Beispiele. • **Poisson Verteilung.** Für $X \sim \mathcal{P}_\lambda$ mit $\lambda > 0$ haben wir $\phi_X(s) = e^{\lambda(s-1)}$, und damit die faktoriellen Momente für alle $k \geq 1$

$$\mathbb{E}[X_{[k]}] = \frac{d^k}{ds^k} e^{\lambda(s-1)}\Big|_{s=1} = \lambda^k e^{\lambda(s-1)}\Big|_{s=1} = \lambda^k \ .$$

Damit bestätigen wir $\mathbb{E}[X] = \lambda$ und $\mathbb{V}[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$.

• Sei $X \in \mathbb{N}$ eine Zufallsvariable mit Verteilungsdichte $\rho(k) = \frac{6}{\pi^2 k^2}$. Dann ist $X \notin \mathcal{L}^1$ und $\mathbb{E}[X] = \infty$. Für die erzeugende Funktion

$$\phi_X(s) = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{s^k}{k^2} \quad \text{ist dann} \quad \phi_X'(1) = \infty \ .$$

• **Geometrische Verteilung.** Für $X \sim \mathcal{G}_p$ mit $p \in (0,1)$ haben wir $\phi_X(s) = \frac{ps}{1-(1-p)s}$. Nach **mühseliger** Rechnung ergibt sich

$$\phi_X'(s) = \frac{p}{(1-(1-p)s)^2} \quad \text{und} \quad \phi_X''(s) = \frac{2p(1-p)}{(1-(1-p)s)^3}$$

und damit $\mathbb{E}[X] = \phi_X'(1) = \frac{1}{p}$ und $\mathbb{V}[X] = \frac{2p(1-p)}{p^3} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$.

Im letzten Beispiel sehen wir, dass die Berechnung von Momenten mittels erzeugenden Funktionen nicht immer vorteilhaft ist, aber das nächste Resultat ist mit Abstand die nützlichste Eigenschaft.

Satz 5.13. Erzeugende Funktionen von Summen.
(a) Sind X, Y unabängige $\mathbb{N}_0$-wertige Zufallsvariablen, so gilt

$$\phi_{X+Y}(s) = \phi_X(s)\phi_Y(s) \ , \quad 0 \leq s \leq 1 \ .$$

(b) Sind $X_1, X_2, \ldots$ iid[a] auf $\mathbb{N}_0$ mit erzeugender Funktion ϕ_X und N eine weitere unabängige $\mathbb{N}_0$-wertige Zufallsvariable, dann gilt für die Summe mit zufällig vielen Termen

$$Z := \sum_{k=1}^{N} X_k \ , \quad \text{dass} \quad \phi_Z(s) = \phi_N(\phi_X(s)) \ .$$

[a] unabhängig, identisch verteilt

$\phi_X(s) = \mathbb{E}[s^X]$

Beweis. (a) Mit Satz 4.8 sind s^X und s^Y für alle $s \in [0,1]$ unabhängig, und damit

$$\phi_{X+Y}(s) = \mathbb{E}[s^{X+Y}] = \mathbb{E}[s^X s^Y] = \mathbb{E}[s^X]\,\mathbb{E}[s^Y] = \phi_X(s)\phi_Y(s) \ .$$

(b) Mit (a) und $\phi_{X_i} = \phi_X(s)$ für alle $i \geq 1$ haben wir

$$\begin{aligned}\phi_Z(s) &= \mathbb{E}\big[s^{X_1+\ldots+X_N}\big] = \sum_{n\geq 0} \mathbb{P}[N=n]\,\mathbb{E}[s^{X_1} \cdots s^{X_n}] \\ &= \sum_{n \geq 0} \mathbb{P}[N=n]\,\phi_X(s)^n = \mathbb{E}\big[(\phi_X(s))^N\big] = \phi_N(\phi_X(s)) \ .\end{aligned}$$

□

Bemerkungen.

- Im Beweis von (b) benutzen wir einen wichtigen Schritt basierend auf dem Satz von Fubini zur Vertauschung der Summations/Integrationsreihenfolge, den wir nochmal kurz zusammenfassen: Sei $f(X, N)$ eine Funktion von zwei unabhängigen diskreten Zufallsvariablen X und N mit $\mathbb{E}[|f(X, N)|] < \infty$, dann ist

$$\mathbb{E}[f(X,N)] = \sum_n \sum_x \underbrace{\mathbb{P}[N = n, X = x]}_{=\mathbb{P}[N=n]\,\mathbb{P}[X=x]} f(x,n) = \sum_n \mathbb{P}[N = n]\,\mathbb{E}[f(X,n)] \ .$$

  Eine analoge Formel gilt auch im kontinuierlichen Fall mit Dichtefunktion $\rho_N(y)$ für N,

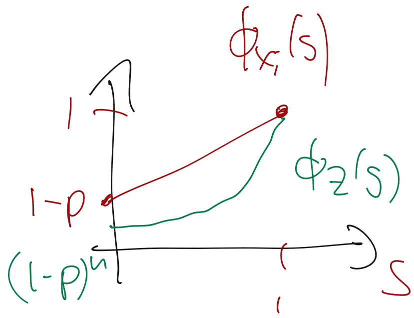
$$\mathbb{E}[f(X,N)] = \int_{\mathbb{R}} \rho_N(y)\mathbb{E}[f(X,y)]\,dy \ .$$

- Die Berechnung der erzeugenden Funktion von Summen unabhängiger Zufallsvariablen ist in der Regel viel einfacher als die Faltung der Verteilungen zu berechnen (siehe Lemma 4.9). Eine analoge Aussage zu Satz 5.13(a) gilt natürlich auch für erzeugende Funktionen von Faltungen. Sind Q_1 und Q_2 Verteilungen auf $\mathbb{N}_0$, dann ist

$$\phi_{Q_1 \star Q_2}(s) = \phi_{Q_1}(s)\,\phi_{Q_2}(s) \ .$$

- Ist die Zahl der Summanden in Satz 5.13(b) deterministisch, d.h. $\mathbb{P}[N = n] = 1$ für ein $n \in \mathbb{N}$, dann ist

$$\phi_N(s) = s^n \quad \text{und} \quad \phi_Z(s) = \phi_X(s)^n \ .$$



Beispiele. • **Binomial Verteilung.** Seien $X_1, \dots, X_n \sim \mathcal{B}_{1,p}$ iid Bernoulli Zufallsvariablen, dann ist die Summe $Z = \sum_{i=1}^n X_i \sim \mathcal{B}_{n,p}$ Binomial verteilt. Für die erzeugende Funktion ergibt sich konsistenterweise

$$\phi_{X_i}(s) = 1 - p + ps \quad \text{und} \quad \phi_Z(s) = (1 - p + ps)^n = \sum_{k=0}^{n} s^k \binom{n}{k} p^k (1-p)^{n-k} \ .$$

- **Die geometrische Summe geometrischer Variablen ist geometrisch.** Seien $X_1, X_2, \ldots \sim \mathcal{G}_p$ und $Y \sim \mathcal{G}_q$ unabhängige geometrisch verteilte Zufallsvariablen mit $p, q \in (0, 1]$. Dann ist

$$Z = \sum_{k=1}^{Y} X_i \sim \mathcal{G}_{pq} \ ,$$

$X_1, X_2 \sim \mathcal{G}_p \qquad X_1 + X_2 \not\sim \mathcal{G}$

  denn mit $\phi_X(s) = \frac{ps}{1-(1-p)s}$ ist

$$\phi_Z(s) = \phi_Y(\phi_X(s)) = \frac{q\frac{ps}{1-(1-p)s}}{1-(1-q)\frac{ps}{1-(1-p)s}} = \frac{qps}{1-(1-p)s-(1-q)ps} = \frac{qps}{1-(1-qp)s} \ .$$

$\Rightarrow Z \sim \mathcal{G}_{qp}$

$\mathbb{E}[Z] = \frac{1}{qp} = \frac{1}{q}\frac{1}{p}$

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

- **Poisson Verteilung.** Für $X_i \sim \mathcal{P}_{\lambda_i}$, $i = 1, \ldots, n$, mit $\lambda_i > 0$ haben wir für $Z := \sum_{i=1}^n X_i$

$$\phi_Z(s) = \prod_{i=1}^{n} \phi_{X_i}(s) = \prod_{i=1}^{n} e^{\lambda_i(s-1)} = e^{\sum_{i=1}^n \lambda_i (s-1)} \ ,$$

  d.h. $Z \sim \mathcal{P}_{\lambda_1+..+\lambda_n}$, wie wir in Korollar 4.10 schon gesehen haben.
  Summen unabängiger Poisson Variablen sind also Poisson, umgekehrt kann man auch

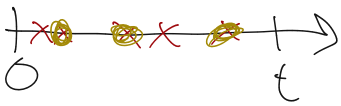
$\phi_{X_i}(s) = e^{\lambda_i(s-1)}$

$X_1 - X_2 \sim \mathcal{P}_{?}$ $\qquad \mathbb{N}_0$

$1 - 2 = -1$ nein

$\mathcal{P}_{\alpha t}$

Poisson Variablen "ausdünnen", wie im nächsten Ergebnis beschrieben.



Lemma 5.14. Ausdünnung von Poisson Variablen.
Seien $X_1, X_2, \ldots \sim \mathcal{B}_{1,p}$, $p \in [0,1]$ iid Bernoulli Variablen, und $N \sim \mathcal{P}_\lambda$, $\lambda > 0$ eine weitere unabhängige Poisson Variable. Dann ist

$X_i \in \{0,1\}$

$$Z = \sum_{i=1}^{N} X_i \sim \mathcal{P}_{p\lambda} \quad \text{Poisson verteilt mit Parameter } p\lambda\,.$$

$p \in (0,1)$

$X \sim \mathcal{P}_{\alpha t}$

Beweis. folgt mit Hilfe erzeugender Funktionen (siehe Übungsblatt).

$\mathbb{E}[Z] = \lambda \underbrace{\mathbb{E}[X_i]}_{p} = \lambda p$

$\emptyset \sim \mathcal{P}_{p\alpha t}$

Lemma 5.15. Wald'sche Identitäten.
Seien $X_1, X_2, \ldots$ iid reellwertige Zufallsvariablen und $N \in \mathbb{N}_0$ unabhängig, alle in $\mathcal{L}^2$. Dann gilt für $S_N = \sum_{i=1}^N X_i$

$$\mathbb{E}[S_N] = \mathbb{E}[N]\,\mathbb{E}[X_1] \quad \text{und} \quad \mathbb{V}[S_N] = \mathbb{E}[N]\mathbb{V}[X_1] + \mathbb{V}[N]\mathbb{E}[X_1]^2\,.$$

Beweis. Für $X_i \in \mathbb{N}_0$ kann das Resultat auch relativ einfach aus Satz 5.13(b) hergeleitet werden (siehe Übungsblatt). Im allgemeinen Fall ist

$$\mathbb{E}[S_N] = \sum_{n=0}^{\infty} \mathbb{P}[N=n]\,\mathbb{E}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{n=0}^{\infty} \mathbb{P}[N=n]\,n\mathbb{E}[X_1] = \mathbb{E}[N]\,\mathbb{E}[X_1]\,.$$

Da die Summanden u.i.v. sind ist

$$\mathbb{V}[S_n] = \mathbb{E}[S_n^2] - n^2\mathbb{E}[X_1]^2 = n\mathbb{V}[X_1] \quad \text{für alle } n \in \mathbb{N}\,,$$

und damit

$\mathbb{E}[S_n]^2 = n\mathbb{V}[X_1] + n^2\mathbb{E}[X_1]^2$

$$\begin{aligned}\mathbb{E}[S_N^2] &= \sum_{n=0}^{\infty} \mathbb{P}[N=n]\,\mathbb{E}[S_n^2] = \sum_{n=0}^{\infty} \mathbb{P}[N=n]\big(n\mathbb{V}[X_1] + n^2\mathbb{E}[X_1]^2\big)\\ &= \mathbb{E}[N]\mathbb{V}[X_1] + \mathbb{E}[N^2]\mathbb{E}[X_1]^2\,.\end{aligned}$$

$- \mathbb{E}[S_N]^2 = \mathbb{E}[N]^2\mathbb{E}[X_1]^2$

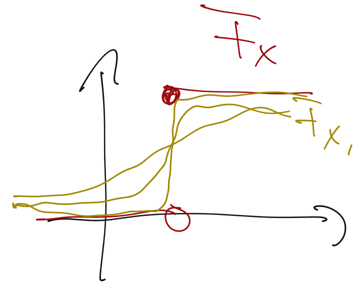
Wir ziehen $\mathbb{E}[S_N]^2 = \mathbb{E}[N]^2\mathbb{E}[X_1]^2$ ab, und erhalten das Ergebnis. □

Bemerkung. Im deterministischen Fall mit $\mathbb{P}[N=n] = 1$ ist $\mathbb{E}[N] = n$ und $\mathbb{V}[N] = 0$, und wir erhalten die zuvor hergeleiteten Formeln. Allgemein wächst die Varianz von S_N mit der von N.

6 Fundamentale Grenzwertsätze (W7-8)

In diesem Kapitel besprechen wir das Gesetz der großen Zahlen und den zentralen Grenzwertsatz, und führen zuvor noch relevante Konvergenzbegriffe für Zufallsvariablen ein.

6.1 Konvergenzbegriffe



Definition 6.1. Eine Folge $X_1, X_2, \ldots \in \mathbb{R}$ von Zufallsvariablen **konvergiert in Verteilung** gegen $X \in \mathbb{R}$, falls die Verteilungsfunktionen

$$F_{X_n}(x) \to F_X(x) \quad \text{für alle } x \in \mathbb{R} \text{ an denen } F_X \text{ stetig ist ,}$$

und wir schreiben $X_n \xrightarrow{d} X$.

Dieser Konvergenzbegriff hängt **nur von den Verteilungen** der Zufallsvariablen ab, und ist unabhängig von den zugrundeliegenden Wahrscheinlichkeitsräumen, auf denen die Variablen definiert sind. Wir sind der Konvergenz in Verteilung schon öfter begegnet:

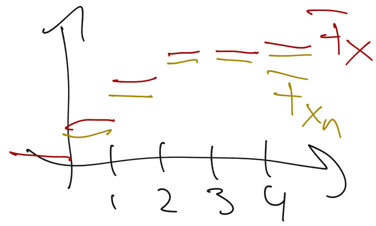
Beispiele.

- Wir erinnern uns an das Beispiel in Satz 2.6: $X_n \sim \mathcal{B}_{n,p_n}$ mit $p_n \to 0$ und $np_n \to \lambda > 0$. Dann konvergiert für alle $k \in \mathbb{N}_0$

  $\mathbb{E}\{X_n\} = n \cdot p_n \to \lambda$

  $$\mathbb{P}[X_n = k] = \underbrace{\binom{n}{k} p_n^k}_{\to \lambda^k / k!} \underbrace{(1 - p_n)^{n-k}}_{\to e^{-\lambda}} \to \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{wenn } n \to \infty ,$$

  d.h. die Zähldichten $\rho_{X_n}(k) \to \rho_X(k)$ konvergieren punktweise. Da die Verteilungsfunktionen eindeutig durch die Zähldichten bestimmt sind, gilt auch $F_{X_n}(x) \to F(x) = \sum_{k \le x} \rho_X(k)$, und damit $X_n \xrightarrow{d} X \sim \mathcal{P}_\lambda$.



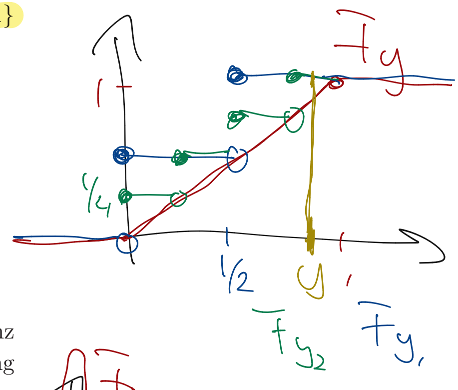
- Im Beweis zu Satz 4.13 haben wir im Rahmen der Binärdarstellung reeller Zahlen in $[0,1]$ gezeigt, dass für unabhängige Bernoulli Variablen $X_1, X_2, \ldots \sim \mathcal{B}_{1,1/2}$ in $\{0,1\}$ gilt

  $$Y_n := \sum_{k=1}^{n} X_k 2^{-k} \to \underbrace{Y}_{\in [0,1]} = \sum_{k=1}^{\infty} X_k 2^{-k} \sim \mathcal{U}_{[0,1]} . \qquad n \to \infty$$

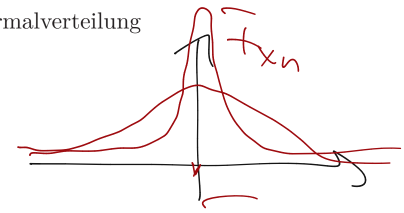
  Für die Verteilungsfunktion haben wir für alle $y \in [0,1]$

  $$\mathbb{P}[Y_n \le y] = \frac{1}{2^n} \left| \{ m \in \mathbb{N}_0 : \frac{m}{2^n} \le y \} \right| = \frac{\lfloor 2^n y \rfloor + 1}{2^n} \to y \quad \text{für } n \to \infty ,$$

  also hält diese Konvergenz in Verteilung.



- Seien $X_1, X_2, \ldots$ normalverteilte Zufallsvariablen mit $X_n \sim \mathcal{N}_{0,1/n}$, d.h. die Varianz konvergiert gegen 0. Dann gilt mit der Skalierungseigenschaft der Normalverteilung $X_n \sim Z/\sqrt{n}$ mit $Z \sim \mathcal{N}_{0,1}$, und damit



  $$\mathbb{P}[X_n \le x] = \mathbb{P}[Z \le x\sqrt{n}] \to \begin{cases} 0 & , \ x < 0 \\ 1/2 & , \ x = 0 \\ 1 & , \ x > 0 \end{cases} .$$

  Dies impliziert $X_n \xrightarrow{d} \underbrace{0}_{= X}$, denn die Verteilungsfunktion der "Zufallsvariablen" 0 ist die Stufenfunktion $\mathbb{1}_{[0,\infty)}(x)$, die nur für alle $x \neq 0$ stetig ist.

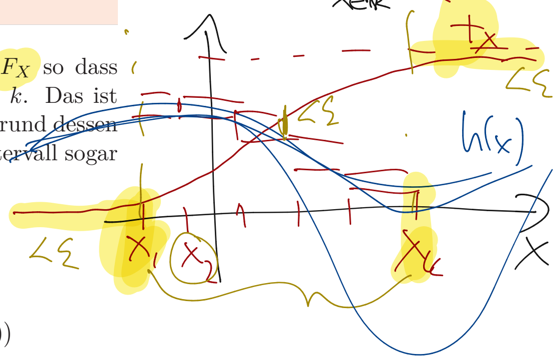
Satz 6.1. Kriterien für Verteilungskonvergenz. Die folgenden Aussagen sind äquivalent für eine Folge $(X_n : n \in \mathbb{N})$ von Zufallsvariablen, wenn $n \to \infty$:

(a) $X_n \xrightarrow{d} X$,

(b) $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$ für alle stetigen, beschränkten Funktionen $h \in C_b(\mathbb{R})$, $= \{h : \mathbb{R} \to \mathbb{R} \mid \text{stetig}, \|h\|_\infty < \infty\}$

$\|h\|_\infty = \sup_{x\in\mathbb{R}} |h(x)|$

(c) $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$ für alle $h \in C_b^k(\mathbb{R})$ mit festem $k \in \mathbb{N}$, d.h. k mal stetig differenzierbar mit beschränkten Ableitungen.



Beweis. (a)⇒(b): Zu $\epsilon > 0$ wählen wir Stetigkeitsstellen $x_1 < \ldots < x_k$ von F_X so dass $F_X(x_1) < \epsilon$, $F_X(x_k) > 1 - \epsilon$ und $|h(x) - h(x_i)| < \epsilon$ für $x_{i-1} \le x \le x_i$, $1 < i \le k$. Das ist möglich für k groß genug wegen der Limeseigenschaften und Monotonie von F_X, aufgrund dessen es nur abzählbar viele Unstetigkeitsstellen gibt, und weil h auf jedem kompakten Intervall sogar gleichmäßig stetig ist. Dann gilt

$$\mathbb{E}[h(X_n)] = \sum_{i=2}^{k} \mathbb{E}[h(X_n) \mathbb{1}_{(x_{i-1},x_i]}(X_n)] + \mathbb{E}[h(X_n) \mathbb{1}_{(-\infty,x_1]\cap(x_k,\infty)}(X_n)]$$

$\mathbb{E}[\ ] = \mathbb{P}[X_n \in (x_{i-1}, x_i]]$

$$\le \sum_{i=2}^{k} (h(x_i) + \epsilon)[F_{X_n}(x_i) - F_{X_n}(x_{i-1})] + \|h\|_\infty (F_{X_n}(x_1) + 1 - F_{X_n}(x_k))$$

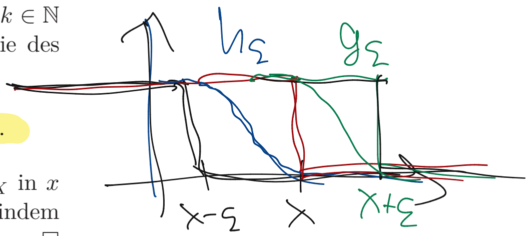
$$\stackrel{(a)}{\to} \sum_{i=2}^{k} (h(x_i) + \epsilon)[F_X(x_i) - F_X(x_{i-1})] + 2\epsilon\|h\|_\infty = \mathbb{E}[h(X)] + O(2\epsilon(1 + 2\|h\|_\infty))$$

für $n \to \infty$, da $F_{X_n}(x_i) \to F_X(x_i)$ für alle i konvergiert. Der Grenzwert stimmt nach Konstruktion bis auf $2\epsilon(1 + 2\|h\|_\infty)$ mit $\mathbb{E}[h(X)]$ überein. Für $\epsilon \to 0$ folgt $\limsup_{n\to\infty} \mathbb{E}[h(X_n)] \le \mathbb{E}[h(X)]$. Mit derselben Ungleichung für $-h$ ergibt sich (b).

$\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]$

(b)⇒(c) ist klar, da $C_b^k(\mathbb{R}) \subset C_b(\mathbb{R})$ für alle $k \in \mathbb{N}$.

(c)⇒(a): Sei $x \in \mathbb{R}$ eine beliebige Stetigkeitsstelle von F_X und $\epsilon > 0$ beliebig. Für jedes $k \in \mathbb{N}$ können wir $h_\epsilon \in C_b^k(\mathbb{R})$ wählen, so dass $\mathbb{1}_{(-\infty,x-\epsilon]} \le h_\epsilon \le \mathbb{1}_{(-\infty,x]}$, und mit Monotonie des Erwartungswerts folgt



$$F_{X_n}(x) = \mathbb{E}[\mathbb{1}_{(-\infty,x]}(X_n)] \ge \mathbb{E}[h_\epsilon(X_n)] \to \mathbb{E}[h_\epsilon(X)] \ge \mathbb{E}[\mathbb{1}_{(-\infty,x-\epsilon]}(X)] = F(x - \epsilon) \ .$$

Somit ist $\liminf_{n\to\infty} F_{X_n}(x) \ge F_X(x - \epsilon)$, und mit $\epsilon \to 0$ folgt wegen Stetigkeit von F_X in x dass $\liminf_{n\to\infty} F_{X_n}(x) \ge F(x)$. Völlig analog zeigen wir $\limsup_{n\to\infty} F_{X_n}(x) \le F(x)$, indem wir $g_\epsilon \in C_{b,\infty}(\mathbb{R})$ wählen mit $\mathbb{1}_{(-\infty,x]} \le g_\epsilon \le \mathbb{1}_{(-\infty,x+\epsilon]}$. □

Es gibt auch noch andere Klassen von Testfunktion h die wie in Satz 6.1 die Verteilungskonvergenz charakterisieren, wie z.B. gleichmäßig stetige Funktionen, oder auch charakteristische Funktionen, die wir später noch einführen. Ist eine Folge von Zufallsvariablen auf demselben Wahrscheinlichkeitsraum definiert, gibt es auch noch stärkere Konvergenzbegriffe.

Definition 6.2. Seien $X, X_1, X_2, \ldots \in \mathbb{R}$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Wir sagen, dass

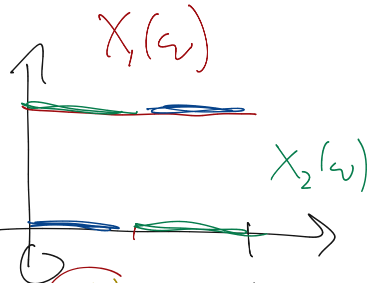
- $(X_n : n \in \mathbb{N})$ konvergiert **($\mathbb{P}$-)fast sicher** gegen X, geschrieben $X_n \xrightarrow{f.s.} X$, wenn

$$\mathbb{P}[X_n \to X] = \mathbb{P}\Big[\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\Big] = 1 \ .$$

- $(X_n : n \in \mathbb{N})$ konvergiert **stochastisch** (oder **in Wahrscheinlichkeit**) gegen X, geschrieben $X_n \xrightarrow{\mathbb{P}} X$, wenn

$$\lim_{n\to\infty} \mathbb{P}[|X_n - X| > \epsilon] = 0 \quad \text{für alle } \epsilon > 0 \,.$$

Bemerkung. Fast sichere Konvergenz entspricht im wesentlichen punktweiser Konvergenz der Zufallsvariablen $\omega \mapsto X_n(\omega)$ also Funktionen auf dem Wahrscheinlichkeitsraum Ω, wobei die Konvergenz auf Nullmengen nicht gegegeben sein muss d.h. $\mathbb{P}[X_n \not\to X] = 0$, da diese bei Integration über Ω keine Rolle spielen.
Die stochastische Konvergenz ist durch einen Grenzwert von Wahrscheinlichkeiten charakterisiert und eine schwächere Form der Konvergenz.



Beispiel. Wir betrachten den Wahrscheinlichkeitsraum $([0,1], \mathcal{B}_{[0,1]}, \mathbb{P} = \lambda_{[0,1]})$ mit Gleichverteilung auf $[0,1]$. Jedes $n \in \mathbb{N}$ besitzt eine eindeutige Darstellung der Form

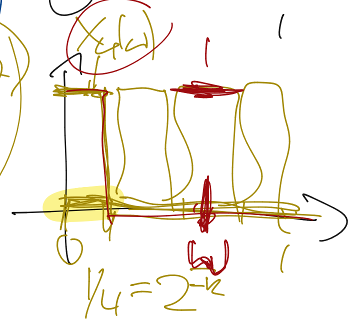
$$n = 2^k + j \quad \text{mit} \quad k \in \mathbb{N}_0 \text{ und } 0 \le j < 2^k \,.$$

$n=1: k=0, j=0$
$n=2: k=1, j=0$
$n=3: k=1, j=1$

Somit wird durch

$$X_n(\omega) := \begin{cases} 1 \,, & \text{falls } j2^{-k} \le \omega \le (j+1)2^{-k} \\ 0 \,, & \text{sonst} \end{cases} \quad , \quad \omega \in [0,1]$$

$n=4,\dots,7$
$k=2$



eine Folge $(X_n : n \in \mathbb{N})$ von Zufallsvariablen auf $\Omega = [0,1]$ definiert. Es gilt $X_n \xrightarrow{\mathbb{P}} X :\equiv 0$, denn für alle $0 < \epsilon < 1$ ist

$$\mathbb{P}[|X_n - X| > \epsilon] = \mathbb{P}[X_n = 1] = 2^{-k} \quad \text{für } 2^k \le n < 2^{k+1} \,.$$

$\to 0$ für $n \to \infty$

Andererseits ist

$$0 = \liminf_{n\to\infty} X_n(\omega) < \limsup_{n\to\infty} X_n(\omega) = 1 \quad \text{für alle } \omega \in [0,1],$$

$(X_n(\omega) : n \in \mathbb{N})$ konvergiert also für **kein** ω. Daher ist $\mathbb{P}[X_n \to X] = 0$ und damit natürlich auch keine fast sichere Konvergenz.

Satz 6.2. Seien $X, X_1, X_2, \ldots \in \mathbb{R}$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann gilt:

(a) Aus $X_n \xrightarrow{f.s.} X$ folgt $X_n \xrightarrow{\mathbb{P}} X$. Die Umkehrung dieser Aussage gilt nur auf diskreten Wahrscheinlichkeitsräumen.

(b) Aus $X_n \xrightarrow{\mathbb{P}} X$ folgt $X_n \xrightarrow{d} X$. Die Umkehrung gilt falls X deterministisch ist, d.h. $\mathbb{P}[X = a] = 1$ für ein $a \in \mathbb{R}$.

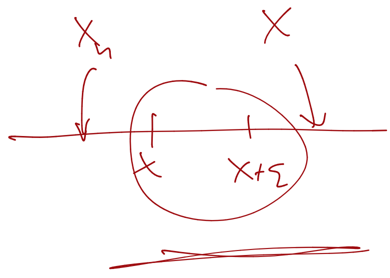
$X_n \xrightarrow{f.s} X$
$\Downarrow \quad \Uparrow$ dislr. Ω
$X_n \xrightarrow{\mathbb{P}} X$
$\Downarrow \quad \Uparrow$ $\mathbb{P}[X=a]=1$ für ein $a \in \mathbb{R}$
$X_n \xrightarrow{d} X$

Beweis. (a) Für alle $\epsilon > 0$ gilt

σ-Stetigkeit

$$\mathbb{P}[|X_n - X| > \epsilon] \le \mathbb{P}[\sup_{k\ge n} |X_k - X| > \epsilon] \xrightarrow{n\to\infty} \mathbb{P}[\limsup_{k\to\infty} |X_k - X| > \epsilon] \le \mathbb{P}[X_k \not\to X] = 0$$

unter Benutzung der σ-Stetigkeit von $\mathbb{P}$. Damit folgt die Aussage, für diskrete Wahrscheinlichkeitsräume siehe Übungsaufgabe.



(b) Für alle $x \in \mathbb{R}$ und $\epsilon > 0$ gilt

$$\overline{F_{X_n}(x)} = \mathbb{P}[X_n \leq x] = \mathbb{P}[X_n \leq x, X \leq x+\epsilon] + \mathbb{P}[X_n \leq x, X > x+\epsilon] \leq F_X(x+\epsilon) + \mathbb{P}[|X_n - X| > \epsilon] \;.$$

und

$$\overline{F_X(x-\epsilon)}\;\mathbb{P}[X \leq x-\epsilon] = \mathbb{P}[X \leq x-\epsilon, X_n \leq x] + \mathbb{P}[X \leq x-\epsilon, X_n > x] \leq F_{X_n}(x) + \mathbb{P}[|X_n - X| > \epsilon] \;.$$

$\to 0$ $\to 0$

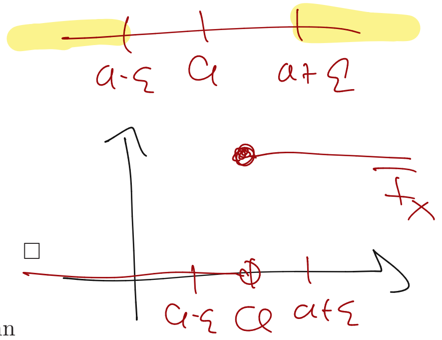
Also $F_X(x-\epsilon) - \mathbb{P}[|X_n - X| > \epsilon] \leq F_{X_n}(x) \leq F_X(x+\epsilon) + \mathbb{P}[|X_n - X| > \epsilon]$,
und für $n \to \infty$ folgt mit stochastischer Konvergenz

$$F_X(x-\epsilon) \leq \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x+\epsilon)$$

für alle $\epsilon > 0$. Wenn F_X in x stetig ist folgt dann

$$F_X(x-\epsilon) \nearrow F_X(x) \quad \text{und} \quad F_X(x+\epsilon) \searrow F_X(x) \quad \text{für } \epsilon \searrow 0 \;,$$

und damit existiert $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$, d.h. Konvergenz in Verteilung.
Wenn $\mathbb{P}[X = a] = 1$ für ein $a \in \mathbb{R}$, dann folgt für alle $\epsilon > 0$



$$\begin{aligned}\mathbb{P}[|X_n - X| > \epsilon] &= \mathbb{P}[|X_n - a| > \epsilon] = \mathbb{P}[X_n < a-\epsilon] + \mathbb{P}[X_n > a+\epsilon] \\ &\leq F_{X_n}(a-\epsilon) + 1 - F_{X_n}(a+\epsilon) \;.\end{aligned}$$

Falls $X_n \xrightarrow{d} X$ konvergiert die rechte Seite für $n \to \infty$ gegen

$$F_X(a-\epsilon) + 1 - F_X(a+\epsilon) = 0 \;,$$

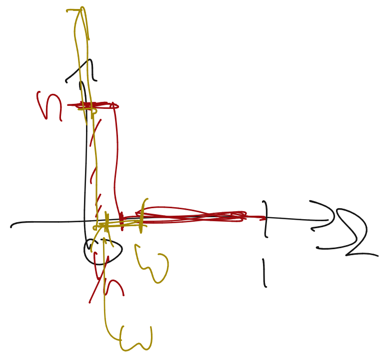
denn $a \pm \epsilon$ sind Stetigkeitspunkte von $F_X(x) = \mathbb{1}_{[a,\infty)}(x)$, und somit $X_n \xrightarrow{\mathbb{P}} X$. □

Beispiele.

- Sei $X \sim \mathcal{B}_{1,1/2}$ eine Bernoulli Variable und $X_n = X$ für alle $n \in \mathbb{N}$. Dann konvergiert $X_n \xrightarrow{d} Y := 1 - X$ wegen der Symmetrie der Verteilung. Aber es ist $|X_n(\omega) - Y(\omega)| = 1$ für alle $\omega \in \Omega$ (beliebiger Wahrscheinlichkeitsraum), also kann X_n nicht gegen Y in einem anderen Sinne konvergieren.

- **Achtung!** Fast sichere Konvergenz $X_n \xrightarrow{f.s.} X$ (oder schwächere) impliziert im Allgemeinen nicht Konvergenz von Momenten wie $\mathbb{E}[X_n] \to \mathbb{E}[X]$ für den Erwartungswert. Betrachte zum Beispiel die Zufallsvariablen

$$X_n(\omega) = n\mathbb{1}_{(0,1/n)}(\omega) \quad \text{auf dem Wahrscheinlichkeitsraum} \quad ((0,1), \mathcal{B}_{(0,1)}, \lambda) \;.$$



Dann ist $\mathbb{E}[X_n] = n\frac{1}{n} = 1$ für alle $n \in \mathbb{N}$, aber für alle $\omega \in (0,1)$ ist $X_n(\omega) \to 0$, d.h. $X_n \xrightarrow{f.s.} 0$ und $\mathbb{E}[X_n] = 1 \not\to 0$.

$\int_0^1 X_n(\omega)\,d\omega$

$P\{X_n \to 0\} = P\{\Omega\} = 1$