effekts von C erhält man $C - AB$, also $0,115 - 0,050 = 0,065$. Hat man nur dieses Ergebnis vorliegen, kann man nicht feststellen, wie groß C und AB einzeln sind. Stellt man sich auf den Standpunkt, dass die Wechselwirkung vernachlässigbar ist, so meint man $C = 0,065$.

Addiert bzw. subtrahiert man jedoch die Ergebnisse aus Tabelle 8.10 und 8.11, so erhält man in Übereinstimmung mit Tabelle 8.8 z. B.:

$$(C + AB) + (C - AB) = 2 \cdot C = 0,165 + 0,065 = 0,230$$

$$(C + AB) - (C - AB) = 2 \cdot AB = 0,165 - 0,065 = 0,100.$$

Einen fraktionellen faktoriellen Versuchsplan kann man auch nachträglich zu einem vollständigen faktoriellen Versuchsplan ergänzen.

8.2.5 Auflösung

Das Risiko einer Fehlinterpretation aufgrund einer Vermengung hängt davon ab, welche Art von Effekten (Effekte der Faktoren, 2FWW, 3FWW usw.) miteinander vermengt sind.

Aus dem Generator können alle Vermengungen bestimmt werden. So erkennt man aus dem Generator $I = ABCD$ z. B. sofort, dass der Effekt von A mit BCD und der Effekt von AB mit CD vermengt sind (oder allgemeiner der Effekt eines Faktors mit einer 3FWW und der Effekt einer 2FWW mit einer anderen 2FWW).

Die kleinste Buchstabenanzahl (= Faktorenanzahl) im Generator (das kürzeste „Wort“) bestimmt die ungünstigste Vermengung. Sie wird als „Auflösung“ (auch „Typ“, englisch: resolution) bezeichnet:

- Der 2^{4-1} -Plan in 8.2.2 hat den Generator $I = ABCD$ und somit Auflösung IV.
- Der 2^{5-1} -Plan in Beispiel 1 von 8.2.3 hat den Generator $I = ABCDE$ und somit Auflösung V.
- Der 2^{5-2} -Plan in Beispiel 2 hat den Generator $I = ABCD = ABE = CDE$ und somit Auflösung III (die kleinste Buchstabenanzahl ist drei).
- Die 2^{3-1} -Pläne in 8.2.4 haben den Generator $I = ABC$ bzw. $I = -ABC$ und somit Auflösung III.

Tabelle 8.12 vergleicht und bewertet die verschiedenen Auflösungen.

| Auflösung | Vermengung | Bewertung |
|-----------|---|------------------|
| III | Faktor mit 2FWW | kritisch |
| IV | Faktor mit 3FWW 2FWW mit 2FWW | weniger kritisch |
| V | Faktor mit 4FWW 2FWW mit 3FWW | unkritisch |
| VI | Faktor mit 5FWW 2FWW mit 4FWW 3FWW mit 3FWW | unkritisch |

TABELLE 8.12

Bewertung der Auflösung eines Versuchsplans

Versuche der Auflösung III sind als kritisch zu bewerten, weil 2FWW die Effekte der Faktoren verfälschen können, ohne dass man irgend einen Hinweis auf die Bedeutung der 2FWW erhält (vgl. die 2^{3-1} -Pläne im Beispiel in 8.2.4). Im Extremfall kann eine 2FWW den Effekt eines Faktors umkehren. Versuche der Auflösung III führen nur dann zu brauchbaren Ergebnissen, wenn die Effekte der 2FWW deutlich kleiner sind als die Effekte der Faktoren. Es genügt nicht, dass man an den Wechselwirkungen nicht interessiert ist.

Versuche der Auflösung IV sind wesentlich weniger kritisch, da 2FWW nicht die Effekte der Faktoren verfälschen können. 2FWW sind immer nur mit anderen 2FWW vermengt. Ist ein solcher Effekt signifikant, so weist dies darauf hin, dass mindestens eine von mehreren 2FWW wichtig ist. Mit technischen Überlegungen lässt sich dann meist die wichtige Wechselwirkung identifizieren. Bleiben Zweifel, so kann mit zusätzlichen Versuchen zwischen zwei plausiblen Möglichkeiten unterschieden werden.

Versuche der Auflösung V oder höher sind als unkritisch zu bewerten, da sämtliche 2FWW getrennt voneinander bestimmt werden können und nur mit 3FWW vermengt sind. Diese 3FWW sind selten von praktischer Bedeutung.

8.2.6 Überblick über 2^{k-p} -Pläne

Je größer die Anzahl der Faktoren k und je kleiner die Anzahl der Faktorstufenkombinationen m in einem Versuchsplan ist, desto niedriger ist die Auflösung, die maximal erreichbar ist. Tabelle 8.13 zeigt diese maximale Auflösung für die häufigsten Fälle.

TABELLE 8.13 Maximal erreichbare Auflösung in Abhängigkeit von der Anzahl der Faktoren k und der Anzahl der Faktorstufenkombinationen m

| $\frac{k}{m}$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---------------|------------------------|------------------|------------------------|-------------------------|--------------------------|-------------------|------------------|-------------------|-------------------|-------------------|
| 4 | 2^{3-1} III | | | | | | | | | |
| 8 | 2^3 vollst. IV | 2^{4-1} III | 2^{5-2} V | 2^{6-3} IV | 2^{7-4} III | | | | | |
| 16 | | 2^4 vollst. | 2^{5-1} V | 2^{6-2} IV | 2^{7-3} IV | 2^{8-4} IV | 2^{9-5} III | 2^{10-6} III | 2^{11-7} III | 2^{12-8} III |
| 32 | | | 2^5 vollst. VI | 2^{6-1} IV | 2^{7-2} IV | 2^{8-3} IV | 2^{9-4} IV | 2^{10-5} IV | 2^{11-6} IV | 2^{12-7} IV |
| 64 | | | | 2^6 vollst. VII | 2^{7-1} V | 2^{8-2} V | 2^{9-3} IV | 2^{10-4} IV | 2^{11-5} IV | 2^{12-6} IV |
| 128 | | | | | 2^7 vollst. VIII | 2^{8-1} VIII | 2^{9-2} VI | 2^{10-3} V | 2^{11-4} V | 2^{12-5} IV |

Umgekehrt kann man auch sagen, je niedriger die Auflösung ist, desto mehr Faktoren können mit einer gegebenen Anzahl Faktorstufenkombinationen untersucht werden:

Bei Auflösung III können

- mit 4 Faktorstufenkombinationen 3 Faktoren
- mit 8 Faktorstufenkombinationen bis zu 7 Faktoren
- mit 16 Faktorstufenkombinationen bis zu 15 Faktoren
- mit 32 Faktorstufenkombinationen bis zu 31 Faktoren untersucht werden, usw.

Ein Versuchsplan, der diese Maximalzahl an Faktoren enthält, heißt „gesättigt“, weil nicht mehr Faktoren untersucht werden können. Alle Effekte sind mit Faktoren belegt. Gesättigte Pläne werden häufig von G. Taguchi genutzt (Kapitel 9).

Bei Auflösung IV können

- mit 8 Faktorstufenkombinationen bis zu 4 Faktoren
- mit 16 Faktorstufenkombinationen bis zu 8 Faktoren
- mit 32 Faktorstufenkombinationen bis zu 16 Faktoren untersucht werden.

Somit benötigt man für Auflösung IV höchstens doppelt so viele Faktorstufenkombinationen wie für Auflösung III. Da das Risiko, das mit der Vermengung verbunden ist, bei Plänen der Auflösung IV weniger kritisch ist als bei III (vgl. Tabelle 8.12), wird empfohlen, soweit möglich mindestens Auflösung IV zu verwenden.

Für Auflösung $\geq V$ nimmt die Anzahl der Faktorstufenkombinationen sehr schnell mit der Anzahl der Faktoren zu. Daher können Versuche mit Auflösung $\geq V$ normalerweise nur für bis zu $k = 6$ Faktoren durchgeführt werden ($m = 32$).

Versuche der Auflösung IV (in Ausnahmefällen III) werden eingesetzt, um aus vielen Faktoren die wichtigsten zu erkennen (Screening). Für genauere Untersuchungen (vgl. Kapitel 11) sollte mindestens Auflösung V verwendet werden.

Zuordnungen

Es gibt viele mögliche Zuordnungen, um einen fraktionellen faktoriellen Versuchsplan mit vorgegebenen Werten von k und m zu erzeugen. Allerdings erhält man nicht mit jeder Zuordnung die maximale Auflösung nach Tabelle 8.13. Vor allem bei mehrfach reduzierten Plänen ist die günstigste Zuordnung nicht offensichtlich. Folgende Aufgabe verdeutlicht dies.



Aufgabe

Durch die Zuordnung von zwei Faktoren E und F zu einem vollständigen faktoriellen 2^4 -Plan soll ein fraktioneller faktorieller 2^{6-2} -Plan erzeugt werden. Vergleichen Sie folgende Möglichkeiten, indem Sie sämtliche Generatoren berechnen und so die Auflösung bestimmen. Was erhalten Sie in Wirklichkeit, wenn Sie den Effekt von A berechnen?

- a) $E = ABCD$ und $F = BCD$
- b) $E = ABC$ und $F = BCD$
- c) $E = -ABC$ und $F = -BCD$

Lösung

- a) Die Generatoren $I = ABCDE$ und $I = BCDF$ ergeben als weiteren Generator
 $I \cdot I = ABCDE \cdot BCDF = AEF$ und damit Auflösung III.
 Statt A erhält man $A + BCDE + ABCDF + EF$, d. h. A ist mit der 2FWW EF vermischt.
- b) Die Generatoren $I = ABCE$ und $I = BCDF$ ergeben als weiteren Generator
 $I \cdot I = ABCE \cdot BCDF = ADEF$ und damit Auflösung IV.
 Statt A erhält man $A + BCE + ABCDF + DEF$, d. h. A ist mit den 3FWW BCE und DEF vermischt, aber mit keiner 2FWW.
- c) Die Generatoren $I = -ABCE$ und $I = -BCDF$ ergeben als weiteren Generator
 $I \cdot I = (-ABCE) \cdot (-BCDF) = ADEF$ und damit Auflösung IV.
 Statt A erhält man $A - BCE - ABCDF + DEF$.

Mit den Zuordnungen b und c erhält man Auflösung IV (wie in Tabelle 8.13 angegeben), mit a nur Auflösung III.

Tabelle 8.14 zeigt Zuordnungen, mit denen für vorgegebene Werte von k und m fraktionelle faktorielle Versuchspläne mit der jeweils maximal erreichbaren Auflösung erzeugt werden. Gibt es mehrere mögliche Zuordnungen mit maximaler Auflösung, so wird diejenige mit der kleinstmöglichen Anzahl kurzer Generatoren genannt (Minimum Aberration [3]).

TABELLE 8.14 Zuordnungen, die die maximal erreichbare Auflösung nach Tabelle 8.13 ergeben [1]:

k = Anzahl der Faktoren

p = Anzahl der Zuordnungen

m = Anzahl der Faktorstufenkombinationen 2^{k-p}

| k | m | p | Aufl. | Geeignete Zuordnungen | | | |
|----|-----|---|-------|-----------------------|-----------------|----------------|------------------------------|
| 3 | 4 | 1 | III | $C = \pm AB$ | | | |
| 4 | 8 | 1 | IV | $D = \pm ABC$ | | | |
| 5 | 16 | 1 | V | $E = \pm ABCD$ | | | |
| 5 | 8 | 2 | III | $D = \pm AB$ | $E = \pm AC$ | | |
| 6 | 32 | 1 | VI | $F = \pm ABCDE$ | | | |
| 6 | 16 | 2 | IV | $E = \pm ABC$ | $F = \pm BCD$ | | |
| 6 | 8 | 3 | III | $D = \pm AB$ | $E = \pm AC$ | $F = \pm BC$ | |
| 7 | 64 | 1 | VII | $G = \pm ABCDEF$ | | | |
| 7 | 32 | 2 | IV | $F = \pm ABCD$ | $G = \pm ABDE$ | | |
| 7 | 16 | 3 | IV | $E = \pm ABC$ | $F = \pm BCD$ | $G = \pm ACD$ | |
| 7 | 8 | 4 | III | $D = \pm AB$ | $E = \pm AC$ | $F = \pm BC$ | $G = \pm ABC$ |
| 8 | 128 | 1 | VIII | $H = \pm ABCDEFG$ | | | |
| 8 | 64 | 2 | V | $G = \pm ABCD$ | $H = \pm ABEF$ | | |
| 8 | 32 | 3 | IV | $F = \pm ABC$ | $G = \pm ABD$ | $H = \pm BCDE$ | |
| 8 | 16 | 4 | IV | $E = \pm ABC$ | $F = \pm ABD$ | $G = \pm ACD$ | $H = \pm BCD$ |
| 9 | 128 | 2 | VI | $H = \pm ACDFG$ | $J = \pm BCEFG$ | | |
| 9 | 64 | 3 | IV | $G = \pm ABCD$ | $H = \pm ACEF$ | $J = \pm CDEF$ | |
| 9 | 32 | 4 | IV | $F = \pm BCDE$ | $G = \pm ACDE$ | $H = \pm ABDE$ | $J = \pm ABCE$ |
| 9 | 16 | 5 | III | $E = \pm ABC$ | $F = \pm BCD$ | $G = \pm ACD$ | $H = \pm ABD$ $J = \pm ABCD$ |
| 10 | 128 | 3 | V | $H = \pm ABCG$ | $J = \pm BCDE$ | $K = \pm ACDF$ | |
| 11 | 128 | 4 | V | $H = \pm ABCG$ | $J = \pm BCDE$ | $K = \pm ACDF$ | $L = \pm ABCDEFG$ |

8.2.7 Praxisbeispiel Reflowlöten

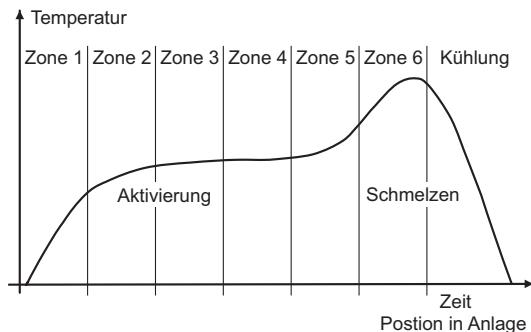
Das folgende Praxisbeispiel wurde im Jahre 1993 in einer Zusammenarbeit zwischen dem Hersteller und einem Anwender einer Reflowlötanlage durchgeführt und hatte wesentlichen Einfluss auf die weitere Entwicklung dieses Anlagentyps. Erstmals wurde ein neues Anlagenkonzept systematisch erprobt. Diese Erprobung war so erfolgreich, dass zwischenzeitlich fast nur noch Anlagen gebaut werden, die auf dem neuen Konzept basieren.

Im Folgenden werden die Einzelschritte der Vorgehensweise aus Kapitel 3 nacheinander behandelt.

1. Ausgangssituation beschreiben

Zum Löten von SMD-Bauelementen (Surface Mounted Devices) auf Leiterplatten wird zunächst Lotpaste aufgetragen. Auf diese Lotpaste werden die Bauelemente platziert (bestückt). In der Reflowlötanlage wird die bestückte Leiterplatte dann erhitzt. Das Lot schmilzt, und die Bauelemente werden dadurch festgelötet.

Der Lötprozess ist kontinuierlich. Die Leiterplatten durchlaufen auf einem Transportsystem verschiedene Temperaturzonen in der Anlage, so dass die Temperatur der Leiterplatte den in Bild 8-2 dargestellten Zeitverlauf hat. Die Leiterplatte wird zunächst auf eine Temperatur von ca. 150 °C erwärmt, um die Lotpaste zu aktivieren. Dadurch wird die Oxidhaut aufgebrochen. Dann wird die Temperatur kurz auf über 200 °C erhöht, um die Lotpaste aufzuschmelzen (Schmelzpunkt ca. 180 °C) und anschließend schnell wieder abgekühlt.

**BILD 8-2**

Temperatur auf der Leiterplatte beim Durchlaufen der Reflowölanlage (Prinzipbild)

Je nach Fabrikat der Lotpaste (und normalerweise auch Bestückung der Leiterplatte) benötigt man etwas unterschiedliche Temperaturverläufe (Profile), um ein optimales Lötgebnis zu erzielen.

Bei Infrarotanlagen erfolgt die Aufheizung durch geregelte Infrarotstrahler. Ein Problem dieser Anlagen ist, dass die Leiterplatte und die Bauelemente deutlich wärmer werden als die Lotpaste. Daher hat man nur ein sehr enges Temperaturfenster zur Verfügung, in dem die Lotpaste überall aufschmilzt, die Leiterplatte jedoch noch nicht geschädigt wird. Je nach Größe der Leiterplatte und Dichte der Bestückung benötigt man ein anderes Profil. Die Änderung der Profile ist mit großem Aufwand verbunden und birgt ein Fehlerrisiko.

Je kleiner der Temperaturunterschied zwischen Leiterplatte und Lotpaste ist, desto größer ist das Fenster. Im Idealfall kann man alle Leiterplattentypen mit einem einzigen Profil löten.

Gegenstand der Untersuchung war eine neue Anlage, bei der in der Schmelzone (Zone 6) die Erwärmung der Leiterplatte durch Anblasen mit Heißluft erfolgte (Konvektionsanlage). Davon erhoffte man sich eine Erniedrigung der Temperaturunterschiede.

2. Untersuchungsziel festlegen

Ziel der Untersuchung war es, das Verhalten dieser völlig neuen Anlage mit möglichst wenig Aufwand quantitativ zu erfassen. Dazu sollte der Temperaturverlauf in Abhängigkeit von den relevanten Anlagenparametern analysiert werden. Insbesondere sollte untersucht werden,

- wie die Temperaturunterschiede zwischen Leiterplatte und Lotpaste minimiert werden können (zur Maximierung des Prozessfensters, Langzeitziel war ein gemeinsames Profil für alle Leiterplattentypen) und

- wie das Temperaturprofil quantitativ von den Anlagenparametern abhängt (um bei evtl. auftretenden Problemen gezielt reagieren zu können und die Besonderheiten verschiedener Lotpasten berücksichtigen zu können).

Die Zielgrößen sind daher Mittelwerte, die Auswertung erfolgt mit klassischen Methoden.¹

3. Zielgrößen und Faktoren festlegen

Für die Untersuchung wurde ein besonders problematischer Leiterplattentyp (mit Gebieten sehr dichter Bestückung und damit besonders niedrigen Temperaturen an den Lötstellen in diesen Gebieten) ausgewählt und eine Leiterplatte als Testplatte mit Thermoelementen an folgenden kritischen Stellen bestückt:

- an der Lotpaste in einem Gebiet niedriger Temperatur T_L
- auf der Leiterplatte in einem Gebiet hoher Temperatur T_P
- in einem Bauelement mit besonders hoher Temperatur T_B
- an einer heißen Stelle auf der Unterseite T_U .

Die 4 Temperaturen wurden jeweils am Ende der 5 Vorheizzonen und im Maximum gemessen. Außerdem wurde die Zeit erfasst, für die die Löttemperatur überschritten ist. D.h. pro Einzelversuch wurden 28 Messwerte erfasst. Alle Einzelversuche wurden mit derselben Testplatte durchgeführt. Dadurch entfiel zwar die Aktivierung, das Lot ist aber bei jedem Einzelversuch erneut geschmolzen.

Jede dieser 28 Größen kann als Zielgröße verwendet werden, aber auch daraus abgeleitete Größen sind möglich. Für die Minimierung der Temperaturdifferenz ist keiner der Messwerte direkt geeignet. Nach Durchführung der Versuche war jedoch klar, dass von den 4 Temperaturen die Temperatur an der Lötstelle immer die niedrigste und die Temperatur auf der Leiterplatte immer die höchste war.

Als Zielgröße bietet sich daher die Temperaturdifferenz $\Delta = T_P - T_L$ an. Statt der Differenz wird im Folgenden das Verhältnis $V = \Delta / (T_L - 150)$ verwendet. Dahinter steckt die Überlegung, dass die Differenz natürlich umso höher ist, je stärker die Erwärmung in der Zone 6 ist (150 °C ist die Ausgangstemperatur im Plateau vor der Schmelzzone). Ein minimaler Wert von V bedeutet, dass eine vorgegebene Erwärmung in der Schmelzzone mit minimaler Temperaturdifferenz erreicht wird.

V ist nicht die einzige mögliche Zielgröße, Alternativen wie z. B. Δ sind denkbar. Wichtig ist nur, dass die verwendete Zielgröße das Untersuchungsziel abbildet. Die wesentlichen Schlussfolgerungen sind weitgehend unabhängig von der genauen Definition der Zielgröße.

¹ Für die Minimierung der Temperaturunterschiede hätte man auch Ideen von G. Taguchi verwenden können (Kapitel 9). Die Grenzen zwischen den Verfahren sind fließend.

Der Einfluss folgender Faktoren auf die 28 Messwerte sollte untersucht werden:

| | | | | |
|---|--|-------|-------|-------|
| A | Transportgeschwindigkeit | 0,7 | 0,9 | m/min |
| B | (Soll-)Temperatur in Zone 1 | 170 | 190 | °C |
| C | (Soll-)Temperatur in Zone 2 | 160 | 180 | °C |
| D | (Soll-)Temperatur in Zone 3 | 140 | 160 | °C |
| E | (Soll-)Temperatur in Zone 4 | 130 | 150 | °C |
| F | (Soll-)Temperatur in Zone 5 | 160 | 180 | °C |
| G | (Soll-)Temperatur in Zone 6 | 225 | 245 | °C |
| H | Heizung auch von unten in Zone 2 bis 5 | aus | ein | |
| J | Lüfter oben in Zone 6 | 60 | 100 % | |
| K | Lüfter unten in Zone 6 | 60 | 100 % | |
| L | Heizung auch von unten in Zone 6 | aus | ein | |
| M | Messgerät | Nr. 1 | Nr. 2 | |

Um die Untersuchung schneller durchführen zu können, sollten abwechselnd zwei verschiedene Messgeräte eingesetzt werden. Diese wurden als zusätzlicher Faktor M in den Versuchsplan aufgenommen. Die 12 Faktoren A–M wurden auf je zwei Stufen untersucht (die Festlegung der Stufenwerte erfolgte aufgrund der Erfahrung mit früheren Anlagen, ergänzt um einige grobe Vorversuche).

4. Versuchsplan aufstellen

Zunächst wurde der Versuchsumfang festgelegt. Bei gleichen Bedingungen streuten Messwerte für die Temperaturen erfahrungsgemäß mit einer Standardabweichung von ca. 2 °C. Effekte von ca. 3 °C sollten erkannt werden. Daraus ergibt sich gemäß (7.15) ein Versuchs-umfang von ca. 27.

Andererseits benötigt man für einen Versuchsplan der Auflösung IV mit 12 Faktoren 32 Faktorstufenkombinationen (vgl. Tabelle 8.13 – $k = 12$ Faktoren in $m = 2^{12-7} = 32$ Kombinationen). $m = 32$ ist bereits mehr als die errechnete Anzahl von 27, daher genügt eine Realisierung ($n = 1$). Zur Abschätzung der Zufallsstreuung wurden vier der Faktorstufen-kombinationen wiederholt.

Somit wurden insgesamt 36 Einzelversuche durchgeführt, d.h. die Testleiterplatte lief insgesamt 36-mal durch die Anlage. Die Durchlaufzeit der Leiterplatte durch die Anlage betrug nur wenige Minuten. Zwischen den Durchläufen musste die Leiterplatte aber wieder gekühlt werden, und die Anlage brauchte Zeit, bis sich die veränderten Bedingungen eingestellt hatten. Daher betrug die Gesamtzeit zur Durchführung eines Einzelversuchs ca. 15 min. Für den gesamten Versuch wurde daher ein Arbeitstag veranschlagt. Der damit verbundene Aufwand erschien angemessen.

Die Einzelversuche wurden in randomisierter Reihenfolge durchgeführt, wobei abwechselnd die Messgeräte 1 und 2 verwendet wurden. Auf Blockbildung wurde verzichtet, da alle Einzelversuche unmittelbar hintereinander durchgeführt werden konnten. Tabelle 8.15 zeigt den Versuchsplan.

TABELLE 8.15 Faktorstufenkombinationen in der Reihenfolge der Versuchsdurchführung

| Vers-Nr. | syst-Nr. | Transp.-geschw. [m/min] | Soll-temp. Zone 1 [°C] | Soll-temp. Zone 2 [°C] | Soll-temp. Zone 3 [°C] | Soll-temp. Zone 4 [°C] | Soll-temp. Zone 5 [°C] | Soll-temp. Zone 6 [°C] | Heizg-unten Zone 2–5 | Lüfter-oben Zone 6 [%] | Lüfter-unten Zone 6 [%] | Heizg-unten Zone 6 | Mess-gerät-Nr. |
|----------|----------|----------------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|-------------------------|---------------------------|----------------------------|-----------------------|----------------|
| 1 | 24 | 0,9 | 170 | 180 | 160 | 150 | 160 | 225 | aus | 100 | 100 | ein | Nr. 1 |
| 2 | 31 | 0,9 | 190 | 180 | 160 | 130 | 180 | 225 | ein | 100 | 60 | aus | Nr. 2 |
| 3 | 23 | 0,9 | 170 | 180 | 160 | 130 | 160 | 245 | aus | 100 | 60 | aus | Nr. 1 |
| 4 | 3 | 0,7 | 170 | 160 | 160 | 130 | 160 | 225 | ein | 100 | 60 | ein | Nr. 2 |
| 5 | 28 | 0,9 | 190 | 160 | 160 | 150 | 160 | 245 | ein | 60 | 60 | ein | Nr. 1 |
| 6 | 19 | 0,9 | 170 | 160 | 160 | 130 | 180 | 245 | aus | 60 | 100 | aus | Nr. 2 |
| 7 | 8 | 0,7 | 170 | 180 | 160 | 150 | 180 | 245 | ein | 60 | 60 | aus | Nr. 1 |
| 8 | 22 | 0,9 | 170 | 180 | 140 | 150 | 160 | 225 | ein | 60 | 100 | aus | Nr. 2 |
| 9 | 13 | 0,7 | 190 | 180 | 140 | 130 | 160 | 245 | ein | 100 | 100 | aus | Nr. 1 |
| 10 | 10 | 0,7 | 190 | 160 | 140 | 150 | 180 | 225 | ein | 60 | 100 | ein | Nr. 2 |
| 11 | 30 | 0,9 | 190 | 180 | 140 | 150 | 180 | 245 | aus | 60 | 100 | aus | Nr. 1 |
| 12 | 25 | 0,9 | 190 | 160 | 140 | 130 | 160 | 225 | aus | 100 | 100 | ein | Nr. 2 |
| 13 | 29 | 0,9 | 190 | 180 | 140 | 130 | 180 | 225 | aus | 60 | 60 | ein | Nr. 1 |
| 14 | 19 | 0,9 | 170 | 160 | 160 | 130 | 180 | 245 | aus | 60 | 100 | aus | Nr. 2 |
| 15 | 7 | 0,7 | 170 | 180 | 160 | 130 | 180 | 225 | ein | 60 | 100 | ein | Nr. 1 |
| 16 | 9 | 0,7 | 190 | 160 | 140 | 130 | 180 | 245 | ein | 60 | 60 | aus | Nr. 2 |
| 17 | 1 | 0,7 | 170 | 160 | 140 | 130 | 160 | 225 | aus | 60 | 60 | aus | Nr. 1 |
| 18 | 32 | 0,9 | 190 | 180 | 160 | 150 | 180 | 245 | ein | 100 | 100 | ein | Nr. 2 |
| 19 | 14 | 0,7 | 190 | 180 | 140 | 150 | 160 | 225 | ein | 100 | 60 | ein | Nr. 1 |
| 20 | 15 | 0,7 | 190 | 180 | 160 | 130 | 160 | 245 | aus | 60 | 100 | ein | Nr. 2 |
| 21 | 27 | 0,9 | 190 | 160 | 160 | 130 | 160 | 225 | ein | 60 | 100 | aus | Nr. 1 |
| 22 | 21 | 0,9 | 170 | 180 | 140 | 130 | 160 | 245 | ein | 60 | 60 | ein | Nr. 2 |
| 23 | 17 | 0,9 | 170 | 160 | 140 | 130 | 180 | 245 | ein | 100 | 100 | ein | Nr. 1 |
| 24 | 16 | 0,7 | 190 | 180 | 160 | 150 | 160 | 225 | aus | 60 | 60 | aus | Nr. 2 |
| 25 | 18 | 0,9 | 170 | 160 | 140 | 150 | 180 | 225 | ein | 100 | 60 | aus | Nr. 1 |
| 26 | 6 | 0,7 | 170 | 180 | 140 | 150 | 180 | 245 | aus | 100 | 60 | ein | Nr. 2 |
| 27 | 8 | 0,7 | 170 | 180 | 160 | 150 | 180 | 245 | ein | 60 | 60 | aus | Nr. 1 |
| 28 | 25 | 0,9 | 190 | 160 | 140 | 130 | 160 | 225 | aus | 100 | 100 | ein | Nr. 2 |
| 29 | 11 | 0,7 | 190 | 160 | 160 | 130 | 180 | 245 | aus | 100 | 60 | ein | Nr. 1 |
| 30 | 26 | 0,9 | 190 | 160 | 140 | 150 | 160 | 245 | aus | 100 | 60 | aus | Nr. 2 |
| 31 | 1 | 0,7 | 170 | 160 | 140 | 130 | 160 | 225 | aus | 60 | 60 | aus | Nr. 1 |
| 32 | 5 | 0,7 | 170 | 180 | 140 | 130 | 180 | 225 | aus | 100 | 100 | aus | Nr. 2 |
| 33 | 2 | 0,7 | 170 | 160 | 140 | 150 | 160 | 245 | aus | 60 | 100 | ein | Nr. 1 |
| 34 | 4 | 0,7 | 170 | 160 | 160 | 150 | 160 | 245 | ein | 100 | 100 | aus | Nr. 2 |
| 35 | 12 | 0,7 | 190 | 160 | 160 | 150 | 180 | 225 | aus | 100 | 100 | aus | Nr. 1 |
| 36 | 20 | 0,9 | 170 | 160 | 160 | 150 | 180 | 225 | aus | 60 | 60 | ein | Nr. 2 |

5. Versuche durchführen

Die Umgebungsbedingungen wurden möglichst konstant gehalten (und zur Absicherung erfasst). Alle Einzelversuche wurden vom gleichen Zwei-Mann-Team an einem Tag durchgeführt. Tabelle 8.16 zeigt die Messwerte für 8 der insgesamt 28 erfassten Größen und die berechneten Werte für V.

TABELLE 8.16 Ergebnisse für ausgewählte Messgrößen und die berechnete Zielgröße V, in der Reihenfolge der Versuchsdurchführung

| Vers. Nr. | syst. Nr. | T _p Zone 1 [°C] | T _p Zone2 [°C] | T _L Zone2 [°C] | T _p Zone3 [°C] | T _p Zone4 [°C] | T _p Zone5 [°C] | T _p max [°C] | T _L max [°C] | V |
|--------------|--------------|----------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-------------------------------|-------------------------------|-------|
| 1 | 24 | 96 | 141 | 113 | 154 | 154 | 170 | 211 | 197 | 0,298 |
| 2 | 31 | 109 | 146 | 117 | 152 | 148 | 174 | 207 | 191 | 0,390 |
| 3 | 23 | 98 | 143 | 113 | 152 | 146 | 162 | 215 | 199 | 0,327 |
| 4 | 3 | 109 | 141 | 117 | 154 | 145 | 164 | 209 | 199 | 0,204 |
| 5 | 28 | 115 | 141 | 111 | 156 | 154 | 164 | 213 | 193 | 0,465 |
| 6 | 19 | 100 | 131 | 105 | 150 | 148 | 178 | 203 | 189 | 0,359 |
| 7 | 8 | 113 | 154 | 129 | 164 | 158 | 184 | 217 | 203 | 0,264 |
| 8 | 22 | 98 | 141 | 109 | 145 | 150 | 158 | 188 | 176 | 0,462 |
| 9 | 13 | 127 | 162 | 135 | 146 | 143 | 168 | 223 | 211 | 0,197 |
| 10 | 10 | 123 | 148 | 125 | 145 | 156 | 178 | 203 | 195 | 0,178 |
| 11 | 30 | 109 | 150 | 119 | 146 | 156 | 182 | 209 | 193 | 0,372 |
| 12 | 25 | 105 | 135 | 107 | 135 | 137 | 160 | 207 | 189 | 0,462 |
| 13 | 29 | 111 | 150 | 121 | 145 | 143 | 180 | 205 | 189 | 0,410 |
| 14 | 19 | 98 | 103 | 105 | 150 | 146 | 176 | 205 | 191 | 0,341 |
| 15 | 7 | 115 | 154 | 129 | 162 | 154 | 180 | 209 | 199 | 0,204 |
| 16 | 9 | 125 | 148 | 123 | 145 | 141 | 172 | 211 | 199 | 0,245 |
| 17 | 1 | 115 | 146 | 127 | 145 | 141 | 164 | 199 | 188 | 0,289 |
| 18 | 32 | 109 | 148 | 119 | 154 | 156 | 184 | 227 | 209 | 0,305 |
| 19 | 14 | 125 | 162 | 137 | 152 | 156 | 172 | 217 | 203 | 0,264 |
| 20 | 15 | 125 | 160 | 133 | 162 | 152 | 164 | 213 | 203 | 0,189 |
| 21 | 27 | 113 | 141 | 115 | 154 | 148 | 160 | 193 | 180 | 0,433 |
| 22 | 21 | 100 | 141 | 113 | 143 | 139 | 156 | 207 | 189 | 0,462 |
| 23 | 17 | 96 | 129 | 105 | 135 | 139 | 176 | 227 | 207 | 0,351 |
| 24 | 16 | 125 | 160 | 135 | 162 | 156 | 168 | 201 | 191 | 0,244 |
| 25 | 18 | 100 | 133 | 109 | 141 | 150 | 178 | 209 | 193 | 0,372 |
| 26 | 6 | 109 | 154 | 131 | 145 | 154 | 188 | 229 | 213 | 0,254 |
| 27 | 8 | 117 | 158 | 135 | 166 | 162 | 186 | 223 | 209 | 0,237 |
| 28 | 25 | 104 | 135 | 109 | 137 | 137 | 160 | 207 | 191 | 0,390 |
| 29 | 11 | 123 | 150 | 129 | 162 | 150 | 188 | 234 | 219 | 0,217 |
| 30 | 26 | 107 | 137 | 113 | 141 | 150 | 164 | 215 | 199 | 0,327 |
| 31 | 1 | 109 | 143 | 121 | 141 | 141 | 168 | 199 | 188 | 0,289 |

| Vers. Nr. | syst. Nr. | T_p Zone1 [°C] | T_p Zone2 [°C] | T_L Zone2 [°C] | T_p Zone3 [°C] | T_p Zone4 [°C] | T_p Zone5 [°C] | T_p max [°C] | T_L max [°C] | V |
|--------------|--------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|----------------------|----------------------|-------|
| 32 | 5 | 115 | 156 | 131 | 143 | 139 | 180 | 213 | 203 | 0,189 |
| 33 | 2 | 113 | 145 | 123 | 145 | 156 | 168 | 217 | 205 | 0,218 |
| 34 | 4 | 111 | 143 | 121 | 156 | 152 | 170 | 221 | 209 | 0,203 |
| 35 | 12 | 125 | 150 | 129 | 160 | 160 | 186 | 213 | 203 | 0,189 |
| 36 | 20 | 96 | 131 | 107 | 152 | 152 | 184 | 205 | 189 | 0,410 |

6. Versuchsergebnisse auswerten

Der typische Ablauf einer Auswertung wird dargestellt.

- Suche nach Ausreißern

Da die 4 Messstellen auf der Leiterplatte in derselben Zone jeweils ein ähnliches Verhalten zeigen müssen, kann man durch paarweises Auftragen der Messergebnisse an zwei solchen Messstellen gegeneinander Ausreißer in den Daten gut erkennen. Bild 8-3 zeigt ein Beispiel mit einem offensichtlichen Ausreißer (der Wert Nr. 14 der Temperatur auf der Platte in Zone 2 (103) ist ein Tippfehler und sollte 130 sein). Diese Art von Fehlern wird zunächst korrigiert.

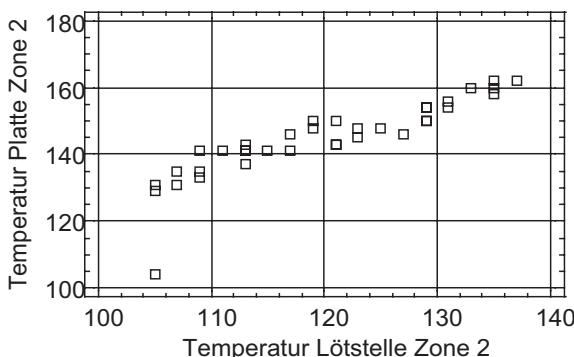


BILD 8-3

Ausreißer können z. B. erkannt werden, wenn zwei Werte gegeneinander aufgetragen werden, die einen ähnlichen Verlauf haben sollten.

- Berechnung der Effekte und ihrer Signifikanz

Für jede Zielgröße einzeln können mit (7.11 ff) die Effekte berechnet und in ihrer Signifikanz beurteilt werden.

Dabei müssen nicht die Messwerte selbst als Zielgrößen verwendet werden. Man sollte vielmehr darauf achten, dass die Zielgröße möglichst gut das Ziel des Versuchs ausdrückt. Hier wird als Beispiel das Verhältnis $V = \Delta / (T_L - 150)$ als Zielgröße verwendet. Die Rechnung ist nicht schwierig, aber für die praktische Durchführung von Hand zu aufwändig. Mit Software erhält man die gesuchten Ergebnisse jedoch sehr einfach.

In Bild 8-4 sind nur die Effekte von A (Geschwindigkeit), J (Lüfter oben in Zone 6) und K (Lüfter unten in Zone 6) und der Wechselwirkungseffekt 8 größer als die Breite des 95 %-Vertrauensbereichs. Nur diese werden daher weiter betrachtet.

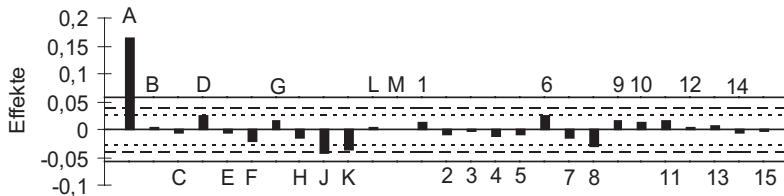


BILD 8-4 Effekte der Faktoren und der 2FWW für die Zielgröße V (Verhältnis) verglichen mit den 95%--, 99%- und 99,9%-Vertrauensbereichen

A-M: Effekte der Faktoren (vermengt mit 3FWW und höheren WW)
1-15: Summen von mehreren 2FWW (vermengt mit höheren WW)

Jeder Effekt ist $2^7 = 128$ fach vermengt. Da ein Versuchsplan der Auflösung IV verwendet wurde, sind die Effekte der Faktoren jedoch nur mit 3FWW und höheren WW vermengt. Sie werden als reine Effekte der Faktoren interpretiert.

Der Effekt 8 ist die Summe der 2FWW AJ + BM + CD + FG + KL und höheren WW. Statistik hilft nicht bei der Entscheidung, welche dieser Wechselwirkungen wichtig ist, dazu ist technisches Wissen erforderlich. In diesem Beispiel haben die Faktoren A und J signifikante Effekte, daher ist naheliegend, dass 8 im Wesentlichen die Wechselwirkung AJ ist.

Eine Wechselwirkung AJ bedeutet, dass die Faktoren A (Geschwindigkeit) und J (Lüfter oben) gemeinsam betrachtet werden müssen. Bild 8-5 stellt die Zielgröße V in Abhängigkeit von diesen Faktoren dar (Wechselwirkung = Abweichung von der Parallelität). Die WW ist technisch plausibel: Sie drückt nur aus, dass sich der Lüfter bei großen Temperaturunterschieden (großem V) stärker auswirkt als bei kleinen. Dies unterstützt die Interpretation der Wechselwirkung 8 als AJ.

Da man ein möglichst kleines Verhältnis V (und damit einen kleinen Unterschied zwischen der Temperatur auf der Platte und an der Lötstelle) anstrebt, sind eine hohe Lüfterdrehzahl oben und eine niedrige Geschwindigkeit günstig.

Der negative Effekt des Faktors K (Lüfter unten) bedeutet, dass eine hohe Lüfterdrehzahl unten die Temperaturunterschiede reduziert (Bild 8-5).

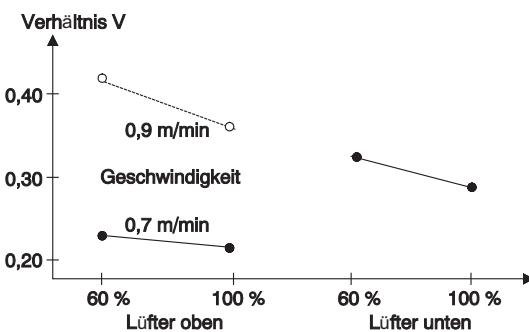


BILD 8-5

Grafische Darstellung der signifikanten Effekte von Faktoren und 2FWW auf die Zielgröße Verhältnis. Sie soll minimiert werden. Die beste Faktorstufenkombination ist daher:
Geschwindigkeit 0,7 m/min
Lüfter oben 100%
Lüfter unten 100%

Insgesamt ist das Ergebnis in Bild 8-5 technisch plausibel:

Eine höhere Lüfterdrehzahl oben und unten bewirkt einen besseren Temperaturausgleich und damit geringere Temperaturdifferenzen. Eine geringere Transportgeschwindigkeit

bedeutet eine längere Verweilzeit in der Schmelzzone und damit ebenfalls kleinere Temperaturdifferenzen.

Zur Absicherung der Ergebnisse wurden die Residuen wie in Kapitel 6 analysiert. Folgende Darstellungen werden empfohlen:

- Residuen gegen mit dem angepassten Modell berechnete Werte:
Verändert sich die Streuung? Vor allem: Nimmt die Streuung mit dem berechneten Wert zu („Trichterform“, vgl. Bild 10-6)?
Wenn ja: Transformation der Zielgröße (Abschnitt 6.4).
- Residuen gegen Stufen der Einflussgrößen:
Verändert sich die Streuung?
Wenn ja: Möglichkeit zur Verringerung der Streuung (Kapitel 9).
Bei mehrstufigen Plänen (Kapitel 11): Abweichungen von der Linearität?
- Residuen gegen Versuchsnummer:
Gibt es einen Trend oder Sprung in den Residuen?
Wenn ja: Ursache identifizieren und ggf. die Messwerte korrigieren.
- Residuen im Wahrscheinlichkeitsnetz:
Gibt es Abweichungen von der Normalverteilung?
Selten erkennbar; aber wenn: Transformation (Abschnitt 6.4).
Gibt es Ausreißer?
Wenn ja: Ursache identifizieren, dann Einzelversuch wiederholen, ggf. Analyse ohne Ausreißer wiederholen, um zu erkennen, ob der Ausreißer die Ergebnisse beeinflusst; technische Interpretation.

Bild 8-6 zeigt als Beispiel die Darstellung der Residuen gegen den berechneten Wert. Sie liegen in einem waagrechten Band, keine Trichterform ist erkennbar. Daher ist auch keine Transformation der Zielgröße V erforderlich. Der eine deutlich größere Wert (er stammt von Versuchsnummer 12) könnte ein Ausreißer sein. Eine Wiederholung der Analyse ohne diesen Wert führt zu keiner Veränderung der Schlussfolgerungen – bei einer Änderung hätte eine technisch-inhaltliche Analyse folgen müssen.

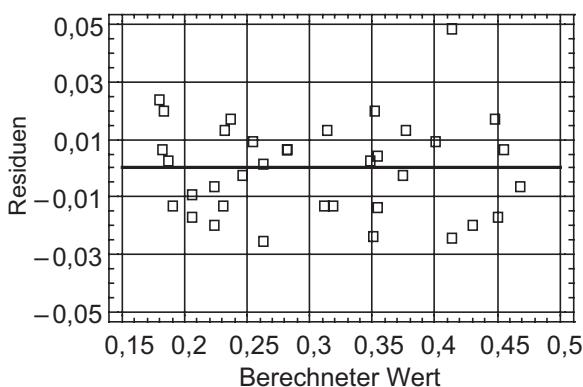


BILD 8-6

Darstellung der Residuen gegen die mit dem angepassten Modell berechneten Werte

Die hier dargestellte Auswertung wird für jede der Zielgrößen durchgeführt. Aus den entsprechenden Auswertungen für die Temperaturen erhält man eine genaue Beschreibung

der Abhängigkeit der Temperatur auf der Leiterplatte von den eingestellten Anlagenparametern (an verschiedenen Stellen in der Anlage – hier nicht dargestellt).

7. Ergebnisse interpretieren und Maßnahmen ableiten

Hauptziel der Untersuchung war es, bei vorgegebener Maximaltemperatur und ausreichend langer Lötzeit möglichst geringe Temperaturunterschiede zwischen der Lötstelle und der Leiterplatte zu erreichen.

- Mit der Einstellung in Bild 8-5 konnten erstmalig **alle Leiterplattentypen mit einem einzigen Profil** gelötet werden.
- Aus den Messwerten der Temperaturen in Tabelle 8.16 wurde zusätzlich die quantitative Abhängigkeit der Temperatur auf der Leiterplatte von den Anlagenparametern ermittelt. Dadurch wurde es möglich, beim folgenden Einsatz der Anlage gezielt auf Änderungen des Anlagenzustandes zu reagieren und die Temperaturprofile auf verschiedene Lotpasten anzupassen.
- Mit Hilfe dieser Ergebnisse konnte die neue Anlage deutlich schneller als bei früheren Anlagen produktiv eingesetzt werden.

Die Breite des 95 %-Vertrauensbereichs für die errechneten Effekte auf die Temperaturen beträgt 2 bis 3 °C. Diese genaue Ermittlung der Effekte ist auf die Mehrfachnutzung der Einzelergebnisse zurückzuführen, die durch die Systematik des Versuchsplans ermöglicht wird.

8. Absicherung, Dokumentation, weiteres Vorgehen

Dies war der erste und bereits sehr erfolgreiche Einsatz einer Konvektionsanlage zum Reflowlöten. Aufgrund der positiven Ergebnisse wurde Konvektion bei Nachfolgeanlagen auch in den anderen Heizzonen eingesetzt, und der Lüfterdurchsatz wurde weiter erhöht.

Bei zu hohem Lüfterdurchsatz ergaben sich natürlich andere Probleme, wie z.B. weggeblasene Bauelemente. Aber aufgrund der Systematik bei der Vorgehensweise mit Versuchsplanung und der statistischen Absicherung der Ergebnisse konnten die Anlagen und die Lötprozesse sehr schnell weiterentwickelt werden. Inzwischen werden zum Reflowlöten fast nur noch Konvektionsanlagen eingesetzt.

■ 8.3 Plackett-Burman-Versuchspläne

Plackett-Burman-Versuchspläne [3] stellen eine Alternative zu den fraktionellen faktoriellen Versuchsplänen dar. Die am weitesten verbreiteten Plackett-Burman-Versuchspläne haben Auflösung III. Wesentliche Unterschiede zu den fraktionellen faktoriellen Versuchsplänen sind:

- Es gibt Plackett-Burman-Pläne auch mit 12, 20, 24, 28 usw. Faktorstufenkombinationen.
- 2FWW sind auf die Effekte von mehreren Faktoren verteilt.

Die Auswertung von Plackett-Burman-Plänen erfolgt wie bei den fraktionellen faktoriellen Plänen.

Durch eine geeignete Verdoppelung der Anzahl Faktorstufenkombinationen kann man Plackett-Burman-Versuchspläne der Auflösung IV erzeugen.

8.3.1 Plackett-Burman-Versuchspläne der Auflösung III

Plackett-Burman-Pläne (PB-Pläne) gibt es für $m = 4, 8, 12, 16, 20, 24$ usw. Faktorstufenkombinationen. Sie erlauben daher eine feine Abstimmung der Anzahl m auf die Anzahl der Faktoren und den vertretbaren Versuchsaufwand. Wirklich neu sind die Pläne mit $m = 12, 20, 24$ usw. Im Folgenden wird beispielhaft der Plan mit $m = 12$ behandelt, die Pläne für $m = 20$ und 24 werden nur angegeben.

Tabelle 8.17 zeigt den PB-Plan für $m = 12$. Theoretisch erlaubt er, die Effekte von bis zu 11 Faktoren zu bestimmen. Es wird jedoch empfohlen, mindestens zwei der Spalten nicht mit Faktoren zu belegen. Sie zeigen dann die Größenordnung der Zufallsstreuung und der 2FWW (wegen Auflösung III).

TABELLE 8.17 PB-Plan zur Untersuchung von bis zu 11 Faktoren in 12 Faktorstufenkombinationen (Empfehlung: max. 9 Spalten für Faktoren verwenden)

| Nr. | A | B | C | D | E | F | G | H | J | K | L |
|-----|---|---|---|---|---|---|---|---|---|---|---|
| 1 | + | - | + | - | - | - | + | + | + | - | + |
| 2 | + | + | - | + | - | - | - | + | + | + | - |
| 3 | - | + | + | - | + | - | - | - | + | + | + |
| 4 | + | - | + | + | - | + | - | - | - | + | + |
| 5 | + | + | - | + | + | - | + | - | - | - | + |
| 6 | + | + | + | - | + | + | - | + | - | - | - |
| 7 | - | + | + | + | - | + | + | - | + | - | - |
| 8 | - | - | + | + | + | - | + | + | - | + | - |
| 9 | - | - | - | + | + | + | - | + | + | - | + |
| 10 | + | - | - | - | + | + | + | - | + | + | - |
| 11 | - | + | - | - | - | + | + | + | - | + | + |
| 12 | - | - | - | - | - | - | - | - | - | - | - |

Wie bei fraktionellen faktoriellen Plänen werden Effekte berechnet, indem man die Versuchsergebnisse mit der jeweiligen Vorzeichenpalte multipliziert, aufaddiert und durch die Anzahl Paare $m/2$ dividiert (hier 6).

Der PB-Plan in Tabelle 8.17 besitzt Auflösung III, d. h. 2FWW sind mit den Effekten der Faktoren vermischt. Während bei fraktionellen faktoriellen Versuchen der Auflösung III jedoch die Vorzeichenpalte einer bestimmten 2FWW identisch zur Vorzeichenpalte eines bestimmten Faktors und damit völlig mit diesem vermischt ist, ist beim PB-Plan z. B. die 2FWW AB „verschmiert“ über alle Faktoren außer A und B. Das folgende Beispiel zeigt dieses „Verschmieren“ anhand konkreter Zahlenwerte.

**Beispiel**

Die Vermengung von 2FWW soll an einem konstruierten Beispiel ohne Zufallsstreuung erläutert werden. Tabelle 8.18 zeigt den Versuchsplan von Tabelle 8.17 mit Versuchsergebnissen y_1 und y_2 . Sie wurden berechnet gemäß:

$$y_1 = 20 + 2 \cdot C \quad \text{d.h. Effekt } C = 22 - 18 = 4, \text{ alle anderen } = 0$$

$$y_2 = 20 + 2 \cdot C + A \cdot B \quad \text{d.h. Effekt } C = 4, \text{ WW AB } = 2, \text{ alle anderen } = 0$$

TABELLE 8.18 Berechnung der Effekte im konstruierten Beispiel:

Σ_1 : Summe aus Spalte y_1 , Effekt C = $24/6 = 4$, andere 0 (richtig)

Σ_2 : Summe aus Spalte y_2 , Effekte C bis L sind um $\pm 4/6 = \pm AB/3$ verfälscht

| Nr. | A | B | C | D | E | F | G | H | J | K | L | y_1 | y_2 |
|------------|---|---|----|---|---|----|----|---|----|----|----|-------|-------|
| 1 | + | - | + | - | - | - | + | + | + | - | + | 22 | 21 |
| 2 | + | + | - | + | - | - | - | + | + | + | - | 18 | 19 |
| 3 | - | + | + | - | + | - | - | - | + | + | + | 22 | 21 |
| 4 | + | - | + | + | - | + | - | - | - | + | + | 22 | 21 |
| 5 | + | + | - | + | + | - | + | - | - | - | + | 18 | 19 |
| 6 | + | + | + | - | + | + | - | + | - | - | - | 22 | 23 |
| 7 | - | + | + | + | - | + | + | - | + | - | - | 22 | 21 |
| 8 | - | - | + | + | + | - | + | + | - | + | - | 22 | 23 |
| 9 | - | - | - | + | + | + | - | + | + | - | + | 18 | 19 |
| 10 | + | - | - | - | + | + | + | - | + | + | - | 18 | 17 |
| 11 | - | + | - | - | - | + | + | + | - | + | + | 18 | 17 |
| 12 | - | - | - | - | - | - | - | - | - | - | - | 18 | 19 |
| Σ_1 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| Σ_2 | 0 | 0 | 20 | 4 | 4 | -4 | -4 | 4 | -4 | -4 | -4 | | |

Tabelle 8.18 zeigt, dass die 2FWW AB die Effekte von 9 Faktoren (alle außer A und B) um $\pm 1/3$ ihres Wertes verfälscht. Sie ist auf 9 Faktoren „verschmiert“.

- Wenn die 2FWW klein bis mittelgroß sind, ist dieses Verschmieren ein Vorteil, weil keiner der Faktoren wesentlich verfälscht wird.
- Ist jedoch eine WW groß, so verfälscht sie fast alle Faktoren, und das Ergebnis des gesamten Versuchs ist unbrauchbar.

Konstruktion von PB-Plänen

Der Versuchsplan in Tabelle 8.17 wurde folgendermaßen erzeugt:

Zeilen 1 – 11: 1. Spalte + + - + + + - - - + - wird zyklisch vertauscht

Letzte Zeile: nur -

Nach demselben Prinzip erhält man PB-Pläne mit $m = 20$ bzw. 24 Faktorstufenkombinationen aus folgenden 1. Spalten:

$m = 20$: 1. Spalte + + - - + + + + - + - - - + + -

$m = 24$: 1. Spalte + + + + + - + - + + - - + + - - + - - - -

Die Signifikanz von Effekten wird wie bei faktoriellen Plänen beurteilt – bei $n > 1$ mit Hilfe der mittleren Varianz s^2 , wie in (7.12), bei $n = 1$ nach Darstellung der Effekte im Wahrscheinlichkeitsnetz oder durch Pooling der „Effekte“ der nicht mit Faktoren belegten Spalten, wie in (7.16). Auch der erforderliche Versuchsumfang $N = nm$ wird wie bei faktoriellen Versuchsplänen ermittelt.

8.3.2 Plackett-Burman-Versuchspläne der Auflösung IV

Aus einem beliebigen Versuchsplan mit Auflösung III erhält man einen Plan mit Auflösung IV, indem man alle Faktorstufenkombinationen mit umgekehrtem Vorzeichen wiederholt (sog. Foldover – die Effekte der Faktoren kehren dabei die Vorzeichen um, die 2FWW nicht, und so werden sie getrennt).

Tabelle 8.19 zeigt als Beispiel den PB-Plan mit Auflösung IV, der durch Foldover aus dem PB-Plan mit Auflösung III von Tabelle 8.17 entsteht.

TABELLE 8.19 PB-Plan der Auflösung IV für bis zu 11 Faktoren:
Nr. 13 – 24 erhält man durch Vorzeichenumkehr aus Nr. 1 – 12 (sog. Foldover).

| Nr. | A | B | C | D | E | F | G | H | J | K | L |
|-----|---|---|---|---|---|---|---|---|---|---|---|
| 1 | + | - | + | - | - | - | + | + | + | - | + |
| 2 | + | + | - | + | - | - | - | + | + | + | - |
| 3 | - | + | + | - | + | - | - | - | + | + | + |
| 4 | + | - | + | + | - | + | - | - | - | + | + |
| 5 | + | + | - | + | + | - | + | - | - | - | + |
| 6 | + | + | + | - | + | + | - | + | - | - | - |
| 7 | - | + | + | + | - | + | + | - | + | - | - |
| 8 | - | - | + | + | + | - | + | + | - | + | - |
| 9 | - | - | - | + | + | + | - | + | + | - | + |
| 10 | + | - | - | - | + | + | + | - | + | + | - |
| 11 | - | + | - | - | - | + | + | + | - | + | + |
| 12 | - | - | - | - | - | - | - | - | - | - | - |
| 13 | - | + | - | + | + | + | - | - | - | + | - |
| 14 | - | - | + | - | + | + | + | - | - | - | + |
| 15 | + | - | - | + | - | + | + | + | - | - | - |
| 16 | - | + | - | - | + | - | + | + | + | - | - |
| 17 | - | - | + | - | - | + | - | + | + | + | - |
| 18 | - | - | - | + | - | - | + | - | + | + | + |
| 19 | + | - | - | - | + | - | - | + | - | + | + |
| 20 | + | + | - | - | - | + | - | - | + | - | + |
| 21 | + | + | + | - | - | - | + | - | - | + | - |
| 22 | - | + | + | + | - | - | - | + | - | - | + |
| 23 | + | - | + | + | + | - | - | - | + | - | - |
| 24 | + | + | + | + | + | + | + | + | + | + | + |

8.3.3 Übersättigte Pläne

Umgekehrt kann man aus den Plackett-Burman-Plänen wie in Tabelle 8.17 die Faktorstufenkombinationen auswählen, in denen z. B. der Faktor L ein „+“ enthält. Man erhält dann den Plan in Tabelle 8.20. Mit diesem Plan können bis zu 10 Faktoren in nur 6 Faktorstufenkombinationen untersucht werden. Man spricht von einem übersättigten Plan, da er mehr Faktoren als Faktorstufenkombinationen enthält. Bei diesen Plänen sind Effekte von Faktoren teilweise miteinander vermengt. Sie erlauben es nur, die größten Effekte zu erkennen.

Vorteil dieser Pläne ist die geringe Anzahl Faktorstufenkombinationen. Nachteil sind die erheblichen Vermengungsrisiken. Es wird empfohlen, diese Pläne möglichst zu vermeiden. Wenn überhaupt, dann ist ihr Einsatz nur für $n = 1$ in Situationen mit extrem hohen Kosten je Einzelversuch vertretbar („Desperado-Pläne“).

TABELLE 8.20 Übersättigter Plan zur Untersuchung von bis zu 10 Faktoren in 6 Faktorstufenkombinationen – die Hälfte des Plans in Tabelle 8.17 (nach Lin [4])

| Nr. | in 8.17 | A | B | C | D | E | F | G | H | J | K |
|-----|---------|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | + | - | + | - | - | - | + | + | + | - |
| 2 | 3 | - | + | + | - | + | - | - | - | + | + |
| 3 | 4 | + | - | + | + | - | + | - | - | - | + |
| 4 | 5 | + | + | - | + | + | - | + | - | - | - |
| 5 | 9 | - | - | - | + | + | + | - | + | + | - |
| 6 | 11 | - | + | - | - | - | + | + | + | - | + |

■ 8.4 Funktionstest

Häufig kann ein Produkt vom Kunden auf viele verschiedene Arten konfiguriert werden bzw. unter vielen verschiedenen Einsatzbedingungen betrieben werden. Verschiedene Konfigurationsmerkmale bzw. Einsatzbedingungen können sich in ihren Auswirkungen gegenseitig beeinflussen und die ungünstigste Kombination ist oft unklar. Andererseits ist der Aufwand für den Test aller denkbaren Kombinationen nicht vertretbar.

Die fraktionellen faktoriellen Versuchspläne von Abschnitt 8.2 bieten eine effiziente Möglichkeit, mit relativ wenigen Einzelversuchen alle Zwei-Faktor-Kombinationen und fast alle Drei- bzw. Vier-Faktor-Kombinationen zu erfassen. So erhält man bei geringem Gesamtaufwand eine gute Absicherung gegen den Ausfall aufgrund einer ungünstigen Kombination.



Beispiel Computertest [5]

Es soll überprüft werden, ob ein neu entwickelter Computer unter allen denkbaren Hardwarekonfigurationen einwandfrei funktioniert. Der Einfluss folgender Komponenten (Faktoren) auf die Funktion soll getestet werden:

| | | | | | | | | |
|---|----------------------------|----|-------|-------|--|--|--|--|
| A | Bildschirm | | Typ X | Typ Y | | | | |
| B | Arbeitsspeicher | | klein | groß | | | | |
| C | Hersteller Stromversorgung | X | | Y | | | | |
| D | Verbindungskabel | | kurz | lang | | | | |
| E | Drucker | | Typ X | Typ Y | | | | |
| F | Festplatte | | klein | groß | | | | |
| G | Modem | ja | | nein | | | | |

Tabelle 8.21 zeigt den fraktionellen faktoriellen 2^{7-4} -Plan mit acht Faktorstufenkombinationen in den sieben Faktoren A – G in systematischer Reihenfolge.

TABELLE 8.21 Versuchsplan für den Funktionstest mit 7 zweistufigen Faktoren

| Nr. | Bildsch. A | Speicher B | Hersteller C | Kabel D | Drucker E | Festplatte F | Modem G | Funktion |
|-----|---------------|---------------|-----------------|------------|--------------|-----------------|------------|----------|
| 1 | Typ X | klein | X | lang | Typ Y | groß | ja | ja |
| 2 | Typ Y | klein | X | kurz | Typ X | groß | nein | ja |
| 3 | Typ X | groß | X | kurz | Typ Y | klein | nein | ja |
| 4 | Typ Y | groß | X | lang | Typ X | klein | ja | ja |
| 5 | Typ X | klein | Y | lang | Typ X | klein | nein | ja |
| 6 | Typ Y | klein | Y | kurz | Typ Y | klein | ja | ja |
| 7 | Typ X | groß | Y | kurz | Typ X | groß | ja | ja |
| 8 | Typ Y | groß | Y | lang | Typ Y | groß | nein | ja |

Bei diesem Funktionstest kann man davon ausgehen, dass eine bestimmte Konfiguration entweder funktioniert oder nicht, d. h. dass keine Zufallsstreuung auftritt. Außerdem hofft man natürlich, dass jede Konfiguration funktioniert – der Zweck dieses Tests ist ja, am Ende einer Entwicklung nachzuweisen, dass alles in Ordnung ist.

Man kann sich leicht davon überzeugen, dass dieser Plan alle möglichen Faktorstufenkombinationen zweier beliebiger Faktoren enthält. Außerdem enthält er 90% aller möglichen Drei-Faktor-Kombinationen und 50% aller Vier-Faktor-Kombinationen, obwohl nur 8 von $2^7 = 128$ insgesamt möglichen Kombinationen untersucht wurden. Man spricht daher von einer Testabdeckung von 100% der Zwei-Faktor-Kombinationen, 90% der Drei-Faktor-Kombinationen und 50% der Vier-Faktor-Kombinationen.

Nach einem solchen Ergebnis kann man die Freigabe mit wesentlich geringerem Risiko erteilen, als wenn man – ausgehend von einer Grundkonfiguration – für jeden der 7 Faktoren einzeln den Wert geändert hätte (one-factor-at-a-time), obwohl man dazu auch 8 Einzelversuche benötigt.

Genauer betrachtet enthält der Plan in Tabelle 8.22 z. B. alle Faktorstufenkombinationen der drei Faktoren ABC, ABE, ABF, ABG, aber nur die Hälfte der Kombinationen der drei Faktoren ABD (nur die vier Kombinationen Typ X-klein-lang, Typ Y-klein-kurz, Typ X-groß-kurz und Typ Y-groß-lang treten auf, die anderen vier fehlen).

Insgesamt gibt es 35 Möglichkeiten, 3 der 7 Faktoren auszuwählen. Für 28 dieser 35 Möglichkeiten werden alle acht Kombinationen untersucht. Nur für die 7 Möglichkeiten ABD, ACE, AFG, BCF, BEG, CDG und DEF wird nur die Hälfte der Kombinationen untersucht.

Durch geeignete Zuordnung der Faktoren kann man erreichen, dass potentiell kritische Kombinationen vollständig getestet werden.

Folgende fraktionellen faktoriellen Versuchspläne sind besonders nützlich für den Funktionstest:

- 2^{7-4} -Plan:

Mit 8 Faktorstufenkombinationen erreicht man für 7 Faktoren eine Testabdeckung von 90 % der Drei-Faktor-Kombinationen und 50 % der Vier-Faktorkombinationen.

- 2^{8-4} -Plan:

Mit 16 Faktorstufenkombinationen erreicht man für 8 Faktoren eine Testabdeckung von 100 % der Drei-Faktor-Kombinationen, 90 % der Vier-Faktorkombinationen und 50 % der Fünf-Faktor-Kombinationen.

- 2^{15-11} -Plan:

Mit 16 Faktorstufenkombinationen erreicht man für 15 Faktoren eine Testabdeckung von 96 % der Drei-Faktor-Kombinationen.

Beim Funktionstest ist die Testabdeckung wichtig, nicht die Auflösung. Es besteht jedoch ein Zusammenhang zwischen den beiden Beurteilungsgrößen:

- Pläne der Auflösung III haben eine Testabdeckung von 100 % der Zwei-Faktor-Kombinationen, aber weniger als 100 % der Drei-Faktor-Kombinationen. Die Generatoren mit 3 Buchstaben geben die Faktoren an, für die nur die Hälfte der möglichen Kombinationen untersucht wird.
- Pläne der Auflösung IV haben eine Testabdeckung von 100 % der Drei-Faktor-Kombinationen, aber weniger als 100 % der Vier-Faktor-Kombinationen. Die Generatoren mit 4 Buchstaben geben die Faktoren an, für die nur die Hälfte der möglichen Kombinationen untersucht wird.

Außer bei verschiedenen Gerätekonfigurationen kann man die hier beschriebene Vorgehensweise auch für den Funktionstest bei verschiedenen Kombinationen von Einsatzbedingungen wie z. B. Umgebungstemperatur, Luftfeuchtigkeit, Versorgungsspannung, Last u. ä. verwenden.

Auch für Zuverlässigkeitsuntersuchungen ist es sinnvoll, solche Kombinationen von Einsatzbedingungen zu verwenden. Dadurch kann man erkennen, wenn es Wechselwirkungen zwischen verschiedenen Einsatzbedingungen gibt und Probleme nur bei bestimmten Kombinationen auftreten.

■ 8.5 Definitive Screening Pläne

Definitive Screening Pläne [6] sind Pläne der Auflösung IV, mit denen auch quadratische Effekte bestimmt werden können. Sie sind nur für quantitative (= numerische) Faktoren einsetzbar, für die drei Stufen in gleichem Abstand voneinander einstellbar sind. Für k Faktoren besteht der Plan aus genau 2^{k+1} Faktorstufenkombinationen. Tabelle 8.22 zeigt beispielhaft einen Definitiven Screening Plan für 6 Faktoren.

TABELLE 8.22 Definitiver Screening Plan für 6 Faktoren

| Nr. | A | B | C | D | E | F |
|-----|----|----|----|----|----|----|
| 1 | 0 | +1 | +1 | +1 | +1 | +1 |
| 2 | 0 | -1 | -1 | -1 | -1 | -1 |
| 3 | +1 | 0 | +1 | +1 | -1 | -1 |
| 4 | -1 | 0 | -1 | -1 | +1 | +1 |
| 5 | +1 | +1 | 0 | -1 | -1 | +1 |
| 6 | -1 | -1 | 0 | +1 | +1 | -1 |
| 7 | +1 | +1 | -1 | 0 | +1 | -1 |
| 8 | -1 | -1 | +1 | 0 | -1 | +1 |
| 9 | +1 | -1 | -1 | +1 | 0 | +1 |
| 10 | -1 | +1 | +1 | -1 | 0 | -1 |
| 11 | +1 | -1 | +1 | -1 | +1 | 0 |
| 12 | -1 | +1 | -1 | +1 | -1 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 |

Mit Ausnahme des Zentrumspunktes (hier Nr. 13) treten die Faktorstufenkombinationen paarweise auf (z. B. Nr. 1 und 2). Ein Faktor steht im Zentrum (0), die anderen auf ± 1 mit jeweils vertauschten Stufenwerten.

Wie für Auflösung IV normal, sind die (Haupt-)Effekte der Faktoren nur mit 3FWW vermengt. Im Gegensatz zu den üblichen Screeningplänen können aus den Ergebnissen auch die quadratischen Effekte (vgl. Kapitel 11) der k Faktoren unabhängig voneinander berechnet werden. Sie sind allerdings auf komplizierte Art mit 2FWW vermengt, und 2FWW sind untereinander vermengt.

Daher wird empfohlen [7], einen Plan zu verwenden, der einige (z. B. zwei) Faktoren mehr enthält, als untersucht werden sollen und die überzähligen Spalten nicht zu belegen. Dieser Plan enthält dann mehr Einzelversuche (z. B. 4 mehr) als die Minimalzahl und erlaubt eine Schätzung der Zufallsstreuung aus den nicht belegten Faktoren. Zusätzlich kann auch noch der Zentrumspunkt mehrfach realisiert werden. Bei der Modellanpassung wird empfohlen, zunächst nur die Effekte der Faktoren zu betrachten und nicht signifikante Effekte aus dem Modell zu entfernen. Erst in einem zweiten Schritt werden Wechselwirkungen und quadratische Effekte betrachtet und einzeln ins Modell aufgenommen, wenn sie signifikant sind.

Diese Vorgehensweise funktioniert gut, wenn insgesamt maximal halb so viele Wechselwirkungen und quadratische Effekte wichtig sind als Faktoren k im Plan sind. Sind es mehr, so wird die Zuordnung zunehmend fehleranfällig. D. h. die Auswertung funktioniert umso besser, je weniger Effekte wirklich aktiv sind.

Zeigt sich bei der Auswertung, dass bei einem Plan mit mindestens 6 Faktoren nur maximal 3 aktiv sind und darf aus inhaltlichen Überlegungen davon ausgegangen werden, dass auch nur Wechselwirkungen und quadratische Effekte dieser 3 Faktoren auftreten, so kann an die Ergebnisse sofort ein quadratisches Modell (vgl. Kapitel 11) angepasst werden.

■ 8.6 Einsatzempfehlungen

Screening-Versuchspläne werden eingesetzt, wenn man mit möglichst wenigen Einzelversuchen erkennen möchte, welche Faktoren (aus einer meist größeren Anzahl k) einen besonders großen Effekt auf die Zielgröße(n) haben und in welche Richtung dieser Effekt geht.

Dazu wird empfohlen:

1. Es genügt zunächst, jeden Faktor auf zwei Stufen zu untersuchen.
2. Den besten Kompromiss zwischen Aufwand und Vermengungsrisiko bieten Pläne der Auflösung IV. Fraktionelle faktorielle Pläne sind bei Auflösung IV besser, da die Vermengungsstruktur angegeben werden kann, und so mit Hilfe von technischen Überlegungen signifikante 2FWW richtig identifiziert werden können (Abschnitt 8.2). Ist eine technisch begründete Zuordnung der Wechselwirkungen nicht möglich, so kann der Versuchsplan nachträglich erweitert werden, um vermengte Wechselwirkungen zu trennen (vgl. Abschnitt 16.1).
3. Pläne der Auflösung III sollten normalerweise **nicht** verwendet werden, da 2FWW die Effekte der Faktoren verfälschen können. Sind sie zur Begrenzung des Aufwands aber unvermeidlich, so sind Plackett-Burman-Pläne besser, da die 2FWW über viele Faktoren verteilt werden (Abschnitt 8.3). Werden einige Spalten nicht mit Faktoren belegt, so bieten auch diese Pläne eine gewisse Absicherung gegen Fehlschlüsse.

Bevor man sich für einen Plan der Auflösung III entscheidet, sollte man bedenken, dass die Anzahl m der Faktorstufenkombinationen für einen Plan der Auflösung IV nur doppelt so groß ist wie für Auflösung III, bei deutlich reduziertem Risiko.

Der Versuchsumfang N ergibt sich wie in (7.15) aus dem technologisch relevanten Unterschied $\Delta\mu$, der mit hoher Wahrscheinlichkeit erkannt werden soll, und der Standardabweichung der Zufallsstreuung σ .

$N = n \cdot m$: In den meisten Fällen ist es bei vorgegebenem Versuchsumfang N günstig, die Anzahl m der Faktorstufenkombinationen möglichst groß zu wählen und jede nur einmal zu realisieren ($n = 1$).

Sind die Faktoren quantitativ, so wird ein zusätzlicher Zentrumspunkt empfohlen (vgl. Abschnitt 16.2). Damit kann erkannt werden, ob es Abweichungen von der Linearität gibt (nur pauschal, ohne Zuordnung zu einem bestimmten Faktor). Eine gute (neue) Alternative sind hier die Definitiven Screening Pläne, die sogar eine Zuordnung zu bestimmten Faktoren erlauben, wenn die Gesamtzahl der zusätzlichen Effekte klein genug ist.

Fraktionelle faktorielle Pläne sind auch für den Funktionstest geeignet. Beim Funktions- test ist die Testabdeckung wichtig. Pläne der Auflösung III können hier durchaus sinnvoll sein (Abschnitt 8.4).

Literatur

- [1] *Box, G. E. P./Hunter, W. G./Hunter, J. S.*: Statistics for Experimenters. John Wiley, New York, 2. Auflage 2005
- [2] Qualitätsverbesserung durch Versuchsmethodik, Schulungsunterlagen der DGQ (Deutsche Gesellschaft für Qualität), Frankfurt am Main 1997
- [3] *Wu, C. F. J./Hamada, M.*: Experiments: Planning, Analysis, and Optimization. John Wiley, New York, 2. Auflage 2009
- [4] *Lin, D. K. J.*: A New Class of Supersaturated Designs. *Technometrics* 35 (Feb. 1993), 28 – 31
- [5] *Breyfogle, F. W.*: Implementing Six Sigma. John Wiley, New York, 2. Auflage 2003
- [6] *Jones, B./Nachtsheim, C.*: A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects. *Journal of Quality Technology* 43 (Jan. 2011), 1 – 15
- [7] *Jones, B./Nachtsheim, C.*: Effective Design-Based Model Selection for Definitive Screening Designs. *Technometrics* 59 (March 2017), 319 – 329

9

Robuste Produkte/ Prozesse

Häufig ist das Hauptziel einer Untersuchung, eine möglichst geringe Abhängigkeit des Ergebnisses von Störgrößen zu erreichen. Im Fall eines Fertigungsprozesses kann dies z.B. bedeuten, dass das Prozessergebnis möglichst wenig von Schwankungen der Prozessparameter, der Umgebungsbedingungen oder des Ausgangsmaterials beeinflusst wird. Im Fall eines Produktes kann dies auch eine möglichst geringe Abhängigkeit von den Einsatzbedingungen (z.B. Versorgungsspannung und Umgebungstemperatur bei einer elektronischen Schaltung) sein. Man spricht dann von robusten Prozessen bzw. Produkten.

Insbesondere G. Taguchi [1], [2] hat sich große Verdienste dadurch erworben, dass er die zentrale Bedeutung der Reduzierung des Streuverhaltens immer wieder betonte und eine Entwicklungsstrategie zusammengestellt hat, in deren Zentrum „robuste Produkte/Prozesse“ stehen.

Im Folgenden wird zunächst diese Entwicklungsstrategie nach G. Taguchi beschrieben.

Zur Umsetzung seiner Entwicklungsstrategie hat G. Taguchi die Methoden der Versuchsplanung vorgeschlagen und popularisiert. Er war dabei sehr erfolgreich und hat der Versuchsplanung viele neue Anwender gewonnen, die z.T. die Begriffe „Taguchi MethodeTM¹“ und „Versuchsplanung“ als Synonyme behandeln. Taguchi baut auf den klassischen Verfahren von Kapitel 8 auf. In diesem Kapitel werden speziell die Ergänzungen durch Taguchi beschrieben [1 – 6].

Viele dieser Ergänzungen sind bei den Anhängern der „klassischen Versuchsplanung“ jedoch heftig umstritten [7 – 8]. Daher soll hier auch der Versuch einer Wertung unternommen werden, und alternative Ansätze zum Erreichen „robuster Produkte/Prozesse“ werden gezeigt [9 – 11].

Das Motto der folgenden Darstellung ist:

Taguchi hat die klassische Versuchsplanung durch einige wesentliche Gesichtspunkte bereichert, vor allem bezüglich der Zielsetzung. Allerdings sollte man nicht jedes seiner Verfahren blind übernehmen, denn in vielen Punkten hat die klassische Versuchsplanung Besseres zu bieten.

¹ Trademark American Supplier Institute (ASI)

■ 9.1 Ziel und Strategie von G. Taguchi

In diesem Abschnitt werden das Qualitätsziel und die Entwicklungsstrategie von Taguchi behandelt, d.h. die konzeptionellen Aspekte. Auch Kritiker von G. Taguchi stimmen mit ihm bezüglich des Ziels überein. Nur der Weg zu diesem Ziel ist umstritten.

9.1.1 Qualitätsziel: Streuung minimieren

Bild 9-1 vergleicht das frühere Toleranzverständnis mit dem Toleranzverständnis von Taguchi. Nach früherem Toleranzverständnis steht der Hersteller im Vordergrund. Ihm entsteht (zumindest auf den ersten Blick) nur dann ein Verlust, wenn ein Qualitätsmerkmal außerhalb der Toleranz liegt. Die Größe des Verlusts ist der Warenwert bzw. die Reparaturkosten. Dieser Verlust aus Herstellersicht steigt an den Toleranzgrenzen des Qualitätsmerkmals sprunghaft an.

Im Toleranzverständnis von Taguchi steht der Kunde bzw. die Gesellschaft als Ganzes im Vordergrund. Für den Kunden reduziert sich der Gebrauchswert eines Produkts kontinuierlich und mit jeder Abweichung des Qualitätsmerkmals vom Zielwert. Die einfachste funktionale Abhängigkeit dieser Form ist eine quadratische Zunahme des Verlusts mit der Abweichung vom Zielwert.

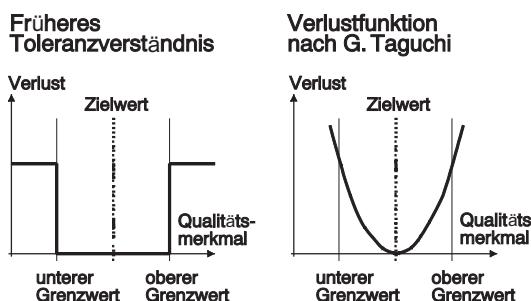


BILD 9-1

Vergleich zwischen frührerem und Taguchs Toleranzverständnis

Aus dem Unterschied im Toleranzverständnis resultieren unterschiedliche Qualitätsziele:
Früher: Es genügt, die Spezifikation einzuhalten.

Taguchi: Auch innerhalb der Spezifikation sollte die Streuung bzw. die Abweichung vom Zielwert möglichst klein sein.

Taguchis Ziel entspricht dem zeitgemäßen Qualitätsverständnis, das den Kunden in den Mittelpunkt stellt. Ein wichtiges **Ziel der Produkt- bzw. Prozessentwicklung ist die Minimierung der Streuung**.

9.1.2 Entwicklungsstrategie: Robuste Produkte/Prozesse

Die Minimierung der Streuung darf nicht dadurch erreicht werden, dass der Entwickler enge Toleranzen vorgibt (eine leider weit verbreitete Vorgehensweise). Das wäre mit ho-

hen Kosten verbunden. Geringe Streuung muss vielmehr durch robuste Produkte bzw. Prozesse erreicht werden.

- Produkte sind robust, wenn ihre Eigenschaften möglichst wenig von Fertigungs- bzw. Einsatzbedingungen abhängen.
- Prozesse sind robust, wenn das Prozessergebnis möglichst wenig von unvermeidlichen Schwankungen der Prozessparameter, Materialeigenschaften, Umgebungsbedingungen u. ä. abhängt.

Bild 9-2 zeigt grafisch, was „robuster Prozess“ bedeutet. Die Bilder 9-3 und 9-4 zeigen, wie der Wert eines Prozessparameters die Robustheit eines Prozesses beeinflussen kann (analog für „robuste Produkte“).

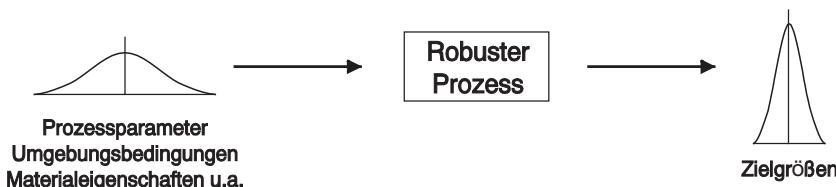


BILD 9-2 Ein robuster Prozess ist ein Prozess, dessen Ergebnis wenig streut, obwohl Prozessparameter, Umgebungsbedingungen, Materialeigenschaften u. a. in vernünftigem Rahmen streuen.

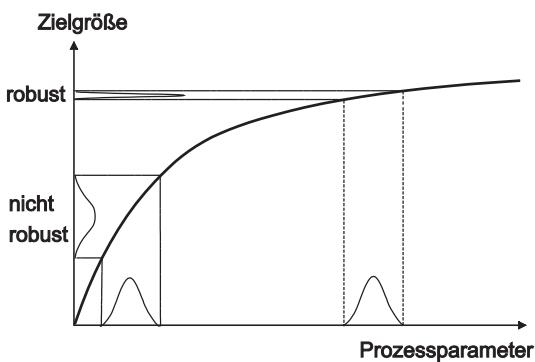


BILD 9-3
Besteht ein nichtlinearer Zusammenhang zwischen einem Prozessparameter und einer Zielgröße, so gibt es Einstellwerte des Parameters, in denen sich eine bestimmte Streuung weniger auf die Zielgröße auswirkt als bei anderen Werten. Diese sucht man für robuste Prozesse.

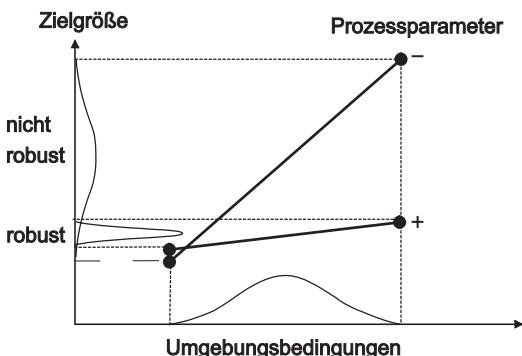


BILD 9-4
Besteht eine Wechselwirkung zwischen einem Prozessparameter und z. B. einer Umgebungsbedingung wie der Temperatur, so wirkt sich dieselbe Streuung der Umgebungsbedingung bei einem Wert des Prozessparameters (hier +) weniger aus als bei einem anderen Wert (hier -).

Um robuste Produkte bzw. Prozesse zu erreichen, unterscheidet G. Taguchi drei Entwicklungsschritte (Designphasen):

1. Systemdesign

Mit Erfahrung, technischem Wissen und neuen Erfindungen wird ein Produkt (Prozess) entwickelt, das (der) unter definierten Bedingungen funktioniert.

2. Parameterdesign

Durch geplante Versuche wird festgestellt, welche Faktoren die Streuung beeinflussen. Dann werden zuerst diese Faktoren auf Werte festgelegt, die die Streuung minimieren (robustes Produkt/Prozess). Erst als nächster Schritt werden die restlichen Faktoren so festgelegt, dass der richtige Zielwert (Mittelwert) erreicht wird. Hinter dieser Vorgehensweise steckt die Überlegung, dass es meist viel leichter ist, den Mittelwert zu verändern, als die Streuung zu reduzieren (das ist auch aus der Prozessregelung bekannt).

3. Toleranzdesign

Erst als Notlösung, wenn robuste Produkte/Prozesse allein nicht ausreichen, um die Streuung zu reduzieren, dürfen im letzten Schritt Toleranzen eingeengt werden.

Der wichtigste Schritt ist das Parameterdesign. Darauf sollte der meiste Entwicklungsaufwand verwendet werden. Dabei helfen Taguchis Versuchspläne.

■ 9.2 Taguchis Versuchspläne und ihre Auswertung

Der wichtigste neue Aspekt von Taguchi ist die Suche nach robusten Produkten/Prozessen. Man sucht nach Einstellwerten von Prozess- bzw. Produktparametern, bei denen sich Störgrößen und zufällige Schwankungen der Parameter möglichst wenig auswirken. Dementsprechend enthalten Taguchis Versuchspläne zwei Arten von Faktoren:

1. Steuerfaktoren (control factors = Steuergrößen im Versuchsplan)

Sie können auch im Fertigungsprozess bzw. im Produkt auf einen bestimmten Wert eingestellt werden (z. B. die Prozessparameter eines Fertigungsprozesses oder die Nominalwerte der Widerstände einer elektronischen Schaltung).

2. Rauschfaktoren (noise factors = Störgrößen im Versuchsplan)

Sie können zwar im Versuch gezielt verändert werden, in der Fertigung bzw. im Produkt soll das Ergebnis aber möglichst unabhängig von ihrem Wert sein (z. B. die Position eines Teils in einem Brennofen, die Umgebungstemperatur bei einer elektronischen Schaltung).

Ziel des Versuchs ist es, Stufenwerte für die Steuerfaktoren zu finden, bei denen sich die Rauschfaktoren möglichst wenig auswirken. Um dieses Ziel zu erreichen, besteht der Versuchsplan aus zwei Teilen:

- einem Versuchsplan für die Steuerfaktoren (Taguchi nennt ihn „inner array“) und
- einem Versuchsplan für die Rauschfaktoren (Taguchi nennt ihn „outer array“).

Für jede Faktorstufenkombination des „inner array“ werden alle Faktorstufenkombinationen des „outer array“ je einmal realisiert. Bild 9-5 zeigt das Prinzip eines solchen Versuchsplans an einem Beispiel. Das „inner array“ in Bild 9-5 ist ein fraktioneller faktorieller 2^{7-4} -Plan, das „outer array“ ein fraktioneller faktorieller 2^{3-1} -Plan (senkrecht lesen).

| | | „inner array“ mit Steuerfaktoren | | „outer array“ mit Rauschfaktoren | | Rauschfaktoren | | 1 | 2 | 3 | 4 |
|-----|---|----------------------------------|---|----------------------------------|---|----------------|---|---|----|----|----|
| Nr. | | Steuerfaktoren | | | | | | | R1 | R2 | R3 |
| | | A | B | C | D | E | F | G | | | |
| 1 | - | - | - | - | + | + | + | - | | | |
| 2 | + | - | - | - | - | - | + | + | | | |
| 3 | - | + | - | - | + | - | + | | | | |
| 4 | + | + | - | + | - | - | - | | | | |
| 5 | - | - | + | + | - | - | + | | | | |
| 6 | + | - | + | - | + | - | - | | | | |
| 7 | - | + | + | - | - | + | - | | | | |
| 8 | + | + | + | + | + | + | + | | | | |

Versuchsergebnisse:
 zugehörige
 Zielgrößenwerte

BILD 9-5 Versuchsplan nach G. Taguchi aus „inner array“ und „outer array“.

Als „inner array“ schlägt G. Taguchi sogenannte „orthogonale Felder“ vor. Für zweistufige Steuerfaktoren sind dies (abgesehen von einer geänderten Bezeichnung) fraktionelle faktorielle Versuchspläne und Plackett-Burman-Pläne der Auflösung III aus Kapitel 8.

Vorteile dieser Pläne sind die geringe Anzahl Faktorstufenkombinationen und die (scheinbar) einfache Interpretation der Ergebnisse, da nur Effekte der Faktoren bestimmt werden können.

Ein Nachteil ist, dass Wechselwirkungen die Ergebnisse völlig verändern können und dann zu Fehlschlüssen führen. Beliebige andere Versuchspläne (z.B. fraktionelle faktorielle Versuchspläne der Auflösung IV, aber auch vollständige faktorielle Pläne und die Regressionspläne aus Kapitel 11) können als „inner array“ verwendet werden. Somit kann dieser Nachteil leicht umgangen werden.

Als „outer array“ können ebenfalls Versuchspläne der Auflösung III verwendet werden – das ist auch sinnvoll. Häufig wird auch ein einzelner Faktor auf vielen Stufen verwendet (z.B. verschiedene Positionen im Brennofen im folgenden Beispiel).

Auswertung

Die verschiedenen Stufen des „outer array“ werden wie mehrmalige Realisierungen der Faktorstufenkombinationen des „inner array“ behandelt (was sie nicht sind, da sich die Rauschfaktoren unterscheiden). Mittelwert und Streuung sind zwei verschiedene, getrennt behandelte Zielgrößen. Für die Streuung schlägt G. Taguchi das „Signal-Rausch-Verhältnis“ (signal-to-noise-ratio S/N nominal-the-best) vor. Das S/N soll möglichst groß sein, dann ist die Streuung klein.

$$\text{Signal-Rausch-Verhältnis } S/N = 10 \cdot \log_{10} \left(\frac{\bar{y}^2}{s^2} \right) \quad (9.1)$$

Taguchis Begründung für die Verwendung gerade dieser Größe als Maß für die Streuung ist fragwürdig und umstritten. Trotzdem kann S/N eine nützliche Zielgröße sein, und zwar dann, wenn problembedingt die Streuung proportional zum Mittelwert zunimmt und man an Änderungen der Streuung interessiert ist, die über diese problembedingte Abhängigkeit hinausgehen. Die logarithmische Transformation im S/N wird auch in der klassischen Versuchsplanung empfohlen, wenn die Standardabweichung einer Größe proportional zum Mittelwert zunimmt (vgl. Absatz 6.4.3).²

S/N ist jedoch nicht die einzige mögliche Zielgröße für die Streuung. Taguchi selbst beschreibt mehrere Varianten [2]. Logothetis und Wynn [10] zeigen, dass es häufig günstiger ist, das

$$\text{Robustheitsmaß} = 10 \cdot \log_{10} \left(\frac{1}{s^2} \right) \quad (9.2)$$

zu verwenden. Franzkowski [9] schlägt vor, die Standardabweichung s zu verwenden. Im Fall von Zufallsstreuung hat s den Vorteil, dass seine Verteilung bekannt ist. Auch die im Praxisbeispiel Absatz 8.2.7 verwendete Zielgröße V ist ein Maß für die Streuung (Rauschfaktor ist dort die Position auf der Leiterplatte).

Es wird empfohlen, mehrere Alternativen zu testen. Normalerweise sind die Schlussfolgerungen weitgehend unabhängig von der genauen Definition der Zielgröße, solange sie die Streuung misst.

Effekte für Mittelwert und S/N werden wie in (7.11) berechnet.

$$\begin{aligned} \text{Effekt (Mittelwert)} &= \frac{2}{m} \cdot \sum_i (\text{Vorzeichen} \cdot \bar{y}_i) \\ \text{Effekt (S / N)} &= \frac{2}{m} \cdot \sum_i (\text{Vorzeichen} \cdot (S/N)_i) \end{aligned} \quad (= 7.11) \quad (9.3)$$

Für die Zielgröße „Mittelwert“ erfolgt die Beurteilung der Signifikanz häufig wie in Kapitel 7 bzw. 8.³ Da die Standardabweichung jedoch aufgrund der gezielten Veränderung der Rauschfaktoren im „outer array“ größer ist als zufällig, wird damit die Signifikanz von Effekten unterschätzt. Alternativ kann man die Mittelwerte als Ergebnisse eines Versuchs mit $n = 1$ behandeln und eine Schätzung für die Zufallsstreuung gemäß (7.16) aus dem Pooling kleiner Effekte ableiten.

Für die Zielgröße S/N gibt es nur einen Wert je Faktorstufenkombination. Daher muss die Beurteilung der Signifikanz wie für Versuchspläne mit $n = 1$ erfolgen. Damit sind (wie in

² Der Faktor 10 in (9.1) ist willkürlich und überflüssig. Die Quadrate im Logarithmus dienen nur dazu, positive Werte zu garantieren. Statt Logarithmen zur Basis 10 könnte auch der natürliche Logarithmus verwendet werden.

³ Taguchi empfiehlt die Varianzanalyse. Varianzanalyse wird in Kapitel 12 behandelt. Für die bisher behandelten Versuchspläne mit Faktoren auf zwei Stufen führen Varianzanalyse und der in Kapitel 6 und 7 behandelte Vergleich der Effekte mit Vertrauensbereichen jedoch zu identischen Ergebnissen.

Abschnitt 7.3) gewisse Risiken verbunden, vor allem wenn das „inner array“ nur wenige Faktorstufenkombinationen enthält.

Aufgrund der Ergebnisse werden die **Steuerfaktoren** unterteilt in:

1. Steuerfaktoren, die **Mittelwert und Streuung** (z. B. S/N) oder **nur die Streuung** beeinflussen:
Sie werden auf die Stufe mit der kleineren Streuung gesetzt.
2. Steuerfaktoren, die **nur den Mittelwert** beeinflussen:
Sie werden auf die Stufe gesetzt, die den richtigen Mittelwert ergibt.
3. Steuerfaktoren, die **weder Mittelwert noch Streuung** beeinflussen:
Sie werden auf die Stufe gesetzt, die die niedrigeren Kosten verursacht oder die aus anderen Gründen günstiger ist.



Beispiel aus der Prozessentwicklung (nach [1]⁴)

Bild 9-6 zeigt die Einzelprozessschritte in der Fertigung von Keramikfliesen. Nach dem Formen werden sie in einem Tunnelofen gebrannt. Beim Brennen schrumpfen die Fliesen. Das Problem besteht darin, dass Fliesen aus dem Innenbereich nach dem Brennen eine andere Größe haben als Fliesen aus dem Außenbereich. Dadurch erhält man eine große Streuung. Ein Teil der Fliesen liegt außerhalb der Grenzwerte.

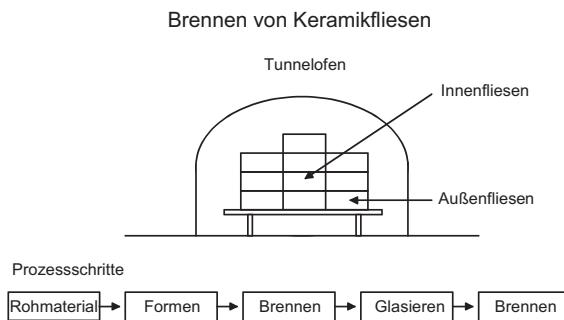


BILD 9-6

Prozessschritte bei der Fertigung von Keramikfliesen und Querschnitt durch den Tunnelofen, in dem die Fliesen gebrannt werden

Die Position der Fliesen im Ofen soll einen möglichst kleinen Einfluss auf die Zielgröße „Größe der Fliesen nach dem Brennen“ haben. Sie wird daher als **Rauschfaktor** behandelt. Fliesen von fünf verschiedenen, festgelegten Positionen werden jeweils gemessen. Tabelle 9.1 zeigt die Prozessparameter, die als Steuerfaktoren untersucht werden und ihre Stufen.⁵

⁴ In der hier verwendeten Form ist das Beispiel Copyright von ASI und MQI Consultants B. V. und wird mit freundlicher Genehmigung von MQI, ASI Consultants in Europe, veröffentlicht.

⁵ Taguchi bezeichnet die Stufen mit 1 und 2 statt – und +.

| Steuerfaktor | Stufe + | Stufe - |
|--------------------------|------------|---------------|
| A: Typ Agalmatolit | alt | neu |
| B: Korngröße Kalk | grob (alt) | feiner |
| C: Menge Kalk | 5 % | 1 % (alt) |
| D: Menge Abfallrecycling | 0 % | 4 % (alt) |
| E: Rohmaterialmenge | 1300 kg | 1200 kg (alt) |
| F: Menge Agalmatolit | 43 % | 53 % (alt) |
| G: Menge Feldspat | 0 % | 5 % (alt) |

TABELLE 9.1

Steuerfaktoren (Prozessparameter) und ihre Stufen

Tabelle 9.2 zeigt den Versuchsplan und die Ergebnisse. Als „inner array“ wird ein fraktioneller faktorieller 2^{7-4} -Versuchsplan mit Auflösung III verwendet (Taguchi nennt diesen Plan L_8). Das „outer array“ besteht aus den fünf Positionen P1 bis P5.

TABELLE 9.2 Versuchsplan und Ergebnisse (Einzelwerte,⁶ Mittelwerte, Standardabweichung und Signal-Rausch-Verhältnis, sowie deren Effekte)

| Nr. | Faktorstufen | | | | | | | Messwerte an Position | | | | | \bar{y} | s | S/N |
|-----------|--------------|------|------|------|------|------|------|-----------------------|-------|-------|-------|-------|-----------|-------|------|
| | A | B | C | D | E | F | G | P1 | P2 | P3 | P4 | P5 | | | |
| 1 | - | - | - | + | + | + | - | 153,0 | 152,0 | 151,3 | 150,0 | 149,5 | 151,16 | 1,433 | 40,5 |
| 2 | + | - | - | - | - | + | + | 154,5 | 153,3 | 151,8 | 150,4 | 149,6 | 151,92 | 2,017 | 37,5 |
| 3 | - | + | - | - | + | - | + | 156,5 | 152,1 | 150,3 | 148,5 | 144,6 | 150,40 | 4,398 | 30,7 |
| 4 | + | + | - | + | - | - | - | 151,5 | 150,8 | 150,6 | 150,2 | 149,7 | 150,56 | 0,673 | 47,0 |
| 5 | - | - | + | + | - | - | + | 152,2 | 151,3 | 151,1 | 150,6 | 150,0 | 151,04 | 0,820 | 45,3 |
| 6 | + | - | + | - | + | - | - | 153,1 | 151,8 | 151,8 | 151,4 | 150,6 | 151,74 | 0,904 | 44,5 |
| 7 | - | + | + | - | - | + | - | 151,5 | 150,8 | 150,0 | 149,4 | 149,1 | 150,16 | 0,991 | 43,6 |
| 8 | + | + | + | + | + | + | + | 151,9 | 151,4 | 150,4 | 150,2 | 149,6 | 150,70 | 0,933 | 44,2 |
| Eff. | | | | | | | | | | | | | | | |
| \bar{y} | 0,5 | -1,0 | -0,1 | -0,2 | 0,1 | 0,0 | 0,1 | | | | | | | | |
| s | -0,8 | 0,5 | -1,2 | -1,1 | 0,8 | -0,4 | 1,0 | | | | | | | | |
| S/N | 3,3 | -0,6 | 5,5 | 5,2 | -3,4 | -0,4 | -4,5 | | | | | | | | |

Signifikanz der Effekte für das Signal-Rausch-Verhältnis S/N:

Durch Pooling der zwei kleinen Effekte (B und F) im Signal-Rausch-Verhältnis erhält man mit (7.16)

$$s_{\bar{d}}^2 = \frac{1}{2} \cdot ((-0,6)^2 + (-0,4)^2) = 0,26 \Rightarrow s_{\bar{d}} = \sqrt{0,26} = 0,51$$

Mit $f = 2$ und Tabelle 6.4 erhält man für die Breite der Vertrauensbereiche:

$$95\%: t \cdot s_{\bar{d}} = 4,3 \cdot 0,51 = 2,2$$

$$99\%: t \cdot s_{\bar{d}} = 9,9 \cdot 0,51 = 5,0$$

$$99,9\%: t \cdot s_{\bar{d}} = 31,6 \cdot 0,51 = 16$$

⁶ Man beachte, dass die Einzelergebnisse von P1 nach P5 abnehmen. Dies bedeutet, dass es echte Unterschiede zwischen den Positionen gibt. Ziel der Untersuchung ist es, Stufenwerte der Steuerfaktoren zu finden, bei denen diese Unterschiede möglichst klein sind.

Die Signifikanz der nicht gepoolten Effekte wird daher mit * bzw. ** beurteilt. A, C, D, E und G sind daher Steuerfaktoren, die die Streuung beeinflussen. Im Parameterdesign werden sie auf die Stufe gesetzt, die das bessere (größere) Signal-Rausch-Verhältnis ergibt. Damit erhält man die optimalen Einstellungen von Tabelle 9.3.

| Steuergröße | opt. Einstellung |
|--------------------------|------------------|
| A: Typ Agalmatolit | alt |
| C: Menge Kalk | 5 % (neu) |
| D: Menge Abfallrecycling | 0 % (neu) |
| E: Rohmaterialmenge | 1200 kg (alt) |
| G: Menge Feldspat | 5 % (alt) |

TABELLE 9.3
Optimale Einstellungen der Steuerfaktoren

B und F haben keinen Einfluss auf das S/N, daher kann man den kostengünstigen Wert wählen. Wenn die Streuung klein ist, kann die Größe der Fliesen leicht durch die Größe der Form festgelegt werden. Auch mit dem Faktor B kann man den Mittelwert zum Sollwert hin verschieben.

Bild 9-7 zeigt das Ergebnis für die Verteilung der Fliesengrößen vor und nach der Optimierung. Durch die Optimierung wurde der Unterschied zwischen Mitte und Rand und auch die Streuung an jeder Stelle wesentlich reduziert. Dies wurde ohne Einengung irgendwelcher Toleranzen erreicht, sondern nur durch geeignete Wahl der Stufen der Steuerfaktoren im Parameterdesign. Der neue Prozess ist robust.

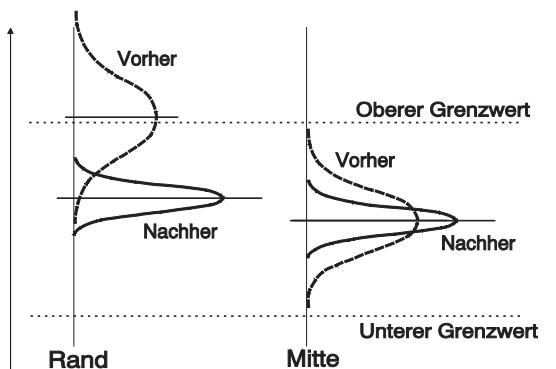


BILD 9-7

Vergleich der Verteilung der Fliesengrößen an Mitte und Rand des Ofens vor und nach der Optimierung

Beispiel aus der Produktentwicklung

Teile dieses Übungsbeispiels sind als Aufgabe formuliert. Es wird empfohlen, diese Teile zu beantworten, bevor Sie die unmittelbar folgende Lösung durchlesen.

Ein elektrischer Schaltkreis enthält u. a. vier Widerstände R1, R2, R3 und R4. Sie sollen so gewählt werden, dass die Ausgangsspannung möglichst wenig streut. Insbesondere soll sie möglichst wenig von der Umgebungstemperatur, der Versorgungsspannung und dem Lastwiderstand am Ausgang abhängen.

Aufgabe

Was sind in diesem Beispiel die Steuer- und Rauschfaktoren und warum?

Lösung

Steuerfaktoren: R1, R2, R3, R4

Sollwerte für diese Faktoren können vom Entwickler vorgegeben werden. Sie werden so vorgegeben, dass die Ausgangsspannung möglichst wenig von den Rauschfaktoren abhängt. Stufenwerte werden aufgrund der Erfahrung des Entwicklers festgelegt, indem er sich fragt: Welche Werte lassen vernünftige Ergebnisse erwarten?

Rauschfaktoren: Umgebungstemperatur, Versorgungsspannung, Lastwiderstand

Die Ausgangsspannung soll möglichst unabhängig von diesen Größen sein. Der Entwickler hat auf diese Größen keinen Einfluss, da sie von der Anwendung abhängen. Der Entwickler kann nur zulässige Bereiche vorschreiben. Diese sollen möglichst breit sein. Innerhalb dieser Bereiche soll die Ausgangsspannung möglichst wenig von diesen Faktoren abhängen. Für den Versuch können diese Faktoren aber gezielt eingestellt werden. Sinnvolle Stufenwerte sind die geplanten Spezifikationsgrenzen.

TABELLE 9.4 Steuer- und Rauschfaktoren mit den jeweiligen Stufenwerten

| Steuerfaktoren | | | | Rauschfaktoren | | | | |
|----------------|----|-------|--------|----------------|----|---------------------|-------|-------|
| A | R1 | 3 kΩ | 5 kΩ | a | T | Temperatur | 0 °C | 50 °C |
| B | R2 | 100 Ω | 200 Ω | b | VE | Versorgungsspannung | 4,5 V | 5,5 V |
| C | R3 | 20 kΩ | 50 kΩ | c | RL | Lastwiderstand | 4 kΩ | 16 kΩ |
| D | R4 | 1 kΩ | 1,5 kΩ | | | | | |

TABELLE 9.5 Inner und outer array mit den Ergebnissen der 32 Einzelversuche, sowie die daraus berechneten Mittelwerte, Standardabweichungen und Signal-Rausch-Verhältnisse. Man beachte, dass S/N groß ist, wenn s klein ist. Die beiden Größen enthalten ähnliche Informationen.

| Nr. | A R1 | B R2 | C R3 | D R4 | a | 0 | 50 | 0 | 50 | \bar{y} | s | S/N |
|-----|---------|---------|---------|---------|------|------|------|------|------|-----------|------|-----|
| | | | | | b | 4,5 | 4,5 | 5,5 | 5,5 | | | |
| | | | | | c | 4 | 16 | 16 | 4 | | | |
| 1 | 3 | 100 | 20 | 1 | 2,85 | 3,69 | 4,35 | 3,07 | 3,49 | 0,675 | 14,3 | |
| 2 | 5 | 100 | 20 | 1,5 | 3,33 | 3,29 | 5,02 | 4,72 | 4,09 | 0,909 | 13,1 | |
| 3 | 3 | 200 | 20 | 1,5 | 4,34 | 4,32 | 4,82 | 4,56 | 4,51 | 0,234 | 25,7 | |
| 4 | 5 | 200 | 20 | 1 | 2,78 | 3,58 | 5,39 | 4,05 | 3,95 | 1,094 | 11,1 | |
| 5 | 3 | 100 | 50 | 1,5 | 5,77 | 5,92 | 6,50 | 6,09 | 6,07 | 0,315 | 25,7 | |
| 6 | 5 | 100 | 50 | 1 | 4,28 | 5,04 | 6,97 | 5,79 | 5,52 | 1,146 | 13,7 | |
| 7 | 3 | 200 | 50 | 1 | 5,30 | 6,00 | 6,85 | 5,53 | 5,92 | 0,685 | 18,7 | |
| 8 | 5 | 200 | 50 | 1,5 | 5,76 | 5,74 | 7,52 | 7,02 | | | | |

Tabelle 9.5 zeigt den Versuchsplan und die Ergebnisse. Das „inner array“ der Steuerfaktoren ist ein 2^{4-1} -Plan mit Auflösung IV. Das „outer array“ der Rauschfaktoren ist ein 2^{3-1} -Plan mit Auflösung III. Für jede der acht Faktorstufenkombinationen des „inner array“ wird eine Schaltung gefertigt. Jede dieser acht Schaltungen wird bei jeder der vier Kombinationen von Einsatzbedingungen des „outer array“ betrieben und die Ausgangsspannung gemessen. D.h. man erhält insgesamt $8 \times 4 = 32$ Messwerte. Tabelle 9.6 zeigt die berechneten Effekte für \bar{y} , s und S/N.

Aufgabe

Berechnen Sie die fehlende Zeile in Tabelle 9.5 (Lösung in Tabelle 9.6) und die fehlenden Effekte in Tabelle 9.6.

TABELLE 9.6 Berechnung der Effekte der Steuerfaktoren und ihrer 2FWW für \bar{y} , s und S/N
(Effekt = Summe/4)

| Nr. | A R1 | B R2 | C R3 | D R4 | AB + CD WW | AC + BD WW | AD + BC WW | \bar{y} | s | S/N |
|----------------|---------|---------|---------|---------|---------------|---------------|---------------|-----------|-------|------|
| 1 | - | - | - | - | + | + | + | 3,49 | 0,675 | 14,3 |
| 2 | + | - | - | + | - | - | + | 4,09 | 0,909 | 13,1 |
| 3 | - | + | - | + | - | + | - | 4,51 | 0,234 | 25,7 |
| 4 | + | + | - | - | + | - | - | 3,95 | 1,094 | 11,1 |
| 5 | - | - | + | + | + | - | - | 6,07 | 0,315 | 25,7 |
| 6 | + | - | + | - | - | + | - | 5,52 | 1,146 | 13,7 |
| 7 | - | + | + | - | - | - | + | 5,92 | 0,685 | 18,7 |
| 8 | + | + | + | + | + | + | + | 6,51 | 0,901 | 17,2 |
| Eff. \bar{y} | 0,02 | 0,43 | | 0,575 | -0,005 | 0,00 | -0,01 | | | |
| s | 0,535 | -0,028 | | -0,310 | 0,0028 | -0,012 | 0,095 | | | |
| S/N | -7,325 | 1,475 | | 5,975 | -0,725 | 0,575 | -3,225 | | | |

Lösung

Effekte von R3: Für \bar{y} : 1,995; s: 0,034; S/N: 2,775

Durch Pooling der jeweils kleinen Effekte erhält man: Signifikante Effekte

für \bar{y} : R2, R3, R4

für S/N (ebenso für s): R1 und R4

Im Parameterdesign werden R1 und R4 auf die Stufen gesetzt, die den größeren Wert von S/N (bzw. den kleineren Wert von s) ergeben:

$$R1 = -- = 3 \text{ k}\Omega \quad R4 = +- = 1,5 \text{ k}\Omega$$

Mit Hilfe von R2 und R3 wird der Mittelwert auf 5 V angepasst. Bild 9-8 zeigt die Abhängigkeit der Ausgangsspannung von R2 und R3, für die o. a. Werte von R1 und R4.

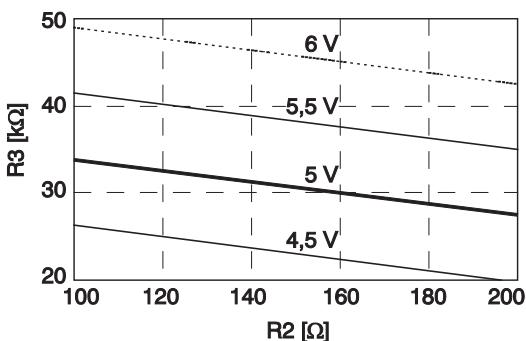


BILD 9-8

Ausgangsspannung in Abhängigkeit von R2 und R3 (für R1 = 3 kΩ und R4 = 1,5 kΩ): Für jede Kombination von R2 und R3 entlang der fetten Linie erhält man als Mittelwert der Ausgangsspannung 5V.

■ 9.3 Alternative Ansätze

Der S/N ist nicht die einzige Möglichkeit, die Robustheit von Produkten und Prozessen zu beurteilen. Im Folgenden werden mehrere Alternativen behandelt. Welcher dieser Ansätze am besten geeignet ist, hängt von der jeweiligen Problemstellung ab. Häufig gelangt man mit verschiedenen Ansätzen zu ähnlichen Schlussfolgerungen.

Ziel der folgenden Darstellung ist die Erweiterung der Betrachtungsweise.

9.3.1 Aus der Differenz von Messwerten abgeleitete Zielgrößen

Häufig werden mehrere Teile gleichzeitig hergestellt, die sich systematisch unterscheiden, oder es gibt systematische Unterschiede innerhalb eines Teils. Ziel des Versuchs ist es, diese Unterschiede zu minimieren.

-  **Beispiel**
- Beim Druckguss wird eine Mehrfachform verwendet. Das Füllverhalten der verschiedenen Formen ist unterschiedlich – daraus resultieren systematische Unterschiede zwischen Teilen aus den verschiedenen Formen.
 - In einem Rohrofen wird eine Schicht gleichzeitig auf vielen Halbleiterscheiben abgeschieden. Da die Reaktionsgase den Ofen in Längsrichtung durchströmen, unterscheiden sich die Bedingungen und damit die Ergebnisse an den verschiedenen Scheiben.
 - Beim Brennen der Fliesen im Beispiel von Abschnitt 9.2 gibt es Unterschiede zwischen Innen- und Außenfliesen (P1 und P5).
 - Beim Reflowlöten ist die Temperatur an der Lötstelle immer niedriger als die Temperatur auf der Leiterplatte. Diese Differenz soll minimiert werden (Praxisbeispiel Absatz 8.2.7).
 - Bei der galvanischen Abscheidung einer Schicht gibt es (bedingt durch die Form des Teils) Stellen mit besonders hoher bzw. niedriger Schichtdicke (Beispiel Absatz 8.2.2).

In diesen Fällen ist die Differenz der Messwerte von Teilen bzw. Stellen mit erfahrungsgemäß besonders hohem und besonders niedrigem Wert eine Zielgröße, die sich für die Prozessoptimierung eignet (vgl. Beispiel 8.2.2). Manchmal ist es noch günstiger, bereits bekannte Effekte bei der Definition der Zielgröße zusätzlich zu berücksichtigen und statt der Differenz eine aus der Differenz errechnete Zielgröße zu verwenden (vgl. Beispiel 8.2.7).

Vorteil solcher aus der Differenz von Messwerten abgeleiteten Zielgrößen ist der geringe Aufwand. Bei einem Einzelversuch werden jeweils mehrere Werte gemessen, aus denen die Zielgröße dann errechnet wird. Aufgrund des geringen Aufwands sind oft echte Wiederholungen der Einzelversuche möglich, und die mit $n = 1$ verbundene Subjektivität kann vermieden werden.

9.3.2 Wechselwirkung zwischen Steuer- und Rauschfaktoren

Bild 9-4 zeigt allgemein den Zusammenhang zwischen Robustheit und Wechselwirkung. Robustheit kann auch erreicht werden, indem man gezielt nach den signifikanten Wechselwirkungen zwischen Steuer- und Rauschfaktoren sucht [11]. Für Steuerfaktoren mit solchen Wechselwirkungen wählt man dann als Einstellung den Stufenwert, bei dem die Abhängigkeit vom Rauschfaktor kleiner ist.

Beispiel

Die Daten in Tabelle 9.5 kann man auch als fraktionellen faktoriellen Versuch mit den sieben Faktoren R1, R2, R3, R4, T, VE und RL und 32 Einzelversuchen mit $n = 1$ auffassen. Für diesen Versuch kann man, wie in Kapitel 8, die Effekte der Faktoren und 2FWW berechnen (soweit sie nicht verengt sind).

Durch Pooling kleiner Effekte, wie in Abschnitt 7.3, erkennt man, dass die 2FWW zwischen R1 und VE sowie zwischen R4 und RL signifikant sind. Bild 9-9 zeigt diese 2FWW in der Darstellung von Kapitel 7.

Man erkennt, dass sich bei $R1 = 3 \text{ k}\Omega$ der Rauschfaktor Versorgungsspannung VE weniger auf die Zielgröße V auswirkt als bei $R1 = 5 \text{ k}\Omega$, und dass sich bei $R4 = 1,5 \text{ k}\Omega$ der Lastwiderstand RL weniger auswirkt als bei $R4 = 1 \text{ k}\Omega$. Man kommt damit zum selben Ergebnis wie mit dem S/N, erhält jedoch Zusatzinformation, die u.U. für die weitere Optimierung nützlich ist.

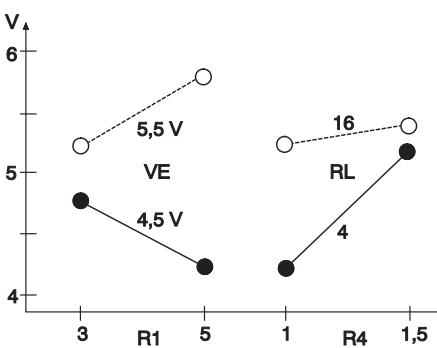


BILD 9-9

Der kleinere S/N bei $R1 = 3 \text{ k}\Omega$ ist darauf zurückzuführen, dass V weniger von der Versorgungsspannung abhängt.

Der kleinere S/N bei $R4 = 1,5 \text{ k}\Omega$ ist darauf zurückzuführen, dass V weniger vom Lastwiderstand abhängt.

Ein solches Ergebnis kann man auch erhalten, wenn man Steuer- und Rauschfaktoren in einem gemeinsamen Versuchsplan mit Auflösung IV oder V verändert, statt sie nach „inner array“ und „outer array“ zu trennen. Die höhere Auflösung erfordert nicht mehr Einzelversuche.

Beispiel

Um 7 Steuerfaktoren im „inner array“ zu untersuchen, benötigt man mindestens 8 Faktorstufenkombinationen. Um 3 Rauschfaktoren zu untersuchen, benötigt man ein „outer array“ mit 4 Faktorstufenkombinationen. Wie in Bild 9-5 dargestellt, benötigt man daher Insgesamt $8 \times 4 = 32$ Einzelversuche, um Steuer- und Rauschfaktoren jeweils in einem Plan der Auflösung III zu verändern.

Fasst man stattdessen alle $7 + 3 = 10$ Faktoren in einem Versuchsplan zusammen, so erhält man mit 32 Einzelversuchen einen Plan der Auflösung IV, mit dessen Hilfe Wechselwirkungen (in Grenzen) bestimmt werden können.

■ 9.4 Anmerkungen zu den „Orthogonalen Feldern“ u. ä.

Speziell für Leser, die über Taguchi zur Versuchsplanung kamen, werden in diesem Abschnitt einige Besonderheiten in der Vorgehensweise von G. Taguchi besprochen, die m. E. problematisch sind.

9.4.1 Orthogonale Felder

Taguchi verwendet eine Liste von Versuchsplänen, die er als „orthogonale Felder“ (orthogonal arrays) bezeichnet. Tabelle 9.7 zeigt am Beispiel des orthogonalen Feldes $L_8(2^7)$, dass, abgesehen von Änderungen

- der Stufen $1 \rightarrow +$ $2 \rightarrow -$
- der Faktoren $1 \rightarrow C$ $2 \rightarrow B$ $3 \rightarrow F$ $4 \rightarrow A$ $5 \rightarrow E$ $6 \rightarrow D$ $7 \rightarrow G$
- und der Reihenfolge der Faktorstufenkombinationen
Vers. Nr. $1 \rightarrow 8a$ $2 \rightarrow 7a$ usw. (Umkehrung),

die orthogonalen Felder für zweistufige Faktoren fraktionelle faktorielle Versuchspläne bzw. Plackett-Burman-Pläne der Auflösung III sind.

In Tabelle 9.8 werden die Bezeichnungen von Taguchi den klassischen Bezeichnungen (FF = fraktionell faktoriell; PB = Plackett-Burman) zugeordnet.

TABELLE 9.7 Vergleich von Taguchis orthogonalem Feld $L_8(2^7)$ mit dem 2^{7-4} -Plan (fraktionell faktoriell)

| Vers. Nr. | Faktor | | | | | | |
|--------------|--------|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| 3 | 1 | 2 | 2 | 1 | 1 | 2 | 2 |
| 4 | 1 | 2 | 2 | 2 | 2 | 1 | 1 |
| 5 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| 6 | 2 | 1 | 2 | 2 | 1 | 2 | 1 |
| 7 | 2 | 2 | 1 | 1 | 2 | 2 | 1 |
| 8 | 2 | 2 | 1 | 2 | 1 | 1 | 2 |

| Vers. Nr. | Faktor | | | | | | |
|--------------|--------|---|---|---|---|---|---|
| | A | B | C | D | E | F | G |
| 1a | - | - | - | + | + | + | - |
| 2a | + | - | - | - | - | + | + |
| 3a | - | + | - | - | + | - | + |
| 4a | + | + | - | + | - | - | - |
| 5a | - | - | + | + | - | - | + |
| 6a | + | - | + | - | + | - | - |
| 7a | - | + | + | - | - | + | - |
| 8a | + | + | + | + | + | + | + |

| Orthogonales Feld | Klassische Bezeichnung |
|-------------------|------------------------|
| $L_4(2^3)$ | FF 2^{3-1} |
| $L_8(2^7)$ | FF 2^{7-4} |
| $L_{12}(2^{11})$ | PB Tabelle 8.17 |
| $L_{16}(2^{15})$ | FF 2^{15-11} |
| $L_{32}(2^{31})$ | FF 2^{31-26} |
| $L_{64}(2^{63})$ | FF 2^{63-57} |

TABELLE 9.8

Klassische Bezeichnungen für die orthogonalen Felder [11]

Die orthogonalen Felder $L_9(3^4)$, $L_{16}(4^5)$ und $L_{25}(5^6)$ für reine mehrstufige Faktoren werden in der klassischen Versuchsplanung als Hyper-Griechisch-Lateinische Quadrate bezeichnet und in Kapitel 13 behandelt. Nur die gemischten Pläne wie z. B. $L_{18}(2^1 \times 3^7)$ sind in der klassischen Versuchsplanung unüblich. Allerdings benutzen ca. 90% der veröffentlichten Taguchi-Beispiele gerade diesen Versuchsplan [10].

Die orthogonalen Felder⁷ von Taguchi sind gesättigte Versuchspläne der Auflösung III, mit all den Risiken, die mit Auflösung III verbunden sind (vgl. Beispiel in Absatz 8.2.4).

Bei gesättigten Plänen ist jeder Effekt mit einem Faktor belegt, der natürlich mit 2FWW, 3FWW usw. vermischt ist. Wegen der großen Anzahl Faktoren ist diese Vermischung bei gesättigten Plänen besonders extrem. So ist z. B. beim Plan $L_{16}(2^{15})$ jeder berechnete Effekt die Summe von 2048 Effekten des vollständigen Plans, davon sieben 2FWW. Die damit verbundenen Risiken sind m.E. nicht mehr überschaubar.

Insbesondere bei mehrstufigen Plänen erscheint es inkonsistent, dass einerseits mehr als zwei Stufen untersucht werden, um Nichtlinearitäten zu erfassen, andererseits aber Wechselwirkungen die Ergebnisse verfälschen können.⁸

9.4.2 Lineare Graphen und Dreieckstabellen

Nach G. Taguchi werden Wechselwirkungen mit Hilfe von Dreieckstabellen oder linearen Graphen berücksichtigt (lineare Graphen enthalten eine Auswahl aus der Information, die in Dreieckstabellen enthalten ist). Dies ist richtig in dem Sinn, dass man damit erreichen kann, dass einzelne 2FWW nicht mit den Effekten der Faktoren vermischt sind. Aber die anderen 2FWW sind weiterhin mit Faktoren vermischt und können diese verfälschen.

Es genügt nicht, dass man an einer Wechselwirkung nicht interessiert ist. Nur wenn sie wirklich klein ist, verfälscht sie nicht andere Effekte (vgl. dazu das Beispiel in Absatz 8.2.4). Und wer weiß schon von vornherein, welche Wechselwirkungen klein sein werden? Wenn man das wüsste, bräuchte man den Versuch oft gar nicht mehr durchzuführen.

Im Übrigen könnte Taguchi auf lineare Graphen und Dreieckstabellen verzichten, hätte er die + - Bezeichnung der Stufen beibehalten. Bild 9-10 zeigt als Beispiel die Dreieckstabelle und die beiden linearen Graphen für $L_8(2^7)$.

Die Zahl 3 in Zeile 1 und Spalte 2 der Dreieckstabelle in Bild 9-10 bzw. an der Verbindungsgeraden zwischen 1 und 2 in den linearen Graphen bedeutet z.B., dass die Wechselwirkung von Faktor 1 mit 2 getrennt von Faktoren ermittelt werden kann, wenn Faktor 3 nicht verwendet wird.

Übersetzt in fraktionelle faktorielle Pläne bedeutet dies, dass die Vorzeichenspalte von Faktor 3 = F das Produkt der Vorzeichenspalten von 1 = C und 2 = B ist, was aus Tabelle 9.7 auch ohne weitere Hilfe sofort erkennbar ist ($F = BC$). Wird BC nicht mit einem Faktor belegt, so kann diese 2FWW natürlich getrennt von Faktoren ermittelt werden.

⁷ Die Kritik richtet sich nicht gegen die Orthogonalität. Alle guten Versuchspläne sind orthogonal (oder weichen zumindest nur wenig von der Orthogonalität ab, siehe dazu auch Kapitel 11). Taguchi verwendet aber nur einen Bruchteil der orthogonalen Pläne, und die Kritik richtet sich gegen die von Taguchi getroffene Auswahl.

⁸ Entwickelt man die Zielgröße in Abhängigkeit von den Faktoren als Potenzreihe, so sind sowohl Wechselwirkungen als auch quadratische Terme die Produkte zweier Faktoren (vgl. Kapitel 11).

| Faktor | Faktor | | | | | | |
|--------|--------|-----|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | (1) | 3 | 2 | 5 | 4 | 7 | 6 |
| 2 | | (2) | 1 | 6 | 7 | 4 | 5 |
| 3 | | | (3) | 7 | 6 | 5 | 4 |
| 4 | | | | (4) | 1 | 2 | 3 |
| 5 | | | | | (5) | 3 | 2 |
| 6 | | | | | | (6) | 1 |
| 7 | | | | | | | (7) |

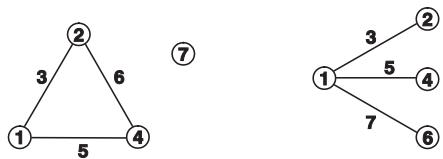


BILD 9-10 Dreieckstabelle und lineare Graphen für $L_8(2^7)$.

Aber Achtung: Selbst wenn man die Spalte 3 nicht mit einem Faktor belegt, erhält man einen 2^{6-3} -Plan mit Auflösung III, der durch die Zuordnungen AB, AC und ABC definiert ist. Verwendet man die Ziffern wie bei Taguchi, so ist z.B. der Effekt von 4 vermengt mit den 2FWW 26 und 15. Dabei genügt es nicht, dass man an diesen Wechselwirkungen nicht interessiert ist. Falls sie existieren, verfälschen sie auf jeden Fall die Schätzung des Effektes 4.

Hier wird vorgeschlagen, Pläne der Auflösung IV zu verwenden. Dann ist kein Faktor mit 2FWW vermengt und man kann auf Dreieckstabellen völlig verzichten. Aus diesem Grund werden sie auch nicht weiter behandelt.

Ordnet man Wechselwirkungen dagegen einzeln mit Hilfe von Dreieckstabellen oder linearen Graphen zu, so besteht das Risiko, dass man einen Plan mit Auflösung III erhält, obwohl mit der gleichen Anzahl Faktoren Auflösung IV möglich wäre. So hat z.B. der 2^{4-1} -Plan, der mit den Faktoren 1, 2, 4 und 6 aus dem rechten linearen Graphen von Bild 9-10 entsteht, nur Auflösung III, obwohl es einen 2^{4-1} -Plan mit Auflösung IV gibt (Tabelle 8.13).

9.4.3 Dummy Levels, Pseudo Factor Designs, Idle Columns

Hinter diesen und ähnlichen Schlagworten verborgen sich diverse Techniken von G. Taguchi, um Faktoren mit einer anderen Anzahl von Stufen in einem Versuchsplan mit zu behandeln. Vom Einsatz dieser Verfahren wird abgeraten.

Werden individuellen Bedürfnissen angepasste Versuchspläne benötigt, so wird empfohlen, optimale Versuchspläne zu verwenden (Abschnitt 11.2).

Literatur

- [1] Taguchi, G./Wu, Y.: Introduction to Off-line Quality Control. Central Japan Quality Control Association, Nagoya 1985
- [2] Taguchi, G./Konishi, S.: Taguchi Methods Orthogonal Arrays and Linear Graphs, Tools for Quality Engineering. ASI Press, 1987

- [3] *Phadke, M. S.*: Quality Engineering Using Robust Design. Prentice Hall, London 1989
- [4] *Fowlkes, W. Y./Creveling, C. M.*: Engineering Methods for Robust Product Design. Addison-Wesley, Reading 1995
- [5] *Ross, P.J.*: Taguchi Techniques for Quality Engineering. McGraw-Hill, New York, 2. Auflage 1996
- [6] *Klein, B.*: Versuchsplanung – DoE: Einführung in die Taguchi/Shainin-Methodik. Oldenbourg Verlag, München 4. Auflage 2014
- [7] Journal of Quality Technology, Vol. 17, 4 (special issue), 1985
- [8] *Montgomery, D. C.*: Design and Analysis of Experiments. John Wiley, New York, 8. Auflage 2012
- [9] *Franzkowski, R.*: „Factorial Designs Estimate Location and Dispersion“, in: Transactions of the ASQC Annual Technical Conference, San Francisco 1981
- [10] *Logothetis, N./Wynn, H. P.*: Quality through Design. Clarendon Press, Oxford 1994
- [11] *Wu, C. F. J./Hamada, M.*: Experiments: Planning, Analysis, and Optimization. John Wiley, New York, 2. Auflage 2009

10

Regressionsanalyse

Mit Hilfe der Regressionsanalyse wird ein mathematisches Modell für den Zusammenhang zwischen Einflussgröße(n) und Zielgröße an vorhandene Daten angepasst (z.B. [1], [2]). Diese Anpassung ist möglich, unabhängig davon, ob eine Einflussgröße gezielt als Faktor verändert wurde, oder ob die Veränderung der Einflussgröße nur beobachtet wurde. Daher wird in diesem Kapitel der Oberbegriff „Einflussgröße“ verwendet.

Die mathematische Form des Zusammenhangs muss vorgegeben sein. Parameter in dieser Form werden dann so angepasst, dass man die „bestmögliche“ Beschreibung der Daten erhält.



Beispiel

Bei der galvanischen Abscheidung einer Schicht erwartet man näherungsweise eine lineare Zunahme der Schichtdicke mit der Abscheidezeit (evtl. abgesehen von einer gewissen Zeitverzögerung am Anfang). Sechs Teile werden nach verschiedenen Zeiten aus der Anlage entnommen. Bild 10-1 zeigt die gemessene Schichtdicke in Abhängigkeit von der Zeit in der Anlage.

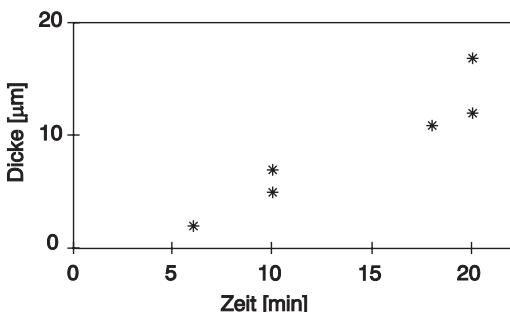


BILD 10-1

Messwerte für die Dicke einer galvanisch abgeschiedenen Schicht für verschiedene Werte der Abscheidezeit

Gesucht ist die Gleichung der Geraden, die die Abhängigkeit der Schichtdicke von der Zeit am besten beschreibt. Sie wird mit einfacher linearer Regression bestimmt.

Die Vorgehensweise kann auf beliebig viele Einflussgrößen und auf nichtlineare Zusammenhänge erweitert werden, soweit sie durch eine geeignete Transformation linearisiert werden können.

■ 10.1 Einfache lineare Regression

Ausgangspunkt der einfachen linearen Regression sind N Messwerte (nummeriert $i = 1, 2, \dots, N$) für eine Zielgröße y in Abhängigkeit von einer Einflussgröße x . Aus technischen Gründen vermutet man einen linearen Zusammenhang zwischen x und y , wie z.B. in Bild 10-1, von dem die Messwerte nur zufällig um ε_i abweichen.

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i \quad (10.1)$$

Im Folgenden wird zunächst gezeigt, wie man Schätzwerte b_0 und b_1 für die Parameter β_0 und β_1 erhält und wie man die erreichte Anpassung grafisch beurteilen kann. Diese deskriptiven Verfahren sind in jedem Fall einsetzbar. Wenn bestimmte Voraussetzungen erfüllt sind, ist darüber hinaus eine statistische Beurteilung möglich. Verfahren dazu werden in Absatz 10.1.4 beschrieben.

10.1.1 Methode der kleinsten Quadrate

Für feste Werte b_0 und b_1 kann man für jeden Wert der Einflussgröße x_i einen Schätzwert \hat{y}_i der Zielgröße berechnen.

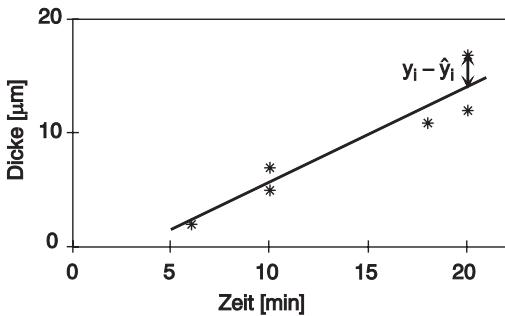
$$\hat{y}_i = b_0 + b_1 \cdot x_i \quad (10.2)$$

b_0 und b_1 werden nun so bestimmt, dass die Summe der quadrierten Abweichungen (S.d.q.A.) zwischen den Schätzwerten \hat{y}_i und den Messwerten y_i über alle N Messwerte so klein wie möglich ist, d.h. dass¹

$$\sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N (y_i - (b_0 + b_1 \cdot x_i))^2 \rightarrow \text{Minimum.} \quad (10.3)$$

Dieses Verfahren heißt Methode der kleinsten Quadrate. Bild 10-2 zeigt dieselben Daten wie Bild 10-1. Zusätzlich ist in Bild 10-2 die so ermittelte Schätzgerade angegeben. Die Bedingung (10.3) bedeutet anschaulich, dass die Gerade so durch die Punkte gelegt wird, dass die Summe der Quadrate der mit einem Pfeil gekennzeichneten Abstände über alle Punkte so klein wie möglich ist – man minimiert die Abstände der Messwerte von der Geraden in **y-Richtung**. Dies ist sinnvoll, wenn man mit der Geradengleichung für vorgegebene x -Werte Schätzwerte für y bestimmen möchte (und nicht umgekehrt).

¹ Die Visualisierung „Lineare Regression“ stellt die Summe der quadrierten Abweichungen grafisch dar und erlaubt es, eine Gerade zu raten und mit der „besten Geraden“ zu vergleichen – vgl. Anhang G.

**BILD 10-2**

Messwerte aus Bild 10-1, zusammen mit der Schätzgeraden: Die Summe der Quadrate der mit einem Pfeil gekennzeichneten Abstände $y_i - \hat{y}_i$ ist minimiert.

Aus der Bedingung (10.3) können folgende Berechnungsformeln für b_0 und b_1 abgeleitet werden:

$$b_1 = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2} = \frac{Q_{xy}}{Q_{xx}} \quad (10.4)$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x} \quad (10.5)$$

wobei

$$\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^N x_i = \text{Mittelwert der } x\text{-Werte} \quad (10.6)$$

$$\bar{y} = \frac{1}{N} \cdot \sum_{i=1}^N y_i = \text{Mittelwert der } y\text{-Werte} \quad (10.7)$$

$$Q_{xx} = \sum_{i=1}^N (x_i - \bar{x})^2 = \text{S.d.q.A der } x\text{-Werte} \quad (10.8)$$

$$Q_{xy} = \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y}) \quad (10.9)$$

Q_{xy} ist ein Maß dafür, wie stark sich x- und y-Werte gemeinsam verändern. Sind sie unabhängig voneinander, so ist $Q_{xy} = 0$.

Der Schätzwert b_1 für die Steigung der Geraden ist somit das Verhältnis aus einer Größe, die angibt, wie stark sich x- und y-Werte gemeinsam verändern und einer Größe, die angibt, wie stark die x-Werte streuen. Den Schätzwert b_0 für den Achsenabschnitt der Geraden erhält man, indem man vom Punkt \bar{x}, \bar{y} mit der Steigung b_1 zu $x = 0$ geht.

**Beispiel**

Tabelle 10.1 zeigt die in Bild 10-2 dargestellten Werte und die Rechenschritte, die zu den Schätzwerten für b_0 und b_1 führen.

TABELLE 10.1 Messwerte von Bild 10-1 bzw. 10-2 mit den daraus berechneten Hilfsgrößen Q

| Zeit x_i [min] | Dicke y_i [μm] | $x_i - \bar{x}$ | $y_i - \bar{y}$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x}) \cdot (y_i - \bar{y})$ | $(y_i - \bar{y})^2$ |
|------------------|-------------------------------|-----------------|-----------------|---------------------|---|---------------------|
| 6 | 2 | -8 | -7 | 64 | 56 | 49 |
| 10 | 5 | -4 | -4 | 16 | 16 | 16 |
| 10 | 7 | -4 | -2 | 16 | 8 | 4 |
| 18 | 11 | 4 | 2 | 16 | 8 | 4 |
| 20 | 17 | 6 | 8 | 36 | 48 | 64 |
| 20 | 12 | 6 | 3 | 36 | 18 | 9 |
| $\Sigma = 84$ | $\Sigma = 54$ | $(\Sigma = 0)$ | $(\Sigma = 0)$ | $\Sigma = 184$ | $\Sigma = 154$ | $\Sigma = 146$ |
| $\bar{x} = 14$ | $\bar{y} = 9$ | | | $= Q_{xx}$ | $= Q_{xy}$ | $= Q_{yy}$ |

Man beachte, dass in diesem Beispiel $(x_i - \bar{x}) \cdot (y_i - \bar{y})$ immer positiv ist, da bei niedrigen x-Werten auch y niedrig ist (und umgekehrt), d.h. x und y verändern sich gemeinsam. Mit (10.4) und (10.5) erhält man

$$b_1 = \frac{Q_{xy}}{Q_{xx}} = \frac{154}{184} = 0,837$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x} = 9 - 0,837 \cdot 14 = -2,72$$

Damit lautet die Gleichung $\hat{y} = -2,72 + 0,837 \cdot x$.

In Bild 10-2 beträgt die Steigung der Geraden $b_1 = 0,837$, d.h. pro Minute nimmt die Schichtdicke (im Mittel) um $0,837 \mu\text{m}$ zu. Der Achsenabschnitt $b_0 = -2,72$ ist der Schnittpunkt der Regressionsgeraden mit der y-Achse ($x = 0$). Der negative Wert bedeutet natürlich nicht, dass die Schichtdicke anfangs negativ ist. Er ergibt sich aus der Zeitverzögerung zwischen dem Anschalten der Anlage und dem eigentlichen Beginn der Abscheidung. Die Gleichung gilt nur im untersuchten Zeitbereich, in dem die berechnete Schichtdicke positiv ist (Extrapolation ist nicht zulässig).

10.1.2 Bestimmtheitsmaß und Korrelationskoeffizient

Mit der Methode der kleinsten Quadrate kann man formal an beliebige Daten eine Gerade anpassen. Um die Güte der Anpassung zu beurteilen, zerlegt man die Summe der quadrierten Abweichungen (S.d.q.A.) der y-Werte in einen Anteil der Regressionsgeraden und eine Abweichung von der Regressionsgeraden:

$$\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (10.10)$$

$$Q_{\text{Gesamt}} = Q_{\text{Regression}} + Q_{\text{Rest}} \quad (10.11)$$

Die Methode der kleinsten Quadrate besteht darin, Q_{Rest} zu minimieren. In Analogie zu (10.8) und (10.9) führt man folgende Bezeichnung ein:

$$Q_{\text{Gesamt}} = Q_{yy} = \sum_{i=1}^N (y_i - \bar{y})^2 \quad (10.12)$$

Bestimmtheitsmaß

Die Anpassung ist umso besser, je größer der Anteil der S.d.q.A. ist, der durch die Regressionsgerade erklärt wird. Dazu verwendet man das Bestimmtheitsmaß B ($= r^2$, englisch r-squared – diese Bezeichnung ist in Software üblich):

$$\text{Bestimmtheitsmaß } B = \frac{Q_{\text{Regression}}}{Q_{\text{Gesamt}}} = \frac{Q_{xy}^2}{Q_{xx} \cdot Q_{yy}} \quad (10.13)$$

Der erste Teil von (10.13) ist die Definition, der zweite Teil erlaubt die praktische Berechnung.

Da $Q_{\text{Regression}} \leq Q_{\text{Gesamt}}$ ist, gilt immer

$$0 \leq B \leq 1 \quad (10.14)$$

Korrelationskoeffizient

Alternativ zum Bestimmtheitsmaß wird zur Beschreibung der Anpassung auch der Korrelationskoeffizient r verwendet:

$$\text{Korrelationskoeffizient } r = (\text{Vorzeichen von } b_1) \cdot \sqrt{B} \quad (10.15)$$

Der Korrelationskoeffizient r liegt immer im Bereich

$$-1 \leq r \leq 1 \quad (10.16)$$



Beispiel

Im Beispiel von Tabelle 10.1 erhält man

$$\text{Bestimmtheitsmaß } B = \frac{Q_{xy}^2}{Q_{xx} \cdot Q_{yy}} = \frac{154^2}{184 \cdot 146} = 0,883.$$

D.h. knapp 90 % der gesamten Summe der quadrierten Abweichungen werden durch die Regression erklärt.

$$\text{Korrelationskoeffizient } r = (\text{Vorzeichen von } b_1) \cdot \sqrt{B} = +\sqrt{0,883} = 0,940.$$

Beurteilung des Korrelationskoeffizienten

$B = 1$ bzw. $r = \pm 1$ bedeutet, dass alle Punkte genau auf einer Geraden liegen. Bei der Beurteilung des Korrelationskoeffizienten ist zu berücksichtigen, wie groß die Abweichung von ± 1 ist und welche Ursache sie hat.

Abweichungen von $r = \pm 1$ können auf Zufallsstreuung zurückzuführen sein, aber auch auf einen nichtlinearen Zusammenhang zwischen x- und y-Werten. Bild 10-3 zeigt dies für verschiedene Werte des Korrelationskoeffizienten. Der Wert des Korrelationskoeffizienten allein macht keine Aussage über die Güte der Anpassung. Daher sollte man die x- und y-Werte immer grafisch darstellen.

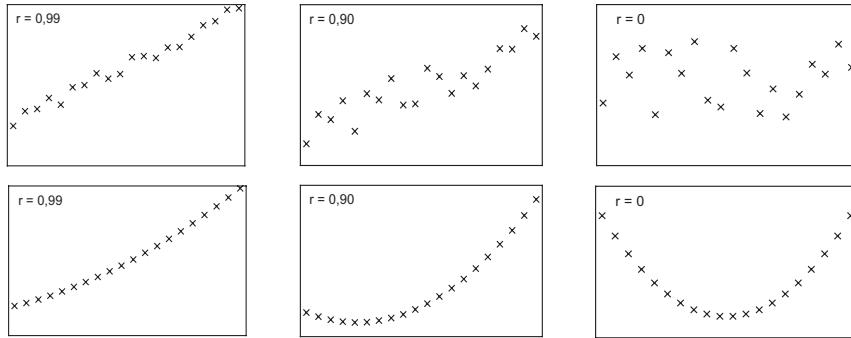


BILD 10-3 Grafische Darstellung von Daten mit Korrelationskoeffizienten $r = 0,99; 0,90$ und 0 .
Obere Zeile: Zufallsstreuung
Untere Zeile: nichtlinearer Zusammenhang

Ein Korrelationskoeffizient nahe 0 kann viele Ursachen haben, wie z. B.:

- die Zielgröße hängt nicht oder nur wenig von der Einflussgröße ab
- der untersuchte Wertebereich der Einflussgröße ist klein
- die Zufallsstreuung der Zielgröße ist groß
- die Abhängigkeit von der betrachteten Einflussgröße wird von anderen Einflussgrößen verdeckt
- die Abhängigkeit ist stark nichtlinear.



Aufgabe

Es wird vermutet, dass die Lebensdauer t (in h) eines elektronischen Bauelements linear vom Leckstrom I_0 (in μA) abhängt. Vier Bauelemente mit unterschiedlichem Leckstrom werden einem beschleunigten Lebensdauertest unterzogen. Folgende Ergebnisse wurden erzielt:²

| Nr. i | $I_0 = x_i$ | $t = y_i$ | $x_i - \bar{x}$ | $y_i - \bar{y}$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x}) \cdot (y_i - \bar{y})$ | $(y_i - \bar{y})^2$ |
|-------------|---------------------------------------|---------------------------------------|-----------------|-----------------|--------------------------------|---|--------------------------------|
| 1 | 40 | 80 | | | | | |
| 2 | 50 | 64 | | | | | |
| 3 | 55 | 62 | | | | | |
| 4 | 75 | 34 | | | | | |
| $N = \dots$ | $\Sigma = \dots$ $\bar{x} = \dots$ | $\Sigma = \dots$ $\bar{y} = \dots$ | | | $\Sigma = \dots$ $= Q_{xx}$ | $\Sigma = \dots$ $= Q_{xy}$ | $\Sigma = \dots$ $= Q_{yy}$ |

² Anmerkung: In einem realistischen Versuch würde man wesentlich mehr Teile untersuchen. Hier soll nur das Prinzip der Rechnung geübt werden.

Vervollständigen Sie die Tabelle, und berechnen Sie die Gleichung der Regressionsgeraden, das Bestimmtheitsmaß und den Korrelationskoeffizienten. Skizzieren Sie die Regressionsgerade und die Versuchsergebnisse.

Lösung

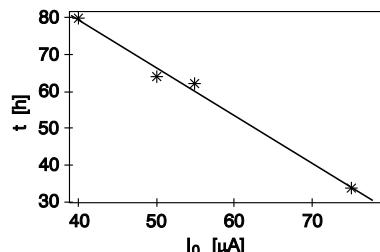
| Nr. i | $I_0 = x_i$ | $t = y_i$ | $x_i - \bar{x}$ | $y_i - \bar{y}$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x}) \cdot (y_i - \bar{y})$ | $(y_i - \bar{y})^2$ |
|--------------|----------------------------------|----------------------------------|-----------------|-----------------|------------------------------|---|-------------------------------|
| 1 | 40 | 80 | -15 | 20 | 225 | -300 | 400 |
| 2 | 50 | 64 | -5 | 4 | 25 | -20 | 16 |
| 3 | 55 | 62 | 0 | 2 | 0 | 0 | 4 |
| 4 | 75 | 34 | 20 | -26 | 400 | -520 | 676 |
| N = 4 | $\Sigma = 220$ $\bar{x} = 55$ | $\Sigma = 240$ $\bar{y} = 60$ | | | $\Sigma = 650$ $= Q_{xx}$ | $\Sigma = -840$ $= Q_{xy}$ | $\Sigma = 1096$ $= Q_{yy}$ |

Gleichung der Regressionsgeraden:

$$b_1 = \frac{Q_{xy}}{Q_{xx}} = \frac{-840}{650} = -1,2923 \quad b_0 = \bar{y} - b_1 \bar{x} = 131,1 \quad \Rightarrow \quad \hat{y} = 131 - 1,29 \cdot x$$

Bestimmtheitsmaß: $B = \frac{Q_{xy}^2}{Q_{xx} Q_{yy}} = \frac{840^2}{650 \cdot 1096} = 0,9905$

Korrelationskoeffizient: $r = (\text{Vorzeichen von } b_1) \cdot \sqrt{B} = -\sqrt{0,9905} = -0,9952$



Versuchsergebnisse und angepasste Regressionsgerade:
Der negative Korrelationskoeffizient (bzw. die negative Steigung b_1) bedeutet, dass die y-Werte mit zunehmenden x-Werten abnehmen.

10.1.3 Grafische Beurteilung der Residuen

Die Abweichungen ($y_i - \hat{y}_i$) der Messwerte von der Regressionsgeraden (oder allgemeiner vom angepassten Modell) heißen Residuen. Die Residuen geben wichtige Hinweise zur Güte der Anpassung. Nach jeder Regression sollten die Residuen auf vier verschiedene Arten dargestellt werden. Gute Software bietet alle diese Darstellungen an³ (von Hand sehr zeitaufwändig). Die grafische Beurteilung der Residuen ist eine Ergänzung der bereits empfohlenen Darstellung der x- und y-Werte.

³ Statt der Residuen werden häufig standardisierte oder studentisierte Residuen dargestellt [3]:

Man erhält sie, indem man die Residuen durch ihre geschätzte Standardabweichung teilt. Bei extern studentisierten bzw. entfernten Residuen wird der betrachtete Punkt nicht bei der Modellbildung verwendet. So werden Ausreißer leichter erkannt. Standardisierte oder studentisierte Residuen, die absolut größer als ca. 3 sind, treten nur selten zufällig auf. Sie verdienen besondere Aufmerksamkeit.

Jede Darstellung beantwortet eine andere Frage:

1. Ist das verwendete mathematische Modell konsistent mit den Daten?

hier: Ist der lineare Ansatz konsistent mit den Daten?

Diese Frage wurde im Prinzip bereits mit der Darstellung der y-Werte gegen die x-Werte wie in Bild 10-3 überprüft. Da die Residuen normalerweise klein sind, kann man auch bereits kleinere Abweichungen erkennen, wenn man statt der y-Werte die Residuen gegen x (den Wert der Einflussgröße) aufträgt.

Bild 10-4 zeigt, worauf man bei der Beurteilung achten sollte. Die obere Darstellung zeigt jeweils, wie das Ergebnis aussehen sollte (eine Punktwolke bzw. ein waagrechtes Band). Die untere Darstellung zeigt, wie das Ergebnis z.B. aussehen kann, wenn das verwendete Modell die Daten nicht richtig beschreibt: Man erkennt eine Abhängigkeit der Residuen von der Einflussgröße x.

Die Verteilung der Werte in x-Richtung ist nicht wichtig. Sie hängt davon ab, welche x-Werte untersucht wurden. Bei geplanten Versuchen treten nur ganz bestimmte x-Werte auf (die Stufen des Faktors x), bei ungeplanten Versuchen treten dagegen beliebige x-Werte auf.

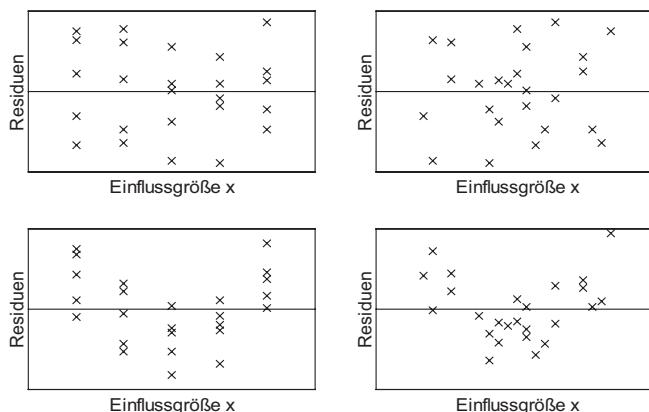


BILD 10-4 Darstellung der Residuen ($y_i - \hat{y}_i$) gegen den Wert der Einflussgröße x:

Links: Geplante Versuche (nur bestimmte x-Werte)

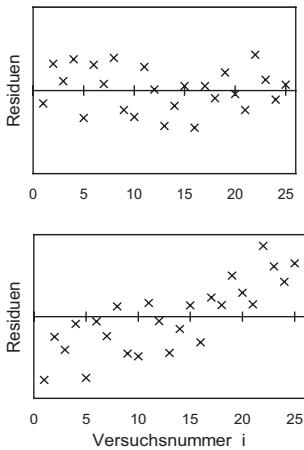
Rechts: Ungeplante Versuche (beliebige x-Werte)

Oben: Nur zufällige Abweichungen vom Modell (waagrechtes Band)

Unten: Modell beschreibt die funktionale Form der Abhängigkeit nicht richtig

2. Gibt es einen Trend in den Versuchsergebnissen?

Dies kann man überprüfen, indem man die Residuen in der Reihenfolge der Versuchsdurchführung aufträgt. Die obere Darstellung in Bild 10-5 zeigt, wie das Ergebnis aussehen sollte: ein waagrechtes Band. Die untere Darstellung zeigt, wie das Ergebnis z.B. aussehen kann, wenn ein Trend vorliegt. Andere, nicht erfasste Einflüsse führen zu einer allmählichen Veränderung der Ergebnisse mit der Zeit (z.B. Werkzeugverschleiß, Änderung der Umgebungstemperatur o.ä.).

**BILD 10-5**

Darstellung der Residuen gegen die Versuchsnummer i
(d. h. in der Reihenfolge der Versuchsdurchführung):

Oben: Kein Trend ist erkennbar, die Residuen liegen in einem waagrechten Band.

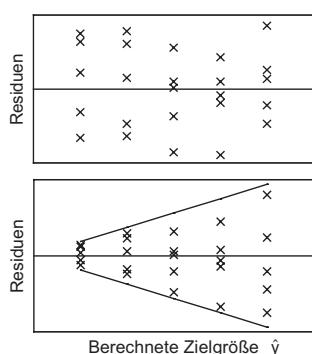
Unten: Trend, die Residuen nehmen mit der Versuchsnummer zu.

Ein Trend kann nur erkannt werden, wenn die Reihenfolge der Einzelversuche zufällig war (Randomisierung). Ohne Randomisierung verfälscht ein evtl. vorhandener Trend die Abhängigkeit der Zielgröße von der Einflussgröße. Er ist dann an den Residuen nicht mehr erkennbar.

3. Hängt die Standardabweichung vom y-Wert ab?

Eine Voraussetzung für die statistischen Analysen im nächsten Abschnitt ist, dass die Standardabweichung unabhängig vom y -Wert ist. Manchmal nimmt die Standardabweichung jedoch mit zunehmendem y -Wert zu (die Standardabweichung kann z. B. ein fester Prozentwert des Mittelwertes sein).

Dies kann man überprüfen, indem man die Residuen gegen den aus dem Modell berechneten y -Wert aufträgt. Die obere Darstellung in Bild 10-6 zeigt, wie das Ergebnis aussehen sollte (eine Punktewolke bzw. ein waagrechtes Band). Die untere Darstellung zeigt, wie das Ergebnis z. B. aussehen kann, wenn die Standardabweichung mit dem y -Wert zunimmt (die Residuen bilden einen Trichter, ihre Standardabweichung nimmt mit dem y -Wert zu). In diesem Fall erzielt man häufig eine Verbesserung, indem man als Zielgröße $\log(y)$ oder eine andere Transformation verwendet (vgl. Abschnitt 6.4).

**BILD 10-6**

Darstellung der Residuen gegen den berechneten Wert der Zielgröße \hat{y}

Oben: Keine Veränderung der Standardabweichung mit dem berechneten y -Wert ist erkennbar, die Residuen liegen in einem waagrechten Band.

Unten: Die Standardabweichung nimmt mit dem berechneten y -Wert zu, die Residuen liegen in einem Trichter.

4. Sind die Residuen normalverteilt?

(genauer: konsistent mit der Normalverteilung)

Dies kann man durch Eintragung der Residuen im Wahrscheinlichkeitsnetz überprüfen. Normalverteilte Merkmalswerte liegen näherungsweise auf einer Geraden (vgl. Absatz 6.2.2 und 6.3.3).

10.1.4 Vertrauensbereiche und Signifikanz

Voraussetzungen

Die Anpassung einer „besten Geraden“ und die grafische Darstellung der Residuen sind immer möglich und nützlich. Unter der Voraussetzung, dass

- für jeden x -Wert die y -Werte normalverteilt sind
- ein linearer Zusammenhang zwischen x und dem Mittelwert der Verteilung der y -Werte besteht, d. h. $\mu_y(x) = \beta_0 + \beta_1 \cdot x$
- die zufälligen Abweichungen ϵ unabhängig voneinander sind (kein Trend o. ä.) und
- die Standardabweichung σ_ϵ konstant (d. h. unabhängig von x) ist,

sind die mit (10.4) und (10.5) berechneten Schätzwerte b_0 und b_1 für die Parameter β_0 und β_1 erwartungstreu (d. h. der Mittelwert für b_0 und b_1 über viele Versuche stimmt mit β_0 und β_1 überein). Außerdem können dann Vertrauensbereiche für b_0 und b_1 berechnet werden, und die Signifikanz des Zusammenhangs zwischen x und y kann überprüft werden. Bild 10-7 stellt die Voraussetzungen grafisch dar.

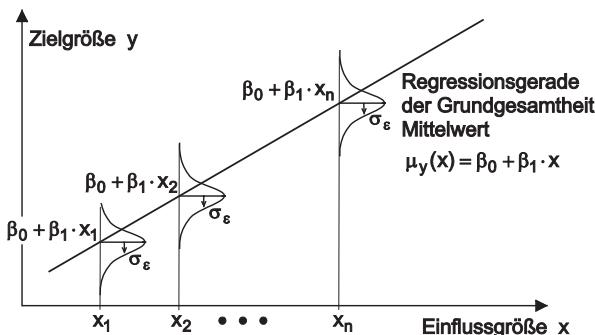


BILD 10-7

Grafische Darstellung der Voraussetzungen für die Berechnung von Vertrauensbereichen und für die Überprüfung der Signifikanz des Zusammenhangs zwischen x und y

Vertrauensbereich für die Steigung

Wie in (6.6) ist die Varianz der Residuen oder

$$\text{Varianz der Reststreuung } s_R^2 = \frac{Q_{\text{Rest}}}{N-2} = \frac{(1-B) \cdot Q_{yy}}{N-2} \quad (10.17)$$

ein Schätzwert für die Varianz σ_ϵ^2 der Streuung um die Regressionsgerade mit

$$\text{Freiheitsgrad } f = N - 2. \quad (10.18)$$

Der Freiheitsgrad beträgt $N - 2$, da zwei Schätzwerte b_0 und b_1 aus den N Daten berechnet werden, statt nur des einen Mittelwerts \bar{y} in (6.6). Mit den t-Werten aus Tabelle 6.4 erhält man als zweiseitigen Vertrauensbereich für die Steigung der Regressionsgeraden zum Vertrauensniveau $1-\alpha$:

$$b_1 - \frac{t \cdot s_R}{\sqrt{Q_{xx}}} \leq \beta_1 \leq b_1 + \frac{t \cdot s_R}{\sqrt{Q_{xx}}} \quad (10.19)$$

Ein Vergleich dieser Formel mit dem Vertrauensbereich für den Mittelwert von N normalverteilten Einzelwerten

$$\bar{y} - \frac{t \cdot s}{\sqrt{N}} \leq \mu \leq \bar{y} + \frac{t \cdot s}{\sqrt{N}} \quad (\equiv 6.9')$$

zeigt eine große Ähnlichkeit und macht (10.19) plausibel:

- Der Vertrauensbereich ist symmetrisch zum berechneten Schätzwert (\bar{y} bzw. b_1) für die gesuchte Größe (μ bzw. β_1).
- s_R bzw. s ist die aus den Daten berechnete Zufallsstreuung.
- Die t-Werte unterscheiden sich nur im Freiheitsgrad.
- Der Faktor \sqrt{N} wird ersetzt durch $\sqrt{Q_{xx}}$, wobei

$$Q_{xx} = \sum_{i=1}^N (x_i - \bar{x})^2 \quad (\equiv 10.8)$$

Q_{xx} enthält N ähnliche Summanden und nimmt somit näherungsweise wie N zu. Die Summanden in Q_{xx} nehmen quadratisch mit der Breite des Bereichs der x -Werte zu, die Wurzel nimmt daher linear mit dieser Breite zu. Wie Bild 10-8 verdeutlicht, ist es plausibel, dass der Vertrauensbereich der Steigung schmäler ist, wenn der untersuchte Bereich der x -Werte breiter ist.

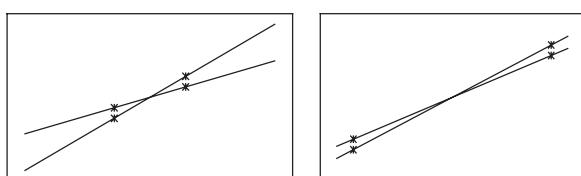


BILD 10-8

Je breiter der Bereich der x -Werte ist, desto kleiner ist der Bereich möglicher Steigungen (bei gleicher Streuung der y -Werte).

Signifikanz

Analog zur Vorgehensweise in Kapitel 6 spricht man von einem zufälligen, indifferenten, signifikanten oder hochsignifikanten Ergebnis, je nachdem, welche Vertrauensbereiche den Wert 0 enthalten (bzw. wie groß die Steigung verglichen mit der Breite des 95 %-, 99 %- und 99,9 %-Vertrauensbereichs für die Steigung ist).⁴

⁴ Häufig wird auch Varianzanalyse (Kapitel 12) für die Beurteilung der Signifikanz verwendet. Sie basiert auf der Streuungszerlegung (10.10) und führt zu identischen Ergebnissen.

**Beispiel**

Im Beispiel von Tabelle 10.1 erhält man:

$$\text{Varianz der Reststreuung } s_R^2 = \frac{(1-B) \cdot Q_{yy}}{N-2} = \frac{(1-0,883) \cdot 146}{6-2} = 4,28$$

und damit $s_R = \sqrt{4,28} = 2,07$.

Mit Tabelle 6.4 erhält man für die Breite der Vertrauensbereiche:

$$95\%: 2,776 \cdot 2,07 / \sqrt{184} = 0,42$$

$$99\%: 4,604 \cdot 2,07 / \sqrt{184} = 0,70$$

$$99,9\%: 8,610 \cdot 2,07 / \sqrt{184} = 1,31.$$

Der Vergleich mit dem Schätzwert $b_1 = 0,837$ für die Steigung ergibt:

Es besteht ein signifikanter Zusammenhang zwischen der Zeit und der Dicke (**).

Der 95%-Vertrauensbereich für die Steigung beträgt: $0,42 \leq b_1 \leq 1,26$.

Achtung

Ein signifikantes Ergebnis ist kein Beweis dafür, dass die Abhängigkeit linear ist. Man hat nur gezeigt, dass ein linearer Anteil existiert. Die hier durchgeführte Analyse unterscheidet nicht zwischen den zufälligen Abweichungen oben in Bild 10-3 und den systematischen Abweichungen unten in Bild 10-3.

Die Linearität wurde hier vorausgesetzt. Soll überprüft werden, ob die Daten konsistent mit der Linearität sind, können die Residuen wie in Bild 10-4 grafisch beurteilt werden. Alternativ kann ein rechnerischer Linearitätstest ([1], nach Fisher, auch Lack-of-Fit-Test genannt) durchgeführt werden, der auf Ideen der Varianzanalyse aufbaut (Kapitel 12).

Vertrauensbereich für den Mittelwert $\mu_y(x)$

Der Vertrauensbereich für den Mittelwert $\mu_y(x)$ der Verteilung der y-Werte für einen beliebigen x-Wert im untersuchten Bereich ist

$$\hat{y} - t \cdot s_R \cdot \sqrt{\frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} \leq \mu_y(x) \leq \hat{y} + t \cdot s_R \cdot \sqrt{\frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} \quad (10.20)$$

Mit $x = 0$ erhält man den Vertrauensbereich für den Achsenabschnitt b_0 .

Auch diese Formel ist plausibel:

- $\hat{y} = b_0 + b_1 \cdot x$ ist der aus den Daten berechnete Schätzwert für den gesuchten Mittelwert $\mu_y(x)$ der Verteilung der y-Werte für diesen Wert x.
- Für s_R und t vgl. (10.19).
- Für $x = \bar{x}$ nimmt die Wurzel ihren kleinsten Wert $1/\sqrt{N}$ an, der Vertrauensbereich für den Mittelwert der y-Werte entspricht damit dem Vertrauensbereich für den Mittelwert (6.9').

- Für $x \neq \bar{x}$ ist der Vertrauensbereich breiter, weil auch die Steigung nur innerhalb ihres Vertrauensbereichs bekannt ist. Wie bei Varianzen werden die Quadrate der beiden Beiträge unter der Wurzel addiert. Bild 10-9 verdeutlicht die Verbreiterung des Vertrauensbereichs für $\mu_y(x)$.

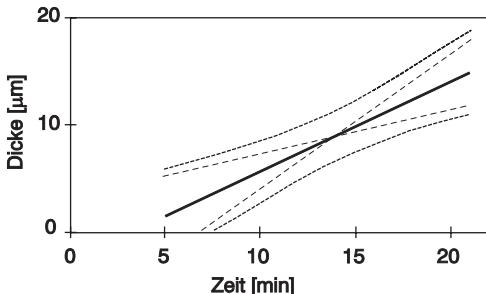


BILD 10-9

95%-Vertrauensbereich für $\mu_y(x)$ zusammen mit den Geraden durch den Punkt $\bar{x} = 14$, $\bar{y} = 9$ mit Steigungen 0,42 und 1,26 (untere und obere Grenze des 95%-Vertrauensbereichs)

Vorhersagebereich für Einzelwerte y

Ein zukünftiger Einzelwert der Zielgröße y liegt mit Wahrscheinlichkeit $1 - \alpha$ im Vorhersagebereich

$$\hat{y} - t \cdot s_R \cdot \sqrt{1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} \leq y(x) \leq \hat{y} + t \cdot s_R \cdot \sqrt{1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} \quad (10.21)$$

Der Vorhersagebereich für Einzelwerte ist breiter als der Vertrauensbereich für den Mittelwert. Mit zunehmender Anzahl N der Einzelversuche wird der Vertrauensbereich immer schmäler (wie $1/\sqrt{N}$), der Vorhersagebereich für Einzelwerte dagegen erreicht eine Grenze, die durch die Standardabweichung der Einzelwerte bestimmt ist.

Bild 10-10 zeigt den Vertrauensbereich für den Mittelwert und den Vorhersagebereich für einen Einzelwert zum Vertrauensniveau von 95%⁵.

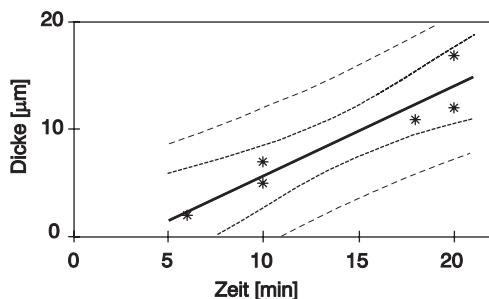


BILD 10-10

Vertrauensbereich für den Mittelwert (innen) und Vorhersagebereich für einen Einzelwert (außen) zum Vertrauensniveau von 95% für das Beispiel in Tabelle 10.1

⁵ Mit der Visualisierung „Lineare Regression“ können Sie Punkte verändern und die Wirkung auf Vertrauens- und Vorhersagebereich beobachten und andere Niveaus einstellen – vgl. Anhang G.

**Beispiel**

Im Beispiel von Tabelle 10.1 erhält man:

95%-Vertrauensbereich für den **Mittelwert** der Schichtdicke nach $x = 10 \text{ min}$:

$$\hat{y} = b_0 + b_1 \cdot x = -2,72 + 0,837 \cdot 10 = 5,7$$

$$t \cdot s_R \cdot \sqrt{\frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} = 2,776 \cdot 2,07 \cdot \sqrt{\frac{1}{6} + \frac{(10 - 14)^2}{184}} = 2,9$$

$$5,7 - 2,9 \leq \mu_y(10) \leq 5,7 + 2,9 \quad \Rightarrow \quad 2,8 \leq \mu_y(10) \leq 8,6.$$

95%-Vorhersagebereich für einen **Einzelwert** der Schichtdicke nach $x = 10 \text{ min}$:

$$t \cdot s_R \cdot \sqrt{1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} = 2,776 \cdot 2,07 \cdot \sqrt{1 + \frac{1}{6} + \frac{(10 - 14)^2}{184}} = 6,4$$

$$5,7 - 6,4 \leq y(10) \leq 5,7 + 6,4 \quad \Rightarrow \quad -0,7 \leq y(10) \leq 12,1.$$

Der Vorhersagebereich für einen Einzelwert ist deutlich breiter als der Vertrauensbereich. Da negative Werte für die Schichtdicke technisch nicht möglich sind, ist ein sinnvoller Bereich $\leq 12,1 \mu\text{m}$.

**Aufgabe**

Berechnen Sie für die Aufgabe in Absatz 10.1.2 (Lebensdauer gegen Leckstrom) die 95%--, 99%- und 99,9%-Vertrauensbereiche für die Steigung, und beurteilen Sie damit die Signifikanz des Zusammenhangs.

In welchem Bereich erwarten Sie bei einem Leckstrom von $50 \mu\text{A}$ den Mittelwert der Lebensdauer, in welchem Bereich erwarten Sie einen Einzelwert (95%-Vertrauensniveau)?

Lösung

$$s_R = \sqrt{\frac{(1-B) \cdot Q_{yy}}{N-2}} = \sqrt{\frac{(1-0,9905) \cdot 1096}{2}} = 2,29 \quad f = N-2 = 4-2 = 2$$

Mit Tabelle 6.4 erhält man für die Breite der Vertrauensbereiche für die Steigung:

$$95\%: \quad 4,303 \cdot 2,29 / \sqrt{650} = 0,39$$

$$99\%: \quad 9,925 \cdot 2,29 / \sqrt{650} = 0,89$$

$$99,9\%: \quad 31,60 \cdot 2,29 / \sqrt{650} = 2,84$$

Der Vergleich mit dem Schätzwert $b_1 = -1,29$ für die Steigung ergibt:

Signifikanter Zusammenhang zwischen Leckstrom und Lebensdauer (**).

95%-Vertrauensbereich für den Mittelwert der Lebensdauer bei $50 \mu\text{A}$:

$$t \cdot s_R \cdot \sqrt{\frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} = 4,303 \cdot 2,29 \cdot \sqrt{\frac{1}{4} + \frac{5^2}{650}} = 5,3$$

$$\hat{y} = 131 - 1,29 \cdot 50 = 66,5 \Rightarrow 61 = 66,5 - 5,3 \leq \mu_y(50) \leq 66,5 + 5,3 = 72$$

95 %-Vorhersagebereich für einen Einzelwert der Lebensdauer bei 50 µA:

$$t \cdot s_{\bar{x}} \cdot \sqrt{1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{Q_{xx}}} = 4,303 \cdot 2,29 \cdot \sqrt{1 + \frac{1}{4} + \frac{5^2}{650}} = 11,2$$

$$\hat{y} = 131 - 1,29 \cdot 50 = 66,5 \Rightarrow 55 = 66,5 - 11,2 \leq y(50) \leq 66,5 + 11,2 = 78$$

Der Vorhersagebereich für einen Einzelwert ist deutlich breiter als der Vertrauensbereich für den Mittelwert.

10.1.5 Zusammenhang lineare Regression – Mittelwertvergleich

In Kapitel 6 wurden beim Vergleich von zwei Mittelwerten der Effekt und der Vertrauensbereich für den Effekt berechnet. Häufig entstehen die beiden Varianten, deren Mittelwerte verglichen werden sollen, indem für einen kontinuierlichen Faktor zwei verschiedene Stufen eingestellt werden. In diesem Fall ist alternativ zum Mittelwertvergleich auch eine Auswertung mit linearer Regression sinnvoll.

Als x-Werte bei der Regression kann man die Stufenwerte des Faktors verwenden. Alternativ kann man die Stufenwerte durch die lineare Transformation

$$\text{normierter Wert} = \frac{\text{Stufenwert} - \text{Mittelwert}}{\text{halbe Differenz}} \quad (10.22)$$

auf die Werte -1 und +1 umrechnen. Diese normierten Werte entsprechen der in der Versuchsplanung üblichen Bezeichnung der Stufen (vgl. Kapitel 7 und 8).

Verwendet man diese normierten x-Werte in der linearen Regression, so erhält man

$$b_1 = \frac{\text{Effekt}}{2} \quad (\text{für Schätzwert und Vertrauensbereich}). \quad (10.23)$$

Diese Umrechnungsformel ergibt sich, weil der Effekt die Differenz der Schätzwerte der Regressionsgeraden für $x = 1$ und $x = -1$ ist.

$$\text{Effekt} = (b_0 + b_1 \cdot 1) - (b_0 + b_1 \cdot (-1)) = 2 \cdot b_1$$

Mit Hilfe der Regressionsgeraden kann man auch Schätzwerte für x-Werte zwischen den Stufenwerten berechnen. Extrapolation sollte man aber vermeiden.

Bei Faktoren mit zwei Stufen sind

- die Berechnung von Effekten und deren Vertrauensbereichen (Kapitel 6),
- die Varianzanalyse (Kapitel 12) und
- die Regressionsanalyse mit Vertrauensbereichen für die Steigung

äquivalent. Der Ablauf der Rechnung ist zwar unterschiedlich, Inhalt und Signifikanz sind aber identisch.

10.1.6 Quasilineare Regression

Durch eine Transformation der x- und/oder y-Werte können viele nichtlineare in lineare Zusammenhänge überführt werden. Die Parameter im linearisierten Zusammenhang können dann wie bei linearer Regression angepasst werden⁶.



Beispiel

- Erwartet man einen quadratischen Zusammenhang der Form

$$z = \beta_0 + \beta_2 \cdot t^2$$

zwischen der Einflussgröße t und der Zielgröße z , so wird daraus durch die Transformationen

$$y = z \quad x = t^2$$

der linearisierte Zusammenhang $y = \beta_0 + \beta_2 \cdot x$.

- Erwartet man einen Zusammenhang der Form

$$z = \alpha \cdot \exp\left(\frac{\beta}{t}\right) \quad (\text{nach Logarithmieren: } \ln(z) = \ln\alpha + \frac{\beta}{t})$$

zwischen der Einflussgröße t und der Zielgröße z , so wird daraus durch die Transformationen

$$y = \ln(z) \quad x = \frac{1}{t}$$

der linearisierte Zusammenhang $y = \ln\alpha + \beta \cdot x$.

■ 10.2 Mehrfache Regression

Beschreibt man die Abhängigkeit einer Zielgröße von mehr als einer Einflussgröße (Faktor), so spricht man von mehrfacher Regression. Die Analyse verläuft ähnlich zur einfachen Regression. Für realistische Daten ist die Rechnung von Hand jedoch sehr aufwändig. Mehrfache Regression wird daher normalerweise mit Rechnerunterstützung durchgeführt. Die folgende Betrachtung beschränkt sich auf die Prinzipien.

⁶ Eine Transformation der y-Werte verändert (wie in Abschnitt 6.4 beschrieben) auch die Verteilungsform und -breite. Wenn dies die Übereinstimmung mit den in Bild 10-7 dargestellten Voraussetzungen verbessert, ist es erwünscht. Die Box-Cox-Darstellung in Bild 6-22 berücksichtigt sowohl die Verbesserung der Modellanpassung als auch die Veränderung der Verteilungsform.

Ist die Veränderung der Verteilungsform nicht erwünscht, so muss auf die Modellierungsansätze in Kapitel 17 zurückgegriffen werden.

10.2.1 Zweifache lineare Regression

Ausgangspunkt der zweifachen linearen Regression sind N Messwerte (nummeriert $i = 1, 2, \dots, N$) für eine Zielgröße y in Abhängigkeit von zwei Einflussgrößen x_1 und x_2 . Man vermutet einen linearen Zusammenhang, von dem die Messwerte nur zufällig um ε_i abweichen.

$$y_i = \beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \varepsilon_i \quad (10.24)$$

Aus den Daten sollen Schätzwerte b_0 , b_1 und b_2 für die Parameter β_0 , β_1 und β_2 ermittelt werden.

Grafisch bedeutet (10.24), dass die Zielgröße y (abgesehen von zufälligen Abweichungen) in einer Ebene liegt, deren Gleichung man bestimmen möchte.

Dazu berechnet man analog zu Q_{xx} , Q_{xy} und Q_{yy} bei der einfachen Regression die Hilfsgrößen

$$Q_{x_1 x_1}, Q_{x_1 x_2}, Q_{x_2 x_2}, Q_{x_1 y}, Q_{x_2 y} \text{ und } Q_{y y} \text{ (alle möglichen Paare).}$$

Wie bei der einfachen Regression sind die xy -Terme ein Maß dafür, wie stark sich x und y gemeinsam verändern, d.h. wie stark y linear von x_1 bzw. x_2 abhängt.

Neu ist der Term $Q_{x_1 x_2}$. Er ist ein Maß dafür, wie stark sich die beiden Einflussgrößen x_1 und x_2 gemeinsam verändern. Diese Größe hängt nur vom Versuchsplan (den x -Werten) ab, nicht von den Ergebnissen (y). Durch eine geeignete Wahl des Versuchsplans kann man erreichen, dass $Q_{x_1 x_2} = 0$ ist.

Dann sind die beiden Einflussgrößen x_1 und x_2 unabhängig voneinander, der Korrelationskoeffizient zwischen x_1 und x_2 ist 0 – man sagt, der Versuchsplan ist orthogonal bzgl. x_1 und x_2 . Bei der Auswertung können die beiden Einflussgrößen dann getrennt voneinander wie bei der einfachen Regression behandelt werden.

Ist der Versuchsplan nicht orthogonal, so ist die Auswertung mit Hilfe von Matrizen immer noch möglich (vgl. z.B. [1] und die Ergänzung am Ende dieses Absatzes; Statistiksoftware erledigt dies automatisch). Bei gleicher Zufallsstreuung der y -Werte sind die Vertrauensbereiche jedoch breiter als bei orthogonalen Plänen.

Bild 10-11 zeigt zwei extreme Versuchspläne (d.h. Anordnungen der x -Werte), um die Bedeutung der Orthogonalität zu verdeutlichen. Der linke Versuchsplan ist orthogonal – er ist offensichtlich gut geeignet, die Gleichung der Ebene zu bestimmen. Beim rechten Versuchsplan dagegen liegen die Versuchspunkte auf einer Geraden (d.h. der Korrelationskoeffizient zwischen x_1 und x_2 ist 1). Viele Ebenen durch die Versuchsergebnisse sind möglich, das Ergebnis ist nicht eindeutig – der Versuchsplan ist offensichtlich ungeeignet. Je näher der Korrelationskoeffizient zwischen x_1 und x_2 bei ± 1 liegt, desto weniger genau ist die Ebene bestimmt.

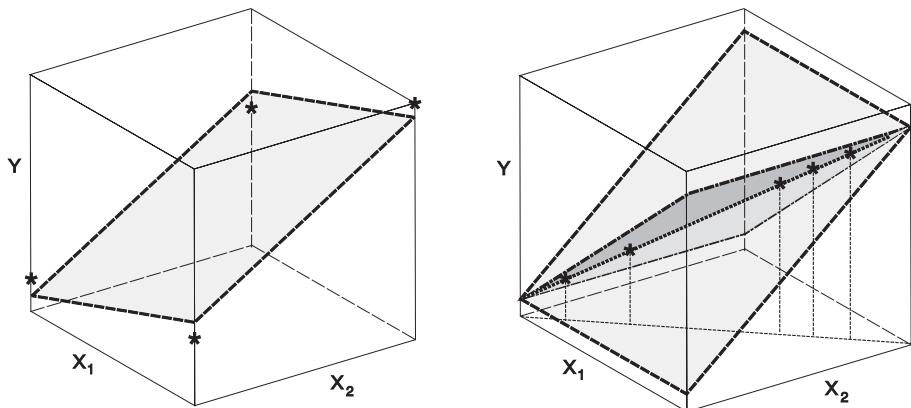


BILD 10-11 Wenn der Korrelationskoeffizient zwischen x_1 und x_2 gleich 0 ist (= orthogonal), sind die x -Werte optimal angeordnet, um später eine Ebene an die Versuchsergebnisse anzupassen (links). Wenn der Korrelationskoeffizient zwischen x_1 und x_2 gleich 1 ist, liegen die x -Werte näherungsweise auf einer Geraden: Es wird später nicht möglich sein, eine eindeutige Ebene durch die Versuchsergebnisse zu legen, sie kann gedreht werden (rechts).



Hinweis

Ergänzungen für mathematisch Interessierte⁷

1. Orthogonalität und Zusammenhang Regression – Effektberechnung

Die Orthogonalität eines faktoriellen Versuchsplans und der Zusammenhang zwischen Regression und Effektberechnung soll an einem 2^2 -Plan ohne Versuchswiederholungen gezeigt werden. Tabelle 10.2 zeigt die Versuchsergebnisse und die Berechnung der Effekte gemäß Kapitel 7.

Dieselben Ergebnisse können aber auch mit zweifacher linearer Regression analysiert werden. Tabelle 10.3 zeigt diese Rechnung für die Effekte der Faktoren A und B.

TABELLE 10.2 Berechnung der Effekte in einem 2^2 -Beispiel ohne Wiederholung

| Faktor A | Faktor B | WW AB | Zielgröße y |
|-------------------|----------|----------|-------------|
| – | – | + | 1 |
| + | – | – | 2 |
| – | + | – | 4 |
| + | + | + | 7 |
| $\Sigma = 4$ | 8 | 2 | |
| Effekt = 2 | 4 | 1 | |

⁷ Für das Verständnis der folgenden Kapitel nicht erforderlich.

TABELLE 10.3 Berechnung der Hilfsgrößen Q für die zweifache lineare Regression:
 Spalten A, B und Zielgröße zeigen Versuchsplan und Ergebnisse aus Tabelle 10.2. Die Faktorstufen von A und B sind die normierten Werte gemäß (10.22). $\bar{x}_1 = 0$ und $\bar{x}_2 = 0$, daher gilt z.B. $(x_{ij} - \bar{x}_1) = x_{ij}$ und die zugehörigen Spalten sind nicht getrennt dargestellt.

| A x_{1i} | B x_{2i} | Zielgr. y_i | $y_i - \bar{y}$ | $(x_{1i} - \bar{x}_1)^2$ | $(x_{1i} - \bar{x}_1) \cdot (x_{2i} - \bar{x}_2)$ | $(x_{2i} - \bar{x}_2)^2$ | $(x_{1i} - \bar{x}_1) \cdot (y_i - \bar{y})$ | $(x_{2i} - \bar{x}_2) \cdot (y_i - \bar{y})$ | $(y_i - \bar{y})^2$ |
|-----------------|-----------------|------------------|-----------------|--------------------------|---|--------------------------|--|--|---------------------|
| -1 | -1 | 1 | -2,5 | 1 | 1 | 1 | 2,5 | 2,5 | 6,25 |
| 1 | -1 | 2 | -1,5 | 1 | -1 | 1 | -1,5 | 1,5 | 2,25 |
| -1 | 1 | 4 | 0,5 | 1 | -1 | 1 | -0,5 | 0,5 | 0,25 |
| 1 | 1 | 7 | 3,5 | 1 | 1 | 1 | 3,5 | 3,5 | 12,25 |
| $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 14$ | $\Sigma = 0$ | $\Sigma = 4 =$ | $\Sigma = 0$ | $\Sigma = 4 =$ | $\Sigma = 4 =$ | $\Sigma = 8 =$ | $\Sigma = 21 =$ |
| $\bar{x}_1 = 0$ | $\bar{x}_2 = 0$ | $\bar{y} = 3,5$ | | $Q_{x_1 x_1}$ | $Q_{x_1 x_2}$ | $Q_{x_2 x_2}$ | $Q_{x_1 y}$ | $Q_{x_2 y}$ | Q_{yy} |

Das Ergebnis $Q_{x_1 x_2} = 0$ bedeutet, dass der Versuchsplan in x_1 und x_2 orthogonal ist.

Analog zur einfachen linearen Regression erhält man folgende Schätzwerte:

$$b_1 = \frac{Q_{x_1 y}}{Q_{x_1 x_1}} = \frac{4}{4} = 1 = \frac{\text{Effekt A}}{2}$$

$$b_2 = \frac{Q_{x_2 y}}{Q_{x_2 x_2}} = \frac{8}{4} = 2 = \frac{\text{Effekt B}}{2}$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}_1 - b_2 \cdot \bar{x}_2 = 3,5 - 1 \cdot 0 - 2 \cdot 0 = 3,5.$$

Wie bei der einfachen Regression ist der Regressionskoeffizient $b = \text{Effekt}/2$, da der Effekt der Unterschied zwischen $x = +1$ und $x = -1$ ist und damit $2b$ beträgt.

Bild 10-12 zeigt die Versuchspunkte und die mit Regression angepasste Ebene. Die Versuchspunkte weichen von ihr ab (vier Punkte liegen i.a. nicht in einer Ebene).

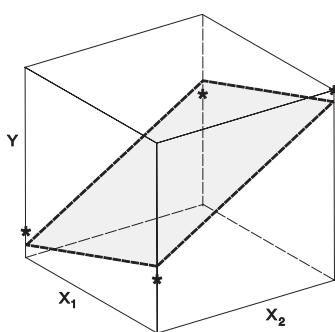


BILD 10-12

Grafische Darstellung der Versuchsergebnisse von Tabelle 10.3 mit der angepassten Ebene (zweifache lineare Regression)

Wie bei der einfachen linearen Regression kann man auch bei der mehrfachen Regression die Summe der quadrierten Abweichungen $Q_{\text{Gesamt}} = Q_{yy}$ in einen Anteil $Q_{\text{Regression}}$ und einen Q_{Rest} zerlegen und das Bestimmtheitsmaß B berechnen. Bei **orthogonalen** Versuchsplänen sind die Beiträge der Einflussgrößen zum Bestimmtheitsmaß additiv, und man erhält analog zu (10.13)

$$B = \frac{Q_{\text{Regression}}}{Q_{\text{Gesamt}}} = \frac{Q_{x_1 y}^2}{Q_{x_1 x_1} \cdot Q_{yy}} + \frac{Q_{x_2 y}^2}{Q_{x_2 x_2} \cdot Q_{yy}} = B_{x_1} + B_{x_2}.$$

Im obigen Beispiel erhält man

$$B = \frac{4^2}{4 \cdot 21} + \frac{8^2}{4 \cdot 21} = \frac{4}{21} + \frac{16}{21} = \frac{20}{21}.$$

Analog zur einfachen Regression kann man Vertrauensbereiche für die Regressionskoeffizienten berechnen und die Signifikanz beurteilen.

Der Freiheitsgrad beträgt $N-3$, weil aus den N Daten jetzt drei Parameter b_0 , b_1 und b_2 berechnet werden.

2. Matrixformalismus für nicht orthogonale Pläne ($Q_{x_1 x_2} \neq 0$)

Wenn der Versuchsplan nicht orthogonal ist, muss die Korrelation zwischen den Einflussgrößen berücksichtigt werden. Die Rechnung für die Einflussgröße x_1 kann nicht mehr unabhängig von x_2 durchgeführt werden. Mit Hilfe von Matrizen erhält man folgende Lösung:

$$Q = \begin{pmatrix} Q_{x_1 x_1} & Q_{x_1 x_2} \\ Q_{x_2 x_1} & Q_{x_2 x_2} \end{pmatrix} \quad q = \begin{pmatrix} Q_{x_1 y} \\ Q_{x_2 y} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = Q^{-1} \cdot q$$

wobei Q^{-1} die zu Q inverse Matrix ist.

Bei orthogonalen Versuchsplänen sind die Matrizen Q und Q^{-1} diagonal und man erhält die o. a. Gleichungen.

Liegen die Versuchspunkte auf einer Geraden, so kann Q nicht invertiert werden, die Regressionskoeffizienten sind nicht bestimmt. Anschaulich bedeutet dies, dass die Ebene wie in Bild 10-11 rechts nicht eindeutig bestimmt ist, sondern um die punktierte Gerade gedreht werden kann. Allgemein kann man zeigen, dass die Regressionsebene am besten bestimmt ist, wenn der Versuchsplan orthogonal ist. Jede Korrelation zwischen x_1 und x_2 verschlechtert die Regression.

Für das Bestimmtheitsmaß erhält man

$$B = \frac{q^T \cdot Q^{-1} \cdot q}{Q_{yy}}.$$

Die obigen Gleichungen können auf beliebig viele Variablen erweitert werden, man erhält nur größere Matrizen.

10.2.2 Transformierte Einflussgrößen

Die Einflussgrößen in der linearen Regression können auch transformierte (d. h. aus den Einflussgrößen berechnete) Größen sein. So fehlt z.B. in Tabelle 10.3 im Vergleich zu Tabelle 10.2 der Effekt der 2FWW AB. Er kann wie die (Haupt-)Effekte der Faktoren bestimmt werden, wenn man eine transformierte Variable

$$x_3 = x_1 \cdot x_2$$

einführt und ein Regressionsmodell der Form

$$y_i = \beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \beta_3 \cdot x_{3i} + \varepsilon_i = \beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \beta_{12} \cdot x_{1i} \cdot x_{2i} + \varepsilon_i$$

anpasst. Die Rechnung verläuft formal wie bei einer dreifachen linearen Regression.

Auch hier ist die Analyse besonders einfach, wenn x_3 orthogonal zu x_1 und x_2 ist. Bei den vollständigen und fraktionellen faktoriellen Versuchsplänen gilt: Unvermengte Effekte sind orthogonal bei der Regression, wenn als Faktorstufen die Werte -1 und +1 verwendet werden.⁸



Hinweis

Ergänzung für mathematisch Interessierte

Zusammenhang Regression – Effektberechnung (Fortsetzung)

Tabelle 10.4 zeigt die Erweiterung der Analyse von Tabelle 10.3 zur Berechnung von $b_3 = b_{12}$.

TABELLE 10.4 Ergänzungen von Tabelle 10.3 zur Berechnung von $b_3 = b_{12}$

| A x_{1i} | B x_{2i} | AB $x_{3i} =$ $x_{1i} \cdot x_{2i}$ | Zielgr. y_i | $y_i - \bar{y}$ | $(x_{1i} - \bar{x}_1) \cdot$ $(x_{3i} - \bar{x}_3)$ | $(x_{2i} - \bar{x}_2) \cdot$ $(x_{3i} - \bar{x}_3)$ | $(x_{3i} - \bar{x}_3)^2$ | $(x_{3i} - \bar{x}_3) \cdot$ $(y_i - \bar{y})$ |
|-----------------|-----------------|---|------------------|-----------------|--|--|--------------------------|---|
| -1 | -1 | 1 | 1 | -2,5 | -1 | -1 | 1 | -2,5 |
| 1 | -1 | -1 | 2 | -1,5 | -1 | 1 | 1 | 1,5 |
| -1 | 1 | -1 | 4 | 0,5 | 1 | -1 | 1 | -0,5 |
| 1 | 1 | 1 | 7 | 3,5 | 1 | 1 | 1 | 3,5 |
| $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 14$ | $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 4 =$ | $\Sigma = 2 =$ |
| $\bar{x}_1 = 0$ | $\bar{x}_2 = 0$ | $\bar{x}_3 = 0$ | $\bar{y} = 3,5$ | | $Q_{x_1 x_3}$ | $Q_{x_2 x_3}$ | $Q_{x_3 x_3}$ | $Q_{x_3 y}$ |

Die Ergebnisse

$$Q_{x_1 x_2} = 0, Q_{x_1 x_3} = 0 \text{ und } Q_{x_2 x_3} = 0$$

bedeuten, dass der Versuchsplan in x_1, x_2 und $x_3 = x_1 \cdot x_2$ orthogonal ist.

Analog zur linearen Regression erhält man folgenden Schätzwert:

$$b_3 = b_{12} = \frac{Q_{x_3 y}}{Q_{x_3 x_3}} = \frac{2}{4} = 0,5 = \frac{\text{Effekt AB (2FWW)}}{2}$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}_1 - b_2 \cdot \bar{x}_2 - b_3 \cdot \bar{x}_3 = 3,5 - 1 \cdot 0 - 2 \cdot 0 - 0,5 \cdot 0 = 3,5$$

Wie bei der linearen Regression gilt $b_3 = b_{12} = \text{Effekt}/2$.

Bild 10-13 zeigt das angepasste Modell als Funktion von x_1 und x_2 . Obwohl es linear in x_1, x_2 und x_3 ist, stellt es eine gekrümmte Fläche in x_1 und x_2 dar, da $x_3 = x_1 \cdot x_2$. Im Höhenliniendiagramm ist die Krümmung deutlicher zu erkennen als in der Wirkungsfläche. Die Wechselwirkung führt dazu, dass die Höhenlinien bei $x_1 = 1$ dichter beieinander liegen als bei $x_1 = -1$.

⁸ Dies gilt nicht für die natürlichen Stufenwerte, daher werden bei der Analyse von geplanten Versuchen immer die Werte -1 und +1 verwendet.

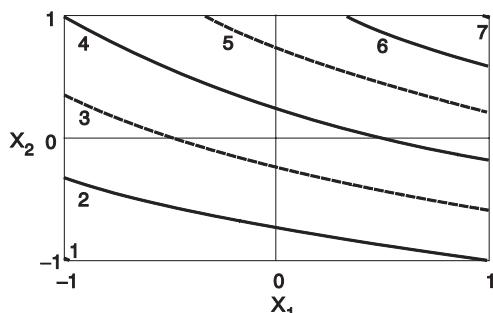
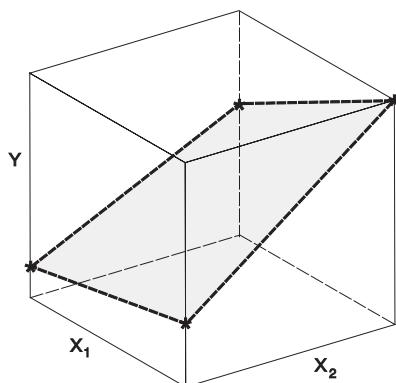


BILD 10-13 Grafische Darstellung der Versuchsergebnisse mit dem angepassten Regressionsmodell (lineare Regression mit dem zusätzlichen Term $x_3 = x_1 \cdot x_2$)

- dreidimensional als Wirkungsfläche (= gekrümmte Fläche im Raum)
- zweidimensional als Höhenliniendiagramm

Bei den Daten in Bild 10-13 ist die Wechselwirkung relativ klein, daher ist auch die Abweichung der angepassten Fläche von einer Ebene nur klein. Bei großen Wechselwirkungen sind auch entsprechend große Abweichungen von einer Ebene möglich. Bild 10-14 zeigt zwei Beispiele.

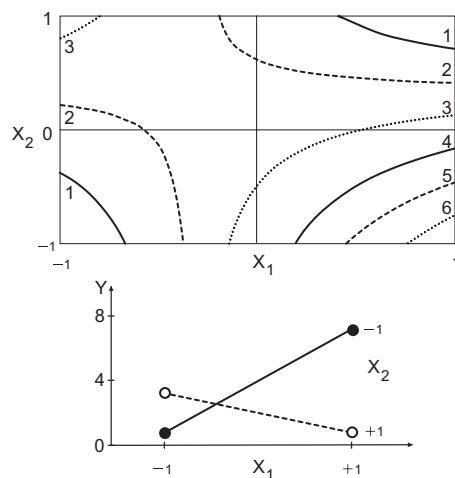
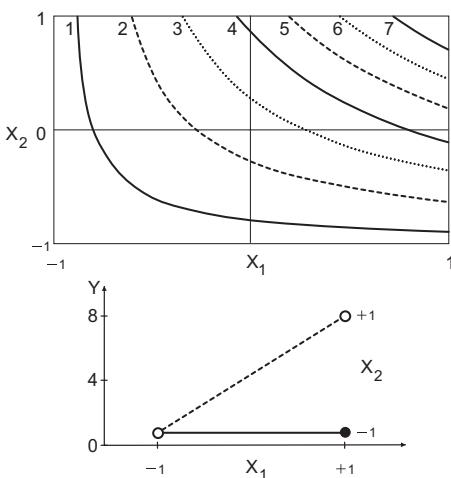


BILD 10-14

Oben: Höhenliniendiagramme mit größerer Wechselwirkung und damit größerer Abweichung von einer Ebene als in Bild 10-13.

Unten: Wechselwirkungsdarstellung derselben Ergebnisse (wie in Kapitel 7).

Die y-Werte der vier Faktorstufenkombinationen sind die Werte in den Ecken des Höhenliniendiagramms; mit dem Höhenliniendiagramm interpoliert man zwischen den Ecken (nur sinnvoll, wenn diese Zwischenwerte für die Faktoren auch möglich sind).

Bei Versuchsplänen, deren Faktoren auf mehr als zwei Stufen untersucht werden, können mit demselben Algorithmus zusätzlich quadratische und höhere Terme im Regressionsmodell angepasst werden. Auch hier ist es wichtig, dass die Versuchspläne möglichst bezüglich aller Terme im Modell orthogonal sind. Geeignete Versuchspläne werden in Kapitel 11 behandelt.

10.2.3 Prinzip der schrittweisen Regression

Soll die Abhängigkeit einer Zielgröße von vielen Einflussgrößen gleichzeitig untersucht werden, so geht man häufig schrittweise vor, um so ein möglichst einfaches Modell zu erhalten. Man nimmt nur diejenigen Einflussgrößen in das Modell auf, von denen die Zielgröße signifikant abhängt.

1. Mit einfacher linearer Regression wird für alle Einflussgrößen einzeln untersucht, welche Einflussgröße das größte Bestimmtheitsmaß B ergibt. Ist die Abhängigkeit signifikant, so wird die Einflussgröße ins Regressionsmodell aufgenommen.
2. Für alle Einflussgrößen, die noch nicht ins Modell aufgenommen wurden, wird untersucht, welche die größte Erhöhung des Bestimmtheitsmaßes B ergibt. Ist die Abhängigkeit signifikant, so wird auch diese Einflussgröße ins Regressionsmodell aufgenommen.
3. Schritt 2 wird so lange wiederholt, bis keine weitere Einflussgröße eine signifikante Abhängigkeit mehr ergibt. Damit ist das Modell erreicht, das die Daten mit den wenigen Parametern beschreibt.

Für nicht orthogonale Versuchspläne kann die Aufnahme eines zusätzlichen Terms (entspricht einem Effekt) ins Regressionsmodell die Signifikanz von Termen, die früher ins Modell aufgenommen wurden, erniedrigen. Daher wird nach Aufnahme eines neuen Parameters überprüft, ob noch alle Terme im Modell signifikant sind. Falls Terme nicht mehr signifikant sind, werden diese schrittweise entfernt, beginnend mit dem Term, der am wenigsten zu B beiträgt.

Alternativ zur hier beschriebenen schrittweisen Aufnahme (forward selection) von Termen ins Modell kann man auch mit einem Modell beginnen, das alle Terme enthält. Aus diesem werden dann schrittweise diejenigen entfernt, die am wenigsten zu B beitragen, bis alle verbliebenen signifikant sind (backward selection).

Bei orthogonalen Versuchsplänen führen beide Vorgehensweisen zum selben Ergebnis. Bei nicht orthogonalen Versuchsplänen oder ungeplanten Versuchen ist es möglich, dass das Endmodell bei forward und backward selection unterschiedliche Terme enthält. Ein Vergleich der beiden Endmodelle ist ein nützlicher empirischer Test für die Vertrauenswürdigkeit der Ergebnisse.

Mit schrittweiser Regression können auch transformierte Einflussgrößen behandelt werden. Dann ist es jedoch sinnvoll, neben der Signifikanz der Terme auch auf ihre Hierarchie zu achten. Wenn z.B. der Term $x_1 \cdot x_2$ aufgenommen werden soll, müssen auch x_1 und x_2 einzeln ins Modell aufgenommen werden.

10.2.4 Beurteilung des Regressionsmodells

Mit Hilfe der mehrfachen linearen Regression wird ein vorgegebenes Modell an die Daten angepasst. Eine solche Anpassung ist formal immer möglich. Das bedeutet jedoch nicht, dass das resultierende Modell die Wirklichkeit beschreibt.

- Bei der Auswertung von ungeplanten Versuchen (wie z. B. Daten aus der laufenden Fertigung) bedeutet ein statistisch signifikanter Zusammenhang nicht notwendigerweise, dass ein Ursache-Wirkungs-Zusammenhang besteht. Ein signifikanter Zusammenhang kann z. B. auch die Folge einer gemeinsamen Abhängigkeit von einer nicht gemessenen anderen Größe sein. Werden z. B. mehrere Prozessparameter in kurzem zeitlichen Abstand immer wieder gemessen, so können durch die langsame Veränderung der Umgebungsbedingungen signifikante Zusammenhänge zwischen diesen Parametern vorgetäuscht werden.
- Ein statistisch signifikanter Zusammenhang bedeutet nicht, dass die mathematische Form des Modells richtig ist. Daraus ergeben sich u. a. folgende Konsequenzen:
 - Die mathematische Form des Modells sollte bereits vor der Anpassung durch physikalisch-technische Überlegungen festgelegt werden.
 - Nach der Anpassung müssen das angepasste Modell und die Residuen grafisch beurteilt werden (ggf. ergänzt durch statistische Tests).
 - Das angepasste Modell darf nicht über den untersuchten Bereich der Einflussgrößen hinaus extrapoliert werden.
- Korrelation zwischen den Einflussgrößen führt immer zu einer Verbreiterung der Vertrauensbereiche der Koeffizienten β_i . Bei geplanten Versuchen kann diese Korrelation klein gehalten werden. Bei orthogonalen Versuchsplänen ist sie 0. Ein direktes Maß für die Verbreiterung der Vertrauensbereiche der Koeffizienten β aufgrund von Korrelationen ist der „Varianzinflationsfaktor“ VIF. Der Vertrauensbereich des Koeffizienten β_i ist um den Faktor $\sqrt{VIF_i}$ größer als er bei einem orthogonalen Plan wäre (ideal ist der Wert VIF = 1) [3].
- Es ist problematisch, wenn ein einzelnes Versuchsergebnis einen besonders großen Einfluss auf das Regressionsmodell hat. Wie stark dieser Einfluss ist, hängt zum Teil von den x-Werten und zum Teil vom Versuchsergebnis (dem y-Wert) ab.
 - Je weiter der x-Wert eines Einzelversuchs von dem der anderen Einzelversuche entfernt liegt, desto größer ist dessen Einfluss auf den Schätzwert eines der Koeffizienten im Regressionsmodell. Ein Maß für diesen Einfluss eines Einzelversuchs ist der „Hebel“ (leverage) [3]. Werden mit N Einzelversuchen Schätzwerte für p Koeffizienten bestimmt, so beträgt der Hebel im Mittel p/N . Ideal ist, wenn der Hebel für jeden Einzelversuch möglichst nahe bei diesem Wert liegt. Da der Hebel nur von den x-Werten abhängt, kann dies durch geeignete Versuchspläne erreicht werden.
 - Aber natürlich hat auch ein besonders ungewöhnliches Versuchsergebnis y einen großen Einfluss auf das Regressionsmodell. Die „Cook-Distanz“ [3, 4] ist ein Maß dafür, um wie viel sich das Regressionsmodell ändert, wenn ein einzelnes Versuchsergebnis weggelassen wird. Bei Versuchsergebnissen mit großer Cook-Distanz ist meist das Residuum oder der Hebel groß (oder beides). Einzelversuche mit besonders großer Cook-Distanz sollten überprüft werden. „DFFITS“ ist ein ähnliches Maß [4].

- Das Regressionsmodell soll mit den Versuchsergebnissen konsistent sein, d.h. es soll keine signifikanten Abweichungen geben. Dazu können die Residuen gegen die x-Werte aufgetragen werden – wie in Bild 10-4. Alternativ oder ergänzend kann die Reststreuung Q_{Rest} in (10.10) weiter zerlegt werden in eine reine Zufallsstreuung $Q_{\text{Pure Error}}$ (aus mehrfachen Realisierungen derselben x-Werte) und einen $Q_{\text{Lack of Fit}}$ (aus den Abweichungen der Mittelwerte von der Regressionsgeraden). Mit Ideen der Varianzanalyse (vgl. Abschnitt 12) kann dann überprüft werden, ob der Lack-of-Fit signifikant größer ist als der Pure-Error. Bei linearer Regression erhält man so den Linearitätstest nach Fisher [1], bei quasilinearer Regression allgemeiner den Anpassungstest oder Lack-of-Fit-Test [3, 4].
- Das Regressionsmodell soll Versuchsergebnisse möglichst gut vorhersagen können. Die Kennzahl PRESS (Prediction Residual Sum of Squares) misst die Prognosefähigkeit des Modells. PRESS wird berechnet wie Q_{Rest} in (10.10), nur dass der gerade betrachtete Punkt bei der Berechnung des Modells jeweils nicht berücksichtigt wird – eine sogenannte Kreuz-Validierung. PRESS ist damit immer größer als Q_{Rest} , der Unterschied sollte jedoch möglichst nicht mehr als einige Prozent betragen [3, 4]⁹. Das aus PRESS statt Q_{Rest} berechnete Bestimmtheitsmaß wird in Software meist mit r^2_{pred} bezeichnet und ist immer kleiner als r^2 . Ein zu großer Unterschied weist auf Überanpassung hin, vgl. Absatz 17.3.1.

Viele potentielle Probleme können durch geeignete Planung der Versuche vermieden werden. Kapitel 11 behandelt Versuchspläne für quantitative Faktoren mit mehr als zwei Stufen, die für die Auswertung mit Regression geeignet sind.

Literatur

- [1] *Graf, U./Henning, H.-J./Stange, K./Wilrich, P.-T.*: Formeln und Tabellen der angewandten mathematischen Statistik. Springer Verlag, Berlin 4. Aufl. 1998
- [2] *Box, G. E. P./Hunter, W. G./Hunter, J. S.*: Statistics for Experimenters. John Wiley, New York, 2. Auflage 2005
- [3] *Fahrmeir, L./Kneib, T./Lang, S.*: Regression: Modelle, Methoden und Anwendungen. Springer Verlag, Berlin, 2. Auflage 2009
- [4] *Montgomery, D. C./Peck, E. A./Vining, G. G.*: Introduction to Linear Regression Analysis. John Wiley, New York, 5. Auflage 2012

⁹ Die Visualisierung „Lineare Regression“ zeigt Hebel, Cook-Distanz, Q_{Rest} , $Q_{\text{Lack of Fit}}$, $Q_{\text{Pure Error}}$ und PRESS – vgl. Anhang G

11

Versuchspläne für nichtlineare Zusammenhänge

Häufig soll die quantitative Abhängigkeit einer (oder auch mehrerer) Zielgröße(n) von einigen wenigen Faktoren im Detail bestimmt werden. Dann begnügt man sich nicht mit der bisher behandelten linearen Näherung. Die Nichtlinearität ist entscheidend, wenn die Lage eines Maximums (z.B. der Ausbeute) oder eines Minimums (z.B. der Anzahl der Fehler) gesucht ist.

Meist verwendet man dann ein quadratisches Modell zur empirischen Beschreibung der Abhängigkeit der Zielgröße y von den Faktoren. Bei zwei Faktoren x_1 und x_2 erhält man z.B.

$$y_i = \beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \beta_{11} \cdot x_{1i}^2 + \beta_{12} \cdot x_{1i} \cdot x_{2i} + \beta_{22} \cdot x_{2i}^2 + \varepsilon_i \quad (11.1)$$

Dieses Modell kann man auf k Faktoren verallgemeinern. Es enthält dann

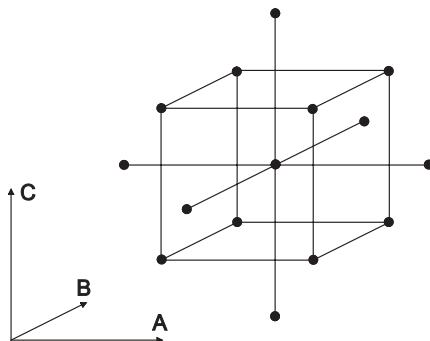
| | |
|------------|---|
| 1 | Koeffizienten β_0 |
| k | Koeffizienten $\beta_1, \beta_2, \dots, \beta_k$ |
| $k(k+1)/2$ | Koeffizienten $\beta_{11}, \beta_{12}, \dots, \beta_{1k}, \beta_{22}, \dots, \beta_{2k}, \dots, \beta_{kk}$ |

 (11.2)

Wenn in (11.1) normierte Stufenwerte verwendet werden, sind (abgesehen vom Umrechnungsfaktor $1/2$) die Koeffizienten β der linearen Terme x_1 und x_2 die Effekte der Faktoren A und B, der Koeffizient von $x_1 \cdot x_2$ ist der Effekt der Wechselwirkung AB, die Koeffizienten von x_1^2 und x_2^2 sind quadratische Effekte der Faktoren (AA und BB). Diese quadratischen Effekte sind hier neu. Um sie zu bestimmen, benötigt man Versuchspläne mit mehr als zwei Faktorstufen. Im Folgenden werden geeignete Versuchspläne behandelt (vgl. z.B. [1-4]).

■ 11.1 Zentral zusammengesetzte Versuchspläne

Ein zentral zusammengesetzter Versuchsplan besteht aus einem vollständigen oder fraktionellen faktoriellen Versuchsplan 2^{k-p} mit Mindestauflösung V („Würfel“), dem ein „Stern“ und ein „Zentrum“ hinzugefügt werden. Bild 11-1 zeigt als Beispiel die Anordnung der Einzelversuche für $k = 3$ Faktoren grafisch.

**BILD 11-1**

Zentral zusammengesetzter Versuchsplan für 3 Faktoren A, B und C

Tabelle 11.1 zeigt eine Liste der Faktorstufenkombinationen für $k=3$. Die Faktorstufen sind dabei in Analogie zur Darstellung in Kapitel 7 und 8 mit den normierten Werten ± 1 , 0 und $\pm \alpha$ bezeichnet. Je nach Zielsetzung können α und n_0 in Tabelle 11.1 verschiedene Werte annehmen.

TABELLE 11.1 Faktorstufenkombinationen eines zentral zusammengesetzten Versuchsplans für $k = 3$ Faktoren, das Zentrum wird hier z. B. viermal realisiert (systematische Reihenfolge, normierte Stufenwerte)

| syst. Nr. | Faktor A | Faktor B | Faktor C | Erläuterung |
|-----------|------------|------------|------------|--------------------------------------|
| 1 | - 1 | - 1 | - 1 | „Würfel“ vollständig faktoriell |
| 2 | 1 | - 1 | - 1 | |
| 3 | - 1 | 1 | - 1 | |
| 4 | 1 | 1 | - 1 | |
| 5 | - 1 | - 1 | 1 | |
| 6 | 1 | - 1 | 1 | |
| 7 | - 1 | 1 | 1 | |
| 8 | 1 | 1 | 1 | |
| 9 | - α | 0 | 0 | „Stern“ jeder Faktor getrennt |
| 10 | α | 0 | 0 | |
| 11 | 0 | - α | 0 | |
| 12 | 0 | α | 0 | |
| 13 | 0 | 0 | - α | |
| 14 | 0 | 0 | α | Zentrum n_0 -mal (hier viermal) |
| 15 | 0 | 0 | 0 | |
| 16 | 0 | 0 | 0 | |
| 17 | 0 | 0 | 0 | |
| 18 | 0 | 0 | 0 | |

11.1.1 Orthogonaler Versuchsplan

Ein zentral zusammengesetzter Versuchsplan wie in Tabelle 11.1 ist **orthogonal** bezüglich aller Terme im Modell (11.1), wenn

$$\text{orthogonal: } \alpha^2 = \frac{1}{2}(\sqrt{N \cdot N_w} - N_w), \quad (11.3)$$

wobei $N_w = 2^{k-p}$ = Anzahl Einzelversuche im Würfel und

$$N = 2^{k-p} + 2k + n_0 = \text{Gesamtzahl Einzelversuche.}$$

Die Vorteile eines orthogonalen Versuchsplans sind:

- Die Schätzwerte für die Koeffizienten β im Modell (11.1) sind unabhängig voneinander (d.h. sie beeinflussen sich nicht gegenseitig), und
- bei vorgegebenen Stufenwerten für -1 und 1 und vorgegebener Anzahl der Einzelversuche erhält man möglichst schmale Vertrauensbereiche für die Koeffizienten β .

Daher wird empfohlen, soweit möglich orthogonale Pläne zu verwenden.



Beispiel

In Tabelle 11.1 beträgt

- die Anzahl der Realisierungen des Zentrums $n_0 = 4$,
- die Anzahl Faktoren $k = 3$, und
- als Würfel wird ein vollständiger faktorieller Plan verwendet ($p = 0$).

Damit erhält man

$$N_w = 2^3 = 8 \text{ und } N = 2^3 + 2 \cdot 3 + 4 = 18 \Rightarrow \alpha^2 = \frac{1}{2}(\sqrt{18 \cdot 8} - 8) = 2$$

\Rightarrow Wählt man für $\alpha = \sqrt{2} = 1,414$, so ist der Versuchsplan orthogonal.

Einen konkreten Versuchsplan erhält man aus Tabelle 11.1, indem man

- für jeden Faktor festlegt, welche natürlichen Stufenwerte -1 und 1 jeweils darstellen,
- die zusätzlichen Stufenwerte 0 und $\pm\alpha$ linear daraus errechnet und
- die Reihenfolge der Einzelversuche randomisiert.

Beispiel

Das Beispiel aus der chemischen Industrie in Kapitel 3 soll zu einem zentral zusammengesetzten Versuch erweitert werden.

Faktor A Temperatur $-1 = 120^\circ\text{C}$ $+1 = 140^\circ\text{C}$

Faktor B Zeit $-1 = 2 \text{ h}$ $+1 = 4 \text{ h}$

Faktor C Katalysator $-1 = 0,1\%$ $+1 = 0,5\%$

Mit diesen Vorgaben erhält man durch lineare Umrechnung die Stufenwerte in Tabelle 11.2.

TABELLE 11.2 Umrechnung der normierten Stufenwerte der Faktoren in die natürlichen Werte

| normierter Wert | - 1,414 | - 1 | 0 | 1 | 1,414 |
|-----------------|---------|-------|-------|-------|---------|
| A Temperatur | 115,9°C | 120°C | 130°C | 140°C | 144,1°C |
| B Zeit | 1,59 h | 2 h | 3 h | 4 h | 4,41 h |
| C Katalysator | 0,017% | 0,1% | 0,3% | 0,5% | 0,583% |

Die Auswertung erfolgt mit Regression, und zwar immer in den normierten Stufenwerten, da der Versuchsplan nur dann orthogonal ist und der einfache Zusammenhang Regressionskoeffizient = Effekt/2 gilt.

Gute Versuchsplanungssoftware führt diese Umrechnung im Hintergrund durch, ohne dass sich der Anwender darum kümmern muss. Aber Achtung, umfassende Statistikpakete unterscheiden zwischen Versuchsplanung und „normaler“ Regression. Mit normaler Regression ist die Auswertung zwar im Prinzip ebenfalls möglich, die Umrechnung muss dann jedoch explizit vom Anwender vorgenommen werden. Daher wird empfohlen, Versuchspläne nur mit den speziell dafür vorgesehenen Routinen auszuwerten.

11.1.2 Technisch bedingte Abweichungen vom Versuchsplan

Je nach weiteren Randbedingungen ist es manchmal erforderlich, vom idealen orthogonalen Versuchsplan abzuweichen. Dies gilt insbesondere,

- wenn es technische oder andere Grenzen für die Stufenwerte der Faktoren gibt oder
- wenn nur bestimmte Stufenwerte einstellbar sind.

Ein solcher veränderter Versuchsplan kann mit entsprechender Software ohne Mehraufwand ausgewertet werden. Da jede Abweichung von der Orthogonalität aber zu einer Verbreiterung der Vertrauensbereiche der Regressionskoeffizienten β führt, sollte man die Veränderung des Versuchsplans so gering wie technisch möglich halten.



Beispiel (Fortsetzung)

Die Maximaltemperatur der Anlage beträgt 140°C. Es ist daher nicht möglich, den idealen Wert von 144,1°C einzustellen.

Andererseits erwartet man aus technischen Gründen bei der Maximaltemperatur auch die maximale Ausbeute. Daher ist es nicht sinnvoll, die Stufen so umzudefinieren, dass $\alpha = 1,414$ dem Maximalwert von 140°C entspricht, da dann nur ein Einzelversuch bei dieser vermutlich optimalen Temperatur durchgeführt würde (Nr. 10 in Tabelle 11.1). Es ist besser, auf die Orthogonalität zu verzichten und für den Faktor A als Stufenwerte z.B. 115,9°C 120°C 130°C 140°C 140°C zu verwenden.

11.1.3 Bekannte nichtlineare Abhängigkeiten

In manchen Anwendungen ist bereits von Anfang an ein bestimmtes, nichtlineares Verhalten der Zielgröße bekannt (z. B. Sättigung). Will man dieses nichtlineare Verhalten gezielt untersuchen, so kann es im Einzelfall sinnvoll sein, die Nichtlinearität bereits bei der Definition der Faktorstufen zu berücksichtigen.

Dies erfolgt, indem man als Faktor im Versuchsplan eine geeignete Funktion der eigentlich eingestellten Größe verwendet (Transformation).



Beispiel (Fortsetzung)

Wenn bereits vor der Versuchsdurchführung bekannt ist, dass die Ausbeute der chemischen Reaktion bei langen Reaktionszeiten gegen einen Grenzwert strebt und bei einer weiteren Verlängerung immer langsamer zunimmt, so ist es sinnvoll, dies bereits bei der Planung des Versuchs zu berücksichtigen.

Ist man speziell an diesem Langzeitverhalten interessiert, so ist es besser, als Faktor statt der Reaktionszeit z. B. den Kehrwert ($1/\text{Reaktionszeit}$) oder den Logarithmus $\ln(\text{Reaktionszeit})$ zu verwenden. Aus der linearen Veränderung des Faktors ergibt sich dann eine nichtlineare Veränderung der eigentlich eingestellten Größe Reaktionszeit. Tabelle 11.3 zeigt diesen Zusammenhang an zwei Beispielen. Im Versuchsplan steht die Reaktionszeit, die Analyse erfolgt aber in der Größe ($1/\text{Reaktionszeit}$) oder $\ln(\text{Reaktionszeit})$.

TABELLE 11.3 Beispiel für die Auswirkung nichtlinearer Faktoren:

Aus der Vorgabe fester Zeiten 2 h und 4 h für die Würfeleckpunkte erhält man mit den Faktoren $1/\text{Zeit}$ bzw. $\ln(\text{Zeit})$ Stufen für die eingestellte Größe „Zeit“, die im erwarteten Sättigungsbereich weiter auseinander liegen als in Tabelle 11.2.

| normierter Wert | -1,414 | -1 | 0 | 1 | 1,414 |
|--|--------|-------|-------|-------|-------|
| Faktor $1/\text{Zeit} [\text{h}^{-1}]$ | 0,198 | 0,25 | 0,375 | 0,5 | 0,552 |
| eingestellte Größe Zeit [h] | 5,05 | 4 | 2,67 | 2 | 1,81 |
| Faktor $\ln(\text{Zeit})$ | 0,550 | 0,693 | 1,040 | 1,386 | 1,530 |
| eingestellte Größe Zeit [h] | 1,73 | 2 | 2,83 | 4 | 4,62 |

11.1.4 Varianten von zentral zusammengesetzten Plänen

Drehbare Pläne

Die Regressionsgleichung (11.1) gibt für jeden Punkt $(x_1, x_2, x_3, \dots, x_k)$ einen Schätzwert für die Zielgröße. Für diesen Schätzwert kann ein Vertrauensbereich berechnet werden. Im Fall der einfachen linearen Regression ist dies der Trichter in Bild 10-9. Beim Modell (11.1) hängt die Breite des Vertrauensbereichs im Allgemeinen von den Werten aller Faktoren ab. Man nennt einen Versuchsplan drehbar, wenn die Breite des Vertrauensbereichs nur vom Abstand des betrachteten Punktes von Zentrum des Würfels abhängt und nicht auch von der Richtung. D.h. ein Versuchsplan ist drehbar, wenn die Breite des Vertrauensbereichs nur von

$$\sqrt{x_1^2 + x_2^2 + \dots + x_k^2}$$

abhängt (normierte Werte von x). Man kann zeigen, dass ein zentral zusammengesetzter Versuchsplan wie in Tabelle 11.1 drehbar ist, wenn

$$\text{drehbar: } \alpha^2 = \sqrt{2^{k-p}}. \quad (11.4)$$



Beispiel 1

Versuchsplan in Tabelle 11.1: $k = 3$; $p = 0$; $n_0 = \text{beliebig}$:

$$\alpha^2 = \sqrt{2^{3-0}} = \sqrt{8} = 2,828 \Rightarrow \alpha = 1,682$$

Beispiel 2

$k = 5$; $p = 1$; $n_0 = \text{beliebig}$:

$$\alpha^2 = \sqrt{2^{5-1}} = \sqrt{16} = 4 \Rightarrow \alpha = 2.$$

Orthogonal und drehbar

Durch geeignete Wahl von n_0 kann man erreichen, dass das α für Orthogonalität aus (11.3) näherungsweise mit dem α für Drehbarkeit aus (11.4) übereinstimmt. Dann ist der Versuchsplan drehbar und (näherungsweise) orthogonal. Tabelle 11.4 zeigt für 2 bis 8 Faktoren den Mindestversuchsumfang mit $n_0=1$, den Wert α für drehbare Pläne und den Versuchsumfang für Pläne, die gleichzeitig drehbar und orthogonal sind.

TABELLE 11.4 Mindestversuchsumfang N_{\min} mit $n_0=1$, α für drehbare Pläne und Versuchsumfang N für drehbare und (näherungsweise) orthogonale Pläne in Abhängigkeit von der Anzahl der Faktoren k

| Anzahl Faktoren k | Würfel | Mindestwert N_{\min} | α (drehbar) | n_0 (für drehbar und orthogonal) | N (für drehbar und orthogonal) |
|-------------------|-----------|------------------------|--------------------|------------------------------------|--------------------------------|
| 2 | 2^2 | 9 | 1,414 | 8 | 16 |
| 3 | 2^3 | 15 | 1,682 | 9 | 23 |
| 4 | 2^4 | 25 | 2,000 | 12 | 36 |
| 5 | 2^{5-1} | 27 | 2,000 | 10 | 36 |
| 6 | 2^{6-1} | 45 | 2,378 | 15 | 59 |
| 7 | 2^{7-1} | 79 | 2,828 | 22 | 100 |
| 8 | 2^{8-2} | 81 | 2,828 | 20 | 100 |

Flächenzentriert

Für manche Faktoren sind aus technischen Gründen nur drei Faktorstufen möglich. Manchmal ist es auch nicht sinnvoll, mit den „Sternpunkten“ über den „Würfel“ hinauszugehen (vgl. Absatz 11.1.2).

In diesen Fällen verwendet man $\alpha = 1$. Dann liegen die Sternpunkte in Bild 11-1 in der Mitte der Würfelflächen. Man nennt diesen Versuchsplan daher flächenzentriert. Er ist nicht orthogonal und sollte daher nur in Ausnahmefällen eingesetzt werden.

Blockbildung und orthogonale Blöcke

Ist es nicht möglich, die Versuchsbedingungen über alle Faktorstufenkombinationen eines zentral zusammengesetzten Versuchsplans homogen zu halten, so können Würfel und Stern (z.B. in Tabelle 11.1) als getrennte Blöcke behandelt werden. Falls erforderlich, kann der Würfel wie in Absatz 8.2.2 noch weiter unterteilt werden. Die n_0 Punkte im Zentrum werden auf die verschiedenen Blöcke aufgeteilt, d.h. $n_w + n_s = n_0$, wobei n_w (bzw. n_s) die Anzahl Zentrumspunkte im Würfel (bzw. Stern) ist.

Die linearen Terme und die Wechselwirkungsterme im Modell (11.1) sind immer orthogonal zum Blockfaktor und daher unabhängig von evtl. Unterschieden zwischen den Blöcken. Allerdings kann ein Unterschied zwischen dem Würfel- und dem Sternblock zu einer Verfälschung der quadratischen Terme in (11.1) führen. Durch eine geeignete Wahl von α kann diese Verfälschung vermieden werden – man erhält sog. orthogonale Blöcke, wenn

$$\text{orthogonale Blöcke: } \alpha^2 = \frac{N_w \cdot (2k + n_s)}{2 \cdot (N_w + n_w)} \quad (11.5)$$

wobei $N_w + n_w = 2^{k-p} + n_w =$ Anzahl Einzelversuche im Würfelblock und

$2k + n_s =$ Anzahl Einzelversuche im Sternblock.



Ergänzung für mathematisch Interessierte

Ableitung der Bedingung (11.3) für Orthogonalität am Beispiel eines Versuchsplans mit $k = 2$ und $n_0 = 4$:

Wie in Absatz 10.2.2 werden zur Anpassung der nichtlinearen Terme in (11.1) transformierte Variablen eingeführt:

$$x_3 = x_1 \cdot x_2 \quad (\text{wie in 10.2.2})$$

$$x_4 = x_1 \cdot x_1 = x_1^2$$

$$x_5 = x_2 \cdot x_2 = x_2^2$$

Tabelle 11.5 zeigt beispielhaft, dass unabhängig von α und n_0 gilt:

$$Q_{x_1 x_2} = Q_{x_1 x_4} = 0$$

d.h., dass x_1 und x_2 sowie x_1 und x_4 orthogonal zueinander sind. Analog kann man zeigen, dass auch fast alle anderen Spalten orthogonal zueinander sind. Nur die Bedingung $Q_{x_4 x_5} = 0$ für die Orthogonalität zwischen x_4 und x_5 ist nicht für beliebige Werte von α und n_0 erfüllt. Einsetzen von \bar{x}_4 und Auflösen nach α^2 ergibt:

$$\alpha^2 = \frac{1}{2}(\sqrt{12 \cdot 4} - 4) = 1,464 \Rightarrow \alpha = 1,21.$$

Die Verallgemeinerung dieser Rechnung liefert (11.3). Ersetzt man A durch den Blockfaktor, so erhält man (11.5).

TABELLE 11.5 Ableitung der Orthogonalitätsbedingung für den Sonderfall $k = 2$ und $n_0 = 4$: Rechnung ähnlich Tabelle 10.4 – Spalten A und B enthalten den zentral zusammengesetzten Versuchsplan für zwei Faktoren, Spalten x_4 und x_5 deren Quadrate.

| A x_{1i} | B x_{2i} | $(x_{1i} - \bar{x}_1) \cdot (x_{2i} - \bar{x}_2)$ | $x_{4i} = x_{1i} \cdot x_{2i}$ | $(x_{4i} - \bar{x}_4)$ | $(x_{1i} - \bar{x}_1) \cdot (x_{4i} - \bar{x}_4)$ | $x_{5i} = x_{2i} \cdot x_{2i}$ | $(x_{5i} - \bar{x}_5)$ | $(x_{4i} - \bar{x}_4) \cdot (x_{5i} - \bar{x}_5)$ |
|-----------------|-----------------|---|----------------------------------|------------------------|---|--------------------------------|------------------------|---|
| -1 | -1 | 1 | 1 | $1 - \bar{x}_4$ | $-(1 - \bar{x}_4)$ | 1 | $1 - \bar{x}_4$ | $(1 - \bar{x}_4)^2$ |
| 1 | -1 | -1 | 1 | $1 - \bar{x}_4$ | $(1 - \bar{x}_4)$ | 1 | $1 - \bar{x}_4$ | $(1 - \bar{x}_4)^2$ |
| -1 | 1 | -1 | 1 | $1 - \bar{x}_4$ | $-(1 - \bar{x}_4)$ | 1 | $1 - \bar{x}_4$ | $(1 - \bar{x}_4)^2$ |
| 1 | 1 | 1 | 1 | $1 - \bar{x}_4$ | $(1 - \bar{x}_4)$ | 1 | $1 - \bar{x}_4$ | $(1 - \bar{x}_4)^2$ |
| $-\alpha$ | 0 | 0 | α^2 | $\alpha^2 - \bar{x}_4$ | $-\alpha(\alpha^2 - \bar{x}_4)$ | 0 | $-\bar{x}_4$ | $-\bar{x}_4(\alpha^2 - \bar{x}_4)$ |
| α | 0 | 0 | α^2 | $\alpha^2 - \bar{x}_4$ | $\alpha(\alpha^2 - \bar{x}_4)$ | 0 | $-\bar{x}_4$ | $-\bar{x}_4(\alpha^2 - \bar{x}_4)$ |
| 0 | $-\alpha$ | 0 | 0 | $-\bar{x}_4$ | 0 | α^2 | $\alpha^2 - \bar{x}_4$ | $-\bar{x}_4(\alpha^2 - \bar{x}_4)$ |
| 0 | α | 0 | 0 | $-\bar{x}_4$ | 0 | α^2 | $\alpha^2 - \bar{x}_4$ | $-\bar{x}_4(\alpha^2 - \bar{x}_4)$ |
| 0 | 0 | 0 | 0 | $-\bar{x}_4$ | 0 | 0 | $-\bar{x}_4$ | \bar{x}_4^2 |
| 0 | 0 | 0 | 0 | $-\bar{x}_4$ | 0 | 0 | $-\bar{x}_4$ | \bar{x}_4^2 |
| 0 | 0 | 0 | 0 | $-\bar{x}_4$ | 0 | 0 | $-\bar{x}_4$ | \bar{x}_4^2 |
| 0 | 0 | 0 | 0 | $-\bar{x}_4$ | 0 | 0 | $-\bar{x}_4$ | \bar{x}_4^2 |
| $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 4 + 2\alpha^2$ | $\Sigma = 0$ | $\Sigma = 0$ | $\Sigma = 4 + 2\alpha^2$ | $\Sigma = 0$ | Summe |
| $\bar{x}_1 = 0$ | $\bar{x}_2 = 0$ | $Q_{x_1 x_2}$ | $\bar{x}_4 = (2 + \alpha^2) / 6$ | | $Q_{x_4 x_4}$ | $\bar{x}_5 = \bar{x}_4$ | | $Q_{x_4 x_5}$ |

11.1.5 Praxisbeispiel Laserschneiden¹

Aluminiumblech der Dicke 1,5 mm wird mit einem CO₂-Laser (maximale Leistung 1,5 kW) geschnitten. Das Schneiden erfolgt durch Aufschmelzen des Materials und Ausblasen des geschmolzenen Materials mit einem Schneidgas (N₂).

Untersuchungsziele:

- Bei nicht optimalen Betriebsparametern erstarrt das Material beim Ausblasen bereits an der Schnittkante und bildet dort einen scharfen Grat oder Bart. Ziel ist es, diesen Bart zu vermeiden. Als Zielgröße wird die Barthöhe gemessen.
- Außerdem soll die Schnittfläche möglichst glatt sein. Als Zielgröße wird die Rautiefe gemessen, sie soll so klein wie möglich sein.
- Aus Kostengründen soll die Schneidgeschwindigkeit möglichst groß sein.

¹ Vereinfachte Darstellung nach der Diplomarbeit von R. Heigl und R. Schauppel, fachliche Betreuung durch Prof. Dr. H.-A. Schertel, Hochschule Aalen

Der Einfluss folgender Faktoren auf die Zielgrößen sollte untersucht werden:

- Schneidgeschwindigkeit (aufgrund von Vorversuchen wurden die besten Ergebnisse zwischen 2,5 und 5 m/min erwartet)
- Druck des Schneidgases (beste Ergebnisse wurden zwischen 10 und 16 bar erwartet)
- Abstand der Düse vom Blech (bei kleinem Abstand wurde ein gutes Schneidergebnis erwartet; aber bei zu kleinem Abstand kann die Düse das Blech berühren und beschädigt werden – der Abstand 0,3 mm darf daher nicht unterschritten werden, er soll aber in vielen Einzelversuchen untersucht werden)
- Fokuslage (von der Oberkante des Blechs (0 mm) bis zur Unterkante (1,5 mm); das Optimum wurde an einer Grenze erwartet, daher $\alpha = 1$) und
- Laserleistung (beste Ergebnisse wurden bei maximaler Leistung erwartet).

Als Versuchsplan wurde ein zentral zusammengesetzter Plan verwendet. Für den Würfelseit wurde ein fraktioneller faktorieller 2^{5-1} -Versuchsplan mit Auflösung V verwendet (16 Einzelversuche). Das Zentrum wurde dreimal realisiert. Insgesamt ergaben sich damit $16 + 10 + 3 = 29$ Einzelversuche. α für einen orthogonalen Versuchsplan beträgt damit nach (11.3)

$$\alpha^2 = \frac{1}{2} \cdot (\sqrt{29 \cdot 16} - 16) = 2,77 \Rightarrow \alpha = 1,664$$

Die Schneidgeschwindigkeit und der Druck wurden auf 5 Stufen untersucht, wobei die Stufenabstände ungefähr der Orthogonalität entsprachen. Die anderen Faktoren wurden aus technischen Gründen nur auf je 3 Stufen untersucht, wie bei flächenzentrierten Plänen. Der Versuchsplan ist damit zwar nicht völlig orthogonal, die Abweichung wurde jedoch in Kauf genommen und beeinträchtigt die Auswertung nicht wesentlich. Tabelle 11.6 zeigt die verwendeten Stufenwerte der $k = 5$ Faktoren, Tabelle 11.7 die Einzelversuche und ihre Ergebnisse (zur besseren Übersicht in systematischer Reihenfolge).

TABELLE 11.6 Stufenwerte für die $k = 5$ Faktoren

| Faktor | | Stufenwerte | | | | |
|-------------------------|-----|-------------|------|-----|-----|--|
| Geschwindigkeit [m/min] | 1,5 | 2,5 | 3,75 | 5,0 | 6,0 | |
| Druck [bar] | 8 | 10 | 13 | 16 | 18 | |
| Abstand [mm] | | 0,3 | 0,7 | 1,1 | | |
| Fokuslage [mm] | | 0,0 | 0,75 | 1,5 | | |
| Leistung [kW] | | 1,0 | 1,25 | 1,5 | | |

TABELLE 11.7 Versuchsplan und Ergebnisse in systematischer Reihenfolge

| Nr. | Geschwindigkeit [m/min] | Druck [bar] | Abstand [mm] | Fokuslage [mm] | Leistung [kW] | Barthöhe [mm] | Rautiefe [µm] |
|-----|-------------------------|-------------|--------------|----------------|---------------|---------------|---------------|
| 1 | 2,5 | 10 | 0,3 | 0,0 | 1,5 | 0,15 | 0,2 |
| 2 | 5,0 | 10 | 0,3 | 0,0 | 1,0 | 2,65 | 10,2 |
| 3 | 2,5 | 16 | 0,3 | 0,0 | 1,0 | 0,75 | 7,0 |
| 4 | 5,0 | 16 | 0,3 | 0,0 | 1,5 | 0,05 | 7,2 |

TABELLE 11.7 Versuchsplan und Ergebnisse in systematischer Reihenfolge (*Fortsetzung*)

| Nr. | Geschwindigkeit [m/min] | Druck [bar] | Abstand [mm] | Fokuslage [mm] | Leistung [kW] | Barthöhe [mm] | Rautiefe [µm] |
|-----|-------------------------|-------------|--------------|----------------|---------------|---------------|---------------|
| 5 | 2,5 | 10 | 1,1 | 0,0 | 1,0 | 1,20 | 9,2 |
| 6 | 5,0 | 10 | 1,1 | 0,0 | 1,5 | 0,85 | 6,7 |
| 7 | 2,5 | 16 | 1,1 | 0,0 | 1,5 | 1,00 | 3,5 |
| 8 | 5,0 | 16 | 1,1 | 0,0 | 1,0 | 2,95 | 13,8 |
| 9 | 2,5 | 10 | 0,3 | 1,5 | 1,0 | 0,70 | 4,9 |
| 10 | 5,0 | 10 | 0,3 | 1,5 | 1,5 | 0,00 | 0,0 |
| 11 | 2,5 | 16 | 0,3 | 1,5 | 1,5 | 0,00 | 0,0 |
| 12 | 5,0 | 16 | 0,3 | 1,5 | 1,0 | 2,10 | 6,3 |
| 13 | 2,5 | 10 | 1,1 | 1,5 | 1,5 | 0,50 | 0,0 |
| 14 | 5,0 | 10 | 1,1 | 1,5 | 1,0 | 1,90 | 4,5 |
| 15 | 2,5 | 16 | 1,1 | 1,5 | 1,0 | 0,45 | 5,8 |
| 16 | 5,0 | 16 | 1,1 | 1,5 | 1,5 | 0,00 | 2,4 |
| 17 | 1,5 | 13 | 0,7 | 0,75 | 1,25 | 0,60 | 2,1 |
| 18 | 6,0 | 13 | 0,7 | 0,75 | 1,25 | 1,95 | 4,3 |
| 19 | 3,75 | 8 | 0,7 | 0,75 | 1,25 | 0,15 | 6,4 |
| 20 | 3,75 | 18 | 0,7 | 0,75 | 1,25 | 0,45 | 7,8 |
| 21 | 3,75 | 13 | 0,3 | 0,75 | 1,25 | 0,30 | 2,6 |
| 22 | 3,75 | 13 | 1,1 | 0,75 | 1,25 | 0,65 | 5,2 |
| 23 | 3,75 | 13 | 0,7 | 0,0 | 1,25 | 0,75 | 5,6 |
| 24 | 3,75 | 13 | 0,7 | 1,5 | 1,25 | 0,00 | 1,8 |
| 25 | 3,75 | 13 | 0,7 | 0,75 | 1,0 | 1,30 | 6,3 |
| 26 | 3,75 | 13 | 0,7 | 0,75 | 1,5 | 0,00 | 0,0 |
| 27 | 3,75 | 13 | 0,7 | 0,75 | 1,25 | 0,30 | 3,7 |
| 28 | 3,75 | 13 | 0,7 | 0,75 | 1,25 | 0,25 | 2,6 |
| 29 | 3,75 | 13 | 0,7 | 0,75 | 1,25 | 0,40 | 5,1 |

Bei der Auswertung werden die beiden Zielgrößen getrennt voneinander behandelt. Bild 11-2 zeigt die Regressionskoeffizienten für die Zielgröße Barthöhe, Bild 11-3 für die Zielgröße Rautiefe. Zur Beurteilung der Signifikanz werden die Koeffizienten jeweils mit der Breite der Vertrauensbereiche verglichen, wie in Kapitel 7.

Die Breite der Vertrauensbereiche für die Koeffizienten ist unterschiedlich. Insbesondere bei den drei flächenzentrierten Faktoren ist die Breite der Vertrauensbereiche der quadratischen Terme deutlich größer als für die anderen Koeffizienten. Dies liegt z.T. am geringeren Wertebereich der Variablen und z.T. an der fehlenden Orthogonalität. Fehlende Orthogonalität führt immer zu einer Verbreiterung der Vertrauensbereiche.

Die Residuen wurden wie in Absatz 10.1.3 untersucht, zeigten aber keine Besonderheiten und werden daher hier nicht dargestellt.

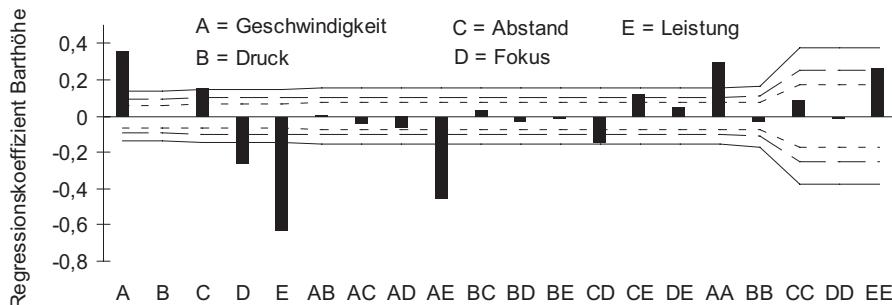


BILD 11-2 Regressionskoeffizienten für die Zielgröße Barthöhe und ihre Vertrauensbereiche:

*** : A, C, D, E, AE, AA

** : CD, CE, EE

Der Druck B hat keinen signifikanten Einfluss auf die Barthöhe.

Anmerkung zur Bezeichnung:

Effekte und Regressionskoeffizienten können ineinander umgerechnet werden. Mit A wird hier der Regressionskoeffizient für den linearen Term in der Geschwindigkeit x_1 (in normierten Stufen), mit AE der Koeffizient des Produktes aus Geschwindigkeit und Leistung (= Wechselwirkung, vgl. Absatz 10.2.2), mit AA der Koeffizient von x_1^2 im Regressionsmodell (11.1) bezeichnet, usw.

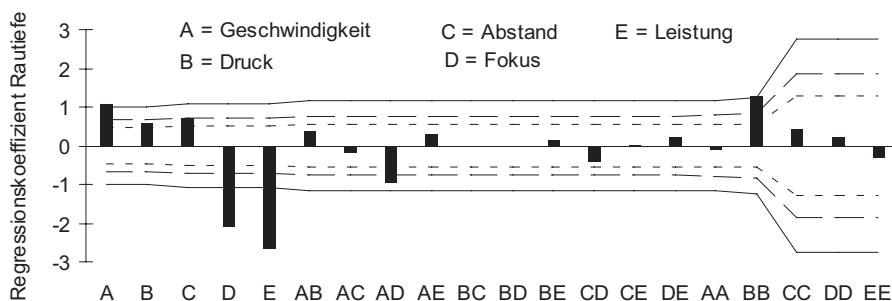


BILD 11-3 Regressionskoeffizienten für die Zielgröße Rautiefe und ihre Vertrauensbereiche:

*** : A, D, E, BB

** : AD

* : B, C

Bild 11-2 und 11-3 enthalten die Information, welche Faktoren einen signifikanten Einfluss auf die jeweilige Zielgröße haben. Für die Optimierung stellt man die signifikanten Effekte wie folgt grafisch dar.

Der Einfluss von zwei beliebigen Faktoren auf eine Zielgröße kann anschaulich als Fläche im Raum (Wirkungsfläche = response surface, wie in Bild 1-6) dargestellt werden. Für eine quantitative Beurteilung ist jedoch eine Darstellung als Höhenliniendiagramm (= contour plot, wie in Bild 11-4) besser geeignet. Die Wirkungsfläche ist wie ein Photo von einer Landschaft, das Höhenliniendiagramm wie eine Landkarte – das Photo gibt zwar einen anschaulichen Eindruck, zur Orientierung ist die Landkarte aber besser geeignet.

Die jeweils nicht dargestellten Faktoren werden auf vorgegebenen Werten festgehalten. Im vorliegenden Beispiel mit 5 Faktoren gibt es für jede der beiden Zielgrößen 10 Möglichkeiten, jeweils 2 Faktoren für die grafische Darstellung auszuwählen. Da die anderen

Faktoren auf verschiedenen Werten festgehalten werden können, gibt es eine unüberschaubare Vielfalt denkbarer Darstellungen.

Im Folgenden wird an diesem Beispiel eine systematische Vorgehensweise gezeigt, mit der eine Faktorstufenkombination gefunden werden kann, die mehrere Zielgrößen gleichzeitig optimiert. Grundlage der Auswahl der Faktorpaare für die grafische Darstellung sind dabei die Zwei-Faktor-Wechselwirkungen.

Prinzipielle Vorgehensweise

1. Nicht signifikante Wechselwirkungen (= Produkte wie $x_1 \cdot x_2$) und quadratische Terme werden schrittweise aus dem Regressionsmodell entfernt (Achtung, dabei kann sich die Signifikanz der verbliebenen Terme verändern, vor allem wenn der Versuchsplan nicht orthogonal war). Lineare Terme (= Effekte der Faktoren) werden nur entfernt, wenn auch alle Wechselwirkungen und quadratischen Terme dieses Faktors bereits entfernt wurden (z.B. der Druck bei der Zielgröße Barthöhe).
2. Ausgehend von Faktoren mit den größten Wechselwirkungen, wird die Abhängigkeit der Zielgrößen von jeweils zwei Faktoren als Höhenliniendiagramm grafisch dargestellt. Die anderen Faktoren werden dabei zunächst auf der mittleren Stufe (0) festgehalten.
3. Schrittweise werden für die Faktoren optimale Stufenwerte festgelegt. Bei grafischen Darstellungen der anderen Faktoren werden dann diese Stufenwerte verwendet.

Anwendung auf das Beispiel

1. Bei der Zielgröße Barthöhe werden alle Terme außer A, C, D, E, AE, AA, EE, CD und CE aus dem Modell entfernt. Bei der Rautiefe werden alle Terme außer A, B, C, D, E, AD und BB entfernt (vgl. Bilder 11-2 und 11-3).
2. Da die größte Wechselwirkung (relativ zu den linearen Termen) bei der Zielgröße Barthöhe zwischen der Geschwindigkeit und der Leistung besteht (AE), wird mit dieser Darstellung begonnen. Bild 11-4 zeigt das Regressionsmodell für die Barthöhe, Bild 11-5 für die Rautiefe in Abhängigkeit von Geschwindigkeit und Leistung, die anderen Faktoren werden jeweils auf der mittleren Stufe festgehalten.

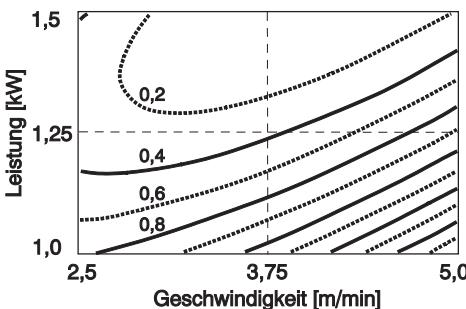
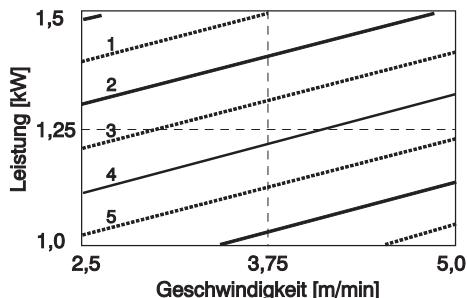


BILD 11-4

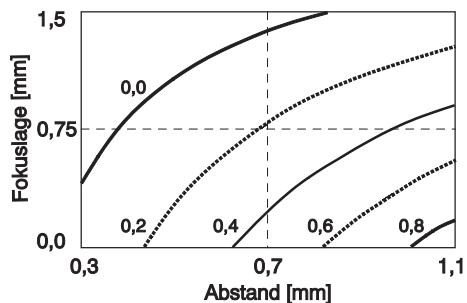
Barthöhe in Abhängigkeit von Geschwindigkeit und Leistung:
Die maximale Leistung 1,5 kW ergibt die kleinste Barthöhe; die optimale Geschwindigkeit hängt von der Leistung ab.

**BILD 11-5**

Rautiefe in Abhängigkeit von Geschwindigkeit und Leistung:

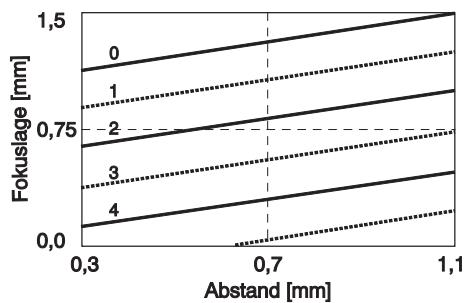
Die maximale Leistung 1,5 kW (und eine niedrige Geschwindigkeit) ergeben die kleinste Rautiefe.

- Da die maximale Leistung von 1,5 kW für beide Zielgrößen günstig ist, wird sie für alle weiteren Darstellungen verwendet. Alle weiteren Wechselwirkungen mit der Leistung E sind damit ebenfalls erfasst (d.h. CE). Die nächstgrößte Wechselwirkung ist CD, daher werden in Bild 11-6 und 11-7 Barthöhe bzw. Rautiefe gegen die Faktoren Abstand (= C) und Fokuslage (= D) dargestellt (für die optimale Leistung von 1,5 kW, die Zielgeschwindigkeit von 5 m/min und einen mittleren Druck von 13 bar).

**BILD 11-6**

Barthöhe in Abhängigkeit von Abstand und Fokuslage:

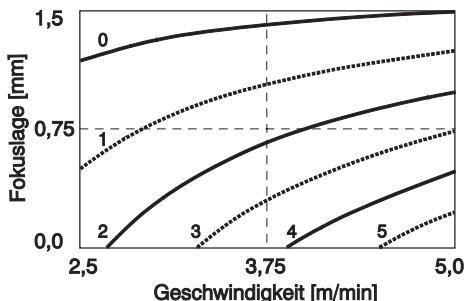
Die Fokuslage 1,5 mm im Blech ergibt die kleinste Barthöhe (Leistung 1,5 kW, Geschwindigkeit 5 m/min).

**BILD 11-7**

Rautiefe in Abhängigkeit von Abstand und Fokuslage:

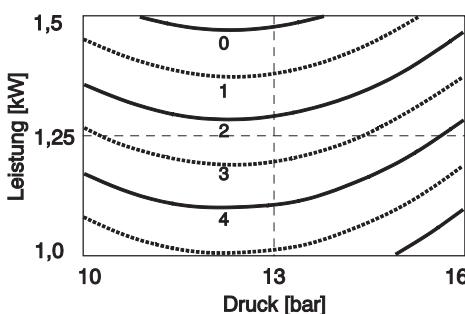
Die Fokuslage 1,5 mm im Blech ergibt die kleinste Rautiefe (Leistung 1,5 kW, Geschwindigkeit 5 m/min).

Da die Fokuslage von 1,5 mm für beide Zielgrößen günstig ist, wird dieser Wert für alle weiteren Darstellungen verwendet. Damit sind bei der Barthöhe alle signifikanten Wechselwirkungen erfasst. Als zusätzliche Absicherung wird in Bild 11-8 die Wechselwirkung AD bei der Rautiefe dargestellt. Bei einer Fokuslage von 1,5 mm hängt die Rautiefe fast nicht von der Geschwindigkeit ab – ein weiterer Vorteil der Fokuslage 1,5 mm.

**BILD 11-8**

Rautiefe in Abhängigkeit von Geschwindigkeit und Fokuslage:
Die Fokuslage 1,5 mm im Blech ergibt die kleinste Rautiefe und eine geringe Abhängigkeit von der Geschwindigkeit (Leistung 1,5 kW, Abstand 1,1 mm, Druck 13 bar).

Der Druck B beeinflusst nur die Rautiefe und hat keine signifikanten Wechselwirkungen. Damit kann er ohne Rücksicht auf andere Faktoren so festgelegt werden, dass die Rautiefe minimiert wird. Bild 11-9 zeigt die Abhängigkeit der Rautiefe vom Druck und als Beispiel von der Leistung. Unabhängig von der Leistung ist ein Druck von ca. 12,5 bar optimal.

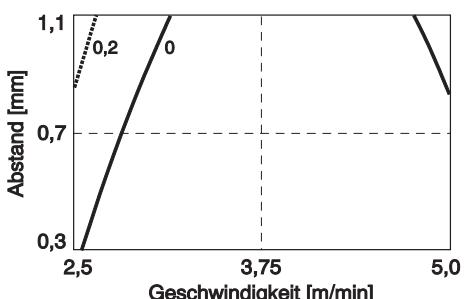
**BILD 11-9**

Rautiefe in Abhängigkeit von Druck und Leistung:
Unabhängig von der Leistung ist ein Druck von ca. 12,5 bar optimal (Geschwindigkeit 5 m/min, Fokuslage 1,5 mm, Abstand 1,1 mm).

Damit sind folgende Faktorwerte festgelegt:

- Leistung 1,5 kW
- Fokuslage 1,5 mm
- Druck 12,5 bar.

Bild 11-10 zeigt die Abhängigkeit der Barthöhe von den verbliebenen Faktoren Geschwindigkeit und Abstand für diese Werte von Leistung, Fokuslage und Druck. Die Rautiefe ist im gesamten Wertebereich optimal.

**BILD 11-10**

Barthöhe in Abhängigkeit von Geschwindigkeit und Abstand
(Leistung 1,5 kW, Fokuslage 1,5 mm, Druck 12,5 bar).

Damit erhält man folgenden optimalen Kompromiss, der bartfreien Schnitt bei gleichzeitig minimaler Rautiefe und geringen Fertigungsproblemen erlaubt:

- Leistung: 1,5 kW (Maximalwert für Anlage)
- Fokuslage: 1,5 mm (an Blechunterkante – dünnflüssige Schmelze)
- Geschwindigkeit: ca. 4,5 m/min (bartfrei, Rautiefe nicht mehr messbar, hoher Fertigungsdurchsatz)
- Druck: 12 – 13 bar (günstig für Rautiefe)
- Abstand: ca. 1 mm (ergibt Fertigungssicherheit, die anderen Ziele sind bereits erreicht)

Mit diesem Beispiel soll gezeigt werden, wie durch quantitative Kenntnis der Zusammenhänge zwischen den Faktoren und den Zielgrößen auch widersprüchliche Ziele miteinander vereinbart werden können. Optimale Kompromisse können gefunden werden.

Oft gelingt es Zielkonflikte zu lösen, einfach weil man die Zusammenhänge genau kennt. Dadurch kann eine wesentliche Verbesserung bestehender Prozesse erreicht werden. Neue Prozesse können von Anfang an optimal gefahren werden. So werden Entwicklungszeiten verkürzt und Kosten gesenkt. Die Einsparungen übersteigen die Versuchskosten oft um ein Vielfaches.

Bei der Interpretation der Ergebnisse beachte man jedoch:

- Die dargestellten Höhenliniendiagramme stellen die besten Schätzwerte aufgrund der Versuchsergebnisse dar. Sie unterliegen zufälligen Streuungen (Vertrauensbereiche).
- Die Vertrauensbereiche geben den Bereich an, der die Schätzwerte (für den Mittelwert) mit vorgegebener Wahrscheinlichkeit enthält. Einzelwerte streuen in einem wesentlich breiteren Bereich (Vorhersagebereich für Einzelwerte, vgl. Absatz 10.1.4).
- Die Versuchsergebnisse werden mit einem empirischen Modell beschrieben. Daher ist keine Extrapolation zulässig. Im untersuchten Bereich erhält man meist eine gute Beschreibung der Zusammenhänge. Manche Formen der Abhängigkeit der Zielgröße von den Faktoren (z. B. Sättigungsverhalten) lassen sich durch das verwendete quadratische Modell nur unvollkommen abbilden. Daher ist technischer Sachverstand bei der Interpretation der Ergebnisse erforderlich (vgl. Abschnitt 11.3).

Bei einer größeren Anzahl von Zielgrößen wird das hier beschriebene Optimierungsverfahren zu unübersichtlich. Dies gilt insbesondere, wenn für eine Zielgröße ein bestimmter Zielwert vorgegeben ist. Dann hilft das Verfahren in Abschnitt 15.3 weiter.

■ 11.2 Alternative Pläne

In diesem Abschnitt werden kurz der Aufbau, sowie die Vor- und Nachteile einiger Alternativen zu zentral zusammengesetzten Plänen behandelt.

11.2.1 3^k - und 3^{k-p} -Pläne

Um eine nichtlineare Abhängigkeit erkennen zu können, sind mindestens drei Stufen je Faktor erforderlich. Vollständige faktorielle Versuchspläne enthalten alle 3^k Faktorstufenkombinationen der k Faktoren. Nachteil dieser Pläne ist, dass die Anzahl der Faktorstufenkombinationen sehr schnell mit der Anzahl der Faktoren zunimmt. Schon für $k = 4$ erhält man 81 Faktorstufenkombinationen und damit einen praktisch kaum mehr anwendbaren Versuchsplan.

Analog zu fraktionellen faktoriellen 2^{k-p} -Plänen gibt es 3^{k-p} -Pläne [3]. Nachteil dieser Pläne ist, dass die Effekte von 2FWW und Faktoren miteinander vermengt sind. Daher sind 3^{k-p} -Pläne für quantitative Faktoren nicht empfehlenswert. Es wäre inkonsistent, im Modell (11.1) die quadratischen Terme in den Faktoren zu ermitteln, die Wechselwirkungen aber zu ignorieren (die 2FWW wie $x_1 \cdot x_2$ sind Produkte von zwei Faktoren, genau wie die quadratischen Terme). 3^{k-p} -Pläne werden daher hier nicht weiter behandelt.

Für Screening-Versuche für mehrstufige qualitative Faktoren sind die 3^{k-p} -Pläne jedoch sinnvoll (vgl. Kapitel 13).

11.2.2 Box-Behnken-Pläne

Ähnlich zu den 3^{k-p} -Plänen sind Box-Behnken-Pläne ([1], [3]) eine Auswahl aus den 3^k Faktorstufenkombinationen eines vollständigen faktoriellen Plans. Die Auswahl ist allerdings so optimiert, dass 2FWW bestimmt werden können.

Tabelle 11.8 zeigt den Box-Behnken-Plan für $k = 3$ Faktoren. Der Versuchsplan besteht aus einem vollständigen faktoriellen Versuchsplan für jedes der drei möglichen Paare von jeweils zwei Faktoren (3×4 Faktorstufenkombinationen) und z.B. drei Realisierungen des Zentrums.

Bei $k = 4$ Faktoren besteht der Box-Behnken-Plan entsprechend aus 6×4 Faktorstufenkombinationen für die sechs möglichen Faktorpaare. Bei sechs und mehr Faktoren setzt sich der Box-Behnken-Plan aus vollständigen faktoriellen Versuchsplänen von Dreiergruppen der Faktoren zusammen.

| Nr. | A | B | C | Erläuterung |
|-----|-----|-----|---|---------------|
| 1 | - 1 | - 1 | 0 | Faktoren A, B |
| 2 | 1 | - 1 | 0 | |
| 3 | - 1 | 1 | 0 | |
| 4 | 1 | 1 | 0 | |

TABELLE 11.8

Box-Behnken-Plan für drei Faktoren A, B und C.

| Nr. | A | B | C | Erläuterung |
|-----|----|----|----|---------------|
| 5 | -1 | 0 | -1 | Faktoren A, C |
| 6 | 1 | 0 | -1 | |
| 7 | -1 | 0 | 1 | |
| 8 | 1 | 0 | 1 | |
| 9 | 0 | -1 | -1 | Faktoren B, C |
| 10 | 0 | 1 | -1 | |
| 11 | 0 | -1 | 1 | |
| 12 | 0 | 1 | 1 | |
| 13 | 0 | 0 | 0 | Zentrum |
| 14 | 0 | 0 | 0 | |
| 15 | 0 | 0 | 0 | |

Vorteile

- Bei Box-Behnken-Plänen werden alle Faktoren nur auf drei Stufen verändert, dies ist in manchen Anwendungen von Vorteil (vgl. flächenzentrierte zentral zusammengesetzte Pläne).
- Die Gesamtzahl der Faktorstufenkombinationen ist (etwas) kleiner als bei zentral zusammengesetzten Versuchsplänen.

Nachteile

- Die quadratischen Effekte der Box-Behnken-Pläne sind nicht orthogonal.
- Box-Behnken-Pläne erlauben keine Überprüfung des quadratischen Modellansatzes (11.1).
- Die Vertrauensbereiche sind breiter als bei zentral zusammengesetzten Plänen mit vergleichbarer Gesamtzahl von Versuchen – dies gilt insbesondere, wenn die Anzahl k der Faktoren groß ist.

Es wird empfohlen, Box-Behnken-Pläne nur in Ausnahmefällen einzusetzen (für $k = 3$ oder 4 Faktoren, wenn die Anzahl der Faktorstufen aus technischen Gründen auf 3 beschränkt ist).

11.2.3 Kleine zusammengesetzte Pläne

Diese Pläne (Draper und Lin [5]) sind ähnlich zu den zentral zusammengesetzten Plänen. Als Würfel wird jedoch (statt des vollständigen faktoriellen Plans oder des Plans der Auflösung V) ein fraktioneller faktorieller 2^{k-p} -Plan der Auflösung III oder ein Plackett-Burman-Plan verwendet. Aufgrund des Sterns kann trotzdem zwischen 2FWW und Faktoren unterschieden werden. Tabelle 11.9 zeigt als Beispiel den kleinen zusammengesetzten Plan für 4 Faktoren.

| Nr. | A | B | C | D | |
|-----|-----------|-----------|-----------|-----------|--------------------------|
| 1 | 1 | 1 | 1 | -1 | Würfel |
| 2 | 1 | 1 | -1 | -1 | |
| 3 | 1 | -1 | 1 | 1 | |
| 4 | -1 | 1 | -1 | 1 | |
| 5 | 1 | -1 | -1 | 1 | |
| 6 | -1 | -1 | 1 | -1 | |
| 7 | -1 | 1 | 1 | 1 | |
| 8 | -1 | -1 | -1 | -1 | |
| 9 | $-\alpha$ | 0 | 0 | 0 | Stern |
| 10 | α | 0 | 0 | 0 | |
| 11 | 0 | $-\alpha$ | 0 | 0 | |
| 12 | 0 | α | 0 | 0 | |
| 13 | 0 | 0 | $-\alpha$ | 0 | |
| 14 | 0 | 0 | α | 0 | |
| 15 | 0 | 0 | 0 | $-\alpha$ | |
| 16 | 0 | 0 | 0 | α | |
| 17 | 0 | 0 | 0 | 0 | Zentrum ($n_0 = 2$) |
| 18 | 0 | 0 | 0 | 0 | |

TABELLE 11.9
Kleiner zusammengesetzter Plan für 4 Faktoren

Der wichtigste Vorteil dieser Pläne ist die kleine Anzahl Faktorstufenkombinationen. Tabelle 11.10 vergleicht den Mindestversuchsumfang für einen normalen zentral zusammengesetzten Plan (aus Abschnitt 11.1) mit dem eines kleinen Plans für die gleiche Faktorenanzahl.

| Anzahl Faktoren k | Mindestversuchsumfang N_{\min} | |
|-------------------|----------------------------------|--------------|
| | Auflösung V | kleiner Plan |
| 2 | 9 | - |
| 3 | 15 | - |
| 4 | 25 | 17 |
| 5 | 27 | 23 |
| 6 | 45 | 29 |
| 7 | 79 | 39 |
| 8 | 81 | 53 |

TABELLE 11.10
Mindestversuchsumfang N_{\min} ($n_0 = 1$) in einem zentral zusammengesetzten Plan mit Würfel der Auflösung V, verglichen mit einem kleinen Plan mit Würfel der Auflösung III

Wichtigster Nachteil der kleinen Pläne ist, dass sie nicht orthogonal sind. Der Parameter α kann zwar mit Gleichung (11.3) so berechnet werden, dass der Plan „orthogonal“ ist. Dies bezieht sich jedoch nur auf die quadratischen Effekte. Insbesondere bei $k = 4$ und 6 korrelieren manche Faktoren sehr stark mit 2FWW (im Plan von Tabelle 11.9 unterscheiden sich z. B. der Faktor A und die 2FWW BD nur im Sternteil des Plans und können daher nur schlecht getrennt werden).

Für $k = 7$ und 8 Faktoren ist die Korrelation zwischen den Effekten so klein und der Aufwand für die normalen Pläne so groß, dass die kleinen Pläne vertretbar sind (wenn man so viele Faktoren überhaupt gleichzeitig untersucht). Für $k = 4$ bis 6 Faktoren werden die zentral zusammengesetzten Pläne aus Abschnitt 11.1 empfohlen.

11.2.4 Optimale Pläne

Optimale Pläne (vgl. z.B. [2, 3, 6, 7]) bieten dem Experimentator im Prinzip volle Freiheit bei der Festlegung des Plans. Er gibt

- eine (meist große) Anzahl von Faktorstufenkombinationen als Kandidaten,
- das Modell, das an die Ergebnisse angepasst werden soll (z.B. 11.1), und
- den Versuchsumfang N

vor. Bei der Vorgabe der Kandidaten können für jeden Faktor beliebige Stufenwerte festgelegt werden. Alle Kombinationen dieser Faktorstufen sind zunächst möglich (vollständig faktoriell). Anschließend können jedoch wieder (fast) beliebig Faktorstufenkombinationen aus der Kandidatenliste entfernt werden (z.B. weil man weiß oder erwartet, dass bestimmte Extremkombinationen der Faktorstufen keine sinnvollen Ergebnisse liefern werden).

Die N Einzelversuche werden iterativ so aus der Kandidatenliste ausgewählt, dass sie optimal für die Anpassung des vorgesehenen Regressionsmodells geeignet sind. Folgende Optimalitätskriterien sind üblich:

A-optimal: Die Breite der Vertrauensbereiche für die Regressionskoeffizienten ist im Mittel so klein wie möglich (Mittelwert der Quadrate)

D-optimal: Das Volumen des gemeinsamen Vertrauensbereichs für die Regressionskoeffizienten ist so klein wie möglich (Produkt)

E-optimal: Die größte Breite des Vertrauensbereichs für einen Regressionskoeffizienten ist so klein wie möglich (maximaler Eigenwert der Informationsmatrix)

G-optimal: Die größte Breite des Vertrauensbereichs für das Regressionsmodell im Prognosebereich ist so klein wie möglich (möglichst gleicher Hebel für alle Versuchspunkte)

I-optimal: Die mittlere Breite des Vertrauensbereichs für das Regressionsmodell im Prognosebereich ist so klein wie möglich (gewichteter Mittelwert der Quadrate) – auch V-, IV- oder Q-optimal genannt.

S-optimal: Die Punkte werden gleichmäßig im Versuchsraum verteilt – der harmonische mittlere Abstand der Punkte von allen anderen wird maximiert.

A-, D- und E-Optimalität bezieht sich auf die Koeffizienten des Regressionsmodells (entsprechend Gleichung 10.19), G- und I-Optimalität dagegen auf das angepasste Modell (entsprechend 10.20). Da diese Optimalitätskriterien vergleichbar gut zur Beurteilung eines Versuchsplans geeignet sind und der Rechenaufwand für D-optimale Pläne am kleinsten ist, sind diese am weitesten verbreitet. S-optimale Pläne eignen sich für die in Abschnitt 17.3 beschriebenen Ansätze, sind aber für die hier beschriebenen quadratischen Modelle weniger geeignet.

Vorteile

- Der Versuchsräum kann beliebig vorgegeben werden (d.h. man kann bereits bei der Planung berücksichtigen, wenn bestimmte Faktorstufenkombinationen nicht möglich oder sinnvoll sind – „constraints“ genannt, vgl. Beispiel 20.1).
- Bereits vorhandene Versuchsergebnisse (z.B. aus Screening-Versuchen) können als so genannte „inclusions“ berücksichtigt werden. Man kann Faktorstufenkombinationen auswählen, die die vorhandenen Ergebnisse möglichst optimal ergänzen.
- Der Modellansatz kann beliebig vorgegeben werden, man muss nicht wie in (11.1) alle Terme bis zur 2. Ordnung (d.h. A, AA, AB, AC, ..., B, BB, BC, ...) berücksichtigen und kann auch höhere Terme aufnehmen.
- Man kann auch mehr als zweistufige quantitative und qualitative Faktoren kombinieren.
- Die Anzahl der Einzelversuche kann beliebig vorgegeben werden (oberhalb einer Mindestzahl, die vom Modell abhängt).

Nachteile

- Der Versuchsplan ist nicht orthogonal, die Abweichungen sind aber meist nur klein.
- Der richtige Modellansatz ist normalerweise nicht im Voraus bekannt, die Optimalität gilt jedoch nur für das gewählte Modell; Abweichungen von diesem Modell können später aus den Versuchsergebnissen nur schlecht erkannt werden.
- Optimale Pläne können nur mit entsprechender Software realisiert werden.

Insgesamt gesehen sind optimale Pläne nützlich, wenn aus technischen Gründen nicht alle Punkte eines zentral zusammengesetzten Plans realisierbar sind (Beispiel 20.1), vorhandene Ergebnisse ergänzt (Kapitel 16) oder verschiedene Arten von Faktoren kombiniert werden sollen. Für Standardsituationen sind zentral zusammengesetzte Pläne m.E. jedoch besser geeignet.

■ 11.3 Grenzen des quadratischen Modells

Bilder 11-4 bis 11-10 sind Beispiele für Abhängigkeiten, die mit dem quadratischen Modell (11.1) beschrieben werden können.

Die Terme im Modell (11.1) sind die ersten Terme einer Potenzreihe. Solange der Wertebereich der Faktoren nicht zu groß ist, erlaubt es meist eine gute empirische Beschreibung des Zusammenhangs zwischen den Faktoren und den Zielgrößen. Es darf jedoch nicht über den untersuchten Bereich hinaus extrapoliert werden.

Wenn der Wertebereich der Faktoren groß ist, kann es vorkommen, dass (11.1) bereits im untersuchten Bereich keine ausreichend gute Näherung ergibt. Im Folgenden werden beispielhaft Zusammenhänge gezeigt, die durch das quadratische Modell nicht vollständig beschrieben werden können.

11.3.1 Beispiele und Lösungsmöglichkeiten

Beispiel 1: Nicht-elliptische Höhenlinien

Bild 11-11 zeigt in Höhenliniendarstellung die Abhängigkeit einer Zielgröße von zwei Faktoren. Ein Maximum in dieser „Bananenform“ kann nicht über den gesamten Wertebereich der Faktoren durch ein quadratisches Modell beschrieben werden. Mit dem quadratischen Modell können nur elliptische Maxima beschrieben werden.

Lösungsmöglichkeiten

- Teilbereiche von Bild 11-11 können mit dem quadratischen Modell beschrieben werden. Wählt man den Abstand der Stufenwerte nur etwa halb so groß wie in Bild 11-11, so erhält man eine ausreichend gute Beschreibung in diesem kleineren Bereich. Eine Extrapolation ist natürlich nicht zulässig.
- Wenn bereits vor der Durchführung der Versuche z.B. aufgrund physikalischer Gesetze eine bestimmte Form der Abhängigkeit erwartet wird, kann man die Beschreibung durch das quadratische Modell verbessern, wenn man die Faktoren wie in Absatz 11.1.3 geeignet definiert.

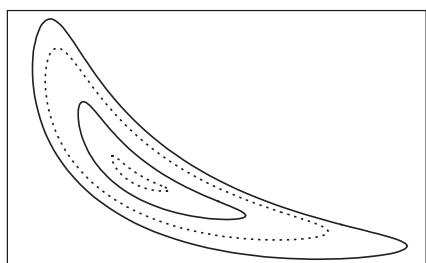


BILD 11-11

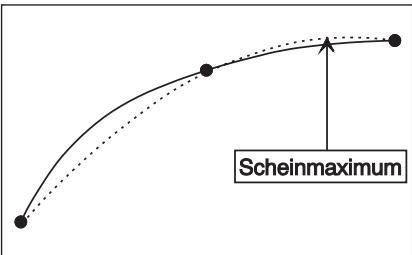
Ein Maximum mit einer Struktur (hier z. B. Bananenform) kann nicht als Ganzes durch ein quadratisches Modell beschrieben werden, Ausschnitte können jedoch ausreichend genau beschrieben werden.

Beispiel 2: Sättigung

Bild 11-12 zeigt die Abhängigkeit einer Zielgröße von einem Faktor, wie sie typischerweise bei einem Sättigungsverhalten auftritt (z.B. die Spannung an einem Kondensator als Funktion der Zeit nach dem Anlegen der Versorgungsspannung, oder die Ausbeute als Funktion der Katalysatormenge oder der Zeit). Mit dem quadratischen Modell wird die Abflachung im Sättigungsbereich nicht richtig beschrieben. Die an drei Punkten angepasste Näherung kann ein Maximum besitzen, das nicht wirklich vorhanden ist.

Lösungsmöglichkeiten

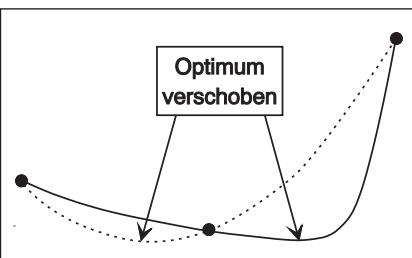
- Zeigt das angepasste Modell bezüglich eines Faktors ein Maximum wie in Bild 11-12, obwohl aus technischen Gründen eine Sättigung erwartet wird, so ist dies vermutlich auf die quadratische Näherung zurückzuführen. Bei der Interpretation ist das Maximum zu ignorieren, das Optimum liegt dann am Rand des untersuchten Bereichs (oder noch weiter in dieser Richtung).
- Wenn bereits vor der Durchführung der Versuche (z.B. aufgrund physikalischer Gesetze) ein Sättigungsverhalten erwartet wird, kann man die Beschreibung durch das quadratische Modell verbessern, wenn man die Faktoren wie in Absatz 11.1.3 geeignet definiert.

**BILD 11-12**

Zielgröße in Abhängigkeit von einem Faktor bei Sättigungsverhalten

— Sättigungskurve

---- quadratische Näherung durch die 3 markierten Punkte

**BILD 11-13**

Zielgröße in Abhängigkeit von einem Faktor mit einer technischen Grenze

— wahre Abhängigkeit

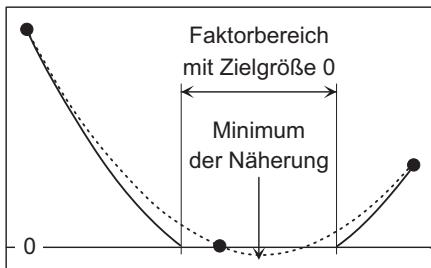
---- quadratische Näherung durch die 3 markierten Punkte

Beispiel 3: Asymmetrie, „Umkippen“

Bild 11-13 zeigt die Abhängigkeit einer Zielgröße von einem Faktor, wie sie typischerweise auftritt, wenn der Faktor eine technische Grenze überschreitet. Gesucht ist z.B. das Minimum der Zielgröße. Über einen weiten Bereich des Faktors erhält man eine allmähliche Verbesserung. Wird jedoch eine bestimmte Grenze überschritten, so tritt eine plötzliche Verschlechterung ein. Ist ein solcher schlechter Wert im untersuchten Wertebereich des Faktors enthalten, so hat dieses Ergebnis einen großen Einfluss auf das angepasste Modell. Da das quadratische Modell immer symmetrisch zum Optimum ist, schiebt ein solcher schlechter Wert das rechnerische Optimum in die entgegengesetzte Richtung.

Lösungsmöglichkeiten

- Weichen Versuchsergebnisse für eine Faktorstufe am Rand des untersuchten Bereichs deutlich von den anderen ab, so kann dies an der Überschreitung einer solchen Grenze liegen. Bei der Interpretation ist zu berücksichtigen, dass das Optimum u.U. näher an diesem Rand liegt, als sich rein rechnerisch ergibt. Zusätzliche Einzelversuche können bei der Suche nach dem Optimum helfen. Manchmal kann es auch sinnvoll sein, solche Ergebnisse als „Ausreißer“ zu löschen (vgl. Beispiel 20.1).
- Wenn bereits vor der Durchführung der Versuche z. B. aufgrund physikalischer Gesetze eine Grenze erwartet wird, ist es besser, mit den Faktorstufen diese Grenze nicht zu überschreiten (vgl. Absatz 3.3.4 – kein neues physikalisches oder technisches Phänomen im Untersuchungsbereich – und Beispiel 20.1).

**BILD 11-14**

Zielgröße, für die keine negativen Werte möglich sind, in Abhängigkeit von einem Faktor
 — wahre Abhängigkeit
 - - - quadratische Näherung durch die 3 markierten Punkte

Beispiel 4: Nicht mögliche (z. B. negative) Werte

Bild 11-14 zeigt die Abhängigkeit einer Zielgröße von einem Faktor, wie sie typischerweise auftritt, wenn aufgrund der Definition der Zielgröße keine negativen Werte möglich sind (vgl. Barthöhe und Rautiefe in Absatz 11.1.5). Der Wert 0 kann dann über einen weiten Bereich des Faktors auftreten. Passt man ein quadratisches Modell an, so errechnet man mit diesem Modell unmögliche, negative Werte für die Zielgröße.

Lösungsmöglichkeit

Bei der Interpretation wird berücksichtigt, dass sich diese unmöglichen Werte rein rechnerisch ergeben und nicht echt sind. Strebt man für die Zielgröße den Wert 0 an und sind technisch keine negativen Werte möglich, so kann man den gesamten Bereich mit negativen Werten als optimalen Bereich interpretieren. Die Faktorstufenkombination, für die man das negative Minimum errechnet, bietet dann (im Rahmen der Genauigkeit des Modells) die größte Fertigungssicherheit.

11.3.2 Allgemeine Lösungsansätze

Neben den bisher vorgeschlagenen Lösungsmöglichkeiten für spezielle Probleme gibt es allgemein folgende Ansätze:

- Wenn der relative Unterschied zwischen dem größten und kleinsten y-Wert mehr als ca. 50 % beträgt, bringt eine Transformation manchmal eine Verbesserung der Modellanpassung, vgl. quasilineare Regression (Absatz 10.1.6). Mit der Box-Cox-Darstellung kann eine geeignete Transformation ausgewählt werden (Abschnitt 6.4). Durch die Transformation wird auch die Verteilung der Daten verändert. Je nach Anwendung kann dies erwünscht oder unerwünscht sein.
- Höhere Potenzen (kubische oder sogar quartische Polynome) erlauben mehr Freiheit bei der Modellanpassung. Dazu werden mehr als drei Stufen für die Faktoren benötigt. Mit zentral zusammengesetzten Plänen können einzelne Zusatzterme angepasst werden, aber im Normalfall werden optimale Pläne speziell für die beabsichtigten Modelle aufgestellt. Achtung, höhere Polynom-Modelle neigen zum Schwingen und können leicht zu einer Überanpassung führen (vgl. Absatz 17.3.1). In einem kubischen Modell nimmt die Anzahl Parameter grob wie $(\text{Anzahl Faktoren})^3$ zu und wird daher schnell unrealistisch groß. Daher sollten höhere Polynommodelle nur in Ausnahmefällen eingesetzt werden. Eine grafische Überprüfung des angepassten Modells auf Plausibilität ist unerlässlich.

- In Kapitel 17 werden Modellansätze vorgestellt, bei denen keine Normalverteilung vorausgesetzt werden muss und die auch komplizierte Wirkzusammenhänge beschreiben können. Die meisten dieser Ansätze erfordern jedoch spezielle Software und viel Erfahrung.
- In Kapitel 18 werden Vorgehensweisen beschrieben, mit denen man sich zunächst schrittweise dem Optimum nähert. In der Nähe des Optimums bieten quadratische Modelle dann eine ausreichend gute Beschreibung des Wirkzusammenhangs.

■ 11.4 Einsatzempfehlungen

Versuchspläne für nichtlineare Zusammenhänge werden eingesetzt, um die quantitative Abhängigkeit der Zielgrößen von wenigen Faktoren im Detail zu bestimmen. In den meisten Anwendungen liegt die Anzahl der Faktoren zwischen drei und fünf, nur in Ausnahmefällen wird man mehr als sechs Faktoren gleichzeitig untersuchen.

Als Versuchspläne sind m.E. für den Normalfall zentral zusammengesetzte Pläne am besten geeignet. Bei bis zu vier Faktoren verwendet man als Würfel einen vollständigen faktoriellen Plan, bei fünf und mehr Faktoren genügt meist ein fraktioneller Plan mit Auflösung V.

Soweit möglich wählt man die fünf Stufenwerte für jeden Faktor so, dass sich ein orthogonaler Plan ergibt (Absatz 11.1.1). Stellt bei einem Faktor ein Stufenwert eine technische Grenze dar oder kann man aus anderen Gründen nur drei Stufenwerte einsetzen, so verwendet man für diesen Faktor $\alpha=1$ (flächenzentriert, Absatz 11.1.4).

Die Faktoren und ihre Stufen sollten so definiert werden, dass sie unabhängig voneinander verändert werden können und im untersuchten Bereich auch sinnvolle Ergebnisse zu erwarten sind. Sind trotzdem einzelne Faktorstufenkombinationen nicht realisierbar, so kann man Ersatzpunkte z.B. mit dem Kriterium der D-Optimalität auswählen. Ist bereits vor dem Versuch bekannt, dass Faktoren sich gegenseitig begrenzen, so eignen sich optimale Pläne (Absatz 11.2.4 und Beispiel 20.1).

Bei vielen Anwendungen wird zur Begrenzung des Aufwands zunächst nur das Zentrum mehrmals realisiert, Würfel und Sternpunkte dagegen nur einmal. Bei großer Zufallsstreuung müssen jedoch alle Faktorstufenkombinationen mehrmals realisiert werden. Mit (7.15) kann man die notwendige Anzahl n der Realisierungen der Würelpunkte berechnen. Es wird dann empfohlen, den gesamten zentral zusammengesetzten Plan n -mal zu realisieren (das Zentrum also $n \cdot n_0$ -mal).

Sollen mehrere Zielgrößen gleichzeitig optimiert werden (der Normalfall), so werden diese zunächst getrennt erfasst und ausgewertet. Durch eine systematische Vorgehensweise wie in Absatz 11.1.5 sucht man anschließend Kompromisse. In komplizierteren Situationen hilft die Vorgehensweise in Abschnitt 15.3.

Literatur

- [1] *Box, G. E. P./Draper, N. R.*: Response Surfaces, Mixtures, and Ridge Analyses. John Wiley, New York, 2. Auflage 2007
- [2] *Myers, R. H./Montgomery, D. C./Anderson-Cook, C. M.*: Response Surface Methodology. John Wiley, New York, 4. Auflage 2016
- [3] *Petersen, H.*: Grundlagen der Statistik und der statistischen Versuchsplanung. ecomed verlagsgesellschaft, Landsberg/Lech 1991
- [4] *Scheffler, E.*: Statistische Versuchsplanung und Auswertung. DVG, Stuttgart, 3. Auflage 1997
- [5] *Draper, N. R./Lin, D. K. J.*: „Small response surface designs“ in: *Technometrics* 32 (1990), 187 – 194
- [6] *Bandemer, B./Bellmann, A.*: Statistische Versuchsplanung. Teubner Verlag, Wiesbaden 1994
- [7] *Goos, P./Jones, B.*: Optimal Design of Experiments – A Case Study Approach. John Wiley, Chichester 2011

12

Varianzanalyse

In den Kapiteln 6–8 wurde die Signifikanz eines Faktors durch Vergleich seines Effekts mit der Breite des Vertrauensbereichs beurteilt. Der Effekt ist die Differenz der Mittelwerte der Versuchsergebnisse bei den beiden Faktorstufen. Dieses Verfahren ist nur bei zweistufigen Faktoren anwendbar.

In Kapitel 10 und 11 wurden mehrstufige Faktoren mit Regression behandelt. Regressionsanalyse ist aber nur bei quantitativen Faktoren direkt einsetzbar.

Manchmal sollen jedoch mehr als zwei Stufen eines qualitativen Faktors miteinander verglichen werden, z.B. wenn

- drei Legierungen bezüglich ihrer Festigkeit verglichen werden sollen oder
- in der Fertigung vier Anlagen parallel eingesetzt werden.

Bei drei Stufen 1, 2, 3 gibt es drei Paare (12, 13 und 23) von je zwei Stufen, zwischen denen die Differenz berechnet werden kann. Bei vier Stufen 1, 2, 3, 4 gibt es sogar sechs Paare (12, 13, 14, 23, 24 und 34). Vergleicht man trotzdem die Stufen paarweise, so erhält man aufgrund der großen Anzahl möglicher Paare mit erhöhter Wahrscheinlichkeit zufällig „signifikante“ Unterschiede. Bei der Varianzanalyse können dagegen beliebig viele Stufen gleichzeitig miteinander verglichen werden, ohne dass das Ergebnis durch die Anzahl möglicher Paare verfälscht wird.

Bei zwei Stufen sind Mittelwertvergleich und Varianzanalyse inhaltlich identische Methoden (auch wenn der Rechengang etwas anders aussieht). Sie führen zu identischen Ergebnissen. In den Kapiteln 6–8 wurden Mittelwertvergleich und Effekte bevorzugt, weil sie anschaulicher sind.

Im Folgenden wird die Varianzanalyse kurz beschrieben, weil sie auch auf qualitative Faktoren mit mehr als zwei Stufen anwendbar ist und in Literatur und Software verbreitet ist. Ab Abschnitt 12.3 werden einige konzeptionell anspruchsvolle Themen kurz erläutert, nämlich

- der Unterschied zwischen Faktoren mit festen und zufälligen Stufen und
- die statistisch saubere Auswertung von Versuchen, bei denen keine volle Randomisierung angewendet wurde.

Nur die prinzipielle Vorgehensweise und die Bedeutung der Ergebnisse werden erklärt. Auf Einzelheiten der Berechnung wird bewusst verzichtet. Für eine detaillierte Beschreibung aller Aspekte der Versuchsplanung für die Varianzanalyse und der Auswertung mit der Varianzanalyse wird auf die Literatur verwiesen [1–5]. Einer weiterer Ansatz, der es erlaubt, mehrere Blockfaktoren zu berücksichtigen, wird in Kapitel 13 beschrieben.

■ 12.1 Einfache balancierte Varianzanalyse

„Einfache“ Varianzanalyse bedeutet, dass nur **ein** Faktor A mit a Stufen betrachtet wird.

„Balanciert“ bedeutet, dass die Anzahl der Einzelversuche auf jeder Stufe gleich ist, $n_1 = n_2 = \dots = n_a = n$. Für diesen Fall ist bei vorgegebenem Versuchsumfang $N = n \cdot a$ die Unterscheidung zwischen den Stufen am besten und die Auswertung am einfachsten.



Beispiel für die Problemstellung

Zur Abscheidung einer SiO_2 -Schicht werden vier Anlagen parallel eingesetzt. Es soll untersucht werden, ob sich die Dicke von unter nominell gleichen Bedingungen abgeschiedenen Schichten unterscheidet. Dazu werden aus jeder Anlage zufällig drei Scheiben entnommen. Tabelle 12.1 zeigt die gemessenen Schichtdicken und die daraus berechneten Mittelwerte und Varianzen. Bild 12-1 zeigt die Einzelwerte und die Mittelwerte grafisch.

| | Anlage 1 | Anlage 2 | Anlage 3 | Anlage 4 |
|------------------------|----------|----------|----------|----------|
| 520 | 590 | 510 | 550 | |
| 560 | 570 | 540 | 500 | |
| 510 | 598 | 525 | 495 | |
| Mittelwert \bar{y}_i | 530 | 586 | 525 | 515 |
| Varianz s_i^2 | 700 | 208 | 225 | 925 |

TABELLE 12.1

Dicke [nm] von unter nominell gleichen Bedingungen auf 4 Anlagen abgeschiedenen Schichten

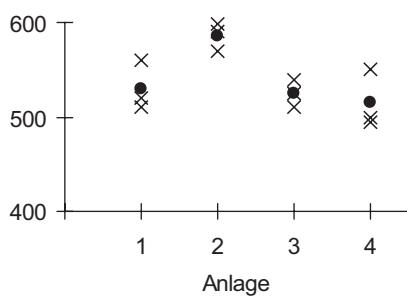


BILD 12-1

Einzelwerte und Mittelwerte der Schichtdicken von den vier untersuchten Anlagen

- x Einzelwerte
- Mittelwerte

Grundidee und Vorgehensweise

Bei der Varianzanalyse wird vorausgesetzt, dass die Einzelwerte innerhalb jeder Stufe normalverteilt sind, und zwar mit dem gleichen σ für alle Stufen. Der Mittelwert der Varianzen innerhalb der Stufen ist dann ein Schätzwert für die Varianz der Einzelwerte

$$s^2 = \frac{1}{a} \cdot \sum_{i=1}^a s_i^2. \quad (12.1)$$

Wenn es keine wahren Unterschiede zwischen den Stufen gibt,¹ ist die Varianz der Stufenmittelwerte $1/n$ der Varianz der Einzelwerte – vgl. (6.7). Die Prüfgröße

¹ Nullhypothese: Alle Mittelwerte sind gleich.

$$F_{\text{Prüf}} = \frac{n \cdot s_{\bar{y}}^2}{s^2} \quad (12.2)$$

unterscheidet sich dann nur zufällig von 1. Die Verteilung der Prüfgröße wird durch die sogenannte F-Verteilung beschrieben. Die Form dieser Verteilung kann berechnet werden und hängt von der Anzahl der Stufen und der Anzahl der Einzelwerte ab – oder, genauer, vom Freiheitsgrad f_1 zur Berechnung der Varianz der Mittelwerte

$$f_1 = a - 1 \quad (12.3)$$

und vom Freiheitsgrad f_2 der Varianz der Einzelwerte

$$f_2 = a \cdot (n - 1) = N - a. \quad (12.4)$$

Je größer die Freiheitsgrade f_1 und f_2 , desto schmäler ist die F-Verteilung. Bild 12-2 zeigt die F-Verteilung für das obige Beispiel. Wenn es keine wahren Unterschiede zwischen den Stufen (hier zwischen den Anlagen) gibt, liegt die Prüfgröße meist nahe bei 1. Werte der Prüfgröße, die größer sind als ein kritischer Wert, treten nur mit geringer Wahrscheinlichkeit α auf (in Bild 12-2 mit 5 %).

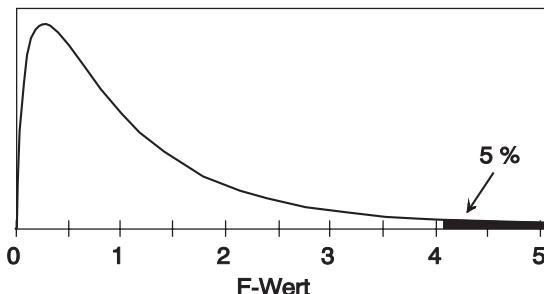


BILD 12-2

F-Verteilung für Freiheitsgrade $f_1 = 3$ und $f_2 = 8$:
95 % der Gesamtfläche liegt unterhalb 4,07, 5 % oberhalb.
D.h. wenn es keine Unterschiede zwischen den Stufen gibt, treten nur mit einer Wahrscheinlichkeit von 5 % F-Werte > 4,07 auf.

Wahre Unterschiede zwischen den Stufen vergrößern die Varianz der Mittelwerte und somit auch die Prüfgröße. Große Werte der Prüfgröße sind daher ein Hinweis auf Unterschiede zwischen den Stufen. Die Tabellen 12.2 bis 12.4 geben kritische Werte von F für ausgewählte Freiheitsgrade. Größere Werte treten nur mit einer Wahrscheinlichkeit von 5 %, 1 % bzw. 0,1 % zufällig auf. Je kleiner die Fläche oberhalb des kritischen Wertes sein soll, desto größer ist der F-Wert.

Das Ergebnis wird wie in den früheren Kapiteln bewertet:

| | |
|---|----------------|
| $F_{\text{Prüf}} \leq F_{f_1; f_2; 0,95}$ | Bewertung – |
| $F_{f_1; f_2; 0,95} < F_{\text{Prüf}} \leq F_{f_1; f_2; 0,99}$ | Bewertung * |
| $F_{f_1; f_2; 0,99} < F_{\text{Prüf}} \leq F_{f_1; f_2; 0,999}$ | Bewertung ** |
| $F_{f_1; f_2; 0,999} < F_{\text{Prüf}}$ | Bewertung ***. |

(12.5)

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 50 | ∞ |
|----------------------|------|------|------|------|------|------|------|------|----------|
| 1 | 161 | 199 | 216 | 225 | 230 | 242 | 248 | 252 | 254 |
| 2 | 18,5 | 19,0 | 19,2 | 19,2 | 19,3 | 19,4 | 19,4 | 19,5 | 19,5 |
| 3 | 10,1 | 9,55 | 9,28 | 9,12 | 9,01 | 8,79 | 8,66 | 8,58 | 8,53 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 5,96 | 5,80 | 5,70 | 5,63 |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,74 | 4,56 | 4,44 | 4,37 |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,06 | 3,87 | 3,75 | 3,67 |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,35 | 3,15 | 3,02 | 2,93 |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 2,98 | 2,77 | 2,64 | 2,54 |
| 20 | 4,35 | 3,49 | 3,10 | 2,87 | 2,71 | 2,35 | 2,12 | 1,97 | 1,84 |
| 50 | 4,03 | 3,18 | 2,79 | 2,56 | 2,40 | 2,03 | 1,78 | 1,60 | 1,44 |
| ∞ | 3,84 | 3,00 | 2,60 | 2,37 | 2,21 | 1,83 | 1,57 | 1,35 | 1,00 |

TABELLE 12.2

Kritische Werte

$F_{f_1; f_2; 0,95}$

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 50 | ∞ |
|----------------------|------|------|------|------|------|------|------|------|----------|
| 1 | 4052 | 4999 | 5404 | 5624 | 5764 | 6056 | 6209 | 6302 | 6366 |
| 2 | 98,5 | 99,0 | 99,2 | 99,3 | 99,3 | 99,4 | 99,4 | 99,5 | 99,5 |
| 3 | 34,1 | 30,8 | 29,5 | 28,7 | 28,2 | 27,2 | 26,7 | 26,4 | 26,1 |
| 4 | 21,2 | 18,0 | 16,7 | 16,0 | 15,5 | 14,5 | 14,0 | 13,7 | 13,5 |
| 5 | 16,3 | 13,3 | 12,1 | 11,4 | 11,0 | 10,1 | 9,55 | 9,24 | 9,02 |
| 6 | 13,7 | 10,9 | 9,78 | 9,15 | 8,75 | 7,87 | 7,40 | 7,09 | 6,88 |
| 8 | 11,3 | 8,65 | 7,59 | 7,01 | 6,63 | 5,81 | 5,36 | 5,07 | 4,86 |
| 10 | 10,0 | 7,56 | 6,55 | 5,99 | 5,64 | 4,85 | 4,41 | 4,12 | 3,91 |
| 20 | 8,10 | 5,85 | 4,94 | 4,43 | 4,10 | 3,37 | 2,94 | 2,64 | 2,42 |
| 50 | 7,17 | 5,06 | 4,20 | 3,72 | 3,41 | 2,70 | 2,27 | 1,95 | 1,68 |
| ∞ | 6,63 | 4,61 | 3,78 | 3,32 | 3,02 | 2,32 | 1,88 | 1,52 | 1,00 |

TABELLE 12.3

Kritische Werte

$F_{f_1; f_2; 0,99}$

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 50 | ∞ |
|----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 1 | 4×10^5 | 5×10^5 | 5×10^5 | 6×10^5 |
| 2 | 998 | 999 | 999 | 999 | 999 | 999 | 999 | 999 | 999 |
| 3 | 167 | 148 | 141 | 137 | 135 | 129 | 126 | 125 | 123 |
| 4 | 74,1 | 61,2 | 56,2 | 53,4 | 51,7 | 48,1 | 46,1 | 44,9 | 44,0 |
| 5 | 47,2 | 37,1 | 33,2 | 31,1 | 29,8 | 26,9 | 25,4 | 24,4 | 23,8 |
| 6 | 35,5 | 27,0 | 23,7 | 21,9 | 20,8 | 18,4 | 17,1 | 16,3 | 15,7 |
| 8 | 25,4 | 18,5 | 15,8 | 14,4 | 13,5 | 11,5 | 10,5 | 9,80 | 9,33 |
| 10 | 21,0 | 14,9 | 12,6 | 11,3 | 10,5 | 8,75 | 7,80 | 7,19 | 6,76 |
| 20 | 14,8 | 9,95 | 8,10 | 7,10 | 6,46 | 5,08 | 4,29 | 3,77 | 3,38 |
| 50 | 12,2 | 7,96 | 6,34 | 5,46 | 4,90 | 3,67 | 2,95 | 2,44 | 2,03 |
| ∞ | 10,8 | 6,91 | 5,42 | 4,62 | 4,10 | 2,96 | 2,27 | 1,73 | 1,00 |

TABELLE 12.4

Kritische Werte

$F_{f_1; f_2; 0,999}$

Auch hier bedeutet die Bewertung „–“ nur, dass die Daten konsistent damit sind, dass kein Unterschied zwischen den Stufen besteht. Nur über einen ausreichend großen Versuchsumfang N kann man gewährleisten, dass ein tatsächlich vorhandener Unterschied auch mit hoher Wahrscheinlichkeit entdeckt wird. Soll ein Unterschied $\Delta\mu$ mit einer Wahrscheinlichkeit von 90% zu einer Bewertung ** oder *** führen, benötigt man einen Versuchsumfang von ca.

$$N = a \cdot n = a \cdot 30 \cdot \left(\frac{\sigma}{\Delta\mu} \right)^2. \quad (12.6)$$



Anwendung auf das Beispiel

Der Mittelwert der Varianzen der Einzelwerte innerhalb der Anlagen ist ein Schätzwert für die Varianz der Einzelwerte

$$s^2 = \frac{1}{4} \cdot (700 + 208 + 225 + 925) = 514,5.$$

Die Varianz der vier Mittelwerte (530, 586, 525 und 515) beträgt

$$s_{\bar{y}}^2 = 1020,7.$$

Daraus erhält man gemäß (12.2) die Prüfgröße

$$F_{\text{Prüf}} = \frac{3 \cdot 1020,7}{514,5} = 5,95.$$

Gemäß (12.3) und (12.4) erhält man für die Freiheitsgrade

$$f_1 = 4 - 1 = 3 \quad f_2 = 4 \cdot (3 - 1) = 8.$$

Aus Tabelle 12.2 bzw. 12.3 erhält man als kritische Werte

$$F_{3;8;0,95} = 4,07 \text{ (vgl. Bild 12-2)}$$

$$F_{3;8;0,99} = 7,59$$

Der Vergleich der Prüfgröße mit den kritischen Werten ergibt die Bewertung *, d. h. es gibt zwar Hinweise auf einen Unterschied zwischen den Anlagen, aber bevor aufwändige Maßnahmen getroffen werden, sollte man mehr Daten sammeln.

Die Varianzanalyse alleine macht keine Aussage darüber, welcher Art die Unterschiede sind. Bild 12-1 legt nahe, dass in diesem Beispiel Scheiben von Anlage 2 eine höhere Schichtdicke aufweisen als von den anderen drei Anlagen.



Aufgabe

Ein Drehautomat fertigt Wellen mit drei parallelen Werkzeugen. Da sämtliche Werkzeuge nominell gleich sind und von demselben Lieferanten stammen, darf angenommen werden, dass sie mit derselben Genauigkeit (Streuung) fertigen. Es ist jedoch möglich, dass sie unterschiedlich justiert sind (unterschiedliche Mittelwerte). Daher werden gezielt aus der Fertigung eines jeden Werkzeugs 10 Wellen entnommen und vermessen. Folgende Tabelle zeigt die Mittelwerte und Standardabweichungen für jedes Werkzeug (in mm). Unterscheiden sich die Mittelwerte der Wellen?

| Werkzeug i | 1 | 2 | 3 |
|--------------------------|-------|-------|-------|
| Mittelwert \bar{y}_i | 5,001 | 5,007 | 5,038 |
| Standardabweichung s_i | 0,011 | 0,020 | 0,022 |

Lösung

Schätzwert für die Varianz der Einzelwerte

$$s^2 = \frac{1}{3} \cdot (0,011^2 + 0,020^2 + 0,022^2) = 0,000335$$

Varianz der drei Mittelwerte (5,001, 5,007 und 5,038)

$$s_{\bar{y}}^2 = 0,000394$$

$$F_{\text{Prüf}} = \frac{10 \cdot 0,000394}{0,000335} = 11,8$$

$$\text{Freiheitsgrade } f_1 = 3 - 1 = 2 \quad f_2 = 3 \cdot (10 - 1) = 27$$

Der Vergleich mit Tabelle 12.4 ergibt die Bewertung ***, d. h. der Unterschied zwischen den Werkzeugen ist hochsignifikant.

Ergänzung

Varianzanalyse kann man auch als eine Zerlegung der Summe der quadrierten Abweichungen (S. d. q. A.) der Einzelwerte vom Gesamtmittelwert in einen Anteil Q_A und einen Rest Q_R auffassen. Wie bei der Zerlegung (10.10) kann Q_A durch die Stufen des Faktors A erklärt werden, Q_R wird zur Schätzung der Zufallsstreuung verwendet. Das Ergebnis wird dann in einer Zerlegungstafel zusammengefasst. Tabelle 12.5 zeigt die Zerlegungstafel in allgemeiner Form mit den üblichen Bezeichnungen und formelmäßigen Zusammenhängen. Tabelle 12.6 zeigt die Zerlegungstafel dann angewendet auf das obige Beispiel.

TABELLE 12.5 Streuungszerlegungstafel der Varianzanalyse, Q_A und Q_R sind Hilfsgrößen – in (12.2) wird die Prüfgröße direkt berechnet.

| Streuungsursache | S. d. q. A. | Freiheitsgrad | Varianz | $F_{\text{Prüf}}$ |
|------------------|-----------------|---------------------------------|---|---|
| Faktor A | Q_A | $f_A = a - 1 (= f_1)$ | $s_A^2 = \frac{Q_A}{f_A} (= n \cdot s_{\bar{y}}^2)$ | $F_{\text{Prüf}} = \frac{s_A^2}{s_R^2}$ |
| Rest | Q_R | $f_R = a \cdot (n - 1) (- f_2)$ | $s_R^2 = \frac{Q_R}{f_R} (= s^2)$ | |
| Gesamt | $Q = Q_A + Q_R$ | $f = f_A + f_R = a \cdot n - 1$ | | (12.7) |

TABELLE 12.6 Streuungszerlegungstafel der Varianzanalyse für das obige Beispiel

| Streuungsursache | S. d. q. A. | Freiheitsgrad | Varianz | $F_{\text{Prüf}}$ | p-Wert |
|------------------|-------------|---------------|---------|-------------------|--------|
| Faktor A | 9186 | 3 | 3062 | 5,95 | 0,02 |
| Rest | 4116 | 8 | 514,5 | | |
| Gesamt | 13302 | 11 | | | |

Die Analyse mit Hilfe der Streuungszerlegungstafel ist offensichtlich äquivalent zur hier vorgeschlagenen Analyse (aber rechnerisch aufwändiger).

Die Spalte „p-Wert“ in Tabelle 12.6 gibt als Zusatzinformation die Wahrscheinlichkeit dafür, dass der beobachtete F-Wert oder ein noch größerer Wert zufällig auftritt, obwohl der Faktor A das Ergebnis nicht beeinflusst. In Bild 12-2 bedeutet dies: Nur 2% der Fläche liegen oberhalb von $F = 5,95$. Software gibt diesen Wert als Entscheidungshilfe an, er ersetzt die Tabellen 12.2 bis 12.4. Man erhält folgende Bewertung:

| | |
|-----------------------------------|----------------|
| $p\text{-Wert} \geq 0,05$ | Bewertung – |
| $0,05 > p\text{-Wert} \geq 0,01$ | Bewertung * |
| $0,01 > p\text{-Wert} \geq 0,001$ | Bewertung ** |
| $0,001 > p\text{-Wert}$ | Bewertung ***. |

In diesem Beispiel sind folgende Aussagen gleichwertig:

$$\begin{array}{lclclcl} F_{3;8;0,95} = 4,07 & < & F_{\text{Prüf}} = 5,95 & \leq & F_{3;8;0,99} = 7,59 & : \text{Bewertung *} \\ 0,05 & > & p\text{-Wert} = 0,02 & \geq & 0,01 & : \text{Bewertung *} \end{array}$$

■ 12.2 Mehrfache Varianzanalyse

Wie der Mittelwertvergleich in Kapitel 7, so kann auch die Varianzanalyse auf mehr als einen Faktor erweitert werden. Man unterscheidet dann zwischen den Einflüssen der Faktoren und der Wechselwirkungen.

Dabei bedeuten z.B.:

Faktor A signifikant: Es gibt Unterschiede zwischen den a Stufen des Faktors A, sie sind nicht alle gleich.

Faktor B signifikant: Es gibt Unterschiede zwischen den b Stufen des Faktors B.

WW AB signifikant: Der Einfluss des Faktors A hängt von der Stufe von B ab (und umgekehrt).

Die Varianzanalyse macht keine Aussage über die Art der Unterschiede. Zur Interpretation betrachtet man am besten die Mittelwerte der Stufen von A bzw. B für die Faktoren, bzw. die $a \cdot b$ Mittelwerte aller Kombinationen der Stufen von A und von B für die WW.

Die Streuungszerlegung ist rechnerisch aufwändig. Sie wird daher normalerweise mit Softwareunterstützung durchgeführt und hier bewusst nicht behandelt. Einzelheiten können in Standardtextbüchern nachgelesen werden, z.B. [1]. Hier soll nur an einem Beispiel der Einfluss der Faktoren und der Wechselwirkungen grafisch dargestellt werden.



Beispiel

Der Einfluss von vier verschiedenen Zusatzstoffen 1, 2, 3 und 4 (Nummerierung willkürlich) auf die Rate einer bestimmten Reaktion bei den Temperaturen 60, 70 und 80°C soll untersucht werden. Jede der 12 Faktorstufenkombinationen wird dreimal realisiert. Tabelle 12.7 zeigt die Einzelergebnisse, zusammen mit den Mittelwerten.

Tabelle 12.8 zeigt das Ergebnis einer zweifachen Varianzanalyse mit diesen Daten. Die S. d. q. A. der Daten wird zerlegt in einen Anteil von jedem Faktor (Unterschiede zwischen den Zeilenmittelwerten 19,4 – 24,1 – 30,9 für die drei Temperaturen bzw. den Spaltenmittelwerten 23,6 – 21,6 – 25,2 – 28,9 für die vier Zusatzstoffe, den sog. Randmittelwerten), einen Anteil von der Wechselwirkung und die Reststreuung.

Beide Faktoren und die WW sind signifikant. Es gibt also Unterschiede zwischen den Faktorstufenkombinationen – Varianzanalyse macht aber keine Aussage über die Art der Unterschiede. Diese erkennt man am besten an einer grafischen Darstellung. Bild 12-3 zeigt die Mittelwerte für alle 12 Faktorstufenkombinationen von Tabelle 12.7. Es stellt damit die Einflüsse der Faktoren und der Wechselwirkungen gleichzeitig dar. Die Wechselwirkung äußert sich in der unterschiedlichen Temperaturabhängigkeit der Reaktionsrate für die verschiedenen Zusatzstoffe. Wie in Kapitel 7 ist die Wechselwirkung eine Abweichung von der Parallelität.

TABELLE 12.7 Drei Einzelergebnisse für die Reaktionsrate für jede Kombination von Zusatzstoff (Spalten) und Temperatur (Zeilen) und die Mittelwerte dieser drei Ergebnisse.

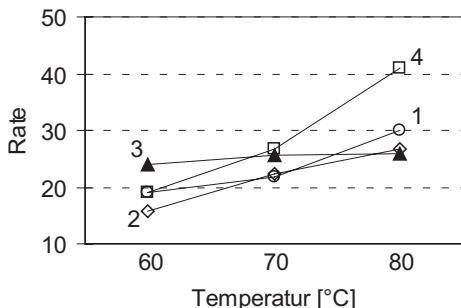
Zeilenmittel: Mittelwert über alle 12 Einzelergebnisse bei einer Temperatur

Spaltenmittel: Mittelwert über alle 9 Einzelergebnisse bei einem Zusatzstoff

| Zusatzstoff Temperatur | 1 | 2 | 3 | 4 | Zeilenmittel |
|---------------------------|------|------|------|------|--------------|
| 60°C | 18 | 11 | 22 | 19 | |
| | 21 | 17 | 23 | 17 | |
| | 18 | 19 | 27 | 21 | |
| Mittelwert | 19,0 | 15,7 | 24,0 | 19,0 | 19,4 |
| 70°C | 14 | 24 | 30 | 20 | |
| | 21 | 25 | 23 | 30 | |
| | 30 | 18 | 24 | 30 | |
| Mittelwert | 21,7 | 22,3 | 25,7 | 26,7 | 24,1 |
| 80°C | 30 | 32 | 26 | 43 | |
| | 27 | 28 | 27 | 41 | |
| | 33 | 20 | 25 | 39 | |
| Mittelwert | 30,0 | 26,7 | 26,0 | 41,0 | 30,9 |
| Spaltenmittel | 23,6 | 21,6 | 25,2 | 28,9 | 24,8 |

TABELLE 12.8 Streuungszerlegungstafel der zweifachen Varianzanalyse für das obige Beispiel

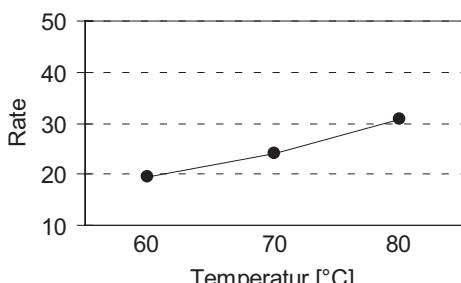
| Streuungsursache | S. d. q. A. | Freiheitsgrad | Varianz | F _{Prüf} | p-Wert | Bewertung |
|------------------|-------------|---------------|---------|-------------------|--------|-----------|
| A: Temperatur | 802,9 | 2 | 401,4 | 23,1 | 0,000 | *** |
| B: Zusatzstoff | 260,7 | 3 | 86,9 | 5,0 | 0,008 | ** |
| WW AB | 334,0 | 6 | 55,7 | 3,2 | 0,019 | * |
| Rest | 418,0 | 24 | 17,4 | | | |
| Gesamt | 1815,6 | 35 | | | | |

**BILD 12-3**

Grafische Darstellung der Mittelwerte der 12 Faktorstufenkombinationen:
Die vier Zusatzstoffe sind durch verschiedene Symbole gekennzeichnet und gemäß Tabelle 12.7 nummeriert.
Wechselwirkung: Die Temperaturabhängigkeit der Rate ist bei den vier Zusatzstoffen unterschiedlich.

Welche Konsequenzen aus dem Ergebnis von Bild 12-3 gezogen werden, hängt von der Zielsetzung ab. Soll ein Zusatzstoff gefunden werden, der die Temperaturabhängigkeit minimiert, so wählt man Zusatzstoff 3 aus (robust gegen Veränderungen der Temperatur – der Unterschied in der Temperaturabhängigkeit wird von der Wechselwirkung beschrieben). Ist das Ziel dagegen eine möglichst hohe Reaktionsrate, so wählt man Zusatzstoff 4 und Temperatur 80°C.

Zur Verdeutlichung des Einflusses der Faktoren zeigt Bild 12-4 die Zeilenmittelwerte für die drei Temperaturen, Bild 12-5 die Spaltenmittelwerte für die vier Zusatzstoffe (aus Tabelle 12.7). Da die Wechselwirkung signifikant ist, enthalten diese Darstellungen jedoch nicht die volle Information der Versuchsergebnisse. Aus ihnen kann man nicht erkennen, dass die Temperaturabhängigkeit bei Zusatzstoff 3 am kleinsten ist.

**BILD 12-4** Einfluss der Temperatur:

Mit zunehmender Temperatur nimmt die Reaktionsrate zu (gemittelt über alle Zusatzstoffe).

**BILD 12-5**

Einfluss des Zusatzstoffes:
Zusatzstoff 2 ergibt die niedrigste, Zusatzstoff 4 die höchste Reaktionsrate (gemittelt über alle Temperaturen).

■ 12.3 Feste und zufällige Effekte

Bei den bisher behandelten Faktoren waren wir genau an den untersuchten Stufenwerten (z. B. an den vier Zusatzstoffen im Beispiel von Abschnitt 12.2) oder an der quantitativen Abhängigkeit von den Stufenwerten im untersuchten Bereich (z. B. von der Temperatur im Beispiel von Abschnitt 12.2 oder von Geschwindigkeit, Druck, Abstand, Fokuslage und Laserleistung im Beispiel von Absatz 11.1.5) interessiert. In diesen Fällen spricht man von **festen Effekten**. Alle bisher behandelten Auswerteverfahren gelten für feste Effekte. Im Beispiel von Abschnitt 12.2 ist jedoch auch eine veränderte Fragestellung denkbar: Wenn die vier „Zusatzstoffe“ nicht vier verschiedene Marken oder Typen sind, von denen wir den besten auswählen wollen, sondern vier verschiedene Chargen desselben Zusatzstoffes, sind wir nicht an der besten Charge interessiert – in der Fertigung werden immer wieder andere Chargen verwendet und die vier Chargen wurden nur zufällig für die Untersuchung ausgewählt. Dann interessiert vielmehr, wie groß der Beitrag der Chargenunterschiede zur Gesamtstreuung der Reaktionsrate ist (Varianzkomponente) und ob die Zunahme der Reaktionsrate mit der Temperatur auch bei Berücksichtigung der Chargenunterschiede signifikant ist. Die Charge ist dann ein Faktor mit **zufälligen Effekten** – man spricht auch von einem zufälligen Faktor oder genauer einem Faktor mit zufälligen Stufenwerten. Bei der Auswertung wird ein **gemischtes Modell** (mixed model) verwendet, das den Faktor Charge mit zufälligen Stufen und den Faktor Temperatur mit festen Stufen enthält.



Verändertes Beispiel (aber mit denselben Zahlenwerten wie in Abschnitt 12.2)

Es soll untersucht werden, ob die Rate einer bestimmten Reaktion sich bei den Temperaturen 60, 70 und 80°C signifikant unterscheidet. Es wird befürchtet, dass die Materialcharge des Zusatzstoffes die Temperaturabhängigkeit beeinflusst, daher werden für den Versuch die Chargen 1, 2, 3 und 4 (Nummerierung willkürlich) zufällig ausgewählt. Jede der 12 Kombinationen von Charge und Temperatur werden dreimal realisiert. Tabelle 12.7 zeigt die Einzelergebnisse, Tabelle 12.8 die Streuungszerlegung für feste Effekte, Tabelle 12.9 die modifizierte Zerlegung für ein gemischtes Modell.

TABELLE 12.9 Modifizierte Streuungszerlegungstafel, wenn der Zusatzstoff ein Faktor mit zufälligen Effekten ist

| Streuungsursache | S. d. q. A. | Freiheitsgrad | Varianz | $F_{\text{Prüf}}$ | p-Wert | Bewertung |
|------------------|-------------|---------------|---------|-------------------|--------|-----------|
| A: Temperatur | 802,9 | 2 | 401,4 | 7,2 | 0,025 | * |
| B: Zusatzstoff | 260,7 | 3 | 86,9 | 1,6 | 0,294 | – |
| WW AB | 334,0 | 6 | 55,7 | | | |

Da der Einfluss der Temperatur von Charge zu Charge unterschiedlich sein kann und die Chargen zufällig ausgewählt wurden, können für die Beurteilung des Effekts der Temperatur nur die Mittelwerte der 12 untersuchten Kombinationen herangezogen werden. Was in Tabelle 12.8 die Wechselwirkung zwischen Temperatur und Zusatzstoff war, wird nun in Tabelle 12.9 zur Schätzung der Zufallsstreuung von Charge zu Charge.

Da nur die Wechselwirkung den Einfluss der Charge auf die Temperaturabhängigkeit erfasst, kann nur sie als Zufallsstreuung verwendet werden – die Streuung innerhalb der Gruppen beinhaltet den

Chargeneinfluss nicht und unterschätzt die Zufallsstreuung daher. F-Werte, Freiheitsgrade und Signifikanz werden ausgewertet wie für einen Versuchsplan mit $n = 1$. Eine Auswertung mit den Mittelwerten alleine liefert hierfür dieselben Ergebnisse (S. d. q. A. und Varianzen werden alle durch 3 geteilt).

Die Charge selbst (der Effekt von B) wirkt hier wie ein Blockfaktor. Die Mittelwerte für die verschiedenen Chargen sind uninteressant (da zufällig), es soll nur vermieden werden, dass Chargenunterschiede die Signifikanz des eigentlich interessierenden Faktors Temperatur verfälschen.

Ergänzend stellt sich manchmal die Frage, wie groß die Varianz σ^2 der Einzelwerte bei vorgegebener Charge und Temperatur ist, sowie die Varianz σ_{AB}^2 der Wechselwirkung zwischen Charge und Temperatur und die Varianz σ_B^2 der Charge (für beide wird Normalverteilung über alle Chargen angenommen). Mit den Varianzen aus Tabelle 12.8 und 12.9 erhält man folgende Schätzwerte:

$$\text{Rest } 17,4 \quad \text{schätzt } \sigma^2 \Rightarrow s^2 = 17,4$$

$$\text{WW AB } 55,7 \quad \text{schätzt } \sigma^2 + 3 \cdot \sigma_{AB}^2 \Rightarrow s_{AB}^2 = \frac{55,7 - 17,4}{3} = 12,8$$

$$\text{B } 86,9 \quad \text{schätzt } \sigma^2 + 3 \cdot \sigma_{AB}^2 + 3 \cdot 3 \cdot \sigma_B^2 \Rightarrow s_B^2 = \frac{86,9 - 55,7}{3 \cdot 3} = 3,5$$

Dies ist die Zerlegung in Varianzkomponenten. Sie ist nur für zufällige Effekte sinnvoll.

Hinweise

- Bei der Auswertung von Versuchen, die zufällige Effekte enthalten, ist die Anzahl der Stufenwerte entscheidend, da sie den Freiheitsgrad bestimmt – auch für die Beurteilung der Signifikanz der festen Effekte. Bei Faktoren mit zufälligen Effekten sollten daher möglichst viele Stufenwerte verwendet werden. Die Anzahl der Realisierungen jeder Faktorstufenkombination hat eine vergleichsweise geringere Bedeutung. Eine mehrfache Realisierung ist nur sinnvoll, wenn der damit verbundene Aufwand klein und die Zufallsstreuung innerhalb der Chargen groß ist ($\sigma^2 > \sigma_{AB}^2$).
- Faktoren mit zufälligen Effekten sind für zweistufige Versuchspläne ungeeignet, die kritischen F-Werte sind viel zu groß. Soll die Materialcharge o.ä. in einem zweistufigen Versuchsplan untersucht werden, sollten die Chargen nicht zufällig ausgewählt werden. Es ist wesentlich effizienter, zwei Chargen zu wählen, die bezüglich ihres Einflusses auf das Versuchsergebnis möglichst extrem sind (z. B. an den beiden Toleranzgrenzen). Als Ergebnis erhält man dann, wie stark sich dieser extreme Unterschied im Material auf die Zielgröße auswirkt – vgl. Beispiel zum Gewinderollen in Abschnitt 7.2. Dies ist dann ein fester Effekt: Der schlimmste zu erwartende Effekt der Charge.
- In gewissem Sinn ist auch ein Blockfaktor ein Faktor mit zufälligen Effekten. Bei der Auswertung wird für den Blockfaktor aber unterstellt, dass er sich nur auf den Mittelwert auswirkt und dass er keine Wechselwirkung mit den anderen Faktoren hat. In der Varianzkomponentenzerlegung wird damit unterstellt, dass $\sigma_{AB}^2 = 0$. Daher ist die Reststreuung ein geeigneter Schätzwert für die Zufallsstreuung und die Freiheitsgrade müssen nicht wie im obigen Beispiel modifiziert werden. Der Unterschied ist technisch begründet, nicht statistisch!

■ 12.4 Nicht vollständige Randomisierung

Für die statistische Auswertung der Versuchsergebnisse wurde bisher immer unterstellt, dass die Einzelversuche innerhalb der Blöcke in randomisierter Reihenfolge durchgeführt werden. Jede Abweichung hiervon birgt das Risiko einer Verfälschung – vgl. Absatz 7.1.4.

Um den Versuchsaufwand zu reduzieren, werden Versuche jedoch häufig in nicht randomisierter Reihenfolge durchgeführt. An zwei modifizierten Reihenfolgen für das Beispiel aus Abschnitt 12.2 wird gezeigt, welche Auswirkungen dies auf die statistisch korrekte Auswertung hat.

12.4.1 Alle Realisierungen einer Kombination gemeinsam

In Abschnitt 12.2 wurde bei der Auswertung unterstellt, dass die 36 Einzelversuche (je drei Realisierungen von 12 Kombinationen aus drei Temperaturen und vier Zusatzstoffen, feste Effekte) in randomisierter Reihenfolge durchgeführt wurden. Dazu müssen von Einzelversuch zu Einzelversuch Temperatur und Zusatzstoff geändert werden.

Dies bedingt einen hohen Änderungsaufwand und ein erhöhtes Verwechslungsrisiko. Daher wird immer wieder vorgeschlagen, die Einzelversuche einer Kombination jeweils unmittelbar hintereinander durchzuführen.

So wird jede Kombination nur einmal eingestellt und dann dreimal durchgeführt. Ungenauigkeiten in der Einstellung werden nicht erfasst. Außerdem sind die Umgebungsbedingungen einheitlich. Die Streuung innerhalb der Gruppen ist im Normalfall kleiner als wenn die Kombination zu verschiedenen Zeiten im Versuch immer wieder neu eingestellt wird. Für die saubere statistische Bewertung ist aber die Zufallsstreuung einschließlich der Ungenauigkeiten in der Einstellung und der Schwankung der Umgebungsbedingungen nötig.

Folgende Vorgehensweisen sind verbreitet:

- Augen zu und durch: Der Unterschied in der Zufallsstreuung wird ignoriert, und so wird die Signifikanz der Effekte überschätzt. Der Einsatz der Statistik gibt nur eine trügerische Sicherheit – Statistik wird missbraucht.
- Nur die Mittelwerte der drei Realisierungen werden verwendet und die Auswertung erfolgt wie bei $n = 1$. Dann wird die Wechselwirkung als Zufallsstreuung (miss-)interpretiert. Man erhält das Ergebnis von Abschnitt 12.3, auch wenn der Zusatzstoff feste Effekte hat. Die Wechselwirkung kann nicht mehr bestimmt werden und die Signifikanz der Faktoren wird unterschätzt. Bei mehr als zwei Faktoren können die höheren Wechselwirkungen jedoch wie in Abschnitt 7.3 zur Schätzung der Zufallsstreuung verwendet werden.

12.4.2 Split-Plot Versuche

Split-Plot ist eine Idee, die ursprünglich aus landwirtschaftlichen Versuchen stammt: Ein Stück Land (plot) wird für den Versuch unterteilt. Manche Faktoren werden dann innerhalb des Plot verändert, andere nur zwischen den Plots. Dies ist eine Abweichung von einer vollen Randomisierung. Die zugehörigen Ideen lassen sich auch auf industrielle Versuche übertragen, z.B. wenn die Veränderung eines Faktors wesentlich aufwändiger ist als die Veränderung anderer Faktoren oder wenn bei manchen Versuchsschritten mehrere Teile gemeinsam bearbeitet werden.

Das folgende Beispiel soll nur einen Eindruck von den Möglichkeiten geben.

In Versuchsreihenfolge verändertes Beispiel 12.2

Wie im Beispiel 12.2 soll der Einfluss von vier verschiedenen Zusatzstoffen 1, 2, 3 und 4 (feste Effekte) auf die Rate einer bestimmten Reaktion bei den Temperaturen 60, 70 und 80°C untersucht werden. Jede der 12 Faktorstufenkombinationen wird dreimal realisiert. Tabelle 12.7 zeigt die Einzelergebnisse.

Allerdings wird nun unterstellt, dass die vier Versuche einer Zeile (vier verschiedene Zusatzstoffe, aber eine feste Temperatur) gemeinsam in einem Ofen gemacht wurden. So sind nur noch neun (statt 36) Ofenfahrten erforderlich. Die neun Ofenfahrten werden in randomisierter Reihenfolge durchgeführt, d.h. die Temperatur wechselt von Ofenfahrt zu Ofenfahrt. Tabelle 12.10 zeigt die zugehörige Streusionszerlegung.

TABELLE 12.10 Streusionszerlegungstafel für das in der Versuchsreihenfolge veränderte Beispiel

| Streuungsursache | S. d. q. A. | Freiheitsgrad | Varianz | F _{Prüf} | f ₁ | f ₂ | p-Wert | Bewertung |
|------------------|-------------|---------------|---------|-------------------|----------------|----------------|--------|-----------|
| A: Temperatur | 802,9 | 2 | 401,4 | 30,1 | 2 | 6 | 0,0007 | *** |
| B: Zusatzstoff | 260,7 | 3 | 86,9 | 4,6 | 3 | 18 | 0,014 | * |
| WW AB | 334,0 | 6 | 55,7 | 3,0 | 6 | 18 | 0,034 | * |
| Ofenfahrt [A] | 80,0 | 6 | 13,3 | 0,7 | 6 | 18 | 0,646 | - |
| Rest | 338,0 | 18 | 18,8 | | | | | |

Im Vergleich zu Tabelle 12.8 wird die Reststreuung weiter zerlegt in einen Beitrag von den Ofenfahrten (neun Fahrten bei drei Temperaturen, Freiheitsgrad daher $9 - 3 = 6$) und einen Rest mit Freiheitsgrad $24 - 6 = 18$.

Um die Signifikanz des Faktors Temperatur zu beurteilen, werden nur die neun Ofenfahrten berücksichtigt (whole plot), deren Varianz zerlegt wird in einen Teil, der durch die Temperatur erklärt wird (mit Freiheitsgrad 2) und einen Rest mit Freiheitsgrad 6. Der p-Wert der Prüfgröße $401,4 / 13,3 = 30,1$ ergibt sich aus den Freiheitsgraden $f_1 = 2$ und $f_2 = 6$.

Der Zusatzstoff (und damit auch die Wechselwirkung) ist ein split-plot Faktor, da er innerhalb einer Ofenfahrt verändert wird. Die zugehörige Reststreuung ist, was von der Reststreuung in Tabelle 12.8 nach Abzug des Unterschieds zwischen den Ofenfahrten (als Blockfaktor) übrig bleibt. Der p-Wert der Prüfgröße $86,9 / 18,8 = 4,6$ ergibt sich aus den Freiheitsgraden $f_1 = 3$ und $f_2 = 18$.

In diesem Beispiel ist die Varianz zwischen den Ofenfahrten ($13,3$) kleiner als innerhalb ($18,8$), entsprechend wird der Unterschied zwischen den Ofenfahrten als zufällig beurteilt. Dies ist nicht überraschend, da hier die Ofenfahrten nur zu Demonstrationszwecken unterstellt wurden, um so die verschiedenen Auswertungen an denselben Daten besser vergleichbar zu machen.

Hinweise

- Bei der Auswertung von Split-Plot-Versuchen ist sorgfältig darauf zu achten, dass die Struktur der Randomisierung richtig abgebildet wird. Die Anzahl verschiedener Möglichkeiten ist groß, entsprechend anspruchsvoll ist die richtige Auswertung.
- Im Versuchsplan werden die drei Ofenfahrten für jede Temperatur zwar mit 1, 2 und 3 bezeichnet, es handelt sich dabei aber trotzdem nicht um dieselben Fahrten (die 1. Fahrt bei 60°C ist schließlich nicht dasselbe, wie die 1. Fahrt bei 70°C und die 1. Fahrt bei 80°C). Man spricht bei der Ofenfahrt daher von einem „**geschachtelten**“ Faktor (**nested**) – im Gegensatz zu den üblichen „**gekreuzten**“ Faktoren (**crossed**) wie Zusatzstoff (Zusatzstoff 1 ist tatsächlich derselbe Zusatzstoff, ob nun bei 60, 70 oder 80°C). In Tabelle 12.10 wird dies durch die Bezeichnung Ofenfahrt [A] angedeutet und ist zu lesen als „Ofenfahrt geschachtelt im Faktor A: Temperatur“. Man spricht auch von hierarchischer Klassifikation – im Gegensatz zur Kreuzklassifikation.

Literatur

- [1] *Graf, U./Henning, H.-J./Stange, K./Wilrich, P.-T.*: Formeln und Tabellen der angewandten mathematischen Statistik. 3. Aufl. Springer Verlag, Berlin 1998
- [2] *Rasch, D./Herrendörfer, G./Bock, J./Victor, N./Guiard, V.*: Verfahrensbibliothek Versuchsprüfung und -auswertung. Oldenbourg, München, 2. Auflage 2007
- [3] *Box, G. E. P./Hunter, W. G./Hunter, J. S.*: Statistics for Experimenters. John Wiley, New York, 2. Auflage 2005
- [4] *Oehlert, G.W.*: A First Course in Design and Analysis of Experiments. Freeman, New York 2000
- [5] *Morris, M.*: Design of Experiments: An Introduction Based on Linear Models. Chapman&Hall, Boca Raton 2010

13

Screening für mehrstufige Faktoren

Versuche mit mehrstufigen qualitativen Faktoren können nur mit Varianzanalyse ausgewertet werden. In Kapitel 12 wurde dieses Verfahren auf vollständige faktorielle Versuchspläne angewendet. Nachteil dieser Pläne ist (noch mehr als bei zweistufigen Plänen), dass die Anzahl der Faktorstufenkombinationen sehr schnell mit der Anzahl der Faktoren zunimmt.

In diesem Kapitel werden daher fraktionelle Pläne für mehrstufige Faktoren behandelt (ähnlich zu Kapitel 8). Diese Pläne heißen in der klassischen Versuchsplanung Lateinische Quadrate bzw. (Hyper-)Lateinisch-Griechische Quadrate. G. Taguchi verwendet dieselben Pläne und nennt sie $L_9(3^4)$, $L_{16}(4^5)$ bzw. $L_{25}(5^6)$. Sie haben Auflösung III.

■ 13.1 Versuchspläne

Tabelle 13.1 zeigt beispielhaft den 3^{4-2} -Versuchsplan für dreistufige Faktoren, Tabelle 13.2 den 4^{5-3} -Plan für vierstufige Faktoren. Die Stufen werden jeweils mit 1, 2, 3 (und 4) bezeichnet. Da es sich im Normalfall um qualitative Faktoren handelt, bedeutet die Nummerierung keine Ordnung der Stufen, sie könnten beliebig umnummeriert werden.

Betrachtet man zwei beliebige Faktoren, so treten jeweils alle Faktorstufenkombinationen dieser beiden Faktoren auf. Die Pläne sind in diesem Sinn ausgewogen.

Mindestens eine der Spalten sollte nicht mit einem Faktor belegt werden, da diese zur Schätzung der Zufallsstreuung (bei $n = 1$) bzw. zur Absicherung gegen Wechselwirkungen (bei n -maliger Realisierung) benötigt wird. Der letzte Faktor in Tabelle 13.1 bzw. 13.2 ist daher eingeklammert.

| Nr. | Faktor A | Faktor B | Faktor C | (Faktor D) |
|-----|----------|----------|----------|------------|
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 2 | 2 | 2 |
| 3 | 1 | 3 | 3 | 3 |
| 4 | 2 | 1 | 2 | 3 |
| 5 | 2 | 2 | 3 | 1 |
| 6 | 2 | 3 | 1 | 2 |
| 7 | 3 | 1 | 3 | 2 |
| 8 | 3 | 2 | 1 | 3 |
| 9 | 3 | 3 | 2 | 1 |

TABELLE 13.1

Screening-Plan für dreistufige Faktoren A, B, C, (D)

| Nr. | Faktor A | Faktor B | Faktor C | Faktor D | (Faktor E) |
|-----|----------|----------|----------|----------|------------|
| 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 2 | 2 | 2 | 2 |
| 3 | 1 | 3 | 3 | 3 | 3 |
| 4 | 1 | 4 | 4 | 4 | 4 |
| 5 | 2 | 1 | 2 | 3 | 4 |
| 6 | 2 | 2 | 1 | 4 | 3 |
| 7 | 2 | 3 | 4 | 1 | 2 |
| 8 | 2 | 4 | 3 | 2 | 1 |
| 9 | 3 | 1 | 3 | 4 | 2 |
| 10 | 3 | 2 | 4 | 3 | 1 |
| 11 | 3 | 3 | 1 | 2 | 4 |
| 12 | 3 | 4 | 2 | 1 | 3 |
| 13 | 4 | 1 | 4 | 2 | 3 |
| 14 | 4 | 2 | 3 | 1 | 4 |
| 15 | 4 | 3 | 2 | 4 | 1 |
| 16 | 4 | 4 | 1 | 3 | 2 |

TABELLE 13.2

Screening-Plan für vierstufige Faktoren A, B, C, D, (E)

■ 13.2 Auswertung

Die Auswertung erfolgt mit Varianzanalyse. Aufgrund der Symmetrie des Plans kann man die Varianz für jeden Faktor wie in (12.2) direkt aus den Mittelwerten der Versuchsergebnisse für die einzelnen Stufen dieser Faktorspalte berechnen (sogenannte Randmittelwerte). Bei $n = 1$ erhält man die Varianz der Zufallsstreuung aus den Spalten, die nicht mit Faktoren belegt waren. Sind mehrere Spalten nicht belegt, verwendet man den Mittelwert der Varianzen und addiert die Freiheitsgrade.



Beispiel

Autoreifen aus vier verschiedenen Materialien sollen bezüglich ihrer Abriebfestigkeit verglichen werden. Das Problem besteht darin, dass der Abrieb vom Fahrer und von der Position des Reifens am Auto abhängen kann. Ohne sorgfältige Planung würden zufällige Unterschiede zwischen Fahrern und Positionen die Zufallsstreuung erhöhen und damit den erforderlichen Versuchsumfang wesentlich erhöhen. Im ungünstigsten Fall können Unterschiede zwischen Fahrern/Positionen sogar das Ergebnis verfälschen. Daher werden sie als Blockfaktoren behandelt.

Tabelle 13.2 ist ein ausgewogener Plan, mit dessen Hilfe systematische Unterschiede zwischen den Fahrern und den Positionen erkannt werden können. Dazu werden je vier Reifen eines jeden Materials gemäß Tabelle 13.3 auf die vier Positionen von vier Autos (mit verschiedenen Fahrern) verteilt.

Faktor A: Material mit Stufen A, B, C, D

Faktor B: Position am Auto mit Stufen VR (vorne rechts), VL, HR, HL

Faktor C: Fahrer mit Stufen Adam, Bert, Chris, Doris

TABELLE 13.3 Anwendung des Plans aus Tabelle 13.2 auf die Untersuchung der Abriebfestigkeit (mit Versuchsergebnissen, n = 1)

| Nr. | Material | Position | Fahrer | Faktor D | Faktor E | Abrieb |
|-----|----------|----------|--------|----------|----------|--------|
| 1 | A | VR | Adam | 1 | 1 | 6 |
| 2 | A | VL | Bert | 2 | 2 | 6 |
| 3 | A | HR | Chris | 3 | 3 | 6 |
| 4 | A | HL | Doris | 4 | 4 | 4 |
| 5 | B | VR | Bert | 3 | 4 | 7 |
| 6 | B | VL | Adam | 4 | 3 | 7 |
| 7 | B | HR | Doris | 1 | 2 | 6 |
| 8 | B | HL | Chris | 2 | 1 | 8 |
| 9 | C | VR | Chris | 4 | 2 | 11 |
| 10 | C | VL | Doris | 3 | 1 | 9 |
| 11 | C | HR | Adam | 2 | 4 | 7 |
| 12 | C | HL | Bert | 1 | 3 | 8 |
| 13 | D | VR | Doris | 2 | 3 | 4 |
| 14 | D | VL | Chris | 1 | 4 | 7 |
| 15 | D | HR | Bert | 4 | 1 | 3 |
| 16 | D | HL | Adam | 3 | 2 | 3 |

Aus jeweils vier Einzelwerten erhält man als Randmittelwerte z. B. für die Stufen des Faktors Material:

$$A: \frac{1}{4} \cdot (6 + 6 + 6 + 4) = 5,5 \quad B: \frac{1}{4} \cdot (7 + 7 + 6 + 8) = 7,0 \quad \text{usw.}$$

Tabelle 13.4 zeigt alle Randmittelwerte im Überblick. Bei den drei „echten“ Faktoren unterscheiden sich die Mittelwerte deutlich, während sich die Mittelwerte bei D und E nur zufällig unterscheiden. Aus den Unterschieden erhält man Schätzwerte für die Varianz. Der Mittelwert der Varianzen aus D und E ergibt die Zufallsstreuung (0,165).

Das Verhältnis der Varianzen gibt beobachtete F-Werte wie in Kapitel 12 (für den Faktor Material erhält man z. B. $15,08/0,165 = 91$).

Der Vergleich mit dem kritischen F-Wert für die Freiheitsgrade $f_1 = 3$ und $f_2 = 3 + 3 = 6$ aus Tabelle 12.4 (23,7) ergibt die Beurteilung:

Alle drei „echten“ Faktoren sind signifikant (Material und Fahrer ***, Position **).

| Faktor | Stufe 1 | Stufe 2 | Stufe 3 | Stufe 4 | 4-Varianz |
|----------|---------|---------|---------|---------|-----------|
| Material | 5,50 | 7,00 | 8,75 | 4,25 | 15,08 |
| Position | 7,00 | 7,25 | 5,50 | 5,75 | 3,08 |
| Fahrer | 5,75 | 6,00 | 8,00 | 5,75 | 4,75 |
| Faktor D | 6,75 | 6,25 | 6,25 | 6,25 | 0,25 |
| Faktor E | 6,50 | 6,50 | 6,25 | 6,25 | 0,08 |

TABELLE 13.4

Randmittelwerte für alle Faktorstufen und daraus berechnete Varianzen

In diesem Beispiel ist man nur am Effekt des Faktors Material interessiert. Der Abrieb soll minimiert werden, daher wählt man das Material D. Die anderen Faktoren dienten nur dazu, die Zufallsstreuung zu reduzieren. Der Faktor Position zeigt, dass der Abrieb vorne größer ist als hinten – ein Ergeb-

nis, das für Vorderradantrieb normal ist. Der Faktor Fahrer zeigt, dass Chris ein Rowdy ist. Bezuglich der eigentlichen Fragestellung sind beide Ergebnisse nicht relevant. Hätte man diese Faktoren aber nicht in die Untersuchung aufgenommen, so hätte man eine größere Zufallsstreuung erhalten.

■ 13.3 Einsatzempfehlungen

Die hier beschriebenen Screening-Pläne haben Auflösung III. Daher können Wechselwirkungen die Effekte der Faktoren verfälschen. Screening-Pläne werden daher vor allem eingesetzt, um wichtige Faktoren von unwichtigen Faktoren zu unterscheiden. Dazu genügen normalerweise zwei Stufen. Daher wird empfohlen, als Screening-Pläne normalerweise die Pläne aus Kapitel 8 zu verwenden.

Für detaillierte Untersuchungen (mehr als zwei Stufen) mit quantitativen Faktoren werden die Regressionspläne aus Kapitel 11 empfohlen.

Für die hier beschriebenen Versuchspläne bleiben daher nur Anwendungen, bei denen, bedingt durch die Fragestellung, mehr als zwei Stufen eines qualitativen Faktors verglichen werden sollen. Die anderen Faktoren haben dann meist eher die Bedeutung von Blockfaktoren, wie im obigen Beispiel. Wechselwirkungen von Blockfaktoren mit dem eigentlich interessierenden qualitativen Faktor sind meist vernachlässigbar.

Typische Fragestellungen sind:

- Vergleich von mehreren Materialien
- Vergleich von mehreren Produktvarianten
- Vergleich von mehreren Prozessvarianten.

Mit dieser Art von Fragestellungen hat sich R. Fisher, der Begründer der Versuchsplanung [1], vor allem auseinandergesetzt. Er verglich verschiedene Pflanzensorten oder Düngemittel. Bodenbedingte und klimatische Unterschiede sollten das Ergebnis nicht verfälschen und die Streuung möglichst nicht erhöhen.

Bei der Entwicklung und Verbesserung neuer Produkte und Fertigungsprozesse treten diese Probleme eher selten auf. Daher wurde hier nur diese eine Variante kurz behandelt. Weitere Einzelheiten und Versuchspläne wie unvollständige Blockpläne finden sich z.B. in [2].

Literatur

[1] Fisher, R. A.: *The Design of Experiments*. Oliver and Boyd, London 1935

[2] Morris, M.: *Design of Experiments: An Introduction Based on Linear Models*. Chapman&Hall, Boca Raton 2010

14

Versuchspläne für Mischungen

In manchen Anwendungen hängt das Versuchsergebnis nicht von der Gesamtmenge, sondern nur vom Mischungsverhältnis ab.



Beispiele

- Schmelzpunkt einer Legierung
- Eigenschaften von Beton in Abhängigkeit von der Menge Kies, Sand und Zement
- Eigenschaften eines Treibstoffes in Abhängigkeit von den Anteilen verschiedener Kohlenwasserstoffe
- Geschmack von Keksen in Abhängigkeit von den Zutaten Mehl, Zucker, Fett, Eier, Milch

Die Anteile der verschiedenen Komponenten sind voneinander abhängig – die Gesamtsumme ergibt immer 1. Daher sind die bisher behandelten Versuchspläne nicht direkt anwendbar. Folgende Möglichkeiten bestehen:

1. Besteht die Mischung (im Wesentlichen) aus zwei Komponenten, so kann man das Mischungsverhältnis als Faktor behandeln.



Beispiele

- das Verhältnis Luftmenge/Ölmenge bei einer Verbrennung
- das Verhältnis Sand/Kies bei Beton

2. Hat eine Komponente aus technischen Gründen immer den bei weitem größten Anteil, so kann man diese Komponente zum Auffüllen auf 100 % verwenden – sie ist kein Faktor. Die anderen Komponenten sind dann unabhängig voneinander.



Beispiel

Bei der Mischung von Beton kann man neben dem Verhältnis Sand/Kies als zweiten Faktor die Menge Zement verwenden, die man einer Tonne Sand/Kies-Mischung zusetzt. Die Mengen von mehreren Arten Zement oder Zement und Kalk kann man dann unabhängig voneinander verändern.

3. Man verwendet spezielle Versuchspläne für Mischungen, die berücksichtigen, dass die Summe der Anteile immer 1 ergibt.

Letztere Versuchspläne sind Gegenstand dieses Kapitels. Sie sind besonders nützlich, wenn der Anteil jeder Komponente im gesamten Bereich zwischen 0 (= nicht enthalten) und 1 (= nur diese Komponente ist vorhanden) verändert werden soll (oder zumindest im größten Teil dieses Bereichs). Sonst sind die 1. und 2. Möglichkeit besser geeignet.

Da nur in wenigen Anwendungen der gesamte Mischungsbereich untersucht werden soll, werden Mischungspläne (Mixture Designs) hier nur kurz behandelt. Weitere Einzelheiten finden sich z.B. in [1 – 4].

■ 14.1 Mischungspläne ohne Begrenzungen

Wenn der Anteil der verschiedenen Mischungskomponenten nicht begrenzt ist, kann jeder Anteil Werte zwischen 0 und 1 annehmen. Die Summe der Anteile aller Komponenten muss jedoch 1 betragen. Dadurch sind die Anteile der einzelnen Komponenten voneinander abhängig. Bild 14-1 zeigt die Folgen dieser Abhängigkeit grafisch am Beispiel von $k = 3$ Komponenten A, B und C. Bezeichnet man die Anteile der drei Komponenten mit x_1 , x_2 und x_3 , so gilt:

$$x_1 + x_2 + x_3 = 1 \quad (14.1)$$

Die möglichen Faktorstufenkombinationen liegen in einem gleichseitigen Dreieck. Bei $k = 4$ Komponenten liegen die möglichen Kombinationen in einem Tetraeder. Dreieck, Tetraeder und die entsprechenden Anordnungen bei mehr als 4 Komponenten heißen Simplexe, daher werden die Mischungspläne auch als Simplexpläne bezeichnet.

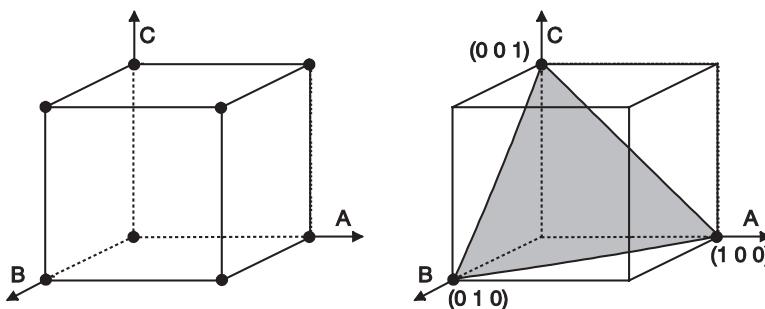
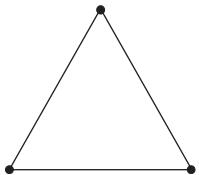


BILD 14-1 Versuchsplan für $k = 3$ unabhängige Faktoren A, B und C (links) und Simplex für $k = 3$ Komponenten einer Mischung (rechts)

Die Bilder 14-2 bis 14-4 zeigen am Beispiel von $k = 3$ Komponenten, wie man durch immer weitere Unterteilung des Simplex eine Hierarchie von sogenannten Simplexgitterplänen (simplex lattice designs) erhält. Die Anzahl der Teile, in die jede Simplexkante zerlegt wird, heißt Grad g des Plans. Die Struktur der Versuchspläne ist am Simplex gut zu erkennen. Die Umrechnung in die Faktorstufenkombinationen erfordert jedoch dreidimensionale Vorstellung, da sie im Würfel in Bild 14-1 abgelesen werden müssen. Ein Simplexgitterplan mit Grad g für k Komponenten enthält

$$m = \binom{g+k-1}{g} = \frac{k \cdot (k+1) \cdot (k+2) \cdot \dots \cdot (k+g-1)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot g} \quad (14.2)$$

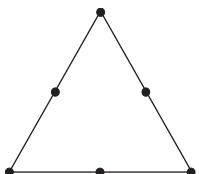
Faktorstufenkombinationen. Tabelle 14.1 zeigt m für ausgewählte Werte von k und g.



| Nr. | A | B | C |
|-----|---|---|---|
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |

BILD 14-2

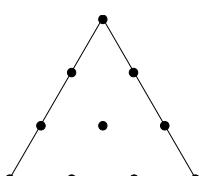
Simplexgitterplan mit Grad 1 für
k = 3 Komponenten



| Nr. | A | B | C |
|-----|-----|-----|-----|
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 1/2 | 1/2 | 0 |
| 5 | 1/2 | 0 | 1/2 |
| 6 | 0 | 1/2 | 1/2 |

BILD 14-3

Simplexgitterplan mit Grad 2 für
k = 3 Komponenten



| Nr. | A | B | C |
|-----|-----|-----|-----|
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 2/3 | 1/3 | 0 |
| 5 | 2/3 | 0 | 1/3 |
| 6 | 1/3 | 2/3 | 0 |
| 7 | 1/3 | 0 | 2/3 |
| 8 | 0 | 2/3 | 1/3 |
| 9 | 0 | 1/3 | 2/3 |
| 10 | 1/3 | 1/3 | 1/3 |

BILD 14-4

Simplexgitterplan mit Grad 3 für
k = 3 Komponenten

| Komponenten | Grad g = 1 | g = 2 | g = 3 | g = 4 |
|-------------|------------|-------|-------|-------|
| k = 2 | 2 | 3 | 4 | 5 |
| k = 3 | 3 | 6 | 10 | 15 |
| k = 4 | 4 | 10 | 20 | 35 |
| k = 5 | 5 | 15 | 35 | 70 |
| k = 6 | 6 | 21 | 56 | 126 |
| k = 7 | 7 | 28 | 84 | 210 |

TABELLE 14.1

Anzahl Faktorstufen-
kombinationen m in
einem Simplexgitter-
plan mit Grad g für k
Komponenten

Durch Weglassen bestimmter Punkte aus den Plänen mit Grad 3 bzw. 4 erhält man so genannte unvollständige Gitterpläne. Durch Hinzufügen weiterer Punkte im Innenbereich des Simplex erhält man so genannte Simplex-Zentroid-Pläne.

■ 14.2 Auswertung von Mischungsplänen

Die Auswertung von Mischungsplänen erfolgt mit Hilfe von Regressionspolynomen ähnlich zu Kapitel 11. Die Bedingung (14.1) führt jedoch dazu, dass einige der Koeffizienten im allgemeinen Ansatz verschwinden. Die Anzahl der verbleibenden Koeffizienten bei Anpassung eines Polynoms mit Grad g entspricht genau der Anzahl der Faktorstufenkombinationen des Simplexgitterplans mit Grad g in Tabelle 14.1.

An die Versuchsergebnisse eines Simplexgitterplans mit Grad 2 (Bild 14-3) kann daher ein Polynom mit Grad 2 angepasst werden (das entspricht dem quadratischen Modell in Gleichung 11.1). Bei einem Plan mit Grad 3 kann ein Polynom mit Grad 3 angepasst werden, usw. Zur Überprüfung des Modells benötigt man zusätzliche Faktorstufenkombinationen (z.B. die Innenpunkte eines Simplex-Zentroid-Plans). Die Zufallsstreuung kann aus der Wiederholung von Faktorstufenkombinationen bestimmt werden.

Die Ergebnisse der Anpassung können als Wirkungsfläche oder Höhenlinien über dem Simplex dargestellt werden. Bild 14-5 zeigt ein Beispiel. Aus solchen grafischen Darstellungen werden dann (wie in Kapitel 11) technische Konsequenzen abgeleitet.

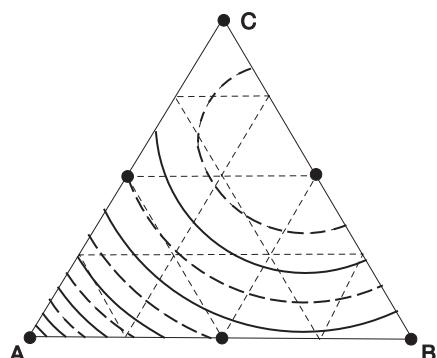
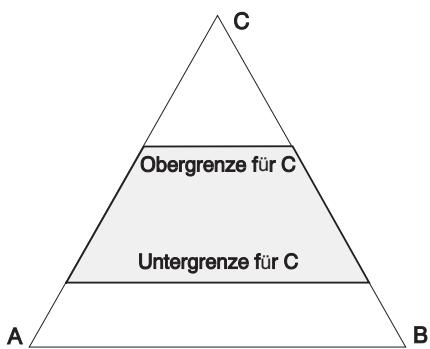


BILD 14-5
Beispiel für Höhenlinien eines Polynoms mit Grad 2

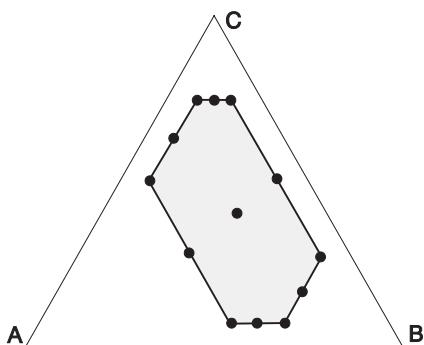
■ 14.3 Mischungspläne mit Begrenzungen

Häufig sind die Anteile der einzelnen Komponenten einer Mischung begrenzt. Bild 14-6 zeigt, wie sich eine Untergrenze und eine Obergrenze für eine einzelne Komponente auswirken – nur der markierte Bereich des Simplex ist für die Untersuchung zugänglich.

Bild 14-7 zeigt die Auswirkung einer Begrenzung der Anteile für alle drei Komponenten. Die Punkte markieren einen Versuchsplan, der es erlaubt, den verbliebenen Bereich zu untersuchen. Bei stark eingeschränkten Bereichen ist es meist günstiger, auf konventionelle Pläne zurückzugreifen (vgl. die Möglichkeiten Nr. 1 und 2 am Beginn dieses Kapitels).

**BILD 14-6**

Auswirkung einer Obergrenze und einer Untergrenze für die Komponente C:
Im markierten Bereich gilt: $0,2 < x_3 < 0,6$.

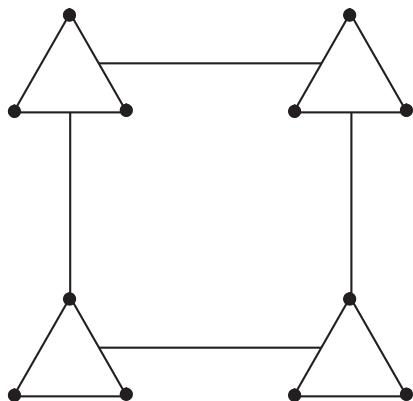
**BILD 14-7**

Wenn für alle Komponenten Ober- und Untergrenzen vorgegeben sind, kann nur ein Teil des Simplex untersucht werden – die Punkte zeigen einen Versuchsplan.

■ 14.4 Kombinierte Versuchspläne

Sollen in einem Versuchsplan gleichzeitig unabhängige Faktoren und Komponenten einer Mischung untersucht werden, so kann man die beiden Versuchspläne miteinander kombinieren. Bild 14-8 zeigt einen solchen kombinierten Versuchsplan für zwei unabhängige Faktoren und drei Mischungskomponenten. In jeder der vier Faktorstufenkombinationen des vollständigen faktoriellen Plans für die unabhängigen Faktoren wird ein Mischungsplan für die Komponenten der Mischung durchgeführt (in Bild 14-8 ein Plan mit Grad 1). Die Gesamtzahl der Faktorstufenkombinationen ist das Produkt der jeweiligen Anzahl für die Einzelpläne (hier $3 \times 4 = 12$). Dieses Prinzip kann auf mehr Faktoren bzw. Komponenten erweitert werden.

Mit Softwareunterstützung sind optimale Versuchspläne (wie in Absatz 11.2.4) für beliebige Kombinationen von Faktoren, beliebige Einschränkungen und beliebige Modelle möglich.

**BILD 14-8**

Kombination aus einem vollständigen faktoriellen Versuchsplan für zwei unabhängige Faktoren (großes Quadrat) und einem Mischungsplan für drei Komponenten einer Mischung (Dreieck an jeder Ecke des Quadrats)

Literatur

- [1] Petersen, H.: Grundlagen der Statistik und der statistischen Versuchsplanung. ecomed verlagsgesellschaft, Landsberg/Lech 1991
- [2] Scheffler, E.: Statistische Versuchsplanung und Auswertung. DVG, Stuttgart, 3. Auflage 1997
- [3] Spenhoff, E.: Prozesssicherheit durch statistische Versuchsplanung in Forschung, Entwicklung und Produktion. gfmt Verlag, München 1991
- [4] Cornell, C.: A Primer on Experiments with Mixtures. J. Wiley, New York 2011

15

Spezielle Zielgrößen

In den bisher behandelten Beispielen war die Zielgröße ein Messwert oder eine aus Messwerten berechnete quantitative Größe. Waren mehrere Zielgrößen wichtig, so wurden sie weitgehend unabhängig voneinander ausgewertet, und anschließend wurde ein Kompromiss zwischen z. T. widersprüchlichen Zielen gesucht. In diesem Kapitel wird gezeigt, wie

- Gut-Schlecht-Ergebnisse ausgewertet,
- Fehlerzahlen ausgewertet und
- mehrere Zielgrößen zu einer Größe zusammengefasst und bei der Suche nach Kompromissen gemeinsam bewertet werden können.

Die hier beschriebenen Methoden sind für alle Versuchspläne anwendbar (Kapitel 7, 8, 11, 13 und 14).

■ 15.1 Gut-Schlecht-Ergebnisse

Manchmal kann das Versuchsergebnis für eine bestimmte Zielgröße nur mit „gut“ oder „schlecht“ bewertet werden. Hier wird die Auswertung solcher Ergebnisse behandelt.

Vorab sei jedoch darauf hingewiesen, dass Gut-Schlecht-Ergebnisse möglichst vermieden werden sollten, da sie nur wenig Information enthalten und daher ein sehr großer Versuchsumfang erforderlich ist (vor allem, wenn der Anteil schlechter Einheiten klein ist, und das ist ja schließlich unser Ziel).

15.1.1 Möglichkeiten zur Vermeidung

Erstes Ziel sollte die Vermeidung von Gut-Schlecht-Ergebnissen sein. Dazu gibt es folgende Möglichkeiten:

1. Messwert verwenden

Manchmal steht hinter einem Gut-Schlecht-Ergebnis ein Messwert, und die Einteilung Gut-Schlecht wurde nur zur Vereinfachung gewählt. Dann sollte zur Bewertung von Versuchen unbedingt der Messwert selbst verwendet werden, auch wenn die Erfassung eines einzelnen Wertes aufwändiger ist. Die Einsparung aufgrund des geringeren Versuchsumfangs ist meist wesentlich größer als die Erhöhung des Aufwands pro Messung.

**Beispiele**

- Für den Durchmesser einer Welle ist folgende Spezifikation angegeben:
 $10,0 \pm 0,1 \text{ mm}$
 Die Einhaltung dieser Spezifikation kann man durch Einlegen der Wellen in zwei Schablonen mit 9,9 und 10,1 mm überprüfen. Dabei erhält man jedoch nur die Information, ob die Spezifikation eingehalten wird und ggf. welche Grenze verletzt wird (gut-schlecht). Besser ist es, den Messwert für den Durchmesser zu erfassen und in der Auswertung zu verwenden.
- Ein Gerät muss bis zu einer Temperatur von -30°C einsatzfähig sein. Dies kann man dadurch überprüfen, dass man das Gerät bei -30°C betreibt. Funktioniert es, so ist es gut, funktioniert es nicht, so ist es schlecht. Besser ist es allerdings, das Gerät langsam abzukühlen und zu messen, bei welcher Temperatur es ausfällt.
- Die Zuverlässigkeit eines neu entwickelten Gerätes soll optimiert werden. Der Ausfallzeitpunkt der einzelnen Geräte ist eine bessere Zielgröße als der Anteil, der eine bestimmte Mindestzeit überlebt.

2. Noten vergeben

Manchmal steht hinter einem Gut-Schlecht-Ergebnis eine subjektive Bewertung, z. B. der Vergleich eines Gesamteindrucks mit einem Grenzmuster. Eine solche subjektive Bewertung lässt sich zwar nicht direkt durch eine Messung ersetzen, häufig ist jedoch eine abgestufte Bewertung möglich. Der Gut-Bereich kann z. B. unterteilt werden in „perfekt“, „gut“ und „gerade noch akzeptabel“, der Schlecht-Bereich kann unterteilt werden in „knapp verfehlt“, „schlecht“ und „extrem schlecht“. In der Auswertung verwendet man dann für diese Bewertungen die Zahlenwerte 1 bis 6 als „Messwert“. Diese Noten sind sicher nicht normalverteilt. Da die üblichen Auswertungsverfahren jedoch nicht sehr empfindlich auf Abweichungen von der Normalverteilung reagieren, ist dies kein gravierendes Problem.

**Beispiele**

- Beurteilung eines in Mehrfarbendruck erzeugten Bildes
- Klang eines Verstärkers oder Lautsprechers

3. Anzahl Fehler oder Fehlstellen (evtl. auch Größe)

Eine feinere Unterteilung von „Schlecht“ erhält man, indem man die Anzahl der Fehler oder Fehlstellen erfasst. Manchmal lassen sich Fehler auch nach ihrer Größe oder Bedeutung unterteilen. Dann kann man auch gewichtete Fehlersummen verwenden (Auswertung im nächsten Absatz)

**Beispiele**

- Bei der Beurteilung eines Lötprozesses erhält man mehr Information, wenn man die Anzahl und Lage der Lötfehler erfasst, als wenn man nur zwischen Leiterplatten mit und ohne Fehler unterscheidet (vgl. dazu auch das Beispiel aus der Leiterplattenfertigung in Abschnitt 4.2).

- Bei der Beurteilung eines Gießprozesses erhält man mehr Information, wenn man die Anzahl, Lage und Größe der Lunker und Einschlüsse erfasst, als wenn man nur zwischen „fehlerfrei“ und „mit Fehlern“ unterscheidet (vgl. dazu auch das Beispiel aus der Gießerei in Abschnitt 4.2).

Die erste dieser drei Möglichkeiten ist die beste, dann folgen die Nummern 2 und 3. Und nur wenn keine dieser Möglichkeiten anwendbar ist, sollte man Gut-Schlecht-Ergebnisse erfassen.

Manchmal wird als Vorteil von Gut-Schlecht-Ergebnissen angeführt, dass man damit die verschiedenen Fehlerarten zusammenfassen kann. Bei der Auswertung von Versuchen sollte eine solche Zusammenfassung jedoch vermieden werden, da sie mit einem weiteren Informationsverlust verbunden ist: Verschiedene Fehlerarten werden meist durch verschiedene Faktoren beeinflusst. Diese Information geht bei einer Zusammenfassung verloren.

15.1.2 Auswertung

Bei jeder Faktorstufenkombination i wird dieselbe Anzahl Teile n gefertigt. Wenn von diesen n Teilen x_i fehlerhaft (defekt) sind, so kann man als Zielgröße den

$$\text{Anteil fehlerhafter Teile } p_i = \frac{x_i}{n} \quad (15.1)$$

verwenden (direkt als Anteil oder in %). Der Anteil p genügt der sogenannten Binomialverteilung. Nachteil der Zielgröße p ist, dass ihre Varianz vom Wert von p abhängt¹. Bild 15-1 zeigt links diese Abhängigkeit. Bei $p = 0$ und 1 ist die Varianz 0, bei $p = 0,5$ ist sie maximal.

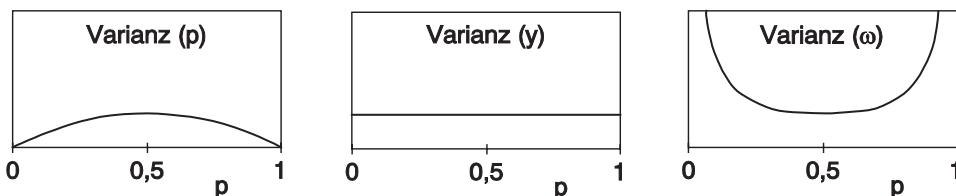


BILD 15-1 Varianz der Größen p , y und ω in Abhängigkeit von p

Für solche Situationen wurde in Abschnitt 6.4 eine varianzstabilisierende Transformation empfohlen. Bei Anteilen p ist die Transformation

$$y_i = \arcsin \sqrt{p_i} \quad (15.2)$$

geeignet [1]. Bild 15-2 stellt diese Transformation dar.

¹ Auch intuitiv bewertet man den Unterschied zwischen $p = 0$ und $0,01$ (0 und 1 %) höher als zwischen $p = 0,50$ und $0,51$ (50 und 51 %). Der Unterschied zwischen $p = 0,99$ und 1 (99 und 100 %) hat wieder dieselbe Bedeutung wie zwischen 0 und 0,01.

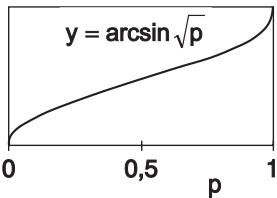


BILD 15-2

Grafische Darstellung der Transformation $y = \arcsin \sqrt{p}$

Bei $p = 0$ und $p = 1$ führt bereits eine kleine Änderung von p zu einer großen Änderung von y . Dadurch wird dort die Varianz von y erhöht. Bild 15-1 zeigt in der Mitte, dass mit Hilfe von (15.2) eine perfekte Kompensation erreicht wird, die Varianz ist konstant, unabhängig von p .

G. Taguchi [2] empfiehlt statt (15.2) die so genannte

$$\omega\text{-Transformation } \omega_i = 10 \cdot \log \frac{p_i}{1 - p_i} \quad (15.3)$$

Der allgemeine Verlauf der ω -Transformation ist ähnlich zu Bild 15-2, in der Nähe von $p = 0$ und $p = 1$ verläuft sie jedoch noch steiler als y . Bild 15-1 zeigt rechts die Varianz von ω in Abhängigkeit von p : Man erkennt, dass überkompensiert wurde und die Varianz von ω bei $p = 0$ und $p = 1$ divergiert.

Abgesehen von einem konstanten Faktor ist (15.3) die Transformation, die in der klassischen Statistik bei der so genannten Logit-Regression [3] verwendet wird. Dort wird sie empirisch damit begründet, dass sie zu einer Linearisierung von Dosis-Wirkungs-Zusammenhängen führt, die von der

$$\text{logistischen Funktion } p = \frac{\exp(\beta_0 + \beta_1 \cdot x)}{1 + \exp(\beta_0 + \beta_1 \cdot x)} \quad (15.4)$$

beschrieben werden. Vorteil der ω - oder Logit-Transformation ist, dass für p nur Werte im sinnvollen Bereich $0 < p < 1$ auftreten können. Nachteil dieser Transformation ist, dass die Varianz der transformierten Größe ω von p abhängt (Bild 15-2) und daher eine gewichtete lineare Regression durchgeführt werden muss.²

Hier wird empfohlen, die Transformation (15.2) zu verwenden. Bei dieser Transformation sind rein rechnerisch Werte $y < 0$ möglich. Wenn ein solches Ergebnis auftritt, ist es als „ p sehr klein“ zu interpretieren. Vorteil der Transformation (15.2) ist, dass die transformierte Zielgröße y näherungsweise normalverteilt ist (mit konstanter Varianz). Die Varianz von y muss nicht aus den Daten ermittelt werden, sondern ergibt sich direkt aus den Eigenschaften der Binomialverteilung. Sie wird daher im Folgenden mit σ statt s bezeichnet. Bei einer Stichprobe vom Umfang n erhält man statt (6.6) als Varianz von y (im Bogenmaß)

$$\sigma^2 \approx \frac{1}{4 \cdot n} \quad (15.5)$$

² Dies ist eine Verallgemeinerung der in Kapitel 10 behandelten linearen Regression, bei der statt der Summe der quadrierten Abweichungen (10.3) die mit $1/\text{Varianz}$ gewichtete Summe minimiert wird. Die Analyse erfolgt iterativ.

Bei zweistufigen Faktoren kann man wie in Kapitel 7 und 8 Effekte für y berechnen. Die Standardabweichung der Effekte beträgt, analog zu (7.13),

$$\sigma_{\bar{d}} = \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{n \cdot m}} \quad (15.6)$$

Der Freiheitsgrad beträgt (formal) $f \rightarrow \infty$ (da die Standardabweichung nicht aus den Daten errechnet wird, sondern aus der Binomialverteilung folgt). Um die Signifikanz der Effekte für y zu beurteilen, vergleicht man sie wie in Kapitel 7 mit der Breite der Vertrauensbereiche. Danach wählt man die Faktorstufenkombination mit dem kleinsten Wert von y . Erst nach dieser Auswahl wird das Ergebnis wieder in p umgerechnet.

Soll bei einem Anteil fehlerhafter Einheiten von p_0 ein wahrer Unterschied Δp mit hoher Wahrscheinlichkeit erkannt werden, so benötigt man analog zu (6.17) bzw. (7.15) einen Versuchsumfang von

$$N = 60 \cdot \frac{p_0 \cdot (1 - p_0)}{(\Delta p)^2} \quad (15.7)$$

Man beachte, dass in (15.7) bei vorgegebenem relativen Unterschied $\Delta p/p_0$ der erforderliche Versuchsumfang wie $1/p_0$ zunimmt. Das bedeutet: Je besser man bereits vor Beginn der Untersuchung ist (je kleiner p_0), desto größer ist der erforderliche Versuchsumfang N .



Beispiel

Ein bestimmtes Antriebsritzel wird durch Trockenfräsen hergestellt. Erfahrungsgemäß enthalten etwa 3 % der Ritzel Beschädigungen an den Flanken („Schmierer“). Um Verbesserungsmöglichkeiten zu erkennen, soll untersucht werden, ob der Anteil beschädigter Ritzel von der Art der Pressluftdüse (zum Wegblasen der Späne), vom Fräsermaterial und vom Spanwinkel abhängt.

| Faktor | Stufe – | Stufe + |
|----------------|---------|----------|
| Pressluftdüse | Typ A | Typ B |
| Fräsermaterial | alt | neu |
| Spanwinkel | 0 Grad | + 5 Grad |

Ein wahrer Unterschied von 2 % soll mit hoher Wahrscheinlichkeit erkannt werden, d. h. man möchte erkennen, wenn z. B. mit Pressluftdüsen vom Typ A nur 2 % der Ritzel beschädigt sind, mit Pressluftdüsen vom Typ B dagegen 4 %. Dazu benötigt man gemäß (15.7) insgesamt

$$N = 60 \cdot \frac{0,03 \cdot 0,97}{(0,02)^2} = 4365 \text{ Teile}^3$$

Als Versuchsplan wird ein vollständiger faktorieller 2^3 -Plan mit $m = 2^3 = 8$ Faktorstufenkombinationen verwendet. Bei jeder Kombination werden daher $4365/8 \approx 500$ Ritzel gefertigt. Tabelle 15.1

³ Man beachte den großen Versuchsumfang, der trotz des relativ bescheidenen Wunsches, den Unterschied zwischen 2 und 4 % zu erkennen, bereits erforderlich ist.

Soll bei einem mittleren Anteil $p_0 = 0,3\%$ der Unterschied zwischen 0,2 und 0,4 % mit hoher Wahrscheinlichkeit erkannt werden, so erhält man $N \approx 45\,000$; soll bereits der Unterschied zwischen 0,25 % und 0,35 % erkannt werden, so erhält man sogar $N \approx 180\,000$. Ein solcher Versuchsumfang ist praktisch nicht mehr realisierbar.

Er unterstreicht die Bedeutung der Empfehlung in Absatz 15.1.1, statt Gut-Schlecht-Werten möglichst Messwerte zu verwenden.

zeigt die Ergebnisse für die Anzahl beschädigter Ritzel x_i (von jeweils 500), den daraus berechneten Anteil p_i und den Wert der transformierten Zielgröße y_i (im Bogenmaß). Mit Hilfe der Vorzeichenspalten werden die Effekte für die Zielgröße y berechnet.

TABELLE 15.1 Versuchsplan und Ergebnisse im Beispiel

x_i = Anzahl beschädigter Ritzel (von jeweils 500)

p_i = daraus berechneter Anteil ($= x_i / 500$)

$y_i = \arcsin \sqrt{p_i}$ = Wert der transformierten Zielgröße

| Nr. | A | B | C | AB | AC | BC | ABC | x_i | p_i | y_i |
|--------------|--------|---------|--------|--------|--------|--------|---------|-------|-------|-------|
| 1 | - | - | - | + | + | + | - | 12 | ,024 | ,156 |
| 2 | + | - | - | - | - | + | + | 16 | ,032 | ,180 |
| 3 | - | + | - | - | + | - | + | 2 | ,004 | ,063 |
| 4 | + | + | - | + | - | - | - | 24 | ,048 | ,221 |
| 5 | - | - | + | + | - | - | + | 15 | ,030 | ,174 |
| 6 | + | - | + | - | + | - | - | 23 | ,046 | ,216 |
| 7 | - | + | + | - | - | + | - | 4 | ,008 | ,090 |
| 8 | + | + | + | + | + | + | + | 34 | ,068 | ,264 |
| Σy_i | 0,398 | -0,088 | 0,124 | 0,266 | 0,034 | 0,016 | -0,002 | | | |
| Effekt | 0,0995 | -0,0220 | 0,0310 | 0,0665 | 0,0085 | 0,0040 | -0,0005 | | | |
| Sign. | *** | - | * | *** | - | - | - | | | |

Aus (15.6) erhält man $\sigma_d = \frac{1}{\sqrt{8 \cdot 500}} = 0,0158$.

Mit Tabelle 6.4 erhält man für die Breite der Vertrauensbereiche (wegen $f \rightarrow \infty^4$)

$$95\%: t \cdot s_d = u \cdot \sigma_d = 1,960 \cdot 0,0158 = 0,0309$$

$$99\%: t \cdot s_d = u \cdot \sigma_d = 2,576 \cdot 0,0158 = 0,0407$$

$$99,9\%: t \cdot s_d = u \cdot \sigma_d = 3,291 \cdot 0,0158 = 0,0520$$

Der Vergleich der Breiten der Vertrauensbereiche mit den Effekten ergibt die Signifikanz in der letzten Zeile von Tabelle 15.1. Da der Effekt der 2FWW AB signifikant ist, müssen die Faktoren A und B gemeinsam betrachtet werden. Tabelle 15.2 zeigt die Mittelwerte der Zielgröße y für die vier Faktorstufenkombinationen dieser Faktoren.

TABELLE 15.2 Mittelwerte der Zielgröße y für die 4 Faktorstufenkombinationen von A und B

| Pressluftdüse | | Fräsermaterial | | Mittelwert Zielgröße y |
|---------------|-------|----------------|-----|---------------------------------------|
| - | Typ A | - | alt | $\frac{1}{2} (0,156 + 0,174) = 0,165$ |
| + | Typ B | - | alt | $\frac{1}{2} (0,180 + 0,216) = 0,198$ |
| - | Typ A | + | neu | $\frac{1}{2} (0,063 + 0,090) = 0,076$ |
| + | Typ B | + | neu | $\frac{1}{2} (0,221 + 0,264) = 0,242$ |

Die Faktorstufenkombination „Pressluftdüse Typ A“ mit „Fräsermaterial neu“ führt zum besten Ergebnis (kleinstes y und damit p). Daher werden in Zukunft Pressluftdüsen vom Typ A und Fräser aus

⁴ Die t-Verteilung geht in die Normalverteilung über, die t-Werte sind identisch zu den u-Werten der Normalverteilung.

dem neuen Material eingesetzt. Der Effekt des Faktors C ist positiv, d. h. für die untersuchte Anlage ist ein Spanwinkel von 0 Grad besser als 5 Grad.

Für die optimale Faktorstufenkombination erwartet man $y = 0,076 - 0,031/2 = 0,06$. Nach Umrechnung erhält man für den erwarteten Anteil Ritzeln mit Beschädigungen $p = (\sin 0,06)^2 = 0,004 = 0,4\%$, eine deutliche Verbesserung.

■ 15.2 Anzahl Fehler

Die Zielgröße „Anzahl Fehler auf einer Einheit“ wurde bereits in Absatz 6.4.2 behandelt. Dort wurde empfohlen, als transformierte Zielgröße

$$y = \sqrt{\text{Anzahl}} \quad (15.8)$$

zu verwenden. Sind die Fehler auf einer Einheit voneinander unabhängig, so ist die Anzahl Fehler poissonverteilt, und die Standardabweichung von y ergibt sich direkt aus dieser Verteilung. Wie in (15.6) erhält man dann für die Standardabweichung der Effekte

$$\sigma_{\bar{d}} = \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{n \cdot m}} \quad (15.9)$$

wobei N die Gesamtzahl der untersuchten Einheiten ist.

In vielen praktischen Anwendungen sind die Fehler auf einer Einheit jedoch voneinander abhängig. Hier wird daher empfohlen, die Größe y wie einen normalverteilten Messwert zu behandeln und die Standardabweichung aus den n Realisierungen jeder Faktorstufenkombination zu ermitteln. Alle bisher behandelten Auswertungsverfahren können dann auf y angewendet werden. Das Ergebnis für die Standardabweichung der Effekte kann dann mit (15.9) verglichen werden. (15.9) stellt eine Untergrenze dar, die nur zufällig unterschritten wird.

Darf man davon ausgehen, dass die Fehler voneinander unabhängig sind, kann man den erforderlichen Versuchsumfang abschätzen. Soll bei einer mittleren Anzahl von Fehlern pro Einheit von μ_0 ein wahrer Unterschied $\Delta\mu$ mit hoher Wahrscheinlichkeit erkannt werden, so benötigt man analog zu (15.7) einen Versuchsumfang von

$$N = 60 \cdot \frac{\mu_0}{(\Delta\mu)^2} \quad (15.10)$$

Wenn die mittlere Anzahl Fehler pro Einheit sehr viel kleiner als 1 ist, gibt es nur sehr wenige Einheiten mit mehr als einem Fehler, und es gilt

$$\mu_0 \approx \mu_0 << 1. \quad (15.11)$$

In diesem Grenzfall sind (15.7) und (15.10) identisch. Es ist dann praktisch gleichgültig, ob man die Anzahl Fehler oder fehlerhafte Einheiten betrachtet. Auch aus (15.10) erhält man dann einen sehr großen Versuchsumfang, der mit abnehmendem μ_0 zunimmt, wenn der relative Unterschied $\Delta\mu/\mu_0$ vorgegeben ist.

■ 15.3 Mehrere Zielgrößen

Bei den meisten Versuchen sollen mehrere Zielgrößen gleichzeitig optimiert werden. Dabei treten häufig Zielkonflikte auf. Einstellungen der Faktoren, die für eine Zielgröße günstig sind, können für andere Zielgrößen ungünstig sein.

Aufgrund der quantitativen Kenntnis der Abhängigkeit aller Zielgrößen von allen Faktoren kann man bei solchen Zielkonflikten nach Kompromissen suchen. In Absatz 11.1.5 wurde folgende Vorgehensweise vorgeschlagen:

1. Zunächst wird für jede Zielgröße getrennt ein Modell an die Daten angepasst. Nur signifikante Abhängigkeiten zwischen Zielgrößen und Faktoren werden dabei berücksichtigt.
2. Anschließend werden schrittweise optimale Einstellungen für alle Faktoren gesucht. Ausgehend von Faktoren mit großen Wechselwirkungen werden die einzelnen Zielgrößen in Abhängigkeit von jeweils zwei Faktoren grafisch dargestellt. Bei der Optimierung wird berücksichtigt,
 - wie stark jede Zielgröße von den einzelnen Faktoren abhängt,
 - wie wichtig jede Zielgröße für das Versuchsziel ist und
 - inwieweit der Zielpunkt für jede Zielgröße bereits durch Optimierung der anderen Faktoren erreicht wurde.

Sollen alle Zielgrößen entweder minimiert oder maximiert werden, so ist diese Vorgehensweise auch bei drei bis vier Zielgrößen noch überschaubar.

Vorteil dieser Vorgehensweise ist, dass man noch während der Optimierung die Ziele anpassen und ggf. die Gewichtung verschieben kann. Bei dieser schrittweisen Vorgehensweise erhält man einen tiefen Einblick in die Ursachen-Wirkungs-Zusammenhänge. Dadurch erhält man oft Anregungen für weitere Verbesserungsmöglichkeiten.

Nachteil dieser Vorgehensweise ist jedoch, dass viele Faktoren und Zielgrößen gleichzeitig betrachtet werden müssen. Insbesondere wenn für manche der Zielgrößen Zielwerte vorgegeben sind, die weder über- noch unterschritten werden sollen, überfordert die gleichzeitige Optimierung mehrerer Zielgrößen die meisten Anwender. Dann ist es hilfreich, die vielen Zielgrößen zu einer Größe zusammenzufassen. Derringer und Suich [4] haben dazu eine einfache, gut einsetzbare Möglichkeit vorgeschlagen. Sie wird im Folgenden beschrieben.

Wie bisher wird zunächst für jede Zielgröße getrennt ein Modell angepasst und auf Signifikanz überprüft. Das resultierende Modell muss für jede Zielgröße einzeln plausibel sein.

Die Zusammenfassung mehrerer Zielgrößen erfolgt mit Hilfe der „Erwünschtheit“. Sie wird durch die so genannte Wunschkfunktion (desirability function) quantitativ beschrieben. Für jede Zielgröße wird einzeln festgelegt, welche Werte erwünscht sind. Die Wunschkfunktion liegt immer zwischen 0 und 1, wobei 0 unerwünscht bzw. nicht zulässig und 1 erwünscht bzw. optimal bedeutet. Im Prinzip kann man eine beliebige Funktion zur Beschreibung der Erwünschtheit in Abhängigkeit vom Wert der Zielgröße verwenden. Bild 15-3 zeigt als einfachste Möglichkeit eine lineare Abhängigkeit.

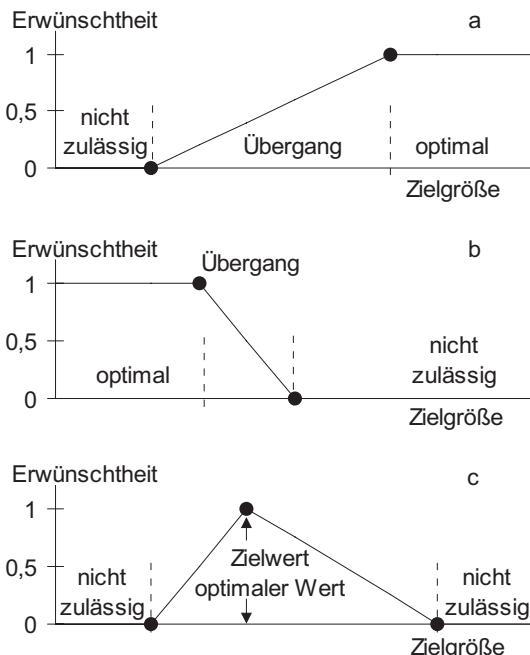


BILD 15-3

Erwünschtheit als Funktion der Zielgröße (Wunschkurve):

- Zielgröße soll maximiert werden
- Zielgröße soll minimiert werden
- Zielgröße soll möglichst nahe bei optimalem Wert liegen.

Zwischen den markierten Fixpunkten ist die Wunschkurve linear.

Für jede Zielgröße kann man mit dem angepassten Modell den erwarteten Wert für jede beliebige Faktorstufenkombination berechnen. Mit der Wunschkurve für diese Zielgröße erhält man zunächst einzeln die Erwünschtheit dieses erwarteten Werts. Die Gesamt-Erwünschtheit ist dann der geometrische Mittelwert der Erwünschtheiten aller Zielgrößen. Sie beschreibt, wie nahe alle Zielgrößen am Optimum liegen:

- Jeder Wert einer Zielgröße im optimalen Bereich wird als gleich gut bewertet.
- Gesamt-Erwünschtheit = 1 bedeutet, dass alle Zielgrößen in ihrem jeweiligen optimalen Bereich liegen.
- Bereits eine einzige Zielgröße im nicht zulässigen Bereich bewirkt, dass die Gesamt-Erwünschtheit = 0 ist (nicht zulässig).
- Je schmäler der Übergangsbereich in der Wunschkurve einer Zielgröße ist, desto höheres Gewicht hat diese Zielgröße bei der Optimierung.
- Die Verwendung des geometrischen Mittelwerts bewirkt, dass das Gewicht einer Zielgröße in der Nähe des nicht zulässigen Bereichs zunimmt.

Die optimale Faktorstufenkombination erhält man mit numerischen Optimierungsverfahren, oder indem man die Gesamt-Erwünschtheit auf einem relativ feinmaschigen Gitter im untersuchten Bereich berechnet. Der Punkt mit der maximalen Gesamt-Erwünschtheit ist die beste Faktorstufenkombination. Diese Optimierung ist praktisch nur mit Rechnerunterstützung möglich.

Die Gesamt-Erwünschtheit einer Faktorstufenkombination hängt von der Wunschkurve für jede einzelne Zielgröße ab. Daher wird empfohlen, mit verschiedenen Fixpunkten in Bild 15-3 zu experimentieren und vor einer endgültigen Bewertung den Einfluss solcher Veränderungen zu beobachten.

In ungünstigen Fällen kann die Gesamt-Erwünschtheit überall 0 sein. Das bedeutet, dass es im untersuchten Bereich der Faktorstufen nicht möglich ist, alle Mindestanforderungen an die Zielgrößen gleichzeitig zu erfüllen. Um in einer solchen Situation trotzdem ein „Optimum“ zu erhalten, müssen einzelne Mindestanforderungen entspannt werden. Dazu sind technische Überlegungen erforderlich.

Liegt das Optimum am Rand des untersuchten Bereichs (oder rechnerisch außerhalb des Bereichs), so sollten zusätzliche Einzelversuche in der Umgebung des rechnerischen Optimums durchgeführt werden.

Auch für Faktorwerte kann eine Erwünschtheit festgelegt werden. Dadurch kann man eine Präferenz für Faktorwerte in bestimmten Bereichen ausdrücken.



Beispiel Laserschneiden (aus Absatz 11.1.5)

In Absatz 11.1.5 wurde ein quadratisches Modell an die Messwerte für Barthöhe und Rautiefe angepasst. Das Prozessoptimum wurde dann mit Hilfe von systematischen grafischen Darstellungen dieses Modells bestimmt.

Hier wird von denselben Daten ausgegangen. Zunächst werden die Regressionskoeffizienten wie in Bild 11-2 bzw. 11-3 bestimmt. Alle Regressionskoeffizienten werden im Modell belassen.

Für jede Zielgröße wird die Wunschfunktion festgelegt. Hier soll die Auswirkung von verschiedenen Wunschfunktionen auf das berechnete Optimum dargestellt werden.

1. Erwünschtheit nur für Barthöhe und Rautiefe

Barthöhe und Rautiefe sollen möglichst klein sein. Technisch können keine Werte < 0 auftreten. Das Regressionsmodell ergibt jedoch formal auch negative Werte. Um eine möglichst große Fertigungssicherheit zu haben, strebt man solche negativen Werte an (im Wissen, dass sie ein Artefakt des Modells sind, vgl. Abschnitt 11.3).

Tabelle 15.3 zeigt mögliche Wunschfunktionen. Zwischen den angegebenen Fixpunkten (Grenzen) verändert sich die Erwünschtheit linear. Das bedeutet z.B., dass eine rechnerische Barthöhe von $-0,5$ eine Erwünschtheit von $0,5$ und eine rechnerische Rautiefe von $-0,5$ eine Erwünschtheit von $0,25$ hat. Positive Werte der Barthöhe und Rautiefe sind bei dieser Festlegung nicht zulässig. Waren die Versuchsergebnisse weniger günstig gewesen, hätte man sich bei der Optimierung natürlich mit niedrigeren Anforderungen begnügt.

| Erwünschtheit | 0 | 1 | 0 | Form |
|---------------|---|-----------|----------|------|
| für Barthöhe | | ≤ -1 | ≥ 0 | b |
| für Rautiefe | | ≤ -2 | ≥ 0 | b |

TABELLE 15.3

Festlegung der Fixpunkte und Form der Wunschfunktionen gemäß Bild 15-3

2. Erwünschtheit zusätzlich für Geschwindigkeit und Abstand

In Absatz 11.1.5 wurde zusätzlich berücksichtigt, dass eine hohe Geschwindigkeit hoher Durchsatz und damit niedrige Fertigungskosten bedeutet und dass ein großer Abstand bessere Fertigungssicherheit ergibt. Hier kann dies durch die zusätzliche Festlegung einer Erwünschtheit für Geschwindigkeit und Abstand erfolgen. Tabelle 15.4 zeigt eine mögliche Festlegung. Danach ist eine Geschwindigkeit unter 2 m/min nicht zulässig und über 6 m/min optimal, sowie ein Abstand unter $0,3 \text{ mm}$ nicht zulässig und über $1,1 \text{ mm}$ optimal.

| Erwünschtheit | 0 | 1 | 0 | Form |
|---------------------|------------|------------|----------|------|
| für Barthöhe | | ≤ -1 | ≥ 0 | b |
| für Rautiefe | | ≤ -2 | ≥ 0 | b |
| für Geschwindigkeit | ≤ 2 | ≥ 6 | | a |
| für Abstand | $\leq 0,3$ | $\geq 1,1$ | | a |

TABELLE 15.4

Festlegung der Fixpunkte und Form der Wunschfunktionen gemäß Bild 15-3, auch für Geschwindigkeit und Abstand

3. Verschärfte Vorgaben für den Abstand

Um zu demonstrieren, dass ein schmälerer Übergangsbereich zu einer größeren Gewichtung der betreffenden Zielgröße führt, wird alternativ der Mindestabstand von 0,3 auf 0,7 mm erhöht. Alle anderen Werte aus Tabelle 15.4 bleiben unverändert.

TABELLE 15.5 Ergebnisse im Überblick: Vergleich der Optima für verschiedene Wunschfunktionen

| | 1. Nur Barthöhe und Rautiefe | 2. Zusätzlich Geschwindigkeit und Abstand | 3. Abstand verschärft |
|-----------------|------------------------------|---|-----------------------|
| Geschwindigkeit | 3,8 m/min | 4,4 m/min | 4,4 m/min |
| Druck | 12,5 bar | 12 bar | 12 bar |
| Abstand | 0,5 mm | 0,75 mm | 0,9 mm |
| Fokus | 1,5 mm | 1,5 mm | 1,5 mm |
| Leistung | 1,5 kW | 1,5 kW | 1,5 kW |
| Barthöhe | -0,27 mm | -0,22 mm | -0,16 mm |
| Rautiefe | -1,10 µm | -0,96 µm | -0,78 µm |
| Erwünschtheit | | | |
| Barthöhe | 0,27 | 0,22 | 0,16 |
| Rautiefe | 0,55 | 0,48 | 0,39 |
| Geschwindigkeit | - | 0,55 | 0,55 |
| Abstand | - | 0,56 | 0,50 |
| Gesamt | 0,39 | 0,42 | 0,36 |

4. Ergebnisse

Tabelle 15.5 zeigt die Ergebnisse für die verschiedenen Wunschfunktionen im Überblick. Die Zeilen Geschwindigkeit bis Leistung geben die Einstellungen dieser Faktoren, die jeweils die Gesamt-Erwünschtheit maximieren. Die Zeilen Barthöhe und Rautiefe geben die für diese Einstellungen errechneten Werte für die Zielgrößen.

Daraus erhält man die Erwünschtheit für jede Zielgröße einzeln und (als geometrischen Mittelwert) die Gesamt-Erwünschtheit.

Ein Vergleich der drei Ergebnisspalten zeigt, in Übereinstimmung mit den Ergebnissen in Absatz 11.1.5:

- Unabhängig von weiteren Randbedingungen ist ein Druck von 12 – 12,5 bar, ein Fokus von 1,5 mm und eine Leistung von 1,5 kW optimal.
- Berücksichtigt man nur die Zielgrößen Barthöhe und Rautiefe, so ist eine Geschwindigkeit von ca. 3,8 m/min und ein relativ kleiner Abstand optimal.
- Berücksichtigt man zusätzlich Fertigungsaspekte, so verschiebt sich das Optimum zu höheren Geschwindigkeiten und größeren Abständen. Dabei werden Barthöhe und Rautiefe weniger günstig, diese Verschlechterung wird aber durch die Verbesserung des Durchsatzes und der Fertigungssicherheit mehr als ausgeglichen.

- Eine schärfere Vorgabe beim Abstand führt zu einer Vergrößerung des optimalen Abstandes. Dafür nimmt man aber eine Verschlechterung bei Barthöhe und Rautiefe in Kauf. Alle Werte der Erwünschtheit werden etwas kleiner.
-

Dieses Beispiel zeigt:

Mit Hilfe der Erwünschtheit kann man direkt nach optimalen Faktorstufenkombinationen suchen. Bei einer geeigneten Festlegung der Erwünschtheit erhält man dieselben Ergebnisse wie bei der grafischen Betrachtung der angepassten Modelle. Das Optimum hängt aber von der (in weiten Bereichen veränderbaren) Definition der Erwünschtheit ab. Daher wird empfohlen, im Rahmen der Optimierung verschiedene Varianten zu testen und die Ergebnisse miteinander zu vergleichen.

Mit Hilfe der Erwünschtheit erhält man auch dann noch ein Optimum, wenn die grafischen Verfahren allein zu unübersichtlich sind. Dies soll das folgende Beispiel von Derringer und Suich [4] zeigen.



Beispiel Gummimischung (nach Derringer und Suich [4])

Die Zusammensetzung der Gummimischung eines Reifens soll optimiert werden. Dazu wird der Einfluss der Faktoren

- Menge hydriertes Silica
- Menge Silan-Verbinder und
- Menge Schwefel

auf die Zielgrößen

- PICO Abriebindex
- 200%-Modul
- Bruchdehnung und
- Härte

untersucht. Als Versuchsplan wird ein drehbarer zentral zusammengesetzter Plan mit 6 Realisierungen des Zentrums verwendet. Tabelle 15.6 zeigt den Versuchsplan und die Messwerte für die Zielgrößen in systematischer Reihenfolge.

Wie in [4] wird zunächst für jede Zielgröße ein volles quadratisches Modell angepasst, um so möglichst leicht nachvollziehbare Ergebnisse zu erhalten. Beachten Sie jedoch die Hinweise zur Modellverbesserung am Ende des Abschnitts.

TABELLE 15.6 Versuchsplan und Ergebnisse zur Optimierung einer Gummimischung

| syst. Nr. | Silica | Silan | Schwefel | Abrieb-index | 200%-Modul | Bruchdehnung | Härte |
|-----------|--------|-------|----------|--------------|------------|--------------|-------|
| 1 | 0,7 | 40 | 2,8 | 102 | 900 | 470 | 67,5 |
| 2 | 1,7 | 40 | 1,8 | 120 | 860 | 410 | 65,0 |
| 3 | 0,7 | 60 | 1,8 | 117 | 800 | 570 | 77,5 |
| 4 | 1,7 | 60 | 2,8 | 198 | 2294 | 240 | 74,5 |
| 5 | 0,7 | 40 | 1,8 | 103 | 490 | 640 | 62,5 |
| 6 | 1,7 | 40 | 2,8 | 132 | 1289 | 270 | 67,0 |
| 7 | 0,7 | 60 | 2,8 | 132 | 1270 | 410 | 78,0 |
| 8 | 1,7 | 60 | 1,8 | 139 | 1090 | 380 | 70,0 |
| 9 | 0,3835 | 50 | 2,3 | 102 | 770 | 590 | 76,0 |
| 10 | 2,0165 | 50 | 2,3 | 154 | 1690 | 260 | 70,0 |
| 11 | 1,2 | 33,67 | 2,3 | 96 | 700 | 520 | 63,0 |
| 12 | 1,2 | 66,33 | 2,3 | 163 | 1540 | 380 | 75,0 |
| 13 | 1,2 | 50 | 1,4835 | 116 | 2184 | 520 | 65,0 |
| 14 | 1,2 | 50 | 3,1165 | 153 | 1784 | 290 | 71,0 |
| 15 | 1,2 | 50 | 2,3 | 133 | 1300 | 380 | 70,0 |
| 16 | 1,2 | 50 | 2,3 | 133 | 1300 | 380 | 68,5 |
| 17 | 1,2 | 50 | 2,3 | 140 | 1145 | 430 | 68,0 |
| 18 | 1,2 | 50 | 2,3 | 142 | 1090 | 430 | 68,0 |
| 19 | 1,2 | 50 | 2,3 | 145 | 1260 | 390 | 69,0 |
| 20 | 1,2 | 50 | 2,3 | 142 | 1344 | 390 | 70,0 |

In diesem Beispiel soll der Abriebindex möglichst groß (mindestens 120, optimal > 170), der 200%-Modul möglichst groß (mindestens 1000, optimal > 1300), die Bruchdehnung möglichst nahe bei 500 (zwischen 400 und 600) und die Härte möglichst nahe bei 67,5 (zwischen 60 und 75) sein. Tabelle 15.7 zeigt die Fixpunkte der Wunschfunktion und die allgemeine Form gemäß Bild 15-3.

| Erwünschtheit | 0 | 1 | 0 | Form |
|------------------|-------------|-------------|------------|------|
| für Abriebindex | ≤ 120 | ≥ 170 | | a |
| für 200%-Modul | ≤ 1000 | ≥ 1300 | | a |
| für Bruchdehnung | ≤ 400 | $= 500$ | ≥ 600 | c |
| für Härte | ≤ 60 | $= 67,5$ | ≥ 75 | c |

TABELLE 15.7

Festlegung der Fixpunkte und Form der Wunschfunktionen gemäß Bild 15-3:

Da bei diesem Beispiel für die beiden Zielgrößen Bruchdehnung und Härte Zielwerte vorgegeben sind, ist es mit grafischen Darstellungen allein fast unmöglich, das Optimum zu finden, zumal es nicht möglich ist, die Optimalwerte der Zielgrößen gleichzeitig zu erreichen. Mit Hilfe der Erwünschtheit erhält man dagegen das Optimum in Tabelle 15.8.

Dieses Optimum ist ein Kompromiss zwischen den widersprüchlichen Zielen. Nur für den 200%-Modul wurde der optimale Wert erreicht (Erwünschtheit 1). Bei den anderen Zielgrößen wurden mehr oder weniger große Abstriche gemacht. Insbesondere beim Abriebindex wird mit 129 nur ein Wert erreicht, der wenig über dem Mindestwert von 120 liegt. Die Erwünschtheit beträgt nur $(129 - 120)/(170 - 120) = 0,18$. Aber eine Verbesserung dieses Wertes wäre mit einer erheblichen Verschlech-

terung der Bruchdehnung und der Härte verbunden. So würde eine Erhöhung des Abriebindex auf 134 z.B. mit einer Bruchdehnung von 450 und einer Härte von 69,7 und damit einer Gesamt-Erwünschtheit von $(0,28 \cdot 1 \cdot 0,50 \cdot 0,71)^{1/4} = 0,56$ erkauft.

| | Optimale Werte | besserer Abrieb |
|-------------------------|----------------|-----------------|
| Silica | 1,17 | 1,17 |
| Silan | 51,5 | 54,5 |
| Schwefel | 1,87 | 1,93 |
| Abriebindex | 129 | 134 |
| 200 % Modul | 1300 | 1304 |
| Bruchdehnung | 466 | 450 |
| Härte | 68 | 69,7 |
| Erwünschtheit Abrieb | 0,18 | 0,28 |
| 200 % Modul | 1,0 | 1,0 |
| Bruchdehnung | 0,66 | 0,50 |
| Härte | 0,93 | 0,71 |
| Gesamt | 0,58 | 0,56 |

TABELLE 15.8

Ergebnisse im Überblick:
Optimale Faktorstufenkombination und Vergleichswert mit besserem Abriebindex, aus den Regressionsmodellen errechnete Werte der Zielgrößen und deren Erwünschtheiten, sowie die Gesamt-Erwünschtheit

Mit Hilfe der Erwünschtheit kann man auch bei mehreren Zielgrößen ohne großen Aufwand ein gemeinsames Optimum berechnen. Voraussetzung ist der Einsatz einer geeigneten Software. Da die Folgen der Zusammenfassung der Zielgrößen zur Erwünschtheit nicht offensichtlich sind, wird empfohlen, verschiedene Fixpunkte für die Wunschfunktionen zu verwenden und die Ergebnisse zu vergleichen.

Eine nützliche Ergänzung ist die Optimierungsgrafik Bild 15.4. Sie zeigt die vier Zielgrößen im Beispiel Gummimischung in Abhängigkeit von den drei Faktoren einzeln. Die anderen Faktoren werden jeweils auf dem oben angegebenen Wert festgehalten. Für die Darstellung wurden die optimalen Werte aus Tabelle 15.8 verwendet. Die Grafik zeigt mit einem Blick sämtliche Abhängigkeiten – welche Zielgröße hängt wie stark und auf welche Art von jedem Faktor ab. Da alle Abhängigkeiten eindimensional dargestellt werden, können sogar Vertrauensbereiche oder Vorhersagebereiche eingezeichnet werden.

Diese Grafik wird (unter verschiedenen Namen) in vielen Versuchsplanungsprogrammen angeboten. Die Einstellwerte für die Faktoren können interaktiv mit der Maus verändert werden und die anderen Darstellungen werden sofort mit verändert. Diese Darstellung ist sehr gut geeignet, um die Optima aus der Wunschfunktion noch weiter den Bedürfnissen anzupassen oder ohne Wunschfunktion von Anfang an zu optimieren.

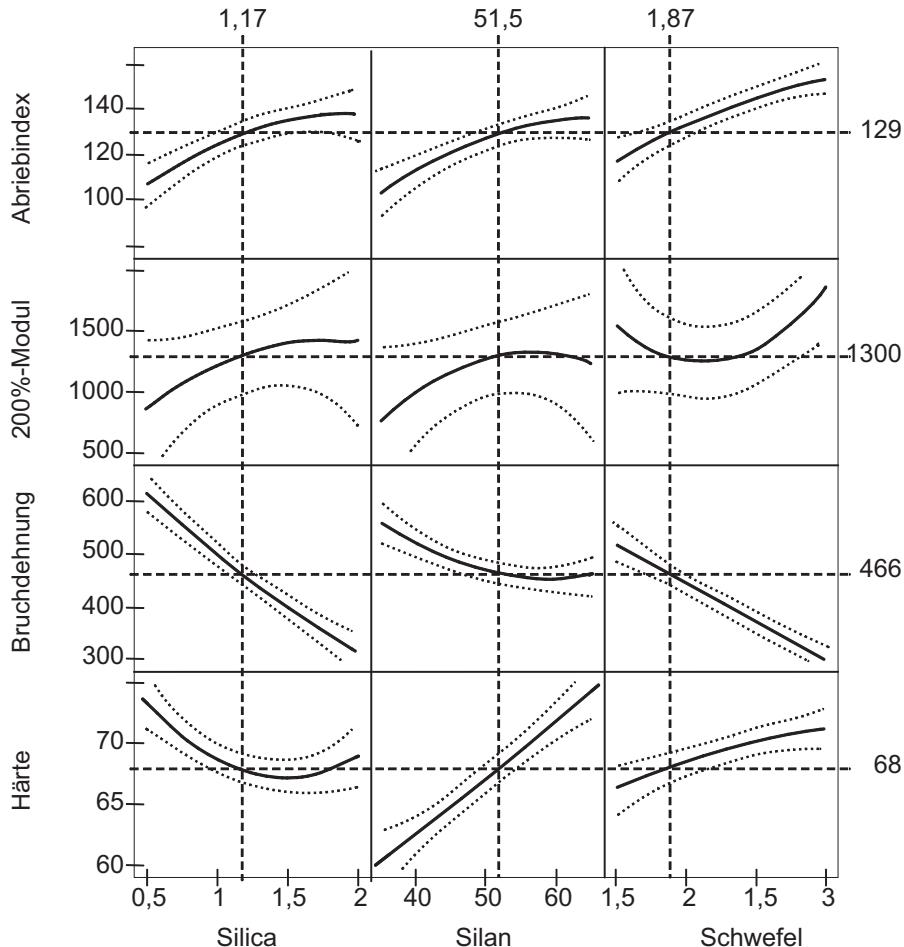


BILD 15.4 Optimierungsgrafik

Grafische Darstellung aller Zielgrößen gegen alle Faktoren im Beispiel Gummimischung:
Die anderen Faktoren werden jeweils auf dem oben angegebenen Wert festgehalten (d. h. die erste Spalte zeigt die Abhängigkeit der Zielgrößen von Silica, wenn Silan = 51,5 und Schwefel = 1,87). Die durchgezogenen Linien zeigen die Schätzwerte, die gestrichelten Linien den 95%-Vertrauensbereich.

Hinweise zur Modellverbesserung

Wie in [4] wurden für die Optimierung und grafische Darstellung der Ergebnisse an alle Zielgrößen volle quadratische Modelle angepasst. So sind die Ergebnisse leicht in verschiedener Software nachvollziehbar. Allerdings sind dies nicht die bestmöglichen Modelle. Daher eignet sich dieses Beispiel auch gut, um verschiedene Software (vgl. Kapitel 19) in Handhabung und Funktionalität miteinander zu vergleichen. Dabei empfehle ich auf folgende Punkte zu achten:

- Liefert die Optimierung mit den Originaldaten und vollen quadratischen Modellen die Ergebnisse in Tabelle 15.8? Ist die Optimierungsgrafik Bild 15.4 möglich? Mit oder nur ohne Vertrauensbereiche?

- Werden nicht signifikante Terme aus dem Modell entfernt, so verbessert sich die Prognosefähigkeit der Modelle (PRESS bzw. r^2_{pred}). Wie leicht ist es, diese Terme zu erkennen und zu entfernen? Gibt es schrittweise Regression? Ist es möglich, für die vier Zielgrößen unterschiedliche Modelle anzupassen und trotzdem zu optimieren?
- Beim 200%-Modul ist der Wert 2184 für syst. Nr. 13 ein Ausreißer oder zumindest ein sehr einflussreicher Punkt. Ohne diesen Punkt sind Lack-of-Fit und r^2_{pred} deutlich verbessert. Finden Sie dies mit den diagnostischen Möglichkeiten der verwendeten Software? Können Sie den Wert deaktivieren oder müssen Sie ihn löschen? Bleiben nach dem Deaktivieren/Löschen jetzt nicht mehr aktuelle Grafiken oder Auswertungen sichtbar (das könnte zu Verwechslungen führen)?
- Für die Zielgröße Bruchdehnung führt eine Wurzeltransformation zu einer weiteren Verbesserung des Modells. Finden Sie dies mit den diagnostischen Möglichkeiten der verwendeten Software? Werden die angepassten Modelle dann trotz Transformation im Originalmaßstab dargestellt?

Literatur

- [1] *Graf, U./Henning, H.-J./Stange, K./Wilrich, P.-T.*: Formeln und Tabellen der angewandten mathematischen Statistik. 3. Aufl. Springer Verlag, Berlin 1998
- [2] *Taguchi, G./Wu, Y.*: Introduction to Off-line Quality Control. Central Japan Quality Control Association, Nagoya 1985
- [3] *Fahrmeir, L./Kneib, T./Lang, S.*: Regression: Modelle, Methoden und Anwendungen. Springer Verlag, Berlin, 2. Auflage 2009
- [4] *Derringer, G./Suich, R.*: „Simultaneous Optimization of Several Response Variables“ in: Journal of Quality Technology 12 (Oct. 1980), 214 – 219

16

Erweiterung von Versuchsplänen

In diesem Kapitel wird gezeigt, wie Versuchspläne schrittweise erweitert werden können („augment“). Der Versuchsraum (d. h. der Bereich, in dem die Faktoren untersucht werden) bleibt unverändert – er wird jedoch genauer ausgeleuchtet.

Bei fraktionellen faktoriellen Plänen können Vermengungen gezielt beseitigt werden und ausgehend von zweistufigen Plänen können auch quadratische Effekte erfasst werden. Außerdem wird gezeigt, wie nicht realisierbare Faktorstufenkombinationen ersetzt werden können.

■ 16.1 Trennung vermengter Wechselwirkungen

Vorteil der Screening-Versuchspläne aus Kapitel 8 ist die geringe Anzahl Faktorstufenkombinationen und damit der geringe Versuchsaufwand. Ihr Nachteil ist jedoch, dass Effekte von Faktoren mit Zwei-Faktor-Wechselwirkungen (2FWW) vermengt sind (bei Auflösung III) oder 2FWW miteinander vermengt sind (bei Auflösung IV).

Screening-Versuchspläne sind gut dazu geeignet, zunächst mit geringem Aufwand zu erkennen, welche Faktoren wichtig sind. Zeigt sich dabei, dass durch Veränderung der Faktorstufen in eine bestimmte Richtung noch Verbesserungspotential besteht, wird man zunächst in Richtung der erwarteten Verbesserung fortschreiten – z. B. mit der Methode des steilsten Anstiegs aus Abschnitt 18.2. Zeigt sich jedoch, dass man sich bereits in der Nähe des Optimums befindet, oder ist die Richtung der Verbesserung aufgrund der Vermengung nicht klar, so ist es sinnvoll, ohne Änderung der Faktorstufen weitere Versuche durchzuführen, um vermengte Effekte zu trennen.

In Absatz 8.3.2 wurde am Beispiel eines Plackett-Burman-Planes gezeigt, wie durch ein Foldover – einer Wiederholung aller Faktorstufenkombinationen mit entgegengesetzten Stufenwerten – aus einem Plan der Auflösung III ein Plan der Auflösung IV entsteht (Tabelle 8.19 im Vergleich zu Tabelle 8.17). Diese Erweiterungsmöglichkeit besteht bei allen Versuchsplänen der Auflösung III – durch die Verdoppelung der Anzahl der Faktorstufenkombinationen entsteht ein Plan der Auflösung IV. Bei einer nachträglichen Erweiterung ist jedoch wichtig, dass die erste und die zweite Versuchsreihe als getrennte Blöcke behandelt werden.

Wurde, wie in Abschnitt 8.6 empfohlen, von Anfang an ein Plan der Auflösung IV verwendet, so sind 2FWW miteinander vermengt. Aus den Daten kann man dann zwar die Summe der vermengten Wechselwirkungen bestimmen, aber nicht deren Aufteilung. Manchmal

erlauben technische Überlegungen die Unterscheidung, wie im Beispiel in Absatz 8.2.7. Ansonsten sind zusätzliche Versuche erforderlich, wenn die Trennung wichtig ist.

Zur Trennung von 2FWW ist der „Foldover on one factor“ geeignet [1]: Dazu werden alle Faktorstufenkombinationen wiederholt, wobei bei **einem** der Faktoren die Stufenwerte vertauscht werden. Dadurch kann man alle 2FWW dieses Faktors unvermengt bestimmen. Es ist sinnvoll, für diesen Foldover den Faktor zu verwenden, der vermutlich die größten Wechselwirkungen hat. Wie beim normalen Foldover („on all factors“) wird die zweite Versuchsreihe als getrennter Block behandelt.



Beispiel

Ein 2^{7-3} -Versuchsplan der Auflösung IV besteht aus 16 Faktorstufenkombinationen. Tabelle 16.1 zeigt den Plan mit den Zuordnungen E = ABC, F = BCD und G = ACD gemäß Tabelle 8.14 in systematischer Reihenfolge (Block 1). Bei diesem Plan sind folgende 2FWW miteinander vermengt, durch „Foldover on factor A“ (Block 2) werden die 2FWW von Faktor A abgetrennt:

| nur Block 1 | | Block 1 + Block 2 | |
|--------------|---------------------|-------------------|--------------|
| AB + CE + FG | wird getrennt in | AB | und CE + FG |
| AC + BE + DG | wird getrennt in | AC | und BE + DG |
| AD + CG + EF | wird getrennt in | AD | und CG + EF |
| AE + BC + DF | wird getrennt in | AE | und BC + DF |
| AF + BG + DE | wird getrennt in | AF | und BG + DE |
| AG + BF + CD | wird getrennt in | AG | und BF + CD |
| BD + CF + EG | bleibt vermengt als | | BD + CF + EG |

TABELLE 16.1 Erweiterung eines 2^{7-3} -Planes (Block 1) durch „Foldover on factor A“ (Block 2) – die Vorzeichen in Spalte A sind vertauscht, der Rest ist unverändert

| Nr. | Block | A | B | C | D | E | F | G |
|-----|-------|---|---|---|---|---|---|---|
| 1 | 1 | – | – | – | – | – | – | – |
| 2 | 1 | + | – | – | – | + | – | + |
| 3 | 1 | – | + | – | – | + | + | – |
| 4 | 1 | + | + | – | – | – | + | + |
| 5 | 1 | – | – | + | – | + | + | + |
| 6 | 1 | + | – | + | – | – | + | – |
| 7 | 1 | – | + | + | – | – | – | + |
| 8 | 1 | + | + | + | – | + | – | – |
| 9 | 1 | – | – | – | + | – | + | + |
| 10 | 1 | + | – | – | + | + | + | – |
| 11 | 1 | – | + | – | + | + | – | + |
| 12 | 1 | + | + | – | + | – | – | – |
| 13 | 1 | – | – | + | + | + | – | – |
| 14 | 1 | + | – | + | + | – | – | + |
| 15 | 1 | – | + | + | + | – | + | – |
| 16 | 1 | + | + | + | + | + | + | + |

| Nr. | Block | A | B | C | D | E | F | G |
|-----|-------|---|---|---|---|---|---|---|
| 17 | 2 | + | – | – | – | – | – | – |
| 18 | 2 | – | – | – | – | + | – | + |
| 19 | 2 | + | + | – | – | + | + | – |
| 20 | 2 | – | + | – | – | – | + | + |
| 21 | 2 | + | – | + | – | + | + | + |
| 22 | 2 | – | – | + | – | – | + | – |
| 23 | 2 | + | + | + | – | – | – | + |
| 24 | 2 | – | + | + | – | + | – | – |
| 25 | 2 | + | – | – | + | – | + | + |
| 26 | 2 | – | – | – | + | + | + | – |
| 27 | 2 | + | + | – | + | + | – | + |
| 28 | 2 | – | + | – | + | – | – | – |
| 29 | 2 | + | – | + | + | + | – | – |
| 30 | 2 | – | – | + | + | – | – | + |
| 31 | 2 | + | + | + | + | – | + | – |
| 32 | 2 | – | + | + | + | + | + | + |

Eine Alternative zum hier beschriebenen Foldover ist die Erweiterung des Versuchsplans als (D-)optimaler Plan (vgl. Absatz 11.2.4).

Dazu werden als Modell mindestens alle vermuteten 2FWW angegeben. Noch sicherer ist es, einfach alle 2FWW anzugeben – dann werden aber auch mehr zusätzliche Faktorstufenkombinationen benötigt. Die bereits durchgeführten Versuche werden als „inclusions“ behandelt, d. h. sie sind auf jeden Fall im neuen Plan enthalten. Die Software wählt (normalerweise aus einem vollständigen faktoriellen Plan) die Faktorstufenkombinationen aus, die die vorhandenen Punkte möglichst gut ergänzen.

Diese Vorgehensweise ist sehr flexibel. Man erhält jedoch einen Plan, der nicht orthogonal ist. Das bedeutet, dass die berechneten Schätzwerte für die Effekte davon abhängen, welche anderen Terme im Modell enthalten sind.

■ 16.2 Zentrumspunkt

Mit 2^{k-p} - und 2^{k-p} -Plänen können nur lineare Effekte und Wechselwirkungen ermittelt werden. Bei quantitativen Faktoren, die auch Zwischenwerte annehmen können, wird empfohlen, den faktoriellen Plan um einen Versuchspunkt in der Mitte des untersuchten Bereichs zu ergänzen, im Zentrum des Würfels in Bild 16-1.

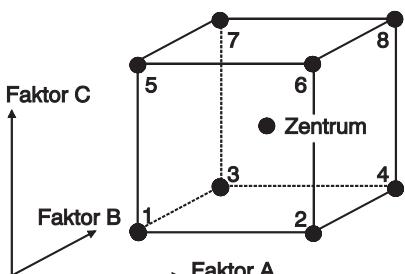


BILD 16-1
Faktorstufenkombinationen eines 2^3 -Plans mit Zentrumspunkt

Vorteile des Zentrumspunktes

- Mit diesem Zentrumspunkt können Abweichungen von der Linearität erkannt werden – sie können allerdings nicht einem bestimmten Faktor zugeordnet werden, siehe dazu Abschnitt 16.3.
- Der Zentrumspunkt kann mehrfach realisiert werden und so auch bei Versuchsplänen mit $n = 1$ ein Maß für die Zufallsstreuung liefern.
- Häufig ist das Zentrum die bisherige Einstellung und erlaubt so als Referenzpunkt den direkten Vergleich.
- Wird der Zentrumspunkt in regelmäßigen Abständen im Versuchsplan wiederholt, so kann man aus einer systematischen Veränderung der Ergebnisse auf einen Trend schließen und die Daten ggf. korrigieren.

Wichtig ist, dass die Versuche im Zentrum gemeinsam mit den anderen Versuchen durchgeführt und nicht erst später nachgeholt werden. Nur so kann zwischen einer Abweichung von der Linearität und einer Veränderung mit der Zeit unterschieden werden. Ist der Versuch in mehrere Blöcke unterteilt, so sollte jeder Block Zentrumspunkte enthalten.



Beispiel aus der chemischen Industrie (Ergänzung)

Im Beispiel aus der chemischen Industrie aus Kapitel 3 wurde, zusätzlich zu den Versuchen in Tabelle 3.3, der Punkt im Zentrum des untersuchten Bereichs (syst. Nr. 9 bei Temperatur 130°C, Zeit 3h und Katalysator 0,3%) in jedem Block zweimal realisiert. Tabelle 16.2 zeigt die Ergebnisse.

TABELLE 16.2 Versuchsergebnisse für Versuch mit Zentrumspunkt (in Tabelle 3.3 waren die Einzelversuche mit syst. Nr. 9 aus didaktischen Gründen weggelassen worden)

| Vers. Nr. | syst. Nr. | Realisierung (Block) | Temperatur [°C] | Zeit [h] | Katalysator [%] | Ausbeute [%] |
|-----------|-----------|----------------------|-----------------|----------|-----------------|--------------|
| 1 | 9 | 1 | 130 | 3 | 0,3 | 62,8 |
| 2 | 8 | 1 | 140 | 4 | 0,5 | 68,5 |
| 3 | 3 | 1 | 120 | 4 | 0,1 | 56,7 |
| 4 | 1 | 1 | 120 | 2 | 0,1 | 52,8 |
| 5 | 7 | 1 | 120 | 4 | 0,5 | 56,5 |
| 6 | 9 | 1 | 130 | 3 | 0,3 | 60,5 |
| 7 | 6 | 1 | 140 | 2 | 0,5 | 62,2 |
| 8 | 2 | 1 | 140 | 2 | 0,1 | 61,5 |
| 9 | 4 | 1 | 140 | 4 | 0,1 | 67,9 |
| 10 | 5 | 1 | 120 | 2 | 0,5 | 53,6 |
| 11 | 4 | 2 | 140 | 4 | 0,1 | 70,2 |
| 12 | 9 | 2 | 130 | 3 | 0,3 | 62,9 |
| 13 | 3 | 2 | 120 | 4 | 0,1 | 55,2 |
| 14 | 8 | 2 | 140 | 4 | 0,5 | 67,2 |
| 15 | 5 | 2 | 120 | 2 | 0,5 | 54,1 |
| 16 | 1 | 2 | 120 | 2 | 0,1 | 54,1 |
| 17 | 6 | 2 | 140 | 2 | 0,5 | 62,9 |
| 18 | 7 | 2 | 120 | 4 | 0,5 | 54,6 |
| 19 | 2 | 2 | 140 | 2 | 0,1 | 61,8 |
| 20 | 9 | 2 | 130 | 3 | 0,3 | 61,8 |

Die zusätzlichen Punkte haben keinen Einfluss auf die berechneten Effekte. Sie gehen aber in die Zufallsstreuung ein und können benutzt werden, um eine Abweichung von der Linearität zu erkennen. Dazu wird der Mittelwert der Versuchsergebnisse im Zentrum verglichen mit dem Mittelwert aller anderen Ergebnisse:

$$\text{Mittelwert Zentrum} = \frac{1}{4}(62,8 + 60,5 + 62,9 + 61,8) = 62,0$$

$$\text{Mittelwert Rest} = \frac{1}{16}(68,5 + 56,7 + 52,8 + \dots + 61,8) = 59,9875$$

Die Differenz der beiden Mittelwerte von 2,0125 % ist ein Maß für die Abweichung von der Linearität (Lack-of-Fit), dessen Signifikanz statistisch beurteilt werden kann (ähnlich zu Abschnitt 6.3, aber unter Berücksichtigung der unterschiedlichen Stichprobenumfänge). Die hier beobachtete Differenz ist signifikant ($p = 0,005$).

Dies bedeutet, dass die Ausbeute von mindestens einem der Faktoren nichtlinear abhängt. Aufgrund dieses Ergebnisses allein kann man jedoch nicht unterscheiden, ob die Nichtlinearität bei Temperatur, Zeit, Katalysator, oder mehreren dieser Faktoren auftritt (natürlich kann man hier vermuten, dass die Abhängigkeit von der Zeit aufgrund eines Sättigungsverhaltens nichtlinear ist, aber das ist technisches Zusatzwissen).

Bild 16-2 zeigt eine Möglichkeit zur grafischen Darstellung der Abweichung von der Linearität. Die Abweichung des zusätzlichen Punktes von der Mitte ist ein Maß für die Nichtlinearität.

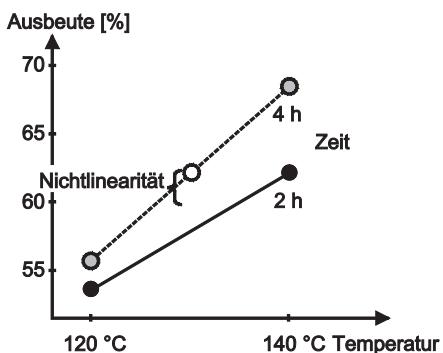


BILD 16-2

Mittelwert der Ergebnisse im Zentrum (62%) liegt um ca. 2 % über dem Mittelwert aller Ergebnisse – ein Maß für die Nichtlinearität (vgl. Bild 3-3)

■ 16.3 Zuordnung quadratischer Effekte

Der Zentrumspunkt zeigt, ob es quadratische Effekte gibt, aber er erlaubt keine Zuordnung zu einem bestimmten Faktor. Für diese Zuordnung sind zusätzliche Versuche erforderlich.

Hatte der faktorielle Plan mindestens Auflösung V, so kann er um den „Stern“ erweitert werden, wie z.B. in Tabelle 11.1 dargestellt. Dadurch entsteht ein zentral zusammengesetzter Versuchsplan.



Beispiel aus der chemischen Industrie (Ergänzung 2)

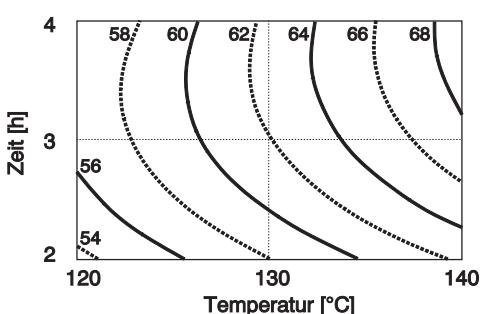
Für das Beispiel aus der chemischen Industrie wurde in Abschnitt 16.2 eine signifikante Abweichung von der Linearität festgestellt. Tabelle 16.3 zeigt die Sternversuche, mit denen der Versuchsplan aus Tabelle 16.2 nachträglich zu einem zentral zusammengesetzten Plan erweitert werden kann. Wie in Tabelle 16.2 wurde jede Faktorstufenkombination im Stern zweimal in getrennten Blöcken in randomisierter Reihenfolge realisiert, der Punkt im Zentrum des untersuchten Bereichs (syst. Nr. 9) wurde in jedem Block zweimal realisiert.

TABELLE 16.3 Versuchsergebnisse für Sternpunkte (zusätzlich zu den Versuchen in Tabelle 16.2)

| Vers. Nr. | syst. Nr. | Block | Temperatur [°C] | Zeit [h] | Katalysator [%] | Ausbeute [%] |
|-----------|-----------|-------|-----------------|----------|-----------------|--------------|
| 21 | 9 | 3 | 130 | 3 | 0,3 | 61,9 |
| 22 | 11 | 3 | 140 | 3 | 0,3 | 68,2 |
| 23 | 12 | 3 | 130 | 1,6 | 0,3 | 55,3 |
| 24 | 13 | 3 | 130 | 4,4 | 0,3 | 62,8 |
| 25 | 14 | 3 | 130 | 3 | 0,1 | 61,7 |
| 26 | 15 | 3 | 130 | 3 | 0,5 | 62,1 |
| 27 | 10 | 3 | 120 | 3 | 0,3 | 56,4 |
| 28 | 9 | 3 | 130 | 3 | 0,3 | 61,9 |
| 29 | 14 | 4 | 130 | 3 | 0,1 | 63,0 |
| 30 | 11 | 4 | 140 | 3 | 0,3 | 68,6 |
| 31 | 9 | 4 | 130 | 3 | 0,3 | 61,9 |
| 32 | 13 | 4 | 130 | 4,4 | 0,3 | 62,0 |
| 33 | 15 | 4 | 130 | 3 | 0,5 | 61,1 |
| 34 | 10 | 4 | 120 | 3 | 0,3 | 55,4 |
| 35 | 12 | 4 | 130 | 1,6 | 0,3 | 55,3 |
| 36 | 9 | 4 | 130 | 3 | 0,3 | 62,5 |

Für die Faktoren Temperatur und Katalysator wurde $\alpha = 1$ verwendet, da die Anlage 140 °C nicht überschreiten kann und bei weniger als 0,1% Katalysator Instabilitäten befürchtet wurden (flächenzentriert). Für den Faktor Zeit wurde $\alpha = 1,4$ verwendet (näherungsweise orthogonal).

Für die Auswertung werden Tabellen 16.2 und 16.3 zusammen als ein zentral zusammengesetzter Plan behandelt. Wie vermutet ist die quadratische Abhängigkeit von der Zeit signifikant, die anderen quadratischen Terme sind nicht signifikant. Bild 16-3 zeigt das angepasste quadratische Modell in Höhenliniendarstellung.

**BILD 16-3**

Ausbeute in Abhängigkeit von Temperatur und Zeit (nichtlineare Abhängigkeit von der Zeit)

Deutlich ist zu erkennen, dass bei Zeiten bis knapp 3 h die Ausbeute schnell mit der Zeit zunimmt, während eine weitere Verlängerung keine wesentliche Erhöhung der Ausbeute mehr bringt. Die scheinbare Abnahme der Ausbeute mit zunehmender Zeit bei niedriger Temperatur und Zeiten um 4 h ist ein Artefakt des quadratischen Modells (vgl. Abschnitt 11.3) – in Wirklichkeit liegt ein Sättigungsverhalten vor.

Bei einer Reaktionszeit von 2,5 bis 3 h erhält man den besten Kompromiss zwischen Ausbeute und Anlagenkosten.

Vermutet man, dass bestimmte Faktoren für die Nichtlinearität verantwortlich sind, so genügen Sternpunkte in diesen Faktoren (im obigen Beispiel die Versuche mit den systematischen Nummern 9, 12 und 13). Dann kann allerdings auch nur deren Nichtlinearität ermittelt werden, alle anderen Nichtlinearitäten gehen nur in den „Lack of fit“, d.h. die Abweichung vom angepassten Modell, ein.

■ 16.4 Nicht realisierbare Faktorstufenkombinationen

Um den Versuchsplan (näherungsweise) orthogonal und damit gut auswertbar zu halten, sollten die Versuchspunkte möglichst so realisiert werden, wie sie im Versuchsplan vorgesehen sind.

Treten geringfügige Abweichungen zwischen den geplanten Sollwerten der Faktoren und den realisierten Istwerten auf, so werden bei der Auswertung immer die Istwerte verwendet, nicht die Sollwerte. Die resultierende Abweichung von der Orthogonalität wirkt sich weniger aus als eine evtl. denkbare Verfälschung der Ergebnisse und Erhöhung der Zufallsstreuung durch „falsche“ Faktorstufenwerte in der Auswertung.

Als Vorbeugung gegen nicht realisierbare Faktorstufenkombinationen wurde in Kapitel 3 empfohlen, in einem Pilotversuch extreme Kombinationen zu testen. Falls dabei Probleme erkannt werden, ist es meist sinnvoll, die Abstände der kritischen Faktorstufen zu reduzieren, damit alle Kombinationen realisierbar sind.

Manchmal sprechen jedoch technische Argumente gegen eine Verringerung des Stufenabstands, und manchmal werden Probleme erst bei der Versuchsdurchführung erkannt. Dann stellt sich die Frage, wie man am besten mit solchen nicht realisierbaren Faktorstufenkombinationen umgeht.

Bei qualitativen Faktoren und Auswertung mit Varianzanalyse gibt es keinen einfachen Ersatz für nicht realisierbare Faktorstufenkombinationen. Aufgrund des fehlenden Punktes kann die höchste Wechselwirkung nicht mehr berechnet werden. Bei drei und mehr Faktoren ist dies aber normalerweise kein ernstes Problem, weil diese höchste Wechselwirkung meist vernachlässigbar ist.

Bei quantitativen Faktoren und Auswertung mit Regressionsanalyse zerstören nicht realisierbare Faktorstufenkombinationen die Orthogonalität des Versuchsplans. Um die Abweichung von der Orthogonalität möglichst klein zu halten, wird empfohlen, einen oder mehrere Ersatzversuche am Rand des realisierbaren Bereiches durchzuführen. Dazu bieten sich zwei Strategien an, die in Bild 16-4 am Beispiel von zwei Faktoren grafisch dargestellt werden:

- Ersatz der nicht realisierbaren Faktorstufenkombination durch die nächstmögliche realisierbare Kombination:
Der Versuchsaufwand bleibt gleich, aber die Genauigkeit des angepassten Modells nimmt ab (insbesondere in der Nähe des nicht realisierbaren Bereiches).
- Ersatz der nicht realisierbaren Faktorstufenkombination durch mehrere benachbarte Kombinationen („extreme vertices“):

Der Versuchsaufwand nimmt zu (bei k sich gegenseitig beschränkenden Faktoren wird eine Kombination durch k neue Kombinationen ersetzt), aber die Genauigkeit des angepassten Modells wird aufgrund der größeren Anzahl Einzelversuche manchmal sogar besser.

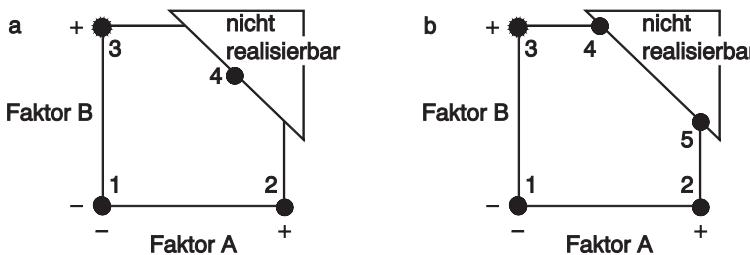


BILD 16-4 Strategien zur Auswahl von Ersatzpunkten, wenn manche Faktorstufenkombinationen nicht realisierbar sind, am Beispiel von zwei Faktoren:

- nächstmögliche realisierbare Kombination (4)
- „extreme vertices“ (4 + 5)

Eine Alternative zu dieser individuellen Modifizierung von Standardplänen von Hand sind optimale Versuchspläne (vgl. Absatz 11.2.4 und [1, 2]), bei denen von Anfang an bestimmte Bereiche von Faktorstufenkombinationen ausgenommen werden können – man spricht von „constraints“, d.h. Faktoren begrenzen sich gegenseitig (vgl. dazu auch das Beispiel in Abschnitt 20.1).

Zur Beurteilung der Qualität des modifizierten Plans eignen sich [1]:

- Varianzinflationsfaktor VIF

Faktor, um den die Varianz der Koeffizienten im Regressionsmodell größer ist als bei einem orthogonalen Plan – er sollte möglichst nahe bei 1 sein, Werte größer als ca. 4 sind problematisch

- A-, D- bzw. G-Effizienz des Versuchsplans

Maßzahlen zwischen 0 und 100 %, die angeben, wie gut der Versuchsplan im Vergleich zu einem (hypothetischen) orthogonalen Plan bezüglich des A-, D- bzw. G-Optimalitätskriteriums ist (vgl. Absatz 11.2.4) – je größer, desto besser; da mit „constraints“ 100 % jedoch nicht erreichbar sind, sind die Werte nur für den Vergleich von Alternativen geeignet, nicht als Absolutwerte.

Alternativ kann man wie in (10.19) die Standardabweichung der Regressionskoeffizienten („standard error“ der Koeffizienten) oder wie in (10.20) die Standardabweichung des angepassten Modells („standard error“ des Modells) in Abhängigkeit von der Position im Versuchraum berechnen – für eine vorgegebene Standardabweichung der Daten (meist $\sigma = 1$).

Die meisten der in Kapitel 19 besprochenen Softwarepakete bieten zumindest einige dieser Beurteilungsmöglichkeiten an.

Literatur

- [1] Myers, R. H./Montgomery, D. C./Anderson-Cook, C. M.: Response Surface Methodology. John Wiley, New York 3. Auflage 2009
- [2] Goos, P./Jones, B.: Optimal Design of Experiments – A Case Study Approach. John Wiley, Chichester 2011

Bisher wurde vorausgesetzt, dass die Daten ausreichend gut normalverteilt sind und der Wirkzusammenhang $y = f(x_1, x_2, \dots)$ im gesamten Versuchraum ausreichend gut durch ein Polynommodell (meist linear oder quadratisch) beschrieben werden kann. Mit einer Transformation der y -Werte kann die Erfüllung dieser Voraussetzungen zwar verbessert werden, es gibt aber Situationen, in denen dies nicht ausreicht (vgl. Abschnitt 11.3). Hier sollen nun kurz alternative Ansätze beschrieben werden.

■ 17.1 Andere Verteilungen: Maximum Likelihood Prinzip

Bisher wurde zur Modellanpassung die Methode der kleinsten Quadrate verwendet (Absatz 10.1.1). Für die darauf aufbauenden statistischen Berechnungen musste Normalverteilung mit konstantem σ_e vorausgesetzt werden. Für andere Verteilungen bietet das Maximum Likelihood Prinzip [1] einen systematischen Ansatz:

Die Parameter im Regressionsmodell werden so gewählt, dass die Wahrscheinlichkeit (= Likelihood) für die gerade beobachteten Daten maximal ist. Diese Wahrscheinlichkeit wird durch die Verteilung der Daten beschrieben.

Für normalverteilte Daten entspricht die Maximierung der Likelihood der Minimierung der Summe der quadrierten Abweichungen. Für andere Verteilungen erhält man automatisch die richtigen Gewichtungen. Man kann dann zwar keine geschlossenen Formeln wie in (10.4) und (10.5) für die besten Schätzwerte mehr angeben, aber mit Rechnereinsatz ist eine iterative Berechnung problemlos¹.

Generalisierte lineare Modelle (GLM) [1]

Linear bedeutet hier linear in den Parametern β , nicht in x . In diesem Sinn ist (11.1) ein lineares Modell.

Bei den generalisierten linearen Modellen können auch andere Verteilungen vorgegeben werden, wie z. B. Binomialverteilung oder Poissonverteilung. Die Modellanpassung erfolgt mit dem Maximum Likelihood Prinzip.

¹ Ein technisches Detail ist, dass Maximum Likelihood-Schätzungen der Varianz immer etwas zu klein sind (im Nenner von (6.6) erhält man n statt $(n-1)$). Mit Restricted Maximum Likelihood (REML) wird dieser Bias vermieden [1]. REML ist daher das üblicherweise verwendete Schätzverfahren.

Damit kann z. B. die logistische Funktion (15.4) auch direkt mit den Versuchsergebnissen 0 (fehlerfrei) und 1 (beschädigt) angepasst werden, ohne den Umweg über Anteile. So ist eine Auswertung auch noch möglich, wenn die x-Werte nicht gruppiert sind. Ähnlich kann ein lineares oder log-lineares Modell für poissonverteilte Zähldaten angepasst werden, u. v. a.

■ 17.2 Robuste Regression

Sowohl die Methode der kleinsten Quadrate als auch der Maximum Likelihood Ansatz reagieren empfindlich auf Ausreißer. Sie sollten daher durch Sorgfalt möglichst vermieden werden. Die nächste Barriere gegen Ausreißer sind die kritische Sichtung der Daten, Residuendarstellungen und Maßzahlen für den Einfluss von Versuchsergebnissen auf das Regressionsmodell (vgl. Abschnitt 3.6 und 10.2.4).

Trotz aller Vorsichtsmaßnahmen können Ausreißer nicht immer zuverlässig erkannt und beseitigt werden. Daher gibt es zahlreiche Ansätze, die weniger empfindlich auf Ausreißer reagieren [2–4]. So kann z. B. statt des Mittelwerts der Median der Abstandsquadrate oder der Absolutwert der Abstände minimiert werden. Oder extreme Residuen werden getrimmt.

Für Daten ohne Ausreißer sind diese Ansätze immer mit einen Verlust an Effizienz verbunden, für Daten mit Ausreißern, die nicht identifiziert und entfernt werden können, bieten sie jedoch Vorteile. Allerdings unterstützen die leicht zu bedienenden Standardprogramme in Kapitel 19 die robusten Methoden nicht, aber für R gibt es Packages zur robusten Regression (<http://www.r-project.org/>).

■ 17.3 Andere Modellansätze

Neben der Versuchsplanung stellt sich auch in vielen anderen Anwendungsgebieten das Problem, ein Modell für den Wirkzusammenhang $y = f(x_1, x_2, \dots)$ an vorhandene Daten anzupassen. Entsprechend wurde in Gebieten wie Data Mining [5], Machine Learning [6, 7] oder Metamodellierung [8] eine fast unüberschaubare Vielfalt verschiedener Modellansätze entwickelt. Nach einer allgemeinen Beschreibung des Problems von Über- und Unteranpassung werden im Folgenden ausgewählte Modellansätze kurz beschrieben.

17.3.1 Überanpassung und Unteranpassung

Je komplexer ein Modell ist, desto besser kann es sich den vorliegenden Daten anpassen. Ist das Modell zu einfach, kann es nicht alle Aspekte des tatsächlichen Wirkzusammenhangs beschreiben, man spricht von Unteranpassung. Ist das Modell jedoch zu komplex, erfassst es auch einen Teil der Zufallsstreuung in den Daten mit.

Bild 17-1 zeigt simulierte Daten, bei denen der wahre Zusammenhang quadratisch ist, mit überlagerter Zufallsstreuung. Das lineare Modell ist eine Unteranpassung – man erhält eine systematische Abweichung, genannt Bias oder Lack-of-Fit. Das quadratische Modell ist richtig. Das Polynommodell höherer Ordnung ist eine Überanpassung. Formal beschreibt es die zur Modellanpassung verwendeten Daten zwar besser (das Bestimmtheitsmaß wird größer, die Abweichung vom Modell wird kleiner). Das sagt jedoch nichts über die Prognosefähigkeit für neue Daten aus. Diese wird nur besser, solange Unteranpassung beseitigt wird. Sobald das angepasste Modell auf zufällige Besonderheiten der verwendeten Daten reagiert, wird die Prognosefähigkeit für neue Daten schlechter. Bild 17-2 zeigt dies im Prinzip.

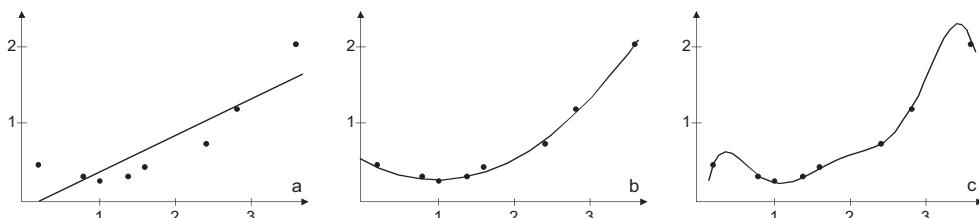


BILD 17-1 Anpassung verschiedener Polynommodelle an simulierte Daten für einen quadratischen Zusammenhang mit Zufallsstreuung

- a) lineares Modell: Unteranpassung
- b) quadratisches Modell: optimale Anpassung
- c) Polynommodell höherer Ordnung: Überanpassung

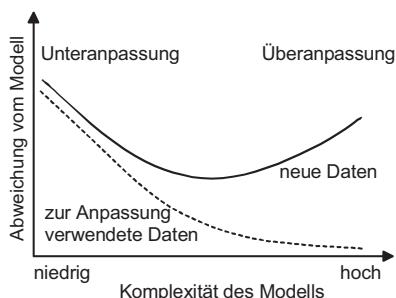


BILD 17-2

Prinzipdarstellung: Abweichung vom Modell für zur Anpassung verwendete Daten und für neue Daten gegen die Komplexität des Modells (z. B. Anzahl Parameter im Modell)

Wenn das „richtige“ Modell verwendet wird, ist die Prognosefähigkeit optimal. „Richtig“ heißt hier:

- Flexibel genug sein, um den Wirkzusammenhang zu beschreiben:
Im Folgenden werden Ansätze beschrieben, die flexibler als die Polynommodelle wie (11.1) sind und damit die in Abschnitt 11.3 beschriebenen Grenzen überwinden können.
- Überanpassung muss aber vermieden werden:
Je flexibler das Modell, desto größer ist das Risiko der Überanpassung. Überanpassung vermeiden heißt: Erkennen und Gegenmaßnahmen treffen.

Überanpassung erkennen

Die Kennzahl PRESS in Absatz 10.2.4 dient genau diesem Ziel: Jeder Punkt wird einzeln weggelassen, an die anderen Punkte wird ein Modell angepasst. Damit wird eine Prognose für den weggelassenen Punkt berechnet und mit dem beobachteten Wert verglichen. Für jeden der N Punkte muss dazu ein neues Modell berechnet werden. Für die im Folgenden beschriebenen Modelle wäre dies zu aufwändig. Stattdessen werden sie zufällig in z.B. fünf Gruppen unterteilt, an vier Gruppen wird jeweils ein Modell angepasst und die weggelassene Gruppe wird wie bei PRESS zur Beurteilung der Prognosefähigkeit verwendet. Dies erfolgt reihum, sodass insgesamt fünf Modelle angepasst werden (statt N bei PRESS). Dieses Verfahren heißt 5-fache Kreuzvalidierung (allgemein n-fach).

Gegenmaßnahmen

Bei der schrittweisen Regression (Absatz 10.2.3) werden nicht-signifikante Terme aus dem Modell entfernt. Statt der Signifikanz können zur Auswahl der Terme im Modell auch Maßzahlen wie AIC (Akaike information criterion) und BIC (Bayes information criterion) verwendet werden. AIC und BIC sind ähnlich, wobei BIC in der Regel zu Modellen mit weniger Termen führt und näherungsweise das Modell mit der größten Posteriori-Wahrscheinlichkeit auswählt [1]. Auf jeden Fall wird durch das Entfernen von Termen das Modell vereinfacht und die Prognosefähigkeit verbessert.

In Machine Learning [5–7] ist der Begriff Regularisierung verbreitet. Die Modellanpassung erfolgt als iterative Lösung einer Optimierungsaufgabe, bei der ein zusätzlicher Strafterm für Modellkomplexität eingeführt wird. Die Größe des Strafterms wird meist durch Kreuzvalidierung angepasst. Bei der Regression führt dies je nach mathematischer Form des Strafterms zu Ridge Regression oder LASSO, es gibt aber noch viele weitere Regularisierungen.

Je größer der Versuchsumfang N relativ zur Anzahl Modellparameter p ist, desto kleiner ist das Risiko einer Überanpassung. In Machine Learning wird als Daumenregel ein Verhältnis von mindestens 5 : 1, besser 10 : 1 genannt.

17.3.2 Lokale Modelle

Die Potenzen der bisher verwendeten Polynommodelle wie (11.1) wirken immer über den gesamten Versuchsraum. Für manche Anwendungen ist dies zu unflexibel, da Polynome mit höheren Potenzen zum Schwingen neigen (vgl. Bild 17-1). Ein möglicher Ausweg sind lokale Modelle. Sie können auf verschiedene Art konstruiert werden, z.B.:

- Statt der über den gesamten Versuchsraum wirkenden Potenzen in (11.1) werden nur lokal wirkende Basisfunktionen verwendet, z.B. an einem Raster über den Versuchsraum zentrierte yieldimensionale Gaussglocken, andere sogenannte Radiale Basisfunktionen (RBF) oder Wavelets. y ist dann die Summe über alle RBFs, die Koeffizienten werden wie bei der Regression angepasst [6].
- Der Versuchsraums wird in mehrere Bereiche unterteilt, in denen unterschiedliche Polynommodelle angepasst werden, z.B. Splines [5] oder Regressionsbäume [5].
- Für jeden Punkt im Versuchsraum wird ein lokales Polynommodell berechnet, für das nur die Versuchsergebnisse berücksichtigt werden, deren x-Werte in der Nähe liegen (meist mit einer Gewichtung, die mit dem Abstand abnimmt – Kernel Glättung) [5].

17.3.3 Neuronale Netze

Neuronale Netze [5, 6, 8] sind vom biologischen Nervensystem inspiriert. Jedes „Neuron“ ist ähnlich zur logistischen Funktion (15.4): Der Eingang ist eine lineare Funktion aller Faktoren bzw. aller Ausgänge der vorgelagerten Ebene, der Ausgang schaltet s-förmig zwischen zwei Niveaus um (0 und 1, oder -1 und 1, oder „nicht feuern“ und „feuern“). Bild 17-3 zeigt die allgemeine Struktur.

Faktoren versteckte Ebenen Zielgröße

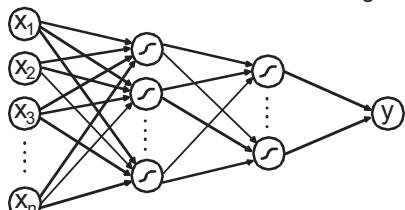


BILD 17-3

Prinzipdarstellung eines neuronalen Netzes mit zwei verdeckten Ebenen

Die Anzahl der verdeckten Ebenen (meist eine oder zwei) und die Anzahl „Neuronen“ in jeder Ebene werden vom Anwender vorgegeben. Jeder Pfeil in Bild 17-3 entspricht einem Koeffizienten β in (15.4) bzw. einem Gewicht in der bei neuronalen Netzen üblichen Bezeichnung. Für vorgegebene Gewichte kann für jede Kombination der Faktorstufenwerte x_1, x_2, \dots, x_n der Ausgang jedes „Neurons“ der 1. verdeckten Ebene berechnet werden, diese sind Eingänge der nächsten Ebene (mit neuen Gewichten), usw. So erhält man in Pfeilrichtung (vorwärts) schließlich eine Prognose für die Zielgröße y (feed forward). Aus der Abweichung vom beobachteten Wert wird rückwärts eine Korrektur für alle verwendeten Gewichte berechnet und iterativ durchgeführt (back propagation).

Aufgrund der großen Anzahl Gewichte ist das neuronale Netz sehr anpassungsfähig und neigt daher auch zur Überanpassung. Zur Regularisierung werden die Gewichte daher mit Straftermen versehen. Die Größe der Strafterme wird mit Kreuzvalidierung optimiert.²

17.3.4 Gaussian Process Models

Bei Gaussian Process Models [7, 9] wird keine bestimmte Form für den Wirkzusammenhang vorgegeben. Stattdessen wird mit einer Kovarianzfunktion die Änderung von y mit x begrenzt. Die Form der Kovarianzfunktion wird anhand von inhaltlichen Überlegungen vorgegeben. In einem Bayes-Ansatz bestimmt sie die a priori Verteilung, die dann anhand der vorliegenden Ergebnisse zu einer a posteriori Verteilung modifiziert wird, mit der neue y -Werte prognostiziert werden.

Im Normalfall nimmt die Kovarianzfunktion mit zunehmendem Abstand zweier Punkte im Versuchsraum ab (evtl. unterschiedlich schnell für verschiedene Faktoren). Je lang-

² Achtung: „Neuronale Netze“ ist ein Modebegriff und wird gelegentlich auch mit völlig anderen Inhalten belegt. Die in verschiedener Software als neuronale Netze bezeichneten Algorithmen unterscheiden sich wesentlich voneinander in Art und Eignung.

samer sie abnimmt, desto weniger können sich benachbarte y-Werte unterscheiden und desto glatter wird damit das Modell. Je schneller sie abnimmt, desto anpassungsfähiger wird das Modell – damit werden aber automatisch auch mehr Daten für die Modellanpassung benötigt und das Risiko von Überanpassung steigt.

Der Maßstab in der Kovarianzfunktion bestimmt das Verhalten des Modells. Ein passender Wert kann mit einem Likelihood-Ansatz gewählt werden. Auch zwischen verschiedenen funktionalen Formen der Kovarianzfunktion kann so unterschieden werden.

Je nach verwendeter Kovarianzfunktion können Kriging-Modelle, Kernel-Ridge-Regression oder kubische Splines als Sonderfälle von Gaussian Process Models betrachtet werden.

17.3.5 Einsatzempfehlungen

All diese Ansätze benötigen sehr viele Daten. Die erforderliche Anzahl Einzelversuche nimmt bei vorgegebener Dichte exponentiell mit der Anzahl Faktoren zu. Daher sind sie vor allem in Situationen verbreitet, in denen der Aufwand für die Durchführung einzelner Versuche relativ klein ist (oder bereits viele Daten vorliegen) und ein geschlossenes Modell über einen breiten Wertebereich der Faktoren erforderlich ist, wie z.B.

- bei der Auswertung geografischer Daten wie der Temperaturverteilung über ganz Deutschland (Faktoren sind der Längen- und der Breitengrad, evtl. zusätzlich die Höhe)
- für die Interpolation von Simulationsergebnissen (die „Versuche“ sind Rechnungen, die Parameter dieser Simulationsrechnungen können über einen breiten Bereich verändert werden)
- für die Bedatung von Steuergeräten für moderne Otto- und Dieselmotoren, bei der optimale Einspritzparameter über einen breiten Bereich von Drehzahl und Last festgelegt werden.

Als Versuchspläne werden (vor allem bei Simulationsrechnungen) Space Filling Designs verwendet [8, 10], bei denen auf eine gleichmäßige Verteilung der Punkte im Versuchsräum geachtet wird. Zur Optimierung der Verteilung wird z.B. der minimale Abstand zweier Punkte maximiert, die Entropie maximiert oder der integrierte mittlere quadratische Fehler minimiert.

Als Software gibt es diverse, oft anwendungsspezifische Programme, z.B. Zusätze für MATLAB® (<http://www.mathworks.com/>), wie die Model-Based Calibration Toolbox für die Motorenbedatung (auch kostenlose Zusätze sind erhältlich, z.B. <http://www.gaussianprocess.org/gpml/code/matlab/doc/>), oder für R (<http://www.r-project.org/>) – [11] gibt eine Einführung in Versuchsplanung mit R. ETAS ASCMO® (<http://www.etas.com/de/products/ascmo.php>) bietet eine besonders effiziente und einfach anzuwendende Umsetzung von Gaussian Process Models, mit dem Anwendungsschwerpunkt Motorbedatung. Auch manche der in Kapitel 19 beschriebenen Programme bieten neuronale Netze oder Gaussian Process Models als Alternative zu quadratischen und kubischen Modellen an. Dabei verbergen sich insbesondere hinter dem Begriff „neuronale Netze“ jedoch sehr unterschiedliche Ansätze mit entsprechend unterschiedlicher Eignung.

Aufgrund der Verschiedenheit der Anwendungen und der Komplexität der Modelle kann die Thematik hier nicht genauer behandelt werden. Für Einzelheiten wird auf die Literatur verwiesen. Die Anwendung dieser und ähnlicher Modellansätze erfordert viel Erfahrung.

Wenn „nur“ nach einem Optimum gesucht wird, ist eine pragmatische und weniger aufwändige Alternative, den Faktorbereich mit Erfahrung und Vorversuchen zunächst so weit einzuschränken, dass in diesem eingeschränkten Bereich ein quadratischer Modellansatz wie in Kapitel 11 anwendbar ist. Dazu kann es nötig sein, den Wertebereich der Faktoren schrittweise einzuschränken oder sich zunächst sequentiell an den optimalen Bereich heranzuarbeiten, wie in den Abschnitten 18.1 bis 18.3 beschrieben wird.

Literatur

- [1] Fahrmeir, L./Kneib, T./Lang, S.: Regression: Modelle, Methoden und Anwendungen. Springer Verlag, Berlin, 2. Auflage 2009
- [2] Rousseeuw, P.J./Leroy, A. M.: Robust Regression and Outlier Detection. John Wiley, Hoboken, 2. Auflage 2003
- [3] Andersen, R.: Modern Methods of Robust Regression. Sage Publications, Thousand Oaks, 2008
- [4] Huber, P.J./Ronchetti, E. M.: Robust Statistics. John Wiley, Hoboken, 2. Auflage 2009
- [5] James, G./Witten, D./Hastie, T./Tibshirani, R.: An Introduction to Statistical Learning. Springer, New York, 2013
- [6] Bishop, C. M.: Pattern Recognition and Machine Learning. Springer Verlag, New York, 2006
- [7] Murphy, K. P.: Machine Learning – a Probabilistic Perspective. MIT Press, Cambridge Mass., 2012
- [8] Siebertz, K./Bebber, D. v./Hochkirchen, T.: Statistische Versuchsplanung. Springer Vieweg, Berlin, 2. Auflage 2017
- [9] Rasmussen, C. E./Williams, C. K. I.: Gaussian Processes for Machine Learning, MIT Press, Cambridge Mass., 2006
- [10] Santner, T.J./Williams, B.J./Notz, W.I.: The Design and Analysis of Computer Experiments, Springer, New York, 2003
- [11] Lawson, J.: Design and Analysis of Experiments with R, CRC Press, Boca Raton, 2015

18

Sequentielle Optimierungsverfahren

In Kapitel 7 bis 17 wurde zunächst der Versuchsplan aufgestellt, dann wurden alle Faktorstufenkombinationen untersucht und erst danach wurden die Ergebnisse ausgewertet und Maßnahmen abgeleitet. Sequentielle Gesichtspunkte flossen nur am Rande in die Überlegungen ein: Es wurde darauf hingewiesen, dass am Anfang, wenn noch wenig bekannt ist, Screening-Versuchspläne besser geeignet sind, für die abschließende Optimierung jedoch mehrstufige Pläne sinnvoll sind.

In diesem Kapitel steht nun die sequentielle Vorgehensweise im Vordergrund. Die bisher behandelten Versuchspläne werden dabei z.T. als Bausteine eingesetzt.

Der Ausgangspunkt ist eine Faktorstufenkombination, die weit vom Optimum entfernt ist, wie in Bild 18-1 beispielhaft für den einfachen Fall von zwei Faktoren und einer Zielgröße dargestellt. Bei mehr Faktoren und mehr Zielgrößen gelten dieselben Prinzipien.

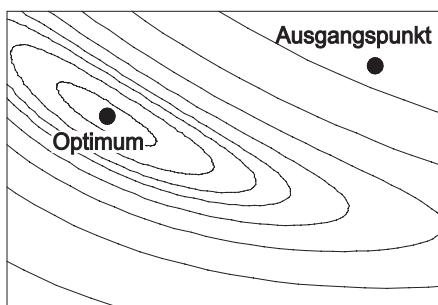


BILD 18-1

Höhenliniendiagramm einer Zielgröße in Abhängigkeit von zwei Faktoren:
Ziel der Optimierungsverfahren ist es, in möglichst wenigen Schritten vom Ausgangspunkt zum Optimum zu gelangen.

Ziel der Optimierungsverfahren ist es, in möglichst wenigen Schritten das Optimum zu finden, und gleichzeitig evtl. vorhandene Risiken zu begrenzen. Welches der folgenden Verfahren am besten geeignet ist, hängt von der Situation ab:

1. Evolutionary Operations EVOP [1] ist ein vorsichtiges Verfahren zur Optimierung der laufenden Fertigung. Zwei oder höchstens drei Prozessparameter (Faktoren) werden in einem vollständigen faktoriellen Versuchspläne innerhalb ihrer Spezifikation um den Ausgangspunkt verändert. Ziel ist, aus kleinen Veränderungen der Prozessergebnisse Wissen für eine kontinuierliche Verbesserung der Fertigung abzuleiten.
2. In der Entwicklung kann man sich größere Abstände zwischen den Faktorstufen und ein schnelleres Fortschreiten in eine optimale Richtung erlauben, auch mit dem Risiko, zunächst über das Optimum hinauszuschießen. Hier eignet sich die Methode des steilen Anstiegs [2].

3. Ist die Zufallsstreuung sehr klein im Vergleich zu den wahren Unterschieden zwischen den untersuchten Faktorstufenkombinationen, so kann man (wie in Kapitel 5) auf mehrmalige Realisierung und die Mehrfachnutzung von Ergebnissen verzichten. Dann führen Simplexverfahren [2] schnell in die Nähe des Optimums. In der Nähe des Optimums versagen sie allerdings, dort eignen sich die Pläne aus Kapitel 11 als Ergänzung.
4. In neuen Ansätzen zur Optimierung [3, 4] versucht man die Vorteile verschiedener Verfahren miteinander zu verbinden.

Im Folgenden wird die Vorgehensweise für den Fall beschrieben, dass die Zielgröße maximiert werden soll. Mit denselben Strategien ist natürlich auch eine Minimierung möglich, wenn man immer in Richtung abnehmender statt zunehmender Werte fortschreitet.

■ 18.1 Evolutionary Operations (EVOP)

Fertigungsverfahren werden normalerweise zunächst in einer Pilotlinie entwickelt und optimiert. Danach werden sie in die Produktionslinie übernommen. Das Optimum für die Prozessparameter (z.B. Druck, Temperatur, Durchflussrate, Konzentration) in der Pilotlinie muss nicht mit dem Optimum in der Produktionslinie übereinstimmen.

In der Produktion konzentriert man sich dann häufig auf die Herstellung des Produktes. Versuche werden nicht mehr durchgeführt, weil man davon ausgeht, dass die Prozessparameter bereits in der Pilotlinie optimiert wurden und weil Versuche die Produktion stören könnten.

Der Ansatz von EVOP ist nun, dass die Produktion nicht nur das Produkt selbst herstellen sollte, sondern gleichzeitig Wissen für eine ständige Verbesserung der Produktion. Dazu geht man wie folgt vor:

- Zwei (maximal drei) Prozessparameter werden ausgewählt und während der laufenden Produktion in einem vollständigen faktoriellen Versuchsplan verändert (mit Zentrum zusätzlich, Bild 18-2).
- Der Stufenabstand wird so klein gewählt, dass die Produktion nicht nennenswert gestört werden kann. Als Stufenwerte werden normalerweise die obere und untere Grenze der momentan gültigen Spezifikation verwendet.
- Aufgrund des kleinen Stufenabstands sind auch die Effekte klein. Da der Versuchsumfang jedoch groß ist (die gesamte Produktion), können auch kleine Effekte erkannt werden.
- Die gesamte Produktion durchläuft zyklisch die Faktorstufenkombinationen des Versuchsplans, bis signifikante Effekte gefunden wurden oder bis die Entscheidung gefällt wird, mit anderen Prozessparametern oder anderen Stufen fortzufahren.
- Der Produktionsleiter entscheidet darüber,
 - welche Prozessparameter auf welchen Stufen verändert werden,
 - wann er ausreichend überzeugt ist, dass durch eine Änderung der Spezifikation eine Prozessverbesserung erreicht wird, oder
 - wann er meint, man solle EVOP mit anderen Prozessparametern oder anderen Stufen fortsetzen.

Er wird dabei von einem Team beraten, in dem auch ein Statistikexperte sitzt, aber der Produktionsleiter entscheidet und trägt die Verantwortung. Er gibt also keine Entscheidungskompetenz aus der Hand, er erhält nur zusätzliche Information als Grundlage für seine Entscheidungen.

Bild 18-2 zeigt den Versuchsplan. Die fünf Faktorstufenkombinationen werden durchlaufen, bis man erkennt, in welche Richtung das Prozessergebnis besser wird – die Auswertung erfolgt mit den Verfahren aus Kapitel 7. Die Spezifikation wird dann in die Richtung der Verbesserung verschoben.

Dabei gibt es keine feste Regel, wie weit verschoben werden soll; die Entscheidung liegt beim Produktionsleiter. Ist er vorsichtig, so wird er, ausgehend von der Situation in Bild 18-2, den Punkt 2 zum neuen Mittelpunkt und den Punkt 1 zum neuen Eckpunkt machen. Aber auch eine weitere Verschiebung ist möglich.¹

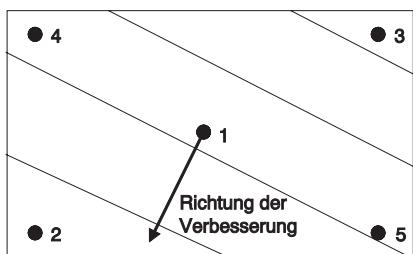


BILD 18-2

Die Faktorstufenkombinationen 1–5 werden durchlaufen, bis erkennbar wird, in welche Richtung sich das Prozessergebnis verbessert. In diese Richtung wird die Spezifikation verschoben.

Probleme beim Einsatz in der laufenden Fertigung

In der Praxis hat EVOP bisher nicht die Verbreitung gefunden, die man erwarten könnte. Dies liegt z.T. daran, dass nicht jeder Fertigungsprozess gleich gut für den Einsatz von EVOP geeignet ist, z.T. liegt es aber auch daran, dass gewisse Ängste überwunden werden müssen. Grundlage der Probleme und Ängste ist, dass jede Veränderung eines Fertigungsprozesses Risiken beinhaltet. Wichtig für die Akzeptanz von EVOP ist, dass diese Risiken überschaubar bleiben. Das bedeutet für den Einsatz von EVOP:

- Ein Fertigungsprozess ist umso besser geeignet für den Einsatz von EVOP, je kürzer die Zeit zwischen der Veränderung der Prozessparameter und der Messung der Ergebnisse ist. Ungeeignet sind Prozesse mit einer langen Zeitverzögerung (langer Leadtime), wie z.B. die IC-Fertigung (die Durchlaufzeit eines Loses beträgt mehrere Wochen und erst am Ende sind die Folgen einer Prozessveränderung erkennbar), oder wenn die Zuverlässigkeit eine wesentliche Zielgröße ist.
- Jede Prozessveränderung ist mit Risiken verbunden. Daher vermeidet die Fertigung normalerweise Änderungen soweit möglich. EVOP dagegen sieht kleine Prozessveränderungen als Chance für Verbesserungen.

¹ „EVOP – Evolutionary Operations“ simuliert die Anwendung auf einen Spritzgießprozess mit zwei Zielgrößen: Sie sind Produktionsleiter und können den Prozess verbessern – vgl. Anhang G. Die Visualisierung illustriert gleichzeitig die Methode des steilsten Anstiegs (Abschnitt 18.2).

Bild 18-3 vergleicht die Chancen und Risiken von Prozessveränderungen: Wenn der momentane Arbeitspunkt weit vom Optimum entfernt liegt, führt die Veränderung eines Faktors in die eine Richtung zwar zu einer Verschlechterung (- in Bild 18-3), in die andere Richtung aber zu einer Verbesserung (+ in Bild 18-3). Wenn das Prozessergebnis linear vom Wert des Faktors abhängt, bleibt der Mittelwert während des Versuchs unverändert, da die Verschlechterung in die eine Richtung durch die Verbesserung in die andere Richtung kompensiert wird. Bei der Risikobewertung ist vielmehr die kleine Erniedrigung des Mittelwertes aufgrund der Abweichung von der Linearität mit der großen potentiellen Verbesserung durch eine langfristige Optimierung zu vergleichen.

Um die Angst vor Prozessveränderungen zu reduzieren, ist es wichtig, dass der Fertigungsleiter über die Vorgehensweise entscheidet. Dabei berücksichtigt er die Ergebnisse der Versuche, aber auch seine Erfahrung.

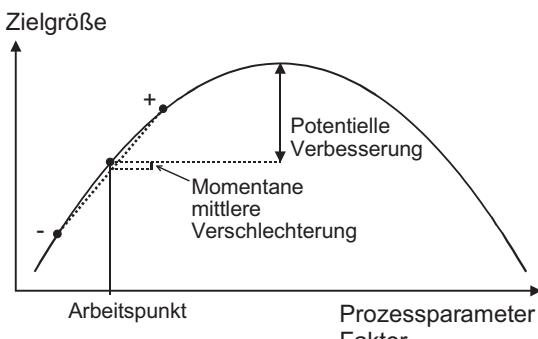


BILD 18-3

Beispiel für eine Zielgröße, die maximiert werden soll (z. B. Schlagzähigkeit, Ausbeute): Aus der Nichtlinearität ergibt sich im Versuch eine kleine momentane mittlere Verschlechterung. Sie ist mit der großen potentiellen Verbesserung zu vergleichen.

■ 18.2 Methode des steilsten Anstiegs

Die Methode des steilsten Anstiegs ist in vieler Hinsicht ähnlich zu EVOP. Da der steilste Anstieg jedoch normalerweise nicht in der laufenden Fertigung gesucht wird, sondern in der Entwicklung oder Fertigungsvorbereitung, wagt man größere Schritte:

- Wie bei EVOP beginnt man mit zwei Faktorstufen. Der Abstand der Stufen ist jedoch größer, dadurch werden auch die Effekte größer – weniger Realisierungen sind erforderlich.
- Meist werden mehrere Faktoren gleichzeitig verändert. Bei fünf oder mehr Faktoren ist es sinnvoll, auch fraktionelle faktorielle Pläne zu verwenden.
- Wenn die Richtung des steilsten Anstiegs ermittelt wurde, sucht man zunächst mit Einzelversuchen immer weiter in diese Richtung, bis das Ergebnis wieder schlechter wird. In der Umgebung des Maximums wird ein neuer Versuchplan durchgeführt. Dies kann wieder ein zweistufiger Plan zur Ermittlung einer neuen Richtung des steilsten Anstiegs sein. Wenn man vermutet, dass man sich bereits in der Nähe des Maximums oder auf einem Kamm befindet, verwendet man einen Plan zur Bestimmung nichtlinearer Zusammenhänge (Kapitel 11).

Bild 18-4 zeigt die Vorgehensweise am Beispiel aus Bild 18-1. Die Vorteile der Methode des steilsten Anstiegs sind:

- Zur Ermittlung der Richtung des steilsten Anstiegs werden die Versuchspläne aus Kapitel 7, 8 und 11 eingesetzt. In diesen Plänen werden alle Versuchsergebnisse zur Ermittlung der Effekte der Faktoren genutzt. Durch diese Mehrfachnutzung erreicht man gute statistische Absicherung bei niedriger Anzahl von Einzelversuchen. Dies ist vor allem wichtig, wenn die Zufallsstreuung vergleichbar mit den Effekten ist.
- Müssen mehrere Zielgrößen betrachtet werden, so können die in Abschnitt 11.1 bzw. 15.3 beschriebenen Verfahren zum Auffinden von Kompromissen bei Zielkonflikten eingesetzt werden.

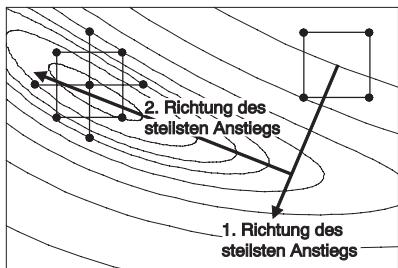


BILD 18-4

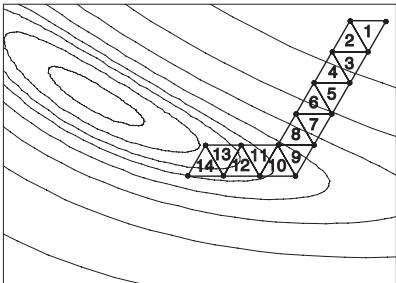
Methode des steilsten Anstiegs:
Zur Ermittlung der 1. Richtung des steilsten Anstiegs genügt ein zweistufiger Plan. Mit Einzelversuchen bewegt man sich in diese Richtung, bis die Ergebnisse wieder schlechter werden.
Zur Ermittlung der 2. Richtung (am Kamm) und des Maximums sind mehrstufige Pläne erforderlich.

■ 18.3 Simplexverfahren²

Das Simplexverfahren dient dazu, die Richtung der Verbesserung mit möglichst wenigen Faktorstufenkombinationen zu erkennen. Bei k Faktoren besteht der Simplex aus $k + 1$ Faktorstufenkombinationen. Bei zwei Faktoren sind dies die Ecken eines gleichseitigen Dreiecks, bei drei Faktoren eines Tetraeders, usw.

Bild 18-5 zeigt die Vorgehensweise am Beispiel aus Bild 18-1 (mit zwei Faktoren). Ausgehend vom Arbeitspunkt werden zunächst die Versuchsergebnisse an den Eckpunkten des Dreiecks mit der Nummer 1 bestimmt (Simplex 1). Die Faktorstufenkombination mit dem schlechtesten Ergebnis wird nun an der Linie durch die anderen Faktorstufenkombinationen gespiegelt. Dadurch erhält man eine neue Faktorstufenkombination. Diese ergibt zusammen mit den beiden besseren Kombinationen aus Simplex 1 einen neuen Simplex 2. Spiegelt man dessen schlechteste Faktorstufenkombination an der Linie durch die beiden anderen Faktorstufenkombinationen, so erhält man einen neuen Simplex 3, usw. Bei jedem neuen Versuch wird die schlechteste Faktorstufenkombination des letzten Simplex durch eine hoffentlich bessere Kombination ersetzt. So wird das Ergebnis Schritt für Schritt verbessert.

² Das hier beschriebene Simplexverfahren darf nicht mit den Simplexgitterplänen aus Kapitel 14 verwechselt werden. Bei der Untersuchung von Mischungen ergibt sich die Simplexform daraus, dass die Summe der Komponenten immer 1 sein muss. Auch Nichtlinearitäten sollen erfasst werden, daher wird meist ein Grad > 1 verwendet. Im hier beschriebenen Simplexverfahren werden nur Simplexpläne vom Grad 1 verwendet. Für unabhängige Faktoren erlauben sie die Bestimmung der linearen Abhängigkeit mit der geringsten Anzahl Faktorstufenkombinationen.

**BILD 18-5**

Simplexverfahren am Beispiel von zwei Faktoren:
Ausgehend von den drei Faktorstufenkombinationen an den Eckpunkten des Dreiecks 1 wird jeweils die Faktorstufenkombination mit dem schlechtesten Ergebnis an der Linie durch die anderen gespiegelt. Dadurch erhält man bei jedem Schritt eine Verbesserung.

Ein Vergleich der Bilder 18-4 und 18-5 zeigt, dass die Verbesserung etwa derselben Linie folgt. Insofern ist eine gewisse Ähnlichkeit der Verfahren erkennbar. Sie unterscheiden sich jedoch in einigen wesentlichen Punkten, die im Folgenden als Vorteile und Nachteile des Simplexverfahrens beschrieben werden.

Vorteile des Simplexverfahrens:

- Das Verfahren erfordert keine aufwändige Auswertung.
- Nach einem Startsimplex mit $k + 1$ Punkten schreitet die Optimierung mit jedem weiteren Einzelergebnis fort, d. h. die Optimierung ist streng sequentiell.

Nachteile des Simplexverfahrens:

- Das Verfahren funktioniert nur, wenn die Zufallsstreuung wesentlich kleiner ist als die wahren Unterschiede zwischen den Faktorstufenkombinationen. Nur dann kann die Zufallsstreuung die Ergebnisse nicht verfälschen und man kann auf eine statistische Auswertung verzichten.
- Versucht man, die Zufallsstreuung durch mehrmalige Realisierung der Faktorstufenkombinationen zu reduzieren, wird das Simplexverfahren schnell sehr ineffizient, weil es keine Mehrfachnutzung der Versuchsergebnisse erlaubt.
- Nur bei einer Zielgröße erhält man eine eindeutige Richtung der Verbesserung. Mehrere Zielgrößen müssen gemeinsam betrachtet werden. Vor allem wenn für Zielgrößen ein bestimmter Wert (Zielwert) erreicht werden soll, ist die optimale Richtung keineswegs offensichtlich.
- Das Simplexverfahren ist linear und funktioniert daher nur in den Flanken der Wirkungsflächen zuverlässig (Simplexe 1 bis 9 in Bild 18-5). Bei Kämmen und in der Nähe von Extremwerten neigt das Simplexverfahren zum Pendeln bzw. zum Kreisen zwischen einigen wenigen Punkten (in Bild 18-5 pendelt man zwischen Simplex 13 und 14 hin und her – ohne Veränderung der Regeln wird kein weiterer Fortschritt mehr erzielt).

Um Pendeln zu vermeiden, kann man ausnahmsweise den zweitschlechtesten Punkt im Simplex spiegeln. Kreist man um einen Punkt, so kann man die Abstände zwischen den Stufen verringern. Dies verringert jedoch auch die wahren Unterschiede und vergrößert damit die Bedeutung der Zufallsstreuung. Je schmäler der Kamm zum Optimum ist, desto kleiner ist der mögliche Stufenabstand.

Bei der Produkt- und Prozessverbesserung funktioniert das Simplexverfahren somit nur weit entfernt vom Optimum ohne Probleme. Dort kann es als besonders einfaches Verfahren jedoch sehr nützlich sein (solange die Zufallsstreuung vernachlässigbar ist).³

■ 18.4 Neuere Entwicklungen

Ein wesentlicher Nachteil des Simplexverfahrens ist, dass es nur die lineare Abhängigkeit der Zielgröße von den Faktoren berücksichtigt. In der Nähe des Optimums ist die Abhängigkeit jedoch nichtlinear.

Mit den Versuchsplänen von Kapitel 11 (z.B. zentral zusammengesetzte Pläne) werden auch nichtlineare Abhängigkeiten erfasst. In den Augen vieler Anwender ist allerdings ihr Nachteil, dass eine relativ große Anzahl von Einzelversuchen durchgeführt werden muss, bevor die Ergebnisse ausgewertet werden und damit eine Verbesserung erzielt wird.

Daher gibt es verschiedene Ansätze [3,4] mit dem Ziel, die sequentielle Vorgehensweise mit nichtlinearen Modellen zu vereinen. Im Prinzip kann man dies dadurch erreichen, dass man nicht nur die letzten $k+1$ Ergebnisse berücksichtigt, sondern in jedem Stadium ausreichend viele weiter zurückliegende Ergebnisse mit einbezieht und versucht, an diese Ergebnisse ein quadratisches Modell wie in (11.1) bzw. (11.2) anzupassen. Aus diesem quadratischen Modell errechnet man dann ein theoretisches Optimum. In diese Richtung schreitet man fort.

Falls das Optimum zu weit vom momentanen Zustand entfernt liegt, begrenzt man die Schrittweite, um zu große Veränderungen zu vermeiden. Würde andererseits der neue Versuchspunkt zu nahe an bereits bekannten Punkten liegen, untersucht man die Umgebung, in der Hoffnung, dort eventuell noch bessere Ergebnisse zu finden.

Von der Idee her ist diese Vorgehensweise bestechend. Die praktische Umsetzung ist jedoch m.E. noch nicht ausreichend ausgereift. Dies liegt an der Komplexität des Optimierungsproblems. Einige der Schwierigkeiten sind:

- Liegen die Stufen für einen Faktor zu weit auseinander, so reicht das quadratische Modell nicht zur Beschreibung aus. Liegen die Stufen zu dicht beieinander, so werden die wahren Unterschiede von Zufallsstreuung überdeckt.
- Im Verlauf der Untersuchung kann sich das System allmählich verändern (Trend). Bei einer sequentiellen Vorgehensweise verfälscht dies meist die Ergebnisse, weil der Trend nicht durch Randomisierung unschädlich gemacht werden kann.
- Bei der Berechnung des Optimums wird nur eine Zielgröße berücksichtigt. Meist sollen aber mehrere Zielgrößen gleichzeitig optimiert werden. Diese müssen dann zu einer übergeordneten Zielgröße zusammengefasst werden (dazu eignen sich z.B. die Kosten, die Wunschfunktion oder die Verlustfunktion). Bei einer sequentiellen Vorgehensweise liegt das Problem nun darin, dass die Gewichtung der verschiedenen Zielgrößen und die Art der Zusammenfassung zur übergeordneten Größe vorweg festgelegt werden muss.

³ In der numerischen Mathematik ist das Simplexverfahren bei der Suche nach Extremwerten weit verbreitet. Durch Anpassung der Schrittweite kann man dort den Extremwerten beliebig nahe kommen, da keine Zufallsstreuung auftritt.

„Was wäre, wenn“-Untersuchungen, wie sie in Abschnitt 15.3 beschrieben wurden, sind dann nicht mehr möglich. Da aber die richtige Gewichtung der einzelnen Zielgrößen nicht immer offensichtlich ist, kann sich im Nachhinein zeigen, dass eigentlich die falsche übergeordnete Zielgröße optimiert wurde.

Versucht man nun ein Programm zu schreiben, das als „Black box“ alle Möglichkeiten abdecken soll, so stellt man extrem hohe Anforderungen an dieses Programm.

Zum Ansatz, der in [3] beschrieben ist, gibt es die Software ULTRAMAX® (<http://www.ultramax.com>). Da im Programm viele Optionen einstellbar sind, deren Wirkung aber aufgrund des nicht bekanntgegebenen Algorithmus nicht nachvollziehbar ist, ist der Nutzen des Programms in der praktischen Anwendung nur schwer zu bewerten. In einigen Testbeispielen mit bekannten Funktionen und daher bekannten Ergebnissen war das Programm jedoch kaum besser als ein einfaches Simplexverfahren.

In [4] ist ein Optimierungsalgorithmus beschrieben, der, wie das Simplexverfahren, aus der numerischen Mathematik stammt, aber über die lineare Näherung hinausgeht. Im Vergleich zu anderen numerischen Verfahren ist dieser Algorithmus relativ unempfindlich gegenüber Zufallsstreuung. Nach [4] hat er sich in numerischen Testbeispielen bewährt, in denen die Zufallsstreuung nicht zu groß ist. Bei großer Zufallsstreuung und wenn in den Daten ein Trend mit enthalten ist, können m.E. jedoch Probleme auftreten. Außerdem gibt es für diesen Algorithmus noch kein allgemein zugängliches Computerprogramm.

Die Software GlobalOptimize (GO, www.globaloptimize.ch) arbeitet mit unterschiedlichen lokalen und globalen Modellen. Als lokale Modelle werden meist lineare Modelle verwendet, deren Parameter schon mit relativ wenigen Einzelversuchen bestimmt werden können. Aus dem Abgleich zwischen den Zielvorgaben und den Versuchsergebnissen ergibt sich eine Verbesserungsrichtung, in der weitere Versuche vorgeschlagen werden. Iterativ werden neue lokale Modelle angepasst und neue Versuche vorgeschlagen, bis die Zielvorgaben erreicht sind oder abgebrochen wird. Außerdem lassen sich die gesammelten Daten jederzeit mit Hilfe lokaler oder globaler Modelle untersuchen.

Insgesamt betrachtet sind dies vielversprechende Ansätze zu einer Verbesserung des Simplexverfahrens. Weitere Verbesserungen sind jedoch m.E. noch wünschenswert. Als Ansatzpunkte könnte ich mir z.B. vorstellen, dass statistische Auswertungen stärker integriert werden und eine formale Trenderkennung mit eingebaut wird (Blockbildung).

Literatur

- [1] Box, G. E. P./Draper, N. R.: Das EVOP-Verfahren. Oldenbourg Verlag, München 1975
- [2] Scheffler, E.: Statistische Versuchsplanung und Auswertung. DVG, Stuttgart 3. Auflage 1997
- [3] Moreno, C. W./Eilebrecht, B.: „Sequentielle Prozessoptimierung“ in: Qualität und Zuverlässigkeit 37 (Nr. 1, 1992), 53–56
- [4] Elster, C./Neumaier, A.: „A trust region method for the optimization of noisy functions“ in: Computing 58 (1997), 31–34

Dieses Kapitel gibt einen kurzen Überblick über ausgewählte Software zur Versuchsplanung. Die beschriebenen Programme sollen als Beispiele dienen. Die Auswahl ist subjektiv.

■ 19.1 Allgemeine Hinweise

- In diesem Buch wurde Versuchsplanung unabhängig von einer speziellen Software beschrieben. Ziel war die allgemeine Darstellung der Vorgehensweise, der Verfahren und ihrer spezifischen Anwendungsmöglichkeiten.
- Auch wenn die Analysen bis einschließlich Kapitel 10 ohne Computer möglich sind, wird dringend empfohlen, Computer einzusetzen, sobald statistische Auswertungen durchgeführt werden. Gute Software bringt eine dramatische Vereinfachung der Rechnung und erlaubt viele aussagekräftige grafische Darstellungen der Ergebnisse. Der Benutzer kann sich dadurch auf die sorgfältige Planung und Durchführung der Versuche, das Verständnis der Ergebnisse und das Ableiten von Maßnahmen konzentrieren.
- Statistikprogramme gibt es für Großrechner und PCs, und für alle gängigen Betriebssysteme. Aufgrund ihrer weiten Verbreitung werden im Folgenden nur Programme für PCs mit Windows®-Betriebssystem betrachtet. Manche dieser Programme gibt es auch für macOS®.
- Nur Programme mit ihrem Schwerpunkt auf den klassischen statistischen Verfahren werden in diesem Kapitel behandelt.
 - Für die einfachen Verfahren in Kapitel 4 und 5 ist keine spezielle Software erforderlich. Soweit überhaupt Computer eingesetzt werden, genügen normale Tabellenkalkulationsprogramme.
 - Die Verfahren nach G. Taguchi in Kapitel 9 können, soweit sie hier empfohlen wurden, auch mit einigen der beschriebenen Programme eingesetzt werden. Es gibt zwar auch spezielle Taguchi-Programme, diese sind m. E. jedoch zu eingeschränkt auf die Besonderheiten von G. Taguchi.
 - Von den Optimierungsverfahren in Kapitel 18 verwendet EVOP faktorielle Versuchspläne. Die Richtung des steilsten Anstiegs wird errechnet oder aus Höhenliniendiagrammen abgelesen. Software für andere Verfahren wird direkt in Kapitel 17 und 18 behandelt (vgl. auch Anhang J).
- Je mehr Möglichkeiten ein Programm bietet, desto mehr Aufwand ist für die Einarbeitung erforderlich. Einige besonders umfassende Statistikprogramme zeichnen sich da-

durch aus, dass sie beliebig große Datenmengen verarbeiten und immer wiederkehrende Abläufe durch eine eigene Sprache zur Programmierung der Auswertung automatisiert werden können. Diese Möglichkeiten werden durch einen wesentlich erhöhten Einarbeitungsaufwand erkauft. Da bei der Versuchsplanung die Datenmengen überschaubar sind und jede Auswertung anders abläuft, kommen die Vorteile dieser Programme hier nicht zum Tragen. Sie werden daher im Folgenden nicht betrachtet.

- Bei der Versuchsplanung ist die einfache und gleichzeitig möglichst flexible Bedienung besonders wichtig. Die hier beschriebenen Programme unterstützen (fast) alle Verfahren und Versuchspläne aus diesem Buch (und meist noch einiges mehr). Sie stellen einen Kompromiss zwischen relativ einfacher Bedienbarkeit und trotzdem möglichst umfassenden Auswertemöglichkeiten dar.
- Manche der vorgestellten Programme sind besonders gut bei der Behandlung der Standardversuchspläne, andere sind bei optimalen Plänen besonders stark. Inzwischen können alle besprochenen Programme optimale Pläne für quantitative Faktoren aufstellen. Aber nur die „Spezialisten“ können dies auch, wenn einige Faktoren quantitativ und andere qualitativ (kategorisch) sind. Auch Mischungsfaktoren werden nicht von allen Programmen beherrscht. Dafür haben diese „Spezialisten“ meist Probleme mit vermengten Wechselwirkungen bei fraktionellen faktoriellen Plänen (bei optimalen Plänen treten keine Vermengungen zwischen den Effekten auf, für die der Plan aufgestellt wurde).
- Die im nächsten Abschnitt beschriebenen Programme lassen sich in zwei Gruppen einteilen:
 - Spezielle Versuchsplansprogramme (VP) bieten eine bessere Benutzerführung.
 - Umfassende Statistikpakete mit einem umfangreichen Versuchsplanauswertungsteil (SP) bieten mehr Flexibilität und weitere Auswertungsverfahren.

■ 19.2 Beschreibung ausgewählter Programme

Tabelle 19.1 gibt einen Überblick über die getesteten Programme, die folgenden Tabellen zeigen, welche Verfahren bei den verschiedenen Programmen implementiert sind.

TABELLE 19.1 Getestete Programme (in alphabetischer Reihenfolge)

| Name | Version | Typ | deutsch | Internetadresse mit Informationen |
|-------------------------|---------|-----|---------|--|
| Cornerstone® (CO) | 7 | SP | nein | www.camline.com/ |
| Design-Expert® (DX) | 12 | VP | nein | www.statease.com/ |
| JMP® | 14 | SP | ja | www.jmp.com/ |
| Minitab® (MI) | 19 | SP | ja | www.minitab.com/ |
| MODDE® (MO) | 12 | VP | nein | www.umetrics.com/ |
| Statgraphics® (SG) | 18 | SP | ja | www.statgraphics.com/ |
| TIPCO® Statistica® (ST) | 13 | SP | ja | www.statsoft.de/ |
| STAVEX® (SV) | 6 | VP | ja | www.aicos.com/ |
| Visual-XSel® (XS) | 15 | SP | ja | www.crgraph.de/ |

TABELLE 19.2 Implementierte Versuchspläne

| Standard-Versuchspläne | Kapitel | CO | DX | JMP | MI | MO | SG | ST | SV | XS |
|---|---------|----|----|-----|----|----|----|-----|-----|-----|
| vollständige faktorielle 2^k | 7 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| fraktionelle faktorielle 2^{k-p} | 8.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Anzeige Vermengungen | 8.2 | | ✓ | ✓ | ✓ | | ✓ | ✓ | | (✓) |
| Blockbildung | 7.1.4 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | |
| Plackett–Burman | 8.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Definitive Screening Pläne | 8.4 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| Taguchi-Pläne | 9 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Taguchi inner + outer array | 9 | | | ✓ | | | ✓ | ✓ | | |
| zentral zusammengesetzte | 11.1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Box–Behnken-Pläne | 11.2.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| vollständig faktoriell mehrstufig | 12 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Lateinische Quadrate | 13.1 | ✓ | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Mischungen | 14.1 | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) |
| Optimale Versuchspläne | | | | | | | | | | |
| quantitative Faktoren | 11.2.4 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| quant. + qual. Faktoren 2 Stufen | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| quant. + qual. Faktoren > 2 Stufen | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Split-Plot-Pläne 2 Stufen | 12.4.2 | | ✓ | ✓ | ✓ | | | ✓ | | |
| Split-Plot-Pläne > 2 Stufen | 12.4.2 | | ✓ | ✓ | | | | ✓ | | |
| Mischungsfaktoren (allein) | 14.4 | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| quant. + qual. + Mischungsfaktoren | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| „, mehrstufige Faktoren | | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| auch für kubische Modelle | 11.3.2 | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ |
| auch 3-Faktor-WW | 11.3.2 | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | |
| auch mit Blockbildung | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ |
| Planerzeugung in einem Schritt ¹ | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Erweiterung vorhandener Pläne ² | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | ✓ | ✓ |
| Begrenzungen als Ungleichungen | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Begrenzungen einfach eingeben ³ | 16.4 | | ✓ | | | ✓ | | | | ✓ |
| Vorgabe eines Kandidatensets ⁴ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | | ✓ |

¹ ohne vorher eine Liste mit möglichen Punkten erzeugen zu müssen² wenn bereits Versuchsergebnisse vorliegen, können diese zu einem optimalen Plan erweitert werden (sogenannte Inclusions)³ Vorgabe von begrenzenden Punkten, statt Ungleichungen oder Kandidatenset⁴ Alternative zur Vorgabe von Begrenzungen: Bietet große Freiheit, ist aber sehr aufwändig

Tabelle 19.3 Implementierte Auswertungsverfahren, Kennzahlen, Diagnosegrafiken

| Auswertungsverfahren | | CO | DX | JMP | MI | MO | SG | ST | SV | XS |
|---|----------|----|----|-----|----|----|-----|-----|-----|-----|
| Signifikanz $n \geq 2$ | 7.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $n = 1$ (Pooling) | 7.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Regression | 10 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| „, schrittweise | 10.2.3 | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | (✓) | (✓) | ✓ |
| „, 2FWW | 7.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| „, quadratisches Modell | 11 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| „, 3FWW | 7.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | | |
| „, kubisches Modell | 11.3.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | | ✓ |
| „, AIC/BIC ¹ | 17.3.1 | | ✓ | ✓ | ✓ | | | | | |
| Varianzanalyse, feste Effekte | 12 + 13 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) |
| „, gemischte Modelle | 12.3 | | | ✓ | ✓ | | ✓ | (✓) | | |
| „, split-plot | 12.4 | | ✓ | ✓ | ✓ | | | (✓) | | |
| „, geschachtelt | 12.4 | | | ✓ | ✓ | | ✓ | (✓) | | |
| Transformationen | 6.4+15.1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Blockfaktor ein-/ausschalten | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| Zielgröße Standardabweichung ² | | ✓ | | | ✓ | | | | | ✓ |
| Signal-Rausch-Verhältnis | 9.2 | | | ✓ | ✓ | | ✓ | ✓ | | |
| Direktauswertung Excel-Datei | | ✓ | | ✓ | ✓ | | ✓ | ✓ | | ✓ |
| Kennzahlen | | | | | | | | | | |
| Bestimmtheitsmaß $B = r^2$ | 10.1.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| PRESS bzw. $r^2(\text{pred})$ | 10.2.4 | | ✓ | ✓ | ✓ | ✓ | (✓) | | | ✓ |
| VIF | 10.2.4 | ✓ | ✓ | ✓ | ✓ | | ✓ | (✓) | | ✓ |
| Lack of Fit | 10.2.4 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Diagnosegrafiken | | | | | | | | | | |
| Residuen gg. Versuchsnummer | 10.1.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| „ im Wahrscheinlichkeitsnetz | 6.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| „ gg. Faktor | 10.1.3 | ✓ | ✓ | | ✓ | ✓ | ✓ | | | ✓ |
| „ gg. berechnete Zielgröße | 10.1.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| entfernte/student. Residuen | 10.1.3 | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | |
| Hebel | 10.2.4 | ✓ | ✓ | ✓ | | | | (✓) | | |
| Cook-Distanz oder DFIT | 10.2.4 | | ✓ | ✓ | | | | (✓) | | |
| Box-Cox-Darstellung Transform. | 6.4 | ✓ | ✓ | ✓ | | ✓ | | ✓ | | ✓ |
| „ mit Vertrauensbereich | 6.4 | ✓ | ✓ | ✓ | | ✓ | | | | ✓ |
| Effekte im WN | 7.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | (✓) | |

¹ Alternativ zum p-Wert kann AIC/BIC als Kriterium zur Modellwahl verwendet werden² Aus Wiederholungen derselben Faktorstufenkombination wird Standardabweichung berechnet, die dann direkt als weitere Zielgröße verwendet werden kann

Tabelle 19.4 Grafische Darstellung der angepassten Modelle und Ergebnisse

| Eine Zielgröße | Kapitel | CO | DX | JMP | MI | MO | SG | ST | SV | XS |
|--|-----------|----|----|-----|----|----|----|----|-----|----|
| Effekte der Faktoren | 7 + 8 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Wechselwirkungen | 7 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Wirkungsflächen | 10.2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Höhenlinien | 10.2+11.1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mischungen wie Bild 14-5 | 14.2 | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Faktoren interaktiv einstellbar ¹ | | | ✓ | ✓ | | ✓ | | | (✓) | |
| Zielgröße rücktransformiert ² | | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ |
| Mehrere Zielgrößen | | | | | | | | | | |
| Unterschiedliche Modelle für verschiedene Zielgrößen | 15.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| Wunschfunktion | 15.3 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 3 |
| Begrenzungen berücksichtigt ⁴ | | ✓ | ✓ | ✓ | | ✓ | | | ✓ | |
| Optimierungsgrafik ⁵ | 15.3 | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | | ✓ |
| " mit Vertrauensbereichen | 15.3 | ✓ | ✓ | ✓ | | ✓ | | ✓ | | ✓ |

Alle hier beschriebenen Programme sind zur Planung und Auswertung von Versuchen geeignet. Aber es gibt kein Programm, das in jeder Hinsicht ideal ist (die „eierlegende Wollmilchsau“ existiert nicht).

Die Tabellen 19.2, 19.3 und 19.4 geben einen Überblick über die implementierten Versuchspläne, Auswertungsverfahren und Ergebnisdarstellungen der Programme. Die Auswahl der Kriterien ist notwendigerweise subjektiv. Sie ist einerseits bestimmt durch die in diesem Buch behandelten Themen und andererseits durch die Art und Weise, wie ich persönlich Versuchspläne aufstelle und worauf ich bei der Auswertung achte. Keines der Programme erfüllt alle Kriterien.

Die Häkchen zeigen vorhandene Funktionen für die getestete Version. Gerade auf dem Gebiet der Versuchsplanung wurden die Programme in den letzten Jahren wesentlich weiterentwickelt. Da zu erwarten ist, dass sich diese Entwicklung fortsetzt, werden die momentan vorhandenen Lücken sicher Schritt um Schritt gefüllt werden. Häkchen in Klammern bedeuten, dass die Funktion nur eingeschränkt oder in manchen Situationen angeboten wird.

¹ Bei Wirkungsflächen und Höhenlinien kann immer nur die Abhängigkeit einer Zielgröße von zwei Faktoren dargestellt werden, für die anderen Faktoren müssen feste Werte vorgegeben werden. Sind diese Werte in der Grafik selbst einstellbar und sieht man dann direkt die Wirkung ihrer Veränderung?

² Obwohl die Zielgröße zur Modellanpassung transformiert wird, werden die Ergebnisse im ursprünglichen Maßstab dargestellt – dies erleichtert die Interpretation der Ergebnisse wesentlich.

³ Alternativer Optimierungsansatz

⁴ Werden die Begrenzungen der Faktoren bei der Optimierung berücksichtigt oder kann das rechnerische Optimum in einem verbotenen Bereich liegen?

⁵ Gibt es die Möglichkeit, die angepassten Modelle für mehrere Zielgrößen in einer Grafik darzustellen, wie in Bild 15.4, mit einstellbaren Werten für die Faktoren? Dies erleichtert die grafische Suche nach Kompromissen bei Zielkonflikten.

Vor der Entscheidung für ein bestimmtes Programm wird dringend ein eigener Test empfohlen, da die Bedürfnisse der Anwender sehr unterschiedlich sind. Um den Vergleich der Programme zu erleichtern, stehen auf der Webseite <https://www.hanser-fachbuch.de/buch/Versuchsplanung/9783446461468> unter dem Reiter Extras die Daten für die Beispiele dieses Buchs im Format der hier behandelten Programme zum Download bereit (zum Test der Importmöglichkeiten auch im ASCII-Format und als Excel®-Tabelle, siehe Anhang H).

Die Software selbst erhalten Sie direkt vom Hersteller (vgl. Tab. 19.1 und Anhang J). Die meisten Hersteller bieten zeitlich limitierte Vollversionen als Download an, bei den anderen kann man sich direkt an den Hersteller wenden. So bestehen sehr gute Möglichkeiten, sich zunächst unverbindlich von der Eignung eines Programms für den geplanten Einsatz zu überzeugen bzw. verschiedene Programme zu vergleichen. Dies wird dringend empfohlen.

Auf den folgenden Seiten finden Sie ergänzende Hinweise zu den einzelnen Programmen.

Cornerstone®

- Cornerstone ist ein umfassendes Statistikpaket mit einem umfangreichen Versuchsplanungsteil.
- Im Versuchsplanungsteil liegt der Schwerpunkt bei der Erzeugung und Auswertung von D-optimalen Plänen. Geeignete Pläne können auch für mehrstufige quantitative und qualitative Faktoren mit Wechselwirkungen aufgestellt werden. Mischungsexperimente erfordern einen Füllerfaktor. Split-Plot-Pläne werden nicht unterstützt.
- Viele der in diesem Buch beschriebenen Standardversuchspläne werden angeboten. Bei der Auswertung von fraktionellen faktoriellen Plänen zeigt Cornerstone die vermengten Effekte jedoch nicht direkt an.
- Die Darstellung der Residuen ist einfach. Brushing bietet gute Unterstützung bei der Analyse der Ursachen für ungewöhnliche oder auffallende Einzelergebnisse.

Design-Expert®

- Design-Expert ist ein spezielles Versuchsplanungsprogramm. Als solches ist es auf das Aufstellen und Auswerten von Versuchen hin optimiert.
- Daraus ergibt sich eine besonders klare Benutzerführung und einfache Bedienung, ohne dass die Freiheit des Nutzers dadurch wesentlich eingeschränkt wird. Die Reihenfolge der Menüpunkte gibt die beste Reihenfolge zur Bearbeitung vor. Das Handbuch führt anhand von Beispielen sehr gut in die Bedienung des Programms ein. Innerhalb weniger Stunden kann man sich vollständig einarbeiten.
- Design-Expert enthält die für die Prozess- und Produktoptimierung wichtigsten Versuchspläne. Optimale Pläne decken praktisch alle Sonderfälle ab. Es gibt sehr gute Möglichkeiten zur Beurteilung der statistischen Eigenschaften der Versuchspläne. Fast alle üblichen Auswertungs- und Darstellungsverfahren werden angeboten.

JMP®

- JMP ist ein umfassendes Statistikpaket mit einem umfangreichen Versuchsplanungsteil.

- JMP hat eine ansprechende Benutzeroberfläche, in der man sich nach kurzer Eingewöhnung auch gut zurechtfindet.
- Bei der Versuchsplanung bietet JMP fast alle hier behandelten Standardpläne an (und noch viele andere). Auch Taguchi-Pläne für zwei- und dreistufige Faktoren stehen zur Verfügung, als „inner und outer arrays“. Optimale Versuchspläne können für beliebige Kombinationen von quantitativen, qualitativen und Mischungsfaktoren aufgestellt werden. Beliebige Blockgrößen können vorgegeben werden.
- JMP zeichnet sich durch eine umfassende Sammlung von Auswertungs- und Darstellungsverfahren aus, die weit über die in diesem Buch beschriebenen Verfahren hinausgehen (z.B. auch neuronale Netze und Gaussian Process Models). Auch gemischte Modelle und Split-Plot-Versuche können komfortabel ausgewertet werden.

Minitab®

- Minitab ist ein umfassendes Statistikpaket mit einem umfangreichen Versuchsplanungsteil und für SixSigma besonders weit verbreitet.
- Die Bedienung von Minitab ist einfach und unproblematisch.
- Die meisten der hier behandelten Standardversuchspläne und Auswertungsverfahren werden unterstützt.
- D-optimale Pläne können für quantitative Faktoren und Mischungsfaktoren getrennt aufgestellt werden. Mischungsfaktoren können nur mit zweistufigen quantitativen und qualitativen Faktoren kombiniert werden.

MODDE®

- MODDE Pro ist ein spezielles Versuchsplanungsprogramm. Als solches ist es auf das Aufstellen und Auswerten von Versuchen hin optimiert.
- Die Bedienung des Programms ist einfach und weitgehend selbsterklärend.
- Der Schwerpunkt des Programms liegt auf der Erzeugung und Auswertung D-optimaler Versuchspläne. Mit Ausnahme von Split-Plot-Plänen sind beliebige Kombinationen von quantitativen, qualitativen und Mischungsfaktoren möglich. Die meisten Standardversuchspläne können aufgestellt werden, auch einige hier nicht behandelte.
- Ein Analysis Wizard führt durch die Auswertung und macht Vorschläge zur Modellverbesserung. Falls gewünscht erfolgt die Modellanpassung auf Knopfdruck.

Statgraphics®

- Statgraphics Centurion 18 ist ein umfassendes Statistikpaket. Die meisten der hier beschriebenen Verfahren finden sich im Menüpunkt DOE (nur die speziellen Varianzanalyseverfahren in Allgemeine lineare Modelle = GLM).
- Die Bedienung von Statgraphics ist einfach und unproblematisch.
- Die meisten der hier behandelten Versuchspläne und Auswertungsverfahren werden unterstützt. Taguchi-Pläne mit „inner und outer array“ können aufgestellt und ausgewertet werden (Signal-Rausch-Verhältnis).

TIPCO® Statistica

- Statistica besitzt eine umfangreiche Versuchsplansammlung. Alle in diesem Buch behandelten Standardpläne sind enthalten. Mit dem Experimental Design Builder können optimale Pläne für eine breite Palette von Anwendungsfällen aufgestellt werden.
- Fast alle in diesem Buch beschriebenen Auswertungsverfahren, Kennzahlen und Diagnosegrafiken werden angeboten, allerdings sind sie auf verschiedene Module (vor allem DOE und GLM = General linear model) verteilt und daher schwer zu kombinieren.

STAVEX®

- STAVEX steht für „Statistische Versuchsplanung mit Expertensystem“. Es ist ein spezielles Versuchsplanungsprogramm. STAVEX gibt es in zwei Versionen (mit und ohne Mischungsfaktoren). Die folgenden Aussagen beziehen sich auf die Vollversion mit Mischungsfaktoren.
- Zielgruppe von STAVEX sind Anwender mit wenig Erfahrung in Versuchsplanung, die dann durch ein Expertensystem geführt werden. Dem Anwender werden Fragen gestellt und je nach Antwort wird ein bestimmtes weiteres Vorgehen empfohlen. Zahlreiche Alternativen werden angeboten, ich vermisste jedoch Informationen zur Auflösung und Vermengungsstruktur der vorgeschlagenen Pläne. Blockbildung ist nur für faktorielle Pläne möglich.
- Als gut empfinde ich, dass bei vielen Faktoren zunächst mit Screening-Plänen die Anzahl der Faktoren reduziert wird und erst mit wenigen Faktoren eine Optimierung versucht wird.
- Die Auswertung der Ergebnisse ist vom Expertensystem vorgegeben – hier wünsche ich mir mehr Möglichkeiten zur Beeinflussung, insbesondere bei der Modellbildung. Die grafische Darstellung der Ergebnisse und die gleichzeitige Optimierung mehrerer Zielgrößen mit der Wunschfunktion sind gut.

Visual-XSel®

- XSel ist ein allgemeines Statistikpaket, das einen großen Versuchsplanungsteil enthält.
- Die Benutzeroberfläche lehnt sich eng an Excel® an und Excel-Dateien können direkt gelesen und sehr einfach ausgewertet werden.
- Das Programm enthält ein Modul Systemanalyse, das die Auswahl der Faktoren und die Festlegung der Stufen unterstützt. Bei der Versuchsplanung liegt der Schwerpunkt auf D-optimalen Plänen für quantitative Faktoren. Vermengungen werden nur bis 2FWW dargestellt. Für Mischungen ist XSel m. E. nicht geeignet.
- Bei der Auswertung helfen attraktive Grafiken und Experteninformationen in Form von Sprechblasen bei der Interpretation der Ergebnisse. Die gleichzeitige Optimierung mehrerer Zielgrößen ist möglich, allerdings wird ein etwas anderer Ansatz als die Wunschfunktion verwendet.
- Für Standardsituationen gibt es fertige Auswertedateivorlagen, bei denen die Berechnungsverfahren als Flussdiagramme vorliegen. XSel verwendet diese als Makrosprache. Die Verfahren können dadurch in ihrem Ablauf eingesehen und von erfahrenen Anwendern auch verändert und erweitert werden.

- Der Anwender hat Zugriff auf diverse Hilfstabellen der Auswertungen – im Prinzip kann man damit weitere Auswertungen durchführen. Allerdings leidet durch die Fülle der Tabellenseiten die Übersichtlichkeit und der normale Anwender wird diese Informationen nicht nutzen.
- Insgesamt besitzt XSel eine sehr offene Struktur. Der Anwender kann an vielen Stellen eingreifen und damit bei unbedachter Vorgehensweise leider auch irreführende Ergebnisse erzeugen.
- XSel enthält auch mehrere Auswertedateivorlagen für Verfahren nach D. Shainin.

■ 19.3 Spezielle Anwendungsgebiete

Mit zunehmender Verbreitung der Versuchsplanung wird inzwischen auch Software für spezielle Anwendungsgebiete entwickelt. Dies geschieht natürlich vor allem für Anwendungen,

- bei denen Versuchsplanung sehr viel eingesetzt wird und daher ein ausreichender Markt vorhanden ist,
- die Möglichkeiten zur Automatisierung von Abläufen bieten, weil immer wieder dieselben Versuchseinrichtungen verwendet werden und
- die besondere Anforderungen bei der Auswertung der Ergebnisse und der Optimierung stellen.

Im Folgenden werden beispielhaft vier solche Programme vorgestellt.

CAMEO® zur Motoroptimierung

An modernen Pkw- und Nutzfahrzeug-Motoren können sehr viele Parameter eingestellt werden, wie z. B. Menge und Zeitpunkt von Vor-, Haupt- und Nacheinspritzung, Einspritzdruck, Aufladung und Abgasrückführung. Für jeden Betriebszustand (Drehzahl, Last, Motortemperatur usw.) sind im elektronischen Steuergerät optimale Werte für diese Parameter hinterlegt, die minimalen Kraftstoffverbrauch bei gleichzeitig günstigen Abgaswerten und Fahreigenschaften gewährleisten.

Bei der Entwicklung und Optimierung neuer Motoren und dann insbesondere zur Bedeutung dieses Steuergeräts (Festlegung der optimalen Einspritzparameter für jeden Betriebszustand) werden umfangreiche Versuche am Motorenprüfstand durchgeführt.

CAMEO koppelt nun die Versuchsplanung und Auswertung direkt mit dem Prüfstand. Damit bietet es z. B. folgende Möglichkeiten:

- Versuchspläne können automatisch abgefahren werden, d. h. jede vorgesehene Faktorstufenkombination wird definiert vom Prüfstand angefahren, eine Stabilisierungszeit wird abgewartet und danach werden die Messwerte für alle Zielgrößen, Umgebungsbedingungen und alle weiteren relevanten Parameter erfasst.
- Vor dem Aufstellen des eigentlichen Versuchsplans können die Grenzen der Faktoren (des Versuchsräums) zunächst automatisch ermittelt werden, indem Faktoren schrittweise verändert werden bis z. B. die Verbrennung instabil oder die maximal zulässige

Abgastemperatur überschritten wird – im so abgegrenzten Bereich kann dann automatisiert z. B. ein D-optimaler Versuchsplan aufgestellt und durchgeführt werden.

- Bei der Auswertung der Ergebnisse können direkt die Informationen ermittelt werden, die für die Bedatung des Steuergeräts erforderlich sind.

So wird die Anwendung der Versuchsplanung zur Motoroptimierung erleichtert. Die Internetadresse www.avl.com/cameo/ gibt weitere Informationen.

vidara®/destra® zur Fertigungsoptimierung

Vor allem in den Fertigungsbereichen der Automobilindustrie und bei Zulieferern ist qs-STAT® verbreitet. Viele Daten zur Überwachung der Fertigung werden mit speziellen Schnittstellen automatisiert erfasst und in einem speziellen Datenformat abgelegt und analysiert.

vidara®/destra® erweitern nun die Funktionalität in Richtung Versuchsplanung und erlauben es, dabei auf bereits vorhandene Daten und Schnittstellen zurückzugreifen.

Als Versuchspläne werden die in der Fertigung besonders wichtigen vollständigen und fraktionellen faktoriellen Versuchspläne, die zentral zusammengesetzten Pläne und die D-optimalen Pläne angeboten. Bei der Auswertung werden die übliche Regressionsanalyse einschließlich Wirkungsflächen, Höhenlinien und Optimierung mit der Wunschfunktion, sowie die Varianzanalyse mit festen und zufälligen Faktoren unterstützt.

Die Internetadresse www.q-das.de/de/software/statistik/ gibt weitere Informationen.

Weibull++® zur Zuverlässigkeitserhöhung

Bei Zuverlässigkeitsuntersuchungen ist die Zielgröße häufig die Zeit bis zum Ausfall. Statt der Normalverteilung beschreibt die Weibullverteilung die Verteilung dieser Ausfallzeiten. Und Zuverlässigkeitsuntersuchungen werden (vor allem für gute Teile) häufig schon vor dem Ausfall abgebrochen. Oder der Ausfallzeitpunkt ist nicht genau bekannt, sondern nur das Intervall zwischen einem Test und dem nächsten. Dadurch entstehen sogenannte censurierte Daten. Weibull++® ist auf diese Besonderheiten von Ausfallzeiten spezialisiert.

Als Versuchspläne bietet es vollständige und fraktionelle faktorielle Versuchspläne, zentral zusammengesetzte Pläne, Box-Behnken Pläne, sowie Taguchi-Pläne an. Neben den erwähnten Besonderheiten bietet es die üblichen Auswertungsverfahren und Darstellungen für normalverteilte Daten an.

Die Internetadresse www.reliasoft.com gibt weitere Informationen.

Fusion QbD® in der Nahrungsmittel- und Pharma industrie

In der Nahrungsmittelindustrie, Biotechnologie und pharmazeutischen Industrie werden hohe Anforderungen an Genauigkeit, Reproduzierbarkeit, Linearität und Spezifität von Analysemethoden und Messsystemen wie Gas- und Flüssigkeitschromatographen gestellt. Diese Anforderungen sind z. B. in Richtlinien der FDA (US Food and Drug Administration) und ICH (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use) festgelegt.

Um die Einhaltung dieser Anforderungen nachzuweisen, müssen sämtliche Versuche sorgfältig geplant, statistisch fundiert ausgewertet und dokumentiert werden. Da es sich

um immer wieder ähnliche Versuche an ähnlichen Chromatographen handelt, ergibt sich ein erhebliches Einsparungspotential durch die Automatisierung von Versuchsplanung, Datenaustausch, Analyse und Dokumentation.

Fusion QbD hat sich auf diese Automatisierung spezialisiert und bietet Module zur Entwicklung und Validierung von Analyseverfahren. Die Internetadresse www.smatrix.com gibt weitere Informationen.

20

Beispiele

Das folgende Beispiel aus dem Bereich Motoroptimierung zeigt, wie Versuchspläne an eine spezielle Situation angepasst werden können. Der anschließende Abschnitt gibt Hinweise darauf, wie Sie Literaturquellen für Beispiele aus Ihrem Anwendungsgebiet finden können.

Versuchsplanung lebt von der Anwendung. Dazu empfehle ich, mit einem einfachen, überschaubaren Beispiel mit zwei bis drei Faktoren zu beginnen. Wichtig ist, dass das Ziel und dafür geeignete, messbare Zielgrößen klar sind.

Falls Sie noch etwas Ermutigung brauchen, bevor Sie sich an einen echten Fertigungsprozess oder ein echtes Produkt heranwagen, helfen vielleicht die einfachen, risikolosen Übungsbeispiele am Ende des Kapitels.

■ 20.1 Beispiel Motoroptimierung

Im Steuergerät von Diesel- und Ottomotoren moderner Kraftfahrzeuge sind, abhängig von Drehzahl, Last, Motortemperatur usw. optimale Werte für die Motorsteuerung hinterlegt. Diese werden in Versuchen am Motorenprüfstand festgelegt. Hier wird ein einfaches Beispiel für einen solchen Versuch an einem 4-Zylinder-Ottomotor mit 2l Hubraum und 120kW Leistung behandelt.¹

Der Versuch wurde bei Betriebstemperatur durchgeführt. Da der Motor in einem breiten Bereich von Drehzahl und Last betrieben wird, ist es nicht möglich, ein quadratisches Modell in Drehzahl und Last anzupassen, das über den gesamten Bereich gilt – die Abweichungen wären zu groß (vgl. Abschnitt 11.3). Im hier beschriebenen Versuch wurden daher 20 Drehzahl-Last-Kombinationen ausgewählt, die über den gesamten Drehzahl-Last-Bereich verteilt sind. Für jede dieser Kombinationen wurden am Prüfstand die von der Steuerung einstellbaren Parameter in einem lokalen Versuchsplan verändert und aus den Ergebnissen die optimalen Einstellungen ermittelt. Für die Bedeutung des Steuergeräts wird dann zwischen den so ermittelten Werten interpoliert.

Im Folgenden wird beispielhaft eine dieser Drehzahl-Last-Kombinationen betrachtet, nämlich 950 U/min bei einem effektiven Mitteldruck PME von 3,9bar (bestimmt das Dreh-

¹ Die Daten sind Teil eines größeren Beispiels der Firma AVL List GmbH, Graz, veröffentlicht mit freundlicher Genehmigung.

moment). Die Drehzahl wurde am Bremsenregler eingestellt und die Fahrpedalstellung an der Motorsteuerung wurde so geregelt, dass das Sollmoment erreicht wurde. Die benötigte Kraftstoffmenge wurde gemessen und ergab den Verbrauch.

Ziel des Versuchs war es, die den Verbrauch bestimmenden Stellungen der Einlassnockenwelle und der Auslassnockenwelle so festzulegen, dass der Verbrauch möglichst niedrig ist. Die Verbrennung darf dabei jedoch nicht instabil werden und Abgasgrenzwerte müssen eingehalten werden. Am untersuchten Motor kann zwischen einem langen und kurzen Saugrohr umgeschaltet werden – es sollte festgelegt werden, welches Saugrohr bei der untersuchten Drehzahl-Last-Kombination besser ist. Der Zündwinkel wurde so eingestellt, dass der Verbrennungsschwerpunkt 8° hinter dem oberen Totpunkt liegt – der resultierende Zündwinkel ergibt erfahrungsgemäß den niedrigsten Verbrauch.

Als Zielgrößen wurden erfasst:

- Der effektive Verbrauch $beff$ [g/kWh] – er soll so klein wie möglich sein, wobei für die anderen Zielgrößen die zulässigen Grenzwerte eingehalten werden müssen.
- Die Konzentration der Schadstoffe NO_x und HC im Abgas – sie sollen nicht höher als ca. 600 ppm sein.
- Die Standardabweichung SigmaPMI des Mitteldrucks über 100 Motorarbeitsspiele – sie ist ein Maß für die Stabilität der Verbrennung und soll nicht höher als ca. 0,1 bar sein.

Zur Beurteilung der Stabilität der Versuchsbedingungen wurden u. a. erfasst:

- Umgebungsluftdruck [mbar]
- Umgebungstemperatur [$^\circ C$]
- Öltemperatur [$^\circ C$]

Als Faktoren wurden verändert:

- A: Saugrohr mit den Stufen kurz und lang – Schalterstellungen 0 und 1.
- B: Stellung der Einlassnockenwelle im Bereich 0 bis 50° Kurbelwinkel (KW).
- C: Stellung der Auslassnockenwelle im Bereich 0 bis 50° KW.

Dabei ist jedoch bekannt, dass die Stellung der Auslassnockenwelle nicht um mehr als ca. 10 bis 20° KW größer sein darf als die Einlassnockenwelle, da sonst die Verbrennung sehr schnell instabil wird. Dies würde bei der Modellierung die in Bild 11-13 beschriebenen Probleme verursachen.

Da sich die Faktoren Einlass- und Auslassnockenwelle somit gegenseitig begrenzen („constraints“), sind die Standardversuchspläne nicht geeignet. Wie in Absatz 11.2.4 beschrieben, wurde daher ein D-optimaler Plan verwendet.

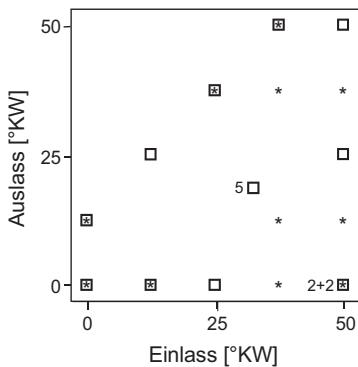
Da vor den Versuchen nicht bekannt war, ob ein quadratisches Modell zur Beschreibung der Zusammenhänge ausreicht, wurde zur Sicherheit ein Versuchspläne für ein kubisches Modell in den Faktoren Einlass- und Auslassnockenwelle aufgestellt (mit fünf Stufen je Faktor, für ein kubisches Modell sind mindestens vier Stufen erforderlich). Der Faktor Saugrohr ist diskret (qualitativ) und kann nur die zwei Stufen kurz und lang bzw. 0 und 1 annehmen – er geht formal linear in das Modell ein, praktisch erhält man verschiedene Höhenliniendarstellungen für kurzes und langes Saugrohr.

Das kubische Modell enthält neben der Konstanten die 15 Terme A, B, C, AB, AC, BC, BB, CC, ABC, ABB, ACC, BBB, BBC, BCC und CCC (A tritt höchstens einmal auf – linear, B und C bis zu dreimal – kubisch). Als Versuchsplan wurde ein D-optimaler Plan verwendet, bei dem die Kombination A=0, B=32,5 und C=18,75 °KW als Wiederhol- und Referenzpunkt fünfmal realisiert wurde, 20 weitere Versuchspunkte je einmal und die beiden Punkte A=0 und 1, B=50 und C=0 °KW je zweimal (vgl. Tabelle 20.1 und Bild 20-1).

TABELLE 20.1 Versuchsplan, Ergebnisse für die Zielgrößen und ausgewählte Versuchsbedingungen

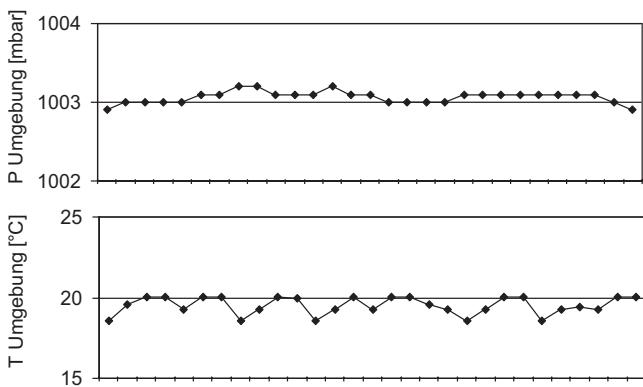
| Nr. | Saugrohr | Einlass [°KW] | Auslass [°KW] | beff [g/kWh] | NOx [ppm] | HC [ppm] | Sigma PMI [bar] | P Umgeb. [mbar] | T Umgeb. [°C] | T Öl [°C] |
|-----|----------|------------------|------------------|-----------------|--------------|-------------|--------------------|--------------------|------------------|--------------|
| 1 | 0 | 32,5 | 18,75 | 290,4 | 1868,1 | 616,6 | 0,10 | 1002,9 | 18,5 | 87,9 |
| 2 | 1 | 50 | 12,5 | 295,3 | 2264,8 | 728,4 | 0,08 | 1003,0 | 19,6 | 86,9 |
| 3 | 1 | 37,5 | 37,5 | 278,4 | 907,9 | 544,9 | 0,11 | 1003,0 | 20,0 | 87,2 |
| 4 | 1 | 0 | 0 | 291,3 | 1194,2 | 558,8 | 0,12 | 1003,0 | 20,0 | 86,8 |
| 5 | 0 | 25 | 0 | 307,4 | 2309,3 | 750,0 | 0,11 | 1003,0 | 19,3 | 84,5 |
| 6 | 1 | 50 | 50 | 282,3 | 608,0 | 570,1 | 0,11 | 1003,1 | 20,0 | 83,9 |
| 7 | 1 | 50 | 0 | 302,2 | 2442,9 | 812,3 | 0,09 | 1003,1 | 20,0 | 84,8 |
| 8 | 0 | 32,5 | 18,75 | 287,4 | 1820,4 | 659,4 | 0,10 | 1003,2 | 18,5 | 86,2 |
| 9 | 0 | 37,5 | 50 | 283,4 | 319,8 | 639,9 | 0,12 | 1003,2 | 19,3 | 86,0 |
| 10 | 0 | 25 | 37,5 | 281,3 | 414,4 | 579,7 | 0,14 | 1003,1 | 20,0 | 85,0 |
| 11 | 1 | 37,5 | 0 | 298,1 | 2484,1 | 769,3 | 0,10 | 1003,1 | 19,9 | 86,7 |
| 12 | 0 | 50 | 0 | 303,8 | 2415,4 | 786,0 | 0,09 | 1003,1 | 18,5 | 85,4 |
| 13 | 0 | 12,5 | 0 | 298,2 | 1931,6 | 667,3 | 0,10 | 1003,2 | 19,3 | 85,0 |
| 14 | 1 | 50 | 37,5 | 279,4 | 1268,5 | 609,2 | 0,11 | 1003,1 | 20,0 | 86,0 |
| 15 | 0 | 32,5 | 18,75 | 295,8 | 1803,3 | 642,7 | 0,11 | 1003,1 | 19,3 | 86,5 |
| 16 | 0 | 12,5 | 25 | 280,7 | 431,7 | 579,5 | 0,14 | 1003,0 | 20,0 | 86,5 |
| 17 | 1 | 50 | 0 | 296,8 | 2403,3 | 789,5 | 0,10 | 1003,0 | 20,0 | 85,7 |
| 18 | 1 | 0 | 12,5 | 282,7 | 587,1 | 564,4 | 0,14 | 1003,0 | 19,6 | 86,7 |
| 19 | 1 | 37,5 | 50 | 283,7 | 318,8 | 652,1 | 0,15 | 1003,0 | 19,3 | 84,3 |
| 20 | 0 | 0 | 0 | 287,1 | 1259,0 | 581,2 | 0,10 | 1003,1 | 18,5 | 86,1 |
| 21 | 0 | 32,5 | 18,75 | 286,6 | 1784,8 | 641,2 | 0,09 | 1003,1 | 19,3 | 86,9 |
| 22 | 1 | 25 | 37,5 | 287,0 | 208,1 | 629,0 | 0,18 | 1003,1 | 20,0 | 85,7 |
| 23 | 1 | 37,5 | 12,5 | 295,6 | 2236,3 | 727,7 | 0,10 | 1003,1 | 20,0 | 86,4 |
| 24 | 0 | 50 | 0 | 299,3 | 2495,0 | 791,2 | 0,08 | 1003,1 | 18,5 | 85,9 |
| 25 | 0 | 50 | 25 | 284,2 | 1940,9 | 668,6 | 0,09 | 1003,1 | 19,3 | 86,4 |
| 26 | 1 | 12,5 | 0 | 299,3 | 1968,9 | 652,6 | 0,11 | 1003,1 | 19,4 | 86,7 |
| 27 | 0 | 0 | 12,5 | 279,9 | 625,7 | 578,0 | 0,13 | 1003,1 | 19,3 | 86,9 |
| 28 | 0 | 50 | 50 | 284,3 | 546,9 | 618,0 | 0,10 | 1003,0 | 20,0 | 84,9 |
| 29 | 0 | 32,5 | 18,75 | 285,5 | 1810,9 | 651,7 | 0,09 | 1002,9 | 20,0 | 86,2 |

In Bild 20-1 ist deutlich der abgeschnittene Bereich zu erkennen, in dem eine instabile Verbrennung erwartet wird. Das Abschneiden führt zu einer deutlichen Korrelation zwischen den Modellkoeffizienten, insbesondere mit kubischen Termen. Das angepasste Modell gilt daher auch nur im untersuchten Bereich und darf nicht in den abgeschnittenen Bereich extrapoliert werden.

**BILD 20-1**

D-optimaler Versuchsplan für die Faktoren Einlass- und Auslassnockenwelle – die Auslassnockenwelle darf höchstens 10 bis 20 °KW größer sein als die Einlassnockenwelle
 □ Saugrohr 0 * Saugrohr 1
 Zahlen bei den Versuchspunkten geben die Anzahl der Realisierungen

Der erste Schritt der Auswertung ist, sich davon zu überzeugen, dass die Versuchsbedingungen stabil waren. Dazu werden in Bild 20-2 beispielhaft der Umgebungsdruck P und die Umgebungstemperatur T gegen die Versuchsnummer aufgetragen: Die Versuchsbedingungen sind ausreichend stabil und kein Trend ist erkennbar. Man kann daher von konstanten Bedingungen ausgehen. Ist dies nicht der Fall, so kann es erforderlich sein, Umgebungsbedingungen als Parameter linear mit ins Modell aufzunehmen, um so den Einfluss der Umgebungsbedingungen zu berücksichtigen.

**BILD 20-2**

Umgebungsdruck P und Umgebungstemperatur T in Abhängigkeit von der Versuchsnummer – die Versuchsbedingungen sind stabil

Der nächste Schritt ist eine kritische Überprüfung der Versuchsergebnisse, denn grobe Ausreißer verfälschen die angepassten Modelle und sollten entfernt werden. Die vorliegenden Daten enthalten keine groben Ausreißer, allerdings fällt auf, dass die NO_x -Werte sich über einen wesentlich breiteren Bereich (von ca. 200 bis 2500 ppm) verändern als die anderen Zielgrößen. SigmaPMI ist über den größten Teil des Versuchsraums fast konstant, in der Nähe der Grenze des untersuchten Bereichs steigen die Werte jedoch, wie erwartet, dramatisch an. Insbesondere bei Nr. 22 ist SigmaPMI sehr hoch – so hoch, dass

diese Kombination von Einlass- und Auslassnockenwelle sicher nicht verwendbar ist. Um eine Verfälschung des angepassten Modells wie in Bild 11-13 zu vermeiden, wird empfohlen, diesen Punkt nachträglich aus dem Versuchsplan zu entfernen. Da dieser Punkt auch bei NO_x extrem ist, wird dadurch auch die Modellanpassung für NO_x im eigentlich interessierenden Bereich verbessert.

Das Saugrohr hat keinen signifikanten Einfluss auf beff , NO_x und HC (keine Terme mit A im Modell), nur auf SigmaPMI hat das Saugrohr einen signifikanten Einfluss – das kurze Saugrohr ($A = 0$) ist günstiger.

Für beff und SigmaPMI beschreiben quadratische Modelle die Abhängigkeit von Einlass- und Auslassnockenwelle sehr gut, für HC liegt der Lack-of-Fit an der Grenze des Akzeptablen, trotzdem wird im Folgenden ein quadratisches Modell verwendet. Für NO_x benötigt man ein kubisches Modell, das quadratische Modell weicht signifikant von den Messwerten ab. In diesem Beispiel bringen Transformationen der Zielgrößen keine wesentliche Verbesserung.

Bild 20-3 zeigt die Abhängigkeit aller Zielgrößen von der Einlassnockenwelle und der Auslassnockenwelle, für SigmaPMI für das kurze Saugrohr mit den besseren Ergebnissen. In diesem einfachen Beispiel kann man leicht erkennen, dass die Einlass- und Auslassnockenwelle von jeweils ca. 50°KW den besten Kompromiss liefert. Eine formale Optimierung führt zum selben Ergebnis.

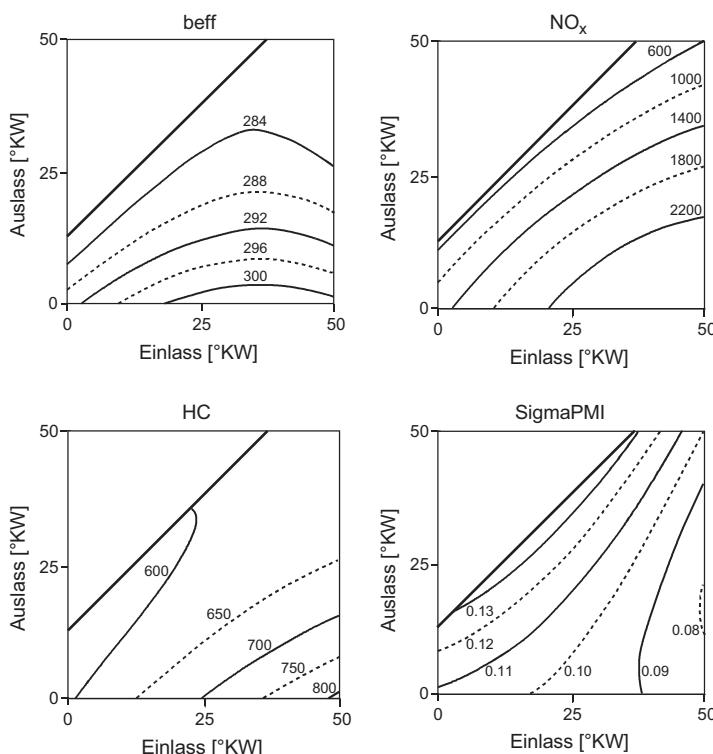


BILD 20-3
beff, NO_x , HC und
SigmaPMI in
Abhängigkeit von
Einlass- und Auslass-
nockenwelle – das
Optimum befindet sich
im rechten oberen Eck

Die mit den angepassten Modellen berechneten Werte der Zielgrößen stimmen gut mit den Messwerten im Punkt Nr. 28 überein (Nr. 6 unterscheidet sich nur im Saugrohr, das wiederum nur SigmaPMI beeinflusst). Normalerweise liegt das berechnete Optimum bei einer Faktorstufenkombination, an der noch keine Messwerte vorliegen. Diese Ergebnisse müssen dann in einem Bestätigungsversuch überprüft werden.

Ähnliche Auswertungen werden getrennt für alle Drehzahl-Last-Kombinationen durchgeführt. So erhält man die optimalen Einstellungen, die dann im Steuergerät des Motors hinterlegt werden.

Normalerweise werden deutlich mehr als drei Faktoren verändert – der Einzug der Elektronik in der Motorsteuerung führt zu immer mehr Faktoren, die optimiert werden können und müssen. Versuchsplanung wird daher verstärkt eingesetzt, um den sonst überproportional wachsenden Aufwand zu begrenzen. Hier wurde bewusst ein leicht überschaubares Beispiel ausgewählt, um zu zeigen, wie mit D-optimalen Versuchsplänen auch eingeschränkte Versuchsräume und Kombinationen von quantitativen (kontinuierlichen) und qualitativen (diskreten) Faktoren behandelt werden können.

■ 20.2 Literaturbeispiele

Weltweit gibt es mehrere tausend Veröffentlichungen mit Anwendungen der Versuchsplanung, vor allem in Fachzeitschriften und als Konferenzbeiträge (meist auf Englisch). Es ist unmöglich, hier einen angemessenen Überblick über diese Vielfalt zu geben, zumal jeder Leser natürlich vor allem an Anwendungsbeispielen aus seinem Arbeitsgebiet interessiert ist. Daher wird eine Literatursuche in Datenbanken empfohlen.

Eine gute Möglichkeit zu einer solchen Suche bietet STN, das wissenschaftliche und technische Informationsnetzwerk von CAS (Chemical Abstracts Service) in Nordamerika, FIZ Karlsruhe (Fachinformationszentrum) in Europa und JST in Japan (<http://stneasy.fiz-karlsruhe.de>).

Die „Einfache Suche“ in sechs Standarddatenbanken nach dem Begriff „design of experiment“ liefert über 4000 Suchergebnisse (davon ca. 2/3 in der Datenbank INSPEC). Die Mehrzahl dieser Suchergebnisse beziehen sich tatsächlich auf Versuchsplanung, und davon sind die Mehrzahl wiederum Anwendungsbeispiele aus den verschiedensten Gebieten (Achtung, die Anführungszeichen sind wichtig, da sonst nach den Einzelbegriffen gesucht wird, mit über 50 000 zum größten Teil irrelevanten Suchergebnissen).

Es wird empfohlen, im Anschluss an diese Grobsuche die Suche mit zusätzlichen Begriffen zu verfeinern. Dazu kann man ergänzend nach bestimmten Versuchsplantypen bzw. Auswertungsverfahren suchen (z. B. factorial, screening, response surface, central composite, Taguchi, simplex) und/oder nach Anwendungsgebieten (z. B. electronic, lithography, etch, laser, alloy, chemical). Je nach den verwendeten Zusatzbegriffen reduziert sich die Anzahl der Suchergebnisse damit deutlich. Dabei sollte man mit verschiedenen Begriffskombinationen experimentieren, um eine überschaubare Anzahl relevanter Zitate zu erhalten.

Insgesamt gibt es sehr viele veröffentlichte Anwendungsbeispiele zu den klassischen Verfahren und zu Taguchi, während zu den Ideen von Shainin nur wenige Beispiele veröffent-

licht wurden. Dies ist jedoch keine Aussage über die Nützlichkeit der Verfahren. Shainin schränkt die Veröffentlichung bewusst ein. Daher wird bezüglich Shainin auf die in Kapitel 4 zitierten Bücher verwiesen.

■ 20.3 Übungsbeispiele

Für den Papier-Rotor in Absatz 20.3.1 benötigen Sie nur Papier, Stift, Lineal, Schere und Stoppuhr. Die Versuchsdurchführung ist einfach. Allerdings ist die Zufallsstreuung erfahrungsgemäß groß im Vergleich zu den Effekten. Daher sind viele Realisierungen und eine große Flughöhe erforderlich, um signifikante Effekte zu finden.

Für den Trichter in Absatz 20.3.2 sind mehr Vorbereitungen erforderlich. Die Effekte sind dann aber größer und physikalisch leichter zu interpretieren als beim Rotor.

101 weitere Ideen/Vorschläge von William Hunter finden Sie unter
<http://curiouscat.com/bill/101doe.cfm>

Spezielle Übungsbeispiele für Chemiker beschreibt [1].

20.3.1 Papier-Rotor

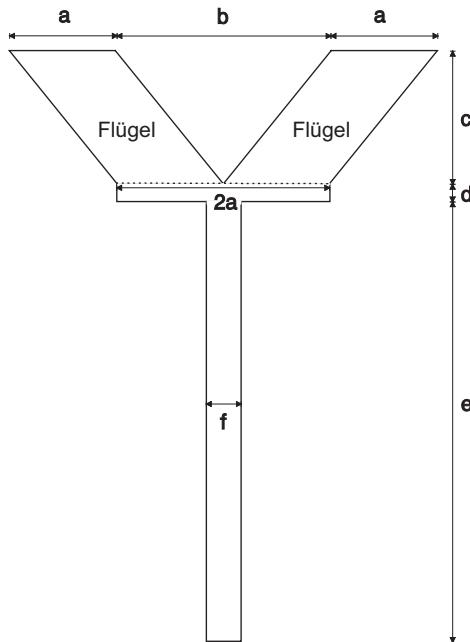


BILD 20-4

Aus einem DIN A4-Blatt ausgeschnittener Papier-Rotor [2]

Dieses Beispiel kann als Modell für eine Produktoptimierung betrachtet werden: Aus einem DIN A4-Blatt Papier soll ein Rotor gemäß Bild 20-4 ausgeschnitten werden. Wird einer der Flügel entlang der punktierten Linie um 90° nach vorne, der andere um 90° nach hinten geknickt, so beginnt sich dieser Rotor zu drehen, wenn man ihn fallen lässt.

Bei günstiger Wahl der Dimensionen kann man erreichen, dass er stabil rotiert und dabei nur langsam zu Boden sinkt (je nach verwendetem Papier kann es erforderlich sein, den Rotor am unteren Ende z. B. mit einer Büroklammer zu beschweren).

Aufgabe ist es, die Dimensionen a, b, c, d, e und f so zu wählen, dass die Zeit maximiert wird, die der Rotor benötigt, um aus einer vorgegebenen Höhe zu Boden zu sinken (= Zielgröße). Als weitere Zielgrößen sind z. B. die Absturzneigung (wie auch immer zu messen) oder die Rotationsfrequenz denkbar. Als Randbedingung sei vorgegeben, dass das DIN A4-Blatt vollständig genutzt wird, d.h.

$$2a + b = 21,0 \text{ cm} = \text{Breite} \quad (20.1)$$

$$c + d + e = 29,7 \text{ cm} = \text{Länge} \quad (20.2)$$

Damit erhält man vier Konstruktionsparameter, die frei wählbar sind. Die anderen beiden ergeben sich aus den Randbedingungen (20.1) und (20.2). Im Prinzip ist es gleichgültig, welche vier Parameter als Faktoren gewählt werden. Um ein möglichst einfaches und technisch interpretierbares Modell zu erhalten, sollte man gemäß Absatz 3.3.2 überlegen, welche Faktoren einen direkten Zusammenhang zur Zielgröße „Sinkzeit“ haben. Dabei kann man z. B. folgende Überlegungen anstellen:

- Die Fläche der Flügel beträgt $F=2ac$. Sie behindert das Sinken.
- Das Gewicht des Rotors hängt von der Gesamtfläche $2a(c + d) + ef$ und ggf. dem Gewicht der Büroklammer ab. Das Gewicht zieht einerseits den Rotor nach unten, andererseits bewirkt es aber auch das Drehmoment, das die Rotation in Gang setzt.
- Zumindest bei großer Fallhöhe wird die Rotation durch den Luftwiderstand behindert; dieser ist umso größer, je größer a, d, e und f sind.
- Wenn d zu klein ist, wird der Rotor instabil und knickt, vor allem wenn die Flügelfläche groß ist.
- Wenn der Schwerpunkt des Rotors zu hoch liegt, trudelt er (hier hilft ggf. die Büroklammer).

Aufgrund dieser Überlegungen wird klar, dass das Flugverhalten nicht analytisch berechnet werden kann. Es muss empirisch bestimmt werden.

Als unabhängige Faktoren kommen z. B. die Längen a, c, d und f in Frage. Man könnte jedoch auch F, a, d und f oder andere Kombinationen verwenden. Da e vermutlich weniger direkten Einfluss auf die Eigenschaften des Rotors hat, dürfte e weniger gut als unabhängiger Faktor geeignet sein als c und d (wegen der Bedingung (20.2) sind nur zwei der drei Faktoren unabhängig).

Es ist durchaus typisch, dass nicht klar ist, welche Faktoren am besten geeignet sind. Trotzdem wird empfohlen, solche Überlegungen anzustellen, um offensichtlich ungeeignete Faktoren (wie e) zu vermeiden. Welche der sinnvollen Faktoren letztlich verwendet werden, ist nicht entscheidend, wenn mit dem Versuchsplan auch Wechselwirkungen bestimmt werden können.

Da die Effekte von vier Faktoren und ihren Wechselwirkungen untersucht werden sollen, bietet sich ein vollständiger faktorieller 2^4 -Plan an. Der Versuchsplan enthält 16 Faktorstufenkombinationen. Jede Faktorstufenkombination ist ein anderer Rotor.

Aufgrund der großen Zufallsstreuung muss jede dieser Kombinationen mehrmals realisiert werden. Was bedeutet hier mehrmalige Realisierung? Genügt es, denselben Rotor mehrmals fallen zu lassen und jeweils die Zeit zu messen? Oder muss für jede Realisierung ein neuer Rotor ausgeschnitten und gefaltet werden, der dann jeweils nur einmal fällt?

Statistisch ideal wäre natürlich, für jede Realisierung einen neuen Rotor auszuschneiden. Die zufälligen Unterschiede zwischen den Sinkzeiten von nominell identischen Rotoren beinhalten dann u. a.:

- zufällige Unterschiede im Material (Papier)
- zufällige Unterschiede in der ausgeschnittenen Form und Größe
- zufällige Unterschiede in der Faltung
- zufällige Unterschiede in Höhe und Haltung beim Loslassen des Rotors
- zufällige Unterschiede in der Luftströmung
- zufällige Unterschiede in der Zeitmessung (Stoppuhr).

Ein einfacher Test zeigt, dass bei entsprechender Sorgfalt die Unterschiede zwischen mehreren nominell identischen Rotoren wesentlich kleiner sind als die Unterschiede zwischen Einzelversuchen mit demselben Rotor. Daher genügt es in diesem Fall, 16 Rotoren nach dem Versuchsplan auszuschneiden und zu falten. Diese kann man z. B. mit der systematischen Nummer durchnumerieren und dann in den Einzelversuchen immer wieder verwenden.

Dadurch wird der Versuchsaufwand wesentlich reduziert. Man muss sich jedoch im Klaren sein, dass die Zufallsstreuung damit evtl. etwas unterschätzt wird, da die ersten drei Streuungsursachen nicht erfasst werden.

Die Einzelversuche können jedoch ohne großen Aufwand in randomisierten Blöcken durchgeführt werden:

- Zunächst lässt man jeden der 16 Rotoren einmal in zufälliger (aber festgelegter) Reihenfolge fallen
- anschließend wieder jeden einmal in einer anderen zufälligen Reihenfolge, usw.

Ändert sich z. B. aufgrund zunehmender Übung die Reaktionszeit, so kann diese Änderung als Unterschied zwischen den Blöcken erkannt und herausgerechnet werden.

Diese Hinweise sollten als Einstieg genügen. Es bleibt dem Leser nun überlassen, die Faktoren festzulegen, konkrete Stufenwerte für die Faktoren festzulegen, den Versuchsplan aufzustellen, die Einzelversuche durchzuführen und auszuwerten und Maßnahmen abzuleiten.

Erweiterungen über die beschriebene Aufgabenstellung hinaus sind natürlich jederzeit möglich. So sind z. B. weitere Faktoren denkbar, deren Einfluss auf die Sinkzeit des Rotors untersucht werden kann, wie z. B.

- Form der Flügel
- Abschrägungen an der Versteifung
- Papierart
- unterschiedliche Gewichte (z. B. Büroklammer, Heftklammern).

Nichtlinearitäten können mit zentral zusammengesetzten Versuchsplänen untersucht werden. Die Standardabweichung der Sinkzeit kann man als weitere Zielgröße behandeln (oder das Signal-Rausch-Verhältnis). Sie ist ein Maß für die Anfälligkeit des Rotors gegenüber Störungen.

20.3.2 Nürnberger Trichter

Dieses Beispiel (nach [3]) kann als Modell für eine Prozessoptimierung betrachtet werden: Eine Kugel rollt durch ein schräg stehendes Rohr in einen Trichter. Dort kreist sie zunächst. Allmählich verliert sie Energie und fällt schließlich aus dem Trichter heraus. Die Verweildauer der Kugel im Trichter (=Zielgröße) wird mit einer Stoppuhr gemessen. Sie soll maximiert werden.

Bild 20-5 zeigt die Anordnung schematisch. Für die Durchführung wird empfohlen, einen Stand zu bauen, an dem Rohr und Trichter befestigt werden. Je starrer die Anordnung ist, desto größer ist die Verweildauer der Kugel. Je mehr die Struktur zu Schwingungen neigt, desto überraschendere Effekte treten auf.

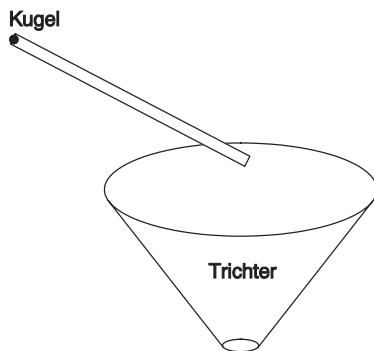


BILD 20-5

Eine Kugel rollt durch ein Rohr in den Trichter und kreist dann im Trichter. Die Verweildauer der Kugel im Trichter soll maximiert werden.

Viele Einflussgrößen können die Verweildauer beeinflussen, wie z.B.

- die Lage des unteren Rohrendes relativ zum Trichter
- die Steigung, die Länge und das Material des Rohrs
- der Winkel des Rohrs tangential zum Trichter
- die Größe und das Material der Kugel
- die Form, die Größe und das Material des Trichters.

Für den Anfang wird empfohlen, nur einige dieser Größen als Faktoren auszuwählen, z.B. den Abstand des Rohrendes vom oberen Trichterrand, die Steigung des Rohrs, sowie die Größe der Kugel.

Man erwartet eine große Verweildauer, wenn die Kugel schnell kreist und möglichst nahe am oberen Rand des Trichters beginnt. Dann ist aber auch das Risiko am größten, dass die Kugel oben wieder herausgeschleudert wird. Hier sollte bei der Optimierung auch die „Prozesssicherheit“ berücksichtigt werden.

Die Zielsetzung kann weiter verfeinert werden, indem man verschiedene Kugelgrößen mit unterschiedlichen Kosten belegt (je größer, desto teurer). Wird jeder Sekunde Verweildauer eine Einnahme zugeordnet, kann man den Nettogewinn maximieren.

Eine weitere Möglichkeit besteht darin, Kugeln aus unterschiedlichem Material zu verwenden und als Ziel festzulegen, dass der relative Unterschied in der Verweildauer der Kugeln möglichst klein sein soll. Diese Vorgehensweise simuliert die Entwicklung eines robusten Prozesses, dessen Ergebnis nur wenig vom verwendeten Ausgangsmaterial abhängt.

Je nach Zielsetzung sind unterschiedliche Faktorstufenkombinationen optimal. Dieses Beispiel unterstreicht damit auch, wie wichtig es ist, dass man sich zunächst über die Zielsetzung einig ist. In derselben Situation sind oft unterschiedliche Zielsetzungen möglich.

Literatur

- [1] *Kopas, D. A./McAllister, P. R.*: „Process Improvement Exercises for the Chemical Industry“ in: *The American Statistician* 46 (1992), 34 – 41
- [2] *Fowlkes, W. Y./Creveling, C. M.*: *Engineering Methods for Robust Product Design*. Addison-Wesley, Reading 1995
- [3] *Gunter, B.*: „Through a Funnel Slowly with Ball Bearing and Insight to Teach Experimental Design“ in: *The American Statistician* 47 (1993), 265 – 269

Anhang

■ Anhang A – Abkürzungen und Formelzeichen

| | |
|---------------|---|
| 2FWW | Zwei-Faktor-Wechselwirkung |
| 3FWW | Drei-Faktor-Wechselwirkung |
| b | Regressionskoeffizient (Schätzwert aus Stichprobe, Daten) |
| β | Regressionskoeffizient (wahrer Wert) |
| B | Bestimmtheitsmaß, Anteil der S. d. q. A., der durch Regression erklärt wird, vgl. (10.13) |
| \bar{d} | Differenz zweier Mittelwerte (aus Stichproben = Effekt), vgl. (6.12) |
| $\Delta\mu$ | technologisch relevanter Unterschied zweier Mittelwerte (Effekt) |
| EVOP | <u>E</u> nvolutionary <u>O</u> perations |
| f | Freiheitsgrad |
| FF | fraktioneller faktorieller (Versuchsplan) |
| IC | <u>I</u> ntegrated <u>C</u> ircuit (integrierte Schaltung) |
| m | Anzahl der Faktorstufenkombinationen eines Versuchsplans |
| μ | Mittelwert der Verteilung (wahrer Wert) |
| n | Stichprobenumfang (Anzahl der Realisierungen einer bestimmten Faktorstufen-kombination) |
| N | Versuchsumfang (Gesamtzahl der Einzelversuche) |
| p | Anteil fehlerhafter Einheiten |
| PB | <u>P</u> lackett- <u>B</u> urman (Versuchsplan) |
| PC | <u>P</u> ersonal <u>C</u> omputer |
| Q | Hilfsgrößen bei Regression, vgl. (10.8) – (10.12) |
| r | Korrelationskoeffizient, vgl. (10.15) |
| s | Stichprobenstandardabweichung (Stichprobenergebnis, Varianz s^2) |
| σ | Standardabweichung der Verteilung (wahrer Wert, Varianz σ^2) |
| $S_{\bar{d}}$ | Standardabweichung für Differenz zweier Mittelwerte (z. B. Effekt), vgl. (6.14) |
| S_R | Standardabweichung der Reststreuung (um Regressionsmodell), vgl. (10.17) |
| $S_{\bar{y}}$ | Standardabweichung des Mittelwerts \bar{y} , vgl. (6.8) |

- S.d.q.A. Summe der quadrierten Abweichungen (vgl. Q)
- S/N Signal-to-noise-ratio, Signal-Rausch-Verhältnis, vgl. (9.1)
- t t-Wert (aus Tabelle 6.4, abhängig von Freiheitsgrad f)
- x Wert einer Faktorstufe oder Einflussgröße (unabhängige Variable)
- \bar{x} Mittelwert der x-Werte (Stichprobenmittelwert), vgl. (10.6)
- y Wert einer Zielgröße, Versuchsergebnis (abhängige Variable)
- \bar{y} Mittelwert der y-Werte (Stichprobenmittelwert), vgl. (6.5 und 10.7)
- \hat{y} Schätzwert für y (mit einem an Daten angepassten Modell berechnet), vgl. (10.2)

Anhang B – Statistische Tabellen

TABELLE 6.4 t-Werte zur Berechnung zweiseitiger Vertrauensbereiche

| Freiheitsgrad f | t für Vertrauensniveau | | |
|-----------------|------------------------|-------|--------|
| | 95 % | 99 % | 99,9 % |
| 1 | 12,71 | 63,66 | 636,62 |
| 2 | 4,303 | 9,925 | 31,60 |
| 3 | 3,182 | 5,841 | 12,92 |
| 4 | 2,776 | 4,604 | 8,610 |
| 5 | 2,571 | 4,032 | 6,869 |
| 6 | 2,447 | 3,707 | 5,959 |
| 7 | 2,365 | 3,499 | 5,408 |
| 8 | 2,306 | 3,355 | 5,041 |
| 9 | 2,262 | 3,250 | 4,781 |
| 10 | 2,228 | 3,169 | 4,587 |
| 12 | 2,179 | 3,055 | 4,318 |
| 15 | 2,131 | 2,947 | 4,073 |
| 20 | 2,086 | 2,845 | 3,850 |
| 30 | 2,042 | 2,750 | 3,646 |
| 40 | 2,021 | 2,704 | 3,551 |
| 50 | 2,009 | 2,678 | 3,496 |
| 70 | 1,994 | 2,648 | 3,435 |
| 100 | 1,984 | 2,626 | 3,390 |
| 1000 | 1,962 | 2,581 | 3,300 |
| ∞ | 1,960 | 2,576 | 3,291 |

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 50 | ∞ |
|----------------------|------|------|------|------|------|------|------|------|----------|
| 1 | 161 | 199 | 216 | 225 | 230 | 242 | 248 | 252 | 254 |
| 2 | 18,5 | 19,0 | 19,2 | 19,2 | 19,3 | 19,4 | 19,4 | 19,5 | 19,5 |
| 3 | 10,1 | 9,55 | 9,28 | 9,12 | 9,01 | 8,79 | 8,66 | 8,58 | 8,53 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 5,96 | 5,80 | 5,70 | 5,63 |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,74 | 4,56 | 4,44 | 4,37 |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,06 | 3,87 | 3,75 | 3,67 |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,35 | 3,15 | 3,02 | 2,93 |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 2,98 | 2,77 | 2,64 | 2,54 |
| 20 | 4,35 | 3,49 | 3,10 | 2,87 | 2,71 | 2,35 | 2,12 | 1,97 | 1,84 |
| 50 | 4,03 | 3,18 | 2,79 | 2,56 | 2,40 | 2,03 | 1,78 | 1,60 | 1,44 |
| ∞ | 3,84 | 3,00 | 2,60 | 2,37 | 2,21 | 1,83 | 1,57 | 1,35 | 1,00 |

TABELLE 12.2

Kritische Werte

$F_{f_1; f_2; 0,95}$

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 50 | ∞ |
|----------------------|------|------|------|------|------|------|------|------|----------|
| 1 | 4052 | 4999 | 5404 | 5624 | 5764 | 6056 | 6209 | 6302 | 6366 |
| 2 | 98,5 | 99,0 | 99,2 | 99,3 | 99,3 | 99,4 | 99,4 | 99,5 | 99,5 |
| 3 | 34,1 | 30,8 | 29,5 | 28,7 | 28,2 | 27,2 | 26,7 | 26,4 | 26,1 |
| 4 | 21,2 | 18,0 | 16,7 | 16,0 | 15,5 | 14,5 | 14,0 | 13,7 | 13,5 |
| 5 | 16,3 | 13,3 | 12,1 | 11,4 | 11,0 | 10,1 | 9,55 | 9,24 | 9,02 |
| 6 | 13,7 | 10,9 | 9,78 | 9,15 | 8,75 | 7,87 | 7,40 | 7,09 | 6,88 |
| 8 | 11,3 | 8,65 | 7,59 | 7,01 | 6,63 | 5,81 | 5,36 | 5,07 | 4,86 |
| 10 | 10,0 | 7,56 | 6,55 | 5,99 | 5,64 | 4,85 | 4,41 | 4,12 | 3,91 |
| 20 | 8,10 | 5,85 | 4,94 | 4,43 | 4,10 | 3,37 | 2,94 | 2,64 | 2,42 |
| 50 | 7,17 | 5,06 | 4,20 | 3,72 | 3,41 | 2,70 | 2,27 | 1,95 | 1,68 |
| ∞ | 6,63 | 4,61 | 3,78 | 3,32 | 3,02 | 2,32 | 1,88 | 1,52 | 1,00 |

TABELLE 12.3

Kritische Werte

$F_{f_1; f_2; 0,99}$

| $f_2 \backslash f_1$ | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 50 | ∞ |
|----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 1 | 4×10^5 | 5×10^5 | 5×10^5 | 6×10^5 |
| 2 | 998 | 999 | 999 | 999 | 999 | 999 | 999 | 999 | 999 |
| 3 | 167 | 148 | 141 | 137 | 135 | 129 | 126 | 125 | 123 |
| 4 | 74,1 | 61,2 | 56,2 | 53,4 | 51,7 | 48,1 | 46,1 | 44,9 | 44,0 |
| 5 | 47,2 | 37,1 | 33,2 | 31,1 | 29,8 | 26,9 | 25,4 | 24,4 | 23,8 |
| 6 | 35,5 | 27,0 | 23,7 | 21,9 | 20,8 | 18,4 | 17,1 | 16,3 | 15,7 |
| 8 | 25,4 | 18,5 | 15,8 | 14,4 | 13,5 | 11,5 | 10,5 | 9,80 | 9,33 |
| 10 | 21,0 | 14,9 | 12,6 | 11,3 | 10,5 | 8,75 | 7,80 | 7,19 | 6,76 |
| 20 | 14,8 | 9,95 | 8,10 | 7,10 | 6,46 | 5,08 | 4,29 | 3,77 | 3,38 |
| 50 | 12,2 | 7,96 | 6,34 | 5,46 | 4,90 | 3,67 | 2,95 | 2,44 | 2,03 |
| ∞ | 10,8 | 6,91 | 5,42 | 4,62 | 4,10 | 2,96 | 2,27 | 1,73 | 1,00 |

TABELLE 12.4

Kritische Werte

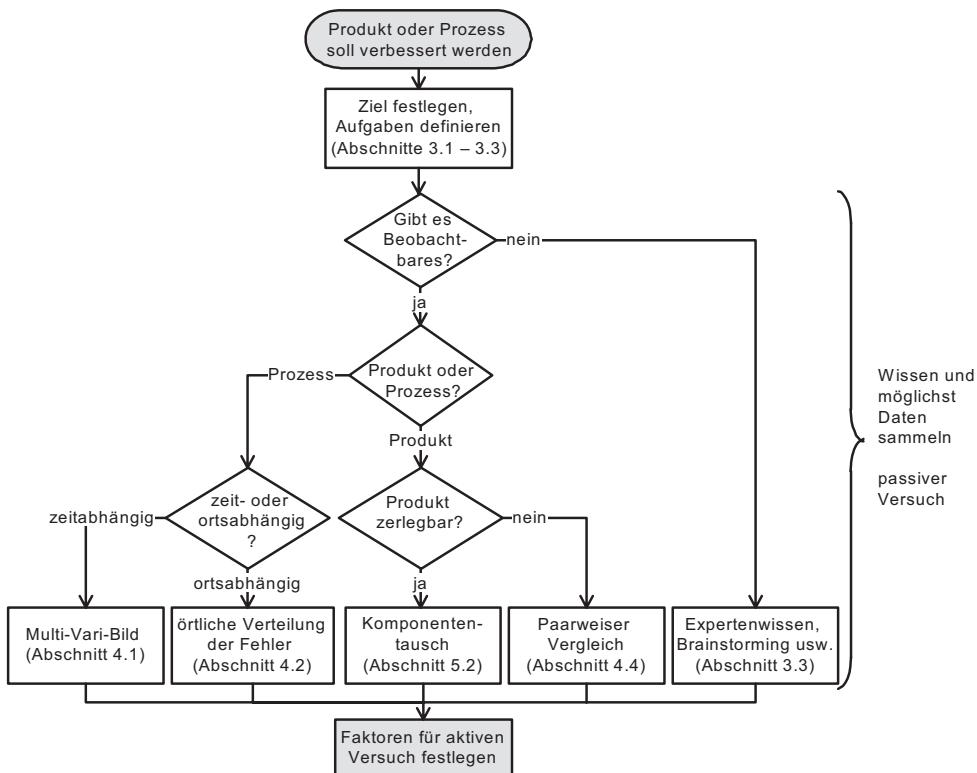
$F_{f_1; f_2; 0,999}$

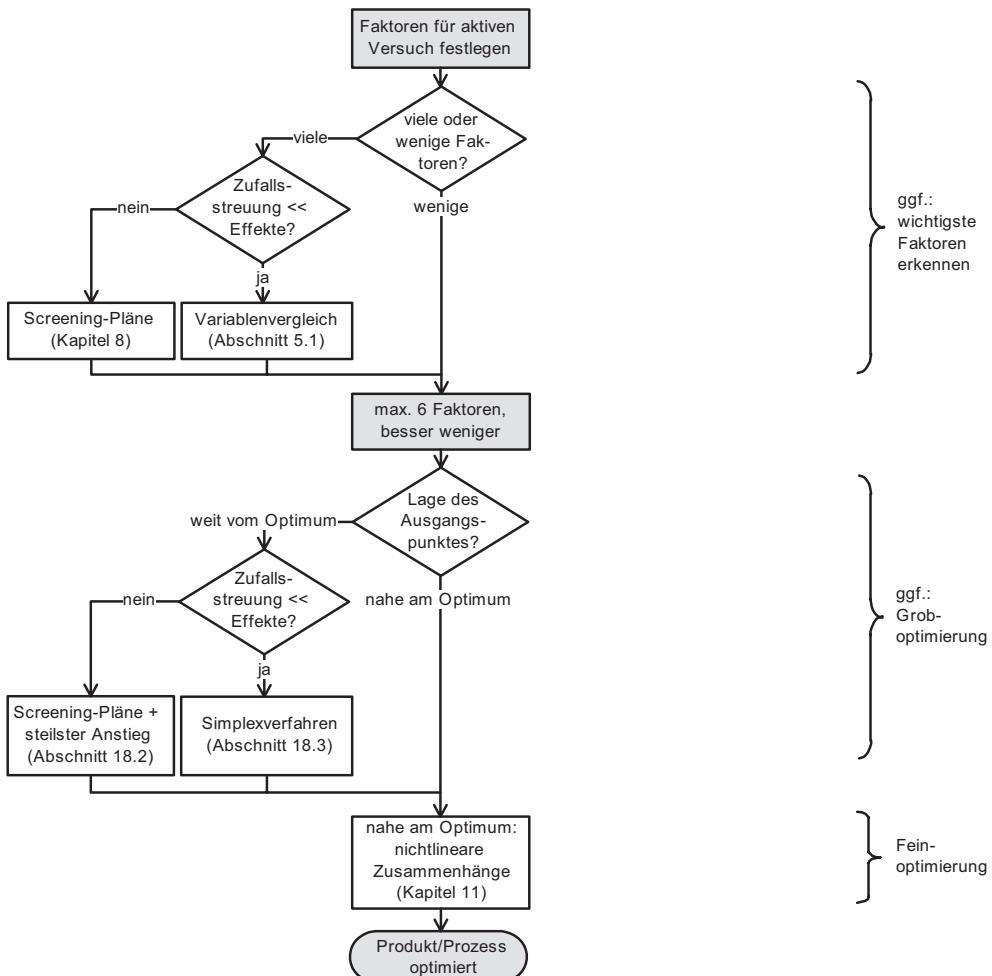
■ Anhang C – Wegweiser durch die Verfahren

Produkt- und Prozessverbesserung ist meist ein mehrstufiger Ablauf. Das folgende Flussdiagramm gibt einen Überblick über den Gesamtablauf und soll bei der Auswahl des jeweils am besten geeigneten Verfahrens helfen. Der Ablauf gliedert sich grob in drei Stufen:

1. Auswahl der Faktoren für den aktiven Versuch durch systematische Beobachtung und Expertenwissen
2. Auswahl der wichtigsten Faktoren aus einer größeren Anzahl (falls erforderlich)
3. Optimierung

Bei den Stufen 2 und 3 hängt das geeignete Verfahren davon ab, ob die Zufallsstreuung vernachlässigbar ist oder nicht.





■ Anhang D – Ablauf einer Versuchsplanung

- Ausgangssituation beschreiben (Abschnitt 3.1)
 - Wer ist Kunde und was braucht er?
 - Wie viel Zeit und Geld stehen maximal zur Verfügung?
 - Wer ist betroffen? – Alle einbinden!
 - Wer ist verantwortlich?
 - Was ist bereits bekannt? (Prozessablauf, systematische Beobachtung)
 - Untersuchungsziel in Abstimmung festlegen (Abschnitt 3.2)
 - Zielgrößen und Faktoren festlegen (Abschnitt 3.3)
 - Zielgrößen vor allem nach Kundenrelevanz festlegen – bei Ersatzgrößen muss Zusammenhang mit der eigentlichen Zielgröße abgesichert sein

- Auswahl der Faktoren ist entscheidend für den Erfolg – Brainstorming, Einflussgrößen-Zielgrößen-Matrix, Fachwissen, systematische Beobachtung (Kapitel 4), Auswertung bereits vorhandener Daten (Grafiken + mehrfache Regression, Kapitel 10)
- Sinnvoller Wertebereich – nicht zu schmal und nicht zu breit, Vorversuche
- Nicht im Versuchsplan veränderte Größen möglichst konstant halten, evtl. Blockfaktoren verwenden
- Versuchsplan aufstellen (Abschnitt 3.4)
 - Für den Einstieg eignen sich z. B. 2^k -Pläne – möglichst mit Zentrumspunkt (Kapitel 7 bzw. 16.2) und zentral zusammengesetzte Pläne (für quantitative Zusammenhänge im Detail – Kapitel 11)
 - Auf ausreichend großen Versuchsumfang achten, möglichst mit Blockbildung und Randomisierung
 - Aufwand bzw. Kosten-Nutzen-Verhältnis beachten
- Versuche vorbereiten (Absatz 3.5.1)
 - Ressourcen planen
 - Messgeräte und Messverfahren überprüfen
 - Teile auswählen, kennzeichnen, zuordnen
 - Versuchs- und Messablauf festlegen
 - Personal einweisen
 - Pilotversuch durchführen (Versuchsablauf und Extremkombinationen im Plan testen)
- Versuche durchführen (Absatz 3.5.2)
 - Sorgfalt, Streuung minimieren, ausreichend Zeit für Einschwingen lassen
 - Besonderheiten und Abweichungen dokumentieren, selbst dabei bleiben
 - Auswertung mit Ist-Werten, nicht mit Soll-Werten
 - Umgebungsbedingungen erfassen

■ Anhang E – Ablauf einer Datenauswertung

Zunächst für jede Zielgröße einzeln

- Versuchsergebnisse auf Vollständigkeit und Plausibilität überprüfen, Fehler bei Dateneingabe, Zuordnung von Ergebnissen zu Faktorstufenkombinationen, Messfehler und andere direkt erkennbare Probleme sofort beheben
- Geeignetes Modell an die Daten anpassen – ausgehend von einem umfassenden Modell (z. B. quadratisch) nicht signifikante Terme entfernen, dabei auf Hierarchie achten (Faktoren erst entfernen, wenn alle WW mit diesem Faktor und der quadratische Term in diesem Faktor bereits entfernt wurden)
- Diagnose der Residuen
 - Ausreißer im Wahrscheinlichkeitsnetz identifizieren, auf mögliche Ursachen überprüfen und ggf. entfernen (Absatz 6.2.2)

- Residuen gegen Versuchsnummer auftragen, Ursache für evtl. vorhandenen Trend identifizieren und ggf. korrigieren, evtl. die Ursache für den Trend als weiteren Faktor behandeln (Absatz 10.1.3)
 - Residuen gegen alle Faktoren auftragen und auf systematische Abweichungen vom Modell überprüfen, Modell ggf. erweitern (Absatz 10.1.3)
 - Transformation der Zielgröße erwägen: Bei großem Wertebereich und nur positiven Werten (z.B. Lebensdauern) Logarithmus, bei Zähldaten Wurzel, Box-Cox-Transformation bietet Entscheidungshilfe (Abschnitt 6.4)

■ Falls eine Transformation durchgeführt wurde, werden Modellanpassung und Residuendiagnose wiederholt

■ Grafiken des angepassten Modells mit technischem Wissen auf Plausibilität überprüfen, Zusammenhänge explorieren, Ergebnisse verstehen

Nachdem für jede Zielgröße ein Modell angepasst wurde, das die Daten gut beschreibt und verstanden ist, folgt die Optimierung, die meist eine Kompromisssuche bei widersprüchlichen Zielen beinhaltet (Abschnitte 15.3, 3.7 und 3.8).

- Wunschfunktion für jede Zielgröße definieren (ggf. auch für die Faktoren)
 - Numerisch optimieren, erreichtes Optimum inhaltlich bewerten
 - Ggf. Wunschfunktion anpassen, entspannen oder verschärfen
 - Iterativ suchen, bis ein technisch befriedigender Kompromiss gefunden ist (oder klar ist, dass noch weitere Verbesserungen erforderlich sind – dann folgt eine neue Versuchsplanung)
 - Zielgrößen in Abhängigkeit von den wichtigsten Faktoren grafisch darstellen, andere Faktoren dabei auf Optimalwert halten – am einfachsten mit der Optimierungsgrafik Bild 15-4, ggf. von Hand nachoptimieren
 - Vorschlag für Prozessverbesserung formulieren
 - Bestätigungsversuch bzw. Probeumstellung durchführen, Ergebnisse mit den berechneten Werten vergleichen (dazu Vorhersagebereich verwenden), Dokumentation

■ Anhang F – Glossar deutsch/englisch

Als Arbeitshilfe für das Buch werden im Folgenden die wichtigsten Begriffe aus Versuchsplanung und zugrunde liegender Statistik kurz erläutert. Um den Umgang mit deutscher und englischer Software und Literatur zu erleichtern werden auch häufig verwendete Alternativbegriffe auf Deutsch und Englisch genannt. Der Verweis führt zur wichtigsten Stelle im Buch mit weiteren Informationen.

Auflösung eines Versuchsplans (Lösungstyp; resolution, type) Absatz 8.2.5

Beschreibt die schlimmste Vermengung von Effekten und damit das Risiko einer Verfälschung. Pläne der Auflösung III sind kritisch, V und mehr unkritisch.

| | |
|--|--------------------------|
| Ausreißer (outlier) | Abschnitt 3.6 |
| Einzelnes Versuchsergebnis, das nicht zu den anderen bzw. zum angepassten Modell passt, z. B. aufgrund eines Versuchs-, Mess- oder Übertragungsfehlers. | |
| Bestimmtheitsmaß ($B = r^2$; r-squared, R-Squared) | Absatz 10.1.2 |
| Anteil der Summe der quadrierten Abweichungen der Versuchsergebnisse vom Mittelwert, der durch das angepasste Modell erklärt wird. Je näher an 1 bzw. 100 %, desto besser. | |
| Blockbildung (blocking) | Absätze 3.4.3 und 7.1.4 |
| Einteilung der Einzelversuche in möglichst homogene Gruppen, um so die Zufallsstreuung innerhalb der Gruppen klein zu halten. Ein Blockfaktor beschreibt die Gruppenzugehörigkeit. | |
| Effekt (effect) | Absätze 7.1.1 und 10.1.5 |
| Maß für die Wirkung eines Faktors bzw. einer Wechselwirkung. Der Effekt eines Faktors gibt an, um wie viel sich die Mittelwerte der Versuchsergebnisse bei der einen Faktorstufe von denen der anderen unterscheiden. Alternativ werden häufig Regressionskoeffizienten in normierten Stufenwerten angegeben, diese sind genau Effekt/2. | |
| Einflussgröße (Einflussparameter; input variable, cause) | Abschnitte 2.2 und 3.3 |
| Größe, die die Versuchsergebnisse (Zielgrößen) möglicherweise beeinflusst. Im Versuch werden Einflussgrößen entweder als Faktoren gezielt verändert oder auf einem definierten Wert konstant gehalten (falls dies nicht möglich ist, wenigstens Wert für jeden Einzelversuch erfassen und aufzeichnen). | |
| Hinweis: In manchen Büchern werden die Begriffe Einflussgröße und Faktor als Synonyme behandelt oder zur Unterscheidung quantitativ/qualitativ verwendet. | |
| Einzelversuch (run, experimental treatment, trial) | Tabelle 3.3 |
| Einzelne Zeile im Versuchsplan, einzelne Faktorstufenkombination. | |
| Faktor (Variation, Versuchsparameter; factor) | Abschnitte 2.5 und 3.3 |
| Einflussgröße, die im Versuch gezielt verändert wird (auf mindestens zwei Stufen). | |
| Faktorstufe (Einstellung; factor level) | Abschnitt 2.6 |
| Ausgewählte Werte, die der Faktor im Versuch annehmen soll. | |
| Fehler 1. Art (α , α -Fehler; type I error, α -risk) | Abschnitt 6.3 |
| Die Nullhypothese wird verworfen, obwohl sie zutrifft. Ein Unterschied oder Effekt wird als signifikant beurteilt, obwohl er nicht vorhanden ist. | |
| Fehler 2. Art (β , β -Fehler; type II error, β -risk) | Abschnitt 6.3 |
| Die Nullhypothese wird beibehalten, obwohl sie nicht zutrifft. Ein tatsächlich vorhandener Unterschied oder Effekt wird nicht entdeckt (geht in der Zufallsstreuung unter). | |

| | |
|--|-------------------------|
| Mittelwert (Durchschnitt; mean, average) | Abschnitte 6.1 und 6.2 |
| Kennzahl für die Lage von Werten. Der Mittelwert μ einer Verteilung ist der Schwerpunkt der Fläche unter der Verteilungskurve (fest, aber normalerweise nicht bekannt). Der Mittelwert \bar{y} einer Stichprobe ist der Mittelwert der beobachteten Einzelwerte. Bei jeder Stichprobe erhält man einen anderen Mittelwert. | |
| Qualitativer Faktor (kategorisch, nominal; categorical, nominal) | Abschnitt 2.7 |
| Für die Stufenwerte gibt es nur Namen, Beschreibungen oder Bezeichnungen. | |
| Quantitativer Faktor (numerisch, kontinuierlich; numeric, continuous) | Absch. 2.7 |
| Die Stufenwerte werden mit Zahlenwerten auf einer Messskala beschrieben. | |
| Randomisierung (randomization) | Absätze 3.4.4 und 7.1.4 |
| Zufällige Reihenfolge der Einzelversuche bzw. zufällige Zuordnung von Teilen zu Versuchsbedingungen, um zu vermeiden, dass unerwartete und unerkannte Veränderungen bzw. Unterschiede die Ergebnisse verfälschen. | |
| Regressionsanalyse (regression) | Kapitel 10 |
| Verfahren zur Anpassung eines vorgegebenen Modells für den Wirkzusammenhang zwischen Faktoren und Zielgröße an vorliegende Daten. In der hier beschriebenen Form wird dazu die Summe der quadrierten Abweichungen der Daten vom Modell in die y-Richtung minimiert. Regressionsanalyse ist nur bei quantitativen Faktoren direkt anwendbar, bei qualitativen Faktoren werden manchmal Hilfskonstruktionen verwendet. | |
| Regressionskoeffizient (regression coefficient) | Kapitel 10 |
| Aus den Daten bestimmte Schätzwerte für die Parameter im vorgegebenen Regressionsmodell (siehe Regressionsanalyse). | |
| Residuum (residual) | Absatz 10.1.3 |
| Abweichung eines einzelnen Versuchsergebnisses vom angepassten Modell (z.B. Regressionsmodell). | |
| Signifikanz (significance) | Abschnitt 6.3 |
| Ein Unterschied, Effekt oder Regressionskoeffizient ist signifikant, wenn er aus der Zufallsstreuung herausragt. Als Entscheidungshilfe zur Beurteilung der Signifikanz wird meist der p-Wert (p-value) verwendet. Er gibt an, mit welcher Wahrscheinlichkeit der beobachtete Effekt oder ein noch größerer Effekt zufällig auftritt. Nicht signifikante Effekte sind nur dann auch nicht relevant, wenn ein ausreichender Versuchsumfang N verwendet wurde. | |
| Standardabweichung (standard deviation) | Abschnitte 6.1 und 6.2 |
| Kennzahl für die Breite einer Verteilung. Die Standardabweichung σ einer Normalverteilung ist die Breite bis zur steilsten Stelle der Verteilungskurve (fest, aber normalerweise nicht bekannt). Die Standardabweichung s einer Stichprobe ist ein Schätzwert für die Breite der Verteilung (nicht erwartungstreu). Bei jeder Stichprobe erhält man eine andere Standardabweichung s . | |

Steuergröße (control parameter) Abschnitt 2.3

Einflussgröße, deren Wert für das Produkt bzw. den Fertigungsprozess auf einen bestimmten Wert eingestellt und dort (in gewissen Grenzen) gehalten werden kann (Konstruktions- bzw. Prozessparameter).

Störgröße (noise parameter) Abschnitt 2.4

Einflussgröße, deren Wert für das Produkt bzw. den Fertigungsprozess nicht vorgegeben werden kann (oder z. B. aus Kostengründen nicht vorgegeben werden soll).

Varianz (variance) Abschnitte 6.1 und 6.2

Kennzahl für die Breite einer Verteilung, Quadrat der Standardabweichung.
Vorteil der Varianz ist, dass die Varianz s^2 der Stichprobe ein erwartungstreuer Schätzwert für die Varianz σ^2 der Verteilung ist. Leider ist die Varianz nicht so anschaulich wie die Standardabweichung.

Varianzanalyse (ANOVA = ANalysis Of VAriance) Kapitel 12

Verfahren zur Beurteilung der Signifikanz von Effekten, das auf einer Streuungszerlegung beruht. Die Varianz der Effekte wird mit der Varianz der Zufallsstreuung verglichen. Sie ist auch bei mehrstufigen qualitativen Faktoren anwendbar.

Vermengung (Alias; confounding, alias) Abschnitt 8.2

Miteinander vermengte Effekte können nicht getrennt voneinander bestimmt werden, man erhält immer nur ihre Summe. Das damit verbundene Risiko wird mit der Auflösung beschrieben.

Versuchsplan (experimental design) Tabelle 3.3

Liste aller geplanten Einzelversuche in definierter Reihenfolge.

Versuchsumfang N (number of experimental runs) Absatz 6.3.2

Gesamtzahl der geplanten Einzelversuche.

Vertrauensbereich (Konfidenzintervall; confidence interval) Absatz 6.2.4

Bereich, der den wahren Wert für die betrachtete Größe (z. B. Mittelwert, Effekt, Regressionskoeffizient) auf dem vorgegebenen Vertrauensniveau einschließt. Aus jedem Versuch erhält man einen anderen Vertrauensbereich, auf lange Sicht enthält der vorgegebene Anteil den wahren Wert. Achtung: Der Vertrauensbereich für den Mittelwert macht keine Aussage über Einzelwerte (s. u.).

Vorhersagebereich (Prognoseintervall; prediction interval) Absatz 10.1.4

Grundsätzlich ähnlich zum Vertrauensbereich für den Mittelwert, aber der Vorhersagebereich macht eine Aussage über den nächsten Einzelwert. Er ist immer breiter als der Vertrauensbereich, da er zusätzlich die Zufallsstreuung der Einzelwerte berücksichtigt. Hinweis: Der Toleranzbereich ist noch breiter und macht eine Aussage über einen vorgegebenen Anteil der Einzelwerte.

Wechselwirkung (interaction) Absatz 7.1.3

Eine signifikante Zwei-Faktor-Wechselwirkung AB bedeutet, dass der Effekt des Faktors A auf die betrachtete Zielgröße vom Wert (Stufe) des Faktors B abhängt und umgekehrt, d. h. dass A und B sich in ihrer Wirkung auf die Zielgröße gegenseitig beeinflussen. Analog für Drei-Faktor-Wechselwirkungen usw.

Wunschfunktion (desirability function) Abschnitt 15.3

Verfahren, um systematisch nach Kompromissen zwischen mehreren Zielgrößen zu suchen. Die Wunschfunktion bildet den Wunsch des Anwenders auf eine Zahl zwischen 0 (nicht zulässig) und 1 (optimal) ab. Dann kann der Rechner nach einem Kompromiss suchen.

Zielgröße (response, output) Abschnitte 2.1 und 3.3

Beschreibt das Ergebnis eines Versuchs. Die Zielgrößenwerte sind Messwerte (auch Noten) oder aus Messwerten berechnet. Meist benötigt man mehrere Zielgrößen um alle Ziele eines Versuchs abzubilden.

■ Anhang G – JavaScript-Visualisierungen zum Herunterladen

Unter dem Reiter Extras können Sie sich die Dateien zum Buch von der folgenden Webseite herunterladen (als zip-Datei, bitte zunächst entpacken):

<https://www.hanser-fachbuch.de/buch/Versuchsplanung/9783446461468>

Viele Anwender finden die statistische Denkweise zunächst abstrakt und ungewohnt. Daher werden im Verzeichnis (Ordner) „Visualisierungen“ fünf JavaScript-Visualisierungen zusammengestellt, die ausgewählte Zusammenhänge aus der Statistik visuell vermitteln, in der Erwartung,

- dass die Verfahren und Konzepte aus der Statistik besser verständlich sind, wenn man sie „sehen“ kann und
- dass eigenes Ausprobieren und Verändern von Parametern das Lernen erleichtern.

Die Visualisierungen wurden von Studierenden des Studiengangs Elektrotechnik der Hochschule Aalen im Rahmen von Projektarbeiten erstellt. Vielen Dank für die Erstellung und die Genehmigung zur Veröffentlichung. Auf der Internetseite

https://www.hs-aalen.de/de/pages/statistik-erleben_applets

finden Sie die Originalarbeiten. Zum Herunterladen für dieses Buch wurden nur die Visualisierungen ausgewählt, die für die hier behandelten Themen relevant sind.

Sicherheitshinweis

JavaScript kann im Prinzip Schäden verursachen. Mir sind keine negativen Auswirkungen der Visualisierungen bekannt, trotzdem können weder ich, noch die Autoren der Visualisierungen, noch der Carl Hanser Verlag irgendwelche Haftung übernehmen.

Start der Visualisierungen

Um die Visualisierungen anzuzeigen, benötigen Sie eine Internetverbindung. Starten Sie zunächst die Datei **start.html** im Verzeichnis „Visualisierungen“. Sie erhalten einen Überblick über die fünf Visualisierungen und können das gewünschte Thema auswählen. Für jedes Thema erhalten Sie dann

- weitere Informationen
- Vorschläge für Übungen
- eine Bedienungsanleitung
- einen Link zum Starten der Visualisierung

Startet die Visualisierung nicht, so kann dies am verwendeten Browser oder an den Sicherheitseinstellungen Ihres Rechners liegen.

Normalverteilte Zufallszahlen (zu Abschnitten 6.1 und 6.2)

Mit Hilfe dieser Visualisierung können Sie viele Stichproben aus derselben normalverteilten Grundgesamtheit erzeugen. Mittelwert μ und Standardabweichung σ der Grundgesamtheit, sowie Umfang n und Anzahl der Stichproben geben Sie vor.

Die Visualisierung zeigt, wie Messwerte von Stichprobe zu Stichprobe streuen. Damit streuen auch aus den Messwerten berechnete Kenngrößen wie Mittelwert und Standardabweichung der Stichprobe. Trotzdem kann die Verteilung dieser Kenngrößen berechnet werden. Der Mittelwert über viele Mittelwerte nähert sich immer mehr dem wahren (hier vorgegebenen) Mittelwert – man sagt, er ist ein erwartungstreuer Schätzwert. Dagegen nähert sich der Mittelwert über viele Standardabweichungen nicht der wahren Standardabweichung – sie ist kein erwartungstreuer Schätzwert.

Die Mittelwerte, die Varianzen und die Standardabweichungen vieler Stichproben werden berechnet und grafisch dargestellt. Die Grafiken zeigen, dass

- die Stichprobenergebnisse streuen
- Mittelwerte weniger streuen als Einzelwerte
- die Verteilung der Einzelwerte, Mittelwerte, Varianzen und Standardabweichungen über viele Stichproben berechnet werden kann
- der Mittelwert (Gl. 6.5) und die Varianz (Gl. 6.6) erwartungstreue Schätzwerte sind, die Varianz mit $1/n$ statt $1/(n-1)$ und die Standardabweichung aber nicht.

Autor: Jan Strohbeck

Vertrauensbereich für den Mittelwert (zu Abschnitt 6.2)

Diese Visualisierung soll die Bedeutung des Vertrauensbereichs für den Mittelwert verdeutlichen.

Versuchsergebnisse, Messwerte usw. streuen aufgrund zufälliger Einflüsse. Daher unterliegen auch die aus mehreren Ergebnissen berechneten Mittelwerte und Standardabweichungen dem Zufall. Aus jeder Messreihe (Stichprobe) erhalten Sie einen etwas anderen Mittelwert (vgl. Normalverteilte Zufallszahlen).

Doch wo liegt der wahre, unbekannte Mittelwert μ der Grundgesamtheit (d. h. der Mittel-

wert, den Sie erhielten, wenn Sie den Versuch unter immer gleichen Bedingungen unendlich oft wiederholen könnten)? Der 95 %-Vertrauensbereich für den Mittelwert ist ein Bereich, der bei durchschnittlich 95 % der Stichproben den wahren Mittelwert der Grundgesamtheit enthält.

Die Visualisierung erzeugt Stichproben mit frei wählbarem Umfang n aus einer Normalverteilung, deren Mittelwert μ und Standardabweichung σ vorgegeben werden. Sie berechnet dann für jede Stichprobe Mittelwert und Standardabweichung dieser Stichprobe, sowie den 95 %- und 99 %-Vertrauensbereich für den Mittelwert. Wahlweise werden nur die Daten oder die Daten zusammen mit dem 95 %- und 99 %-Vertrauensbereich oder die Vertrauensbereiche für 20 Stichproben gleichzeitig grafisch dargestellt.

- Aus jeder Stichprobe erhält man einen anderen Vertrauensbereich, da Mittelwert und Standardabweichung der Stichprobe dem Zufall unterliegen.
- Durchschnittlich 95 % der 95 %-Vertrauensbereiche (bzw. 99 % der 99 %-Vertrauensbereiche) enthalten den wahren Mittelwert (das Maximum der Kurve).
- Der 99 %-Vertrauensbereich ist immer breiter als der 95 %-Vertrauensbereich.
- Je größer der Stichprobenumfang, desto schmäler ist der Vertrauensbereich für den Mittelwert. Die Einzelwerte streuen jedoch in einem breiteren Bereich – der Vertrauensbereich ist eine Aussage über die Lage des Maximums der Verteilung, nicht über die Breite der Verteilung oder die Einzelwerte.

Autor: Jan Strohbeck

Vergleich von zwei Mittelwerten (zu Abschnitt 6.3)

Diese Visualisierung soll die Problematik des Vergleichs von zwei Mittelwerten nahebringen. Häufig sollen zwei Lieferchargen, Fertigungsverfahren, Prozess- oder Produktvarianten verglichen werden. Man möchte wissen, ob sich die wahren aber unbekannten Mittelwerte der beiden Gruppen (= Grundgesamtheiten) unterscheiden oder nicht. Dazu wird aus jeder Gruppe eine Stichprobe entnommen.

Die Differenz der Stichprobenmittelwerte \bar{d} ist ein Schätzwert für die wahre Differenz $\Delta\mu$ der Mittelwerte der Grundgesamtheiten – sie unterliegt aber der Zufallsstreuung. Daher stellt sich die Frage: Unterscheiden sich die Mittelwerte der beiden Grundgesamtheiten wirklich oder ist der beobachtete Unterschied der Stichprobenmittelwerte nur auf die Zufallsstreuung zurückzuführen?

In der Visualisierung wird unterstellt, dass Stichproben vom gleichen Umfang n aus zwei Grundgesamtheiten gezogen werden, deren Mittelwerte sich um die wahre Differenz $\Delta\mu$ unterscheiden. Die wahre Standardabweichung σ ist für beide gleich.

Zwischen zwei Darstellungen kann umgeschaltet werden:

- „Verteilungen“ enthält die Verteilung
 - der Einzelwerte in den beiden Grundgesamtheiten
 - die Verteilung von \bar{d} , wenn $\Delta\mu = 0$ und
 - die Verteilung von \bar{d} für den vorgegebenen Wert für $\Delta\mu$.
- „Stichprobenergebnisse“ enthält die aus den Ergebnissen von je 20 Stichproben berechnete Differenz der Mittelwerte \bar{d} und 95 %- und 99 %-Vertrauensbereiche für die wahre Differenz $\Delta\mu$.

Beide Darstellungen zeigen unten die Wahrscheinlichkeit dafür, dass der Vertrauensbereich für die Differenz der Mittelwerte die 0 enthält, d. h. dass der Unterschied nicht signifikant ist (Fehler 2. Art). Vergleichen Sie die Darstellung der Verteilungen und der Stichprobenergebnisse und beobachten Sie interaktiv, wie sich die Veränderung der Parameter n , $\Delta\mu$ und σ auswirkt. So erleben Sie den Mittelwertvergleich und verstehen besser, worauf bei der Planung zu achten ist und was die Ergebnisse bedeuten.

Autor: Jan Strohbeck

Lineare Regression (zu Abschnitt 10.1 und Absatz 10.2.4)

Mit dieser Visualisierung können Sie Punkte in ein Koordinatensystem einzeichnen – oder Sie verwenden im Menü vorgegebene Beispiele wie in Tabelle 10.1 „Schichtdicke gegen Zeit“ als Ausgangspunkt. Sie können eine Gerade von Hand („by Eye“) einzeichnen. „Optimal“ zeigt die Regressionsgerade.

Die Regressionsgerade minimiert die Summe der quadrierten Abweichungen (S. d. q. A.). „ Q_{Rest} “ zeigt die quadrierten Abweichungen als Flächen. So wird das Prinzip der kleinsten Quadrate und das Bestimmtheitsmaß verdeutlicht.

Auch der Vertrauensbereich für den Mittelwert und der Vorhersagebereich für Einzelpunkte wie in Bild 10-10 kann dargestellt werden. „Residuendiagnose“ zeigt Residuen, Hebel und Cook-Distanz aller Punkte wie in Absatz 10.2.4. „Quadratsummen“ zeigt die Zerlegung von $Q_{yy} = Q_{\text{Gesamt}}$ in $Q_{\text{Regression}}$ und Q_{Rest} wie in Absatz 10.1.2 und zum Vergleich PRESS aus Absatz 10.2.4. „Lack of Fit“ zeigt die Zerlegung von Q_{Rest} in $Q_{\text{Lack of Fit}}$ und $Q_{\text{Pure Error}}$ wie in Absatz 10.2.4.

Sie können Punkte verschieben, löschen oder hinzufügen und die Wirkung auf die Gerade und alle anderen Darstellungen beobachten. Ein einzelner Punkt, dessen x-Wert wesentlich von den anderen abweicht, beeinflusst die Gerade sehr stark – er hat einen großen Hebel und, wenn er von der Geraden durch die anderen Punkte abweicht, auch eine große Cook-Distanz und einen großen Beitrag zu PRESS – aber nur ein kleines Residuum.

Die Beispiele „Beton“ und „Bremsweg gegen Geschwindigkeit“ eignen sich für die Beobachtung des Lack of Fit.

Autor: Philipp Pfund

EVOP – Evolutionary Operations (zu Abschnitten 7.1, 18.1 und 18.2)

Bei EVOP werden während der Fertigung Prozessparameter im Rahmen ihrer Spezifikation systematisch verändert. Ziel ist es, bei laufender Produktion Möglichkeiten zur Verbesserung des Fertigungsprozesses zu erkennen. Als Beispiel wird hier ein Spritzgießprozess betrachtet.

Mit der Visualisierung spielen Sie Fertigungsleiter. Sie verändern zwei Prozessparameter (die Faktoren Einspritzgeschwindigkeit und Nachdruck), beobachten die Veränderung der Zielgrößen Verzug und Schlagzähigkeit und entscheiden, wie Sie Ihren Fertigungsprozess verändern wollen, um noch besser zu werden. Erleben Sie selbst, wie die Produktion nicht nur Produkte sondern gleichzeitig Wissen über Verbesserungsmöglichkeiten produziert.

Als Versuchsplan wird ein 2²-Plan mit Zentrumspunkt verwendet. Pfeile markieren die Richtung der steepesten Verbesserung, sobald sie signifikant ist (wie in Abschnitt 18.2).

Autor: Üzeyir Tektaş

■ Anhang H – Beispiele zum Herunterladen

Um den Einstieg in Versuchsplanung mit Softwareunterstützung zu erleichtern, können Sie sich unter dem Reiter Extras hier folgende Dateien als zip-Datei herunterladen:

<https://www.hanser-fachbuch.de/buch/Versuchsplanung/9783446461468>

- die Daten der Beispiele im Dateiformat verschiedener Programme,
- die in Anhang G beschriebenen Visualisierungen
- die Datei Informationen.html.

Informationen.html

= html-Datei mit der Inhaltsangabe des Downloads und Links zu Informationen der Hersteller über ihre Programme (siehe Anhang J)

Besteht eine Verbindung zum Internet, so führen diese Links direkt zu den Herstellern. Bei vielen Programmen gibt es Downloads von zeitlich limitierten Vollversionen oder leicht eingeschränkten Testversionen, für fast alle Programme können solche Versionen kostenlos bestellt werden.

Buchbeispiele

= Verzeichnis mit den Daten der Beispiele und Aufgaben in diesem Buch

Für jedes Programm gibt es einen Unterordner mit dem Programmnamen, der die Daten **typischer** Beispiele enthält. Zusätzlich gibt es die Unterordner „ASCII“ und „EXCEL“, die die Daten **aller** Aufgaben und Beispiele als ASCII-Datei bzw. als EXCEL®-2003-Tabelle enthalten. Diese Form ermöglicht es

- mit denselben Daten die Auswertung und Darstellung der Ergebnisse durch verschiedene Programme direkt zu vergleichen
- den Import weiterer Daten in diese Programme zu vergleichen – aus EXCEL® am leichtesten über die Zwischenablage, manche Programme haben Importfilter für EXCEL- und/oder ASCII-Dateien (bei manchen Programmen muss die Ländereinstellung auf „Englisch“ umgestellt werden, um eine fehlerfreie Übernahme zu erreichen)
- die Daten auch in andere, hier nicht besprochene Statistikprogramme zu importieren.

Das ASCII- und das alte EXCEL-Format wurden gewählt, damit die Daten möglichst von allen anderen Programmen gelesen werden können. Die Dateinamen haben folgende Bedeutung: Beispiel3_5.xls enthält die Daten für das Beispiel aus der chemischen Industrie in Abschnitt 3.5 im EXCEL-Format (Endungen je nach Programm).

■ Anhang J – Software/Demos im Internet

Die Inhalte dieses Anhangs finden Sie auch in der Datei Informationen.html im Download, um Ihnen den Zugang zu den Links zu erleichtern. Außerdem kann die Download-Datei

auch zwischen Neuauflagen des Buches aktualisiert werden. Die Programme sind nach Themen und dann alphabetisch sortiert, die Reihenfolge stellt keine Wertung dar.

Umfassende Statistikpakete mit Versuchsplanung – Abschnitt 19.2

Cornerstone®

Informationen zu Cornerstone: <https://www.camline.com/>

Um einen Download für einen kostenlosen Test anzufordern, genügt ein E-Mail an info-cornerstone@camline.com

JMP®

Informationen zu JMP und Download für einen kostenlosen Test:

<https://www.jmp.com>

Minitab®

Informationen zu Minitab und Download für einen kostenlosen Test:

<https://www.minitab.com/>

Distributor in Deutschland ist die ADDITIVE GmbH:

<https://www.additive-net.de/de/software/produkte/minitab/>

Statgraphics®

Informationen zu Statgraphics und Download für einen kostenlosen Test:

<https://www.statgraphics.com/>

Distributoren in Deutschland:

<https://www.dpc-software.de> und <https://www.umex.de/statgraphics/>

TIBCO® Statistica

Informationen zu TIBCO Statistica und Download für einen kostenlosen Test auf Deutsch: <https://www.statsoft.de/> (TIBCO Statistica Partner) und auf Englisch: <https://www.tibco.com/products/data-science/> (Hersteller).

Visual-XSel®

Informationen zu XSel und Download für einen kostenlosen Test:

<https://crgraph.de/>

Spezielle Versuchsplanungsprogramme – Abschnitt 19.2

Design-Expert®

Informationen zu Design-Expert und Download für einen kostenlosen Test:

<https://www.statease.com/>

Distributor in Deutschland: <https://www.statcon.de>

MODDE®

Informationen zu MODDE Pro und Download für einen kostenlosen Test:

Englisch: <https://www.umetrics.com/> (Hersteller, Sartorius Stedim Data Analytics)

Deutsch: <https://www.umesoft.de/> (Vertriebspartner)

STAVEX®

Informationen zu STAVEX und Download für einen kostenlosen Test:
<https://www.aicos.com/>

Versuchsplanung für spezielle Anwendungsgebiete – Abschnitt 19.3**CAMEO® zur Motoroptimierung**

Informationen zu CAMEO:
<https://www.avl.com/cameo/>

vidara®/destra® zur Fertigungsoptimierung

Informationen zu vidara und destra:
<https://www.q-das.com/de/software/statistik/>

Weibull++® zur Zuverlässigkeitsermittlung

Informationen zu Weibull++ und Download für einen kostenlosen Test:
<https://www.reliasoft.com/products/reliability-analysis/weibull>

Fusion Pro/QbD® in der Nahrungsmittel- und Pharmaindustrie

Informationen zu Fusion Pro/QbD und Download für einen kostenlosen Test:
www.smatrix.com/

Sequentielle Optimierung – Abschnitte 18.3 und 18.4**GlobalOptimize®**

Informationen zu GlobalOptimize (GO):
www.globaloptimize.ch

ULTRAMAX®

Informationen zu ULTRAMAX:
www.ultramax.com

Spezielle Taguchi-Software – zu Kapitel 9**ANOVA-TM®**

Informationen zu ANOVA-TM:
<https://www.anovatm.com>

Qualitek-4®

Informationen zu Qualitek-4:
nutek-us.com/wp-q4w.html

Auswahl weiterer Statistikpakete und Versuchsplanungsprogramme**DOE PRO XL®**

Versuchsplanung als Excel®-Add-in
<https://sigmaxzone.com/doepro/>

DOE Wisdom®

Spezielles Versuchsplanungsprogramm
www.launsby.com/software_main.html

NCSS®

Umfassendes Statistikpaket mit Versuchsplanung
<https://www.ncss.com/software/ncss/>

SYSTAT®

Umfassendes Statistikpaket mit Versuchsplanung
<https://systatsoftware.com/products/systat/>

Unscrambler®

Schwerpunkt multivariate Verfahren/Spektroskopie, mit Versuchsplanung
<https://www.camo.com/unscrambler/>

Spezielle Anforderungen und alternative Modellierungsansätze – Kapitel 17

ETAS ASCMO®

Schwerpunkt auf Gaussian Process Models
<https://www.etas.com/de/portfolio/ascmo.php>

MATLAB®

Mathematikprogramm mit Toolboxen für verschiedene Anwendungen, auch für alternative Modellierungsansätze
<https://de.mathworks.com/>

R

Free Software, umfassende Sammlung statistischer Methoden mit vielen Einzelbeiträgen, fast beliebig anpassbar, erfordert aber viel Detailwissen
<https://www.r-project.org/>

Aktuelle Informationen und Links

Die angegebenen Links wurden im September 2019 überprüft. Erfahrungsgemäß ändern sich Internetadressen gelegentlich, die zugehörigen Links laufen dann ins Leere. Daher werden die Links im Download unter dem Reiter Extras folgender Seite regelmäßig aktualisiert:

<https://www.hanser-fachbuch.de/buch/Versuchsplanung/9783446461468>

Falls diese Links auch nicht funktionieren sollten oder falls Sie Änderungs- und Ergänzungswünsche haben, informieren Sie mich bitte per E-Mail:

wilhelm.kleppmann@hs-aalen.de

Index

A

- Ablauf
 - Datenauswertung 332
 - Produkt-/Prozessoptimierung 330
 - Versuchsplanung 17, 331
- A-Effizienz 284
- AIC 290
- Alias 132
- Alternativhypothese 84
- Anpassungstest *Siehe* Lack-of-Fit-Test
- A-optimal 229
- array
 - inner 170
 - orthogonal 180
 - outer 171
- Aufgabe
 - Ausbeute Chemie 110, 120
 - Ausbeute Halbleiterfertigung 126
 - Drehautomat 241
 - Druckgießen 85, 91, 94
 - Gewinderollen 118
 - Gießerei 47
 - Längenmessgerät 72
 - Leckstrom 190
 - Motoren 77, 81
 - Papier-Rotor 321
 - Schaltkreis 175
 - Trichter 324
- Auflösung 143, 277, 333
- augment 277
- Ausreißer
 - Begriff 37, 334
 - Beispiel 153, 276, 319
 - Tests 75
 - Umgang mit 37, 288
- Auswertungsverfahren
 - Effektberechnung 101
 - Mittelwertvergleich 82
 - Regressionsanalyse 185
 - Varianzanalyse 237
 - Vertrauensbereich 77

B

- backward selection 207
- Beispiel
 - Abscheidung in vier Anlagen 238
 - Ankerwellen 46
 - Antriebsritzel 265
 - Anzahl Lötfehler 97
 - Ausbeute Halbleiterfertigung 121
 - chemische Industrie 25, 116, 213, 280 f.
 - Computertest 160
 - Dickendifferenz Metallschicht 135
 - Dickenzunahme 185
 - Elektronikmodul 54
 - EVOP 340
 - Galvanik 66
 - Gewinderollen 141
 - Gießerei 51
 - Gummimischung 272
 - Halbleiterfertigung 53
 - Keramikfliesen 173
 - Laserschneiden 218, 270
 - Lebensdauer 95
 - Leiterplattenfertigung 49
 - Literatur 320
 - Metallpresse 59
 - Motoroptimierung 315
 - Papier-Rotor 321
 - Reaktionsrate 243, 246, 249
 - Reflowlöten 147
 - Reifenabrieb 252
 - Schaltkreis 175, 179
 - Schichtdicke 93
 - SiO₂-Abscheidung 102
 - Spritzgießprozess 340
 - Stahlsorten 84
 - Stoßdämpfer 54
 - Trichter 324
 - Vakuumverpackung 52
 - Wafermap 50
 - Zeitzähler 62
- Bestimmtheitsmaß 188, 289, 334

- Bias 289
 - BIC 290
 - Binomialverteilung 263, 287
 - Blockbildung
 - Aufwand 33
 - Auswertung 109
 - Begriff 31, 334
 - bei Erweiterung 278
 - Beispiel 32, 108
 - Blockfaktor 134
 - mehrere Blockfaktoren 252
 - zentral zusammengesetzt 217
 - Box-Cox-Transformation 98
 - Brainstorming 24, 330
- C**
- CAMEO® 311, 343
 - constraints 230, 284, 316
 - contour plot *Siehe* Höhenliniendiagramm
 - control factor 170
 - Cook-Distanz 208, 340
 - Cornerstone® 304, 342
 - crossed 250
- D**
- Data Mining 288
 - Datenauswertung 332
 - D-Effizienz 284
 - Design-Expert® 304, 342
 - Design of Experiments 1
 - desirability function 268
 - destra® 312, 343
 - DFFITS 208
 - DMAIC 8
 - DOE *Siehe* Versuchsplanung 1
 - DOE++® 312, 343
 - Dokumentation 42
 - D-optimal 229, 279, 284, 316
 - Dosis-Wirkungs-Zusammenhang 264
 - Dreieckstabelle 181
 - Drei-Faktor-Wechselwirkung *Siehe* Wechselwirkung
 - dummy levels 182
- E**
- Effekt
 - als Mittelwertdifferenz 83
 - Anzahl 115
 - Begriff 334
 - Berechnung 115
 - einer Wechselwirkung 102
 - eines Faktors 38, 102
 - fest/zufällig 246
 - für Mittelwert 115
 - Signifikanz 104
 - S/N (signal/noise ratio) 172
 - Vertrauensbereich 105
 - Wahrscheinlichkeitsdarstellung 121
 - Wechselwirkung bei Regression 204
 - Zusammenhang mit Regression 199
- Einflussgröße
 - Begriff 12, 334
 - Ideenfindung 24
 - nicht untersucht 28
 - Regression 185
- Einflussgrößen-/Zielgrößenmatrix 24
- Einsatzempfehlungen
 - alternative Modellansätze 292
 - Funktionstest 164
 - nichtlineare Zusammenhänge 234
 - Screening mehrstufig 254
 - Screening zweistufig 164
- Einzelversuch 334
- E-optimal 229
- Ergebnisse interpretieren 40
- erwartungstreu 76, 338
- Erwünschtheit 268
- ETAS ASCMO® 292, 344
- Evolutionary Operations *Siehe* EVOP
- EVOP
 - Faktorstufen 27
 - Probleme 297
 - Untersuchungsziel 22
 - Vorgehensweise 296, 340
- extreme vertices 283
- F**
- Faktor
 - Auswahl 25, 45, 330
 - Begriff 14, 334
 - gegenseitige Beschränkung 318
 - geschachtelt/gekreuzt 250
 - qualitativ (= kategorisch) 15, 27, 335
 - quantitativ (= numerisch) 15, 27, 335
 - Faktorstufen
 - Anzahl festlegen 26
 - Begriff 14, 334
 - fest/zufällig 246
 - natürlich 214
 - nichtlinear 215
 - normiert 214
 - Werte festlegen 27
 - Faktorstufenkombination 29, 113
 - nicht realisierbar 283

Fehler
 - 1./2. Art 90, 334, 339
 - Anzahl 96, 267
 feste Effekte 246
 Foldover 159, 277
 forward selection 207
 Freiheitsgrad
 - bei Pooling 125
 - einfache Regression 195
 - einfache Varianzanalyse 239
 - faktorieller Plan 116
 - Lateinische Quadrate 252
 - Mittelwertvergleich 84
 - Stichprobe 79
 - zweifache Regression 204
 F-Test 94
 Funktionstest 160
 FUSION® 312, 343
 F-Verteilung 239

G

Gaussian Process Models 291
 G-Effizienz 284
 gemischtes Modell 246
 Generator 133
 GLM 287
 GlobalOptimize® 302, 343
 G-optimal 229
 grafische Darstellung
 - Effekt Faktor 103
 - Effekt Wechselwirkung 104
 - Fehlerverteilung 49
 - Histogramm 68
 - Höhenlinien Beispiel 222, 319
 - Höhenlinien Mischung 258
 - Komponententausch 63
 - Masern-Diagramm 51
 - Messwerte 66
 - Mittelwertvergleich 2
 - Multi-Vari-Bild 45
 - Normalverteilung 70
 - Signifikanz 87, 106, 117, 221
 - Variablen tausch 59
 - Wahrscheinlichkeitsnetz Effekte 124
 - Wahrscheinlichkeitsnetz Messwerte 74
 - Wahrscheinlichkeitsnetz Residuen 94
 - Wechselwirkung bei Regression 206
 - Wechselwirkung mehrstufig 245
 - Wirkungsfläche 6
 Grundgesamtheit 68
 Grundlagen Statistik 65
 Gut-Schlecht-Ergebnisse 261

H

Half-Normal-Plot 127
 Häufigkeit
 - absolut 66
 - kumuliert 67
 - relativ 66
 Häufigkeitsverteilung 65
 Haupteffekt 38
 Hebel 208, 340
 Histogramm 67
 Höhenliniendiagramm 206, 221, 258, 319

I

Ideenfindung 24
 idle columns 182
 inclusions 230, 279
 indifferent 84
 inner array 170
 I-optimal 229
 Ishikawa-Diagramm 24

J

JAVA-Visualisierungen 337
 JMP® 304, 342

K

Kernel 290, 292
 Klasse 66
 klassisch 7, 20
 Kolmogoroff-Smirnoff-Test 75
 Komponententausch 61, 330
 Korrelationskoeffizient 188
 Korrelation zwischen Einflussgrößen 201, 208
 Kreuzvalidierung 290
 Kriging 292
 kubisches Modell 316

L

Lack-of-Fit-Test 196, 209, 281, 340
 Lateinische Quadrate 251
 Lenth ME 127
 leverage Siehe Hebel
 lineare Regression 185, 340
 linearer Graph 181
 Linearitätstest 196, 209
 Literatursuche 320
 logistische Funktion 264, 288
 Logit-Regression 264
 lokale Modelle 290
 loss function 168

M

- Machine Learning 288
- Masern-Diagramm 51
- Maßnahmen ableiten 41
- MATLAB® 292, 344
- Metamodellierung 288
- Metaplantechnik 24
- Minitab® 304, 342
- Mischungspläne
 - kombiniert 259
 - mit Begrenzungen 258
 - ohne Begrenzungen 256
- Mittelwert
 - Begriff 335
 - Effekt 115
 - Lage 19
 - Normalverteilung 70
 - Rand 244, 253
 - Stichprobe 76
 - Vertrauensbereich für 77, 338
- Mittelwertvergleich 82, 339
- mixed model 246
- MODDE® 304, 342
- Modellverbesserung 275
- Multi-Vari-Bild 45, 330

N

- nested 250
- Neuronale Netze 291
- Nichtlinearität 279
- noise factor 170
- normal probability plot *Siehe Wahrscheinlichkeitsnetz*
- Normalverteilung 70
 - Abweichung von 287
 - Simulation 338
 - standardisiert 70
- Nullhypothese 84

O

- One-factor-at-a-time 4
- Optimierung
 - mehrere Zielgrößen 222, 268, 319
 - sequentiell 295
- Optimierungsgrafik 274
- orthogonales Feld 171, 180
- Orthogonalität
 - Begriff 202
 - Beispiel 217
- outer array 171
- Outlier *Siehe Ausreißer*

P

- Paarweiser Vergleich 54, 330
- Parameterdesign 170
- Pareto-Prinzip 64
- Plackett-Burman 156
- Plan *Siehe Versuchsplan*
- Poissonverteilung 96, 267, 287
- Pooling 125
- P-P-Plot 75
- PRESS 209, 276, 290, 340
- probability level 88
- probability plot *Siehe Wahrscheinlichkeitsnetz*
- Prognosefähigkeit 209, 276, 289
- Prozessablaufdiagramm 24
- Prozessvergleich 52
- pseudo factor designs 182
- Pure Error 209
- p-Wert 243, 249

Q

- Q-optimal 229
- Q-Q-Plot 75
- quadratisches Modell
 - Erweiterung auf 281
 - Gleichung 211
 - Grenzen 230, 287

R

- R 288, 292, 344
- Radiale Basisfunktionen 290
- Randomisierung
 - Aufwand 33
 - Begriff 31, 335
 - Beispiel 32
 - nicht vollständig 248
 - Wirkung an Beispiel 108
- Rauschfaktor
 - Begriff 14
 - Faktorstufen 27
 - outer array 170
- Realisierung
 - Anzahl n berechnen 30, 120
 - n = 1 auswerten 121
- Regression
 - Beurteilung 208
 - einfache linear 186, 340
 - lokale Modelle 290
 - Matrixformalismus 204
 - mehrfach 200
 - quadratischer Effekt 211

- quasilinear 200
 - robust 288
 - schrittweise 207
 - Wechselwirkungseffekt 204, 211, 221
 - Zusammenhang mit Effekt 199, 202, 205
 - Regressionsanalyse 185, 335
 - Regressionskoeffizient 335
 - Regularisierung 290
 - Reihenfolge
 - Einzelversuche 92, 108
 - systematisch 37
 - zufällig 32
 - Residuen
 - Begriff 335
 - Berechnung 93
 - grafische Beurteilung 191, 340
 - standardisiert 191
 - studentisiert 191
 - Wahrscheinlichkeitsnetz 94
 - resolution *Siehe* Auflösung
 - response surface 221
 - Ridge Regression 290
 - Robustheit
 - als Ziel 167
 - alternative Ansätze 178
 - Begriff 20
 - grafische Darstellung 169
 - Maßzahlen 172
 - Rosa X 58
 - Rotes X 58
- S**
- Sättigung 215, 231
 - Screening
 - Erweiterung 277
 - mehrstufige Faktoren 251
 - zweistufige Faktoren 129, 331
 - S.d.q.A. 186, 340
 - sequentielle Optimierung
 - EVOP 296, 340
 - neuere Entwicklungen 301
 - Simplexverfahren 299
 - Software 343
 - steilster Anstieg 298
 - Shainin
 - Komponententausch 61
 - Multi-Vari-Bild 45
 - Paarweiser Vergleich 54
 - Prozessvergleich 52
 - sieben Werkzeuge 64
 - Überblick 64
 - Untersuchungsziel 21
 - Variablenvergleich 57
 - Shapiro-Wilk-Test 75
 - Signal-Rausch-Verhältnis 171
 - significance level 88
 - Signifikanz
 - Begriff 335
 - Definition 84
 - grafisch 87
 - Regression 195
 - Signifikanzniveau 88
 - Simplexgitterplan 256
 - Simplexverfahren 299, 331
 - Simplex-Zentroid-Pläne 257
 - Six-Sigma 8
 - S/N signal-to-noise-ratio 171
 - Software 303, 341
 - Soll-/Istwerte 283
 - S-optimal 229
 - Space Filling Plan 292
 - Spline 290, 292
 - Split-Plot 249
 - Standardabweichung
 - Begriff 335
 - eines Effekts 84, 105, 115, 284
 - eines Mittelwerts 78
 - Normalverteilung 70
 - Stichprobe 76
 - standard error *Siehe* Standardabweichung
 - z.B. eines Effekts
 - STATGRAPHICS® 304, 342
 - STATISTICA® 304, 342
 - STAVEX® 304, 343
 - steepest ascent *Siehe* steilster Anstieg
 - steilster Anstieg 298, 331, 340
 - Steuerfaktor
 - Begriff 14
 - inner array 170
 - Unterteilung 173
 - Steuergröße 24, 336
 - Stichprobe
 - Begriff 65, 69
 - repräsentativ 71, 92
 - Stichprobenumfang 78 *Siehe auch* Versuchs-umfang
 - Störgröße 24, 170, 336
 - Streudiagramm 64
 - Stufe *Siehe* Faktorstufen 27
 - Summenhäufigkeit 67
 - Systemdesign 170

T

Taguchi
 - Alternativen zu 178
 - Designphasen 170
 - Kritik an 180
 - mehrstufige Pläne 251
 - Robustheit 167
 - Software 343
 - Strategie 168
 - Untersuchungsziel 20
 - Versuchspläne 170

Team 24

Testabdeckung 162

Toleranzdesign 170

Transformation

- Anwendung 193
- arcsin 263
- Beispiel 276
- Box-Cox 98
- Einflussgrößen 204
- logarithmisch 95
- Messwerte 95
- varianzstabilisierend 263
- Wurzel 97, 267
- ω - 264

Trend 108, 192

t-Wert 79

Typ Siehe Auflösung

U

Über-/Unteranpassung 288
 ULTRAMAX® 343
 Umgebungsbedingungen 318
 Unterschied
 - (hoch)signifikant 84
 - indifferent 84
 Untersuchungsziel 19
 Ursache-Wirkungs-Diagramm 24

V

Variablenvergleich 57
 Varianz
 - Begriff 336
 - Normalverteilung 70
 - Residuen 194
 - Stichprobe 76
 Varianzanalyse
 - Begriff 336
 - einfache 238
 - mehrfach 243
 - Streuungszerlegung 242

Varianzinflationsfaktor 208, 284
 Varianzkomponente 246
 Verbesserung absichern 42
 Verbesserungszyklus 43
 Verlustfunktion 168
 Vermengung

- Begriff 132, 336
- Risiken 140
- trennen 277

 Versuchsaufwand abschätzen 33
 Versuchsplan

- 3k und 3k-p 226
- Begriff 336
- Box-Behnken 226
- Definitiver Screening 162
- Desperado 160
- D-optimal 229, 279, 316
- drehbar 215
- Erweiterung 277
- flächenzentriert 216
- foldover 277
- fraktionell faktoriell 130
- gemischt 181
- gesättigt 145, 181
- klein zusammengesetzt 227
- Minimum Aberration 146
- Mischungen 255
- nichtlineare Zusammenhänge 211
- nicht orthogonal 230, 279
- optimal 229
- orthogonal 208, 213
- orthogonale Blöcke 217
- orthogonales Feld 171
- orthogonal und drehbar 216
- Plackett-Burman 156
- Screening 129
- Simplexgitterplan 256
- Simplexverfahren 299
- Simplex-Zentroid-Pläne 257
- Space Filling 292
- Split-Plot 249
- übersättigt 160
- vollständig faktoriell 101
- Wechselwirkungen trennen 277
- zentral zusammengesetzt 211
- Zentrumspunkt 279

 Versuchsplanung

- Einsatzgebiete 1
- Ergebnisse 6
- klassisch 7, 20, 101, 211
- Prinzip 4
- Shainin 7, 21, 45, 64
- Software 303
- Taguchi 7, 20, 167

- Vorgehensweise 17, 331
- Vorteile 5
- Versuchsumfang
 - Anzahl Fehler 267
 - Auswirkung 89
 - Begriff 4, 83, 336
 - gut-schlecht 265
 - Mittelwertvergleich 90
 - nichtlineare Zusammenhänge 234
 - Plackett-Burman 159
 - Screeningpläne 164
 - Varianzanalyse 241
 - vollständig faktoriell 120
- Verteilung
 - Binomial 263, 287
 - F- 239
 - Fehler örtlich 49
 - Häufigkeit 65
 - logarithmisch normal 95
 - nicht normal 287
 - normal 70, 338
 - Poisson 96, 287
 - t- 78
- Verteilungsdichte 68
- Verteilungsfunktion 68
- Vertrauensbereich
 - Begriff 336
 - für Differenz 83
 - für Effekt 105
 - für Mittelwert 77, 338
 - für Regression 195, 340
 - Hinweise zu 81
 - Simulation 338
- Vertrauensniveau 78, 338
- VIF 208, 284
- Vorhersagebereich 197, 336, 340

W

- Wahrscheinlichkeitsnetz
 - Effekte 121
 - Entstehung 73
 - Residuen 93
- Wavelets 290
- Wechselwirkung
 - Begriff 337
 - Drei-Faktor 115
 - Formel 102
 - Grafik 104
 - Höhenliniendiagramm 206
 - Interpretation 106, 117
 - Varianzanalyse 243
 - Vermengung 132, 141
 - Vermengung trennen 277
 - Zwei-Faktor 106, 115
- Wirkungsfläche 6, 206, 221
- Wunschfunktion 268, 337

X

- XSel® 304, 342

Z

- Zentrumspunkt 279, 340
- Zielgröße
 - Anzahl Fehler 267
 - Auswahl 22
 - Begriff 337
 - gut-schlecht 261
 - mehrere 222, 268, 316
 - zufällige Effekte 246
- Zuordnung 131
- Zuverlässigkeitstests 162
- Zwei-Faktor-Wechselwirkung Siehe Wechselwirkung

Stellen Sie sich auf die Zukunft um



Ross
Funktionale Sicherheit im Automobil
**Die Herausforderung für Elektromobilität und
automatisiertes Fahren**
2., vollständig überarbeitete Auflage
499 Seiten
€ 99,99. ISBN 978-3-446-45841-3
Auch einzeln als E-Book erhältlich

Die Elektromobilität ist im Kommen, die Tage des Verbrennungsmotors scheinen gezählt. Selbst die Rolle des Fahrers wird neu definiert oder gar infrage gestellt werden müssen.

Dieses vollständig überarbeitete Buch beschäftigt sich mit vielen Fragen um die Sicherheit, die grundlegende Veränderungen mit sich bringen.

Alle diese Fragen gehen weit über den Sicherheitsbegriff der ISO 26262 hinaus. Zahlreiche Aspekte, die Hersteller, Betreiber und Service-Anbieter in Betracht ziehen sollten, wenn sie für das Post-Automobilzeitalter sichere Mobilität anbieten wollen werden aufgezeigt.

Automatisieren, optimieren, profitieren



Müller, Franke, Henrich, Kuhlenkötter, Raatz, Verl

Handbuch Mensch-Roboter-Kollaboration

480 Seiten. E-Book inside. Komplett in Farbe

€ 199,–. ISBN 978-3-446-45016-5

Auch einzeln als E-Book erhältlich

Das Handbuch beschreibt alle wichtigen Aspekte, die beim Einsatz von kollaborativen Robotern eine Rolle spielen:

- Das Geschäfts- und Wettbewerbsumfeld – Wo und wann lohnt sich der Einsatz von kollaborativen Robotern überhaupt?
- Der vorhandene Maschinenpark im Unternehmen – Passen Roboter da hinein oder muss man zusätzliche Investitionen einplanen?
- Arbeitsschutz – Sind Roboter unter allen Umständen sicher?
- Technik – Welche Typen gibt es, welche Steuerungskonzepte gibt es?
- Produktionsprozesse – Wie werden Roboter auf allen Ebenen erfolgreich integriert, ohne Menschen zu benachteiligen?

Den digitalen Wandel meistern



Gassmann, Sutter

Digitale Transformation gestalten

Geschäftsmodelle Erfolgsfaktoren Checklisten

2., überarbeitete und erweiterte Auflage

356 Seiten. E-Book inside

€ 34,90. ISBN 978-3-446-45868-0

Auch einzeln als E-Book erhältlich

Digitalisierung hat sich rapide vom Schlagwort zum strategischen Wettbewerbsfaktor entwickelt. Auch wenn die digitale Transformation alle Branchen in unterschiedlicher Geschwindigkeit erfasst, kann sich keine Industrie dieser Entwicklung entziehen. Um wettbewerbsfähig zu bleiben, müssen Unternehmen den digitalen Wandel aktiv gestalten und entstehende Chancen nutzen.

Die Gestaltung der digitalen Transformation betrifft nicht nur IT-Verantwortliche, sondern ist Aufgabe des gesamten Unternehmens. Führende Autoren aus Wissenschaft und Praxis zeigen mit zahlreichen Fallstudien, wie die digitale Transformation erfolgreich gestaltet und konkret umgesetzt werden kann.

Mit Flexibilität Herausforderungen meistern



Lang, Wagner

Der Weg zum projektorientierten Unternehmen
Wissen für Entscheider

336 Seiten

€ 39,90. ISBN 978-3-446-45837-6

Auch einzeln als E-Book erhältlich

Unternehmen stehen heute vor gravierenden Herausforderungen: digitale Transformation, zunehmende Komplexität und disruptive Veränderungen. Nur flexible Organisationen können unter diesen neuen Rahmenbedingungen dauerhaft erfolgreich sein. Somit wird Projektarbeit immer bedeutender für Unternehmen. Dabei reicht es aber nicht mehr aus, sich auf einzelne Projekte und »klassisches« Projektmanagement zu beschränken. Vielmehr wird die Projektorientierung des gesamten Unternehmens immer mehr zum entscheidenden Erfolgsfaktor und Wettbewerbsvorteil.

Dieses Werk zeigt, wie sich ein Unternehmen ganz auf Projektarbeit einstellt und sich zu einem projektorientierten Unternehmen entwickelt. Dabei werden alle relevanten Aspekte praxisorientiert zusammengefasst und ein konkreter Umsetzungsweg vorgestellt.

Ihr Weg zur gültigen Konformitätserklärung



Kuntz, Rost
EU-Konformität für elektrotechnische und elektronische Produkte
Leitfaden für Hersteller, Importeure und Händler
215 Seiten. E-Book inside
€ 69,-. ISBN 978-3-446-45920-5
Auch einzeln als E-Book erhältlich

Dieses Buch bietet die zur CE-Kennzeichnung notwendige Orientierung:

- Fundiertes Basiswissen, um gegenüber Herstellern, Kunden, Überwachungsbehörden und Prüfstellen ein zuverlässiger Partner zu sein
- Führung durch die Konformitätsvielfalt elektronischer Produkte, Klärung der Verantwortlichkeiten bei Zuliefer- und Handelsketten
- Wege zur Konformitätserklärung anhand von Praxisbeispielen
- Auflistung der notwendigen Begleitunterlagen
- Herausfinden der zulässigen Normenstände, um die Gültigkeit von Konformitätserklärungen prüfen zu können
- Anhand der meist geltenden Richtlinien wird das Prinzip der Konformitätsbewertung vermittelt.

Versuchsplanung

Produkte und Prozesse optimieren

Jede Neu- oder Weiterentwicklung von Produkten und Fertigungsprozessen durchläuft eine Vielzahl von Versuchen. Aber Versuche kosten Zeit und Geld, und manche Ergebnisse sind nicht reproduzierbar. Versuchsplanung (auch DOE = Design of Experiments genannt) hilft dabei, mit möglichst geringem Aufwand reproduzierbare Ergebnisse zu erhalten.

Dieses Buch zeigt Praktikern in Entwicklung, Konstruktion und Fertigung, sowie Studenten, wie mit Versuchsplanung Produkte und Prozesse optimiert werden können. Versuchsplanung ist ein wesentlicher Bestandteil der Six-Sigma-Strategie. Anhand von vielen Beispielen wird beschrieben, wie man

HIGHLIGHTS!

- ▶ durch systematische Beobachtung Informationen sammeln,
- ▶ durch einfache Versuche einfache Probleme lösen,
- ▶ mit wenigen Einzelversuchen die Bedeutung vieler Faktoren gleichzeitig beurteilen und so die wichtigen von den unwichtigen trennen,
- ▶ robuste Produkte und Prozesse entwickeln,
- ▶ die Abhängigkeit mehrerer Zielgrößen von mehreren Faktoren gleichzeitig erfassen und ein gemeinsames Optimum suchen sowie
- ▶ schrittweise Fertigungsprozesse und Produkte verbessern kann.

In der 10. Auflage wurde insbesondere das Software-Kapitel wesentlich erweitert, um neue Entwicklungen und Tendenzen zu berücksichtigen. Andere Kapitel wurden aktualisiert und ergänzt.

Im Download-Bereich des Hanser-Verlags zu diesem Buch werden die Daten aller Beispiele und Links zu geeigneter Software bereitgestellt. Mit diesen Daten können Sie die Programme anhand bekannter Beispiele selbst testen und das für Ihre Bedürfnisse am besten geeignete auswählen. Visualisierungen in JavaScript erlauben es, statistische Grundlagen spielerisch zu erleben und besser zu verstehen.

HANSER

www.hanser-fachbuch.de

€ 34,99 [D] | € 36,00 [A]

ISBN 978-3-446-46146-8



9 783446 461468