INHALTDIS

EINFÜHRUNG IN DIE EXPERIMENTPLANUNG UND DIE BEDEUTUNG DER STATISTIK

EINFÜHRUNG UND PROBLEMBESCHREIBUNG

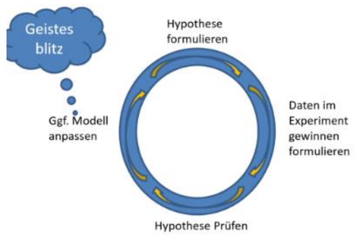
WARUM VERSUCHE/EXPERIMENTE

- ***Funktionsumfang*** der Produkte ***erhöhen***(Kundenanforderungen gerecht werden)
- ***Kosten*** müssen ***gesenkt*** werden
- **Entwicklungszeit**/Fertigung, etc. **verkürzen**
- Dazu Zusammenhänge/Parameter untersuchen(analysieren + verbessern)
- Verbesserungen anwenden – überprüfen – bewerten
- Gezielte Versuche/Expertimente sind erforderlich

WISSENSERWERB NACH SHEWHART

Wissenserwerb durch:

- Wahrnehmung
- Beobachten
- Lernen
- Denken und Schliessen
- Erklärungen(Modelle) finden und beschreiben
- Hypothesen formulieren & validieren



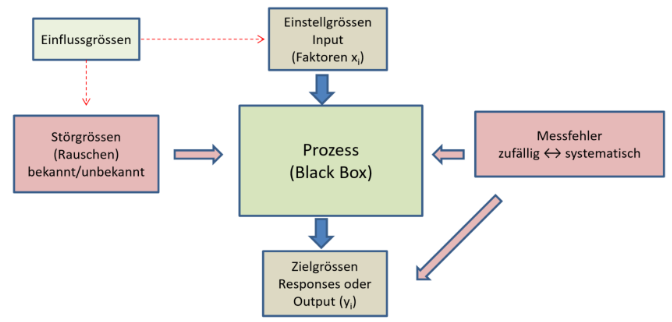
→dafür braucht es Experimente(am realen System oder Modell)

ASPEKTE DER INDUSTRIELLEN VERSUCHSPLANUNG

Die Hindernisse beim Experimentieren/Erkenntnisgewinn:

- ***Komplexität*** – Vielgestaltigkeit(viele Faktoren)
- ***Kompliziertheit*** – Schwierigkeit
  - Schwierige/nicht verstandenen/unbekannte/unerklärbare Zusammenhänge oder Faktoren
- ***Rauschen, Dynamik*** – Fehlereinflüsse → Messungenauigkeit
  - Unterschiedliche Ergebnisse bei gleiche Faktoren

DAS PROZESSMODELL DES EXPERIMENTS



Experimente werden immer nach einem bestimmen Schema durchgeführt:
1. Ausgangssituation beschreiben
2. Untersuchungsziele festlegen / Zielgrössen definieren
3. Faktoren auswählen und gewichten
4. Versuchsplanung erstellen
5. Versuche durchführen
6. Ergebnisse auswerten und Vertrauensintervalle bestimmen
7. Ergebnisse interpretieren und Massnahmen ableiten
8. Überprüfen der «Verbesserungen»

BEGRIFFE

Systematischer Fehler – Fehler, welcher «mit System» auftritt. Kann deshalb rausgerechnet werden

- Messapparatur falsch geeicht
- verunreinigte Salzlösung

Zufälliger Fehler –zufällig entstandener Fehler, kann nicht rausgerechnet(vorhergesehen) werden.

- Thermisches Rauschen eines Sensors

DOE: DESIGN OF EXPERIMENT

VERSUCHSPLANUNG

So wenig Versuche wie möglich.
Zwei verschiedene Ansätze:

- Immer nur eine der Grösse(Faktoren) verändern
- Alle möglichen Kombinationen der Faktoren verändern
  - Kein Probieren – strukturelles Vorgehen

BEGRIFFE

Zielgrössen – beschreiben das Ergebnis eines Versuchs(Messwerte oder errechnete Grössen[Achtung! Fehlerfortplanzung]

Einflussgrössen – Grössen die die Zielgrössen beinflussen können

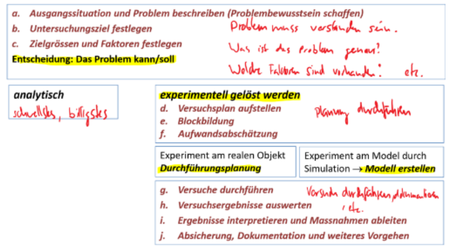
- Steuergrössen – einstellbar evtl. schwankend
- Störgrössen – nicht vorgegeben/vorhersehbar

Faktoren – für den Versuch wesentlichen Einflussgrössen

Faktorstufen – Werte die die Faktoren im Versuch annehmen sollen:

- Quantitative Faktoren – Stufen mit Zahlenwerten
- Qualitative Faktoren – Namen, Beschreib./Bezeic.

VORGEHENSWEISE



<u>Ausgangssituation spezifizieren / Problem beschreiben / Ziel definieren</u>

- Kunde und dessen Bedürfnisse definieren
- Liegen bereits Daten vor?
- Welche Probleme müssen gelöst werden
- Welche Ressourcen (Zeit und Geld) stehen zur Verfügung (Kosten/Nutzen Analyse)
- Betroffene Gruppen und deren Beziehung untereinander listen (Wiederstände, Supporter, Wissensträger)
- ***Ist das Problem verstanden?***
- ***Was ist das Problem genau?***

<u>Untersuchungsziel festlegen</u>

- Wertdefinition von Prozessergebnissen oder -parameter
- Versuch Streuung/Robustheit zu reduzieren
- Wichtigsten Störgrössen erkennen/definieren
- Prozesse optimieren durch fertigen + lernen
- Funktion/Zuverlässigkeit nachweisen

<u>-Zielgrössen/Faktoren festlegen</u>

- Zielgrössen beschreiben(Kunden- & Zielorientiert)
- Quantifizierung → Zielgrössen müssen messbar sein
- Vollständige Ergebnisse & Eigenschaften berücksichtigen
- Jede Zielgrösse möglichst grundlegenden Zusammenhang
- ***Erst möglich viele(alle) Einflussgrössen sammeln und dann auf die (wesentlichen) Faktoren reduzieren***

Unterstützung durch Diagramme, Tabellen, etc.

<u>Lösen des Problems</u>

- ***Analytisch*** – schnellstes, billigstes
- ***Experimentell***
  - Planung durchführen
  - Realen Objekt oder Simulation?
  - Versuch durchführen/dokumentieren/auswerten

→Experiment auswerten → Massnahmen/Weiteres ableiten

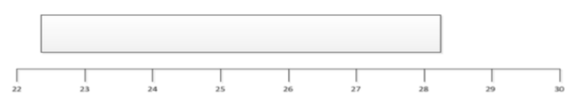
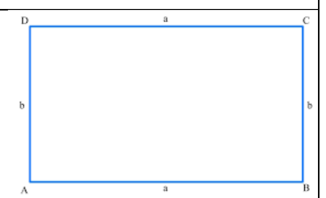
FEHLERFORTPFLANZUNG

Weiterziehen eines gemessenen Fehlers in einer Rechnung. Auch das Resultat wird den Messfehler enthalten.

FEHLERRECHNUNG

- Der Begriff ist irreführend denn man kann den Fehler nicht berechnen sondern nur schätzen.
- Benötigt um ***Bereich abzuschätzen***, in denen der tatsächliche Wert mit einer gewisser Wahrscheinlichkeit liegt

Daher wird immer ein Angabe über die Genauigkeit des Wertes angegeben!

<table>
<tr><td>Absoluter Fehler Δx</td><td>Mit dem absoluten Fehler wird das Intervall berechnet oder geschätzt, in welchem sich das Ergebnis befinden sollte. Der absolute Fehler wird immer mit einer Masseinheit angegeben.
Formel: relativer Fehler * Messung(Wert)
Bei Summen und Differenzen addieren sich die absoluten Fehler.
Gut für Ergebnisdarstellungen(z.B 5.8cm ± 0.1cm)</td></tr>
<tr><td>Implizite Fehlerannahme</td><td>Oft gibt man den Fehlereingabe nicht für jeden Wert an. Es hat sich daher eine Regel eingebürgert. Fehler bei Nennwerteingabe:
Mindestens einer halben Einheit der letzten(rechten) Stelle
Höchstens 4 Einheiten der letzten bedeutsamen Stelle
<table>
<tr><td>t = 15.32 s</td><td>bedeutet möglicher Fehler</td><td>Δt = ± 0.005 s</td><td>bis etwa</td><td>± 0.04 s</td></tr>
<tr><td>t = 15.3 s</td><td>bedeutet möglicher Fehler</td><td>Δt = ± 0.05 s</td><td>bis etwa</td><td>± 0.4 s</td></tr>
<tr><td>t = 15.320 s</td><td>bedeutet möglicher Fehler</td><td>Δt = ± 0.0005 s</td><td>bis etwa</td><td>± 0.004 s</td></tr>
</table></td></tr>
<tr><td>Relativer Fehler $\frac{\Delta x}{x}$</td><td>Mit dem Relativen Fehler wird prozentuale Abweichung des Ergebnis berechnet, daher ist er auch ohne einheitenlos(in Prozent):
Relativer Fehler: $\frac{Absoluter\ Fehler}{Messung}$
Mit Hilfe des relativen Fehler lässt sich gut Abschätzen, welcher Faktor verbessert werden sollte. (der mit dem grösseren Fehleranteil)
Bei Produkten und Quotienten addieren sich die relativen Fehler.
Gut um Gefühl der Messgenauigkeit zu erhalten(Bei Messgeräten relativer Fehler bezogen auf eingestellten Messbereich)
Beispiel: Voltmeter hat bei 20-100V eine Abweichung(relativer Fehler) von 0.001=0.1%</td></tr>
<tr><td>Beispiel Absoluter Fehler Lineal</td><td>Auf beiden Seiten wird abgelesen und die Fehlertoleranzen addiert ergeben den absoluten Fehler.

Länge wird abgelesen: L = 28.15cm – 22.35cm = 5.8cm
Abgelesene Unsicherheit der Werte ist ± 0.05cm
//Differenz→Addieren des <u>absoluten</u> Fehlers:
Addieren des absoluten Fehlers: ΔL = 0.05 cm + 0.05 cm = 0.1 cm
Resultat notieren: 5.8cm ± 0.1cm
Der Fehler befindet sich also im Intervall [5.7 ; 5.9]</td></tr>
<tr><td>Beispiel Relativer Fehler</td><td>Länge: L = 5.8cm
absoluter Fehler: ΔL = 0.1cm
Einfügen in die Formel: $\frac{Absoluter\ Fehler}{Messung} = \frac{0.1}{5.8} = 0.017 = 1.7\%$</td></tr>
<tr><td>Beispiel Relativer Fehler Fläche</td><td>Es soll die Fläche bestimmt werden.
Länge: 5.8cm ± 0.1cm
Breite: 0.9 ± 0.1 cm
Fläche: 5.8 * 0.9 = 5.22 cm^2
Relativer Fehler Breite: ΔB = 0.1/0.9 = 11.1%
Relativer Fehler Länge: ΔL = 0.1/5.8cm = 1.7%
//Multiplikation: Addieren des <u>relativen</u> Fehlers
Relativer Fehler Fläche: ΔB + ΔL = 12.8%
Absoluter Fehler Fläche: ΔA = relativer Fehler * Messung(Wert) = 0.128 * 5.22 = 0.668cm^2
Ergebnis: 5.2cm^2 ± 0.7cm^2.
</td></tr>
<tr><td>Methode mehr als zwei Messgrössen</td><td>$E = \frac{x\,y^2}{z} \rightarrow \frac{\Delta E}{E} = \frac{\Delta x}{x} + 2\frac{\Delta y}{y} + \frac{\Delta z}{z}$ $\Delta E = \left|\frac{\delta E}{\delta x}\Delta x\right| + \left|\frac{\delta E}{\delta y}\Delta y\right| + \left|\frac{\delta E}{\delta z}\Delta z\right|.$
Sobald Exponenten vorkommen → diese im Produkt nach vorne nehmen</td></tr>
</table>

Methode min/max	$E = \frac{u-v}{x-y} \rightarrow E_{max} = \frac{u_{max}-v_{min}}{x_{mi}\ -y_{max}} \rightarrow \Delta E = E_{max} - E$ h gewählter Punkt. Somit kann man ten den Abstand(Intervall) messen, jedoch kein Verhältnis bilden werden(8:00 nicht doppelt 4:00): Wenn mann nur Fehlergrenzen berechnen muss(nur absoluter Fehler)

ABLEITUNGSREGELN

Summenregel	$(f(x)+g(x))' = f'(x)+g'(x)$
Produktregel	$(f(x)\cdot g(x))' = f'(x)\cdot g(x) + f(x)\cdot g'(x)$
Kettenregel	$(f(g(x)))' = f'(g(x))\cdot g'(x)$
Quotientenregel	$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)\cdot g(x) - f(x)\cdot g'(x)}{(g(x))^2}$

STATISTISCHE GRUNDBEGRIFFE

Merkmalsträger: Gegenstand der statistischen Untersuchung. «Träger der Merkmale» z.B Werkzeugmaschine

Grundgesamtheit: Menge aller Merkmalsträger mit übereinstimmenden Abgrenzungsmerkmalen(zeitlich, räumlich,sachlich).
Z.B. Menge aller Werkzeugmaschinen Standort Zürich
Alle Maschinen Typ B-299
Alle Ausfälle in den letzten 30 Tagen

Merkmal(Information): Eigenschaft des Merkmalträgers.
Z.B. Ausfallraten, Servicedauer, Leistung, etc.

Merkmalswert: Wert des eines Merkmals des Merkmalträgers.
Z.B. Maschine XY hat Leistung 4000W oder Ausfallrate 0.01%

SKALENE

In Skalen sind mögliche Merkmalswerte nach einer Ordnung abgelegt.

NOMINALSKALA

Auf der Nominalskala sind als Skalenwerte ***Namen*** abgetragen, die ***gleichberechtigt*** bzw. ***gleichbedeutend*** nebeneinander angeordnet sind.

Merkmal	Merkmalswert
Geschlecht	männlich, weiblich
Familienstand	ledig, verheiratet, geschieden, verwitwet
Religion	katholisch, evangelisch
Rebsorte	Silvaner, Riesling, Portugieser, Traminer, ...

ORDINALSKALA

Auf der Ordinalskala sind als Skalenwerte ***Klassenbezeichnungen*** abgetragen. Die Skalenwerte stehen jetzt ***nicht mehr gleichberechtigt*** nebeneinander, sondern sind entsprechend ihrer Klasse in ***auf- oder absteigender Folge*** auf der Skala angeordnet.

METRISCHE SKALA(KARDINALSKALA)

Schulnote	Sehr gut, gut, befriedigend, ausreichend, mangelhaft
Qualitätsstufe	Standard, Business, First Class

Die Metrische Skala beinhaltet reelle Zahlen entsprechend auf- oder absteigender Folge.

INTERVALLSKALA

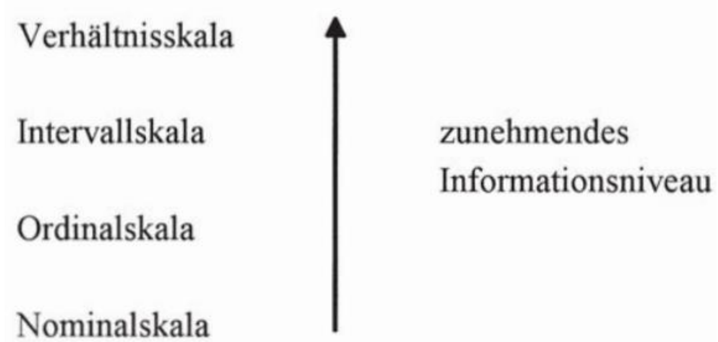
Temperatur	-12, .. 0, ..42, ..
Uhrzeit	20:00, .. 0.00, .. 10:00

VERHÄLTNISSKALA

Auf der Verhältnisskala entspricht der Skalenwert Null dem natürliche, absoluten Null Punkt. Negative Werte somit nicht möglich, dafür kann man das Verhältnis zwischen zwei Werten messen.(20Jahre ist doppel so viel wie 10Jahre)

Gewicht (kg)	0, .. 10, ...20,... 40, ..
Alter (Jahre)	0, 1, .. 40, .. 89, ..

INFORMATIONSNIVEAU



ABLAUF DER STATISTISCHEN UNTERSUCHUNG

1. ***Planung***
   I. **Festlegung der folgenden Punkte**: Merkmal, Merkmalträger, Erhebungstechnik, Aufbereitungsverfahren, Darstellungsform, statistische Analyseverfahren
2. ***Datenerhebung***
   I. **Konkretisierung Untersuchungsziel**: Abgrenzungsmerkmale, Untersuchungsmerkmale, Datenumfang, erwartetes Ergebnis (Ziel/Problem festlegen)
   II. **Erhebungstechniken**: Messungen, Zählung, Befragung, Beobachtung
   III. **Herkunft der Daten**:
      - Primärstatistik
         - Daten erstmalig erhoben
         - Exakt fürs Ziel gerichtet/erzeugt
      - Sekundärstatistik
         - Bereits vorliegendes Material
         - Kostengünstig
         - Nicht aktuell & nicht exakt
   IV. **Erhebungsumfang**: Teil- oder Vollerhebung
         - Bei experiment immer nur Teilerhebung(Stichprobe)→schliessende Statis.
3. ***Datenaufbereitung und -darstellung***
   I. **Daten von Urliste überführen in**: Strichliste, Häufigkeitstabelle
4. ***Analyse und Interpretation***

DATAMINING & DATAFARMING

DATAMINING

Vereinfacht gesagt ist Data-Miniing das Auswerten von grossen Daten und daraus Rückschlüsse wie Regelmäßigkeiten, Gesetzmäßigkeiten und verborgener Zusammenhänge zu finden.

DATAFARMING

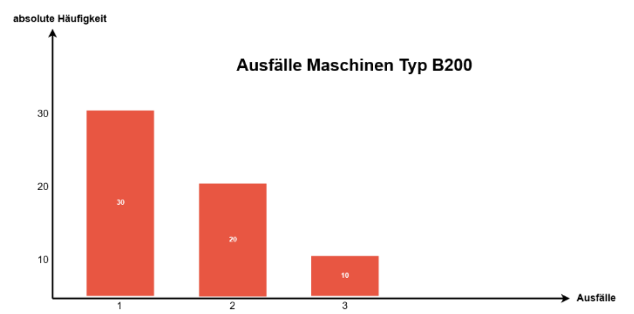
Vereinfacht gesagt will man mittels Data-Farming Daten erzeugen, welche dann wieder für Experimente gebraucht werden können.

HÄUFIGKEITSVERTEILUNGEN

EINFACHE HÄUFIGKEIT

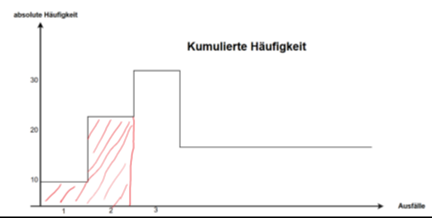
Gibt an, ***wie häufig ein Merkmalswert*** xi aufgetreten ist. Dieser kann ***absolut oder relativ*** ausgedrückt werden.

absolute einfache Häufigkeit h_i	**Anzahl** der Merkmalsträger mit dem Merkmals-wert x_i (i = 0, .. v) Wie viel Mal kommt der Wert xi gesamthaft vor?
relative einfache Häufigkeit f_i	**Anteil** der Merkmalsträger mit dem Merkmals-wert i (i: 1., .. , v) Wie ist das Verhältnis/Anteil dieses Merkmalswert gegenüber allen Werten?
Gesamtzahl aller Merkmalsträger n	Summe aller Merkmalsträger, welche vorkommen. Wie viele Merkmalsträger gibt es insgesamt?
Anzahl verschiedener Merkmalswerte v	Summe aller verschiedener Merkmalsträger, welche vorkommen. Wie viele <u>verschiedene</u> Merkmalsträger gibt es?
Formeln	$n = \sum_{i=1}^{v} h_i$ $f_i = \frac{h_i}{n}$ $\sum_{i=1}^{n} f_i = 1$
Beispiel	**<u>Ausfälle Maschinen B200</u>** **Absolute Häufigkeit**(ablesbar) $h_1 = 30$, $h_2 = 20$, $h_3 = 10$ **Gesamtanzahl** n = 30+20+10 = 60 **Relative Häufigkeit** $f_1 = 30/60 = 0.5$ $f_2 = 20/60 = 0.33$ $f_3 = 10/60 = 0.16$ **Anzahl verschiedene Merkmalsträger** = 3



KUMULIERTE HÄUFIGKEIT(SUMMENHÄUFIGKEIT)

Absolute kumulierte Häufigkeit H_i	**Anzahl** der Merkmalsträger mit einem Merkmalswert der **kleiner oder gleich** x_i ist
Relative kumulierte Häufigkeit	**Summe der Anteile** der Merkmalsträger mit dem Merkmalswert kleiner gleich i.
Formeln	$H_i = h_1 + h_2 + \ldots + h_i = \sum_{a=1}^{i}$ $F_i = f_1 + f_2 + \ldots + f_i = \sum_{a=1}^{i} f_a = \frac{H_i}{n} = 1\,, denn$ $\sum_{i=1}^{n} f_i = 1$



KLASSIFIZIERTE HÄUFIGKEIT

Bei mehr als 10 bis 15 verschiedenen Merkmalswerten ist die Darstellung nicht mehr überschaubar, man wechselt darum auf die Klassifizierte Häufigkeit. Dabei werden ***Klassen von Merksmalswerten(Intervalle)*** gebildet.

j = Klassenzahl, Vorschlag für $j_{max} = \sqrt{n}$
x_j^u = Untere Klassengrenze für Merkmalswert
x_j^o = Obere Klassengrenze für Merkmalswert

J	Rechnungsbetrag $x_j^u \leq x_i < x_j^o$	h_j	H_j	f_j	F_j
1	0 bis 20	10	10	0.09	0.09
2	20 bis 40	20	30	0.18	0.27
3	40 bis 90	80	110	0.73	1
		110		1	

GRUNDLAGEN DER BESCHREIBENDEN STATISTIK

EINFÜHRUNG UND PROBLEMBESCHREIBUNG

Mit ***Mittelwerte*** können typische Eigenschaften der Häufigkeitsverteilung beschrieben werden. Sie sind zudem sind kompakter als Tabellierte Häufigkeiten. Mithilfe von Kenngrössen, den Parametern, können aussagekräftige Grössen definiert werden, welche die grosse Menge von Messdaten verdichten. Diese Menge nennt man ***Streuungsmasse***.
Anhand von empirischen Daten versucht man Kenngrössen Daten zu beschrieben/charakterisieren.

MITTELWERTE

MODUS

Grundlagen

Der Modus ist der Wert, welcher ***am häufigsten beobachtet*** wird.

- Für jede Verteilung bestimmbar.
- Von Ausreissern unbeeinflussbar
- Schnelle und einfache Ermittlung
- Als Mittelwert geeignet wenn seine Häufigkeit die anderen Häufigkeiten dominiert

Formeln bei gleicher Klassen-breite

$$Mo = x_m^u + \frac{h_m - h_{m-1}}{(h_m - h_{m-1}) + (h_m - h_{m+1})} \times (x_m^o - x_m^u)$$

- x_m^u: Unterer Klassengrenze
- x_m^o: oberere Klassengrenze
- h_m: Modusklasse
- h_{m-1}: Klasse links von Modus
- h_{m+1}: Klasse rechts von Modus

Formel bei ungleicher Klassen-breite

$$Mo = x_m^u + \frac{d_m - d_{m-1}}{(d_m - d_{m-1}) + (d_m - d_{m+1})} \times (x_m^o - x_m^u)$$

- x_m^u: Unterer Klassengrenze
- x_m^o: oberere Klassengrenze
- d_m: Dichte Modusklasse
- d_{m-1}: Dichte Klasse links von Modus
- d_{m+1}: Dichteklasse rechts von Modus

Dichte Formel

$$Häufigkeitsdichte = \frac{absolute\ Häufigkeit}{x_m^o - x_m^u}$$

MEDIAN

Grundlagen

Derjenige Merkmalswert, dessen Merkmalsträger in der Rangordnung aller Merkmalsträger genau ***in der Mitte*** ist.

- Mindestens Ordinalskala(weil die Merkmalswerte eine Rangordnung aufzeigen müssen)
- Unbeeinflusst von Ausreissern
- Gut geeignet für schiefe Verteilungen
- Vermittelt einen guten Eindruck für die «Mitte»

Formeln

$$Me = x_m^u + \frac{\frac{n}{2} - H_{m-1}}{H_m - H_{m-1}} \times (x_m^o - x_m^u)$$

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & \text{für } n \text{ ungerade} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) & \text{für } n \text{ gerade} \end{cases}$$

• H_m: Kumulierte Klassenhäufigkeit	n: Gesamtzahl aller Merkmalsträger

QUANTIL

Grundlagen

Ein Quantil ist ein Merkmalswert, durch den die Gesamtheit in **zwei Teile** zerlegt. Zusätzlich gibt es noch

Quartil = 4 Teile	***Dezile*** = zehn Teile	***Perzentile*** = 100

Formeln

Berechnung erfolgt analog des Medians(anstatt n/2→n/4 oder n/10, etc.):
Beispiel: 3, Quantil(3n/4 anstatt n/2):

$$3.Quantil = x_m^u + \frac{\frac{3n}{4} - H_{m-1}}{H_m - H_{m-1}} \times (x_m^o - x_m^u)$$

ARITHMETISCHES MITTEL

Grundlagen

Das arithmetische Mittel ist die ***Summe*** aller Merkmalswerte ***geteilt*** durch die ***Anzahl*** der Merkmalswerte. Das ist der Durchschnitt, wie man in «im Volksmund» kennt.

- , da Addition notwendig von Merkmalswerten
- In der Praxis sehr oft benutzt - Geeignet für symmetrische, nicht für schiefe Verteilungen
- Ausreisser können zu starken Verzerrungen füren(weil «Durchschnitt verfälscht»)

Formel

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{v} x_i \cdot h_i \quad bzw.\ \bar{x} = \sum_{i=1}^{v} x_i \cdot f_i \qquad \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

Bei der Klassifizierten Häufigkeit ergibt sich xj' aus $\frac{x_m^o - x_m^u}{2}$(Klassenmittelwert)

HARMONISCHES MITTEL

Grundlagen

Man bildet den Durchschnitt zu Merkmalswerten, welche in Summe gesehen relativ gleich weit entfernt sind.

- Merkmal muss verhältnisskaliert sein(Verhältnissskala , wegen Bildung Quotienten)
- Merkmalswerte alle positiv oder alle negativ
- Wird oft verwendet für ***Durchschnittgeschwindigkeit***(hm/h oder z.B Flasche/h)

Formeln

$$\overline{MH} = \frac{\sum_{i=1}^{v} h_i}{\sum_{i=1}^{v} \frac{h_i}{x_i}} \qquad MH = \frac{h_1 + h_2}{\frac{h_1}{x_1} + \frac{h_2}{x_2}}$$

Beispiel

Teilstrecke i	1	2
Länge g_i in km	100	100
Geschwindigkeit x_i in km/h	150	50

$$\bar{x}_{\text{harm}} = \frac{100 + 100}{\frac{100}{150} + \frac{100}{50}} = 75$$

GEOMETRISCHES MITTEL

Grundlagen

Ist die n-te Wurzel aus dem Produkten aller Merkmalswerte.

- Wird verwendet um durcschnittliche ***Entwicklungen*** oder Wachstum zu berechnen z.B Veränderungen von Geldbeträgen über Jahre oder Kappenlänge nach Anzahl Waschgängen
- Merkmalswerte > 0(wegen Division)

Formel/Beispiel

$$\sqrt[n]{\frac{endwert}{anfangswert}}$$

$$40 \xrightarrow{\cdot 1.2} 48 \xrightarrow{\cdot 1.25} 60 \xrightarrow{0.95} 57$$

$$\sqrt[3]{\frac{57}{40}} = 1.125$$

Beispiel gleicher Klassen

Klassenbreite	***absolute Häufigkeit(h_i)***	***relative Häufigkeit(f_i)***	***Häufigkeitsdichte***	***Kumulierte absolute Häufigkeit***
0 - 100	100	0.083	100/100 =1	100
100 – 200	300	0.25	300/100 =3	400
300 - 400	500	0.417	500/100 = 5	900
500 - 600	300	0.25	300/1000 = 3	1200

$$Mo = x_m^u + \frac{h_m - h_{m-1}}{(h_m - h_{m-1}) + (h_m - h_{m+1})} \times (x_m^o - x_m^u) = 300 + \frac{500 - 300}{(500 - 300) + (500 - 300)} \times (400 - 300) = 350$$

$$Me = x_m^u + \frac{\frac{n}{2} - H_{m-1}}{H_m - H_{m-1}} \times (x_m^o - x_m^u) = 300 + \frac{\frac{1200}{2} - 400}{900 - 400} \times (400 - 300) = 420$$ //600 in Klasse 3

Beispiel ungleicher Klassen-breite

Klassenbreite	***absolute Häufigkeit***	***relative Häufigkeit***	***Häufigkeitsdichte***	***Kumulierte absolute Häufigkeitsdichte***
0 - 100	100	0.01	100/100 =1	100
100 - 500	2400	0.24	2400/400 =6	2500
500 - 1000	4500	0.45	4500/500 = 9	7000
1000 - 2000	3000	0.30	3000/1000 = 3	10000

$$Mo = x_m^u + \frac{d_m - d_{m-1}}{(d_m - d_{m-1}) + (d_m - d_{m+1})} \times (x_m^o - x_m^u) = 500 + \frac{9-6}{(9-6)+(9-3)} \times (1000 - 500) = 666.667$$

$$Me = x_m^u + \frac{\frac{n}{2} - H_{m-1}}{H_m - H_{m-1}} \times (x_m^o - x_m^u) = 500 + \frac{\frac{10000}{2} - 2500}{7000 - 2500} \times (1000 - 500) = 777.8$$

Beispiel 3. Quantil

Forderung Von .... Bis CHF	h_j	H_j
50 100	15	15
100 200	50	65
200 ... 300	80	145
300 ... 400	40	185
400 ... 600	40	225
600 ... 1000	20	245

4. Klasse muss dazwischen liegen

183.75 ist zwischen 145 & 185

1. ***Bestimmen der Klasse, in welchem das 3. Quantil liegt(n = 245)***
   - *I.* $\frac{3n}{4} = 183.75$
   - *II.* 183.75 > als 145 → 4. Klasse muss genommen werden
2. ***Berechnen:***

$$300 + \frac{183.75 - 145}{185 - 140} \times (400 - 300) = \underline{\underline{396.60\ CHF}}$$

STREUUNGSMASSE

SPANNWEITE R

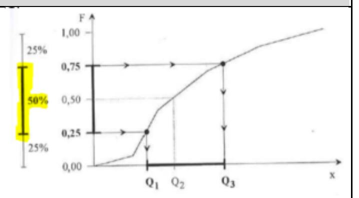
Grundlage	Die Spannweite R ist die Differenz aus dem grössten und dem kleinsten beobachteten Merkmalswert. • Braucht mindestens Intervallskala • Extrem empfindlich auf Ausreisser
Formel	R = größter Merkmalswert – kleinster Merkmalswert, $R = x[n] - x[1]$ $R = x_v^o - x_1^u$

ZENTRALER QUARTILSABSTAND(ZQA)

Grundlagen

Der zentrale Quartilsabstand ist die Entfernung zwischen den zwei Merkmalswerten, welche die zentral gelegenen 50% der Merkmalsträger einordnen.

- Mindestens Intervallskala
- Ausreisser-Problem tritt nicht auf
- Geeignet wenn Kernbereich(50%) intressiert



Formel

$ZQA = Q_3 - Q_1$

Beispiel

Ausfallzeit (h)	0	2	5	6	7	11	12	14
hi	4	2	2	2	4	3	2	1
Hi	4	6	8	10	14	17	19	20

$Q_1 = x_{\left[\frac{1}{4} \cdot n\right]} = x_{[5]} = 2\,h$

$Q_3 = x_{\left[\frac{3}{4} \cdot n\right]} = x_{[15]} = 11\,h$

ZQA = 11h -2h = 9h,

d. h.: Die mittleren 50% der Ausfallzeiten streuen um 9 Stunden
oder anders:
die mittleren 50% der Maschinen sind zwischen 2 und 11 Stunden ausgefallen

MITTLERE ABSOLUTE ABWEICHUNG

Grundlagen

Die mittlere absolute Abweichung ist die ***durchschnittliche Entfernung*** aller beobachteten Merkmalswerte ***vom arithmetischen Mittel(oder Median).*** Man rechnet die Differenz von jedem Punkt und dem Mittelwert und teilt diese nachher durch die Anzahl Werte.

- Mindestens Intervallskala
- Ausreisser können auch hier das Bild verzerren

Formel

$\delta = \frac{1}{n} \sum_{i=1}^{v} |x_i - \bar{x}|\, h_i$ x-Mittel hier arith. Mittel oder Median.

Beispiel

x_i	h_i	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert\, h_i$
1	3	2.04	6.12
2	10	1.04	10.40
3	4	0.04	0.16

n = 3 + 10 + 4 = 17
δ = 1/17 * (6.12 + 10.40 + 0.16) = 0.98

VARIANZ

Grundlagen

Varianz ist die ***Summe der quadierten Abweichungen***, ***dividiert durch die Anzahl der Merkmalsträger***. Der Unterschied zur mittleren absoluten Abweichung ist das Quadrat. Somit ***keine negativen Werte.***

Formel

$\sigma^2 = \frac{1}{n} \sum_{i=1}^{v} (x_i - \bar{x})^2 h_i$ $\bar{x} = \frac{1}{n} \sum_{i=1}^{v} x_i \cdot h_i$

STANDARDABWEICHUNG

Grundlagen

Die Standardabweichung ist die ***Quadratwurzel aus der Varianz***. Die Berchnungen werden meistens mit der Standardabweichungen gerechnet.

Formel

$$\sigma = \sqrt{Var(X)} = \underbrace{\sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}}}_{\text{Standardabweichung der Grundgesamtheit}}$$

Beispiel

x_i	h_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 h_i$
0	3	4.16	12.48
1	10	1.08	10.80
2	4	0.0016	0.0064
3	3	0.92	2.76
4	2	3.85	7.7
12	1	99.2	99.2
	23		132.95

$\sigma^2 = \frac{1}{n} \sum_{i=1}^{v} (x_i - \bar{x})^2 h_i = 5.78$

$\sigma = \mp 2.40$

VARIATIONSKOEFFIZIENT

Grundlagen

Der Variationskoeffizient misst **die relative Streuung**. Er setzt die Streung in Relation zu Lage der Häufigkeitsverteilung.
Er ist der Quotient aus Standardabweichung und arithmetischem Mittel multipliziert mit 100.

Formel

$VK = \frac{\sigma}{\bar{x}} \cdot 100$

Beispiel

Beispiel 1

Die Preisuntersuchung für die Güter A und B hat zu folgenden Ergebnissen geführt:

$\bar{x}_A = 7\ CHF, \quad \sigma_A = 2.80\ CHF$
$\bar{x}_B = 750\ CHF, \quad \sigma_B = 20.40\ CHF$

$VK_A = \frac{2.80}{7} \cdot 100 = 40\%,$

$VK_B = \frac{20.40}{750} \cdot 100 = 2.72\%,$

- Die Streuung der Preise für Gut B ist also relativ geringer als die für Gut A.
- Eine weitere Interpretation ist nicht zulässig
- würde statt der Standardabweichung die mittlere absolute Abweichung verwendet, so könnte man sagen, um welchen Faktor sich die mittleren relative Abweichungen unterscheiden

Beispiel 2

Vergleich der Streuung der Leistungen eines Weitspringer W und eines Langstreckenläufer L (Achtung: **unterschiedliche Dimension** daher ist die Verwendung der **absoluten Streuung nicht erlaubt**)

$\bar{x}_W = 7.20\ m, \quad \sigma_W = 0.24\ m$
$\bar{x}_L = 29.4\ min, \quad \sigma_L = 0.89\ min$

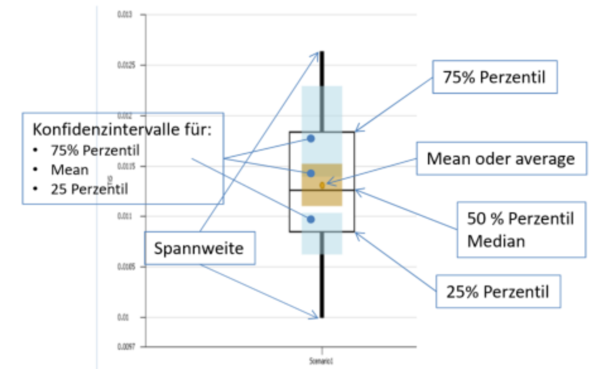
$VK_W = \frac{0.24}{7.20} \cdot 100 = 3.3\ \%,$

$VK_B = \frac{0.89}{29.40} \cdot 100 = 3.0\%,$

Der Langstreckenläufer und der Weitspringer erbringen- relativ gesehen nahezu gleichmäßige Leistungen.

Ein Vergleich der absoluten Streuung ist wegen der unterschiedlichen Dimension der Merkmale nicht möglich!

Überblick - Boxplot



Boxplot = Darstellung von Mittelwerten und Streuungsmassen um einen schnellen Einblick zu bekommen, schnelle Übersicht wo die Daten liegen und wie diese verteilt sind.

ZEITREIHEN, REGRESSION, KORRELATION

MOTIVATION

1. Bei Experimenten entstehen Tupel, ein Faktorwert(control) und ein Ergebniswert(response).
2. Mit diesen Tupel(Punkten) kann man in ein Streungsdiagramm (Punktwolke) erstellen
3. Nun wollen wir eine Funktion finden, welche diese Punkte bestmöglich annähert.
4. Achtung! Eine Extrpolation kann gefährlich sein, wenn man zu wenig Werte hat, kann es falsche Funktion liefern.

x	1	3	6
y	0	2.5	5

REGRESSIONSANALYSE

Grundlagen

- Man will einen Zusammehnag zwischen einer unabhängigen Grösse(x-Werte) und abhängigen Werten(y-Werte) anhand von Einzelmessungen modellieren.
- Die Form bzw. Tendenz des Zusammenhangs durch eine mathematische Funktion(Regressionsfunktion) beschreiben

Lineare Regression

In der linearen Regressionsanalyse wird die Tendent ***durch die Funktion(Regressionsgerade)*** bestimmt***:***

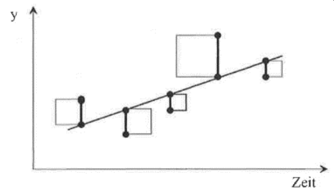
- ***Geradengleichung: y =a + b·x*** //y wird duch Parameter bestimmt(Standard)
- ***Proportionalität: y = c·x*** //bei unbekannter Abhängigkeit zusätzlich

Ansatz der kleinsten Quadrate

1) partielles Ableiten des Ausdrucks nach a und nach b
2) nullsetzen der beiden partiellen Ableitungen
3) auflösen der beiden Gleichungen nach a und b

Ergibt für: $\mathbf{y = a_1 + b_1\, x}$:

$$\boldsymbol{a_1 = \bar{y} - b_1 \bar{x}} \quad \textbf{und} \quad \boldsymbol{b_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}}$$



Analog folgt für: $\mathbf{x = a_2 - b_2\, y}$

$$a_2 = \bar{x} - b_2\bar{y} \quad \text{und} \quad b_2 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2}$$

Beispiel

Student muss 6h arbeiten. Wie viel Zeit hat er für das Studium?

Student	A	B	C	D	E	F	G	H	I	J	K	L
Erwerb.	1	2	2	3	3	4	5	6	8	12	15	33
Studium	39	37	36	40	36	37	34	36	33	33	32	27

$$a_1 = \bar{y} - b_1\bar{x} \quad \text{und} \quad b_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

Stdent	xi	yi	xi yi	xi^2	yi^2
A	1	39	39	1	1521
B	2	37	74	4	1369
C	2	36	72	4	1296
D	3	40	120	9	1600
E	3	36	108	9	1296
F	4	37	148	16	1369
G	5	34	170	25	1156
H	6	36	216	36	1296
I	8	33	264	64	1089
J	12	33	396	144	1089
K	15	32	480	225	1024
L	23	27	621	529	729
Summe	84	420	2708	1066	14834

Werte einsetzen:

$$b_1 = \frac{2708 - 12 \cdot \frac{84}{12} \cdot \frac{420}{12}}{1066 - 12 \cdot \left(\frac{84}{12}\right)^2} \rightarrow \mathbf{b_1 = -0.49}$$

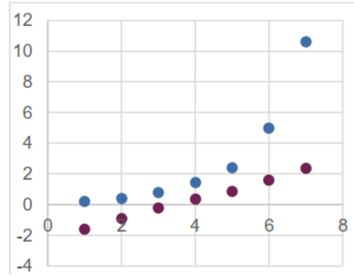
$$a_1 = \frac{420}{12} - (-0.49) \cdot \frac{84}{12} \rightarrow \mathbf{a_1 = 38.43}$$

y = 38.43 - 0.49 · x mit y(6) = <u>35.5 h</u> → <u>der Student kann 35.5h für das Studium aufwenden, wenn er 6 Stunden(y=6) arbeitet.</u>

Linearisierung

Allgemeine Exponentialfunktion: y = a e^x
Allgemeine Logarithmusfunktion: y = a + b·lnx
Wenn der Verlauf Exponentiell und nicht linear ist müssen die ***y-Werte logarithmiert*** werden.
Dann wie oben weiterfahren und am Schluss Logarithmieren und Kehrwert bilden:

$$\ln(f(x)) = -2.11 + 0.621x \;\rightarrow f(x) = e^{-2.11+0.621x} = \frac{e^{0.621x}}{e^{2.11}}$$



Newton-Algorithmus

Polynome: $y = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0$

Ansatz: $f(x) = a_0 + a_1(x-x_1) + a_2(x-x_1)(x-x_2) + \ldots + a_n(x-x_1)(x-x_2)\cdots(x-x_n)$

Einsetzen von x_1 in die Funktion f(x) ergibt:

- $y_1 = f(x_1) = a_0 \qquad \Rightarrow a_0 = y_1$
- $y_2 = f(x2) = a_0 + a_1(x_2 - x_1) \qquad \Rightarrow a_1 = \frac{y_2 - y_1}{x_2 - x_1} = D_{2,1}$
- $y_3 = f(x_3) = a_0 + a_1(x_3 - x_1) + a_2(x_3 - x_1)(x_3 - x_2) \qquad \Rightarrow a_2 = \frac{y_3 - a_0 - a_1(x_3 - x_1)}{(x_3 - x_1)(x_3 - x_2)}$

durch umformen folgt: $a_2 = \frac{1}{(x_3 - x_1)}\left(\frac{y_3 - y_2}{(x_3 - x_2)} - \frac{y_2 - y_1}{(x_2 - x_1)}\right)$

$$a_2 = \frac{D_{3.2} - D_{2.1}}{(x_3 - x_1)} = D_{3,2,1}$$

Koeffizienten können iterativ bestimmt werden:

Schema Newton-Algorithmus

x_k	y_k	
x_1	y_1	a_0
x_2	y_2	$D_{2,1} = \frac{y_2 - y_1}{x_2 - x_1} = a_1$
x_3	y_3	$D_{3,2} = \frac{y_3 - y_2}{x_3 - x_2} \rightarrow D_{3,2,1} = \frac{D_{3,2} - D_{2,1}}{x_3 - x_1} = a_2$
x_4	y_4	$D_{4,3} = \frac{y_4 - y_3}{x_4 - x_3} \rightarrow D_{4,3,2,} = \frac{D_{4,3} - D_{3,2}}{x_4 - x_2} \rightarrow D_{4,3,2,1} = \frac{D_{4,3,2} - D_{3,2,1}}{x_4 - x_1} = a_3$
x_5	y_5	$D_{5,4} = \frac{y_5 - y_4}{x_5 - x_4} \rightarrow D_{5,4,3,} = \frac{D_{5,4} - D_{4,3}}{x_5 - x_3} \rightarrow D_{5,4,3,2} = \frac{D_{5,4,3} - D_{4,3,2}}{x_5 - x_{2,}} \rightarrow D_{5,4,3,2,1} = \frac{D_{5,4,3,2} - D_{4,3,2,1}}{x_5 - x_1} = a_4$

Beispiel

Warum nicht bechtet?

k	x_k	y_k	
1	1	6	$a_0 = 6$
2	2	11	$D_{2,1} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{11-6}{2-1} = a_1 = 5$
3	3	18	$D_{3,2} = \frac{y_3 - y_2}{x_3 - x_2} \rightarrow D_{3,2,1} = \frac{D_{3,2} - D_{2,1}}{x_3 - x_1} = a_2$
			$D_{3,2} = \frac{18 - 11}{2 - 1} \rightarrow D_{3,2,1} = \frac{7 - 5}{3 - 1} = a_2 = 1$
4	4	27	Warum nicht betrachtet?

Mit:

$f(x) = a_0 + a_1(x-x_1) + a_2(x-x_1)(x-x_2) + \ldots + a_n(x-x_1)(x-x_2)\cdots(x-x_n)$

$f(x) = 6 + 5(x-1) + 1(x-1)(x-2) = 6 + 5x - 5 + x^2 - 2x - x + 2$

$\rightarrow f(x) = x^2 + 2x + 3$

ZEITREIHE

Definition

Zusammenhang zwischen Merkmalsträger x und Zeitpunkten ti.
Zeitreihe ist ***zeitlich geordnete Folge von Merkmalswerten***

- Kosten/Gewinnentwicklung
- Verbrauch von Ressourcen
- Auftragseingang
- Bearbeitungs- oder Durchlaufzeiten

Aufgaben/Ziel

erkennen der Struktur und der Gesetzmäßigkeiten einer Zeitreihe
Mittel zu Prognose - ***erkennen eines Trends***

Trend

Blick in die Zukunft → beschreibt **langfristige Grundrichtung** einer Zeitreihe
Um ihn streuen die Zeitreihenwerte im Zeitablauf

Probleme

- periodische Schwankungen (Schweinezyklus)
- Besondere Ereignisse → Ausreisser(unvorhersehbar)

→ ***Lösung*** durch sogenanntes ***Glätten*** des Verlaufs mittels linearer Regression

Zusammenhang

Das ***Erkennen des Zusammenhangs*** zwischen zwei oder mehr Merkmalen nennt man ***Korrelation***:
- Kraftstoffart – Leistung - Autounfälle – Alter -Lernaufwand – Lernerfolg
Dabei Untersucht man die Merkmale x & . Dabei intressiert man sich für folgende Fragen:

a. besteht ein Zusammenhang zwischen X und Y?
b. von welcher Form ist der Zusammenhang?
c. von welcher Stärke (Intensität) ist der Zusammenhang?

Abhängigkeit

- Zwei Merkmale sind statistisch ***unabhängig***, ***wenn* der Wert des einen Merkmals nicht davon abhängt, welchen Wert das andere Merkmal besitzt**, sonst sind sie abhängig
- **Formale Abhängigkeit - *zahlenmäßig begründete Abhängigkeit***
- **Sachliche Abhängigkeit –** ist der Wert des einen Merkmales ***ursächlich***(***kausal***) für den Wert des anderen?(Ursache – Wirkung kann nicht bestimmt werden)

Korrelationsanalyse

Stärke (Intensität) des Zusammenhangs feststellen → wie stark ist der Einfluss von x auf y

Korrelationskoeffizient von Bravais Pearson

1): ***Kovarianz → Streuung der Merkmalsträger***(kombinationen) um den Mittelpunkt/Durschnitt
Berechnung analog Varianz:

$$\sigma_{XY} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \quad oder \quad \sigma_{XY} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$$

2.) Normierung der Kovarianz es entsteht der **Korrelationskoeffizient r** zu:

$$r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad mit \quad |\sigma_{XY}| \leq \sigma_X \sigma_Y \quad \frac{Kovarianz}{Varianz} = Korrelationskoeffizient$$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum(x_i - \bar{x})^2\right)\left(\sum(y_i - \bar{y})^2\right)}}$$

Daraus folgt:

- → ***Positivem r*** → der lineare Zusammenhang der Merkmale X und Y ist ***gleichläufig***
- → ***Negativem r*** → der lineare Zusammenhang der Merkmale X und Y ist ***gegenläufig***
- → 0: kein **linearer** Zusammenhang
- → ±1: sehr starker **linearer** Zusammenhang

WAHRSCHEINLICHKEIT

Zufallsexperimente

Ein Zufallsexperiment ist ein Vorgang, der als ***beliebig oft wiederholbar*** angesehen werden kann und dessen Endzustand oder ***Ergebnis vom Zufall abhängt***.

Wahrscheinlichkeitsräume

- ***Ergebnismenge*** Ω → alle mögliche Ergebnisse des Experiments
- ***System*** $\mathcal{A}$ ***der Ereignisse*** → Elemente Teilmenge von Ω
- ***Wahrscheinlichkeitsmass P*** → jedem Ereignis A wird Wahrscheinlichkeit P(a) zugeordnet

System der Ereignisse

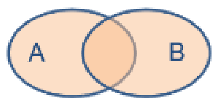
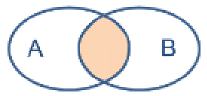
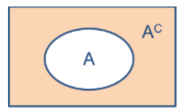
weisst gewisse ***Eigenschaften*** auf, die wir aus der ***Mengenlehre*** kennen:

- dass beide Ereignisse zugleich **$A \cap B$**
- zumindest eines von beiden Ereignissen **$A \cup B$**
- das Gegenereignis (komplementäre) zu einem Ereignis **A^C**

Es seien A, B $\subset \Omega$ zwei Ereignisse die ein System $\mathcal{A}$ der Ereignisse bilden.

Das Ereignis „A oder B" tritt ein, wenn ein Ergebnis $\omega \in A \cup B$ auftritt.

Das Ereignis „A und B" tritt ein, wenn ein Ergebnis $\omega \in A \cap B$ auftritt.

A oder B	*A und B*	*A komplementär*
		

Speziell gilt:
Die ***leere Menge*** $\emptyset$ bildet das ***unmögliche Ereignis***.
Die ***Ergebnismenge*** Ω bildet das ***sichere Ereignis***
Ein-Elementige Teilmengen der Ergebnismenge Ω werden als ***Elementarereignisse*** bezeichnet.
Zwei Ereignisse A und B heißen ***unvereinbar***, wenn $A \cap B = \emptyset$

Es sei Ω eine nichtleere Menge und $A_i \subset \Omega$ sei ein Mengensystem.
Das Mengensystem $\mathcal{A}$ bildet eine σ-Algebra, falls:

(1) $\Omega \in \mathcal{A}$.

(2) Für $A \in \mathcal{A}$ gilt auch $A^C \in \mathcal{A}$.

(3) Aus $A_i \in \mathcal{A}$ für alle $i \in \mathbb{N}$ folgt $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Die leere Menge ist ein Ereignis.
Endliche Vereinigungen, endliche Durchschnitte und Differenzen sind Ereignisse.
Auf der Menge der Ergebnisse eines Experiments werden Teilmengen definiert, jede Teilmenge stellt ein Ereignis dar und ihr kann eine Wahrscheinlichkeit p zugeordnet werden.

Wahrscheinlichkeitsmass

Ist Ω die Ergebnismenge eines Zufallsexperiments und
ist $\mathcal{A}$ eine σ-Algebra von Ereignissen über Ω,
so ist die Abbildung

$P : \mathcal{A} \to [0,1]$ Wahrscheinlichkeit von 0-1

ein Wahrscheinlichkeitsmaß, falls

(1) $P(A) \geq 0$ für alle $A \in \mathcal{A}$

(2) $P(\Omega) = 1$

(3) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ für paarweise unvereinbare Ereignisse $A_1, A_2, \ldots \in \mathcal{A}$.

Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum sowie $A, B \in \mathcal{A}$.

(1) $P(A^C) = 1 - P(A)$ (Gegenwahrscheinlichkeit)

(2) $P(\emptyset) = 0$ (Unmögliches Ereignis)

(3) $0 \leq P(A) \leq 1$ (Wertebereich)

(4) $A \subset B \Rightarrow P(A) \leq P(B)$ (Monotonieeigenschaft)

(5) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Additionssatz)

(6) $P(A \cup B) = P(A) + P(B)$, falls A und B unvereinbar (Additionssatz)

Laplace-Experiment

Ist die ***Ergebnismenge*** $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ ***endlich*** und wird jedem Elementarereignis aus Ω ***die gleiche Wahrscheinlichkeit,*** $P(\{\omega_i\}) = \frac{1}{n}$ zugeordnet, so gilt für die Wahrscheinlichkeit eines beliebigen Ereignisses:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der für } A \text{ günstigen Ergebnisse}}{\text{Anzahl aller möglichen Elementarereignisse}}$$

Beispiele

Werfen eines Würfels mit $\Omega = \{1, 2, \ldots, 6\}$ und $P(\{\omega_i\}) = 1/6$.
Werfen einer Münze mit $\Omega = \{\{\text{Wappen}\}, \{\text{Zahl}\}\}$ und $P(\{\text{Wappen}\}) = P(\{\text{Zahl}\}) = 1/2$.

Omega Algebra

Eine Omega Algebra lieht vor, wenn zum System der Ereignisse noch die leere Menge und der Ergebnisraum zugeordnet werden und ferner die Regeln von Kolmogorov gelten.

KOMBINATORIK UND WAHRSCHEINLICHKEIT

KOMBINATORIK – REGELN FÜR DAS ZÄHLPRINZIP

Grundsatz

- ***Mengen → Reihenfolge unwichtig***
- ***Tupel → Reihenfolge wichtig***

Zählprinzip

Gibt es in einem n-Tupel für die Besetzung

- der ersten Stelle **k_1** Möglichkeiten
- der zweiten Stelle **k_2** Möglichkeiten
- der dritten Stelle **k_3** Möglichkeiten
- der n-ten Stelle **k_n** Möglichkeiten

dann gibt es insgesamt $k_1 \cdot k_2 \cdot k_3 \cdot \ldots \cdot k_n$ verschiedene n-Tupel

Beispiel

Beispiele sind Pferde oder Autorennen, da es bei jeder Stelle ein Auto weniger hat:
Die Reihenfolge der ersten 5 Plätze eines Pferderennen mit 18 gestarteten Pferden:
18 • 17 • 16 • 15 • 14 = <u>1.028.160</u> → entspricht Geordnete Probe ohne zurücklegen

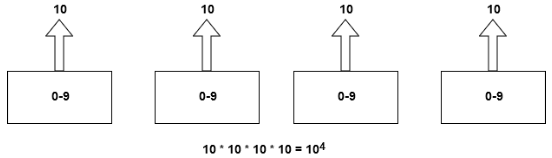
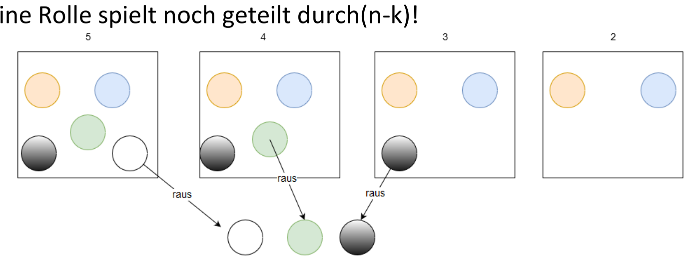
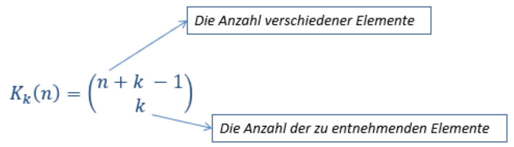
Urnenmodell der Kombinatorik

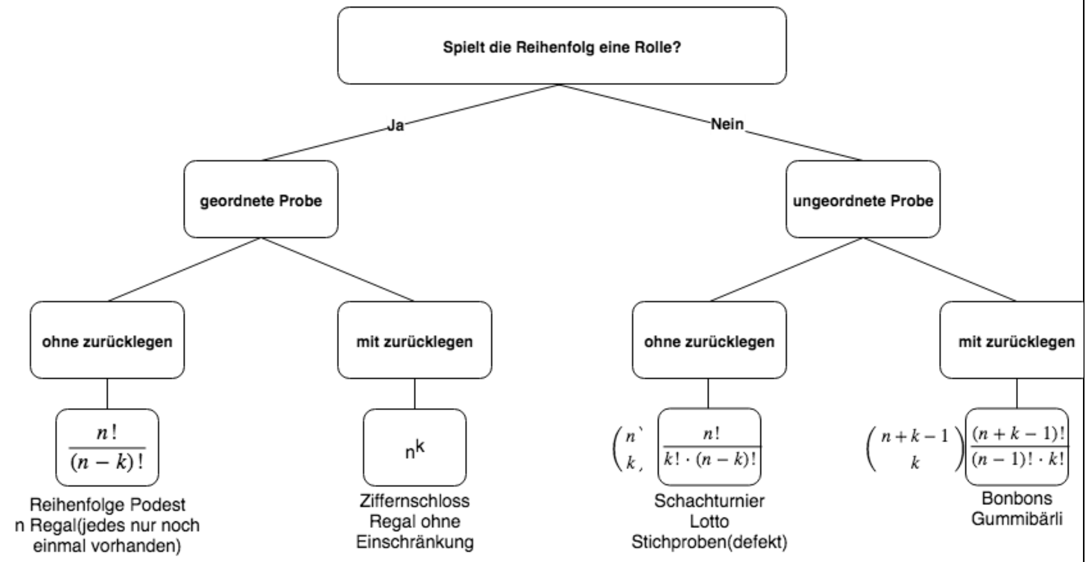
In der Kombinatorik wird viel mit einem Urnenmodell gearbeitet. Es werden daraus beliebig viele Kugeln gezogen. Dabei gibt es folgende Unterscheidungen:
- Die Kugeln werden zurückgelegt(mit Wiederholungen) oder nicht(ohne Wiederholungen)
- Die Reihenfolge wird beachtet(geordnet) oder nicht(ungeordnet)

Bei der ungeordneten Kugel Ziegung entspricht (1;2) → (2;1), etc.

	geordnet	ungeordnet
Mit zurücklegen	(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3),	(1,1), (1,2), (1,3), (2,2), (2,3), (3,3),
Ohne zurücklegen	(1,2), (1,3), (2,1), (2,3), (3,1), (3,2),	(1,2), (1,3), (2,3),

Ziehung von k Kugeln aus einer Urne mit n = 3 Kugeln

Geordnete Proben ***Für Tupel*** ***Reihenfolge wichtig***	Bei geordneten Proben wird auf die Reihenfolge geachtet. Es gibt immer eine Menge m mit ***n Elementen***. Bilden wir nun ***k-Tupel, deren Einträge aus der Menge M*** stammen, so zwei Möglichkeiten: ***1.) Mit Wiederholung*** Es kann jedes Mal wieder die gleiche Zahl(oder gleiches Element) vorkommen  n^k ***2.) Ohne Wiederholen(ohne zurücklegen)*** Bei jedem entnehmen gibt es eine Kugel weniger(n!-Möglichkeiten), da die Reihenfolge eine Rolle spielt noch geteilt durch(n-k)!  $\frac{n!}{(n-k)!}$ Mehr Ergebnisse da Reihenfolge wichtig: 
Permutationen	Die Zahl der Permutationen einer n-Menge ist ***n!*** Das ist ein Spezialfall von einer geordneten Probe ohne Wiederholung(n = k) Beispiel: Festlegung der Sitzordnung von 10 Gästen ist: 10! = 3628800
Permutation vs Kombinatorik	Permutation Mass für alle Kombinationsmöglichkeiten. Kombinatorik Regel für Berechnung der Möglichkeiten, wenn Bedingungen die Möglichkeiten einschränken
Ungeordnete Proben ***für Mengen*** ***Reihenfolge unwichtig***	Bei der ungeordneten Proben wird nicht auf die Reihenfolge geachtet. Es gibt ebenfalls eine Menge m mit n Elementen. Es gibt wieder die 2 Fälle: ***1.) Mit Wiederholung*** Sind aus ***n verschiedenen Elementen k Elemente mit Wiederholung*** auszuwählen und ist die ***Anordnung ohne Bedeutung***. $\frac{(n+k-1)!}{(n-1)!\cdot k!}$  ***2.) Ohne Wiederholung(ohne zurücklegen)*** ***Die Anzahl der k-elementigen Teilmengen aus einer n-elementigen Menge.*** ***Gleich wie bei der geordneten Probe, jedoch da die Ordnung keine Rolle spielt muss man noch durch k! teilen.*** $\binom{n}{k}$ $\frac{n!}{k!\cdot(n-k)!}$ ***Anzahl aller möglichen Teilmengen k aus n.*** k: Anzahl der zu entnehmenden Elemente ***Anzahl aller möglicher Element-Kombinationen mit Wiederholung*** Da die Reigenfolge unwichtig ist gibt es weniger Ergebnisse: 

Übersicht	
Beispiel Geordnete Probe ohne Wiederholung	Man 10 verschiedene Fruchtsorten und will jeweils eine Anordnung mit 4 machen. Dabei ist jede Sorte nur einmal vorhanden. Wie viele Kombinationen gibt es? $\frac{n!}{(n-k)!} = \frac{10!}{(10-4)!} = 5040\ verschiedene\ Anordnungen$
Beispiel Geordnete Probe mit Wiederholungen	Man hat wieder 10 verschiedene Früchte. Man will genau 7 Früchte aufstellen(also kann auch die gleiche Frucht mehrmals vorkommen): $n^k = 10^7$ = 10'000'000 Kombinationen
Beispiel Geordnete Probe mit Wiederholungen Schwierig	Man hat nun verschiedene Früchte einmal und eine Frucht zwei Mal. Nun will man alle Elemente anordnen: $\frac{n!}{(n-k)!} : doppelte\ Elemente = \frac{6!}{(6-4!)} : 2 = 360 : 2 = 180$ //geteilt durch zwei, weil ein Element zwei Mal vorkommt→ Es können nur halb soviel Kombinationen auftreten.(Wäre das Element 3x vorgekommen → :3)
Beispiel Ungeordnete Proben ohne Wiederholung	Aus einer 30 köpfigen Schulklasse gewinnen 4 Schüler einen Preis. Wie viele Auswahlmöglichkeiten gibt es? $\binom{n}{k} = \binom{30}{4} = \frac{30!}{4! * (30-4)!} = 27405$
Beispiel Schachturnier	Klasse mit 30 Schüler möchte Spiel spielen, wobei alle gegen alle einmal spielen: $\frac{30 * 29}{2!} = 435\ Spiele(bei\ K.O -\!\!\to 30\ Spiele)$
Beispiel Ungeordnete Probe mit Wiederholung	Eine Firma stellt Gummibärli mit fünf verschiedenen Geschmackssorten her. Die Tüten, welche Sie verkaufen beinhalten jeweils 12 Gummibärli. Wie viele verschiedene Gummibärlimischungen gibt es? $\binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)! * k!} = \binom{5+12-1}{12} = \frac{(5+12-1)!}{(5-1)! * 12!} = 1820$ //beachte n ist die Anzahl verschiedener Gummibärli und k die Anzahl, welche in einen Pack hineingefüllt wird.
Beispiel Ungeordnete Probe mit Wiederholung	Die 6 grossen Parteien in der Schweiz kämpfen um 3 Sitze im Nationalrat. Wie viele Kombis gibt's $\binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)! * k!} == \frac{(6+3-1)!}{(6-1)! * 3!} = 56\ Kombinationen$
Beispiel Stichproben	Eine Lieferung besteht aus 30 Glühbirnen, es werden jeweils 3 ohne zurücklegen entnommen. Wie viele Strichproben sind möglich? (30!)/(3! *(30-3)!) =4060 Proben Wie gross ist die Wahrscheinlichkeit, dass in einer Strichprobe genau 2 defekte Glühbirnen enthalten sind wenn von 30 genau 6 defekt sind= P(2 defekte) = $\frac{\binom{Gesamtzahl\ defekte\ Birnen}{Anzahl\ defekte\ Birnen\ soll(pro\ Ziehung)} * \binom{Gesamtzahl\ gute\ Birnen}{Anzahl\ gute\ Birnen(pro\ Ziehung)}}{Anzahl\ Gesamtproben} = \frac{\binom{6}{2}\binom{24}{1}}{4060} = \frac{\frac{6!}{2! * (6-2)!} * \frac{24!}{1! * (24-1)!}}{4060} = 0.0887$
Beispiel Lotto	Wie gross ist es im Lotto(5 aus 45) genau zwei richtige zu erhalten? P(2 richtige) = $\frac{\binom{Anzahl\ gezogene\ Kugeln}{Anzahl\ treffende\ Kugeln\ soll(pro\ Ziehung)} * \binom{Gesamtzahl\ nichtgezogene\ Kugeln}{Anzahl\ nicht-Treffer\ Kugel(pro\ Ziehung)}}{Gesamtanzahl\ mögliche\ Ziehungen} = \frac{\binom{5}{2}\binom{40}{3}}{\binom{45}{5}} = \frac{\frac{5!}{2! * (5-2)!} * \frac{40!}{3! * (40-3)!}}{\frac{45!}{5! * (45-5)!}} = 0.081$

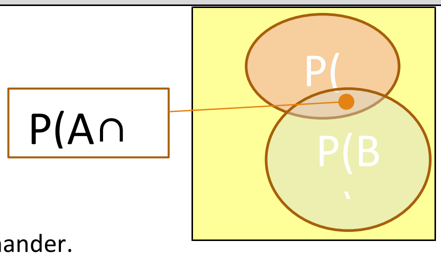
BEDINGTE WAHRSCHEINLICHKEIT

Grundlagen

Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum.
Sind $A, B \in \mathcal{A}$ Ereignisse und gilt $P(B) > 0$, so heißt

$$\boldsymbol{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

Die bediente Wahrscheinlichkeit von A unter der Bedingung B.



Man vergleicht also zwei Wahrscheinlichkeiten miteinander.

Beispiel Impfungen

Versucht, ob eine Impfung gegen eine Erkrankung helfen kann:

	A	$\bar{A}$	Summe
I	87	5.671	5.758
$\bar{I}$	312	4.390	4.702
Summe	399	10.061	10.460

Wahrscheinlichkeit einer Erkrankung

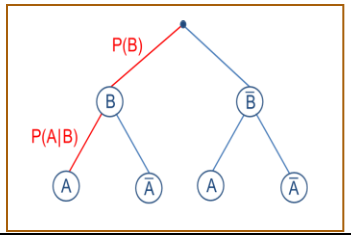
- ohne Impfung
  $P(A|\bar{I}) = \frac{|A \cap \bar{I}|}{|\bar{I}|} = \frac{312}{4702} = 6.6\%$
- mit Impfung
  $P(A|I) = \frac{|A \cap I|}{|I|} = \frac{87}{5758} = 1.5\%$

- Die Erkrankung hängt also davon ab von der Bedingung «geimpft» oder «nicht geimpft»

Bemerkung

Die bedingte Wahrscheinlichkeit tritt bei mehrstufigen Experimenten auf, wenn man in den einzelnen Stufen Ereignisse protokolliert.

- Die Pfadregel besagt für den rot gezeichneten Pfad
  $\boldsymbol{P(A \cap B) = P(B) \cdot P(A|B)}$



Vorgehen

Unter den 20 Schülern einer Klasse sind 4 Raucher. Von den 12 männlichen Schülern sind 3 Raucher. Gesucht wird ein neuer Klassensprecher.
Wie groß ist der Anteil der Männer unter der Bedingung, dass es sich um einen Raucher handelt?

1- Bezeichnungen auswählen

R: „Der ausgewählte Schüler ist Raucher."
M: „Der ausgewählte Schüler ist männlich."
¯R: „Der ausgewählte Schüler ist Nichtraucher."
¯M: „Der ausgewählte Schüler ist weiblich."

2- Gesucht definieren: Gesucht ist $P(R|M)$
3- Tabelle ausfüllen:
4- Wahrscheinlichkeit rauchender Mann berechnen

	M	$\bar{M}$	
R	3	1	4
$\bar{R}$	9	7	16
	12	8	20

$P(R|M) = \frac{|R \cap M|}{|R|} = \frac{3}{4} = 0.75$

UNABHÄNGIGE WAHRSCHEINLICHKEIT

Grundlagen

Zwei Ereignisse A und B heißen stochastisch unabhängig, falls P(A\B) = P(A) gilt:

$$\boldsymbol{P(A \cap B) = P(A) \cdot P(B)}$$

P(A) ist nicht von P(B) abhängig.

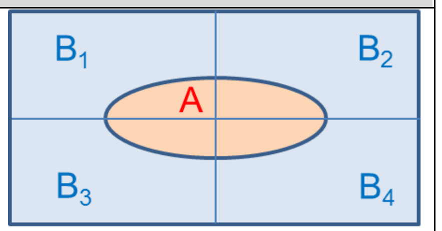
- Wenn also keine bedingte Wahrscheinlichkeit vorliegt, dann eine unabhängige.

VOLLSTÄNDIGE WAHRSCHEINLICHKEIT

Grundlagen

Seien $B_1,\ldots,B_n$ paarweise disjunkte Ereignisse, dann gilt:
$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$

Die Wahrscheinlichkeit A ist die Wahrscheinlichkeit A in Abhängigkeit Bi multipliziert mit der Wahrscheinlichkeit Bi
Wobei i der Index für die Anzahl disjunkten Ereignisse ist.



Rechenregeln allgemein

$P(A \cap B) = P(B) \cdot P(A|B)$

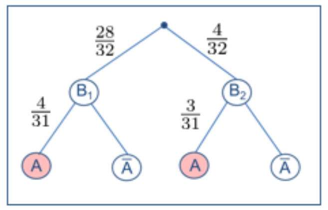
$P(B \cap A) = P(A) \cdot P(B|A)$

$P(A \cap B) = P(B \cap A)$

$P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

Beispiel Karten

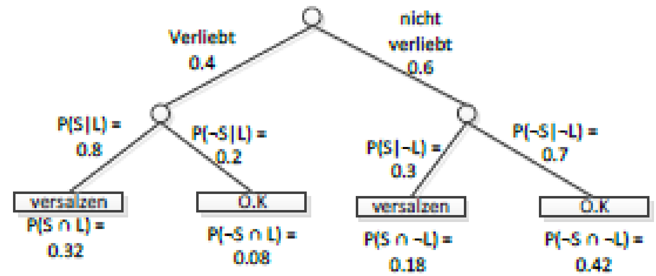
Aus einem Skatblatt (32 Karten, davon 4 Asse) werden nacheinander zwei Karten gezogen, ohne dass die erste Karte wieder zurückgelegt wird.
Wie groß ist die Wahrscheinlichkeit beim zweiten Zug ein Ass zu ziehen?



$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)$$
$$= \frac{4}{31} \cdot \frac{28}{32} + \frac{3}{31} \cdot \frac{4}{32} = \frac{1}{8}$$

Beispiel

Ein Koch- Lehrling versalzt seine Suppe mit einer Wahrscheinlichkeit von 0.5. Wenn er verliebt ist –ein Zustand, in dem er sich mit Wahrscheinlichkeit 0.4 befindet – ist es ganz schlimm: Dann versalzt er nämlich 80 % seiner Suppen. §



P(S) allgemein = 0.5 P(L) = 0.4 → P(¬L) = 0.6
P(S|L) = 0.8 → P(S|¬L) = 0.2 → P(S∩L) = 0.8 * 0.4 = 0.32 → P(S∩¬L) = 0.2 * 0.4 = 0.08
P(S) = P(S∩L) + P(S∩¬L) → 0.5 = 0.32 + P(S∩¬L) → P(S∩¬L) = 0.5 – 0.32 = 0.18
P(¬S∩¬L) = 1 - P(S∩L) - P(¬S∩L) - P(S∩¬L) = 1 – 0.32 – 0.08 – 0.18 = 0.42

ZUFALLSVARIABLE UND DIE WICHTIGSTEN THEORETISCHEN VERTEILUNGEN

ZUFALLSVARIABELN

Grundlagen

Zufallsvariable ist:

- Keine Variable
- Ist nicht zufällig

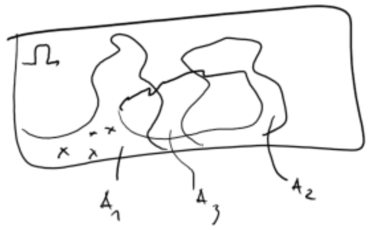
Der Name ist völlig irreführend:

Die Zufallsvariable X ist eine Abbildung/Funktion die jedem Ergebnis ω eines Zufallsexperimentes einen Wert(aus dem Ereignisraum Ω) zuordnet.

Für das gilt:

$$A_I = \{\omega \in \Omega \mid X(\omega) \in I\} \subset \Omega$$

$$P(X \in I)$$



Jede Menge Ai muss eine gewisse Wahrscheinlichkeit beinhalten:

Beispiele

$$X : \{\text{Wappen, Zahl}\} \to \mathbb{R}\,, \quad X(\omega) = \begin{cases} 1 & , \text{falls } \omega = \text{Wappen} \\ -1 & , \text{falls } \omega = \text{Zahl} \end{cases}$$

Ein Roulett Spieler setzt einen Chip «Ereignis 1. Dutzend»:

Wenn die Zahlen 1-12 kommen, gibt es 3 Geldstücke, ansonsten wird das gesetzte weggenommen:

Der Reingewinn ist eine Funktion des bei der Ausspielung auftretenden Ergebnisses ω:

$$\underset{\text{Ergebnismenge}}{X} : \{0,1,\ldots,36\} \to \mathbb{R} \quad \text{mit} \quad X(\omega) = \begin{cases} 3 & , \text{für } \omega \in \{1,2,\ldots,12\} \\ -1 & , \text{für } \underset{\text{Ergebnis}}{\omega} \in \underset{\text{Ereigniss}}{\{0,13,14,\ldots,36\}} \end{cases}$$

Jede Gleichung X(ω) = x mit x ∈ {-1,3}
führt eindeutig auf ein Ereignis (eine Teilmenge von Ω).

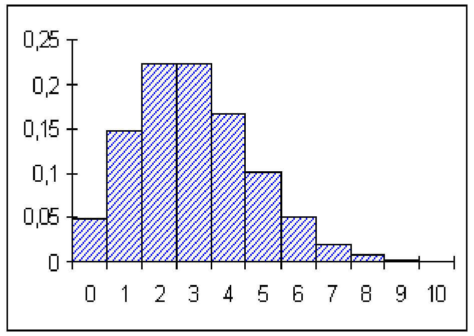
DISKRETE WAHRSCHEINLICHKEITSFUNKTION

Grundlagen

Wie der Name bereits verrät, weist diese Funktion beliebigen Werten die Wahrscheinlichkeit zu.

- Die Summe aller Einzelwahrscheinlichkeit ist 1, entspricht der Fläche
- Die Wahrscheinlichkeit P, dass ein Ereignis xi eintritt kann unmittelbar abgelesen werden
- Die Wahrscheinlichkeit eines Intervalles xi, xi+1,.. xn ergibt sich:

$$P(x_i \le X \le x_n) = \sum_{j=i}^{n} f(x_j)$$



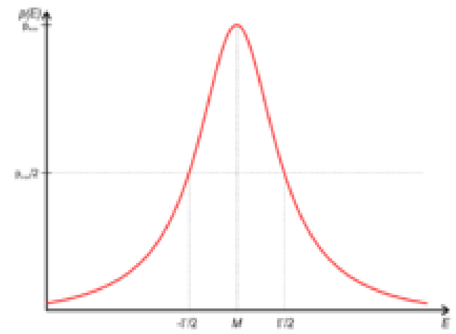
KONTINUIERLICHE VERTEILUNGSDICHTEFUNKTION

Grundlagen

Zeigt die Verteilung der Wahrscheinlichkeiten

- Die Fläche unter der Kurve entsprich 1, Summe aller Wahrscheinlichkeiten(Integral)
- Es kann nur die Wahrscheinlichkeit eines Bereiches berechnet werden(Weil nur Fläche(Integral) bestimmbar:

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$



Einfache Mittelwert

Der Mittelwert einer Verteilungsfunktion kann berechnet werden:

$$E[X] = \mu = \sum_i x_i f(x_i) \quad oder = \int_{-\infty}^{\infty} x f(x) dx$$

Beispiel

Augensumme eines Würfels mit zwei Würfen:

Gesamtanzahl n an Möglichkeiten: 6*6 = 36

Augensumme

xi	2	3	4	5	6	7	8	9	10	11	12	Σ
Pi	0.028	0.056	0.083	0.111	0.139	0.167	0.139	0.111	0.083	0.056	0.028	1
xi*pi	0.056	0.167	0.333	0.556	0.833	1.167	1.111	1.000	0.833	0.611	0.333	7

f(xi)|⇒ xi →pi

$E[X]$ oder μ

Varianz(Streuwert)

Auch der Streuwert – die Varianz – kann berechnet werden:

$$\sigma^2 = \sum_{i=1}^{v} (x_i - \bar{x})^2 f_i \quad \text{oder}$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x_i - \bar{x})^2 f_i \, dx$$

Beispiel

Augensumme eines Würfels mit zwei Würfen:

Gesamtanzahl n an Möglichkeiten: 6*6 = 36

xi	2	3	4	5	6	7	8	9	10	11	12		
Pi	0.028	0.056	0.083	0.111	0.139	0.167	0.139	0.111	0.083	0.056	0.028		
xi*pi	0.056	0.167	0.333	0.556	0.833	1.167	1.111	1	0.833	0.611	0.333		7
(xi-x)2 fi	0.7	0.896	0.747	0.444	0.139	0	0.139	0.444	0.747	0.896	0.7		**5.852**

n = 36 Beispiel für i = 2 $(2-7)^2 * 1/36 = 0.7$

WICHTIGE DISKRETE VERTEILUNGEN

Bernoulli Prozess

Als Bernoulli-Prozess wird eine ***wiederholte Durchführung eines Zufallsvorgangs*** bezeichnet, wobei gilt:

- Bei jeder Wiederholung interessiert nur, ob ein bestimmtes Ereignis eintritt oder nicht
- Die Wiederholungen sind unabhängig.
- Die ***Erfolgswahrscheinlichkeit p*** bleibt gleich
- Die ***Misserfolgswahrscheinlichkeit q*** ist dann: ***q = 1-p***
- ***Erwartungswert: E(X) = p***
- ***Varianz: σ^2 = pq***

Binominal-Verteilung

Ist die ***bekannteste und wichtigste*** diskrete Verteilung:

Wahrscheinlichkeitsfunktion:

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Erwartungswert:

$$E(X) = np$$

Varianz

$$Var(X) = np(1-p)$$

Anwendung:

- Das Zufallsexperiment unterscheidet nur zwei Ergebnisse
- Das Experiment wird ***n-mal wiederholt*** (Zufallsstichprobe vom Umfang n)
- ***Gesucht***: Die ***Wahrscheinlichkeit*****, dass bei n-maliger Durchführung des Experimentes das Ereignis**
  - **genau**
  - **mindestens**
  - **höchstens**

  x-mal eintrifft

Beispiel

Gegen eine Krankheit wurde ein neues Medikament entwickelt. Die Heilungschance liegt bei 90%. Wie gross ist die Wahrscheinlichkeit, dass bei 5 zufällig gewählten Patienten mindestens 4 geheilt werden?(Addieren wegen midestens!!)

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$E(X) = np$$

$$Var(X) = np(1-p)$$

P(4) = 0.32805
P(5) = 0.59049
P(5+4)= 0.91854

E(X) = 4.5

Var(X) = 0.45

Poissonverteilung

Man interessiert sich dafür, ***wie hoch die Wahrscheinlichkeit*** ist, dass das ***Ereignis E in einem Intervall genau oder höchstens x-mal eintritt***, wenn bekannt ist, dass in diesem Intervall das Ereignis im Mittel µ - mal auftritt.
µ gibt also eine Rate pro Zeitintervall an(im ***Durchschnitt***)
Wahrscheinlichkeitsfunktion:

$$f(x) = P(X = x) = \frac{\mu^x}{x!} e^{-\mu}$$

Erwartungswert = Varianz:

$$E(X) = Var(X) = \mu$$

Unabhängige Ereignisse → Poisson Verteilt(Errinerungsfrei)

Beispiel

Die Anzahl X der Telefonanrufe, die in einer Telefonvermittlung im Mittel pro Minute ankommen sei Ps(1) – verteilt.

- Wie gross ist die Wahrscheinlichkeit, dass in einer Minute ***µ = 1***
  - Genau ein Anruf ankommt → P(1) = 0.368
  - Höchstens ein Anruf → p(0) + p(1) = 0736
  - Mindestens ein Anruf 1 – p(1) = 1 - 0.368 = 0.632
  - Zwei oder drei Anrufe ankommen P(2) + p(3) = 0.245
- Wie gross ist die Wahrscheinlichkeit, dass in 5 Minuten genau 6 Anrufe ankommen: ***µ = 5***
- P(6) = 0.146

WICHTIGSTEN STETIGEN THEORETISCHEN VERTEILUNGEN

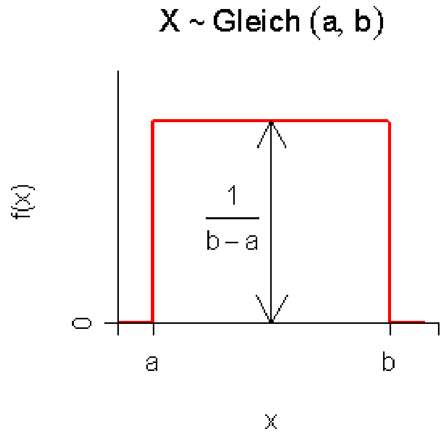
Rechteckverteilung

- Rechteckverteilungen eignen sich zur Beschreibung von Vorgängen, bei denen die ***Ergebnisse nur Zahlen eines bestimmten Intervalls [a, b]*** sein können.
- Die ***Wahrscheinlichkeit***, dass ein Ergebnis in ein ***bestimmtes Teilintervall*** fällt, wird nur ***durch dessen Länge*** bestimmt.
- ***Alle Ergebnisse*** eines bestimmten Intervalls [a, b] sind ***gleich wahrscheinlich*** →Gleichverteilung im Intervall [a, b]

E(X) = Median = $\frac{a+b}{2}$

Var(X) = $\sigma^2 = \frac{1}{12}(b-a)^2$

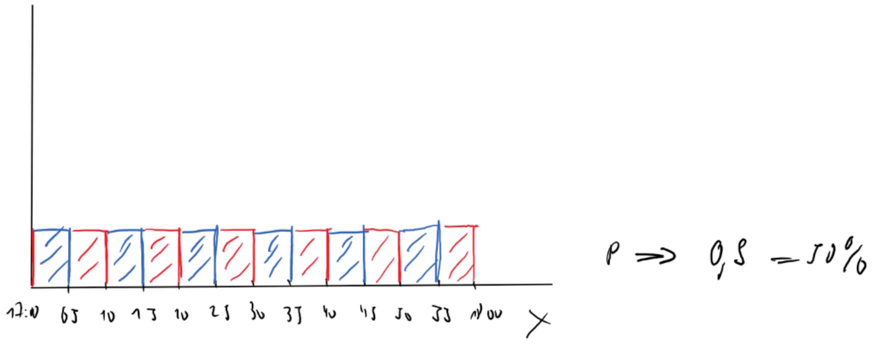
$\sigma \approx 0.289\,(b-a)$



Beispiel

Zwei Linien

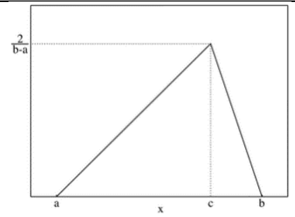
Sie besuchen Freunde an der Uni Erfurt und wollen von dort mit der nächsten S-Bahn weiterfahren und zwar zwischen 17 und 18 Uhr. Zwischen 17 und 18 Uhr fahren die Straßenbahnen der Linien 3 und 6 ab Haltestelle Universität Richtung „Urbicher Kreuz“ bzw. „Rieth“ jeweils im 10-Minuten Takt. Der erste Zug der Linie 3 fährt um 17:05 Uhr, der erste Zug der Linie 6 um 17:00 Uhr. Wenn ein Sie zufällig zu einer gleichverteilten Zeit zwischen 17:00 Uhr und 18:00 Uhr die Haltestelle erreichen und einfach in die nächste Straßenbahn einsteigt, wie groß ist die dann die Wahrscheinlichkeit, dass Sie eine Bahn in Richtung „Rieth“ nehmen?



P(blau) = Fläche blau = 50%

Dreieckverteilung

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)}, & \text{wenn } a \le x < c \\ \frac{2}{b-a}, & \text{wenn } x = c \\ \frac{2(b-x)}{(b-a)(b-c)}, & \text{wenn } c < x \le b. \end{cases}$$



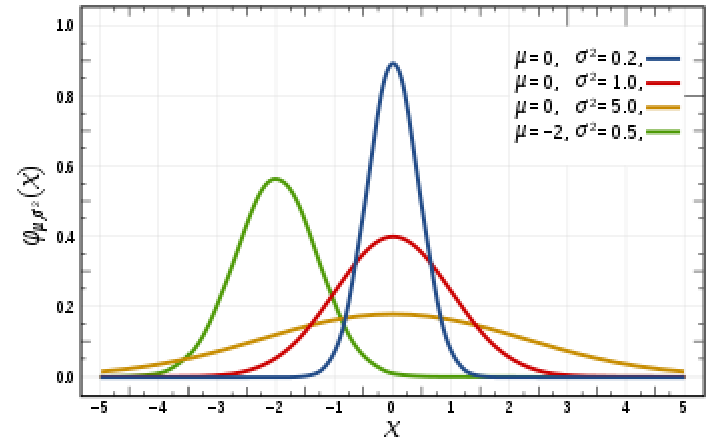
Erwartungswert oder Median: $E[x] = \frac{a+b+c}{3}$

Varianz: $E[x] = \frac{a^2+b^2+c^2-ab-ac-bc}{18}$

Beispiele

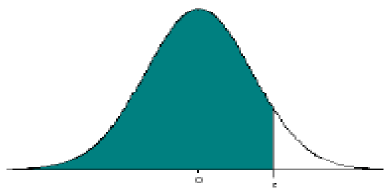
1) Beschreibung von Bedienprozessen
   Ziel ist die Ermittlung oder Beschreibung der Bearbeitungszeit, einer Dienstleistung.
   z.B. Servicezeit an einer Kasse, Zeit eines Produktionsprozesses

Exponentialverteilung	**Exponentialverteilung mit Parameter λ > 0:** **λ** = Irgendeine Rate z.B Anrufsrate $F(x) = \begin{cases} 1 - e^{-\lambda x} & , \text{für } x \geq 0 \\ 0 & , \text{für } x < 0 \end{cases}$ $f(t) = \begin{cases} \lambda t\, e^{-\lambda t} & , \text{für } t \geq 0 \\ 0 & , \text{sonst} \end{cases}$ • E(X) = $\frac{1}{\lambda}$ • Median = $\frac{\ln 2}{\lambda}$ • $Modus = 0$ • $Var(X) = \sigma^2 = \frac{1}{\lambda^2}$
Beispiele	• Zeitspanne zwischen zwei Anrufen in einer Telefonzentrale. • Dauer eines Telefongesprächs. • **Lebensdauer eines Geräts**, wenn Defekte durch äußere Einflüsse und **nicht durch Verschleiß** verursacht werden.
Beispiel Anruf	Ein Software-Hersteller hat für seine Kunden in Norddeutschland die Hotline ND eingerichtet. An Werktagen rufen zwischen 20.00 und 21 .00 Uhr durchschnittlich 5 Kunden an. ***Wie groß ist die Wahrscheinlichkeit, dass zwischen zwei Anrufen höchstens 5 bzw. 12 Minuten vergehen***? **λ** = 5 Anrufe pro Stunde → 5/60 → 1/12 Anruf/Minute -> höchstens 5 Minuten: $1 – e^{-1/12 * 5}$= 0.34 und -> höchstens 12 Minuten: 0.63 → **Beachte: Immer die gleiche Masseinheit verwenden!**
Weilbull-Verteilung ***W(α, β)-Verteilung***	**Weibull-Verteilung mit Parametern α (scale) > 0 und β (shape) > 0:** $F(x) = \begin{cases} 1 - e^{-\alpha x^{\beta}} & , \text{für } x \geq 0 \\ 0 & , \text{für } x < 0 \end{cases}$ E(X) = $\frac{1}{\alpha}\Gamma\square\left(1+\frac{1}{\beta}\right)$ Var(X) = $\sigma^2 = \frac{1}{\alpha^2}\left(\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right)$ • $\Gamma(n+1) = n!\ \ für\ jede\ natürliche\ Zahl\ \ n$ • $\Gamma(x) = \int_0^{\infty} t^{x-1}\ e^{-1}\ \ für\ positive\ reelle\ Zahlen\ x$
Anwendung	Beschreibung der **Lebensdauer von Geräten oder Materialien mit Abnutzungserscheinungen** • α ist ein Skalierungsparameter, β ist ein Formparameter. **Interpretation des Formparameters β > 0:** β < 1: Ausfallrate nimmt mit der Zeit ab (Ausfälle finden frühzeitig statt). β = 1: Ausfallrate konstant (zufällige äußere Einflüsse sind Ursache des Versagens). β > 1: Ausfallrate nimmt mit der Zeit zu (Alterungsprozesse). **Bemerkung** Der Parameterwert β = 1 führt auf die Exponentialverteilung, welche somit einen Spezialfall der Weibull-Verteilung darstellt. Passt oft besser als normaverteilt.
Zentrale Grenzwertsatz	Der Zentrale Grenzwertsatz besagt, dass irgendwann, wenn man genug Zufallsvariablen(Werte) hat, alles näherungsweise normalverteilt ist.
Normalverteilung	Die wichtigste stetige Verteilung. Spielt in der schliessenden Statistik eine wichtige Rolle. **<u>Wahrscheinlichkeitsdichteverteilung</u>** $f(x) == \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \sigma > 0$

	<u>Verteilungsfunktion</u>(Integral lässt sich nicht lösen) **<u>Erwartungswert</u>** **<u>Varianz</u>** $Var\ (X) = \sigma^2$ $F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt, \sigma > 0$ $E(X) = \mu$
Wichtige Anmerkungen	• Die Normalverteilung ist eine ***stetige symmetrische Verteilung*** • µ enspricht dem Erwartungswert(x-Wert des Scheitelpunktes) →xmax = µ • Wird ***µ verändert, so hat es eine Verschiebung nach links oder rechts*** zu Folge • σ entspricht der Standardabweichung • Wird ***σ verändert, so hat es eine Veränderung in der Breite der Kurve*** zu Folge • Die Fläche zwischen zwei Werten entspricht der Wahrscheinlichkeit, dass x in diesem Bereich liegt. 
Wichtige Beziehungen	Es gibt folgende wichtige Beziehungen ***zwischen der Verteilfunktion F(X) und der Normal-verteilung:*** $F(x) = F_0(\frac{x-\mu}{\sigma})$ $z = \frac{x-\mu}{\sigma}$
Beispiele	σ= 1 und m = 0, in dem Fall ist z = x – z = -∞ und z = 0 → 0.5 – z = 0 und z = 1.2 →0.385 – z = -1.2 und 0 →0.385 – z = 0.81 und z = 1.94 = 0.183

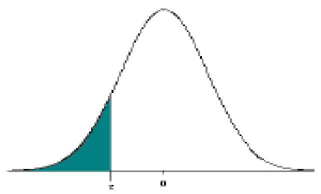
Positive z Werte

Table of Standard Normal Probabilities for Positive Z-scores



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Negative z Werte



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Beispiel Zucker Abfüllanlage

Messungen ergeben, dass das Gewicht der Zucker- Pakete in Gramm um den Mittelwert μ= 1000 g um die Standardabweichung σ = 5g schwankt. Es kann von einer Normalverteilung ausgegangen werden.

- Wie gross ist die Wahrscheinlichkeit, dass der tatsächliche Inhalt der Packung zwischen 990 Gramm und 1010 Gramm liegt

z1 = (x1- μ) / σ = (990-1000)/5 = -2
z1 = (x2- μ) / σ = (1010-1000)/5 = 2
p(2) = 0.9772
p(-2) = 0.0228
P(2) – P(-2) = 0.9544

- Wie gross ist die Wahrscheinlichkeit das in einer Packung mehr als 1015 Gramm enthalten sind

Z = (x- μ) / σ = (1015-1000)/5 =
P(3) = 0.9987
1 – p(3) = 0.0013

- Wie viel wiegt zu 90% ein zufällig entnommenes Paket höchstens?

0.9 aus Tabelle gelesen → z=1.285
1.285 = (x- 1000) / 5 → x = 1006.43 g

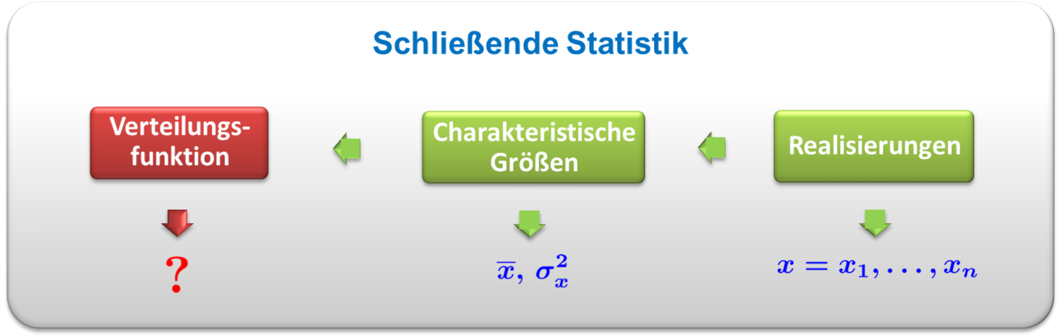
- In welchem Bereich schwankt das Gewicht eines Paketes in 98% Prozent aller Fälle symmetrisch zum Mittelwert?

P(2.3) = 0.99
P(-2.3) = 0.01
2.3 = (x- 1000) / 5 = = 1011.5
-2.3 = (x- 1000) / 5 = 988.5
Es schwankt zwischen 988.5g und 1011.5 g
Dabei ist die Schwankung vom Mittelwert aus: $\frac{1011.5-988.5}{2} = 11.65$

Beispiel Flugzeug

Bei einem bestimmten Grossraumflugzeug ist die Auslastung pro Flug näherungsweise normalverteilt, mit den Parametern μ = 150 Passagiere und σ = 25 Passagiere.
Mit welcher Anzahl von Passagieren ist mit einer Wahrscheinlichkeit von 90% **mindestens** zu rechnen?
P(1.285) = 0.9 //aus der Tabelle gelesen
1.285 = (x – 150) / 25 = 117.875
Es ist mindestens mit 118 Personen zu rechnen.

SCHLIESSENDE STATISTIK

Grundlagen	 Folgende Fakten sind wichtig: • Die schließende Statistik befasst sich mit ***Messreihen***, deren ***konkrete Ausprägungen vom Zufall*** beeinflusst werden. • Diesen Messreihen können ***ebenfalls charakteristische Größen*** etwa in Form des ***arithmetischen Mittels*** und der ***Varianz*** zugeordnet werden. • Hier stellt sich die ***Frage*** nach der ***Gestalt des unbekannten Verteilungsgesetzes*** unter dem die vorliegende Realisierung zustande kam. • ***Anhand der charakteristischen Größen*** kann man die eingesetzte ***Verteilungs-funkton annehmen*** oder ***ablehnen***. • Man hat eine angenommene Verteilungsfunktion und eine theoretische Verteilung, diese werden verglichen um eine passende Funktion zu fin-den.
Beispiel	***Ziel der Schliessenden Statistik ist es aus einer Teilmenge einer Menge auf die ganze Menge zu schliessen***: Beispielsweise nimmt man 1000 Studenten und untersucht, ob und wie sich der Alkohol-konsum von Studenten auf ihre Noten auswirkt. Dann schliesst man aus den erhobenen Daten auf alle Studenten.
ZUFALLSSTICHPROBEN	
Vorgehen / Problembereiche	1) Die Stichproben- oder Datenanalyse 2) Die Auswertung oder Analyse der erhobenen Daten
Einfache Zufallsstichprobe	- Bekannteste Form der Stichprobe - Es werden Elemente(Proben) so ausgewählt, dass ***die Wahrscheinlichkeit für alle Elemente gleich gross ist.*** **<u>Beispiel</u>**: Ich befrage einzelne Studenten aus allen Hochschulen in der Schweiz über ihre Motivation zum Lernen.
Geschichtete Stichprobe	- Man ***bildet Gruppen***(Schichten) und ***entnimmt*** jeweils verschiedene ***Elemente aus*** den jeweiligen ***Gruppen***(Schichten) - Dabei haben ***Elemente in den*** gleichen ***Gruppen*** ähnliche oder noch besser ***gleiche Merkmale***(Homogenität der Elemente) **<u>Beispiel</u>**: Ich gruppiere die Studenten der Schweiz anhand der Studienrichtigung und be-frage sie über die Motivation zum Lernen. Dabei werden aus jeder Schicht(Studiengang) einige Personen befragt.
Klumpen Stichprobe	Es ***existieren Teilmengen***(Klumpen), ***welche*** die ***Grundgesamtheit*** relativ ***gut abbilden***. Dann werden Proben aus den Klumpen entnommen, die dann ausgewertet werden. **<u>Beispiel:</u>** Die HSR und die HTW haben in den letzten Jahren in der Umfrage «Motivation beim Lernen» ähnlich abgestummen wie die Grundgesamtheit(alle Studenten Schweiz). Deshalb werden jetzt nur noch Proben aus dem Klumpen HTW und HSR genommen.
Systematische Stichprobe	Es werden ***Regeln***(eine ***Systematik***) ***definiert***, die auf wiederholbare Weise eine Zufalls-ähnliche Stichprobenauswahl ermöglicht. **<u>Beispiel:</u>** Bei der Umfrage werden jeweils 10 Studenten aus jedem Kanton befragt.
Mehrstufige Stichprobe	Ist letztlich eine ***Kombination der*** oben benannten ***Verfahren***
MODELLBILDUNG	

Vorgehen	a. **Isolierende Abstraktion** Ermitteln der Merkmalsträger und Merkmale, die zur Modellbildung erforderlich sind. ***Suche den Kern des Experiments*** und blende Unwichtiges aus. b. Verteilungsfunktion bestimmen. Daraus lassen sich beliebig viele Daten für die Experimentdurchführung generieren. c. Daten für Modellverifikation und Modellvalidierung festlegen d. Festlegen der control/respose-Daten
Anforderungen der Experi-mentauswertung	• Wie sieht die Stichprobenfunktion und ihre Wahrscheinlichkeits-verteilung aus? • Wie kann ein Vertrauensintervall für den Mittelwert bzw. die Varianz einer Messreihe ermittelt werden? • Wie kann der Stichprobenumfang, d.h. die Anzahl der Experi-mente, festgelegt werden? • Welche etablierten Testverfahren gibt es?
TESTVERTEILUNGEN	
Grundlagen	Sind Verteilungen, die bei vielen statistischen Tests Verwendung finden. - ***Normalverteilung***, wenn die Stichprobe $n \geq 30$ ist - ***t-Verteilung*** (auch Student t-Verteilung), wenn die ***Stichprobe*** < 30 ist - Chi Quadrat Verteilung - Viele Stichproben sind annähernd Normalverteilt(Zentraler Grenzwertsatz)
NORMALVERTEILUNG	
Konfidenzintervalle für den Wert m der	Liegt eine konkrete Stichprobe {x1, x2 , x3, …, xn} vor, so gilt für das **<u>Stichprobenmittel</u>**(der Erwartungswert = der Mittelwert) $$E[X] = \mu = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$ **<u>Stichprobenfunktion(= Stichprobenmittelwert)</u>** $$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
Satz 1	Es sei X ein Zufallsvariable mit dem Erwartungswert µ und der Varianz σ^2. **<u>Stichprobenmittelwert</u>** $$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Erwartungswert

$$\overline{X} = \mu_{\overline{x}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \mu$$

Varianz

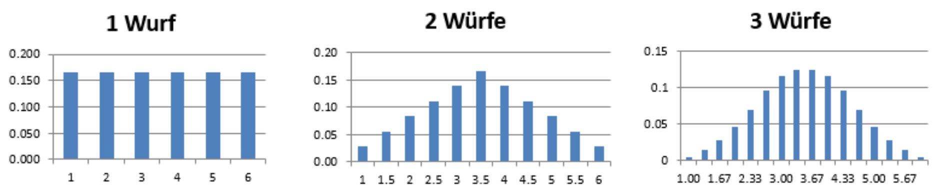
$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}$$

Interpretation Erwartungswert:
Der Stichprobenmittelwert hat den gleichen Erwartungswert, wie die Grundgesamtheit. Dieser Mittelwert ist selbst wieder eine Zufallsvariable oder Stichprobenfunktion und streut um den Mittelwert der Grundgesamtheit. Je grösser n ist, desto kleiner die Streuung oder, um so besser die Näherung.

Grafische Darstellung

Je mehr n man hat, desto näher kommt man an die Normalverteilung (Zentraler Grenzwertsatz):

	1 WURF	2 WURF	3 WURF
Mittelwert µ:	3.5	3.5	3.5
Varianz σ²:	2.92	2.92/2 = 1.46	2.92/3 = 0.97
Standab. σ :	1.71	1.21	0.98



Satz 2

Es gelten die Voraussetzungen von Satz 1. Ist *X* darüber hinaus normalverteilt, ist auch der Stichprobenmittelwert Normalverteilt.

$$\overline{X} \Rightarrow N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$$

σ die Standardabweichung
µ der Mittelwert
n Anzahl der Stichproben

Stichprobenwert

$$Z_{\text{stichpro}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Beispiel

Sie haben das Simulationsmodell einer Produktionseinrichtung erstellt, vom realen System wissen sie, dass diese Maschine im ***normalverteilten Mittel 10 Stück pro Sekunde*** produziert und die ***Standardabweichung von 1 Stück pro Sekunde*** besitzt. Sie führen in der Simulationsumgebung ***25 Experimente*** (Replications) durch.
Gegeben: µ = 10 Stk./s und σ = 1 Stk./s und n=25

a. Mit welcher Wahrscheinlichkeit wird der mittlere Ausstoss zwischen 9.8 und 10.2 Stück pro Sekunde liegen?

Gesucht:

$$P(9.8 \leq \overline{x} \leq 10.2)$$

Z1 = $\frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{9.8-10}{\frac{1}{\sqrt{25}}} = -1$ → P(-1) = 0.1587(aus Tabelle)
Z2 = $\frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{10.2-10}{\frac{1}{\sqrt{25}}} = 1$ → P(1) = 0.8413(aus Tabelle)
P(-1 < x < 1) = 1 – (0.8413 – 0.1587) = 0.3174 = 31.74%

b. Mit welcher Wahrscheinlichkeit wird der Ausstoss über 10.2 liegen

Z = $\frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{10.2-10}{\frac{1}{\sqrt{25}}} = 1 \rightarrow \mathrm{P}(1) = 0.8413(\text{aus Tabelle})$
1 – 0.8413 = 0.1587 = 15.87%

c. Wie verändert sich das Ergebniss, wenn sie statt dessen 100 Experimente machen?

Z = $\frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{10.2-10}{\frac{1}{\sqrt{100}}} = 2 \rightarrow \mathrm{P}(2)\ 0.9772(\text{aus Tabelle})$
1 - 0.9772= 0.0228 → weniger Varianz - annähernd normalverteilt(Zentraler Grenzwertsatz)

Konfidenzintervall

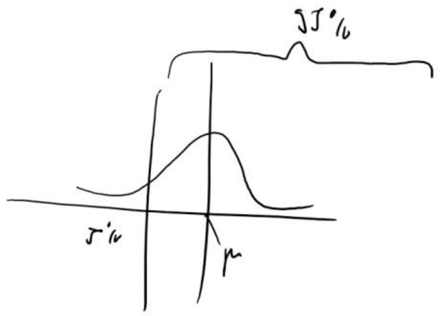
$$\overline{Z} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\mathrm{P}\left(-z \leq \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z\right) = 1 - \alpha$$

$$\mathrm{P}\left(-z\frac{\sigma}{\sqrt{n}} \leq \overline{X} - \mu \leq z\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$\mathrm{P}\left(\mu - z\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq z\frac{\sigma}{\sqrt{n}} + \mu\right) = 1 - \alpha$$

Konfidenzintervall
$P(\mu - \Delta x \leq \mu \leq \mu + \Delta x) = gesuchte\%$
Beispiel: $P(1000.55 - 0.39 \leq \mu \leq 1000.55 + 0.39) = 95\%$

Nach unten begrenzt: Die Grenze ist unten. Mindestens:	Nach oben begrenzt z,B (1-p(untere Grenze)):
	

T-VERTEILUNG

Grundlagen	T-Verteilungen werden eingesetzt, wenn n< 30. T-Verteilungen haben, wie Normalverteilungen ein glockenförmigen Verlauf, sind aber flacher und breiter.
Formeln	$f(t) = \frac{1}{\sqrt{r} \cdot B\left(\frac{1}{2}, \frac{r}{2}\right)} \cdot \left(1 + \frac{t^2}{r}\right)^{-\frac{1}{2}(r+1)}, \quad t \in \mathbb{R}$ $t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$ $E(X) = 0 \text{ (für } r > 1),$ $Var(X) = \frac{r}{r-2} \text{ (für } r > 2)$
Freiheitsgrade r	Der **Freiheitsgrad r** bestimmt bei einer Gleichung, wie viele Parameter «**frei**» **wählbar** sind.

Tabelle 1: Vertrauensfaktor t der Student t-Verteilung

Anahl Freiheits-grade	Anzahl Messungen n in der Messreihe	Student t-Verteilung: **Vertrauensfaktor t**						
		Vertrauensniveau [%]						
		68.27	90.00	95.00	95.45	99.00	99.727	99.9937
1	2	1.84	6.31	12.71	13.97	63.66	233.19	10105.08
2	3	1.32	2.92	4.30	4.53	9.92	19.10	125.98
3	4	1.20	2.35	3.18	3.31	5.84	9.18	32.68
4	5	1.14	2.13	2.78	2.87	4.60	6.60	17.47
5	6	1.11	2.02	2.57	2.65	4.03	5.49	12.30
6	7	1.09	1.94	2.45	2.52	3.71	4.89	9.85
7	8	1.08	1.89	2.36	2.43	3.50	4.52	8.47
8	9	1.07	1.86	2.31	2.37	3.36	4.27	7.60
9	10	1.06	1.83	2.26	2.32	3.25	4.09	7.00
19	20	1.03	1.73	2.09	2.14	2.86	3.44	5.10
29	30	1.02	1.70	2.05	2.09	2.76	3.28	4.67
49	50	1.01	1.68	2.01	2.05	2.68	3.16	4.38
99	100	1.01	1.66	1.98	2.03	2.63	3.07	4.18
199	200	1.00	1.65	1.97	2.01	2.60	3.03	4.09
$n \to \infty$	$n \to \infty$	**1.00**	1.65	1.96	**2.00**	2.58	**3.00**	**4.00**

SCHÄTZVERFAHREN & **ABSCHÄTZUNG** DER UNBEKANNTEM PARAMETER μ UND σ

SCHÄTZVERFAHREN

Grundlagen

Haben die Aufgabe den oder die ***unbekannten Parameter*** der Verteilung eines Merkmals in der Grundgesamtheit ***anhand einer Stichprobe zu schätzen***.
Die Schätzung kann

- durch die Angabe eines einzigen Wertes erfolgen → **Punktschätzung**
- durch die Angabe eines Intervalls → **Intervallschätzung**

SCHÄTZFUNKTION

Grundlagen

- mathematisches Instrument
- Bindeglied zwischen Grundgesamtheit und Stichproben
- Verteilung der Schätzfunktion approximativ bekannt
- ***Ziel*** der Schätzfunktionen ist,
  - ***von einer Stichprobe auf die Grundgesamtheit zu schliessen***,
  - Den Fehler einer falschen Schätzung zu minimieren oder zu bestimmen
  - **Punktschätzung**,
    - ***Parameterschätzung*** für den ***Mittelwert*** μ, die ***Varianz*** σ^2 und für ***unbekannte Wahrscheinlichkeiten*** (auch Anteilswerte genannt) p
  - **Intervallschätzung**
    - ***Bestimmung von Vertrauensintervallen*** für die oben aufgeführten Parameter und das damit verbundene ***Risiko*** einer ***Fehlentscheidung*** oder ***Fehlinterpretation*** ,
    - Die ***Angabe*** des ***Fehlers*** oder der ***Genauigkeit einer Schätzung*** wird auch als ihre Zuverlässigkeit bezeichnet

Gütekriterien für Schätzfunktionen

$$E[(\hat{T}-T)^2]$$

$$\mathrm{E}\left[\left(\hat{T}-T\right)^2\right] = VAR(\hat{T}) + \left[E[\hat{T}]-T\right]^2$$

Ist $E(\hat{T})-T=0$ *so spricht man von* ***Erwartungstreu***
Ein kleiner Restfehler ist ein Gütekriterium für eine Schätzfunktion.

Konstruktion der Schätzfunktion für μ

- Minimiere $E[(\hat{T}-T)^2]$
- Also minimiere $\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$
- Differenzieren nach $\hat{\mu}$ und 0 setzen ergibt die Schätzfunktion zu

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

Punktschätzung

Unbekannter Parameter der Grundmenge	Schätzfunktion (Ergebnis der Stichprobe)
Mittelwert μ	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Varianz σ^2	o.B. $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$
Anteilswert/ Wahrscheinlichkeit p	$\bar{p} = \frac{k}{n}$

Es kann immer die t-Verteilung genommen werden, weil diese bei n > 30 ungefähr normalverteilt ist & bei n>30 besser als die Normalverteilung ist.

Intervallsschätzung für das Stichprobenmittel $\bar{X}$

Stichprobe	Varianz σ^2 bekannt	Varianz σ^2 unbekannt
Mit Zurücklegen (unendliche Grundgesamtheit)	$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$	$\hat{\sigma}_{\bar{X}}^2 = \frac{s^2}{n}$
ohne zurücklegen $\frac{n}{N} < 0.05$	$\sigma_{\bar{X}}^2 \approx \frac{\sigma^2}{n}$	$\hat{\sigma}_{\bar{X}}^2 \approx \frac{s^2}{n}$
ohne zurücklegen $\frac{n}{N} \geq 0.05$	$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\cdot\frac{N-n}{N-1}$	$\hat{\sigma}_{\bar{X}}^2 = \frac{s^2}{n}\cdot\frac{N-n}{N}$

N: Grundmenge
n: Grösse der Stichprobe

Korrekturfaktor

Erstellung eines Konfidenzintervalls
Bei konkreten Werten

1. Feststellung der Verteilungsform von $\bar{X}$
   $\frac{n}{N}$ entscheiden ob, kleiner oder grössergleich 0.05
2. Feststellung der Varianz von $\bar{X}$ ggf. schätzen mit s^2
3. Ermittlung des Quantilwertes ***z aus Tabelle*** oder Rechner
4. Berechnung des maximalen Schätzfehlers
   *Der **maximale Schätzfehler** ist das **Produkt** aus **Quantilswert** und **Standardabweichung** von X .*
5. Ermittlung der Konfidenzgrenzen
   *Die **untere** und die **obere Konfidenzgrenze** ergeben sich durch **Subtraktion** bzw. **Addition** des **maximalen Schätzfehlers** vom bzw. **zum Stichprobenmittel** $\bar{X}$*

Beispiel Wurstfabrik

In einer Wurstfabrik werden u.a. Leberwürste hergestellt. Aus langjährigen Messreihen ist bekannt, dass das Füllgewicht der Leberwürste normalverteilt ist. Das ***Soll-Mindestgewicht*** der Würste ***beträgt 125 g.*** Aus der ***Tagesproduktion von 600 Würsten wurden 26 Würste*** zufällig ohne Zurücklegen ***entnommen*** und gewogen. Die Messergebnisse für das Füllgewicht (in g) betrugen dabei:

128.4	123.8	123.5	126.9	125.5	123.1	124.9
123.1	126.6	121.9	125.3	123.4	122.1	124
123.3	123.2	123.2	124	122.8	127.1	125.7
127.1	125.8	123.7	125.9	124.9		

$$\bar{X} = 124.5\text{ g},\ s = 1.72\text{ g}$$

- Erstellung des zentralen 95%-Konfidenzintervalls für μ

1. Feststellung der Verteilungsform von $\bar{X}$:

$\frac{n}{N} = \frac{26}{600} = 0.043 < 0.05$

2. Feststellung der Varianz von $\bar{X}$ ggf. schätzen mit s^2

$s = \sqrt{1.72} = 1.31$

3. Ermittlung des Quantilwertes z aus Tabelle oder Rechner(97.5%)

Z = 1.96

4. Berechnung des maximalen Schätzfehlers

Der maximale Schätzfehler ist das Produkt aus Quantilswert und Standardabweichung von X .

E = 1.96 * 1.31 = 2.57

Ermittlung der Konfidenzgrenzen:

124.5g ± 2.57 = [121.93, 127.07]

Varianzen für die Schätzfunktion

$$P = \frac{k}{n} = q, \; für \; n\,P\,(1-P) > 9 \; aproximativ \; normalverteilt$$

Stichprobe	Varianz σ^2 bekannt	Varianz σ^2 Unbekannt
Mit Zurücklegen	$\hat{\sigma}_P^2 = \frac{P\,(1-P)}{n}$	
ohne zurücklegen $\frac{n}{N} < 0.05$	$\hat{\sigma}_P^2 \approx \frac{P\,(1-P)}{n}$	
ohne zurücklegen $\frac{n}{N} \geq 0.05$	$\hat{\sigma}_P^2 = \frac{P\,(1-P)}{n} \cdot \frac{N-n}{N-1}$	$\hat{\sigma}_P^2 = \frac{P\,(1-P)}{n} \cdot \frac{N-n}{N}$

N: Grundmenge
n: Grösse der Stichprobe

Korrekturfaktor

Erstellung Konfidenzintervall Bei Prozentsätzen

1. Feststellung der Verteilungsform von P
   die Schätzfunktion ist aprox. normalverteilt, wenn n*P*(1-P) > 9 normalverteilt
2. Feststellung der Varianz von P
   siehe vorherige Folie
3. Ermittlung des Quantilswertes z
4. Berechnung des maximalen Schätzfehlers
   Der maximale Schätzfehler ist das Produkt aus Quantilswert und Standardabweichung von P
5. Ermittlung der Konfidenzgrenzen
   Die untere und die obere Konfidenzgrenze ergeben sich durch Substraktion bzw. Addition des maximalen Schätzfehlers vom bzw. zum Stichprobenmittel P

Beispiel

Ein Chemieunternehmen möchte den Bekanntheitsgrad eines von ihm hergestellten Waschmittels in Erfahrung bringen. Dazu werden ***400 Personen*** zufällig ausgewählt und befragt. Das Waschmittel war bei ***30 % der Befragten bekannt***.

- Erstellen des zentralen 95%-Konfidenzintervalls

1) Feststellung der Verteilungsform von P

N*P*(1-P) = 400 * 0.3 * 0.7 = 84 → 84>9 →normalverteilt

2) Feststellung der Varianz von P

$$\sigma^2 = \frac{P(1-P)}{n} = \frac{0.3 * (1-0.3)}{400} = 0.0005353 \rightarrow \sqrt{0.0005353} = 0.023$$

3) Ermittlung des Quantilswerts(0.975)

Z = 1.96

4) Berechnung des maximalen Schätzfehlers

e = 1.96 * 0.023 = 0.045 = 4.5 %

5) Ermittlung der Konfidenzgrenzen

30% ± 4.5 = [25.5, 34.5]

Anzahl Proben mit zurücklegen

$$\mu \pm e = \mu \pm z\frac{\sigma}{\sqrt{n}} \; mit \; der \; Wahrschinlichkeit \; 1 - \alpha$$

$$e \leq \pm z\frac{\sigma}{\sqrt{n}}$$

$\mathrm{n} \geq \frac{Z^2\sigma^2}{e^2}$ Bestimmung des Z-Wertes***<u>(Varianz nicht doppelt quadrieren)</u>***

Beispiel

Es ist bei einer ***Konfidenz von 95 %*** und einer bekannten ***Varianz von1.44 g²*** die ***erforderliche Anzahl von Proben zu ermitteln***, wenn eine ***Genauigkeit von 0.2 g*** gefordert wird. Annahme Normalverteilung, dann folgt für z = 1.96!

$$\mathrm{n} \geq \frac{Z^2\sigma^2}{e^2} = \frac{1.96^2 * 1.44}{0.2^2} = mindestens \; 139 \; Proben$$

Achtung! Die Varianz ist = σ^2 muss also nicht nochmals quadriert werden!

Anzahl Proben ohne Zurücklegen

$$e \leq \pm z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$n \geq \frac{z^2\,N\,\sigma^2}{e^2(N-1) + z^2\,\sigma^2}$$

Beispiel

Einer Lieferung von ***1000 Paketen Zucker*** ist mit einer ***95 % Konfidenz*** bei einem Fehler von ***e = 0.2*** g zu untersuchen, ob bei einer bekannten Standardabweichung von 1.2 g der garantierte Mittelwert eingehalten wird. Wie viele Proben sind aus der Lieferung mindestens zu entnehmen?

Z Aus Tabelle = 1.96

$$n \geq \frac{z^2\,N\,\sigma^2}{e^2(N-1) + z^2\,\sigma^2} = \frac{1.96^2 \cdot 1000 \cdot 1.2^2}{0.2^2 \cdot (1000-1) + 1.96^2 \cdot 1.2^2} = 121.6 = 122 \; Prb$$

Konfidenzintervall für die Varianz I

$$s^2 = \frac{\sigma^2}{n-1}\sum_{i=1}^{n} Z_i^{\,2} \rightarrow \sum_{i=1}^{n} Z_i^{\,2} = \frac{s^2\,(n-1)}{\sigma^2}$$

$$Y = \frac{s^2\,(n-1)}{\sigma^2}$$

$$P\left(y_{\frac{\alpha}{2};\, r=n-1} \leq \frac{(n-1)\,s^2}{\sigma^2} \leq y_{1-\frac{\alpha}{2};\, r=n-1}\right)$$

$$P\left(\frac{(n-1)\,s^2}{y_{1-\frac{\alpha}{2};\, r=n-1}} \leq \sigma^2 \leq \frac{(n-1)\,s^2}{y_{\frac{\alpha}{2};\, r=n-1}}\right)$$

Beispiel

Gegeben sind die folgenden Werte aus dem vorangegangenen Beispiel der Wurstfabrik: $n = 26$, $\bar{x} = 124.58$, $s^2 = 2.95$

a. *Erstellung des (zentralen) 95% Konfidenzintervalls für σ^2*
b. *Erstellung des noch oben begrenzten 95 % Konfidenzintervalls*

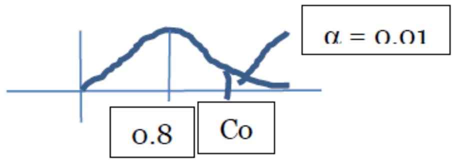
$$\frac{\alpha}{2} = 0.025, \; 1 - \frac{\alpha}{2} = 0.975 \rightarrow y_{0.025;25} = 13.12 \; ;$$

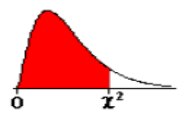
$$y_{0.975;\,25} = 40.65$$

$$W\left(\frac{(n-1)\,s^2}{y_{1-\frac{\alpha}{2};\, r=n-1}} \leq \sigma^2 \leq \frac{(n-1)\,s^2}{y_{\frac{\alpha}{2};\, r=n-1}}\right) = 0.95 \rightarrow$$

$$W\left(\frac{25 \cdot 2.95}{40.64631} \leq \sigma^2 \leq \frac{25 \cdot 2.95}{13.1197}\right) = 0.95$$

$$W(1.81 \leq \sigma^2 \leq 5.62)$$

	$\alpha = 0.05 \rightarrow \quad = 14.6114,$ $W\left(\sigma^2 \leq \frac{25 \cdot 2.9}{14.6114}\right) \rightarrow W(\sigma^2 \leq 4.96) = 0.95$
Länge des Konfidenzintervalls	Die Länge des Konfidenzintervalls variiert mit der Anzahl der Stichproben: ***Wenige Stichproben → Weniger Vertrauen → Grosses Konfidenzintervall*** ***Viele Stichproben → Mehr Vertrauen → Kleiners Konfidenzintervall*** ***Mehr Vetrauen/Sicherheit durch:*** ***1. Mehr Stichproben*** ***2. Grösseres Konfidenzintervall***
Vorgehen wenn Sigma nicht vorgegeben	**<u>μ bestimmen:</u>** $\mu = p * n$ ***<u>Standardabweichung bestimmen</u>*** $Var(x) = n * p * (1-p)$ $\sigma = \sqrt{n * p * (1-p)}$ ***<u>Z bestimmen</u>*** $Z = \frac{x - \mu}{\sigma}$
Hypothesentest	**$Co = \mu + z * \sqrt{\frac{p*(1-p)}{n}}$ //Maximalwert des Hypothese** **Z in diesem Fall 1-0.01 = 0.99** 



Lesebeispiel: Gesucht sei der c²-Wert, unter dem bei df=17 Freiheitsgraden 95% aller möglichen Werte einer c²-verteilten Zufallsvariablen X² liegen. In der Zeile für df=17 finden Sie in der Spalte (1-a)=0,95 den gesuchten Wert c²=27,59.

df	(rote/dunkle) Fläche (1-a)								
	0,7	**0,75**	**0,8**	**0,85**	**0,9**	**0,95**	**0,975**	**0,99**	**0,995**
1	1,07	1,32	1,64	2,07	2,71	3,84	5,02	6,63	7,88
2	2,41	2,77	3,22	3,79	4,61	5,99	7,38	9,21	10,60
3	3,66	4,11	4,64	5,32	6,25	7,81	9,35	11,34	12,84
4	4,88	5,39	5,99	6,74	7,78	9,49	11,14	13,28	14,86
5	6,06	6,63	7,29	8,12	9,24	11,07	12,83	15,09	16,75
6	7,23	7,84	8,56	9,45	10,64	12,59	14,45	16,81	18,55
7	8,38	9,04	9,80	10,75	12,02	14,07	16,01	18,48	20,28
8	9,52	10,22	11,03	12,03	13,36	15,51	17,53	20,09	21,95
9	10,66	11,39	12,24	13,29	14,68	16,92	19,02	21,67	23,59
10	11,78	12,55	13,44	14,53	15,99	18,31	20,48	23,21	25,19
11	12,90	13,70	14,63	15,77	17,28	19,68	21,92	24,73	26,76
12	14,01	14,85	15,81	16,99	18,55	21,03	23,34	26,22	28,30
13	15,12	15,98	16,98	18,20	19,81	22,36	24,74	27,69	29,82
14	16,22	17,12	18,15	19,41	21,06	23,68	26,12	29,14	31,32
15	17,32	18,25	19,31	20,60	22,31	25,00	27,49	30,58	32,80
16	18,42	19,37	20,47	21,79	23,54	26,30	28,85	32,00	34,27
17	19,51	20,49	21,61	22,98	24,77	27,59	30,19	33,41	35,72
18	20,60	21,60	22,76	24,16	25,99	28,87	31,53	34,81	37,16
19	21,69	22,72	23,90	25,33	27,20	30,14	32,85	36,19	38,58
20	22,77	23,83	25,04	26,50	28,41	31,41	34,17	37,57	40,00

x^2 - Verteilung

v	α =											
	0.001	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995	0.999
1	0.000	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879	10.828
2	0.002	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597	13.816
3	0.024	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838	16.266
4	0.091	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860	18.467
5	0.210	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750	20.515
6	0.381	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548	22.458
7	0.598	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278	24.322
8	0.857	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955	26.125
9	1.152	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589	27.877
10	1.479	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188	29.588
11	1.834	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757	31.264
12	2.214	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300	32.910
13	2.617	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819	34.528
14	3.041	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319	36.123
15	3.483	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801	37.697
16	3.942	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267	39.252
17	4.416	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718	40.790
18	4.905	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156	42.312
19	5.407	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582	43.820
20	5.921	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997	45.315
21	6.447	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401	46.797
22	6.983	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796	48.268
23	7.529	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181	49.728
24	8.085	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559	51.179
25	8.649	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928	52.620
26	9.222	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290	54.052
27	9.803	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645	55.476
28	10.391	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993	56.892
29	10.986	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336	58.301
30	11.588	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672	59.703
31	12.196	14.458	15.655	17.539	19.281	21.434	41.422	44.985	48.232	52.191	55.003	61.098
32	12.811	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328	62.487
33	13.431	15.815	17.074	19.047	20.867	23.110	43.745	47.400	50.725	54.776	57.648	63.870
34	14.057	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964	65.247
35	14.688	17.192	18.509	20.569	22.465	24.797	46.059	49.802	53.203	57.342	60.275	66.619