# Skriptum zur Vorlesung Wahrscheinlichkeitslehre und Statistik

**Gilbert Helmberg / Peter Wagner**

WS 2002/03
Letzte Änderung: 2. 2. 2010

http://techmath.uibk.ac.at/wagner/lehre/

**Institut für Technische Mathematik, Geometrie und Bauinformatik**
**Baufakultät, Universität Innsbruck**

# Inhaltsverzeichnis

# Kapitel 1

# Wahrscheinlichkeitsrechnung

## § 1 Begriff der Wahrscheinlichkeit

Unterscheide allgemein:

| N(atur) | M(odell) |
| --- | --- |
| Bereich der Wirklichkeit und Messungen darin | Math. Modell und Rechnungen darin |

Mögliche Fehler: a) in N: Messfehler;
b) in M: Rechenfehler;
c) in der Modellwahl, d.h. M beschreibt N schlecht.

**Bsp.:** 1) **Math. A**

N: Hängendes Seil;
Messen der Schnittkraft,
und des Durchhangs
in verschiedenen Punkten

$$\text{M: } \begin{cases} y = f(x) \\ \text{Schnittkraft} \parallel \text{Tangente} \end{cases}$$

$$\text{Rechnung: } \begin{cases} Hy'' = \sigma\sqrt{1 + y'^2} \\ y = \dfrac{H}{\sigma} \operatorname{ch}\left(\dfrac{\sigma x}{H}\right) \end{cases}$$

2) **Wahrscheinlichkeitstheorie**

N: Ein Würfel
Experiment: Würfeln

$$\text{M: } \begin{cases} \Omega = \{1, 2, \cdots, 6\} \\ p_1 = p_2 = \cdots = p_6 = \dfrac{1}{6} \end{cases}$$

**Unterschied 1)/2):** In den Vorlesungen zu Mathematik A, B und Wissenschaftliches Rechnen werden „determinierte“ Systeme betrachtet: Die **Ergebnisse** sind vorhersagbar und berechenbar. In Wahrscheinlichkeitstheorie und Statistik wird nur die **Wahrscheinlichkeit von Ergebnissen** berechnet.

Ein wiederholbares Experiment sei gegeben. Wir bilden ein mathematisches Modell dazu.

| | **Definition** | **Bsp.: Würfeln mit einem Würfel** |
|---|---|---|
| 1) | Jedes mögliche Ergebnis des Experiments heißt **Elementarereignis**. | 5 ist ein Elementarereignis<br>7 ist kein Elementarereignis |
| 2) | Die Menge aller Elementarereignisse heißt **Ereignisraum** $\Omega$. | $\Omega = \{1,2,3,4,5,6\}$ |
| 3) | Ein **Ereignis** ist eine Menge $A$ von Elementarereignissen, d.h. $A \subset \Omega$. | Z.B. $A =$ „Augenzahl gerade“,<br>d.h. $A = \{2,4,6\}$ |
| 4) | $\Omega = \{x_1, x_2, \cdots\}$ heißt **diskreter Wahrscheinlichkeitsraum**, wenn für jedes Elementarereignis $x_i$ eine Zahl $p_i$<br>mit a) $0 \le p_i \le 1$ und<br>b) $\sum_i p_i = 1$<br>gegeben ist.<br>*Experimentelle Bedeutung:*<br>$p_i = \lim_{N\to\infty} \frac{\text{Anzahl der } x_i \text{ bei } N\text{-mal Exp.}}{N}$<br>(Bemerkung: Manchmal ist auch $\Omega = \{x_0, x_1, x_2, \cdots\}$; oft ist $x_i = i$) | $p_1 = p_2 = \cdots = p_6 = \frac{1}{6}$<br>*Experimentelle Bedeutung,* z.B.<br>von $p_5 = \frac{1}{6}$ : Bei oftmaligem Würfeln tritt 5 in $\approx \frac{1}{6}$ der Fälle, d.h. in $\frac{1}{6} \cdot 100\% \approx 16.7\%$ der Fälle auf. |
| 5) | Wenn $A \subset \Omega$ ein Ereignis ist, so heißt $P(A) = \sum_{x_i \in A} p_i$ seine **Wahrscheinlichkeit** (Englisch: Probability).<br>*Experimentelle Bedeutung:* $P(A) =$<br>$= \lim_{N\to\infty} \frac{\text{Anzahl von } A \text{ bei } N\text{-mal Exp.}}{N}$ | Z.B. $P$ („Augenzahl gerade“)<br>$= P\big(\{2,4,6\}\big) = p_2 + p_4 + p_6 = 3 \cdot \frac{1}{6} = \frac{1}{2}$<br>*Experimentelle Bedeutung:*<br>Bei oftmaligem Würfeln ist die Augenzahl gerade in<br>$\approx \frac{1}{2} = \frac{1}{2} \cdot 100\% = 50\%$ der Fälle. |

**Beispiele:** 1) Werfe 3 verschiedene Münzen zugleich (oder 3-mal hintereinander dieselbe Münze) hoch. Berechne $P$ („2-mal Zahl, 1-mal Kopf")!

$$\begin{aligned}
\Omega &= \{\underbrace{\text{KKK}}_{x_1}, \underbrace{\text{KKZ}}_{x_2}, \text{KZK}, \text{KZZ}, \text{ZKK}, \text{ZKZ}, \text{ZZK}, \underbrace{\text{ZZZ}}_{x_8}\} \\
p_1 = p_2 = \cdots &= p_8 = \frac{1}{8} \\
A &= \{\underbrace{\text{KZZ}}_{x_4}, \underbrace{\text{ZKZ}}_{x_6}, \underbrace{\text{ZZK}}_{x_7}\} \\
&\Longrightarrow P(A) = p_4 + p_6 + p_7 = 3 \cdot \frac{1}{8} = \frac{3}{8} \cdot 100\% = 37.5\%
\end{aligned}$$

**Def.:** Ein Wahrscheinlichkeitsraum, in dem alle Elementarereignisse gleich wahrscheinlich sind, heißt **symmetrischer Wahrscheinlichkeitsraum**.

In einem symmetrischen Wahrscheinlichkeitsraum gilt

$$P(A) = \frac{\text{Anzahl der Elementarereignisse in } A}{\text{Anzahl aller Elementarereignisse}},$$

d.h. $$\boxed{P(A) = \frac{\text{Anzahl der „günstigen" Möglichkeiten}}{\text{Anzahl aller Möglichkeiten}}}$$

Mathematische Schreibweise: $P(A) = \dfrac{\#A}{\#\Omega}$, wobei $\#M$ = Anzahl der Elemente der Menge $M$. Z.B. oben: $P(A) = \dfrac{\#A}{\#\Omega} = \dfrac{3}{8}$.

**Vorsicht:** Wenn mit 3 **gleichen** Münzen **zugleich** geworfen wird, so ist

$$\Omega = \{ \underset{\substack{\downarrow \\ \hat{=}\text{„0-mal} \\ \text{Zahl"}}}{0}, \underset{\substack{\downarrow \\ \hat{=}\text{„1-mal} \\ \text{Zahl"}}}{1}, \overset{\substack{A \text{ oben} \\ \uparrow}}{2}, 3\}$$

Die Wahrscheinlichkeiten müssen dieselben wie vorher sein, d.h.
$p_1 = \frac{1}{8},\ p_2 = \frac{3}{8} = p_3,\ p_4 = \frac{1}{8}$. Nun ist also $\Omega$ **nicht** symmetrisch.

2) Würfle mit 2 (verschiedenen) Würfeln. $A =$ „Augensumme $= 8$ ", $B =$ „6 kommt vor", $\not{B} =$ „6 kommt nicht vor". Berechne $P(A), P(B), P(\not{B})$!

$$\Omega = \big\{(1,1), (1,2), \cdots, (6,6)\big\} = \big\{(\underset{\substack{\swarrow \\ \text{Ergebnis} \\ \text{1. Würfel}}}{i}, \underset{\substack{\searrow \\ \text{Ergebnis} \\ \text{2. Würfel}}}{j}) :\ 1 \leq i, j \leq 6\big\}$$

$\Omega$ ist symmetrisch.

$A = \{(2,6),(3,5),(4,4),(5,3),(6,2)\} = \{(i,j) \in \Omega : i+j=8\}$
$B = \{(1,6),(6,1),(2,6),(6,2),\cdots,(5,6),(6,5),(6,6)\}$
$\not{B} = \{(1,1),\cdots\cdots,(5,5)\} = \{(i,j) \in \Omega : 1 \le i,j \le 5\}$

$$P(A) = \frac{\#A}{\#\Omega} = \frac{5}{36} \approx 13.9\%$$

$$P(B) = \frac{\#B}{\#\Omega} = \frac{11}{36} \approx 30.6\%$$

$$P(\not{B}) = \frac{\#\not{B}}{\#\Omega} = \frac{25}{36} \approx 69.4\%$$

**Def.:** Wenn $\not{B}$ = „Gegenteil von $B$“, so heißt $\not{B}$ **komplementäres Ereignis** zu $B$. Es gilt dann $P(\not{B}) = 1 - P(B)$.

(**Beweis:** $\not{B}$ =„Gegenteil $B$“ $\Longrightarrow$ $\not{B} \cap B = \{\}$, $\not{B} \cup B = \Omega \Longrightarrow 1 = P(\Omega) = \sum_{x_i \in \Omega} p_i =$
$= \sum_{x_i \in B} p_i + \sum_{x_i \in \not{B}} p_i = P(B) + P(\not{B})$ $\square$)

**Bemerkung:** Manchmal ist $P(\not{B})$ leichter zu bestimmen als $P(B)$.

3) **Lotto:** 6 Zahlen von $1,2,\cdots,45$ ankreuzen.
$A$ = „1. Rang“= „alle richtig“, $B$ = „genau 2 Richtige“. **Gesucht:** $P(A), P(B)$.
$\Omega = \{(1,2,3,4,5,6),(1,2,3,4,5,7),\cdots\cdots\} =$
$= \{(i,j,k,l,m,n) : 1 \le i < j < k < l < m < n \le 45\}$
$\#\Omega$ =Anzahl der Möglichkeiten, eine Menge von 6 verschiedenen Zahlen aus 45 Zahlen auszuwählen.

I) Es gibt $n!$ Möglichkeiten, $n$ Dinge **anzuordnen**.

**Beweis:** Es gibt: $n$ Möglichkeiten, was als erstes genommen wird,
$(n-1)$ Möglichkeiten, was als zweites genommen wird,
$\ldots$
2 Möglichkeiten, was als vorletztes genommen wird,
1 Möglichkeit, was als letztes genommen wird.

Ergibt $n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1 = n!$ $\square$

**Bsp.:** Ordne die Buchstaben a,b,c an!

$$\text{3 Möglichkeiten, was als erstes genommen wird} \left\{ \underbrace{\begin{matrix} \text{a b c}, & \text{a c b} \\ \text{b a c}, & \text{b c a} \\ \text{c a b}, & \text{c b a} \end{matrix}}_{\text{2 Möglichkeiten, was als zweites genommen wird}} \right. \qquad 6 = 3 \cdot 2 = 3!$$

II) Es gibt $\binom{n}{k}$ Möglichkeiten, eine **Menge** von $k$ Dingen aus $n$ Dingen **auszuwählen**. $(0 \leq k \leq n)$

**Beweis:** Wir wählen nacheinander $k$ Dinge aus.

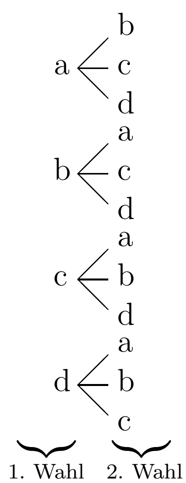
Es gibt: $n$ Möglichkeiten, was als erstes genommen wird,
$n-1$ Möglichkeiten, was als zweites genommen wird,
$\ldots$
$(n-k+1)$ Möglichkeiten, was als $k$-tes genommen wird.

Wenn man also **hintereinander** $k$ Dinge auswählt, gibt es $n(n-1)\cdots(n-k+1)$ Möglichkeiten. Die ausgewählten Dinge sind dann in einer ihrer $k!$ Anordnungen. Also gibt es $\dfrac{n(n-1)\cdots(n-k+1)}{k!}$ mögliche Auswahlmengen.

$$\begin{aligned}\frac{n(n-1)\cdots(n-k+1)}{k!} &= \frac{n(n-1)\cdots(n-k+1)(n-k)\cdots 1}{k!(n-k)\cdots 1} \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k}\end{aligned}$$

$\square$

**Bsp.:** Wähle 2 Buchstaben aus a,b,c,d aus.




$4 \cdot 3 = 12$ Möglichkeiten **hintereinander** zwei Buchstaben auszuwählen.
Weil {a,b}={b,a} etc.
$\Longrightarrow \dfrac{12}{2!} = 6$ d.h.
6 Auswahlmengen sind möglich, nämlich
{a,b},{a,c},{a,d},{b,c},{b,d},{c,d}.

Zurück zum Lotto: $\#\Omega = \binom{45}{6} = \dfrac{45 \cdot 44 \cdot 43 \cdot 42 \cdot 41 \cdot 40}{6!} = 8\,145\,060$

6 Lottozahlen sind richtig, 39 sind falsch.

$A =$„alle richtig“ $\Longrightarrow$ $\#A = 1$ $\Longrightarrow$ $P(A) = \dfrac{1}{8145060} \approx$„1 zu 8 Millionen“.

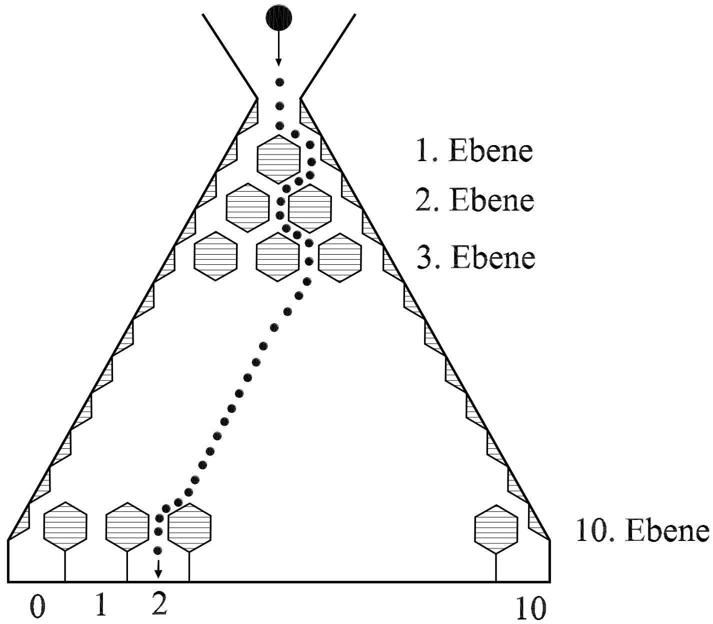
$B =$„2 Richtige“

$$\#B = \text{Anzahl der Möglichkeiten, die zwei Richtigen auszuwählen} \times \text{Anzahl der Möglichkeiten, vier falsche Zahlen auszuwählen}$$
$$= \binom{6}{2} \cdot \binom{39}{4} = \frac{6 \cdot 5}{2!} \cdot \frac{39 \cdot 38 \cdot 37 \cdot 36}{4!} = 1\,233\,765 \implies$$
$$P(B) = \frac{1233765}{8145060} \approx 0.15147 = 15.147\% \approx 1 \text{ zu } 7$$

# § 2 Einige diskrete Wahrscheinlichkeitsräume

## A) Die Binomialverteilung

**Aufgabe:** Eine Münze wird 10-mal hintereinander aufgeworfen. Wie wahrscheinlich ist es, dass genau 5-mal Zahl kommt?
**Lösung:**
Setze $0 =$ Kopf, $1 =$ Zahl
$$\Omega = \Big\{(b_1, \underset{\substack{\uparrow \\ \text{Ergebnis des} \\ \text{2. Aufwurfs}}}{b_2}, \cdots, b_{10}) : b_j \in \{0,1\}\Big\}$$
Z.B. ist $x = (1,0,0,1,0,1,1,1,0,1)$ das Elementarereignis „beim ersten Mal Zahl, beim zweiten Mal Kopf, $\cdots$, beim zehnten Mal Zahl".
$$\#\Omega = \underbrace{2 \cdot 2 \cdot \cdots \cdot 2}_{10-\text{mal}} = 2^{10} = 1024$$
$A =$ „5-mal Zahl" $=$ „5 Einser", d.h. $A = \Big\{(b_1, \cdots, b_{10}) : \sum_{j=1}^{10} b_j = 5\Big\}$
Z.B. für $x$ oben gilt $x \notin A$.
$\#A =$ Anzahl der Möglichkeiten, 5 Stellen (dort wo 1 ist) aus 10 Stellen auszuwählen $=$
$$= \binom{10}{5} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 252 \implies P(A) = \frac{252}{1024} \approx 0.246 = 24.6\%$$
Also in 24.6% der Fälle kommt genau 5-mal Zahl. Wenn also 1000 Menschen je eine Münze 10-mal aufwerfen, werden $\approx 230 - 260$ genau 5-mal Zahl haben, die übrigen öfter oder weniger als 5-mal Zahl haben.
**Bsp. 1:** (Das gerade Galtonbrett) Eine Kugel kann in jeder Ebene links oder rechts gehen. Links=0, rechts=1. Die Kugel landet im Fach $i \iff$ sie geht $i$-mal rechts. (Die strichlierte Bahn z.B. geht 2-mal rechts, 8-mal links und landet in Fach 2). Das ist gleich wahrscheinlich, wie dass eine 10-mal aufgeworfene Münze $i$-mal Zahl zeigt, d.h.
Anteil der Kugeln im Fach $i \approx \binom{10}{i} \cdot 2^{-10}$




Z.B. in Fach 5 $\approx 24.6\%$

in Fach 0 $\approx \underbrace{\binom{10}{0}}_{1} \cdot 2^{-10} = \dfrac{1}{1024} \approx 1{}^\circ/_{\circ\circ}$

**Verallgemeinerung:** Die Münze wird $n$-mal aufgeworfen. Wie wahrscheinlich ist $i$-mal Zahl $(0 \leq i \leq n)$?

Antwort ebenso: $\dfrac{\binom{n}{i}}{2^n} = \binom{n}{i} \cdot 2^{-n}$.

Oft interessiert man sich nur für die **Anzahl** $i$, nicht dafür, **wann** Zahl bzw. Kopf kommt. Man setzt dann

$\Omega = \{0, 1, 2, \cdots, n\}$ mit $p_i =$ Wahrscheinlichkeit von $i = \binom{n}{i} \cdot 2^{-n}$.

($\Omega$ entspricht dem gleichzeitigen Aufwurf von $n$ gleichen Münzen, vgl. § 1, Bsp. 1)

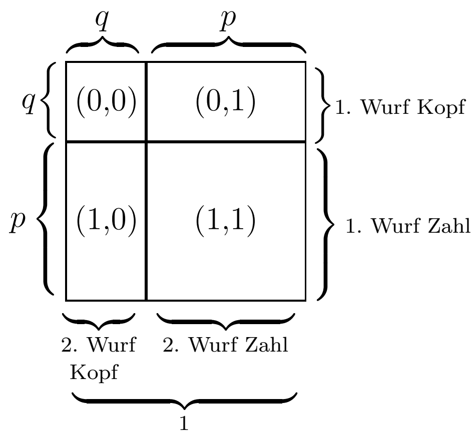
**Weitere Verallgemeinerung:** Die Münze falle bei jedem einzelnen Wurf mit Wahrscheinlichkeit $p$ auf Zahl und mit Wahrscheinlichkeit $q = 1 - p$ auf Kopf. (D.h. bei $N$-maligem Aufwurf fällt die Münze $\approx p \cdot N$-mal auf Zahl, $N \to \infty$. Normalerweise ist $p = q = \frac{1}{2}$.)

**Frage:** Wie wahrscheinlich ist $i$-mal Zahl bei $n$-maligem Aufwurf?

**Zuerst** $n = 2$ :

$\Omega = \big\{(0,0), (0,1), (1,0), (1,1)\big\}$ ist nun **kein** symmetrischer Wahrscheinlichkeitsraum. Wenn z.B. $p > q$, so ist $(1,1)$ (d.h. „Zahl,Zahl“) wahrscheinlicher als $(0,1)$ (d.h. „Kopf,Zahl“). Genauer: $P\big((1,1)\big) = P($„Zahl,Zahl“$) = p \cdot p$, denn im Anteil $p$ der Fälle kommt zuerst Zahl, und davon wieder im Anteil $p$ noch einmal Zahl.

Bild:




Also: $P\big((1,1)\big) = p^2$
$P\big((1,0)\big) = pq = P\big((0,1)\big)$
$P\big((0,0)\big) = q^2$

$\Longrightarrow\ p_0 = q^2,\ p_1 = 2pq,\ p_2 = p^2$

**Allgemein:** $\Omega = \big\{(b_1, \cdots, b_n) : b_j \in \{0,1\}\big\}$,

$A = \Big\{(b_1, \cdots, b_n) : \sum\limits_{j=1}^{n} b_j = i\Big\},\ \#A = \binom{n}{i}$

$(b_1, \cdots, b_n) \in A \implies P\big((b_1, \cdots, b_n)\big) = \underbrace{p \cdots p}_{i\text{-mal}} \cdot \underbrace{q \cdots q}_{(n-i)\text{-mal}} = p^i q^{n-i}$

Daher $p_i = P(A) = P($„$i$-mal Zahl“$) = \binom{n}{i} p^i q^{n-i}$

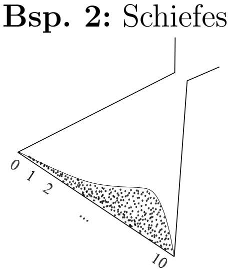
**Def.:** Es sei $0 < p < 1,\ q = 1 - p$.

$\Omega = \{0, 1, \cdots, n\}$ mit $\boxed{p_i = \binom{n}{i} p^i q^{n-i}}$,

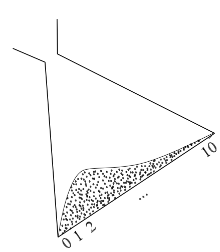
$0 \leq i \leq n$, heißt Wahrscheinlichkeitsraum der **Binomialverteilung** mit den Parametern $n, p$, kurz Bin $(n, p)$.

(Dabei ist für $n = 1:\ p = p_1 = P($„1“$),\ q = p_0 = P($„0“$)$.)
Kontrolle von $\sum p_i = 1$ :
$\sum\limits_{i=0}^{n} p_i = \sum\limits_{i=0}^{n} \binom{n}{i} p^i q^{n-i} = (p+q)^n = 1^n = 1$

**Bsp. 2:** Schiefes Galtonbrett




$p = P($„1“$) = P($„rechts“$) > q$




$p = P($„1“$) < q = P($„0“$)$

**Bsp. 3:** Wir ziehen blind aus einem Topf mit 5 roten und 10 blauen Kugeln 5-mal **mit Zurücklegen** der Kugel zwischen den Zügen. Wie wahrscheinlich ist es, 3-mal rot und 2-mal blau zu ziehen?
Es sei z.B. rot=1, blau=0; $n = 5,\ p = P(„1“) = \frac{1}{3},\ q = P(„0“) = \frac{2}{3}$
$p_3 = P(„\text{3-mal rot}“) = \binom{5}{3} \cdot \left(\frac{1}{3}\right)^3 \cdot \left(\frac{2}{3}\right)^2 \approx 0.16461$

## B) Die hypergeometrische Verteilung

**Bsp. 4:** Wie wahrscheinlich ist es, im vorigen Beispiel 3-mal rot **ohne Zurücklegen** zu ziehen?
Anschaulich: Je mehr rote Kugeln weg sind, umso unwahrscheinlicher wird es, rot zu ziehen (nicht mehr $\frac{1}{3}$). Die Wahrscheinlichkeit wird also etwas abnehmen.
Genau: Anzahl der Zugmöglichkeiten = Anzahl der Möglichkeiten, 5 Dinge aus 15 Dingen auszuwählen = $\binom{15}{5}$.
Anzahl der günstigen:

$$(\text{3 rote aus 5}) \cdot (\text{2 blaue aus 10}) = \binom{5}{3} \cdot \binom{10}{2}$$
$$\Longrightarrow p_3 = \frac{\binom{5}{3} \cdot \binom{10}{2}}{\binom{15}{5}} \approx 0.14985$$

**Allgemein:** $N$ Kugeln, $F$ rote, $N - F$ blaue, $n$ werden gezogen, $i$ sollen rot sein, $n - i$ blau.
Anzahl der Zugmöglichkeiten: $\binom{N}{n}$

Anzahl der günstigen: $\binom{F}{i} \cdot \binom{N-F}{n-i}$

**Def.:** Es gelte $n \leq N,\ n \leq F,\ n \leq N - F$.

$$\Omega = \{0, \cdots, n\} \quad \text{mit} \quad \boxed{p_i = \frac{\binom{F}{i}\binom{N-F}{n-i}}{\binom{N}{n}}},\ 0 \leq i \leq n,$$

heißt Wahrscheinlichkeitsraum der **hypergeometrischen Verteilung** mit den Parametern $n, N, F$, kurz Hyp $(n, N, F)$.

Warum stimmen die Ergebnisse von Bsp. 3 und Bsp. 4 fast überein?

$$\begin{aligned}
p_i^{\text{Hyp}} &= \frac{\binom{F}{i}\binom{N-F}{n-i}}{\binom{N}{n}} = \frac{\frac{F(F-1)\cdots(F-i+1)}{i!}\cdot\frac{(N-F)\cdots\big(N-F-(n-i)+1\big)}{(n-i)!}}{\frac{N(N-1)\cdots(N-n+1)}{n!}} \\
&= \frac{n!}{i!(n-i)!}\cdot\frac{F(F-1)\cdots(F-i+1)\cdot(N-F)\cdots\big(N-F-(n-i)+1\big)}{N(N-1)\cdots(N-i+1)\cdot(N-i)\cdots\underbrace{\big(N-i-(n-i)+1\big)}_{=N-n+1}} \\
&= \binom{n}{i}\cdot\underbrace{\frac{F}{N}\cdot\frac{F-1}{N-1}\cdots\frac{F-i+1}{N-i+1}}_{i\text{ Faktoren}}\cdot\underbrace{\frac{N-F}{N-i}\cdots\frac{N-F-(n-i)+1}{N-i-(n-i)+1}}_{n-i\text{ Faktoren}}
\end{aligned}$$

Falls $F\to\infty$ und $N\to\infty$ mit $\dfrac{F}{N}\to p$, so ist auch $\dfrac{F-1}{N-1}\to p$ etc. und $\dfrac{N-F}{N-i}=1-\dfrac{F-i}{N-i}\to 1-p=q$ etc. (für festes $n$ und $i$), d.h. $p_i^{\text{Hyp}}\to\binom{n}{i}p^iq^{n-i}=p_i^{\text{Bin}}$. Somit:

**Satz:**

$$\boxed{\lim_{\substack{N\to\infty\\ \frac{F}{N}\to p}}\text{Hyp}\,(n,N,F)=\text{Bin}\,(n,p)}\qquad (n,p\text{ fest})$$

Für „großes“ $N$ (bzgl. $n$) ist daher $\text{Hyp}\,(n,N,F)\approx\text{Bin}\,(n,\frac{F}{N})$.

**Bsp. 5:** In einer Box sind 500 Kugeln, 75 farbige, 425 weiße. Nach Schütteln erscheinen 10 im Sichtfenster. Wie wahrscheinlich ist es, dass genau 3 farbige unter diesen 10 sind?

Exakt hypergeometrisch (denn der Vorgang entspricht einem „Ziehen ohne Zurücklegen“):

$$N=500,\ F=75,\ n=10,\ p_3^{\text{Hyp}}=\frac{\binom{75}{3}\binom{425}{7}}{\binom{500}{10}}\approx 0.1298932$$

Angenähert binomial: $p=\dfrac{F}{N}=\dfrac{75}{500}=0.15,\ q=0.85,\ p_3^{\text{Bin}}=\binom{10}{3}\cdot 0.15^3\cdot 0.85^7\approx 0.1298337$

## C) Die Poissonverteilung

**Problem:** Wir werfen eine Münze, die mit Wahrscheinlichkeit $p$ auf Zahl fällt, $n$-mal auf. Dann ist $P$ („$i$-mal Zahl“)$=p_i=\binom{n}{i}p^iq^{n-i}$.
Was passiert, wenn wir doppelt so oft aufwerfen dürfen, aber die Wahrscheinlichkeit auf

Zahl zu fallen, nur halb so groß ist? Dann ist $p_i = \binom{2n}{i}\left(\frac{p}{2}\right)^i\left(1-\frac{p}{2}\right)^{2n-i}$. Was passiert, wenn wir das iterieren, d.h. wenn $n$ vergrößert und $p$ verkleinert wird, so, dass $n\cdot p$ gleich bleibt oder konvergiert?

Somit: Was ist $\lim\limits_{\substack{n\to\infty\\ p\to 0\\ n\cdot p\to\lambda}} p_i^{\text{Bin}}$?

**Lösung:** $\lambda$ und $i$ werden festgehalten, $n\to\infty,\ p\to 0,\ n\cdot p\to\lambda,\ q=1-p \Longrightarrow$

$$\begin{aligned}
\Longrightarrow p_i^{\text{Bin}} &= \binom{n}{i}p^i q^{n-i} = \frac{n(n-1)\cdots(n-i+1)}{i!}\,p^i\cdot\frac{(1-p)^n}{(1-p)^i} = \\
&= \frac{(n\cdot p)\cdot\big((n-1)\cdot p\big)\cdot\dots\cdot\big((n-i+1)\cdot p\big)}{i!}\cdot\frac{\left(1-\frac{n\cdot p}{n}\right)^n}{\big(1-p\big)^i} \\
&\to \frac{\lambda\cdot\lambda\cdots\lambda}{i!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = \frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}
\end{aligned}$$

**Def.:** Es sei $\lambda > 0$.

$\Omega = \{0,1,2,\cdots\}$ mit $\boxed{p_i = \frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}}$, $\quad 0\le i<\infty$,

heißt Wahrscheinlichkeitsraum der **Poissonverteilung** zum Parameter $\lambda$, kurz Poi $(\lambda)$.

Dann gilt

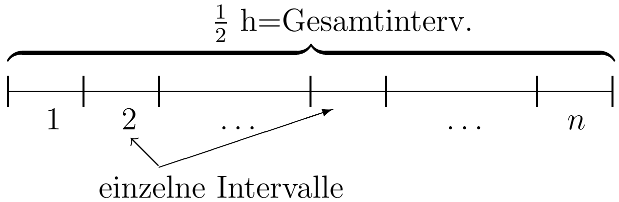
**Satz:** $\boxed{\lim\limits_{\substack{n\to\infty\\ n\cdot p\to\lambda}} \text{Bin}\,(n,p) = \text{Poi}\,(\lambda)}$ $\quad$ ($\lambda$ fest)

Kontrolle von $\sum\limits_{i=0}^{\infty} p_i^{\text{Poi}} = 1 : \sum\limits_{i=0}^{\infty}\frac{\lambda^i}{i!}\mathrm{e}^{-\lambda} = \mathrm{e}^{\lambda}\cdot\mathrm{e}^{-\lambda} = 1$

**Bsp. 6:** In Bsp. 5 ist $\lambda = n\cdot p = 10\cdot 0.15 = 1.5$ und $p_3^{\text{Poi}} = \frac{1.5^3}{3!}\,\mathrm{e}^{-1.5} \approx 0.1255107$.

**Bsp. 7:** In einem Büro treffen im Schnitt pro Stunde fünf Telefongespräche ein (d.h. $\lim\limits_{N\to\infty}\frac{\text{Anrufezahl in } N \text{ h}}{N} = 5$). Wie wahrscheinlich ist es, dass schon in einer halben Stunde genau 5 Anrufe einlangen? Im Schnitt kommen pro $\frac{1}{2}$ h $:\ \lambda = 2.5$ Anrufe, d.h. $\lim\limits_{N\to\infty}\frac{\text{Anzahl der Anrufe in } N\cdot\left(\frac{1}{2}\right) \text{ h}}{N} = 2.5$.

Wir teilen die $\frac{1}{2}$ h (=Gesamtintervall) in $n$ Teile (=Intervalle) ein:




(Z.B. $n = 30$, 1 Teil = 1 min)

Für großes $n$ ist es sehr unwahrscheinlich, dass in einem Intervall zwei oder mehr Anrufe kommen. Wir nehmen außerdem an, dass die Anrufe voneinander „unabhängig" sind, d.h., ob in einem Intervall ein Anruf eintrifft oder nicht, hat keinen Einfluss auf die übrigen Intervalle.
Daher ist die Anzahl der Anrufe in etwa $\mathrm{Bin}\,(n, p)$-verteilt. Dabei ist

$$\begin{aligned} p &= P(\text{„Ein Anruf im Intervall“}) \\ &= \lim_{M\to\infty} \frac{\text{Anzahl der Anrufe in } M \text{ Intervallen}}{M} \qquad \text{(setze } M = nN) \\ &= \lim_{N\to\infty} \frac{\text{Anzahl der Anrufe in } N \cdot \left(\frac{1}{2}\text{ h}\right)}{n \cdot N} = \frac{\lambda}{n} \end{aligned}$$

Je größer $n$, umso besser stimmt unser Modell, d.h. die „wahre“ Verteilung ist $\lim\limits_{n\to\infty} \mathrm{Bin}\left(n, \frac{\lambda}{n}\right) = \mathrm{Poi}\,(\lambda)$.

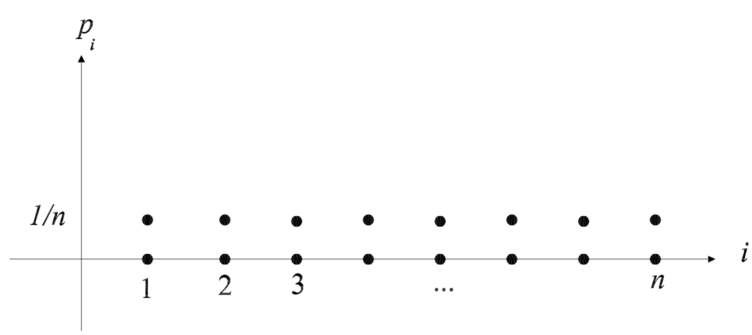
**Daher:** $p_i = P(\text{„}i\text{ Anrufe im Gesamtintervall“}) = \dfrac{\lambda^i}{i!}\,\mathrm{e}^{-\lambda}$, wobei $\lambda =$ durchschnittliche Anzahl der Anrufe im Gesamtintervall.

**Speziell:** $\lambda = 2.5 \Longrightarrow p_5 = \dfrac{2.5^5}{5!}\,\mathrm{e}^{-2.5} \approx 6.7\%$

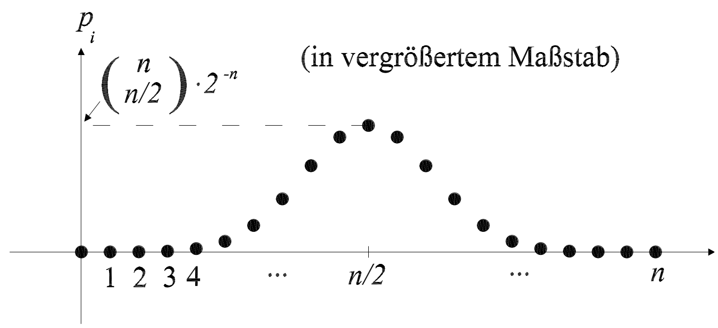
D) **Bildliche Darstellung einiger Wahrscheinlichkeitsräume** $\Omega \subset \mathbb{R}$ (man sagt dazu auch “Verteilung”)

a) $\mathrm{Equ}\,(n): \;\; \Omega = \{1, \cdots, n\},\;\; p_i = \dfrac{1}{n}$

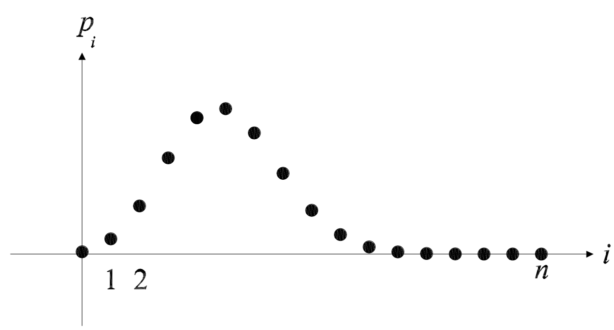
„Gleichverteilung“



b) Bin $\left(n, \frac{1}{2}\right)$, $n$ gerade

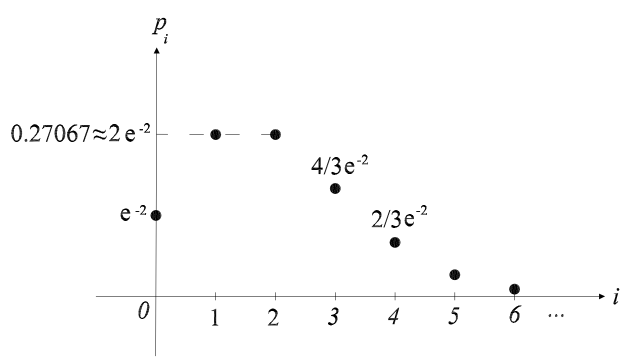



Bin $\left(n, \frac{1}{4}\right)$

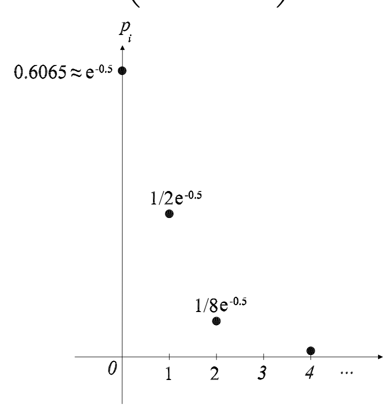



Zum „Glockenkurven“-artigen siehe § 3.

c) Poi $(2)$ $\quad \left(p_i = \dfrac{2^i}{i!}\, \mathrm{e}^{-2}\right)$



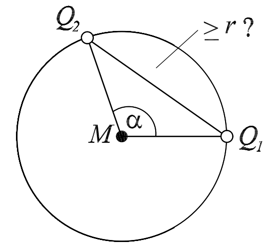

Poi $(0.5)$ $\quad \left(p_i = \dfrac{\mathrm{e}^{-0.5}}{2^i i!}\right)$



# § 3 Kontinuierliche Wahrscheinlichkeitsräume

## A) Die Gleichverteilung

**Aufgabe:** Auf einem Kreis werden „zufällig“ und „unabhängig“ voneinander 2 Punkte $Q_1, Q_2$ gewählt. Wie wahrscheinlich ist es, dass ihr Abstand $\geq$ dem Kreisradius ist?




**Lösung:** Nach Drehen des Kreises können wir uns $Q_1$ fixiert denken. Die Wahl des 2. Punktes $Q_2$ bedeutet, dass man einen Winkel $\alpha \in [0, 2\pi[$ wählt.
Also: $\Omega$ = Menge der Elementarereignisse $= [0, 2\pi[$
Beachte: $\Omega$ ist nun nicht mehr abzählbar wie in §§ 1,2.
Wir betrachten jeden Winkel als gleich wahrscheinlich, d.h.

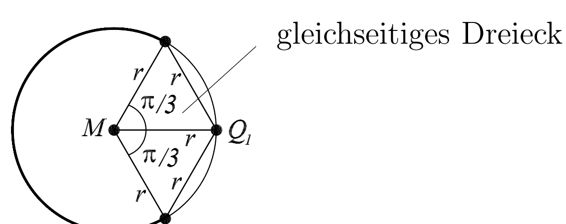
$$P(\alpha \in \text{ Intervall}) \text{ ist proportional zur Intervalllänge.}$$

Ähnlich zum diskreten symmetrischen Wahrscheinlichkeitsraum in § 1 definieren wir

**Def.:** Es seien $a < b \in \mathbb{R}$.
$\Omega = [a, b]$ $\big($ oder $]a, b[,\ [a, b[,\ ]a, b]\big)$ heißt **symmetrischer (kontinuierlicher) Wahrscheinlichkeitsraum** (oder Wahrscheinlichkeitsraum der **Gleichverteilung**), wenn

$$\begin{aligned} P(\text{Intervall}) &= \frac{\text{Intervalllänge}}{\text{Länge von } \Omega}, \text{ d.h.} \\ P\big([c, d]\big) &= \frac{d-c}{b-a}. \end{aligned}$$

**Speziell oben:**




wenn $Q_2$ hier, so ist $\overline{Q_1Q_2} \geq r$

$\overline{Q_1Q_2} \geq r \iff \alpha \in \left[\frac{\pi}{3}, 2\pi - \frac{\pi}{3}\right] = \left[\frac{\pi}{3}, \frac{5\pi}{3}\right]$, d.h. $P(\text{„Abstand} \geq \text{Radius“}) =$

$$= P\left(\left[\underbrace{\frac{\pi}{3}}_{c}, \underbrace{\frac{5\pi}{3}}_{d}\right]\right) = \frac{\overbrace{4\pi/3}^{d-c}}{\underbrace{2\pi}_{b-a}} = \frac{2}{3} = 66,\dot{6}\,\%$$

**Beachte:** Jedes Elementarereignis hat nun Wahrscheinlichkeit 0; denn
$P(x) = P\big([x,x]\big) = \dfrac{x-x}{b-a} = 0.$

A.H. Колмогоров fand 1933 ein mathematisches Modell, um sowohl diskrete als auch kontinuierliche Wahrscheinlichkeitsräume zu behandeln:

| | **Definition** | **Bsp.: Gleichverteilung auf $[0, 2\pi[$** |
|---|---|---|
| 1) | Jedes mögliche Ergebnis des Experiments heißt **Elementarereignis** | 1, $\pi$, $\frac{5}{3}$, $\sqrt{2}$ sind Elementarereignisse. 7 ist kein Elementarereignis. |
| 2) | Die Menge aller Elementarereignisse heißt **Ereignisraum** $\Omega$. | $\Omega = [0, 2\pi[$ |
| 3) | Eine Menge $\mathcal{S}$ von Teilmengen von $\Omega$ sei gegeben mit<br>a) $\Omega \in \mathcal{S}$<br>b) $A \in \mathcal{S} \Longrightarrow \not{A} \in \mathcal{S}$<br>c) $A_k \in \mathcal{S} \Longrightarrow A_1 \cup A_2 \cup \cdots \in \mathcal{S}$.<br>$\mathcal{S}$ heißt **Sigma-Algebra der Ereignisse**. | $\mathcal{S}$ = kleinste Menge von Teilmengen, die a), b), c) erfüllt und alle Intervalle enthält. |
| 4/5) | $\Omega$ heißt **Wahrscheinlichkeitsraum**, wenn für jedes $A \in \mathcal{S}$ eine Zahl $P(A)$ (die **Wahrscheinlichkeit** von $A$) gegeben ist mit a) $0 \leq P(A) \leq 1$<br>b) $P(\Omega) = 1$ | $P(\text{Intervall}) = \dfrac{\text{Intervalllänge}}{2\pi}$<br>$P\big([0,2] \cup [3,4]\big) = \dfrac{3}{2\pi}$ |

c) $A, B$ schließen einander aus $\big($d.h. $A \cap B = \{\}\big) \Longrightarrow P(A \cup B) = P(A) + P(B)$
(und analog für abzählbar viele Ereignisse:
$A_j \cap A_k = \{\}$ für alle $j \neq k \Rightarrow P(A_1 \cup A_2 \cup \cdots) = \sum\limits_{k=1}^{\infty} P(A_k))$.

**Bemerkung:** Für abzählbare $\Omega$ (wie in § 1,2) ist $\mathcal{S}$ die Menge aller Teilmengen von $\Omega$; bei überabzählbaren $\Omega$ muss man gewisse „pathologische" Mengen ausschließen, damit 4c) sinnvoll wird.

**Wesentlich ist:** 1) Statt $p_i$ wie in § 1 (wäre im kontinuierlichen Fall 0) wird nun $P(A)$ als grundlegend betrachtet;
2) wenn $A, B$ einander ausschließen, so ist

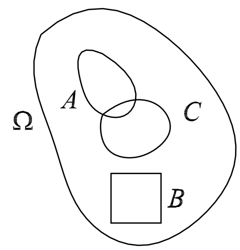
$$P(A \cup B) = P(A) + P(B).$$

**Bsp.** Experiment: Einmal Würfeln, $\Omega = \{1, 2, \cdots, 6\}$, $A$ = „Augenzahl gerade" = $\{2, 4, 6\}$, $B = \{5\}$, $C$ = „Augenzahl durch 3 teilbar" = $\{3, 6\}$.

$A, B$ schließen einander aus, $P(A \cup B) = \dfrac{4}{6} = \dfrac{1}{2} + \dfrac{1}{6} = P(A) + P(B)$

$A, C$ schließen einander nicht aus, $P(A \cup C) = \dfrac{4}{6} < \dfrac{1}{2} + \dfrac{1}{3} = P(A) + P(C)$

Bild:




$$P(A \cup B) = P(A) + P(B)$$
$$P(A \cup C) \leq P(A) + P(C)$$

(Allgemein gilt: $P(A \cup C) = P(A) + P(C) - P(A \cap C)$)

## B) Die Exponentialverteilung

**Aufgabe:** Die Sekretärin erscheint um 8 Uhr im Büro (durchschnittlich $\mu = 5$ Anrufe/Stunde). Wie wahrscheinlich ist es, dass sie mehr als eine halbe Stunde bis zum ersten Anruf warten muss?

**Lösung:**

$P$(„mehr als 1/2 h warten")$=P$(„in 1/2 h 0 Anrufe")$= p_0^{\text{Poi mit } \lambda=2.5} = \mathrm{e}^{-2.5} \approx 8.21\,\%$.

**Allgemein:** Wie wahrscheinlich ist es, dass die Wartezeit bis zum 1. Anruf im Intervall $]t_1, t_2]$ liegt bei durchschnittlich $\mu$ Anrufen/h?

$\left(t_1, t_2 \text{ werden in Stunden gerechnet, } \mu \left[\frac{1}{\mathrm{h}}\right]\right)$

Somit: Elementarereignis = Wartezeit $t$ zum 1. Anruf, $\Omega = ]0, \infty[$.

schließen sich aus

$$\begin{aligned}
P\big(]t_1, \infty[\big) &= P\big(]t_1, t_2] \cup ]t_2, \infty[\big) \\
&= P\big(]t_1, t_2]\big) + P\big(]t_2, \infty[\big) \\
\Longrightarrow P\big(]t_1, t_2]\big) &= P\big(]t_1, \infty[\big) - P\big(]t_2, \infty[\big); \\
P\big(]t_1, \infty[\big) &= P(\text{„mehr als } t_1 \text{ Stunden warten"}) \\
&= P(\text{„in } t_1 \text{ Stunden 0 Anrufe"}) \\
&= p_0^{\text{Poi mit } \lambda=\mu t_1} = \mathrm{e}^{-\mu t_1} \\
\Longrightarrow P\big(]t_1, t_2]\big) &= \mathrm{e}^{-\mu t_1} - \mathrm{e}^{-\mu t_2}
\end{aligned}$$

Z.B. für $\mu = 5$ so wie oben:

$P$(„1. Anruf in den ersten 10 Minuten")

$$= P\Big(\big]0, \tfrac{1}{6}\big]\Big) = \mathrm{e}^{-0} - \mathrm{e}^{-5/6} \approx 56.54\,\%$$

$P(„1.$ Anruf in den zweiten 10 Minuten“)

$$= P\Big(\big]\tfrac{1}{6}, \tfrac{1}{3}\big]\Big) = \mathrm{e}^{-5/6} - \mathrm{e}^{-5/3} \approx 24.57\,\%$$

$P(„1.$ Anruf in den dritten 10 Minuten“)

$$= P\Big(\big]\tfrac{1}{3}, \tfrac{1}{2}\big]\Big) = \mathrm{e}^{-5/3} - \mathrm{e}^{-5/2} \approx 10.68\,\%$$

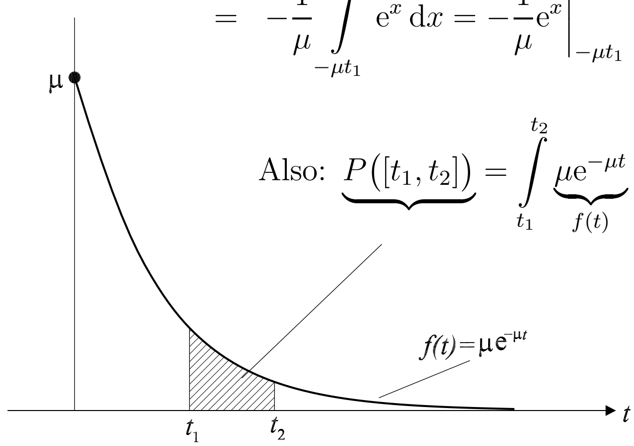
Der Rest sind die am Anfang berechneten 8.21 %.

**Def.:** Es sei $\mu > 0$.
$\Omega = ]0, \infty[$ mit $P\big(]t_1, t_2]\big) = \mathrm{e}^{-\mu t_1} - \mathrm{e}^{-\mu t_2}$ heißt Wahrscheinlichkeitsraum der **Exponentialverteilung** mit Parameter $\mu$, kurz Expo$\,(\mu)$.

(**Bemerkung:** $P\big([t_1, t_2]\big) = P\big(]t_1, t_2]\big) = P\big(]t_1, t_2[\big)$, da $P\big(\{t\}\big) = \lim\limits_{t' \nearrow t} P\big(]t', t]\big) =$
$= \lim\limits_{t' \nearrow t} (\mathrm{e}^{-\mu t'} - \mathrm{e}^{-\mu t}) = 0.$)

**Bildliche Darstellung**

$$\begin{aligned}\int_{t_1}^{t_2} \mathrm{e}^{-\mu t}\,\mathrm{d}t &= \Big(\text{Subst.: } x = -\mu t,\ \frac{\mathrm{d}x}{\mathrm{d}t} = -\mu,\ \mathrm{d}t = -\frac{\mathrm{d}x}{\mu}\\ &= -\frac{1}{\mu}\int_{-\mu t_1}^{-\mu t_2} \mathrm{e}^x\,\mathrm{d}x = -\frac{1}{\mu}\mathrm{e}^x\Big|_{-\mu t_1}^{-\mu t_2} = \frac{1}{\mu}(\mathrm{e}^{-\mu t_1} - \mathrm{e}^{-\mu t_2})\end{aligned}$$




Also: $\underbrace{P\big([t_1, t_2]\big)} = \int_{t_1}^{t_2} \underbrace{\mu \mathrm{e}^{-\mu t}}_{f(t)}\,\mathrm{d}t$

**Satz und Def.:** Wenn $f : \mathbb{R} \longrightarrow \mathbb{R}$ integrierbar, $\forall t : f(t) \geq 0$, und $\int\limits_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$, so ist $\Omega = \mathbb{R}$ ein Wahrscheinlichkeitsraum mit $P\big([t_1, t_2]\big) = \int\limits_{t_1}^{t_2} f(t)\,\mathrm{d}t$.

$\Omega$ heißt **absolut stetiger Wahrscheinlichkeitsraum** mit der **Dichtefunktion** $f$.

**Beweis:**

a) $0 \le P(A) = \int\limits_A f(t)\,\mathrm{d}t \le \int\limits_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$;

b) $P(\Omega) = P(\mathbb{R}) = \int\limits_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$;

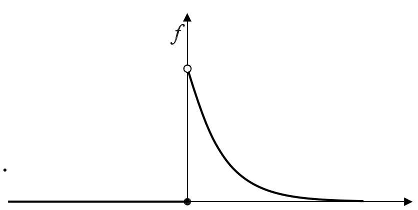
c) $A \cap B = \{\} \Longrightarrow P(A \cup B) = \int\limits_{A\cup B} f(t)\,\mathrm{d}t$

$= \int\limits_A f(t)\,\mathrm{d}t + \int\limits_B f(t)\,\mathrm{d}t = P(A) + P(B)$ $\square$

**Bsp.:** 1) Exponentialverteilung mit Parameter $\mu$ :

Als Dichtefunktion nehmen wir

$$f(t) = \begin{cases} \mu \mathrm{e}^{-\mu t} & : \quad t > 0 \\ 0 & : \quad t \le 0. \end{cases}$$

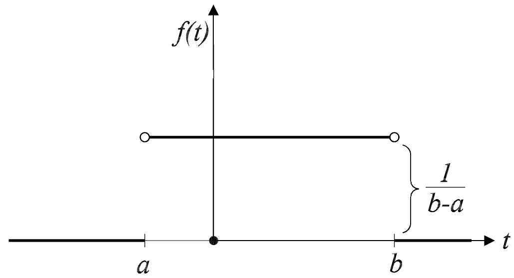



Eigentlich sind $t \le 0$ gar keine Elementarereignisse. Ebenso gut können wir aber auch $t \le 0$ zulassen, wenn wir allen Ereignissen $A \subset\, ]-\infty, 0]$ die Wahrscheinlichkeit $P(A) = 0$ zuweisen, d.h. wenn wir $f(t) = 0$ für $t < 0$ setzen.

2) Gleichverteilung auf $[a, b]$ :

Setze $f(t) = \begin{cases} \frac{1}{b-a} & : \quad a < t < b, \\ 0 & : \quad \text{sonst.} \end{cases}$

Dann gilt $\int\limits_c^d f(t)\,\mathrm{d}t = \dfrac{d-c}{b-a}$ für $[c, d] \subset [a, b]$ wie es sein soll (vergl. Seite 15)




## C) Die Normalverteilung

**Aufgabe:** Bei 100-maligem Würfeln erwarten wir jede Augenzahl (wie z.B. 6) etwa $\frac{100}{6} = 16.\dot{6}$ mal. Wie wahrscheinlich ist es, dass bei 100-mal Würfeln zwischen 12 und 22 Sechser kommen? (Gegenteil = „Ausreißer“)
Also: Bestimme $P\big(\text{Anzahl der 6-er bei 100 Würfen} \in [12, 22]\big)$
**Exakte Lösung:** Entspricht Bin $\left(100, \frac{1}{6}\right)$
$\Longrightarrow P(\text{„12 bis 22 Sechser“}) = p_{12} + p_{13} + \cdots + p_{22}$

$$= \sum_{i=12}^{22} p_i = \sum_{i=12}^{22} \binom{100}{i} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{100-i}.$$

Eine mühsame Rechnung liefert 85.923045 %.
Die Approximation durch die Poissonverteilung ist nicht sehr gut, weil $p$ nicht klein genug ist: $\lambda = np = \dfrac{100}{6} = 16.\dot{6}$, $\sum\limits_{i=12}^{22} \dfrac{\lambda^i}{i!}\,\mathrm{e}^{-\lambda} \approx 82.121186\,\%$.
**Problem:** Wie lässt sich Bin$\,(n, p)$ für $n \to \infty$, $p$ fest (nicht $\to 0$ wie in Seite 12) approximieren?

**Def.:** Es seien $\mu \in \mathbb{R}$, $\sigma > 0$. Dann heißt $\Omega = \mathbb{R}$ mit der Dichtefunktion

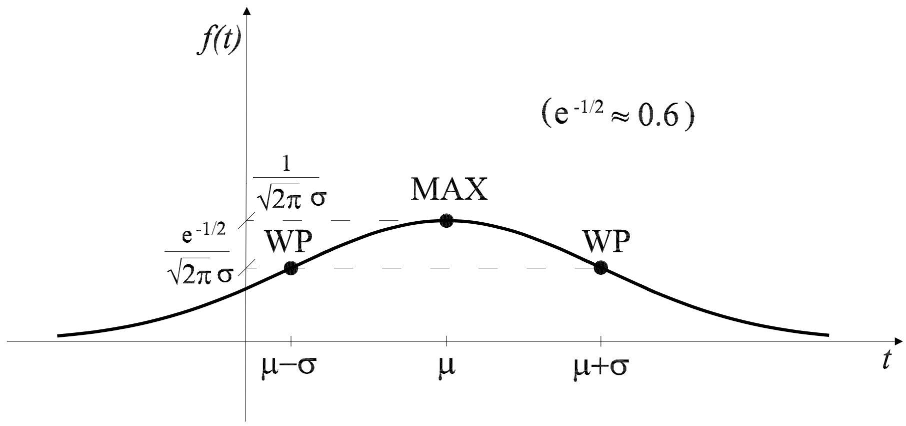
$$\boxed{f(t) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \mathrm{e}^{-(t-\mu)^2/(2\sigma^2)}}$$

**Normalverteilung** mit den Parametern $\mu, \sigma$, kurz $N(\mu, \sigma)$.

(Beweis von $\displaystyle\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$ siehe Seite 24;
später $\mu$ = „Erwartungswert“, $\sigma$ = „Streuung“)

a) **Kurvendiskussion von** $f$ :

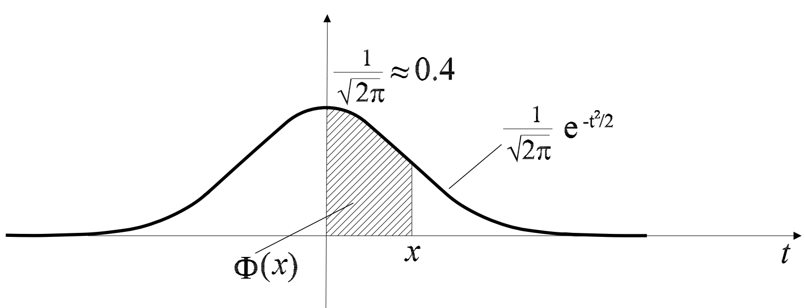
$$\begin{aligned}
f'(t) &= \frac{\mu - t}{\sqrt{2\pi}\sigma^3} \cdot \mathrm{e}^{-(t-\mu)^2/(2\sigma^2)} \\
f'(t) &= 0 \Longleftrightarrow t = \mu \\
f''(t) &= \frac{(\mu - t)^2 - \sigma^2}{\sqrt{2\pi}\sigma^5} \cdot \mathrm{e}^{-(t-\mu)^2/(2\sigma^2)} \\
f''(t) &= 0 \Longleftrightarrow (\mu - t)^2 = \sigma^2 \Longleftrightarrow t = \mu \pm \sigma
\end{aligned}$$




b) **Def.:** $N(0,1)$ heißt **Standardnormalverteilung**.

Die Tabelle auf Seite 22 gibt $\boxed{\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x \mathrm{e}^{-t^2/2}\,\mathrm{d}t}$ für $x$ zwischen 0 und 3.99.

Für $x \geq 0$ gilt $\Phi(x) = P\big([0,x]\big)$ bzgl. $N(0,1)$.

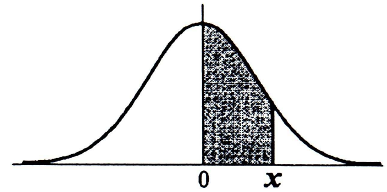



Weiters gilt $\Phi(-x) = -\Phi(x)$

(denn $\int\limits_0^{-x} \mathrm{e}^{-t^2/2}\,\mathrm{d}t = \big(\text{Subst. } u = -t\big) = \int\limits_0^{x} \mathrm{e}^{-u^2/2}\,(-\mathrm{d}u) = -\Phi(x)\big)$

# *Appendix II*

**$\Phi(x)$= AREA under the STANDARD NORMAL CURVE from 0 to $x$**




| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0754 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2258 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2518 | .2549 |
| 0.7 | .2580 | .2612 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2996 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |
| 1.1 | .3643 | .3665 | .3686 | .3708 | .3729 | .3749 | .3770 | .3790 | .3810 | .3830 |
| 1.2 | .3849 | .3869 | .3888 | .3907 | .3925 | .3944 | .3962 | .3980 | .3997 | .4015 |
| 1.3 | .4032 | .4049 | .4066 | .4082 | .4099 | .4115 | .4131 | .4147 | .4162 | .4177 |
| 1.4 | .4192 | .4207 | .4222 | .4236 | .4251 | .4265 | .4279 | .4292 | .4306 | .4319 |
| 1.5 | .4332 | .4345 | .4357 | .4370 | .4382 | .4394 | .4406 | .4418 | .4429 | .4441 |
| 1.6 | .4452 | .4463 | .4474 | .4484 | .4495 | .4505 | .4515 | .4525 | .4535 | .4545 |
| 1.7 | .4554 | .4564 | .4573 | .4582 | .4591 | .4599 | .4608 | .4616 | .4625 | .4633 |
| 1.8 | .4641 | .4649 | .4656 | .4664 | .4671 | .4678 | .4686 | .4693 | .4699 | .4706 |
| 1.9 | .4713 | .4719 | .4726 | .4732 | .4738 | .4744 | .4750 | .4756 | .4761 | .4767 |
| 2.0 | .4772 | .4778 | .4783 | .4788 | .4793 | .4798 | .4803 | .4808 | .4812 | .4817 |
| 2.1 | .4821 | .4826 | .4830 | .4834 | .4838 | .4842 | .4846 | .4850 | .4854 | .4857 |
| 2.2 | .4861 | .4864 | .4868 | .4871 | .4875 | .4878 | .4881 | .4884 | .4887 | .4890 |
| 2.3 | .4893 | .4896 | .4898 | .4901 | .4904 | .4906 | .4909 | .4911 | .4913 | .4916 |
| 2.4 | .4918 | .4920 | .4922 | .4925 | .4927 | .4929 | .4931 | .4932 | .4934 | .4936 |
| 2.5 | .4938 | .4940 | .4941 | .4943 | .4945 | .4946 | .4948 | .4949 | .4951 | .4952 |
| 2.6 | .4953 | .4955 | .4956 | .4957 | .4959 | .4960 | .4961 | .4962 | .4963 | .4964 |
| 2.7 | .4965 | .4966 | .4967 | .4968 | .4969 | .4970 | .4971 | .4972 | .4973 | .4974 |
| 2.8 | .4974 | .4975 | .4976 | .4977 | .4977 | .4978 | .4979 | .4979 | .4980 | .4981 |
| 2.9 | .4981 | .4982 | .4982 | .4983 | .4984 | .4984 | .4985 | .4985 | .4986 | .4986 |
| 3.0 | .4987 | .4987 | .4987 | .4988 | .4988 | .4989 | .4989 | .4989 | .4990 | .4990 |
| 3.1 | .4990 | .4991 | .4991 | .4991 | .4992 | .4992 | .4992 | .4992 | .4993 | .4993 |
| 3.2 | .4993 | .4993 | .4994 | .4994 | .4994 | .4994 | .4994 | .4995 | .4995 | .4995 |
| 3.3 | .4995 | .4995 | .4995 | .4996 | .4996 | .4996 | .4996 | .4996 | .4996 | .4997 |
| 3.4 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4998 |
| 3.5 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 |
| 3.6 | .4998 | .4998 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.7 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.8 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.9 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 |

c) **Berechnung** von $P\big([a,b]\big)$ für $N(\mu,\sigma)$ mittels $\Phi$ :

Bezüglich $N(\mu,\sigma)$ gilt:

$$P\big([a,b]\big) = \frac{1}{\sqrt{2\pi}\sigma}\int\limits_a^b \mathrm{e}^{-(t-\mu)^2/2\sigma^2}\,\mathrm{d}t =$$

(Subst. $u = \frac{t-\mu}{\sigma}$, $\mathrm{d}u = \frac{\mathrm{d}t}{\sigma}$, $\mathrm{d}t = \sigma\,\mathrm{d}u$)

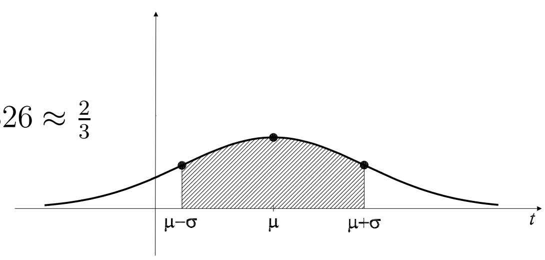
$$= \frac{1}{\sqrt{2\pi}}\int\limits_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \mathrm{e}^{-u^2/2}\,\mathrm{d}u = \frac{1}{\sqrt{2\pi}}\int\limits_0^{\frac{b-\mu}{\sigma}} \mathrm{e}^{-u^2/2}\,\mathrm{d}u - \frac{1}{\sqrt{2\pi}}\int\limits_0^{\frac{a-\mu}{\sigma}} \mathrm{e}^{-u^2/2}\,\mathrm{d}u =$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Also: $$\boxed{P_{N(\mu,\sigma)}\big([a,b]\big) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

Z.B.: $P\big([\mu-\sigma,\mu+\sigma]\big) = \Phi(1) - \Phi(-1) = 2\Phi(1) =$ (Tabelle) $\approx 68.26\,\%$.

schraffierte Fläche $\approx 0.6826 \approx \frac{2}{3}$




d) **Kontrolle**, dass $\int\limits_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$ :

$$\int\limits_{-\infty}^{\infty} f(t)\,\mathrm{d}t = P\big(]-\infty,\infty[\big) = \Phi\left(\frac{\infty-\mu}{\sigma}\right) - \Phi\left(\frac{-\infty-\mu}{\sigma}\right) = \Phi(\infty) - \Phi(-\infty) =$$

$$= 2\Phi(\infty) = \frac{2}{\sqrt{2\pi}}\int\limits_0^{\infty} \mathrm{e}^{-t^2/2}\,\mathrm{d}t.$$

Zeige also: $\Phi(\infty) = \dfrac{1}{2}$ bzw. $I := \int\limits_0^{\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = \sqrt{\frac{\pi}{2}}$ bzw. $I^2 = \dfrac{\pi}{2}$.

Beweis davon:

$$\begin{aligned}
I^2 &= \int_0^\infty \mathrm{e}^{-x^2/2}\,\mathrm{d}x \cdot \int_0^\infty \mathrm{e}^{-y^2/2}\,\mathrm{d}y = \\
&= \int_0^\infty \int_0^\infty \mathrm{e}^{-(x^2+y^2)/2}\,\mathrm{d}x\mathrm{d}y = \text{ (Polarkoordinaten)} \\
&= \int_{r=0}^\infty \int_{\varphi=0}^{\pi/2} \mathrm{e}^{-r^2/2} r\,\mathrm{d}\varphi\mathrm{d}r = \frac{\pi}{2}\int_0^\infty \mathrm{e}^{-r^2/2} r\,\mathrm{d}r \\
&= \left(\text{Subst. } u = \frac{r^2}{2},\ \mathrm{d}u = r\,\mathrm{d}r\right) = \frac{\pi}{2}\int_0^\infty \mathrm{e}^{-u}\,\mathrm{d}u \\
&= \frac{\pi}{2}\left[-\mathrm{e}^{-u}\right]_{u=0}^\infty = -\frac{\pi}{2}\left[0-1\right] = \frac{\pi}{2} \checkmark
\end{aligned}$$

e) **Satz: (Jacob Bernoulli, ″Goldenes Theorem”, ≈ 1700)**

Für $p, C$ fest, $n \to \infty, i \to \infty$, $\left|\frac{i-np}{\sqrt{npq}}\right| \le C$ gilt

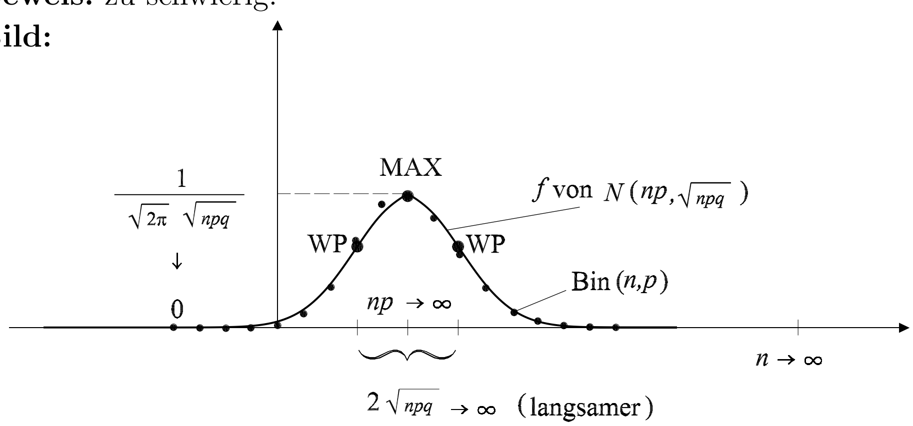
$$p_i^{\mathrm{Bin}\,(n,p)} = \binom{n}{i} p^i q^{n-i} \approx \frac{\mathrm{e}^{-\frac{(i-np)^2}{2npq}}}{\sqrt{2\pi}\,\sqrt{npq}} = f(i),$$

wenn $\mu = np$, $\sigma = \sqrt{npq}$ und $f(t)$ die Dichtefunktion von $N(\mu,\sigma)$ ist.

$\approx$ heißt exakt: $\lim\limits_{\substack{i,n\to\infty \\ \left|\frac{i-np}{\sqrt{npq}}\right| \le C}} \frac{p_i}{f(i)} = 1$

**Beweis:** zu schwierig.

**Bild:**



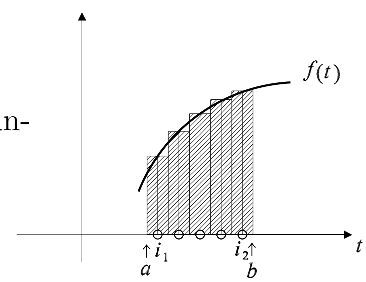
**Beachte:** Für $p$ fest, $n \to \infty$ konvergieren alle $p_i$ gegen 0. Da $\lim\limits_{n\to\infty} p_i = 0$, betrachtet man $\lim\limits_{n\to\infty} \frac{p_i}{f(i)}$. Die Bedingung $\left|\frac{i-np}{\sqrt{npq}}\right| \le C$ heißt, dass $i$ höchstens $C$-mal soweit wie die zwei Wendepunkte vom Maximum weg liegt.
**Salopp zusammengefasst:** $p$ fest, $n \to \infty$

$$\Longrightarrow \qquad \boxed{\mathrm{Bin}\,(n,p) \approx N\big(np, \sqrt{npq}\,\big)}$$

f) **Anwendung:** Setze $\mu = np,\ \sigma = \sqrt{npq}$.
Für $n$ groß genug gilt dann

$$\sum_{i=i_1}^{i_2} p_i \approx \sum_{i=i_1}^{i_2} f(i) \approx \int\limits_{\left.i_1-1/2\right\}a}^{\left.i_2+1/2\right\}b} f(t)\,\mathrm{d}t = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Warum $a = i_1 - \frac{1}{2},\ b = i_2 + \frac{1}{2}$?
$\sum\limits_{i=i_1}^{i_2} f(i) =$ Fläche der Rechtecke $\approx$ Fläche unter der Kurve $= \int\limits_{i_1-1/2}^{i_2+1/2} f(t)\,\mathrm{d}t$




g) **Fortführung der Aufgabe**

$\mu = np = \frac{100}{6} = 16.\dot{6},\ \sigma = \sqrt{npq} = \sqrt{100 \cdot \frac{1}{6} \cdot \frac{5}{6}} \approx 3.727$

$P(\text{„12 bis 22 Sechser“}) = \sum\limits_{i=12}^{22} p_i \overset{\text{Satz}}{\approx} \sum\limits_{i=12}^{22} f(i) \approx \int\limits_{11.5}^{22.5} f(t)\,\mathrm{d}t = P_{N(\mu,\sigma)}\big([11.5, 22.5]\big) =$
$= \Phi\left(\frac{22.5-\mu}{\sigma}\right) - \Phi\left(\frac{11.5-\mu}{\sigma}\right) \approx \Phi(1.565) - \Phi(-1.386) =$
$= \Phi(1.565) + \Phi(1.386) \underset{\text{Tabelle}}{\approx} 0.4412 + 0.4171 = \underline{\underline{0.8583}}$
in guter Übereinstimmung mit Seite 20.

h) **Bemerkung**
Längen, Maße, etc. von **ähnlichen** Objekten in Natur und Technik ergeben sich als Summe vieler kleiner schwankender Größen, ähnlich wie Bin $(n,p)$ und sind daher $\approx$ normalverteilt.
Genauer: siehe zentraler Grenzwertsatz (Seite 54)

i) **Beispiel** Eine Maschine produziert Butterpakete, deren Gewicht normalverteilt ist mit $\mu = 25$ dag, $\sigma = 1$ dag.

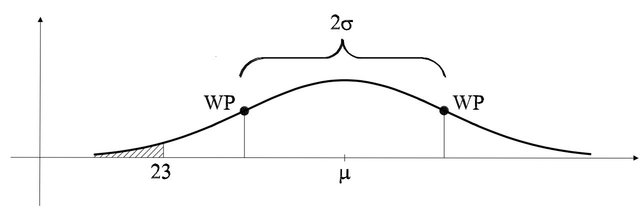
1) Wieviel % der Pakete haben Gewicht $\leq 23$ dag?
2) Wie müsste $\mu$ verändert werden (bei festem $\sigma = 1$ dag), dass nur 1 % der Pakete $\leq 23$ dag wiegt?
3) Wie müsste $\sigma$ verändert werden (bei festem $\mu = 25$ dag), dass nur 1 % der Pakete $\leq 23$ dag wiegt?

**Lösung** Ad 1): $P\big(]\underbrace{-\infty}_{a}, \underbrace{23}_{b}]\big) = \Phi\left(\frac{23-25}{1}\right) - \Phi\left(\frac{-\infty-25}{1}\right) =$
$= \Phi(-2) - \Phi(-\infty) = -\Phi(2) + \Phi(\infty) = -0.4772 + 0.5 = 0.0228$
Also: 2.28 % der Pakete haben Gewicht $\leq 23$ dag.

Ad 2): $P\big(]-\infty, 23]\big) = \Phi\left(\frac{23-\mu}{1}\right) - \Phi\left(\frac{-\infty-\mu}{1}\right) = \Phi(23-\mu) - \Phi(-\infty) =$
$= -\Phi(\mu - 23) + \Phi(\infty) = \frac{1}{2} - \Phi(\mu - 23) \stackrel{!}{=} 0.01$
$\Longrightarrow \Phi(\mu - 23) = 0.49 \Longrightarrow$ (Tabelle) $\mu - 23 = 2.327$
Also: $\mu = 25.327$ dag.

Ad 3): $P\big(]-\infty, 23]\big) = \Phi\left(\frac{23-25}{\sigma}\right) - \Phi\left(\frac{-\infty-25}{\sigma}\right) = \Phi\left(\frac{-2}{\sigma}\right) - \Phi(-\infty) =$
$= \frac{1}{2} - \Phi\left(\frac{2}{\sigma}\right) \stackrel{!}{=} 0.01 \Longrightarrow \Phi\left(\frac{2}{\sigma}\right) = 0.49 \Longrightarrow \frac{2}{\sigma} = 2.327,\ \sigma \approx 0.859$ dag

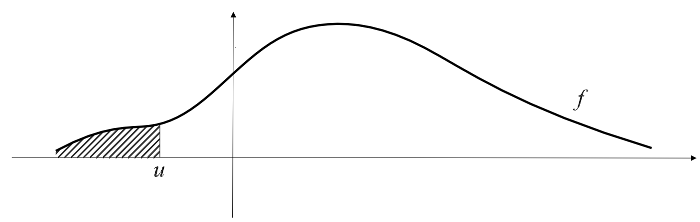
**Ergebnis von 2/3:** Um nur 1 % Pakete $\leq 23$ dag zu produzieren, kann man $\mu$ vergrößern (im Mittel Butter zuschießen) oder $\sigma$ verkleinern (d.h. die Maschine verbessern).
**Skizze zu 2/3:**




Forderung: schraffierte Fläche = 0.01= $\alpha$

**Def.:** Wenn $P$ absolut stetig ist mit Dichtefunktion $f$ und $0 \leq n \leq 100$, so heißt $u$ ***n* %-Fraktil** oder ***n* %-Quantil** von $P \iff P\big(]-\infty, u]\big) = \dfrac{n}{100} = \alpha \iff \displaystyle\int\limits_{-\infty}^{u} f(t)\,\mathrm{d}t = \frac{n}{100} = \alpha$.

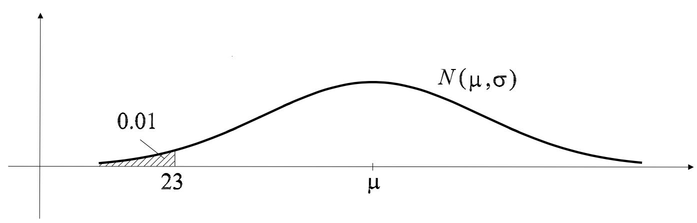
**Bild:**

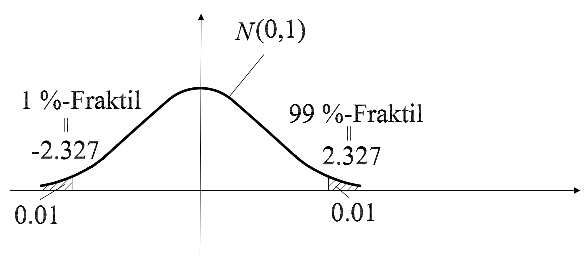



$$\text{schraffierte Fläche} = \frac{n}{100} = \alpha$$

**Bsp. i) 2/3:** Bestimme $\mu$ bzw. $\sigma$ so, dass das 1 %-Fraktil von $N(\mu, \sigma)$ bei 23 dag liegt $\iff \dfrac{23-\mu}{\sigma} = 1\,\%\text{-Fraktil von } N(0,1) \iff \dfrac{23-\mu}{\sigma} = -2.327$

**Bild:**



# § 4 Bedingte Wahrscheinlichkeit, Unabhängigkeit

**Bsp. 1:** 52 Karten, 13 $\heartsuit$, 13 $\diamondsuit$, 13 $\spadesuit$, 13 $\clubsuit$ von 2-10, Bube, Dame, König, Ass.
Wie wahrscheinlich ist es, ein Ass zu ziehen?
Antwort: Wir nummerieren die Karten und geben den Assen die Nummern 1,2,3,4.
$\Omega = \{1, 2, \cdots, 52\}$, $A =$„Ass"$= \{1, 2, 3, 4\}$.
$P(A) = \dfrac{\#A}{\#52} = \dfrac{4}{52} = \dfrac{1}{13} \approx 7.7\,\%$.
Ein Experiment ergibt eine höhere Wahrscheinlichkeit, denn 4 Karten haben blaue Rücken, davon 3 Assen. Die blauen Karten sollen die Nummern 1,2,3,5 haben $\Longrightarrow$ $B =$„blau"$= \{1, 2, 3, 5\}$.
Wenn nur aus den blauen Karten gezogen wird, so ist

$$P(\text{„Ass bei Ziehen aus blau"}) = \frac{\text{Anzahl der günstigen}}{\text{Anzahl der möglichen}} = \frac{\#\{1,2,3\}}{\#B} = \frac{3}{4} = 75\,\%$$

**Def.:** $A, B$ seien Ereignisse in einem Wahrscheinlichkeitsraum $\Omega$ und $P(B) \neq 0$.
$P(A|B) =$ Wahrscheinlichkeit von $A$ unter der Voraussetzung (=Hypothese), dass $B$ eingetroffen ist.
$P(A|B)$ heißt **bedingte Wahrscheinlichkeit von *A* unter der Hypothese *B***.

**Satz:**

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

(Das wird manchmal **Bayes'sche Formel** genannt.)

**Beweis:** Führe das zu $\Omega$ gehörige Experiment $N$-mal aus. In $K_N$ Fällen tritt $B$ ein, in $L_N$ Fällen davon treten $A$ und $B$ ein $\Longrightarrow$

$$\Longrightarrow\ P(A|B) = \lim_{N\to\infty} \frac{L_N}{K_N} = \lim_{N\to\infty} \frac{\frac{L_N}{N}}{\frac{K_N}{N}} = \frac{\lim\limits_{N\to\infty} \frac{L_N}{N}}{\lim\limits_{N\to\infty} \frac{K_N}{N}} = \frac{P(A \cap B)}{P(B)}. \qquad \square$$

**Bsp. 1 von vorher:**
$P(A \cap B) = P\big(\{1, 2, 3\}\big) = \dfrac{3}{52}, P(B) = \dfrac{4}{52} \Longrightarrow P(A|B) = \dfrac{3/52}{4/52} = 75\,\%$,
wie wir oben schon direkt überlegten.

**Bsp. 2:** In einer Fabrik produzieren zwei Maschinen die gleichen Werkstücke. Der Ausstoß der ersten Maschine ist doppelt so hoch wie der der zweiten, jedoch sind 12 % der Produkte der ersten Maschine defekt, hingegen nur 6 % bei der zweiten. Ein Kunde erhält ein defektes Werkstück. Mit welcher Wahrscheinlichkeit kommt es von Maschine 1 bzw. Maschine 2?

**Mathematische Formulierung:**
$D =$ „Werkstück defekt“,
$A_i =$ „Werkstück kommt von Maschine $i$“, $i = 1, 2$
$P(A_1) = \frac{2}{3},\ P(A_2) = \frac{1}{3},\ P(D|A_1) = 0.12,\ P(D|A_2) = 0.06$

Gesucht: $P(A_1|D),\ P(A_2|D)$.

**Lösung:** a) $P(D|A_1) = \frac{P(D \cap A_1)}{P(A_1)} \implies$
$\implies P(D \cap A_1) = P(A_1) \cdot P(D|A_1) = \frac{2}{3} \cdot 0.12 = 0.08$
Ebenso $P(D \cap A_2) = P(A_2) \cdot P(D|A_2) = \frac{1}{3} \cdot 0.06 = 0.02$

b) $P(D) = P\big((D \cap A_1) \cup (D \cap A_2)\big) = P(D \cap A_1) + P(D \cap A_2) = 0.1$
schließen sich aus

c) $P(A_1|D) = \frac{P(A_1 \cap D)}{P(D)} = \frac{P(D \cap A_1)}{P(D)} = \frac{0.08}{0.1} = 0.8 = 80\,\%$
$P(A_2|D) = \frac{P(D \cap A_2)}{P(D)} = \frac{0.02}{0.1} = 0.2 = 20\,\%$
Also: Ein defektes Produkt kommt zu $80\,\%$ von $A_1$, zu $20\,\%$ von $A_2$.
Die obige Vorgangsweise wird formalisiert im

**Satz: (Satz von Bayes)**
$D, A_1, \cdots, A_n$ seien Ereignisse, sodass sich $A_i$ und $A_j$ für $i \neq j$ ausschließen und mit Sicherheit eines der Ereignisse $A_1, \cdots, A_n$ eintrifft (d.h. $\Omega = A_1 \cup \cdots \cup A_n \wedge \forall i \neq j : A_i \cap A_j = \{\}$). Weiters seien $P(D) \neq 0,\ P(A_1) \neq 0, \cdots, P(A_n) \neq 0$. Dann gilt:

$$P(A_k|D) = \frac{P(D|A_k) \cdot P(A_k)}{\sum\limits_{i=1}^{n} P(D|A_i) \cdot P(A_i)}.$$

**Beweis:** Analog zu a) - c) oben:

$$P(A_k|D) = \frac{P(A_k \cap D)}{P(D)} = \frac{P(D|A_k)P(A_k)}{\sum\limits_{i=1}^{n} P(D \cap A_i)} = \frac{P(D|A_k) \cdot P(A_k)}{\sum\limits_{i=1}^{n} P(D|A_i) \cdot P(A_i)} \qquad \square$$

**Def.:** Zwei Ereignisse $A, B$ heißen **unabhängig** genau dann, wenn

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

**Bemerkung:** Wenn diese Gleichung gilt, so ist entweder $P(B) = 0$ oder $P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$, d.h. die Wahrscheinlichkeit von $A$ wird durch das Wissen, dass $B$ eingetroffen ist, nicht verändert. Das entspricht der anschaulichen Bedeutung, die man mit dem Begriff der Unabhängigkeit des Eintreffens zweier Ereignisse verbindet.

**Bsp. 3:** Würfeln mit zwei (verschiedenen) Würfeln:
$A_i =$ „1. Würfel hat Augenzahl $i$",
$B_j =$ „2. Würfel hat Augenzahl $j$";
1. Würfel regulär $\implies \forall i: \ P(A_i) = \frac{1}{6}$,
2. Würfel regulär $\implies \forall j: \ P(B_j) = \frac{1}{6}$;
vom Experiment her postulieren (=fordern) wir, dass $A_i$ und $B_j$ für alle $i, j$ unabhängig sind (d.h. die Wahrscheinlichkeit, dass der 2. Würfel Augenzahl $j$ hat, wird durch das Wissen, dass der 1. Würfel Augenzahl $i$ hat, nicht verändert)

$$\implies P\big((i,j)\big) = P(A_i \cap B_j) = P(A_i) \cdot P(B_j) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

So sieht man, dass $\Omega = \big\{(i,j) : 1 \leq i, \ j \leq 6\big\}$ ein symmetrischer Wahrscheinlichkeitsraum ist, wie dies in Seite 4 unten vorausgesetzt wurde.
Etwas Analoges wurde verwendet auf Seite 4, Bsp. 1, und auf Seite 9.

# Kapitel 2

# Zufallsgrößen

## § 5 Zufallsgrößen und ihre Verteilungen

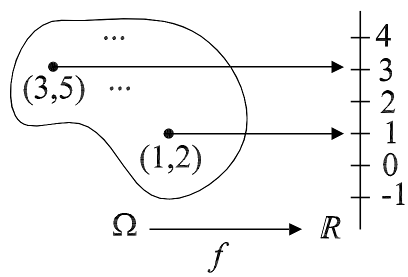
**Problem:** Würfeln mit zwei verschiedenen Würfeln, $\Omega = \big\{(i,j) : 1 \leq i, j \leq 6\big\}$. Welchem mathematischen Objekt entspricht

$$f = \text{„die Augenzahl des 1. Würfels“?}$$

Antwort:

$$f : \Omega \longrightarrow \mathbb{R} : (i,j) \longmapsto i$$

Bild:



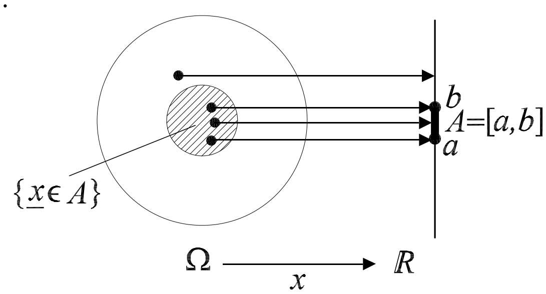

**Def.:** $\Omega$ sei ein Wahrscheinlichkeitsraum. Eine Funktion $f : \Omega \longrightarrow \mathbb{R}$ heißt **Zufallsgröße** (oder **Zufallsvariable**).

**Experimentelle Bedeutung:** Eine Zufallsgröße $f$ ordnet einem Elementarereignis $\omega$ die reelle Zahl $f(\omega)$ zu. In dem $\Omega$ entsprechenden Experiment ist $f$ eine Messgröße oder eine aus Messgrößen herleitbare Größe.
**Mathematische Bemerkung:** $\mathcal{B}$ sei die kleinste Sigma-Algebra auf $\mathbb{R}$, die alle Intervalle enthält. ($\mathcal{B}$ heißt **Borel-Sigma-Algebra**.) Für Zufallsgrößen verlangt man genaugenommen:

$$\forall A \in \mathcal{B} : \{\omega \in \Omega : f(\omega) \in A\} \text{ ist ein Ereignis.}$$

Das ist in der Praxis immer erfüllt.

**Schreibweisen:** 1) Statt $f$ schreibt man $\underline{x}, \underline{y}$ etc.
2) Für das Ereignis $\{\omega \in \Omega : \underline{x}(\omega) \in A\}$ schreibt man $\{\underline{x} \in A\}$ (wobei $A \in \mathcal{B}$)
Analog $\{\omega \in \Omega : \underline{x}(\omega) = 5\} = \{\underline{x} = 5\}$ und $\{\omega \in \Omega : \underline{x}(\omega) \geq 7\} = \{\underline{x} \geq 7\}$ etc.
Bild:




**Bemerkung:** Wenn $\underline{x}, \underline{y}$ Zufallsgrößen auf $\Omega$ und $c, d \in \mathbb{R}$, so sind auch $c\underline{x}$, $\underline{x}+\underline{y}$, $c\underline{x} + d\underline{y}$, $\underline{x} \cdot \underline{y}$, $\underline{x}/\underline{y}$ (falls $\forall \omega \in \Omega : \underline{y}(\omega) \neq 0$) Zufallsgrößen.
Für $h : \mathbb{R} \longrightarrow \mathbb{R}$ ist $h \circ \underline{x}$ eine Zufallsgröße, z.B. $\underline{x} - c$, $(\underline{x} - c)^2$, $\mathrm{e}^{\underline{x}}$, $|\underline{x}|$, $\arctan \underline{x}$ etc.

**Bsp. 1:** Würfeln mit zwei verschiedenen Würfeln,
$\underline{x}$ = „Augenzahl 1. Würfel", $\underline{y}$ = „Augenzahl 2. Würfel"
$\Longrightarrow$ $\underline{x} + \underline{y}$ = „Augensumme", $\underline{x}/\underline{y}$ = „Quotient der Augenzahlen";
$\{\underline{x} + \underline{y} = 6\} = \big\{(1,5),(5,1),(2,4),(4,2),(3,3)\big\}$
$\{\underline{x}/\underline{y} = 1\} = \big\{(1,1),(2,2),\cdots,(6,6)\big\} = \{\underline{x} - \underline{y} = 0\}$

**Def. und Satz:** $\underline{x}$ sei eine Zufallsvariable auf $\Omega$. Dann ist $\mathbb{R}$ mit

$$P_{\underline{x}}(A) = P\big(\{\underline{x} \in A\}\big) = P\big(\{\omega \in \Omega : \underline{x}(\omega) \in A\}\big)$$

(für $A \in \mathcal{B}$) ein Wahrscheinlichkeitsraum.
$P_{\underline{x}}$ heißt die **Wahrscheinlichkeitsverteilung von $\underline{x}$**.

**Beweis:** a) $0 \leq P_{\underline{x}}(A) \leq 1 \checkmark$
b) $P_{\underline{x}}(\mathbb{R}) = P\big(\{\underline{x} \in \mathbb{R}\}\big) = P(\Omega) = 1$
c) $A \cap B = \{\} \implies \{\underline{x} \in A\} \cap \{\underline{x} \in B\} = \{\} \implies$
$\implies P_{\underline{x}}(A \cup B) = P\big(\{\underline{x} \in A \cup B\}\big) = P\big(\{\underline{x} \in A\} \cup \{\underline{x} \in B\}\big)$
$= P\big(\{\underline{x} \in A\}\big) + P\big(\{\underline{x} \in B\}\big) = P_{\underline{x}}(A) + P_{\underline{x}}(B)$. $\square$

**Bsp. 1:** $\underline{x}, \underline{y}$ wie oben bei zwei Würfeln;

$$P_{\underline{x}}\big(\{1\}\big) = P\Big(\big\{\underline{x} \in \{1\}\big\}\Big) = P\big(\{\underline{x} = 1\}\big)$$
$$= P\Big(\big\{(1,1),(1,2),\cdots,(1,6)\big\}\Big) = \frac{6}{36} = \frac{1}{6};$$

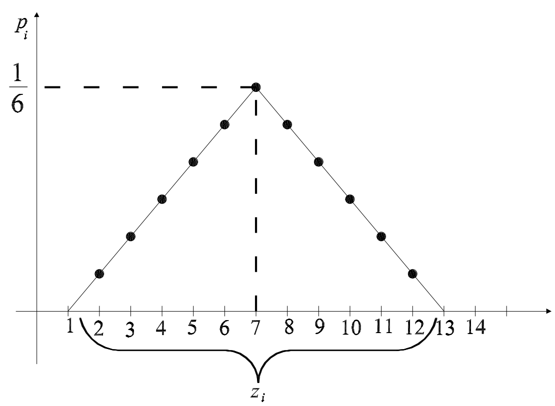
ebenso $p_i = P_{\underline{x}}\big(\{i\}\big) = \dfrac{1}{6}$ für $i = 1, 2, \cdots, 6$, d.h. $P_{\underline{x}}$ ist die Gleichverteilung auf $\{1, 2, \cdots, 6\}$.
(In §§ 1,2 wurde in diesem Fall nur $\{1, 2, \cdots, 6\}$ als Ereignisraum betrachtet. Nun ist $\mathbb{R}$ der Ereignisraum von $P_{\underline{x}}$, aber wegen $P_{\underline{x}}(A) = 0$, wenn $A \cap \{1, 2, \cdots, 6\} = \{\}$, macht das keinen Unterschied. Vergleiche auch Seite 19.)
Aus Symmetriegründen ist $P_{\underline{x}} = P_{\underline{y}}$, obwohl $\underline{y} \neq \underline{x}$, z.B. ist $\underline{x}\big((1,2)\big) = 1$, $\underline{y}\big((1,2)\big) = 2$.
Für $\underline{z} = \underline{x} + \underline{y}$ gilt z.B.

$$P_{\underline{z}}\big(\{4\}\big) = P\big(\{\underline{z} = 4\}\big) = P\Big(\big\{(1,3), (3,1), (2,2)\big\}\Big) = \frac{3}{36}$$

Allgemein:

| $z_i = i,\ i = 2, \cdots, 12$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_i = P_{\underline{z}}\big(\{i\}\big)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

Bild:




**Def.:** Eine Zufallsgröße $\underline{x}$ heißt **diskret**, wenn $\underline{x}$ nur endlich oder abzählbar viele Werte $x_i$ annimmt (oft ist $x_i = i$). Dann sei $p_i = P_{\underline{x}}\big(\{x_i\}\big) = P\big(\{\underline{x} = x_i\}\big)$.
$\underline{x}$ heißt **absolutstetig** $\iff$ $P_{\underline{x}}$ hat eine Dichtefunktion $f_{\underline{x}}$, d.h.

$$\int_a^b f_{\underline{x}}(t)\,\mathrm{d}t = P_{\underline{x}}\big([a,b]\big) = P\Big(\big\{\underline{x} \in [a,b]\big\}\Big)$$

**Bsp. 2:** $n$-mal Münze werfen, die mit Wahrscheinlichkeit $p$ auf Zahl fällt,
$\underline{x} =$ „Anzahl von Zahl“
$\Longrightarrow\ \Omega = \big\{(b_1, \cdots, b_n) : b_j \in \{0,1\}\big\}$, $1 \hat{=}$ „Zahl“,
$\underline{x} : \Omega \longrightarrow \mathbb{R} : (b_1, \cdots, b_n) \longmapsto \sum\limits_{j=1}^{n} b_j$ ist diskret

$x_i = i,\ i = 0, \cdots, n,\ \ p_i = P\big(\{\underline{x} = i\}\big) = \binom{n}{i} p^i q^{n-i}$ (siehe Seite 9),
$P_{\underline{x}} = \mathrm{Bin}\,(n, p)$.

**Bsp. 3:** Experiment: Passierzeiten von Fahrzeugen an einer Straße,
$\underline{x}$ = „Anzahl der Fahrzeuge im Zeitraum $[0, t]$“ ($t$ fest),
$\underline{y}$ = „Wartezeit bis zum 1. Fahrzeug“
$\Longrightarrow\ \Omega = \big\{(z_1, z_2, z_3, \cdots) : 0 < z_1 < z_2 < \cdots\big\}$
$z_i$ = Zeitpunkt, wann $i$-tes Fahrzeug kommt,
$\underline{x}\big((z_1, z_2, \cdots)\big) = \#\{i : z_i \le t\},\ \underline{y}\big((z_1, z_2, \cdots)\big) = z_1$
Im Mittel seien $\mu$ Fahrzeuge/h, $\lambda = \mu t$
$\Longrightarrow\ \ \underline{x}$ diskret, $x_i = i,\ \ i = 0, 1, 2, \cdots,\ \ p_i = \dfrac{\lambda^i}{i!} \mathrm{e}^{-\lambda},\ \ P_{\underline{x}} = \mathrm{Poi}(\lambda)$ und $\underline{y}$ absolutstetig,
$P\Big(\big\{\underline{y} \in [t_1, t_2]\big\}\Big) = \int\limits_{t_1}^{t_2} \underbrace{\mu \mathrm{e}^{-\mu t}}_{f_{\underline{y}}(t)}\,\mathrm{d}t$, d.h. $P_{\underline{y}} = \mathrm{Expo}\,(\mu)$, siehe Seiten 13 und 17.

# § 6 Erwartungswert und Varianz

**Problem:** Ein wiederholbares Experiment sei gegeben, $\Omega$ sei der dazugehörige Wahrscheinlichkeitsraum, $\underline{x} : \Omega \longrightarrow \mathbb{R}$ eine Zufallsgröße. Als **Erwartungswert** von $\underline{x}$, kurz $\mathcal{E}(\underline{x})$, versteht man den **Grenzwert der Mittelwerte** von $\underline{x}$ für $N \to \infty$, d.h.

$$\mathcal{E}(\underline{x}) = \lim_{N \to \infty} \frac{\sum\limits_{j=1}^{N} \text{Wert von } \underline{x} \text{ bei } j\text{-ter Ausführung des Experiments}}{N}$$

Wie lässt sich $\mathcal{E}(\underline{x})$ ohne Experimente festlegen und berechnen?

**Def.:** $\underline{x} : \Omega \longrightarrow \mathbb{R}$ sei diskret oder absolutstetig. Dann heißt

$$\mathcal{E}(\underline{x}) = \begin{cases} \sum\limits_i x_i p_i & : \ \underline{x} \text{ diskret,} \\ \int\limits_{-\infty}^{\infty} t f_{\underline{x}}(t)\,\mathrm{d}t & : \ \underline{x} \text{ absolutstetig} \end{cases}$$

**Erwartungswert** von $\underline{x}$. (Zu $x_i, p_i, f_{\underline{x}}$ siehe Seite 33.)

**Bemerkung:** $\mathcal{E}(\underline{x})$ entspricht dem Problem oben:
a) $\underline{x}$ diskret $\Longrightarrow$

$$\lim_{N\to\infty}\frac{\sum\limits_{j=1}^{N}\text{ Wert von }\underline{x}\text{ bei }j\text{-ter Ausführung des Experiments}}{N}=$$

$$=\lim_{N\to\infty}\frac{\sum\limits_i x_i\cdot\text{ Anzahl von }x_i\text{ bei }N\text{-maligem Experiment}}{N}=$$

$$=\sum_i x_i\lim_{N\to\infty}\frac{\text{Anzahl von }x_i\text{ bei }N\text{-maligem Experiment}}{N}=$$

$$=\sum_i x_i P\big(\{\underline{x}=x_i\}\big)=\sum_i x_i p_i$$

b) $\underline{x}$ absolutstetig mit Werten in $[a,b]$, $Z$ Zerlegung, d.h. $a=t_0<t_1<\cdots<t_k=b\Longrightarrow$

$$\begin{array}{ccccc}
\dfrac{\sum\limits_i t_i\cdot\text{Anz. }\underline{x}\in[t_i,t_{i+1}[}{N} & \le & \dfrac{\sum\limits_{j=1}^{N}\text{Wert }\underline{x}\text{ b. }j\text{-ter Ausf.}}{N} & \le & \dfrac{\sum\limits_i t_{i+1}\cdot\text{Anz. }\underline{x}\in[t_i,t_{i+1}[}{N}\\
\downarrow & & & & \downarrow\\
\sum\limits_i t_iP_{\underline{x}}\big([t_i,t_{i+1}]\big) & & & & \sum\limits_i t_{i+1}P_{\underline{x}}\big([t_i,t_{i+1}]\big)\\
\| & & & & \|\\
\sum\limits_i t_i\int\limits_{t_i}^{t_{i+1}}f_{\underline{x}}(t)\,\mathrm{d}t & & & & \sum\limits_i t_{i+1}\int\limits_{t_i}^{t_{i+1}}f_{\underline{x}}(t)\,\mathrm{d}t
\end{array}$$

Wenn $\epsilon=\varphi(Z)=\max\limits_i(t_{i+1}-t_i)$, so folgt

$$\int_a^b(t-\epsilon)f_{\underline{x}}(t)\,\mathrm{d}t\le\lim_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}\text{Wert }\underline{x}\text{ b. }j\text{-ter Ausf. }\le\int_a^b(t+\epsilon)f_{\underline{x}}(t)\,\mathrm{d}t$$

und $\epsilon\searrow 0$ ergibt $\int_a^b tf_{\underline{x}}(t)\,\mathrm{d}t=\lim_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}$ Wert $\underline{x}$ bei $j$-ter Ausführung.

**Bsp. 1:** $P_{\underline{x}}=\operatorname{Bin}(n,p):x_i=i\ (i=0,1,\cdots,n),\ p_i=\binom{n}{i}p^iq^{n-i}\Longrightarrow$

$\Longrightarrow\mathcal{E}(\underline{x})=\sum x_ip_i=\sum\limits_{i=\cancel{0}\,1}^{n}i\binom{n}{i}p^iq^{n-i};$

$$i\cdot\tbinom{n}{i}=i\frac{n!}{i!(n-i)!}=\frac{n!}{(i-1)!(n-i)!}=\frac{n\cdot(n-1)!}{(i-1)!\big(n-1-(i-1)\big)!}=n\cdot\binom{n-1}{i-1}$$

Setze $i-1=l,\ i=l+1;\ i=1\implies l=0,\ i=n\implies l=n-1$

$$\begin{aligned}\Longrightarrow\ \mathcal{E}(\underline{x})\ &=\ \sum_{l=0}^{n-1}n\cdot\binom{n-1}{l}p^{l+1}q^{n-(l+1)}\\
&=\ n\cdot p\sum_{l=0}^{n-1}\binom{n-1}{l}p^lq^{n-1-l}\\
&=\ n\cdot p\cdot\underbrace{(p+q)}_{=1}{}^{n-1}=np\end{aligned}$$

Somit: $\boxed{P_{\underline{x}}=\operatorname{Bin}(n,p)\implies\mathcal{E}(\underline{x})=np}$

**Z.B.:** Experiment: 100-mal Würfeln, $\underline{x}$ = Anzahl der Sechser, $P_{\underline{x}} = \text{Bin}\left(100, \frac{1}{6}\right)$, $\mathcal{E}(\underline{x}) = \frac{100}{6} = 16.\dot{6}$; wenn also $N$ Leute je 100-mal würfeln und jeweils $a_1, a_2, \cdots, a_N$ Sechser dabei erhalten, so ist $\lim\limits_{N\to\infty} \dfrac{1}{N} \sum\limits_{j=1}^{N} a_j = 16.\dot{6}$.

**Berechnung von $\mathcal{E}(\underline{x})$ mit Differenzieren:**

$$
\begin{array}{rlcl}
 & (x+y)^n & = & \sum\limits_{i=0}^{n} \binom{n}{i} x^i y^{n-i} \qquad \Big| \dfrac{\partial}{\partial x} \\
\Longrightarrow & n(x+y)^{n-1} & = & \sum\limits_{i=0}^{n} \binom{n}{i} i x^{i-1} y^{n-i} \ (*) \\
\Longrightarrow & \sum\limits_{i=0}^{n} \binom{n}{i} i p^i q^{n-i} & = & p \cdot \sum\limits_{i=0}^{n} \binom{n}{i} i p^{i-1} q^{n-i} \stackrel{(*)}{=} pn(p+q)^{n-1} = np
\end{array}
$$

**Bsp. 2:** $P_{\underline{x}} = \text{Expo}\,(\mu): \ f_{\underline{x}}(t) = \begin{cases} \mu \mathrm{e}^{-\mu t} & : \ t > 0, \\ 0 & : \ t \leq 0. \end{cases}$

**Berechnung von $\mathcal{E}(\underline{x})$ mit Differenzieren:**

$$
\begin{array}{rl}
 & \int\limits_0^\infty \mu \mathrm{e}^{-\mu t} \,\mathrm{d}t = \int\limits_{-\infty}^{\infty} f_{\underline{x}}(t) \,\mathrm{d}t = 1 \qquad \Big| \dfrac{\partial}{\partial \mu} \\
\Longrightarrow & \underbrace{\int\limits_0^\infty \mathrm{e}^{-\mu t} \,\mathrm{d}t}_{\frac{1}{\mu}} - \underbrace{\int\limits_0^\infty \mu t \mathrm{e}^{-\mu t} \,\mathrm{d}t}_{\int\limits_{-\infty}^{\infty} t f_{\underline{x}}(t) \,\mathrm{d}t = \mathcal{E}(\underline{x})} = 0 \\
\Longrightarrow & \mathcal{E}(\underline{x}) = \dfrac{1}{\mu}.
\end{array}
$$

**Z.B.:** $\underline{x}$ = Lebensdauer von Glühlampen, $P_{\underline{x}} = \text{Expo}\,(\mu)$ mit $\mu$ in $\left[\frac{1}{h}\right]$; wenn also $N$ Glühlampen jeweils $t_1, t_2, \cdots, t_N$ Stunden brennen, so ist $\dfrac{1}{\mu} = \mathcal{E}(\underline{x}) = \lim\limits_{N\to\infty} \dfrac{1}{N} \sum\limits_{j=1}^{N} t_j$, d.h. $\dfrac{1}{\mu}$ = durchschnittliche Lebensdauer.

**Problem:** Beschreibe quantitativ, wie weit $\underline{x}$ „im Schnitt“ von $\mathcal{E}(\underline{x})$ abweicht, d.h., wie konzentriert die Verteilung $P_{\underline{x}}$ ist! Wenn die $N$-malige Ausführung des Experiments für $\underline{x}$ die Werte $a_1, \cdots, a_N$ liefert, so ist $(a_1, \cdots, a_N)$ ein Vektor in $\mathbb{R}^N$ und das Quadrat seiner Entfernung vom „Idealwert“ $\big(\mathcal{E}(\underline{x}), \mathcal{E}(\underline{x}), \cdots, \mathcal{E}(\underline{x})\big) \in \mathbb{R}^N$ ist $\sum\limits_{j=1}^{N} \big(a_j - \mathcal{E}(\underline{x})\big)^2$.

Daher ist $\lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{j=1}^{N} (a_j - \mathcal{E}(\underline{x}))^2 = \mathcal{E}\Big(\big(\underline{x} - \mathcal{E}(\underline{x})\big)^2\Big)$ ein Maß für die Konzentration von $P_{\underline{x}}$.

(Hingegen würde $\lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{j=1}^{N} \big|a_j - \mathcal{E}(\underline{x})\big| = \mathcal{E}\big(\big|\underline{x} - \mathcal{E}(\underline{x})\big|\big)$ der 1-Norm auf $\mathbb{R}^N$ entsprechen und somit nicht dem üblichen „Abstandsgefühl".)

**Def.:** $\underline{x} : \Omega \longrightarrow \mathbb{R}$ sei diskret oder absolutstetig. Dann heißt $\mathcal{V}(\underline{x}) = \mathcal{E}\Big(\big(\underline{x} - \mathcal{E}(\underline{x})\big)^2\Big)$ **Varianz** von $\underline{x}$. $\sigma(\underline{x}) = \sqrt{\mathcal{V}(\underline{x})}$ heißt **Streuung** oder **Standardabweichung** von $\underline{x}$.

Zur Berechnung von $\mathcal{V}(\underline{x})$ brauchen wir

**Satz 0:** $c$ konstante Zufallsgröße $\Longrightarrow$ $\mathcal{E}(c) = c$

**Satz 1:** $\mathcal{E}$ ist linear, d.h. $\mathcal{E}(\underline{x} + \underline{y}) = \mathcal{E}(\underline{x}) + \mathcal{E}(\underline{y}),\ \mathcal{E}(c\underline{x}) = c\mathcal{E}(\underline{x})$

**Satz 2:** $h : \mathbb{R} \longrightarrow \mathbb{R} \implies \mathcal{E}(h \circ \underline{x}) = \begin{cases} \sum\limits_i h(x_i)p_i & : \ \underline{x} \text{ diskret,} \\ \int\limits_{-\infty}^{\infty} h(t) f_{\underline{x}}(t)\,\mathrm{d}t & : \ \underline{x} \text{ absolutstetig} \end{cases}$

**Beweis:** 0) $x_1 = c,\ p_1 = 1 \implies \mathcal{E}(c) = x_1 \cdot p_1 = c$

1) $\mathcal{E}(c\underline{x} + \underline{y}) = \lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{j=1}^{N}$ Wert von $(c\underline{x} + \underline{y})$ bei $j$-ter Ausführung des Experiments

$= \lim\limits_{N\to\infty} \left[ \frac{c}{N} \sum\limits_{j=1}^{N} \text{Wert von } \underline{x} \text{ bei } j\text{-ter Ausf.} + \frac{1}{N} \sum\limits_{j=1}^{N} \text{Wert } \underline{y} \text{ bei } j\text{-ter Ausf.} \right] = c\mathcal{E}(\underline{x}) + \mathcal{E}(\underline{y}).$

2) Z.B. für $\underline{x}$ diskret: $h \circ \underline{x}$ nimmt die Werte $h(x_i)$ mit Wahrscheinlichkeit $p_i$ an $\Longrightarrow$ $\mathcal{E}(h \circ \underline{x}) = \sum\limits_i h(x_i)p_i$. □

**Folgerung:** Es gilt $\mathcal{V}(\underline{x}) = \mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2$ und

$$\mathcal{E}(\underline{x}^2) = \begin{cases} \sum\limits_i x_i^2 p_i & : \ \underline{x} \text{ diskret,} \\ \int\limits_{-\infty}^{\infty} t^2 f_{\underline{x}}(t)\,\mathrm{d}t & : \ \underline{x} \text{ absolutstetig.} \end{cases}$$

**Beweis:** $\mathcal{V}(\underline{x}) = \mathcal{E}\Big(\big(\underline{x} - \mathcal{E}(\underline{x})\big)^2\Big)$;

$\big(\underline{x} - \mathcal{E}(\underline{x})\big)^2 = \underline{x}^2 - 2\mathcal{E}(\underline{x}) \cdot \underline{x} + \overset{\text{konstante ZG}}{\mathcal{E}(\underline{x})^2} \quad \Longrightarrow$ (Sätze 0,1)

$\Longrightarrow \mathcal{V}(\underline{x}) = \mathcal{E}(\underline{x}^2) - 2\mathcal{E}(\underline{x}) \cdot \mathcal{E}(\underline{x}) + \mathcal{E}(\underline{x})^2 = \mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2.$

Die Formel für $\mathcal{E}(\underline{x}^2)$ folgt aus Satz 2. □

**Bsp. 3:** $P_{\underline{x}} = \mathrm{Bin}\,(n,p),\ x_i = i\ (i=0,\cdots,n),\ p_i = \binom{n}{i} p^i q^{n-i} \implies$
$\implies \mathcal{E}(\underline{x}^2) = \sum\limits_{i=0}^{n} i^2 \binom{n}{i} p^i q^{n-i}$

**Berechnung von $\boldsymbol{\mathcal{E}(\underline{x}^2)}$ mit Differenzieren:**

Nach Seite 36: $\sum\limits_{i=0}^{n} i\binom{n}{i} x^i y^{n-i} = xn(x+y)^{n-1} \quad \Big| \dfrac{\partial}{\partial x}$

$$\implies \sum_{i=0}^{n} i^2 \binom{n}{i} x^{i-1} y^{n-i} = n(x+y)^{n-1} + xn(n-1)(x+y)^{n-2}$$

$$\begin{aligned}
\implies \mathcal{E}(\underline{x}^2) &= p\big[n\underbrace{(p+q)^{n-1}}_{1} + pn(n-1)\underbrace{(p+q)^{n-2}}_{1}\big] \\
&= pn[1+pn-p] = np[q+np] = npq + (np)^2 \\
\implies \mathcal{V}(\underline{x}) &= \mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2 = npq + (np)^2 - (np)^2 = npq \\
\implies \sigma(\underline{x}) &= \sqrt{npq}.
\end{aligned}$$

Somit: $\boxed{P_{\underline{x}} = \mathrm{Bin}\,(n,p) \implies \mathcal{V}(\underline{x}) = npq}$

**Bsp. 4:** $P_{\underline{x}} = N(\mu,\sigma),\ f_{\underline{x}}(t) = \dfrac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}$

**Berechnung von $\boldsymbol{\mathcal{E}(\underline{x})}$, $\boldsymbol{\mathcal{V}(\underline{x})}$ mit Differenzieren:**

Nach Seite 23: $\dfrac{1}{\sqrt{2\pi}\,\sigma}\displaystyle\int_{-\infty}^{\infty} \mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t = P(\mathbb{R}) = 1 \qquad \Big| \dfrac{\partial}{\partial\mu}$

$$\implies \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} -\frac{\cancel{2}(\mu-t)}{\cancel{2}\sigma^2}\,\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t = 0 \qquad \Big|\cdot\sigma^2$$

$$\implies \mathcal{E}(\underline{x}) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} t\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t = \mu\cdot\underbrace{\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t}_{1} = \mu \qquad \Big|\frac{\partial}{\partial\mu}$$

$$\implies \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} t\cdot\frac{-\cancel{2}(\mu-t)}{\cancel{2}\sigma^2}\,\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t = 1 \qquad \Big|\cdot\sigma^2$$

$$\implies \mathcal{E}(\underline{x}^2) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} t^2\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t = \sigma^2 + \mu\cdot\underbrace{\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} t\mathrm{e}^{-(\mu-t)^2/(2\sigma^2)}\,\mathrm{d}t}_{\mu}$$

$\Longrightarrow\ \mathcal{V}(\underline{x}) = \mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$
$\Longrightarrow\ \sigma(\underline{x}) = \sigma$

Die Normalverteilung $N(\mu,\sigma)$ hat also Erwartungswert $\mu$ und Streuung $\sigma$.

Unter der Voraussetzung, dass Bin $(n,p)$ für $p$ fest, $n \to \infty$ durch $N(\mu,\sigma)$ approximiert wird, ist mit Bsp. 3/4 verständlich, warum man $\mu = np$ und $\sigma = \sqrt{npq}$ wählen muss, siehe Seite 24.

**Bemerkung:** $\mathcal{E}(\underline{x})$, $\mathcal{V}(\underline{x})$ hängen nur von $P_{\underline{x}}$, nicht von $\underline{x}$ selber ab. Daher ist die Redeweise „Erwartungswert/Varianz der $\cdots$ **Verteilung**“ gerechtfertigt.

**Tabelle:**

| | $P_{\underline{x}}$ | $P_{\underline{x}}\big(\{i\}\big)$ bzw. $f_{\underline{x}}(t)$ | $\mathcal{E}(\underline{x})$ | $\mathcal{V}(\underline{x})$ | $\sigma(\underline{x}) = \sqrt{\mathcal{V}(\underline{x})}$ |
|---|---|---|---|---|---|
| diskret | Equ $(n)$ | $\frac{1}{n}\ (i = 1, \cdots, n)$ | $\frac{n+1}{2}$ | $\frac{n^2-1}{12}$ | $\frac{\sqrt{n^2-1}}{2\sqrt{3}}$ |
| | Bin $(n,p)$ | $\binom{n}{i} p^i q^{n-i}\ (i = 0, \cdots, n)$ | $np$ | $npq$ | $\sqrt{npq}$ |
| | Hyp $(n,N,F)$ | $\dfrac{\binom{F}{i}\binom{N-F}{n-i}}{\binom{N}{n}}\ (i = 0, \cdots, n)$ | $n\frac{F}{N}$ | $\frac{N-n}{N-1} n \frac{F}{N} \frac{N-F}{N}$ | $\sqrt{\frac{N-n}{N-1} n \frac{F}{N} \frac{N-F}{N}}$ |
| | Poi $(\lambda)$ | $\dfrac{\lambda^i}{i!} \mathrm{e}^{-\lambda}\ (i = 0, 1, \cdots)$ | $\lambda$ | $\lambda$ | $\sqrt{\lambda}$ |
| absolutstetig | Equ $(a,b)$ | $\frac{1}{b-a}\ (a < t < b)$ | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ | $\frac{b-a}{2\sqrt{3}}$ |
| | Expo $(\mu)$ | $\mu \mathrm{e}^{-\mu t}\ (t > 0)$ | $\frac{1}{\mu}$ | $\frac{1}{\mu^2}$ | $\frac{1}{\mu}$ |
| | $N(\mu,\sigma)$ | $\frac{1}{\sqrt{2\pi}\,\sigma}\ \mathrm{e}^{-(t-\mu)^2/(2\sigma^2)}$ | $\mu$ | $\sigma^2$ | $\sigma$ |
| | $t -$ Vert $(\nu)$ | $\frac{\Gamma\big((\nu+1)/2\big)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)} \cdot \big(1 + \frac{t^2}{\nu}\big)^{-(\nu+1)/2}\ (\nu \geq 2)$ | $0$ | $\frac{\nu}{\nu-2}$ | $\sqrt{\frac{\nu}{\nu-2}}$ |
| | $\chi^2 -$ Vert$(\nu)$ | $\dfrac{t^{-1+\nu/2} \mathrm{e}^{-t/2}}{2^{\nu/2}\Gamma(\nu/2)}\ (t > 0)$ | $\nu$ | $2\nu$ | $\sqrt{2\nu}$ |
| | $F -$ Vert$(\nu_1,\nu_2)$ | $\dfrac{\nu_1^{\nu_1/2}\nu_2^{\nu_2/2}\Gamma\big(\frac{\nu_1+\nu_2}{2}\big) t^{\nu_1/2-1}}{\Gamma\big(\frac{\nu_1}{2}\big)\Gamma\big(\frac{\nu_2}{2}\big)(\nu_2 + \nu_1 t)^{(\nu_1+\nu_2)/2}}\ (t > 0)$ | $\frac{\nu_2}{\nu_2-2}$ | $\frac{2(\nu_1+\nu_2-2)\nu_2^2}{\nu_1(\nu_2-2)^2(\nu_2-4)}$ | $\sqrt{\frac{2(\nu_1+\nu_2-2)\nu_2^2}{\nu_1(\nu_2-2)^2(\nu_2-4)}}$ |

# § 7 Korrelation

$\underline{x}, \underline{y} : \Omega \longrightarrow \mathbb{R}$ seien Zufallsgrößen.
**Problem 1:** Welche Eigenschaften hat $\mathcal{V}$? Was ist $\mathcal{V}(\underline{x}+\underline{y})$?
**Problem 2:** Was lässt sich aus der Kenntnis von $\underline{x}$ über $\underline{y}$ sagen?
Wie wir sehen werden, hängen diese Probleme zusammen.

**Zu Problem 1:**

a) $$\mathcal{V}(\underline{x}) = \mathcal{E}\Big(\big(\underline{x}-\mathcal{E}(\underline{x})\big)^2\Big) = \left\{ \begin{array}{l} \sum\limits_i \big(x_i - \mathcal{E}(\underline{x})\big)^2 p_i \\ \int \big(t-\mathcal{E}(\underline{x})\big)^2 f_{\underline{x}}(t)\,\mathrm{d}t \end{array} \right\} \geq 0$$

b) $$\mathcal{V}(c\underline{x}) = \mathcal{E}\big((c\underline{x})^2\big) - \mathcal{E}(c\underline{x})^2 = c^2\big[\mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2\big] = \underline{\underline{\underline{c^2}}}\mathcal{V}(\underline{x})$$

c) $$\begin{aligned} \mathcal{V}(\underline{x}+\underline{y}) &= \mathcal{E}\big((\underline{x}+\underline{y})^2\big) - \mathcal{E}(\underline{x}+\underline{y})^2 \\ &= \mathcal{E}(\underline{x}^2 + 2\underline{x}\underline{y} + \underline{y}^2) - \big(\mathcal{E}(\underline{x}) + \mathcal{E}(\underline{y})\big)^2 \\ &= \mathcal{E}(\underline{x}^2) + 2\mathcal{E}(\underline{x}\underline{y}) + \mathcal{E}(\underline{y}^2) - \big[\mathcal{E}(\underline{x})^2 + 2\mathcal{E}(\underline{x})\mathcal{E}(\underline{y}) + \mathcal{E}(\underline{y})^2\big] \\ &= \underbrace{\mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2}_{\mathcal{V}(\underline{x})} + \underbrace{\mathcal{E}(\underline{y}^2) - \mathcal{E}(\underline{y})^2}_{\mathcal{V}(\underline{y})} + 2\big[\underbrace{\mathcal{E}(\underline{x}\underline{y}) - \mathcal{E}(\underline{x})\mathcal{E}(\underline{y})}_{\mathrm{cov}\,(\underline{x},\underline{y})}\big] \end{aligned}$$

**Def.:** $\underline{x}, \underline{y} : \Omega \longrightarrow \mathbb{R}$.
Dann heißt $\mathcal{E}(\underline{x}\cdot\underline{y}) - \mathcal{E}(\underline{x})\cdot\mathcal{E}(\underline{y})$ **Kovarianz** von $\underline{x}$ und $\underline{y}$, kurz $\mathrm{cov}\,(\underline{x},\underline{y})$.

Somit gilt: $$\boxed{\mathcal{V}(\underline{x}+\underline{y}) = \mathcal{V}(\underline{x}) + \mathcal{V}(\underline{y}) + 2\mathrm{cov}\,(\underline{x},\underline{y})}$$

Weiters gilt:

d) $\mathrm{cov}\,(\underline{x},\underline{x}) = \mathcal{E}(\underline{x}\cdot\underline{x}) - \mathcal{E}(\underline{x})\cdot\mathcal{E}(\underline{x}) = \mathcal{V}(\underline{x})$

e) $\mathrm{cov}(\underline{x},\underline{y}) = \mathcal{E}\Big[\big(\underline{x}-\mathcal{E}(\underline{x})\big)\cdot\big(\underline{y}-\mathcal{E}(\underline{y})\big)\Big]$, denn das letztere ist
$\mathcal{E}\big(\underline{x}\cdot\underline{y} - \mathcal{E}(\underline{x})\cdot\underline{y} - \mathcal{E}(\underline{y})\cdot\underline{x} + \mathcal{E}(\underline{x})\cdot\mathcal{E}(\underline{y})\big) = \mathcal{E}(\underline{x}\cdot\underline{y}) - \mathcal{E}(\underline{x})\cdot\mathcal{E}(\underline{y}) - \cancel{\mathcal{E}(\underline{y})\cdot\mathcal{E}(\underline{x})} + \cancel{\mathcal{E}(\underline{x})\cdot\mathcal{E}(\underline{y})} =$
$\mathrm{cov}\,(\underline{x},\underline{y})$

**Zu Problem 2:** $\underline{x}, \underline{y} : \Omega \longrightarrow \mathbb{R}$ sind gegeben. Bestimme die Gerade $y = kx + d$ im $\mathbb{R}^2$ so, dass die Zufallsvariable $\underline{y} - k\underline{x} - d$ im Durchschnitt möglichst klein ist.
Wenn dann $\underline{x}$ den Wert $w$ hat, so ist $kw + d$ **die beste lineare Vorhersage** für den Wert von $\underline{y}$.

**Def.:** $\underline{x}, \underline{y} : \Omega \longrightarrow \mathbb{R}$. Die Gerade $y = kx + d$ im $\mathbb{R}^2$ für die $\mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big) = \min$ heißt **Regressionsgerade** (=RG) von $\underline{y}$ nach $\underline{x}$.

**Bestimmung der Regressionsgeraden:**

$$\begin{aligned}
\mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big) &= \mathcal{E}(\underline{y}^2 + k^2\underline{x}^2 + d^2 - 2k\underline{x}\underline{y} - 2d\underline{y} + 2kd\underline{x}) \\
&= \mathcal{E}(\underline{y}^2) + k^2\mathcal{E}(\underline{x}^2) + d^2 - 2k\mathcal{E}(\underline{x} \cdot \underline{y}) - 2d\mathcal{E}(\underline{y}) + 2kd\mathcal{E}(\underline{x}) \\
\frac{\partial}{\partial k}\Big(\mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big)\Big) = 0 &\implies 2k\mathcal{E}(\underline{x}^2) - 2\mathcal{E}(\underline{x} \cdot \underline{y}) + 2d\mathcal{E}(\underline{x}) = 0 \\
\frac{\partial}{\partial d}\Big(\mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big)\Big) = 0 &\implies 2d - 2\mathcal{E}(\underline{y}) + 2k\mathcal{E}(\underline{x}) = 0
\end{aligned}$$

$$\left.\begin{aligned}
\implies \mathrm{I} &: k\mathcal{E}(\underline{x}^2) + d\mathcal{E}(\underline{x}) = \mathcal{E}(\underline{x} \cdot \underline{y}) \\
\mathrm{II} &: k\mathcal{E}(\underline{x}) + d = \mathcal{E}(\underline{y}) \quad \big| \cdot \mathcal{E}(\underline{x})
\end{aligned}\right\} -$$

$$\implies k\big(\underbrace{\mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x})^2}_{\mathcal{V}(\underline{x})}\big) = \underbrace{\mathcal{E}(\underline{x}\underline{y}) - \mathcal{E}(\underline{x}) \cdot \mathcal{E}(\underline{y})}_{\operatorname{cov}(\underline{x},\underline{y})}$$

$$\implies \boxed{k = \frac{\operatorname{cov}(\underline{x}, \underline{y})}{\mathcal{V}(\underline{x})}}$$

(Für $\mathcal{V}(\underline{x}) = 0$ ist $\underline{x}$ = konstant (vgl. Seite 42, 3) und $k$ unbestimmt. Die RG wird definiert als $y = \mathcal{E}(\underline{y})$.)

Die Gleichung II sagt, dass der „Schwerpunkt“ $\big(\mathcal{E}(\underline{x}), \mathcal{E}(\underline{y})\big)$ auf der RG $y = kx + d$ liegt. Daraus wird $d$ berechnet.
Falls $\operatorname{cov}(\underline{x}, \underline{y}) = 0$, so ist $k = 0,\ d = \mathcal{E}(\underline{y})$, und die Regressionsgerade von $\underline{y}$ nach $\underline{x}$ ist $y = \mathcal{E}(\underline{y})$, d.h. die Kenntnis von $\underline{x}$ nützt nichts für eine lineare Voraussage von $\underline{y}$.

**Def.:** $\underline{x}, \underline{y}$ heißen **unkorreliert** $\iff \operatorname{cov}(\underline{x}, \underline{y}) = 0$.

Aus der Definition ergibt sich unmittelbar der folgende

**Satz:** Äquivalent sind:

1) $\underline{x}, \underline{y}$ unkorreliert;
2) $\operatorname{cov}(\underline{x}, \underline{y}) = 0$;
3) $\mathcal{E}(\underline{x} \cdot \underline{y}) = \mathcal{E}(\underline{x}) \cdot \mathcal{E}(\underline{y})$;
4) $\mathcal{V}(\underline{x} + \underline{y}) = \mathcal{V}(\underline{x}) + \mathcal{V}(\underline{y})$;
5) Die Regressionsgerade von $\underline{y}$ nach $\underline{x}$ ist $y = \mathcal{E}(\underline{y})$ (d.h. $\parallel$ $x$-Achse)
6) Die Regressionsgerade von $\underline{x}$ nach $\underline{y}$ ist $x = \mathcal{E}(\underline{x})$ (d.h. $\parallel$ $y$-Achse)

**Problem 3:** Quantifiziere (=Beziffere) den Zusammenhang von $\underline{x}$ und $\underline{y}$ durch eine Zahl $\varrho \in [-1,1]$, sodass $\varrho = 0 \iff \underline{x}, \underline{y}$ unkorreliert und $\varrho = \pm 1 \iff \underline{y} = k\underline{x} + d$, d.h. $\underline{y}$ ist linear durch $\underline{x}$ gegeben.

**Lösung:** $k, d$ seien die Lösungen von I,II in Seite 41

$$\Longrightarrow 0 \underset{\substack{\uparrow \\ \text{(vgl. p. 40 a))}}}{\leq} \mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big) \overset{\text{(vgl. p. 41)}}{=}$$

$$= k\,\big[\underbrace{k\mathcal{E}(\underline{x}^2) - \mathcal{E}(\underline{x}\cdot\underline{y}) + d\mathcal{E}(\underline{x})}_{\text{0 wegen I}}\big] + d\,\big[\underbrace{d - \mathcal{E}(\underline{y}) + k\mathcal{E}(\underline{x})}_{\text{0 wegen II}}\big] + \mathcal{E}(\underline{y}^2) - k\mathcal{E}(\underline{x}\cdot\underline{y}) - d\mathcal{E}(\underline{y}) =$$

$$\overset{\text{(II)}}{=} \mathcal{E}(\underline{y}^2) - k\mathcal{E}(\underline{x}\cdot\underline{y}) - \big[\mathcal{E}(\underline{y}) - k\mathcal{E}(\underline{x})\big]\cdot\mathcal{E}(\underline{y})$$

$$= \mathcal{V}(\underline{y}) - k\,\mathrm{cov}\,(\underline{x},\underline{y}) = \mathcal{V}(\underline{y}) - \frac{\mathrm{cov}^2(\underline{x},\underline{y})}{\mathcal{V}(\underline{x})}$$

$$\Longrightarrow 0 \leq \frac{\mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big)}{\mathcal{V}(\underline{y})} = 1 - \left(\frac{\mathrm{cov}\,(\underline{x},\underline{y})}{\sigma(\underline{x})\sigma(\underline{y})}\right)^2$$

**Def.:** $\boxed{\varrho(\underline{x},\underline{y}) = \dfrac{\mathrm{cov}\,(\underline{x},\underline{y})}{\sigma(\underline{x})\sigma(\underline{y})}}$ heißt **Korrelationskoeffizient** von $\underline{x}$ und $\underline{y}$.

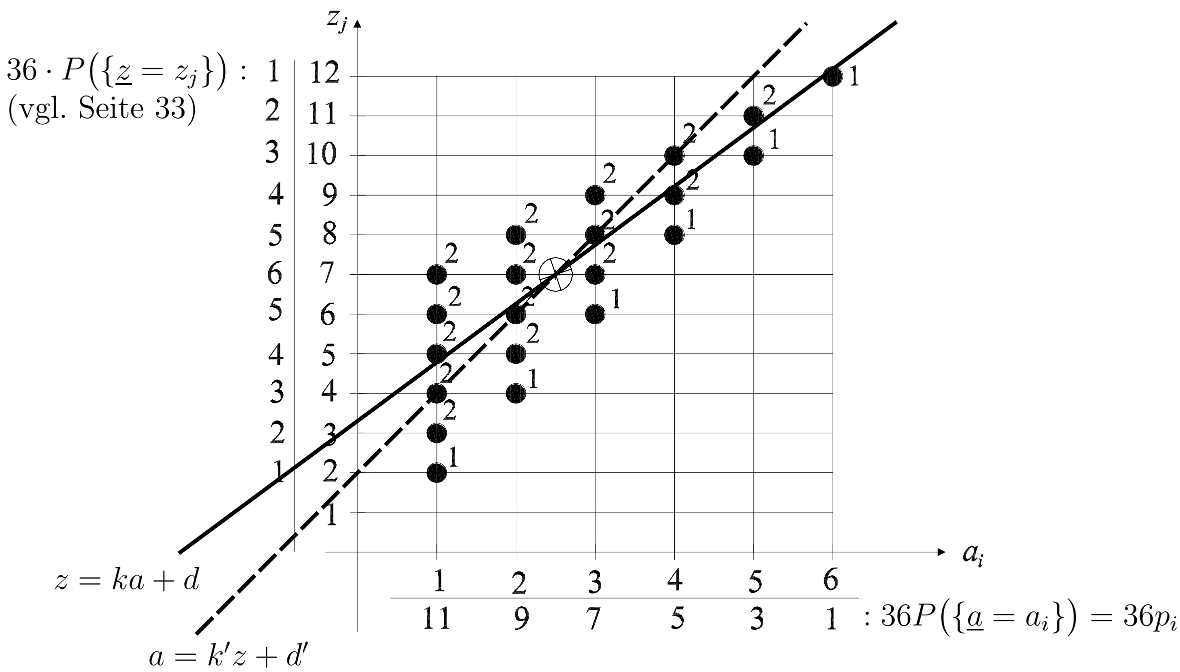
**Satz:** 1) $-1 \leq \varrho(\underline{x},\underline{y}) \leq 1$; $\mathrm{sign}\,\varrho = \mathrm{sign}\,k$
2) $\varrho(\underline{x},\underline{y}) = 0 \iff \underline{x}, \underline{y}$ unkorreliert;
3) $\varrho(\underline{x},\underline{y}) = \pm 1 \iff \underline{y} = k\underline{x} + d$.

**Beweis:** 1) folgt aus $0 \leq 1 - \varrho(\underline{x},\underline{y})^2$ und wegen $k = \dfrac{\mathrm{cov}\,(\underline{x},\underline{y})}{\mathcal{V}(\underline{x})} = \varrho(\underline{x},\underline{y}) \cdot \dfrac{\sigma(\underline{y})}{\sigma(\underline{x})}$.
2) $\varrho(\underline{x},\underline{y}) = 0 \iff \mathrm{cov}\,(\underline{x},\underline{y}) = 0$
3) $\varrho(\underline{x},\underline{y}) = \pm 1 \iff \mathcal{E}\big((\underline{y} - k\underline{x} - d)^2\big) = 0$; der Erwartungswert einer positiven Zufallsgröße $\underline{u}$ kann nur dann 0 sein, wenn $\underline{u} = 0$; denn

$$\mathcal{E}(\underline{u}) = \begin{cases} \sum\limits_i \overbrace{u_i}^{\geq 0}\overbrace{p_i}^{>0} = 0 \Longrightarrow & \text{nur } i = 1 \text{ und } u_1 = 0 \\ \int\limits_0^\infty f_{\underline{u}}(t)t\,\mathrm{d}t = 0 \text{ unmöglich, da } \int\limits_0^\infty f_{\underline{u}}(t)\,\mathrm{d}t = 1 \end{cases}$$

$\square$

**Bsp. 1:** Würfeln mit 2 Würfeln, $\underline{x} =$ „Augenzahl 1. Würfel“, $\underline{y} =$ „Augenzahl 2. Würfel“, $\underline{z} = \underline{x} + \underline{y}$, $\underline{a} = \min(\underline{x}, \underline{y})$. Wie sind $\underline{a}$ und $\underline{z}$ korreliert?




1 bzw. 2 = Anzahl der Elementarereignisse bei denen $\underline{a} = a_i$, $\underline{z} = z_j$

$\otimes$ = Schwerpunkt = $\big(\mathcal{E}(\underline{a}), \mathcal{E}(\underline{z})\big) \approx (2.53, 7)$

$$\begin{aligned}
\mathcal{E}(\underline{z}) &= \mathcal{E}(\underline{x}) + \mathcal{E}(\underline{y}) = 2 \cdot \tfrac{7}{2} = 7,\ \mathcal{V}(\underline{z}) = \tfrac{35}{6} \text{ (vgl. Übung 27)},\ \sigma(\underline{z}) \approx 2.4152 \\
\mathcal{E}(\underline{a}) &= \textstyle\sum a_i p_i = 1 \cdot \tfrac{11}{36} + 2 \cdot \tfrac{9}{36} + \cdots = \tfrac{91}{36} \approx 2.53 \\
\mathcal{E}(\underline{a}^2) &= \textstyle\sum a_i^2 p_i = 1 \cdot \tfrac{11}{36} + 4 \cdot \tfrac{9}{36} + \cdots = \tfrac{301}{36},\ \mathcal{V}(\underline{a}) = \mathcal{E}(\underline{a}^2) - \mathcal{E}(\underline{a})^2 = \tfrac{2555}{36^2},\ \sigma(\underline{a}) \approx 1.4041 \\
\mathcal{E}(\underline{a} \cdot \underline{z}) &= \sum_{i,j} a_i z_j \cdot P\big(\{\underline{a} = a_i \wedge \underline{z} = z_j\}\big) = 1 \cdot 2 \cdot \tfrac{1}{36} + 1 \cdot 3 \cdot \tfrac{2}{36} + 1 \cdot 4 \cdot \tfrac{2}{36} + \cdots + 6 \cdot 12 \cdot \tfrac{1}{36} \\
&= \tfrac{742}{36} \Longrightarrow \operatorname{cov}(\underline{a}, \underline{z}) = \mathcal{E}(\underline{a} \cdot \underline{z}) - \mathcal{E}(\underline{a}) \cdot \mathcal{E}(\underline{z}) = \tfrac{35}{12} \Longrightarrow \varrho(\underline{a}, \underline{z}) = \tfrac{\operatorname{cov}(\underline{a},\underline{z})}{\sigma(\underline{a})\sigma(\underline{z})} \\
&= \sqrt{\tfrac{54}{73}} \approx 0.86 \qquad \text{(d.h. } \underline{a}, \underline{z} \text{ sind stark korreliert)}
\end{aligned}$$

**Regressionsgerade von $\underline{z}$ nach $\underline{a}$ :** $k = \dfrac{\operatorname{cov}(\underline{a}, \underline{z})}{\mathcal{V}(\underline{a})} = \dfrac{108}{73} \approx 1.48,$

$d$ ist so, dass $\mathcal{E}(\underline{z}) = k\mathcal{E}(\underline{a}) + d \implies d = \dfrac{238}{73} \approx 3.26$

Also: $z = ka + d \approx 1.48a + 3.26$

**Regressionsgerade von $\underline{a}$ nach $\underline{z}$ :** $k' = \dfrac{\operatorname{cov}(\underline{a}, \underline{z})}{\mathcal{V}(\underline{z})} = \dfrac{1}{2},$

$d'$ ist so, dass $\mathcal{E}(\underline{a}) = k'\mathcal{E}(\underline{z}) + d' \implies d' = -\dfrac{35}{36}$

Also: $a = k'z + d' = \dfrac{1}{2}z - \dfrac{35}{36}$

**Warum unterscheiden sich die zwei Regressionsgeraden?**

$z = ka + d$ liefert uns die beste lineare Schätzung von $z$, wenn $a$ bekannt ist.
$a = k'z + d'$ liefert uns hingegen die beste lineare Schätzung von $a$, wenn $z$ bekannt ist.
Dies ist ein ähnliches Verhältnis wie zwischen $P(A|B)$ und $P(B|A)$.

**Zusammenhang Regressionsgerade/Ausgleichsgerade:**

Wenn Daten $(x_1, y_1), \cdots, (x_n, y_n)$ gegeben sind, so nennt man $y = kx + d$ **Ausgleichsgerade** dazu, wenn $\sum\limits_{i=1}^{n}(y_i - kx_i - d)^2 = \min$. D.h. wenn $\Omega_n = \{1, \cdots, n\}$ symmetrisch,

$$\underline{X} : \Omega_n \longrightarrow \mathbb{R} : i \longmapsto x_i,\ \underline{Y} : \Omega_n \longrightarrow \mathbb{R} : i \longmapsto y_i$$
$$\Longrightarrow\ \mathcal{E}\big((\underline{Y} - k\underline{X} - d)^2\big) = \frac{1}{n}\sum_{i=1}^{n}(y_i - kx_i - d)^2 = \min.$$

Die Ausgleichsgerade ist also ein Spezialfall der Regressionsgeraden und es können dieselben Formeln verwendet werden.
Wenn $(x_i, y_i)$ Messwerte von Zufallsgrößen $\underline{x}, \underline{y}$ sind, so ist das Verhältnis Ausgleichsgerade/Regressionsgerade so wie Mittelwert/Erwartungswert.

**Bsp. 2:** 55 Messversuche mit Betonproben nach 28 Tagen ergaben:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ (Dichte) | 2.48 | 2.48 | 2.48 | 2.48 | 2.49 | 2.47 | 2.4 | 2.49 | 2.48 | 2.45 |
| $y_i$ (Druckfestigkeit) | 460 | 406 | 363 | 353 | 387 | 398 | 298 | 388 | 377 | 358 |

| $i$ | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 2.47 | 2.48 | 2.48 | 2.47 | 2.51 | 2.41 | 2.4 | 2.5 | 2.42 | 2.52 |
| $y_i$ | 396 | 419 | 433 | 385 | 360 | 340 | 312 | 397 | 325 | 396 |

| $i$ | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 2.42 | 2.45 | 2.42 | 2.42 | 2.44 | 2.43 | 2.46 | 2.46 | 2.37 | 2.37 |
| $y_i$ | 333 | 383 | 313 | 325 | 362 | 382 | 384 | 366 | 338 | 340 |

| $i$ | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 2.46 | 2.43 | 2.39 | 2.37 | 2.45 | 2.42 | 2.42 | 2.44 | 2.45 | 2.47 |
| $y_i$ | 393 | 318 | 303 | 263 | 372 | 339 | 320 | 324 | 339 | 370 |

| $i$ | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 2.46 | 2.43 | 2.44 | 2.44 | 2.45 | 2.45 | 2.45 | 2.41 | 2.43 | 2.43 | 2.45 | 2.43 | 2.43 | 2.46 | 2.37 |
| $y_i$ | 368 | 320 | 345 | 353 | 385 | 357 | 383 | 371 | 356 | 381 | 358 | 328 | 332 | 323 | 235 |

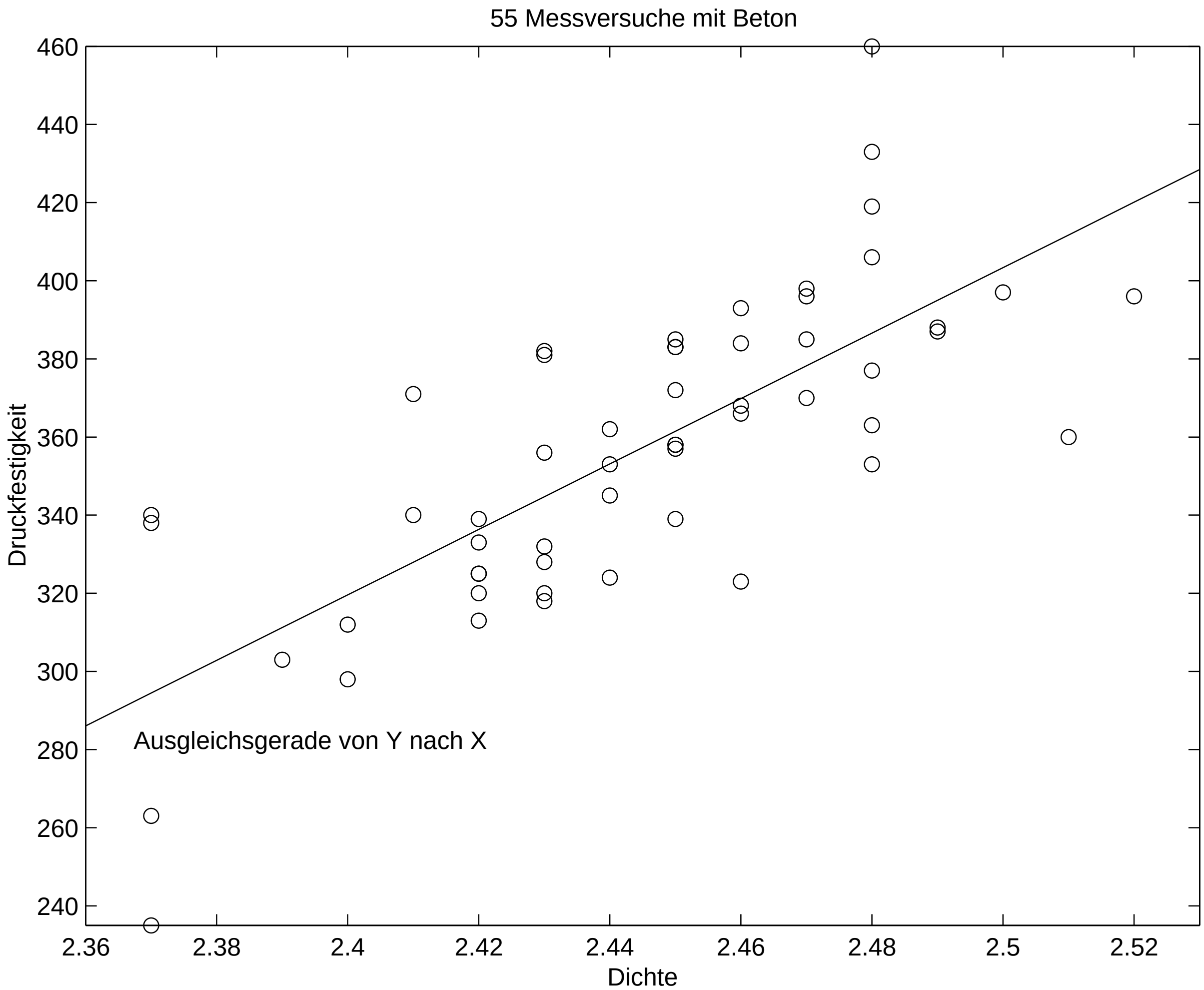
$$\begin{aligned}
\mathcal{E}(\underline{X}) &= \text{ Mittelwert der } x_i = \frac{1}{55}[2.48 + 2.48 + \cdots + 2.37] \approx 2.44418, \\
\mathcal{E}(\underline{X}^2) &= \frac{1}{55}[2.48^2 + 2.48^2 + \cdots + 2.37^2] \approx 5.97524, \\
\mathcal{V}(\underline{X}) &= \mathcal{E}(\underline{X}^2) - \mathcal{E}(\underline{X})^2 \approx 0.00121342, \ \sigma(\underline{X}) \approx 0.0348, \\
\mathcal{E}(\underline{Y}) &= \text{ Mittelwert der } y_i = \frac{1}{55}[460 + 406 + \cdots + 235] = 356.6, \\
\mathcal{V}(\underline{Y}) &\approx 1539.1127, \ \sigma(\underline{Y}) = \sqrt{\mathcal{V}(\underline{Y})} \approx 39.2315, \\
\mathcal{E}(\underline{X} \cdot \underline{Y}) &= \frac{1}{55}[2.48 \cdot 460 + \cdots + 2.37 \cdot 235] \approx 872.6116, \\
\operatorname{cov}(\underline{X}, \underline{Y}) &= \mathcal{E}(\underline{X} \cdot \underline{Y}) - \mathcal{E}(\underline{X}) \cdot \mathcal{E}(\underline{Y}) = 1.0164, \\
\varrho(\underline{X}, \underline{Y}) &= \frac{\operatorname{cov}(\underline{X}, \underline{Y})}{\sigma(\underline{X})\sigma(\underline{Y})} \approx 0.743744.
\end{aligned}$$

Ausgleichsgerade = RG von $\underline{Y}$ nach $\underline{X} : y = kx + d$,

$$k = \frac{\operatorname{cov}(\underline{X}, \underline{Y})}{\mathcal{V}(\underline{X})} \approx 837.6314$$

$d$ so, dass $\mathcal{E}(\underline{Y}) = k\mathcal{E}(\underline{X}) + d \implies d \approx -1690.72356$

Bild:



55 Messversuche mit Beton
Druckfestigkeit
460
440
420
400
380
360
340
320
300
280
260
240
Ausgleichsgerade von Y nach X
2.36
2.38
2.4
2.42
2.44
2.46
2.48
2.5
2.52
Dichte

## Zusatz zu § 7:

Im Bild auf Seite 43 werden die zwei Zufallsvariablen $\underline{a}, \underline{z}$ sozusagen „zugleich" betrachtet. $(\underline{a}, \underline{z})$ ist ein Zufallsvektor.

**Def.:** $\Omega$ sei ein Wahrscheinlichkeitsraum. Ein **Zufallsvektor** ist eine Abbildung $\underline{\vec{v}} : \Omega \longrightarrow \mathbb{R}^n$.

**Experimentelle Bedeutung:** Der Zufallsvektor $\underline{\vec{v}}$ ordnet jedem Elementarereignis $\omega \in \Omega$ die $n$ reellen Zahlen $\underline{v}_1(\omega), \cdots, \underline{v}_n(\omega)$ zu und entspricht also $n$ Zufallsvariablen am selben Wahrscheinlichkeitsraum.

**Def.:** Wenn alle $\underline{v}_k$ diskret sind und nur die Werte $x_1, \cdots$ annehmen, so heißt $\underline{\vec{v}}$ **diskret** und man setzt $p_{i_1 \cdots i_n} = P\big(\{\underline{v}_1 = x_{i_1} \wedge \cdots \wedge \underline{v}_n = x_{i_n}\}\big)$. $\underline{\vec{v}}$ heißt **absolutstetig** $\Longleftrightarrow P_{\underline{\vec{v}}}$ hat eine Dichtefunktion $f_{\underline{\vec{v}}}$, d.h.

$$P\Big(\big\{\underline{v}_1 \in [a_1, b_1] \wedge \cdots \wedge \underline{v}_n \in [a_n, b_n]\big\}\Big) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f_{\underline{\vec{v}}}(t_1, \cdots, t_n)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

**Bemerkung:** $\mathbb{R}^n$ ist mit $P_{\underline{\vec{v}}}(A) = P\big(\{\underline{\vec{v}} \in A\}\big)$ (für $A \subset \mathbb{R}^n$ messbar) ein Wahrscheinlichkeitsraum und $P_{\underline{\vec{v}}}$ heißt die **gemeinsame Verteilung** von $\underline{v}_1, \cdots, \underline{v}_n$.

**Satz:** Es seien der Zufallsvektor $\underline{\vec{v}}$ und $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ gegeben. Dann gilt

$$\mathcal{E}\big(h(\underline{v}_1, \cdots, \underline{v}_n)\big) = \begin{cases} \sum h(x_{i_1}, \cdots, x_{i_n}) p_{i_1 \cdots i_n} : \underline{\vec{v}} \text{ diskret,} \\ \int h(t_1, \cdots, t_n) f_{\underline{\vec{v}}}(t_1, \cdots, t_n)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n : \vec{v} \text{ absolutstetig} \end{cases}$$

(Vgl. § 6, Satz 2, Seite 37)

**Bsp. 1:** Wenn $\underline{\vec{v}} = (\underline{a}, \underline{z})$, d.h. $\underline{v}_1 = \underline{a}$, $\underline{v}_2 = \underline{z}$ entsprechend Seite 42/43, so ist $p_{ij} = \frac{1}{36}$ für $(i, j) = (1, 2), (2, 4), (3, 6), (4, 8), (5, 10), (6, 12)$ und $p_{ij} = \frac{2}{36}$ für $(i, j) = (1, 3), \cdots, (1, 7), (2, 5), \cdots, (2, 8), (3, 7), (3, 8), (3, 9), (4, 9), (4, 10), (5, 11)$ und $p_{ij} = 0$ sonst

$$\begin{aligned} \Longrightarrow \mathcal{E}(\underline{a} \cdot \underline{z}) &= \sum a_i z_j P\big(\{\underline{a} = a_i \wedge \underline{z} = z_j\}\big) \\ &= \sum_{i,j} ij p_{ij} = 1 \cdot 2 \cdot \tfrac{1}{36} + 1 \cdot 3 \cdot \tfrac{2}{36} + \cdots = \tfrac{742}{36}. \end{aligned}$$

Hier ist $h(x_1, x_2) = x_1 \cdot x_2,\ n = 2,\ i_1 = i,\ i_2 = j,\ x_{i_1} = i,\ x_{i_2} = j$ im Satz oben.

**Bsp. 3:** Der Zufallsvektor $\underline{\vec{v}} = (\underline{v}_1, \cdots, \underline{v}_n)$ heißt **normalverteilt**, wenn seine Dichte durch

$$f_{\underline{\vec{v}}}(t_1, \cdots, t_n) = \sqrt{\frac{\det A}{\pi^n}} \cdot \exp\left(-\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(t_i - \mu_i)(t_j - \mu_j)\right)$$

gegeben ist, wobei $\exp(s) = \mathrm{e}^s$ für $s \in \mathbb{R}$ und $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ eine positiv definite symmetrische Matrix ist und $\mu_1, \cdots, \mu_n \in \mathbb{R}$.

Dann gilt $\int\limits_{\mathbb{R}^n} f_{\vec{v}}(t_1, \cdots, t_n)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n = 1$, denn

$$\begin{aligned}
&\int\limits_{\mathbb{R}^n} \exp\big(-(\vec{t} - \vec{\mu})^T A (\vec{t} - \vec{\mu})\big)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n = \ (\text{Substitution } \vec{s} = \vec{t} - \vec{\mu}) \\
= &\int\limits_{\mathbb{R}^n} \exp(-\vec{s}^T A \vec{s})\, \mathrm{d}s_1 \cdots \mathrm{d}s_n \\
= &\ (\text{HA-Transformation } \vec{y} = T^{-1}\vec{s};\ \left|\frac{\partial(y_1, \cdots, y_n)}{\partial(s_1, \cdots, s_n)}\right| = |\det T^{-1}| = 1,\ \text{da } T \text{ orthogonal ist}) \\
= &\int\limits_{\mathbb{R}^n} \exp(-\lambda_1 y_1^2 - \cdots - \lambda_n y_n^2)\, \mathrm{d}y_1 \cdots \mathrm{d}y_n \\
= &\int\limits_{-\infty}^{\infty} \mathrm{e}^{-\lambda_1 y_1^2}\, \mathrm{d}y_1 \cdots \int\limits_{-\infty}^{\infty} \mathrm{e}^{-\lambda_n y_n^2}\, \mathrm{d}y_n \\
= &\sqrt{\frac{\pi}{\lambda_1}} \cdot \sqrt{\frac{\pi}{\lambda_2}} \cdots \sqrt{\frac{\pi}{\lambda_n}} = \sqrt{\frac{\pi^n}{\det A}} \quad (\text{siehe Skriptum Lineare Algebra, Seite 88}).
\end{aligned}$$

Weiters ist

$$\begin{aligned}
\mathcal{E}(\underline{v}_i) &= \sqrt{\frac{\det A}{\pi^n}} \int\limits_{\mathbb{R}^n} \overbrace{t_i}^{s_i + \mu_i} \exp\big(-(\vec{t} - \vec{\mu})^T A (\vec{t} - \vec{\mu})\big)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n \\
&= \sqrt{\frac{\det A}{\pi^n}} \underbrace{\int\limits_{\mathbb{R}^n} s_i \exp(-\vec{s}^T A \vec{s})\, \mathrm{d}s_1 \cdots \mathrm{d}s_n}_{0} + \mu_i \underbrace{\int\limits_{\mathbb{R}^n} f_{\vec{v}}(\vec{t})\, \mathrm{d}t_1 \cdots \mathrm{d}t_n}_{1} \\
&= \mu_i \qquad \text{und}
\end{aligned}$$

$$\begin{aligned}
\operatorname{cov}(\underline{v}_i, \underline{v}_j) &= \int\limits_{\mathbb{R}^n} (t_i - \mu_i)(t_j - \mu_j) f_{\vec{v}}(t)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n \\
&= \sqrt{\frac{\det A}{\pi^n}} \int\limits_{\mathbb{R}^n} s_i s_j \exp(-\vec{s}^T A \vec{s})\, \mathrm{d}s_1 \cdots \mathrm{d}s_n \\
&= \sqrt{\frac{\det A}{\pi^n}} \left(-\frac{\partial}{\partial a_{ij}}\right) \int\limits_{\mathbb{R}^n} \exp(-\vec{s}^T A \vec{s})\, \mathrm{d}s_1 \cdots \mathrm{d}s_n \\
&= \sqrt{\frac{\det A}{\pi^n}} \left(-\frac{\partial}{\partial a_{ij}}\right) \sqrt{\frac{\pi^n}{\det A}} \\
&= \sqrt{\det A} \cdot \frac{1}{2} (\det A)^{-3/2} \cdot \frac{\partial \det A}{\partial a_{ij}}
\end{aligned}$$

$$= \quad \frac{1}{2\det A} \cdot A^{\mathrm{ad}}_{ji} \quad \text{(vgl. Skriptum Lineare Algebra § 5, Satz 8, 2), S. 59)}$$

$$= \quad \frac{1}{2}(A^{-1})_{ij} \quad (\text{weil } A = A^T \implies A^{-1} \text{ symmetrisch})$$

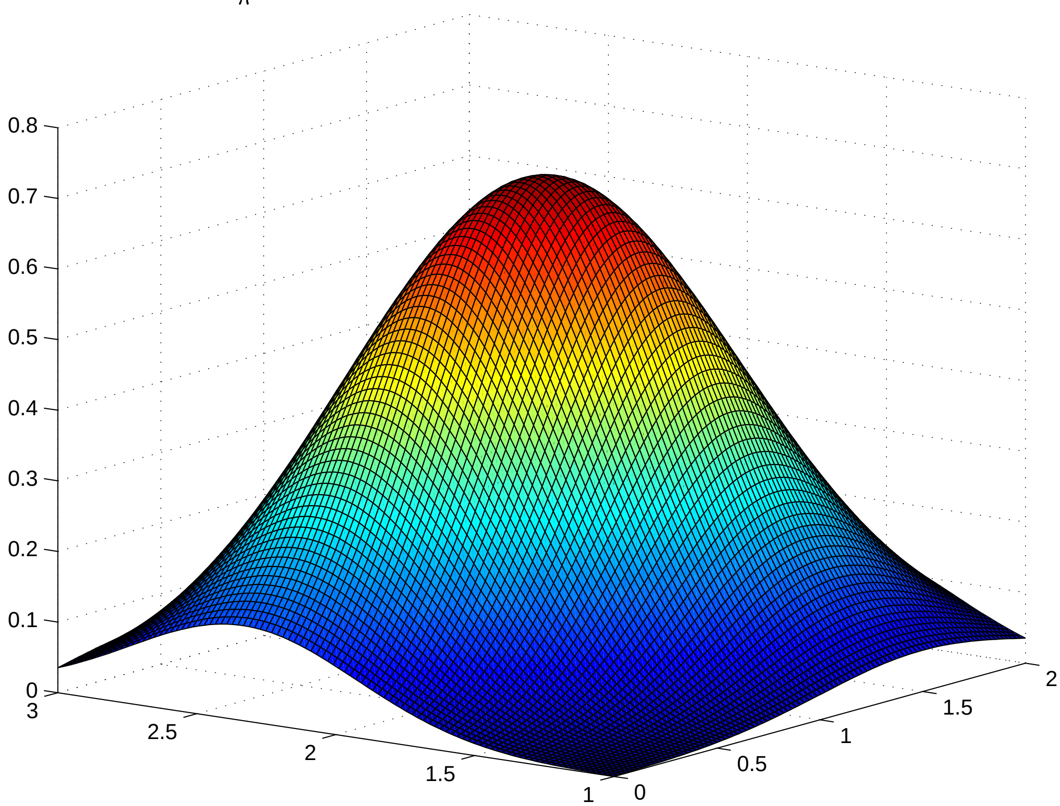
Man nennt $\big(\operatorname{cov}(\underline{v}_i, \underline{v}_j)\big)_{i,j}$ die **Kovarianzmatrix** des Zufallsvektors $\underline{\vec{v}}$. Im Fall einer normalverteilten Zufallsvariablen ist sie also $\frac{1}{2}A^{-1}$. Der Spezialfall $n = 1$ ergibt sich wieder mit $a_{11} = \frac{1}{2\sigma^2}$ und $\operatorname{cov}(\underline{v}_1, \underline{v}_1) = \mathcal{V}(\underline{v}_1) = \frac{1}{2} \cdot \frac{1}{a_{11}} = \sigma^2$, vgl. auch die Rechnung in § 6, Bsp. 4, S. 38.

**Bild für $n = 2$ :**

Dichtefunktion eines normalverteilten Zufallsvektors $\underline{\vec{v}} = (\underline{v}_1, \underline{v}_2)$ mit

$\mu_1 = 1, \mu_2 = 2,\ A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$

$$z = f_{\underline{\vec{v}}}(t_1, t_2) = \frac{\sqrt{5}}{\pi} \exp\big(-2(t_1 - 1)^2 - 3(t_2 - 2)^2 - 2(t_1 - 1)(t_2 - 2)\big)$$



# § 8 Der zentrale Grenzwertsatz

**Problem:** Beim Würfeln mit zwei Würfeln sei $\underline{d} = \underline{x} - \underline{y} =$ „Differenz der Augenzahlen des 1. bzw. 2. Würfels “, $\underline{s} =$ „Anzahl der Sechser“. Dann sind $\underline{s}$, $\underline{d}$ unkorreliert (siehe Übung 30), d.h. die Kenntnis von $\underline{s}$ liefert nichts für eine **lineare** Schätzung von $\underline{d}$. Dennoch gilt: $\underline{s} = 2 \implies \underline{d} = 0$ und $\underline{s} = 0 \implies |\underline{d}| \leq 4$, d.h. $\underline{s}$ liefert Informationen über $\underline{d}$.

**Def.:** $\underline{x}, \underline{y} : \Omega \longrightarrow \mathbb{R}$ heißen **unabhängig** $\iff \forall A, B \subset \mathbb{R}$ Intervalle: $\{\underline{x} \in A\}$, $\{\underline{y} \in B\}$ sind unabhängig$\iff \forall A, B \subset \mathbb{R}$ Intervalle: $P\big(\{\underline{x} \in A\} \cap \{\underline{y} \in B\}\big) =$
$= P\big(\{\underline{x} \in A\}\big) \cdot P\big(\{\underline{y} \in B\}\big)$.

**Bemerkung:** Die letzte Gleichung heißt: Entweder $P\big(\underline{y} \in B\}\big) = 0$ oder $P\big(\{\underline{x} \in A\}\big|\{\underline{y} \in B\}\big) = P\big(\{\underline{x} \in A\}\big)$, d.h. das Wissen, dass $\underline{y}$ Werte in $B$ hat, verändert die Wahrscheinlichkeit, dass $\underline{x}$ Werte in $A$ hat, **nicht**.
Bzw.: Aus $\underline{y}$ lässt sich **nichts** über $\underline{x}$ schließen.

**Satz 1:** $\underline{x}, \underline{y}$ unabhängig $\implies$ $\underline{x}, \underline{y}$ unkorreliert.

**Beweis:** Z.B. für $\underline{x}, \underline{y}$ diskret: $\underline{x}$ habe Werte $x_i$ mit Wahrscheinlichkeit $p_i$, $\underline{y}$ habe Werte $y_j$ mit Wahrscheinlichkeit $q_j$; $\underline{x}, \underline{y}$ unabhängig

$$\implies P\big(\{\underline{x} = x_i \wedge \underline{y} = y_j\}\big) = P\big(\{\underline{x} = x_i\}\big) \cdot P\big(\{\underline{y} = y_j\}\big) = p_i q_j$$

$$\implies \mathcal{E}(\underline{x} \cdot \underline{y}) = \sum_i \sum_j \underbrace{x_i y_j}_{\substack{\text{Werte von} \\ \underline{x} \cdot \underline{y}}} \cdot \underbrace{p_i q_j}_{\text{ihre W.}} = \big(\sum x_i p_i\big) \cdot \big(\sum y_j q_j\big)$$

$$= \mathcal{E}(\underline{x}) \cdot \mathcal{E}(\underline{y}) \implies \underline{x}, \underline{y} \text{ unkorreliert.}$$ □

**Bsp. 1:** Zweimal Würfeln, $\underline{x}, \underline{y}, \underline{s}, \underline{d}$ wie oben. $\underline{x}, \underline{y}$ sind unabhängig, denn

$$P\big(\{\underline{x} = i\} \cap \{\underline{y} = j\}\big) = P\big((i, j)\big) = \tfrac{1}{36} = \tfrac{1}{6} \cdot \tfrac{1}{6} =$$
$$= P\big(\{\underline{x} = i\}\big) \cdot P\big(\{\underline{y} = j\}\big) \quad \text{vgl. auch S. 30.}$$

Aber $\underline{d}, \underline{s}$ sind nicht unabhängig, denn

$$P\big(\{\underline{s} = 2\} \cap \{\underline{d} = 5\}\big) = 0 \neq P\big(\{\underline{s} = 2\}\big) \cdot P\big(\{\underline{d} = 5\}\big) = \frac{1}{36} \cdot \frac{1}{36}.$$

Verallgemeinerung:

**Def.:** 1) $\underline{x}_1, \cdots, \underline{x}_n : \Omega \longrightarrow \mathbb{R}$ heißen **unabhängig**

$\iff$ aus der Kenntnis von $\underline{x}_1, \cdots, \underline{x}_{j-1}, \underline{x}_{j+1}, \cdots, \underline{x}_n$ lässt sich **nichts** über $\underline{x}_j$ schließen $(j = 1, \cdots, n)$

$\Longleftrightarrow \forall A_1, \cdots, A_n \subset \mathbb{R}$ Intervalle: $P\big(\{\underline{x}_1 \in A_1\} \cap \cdots \cap \{\underline{x}_n \in A_n\}\big) = \prod_{j=1}^{n} P\big(\{\underline{x}_j \in A_j\}\big)$

2) $\underline{x}_1, \underline{x}_2, \cdots$ heißen unabhängig $\Longleftrightarrow \forall n : \underline{x}_1, \cdots, \underline{x}_n$ unabhängig.

**Bemerkung:** Bei Experimenten wird die Unabhängigkeit aus den Versuchsbedingungen geschlossen. Z.B. 55 Betonproben, wobei die Herstellungsanordnung zwischendurch in den Originalzustand gebracht wird, $\underline{x}_j =$ Dichte beim $j$-ten Mal, $j = 1, \cdots, 55$.
Oder: $\infty$-oft würfeln, $\underline{x}_j =$ Ergebnis beim $j$-ten Mal, $j = 1, 2, 3, \cdots$.

**Satz 2:** $\underline{x}_1, \cdots, \underline{x}_n$ unabhängig $\Longrightarrow \mathcal{V}(\underline{x}_1 + \cdots + \underline{x}_n) = \mathcal{V}(\underline{x}_1) + \cdots + \mathcal{V}(\underline{x}_n)$.

**Beweis:**

$$\begin{aligned}
\mathcal{V}(\underline{x}_1 + \cdots + \underline{x}_n) &= \mathcal{E}\big((\underline{x}_1 + \cdots + \underline{x}_n)^2\big) - \mathcal{E}(\underline{x}_1 + \cdots + \underline{x}_n)^2 \\
&= \mathcal{E}\Big(\sum_{j=1}^{n}\sum_{k=1}^{n} \underline{x}_j\underline{x}_k\Big) - \Big(\sum_{j=1}^{n}\mathcal{E}(\underline{x}_j)\Big)\cdot\Big(\sum_{k=1}^{n}\mathcal{E}(\underline{x}_k)\Big) \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n}\big[\mathcal{E}(\underline{x}_j\underline{x}_k) - \mathcal{E}(\underline{x}_j)\mathcal{E}(\underline{x}_k)\big] = \sum_{j=1}^{n}\sum_{k=1}^{n} \operatorname{cov}(\underline{x}_j, \underline{x}_k);
\end{aligned}$$

$\underline{x}_j, \underline{x}_k$ unabhängig für $j \neq k \Longrightarrow \operatorname{cov}(\underline{x}_j, \underline{x}_k) = \begin{cases} 0 & : \ j \neq k \\ \mathcal{V}(\underline{x}_j) & : \ j = k \end{cases} \Longrightarrow$

$\Longrightarrow \mathcal{V}(\underline{x}_1 + \cdots + x_n) = \sum_{j=1}^{n} \mathcal{V}(\underline{x}_j)$. $\square$

**Bsp. 2:** Varianz der Binomialverteilung
Wir werfen $n$-mal eine Münze, die mit Wahrscheinlichkeit $p$ auf Zahl fällt,

$$\underline{x}_j = \begin{cases} 1: & \text{beim } j\text{-ten Wurf Zahl,} \\ 0: & \text{beim } j\text{-ten Wurf Kopf.} \end{cases}$$

$\underline{x} = \underline{x}_1 + \cdots + \underline{x}_n =$ „Anzahl von Zahl“, $P_{\underline{x}} = \operatorname{Bin}(n, p)$, vgl. § 5, Bsp. 2, Seite 33.

$$\begin{aligned}
\mathcal{E}(\underline{x}_j) &= 1 \cdot p + 0 \cdot q = p \quad \Longrightarrow \quad \mathcal{E}(\underline{x}) = np, \text{ vgl. § 6, Bsp. 1, Seite 35} \\
\mathcal{E}(\underline{x}_j^2) &= 1^2 p + 0^2 q = p \quad \Longrightarrow \quad \mathcal{V}(\underline{x}_j) = p - p^2 = p(1-p) = pq,
\end{aligned}$$

$\underline{x}_1, \cdots, \underline{x}_n$ unabhängig $\Longrightarrow \mathcal{V}(\underline{x}) = \mathcal{V}(\underline{x}_1) + \cdots + \mathcal{V}(\underline{x}_n) = npq$, vgl. Seite 38.

**Satz 3:** (ohne Beweis)
$\underline{x}_1, \cdots, \underline{x}_n : \Omega \longrightarrow \mathbb{R}$ seien unabhängig und normalverteilt, $c_1, \cdots, c_n \in \mathbb{R}$.
Dann ist auch $\underline{x} = c_1\underline{x}_1 + \cdots + c_n\underline{x}_n$ normalverteilt.

**Bemerkung:** Aus den Parametern von $\underline{x}_1, \cdots, \underline{x}_n$ erhalten wir die von $\underline{x}$ :
$P_{\underline{x}_j} = N(\mu_j, \sigma_j) \Longrightarrow \mathcal{E}(\underline{x}) = \mathcal{E}(c_1\underline{x}_1 + \cdots + c_n\underline{x}_n) = c_1\mu_1 + \cdots + c_n\mu_n$
und $\mathcal{V}(\underline{x}) = \mathcal{V}(c_1\underline{x}_1 + \cdots + c_n\underline{x}_n) = \mathcal{V}(c_1\underline{x}_1) + \cdots + \mathcal{V}(c_n\underline{x}_n) = c_1^2\sigma_1^2 + \cdots + c_n^2\sigma_n^2$
(weil $c_1\underline{x}_1, \cdots, c_n\underline{x}_n$ unabhängig)

$$\Longrightarrow \; P_{\underline{x}} = N(\mu,\sigma) \text{ mit } \mu = c_1\mu_1 + \cdots + c_n\mu_n$$
$$\text{und } \sigma = \sqrt{c_1^2\sigma_1^2 + \cdots + c_n^2\sigma_n^2}.$$

**Bsp. 3:** In einer Zuckerraffinerie wird maschinell Zucker verpackt. Das Gewicht der Pakete ist normalverteilt mit $\sigma = 5\,\text{g}$. Bei Kontrollen entnimmt das Marktamt Stichproben von 30 Paketen Zucker. Die Raffinerie wird bestraft, wenn das mittlere Gewicht der Stichprobe $< 999\,\text{g}$ ist. Wie ist $\mu$ bei der Abfüllmaschine einzustellen, damit das Risiko einer Bestrafung $1\,\%$ ist?

**Lösung:** $\underline{x}_1, \cdots, \underline{x}_{30}$ seien die Gewichte der 30 Pakete; sie werden als unabhängig angenommen und sind alle $N(\mu, 5)$-verteilt. Nach Satz 3 ist $\underline{\overline{x}}$ = „mittleres Gewicht der Stichprobe“$= \frac{1}{30}(\underline{x}_1 + \cdots + \underline{x}_{30})$ normalverteilt mit

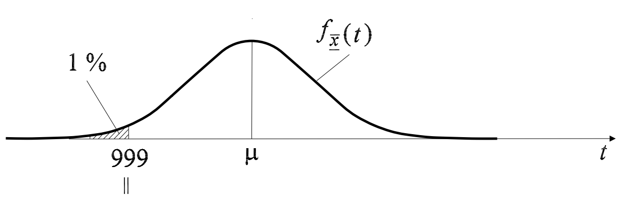
$$\mathcal{E}(\underline{\overline{x}}) = \tfrac{1}{30}(\mu + \cdots + \mu) = \mu \qquad \big(c_j = \tfrac{1}{30},\ \mu_j = \mu,\ \sigma_j = 5\big)$$
$$\sigma(\underline{\overline{x}}) = \sqrt{\tfrac{5^2}{30^2} + \cdots + \tfrac{5^2}{30^2}} = \tfrac{5}{30}\cdot\sqrt{30} = \tfrac{5}{\sqrt{30}}$$

Also: $\underline{\overline{x}}$ ist $N\big(\mu, \frac{5}{\sqrt{30}}\big)$-verteilt.
(Die Schwankung von $\underline{\overline{x}}$ ist kleiner als die der einzelnen $\underline{x}_j$!)

**Forderung:** $P_{\underline{\overline{x}}}\big(]-\infty, 999]\big) \overset{!}{=} 1\,\%$

$$\Longrightarrow \; \Phi\left(\frac{999-\mu}{5/\sqrt{30}}\right) - \Phi(-\infty) = 0.01$$
$$\Longrightarrow \; -\Phi\left(\frac{\mu-999}{5/\sqrt{30}}\right) + \frac{1}{2} = 0.01 \;\Longrightarrow\; \Phi\left(\frac{\mu-999}{5/\sqrt{30}}\right) = 0.49$$
$$\overset{\text{s.S. 22}}{\Longrightarrow} \; \frac{\mu-999}{5/\sqrt{30}} = 2.327 \;\Longrightarrow\; \mu = 999 + \frac{5\cdot 2.327}{\sqrt{30}} \approx 1001.12\,[\text{g}].$$

Bild:




$1\,\%$-Fraktil von $P_{\underline{\overline{x}}}$

**Def.:** Die Zufallsgrößen $\underline{x}_1, \cdots, \underline{x}_n$ ($n \in \mathbb{N}$ oder $n = \infty$) heißen **Stichprobe vom Umfang $n$** zur Verteilung $P \Longleftrightarrow$ a) $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n$ unabhängig;
b) $\forall j : P_{\underline{x}_j} = P$, d.h. alle $\underline{x}_j$ haben dieselbe Verteilung $P$.
Dann heißt $\overline{\underline{x}}_{(n)} = \dfrac{1}{n} \sum\limits_{j=1}^{n} \underline{x}_j$ $n$-tes **Stichprobenmittel**.

**Schreibweise:** Statt $\overline{\underline{x}}_{(n)}$ schreibt man meist nur $\overline{\underline{x}}$. Für Erwartungswert bzw. Varianz von $\underline{x}_j$, schreibt man meist $\mu$ bzw. $\sigma^2$, d.h. $\mathcal{E}(\underline{x}_j) = \mu$, $\mathcal{V}(\underline{x}_j) = \sigma^2$.
(Falls $\underline{x}_j$ Expo$\,(\mu_{\text{alt}})$-verteilt sind, wäre $\mu = \frac{1}{\mu_{\text{alt}}}$.)

**Bemerkung:** $\mathcal{E}(\overline{\underline{x}}) = \dfrac{1}{n} \sum\limits_{j=1}^{n} \mathcal{E}(\underline{x}_j) = \dfrac{1}{n} \cdot n\mu = \mu$,

$$\mathcal{V}(\overline{\underline{x}}) = \frac{1}{n^2} \mathcal{V}\Big(\sum_{j=1}^{n} \underline{x}_j\Big) \underset{\text{a)}}{=} \frac{1}{n^2} \cdot \sum_{j=1}^{n} \mathcal{V}(\underline{x}_j) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

(Beachte, dass $\mathcal{V}(c\underline{y}) = c^2 \mathcal{V}(\underline{y})$!)

Also: $\boxed{\sigma(\overline{\underline{x}}) = \dfrac{\sigma}{\sqrt{n}}}$, d.h. $\overline{\underline{x}}$ streut weniger als die einzelnen $\underline{x}_j$.

Weiters gilt:

| | $\mathcal{E}$ | $\mathcal{V}$ |
|---|---|---|
| $\overline{\underline{x}}$ | $\mu$ | $\dfrac{\sigma^2}{n}$ |
| $\overline{\underline{x}} - \mu$ | $0$ | $\dfrac{\sigma^2}{n}$ |
| $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}}$ | $0$ | $1$ |

**Def.:** $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}}$ heißt $n$-tes **standardisiertes Stichprobenmittel**.

**Bemerkung:** Falls die $\underline{x}_j$ normalverteilt sind (d.h. $P = P_{\underline{x}_j} = N(\mu, \sigma)$), so auch $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}}$ nach Satz 3. Daher ist dann $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}}$ $N(0,1)$-verteilt. Der zentrale Grenzwertsatz sagt, dass $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}}$ für beliebige Stichproben **im Limes $n \to \infty$** $N(0,1)$-verteilt ist.

**Satz 4: (Zentraler Grenzwertsatz, Gauß, ohne Beweis)**
$\underline{x}_1, \underline{x}_2, \cdots$ unendliche Stichprobe, $\mu = \mathcal{E}(\underline{x}_j)$, $\sigma = \sqrt{\mathcal{V}(\underline{x}_j)}$. Dann ist $\underline{y}_n = \dfrac{\overline{\underline{x}}_{(n)} - \mu}{\sigma/\sqrt{n}}$ im Limes $N(0,1)$-verteilt, d.h. $\lim\limits_{n\to\infty} P_{\underline{y}_n} = N(0,1)$, d.h.

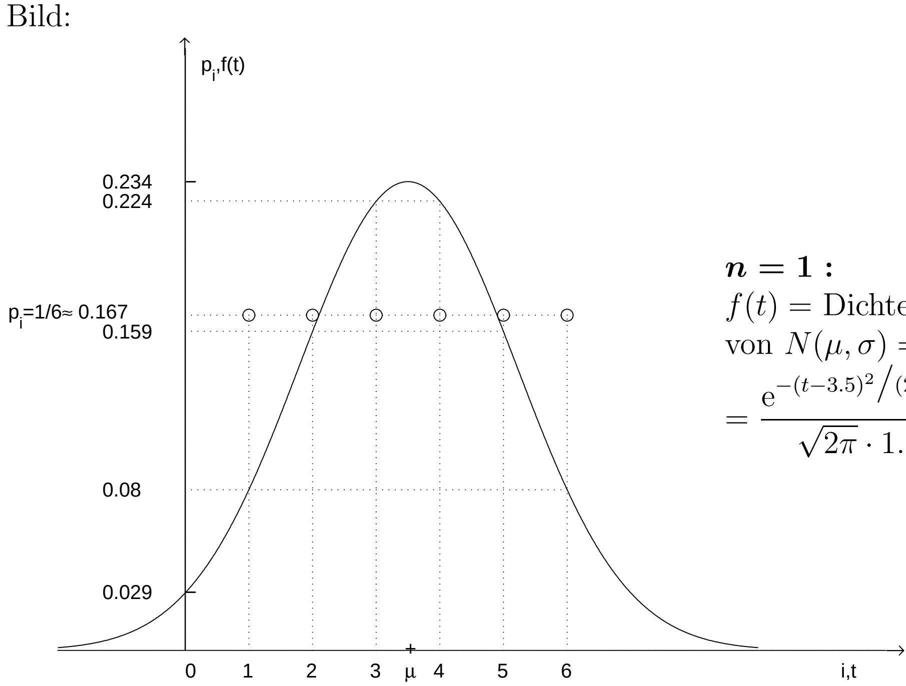
$$\forall a < b : \lim_{n\to\infty} P\left(\left\{a \le \frac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}} \le b\right\}\right) = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-t^2/2}\,\mathrm{d}t.$$

**Bsp. 4:** $\infty$-oft Würfeln, $\underline{x}_j$ = Augenzahl beim $j$-ten Wurf, $\mu = 3.5$, $\sigma = \sqrt{\frac{35}{12}} \approx 1.708$

$$\frac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}} = \frac{\frac{\underline{x}_1 + \cdots + \underline{x}_n}{n} - \mu}{\sigma/\sqrt{n}} = \frac{\underline{x}_1 + \cdots + \underline{x}_n - n\mu}{\sigma \cdot \sqrt{n}} \approx N(0,1)\text{-verteilt}$$

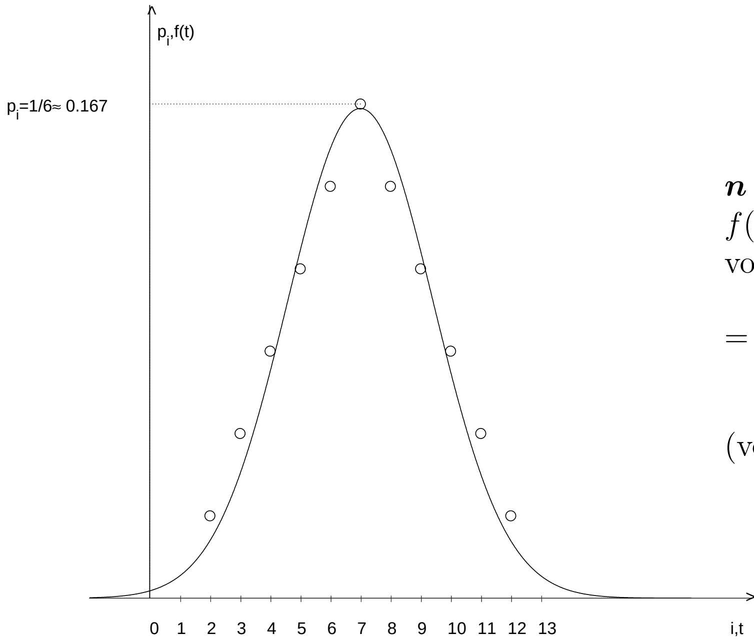
$\Longrightarrow$ $\underline{x}_1 + \cdots + \underline{x}_n$ = Augensumme $\approx N(n\mu, \sqrt{n}\sigma)$-verteilt, $n \to \infty$.

Bild:




***n* = 1 :**
$f(t)$ = Dichtefunktion
von $N(\mu, \sigma)$ =

$$= \frac{\mathrm{e}^{-(t-3.5)^2 / (2\cdot 35/12)}}{\sqrt{2\pi} \cdot 1.708}$$




$\boldsymbol{n = 2:}$
$f(t)$ = Dichtefunktion
von $N(2\mu, \sqrt{2}\,\sigma) =$

$$= \frac{\mathrm{e}^{-(t-7)^2 \big/ (2\cdot 2\cdot 35/12)}}{\sqrt{2\pi}\cdot\sqrt{2}\cdot 1.708}$$

(vergl. Seite 33)

$\boldsymbol{n = 3:}$ siehe Übung 35.

**Bsp. 5:** Den Satz von Bernoulli, (Seite 24), erhalten wir aus Satz 4 so:
$\infty$-oft Münze werfen, die mit Wahrscheinlichkeit $p$ auf Zahl fällt,

$$\underline{x}_j = \begin{cases} 1 & : \text{ Zahl beim } j\text{-ten Mal,} \\ 0 & : \text{ Kopf beim } j\text{-ten Mal} \end{cases}$$

$\mu = p,\ \sigma = \sqrt{pq}$, vgl. Bsp. 2, Seite 51

$\Longrightarrow\ \underbrace{\underline{x}_1 + \cdots + \underline{x}_n}_{\text{Bin}\,(n,p)\text{-verteilt}} \approx N\big(n\mu, \sqrt{n}\,\sigma\big) = N\big(np, \sqrt{npq}\big)$-verteilt für $n \to \infty$

$\Longrightarrow\ \text{Bin}\,(n,p) \approx N\big(np, \sqrt{npq}\big)$

**Bemerkung:** Der noch allgemeinere Satz von A.M. Ljapunow sagt:
$\underline{x}_1, \underline{x}_2, \cdots$ seien unabhängige Zufallsgrößen (nicht notwendig mit gleicher Verteilung), $\mathcal{E}(\underline{x}_j) = \mu_j,\ \mathcal{V}(\underline{x}_j) = \sigma_j^2$

$$\Longrightarrow\ \underline{y}_n = \frac{\underline{x}_1 + \cdots + \underline{x}_n - \mu_1 - \cdots - \mu_n}{\sqrt{\sigma_1^2 + \cdots + \sigma_n^2}}$$

ist im Limes $n \to \infty$ $N(0,1)$-verteilt, falls $\sum_{j=1}^{\infty} \sigma_j^2 = \infty$. Dies ist der Grund dafür, dass in Natur und Technik Zufallsgrößen, die sich durch Überlagerung vieler Ursachen, d.h. als Summe vieler unabhängiger Einzelgrößen, ergeben, annähernd normalverteilt sind. Wesentlich ist dabei, dass die einzelnen $\sigma_j^2$ im Verhältnis zur gesamten Varianz $\sigma_1^2 + \cdots + \sigma_n^2$ nicht allzu groß sind.

# Kapitel 3

# Statistik

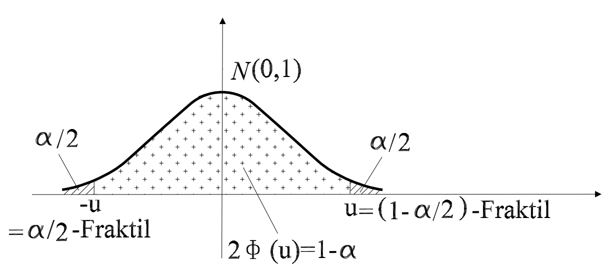
## § 9 Konfidenzintervalle für $\mu$

**Aufgabe:** Eine Stichprobe liefert Werte $x_1, \cdots, x_n$ mit Mittelwert $\overline{x} = \frac{1}{n}\sum\limits_{j=1}^{n} x_j$
(z.B. $\overline{x} = 2.44418$ auf Seite 45).
Bestimme ein Intervall $[a, b]$ um $\overline{x}$ (das **Konfidenzintervall**), in dem $\mu = \mathcal{E}(\underline{x}_j)$ mit Wahrscheinlichkeit $1 - \alpha$ liegt. $\alpha =$**Signifikanzniveau** wird vorgegeben (meist 5 %, 2 % oder 1 %).

### A) 2-seitige Konfidenzintervalle bei bekanntem $\sigma$

**Annahmen:** $\underline{x}_1, \cdots, \underline{x}_n$ Stichprobe mit bekannter Varianz $\mathcal{V}(\underline{x}_j) = \sigma^2$ und unbekanntem $\mu = \mathcal{E}(\underline{x}_j)$. Dann ist $\mathcal{E}(\overline{\underline{x}}) = \mu,\ \mathcal{V}(\overline{\underline{x}}) = \dfrac{\sigma^2}{n}$. Entweder seien $\underline{x}_j$ normalverteilt oder $n$ so groß ($n \geq 10$ ist meist genug), dass $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$-verteilt (siehe § 8).

**Berechnung:** Es sei $u = \left(1 - \frac{\alpha}{2}\right)$-Fraktil von $N(0,1)$, d.h. $2\Phi(u) = 1 - \alpha$.




Dann gilt:

$$\begin{aligned} & P\left(\left\{-u \leq \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} \leq u\right\}\right) = 1-\alpha \\ = \; & P\left(\left\{-u \leq \frac{\mu-\overline{x}}{\sigma/\sqrt{n}} \leq u\right\}\right) \\ = \; & P\left(\left\{-\frac{\sigma u}{\sqrt{n}} \leq \mu-\overline{x} \leq \frac{\sigma u}{\sqrt{n}}\right\}\right) \\ = \; & P\left(\left\{\underbrace{\overline{x}-\frac{\sigma u}{\sqrt{n}}}_{\underline{a}} \leq \mu \leq \underbrace{\overline{x}+\frac{\sigma u}{\sqrt{n}}}_{\underline{b}}\right\}\right) \end{aligned}$$

**Ergebnis:** $\mu$ liegt mit Wahrscheinlichkeit $1-\alpha$ im Intervall $[\underline{a}, \underline{b}] = \left[\overline{x}-\frac{\sigma u}{\sqrt{n}},\ \overline{x}+\frac{\sigma u}{\sqrt{n}}\right]$.

**Vorsicht:** $\mu$ ist **keine** Zufallsgröße! Die Aussage ist: Bei wiederholter Bestimmung von $\overline{x}$ und entsprechenden $a, b$ ist die Wahrscheinlichkeit, dass $\mu \in [a, b]$, gleich $1-\alpha$.

**Bemerkung:** Falls $\sigma^2$ nicht bekannt ist, nimmt man für $\sigma^2$ die **erwartungstreue Stichprobenvarianz** $\hat{s}^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j-\overline{x})^2 = \frac{n}{n-1}s^2$ wobei $s^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j-\overline{x})^2$.

Kurze Erklärung: $\mathcal{V}(\underline{x}_j) = \sigma^2 = \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)^2$, siehe Seite 37. Weil $\overline{x}$ statt $\mu$ genommen wird, muss der Faktor $\frac{1}{n}$ durch $\frac{1}{n-1}$ ersetzt werden. Näheres dazu in § 10; siehe auch Übung (Z9).

**Bsp.:** Betonproben auf Seite 44, $\alpha = 5\,\%$ vorgegeben $\implies u = 97.5\,\%\text{-Fraktil} \overset{\text{s.S. 22}}{=} 1.96$
$\overline{x} = \mathcal{E}(\underline{X}) = 2.44418$, $s^2 = \mathcal{V}(\underline{X}) = 0.00121342$, $\hat{s}^2 = \frac{n}{n-1}s^2 = \frac{55}{54}\cdot 0.00121342 \approx 0.00123589$
Annahme: $\sigma = \sqrt{\hat{s}^2} \approx 0.035155$.
(Bei konkreten Werten lässt man die Striche unter den Zufallsgrößen weg!)

$$\begin{aligned} \implies \quad a &= \overline{x} - \tfrac{\sigma u}{\sqrt{55}} \approx 2.4349 \\ b &= \overline{x} + \tfrac{\sigma u}{\sqrt{55}} \approx 2.4535 \end{aligned}$$

Also: Mit 95 % Sicherheit liegt $\mu$ in [2.4349, 2.4535] (Einmal unter 20 Testserien würde $\mu$ nicht im Konfidenzintervall liegen.)
Bei $\alpha = 1\,\%$ ist $u = 2.575$ und das Konfidenzintervall vergrößert sich ein wenig:
$[a, b] = [2.432,\ 2.4564]$.

## B) Einseitige Konfidenzintervalle bei bekanntem $\sigma$

Oft interessiert nur, dass $\mu \geq$ bzw. $\leq$ eine bestimmte Schranke ist. Gesucht ist dann ein **einseitiges Konfidenzintervall** $[a, \infty[$ bzw. $]-\infty, b]$, in dem $\mu$ mit Wahrscheinlichkeit $1 - \alpha$ liegt.

Z.B. $[a, \infty[$: $u = (1-\alpha)$-Fraktil von $N(0,1) \implies 1 - \alpha = P\left(\left\{\frac{\underline{\overline{x}} - \mu}{\sigma/\sqrt{n}} \leq u\right\}\right) =$

$$= P\left(\left\{\underline{\overline{x}} - \mu \leq \frac{\sigma u}{\sqrt{n}}\right\}\right) = P\left(\left\{\underline{\overline{x}} - \frac{\sigma u}{\sqrt{n}} \leq \mu\right\}\right) \implies \underline{a} = \underline{\overline{x}} - \frac{\sigma u}{\sqrt{n}}$$

**Bsp.:** Betonproben, $\alpha = 5\,\%$, $u = 95\,\%$-Fraktil $\overset{\text{(s.S. 22)}}{=} 1.645 \implies a = \overline{x} - \frac{\sigma u}{\sqrt{n}} = 2.4364$

Also: Mit $95\,\%$ Sicherheit ist $\mu \geq 2.4364$.

## C) Konfidenzintervalle bei unbekanntem $\sigma$

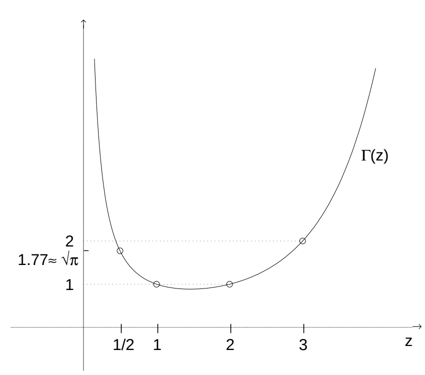
Dies war im Beispiel oben eigentlich auch der Fall. Wir griffen zur Notlösung $\sigma$ durch $\hat{s} = 0.035155$ zu ersetzen.

**Korrekte Lösung:** Betrachte die Zufallsgröße $\underline{t} = \frac{\underline{\overline{x}} - \mu}{\underline{\hat{s}}/\sqrt{n}}$.

Für großes $n$ ist $\underline{\hat{s}} \approx \sigma$ und $\underline{t} \approx N(0,1)$-verteilt.
Exakt gilt: $\underline{t}$ ist „$t$-verteilt mit $n-1$ Freiheitsgraden“. Die Dichte der **(Student'schen) $\boldsymbol{t}$-Verteilung** mit $\nu = n-1$ Freiheitsgraden ist (ohne Beweis):

$$\begin{aligned}
f(t) &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, \text{ wobei} \\
\Gamma(z) &= \int_0^\infty x^{z-1}\mathrm{e}^{-x}\,\mathrm{d}x,\ \Gamma(n) = (n-1)! \text{ für } n \in \mathbb{N} \\
\Gamma\left(\tfrac{1}{2}\right) &= \sqrt{\pi},\ \Gamma(z+1) = z \cdot \Gamma(z) \quad \text{(„Funktionalgleichung“)}
\end{aligned}$$

$\Gamma$ heißt **Euler'sche Gammafunktion**.



Wir erhalten z.B. für $\nu = 3$ :

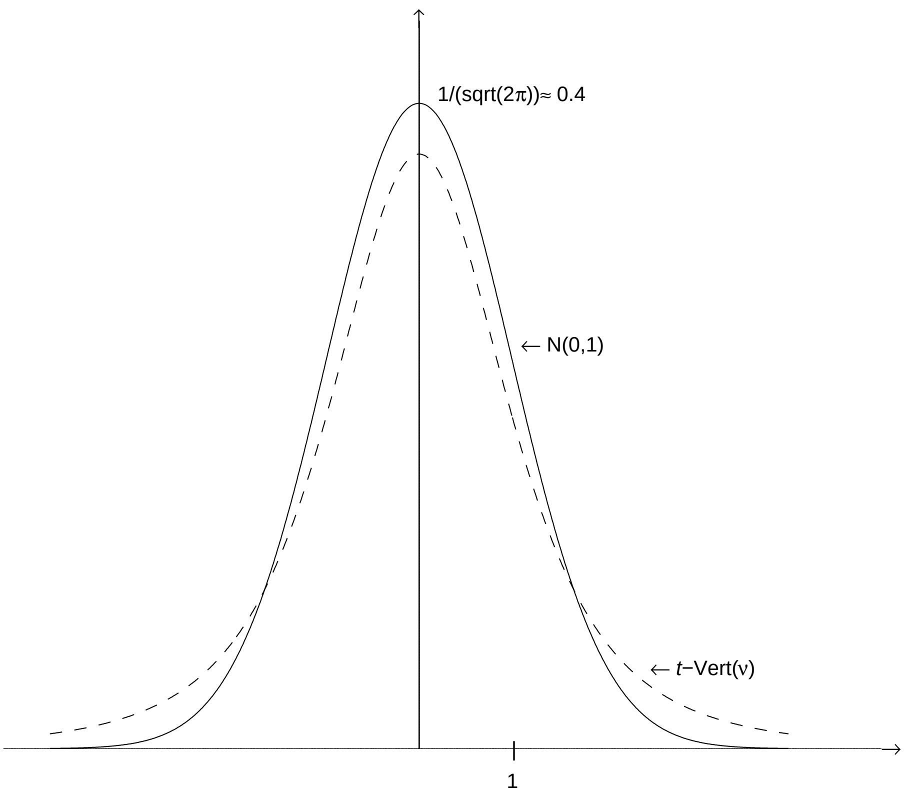
$$f(t) = \frac{(1+t^2/3)^{-2}}{\sqrt{3\pi}\cdot\frac{1}{2}\cdot\sqrt{\pi}}$$

Es gilt: $\mathcal{E}(\underline{t}) = 0 \qquad$ (Symmetrie)

$\qquad\qquad \mathcal{V}(\underline{t}) = \dfrac{\nu}{\nu-2} > 1$

und $t$-Vert $(\nu) \longrightarrow N(0,1)$ für $\nu \to \infty \qquad (*)$

Bild:

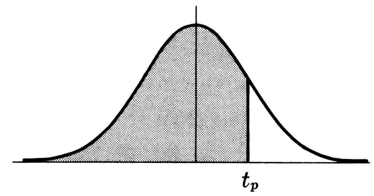



(Beweis von $(*)$ : $\displaystyle\lim_{\nu\to\infty}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} = \lim_{\nu\to\infty}\frac{\overbrace{\left(1+\frac{t^2}{\nu}\right)^{-1/2}}^{\to 1}}{\sqrt{\underbrace{\left(1+\frac{t^2}{\nu}\right)^{\nu}}_{\to \mathrm{e}^{t^2}}}} = \mathrm{e}^{-t^2/2}$;

der Vorfaktor ist so, dass $\displaystyle\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$ und konvergiert daher gegen $\dfrac{1}{\sqrt{2\pi}}$.)

# *Appendix III*

**PERCENTILE VALUES ($t_p$)**
**for**
**STUDENT'S $t$ DISTRIBUTION**
**with $\nu$ degrees of freedom**
(shaded area $= p$)

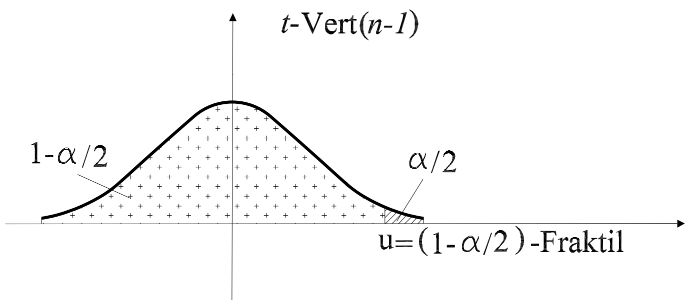



| $\nu$ | $t_{.995}$ | $t_{.99}$ | $t_{.975}$ | $t_{.95}$ | $t_{.90}$ | $t_{.80}$ | $t_{.75}$ | $t_{.70}$ | $t_{.60}$ | $t_{.55}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 63.66 | 31.82 | 12.71 | 6.31 | 3.08 | 1.376 | 1.000 | .727 | .325 | .158 |
| 2 | 9.92 | 6.96 | 4.30 | 2.92 | 1.89 | 1.061 | .816 | .617 | .289 | .142 |
| 3 | 5.84 | 4.54 | 3.18 | 2.35 | 1.64 | .978 | .765 | .584 | .277 | .137 |
| 4 | 4.60 | 3.75 | 2.78 | 2.13 | 1.53 | .941 | .741 | .569 | .271 | .134 |
| 5 | 4.03 | 3.36 | 2.57 | 2.02 | 1.48 | .920 | .727 | .559 | .267 | .132 |
| 6 | 3.71 | 3.14 | 2.45 | 1.94 | 1.44 | .906 | .718 | .553 | .265 | .131 |
| 7 | 3.50 | 3.00 | 2.36 | 1.90 | 1.42 | .896 | .711 | .549 | .263 | .130 |
| 8 | 3.36 | 2.90 | 2.31 | 1.86 | 1.40 | .889 | .706 | .546 | .262 | .130 |
| 9 | 3.25 | 2.82 | 2.26 | 1.83 | 1.38 | .883 | .703 | .543 | .261 | .129 |
| 10 | 3.17 | 2.76 | 2.23 | 1.81 | 1.37 | .879 | .700 | .542 | .260 | .129 |
| 11 | 3.11 | 2.72 | 2.20 | 1.80 | 1.36 | .876 | .697 | .540 | .260 | .129 |
| 12 | 3.06 | 2.68 | 2.18 | 1.78 | 1.36 | .873 | .695 | .539 | .259 | .128 |
| 13 | 3.01 | 2.65 | 2.16 | 1.77 | 1.35 | .870 | .694 | .538 | .259 | .128 |
| 14 | 2.98 | 2.62 | 2.14 | 1.76 | 1.34 | .868 | .692 | .537 | .258 | .128 |
| 15 | 2.95 | 2.60 | 2.13 | 1.75 | 1.34 | .866 | .691 | .536 | .258 | .128 |
| 16 | 2.92 | 2.58 | 2.12 | 1.75 | 1.34 | .865 | .690 | .535 | .258 | .128 |
| 17 | 2.90 | 2.57 | 2.11 | 1.74 | 1.33 | .863 | .689 | .534 | .257 | .128 |
| 18 | 2.88 | 2.55 | 2.10 | 1.73 | 1.33 | .862 | .688 | .534 | .257 | .127 |
| 19 | 2.86 | 2.54 | 2.09 | 1.73 | 1.33 | .861 | .688 | .533 | .257 | .127 |
| 20 | 2.84 | 2.53 | 2.09 | 1.72 | 1.32 | .860 | .687 | .533 | .257 | .127 |
| 21 | 2.83 | 2.52 | 2.08 | 1.72 | 1.32 | .859 | .686 | .532 | .257 | .127 |
| 22 | 2.82 | 2.51 | 2.07 | 1.72 | 1.32 | .858 | .686 | .532 | .256 | .127 |
| 23 | 2.81 | 2.50 | 2.07 | 1.71 | 1.32 | .858 | .685 | .532 | .256 | .127 |
| 24 | 2.80 | 2.49 | 2.06 | 1.71 | 1.32 | .857 | .685 | .531 | .256 | .127 |
| 25 | 2.79 | 2.48 | 2.06 | 1.71 | 1.32 | .856 | .684 | .531 | .256 | .127 |
| 26 | 2.78 | 2.48 | 2.06 | 1.71 | 1.32 | .856 | .684 | .531 | .256 | .127 |
| 27 | 2.77 | 2.47 | 2.05 | 1.70 | 1.31 | .855 | .684 | .531 | .256 | .127 |
| 28 | 2.76 | 2.47 | 2.05 | 1.70 | 1.31 | .855 | .683 | .530 | .256 | .127 |
| 29 | 2.76 | 2.46 | 2.04 | 1.70 | 1.31 | .854 | .683 | .530 | .256 | .127 |
| 30 | 2.75 | 2.46 | 2.04 | 1.70 | 1.31 | .854 | .683 | .530 | .256 | .127 |
| 40 | 2.70 | 2.42 | 2.02 | 1.68 | 1.30 | .851 | .681 | .529 | .255 | .126 |
| 60 | 2.66 | 2.39 | 2.00 | 1.67 | 1.30 | .848 | .679 | .527 | .254 | .126 |
| 120 | 2.62 | 2.36 | 1.98 | 1.66 | 1.29 | .845 | .677 | .526 | .254 | .126 |
| $\infty$ | 2.58 | 2.33 | 1.96 | 1.645 | 1.28 | .842 | .674 | .524 | .253 | .126 |

*Source:* R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research* (5th edition), Table III, Oliver and Boyd Ltd., Edinburgh, by permission of the authors and publishers.

**Berechnung:**

Es sei $u = \left(1 - \dfrac{\alpha}{2}\right)$-Fraktil der $t$-Vert $(n-1)$, d.h. in der Tafel auf Seite 60 $u = t_{1-\frac{\alpha}{2}}$.




$$\Longrightarrow\ P\left(\left\{-u \leq \frac{\underline{\overline{x}} - \mu}{\underline{\hat{s}}/\sqrt{n}} \leq u\right\}\right) = 1 - \alpha$$

$$= P\left(\left\{\underbrace{\underline{\overline{x}} - \frac{\hat{s}u}{\sqrt{n}}}_{\underline{a}} \leq \mu \leq \underbrace{\underline{\overline{x}} + \frac{\hat{s}u}{\sqrt{n}}}_{\underline{b}}\right\}\right)$$

**Bsp.:** Betonproben, $\alpha = 5\,\%$, $u = \big[97.5\,\%$-Fraktil von $t$-Vert $(54)\big] = 2.006$ (denn Seite 60: $\left.\begin{array}{l} \nu = 40\ :\ t_{.975} = 2.02 \\ \nu = 60\ :\ t_{.975} = 2 \end{array}\right\} \Longrightarrow\ \nu = 54\ :\ t_{.975} = 2.006$) $\Longrightarrow$

$$\Longrightarrow\ \left.\begin{array}{c} a \\ b \end{array}\right\} = 2.44418 \mp \frac{0.035155 \cdot 2.006}{\sqrt{55}} \approx \left\{\begin{array}{l} 2.4347 \\ 2.4537 \end{array}\right.$$

Das korrekte Konfidenzintervall [2.4347, 2.4537] ist also etwas größer als in Seite 57, da durch die Unkenntnis von $\sigma$ die Unsicherheit etwas größer ist.

Für kleine $n$ unterscheiden sich $\left\{\begin{array}{l} A : \sigma, u \text{ aus } N(0,1) \\ C : \hat{s}, u \text{ aus } t\text{-Vert}\,(n-1) \end{array}\right\}$ mehr:

Die ersten 5 Betonproben geben mit $\alpha = 5\,\%$ für $\mathcal{E}(\underline{y})$ :

$\overline{y} = 393.8,\ s^2 = 1438.16,\ \hat{s} = \sqrt{\frac{5}{4}}\, s \approx 42.4$

$A$ : Annahme $\sigma = \hat{s},\ u = 1.96\ (\nu = \infty),\ [a,b] = [356.64,\ 430.96]$

$C$ : $\sigma$ unbekannt, $u = 2.78\ (\nu = 4),\ [a,b] = [341.09,\ 446.51]$.

# § 10 Punktschätzungen; Konfidenzintervalle für $\sigma^2$

**A) Idee:** Eine **Punktschätzung** ist eine Formel $f_n(\underline{x}_1, \cdots, \underline{x}_n)$ $(\forall n)$, um einen Verteilungsparameter (z.B. $\mu, \sigma^2, \lambda$ etc.) mittels einer Stichprobe zu schätzen.

**1. Forderung:** Wenn die Stichprobengröße $n \to \infty$, so sollte der Schätzwert gegen den zu schätzenden Parameter konvergieren (mit Wahrscheinlichkeit 1).

**Bsp. 1:** $\overline{\underline{x}} = \frac{1}{n} \sum_{j=1}^{n} \underline{x}_j$ ist eine Punktschätzung für $\mu = \mathcal{E}(\underline{x}_j)$

$$\begin{aligned} \text{Denn: } \mu &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \text{ Werte bei } j\text{-ter Ausführung} \\ &= \lim_{n\to\infty} \overline{\underline{x}}_{(n)}(\omega), \text{ vgl. Seite 34} \end{aligned}$$

(genaugenommen: $P\big(\{\lim_{n\to\infty} \overline{\underline{x}}_{(n)}(\omega) = \mu\}\big) = 1$, „starkes Gesetz der großen Zahlen”)

**Vorsicht:** $\overline{\underline{x}}_{\text{Unsinn}} = \frac{17}{\sqrt[3]{n}} + \frac{13}{n^2} + \frac{1}{n} \sum_{j=1}^{n} \underline{x}_j$ wäre auch eine Punktschätzung für $\mu$.

**Unterschied:** $\mathcal{E}(\overline{\underline{x}}) = \mu,\ \mathcal{E}(\overline{\underline{x}}_{\text{Unsinn}}) = \mu + \frac{17}{\sqrt[3]{n}} + \frac{13}{n^2} \neq \mu$, d.h. wenn **für festes $\boldsymbol{n}$** $\overline{\underline{x}}$ und $\overline{\underline{x}}_{\text{Unsinn}}$ immer wieder berechnet werden, liefert $\overline{\underline{x}}_{\text{Unsinn}}$ Unsinn.

**Def.:** Eine Punktschätzung heißt **erwartungstreu**, wenn
$\forall n : \mathcal{E}(\text{Schätzgröße}) = \text{zu schätzender Parameter}$.

**Ergebnis:** Das Stichprobenmittel $\overline{\underline{x}}$ ist eine erwartungstreue Punktschätzung für $\mu = \mathcal{E}(\underline{x}_j)$.

**Bsp. 2:** $\mathcal{V}(\underline{x}) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (a_j - \mu)^2$, siehe Seite 37
$\Longrightarrow\ \frac{1}{n} \sum_{j=1}^{n} (\underline{x}_j - \mu)^2$ ist eine Schätzgröße für $\sigma^2$.

Diese Punktschätzung ist erwartungstreu:

$$\mathcal{E}\left(\frac{1}{n} \sum_{j=1}^{n} (\underline{x}_j - \mu)^2\right) = \frac{1}{n} \sum_{j=1}^{n} \underbrace{\mathcal{E}\big((\underline{x}_j - \mu)^2\big)}_{=\mathcal{V}(\underline{x}_j)} = \frac{1}{n} \cdot n \cdot \sigma^2 = \sigma^2$$

Problem: Oft ist $\mu$ nicht bekannt.

Idee: Ersetze $\mu$ durch $\underline{\overline{x}}_{(n)}$, d.h. nehme

$$\begin{aligned}
\underline{s}^2 &= \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j - \underline{\overline{x}}_{(n)})^2 \\
&= \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j^2 - 2\underline{x}_j\underline{\overline{x}} + \underline{\overline{x}}^2) \quad = \quad \frac{1}{n}\sum \underline{x}_j^2 - 2\underline{\overline{x}}\cdot\overbrace{\frac{1}{n}\sum \underline{x}_j}^{\underline{\overline{x}}} + \underline{\overline{x}}^2 \\
&\qquad\qquad\qquad\qquad = \left(\frac{1}{n}\sum_{j=1}^{n}\underline{x}_j^2\right) - \underline{\overline{x}}^2 \qquad (*) \\
&(\text{für } n\to\infty) \quad \to \quad \mathcal{E}(\underline{x}_j^2) - \mathcal{E}(\underline{x}_j)^2 = \sigma^2
\end{aligned}$$

Also: $\underline{s}^2$ ist eine Punktschätzung für $\sigma^2$.

Aber: $\underline{s}^2$ ist nicht erwartungstreu (d.h. $\underline{s}^2$ entspricht $\underline{\overline{x}}_{\text{Unsinn}}$ in Bsp. 1), denn

$$\mathcal{E}(\underline{s}^2) \overset{(*)}{=} \frac{1}{n}\sum_{j=1}^{n}\mathcal{E}(\underline{x}_j^2) - \mathcal{E}(\underline{\overline{x}}^2) = \mathcal{E}(\underline{x}_j^2) - \mathcal{E}(\underline{\overline{x}}^2)$$

Allgemein gilt $\mathcal{E}(\underline{y}^2) = \mathcal{V}(\underline{y}) + \mathcal{E}(\underline{y})^2$

$$\begin{aligned}
\Longrightarrow \mathcal{E}(\underline{s}^2) &= \mathcal{V}(\underline{x}_j) + \mathcal{E}(\underline{x}_j)^2 - \mathcal{V}(\underline{\overline{x}}) - \mathcal{E}(\underline{\overline{x}})^2 \\
&= \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 = \frac{n-1}{n}\sigma^2
\end{aligned}$$

**Def.:**

$$\begin{aligned}
\underline{\hat{s}}^2 = \frac{1}{n-1}\sum_{j=1}^{n}(\underline{x}_j - \underline{\overline{x}})^2 \quad &= \quad \frac{n}{n-1}\,\underline{s}^2 \\
&\overset{(*)}{=} \left(\frac{1}{n-1}\sum_{j=1}^{n}\underline{x}_j^2\right) - \frac{n}{n-1}\,\underline{\overline{x}}^2
\end{aligned}$$

heißt **erwartungstreue Stichprobenvarianz** (engl.: unbiased estimate for $\sigma^2$).

**Bsp. 3:** $\underline{x}_1, \cdots, \underline{x}_n$; $\underline{y}_1, \cdots, \underline{y}_n$ seien zwei Stichproben von verschiedenen Größen (d.h. zu verschiedenen Verteilungen). Wieder ist $\underline{c} = \left(\frac{1}{n}\sum_{j=1}^{n}\underline{x}_j\underline{y}_j\right) - \underline{\overline{x}}\cdot\underline{\overline{y}}$ keine erwartungstreue Schätzgröße für $\gamma = \operatorname{cov}(\underline{x}_j, \underline{y}_k)$.
Stattdessen ist $\underline{\hat{c}} = \frac{n}{n-1}\,\underline{c}$ zu nehmen.

**B) Aufgabe:** Für gegebene Messdaten liefert $\hat{s}^2$ eine Schätzung für $\sigma^2$. Bestimme zu gegebenem $\alpha$ ein $(1-\alpha)$-Konfidenzintervall für $\sigma^2$ (zwei- oder einseitig).

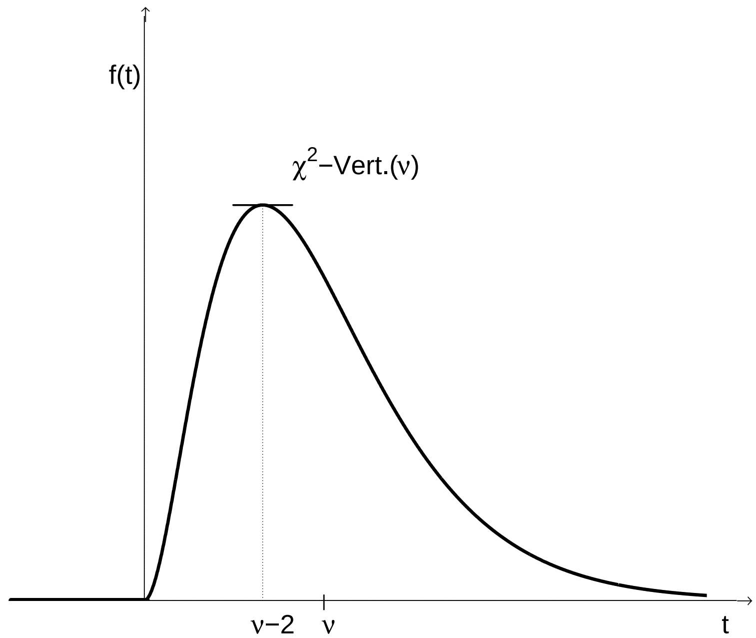
**Satz:** (ohne Beweis)
$\underline{x}_1, \cdots, \underline{x}_n$ Stichprobe mit $P_{\underline{x}_j} = N(\mu, \sigma)$, $\overline{\underline{x}} = \frac{1}{n}\sum\limits_{j=1}^{n} \underline{x}_j$, $\underline{\hat{s}}^2 = \frac{1}{n-1}\sum\limits_{j=1}^{n}(\underline{x}_j - \overline{\underline{x}})^2$.

Dann gilt: $\underline{z} := \frac{(n-1)\underline{\hat{s}}^2}{\sigma^2}$ hat die $\chi^2$-Vert.$(\overbrace{n-1}^{=\nu})$, wobei die Dichte der $\boldsymbol{\chi^2}$**-Verteilung mit $\boldsymbol{\nu}$ Freiheitsgraden** durch

$$f(t) = \begin{cases} \dfrac{t^{-1+\nu/2}\mathrm{e}^{-t/2}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} & : \quad t > 0 \\ 0 & : \quad t \leq 0 \end{cases}$$

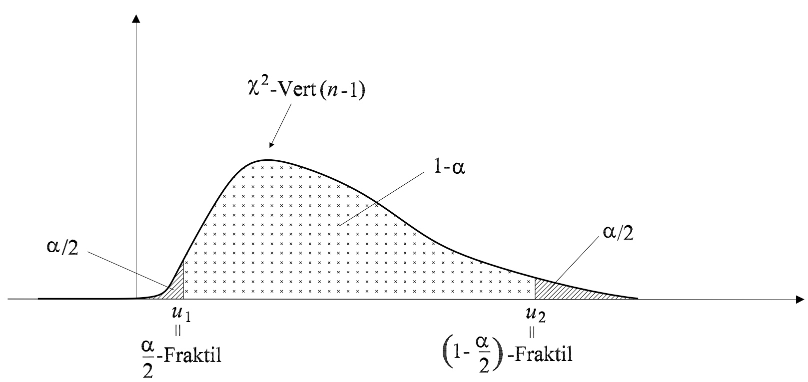
gegeben ist.

$\big(f(t)$ ist eigentlich die Dichte von $\underline{\chi}^2 = \underline{y}_1^2 + \cdots + \underline{y}_\nu^2$, wenn $\underline{y}_1, \cdots, \underline{y}_\nu$ unabhängig und $N(0,1)$-verteilt sind. Es gilt:
$\mathcal{E}(\underline{\chi}^2) = \nu$, $\mathcal{V}(\underline{\chi}^2) = 2\nu.\big)$




**Berechnung zweiseitiger Konfidenzintervalle:**

Ähnlich wie in § 9, A/C. Aber: Die $\chi^2$-Verteilung ist nicht symmetrisch und daher braucht man **zwei** Fraktilwerte.
Es seien $u_1$ das $\frac{\alpha}{2}$-Fraktil und $u_2$ das $\left(1-\frac{\alpha}{2}\right)$-Fraktil der $\chi^2$-Vert$(n-1)$, d.h. in der Tafel auf Seite 66 $u_1 = \chi^2_{\alpha/2}$ und $u_2 = \chi^2_{1-\frac{\alpha}{2}}$



$$\underset{\text{Satz}}{\Longrightarrow} P\left(\left\{u_1 \le \frac{(n-1)\underline{\hat{s}}^2}{\sigma^2} \le u_2\right\}\right) = 1-\alpha$$

$$= P\big(\{\sigma^2 u_1 \le (n-1)\underline{\hat{s}}^2 \le \sigma^2 u_2\}\big)$$

$$= P\left(\left\{\underbrace{\frac{(n-1)\underline{\hat{s}}^2}{u_2}}_{\underline{a}} \le \sigma^2 \le \underbrace{\frac{(n-1)\underline{\hat{s}}^2}{u_1}}_{\underline{b}}\right\}\right)$$

Somit: Mit Sicherheit $1-\alpha$ liegt $\sigma^2$ im Konfidenzintervall $[\underline{a}, \underline{b}] = \left[\frac{(n-1)\underline{\hat{s}}^2}{u_2}, \frac{(n-1)\underline{\hat{s}}^2}{u_1}\right]$.

**Bsp.** Dichtemessungen bei den Betonproben auf Seite 44, $\alpha = 5\,\%$.
$u_1 = \big[2.5\,\%\text{-Fraktil von } \chi^2\text{-Vert}(54)\big] = 35.64$, denn nach Tabelle auf Seite 66 ist

$$\begin{array}{l} \nu = 50 \,:\, \chi^2_{0.025} = 32.4 \\ \nu = 60 \,:\, \chi^2_{0.025} = 40.5 \end{array} \Longrightarrow \nu = 54 : \chi^2_{0.025} \approx 35.64$$

$u_2 = \big[97.5\,\%\text{-Fraktil von } \chi^2\text{-Vert}(54)\big] = 76.16$, denn

$$\begin{array}{l} \nu = 50 \,:\, \chi^2_{0.975} = 71.4 \\ \nu = 60 \,:\, \chi^2_{0.975} = 83.3 \end{array} \Longrightarrow \nu = 54 : \chi^2_{0.975} \approx 76.16$$

$\hat{s}^2 \approx 0.00123589$ (vgl. Seite 57)

$$\Longrightarrow a = \frac{54 \cdot \hat{s}^2}{u_2} \approx 0.00087629, \; b = \frac{54 \cdot \hat{s}^2}{u_1} \approx 0.00187256$$

$\Longrightarrow$ $\sigma$ liegt mit 95 % Sicherheit im Intervall $\big[\sqrt{a}, \sqrt{b}\,\big] = [0.0296, 0.0433]$.

**Bemerkung:** Für einseitige Konfidenzintervalle geht man analog zu § 9 B) vor und braucht nur **einen** Fraktilwert.

**Zusammenfassung:**

$\underline{x}_1, \cdots, \underline{x}_n$ Stichprobe mit $P_{\underline{x}_j} = N(\mu, \sigma) \Longrightarrow$

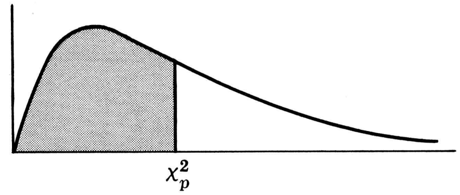
(i) $\dfrac{\overline{\underline{x}} - \mu}{\sigma/\sqrt{n}}$ $N(0,1)$-verteilt

(ii) $\dfrac{\overline{\underline{x}} - \mu}{\underline{\hat{s}}/\sqrt{n}}$ $t$-verteilt mit $n-1$ Freiheitsgraden

(iii) $\dfrac{(n-1)\underline{\hat{s}}^2}{\sigma^2}$ $\chi^2$-verteilt mit $n-1$ Freiheitsgraden

# Appendix IV

**PERCENTILE VALUES ($\chi^2_p$)**
**for**
**THE CHI-SQUARE DISTRIBUTION**
**with $\nu$ degrees of freedom**
(shaded area $= p$)




| $\nu$ | $\chi^2_{.995}$ | $\chi^2_{.99}$ | $\chi^2_{.975}$ | $\chi^2_{.95}$ | $\chi^2_{.90}$ | $\chi^2_{.75}$ | $\chi^2_{.50}$ | $\chi^2_{.25}$ | $\chi^2_{.10}$ | $\chi^2_{.05}$ | $\chi^2_{.025}$ | $\chi^2_{.01}$ | $\chi^2_{.005}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 7.88 | 6.63 | 5.02 | 3.84 | 2.71 | 1.32 | .455 | .102 | .0158 | .0039 | .0010 | .0002 | .0000 |
| 2 | 10.6 | 9.21 | 7.38 | 5.99 | 4.61 | 2.77 | 1.39 | .575 | .211 | .103 | .0506 | .0201 | .0100 |
| 3 | 12.8 | 11.3 | 9.35 | 7.81 | 6.25 | 4.11 | 2.37 | 1.21 | .584 | .352 | .216 | .115 | .072 |
| 4 | 14.9 | 13.3 | 11.1 | 9.49 | 7.78 | 5.39 | 3.36 | 1.92 | 1.06 | .711 | .484 | .297 | .207 |
| 5 | 16.7 | 15.1 | 12.8 | 11.1 | 9.24 | 6.63 | 4.35 | 2.67 | 1.61 | 1.15 | .831 | .554 | .412 |
| 6 | 18.5 | 16.8 | 14.4 | 12.6 | 10.6 | 7.84 | 5.35 | 3.45 | 2.20 | 1.64 | 1.24 | .872 | .676 |
| 7 | 20.3 | 18.5 | 16.0 | 14.1 | 12.0 | 9.04 | 6.35 | 4.25 | 2.83 | 2.17 | 1.69 | 1.24 | .989 |
| 8 | 22.0 | 20.1 | 17.5 | 15.5 | 13.4 | 10.2 | 7.34 | 5.07 | 3.49 | 2.73 | 2.18 | 1.65 | 1.34 |
| 9 | 23.6 | 21.7 | 19.0 | 16.9 | 14.7 | 11.4 | 8.34 | 5.90 | 4.17 | 3.33 | 2.70 | 2.09 | 1.73 |
| 10 | 25.2 | 23.2 | 20.5 | 18.3 | 16.0 | 12.5 | 9.34 | 6.74 | 4.87 | 3.94 | 3.25 | 2.56 | 2.16 |
| 11 | 26.8 | 24.7 | 21.9 | 19.7 | 17.3 | 13.7 | 10.3 | 7.58 | 5.58 | 4.57 | 3.82 | 3.05 | 2.60 |
| 12 | 28.3 | 26.2 | 23.3 | 21.0 | 18.5 | 14.8 | 11.3 | 8.44 | 6.30 | 5.23 | 4.40 | 3.57 | 3.07 |
| 13 | 29.8 | 27.7 | 24.7 | 22.4 | 19.8 | 16.0 | 12.3 | 9.30 | 7.04 | 5.89 | 5.01 | 4.11 | 3.57 |
| 14 | 31.3 | 29.1 | 26.1 | 23.7 | 21.1 | 17.1 | 13.3 | 10.2 | 7.79 | 6.57 | 5.63 | 4.66 | 4.07 |
| 15 | 32.8 | 30.6 | 27.5 | 25.0 | 22.3 | 18.2 | 14.3 | 11.0 | 8.55 | 7.26 | 6.26 | 5.23 | 4.60 |
| 16 | 34.3 | 32.0 | 28.8 | 26.3 | 23.5 | 19.4 | 15.3 | 11.9 | 9.31 | 7.96 | 6.91 | 5.81 | 5.14 |
| 17 | 35.7 | 33.4 | 30.2 | 27.6 | 24.8 | 20.5 | 16.3 | 12.8 | 10.1 | 8.67 | 7.56 | 6.41 | 5.70 |
| 18 | 37.2 | 34.8 | 31.5 | 28.9 | 26.0 | 21.6 | 17.3 | 13.7 | 10.9 | 9.39 | 8.23 | 7.01 | 6.26 |
| 19 | 38.6 | 36.2 | 32.9 | 30.1 | 27.2 | 22.7 | 18.3 | 14.6 | 11.7 | 10.1 | 8.91 | 7.63 | 6.84 |
| 20 | 40.0 | 37.6 | 34.2 | 31.4 | 28.4 | 23.8 | 19.3 | 15.5 | 12.4 | 10.9 | 9.59 | 8.26 | 7.43 |
| 21 | 41.4 | 38.9 | 35.5 | 32.7 | 29.6 | 24.9 | 20.3 | 16.3 | 13.2 | 11.6 | 10.3 | 8.90 | 8.03 |
| 22 | 42.8 | 40.3 | 36.8 | 33.9 | 30.8 | 26.0 | 21.3 | 17.2 | 14.0 | 12.3 | 11.0 | 9.54 | 8.64 |
| 23 | 44.2 | 41.6 | 38.1 | 35.2 | 32.0 | 27.1 | 22.3 | 18.1 | 14.8 | 13.1 | 11.7 | 10.2 | 9.26 |
| 24 | 45.6 | 43.0 | 39.4 | 36.4 | 33.2 | 28.2 | 23.3 | 19.0 | 15.7 | 13.8 | 12.4 | 10.9 | 9.89 |
| 25 | 46.9 | 44.3 | 40.6 | 37.7 | 34.4 | 29.3 | 24.3 | 19.9 | 16.5 | 14.6 | 13.1 | 11.5 | 10.5 |
| 26 | 48.3 | 45.6 | 41.9 | 38.9 | 35.6 | 30.4 | 25.3 | 20.8 | 17.3 | 15.4 | 13.8 | 12.2 | 11.2 |
| 27 | 49.6 | 47.0 | 43.2 | 40.1 | 36.7 | 31.5 | 26.3 | 21.7 | 18.1 | 16.2 | 14.6 | 12.9 | 11.8 |
| 28 | 51.0 | 48.3 | 44.5 | 41.3 | 37.9 | 32.6 | 27.3 | 22.7 | 18.9 | 16.9 | 15.3 | 13.6 | 12.5 |
| 29 | 52.3 | 49.6 | 45.7 | 42.6 | 39.1 | 33.7 | 28.3 | 23.6 | 19.8 | 17.7 | 16.0 | 14.3 | 13.1 |
| 30 | 53.7 | 50.9 | 47.0 | 43.8 | 40.3 | 34.8 | 29.3 | 24.5 | 20.6 | 18.5 | 16.8 | 15.0 | 13.8 |
| 40 | 66.8 | 63.7 | 59.3 | 55.8 | 51.8 | 45.6 | 39.3 | 33.7 | 29.1 | 26.5 | 24.4 | 22.2 | 20.7 |
| 50 | 79.5 | 76.2 | 71.4 | 67.5 | 63.2 | 56.3 | 49.3 | 42.9 | 37.7 | 34.8 | 32.4 | 29.7 | 28.0 |
| 60 | 92.0 | 88.4 | 83.3 | 79.1 | 74.4 | 67.0 | 59.3 | 52.3 | 46.5 | 43.2 | 40.5 | 37.5 | 35.5 |
| 70 | 104.2 | 100.4 | 95.0 | 90.5 | 85.5 | 77.6 | 69.3 | 61.7 | 55.3 | 51.7 | 48.8 | 45.4 | 43.3 |
| 80 | 116.3 | 112.3 | 106.6 | 101.9 | 96.6 | 88.1 | 79.3 | 71.1 | 64.3 | 60.4 | 57.2 | 53.5 | 51.2 |
| 90 | 128.3 | 124.1 | 118.1 | 113.1 | 107.6 | 98.6 | 89.3 | 80.6 | 73.3 | 69.1 | 65.6 | 61.8 | 59.2 |
| 100 | 140.2 | 135.8 | 129.6 | 124.3 | 118.5 | 109.1 | 99.3 | 90.1 | 82.4 | 77.9 | 74.2 | 70.1 | 67.3 |

*Source:* Catherine M. Thompson, *Table of percentage points of the $\chi^2$ distribution,* Biometrika, Vol. 32 (1941), by permission of the author and publisher.

Erwartungswert und Streuung von $\underline{\overline{x}}$ bzw. $\underline{\hat{s}}^2$ :

a) $\mathcal{E}(\underline{\overline{x}}) = \mu,\ \mathcal{V}(\underline{\overline{x}}) = \dfrac{\sigma^2}{n},\ \sigma(\underline{\overline{x}}) = \dfrac{\sigma}{\sqrt{n}}$

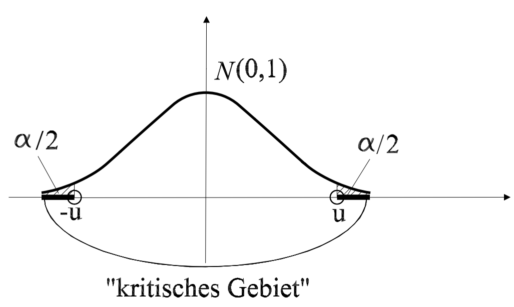
b) $\mathcal{E}(\underline{\hat{s}}^2) = \sigma^2, \quad \mathcal{V}(\underline{\hat{s}}^2) = \dfrac{\sigma^4}{(n-1)^2}\mathcal{V}(\underline{z}) = \dfrac{2\sigma^4}{n-1}, \quad \sigma(\underline{\hat{s}}^2) = \sqrt{\dfrac{2}{n-1}}\,\sigma^2.$

# § 11 Parametertests

**Zusammenfassung:** Eine Vermutung (=**Nullhypothese $\boldsymbol{H_0}$**) soll mittels Stichproben auf Plausibilität getestet werden. Dies nennt man **Hypothesentest**. Wir behandeln davon zwei Arten:

$$\begin{cases} \text{§ 11:} & \text{Parametertests} \\ \text{§ 12:} & \text{Anpassungstests} \end{cases}$$

**Einfachstes Beispiel eines Parametertests:** Test für $\mu$ bei bekanntem $\sigma$.
Annahme: $\underline{x}_1, \ldots, \underline{x}_n$ Stichprobe mit bekannter Varianz $\sigma^2$, unbekanntem Erwartungswert $\mu$.
Zu vorgegebenem Signifikanzniveau $\alpha$ (meist 5 %, 2 % oder 1 %) und vorgegebenem $\mu_0$ soll getestet werden, ob $\mu = \mu_0$ plausibel ist.
Also: „Nullhypothese" $H_0 : \mu = \mu_0$;
**falls $\boldsymbol{H_0}$ gilt**, ist $P_{\underline{\overline{x}}} = N\left(\mu_0, \dfrac{\sigma}{\sqrt{n}}\right)$ (nach dem zentralen Grenzwertsatz zumindest näherungsweise)
$\Longrightarrow$ die „Testgröße" $\underline{z} = \dfrac{\underline{\overline{x}} - \mu_0}{\sigma/\sqrt{n}}$ ist $N(0,1)$-verteilt; sei $u = \left(1 - \dfrac{\alpha}{2}\right)$-Fraktil von $N(0,1)$.




Wir „verwerfen" $H_0$, falls sich $z$ im „kritischen Gebiet" befindet, d.h. falls $|z| > u$ bzw. $|\overline{x} - \mu_0| > \dfrac{u\sigma}{\sqrt{n}}$. (Dann ist $\overline{x}$ so weit von $\mu_0$ weg, dass $H_0$ nicht plausibel ist.) Ansonsten wird $H_0$ „angenommen". Man sagt dann: $\mu$ und $\mu_0$ sind „nicht signifikant verschieden".

Zusammenhang mit dem Konfidenzintervall:

$[\underline{a}, \underline{b}] = \left[\overline{\underline{x}} - \frac{u\sigma}{\sqrt{n}}, \overline{\underline{x}} + \frac{u\sigma}{\sqrt{n}}\right]$, vgl. Seite 57;

$H_0$ annehmen $\iff |\underline{z}| \leq u \iff \left|\frac{\overline{\underline{x}} - \mu_0}{\sigma/\sqrt{n}}\right| \leq u \iff |\overline{\underline{x}} - \mu_0| \leq \frac{u\sigma}{\sqrt{n}} \iff \mu_0 \in$ Konfidenzintervall.

**Bsp.** $\alpha = 1\,\%,\ x_1, \cdots, x_{55}$ aus Seite 44, es sei $H_0 : \mu = \underbrace{2.43}_{\mu_0}$

Rechnung: $u = \big[99.5\,\%\text{-Fraktil von } N(0,1)\big] = 2.575$,

Annahme: $\sigma = \hat{s} = 0.035155$

$$\implies z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.44418 - 2.43}{0.035155/\sqrt{55}} \approx 2.9914$$

$|z| > u \implies$ wir verwerfen $H_0$, d.h. $\overline{x} = 2.44418$ ist signifikant anders, als nach $H_0$ zu erwarten wäre.

(Alternativ: $\mu_0 = 2.43 \notin$ Konfidenzintervall $= [2.432, 2.4564]$, siehe Seite 57 unten $\implies H_0$ ist zu verwerfen)
Hingegen wird z.B. die Hypothese $\tilde{H}_0 : \mu = 2.44$ „angenommen", d.h. sie kann von den Daten her nicht abgelehnt werden. $\tilde{H}_0$ muss deshalb noch lange nicht stimmen.

**Allgemeines:** $H_1$ = „Gegenteil von $H_0$" heißt **Alternativhypothese** (Oben wäre $H_1 : \mu \neq 2.43$). Es gibt zwei Fehlertypen:

a) **Fehler 1. Art:** $H_0$ wird verworfen, obwohl es stimmt
(oben: $\mu = 2.43,\ z \in$ kritischem Gebiet)

b) **Fehler 2. Art:** $H_0$ wird angenommen, obwohl es nicht stimmt
(oben: $\mu \neq 2.43,\ |z| \leq u$)

Tabelle:

| | $H_0$ wahr | $H_0$ falsch |
|---|---|---|
| $H_0$ verwerfen | Fehler 1. Art | richtige Entscheidung |
| $H_0$ annehmen | richtige Entscheidung | Fehler 2. Art |

Wesentlich und bei allen Tests gültig ist

$$\boxed{P(\text{„Fehler 1. Art“}) = \alpha}$$

denn unter der Voraussetzung, dass $H_0$ wahr ist, liegt $\underline{z}$ mit Wahrscheinlichkeit $\alpha$ im kritischen Gebiet, d.h. $P(\text{Ablehnung von } H_0 | H_0 \text{ wahr}) = \alpha$.

## A) $N$-Test

$(\mu_1 \stackrel{?}{=} \mu_2,\ \sigma_1 = \sigma_2 \text{ bekannt})$

**Gegeben:** Zwei unabhängige, normalverteilte Stichproben $\underline{x}_1, \cdots, \underline{x}_{n_1}$ bzw. $\underline{x}'_1, \cdots, \underline{x}'_{n_2}$ mit bekannter und gleicher Varianz $\sigma^2$. **Gefragt:** $\mathcal{E}(\underline{x}_j) \stackrel{?}{=} \mathcal{E}(\underline{x}'_j)$.
Also: $\underline{x}_1, \cdots, \underline{x}_{n_1},\ \underline{x}'_1, \cdots, \underline{x}'_{n_2}$ unabh., $P_{\underline{x}_j} = N(\mu_1, \sigma),\ P_{\underline{x}'_j} = N(\mu_2, \sigma),\ H_0 : \mu_1 = \mu_2$.

**Vorgangsweise:**

$$\overline{\underline{x}} = \frac{1}{n_1} \sum_{j=1}^{n_1} \underline{x}_j \cdots\cdots N\left(\mu_1, \frac{\sigma}{\sqrt{n_1}}\right)\text{-verteilt}$$

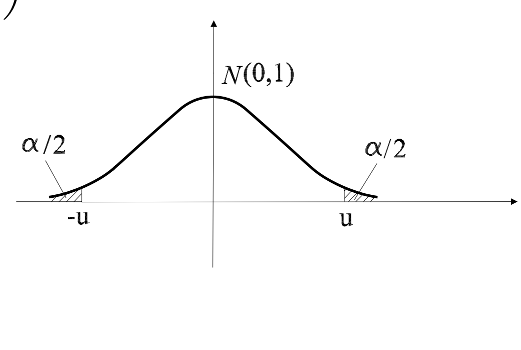
$$\overline{\underline{x}}' = \frac{1}{n_2} \sum_{j=1}^{n_2} \underline{x}'_j \cdots\cdots\cdots N\left(\mu_2, \frac{\sigma}{\sqrt{n_2}}\right)\text{-verteilt}$$

$$\Longrightarrow \text{(Satz 3, Seite 51)} \quad \Longrightarrow \overline{\underline{x}} - \overline{\underline{x}}' \cdots N\left(\mu_1 - \mu_2, \sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)\text{-verteilt}$$

$$\Longrightarrow \frac{\overline{\underline{x}} - \overline{\underline{x}}' - (\mu_1 - \mu_2)}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \cdots N(0,1)\text{-verteilt}$$

**falls $H_0$ gilt**, ist also $\underline{z} = \dfrac{\overline{\underline{x}} - \overline{\underline{x}}'}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $N(0,1)$-verteilt.

Setze $u = \left(1 - \dfrac{\alpha}{2}\right)$-Fraktil von $N(0,1)$



Wenn $|z| \leq u$, nehme $H_0$ an; wenn $|z| > u$, lehne $H_0$ ab.
Dann ist wieder $P$(„Fehler 1. Art“) =
$= P(H_0 \text{ ablehnen} \,|\, H_0 \text{ wahr}) = P\Big(\big\{|\underline{z}| > u\big\}\big|\mu_1 = \mu_2\Big)$
$= \alpha$ wie es sein soll.

**Bsp.** Betonproben in Seite 44. Ist der Mittelwert der ersten 25 Dichtemessungen signifikant von dem der folgenden 30 verschieden? ($\alpha = 5\,\%$)

Also $n_1 = 25,\ n_2 = 30,\ \overline{x} = \dfrac{1}{25}(x_1 + \cdots + x_{25}) \approx 2.4604$,

$\overline{x}' = \dfrac{1}{30}(\overbrace{x_{26}}^{=x_1'} + \cdots + \overbrace{x_{55}}^{=x_{30}'}) \approx 2.4307$,

$u = \big[97.5\,\%\text{-Quantil von } N(0,1)\big] = 1.96$,

Annahme: $\sigma = \hat{s} = 0.035155$

$$\Longrightarrow \quad z = \frac{\overline{x} - \overline{x}'}{\sigma \cdot \sqrt{\frac{1}{25} + \frac{1}{30}}} \approx 3.1232;$$

$|z| > u \implies$ wir lehnen die Nullhypothese $\mu_1 = \mu_2$ ab, d.h. die Mittelwerte sind signifikant verschieden, d.h. die Herstellungs- und Messbedingungen dürften sich im Verlauf der Messreihe etwas verändert haben.

## B) $t$-Test

($\mu_1 \overset{?}{=} \mu_2,\ \sigma_1 = \sigma_2$ unbekannt)

Vorgaben wie in A), aber nun ist $\sigma$ nicht bekannt. ($\sigma$ wird aber für $\underline{x}_1, \cdots, \underline{x}_{n_1}$ und $\underline{x}_1', \cdots \underline{x}_{n_2}'$ als gleich angenommen.)
Man nimmt als erwartungstreue Schätzgröße für $\sigma^2$ :

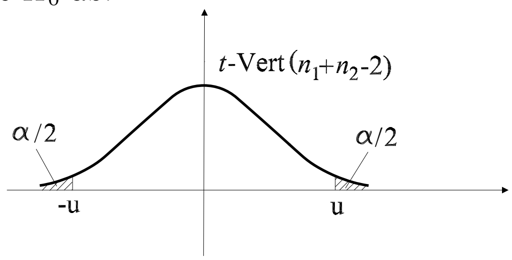
$$\underline{\hat{s}}_k^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_{j=1}^{n_1}(\underline{x}_j - \overline{\underline{x}})^2 + \sum_{j=1}^{n_2}(\underline{x}_j' - \overline{\underline{x}}')^2\right]$$

(vgl. Übung Z 12). Dies heißt **kombinierte Varianzschätzung**.
Weiters gilt (ohne Beweis)

$$\frac{\overline{\underline{x}} - \overline{\underline{x}}' - (\mu_1 - \mu_2)}{\underline{\hat{s}}_k \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{ist } t\text{-verteilt } (n_1 + n_2 - 2)$$

$\Longrightarrow$ **falls $\boldsymbol{H_0}$ gilt**, ist $\underline{z} = \dfrac{\underline{\overline{x}} - \underline{\overline{x}}'}{\underline{\hat{s}}_k \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $t$-vert$(n_1 + n_2 - 2)$;

$u = \left(1 - \dfrac{\alpha}{2}\right)$-Fraktil von $t$-Vert$(n_1 + n_2 - 2)$
Wenn $|z| \leq u$, nehme $H_0$ an,
wenn $|z| > u$, lehne $H_0$ ab.




**Bsp.** Betonproben wie Seite 70, eigentlich ist $\sigma$ dort unbekannt und man muss korrekterweise den $t$-Test nehmen.
$\overline{x} = 2.4604,\ \overline{x}' = 2.4307$;
Kombinierte Varianzschätzung:

$$\hat{s}_k^2 = \frac{1}{25 + 30 - 2}\left[\sum_{j=1}^{25}(x_j - \overline{x})^2 + \sum_{j=26}^{55}(x_j - \overline{x}')^2\right]$$
$$\approx 0.0010317 \Longrightarrow \hat{s}_k \approx 0.032121$$

$$u = \left[97.5\,\%\text{-Fraktil von } t\text{-Vert}(53)\right] \overset{\text{s.S.60}}{=} 2.007$$

$$z = \frac{\overline{x} - \overline{x}'}{\hat{s}_k \cdot \sqrt{\frac{1}{25} + \frac{1}{30}}} \approx 3.4183$$

$|z| > u \Longrightarrow$ lehne $\mu_1 = \mu_2$ ab.

## C) $F$-Test

$(\sigma_1 \overset{?}{=} \sigma_2,\ \mu_1, \mu_2$ unbekannt)

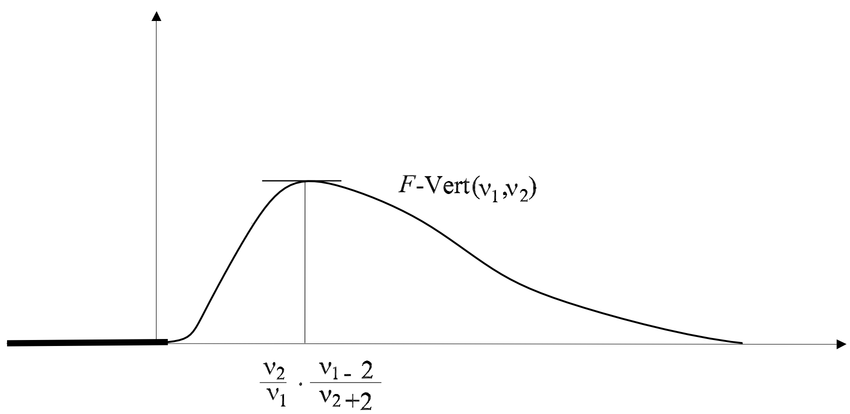
**Gegeben:** Zwei unabhängige, normalverteilte Stichproben $\underline{x}_1, \cdots, \underline{x}_{n_1}$ bzw. $\underline{x}'_1, \cdots, \underline{x}'_{n_2}$.
$H_0: \ \mathcal{V}(\underline{x}_j) = \mathcal{V}(\underline{x}'_j)$

**Vorgangsweise:** $\underline{\overline{x}} = \dfrac{1}{n_1}\displaystyle\sum_{j=1}^{n_1} \underline{x}_j, \quad \underline{\overline{x}}' = \dfrac{1}{n_2}\displaystyle\sum_{j=1}^{n_2} \underline{x}'_j,$

$$\underline{\hat{s}}^2 = \frac{1}{n_1 - 1}\sum_{j=1}^{n_1}(\underline{x}_j - \overline{\underline{x}})^2, \quad \underline{\hat{s}}'^2 = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}(\underline{x}'_j - \overline{\underline{x}}')^2$$

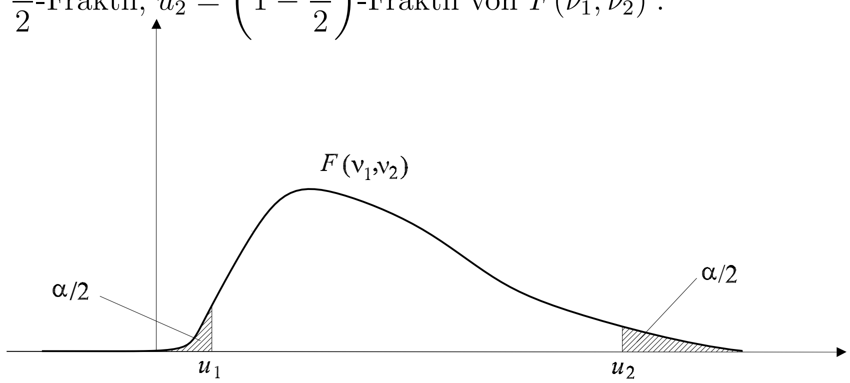
**Falls $\boldsymbol{\sigma_1 = \sigma_2}$,**

ist $\underline{\hat{s}}^2/\underline{\hat{s}}'^2$ ***F*-verteilt** (nach R.A. Fisher) mit $(\underbrace{n_1 - 1}_{\nu_1}, \underbrace{n_2 - 1}_{\nu_2})$ Freiheitsgraden.




Dichte: siehe Seite 39.

Dabei sind $\nu_1 = n_1 - 1$ die **Freiheitsgrade des Zählers** (ndld. = teller) und $\nu_2 = n_2 - 1$ die **Freiheitsgrade des Nenners** (ndld. = noemer).
Es sei $u_1 = \frac{\alpha}{2}$-Fraktil, $u_2 = \left(1 - \frac{\alpha}{2}\right)$-Fraktil von $F(\nu_1, \nu_2)$ :




Aus Symmetriegründen gilt

$$\left[\frac{1}{u_1} \text{ von } F(\nu_1, \nu_2)\right] = \left[u_2 \text{ von } F(\nu_2, \nu_1)\right].$$

Daher ist in Seite 74 nur $u_2$ tabelliert.

**Wenn $\boldsymbol{\sigma_1 = \sigma_2}$**, so ist $P\left(\left\{\frac{\underline{\hat{s}}^2}{\underline{\hat{s}}'^2} > u_2\right\}\right) = \frac{\alpha}{2}$ und

$$\frac{\alpha}{2} = P\left(\left\{\frac{\underline{\hat{s}}^2}{\underline{\hat{s}}'^2} < u_1\right\}\right) = P\left(\left\{\frac{\underline{\hat{s}}'^2}{\underline{\hat{s}}^2} > \frac{1}{u_1} = \big[u_2 \text{ von } F(n_2-1, n_1-1)\big]\right\}\right).$$

Wenn wir daher festlegen:

$$\text{Testgröße } \underline{z} = \begin{cases} \dfrac{\underline{\hat{s}}^2}{\underline{\hat{s}}'^2} & : \quad \text{falls } \underline{\hat{s}} \geq \underline{\hat{s}}' \\[2ex] \dfrac{\underline{\hat{s}}'^2}{\underline{\hat{s}}^2} & : \quad \text{falls } \underline{\hat{s}}' > \underline{\hat{s}} \end{cases}$$

$$\text{und } u = \begin{cases} u_2 & = & \left[\left(1-\dfrac{\alpha}{2}\right)\text{-Fraktil } F(n_1-1, n_2-1)\right] & : \quad \text{falls } \underline{\hat{s}} \geq \underline{\hat{s}}' \\[2ex] \dfrac{1}{u_1} & = & \left[\left(1-\dfrac{\alpha}{2}\right)\text{-Fraktil } F(n_2-1, n_1-1)\right] & : \quad \text{falls } \underline{\hat{s}}' > \underline{\hat{s}} \end{cases}$$

$$\text{und } \begin{cases} \underline{z} \leq u \implies \text{nehme} & [H_0 : \sigma_1^2 = \sigma_2^2] \text{ an} \\ \underline{z} > u \implies \text{lehne} & [H_0 : \sigma_1^2 = \sigma_2^2] \text{ ab,} \end{cases}$$

so gilt, wie es sein muss,

$P(\text{„Fehler 1. Art“}) = P(\text{Ablehnung}|\sigma_1 = \sigma_2) =$

$$\begin{aligned} = P\big(\{\underline{z} > u\}|\sigma_1 = \sigma_2\big) \quad &= \quad P\left(\left\{\frac{\underline{\hat{s}}^2}{\underline{\hat{s}}'^2} > u_2 \text{ oder } \frac{\underline{\hat{s}}^2}{\underline{\hat{s}}'^2} < u_1\right\}\middle|\sigma_1 = \sigma_2\right) \\ &= \quad \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha \end{aligned}$$

**Bsp.** Betonproben. In Seite 70/71 wurde $\sigma_1 = \sigma_2$ vorausgesetzt. Dies wollen wir mit $\alpha = 5\,\%$ testen.

Es ist $\overline{x} = 2.4604$, $\overline{x}' = 2.4307$,

$\hat{s}^2 = \frac{1}{24}\sum_{j=1}^{25}(x_j - \overline{x})^2 = 0.001220\dot{6}$, $\hat{s}'^2 = \frac{1}{29}\sum_{j=26}^{55}(x_j - \overline{x}')^2 \approx 0.0008754$

$\hat{s} \geq \hat{s}' \implies z = \frac{\hat{s}^2}{\hat{s}'^2} \approx 1.3944$

$\hat{s} \geq \hat{s}' \implies u = \big[97.5\,\%\text{-Fraktil von } F(\overset{\text{teller}}{\downarrow}\!24,\ \overset{\text{noemer}}{\downarrow}\!29\ )\big]$

$$\text{(Seite 74): } \left.\begin{cases} \text{zu } F(24, 25): & u = 2.24 \\ \text{zu } F(24, 30): & u = 2.14 \end{cases}\right\} \implies u = 2.16$$

$z < u \implies$ nehme $[H_0 : \sigma_1^2 = \sigma_2^2]$ an.

# Kritieke waarden $\alpha/2 = 0.025$

$P(F_{10}^{7} < 3.95) = 0.975$

| Vrijheidsgraden van de noemer | Vrijheidsgraden van de teller: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | $\infty$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 648 | 800 | 864 | 900 | 922 | 937 | 948 | 957 | 963 | 969 | 977 | 985 | 993 | 997 | 1,001 | 1,006 | 1,010 | 1,014 | 1,018 |
| 2 | 38.5 | 39.0 | 39.2 | 39.2 | 39.3 | 39.3 | 39.4 | 39.4 | 39.4 | 39.4 | 39.4 | 39.4 | 39.4 | 39.5 | 39.5 | 39.5 | 39.5 | 39.5 | 39.5 |
| 3 | 17.4 | 16.0 | 15.4 | 15.1 | 14.9 | 14.7 | 14.6 | 14.5 | 14.5 | 14.4 | 14.3 | 14.3 | 14.2 | 14.1 | 14.1 | 14.0 | 14.0 | 13.9 | 13.9 |
| 4 | 12.2 | 10.6 | 9.98 | 9.60 | 9.36 | 9.20 | 9.07 | 8.98 | 8.90 | 8.84 | 8.75 | 8.66 | 8.56 | 8.51 | 8.46 | 8.41 | 8.36 | 8.31 | 8.26 |
| 5 | 10.0 | 8.43 | 7.76 | 7.39 | 7.15 | 6.98 | 6.85 | 6.76 | 6.68 | 6.62 | 6.52 | 6.43 | 6.33 | 6.28 | 6.23 | 6.18 | 6.12 | 6.07 | 6.02 |
| 6 | 8.81 | 7.26 | 6.60 | 6.23 | 5.99 | 5.82 | 5.70 | 5.60 | 5.52 | 5.46 | 5.37 | 5.27 | 5.17 | 5.12 | 5.07 | 5.01 | 4.96 | 4.90 | 4.85 |
| 7 | 8.07 | 6.54 | 5.89 | 5.52 | 5.29 | 5.12 | 4.99 | 4.90 | 4.82 | 4.76 | 4.67 | 4.57 | 4.47 | 4.42 | 4.36 | 4.31 | 4.25 | 4.20 | 4.14 |
| 8 | 7.57 | 6.06 | 5.42 | 5.05 | 4.82 | 4.65 | 4.53 | 4.43 | 4.36 | 4.30 | 4.20 | 4.10 | 4.00 | 3.95 | 3.89 | 3.84 | 3.78 | 3.73 | 3.67 |
| 9 | 7.21 | 5.71 | 5.08 | 4.72 | 4.48 | 4.32 | 4.20 | 4.10 | 4.03 | 3.96 | 3.87 | 3.77 | 3.67 | 3.61 | 3.56 | 3.51 | 3.45 | 3.39 | 3.33 |
| 10 | 6.94 | 5.46 | 4.83 | 4.47 | 4.24 | 4.07 | 3.95 | 3.85 | 3.78 | 3.72 | 3.62 | 3.52 | 3.42 | 3.37 | 3.31 | 3.26 | 3.20 | 3.14 | 3.08 |
| 11 | 6.72 | 5.26 | 4.63 | 4.28 | 4.04 | 3.88 | 3.76 | 3.66 | 3.59 | 3.53 | 3.43 | 3.33 | 3.23 | 3.17 | 3.12 | 3.06 | 3.00 | 2.94 | 2.88 |
| 12 | 6.55 | 5.10 | 4.47 | 4.12 | 3.89 | 3.73 | 3.61 | 3.51 | 3.44 | 3.37 | 3.28 | 3.18 | 3.07 | 3.02 | 2.96 | 2.91 | 2.85 | 2.79 | 2.72 |
| 13 | 6.41 | 4.97 | 4.35 | 4.00 | 3.77 | 3.60 | 3.48 | 3.39 | 3.31 | 3.25 | 3.15 | 3.05 | 2.95 | 2.89 | 2.84 | 2.78 | 2.72 | 2.66 | 2.60 |
| 14 | 6.30 | 4.86 | 4.24 | 3.89 | 3.66 | 3.50 | 3.38 | 3.29 | 3.21 | 3.15 | 3.05 | 2.95 | 2.84 | 2.79 | 2.73 | 2.67 | 2.61 | 2.55 | 2.49 |
| 15 | 6.20 | 4.77 | 4.15 | 3.80 | 3.58 | 3.41 | 3.29 | 3.20 | 3.12 | 3.06 | 2.96 | 2.86 | 2.76 | 2.70 | 2.64 | 2.59 | 2.52 | 2.46 | 2.40 |
| 16 | 6.12 | 4.69 | 4.08 | 3.73 | 3.50 | 3.34 | 3.22 | 3.12 | 3.05 | 2.99 | 2.89 | 2.79 | 2.68 | 2.63 | 2.57 | 2.51 | 2.45 | 2.38 | 2.32 |
| 17 | 6.04 | 4.62 | 4.01 | 3.66 | 3.44 | 3.28 | 3.16 | 3.06 | 2.98 | 2.92 | 2.82 | 2.72 | 2.62 | 2.56 | 2.50 | 2.44 | 2.38 | 2.32 | 2.25 |
| 18 | 5.98 | 4.56 | 3.95 | 3.61 | 3.38 | 3.22 | 3.10 | 3.01 | 2.93 | 2.87 | 2.77 | 2.67 | 2.56 | 2.50 | 2.44 | 2.38 | 2.32 | 2.26 | 2.19 |
| 19 | 5.92 | 4.51 | 3.90 | 3.56 | 3.33 | 3.17 | 3.05 | 2.96 | 2.88 | 2.82 | 2.72 | 2.62 | 2.51 | 2.45 | 2.39 | 2.33 | 2.27 | 2.20 | 2.13 |
| 20 | 5.87 | 4.46 | 3.86 | 3.51 | 3.29 | 3.13 | 3.01 | 2.91 | 2.84 | 2.77 | 2.68 | 2.57 | 2.46 | 2.41 | 2.35 | 2.29 | 2.22 | 2.16 | 2.09 |
| 21 | 5.83 | 4.42 | 3.82 | 3.48 | 3.25 | 3.09 | 2.97 | 2.87 | 2.80 | 2.73 | 2.64 | 2.53 | 2.42 | 2.37 | 2.31 | 2.25 | 2.18 | 2.11 | 2.04 |
| 22 | 5.79 | 4.38 | 3.78 | 3.44 | 3.22 | 3.05 | 2.93 | 2.84 | 2.76 | 2.70 | 2.60 | 2.50 | 2.39 | 2.33 | 2.27 | 2.21 | 2.14 | 2.08 | 2.00 |
| 23 | 5.75 | 4.35 | 3.75 | 3.41 | 3.18 | 3.02 | 2.90 | 2.81 | 2.73 | 2.67 | 2.57 | 2.47 | 2.36 | 2.30 | 2.24 | 2.18 | 2.11 | 2.04 | 1.97 |
| 24 | 5.72 | 4.32 | 3.72 | 3.38 | 3.15 | 2.99 | 2.87 | 2.78 | 2.70 | 2.64 | 2.54 | 2.44 | 2.33 | 2.27 | 2.21 | 2.15 | 2.08 | 2.01 | 1.94 |
| 25 | 5.69 | 4.29 | 3.69 | 3.35 | 3.13 | 2.97 | 2.85 | 2.75 | 2.68 | 2.61 | 2.51 | 2.41 | 2.30 | 2.24 | 2.18 | 2.12 | 2.05 | 1.98 | 1.91 |
| 30 | 5.57 | 4.18 | 3.59 | 3.25 | 3.03 | 2.87 | 2.75 | 2.65 | 2.57 | 2.51 | 2.41 | 2.31 | 2.20 | 2.14 | 2.07 | 2.01 | 1.94 | 1.87 | 1.79 |
| 40 | 5.42 | 4.05 | 3.46 | 3.13 | 2.90 | 2.74 | 2.62 | 2.53 | 2.45 | 2.39 | 2.29 | 2.18 | 2.07 | 2.01 | 1.94 | 1.88 | 1.80 | 1.72 | 1.64 |
| 60 | 5.29 | 3.93 | 3.34 | 3.01 | 2.79 | 2.63 | 2.51 | 2.41 | 2.33 | 2.27 | 2.17 | 2.06 | 1.94 | 1.88 | 1.82 | 1.74 | 1.67 | 1.58 | 1.48 |
| 120 | 5.15 | 3.80 | 3.23 | 2.89 | 2.67 | 2.52 | 2.39 | 2.30 | 2.22 | 2.16 | 2.05 | 1.94 | 1.82 | 1.76 | 1.69 | 1.61 | 1.53 | 1.43 | 1.31 |
| $\infty$ | 5.02 | 3.69 | 3.12 | 2.79 | 2.57 | 2.41 | 2.29 | 2.19 | 2.11 | 2.05 | 1.94 | 1.83 | 1.71 | 1.64 | 1.57 | 1.48 | 1.39 | 1.27 | 1.00 |

THEORIE PAGINA 12, 13 - UITVOERIGE TABEL IN [9] [3] [1]

## D) Ergänzungen

a) **Einseitige Tests:** In gewissen Fällen ist z.B. $\mu_1 \geq \mu_2$ von vornherein bekannt. Man setzt dann immer noch $H_0: \ \mu_1 = \mu_2$, aber $H_1: \ \mu_1 > \mu_2$, d.h. also $\mu_1 < \mu_2$ wird von vornherein ausgeschlossen.
Dann ist der kritische Bereich von

$$\underline{z} = \frac{\overline{\underline{x}} - \overline{\underline{x}}'}{\underline{\hat{s}}_k \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \ \big(\text{vgl. B)}\big) \text{ nur } ]u, \infty[ \ \text{ da } \underline{z} < -u \text{ zum Schluss } \mu_1 < \mu_2 \text{ führen}$$

würde. Man setzt daher $u = (1-\alpha)$-Fraktil von $t$-Vert$(n_1+n_2-2)$ und entscheidet so:

$$\begin{cases} z > u \implies \text{lehne} & H_0 \text{ ab} \\ z \leq u \implies \text{nehme} & H_0 \text{ an} \end{cases}$$

Dann ist wieder $P(\text{„Fehler 1. Art“}) = \alpha$.

b) **Welch-Test:** $\mu_1 \stackrel{?}{=} \mu_2, \ \sigma_1 \neq \sigma_2$ unbekannt.
Falls der $F$-Test $\big(\text{C)}\big)$ ergibt, dass $\sigma_1 \neq \sigma_2$, so ist der $t$-Test $\big(\text{B)}\big)$ nicht anwendbar;

$$\mathcal{V}(\overline{\underline{x}} - \overline{\underline{x}}') = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \approx \frac{\underline{\hat{s}}^2}{n_1} + \frac{\underline{\hat{s}}'^2}{n_2}, \text{ wobei}$$

$$\underline{\hat{s}}^2 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (\underline{x}_j - \overline{\underline{x}})^2, \ \underline{\hat{s}}'^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (\underline{x}'_j - \overline{\underline{x}}')^2.$$

Unter der Hypothese $H_0 : \mu_1 = \mu_2$ ist dann

$$\underline{z} = \frac{\overline{\underline{x}} - \overline{\underline{x}}'}{\sqrt{\frac{\underline{\hat{s}}^2}{n_1} + \frac{\underline{\hat{s}}'^2}{n_2}}} \approx t\text{-Vert}\,(f), \text{ wobei}$$

$$f \approx \frac{1}{\frac{c^2}{n_1-1} + \frac{(1-c)^2}{n_2-1}} \qquad \text{und} \qquad c = \frac{\hat{s}^2 / n_1}{\hat{s}^2 / n_1 + \hat{s}'^2 / n_2}.$$

c) Für A), B), C) wurde vorausgesetzt, dass (wie in der Technik meist der Fall) $\underline{x}_j, \underline{x}'_j$ normalverteilt sind. (In § 12 werden wir dazu einen Test angeben.) Falls $\underline{x}_j, \underline{x}'_j$ nicht normalverteilt sind, verwendet man zum Test $\mu_1 \stackrel{?}{=} \mu_2$ sog. „parameterfreie Prüfverfahren“: $U$-Test, Kolmogorov-Smirnov-Test etc.

# § 12 Anpassungstests

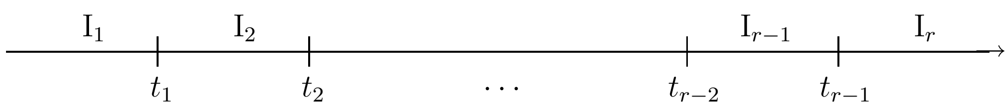
Bei Parametertests wird **vorausgesetzt**, **dass** die Stichproben in gewisser Weise (meist normal) verteilt sind, und es werden Hypothesen über die Parameter dieser Verteilungen getestet.
Bei einem Anpassungstest wird **untersucht**, **ob** eine Stichprobe einer vorgegebenen Verteilung $P_0$ entspricht.

**Bsp. 1:** 300 Würfe mit einem Würfel ergaben für die Augenzahlen $1, 2, \cdots, 6$ die jeweiligen Häufigkeiten 40,55,51,49,46,59. Ist der Würfel in Ordnung, d.h. ist das mit $H_0: \ p_1 = p_2 = \cdots = p_6 = \dfrac{1}{6}$ (siehe Seiten 2/3) vereinbar?

**Zum $\chi^2$-Test führende Prozedur:**

$\underline{x}_1, \cdots, \underline{x}_n$ Stichprobe, $\alpha$, $P_0$ vorgegeben, $H_0: \ P_{\underline{x}_j} = P_0$. Bestimme $\underline{z}$, $u$ so, dass
$P(\text{„Fehler 1. Art“}) = P\big(\{\underline{z} > u\} | H_0\big) = \alpha$.

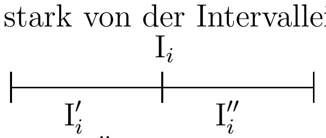
Teile $\mathbb{R}$ in $r$ Intervalle




Es sei $p_i = P_0(\mathrm{I}_i) =$ Wahrscheinlichkeit von $\mathrm{I}_i$ nach Verteilung $P_0$,
$\underline{n}_i = \#\{j : \underline{x}_j \in \mathrm{I}_i\} =$ Anzahl der Messwerte in $\mathrm{I}_i$

Idee: Vergleiche $p_i$ mit $\dfrac{\underline{n}_i}{n}$.

1. Versuch: $\displaystyle\sum_{i=1}^{r} \left(\frac{\underline{n}_i}{n} - p_i\right)^2$

   Nachteil: Das hängt stark von der Intervalleinteilung ab.
   (Wenn $\mathrm{I}_i = \mathrm{I}_i' \cup \mathrm{I}_i''$, ,
   
   $\implies \dfrac{n_i}{n} - p_i = \dfrac{n_i'}{n} - p_i' + \dfrac{n_i''}{n} - p_i''$; wenn z.B.
   $\dfrac{n_i'}{n} - p_i' = \dfrac{n_i''}{n} - p_i'' = \dfrac{1}{2}\left(\dfrac{n_i}{n} - p_i\right) \implies \left(\dfrac{n_i'}{n} - p_i'\right)^2 + \left(\dfrac{n_i''}{n} - p_i''\right)^2 = \dfrac{1}{2}\left(\dfrac{n_i}{n} - p_i\right)^2 \implies$
   die Summe wird viel kleiner)

2. Versuch: $\sum\limits_{i=1}^{r} \dfrac{\left(\frac{\underline{n}_i}{n} - p_i\right)^2}{p_i} = \dfrac{1}{n} \sum\limits_{i=1}^{r} \dfrac{(\underline{n}_i - np_i)^2}{np_i}$

Nachteil: Falls $H_0$ gilt $\Longrightarrow$ $\dfrac{\underline{n}_i}{n} \to p_i$ für $n \to \infty$ $\Longrightarrow$ $\sum\limits_{i=1}^{r} \cdots \to 0$

3. Versuch: Lasse $\dfrac{1}{n}$ vor der letzten $\sum\limits_{i=1}^{r}$ weg. Das führt zum Ziel!

**Satz:** (ohne Beweis) Falls $H_0$ gilt, so ist

$$\text{für festes } \mathrm{I}_1, \cdots, \mathrm{I}_r \qquad \boxed{\underline{z} = \sum_{i=1}^{r} \frac{(\underline{n}_i - np_i)^2}{np_i}}$$

im Limes $n \to \infty$ $\chi^2$-verteilt mit $r-1$ Freiheitsgraden, d.h.

$$\lim_{n\to\infty} P_{\underline{z}(n)} = \chi^2\text{-Vert}\,(r-1).$$

**Zusatz:** Wenn $P_0$ nur bis auf Parameter gegeben ist (z.B. $P_0 = N(\mu, \sigma),\ \mu, \sigma$ unbekannt) und diese Parameter aus der Stichprobe geschätzt werden (z.B. $\mu \approx \overline{\underline{x}},\ \sigma \approx \hat{\underline{s}}$), so gehen die Stichprobenwerte auch in die Berechnung der $p_i$ ein. Dann verringert sich der Freiheitsgrad $\nu$ der $\chi^2$-Verteilung pro Parameter um 1.

## $\chi^2$-Test

Gegeben $\alpha$, $P_0$ und $\underline{x}_1, \cdots, \underline{x}_n$. Wenn nötig, schätze $f$ fehlende Parameter aus $\underline{x}_1, \cdots, \underline{x}_n$. Wähle $\mathrm{I}_1, \cdots, \mathrm{I}_r$ klein, aber so, dass $\boxed{\forall i:\ np_i \geq 5}$. (Sonst ist $P_{\underline{z}(n)}$ noch zu weit von $\lim\limits_{n\to\infty} P_{\underline{z}(n)}$ entfernt.)

Bilde $\underline{z} = \sum\limits_{i=1}^{r} \dfrac{(\underline{n}_i - np_i)^2}{np_i}$, wobei $\underline{n}_i = \#\{j : \underline{x}_j \in \mathrm{I}_i\}$ und $p_i = P_0(\mathrm{I}_i)$.

**Falls $\boldsymbol{H_0 : P_{\underline{x}_j} = P_0}$ gilt** $\underset{\text{Satz}}{\Longrightarrow}$ $\underline{z} \approx \chi^2\text{-vert}(r-1-f)$
Bestimme $u = \big[(1-\alpha)\text{-Fraktil von } \chi^2\text{-Vert}(r-1-f)\big]$
$\underline{z} > u \implies$ lehne $H_0$ ab
$\underline{z} \leq u \implies$ nehme $H_0$ an.
Dann gilt wieder $P(H_0 \text{ abgelehnt} | H_0 \text{ wahr}) = \alpha$.

**Bsp. 1:** $n = 300$ mal würfeln wie in Seite 76, $\alpha = 5\,\%$, $P_0 =$ Gleichverteilung auf $\{1, 2, \cdots, 6\}$.

Setze

$$\begin{array}{cccccccccccc} & \mathrm{I}_1 & & \mathrm{I}_2 & & \mathrm{I}_3 & & \mathrm{I}_4 & & \mathrm{I}_5 & & \mathrm{I}_6 \\ \hline & & 1.5 = t_1 & & 2.5 & & 3.5 & & 4.5 & & 5.5 & \end{array}$$

$\implies p_1 = \cdots = p_6 = \dfrac{1}{6};\ n \cdot p_i = 300 \cdot \dfrac{1}{6} = 50 \geq 5\surd$

$n_1 = 40,\ n_2 = 55,\ n_3 = 51,\ n_4 = 49,\ n_5 = 46,\ n_6 = 59$

$$\implies z = \sum_{i=1}^{6} \frac{(n_i - np_i)^2}{np_i} = \frac{\left(40 - \frac{300}{6}\right)^2}{\frac{300}{6}} + \cdots + \frac{\left(59 - \frac{300}{6}\right)^2}{\frac{300}{6}} =$$

$$= \frac{1}{50}\left[10^2 + 5^2 + 1^2 + 1^2 + 4^2 + 9^2\right] = \frac{224}{50} = 4.48$$

$u = [95\,\%\text{-Fraktil von } \chi^2\text{-Vert}(5)] \overset{\text{s.S. 66}}{=} 11.1$

$4.48 < 11.1 \implies$ nehme $P = P_0$ an.

($\chi^2_{.50} = 4.35 \implies H_0 : P = P_0$ lässt sich nur mit $\approx 50\,\%$ Irrtumswahrscheinlichkeit ablehnen.)

**Bsp. 2:** Sind die $x_i$ aus Seite 44 normalverteilt?

Zuerst schätze $\mu \approx \overline{x} = 2.44418,\ \sigma \approx \hat{s} = 0.035155 \implies \nu =$ Anzahl der Freiheitsgrade $= r - 1 - 2 = r - 3$.

**1. Versuch der Intervalleinteilung:**

| $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I_9$ | $I_{10}$ |
|---|---|---|---|---|---|---|---|---|---|
| $t_1 =$ 2.405 | 2.415 | 2.425 | 2.435 | 2.445 $\approx \mu$ | 2.455 | 2.465 | 2.475 | 2.485 $= t_9$ | |

$$p_1 = P_{N(\mu,\sigma)}\big(]-\infty, 2.405]\big) = \Phi\left(\frac{2.405 - \mu}{\sigma}\right) - \Phi(-\infty)$$

$$= \frac{1}{2} - \Phi\left(\frac{2.44418 - 2.405}{0.035155}\right) \overset{\text{s.S. 22}}{=} \frac{1}{2} - 0.3675 = 0.1325$$

$np_1 = 55 \cdot p_1 = 7.288 > 5\surd$

$$p_2 = P_{N(\mu,\sigma)}\big([2.405, 2.415]\big) = 0.3675 - \underbrace{\Phi(0.83)}_{0.2967} = 0.0708$$

$np_2 = 3.894 < 5$ !!!!

**Daher 2. Versuch der Intervalleinteilung:**

Wir legen $I_2$ und $I_3$ sowie $I_8$ und $I_9$ zusammen:

| $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ |
|---|---|---|---|---|---|---|---|
| $t_1 =$ 2.405 | 2.425 | 2.435 | 2.445 $\approx \mu$ | 2.455 | 2.465 | 2.485 $= t_7$ | |

$\implies p_2 = P_0\big([2.405, 2.425]\big) = 0.3675 - \underbrace{\Phi(0.546)}_{0.2073} = 0.1602$

$np_2 = 8.809 > 5\checkmark$

Nun läuft alles glatt und wir erhalten:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $p_i = P(\mathrm{I}_i)$ | 0.1325 | 0.1602 | 0.1043 | 0.1123 | 0.1116 | 0.1023 | 0.1541 | 0.1228 |
| $55p_i$ | 7.288 | 8.809 | 5.737 | 6.177 | 6.136 | 5.625 | 8.475 | 6.754 |
| $n_i = \#\{j : x_j \in \mathrm{I}_i\}$ | 7 | 8 | 7 | 4 | 8 | 5 | 11 | 5 |
| $\dfrac{(n_i - 55p_i)^2}{55p_i}$ | 0.0114 | 0.0742 | 0.2782 | 0.7674 | 0.5663 | 0.0694 | 0.7526 | 0.4554 |

$$z = \sum_{i=1}^{8} \frac{(n_i - 55p_i)^2}{55p_i} = 2.9748$$

$\nu =$ Anzahl der Freiheitsgrade$= r - 3 = 8 - 3 = 5$,
$\alpha = 5\,\% \implies u = \big[95\,\%\text{-Fraktil von } \chi^2\text{-Vert}(5)\big]$
$= 11.1$
$z = 2.9748 < 11.1 \implies$ wir nehmen die Hypothese „$H_0$ : Dichtemessungen normalverteilt“ an.

# Literaturliste zur Vorlesung Wahrscheinlichkeitstheorie und Statistik WS 2002/03


| | |
|---|---|
| Anderson, Popp, etc.: | Schätzen und Testen, HTB 177, Springer, Berlin 1976. |
| Bandemer, Bellmann: | Statistische Versuchsplanung, Teubner, Leipzig 1988. |
| Blaut: | Statistische Verfahren für die Gütesicherung von Beton, Bauverlag, Berlin 1968. |
| Blume: | Statistische Methoden für Ingenieure und Naturwissenschaftler I,II, VDI, Düsseldorf 1970, 1974. |
| Bosch: | Aufgaben und Lösungen zur angewandten Statistik, Vieweg, Braunschweig 1983. |
| Brandt: | Statistische Methoden der Datenanalyse, BI 816, Mannheim 1968. |
| Clauss, Ebner: | Grundlagen der Statistik für Psychologen etc., H.Deutsch, Frankfurt 1972. |
| Dorfwirth: | Einführung in die Verkehrsstatistik, dbv-Verlag, Graz 1982. |
| Gräff: | Statistik in der Klimatechnik, C.F. Müller-Verlag, Karlsruhe 1974. |
| Hartung, Elpelt: | Multivariate Statistik, Oldenbourg, München 1984. |
| Hein, O.: | Statistische Verfahren der Ingenieurpraxis, BI 119, Mannheim 1972. |
| Heinhold, Gaede: | Zufall und Gesetz, Oldenbourg, München 1974. |
| Kreyszig: | Statistische Methoden und ihre Anwendungen, 3. Aufl., Vandenhoek, Göttingen 1968. |
| Lindner, Berchtold: | Elementare statistische Methoden, UTB 796, Birkhäuser, Basel 1978. |
| Müller, Neumann, Storm: | Tafeln der mathematischen Statistik, VEB, Leipzig 1979. |
| Pfanzagl: | Elementare Wahrscheinlichkeitsrechnung, de Gruyter, Berlin 1989. |
| Pfanzagl: | Methodenlehre der Statistik I,II, Göschen 746/7, Berlin 1968. |

Plate: Statistik und angewandte Wahrscheinlichkeitslehre für Bauingenieure, Ernst, Berlin 1993.

Renyi: Wahrscheinlichkeitsrechnung, 6. Aufl., Dt. Verl. Wiss., Berlin 1979.

Sachs: Statistische Methoden, Springer, Berlin 1972.

Schmetterer: Einführung in die mathematische Statistik, Springer, Wien 1966.

Schneider: Sicherheit und Zuverlässigkeit im Bauwesen, Teubner, Stuttgart 1994.

Schuëller: Einführung in die Sicherheit und Zuverlässigkeit von Tragwerken, Ernst, Berlin 1981.

Spiegel: Statistics, Schaum, New York 1961.

Storm: Wahrscheinlichkeitsrechnung, Math. Statistik, Statistische Qualitätskontrolle, Fachbuchverlag, Leipzig, 1995.

Waerden, B.L.: Mathematische Statistik, Grundlehren 87, 2. Aufl., Springer, Berlin 1965.

Weber: Einführung in die Wahrscheinlichkeitsrechnung & Statistik für Ingenieure, Teubner, Stuttgart 1983.

Wilks, D.S.: Statistical Methods in the Atmospheric Sciences, Academic Press, Elsevier, 2006.