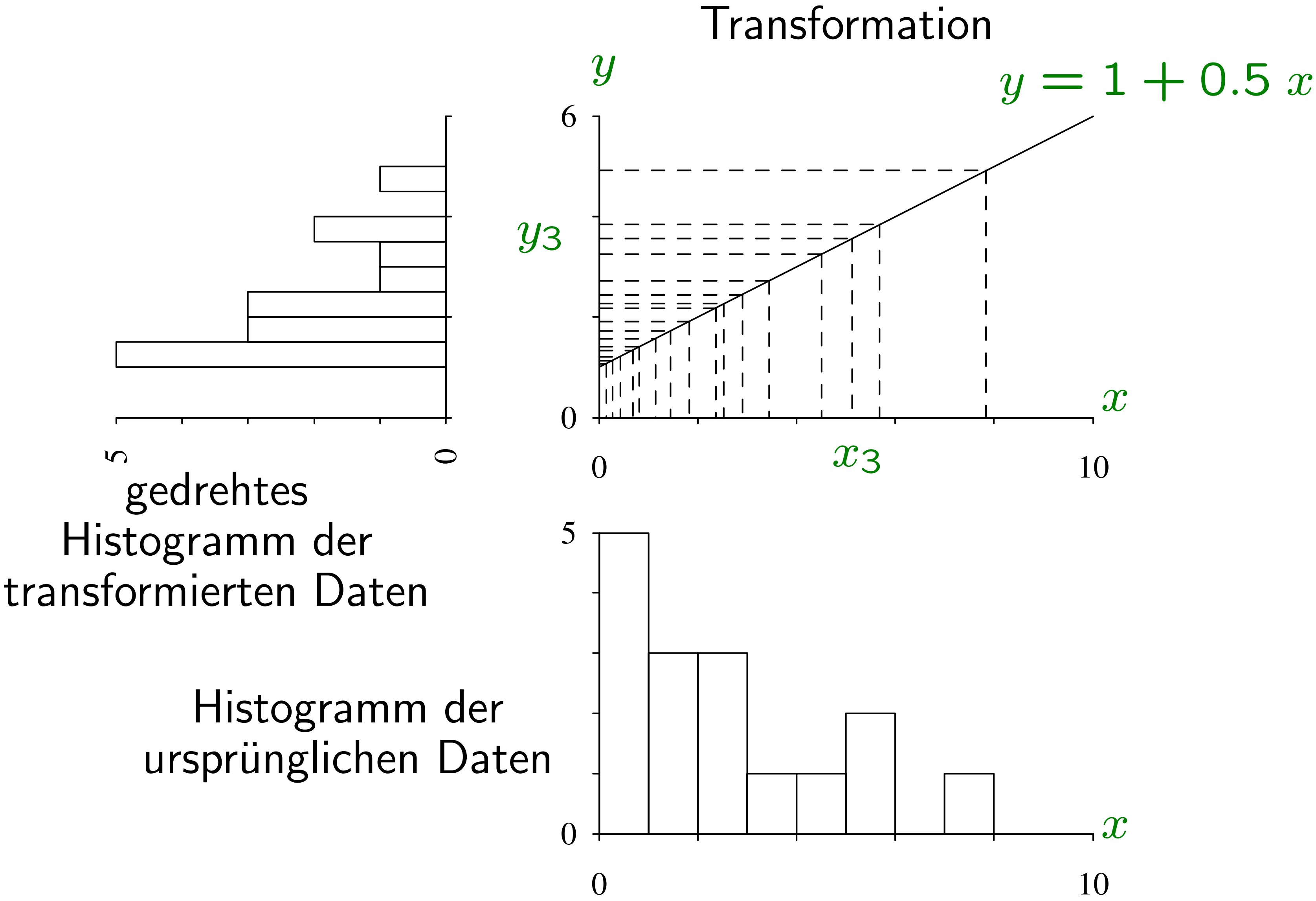
2.6 Transformationen von Beobachtungen

b Umrechnung von ^{0}F in ^{0}C:

$x \mapsto y = a + b \cdot x$ oder $x \mapsto y = b \cdot (x + c)$

lineare Transformation.


Transformation
$y = 1 + 0.5\,x$
y
y_3
x
x_3
6
0
10
gedrehtes
Histogramm der
transformierten Daten
5
Histogramm der
ursprünglichen Daten

d
$$\begin{aligned}\text{Lage}_Y &= a + b \cdot \text{Lage}_X\\ \text{Streuung}_Y &= |b| \cdot \text{Streuung}_X\ .\end{aligned}$$

Form unverändert.

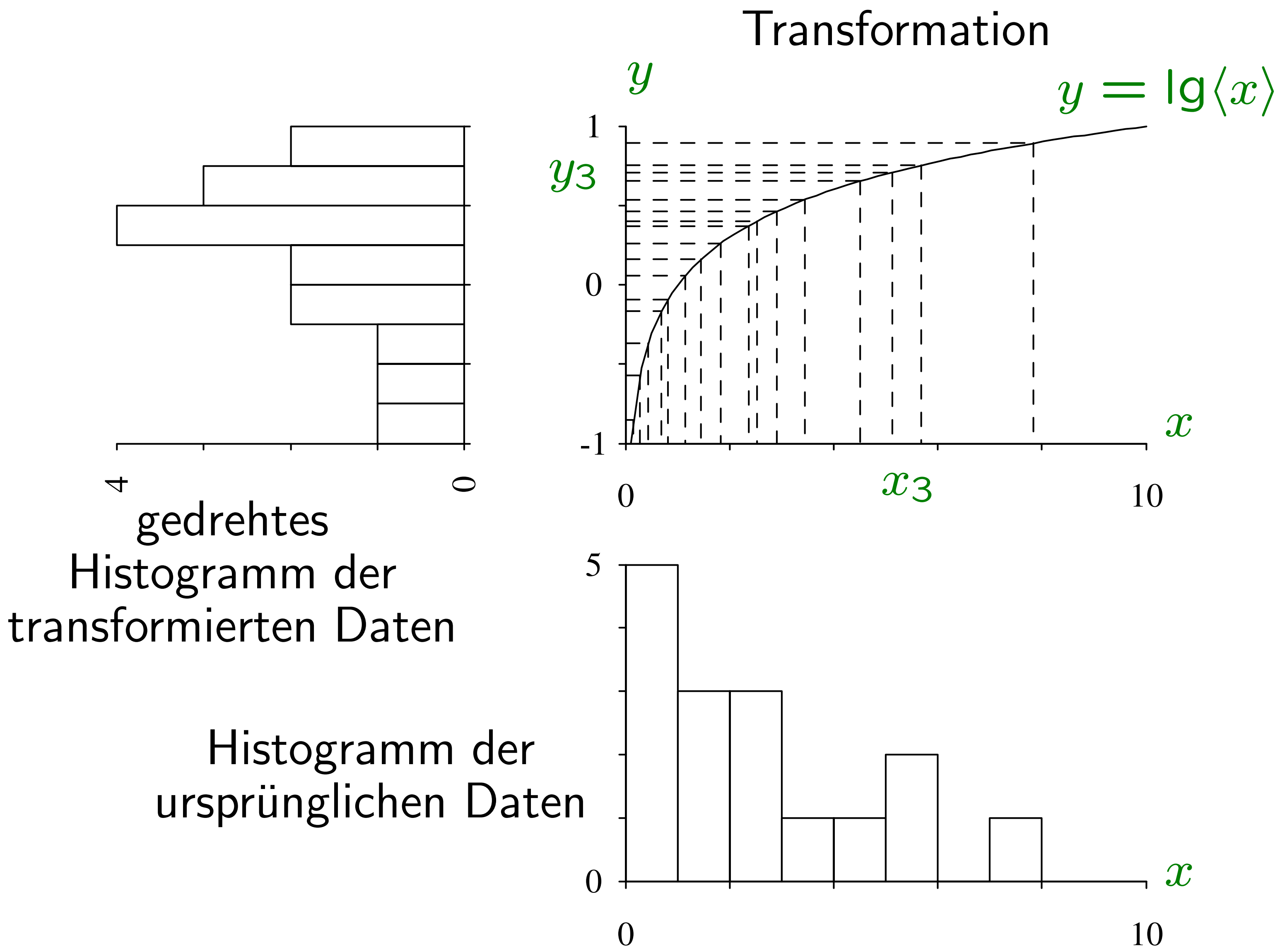
e Standardisierte Stichprobe $z_i = (x_i - \bar{x}) / \text{sd}_X$

$\longrightarrow\ \bar{z} = 0,\ \text{var}_Z = 1\,.$

f Form:

$$\begin{aligned} \text{Schiefe}_X &= \text{Schiefe}_Z = \frac{1}{n}\sum_{i=1}^{n} z_i^3 \\ \text{Langschwänzigkeit} = \text{Exzess}_X &= \text{Exzess}_Z = \frac{1}{n}\sum_{i=1}^{n} z_i^4 \; - \; 3 \end{aligned}$$

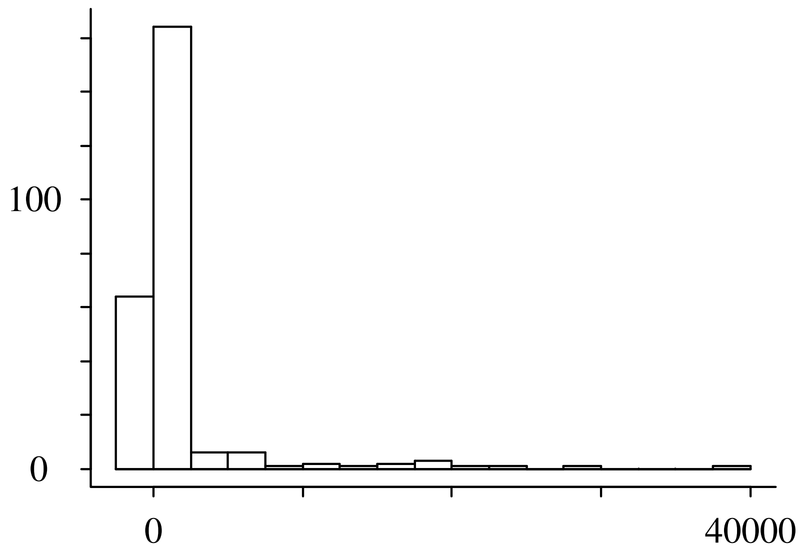
g Nicht-lineare, monotone Transformationen $x \mapsto g\langle x\rangle$ verändern Lage, Streuung und Schiefe.

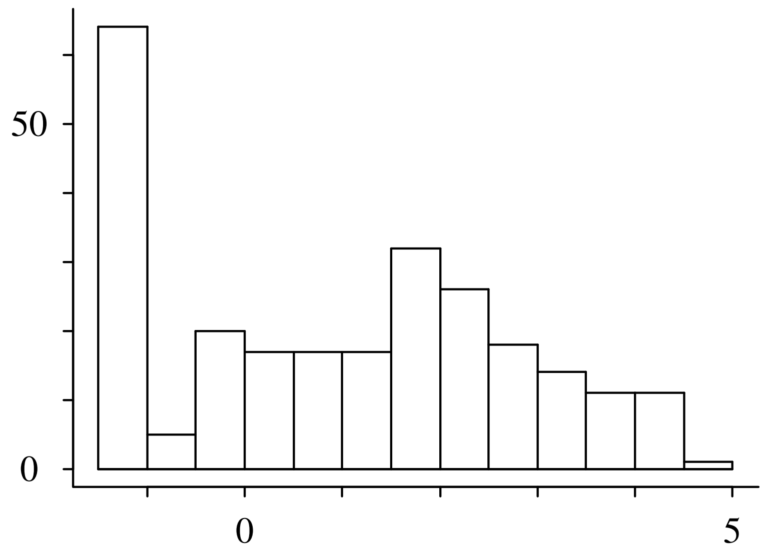


Median und Quantile sind „äquivariant”, med $\mapsto g\langle$med$\rangle$

h **Logarithmus-** und **Wurzel-Transformation**:
Positive Schiefe nimmt ab. $\longrightarrow$ Nützlich!

Beispiel Hagel-Energien („Grossversuch IV" zur Hagelbekämpfung im Napfgebiet, 1980-85)





2.6

i Hagelenergie $= 0$, $\lg\langle 0\rangle = -\infty$! Deshalb $x \mapsto \lg\langle x + c\rangle$.
$c =?$ Nicht einfach $c = 1$! $c = q_{1/4}^{*\,2}/q_{3/4}^{*}$.

j Basis des Logarithmus: Mess-Einheit.

2.7 Wertebereiche, Datensorten

b Der Bereich der möglichen Werte:

1. „alle" reellen Zahlen;
2. die positiven Zahlen, mit oder ohne Null: Beträge;
3. Intervall $[0, 1]$ oder $[0\%, 100\%]$: Anteile, Prozentzahlen;
4. nicht-negative ganze Zahlen: Anzahlen, Zähldaten;
5. $\{0, 1/n, 2/n, \ldots, (n-1)/n, 1\}$ für geg. n : Anteile aus Anzahlen;
6. $\{0, 1\}$ (0 oder 1) : binäre Daten;
7. $\{0, 1, 2, \ldots, n\}$ als Codes für Herkunft, Typ, Farbe,... : nominale Daten;
8. Intervall $[0^0, 360^0)$ oder $[0^0, 180^0)$: Winkel, Richtungen, zirkuläre Daten.

c Zensierte Daten.
Überlebenszeiten von Patienten nach einer Herzoperation,
Erholungszeiten nach einer Erkrankung,
Brenndauern von Glühbirnen,
Verweildauern von Zugvögeln:

Beob. „zensuriert": überlebenszeit $\geq$ Beobachtungsdauer c_i.

d Ordnungsstruktur.
Nominale Daten: Codes 1, 2, 3, ...,
Ordnung $1 < 2 < 3 \ldots$ nicht sinnvoll interpretierbar.
Unterschiede $\neq$ Differenzen zwischen Beobachtungen.

2.7

e Zustand v. Bäumen: Klassen v. 0 (gesund) bis 4 (abgestorben).
„Bonitierung". Ordnung von Bedeutung.
Aber wieder: Unterschiede $\neq$ Differenzen. Ordinalskala.

Temperaturen: Unterschiede $=$ Differenzen. Differenzenskala
Konzentrationen: Unterschiede $=$ Verhältnisse. Quotientenskala

2.8 * Transformationen und Unterschiede zwischen Beobachtungen

a Daten so transformieren, dass Unterschiede = Differenzen !
Oder: Unterschiede = Rang-Differenzen.
Rang-Transformation $x_i \mapsto \text{Rang}\langle x_i \rangle$
hängt von der ganzen Stichprobe ab!

Stetige Verteilungen

6.1 Einleitung

- Repetition diskrete Verteilungen:

  1. Verteilung gegeben durch die $P\langle X = x\rangle$
     für alle mögl. Werte x (=0,1,2,... f. Zähldaten)

  2. oder: durch (theoretische) kumulative Verteilungsfn.

     $$F\langle x\rangle = P\langle X \leq x\rangle \qquad .$$

  3. Zufallszahlen mit einer bestimmten Verteilung:
     z_ℓ uniform verteilte Zz. in $(0, 1)$
     $x_\ell = F^{-1}\langle z_\ell\rangle$: Zz. mit Verteilung F.

Punkte 2 und 3 genau gleich für stetige Verteilungen.

- Wir brauchen noch etwas mehr
  Theorie = Wahrscheinlichkeitsrechnung:

  - Transformationen von Zufallsvariablen,
  - mehrere Zufallsvariable, gemeinsame Verteilung
  - Modell für eine Zufalls-Stichprobe
    = mehrere Messungen „der gleichen Grösse"

Bitte repetieren Sie die Grundbegriffe
Ableitung und Integral (Stammfunktion) *Buch 6.1.d-e*

6.2 Grundbegriffe, Exponential- und uniforme Vert.

d Beispiel Regentropfen. X-Koordinate.

$$F\langle x\rangle = \begin{cases} 0 & \text{für } x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{für } \alpha \le x \le \beta \\ 1 & \text{für } x > \beta \,. \end{cases}$$

$X \sim \mathcal{U}\langle\alpha, \beta\rangle$: uniforme Verteilung, Rechtecks-Vert.

Anwendung: Rundungsfehler: Mess-Ergebnis 1.2,
eigentlicher Messwert $X \sim \mathcal{U}\langle 1.15, 1.25\rangle$.
(„Wahrer" Wert: weiter weg!)

e Exponential-Verteilung:

$$F\langle x\rangle = \begin{cases} 1 - e^{-\lambda x} & \text{für} \quad x \geq 0\,, \\ 0 & \text{für} \quad x < 0 \end{cases}$$

Anwendungen:

- Zeit bis zum Eintreffen eines bestimmten „Ereignisses"
  = „Wartezeit".

- Speziell: Zeit bis zu einem Defekt bei einem Gerät
  bis zum Tod bei Lebewesen = Lebensdauer, Überlebenszeit

f* Zusammenhang mit Poisson-Verteilung:

$\mathcal{P}$ zählt Ereignisse in einem bestimmten Zeitabschnitt.

$X =$ Wartezeit bis zum nächsten Ereignis,

$Y =$ Anzahl Ereignisse in der nächsten Stunde

$$P\langle Y = 0\rangle = \frac{\lambda^0}{0!}\exp\langle -\lambda\rangle = P\langle X > 1\rangle = F_X\langle 1\rangle$$

Poisson-Prozess

g **Beispiel Atomzerfall**. Exponential-Verteilung.

Für welchen Zeitpunkt wird die Wahrscheinlichkeit, dass das Isotop bis dahin zerfällt, gleich $1/2$?

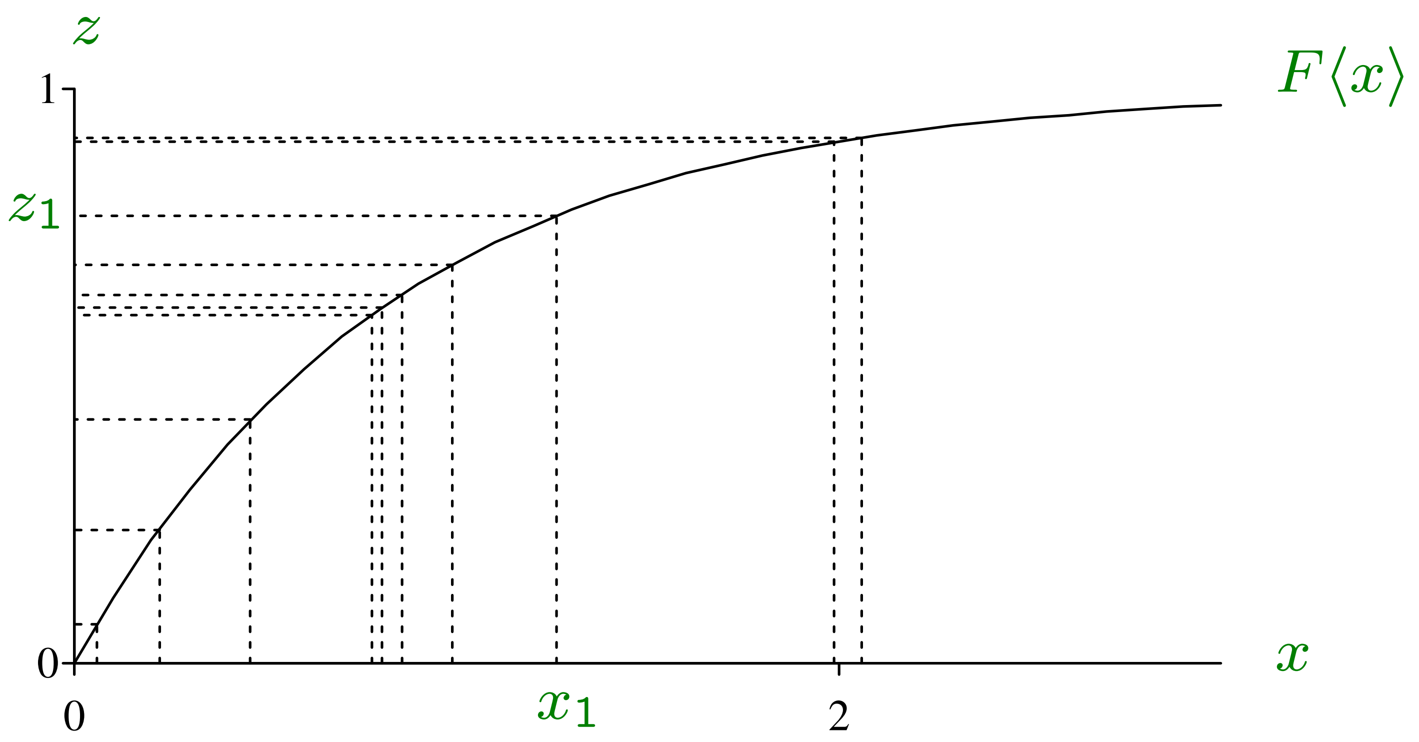
Antwort: Median

$$\begin{aligned} F\langle x\rangle &= 1 - e^{-\lambda x} = \frac{1}{2} \quad \Rightarrow \quad -\lambda x = \log_e \left\langle \frac{1}{2} \right\rangle \\ \Rightarrow \quad x &= \log_e\langle 2\rangle/\lambda = 0.693/\lambda \end{aligned}$$

Der Wahrscheinlichkeit $1/2$ des Zerfalls eines einzelnen Isotops entspricht die relative Häufigkeit $1/2$ der zerfallenen Isotope bis zum Zeitpunkt $0.693/\lambda$.

Strahlung halbiert, **Halbwertszeit**.

i Zufallszahlen, exponential-verteilt $x = F^{-1}\langle z\rangle = -\frac{1}{\lambda}\,\mathsf{log}_e\langle 1 - z\rangle$



j Darstellung von Verteilungen.

Für diskrete Zufallsv. $P\langle X = x\rangle$, Stabdiagramm (Histogramm).

Der rel. Häufigkeit für Klasse $(c, c + \Delta c)$ entspricht

$$P\langle c \leq X \leq c + \triangle c\rangle = F\langle c + \triangle c\rangle - F\langle c\rangle$$

Klasseneinteilung verfeinern! – Flächen proportional zu Anteilen:

$P\langle c < X \leq c + \triangle c\rangle / \triangle c = (F\langle c + \triangle c\rangle - F\langle c\rangle) / \triangle c$

Grenzübergang $\triangle c \rightarrow 0$:   Höhe $=$ Ableitung von F.

k Definition. Dichte f einer stetigen Zufallsvariablen X $=$ Ableitung ihrer Verteilungsfunktion F,

$$f\langle x\rangle = F'\langle x\rangle \ .$$

Umkehrung:

$$F\langle x\rangle = \int_{-\infty}^{x} f\langle t\rangle dt$$

Stammfunktion von f mit $F\langle -\infty\rangle = 0$.

Wahrscheinlichkeiten für Intervalle:

$$P\langle a < X \le b\rangle = F\langle b\rangle - F\langle a\rangle = \int_a^b f\langle t\rangle dt \ .$$

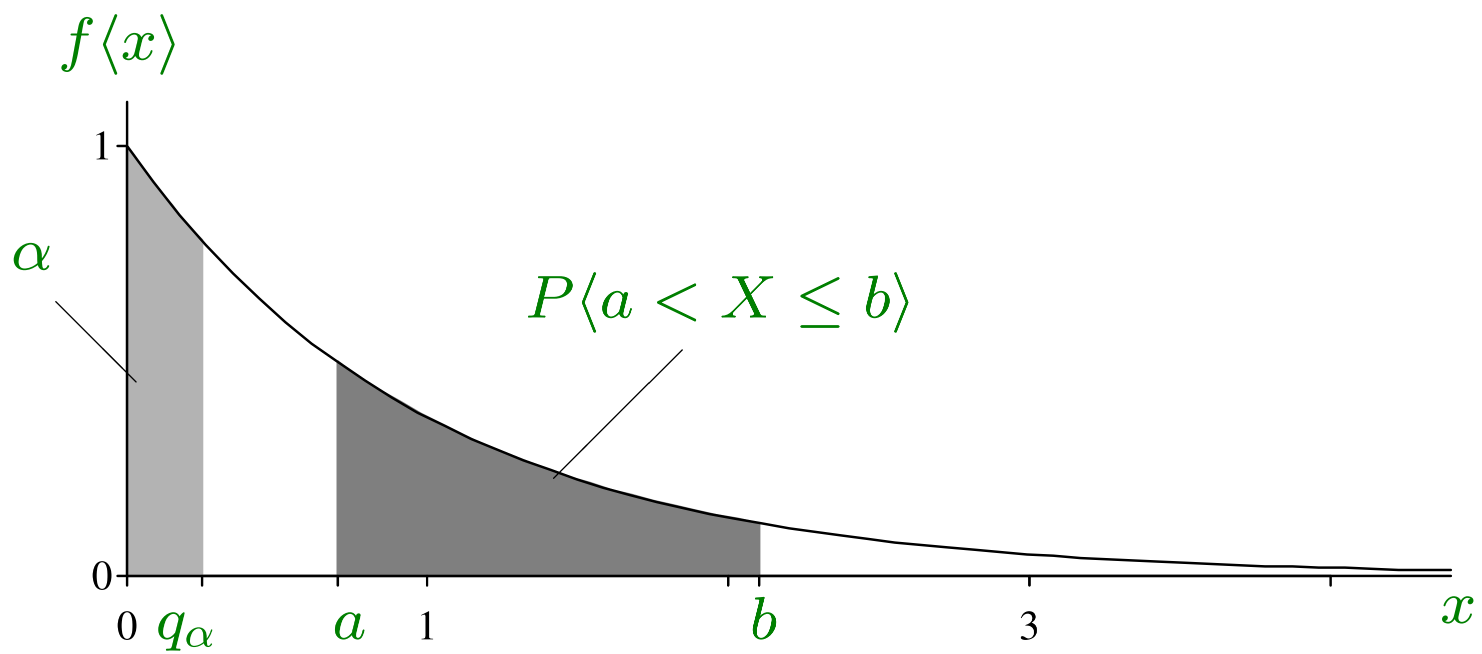
$f\langle x\rangle \ge 0$ für alle x, $\quad \int_{-\infty}^{\infty} f\langle x\rangle dx = 1$.

l Exponential-Verteilung

$$F\langle x\rangle = 1 - e^{-\lambda x}$$
$$f\langle x\rangle = \lambda e^{-\lambda x}$$

für $x > 0$.

Wahrsch. für ein Intervall: Fläche unter der Dichtekurve.



m uniforme Verteilung auf (α, β). $f\langle x\rangle =?$

6.3 Kennzahlen für stetige Verteilungen

a Quantile: $q_\alpha = F^{-1}\langle\alpha\rangle$.

b Für klassierte Daten war das arithmetische Mittel $\bar{x} \approx \sum_k z_k f_k$.
z_k: Mitte der $k$ten „Klasse" $\mathcal{C}_k$ f_k: relative Häufigkeit.

$$\begin{aligned}\mathcal{E}\langle X\rangle &\approx \sum_k z_k P\langle c_{k-1} \le X \le c_k\rangle \\ &= \sum_k \left(c_{k-1} + \frac{\triangle c}{2}\right) \frac{F\langle c_{k-1} + \triangle c\rangle - F\langle c_{k-1}\rangle}{\triangle c} \cdot \triangle c,\end{aligned}$$

$\triangle c$: Klassenbreite. – Grenzübergang:

$$\frac{F\langle c_{k-1} + \triangle c\rangle - F\langle c_{k-1}\rangle}{\triangle c} \longrightarrow f\langle c_{k-1}\rangle$$

$$\begin{aligned}\mathcal{E}\langle X\rangle &\approx \sum_k c_{k-1} f\langle c_{k-1}\rangle \triangle c \\ \mathcal{E}\langle X\rangle &= \int c\, f\langle c\rangle\, dc = \int_{-\infty}^{\infty} x\, f\langle x\rangle\, dx\end{aligned}$$

d Varianz:

$$\begin{aligned} \mathsf{var} &\approx \sum_k (z_k - \widetilde{x})^2 P\langle c_{k-1} \le X < c_k \rangle \\ \mathsf{var}\langle X \rangle &= \int_{-\infty}^{\infty} \big(x - \mathcal{E}\langle X \rangle\big)^2 f\langle x \rangle \, dx \\ &= \mathcal{E}\left\langle (X - \mu)^2 \right\rangle \end{aligned}$$

$\sigma^2 = \mathsf{var}\langle X \rangle$.

σ: Standardabweichung.

e **Exponential-Verteilung**: $f\langle x\rangle = \lambda e^{-\lambda x}$ für $x \geq 0$.

$$\begin{aligned}
\mathcal{E}\langle X\rangle &= \int_0^\infty x\,\lambda e^{-\lambda x}\,dx \\
&= \lambda \int_0^\infty x e^{-\lambda x}\,dx \\
&= \lambda \cdot \left(\left[x \left(\frac{-1}{\lambda e^{\lambda x}} \right) \right]_0^\infty - \int_0^\infty 1 \cdot \left(-\tfrac{1}{\lambda} e^{-\lambda x}\right) dx \right) \\
&= \int_0^\infty e^{-\lambda x} dx = -\frac{1}{\lambda} \left[e^{-\lambda x} \right]_0^\infty = \frac{1}{\lambda}. \\
\mathrm{var}\langle X\rangle &= \int_0^\infty \left(x - \frac{1}{\lambda} \right)^2 \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2}
\end{aligned}$$

6.4 Transformationen von Zufallsvariablen

a Zufallsvariable X, Verteilung gegeben.
Zufallsvariable Y: $Y\langle\omega\rangle = g\big\langle X\langle\omega\rangle\big\rangle$.
Verteilung von Y?

$$\begin{aligned} g\langle x\rangle &= a + b\cdot x & &\text{lineare T.}\\ g\langle x\rangle &= \lg\langle x\rangle & &\text{Logarithmus-T.}\end{aligned}$$

$X = 5.1 \quad \Rightarrow \quad Y = \lg\langle 5.1\rangle = 0.71$.

6.4

c
$$\begin{aligned} F^{(Y)}\langle 0.71\rangle &= P\langle Y \le 0.71\rangle = P\langle \lg\langle X\rangle \le \lg\langle 5.1\rangle\rangle \\ &= P\langle X \le 5.1\rangle = F^{(X)}\langle 5.1\rangle. \end{aligned}$$

Allgemein:

$$\begin{aligned} F^{(Y)}\langle y\rangle &= P\langle Y \le y\rangle = P\langle g\langle X\rangle \le g\langle x\rangle\rangle \\ &= P\langle X \le x\rangle = F^{(X)}\langle x\rangle. \end{aligned}$$

Schwierigkeit: Zufallsvariablen X, Y
und die Zahlen x, y auseinanderhalten.

d
$$F^{(Y)}\langle y\rangle = F^{(X)}\left\langle g^{-1}\langle y\rangle\right\rangle$$

$\lg^{-1}\langle y\rangle = \exp\langle y\rangle$, $\quad \lg^{-1}\langle 0.71\rangle = 5.1$, $\quad F^{(Y)}\langle 0.71\rangle = F^{(X)}\langle 5.1\rangle$.

e

$$\begin{aligned} y &= g\langle x\rangle \\ F^{(Y)}\langle y\rangle &= P\langle Y \le y\rangle = P\langle g\langle X\rangle \le g\langle x\rangle\rangle \\ &= P\langle X \le x\rangle \;=\; F^{(X)}\langle x\rangle. \end{aligned}$$

Schwierigkeit: Zufallsvariablen X, Y
und die Zahlen x, y auseinanderhalten.

f Wie wird die Dichte transformiert? – Kettenregel: ...

g Beispiel Quadrat. $X \sim \mathcal{U}\langle 0, 2\rangle$
Z Fläche eines Quadrates mit Seitenlänge X. $Y = X^2$

h Skalen-Änderung: $Y = bX$
$f^{(Y)}\langle y\rangle = \frac{1}{b} f^{(X)}\langle x\rangle$

k Kennzahlen: Quantile, Median von Y:

$q_\alpha^{(Y)} = g\left\langle q_\alpha^{(X)}\right\rangle$, falls g monoton zunehmend ist.

$$\alpha = P\langle X \le q_\alpha^{(X)}\rangle = P\left\langle g\langle X\rangle \le g\langle q_\alpha^{(X)}\rangle\right\rangle = P\left\langle Y \le g\langle q_\alpha^{(X)}\rangle\right\rangle \;.$$

l Erwartungswert, Varianz: keine wesentliche Vereinfachung.

m Lineare Transformation: $y = a + bx$

$$\mathcal{E}\langle Y\rangle \;=\; a + b\cdot\mathcal{E}\langle X\rangle, \qquad \text{var}\langle Y\rangle = b^2\cdot\text{var}\langle X\rangle.$$

- Wie wird die Dichte transformiert? – Kettenregel:

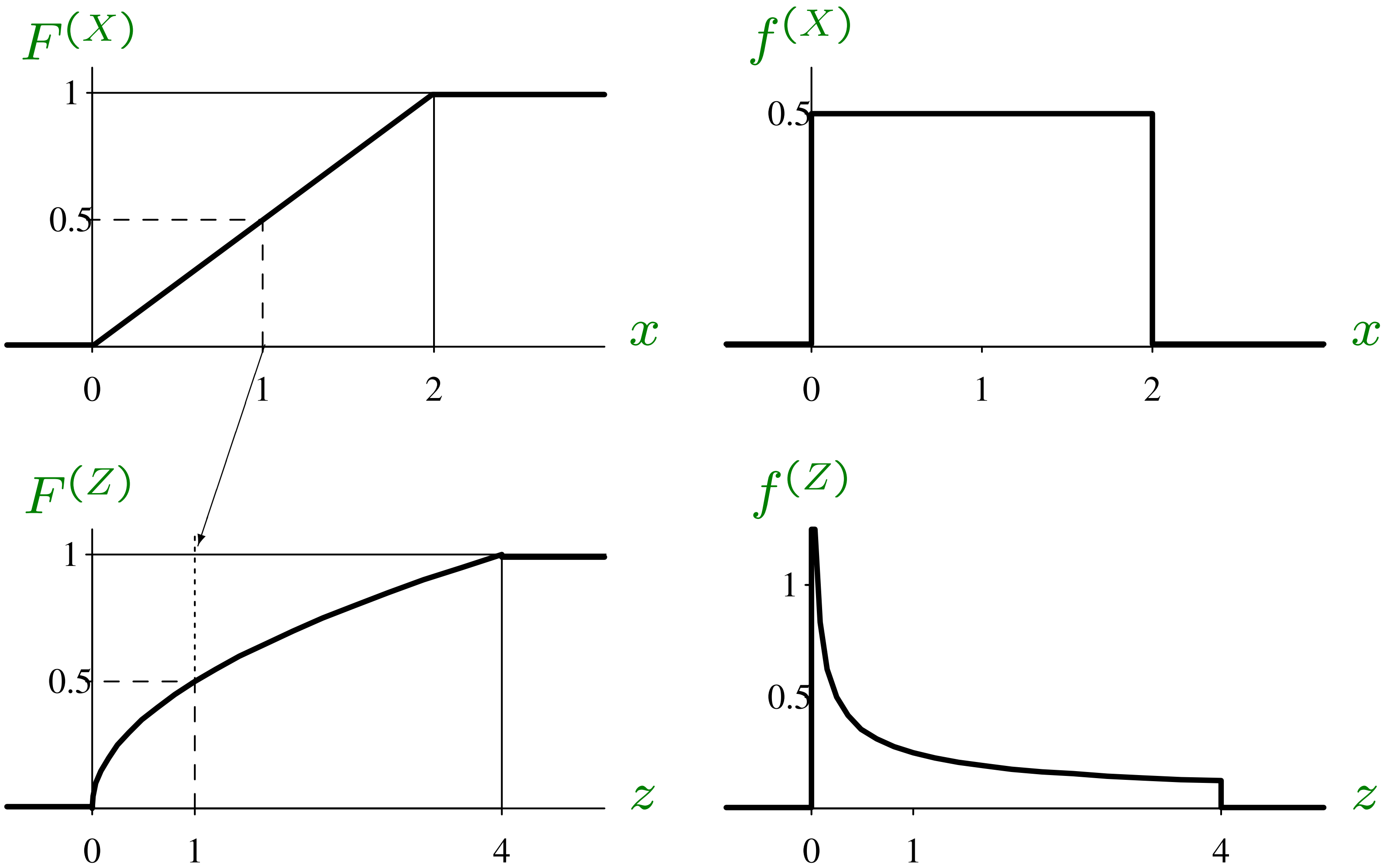
$$f^{(Y)}\left\langle g\langle x\rangle\right\rangle \cdot g'\langle x\rangle = f^{(X)}\langle x\rangle$$

$$f^{(Y)}\langle y\rangle = f^{(X)}\left\langle g^{-1}\langle y\rangle\right\rangle \Big/ g'\left\langle g^{-1}\langle y\rangle\right\rangle$$

- **Beispiel Quadrat.** $X \sim \mathcal{U}\langle 0,2\rangle$
  Z Fläche eines Quadrates mit Seitenlänge X.

$$\begin{aligned} F^{(X)}\langle x\rangle &= \frac{x}{2}, \qquad f^{(X)}\langle x\rangle = \frac{1}{2} \\ F^{(Z)}\langle z\rangle &= F^{(X)}\langle\sqrt{z}\rangle = \frac{\sqrt{z}}{2} \\ f^{(Z)}\langle x^2\rangle \cdot 2x &= f^{(X)}\langle x\rangle = \frac{1}{2} \quad \Longrightarrow \quad f^{(Z)}\langle z\rangle = \frac{1}{4\sqrt{z}}. \end{aligned}$$

$f^{(Z)}\langle z\rangle = (F^{(Z)})'\langle z\rangle$!


$F^{(X)}$
1
0.5
x
0
1
2
$f^{(X)}$
0.5
x
0
1
2
$F^{(Z)}$
1
0.5
z
0
1
4
$f^{(Z)}$
1
0.5
z
0
1
4

p **Skalen-Änderung**.

$X \sim \mathcal{E}xp\langle\lambda\rangle$, $g\langle x\rangle = b \cdot x$ mit $b > 0$.

$g^{-1}\langle y\rangle = y/b$, $g'\langle x\rangle = b$.

$$\begin{aligned} F^{(Y)}\langle y\rangle &= F^{(X)}\langle y/b\rangle = 1 - e^{-\lambda y/b} \\ f^{(Y)}\langle b \cdot x\rangle \cdot b &= f^{(X)}\langle x\rangle = \lambda\, e^{-\lambda x} \\ f^{(Y)}\langle y\rangle &= \frac{\lambda}{b}\, e^{-\lambda y/b}\,. \end{aligned}$$

$\mu = \frac{\lambda}{b} \quad \Rightarrow \quad f^{(Y)}\langle y\rangle = \mu e^{-\mu y}$.

$Y \sim \mathcal{E}xp\langle\lambda/b\rangle$

s Kennzahlen: Quantile, Median von Y:
$q_\alpha^{(Y)} = g\left\langle q_\alpha^{(X)} \right\rangle$, falls g monoton zunehmend ist.

$$\alpha = P\langle X \le q_\alpha^{(X)}\rangle = P\left\langle g\langle X\rangle \le g\langle q_\alpha^{(X)}\rangle \right\rangle = P\left\langle Y \le g\langle q_\alpha^{(X)}\rangle \right\rangle .$$

t Erwartungswert: keine wesentliche Vereinfachung. Beispiel Quadrat
$\mathcal{E}\langle Z\rangle = \int_0^4 z\, f^{(Z)}\langle z\rangle\, dz = \int_0^4 z \frac{1}{4\sqrt{z}}\, dz = \frac{1}{4}\int_0^4 \sqrt{z}\, dz = \frac{1}{4}\left[\frac{2}{3} z^{3/2}\right]_0^4 = \frac{4}{3}$
oder $\mathcal{E}\langle Z\rangle = \int g\langle x\rangle\, f^{(X)}\langle x\rangle\, dx = \int_0^2 x^2 \frac{1}{2}\, dx = \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 = \frac{4}{3}$
(Dagegen ist $g\langle\, \mathcal{E}\langle X\rangle\rangle = 1^2 = 1$.)

- **Lineare Transformation:** $y = a + bx$

$$\mathcal{E}\langle Y\rangle = a + b\cdot\mathcal{E}\langle X\rangle, \qquad \mathrm{var}\langle Y\rangle = b^2\cdot\mathrm{var}\langle X\rangle.$$

- Herleitung:

$$\begin{aligned}
\mathcal{E}\langle Y\rangle &= \int y\, f^{(Y)}\langle y\rangle dy \\
y &= a + bx, \quad dy = b\,dx \\
\int (a+bx) f^{(Y)}\langle a+bx\rangle\, b\,dx &= \int (a+bx)\frac{f^{(X)}\langle x\rangle}{b}\, b\,dx \\
\mathcal{E}\langle Y\rangle &= a\int f^{(X)}\langle x\rangle\; dx + b \\
\int x f^{(X)}\langle x\rangle\; dx &= a + b\cdot\mathcal{E}\langle X\rangle
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}\langle Y\rangle &= \mathcal{E}\left\langle (Y-\mathcal{E}\langle Y\rangle)^2\right\rangle = \mathcal{E}\left\langle (a+bX-(a+b\,\mathcal{E}\langle X\rangle))^2\right\rangle \\
&= \mathcal{E}\left\langle b^2\,(X-\mathcal{E}\langle X\rangle)^2\right\rangle = b^2\;\mathrm{var}\langle X\rangle\,.
\end{aligned}$$

w **Standardisierte Zufallsvariable.**

Zufallsvariable X mit Erwartungswert μ und Varianz σ^2.

Transformation: Z so, dass $\mathcal{E}\langle Z\rangle = 0$ und $\mathrm{var}\langle Z\rangle = 1$:

$g\langle x\rangle = (x-\mu)/\sigma$.

$$\mathcal{E}\langle Z\rangle = \frac{1}{\sigma}(\mathcal{E}\langle X\rangle - \mu) = 0, \quad \mathrm{var}\langle Z\rangle = \frac{1}{\sigma^2}\mathrm{var}\langle X\rangle = 1 .$$

6.5 Die Normalverteilung

a Normalverteilung, „Gausssche Glockenkurve" spielt eine zentrale Rolle in der Wahrscheinlichkeitstheorie und Statistik.

- Mathematisch einfache Gesetze

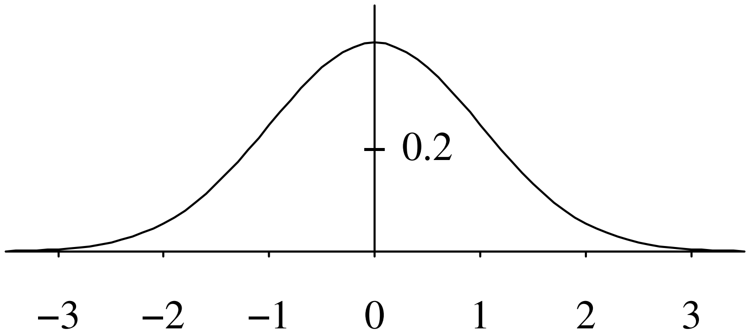
- Passt in vielen Situationen nicht schlecht.
  Grund: Wenn sich viele kleine, unabhängige Zufallseffekte summieren,
  ist die Summe näherungsweise normalverteilt nach dem „Zentralen Grenzwertsatz".
  $\longrightarrow$ „Hypothese der Elementarfehler"

b Standard-Normalverteilung $s\mathcal{N}$: Dichte

$$f\langle z\rangle = \phi\langle z\rangle = c \cdot e^{-\frac{1}{2}z^2}$$

„Normierungskonst." c so, dass $\int_{-\infty}^{\infty} f\langle z\rangle\, dz = 1$. $\quad c = 1/\sqrt{2\pi}$

$f\langle z\rangle$



z

Kumulative Verteilungsfunkt. Φ. Keine „geschlossene" Formel.

c Wegen der Symmetrie ist $\mathcal{E}\langle Z\rangle = 0$.

$$\text{var}\langle Z\rangle = c\int z^2 \exp\left\langle -\frac{1}{2}z^2\right\rangle dz = 1$$

e Lage und Streuung soll sich an die Daten anpassen.
Lineare Transformation von $Z \sim s\mathcal{N}$: $X = \mu + \sigma Z$
$\mathcal{E}\langle X\rangle = ...,\quad \text{var}\langle X\rangle = ...,\quad \text{St.abw.}\langle X\rangle = ...,$ Dichte:

$$f^{(X)}\langle x\rangle = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\langle -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\rangle$$

X normalvert. m. Erwartungswert μ und Var. σ^2. $X \sim \mathcal{N}\langle \mu, \sigma^2\rangle$.

f Kumulative Verteilungsfunktion

$$F\langle x\rangle = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\exp\left\langle -\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right\rangle dt$$

$1 - \Phi\langle z \rangle$

$z = \downarrow + \rightarrow$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233

Tabelle! Standardisieren!

$$F^{(X)}\langle x\rangle = F^{(Z)}\left\langle \frac{x-\mu}{\sigma} \right\rangle = \Phi\left\langle \frac{x-\mu}{\sigma} \right\rangle$$

Beispiel: $X \sim \mathcal{N}\langle 10, 2^2\rangle$. Wie gross ist $F^{(X)}\langle 14\rangle$?

$z = (x-\mu)/\sigma = 2$, $\qquad F^{(s\mathcal{N})}\langle 2\rangle = \Phi\langle 2\rangle = ?$

Tabelle: $1 - \Phi\langle 2\rangle = 0.023$, also $F^{(X)}\langle 14\rangle = \Phi\langle 2\rangle = 0.977$.

Für negative z benützt man die Symmetrie:

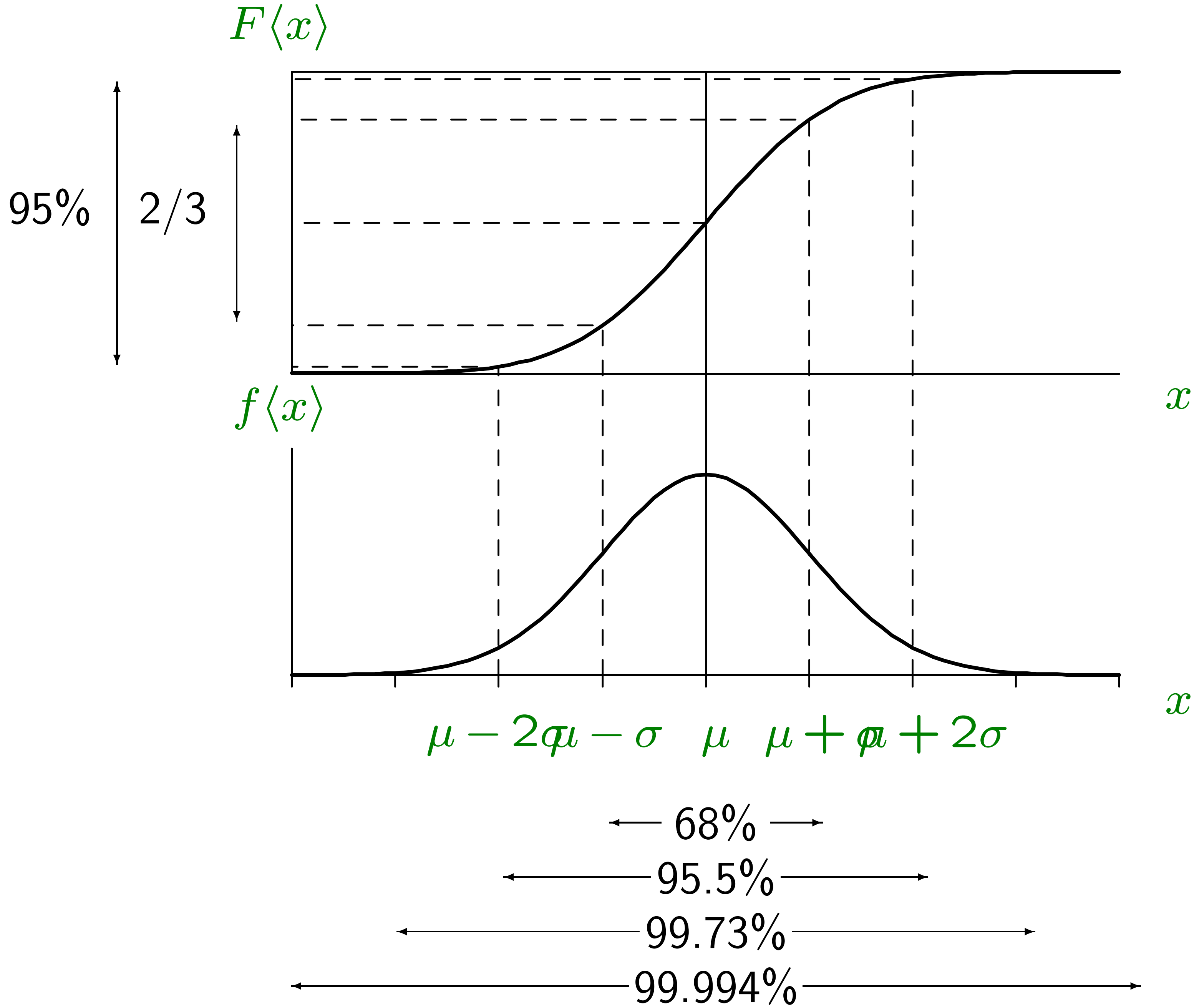
$$\Phi\langle z\rangle = P\langle Z \le z\rangle = P\langle Z \ge -z\rangle = 1 - P\langle Z \le -z\rangle = 1 - \Phi\langle -z\rangle$$

$x = 8$, $\mu = 10$, $\sigma = 2$:

$z = (8 - 10)/2 = -1$, $\quad \Phi\langle z\rangle = 1 - \Phi\langle -z\rangle = 1 - \Phi\langle 1\rangle = 0.1587$, $\quad F^{(X)}\langle 8\rangle = 0.1587$.

g Wahrscheinlichkeiten für Intervalle

Intervall	Wahrscheinlichkeit
$(\mu - \sigma, \mu + \sigma)$	ca. 2/3;
$(\mu - 2\sigma, \mu + 2\sigma)$	ca. 95%;
$(\mu - 3\sigma, \mu + 3\sigma)$	ca. 100% (99.7%).

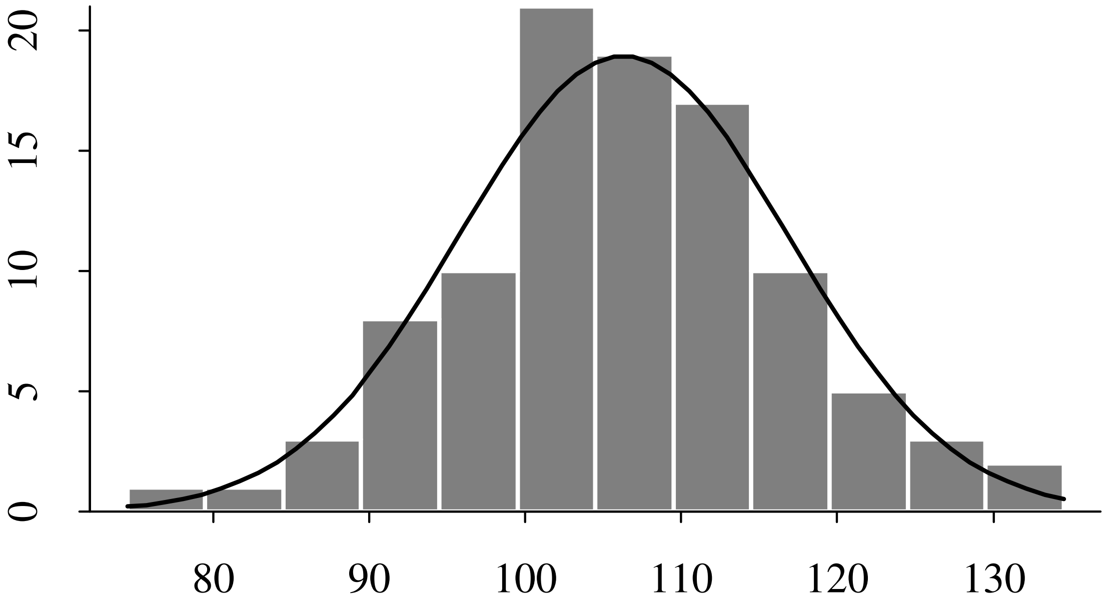

$F\langle x \rangle$
95%
2/3
$f\langle x \rangle$
x
x
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
68%
95.5%
99.73%
99.994%

h **Beispiel Waage**. Messfehler X normalverteilt mit $\mu = 0$ und $\sigma = 0.63\,\text{mg}$.
Wie gross ist die Wahrscheinlichkeit, dass eine Messung um $\leq 0.63\,\text{mg}$, $1.26\,\text{mg}$ oder $1\,\text{mg}$ vom wahren Wert abweicht?

Für $1\,\text{mg}$ rechnen wir

$$\begin{aligned} P\langle |X| \leq 1\rangle &= P\langle -1 \leq X \leq 1\rangle = 2P\langle 0 \leq X \leq 1\rangle \\ &= 2P\left\langle 0 \leq Z < \frac{1-0}{0.63}\right\rangle = 2\left(\Phi\langle 1.587\rangle - \frac{1}{2}\right) = 0.888\,. \end{aligned}$$

¡ Beispiel Küken: Normalvert. passt recht gut auf das Histogramm.
Welches μ, σ? $\longrightarrow$ Schätzung der Parameter.



j Beträge: $X > 0$, schiefe Verteilungen. Also nicht normalverteilt.

Aber:

$$Y = \log\langle X\rangle \sim \mathcal{N}\langle \mu, \sigma^2\rangle \ .$$

X hat eine Lognormal-Verteilung mit Parametern μ und σ^2.

Dichte und Kennzahlen der Lognormal-Vert. nicht wichtig.

Arbeiten Sie mit logarithmierten Daten!

Beschreibende Statistik mehrdimensionaler Daten

3.1 Grafische Darstellungen einer Stichprobe von Zahlenpaaren

a Beispiele

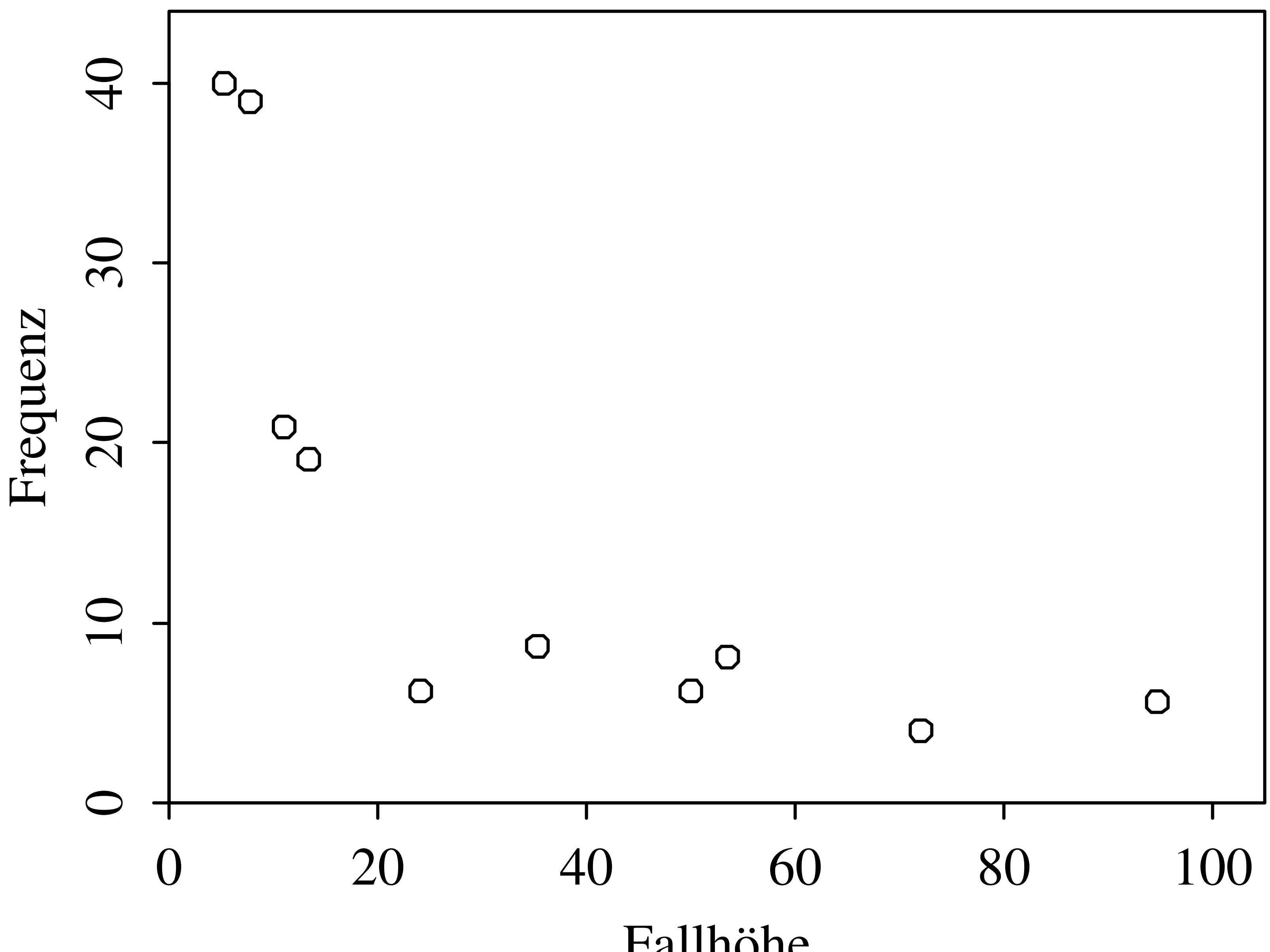
- Druck und Volumen bei einem physikalischen Experiment,
- Sauerstoffzufuhr und Ausbeute bei einem chem. Versuch,
- Blattfläche und mittl. Fotosynthese-Rate bei einer Pflanze,
- Spraydosen-Treibgasverbrauch und Ozongehalt,
- Geschwindigkeit und Bremsweg eines Autos,
- Grösse des Vaters und Grösse des Sohnes.

Zusammenhang zwischen den zwei Variablen?

b Notation: $[x_i, y_i]$. Bivariate Stichprobe.
Streudiagramm (scatterplot).
Streudiagramm funktioniert oft nicht auf Anhieb:

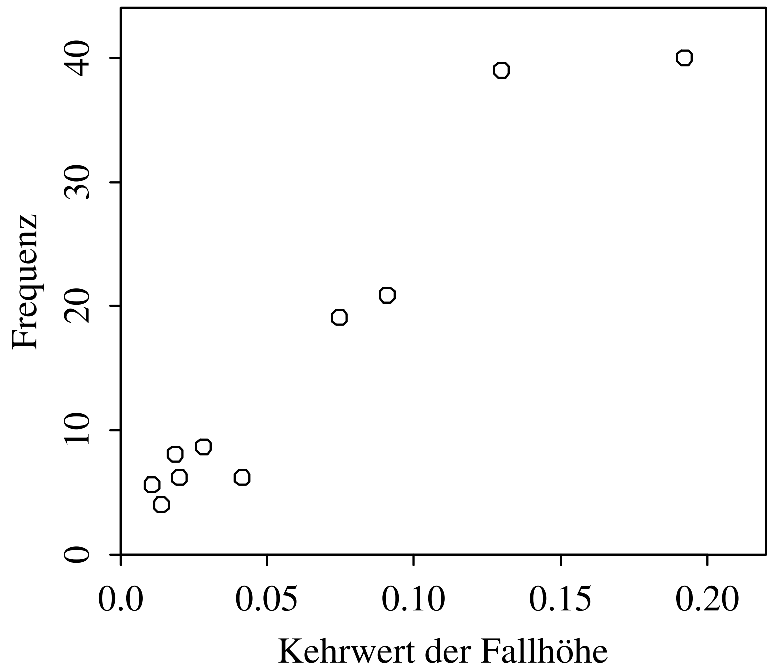
- verdeckte Punkte
- zuviele Punkte
- Ausreisser

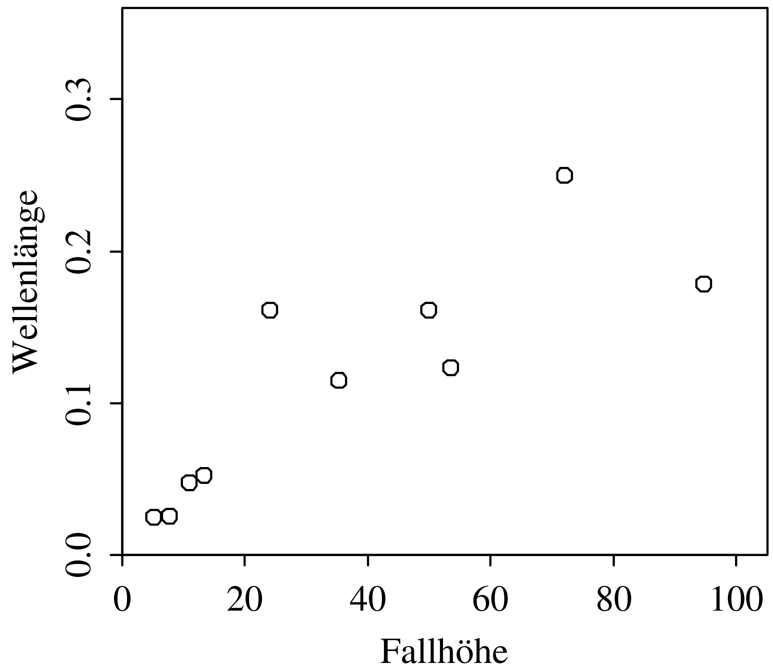
c



d Einfachste und anschaulichste Art einer Beziehung ist ein linearer Zusammenhang.
Oft durch Transformation einer oder beider Var. zu erreichen.
Oder „logarithmierte Achsen", „Log-Skala".

e





3.2 Die Pearsonsche Produktmomenten-Korrelation („gewöhnliche Korrelation”)

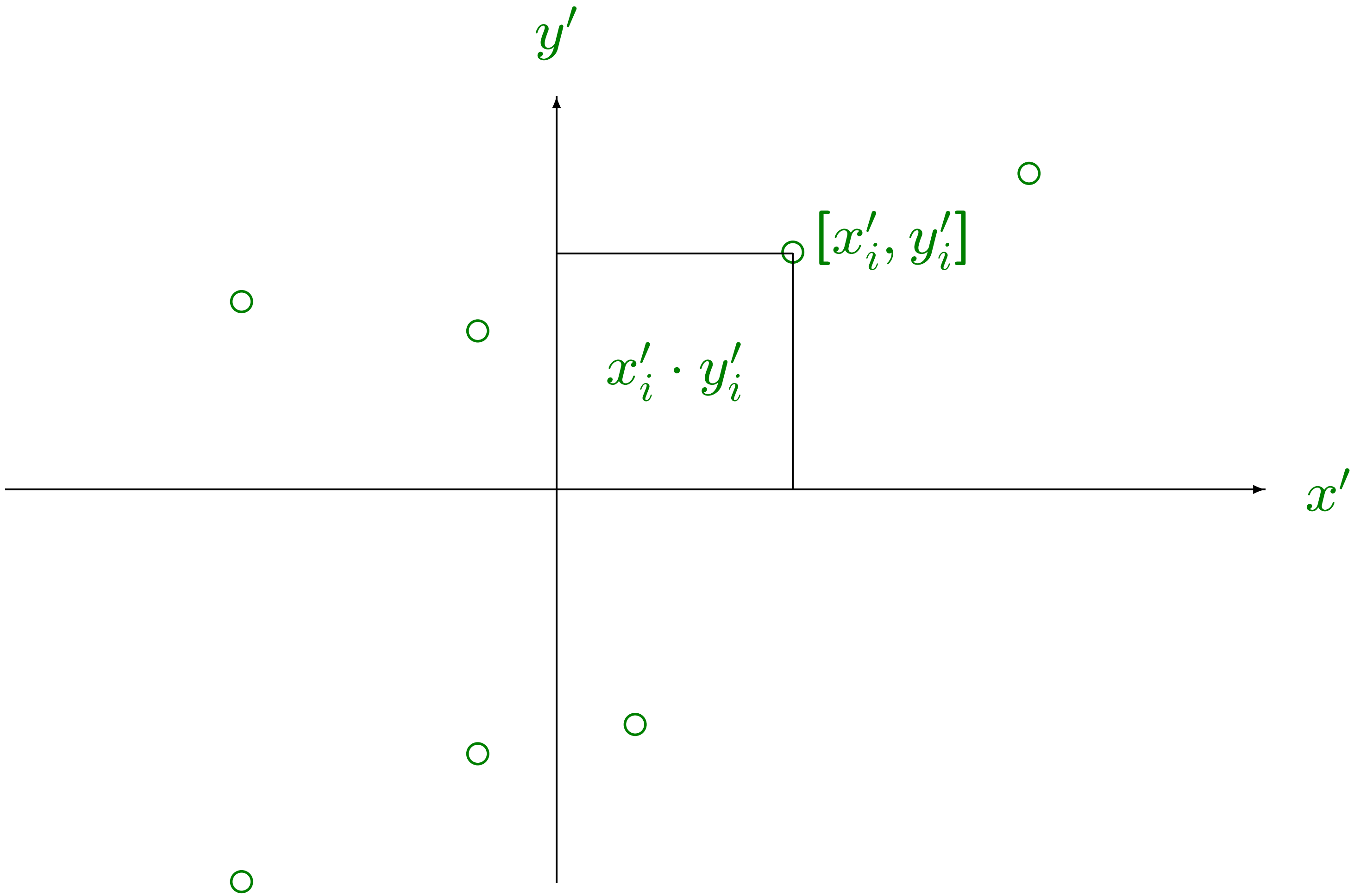
a Kennzahl für den Zusammenhang: Korrelation.

b

$$
\begin{aligned}
x_i' &= (x_i - \bar{x}) / \mathrm{sd}_X, \qquad y_i' = (y_i - \bar{y}) / \mathrm{sd}_Y \\
r_{XY} &= \frac{1}{n-1} \sum_{i=1}^{n} x_i' y_i' \\
r_{XY} &= \frac{s_{XY}}{\mathrm{sd}_X \, \mathrm{sd}_Y} \\
s_{XY} &= \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})
\end{aligned}
$$

r_{XY}: Produktmomenten-Korrelation oder Pearson correlation.

s_{XY}: Kovarianz.


y'
$[x'_i, y'_i]$
$x'_i \cdot y'_i$
x'

d Spezialfälle:

- $x_i = y_i$: $\quad s_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \mathrm{sd}_X^2 = \mathrm{sd}_Y^2$,
  $r_{XY} = 1$.

- $y_i = a + bx_i$:
  $y_i - \bar{y} = b \cdot (x_i - \bar{x})$, $\qquad s_{XY} = \dfrac{1}{n-1}\, b \sum_{i=1}^{n}(x_i - \bar{x})^2 = b$ s
  $\mathrm{sd}_Y^2 = b^2\ \mathrm{sd}_X^2$, $\qquad r_{XY} = b\ \mathrm{sd}_X^2 / (\mathrm{sd}_X \cdot |b|\, \mathrm{sd}_X) = b/|b|$
  $r_{XY} = 1$, falls $b > 0$ ist, $\quad r_{XY} = -1$ falls $b < 0$

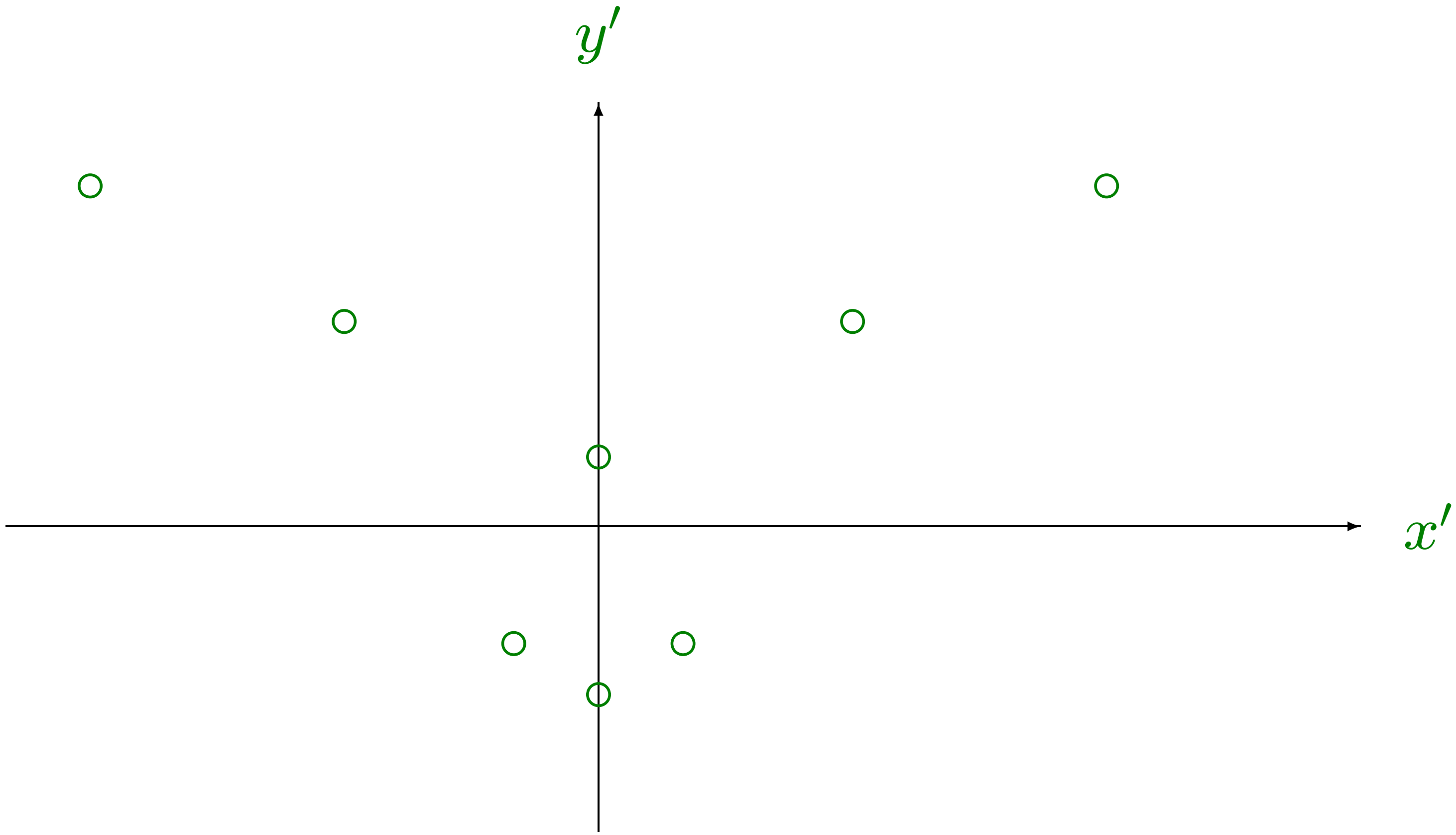
- Punkte eng um eine Gerade: r_{XY} nahe bei ± 1.

Die Produktmomenten-Korrelation misst Stärke & Richtung des linearen Zusammenhanges zwischen X und Y.

e Beispiel Wasserfälle $r_{XY} = -0.74$, $r_{1/X,Y} = +0.96$.

h

- Symmetrische Stichprobe: $r_{XY} = 0$, kein linearer Zus.hang.



i Streudiagramme mit gleichem Korrelations-Koeffizienten können sehr verschieden aussehen!

j r_{XY} nicht robust!

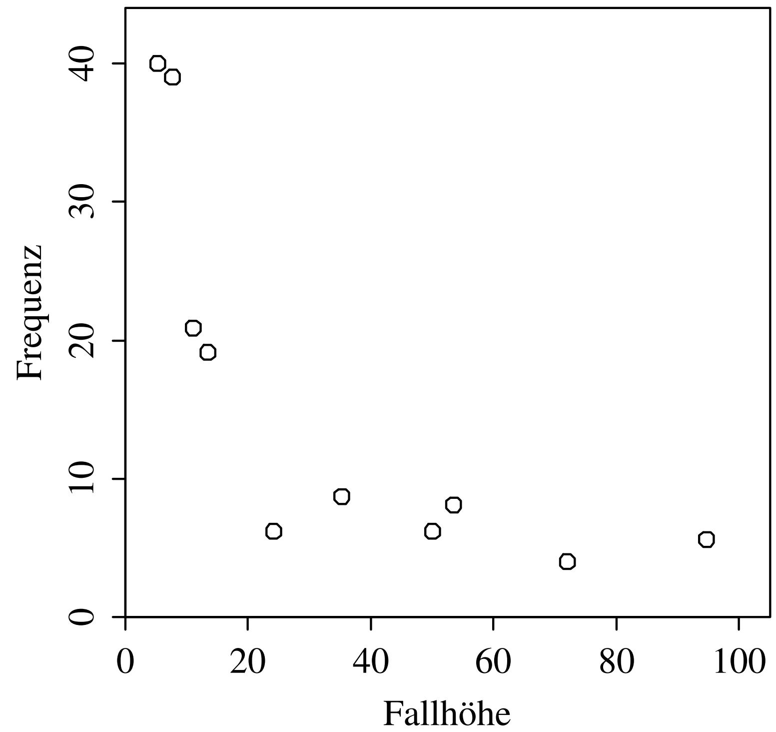
3.3 Rangkorrelationen

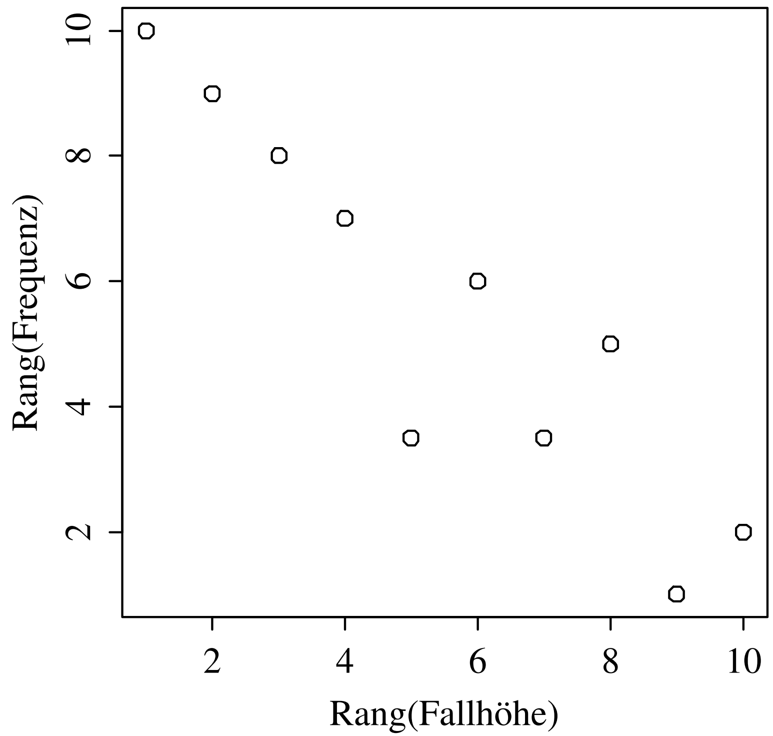
a Spearmansche Rangkorrelation.

Korrelation zwischen den Rängen der x_i und den Rängen der y_i,

$$\begin{aligned} r_{XY}^{(Sp)} &= \frac{s_{\text{Rang}\langle X\rangle,\text{Rang}\langle Y\rangle}}{\text{sd}_{\text{Rang}\langle X\rangle}\,\text{sd}_{\text{Rang}\langle Y\rangle}} \\ &= 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{n} \left(\text{Rang}\langle x_i\rangle - \text{Rang}\langle y_i\rangle\right)^2 , \end{aligned}$$

falls keine Bindungen (geteilten Ränge) auftreten.





c Beispiel der Wasserfälle: $r_{XY}^{(Sp)} = -0.9\mathbf{3}$.

d Wenn die x_i monoton wachsend transformiert werden, ändert sich $r_{XY}^{(Sp)}$ nicht, „bleibt (invariant)".

$r_{XY}^{(Sp)} = \pm 1$, falls Reihenfolgen der X- und der Y-Werte genau gleich oder entgegengesetzt sind.

Rangkorrelation misst Stärke und Richtung des monotonen Zusammenhangs.

3.4 Zur Interpretation von Korrelationen

- Zufall.
- Ursache – Wirkung - Zus'hang. Oft das Gewünschte.
- Indirekter Zusammenhang: X und Y hängen von Z ab.
- Inhomogenität:

  Hämoglobingehalt und Oberfläche der roten Blutkörperchen:

  Für Frauen und Männer kein Zusammenhang.

  Zulassung zur Hochschule.

- „Technische" K.: Standardisierung mit gleicher, ungenau erfasster Grösse.

- Schein-K. zwischen Zeitreihen:

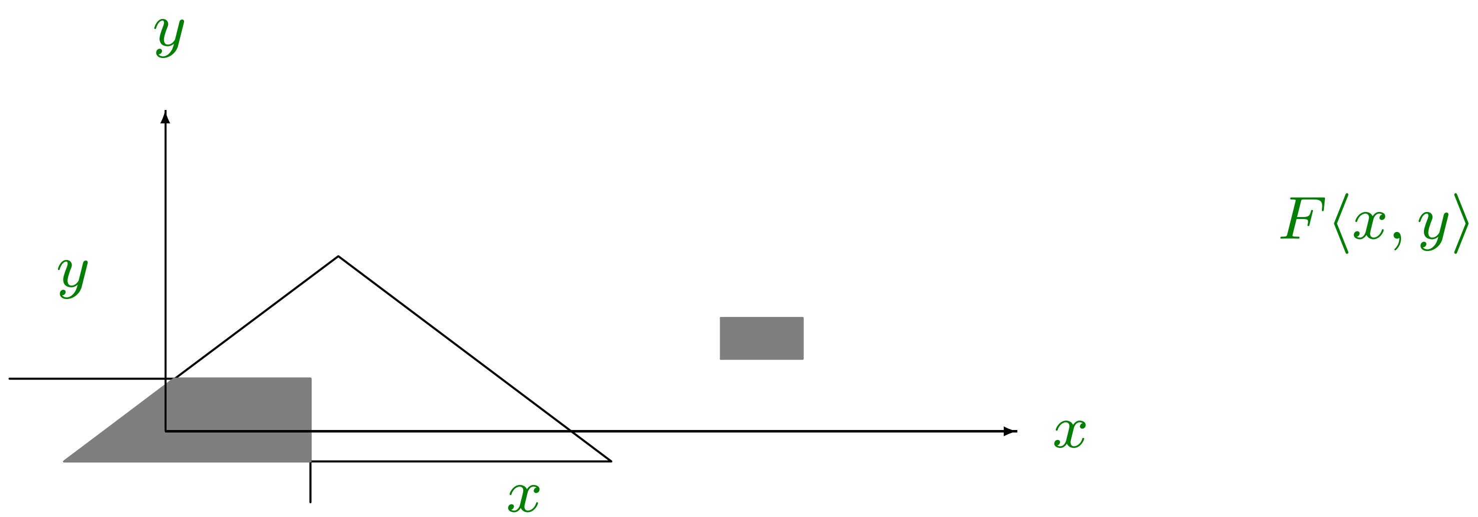
  Störche und Geburten;

  Zahl der Fernsehapparate und der Zahl der Verbrechen (?)

Regression: Wichtig! Später.

6.7 Gemeinsame und bedingte Verteilung

a Beispiel Regentropfen X, Y Koordinaten.
Uniforme Verteilung auf einer Platte: W. eines Ereignisses prop. zur Fläche.



b Gemeinsame Verteilung von 2 Zufallsvariablen $X^{(1)}, X^{(2)}$ gegeben durch kumulative Verteilungsfunktion

$$F\langle x_1, x_2\rangle = P\langle X^{(1)} \le x_1,\; X^{(2)} \le x_2\rangle$$

Wahrscheinlichkeiten für Rechtecke:

$$P\langle a_1 < X^{(1)} \le b_1,\; a_2 < X^{(2)} \le b_2\rangle$$
$$= F\langle b_1, b_2\rangle - F\langle b_1, a_2\rangle - F\langle a_1, b_2\rangle + F\langle a_1, a_2\rangle .$$

c **Dichte**

$$f\langle x_1,x_2\rangle = \lim_{h_1\to 0,h_2\to 0} \frac{P\left\langle x_1-h_1 < X^{(1)} \le x_1\ ,\ x_2-h_2 < X^{(2)} \le x_2\right\rangle}{h_1 h_2} = \frac{\partial^2 F\langle x_1,x_2\rangle}{\partial x_1\,\partial x_2}.$$

Regentropfen-Beispiel:

$$f\langle x_1,x_2\rangle = \begin{cases} 1/\alpha, & \text{falls } [x_1,x_2] \in \text{Platte};\\ 0 & \text{sonst}\ , \end{cases}$$

α: Fläche der Platte.

d

$$P\left\langle [X^{(1)},X^{(2)}] \in A \right\rangle = \int_A f\langle x_1,x_2\rangle\, dx_1\, dx_2$$

e Randverteilung:

$$\begin{aligned} P\langle a_1 < X^{(1)} \le b_1\rangle &= P\langle a_1 < X^{(1)} \le b_1,\ X^{(2)} \text{ beliebig}\rangle \\ &= \int_{x_1=a_1}^{b_1} \left(\int_{x_2=-\infty}^{\infty} f\langle x_1, x_2\rangle\, dx_2 \right) dx_1 \\ f^{(1)}\langle x_1\rangle &= \int_{x_2=-\infty}^{\infty} f\langle x_1, x_2\rangle\, dx_2 \end{aligned}$$

f Bedingte Verteilung, geg. Ereignis A mit $P\langle A\rangle > 0$: wie gehabt.

Exponential-Verteilung:

X Zeit bis zum ersten Impuls des Geigerzählers.

1. Sekunde kein Ton.

Wie ist dann der Rest $Y = X - 1$ der „Wartezeit" verteilt?

$$\begin{aligned} P\langle X > x\rangle &= 1 - F\langle x\rangle = e^{-\lambda x}; \\ P\langle Y > y \mid X > 1\rangle &= \frac{P\langle X > 1 + y\rangle}{P\langle X > 1\rangle} = \frac{\exp\langle -\lambda(1+y)\rangle}{\exp\langle -\lambda\rangle} \\ &= \exp\langle -\lambda y\rangle = P\langle X > y\rangle. \end{aligned}$$

$P\langle X > y\rangle = P\langle Y > y \mid X > 1\rangle$. aber Ereignisse verschieden!

Vt. von $Y = X - 1$, geg. $X > 1$ (also $Y > 0$), $=$ Vt. von X.

$X - 1 | X > 1 \sim \mathcal{E}xp\langle\lambda\rangle$ „Gedächnislosigkeit"

6.8 Unabhängige Zufallsvariable und Korrelation

a Zwei Zufallsvariable $X^{(1)}$ und $X^{(2)}$ heissen unabhängig, falls

$$P\langle X^{(1)} \in \widetilde{A}, X^{(2)} \in \widetilde{B}\rangle = P\langle X^{(1)} \in \widetilde{A}\rangle \cdot P\langle X^{(2)} \in \widetilde{B}\rangle$$

für alle $\widetilde{A}$ und $\widetilde{B}$. – Dichten multiplizieren sich auch:

$$\begin{aligned} f\langle x_1, x_2\rangle &= \frac{\partial^2 F\langle x_1, x_2\rangle}{\partial x_1\, \partial x_2} \\ &= \lim_{h_1 \to 0, h_2 \to 0} \frac{P\left\langle x_1 - h_1 < X^{(1)} \le x_1 \ , \ x_2 - h_2 < X^{(2)} \le x_2\right\rangle}{h_1 h_2} \\ &= \lim_{h_1 \to 0, h_2 \to 0} \frac{P\left\langle x_1 - h_1 < X^{(1)} \le x_1\right\rangle}{h_1} \cdot \frac{P\left\langle x_2 - h_2 < X^{(2)} \le x_2\right\rangle}{h_2} \\ &= f^{(1)}\langle x_1\rangle \cdot f^{(2)}\langle x_2\rangle. \end{aligned}$$

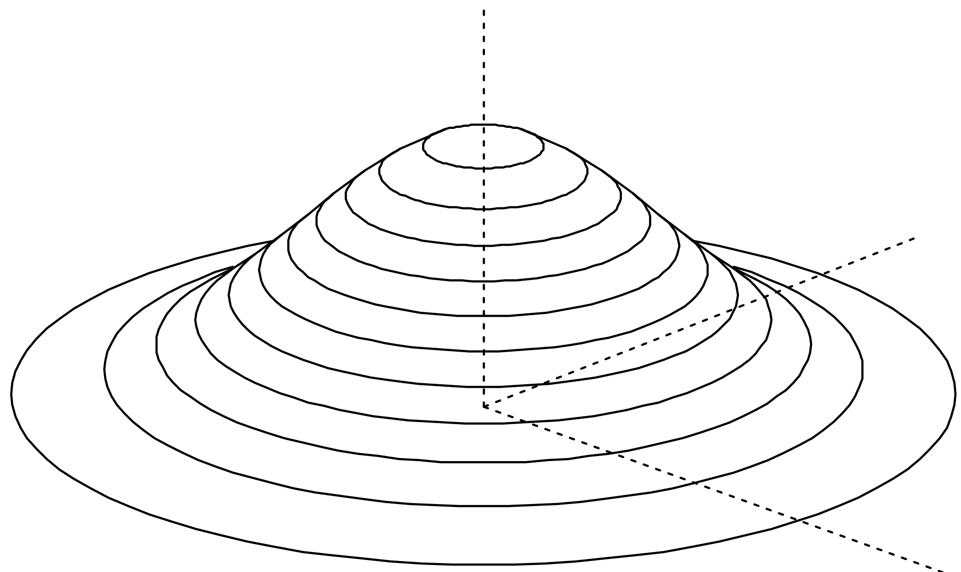
b $X_1 \sim \mathcal{N}\langle \mu_1, \sigma_1^2 \rangle, \quad X_2 \sim \mathcal{N}\langle \mu_2, \sigma_2^2 \rangle$ unabhängig.

Gemeinsame Dichte

$$f\langle x_1, x_2 \rangle = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\langle -\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 \right\rangle \frac{1}{\sqrt{2\pi}\sigma_2} \cdot \exp\left\langle -\frac{1}{2}\left(\frac{x_2-\mu\right.$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\langle -\frac{1}{2}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right) \right\rangle$$

Spezialfall der zweidimensionalen Normalverteilung.



c Mass für die Abhängigkeit: Korrelation $r_{12} = s_{12}/(\mathsf{sd}_1\, \mathsf{sd}_2)$.

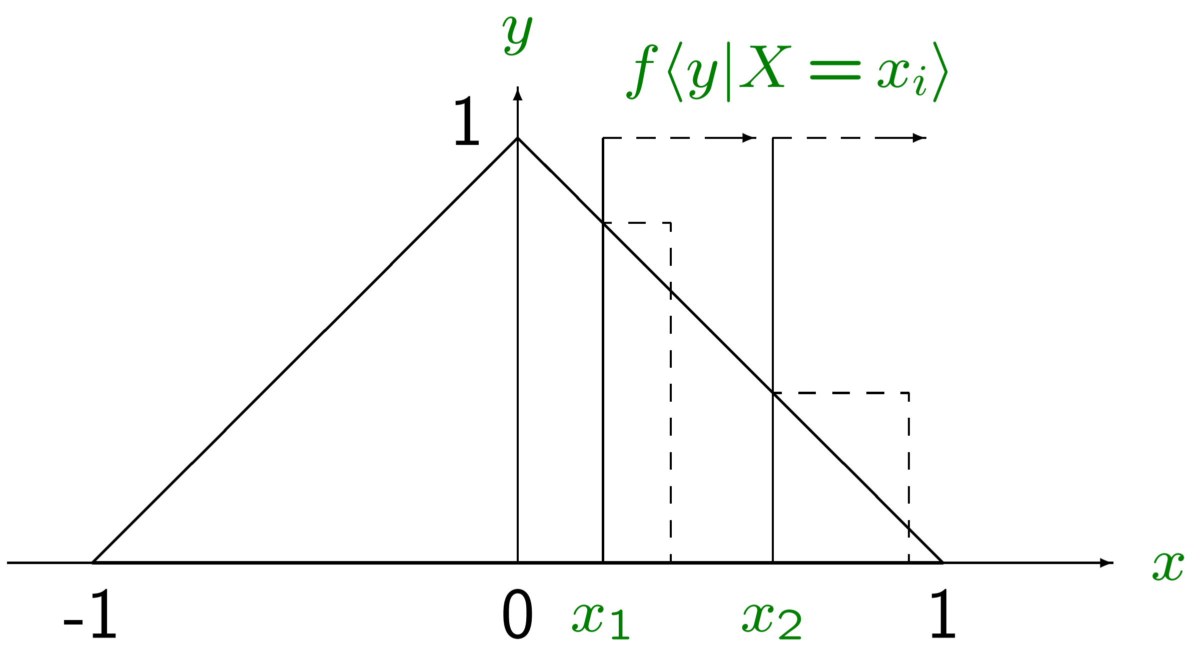
Modell-Grössen einsetzen! $\mu_1 = \mathcal{E}\langle X^{(1)}\rangle,\ \mu_2 = \mathcal{E}\langle X^{(2)}\rangle$.

$$\begin{aligned}
\mathsf{cov}\langle X^{(1)},\, X^{(2)}\rangle &= \mathcal{E}\langle (X^{(1)} - \mu_1)\,(X^{(2)} - \mu_2)\rangle \\
\mathsf{cov}\langle X^{(1)}, X^{(2)}\rangle &= \int\!\!\int (x_1 - \mu_1)\,(x_2 - \mu_2)\; f\langle x_1, x_2\rangle\, dx_1 dx_2 \\
\langle X^{(1)}, X^{(2)}\rangle &= \frac{\mathsf{cov}\langle X^{(1)}, X^{(2)}\rangle}{\sqrt{\mathsf{var}\langle X^{(1)}\rangle \cdot \mathsf{var}\langle X^{(2)}\rangle}}
\end{aligned}$$

d Wenn zwei Zufallsvariable unabhängig sind, sind sie unkorreliert

Unkorrel. Zufallsvar. sind nicht notwendigerweise unabhängig.
Korrelation erfasst nur „lineare Komponente” der Abhängigkeit.

e



f Zweidimensionale Normalverteilung: Dichte $f\langle x_1, x_2\rangle =$

$c{\cdot}\mathrm{exp}\left\langle \alpha_1(x_1-\mu_1)^2 + \alpha_2(x_2-\mu_2)^2 + \alpha_{12}(x_1-\mu_1)(x_2-\mu_2)\right\rangle$

$[\mu_1, \mu_2] = [\mathcal{E}\langle X^{(1)}\rangle, \mathcal{E}\langle X^{(2)}\rangle]$:

Mittelpunkt der Ellipsen konstanter Dichte.

α-s: $\mathrm{var}\langle X^{(1)}\rangle$, $\mathrm{var}\langle X^{(2)}\rangle$, $\mathrm{cov}\langle X^{(1)}, X^{(2)}\rangle$,

Länge und Orientierung der Hauptachsen.

$\alpha_{12} = 0$ unabhängige, also unkorrelierte Normalverteilungen.

Multivariate Statistik!

5.7 Summen von Zufallsvariablen

Wir betrachten der Einfachheit halber diskrete Zufallsvariable.

a Verteilung der Anzahl Erfolge in n Versuchen: Binomial-V.
hergeleitet als V. der Summe von unabh. $X_i \sim \mathcal{B}\text{ernoulli}\langle\pi\rangle$.

Anzahl Unfälle innerhalb einer Woche in zwei Stadtteilen
oder Anzahl Bakterien-Kolonien auf 2 Schalen oder
Anzahl fehlerhafte Gläser in zwei Produktions-Losen.
$X_i \sim \mathcal{P}\langle\lambda_i\rangle$, unabhängig.
Verteilung von $X^{(1)} + X^{(2)} \sim ?$

Gleiche Frage für beliebige gemeinsame Verteilung von $X^{(1)}, X^{(2)}$.

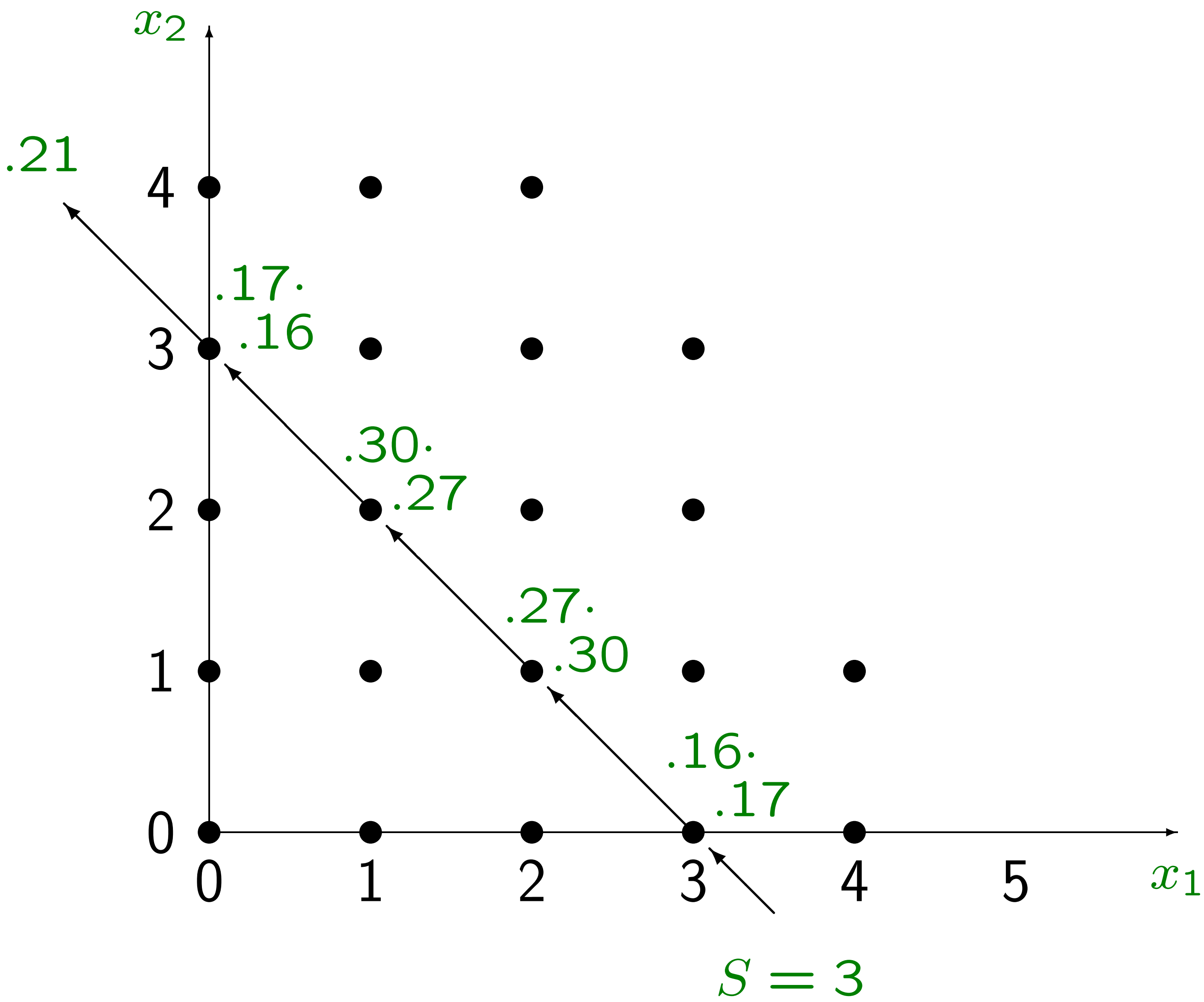
5.7

b Mit Hilfe von Zufallszahlen: siehe Übungen!

Liefert aber keine übersichtlichen Zusammenhänge.

c

$$\begin{aligned}
P\langle S=0\rangle &= P\langle X_1=0,\ X_2=0\rangle = e^{-\lambda_1}\cdot e^{-\lambda_2} = e^{-(\lambda_1+\lambda_2)} \\
P\langle S=1\rangle &= P\langle X_1=0,\ X_2=1\rangle + P\langle X_1=1,\ X_2=0\rangle \\
&= e^{-\lambda_1}\frac{\lambda_2^1}{1!}e^{-\lambda_2} + \frac{\lambda_1^1}{1!}e^{-\lambda_1}e^{-\lambda_2} \\
&= \frac{(\lambda_1+\lambda_2)^1}{1!}\,e^{-(\lambda_1+\lambda_2)}
\end{aligned}$$



x_2
.21
.17·
.16
.30·
.27
.27·
.30
.16·
.17
0
1
2
3
4
0
1
2
3
4
5
x_1
$S = 3$

d Resultat: Wenn man zwei (oder mehrere) unabhängige, Poisson-verteilte Zufallsvariable zusammenzählt, erhält man wieder eine Poisson-verteilte Zufallsvariable. Die Parameter λ_j addieren sich.

$$S = X_1 + X_2 \sim \mathcal{P}\langle \lambda_1 + \lambda_2 \rangle \ .$$

(Anschaulich klar!)

e Binomial-Verteilung

$$X_1 \sim \mathcal{B}\langle n_1,\ \pi\rangle,\ X_2 \sim \mathcal{B}\langle n_2,\ \pi\rangle, \quad X_1,\ X_2 \ \text{unabhängig}$$
$$\Longrightarrow \quad X_1 + X_2 \sim \mathcal{B}\langle n_1 + n_2,\ \pi\rangle.$$

Anschaulich klar: $X_1 + X_2$ = Anz. Erfolge in $n_1 + n_2$ Versuchen.

$X_1 \sim \mathcal{B}\langle n_1, \pi_1\rangle, \quad X_2 \sim \mathcal{B}\langle n_2, \pi_2\rangle, \quad \pi_1 \neq \pi_2$: keine $\mathcal{B}$!

f Allgemein:

$$\begin{aligned} P\langle S=s\rangle &= \sum_{k=0}^{s} P\langle X_1=k, X_2=s-k\rangle \\ &= \sum_{k=0}^{s} P\langle X_1=k\rangle \cdot P\langle X_2=s-k\rangle \end{aligned}$$

i Kennzahlen von S?

$$
\begin{aligned}
\mathcal{E}\langle S\rangle &= \sum_{s=0}^{\infty} s\,P\langle S=s\rangle \\
&= \sum_{s=0}^{\infty}\sum_{x_1=0}^{s}(x_1+s-x_1)\cdot P\langle X_1=x_1, X_2=s-x_1\rangle \\
&= \sum_{x_2=0}^{\infty}\sum_{x_1=0}^{\infty}(x_1+x_2)\cdot P\langle X_1=x_1, X_2=x_2\rangle \\
&= \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} x_1\,P\langle X_1=x_1, X_2=x_2\rangle + \sum_{x_2=0}^{\infty}\sum_{x_1=0}^{\infty} x_2 \,.. \\
&= \sum_{x_1=0}^{\infty}\left(x_1\left(\sum_{x_2=0}^{\infty} P\langle X_1=x_1, X_2=x_2\rangle\right)\right) + \\
&\quad + \sum_{x_2=0}^{\infty}\left(x_2\left(\sum_{x_1=0}^{\infty} P\langle X_1=x_1, X_2=x_2\rangle\right)\right) \\
\mathcal{E}\langle S\rangle &= \mathcal{E}\langle X_1\rangle + \mathcal{E}\langle X_2\rangle
\end{aligned}
$$

$\mathcal{E}\langle S\rangle = \mathcal{E}\langle X_1\rangle + \mathcal{E}\langle X_2\rangle$

j Varianz von S

$$\text{var}\langle S\rangle = \text{var}\langle X_1\rangle + \text{var}\langle X_2\rangle + 2\ \text{cov}\langle X_1, X_2\rangle$$

Falls unabhängig:

$$\text{var}\langle S\rangle = \text{var}\langle X_1\rangle + \text{var}\langle X_2\rangle$$

Der Erwartungswert der Summe von zwei (oder mehreren) Zufallsvariablen ist gleich der Summe ihrer Erwartungswerte.

Die Varianz der Summe von zwei (oder mehreren) unabhängigen Zufallsvariablen ist gleich der Summe ihrer Varianzen.

k **Beispiel**.

Je 100 Würste aus 2 Fabriken werden auf Haltbarkeit geprüft.
Die W., dass sie länger als 2 Wochen geniessbar bleiben,
sei für jede Wurst der einen Fabrik 0.6, bei der anderen 0.3.
Wie gross sind $\mathcal{E}$ und Varianz der Anzahl geniessbarer Würste
unter allen 200 Würsten?

$X_1 \sim \mathcal{B}\langle 100, 0.6\rangle, \quad X_2 \sim \mathcal{B}\langle 100, 0.3\rangle, \quad S = X_1 + X_2$

$$\begin{aligned} \mathcal{E}\langle S\rangle &= \mathcal{E}\langle X_1\rangle + \mathcal{E}\langle X_2\rangle = 100 \cdot 0.6 + 100 \cdot 0.3 = 90 \\ \text{var}\langle S\rangle &= \text{var}\langle X_1\rangle + \text{var}\langle X_2\rangle = 100 \cdot 0.6 \cdot 0.4 + 100 \cdot 0.3 \cdot 0.7 = 45 \end{aligned}$$

Wenn 200 Würste „zufällig" aus der ganzen Produktion
der beiden Fabriken (mit gleichen Anteilen der beiden Teile)
ausgewählt würden, wäre $S \sim \mathcal{B}\langle n = 200, \pi = 0.45\rangle$.

$\mathcal{E}\langle S\rangle = 200 \cdot 0.45 = 90, \quad \text{var}\langle S\rangle = 200 \cdot 0.45 \cdot 0.55 = 49.5 > 45.$

- Einfache Zufalls-Stichprobe:

  $X_1, ..., X_n$ gleich verteilt und unabhängig (i.i.d.)

$$\mathcal{E}\langle X_i\rangle = \mu, \quad \text{var}\langle X_i\rangle = \sigma^2, \quad i = 1, 2, \ldots, n$$

$= \sum_{i=1}^{n} X_i$, $\quad \bar{X} = S/n$ arithmetisches Mittel

$$\begin{aligned} \mathcal{E}\langle S\rangle &= n\mu & \text{var}\langle S\rangle &= n\sigma^2 \\ \mathcal{E}\langle \bar{X}\rangle &= \mu & \text{var}\langle \bar{X}\rangle &= \sigma^2/n. \end{aligned}$$

$$\bar{X} = S/n, \;\longrightarrow\; \mathcal{E}\langle \bar{X}\rangle = \mathcal{E}\langle S\rangle/n, \qquad \text{var}\langle \bar{X}\rangle = \text{var}\langle S\rangle/n^2$$

In Worten:

Der Erwartungswert des arithmetischen Mittels von mehreren gleich verteilten Zufallsvariablen ist gleich dem Erwartungswert der einzelnen Zufallsvariablen.

Die Varianz des arithmetischen Mittels von n unabhängigen gleich verteilten Zufallsvariablen ist gleich der Varianz der einzelnen Zufallsvariablen, dividiert durch n.
Die Standardabweichung des arithmetischen Mittels ist die Standardabweichung der einzelnen Zufallsvariablen, dividiert durch $\sqrt{n}$, die Wurzel aus n.

m Unvermeidliche Verwirrung:
$\mathcal{E}\langle \overline{X} \rangle$ = Erwartungswert eines arithmetischen Mittels,
Kennzahl einer Kennzahl !? Was soll das?

Einfachen Zufalls-Stichprobe = Modell für „Stichprobe" aus Kap.2.

Verteilung von anderen Stichproben-Kennzahlen bestimmen,
z. B. die Verteilung der (empirischen) Varianz?
Verteilung hat Erwartungswert, Median, (theoretische) Varianz,
...

- Simulation, Zufallszahlen.

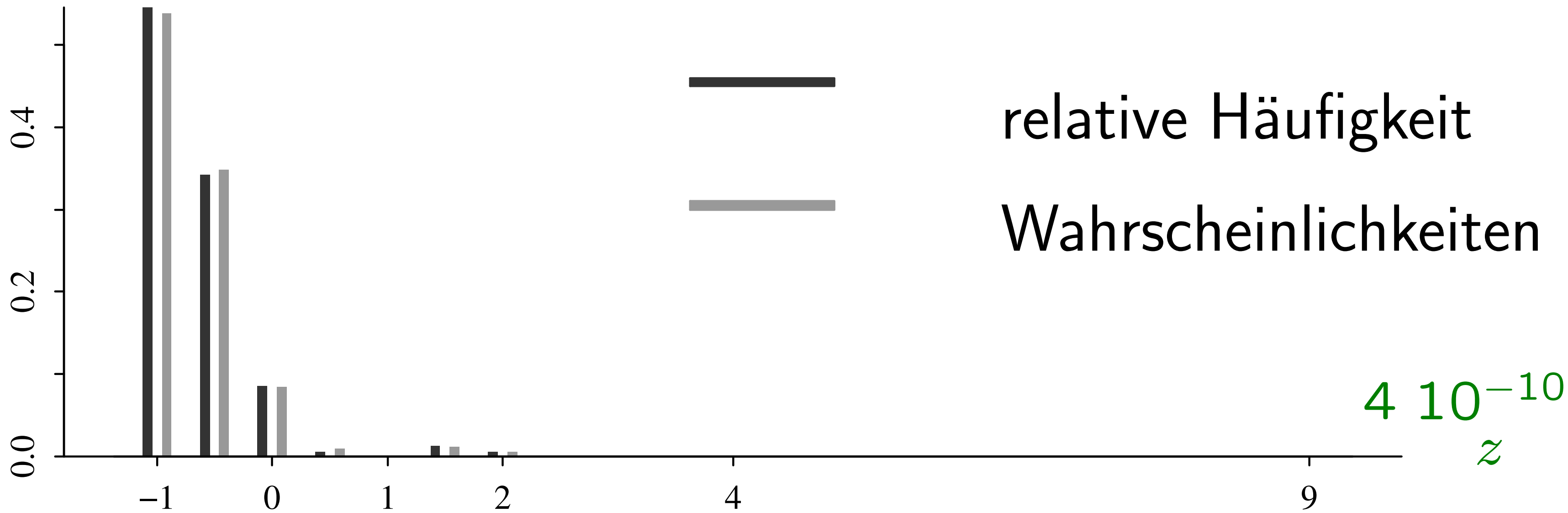
  Je n Zufallszahlen zeigen, was eine Zufalls-Stichprobe als Resultate ergeben könnte:

  n Einzelwerte, Summe, arithmetisches Mittel, Median, emp. Varianz, ...

  1000 mal durchführen $\longrightarrow$ 1000 Werte für med, ...

  simulierte Verteilung des Medians.

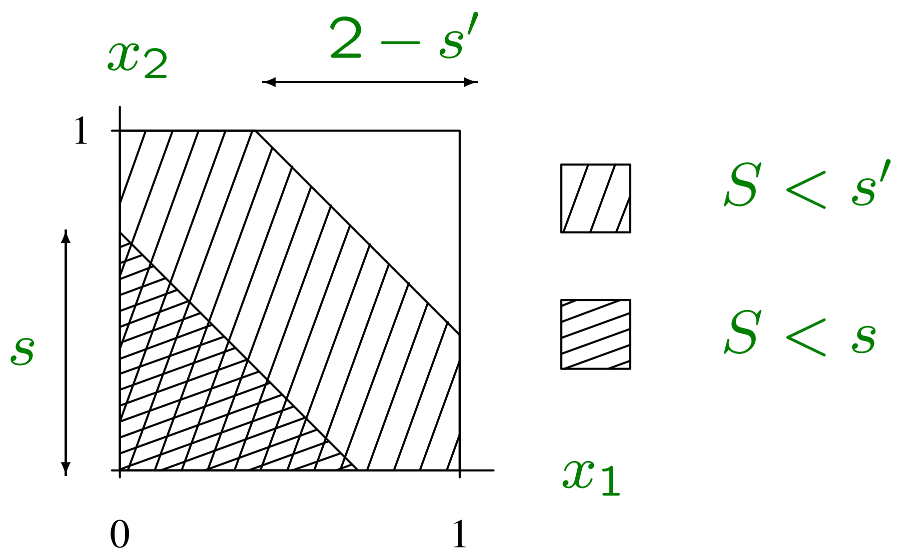
o Beispiel Kartenziehen.

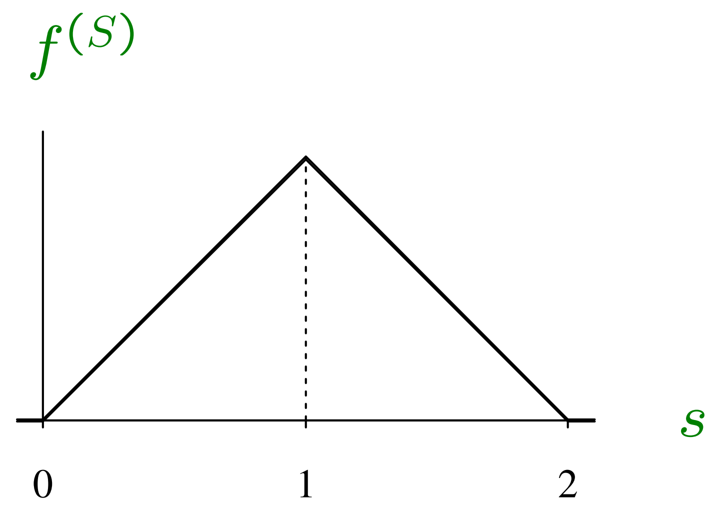


p Arithmetisches Mittel der simulierten $\bar{Z}^{(k)}$:
$\frac{1}{m}\sum_k \bar{z}^{(k)} = -0.6975 \approx \mathcal{E}\langle\bar{Z}\rangle = -0.6774$
$\widehat{\mathrm{var}}_k\langle\bar{z}^{(k)}\rangle = 0.2165 \approx \mathrm{var}\langle\bar{Z}\rangle = \mathrm{var}\langle Z_i\rangle/4 = 0.2258$.
Simulation der Kennzahlen der Verteilung.

6.9 Funktionen von mehreren Zv.: Summen

e $X_1 \sim \mathcal{U}\langle 0,1\rangle,\ X_2 \sim \mathcal{U}\langle 0,1\rangle$, unabhängig. $S = X_1 + X_2$.





$$F^{(S)}\langle s\rangle = \begin{cases} 0 & \text{falls } s \leq 0 \\ s^2/2 & \text{falls } 0 < s \leq 1, \\ 1-(2-s)^2/2 & \text{falls } 1 < s \leq 2, \\ 1 & \text{falls } s > 2\ . \end{cases}$$

$$f^{(S)}\langle s) = \begin{cases} s & \text{falls } 0 \leq s \leq 1, \\ 2-s & \text{falls } 1 < s \leq 2, \\ 0 & \text{sonst}\ . \end{cases}$$

Dreiecks-Verteilung.

f Summen von unabhängigen stetigen Zufallsvariablen $X^{(1)}, X^{(2)}$:

$$\begin{aligned}
F^{(S)}\langle s\rangle &= P\langle S \leq s\rangle = P\langle X^{(1)} + X^{(2)} \leq s\rangle \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{s-x_1} f\langle x_1, x_2\rangle dx_2\right) dx_1 \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{s-x_1} f^{(2)}\langle x_2\rangle dx_2\right) f^{(1)}\langle x_1\rangle dx_1 \\
&= \int_{-\infty}^{\infty} F^{(2)}\langle s-x\rangle\; f^{(1)}\langle x\rangle dx \\
f^{(S)}\langle s\rangle &= \tfrac{d}{ds}\int_{-\infty}^{\infty} F^{(2)}\langle s-x\rangle f^{(1)}\langle x\rangle dx \\
&= \int \frac{d}{ds} F^{(2)}\langle s-x\rangle f^{(1)}\langle x\rangle dx \\
&= \int f^{(2)}\langle s-x\rangle f^{(1)}\langle x\rangle dx.
\end{aligned}$$

i 2 unabhängige, standard-normalverteilte Zufallsvariable:

$$f^{(S)}\langle s\rangle = \int \frac{1}{\sqrt{2\pi}} \exp\langle -\frac{1}{2}x^2\rangle \frac{1}{\sqrt{2\pi}} \exp\langle -\frac{1}{2}(s-x)^2\rangle \, dx$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^2 \int \exp\left\langle -\frac{1}{2}(2x^2 - 2xs + s^2)\right\rangle \, dx$$

$$f^{(S)}\langle s\rangle = \left(\frac{1}{\sqrt{2\pi}}\right)^2 \int \exp\left\langle -\frac{1}{2}\left(2\left(t-\frac{s}{2}\right)^2 - 2\left(\frac{s}{2}\right)^2 + s^2\right)\right\rangle \, dt$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^2 \int \exp\left\langle -\frac{1}{2}\left(\frac{t-s/2}{1/\sqrt{2}}\right)^2\right\rangle dt \cdot \exp\left\langle -\frac{1}{2}\left(\frac{s}{\sqrt{2}}\right)^2\right\rangle$$

$$= \frac{1}{\sqrt{2\pi}\,\sqrt{2}} \exp\left\langle -\frac{1}{2}\left(\frac{s}{\sqrt{2}}\right)^2\right\rangle \int \frac{1}{\sqrt{2\pi}\,(1/\sqrt{2})} \exp\left\langle -\frac{1}{2}\left(\frac{t-s/2}{1/\sqrt{2}}\right)^2\right\rangle \, d$$

Integral integriert Dichte von $\mathcal{N}\langle s/2,\ (1/\sqrt{2})^2\rangle$. Deshalb $=1$.

Es bleibt $S \sim \mathcal{N}\langle 0, 2\rangle$.

j Wenn $X_1 \sim \mathcal{N}\langle \mu_1, \sigma_1^2 \rangle$ und $X_2 \sim \mathcal{N}\langle \mu_2, \sigma_2^2 \rangle$ unabhängig sind, hat die Summe wieder eine Normalverteilung,

$$X_1 + X_2 \sim \mathcal{N}\langle \mu_1 + \mu_2,\ \sigma_1^2 + \sigma_2^2 \rangle$$

Die Erwartungswerte und die Varianzen addieren sich.

m Kennzahlen? Wie für diskrete Verteilungen

$$\begin{aligned} \mathcal{E}\langle X_1 + X_2 \rangle &= \mathcal{E}\langle X_1 \rangle + \mathcal{E}\langle X_2 \rangle \\ \mathrm{var}\langle X_1 + X_2 \rangle &= \mathrm{var}\langle X_1 \rangle + \mathrm{var}\langle X_2 \rangle + 2\ \mathrm{cov}\langle X_1, X_2 \rangle \\ &= \mathrm{var}\langle X_1 \rangle + \mathrm{var}\langle X_2 \rangle \quad \text{falls unabh.} \end{aligned}$$

- **Beispiel Waage**. Messfehler $X \sim \mathcal{N}\langle 0, 0.63^2\rangle$.
  Wie kann man ein Gewicht „auf $0.5\,\mathrm{mg}$ genau" bestimmen?
  Genauer: Messergebnis soll mit Wahrscheinlichkeit von 95% um nicht mehr als $0.5\,\mathrm{mg}$ vom wahren Wert abweichen.

  Mehrmals unabhängig wägen. Resultat $= \bar{X}$.
  Für n unabhängige Messwerte gilt $\bar{X} \sim \mathcal{N}\langle 0, (0.63\,\mathrm{mg})^2/n\rangle$.
  Standardisiertes $\bar{X}$: $Z = \bar{X}/\sqrt{(0.63\,\mathrm{mg})^2/n} \sim \mathcal{N}\langle 0, 1\rangle$.
  Es soll $P\langle -0.5 \le \bar{X} \le 0.5\rangle = 0.95$ sein.
  Es gilt $P\langle -1.96 \le Z \le 1.96\rangle = 0.95$.
  Also $1.96 = 0.5/\sqrt{(0.63\,\mathrm{mg})^2/n}$ oder
  $n = \left(\frac{1.96}{0.5}\, 0.63\right)^2 = 6.09$.

6.9 Funktionen von mehreren Zufallsvariablen

a $T = \min\langle X^{(1)}, X^{(2)}\rangle$. $F^{(T)}\langle t\rangle = P\langle T \le t\rangle = ?$

b Allgemein: $X^{(1)}, X^{(2)}, \ldots, X^{(m)}$ stetig mit Dichte f.

$T = g\langle X^{(1)}, X^{(2)}, \ldots, X^{(m)}\rangle$

$$P\langle T \le t\rangle = \int_{x_1=-\infty}^{\infty} \ldots \int_{x_{m-1}}^{\infty} \int_{x_m \in A_t} f\langle x_1, \ldots, x_m\rangle \, dx_m \, dx_{m-1} \ldots dx_1$$

$A = \{x_m \mid g\langle x_1, x_2, \ldots, x_m\rangle \le t\}$.

6.9

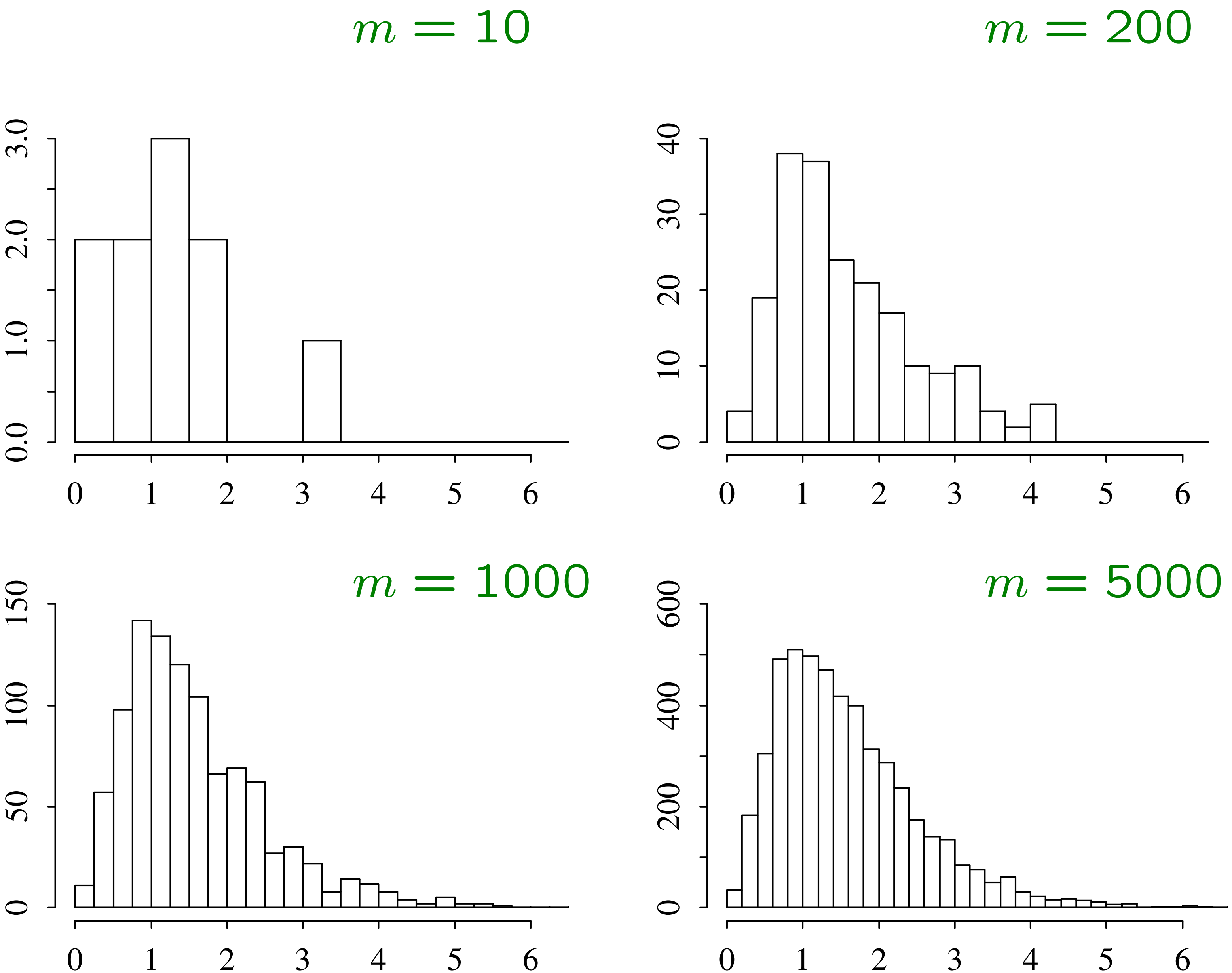
c $X_1, X_2, X_3, X_4, X_5,\ X_i \sim \mathcal{E}xp\langle\lambda\rangle$. Verteilung des Medians?

$$T = \widehat{\text{med}}\langle X_1, X_2, X_3, X_4, X_5\rangle$$

5-faches Integral!

d Simulation der Verteilung des Medians von 5 exp.-vert. X_i.

ℓ	X_1	X_2	X_3	X_4	X_5	T_ℓ
1	0.044	1.422	1.795	0.155	0.414	0.414
2	0.390	0.357	1.844	0.034	3.850	0.390
3	0.048	3.887	1.385	5.600	0.379	1.385
4	0.130	0.663	2.024	0.411	0.806	0.663
5	0.444	2.385	0.506	3.840	1.176	1.176
6	0.606	4.009	0.879	0.591	1.381	0.879
7	1.062	3.855	8.615	3.301	0.409	3.301
8	1.105	5.576	2.443	1.848	0.734	1.848
9	0.866	1.040	4.448	0.237	2.044	1.040
10	1.159	1.852	1.998	4.082	1.799	1.852


$m = 10$
$m = 200$
$m = 1000$
$m = 5000$
0.0
1.0
2.0
3.0
0
10
20
30
40
50
100
150
200
400
600
0 1 2 3 4 5 6

Schliessende Statistik

7 Schätzungen

7.1 Drei Grundfragestellungen der Schliessenden Statistik

a Aufgabe der Schliessenden oder Analytischen Statistik:
Brücke zwischen Modellen und konkreten Daten schlagen.

Beispiel Asbest.

7.1

b **Beispiel Schlafverlängerung.**

Unterschied X_i in der Schlafdauer bei Verwendung von Schlafmittel A gegenüber Mittel B.

Modell: $X_i \sim \mathcal{N}\langle \mu, \sigma^2 \rangle$, unabhängig.

Beobachtungen: 1.2, 2.4, 1.3, 1.3, 0.0, 1.0, 1.8, 0.8, 4.6, 1.4.

Fragen:

1 Um wieviel wird der Schlaf verlängert? Wie gross ist wohl μ?

2 Ist es möglich, dass Mittel A nicht wirksamer ist als B? Ist $\mu \leq 0$ mit den Beobachtungen vereinbar?

3 In welchen Grenzen liegen die Werte μ die aufgrund der Daten noch plausibel erscheinen?

c Methoden:

1 Schätzungen,
2 Tests,
3 Vertrauensintervalle.

d Modell der Zufalls-Stichprobe:

Daten werden unter gleichbleibenden Verhältnissen und unabhängig voneinander gewonnen.

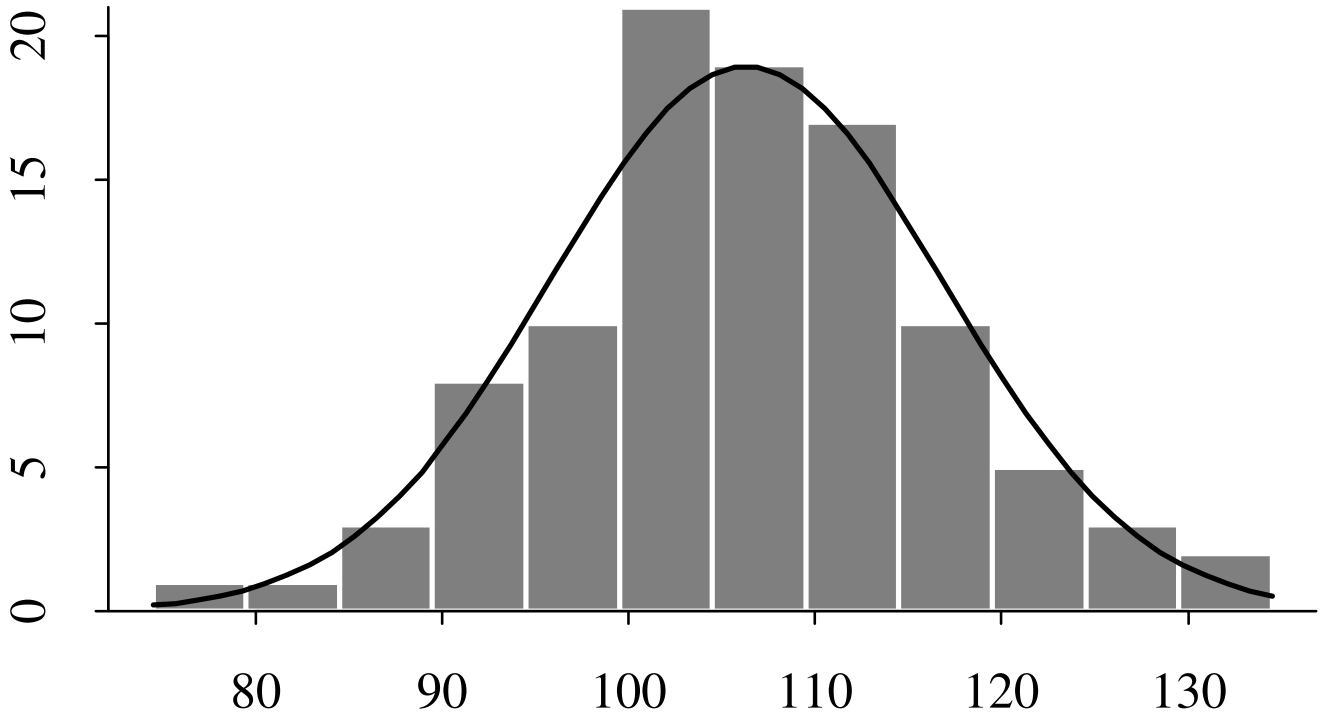
Unabhängig und gleich verteilt
(independent and identically distributed, i.i.d.)

7.2 Schätzungen für ... Normalverteilung

a Problem: Man hat Daten und ein Modell für sie
in Form einer parametrischen Verteilungsfamilie,
$X_i \sim \mathcal{N}\langle\mu, \sigma^2\rangle$ (i.i.d.).
Werte der Parameter festlegen, und zwar so,
dass sie möglichst gut zu den Daten passen.

d Bsp. Schlafverlängerung.
Erwartungswert μ schätzen durch das arithm. Mittel $\overline{X}$
Varianz σ^2 schätzen durch die empir. Varianz $\widehat{\text{var}} = S^2$.
$\overline{X} = 1.58, \quad S^2 = \frac{1}{9}\sum_{i=1}^{10}(X_i - \overline{X})^2 = 1.51 = 1.23^2$.
Modell, das die Daten gut beschreibt: $\mathcal{N}\langle 1.58, 1.23^2\rangle$.

e **Beispiel Küken** : $\overline{X} = 106.25, \quad S^2 = 111.8 = 10.6^2$,
$X_i \sim \mathcal{N}\langle 106.25, 10.6^2 \rangle$.



f Prinzip: Kennzahlen der Stichprobe
mit Kennzahlen der Verteilung identifizieren!
$\longrightarrow$ Momenten-Methode.

g μ ist (auch) Median der Normalverteilung.
Also durch den Median der Stichprobe schätzen.

h $\overline{X}$ und $\widehat{\mathrm{med}}$ sind Schätzungen
für den Parameter μ der Normalverteilung.
Schätzungen sind Funktionen, die
den n Daten eine Zahl und damit
den n Zufallsvariablen $X_1, \ldots, X_n$ eine Zufallsvariable zuordnen.

Schätzungen sind selbst Zufallsvariable.
Bazeichnung: Grossbuchstaben: $\overline{X}$, T
oderHut über Parameter, $\widehat{\mu}$, $\widehat{\sigma}$, allgemein $\widehat{\theta}$.

7.3 Eigenschaften von Schätzungen

a Eigenschaften v. Schätzungen m. Hilfe des W.modells studieren.
Dazu vergessen wir f. den Moment die konkreten Daten wieder.
Legen Modell für die Beobachtungen fest.

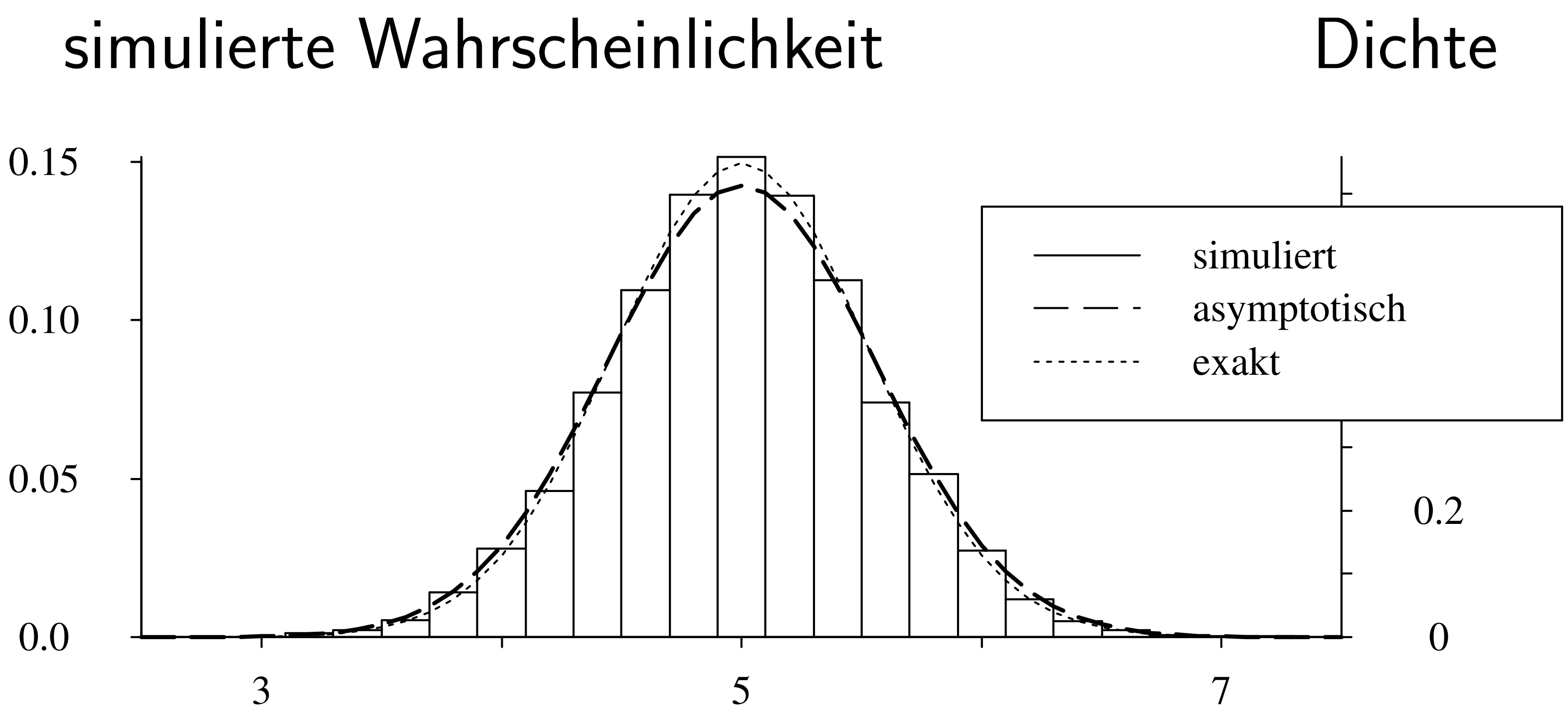
b $X_i \sim \mathcal{N}\langle \mu, \sigma^2 \rangle,\ i = 1, 2, ..., n$ (i.i.d.).

Schätzung $\overline{X}$ oder $\widehat{\text{med}}$? Verteilungen studieren!

Verteilung von $\overline{X}$?

Verteilung von $\widehat{\text{med}}$?
z.B. Simulation. Übungen. Ergebnisse?

Empirischer Median $\widehat{\text{med}}$ von $n = 5$ $X_i \sim \mathcal{N}\langle 5, 1 \rangle$.


simulierte Wahrscheinlichkeit
Dichte
0.15
0.10
0.05
0.0
3
5
7
0.2
0
simuliert
asymptotisch
exakt

d Welche Schätzung soll man wählen?

Die T soll möglichst nahe bei θ liegen!
Ein Mass für die Grösse der Abweichungen ist der
„mittlere quadratische Fehler"

$$\text{MSE} = \mathcal{E}\left\langle (T-\theta)^2 \right\rangle = \text{var}\langle T\rangle + (\mathcal{E}\langle T\rangle - \theta)^2$$

$b = \mathcal{E}\langle T\rangle - \theta$ misst Unterschied zw. dem Erwartungswert von T und dem „Sollwert" θ.
Systematischer Fehler oder Bias b der Schätzung.
Bias $b = 0$: „erwartungstreu"

b^2 und $\text{var}\langle T\rangle$ sollen klein sein!
Oder: Erwartungstreu $(b=0)$ und $\text{var}\langle T\rangle$ möglichst klein!

e $X_i \sim \mathcal{N}\langle 5, 1\rangle,\ i = 1, 2, ..., n$.

$\mathcal{E}\langle \overline{X}\rangle = \mathcal{E}\langle X_i\rangle = \mu = 5$. $\mathcal{E}\langle \widehat{\text{med}}\rangle = \mu = 5$.

$\text{var}\langle \overline{X}\rangle = \frac{1}{n}\text{var}\langle X_i\rangle = \frac{1}{5}$. $\text{var}\langle \widehat{\text{med}}\rangle$? Simulation: ≈ 0.294.

Standardabweichungen („Standardfehler") 0.447 resp. 0.542.

$\overline{X}$ gewinnt! ... auch gegenüber allen anderen.

Für normalvert. Beob. ist $\overline{X}$ die beste Schätzung von μ!

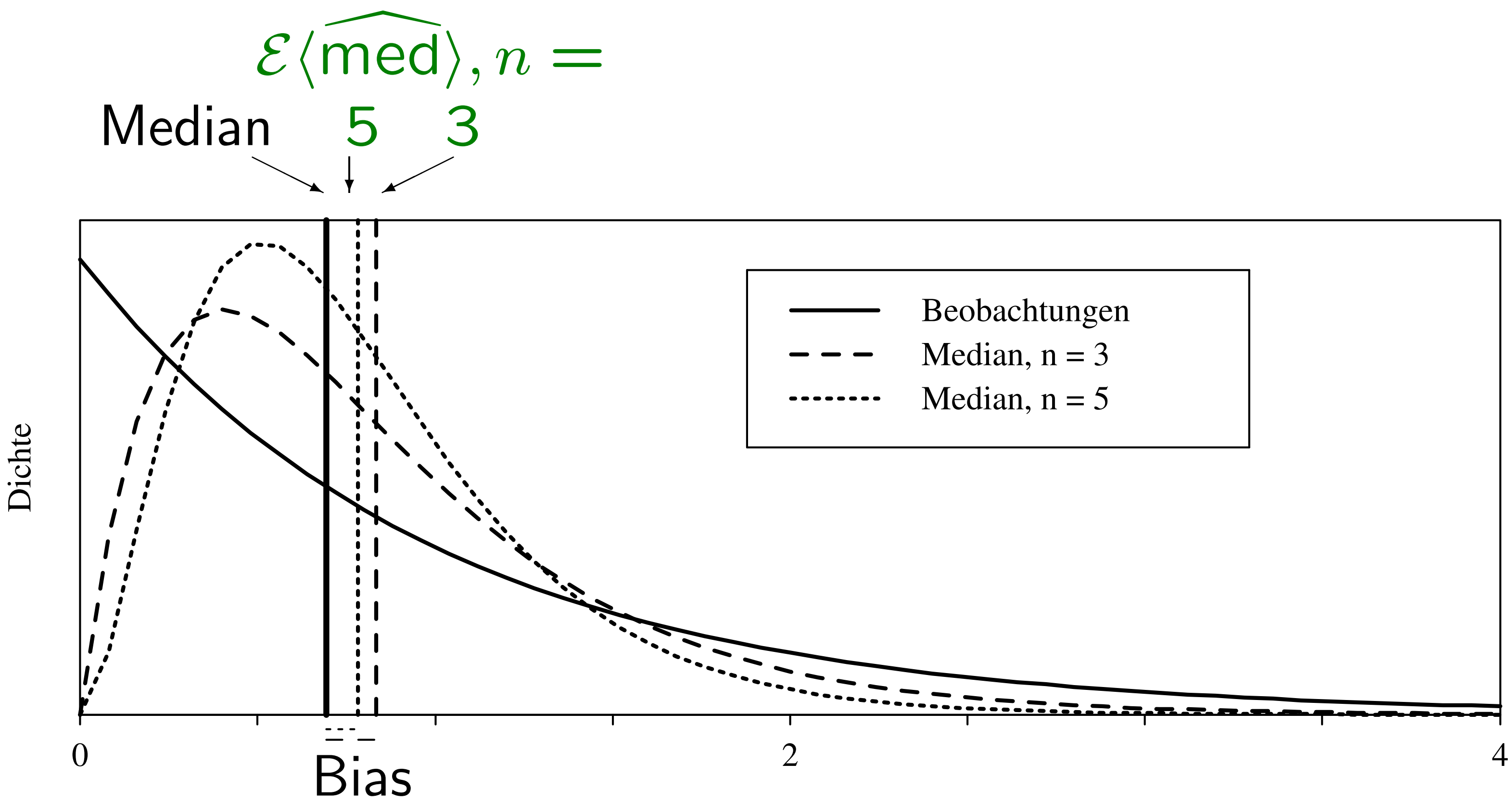
f Exponential-Vert. Schätzung der Halbwertszeit (Median)

$\tau = \log_e\langle 2\rangle/\lambda$. $X_i \sim \mathcal{E}xp\langle\lambda\rangle$.

Empirischer Medians von $n = 3$ und $n = 5$ Beobachtungen.

$\mathcal{E}\langle \widehat{\text{med}} \neq \tau$! $b \neq 0$.

b nimmt mit zunehmendem n ab. Die Verteilung wird schmaler.



$\mathcal{E}\langle\widehat{\text{med}}\rangle, n =$
Median
5
3
Beobachtungen
Median, n = 3
Median, n = 5
Dichte
0
2
4
Bias

g $b = 0$: biasfrei oder erwartungstreu (englisch unbiased).
Viele übliche Schätzungen sind erw.treu,
wenn die Modell-Annahmen „stimmen".

j Varianz $\sigma^2 = \text{var}\langle X_i \rangle$ schätzen!
$S^2 = \frac{1}{n-1} \sum_i (X_i - \overline{X})^2$. Gute Schätzung?
Erwartungstreu? Es sei $\mathcal{E}\langle X_i \rangle = 0$.

$$\begin{aligned}
\mathcal{E}\left\langle \sum_{i=1}^n (X_i - \overline{X})^2 \right\rangle &= \mathcal{E}\left(\sum_{i=1}^n X_i^2 - n\overline{X}^2 \right) \\
&= \sum_{i=1}^n \mathcal{E}\langle X_i^2 \rangle - n\, \mathcal{E}\langle \overline{X}^2 \rangle \\
&= n\,\sigma^2 - n \cdot \frac{1}{n}\sigma^2 = (n-1)\sigma^2
\end{aligned}$$

Deshalb der ominöse Faktor $1/(n-1)$ in S^2.

k Verteilung von S^2 hängt von der Verteilung der X_i ab!

8. Tests

8.3 Weitere Beispiele und Begriffe

a Beispiel Schlafverlängerung
Nullhypothese: kein Unterschied

$$X_i \sim \mathcal{N}\langle 0, 1.5\rangle; \quad X_i \text{ unabhängig}$$

b Alternativen: $(H_A)_\mu : X_i \sim \mathcal{N}\langle \mu, 1.5\rangle$ mit $\mu > 0$ (einseitig).

Beobachtungen zusammenfassen! $\overline{X}$ „Test-Statistik" U.

8.3

c Verwerfungsbereich?

$$U = \overline{X} \sim \mathcal{N}\langle 0, 1.5/10\rangle$$

falls H_0 gilt.

$T = \overline{X}/\sqrt{0.15} \sim \mathcal{N}\langle 0, 1\rangle$ $\longrightarrow$ Extreme Werte: $T \geq 1.64$.

d $\bar{x} = 1.58,\ t = 1.58/1.5 = 1.05 < 1.64$.

Verwerfung. Effekt statistisch nachgewiesen.

f Modern: Statt t mit c zu vergleichen, können wir

$$P\langle T > 1.05\rangle = P\langle \overline{X} > 1.58\rangle = 0.15$$

ausrechnen.

W. von mindestens so extremen Werten wie der beobachtete.

P-Wert p

$$t > 1.64 \quad \Longleftrightarrow \quad p < 0.05$$

8.4 Vorgehen bei einem statistischen Test (Rezept)

a Problem formulieren, in Worten.

b Nullhypothese H_0. Modell für die Beobachtungen.
Oft: Parametrische Familie, H_0 legt Parameter fest.
In der Regel möchte man H_0 gerne widerlegen.

c Alternativen H_A werden in Betracht ziehen
(einseitig oder zweiseitig)

d Test-Statistik.
Verteilung von U unter Alternativen soll möglichst stark verschieden sein von der Vt. unter der Nullhypothese (Macht!).
Oft: Parameter testen: $H_0: \ \theta = \theta_0$.
$\longrightarrow$ Test-Statistik $U = \widehat{\theta}$.

Art des Experimentes, Beobachtungs-Situation [a]

↓

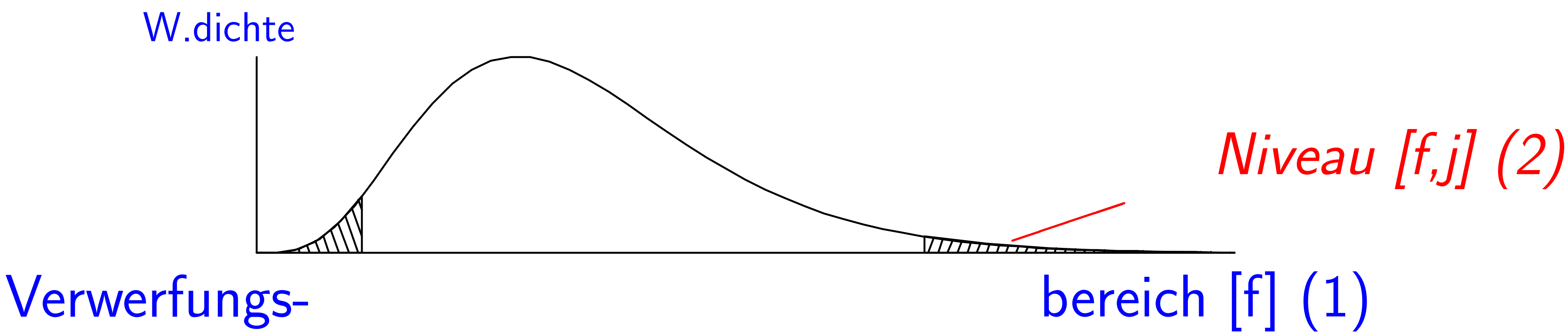
Modelle für eine Beobachtung:

Nullhypothese [b] | Alternativen [c]

Alternativen [c] → Test-Statistik [d]

(1)

Verteilung der Test-Statistik unter der Nullhypothese [e]



e Verteilung $\mathcal{F}_0\langle U\rangle$ von U unter H_0
Standardisierte Test-Statistik T.

f Klassische Variante: Verwerfungsbereich bestimmen.
Willkürlich: Niveau α wählen.
„Kritischer Wert", oft aus Tabelle oder Diagramm.

g Bis hierher: Test-Vorschrift „fertig verpackt"
in den Lehr- und Nachschlagebüchern nachzulesen.

Art des Experimentes, Beobachtungs-Situation [a]

Modelle für eine Beobachtung:

Nullhypothese [b]

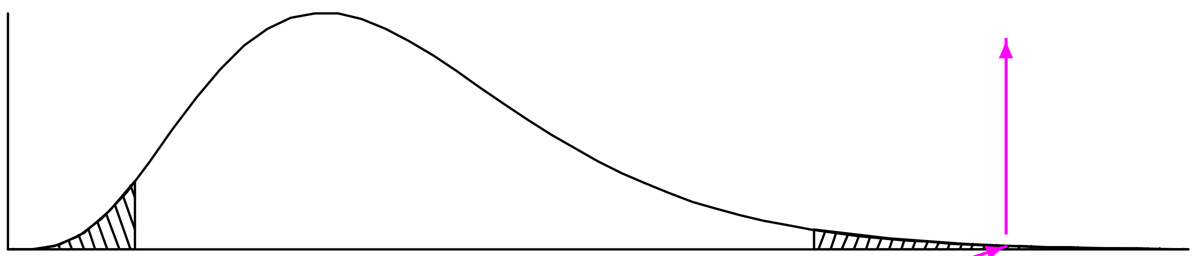
Alternativen [c]

Test-Statistik [d]

(1)

Verteilung der Test-Statistik unter der Nullhypothese [e]

Nullhypothese verworfen! [i] (3)



Ausrechnen des Wertes der Test-Statistik [h] (3) h Beobachtungen $\longrightarrow$ Wert t der (standard.) Test-Statistik T

i Klassisch: Entscheidung:
Falls $t \in K$, wird die Nullhypothese verworfen.

j Modern: P-Wert bestimmen. Entscheidung:
Falls $p < 5\%$, wird die Nullhypothese verworfen.

8.5 Tests für eine St. oder zwei gepaarte Stichproben

a Problem „eine Stichprobe": $X_i,\ i = 1, 2, \ldots, n$, unabhängig.

Kann ein bestimmter Lageparameter (Erwartungswert, Median) gleich einem vorgegebenen Wert μ_0 sein?

$H_0: \quad \mu = \mu_0$.

μ_0: physikalische Konstante, chemische Konzentration, Grenzwert.

b Problem „zwei gepaarte oder verbundene Stichproben”:

- Blutdruck vor und nach Medikamenten-Einnahme,
- Schlaflänge bei zwei Schlafmitteln,
- Grösse von Vater und Sohn (Beobachtungseinh = Familie)
- Anzahl beobachteter Verhaltenselemente von Mäusen bei Licht und bei Dunkelheit. „Verbundene Stichproben”

$X_i =$ Differenz der Zielgrösse in den beiden Zuständen.
(evtl. nach Transformation).

Frage: Differenz „im wesentlichen gleich null”?

$H_0: \quad \mu = 0.$

d z-Test: besprochen.

H_0 : $X_i \sim \mathcal{N}\langle \mu_0, \sigma_0^2 \rangle,\ i = 1, \ldots, n$, unabhängig;
die Varianz σ_0^2 sei bekannt

H_A : $X_i \sim \mathcal{N}\langle \mu, \sigma_0^2 \rangle, \quad i = 1, \ldots, n$, unabhängig, mit
(a) $\mu > \mu_0$; (b) $\mu < \mu_0$; (c) $\mu \neq \mu_0$

U : $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

$\mathcal{F}_0\langle U \rangle$: $\overline{X} \sim \mathcal{N}\langle \mu_0, \sigma_0^2/n \rangle$

K : Testgrösse standarisieren: $Z = \frac{\overline{X} - \mu_0}{\sigma_0/\sqrt{n}} \sim \mathcal{N}\langle 0, 1 \rangle$

Kritische Werte für Z:

(a) c so, dass $1 - \Phi\langle c\rangle = 0.05 \Rightarrow c = 1.64\,;\quad K = \{Z \geq 1.64\}$

(b) c so, dass $\Phi\langle c\rangle = 0.05 \Rightarrow c = -1.64;\quad K = \{Z \leq -1.64\}$

(c) c_1 so, dass $1 - \Phi\langle c_1\rangle = 0.025 \Rightarrow c_1 = 1.96;$
c_0 so, dass $\Phi\langle c_0\rangle = 0.025 \Rightarrow c_0 = -c_1;\quad K = \{|Z| \geq 1.96\}$

Kritische Werte für standardisierte Testgrösse unabhängig von n, μ_0, σ_0.

e Beispiel Reifen (nach Lehn+Wegmann, 1992).

Vergleich von 2 Profilen auf Winterreifen bezügl. Bremswirkung

10 Testfahrzeuge. $\sigma_0 = 3$ aus früheren Versuchen.

i	Profil A	Profil B	Differenz	Vorzeichen
1	44.5	44.9	0.4	+
2	55.0	54.8	− 0.2	−
3	52.5	55.6	3.1	+
4	50.2	55.2	5.0	+
5	45.3	55.6	10.3	+
6	46.1	47.7	1.6	+
7	52.1	53.0	0.9	+
8	50.5	49.1	− 1.4	−
9	50.6	52.3	1.7	+
10	49.2	50.7	1.5	+
mittlere Differenz			2.29	

$z = 2.29/(3/\sqrt{10}) = 2.41 \in K$:

$\longrightarrow$ Die Reifensorten unterscheiden sich statistisch signifikant in der Länge des Bremsweges.

f Meistens: Nullhypothese $X_i \sim \mathcal{N}\langle\mu_0, \sigma^2\rangle$, unabhängig,

σ unbekannt.

Nicht eine einzige Verteilung, sondern

alle Normalverteilungen mit Erwartungswert μ_0.

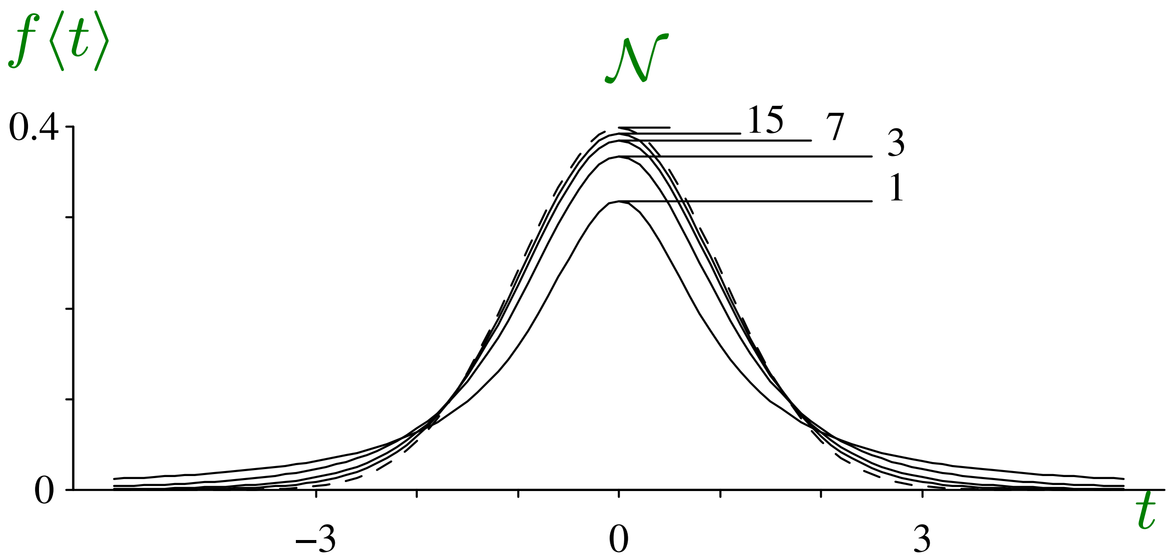
σ^2: Stör-Parameter, nuisance parameter.

σ durch den Schätzwert $\widehat{\sigma} = S$ ersetzen!

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{\overline{X} - \mu_0}{\sqrt{\frac{1}{n(n-1)} \sum (X_i - \overline{X})^2}}$$

g S schwankt zufällig. T nicht normalverteilt.
Verteilung von T hängt nur vom Stichprobenumfang n ab.
$m = n - 1$ „Freiheitsgrade".

Verteilung von T: grössere Varianz und langschwänziger.



t-Test von Student.

i $S/\sqrt{n}$: Schätzung der Standard-Abweichung $\sigma_0/\sqrt{n}$
des arithmetischen Mittels $\bar{X}$,
„Standardfehler", standard error se.

j Beispiel Reifen: $S^2 =$
$\frac{1}{9}\left((0.4-2.29)^2+(0.2-2.29)^2+\ldots+(1.5-2.29)^2 \right) = 3.31^2$,
Standardfehler
se $= 3.31/\sqrt{10} = 1.048$, $t = 2.29/1.048 = 2.18$.
$m = 10 - 1$. Krit.Wert $2.26 > |t|$. H_0 nicht verworfen!

Einseitiger Test: Kritischer Wert 1.83. Signifikant!
Interpretation?

k Vorzeichen-Test (oder Zeichen-Test).
Testgrösse = Anzahl Beobachtungen grösser als μ_0
= Anzahl positive Vorzeichen von $X_i - \mu_0$.

l H_0 : $X_i \sim \mathcal{F}_0$ mit Median μ_0, unabhängig.

H_A : $X_i \sim \mathcal{F}$ mit Median μ;
(a) $\mu > \mu_0$, (b) $\mu < \mu_0$, (c) $\mu \neq \mu_0$.

U : $U =$ Anzahl $\{i \mid X_i > \mu_0\}$

$\mathcal{F}_0\langle U\rangle$: $P_0\langle X_i > \mu_0\rangle = 1/2$, $U \sim \mathcal{B}\langle n,\ 1/2\rangle$.

K : Tabelle!

m Beispiel Reifen: 8 positive Vorzeichen von $n = 10$.
Tabelle: $c_0 = 1$, $c_1 = n - c_0 = 9$.
Nullhypothese durch den Vorzeichentest nicht abgelehnt.

o Keine Normalverteilung vorausgesetzt!
Aber: Information „verschenkt".

p Rangsummen-Test von Wilcoxon für gepaarte Stichproben
oder Vorzeichen-Rangsummen-Test
(signed rank test, one-sample Wilcoxon test).

H_0 : $X_i \sim \mathcal{F}_0$, unabh., $\mathcal{F}_0$ stetig und symmetrisch bez. μ_0

H_A : $X_i \sim \mathcal{F}$, unabhängig, $\mathcal{F}$ symmetrisch bez.
(a) $\mu > \mu_0$; (b) $\mu < \mu_0$; (c) $\mu \neq \mu_0$.

U : Nach folgendem Rezept zu bilden:

1. Bilde $X_i' = X_i - \mu_0$, streiche die negativen Vorzeichen, bilde die Ränge R_i.

$$R_i = \text{rank} \left\langle |X_i'| \,\middle|\, |X_1'|, |X_2'|, \ldots, |X_n'| \right\rangle .$$

Beispiel Reifen: $\mu_0 = 0$, $X_i - \mu_0$ Differenz der Bremswege.
Ränge R_i: 2, 1, 8, 9, 10, 6, 3, 4, 7, 5.

2. Summiere die R_i, für die $X_i - \mu_0 > 0$
Beispiel: 2+8+9+10+6+3+7+5=50.
Sei $V_i = 1$, falls $X_i - \mu_0 > 0$ ist, $V_i = 0$ sonst.

$$U^+ = \sum_{i=1}^n V_i R_i.$$

Beispiel: $u^+ = 1 \cdot 2 + 0 \cdot 1 + 1 \cdot 8 + 1 \cdot 9 + 1 \cdot 10 + 1 \cdot 6 + 1 \cdot 3 + 0 \cdot 4 + 1 \cdot 7 + 1 \cdot 5 = 50$.
$U^- = \sum_{i=1}^n (1 - V_i) R_i$. $U^+ + U^- = \sum_{i=1}^n R_i = n(n+1)/2$.

$\mathcal{F}_0\langle U\rangle$: Verteilung von U^+ oder U^- hängt nicht von der Verteilung $\mathcal{F}$ ab (solange sie stetig ist).

K : Bis $n = 30$ aus Tabelle:

(a) c aus der Zeile „c, einseitig";
$K = \{U^+ \geq n(n+1)/2 - c\} = \{U^- \leq c\}$;

(b) c aus der Zeile „c, einseitig"; $K = \{U^+ \leq c\}$;

(c) c_0 aus der Zeile „c_0, zweiseitig";
$$K = \{U^+ \leq c_0\} \cup \{U^+ \geq n(n+1)/2 - c_0\}$$
$$= \{\min\langle U^+, U^-\rangle \leq c_0\} \ .$$
Für grössere n: Approximation durch die Normalvert.

n	1	2	3	4	5	6	7	8	9	10
c, einseitig	-	-	-	-	0	2	3	5	8	10
c_0, zweiseitig	-	-	-	-	-	0	2	3	5	8
n	11	12	13	14	15	16	17	18	19	20
c, einseitig	13	17	21	25	30	35	41	47	53	60
c_0, zweiseitig	10	13	17	21	25	29	34	40	46	53
n	21	22	23	24	25	26	27	28	29	30
c, einseitig	58	65	73	81	89	98	107	116	126	137
c_0, zweiseitig	67	75	83	91	100	110	119	130	140	151

Beispiel Reifen: $u^+ = 50$, $u^- = 5$.

Tabellenwert $c_0 = 8$, $n(n+1)/2 - c_0 = 47 < u^+$ $(u^- \leq 8)$.

Nullhypothese verworfen.

U ist asymptotisch normalverteilt.

$$\mathcal{E}\langle U^+\rangle = \frac{n(n+1)}{4}, \quad \mathrm{var}\langle U^+\rangle = \frac{n(n+1)(2n+1)}{24}$$

$$Z^+ = \frac{U^+ - \mathcal{E}\langle U^+\rangle}{\sqrt{\mathrm{var}\langle U^+\rangle}} = \frac{U^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \approx\sim \mathcal{N}\langle 0, 1\rangle$$

Verwerfungsbereich (zweiseitig): $|Z^+| \geq 1.96$.

q Keine („Bindungen”), ties) oder Nullen!
Nullen $(X_i = \mu_0)$ weglassen.
(Vert. von T = bedingte Vert., gegeben die Anzahl Nullen).
Bindungen: Ränge aufteilen. Asymptot. Näherung mit korrig. Varianz.

Herleitung der Verteilung von U^+:
Wie verteilen sich die Vorzeichen auf die Ränge, wenn H_0 gilt?

$\lvert x_i' \rvert$	0.2	0.4	0.9	1.4	1.5	1.6	1.7	3.1	5.0	10.3
Rang$\langle \lvert x_i' \rvert \rangle$	1	2	3	4	5	6	7	8	9	10
Vorz$\langle x_i' \rangle$	−	+	+	−	+	+	+	+	+	+
v_i	0	1	1	0	1	1	1	1	1	1

3. Zeile zufällig, Bernoulli-vert.
Jede Folge von 10 Vorzeichen hat gleiche Wahrscheinlichkeit $1/2^{10}$.
$\longrightarrow$ Verteilung von U^+.

r Test-Statistiken, deren Verteilung nicht von einem konkreten parametrischen Modell für die Beobachtungen abhängt, und die entsprechenden Tests nennt man nicht-parametrisch (oder verteilungsfrei). Rangtest.

Genaueres, Allgemeineres unter den Stichworten Randomisierungs-Tests oder Permutations-Tests.

s Wilcoxon-Test nützt absolute Grösse der positiven gegenüber den negativen $X_i - \mu_0$ aus.
Nullhypothese ist „schärfer" als beim Vorzeichnetest:
Es wird Symmetrie vorausgesetzt.
Für Differenzen der Paare bei verbundenen Stichpr. naheliegend.

t Welchen Test wählen?

8.6 Interpretation von Testergebnissen

a Verwerfung der Nullhypothese. Interpretation?

(1) Ein Effekt ist nachgewiesen, eine „Alternative" ist richtig,

(2) Nullhypothese auf andere Weise verletzt:

- Daten enthalten systematischen Fehler,
- X_i sind nicht unabhängig,
- nicht normalverteilt (beim t-Test),
  nicht symmetrisch verteilt (beim Rangsummen-Test),
  nicht alle gleich verteilt ($\text{med}\langle X_i \rangle$ versch. beim Vorz.t.),

(3) es ist zufällig das unwahrscheinl. Ereignis K eingetreten.

(3) ist unvermeidbar. W. durch Signifikanzniveau kontrolliert.
Man möchte (2) vermeiden, um (1) schliessen zu können.
Aber wie?

b (A) $X_i \sim \mathcal{N}\langle\mu, \sigma_0^2\rangle$, σ_0^2 bekannt $\longrightarrow$ z-Test.

(B) $X_i \sim \mathcal{N}\langle\mu, \sigma_0^2\rangle$, σ_0^2 unbekannt $\longrightarrow$ t-Test.

(C) $X_i \sim \mathcal{F}$ symmetrisch um μ, sonst bel. $\longrightarrow$ Rangsummen-Test.

(D) $X_i \sim \mathcal{F}_i$ mit Median μ, sonst beliebig $\longrightarrow$ Vorzeichen-Test.

Unabhängigkeit der X_i.

c • Annahme, dass die Nullhypothese oder eine Alternative „gilt", ist Voraussetzung für die Anwendung des Tests.

d • Gesamte Nullhypothese führt zu „statistischem Widerspruch".
Abweichungen von den anderen Annahmen wegdiskutieren (soweit möglich überprüfen).

e Tests verwenden, die möglichst wenige Voraussetzungen brauchen!
Also immer Rangsummen-Test statt t- oder z-Test,
immer Vorzeichen-Test statt Rangsummen-Test.

8.6

f Nein! Information ausnützen. Macht = W. für Fehler 2. Art.
Die allzu grosszügige Lockerung der Voraussetzungen
bezahlt man mit einem Verlust an Macht.

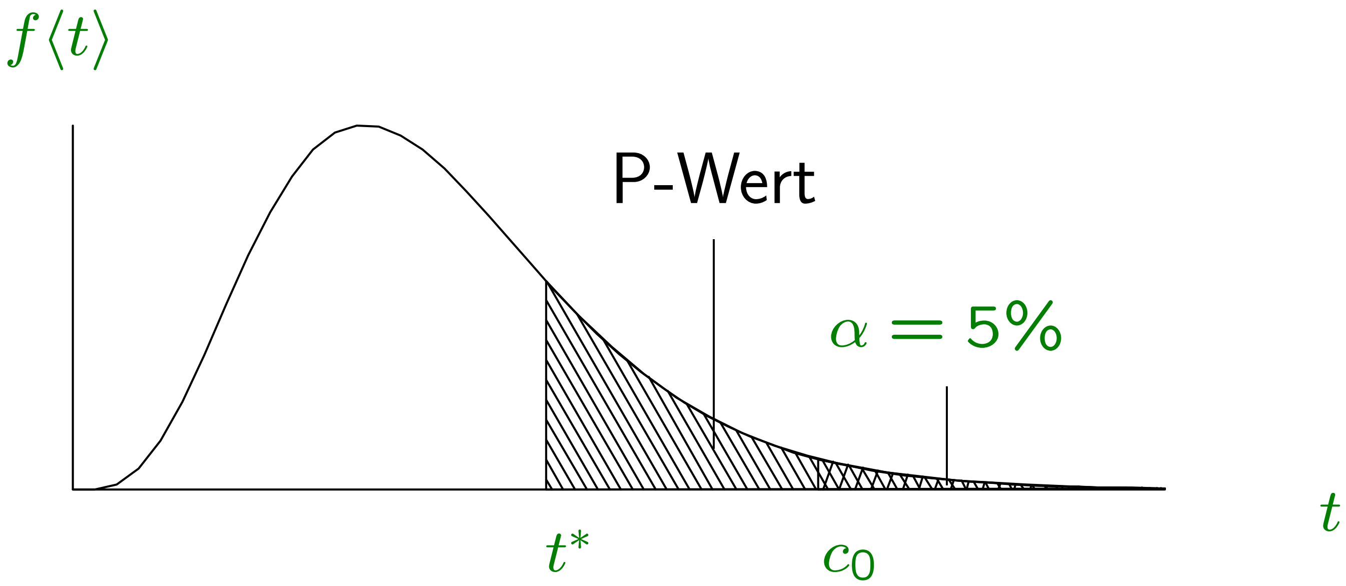
g Konkrete Empfehlung für das Testen eines Lageparameters:

- Rangsummen-Test von Wilcoxon anwenden, falls
  Verteilung symmetrisch ist wegen
  – theoretischen überlegungen (z.B. Differenzen),
  – empirischen Resultaten in grossen Datensätzen,
  – Vorwissen (frühere Studien mit ähnlichen Daten);

- sonst Vorzeichen-Test durchführen;

- t-Test (und z-Test) vermeiden.
  Gewinn an Macht für normalverteilte Beob. minim.
  Wenn nicht genau normalvert., hat der Rangsummen-Test meistens grössere Macht als der t-Test (Ausnahme: Sehr kleine Stichproben).

8.7 Der P-Wert



c Signifikanz: Test-Statistik rechts von c_0 $(\in K)$
$\Longleftrightarrow$ P-Wert kleiner als 5% – und umgekehrt.

P-Wert: verfeinertes „Mass der Signifikanz".
$P = 6\%$: knapp nicht signifikant, Nachdenken „erlaubt".
$P = 74\%$: kein Hinweis auf Abweichung von der Nullhypothese.

d Literatur, ältere Konvention:

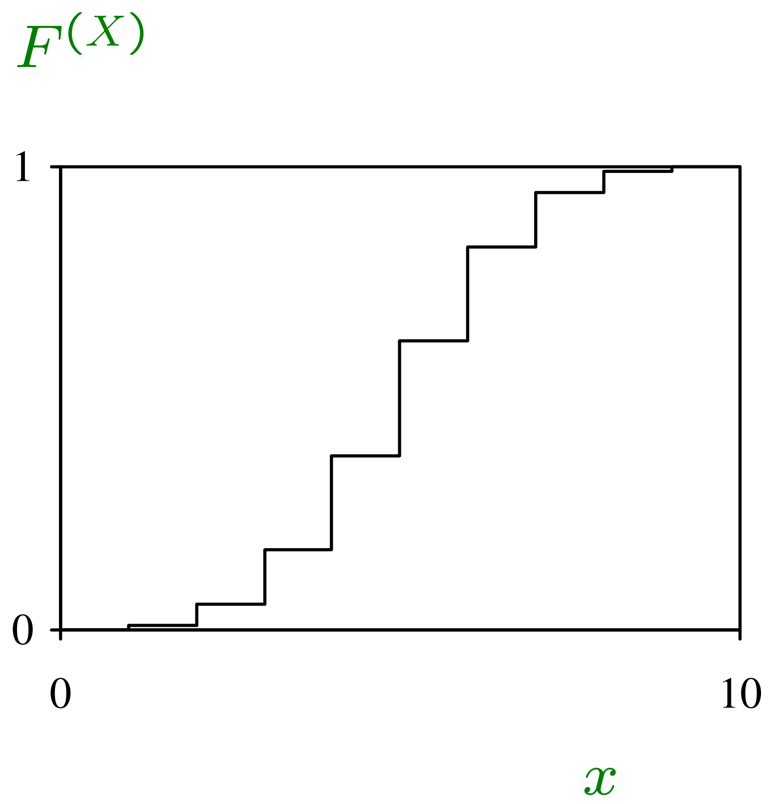
$P > 0.05$	nicht signifikant	(n.s.)
$0.05 \geq P > 0.01$	schwach signifikant	*
$0.01 \geq P > 0.001$	stark signifikant	**
$0.001 \geq P$	sehr stark signifikant	***

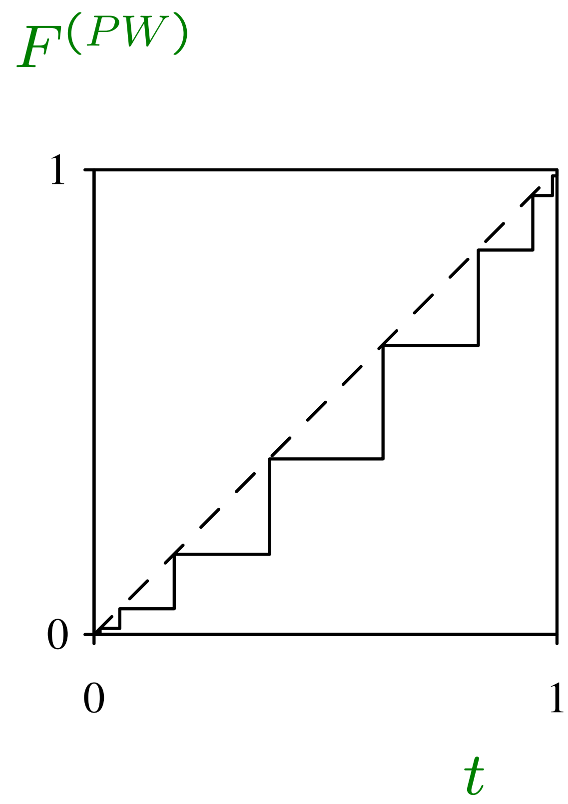
$z = 2.63^{**}$. $z = 1.46$ (n.s.).

e Hinweise:

- P-Wert = transformierte Test-Statistik mit uniformer Vert.
- Mit P-Wert kann man einfacher entscheiden als mit t.
  Computer muss mehr arbeiten.
- Achtung: Wegen Faulheit der Programmierer oft unexakte Tests!
  (Asymptotische Näherung auch bei kleinen Stichproben!)
- P-Wert aus verschiedenen Tests ?!
- P-Wert und Wahrscheinlichkeit!
  „Die Wahrscheinlichkeit für die Nullhypothese ist 10%."
  BITTE NICHT!
  „Die Irrtums-Wahrscheinlichkeit ist 3%."(auch nicht.)

- Bei diskret vert. T ist der P-Wert auch diskret verteilt.





8.8 Vergleich von zwei quantitativen Stichproben

a Einige Fragestellungen:

- Unterscheiden sich Puppengewichte von Fliegen (der gl. Art), die sich in gleich grossen Tieren verschiedener Wirtsarten entwickeln?
- Haben Raucher und Nichtraucher unterschiedl. Blutdruck?
- Unterscheiden sich Schnecken auf Mager- & Fettwiesen?
- Führt Düngersorte A zu höherem Ertrag als Dünger B?

Je eine Stichprobe aus zwei verschiedenen Grundgesamtheiten. Zielgrössse Y messen.

8.8

b Unterscheidung von gepaarten Stichproben:
Stichprobenumfänge n_1 und n_2 können
für unabh. Stichproben verschieden sein;
bei gepaarten Stichproben immer gleich gross.

d **Modell**: $Y_{1,1},\ Y_{1,2},\ \ldots,\ Y_{1,n_1} \qquad Y_{2,1},\ Y_{2,2},\ \ldots,\ Y_{2,n_2}$

$$Y_{1,i} \sim \mathcal{F}_1\ , \quad i = 1,2,\ldots,n_1;$$

$$Y_{2,i} \sim \mathcal{F}_2\ , \quad i = 1,2,\ldots,n_2;$$

$Y_{1,1}, Y_{1,2}, \ldots, Y_{1,n_1}, Y_{2,1}, \ldots, Y_{2,n_2}$ alle unabhängig.

$H_0:\quad \mathcal{F}_1 = \mathcal{F}_2$. Alle Y haben gleiche Verteilung.

$H_A:\quad \mathcal{F}_1$ und $\mathcal{F}_2$ unterscheiden sich
durch eine Verschiebung δ: $F_2\langle x\rangle = F_1\langle x-\delta\rangle$

Speziell: $Y_{1,i} \sim \mathcal{N}\langle \mu, \sigma^2\rangle \quad Y_{2,i} \sim \mathcal{N}\langle \mu+\delta, \sigma^2\rangle$.

δ interessierender Parameter, μ, σ Stör-Parameter.

e $\widehat{\delta} = \overline{Y}_{2,\cdot} - \overline{Y}_{1,\cdot} = (Y_{2,1}+\ldots+Y_{2,n_2})/n_2 - (Y_{1,1}+\ldots+Y_{1,n_1})/n_1$

f Zwei-Stichproben-z-Test.

Beispiel Mastochsen. 2 Futterarten vergleichen.

Zielgrösse: mittlere wöchentliche Gewichtszunahme in 1 Monat.

Erfahrung: Solche Gewichtszunahmen $\approx\sim \mathcal{N}\langle \ldots, (\sigma_0 = 1.8\text{kg})^2 \rangle$.

Einseitige Alternative.

g $H_0 : \quad Y_{k,i} \sim \mathcal{N}\langle \mu, \sigma_0^2 \rangle \quad (i.i.d.)$;

$H_A : \quad Y_{1,i} \sim \mathcal{N}\langle \mu_1, \sigma_0^2 \rangle; \quad Y_{2,i} \sim \mathcal{N}\langle \mu_1 + \delta, \sigma_0^2 \rangle$

einseitig: $\delta > 0$.

$U : \qquad U = \overline{Y}_{2,\cdot} - \overline{Y}_{1,\cdot}$

$\mathcal{F}_0\langle U \rangle : \quad \overline{Y}_{k,\cdot} \sim \mathcal{N}\langle \mu, \sigma_0^2 / n_k \rangle$, also

$$U = \overline{Y}_{2,\cdot} - \overline{Y}_{1,\cdot} \sim \mathcal{N}\left\langle 0, \sigma_0^2 \left(1/n_1 + 1/n_2\right) \right\rangle$$

$$Z = \frac{\overline{Y}_{2,\cdot} - \overline{Y}_{1,\cdot}}{\sqrt{\sigma_0^2 (1/n_1 + 1/n_2)}} \sim \mathcal{N}\langle 0, 1 \rangle$$

$K : \qquad$ einseitig: $K = \{Z \geq 1.64\}$.

h Mastochsen:

„extensiv"	2.7	2.7	1.1	3.0	1.9	3.0	3.8	3.8	0.3	1.9	1.9
„intensiv"	6.5	5.4	8.1	3.5	0.5	3.8	6.8	4.9	9.5	6.2	4.1

$\bar{y}_{2,\cdot} - \bar{y}_{1,\cdot} = 5.39\text{kg} - 2.37\text{kg} = 3.02\text{kg}$

$z = 3.02/\sqrt{1.8^2 \cdot 2/11} = 3.93 > 1.64.$

Nullhypothese (wuchtig) verworfen.

(Sinn eines solchen Tests? Schätzproblem!)

i **t-Test.**

$H_0: \quad Y_{k,i} \sim \mathcal{N}\langle\mu, \sigma^2\rangle \quad (i.i.d)$

$H_A: \quad Y_{1,i} \sim \mathcal{N}\langle\mu, \sigma^2\rangle,\ Y_{2,i} \sim \mathcal{N}\langle\mu+\delta, \sigma^2\rangle,$
$\delta \neq 0$ (oder $\delta > 0$ oder $\delta < 0$),
unabhängig.

Man ersetzt in der standardisierten Statistik für den z-Test die Varianz σ_0^2 durch eine Schätzung,

$$\widehat{\sigma}^2 = \frac{1}{n_1+n_2-2}\left(\sum_{i=1}^{n_1}(Y_{1,i}-\overline{Y}_{1,\cdot})^2 + \sum_{i=1}^{n_2}(Y_{2,i}-\overline{Y}_{2,\cdot})^2\right) .$$

t-Verteilung mit n_1+n_2-2 Freiheitsgraden.

j Test mit weniger Voraussetzungen!

Rangsummen-Test von Wilcoxon, Mann und Whitney, U-Test

H_0 : $Y_{k,i} \sim \mathcal{F}\,(i.i.d)$; $\mathcal{F}$ bel. Vert.

H_A : $Y_{1,i} \sim \mathcal{F}_1$, $Y_{2,i} \sim \mathcal{F}_2$, $F_2\langle x\rangle = F_1\langle x-\delta\rangle$, $\delta \neq 0$.

U: 1. Bestimme Rang $R_{k,i}$ bezügl. „vereinigten Stichproben"

2. $U^{(1)} = \sum_{i=1}^{n_1} R_{1,i}$ (oder $U^{(2)} = \sum_{i=1}^{n_2} R_{2,i}$)

$$(U^{(1)} + U^{(2)} = (n_1+n_2)(n_1+n_2+1)/2.)$$

$$T^{(k)} = U^{(k)} - n_k(n_k+1)/2$$

$$(T^{(1)} + T^{(2)} = n_1 n_2.)$$

Zweiseitige Fragestellung: $T = \min\langle T^{(1)}, T^{(2)}\rangle$.

$\mathcal{F}_0\langle U\rangle$: Verteilung von $T^{(k)}$ hängt nur von n_1, n_2 ab.

Grosse n_1, n_2: Asymptotische Näherung.

K : Tabelle.

n_1 \ n_2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	0	0	0	0	1	1	1	1	1	2	2	2	2
3	–	–	–	–	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	–	–	–	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
5	–	–	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	–	–	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	–	–	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	–	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	–	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	–	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	–	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	–	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	–	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	–	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	69	74	78	83
15	–	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	–	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	–	2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87	93	99	105
18	–	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	–	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	–	2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

k **Beispiel Mastochsen.**

	Daten·10																						
ext.	3		11	19	19	19	27	27	30	30		38	38										
int.		5									35			38	41	49	54	62	65	68	81	95	
	Ränge																						$\sum$
ext.	1		3	5	5	5	7.5	7.5	9.5	9.5		13	13										79
int.		2									11			13	15	16	17	18	19	20	21	22	174

$u^{(1)} = 79$ und $u^{(2)} = 174$. (Kontrolle: $u^{(1)} + u^{(2)} = 22 \cdot 23/2$.)

$t^{(1)} = 79 - 11 \cdot 12/2 = 13$, $t^{(2)} = 174 - 11 \cdot 12/2 = 108$,

$t = \min\langle t^{(1)}, t^{(2)} \rangle = 13$.

$K = \{T \leq 30\}$. Nullhypothese wird verworfen.

m Vergleich der Streuungen

n Grafischer Vergleich mit box-plots.

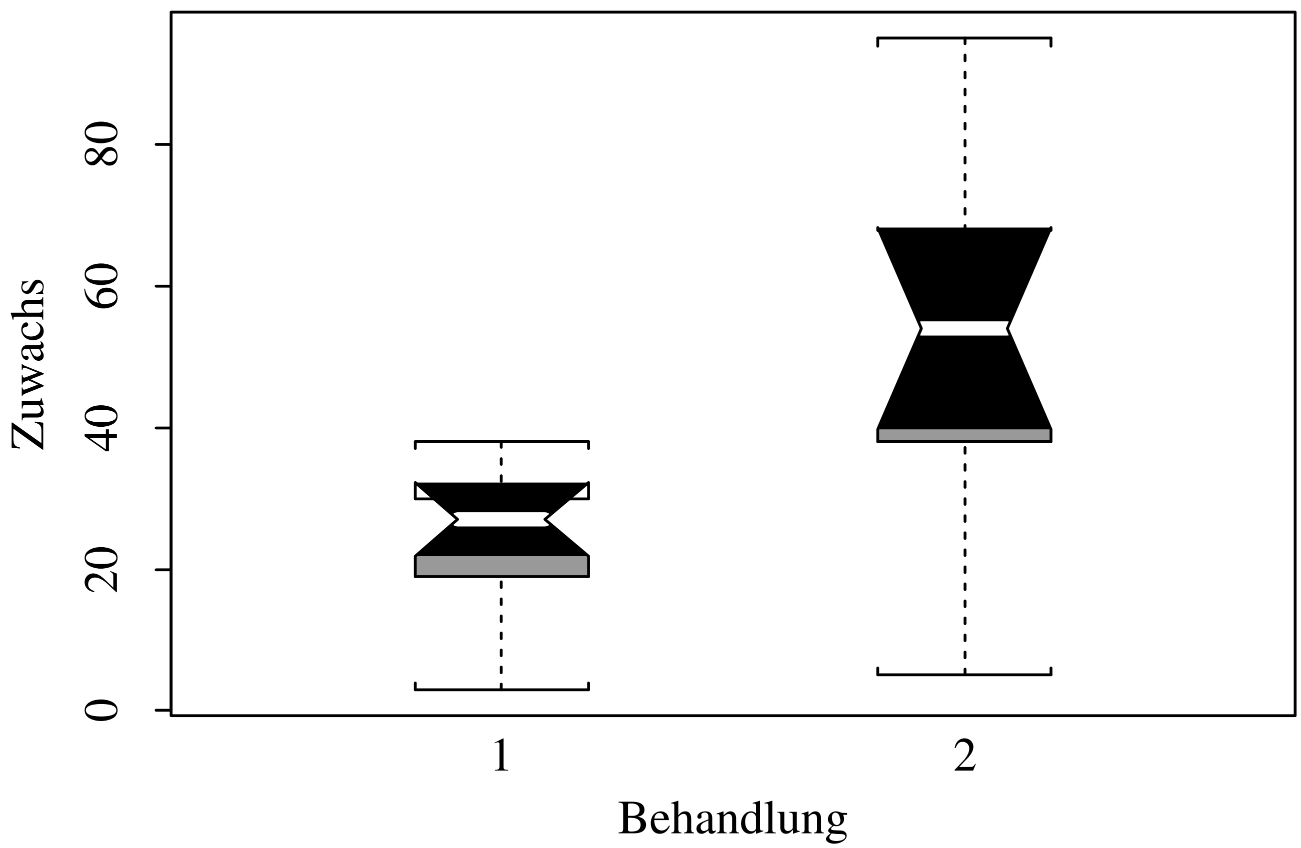
Unterschiede zwischen den Medianen von 2 Gruppen signifikant?

Kerben. Es soll gelten:

> Wenn sich die Kerben von zwei Kisten nicht überschneiden,
> dann ist der Unterschied zwischen den Gruppen signifikant.

$\longrightarrow$ „Notched box plots", gekerbte Kisten-Diagramme.

(Testregel nicht nach dem üblichen Schema!)


Zuwachs
0
20
40
60
80
1
2
Behandlung

8.9 Macht

a Bisher: H_0 vorausgesetzt.
Wahrscheinlichkeiten nur unter dieser Voraussetzung berechnet.
$P^{(0)}\langle \text{Verwerfen von } H_0 \rangle = \alpha$

Jetzt: Alternative voraussetzen.
Fehler: H_0 nicht verwerfen. = Fehler 2. Art
Wahrscheinlichkeit berechnen!

Berechnung von Wahrscheinlichkeiten setzt genau spezifizierte Hypothese voraus!

$$P^{(A)}\langle T \notin K \rangle = 1 - P^{(A)}\langle T \in K \rangle = \beta^{(A)}.$$

$1 - \beta^{(A)}$ = W. der „richtigen Entscheidung = Macht
soll möglichst hoch sein.

(Aufgepasst: In manchen Büchern wird
die Macht mit β bezeichnet.)

b K ist der Verwerfungsbereich, hergeleitet aus der
Verteilung der Test-Statistik unter der Nullhypothese
Der Test (die Entscheidungsregel) bleibt.

8.9

c Beispiel Mastochsen: 2 Fütterungsarten vergleichen.
„Wahrer" Unterschied $\delta = 1.0$ kg/Woche (Vermutung).
Wie gross ist die Wahrscheinlichkeit, dass dieser Effekt, falls er richtig ist, sich durch den Versuch statistisch nachweisen lässt?
... d.h., dass die Testregel zur Verwerfung von H_0 führt?

Entscheidungsregel: H_0 verwerfen, falls $Z > 1.64$.
$\longrightarrow$ $U = Z \cdot \sqrt{1.8^2 \cdot 2/11} = Z \cdot 0.77$, also
$Z > 1.64 \quad \longrightarrow \; U > 1.64 \cdot 0.77 = 1.26$.

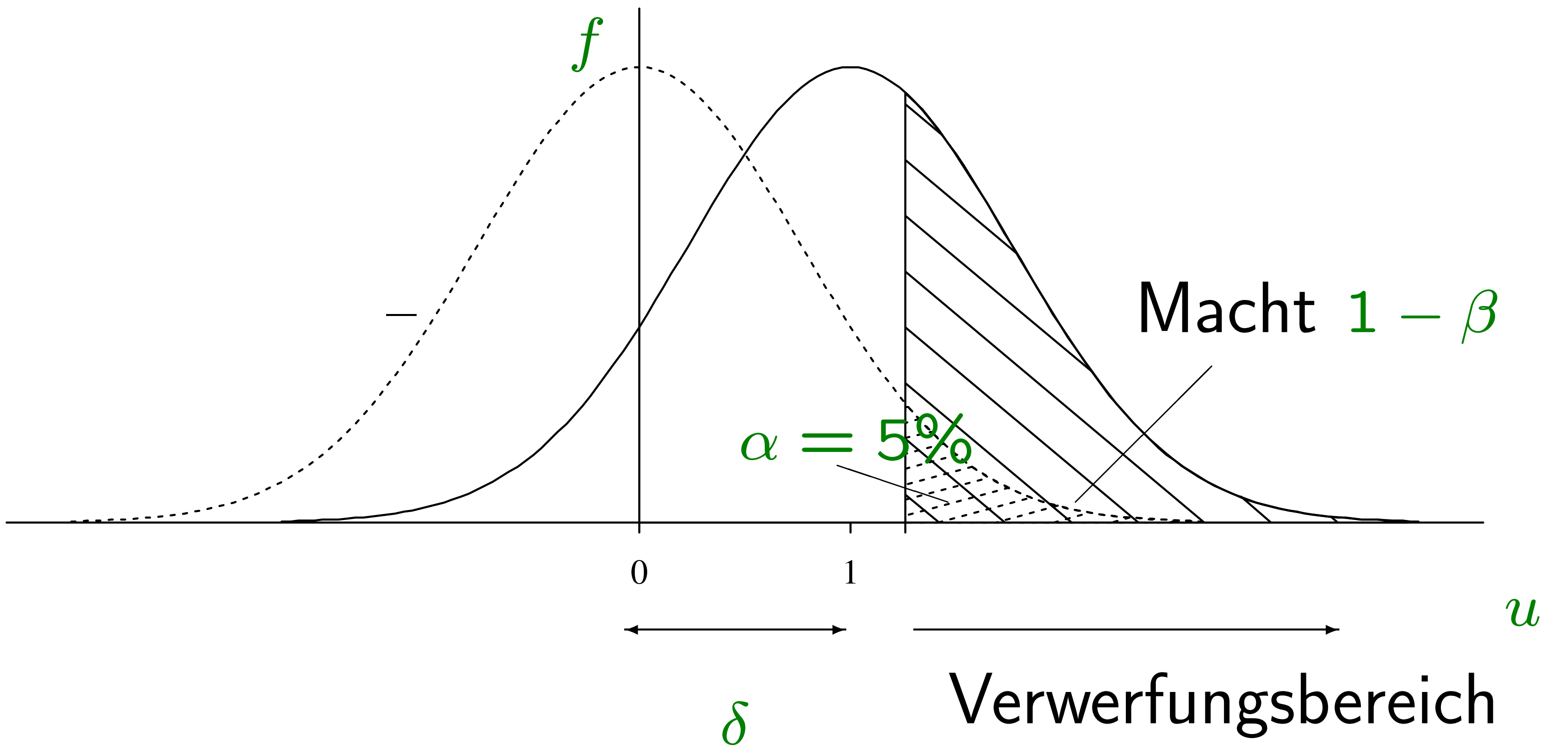
Vt. von T unter H_A mit $\delta = 1.0$kg ist

$$U = \overline{Y}_{2\cdot} - \overline{Y}_{1\cdot} \sim \mathcal{N}\left\langle \delta, \sigma_0^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right\rangle = \mathcal{N}\langle 1.0, 0.77^2 \rangle$$

Macht $1 - \beta^{(A)} = P^{(A)}\langle U > 1.26 \rangle$
$= 1 - \Phi \left\langle \frac{1.26-1}{0.77} \right\rangle = 1 - \Phi\langle 0.34 \rangle = 0.37$
$\longrightarrow$ Wahrsch. eines „erfolgreichen Ausgangs" des Versuchs (statistischer „Beweis" eines Effekts gelungen) $= 37\%$. (!)

Verteilung der Test-Statistik U ...

... unter H_0 unter H_A



d Oft keine Vermutung betreffend Effekt δ.

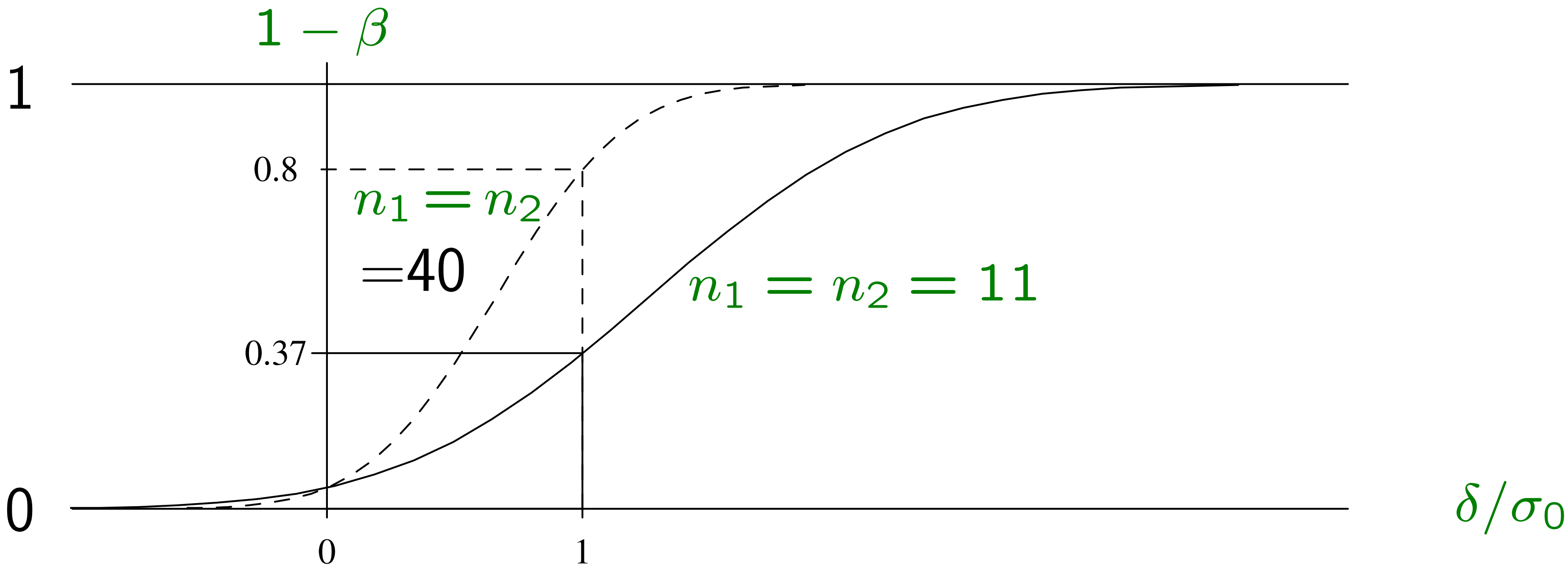
$\longrightarrow$ Macht als Funktion von δ.

$n_1 = n_2 = n$. $\quad H_0$ verworfen, wenn $U > 1.64 \cdot \sigma_0 \sqrt{2/n}$

Unter H_A gilt $U \sim \mathcal{N} \left\langle \delta,\, \sigma_0^2\, 2/n \right\rangle$

$$\begin{aligned} 1 - \beta^{(A)} &= P^{(A)} \left\langle U > 1.64 \cdot \sigma_0 \sqrt{2/n} \right\rangle \\ &= 1 - \Phi \left\langle \frac{1.64 \cdot \sigma_0 \sqrt{2/n} - \delta}{\sigma_0 \sqrt{2/n}} \right\rangle \\ &= 1 - \Phi \left\langle 1.64 - \frac{\delta}{\sigma_0} \sqrt{n/2} \right\rangle . \end{aligned}$$

Die Macht nimmt zu mit „standardisiertem Effekt" δ/σ_0 und Stichprobenumfang n.



Macht $1-\beta^{(A)}$ als Funktion von δ heisst auch Gütefunktion.
$\beta^{(A)}$, als Funktion von δ heisst auch Operations-Charakteristik.

e Wie gross muss n sein, damit bei $\delta = 1.0$ und $\sigma_0 = 1.8$ $1 - \beta^{(A)} = 0.80$ wird?

$$\begin{aligned} 1 - \Phi\left\langle 1.64 - \frac{1.0}{1.8}\sqrt{n/2}\right\rangle &= 0.8 \quad \Longrightarrow \\ 1.64 - 0.56\sqrt{n/2} &= -0.84 \\ n = 2 \cdot \left(\frac{2.48}{0.56}\right)^2 &= 39.2 \end{aligned}$$

Es braucht (etwa) 40 Tiere pro Gruppe.

$\longrightarrow$ Versuchsplanung, Berechnung des Stichprobenumfangs.

f* Man muss σ_0 ungefähr kennen und δ festlegen.
Vorversuch?
Nichtparametrischer Test?

g* Macht (-funktion) als Qualitätskriterium zur Wahl eines Tests.
Optimieren! $\longrightarrow$ Mathematische Statistik.
Optimale Schätzungen $\longrightarrow$ Testgrössen für optimale Tests.

8.11 Sinn und Unsinn statistischer Tests

a Fachzeitschriften: Man darf nur über Effekte schreiben,
die statistisch auf dem 5%-Niveau signifikant sind.
$\longrightarrow$ Test dient als Filter gegen wilde Spekulationen.

Perversion: Untersuche Fragen, für die ein signifikantes Ergebnis erwartet werden kann – auch uninteressante ...

Statistiker unterstützen diesen Unsinn nicht.
Relevanz wichtiger als Signifikanz!

b **Unterschied zwischen (statistischer) „Signifikanz" und (praktischer) „Relevanz"**

Zu kleine und zu grosse Stichproben!

Mit grossen Stichproben kann man Bagatelleffekte, die eigentlich niemanden interessieren, stolz als „statistisch signifikant" nachweisen und publizieren.

c* Konsequenz? Statistische Tests sind unsinnig!

$\longrightarrow$ Man müsste $H_0 : \theta \leq$ Schwellenwert γ prüfen!

Regel für Entscheidung zwischen $H_0 : \theta \leq \gamma$ und $H_1 : \theta > \gamma$.

8.11

d Wieso Tests ausführlich behandeln?

- Filter gegen wilde Spekulationen
- Grundlage für das Verständnis der Beziehung zwischen Wahrscheinlichkeits-Modellen und empirischen Daten.
- Grundlage für Vertrauensintervall.

9. Vertrauensintervalle

9.3 Vertrauensintervalle für Lageparameter

Grundfrage?

a Jedem Test für μ einer Stichprobe entspricht
eine Regel zur Konstruktion von Vertrauensintervallen.

z-Test. $X_i \sim \mathcal{N}\langle \mu, \sigma_0^2 \rangle$, σ_0^2 bekannt.
Vertrauensintervall für μ?

$H_0: \ \mu = \mu_0$. $\longrightarrow$ Annahmebereich

$$\{|\overline{X} - \mu_0|/(\sigma_0/\sqrt{n}) \le 1.96\} =$$
$$\{\overline{X} - \mu_0 \ge -1.96\,\sigma_0/\sqrt{n} \text{ und } \overline{X} - \mu_0 \le 1.96\,\sigma_0/\sqrt{n}\}$$
$$V_0 = \overline{X} - 1.96\,\sigma_0/\sqrt{n} \quad \text{respektive} \quad V_1 = \overline{X} + 1.96\,\sigma_0/\sqrt{n}$$
$$\overline{X} \pm 1.96\,\sigma_0/\sqrt{n}$$

b **Beispiel Reifen**

$v_0 = 2.29 - 1.96 \cdot 3/\sqrt{10} = 2.29 - 1.86 = 0.43$

$v_1 = 2.29 + 1.86 = 4.15$.

Vertrauensintervall für die Differenz des Bremsweges:

$(0.43, 4.15) = 2.29 \pm 1.86$.

Enthält 0 nicht. Nullhypothese $\mu_0 = 0$ verworfen.

9.3

c **t-Test**

$$V_0 = \overline{X} - q_{0.975}^{(t_{n-1})} \cdot S/\sqrt{n}\,, \qquad V_1 = \overline{X} + q_{0.975}^{(t_{n-1})} \cdot S/\sqrt{n}\,.$$

S^2: empirische Varianz,

$q_{0.975}^{(t_{n-1})}$: 0.975-Quantil der t-Vert. mit $n-1$ Freiheitsgraden (kritische Grenze für den zweiseitigen t-Test).

d Vertrauensintervall, das dem Wilcoxon-Test entspricht:
Besser, aber kaum in Programmen anzutreffen.

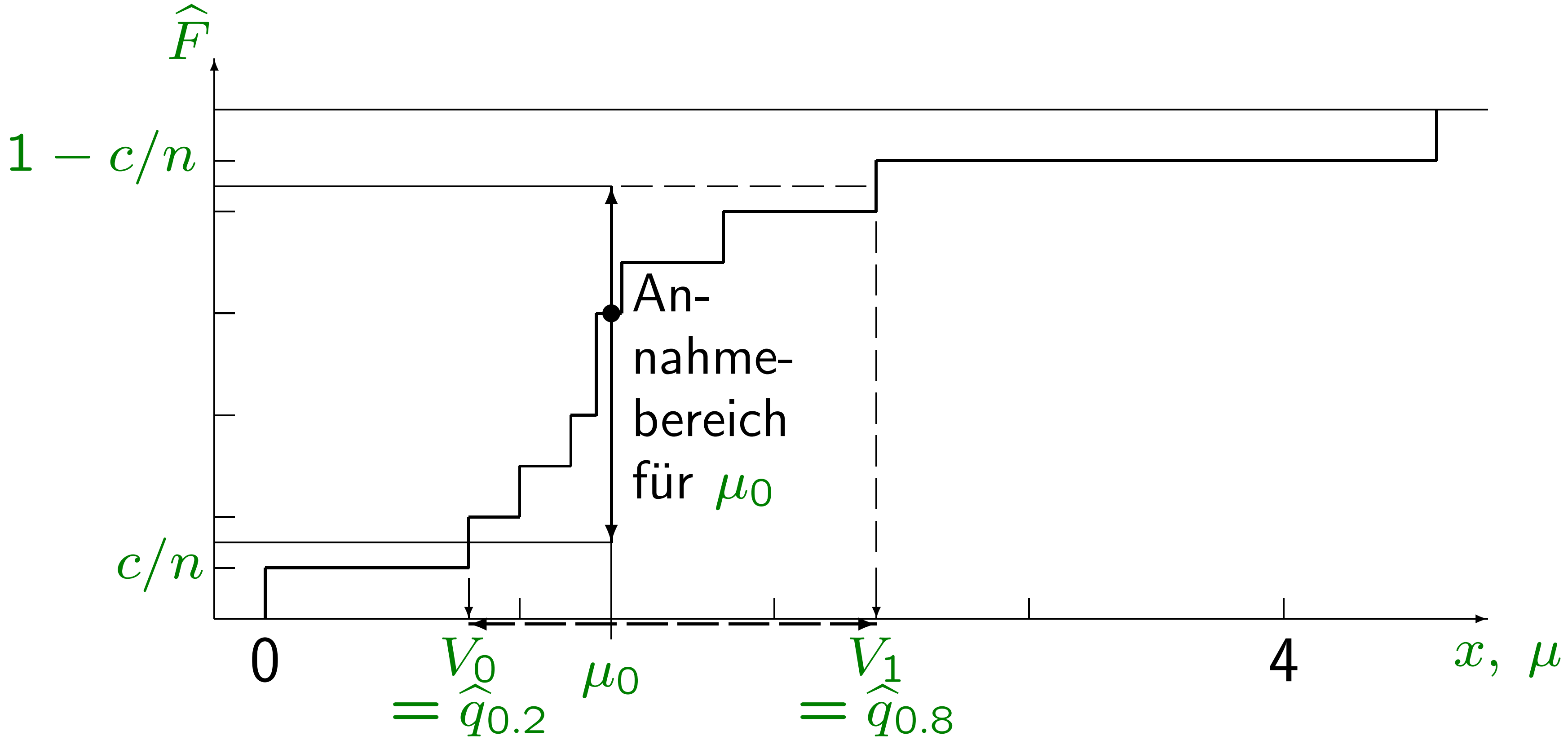
e **Vorzeichentest** Annahmebereich

$$\begin{aligned} & \{c < \text{Anzahl}\{\, i \mid X_i > \mu_0\} < n - c\} \\ = \; & \{c/n < \text{Anzahl}\{\, i \mid X_i \le \mu_0\}/n < 1 - c/n\} \\ = \; & \{c/n < \widehat{F}\langle\mu_0\rangle < 1 - c/n\} \,. \end{aligned}$$

$\widehat{F}$: empirische Verteilungsfunktion.

Untere Grenze: $c/n = \widehat{F}\langle\mu_0\rangle \Longrightarrow \mu_0 = (c/n)$-Quantil $\widehat{q}_{c/n}$.

$(\widehat{q}_\gamma, \widehat{q}_{1-\gamma})$ mit $\gamma = c/n$.


$\widehat{F}$
$1 - c/n$
c/n
An-
nahme-
bereich
für μ_0
0
V_0
$= \widehat{q}_{0.2}$
μ_0
V_1
$= \widehat{q}_{0.8}$
4
$x,\ \mu$

9.3

f Lage-Differenz δ zwischen zwei unabhängigen Stichproben:

Vertrauensintervalle aus den in 8.8 besprochenen Tests.

g Welches Vertrauensintervall wählen?

9.5 Vertrauens- und andere Intervalle

a Vertrauensintervall: zufälliges Intervall,
abhängig von den Daten.
Enthält, wenn man von einem Modell ausgeht,
den wahren Parameterwert mit Wahrscheinlichkeit 95%.

b Annahmebereich: festes Intervall,
abhängig vom zu testenden Parameterwert.
Enthält den Wert der Test-Statistik mit W. 95%,
falls die Nullhypothese richtig ist.

9.5

c Streubereich: Intervall, in dem die meisten Daten
enthalten sind.

Kiste im Box-Plot enthält 50% der Daten.

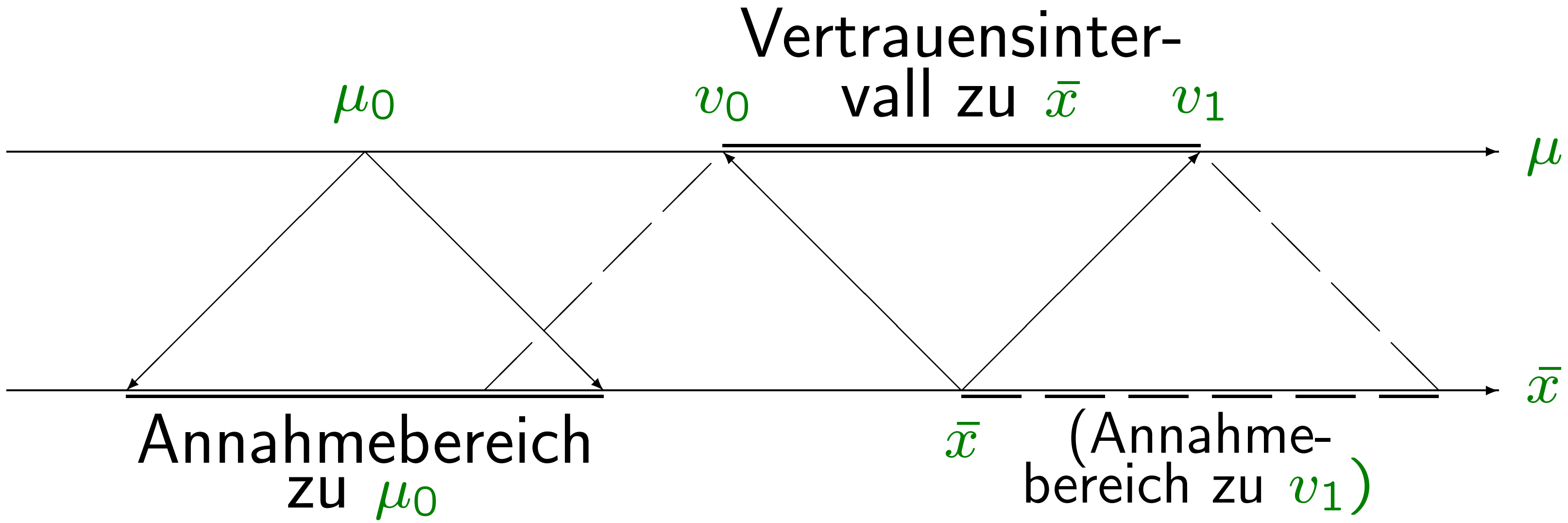
$\bar{X} \pm 2\widehat{\sigma}$: enthält ungefähr 95% der Daten.

e Kerbenbereich in gekerbten Kisten-Diagrammen:
ermöglicht einen approximativen Test für den Unterschied zwischen zwei unabhängigen Stichproben.

Genäherte Vertrauensintervalle für die Mediane,
für Vertrauenskoeffizienten $\approx 90\%$.

g Der Annahmebereich und das Vertrauensintervall werden sehr oft verwechselt.

$$\begin{aligned} \text{Annahmebereich} &= \{\mu_0 \pm c \cdot \text{se}\} \\ \text{Vertrauensintervall} &= \{\bar{X} \pm c \cdot \text{se}\} \end{aligned}$$

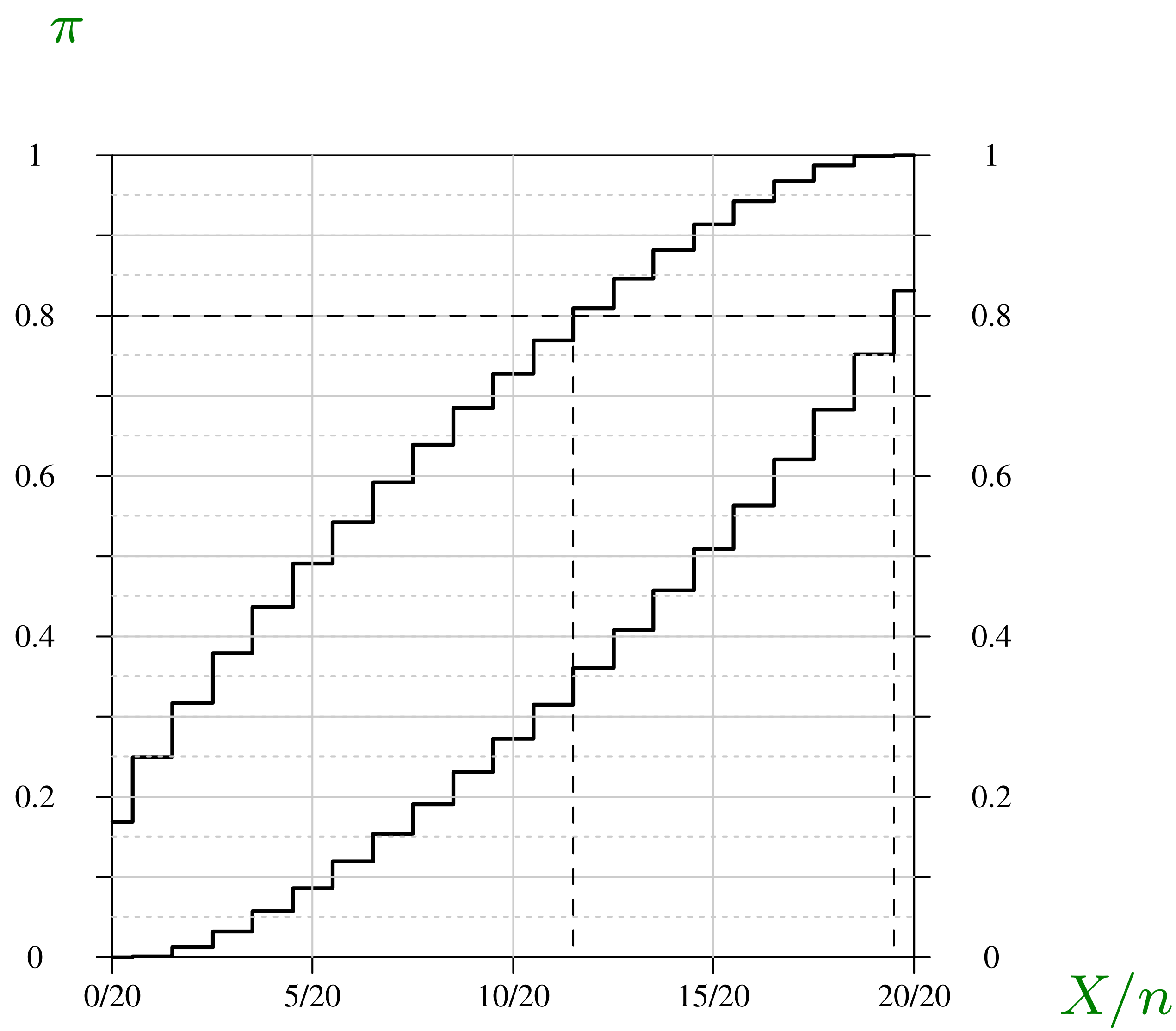


h Dualität.

- Der Annahmebereich umfasst die Schätzwerte $\widehat{\theta}$, die mit einem bestimmten $\theta = \theta_0$ verträglich sind.

- Das Vertrauensintervall umfasst die Parameterwerte θ, die mit einem bestimmten Schätzwert $\widehat{\theta}$ verträglich sind.

vgl. Diagramm für die Binomialverteilung.



Annahmebereich nicht symmetrisch um θ_0
Vertrauensintervall nicht symmetrisch um $\widehat{\theta}$.

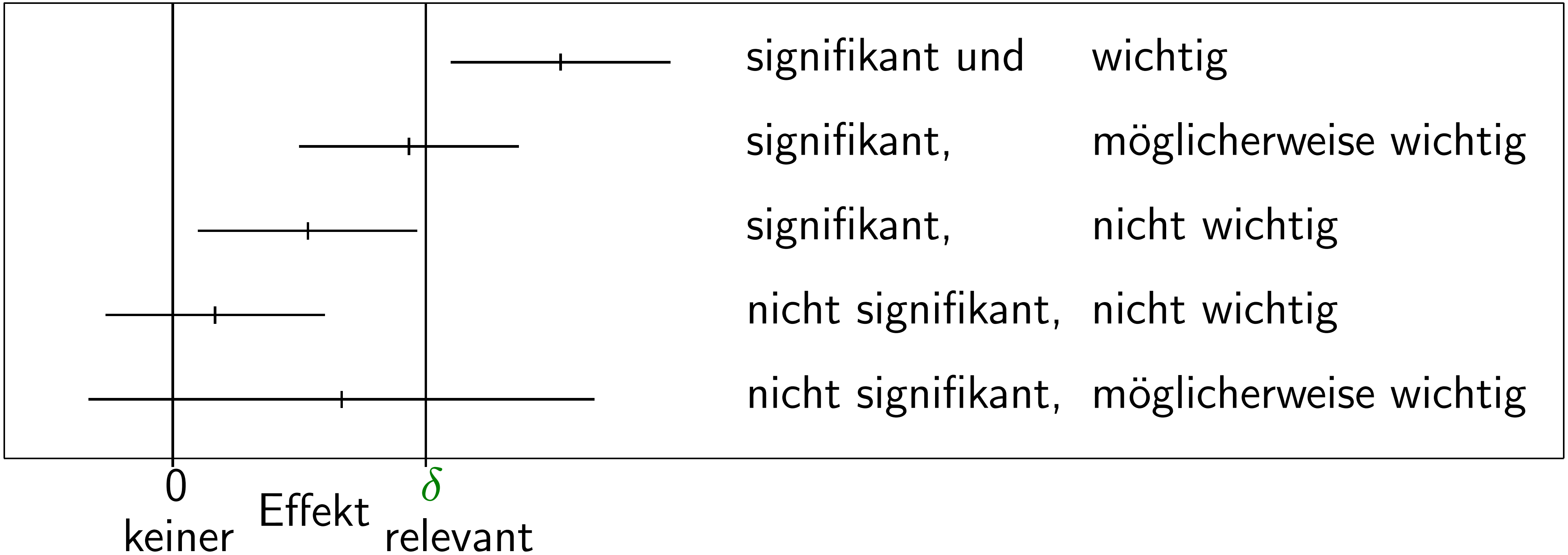
9.6 Schätzungen, Tests und Vertrauensintervalle

a Vertrauensintervalle sind nützlicher als Punktschätzungen.
Vertrauensintervall enthält Punktschätzung.
Punktschätzung allein sagt nichts über ihre Genauigkeit,
und eine Zahl ohne Genauigkeitsangabe sagt nichts aus!

c Vertrauensintervalle sind nützlicher als Tests

e Vertrauensintervall oder P-Wert?
Beide zeigen, „wie signifikant" ein Effekt ist.

f Achtung: Relevanz vs. Signifikanz!



b Punktschätzungen sind Hilfsmittel:

- als Test-Statistik zur Konstruktion von Tests
  und dadurch von Vertrauensintervallen;
- Daten beschreiben, Dichte an Histogramm anpassen.

d Tests

- für Hypothesen, die nicht einen Parameter festlegen.
  Anpassungstests, Test auf Unabhängigkeit
- als Nachweis eines Effekts,
  ohne Grösse des Effekts zu quantifizieren