



Statistik

Andy Dunkel

E-Mail: andy.dunkel@ekiwi.de

Homepage: <http://da.ekiwi.de>

Inhaltsverzeichnis

1 Grundlagen.....	4
1.1 Statistische Begriffe.....	4
1.2 Statistische Maßzahlen.....	4
2 Beschreibende Statistik.....	4
2.1 Häufigkeitsverteilungen.....	5
2.1.1 Absolute und relative Häufigkeit.....	5
2.2 Absolute und relative Summenhäufigkeit.....	6
2.3 Statistische Maßzahlen.....	6
2.3.1 Modalwert (Modus) D.....	6
2.3.2 Median / Zentralwert Z.....	6
2.3.3 Mittelwerte.....	7
2.3.3.1 Arithmetischer Mittelwert.....	7
2.3.3.2 Geometrischer Mittelwert.....	7
2.3.3.3 Harmonischer Mittelwert.....	8
2.3.4 Streuungsparameter.....	8
2.3.4.1 Spannweite d.....	8
2.3.4.2 Mittlere lineare (absolute) Abweichung.....	8
2.3.4.3 Varianz und Standardabweichung.....	9
2.3.4.4 Varianzkoeffizient.....	9
2.3.5 Konzentrationsmaße.....	9
2.3.5.1 Lorenzkurve.....	9
2.3.6 Ginikoeffizient und Konzentrationsfläche.....	10
2.3.7 Fehlerrechnung – Messfehler in der Physik.....	10
2.3.7.1 Fehlerrechnung für zufällige Fehler.....	11
2.3.7.2 Gaußsche Glockenkurve.....	11
2.3.7.3 Messunsicherheit und Vertrauensbereich.....	12
2.3.7.4 Fehlerfortpflanzung.....	12
2.3.7.5 Größtfehlerabschätzung.....	12
2.3.7.6 Relative und prozentuale Fehlerangabe.....	13
2.3.7.7 Angabe des Ergebnisses.....	14
2.3.7.8 Fehlermessung bei analogen Messgeräten.....	14
3 Wahrscheinlichkeitsrechnung.....	14
3.1 Grundbegriffe.....	14
3.1.1 Axiome der Wahrscheinlichkeitsrechnung.....	15
3.1.2 Gesetz der großen Zahlen.....	15
3.1.2.1 Erwartungswert E.....	15
3.1.3 Laplace-Experiment und die Wahrscheinlichkeit.....	15
3.1.3.1 Stochastische Unabhängigkeit.....	17
3.1.4 Bedingte Wahrscheinlichkeit.....	17
3.1.5 Bedingte Wahrscheinlichkeit im Baumdiagramm.....	18
3.1.6 Formel von Bayes.....	18
3.1.7 Vierfeldertafel.....	18
3.2 Kombinatorik.....	19
3.2.1 Anzahl der k-Tupel aus einer n-Menge.....	19
3.2.2 Anzahl der Permutationen einer n-Menge.....	19
3.2.3 Anzahl der k-Teilmengen aus einer n-Menge.....	21
3.2.4 Permutation mit Wiederholung.....	21
3.3 Binomialverteilung und Binomialkoeffizient.....	22
3.4 Normalverteilung.....	23
3.4.1 Dichtefunktion $d(x)$ der Normalverteilung.....	24
3.4.2 Verteilungsfunktion $F(x)$ der Normalverteilung.....	24
3.4.3 Standardnormalverteilung.....	25
3.4.4 Poissonverteilung.....	25
3.4.5 Approximation von Verteilungen.....	26
3.4.5.1 Approximation der Binomialverteilung.....	26
3.4.5.2 Approximation der Poissonverteilung.....	26

4 Beurteilende Statistik.....	26
4.1 Grundlagen.....	26
4.1.1 Anwendung.....	26
4.1.2 Begriffe.....	26
4.1.3 Methoden zur Stichprobengewinnung.....	26
4.2 Stichprobenverteilung.....	27
4.3 Intervallschätzung.....	28
4.4 Testverfahren.....	28
4.4.1 Grundlagen.....	28
4.4.2 Parametertest.....	28
4.4.3 Verteilungstest	29
4.4.4 Berechnung des notwendigen Stichprobenumfangs.....	30
4.4.4.1 Berechnung des Vertrauensintervalls.....	30
5 Anlagen.....	31

1 Grundlagen

1.1 Statistische Begriffe

Bestandsmasse	Zeitpunkt, z.B. Lagerbestand Kontostand
Bewegungsmasse	Zeitraum, z.B. Anzahl der Auslagerungen aus Lagerbestand innerhalb einer Stunde

1.2 Statistische Maßzahlen

Nominalskala	Direkte Verschiedenheit ohne Reihenfolge, z.B. männlich/weiblich
Ordinalskala	Natürliche Rangordnung, Abstände unwichtig, z.B. Rangliste, Noten (1,2,3,4,5)
Intervallskala	Immer Zahlen, Abstände vergleichbar, 0-Punkt kann/wird von Menschen festgelegt.
Verhältnisskala	Intervallskala mit natürlichen Nullpunkt, z.B. Alter, Gewicht, Einkommen.
Absolutskala	Verhältnisskala mit natürlicher Einheit, z.B. Stückzahlen

2 Beschreibende Statistik

- Aufnahme und Charakterisierung von Messwerten einer bestimmten **Grundgesamtheit** (Population, statische Masse, Merkmalsträger); e_k
- Menge von Elementen, die unter einen vom Untersuchungsziel her gesehenen Gesichtspunkt (Merkmal) gleichartig sind (a_k)

1. Beispiel:

- Grundgesamtheit Inhaber von KFZ e_k
- mit Merkmal Fahrzeugtyp a_k

2. Beispiel:

- von Elementen einer Grundgesamtheit lassen sich auch mehrere Merkmale (a, b, c, d) untersuchen:
- Grundgesamtheit Bevölkerung BRD: e_1, e_2, e_3, \dots
- Merkmale: $a_k=1 \dots N$ Geschlecht
 $b_k=1 \dots N$ Familienstand

Unterscheidung in:

Qualitative Merkmale	Eigenschaft wie Geschlecht, Familienstand
Quantitative Merkmale	Zahlen wie Alter, Gewicht, müssen sich auf numerischer Skala darstellen lassen

2.1 Häufigkeitsverteilungen

2.1.1 Absolute und relative Häufigkeit

Messwerte von quantitativen Merkmalen fasst man in Kategorien zusammen

Relative Häufigkeit: $h_i = \frac{f_i}{N}$ $f_1 + f_2 + f_3 + \dots + f_m = N$	Summe aller relativen Häufigkeiten ist 1: $\sum_{i=1}^n h_i = 1$
---	---

Beispiel:

Nr _i Kategorie	x _i Merkmalsrealisation	f _i Absolute Häufigkeit	h _i Relative Häufigkeit	h _i in Prozent
1	Raucher	5	$h_i = 5 \div 31 = 0,161$	$= h_i \cdot 100 = 16,1\%$
2	Nichtraucher	26	0,839	84%
		$\sum = 31 = N$	$\sum = 1$	$\sum = 100\%$

Bei quantitativen Merkmalen fasst man gleiche Messwerte zu einer Merkmalsrealisation zusammen, z.B. Zusammenfassung der Personen mit einem bestimmten Alter.	Beispiel: Zeitdauer Testaufgabe von 32 Versuchspersonen		
	Zeit (Minuten)	absolute Häufigkeit f _i	relative Häufigkeit h _i
	1	2	$= 2 \div 32 = 0,0625$
	2	5	0,156
	3	11	0,344
	4	9	0,281
	5	2	0,0625
6	3	0,094	
		N=32	$\sum = 1$

Häufigkeitsverteilung mit 2 Merkmalen: zweidimensional / bivariational	Beispiel: Zusammenhang zwischen Note in Analysis und Statistik						
	<i>Note in Statistik</i>						
			2	3	4	5	6
	<i>Note in Analysis</i>	1					2
		2	3	5	2		10
		3	2	3	1		6
		4	1	1	5	3	10
		5		1	1	2	5
		6			1		1
			6	10	11	5	1

2.2 Absolute und relative Summenhäufigkeit

Absolute Summenhäufigkeit: $F_i = f_1 + f_2 + \dots + f_i = \sum_{j=1}^i f_j$	Relative Summenhäufigkeit: $H_i = h_1 + h_2 + \dots + h_i = \sum_{j=1}^i h_j$
--	--

Beispiel: Altersverteilung eines Kurses

Alter Merkmals- realisation	f_i	h_i	h_i in %	i	F_i Absolute Summenh.	H_i Relative Summeh.
19	1	0,032	3,2	1	1	0,032
20	5	0,161	16,1	2	6	0,193
21	12	0,387	38,7	3	18	0,58
22	9	0,290	29	4	27	0,87
23	1	0,032	3,2	5	28	0,902
24	1	0,032	3,2	6	29	0,934
25	1	0,032	3,2	7	30	0,966
26	1	0,032	3,2	8	31 = N	0,998 = 1
	$\sum = 31$	$\sum = 1$	$\sum = 100\%$			

Anwendungsbeispiele:

Relative Häufigkeit der Studenten zwischen 20 und 24	$H(20 \leq x \leq 24) = f_2 + f_3 + f_4 + f_5 + f_6 =$ $= 0,161 + 0,387 + 0,290 + 0,032 + 0,032 = \underline{0,896}$
Rel. Häufigkeit der Studenten die mindestens 24 Jahre sind	$H(x \geq 24) = 1 - H(x \leq 23) = 1 - 0,902 = \underline{0,098}$

2.3 Statistische Maßzahlen

2.3.1 Modalwert (Modus) D

- Wert der am **Häufigsten vorkommt**
- Beispiel: 3, 4, 4, 5, 5, 5, 5, 6, 6 → D = 5
- Beispiel: 3, 4, 4, 5, 5, 5, 6, 6, 6 → D = 5,5 (5 und 6 kommen gleich oft vor)

2.3.2 Median / Zentralwert Z

- Anordnung der Messwerte **der Größe nach**, anschließend Bestimmung mittlerer Wert
- Dabei gilt:

Anzahl der Messwerte „n“	Stelle des mittleren Wertes
n – ungeradzahlig	$\frac{n+1}{2}$
n – geradzahlig	zwischen $\frac{n}{2}$ und $\frac{n}{2} + 1 \rightarrow \frac{(x_{n/2} + x_{n/2+1})}{2}$

Beispiel 1 Teststrecke:	15 min, 18 min, 19 min, 20 min, 22 min, 23 min, 89 min → 7 Werte → ungerade → $\frac{7+1}{2} = 4$ → z = 20 min
Beispiel 2:	1, 1, 2, 2, 2, 3, 4, 7, 8, 8, 9, 10 → 12 Werte → gerade → 6. und 7. Wert → z = $\frac{(3+4)}{2} = 3,5$

2.3.3 Mittelwerte

2.3.3.1 Arithmetischer Mittelwert

$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$	$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n f_i \cdot x_i = \sum_{i=1}^n h_i \cdot x_i$
--	---

Beispiel: Noten Analysisklausur				$\bar{x} = \frac{1}{30} \cdot (9 \cdot 1 + 6 \cdot 2 + 6 \cdot 3 + 6 \cdot 4 + 3 \cdot 5 + 0 \cdot 6)$ $= \underline{\underline{2,6}}$ $\bar{x} = 0,3 \cdot 1 + 0,2 \cdot 2 + 0,2 \cdot 3 + 0,2 \cdot 4 + 0,1 \cdot 5 + 0,0 \cdot 6$ $= \underline{\underline{2,6}}$
i	x_i	f_i	h_i	
1	1	9	0,3	
2	2	6	0,2	
3	3	6	0,2	
4	4	6	0,2	
5	5	3	0,1	
6	6	0	0	
		$\sum 30$	$\sum 1$	

Eigenschaften des arithmetischen Mittelwertes:

- Summe der Abweichung der Merkmalswerte x_i vom Mittelwert ist immer 0:
- hohe Empfindlichkeit gegen „Ausreißer“

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

2.3.3.2 Geometrischer Mittelwert

Anwendung:

- Wachstumsvorgänge, z.B. Verzinsung von Kapital, Bevölkerungswachstum

Multiplikation der einzelnen Werte: $\bar{x}_g = \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n}$ Merkmalswert hoch absolute Häufigkeit: $\bar{x}_g = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}}$	Beispiel: Noten Analysisklausur $\bar{x}_g = \sqrt[30]{1^9 \cdot 2^6 \cdot 3^6 \cdot 4^6 \cdot 5^3 \cdot 6^0} = \underline{\underline{2,22}}$
---	---

2.3.3.3 Harmonischer Mittelwert

Einzelne Werte: $\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{1}{a_i}}$	Merkmalswert / Absolute Häufigkeit: $\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{f_i}{x_i}}$
Beispiel Analysisklausur: $\bar{x}_h = \frac{n}{\frac{9}{1} + \frac{6}{2} + \frac{6}{3} + \frac{6}{4} + \frac{3}{5} + \frac{0}{6}} = \underline{1,86}$	Anwendung z.B. bei Geschwindigkeiten

2.3.4 Streuungsparameter

2.3.4.1 Spannweite d

d = Differenz zwischen größten und kleinsten Merkmalswert	$d = a_{imax} - a_{imin}$
---	---------------------------

2.3.4.2 Mittlere lineare (absolute) Abweichung

Einzelne Werte und Mittelwert: $\delta = \frac{1}{n} \sum_{i=1}^k a_i - \bar{x} $ Absolute Häufigkeit / Merkmalswert: $\delta = \frac{1}{n} \sum_{i=1}^k f_i \cdot x_i - \bar{x} $ Relative Häufigkeit / Merkmalswert: $\delta = \sum_{i=1}^k h_i \cdot x_i - \bar{x} $	Arithmetischer Mittelwert der absoluten Abweichung zwischen Merkmalswert und dem Mittelwert.
Beispiel: Merkmalswerte: 3, 7, 8, 9, 13 $\bar{x} = 8$ $\delta = \frac{1}{n} \sum_{i=1}^k a_i - \bar{x} $ $= \frac{1}{5} \cdot (5 + 1 + 0 + 1 + 5) = \underline{2,4}$	

2.3.4.3 Varianz und Standardabweichung

$$\sigma^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^k (a_i - \bar{x})^2 \quad \text{einzelne Werte}$$

$$\sigma^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^k f_i \cdot (x_i - \bar{x})^2 \quad \text{absolute H\u00e4ufigkeit / Merkmalswert}$$

Wichtigster Streuungsparameter!

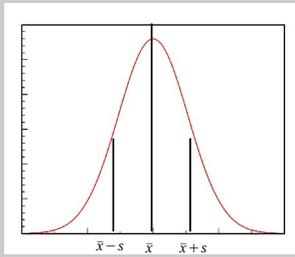
Varianz σ^2, s^2 = arithmetischer Mittelwert der quadratischen Abweichung der Merkmalswerte vom arithmetischen Mittelwert

Beispiel: Lieferung von Tabletten, 50 werden zuf\u00e4llig herausgenommen und die Masse in Gramm gewogen

Masse x_i	Anzahl f_i	$x_i \cdot f_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
0,9	3	2,7	-0,2	0,04	0,12
1,0	9	9	-0,1	0,01	0,09
1,1	24	26,4	0	0	0
1,2	12	14,4	0,1	0,01	0,12
1,3	2	2,6	0,2	0,04	0,08
	$\Sigma = 50$				$\Sigma = 0,41$

$\bar{x} = \frac{1}{50} \sum_{i=1}^5 x_i \cdot f_i = 1,102 \approx 1,1$

Sinnvoll Aufstellen einer Tabelle und Ausrechnen in Einzelschritten

$$\sigma^2 = s^2 = \frac{1}{50-1} \cdot 0,41 = \underline{0,00837}$$


2.3.4.4 Varianzkoeffizient

$v = \frac{s}{\bar{x}} = \frac{\sigma}{\bar{x}}$	Gibt die relative Streuung an
--	-------------------------------

2.3.5 Konzentrationsma\u00dfe

- Konzentrationsmerkmal = Merkmal bei dem Bildung einer **Merkmalssumme** m\u00f6glich bzw. sinnvoll ist

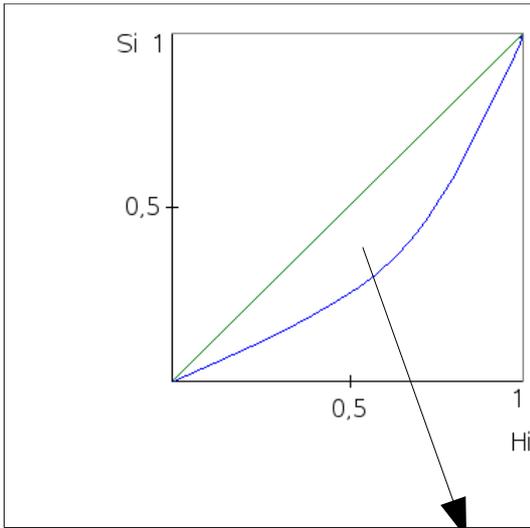
2.3.5.1 Lorenzkurve

- zur grafischen Darstellung von statistischen Verteilungen und der Veranschaulichung des Ausma\u00dfes an Konzentration bzw. Ungleichheit Sie wird insbesondere zur Analyse der Einkommensverteilung verwendet.

Beispiel: Einkommenssituation

i	Bruttoeinkommen x_i	f_i	h_i	$f_i \cdot x_i$	H_i^1	$\sum f_i \cdot x_i$	S_i^2
1	30.000	25	0,25	750.000	0,25	750.000	0,15
2	40.000	30	0,30	1.200.000	0,55	1.950.000	0,39
3	50.000	25	0,25	1.250.000	0,8	3.200.000	0,64
4	80.000	10	0,10	800.000	0,9	4.000.000	0,80
5	100.000	10	0,10	1.000.000	1,0	5.000.000	1,00
		$\Sigma 100$	$\Sigma 1$	$\Sigma 5.000.000$			

1 relative Summenh\u00e4ufigkeit (Aussummierung von h_i)
 2 kumulierter relativer Merkmalswert (hier aufsummierter Wert durch Gesamtgeld, Spalte davor)



Anwendung der Lorenzkurve:

- Veranschaulichung der Einkommensverteilung – z.B. verfügen 50% der Haushalte (und zwar die ärmeren) über ca. 25 % aller Einkommen, die (reicheren) 50% verfügen hier über ca.. 75% der Einkommen

2.3.6 Ginikoeffizient und Konzentrationsfläche

- je größer die Konzentrationsfläche, desto größer die Konzentration, max. Fläche = max. Konzentration, im Beispiel würde einer alles verdienen

$\text{Ginikoeffizient} = \frac{\text{Konzentrationsfläche}}{F_{\max}}$ $G = 1 - \sum_{i=1}^k h_i \cdot (s_i - s_{(i-1)}) \quad 0 \leq G \leq 1$	$F_{\max} = \frac{1}{2}$ $F_{\min} = 0$
Beispiel: $G = 1 - [0,25 \cdot (0,15 - 0) + 0,3 \cdot (0,39 - 0,15) + 0,25 \cdot (0,64 - 0,39) + 0,1 \cdot (0,8 - 0,64) + 0,1 \cdot (1,0 - 0,8)] = \underline{0,22}$	

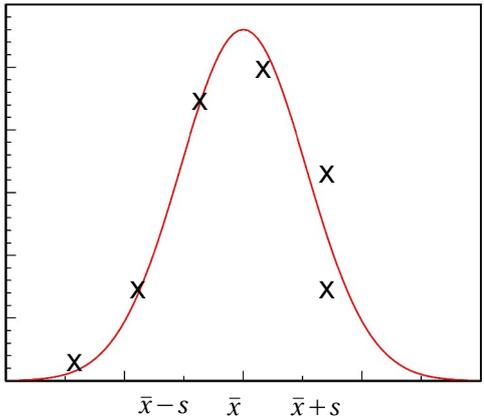
2.3.7 Fehlerrechnung – Messfehler in der Physik

Systematische Fehler	Zufällige Fehler
<ul style="list-style-type: none"> ● beeinflussen Messergebnisse immer gleich, z.B. Vorzeichenfehler 	<ul style="list-style-type: none"> ● statistischer Charakter ● schwanken bei gleicher Experimentieranordnung innerhalb eines Intervalls um einen Mittelwert
Ursachen: <ul style="list-style-type: none"> ● falsche Eichung Messgerät ● Alterung, Überlastung, Streufelder 	Ursachen: <ul style="list-style-type: none"> ● ungenaues Ablesen, Reaktionszeit des Beobachters ● Umwelteinflüsse, z.B. Temperaturschwankungen, Luftdruck, Spannungsschwankungen

2.3.7.1 Fehlerrechnung für zufällige Fehler

- Größe x wird n-mal gemessen → Messreihe x_1, x_2, x_3, \dots
- Annäherung an den wahren Wert: arithmetischer Mittelwert \bar{x} (siehe S. 7)
- Bewertung des Resultats durch Charakterisierung des Intervalls in dem die einzelnen Messwerte liegen und um den Mittelwert streuen

2.3.7.2 **Gaußsche Glockenkurve**

	$f(x) = \frac{1}{\sqrt{2\pi} \cdot s} \cdot e^{-\frac{1}{2} \cdot \frac{(x-\bar{x})^2}{s^2}}$ <p>s = Standardabweichung der Messreihe</p> <p>Eigenschaften der Glockenkurve:</p> <ul style="list-style-type: none"> ● Maximum bei \bar{x} ● Wendepunkt bei $\bar{x}-s$; $\bar{x}+s$ <p>Häufigkeitsverteilungsfunktion $f(x)$ ist auf 1 normiert:</p> $\int_{-\infty}^{+\infty} f(x) dx = 1$ <p>d.h. im Intervall $[-\infty, \infty]$ muss die Gesamtheit aller Messwerte liegen.</p>
<p>Wieviele Messwerte liegen im Intervall $[\bar{x}-s; \bar{x}+s]$?</p> $P = \int_{\bar{x}-s}^{\bar{x}+s} f(x) dx$	<p>$[\bar{x}-s; \bar{x}+s]$: P=68,3%</p> <p>$[\bar{x}-2s; \bar{x}+2s]$: P=95,4% (Industriestandard)</p> <p>$[\bar{x}-3s; \bar{x}+3s]$: P=99,7% (z.B. für Konstanten)</p> <p>z.B. $P = 68,3\%$, d.h. 68,3 % aller Messwerte liegen im Intervall, bzw. beliebiger Messwert liegt mit Wahrscheinlichkeit von 68,3% in dem Intervall.</p>
<p>Standardabweichung berechnen:</p> $s = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$	<p>x_i = Einzelwert</p> <p>\bar{x} = arith. Mittelwert</p> <p>n = Anzahl Messwerte</p>

2.3.7.3 **Messunsicherheit und Vertrauensbereich**

<p>Messunsicherheit $u = \frac{s}{\sqrt{n}}$</p>	<p>Mit s kann Aussage gemacht werden, in welchem Bereich um \bar{x} der wahre Wert x_u zu finden sein wird.</p>
<p>Vertrauensgrenzen:</p> <p>$\bar{x} - u$: untere Vertrauensgrenze</p> <p>$\bar{x} + u$: obere Vertrauensgrenze</p>	<p>Mit steigender Anzahl der Messwerte n wird Differenz zwischen \bar{x} und x_u kleiner.</p>
<p>$[\bar{x} - u; \bar{x} + u]$ Vertrauensbereich der Mittelwerte</p>	

Praxis: Für praktische Messungen sollte $n \geq 5$ sein.

- $[\bar{x} - u; \bar{x} + u]$: P=68,3% übliche naturwissenschaftliche Experimente
- $[\bar{x} - 2u; \bar{x} + 2u]$: P=95,4% Industriestandard
- $[\bar{x} - 3u; \bar{x} + 3u]$: P=99,7% internationale physikalische Konstanten

Beispiel Messprotokoll für Zeitmessung:		Gesucht: Intervallgrenzen mit denen der wahre Wert mit einer statistischen Sicherheit von 68,3 % abweicht. $\bar{t} = 50,6$ $s = 1,483$ $u = 0,663$ $\bar{x} - u = 49,94$ $\bar{x} + u = 51,26$
n	x_i t in s	
1	52,5	
2	49,6	
3	50,3	
4	51,7	
5	48,9	

2.3.7.4 Fehlerfortpflanzung

Z.B. Volumen Kreiskegelstumpf: $V = \frac{\pi}{3} h \cdot (R^2 + R \cdot r + r^2)$	Fehler können hier bei mehreren Parametern als Messfehler auftreten → Fehlerfortpflanzung.
---	--

2.3.7.5 Größtfehlerabschätzung

Sofern gilt: $\Delta x_i \ll x_i \rightarrow f(x_1 + \Delta x_1; \dots; x_m + \Delta x_m) \approx f(x_1 \dots x_m) + \sum_{i=1}^m \frac{df(x)}{dx_i} \cdot \Delta x_i$

Beispiel: Messung von g auf Fallhöhe h und Fallzeit: $g = \frac{2h}{t^2}$ Δh : Schätzfehler Meterstab Δt : Schätzfehler Stoppuhr	Rechnung: $\frac{dg}{dh} = \frac{2}{t^2}$ $\frac{dg}{dt} = -\frac{4h}{t^3}$ $\rightarrow \Delta g = \mp \left(\frac{2}{t^2} \cdot \Delta h + \frac{4h}{t^3} \cdot \Delta t \right)$
---	---

2.3.7.6 Relative und prozentuale Fehlerangabe

Absolute Fehler: <ul style="list-style-type: none"> ● Standardabweichung S ● Messunsicherheit u 	Relative Fehler: $\frac{\text{absolute Fehler}}{\text{Mittelwert}}$ $\delta f(x) = \frac{u_{f(x)}}{f(x)}$ $\delta y = \frac{u_x}{\bar{y}}$ Prozentuale Fehler: $= \delta f(x) \cdot 100\%$
--	---

Ist Funktion $f(x)$ ein **Potenzprodukt (und nur dann)**, z.B. $y = f(x_1, x_2, x_3) = const \cdot x_1^\alpha \cdot x_2^\beta \cdot x_3^\gamma$:

$$\delta y = \delta f(x_1, x_2, x_3) = \pm (|\alpha \cdot \delta x_1| + |\beta \cdot \delta x_2| + |\gamma \cdot \delta x_3|)$$

2.3.7.7 Angabe des Ergebnisses

- Absolutfehlerangabe : Resultat = $\bar{y} \pm u_y$ bzw. $\bar{y} \pm \Delta y$
- Relativfehlerangabe : Resultat = $\bar{y} \cdot (1 \pm \delta y \cdot 100)$

2.3.7.8 Fehlermessung bei analogen Messgeräten

Fehlerklasse = relativer Fehler in % bezogen auf Endausschlag

Beispiel: Strommessgerät, Messbereich 6 A, Fehlerklasse 1,5

ΔI bei 6 A $\rightarrow \Delta I = \pm 0,09A$ relativer Messfehler über gesamten Messbereich

ΔI bei 2 A $\rightarrow \Delta I = \frac{0,09A}{2A} = 0,045 = 4,5\%$

ΔI bei 0,5 A $\rightarrow \Delta I = \frac{0,09A}{0,5} = 0,18 = 18\%$

3 Wahrscheinlichkeitsrechnung

3.1 Grundbegriffe

Zufallsexperiment	<ul style="list-style-type: none"> ● kann unter gleichen Bedingungen beliebig oft wiederholt werden ● mehrere Ergebnisse möglich und Ergebnis hängt vom Zufall ab
Elementarereignisse	<ul style="list-style-type: none"> ● mögliche Ergebnisse des Zufallsexperimentes
Ergebnismenge	<ul style="list-style-type: none"> ● Gesamtheit aller möglichen Ergebnisse, z.B. Würfel $\Omega = \{1,2,3,4,5,6\}$
Ereignis	<ul style="list-style-type: none"> ● Teilmenge A der Ergebnismenge, z.B. gerade Zahl Würfeln $A = \{2,4,6\}$
Sicheres Ereignis	<ul style="list-style-type: none"> ● Menge die alle Ergebnisse des Zufallsexperimentes enthält z.B. Ω
Unmögliches Ereignis	<ul style="list-style-type: none"> ● Ereignis das nicht eintreten kann, z.B. 7 oder 8 beim Würfeln

Beispiel: Urne mit roter (r), weißer (w) und zwei grünen Kugeln (g_1, g_2), Zufallsexperiment: eine Kugel aus Urne ziehen

Ergebnismenge	$\Omega = \{r, w, g_1, g_2\}$
Elementarereignisse	$\{r\}, \{w\}, \{g_1\}, \{g_2\}$
sicheres Ereignis	Ω
weitere Ereignisse	grüne Kugel $\{g_1, g_2\}$; schwarze Kugel $\{\}$, grün, weiße Kugel $\{w, g_1, g_2\}$

3.1.1 Axiome der Wahrscheinlichkeitsrechnung

- Jedem Ereignis eines Zufallsexperimentes ist eine eindeutige bestimmte Zahl $P(A) = R^+ \cup \{0\}$ zugeordnet. Diese Zahl wird **Wahrscheinlichkeit P** für das Ereignis A genannt.
- Wahrscheinlichkeit für ein sicheres Ereignis ist immer 1.
- $A \cap B = \{\}$ → $P(A \cup B) = P(A) + P(B)$ (Additivität)
- Rechenregeln:

$P(\bar{A}) = 1 - P(A)$	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(A_1 \cup A_2 \dots A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$ falls $A_1, A_2 \dots A_n$ unvereinbare Ereignisse sind, d.h. $A_1 \cap A_2 \cap \dots A_n = \{\}$	A, B sind unabhängige Ereignisse: $P(A \cap B) = P(A) \cdot P(B)$

3.1.2 Gesetz der großen Zahlen

- Je mehr Versuche durchgeführt werden, desto genauer wird die relative Häufigkeit h der Wahrscheinlichkeit P
- Beispiel Würfeln: je großer Anzahl Würfe, desto mehr stellt sich relative Häufigkeit für eine bestimmte Zahl bei $P = \frac{1}{6}$ ein

3.1.2.1 Erwartungswert E, μ

Bei Wahrscheinlichkeit P gibt es ebenfalls einen arithmetischen Mittelwert, der Erwartungswert (μ, E),	Beispiel Würfel: $P(1) = \frac{1}{6}; P(2) = \frac{1}{6}; \dots; P(6) = \frac{1}{6}$ $E = \mu = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6$ $= \frac{21}{6} = \underline{\underline{3,5}}$
--	--

3.1.3 Laplace-Experiment und die Wahrscheinlichkeit

Ein Laplace-Experiment meint ein Zufallsexperiment bei dem davon ausgegangen wird, dass jeder Versuchsausgang gleichwahrscheinlich ist.	$P(A) = \frac{\text{Anzahl der Elemente aus A}}{\text{Anzahl der Elemente aus } \mu}$ <p>Beispiel Würfeln 1: Ergebnismenge ungerade Augenzahl: $A = \{1, 3, 5\}$</p> $P(A) = g \frac{(A)}{\mu} = \frac{3}{6} = \underline{\underline{\frac{1}{2}}}$ <p>Beispiel Würfeln 2: Ergebnismenge ungerade Augenzahl oder 6: $A = \{1, 3, 5, 6\}$</p> $B \{1, 3, 5\}; P(B) = \frac{1}{2}$ $C \{6\}; P(C) = \frac{1}{6}$ $P(A) = P(B \cup C) = P(B) + P(C) = \frac{1}{2} + \frac{1}{6} = \underline{\underline{\frac{2}{3}}}$
---	---

3.1.3.1 Stochastische Unabhängigkeit

Definition: 2 Ereignisse A und B sind stochastisch unabhängig, wenn gilt: $P(A \cap B) = P(A) \cdot P(B)$	Zwei Ereignisse sind stochastisch unabhängig, wenn die Wahrscheinlichkeit des zweiten Ereignisses nicht vom ersten abhängt. Beispiel 1: Zweimal Würfeln, mit den Ereignissen A = 3 und B = 5, der zweite Wurf ist nicht abhängig vom Ergebnis des ersten. Beispiel 2: Dreimal Würfeln, A = gerade Zahl, B = 5, C = Summe von A und B = 12. A und B sind stochastisch unabhängig, C ist von A und B abhängig.
--	--

3.1.4 Bedingte Wahrscheinlichkeit

Definition: Wahrscheinlichkeit für Eintreten von B, falls bekannt ist, dass A eingetreten ist, wird als bedingte Wahrscheinlichkeit $P(B A)$ bezeichnet.	B unter der Bedingung das A eingetreten ist: $P(B A) = \frac{P(A \cap B)}{P(A)}$ mit $P(A) \neq 0$ A unter der Bedingung das B eingetreten ist: $P(A B) = \frac{P(A \cap B)}{P(B)}$ mit $P(B) \neq 0$ Produktregel: $P(A \cap B) = P(A) \cdot P(B A)$ $P(A \cap B) = P(B) \cdot P(A B)$
Beispiel: Kartenspiel mit 32 Karten, 2 werden nacheinander gezogen A : erster Zug wird ein König B : zweiter Zug wird ein König gesucht: $P(A \cap B)$	Lösung: 4 Könige in 32 Karten: $P(A) = \frac{1}{8}$ 3 Könige aus 31 Karten sind noch da: $P(B A) = \frac{3}{31}$ $P(A \cap B) = P(A) \cdot P(B A) = \frac{1}{8} \cdot \frac{3}{31} = \frac{3}{248}$
Es gilt auch: $P(A \cap B \cap C) = P(A \cap B) \cdot P(C A \cap B) = P(A) \cdot P(B A) \cdot P(C (A \cap B))$	
Beispiel: Lotto „6 aus 49“, Ziehung in 6 Teilschritten gesucht: P das 6 richtige gezogen werden Lösung: A ₁ : Erste Zahl richtig A ₂ : Zweite Zahl richtig ... A ₆ : Sechste Zahl richtig $P(A_1) = \frac{6}{49}$ $P(A_2) = \frac{5}{48}$... $P(A_6) = \frac{1}{44}$	$P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6)$ $= \frac{6}{49} \cdot \frac{5}{48} \cdot \frac{4}{47} \cdot \frac{3}{46} \cdot \frac{2}{45} \cdot \frac{1}{44}$ $\approx \underline{\underline{0,0000072\%}}$

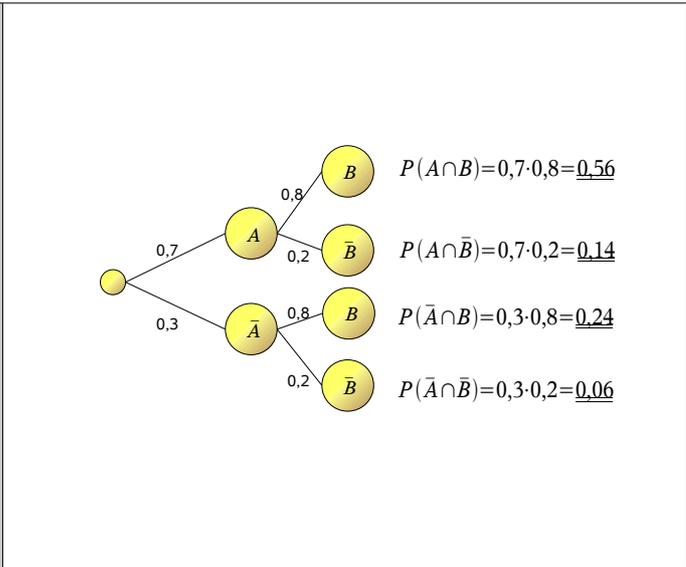
3.1.5 Bedingte Wahrscheinlichkeit im Baumdiagramm

Beispiel:

$P(A) = 0,7$
 $P(B) = 0,8$

gesucht:
 Wahrscheinlichkeit, dass A und B oder A und nicht B auftreten:

$$P(C) = (P(A \cap B) \cup P(A \cap \bar{B}))$$

$$= P(A) \cdot P(B) + P(A) \cdot P(\bar{B}) = 0,56 + 0,14 = \underline{0,7}$$


3.1.6 Formel von Bayes

Beispiel:

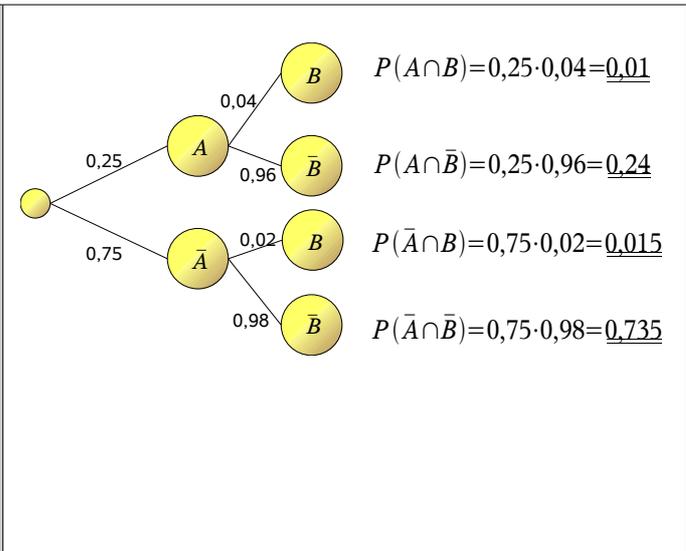
- Maschine M1 stellt Widerstände her mit Ausschuss von 4%.
- Maschine M2 stellt 3 mal soviel Widerstände her mit Ausschuss von 2%
- Wie groß ist P, dass ein defekter Widerstand von M1 stammt?

Ereignisse:

- A : Widerstand von M1
- B : Widerstand ist defekt

$$P(B) = P(A \cap B) + P(\bar{A} \cap B) = \underline{0,025}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})}$$

$$= 0,01 \div 0,025 = \underline{0,4}$$


Satz von Bayes:

$$P(A|B) = \frac{P(A \cap B)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})} \quad \text{mit} \quad P(A \cap B) = P(A) \cdot P(B|A)$$

3.1.7 Vierfeldertafel

Beispiel Fremdsprachen:

Ereignisse:

- E : Student spricht Englisch $P(E) = 0,7$
- F : Student spricht Französisch $P(F) = 0,6$
- S : Student spricht nur Spanisch $P(S) = 0,1$

$P(S) \rightarrow P(\bar{E} \cap \bar{F})$
 $P(\bar{E} \cap \bar{F} \cap S) = 0$

	F	\bar{F}	
E	0,4*	0,3*	0,7
\bar{E}	0,2*	0,1	0,3*
	0,6	0,4*	$\sum 1$

Vorgehen, gegebene Wahrscheinlichkeiten Eintragen und Rest ermitteln*.

Gesucht: P von Student spricht Englisch und Fransösich
 $P(E \cap F) = \underline{0,4}$ → kann aus Tabelle einfach abgelesen werden

Weitere Informationen:

- $P(E) = 0,25$ (nur Englisch)
- $P(F \cap S) = 0,15$ (Franz. und Spanisch)
- $P(E \cap F \cap S) = 0,3$ (alle 3 Sprachen)

	<i>F</i>	<i>F</i>	\bar{F}	\bar{F}	
<i>E</i>	0,3	0,1*	0,25	0,05*	0,7
\bar{E}	0,15	0,05*	0	0,1	0,3
	<i>S</i>	\bar{S}	\bar{S}	<i>S</i>	
	0,6		0,4		$\Sigma 1$

3.2 Kombinatorik

3.2.1 Anzahl der k-Tupel aus einer n-Menge

Reihenfolge	mit
Wiederholung/Zurücklegen	mit

n^k
 n: verschiedene Elemente
 k - Plätze

Beispiel 1: 5-stelliges Zahlenschloss

5	3	2	7	2
---	---	---	---	---

→ 5 Plätze, immer 10 Möglichkeiten

$10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 10^5$ Möglichkeiten

Beispiel 2: Spiel 77

7 – Plätze; 10 Elemente → $10^7 = 10$ Mio. Möglichkeiten

3.2.2 Anzahl der Permutationen einer n-Menge

Reihenfolge	mit
Wiederholung/Zurücklegen	ohne

$n!$ für $k = n$; $0! = 1$
 $n \frac{n!}{(n-k)!}$ für $k < n$
 n: verschiedene Elemente
 k - Plätze

Beispiel 1: 5 Redner, wie viele Möglichkeiten der Reihenfolge gibt es?

- 1. Rede : 5 Möglichkeiten
- 2. Rede : 4 Möglichkeiten
- ...
- 5. Rede : 1 Möglichkeit

$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 6!$ Möglichkeiten

Beispiel 2: 5 Kandidaten für ein Gewinnspiel, 3 werden gezogen, 1. Platz, 2. Platz, 3. Platz

- 1. Platz : 5 Möglichkeiten
- 2. Platz : 4 Möglichkeiten
- 3. Platz : 3 Möglichkeiten

$5 \cdot 4 \cdot 3 = 60$ Möglichkeiten
 $= \frac{n!}{(n-k)!} = \frac{5!}{(5-2)!} = \frac{120}{2} = \underline{60}$

3.2.3 Anzahl der k-Teilmengen aus einer n-Menge

Reihenfolge	ohne	Beispiel: 5 Kandidaten für ein Gewinnspiel, alle gewinnen den selben Preis
Wiederholung/Zurücklegen	ohne	
$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$ <p>n: verschiedene Elemente k - Plätze</p>		$\frac{5!}{(5-3)! \cdot 3!} = \underline{10}$

3.2.4 Permutation mit Wiederholung

<p>Besteht ein n-Tupel aus k-verschiedenen Elementen, die n_1, n_2, \dots, n_k-mal vorkommen mit: $n_1 + n_2 + \dots + n_k = n$ so gibt es</p> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 10px auto;"> $\frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!} = \text{verschiedene Tupel}$ </div>	<p>Beispiel: Wort „Mississippi“ Auf wieviele Arten kann man die Buchstaben, dass neue Wörter entstehen?</p> <p> $1 \times M : \binom{11}{1} = 11$ Möglichkeiten $4 \times I : \binom{10}{4} = \frac{10!}{6! \cdot 4} = 11$ $4 \times S : \binom{6}{4} = 15$ $2 \times P : \binom{2}{2} = 1$ $\binom{11}{1} \cdot \binom{10}{4} \cdot \binom{6}{4} \cdot \binom{2}{2} = \underline{34650}$ </p> <p style="text-align: center;">→ $\frac{11!}{1! \cdot 4! \cdot 4! \cdot 2!} = \underline{34650}$</p>
---	--

3.3 Binomialverteilung und Binomialkoeffizient

Anwendung	<ul style="list-style-type: none"> Wie viele Treffer bei wie vielen Versuchen erzielt werden.
Voraussetzung	<ul style="list-style-type: none"> Zufallsexperiment darf nur 2 Ausgänge (Ereignisse) haben (A und \bar{A}) muss stochastisch unabhängig sein

<p>Binomialverteilung:</p> $B(n, p, k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$ $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$	<p>n = Anzahl der Einzelversuche p = Wahrscheinlichkeit für Erfolgstreffer der Einzelversuche k = Anzahl der Erfolge (Treffer) q = 1 - p (Gegenwahrscheinlichkeit)</p>
<p>Beispiel: Werfen einer Münze P(K) = 0,5 P(Z) = 0,5</p> <p>Gesucht: Wahrscheinlichkeit, dass bei 3-fachen Werfen Kopf geworfen wird.</p> <p>Ereignis A {(Kopf)}; P(A) = 0,5 ; P(B) = 0,5</p>	<p>Lösung: n = 3 (3 mal werfen)</p> $B(3; 0,5; k) = \binom{3}{k} \cdot p^k \cdot (1-p)^{n-k}$ <p>0 mal Kopf, k = 0 : $B(3; 0,5; 0) = \binom{3}{0} \cdot 0,5^0 \cdot 0,5^3 = 0,125$ 1 mal Kopf, k = 1 : $B(3; 0,5; 1) = \binom{3}{1} \cdot 0,5^1 \cdot 0,5^2 = 0,375$ 2 mal Kopf, k = 2 : $B(3; 0,5; 2) = \binom{3}{2} \cdot 0,5^2 \cdot 0,5^1 = 0,375$ 3 mal Kopf, k = 3 : $B(3; 0,5; 3) = \binom{3}{3} \cdot 0,5^3 \cdot 0,5^0 = 0,125$</p>

Es gilt bei der Binomialverteilung:	
Erwartungswert: $\mu = E = n \cdot p$	Verteilungsfunktion: $F_{n,p}(k) = \sum_{i=0}^k B(n; p; i) = \sum_{i=0}^k \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$
Varianz: $s^2 = n \cdot p \cdot q = n \cdot p \cdot (1-p)$	

3.4 Normalverteilung

<p>Eine Zufallsvariable x, die sich als Summe der n-Zufallsvariablen $x_1 \dots x_n$ ergibt, ist näherungsweise normalverteilt, wenn:</p> <ol style="list-style-type: none"> Anzahl der Zufallsvariablen hinreichend groß ist Zufallsvariablen x_1, x_2, \dots, x_n unabhängig sind nicht eine, aber einige wenige Zufallszahlen $x_1 \dots x_n$ gegenüber den anderen dominieren 	<p>Begriffe: $N(\mu, s)$ – Verteilung</p> $Var(x) = s^2$ <p>Es gilt: $\mu = n \cdot p$ $s = \sqrt{n \cdot p \cdot q}$</p>
--	---

3.4.1 Dichtefunktion d(x) der Normalverteilung

$$d(x) = \frac{1}{\sqrt{2\pi} \cdot s} \cdot e^{-\frac{1}{2} \left(\frac{x-\mu}{s}\right)^2} \quad x \in \mathbb{R}; s > 0$$

1. $d(x) \geq 0$
2. $\int_{-\infty}^{\infty} d(x) dx = 1$
3. Maximum: $x_m = \mu$ (μ = Mittelwert)
4. Wendepunkte: $x_w = \mu \pm s$
5. symmetrisch zu $x = \mu$
6. $d(x)$ ist stetig

Dichtefunktionen von normalverteilten Zufallsgrößen mit unterschiedlichen Parametern

Fläche, welche von der Dichtefunktion $d(x)$ der Normalverteilung zwischen den Grenzen $x_1 = \mu - k \cdot s$ und $x_2 = \mu + k \cdot s$ ($k > 0$) eingeschlossen wird, ist für alle Normalverteilungen gleich groß!

k	Intervallgrenzen	Anteil der Teilfläche
1	$\mu \pm 1 \cdot s$	68,26 %
2	$\mu \pm 2 \cdot s$	95,44 %
3	$\mu \pm 3 \cdot s$	99,73 %
4	$\mu \pm 4 \cdot s$	99,99 %
1,64	$\mu \pm 1,64 \cdot s$	90 %
1,96	$\mu \pm 1,96 \cdot s$	95 %
2,58	$\mu \pm 2,58 \cdot s$	99 %
2,81	$\mu \pm 2,81 \cdot s$	99,5 %

3.4.2 Verteilungsfunktion F(x) der Normalverteilung

$F(x)$ gibt die Wahrscheinlichkeit dafür an, dass die Zufallsvariable **höchstens den Wert x** annimmt.

Es gilt:

1. $F(x)$ = monoton steigend
2. $\lim_{x \rightarrow \infty} F(x) = 1$
3. $\lim_{x \rightarrow -\infty} F(x) = 0$

Anwendung:
Wahrscheinlichkeiten wie z.B. mindestens 3, höchstens 3, $P(x \leq 3)$, hier werden die Einzelwahrscheinlichkeiten aufsummiert.

3.4.3 Standardnormalverteilung

Wenn $\mu=0$; $s=1$; $N(0,1)$ (Standardnormalverteilung)

Jede Normalverteilung $N(\mu, s)$ kann auf die Standardnormalverteilung transformiert werden.

Es sei x eine $N(\mu, s)$ verteilte Zufallsvariable. Dann ist die Zufallsvariable:

$$Z = \frac{x - \mu}{s} \quad N(0,1) \text{ verteilt.}$$

Folgende Beziehung gilt zwischen $F(x)$ der $N(\mu, s)$ Verteilung und $F_0(x)$ der $N(0,1)$ Verteilung:

$$F(x) = F_0\left(\frac{x - \mu}{s}\right)$$

→ die Werte der Standardnormalverteilung können aus Tabellen entnommen werden!

Beispiel: Zufallsvariable x ist $N(10,3)$ verteilt
 $10 =$ Erwartungswert
 $3 =$ Standardabweichung

Ermittlung von folgenden Wahrscheinlichkeiten:

$$P(x \leq 16) = F(16) = F_0\left(\frac{x - \mu}{s}\right) =$$

$$F_0\left(\frac{16 - 10}{3}\right) = F_0(2) = \underline{0,9772}$$

(stetige Funktion, daher wird der Grenzwert immer mit eingesetzt)

$$P(x \geq 12) = 1 - F(12) = 1 - F_0\left(\frac{x - \mu}{s}\right) =$$

$$1 - F_0\left(\frac{12 - 10}{3}\right) = F_0(2) = \underline{0,2514}$$

$$P(7 \leq x \leq 13) = F(13) - F(7) =$$

$$F_0(1) - F_0(-1) = \underline{0,6826}$$

3.4.4 Poissonverteilung

Ereignisse die relativ selten auftreten, d.h. geringe Wahrscheinlichkeit genügen der Poissonverteilung.

Ein Ereignis E tritt in einem gegebenen Zeitraum μ mal auf → $P_s(\mu)$

Wahrscheinlichkeit $f(x)$ und Verteilungsfunktion $F(x)$ sind in Tabellen festgehalten.

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

Beispiel: ein Telefon klingelt pro Minute einmal (Poisson verteilt), $P_s(1)$

a) genau ein Anruf:

$$P_s(x=1) = f(1) = \frac{\mu^1}{1!} \cdot e^{-\mu} = \frac{1^1}{1} \cdot e^{-1} = \underline{0,3679}$$

b) höchstens ein Anruf:

$$P_s(x \leq 1) = F(1) = \underline{0,7358}$$

c) mindestens ein Anruf:

$$P_s(x \geq 1) = 1 - P_s(x=0) = 1 - F(0) = 1 - 0,3679 = \underline{0,6331}$$

d) zwei oder drei Anrufe

$$P_s(x=2) \vee P_s(x_3) = f(2) + f(3) = 0,1839 + 0,0613 = \underline{0,2452}$$

e) in 5 min klingelt das Telefon genau 6 mal

$$P_s(5) ; \mu=5 ; p=6$$

$$P(x=6) = \underline{0,1462}$$

3.4.5 Approximation von Verteilungen

3.4.5.1 Approximation der Binomialverteilung

<p>1. für $n \cdot p \leq 10$ und $n \geq 1500 \cdot p$ ist eine $B(n,p,k)$ verteilte Zufallsvariable näherungsweise $P_s(m,p)$ verteilt</p> <p>2. für $n \cdot p \cdot (1-p) > 9$ ist eine $B(n,p,k)$ verteilte Zufallsvariable näherungsweise $N(n \cdot p; \sqrt{n \cdot p \cdot (1-p)})$ verteilt</p>	<p>Beispiel: 100 mal Münze werfen, wie groß ist die Wahrscheinlichkeit, dass höchstens 40 mal „Wappen“ auftritt.</p> <p>Lösung: Zufallsvariable ist $B(100; 0,5; k)$ verteilt ($k = 0$ bis 40)</p> $n \cdot p \cdot (1-p) = 100 \cdot 0,5 \cdot (1-0,5) = 2,5 > 9$ <p>→ kann angenähert werden durch Normalverteilung</p> $\mu = n \cdot p = 100 \cdot 0,5 = 50 \quad ; \quad s = \sqrt{100 \cdot 0,5 \cdot 0,5} = 5$ <p>→ $N(50,5) = F(40) = F_0\left(\frac{40-50}{5}\right) = F_0(-2) = \underline{\underline{0,0228}}$</p>
---	---

3.4.5.2 Approximation der Poissonverteilung

<p>Für $\mu \geq 10$ ist $P_s(\mu)$ -verteilte Zufallsvariable $N(\mu; \sqrt{\mu})$ verteilt.</p>	<p>Beispiel: In eine Apotheke kommen pro Stunde im Durchschnitt 25 Kunden. Zufallsvariable ist Poissonverteilt. Wie groß ist Wahrscheinlichkeit, dass in einer bestimmten Stunde maximal 20 Kunden kommen.</p> <p>Lösung:</p> $\mu = 25 \quad ; \quad p = \frac{25}{20} = 0,8$ $N(25; 5) = F(20) = F_0\left(\frac{20-25}{5}\right) = F_0(-1) = \underline{\underline{0,1587}}$
--	--

4 Beurteilende Statistik

4.1 Grundlagen

4.1.1 Anwendung

- Lieferant gibt an, dass nur 5% seiner gelieferten Geräte fehlerhaft sind, diese Aussage wird durch **Stichproben** untersucht

4.1.2 Begriffe

Grundgesamtheit	Gesamtmenge des Untersuchungsgegenstandes, z.B. alle Autos einer Lieferung
Stichprobe	konkret untersuchte Menge
Stichprobenumfang	Anzahl der Elemente einer Stichprobe, z.B. 5 Autos aus einer Lieferung von 100

4.1.3 Methoden zur Stichprobengewinnung

Grundprinzip: ZUFALL!

Einfache Zufallsstichprobe	Auswahlwahrscheinlichkeit für alle Elemente ist gleich groß, unabhängig welche Elemente bereits gezogen wurden, z.B. Ziehung der Lottozahlen
Klumpenstichprobe	Grundgesamtheit in Teilmengen (Klumpen), Klumpen = Abbild der Grundgesamtheit, z.B. Konsumverhalten der Studenten

Systematische Stichprobenverfahren

Auswahl nach ganz bestimmten Regeln, z.B. Geburtstag-, Buchstaben-, Schlussziffernverfahren, z.B. Untersuchung 1/30 der Grundgesamt, also nimmt man alle Personen die am 10. eines Monats Geburtstag haben

4.2 Stichprobenverteilung

In welcher Weise hängt die Abweichung des Erwartungswertes vom arithmetischen Mittelwert der Stichprobe, vom Umfang der Stichprobe ab?

- Beispiel:** Schraube soll 100 mm lang sein (Erwartungswert $\mu = 100$)
geg: normalverteilte Grundgesamtheit mit $\mu = 100$ und $s = 10$ (s : Standardabweichung)
Stichprobe: Folgende Werte der Standardabweichung ergeben sich in Abhängigkeit vom Stichprobenumfang n :

n	4	10	25	50	100	400	1000
$s(\bar{x})$	5	3,16	2	1,41	1,0	0,5	0,316

\bar{x} : Stichprobe

$$s(\bar{x}) = \frac{s(x)}{\sqrt{n}}$$

$s(\bar{x})$: Standardabweichung des arithmetischen Mittelwertes der Stichprobe zum Erwartungswert.

→ Je größer der Stichprobenumfang, desto größer ist die Wahrscheinlichkeit, dass der Stichprobenmittelpunkt nahe beim Erwartungswert μ ist

Ist x normalverteilt, so ist auch die Stichprobe \bar{x} normalverteilt:

$$N\left(\mu; \frac{s}{\sqrt{n}}\right)$$

Beispiel:

- mittleres Monatseinkommen 2000 €
- Standardabweichung $s = 750$ €

Gesucht:

- Wahrscheinlichkeit p , dass bei einer zufällig ausgewählten Gruppe von 100 Arbeitern, dass Einkommen höher als 2150 € ist (Normalverteilung)

Lösung:

- $\mu = 2000$; $s(\bar{x}) = \frac{s}{\sqrt{n}} = \frac{750}{\sqrt{100}} = 75$
- $P(\bar{x} > 2150) = 1 - P(\bar{x} \leq 2150) = 1 - F_0\left(\frac{2150 - 2000}{750}\right) = 1 - F_0(2) = \underline{\underline{0,0228}}$

4.3 Intervallschätzung

<p>Gegeben:</p> <ul style="list-style-type: none"> normalverteilte Grundgesamtheit mit Mittelwert μ und der Standardabweichung <p>Gesucht:</p> <ul style="list-style-type: none"> Stichprobenverfahren, dass mit hoher Wahrscheinlichkeit (z.B. $P_k=95\%$) ermöglicht, ein Intervall anzugeben, innerhalb dessen der wahre Mittelwert μ der Grundgesamtheit liegt. <p>Lösung:</p> <ul style="list-style-type: none"> man erhält dies mit P_k (z.B. 95%) indem man <ol style="list-style-type: none"> eine Zufallsstichprobe aus der Grundgesamtheit zieht daraus Stichprobenmittelwert \bar{x} ermittelt für dieses Schätzintervall die Grenzen festlegt <div style="text-align: center; border: 1px solid black; padding: 2px; width: fit-content; margin: 5px auto;"> $\bar{x} \pm \frac{k \cdot s}{\sqrt{n}}$ </div> <p>hier bei 95% $\rightarrow k$ bzw. $z = 1,96$ (siehe S. 20)</p> ein so erhaltenes Schätzintervall wird auch als Konfidenzintervall bezeichnet 	<p>Beispiel: Widerstände</p> <p>$\bar{x} = 203 \Omega$; $n = 25$</p> <p>$P_k = 95 \rightarrow \bar{x} \pm \frac{k \cdot s}{\sqrt{n}}$</p> <p>Bestimmung Konfidenzintervall:</p> <p>Untergrenze: $\bar{x} - \frac{1,96 \cdot 10}{\sqrt{25}} \approx 194 \Omega$</p> <p>Obergrenze: $\bar{x} + \frac{1,96 \cdot 10}{\sqrt{25}} \approx 207 \Omega$</p> <p>$\rightarrow 95\%$ Wahrscheinlichkeit, dass Widerstandswerte zwischen 199Ω und 207Ω liegen.</p>
--	---

4.4 Testverfahren

4.4.1 Grundlagen

Beispiel:

- von einer Eierlieferung behauptet Zulieferer, dass das mittlere Gewicht bei 60g und die Streuung (normalverteilt) bei $s = 10g$ liegt
- Abnehmer möchte dies durch Stichprobe prüfen
- Hypothese:** zu überprüfende Hypothese: **Nullhypothese** ; $m = 60g$ und $s = 10g$
 Negation der Nullhypothese: **Gegenhypothese** ; $s \gg 10g$?
 - falls Nullhypothese richtig ist \rightarrow Annahme Lieferung
 - falls Gegenhypothese richtig ist \rightarrow Verweigerung der Lieferung
- mögliche Fehler:**
 - Fehler 1 Art: Stichprobenergebnis \rightarrow Hypothese wird zurückgewiesen, obwohl richtig, z.B. man die bei der Stichprobe entnommenen Eier sind zufällig alle zu klein
 - Fehler 2 Art: Stichprobenergebnis \rightarrow Hypothese wird akzeptiert, obwohl sie falsch ist

4.4.2 Parametertest

<ol style="list-style-type: none"> Schätzintervall bestimmen, z.B. 99 % Unter- und Obergrenze bestimmen (Annahmegerenzen) Mittelwert \bar{x} muss berechnet werden 	<p>falls \bar{x} in Annahmebereich, d.h. $C_u \leq \bar{x} \leq C_o$ \rightarrow Annahme der Hypothese</p> <p>falls \bar{x} nicht in Annahmebereich \rightarrow Ablehnung der Hypothese</p>
--	--

Beispiel:

- Qualitätskontrolle, stündlich eine Stichprobe; Sollwert Werkstück $\mu_0 = 10\text{ cm}$; $s = 0,1$
- Schätzintervall bei 99% (**Signifikanzzahl** $\alpha = 0,01$)

Stichproben:

1. Stichprobe	9,91	10,00	10,05	10,10	10,01
2. Stichprobe	10,18	9,97	10,00	10,05	10,06
3. Stichprobe	9,80	9,80	9,42	10,03	9,8

- Schritt: $\alpha = 0,01$; $1 - \alpha = 0,99 \rightarrow k = 2,58$
- Schritt: $C_u = \mu_0 - \frac{k \cdot s}{\sqrt{n}} = 10 - \frac{2,58 \cdot 0,1}{\sqrt{5}} = \underline{9,885}$
 $C_o = \mu_0 + \frac{k \cdot s}{\sqrt{n}} = 10 + \frac{2,58 \cdot 0,1}{\sqrt{5}} = \underline{10,115}$
- Schritt: Mittelwert \bar{x}
 für 1. Stichprobe $\bar{x} = 10,014 \rightarrow$ **Annahme**
 für 2. Stichprobe $\bar{x} = 10,052 \rightarrow$ **Annahme**
 für 3. Stichprobe $\bar{x} = 9,870 \rightarrow$ **Verweigerung, da nicht in Grenzen**

4.4.3 Verteilungstest χ^2 – chi-Quadrat-Test

<p>Beispiel: Würfel wird 60 x geworfen</p> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tr> <td style="width: 10%;">Zahl</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td>Anzahl</td> <td>7</td> <td>8</td> <td>13</td> <td>8</td> <td>9</td> <td>15</td> </tr> </table> <p>Ziel ist es zu prüfen, ob der Würfel in Ordnung ist, d.h. nicht gezinkt ist.</p>	Zahl	1	2	3	4	5	6	Anzahl	7	8	13	8	9	15	$\chi^2 = \sum_{i=1}^m \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$																										
Zahl	1	2	3	4	5	6																																			
Anzahl	7	8	13	8	9	15																																			
<p>Handelt es sich um einen fairen Würfel?</p> <p><u>Nullhypothese:</u> $H_0 = p_i = \frac{1}{6}$ für $(1 \dots 6)$ $m = 6$</p> <p>1. Schritt: $\alpha = 0,05$ (\rightarrow 95 %ige Sicherheit, dass der Würfel nicht getürkt ist) $\rightarrow 1 - \alpha = 0,95$ <u>Freiheitsgrade:</u> $m - 1 = 6 - 1 = 5 \rightarrow$ aus Tabelle folgt $C_0 = 11,0705$</p> <p>2. Schritt: Berechnung von χ^2</p> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr style="background-color: #e0e0e0;"> <th>Augenzahl</th> <th>beobachtete Häufigkeit h_i</th> <th>erwartete Häufigkeit $(n \cdot p_i)$</th> <th>$n_i - p_i \cdot n$</th> <th>$\frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$</th> </tr> </thead> <tbody> <tr><td>1</td><td>7</td><td>10</td><td>9</td><td>0,9</td></tr> <tr><td>2</td><td>8</td><td>10</td><td>4</td><td>0,4</td></tr> <tr><td>3</td><td>13</td><td>10</td><td>9</td><td>0,9</td></tr> <tr><td>4</td><td>8</td><td>10</td><td>4</td><td>0,4</td></tr> <tr><td>5</td><td>9</td><td>10</td><td>1</td><td>0,1</td></tr> <tr><td>6</td><td>15</td><td>10</td><td>25</td><td>2,5</td></tr> <tr style="border-top: 1px solid black;"> <td colspan="4"></td> <td><u>$\chi^2 = 5,20$</u></td> </tr> </tbody> </table>		Augenzahl	beobachtete Häufigkeit h_i	erwartete Häufigkeit $(n \cdot p_i)$	$n_i - p_i \cdot n$	$\frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$	1	7	10	9	0,9	2	8	10	4	0,4	3	13	10	9	0,9	4	8	10	4	0,4	5	9	10	1	0,1	6	15	10	25	2,5					<u>$\chi^2 = 5,20$</u>
Augenzahl	beobachtete Häufigkeit h_i	erwartete Häufigkeit $(n \cdot p_i)$	$n_i - p_i \cdot n$	$\frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$																																					
1	7	10	9	0,9																																					
2	8	10	4	0,4																																					
3	13	10	9	0,9																																					
4	8	10	4	0,4																																					
5	9	10	1	0,1																																					
6	15	10	25	2,5																																					
				<u>$\chi^2 = 5,20$</u>																																					

3. Schritt:

$\chi^2 < C_0 \rightarrow$ Hypothese wird angenommen \rightarrow Würfel ist fair!

bei $\chi^2 > C_0$ muss Hypothese abgelehnt werden!

4.4.4 Berechnung des notwendigen Stichprobenumfangs

Beispiel:

Umfrage unter 200 Wählern, 98 wählen bei der nächsten Wahl CDU, der Rest wählt andere Parteien.

Gesucht:

95 % Vertrauensintervall für den Anteil (P) der CDU-Wähler

1. Schritt:

Vertrauensniveau: $1 - \alpha = 0,95 \rightarrow z = 1,96$

2. Schritt:

$$\bar{p} = \frac{98}{200} = 0,49 \quad (\text{Wahrscheinlichkeit der Stichprobe})$$

3. Schritt:

Vertrauensintervall, gilt $n \cdot \bar{p} \cdot (1 - \bar{p}) > 9$ so lautet Vertrauensintervall:

$$\bar{p} - z \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}} \leq p \leq \bar{p} + z \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}}$$

$$0,421 \leq p \leq 0,559$$

Wahrscheinlichkeit von 95%, dass CDU bei der Wahl zwischen diesen Werten ist

4.4.4.1 Berechnung des Vertrauensintervalls

	Untergrenze	Obergrenze	Länge des Intervalls
Wenn Standardabweichung bekannt	$\bar{x} - z \cdot \frac{s}{\sqrt{n}}$	$\bar{x} + z \cdot \frac{s}{\sqrt{n}}$	$2 \cdot z \cdot \frac{s}{\sqrt{n}}$
Wenn Wahrscheinlichkeit bekannt	$\bar{p} - z \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}}$	$\bar{p} + z \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}}$	$2 \cdot z \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}}$

Halbe Intervalllänge bezeichnet man auch als **Schätzfehler** $\Delta \mu$ oder Δp bei dem Intervall $1 - \alpha$

Gibt man den maximalen Schätzfehler vor:

$z \cdot \frac{s}{\sqrt{n}} \leq \Delta \mu$	$z \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}} \leq \Delta p$
$n \geq z^2 \cdot \frac{s^2}{\Delta \mu^2}$	$n \geq z^2 \cdot \frac{\bar{p} \cdot (1 - \bar{p})}{\Delta p^2}$

Wenn vorgegeben Genauigkeit eingehalten werden soll → **Mindestumfang der Stichprobe kann bestimmt werden!**

Problem:

- \bar{p} nicht bekannt (\bar{p} abhängig vom Stichprobenumfang)
- oft auch s (Standardabweichung) nicht bekannt

Will man ganz sicher bei Intervallschätzung sein, so wird man $\bar{p} = 1/2$ setzen.

Weiterführung Beispiel Wähler:

- weitere Umfrage soll gestartet werden mit größeren Stichprobenumfang
- **Ziel:** CDU-Anteil soll mit max. Abweichung von 1% genau geschätzt werden.

$$\Delta p = 0,01 \quad \text{man setzt} \quad \bar{p} = \frac{1}{2}$$

$$n \geq z^2 \cdot \frac{\bar{p} \cdot (1 - \bar{p})}{\Delta p^2}$$

$$n \geq 1,96^2 \cdot \frac{0,5 \cdot (1 - 0,5)}{0,01^2} \rightarrow \text{größerer Stichprobenumfang für genauere Schätzung}$$
$$\underline{\underline{n \geq 9604}}$$

5 Anlagen

Für dieses Dokument werden folgende Tabellen benötigt:

- Binomialverteilung
- Poissonverteilung
- Standardnormalverteilung
- χ^2 -Test