Manuskript zur Vorlesung

Stochastik

Teil 1

Prof. F. Stockmayer

Inhalt

Das Manuskript stützt sich hauptsächlich auf folgende Literatur:
Lothar Papula, Mathematik für Ingenieure und Naturwissenschaftler Band 3, Vieweg und Teubner

1 Geburtstagsparadoxon

Frage: Wie hoch ist die Wahrscheinlichkeit, dass bei 23 Personen mindestens zwei von ihnen am selben Tag Geburtstag haben?
Antwort: $p > 50\%$

1.1 Überlegung

In einer Gruppe mit n-Personen können n(n-1)/2 Paare gebildet werden.

Übertragung der Fragestellung in das sog. Urnenmodell:
Eine Urne enthält 365 durchnummerierte Kugeln. Daraus werden n-Kugeln mit Zurücklegen gezogen.

Anzahl aller möglichen Geburtstagskombinationen für n-Personen: $\mathbf{m = 365^n}$
Z.B. für 2 Personen $m = 365^2 = 133225$
Darin enthalten sind auch identische Geburtstage
Somit gibt es in der Gruppe
u = 365*364*363* ... *(365-(n-1)) unterschiedliche Geburtstage

Für die erste Person gibt es noch freie Auswahl, für die zweite Person bleiben dann noch 364 Tage übrig etc.

Die W., dass alle n-Personen an unterschiedlichen Tagen Geburtstag haben, ergibt sich aus der sog. Laplace-Formel:

1-1-1

$$\frac{u}{m} = \frac{365 \cdot 364 \cdot \ldots \cdot (365-(n-1))}{365^n}$$

Für die Umkehrung, also die W. mindestens einen doppelten Geburtstag:

1-1-2

$$p = 1-\frac{u}{m} = 1-\frac{365 \cdot 364 \cdot \ldots \cdot (365-(n-1))}{365^n}$$

Für n = 23:

1-1-3

$$p = 1-\frac{u}{m} = 1-\frac{365 \cdot 364 \cdot \ldots \cdot (365-(23-1))}{365^{23}} = 0,5073 = 50,73\%$$

Es müssen dabei n(n-1)/2 = 23*22/2 = 253 Paare geprüft werden.

1-1-4

$$p = 1 - \frac{u}{m} = 1 - \frac{365!}{(365-n)! \cdot 365^n} = 1 - \frac{n! \cdot \binom{365}{n}}{365^n}$$

1.2 Erweiterung: Wahrscheinlichkeit für einen bestimmten Tag

Die W. für einen Geburtstag an einem bestimmten Datum:

1-2-1 $p = 1/365 = 0{,}0027$ Aussage: A

Das Gegenteil davon, also nicht am ausgewählten Tag Geburtstag zu haben:

1-2-2 $q = 1 - p = 1 - 1/365 = 0{,}997$ Aussage: $\overline{A}$

Dass zwei Personen beide nicht am ausgewählten Tag Geburtstag haben:

1-2-3 $p = q_1 * q_2 = q * q = q^2$ Aussage: $\overline{A} \wedge \overline{B}$

Mindestens eine der beiden Personen hat an einem bestimmten Tag Geburtstag:

1-2-4 $p = 1 - q^2$ Aussage inv.: $\overline{\overline{A} \wedge \overline{B}} = A \vee B$

Übertragung auf n - Personen:

1-2-5 $p = 1 - q^n$

Wieviel Personen n benötigt man, um eine bestimmte W. p zu erreichen, dass mindestens eine Person an einem bestimmten Tag Geburtstag hat?

1-2-6

$$1-p = q^n = \left(1 - \frac{1}{365}\right)^n$$

1-2-7

$$\log(1-p) = n \cdot \log q = n \cdot \log\left(1 - \frac{1}{365}\right)$$

1-2-8

$$n = \frac{\log(1-p)}{\log\left(1 - \frac{1}{365}\right)}$$

z.B. für p = 50%: n = 253

Und für den Fall, dass eine der n anwesenden Personen Geburtstag hat (z.B. „Heute“) und von den übrigen (n-1) Personen mindestens eine am gleichen Tag Geburtstag hat:

1-2-9

$$p = 1 - q^{n-1}$$

2 Hilfsmittel aus der Kombinatorik

2.1 Urnenmodell

Es befinden sich n verschiedene Kugeln (Farbe/Größe/Beschriftung/...) in einer Urne.
a) Wieviele Möglichkeiten, die Kugeln unterschiedlich anzuordnen? -> Permutation

Bei einer zufälligen Entnahme (Stichprobe vom Umfang k) kann unterschieden werden, mit/ohne Berücksichtigung der Reihenfolge (geordnet/ungeordnet), sowie

b) Ziehen ohne zurücklegen
c) Ziehen mit zurücklegen

2.2 Permutation

Urne mit n verschiedenfarbigen Kugeln. Wieviele unterschiedliche Anordnungen?

Beispiel: 3 Kugeln

2-2-1
Permutationen: Perm(n) = n(n-1)(n-2) ... 1 = n!
Befinden sich darunter n_1 identische Kugeln: Perm(n, n_1) = n!/n_1!

Überlegung:

Zuerst 4 unterschiedliche Objekte $\Omega = \{1,2,3,4\}$
Hierfür gibt es n! = 4! = 25 Permutationen

Anschließend wird z.B. Objekt 3 durch ein identisches Objekt 2 ersetzt: $\Omega = \{1,2,2,4\}$
Jetzt gibt es nur noch n!/n1! = 4!/2! = 12 Permutationen:

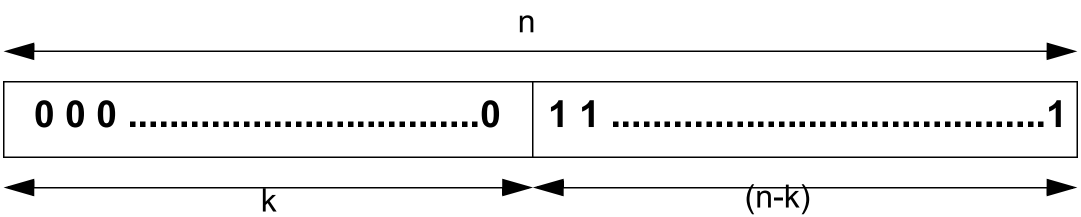
1 2 2 4	4 1 2 2	2 1 2 4	2 4 2 1
1 2 4 2	4 2 1 2	2 1 4 2	2 2 4 1
1 4 2 2	4 2 2 1	2 4 1 2	2 2 1 4

Oder allgemein: Gibt es von jeder Kugelart mehrere Kugeln:
2-2-2 Permutation mit Wiederholung:

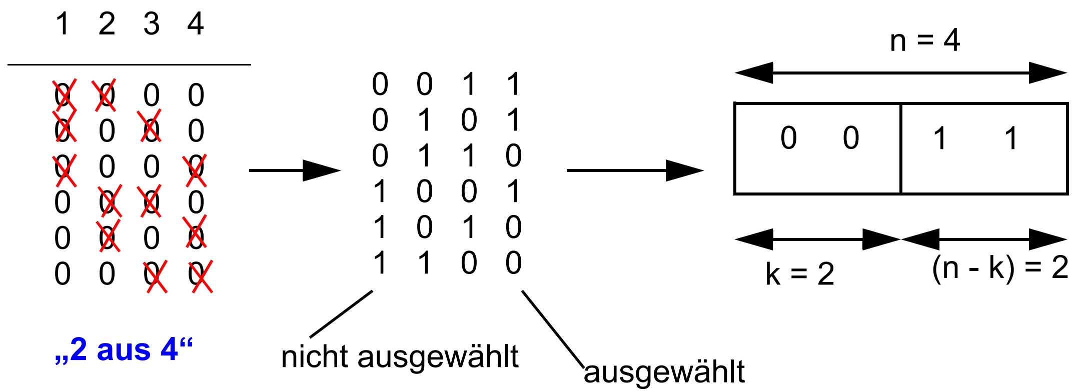
2.3 Kombination ohne Wiederholung

Auf wieviel verschiedene Arten lassen sich k Kugeln aus einer Urne mit n verschiedenen Kugeln ohne zurücklegen und **ohne Berücksichtigung der Reihenfolge** ziehen?

Lösung: Alle n Kugeln aus der Urne in einer Reihe anordnen und dabei die k gezogenen Kugeln mit ‚0', die anderen mit ‚1' kennzeichnen. Anschließend die entsprechend markierten Kugeln gruppieren:



Beispiel:
Gegeben seien 4 Objekte nummeriert mit 1 ... 4
Es sollen 2 Objekte ausgewählt werden. Wieviel unterschiedliche Wahlmöglichkeiten gibt es?



Durch die Markierung werden die ursprünglich unterschiedlichen Objekte in 2 Klassen gruppiert: Ausgewählt/Nicht ausgewählt. Beide Klassen zeichnen sich dadurch aus, dass Elemente darin mehrfach vorkommen (k- bzw. (n-k)-mal)
Die Frage lautet jetzt: Wieviel unterschiedliche Permutationen sind möglich?
Für die Beantwortung kann nun Gl. 2-2-2 angewendet werden.

Mit Gl. 2-2-2 berechnet sich die Anzahl der Möglichkeiten, aus einer Urne mit n verschiedene Kugeln, k Kugeln ohne zurücklegen und ohne Berücksichtigung der Reihenfolge herauszugreifen:

2-3-1

Beispiel: Lotto 6 aus 49

2.4 Kombination mit Wiederholung

Ziehung mit Zurücklegen der Kugeln: n Kugeln, k Ziehungen

2-4-1

$$\frac{(n+k-1)!}{(n-1)!\cdot k!} = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

2.5 Wie kommt Gl. 2-4-1 zustande?

Sollen aus einer Menge von n Elementen k Elemente ausgewählt werden, mit Wiederholung und ohne Berücksichtigung der Reihenfolge (Ziehung mit Zurücklegen), so kann man sich die Formel folgendermaßen verdeutlichen:

Jedes Ergebnis von k ausgewählten Elementen aus n möglichen Elementen wird durch eine Folge von n+k Symbole dargestellt.
Dabei stellen n Symbole („N“) die Elemente der Auswahlmenge und k Symbole („K“) die ausgewählten Elemente dar. Die Folge beginnt immer mit einem N-Symbol. Die Anzahl der K-Symbole vor dem zweiten N-Symbol entspricht der Häufigkeit, mit der das erste der n Elemente gezogen wurde, die Anzahl der K-Symbole zwischen dem zweiten und dritten N-Symbol entsprechen der Häufigkeit, mit der das zweite der n Elemente gezogen wurde etc. Da bis auf das erste „N“ alle Symbole frei kombiniert werden können, entspricht die Anzahl der Kombinationen und damit die Anzahl der Zugmöglichkeiten der angegebenen Formel. Jetzt kann entsprechend dem Fall „ohne Zurücklegen“ argumentiert werden.

Beispiel:
Auswahl 3 aus 5 Elementen mit Zurücklegen, Ω = {1,2,3,4,5}
Das Ergebnis A={1,3,3} führt zur Symbolfolge „NKNNKKNN“
Das Ergebnis B = {5,5,5} führt zur Symbolfolge „NNNNNKKK“

$$\binom{n+k-1}{k} = \binom{7}{3} = 35$$

2.6 Variationen ohne Wiederholung

Ziehung von k Kugeln ohne Zurücklegen, aber mit Berücksichtigung der Reihenfolge
Wieviele Variationen sind möglich?

Lösung:
Für jede Ziehung von k-Kugeln sind k! Permutationen (Reihenfolgen) möglich, wobei es mit Gl. 2-3-1 (Anzahl Kombinationen) dann insgesamt folgende Anzahl Variationen gibt:

2-6-1

2.7 Variation mit Wiederholung

Ziehung von k Kugeln mit Zurücklegen und Berücksichtigung der Reihenfolge:

2-7-1

Beispiel: Zahlensystem, 3 Ziffern, je Ziffer 0 ... 9
Anzahl Variationen: $V = 10^3 = 1000$ (000, 001, 002, ... 999)

Begriffsabgrenzung Kombinatorik

Permutation	Variation	Kombination
Es werden alle verfügbaren Objekte ausgewählt (n=k)	Auswahl von k-Objekten aus einer Menge von n-Objekten. Reihenfolge spielt eine Rolle mit/ohne Zurücklegen	Bei der Auswahl der Objekte spielt die Reihenfolge keine Rolle mit/ohne Zurücklegen

Variation ohne Wiederholung: „k aus n mit Reihenfolge“

Ausgehend von der maximal möglichen Anzahl der Permutationen muss aufgrund der partiellen Auswahl „k aus n“ durch die Anzahl nicht vorhandener Anordnungen dividiert werden.

$$\frac{n!}{(n-k)!} = \frac{n(n-1)(n-2)\dots(n-k+1)(n-k)(n-k-1)\dots(1)}{(n-k)!}$$

$$\frac{n!}{(n-k)!} = n(n-1)(n-2)\dots(n-k+1)$$

Beispiel:
6 Wettkämpfer kommen in die Endrunde. Die ersten 3 Plätze werden prämiert. Wieviele Möglichkeiten gibt es, die „TOP3“ uu besetzen?

$$\frac{n!}{(n-k)!} = \frac{6!}{(6-3)!} = \frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1} = 120$$

Variation mit Wiederholung
Jedes der n-Objekte kann auf jedem der k-Plätze erscheinen: n*n*n* ....*n = n^k

Beispiele:
4stelliger PIN, oder zahlenschloss 0000 ... 9999 => 10^4 = 10000 Variationen
8stellige Binärzahl: 00000000 ... 11111111 => 2^8 = 256 Variationen

Kombination ohne Wiederholung: „k aus n ohne Reihenfolge“

Ausgehend von Variationen ohne Wdh. (mit Reihenfolge) muss das Ergebnis durch die Anzahl der verschiedenen Anordnungen aufgrund der Reihenfolge geteilt werden.
Die k ausgewählten Elemente können auf k! verschiedene Weisen angeordnet werden, die aber allesamt nicht unterschieden werden sollen und als gleiche Auswahl der Elemente gelten. Daraus entsteht der sog. Binomialkoeffizient:

$$\frac{n!}{(n-k)!\cdot k!} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} = \binom{n}{n-k} = \binom{n}{k}$$

Beispiel: Lotto „6 aus 49“:

$$\frac{49!}{43!\cdot 6!} = \binom{49}{43} = \binom{49}{6} = 13983816$$

Kombination mit Wiederholung (s. Abschn. 2.5)

Beispiel:
Gummibärchen Orakel. Aus einer Tüte Gummibärchen mit n=5 Farben werden k=5 Gummibärchen entnommen. Wieviele Farbkombinationen gibt es?

$$\frac{(n+k-1)!}{(n-1)!k!} = \binom{n+k-1}{k} = \binom{n+k-1}{n-1} = \binom{9}{5} = \frac{9!}{5!\cdot 4!} = 126$$

2.8 Zusammenfassung: Formeln zur abzählenden Kombinatorik

		Menge	Reihenfolge
Permutation ohne Wiederholung	$n!$	n aus n	wird beachtet
Permutation mit Wiederholung	$\frac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_s!}$	n aus n	wird beachtet
Variation ohne Wiederholung	$\frac{n!}{(n-k)!}$	k aus n	wird beachtet
Variation mit Wiederholung	n^k	k aus n	wird beachtet
Kombination ohne Wiederholung	$\binom{n}{k}$	k aus n	wird nicht beachtet
Kombination mit Wiederholung	$\binom{n+k-1}{k}$	k aus n	wird nicht beachtet

3 Grundbegriffe

3.1 Beispiel 1

Augenzahlen beim Wurf eines homogenen Würfels

Ergebnismenge $\Omega = \{1,2,3,4,5,6\}$ mit den Elementarereignissen ω_i
1 Wurf (Experiment) entspricht einem Zufallsexperiment
A = {5} beschreibt z.B. eine Teilmenge von Ω, bestehend aus einem Elementarereignis
C = {1,3,5} ist ebenfalls Teilmenge von Ω und beschreibt die Teilmenge bzw. Ereignis „Würfeln einer ungeraden Zahl“

3.2 Beispiel 2

Wurf mit zwei unterscheidbaren homogenen Würfeln

Ergebnismenge Ω = {(1;1), (1;2), ... (6;6)}
Teilmenge für „Summe der Augenzahl ist 4“: B = {(1;3), (2;2), (3;1)}

3.3 Beispiel 3

Ziehung von drei Kugeln aus einer Urne (3 Weiße, 3 Schwarze) mit Zurücklegen
Ergebnismenge Ω = {(WWW), (WWS), (WSW), (SWW), (WSS), (SWS), (SSW), (SSS)}
Mögliche Teilmengen: „Zwei der Kugeln sind Weiß“, „Alle Kugeln haben gleiche Farbe“

3.4 Zufallsexperiment

- unter gleichen Bedingungen beliebig wiederholbar
- Ergebnisse, die sich gegenseitig ausschließen
- Ergebnis eines konkreten Experiments ist zufallsbedingt

Die sich gegenseitig ausschließenden Ereignisse heißen Elementarereignisse ω_i
Die Menge aller Elementarereignisse heißt Ergebnismenge Ω

Unterscheidung:

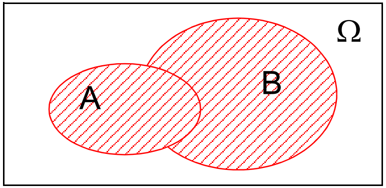
„endliche Ergebnismenge“
$\Omega = \{\omega_1, \omega_2, ... \omega_n\}$

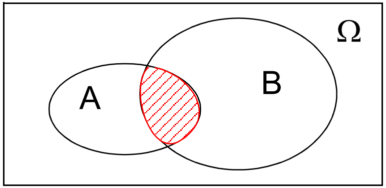
„abzählbar unendl. Ergebnismenge“
$\Omega = \{\omega_1, \omega_2, ... \omega_n, ...\}$

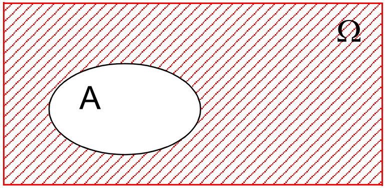
Beschreibung der Ergebnisse eines Zufallsexp. mit Teilmengen bzw. Ereignissen:

$A = \{\omega_1\}$
$B = \{\omega_1, \omega_3, \omega_5\}$
...

3.5 Verknüpfung von Ereignissen







4 Wahrscheinlichkeit

4.1 Laplace Experiment

Zufallsexperiment mit endlicher Ergebnismenge $\Omega = \{\omega_1,\omega_2,...., \omega_m\}$
Alle Elementarereignisse ω_i gleichmöglich, d.h. sie treten mit gleicher Häufigkeit auf
absolute Häufigkeit: $n(A)$
relative Häufigkeit: $h(A) = n(A) / n$

Laplace-Raum: Ergebnismenge Ω besitzt m gleichmögliche Elementarereignisse ω_i

Wahrscheinlichkeit:
4-1-1

Wahrscheinlichkeit für ein Ereignis A:
4-1-2

Beispiel: In einer Warenlieferung sind von insgesamt 100 Artikel 20 defekt. Zur Kontrolle wird der Lieferung wahllos ein Artikel entnommen. Zufallsereignis A: „Ziehung eines defekten Artikels“. $P(A) = g(A)/m = 20/100 = 0{,}2$

4.2 Eigenschaften der rel. Häufigkeiten

-
-
-
-

Wahrscheinlichkeitsaxiome:

A1:

A2:

A3:

Folgerungen:

1.

2.

3.

4.3 Statistische Definition der Wahrscheinlichkeit

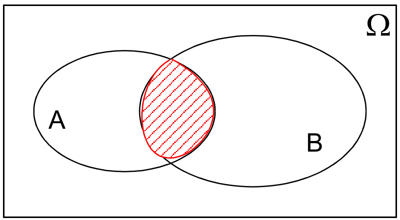
oder: empirisch bestimmte Wahrscheinlichkeit

4-3-1

4.4 Additionssatz für beliebige Ereignisse

4-4-1

Venn-Diagramm:



Die Ereignisse A, B schließen sich nicht gegenseitig aus. Es gibt eine Schnittmenge

Beispiel: Ein homogener Würfel wird 2x geworfen.
Wie groß ist die Wahrscheinlichkeit für mindestens einmal eine 6?

Lösung: A: 6 beim ersten Wurf, B: 6 beim zweiten Wurf

P(A) = P(B) = 1/6

Aber: Überschneidung bei Augenzahl 6 im ersten **und** zweiten Wurf

$$P(A \cap B) = \frac{1}{36}$$

Experiment A, B -> 1. Wurf, 2. Wurf

6,1	6,2	6,3	6,4	6,5	6,6
1,6	2,6	3,6	5,6	5,6	6,6

6,X -> P(A) = 1/6

X,6 -> P(B) = 1/6

identisches Ergebnis
in P(A) + P(B) doppelt berücksichtigt

4-4-2

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

Kontrolle:

Ergebnispaare

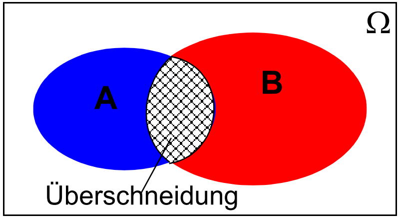
1,1	2,1	3,1	4,1	5,1	6,1
1,2	2,2	3,2	4,2	5,2	6,2
1,3	2,3	3,3	4,3	5,3	6,3
...	...	...	...	...	6,4
...	...	...	...	...	6,5
1,6	2,6	3,6	4,6	5,6	6,6

Insgesamt 36 Ergebnismöglichkeiten
davon 11 günstige Ergebnisse (mindestens eine 6)

P = 11/36 = 1/6 - 1/6 - 1/36

Beweis von Gl. 4-4-2 :

Zerlegen von (A v B) in zwei sich ausschließende Ereignisse



$$A \vee B = A \vee (\bar{A} \wedge B)$$

4.5 Bedingte Wahrscheinlichkeit

Wahrscheinlichkeit für das Eintreten eines best. Ereignisses B unter der Voraussetzung bzw. Bedingung, dass A bereits eingetreten ist.

1. Beispiel
Urne mit 3 weißen u. 3 schwarzen Kugeln.
Wie groß ist die W., bei der zweiten Ziehung eine weiße Kugel zu erhalten
(Ziehung nacheinander, ohne zurücklegen)
A: Weiß in der ersten Ziehung, B: Weiß in der zweiten Ziehung

Fall 1: Erste Ziehung Weiß: $P(A) = 3/6$, zweite Ziehung Weiß: $P(B|A) = 2/5$
Fall 2: Erste Ziehung Schwarz: $P(\bar{A}) = 3/6$, zweite Ziehung Weiß: $P(B|\bar{A}) = 3/5$

-> unterschiedliche Ergebnisse, je nach Vorbedingung

4-5-1

2. Beispiel
Gleichzeitiger Wurf zweier Würfel
Wahrscheinlichkeit für Augensumme = 8 und beide Augenzahlen gerade.

A: Augensumme = 8
B: Beide Augenzahlen sind gerade
Gesucht: $P(B|A) = ?$

1. Lösung
Ω enthält 36 Elementarereignisse, davon erfüllen 5 Ereignisse „Augensumme=8"
$A = \{(2;6),(3;5),(4;4),(5;3),(6;2)\}$
$B|A = \{(2;6),(4;4),(6;2)\}$
$P(B|A) = 3/5$

2. Lösung

4.6 Multiplikationssatz

Mit Gl. 4-5-1 :

4-6-1

Beispiel

Urne mit 4 W und 2 S, zwei Ziehungen nacheinander, ohne zurücklegen.
Gesucht: W. für zweimal Weiß

Lösung:
A: 1. Ziehung ergibt W: P(A) = 2/3
B: 2. Ziehung ergibt W: P(B|A) = 3/5

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{2}{3} \cdot \frac{3}{5} = \frac{2}{5} = 0,4 = 40\%$$

Sonderfall: Stochastisch unabhängige Ereignisse

4-6-2

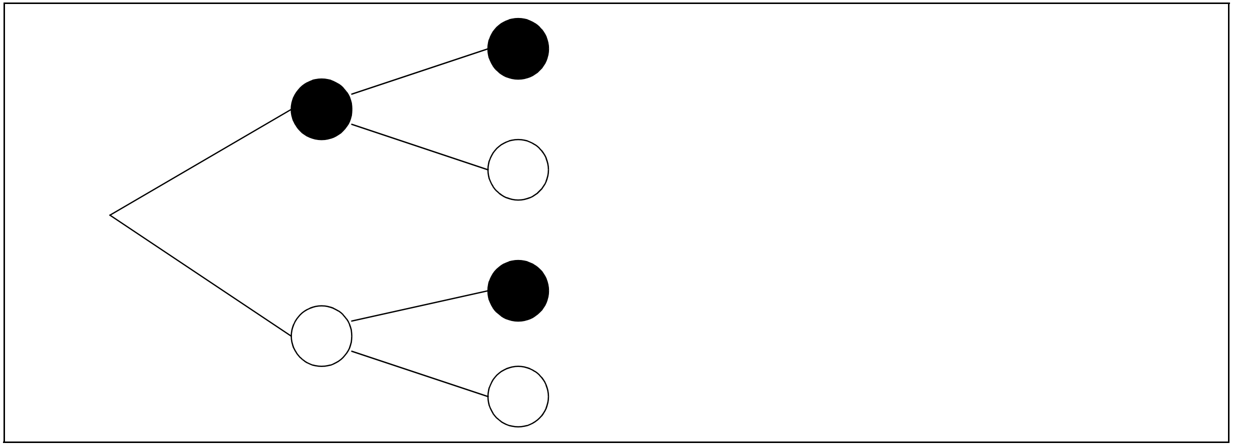
d.h., das Ereignis B ist völlig unabhängig von Ereignis A

Aufgabe:
Machen Sie sich den Unterschied zw. Gl. 4-6-2 und Gl. 4-6-1 an einem Urnenbeispiel mit und ohne Zurücklegen deutlich.

4.7 Ereignisbaum

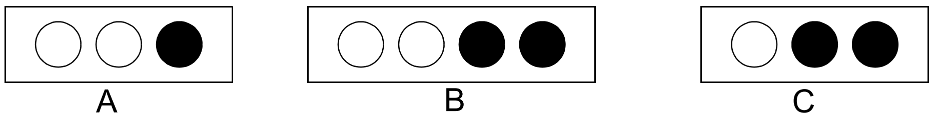
Ereignisbäume bilden ein anschauliches Hilfsmittel zur Berechnung von Wahrscheinlichkeiten bei mehrstufigen Zufallsexperimenten.

Beispiel: Urne mit 2W und 4S Kugeln. Ziehung von 2 Kugeln nacheinander ohne zurücklegen. Wie groß ist die W. für
a) zwei gleichfarbige
b) zwei verschiedenfarbige Kugeln?

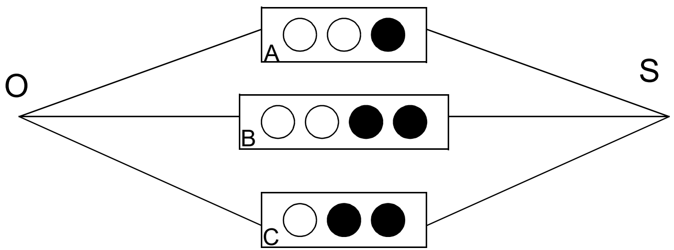


4.8 Totale Wahrscheinlichkeit, Bayes‘sche Formel

Beispiel: 3 Urnen mit weißen und schwarzen Kugeln. Entnahme einer Kugel aus einer zufällig ausgewählten Urne.
Frage:
a) W. für eine schwarze Kugel
b) W., dass gezogene schwarze Kugel aus C stammt



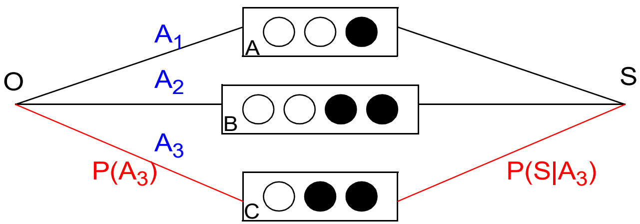
a) Ereignis S: Ziehung einer schwarzen Kugel P(S)



4-8-1: Totale Wahrscheinlichkeit

$$P(S) = \sum_i^n P(A_i) \cdot P(S|A_i)$$

b) Spezialisierung:
Wie groß ist die Wahrscheinlichkeit, dass das eingetretene Ereignis mit einem bestimmten Zwischenergebnis (Pfad) verbunden ist?



4-8-2: Bayessche Formel

$$P(A_j|S)) = \frac{P(A_j) \cdot P(S|A_j)}{\sum_{i=1}^{3} P(A_i) \cdot P(S|A_i)}$$

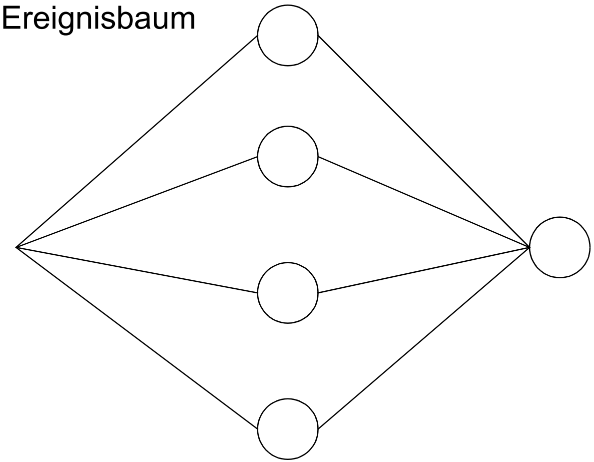
Beispiel:
Auf 4 Maschinen werden Bauteile hergestellt. Ihr Anteil an der Gesamtproduktion beträgt 10%, 20%, 30%, 40% und die Ausschussanteile betragen 2%, 1%, 4%, 2%
Aus der Gesamtproduktion wird zufällig ein Bauteil entnommen.
a) W. für ein defektes Bauteil
b) Vorausgesetzt das Bauteil ist defekt: W., dass das Bauteil von Maschine M_3 stammt

Ereignisse:

A_i: Das entnommene Bauteil stammt von $M_{i,}$

B: Das entnommene Bauteil ist defekt

Ereignisbaum



Lösung a)

Lösung b)

5 Wahrscheinlichkeitsverteilung einer Zufallsvariablen

5.1 Zufallsvariable

Unter einer Zufallsvariablen X versteht man eine Funktion, die jedem Elementarereignis ω aus der Ergebnismenge Ω eines Zufallsexperiments genau eine reelle Zahl $X(\omega)$ zuordnet $X = X(\omega)$

Beispiel: $\Omega = \{$„1“, „2“, „3“, „4“, „5“, „6“,$\}$

X = 1 für Augenzahl „6“, X = 0 sonst -> $\Omega_X = \{0,1\}$

Y = Anzahl gewürfelter 6er -> $\Omega_Y = \{0,1,2,3, ...\}$

Zufallsvariable werden i.a. groß geschrieben, ihre Werte mit kleinen Buchstaben.
Eigenschaften einer Zufallsvariable:

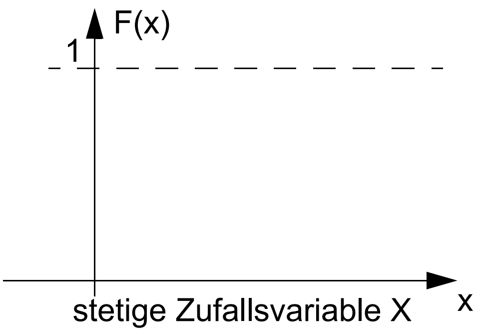
- Diskrete Zufallsvariable: endlich viele bzw. abzählbar unendlich viele Werte
- Stetige Zufallsvariable: Kann jeden bel. Wert aus einem reellen endlichen oder unendlichen Intervall annehmen
- Wahrscheinlichkeit P, dass die Zufallsvariable X einen bestimmten Wert annimmt oder in einem best. Intervall liegt

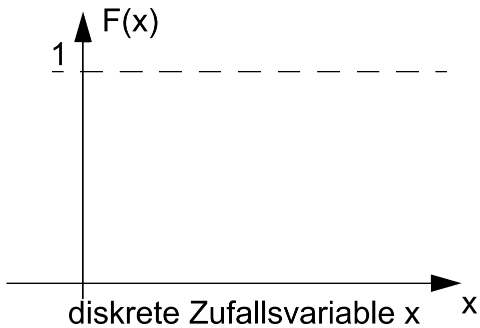
5.2 Verteilungsfunktion einer Zufallsvariablen

Die Verteilungsfunktion F(x) einer Zufallsvariablen X gibt die Wahrscheinlichkeit an für den Fall, dass die Zufallsvariable X einen Wert kleiner oder gleich der vorgegebenen Zahl x hat.

Definition Verteilung
Bezeichnung für eine empirische Häufigkeitsverteilung oder für die Wahrscheinlichkeits-verteilung einer Zufallsvariablen, die durch eine Verteilungsfunktion, einer Dichtefunktion oder einer Wahrscheinlichkeitsfunktion f(x) angegeben wird.
5-2-1

$$F(x) = P(X \leq x)$$



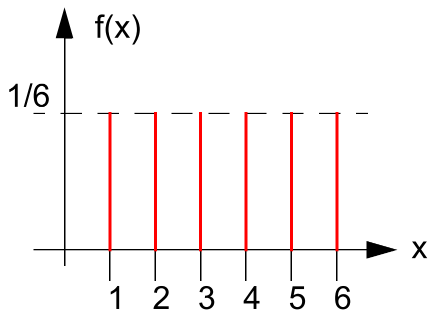


Beispiel: Zufallsexperiment „Wurf eines Würfels", X = Gewürfelte Augenzahl

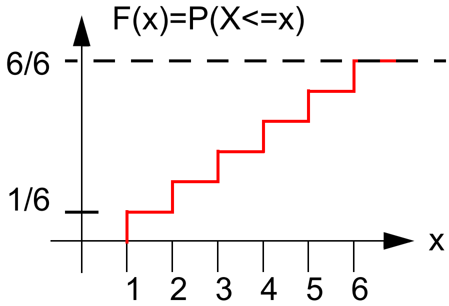
Diskrete Zufallsvariable mit der Dichte/Wahrscheinlichkeitsfunktion $f(x_i)$:

x_i	1	2	3	4	5	6
$f(x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

Stabdiagramm:



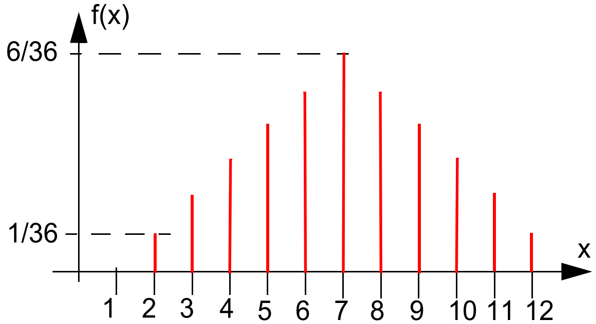
Verteilungsfunktion:

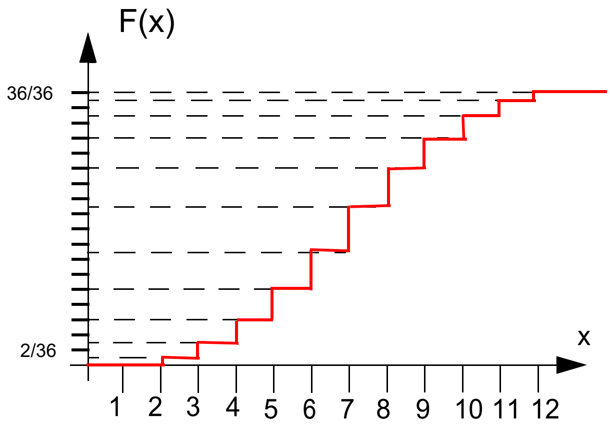


Beispiel: Zufallsexperiment „Gleichzeitiger Wurf zweier Würfel“,
X = Gewürfelte Augenzahl

Diskrete Zufallsvariable mit der Dichte/Wahrscheinlichkeitsfunktion $f(x_i)$:

x_i	1	2	3	4	5	6	7	8	9	10	11	12
$f(x_i)$	0	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36





Eigenschaften einer Verteilungsfunktion F(x):

-
-
-

5-2-2

$$P(a \leq x \leq b) = F(b) - F(a)$$

5.3 Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariablen

5-3-1 Verteilungsfunktion

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

wobei f(x) die Wahrscheinlichkeitsfunktion einer diskreten Verteilung beschreibt

$$f(x) = \begin{cases} p_i & \text{für} \quad x = x_i \\ 0 & \text{sonst} \end{cases}$$

Beispiele

5.4 Wahrscheinlichkeitsverteilung einer stetigen Zufallsvariablen

5-4-1 Verteilungsfunktion

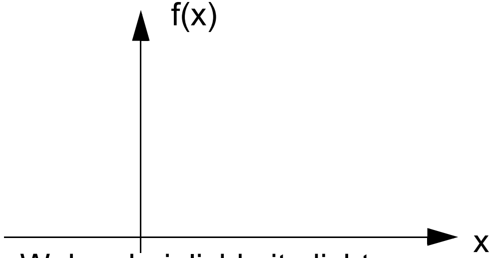
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

wobei f(x) die Wahrscheinlichkeitsdichtefunktion einer stetigen Verteilung beschreibt

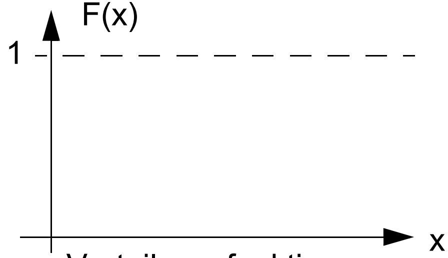
a)

b)

c)



Wahrscheinlichkeitsdichte



Verteilungsfunktion

Für b > a gilt:
5-4-2

$$P(a \le x \le b) = F(b) - F(a) = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx$$

Die Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert zwischen a und b annimmt, entspricht dem Flächeninhalt unter der Dichtefunktion f(x) zwischen den Grenzen x=a und x=b (Gleichheitszeichen, falls f(x) an a bzw. b stetig ist).

Beispiel: Die Lebensdauer eines el. Bauelements sei eine exponentielle Zufallsgröße mit der Dichtefunktion

$$f(t) = \begin{cases} 0 & \text{für} \quad (t < 0) \\ Ce^{-0,1t} & \quad (t \ge 0) \end{cases}$$

a) Wie lautet die zugeh. Verteilungsfunktion F(t)?
b) Wie groß ist der Anteil der Bauelemente, deren Lebensdauer den Wert t > 10 überschreitet?

Lösung: Zunächst Bestimmung der Konstanten C
aus Gl. 5-4-1 c) folgt:

$$\int_{-\infty}^{\infty} f(t) = \int_0^{\infty} Ce^{-0,1t}dt = C[-10 \cdot e^{-0,1t}]\Big|_0^{\infty} = 10C = 1$$

$$f(t) = 0,1 \cdot e^{-0,1t}$$

a) Verteilungsfunktion:

$$F(t) = \int_{-\infty}^{t} f(u)du = \int_0^t 0,1 \cdot e^{-0,1u}du = 0,1[-10e^{-0,1u}]\Big|_0^t$$

$$F(t) = -[e^{-0,1u}]\Big|_0^t = 1 - e^{-0,1t}$$

$$F(t) = \begin{cases} 0 & \text{für} \quad (t < 0) \\ 1 - e^{-0,1t} & \quad (t \geq 0) \end{cases}$$

b) Die gesuchte W.-Funktion P(T >= 10) entspricht der Fläche unter f(t) f. t >= 10

Noch ein **Beispiel**: Stetige Gleichverteilung

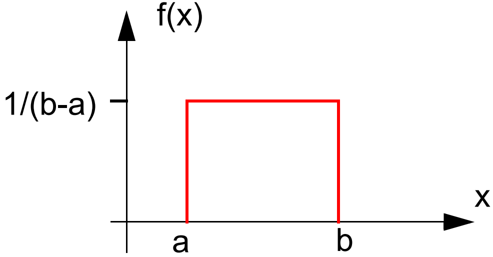
Dichtefunktion f(x) einer Gleichverteilung (z.B. für Zufallszahlen):

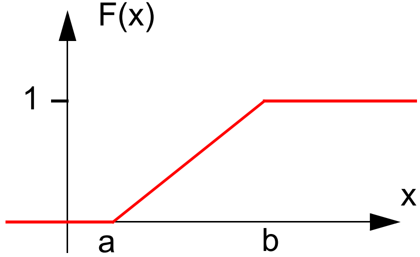
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für a} \le x \le b \\ 0 & \text{für x } < a \end{cases}$$

Verteilungsfunktion F(x):

$$F(x) = \int_{-\infty}^{\infty} f(u)du = \int_{a}^{x} \frac{1}{b-a}du = \frac{1}{b-a}[u]\Big|_a^x = \frac{x-a}{b-a}$$

$$F(x) = \begin{cases} 0 & \text{für x} < a \\ \frac{x-a}{b-a} & \text{für a} \le x \le b \\ 1 & \text{für x} > b \end{cases}$$

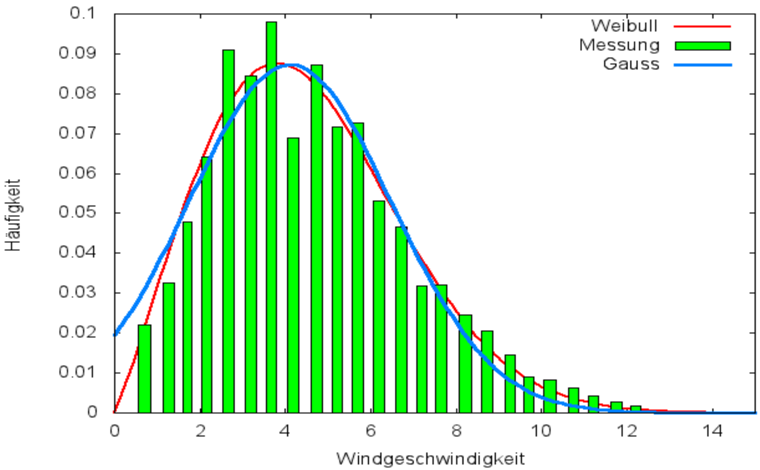




Beispiel: Weibull Verteilung
z.B. zur Auswertung von Windgeschwindigkeiten und Planung von Windkraftanlagen
Messung der mittleren Windgeschwindigkeit mit einem Anemometer alle 10 Minuten
Mit den so ermittelten Häufigkeiten der Windgeschwindigkeit lässt sich der Energiegehalt berechnen.

$$f(v) = \frac{k}{A}\left(\frac{v}{A}\right)^{k-1} e^{-\left(\frac{v}{A}\right)^k}$$

Mit den Parameter k, A lässt sich die Verteilung sehr vielseitig anpassen.



Bildquelle: https://de.wikipedia.org/wiki/Weibull-Verteilung

6 Kennwerte bzw. Maßzahlen einer W.-verteilung

Die Wahrscheinlichkeitsverteilung einer Zufallsvariablen X lässt sich eindeutig/vollständig anhand der Verteilungsfunktion F(X) oder der Dichtefunktion f(x) beschreiben. In der Praxis werden weitere Parameter zur Charakterisierung einer Verteilung herangezogen.

6.1 Erwartungswert einer diskreten Zufallsvariablen

Sei X die erzielte Augenzahl beim Wurf eines Würfels, dann treten die 6 möglichen Werte x_i mit gleicher Wahrscheinlichkeit auf.
Man kann hier „erwarten", dass die mittlere Augenzahl pro Wurf (bei einer großen Anzahl Versuche) dem arithmetischen Mittel entspricht.

$$\bar{x} = \frac{1+2+3+4+5+6}{6} = 3,5$$

Allgemein für eine diskrete Zufallsvariable:
6-1-1

$$E(X) = \sum_i x_i \cdot f(x_i)$$

Gewichtung

Hinweis: E(X) wird aus allen Werten der Grundgesamtheit gebildet. Es handelt sich hierbei nicht um den Schätzwert für μ einer Stichprobe (s. Abschn. 12)!

6.2 Erwartungswert einer stetigen Zufallsvariablen

6-2-1

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

Beispiel: Lebensdauer eines elektronischen Bauelements

$$f(t) = \begin{cases} 0 & \text{für} \quad (t < 0) \\ \lambda e^{-\lambda t} & \quad (t \geq 0) \end{cases}$$

Erwartungswert für mittlere Lebensdauer:

Ansatz zur Integration von E(X):

$$\int u \cdot dv = u \cdot v - \int v \cdot du$$

$$E(T) = \int_{-\infty}^{\infty} t \cdot f(t)dt = \int_{-\infty}^{\infty} t \cdot \lambda e^{-\lambda t} dt = \lambda \cdot \int_{0}^{\infty} t \cdot e^{-\lambda t} dt = \lambda \cdot \left[\frac{-\lambda t - 1}{\lambda^2} \cdot e^{-\lambda t}\right]\Bigg|_0^{\infty}$$

$$E(T) = \frac{1}{\lambda} \cdot [(-\lambda t - 1) \cdot e^{-\lambda t}]\Big|_0^{\infty} = \frac{1}{\lambda}$$

6.3 Erwartungswert einer Funktion

Eine Zufallsvariable X kann mit Hilfe einer Funktion g(x) in eine neue Zufallsvariable abgebildet werden mit Z = g(x)
Für den Erwartungswert folgt:
6-3-1(diskret)

$$E(Z) = \sum_i g(x_i) f(x_i)$$

6-3-2 (stetig)

$$E(Z) = \int_{-\infty}^{\infty} g(x_i) f(x_i) dx$$

Beispiel:
X sei eine stetige Zufallsvariable mit der Dichtefunktion f(x) und der Abbildung g(x)

Streuungsmaß

Da die einzelnen Werte x_i der Zufallsvariablen X um ihren Mittelwert μ streuen, definiert man ein sog. Streuungsmaß

1. Idee:

Summe der Abweichungen $(x_i - \mu)$ scheidet aus, da das Ergebnis immer 0 ist

$$\sum_{i=1}^{n} (x_i - \mu) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \mu = n \cdot \mu - n \cdot \mu = 0$$

weil:

$$\mu = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

2. Idee:

Summe der Abweichungsquadrate $(x_i - \mu)^2$
Funktioniert, da nun alle Terme positiv

6.4 Mittelwert, Varianz, Standardabweichung (diskret)

Mittelwert:

Varianz:

Standardabweichung:

Anwendung:

- Die Varianz σ^2 kann als Erwartungswert der Zufallsvariable $Z = (X-\mu)^2$ aufgefasst werden, welche die mittlere quadratische Abweichung vom Mittelwert μ bzw. Erwartungswert $E(X)$ mit $E(X)=\mu$ beschreibt.

- σ hat als Wurzel der Varianz dieselbe Einheit wie die Zufallsvariable und ist ein Maß für die Streuung der Zufallsvariable um den Erwartungswert $E(X)$
  Im Intervall $[\mu-\sigma, \mu+\sigma]$ liegen ca. 68% aller Werte (1-Sigma Intervall).

6.5 Mittelwert, Varianz, Standardabweichung (stetig)

Mittelwert:

Varianz:

Standardabweichung:

Hinweis (gilt für diskret & stetig):

$$\sigma^2 = E[(X-\mu)^2] = E(X^2 - 2\mu X + \mu^2)$$

$$\sigma^2 = E(X^2) - 2\mu E(X) + \mu^2 E(1)$$

Beispiel 1:

Wurf mit 2 Würfel, X = Erzielte Augenzahl

x_i	2	3	4	5	6	7	8	9	10	11	12
$f(x_i)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Mittelwert μ =

Varianz σ^2 =

Standardabweichung σ =

Beispiel 2: (Variation mit Wiederholung u. Reihenfolge)

Urne mit 3W, 2S, Ziehung von 3 Kugeln mit Zurücklegen, X = Anzahl schwarzer Kugeln

x_i	0	1	2	3
$n^k = 2^3$ Möglichkeiten	○ ○ ○	● ○ ○ ○ ● ○ ○ ○ ●	● ● ○ ● ○ ● ○ ● ●	● ● ●
$f(x_i)$				

Mittelwert μ =

Varianz σ^2 =

Beispiel 3:

X sei eine stetige Zufallsvariable mit der linearen
W.-Dichtefunktion f(x) = 0.02x für 0 <= x <= 10

Erwartungswert/Mittelwert μ:

Varianz σ^2:

6.6 Mittelwert und Varianz einer linearen Funktion

X: diskrete/stetige Zufallsvariable mit Mittelwert $E(X)=\mu_X$ und Varianz σ_X^2

Z: Neue Zufallsvariable, von X abhängig, mit
$Z = g(x) = aX + b$ $\quad\quad\quad$ a,b = const.

mit $\mu_Z = E(Z) = E(aX + b) = aE(X) + bE(1) = a\,\mu_X + b$

$\sigma_Z^2 = E[(Z - \mu_Z)^2] = E(Z^2) - \mu_Z^2 = a^2\,\sigma_X^2$

Herleitungen s. Papula S. 348

Beispiel:
Eine Zufallsvariable X besitze die Kennwerte $\mu_X = 10$, $\sigma_X^2 = 1$
Für $Z = 2X + 1$ folgt ($a = 2$, $b = 1$)

$\mu_Z = a\,\mu_X + b = 2\,\mu_X + 1 = 21$
$\sigma_Z^2 = a^2\,\sigma_X^2 = 4$

Beispiel:

X sei eine Zufallsvariable mit $\mu_X = \mu$ und $\sigma^2_X = \sigma^2$

Durch die lineare Transformation

$$Z = \frac{X-\mu}{\sigma} = \frac{1}{\sigma}\cdot(X-\mu) = \frac{1}{\sigma}\cdot X - \frac{\mu}{\sigma} = aX+b$$

wird die Zufallsvariable X in eine Zufallsvariable Z mit $\mu_Z = 0$ und $\sigma^2_Z = 1$ transformiert

Kontrolle:

$$\mu_Z = a\cdot\mu_X + b = \frac{1}{\sigma}\cdot\mu - \frac{\mu}{\sigma} = 0$$

$$\sigma_Z^2 = a^2\cdot\sigma_X^2 = \left(\frac{1}{\sigma^2}\right)\cdot\sigma^2 = 1$$

Standardisierung: Z ist die zu X gehörende standardisierte Zufallsvariable

7 Spezielle Wahrscheinlichkeitsverteilungen

7.1 Bernoulli Experiment

Es existieren nur **zwei**, sich gegenseitig ausschließende, Ereignisse mit konstanten Wahrscheinlichkeiten. Bernoulli Experimente führen zur sog. **Bi**nomialverteilung
Hinweis: Bei Laplace Experimenten haben alle Ereignisse dieselbe Wahrscheinlichkeit

7.2 Binomialverteilung (diskret)

n-fache Ausführung eines Bernoulli-Experiments mit dem Ereignis A und P(A) = const. = p

Zufallsvariable X: Anzahl der Versuche, in denen das Ereignis nach n-facher Ausführung eingetreten ist. X kann somit die Werte 0,1,2,3,... n annehmen.
Gesucht: Wahrscheinlichkeit, dass das Ereignis A bei n-Versuchen genau x-mal eintritt.

Herleitung der Binomialverteilung

Beispiel: Urnenmodell, weiße und schwarze Kugeln
p: Wahrscheinlichkeit eine weiße Kugel zu ziehen, q: W. einer schw. Kugel mit q = 1-p
Ereignis A: Ziehen einer weißen Kugel, Ereignis $\bar{A}$: Ziehen einer schwarzen Kugel
n-Ziehungen **mit** zurücklegen, p,q sind const.
X: Anzahl der gezogenen weißen Kugeln nach n-Versuchen mit zurücklegen

Ergebnis nach n-Versuchen:

x-mal (n-x)mal

7-2-1

p * p * p * .... * p * q * q * q * * q = p^x q$^{(n-x)}$

Die Anzahl der Permutationen gibt an, wieviel Möglichkeiten es gibt, x-weiße Kugeln in n-Versuchen zu erhalten. Jede Permutation hat die Wahrscheinlichkeit gemäß Gl. 7-2-1 .

Somit ergibt sich für die **Wahrscheinlichkeitsfunktion** der diskreten Binomialverteilung:
7-2-2

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

Parameter/Kennwerte der Binomialverteilung: n,p

Verteilungsfunktion der Binomialverteilung:
7-2-3

$$F(X) = P(X \le x) = \sum_{k \le x} \binom{n}{k} p^k q^{n-k}$$

Kennwerte der Binomialverteilung:

Mittelwert: μ = np, Varianz: σ^2 = npq, Standardabweichung: σ

7.3 Herleitung des Mittelwerts und Varianz der Binomialverteilung:

Für Binomialkoeffizienten gilt:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} = \frac{n}{k} \cdot \binom{n-1}{k-1} \qquad (1)$$

Mit dem binomischen Lehrsatz:

$$(p+q)^n = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \quad \text{bzw.} \quad (p+q)^{n-1} = \sum_{i=0}^{n-1} \binom{n-1}{i} p^i q^{n-1-i} \qquad (2)$$

Daraus ergibt sich für den **Erwartungswert** einer Binomialverteilung:

$$\mu = \sum_{k=0}^{n} k \cdot P(X=k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} k \cdot \binom{n}{k} p^k q^{n-k}$$

k=0 liefert keinen Beitrag

mit (1):

$$\sum_{k=1}^{n} k \cdot \frac{n}{k} \cdot \binom{n-1}{k-1} p^k q^{n-k} = \sum_{k=1}^{n} n \cdot \binom{n-1}{k-1} p^k q^{n-k} = np \cdot \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k}$$

Substitution: k = i+1 bzw. i = k-1:

$$\mu = np \cdot \sum_{i=0}^{n-1} \binom{n-1}{i} p^i q^{n-(i+1)} = np \cdot \sum_{i=0}^{n-1} \binom{n-1}{i} p^i q^{n-1-i}$$

mit (2):

$$\boxed{\mu = n \cdot p \cdot (p+q)^{n-1} = n \cdot p \cdot (1)^{n-1} = n \cdot p}$$

Es gilt also:

$$\sum_{i=0}^{n} i \cdot \binom{n}{i} p^i q^{n-i} = n \cdot p$$

bzw.:

$$\sum_{i=0}^{n-1} i \cdot \binom{n-1}{i} p^i q^{n-1-i} = (n-1) \cdot p \qquad (3)$$

Damit erhält man die **Varianz der Binomialverteilung**:

$$\sigma^2 = \sum_{k=0}^{n} (k-\mu)^2 \cdot P(X=k)$$

$$= \sum_{k=0}^{n} (k^2 - 2k\mu + \mu^2) P(X=k)$$

$$= \sum_{k=0}^{n} k^2 P(X=k) - 2\mu \sum_{k=0}^{n} k \cdot P(X=k) + \mu^2 \sum_{k=0}^{n} P(X=k)$$

$$= \sum_{k=1}^{n} k^2 P(X=k) - 2 \cdot \mu \cdot \mu + \mu^2 \cdot 1 = \sum_{k=1}^{n} k^2 \cdot P(X=k) - \mu^2$$

mit (3):

$$= \sum_{k=1}^{n} k^2 \binom{n}{k} p^k q^{n-k} - (np)^2$$

und mit (1):

$$= \sum_{k=1}^{n} k^2 \cdot \frac{n}{k} \cdot \binom{n-1}{k-1} p^k q^{n-k} - n^2 p^2 = \sum_{k=1}^{n} k \cdot n \cdot \binom{n-1}{k-1} p^k q^{n-k} - n^2 p^2$$

$$= np \cdot \left[\sum_{k=1}^{n} k \cdot \binom{n-1}{k-1} p^{k-1} q^{n-k} - np\right]$$

Substitution: k = i+1, i = k-1:

$$= np \cdot \left[\sum_{i=0}^{n-1} (i+1) \cdot \binom{n-1}{i} p^{i} q^{n-(i+1)} - np\right]$$

$$= np \cdot \left[\sum_{i=0}^{n-1} i \cdot \binom{n-1}{i} p^{i} q^{n-1-i} + \sum_{i=0}^{n-1} \binom{n-1}{i} p^{i} q^{n-1-i} - np\right]$$

zum Schluss noch mit (1) und (3):

$$= n \cdot p \cdot [(n-1) \cdot p + (p+q)^{n-1} - n \cdot p] = n \cdot p \cdot [n \cdot p - p + 1 - n \cdot p]$$

$$\boxed{\sigma^2 = n \cdot p \cdot [1-p] = n \cdot p \cdot q}$$

Beispiel 1: **Qualitätskontrolle**
Es werden Bauteile mit einem Ausschussanteil von 2% produziert. Mit welchen Wahrscheinlichkeiten findet man in einer Zufallsstichprobe von 5 Bauteilen genau 0,1,2,3,4,5 defekte Bauteile?

Lösung: Es handelt sich hierbei um ein Bernoulli Experiment mit p = 0.02
Bei sehr großen Stückzahlen spielt Zurücklegen oder nicht keine Rolle, sonst siehe Abschnitt Hypergeometrische Verteilung (Abschn. 7.4).

X = Anzahl der defekten Bauteile unter den 5 entnommenen
X sei binomialverteilt mit den Kennwerten n=5, p=0.02

$$f(x) = P(X = x) = \binom{5}{x} 0,02^{x} \cdot 0,98^{5-x}$$

X	0	1	2	3	4	5
P(X=x)	0,9039	0,0922	0,0038	0,00007	0	0

-> P(X=1) = f(1) = 0,0922 = 9,2%

Beispiel 2: Statistisch gesehen, haben 97% aller Jugendlichen ein Smartphone.
a) Wie groß ist die Wahrscheinlichkeit, dass in einer Gruppe von 100 Jugendlichen alle ein Smartphone besitzen?

Lösung:
X = Anzahl der Smartphone-Besitzer

Bernoulli Experiment mit p = 0,97, n = 100, k = 100

$$P(X = 100) = \binom{100}{100} 0,97^{100} \cdot 0,03^{0} = 0,476 = 47,6\%$$

b) Wieviele Jugendliche müssen mindestens befragt werden, um mit 99% Sicherheit mindestens eine Person zu finden, die kein Smartphone besitzt?

7.4 Hypergeometrische Verteilung (diskret)

In best. Anwendungen werden Stichproben aus einer überschaubaren Menge **ohne** zurücklegen gezogen (z.B. Qualitätskontrolle).

Herleitung am Urnenmodell: N Kugeln, M weiße, N-M schwarze
Ziehung von n Kugeln ohne zurücklegen. Hierfür gibt es

$$m = \binom{N}{n}$$

verschiedene Möglichkeiten.

Unter den n gezogenen Kugeln sollen sich x weiße K. befinden, die aus der Gesamtmenge M der Urne stammen.

Aus M-Elementen kann eine Auswahl x auf

$$g_1 = \binom{M}{x}$$

verschiedene Arten erfolgen.

Unter den n gezogenene Kugeln befinden sich dann (n-x) schwarze K., die aus der Gesamtmenge (N-M) schwarze Kugeln der Urne stammen. Dies kann auf

$$g_2 = \binom{N-M}{n-x}$$

verschiedene Arten erfolgen

Somit gibt es

$$g = g_1 \cdot g_2 = \binom{M}{x}\binom{N-M}{n-x}$$

verschieden Möglichkeiten, der Urne n Kugeln so zu entnehmen, dass sich darunter genau x weiße Kugeln (bzw. n-x schwarze) befinden.

Für die Wahrscheinlichkeit erhält man
7-4-1

$$f(x) = \frac{\text{günstige Fälle}}{\text{alle Fälle}} = P(X = x) = \frac{g}{m} = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$$

7-4-2

$$F(x) = \sum_{k \le x} f(k)$$

Kennwerte:

Mittelwert: $\mu = n \cdot \frac{M}{N}$ Varianz $\sigma^2 = \frac{nM(N-M)(N-n)}{N^2(N-1)}$

Die Parameter der hypergeometrischen Verteilung haben folgende Bedeutung:

N: Anzahl der Elemente in der Grundgesamtheit
M: Anzahl der Elemente in der Grundgesamtheit mit der Eigenschaft A
n: Anzahl der entnommenen Elemente (Umfang der Stichprobe)
x: Anzahl der Elemente in der Stichprobe mit der Eigenschaft A

Beispiel: Eine Lieferung enthält N=100 Transistoren aus einer Massenproduktion mit 5% Ausschuss. Bei einer Kontrolle wird eine Stichprobe vom Umfang n=4 ohne Zurücklegen entnommen. Mit welcher Wahrscheinlichkeit enthält die Stichprobe nur einwandfreie Bauteile?

Lösung:
X = Anzahl defekter Transistoren in der Stichprobe mit n=4

$$f(x) = P(X=x) = \frac{\binom{5}{x} \cdot \binom{100-5}{4-x}}{\binom{100}{4}} = \frac{\binom{5}{x} \cdot \binom{95}{4-x}}{\binom{100}{4}}$$

Für x=0: P(X=0) = f(0) = 0,8119 = 81,2%

Voriges Beispiel kann auch mit Hilfe der Binomialverteilung gelöst werden. Die Ergebnisse sind vergleichbar, da mit N=100 eine genügend hohe Stückzahl vorliegt und der Unterschied zwischen mit/ohne Zurücklegen sich dann nicht mehr bemerkbar macht.

Hypergeometr. Verteilung:

$$f(x) = \frac{\binom{5}{x} \cdot \binom{95}{4-x}}{\binom{100}{4}}$$

Binomialverteilung:

$$f(x) = \binom{4}{x} \cdot 0,05^{x} \cdot 0,95^{4-x}$$

Vergleich:

Hypergeometr. Verteilung		Binomialverteilung	
x = 0:	f(0) = 0,8118	x = 0:	f(0) = 0,8145
x = 1:	f(1) = 0,176	x = 1:	f(1) = 0,176
x = 2:	f(2) = 0,005	x = 2:	f(2) = 0,013

Beispiel:

In einem Kurs befinden sich 30 Studierende, darunter 3 aus derselben Firma X-GmbH.
Für eine Teamarbeit werden zufällig 5 Mitglieder ausgewählt.
Wie groß sind die Wahrscheinlichkeiten, dass sich in der Gruppe 0,1,2 oder 3 Studierende der Firma X-GmbH befinden?

Parameter: N=30, M=3, n=5, x=[0,1,2,3]

Berechnung mit R

```
> x = seq(0,3,by=1)
> dhyper(x,3,27,5)
[1] 0.566502463 0.369458128 0.061576355 0.002463054
```

Herleitung der Mittelwertes der Hypergeom. Verteilung

Prinzipiell ergibt sich der Erwartunsgwert wieder aus der Summe aller in Frage kommenden Werte, gewichtet mit den zugehörigen Wahrscheinlichkeiten:

$$\mu = \sum_{x=0}^{n} x \cdot \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} = \sum_{x=0}^{n} x \cdot \frac{\frac{M}{x} \cdot \binom{M-1}{x-1} \cdot \binom{N-M}{n-x}}{\frac{N}{n} \cdot \binom{N-1}{n-1}}$$

$$\mu = \frac{n \cdot M}{N}\left[\sum_{x=0}^{n} \frac{\binom{M-1}{x-1} \cdot \binom{N-M}{n-x}}{\binom{N-1}{n-1}}\right]$$

Substitution: x = y + 1

$$\mu = \frac{n \cdot M}{N}\left[\sum_{y=0}^{n-1} \frac{\binom{M-1}{y} \cdot \binom{N-1-(M-1)}{n-1-y}}{\binom{N-1}{n-1}}\right] = \frac{n \cdot M}{N}$$

Die eckige Klammer stellt wieder eine Hypergeom. Verteilung F(y) dar, sie könnte durch Variablensubstitution wieder in die urspr. Form gebracht werden. Die Summe über alle Werte muss 1 ergeben (Eigenschaft der Verteilungsfunktion).

Auch die **Herleitung der Varianz einer Hypergeom. Verteilung** erfolgt ähnlich wie für die Binomialverteilung gezeigt:

$$Var(x) = \sum_{x=0}^{n}\left(x - \frac{nM}{N}\right)^2 \cdot \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$$

und ausmultipliziert:

$$Var(x) = \sum_{x=0}^{n}\left[(x)^2 \cdot \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} - 2x \cdot \frac{nM}{N} \cdot \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} + \left(\frac{nM}{N}\right)^2 \cdot \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}\right]$$

der zweite Term in der Klammer entspricht 2nM/N * μ (s. oben) = 2(nM/N)^2, der dritte Term ergibt (nM/N)^2 * 1 und zusammengefasst:

$$Var(x) = -\left(\frac{nM}{N}\right)^2 + \sum_{x=0}^{n}(x)^2 \cdot \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} = \ldots = \frac{nM}{N}\left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1}$$

Die algebraischen Umformungen werden hier nicht weiter ausgeführt (s.o.).

7.5 Poisson Verteilung (diskret)

Die Poisson Verteilung entsteht aus der Binomialverteilung durch den Grenzübergang:

n -> unendl., p -> 0 mit μ = np = const.

D.h. der Umfang der Stichprobe n sei sehr groß und die Wahrscheinlichkeit des Ereignisses (Bernoulli) sehr klein.

7-5-1

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

7-5-2

$$F(x) = P(X \le x) = e^{-\mu} \cdot \sum_{k \le x} \frac{\mu^k}{k!}$$

Mittelwert: μ
Varianz: $\sigma^2 = \mu$

Der einzige Parameter der Poissonverteilung ist der Mittel- oder Erwartungswert.

Beispiel: **Radioaktiver Zerfall**

X = Anzahl der Atomkerne, die in einer Sekunde zerfallen
μ = Durchschnittliche Anzahl zerf. Atomkerne pro Sekunde
Die Poissonvert. kann hier angenommen werden, da eine sehr große Zahl Atome betrachtet wird, die Wahrscheinlichkeit für den Zerfall eines einzelnen Atoms dabei aber sehr gering ist.

Sei μ = 2 Zerf./Sekunde, wie groß ist die Wahrscheinlichkeit, dass ein Zählgerät mehr als zwei Zerfälle pro Sekunde registriert?

Lösung:

$$f(x) = P(X = x) = \frac{2^x}{x!} \cdot e^{-2}$$

für X > 2:

$$P(X > 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

$$P(X > 2) = 1 - \frac{2^0}{0!} \cdot e^{-2} - \frac{2^1}{1!} \cdot e^{-2} - \frac{2^2}{2!} \cdot e^{-2}$$

$$P(X > 2) = 1 - (1 + 2 + 2)e^{-2} = 1 - 5e^{-2} = 0,323$$

Ergebnis/Erwartung: Bei 100 Messungen zeigt das Messgerät in 32 Fällen mehr als 2 Zerfälle an.

7.6 Von der diskreten zur stetigen Verteilung

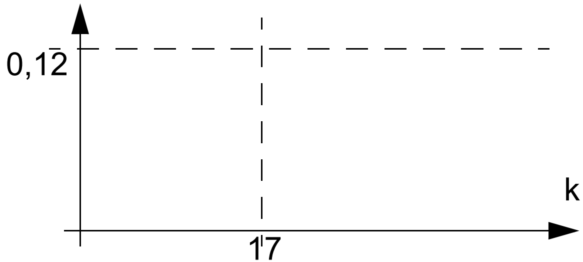
Die Binomialverteilung beschreibt die Anzahl Erfolge (k) in einer Serie von (n) gleichartigen und unabhängigen Versuchen, die jeweils genau zwei mögliche Ergebnisse haben (Erfolg/Misserfolg). D.h. ein Bernoulli-Experiment mit der Wahrscheinlichkeit p.

Die Wahrscheinlichkeitsfunktion f(x) besteht dabei aus diskreten Wertepaaren (x, f(x)) „Anzahl Erfolge/Wahrscheinlichkeit"

Beispiel: n=50, p=1/3 -> μ = np = 50/3, σ^2 = npq, σ = 10/3

Wahrscheinlichkeitsfunktion zeichnen mit R:

```
> n = 50
> k = seq(0,n,by=1)
> plot(k, dbinom(k,n,1/3))
```



Die Binomialverteilung lässt nur dieskrete Werte für k zu. Messwerte in vielen natur-, wirtschafts- und ingenieurwissenschaftl. Vorgängen sind hingegen kontinuierlich bzw. stetig. Sind die Messfehler bzw. Abweichungen vom Messwert zufälliger Natur, lassen sie sich mit Hilfe der Normal- bzw. Gaußverteilung beschreiben.

Versuch die Binomialverteilung mit einer stetigen Funktion zu approximieren:

Ansatz:
Die Funktion

$$y_1(x) = u \cdot e^{vx^2}$$

ist achsensymmetrisch und geht für **negative v** mit wachsendem |x| gegen 0.

1. Forderung: Lage des Maximus $x_{max} = \mu$

$$y_2(x) = u \cdot e^{v(x-\mu)^2}$$

2. Forderung: Wendepunkte $x_{WP} = \mu$ +/- σ
   Die 2. Ableitung von y_2 muss an den Stellen x_{WP} Null werden

$$x_{WP} = \mu \pm \sigma = \mu \pm \frac{1}{\sqrt{-2v}}$$

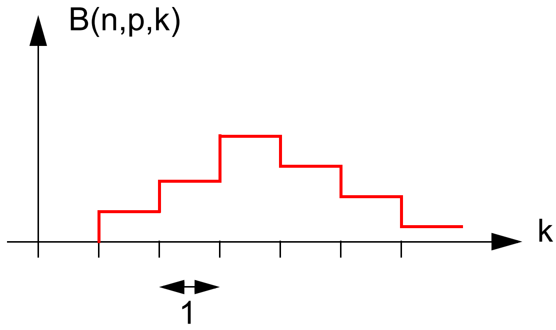
$$\sigma = \frac{1}{\sqrt{-2v}}$$

$$v = -\frac{1}{2\sigma^2}$$

$$y_3(x) = u \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

3. Forderung: Für die Binomialverteilung gilt:

$$\sum_{k=0}^{n} B(n,p,k) = 1$$



$$\sum_{k=0}^{n} B(n,p,k) = 1$$

$$\int_{-\infty}^{\infty} u e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

Substitution: $t = (x-\mu)/\sigma$, $dt/dx=1/\sigma$

$$u\sigma\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t)^2} dt = 1$$

Integral ist nicht geschl. lösbar mit

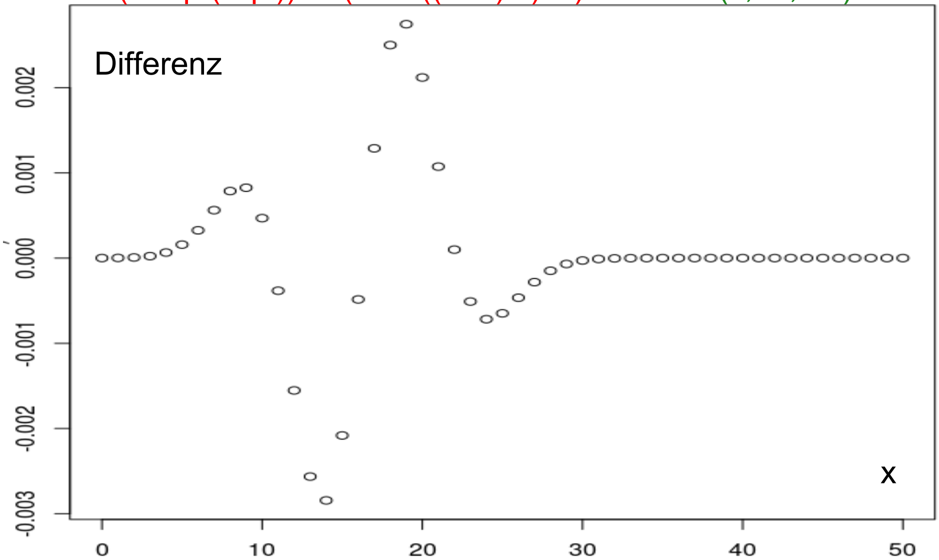
$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t)^2} dt = \sqrt{2\pi} = \frac{1}{u\sigma}$$

7-6-1

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

Vergleich: Approx. nach Gl. 7-6-1 mit exakter Binomialverteilung

1/(s*sqrt(2*pi))*e^(-0.5*((x-m)/s)^2) - dbinom(x,50,1/3)



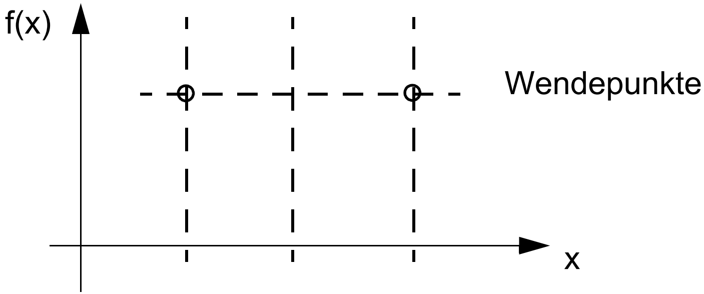

```
> e=2.71828
> pi=3.1415
> m=50/3
> s=10/3
> x=seq(0,50,by=1)
> plot(x,1/(s*sqrt(2*pi))*e^(-0.5*((x-m)/s)^2)-dbinom(x,50,1/3))
```

7.7 Gaußsche Normalverteilung (stetig)

Dichtefunktion:

7-7-1

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{für } -\infty < x < \infty$$



Verteilungsfunktion:

7-7-2

$$F(x) = P(-\infty < X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

7.8 Standardnormalverteilung

Die Gaußsche Normalverteilung mit den Parametern μ und σ lässt sich mit einer Transformation der Zufallsvariablen X in die sog. Standardnormalverteilung mit μ=0 und σ=1 überführen.

$$u = \frac{t-\mu}{\sigma}$$

$$dt = \sigma \cdot du$$

dabei transformieren sich die Integrationsgrenzen zu:

Untere Grenze: t = -unendl. -> u = -unendl.

Obere Grenze: t = x -> u = (x-μ)/σ

somit wird aus Gl. 7-7-2

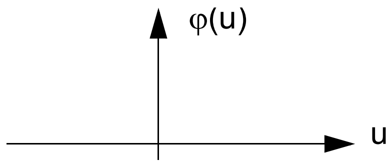
7-8-1

$$\phi(u) = P(U \le u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}(u)^2} du = \phi\left(\frac{x-\mu}{\sigma}\right)$$

hierzu gehört die Dichtefunktion:

7-8-2

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(u)^2}$$



Beispiel:
Die Zufallsvariable X sei normalverteilt mit dem Mittelwert μ = 10 und der Varianz σ^2 = 4
Wieviel Prozent aller Werte liegen dann im Intervall [5, 12]?

Lösung:
Transformation in Standardnormalverteilung

Untere Grenze: a -> a* = (a-μ)/σ = (5-10)/2 = -2.5
Obere Grenze: b -> b* = (a-μ)/σ = (12-10)/2 = 1

$$P(5 \leq X \leq 12) \;=\; P(-2,5 \leq U \leq 1) \;=\; \Phi(1) - \Phi(-2,5)$$

s. Tabelle

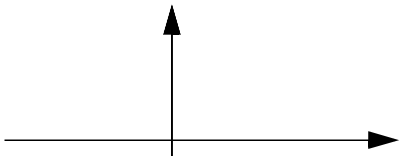
$$\Phi(1) - \Phi(-2,5) \;=\; 0,8413 - (1 - 0,9938) \;=\; 0,8351 \;=\; 83,5\%$$

7.9 Praktische Anwendung der Standardnormalverteilung

Da Gl. 7-8-1 nicht geschlossen lösbar ist, werden Tabellen mit Näherungswerten benutzt

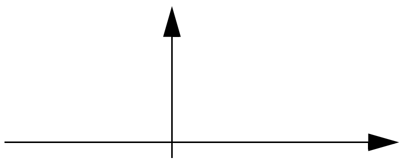
$$\phi(u) = P(U \le u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-\frac{1}{2}(t)^2}\, dt$$

a) u >= 0



$\phi(u)$ direkt aus Tabelle

b) u < 0



$\phi(u) = 1 - \phi(|u|)$
wg. Achsensymmetrie

c) Berechnung von P(X <= x)

$$F(x) = P(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

Substitution

$$u = \frac{t-\mu}{\sigma}$$

$$dt = \sigma \cdot du$$

dabei transformieren sich die Integrationsgrenzen zu:

Untere Grenze: t = -unendl. -> u = -unendl.

Obere Grenze: t = x -> u = (x-μ)/σ

$$F(x) = P(X \le x) = \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}(u)^2} du = \phi\left(\frac{x-\mu}{\sigma}\right) = \phi(u)$$

d) Berechnung von P(X >= x)

$$P(X \ge x) = 1 - P(X \le x) = 1 - \phi\left(\frac{x-\mu}{\sigma}\right) = 1 - \phi(u)$$

e) Berechnung von P(a <= X <= b)

$$P(a \le X \le b) = F(b) - F(a) = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right) = \phi(b^*) - \phi(a^*)$$

f) Intervalle symmetrisch zum Mittelwert μ

$$P(|X-\mu| \le k\sigma) = \phi(k) - \phi(-k) = \phi(k) - [1 - \phi(k)] = 2\phi(k) - 1$$

Zahlenwerte:

k=1

$$P(|X-\mu| \le \sigma) = P(-1 \le U \le 1) = 2\phi(1) - 1 = 0,6826$$

d.h. 68,3% aller Werte der Zufallsvariablen X liegen im Intervall $[\mu-\sigma, \mu+\sigma]$

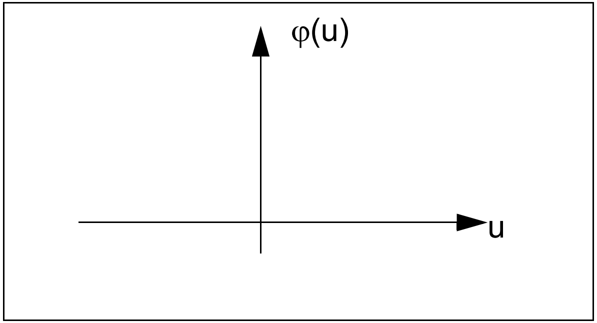
k=2

$$P(|X-\mu| \le 2\sigma) = P(-2 \le U \le 2) = 2\phi(2) - 1 = 0,9544$$

k=3

$$P(|X-\mu| \le 3\sigma) = P(-3 \le U \le 3) = 2\phi(3) - 1 = 0,9974$$

7.10 Quantile der Standardnormalverteilung



Zu einer vorgegebenen Wahrscheinlichkeit p gehört eine eindeutig bestimmte Schranke U_p, die als Quantil der Standardnormalverteilung bezeichnet wird.

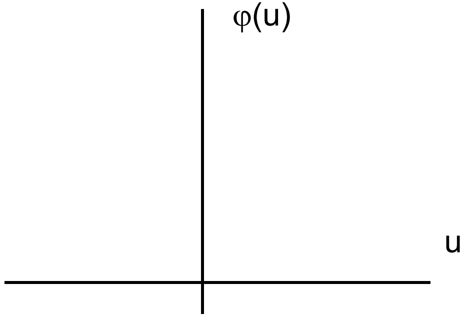
$$\boxed{P(u \le u_p) = \phi(u_P) = p}$$

Beispiel 1:
Die standardnormalverteilte Zufallsvariable U soll mit der Wahrscheinlichkeit p=0.9 einen Wert aus dem Intervall U <= c annehmen.
Gesucht ist das zum Wahrscheinlichkeitswert p = 0.9 gehörende Quantil $u_p = u_{0.9} = 1.282$ (s. Tabelle).

Beispiel 2:
Die Wete der standardnormalverteilten Zufallsvariablen U sollen mit p = 0.9 oberhalb der Schranke c liegen.

$P(U > c) = 1 - P(U <= c) = 1 - \Phi(c) = 0.9$
Daraus folgt: $\Phi(c) = 0.1 \rightarrow c = u_{0.1} = -1.282$

Beispiel 3:
Bei einer zweiseitigen Abgrenzung sollen die Werte der standardnormalverteilten Zufallsvariable U mit p = 0.95 im symmetrischen Intervall -c <= U <= c liegen.
Gesucht: Werte für c so, dass P(-c <= U <= c) = 0.95



7.11 Exponentialverteilung

Die Funktions- oder Lebensdauer eines techn. Systems lässt sich durch eine Zufallsvariable T beschreiben mit der Dichtefunktion:

7-11-1

$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda t} & \text{für t } \geq 0 \\ 0 & \text{sonst} \end{cases}$$

bzw. Verteilungsfunktion:
7-11-2

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{für t } \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Kennwerte:

Mittelwert: $\mu = 1/\lambda$
Varianz: $\sigma^2 = 1/\lambda^2$

Ausfallwahrscheinlichkeit:

Wahrscheinlichkeit, dass die Lebensdauer T des Systems den Zeitpunkt t nicht überschreitet.

$$F(t) = P(T \leq t) = 1 - e^{-\lambda t}$$

Zuverlässigkeitsfunktion:

$$R(t) = P(T \geq t) = 1 - P(T \leq t) = 1 - F(t) = e^{-\lambda t}$$

Beispiel:
Die mittlere Lebensdauer eines Bauteils betrage $\overline{T}$=1000h
$\overline{T} = 1/\lambda = 1000h \rightarrow \lambda = 1/1000h = 0{,}001h^{-1}$

Ausfallwahrscheinlichkeit

$$F(t) = 1 - e^{-0,001\frac{t}{h}}$$

Überlebenswahrscheinlichkeit

$$R(t) = e^{-0,001\frac{t}{h}}$$

a) Wie groß ist die Wahrscheinlichkeit, dass das Bauteil innerhalb der ersten 2000 Betriebsstunden ausfällt?

b) Mit welcher Wahrscheinlichkeit fällt das Bauelement zwischen der 1000. und 2000. Betriebsstunde aus?

7.12 Zusammenhang zw. Binomial- und Normalverteilung

vgl. hierzu auch Abschn. 7.6

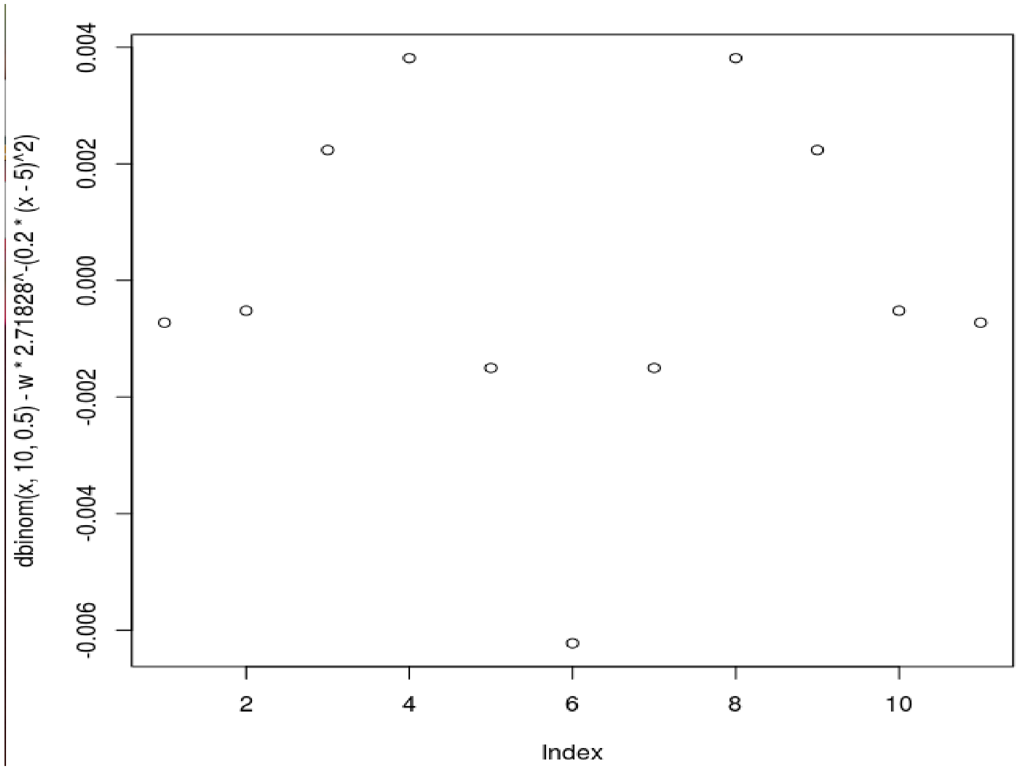
$$f_B(x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$$

für große n ist eine Approx. durch Gauß mit μ = np, σ^2 = npq möglich:

$$\boxed{f_B(x) = \binom{n}{x} \cdot p^x \cdot q^{n-x} \approx \frac{1}{\sqrt{2\pi} \cdot \sqrt{npq}} \cdot e^{-\frac{1}{2}\left(\frac{x-np}{\sqrt{npq}}\right)^2}}$$

Beispiel: n = 10, p = 1/2, q = 1/2

Vergleich


dbinom(x, 10, 0.5) - w * 2.71828^-(0.2 * (x - 5)^2)
0.004
0.002
0.000
-0.002
-0.004
-0.006
2
4
6
8
10
Index

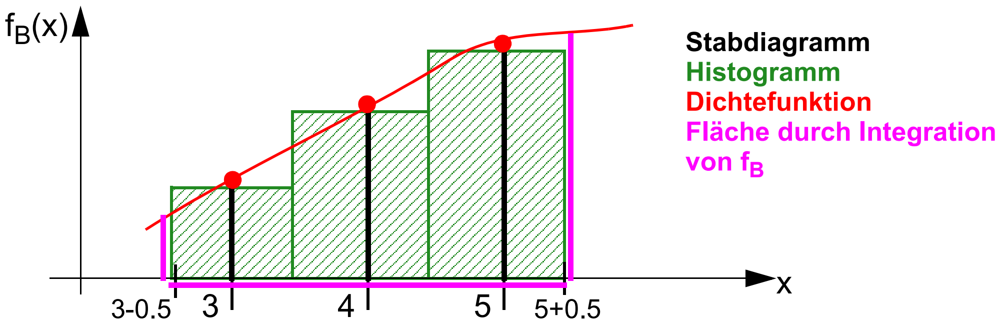
Anwendung und Notwendigkeit der sog. Stetigkeitskorrektur

$$P(a \le X \le b) = \sum_{x=a}^{b} \binom{n}{x} p^x q^{n-x} \approx F(b+0,5) - F(a-0,5)$$

$$P(a \le X \le b) \approx P(a^* \le U \le b^*) = \Phi(a^*) - \Phi(b^*)$$

$$a^* = \frac{(a-0,5)-\mu}{\sigma} \qquad b^* = \frac{(b+0,5)-\mu}{\sigma}$$

Stetigkeits-korrektur



Die Rechtecke des Histogramms haben die Breite $\Delta x = 1$, die Höhe entspricht der Dichtefunktion $f_B(x)$, die Gesamtfläche der Rechtecke im Histogramm entspricht der gesuchten Wahrscheinlichkeit P(a <= X <= b).
Bei einer Approximation durch die Normalverteilung müssen die Intervallgrenzen um 0,5 nach aussen verschoben werden (sog. Stetigkeitskorrektur).

$$P(a \le X \le b) \approx \int_{2,5}^{5,5} \frac{1}{\sqrt{5\pi}} e^{-\frac{1}{5}(x-5)^2} dx = F(5,5) - F(2,5)$$

mit

$$a^* = \frac{2,5-5}{\sqrt{2,5}} \qquad b^* = \frac{5,5-5}{\sqrt{2,5}}$$

$$a^* = -1,581 \qquad b^* = 0,316$$

$$P(3 \le X \le 5) \approx P(-1,581 \le U \le 0,316) = \phi(0,316) - \phi(-1,581)$$

$$\phi(0,316) - \phi(-1,581) = \phi(0,316) - [1 - \phi(1,581)] = \phi(0,316) + \phi(1,581) - 1 = 0,567$$

7.13 Stetige Gleichverteilung

Das einfachste Beispiel einer stetigen Verteilung ist die stetige Gleichverteilung über dem Intervall (0,1). Eine Zufallsvariable X mit der Verteilungsfunktion F(x) heißt gleichverteilt über dem Intervall (0,1) bzw. X ~ U(0,1)

$$F(x) = \begin{cases} 0 & \mathrm{x} \geq 1 \\ x & 0 \leq x < 1 \\ 1 & \mathrm{x} \geq 1 \end{cases} \leftrightarrow f(x) = \begin{cases} 0 & \mathrm{x} < 0 \\ 1 & 0 \leq x < 1 \\ 0 & \mathrm{x} \geq 1 \end{cases}$$

Anwendung: Das Grundmodell für viele computererzeugte (Pseudo-)Zufallszahlen bildet die U(0,1)-Verteilung. Eine gleichverteilte Zufallszahl kann mit Hilfe der Umkehrfunktion so transformiert werden, dass das Ergebnis aus einer anderen Verteilung erscheint.

Beispiel: Monatsrenditen

Für das Management von Geldanlagen ist die zukünftige Kursentwicklung von besonderem Interesse, wobei eine Vorhersage kaum möglich ist. Um dennoch Aussagen treffen zu können, werden Modelle erstellt, mit deren Hilfe eine Simulation möglich ist.
Für eine spezielle Monatsrendite wird aufgrund der Erfahrung eine symmetrische Dreiecksverteilung angenommen. Die Renditen zentrieren sich dabei um den Wert 1:

$$f(x) = \begin{cases} \frac{1}{3} - \frac{1}{9}|x-1| & -2 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

Die Wahrscheinlichkeit, dass die Monatsrendite z.B. zwischen 1,5% und 3,5% liegt beträgt:

$$P(1,5 < x < 3,5) = \int_{1,5}^{3,5} f(x)dx = \int_{1,5}^{3,5} \left(\frac{1}{3} - \frac{1}{9} \cdot (x-1)\right) dx = \left[\frac{4}{9}x - \frac{x^2}{18}\right]\Bigg|_{1,5}^{3,5} = \frac{1}{3}$$

Zu der Wahrscheinlichkeitsdichtefunktion f(x) gehört die Verteilung F(x):

$$F(x) = \begin{cases} 0 & x \le -2 \\ \frac{4}{18} + \frac{4x}{18} + \frac{x^2}{18} & -2 < x \le 1 \\ \frac{2}{18} + \frac{8x}{18} - \frac{x^2}{18} & 1 < x \le 4 \\ 1 & x > 4 \end{cases}$$

mit der zugehörigen Umkehrfunktion:

$$x = F^{-1}(u) = \begin{cases} \sqrt{18u} - 2 & 0 < u \le 0,5 \\ -\sqrt{18(1-u)} + 4 & 0,5 < u < 1 \end{cases}$$

Mit einer gleichverteilten Zufallsvariablen u können nun Zufallszahlen erzeugt werden, die der Verteilungsfunktion F(x) genügen.

Hinweis zur Berechnung der Umkehrfunktion:

$$u = \frac{4}{18} + \frac{4x}{18} + \frac{x^2}{18}$$

$$18u = x^2 + 4x + 4 = (x+2)^2$$

$$\sqrt{18u} - 2 = x$$

- Weitere Beispiele zur Inversionsmethode s. Wiki
- Erzeugung von gleichverteilten Zufallszahlen mit Xor Shifter
- Erzeugung von gleichverteilten Zufallszahlen Kongruenzgenerator

Beispiel Bogenschütze
Ein ungeübter Bogenschütze schießt auf eine Zielscheibe mit dem Durchmesser 1m. Der Abstand Schütze - Zielscheibe wurde so gewählt, dass die Scheibe auf jeden Fall getroffen wird. Die Zufallsvariable X beschreibt den Abstand des Treffers zum Mittelpunkt.

$$P(X \le x) = \frac{\text{Flächeninhalt des Kreises mit Radius x}}{\text{Flächeninhalt der Zielscheibe}} = \frac{\pi x}{\pi \cdot 1^2} = x^2$$

das führt unmittelbar zur Verteilungsfunktion F(x):

$$P(X \leq x) = \begin{cases} 0 & x \leq 0 \\ x^2 & 0 < x \leq 1 \\ 1 & x > 1 \end{cases} = F(x)$$

mit der Dichtefunktion f(x):

$$f(x) = \begin{cases} 2x & 0 < x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

In vielen Anwendungen sind Intervalle interessant, die durch Paare von Quantilen festgelegt werden, welche jeweils die gleiche Wahrscheinlichkeitsmasse links und rechts abspalten. Intervalle $[x_\alpha, x_{1-\alpha}]$, bei denen die Zufallsvariable X mit gleicher Wahrscheinlichkeit α einen Wert links von der Untergrenze bzw. rechts von der Obergrenze annimmt, werden als zentrale Schwankungsintervalle $(1-2\alpha)$ bezeichnet.

$$0,1 = F(x_{q(\alpha)}) = x^2_{q(\alpha)} \Rightarrow x_{q(\alpha)} = \sqrt{\alpha} = 0,3162$$

$$0,9 = F(x_{q(1-\alpha)}) = x^2_{q(1-\alpha)} \Rightarrow x_{q(1-\alpha)} = \sqrt{(1-\alpha)} = 0,9486$$

Das zentrale Schwankungsintervall z.B. für $\alpha = 0.1$ ist demnach [0.3162, 0.9486].
D.h. mit einer Wahrscheinlichkeit von 10% liegen die Treffer unter 31cm bzw. über 95cm, oder mit p=80% innerhalb des Bereichs.

8 Verteilungen von mehreren Zufallsvariablen

- Zufallsexperimente, bei denen gleichzeitig zwei oder mehr Zufallsvariable beobachtet werden.
- Fehlerrechnung, Eigenschaften von Summen/Produkten von unabhängigen Zufallsvariablen (Mittelwerte, Varianz, Standardabweichung)

Beispiel: Gleichzeitiges Werfen einer Münze und eines Würfels
X = Anzahl Wappen
Y= Augenzahl des Würfels
-> Zweidimensionale Zufallsvariable (X;Y)

Wertetabelle für die 12 Elementarereignisse -> Laplace Experiment

X\Y	1	2	3	4	5	6	$f_1(x)$
0 = W	0;1 (p=1/12)	0;2	0;3	0;4	0;5	0;6	1/2
1 = Z	1;1 (p=1/12)	1;2	1;3	1;4	1;5	1;6	1/2
$f_2(y)$	1/6	1/6	1/6	1/6	1/6	1/6	

Addiert man in der Verteilungstabelle die W.-werte zeilenweise:

$f_1(0)$
$f_1(1)$

bzw. spaltenweise:

$f_2(1)$
$f_2(2)$
...

Randverteilung
$f_1(x)$ bzw. $f_2(x)$ beschreiben die Verteilungen der Zufallsvariablen X bzw. Y unabhängig von der jew. anderen Variable

Für obiges Beispiel gilt: $f(x;y) = f_1(x) * f_2(y)$

8.1 Zweidimensionale Verteilungen

Für eine einzelne Zufallsvariable X sei x ein Wert den X annimmt und der einem Punkt auf der X-Achse entspricht.
Ist für jedes Intervall [a,b] die zugehörige Wahrscheinlichkeit bekannt mit $P(a < X \le b)$, so ist die Wahrscheinlichkeitsverteilung der Zufallsvariable X bekannt.

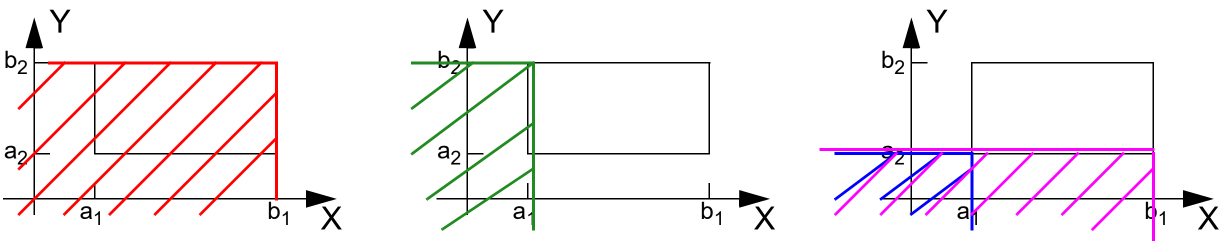
$$F(x) = P(X \le x)$$
$$P(a < X \le b) = F(b) - F(a)$$

Werden gleichzeitig zwei Größen X,Y beobachtet, so ist das Ergebnis eines Experiments ein geordnetes Zahlenpaar X=x, Y=y, was einem Punkt auf der XY-Ebene entspricht.
Kennt man für das Rechteck $a_1 < X \le b_1$ und $a_2 < Y \le b_2$ die zugehörige Wahrscheinlichkeit $P(a_1 < X \le b_1 , a_2 < Y \le b_2)$, so ist die Wahrscheinlichkeitsverteilung der zweidimensionalen Zufallsvariable (X,Y) bekannt.

$$F(x,y) = P(X \le x, Y \le y)$$

bzw. für $F(x,y) = P(a_1 < X \le b_1 , a_2 < Y \le b_2)$

$$F(x,y) = F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2)$$



8.2 Diskrete zweidimensionale Verteilung

$$f(x, y) = \begin{cases} p_{ik} & \text{für } x = x_i, y = y_k \\ 0 & \text{sonst} \end{cases}$$

$$F(x, y) = \sum_{x_i \le x} \sum_{y_k \le y} f(x_i, y_k)$$

mit den Randverteilungen:

$$f_1(x) = \begin{cases} p_i & \text{für } x = x_i \\ 0 & \text{sonst} \end{cases}$$

$$f_2(y) = \begin{cases} p_k & \text{für y} = y_k \\ 0 & \text{sonst} \end{cases}$$

Beispiel: 3maliger Wurf einer homogenen Münze (Zahl, Wappen)
Elementarereignisse: ZZZ, ZZW, ZWZ, WZZ, ZWW, WZW, WWZ, WWW

X=Anzahl „Zahl“ beim ersten Wurf (x = 0,1)
Y=Anzahl „Zahl“ bei drei Würfen

X=0 Wappen beim ersten Wurf

Y=0:	WWW	P(X=0;Y=0) = 1/8
Y=1:	WWZ,WZW	P(X=0;Y=1) = 2/8
Y=2:	WZZ	P(X=0;Y=2) = 1/8
Y=3:	ZZZ = unmöglich	P(X=0;Y=3) = 0

X=1 Zahl beim ersten Wurf

Y=0:	WWW = unmöglich	P(X=1;Y=0) = 0
Y=1:	ZWW	P(X=1;Y=1) = 1/8
Y=2:	ZZW, ZWZ	P(X=1;Y=2) = 2/8
Y=3:	ZZZ	P(X=1;Y=3) = 1/8

2dimensionale Verteilungstabelle

X \ Y	0	1	2	3	
0	1/8	2/8	1/8	0	4/8
1	0	1/8	2/8	1/8	4/8
	1/8	3/8	3/8	1/8	

8.3 Stetige zweidimensionale Verteilung

$$F(x, y) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{y} f(u, v) dv du$$

mit der Normierung:

$$\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f(x, y) dy dx = 1$$

Dichtefunktionen der Randverteilungen:

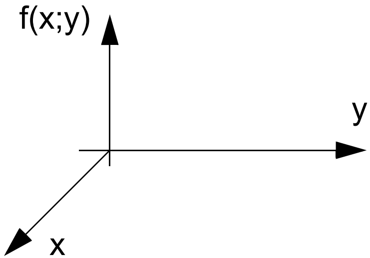
$$f_1(x) = \int_{y=-\infty}^{\infty} f(x, y) dy$$

$$f_2(y) = \int_{x=-\infty}^{\infty} f(x, y) dx$$

Beispiel: Gegeben ist die zweidimensionale Dichtefunktion einer zweidimensionalen Gleichverteilung

$$f(x,y) = \begin{cases} c = const & \text{für } (0 \le x \le 2),(0 \le y \le 5) \\ 0 & \text{sonst} \end{cases}$$

Gesucht: P(X <= 1; Y <= 3) = F(1;3)

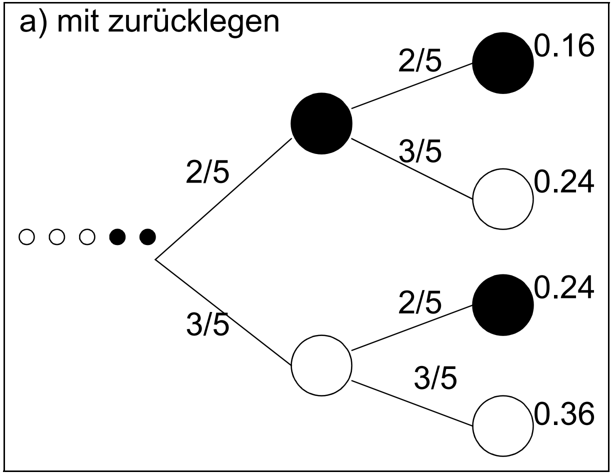


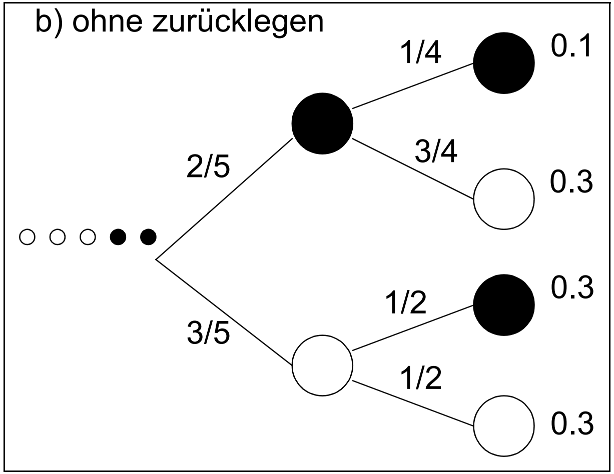
8.4 Stochastisch unabhängige Zufallsvariable

Zufallsvariable X,Y mit den Wahrscheinlichkeits- bzw. Dichtefunktionen $f_1(x)$, $f_2(y)$ und den zugehörigen Verteilungsfunktionen $F_1(x)$, $F_2(y)$ sowie einer gemeinsamen zweidimensionalen Verteilungsfunktion F(x;y) heißen stochastisch unabhängig, falls

$$f(x;y) = f_1(x) * f_2(y)$$
$$F(x;y) = F_1(x) * F_2(y)$$

Beispiel:





Verteilungstabellen:

a) mit zurücklegen

X \ Y	○	●	$f_1(x)$
○	0.36	0.24	0.6
●	0.24	0.16	0.4
$f_2(y)$	0.6	0.4	

b) ohne zurücklegen

X \ Y	○	●	$f_1(x)$
○	0.3	0.3	0.6
●	0.3	0.1	0.4
$f_2(y)$	0.6	0.4	

Ergebnis:

$$f(x;y) = f_1(x) \cdot f_2(y)$$

$$f(x;y) \neq f_1(x) \cdot f_2(y)$$

8.5 Funktionen von mehreren Zufallsvariablen (diskret/stetig)

- Mittelwert einer **Summe** von Zufallsvariablen (Additionssatz)

8-5-1

Sei $Z = a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n$

$E(Z) = E(X_1 + X_2 + X_3 + ... + X_n) = a_1E(X_1) + a_2E(X_2) + a_3E(X_3) + ... + a_nE(X_n)$

$\mu = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + ... + a_n\mu_n$

- Mittelwert eines **Produkts** von stochastisch unabh. Zufallsvariablen

8-5-2

Sei $Z = X_1 * X_2 * X_3 ... * X_n$

$E(Z) = E(X_1 * X_2 * X_3 * ..... * X_n) = E(X_1) * E(X_2) * E(X_3) * .... * E(X_n)$

$\mu = \mu_1 * \mu_2 * \mu_3 * .... * \mu_n$

- Additionssatz für Varianzen

8-5-3

Sei $Z = X + Y$

$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{X,Y}$

σ_X^2, σ_Y^2: Varianzen von X, Y

$\sigma_{X,Y}$: Kovarianz X,Y

8.6 Einschub: Kovarianz

Die Kovarianz (s. Abschn. 14.1) ist ein Zusammenhangsmaß für einen monotonen Zusammenhang zweier Zufallsvariablen mit gemeinsamer Wahrscheinlichkeitsverteilung. Es ist damit eine tendenzielle Aussage darüber möglich, ob hohe Werte der einen Zufallsvariable eher mit hohen oder eher mit niedrigen Werten der anderen Zufallsvariable einhergehen.

Kovarianz positiv: monotoner Zusammenhang hoch -> hoch bzw. niedrig -> niedrig
Kovarianz negativ: gegensinnig monoton hoch -> niedrig, niedrig -> hoch
Kovarianz ist null: kein Zusammenhang

8-6-1Definition:

mit R:

```
> x = rnorm(1000, mean=3.5)          //1000 Zufallszahlen, normalverteilt
> y = rnorm(1000, mean=3.5)
                                     // Mittelwert = 3.5
> var(x)
[1] 1.0216
> cov(x,x)                           //var(x) und cov(x) identisch für große n
[1] 1.0216

> mean((x-mean(x))^2)                // E[(X - E(X))^2]
[1] 1.020579                         // Erwartungswert = Mittelwert für große n
>
> 1000/999*mean((x-mean(x))^2)       // mit n/(n-1) korrigierte Stichproben-
[1] 1.0216                           // varianz, da E(X) nicht bekannt
>
> cov(x,y)
[1] 0.06002309
> mean(x*y) - mean(x)*mean(y)        // E(XY) - E(X) * E(Y)
[1] 0.05996307
> mean((x-mean(x))*(y-mean(y)))      // E[(X-E(X))(Y-E(Y))]
[1] 0.05996307
```

8.7 Beziehung zwischen Varianz und Kovarianz

Die Varianz ist ein Spezialfall der Kovarianz (s.a. Wiki):

Var(X) = Cov(X,X)

Cov(X,X) = E((X - E(X))^2) = Var(X)

- Varianz der Summe von Zufallsvariablen

$$Var\sum_{i=1}^{n} X_i = \sum_{i,j=1}^{n} Cov(X_i, X_j) = \sum_{i=1}^{n} Var(X_i) + \sum_{i,j=1,\, i \neq j}^{n} Cov(X_i, X_j)$$

$$Var\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Var(X_i) + 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Cov(X_i, X_j)$$

- Für die Summe zweier Zufallsvariablen (n=2)

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)$$

Beispiel: Parallelschaltung von Kondensatoren
$C_1 = 30\mu F$, $C_2 = 50\mu F$, $C_3 = 20\mu F$
$\sigma_1 = 2\mu F$, $\sigma_2 = 2\mu F$, $\sigma_3 = 1\mu F$
$C_{ges} = C_1 + C_2 + C_3$
Die Werte der einzelnen Kondensatoren können als unabh. Zufallsvariable aufgefasst werden.

Mittelwert: $\mu = \mu_1 + \mu_2 + \mu_3 = 100\mu F$
Varianz: $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 9(\mu F)^2$
Standardabweichung: $\sigma = 3\mu F$

8.8 Eigenschaften einer Summe von stochastisch unabh. und normalverteilten Zufallsvariablen

8.9 Zentraler Grenzwertsatz

Die Wahrscheinlichkeit einer Summe von Zufallsvariablen mit unbekannter (aber identischer) Verteilung kann unter bestimmten Voraussetzungen als annähernd normalverteilt betrachtet werden.

Seien X_1, X_2, ... ,X_n stochastisch unabh. Zufallsvariable, die alle der gleichen Verteilungsfunktion mit Mittelwert μ und Varianz σ^2 genügen, dann konvergiert die Verteilungsfunktion $F_n(u)$ der standardisierten Zufallsvariable U_n mit

$$U_n = \frac{(X_1 + X_2 + \ldots\ldots + X_n) - (n \cdot \mu)}{\sqrt{n} \cdot \sigma}$$

im Grenzfall n -> unendl. gegen die Verteilungsfunktion $\Phi(u)$ der Standardnormalverteilung

8-9-1

$$\lim_{n \to \infty} F_n(u) = \phi(u) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^u e^{-\frac{1}{2} \cdot t^2} dt$$

Interpretation:
Die Zufallsvariable

$$U_n = \frac{(X_1 + X_2 + \ldots\ldots + X_n) - (n \cdot \mu)}{\sqrt{n} \cdot \sigma} = \frac{Z_n - (n \cdot \mu)}{\sqrt{n} \cdot \sigma}$$

ist für hinreichend großes n (Faustregel n > 30) annähernd standardnormalverteilt.

Oder: Die aus n-Summanden mit der gleichen Verteilungsfunktion bestehende Summe $Z_n = X_1 + X_2 + X_3 + \ldots X_n$ ist für hinreichend großes n annähernd normalverteilt mit

Mittelwert $\mu_Z = E(Z_n) = n * \mu$ und Varianz $\sigma_Z{}^2 = Var(Z_n) = n * \sigma^2$

Allgemein:
Wahrscheinlichkeitsverteilung einer Summe von Zufallsvariablen mit jew. **bel**. Verteilung:
$Z_n = X_1 + X_2 + X_3 + \ldots X_n$ $E(Z_n)=\mu_Z=\mu_1+\mu_2+ \ldots +\mu_n$ $\sigma_Z{}^2=\sigma^2=\sigma_1{}^2+\sigma_2{}^2+\ldots +\sigma_n{}^2$

$$u_n = \frac{Z_n - \mu_Z}{\sigma_Z}$$

$$\lim_{n \to \infty} F_n(u) = \phi(u)$$

Definition: Eine bel. Zufallsgröße X mit dem Erwartungswert μ und der Standardabweichung σ heißt „annähernd normalverteilt“, wenn für ihre Verteilungsfunktion F(x) gilt:

$$F(x) = P(X \le x) \approx \phi\left(\frac{X-\mu}{\sigma}\right)$$

Beispiel: Werfen eines Laplace-Würfels, X = Augenzahl

x_i	1	2	3	4	5	6
$f(x_i) = P(X{=}x)$	1/6	1/6	1/6	1/6	1/6	1/6

$$\mu = \sum_{i=1}^{6} x_i f(x_i) = \frac{21}{6} = 3,5$$

$$\sigma^2 = \sum_{i=1}^{6} (x_i - \mu)^2 f(x_i) = 2,916$$

Erweiterung: Werfen 2er Laplace-Würfel, X = Augensumme $X_1 + X_2$

x_i	2	3	4	5	6	7	8	9	10	11	12
$f(x_i) = P(X{=}x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$$\mu = E(X) = E(X_1) + E(X_2) = 2 \cdot 3,5 = 7$$

$$\sigma^2 = \sigma^2(X_1) + \sigma^2(X_2) = 2 \cdot 2,916 = 5,833$$

Wie groß ist die Wahrscheinlichkeit eine Augensumme größer oder gleich 3 und kleiner gleich 6 zu erhalten?

$$P(3 \le X \le 6) = \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} = \frac{14}{36} = 38,89\%$$

Annäherung durch Normalverteilung:

$$P(3 \le X \le 6) = \phi\left(\frac{6-7+0,5}{\sqrt{5,833}}\right) - \phi\left(\frac{3-7-0,5}{\sqrt{5,833}}\right) = \phi(-0,21) - \phi(-1,86) = 38,54\%$$

Beispiel:

Mit welcher Wahrscheinlichkeit erhält man beim Werfen von 1000 Würfel
a) eine Augensumme unter 3485
b) einen Mittelwert von höchstens 3,51
c) genau die Augensumme 3500

Lösung: X = Augensumme, mit $X = X_1 + X_2 + \dots . X_{1000}$
$\sigma^2 = 1000*2.916 = 2916$ E(X) = 1000*3.5 = 3500

zu a)

$$P(X < 3485) = P(X \le 3484) = \phi\left(\frac{3484 - 3500 + 0,5}{\sqrt{2916}}\right) = \phi(-0,29) = 0,3859$$

zu b)

$$P(X \le 3510) = \phi\left(\frac{3510 - 3500 + 0,5}{\sqrt{2916}}\right) = \phi(0,194) = 0,5754$$

zu c)

$$P(X = 3500) = \phi\left(\frac{3500 - 3500 + 0,5}{\sqrt{2916}}\right) - \phi\left(\frac{3500 + -3500 - 0,5}{\sqrt{2916}}\right) = 2\phi(0,01) - 1 = 0,008$$

Beispiel:

Wie oft muss man würfeln, damit der Mittelwert der Augensumme mit mehr als 90% W. um nicht mehr als 1% vom Erwartungswert abweicht?
Lösung:

	Einzelwurf	n*Einzelwurf
Mittelwert	$\mu = 3.5$	$\mu = n * 3.5$
Abweichung	$\varepsilon = 3.5*0.01$	$\Delta = n * 3.5 * 0.01$

Ansatz:

$$P((3,5n \cdot 0,99) \le X \le (3,5n \cdot 1,01)) > 0,9$$

$$\phi\left(\frac{0,01n}{\sqrt{n \cdot 2,916}}\right) - \phi\left(\frac{-0,01n}{\sqrt{n \cdot 2,916}}\right) > 0,9$$

$$2 \cdot \phi\left(\frac{0,01n}{\sqrt{n \cdot 2,916}}\right) - 1 > 0,9$$

$$\phi\left(\frac{0,01n}{\sqrt{n \cdot 2,916}}\right) > 0,95$$

Tabelle:

$$\phi(1,65) > 0,95$$

$$\frac{0,01n}{\sqrt{n \cdot 2,916}} > 1,65$$

n > 79388

8.10 Gesetz der großen Zahlen

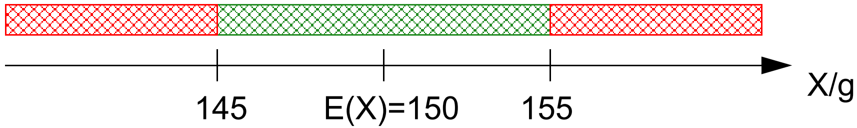
Die rel. Häufigkeit eines Zufallsprozesses nähert sich immer weiter an die theoretische Wahrscheinlichkeit für dieses Ergebnis, je häufiger das Experiment durchgeführt wird.

Anzahl Würfe	davon Kopf		Verhältnis		absoluter Abstand
	theoretisch	beobachtet	theoretisch	beobachtet	
100	50	48	0.5	**0.480**	**2**
1000	500	491	0.5	**0.491**	**9**
10000	5000	4970	0.5	**0.497**	**30**

Der absolute Abstand kann wachsen, obwohl die rel. Häufigkeit gegen 0.5 konvergiert
Die Schlussfolgerung wäre falsch, wonach ein bisher selteneres Ereignis den Rückstand aufholen muss.
Anwendung:
- Versicherungen -> ungefähre Vorhersage des künftigen Schadenverlaufs. Je größer die Zahl der versicherten Personen(Sachwerte, die von der gleichen Gefahr bedroht sind, desto geringer ist der Einfluss des Zufalls. Aber keine Aussage wer betroffen ist.
- Medizin: Wirksamkeitsnachweis, Zufallseinflüsse eliminieren
- Naturwissenschaften: Einfluss von Messfehlern reduzieren

8.11 Einschub: Tschebyscheff Ungleichung

Beobachtet wird das Abfüllgewicht eines Produkts:



Wie groß ist die Wahrscheinlichkeit, dass das Abfüllgewicht nicht in Ordnung ist?
Für die Zufallsvariable X ist keine Wahrscheinlichkeitsverteilung bekannt.

Gesucht: P(|X-150| >= 6) -> keine exakte Lösung oder Abschätzung

Lösung: Ausgehend von der Varianz von X:

$$Var(X) = \sum_i (X_i - E(X))^2 \cdot P(X = x_i) = \sigma^2$$

nur roter Bereich (grüner Bereich einfach weggelassen):
8-11-1

$$\sigma^2 \geq \sum_{i,\, rot} (X_i - E(X))^2 \cdot P(X = x_i)$$

mit

$$|X - 150| \geq 6$$

bzw. quadriert, damit keine Betragsbetrachtung mehr

$$(X-150)^2 \geq 6^2$$

und mit einer allg. Schranke c:

$$(X-150)^2 \geq c^2$$

bzw.
8-11-2

$$(X_i - E(X))^2 \geq c^2$$

Gl. 8-11-2 in Gl. 8-11-1 eingesetzt:

$$\sigma^2 \geq \sum_{i,\,rot} c^2 \cdot P(X = x_i) \;=\; c^2 \cdot \underbrace{\sum_{i,\,rot} P(X = x_i)}_{P(|X - E(X)| >= c)}$$

auf der rechten Seite stehen genau die Wahrscheinlichkeiten, die interessieren:

$$P(|X - E(X)| \geq c) \leq \frac{\sigma^2}{c^2}$$

Anwendung auf Gesetz der großen Zahlen: Bei einem Bernoulli Experiment (Erfolg/Misserfolg) werde n-mal unabhängig wiederholt. Bezeichnet p aus [0,1] die Erfolgswahrscheinlichkeit einer einzelnen Durchführung, dann ist die Anzahl X_n der Erfolge binomialverteilt mit Parameter n,p, Erwartungswert E(X)=np, Varianz σ^2 =np(1-p)

8.12 Schwaches Gesetz der großen Zahlen für rel. Häufigkeiten

Die rel. Häufigkeit R_n konvergiert gegen die Wahrscheinlichkeit p für n -> unendl.
D.h. für jedes noch so kleine $\varepsilon > 0$ geht die Wahrscheinlichkeit, dass die rel. Häufigkeit der Erfolge **nicht** im Intervall [(p-ε, p+ε) liegt gegen Null, für n -> unendl.

$$R_n = \frac{1}{n} \cdot X_n$$

$$\sigma_{R_n}^2 = \frac{1}{n^2} \cdot \sigma_{X_n}^2 = \frac{1}{n^2} \cdot np(1-p) = \frac{p(1-p)}{n}$$

Anwendung der Tschebyscheff-Ungleichung, mit p als Erwartungswert für R_n

$$P(|R_n - p| \geq \varepsilon) \leq \frac{\sigma_{R_n}^2}{\varepsilon^2} = \frac{p(1-p)}{n \cdot \varepsilon^2}$$

$$\lim_{n \to \infty} P(|R_n - p| \geq \varepsilon) = 0$$

Folgerung:

$$E(R_n) = p$$

Wiki Links: Schwaches bzw. starkes Gesetz der großen Zahlen.

9 Prüf und Testverteilungen

9.1 Chi-Quadrat Verteilung

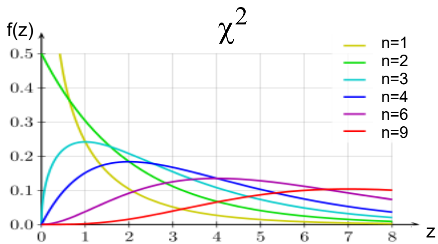
Sei X_1, X_2, ... X_n n stochastisch unabh. Zufallsvariable, die alle der Standardnormalverteilung genügen, mit $\mu = 0$, $\sigma^2 = 1$

$Z = \chi^2 = X_1^2 + X_2^2 + \ldots + X_n^2$ folgt dann einer Verteilung mit der Dichtefunktion

9-1-1

$$f(z) = \begin{cases} A_n \cdot z^{\left(\frac{n-2}{2}\right)} e^{-\frac{z}{2}} & \text{für } z > 0 \\ 0 & \text{sonst} \end{cases}$$

Bestimmung von A_n durch Normierung von f(z)



Der Parameter n bestimmt die Anzahl der sog. Freiheitsgrade f der Verteilung (f = n).

Hinweis: Kann eine der n-Zufallsvariablen X_i durch die restlichen (n-1) Variablen bestimmt werden, reduziert sich der Freiheitsgrad auf f = (n-1)

Normierung, Bstimmung von A_n
9-1-2

$$\int_0^\infty f(z)dz = A_n \cdot \underbrace{\int_0^\infty z^{\left(\frac{n-2}{2}\right)} \cdot e^{-\frac{z}{2}}}_{\text{sog. Gammafunktion } \Gamma} = 1$$

Gammafunktion:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \cdot e^{-t} dt$$

$$A_n = \frac{1}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)}$$

Verteilungsfunktion der Chi-Quadrat-Verteilung
9-1-3

$$F(Z) = A_n \cdot \int_0^z u^{\left(\frac{n-2}{2}\right)} \cdot e^{-\frac{u}{2}} du$$

Mittelwert: $\mu = n$, Varianz: $\sigma^2 = 2n$

9.2 t-Verteilung von Student

(Student ist ein Pseudonym von W.S. Gosset)

Die Verteilung dient als Grundlage für statistische Schätzmethoden und Verteilungstests (Abschn. 12 und Abschn. 13).

Notwendigkeit:
Ist die zur Standardisierung des Mittelwertes einer Stichprobe die benötigte Varianz nicht bekannt, ist der Stichprobenmittelwert nicht mehr normalverteilt (v.a.für rel. kleinen Stichprobenumfang n), da über die Schätzfunktion der Varianz eine zus. Streuung (Unsicherheit) hinzukommt (vgl. Abschn. 12.8).

Die t-Verteilung erlaubt die Berechnung der Verteilung der Differenz vom Mittelwert der Stichprobe zum wahren Mittelwert der Grundgesamtheit.
Die t-Werte hängen ab vom Signifikanzniveau und vom Stichprobenumfang n und bestimmen das Vertrauensintervall und damit die Aussagekraft der Schätzung des Mittelwertes.
Die t-Verteilung wird mit wachsendem n schmaler und geht für $n \to \infty$ in die Normalverteilung über.

Unter dem Begriff **Freiheitsgrad** versteht man die Anzahl unabhängiger Beobachtungen (z.B. Stichprobenumfang n). Werden mithilfe einer Stichprobe u unbekannte Parameter geschätzt, reduziert sich der Freiheitsgrad auf n-u.

X,Y seien zwei stochastisch unabh. Zufallsvariable

X: Standardnormalverteilt
Y: Chi-Quadrat verteilt mit Freiheitsgrad n

9-2-1

$$T = \frac{X}{\sqrt{Y/n}}$$

T ist eine stetige Zufallsvariable mit der Dichtefunktion
9-2-2

$$f(t) = A_n \cdot \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}$$

n: Freiheitsgrad der t-Verteilung
A_n: Normalisierungskonstante

$$\int_{-\infty}^{\infty} f(t)dt = A_n \cdot \int_{-\infty}^{\infty} \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}} dt = 1$$

Daraus kann A_n berechnet werden, mit

9-2-3

$$A_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \cdot \Gamma\left(\frac{n}{2}\right)}$$

Verteilungsfunktion der t-Verteilung
9-2-4

$$F(t) = A_n \cdot \int_0^t \frac{1}{\left(1 + \frac{u^2}{n}\right)^{\frac{n+1}{2}}} du$$

Mittelwert: μ = 0 für n >= 2 (f. n=1 existiert kein μ)

Varianz: σ^2 = n/(n-2) für n>= 3 (f. n=1 u. n=2 existiert kein σ^2)

Beispiele

Manuskript zur Vorlesung

Stochastik

Teil 2

Prof. F. Stockmayer

10 Mathematische Statistik

Eine grundlegende Aufgabe der Statistik ist, Kenntnisse und Informationen über Eigenschaften oder Merkmale einer bestimmten Menge von Objekten/Elementen zu gewinnen. Oft können dabei nicht alle Objekte der Grundgesamtheit einbezogen werden. Die Untersuchung muss auf eine Teilmenge mit n-Elementen beschränkt werden.

10.1 Stichprobe vom Umfang n

- Zufällige Entnahme der n-Elemente, voneinander unabhängig, alle Elemente haben die gleiche Chance gezogen zu werden.
- Stichprobe repräsentiert Grundgesamtheit
- Ergebnis einer Stichprobe: X = Interessierendes Merkmal, Zufallsvariable
  $x_i = \{x_1, x_2, x_3, \dots , x_n\}$ Merkmalsausprägungen, Stichprobenwerte
- Quantitatives Merkmal: z.B. Gemessene Laufzeit eines Programms
- Qualitatives Merkmal: Programm funktioniert Ja/Nein

10.2 Häufigkeitsverteilung einer Stichprobe

In einer Stichprobe vom Umfang n sind k (k < n) unterschiedliche Stichprobenwerte enthalten.

10-2-1

$$f(x) = \begin{cases} h_i & \text{für } x = x_i \text{ (i=1,2,3, ..., n)} \\ 0 & \text{sonst} \end{cases}$$

mit der rel. Häufigkeit:

$$h_i = \frac{n_i}{n}$$

Beispiel: Stichprobe mit n = 25
X = Durchmesser einer Gewindeschraube

x_i/mm	4,7	4,8	4,9	5,0	5,1	5,2
n_i	1	3	6	9	4	2
h_i	0,04	0,12	0,24	0,36	0,16	0,008

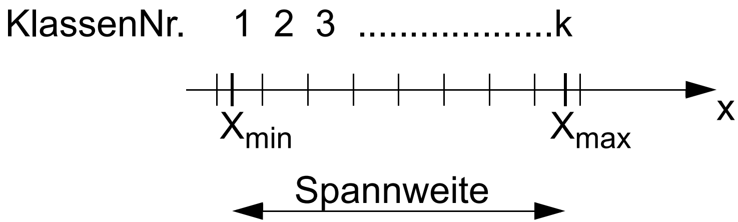
10.3 Verteilungsfunktion einer Stichprobe

10-3-1

$$F(x) = \sum_{x_i \leq x} f(x_i)$$

10.4 Gruppierung der Stichprobe bei großem Umfang

Einteilung in Klassen:



11 Kennwerte oder Maßzahlen einer Stichprobe

11.1 Statistische Kennwerte einer Stichprobe

Mittelwert:
11-1-1

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

Median: Stichprobenwert, der genau in der Mitte liegt, links/rechts befinden sich gleich viele Werte

Modalwert: Stichprobenwert, der am häufigsten vorkommt

Streuungsmaß/Varianz:
11-1-2

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n-1} \cdot \left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$$

oder mit Hilfe der Häufigkeiten, falls f(x) bekannt:
11-1-3

$$s^2 = \frac{n}{n-1} \cdot \sum_{i=1}^{k} (x_i - \bar{x})^2 \cdot f(x_i) = \frac{n}{n-1} \cdot \left[\sum_{i=1}^{k} x_i^2 \cdot f(x_i) - \bar{x}^2\right]$$

Standardabweichung:
11-1-4

$$s = \sqrt{s^2}$$

11.2 Einschub: Woher kommt der Nenner (n-1) in Gl. 11-1-2 ?

1. Ansatz: Die Varianz entspricht der Hälfte der mittleren quadr. Abweichung aller Messwerte voneinander. Für n Messwerte gibt es n^2 Paardifferenzen
11-2-1

$$\Delta_m = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)^2 = 2 \cdot \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 2s^2$$

s. nachfolgende Rechnung

Beispiel: n=2

Paardifferenzen:	$x_1 - x_1$	= 0
	$\mathbf{x_1 - x_2}$	**Da diese Paare doppelt enthalten sind,**
	$\mathbf{x_2 - x_1}$	**führt dieser Ansatz zur doppelten Varianz**
	$x_2 - x_2$	= 0

Die Gewichtung $1/n^2$ im Ansatz berücksichtigt alle Paardifferenzen, auch die, die von vornherein Null ergeben. Davon gibt es genau n-Paare (x_1-x_1, ... x_n-x_n). D.h. die Quadratsumme wird durch einen zu hohen Wert dividiert. Besser wäre es (und damit „Erwartungstreu“), wenn man von n^2 die n nicht relevanten Paare abzieht: $n^2 - n = n(n - 1)$ und den Ansatz etwas modifiziert:

11-2-2

$$\Delta_m^* = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)^2 = 2 \cdot \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 2(s^*)^2$$

s. nachfolgende Rechnung

Bessel Korrektur:

11-2-3

$$\boxed{(s^*)^2 = \frac{n}{n-1} \cdot s^2}$$

und für großes n:

11-2-4

$$\boxed{(s^*)^2 \approx s^2}$$

Rechnung:

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_i - x_j\right)^2$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left((x_i - \bar{x}) + (\bar{x} - x_j)\right)^2$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left((x_i - \bar{x})^2 + 2(x_i - \bar{x})(\bar{x} - x_j) + (\bar{x} - x_j)^2\right)$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x})^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x})(\bar{x} - x_j) + \sum_{i=1}^{n}\sum_{j=1}^{n}(\bar{x} - x_j)^2\right]$$

$$= \frac{1}{n^2}\left[2\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x})^2 + 2\sum_{i=1}^{n}\left((x_i - \bar{x})\sum_{j=1}^{n}(\bar{x} - x_j)\right)\right]$$

$$= \frac{1}{n^2}\left[2n\sum_{i=1}^{n}(x_i - \bar{x})^2 + 2\sum_{i=1}^{n}(x_i - \bar{x})\sum_{j=1}^{n}(\bar{x} - x_j)\right]$$

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

Einschub II: Geometrische Interpretation der Varianz

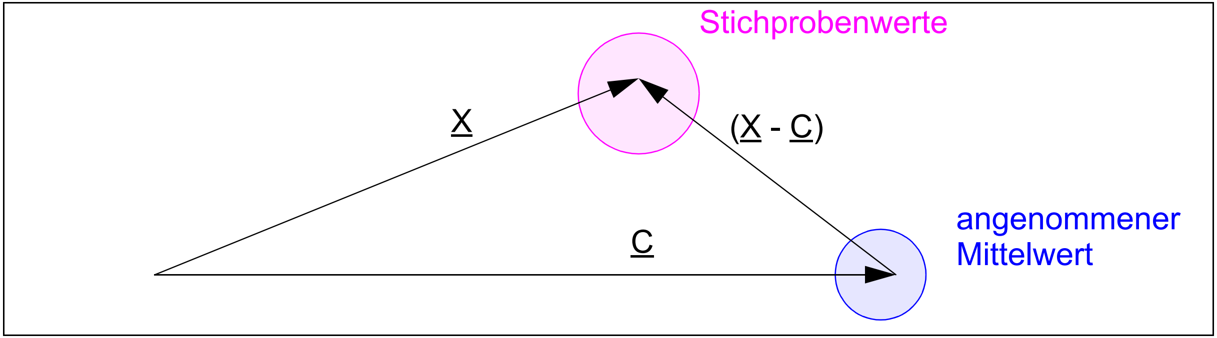
Die einzelnen Werte einer Stichprobe können als Vektor aufgefasst werden:

$$\underline{x} = (x_1 \ldots x_n)^T$$

Der Vektor x̲ besteht aus einem deterministischen Anteil c̲ = (c, ..., c)^T , dessen Komponenten alle identisch sind und aus einem einen Streuungsanteil (x̲ - c̲):

$$\underline{x} = \underline{c} + (\underline{x} - \underline{c})$$

mit c̲ = c * E̲ (E̲ = Einheitsvektor)



Soll der deterministische Anteil c̲ so gewählt werden, dass die Differenz (x̲ - c̲) minimal ist, müssen die beiden Vektoren c̲ und (x̲ - c̲) orthogonal sein. Dann liegt c̲ im Zentrum der Stichprobenwerte.

D.h. das Skalarprodukt $\underline{c} * (\underline{x} - \underline{c}) = 0$, bzw.

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

daraus folgt:

$$c = \bar{x}$$

somit:

$$\underline{x} = \underline{\bar{x}} + (\underline{x} - \underline{\bar{x}})$$

Die drei Vektoren haben jew. die Länge (Beträge):

$$|\underline{x}|^2 = \sum_1^n x_i^2 \qquad |\underline{\bar{x}}|^2 = n\bar{x}^2 \qquad |\underline{x} - \underline{\bar{x}}|^2 = \sum_1^n (x_i - \bar{x})^2$$

Da $\underline{\bar{x}}$ und $(\underline{x} - \underline{\bar{x}})$ orthogonal sind, folgt mit Pythagoras:

$$\sum_1^n x_i^2 = n\bar{x}^2 + \sum_1^n (x_i - \bar{x})^2$$

$$\sum_1^n x_i^2 - n\bar{x}^2 = \sum_1^n (x_i - \bar{x})^2$$

entspricht Quadrat der Länge des Sreuungsanteils

$$\frac{1}{n-1} \cdot \left[\sum_1^n x_i^2 - n\bar{x}^2\right] = \frac{1}{n-1} \cdot \sum_1^n (x_i - \bar{x})^2 = s^2$$

vgl. Gl. 11-1-2

12 Statistische Schätzmethoden

- Die Grundgesamtheit besitzt die Verteilungsfunktion F(x)
- Die Wahrscheinlichkeitsfunktion einer Zufallsvariablen X sei zwar vom Typ her bekannt, meist sind die spezifischen Parameter jedoch unbekannt.
- Bestimmung von Schätz- und Näherungswerte für die unbekannten Parameter der Verteilung mit Hilfe einer konkreten Stichprobe aus der betreffenden Grundgesamtheit.
- Konstruktion sog. Konfidenz- oder Vertrauensintervalle, in denen die Parameter mit einer vorgegebenen Wahrscheinlichkeit (Sicherheit) vermutet werden.

Beispiel: X = Anzahl der Datenpakete, die in einem best. Zeitintervall Δt an einer Schnittstelle beobachtet werden. Die diskrete Zufallsvariable X genügt der Poisson-Verteilung mit dem Mittelwert μ und

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

Intervall i	1	2	3	...	n
Datenpakete $x_i/\Delta t$	x_1	x_2	x_3	...	x_n

Das arithm. Mittel aus der Stichprobe kann als Schätzwert für μ herangezogen werden:

$$\mu = \bar{x} = \frac{1}{n}\sum_1^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

12.1 Schätz- und Stichprobenfunktionen

Spezielle Funktionen zur Schätzung unbekannter Parameter einer W.-verteilung mit Hilfe einer konkreten Stichprobe $x_1, x_2, \ldots, x_n$
Abschn. 12.4 zeigt ein Verfahren zur Gewinnung von Schätzfunktionen.

Schätzung für den Mittelwert μ
12-1-1

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

$$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

Der Schätzwert aufgrund der konkreten Werte x_i einer Stichprobe kann als spezieller Wert bzw. Realisierung der Funktion $\bar{X}$ aufgefasst werden, wobei X_i jew. unabh. Zufallsvariable darstellen, die alle die gleiche Verteilung wie die Zufallsvariable X besitzen (μ, σ^2).
$\bar{X}$ ist damit ebenfalls eine Zufallsvariable und wird als Schätzfunktion/Stichprobenfunktion für den unbekannten Mittelwert μ bezeichnet.

n-dimensionale Zufallsgröße $(X_1; X_2; \ldots ; X_n)$

$\downarrow \quad \downarrow \quad \downarrow$

konkrete Stichprobe $(x_1; x_2; \ldots ; x_n)$

12.2 Eigenschaften der Schätzfunktion $\overline{X}$

Verschiedene Stichproben führen zu verschiedenen Stichprobenmittelwerte $\overline{x}$ bzw. verschiedenen Schätzwerten für den gesuchten Mittelwert der Grundgesamtheit, wobei

- Die Schätzfunktion $\overline{X}$ besitzt den Erwartungswert μ: $E(\overline{X}) = \mu$
- $Var(\overline{X}) = (1/n^2)n\sigma^2 = \sigma^2/n$
  (Transformation mit 1/n und Summe, vgl. Abschn. 8.5 Abschn. 8.8 Abschn. 6.6)

Die Varianz der Schätzfunktion $\overline{X}$ wird mit zunehmendem Stichprobenumfang n kleiner (gegen Null für n gegen unendl.)

Kriterien für eine optimale Schätzfunktion:
Die Schätzung eines unbekannten Parameters ϑ erfolgt mit Hilfe einer geeigneten **Schätzfunktion**:

$$\Theta = g(X_1;X_2 \ldots ;X_n)$$

die für jede konkrete Stichprobe x_1, x_2, ..., x_n einen Schätzwert liefert:

$$\hat{\vartheta} = g(x_1;x_2 \ldots ;x_n)$$

- sie ist erwartungstreu
- sie ist konsistent, d.h. Θ konvergiert gegen ϑ
- sie ist effizient, d.h. es gibt für gleichen Umfang n keine andere erwartungstreue Schätzfunktion mit kleinerer Varianz

Schätzung für die Varianz σ^2

$$\sigma^2 \approx \hat{\sigma}^2 = s^2 = \frac{1}{n-1}\sum_i^n (x_i - \bar{x})^2$$

Die zugehörige erwartungstreue Schätzfunktion ist die Stichprobenfunktion:

12-2-1

$$S^2 = \frac{1}{n-1}\sum_i^n (X_i - \bar{X})^2$$

Schätzung für den Parameter p (Erfolgswahrscheinlichkeit) einer Binomialverteilung (s. Abschn. 12.5):

Die Stichprobe wird als Bernoulli-Experiment n-mal durchgeführt und die Anzahl der Erfolge k festgestellt. Dann ist die rel. Häufigkeit, mit der das Ereihgnis A (Erfolg) eingetreten ist, ein Schätzwert für den unbekannten Parameter p der binomialverteilten Grundgesamtheit.

$$p \approx \hat{p} = h(A) = \frac{k}{n}$$

Die zugehörige Schätzfunktion ist

$$\hat{P} = \frac{X}{n}$$

Die Zufallsvariable X genügt einer Binomialverteilung mit dem Erwartungswert $E(X) = \mu = np$ und der Varianz $\sigma^2 = np(1-p)$. Die Schätzfunktion ist ebenfalls binomialverteilt mit dem Mittelwert p und Varianz p(1-p)/n

12.3 Schätzwerte für Paramter spezieller Verteilungen

Verteilung [13]	Schätzwert für ...	Bemerkungen
Binomialverteilung $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ $(x = 0, 1, \ldots, n)$	Parameter p: $\hat{p} = \frac{k}{n}$	k: Anzahl der „Erfolge" bei einer n-fachen Ausführung des Bernoulli-Experiments
Poisson-Verteilung $f(x) = \frac{\mu^x}{x!} \cdot e^{-\mu}$ $(x = 0, 1, 2, \ldots)$	Mittelwert μ: $\hat{\mu} = \bar{x}$	$\bar{x}$: Mittelwert der Stichprobe
Exponentialverteilung $f(x) = \lambda \cdot e^{-\lambda x}$ $(x \geq 0)$	Parameter λ: $\hat{\lambda} = \frac{1}{\bar{x}}$	$\bar{x}$: Mittelwert der Stichprobe
Gaußsche Normalverteilung $f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ $(-\infty < x < \infty)$	a) Mittelwert μ: $\hat{\mu} = \bar{x}$ b) Varianz σ^2: $\hat{\sigma}^2 = s^2$	$\bar{x}$: Mittelwert der Stichprobe s^2: Varianz der Stichprobe

Quelle: Papula Bd3.

12.4 Maximum Likelihood Methode

- Verfahren zur Gewinnung von Schätzfunktionen

Überlegung:
X ist eine diskrete/stetige Zufallsvariable deren Wahrscheinlichkeits- bzw. Dichtefunktion f(x) noch einen unbekannten Parameter ϑ enthalte, der aus einer Stichprobe mit Umfang n geschätzt werden soll. Die Zufallsvariable X nimmt dabei die einzelnen Stichprobenwerte x_i mit den folgenden Wahrscheinlichkeiten an:

x_i	x_1	x_2	...	x_n
$P(X=x)=f(x_i)$	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

Die Wahrscheinlichkeit eine Stichprobe mit genau den Werten $x_1, x_2, ..., x_n$ zu erhalten, ist wg. der Unabhängigkeit der Werte xi:

12-4-1

$$L = f(x_1)f(x_2) \ldots f(x_n)$$

DieseFunktion hängt von den n Stichprobenwerten ab und von einem unbekannten Parameter ϑ der Wahrscheinlichkeitsfunktion f(x)

Likelihood Funktion L
12-4-2

$$L = L(\vartheta) = f(x_1;\vartheta)f(x_2;\vartheta)\ \dots\ f(x_n;\vartheta)$$

Man bestimmt den Parameter ϑ so, dass die Wahrscheinlichkeit für das Eintreffen der Stichprobenwerte unter Berücksichtigung der Wahrscheinlichkeitsfunktion f(x) und unter dem Einfluss des unbekannten Parameters ϑ maximal wird.
Bedingung für Maximum:

$$\frac{\partial L}{\partial \vartheta} = 0$$

Möglichkeit zur Vereinfachung durch logarithmierte Likelihood Funktion:

$$L^* = \ln L$$

12.5 Anwendungen auf spezielle Wahrscheinlichkeitsverteilungen

1. Binomialverteilung

$$f(x) = \binom{n_0}{x} p^x \cdot (1-p)^{n_0 - x}$$

x sei dabei die Anzahl Erfolge eines Ereignisses mit der (unbekannten) Einzelwahrscheinlichkeit p in einem Experiment vom Umfang n_0. Führt man das Experiment N-mal durch, erhält man N Ergebnisse x_i (i = 1...N). Die Ergebnisse x_i können nun als eine „Stichprobe“ aufgefasst werden, um Gl. 12-4-2 anwenden zu können.

Die Likelihood Funktion lautet gemäß Gl. 12-4-2 dann hier:

$$L(x_i, p) = \prod_{i=1}^{N} \binom{n_0}{x_i} p^{x_i} \cdot (1-p)^{n_0 - x_i}$$

Die L-Funktion kann jedoch stark vereinfacht werden: da später nach dem unbekannten Parameter p abgeleitet und nullgesetzt wird, können die bzgl. p konstanten Binomialkoeffizienten ignoriert werden. Übrig bleibt dann:

$$L(x_i, p) = \prod_{i=1}^{N} p^{x_i} \cdot (1-p)^{n_0 - x_i} = p^{x_1 + \ldots + x_N} \cdot (1-p)^{Nn_0 - (x_1 + \ldots + x_N)}$$

Die rechte Seite der Gleichung kann als eine gemeinsame große Stichprobe vom Umfang $n = Nn_0$ und der Gesamtanzahl der Erfolge $k = (x_1 + x_2 + \ldots + x_N)$ aufgefasst werden. Dann bleibt als L-Funktion übrig:

$$L(p) = p^k \cdot (1-p)^{n-k}$$

Übergang zur logarithmierten Funktion:

$$\ln L(p) = \ln[p^k \cdot (1-p)^{n-k}] = k\ln p + (n-k)\ln(1-p)$$

und partiell nach p differenziert:

$$\frac{\mathrm{d}}{\mathrm{d}p}(k\ln p + (n-k)\ln(1-p)) = k \cdot \frac{1}{p} + (n-k) \cdot \frac{-1}{1-p} = \frac{k}{p} - \frac{n-k}{1-p}$$

Nullsetzen obiger Gleichung liefert das Maximum der L-Funktion:

$$\frac{k}{\hat{p}} - \frac{n-k}{1-\hat{p}} = 0$$

bzw. den Schätzwert für den gesuchten Parameter p:

$$\boxed{\hat{p} = \frac{k}{n}}$$

2. Poisson Verteilung

Aufgabe:
Es soll der Mittelwert μ einer Poisson Verteilung mit Hilfe der Maximum-Likelihood geschätzt werden.

a) Geben Sie eine geeignete L-Funktion an.
b) Leiten Sie daraus eine Formel zur Berechnung des Schätzwertes ab.

3. Gaußsche Normalverteilung

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$L(\mu, \sigma) = f(x_1, \mu, \sigma) \ldots f(x_n, \mu, \sigma)$$

$$L(\mu, \sigma) = \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}}\right] \ldots \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}\right]$$

$$L(\mu, \sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}[(x_1-\mu)^2 + \ldots + (x_n-\mu)^2]}$$

$$L(\mu, \sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right) e^{-\frac{\alpha}{2\sigma^2}}$$

$$L^* = \ln L(\mu, \sigma) = -n\ln(\sigma\sqrt{2\pi}) - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

die benötigten partiellen Ableitungen:

$$\frac{\partial}{\partial\mu}L^* = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)$$

$$\frac{\partial}{\partial\sigma}L^* = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2$$

notwendige Bedingungen für das ges. Maximum:

$$\frac{1}{\hat{\sigma}^2} \cdot \sum_{i=1}^{n}(x_i-\hat{\mu}) = 0 \qquad (1)$$

$$-\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = 0 \qquad (2)$$

aus (1) folgt:

$$\sum_{i=1}^{n}(x_i-\hat{\mu}) = \sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\hat{\mu} = \sum_{i=1}^{n}x_i - n\hat{\mu} = 0$$

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

aus (2) folgt:

$$\boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

die Schätzung für σ^2 liefert jedoch **keine** erwartungstreue Funktion (bzw. nur für große n), vgl. Gl. 11-2-3

12.6 Vertrauens- und Konfidenzintervalle

Überlegung: Kann man ein Intervall $[c_u; c_o]$ angeben, das den unbekannten (bzw. geschätzten) Parameter ϑ ***mit 100%-Sicherheit*** enthält?
Falls ja, wäre die Aussage

$$c_u \le \vartheta \le c_o$$

entweder wahr oder falsch:

$$P(c_u \le \vartheta \le c_o) = 1 \qquad\qquad P(c_u \le \vartheta \le c_o) = 0$$

bzw. eine Aussage mit einer Wahrscheinlichkeit zwischen 0 und 1 nicht möglich.
Einen Ausweg liefert die Bestimmung zweier Zufallsvariablen Θ_u und Θ_o so, dass sie mit einer gewählten Wahrscheinlichkeit γ (sog. statistische Sicherheit oder Vertrauensniveau) den wahren, aber unbekannten Parameterwert einschließen.

$$P(\Theta_u \le \vartheta \le \Theta_o) = \gamma = 1 - \alpha$$

α: Irrtumswahrscheinlichkeit
γ: Vertrauensniveau

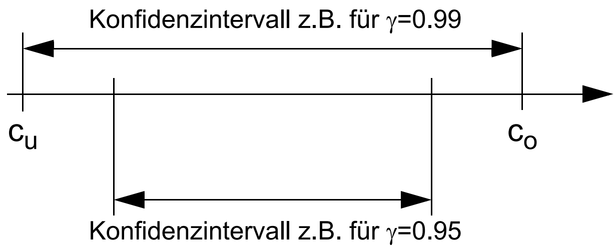
Die Werte für Θ_u und Θ_o müssen sich aus den Stichprobenwerten $x_1, x_2, ..., x_n$ einer konkreten Stichprobe berechnen lassen, unterscheiden sich jedoch bei unterschiedlichen Stichproben. Man kann Θ_u und Θ_o auch als Stichproben- oder Schätzfunktionen der n unabhängigen Zufallsvariablen $X_1, X_2, ... X_n$ des Parameters ϑ verstehen.

$$\Theta_u = g(X_1;X_2;\ldots;X_n) \qquad\qquad \Theta_o = g(X_1;X_2;\ldots;X_n)$$

und durch Einsetzen der Stichprobenwerte:

$$c_u = g(x_1;x_2;\ldots;x_n) \qquad\qquad c_o = g(x_1;x_2;\ldots;x_n)$$

Die Grenzen c_u bzw. c_o werden mit Hilfe einer geeigneten Schätzfunktion gewonnen. Die Grenzen können als Werte einer Zufallsvariablen (Quantile) zu einer vorgegebenen Wahrscheinlichkeit γ aufgefasst werden.



Beispiel:
In einem konkreten Anwendungsfall wählt man z.B. ein Vertrauensniveau von γ=0.95=95%. Entnimmt man der betreffenden Grundgesamtheit 100 voneinander unabhängige Stichproben, so kann man zu 95% sicher sein, dass im Mittel für ca. 95 Stichproben der unbekannte Parameter ϑ in das zugehörige Vertrauensintervall fällt und nur in etwa 5 Fällen außerhalb des Vertrauensintervalls liegt. Die Irrtumswahrscheinlichkeit beträgt dann 5%.

12.7 Vertrauensintervall für den unbekannten Mittelwert μ

- Voraussetzung: normalverteilte Variable X mit <u>bekannter</u> Varianz

Schätzfunktion für den Mittelwert μ, vgl. Gl. 12-1-1 :

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

$\bar{X}$ ist eine normalverteilte Zufallsvariable mit dem Mittelwert $E(\bar{X}) = \mu$ und der Varianz $Var(\bar{X}) = \sigma^2/n$

Die zugehörige standardisierte Zufallsvariable U besitzt den Mittelwert E(U)=0 und Varianz Var(U)=1

$$U = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

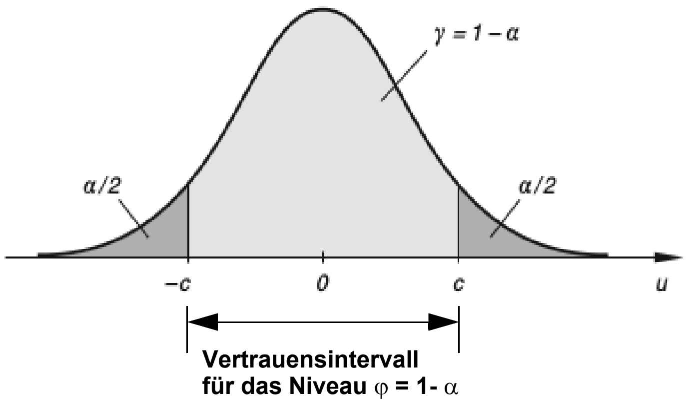
Konstruktion des Vertrauensintervalls:

1. Wahl des Vertrauensniveaus $\gamma = 1 - \alpha$ (z.B. γ=0.95 oder γ=0.99). α ist dabei die Irrtumswahrscheinlichkeit.

2. Für die standardisierte Variable U soll gelten:
12-7-1

$$P(-c \leq U \leq c) = P\left(-c \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq c\right) = \gamma = 1 - \alpha$$

Den zugehörigen Wert für c entnimmt man einer Tabelle



3. Herleitung einer Berechnungsvorschrift für μ:

aus Gl. 12-7-1

$$-c \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le c$$

$$-c \cdot \frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le c \cdot \frac{\sigma}{\sqrt{n}}$$

$$c \cdot \frac{\sigma}{\sqrt{n}} \ge \mu-\bar{X} \ge -c \cdot \frac{\sigma}{\sqrt{n}}$$

$$-c \cdot \frac{\sigma}{\sqrt{n}} \le \mu-\bar{X} \le c \cdot \frac{\sigma}{\sqrt{n}}$$

$$\bar{X} - c \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + c \cdot \frac{\sigma}{\sqrt{n}}$$

Gl. 12-7-1 wird damit zu:

$$P\left(\bar{X} - c \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + c \cdot \frac{\sigma}{\sqrt{n}}\right) = \gamma = 1 - \alpha$$

bzw.

$$P(\Theta_u \le \mu \le \Theta_o) = \gamma = 1 - \alpha$$

Einsetzen des Mittelwerts $\bar{x}$ aus der Stichprobe:

$$\bar{x} - c \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + c \cdot \frac{\sigma}{\sqrt{n}}$$

Untere und obere Grenze des Konfidenzintervalls:

$$\boxed{c_u = \bar{x} - c \cdot \frac{\sigma}{\sqrt{n}}}$$

$$\boxed{c_o = \bar{x} + c \cdot \frac{\sigma}{\sqrt{n}}}$$

Beispiele (1) und (2) ab Seite 523

12.8 Vertrauensintervall für den unbekannten Mittelwert μ

- Voraussetzung: normalverteilte Variable X mit unbekannter Varianz σ^2

Im Gegensatz zum Abschn. 12.7 sind hier die Parameter μ und σ^2 unbekannt und müssen geschätzt werden. Bei der Herleitung eines Vertrauensintervalls für den Mittelwert μ ersetzt man die vorig benutzte standardnormalverteilte Variable U durch die t-verteilte Zufallsgröße.

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Nach Abschn. 9.2 und Gl. 9-2-1 wird die Variable T dort durch die Division zweier Zufallsvariablen X (normalverteilt) und Y (Chi-Quadrat verteilt) gebildet. Dies lässt sich auf die vorliegende Anwendung übertragen: $\bar{X}$ sei eine normalverteilte Zufallsvariable bzw. Schätzfunktion für den Mittelwert und S eine Chi-Quadrat verteilte Schätzfunktion für die Varianz (Quadratsumme) bzw. Standardabweichung mit dem Freiheitsgrad n-1 (wg. Erwartungstreue).

$$P(-c \le T \le c) = \gamma = 1 - \alpha$$

Mit den Werten für die Grenze c aus einer Tabelle der Quantile der t-Verteilung.
Durch die Anwendung der t-Verteilung sind die resultierenden Vertrauensintervalle im Vergleich zur Normalverteilung etwas größer. Die t-Verteilung strebt für einen großen Stichprobenumfang gegen die Standardnormalverteilung, d.h. ab ca. n=30 kann die t-Verteilung auch durch die Standardnormalverteilung ersetzt werden. Beispiele S. 528ff.

12.9 Weitere Vertrauensintervalle

1. Vertrauensintervall für die unbekannte Varianz σ^2 einer Normalverteilung

Um die Gleichmäßigkeit eines Massenproduktes beurteilen zu können, benötigt man neben dem Mittelwert μ der interessierenden Zufallsgröße X noch zusätzliche Kenntnisse über deren Varianz σ^2, die ja als Streuungsmaß Aufschluss über die Größe der Schwankungen gibt, denen das Merkmal X unterworfen ist. Dieser Parameter ist jedoch in der Praxis oft unbekannt und muss daher geschätzt werden. ähnlich wie für den Mittelwert μ lässt sich auch für die Varianz σ^2 einer normalverteilten Grundgesamtheit ein Vertrauens- oder Konfidenzintervall bestimmen. Die Überlegungen gehen wiederum von einer normalverteilten Grundgesamtheit mit dem (bekannten oder unbekannten) Mittelwert μ und der unbekannten Varianz σ^2 aus. Die Aufgabe lautet, aus einer konkreten Zufallsstichprobe $x_1, x_2, \ldots, x_n$ ein Vertrauens- oder Konfidenzintervall für die unbekannte Varianz σ^2 zu bestimmen. Als Schätzfunktion für die Varianz σ^2 verwendet man die bereits bekannte Stichprobenfunktion S^2 und bildet mit ihr die Zufallsvariable Z einer Chi-Quadrat-Verteilung mit f = n-1 Freiheitsgraden.

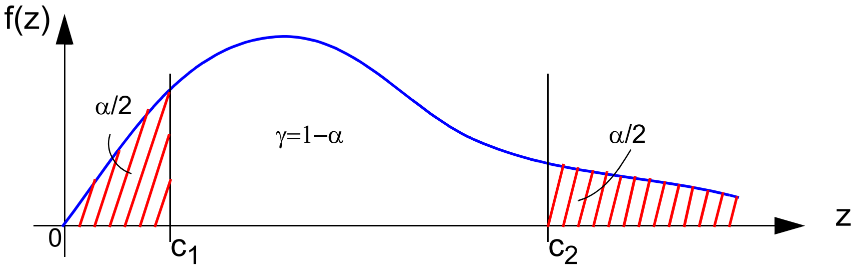
12-9-1

$$Z = (n-1) \cdot \frac{S^2}{\sigma^2}$$

Nach Vorgabe eines gewählten Vertrauensniveaus $\gamma = 1 - \alpha$
lassen sich dann zwei Schranken c1 und c2 so bestimmen, dass die Zufallsvariable Z mit der gewählten Wahrscheinlichkeit γ Werte aus dem Intervall $[c_1, c_2]$ annimmt:

$$P(c_1 \le Z \le c_2) = \gamma = 1 - \alpha$$

Der Flächenanteil der Irrtumswahrscheinlichkeit α wird dabei gleichmäßig auf beide Seiten der Dichtefunktion f(z) verteilt:



Die Intervallgrenzen c_1 und c_2 werden dann folgendermaßen ermittelt:

$$P(Z \le c_1) = F(c_1) = \frac{\alpha}{2} = \frac{1}{2}(1 - \gamma)$$

$$P(Z \le c_2) = F(c_2) = 1 - \frac{\alpha}{2} = 1 - \frac{1}{2} \cdot (1 - \gamma) = \frac{1}{2}(1 + \gamma)$$

aus

$$c_1 \le Z \le c_2$$

und Gl. 12-9-1

$$c_1 \le (n-1) \cdot \frac{S^2}{\sigma^2} \le c_2$$

folgt nach Umstellung für die gesuchte Varianz und Einsetzen des Schätzwertes s^2:
12-9-2

$$\boxed{(n-1) \cdot \frac{s^2}{c_2} \le \sigma^2 \le (n-1) \cdot \frac{s^2}{c_1}}$$

Beispiel: s.S. 532

Einschub: Bedeutung Chi-Quadrat-Verteilung für die Varianz

Die erwartungstreue Schätzfunktion für die Varianz einer Stichprobe vom Umfang n lautet nach Gl. 12-2-1 :

$$S^2 = \frac{1}{n-1}\sum_i^n (X_i - \bar{X})^2 \qquad \text{Gl. I}$$

Sind die Zufallsvariablen X_i normalverteilt, lassen sie sich mit

$$Z_i = \frac{X_i - \bar{X}}{\sigma} \qquad \text{Gl. II}$$

in eine Standardnormalverteilung transformieren (s. Abschn. 7.8).
Setzt man Gl. II in Gl. I ein

$$S^2 = \frac{1}{n-1}\sum_i^n (\sigma \cdot Z_i)^2 \qquad \text{Gl. III}$$

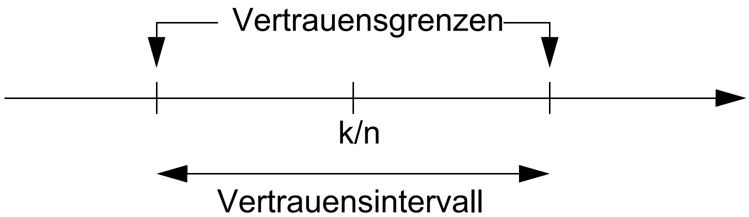
und zieht die Varianz σ^2 vor die Summe, so erhält man nach Umstellung von Gl. III:

$$Z = \frac{(n-1)\cdot S^2}{\sigma^2} = \sum_i^n Z_i^2 \sim \chi^2(n-1)$$

Die Quadratsumme der Zufallsvariablen Z_i folgt jedoch der Chi-Quadrat-Verteilung mit dem Freiheitsgrad (n-1), wobei der reduzierte Freiheitsgrad (n-1) durch die Korrektur der Schätzfunktion zu einer erwartungstreuen Varianz begründet ist (vgl. Abschn. 11.2).

2. Vertrauensintervall für einen unbekannten Parameter p der Binomialverteilung

Auch für den unbekannten Parameter p einer binomialverteilten Grundgesamtheit lässt sich auf dem vorgegebenen Vertrauensniveau $\gamma = 1 - \alpha$ aus einer konkreten Stichprobe ein Vertrauens- oder Konfidenzintervall bestimmen, das mit einem Vertrauen von γ*100 % den wahren Wert des Parameters p enthält.
p sei die unbekannte Wahrscheinlichkeit dafür, dass bei einem Bernoulli-Experiment das Ereignis A eintrete. Die Wahrscheinlichkeit für das Eintreten des komplementären Ereignisses $\overline{A}$ ist dann q = 1 - p.
Eine Stichprobe besteht in einer n-fachen Ausführung des Bernoulli-Experiments, wobei die Anzahl k der dabei erzielten „ Erfolge “ festgestellt wird (das Eintreten des Ereignisses A sei ein „ Erfolg “, das Eintreten des komplementären Ereignisses $\overline{A}$ ein „ Misserfolg“).
Dann ist p‘ = k/n ein Schätz- oder Näherungswert für die unbekannte Erfolgswahrscheinlichkeit (den Anteilswert) p. Für diesen Parameter lässt sich nun unter gewissen Voraussetzungen auch ein Vertrauens- oder Konfidenzintervall bestimmen, dessen Grenzen symmetrisch zum Schätzwert p‘ = k/n angeordnet sind.



Die Zufallsvariable X = „Anzahl der Erfolge bei einer n-fachen Ausführung des Bernoulli-Experiments“ sei binomialverteilt mit den Parametern n und p. Für umfangreiche Stichproben folgt aus dem Grenzwertsatz vom Moivre und Laplace, dass diese Zufallsvariable als nahezu normalverteilt betrachtet werden kann mit dem Mittelwert bzw. Erwartungswert $E(X) = \mu = np$ und der Varianz $Var(X) = \sigma^2 = np(1 - p)$

Für die zugehörige standardisierte Zufallsvariable U gilt:
12-9-3

$$U = \frac{X - \mu}{\sigma} = \frac{X - np}{\sqrt{np(1-p)}}$$

mit dem Schätzwert für den unbekannten Anteilswert p

$$\hat{P} = \frac{X}{n}$$

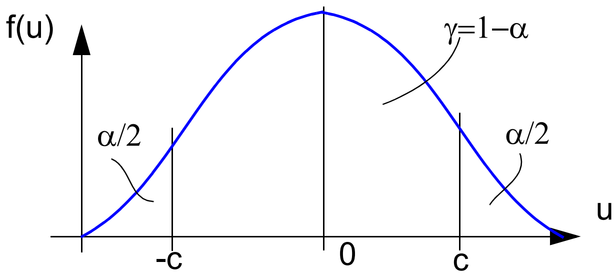
bzw.

$$X = n \cdot \hat{P}$$

eingesetzt in Gl. 12-9-3

$$U = \frac{X - np}{\sqrt{np(1-p)}} = \frac{n\hat{P} - np}{\sqrt{np(1-p)}}$$

Die Grenzen des gesuchten Vertrauensintervall werden nun so bestimmt, dass die Zufallsvariable U mit der gewählten Wahrscheinlichkeit $\gamma = 1 - \alpha$ Werte aus dem symmetrischen Intervall [-c; c] annimmt.



$$P(-c \leq U \leq c) = \gamma = 1-\alpha$$

$$\hat{P} - \frac{c}{n}\sqrt{np(1-p)} \leq p \leq \hat{P} + \frac{c}{n}\sqrt{np(1-p)}$$

$$-c \leq \frac{n\hat{P} - np}{\sqrt{np(1-p))}} \leq c$$

12-9-4

$$\boxed{\hat{p} - \frac{c}{n}\sqrt{n\hat{p}(1-\hat{p})} \leq p \leq \hat{p} + \frac{c}{n}\sqrt{n\hat{p}(1-\hat{p})}}$$

Beispiel: s.S. 538

12.10 Beispiel aus der Informatik: Bitfehlermessung

Für ein Übertragungssystem ist lt. Datenblatt eine Bitfehlerwahrscheinlichkeit (Bit Error Rate, BER) von BER = 1E-6 spezifiziert. Die Fehler treten zufällig und unabhängig voneinander auf. In einer Simulation des Systems werden 1E7 Bit übertragen und dabei die Anzahl der Bitfehler gemessen.

Wie groß ist die Wahrscheinlichkeit, dass in der Simulation

a) genau 10 Fehler auftreten?

b) weniger als 5 Bitfehler auftreten?

c) mehr als 15 Bitfehler auftreten?

Lösung:

a) Die Wahrscheinlichkeit genau N_e=10 Bitfehler in n=1E7 Bit zu messen, ist durch die Binomialverteilung B(n, p, N_e) gegeben mit p = BER = 1E-6

Aufgrund des großen Wertes für n und der sehr kleinen „Erfolgswahrscheinlichkeit" für einen Bitfehler, bereitet die Auswertung des Binomialkoeffzienten Probleme. Einfacher ist die Verwendung der Poisson-Verteilung (Näherung):

$$P(X = N_e) \approx \frac{(np)^{N_e}}{N_e!} \cdot e^{-np} = \frac{(10)^{10}}{10!} \cdot e^{-10} = 0,1251$$

b)

$$P(X < 5) = P(X \le 4) = \sum_{k=0}^{4} P(k) \approx \sum_{k=0}^{4} \frac{(np)^k}{k!} \cdot e^{-np} = 0,02925$$

c)

$$P(X > 15) = 1 - P(X \le 15) \approx 1 - \sum_{k=0}^{15} \frac{(np)^k}{k!} \cdot e^{-np} = 0,04874$$

12.11 Erweiterung: Schätzung der Bitfehlerwahrscheinlichkeit

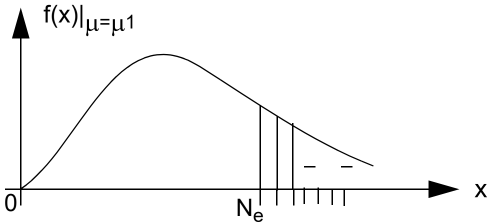
Bei der Bitfehlermessung eines Übertragungssystems werden n=1E7 Bit übertragen und N_e=10 Bitfehler gemessen. Es soll ein Vertrauensintervall für die geschätzte Bitfehlerwahrscheinlichkeit p angegeben werden.
Für die Verteilungsfunktion der Bitfehleranzahl erhält man aus der Poisson-Näherung:

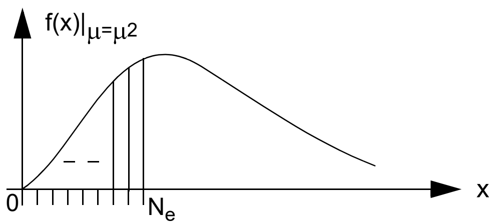
$$P(X \le N_e) \approx \sum_{k=0}^{N_e} \frac{(n\hat{p})^k}{k!} \cdot e^{-n\hat{p}} = \sum_{k=0}^{N_e} \frac{\mu^k}{k!} \cdot e^{-\mu}$$

Ein Vertrauensintervall für die Bitfehlerwahrscheinlichkeit p ergibt sich aus den Lösungen der beiden Verteilungen, d.h. gesucht sind die beiden Parameter μ_1 bzw. μ_2 so dass:

$$P(X \ge N_e) = \sum_{k=N_e}^{n} \frac{(\mu_1)^k}{k!} e^{-\mu_1} = \frac{\alpha}{2} \qquad P(X \le N_e) = \sum_{k=0}^{N_e} \frac{(\mu_2)^k}{k!} e^{-\mu_2} = \frac{\alpha}{2}$$

$$\mu_1 = np_1 < \mu_2 \qquad \mu_2 = np_2 > \mu_1$$





Die u.a. Tabelle zeigt auszugsweise die numerische Lösung der beiden Gleichungen für ein Konfidenzniveau von α = 0,05 (95% Vertrauensintervall) der Poissonverteilung.
Die Bitfehlerwahrscheinlichkeit liegt mit einer W. von 1-α im Intervall [$p_1=\mu_1/n$; $p_2=\mu_2/n$].

Für N_e =10 erhält man aus der Tabelle: μ_1 = 4,795 bzw. μ_2 = 18,39
Und somit: p_1 = 4.795E-7 und p_2 = 1.839E-6

N_e	1-α	
	untere Grenze μ_1	obere Grenze μ_2
1	0.025	5.572
2	0.242	7.255
3	0.619	8.767
4	1.090	10.24
5	1.623	11.67
6	2.202	13.06
7	2.814	14.42
8	3.454	15.76
9	4.115	17.09
10	4.795	18.39

Zahlenbeispiele:

N_e = 10
vermutete Fehlerrate: 1E-6
mit Poisson abgesicherte Fehlerrate:

$$4,8\times10^{-7} \le p \le 1,84\times10^{-6}$$

N_e = 1
vermutete Fehlerrate: 1E-7
mit Poisson abgesicherte Fehlerrate:

$$2,5\times10^{-9} \le p \le 5,57\times10^{-7}$$

Weitere Interpretation:

Mit $\mu_1 = np_1$ bzw. $\mu_2 = np_2$ kann das Intervall auch so geschrieben werden:

$$[np_1 = \mu_1; np_2 = \mu_2]$$

Die tatsächliche Anzahl Bitfehler liegt bei einer gemessenen Bitfehlerzahl N_e mit einer Wahrscheinlichkeit von 95% im Intervall $[\mu_1; \mu_2]$

N_e = 10 Fehler lässt eine Fehlerrate von ^p = N_e/n = 1E-6 vermuten. Die tatsächlich auftretenden Fehler liegen jedoch im Intervall [4.795; 18.39]. Somit liegt die abgesicherte Fehlerrate im Bereich:

$$\frac{4,795}{10}\hat{p} \le p \le \frac{18,39}{10}\hat{p}$$

Den „Worst-Case“ erhält man somit, indem man die vermutete Fehlerrate mit dem Faktor 1.839 multipliziert.

Aus dieser Überlegung lässt sich ein „Korrekturfaktor“ ableiten:

Gezählte Fehler	Für 95% Konfidenz Multiplikation mit
1	5.572 : 1 = **5.572**
2	7.225 : 2 = **3.6125**
5	11.67 : 5 = **2.334**
10	18.39 : 10 = **1.839**
15	24.74 : 15 = **1.65**
20	30.89 : 20 = **1.54**
25	36.9 : 25 = **1.47**

Anwendung: Bei einer Bitrate von 2,048 MBit/s werden 25 Fehler in der Messzeit von 30s gezählt, was eine Bitfehlerrate von ^p = 25/(2.048x1E6x30s) = 4.07x1E-7 vermuten lässt.

Die abgesicherte Bitfehlerrate beträgt dann mit einer Wahrscheinlichkeit von 95% maximal 4.07x1E-7 X **1.47** = 5.98x1E-7

13 Anpassungs- und Verteilungstests

Anpassungs- oder Verteilungstests dienen als Statistische Prüfverfahren zur Überprüfung einer Hypothese über die Art der (unbekannten) Wahrscheinlichkeitsverteilung. Die Grundgesamtheit mit einer unbekannten Verteilungsfunktion F(x) wird dabei mit einer bekannten Verteilungsfunktion $F_0(x)$ verglichen.
Die sog. Nullhypothese vermutet, dass die Zufallsvariable X einer Wahrscheinlichkeitsfunktion $F_0(x)$ genügt. Der Nullhypothese wird eine alternative Hypothese gegenübergestellt, die besagt, dass $F_0(x)$ nicht die gesuchte Verteilungsfunktion der Zufallsvariablen X ist.

Nullhypothese H_0 $$F(x) = F_0(x)$$

Alternative Hypothese H_1 $$F(x) \neq F_0(x)$$

Grundlage des Tests ist eine Stichprobe vom Umfang n aus der Grundgesamtheit.
Der Test ermöglicht eine Entscheidung, ob man die Nullhypothese akzeptieren kann, oder zugunsten der Alternativhypothese verwerfen muss.

13.1 Chi-Quadrat Test

Die Chi-Quadrat-Verteilung ist für einen Vergleich deshalb geeignet, weil sie die Verteilung einer Summe von quadrierten Zufallsvariablen beschreibt. Der Vergleich von beobachteten Häufigkeiten mit den theoretisch erwarteten Häufigkeiten lässt sich mit Hilfe einer Maßzahl beschreiben, die aus dem Quadrat der Differenzen gebildet wird. Die Quadrierung verhindert, dass sich pos. und neg. Abweichungen gegenseitig kompensieren. Man erhält dadurch eine neue Variable Z als eine Art Streuungsmaß, die der Chi-Quadrat-Verteilung folgt. Für eine gegebene Stichprobe lässt sich die Maßzahl berechnen womit man dann wie gewohnt prüfen kann, ob der Zahlenwert unterhalb einer durch die vorgegebene Irrtumswahrscheinlichkeit α liegt.

Vorgehen:

- Unterteilung der Stichprobe in Klassen und Feststellung der absoluten Klassenhäufigkeiten:
  Die n Stichprobenwerte $x_1, x_2, ..., x_n$ werden in k-Klassen (Intervalle, $I_1, I_2, ...I_k$) $k < n$ unterteilt. Jede Klasse sollte erfahrungsgemäß mindestens 5 Stichprobenwerte enthalten. Dann werden die absoluten Klassenhäufigkeiten $n_1, n_2, ..., n_k$ festgestellt.
- Berechnung der theoretisch erwarteten absoluten Klassenhäufigkeiten:
  Aus der als wahr angenommenen Verteilungsfunktion $F_0(x)$ berechnet man für jede Klasse I_i die zugeh. Klassenwahrscheinlichkeit p_i und daraus die hypothetische (theoretisch erwartete) Anzahl der Stichprobenwerte in I_i.

Chi-Quadrat Test:

Klasse Nr. i	n_i (beobachtet)	p_i (beobachtet)	$n_i^* = n * p_i$ (theoretisch)	$\Delta n_i = n_i - n_i^*$	$\frac{(\Delta n_i)^2}{n_i^*}$
1	n_1	p_1	$n_1^* = n\, p_1$	$n_1 - n_1^*$	$\frac{(n_1 - n_1^*)^2}{n_1^*}$
2	n_2	p_2	$n_2^* = n\, p_2$	$n_2 - n_2^*$	$\frac{(n_2 - n_2^*)^2}{n_2^*}$
.....	.....	.....	.....	.....	.....
k	n_k	p_k	$n_k^* = n\, p_k$	$n_k - n_k^*$	$\frac{(n_k - n_k^*)^2}{n_k^*}$
Σ	n	1	n	0	$\hat{\chi}^2$

- Festlegung eines geeigneten Maßes für die Abweichung der beobachteten und theoretischen Häufigkeiten

13-1-1

$$\hat{\chi}^2 = \sum_{i=1}^{k} \frac{(\Delta n_i)^2}{n_i^*} = \sum_{i=1}^{k} \frac{(n_i - n_i^*)^2}{n_i^*}$$

Diese Maßzahl ist als spezieller Wert der Testvariablen (Prüfgröße) Z mit

$$Z = \chi^2 = \sum_{i=1}^{k} \frac{(N_i - n_i^*)^2}{n_i^*} = \frac{(N_i - n \cdot p_i)^2}{n \cdot p_i}$$

zu verstehen, die aus der Zufallsvariable N_i hervorgeht:
N_i = Beobachtete Anzahl der Stichprobenwerte in der i-ten Klasse.

Die Testvariable $Z = \chi^2$ genügt für großes n (n > 50) näherungsweise einer Chi-Quadrat-Verteilung mit f = k-1 Freiheitsgrade, wenn alle Parameter in der als wahr angenommenen Verteilungsfunktion $F_0(x)$ bekannt sind. Für jeden geschätzten Parameter muss die Anzahl der Freiheitsgrade um 1 reduziert werden.

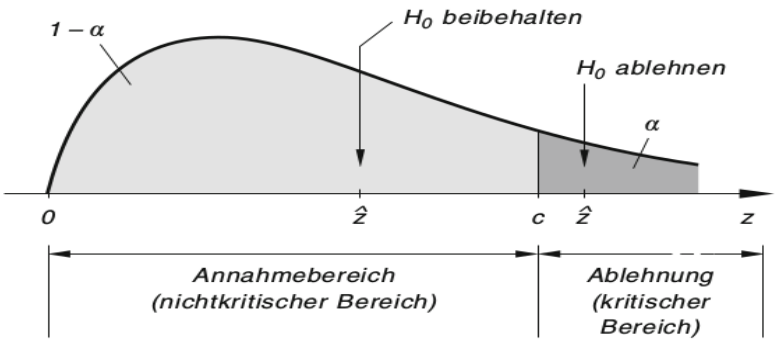
- Wahl einer Signifikanzzahl α und Berechnung der kritischen Grenze c
  Durch Vorgabe der Signifikanzzahl α (z.B. α=0.05 =5%, oder α=0.01=1%) kann die kritische Grenze c so bestimmt werden, dass die Werte der Testvariablen $Z = \chi^2$ mit der Wahrscheinlichkeit $\gamma = 1 - \alpha$ unterhalb der kritischen Grenze liegt.

$$P(Z \leq c)_{H_0} = \gamma = 1 - \alpha$$

- Berechnung des Testwertes und Testentscheidung
  Liegt der aus der Stichprobe berechnete Testwert unterhalb der kritischen Grenze c, d.h.

$$\hat{z} = \hat{\chi}^2 \leq c$$

  wird die Nullhypothese angenommen, ansonsten zugunsten der Alternativhypothese abgelehnt.



1. Beispiel: Fairer Würfel?

Ein Würfel soll mit Hilfe des Chi-Quadrat-Tests überprüft werden, ob die Nullhypothese
H0: Alle 6 möglichen Augenzahlen sind gleichwahrscheinlich
angenommen werden kann

Klasse Nr. i	n_i	p_i	$n_i^* = n * p_i$	$\Delta n_i = n_i - n_i^*$	$\frac{(\Delta n_i)^2}{n_i^*}$
1	15	1/6	20	-5	25/20
2	19	1/6	20	-1	1/20
3	22	1/6	20	2	4/20
4	21	1/6	20	1	1/20
5	17	1/6	20	-3	9/20
6	26	1/6	20	6	36/20
Σ	120	1	120	0	76/20=3.8

Die Irrtumswahrscheinlichkeit soll α = 0.05 betragen.
Die Berechnung der kritischen Grenze c erfolgt dann aus der Bedingung:

$$P(Z \leq c)_{H_0} = P(\chi^2 < c)_{H_0} = 1 - \alpha = 0,95$$

Mit Hilfe der Tabelle und dem Freiheitsgrad f = k - 1 = 5:

$$P(Z \leq c)_{H_0} = F(c) = 0,95 \rightarrow c = z_{(0,95)} = 11,07$$

$$\hat{z} = 3,8 = \hat{\chi}^2 \leq 11,07$$

Ergebnis: Die Nullhypothese kann somit angenommen werden, es handelt sich um einen fairen Würfel mit einer Gleichverteilung der Augenzahlen und somit um ein Laplace-Experiment.

2. Beispiel: Test auf Unabhängigkeit zweier Variablen

Zwei Zufallsvariable X und Y sind unabhängig, wenn die Wahrscheinlichkeitsverteilung einer Variablen die Verteilung der anderen Variable nicht beeinflusst. Der in diesem Abschnitt beschriebene Chi-Quadrat Test kann auf eine zweidimensionale Kontingenztabelle zur Darstellung der beiden Variablen erweitert werden. In diesem Beispiel soll untersucht werden, ob die Rauchgewohnheiten (Heavy, Never, Occasionally, Regularly) mit dem Übungslevel (Frequently, Some, None) zusammenhängen.

Stichprobe f_{ij}:

Smoke	Exercise Level			Σ
	Freq	None	Some	
Heavy	7	1	3	11
Never	87	18	84	189
Occas	12	3	4	19
Regul	9	1	7	17
Σ	115	23	98	236

Erwartung e_{ij}:
Die Variablen „Smoke“ und Exercise Level sind gleichverteilt:

Smoke	Exercise Level			$n_i = \Sigma_i$
	Freq	None	Some	
Heavy	11*115/236= **5,36**	11*23/236= **1,07**	11*98/236= **4,56**	**11**
Never	189*115/236= **92,09**	189*23/236= **18,42**	189*98/236= **78,43**	189
Occas	19*115/236= **9,25**	19*23/236= **1,85**	19*98/236= **7,89**	19
Regul	17*115/236= **8,28**	17*23/236= **1,65**	17*98/236= **7,05**	17
$n_j = \Sigma j$	**115**	23	98	**236**

Die Erwartungswerte berechnen sich mit:

$$e_{ij} = \frac{n_i \cdot n_j}{\sum_i n_i}$$

Die Maßzahl zur Kontingenztabelle berechnet sich mit:

$$\hat{\chi}^2 = \sum_{i,j} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \sim 5,49$$

Sowie dem Freiheitsgrad f anhand der Anzahl Variablenklassen in Spalte (col) und Zeile (row):

$$f = (col - 1)(row - 1) = 2 \cdot 3 = 6$$

Hinweis: Für jeden geschätzten Parameter der Verteilung F0 muss ein Freiheitsgrad abgezogen werden. Hier wird eine Gleichverteilung zugrunde gelegt ohne Parameter.

Nullhypothese H_0: Die Variablen Smoke und Exercise Level sind unabhängig

Bei einem Signifikanzlevel von α=0.05 und dem Freiheitsgrad f = 6 erhält man aus der Tabelle (s. Anhang) das Quantil für die kritische Grenze: 12.59. Der oben berechnete Wert 5.49 ist kleiner und liegt somit im unkritischen Bereich.

Die Nullhypothese kann demnach angenommen werden.

Berechnung mit R (Die Stichprobe bzw. Umfrage ist in der Library MASS enthalten)

```
> library(MASS)
> tbl=table(survey$Smoke, survey$Exer)
> tbl

        Freq None Some
  Heavy    7    1    3
  Never   87   18   84
  Occas   12    3    4
  Regul    9    1    7

> chisq.test(tbl)

        Pearson's Chi-squared test

data:  tbl
X-squared = 5.4885, f = 6, p-value = 0.4828

Warnmeldung:
In chisq.test(tbl) : Chi-Quadrat-Approximation kann inkorrekt sein
```

R produziert noch eine Warnung, da manche Werte in der Kontingenztabelle kleiner 5 sind. Dies lässt sich beheben, wenn die beiden Spalten None und Some zusammengefasst werden.

```
> ctbl = cbind(tbl[,"Freq"], tbl[,"None"] + tbl[,"Some"])
> ctbl

      [,1] [,2]
Heavy    7    4
Never   87  102
Occas   12    7
Regul    9    8
>
>
> chisq.test(ctbl)

        Pearson's Chi-squared test

data:  ctbl
X-squared = 3.2328, f = 3, p-value = 0.3571
```

3. Beispiel: Chi-Quadrat Anpassungstest für eine vermutete Poisson Verteilung

In einer Telefonzentrale wurden in 100 Intervallen (je 5min) die Anruferzahlen beobachtet.

Anzahl Anrufe	0	1	2	3	4	5	Σ
n_i = Gezählte Intervalle mit jew. Anruferzahl	6	9	19	16	32	18	100
	0*6=0	1*9=9	2*19=38	3*16=48	4*32=128	5*18=90	313
n_i^* = 100*f(x): Intervalle mit jew. Anruferzahl	4,4	13,7	21,41	22,34	17,48	10,94	μ= 313/100

Der Mittelwert μ einer angenommenen Poissonverteilung kann anhand der Stichprobe geschätzt werden: μ = 3,13 Anrufer/Intervall

Nullhypothese H_0: Die Anruferzahl folgt einer Poissonverteilung mit μ = 3,13

Der Signifikanzlevel betrage α = 0.05
Die theoretisch erwartete Anruferzahl berechnet sich mit Gl. 7-5-1

$$n_i^* = 100 \cdot f(x) = \frac{\mu^x}{x!} \cdot e^{-\mu}$$

Die Anzahl der Freiheitsgrade beträgt bei k = 6 Klassen:

f = (k - 1) - 1 = 4

Da der Parameter μ geschätzt wurde, muss ein zus. Freiheitsgrad abgezogen werden.

Die Maßzahl $\chi^2_{(0.95)}$ bei 4 Freiheitsgraden beträgt 9.49 (s. Anhang)

Aus der o.a. Tabelle kann die spez. Maßzahl mit Hilfe von Gl. 13-1-1 bestimmt werden:

$$\hat{\chi}^2 = \sum_{i=1}^{k} \frac{(n_i - n_i^*)^2}{n_i^*} = 20,882 > \chi_{krit}^2 = 9,49$$

Entscheidung:

Die Nullhypothese kann hier somit nicht angenommen werden.

Aufgabe:

Eine Stichprobe erfasst die Verspätungen von Zügen. Es wurden insgesamt n=800 Zugverbindungen beobachtet, ob sie pünktlich, oder mit einer Verspätung von 1-15 Minuten bzw. mehr als 15 Minuten unterwegs waren. Zusätzlich wurden die Verbindungen getrennt nach Werktag oder Wochenende notiert.

	Pünktlich	1-15 min	> 15 min	Σ
Mo-Fr	58	46	16	120
Sa/So	32	14	14	60
Σ	90	60	30	180

Zu prüfen ist die Nullhypothese

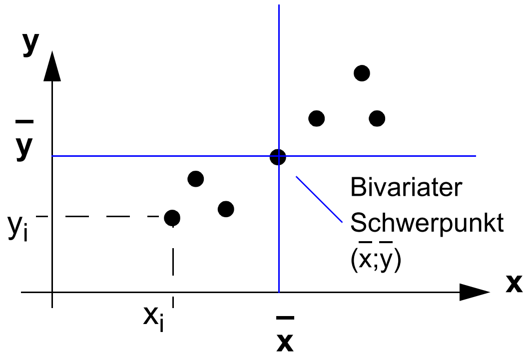
H_0: Die Verspätungen sind unabhängig vom Wochentag

auf dem Signifikanzlevel α = 0.05

14 Korrelation

Erfasst man bei einem Zufallsexperiment gleichzeitig zwei Merkmale X und Y, ist häufig interessant, ob es zwischen den Größen x, y einen **linearen** Zusammenhang (empirische Korrelation) gibt.

Die Stichprobe aus der zweidimensionalen Grundgesamtheit besteht nun aus n geordneten Wertepaaren: $(x_1;y_1)$, $(x_2;y_2)$, ..., $(x_n;y_n)$, die sich in einem kartesischen Koordinatensystem als Punkte bzw. Punktwolke des Streudiagramms darstellen lassen.



Kennwerte:

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

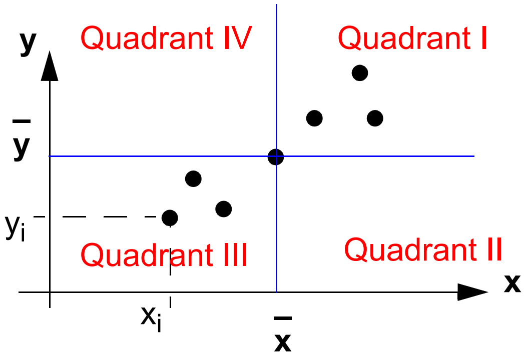
$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2$$

14.1 Kovarianz

Setzt man den bivariaten Schwerpunkt (x,y) in den Ursprung eines neuen Koordinatensystems, entsteht ein 4-Quadranten Schema:



Das Produkt (bzw. die Fläche):

$$(x_i - \bar{x}) \cdot (y_i - \bar{y})$$

kann als Maß für den Abstand eines Punktes (x_i, y_i) vom bivariaten Schwerpunkt interpretiert werden.
Das Produkt liefert in Quadrant I und III stets positive Werte, in Quadrant II und IV stets negative Werte.

Für die Berechnung einer Stärke des Zusammenhangs zwischen den beiden Merkmalen X und Y kommt es darauf an, wie groß die Summe der positiven Abstände in den Quadranten I und III im Vergleich zur Summe der negativen Abstände in den Quadranten II und IV ist. Bei einer rein zufälligen Verteilung der Punkte über alle Quadranten, steht der positiven Summe der Abstände aus Q-I/III ungefähr die selbe negative Summe aus Q-II/IV gegenüber. Befinden sich die Punkte hauptsächlich auf einer Geraden (z.B. mit pos. Steigung) überwiegt die positive Summe aus Q-I/III. Bildet man die Summe aus allen Abstandsprodukten (Flächen) und teilt sie durch die Anzahl der gebildeten Flächen (n-1) , erhält man die durchschnittliche Abweichung vom bivariaten Schwerpunkt:

14-1-1

$$s_{xy} = cov(x;y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}) \cdot (y_i - \bar{y}) = \frac{1}{n-1}\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$$

Ist die Kovarianz ungleich Null, lässt sich daraus ableiten, ob ein pos. oder neg. Zusammenhang besteht. Ein Nachteil hierbei ist, dass die Größe des Zahlenwertes der Kovarianz mit der Anzahl der Punkte wächst und beliebig große Werte annehmen kann. Deshalb teilt man die Kovarianz durch die Standardabweichungen s_x und s_y (Normierung) und erhält so den Korrelationskoeffizienten:

14-1-2

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

mit $-1 \le r \le 1$

Beschreibt der Zähler in Gl. 14-1-2 ein resultierendes Abstandsmaß aus der Summe pos. und neg. Flächen, kann der Nenner in ähnlicher Weise interpretiert werden. Aus den Abständen in x- und y-Richtung zum Schwerpunkt werden jew. Flächen**quadrate** gebildet

und aufsummiert. Zieht man davon jew. die Wurzel entstehen äquivalente Seitenlängen in x- und y-Richtung der Gesamtfläche. Das Produkt der beiden Standardabweichungen s_x und s_y bildet somit die max. mögliche Fläche aller Abweichungen vom bivariaten Schwerpunkt (alle Flächen werden positiv gezählt, es wird nichts abgezogen).
Zur Kontrolle kann in Gl. 14-1-2 angenommen werden, dass alle Punkte auf einer Geraden mit der Steigung 1 liegen ($x_i = y_i$). Dann sind Zähler und Nenner identisch und $r = 1$.

Interpretation des Korrelationskoefizienten:

$\|r\| < 0.5$	schwacher Zusammenhang
$0.5 < r < 0.8$	mittlerer linearer Zusammenhang
$\|r\| > 0.8$	großer linearer Zusammenhang

14.2 Korrelationskoeffizient einer Grundgesamtheit

Der in Abschn. 14.1 aus einer Stichprobe gefundene empirische Korrelationskoeffizient r kann als Schätzwert für den unbekannten Korrelationskoeffizient ρ der Zufallsvariablen X und Y einer Grundgesamtheit aufgefasst werden.

$$\rho = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{E[(X-\mu_x)\cdot(Y-\mu_y)]}{\sigma_x \cdot \sigma_y} = \frac{E(X\cdot Y)-E(X)\cdot E(Y)}{\sigma_x \cdot \sigma_y}$$

Im Falle einer Unabhängigkeit gilt:

$$E(X\cdot Y) = E(X)\cdot E(Y)$$

und damit $\rho = 0$
Stochastisch unabhängige Zufallsgrößen sind stets unkorreliert. Umgekehrt gilt diese Aussage i.a. nicht, aus $\rho = 0$ kann nicht auf eine stochastische Unabhängigkeit geschlossen werden. Es sei denn, X und Y sind beide normalverteilt.
Ist $\rho = 0$, so bedeutet das, dass zwischen den beiden Variablen X und Y keine lineare Abhängigkeit besteht, eine nichtlineare Bindung wäre aber durchaus möglich.

15 Run Test

Betrachtet man eine Folge von binären Zufallszahlen (z.B. 001011101101111000100110), so kann man die Frage stellen, wie viele sog. „Runs“ (Teilfolgen gleicher Werte) in einer Folge mit einer bestimmten Anzahl ‚0‘en und ‚1‘en normalerweise auftreten, bzw. unter der Annahme einer Zufälligkeit zu erwarten sind. Zu wenige Runs deuten auf eine zu geringe Durchmischung hin, zu viele auf eine zu starke Durchmischung. Weicht die Zahl der tatsächlich gefundenen Runs von der erwarteten Häufigkeit ab, so liegt der Verdacht nahe, dass die Folge nicht zufällig entstanden ist.
Daraus lassen sich sehr vielseitige Anwendungen ableiten, da sich viele Stichprobenwerte dichotomisieren (auf zwei Skalenwerte reduzieren) lassen:

- In einer Folge von Zufallszahlen kann man z.B. jeden Zahlenwert mit dem Mittelwert (oder auch Median) vergleichen, so dass bei (z-z)<0 der Binärwert 0 gesetzt wird, ansonsten 1.
- In einer Zahlenreihe werden hintereinander aufsteigende Zahlenfolgen mit 1 und absteigende Zahlenfolgen mit 0 markiert.
- Oder es lassen sich zwei Stichproben auf Gleichheit ihrer Verteilungen prüfen, indem die Zahlenwerte ihrer Größe nach sortiert werden und dabei z.B. farblich (rot/blau) markiert werden, aus welcher Stichprobe der jew. Wert stammt. Durch die (binäre) Farbgebung können dann wieder Runs identifiziert werden.
- ...

Die eingangs erwähnte Beispielfolge enthält r = 13 Runs bei insgesamt n=23 Zufallszahlen, 11x Null (n_1) und 12 mal Eins (n_2) mit $n = n_1 + n_2$.

00 1 0 111 0 11 0 111 000 1 00 11 0

Um zu entscheiden, ob die gefundene Anzahl der Runs innerhalb des Erwartungsbereiches liegt, muss man die Verteilung der Runs in Abhängigkeit der Anzahl der beiden Binärwerte kennen. Die Verteilung lässt sich aus kombinatorischen Überlegungen herleiten (s.a. Hypergeometrische Verteilung) und es ergibt sich folgender Zusammenhang:

für ungerades r:
k=(r-1)/2

$$p(r) = \frac{\binom{n_1}{k} \cdot \binom{n_2}{k-1} + \binom{n_1 - 1}{k-1} \cdot \binom{n_2 - 1}{k}}{\binom{n_1 + n_2}{n_1}}$$

für gerades r:
k = r/2

$$p(r) = \frac{2 \cdot \binom{n_1 - 1}{k-1} \cdot \binom{n_2 - 1}{k-1}}{\binom{n_1 + n_2}{n_1}}$$

Für größere Zahlenwerte werden die Binomialkoeffizienten jedoch sehr groß und die Berechnung dadurch sehr unpraktisch. In praktischen Anwendungen interessiert ja auch mehr die Verteilung bzw. P(X <= c), also die Aufsummierung vieler Wahrscheinlichkeitswerte (Rechenaufwand).
Zur Vereinfachung kann angenommen werden, dass für n_1, n_2 > 20 die Wahrscheinlichkeiten der Runs durch eine Normalverteilung mit den Parametern μ und σ^2 approximiert werden kann. Wichtig dabei ist, dass n_1 und n_2 annähernd gleich sein müssen.

Herleitung des Erwartungswertes:
Für eine gegebene Zufallsfolge Z := (Z_0, Z_1, Z_2, ..., Z_n) mit Z_i = 0 oder 1 gilt:

$P\{Z_i = 1\} = p = n_1/n$ bzw. $P\{Z_i = 0\} = q = (1 - p) = n_2/n$

Daraus lässt sich eine neue Folge von Zufallsvariablen Y_i bilden, welche den Beginn eines Runs markieren, wobei die Summe aller Y_i die Anzahl der Runs ergeben.

$$Y_i = \begin{cases} 1 & \text{falls bei i ein Run beginnt, i > 0} \\ 0 & \text{sonst} \end{cases}$$

$$Y_0 = 1$$

$$R = Y_1 + Y_2 + \ldots + Y_n = \sum_{i=1}^{n} Y_i$$

$$\{Y_i = 1\} = \{(Z_{i-1}, Z_i) = (0,1)\text{oder } (1,0)\} \qquad (i > 0)$$

Wahrscheinlichkeit für den Start eines Runs

$$P(Y_i = 1) = qp + pq \qquad (i > 0)$$

Erwartungswert für den Start eines Runs = Wahrsch.

$$E(Y_i) = 2pq \qquad (i > 0)$$

Erstes Zeichen ist immer Start eines Runs

$$E(Y_0) = 1$$

$$E(R) = E(Y_0) + E(Y_1) + E(Y_2) + \ldots + E(Y_n)$$

$$\boxed{\mu = E(R) = 1 + 2pqn = 1 + 2 \cdot \frac{n_1 \cdot n_2}{n}}$$

$$\boxed{\sigma^2 = \frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)} = \frac{(\mu-1)(\mu-2)}{n-1}}$$

Die Herleitung der Varianz erfolgt ähnlich zur Varianz der Hypergeometrischen Verteilung, oder s. Literatur (M. Fisz, 1986, Wahrscheinlichkeitsrechnung und Math. Statistik, VEB)

Normalapproximation, Standardisierung:

$$U = \frac{R - \mu}{\sqrt{\sigma^2}}$$

Die Hypothese H_0 kann angenommen werden, wenn auf einem Signifikanzniveau α gilt:

$$P(-c \le U \le c) = 1 - \alpha$$

$$-c \le \frac{R - \mu}{\sqrt{\sigma^2}} \le c$$

d.h. die Anzahl der Runs bewegt sich im erwarteten Intervall:

$$\boxed{(-c \cdot \sqrt{\sigma^2} + \mu) - 0,5 \le R \le (c \cdot \sqrt{\sigma^2} + \mu) + 0,5}$$

ggfs. Stetigkeitskorrektur

Für obiges Beispiel ergeben sich folgende Zahlenwerte:
Erwartungswert: μ = 12,47, Varianz: σ^2 = 5.467, Signifikanzniveau α = 0.05 (c = 1.645)
Intervall für R = [8.6;16.3], bzw. nach Stetigkeitskorrektur: R = [9.1;15.8]
Die beobachtete Anzahl Runs befindet sich mit r = 13 innerhalb des Intervalls.
Mit 95% Wahrscheinlichkeit handelt sich dabei um ein zufällige Folge.

Beispiel:

In zwei Kursen wurde gleichzeitig dieselbe Klausur geschrieben. Der Dozent erwartet zunächst, dass die Ergebnisse aufgrund identischer Voraussetzungen (Vorlesung, Übungen, Aufgaben etc.) in beiden Kursen vergleichbar sind und derselben Verteilung folgen. Die Vermutung soll statistisch untermauert werden und die Hypothese
H0: „Die Klausurergebnisse in beiden Kursen sind ähnlich verteilt“
auf einem Signifikanzniveau von 5% (c = 1,645) überprüft werden.
Die Klausurergebnisse werden sortiert und farblich den beiden Kursen (blau/rot) zugeordnet ($n_1 = 26$, $n_2 = 25$, n = 51):

8 9 13 14 15 16 17 18 19 20 21 23 25 26 27 28 30 31 31 32 33 34 35 35 36 36 36
37 38 38 39 39 40 41 41 42 43 44 44 46 48 48 50 51 51 51 51 54 55 63 65

Beobachtet werden 17 Runs.
Der Erwartungswert beträgt: $\mu_R = 26{,}49$ bei einer Varianz von: $\sigma^2 = 12{,}485$, womit

$$21 \leq R \leq 31$$

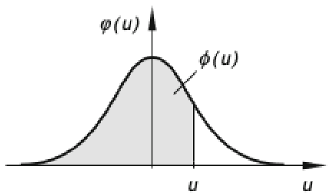
Interpretation: Die beobachtete Anzahl Runs (17) liegt deutlich außerhalb des Bereichs. Die Hypothese kann somit nicht bestätigt werden, die beiden Kurse scheinen ein unterschiedliches Leistungsniveau zu haben.

16 Anhang: Tabellen

- Tabelle 1: Verteilungsfunktion der Standardnormalverteilung
- Tabelle 2: Quantile der Standardnormalverteilung
- Tabelle 3: Quantile der Chi-Quadrat-Verteilung
- Tabelle 4: Quantile der t-Verteilung

Quelle: L. Papula, Mathematik für Ingenieure, Bd. 3, Vieweg

Alternativ in R:	Beispiele
Verteilungsfunktion der Standardnormalverteilung:	pnorm(0.1)
Quantile der Standardnormalverteilung:	qnorm(0.975)
Quantile der Chi-Quadrat-Verteilung:	qchisq(0.95, df=5)
Quantile der t-Verteilung:	qt(0.9,df=99)



Schrittweite: $\Delta u = 0{,}01$

Für *negative* Argumente verwende man die Formel

$$\phi(-u) = 1 - \phi(u) \qquad (u > 0)$$

Für $u \geq 4$ ist $\phi(u) \approx 1$.

u	0	1	2	3	4	5	6	7	8	9
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	0,5639	0,5675	0,5714	0,5754
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	0,7258	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7518	0,7549
0,7	0,7580	0,7612	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	0,7996	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	0,8398
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
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

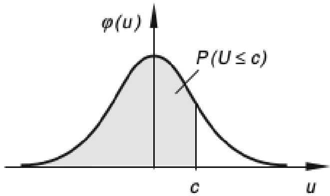
(1) $\phi(1{,}32) = 0{,}9066$

(2) $\phi(1{,}855) = 0{,}9682$ (durch lineare Interpolation)

(3) $\phi(-2{,}36) = 1 - \phi(2{,}36) = 1 - 0{,}9909 = 0{,}0091$

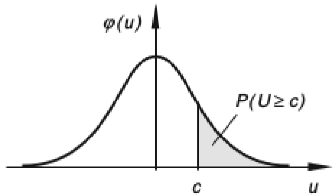
Formeln zur Berechnung von Wahrscheinlichkeiten

(1) *Einseitige* Abgrenzung nach *oben* $(c > 0)$



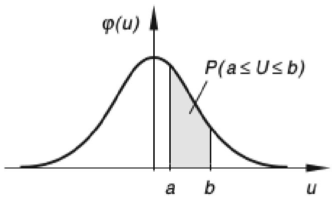
$$P(U \leq c) = \phi(c)$$

(2) *Einseitige* Abgrenzung nach *unten* $(c > 0)$



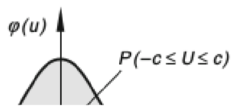
$$P(U \geq c) = 1 - P(U \leq c) = 1 - \phi(c)$$

(3) *Zweiseitige* (unsymmetrische) Abgrenzung $(a < b)$



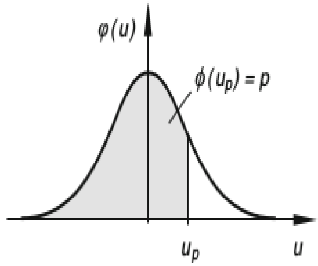
$$P(a \leq U \leq b) = \phi(b) - \phi(a)$$

(4) *Zweiseitige* (symmetrische) Abgrenzung $(c > 0)$



$$P(-c < U < c) = P(|U| < c) =$$

Tabelle 2: Quantile der Standardnormalverteilung



p: Vorgegebene Wahrscheinlichkeit

$(0 < p < 1)$

u_p: Zur Wahrscheinlichkeit p gehöriges Quantil (*obere* Schranke)

Die Tabelle enthält für spezielle Werte von p das jeweils zugehörige Quantil u_p (*einseitige* Abgrenzung nach *oben*).

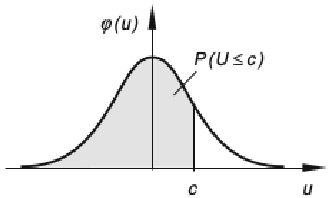
p	u_p	p	u_p
0,90	1,282	0,1	−1,282
0,95	1,645	0,05	−1,645
0,975	1,960	0,025	−1,960
0,99	2,326	0,01	−2,326
0,995	2,576	0,005	−2,576
0,999	3,090	0,001	−3,090

Formeln:

$u_{1-p} = -u_p$

$u_p = -u_{1-p}$

(1) *Einseitige* Abgrenzung nach *oben*

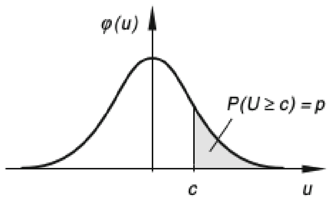


$$P(U \le c) = \phi(c) = p$$
$$\phi(c) = p \rightarrow c = u_p$$

Zahlenbeispiel:

$$P(U \le c) = \phi(c) = 0{,}90 \rightarrow c = u_{0{,}90} = 1{,}282$$

(2) *Einseitige* Abgrenzung nach *unten*

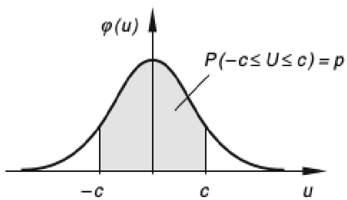


$$P(U \ge c) = 1 - P(U \le c) =$$
$$= 1 - \phi(c) = p$$
$$\phi(c) = 1 - p \rightarrow c = u_{1-p}$$

Zahlenbeispiel:

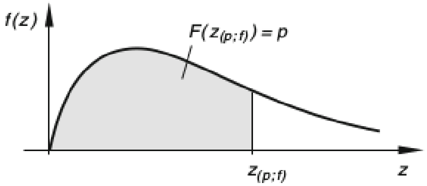
$$P(U \ge c) = 1 - P(U \le c) = 1 - \phi(c) = 0{,}90$$
$$\phi(c) = 1 - 0{,}90 = 0{,}10 \rightarrow c = u_{0{,}10} = -1{,}282$$

(3) *Zweiseitige* (symmetrische) Abgrenzung



$$P(-c \le U \le c) = 2 \cdot \phi(c) - 1 = p$$
$$\phi(c) = \frac{1}{2}(1 + p) \rightarrow c = u_{(1+p)/2}$$

Zahlenbeispiel:



p: Vorgegebene Wahrscheinlichkeit ($0 < p < 1$)

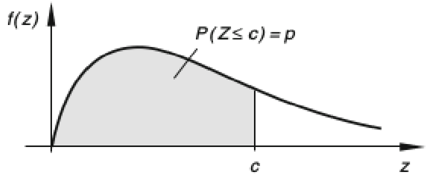
f: Anzahl der Freiheitsgrade

$z_{(p;f)}$: Zur Wahrscheinlichkeit p gehöriges Quantil bei f Freiheitsgraden (*obere* Schranke)

Die Tabelle enthält für spezielle Werte von p das jeweils zugehörige Quantil $z_{(p;f)}$ in Abhängigkeit vom Freiheitsgrad f (*einseitige* Abgrenzung nach *oben*).

	p									
f	0,005	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99	0,995
1	0,000	0,000	0,001	0,004	0,016	2,71	3,84	5,02	6,63	7,88
2	0,01	0,020	0,051	0,103	0,211	4,61	5,99	7,38	9,21	10,60
3	0,07	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,35	12,84
4	0,21	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,41	0,554	0,831	1,15	1,16	9,24	11,07	12,83	15,09	16,75
6	0,68	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	12,02	14,06	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,28	19,67	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	5,70	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
22	8,6	9,5	11,0	12,3	14,0	30,8	33,9	36,8	40,3	42,8
24	9,9	10,9	12,4	13,8	15,7	33,2	36,4	39,4	43,0	45,6
26	11,2	12,2	13,8	15,4	17,3	35,6	38,9	41,9	45,6	48,3
28	12,5	13,6	15,3	16,9	18,9	37,9	41,3	44,5	48,3	51,0
30	13,8	15,0	16,8	18,5	20,6	40,3	43,8	47,0	50,9	53,7
40	20,7	22,2	24,4	26,5	29,1	51,8	55,8	59,3	63,7	66,8
50	28,0	29,7	32,4	34,8	37,7	63,2	67,5	71,4	76,2	79,5
60	35,5	37,5	40,5	43,2	46,5	74,4	79,1	83,3	88,4	92,0
70	43,3	45,4	48,8	51,7	55,3	85,5	90,5	95,0	100,4	104,2
80	51,2	53,5	57,2	60,4	64,3	96,6	101,9	106,6	112,3	116,3

(1) *Einseitige* Abgrenzung nach *oben* $(c > 0)$



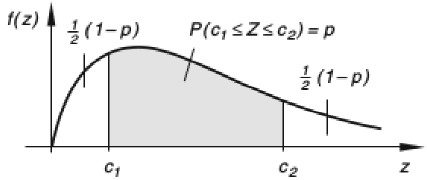
$$P(Z \le c) = F(c) = p$$

$$F(c) = p \rightarrow c = z_{(p;f)}$$

Zahlenbeispiel (bei $f = 10$ Freiheitsgraden):

$$P(Z \le c) = F(c) = 0{,}90 \xrightarrow{f=10} c = z_{(0{,}90;\,10)} = 15{,}99$$

(2) *Zweiseitige* (symmetrische) Abgrenzung $(c_1 < c_2)$



$$P(c_1 \le Z \le c_2) = p$$

Bestimmung der Schranken (Grenzen) c_1 und c_2:

$$P(Z \le c_1) = F(c_1) = \frac{1}{2}(1 - p)$$

$$F(c_1) = \frac{1}{2}(1 - p) \rightarrow c_1 = z_{((1-p)/2;\,f)}$$

$$P(Z \ge c_2) = 1 - P(Z \le c_2) = 1 - F(c_2) = \frac{1}{2}(1 - p)$$

$$F(c_2) = \frac{1}{2}(1 + p) \rightarrow c_2 = z_{((1+p)/2;\,f)}$$

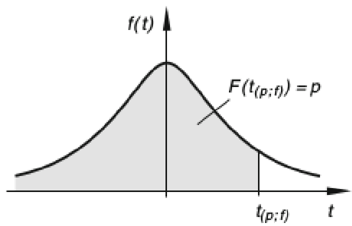
Zahlenbeispiel (bei $f = 10$ Freiheitsgraden):

$$P(c_1 \le Z \le c_2) = 0{,}90$$

$$P(Z \le c_1) = F(c_1) = \frac{1}{2}(1 - 0{,}90) = 0{,}05$$

$$F(c_1) = 0{,}05 \xrightarrow{f=10} c_1 = z_{(0{,}05;\,10)} = 3{,}94$$

$$P(Z \ge c_2) = 1 - P(Z \le c_2) = 1 - F(c_2) = \frac{1}{2}(1 - 0{,}90) = 0{,}05$$



p: Vorgegebene Wahrscheinlichkeit ($0 < p < 1$)

f: Anzahl der Freiheitsgrade

$t_{(p;f)}$: Zur Wahrscheinlichkeit p gehöriges Quantil bei f Freiheitsgraden (*obere* Schranke)

Die Tabelle enthält für spezielle Werte von p das jeweils zugehörige Quantil $t_{(p;f)}$ in Abhängigkeit vom Freiheitsgrad f (*einseitige* Abgrenzung nach *oben*).

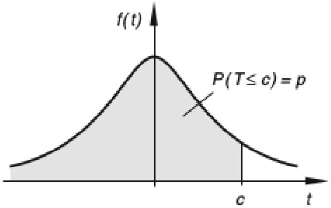
	p				
f	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,707	31,820	63,654
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
22	1,321	1,717	2,074	2,508	2,819
24	1,318	1,711	2,064	2,492	2,797
26	1,315	1,706	2,056	2,479	2,779
28	1,313	1,701	2,048	2,467	2,763
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
50	1,299	1,676	2,009	2,403	2,678
60	1,296	1,671	2,000	2,390	2,660
100	1,290	1,660	1,984	2,364	2,626

Formeln:

$$t_{(1-p;f)} = -t_{(p;f)}$$

$$t_{(p;f)} = -t_{(1-p;f)}$$

(1) *Einseitige* Abgrenzung nach *oben*



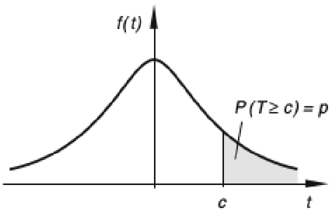
$$P(T \leq c) = F(c) = p$$

$$F(c) = p \rightarrow c = t_{(p;f)}$$

Zahlenbeispiel (bei $f = 10$ Freiheitsgraden):

$$P(T \leq c) = F(c) = 0{,}90 \xrightarrow{f=10} c = t_{(0{,}90;\,10)} = 1{,}372$$

(2) *Einseitige* Abgrenzung nach *unten*



$$P(T \geq c) = 1 - P(T \leq c) =$$

$$= 1 - F(c) = p$$

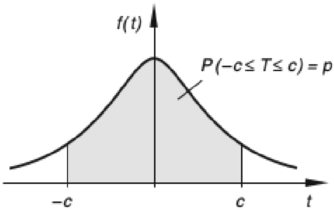
$$F(c) = 1 - p \rightarrow c = t_{(1-p;f)}$$

Zahlenbeispiel (bei $f = 10$ Freiheitsgraden):

$$P(T \geq c) = 1 - P(T \leq c) = 1 - F(c) = 0{,}90$$

$$F(c) = 1 - 0{,}90 = 0{,}10 \xrightarrow{f=10} c = t_{(0{,}10;\,10)} = -t_{(0{,}90;\,10)} = -1{,}372$$

(3) *Zweiseitige* (symmetrische) Abgrenzung



$$P(-c \leq T \leq c) = 2 \cdot F(c) - 1 = p$$

$$F(c) = \frac{1}{2}(1 + p) \rightarrow c = t_{((1+p)/2;f)}$$

Zahlenbeispiel (bei $f = 10$ Freiheitsgraden):