Kapitel 3

Erwartungswert, Varianz und Kovarianz

Der Erwartungswert ist eine wichtige Kenngröße der Verteilung einer reellwertigen Zufallsvariable X, er gibt eine Antwort auf die – etwas salopp formulierte – Frage „Wie groß ist X typischerweise?“

3.1 Diskreter Fall

Sei X reelle ZV mit abzählbarem Wertebereich (auf einem W'raum $(\Omega, \mathcal{F}, P)$ definiert), d.h. es gibt eine abzählbare Menge $S = S_X \subset \mathbb{R}$ mit $P(X \in S) = 1$ und $\mathscr{L}_P(X)$ hat Gewichte $P(X = x), x \in S$.

Definition 3.1. X besitzt einen Erwartungswert, wenn

$$\sum_{x \in S_X} |x| P(X = x) < \infty$$

gilt, man schreibt dies auch als $X \in \mathcal{L}^1$ (bzw. $X \in \mathcal{L}^1(P)$, wenn das zugrundeliegende W'maß P nicht aus dem Kontext klar ist).
In diesem Fall heißt

$$\mathbb{E}[X] := \sum_{x \in S_X} x P(X = x)$$

der Erwartungswert von X.

ZV $X \in \mathcal{L}^1 :\Longleftrightarrow \sum_{x \in S_X} |x| P(X = x) < \infty$, dann setzt man $\mathbb{E}[X] = \sum_{x \in S_X} x P(X = x)$

Bemerkung. Die Summe $\sum_{x \in S_X} x P(X = x)$ ist dann wohldefiniert (unabhängig von der Summationsreihenfolge) und es gilt $|\mathbb{E}[X]| \leq \sum_{x \in S_X} |x| P(X = x) < \infty$.

(Aussage aus der Analysis: absolut konvergente Reihen dürfen in bel. Reihenfolge summiert werden.)

Ein „Gegenbeispiel“:

$$P(X = n) = P(X = -n) = \frac{1}{2n(n-1)} \quad \text{für } n = 2, 3, \ldots$$

(d.h. $P(X=n) = \frac{1}{2|n|(|n|-1)}$ falls $|n| \geq 2$)

(es ist $\sum_{n=2}^{\infty} 2 \frac{1}{2n(n-1)} = \sum_{n=2}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n} \right) = 1$, d.h. dies sind W'gewichte), wenn man die Werte durchnummerierte mit $x_{2i} = i + 1$, $x_{2i-1} = -i - 1$, $i \in \mathbb{N}$, so wäre

(bel. Reihenfolge $-2, 2, -3, 3, -4, 4, \ldots$)

$$\sum_{j=1}^{\infty} x_j P(X = x_j) = \lim_{N \to \infty} \sum_{j=1}^{N} x_j P(X = x_j) = 0,$$

andererseits ist $\sum_{j=1}^{\infty} |x_j| P(X = x_j) = \sum_{n=2}^{\infty} \frac{n}{n(n-1)} = \sum_{k=1}^{\infty} \frac{1}{k} = \infty$.

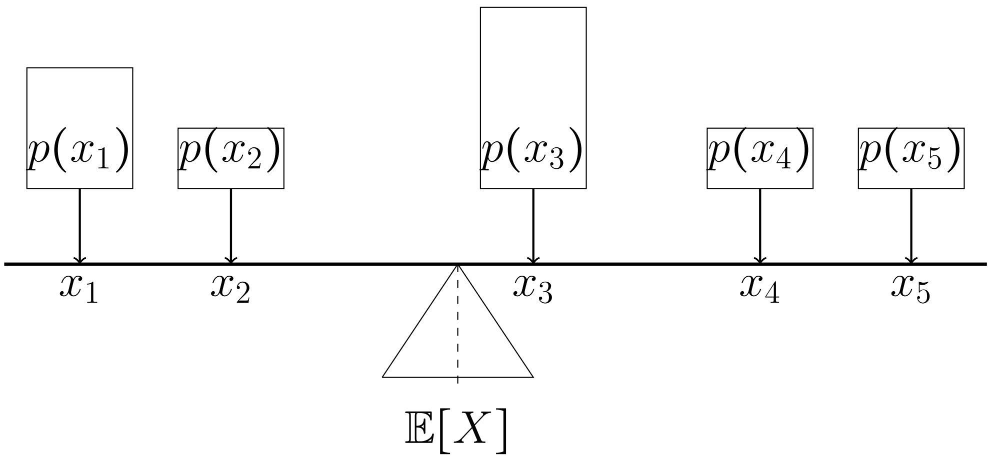
Dieses X besitzt keinen Erwartungswert (im Sinne unserer Def.).

Bemerkung 3.2. 1. Falls $|\Omega| < \infty$, so besitzt jede ZV eine Erwartungswert und es gilt

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} P(\{\omega\}) X(\omega).$$

$$\left(\sum_{x \in S_X} x P(X = x) = \sum_{x \in S_X} x \sum_{\omega : X(\omega) = x} P(\{\omega\}) = \sum_{x \in S_X} \sum_{\omega : X(\omega) = x} X(\omega) P(\{\omega\}) \right)$$

2. Wenn X endlich viele mögliche Werte $x_1, \dots, x_n$ (mit Gewichten $p(x_i) = P(X = x_i)$) hat, so besitzt es einen Erwartungswert und man kann $\mathbb{E}[X]$ als den „Massenschwerpunkt" interpretieren.



Auf einer Balkenwaage (deren Balken Eigengewicht 0 habe) liege an der Position x_i das Gewicht $p(x_i)$. Damit der Balken in Ruhelage ist, muss man ihn an der Stelle $\sum_{i=1}^n x_i p(x_i) = \mathbb{E}[X]$ unterstützen, denn dann ist das Gesamtdrehmonent (proportional zu)

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

3. Wenn $X \geq 0$ (bzw. $P(X \geq 0) = 1$), so ist $\sum_x x\, P(X = x)$ stets wohldefiniert (möglicherweise mit Wert $+\infty$, den man dann formal zulässt).

4. Der Erwartungswert von X muss nicht notwendigerweise ein möglicher Wert von X sein: $P(X = \mathbb{E}[X]) = 0$ ist durchaus möglich, beispielsweise gilt für W das Ergebnis eines fairen Würfelwurfs

$$\mathbb{E}[W] = \sum_{w=1}^{6} \frac{1}{6} w = \frac{7}{2} \notin \{1, 2, \ldots, 6\}$$

Daher kann man die Interpretation von $\mathbb{E}[X]$ als „typischer Wert von X" i.A. nicht wörtlich nehmen.

(mit Werteber. S_X)

Es gilt aber: Sind $X_1, X_2, \ldots$ unabhängig mit derselben Verteilung wie X, so konvergiert

$$M_n := \frac{X_1 + X_2 + \cdots + X_n}{n} \underset{n\to\infty}{\longrightarrow} \mathbb{E}[X] = \sum_x x P(X = x)$$

(in geeignetem Sinn), dies ist die Aussage des *Gesetzes der großen Zahlen*, das wir später sehen werden.

Es ist nämlich

$$M_n = \sum_{x \in S_X} x \cdot \frac{\#\{i \le n : X_i = x\}}{n}$$

und $\#\{i \le n : X_i = x\}/n \underset{n\to\infty}{\longrightarrow} P(X = x)$.

Illustration: $X_1, X_2, \ldots$ uniform auf $\{1, 2, 3, 4, 5, 6\}$,
X_n sind die schwarzen Punkte,
$M_n = (X_1 + X_2 + \cdots + X_n)/n$ die blaue Linie

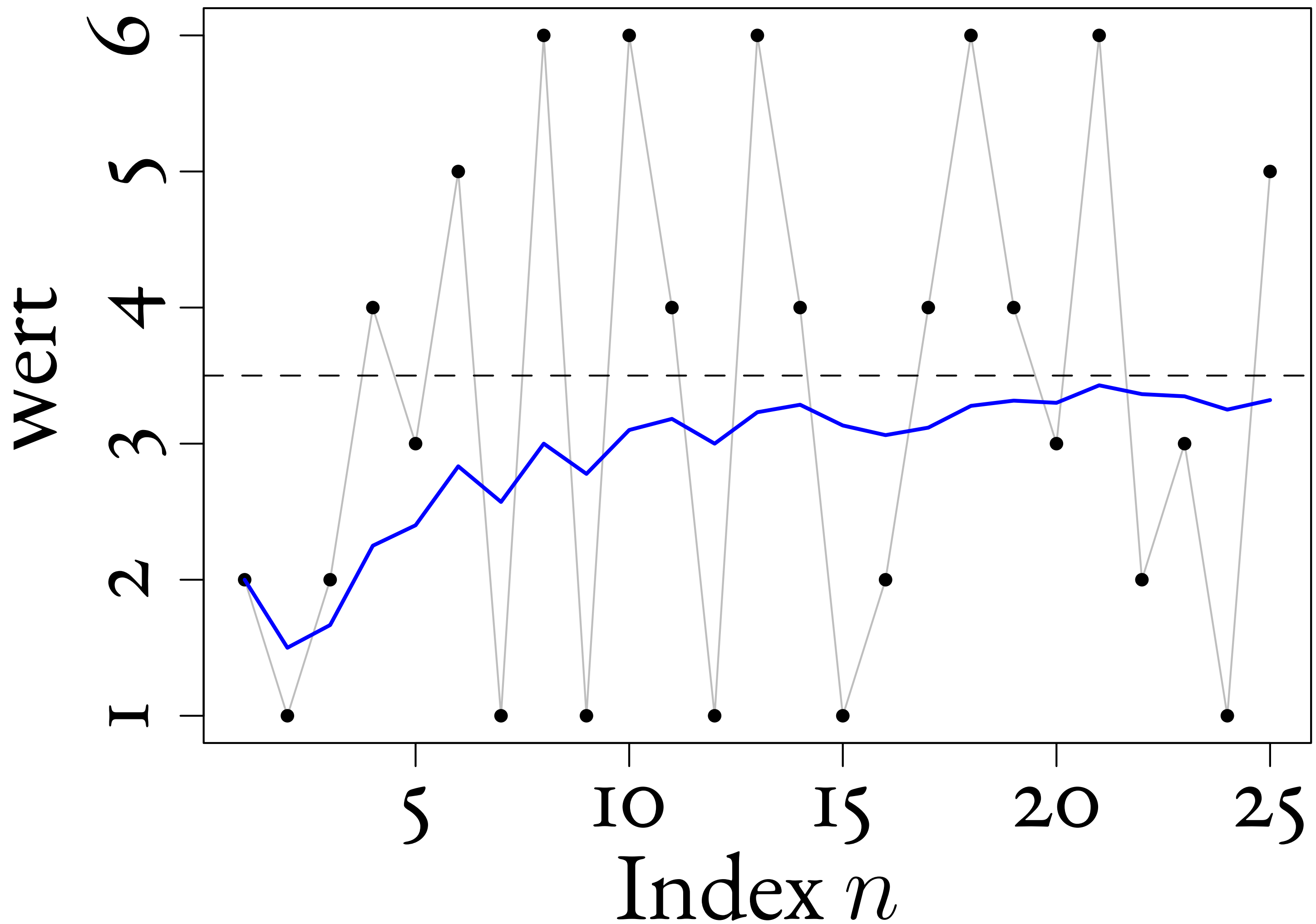


Illustration: $X_1, X_2, \ldots$ uniform auf $\{1, 2, 3, 4, 5, 6\}$,
X_n sind die schwarzen Punkte,
$M_n = (X_1 + X_2 + \cdots + X_n)/n$ die blaue Linie

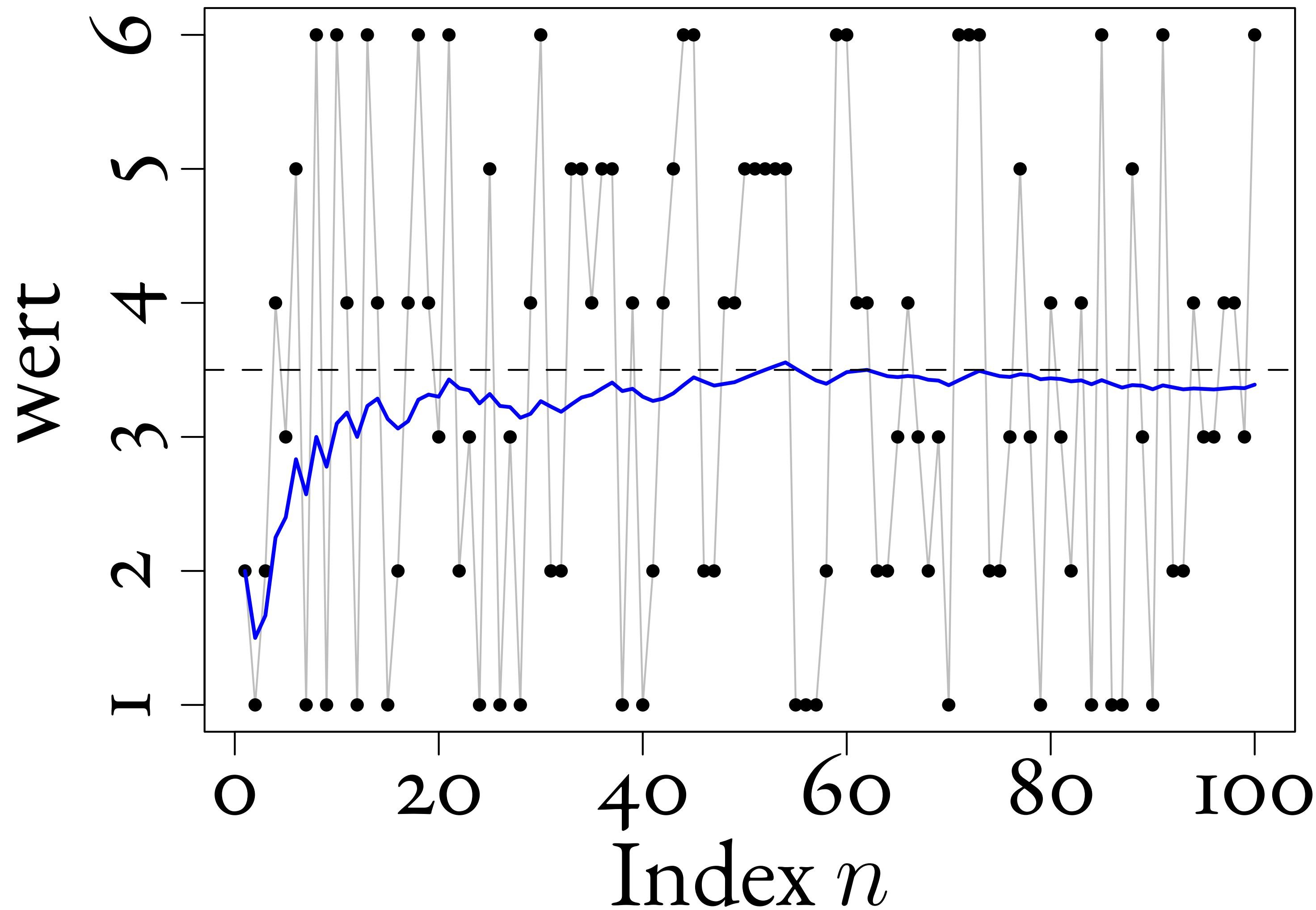
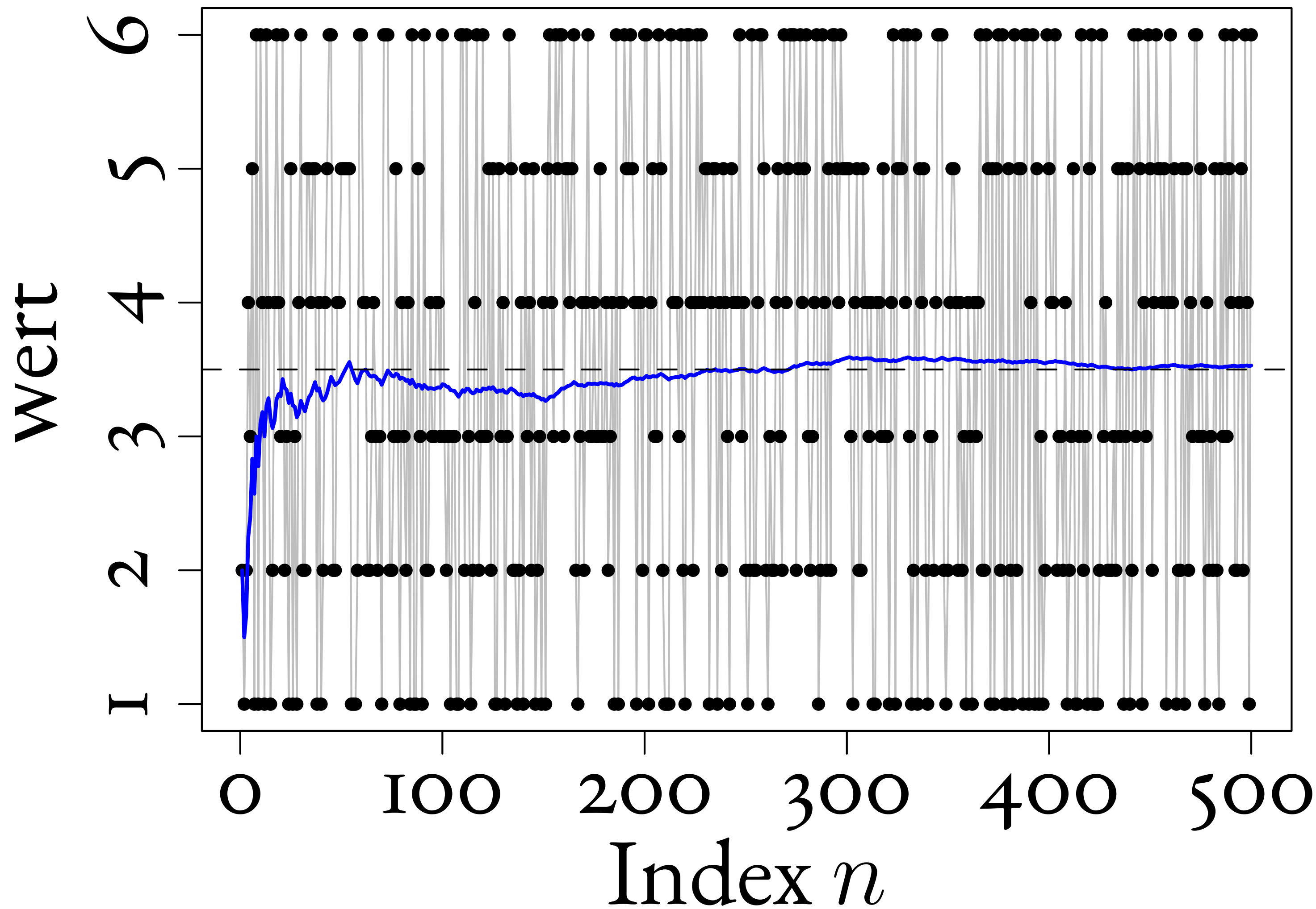


Illustration: $X_1, X_2, \ldots$ uniform auf $\{1, 2, 3, 4, 5, 6\}$,
X_n sind die schwarzen Punkte,
$M_n = (X_1 + X_2 + \cdots + X_n)/n$ die blaue Linie



5. Man kann $\mathbb{E}[X]$ als den erforderlichen Einsatz in einem „fairen Spiel" interpretieren, bei dem man eine zufällige Auszahlung X erhält.

6. Der Erwartungswert ist eine Eigenschaft der Verteilung: $\mathscr{L}(X) = \mathscr{L}(Y)$ impliziert $\mathbb{E}[X] = \mathbb{E}[Y]$.

Beispiel 3.3. 1. A Ereignis, so ist $\mathbb{E}[\mathbf{1}_A] = 1 \cdot \underbrace{P(\mathbf{1}_A = 1)}_{=P(A)} + 0 \cdot P(\mathbf{1}_A = 0) = P(A)$.

2. Sei $X \sim \mathrm{Bin}_{n,p}$, $n \in \mathbb{N}$, $p \in [0,1]$:

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

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=0}^{n} kP(X = k) = \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\ &= np\underbrace{\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-1-(k-1)}}_{=1} = np\mathrm{Bin}_{n-1,p}(\{0,1,\dots,n-1\}) = np\end{aligned}$$

3. Sei $X \sim \mathrm{Geom}_p$, $p \in (0,1]$: $\mathbb{E}[X] = \dfrac{1-p}{p}$

denn $\mathbb{E}[X] = \sum_{n=0}^{\infty} n\,p(1-p)^n = p(1-p)\underbrace{\sum_{n=1}^{\infty} n(1-p)^{n-1}}_{=\frac{1}{p^2}}$

(verwende $f(t) = \sum_{n=0}^{\infty} t^n = \frac{1}{1-t}$ (für $|t|<1$)

erfüllt $f'(t) = \sum_{n=1}^{\infty} n t^{n-1} = \frac{1}{(1-t)^2}$ (— " —))

4. Sei $X \sim \mathrm{Poi}_\alpha$, $\alpha > 0$: $\mathbb{E}[X] = \sum_{n=0}^{\infty} ne^{-\alpha}\dfrac{\alpha^n}{n!} = \alpha\underbrace{\sum_{n=1}^{\infty} e^{-\alpha}\dfrac{\alpha^{n-1}}{(n-1)!}}_{=1} = \alpha$

Satz 3.4 (Rechenregeln für Erwartungswerte). *Seien* $X, Y, X_1, X_2, \ldots, Y_1, Y_2, \ldots \in \mathcal{L}^1(P)$.

1. *(Linearität) Für* $a, b \in \mathbb{R}$ *gilt* $aX + bY \in \mathcal{L}^1(P)$ *und*

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + bE[Y].$$

2. *(Monotonie) Wenn* $X \geq Y$ *(es genügt* $P(X \geq Y) = 1$*), so gilt*

$$\mathbb{E}[X] \geq \mathbb{E}[Y]$$

insbesondere gilt $\mathbb{E}[X] \geq 0$ *für* $X \geq 0$.

3. $P(X \geq 0) = 1$ *und* $\mathbb{E}[X] = 0 \Rightarrow P(X = 0) = 1$.

4. *(Faktorisierung für unabhängige Produkte) Wenn* X *und* Y *unabhängig sind, so ist* $XY \in \mathcal{L}^1(P)$ *und*

$$\mathbb{E}[XY] = \mathbb{E}[X]\, E[Y].$$

$\mathbb{E}[Y]$

5. *(Monotone Konvergenz) Sei* $X_n \nearrow_{n\to\infty} X$, *so gilt*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X];$$

insbesondere für $Y_n \geq 0$, $Y := \sum_{n=1}^{\infty} Y_n$ *gilt* $\mathbb{E}[Y] = \sum\limits_{n=1}^{\infty} \mathbb{E}[Y_n]$ *(möglicherweise als* $\infty = \infty$*).*

Beweis. 1. Z.z.: $aX + bY \in \mathcal{L}^1(P)$ und $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + bE[Y]$ ✓

Beob.: $aX+bY$ ist ebenfalls diskret, Wertebereich $\{ax+by : x \in S_X, y \in S_Y\}$ (ist abzählbar)

$$\sum_z |z| P(aX+bY=z)$$

$$= \sum_{x\in S_X,\, y\in S_Y} \underbrace{|ax+by|}_{\le |a|\cdot|x| + |b||y|} P(X=x, Y=y)$$

$$= |a| \underbrace{\sum_{x,y} |x| P(X=x, Y=y)}_{= \sum_x |x| P(X=x) < \infty} + |b| \underbrace{\sum_{x,y} |y| P(X=x, Y=y)}_{= \sum_y |y| P(Y=y) < \infty} < \infty$$

Analog

$$\sum_z z P(aX+bY=z) = \sum_{x,y} (ax+by) P(X=x, Y=y)$$

$$= a \sum_x x P(X=x) + b \sum_y y P(Y=y) = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

2. Z.z.: $P(X \geq Y) = 1 \Longrightarrow \mathbb{E}[X] \geq \mathbb{E}[Y]$

denn: $\mathbb{E}[X] = \sum_x x P(X=x) = \sum_{x,y} x \underbrace{P(X=x, Y=y)}_{=0 \text{ falls } x<y}$

$\geq \sum_{x,y} y P(X=x, Y=y) = \sum_y y P(Y=y)$

$= \mathbb{E}[Y]$

3. Z.z.: $X \geq 0$ und $\mathbb{E}[X] = 0 \Longrightarrow P(X = 0) = 1$

Ang., es gälte $P(X = x_*) > 0$ für ein $x_* > 0$,

dann wäre $\sum_{x \in S_X} x \underbrace{P(X = x)}_{\geq 0} \geq x_* \underbrace{P(X = x_*)}_{>0}$

$S_X \subset [0, \infty)$

$\lightning$

4. Z.z: X und Y unabhängig $\Longrightarrow XY \in \mathcal{L}^1(P)$ und $\mathbb{E}[XY] = \mathbb{E}[X]\,E[Y]$.

$$\sum_z |z| P(X \cdot Y = z) = \sum_{x,y} \underbrace{|xy|}_{= |x| \cdot |y|} \underbrace{P(X = x, Y = y)}_{= P(X = x) \cdot P(Y = y)}$$

$$= \underbrace{\Big(\sum_x |x| P(X = x)\Big)}_{< \infty} \underbrace{\Big(\sum_y |y| P(Y = y)\Big)}_{< \infty} < \infty$$

dieselbe Rechnung (ohne Beträge) zeigt $\mathbb{E}[X \cdot Y]$

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

5. Betrachte zunächst die Situation $Y_n \geq 0, Y := \sum_{n=1}^{\infty} Y_n$ und zeige $\mathbb{E}[Y] = \sum_{n=1}^{\infty} \mathbb{E}[Y_n]$ (möglicherweise als $\infty = \infty$).

$Y \geq \sum_{n=1}^{N} Y_n$, also

$$\mathbb{E}[Y] \overset{2.}{\geq} \mathbb{E}\Big[\sum_{n=1}^{N} Y_n\Big] \overset{1.}{=} \sum_{n=1}^{N} \mathbb{E}[Y_n]$$

für jedes $N \in \mathbb{N}$, somit $\mathbb{E}[Y] \geq \sum_{n=1}^{\infty} \mathbb{E}[Y_n]$.

Sei $\varepsilon \in (0,1)$,

$\mathbb{N}$

$$\tau := \inf\Big\{N \in \cancel{N} : \sum_{n=1}^{N} Y_n \geq (1-\varepsilon)Y\Big\},$$

n. Vor. ist (auf $\{Y < \infty\}$) $P(\tau < \infty) = 1$. Setze $S := \sum_{n=1}^{\tau} Y_n$.

$$\begin{aligned}
(1-\varepsilon)\mathbb{E}[Y] \leq \mathbb{E}[S] &= \sum_{N,s} sP(S = s, \tau = N) = \sum_{N,s} sP\Big(\tau = N, \sum_{n=1}^{N} Y_n = s\Big) \\
&= \sum_{N} \mathbb{E}\Big[\mathbf{1}_{\{\tau=N\}} \sum_{n=1}^{N} Y_n\Big] = \sum_{N\in\mathbb{N}} \sum_{1\leq n\leq N} \mathbb{E}[Y_n \mathbf{1}_{\{\tau=N\}}] \\
&= \sum_{n\in\mathbb{N}} \sum_{N\geq n} \mathbb{E}[Y_n \mathbf{1}_{\{\tau=N\}}] = \sum_{n\in\mathbb{N}} \sum_{N\geq n} \sum_{y} yP(Y_n = y, \tau = N) \\
&= \sum_{n\in\mathbb{N}} \sum_{y} yP(Y_n = y, \tau \geq n) = \sum_{n\in\mathbb{N}} \mathbb{E}[Y_n \mathbf{1}_{\{\tau\geq n\}}] \leq \sum_{n\in\mathbb{N}} \mathbb{E}[Y_n],
\end{aligned}$$

mit $\varepsilon \downarrow 0$ folgt der 2. Teil der Beh.

Für den ersten Teil der Beh. ($X_n \nearrow X \Longrightarrow \mathbb{E}[X_n] \nearrow \mathbb{E}[X]$) schreibe $Y_n := X_n - X_{n-1}$ (mit $X_0 := 0$), dann verwende obiges. $\square$

und $\sum_{n=0}^{N} Y_n = X_N (- X_0)$ ≥ 0