Stochastik I

10. Große Übung

Niklas Dexheimer

27.11.2019

Erinnerung: Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ $(\mathbb{P})$-fast sicher gegen X, falls

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = \mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{\text{f.s.}} X.$$

- Die fast sichere Konvergenz von Zufallsvariablen ist dazu äquivalent, dass eine $\mathbb{P}$-Nullmenge $N \in \mathcal{A}$ existiert, sodass für alle $\omega \in N^C$

  $$\lim_{n\to\infty} X_n(\omega) = X(\omega)$$

  gilt.
- Die fast sichere Konvergenz entspricht somit der punktweisen Konvergenz von Funktionen fast überall.

Erinnerung: Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{E}[|X|^p] < \infty, \mathbb{E}[|X_n|^p] < \infty, \forall n \in \mathbb{N}$, für ein $p \geq 1$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ im p-ten Mittel (oder in $\mathcal{L}^p$) gegen X, falls

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}^p} X.$$

- Ihr kennt die $\mathcal{L}^p$-Konvergenz schon aus dem Kapitel zu $\mathcal{L}^p$-Räumen, nur wird sie hierbei explizit auf Zufallsvariablen angewendet (denkt daran Erwartungswerte sind Integrale!).

Erinnerung: Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ stochastisch (oder in Wahrscheinlichkeit) gegen X, falls für alle $\epsilon > 0$

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

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{P} X.$$

- Man folgert stochastische Konvergenz häufig direkt aus der Konvergenz im p-ten Mittel mit der Markov-Ungleichung (oder Spezialfall Chebyshev-Ungleichung).

Erinnerung: Konvergenz von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ in Verteilung gegen X, falls für alle stetigen, beschränkten, reell-wertigen Funktionen f (oder kurz $\forall f \in C_b(\mathbb{R})$)

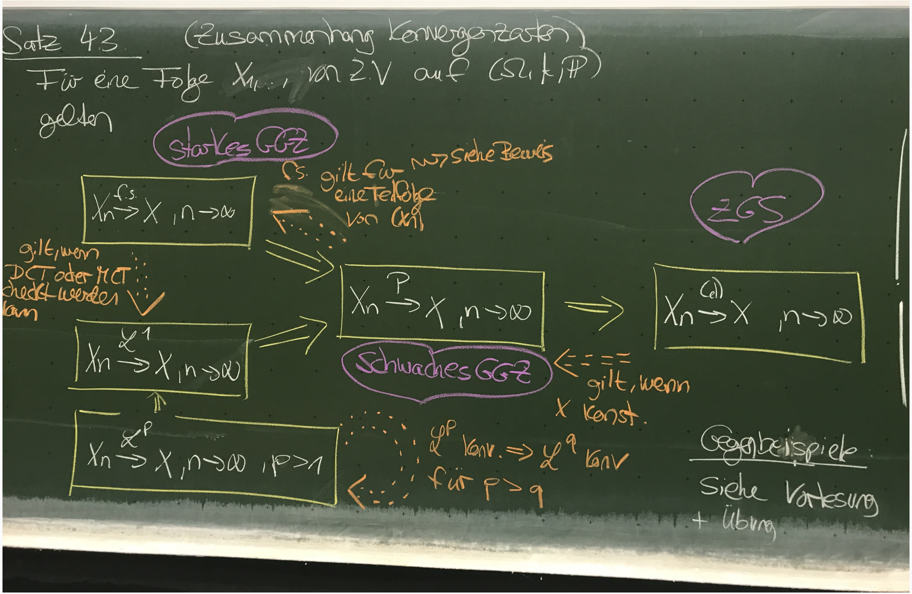
$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$$

gilt. Man schreibt

$$X_n \xrightarrow[n\to\infty]{(d)} X.$$

- Die Konvergenz in Verteilung ist die schwächste (hier behandelte) Art der Konvergenz von Zufallsvariablen.
- Eigentlich ist sie eine Konvergenz von Wahrscheinlichkeitsmaßen und nicht von Zufallsvariablen!

Bildchen (Konvergenzdiagramm)



Verteilungskonvergenz und momenterzeugende Funktionen

Sei X eine Zufallsvariable und $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$, sodass die momenterzeugenden Funktionen $\mathcal{M}_X, \mathcal{M}_{X_n}$ für alle $n \in \mathbb{N}$ existieren, d.h.

$$\mathbb{E}[e^{tX}] < \infty, \quad \mathbb{E}[e^{tX_n}] < \infty, \quad \forall t \in \mathbb{R}, n \in \mathbb{N}.$$

Dann lässt sich zeigen, dass

$$X_n \xrightarrow[n\to\infty]{(d)} X \quad \Leftrightarrow \quad \lim_{n\to\infty} \mathcal{M}_{X_n}(t) = \mathcal{M}_X(t) \quad \forall t \in \mathbb{R}.$$

Mehr dazu nächste Woche in der Vorlesung.

Erinnerung: Hufschlagtote in preußischen Kavallerieregimenten

Ein in der Literatur häufig zitiertes Beispiel zur Poisson-Verteilung ist die Anzahl der Soldaten eines preußischen Kavallerieregiments, die innerhalb eines Jahres an den Folgen eines Huftritts starben. Die Daten wurden für 10 Regimenter über 20 Jahre erfasst.

Anzahl der Todesfälle	0	1	2	3	4	≥ 5
Beobachtete Häufigkeit	109	65	22	3	1	0

Die mittlere Anzahl beträgt dann 0.61. Es ergibt sich dann eine sehr gute Übereinstimmung der relativen Häufigkeiten mit den Wahrscheinlichkeiten der Poi(0.61)-Verteilung.

Anzahl der Todesfälle	0	1	2	3	4
Poi(0.61)($\{k\}$)	0.545	0.325	0.110	0.015	0.005
Beob. relative Häufigkeit	0.543	0.331	0.101	0.021	0.003

Warum Poisson und nicht Binomial?

- Aber warum sollte die Anzahl der Hufschlagtoten Poissonverteilt sein?
- Falls beispielsweise 200 Soldaten in einem Kavallerieregiment sind, beschreibt die Anzahl der Hufschlagtoten die Summe von 200 Experimenten mit Ausgang Tod durch Hufschlag oder Überleben.
- Also sollten wir es eigentlich mit einer Binomialverteilung zu tun haben!
- Die Antwort liefert der Satz von Poisson.

Satz von Poisson

Satz

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum auf dem $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen und X eine Zufallsvariable mit

$$X_n \sim \mathrm{Bin}(n, p_n), \quad X \sim \mathrm{Poi}(\lambda)$$

seien, wobei $(p_n)_{n\in\mathbb{N}}$ eine Folge in $(0, 1)$ sei mit

$$\lim_{n\to\infty} np_n = \lambda$$

für ein $\lambda \in \mathbb{R}$. Dann gilt

$$X_n \xrightarrow[n\to\infty]{(d)} X.$$

Für große n gilt also ungefähr

$$\mathrm{Bin}(n, p) \approx \mathrm{Poi}(np).$$

Aufgabe 1

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum auf dem $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen und X eine Zufallsvariable mit

$$X_n \sim \mathrm{Bin}(n, p_n), \quad X \sim \mathrm{Poi}(\lambda)$$

seien, wobei $(p_n)_{n\in\mathbb{N}}$ eine Folge in $(0, 1)$ sei mit

$$\lim_{n\to\infty} np_n = \lambda$$

für ein $\lambda \in \mathbb{R}$. Zeige, dass

$$X_n \xrightarrow[n\to\infty]{(d)} X$$

gilt. Berechne dazu zunächst $\mathcal{M}_{X_n}, \mathcal{M}_X$ und nutze die Aussage zur Konvergenz von momenterzeugenden Funktionen aus und, dass für eine Folge $(a_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = a$

$$\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = e^a$$

gilt.

Beweis des Hinweises

Sei $\epsilon > 0$ gegeben. Dann existiert ein $N_0 \in \mathbb{N}$, sodass

$$\left(1 + \frac{a-\epsilon}{n}\right)^n \leq \left(1 + \frac{a_n}{n}\right)^n \leq \left(1 + \frac{a+\epsilon}{n}\right)^n, \quad \forall n \geq N_0.$$

Also gilt

$$\liminf_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n \geq \liminf_{n\to\infty} \left(1 + \frac{a-\epsilon}{n}\right)^n = e^{a-\epsilon}$$

und

$$\limsup_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n \leq \limsup_{n\to\infty} \left(1 + \frac{a+\epsilon}{n}\right)^n = e^{a+\epsilon}.$$

Da $\epsilon > 0$ beliebig gewählt war folgt somit

$$e^a \leq \liminf_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n \leq \limsup_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n \leq e^a$$

und daraus folgt

$$\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = e^a.$$

Lösung

Per Definition gilt für $n \in \mathbb{N}$, $t \in \mathbb{R}$

$$\begin{aligned} \mathcal{M}_{X_n}(t) = \mathbb{E}[e^{tX_n}] &= \sum_{k=0}^{n} \binom{n}{k} e^{tk} p_n^k (1 - p_n)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (e^t p_n)^k (1 - p_n)^{n-k} \\ &= (e^t p_n + 1 - p_n)^n \\ &= (1 + p_n(e^t - 1))^n, \end{aligned}$$

wobei wir den Binomischen Lehrsatz benutzt haben.

Lösung

Des Weiteren gilt für $t \in \mathbb{R}$

$$\begin{aligned}
\mathcal{M}_X(t) = \mathbb{E}[e^{tX}] &= e^{-\lambda} \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!} \\
&= e^{-\lambda} e^{e^t \lambda} \\
&= e^{\lambda(e^t - 1)},
\end{aligned}$$

wobei wir die Definition der Exponentialreihe benutzt haben.

Lösung

Also folgt für $t \in \mathbb{R}$

$$\begin{aligned}\mathcal{M}_{X_n}(t) &= (1 + p_n(e^t - 1))^n \\ &= \left(1 + \frac{np_n(e^t - 1)}{n}\right)^n \\ &\underset{n\to\infty}{\longrightarrow} e^{\lambda(e^t - 1)} \\ &= \mathcal{M}_X(t)\end{aligned}$$

da per Voraussetzung

$$\lim_{n\to\infty} np_n(e^t - 1) = \lambda(e^t - 1)$$

gilt und der letzte Schritt aus dem Hinweis folgt. Also folgt die Behauptung, da $\mathcal{M}_{X_n} < \infty, \mathcal{M}_X < \infty$ gilt und somit die Konvergenz der momenterzeugenden Funktionen die Konvergenz in Verteilung impliziert.

Aufgabe 2

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum auf dem die Folgen von Zufallsvariablen $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ und eine Zufallsvariable Z gegeben seien mit

$$X_n \sim \mathrm{Exp}(n), \quad \mathbb{E}[Y_n] = \frac{1}{\sqrt{n}}, \quad \mathbb{V}[Y_n] = \frac{\sigma^2}{n},$$

für $\sigma > 0$. Zeige

$$X_n \xrightarrow[n\to\infty]{P} 0, \quad Y_n + Z \xrightarrow[n\to\infty]{P} Z.$$

Sei $\epsilon > 0$. Dann gilt für $n \in \mathbb{N}$

$$\begin{aligned}\mathbb{P}(|X_n - 0| > \epsilon) &= \mathbb{P}(X_n > \epsilon)\\ &= 1 - \mathbb{P}(X_n \leq \epsilon)\\ &= 1 - F_{X_n}(\epsilon)\\ &= 1 - (1 - e^{-n\epsilon})\\ &= e^{-n\epsilon},\end{aligned}$$

und somit folgt

$$\lim_{n\longrightarrow\infty} \mathbb{P}(|X_n - 0| > \epsilon) = 0,$$

für alle $\epsilon > 0$, was per Definition

$$X_n \xrightarrow[n\longrightarrow\infty]{P} 0$$

impliziert.

Lösung

Sei $\epsilon > 0$. Dann gilt für $n \in \mathbb{N}$ mit der Chebyshev-Ungleichung und der Verschiebungsformel

$$\begin{aligned}
\mathbb{P}(|Z + Y_n - Z| > \epsilon) = \mathbb{P}(|Y_n| > \epsilon) &\leq \frac{\mathbb{E}[|Y_n|^2]}{\epsilon^2} \\
&= \frac{\mathbb{V}(Y_n) + \mathbb{E}[Y_n]^2}{\epsilon^2} \\
&= \frac{\sigma^2 + 1}{n\epsilon^2}
\end{aligned}$$

und somit folgt

$$\lim_{n\to\infty} \mathbb{P}(|Z + Y_n - Z| > \epsilon) = 0$$

für alle $\epsilon > 0$, was per Definition

$$Z + Y_n \xrightarrow[n\to\infty]{P} Z$$

impliziert.