Vorlesung 16 | 18.12.2020 | 14:15–16:00 via Zoom

Weitere information der Fachschaft: Am **22. Dezember 2020 um 19:15** findet ein

Treffen für die Mathe-Lehramtsstudierenden

auf Zoom statt. Die Zugangsdaten findet ihr auf der Fachschaftswebsite (www.fsmath.uni-bonn.de). Wir sind gespannt auf eure Erfahrungen und Eindrücke! Bitte erscheint zahlreich, damit wir viel Rückmeldung bekommen, um das Studium für kommende Generationen zu optimieren.

6 Das Gesetz der großen Zahlen (fortsetzung)

(Kapitel 6 in Bovier Skript)

Folgerung 9. *Sei X eine reelle Z.V.. Dann*

$$\mathbb{P}(X \geqslant a) \leqslant \inf_{\lambda \geqslant 0} [e^{-\lambda a} \mathbb{E}[e^{\lambda X}]]$$

Ziemlich effektiv (gut) ist die Ungleichung von Korollar 9 angewendet an Summe unabhängiger (noch besser iid) Z.V.

Folgerung 10. *Sei $(X_n)_{n \geqslant 1}$ eine Folge i.i.d. Z.V.. Dann*

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geqslant a\right) \leqslant \inf_{\lambda \geqslant 0} [e^{-\lambda n a} (\mathbb{E}[e^{\lambda X_1}])^n]$$

Beweis. Einfach. Beachten dass

$$\mathbb{E}[e^{\lambda \sum_{k=1}^{n} X_k}] = \mathbb{E}\left[\prod_{k=1}^{n} e^{\lambda X_k}\right] \underset{\text{unab.}}{=} \prod_{k=1}^{n} \mathbb{E}[e^{\lambda X_k}] \underset{\text{id. Vert.}}{=} (\mathbb{E}[e^{\lambda X_1}])^n$$

□

Beispiel. Sei X_n mit $\mathbb{P}(X_n = \pm 1) = 1/2$, unabhängiger Z.V.

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geqslant a\right) \leqslant \inf_{\lambda \geqslant 0} [e^{-\lambda n a} (\cosh(\lambda))^n]$$

$$= \exp\left(n \inf_{\lambda \geqslant 0} (-a\lambda + \log\cosh(\lambda))\right) = \exp(-nI(a))$$

wobei

$$I(a) := \frac{1-a}{2}\log(1-a) + \frac{1+a}{2}\log(1+a).$$

Bedingung:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}(-a\lambda + \log\cosh(\lambda))\bigg|_{\lambda=\lambda_*} = -a + \frac{\sinh(\lambda_*)}{\cosh(\lambda_*)} = 0 \Leftrightarrow \tanh \lambda_* = a \Leftrightarrow \lambda_* = \operatorname{arctanh}(a) = \frac{1}{2}\log\frac{1+a}{1-a}$$

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k \geqslant a\right) \leqslant \exp(-nI(\lambda_*))$$

Die Abschatzung exponentiell fallen.

6.3 Das schwache Gesetz der großen Zahlen

Zuerst bewisen wir schwächene Versionen von das GGZ.

Wir haben gesehen, dass die Varianz ein (quadratisches) Maß für die Dispersion ist. Die Varianz einer Summe von N unabhängigen Z.Variablen kann manchmal viel langsamer wachsen als N^2.

Sei

$$S_n = X_1 + \cdots + X_n.$$

Lemma 11. *Seien $X_1, \ldots, X_n$ reelle Z.V. und $S_n = X_1 + \cdots + X_n$. Dann*

$$\mathrm{Var}(S_n) = \sum_{k=1}^{n} \mathrm{Var}(X_k) + 2 \sum_{1 \leqslant k < \ell \leqslant n} \mathrm{Cov}(X_k, X_\ell).$$

Beweis. Wegen Bilinearität und Symetrie von $(X, Y) \mapsto \mathrm{Cov}(X, Y)$ gilt

$$\mathrm{Var}(S_n) = \mathrm{Cov}(S_n, S_n) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \mathrm{Cov}(X_k, X_\ell) = \sum_{k=1}^{n} \mathrm{Var}(X_k) + 2 \sum_{1 \leqslant k < \ell \leqslant n} \mathrm{Cov}(X_k, X_\ell).$$

□

Betrachten wir die Annahme „Schnelles Abklingen der positiven Korrelationen“

Annhame. *(A) $\exists (c_n)_{n \geqslant 1}$, $c_n \in \mathbb{R}_+$ mit $\sum_{n \geqslant 0} c_n < \infty$ und*

$$\mathrm{Cov}(X_n, X_\ell) = \mathbb{E}[(X_n - \mathbb{E}[X_n])(X_\ell - \mathbb{E}[X_\ell])] = \mathbb{E}[X_n X_\ell] - \mathbb{E}[X_n]\mathbb{E}[X_\ell] \leqslant c_{|n-\ell|}, \qquad k, \ell \geqslant 1.$$

Fall A1) *Kovarianzen exponentiell abfallen:*

$$|\mathrm{Cov}(X_n, X_\ell)| \leqslant c\, \alpha^{|n-\ell|}$$

für einen $\alpha \in (0,1)$, $c \in \mathbb{R}_+$.

Fall A2) *$X_1, X_2, \ldots$ unkorreliert und beschränken Varianzen: d.h.*

$$\mathrm{Cov}(X_k, X_\ell) = 0, \quad k, \ell \geqslant 1, \qquad \nu = \sup_{k \geqslant 1} \mathrm{Var}(X_k) < \infty.$$

Annhame (A) erffült mit $c_0 = \nu$, $c_n = 0$, $n \geqslant 1$.

Satz 12. *(Schwaches Gesetz der großer Zahlen und L^2-Version) Under der Voraussetzung (A), für alle $n \geqslant 1$, $\varepsilon > 0$, gilt:*

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \frac{\mathbb{E}[S_n]}{n} \right| \geqslant \varepsilon \right) \leqslant \frac{C}{\varepsilon^2 n}$$

und

$$\mathrm{Var}\left(\frac{S_n}{n} \right) = \mathbb{E}\left[\left(\frac{S_n}{n} - \frac{\mathbb{E}[S_n]}{n} \right)^2 \right] \leqslant \frac{C}{n}$$

für $C = c_0 + 2\sum_{n \geqslant 1}^{\infty} c_n < \infty$. Falls dazu $\mu = \mathbb{E}[X_k]$, $k \geqslant 1$ dann

$$\frac{S_n}{n} \xrightarrow[n \to \infty]{} \mu \quad \textit{in } L^2 \textit{ und in W-keit.}$$

Beweis.

$$\mathbb{E}\left[\left(\frac{S_n}{n}-\frac{\mathbb{E}[S_n]}{n}\right)^2\right]=\mathrm{Var}\left(\frac{S_n}{n}\right)=\frac{1}{n^2}\mathrm{Var}(S_n)$$

$$=\frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(X_k)+2\frac{1}{n^2}\sum_{1\leqslant k<\ell\leqslant n}\mathrm{Cov}(X_k,X_\ell)$$

$$\underset{(A)}{\leqslant}\frac{1}{n}c_0+2\frac{1}{n^2}\sum_{1\leqslant k<\ell\leqslant n}c_{|k-\ell|}\leqslant\frac{1}{n}c_0+2\frac{1}{n^2}\sum_{1\leqslant\ell\leqslant n}\sum_{m\geqslant 1}c_m=\frac{1}{n}\left(c_0+2\sum_{m\geqslant 1}c_m\right)=\frac{C}{n}.$$

$$\mathbb{P}\left(\left|\frac{S_n}{n}-\frac{\mathbb{E}[S_n]}{n}\right|\geqslant\varepsilon\right)\underset{\text{Tchebichev}}{\leqslant}\frac{1}{\varepsilon^2}\mathbb{E}\left[\left(\frac{S_n}{n}-\frac{\mathbb{E}[S_n]}{n}\right)^2\right]\leqslant\frac{C}{n\varepsilon^2}.$$

□

Bemerkung. Hier haben wir nicht angenommen, dass die Z.V. unabhängig sind. Dazu haben wir angenommen insbesonderes dass die Varianz änhlich sind.

Die Konvergenz ist nicht f.s., aber das ist nur technische Beweis bedingt.

Das Folgende ist eine Version des starken Gesetzes der großen Zahlen unter der Annahme (A). Wie wir sehen werden, ist dies nicht optimal, aber der Beweis ist recht einfach und eine interessante Anwendung von Borel-Cantelli I.

Satz 13. *Unter der Voraussetzung (A) folgt*

$$\lim_{n\to\infty}\left(\frac{S_n}{n}-\frac{\mathbb{E}[S_n]}{n}\right)=0,\qquad \mathbb{P}-f.s.$$

Falls $\mu=\mathbb{E}[X_k]$, $k\geqslant 1$ *dann* $\lim_{n\to\infty}\frac{S_n}{n}=\mu$ *f.s.*

Beweis. O.B.d.A. $\mathbb{E}[X_k]=0$, $k\geqslant 1$. Nehmen wir Teilfolgen $n_k=k^2$, Tchebichev gibt

$$\mathbb{P}\left(\left|\frac{S_{k^2}}{k^2}\right|\geqslant\varepsilon\right)\leqslant\frac{C}{\varepsilon k^2}.$$

Wegen Borel–Cantelli I, $S_{k^2}/k^2\to 0$ f.s. weil $\sum_{k\geqslant 1}k^{-2}<\infty$ (punktweise bis auf eine Nullmenge $\mathcal{N}_1$, d.h. $\mathbb{P}(\mathcal{N}_1)=0$). Teilen $(S_n)_n$ zwischen den Werten der Teilfolge. Sei

$$D_k:=\max_{k^2\leqslant\ell\leqslant(k+1)^2}|S_\ell-S_{k^2}|$$

Dann

$$\mathbb{P}\left(\frac{D_k}{k^2}\geqslant\varepsilon\right)=\mathbb{P}\left(\bigcup_{k^2\leqslant\ell\leqslant(k+1)^2}\{|S_\ell-S_{k^2}|\geqslant\varepsilon k^2\}\right)\leqslant\sum_{k^2\leqslant\ell\leqslant(k+1)^2}\mathbb{P}(|S_\ell-S_{k^2}|\geqslant\varepsilon k^2)$$

$$\underset{\text{Tcheb.}}{\leqslant}\sum_{k^2\leqslant\ell\leqslant(k+1)^2}\frac{\mathbb{E}((S_\ell-S_{k^2})^2)}{\varepsilon^2k^4}\leqslant\sum_{k^2\leqslant\ell\leqslant(k+1)^2}\frac{C}{\varepsilon^2k^4}\underbrace{(\ell-k^2)}_{\leq k}\leqslant\frac{C'}{k^3}.$$

und $D_k/k^2\to 0$ f.s. (wegen Borel–Cantelli I, punktweise bis auf eine Nullmenge $\mathcal{N}_2$).

Für $n\geqslant 1$ wahlen k, s.d. $k^2\leqslant n<(k+1)^2$, dann

$$\left|\frac{S_n}{n}\right|\leqslant\frac{|S_{k^2}|+D_k}{n}\leqslant\frac{|S_{k^2}|+D_k}{k^2}\to 0\quad f.s.$$

(punktweise bis auf eine Nullmenge $\mathcal{N}_1\cup\mathcal{N}_2$). □

Anwendung. Ein probabilistischer Beweis von Weierstraß Theorem.

Der Satz von (Stone–)Weierstrass besagt, dass stetige Funktionen auf kompakten Mengen über Polynome gleichmässig approximiert werden können.

Satz 14. *Sei $f:[0,1]\to\mathbb{R}$ stetig. Dann für alle $\varepsilon>0$, $\exists n=n(\varepsilon)\geqslant 1$, $\exists$ Polynom B_n vom Grad n, s.d.*

$$\sup_{x\in[0,1]}|f(x)-B_n(x)|\leqslant\varepsilon.$$

Beweis. Sei

$$B_n(x):=\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}f\left(\frac{n}{k}\right).$$

Dann

$$B_n(x)=\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right]$$

wobei $S_n=X_1+\cdots+X_n\sim\mathrm{Bin}(n,x)$, $(X_k)_k$ i.i.d. $X_k\sim\mathrm{Ber}(x)$. Dann

$$|f(x)-B_n(x)|=\left|\mathbb{E}\left[f(x)-f\left(\frac{S_n}{n}\right)\right]\right|\leqslant\mathbb{E}\left[\left|f(x)-f\left(\frac{S_n}{n}\right)\right|\right]$$

Sei $M:=\sup_{x\in[0,1]}|f(x)|$. Da f gleichmässig stetig ist $\exists\delta>0$ s.d. $\forall x,y$ mit $|x-y|<\delta$, $|f(x)-f(y)|\leqslant\varepsilon/2$.

$$|f(x)-B_n(x)|\leqslant\mathbb{E}\left[\left|f(x)-f\left(\frac{S_n}{n}\right)\right|\right]=\mathbb{E}\left[\underbrace{\left|f(x)-f\left(\frac{S_n}{n}\right)\right|}_{\leqslant\varepsilon/2}\mathbb{1}_{|S_n/n-x|<\delta}\right]+\mathbb{E}\left[\underbrace{\left|f(x)-f\left(\frac{S_n}{n}\right)\right|}_{\leqslant 2M}\mathbb{1}_{|S_n/n-x|\geqslant\delta}\right]$$

$$\leqslant\frac{\varepsilon}{2}\underbrace{\mathbb{P}(|S_n/n-x|<\delta)}_{\leqslant 1}+2M\,\mathbb{P}(|S_n/n-x|\geqslant\delta)\leqslant\frac{\varepsilon}{2}+2M\frac{\mathrm{Var}(S_n/n)}{\delta^2}=\frac{\varepsilon}{2}+2M\frac{\overbrace{x(1-x)}^{\leqslant\frac{1}{4}}}{n\delta^2}$$

$$\leqslant\frac{\varepsilon}{2}+\frac{M}{2n\delta^2}\leqslant\varepsilon$$

für $n\geqslant M/\delta^2\varepsilon$. □

6.4 Kolmogorov'sche Ungleichung

Statt $\mathbb{P}(S_n\geqslant a)$, wollen wir eine Ungleichung für

$$\mathbb{P}(S_n^*\geqslant a)$$

wo $S_n^*:=\max_{k\leqslant n}S_k$ ist die laufendes Maximum von die Stochastiche Prozess $(S_n)_{n\geqslant 1}$.

Lemma 15. *(Kolmogorov'sche Ungleichung) Seien $(X_n)_{n\geqslant 1}$ unabhängige Z.V. mit $\mu_k=\mathbb{E}[X_k]$, $\sigma_k^2=\mathrm{Var}(X_k)$. Setze*

$$S_n:=\sum_{k=1}^{n}X_k,\quad m_n:=\sum_{k=1}^{n}\mathbb{E}[X_k]=\mathbb{E}[S_n],$$

und

$$s_n^2:=\sum_{k=1}^{n}\sigma_k^2=\mathrm{Var}(S_n).$$

Dann, für alle $t>0$:

$$\mathbb{P}\left(\exists k\leqslant n\text{ s.d. }|S_k-m_k|\geqslant t s_n\right)\leqslant t^{-2}.$$

Beweis. OBdA: $\mu_k=0$, $k\geqslant 1$, dann $m_k=0$, weil wir können setzen $\tilde{X}_k:=X_k-\mathbb{E}[X_k]$ und $\mathbb{E}[\tilde{X}_k]=0$. Und wir benutzen diese neue Folge. $S_n-\mathbb{E}[S_n]=\tilde{S}_n$.

Wir werden eine sehr interessante Technik der Kürzung anwenden. Sei

$$Y_k := \mathbb{1}_{|S_k| > t\mathfrak{s}_n} \prod_{1 \leqslant \ell < k} \mathbb{1}_{|S_\ell| < t\mathfrak{s}_n} = \mathbb{1}_{\{k = \min\{\ell \geqslant 1 \mid |S_\ell| \geqslant t\mathfrak{s}_n\}\}} \in \{0,1\}$$

$Y_k \in \{0,1\}$ und falls $Y_\ell = 1$ dann $Y_k = 0$ für $k \neq \ell$.

$$Z_n = \mathbb{1}_{\{\exists \ell \leqslant n : |S_\ell| \geqslant t\mathfrak{s}_n\}} = \mathbb{1}_{\{\min\{\ell \geqslant 1 \mid |S_\ell| \geqslant t\mathfrak{s}_n\} \leqslant n\}} = \sum_{k=1}^{n} Y_k \in \{0,1\}.$$

Wir müssen abschätzen

$$\mathbb{P}\left(\exists k \leqslant n \text{ s.d. } |S_k - m_k| \geqslant t\mathfrak{s}_n\right) = \mathbb{E}[Z_n] = \sum_{k=1}^{n} \mathbb{E}[Y_k] \leqslant \sum_{k=1}^{n} \mathbb{E}\left[Y_k \underbrace{\frac{S_k^2}{t^2 \mathfrak{s}_n^2}}_{\geqslant 1 \text{ wenn } Y_k \neq 0} \right] \tag{1}$$

weil, falls $Y_k \neq 0$, dann $|S_k| \geqslant t\mathfrak{s}_n$. Sei $U_k := S_n - S_k = \sum_{\ell=k+1}^{n} X_\ell$ für $k \leqslant n$. Beachten, dass dies unabhängig von S_k, Y_k ist. (hier wir benutzen die unab. von die Folge $(X_n)_n$).

$$\mathbb{E}[Y_k S_n^2] = \mathbb{E}[Y_k (U_k + S_k)^2] = \underbrace{\mathbb{E}[Y_k U_k^2]}_{\geqslant 0} + 2 \underbrace{\mathbb{E}[Y_k S_k U_k]}_{\overset{}{\underset{\text{unab,}}{=}} \mathbb{E}[Y_k S_k]\mathbb{E}[U_k] = 0} + \mathbb{E}[Y_k S_k^2] \geqslant \mathbb{E}[Y_k S_k^2].$$

Zusammen mit Ungleichung (1) das gilt

$$\mathbb{E}[Z_n] \leqslant \frac{1}{t^2 \mathfrak{s}_n^2} \sum_{k=1}^{n} \mathbb{E}[Y_k S_k^2] \leqslant \frac{1}{t^2 \mathfrak{s}_n^2} \sum_{k=1}^{n} \mathbb{E}[Y_k S_n^2] \leqslant \frac{1}{t^2 \mathfrak{s}_n^2} \mathbb{E}\left[\left(\sum_{k=1}^{n} Y_k\right) S_n^2\right] = \frac{1}{t^2 \mathfrak{s}_n^2} \mathbb{E}[Z_n S_n^2] \leqslant \frac{1}{t^2 \mathfrak{s}_n^2} \mathbb{E}[S_n^2] = \frac{1}{t^2}.$$

□

Bemerkung. Vergleiche diese Aussage mit Tchebichev für $S_n - m_n$.
