Konzentrationsungleichungen

Übungsblatt 10

TU Dortmund, Sommersemester 2021

Prof. Dr. Ivan Veselić

Übung 1 (Penrose's square root law). Erinnerung: *Seien $X_1, \ldots, X_n$ i.i.d. Rademacher-Zufallsvariablen (d.h. $\mathbb{P}(X_i = \pm 1) = 1/2$). Dann ist der Vektor $X = (X_1, \ldots, X_n)$ gleichverteilt auf dem binären Hyperwürfel $BHW := \{-1, +1\}^n$. Bezeichne $\overline{X}^{(i)} = (X_1, \ldots, X_{i-1}, -X_i, X_{i+1}, \ldots, X_n)$ den Vektor X mit geflippter i-ter Koordinate. Für $A \subset BHW$ ist der* Einfluss der i-ten Variable auf das Ereignis A *definiert als*

$$I_i(A) := \mathbb{P}\left(\mathbb{1}_{\{X \in A\}} \neq \mathbb{1}_{\{\overline{X}^{(i)} \in A\}}\right).$$

a) Sei $A := \{x \in BHW : \sum_{i=1}^n x_i \geq 0\}$. Zeigen Sie, dass für jedes $i \in \{1, \ldots, n\}$ gilt: $I_i(A) \approx \frac{\text{Const}}{\sqrt{n}}$ für $n \to \infty$.
(Hinweis: Sie dürfen sich der Einfachkeit halber auf ungerade n beschränken.)

b) Interpretieren Sie das Ergebnis aus Teil a) im Kontext demokratischer Abstimmungen.

Beweis. a) Sei oBdA $n = 2m + 1$. Dann gilt

$$I_i(A) = \mathbb{P}\left(\sum_{k \neq i} X_i = 0 \text{ or } -1\right) = \mathbb{P}\left(\sum_{k \neq i} X_i = 0\right),$$

denn es gibt genau $2m$ von i verschiedene Indizes und die Summe der entsprechenden X_i muss gerade sein. Dies ist aber gleich

$$\left(\frac{1}{2}\right)^{2m} \frac{(2m)!}{m!m!} \approx \left(\frac{1}{2}\right)^{2m} \frac{\sqrt{2\pi \cdot 2m}(2m/\mathrm{e})^{2m}}{\sqrt{2\pi m}^2 (m/\mathrm{e})^{2m}} = \left(\frac{1}{2}\right)^{2m} 2^{2m} \frac{\sqrt{2}}{\sqrt{2\pi m}}$$
$$= \sqrt{1/\pi} \frac{1}{\sqrt{m}} \approx \sqrt{2/\pi} \frac{1}{\sqrt{n}}.$$

wobei wir die Stirlingformel $k! \approx \sqrt{2\pi k}(k/\mathrm{e})^k$ verwendet haben.

b) Bei Ja-Nein-Abstimmungen in Demokratien ist der Einfluss von Wählern proportional zu $1/\sqrt{n}$ (Wurzelgesetz von Penrose). Dies ist für $n \gg 1$ größer, als $1/n$. Beispiele:

- Klassensprecherwahl: 31 Schüler und Schülerinnen. Dann ist $\sqrt{2/\pi}/\sqrt{m} \approx 5^{-1}$.
- Bürgermeister in Kleinstadt mit $n = 30.001$ Einwohnern/Hochschulwahl: Dann ist $\sqrt{2/\pi}/\sqrt{m} \approx 150^{-1}$
- Bürgermeister von Hamburg: $n = 1.300.001$ Wähler. Dann ist $\sqrt{2/\pi}/\sqrt{m} \approx 1.010^{-1} \gg 1.300.000^{-1}$.
- Bundestagswahl: $n = 60.000.001$ Wähler. Dann ist $\sqrt{2/\pi}/\sqrt{m} \approx 6.864^{-1}$

$\square$

Übung 2. *(Wir verwenden weiterhin die Notation aus der vorherigen Aufgabe). Der* totale Einfluss auf *ein Ereignis $A \subset BHW$ (oder die* Instabilität von A*) ist*

$$I(A) = \sum_{i=1}^n I_i(A).$$

Zeigen Sie:

$$I(A) \geq 4\mathbb{P}(A)(1-\mathbb{P}(A))$$

Beweis. Wir betrachten die Zufallsvariable $Y := 2\mathbb{1}_{X\in A}$. Beachte dass $2Y = Y^2$. Dann

$$\begin{aligned}\mathrm{Var}(Y) &= \left[\mathbb{E}[Y^2] - \mathbb{E}[Y]^2\right] = \left[2\mathbb{E}[Y] - \mathbb{E}[Y]^2\right] \\ &= 4\left[\mathbb{P}(A) - \mathbb{P}(A)^2\right] = 4\mathbb{P}(A)(1-\mathbb{P}(A)).\end{aligned}$$

Nach Efron-Stein gilt aber auch

$$\mathrm{Var}(Y) \leq \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\left[(Y-Y')^2\right] = 2\sum_{i=1}^{n} \mathbb{E}\left[(\mathbb{1}_{X\in A} - \mathbb{1}_{\tilde{X}^{(i)}\in A})^2\right],$$

wobei $\tilde{X}^{(i)}$ eine Kopie von X ist, bei der die i-te Koordinate unabhängig neu ausgewürfelt wurde. Nun ist

$$(\mathbb{1}_{X\in A} - \mathbb{1}_{\tilde{X}^{(i)}\in A)})^2 = \begin{cases} 1 & \text{falls} \quad \mathbb{1}_{X\in A} \neq \mathbb{1}_{\tilde{X}^{(i)}\in A}, \\ 0 & \text{sonst.}\end{cases}$$

Also ist

$$\begin{aligned}\mathbb{E}\left[(\mathbb{1}_{X\in A} - \mathbb{1}_{\tilde{X}^{(i)}\in A})^2\right] &= \mathbb{P}\left(\mathbb{1}_{X\in A} \neq \mathbb{1}_{\tilde{X}^{(i)}\in A}\right) \\ &= \frac{1}{4}\sum_{(a,b)\in\{(\pm 1,\pm 1)\}} \mathbb{P}\left(\mathbb{1}_{X\in A} \neq \mathbb{1}_{\tilde{X}^{(i)}\in A}, (X_i, \tilde{X}_i) = (a,b)\right) \\ &= \frac{1}{4}\sum_{c=\pm 1} \mathbb{P}\left(\mathbb{1}_{X\in A} \neq \mathbb{1}_{\tilde{X}^{(i)}\in A}, (X_i, \tilde{X}_i) = (c,-c)\right) \\ &= \frac{1}{2}\mathbb{P}(\mathbb{1}_{X\in A} \neq \mathbb{1}_{\overline{X}^{(i)}\in A}) \\ &= \frac{1}{2} I_i(A)\end{aligned}$$

□

und es folgt die Behauptung.

Übung 3. *Seien $\mathcal{X}$ und $\mathcal{Y}$ abzählbar und P eine Wahrscheinlichkeitsverteilung auf $\mathcal{X}\times\mathcal{Y}$ mit Randverteilungen P_X und P_Y. Seien Q_X und Q_Y weiterhin Wahrscheinlichkeitsverteilungen auf $\mathcal{X}$ bzw. $\mathcal{Y}$. Zeigen sie*

$$D(P\|Q_X\otimes Q_Y) = D(P\|P_X\otimes P_Y) + D(P_X\|Q_X) + D(P_Y\|Q_Y)$$

Beweis. Sei p die Dichte von P, p_X und p_X die Dichten von P_X und P_Y sowie q_X und q_Y die Dichten von Q_X bzw. Q_Y. Nach der Definition von $D(P\|Q_X\otimes Q_Y)$ und da $p_X(x) = 0$

oder $p_Y(y) = 0$ bereits $p(x,y) = 0$ impliziert gilt

$$
\begin{aligned}
D(P\|Q_X \otimes Q_Y) &= \sum_{(x,y)\in(\mathcal{X},\mathcal{Y})} p(x,y) \log\left(\frac{p(x,y)}{q_X(x)q_Y(y)}\right) \\
&= \sum_{(x,y)\in(\mathcal{X},\mathcal{Y})} p(x,y) \log\left(\frac{p(x,y)}{p_X(x)p_Y(y)} \frac{p_X(x)}{q_X(x)} \frac{p_Y(y)}{q_Y(y)}\right) \\
&= \sum_{(x,y)\in(\mathcal{X},\mathcal{Y})} p(x,y) \log\left(\frac{p(x,y)}{p_X(x)p_Y(y)}\right) + \sum_{(x,y)\in(\mathcal{X},\mathcal{Y})} p(x,y) \log\left(\frac{p_X(x)}{q_X(x)}\right) \\
&\quad + \sum_{(x,y)\in(\mathcal{X},\mathcal{Y})} p(x,y) \log\left(\frac{p_Y(y)}{q_Y(y)}\right) \\
&= D(P\|P_X \otimes P_Y) + \sum_{x\in\mathcal{X}} \log\left(\frac{p_X(x)}{q_X(x)}\right)\left(\sum_{y\in\mathcal{Y}} p(x,y)\right) \\
&\quad + \sum_{y\in\mathcal{Y}} \log\left(\frac{p_Y(y)}{q_Y(y)}\right)\left(\sum_{x\in\mathcal{X}} p(x,y)\right) \\
&= D(P\|P_X \otimes P_Y) + \sum_{x\in\mathcal{X}} p_X(x) \log\left(\frac{p_X(x)}{q_X(x)}\right) + \sum_{y\in\mathcal{Y}} p_Y(y) \log\left(\frac{p_Y(y)}{q_Y(y)}\right) \\
&= D(P\|P_X \otimes P_Y) + D(P_X\|Q_X) + D(P_Y\|Q_Y)
\end{aligned}
$$

□

Übung 4. *In der Vorlesung wird gezeigt werden:*

Theorem 0.1 (Subadditivität der Entropie)**.** *Sei $\Phi : [0,\infty) \to \mathbb{R}$ definiert durch $\phi(0) = 0$ und $\phi(x) = x \ln x$ sonst. Seien $X_1, \dots, X_n$ unabhängig mit Werten in einer abzählbaren Menge $\mathcal{X}$ und sein $f : \mathcal{X}^n \to [0,\infty)$. Wir definieren $Z := f(X_1, \dots, X_n)$. Dann gilt:*

$$
\mathrm{Ent}(Z) := \mathbb{E}\Phi(Z) - \Phi(\mathbb{E}Z) \leq \sum_{i=1}^{n} \left[\mathbb{E}^{(i)}\Phi(Z) - \Phi(\mathbb{E}^{(i)}Z)\right].
$$

Führen sie einen ersten Schritt des Beweis aus, indem Sie zeigen, dass es genügt, die Aussage im Fall $\mathbb{E}Z = 1$ zu zeigen.

Beweis. Gelte die Aussage für Z und sei $c > 0$. Wir müssen zeigen, dass sie dann auch für cZ gilt. Dazu rechnen wir:

$$
\begin{aligned}
\Phi(cZ) &= cZ(\ln c + \ln Z) = cZ \ln Z + \Phi(c)Z = c\Phi(Z) + \Phi(c)Z, \\
\Phi(c\mathbb{E}Z) &= c\Phi(\mathbb{E}Z) + \Phi(c)\mathbb{E}Z
\end{aligned}
$$

(Das gleiche gilt mit $\mathbb{E}^{(i)}$ statt $\mathbb{E}$). Damit:

$$\begin{aligned}
\mathbb{E}\Phi(cZ) - \Phi(\mathbb{E}(cZ)) &= c\mathbb{E}\Phi(Z) + \Phi(c)\mathbb{E}Z - c\Phi(\mathbb{E}Z) - \Phi(c)\mathbb{E}Z \\
&= c\left(\mathbb{E}\Phi(Z) - \Phi(\mathbb{E}Z)\right) \\
&\leq c\sum_{i=1}^{n}\left[\mathbb{E}^{(i)}\Phi(Z) - \Phi(\mathbb{E}^{(i)}Z)\right] \\
&= \sum_{i=1}^{n}\left[c\mathbb{E}^{(i)}\Phi(Z) + \Phi(c)\mathbb{E}^{(i)}Z - \Phi(\mathbb{E}^{(i)}Z) - \Phi(c)\mathbb{E}^{(i)}Z\right] \\
&= \sum_{i=1}^{n}\left[\mathbb{E}^{(i)}(c\Phi(Z)) - \Phi(\mathbb{E}^{(i)}(cZ))\right].
\end{aligned}$$

Nichtnegative Zufallsvariablen mit Erwartungswert 0 sind fast sicher konstant 0, sodass die Gleichung automatisch gilt. □

Abgabe bis zum 28.06.2021