Dr. Stefan Richter, Timo Dörzbach
Wahrscheinlichkeitstheorie 2
Sommersemester 2023

Präsenzblatt 6 - Lösungen

Aufgabe P1 (Maximalungleichungen für normalverteilte Zufallsvariablen).
In dieser Aufgabe sollen weitere Maximalungleichungen gezeigt werden.

(a) Seien $Z_1, ..., Z_M$ Zufallsvariablen mit

$$\max_{j=1,\dots,M} \mathbb{E}[\exp(\alpha Z_j^2)] \le C$$

für gewisse $\alpha, C > 0$. Dann gilt:

$$\mathbb{E}\left[\max_{j=1,\dots,M} Z_j^2\right] \le \frac{1}{\alpha}\log(MC).$$

(b) Zeigen Sie, dass für eine normalverteilte Zufallsvariable $Z \sim N(0, \sigma^2)$ und $\alpha < \frac{1}{2\sigma^2}$ gilt:

$$\mathbb{E}\exp(\alpha \cdot Z^2) = \frac{1}{\sqrt{1 - 2\alpha\sigma^2}}.$$

(c) Seien $Z_1, ..., Z_M$, $Z_j = (Z_{j1}, ..., Z_{jd})^T$ d-dimensional normalverteilt mit $\mathbb{E}[Z_j] = 0$ und

$$\max_{j=1,\dots,M} \max_{k=1,\dots,d} \mathbb{E}[Z_{jk}^2] =: \sigma_{max}^2 < \infty.$$

Dann gilt:

$$\mathbb{E}\Big[\max_{j=1,\dots,M} \|Z_j\|^2\Big] \le 4d \cdot \sigma_{max}^2 \log(\sqrt{2}Md).$$

wobei $\|\cdot\|$ die euklidische Norm bezeichnet.
Hinweis: Verwenden Sie (a) und (b) mit $\alpha = \frac{1}{4\sigma_{max}^2}$.

Lösung:

(a) Es ist

$$\begin{aligned}
\mathbb{E}\left[\max_{j=1,\ldots,M} Z_j^2\right] &= \frac{1}{\alpha}\mathbb{E}\left[\log\Big(\exp(\alpha \max_{j=1,\ldots,M} Z_j^2)\Big)\right] \\
&\overset{\text{Jensen}}{\leq} \frac{1}{\alpha}\cdot\log\mathbb{E}\left[\exp\Big(\alpha \max_{j=1,\ldots,M} Z_j^2\Big)\right] \\
&\overset{\text{exp mon. wachs.}}{\leq} \frac{1}{\alpha}\cdot\log\mathbb{E}\left[\max_{j=1,\ldots,M}\Big(\exp(\alpha Z_j^2)\Big)\right] \\
&\leq \frac{1}{\alpha}\log\left(\mathbb{E}\left[\sum_{j=1}^{M}\exp(\alpha Z_j^2)\right]\right) \\
&\leq \frac{1}{\alpha}\log\left(M\cdot \underbrace{\max_{j=1,\ldots M}\mathbb{E}[\exp(\alpha Z_j^2)]}_{\leq C}\right) \\
&\leq \frac{1}{\alpha}\log(MC).
\end{aligned}$$

(b) Es gilt

$$\begin{aligned}
\mathbb{E}[\exp(\alpha Z_{jk}^2)] &= \int e^{\alpha z^2}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{z^2}{2\sigma^2}}\,\mathrm{d}z \\
&= \frac{1}{\sqrt{\sigma^2}\cdot\sqrt{\frac{1}{\sigma^2}-2\alpha}}\cdot \underbrace{\int \frac{1}{\sqrt{2\pi\left(\frac{1}{\sigma^2}-2\alpha\right)^{-1}}}\cdot e^{-\frac{z^2}{2}\cdot\left(\frac{1}{\sigma^2}-2\alpha\right)}\,\mathrm{d}z}_{=1,\text{ da Dichte der } N\left(0,\left(\frac{1}{\sigma^2}-2\alpha\right)^{-1}\right)\text{-Verteilung}} \\
&= \frac{1}{\sqrt{1-2\alpha\sigma^2}}
\end{aligned}$$

(b) Wir wollen (a) benutzen. Die Zufallsvariablen $\|Z_j\|^2$ sind aber zu kompliziert, um (a) direkt anwenden zu können. Unter Beachtung von

$$\|Z_j\|^2 = \sum_{k=1}^{d} Z_{jk}^2 \leq d\cdot\max_{k=1,\ldots,d} Z_{jk}^2$$

sehen wir aber, dass gilt:

$$\mathbb{E}[\max_{j=1,\ldots,M}\|Z_j\|^2] \leq d\cdot\mathbb{E}[\max_{j=1,\ldots,M;k=1,\ldots,d} Z_{jk}^2].$$

Wir wenden nun das Lemma 3.11.(a) auf die $M\cdot d$ Zufallsvariablen $(Z_{jk})_{j=1,\ldots,M;k=1,\ldots,d}$ an. Können wir also $\alpha, C>0$ finden, sodass

$$\max_{j=1,\ldots,M;k=1,\ldots,d}\mathbb{E}[\exp(\alpha Z_{jk}^2)] \leq C \tag{1}$$

gilt, so folgt:

$$\mathbb{E}[\max_{j=1,\ldots,M;k=1,\ldots,d} Z_{jk}^2] \leq \frac{1}{\alpha}\log((M\cdot d)\cdot C)$$

und damit:
$$\mathbb{E}[\max_{j=1,...,M} \|Z_j\|^2] \le \frac{d}{\alpha}\log((M\cdot d)\cdot C).$$
Ein Vergleich mit der Schranke aus der Aufgabenstellung erlaubt die Vermutung, dass $\alpha = \frac{1}{4\sigma_{max}^2}$ und $C = \sqrt{2}$ gewählt werden muss. Wir verifizieren (**??**) mit diesen Setzungen. Mit $\sigma_{jk}^2 = \mathbb{E}[Z_{jk}^2]$ ist für beliebiges $j = 1, ..., M$, $k = 1, ..., d$:

$$\begin{aligned}
\mathbb{E}[\exp(\alpha Z_{jk}^2)] &\overset{(b)}{=} \frac{1}{\sqrt{1-2\alpha\sigma_{jk}^2}} \\
&= \frac{1}{\sqrt{1-\frac{1}{2}\underbrace{\frac{\sigma_{jk}^2}{\sigma_{max}^2}}_{\le 1}}} \\
&\le \frac{1}{\sqrt{1-\frac{1}{2}}} = \sqrt{2} = C.
\end{aligned}$$

Damit ist alles gezeigt.

Aufgabe P2 (Maximalungleichungen bei nur polynomiellen Momenten).

Sinn dieser Aufgabe ist es, dass Sie sehen, dass auch Maximalungleichungen bewiesen werden können, wenn nur polynomielle Momente vorliegen. In diesem Fall ist allerdings die Qualität der oberen Schranke entsprechend schlechter.

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $(\mathcal{X}, d)$ ein metrischer Raum und $X_i : \Omega \to \mathcal{X}$, $i \in \mathbb{N}$ i.i.d. Zufallsvariablen. Sei $\mathcal{F} = \{f : \mathcal{X} \to \mathbb{R} \text{ messbar}\}$ mit der Eigenschaft, dass ein $M > 1$ und ein $\alpha > 2$ existiert, so dass
$$\forall f \in \mathcal{F}: \quad \|f(X_1)\|_\alpha^\alpha := \mathbb{E}[|f(X_1)|^\alpha] \le M.$$
Sei $\mathcal{F}$ endlich. Zeigen Sie, dass für alle $\beta < \alpha$ gilt:
$$\mathbb{E}\max_{f\in\mathcal{F}} |\mathbb{G}_n(f)| \le 2(\alpha-1)^{1/2} M^{1/\alpha}\cdot|\mathcal{F}|^{1/\alpha}$$
Hinweis: Nutzen Sie ohne Beweis, dass für $\alpha > 2$ gilt: $\|\mathbb{G}_n(f)\|_\alpha \le 2\sqrt{\alpha-1}\|f(X_1)\|_\alpha$ (sog. Burkholder-Ungleichung für i.i.d. Variablen).

Lösung: Definiere die Funktion
$$\psi(x) = x^\alpha.$$
Diese ist konvex und streng monoton wachsend. Es gilt:
$$\mathbb{E}\psi(|\mathbb{G}_n(f)|) = \|\mathbb{G}_n(f)\|_\alpha^\alpha \overset{\text{Hinweis}}{\le} 2^\alpha(\alpha-1)^{\alpha/2}M.$$
Es folgt
$$\psi\Big(\mathbb{E}\max_{f\in\mathcal{F}}|\mathbb{G}_n(f)|\Big) \overset{\psi \text{ konvex}}{\le} \mathbb{E}\max_{f\in\mathcal{F}}\psi(|\mathbb{G}_n(f)|) \le \sum_{f\in\mathcal{F}}\mathbb{E}\psi(|\mathbb{G}_n(f)|) \le |\mathcal{F}|\cdot 2^\alpha(\alpha-1)^{\alpha/2}M,$$
also wegen $\psi^{-1}(y) = y^{1/\alpha}$:
$$\mathbb{E}\max_{f\in\mathcal{F}}|\mathbb{G}_n(f)| \le \big(|\mathcal{F}|\cdot 2^\alpha(\alpha-1)^{\alpha/2}M\big)^{1/\alpha} \le |\mathcal{F}|^{1/\alpha}\cdot 2(\alpha-1)^{1/2}M^{1/\alpha}.$$

Aufgabe P3 (Anwendungen des Satzes von Glivenko-Cantelli).

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X_i : \Omega \to \mathcal{X} = \mathbb{R}$, $i \in \mathbb{N}$ i.i.d. Zufallsvariablen mit Verteilungsfunktion $F(x) = \mathbb{P}(X \leq x)$. Es sei

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}}$$

die empirische Verteilungsfunktion.

(a) Zeigen Sie mit Hilfe des Satzes von Glivenko-Cantelli (Blatt 6, Aufgabe 3):

$$\sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F(x) \right| \to 0 \quad \text{f.s.}$$

(b) Sei nun $p \in (0,1)$ und F stetig und streng monoton wachsend, $x_p = F^{-1}(p)$ das p-Quantil der Verteilung von X.
Bezeichne $X_{\lfloor np \rfloor:n}$ das empirische p-Quantil der Stichprobe (hierbei sind $X_{1:n} \leq ... \leq X_{n:n}$ die Ordnungsstatistiken von $X_1, ..., X_n$). Zeigen Sie, dass

$$X_{\lfloor np \rfloor:n} \to x_p \quad \text{f.s.}$$

Hinweis: Da F stetig und streng monoton wachsend, ist auch F^{-1} stetig.

Lösung:

(a) Laut Vorlesung gilt für die Funktionenklasse

$$\mathcal{F} = \{f_x = \mathbb{1}_{\{\cdot \leq x\}} : x \in \mathbb{R}\},$$

für alle $\varepsilon > 0$, dass $N(\varepsilon, \mathcal{F}, \mathbb{P}^{X_1}) \leq \frac{2}{\varepsilon^2} < \infty$.
Außerdem ist für alle $x \in \mathbb{R}$:

$$\mathbb{E}|f_x(X_1)| = \mathbb{E}[\mathbb{1}_{\{X_1 \leq x\}}] = \mathbb{P}(X_1 \leq x) = F(x) \leq 1 < \infty.$$

Damit ist der Satz von Glivenko-Cantelli (Blatt 6, Aufgabe 3) auf diese Funktionenklasse anwendbar und liefert

$$\sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| = \sup_{x \in \mathbb{R}} |\hat{P}_n f_x - P f_x| = \sup_{f \in \mathcal{F}} |\hat{P}_n f - P f| \to 0 \quad \text{f.s.}$$

(b) Es gilt

$$\hat{F}_n(X_{\lfloor np \rfloor:n}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq X_{\lfloor np \rfloor:n}} = \frac{\lfloor np \rfloor}{n}$$

(Anzahl der X_i kleiner als $X_{\lfloor np \rfloor:n}$ ist $\lfloor np \rfloor$).
Damit folgt (beachte: $F(x_p) = p$):

$$\begin{aligned}
\left|F(X_{\lfloor np \rfloor:n}) - F(x_p)\right| &\leq \left|F(X_{\lfloor np \rfloor:n}) - \hat{F}_n(X_{\lfloor np \rfloor:n})\right| + \left|\hat{F}_n(X_{\lfloor np \rfloor:n}) - F(x_p)\right| \\
&\leq \underbrace{\sup_{x \in \mathbb{R}} \left|\hat{F}_n(x) - F(x)\right|}_{\to 0} + \underbrace{\left|\frac{\lfloor np \rfloor}{n} - p\right|}_{= \frac{|\lfloor np \rfloor - np|}{n} \leq \frac{1}{n} \to 0} \\
&\to 0 \quad \text{f.s.}
\end{aligned}$$

also

$$F(X_{\lfloor np \rfloor : n}) \to F(x_p) \quad \text{f.s.}$$

Da F^{-1} stetig ist, folgt

$$X_{\lfloor np \rfloor : n} = F^{-1}(F(X_{\lfloor np \rfloor : n})) \to F^{-1}(F(x_p)) = x_p \quad \text{f.s.}$$

Vorlesungswebseite:
`https://ssp.math.uni-heidelberg.de/wt2-ss23/`