Satz 1.77

Sei X reelle ZV, $f : [0,\infty) \to [0,\infty)$ monoton wachsend.

1. *Für $a > 0$ mit $f(a) > 0$ gilt*

$$P\big(|X| \geq a\big) \leq \frac{1}{f(a)} \mathbb{E}\big[f(|X|)\big] \qquad \text{(Markov[1]-Ungleichung).} \tag{1}$$

2. *Für $X \in \mathcal{L}^2$ gilt*

$$P\big(|X - \mathbb{E}[X]| \geq a\big) \leq \frac{\mathrm{Var}[X]}{a^2} \qquad \text{(Chebyshev[2]-Ungleichung).} \tag{2}$$

z.B. wähle $a = 10 \cdot \sqrt{\mathrm{Var}[X]}$, dann ist

r.S. $= \frac{1}{100}$

[2]Andrei Andrejewich Markov, 1856–1922.
[2]Pafnuty Lvovich Chebyshev, 1821–1894.

1. $P\big(|X| \geq a\big) \leq \dfrac{1}{f(a)}\mathbb{E}\big[f(|X|)\big]$

2. $P\big(|X - \mathbb{E}[X]| \geq a\big) \leq \dfrac{\mathrm{Var}[X]}{a^2}$

1. $P\big(|X| \geq a\big) \leq \dfrac{1}{f(a)}\mathbb{E}\big[f(|X|)\big]$

2. $P\big(|X - \mathbb{E}[X]| \geq a\big) \leq \dfrac{\operatorname{Var}[X]}{a^2}$

Beweis.

1. Sei $Y := f(a) I_{\{|X| \geq a\}}$, so ist $Y \leq f(|X|)$ und

$$\mathbb{E}[Y] = f(a) P\big(|X| \geq a\big) \leq \mathbb{E}\big[f(|X|)\big]$$

$(+0 \cdot P(|X| < a))$

nach Satz 1.70, 2.

1. $P\big(|X| \geq a\big) \leq \dfrac{1}{f(a)}\mathbb{E}\big[f(|X|)\big]$

2. $P\big(|X - \mathbb{E}[X]| \geq a\big) \leq \dfrac{\mathrm{Var}[X]}{a^2}$

Beweis.

1. Sei $Y := f(a) I_{\{|X| \geq a\}}$, so ist $Y \leq f(|X|)$ und

$$\mathbb{E}[Y] = f(a) P\big(|X| \geq a\big) \leq \mathbb{E}\big[f(|X|)\big]$$

nach Satz 1.70, 2.

2. Wende *1.* an auf $\widetilde{X} := X - \mathbb{E}[X]$ und $f(a) = a^2$. $\square$

1. $P\big(|X| \geq a\big) \leq \dfrac{1}{f(a)}\mathbb{E}\big[f(|X|)\big]$

2. $P\big(|X - \mathbb{E}[X]| \geq a\big) \leq \dfrac{\mathrm{Var}[X]}{a^2}$

Beweis.

1. Sei $Y := f(a) I_{\{|X| \geq a\}}$, so ist $Y \leq f(|X|)$ und

$$\mathbb{E}[Y] = f(a)P\big(|X| \geq a\big) \leq \mathbb{E}\big[f(|X|)\big]$$

nach Satz 1.70, 2.

2. Wende *1.* an auf $\widetilde{X} := X - \mathbb{E}[X]$ und $f(a) = a^2$. □

Insbesondere (wähle $a = b\sigma_X$ in (2)):
Die W'keit $P\big(|X - \mathbb{E}[X]| \geq b\,\sigma_X\big)$, dass X von $\mathbb{E}[X]$ um mehr als das b-fache von σ_X abweicht, ist höchstens $1/b^2$.

Beobachtung 1.78

1. *Wegen* $|XY| \leq X^2 + Y^2$ *ist die Kovarianz wohldefiniert. Es gilt (offensichtlich)*

$$\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X].$$

Beobachtung 1.78

1. *Wegen* $|XY| \leq X^2 + Y^2$ *ist die Kovarianz wohldefiniert. Es gilt (offensichtlich)*

$$\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X].$$

2. *Es ist*

$$\begin{aligned}&\mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big]\\ &= \mathbb{E}\big[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]\big]\\ &= \mathbb{E}[YX] - \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[Y]\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]\\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

(und analog für $\mathrm{Var}[X] = \mathrm{Cov}[X, X]$*).*

Beobachtung 1.78 (Fortsetzung)

3. $\mathrm{Var}[X] = 0 \iff P\big(X = \mathbb{E}[X]\big) = 1$
   („$\Leftarrow$“ ist klar, für „$\Rightarrow$“ wende Satz 1.70, 3. an auf die ZV $(X - \mathbb{E}[X])^2$*)*

Beobachtung 1.78 (Fortsetzung)

3. $\mathrm{Var}[X] = 0 \iff P\big(X = \mathbb{E}[X]\big) = 1$
   („$\Leftarrow$“ ist klar, für „$\Rightarrow$“ wende Satz 1.70, 3. an auf die ZV $(X - \mathbb{E}[X])^2$*)*

4. $\mathrm{Var}[X]$ *ist eine Eigenschaft der Verteilung von* X*,*
   $\mathrm{Cov}[X, Y]$ *ist eine Eigenschaft der gemeinsamen Verteilung von* X *und* Y*.*

Beispiel 1.79

1. $X \sim \mathrm{Ber}_p$, $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - p^2 = p(1-p)$.

Beispiel 1.79

1. $X \sim \mathrm{Ber}_p$, $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - p^2 = p(1-p)$.

2. $X \sim \mathrm{Poi}_\alpha$,

$$\mathbb{E}[\overbrace{X(X-1)}^{=X^2-X}] = \sum_{k=0}^{\infty} k(k-1)e^{-\alpha}\frac{\alpha^k}{k!} = \alpha^2 \sum_{k=2}^{\infty} e^{-\alpha}\frac{\alpha^{k-2}}{(k-2)!} = \alpha^2,$$

Beispiel 1.79

1. $X \sim \mathrm{Ber}_p$, $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - p^2 = p(1-p)$.

2. $X \sim \mathrm{Poi}_\alpha$,

$$\mathbb{E}[X(X-1)] = \sum_{k=0}^{\infty} k(k-1)e^{-\alpha}\frac{\alpha^k}{k!} = \alpha^2 \sum_{k=2}^{\infty} e^{-\alpha}\frac{\alpha^{k-2}}{(k-2)!} = \alpha^2,$$

also

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}\left[X^2\right] - \left(\mathbb{E}[X]\right)^2 = \mathbb{E}\left[X(X-1)\right] + \mathbb{E}[X] - \left(\mathbb{E}[X]\right)^2 \\ &= \alpha^2 + \alpha - \alpha^2 = \alpha\end{aligned}$$

2b.: W ... Augenzahl beim fairen Würfelwurf

$(\mathbb{E}[W] = 3{,}5)$. $\mathrm{Var}[W] = \mathbb{E}[W^2] - (\mathbb{E}[W])^2$

$= \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2 = \frac{35}{12} \approx 2{,}91\overline{6}$

Beispiel 1.79 (Fortsetzung)

3. $X \sim \mathrm{Bin}_{n,p}$,

$$\begin{aligned}\mathbb{E}[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k} \\ &= n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{(n-2)-(k-2)} \\ &= n(n-1)p^2\end{aligned}$$

also

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}\big[X(X-1)\big] + \mathbb{E}[X] - \big(\mathbb{E}[X]\big)^2 \\ &= n(n-1)p^2 + np - (np)^2 = -np^2 + np \\ &= np(1-p)\end{aligned}$$

Beispiel 1.79 (Fortsetzung)

4. $X \sim \mathrm{Geom}_p$, $p \in [0,1]$ (wir hatten gesehen, dass $\mathbb{E}[X] = (1-p)/p$, siehe Bsp. 1.69, 2)

$P(X=n) = (1-p)^n \cdot p \;, n = 0,1,2,\ldots$

Es ist

$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$$\mathbb{E}[X(X-1)] = \sum_{n=0}^{\infty} n(n-1)p(1-p)^n$$

$$= p(1-p)^2 \sum_{n=2}^{\infty} n(n-1)p(1-p)^{n-2} = 2\frac{(1-p)^2}{p^2}$$

(verwende, dass $f(t) := \sum_{n=0}^{\infty} t^n = \frac{1}{1-t}$ (für $|t| < 1$) erfüllt $\frac{d^2}{dt^2} f(t) = \frac{2}{(1-t)^3} = \sum_{n=2}^{\infty} n(n-1)t^{n-2}$), somit

$$\mathrm{Var}[X] = \mathbb{E}[X(X-1)] + \mathbb{E}[X]\big(1 - \mathbb{E}[X]\big)$$

$$= 2\frac{(1-p)^2}{p^2} - \frac{1-p}{p} \cdot \frac{2p-1}{p} = \frac{1-p}{p^2}.$$

Beispiel 1.79 (Fortsetzung)

5. $X \sim \mathcal{N}_{\mu,\sigma^2}$ hat $\mathrm{Var}[X] = \sigma^2$ (wir hatten in Bsp. 1.74 bereits gesehen, dass $\mathbb{E}[X] = \mu$):

$$\mathrm{Var}[X] = \mathbb{E}\big[(X-\mu)^2\big] = \int_{\mathbb{R}} (x-\mu)^2 \overbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)}^{\text{Dichte von } X} dx$$

$\frac{x-\mu}{\sigma} = z$

$dz = \frac{1}{\sigma} dx$

$$= \int_{\mathbb{R}} \sigma^2 z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{z^2 e^{-z^2/2}}_{=z\left(-\frac{d}{dz} e^{-z^2/2}\right)} dz$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(\Big[z\Big(-\frac{d}{dz} e^{-z^2/2}\Big)\Big]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -e^{-z^2/2}\, dz \right)$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \big(0 + \sqrt{2\pi}\big) = \sigma^2$$

(Wir haben Bericht 1.75, 2. verwendet, dann im Integral $z = (x - \mu/\sigma)$ substituiert und partiell integriert.)

Satz 1.80 (Rechenregeln für Varianz und Kovarianz)

Seien $X, Y, X_1, X_2, \dots, X_n \in \mathcal{L}^2$, $a, b, c, d \in \mathbb{R}$.

1. $aX + b, cY + d \in \mathcal{L}^2$ und

$$\mathrm{Cov}[aX + b, cY + d] = ac\mathrm{Cov}[X, Y],$$

insbesondere $\mathrm{Var}[aX + b] = a^2\mathrm{Var}[X]$ (die Kovarianz ist eine Bilinearform, die Varianz ein quadratisches Funktional).

Satz 1.80 (Rechenregeln für Varianz und Kovarianz)

$= \sum_{i,j=1}^{n} \mathrm{Cov}[X_i, X_j]$

Seien $X, Y, X_1, X_2, \ldots, X_n \in \mathcal{L}^2$, $a, b, c, d \in \mathbb{R}$.

1. $aX + b, cY + d \in \mathcal{L}^2$ *und*

$$\mathrm{Cov}[aX + b, cY + d] = ac\mathrm{Cov}[X, Y],$$

insbesondere $\mathrm{Var}[aX + b] = a^2\mathrm{Var}[X]$ *(die Kovarianz ist eine Bilinearform, die Varianz ein quadratisches Funktional).*

2. $$\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{\substack{1 \le i,j \le n \\ i \ne j}} \mathrm{Cov}[X_i, X_j],$$

hat $\mathbb{E}[\ldots] = 0$

insbesondere gilt für paarweise unkorrelierte $X_1, \ldots, X_n$ *also*

$$\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mathrm{Var}[X_i].$$

$$\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] = \mathrm{Var}\Big[\sum_{i=1}^{n} (X_i - \mathbb{E}[X_i])\Big]$$

$$= \mathbb{E}\Big[\Big(\sum_{i=1}^{n} (X_i - \mathbb{E}[X_i])\Big)^2\Big] - 0^2 = \mathbb{E}\Big[\sum_{i,j=1}^{n} (X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\Big]$$

Satz 1.80 (Rechenregeln für Varianz und Kovarianz, 2)

3. Sind X und Y unabhängig, so gilt $\mathrm{Cov}[X, Y] = 0$.

$$\underbrace{\mathbb{E}[X \cdot Y]}_{= \mathbb{E}[X] \cdot \mathbb{E}[Y]} - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

für X, Y u.a.

[3] nach Augustin-Louis Cauchy (1789–1857) und Hermann Amandus Schwarz (1843–1921)

Satz 1.80 (Rechenregeln für Varianz und Kovarianz, 2)

3. Sind X und Y unabhängig, so gilt $\mathrm{Cov}[X, Y] = 0$.

4. Es gilt

$$\left|\mathrm{Cov}[X, Y]\right| \leq \sqrt{\mathrm{Var}[X]}\,\sqrt{\mathrm{Var}[Y]}$$

(die Cauchy-Schwarz-Ungleichung[3])

[3] nach Augustin-Louis Cauchy (1789–1857)
und Hermann Amandus Schwarz (1843–1921)

Beweis.

1. Es ist

$$\mathrm{Cov}[aX + b, cY + d] = \mathrm{Cov}[aX, cY]$$

(denn $\mathbb{E}[aX + b] = \mathbb{E}[aX] + b$ und $\mathbb{E}[cY + d] = \mathbb{E}[cY] + d$)

$$\begin{aligned} &= \mathbb{E}[aX\,cY] - \mathbb{E}[aX]\,\mathbb{E}[cY] \\ &= ac\big(\mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]\big) \\ &= ac\mathrm{Cov}[X, Y]. \end{aligned}$$

$\square$

Beweis.

2. Dies folgt etwa per Induktion über n aus 1., oder direkt folgendermaßen:

Sei o.E. $\mathbb{E}[X_1] = \cdots = \mathbb{E}[X_n] = 0$ (sonst ziehe jeweils die Erwartungswerte ab, verwende 1.), dann ist

$$\begin{aligned}
\mathrm{Var}\Big[\sum_{i=1}^{n} X_i\Big] &= \mathbb{E}\Big[\Big(\sum_{i=1}^{n} X_i\Big)^2\Big] = \sum_{i,j=1}^{n} \mathbb{E}[X_i X_j] \\
&= \sum_{i=1}^{n} \mathbb{E}[X_i^2] + \sum_{i\neq j}^{n} \mathbb{E}[X_i X_j] \\
&= \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{i\neq j}^{n} \mathrm{Cov}[X_i, X_j]
\end{aligned}$$

$\square$

Beweis.

3. Klar: für X und Y unabhängig ist $\mathbb{E}[X\,Y] = \mathbb{E}[X]\,\mathbb{E}[Y]$ nach Satz 1.70, 4, also $\mathrm{Cov}[X, Y] = \mathbb{E}[X\,Y] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0$

Beweis.

3. Klar: für X und Y unabhängig ist $\mathbb{E}[X\,Y] = \mathbb{E}[X]\,\mathbb{E}[Y]$ nach Satz 1.70, 4, also $\mathrm{Cov}[X, Y] = \mathbb{E}[X\,Y] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0$

4. Falls $\mathrm{Var}[Y] = 0$, so ist die Ungleichung $\big|\mathrm{Cov}[X, Y]\big| \le \sqrt{\mathrm{Var}[X]}\,\sqrt{\mathrm{Var}[Y]}$ (als $0 \le 0$) erfüllt

(denn dann ist $P(Y - \mathbb{E}[Y] = 0) = 1$ nach Beob. 1.78, 3. und somit auch $\mathrm{Cov}[X, Y] = 0$).

Beweis.

3. Klar: für X und Y unabhängig ist $\mathbb{E}[X\,Y] = \mathbb{E}[X]\,\mathbb{E}[Y]$ nach Satz 1.70, 4, also $\mathrm{Cov}[X, Y] = \mathbb{E}[X\,Y] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0$

4. Falls $\mathrm{Var}[Y] = 0$, so ist die Ungleichung $\left|\mathrm{Cov}[X, Y]\right| \leq \sqrt{\mathrm{Var}[X]}\,\sqrt{\mathrm{Var}[Y]}$ (als $0 \leq 0$) erfüllt

(denn dann ist $P(Y - \mathbb{E}[Y] = 0) = 1$ nach Beob. 1.78, 3. und somit auch $\mathrm{Cov}[X, Y] = 0$).

Falls $\mathrm{Var}[Y] > 0$, setze $\alpha := -\dfrac{\mathrm{Cov}[X, Y]}{\mathrm{Var}[Y]}$, es ist

$$\begin{aligned}
0 &\leq \mathrm{Var}[X + \alpha Y]\,\mathrm{Var}[Y] \\
&\overset{1.}{=} \big(\mathrm{Var}[X] + 2\alpha\mathrm{Cov}[X, Y] + \alpha^2\mathrm{Var}[Y]\big)\mathrm{Var}[Y] \\
&= \mathrm{Var}[X]\,\mathrm{Var}[Y] - \big(\mathrm{Cov}[X, Y]\big)^2.
\end{aligned}$$

$\square$

Bemerkung 1.81

Es gilt Gleichheit in der Cauchy-Schwarz-Ungleichung g.d.w.

es gibt $a, b, c \in \mathbb{R}$ (mit $a \neq 0$ oder $b \neq 0$), so dass
$$P\left(aX + bY + c = 0\right) = 1.$$

In diesem Fall heißen X und Y *perfekt korreliert*.

Bemerkung 1.81

Es gilt Gleichheit in der Cauchy-Schwarz-Ungleichung g.d.w.

es gibt $a, b, c \in \mathbb{R}$ (mit $a \neq 0$ oder $b \neq 0$), so dass

$$P\big(aX + bY + c = 0\big) = 1.$$

In diesem Fall heißen X und Y *perfekt korreliert*.

(Denn wir sehen aus dem Beweis, dass Gleichheit genau dann eintritt, wenn $\mathrm{Var}[Y] = 0$ oder $\mathrm{Var}[X + \alpha Y] = 0$.)

Beispiel 1.82

1. $X \sim \mathrm{Bin}_{n,p}$, schreibe $X = Y_1 + \cdots + Y_n$ mit Y_i u.i.v. $\sim \mathrm{Ber}_p$, so ist (mit Satz 1.80, 2.)

$$\mathrm{Var}[X] = \sum_{i=1}^{n} \mathrm{Var}[Y_i] = n\mathrm{Var}[Y_1] = np(1-p)$$

(vgl. auch Bsp. 1.79, 3.).

Beispiel 1.82 (Fortsetzung)

2. $X \sim \mathrm{Hyp}_{s,w,n}$, stelle dar als $X = Y_1 + \cdots + Y_n$ mit $Y_i = \mathbf{1}_{A_i}$,
   $A_i = \{i\text{-te gezogene Kugel ist schwarz}\}$
   (bei n-fachem Ziehen ohne Zurücklegen aus einer Urne mit s schwarzen und w weißen Kugeln)

Beispiel 1.82 (Fortsetzung)

2. $X \sim \mathrm{Hyp}_{s,w,n}$, stelle dar als $X = Y_1 + \cdots + Y_n$ mit $Y_i = \mathbf{1}_{A_i}$, $A_i = \{i\text{-te gezogene Kugel ist schwarz}\}$
(bei n-fachem Ziehen ohne Zurücklegen aus einer Urne mit s schwarzen und w weißen Kugeln)

Erinnerung: Für $y_1, \ldots, y_n \in \{0, 1\}$ mit $y_1 + \cdots + y_n = k$ gilt

$$P(Y_1 = y_1, \ldots, Y_n = y_n) \\ = \frac{s(s-1)(s-2)\cdots(s-k+1) \cdot w(w-1)\cdots(w-n+k+1)}{(s+w)(s+w-1)(s+w-2)\cdots(s+w-n+1)},$$

was nicht von der Reihenfolge abhängt – die Y_i sind „austauschbar“.

Beispiel 1.82 (Fortsetzung)

2. Es ist $\mathbb{E}[Y_i] = P(A_i) = P(A_1) = \dfrac{s}{s+w} =: p$,
   $\mathrm{Var}[Y_i] = p(1-p)$; für $i \neq j$ ist

$$\begin{aligned}
\mathbb{E}[Y_i\, Y_j] &= \mathbb{E}[Y_1\, Y_2] = P(A_1 \cap A_2) = \frac{s}{s+w}\,\frac{s-1}{s+w-1},\\
\mathrm{Cov}[Y_i, Y_j] &= \mathbb{E}[Y_i\, Y_i] - \mathbb{E}[Y_i]\,\mathbb{E}[Y_j]\\
&= \frac{s}{s+w}\,\frac{s-1}{s+w-1} - \Big(\frac{s}{s+w}\Big)^2\\
&= \frac{s}{s+w}\Big(\underbrace{\frac{s-1}{s+w-1} - \frac{s}{s+w}}_{=-\frac{w}{s+w}\,\frac{1}{s+w-1}}\Big)\\
&= -p(1-p)\,\frac{1}{s+w-1},
\end{aligned}$$

Beispiel 1.82 (Fortsetzung)

2. $X = Y_1 + \cdots + Y_n \sim \mathrm{Hyp}_{s,w,n}$
   mit $Y_i = \mathbf{1}_{A_i}$, $A_i = \{i\text{-te gezogene Kugel ist schwarz}\}$

Beispiel 1.82 (Fortsetzung)

2. $X = Y_1 + \cdots + Y_n \sim \mathrm{Hyp}_{s,w,n}$
   mit $Y_i = \mathbf{1}_{A_i}$, $A_i = \{i\text{-te gezogene Kugel ist schwarz}\}$
   Also (mit $p = s/(s+w)$)

$$\begin{aligned}
\mathrm{Var}[X] &= \sum_{i=1}^{n} \mathrm{Var}[Y_i] + \sum_{i \neq j}^{n} \mathrm{Cov}[Y_i, Y_j] \\
&= np(1-p) - n(n-1)\Big(- p(1-p)\frac{1}{s+w-1}\Big) \\
&= np(1-p)\Big(1 - \frac{n-1}{s+w-1}\Big)
\end{aligned}$$

Beispiel 1.82 (Fortsetzung)

2. $X = Y_1 + \cdots + Y_n \sim \mathrm{Hyp}_{s,w,n}$
   mit $Y_i = \mathbf{1}_{A_i}$, $A_i = \{i\text{-te gezogene Kugel ist schwarz}\}$
   Also (mit $p = s/(s+w)$)

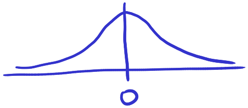
$$\begin{aligned}\mathrm{Var}[X] &= \sum_{i=1}^{n} \mathrm{Var}[Y_i] + \sum_{i\neq j}^{n} \mathrm{Cov}[Y_i, Y_j] \\ &= np(1-p) - n(n-1)\Big(-p(1-p)\frac{1}{s+w-1}\Big) \\ &= np(1-p)\Big(1 - \frac{n-1}{s+w-1}\Big)\end{aligned}$$

Wir sehen: Die Varianz ist kleiner im Fall ohne Zurücklegen als im Fall mit Zurücklegen – insbes. ist sie natürlich $= 0$ im Fall $n = s + w$.

Beispiel 1.82 (Fortsetzung)

3. Z reelle ZV mit $\mathbb{E}\big[|Z|^3\big] < \infty$ und *symmetrischer* Verteilung, d.h. es gilt $P(Z > z) = P(Z < -z)$ für alle $z \geq 0$ (z.B. $Z \sim \mathcal{N}_{0,1}$), setze

$$Y := Z^2$$



Beispiel 1.82 (Fortsetzung)

3. Z reelle ZV mit $\mathbb{E}\big[|Z|^3\big] < \infty$ und *symmetrischer* Verteilung, d.h. es gilt $P(Z > z) = P(Z < -z)$ für alle $z \geq 0$ (z.B. $Z \sim \mathcal{N}_{0,1}$), setze

$$Y := Z^2,$$

dann gilt

$$\begin{aligned}\mathrm{Cov}[Y, Z] &= \mathbb{E}[Z^2 Z] - \mathbb{E}[Z^2]\mathbb{E}[Z] \\ &= \mathbb{E}[Z^3] - \mathbb{E}[Z^2]\mathbb{E}[Z] = 0 - \mathbb{E}[Z^2] \cdot 0 = 0.\end{aligned}$$

Beispiel 1.82 (Fortsetzung)

3. Z reelle ZV mit $\mathbb{E}\big[|Z|^3\big] < \infty$ und *symmetrischer* Verteilung, d.h. es gilt $P(Z > z) = P(Z < -z)$ für alle $z \geq 0$ (z.B. $Z \sim \mathcal{N}_{0,1}$), setze
$$Y := Z^2,$$
dann gilt
$$\begin{aligned}\mathrm{Cov}[Y, Z] &= \mathbb{E}[Z^2 Z] - \mathbb{E}[Z^2]\mathbb{E}[Z]\\ &= \mathbb{E}[Z^3] - \mathbb{E}[Z^2]\mathbb{E}[Z] = 0 - \mathbb{E}[Z^2]\cdot 0 = 0.\end{aligned}$$
Z und Y sind also unkorreliert, aber i.A. *nicht* unabhängig.

Definition 1.83

Seien $X, Y \in \mathcal{L}^2$.

$$\kappa_{X,Y} := \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}} \in [-1, 1]$$

heißt *Korrelationskoeffizient* von X und Y (manche Autoren schreiben auch $\rho_{X,Y}$).

(Die Cauchy-Schwarz-Ungleichung (Satz 1.80, 4.) zeigt, dass $|\kappa_{X,Y}| \leq 1$.)

Insbes.: $a, b, c, d \in \mathbb{R}$, $a, c > 0$

$$\kappa_{aX+b,\,cY+d} = \kappa_{X,Y}$$

Beobachtung 1.84 (Interpretation des Korrelationskoeffizienten via „beste lineare Vorhersage“)

Es ist

$$\min_{\beta_0,\beta_1\in\mathbb{R}} \mathbb{E}\big[\overbrace{(Y-\beta_1 X-\beta_0)^2}^{=(Y-(\beta_1 X+\beta_0))^2}\big]$$

$$= (1-\kappa^2_{X,Y}) \min_{\beta_0\in\mathbb{R}} \mathbb{E}\big[(Y-\beta_0)^2\big] \quad \Big(= (1-\kappa^2_{X,Y})\mathrm{Var}[Y]\Big),$$

Denn

$$
\begin{aligned}
&\mathbb{E}\big[(Y - \beta_1 X - \beta_0)^2\big] \\
&= \mathrm{Var}[Y - \beta_1 X - \beta_0] + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2 \\
&= \mathrm{Var}[Y] - 2\beta_1 \mathrm{Cov}[X, Y] + \beta_1^2 \mathrm{Var}[X] + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2 \\
&= \sigma_Y^2 - 2\beta_1 \sigma_X \sigma_Y \kappa_{X,Y} + \beta_1^2 \sigma_X^2 + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2 \\
&= \sigma_Y^2\big(1 - \kappa_{X,Y}^2\big) + \sigma_X^2\Big(\beta_1 - \frac{\sigma_Y}{\sigma_X}\kappa_{X,Y}\Big)^2 + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2
\end{aligned}
$$

Denn

$$
\begin{aligned}
&\mathbb{E}\big[(Y - \beta_1 X - \beta_0)^2\big] \\
&= \mathrm{Var}[Y - \beta_1 X - \beta_0] + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2 \\
&= \mathrm{Var}[Y] - 2\beta_1 \mathrm{Cov}[X, Y] + \beta_1^2 \mathrm{Var}[X] + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2 \\
&= \sigma_Y^2 - 2\beta_1 \sigma_X \sigma_Y \kappa_{X,Y} + \beta_1^2 \sigma_X^2 + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2 \\
&= \sigma_Y^2\big(1 - \kappa_{X,Y}^2\big) + \sigma_X^2\Big(\beta_1 - \frac{\sigma_Y}{\sigma_X}\kappa_{X,Y}\Big)^2 + \big(\mathbb{E}[Y] - \beta_1 \mathbb{E}[X] - \beta_0\big)^2
\end{aligned}
$$

was offensichtlich minimal wird für die Wahl

$$
\beta_1 = \beta_1^* := \frac{\sigma_Y}{\sigma_X}\kappa_{X,Y}, \quad \beta_0 = \beta_0^* := \mathbb{E}[Y] - \beta_1^* \mathbb{E}[X]
$$

und dann den Wert $(1 - \kappa_{X,Y}^2)\sigma_Y^2$ hat.

Für den Zusatz beachte analog:

$$\mathbb{E}\left[(Y-\beta_0)^2\right] = \mathbb{E}[Y^2] - 2\beta\mathbb{E}[Y] + \beta^2 = \mathrm{Var}[Y] + \big(\beta - \mathbb{E}[Y]\big)^2$$

ist minimal für die Wahl $\beta = \mathbb{E}[Y]$.

Für den Zusatz beachte analog:

$$\mathbb{E}\big[(Y-\beta_0)^2\big] = \mathbb{E}[Y^2] - 2\beta\mathbb{E}[Y] + \beta^2 = \mathrm{Var}[Y] + \big(\beta - \mathbb{E}[Y]\big)^2$$

ist minimal für die Wahl $\beta = \mathbb{E}[Y]$.

Im Sinne einer möglichst kleinen quadratischen Abweichung ist $\mathbb{E}[Y]$ die beste konstante „Vorhersage“ von Y. Man kann demnach um einen Faktor $(1-\kappa^2_{X,Y})$ besser vorhersagen, wenn man stattdessen eine affin-lineare Funktion von X verwenden darf.

Für den Zusatz beachte analog:

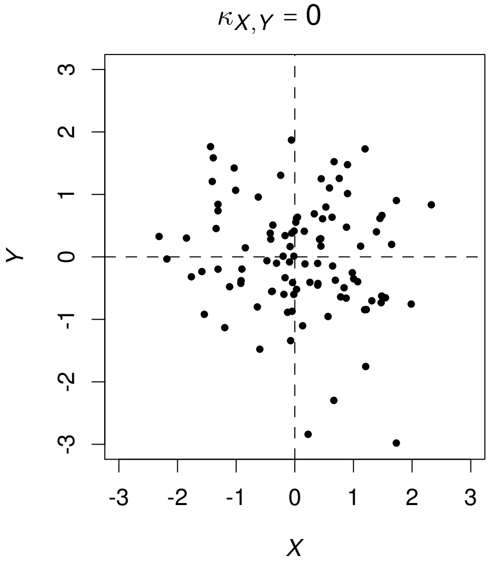
$$\mathbb{E}\big[(Y-\beta_0)^2\big] = \mathbb{E}[Y^2] - 2\beta\mathbb{E}[Y] + \beta^2 = \mathrm{Var}[Y] + \big(\beta - \mathbb{E}[Y]\big)^2$$

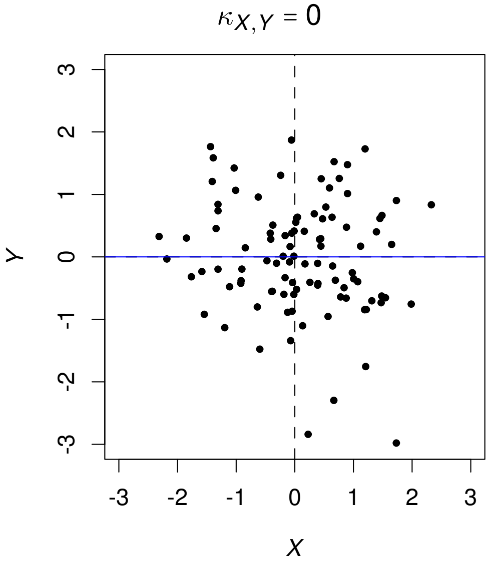
ist minimal für die Wahl $\beta = \mathbb{E}[Y]$.

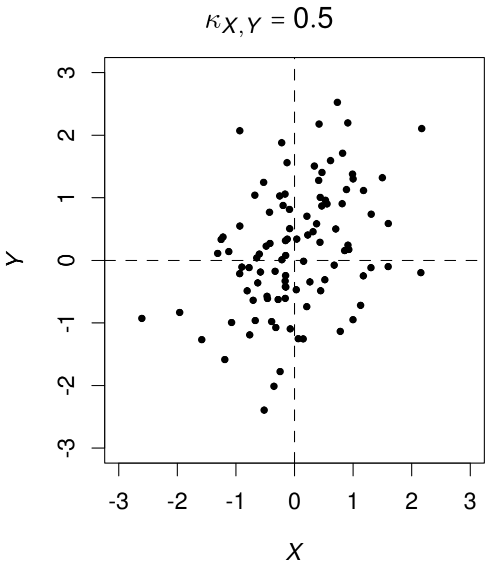
Im Sinne einer möglichst kleinen quadratischen Abweichung ist $\mathbb{E}[Y]$ die beste konstante „Vorhersage" von Y. Man kann demnach um einen Faktor $(1-\kappa_{X,Y}^2)$ besser vorhersagen, wenn man stattdessen eine affin-lineare Funktion von X verwenden darf.

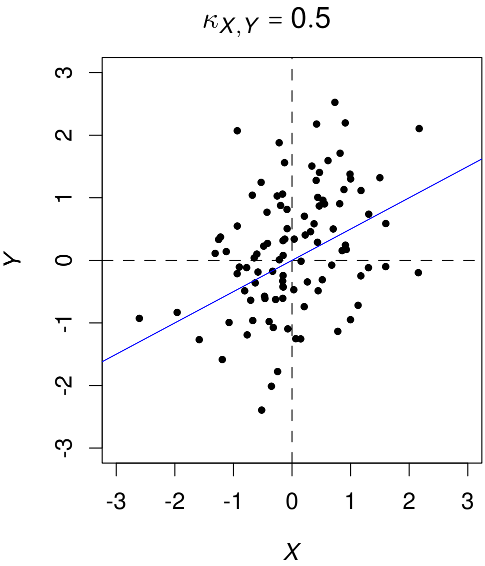
Die folgenden Scatterplots zeigen jeweils 100 simulierte Paare (X, Y), wobei $\sigma_X = \sigma_Y = 1$ und $\kappa_{X,Y}$ den angegebenen Wert hat.

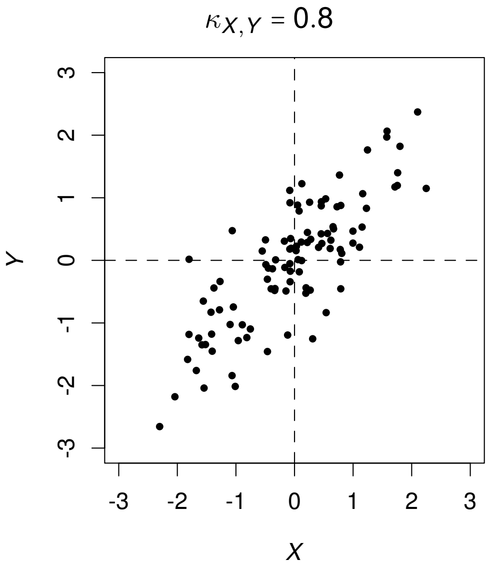
(Blau eingezeichnet ist die „Vorhersagegerade" $x \mapsto \beta_1^* x + \beta_0^*$.)

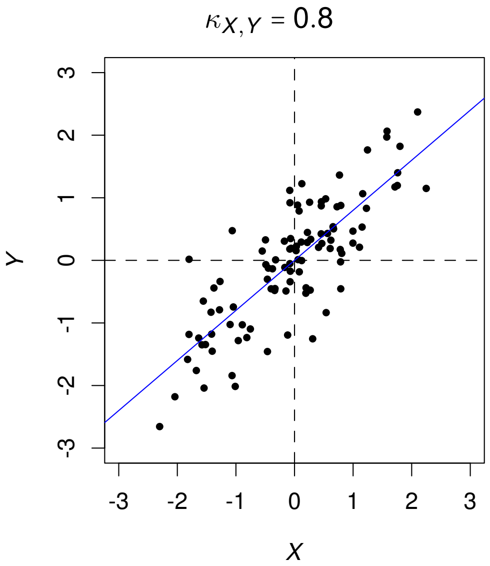

$\kappa_{X,Y} = 0$
3
2
1
0
-1
-2
-3
Y
-3
-2
-1
0
1
2
3
X

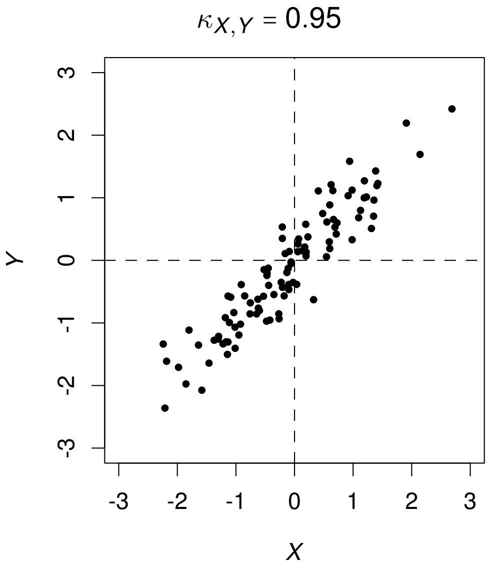

$\kappa_{X,Y} = 0$
Y
X
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3

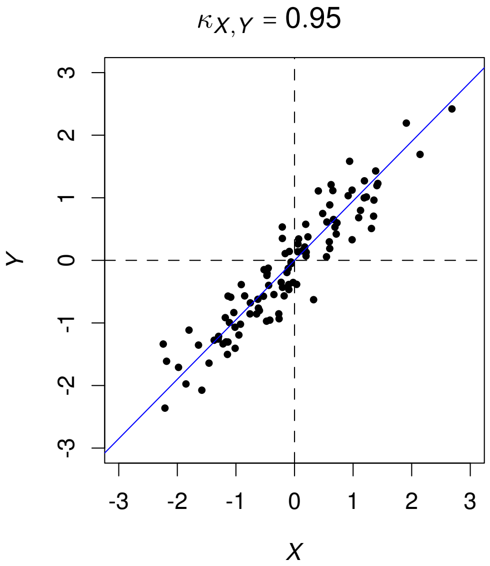

$\kappa_{X,Y} = 0.5$
Y
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
X


$\kappa_{X,Y} = 0.5$
3
2
1
0
-1
-2
-3
Y
-3
-2
-1
0
1
2
3
X

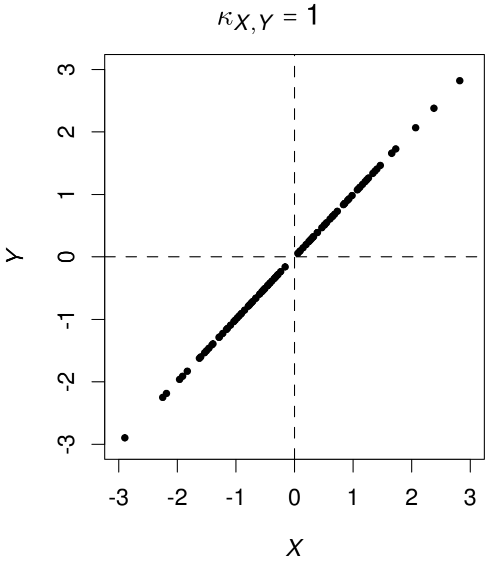

$\kappa_{X,Y} = 0.8$
Y
X
-3
-2
-1
0
1
2
3

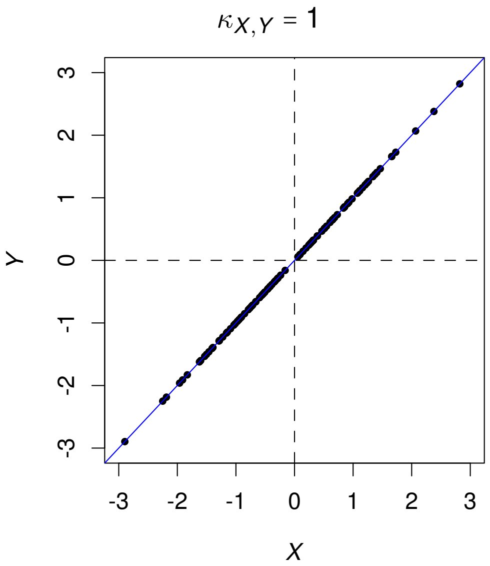

$\kappa_{X,Y} = 0.8$
Y
X
-3
-2
-1
0
1
2
3

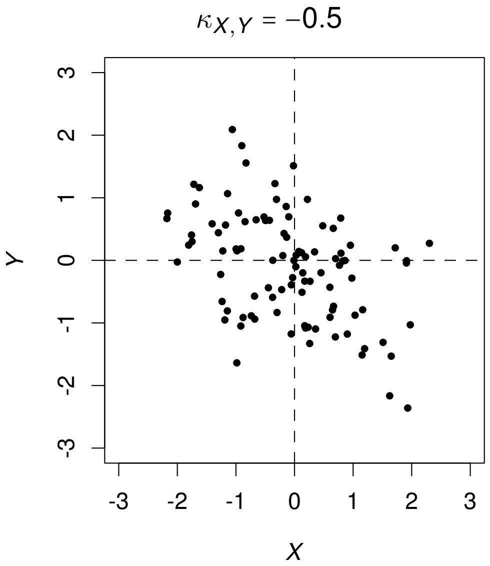

$\kappa_{X,Y} = 0.95$
Y
X
-3
-2
-1
0
1
2
3

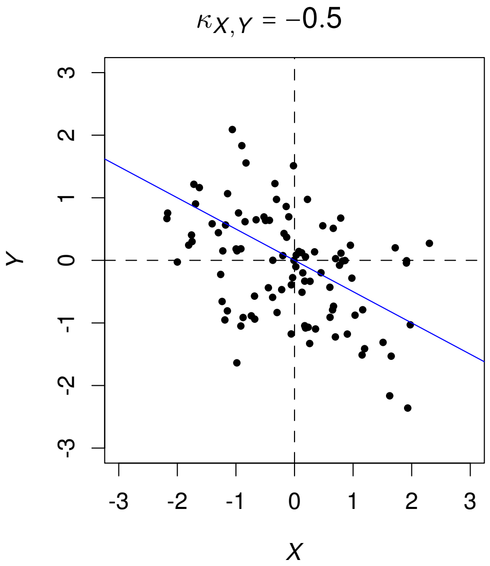


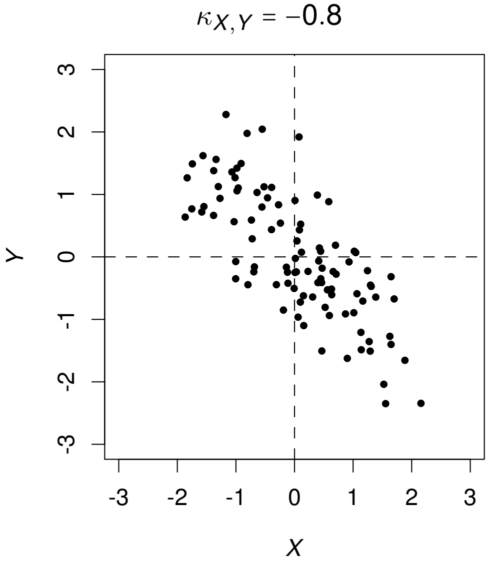
$\kappa_{X,Y} = 0.95$
3
2
1
0
-1
-2
-3
Y
-3
-2
-1
0
1
2
3
X

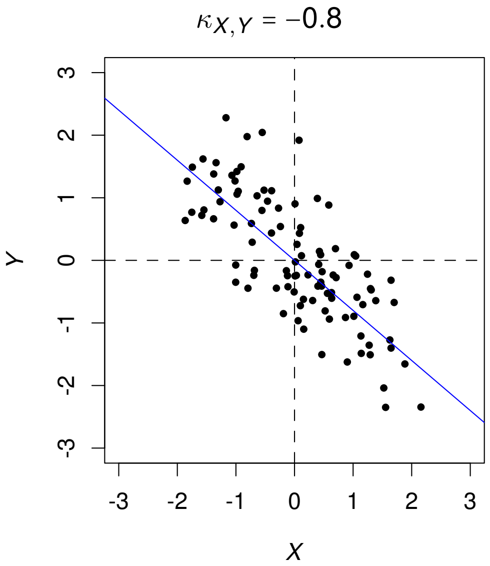

$\kappa_{X,Y} = 1$
3
2
1
0
-1
-2
-3
Y
-3
-2
-1
0
1
2
3
X

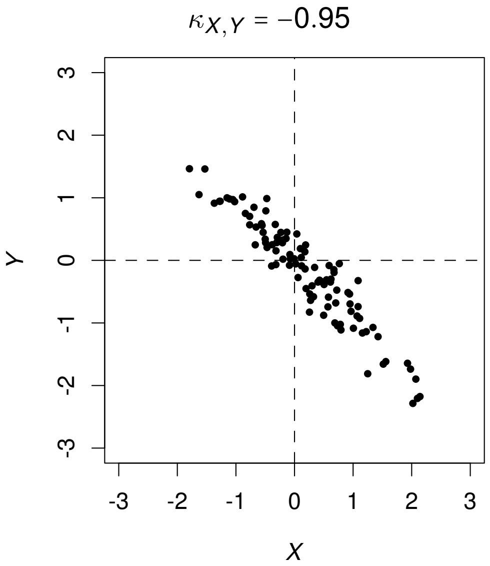

$\kappa_{X,Y} = 1$
Y
-3
-2
-1
0
1
2
3
X

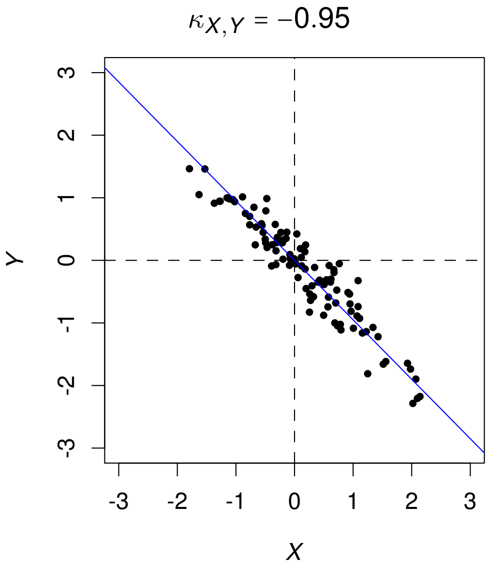

$\kappa_{X,Y} = -0.5$
Y
X
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3

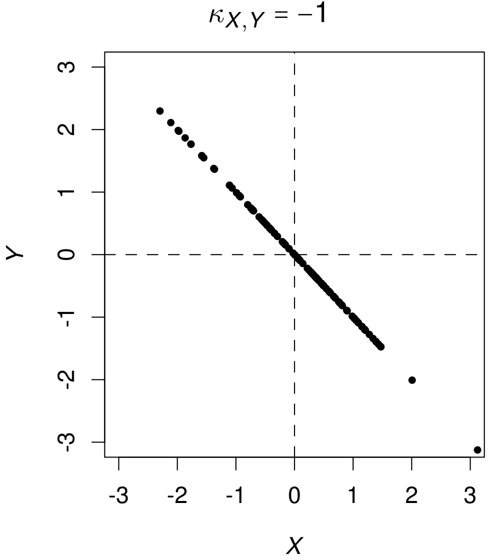

$\kappa_{X,Y} = -0.5$
Y
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
X

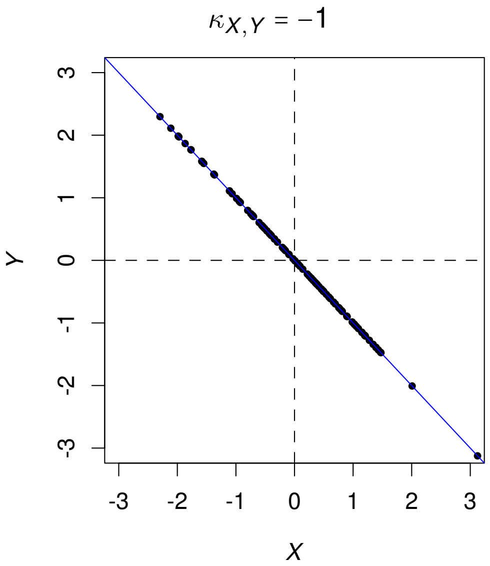

$\kappa_{X,Y} = -0.8$
Y
X
-3
-2
-1
0
1
2
3

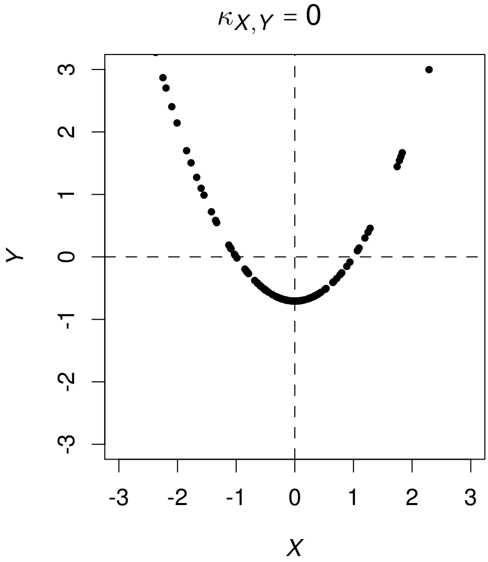

$\kappa_{X,Y} = -0.8$
Y
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
X

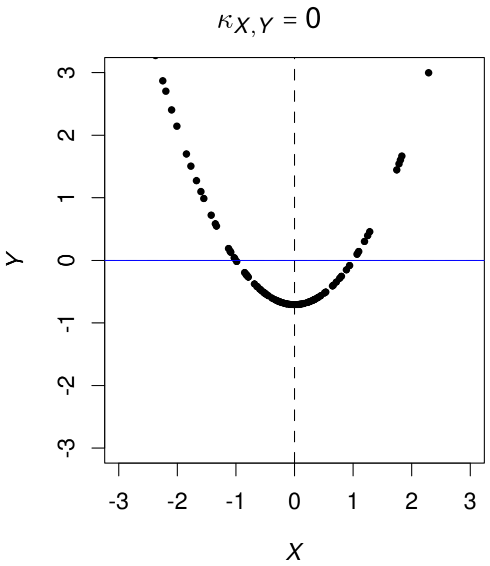

$\kappa_{X,Y} = -0.95$
Y
X
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3


$\kappa_{X,Y} = -0.95$
Y
3
2
1
0
-1
-2
-3
-3 -2 -1 0 1 2 3
X


$\kappa_{X,Y} = -1$
Y
X
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3


$\kappa_{X,Y} = -1$
3
2
1
0
-1
-2
-3
Y
-3
-2
-1
0
1
2
3
X


$\kappa_{X,Y} = 0$
Y
X
-3
-2
-1
0
1
2
3


$\kappa_{X,Y} = 0$
Y
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
X

Demnach (vgl. auch Bem. 1.81)

$|\kappa_{X,Y}| = 1 \quad \leftrightarrow$ perfekter linearer Zusammenhang zwischen X und Y

$\kappa_{X,Y} = 1 \quad \leftrightarrow$ perfekter linearer Zusammenhang zwischen X und Y mit positivem Koeffizienten
(X größer als $\mathbb{E}[X] \iff Y$ größer als $\mathbb{E}[Y]$)

$\kappa_{X,Y} = -1 \quad \leftrightarrow$ perfekter linearer Zusammenhang zwischen X und Y mit negativem Koeffizienten
(X größer als $\mathbb{E}[X] \iff Y$ kleiner als $\mathbb{E}[Y]$)

Nicht-lineare Zusammenhänge erfasst der Korrelationskoeffizient möglicherweise nicht korrekt (oder gar nicht), vgl. Bsp. 1.82, 3.

Median(e)

Anschaulich ist der Median einer reellen Zufallsvariable X der Wert m, so dass

$$„P(X \leq m) = \frac{1}{2} = P(X \geq m)“$$

gilt.

Median(e)

Anschaulich ist der Median einer reellen Zufallsvariable X der Wert m, so dass

$$„P(X \leq m) = \frac{1}{2} = P(X \geq m)“$$

gilt.

Da man diese Gleichheit (zumal im diskreten Fall) nicht immer genau einstellen kann, definiert man formal folgendermaßen:

Definition 1.85

X reelle ZV, m heißt (ein) Median von X (auch „Zentralwert“, manchmal auch m_X geschrieben), wenn gilt

$$P(X \geq m) \geq \frac{1}{2} \quad \text{und} \quad P(X \leq m) \geq \frac{1}{2}.$$

m Median von X, wenn

$$P(X \geq m) \geq \frac{1}{2} \quad \text{und} \quad P(X \leq m) \geq \frac{1}{2}.$$

Ein Median existiert stets, auch wenn X keinen Erwartungswert besitzt.

m Median von X, wenn

$$P(X \geq m) \geq \frac{1}{2} \quad \text{und} \quad P(X \leq m) \geq \frac{1}{2}.$$

Ein Median existiert stets, auch wenn X keinen Erwartungswert besitzt.

Man kann den Median als eine „robustere“ Antwort auf die Aufgabe, für eine ZV *nur einen* „typischen Wert“ anzugeben, ansehen (im Gegensatz zum Erwartungswert besitzt ja jede Verteilung einen Median).

m Median von X, wenn

$$P(X \geq m) \geq \frac{1}{2} \quad \text{und} \quad P(X \leq m) \geq \frac{1}{2}.$$

Ein Median existiert stets, auch wenn X keinen Erwartungswert besitzt.

Man kann den Median als eine „robustere“ Antwort auf die Aufgabe, für eine ZV *nur einen* „typischen Wert“ anzugeben, ansehen (im Gegensatz zum Erwartungswert besitzt ja jede Verteilung einen Median).

Allerdings gibt es für Mediane keine so angenehmen Rechenregeln, wie sie Satz 1.70 für den Erwartungswert liefert.

m Median von X, wenn

$$P(X \geq m) \geq \frac{1}{2} \quad \text{und} \quad P(X \leq m) \geq \frac{1}{2}.$$

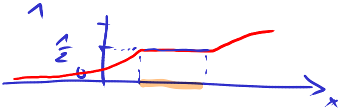
Bemerkung

Falls der Median von X (im Sinne von Def. 1.85) uneindeutig ist, d.h. wenn die Verteilungsfunktion F_X von X den Wert $\frac{1}{2}$ auf einem nicht-trivialen Intervall annimmt, so betrachtet man gelegentlich „pragmatisch“

$$\frac{1}{2}\Big(\inf\{x \in \mathbb{R} : F_X(x) = \tfrac{1}{2}\} + \sup\{x \in \mathbb{R} : F_X(x) = \tfrac{1}{2}\}\Big),$$

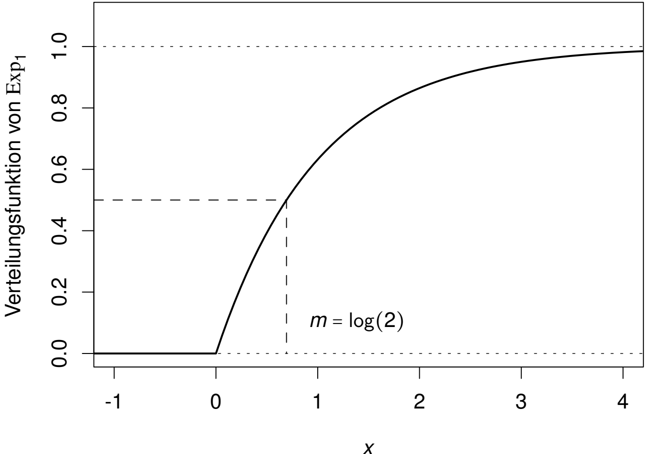
das arithmetische Mittel des kleinst- und des größtmöglichen Medians, als „den“ Median.

(So berechnet es beispielsweise R.)



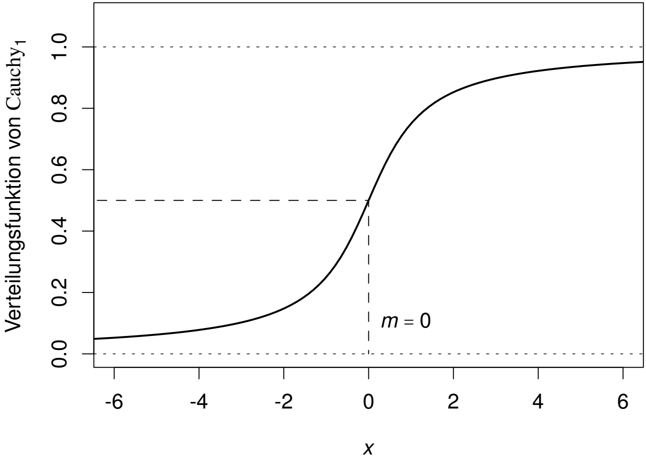
Beispiel 1.86

1. $X \sim \mathrm{Exp}_\theta$ hat Dichte $\theta e^{-\theta x}\mathbf{1}_{[0,\infty)}(x)$, Verteilungsfunktion $(1 - e^{-\theta x})\mathbf{1}_{[0,\infty)}(x)$, demnach ist der (eindeutig bestimmte) Median $m = \frac{1}{\theta}\log 2$.



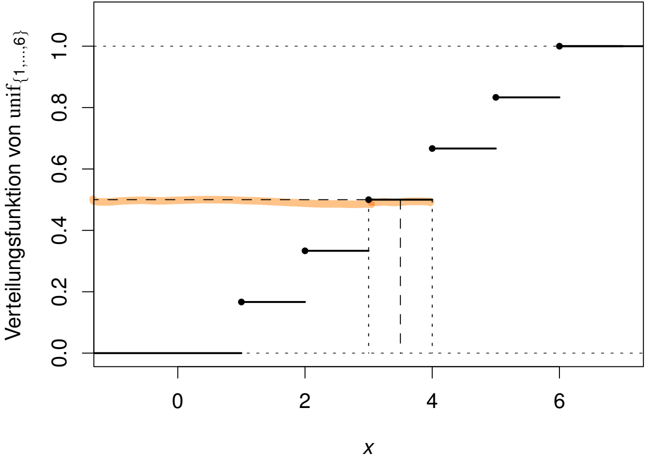
Beispiel 1.86 (Fortsetzung)

2. X Cauchy-verteilt mit Dichte $\frac{1}{\pi}\frac{1}{1+x^2}$, Verteilungsfunktion $\frac{1}{2}\frac{1}{\pi}\,\mathrm{arctan}(x)$, der (eindeutig bestimmte) Median ist $m = 0$.



Beispiel 1.86 (Fortsetzung)

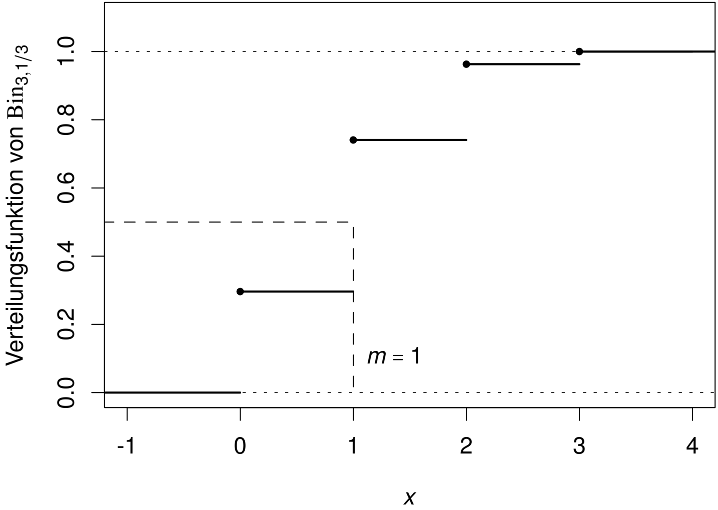
3. $X \sim \mathrm{unif}_{\{1,2,\ldots,6\}}$



Jeder Wert $m \in [3, 4]$ ist ein Median (und die vielleicht „kanonischste" Wahl wäre $m = 3{,}5$).

Beispiel 1.86 (Fortsetzung)

4. $X \sim \mathrm{Bin}_{3,1/3}$ hat Median 1



Bemerkung 1.87

Sei $X \in \mathcal{L}^1$.

$\operatorname{argmin} \mathbb{E}[(X-a)^2] = \mathbb{E}[X]$

1. Jeder Median von X ist ein Minimierer von
   $a \mapsto \mathbb{E}\big[|X-a|\big]$.

2. Für jeden Median m ist $\big|\mathbb{E}[X] - m\big| \leq \sqrt{\mathrm{Var}[X]}$.

Bemerkung 1.87

Sei $X \in \mathcal{L}^1$.

1. Jeder Median von X ist ein Minimierer von $a \mapsto \mathbb{E}\big[|X - a|\big]$.

2. Für jeden Median m ist $\big|\mathbb{E}[X] - m\big| \le \sqrt{\mathrm{Var}[X]}$.

Beweis. 1. Sei m ein Median. Falls $a > m$:

$$|X - a| - |X - m| \ge (a - m)\mathbf{1}_{\{X \le m\}} - (a - m)\mathbf{1}_{\{X > m\}},$$

also

$$\mathbb{E}\big[|X - a|\big] - \mathbb{E}\big[|X - m|\big] \ge (a - m)\big(\underbrace{P(X \le m)}_{\ge 1/2} - \underbrace{P(X > m)}_{\le 1/2}\big) \ge 0,$$

analog im Fall $a < m$.

Bemerkung 1.87

Sei $X \in \mathcal{L}^1$.

1. Jeder Median von X ist ein Minimierer von $a \mapsto \mathbb{E}\big[|X - a|\big]$.
2. Für jeden Median m ist $\big|\mathbb{E}[X] - m\big| \leq \sqrt{\mathrm{Var}[X]}$.

Bemerkung 1.87

Sei $X \in \mathcal{L}^1$.

1. Jeder Median von X ist ein Minimierer von $a \mapsto \mathbb{E}\big[|X-a|\big]$.
2. Für jeden Median m ist $\big|\mathbb{E}[X]-m\big| \leq \sqrt{\mathrm{Var}[X]}$.

Beweis. 2. Es ist

$$\big|\mathbb{E}[X]-m\big| = \big|\mathbb{E}[X-m]\big| \leq \mathbb{E}\big[|X-m|\big]$$

$$\overset{1.}{\leq} \mathbb{E}\big[|X-\mathbb{E}[X]|\big] = \sqrt{\Big(\mathbb{E}\big[|X-\mathbb{E}[X]|\big]\Big)^2} \leq \sqrt{\mathbb{E}\big[|X-\mathbb{E}[X]|^2\big]},$$

wobei für die erste Ungleichung die Monotonie des Erwartungswerts (beachte: $X-m \leq |X-m|$ und $-(X-m) \leq |X-m|$) und für die letzte Ungleichung die Cauchy-Schwarz-Ungleichung verwenden.

Statistik für Informatiker, SS 2021

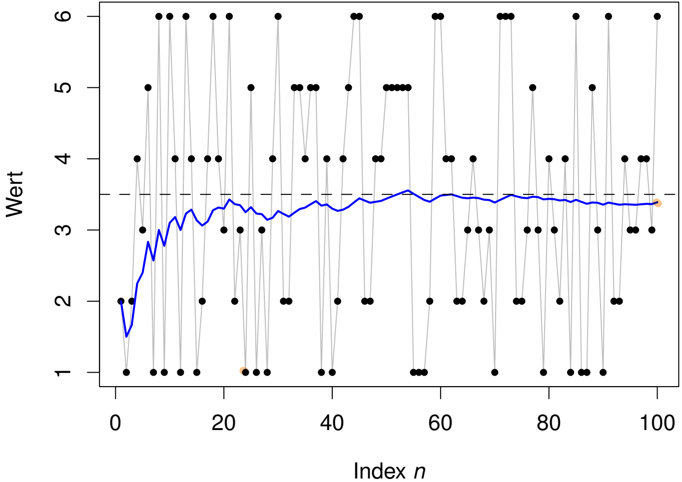
1 Grundlagen aus der Wahrscheinlichkeitstheorie

1.5 Gesetz der großen Zahlen und zentraler Grenzwertsatz

Matthias Birkner

Erinnerung. Wir hatten bereits in Bem. 1.68, 3 das Gesetz der großen Zahlen illustriert:

$X_1, X_2, \dots$ unabh., uniform auf $\{1, 2, 3, 4, 5, 6\}$, X_n sind die schwarzen Punkte, $(X_1 + \cdots + X_n)/n$ die blaue Linie



Satz 1.88 ((Schwaches) Gesetz der großen Zahlen)

Seien $X_1, X_2, \ldots$ unabhängige und identisch verteilte (u.i.v.) reellwertige ZVn mit $\mathbb{E}[X_1] = \mu$ und $\mathrm{Var}[X_1] < \infty$, dann gilt für jedes $\varepsilon > 0$

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{\mathrm{Var}[X_1]}{\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0$$

Satz 1.88 ((Schwaches) Gesetz der großen Zahlen)

Seien $X_1, X_2, \dots$ unabhängige und identisch verteilte (u.i.v.) reellwertige ZVn mit $\mathbb{E}[X_1] = \mu$ und $\mathrm{Var}[X_1] < \infty$, dann gilt für jedes $\varepsilon > 0$

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \varepsilon\right) \le \frac{\mathrm{Var}[X_1]}{\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0$$

Beweis. Sei $Y_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)$, es ist

$$\mathrm{Var}[Y_n] = \frac{1}{n^2}\left(\underbrace{\sum_{i=1}^{n}\mathrm{Var}[X_i - \mu]}_{= n\cdot\mathrm{Var}[X_1]} + \cancel{\sum_{\substack{1\le i,j\le n\\ i\neq j}} \mathrm{Cov}[X_i - \mu, X_j - \mu]}\right)$$

$$= n\cdot\frac{1}{n^2}\mathrm{Var}[X_i] = \frac{1}{n}\mathrm{Var}[X_1]$$

Satz 1.88 ((Schwaches) Gesetz der großen Zahlen)

Seien $X_1, X_2, \ldots$ unabhängige und identisch verteilte (u.i.v.) reellwertige ZVn mit $\mathbb{E}[X_1] = \mu$ und $\mathrm{Var}[X_1] < \infty$, dann gilt für jedes $\varepsilon > 0$

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{\mathrm{Var}[X_1]}{\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0$$

$$Y_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \text{ hat } \mathrm{Var}[Y_n] = \frac{1}{n}\mathrm{Var}[X_1]$$

Satz 1.88 ((Schwaches) Gesetz der großen Zahlen)

Seien $X_1, X_2, \dots$ unabhängige und identisch verteilte (u.i.v.) reellwertige ZVn mit $\mathbb{E}[X_1] = \mu$ und $\mathrm{Var}[X_1] < \infty$, dann gilt für jedes $\varepsilon > 0$

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{\mathrm{Var}[X_1]}{\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0$$

$$Y_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \text{ hat } \mathrm{Var}[Y_n] = \frac{1}{n}\mathrm{Var}[X_1],$$

somit

$$P\big(|Y_n| \geq \varepsilon\big) \leq \frac{\mathrm{Var}[Y_n]}{\varepsilon^2} = \frac{\mathrm{Var}[X_1]}{\varepsilon^2 n}$$

gemäß Chebyshev-Ungleichung.

Bemerkung 1.89

1. Wir entnehmen dem Beweis von Satz 1.88 folgende kleine Verallgemeinerung:
   Sind $X_1, X_2, \dots \in \mathcal{L}^2$ seien paarweise unkorreliert mit

$$\sup_n \mathrm{Var}[X_n] \leq \theta < \infty,$$

Bemerkung 1.89

1. Wir entnehmen dem Beweis von Satz 1.88 folgende kleine Verallgemeinerung:
   Sind $X_1, X_2, \dots \in \mathcal{L}^2$ seien paarweise unkorreliert mit

   $$\sup_n \mathrm{Var}[X_n] \leq \theta < \infty,$$

   dann gilt für $\varepsilon > 0$

   $$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\big(X_i - \mathbb{E}[X_i]\big)\right| > \varepsilon\right) \leq \frac{\theta}{\varepsilon^2 n} \; \big(\xrightarrow[n\to\infty]{} 0\big).$$

   (Das Argument geht genauso wie im Beweis von Satz 1.88, wenn wir $Y_n := \frac{1}{n}\sum_{i=1}^{n}\big(X_i - \mathbb{E}[X_i]\big)$ setzen.)