

Entwicklung Effizienter Multivariater Lebensdauertests

Efficient Multivariate Lifetime Testing

Der Fakultät Konstruktions-, Produktions- und Fahrzeugtechnik
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) vorgelegte
Abhandlung

von
Marco Arndt, M.Sc.
aus Ravensburg

Hauptberichter: PD Dr.-Ing. habil. Martin Dazer
Mitberichter: Univ.-Prof. Dr.-Ing. Mit Berichter

Tag der mündlichen Prüfung:

Institut für Maschinenelemente der Universität Stuttgart

2026

Vorwort

Thanks for your service.

Kurzfassung/Abstract

-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-

TODO:

-
- Abstract / Kurzfassung / Danksagung
- Bilder prüfen
- "mark changes after your Oral presentation for review"
- "textcolor blue in definitions ändern
- "Print" Funktion aktivieren
- "nocite" Funktion deaktivieren in main Z.269

-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-

Inhaltsverzeichnis

Vorwort	iii
Kurzfassung/Abstract	v
Nomenklatur	ix
Abkürzungen	ix
Indizes	x
Formelzeichen	xi
Abbildungsverzeichnis	xv
Tabellenverzeichnis	xvii
1 Einleitung	1
1.1 Forschungsperspektive und Problembeschreibung	2
1.2 Beitrag dieser Arbeit	3
1.3 Aufbau der Arbeit	4
2 Stand der Forschung und Technik	7
2.1 Zuverlässigkeitstechnik und Wahrscheinlichkeitstheorie	7
2.1.1 Begriffe und Definitionen	8
2.1.2 Deskriptive Statistik für Lebensdauerdaten	9
2.1.3 Parametrische Lebensdauermodelle	11
2.1.4 Parameterschätzverfahren	13
2.2 Statistische Versuchsplanung und Modellbildung	18
2.2.1 Grundlagen zur statistischen Versuchsplanung	19
2.2.2 Statistische Versuchsplanung zur Lebensdauererprobung	24
2.2.3 Statistische Modellbildung	41

3	Effiziente Testplanung für die multivariate Lebensdauererprobung .	49
3.1	Bewertung des Standes der Forschung und Technik	49
3.2	Forschungsfragen und Aufbau der Arbeit	52
4	Parameter-Screening für multifaktorielle Lebensdauerests	53
4.1	Identifikation potenzieller Einflussgrößen	53
4.2	Kreativmethoden zum Auswahlprozesse im Parameter-Screening	53
4.3	Randbedingungen in der Parameterauswahl für die Zuverlässigkeits- modellierung	53
4.4	Vorgehen zum heuristische Screening für die Zuverlässigkeitsmodel- lierung	53
4.5	Zusammenfassung	53
5	Effiziente multivariate Versuchspläne für Lebensdaueruntersuchun- gen	55
6	Fallstudie	57
7	Zusammenfassung und Ausblick	59
7.1	Key Findings	59
7.2	Diskussion	59
7.3	Ausblick	59
	Literatur	61
	Anhang	71
A	Ableitungen zur Maximum-Likelihood-Schätzung	71

Nomenklatur

Abkürzungen

<i>cdf</i>	Cumulative Distribution Function
<i>pdf</i>	Probability Density Function
ALT	Accelerated Lifetime Testing
ANOVA	Varianzanalyse, engl. Analysis of Variance
BEV	Battery Electric Vehicle
CCD	Central Composite Design
CI	Confidence Interval
CP	Center Point
DfR	Design for Reliability
DoE	Design of Experiments
DSM	Design-Structure-Matrix
ECU	Electronic Control Unit
EoL	End-of-Life
FDS	Fraction of Design Space
GLL	Generalized Log-Linear
GLM	Generalized Linear Model
IMA	Institut für Maschinenelemente
L-DoE	Lifetime Design of Experiments
LR-Test	Likelihood-Ratio Test
MLE	Maximum-Likelihood-Estimation

MMR	Median-Rank-Regression-Methode
OFAT	One Factor At Time
OLS	Ordinary Least Squares
PH	Proportional Hazard
RSD	Response Surface Design
RSM	Response Surface Methodology
SPV	Scaled Prediction Variance
UPV	Unscaled Prediction Variance
VDG	Variance Dispersion Graph
VIF	Varianz-Inflations-Faktor

Indizes

0	Zeitpunkt $t = 0$ Zentralpunkt Regressionskonstante Position im Versuchsraum
BC	Box-Cox-Transformation
C	Zentralpunkt
D	Sternpunktabstand CCD
F	Faktorieller Punkt
MC	Monte-Carlo-Simulation
O	Beobachtet (Observed)
S	Sternpunkt
V	Transformation
error	Fehler 2. Art (Wahrscheinlichkeit für ein falsch-negatives Ergebnis)
i	Laufvariable

j	Laufvariable
l	Laufvariable
o	Obere Grenze des Vertrauensbereichs
q	Quantilwert
u	Untere Grenze des Vertrauensbereichs
z	Zensierungsindikator

Formelzeichen

A_{opt}	A-Optimalitätskriterium (Spur-Minimierung)
α	Signifikanzniveau
	Abstand der Sternpunkte (Axial Points)
b	Weibull-Formparameter (Weibull-Modul)
β	Koeffizientenmatrix
β	Regressionskoeffizient, Modellparameter
	Fehler 2.Art
$\chi^2(\cdot)$	Chi-Quadrat-Verteilung
$\text{Cov}[\cdot]$	Covarianz-Operator
COVRATIO	Einflussstatistik auf Schätzwerte (Covariance Ratio)
D	Cook's Distanz
d_f	Anzahl der Freiheitsgrade
D_{opt}	D-Optimalitätskriterium (Determinanten-Maximierung)
δ	Statusindikator (1=Ausfall, 0=Zensiert)
DFBETAS	Einflussstatistik auf Koeffizienten (Difference in Betas)
DFFITS	Einflussstatistik auf Prädiktionen (Difference in Fits)
$E(\cdot)$	Erwartungswert
E	Effekt
A_{eff}	A-Effizienz

D_{eff}	D-Effizienz (Normierte Determinante)
G_{eff}	G-Effizienz (Verhältnis Modellparameter zu max. Varianz)
I_{eff}	I-Effizienz
V_{eff}	V-Effizienz
λ	Eigenwert (einer Matrix)
	Transformationsparameter (Box-Cox)
ϵ	Fehlerterm-Matrix
ϵ	Fehlerterm
$E(\cdot)$	Erwartungswert
η	Linearer Prädiktor (Logarithmus der charakteristischen Lebensdauer)
$F(\cdot)$	Ausfallwahrscheinlichkeit, Verteilungsfunktion
$f(\cdot)$	Dichtefunktion
\mathbf{F}	Fisher-Informationsmatrix
$G(\cdot)$	Verteilungsfunktion der kleinsten Extremwerte
$g(\cdot)$	Abgeleitete Funktion (für Delta-Methode)
\mathbf{g}'	Gradientenvektor der Funktion g
G_{opt}	G-Optimalitätskriterium (Minimierung max. Prädiktionsvarianz)
$\Gamma(\cdot)$	Gamma-Funktion
\mathbf{H}_{f}	Hessian-Matrix (der Log-Likelihood-Funktion)
h	Eintrag von \mathbf{H} , Hebelwert
H_0	Nullhypothese
H_1	Alternativhypothese
\mathbf{H}	Orthogonale Projektionsmatrix, Prädiktionsmatrix, Hat-Matrix
I_{opt}	I-Optimalitätskriterium (Minimierung durchschn. Prädiktionsvarianz)
k	Anzahl der Modellparameter

$\mathcal{L}(\cdot)$	Likelihood-Funktion
$\Lambda(\cdot)$	Log-Likelihood-Funktion
$\lambda(t)$	Ausfallrate (zeitabhängig)
LR	Likelihood-Ratio
M	Informationsmatrix $\mathbf{X}'\mathbf{X}$
m	Replikation
μ	Lageparameter
N	Versuchsanzahl
n	Stichprobenumfang Anzahlvariable
p	Anzahl der Regressionskoeffizienten
p -Wert	Überschreitungs-, Rest-Irrtumswahrscheinlichkeit
p_f	Fraktionsgrad
$power$	Trennschärfe, Power
\Pr	Wahrscheinlichkeit
$\nu(\mathbf{x}_0)$	Prädiktionsvarianz an der Stelle \mathbf{x}_0 (skaliert)
q	Quantilwert
R	Zuverlässigkeit
r	Studentisiertes Residuum
r_{CS}	Cox-Snell-Residuum
r_{Dev}	Standardisiertes Devianz-Residuum
r	Standardisiertes Residuum (bezogen auf die SEV-Verteilung)
\mathbb{R}	Menge der reellen Zahlen
s	Empirische Standardabweichung (von x)
s^2	Empirische Varianz (von x)
$SEV(\cdot, \cdot)$	Verteilung der kleinsten Extremwerte, engl. Smallest Extreme Value

σ	Standardabweichung (theoretisch) Skalenparameter
σ^2	Varianz der Lebensdauer (theoretisch)
T	Charakteristische Lebensdauer (Skalenparameter)
t	Zeit, Lebensdauermerimal Teststatistik, Signal-to-Noise Ratio
\mathbf{t}	Vektor der Ausfallzeiten
τ	kontinuierliche Zufallsvariable
θ	Modellparameter, Schätzung
$\boldsymbol{\theta}$	Vektor der Modellparameter
v	Integrationsvariable (Gamma-Funktion)
\mathbf{V}	Varianz-Kovarianz-Matrix
V_{opt}	V-Optimalitätskriterium (Punktueller Varianz-Minimierung)
$\text{Var}[\cdot]$	Varianz-Operator
$\mathcal{W}(\cdot, \cdot)$	Weibull-Verteilung
w	Hilfsgröße zur Berechnung von Vertrauensbereichen (standardisierte Log-Lebensdauer)
\mathbf{X}	Versuchsplan-Matrix oder Modellmatrix
x	Messwert, Variable, Parameter
\bar{x}	Arithmetischer Mittelwert (empirisch)
\mathbf{x}	Position im Parameterraum
\mathbf{y}	Antwortmatrix
y	Systemantwort, -variable
z	Quantil der Standardnormalverteilung
z	Hilfsvariable für die Log-Likelihood-Berechnung
z_e	Standardisierte Variable der Kleinstwertverteilung (SEV); Hilfsgröße für Wald-Konfidenzintervalle (äquivalent zu w)

Abbildungsverzeichnis

2.1	Weibull $f(t)$ für ausgewählte Werte von T und b	13
2.2	Design of Experiments (DoE) Steps gemäß [41]	20
2.3	Parameter-Diagramm (P-Diagramm)	21
2.4	Schematische Darstellung der Effekte: (a) positiver Haupteffekt von Faktor x_1 , (b) Wechselwirkungseffekte zwischen x_1 und x_2	22
2.5	Standard voll-faktorieller Versuchsplan	25
2.6	Einfluss der Schrittweite auf die Approximation des Effekts E	25
2.7	Zentral zusammengesetzter Versuchsplan (CCD) mit $k = 2$	29
2.8	Strategischer Ansatz zu augmentierter Versuchsplanung für die Lebensdauererprobung nach [51, 52]	30

Tabellenverzeichnis

2.1	Übersicht, Zielsetzung und mathematische Definition verschiedener Optimalitätskriterien für Versuchspläne	36
2.2	Übersicht, Zielsetzung und mathematische Definition verschiedener Effizienzkriterien für Versuchspläne	38
2.3	Übersicht gängiger Transformationen der Belastungsgrößen zur Abbildung physikalischer Lebensdauermodelle	43

1 Einleitung

Die Absicherung technischer Produkte und Systeme hinsichtlich ihrer Funktionalität bildet einen zentralen Bestandteil der ingenieurwissenschaftlichen Verantwortlichkeiten im Produktentwicklungsprozess. Motiviert durch Produktversprechen gegenüber der potenziellen Käuferschaft sowie bestehender Kunden, zur Wahrung des Selbstverständnisses einer Marke oder rein aufgrund regulatorischer Vorgaben soll hier im Kontext des vorgesehenen Einsatzzweckes Zuverlässigkeitsmanagement betrieben werden. So wird insbesondere aus Marktperspektive erwartet, dass ein (technisches) Produkt - ein Fahrzeug, ein Smartphone, eine Photovoltaikanlage - seine Funktionalität mindestens zum Gewährleistungs- oder Garantiezeitraum uneingeschränkt erfüllt.

Über den gesamten Produktlebenszyklus hinweg unterstützen Methoden der Zuverlässigkeitstechnik dabei, diese Anforderungen systematisch zu erfüllen. Verfahren aus dem Bereich „Safety“, explorative Datenanalysen zur Untersuchung der Produktperformance im Betrieb oder Test, effiziente Versuchsplanung zur Analyse oder zum Nachweis der Lebensdauer am Design, Methoden der beschleunigten Versuchsplanung, engl. [Accelerated Lifetime Testing \(ALT\)](#), der Aufbau probabilistischer Lebensdauermodelle sowie das Risikomanagement im Allgemeinen eignen sich für diese Herausforderung. Das zentrale Ergebnis liegt in der Ermittlung der Ausfallwahrscheinlichkeit als Komplement zur Zuverlässigkeit - also der Wahrscheinlichkeit, dass ein Produkt unter den definierten Randbedingungen eine vorgegebene Zeitdauer ohne funktionskritischen Ausfall übersteht [1]. Üblicherweise soll so nachgewiesen werden, dass das Erzeugnis dem Einfluss einer bestimmten Belastung - beispielsweise einer physikalischen oder elektrischen Kraft, einem Wärmeeintrag oder der Exposition gegenüber einer chemischen Beanspruchung - standhält. Gelingt in der Praxis keine hinreichend genaue Quantifizierung dieser probabilistischen Metrik, so liegen die Ursachen jedoch nicht zwangsläufig allein in ökonomischen Einschränkungen wie dem Zeit- und Kostenbudget für ein erforder-

liches Testing oder einem fehlenden methodischen Know-how - vielmehr könnten *mehrere* Einflussfaktoren auf die Zuverlässigkeit einwirken und sogar Wechselwirkungen ausprägen, ohne dass dies adäquat wahrgenommen oder antizipiert wird.

Moderne Produkte können schlichtweg durch multivariat bedingte Fehlermechanismen ausfallen. Die wirtschaftliche Tragweite dieser Erkenntnis wurde bereits 1996 im renommierten *Forbes Magazine* thematisiert, welches die einschlägigen Erfolge multivariabler Testmethoden in der industriellen Praxis hervorhob [2].

1.1 Forschungsperspektive und Problembeschreibung

Da trotz genannter Umstände die Kundenanforderungen und Garantiebedingungen üblicherweise als unveränderlich, teils sogar als zunehmend anspruchsvoller zu verstehen sind, werden Unsicherheiten in der Lebensdauerabsicherung dann meist nur durch präventive Wartungsstrategien, durch Tolerierung von Restrisiken oder durch die Inkaufnahme nachträglicher Schadensbegrenzung behandelt. Der zugrunde liegende Gedanke: ehe ein Produkt, dessen Lebensdauerverhalten nicht quantifizierbar verstanden ist, einen kritischen Verschleißzustand erreicht, wird es im Rahmen eines festgelegten Wartungsintervalls vorsorglich ersetzt. Dabei könnte zugrunde liegen, dass schlichtweg kein physikalisches Modell oder eine ausreichend ausgeprägte empirische Datengrundlage vorhanden ist. Führt auch diese Vorsorge zu erheblichen Regress- oder Kulanzkosten, müssen ausfallschutzorientierte Maßnahmen - bis hin zu Rückrufaktionen - frühzeitig eingeplant und umgesetzt werden, um Image- und Kostenrisiken (wenn auch selten nachhaltig) zu minimieren. Um jenes zu vermeiden, muss also bereits im Vorfeld den verschiedensten Randbedingungen mithilfe der Zuverlässigkeitstechnik begegnet werden. Besonders komplexe Randbedingungen lassen sich beispielsweise durch eine hochgradige Integration von Elektrifizierung und Digitalisierung, verkürzte Entwicklungszyklen, verschärfter Kostendruck, sich per se verändernde Prioritäten aus Marktperspektive, wandelnde Materialauswahl und -komposition, leistungsoptimierte Belastungsszenarien, intensivierte Einsatzbedingungen und nicht zuletzt eine effiziente Ressourcennutzung innerhalb eines Produkts beschreiben - um nur einige zu nennen. Ein einfaches Beispiel verdeutlicht dies: Der komfortable sowie ausfallfreie Betrieb eines Fahr-

zeugs soll einerseits gegenüber verschiedensten Schadensursachen gewährleistet werden; andererseits hängt er maßgeblich vom Funktionserhalt einer inzwischen bis zu dreistelligen Anzahl an [Electronic Control Units \(ECUs\)](#) ab - während in der jüngeren Vergangenheit noch eine geringe bis mittlere zweistellige Anzahl mit nur begrenztem Funktionsumfang üblich war [3]. Oder aber der störungsfreie Betrieb digitaler Services im [Battery Electric Vehicle \(BEV\)](#) setzt bei Zentralisierung von [ECU](#)-Funktionen die stetige Funktionsfähigkeit der Traktionsbatterie voraus, durch deren chemische Alterung, verschiedenst beeinflusst, sich jedoch zeitgleich wiederum weitere Unsicherheiten aus verschiedenen Einflüssen eingliedern können. Ein damit betrachtetes System bildet somit ein vielfältig komplexeres Netzwerk aus teils wechselwirkenden oder direkten Eigenschaften, die zu einem Versagen führen könnten.

Damit kann festgehalten werden: Produktdesigns werden angesichts steigender Kundenanforderungen zunehmend raffinierter und im Sinne des Effizienzgedankens optimiert (z.B. durch mehr Sensorik, Rechenleistung und fortschrittliche Assistenzsysteme), gleichzeitig wird jedoch die Art und Weise eines möglichen Funktionsverlustes zunehmend komplexer. Somit rückt die Fähigkeit, Prognosen über die Lebensdauer und Ausfallwahrscheinlichkeit in Abhängigkeit von mehreren Einflussgrößen treffen zu können, zunehmend in den unternehmerischen Fokus. Über die klassische Testplanung im Rahmen der Zuverlässigkeitstechnik hinaus erfordert dies also Methoden der statistischen Versuchsplanung bei gleichzeitiger Berücksichtigung *mehrerer* Einflussfaktoren auf die Lebensdauer. Methoden wie [ALT](#) und die Lebensdauermodellbildung behalten dabei weiterhin ihre Relevanz und bilden einen unverzichtbaren Bestandteil einer fundierten Teststrategie.

1.2 Beitrag dieser Arbeit

Ausgehend von der beschriebenen Problemstellung lässt sich der übergeordnete Beitrag dieser Arbeit wie folgt formulieren: Liegt ein komplexes technisches System vor und soll dieses hinsichtlich seiner Lebensdauer empirisch untersucht werden, um fundierte Prognosen über die Funktionalität im Betrieb treffen zu können, so müssen mehrdimensionale Lebensdaueruntersuchungen nach dem Prinzip des [DoE](#)

geplant werden. Neben der bloßen Implementierung von mehrdimensionale Tests für die Lebensdauererprobung berücksichtigt dieser Ansatz damit:

- eine effiziente Methodik zur gezielten Vorauswahl relevanter Faktoren aus der Gesamtheit potenzieller Systemparameter - mit dem Ziel, deren signifikanten Einfluss auf die Lebensdauer zu untersuchen;
- die Auswahl geeigneter Strategien und passender Testpläne zur statistisch abgesicherten Quantifizierung von Einflüssen auf die Lebensdauer, in Kombination mit konventionellen Zuverlässigkeitsmethoden wie beispielsweise [ALT](#);
- eine präzise Parameterschätzung zur mathematischen Beschreibung der Effekte auf Basis der als signifikant identifizierten Einflussgrößen;
- die Bilanzierung geeigneter Testpläne im Vergleich zu etablierten, in der Literatur bereits umfangreich diskutierten Versuchsplänen, insbesondere hinsichtlich potenzieller Abweichungen bei nicht-normalverteilten Daten.

1.3 Aufbau der Arbeit

Der allgemeine inhaltliche Aufbau der vorliegenden Arbeit kann [Abb. 1](#) entnommen werden. So folgt auf die in diesem Abschnitt beschriebene Problemstellung sowie Ausführung über den generellen Beitrag der Arbeit weiter in Kapitel 2 der relevante Stand aus aktueller Forschung und Literatur. Kapitel 3 fasst schließlich den Forschungsbedarf zusammen und stellt das Ziel der Arbeit, aus der sich die relevanten Forschungsfragen ergeben, konkret heraus. Kapitel 4 beinhaltet die Vorstellung zu effizienten, qualitativen Screening-Methoden. Hier werden die herausgearbeiteten Vorschläge zu einer Auswahl an heuristischen Methoden für die Selektion der perspektivisch wenig relevanten Faktoren beschrieben, die für die Umsetzung in der experimentellen statistischen Datenerhebung für maßgeblich erachtet werden. Weiter werden in Kapitel 5 darauf die Rahmenbedingungen für die zur statistischen Versuchsplanung neu herausgearbeiteten Versuchsplankonfigurationen für effiziente Lebensdauertests abgeglichen und schließlich bewertet. Als Ergebnis sind neben neuen, effizienten Versuchsplänen auch die relevanten Merkmale beschrieben, die es bedarf, um Versuchspläne im Kontext von Lebensdauertests zu bewerten. [Kapitel 6](#).

Abschließend stellt Kapitel 7 eine Zusammenfassung über die methodische Herangehensweise und die erreichten Ergebnisse der Arbeit zusammen und ordnet diese für künftige Forschungsbestrebungen im Bereich der multivariaten Lebensdauer-Versuchsplanung ein.

2 Stand der Forschung und Technik

Dieses Kapitel stellt die für diese Arbeit erforderlichen technischen und methodischen Grundlagen bereit. Zunächst werde in Abschnitt 2.1 zentrale Begriffe und Konzepte der Zuverlässigkeitstechnik sowie das grundlegende statistische Verfahren zur Lebensdauer-Datenanalyse in Kombination mit Versuchsplänen erläutert. Darauf aufbauend folgen in Abschnitt 2.2 die Einführung und die Einordnung von DoE für Lebensdaueruntersuchungen sowie der multivariaten Lebensdauermodellierung aus dem Stand der Technik und der Wissenschaft, die beide für die Entwicklung effizienter Lebensdauerversuchspläne maßgeblich sind. Im Kontext der Lebensdauererprobung umfasst dies insbesondere typische, statistische Versuchspläne sowie Metriken und Indikatoren zur allgemeinen Bewertung der Versuchspläne.

2.1 Zuverlässigkeitstechnik und Wahrscheinlichkeitstheorie

Die Zuverlässigkeitstechnik befasst sich mit der probabilistischen Beschreibung der Lebensdauer technischer Produkte und Systeme sowie der strategischen und statistischen Planung von Lebensdauertests. Ziel ist die statistische Modellierung des Ausfallverhaltens unter Berücksichtigung der Funktionalität des Produkts bei relevanten Randbedingungen. Eine zentrale Aufgabe besteht somit in der statistischen Charakterisierung des Ausfallbegriffs mithilfe deskriptiver Statistik sowie in der Parametrisierung geeigneter Verteilungen zur Abbildung des Lebensdauerverhaltens. Die Modellierung kann - abhängig von den Randbedingungen - auf Basis *einer einzelnen* Belastungsgröße oder *mehrerer* Beanspruchungsparameter erfolgen, die gemeinsam den Produktausfall determinieren. Ein grundlegendes Verständnis des Umgangs mit zufallsverteilten Lebensdauerereignissen ist daher eine elementare Voraussetzung für die statistische Versuchsplanung im Rahmen der Zuverlässigkeitstechnik. Weiterführende Konzepte und vertiefte methodische Ansätze zur Zu-

verlässigkeitstechnik sowie zur statistischen Testplanung sind allen voran in der Standardliteratur von Bertsche und Dazer [1] dargelegt, an deren Vorgehensweise sich die nachfolgenden Ausführungen orientieren.

2.1.1 Begriffe und Definitionen

Der **Ausfall** eines technischen Produkts bezeichnet den Zeitpunkt innerhalb seiner Lebensdauer, zu dem die geforderte Funktionalität unter definierten Umgebungs- und Randbedingungen nicht mehr erfüllt ist - also das Lebensdauerende - engl. **End-of-Life (EoL)**. Als **Belastung** werden die von außen auf ein Produkt einwirkenden Einflussparameter - Kräfte und Momente im mechanischen Kontext - bezeichnet. *Einzelne* oder zeitgleich *mehrere* Einflussparameter induzieren infolge der Produktgestalt daraus **Beanspruchungen**: innere Kräfte, Momente und lokale Spannungen. Belastung und Beanspruchung sind die maßgeblichen Faktoren, welche die Lebensdauer determinieren. Die **Ausfallzeit**, welche diese Zustandsänderung zeitlich definiert, wird im Allgemeinen als kontinuierliche Zufallsvariable $\tau > 0$ aufgefasst. So ergibt sich die Wahrscheinlichkeit, dass ein Produkt im Zeitraum bis t einen Funktionsverlust erleidet, zu

$$F(t) = \Pr(\tau \leq t) = \int_0^t f(t) dt. \quad (2.1)$$

Diese Funktion beschreibt die **Ausfallwahrscheinlichkeit**. Sie definiert damit die Verteilungsfunktion - engl. **Cumulative Distribution Function (cdf)** - für stochastische **EoL**-Events, während die **Zuverlässigkeit**

$$R(t) = \Pr(\tau > t) = 1 - F(t) = \int_t^\infty f(t) dt, \quad t \geq 0. \quad (2.2)$$

komplementär diejenige Wahrscheinlichkeit $R(t) : \mathbb{R}_{\geq 0} \rightarrow [0, 1] \subset \mathbb{R}$ quantifiziert, zu der das nicht reparierbare Produkt die realisierte Zeit t überlebt: also frei von Funktionsverlust bleibt und funktionsfähig ist [1, 4, 5]. Damit ist die Zuverlässigkeit mathematisch als reellwertige, monoton fallende und stetige Funktion definiert. Gleichwohl ist $R(t)$ keine universelle Eigenschaft, sondern vielmehr eine Funktion der Betriebsbedingungen. Diese Bedingungen umfassen unter anderem eine oder mehrere Belastungsarten und deren Niveaus, Nutzungsverhalten sowie spezifische

Betriebsprofile. Mechanische, elektrische und thermische Belastungen treten dabei am häufigsten auf [6]. Die **Wahrscheinlichkeitsdichtefunktion** - engl. **Probability Density Function (pdf)** - $f(t)$ der Ausfallzeit beschreibt, wie sich die Wahrscheinlichkeiten der Ausfälle über der Zeit verteilen. Sie folgt somit der Ableitung der *cdf*:

$$f(t) = \frac{d}{dt}F(t) = \frac{d}{dt}\Pr(\tau \leq t), \quad t \geq 0. \quad (2.3)$$

Damit repräsentiert $f(t)$ die Ausfallintensität pro Zeiteinheit und ist proportional zur lokalen Änderungsrate der Ausfallwahrscheinlichkeit. Als vierte fundamentale Größe der Zuverlässigkeitsanalyse wird außerdem die **Ausfallrate** (auch Hazard-Funktion) $\lambda(t)(t)$ eingeführt. Sie quantifiziert das momentane Ausfallrisiko eines Produkts zum Zeitpunkt t , bedingt dadurch, dass es bis zu diesem Zeitpunkt überlebt hat ($R(t) > 0$). Mathematisch ist sie als das Verhältnis der *pdf* zur Zuverlässigkeitsfunktion $R(t)$ definiert:

$$\lambda(t)(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < \tau \leq t + \Delta t | \tau > t)}{\Delta t} = \frac{1}{R(t)} \left[-\frac{dR(t)}{dt} \right] = \frac{f(t)}{R(t)}. \quad (2.4)$$

Die Ausfallrate $\lambda(t)(t)$ ist von zentraler Bedeutung, da ihr zeitlicher Verlauf (z.B. konstant, steigend, fallend) direkte Rückschlüsse auf zugrundeliegende Ausfallmechanismen wie Frühausfälle, Zufallsausfälle oder Verschleiß (vgl. "Badewannenkurve") zulässt [1, 6, 7].

2.1.2 Deskriptive Statistik für Lebensdauerdaten

Die im vorherigen Abschnitt definierten Funktionen $F(t)$, $R(t)$, $f(t)$ und $\lambda(t)(t)$ beschreiben das stochastische Ausfallverhalten eines Produktes auf einer theoretischen Populationsebene. Für die praktische Anwendung im Engineering müssen diese Funktionen, respektive die Parameter der ihnen zugrundeliegenden Verteilungsmodelle, auf Basis von empirisch ermittelten Lebensdauerdaten jedoch approximiert werden.

Die deskriptive Statistik stellt die notwendigen Methoden zur initialen Charakterisierung, Quantifizierung und Aufbereitung dieser Stichprobendaten bereit. Zur Beschreibung der Lebensdauerverteilungen sind **Lageparameter** und **Streuungs-**

maße notwendig, die zunächst theoretisch (für die Grundgesamtheit) definiert und anschließend aus der Stichprobe berechnet werden.

Der primäre Lageparameter ist der **Erwartungswert** μ der Zufallsvariable τ . Er repräsentiert den Schwerpunkt von $f(t)$ und wird für kontinuierliche Lebensdauerdaten berechnet als:

$$\mu = E[\tau] = \int_0^{\infty} t \cdot f(t) dt. \quad (2.5)$$

Ein weiterer Lageparameter ist das **Quantil** t_q der Lebensdauer. Es definiert den Zeitpunkt, zu dem $F(t)$ den Anteil q (respektive das **Perzentil** in Prozentpunkten) erreicht:

$$F(t_q) = q, \quad q \in [0, 1]. \quad (2.6)$$

Damit gibt das q - Quantil denjenigen Lebensdauerwert an, unterhalb dessen der Anteil q aller betrachteten Produkte ausgefallen ist. Ein spezieller Fall ist der **Median** $t_{0.5}$, bei dem die Ausfallwahrscheinlichkeit 50% beträgt:

$$F(t_{0.5}) = 0.5. \quad (2.7)$$

Der Median beschreibt somit den Zeitpunkt, zu dem die Hälfte aller Produkte ausgefallen ist. Damit teilt er die Fläche 1 unter der pdf in zwei gleich große Teilflächen [6, 8].

Das primäre Streuungsmaß ist die **theoretische Varianz** σ^2 , welche die mittlere quadratische Abweichung vom Erwartungswert beschreibt:

$$\sigma^2 = \text{Var}[\tau] = E[(\tau - \mu)^2] = \int_0^{\infty} (t - \mu)^2 \cdot f(t) dt. \quad (2.8)$$

Da μ und σ^2 als theoretische Parameter üblicherweise unbekannt sind, werden auch sie durch empirische Statistiken approximiert, die aus einer Stichprobe vom Umfang n (bestehend aus den Messwerten x_1, \dots, x_n) berechnet werden. Diese werden wiederum als Realisierungen der Zufallsvariable τ aufgefasst.

Das gängige empirische Äquivalent für den Erwartungswert μ ist der **arithmetische Mittelwert** \bar{x} :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2.9)$$

Analog wird die theoretische Varianz σ^2 durch die **empirische Varianz** s^2 (eine erwartungstreue Kenngröße) approximiert:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (2.10)$$

Die **empirische Standardabweichung** $s = \sqrt{s^2}$ dient entsprechend als Näherung für die theoretische Standardabweichung $\sqrt{\sigma^2}$.

2.1.3 Parametrische Lebensdauermodelle

Während die deskriptiven Statistiken \bar{x} und s^2 die zentrale Tendenz und die Streuung der vorliegenden Stichprobe quantifizieren, erlauben sie keine Extrapolation oder die Modellierung der zugrundeliegenden Funktionen $F(t)$ und $f(t)$ der Grundgesamtheit. Um eine prädiktive, mathematische Beschreibung des stochastischen Ausfallverhaltens zu erhalten, müssen die in Abschnitt 2.1.1 definierten Lebensdauerfunktionen durch geeignete parametrische Verteilungsmodelle approximiert werden. Andernfalls können nur nichtparametrische Modellierungsansätze zur Schätzung der kumulierten Wahrscheinlichkeit in Überlebensfunktionen wie beispielsweise nach Kaplan und Meier [9] genutzt werden [5, 7]. Die Verteilungsmodelle hingegen bieten eine geschlossene mathematische Form für *cdf* und *pdf* und ermöglichen es, das komplexe Ausfallverhalten durch eine geringe Anzahl von Parametern zu charakterisieren.

Weibull-Verteilung

In der Zuverlässigkeitstechnik hat sich die **Weibull-Verteilung** aufgrund ihrer hohen Flexibilität als das am häufigsten verwendete Modell etabliert. Je nach zugrundeliegendem physikalischen Ausfallmechanismus finden jedoch auch andere statistische Verteilungen Anwendung, wie beispielsweise die **Lognormal-Verteilung** (häufig bei Ermüdungs-, Korrosions- oder Diffusionsprozessen), die **Exponentialverteilung** (zur Modellierung von Zufallsausfällen ohne Alterungseffekte) oder die **Beta-Verteilung** (allgemein zur formenreichen Modellierung von R über dem festen Intervall $[0, 1]$). Für weitere Ausführungen dazu sei an dieser Stelle jedoch auf

bereits ausreichend diskutierte Aufbereitungen von Bertsche und Dazer [1], Birolini [4], Yang [6], Rigdon et al. [7] und Hedderich und Sachs [10] verwiesen.

Die (zweiparametrische) Weibull-Verteilung ist das Standardmodell zur Beschreibung der Lebensdauer von technischen Produkten ohne die Berücksichtigung eines möglichen dritten Parameters - der ausfallfreien Zeit t_0 . Sie wird durch den **Formparameter** $b > 0$ (Weibull-Modul) und die **charakteristische Lebensdauer** $T > 0$ (Skalenparameter), welche dem 63,2-ten Perzentil $t_{0,632}$ entspricht, beschrieben. Unabhängig von b gilt hier: $F(T) = 1 - e^{-1} \approx 63,2\%$. Folgt die Lebensdauer-Zufallsvariable τ dieser Verteilung, wird dies mathematisch als $\tau \sim \mathcal{W}(T, b)$ notiert. Damit ist sie in der Lage, alle drei Phasen der "Badewannenkurve" (Frühausfälle mit $b < 1$, Zufallsausfälle $b \approx 1$, Verschleißausfälle mit $b > 1$) durch die Wahl ihrer Parametrisierung abzubilden, vgl. Bertsche und Dazer [1].

Die Einheit des Skalenparameters T entspricht der Einheit des Messwertes (t in Stunden, Überrollungen, Kilometer, etc.). Die *pdf* der Weibull-Verteilung ist damit definiert als:

$$f(t) = \frac{b}{T^b} t^{b-1} \exp \left[- \left(\frac{t}{T} \right)^b \right], \quad t > 0. \quad (2.11)$$

Die *cdf* ergibt sich durch Integration der *pdf* zu:

$$F(t) = 1 - \exp \left[- \left(\frac{t}{T} \right)^b \right], \quad t > 0. \quad (2.12)$$

Aus $f(t)$ und $R(t) = 1 - F(t)$ leitet sich die **Ausfallrate** $\lambda(t)(t)$ der Weibull-Verteilung ab:

$$\lambda(t)(t) = \frac{b}{T} \left(\frac{t}{T} \right)^{b-1}, \quad t > 0. \quad (2.13)$$

Der Erwartungswert μ (vgl. Gleichung (2.5)) und die Varianz σ^2 (vgl. Gleichung (2.8)) der Weibull-Verteilung lassen sich ebenfalls in geschlossener Form ausdrücken. Sie sind von der **Gamma-Funktion** $\Gamma(\cdot)$ abhängig, welche für $x > 0$ definiert ist als:

$$\Gamma(x) = \int_0^\infty v^{x-1} \exp(-v) dv. \quad (2.14)$$

Der Erwartungswert μ der Weibull-verteilten Lebensdauer τ ergibt sich zu:

$$\mu = E[\tau] = T \cdot \Gamma \left(1 + \frac{1}{b} \right). \quad (2.15)$$

Die Varianz σ^2 ist gegeben durch:

$$\sigma^2 = \text{Var}[\tau] = T^2 \left[\Gamma \left(1 + \frac{2}{b} \right) - \Gamma^2 \left(1 + \frac{1}{b} \right) \right]. \quad (2.16)$$

Die Ausprägung der erwähnten Flexibilität der Weibull-Verteilung ist durch Abbildung 2.1 nachvollziehbar. Ein Spezialfall tritt ein für $b = 1$, da sich in diesem Fall die Weibull-Verteilung zur Exponentialverteilung mit dem Ausfallraten-Parameter $\lambda(t) = 1/T$ und dem Erwartungswert $\mu = T$ reduziert. Für $b = 3,6$ wird die Schiefe der *pdf* annähernd eliminiert, sodass sich die Weibull-Verteilung einer Normalverteilung annähert [11, 12].

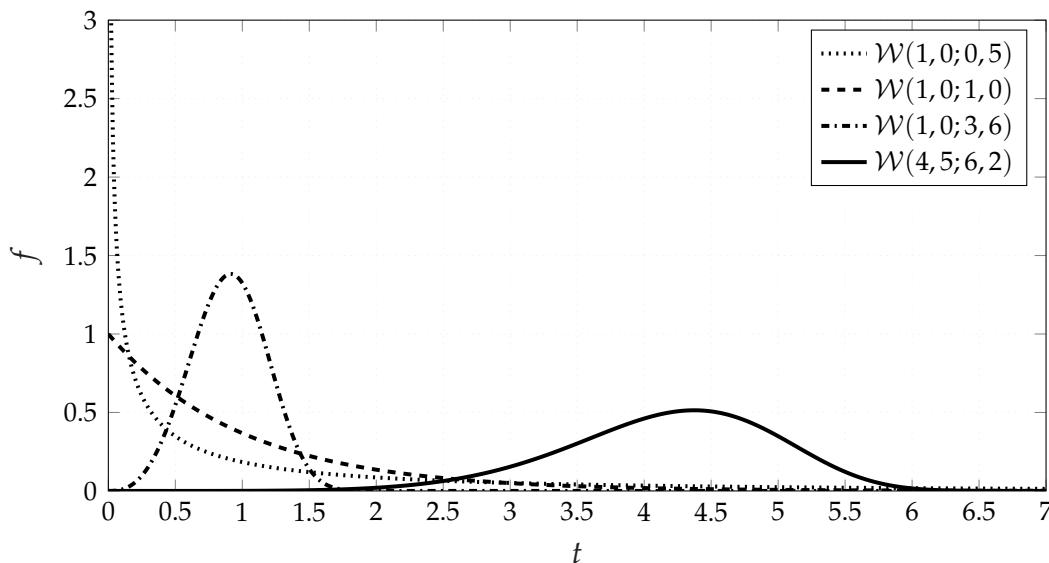


Abbildung 2.1: Weibull $f(t)$ für ausgewählte Werte von T und b .

2.1.4 Parameterschätzverfahren

Soll eine geschlossene mathematische Beschreibung des stochastischen Ausfallverhaltens eines Produktes gefunden werden, ist das im vorherigen Abschnitt 2.1.2

definierte parametrische Verteilungsmodell $\tau \sim \mathcal{W}(T, b)$ zu schätzen. Die Modellparameter der Grundgesamtheit sind in der praktischen Anwendung jedoch unbekannt. Die zentrale Problemstellung der **Parameterschätzung** besteht somit darin, aus der empirischen Stichprobe bestehend aus n Realisierungen t_1, \dots, t_n der Zufallsvariable τ statistisch fundierte Schätzwerte \hat{T} und \hat{b} zu gewinnen. Diese sind Voraussetzung, um das Lebensdauermodell (z.B. Gleichung (2.12)) zu quantifizieren und prädiktive Aussagen zu Quantilen oder der Zuverlässigkeit $R(t)$ zu ermöglichen. Eine wesentliche Komplikation hierbei sind jedoch das mögliche Auftreten von unvollständigen bzw. **zensierten** Daten sowie *multivariate* Abhängigkeiten der Belastungen zur Messgröße. Während für die Schätzung von Verteilungsparametern einfache Verfahren, wie die **Momentenmethode** oder die **Methode der kleinsten Fehlerquadrate** - engl. **Ordinary Least Squares (OLS)**, die beispielsweise bei der **Median-Rank-Regression-Methode (MMR)** im Wahrscheinlichkeitsnetz Anwendung findet, existieren, sind diese für die umfassende Analyse vielschichtiger Lebensdauerdaten in der Regel unzureichend und hier nur der Vollständigkeit wegen erwähnt - vgl. [1, 13]. Das universell anwendbare und robuste Verfahren, das Herausforderungen wie zensierte Daten und multivariate Modelle inhärent behandelt, ist die **Maximum-Likelihood-Estimation (MLE)** [5, 14].

Maximum-Likelihood-Estimation

Das Grundprinzip der **MLE** besteht darin, diejenigen Parameterwerte (z.B. \hat{T}, \hat{b}) als Schätzwerte auszuwählen, welche die Wahrscheinlichkeit (engl. Likelihood) maximieren, die empirisch beobachtete Stichprobe (bestehend aus unabhängigen Ausfällen und Zensierungen) zu erhalten. Mathematisch wird die Wahrscheinlichkeit der Realisierung von $\mathbf{t} = (t_1, \dots, t_n)$ einer Stichprobe durch die **Likelihood-Funktion** \mathcal{L} bestimmt. Diese ist eine Funktion des unbekannten Parametervektors $\boldsymbol{\theta}$, der k zu schätzende Parameter enthält (z.B. $\boldsymbol{\theta} = (T, b)$ mit $k = 2$).

Für den vereinfachten Fall, dass die Stichprobe ausschließlich aus n exakten Ausfallereignissen (vollständige Daten) besteht, ist die Likelihood-Funktion \mathcal{L} das Produkt der einzelnen Wahrscheinlichkeitsdichten $f(\cdot)$:

$$\mathcal{L}(\mathbf{t}|\boldsymbol{\theta}) \propto \prod_{i=1}^n f(t_i|\boldsymbol{\theta}). \quad (2.17)$$

Zur Vereinfachung der numerischen Berechnung wird in der Anwendung die **Log-Likelihood-Funktion** Λ verwendet. Durch die Logarithmierung wird das Produkt (Gleichung (2.17)) in eine äquivalente, leichter zu maximierende Summe überführt:

$$\Lambda := \ln(\mathcal{L}(\boldsymbol{\theta})) \propto \sum_{i=1}^n \ln[f(t_i|\boldsymbol{\theta})]. \quad (2.18)$$

Wie zuvor dargelegt, ist dieser vereinfachte Ansatz für Lebensdauerdaten jedoch oft unzureichend, da er das Auftreten von zensierten Daten vernachlässigt. Für die praktische Anwendung existiert jedoch die entsprechende Erweiterung der Likelihood-Funktion um die Differenzierung etwaiger Testausgänge als *Durchläufer*. Dazu wird die Stichprobe als Paarung von t_i, δ_i für \mathbf{t} definiert, wobei t_i der beobachteten Zeit und δ_i einem Statusindikator ($\delta_i = 1$ für einen exakten Ausfall; $\delta_i = 0$ für eine Rechts-Zensierung) entspricht [5, 15]. \mathcal{L} für rechts-zensierte Lebensdauerdaten lautet somit:

$$\mathcal{L}(\mathbf{t}|\boldsymbol{\theta}) \propto \prod_{i=1}^n [f(t_i|\boldsymbol{\theta})^{\delta_i} \cdot R(t_i|\boldsymbol{\theta})^{1-\delta_i}] \quad (2.19)$$

und definiert die Log-Likelihood Funktion als:

$$\Lambda := \ln(\mathcal{L}(\mathbf{t}|\boldsymbol{\theta})) \propto \sum_{i=1}^n [\delta_i \cdot \ln f(t_i|\boldsymbol{\theta}) + (1 - \delta_i) \cdot \ln R(t_i|\boldsymbol{\theta})]. \quad (2.20)$$

Der Parametervektor $\hat{\boldsymbol{\theta}}$, der den Wert von $\Lambda(\boldsymbol{\theta})$ maximiert, liefert die MLE-Werte. Die Schätzwerte repräsentieren die (asymptotisch) effizientesten Schätzwerte für die Parameter der Grundgesamtheit. Dies erfolgt mathematisch durch Nullsetzen k partieller Ableitungen von Λ , sofern mathematisch entsprechende Schätzwerte in geschlossener Form durch $\partial\Lambda/\partial\boldsymbol{\theta} \stackrel{!}{=} 0$ identifiziert werden können [11, 16]. Andernfalls werden numerische Optimierungsalgorithmen, vgl. Newton-Raphson-Verfahren, Patternsearch und vergleichbare, dafür herangezogen - siehe weiterführend [16, 17] sowie detaillierte Untersuchungen von Kremer und Bertsche [18]. An dieser Stelle sei erwähnt, dass systematische Verzerrungen (engl. **Bias**) in $\hat{\boldsymbol{\theta}}$ aufgrund kleiner Stichprobenumfänge auftreten können [19] - jedoch auch korrigierbar sind, vgl. Arbeiten von Hirose [20] und Ross [21].

Die Qualität der Parameterschätzung beeinflusst daraus nicht nur die Prädiktionsgüte zur Schätzung der Lebensdauer oder Zuverlässigkeit - sie bedingt schließlich

auch die Effizienz des Schätzverfahrens. Wird im Sinne eines effizienten Verfahrens zur multivariaten Lebensdauermodellbildung eine Methodik gesucht, ist auch die Qualität der Parameterschätzung damit entscheidend. Vertrauensbereiche, oder engl. **Confidence Intervals (CIs)**, können eine Metrik für die Qualität der Modellierung einnehmen, da sie die Unsicherheit oder *Unschärfe* in der Prädiktion bemessen.

Die **MLE** liefert nicht nur die Punktschätzer $\hat{\theta}$, sondern auch die Quantifizierung von deren statistischer Unsicherheit (Präzision). Obwohl verschiedene Ansätze, wie die numerisch anspruchsvolleren Berechnungen nach Likelihood-Ratio-Methode, Bootstrap-Perzentil-Methode oder Monte-Carlo-Approximation existieren, ist das gängigste Verfahren zur Berechnung von **CIs** die Approximation mittels asymptotischer Normalverteilung der **MLE**-Schätzer $\hat{\theta}$ [1, 14]. Dies erfolgt über die **Fisher-Informationmatrix** \mathbf{F} , welche die Information der Stichprobe über die Parameter θ gemäß Kremer und Bertsche [22] bezüglich des Rechenaufwands und resultierender Modellqualität vergleichsweise effizient quantifiziert. So wird diese Methodik auch in gängiger Applikationen als Standard angewandt, vgl. [6, 14, 16]. In der praktischen Anwendung wird die Fisher-Informationmatrix auf Basis der resultierenden Schätzwerte $\hat{\theta} = \theta$ so als Schätzung zur Beobachtung nach \mathbf{F}_O verwendet [14, 23]. Diese ist definiert als die negative **Hesse-Matrix** \mathbf{H}_f der Log-Likelihood-Funktion, ausgewertet an der Stelle der **MLE**-Schätzwerte $\hat{\theta}$:

$$\mathbf{F}_O := -\mathbf{H}_f(\hat{\theta}) = - \left[\frac{\partial^2 \Lambda(\theta)}{\partial \theta_j \partial \theta_l} \right]_{\theta=\hat{\theta}}. \quad (2.21)$$

Die Matrix \mathbf{H}_f entspricht der $(k \times k)$ -Matrix der k zweiten partiellen Ableitungen von Λ (vgl. Gl. (2.20)). Eine Invertierung $\hat{\mathbf{F}}_O^{-1}$ ergibt die geschätzte **Varianz-Kovarianz-Matrix** $\hat{\mathbf{V}}$:

$$\hat{\mathbf{V}} \approx \hat{\mathbf{F}}_O^{-1}. \quad (2.22)$$

Die Diagonalelemente dieser Matrix $\hat{\mathbf{V}}_{jj}$ entsprechen den Varianzen $\text{Var}(\hat{\theta}_j)$ der einzelnen Parameterschätzwerte [14]. Die Nicht-Diagonalelemente $\hat{\mathbf{V}}_{jl}$ (für $j \neq l$) repräsentieren die **Kovarianzen** $\text{Cov}(\hat{\theta}_j, \hat{\theta}_l)$ [6, 14]. Diese Kovarianzen sind von entscheidender Bedeutung, da sie die statistische Abhängigkeit zwischen den Schätzwerten (z.B. zwischen \hat{T} und \hat{b}) quantifizieren, welche für die Berechnung der **CIs** von abgeleiteten Funktionen wie $\hat{R}(t)$ erforderlich sind [5]. Basierend auf der An-

nahme der asymptotischen Normalität der Schätzer wird ein zweiseitiges $(1 - \alpha)$ -CI für einen einzelnen Parameter $\hat{\theta}_j$ direkt aus dessen Varianz approximiert durch:

$$[\theta_{j,u}, \theta_{j,o}] = \hat{\theta}_j \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\mathbf{V}}_{jj}}, \quad (2.23)$$

wobei $z_{1-\alpha/2}$ dem $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung entspricht. Da Lebensdauerparameter (z.B. T, b) üblicherweise auf $\mathbb{R}_{>0}$ beschränkt sind, werden CIs robust über eine Log-Transformation der Parameter berechnet, um physikalisch unmögliche (negative) Intervallgrenzen zu vermeiden [5, 6, 14]:

$$[\theta_{j,u}, \theta_{j,o}] = \hat{\theta}_j \exp \left(\pm z_{1-\alpha/2} \cdot \frac{\sqrt{\hat{\mathbf{V}}_{jj}}}{\hat{\theta}_j} \right). \quad (2.24)$$

Die Berechnung dieser CIs für Schätzwerte $g(\hat{\boldsymbol{\theta}})$ zu Größen wie $\hat{R}(t)$ oder \hat{t}_q erfolgt mittels **Delta-Methode** [5, 14]. Dieses auf einer Taylor-Reihenentwicklung basierende Verfahren (Gauß'sche Fehlerfortpflanzung) approximiert die Varianz der Funktion $\hat{g} = g(\hat{\boldsymbol{\theta}})$ unter Einbeziehung der gesamten Varianz-Kovarianz-Matrix. Dazu wird der **Gradientenvektor** \mathbf{g}' der Funktion g (z.B. $g = R(t)$) bezüglich des k -dimensionalen Parametervektors $\boldsymbol{\theta}$ gebildet:

$$\mathbf{g}' := \left[\frac{\partial g(\boldsymbol{\theta})}{\partial \theta_1}, \dots, \frac{\partial g(\boldsymbol{\theta})}{\partial \theta_k} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}^T. \quad (2.25)$$

Die approximierte Varianz $\text{Var}(\hat{g})$ der Funktion ergibt sich dann aus:

$$\text{Var}(\hat{g}) \approx \mathbf{g}'^T \hat{\mathbf{V}} \mathbf{g}'. \quad (2.26)$$

Das Vertrauensintervall für die Funktion \hat{g} wird anschließend unter Verwendung dieser Varianz (bzw. des Standardfehlers $\sqrt{\text{Var}(\hat{g})}$) analog zu Gleichung (2.23) berechnet [6, 14, 24]:

$$[g_u, g_o] = \hat{g} \pm z_{1-\alpha/2} \sqrt{\text{Var}(\hat{g})}. \quad (2.27)$$

Sollen auch hier nur positive Werte für g berücksichtigt werden, kann eine Logarithmierung in der Berechnung der Vertrauensbereiche analog zu Gl. 2.24 erfolgen [6].

2.2 Statistische Versuchsplanung und Modellbildung

Multivariate Lebensdauertests erfordern definitionsgemäß die Betrachtung mehrerer $k \geq 2$ Einflussfaktoren als Versuchsparameter. Dementsprechend entscheidend ist das Verständnis der wesentlichen Grundlagen im Umgang mit statistischer Versuchsplanung (**DoE**) für Lebensdauerdaten - auch unter dem Begriff **Lifetime Design of Experiments (L-DoE)** zusammengefasst - sowie der darauffolgenden Lebensdauermodellbildung. Während detaillierte Übersichten zur Historie von **DoE** von anfänglichen einschlägigen Beschreibungen durch Fisher [25], außerdem maßgebliche Weiterentwicklungen durch Box et al. [26] oder evolutionäre Schritte durch Taguchi [27] ausgiebig in Werken von Rigdon et al. [7], Kleppmann [28] und Montgomery [29] beschrieben sind, wird im Folgenden auf die wesentlichen Inhalte für die Forschungsschwerpunkte eingegangen.

Der primäre Anspruch von **DoE** besteht in der effizienten Planung empirischer Datenerhebungen, um Zielgrößen in Abhängigkeit erklärender Variablen zunächst robust zu modellieren und schließlich zu optimieren. Dieses Paradigma lässt sich auch unter der Begrifflichkeit **Design for Reliability (DfR)** unmittelbar wiedererkennen und so auf die Analyse von Lebensdauer und Zuverlässigkeit übertragen [6, 30]. Da das lokale oder globale Optimum der Lebensdauer- bzw. Zuverlässigkeitsfunktion eines Produktes a priori meist unbekannt ist, erfordert dessen Identifikation eine systematische Exploration des Parameterraumes. Eine besondere Herausforderung stellt hierbei zusätzlich die Integration von **ALT** dar: Die Diskrepanz zwischen dem hochbelasteten Testraum (engl. **Design Space**) und dem regulären **Prädiktionsraum** (engl. Use Space oder **Field Space**) kann eine **Extrapolation** erforderlich machen, welche die Anforderungen an die Daten- und somit auch an die Designqualität deutlich verschärft. Für umfassendere Ausführungen zu Forschungserkenntnissen in **ALT** sei hierbei insbesondere auf Arbeiten von Meeker et al. [5], Nelson [14], Meeker und Escobar [31] und Elsayed und Zhang [32] verwiesen. Das Unwissen zur tatsächlichen Lage optimaler Antwortwerte und die Möglichkeit, per se einen systematischen Offset zwischen Design- und Field-Space durch **ALT** vorzufinden, stellen Teststrategien nach Best-Guess Ansätzen nachteilig. Hier wird in der industriellen Praxis häufig fälschlicherweise ein **One Factor At Time (OFAT)**-Testing Ansatz gewählt - unabhängig, ob vom Vorhandensein von Lebensdauer- oder anderen Daten, welcher schlichtweg die Wahrscheinlichkeit, Optimalstellen im Parameterraum

systematisch zu treffen, senkt und somit gegenüber **L-DoE** nachteilig ist, vgl. [29, 33]. Da nun die geometrische Struktur eines Versuchsplans die erreichbare Modellierungsqualität deterministisch begrenzt, ist eine präzise Bewertung der Plangüte anhand genau dieser Eigenschaft im Vorfeld unerlässlich. Hierfür können objektive **Performance-Indikatoren** sowie mathematische **Optimalitätskriterien** dienen, vgl. [29, 34]. Ergänzend zu rationalen Metriken wie der statistischen **Trennschärfe** (engl. **Power**) und dem Schätzergebnis einer **Koeffizienten-** bzw. Parameterschätzung sind diese Größen damit bestimmend für effiziente multivariate Lebensdauertests.

Vor diesem Hintergrund fokussiert sich dieser Abschnitt auf eine gezielte Auswahl an Grundbegriffen und Metriken für multivariate Testpläne im Kontext von **L-DoE** sowie auf eine Übersicht der für Lebensdauertests geeigneten Versuchspläne, bevor abschließend die statistische Modellbildung beleuchtet wird.

Für eine grundsätzlichere Auseinandersetzung mit konventionellen Methoden und Werkzeugen von **DoE** sei, mit Blick auf den Fokus der vorliegenden Arbeit, hingegen auf die einschlägige Literatur von Kleppmann [28], Siebertz et al. [33], Hinkelmann [35] sowie vornehmlich Montgomery [29] und Myers et al. [36] verwiesen. Diese Werke behandeln intensiv die Inhalte grundsätzlicher statistischer Versuchsplanung, welche um Perspektiven zu **L-DoE** bereits durch am **Institut für Maschinenelemente (IMA)** entstandene Dissertationen von Dazer [37], Herzig [38], Grundler [39] und maßgeblich durch Kremer [40] fortschreitend ergänzt wurden. Konsequenterweise werden Hintergründe zum Umgang mit normalverteilten Daten oder Abweichungen davon im Rahmen des **DoE**, die Diskussion zu einschlägigen Vor- und Nachteilen auch unter Abgrenzung zu Alternativen wie **OFAT**, die Regressionsmodellierung auf Basis der **Varianzanalyse, engl. Analysis of Variance (ANOVA)**, konventionelle Hypothesentests sowie fundamentale Ausführungen zu **ALT** in den nachfolgenden Ausführungen nicht explizit betrachtet, sondern als bekannt vorausgesetzt.

2.2.1 Grundlagen zur statistischen Versuchsplanung

Die Anwendung von **DoE** versteht sich grundsätzlich als Verfahrenskette entlang mehrerer Prozessschritte [29, 41], die beginnend von einer spezifischen Aufgabendeinition in einer statistisch abgesicherten Testentscheidung und Datenmodellierung mündet, vgl. Abbildung 2.2. Das erklärte Ziel ist es, den kausalen Zusammenhang

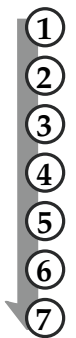
- 
- A vertical grey bar on the left side of the list, with a white circle containing a number inside it, positioned next to each step.
- ① Definition des Untersuchungsziels
 - ② Identifizierung von Einflussparametern und Systemantwort
 - ③ Auswahl von Faktoren, Stufen und Parameterbereich
 - ④ Auswahl des Versuchsplans
 - ⑤ Durchführung der Versuchsreihe gemäß Versuchsplan
 - ⑥ Statistische Datenanalyse
 - ⑦ Schlussfolgerungen und Empfehlungen/Optimierung

Abbildung 2.2: DoE Steps gemäß [41]

zwischen Einflussfaktoren und Systemantwort funktional abzubilden. Darin abgebildete Einflussfaktoren sollen also per se **statistisch signifikant** und somit relevant für das Systemverhalten sein. So kann beispielsweise die zufallsverteilte Lebensdauer τ in Abhängigkeit von $k \geq 2$ technischen Beanspruchungen zunächst empirisch untersucht und anschließend modelliert sowie optimiert werden (*Schritt 1* in Abbildung 2.2).

Im Zentrum der Betrachtung steht damit generell ein technisches **System**, welches abstrakt als Produkt oder Prozess verstanden wird und den Zustand der Ausgangsgröße in Abhängigkeit der **definiert**. Die zu untersuchende oder zu optimierende Ausgangsgröße wird als **Systemantwort** y (engl. **Response**) bezeichnet. Die gezielt kontrollierbaren und variierten Eingangsgrößen sind **Faktoren** (**Steuergrößen**), während nicht kontrollierbare oder unbekannte Einflüsse als **Störgrößen** (engl. **Noise**) klassifiziert werden (*Schritt 2*, vgl. [28]). Eine visuelle Aufstellung des genannten Zusammenspiels der Parameter kann dem Parameterdiagramm, kurz **P-Diagramm**, in Abbildung 2.3 entnommen werden [29]. Um das Systemverhalten zu charakterisieren, werden die Faktoren als kategoriale oder kontinuierliche Parameter im Versuch auf diskreten Werten, den sogenannten **Faktorstufen** (engl. **Level**), variiert (*Schritt 3*). Dies erfolgt in aller Regel in kodierter Darstellung, so entsprechen gemäß der gängigsten Konvention die Stufe -1 der niedrigen und $+1$ der hohen Einstellstufe. Die planerische Kombination verschiedener Faktorstufen äußert sich in spezifischen **Versuchspunkten** innerhalb des Parameterraums und entspricht der Versuchsplan-Matrix [28, 33]. Der **Versuchsraum** (engl. **Design Space**) wird hierbei durch die Gesamtheit der technisch realisierbaren und im Versuch einstellbaren Parameterkombinationen aufgespannt. Die Auswahl geeigneter statistischer

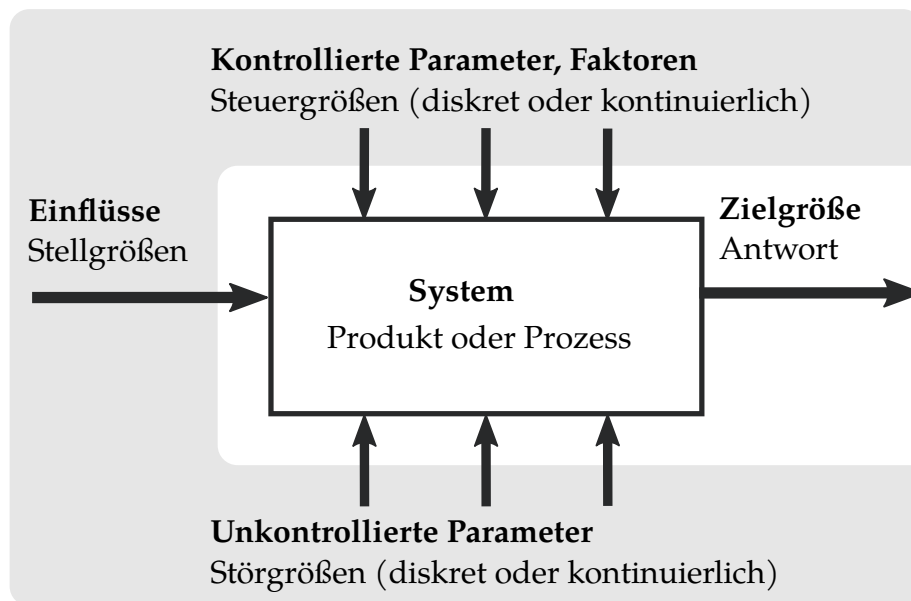


Abbildung 2.3: Parameter-Diagramm (P-Diagramm)

Versuchspläne (*Schritt 4*) für die Durchführung (*Schritt 5*) wird in Abschnitt 2.2.2 detailliert behandelt.

Die aus der Variation resultierende Änderung der Systemantwort quantifiziert den Einfluss des Faktors, der statistisch als **Effekt** E bezeichnet wird und den Mittelwertunterschiede zweier Faktorstufen beschreibt (*Schritte 6-7*). Mittels **Kontrastmethode** wird also die Änderung der Systemantwort über alle durchgeführten Versuche mit jeweiligen Faktorstufen registriert [28, 29]:

$$E = \frac{1}{n|_{+1}} \sum y|_{+1} - \frac{1}{n|_{-1}} \sum y|_{-1} \quad (2.28)$$

So können mittels DoE strukturiert, effizient und verbindlich Informationen gewonnen werden, die über die direkten Effekte hinausgehen und differenziert Aufschluss über **Haupteffekte** sowie etwaige **Wechselwirkungen** der Faktoren auf die Antwort des Systems geben - vergleiche Abb. 2.4 sowie Kleppmann [28], Montgomery [29], Siebertz et al. [33] und Kremer [40]. Abbildung 2.4 visualisiert derartige Effekte. So gibt die Darstellung eines Haupteffekts in Abhängigkeit des Vorzeichens und der Steigung (**positiver** oder **negativer** Haupteffekt) sozusagen die Einflussstärke und -richtung wieder, während bei Wechselwirkung der Effekt in Abhängigkeit der Einstellung eines **Co-Faktors** dargestellt wird. Diese Zusammenhänge können

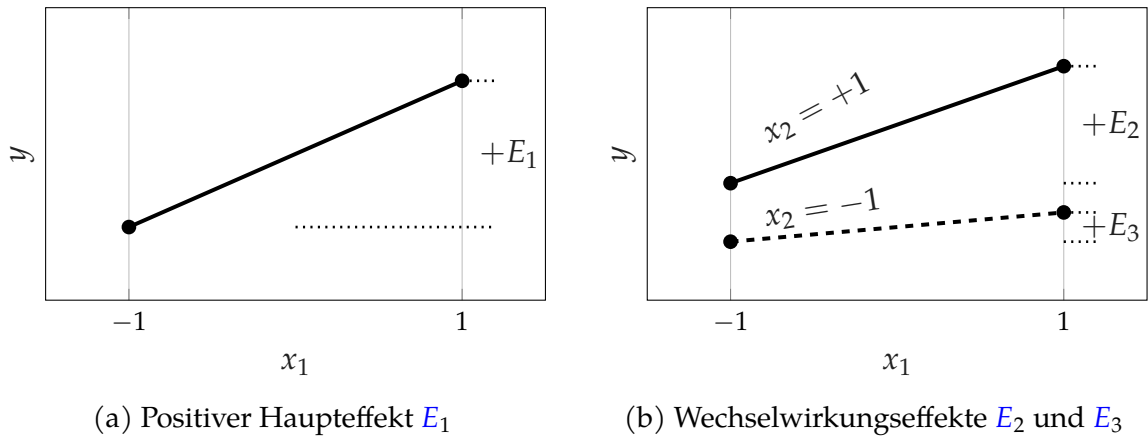


Abbildung 2.4: Schematische Darstellung der Effekte: (a) positiver Haupteffekt von Faktor x_1 , (b) Wechselwirkungseffekte zwischen x_1 und x_2 .

mathematisch positiv oder negativ beschrieben sowie durch Polynomfunktionen höherer Ordnung approximiert werden, um zusätzlich beispielsweise **quadratische Effekte** oder **Mehrfachwechselwirkungen** abzubilden. In einem einfachen Fall wird für die lineare Beschreibung des Einflusses von Faktoren auf eine Antwortvariable ein durchschnittlicher Effekt durch eine Regressionskonstante β_0 sowie Haupteffekte durch die Regressionskoeffizienten $\beta_j, j = 0, 1, \dots, k$ zu einem Regressionsmodell erster Ordnung geschätzt. Falls relevant, erfolgt die Ergänzung um die jeweilige Wechselwirkung und einen Fehlerterm ϵ als Zufallsvariable für Abweichungen durch Mess- und Streufehler [7, 29, 36]. Beispielhaft $k = 2$ Faktoren resultiert daraus:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon. \quad (2.29)$$

Ein Modell zweiter Ordnung enthält zudem quadratische Terme, welche üblicherweise in Optimierungsaufgaben - so auch in der Lebensdauer- und Zuverlässigkeitsanalyse - relevant werden können:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \epsilon. \quad (2.30)$$

Die Gleichung der Modellierung kann so zur einfacheren Handhabung auch in Matrixnotation notiert werden und resultiert in:

$$y = \mathbf{X}\boldsymbol{\beta} + \epsilon, \quad (2.31)$$

wobei β als $p \times 1$ Vektor der Regressionskoeffizienten ($p = k + 1$) durch den $n \times 1$ Vektor y aller Beobachtungen sowie durch die Versuchsplan-Matrix X ($n \times p$ Einträge) unter Zuhilfenahme eines geeigneten Schätzverfahrens (vgl. Abschnitt 2.1.4) zu ermittelten ist [6, 16]. Sollen zunächst perspektivisch relevante Faktoren für eine versuchstechnische Untersuchung identifiziert werden, kann ein **Parameter-Screening** durchgeführt werden. Dessen Durchführung kann sowohl heuristisch als auch versuchstechnisch erfolgen.

Parameter-Screening

Angesichts der potenziell hohen Komplexität durch Wechselwirkungen und Nicht-linearitäten sind die in Abbildung 2.2 beschriebenen *Schritte 2-3* als propädeutische Arbeiten für ein effizientes Testdesign zu interpretieren. Methodisch lassen sich diese unter dem Terminus **Screening** subsumieren. Screening-Schritte sind zwischen der Definition des Untersuchungsziels und der Durchführung der physischen Screening-Experimente angeordnet (vgl. *Schritt 3* in Abbildung 2.2 sowie Abschnitt 2.2.2). Daraus folgend dienen Screening-Methoden und -Versuchspläne dem Ziel, Informationsverluste bei einer minimalen Anzahl an Versuchsläufen zu begrenzen und die vitalen (*Steuergrößen*) von den trivialen (*Störgrößen*) Faktoren zu separieren, vgl. Abbildung 2.3.

Im Hinblick auf die Realisierung eines unter Zeit- und Kostenrestriktionen hochgradig effizienten DoE ist die effiziente Ausgestaltung der Screening-Strategie selbst schon von primärem Interesse. In traditionellen DoE-Ansätzen impliziert dies den Einsatz von **Kreativmethoden**, wie sie Standardliteratur von Montgomery [29] aufführen oder exemplarisch durch Kremer [40] und Gundlach [42] zusammengefasst werden. Hierbei ist ein Rückgriff auf Ergebnisse aus Experimenten, die explizit für das Forschungsziel ausgelegt wären, in dieser Phase unter Umständen noch nicht möglich. Es gilt damit zunächst, qualitativ eine rein rational erlesene Sammlung an potenziellen Einflussparameter zu erstellen, um diese dann anhand ihrer geschätzten Einflüsse auf die Systemantwort zu priorisieren. Ansätze aus der Kreativmethodik können dazu genutzt werden und fundieren auf der technischen **Systemanalyse**, die sowohl mit als auch ohne spezifisches Vorwissen über das System erfolgen kann [1]. Hilfsmittel zur Priorisierung einer hier erstellten Parametersammlung können beispielsweise Entscheidungsfindungsprotokolle, wie die **Design-Structure-**

Matrix (DSM), und Methoden aus dem Komplexitätsmanagement, z.B. **Ishikawa-Diagramm**, sein - siehe hierzu auch weiterführende Werke von Mayers [43], Pahl et al. [44], Wu und Hamada [30] und Daenzer und Haberfellner [45] sowie Lindemann et al. [46]. Das Screening liefert somit eine rational festgestellte Auswahl an möglichst wenigen Einflussparametern, die mutmaßlich den entscheidenden Anteil an statistisch begründeter Manipulation der Systemantwort tragen und sich daher für eine Untersuchung in Versuchsplänen qualifizieren. Entsprechend ist daraufhin ein geeigneter Versuchsplan für die physischen Datenerhebungen zu wählen.

2.2.2 Statistische Versuchsplanung zur Lebensdauererprobung

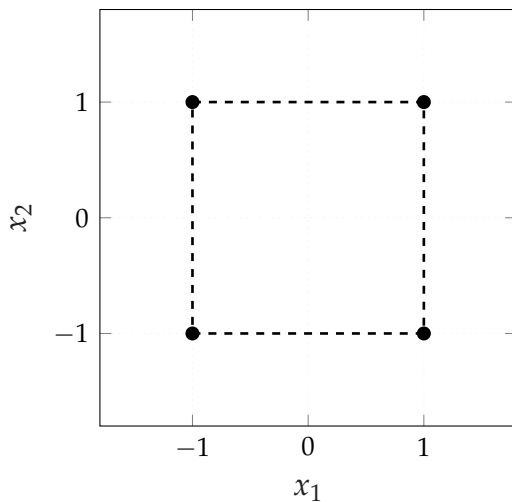
Standardprotokolle aus dem DoE wie der 2^k **voll-faktorielle Versuchsplan** eignen sich grundsätzlich auch für Lebensdaueruntersuchungen, da sich hier analog zu vergleichbar statistisch verteilter Datenlage Effekte stets als (Mittelwert-) Unterschiede in der Beobachtung der Systemantwort aus dem Vergleich zweier Einstellstufen eines oder mehrerer Faktoren ergeben [28].

2^k Faktorielle Versuchspläne

Demzufolge kann auch ein Lebensdauer-beeinflussendes Parameterset - beispielhaft (x_1, x_2) - voll-faktoriell auf zwei Stufen variiert und vollständig kombiniert werden, vgl. Abbildung 2.5. Ein derartiges Setup erlaubt es, die perspektivische Differenz erreichbarer EoL-Werte durch tiefe und hohe Beanspruchungswerte der Faktoren zu beobachten [5, 6]. Der voll-faktorielle Versuchsplan bildet somit den Standard-Versuchsplan im DoE und fordert bei einmaliger Durchführung (**Replikation** $m = 1$)

$$n = 2^k \quad (2.32)$$

Versuche. Dieser Stichprobenumfang stellt sicher, dass das resultierende Gleichungssystem **gesättigt** ist: Mit n Versuchen lassen sich $n - 1$ Effekte für Hauptfaktoren und Wechselwirkungen eindeutig bestimmen. Von entscheidender Bedeutung für die Aussagekraft des Versuchsplans ist die Wahl der Faktorstufen (vgl. Abbildung 2.6). Die Differenz der gewählten Level muss bereits im Vorfeld definiert werden, sodass signifikante Effekte sicher detektiert werden („**Signal-to-Noise**“), wobei gleichzeitig zu geringe Abstände (Rauschen) sowie zu große Intervalle (Gefahr unerkannter



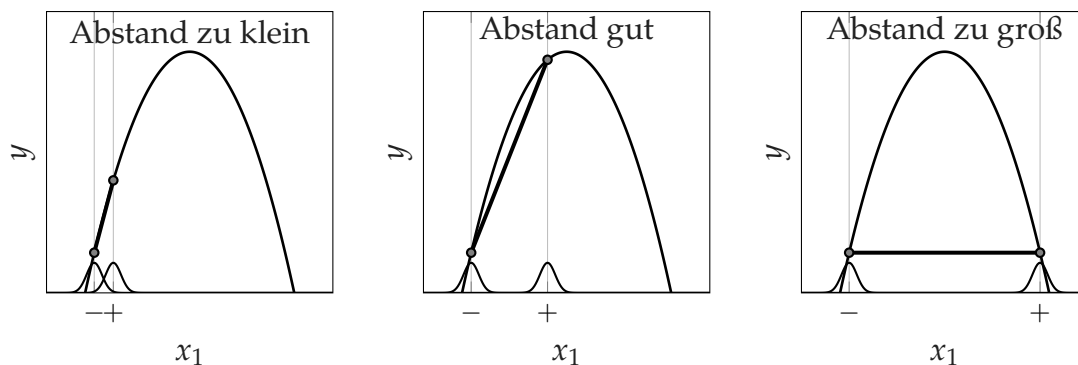
(a) Versuchsplan

Faktorstufen- Kombination	Faktoren und Wechselwirkung		
	#	x_1	x_2
1	1	-1	-1
2	2	+1	-1
3	3	-1	+1
4	4	+1	+1

(b) Versuchsplanmatrix zu Abb. 2.5a

Abbildung 2.5: Standard voll-faktorieller Versuchsplan

Nichtlinearitäten) zu vermeiden sind - vgl. Abbildung 2.6 sowie Kleppmann [28], Wu und Hamada [30] und Siebertz et al. [33]. Auf Basis eines solchen zweistufigen


Abbildung 2.6: Einfluss der Schrittweite auf die Approximation des Effekts E

Setups lässt sich der Zusammenhang zwischen Einflussgrößen und Lebensdauer interpolieren und in einem linearen Modell abbilden, welches zudem die Schätzung von Wechselwirkungen erlaubt. Werden davon abweichende Modellterme zur Abbildung der Systemantwort erwartet, berücksichtigen alternative Versuchspläne typischerweise drei bis fünf Faktorstufen. Der voll-faktorielle Versuchsplan nimmt dabei eine entscheidende Schlüsselrolle in der strategischen Modellbildung ein - insbesondere im Hinblick auf Lebensdauerdaten und Zuverlässigkeitstechnik.

Darin aufgeführte Versuchspunkte können auf Basis der Struktur ihrer Zuordnung üblicherweise ideal aus Voruntersuchungen übernommen oder durch weiterführende Untersuchungen nachfolgend erweitert werden. Ausgehend vom qualitativen Parameter-Screening (vgl. Abschnitt 2.2.1) ist ohne initiale Experimente oft unklar, welche Faktoren die Antwortvariable, also beispielsweise die Lebensdauer eines Systems, nun tatsächlich signifikant beeinflussen. Folglich ist es essenziell, diese Fragestellung vor der eigentlichen Versuchsplanumsetzung effizient zu klären.

2^{k-p_f} Fraktionell Faktorielle Versuchspläne

Sind nach Anwendung der Kreativmethoden (qualitatives Screening) weiterhin so viele Einflussfaktoren als relevant eingestuft, dass ein voll-faktorieller Ansatz gemäß Gleichung 2.32 zu einem wirtschaftlich nicht vertretbaren Versuchsumfang führen würde, muss die Strategie hin zu physikalischen Screening-Tests verschärft werden. Dies empfiehlt sich insbesondere für Systeme mit $k > 5$ Faktoren, um die experimentelle Effizienz zu gewährleisten. Zur Veranschaulichung der Notwendigkeit: Bereits eine einzelne Replikation eines voll-faktoriellen Experiments mit $k = 8$ Faktoren würde $2^8 = 256$ Versuchsdurchläufe erfordern, was in der Lebensdauererprobung meist illusorisch ist.

Für derartige Selektionsaufgaben eignen sich daher **Screening-Versuchspläne**, wie der **teil-faktorielle Versuchsplan** (Fractional Factorial Design) oder alternativ der **Plackett-Burman-Plan** [28, 29, 33]. Bei diesem Ansatz wird lediglich eine selektive Teilmenge (Fraktion) der voll-faktoriellen Versuchsagenda umgesetzt, um mit minimalem Informationsverlust die für den Anwendungsfall signifikanten Effekte zu beschreiben. Mathematisch wird die Anzahl der Versuche dabei auf

$$n = 2^{k-p_f} \quad (2.33)$$

reduziert, wobei p_f den Grad der Fraktionierung (die Anzahl der Generatoren) angibt. Die Validität dieses Vorgehens stützt sich auf zwei fundamentale empirische Postulate [7, 13]:

- Die **Effekthierarchie** besagt, dass Effekte niedrigerer Ordnung – primär Haupteffekte – in der Regel eine größere Amplitude aufweisen und mit höherer Wahrscheinlichkeit signifikant sind als Effekte höherer Ordnung.

- Die **Effektvererbung** impliziert, dass das Auftreten signifikanter Wechselwirkungen oder quadratischer Terme strukturell an die Signifikanz ihrer korrespondierenden Haupteffekte gekoppelt ist.

Eine direkte Konsequenz dieser Reduktion („Der Preis der Einsparung“) ist jedoch, dass sich in teil-faktoriellen Versuchsplänen bestimmte Effekte nicht mehr isoliert betrachten lassen (vgl. Abbildung 2.5a für den Fall einer Fraktionierung). So sind beispielsweise Haupteffekte unter Umständen nicht mehr zweifelsfrei von Wechselwirkungen höherer Ordnung zu unterscheiden. Da sich diese Effekte statistisch überlagern, spricht man von einer **Vermengung** (engl. **Aliasing**). Die Schwere dieser Vermengung wird dabei über die **Auflösung** (engl. Resolution) des Versuchsplans klassifiziert (z. B. Auflösung III, IV oder V). Wird dieser Informationsverlust jedoch bewusst in Kauf genommen und ingenieurwissenschaftlich bewertet, ermöglicht dies eine signifikante Reduktion des Versuchsumfangs, um effizient die dominanten Faktoren aus der initialen Parametermenge zu isolieren. Diese Vorgehensweise ist von hoher Relevanz, da Lebensdauertests – ob als beschleunigte Prüfung mittels **ALT** oder unter Feldbedingungen – durch die inhärente Zeitabhängigkeit der Systemantwort erhebliche Kapazitäten binden und Ergebnisse nicht ad hoc verfügbar sind. Im Sinne einer ressourceneffizienten Gesamtstrategie sollten die Screening-Versuche daher idealerweise so konzipiert sein, dass sie nahtlos in einen nachfolgenden, höher aufgelösten Versuchsplan integriert (**augmentiert**) werden können.

Wirkungsflächenversuchspläne

Die bisher diskutierten Versuchspläne beschränken sich auf die Untersuchung von Faktoren auf jeweils zwei Stufen (± 1). Dies ermöglicht zwar eine effiziente Darstellung linearer Beziehungen und Interaktionen, jedoch ist die Modellierung komplexerer, nicht-linearer Effekte aufgrund fehlender Stützstellen im Versuchsraum damit physikalisch nicht möglich. Perspektivisch ist daher entscheidend, wie die bestehende Datenbasis weitergenutzt und augmentiert werden kann, falls die Systemantwort signifikante **Krümmungen** (engl. **Curvature**) aufweist und die quantitativen Beziehungen zwischen Faktoren und Zielgröße für eine Optimierung detaillierter beschrieben werden müssen.

Die **Response Surface Methodology (RSM)** behandelt als Teildisziplin des **DoE** derartige Herangehensweisen und bietet hierfür spezielle **Wirkungsflächenversuch-**

spläne - engl. **Response Surface Designs (RSDs)** - an. Der erste Schritt zur Detektion von Nichtlinearitäten besteht in der Integration von n_C sogenannten **Zentralpunkten**, engl. **Center Points (CPs)**, in den faktoriellen Basisplan. Hierbei werden alle Faktoren auf die kodierte Stufe 0 (die Mitte des Versuchsraums) gesetzt. Weicht der Mittelwert der Systemantwort in den Zentralpunkten signifikant vom Mittelwert über die faktoriellen Eckpunkte ab, deutet dies auf eine Krümmung der Antwortfläche hin [29].

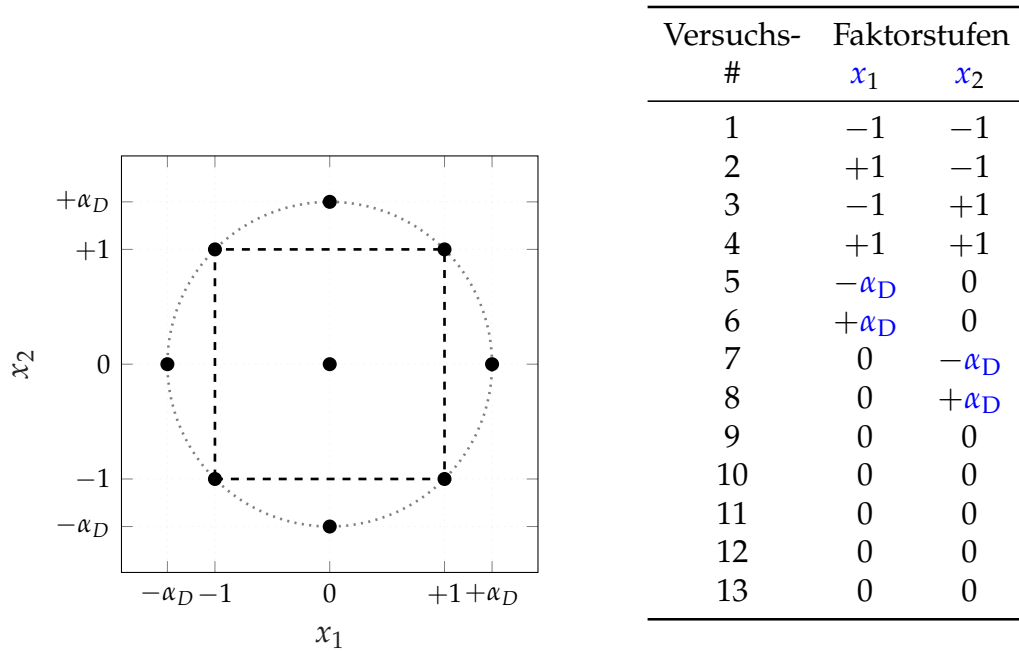
Um diese quadratischen Zusammenhänge explizit zu bestimmen, gilt der **Zentral-Zusammengesetzte-Versuchsplan**, engl. **Central Composite Design (CCD)**, als etablierter Standard. Ein **CCD** entsteht durch die Augmentierung des ursprünglichen voll- (oder teil-)faktoriellen Plans (den n_F Eckpunkten), vgl. Abbildung 2.7, um:

- eine definierte Anzahl n_C an wiederholten Zentralpunkten (üblicherweise $3 \leq n_C \leq 5$ zur Abschätzung des reinen Fehlers) sowie
- $n_S = 2 \cdot k$ zusätzliche **Sternpunkte** (engl. **Axial Points** oder **Star Points**), die auf den Achsen des Koordinatensystems im Abstand $\pm\alpha_D$ vom Zentrum liegen.

Während die Zentralpunkte das Vorhandensein quadratischer Effekte validieren, ermöglichen die Sternpunkte deren Wertbestimmung. Die Wahl des Abstands α_D wird primär durch die Geometrie des interessierenden Versuchsraums (**Region of Interest**) diktiert. Betrachtet man diesen Raum als Kugel (**sphärisch**), ist die **Drehbarkeit** (engl. **Rotatability**) ein wesentliches Qualitätsmerkmal. Sie stellt sicher, dass die Varianz der Vorhersage nur vom Abstand zum Zentrum abhängt und invariant gegenüber einer Rotation des Koordinatensystems ist. Um diese Eigenschaft zu gewährleisten, berechnet sich der Abstand α_D in Abhängigkeit der faktoriellen Versuchspunkte n_F zu [29, 47]:

$$\alpha_D = (n_F)^{1/4}. \quad (2.34)$$

Alternativ kann für sphärische Räume auch $\alpha_D = \sqrt{k}$ gewählt werden (**Sphärisches CCD**), wodurch alle Versuchs- und Sternpunkte auf einer Kugeloberfläche liegen [36]. Ist der Versuchsraum hingegen durch harte physikalische Grenzen (z.B. maximale Temperatur) kubisch beschränkt, bietet sich der **flächenzentrierte CCD** (**Face**



(a) Versuchsplan (CCD)

(b) Versuchsplanmatrix (Stufen) zu Abb. 2.7a

Abbildung 2.7: Zentral zusammengesetzter Versuchsplan (CCD) mit $k = 2$

Centered CCD) an. Hierbei wird $\alpha_D = 1$ gesetzt, sodass die Sternpunkte direkt auf den Flächenmitten des Würfels liegen. Dies vereinfacht die Durchführung, da nur drei Faktorstufen ($-1, 0, +1$) benötigt werden, opfert jedoch die Eigenschaft der Rotierbarkeit. Eine effiziente Alternative zum CCD stellt das **Box-Behnken-Design** dar [48]. Dieses Design kombiniert 2^k -Faktorielle mit unvollständigen Blockplänen und platziert Versuchspunkte auf den Kantenmitten des Versuchsraums, vermeidet jedoch die extremen faktoriellen Versuchspunkte. Dies kann vorteilhaft für Lebensdauertests sein, bei denen extreme Ecken oft zu verfrühten Ausfällen führen können. Zudem ist es bei korrekter Wahl der Zentralpunkte alias-optimal gegenüber kubischen Modellen [49]. Stoßen Standard-Designs (CCD, Box-Behnken) aufgrund von Restriktionen im Versuchsraum (**Constraints**), nicht-standardmäßigen Modellierungszielen oder ungewöhnlichen Stichprobenumfängen an ihre Grenzen, empfiehlt sich der Einsatz von **Optimalen Versuchsplänen**, vergleiche übernächsten Abschnitt. Diese computergenerierten Designs minimieren algorithmisch die durchschnittliche Vorhersagevarianz über den gesamten Designraum und können Standard-Designs in ihrer Prädiktionsgüte oft übertreffen [29, 34].

Strategische Vorgehensweisen

In der Praxis entsteht ein **CCD** häufig im Rahmen einer **sequenziellen Versuchssstrategie**: Zeigt das initiale lineare Modell Anpassungsmängel (**Lack-of-Fit**), wird der bestehende faktorielle Plan um die Sternpunkte augmentiert, um ein Modell zweiter Ordnung zu fitten [36]. Modelle zweiter Ordnung sind insbesondere in der Lebensdauer- und Zuverlässigkeitstechnik vermehrt von Bedeutung, da sie die Existenz von lokalen Extrema (Minima oder Maxima) der Systemantwort im Versuchsraum detaillierter abbilden können [7, 50]. Prinzipielle Vorgehensweisen, die ein erfolgreiches Umsetzen von **DoE** bzw. **L-DoE** begünstigen sollten, wie **Blockbildung** oder **Randomisierung** (vgl. [28, 33]) seien hier bereits vorausgesetzt. Abbildung 2.8 fasst eine derartige strategische Vorgehensweise zur Realisierung eines **CCD** zusammen, wie sie beispielsweise Box et al. [51] und Bisgaard [52] vorschlagen.

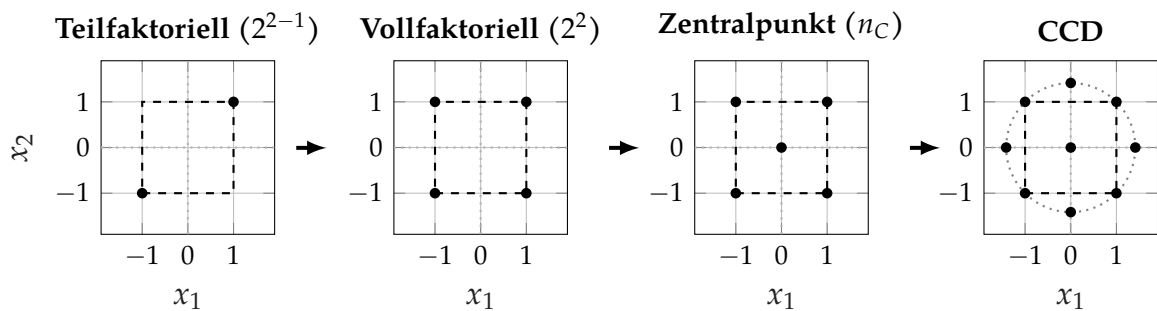


Abbildung 2.8: Strategischer Ansatz zu augmentierter Versuchsplanung für die Lebensdauererprobung nach [51, 52]

Um effizient Versuchspunkte zu allokalieren und für sequentiell fortlaufende Versuche weiterverwendet zu werden, wird in Abgleich mit Abbildung 2.2 zunächst ein **teil-faktorieller Versuchsplan** zur Identifikation signifikanter Faktoren durchgeführt. Anschließend wird der Plan zu einem **voll-faktoriellen Versuchsplan** erweitert, um alle Haupteffekte und Wechselwirkungen erster Ordnung zu schätzen. Daraufhin werden **Zentralpunkte** ergänzt, um das Vorhandensein von Krümmungen zu validieren. Abschließend werden bei Bedarf die **Sternpunkte** hinzugefügt, um ein vollständiges **CCD** zu realisieren. Insbesondere in der Lebensdauererprobung ist eine derartige schrittweise Vorgehensweise sinnvoll, um den Versuchsumfang zu minimieren und dennoch eine fundierte Modellierung der Systemantwort

zu gewährleisten. Gleichzeitig erlaubt diese Strategie eine flexible Anpassung an unerwartete Ergebnisse in den einzelnen Phasen der Untersuchung, während lokale Optima von Zuverlässigkeiten im Sinne von **DfR** im Vorfeld ohnehin nicht bekannt sind. So kann beispielsweise auf Basis der Ergebnisse des teil-faktoriellen Plans entschieden werden, ob eine Erweiterung zum voll-faktoriellen Plan überhaupt notwendig ist, oder ob direkt zu den Zentralpunkten übergegangen wird. Genauso kann die Anzahl der Zentralpunkte an die beobachtete Streuung der Lebensdauerdaten angepasst werden, um eine robuste Schätzung des Fehlers zu gewährleisten. Ist dann eine Extrapolation in Richtung a-priori unbekannter Beanspruchungsniveaus geplant, verfügt der **CCD** über die notwendigen Stützstellen, um eine adäquate Exploration des Versuchsraums zu ermöglichen.

Metriken zur optimalen Versuchsplanung

Wird maximale Flexibilität jenseits starrer Standard-Designs wie dem voll-faktoriellen Setup oder dem **CCD** verlangt, können **optimale Versuchspläne** einen leistungsfähigen, *effizienten* Lösungsansatz für *multivariate* Versuchsplanung bieten [29, 36, 53]. Diese verfolgen das Ziel, die Versuchspunkte algorithmisch so im Versuchsraum zu positionieren, dass spezifische Qualitätskriterien - allen voran die Verteilung der (um die Versuchsanzahl N skalierten) **Prädiktionsvarianz**, engl. **Scaled Prediction Variance (SPV)**,

$$v(\mathbf{x}_0) = \frac{N \text{Var}[\hat{y}(\mathbf{x}_0)]}{\sigma^2} = \mathbf{x}_0' \left(\frac{\mathbf{M}}{N} \right)^{-1} \mathbf{x}_0 = N \mathbf{x}_0' \mathbf{M}^{-1} \mathbf{x}_0 \quad (2.35)$$

optimiert werden [29, 36]. Die Prädiktionsvarianz ist ein dimensionsloses Maß, das ausschließlich über die geometrische Anordnung der N Versuchspunkte (der Matrix **X**) in der **Informationsmatrix** (auch **Momentenmatrix**)

$$\mathbf{M} \stackrel{m=1}{=} \mathbf{X}'\mathbf{X} \quad (2.36)$$

definiert und damit positionsabhängig von einer Referenz $\mathbf{x}_0' = [1, x_{01}, \dots, x_{0k}]$ im Parameterraum ist [36]. Sie ist damit unabhängig von der tatsächlichen Streuung der Messdaten (σ^2). Weiter wird sie über die Anzahl der Versuchspunkte N skaliert, um Vergleiche zwischen Versuchsplänen mit unterschiedlicher Stichpro-

bengröße zu ermöglichen [34, 36] - kann aber auch unskaliert betrachtet werden: als **Unscaled Prediction Variance (UPV)** $= \text{Var}[\hat{y}(\mathbf{x}_0)] / \sigma^2 = \mathbf{x}_0' \mathbf{M}^{-1} \mathbf{x}_0$, vgl. Montgomery [29] und Myers et al. [54]. Zusammen mit der Varianz des Versuchsfehlers σ^2 (vgl. Gleichung 2.38) erlaubt die Prädiktionsvarianz somit eine direkte Quantifizierung der **Unsicherheit in der Modellvorhersage** $\text{Var}[\hat{y}(\mathbf{x}_0)]$ an einer beliebigen Stelle \mathbf{x}_0 im Versuchsraum. Um sie über verschiedene Stellen im Parameterraum zu bewerten, konnten allen voran Zahran et al. [55] und Giovannitti-Jensen und Myers [56] mit dem **Variance Dispersion Graph (VDG)** oder **Fraction of Design Space (FDS)-Plot** Visualisierungsstrategien entwickeln. Eine effiziente Versuchsplanung für multivariate Lebensdaueruntersuchungen muss sich neben pragmatischen Beweggründen maßgeblich an diesen Größen orientieren, weshalb im Folgenden eine Übersicht der gängigen Metriken und Qualitätsmerkmale gegeben wird.

Der Einsatz optimaler Versuchspläne ist zunächst prädestiniert für Szenarien, in denen klassische Pläne an ihre Grenzen stoßen - sei es durch physikalische Randbedingungen welche bestimmte Faktorstufenkombinationen ausschließen, oder durch strikte Limitierungen der verfügbaren Versuchskapazität [28, 34]. Derart klassische Versuchspläne wie der voll-faktorielle Versuchsplan sind aufgrund ihrer Versuchspunktanordnung stets **orthogonal** [7, 29]. Ein Versuchsplan wird als orthogonal bezeichnet, wenn keine Korrelation zwischen jeweils zwei Spalten der Versuchsplan-Matrix vorliegt - deren Skalarprodukte also jeweils null ergeben:

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0 \quad \text{für alle } i \neq j. \quad (2.37)$$

Oder in anderen Worten: entspricht die Informationsmatrix \mathbf{M} einer Diagonalmatrix, ist der Versuchsplan orthogonal [7, 13]. So können Effekte eindeutig identifiziert werden - die Vektoren der Faktorstufenkombinationen sind linear unabhängig und Effekte lassen sich unverzerrt schätzen. Dies stellt einen Versuchsplan also zunächst einmal qualitativ günstig dar.

Zudem liegt **Ausgewogenheit** vor, sofern für einen jeweiligen Faktor alle anderen Faktoreinstellungen gleichmäßig aufgeteilt sind [33]. Damit wird Varianzhomogenität und Gleichbehandlung der Faktoren gewährleistet, sodass die Schätzung der Effekte unverzerrt erfolgt.

Mathematisch fundiert die Bewertung der Schätzgenauigkeit auf der Inversen der Informationsmatrix, welche als **Varianz-Kovarianz-Matrix** (oder **Dispersionsmatrix**) $(\mathbf{X}'\mathbf{X})^{-1}$ der Regressionskoeffizienten definiert ist:

$$\text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}, \quad (2.38)$$

wobei wie gehabt σ^2 die Varianz des Versuchsfehlers darstellt. Die Diagonalelemente der Matrix $(\mathbf{X}'\mathbf{X})^{-1}$ stehen hierbei in direktem Zusammenhang mit dem **Varianz-Inflations-Faktor (VIF)** [28, 33]. Der **VIF** dient als Maßzahl für Multikollinearität und quantifiziert den Faktor, um den sich die Varianz eines geschätzten Koeffizienten im Vergleich zu einem vollständig orthogonalen Design aufgrund von Korrelationen zwischen den Faktoren erhöht (vergleiche auch **Konditionszahl**) [10]. Während bei orthogonalen Plänen (Idealfall) ein **VIF** = 1 vorliegt, deuten hohe Werte (typischerweise > 5 oder > 10) auf eine instabile Modellschätzung hin [13, 28].

Ergänzend zur globalen Bewertung der Multikollinearität durch den **VIF** erlaubt die Betrachtung der sogenannten **Prädiktionsmatrix** (engl. auch **Hat-Matrix**)

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (2.39)$$

eine lokale Diagnose der Versuchspunkte. Die Diagonalelemente $h_{ii} \in \left[\frac{1}{n}, 1\right]$ dieser Projektionsmatrix, bezeichnet als **Hebelwerte** (engl. **Leverage**), quantifizieren den Einfluss eines einzelnen Versuchslaufs auf die Modellvorhersage. Punkte mit hohen Hebelwerten (typischerweise $h_{ii} > 2p/n$) befinden sich geometrisch weit vom Zentrum des Versuchsraums entfernt und dominieren die Regression, was den Plan anfällig für Ausreißer in diesen spezifischen Einstellungen macht [13, 33].

Während Hebelwerte jedoch lediglich das *Potenzial* einer Beobachtung zur Modellbeeinflussung aufgrund ihrer geometrischen Exponiertheit indizieren, quantifizieren Einflussstatistiken die *tatsächliche* Auswirkung auf die Regressionsparameter und die Vorhersagegüte. Die Identifikation solcher Beobachtungen erfolgt methodisch durch den iterativen Ausschluss des i -ten Datensatzes (*Leave-One-Out-Methodik*) und den Vergleich der resultierenden Modellstatistiken mit dem ursprünglichen Modell basierend auf n Beobachtungen [13, 57]. Als globales Maß für den Einfluss der i -ten Beobachtung auf den Vektor aller geschätzten Regressi-

onskoeffizienten $\hat{\beta}$ dient die **Cook-Distanz** (D_i). Sie verknüpft die Information des intern studentisierten Residuums r_i mit dem Hebelwert h_{ii} :

$$D_i = \frac{r_i^2}{p} \cdot \frac{h_{ii}}{1 - h_{ii}}. \quad (2.40)$$

Ein hoher Wert (typischerweise $D_i > 1$ oder $D_i > 4/n$) signalisiert, dass das Entfernen der Beobachtung zu einer signifikanten Verschiebung der Modellparameter führen würde, da der Punkt sowohl weit vom Zentrum liegt als auch einen großen Fehler aufweist [57, 58].

Für eine differenzierte Analyse auf Ebene der einzelnen Modellterme wird die Metrik **DFBETAS** (Difference in Betas, Standardized) herangezogen. Sie misst die standardisierte Änderung eines spezifischen Regressionskoeffizienten β_j bei Ausschluss der i -ten Beobachtung:

$$\text{DFBETAS}_{j,i} = \frac{\hat{\beta}_j - \hat{\beta}_{j(-i)}}{\sqrt{\hat{\sigma}_{(-i)}^2 (\mathbf{X}'\mathbf{X})_{jj}^{-1}}}, \quad (2.41)$$

wobei $\hat{\beta}_{j(-i)}$ den Koeffizienten ohne die i -te Beobachtung und $\hat{\sigma}_{(-i)}^2$ die entsprechende Fehlerquadratschätzung darstellt. Hiermit lässt sich prüfen, ob einzelne Versuche die Schätzung eines Effekt-Terms verzerren (kritischer Schwellenwert oft $> 2/\sqrt{n}$) [13, 57].

Ergänzend beschreibt **DFFITS** (Difference in Fits) die Änderung des Vorhersagewertes an der Stelle i selbst, normiert auf die Standardabweichung der Anpassung:

$$\text{DFFITS}_i = \frac{\hat{y}_i - \hat{y}_{i(-i)}}{\sqrt{\hat{\sigma}_{(-i)}^2 h_{ii}}}. \quad (2.42)$$

Abschließend bewertet die **COVRATIO** den Einfluss auf die Präzision der Schätzung. Sie setzt die Determinante der Varianz-Kovarianz-Matrix ohne die i -te Beobachtung ins Verhältnis zur ursprünglichen Matrix:

$$\text{COVRATIO}_i = \frac{\det((\mathbf{X}'_{(-i)}\mathbf{X}_{(-i)})^{-1})}{\det((\mathbf{X}'\mathbf{X})^{-1})}. \quad (2.43)$$

Werte, die signifikant von 1 abweichen (Grenzen $1 \pm 3p/n$), zeigen an, dass die Beobachtung die Konfidenzbereiche der Koeffizienten unverhältnismäßig beeinflusst [57].

Da in restriktiven Versuchsräumen Orthogonalität oft nicht erreichbar ist, zielen optimale Versuchspläne darauf ab, durch Minimierung spezifischer Eigenschaften der Dispersionsmatrix $(\mathbf{X}'\mathbf{X})^{-1}$ den Informationsgehalt trotz Korrelationen zu maximieren. Die Generierung solcher maßgeschneiderten Pläne erfolgt algorithmisch unter Maximierung spezifischer statistischer Gütekriterien. Eine Übersicht der einschlägigen **Optimalitätskriterien** - oder einfach **Optimalitäten** - welche den quantitativen Vergleich zur Standard-Methodik ermöglichen, ist in Tabelle 2.1 zusammengefasst [29, 34]. Sie bilden mitunter die Grundlage für diverse weitere Optimalitäten und hybride Ansätze und lassen sich prinzipiell in Modellierungskriterien (*A*-, *D*-Optimalität) und Prädiktionskriterien (*G*-, *I*-, *U*-, *V*- **Optimalität**) unterteilen [7, 36].

Jenseits praktischer Kriterien sind hier der Vollständigkeit halber auch eine Teilmenge eher theoretischerer Optimalitäten zu nennen. Die **E-Optimalität**, führt zur Minimierung des maximalen Eigenwerts der Dispersionsmatrix, was eine Worst-Case-Absicherung zur Varianzschätzung begünstigt - vergleiche Boyd und Vandenberghe [59] und Russell [60]. Für explorative Zielsetzungen ohne starre Modellannahmen kann nach Atkinson [61] die **S-Optimalität** relevant sein, da hier durch die Maximierung der euklidischen Abstände benachbarter Punkte eine gleichmäßige Raumfüllung (*Space-Filling*) sichergestellt werden kann. So können für sich ändernde Modelle entlang einzelner Parameter wie beispielsweise $\hat{T} \in \hat{\theta}$ Sensitivitätsanalysen effizient durchgeführt werden. Eine Verallgemeinerung der prädiktionsorientierten Kriterien stellt die **Q-Optimalität** dar, welche analog zur *I*-Optimalität die integrierte Prädiktionsvarianz minimiert, jedoch mittels Gewichtsfunktionen eine differenzierte Priorisierung spezifischer Regionen im Versuchsraum erlaubt, vgl. Goos und Jones [34]. Liegt der Fokus hingegen isoliert auf einer Teilmenge der Modellparameter - etwa zur Trennung von Haupteffekten und Blockeffekten oder von Termen erster und zweiter Ordnung - ermöglicht die **D_S-Optimalität** als Derivat der *D*-Optimalität eine gezielte Maximierung der Schätzgüte für genau dieses jeweilige Subset der Parametergruppe - siehe auch Werke von Goos und Jones [34], Myers et al. [36] und Atkinson et al. [62].

Tabelle 2.1: Übersicht, Zielsetzung und mathematische Definition verschiedener Optimalitätskriterien für Versuchspläne

Kriterium	Zielgröße, Beschreibung und Definition
A-Optimalität	<p>Minimierung der Spur der inversen Informationsmatrix [7, 28, 29].</p> <p><i>Ziel: Maximale Präzision der einzelnen Parameter im Durchschnitt.</i></p> $A_{\text{opt}} = \min \left(\text{spur} \left(\mathbf{M}^{-1} \right) \right) = \min \left(\sum_{j=1}^k \text{Var} \left(\hat{\beta}_j \right) \right) \quad (2.44)$
D-Optimalität	<p>Maximierung der Determinante der Informationsmatrix \mathbf{M}. Da das Volumen des Vertrauensellipsoids der Koeffizientenschätzwerte umgekehrt proportional zur Quadratwurzel der Determinante ($\sqrt{\det(\mathbf{M})}$) ist, minimiert dieses Kriterium das Unsicherheitsvolumen im Parameterraum [28, 29, 36].</p> <p><i>Ziel: Minimierung des Volumens des gemeinsamen Vertrauensbereichs aller Modellparameter.</i></p> $D_{\text{opt}} = \max \left(\det(\mathbf{M}) \right) = \min \left(\det \left(\mathbf{M}^{-1} \right) \right) \quad (2.45)$
G-Optimalität	<p>Minimierung der maximalen skalierten Prädiktionsvarianz $v(\mathbf{x}_0)$ im relevanten Parameterraum [29, 36, 60].</p> <p><i>Ziel: Qualität der Vorhersage an der ungünstigsten Stelle sichern.</i></p> $G_{\text{opt}} = \min \left(\max_{\mathbf{x}_0 \in \mathbb{R}^k} v(\mathbf{x}_0) \right) \quad (2.46)$

Fortsetzung auf der nächsten Seite...

Tabelle 2.1 (Fortsetzung): Übersicht der Optimalitätskriterien

Kriterium	Zielgröße, Beschreibung und Definition
I-Optimalität	<p>Minimierung der durchschnittlichen (<i>integrierten</i>) Prädiktionsvarianz über den Parameterraum - auch als U-Optimalität im ingenieurwissenschaftlich-statistischen Sinn bekannt: Optimierung bzgl. einer <i>uniformen</i> Prädiktionsvarianz [7, 62, 63].</p> <p>Ziel: <i>Optimale Vorhersagegüte (SPV) im Mittel über den gesamten Raum</i> - z.B. bei initialer Parametrisierung eines multivariaten Lebensdauermodells.</p> $I_{\text{opt}} = \min \left(\frac{N}{\int_{\mathbb{R}^k} d\mathbf{x}} \int_{\mathbb{R}^k} v(\mathbf{x}_0) d\mathbf{x} \right) \quad (2.47)$
V-Optimalität	<p>Minimierung der durchschnittlichen Prädiktionsvarianz über ein diskretes Set von N Punkten [29, 34, 54].</p> <p>Ziel: <i>Optimale Vorhersagegüte (SPV) an spezifischen i Stellen</i> - z.B. bei L-DoE-Versuchspunkten oder spezifischen Nennlasten.</p> $V_{\text{opt}} = \min \left(\frac{1}{N} \sum_{i=1}^N v(x_i) \right) \quad (2.48)$

Zur vergleichbaren Bewertung unterschiedlicher Designs unabhängig von der Skalierung werden normierte **Effizienzen** herangezogen [29]. Aus Tabelle 2.1 lassen sich die **A-Effizienz** A_{eff} , **D-Effizienz** D_{eff} , **G-Effizienz** G_{eff} , **I-Effizienz** I_{eff} und **V-Effizienz** V_{eff} ableiten, welche jeweils die Güte $[0, 1]$ eines Versuchsplans quantifizieren - vergleiche Tabelle 2.2.

Tabelle 2.2: Übersicht, Zielsetzung und mathematische Definition verschiedener Effizienzkriterien für Versuchspläne

Kriterium	Zielgröße, Beschreibung und Definition
A-Effizienz	Maß für die durchschnittliche Präzision der Regressionskoeffizienten. [34, 36]. $A_{\text{eff}} = 100 \cdot \frac{p}{\text{spur}(\mathbf{N} \cdot \mathbf{M}^{-1})} \quad (2.49)$
D-Effizienz	Maß für den Informationsgehalt der Informationsmatrix \mathbf{M} , definiert über die Determinante, umgekehrt proportional zum Volumen des Vertrauensellipsoids der Parameterschätzwerte und durch die Potenzierung mit $1/p$ pro Schätzparameter normiert. [34, 36]. $D_{\text{eff}} = 100 \cdot \frac{\det(\mathbf{M})^{1/p}}{N} \quad (2.50)$
G-Effizienz	Maß für die Vorhersagegüte im ungünstigsten Fall innerhalb des Versuchsraums. Da für die maximale SPV die theoretische Untergrenze $v(\mathbf{x}_0)_{\max} \geq p$ gilt, beschreibt dieses Kriterium das Verhältnis zwischen idealer und realisierter maximaler Varianz [36]. $G_{\text{eff}} = 100 \cdot \frac{p}{v(\mathbf{x}_0)_{\max}} \quad (2.51)$
I-Effizienz	Maß für die durchschnittliche Vorhersagegüte über den gesamten Versuchsraum [36]. $I_{\text{eff}} = 100 \cdot \frac{p}{\frac{1}{\int_R d\mathbf{x}} \int_R v(\mathbf{x}_0) d\mathbf{x}} \quad (2.52)$

Fortsetzung auf der nächsten Seite...

Tabelle 2.2 (Fortsetzung): Übersicht der Effizienzkriterien

Kriterium	Zielgröße, Beschreibung und Definition
V-Effizienz	Maß für die durchschnittliche Vorhersagegüte über ein diskretes Set von N Punkten analog zu Gleichung 2.52 [36].
$V_{\text{eff}} = 100 \cdot \frac{1}{N} \sum_{i=1}^N \frac{p}{v(\mathbf{x}_0)(x_i)} \quad (2.53)$	

Signifikanz und Trennschärfe

Zuletzt muss sichergestellt werden, dass der Versuchsplan über eine ausreichende **Trennschärfe** (engl. *power*) verfügt. Die Trennschärfe beschreibt als Metrik diejenige Wahrscheinlichkeit, mit der ein existierender Einfluss durch Faktoren vor dem Hintergrund des experimentellen Rauschens (vergleiche **t-Statistik**, **Signal-to-Noise Ratio**, **Signifikanztests** nach Kleppmann [28], Montgomery [29] und Goos und Jones [34]) korrekt als **signifikanter** Effekt E identifiziert wird [28, 33]. Insofern ist ein Effekt in praktischem Umfeld natürlich problemabhängig (Streuungscharakteristik der Lebensdauer, realisierbare Stichprobengröße, etc.) und meist von Bedeutung für $E \geq 2\sigma$, wobei Trennschärfewerte $\geq 80\%$ als befriedigend gelten [7, 29]. Insbesondere bei reduzierten Versuchsplänen und unter dem Einfluss stochastischer Lebensdauerstreuung ist die Power-Analyse jedoch essenziell, um das Risiko von falsch-negativen Schlussfolgerungen zu minimieren. Letztendlich kann die Trennschärfe somit als Garantimetrik verstanden werden, die sicherstellt, dass ein geplanter Versuchsaufbau in der Lage ist, Effekte von praktischer Relevanz zu detektieren [36, 64] - was vornehmlich in Bezug auf Lebensdauertests aufgrund genannter wirtschaftlicher Randbedingungen von hoher Bedeutung ist. Formalisiert wird dies im Rahmen **statistischer Hypothesentests** [28, 29]. Die Zielsetzung besteht darin, den Nachweis zu erbringen, ob ein bestimmter Faktor x_j einen signifikanten Einfluss auf die Systemantwort y ausübt. Dazu werden zwei konkurrierende Hypothesen formuliert: Die **Nullhypothese** H_0 unterstellt, dass kein Zusammenhang besteht, der

korrespondierende Regressionskoeffizient β_j (vgl. Abschnitt 2.2.3) also den Wert Null annimmt:

$$H_0 : \beta_j = 0 \quad \text{gegen} \quad H_1 : \beta_j \neq 0. \quad (2.54)$$

Die **Alternativhypothese** H_1 postuliert hingegen einen signifikanten Effekt ($\beta_j \neq 0$). Die Entscheidung über die Annahme oder Ablehnung von H_0 basiert auf dem **p-Wert**. Dieser quantifiziert die Wahrscheinlichkeit, unter der Annahme der Gültigkeit von H_0 die beobachteten Daten (oder extremere Ergebnisse) zu erhalten. Unterschreitet der **p-Wert** das a priori definierte **Signifikanzniveau** α (üblicherweise $\alpha = 0,05$), wird H_0 zugunsten der Alternativhypothese H_1 verworfen. Der Fehler, H_0 fälschlicherweise abzulehnen, obwohl kein Effekt vorliegt, wird als Fehler 1. Art bezeichnet; seine Wahrscheinlichkeit ist durch α begrenzt. Die **Trennschärfe** fokussiert hingegen auf den Fehler 2. Art (oft mit β_{error} notiert), welcher das Risiko beschreibt, einen tatsächlich vorhandenen Effekt nicht zu erkennen (falsch-negative Entscheidung). Die Power folglich ist definiert als das Komplement dieses Fehlers ($1 - \beta_{\text{error}}$) und entspricht somit der Wahrscheinlichkeit, H_0 korrekterweise abzulehnen, wenn H_1 wahr ist:

$$\text{power} = \Pr(H_0 \text{ ablehnen} \mid H_1 \text{ ist wahr}) = 1 - \beta_{\text{error}}. \quad (2.55)$$

Im Kontext der Lebensdauerdatenanalyse, bei der nicht-normalverteilte und zensierte Daten vorliegen, lassen sich diese Wahrscheinlichkeiten nicht über klassische **t-Tests** berechnen. Stattdessen kommen Verfahren für **Generalisierte Lineare Modelle**, engl. **Generalized Linear Model (GLM)** zum Einsatz, bei denen die Signifikanz der Koeffizienten β_j vorzugsweise über **Likelihood-Ratio Tests (LR-Tests)** (oder approximativ über **Wald-Tests** - vergleiche hierzu Arbeiten von Meeker et al. [5], Rigdon et al. [7] und Montgomery [29]) ermittelt wird. Da für diese komplexen Verteilungsmodelle bei kleinen Stichprobenumfängen n keine geschlossenen analytischen Lösungen zur Berechnung der Trennschärfe existieren, erfolgt die Untersuchung gerne numerisch mittels **Monte-Carlo-Simulation** [40, 65, 66]. Hierbei wird der geplante Versuchsplan virtuell vielfach durchlaufen (z. B. $n_{\text{MC}} = 10^4$ Simulationsläufe). Für jeden Lauf werden basierend auf einem angenommenen Modell Ausfallzeiten generiert, verrauscht bzw. mit Streuung überlagert und/oder zensiert. Anschließend erfolgt die Modellbildung und Hypothesenprüfung. Der Anteil der

Simulationsläufe, in denen der definierte Effekt korrekt als signifikant ($p\text{-Wert} < \alpha$) erkannt wird, entspricht der geschätzten $power$ des Designs.

2.2.3 Statistische Modellbildung

Die statistische Modellbildung transformiert die durch den Versuchsplan generierte Datenbasis in einen funktionalen, empirisch basierten Zusammenhang zwischen den Einflussfaktoren x_i und der Lebensdauerantwort. Aufgrund der in Abschnitt 2.1 dargelegten Eigenschaften von Lebensdauerdaten (Nicht-Normalverteilung, Zensierung) sind klassische Regressionsverfahren (OLS) hier nicht zulässig. Stattdessen kommt i.d.R. das Framework der GLM zur Anwendung, welches die lineare Prädiktion mit der zugrundeliegenden Verteilung verknüpft [5, 6, 7]. Hinsichtlich der Modellauswahl können in klassischen Ansätzen der Versuchsplanung dabei oft auf Potenztransformationen der Antwortvariablen in Form $y^{(\lambda_{BC})}$ oder $\ln(y)$ (für $\lambda_{BC} = 0$ mit geometrischem Mittelwert der Beobachtung $\bar{y} = \exp\left[\frac{1}{n} \sum_{i=1}^n \ln(y_i)\right]$) zurückgegriffen werden - wie die **Box-Cox-Transformation** nach Montgomery et al. [13] und Box und Cox [67]. So wird Varianzhomogenität und Normalverteilung approximiert zu:

$$y_i^{(\lambda_{BC})} = \begin{cases} \frac{y_i^{\lambda_{BC}} - 1}{\lambda_{BC} \bar{y}^{\lambda_{BC} - 1}}, & \text{für } \lambda_{BC} \neq 0, \\ \bar{y} \ln(y_i), & \text{für } \lambda_{BC} = 0. \end{cases} \quad (2.56)$$

Im Kontext der expliziten Lebensdaueranalyse ist dieser generische Ansatz jedoch weniger gebräuchlich, da die physikalisch motivierte Verteilungsannahme (z.B. Weibull) bereits eine inhärente Transformation der Lebensdauerzeit impliziert - vgl. Abschnitt 2.1.3, Gleichung 2.11 sowie Yang [6] und Wu und Hamada [30]. Wird also für die Lebensdauerverteilung die Weibull-Verteilung (Abschnitt 2.1.3) angenommen, so ergibt sich für $j = 0, \dots, k$ betrachtete Parameter aus beispielsweise Gleichung 2.30 und n Beobachtungen analog zu Gleichung 2.11 der Erwartungswert der charakteristischen Lebensdauer als Funktion der Einflussfaktoren x_i und des unbekannten Koeffizientenvektors $\hat{\beta}$ [6, 30, 68]:

$$E_{\theta_x}(\ln(T)) = y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \dots + \epsilon. \quad (2.57)$$

In der Lebensdauer-Datenanalyse werden insbesondere zwei weitere Modellierungsansätze verwendet: neben den GLM-Ansätzen insbesondere der **Proportional Hazard (PH)-Modell**-Ansatz. Jedoch werden PH-Modelle vornehmlich in der Biostatistik verwendet und spielen hier ihre Stärke bei der Abbildung zeitabhängiger Kovariate aus [69, 70]:

$$f(t, \mathbf{X}) = \lambda(t)(t, \mathbf{X}) \cdot R(t, \mathbf{X}) = b \cdot t^{b-1} \cdot e^{[y-t^b e^y]}. \quad (2.58)$$

Weibull-GLM

Im ingenieurwissenschaftlichen Kontext - maßgeblich gestützt durch Ausführungen in Meeker et al. [5], Yang [6], Wu und Hamada [30], Kremer et al. [70] und Dobson und Barnett [71] - finden hingegen **Log-Location-Scale-Modelle**, also GLMs, überwiegend Anwendung. Obwohl für die Weibull-Verteilung eine mathematische Äquivalenz zwischen beiden Ansätzen besteht, sprechen drei fundamentale Gründe für deren Verwendung. **Physikalischer Bezug:** Gängige physikalische Beschleunigungsmodelle (Arrhenius, Inverses Potenzgesetz) beschreiben die Änderung der Zeit bis zum Ausfall ($\ln(\hat{T})$) und nicht der Ausfallrate. Die Log-Linearisierung dieser Gesetze führt direkt auf die Regressionsgleichung der Location-Scale-Modelle. **Universalität der Auswertung:** Das Modell bietet einen einheitlichen mathematischen Rahmen, der nicht nur für die Weibull-Verteilung, sondern identisch auch für die Lognormal-Verteilung und Normalverteilung gültig ist. Das PH-Modell hingegen ist für Lognormal-Verteilungen mathematisch nicht geschlossen anwendbar. **Fokus auf Lebensdauer-Quantile:** Im Engineering und DfR liegt das Interesse primär auf der Prädiktion von Ausfallzeiten für geringe Ausfallanteile (z.,B. t_{10}). Diese Größen sind direkte Ergebnisse der Location-Scale-Gleichung, während sie im PH-Modell nur indirekt über die Invertierung der Hazard-Funktion zugänglich sind. Folglich wird eine zufallsverteilte Lebensdauer τ für ein Weibull-GLM mit der Dichtefunktion

$$f(t, \mathbf{X}) = b \cdot t^{b-1} e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_j)} \cdot e^{-t^b e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_j)}}, \quad t > 0. \quad (2.59)$$

beschrieben und analog zu Gleichungen 2.17–2.20 in die **Generalized Log-Linear (GLL)-Funktion** überführt, [6, 60, 72]:

$$\Lambda = \ln(\mathcal{L}(t_i; b, \beta_0, \dots, \beta_k)) \\ \propto \sum_{i=1}^n \left[\ln(b) - b \left(\beta_0 + \sum_{j=1}^k \beta_j x_{ij} \right) + (b-1) \ln(t_i) - t_i^b e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})} \right]. \quad (2.60)$$

Werden außerdem (rechts-) zensierte Daten berücksichtigt, so ergibt sich analog zu Gleichung 2.20 über Gleichung 2.60 die erweiterte GLL-Funktion [5, 6]:

$$\Lambda_z \propto \Lambda - (1 - \delta_i) \cdot t_i^b e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})}. \quad (2.61)$$

Sollen über die rein empirische Datenauswertung hinaus bekannte physikalische Gesetzmäßigkeiten im Modell berücksichtigt werden, ist die lineare Struktur des Prädiktors entsprechend anzupassen. Dies erfolgt durch eine Transformation der physikalischen Beanspruchungsgrößen x in den Modellraum. Anstelle der direkten Verwendung des Einflussfaktors in der GLL-Beziehung (Gleichung 2.59) wird eine transformierte Größe x_V eingesetzt, welche die physikalische Natur des Schädigungsmechanismus abbildet. Gängige Transformationen umfassen hierbei die reziproke Anpassung für thermische Lasten oder logarithmische Ansätze für elektrische Potenzmodelle, vgl. Tabelle 2.3. Eine Übersicht verschiedener Modellierungs-

Tabelle 2.3: Übersicht gängiger Transformationen der Belastungsgrößen zur Abbildung physikalischer Lebensdauermodelle

Art	Transformation	Lebensdauermodell	Anwendungsfall
Keine	$x_V = x$	Exponential / Wöhler	mechanisch
Reziprok	$x_V = 1/x$	Arrhenius	thermisch
Logarithmisch	$x_V = e^x$	Potenzmodell	elektrisch

ansätze (z.B. Arrhenius, Inverses Potenzgesetz, etc.) und deren Einbindung in die GLM-Struktur findet sich in den Arbeiten von Meeker et al. [5], Rigdon et al. [7] und Modarres et al. [69]. Die Schätzung der unbekannten Parameter $\hat{\beta}$ und \hat{b} erfolgt anschließend mittels MLE durch Maximierung der Log-Likelihood-Funktion (Gleichung 2.61) mittels beispielsweise **Newton-Raphson-Verfahren** bzw. **Pattern-**

Search-Algorithmen oder **Gradientenverfahren**, [6, 17]. Geeignete numerische Optimierungungsverfahren wurden in Abschnitt 2.1.4 erwähnt und unter anderem durch Vorarbeiten von Kremer und Bertsche [18] evaluiert.

Signifikanzanalyse

Liegen die Parameterschätzungen $\hat{\theta}_i$ vor, erfolgt die Bewertung ihrer statistischen Relevanz mittels **LR-Test**. Dieses Verfahren stellt das methodische Äquivalent zur **ANOVA** bei normalverteilten Daten dar und prüft, ob die Aufnahme eines Parameters die Modellanpassung signifikant verbessert. Der Test basiert auf dem Vergleich der maximalen Likelihoods des vollständigen Modells $\mathcal{L}(\hat{\theta}_i)$ gegenüber einem reduzierten Modell $\mathcal{L}(\hat{\theta}_{-i})$, aus dem der zu untersuchende Parameter eliminiert wurde [22, 40]. Die Teststatistik **LR** quantifiziert den Informationsverlust durch die Modellreduktion und ist definiert als **Devianz-Funktion** [5]:

$$LR = -2 \ln \left(\frac{\mathcal{L}(\hat{\theta}_{-i})}{\mathcal{L}(\hat{\theta}_i)} \right) = 2 \cdot [\ln(\mathcal{L}(\hat{\theta}_i)) - \ln(\mathcal{L}(\hat{\theta}_{-i}))]. \quad (2.62)$$

Unter der Nullhypothese H_0 , dass der ausgeschlossene Parameter keinen Einfluss hat, folgt **LR** asymptotisch einer χ^2 -Verteilung mit d_f Freiheitsgraden (Anzahl der ausgeschlossenen Parameter, hier meist $d_f = 1$). Die Signifikanzentscheidung erfolgt über den **p-Wert**, der die Wahrscheinlichkeit angibt, einen Testwert $\geq LR$ zu beobachten:

$$p\text{-Wert} = \Pr \left(\chi^2(d_f) \geq LR \right). \quad (2.63)$$

Ist dieser **p-Wert** kleiner als das definierte Signifikanzniveau $p\text{-Wert} < \alpha$, wird der Parameter als signifikant in das Modell aufgenommen.

Erwartungswerte und Konfidenzintervalle

Nach der Ermittlung der optimalen Modellparameter $\hat{\theta}$ ist die Quantifizierung der mit diesen Schätzungen verbundenen Unsicherheit essenziell. Analytisch erfolgt dies über die Betrachtung der Krümmung der Log-Likelihood-Funktion Λ im Maximum. Je stärker die Krümmung, desto präziser ist die Schätzung (geringere Varianz). Mathematisch erfolgt dies durch die drei Schritte: Wahl der statistischen Zielgröße (R, F), Berechnung des Standard-Fehlers (via **F** analog zu Gleichung 2.21)

und Vertrauensbereich-Schätzung / Wald-Statistik (Gleichung 2.23 bzw. 2.24). So wird dies analog zu Gleichungen 2.21-2.22 durch die Hesse-Matrix \mathbf{H}_f beschrieben, welche die zweiten partiellen Ableitungen der Log-Likelihood-Funktion nach den Parametern θ_i (also β_j und b) enthält. Der negative Erwartungswert dieser Matrix führt damit zur Erwartung der Fisher-Informationsmatrix \mathbf{F} [25, 73]:

$$\mathbf{F}_O := -\mathbf{H}_f(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} -\frac{\partial^2 \mathcal{L}}{\partial \theta_1^2} & -\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_2} & \cdots & -\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_k} \\ -\frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_1} & -\frac{\partial^2 \mathcal{L}}{\partial \theta_2^2} & \cdots & -\frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial^2 \mathcal{L}}{\partial \theta_k \partial \theta_1} & -\frac{\partial^2 \mathcal{L}}{\partial \theta_k \partial \theta_2} & \cdots & -\frac{\partial^2 \mathcal{L}}{\partial \theta_k^2} \end{bmatrix}. \quad (2.64)$$

Die asymptotische Varianz-Kovarianz-Matrix $\hat{\mathbf{V}}$ der Parameterschätzer wird anschließend durch die Inversion der Fisher-Informationsmatrix für $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$ approximiert [6, 73]:

$$\hat{\mathbf{V}} \approx \mathbf{F}_O^{-1} = \begin{bmatrix} \hat{\text{Var}}(\hat{\theta}_1) & \hat{\text{Cov}}(\hat{\theta}_1, \hat{\theta}_2) & \cdots & \hat{\text{Cov}}(\hat{\theta}_1, \hat{\theta}_k) \\ \hat{\text{Cov}}(\hat{\theta}_2, \hat{\theta}_1) & \hat{\text{Var}}(\hat{\theta}_2) & \cdots & \hat{\text{Cov}}(\hat{\theta}_2, \hat{\theta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\text{Cov}}(\hat{\theta}_k, \hat{\theta}_1) & \hat{\text{Cov}}(\hat{\theta}_k, \hat{\theta}_2) & \cdots & \hat{\text{Var}}(\hat{\theta}_k) \end{bmatrix}. \quad (2.65)$$

Die Hauptdiagonalelemente dieser Matrix liefern die geschätzten Varianzen $\hat{\text{Var}}(\hat{\theta}_j)$ der einzelnen Modellparameter. Unter der Annahme der asymptotischen Normalverteilung der MLE beschreibt Yang [6] neben weiteren Alternativen aus Meeker et al. [5] und Kremer und Bertsche [22] die symmetrischen $(1 - \alpha)$ -Konfidenzintervalle (Wald-Intervalle) wie nachfolgend - vgl. Gleichung 2.24:

$$[\theta_{j,u}, \theta_{j,o}] = \hat{\theta}_j \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\text{Var}}(\hat{\theta}_j)} \quad \text{bzw.} \cdots = \hat{\theta}_j \exp \left(\pm z_{1-\alpha/2} \cdot \frac{\sqrt{\hat{\mathbf{V}}_{jj}}}{\hat{\theta}_j} \right). \quad (2.66)$$

Vertrauensbereiche für Funktionswerte und Weibull-GLM

Oft ist nicht nur die Unsicherheit der Modellparameter selbst, sondern die der daraus abgeleiteten Funktionen - wie der Zuverlässigkeit $R(t)$ oder der Ausfallwahrscheinlichkeit $F(t)$ - von Interesse. Zur analytischen Ermittlung dieser sogenannten

Fisher-Vertrauensbereiche führt Yang [6] zunächst eine Hilfsgröße w ein - vergleiche auch z_e nach Meeker et al. [5]. Diese normiert die logarithmische Abweichung der betrachteten Zeit t von der geschätzten charakteristischen Lebensdauer \hat{T} mit dem Formparameter \hat{b} (vergleiche auch Bain und Engelhardt [24] für weitere Abschätzungen):

$$\hat{w} = \hat{b} \cdot \ln \left(\frac{t}{\hat{T}} \right). \quad (2.67)$$

Die Varianz dieser Größe lässt sich mittels Taylor-Reihenentwicklung (Delta-Methode, vgl. Abschnitt 2.1.4) approximativ aus den Varianzen und Kovarianzen der Parameterschätzer und analog zu Gleichungen 2.25 und 2.26 bestimmen:

$$\hat{\text{Var}}(\hat{w}) = \left(\frac{\hat{b}}{\hat{T}} \right)^2 \hat{\text{Var}}(\hat{T}) + \left(\frac{\hat{w}}{\hat{b}} \right)^2 \hat{\text{Var}}(\hat{b}) - \frac{2\hat{w}}{\hat{T}} \hat{\text{Cov}}(\hat{T}, \hat{b}). \quad (2.68)$$

Basierend auf der asymptotischen Normalverteilung ergeben sich die Vertrauensgrenzen $[\hat{w}_u, \hat{w}_o]$ gemäß:

$$[\hat{w}_u, \hat{w}_o] = \hat{w} \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\text{Var}}(\hat{w})}. \quad (2.69)$$

Die Rücktransformation dieser Grenzen in den Wertebereich der Ausfallwahrscheinlichkeit erfolgt anschließend über die Verteilungsfunktion der kleinsten Extremwerte $G = 1 - \exp[-\exp(\hat{w})]$, woraus sich direkt die Konfidenzintervalle der Zuverlässigkeit ableiten lassen. Im Fall von GLM ist zu berücksichtigen, dass die geschätzte charakteristische Lebensdauer \hat{T} keine unabhängige Konstante ist, sondern funktional entsprechend Gleichung 2.57 von den Kovariaten x und dem Parametervektor $\hat{\beta}$ abhängt. Die für Gleichung 2.68 benötigte Varianz $\hat{\text{Var}}(\hat{T})$ muss folglich unter Berücksichtigung der gesamten Kovarianzmatrix der Regressionskoeffizienten aus Gleichung 2.65 berechnet werden. Dies erfolgt durch die allgemeine Fehlerfortpflanzung nach Gauß:

$$\hat{\text{Var}}(\hat{T}) \approx \sum_{i=1}^k \left(\frac{\partial \hat{T}}{\partial \hat{\beta}_i} \right)^2 \hat{\text{Var}}(\hat{\beta}_i) + \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k \left(\frac{\partial \hat{T}}{\partial \hat{\beta}_i} \right) \left(\frac{\partial \hat{T}}{\partial \hat{\beta}_j} \right) \hat{\text{Cov}}(\hat{\beta}_i, \hat{\beta}_j). \quad (2.70)$$

Modellgüte und Residuenanalyse

Die Validität eines parametrischen Lebensdauermodells (hier Weibull-GLM) beruht auf drei fundamentalen Annahmen, die es nach der Parameterschätzung zu verifizieren gilt: Erstens folgen die Lebensdauern der angenommenen Verteilung. Zweitens sind die Lebensdauern statistisch unabhängig. Drittens besteht ein log-linearer Zusammenhang zwischen dem Skalenparameter und den Einflussgrößen, wobei der Formparameter in einfachen Modellen als konstant über den gesamten Versuchsraum angenommen wird. Die Beurteilung dieser Modellgüte erfordert bei Lebensdauerdaten, bedingt durch Nicht-Normalverteilung und Zensierung, eine Anpassung klassischer Residualkonzepte ($y - \hat{y}$) [28, 33]. Stattdessen werden transformierte Residuen betrachtet, die den Bezug zur Verteilung der kleinsten Extremwerte *SEV* herstellen, welcher der logarithmierten Weibull-Verteilung zugrunde liegt. Analog zur *z*-Transformation bei der Normalverteilung lassen sich für jede Beobachtung t_i standardisierte Residuen r_i berechnen. Dabei ersetzt der logarithmierte Schätzwert der charakteristischen Lebensdauer $\ln(\hat{T}_i)$ (vgl. Gleichung 2.57) den Lageparameter μ und der Kehrwert des Formparameters $1/\hat{b}$ die Skalierung σ . Dies entspricht exakt der in diesem Abschnitt eingeführten Hilfsgröße w :

$$r_i = \frac{\ln(t_i) - \ln(\hat{T}_i)}{1/\hat{b}} = \hat{b} \cdot \ln\left(\frac{t_i}{\hat{T}_i}\right). \quad (2.71)$$

Ist das Modell korrekt spezifiziert, müssen diese Residuen (unter Berücksichtigung der Zensierung) einer Verteilung der kleinsten Extremwerte *SEV* ($\mu = 0, \sigma = 1$) folgen. Eine allgemeinere Form der Validierung bieten die **Cox-Snell-Residuen** r_{CS} [5, 7]. Sie stellen eine Transformation der standardisierten Residuen dar und entsprechen dem Wert der kumulativen Hazard-Funktion an der Stelle der Beobachtung. Für das Weibull-Modell gilt der Zusammenhang:

$$r_{CSi} = -\ln(\hat{R}(t_i)) = \exp(r_i) = \left(\frac{t_i}{\hat{T}_i}\right)^{\hat{b}}. \quad (2.72)$$

Folgt das angepasste Modell den Daten, so verhalten sich die Residuen r_{CS} wie eine Stichprobe aus einer Standard-Exponentialverteilung mit Mittelwert 1. Diese Eigenschaft erlaubt eine grafische Validierung mittels Probability-Plots: Liegen die geplotteten Residuen auf einer Geraden im Exponential-Wahrscheinlichkeitsnetz, kann

von einer adäquaten Modellpassung ausgegangen werden. Zur Identifikation von Ausreißern oder einflussreichen Beobachtungen eignen sich hingegen standardisierte Devianz-Residuen r_{Dev} . Im Gegensatz zu den stets positiven Cox-Snell-Residuen sind diese symmetrischer um Null verteilt und folgen asymptotisch einer Standardnormalverteilung:

$$r_{\text{Dev}i} = \text{sgn}(r_{\text{CS}i} - 1) \cdot \sqrt{2 [r_{\text{CS}i} - \delta_i - \delta_i \ln(r_{\text{CS}i})]}. \quad (2.73)$$

Neben der Verteilungsprüfung dienen Streudiagramme dieser Residuen zur Aufdeckung struktureller Modelldefizite. Dabei werden r_{CS} oder r_{Dev} gegen die prädierten Werte oder gegen einzelne Einflussfaktoren aufgetragen und gegenüber Trendfreiheit und Varianzhomogenität bewertet. Besteht initiale Unsicherheit über die zugrundeliegende Physik, ermöglicht zudem der Vergleich den Wahrscheinlichkeitsnetzen der Residuen eine empirische Selektion: Dasjenige Modell, dessen Punktegraphen eine geringere Krümmung aufweisen und besser der theoretischen Geraden folgen, ist statistisch zu bevorzugen.

3 Effiziente Testplanung für die multivariate Lebensdauererprobung

3.1 Bewertung des Standes der Forschung und Technik

- [36, Kap. 9.1.4]
- [36, Kap. 9.2.1]
- [74, Chapter. 27]
- Khuri - 2006 - Design Issues for Generalized Linear Models A Review
- Yang - 2011 - OPTIMAL DESIGNS FOR GENERALIZED LINEAR MODELS WITH MULTIPLE DESIGN VARIABLES
- [7, Seite 314]
- Monroe, E.M., Pan, R., Anderson-Cook, C. et al. (2011). A generalized linear model approach to designing accelerated life test experiments. Quality and Reliability Engineering International 27: 595-607.
- [60, S.69:]: The minimum number is p , the number of parameters. This is intuitively obvious. At least two separate points are needed to estimate the two parameters, β_0 and β_1 , of a straight line, and at least three separate points are required to β_0 , β_1 and β_2 , of a quadratic. See Figure 3.2 to recall the situation that arises when there are only two points for a quadratic. The maximum number of support points necessary for an optimal design can be shown to be $p(p + 1) + 1$, using Caratheodory's Theorem. The proof is beyond the scope of this book, but is given in Rockafellar (1970, p. 155) or Silvey (1980, p.

77). Pukelsheim (1993, p. 190) showed that, if our interest is in estimating all elements of the parameter vector β , then this upper limit can be reduced to $p(p + 1)/2$.

[25, 30, 35, 51, 54, 66, 69, 73, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88]

Besonderheiten der Effizienz in der Lebensdauerprüfung

Der Begriff der Effizienz erfährt im Kontext der Planung multivariater Lebensdauererprobungen (ALT oder Reliability Demonstration Tests) eine signifikante Erweiterung gegenüber der klassischen linearen Versuchsplanung. Während Standard-Designs primär die Varianz der Parameterschätzer minimieren, unterliegen Lebensdauererprobungen der zusätzlichen Restriktion, dass die Information (der Ausfall eines Bauteils) stochastisch über die Zeit generiert wird und oft durch **Zensierung** limitiert ist.

Eine zentrale Herausforderung besteht in der **Modellabhängigkeit** (engl. Model Dependence) der Informationsmatrix. Bei nicht-linearen Modellen, wie der in der Zuverlässigkeitstechnik omnipräsenten Weibull- oder Lognormal-Regression, ist die Fisher-Informationsmatrix \mathbf{F} nicht mehr allein von der Versuchsplanmatrix \mathbf{X} abhängig, sondern auch von den wahren, aber unbekannten Verteilungsparametern $\boldsymbol{\theta}$ (z. B. Formparameter b):

$$\mathbf{M}(\mathbf{X}, \boldsymbol{\theta}) = E \left[-\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right]. \quad (3.1)$$

Daraus resultiert das Paradoxon, dass zur Konstruktion eines optimalen Plans bereits Kenntnisse über die zu ermittelnden Parameter vorliegen müssen. Klassische Optimalitätskriterien (D-, A-Optimalität) wandeln sich daher zu **lokalen Optimalitäten**, die nur für einen spezifischen Parametervektor $\boldsymbol{\theta}_0$ („Best Guess“) gültig sind. Um Robustheit gegenüber Fehlannahmen dieser Startwerte zu gewährleisten, werden in der Entwicklung effizienter Lebensdauererprobungen häufig **Bayes-Optimale Versuchspläne** eingesetzt, welche die Effizienz über eine A-Priori-Verteilung der Parameter maximieren [5, 34].

Zudem muss die **zeitliche Effizienz** berücksichtigt werden. Ein Versuchsplan gilt im Kontext der Lebensdaueranalyse nur dann als effizient, wenn er unter Berücksichtigung der Zensierungsmechanismen (Typ-I oder Typ-II) die erwartete Anzahl an Ausfällen maximiert oder die **erwartete Testdauer** (Expected Test Duration, ETD)

bei gegebener Präzision minimiert. Die Varianz der Schätzung wird hierbei maßgeblich durch die Anzahl der ausgefallenen Einheiten getrieben, nicht allein durch die Stichprobengröße n [16].

Abgrenzung zu weiteren Themen in der **RSM**

- **GLM Mixed Models, Random Effects** (Generalized Linear Mixed Models) [54, Myers 2010 Kap.7.1.2]
- **Robuste Versuchspläne** (Robust Designs):
- **Adaptive Versuchspläne** (Adaptive Designs):
- **Bayesian Designs**, z.B. in Englert et al. [89]
- **Sequential Designs**:
- **Optimalität für Vorhersagegüte**:
- **Optimalität für Klassifikationsaufgaben**:
- **Optimalität für multiple Zielgrößen**:
- **Robuste Optimale Versuchspläne** (Robust Optimal Designs):
- **Kombinierte Optimalitätskriterien** (Compound Optimal Designs):
- **Optimale Versuchspläne für Mischmodelle** (Optimal Designs for Mixed Models):
- **Optimale Versuchspläne für räumliche Modelle** (Optimal Designs for Spatial Models):
- **Optimale Versuchspläne für Zeitreihenmodelle** (Optimal Designs for Time Series Models):
- **Optimale Versuchspläne für nichtparametrische Modelle** (Optimal Designs for Nonparametric Models):
- **Optimale Versuchspläne für Hochdimensionale Modelle** (Optimal Designs for High-Dimensional Models):

- **Optimale Versuchspläne für dynamische Systeme** (Optimal Designs for Dynamic Systems):
- **Optimale Versuchspläne für Netzwerke** (Optimal Designs for Networks):
- **Optimale Versuchspläne für Big Data Anwendungen** (Optimal Designs for Big Data Applications):
- **Optimale Versuchspläne für maschinelles Lernen** (Optimal Designs for Machine Learning): Çolak et al. [90]
- **Optimale Versuchspläne für Künstliche Intelligenz** (Optimal Designs for Artificial Intelligence):
- **Optimale Versuchspläne für Quantencomputing** (Optimal Designs for Quantum Computing):
- **Optimale Versuchspläne für Blockchain-Technologien** (Optimal Designs for Blockchain Technologies):
- **Optimale Versuchspläne für Internet der Dinge (IoT)** (Optimal Designs for Internet of Things (IoT)):
- **Optimale Versuchspläne für Cyber-Physische Systeme** (Optimal Designs for Cyber-Physical Systems):
- **Optimale Versuchspläne für Nachhaltigkeit und Umweltwissenschaften** (Optimal Designs for Sustainability and Environmental Sciences):

3.2 Forschungsfragen und Aufbau der Arbeit

4 Parameter-Screening für multifaktorielle Lebensdauertests

4.1 Identifikation potenzieller Einflussgrößen

4.2 Kreativmethoden zum Auswahlprozesse im Parameter-Screening

4.3 Randbedingungen in der Parameterauswahl für die Zuverlässigkeitsmodellierung

4.4 Vorgehen zum heuristische Screening für die Zuverlässigkeitsmodellierung

4.5 Zusammenfassung

5 Effiziente multivariate Versuchspläne für Lebensdaueruntersuchungen

6 Fallstudie

7 Zusammenfassung und Ausblick

7.1 Key Findings

7.2 Diskussion

7.3 Ausblick

Literatur

- [1] B. Bertsche und M. Dazer. *Zuverlässigkeit im Fahrzeug- und Maschinenbau: Ermittlung von Bauteil- und System-Zuverlässigkeiten*. 4. Auflage. Lehrbuch. Berlin und Heidelberg: Springer Vieweg, 2022. ISBN: 978-3-662-65023-3. URL: <http://www.springer.com/>.
- [2] R. Koselka. „The New Mantra: MVT“. In: *Forbes* (11.03.1996), S. 114–118.
- [3] *DAT Report 2025: Kurzbericht*. Ostfildern, Deutschland, Februar 2025. URL: https://www.dat.de/fileadmin/protected/DAT_Report/2025/DAT-Report-2025-Kurzbericht.pdf.
- [4] A. Birolini. *Reliability Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017. ISBN: 978-3-662-54208-8. DOI: [10.1007/978-3-662-54209-5](https://doi.org/10.1007/978-3-662-54209-5).
- [5] W. Q. Meeker, L. A. Escobar und F. G. Pascual. *Statistical methods for reliability data*. Second edition. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2022. ISBN: 978-1-118-11545-9.
- [6] G. Yang. *Life cycle reliability engineering*. Hoboken, N.J.: Wiley, 2007. ISBN: 9780471715290. URL: <http://www.loc.gov/catdir/enhancements/fy0740/2006019150-b.html>.
- [7] S. E. Rigdon, R. Pan, D. C. Montgomery und L. J. Freeman. *Design of experiments for reliability achievement*. Hoboken, NJ: Wiley, 2022. ISBN: 9781119237693. DOI: [10.1002/9781119237754](https://doi.org/10.1002/9781119237754).
- [8] L. Fahrmeir, C. Heumann, R. Künstler, I. Pigeot und G. Tutz. *Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016. ISBN: 978-3-662-50371-3. DOI: [10.1007/978-3-662-50372-0](https://doi.org/10.1007/978-3-662-50372-0).
- [9] E. L. Kaplan und P. Meier. „Nonparametric Estimation from Incomplete Observations“. In: *Journal of the American Statistical Association* 53.282 (1958), S. 457. ISSN: 01621459. DOI: [10.2307/2281868](https://doi.org/10.2307/2281868).

- [10] J. Hedderich und L. Sachs. *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2020. ISBN: 978-3-662-62293-3. DOI: [10.1007/978-3-662-62294-0](https://doi.org/10.1007/978-3-662-62294-0).
- [11] H. Rinne. *The Weibull Distribution: A Handbook*. 1st ed. London: CRC Press LLC, 2008. ISBN: 978-1-4200-8743-7. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=1633687>.
- [12] D. Kececioglu. *Reliability & life testing handbook*. Lancaster, PA: DEStech Publications, 2002. ISBN: 1-932078-03-7.
- [13] D. C. Montgomery, E. A. Peck und G. G. Vining. *Introduction to linear regression analysis*. Sixth edition. Wiley series in probability and statistics. Hoboken, New Jersey: Wiley, 2021. ISBN: 978-1-119-57872-7.
- [14] W. Nelson. *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. 1st ed. Bd. v.344. Wiley Series in Probability and Statistics Series. Newark: John Wiley & Sons Incorporated, 1990. ISBN: 978-0471697367. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=469757>.
- [15] J. D. Kalbfleisch. *The statistical analysis of failure time data*. 2nd ed. Wiley series in probability and statistics. Hoboken, N.J.: J. Wiley, 2002. ISBN: 047136357X. DOI: [10.1002/9781118032985](https://doi.org/10.1002/9781118032985).
- [16] W. Nelson. *Applied Life Data Analysis*. Bd. v.577. Wiley series in probability and statistics. Hoboken: John Wiley & Sons Inc, 2005. ISBN: 9780471644620. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=226561>.
- [17] H. Qiao und C. P. Tsokos. „Parameter estimation of the Weibull probability distribution“. In: *Mathematics and Computers in Simulation* 37.1 (1994), S. 47–55. ISSN: 0378-4754. DOI: [10.1016/0378-4754\(94\)90058-2](https://doi.org/10.1016/0378-4754(94)90058-2).
- [18] A. Kremer und B. Bertsche. „A New Approach for Parametrizing Multidimensional Lifetime Models“. In: *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*. Hrsg. von M. Beer und E. Zio. Europe und Singapore: European Safety and Reliability Association und Research Publishing Services, 2019, S. 715–722. ISBN: 978-981-11-2724-3. DOI: [10.3850/978-981-11-2724-3\textunderscore0443-cd](https://doi.org/10.3850/978-981-11-2724-3\textunderscore0443-cd).

-
- [19] R. B. Abernethy. *The new Weibull handbook: Reliability & statistical analysis for predicting life, safety, risk, support costs, failures, and forecasting warranty claims, substantiation and accelerated testing, using Weibull, log normal, Crow-AMSAA, Probit, and Kaplan- meier Models*. 5. ed. North Palm Beach, Fla.: R.B. Abernethy, 2006. ISBN: 0965306232.
- [20] H. Hirose. „Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution“. In: *IEEE Transactions on Dielectrics and Electrical Insulation* 6.1 (1999), S. 66–68. ISSN: 10709878. DOI: [10.1109/94.752011](https://doi.org/10.1109/94.752011).
- [21] R. Ross. „Bias and standard deviation due to Weibull parameter estimation for small data sets“. In: *IEEE Transactions on Dielectrics and Electrical Insulation* 3.1 (1996), S. 28–42. ISSN: 10709878. DOI: [10.1109/94.485512](https://doi.org/10.1109/94.485512).
- [22] A. Kremer und B. Bertsche. „A Methodology for Consideration of Uncertainty in Lifetime Design of Experiments“. In: *2019 Annual Reliability and Maintainability Symposium (RAMS)*. IEEE, 2019, S. 1–6. ISBN: 978-1-5386-6554-1. DOI: [10.1109/RAMS.2019.8768906](https://doi.org/10.1109/RAMS.2019.8768906).
- [23] J. F. Lawless. *Statistical models and methods for lifetime data*. 2. ed. Wiley series in probability and statistics. Hoboken, NJ: Wiley-Interscience, 2003. ISBN: 0471372153. URL: <http://www.loc.gov/catdir/bios/wiley044/2002151805.html>.
- [24] L. J. Bain und M. Engelhardt. *Statistical Analysis of Reliability and Life-Testing Models*. Routledge, 2017. ISBN: 9780203738733. DOI: [10.1201/9780203738733](https://doi.org/10.1201/9780203738733).
- [25] R. A. Fisher. *The Design of Experiments*. Edinburgh: Oliver & Boyd, 1935.
- [26] G. E. P. Box, W. G. Hunter und J. S. Hunter. *Statistics for experimenters: An introduction to design, data analysis, and model building*. Wiley series in probability and mathematical statistics. New York: Wiley, 1978. ISBN: 0471093157. URL: <http://www.loc.gov/catdir/description/wiley033/77015087.html>.
- [27] G. Taguchi. *Taguchi's quality engineering handbook*. Hoboken, NJ: Wiley, 2005. ISBN: 9780471413349. DOI: [10.1002/9780470258354](https://doi.org/10.1002/9780470258354).
- [28] W. Kleppmann. *Versuchsplanung: Produkte und Prozesse optimieren*. 10., überarbeitete Auflage. Praxisreihe Qualität. München: Hanser, 2020. ISBN: 978-3-446-46397-4.

- [29] D. C. Montgomery. *Design and Analysis of Experiments*. 10th. Hoboken, NJ, USA: John Wiley & Sons, Inc, 2020. ISBN: 9781119722106.
- [30] C.-F. Wu und M. Hamada. *Experiments: Planning, analysis and optimization*. Third edition. Wiley series in probability and statistics. Hoboken, NJ, USA: Wiley, 2021. ISBN: 9781119470106. URL: <https://zbmath.org/?q=an%3A1460.62003>.
- [31] W. Q. Meeker und L. A. Escobar. „A Review of Recent Research and Current Issues in Accelerated Testing“. In: *International Statistical Review / Revue Internationale de Statistique* 61.1 (1993), S. 147. ISSN: 03067734. DOI: [10.2307/1403600](https://doi.org/10.2307/1403600).
- [32] E. A. Elsayed und H. Zhang. „Design of PH-based accelerated life testing plans under multiple-stress-type“. In: *Reliability Engineering & System Safety* 92.3 (2007), S. 286–292. ISSN: 09518320. DOI: [10.1016/j.ress.2006.04.016](https://doi.org/10.1016/j.ress.2006.04.016).
- [33] K. Siebertz, D. van Bebber und T. Hochkirchen. *Statistische Versuchsplanung: Design of Experiments (DoE)*. 2. Auflage. VDI-Buch. Berlin und Heidelberg: Springer Vieweg, 2017. ISBN: 978-3-662-55742-6.
- [34] P. Goos und B. A. Jones. *Optimal Design of Experiments: A Case Study Approach*. 1st. Hoboken, NJ: John Wiley & Sons, Inc, 2011. ISBN: 9780470744611.
- [35] K. Hinkelmann. *Design and Analysis of Experiments - Volume 3: Special Designs and Applications*. 1st. John Wiley & Sons, Inc, 2012. ISBN: 978-0470-53068-9.
- [36] R. H. Myers, D. C. Montgomery und C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 4. Aufl. Hoboken, NJ, USA: John Wiley & Sons, Inc, 2016. ISBN: 978-1-118-91601-8.
- [37] M. Dazer. *Zuverlässigkeitstestplanung mit Berücksichtigung von Vorwissen aus stochastischen Lebensdauerberechnungen*. 2019. DOI: [10.18419/OPUS-10501](https://doi.org/10.18419/OPUS-10501).
- [38] T. Herzig. *Anforderungsgerechte Produktauslegung durch Planung effizienter beschleunigter Zuverlässigkeitstests*. 2021. DOI: [10.18419/OPUS-12001](https://doi.org/10.18419/OPUS-12001).
- [39] A. Grundler. *Zuverlässigkeitsnachweis von Systemen durch Komponenten- und Systemtests unter Berücksichtigung von Vorkenntnissen*. 2024. DOI: [10.18419/OPUS-14393](https://doi.org/10.18419/OPUS-14393).

-
- [40] A. Kremer. „Statistische Versuchsplanung in der Lebensdauererprobung mit Vertrauensintervallen“. Diss. 2021. DOI: [10.18419/OPUS-11728](https://doi.org/10.18419/OPUS-11728).
- [41] D. E. Coleman und D. C. Montgomery. „A Systematic Approach to Planning for a Designed Industrial Experiment“. In: *Technometrics* 35.1 (1993), S. 1. ISSN: 0040-1706. DOI: [10.2307/1269280](https://doi.org/10.2307/1269280).
- [42] C. Gundlach. „Entwicklung eines ganzheitlichen Vorgehensmodells zur problemorientierten Anwendung des statistischen Versuchsplanung“. Zugl.: Kassel, Univ., Diss, 2004. Kassel. URL: <http://d-nb.info/971843546/34>.
- [43] B. Mayers. *Prozeß- und Produktoptimierung mit Hilfe der statistischen Versuchsmethodik: Zugl.: Aachen, Techn. Hochsch., Diss., 1997. Als Ms. gedr. Bd. 97,9. Berichte aus der Produktionstechnik. Aachen: Shaker, 1997. ISBN: 3826524977.*
- [44] G. Pahl, W. Beitz, J. Feldhusen und K.-H. Grote. *Engineering design: A systematic approach*. 3. ed. London: Springer, 2007. ISBN: 978-1-84628-318-5. DOI: [10.1007/978-1-84628-319-2](https://doi.org/10.1007/978-1-84628-319-2).
- [45] W. F. Daenzer und R. Haberfellner, Hrsg. *Systems Engineering: Methodik und Praxis*. 11., durchges. Aufl. Zürich: Verl. Industrielle Organisation, 2002. ISBN: 385743998X.
- [46] U. Lindemann, M. Maurer und T. Braun. *Structural Complexity Management: An Approach for the Field of Product Design*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. ISBN: 978-3-540-87888-9. DOI: [10.1007/978-3-540-87889-6](https://doi.org/10.1007/978-3-540-87889-6).
- [47] G. E. P. Box und J. S. Hunter. „Multi-Factor Experimental Designs for Exploring Response Surfaces“. In: *The Annals of Mathematical Statistics* 28.1 (1957), S. 195–241.
- [48] G. E. P. Box und D. W. Behnken. „Some New Three Level Designs for the Study of Quantitative Variables“. In: *Technometrics* 2 (1960), S. 455–475. ISSN: 0040-1706.
- [49] B. Jones und C. J. Nachtsheim. „Efficient Designs With Minimal Aliasing“. In: *Technometrics* 53.1 (2011), S. 62–71. ISSN: 0040-1706. DOI: [10.1198/TECH.2010.09113](https://doi.org/10.1198/TECH.2010.09113).

- [50] A. M. Dean, D. Voss und D. Draguljić. *Design and analysis of experiments*. Second edition. Springer texts in statistics. Cham: Springer, 2017. ISBN: 978-3-319-52248-7. DOI: [10.1007/978-3-319-52250-0](https://doi.org/10.1007/978-3-319-52250-0).
- [51] G. Box, S. Bisgaard und C. Fung. „An explanation and critique of taguchi’s contributions to quality engineering“. In: *Quality and Reliability Engineering International* 4.2 (1988), S. 123–131. ISSN: 0748-8017. DOI: [10.1002/qre.4680040207](https://doi.org/10.1002/qre.4680040207).
- [52] S. Bisgaard. „A Conceptual Framework for the use of Quality Concepts and Statistical Methods in Product Design“. In: *Journal of Engineering Design* 3.1 (1992), S. 31–47. ISSN: 0954-4828. DOI: [10.1080/09544829208914746](https://doi.org/10.1080/09544829208914746).
- [53] W. Kleppmann. *Versuchsplanung: Produkte und Prozesse optimieren*. 9., überarbeitete Auflage. Hanser eLibrary. München: Hanser, 2016. ISBN: 978-3-446-44716-5. DOI: [10.3139/9783446447172?locatt=mode:legacy](https://doi.org/10.3139/9783446447172?locatt=mode:legacy).
- [54] R. H. Myers, D. C. Montgomery, G. G. Vining und T. J. Robinson. *Generalized linear models: With applications in engineering and the sciences*. 2. ed. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2010. ISBN: 9780470454633. DOI: [10.1002/9780470556986](https://doi.org/10.1002/9780470556986).
- [55] A. Zahran, C. M. Anderson-Cook und R. H. Myers. „Fraction of Design Space to Assess Prediction Capability of Response Surface Designs“. In: *Journal of Quality Technology* 35.4 (2003), S. 377–386. ISSN: 0022-4065.
- [56] A. Giovannitti-Jensen und R. H. Myers. „Graphical Assessment of the Prediction Capability of Response Surface Designs“. In: *Technometrics* 31.2 (1989), S. 159–171. ISSN: 0040-1706.
- [57] D. A. Belsley. *Regression diagnostics: Identifying influential data and sources of collinearity*. Wiley series in probability and statistics. New York: Wiley, 2004. ISBN: 9780471725145. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=226535>.
- [58] L. Fahrmeir, T. Kneib und S. Lang. *Regression: Modelle, Methoden und Anwendungen*. 2. Auflage. Statistik und ihre Anwendungen. Heidelberg und Berlin: Springer, 2009. ISBN: 978-3-642-01836-7. DOI: [10.1007/978-3-642-01837-4](https://doi.org/10.1007/978-3-642-01837-4).

-
- [59] S. P. Boyd und L. Vandenberghe. *Convex optimization*. Cambridge, New York, NY und Port Melbourne: Cambridge University Press, 2004. ISBN: 978-0-521-83378-3. URL: <http://www.loc.gov/catdir/description/cam041/2003063284.html>.
- [60] K. G. Russell. *Design of experiments for generalized linear models*. Chapman and Hall/CRC Interdisciplinary Statistics Ser. Boca Raton, London und New York: CRC Press Taylor & Francis Group, 2019. ISBN: 9780429057489. DOI: [10.1201/9780429057489](https://doi.org/10.1201/9780429057489).
- [61] A. C. Atkinson. „The Design of Experiments to Estimate the Slope of a Response Surface“. In: *Biometrika* 57.2 (1970), S. 319. ISSN: 00063444. DOI: [10.2307/2334839](https://doi.org/10.2307/2334839).
- [62] A. C. Atkinson, A. N. Donev und R. D. Tobias. *Optimum Experimental Designs, with SAS*. 1st. New York: Oxford University Press Inc, 2007. ISBN: 9780199296590.
- [63] E. M. Monroe, R. Pan, C. M. Anderson–Cook, D. C. Montgomery und C. M. Borrer. „A generalized linear model approach to designing accelerated life test experiments“. In: *Quality and Reliability Engineering International* 27.4 (2011), S. 595–607. ISSN: 0748-8017. DOI: [10.1002/qre.1143](https://doi.org/10.1002/qre.1143).
- [64] J. Cohen. *Statistical power analysis for the behavioral sciences*. Second edition. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1988. ISBN: 9781134742707. DOI: [10.4324/9780203771587](https://doi.org/10.4324/9780203771587).
- [65] M. Arndt, P. Mell und M. Dazer. „Generic effects of deviations from test design orthogonality on test power and regression modelling of Central-Composite Designs“. In: *Proceedings of the 16th International Probabilistic Safety Assessment and Management Conference (PSAM16)*. Bd. PSAM 16. Honolulu, HI, USA, 2022.
- [66] A. Khuri. *Response Surface Methodology and Related Topics*. 1st. World Scientific Publishing Co Pte Ltd, 2006. ISBN: 9789812564580.
- [67] G. E. P. Box und D. R. Cox. „An Analysis of Transformations“. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 26.2 (1964), S. 211–243. ISSN: 1369-7412. DOI: [10.1111/j.2517-6161.1964.tb00553.x](https://doi.org/10.1111/j.2517-6161.1964.tb00553.x).

- [68] M. V. Wüthrich und M. Merz. *Statistical Foundations of Actuarial Learning and its Applications*. Cham: Springer International Publishing, 2023. ISBN: 978-3-031-12408-2. DOI: [10.1007/978-3-031-12409-9](https://doi.org/10.1007/978-3-031-12409-9).
- [69] M. Modarres, M. Nuri-Amiri und C. Jackson. *Probabilistic physics of failure approach to reliability: Modeling, accelerated testing, prognosis and reliability assessment*. Performability engineering series. Beverly, MA und Hoboken, NJ: Scrivener Publishing und John Wiley & Sons Inc, 2017. ISBN: 9781119388630.
- [70] A. Kremer, L. Dücsö und B. Bertsche. „Reliability Prediction using Design of Experiments“. In: *Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference*. Hrsg. von P. Baraldi, F. Di Maio und E. Zio. Singapore: Research Publishing Services, 2020, S. 1501–1507. ISBN: 978-981-14-8593-0. DOI: [10.3850/978-981-14-8593-0{\textunderscore}3538-cd](https://doi.org/10.3850/978-981-14-8593-0{\textunderscore}3538-cd).
- [71] A. J. Dobson und A. G. Barnett. *An Introduction to Generalized Linear Models*. Fourth edition. Texts in statistical science. Boca Raton, FL: CRC Press, 2018. ISBN: 9781315182780. DOI: [10.1201/9781315182780](https://doi.org/10.1201/9781315182780).
- [72] M. Aitkin und D. Clayton. „The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data Using GLIM“. In: *Journal of the Royal Statistical Society Series C: Applied Statistics* 29.2 (1980), S. 156. ISSN: 0035-9254. DOI: [10.2307/2986301](https://doi.org/10.2307/2986301).
- [73] D. Rasch und D. Schott. *Mathematical statistics*. Hoboken, NJ: John Wiley & Sons Inc, 2018. ISBN: 9781119385288.
- [74] L. Cui, I. B. Frenkel und A. Lisnianski, Hrsg. *Stochastic models in reliability engineering*. Advanced Research in Reliability and System Assurance Engineering. Boca Raton, London und New York: CRC Press Taylor & Francis Group, 2021. ISBN: 9780429331527. DOI: [10.1201/9780429331527](https://doi.org/10.1201/9780429331527).
- [75] M. Risberg Ellekjær und S. Bisgaard. „The use of experimental design in the development of new products“. In: *International Journal of Quality Science* 3.3 (1998), S. 254–274. ISSN: 1359-8538. DOI: [10.1108/13598539810229230](https://doi.org/10.1108/13598539810229230).
- [76] S. Bisgaard. *Time series analysis and forecasting by example*. Wiley series in probability and statistics. Hoboken, N.J: Wiley, 2011. ISBN: 1118056949. DOI: [10.1002/9781118056943](https://doi.org/10.1002/9781118056943).

-
- [77] R. T. Johnson, D. C. Montgomery und B. A. Jones. „An Expository Paper on Optimal Design“. In: *Quality Engineering* 23.3 (2011), S. 287–301. ISSN: 0898-2112. DOI: [10.1080/08982112.2011.576203](https://doi.org/10.1080/08982112.2011.576203).
- [78] A. N. Donev. „Design of experiments in the presence of errors in factor levels“. In: *Statistical Planning and Inference* 126.2 (2004), S. 569–585.
- [79] G. E. P. Box und K. B. Wilson. „On the Experimental Attainment of Optimum Conditions“. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 13.1 (1951), S. 1–45. ISSN: 00359246. URL: <http://www.jstor.org/stable/2983966> (besucht am 23. 03. 2023).
- [80] M. K. Ardakani, D. Das, S. S. Wulf und T. J. Robinson. „Estimation in Second-Order Models with Errors in the Factor Levels“. In: *Commun. in Statistics* 40.9 (2011), S. 1573–1590. DOI: [10.1080/03610921003637421](https://doi.org/10.1080/03610921003637421).
- [81] B. A. Jones und P. Goos. „I-Optimal Versus D-Optimal Split-Plot Response Surface Designs“. In: *Journal of Quality Technology* 44.2 (2012), S. 85–101. ISSN: 0022-4065.
- [82] B. A. Jones, K. Allen-Moyer und P. Goos. „A-optimal versus D-optimal design of screening experiments“. In: *Journal of Quality Technology* 53.4 (2021), S. 411–427. ISSN: 0022-4065.
- [83] L. A. Escobar und W. Q. Meeker. „Planning Accelerated Life Tests with Two or More Experimental Factors“. In: *Technometrics* 37.4 (1995), S. 411–427. ISSN: 0040-1706.
- [84] H. Ahn. „Central Composite Design for the Experiments with Replicate Runs at Factorial and Axial Points“. In: *Industrial Engineering* 349 (2015), S. 969–979.
- [85] H. Xu. „An Algorithm for Constructing Orthogonal and Nearly-Orthogonal Arrays with Mixed Levels and Small Runs“. In: *Technometrics* 44.4 (2002), S. 356–368. ISSN: 0040-1706.
- [86] A. Wald. „On the Efficient Design of Statistical Investigations“. In: *The Annals of Mathematical Statistics* 14.2 (1943), S. 134–140.
- [87] A. C. Rencher und G. B. Schaalje. *Linear models in statistics*. 2. ed. Hoboken, NJ: Wiley-Interscience, 2008. ISBN: 9780471754985. URL: <http://www.loc.gov/catdir/enhancements/fy0826/2007024268-b.html>.

- [88] G. E. P. Box und N. R. Draper. *Response surfaces, mixtures, and ridge analyses*. 2. ed. Wiley series in probability and statistics. Hoboken, New Jersey: Wiley, 2007. ISBN: 9780470053577. URL: <http://www.loc.gov/catdir/enhancements/fy0826/2006043975-b.html>.
- [89] B. R. Englert, S. E. Rigdon, C. M. Borror, D. C. Montgomery und R. Pan. „Optimal Design for Multifactor Life Testing Experiments for Exponentially Distributed Lifetimes“. In: *Frontiers in Statistical Quality Control 10*. Hrsg. von H.-J. Lenz, W. Schmid und P.-T. Wilrich. SpringerLink Bücher. Heidelberg und s.l.: Physica-Verlag HD und Imprint Physica, 2012, S. 303–317. ISBN: 978-3-7908-2845-0. DOI: [10.1007/978-3-7908-2846-7](https://doi.org/10.1007/978-3-7908-2846-7){\textunderscore}20.
- [90] A. B. Çolak, T. N. Sindhu, S. A. Lone, M. T. Akhtar und A. Shafiq. „A comparative analysis of maximum likelihood estimation and artificial neural network modeling to assess electrical component reliability“. In: *Quality and Reliability Engineering International* 40.1 (2024), S. 91–114. ISSN: 0748-8017. DOI: [10.1002/qre.3233](https://doi.org/10.1002/qre.3233).

Anhang A Ableitungen zur Maximum-Likelihood-Schätzung

A.1 Modell Daefinition (Modell zweiter Ordnung)

Für ein System mit $k = 2$ Einflussfaktoren (x_1, x_2) basierend auf einem CCD ist der lineare Prädiktor η_i definiert. Er beschreibt den Logarithmus der charakteristischen Lebensdauer wie folgt:

$$\eta_i = \ln(T_i) = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \beta_{12} x_{i,1} x_{i,2} + \beta_{11} x_{i,1}^2 + \beta_{22} x_{i,2}^2 \quad (.1)$$

Der zugehörige Parametervektor β und der Regressorvektor x_i lauten:

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \\ \beta_{11} \\ \beta_{22} \end{bmatrix}, \quad x_i = \begin{bmatrix} 1 \\ x_{i,1} \\ x_{i,2} \\ x_{i,1} x_{i,2} \\ x_{i,1}^2 \\ x_{i,2}^2 \end{bmatrix} \quad (.2)$$

A.2 Log-Likelihood-Funktion

Wir führen zunächst den Indikator δ_i ein ($\delta_i = 1$ für Ausfall, 0 für Zensierung). Zusätzlich definieren wir die Hilfsvariable z_i :

$$z_i = t_i^b \exp(-b\eta_i) \quad (.3)$$

Damit lautet die Log-Likelihood-Funktion Λ :

$$\Lambda = \sum_{i=1}^n [\delta_i (\ln(b) - b\eta_i + (b-1) \ln(t_i)) - z_i] \quad (.4)$$

A.3 Score-Vektor (Erste Ableitungen)

Die partiellen Ableitungen (Gradient) dienen zur Maximierung der Funktion. Die Ableitung nach den Regressionskoeffizienten für $j \in \{0, 1, 2, 12, 11, 22\}$ lautet:

$$\frac{\partial \Lambda}{\partial \beta_j} = b \sum_{i=1}^n (z_i - \delta_i) x_{i,j} \quad (.5)$$

Die Ableitung nach dem Formparameter b ergibt sich zu:

$$\frac{\partial \Lambda}{\partial b} = \sum_{i=1}^n \left[\frac{\delta_i}{b} + (\delta_i - z_i)(\ln(t_i) - \eta_i) \right] \quad (.6)$$

A.4 Hesse-Matrix (Zweite Ableitungen)

Die Elemente der Hesse-Matrix werden für die Fisher-Information und den Newton-Raphson-Algorithmus benötigt. Die zweite Ableitung nach den Koeffizienten β (für Indizes j und l) ist:

$$\frac{\partial^2 \Lambda}{\partial \beta_j \partial \beta_l} = -b^2 \sum_{i=1}^n z_i x_{i,j} x_{i,l} \quad (.7)$$

Die zweite Ableitung nach dem Formparameter b lautet:

$$\frac{\partial^2 \Lambda}{\partial b^2} = - \sum_{i=1}^n \left[\frac{\delta_i}{b^2} + z_i (\ln(t_i) - \eta_i)^2 \right] \quad (.8)$$

Die gemischte Ableitung nach β und b berechnet sich wie folgt:

$$\frac{\partial^2 \Lambda}{\partial \beta_j \partial b} = \sum_{i=1}^n x_{i,j} [(z_i - \delta_i) + b z_i (\ln(t_i) - \eta_i)] \quad (.9)$$