Vorlesung 13 | 8.12.2020 | 14:15–16:00 via Zoom

(Ein Handzettel für heutige Vorlesung finden Sie auf der Seite der Vorlesung auf meiner Website (URL), siehe Tagebuch)

Information der Fachschaft: Dieses Jahr findet die Mathe-Weihnachtsfeier am Donnerstag, 17.12, ab 18 ct. online via Zoom statt. Alle aktuellen Informationen sind auf https://fsmath.uni-bonn.de/veranstaltungen-detail/events/virtuelle-mathe-weihnachtsfeier.html zu finden. Schaut vorbei!

Information from the Fachschaft: This year's Math Christmas party will take place at Thursday, the 17.12. starting 18 ct. online via zoom. All current information can be found on https://fsmath.uni-bonn.de/events-detail/events/virtual-christmas-party.html. Swing by!

5 Konvergenzbegriffe

(Kapitel 5 in Bovier Skript)

In der letzten Vorlesungen haben wir gesehen: schwache Konvergenz von Verteilungsfuktionen und von W-Maße, Konvergenz in Verteilung von Z.V., Konvergenz in W-keit von Z.V.

Heutigen Vorlesung: Konvergenz in L^p, Fast sichere Konvergenz, Borel–Cantelli Lemmata.

5.1 Konvergenz von Verteilungsfunktionen

Definition 1. *Sei $(F_n)_{n\geqslant 1}$ eine Folge von Verteilungsfunktionen. Dann konvergiert F_n schwach gegen eine Verteilungsfunktion F, falls*

$$\lim_{n\to\infty} F_n(x) = F(x)$$

für alle $x\in\mathbb{R}$ für welche F stetig ist.

Definition 2. *Sei $(\mathbb{P}_n)_{n\geqslant 1}$ eine Folge von W-Maße auf $(\Omega,\mathscr{B}(\Omega))$ mit Ω topologische Raum. Dann konvergiert $(\mathbb{P}_n)_n$ schwach gegen $\mathbb{P}$ falls für alle beschränkten stetigen Funktionen $g\colon\Omega\to\mathbb{R}$,*

$$\lim_{n\to\infty}\int_\Omega g(\omega)\mathbb{P}_n(\mathrm{d}\omega)\to\int_\Omega g(\omega)\mathbb{P}(\mathrm{d}\omega).$$

Satz 3. *Sei $(\mathbb{P}_n)_n$ eine Folge W-maße auf $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ und $F_n(x):=\mathbb{P}_n((-\infty,x])$. Dann konvergiert $(\mathbb{P}_n)_n$ schwach gegen ein W-maß $\mathbb{P}$ mit Verteilungsfunktion $F(x):=\mathbb{P}_n((-\infty,x])$ dann und nur dann, wenn*

$$F_n\xrightarrow[\text{schwach}]{}F.$$

Beispiel. Wir haben die folgenden schwachen Konvergenzen von W-Maße:

- $\lambda>0$ (diskret→diskret)

$$\mathrm{Bin}(n,\lambda/n)\xrightarrow[n\to\infty]{}\mathrm{Poi}(\lambda)$$

- (diskret→abs. stetig)

$$(1-q)\mathrm{Geo}(q)\xrightarrow[q\to 1]{}\mathrm{Exp}(1)$$

- $$\left(1-\frac{1}{M}\right)\delta_0+\frac{1}{M}\delta_M\xrightarrow[M\to\infty]{}\delta_0$$

- $$\mathbb{P}_M = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_M \xrightarrow[M\to\infty]{} ??? \quad \text{keine Konvergenz hier}$$

$$\int_\Omega g(\omega)\mathbb{P}_M(\mathrm{d}\omega) = \frac{1}{2}g(0) + \frac{1}{2}g(M)$$

Betrachten $g(x) = \sin(x)$.

(Siehe oben)

5.2 Konvergenz von Z.V.

Definition 4. *(Konvergenz in Verteilung) Sei $(X_n)_{n\geqslant 1}$ eine Folge von Z.V. wobei X_n auf $(\Omega_n, \mathscr{F}_n, \mathbb{P}_n)$ definiert ist. Dann konvergiert X_n in Verteilung gegen eine Z.V. X,*

$$X_n \xrightarrow[n\to\infty]{\mathscr{D}} X,$$

falls die Folge $(F_{X_n})_{n\geqslant 1}$ schwach konvergiert gegen F_X, d.h.

$$F_{X_n}(x) = \mathbb{P}(X_n \leqslant x) \xrightarrow[\textit{schwach}]{} F_X(x) = \mathbb{P}(X \leqslant x).$$

Bemerkung. Die Z.V. X_n müssen nicht auf dasselbe W-raum definiert sein!

Definition 5. *(Konvergenz in W-keit) Seien $X, (X_n)_n$ Z.V. auf $(\Omega, \mathscr{F}, \mathbb{P})$. Die Folge $(X_n)_n$ konvergiert in W-keit gegen X, falls $\forall \varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\big(|X_n - X| > \varepsilon\big) = 0.$$

Bemerkung. Hier sind die $(X_n)_n$ und X definiert auf dasselbe W-raum!

Zusammenhänge:

Lemma 6.

$$\boxed{X_n \xrightarrow[n\to\infty]{\mathbb{P}} X} \overset{\Longrightarrow}{\not\Longleftarrow} \boxed{X_n \xrightarrow[n\to\infty]{\mathscr{D}} X}$$

Beweis. ($\Rightarrow$) Konvergenz in W-keit, bedeutet dass $\forall \varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$.
Sei $c \in \mathbb{R}$ s.d. $F_X(x) = \mathbb{P}(X \leqslant x)$ stetig in c ist und wir setzen $F_n(x) = \mathbb{P}(X_n \leqslant x)$. Dann

$$|F(c) - F_n(c)| = |\mathbb{P}(X \leqslant c) - \mathbb{P}(X_n \leqslant c)| = |\mathbb{P}(X \leqslant c, X_n > c) + \mathbb{P}(X \leqslant c, X_n \leqslant c) - \mathbb{P}(X_n \leqslant c)|$$

$$= \Big|\mathbb{P}(X \leqslant c, X_n > c) + \underbrace{\mathbb{P}(X \leqslant c, X_n \leqslant c) - \mathbb{P}(X_n \leqslant c)}_{= -\mathbb{P}(X_n \leqslant c, X > c)}\Big| = |\mathbb{P}(X \leqslant c, X_n > c) - \mathbb{P}(X_n \leqslant c, X > c)|$$

$$\leqslant \mathbb{P}(X \leqslant c, X_n > c) + \mathbb{P}(X_n \leqslant c, X > c)$$

$$\leqslant \mathbb{P}(X \leqslant c - \varepsilon, X_n > c) + \mathbb{P}(X_n \leqslant c, X > c + \varepsilon) + \mathbb{P}(c - \varepsilon \leqslant X \leqslant c + \varepsilon)$$

$$\leqslant \underbrace{\mathbb{P}(|X - X_n| > \varepsilon)}_{\xrightarrow[n\to\infty]{} 0 \text{ (Annahme)}} + \underbrace{\mathbb{P}(c - \varepsilon \leqslant X \leqslant c + \varepsilon)}_{= F(c+\varepsilon) - F(c-\varepsilon) \xrightarrow[\varepsilon\to 0]{} 0 \quad \text{weil } F \text{ stetig in } c \text{ ist.}}$$

Dann für alle $\varepsilon > 0$:

$$\limsup_{n\to\infty} |F(c) - F_n(c)| \leqslant F(c+\varepsilon) - F(c-\varepsilon) \Rightarrow \limsup_{n\to\infty} |F(c) - F_n(c)| \leqslant \liminf_{\varepsilon\to 0} (F(c+\varepsilon) - F(c-\varepsilon)) = 0$$

Wir haben dass $F_n(x) \to F(x)$ für alle $x \in \mathbb{R}$ wo F stetig ist.

($\not\Leftarrow$) Gegenbeispiel. Sei $\Omega = \{1, 2\}$ mit Gleichverteilung. Setze $X_n(1) = 1$ und $X_n(2) = 0$ für alle $n \geqslant 1$, und $X(1) = 0$, $X(2) = 1$. Dann

$$F_n(x) = \begin{cases} 0, & x < 0 \\ 1/2, & x \in [0, 1) \\ 1, & x \geqslant 1 \end{cases} = \mathbb{P}(X \leqslant x) = F(x)$$

aber $|X_n - X| = 1$ für alle $n \geqslant 1$ und $X_n \overset{\mathbb{P}}{\underset{n\to\infty}{\not\longrightarrow}} X$.

□

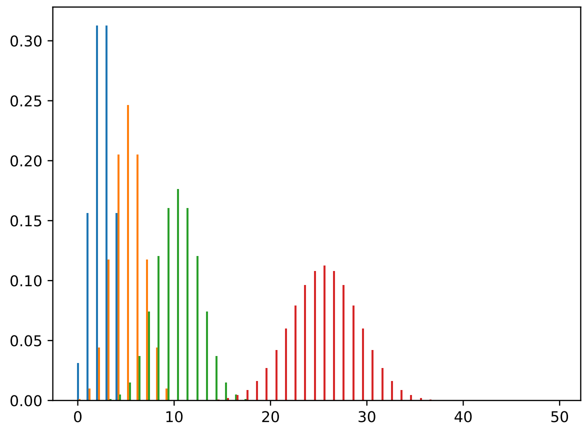
Jetz werden wir auschauen wie Summe i.i.d Bernoulli Z.V. sich Verhalten.

Machen wir eine Simulation. Betrachten wir

$$S_n = \sum_{k=1}^{n} X_n \sim \text{Bin}(n, 1/2), \qquad X_n \sim \text{Ber}(1/2),$$

mit $n = 1, 5, 10, 20, 50$. Die Python Funktion `scipy.stats.binom.pmf(x,n,p)` gibt die Werten von die W-keit $\mathbb{P}(\text{Bin}(n, p) = x)$ zuruck (wir verwenden die Programmbibliotheken `matplotlib`, `scipy` und `numpy`).

```
>>> from scipy.stats import binom
    import numpy as np
    import matplotlib.pyplot as plt
    # setting the values of n and p
    n = [1, 5, 10, 20, 50]; p = 0.5
>>> plt.clf()
    for i in range(1,5):
        r_values = np.arange(0, n[i],1)
        dist = [binom.pmf(r, n[i], p) for r in r_values]
        plt.bar(r_values+0.2*(i-1), dist, 0.2)
    pdf_out(plt.gcf())
```




```
>>>
```

Wir haben

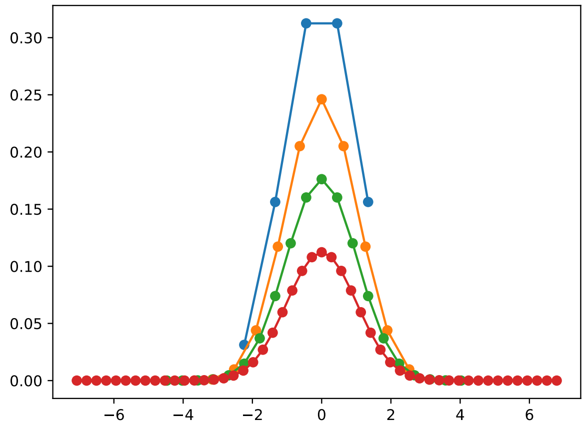
$$\mathbb{E}[S_n] = np, \qquad \text{Var}(S_n) = np(1-p).$$

Betrachten wir jetz die Verteilung von

$$Y_n = \frac{(S_n - n \cdot p)}{\sqrt{np(1-p)}}$$

mit $S_n \sim \mathrm{Bin}(n, 1/2)$, $p = 1/2$, $n = 1, 5, 10, 20, 50$.

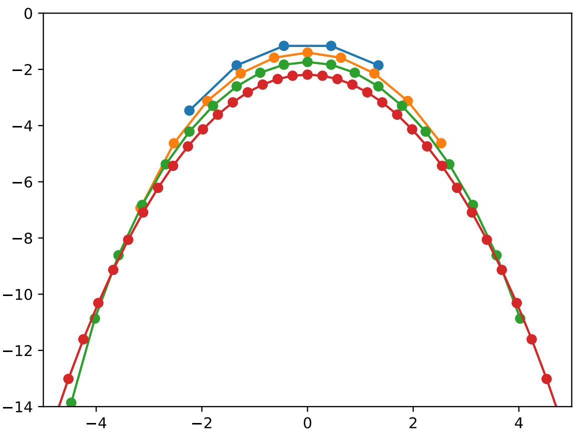
```
>>> plt.clf()
    for i in range(1,5):
        r_values = np.arange(0, n[i],1)
        dist = [binom.pmf(r, n[i], p) for r in r_values]
        xpos = ((r_values - n[i]*p)/np.sqrt(n[i]*p*(1-p)))
        plt.plot(xpos, dist, "o-")
    pdf_out(plt.gcf())
```




```
>>>
```

Das gleiche Diagramm mit einer logarithmischen Skala:

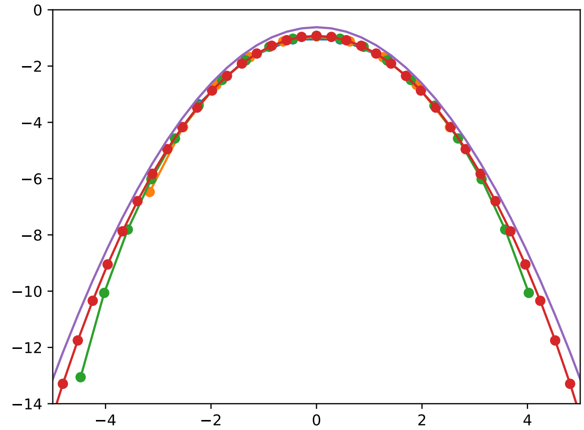
```
>>> plt.clf()
    for i in range(1,5):
        r_values = np.arange(0, n[i],1)
        dist = [np.log(binom.pmf(r, n[i], p)) for r in r_values]
        xpos = ((r_values - n[i]*p)/np.sqrt(n[i]*p*(1-p)))
        plt.plot(xpos, dist, "o-")
    plt.xlim(-5, 5); plt.ylim(-14,0)
    pdf_out(plt.gcf())
```



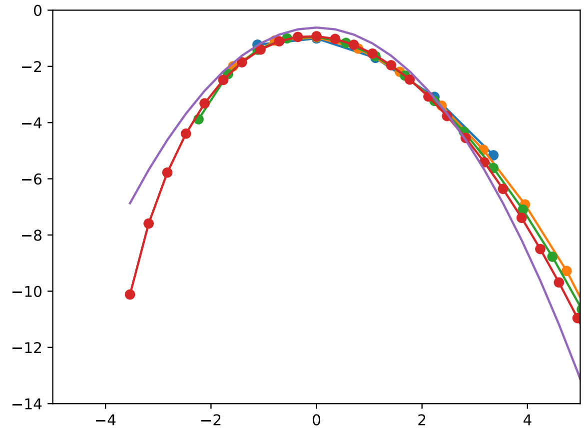
```
>>>
```

Wir sehen, dass sich die Verteilung von $(Y_n)_n$ einer Grenze nähert. Der Unterschied in der vertikalen Skala ist auf die unterschiedliche Dichte der Punkte in der horizontalen Achse zurückzuführen. Um dies zu berücksichtigen, können wir die Wahrscheinlichkeitswerte mit dem Faktor $(np(1-p))^{1/2}$ multiplizieren.

```
>>> p=0.5;
    plt.clf()
    for i in range(1,5):
        r_values = np.arange(0, n[i],1)
        dist = [np.log(np.sqrt(n[i]*p*(1-p))*binom.pmf(r, n[i], p)) for r in r_values]
        xpos = ((r_values - n[i]*p)/np.sqrt(n[i]*p*(1-p)))
        plt.plot(xpos, dist, "o-")
    plt.plot(xpos, [0.3-np.log(2*np.pi)*0.5-x*x*0.5 for x in xpos], "-")
    plt.xlim(-5, 5); plt.ylim(-14,0)
    pdf_out(plt.gcf())
```




```
>>>
>>> p=0.2;
    plt.clf()
    for i in range(1,5):
        r_values = np.arange(0, n[i],1)
        dist = [np.log(np.sqrt(n[i]*p*(1-p))*binom.pmf(r, n[i], p)) for r in r_values]
        xpos = ((r_values - n[i]*p)/np.sqrt(n[i]*p*(1-p)))
        plt.plot(xpos, dist, "o-")
    plt.plot(xpos, [0.3-np.log(2*np.pi)*0.5-x*x*0.5 for x in xpos], "-")
    plt.xlim(-5, 5); plt.ylim(-14,0)
    pdf_out(plt.gcf())
```



>>>

$$\log \mathbb{P}(Y_n = x) \approx -\frac{1}{2}\log(2\pi) - \frac{x^2}{2}$$

Das ist (die log der Dichte) eine Normalverteilung $\mathcal{N}(0,1)$. Der folgende Satz präzisiert diese Beobachtungen.

Satz 7. *(Satz von de Moivre–Laplace) Sei $(X_n)_{n\geqslant 1}$ eine Folge iid* Ber(p) *Z.V.. Dann konvergiert die Folge*

$$Z_n := \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left(X_k - p\right)$$

in Verteilung gegen eine $\mathcal{N}(0, p(1-p))$ Z.V.

Bemerkung. Wir haben $\mathbb{E}[Z_n] = 0$, $\operatorname{Var}(Z_n) = p(1-p)$.

Beweis. Sei $S_n := \sum_{k=1}^{n} X_k$. Dann

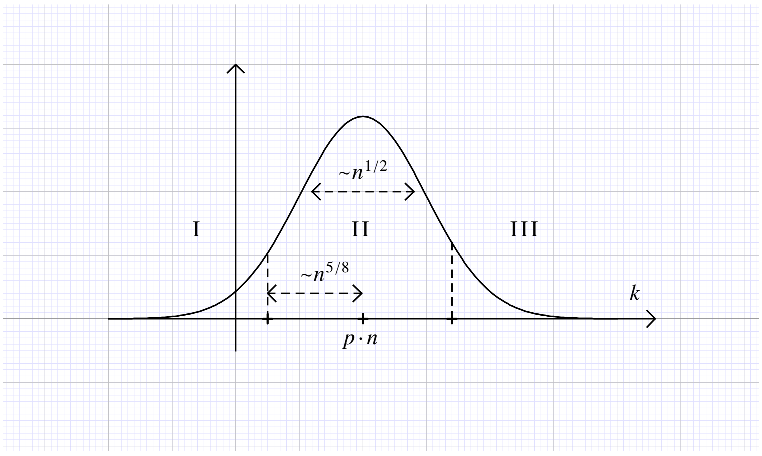
$$\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Wir wollen zeigen, dass

$$\mathbb{P}(Z_n \in (a,b]) \longrightarrow \mathbb{P}(\mathcal{N}(0, p(1-p)) \in (a,b]) = \int_a^b \exp\left(-\frac{x^2}{2p(1-p)}\right)\frac{\mathrm{d}x}{(2\pi p(1-p))^{1/2}}$$

als $n \to \infty$, fur alle $a < b$. Aber

$$\mathbb{P}(Z_n \in (a,b]) = \sum_{k \in (n^{1/2}a + np, n^{1/2}b + np]} \mathbb{P}(S_n = k).$$



Sei $k = x \cdot n \in \mathbb{Z}$ mit $|x - p| \leqslant n^{-3/8}$ (Region II)

$$\mathbb{P}(S_n = x \cdot n) = \frac{n!}{(nx)!(n(1-x))!}\underbrace{\left(\frac{p}{1-p}\right)^{nx}(1-p)^n}_{\exp(n(x\log p - (1-x)\log(1-p)))}$$

Stirling formel:

$$n! = (2\pi n)^{1/2} n^n e^{-n}\left[1 + O\left(\frac{1}{n}\right)\right]$$

$$\frac{n!}{(nx)!(n(1-x))!} = \frac{\exp(-nI(x))}{(2\pi n)^{1/2}(x(1-x))^{1/2}}\left[1 + O\left(\frac{1}{n}\right)\right]$$

mit

$$I(x) = x\log(x) + (1-x)\log(1-x).$$

$$\Rightarrow \mathbb{P}(S_n = nx) = \frac{1}{(2\pi nx(1-x))^{1/2}} e^{-nI(p,x)}\left[1 + O\left(\frac{1}{n}\right)\right]$$

wobei

$$I(p,x) = x\log(x) + (1-x)\log(1-x) - x\log(p) - (1-x)\log(1-p)$$

Eigenschaffen von $x \mapsto I(p,x)$.

$$\begin{array}{ll} (a) & I(p,p) = 0 \\ (b) & \frac{\partial^2}{\partial x^2} I(p,x) = \frac{1}{x(1-x)} \geqslant 4 > 0 \Rightarrow I(p,\cdot) \text{ konvex} \\ (c) & \frac{\mathrm{d}}{\mathrm{d}x} I(p,x) = \log\left(\frac{x}{p}\right) + \log\left(\frac{1-p}{1-x}\right) = 0 \Rightarrow x = p \end{array}$$

Taylorformel dritte Ordnung zeigen dass: $(x-p)^3 \approx (n^{-3/8})^3$ (Region II)

$$I(p,x) = \frac{(x-p)^2}{2p(1-p)} + O(n^{-9/8}) \qquad \text{für } |x-p| \leqslant n^{-3/8}.$$

$$\Rightarrow \mathbb{P}(S_n = p\cdot n + yn^{1/2}) \overset{|y|\leqslant n^{1/8}}{=\!=\!=} \frac{e^{-x^2/(2p(1-p))}}{(2\pi np(1-p))^{1/2}}\left[1 + O(n^{-1/8})\right]$$

Es folgt dass

$$\mathbb{P}(S_n \leqslant pn + yn^{1/2}) = \underbrace{\mathbb{P}(S_n < pn - n^{5/8})}_{(\mathrm{I})} + \underbrace{\mathbb{P}(pn - n^{5/8} \leqslant S_n \leqslant pn + yn^{1/2})}_{(\mathrm{II})}$$

Aber da

$$2\mathbb{P}(S_n < pn - n^{5/8}) \leqslant 1 - \mathbb{P}(pn - n^{5/8} \leqslant S_n \leqslant pn + n^{5/8}).$$

weil $\mathrm{I} + \mathrm{II} + \mathrm{III} = \mathbb{R}$. Z.z.

$$\begin{array}{ll} (a) & \mathbb{P}(pn - n^{5/8} \leqslant S_n \leqslant pn + n^{5/8}) \to 1 \\ (b) & \mathbb{P}(pn - n^{5/8} \leqslant S_n \leqslant pn + yn^{1/2}) \to \int_{-\infty}^{y} \frac{e^{-x^2/(2p(1-p))}}{(2\pi np(1-p))^{1/2}}\mathrm{d}y \end{array}$$

(b) gibt, weil

$$\mathbb{P}(pn - n^{5/8} \leqslant S_n \leqslant pn + yn^{1/2}) = \sum_{k = pn - n^{5/8}}^{pn + yn^{5/8}} \mathbb{P}(S_n = k)$$

$$= \sum_{z \in \mathcal{I}_n, z \leqslant y} \mathbb{P}(S_n = pn + zn^{1/2})$$

wo $\mathcal{I}_n := n^{-1/2}(\mathbb{Z} - pn) \cap [-n^{1/8}, n^{1/8}]$.

$$= \frac{1}{n^{1/2}} \sum_{z \in \mathcal{I}_n} \frac{e^{-z^2/(2p(1-p))}}{(2\pi np(1-p))^{1/2}}\left[1 + O(n^{-1/8})\right]$$

Dazu: $(e^{-c(z+n^{-1/2})^2} - e^{-cz^2} = e^{-cz^2}(1 + O(n^{-3/8}))$ gleichmässig für $|z| \leqslant n^{1/8})$

$$= (1 + O(n^{-1/8}))\int_{-n^{1/8}}^{y} \frac{e^{-z^2/(2p(1-p))}}{(2\pi np(1-p))^{1/2}}\mathrm{d}z$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{P}(pn - n^{5/8} \leqslant S_n \leqslant pn + yn^{1/2}) = \Phi_{0,p(1-p)}(y),$$

Aus dem gleichem Gründe,

$$\lim_{n\to\infty}\mathbb{P}(S_n < pn - n^{5/8}) = \lim_{n\to\infty}\frac{1}{2}(1-\mathbb{P}(pn-n^{5/8}\leqslant S_n \leqslant pn+n^{5/8})) = \frac{1}{2}(1-\Phi_{0,p(1-p)}(+\infty)) = 0$$

Im Ende wir haben: für alle $y\in\mathbb{R}$

$$\lim_{n\to\infty}\mathbb{P}(Z_n\leqslant y) = \lim_{n\to\infty}\mathbb{P}(S_n\leqslant pn+yn^{1/2}) = \Phi_{0,p(1-p)}(y) = \mathbb{P}(\mathcal{N}(0,p(1-p))\leqslant y).$$

□

Definition 8. *(Konvergenz in L^p) Seien $X,(X_n)_n$ Z.V. auf $(\Omega,\mathscr{F},\mathbb{P})$. Sei $p\geqslant 1$ und nehmen wir an $(X_n)_{n\geqslant 1}\subseteq \mathscr{L}^p$, $X\in\mathscr{L}^p$ d.h.*

$$\|X\|_p = \left[\mathbb{E}|X|^p\right]^{1/p} < \infty.$$

Dann konvergiert X_n gegen X in $\mathscr{L}^p$,

$$X_n \xrightarrow[n\to\infty]{\mathscr{L}^p} X,$$

falls

$$\lim_{n\to\infty}\|X_n - X\|_p = 0.$$

Bemerkung.

$$L^p = \mathscr{L}^p \text{ modulo Äquivalenzklassen } (X\sim Y \Leftrightarrow \|X-Y\|_p = 0)$$

ist ein Banach Raum. Für $p=2$ ist ein Hilbertraum mit $\langle X,Y\rangle = \mathbb{E}[XY]$.

Lemma 9.

$$\boxed{X_n \xrightarrow{\mathscr{L}^p} X} \underset{\not\Leftarrow}{\Rightarrow} \boxed{X_n \xrightarrow{\mathbb{P}} X}$$

Beweis. $(\Rightarrow)$ Sei $p\geqslant 1$ und $X_n \xrightarrow{\mathscr{L}^p} X$, d.h.

$$\|X_n - X\|_p^p = \mathbb{E}[|X_n - X|^p] \to 0.$$

Dann $\forall\varepsilon > 0$

$$\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{E}[\mathbb{1}_{|X_n - X|>\varepsilon}] \underset{\text{Markov ungl.}}{\leqslant} \frac{\mathbb{E}[|X_n - X|^p]}{\varepsilon^p} \to 0$$

(Markov ungleichung)

$(\not\Leftarrow)$ Gegenbeispiel. Sei $(X_n)_{n\geqslant 1}$ mit $\mathbb{P}(X_n = n) = n^{-1}$ und $\mathbb{P}(X_n = 0) = 1 - n^{-1}$. Dann $\forall\varepsilon > 0$

$$\mathbb{P}(|X_n| > \varepsilon) = n^{-1} \to 0$$

als $n\to\infty$, aber $\forall p\geqslant 1$

$$\mathbb{E}|X_n - 0|^p = n^{p-1} \underset{n\to\infty}{\not\longrightarrow} 0.$$

□

Lemma 10. *(Markov Ungleichung) Sei $X\geqslant 0$ eine Z.V. dann für alle $\lambda > 0$*

$$\mathbb{P}(X > \lambda) \leqslant \frac{\mathbb{E}[X]}{\lambda}$$

Beweis. Ist klar dass

$$\mathbb{1}_{\{X>\lambda\}}(\omega) \leqslant \frac{X(\omega)}{\lambda}, \qquad \omega \in \Omega$$

dann nehmen die $\mathbb{E}$:

$$\mathbb{P}(X>\lambda) = \mathbb{E}[\mathbb{1}_{\{X>\lambda\}}] \leqslant \frac{\mathbb{E}[X]}{\lambda}, \qquad \square$$

Bemerkung. Für alle monotone steigende Funktionen $\Phi: \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ wir haben

$$\mathbb{P}(X>\lambda) \leqslant \mathbb{P}(\Phi(X) > \Phi(\lambda)) \underset{\text{Markov}}{\leqslant} \frac{\mathbb{E}[\Phi(X)]}{\Phi(\lambda)}.$$

Z.B. $\Phi(x) = x^p$ für alle $p>0$.
