Beobachtung 3.5 (Erwartungswerte für Kompositionen)**.** X (diskrete) reelle ZV, $g : \mathbb{R} \to \mathbb{R}$, $Y := g(X)$. Dann gilt $Y \in \mathcal{L}^1(P)$ g.d.w. $\sum_x |g(x)| P(X = x) < \infty$ und in diesem Fall ist

$$\mathbb{E}[Y] = \sum_x g(x) P(X = x).$$

d.h. $\sum_{y \in S_Y} |y| P(Y=y) < \infty$

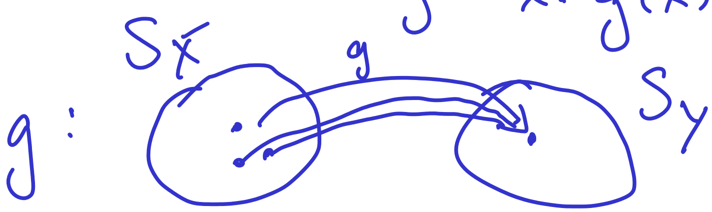
summiere alle $x \in S_X$ (S_X = Werteber. von X)

Denn: (nehme zu. an, dass $\sum_y |y| P(Y=y) < \infty$)

$$\mathbb{E}[Y] = \sum_{y \in S_Y} y P(Y=y) = \sum_{y \in S_X} y \sum_{\substack{x \in S_X : \\ g(x)=y}} P(X=x)$$

$$= \sum_y \sum_{x : g(x)=y} \underbrace{y}_{=g(x)} P(X=x)$$

$$= \sum_y \sum_{x : g(x)=y} g(x) P(X=x) = \sum_{x \in S_X} g(x) P(X=x)$$



Beispiel 3.6. 1. Seien $X_1, \dots, X_n$ u.i.v., $\sim \mathrm{Ber}_p$, so ist $X := X_1 + \dots + X_n \sim \mathrm{Bin}_{n,p}$ und

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n] = np.$$

(Wir hatten den Erwartungswert einer binomialverteilten ZV bereits in Bsp. 3.3, 2. bestimmt, hier kommen wir allerdings ohne explizite Rechnung aus).

2. Sei $X \sim \mathrm{Hyp}_{s,w,k}$ hypergeometrisch verteilt, vgl. Bsp. 1.19. Wir denken an eine Urne mit s schwarzen und w weißen Kugeln, aus der k mal ohne Zurücklegen gezogen wird und notieren

$$A_i = \{i\text{-te gezogene Kugel ist schwarz}\}$$

$$\mathbb{1}_{A_1} + \mathbb{1}_{A_2} + \dots + \mathbb{1}_{A_k} \overset{d}{=} X \quad (\sim \mathrm{Hyp}_{s,w,k})$$

$$P(A_1) = \frac{s}{s+w} = P(A_2) = P(A_3) = \dots = P(A_k)$$

($(\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_k})$ ist austauschbar, vgl. Bsp. 2.7 (Pólya-Urne))

$$\Rightarrow \mathbb{E}[X] = \mathbb{E}[\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_k}] = P(A_1) + \dots + P(A_k)$$

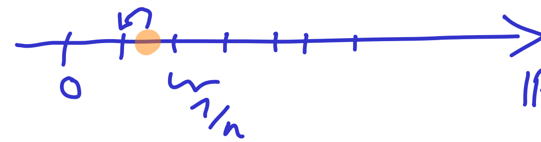
$$= k \cdot \frac{s}{s+w}$$

Bericht 3.7 (Allgemeiner Fall: Erwartungswerte für nicht notwendig diskrete ZV). Sei X eine reelle ZV, setze

$$X_{(n)} := \frac{1}{n}\lfloor nX \rfloor \quad \text{(Diskretisierung von } X \text{ auf das Gitter } \tfrac{1}{n}\mathbb{Z})$$

$\lfloor x \rfloor = \max\{z \in \mathbb{Z} : z \leq x\}$

mit $\lfloor x \rfloor := \max\{z : z \in \mathbb{Z}, z \leq x\}$, offenbar ist $X_{(n)}$ eine diskrete ZV (für jedes $n \in \mathbb{N}$).



Wenn $X_{(n)} \in \mathcal{L}^1$ (für ein, und dann automatisch für alle n) gilt, so existiert

$$\mathbb{E}[X] := \lim_{n\to\infty} \mathbb{E}[X_{(n)}]$$

und die Rechenregeln aus Satz 3.4 gelten ebenso.

Beweisidee. Wir betrachten hier nur eine Skizze, siehe z.B. [G, Kap. 4.1.2] (oder die Vorlesung Stochastik I) für das „volle Argument“.

Es ist $X_{(n)} \leq X < X_{(n)} + \frac{1}{n}$, also $X_{(m)} < X_{(n)} + \frac{1}{n}$ und $X_{(n)} < X_{(m)} + \frac{1}{m}$ für alle $m, n \in \mathbb{N}$.

Demnach

$$|X_{(n)}| \leq |X_{(m)}| + \Big(\frac{1}{m} \vee \frac{1}{n}\Big),$$

insbesondere gilt $X_{(n)} \in \mathcal{L}^1$ g.w.d. $X_{(m)} \in \mathcal{L}^1$ und

$$\mathbb{E}\big[|X_{(n)}|\big] \leq \mathbb{E}\big[|X_{(m)}|\big] + \frac{1}{m}, \quad \mathbb{E}\big[|X_{(m)}|\big] \leq \mathbb{E}\big[|X_{(m)}|\big] + \frac{1}{n},$$

(dasselbe ohne Beträge, d.h. $\mathbb{E}[X_{(n)}] \leq \mathbb{E}[X_{(m)}] + \frac{1}{m}$

also

$$\Big|\mathbb{E}\big[|X_{(n)}|\big] - \mathbb{E}\big[|X_{(m)}|\big]\Big| \leq \Big(\frac{1}{m} \vee \frac{1}{n}\Big),$$

d.h. $\big(\mathbb{E}[X_{(n)}]\big)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge, somit existiert ihr Grenzwert.

Um die Rechenregeln aus Satz 3.4 auf den allgemeinen Fall zu übertragen, muss man jeweils prüfen, dass sie mit der Grenzwertbildung in der diskreten Approximation verträglich sind. □

Satz 3.8 (Jensen'sche Ungleichung[1]). *Sei $\varphi : \mathbb{R} \to \mathbb{R}$ konvex, $X \in \mathcal{L}^1$ und $\varphi(X) \in \mathcal{L}^1$, dann gilt*

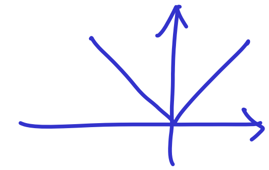
$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$$

insbesondere gilt für $X \in \mathcal{L}^1$

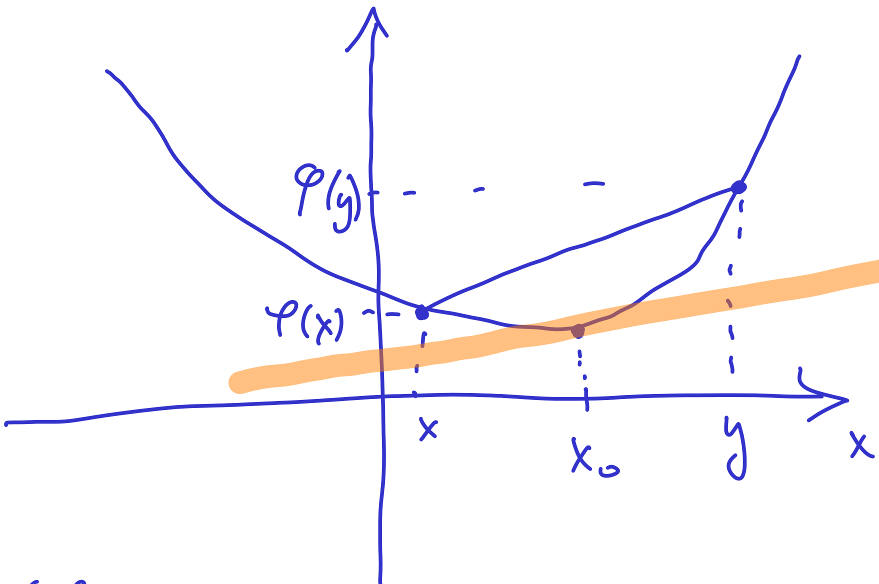
$$|\mathbb{E}[X]| \leq \mathbb{E}[|X|] \quad \text{und} \quad (\mathbb{E}[X])^2 \leq \mathbb{E}[X^2]$$

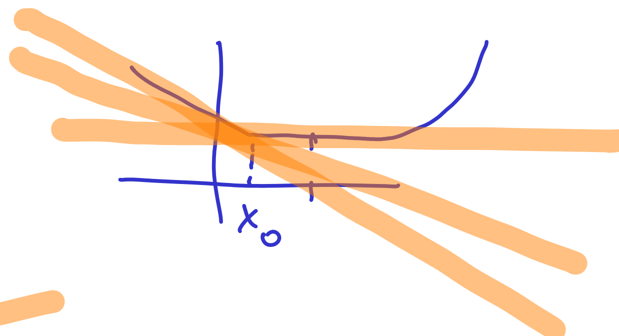
(letzeres möglicherweise im Sinne von $(\mathbb{E}[X])^2 \leq \infty$).

(bedr. $x \mapsto |x|$



$x \mapsto x^2$)





Bew.:
Zu $x_0 \in \mathbb{R}$ gibt es ein $a \, (= a(x_0))$,
so dass
$$\varphi(x_0) + a(x - x_0) \leq \varphi(x)$$
für alle $x \in \mathbb{R}$ gilt,
insbesondere:
$$\varphi(x_0) + a(X - x_0) \leq \varphi(X)$$

(φ konvex $:\Leftrightarrow$
$\forall x, y \in \mathbb{R}, \ \alpha \in [0,1]$
$$\varphi(\alpha x + (1-\alpha) y) \leq \alpha \varphi(x) + (1-\alpha)\varphi(y) \,)$$

$$\Rightarrow \quad \varphi(x_0) + a(\mathbb{E}[X] - x_0) \leq \mathbb{E}[\varphi(X)],$$

wähle $x_0 = \mathbb{E}[X]$

[1] nach Johan Ludvig Jensen, 1859–1925 benannt

3.2 Der Fall mit Dichte

Definition 3.9. X reelle ZV mit Dichte f_X, dann ist

$$X \in \mathcal{L}^1 \quad :\Leftrightarrow \quad \int_{\mathbb{R}} |x| f_X(x)\, dx < \infty$$

$$\int_{-\infty}^{\infty} |x| f_X(x) dx$$

(man sagt dann, X besitzt einen Erwartungswert) und

$$\mathbb{E}[X] := \int_{\mathbb{R}} x f_X(x)\, dx.$$

Dichte $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$\frac{d}{dx} e^{-\frac{x^2}{2}} = e^{-\frac{x^2}{2}} \left(-\frac{1}{2} \cdot 2x\right) = -x e^{-x^2/2}$$

Beispiel 3.10. 1. $X \sim \mathcal{N}_{0,1}$ hat $\mathbb{E}[X] = 0$.

$$\int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \underbrace{\left[-e^{-\frac{x^2}{2}}\right]_0^{\infty}}_{-0-(-1)=1} < \infty$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \left[e^{-\frac{x^2}{2}}\right]_{-\infty}^{\infty} = 0$$

Bem.:

$X \sim \mathcal{N}_{0,1}$, so ist $X + \mu \sim \mathcal{N}_{\mu,1}$ und hat $\mathbb{E}[X+\mu] = \mu$

2. Die Gammaverteilung $\Gamma_{a,\nu}$ $(a, \nu > 0)$ hat Dichte

$$\frac{a^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-ax} \mathbf{1}_{(0,\infty)}(x)$$

(mit $\Gamma(\nu) = \int_0^\infty x^{\nu-1} e^{-x}\, dx$).
Für $X \sim \Gamma_{a,\nu}$ ist $\mathbb{E}[X] = \nu/a$.

$$\mathbb{E}[X] = \int_0^\infty x \cdot \frac{a^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-ax}\, dx$$

$$= \frac{\Gamma(\nu+1)}{a\Gamma(\nu)} \cdot \underbrace{\int_0^\infty \frac{a^{\nu+1}}{\Gamma(\nu+1)} x^{(\nu+1)-1} e^{-ax}\, dx}_{=1} = \frac{\overbrace{\Gamma(\nu+1)}^{=\nu\cdot\cancel{\Gamma(\nu)}}}{\cancel{\Gamma(\nu)}} \cdot \frac{1}{a}$$

3. Die Cauchy-Verteilung mit Dichte $\frac{1}{\pi}\frac{1}{1+x^2}$ besitzt keinen Erwartungswert:

$$\int_{-\infty}^\infty \frac{1}{\pi} \underbrace{\frac{|x|}{1+x^2}}_{\approx 1/|x|}\, dx = \frac{2}{\pi} \int_0^\infty \frac{1}{\pi} \frac{x}{1+x^2}\, dx = \left[\frac{1}{\pi} \log\left(1 + x^2\right)\right]_0^\infty = \infty.$$

Bericht 3.11. 1. (Zur Herleitung des Falls mit Dichte aus der allgemeinen Situation aus Bericht 3.7)

$X_{(n)} = \frac{1}{n}\lfloor nX \rfloor$ nimmt den Wert $\frac{k}{n}$, $k \in \overset{\mathbb{Z}}{\cancel{\mathbb{N}}}$ an mit

$$P\Big(X_{(n)} = \frac{k}{n}\Big) = \int_{k/n}^{(k+1)/n} f_X(x)\, dx,$$

$= P\left(\frac{k}{n} \le X < \frac{k+1}{n}\right)$,

X besitze die Dichte f_X

also ist (sofern die Reihe absolut konvergiert, was man analog überprüft)

$$\begin{aligned}\mathbb{E}\big[X_{(n)}\big] &= \sum_{k\in\mathbb{Z}} \frac{k}{n} \int_{k/n}^{(k+1)/n} f_X(x)\, dx \\ &= \int_{\mathbb{R}} \frac{1}{n}\lfloor nx \rfloor f_X(x)\, dx \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} x f_X(x)\, dx\end{aligned}$$

(die Konvergenz folgt, falls $x f_X(x)$ [uneigentlich] Riemann-integrierbar ist, aus der Riemann-Approximation, für den allgemeinen Fall mit messbarem f_X aus dem Konvergenzsatz von Lebesgue).

2. (Analogon zu Beob. 3.5 im Fall mit Dichte)
Sei $X = (X_1, \dots, X_d)$ $\mathbb{R}^d$-wertig mit Dichte $f_X : \mathbb{R}^d \to [0, \infty]$, $g : \mathbb{R}^d \to \mathbb{R}$, $Y := g(X)$. Dann gilt $Y \in \mathcal{L}^1$ g.d.w.

$$\int_{\mathbb{R}^d} |g(x_1, \dots, x_d)| f_X(x_1, \dots, x_d)\, dx_1 \dots dx_d < \infty$$

und in diesem Fall

$$\mathbb{E}[Y] = \int_{\mathbb{R}^d} g(x_1, \dots, x_d) f_X(x_1, \dots, x_d)\, dx_1 \dots dx_d < \infty.$$

(Siehe z.B. [G, Korollar 4.13])

3.3 Varianz und Kovarianz

Definition 3.12. X reelle ZV, $p > 0$. X besitzt p-tes Moment (auch geschrieben $X \in \mathcal{L}^p$), wenn $\mathbb{E}\big[|X|^p\big] < \infty$. $\mathbb{E}\big[X^p\big]$ heißt das p-te Moment von X.

Bemerkung 3.13. Für $p \geq p' \geq 1$ ist $\mathcal{L}^p \subset \mathcal{L}^{p'}$ (denn $|x|^{p'} \leq 1 + |x|^p$).

Definition 3.14. Für $X, Y \in \mathcal{L}^2$ heißt

1. $\mathrm{Var}[X] := \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}\big[X^2\big] - \big(\mathbb{E}[X]\big)^2$ die *Varianz* von X (manchmal schreibt man auch $\sigma_X^2 := \mathrm{Var}[X]$),

   $\sqrt{\mathrm{Var}[X]}$ die *Standardabweichung* (oder *Streuung*) von X (manchmal auch $\sigma_X = \sqrt{\sigma_X^2}$ geschrieben),

2. $\mathrm{Cov}[X, Y] := \mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ die *Kovarianz* von X und Y.

X und Y heißen *unkorreliert*, wenn $\mathrm{Cov}[X, Y] = 0$.