Konzentrationsungleichungen

Übungsblatt 1

TU Dortmund, Sommersemester 2021

Prof. Dr. Ivan Veselić

Wichtig: Für diesen Zettel ist es nützlich, dass zusätzlich zu Aufgabe 3 c) i) die Funktion ψ auf $[0, b)$ sogar C^∞ ist, falls $X \geq 0$ oder alle absoluten Momente von X endlich sind!

Übung 1. *Sei Y binomialverteilt mit Parametern (n, p), wobei $p \geq 1/2$. Zeigen Sie:*

$$\mathbb{P}\left(Y - np \geq n\epsilon\right) \leq \mathrm{e}^{-\frac{n\epsilon^2}{2p(1-p)}}, \quad 0 < \epsilon < 1 - p.$$

Beweis. Wie in der Vorlesung gesehen gilt für die KEF einer zentrierten Bernoulli-ZV Z

$$\psi_Z(\lambda) = -\lambda p + \ln\left(p\mathrm{e}^\lambda + 1 - p\right)$$

und somit auch

$$\psi_Z'(\lambda) = -p + \frac{1}{1 + \frac{1-p}{p}\mathrm{e}^{-\lambda}}.$$

Über $t \overset{!}{=} \psi_X'(\lambda_t)$ folgt

$$\lambda_t = -\ln\left(\frac{1-p-t}{1-p} \cdot \frac{p}{t+p}\right).$$

Daraus folgt

$$\begin{aligned}
\psi_Z^*(t) &= t\lambda_t - \psi_Z(\lambda_t) \\
&= -(p+t) \cdot \ln\left(\frac{1-p-t}{1-p}\frac{p}{t+p}\right) - \ln\left(\frac{1-p}{1-p-t} \cdot (p+t) + 1 - p\right) \\
&= (1-p-t)\ln\left(\frac{1-p-t}{1-p}\right) + (p+t)\left(\frac{p+t}{p}\right)
\end{aligned}$$

Zusammen mit dem Ergebnis zu Summen von unabhängigen ZV aus der Vorlesung folgt nun für Y

$$\mathbb{P}\left(Y - np \geq n\epsilon\right) \leq \mathrm{e}^{-nh_p\left(\frac{\epsilon n}{n}+p\right)} = \mathrm{e}^{-nh_p(\epsilon+p)}$$

mit der Kullbach-Leibler-Divergenz

$$h_p(a) = a\ln(a/p) + (1-a)\ln((1-a)/(1-p)).$$

Für diese gilt:

Lemma 0.1. $h_p(\epsilon + p) \geq \frac{\epsilon^2}{2p(1-p)}$*, falls $\epsilon + p < 1$ und $p \geq 1/2$.*

Beweis des Lemma. Für $\epsilon = 0$ gilt die Ungleichung, denn

$$p\ln(1) + (1-p)\ln(1) = 0 = \frac{0^2}{2p(1-p)}.$$

Wir leiten nun beide Seiten nach ϵ ab und müssen nun zeigen:

$$h_p'(\epsilon+p) = \ln\left(\frac{\epsilon+p}{p}\right) + 1 - \ln\left(\frac{1-\epsilon-p}{1-p}\right) - 1 = \ln\left(\frac{\epsilon+p}{p}\right) - \ln\left(\frac{1-\epsilon-p}{1-p}\right) \geq \frac{\epsilon}{p(1-p)}.$$

Bei $\epsilon = 0$ ist dies wieder korrekt und wir müssen nur noch zeigen, dass

$$h_p''(\epsilon + p) = \frac{1}{(\epsilon + p)(1 - \epsilon - p)} \geq \frac{1}{p(1-p)}.$$

oder entsprechend

$$(\epsilon + p)(1 - \epsilon - p) \leq p(1-p).$$

Da $a \mapsto a(1-a)$ auf $[p, 1] \subset [1/2, 1]$ monoton fällt, ist dies korrekt. □

Aus dem Lemma folgt dann

$$\mathbb{P}\left(Y - np \geq n\epsilon\right) \leq \mathrm{e}^{-n\frac{\epsilon^2}{p(1-p)}}.$$

□

Übung 2. *Sei X eine normalverteilte Zufallsvariable mit Erwartungswert μ und Varianz ν, d.h. mit Dichte $f(x) = \frac{1}{\sqrt{2\pi\nu}} \exp\left(-\frac{(x-\mu)^2}{2\nu}\right)$.*

- *Zeigen Sie*

$$M_X(\lambda) = \mathrm{e}^{\mu\lambda + \frac{\nu\lambda^2}{2}}.$$

- *Sei nun X eine zentrierte, normalverteilte Zufallsvariable mit Varianz ν. Zeigen sie:*

$$\sup_{t>0} f(t) := \sup_{t>0} \left(\mathbb{P}(X \geq t) \exp\left(\frac{t^2}{2\nu}\right)\right) = \frac{1}{2}.$$

Beweis.

- Wir setzen die Definition der MEF ein und berechnen

$$\begin{aligned}
\mathbb{E}\left(\mathrm{e}^{\lambda X}\right) &= \int_{-\infty}^{\infty} \exp(\lambda x) \frac{1}{\sqrt{2\pi\nu}} \exp\left(-\frac{(x-\mu)^2}{2\nu}\right) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\nu}} \exp\left(-\frac{x^2 - (2\mu + 2\nu\lambda)x + \mu^2}{2\nu}\right) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\nu}} \exp\left(-\frac{(x - (\mu + \nu\lambda))^2 - (\mu + \nu\lambda)^2 + \mu^2}{2\nu}\right) \mathrm{d}x \\
&= \exp\left(\mu\lambda + \frac{\nu\lambda^2}{2}\right)
\end{aligned}$$

- Falls X zentriert ist, ist somit die Cramér-Transformierte

$$\psi_X^*(t) = \sup_{\lambda \geq 0} \left(\lambda t - \psi_X(\lambda)\right) = \sup_{\lambda \geq 0} \left(\lambda t - \frac{\nu\lambda^2}{2}\right) = \frac{t^2}{2\nu}.$$

  Nach der Chernoff-Ungleichung ist damit

$$\mathbb{P}(X \geq t) \leq \exp\left(\psi_X^*(t)\right) = \exp\left(\frac{t}{2\nu}\right).$$

  Also bekommt man mit der Chernoff-Abschätzung die Schranke 1. Die Aufgabe zeigt nun im wesentlichen, dass diese Schranke “gar nicht so schlecht” ist (anstatt 1/2, wie es sein sollte, bekommt man mit Chernoff eben 1).

Wir unterscheiden zwei Fälle: Fall 1: $t \geq \frac{2\sqrt{\nu}}{\sqrt{2\pi}}$.
Dann haben wir

$$\begin{aligned}\mathbb{P}(X \geq t) &= \frac{1}{\sqrt{2\pi\nu}} \int_t^\infty \exp\left(-\frac{x^2}{2\nu}\right) \mathrm{d}x \leq \frac{1}{\sqrt{2\pi\nu}} \int_t^\infty \frac{x}{t} \exp\left(-\frac{x^2}{2\nu}\right) \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi\nu}} \frac{\nu}{t} \exp\left(-\frac{t^2}{2\nu}\right) \leq \frac{\exp\left(-\frac{t^2}{2\nu}\right)}{2}.\end{aligned}$$

Wir müssen also noch kleine t verstehen.
Fall 2: $0 \leq t \leq \frac{2\sqrt{\nu}}{\sqrt{2\pi}}$
Definiere $f(t) := \mathbb{P}(X \geq t) \exp\left(\frac{t^2}{2\nu}\right)$. Es ist $f(0) = 1/2$. Es bleibt also zu zeigen:

$$f(t) \leq \frac{1}{2} = f(0), \quad 0 \leq t \leq \frac{2\sqrt{\nu}}{\sqrt{2\pi}}.$$

Wir berechnen

$$f'(t) = \frac{-1}{\sqrt{2\pi\nu}} + \frac{t}{\nu} f(t).$$

Nach der in obiger Vorbemerkung berechneten Chernoff-Schranke, haben wir $f(t) \leq 1$, womit

$$f'(t) \leq \frac{-1}{\sqrt{2\pi\nu}} + \frac{t}{\nu}$$

Nun ist

$$f(t) - f(0) = \int_0^t f'(s) \mathrm{d}s \leq \int_0^t \frac{-1}{\sqrt{2\pi\nu}} + \frac{s}{\nu} \mathrm{d}s = -\frac{t}{\sqrt{2\pi\nu}} + \frac{t^2}{2\nu} = \frac{1}{2\nu} \cdot t \cdot (t - \frac{2\sqrt{\nu}}{\sqrt{2\pi}})$$

Dies ist ≤ 0 für $t \in [0, \frac{2\sqrt{\nu}}{\sqrt{2\pi}}]$.

$\square$

Übung 3. *Zeigen Sie:*

i) Ist $X \in \mathcal{G}(\nu)$ und $\lambda \in \mathbb{R}$, so folgt $\mathbb{E}(\mathrm{e}^{\lambda X}) \geq 1$.

ii) Ist $X \in \mathcal{G}(\nu)$, so ist $\mathrm{Var}(X) \leq \nu$.

Beweis. i) Wir wenden die Jensen-Ungleichung an:

$$1 = \exp(0) = \exp\left(\mathbb{E}[\lambda X]\right) \leq \mathbb{E}\left[\exp(\lambda X)\right].$$

Bem: Wir haben am Ende von Satz 2.8 genutzt, dass $\mathbb{E}[\mathrm{e}^{-\lambda X}] \geq 1$.

ii) Da X zentriert ist und wie in der VL gezeigt alle absoluten Momente existieren, ist $M \in C^\infty$ und es gilt $\mathrm{Var}(X) = \mathbb{E}(X^2) = \frac{\partial^2}{\partial\lambda^2} M_X(\lambda) \mid_{\lambda=0}$ sowie $M_X(0) = 0$ und $\frac{\partial}{\partial\lambda} M_X(\lambda) \mid_{\lambda=0} = 0$.
Aufgrund von $\Psi_X''(\lambda) = -\frac{(M_X'(\lambda))^2}{(M_X(\lambda))^2} + \frac{M_X''(\lambda)}{M_X(\lambda)}$ folgt damit

$$\mathrm{Var}(X) = \mathbb{E}(X^2) = \frac{\partial^2}{\partial\lambda^2} M_X(\lambda) \mid_{\lambda=0} = \frac{\partial^2}{\partial\lambda^2} \psi_X(\lambda) \mid_{\lambda=0}$$

Wir wählen zunächst ein $b > 0$, so dass $\psi_X(\lambda)$ eine C^∞-Funktion auf $[0, b]$ ist (solch ein b existiert wie oben diskutiert). Wie in der Vorlesung gezeigt gilt weiterhin: $\psi_X(0) = \psi'_X(0) = 0$. Taylorentwicklung führt dann zu:

$$\psi(\lambda) = \psi(0) + \lambda\psi(0) + \frac{\lambda^2}{2}\psi''(0) + \frac{\lambda^3}{6}\psi'''(\theta)$$

für ein $\theta \in [0, b]$. Damit gilt

$$\psi''(0) = \frac{2\psi(\lambda)}{\lambda^2} - \frac{\lambda}{3}\psi'''(\theta) \leq \frac{2\nu\lambda^2}{2\lambda^2} + \frac{\lambda}{3}\max_{\theta\in[0,b]}|\psi'''(\theta)| \to \nu$$

für $\lambda \to 0$.

□

Übung 4. *Eine Zufallsvariable X heißt* exponentialverteilt mit Parameter $a > 0$, *wenn ihre Verteilung die Dichte $a\cdot\mathrm{e}^{-ax}$, $x \geq 0$, hat. Eine Zufallsvariable Y heißt* sub-exponentialverteilt, *wenn ein $a > 0$ existiert, so dass für alle $\lambda \in (0, a)$ gilt:*

$$\mathbb{E}\left(\mathrm{e}^{\lambda Y}\right) \leq \frac{1}{1 - \lambda/a}.$$

a) Sei X exponentialverteilt mit Parameter a. Zeigen Sie, dass für alle $q \in \mathbb{N}$ gilt:

$$\mathbb{E}\left(X^q\right) = \frac{q!}{a^q}.$$

b) Sei $Y \geq 0$ sub-exponentialverteilt. Zeigen Sie, dass für jedes $q \in \mathbb{N}$ gilt:

$$\mathbb{E}\left(Y^q\right) \leq 2^{q+1}\frac{q!}{a^q},$$

wobei $a > 0$ die Konstante aus obiger Definition ist.

c) Sei X exponentialverteilt mit Parameter a. Kann $X - \mathbb{E}[X]$ subgaußsch sein?

Beweis. a) Es gilt

$$\mathbb{E}[X^q] = \frac{\partial^q}{\partial\lambda^q}\mathbb{E}[\exp(\lambda X)]\mid_{\lambda=0}$$

da $X \geq 0$ und somit $M_X(\lambda)$ unendlich oft stetig differenzierbar ist.
Wir berechnen zunächst für $\lambda < a$:

$$\begin{aligned}\mathbb{E}[\exp(\lambda X)] &= a\int_0^\infty \exp(\lambda x)\exp(-ax)\mathrm{d}x = a\int_0^\infty \exp((-a+\lambda)x)\mathrm{d}x \\ &= \frac{a}{a-\lambda} = \frac{1}{1-\lambda/a}.\end{aligned}$$

Weiterhin folgt mit Induktion

$$\frac{\partial^q}{\partial\lambda^q}\frac{1}{1-\lambda/a} = \frac{q!}{a^q}\frac{1}{(1-\lambda/a)^{q+1}}$$

woraus das Ergebnis folgt.

b) Für $0 < \lambda < a$ gilt (wobei wir $Y \geq 0$ nutzen, um die Reihendarstellung der Exponentialfunktion zu erhalten)

$$\mathbb{E}[Y^q] = \frac{q!}{\lambda^q}\mathbb{E}\left[\frac{(\lambda Y)^q}{q!}\right] \leq \frac{q!}{\lambda^q}\mathbb{E}[\mathrm{e}^{\lambda Y}] \leq \frac{q!}{\lambda^q}\frac{1}{1-\lambda/a}.$$

Die Wahl $\lambda = a/2$ führt dann auf

$$\mathbb{E}[Y^q] \leq 2^q \frac{q!}{a^q}\frac{1}{1-1/2}.$$

c) Es gilt $\mathbb{E}[Y] = 1/a$. Für $\lambda \geq 0$ gilt:

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{\lambda(Y-1/a)}] &= \mathrm{e}^{-\lambda/a}\int_0^\infty a\mathrm{e}^{\lambda x}\cdot \mathrm{e}^{-ax}\mathrm{d}x = \mathrm{e}^{-\lambda/a}a\int_0^\infty \mathrm{e}^{-(a-\lambda)x}\mathrm{d}x \\ &= \begin{cases}\mathrm{e}^{-\lambda/a}\frac{a}{a-\lambda} & 0 \leq \lambda < a, \\ \infty & \lambda \geq a.\end{cases}\end{aligned}$$

Somit folgt

$$\psi(\lambda) = \begin{cases}\ldots & 0 \leq \lambda < a, \\ \infty & \lambda \geq a.\end{cases}$$

Daher kann $\psi(\lambda) \leq \lambda^2\nu/2$ nie gelten (ψ ist ∞ in $[a,\infty)$), und daher ist *keine* exponentialverteilte ZV sub-Gaußsch.

□

Abgabe bis zum 03.05.2021