# Wahrscheinlichkeit und Statistik

## Serie 8

Version 2 (29. April: Hinweise in Aufgabe 8.4 ergänzt.), Version 1 (25. April)
Bitte stellt Fragen in den Übungen und/oder im Forum.
Bitte stell sicher, dass du die Webseite `https://kahoot.it/` in der Übung am **30. April** öffnen kannst.
Freiwillige Abgabe bis **02. Mai 8:00**. Nachher kann selbstständig mit der Lösung verglichen werden.

**Aufgabe 8.1 [Korrelation & Unabhängigkeit]**
Für zwei unabhängige Zufallsvariablen $X, Y$ ist aus der Vorlesung bekannt, dass

$$\mathrm{Cov}(X, Y) = 0,$$

d.h. die Zufallsvariablen sind unkorreliert. In dieser Aufgabe zeigen wir, dass die Umkehrung dieser Aussage im Allgemeinen nicht gilt.

(a) Sei $X \sim \mathcal{U}([-\pi, \pi])$. Zeige, dass $Y := \cos(X)$ und $Z := \sin(X)$ unkorreliert sind, d.h.

$$\mathrm{Cov}(Y, Z) = 0.$$

(b) Zeige, dass $Y$ und $Z$ nicht unabhängig sind.

**Aufgabe 8.2 [Gesetz der grossen Zahlen I: Empirische Verteilungsfunktion]**
Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit Verteilungsfunktion $F$. Die empirische Verteilungsfunktion $F_n : \mathbb{R} \times \Omega \to [0, 1]$ ist gegeben durch

$$F_n(t, \omega) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i(\omega) \leq t}.$$

Der Wert $F_n(t, \omega)$ beschreibt also die relative Häufigkeit der $X_i(\omega)$ mit Werten $\leq t$ unter den ersten $n$. Damit ist $F_n(t) := F_n(t, \cdot)$ selbst eine Zufallsvariable für jedes $t \in \mathbb{R}$.

(a) Sei $t \in \mathbb{R}$ beliebig und $Y_i := \mathbb{1}_{X_i \leq t}$ für $i \in \mathbb{N}$. Zeige, dass $Y_1, Y_2, \ldots$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen ist. Was ist $\mathbb{E}[Y_1]$?

(b) Zeige, dass für jedes $t \in \mathbb{R}$ die empirische Verteilungsfunktion $F_n(t)$ fast sicher gegen $F(t)$ konvergiert für $n \to \infty$.

**Aufgabe 8.3** We define[1] the *variance* of a random variable $X$ as

$$\mathbb{V}(X) := \begin{cases} \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2, & \mathbb{E}[|X|] < \infty, \\ \infty, & \text{else.} \end{cases}$$

(a) Suppose that $X$ is a random variable. Show that $\mathbb{E}\left[X^2\right] < \infty$ if and only if $\mathbb{V}(X) < \infty$.

(b) Suppose that $\mathbb{V}(X) < \infty$. Show that $\mathbb{E}[X]$ minimizes the function

$$a \mapsto E[(X - a)^2] \quad (a \in \mathbb{R}). \tag{1}$$

[1] In (a) we will show that this definition of the variance coincides with the definition of the variance in Notes 9 for every $X \in L^2$ and is $\infty$ for every $X \notin L^2$.

(c) Appreciate the fact that the median $F_X^{-1}(0.5)$ of $X$ minimizes the function

$$a \mapsto \mathbb{E}\left[|X-a|\right] \quad (a \in \mathbb{R}). \tag{2}$$

If you want you can prove (c), but the proof is rather technical, long (and maybe boring). A proof will be provided in the solution next week.

**Aufgabe 8.4 [Markow & Chebychev Ungleichung]**
Sei $n \in \mathbb{N}$ ungerade. Wir definieren die Mehrheitsfunktion

$$\begin{aligned} \mathrm{MAJ}_n\colon \{0,1\}^n &\to \{0,1\} \\ (x_1,\dots,x_n) &\mapsto \begin{cases} 0, & \text{falls } \sum_{i=1}^n x_i < \frac{n}{2}, \\ 1, & \text{falls } \sum_{i=1}^n x_i > \frac{n}{2}. \end{cases} \end{aligned}$$

Sei $(X_i)_{i\geq 1}$ eine Folge von unabhängigen, $\mathrm{Ber}(p)$-verteilten Zufallsvariablen mit Parameter $p \in [0,1]$. Definiere

$$Y_n := \mathrm{MAJ}_n(X_1,\dots,X_n).$$

(a) Berechne $\mathrm{P}(Y_n = 1)$, $\mathbb{E}[Y_n]$ und $\mathbb{V}(Y_n)$ für $p = 1/2$.

(b) Zeige, dass für alle $p \in [0,1]$,

$$\mathbb{E}[Y_n] \leq 2p \quad \text{und} \quad \mathbb{V}(Y_n) \leq \frac{1}{4}.$$

*Hinweis: Nutze die Markow Ungleichung aus Notizen 8 (Seite 27), um* $\mathrm{P}\left(\sum_{i=1}^n X_i > \frac{n}{2}\right)$ *abzuschätzen.*

(c) Zeige, dass

$$\mathbb{E}[Y_n] = \mathrm{P}(Y_n = 1) \xrightarrow{n\to\infty} \begin{cases} 0, & \text{falls } p < 1/2, \\ 1, & \text{falls } p > 1/2. \end{cases}$$

*Hinweis: Nutze die Chebychev[2] Ungleichung aus Notizen 9 (Seite 9).*

(d) Schlussfolgere aus (c), dass $\mathbb{V}(Y_n) \xrightarrow{n\to\infty} 0$ für $p \neq 1/2$.

**Aufgabe 8.5** Sei $X_1, X_2, \dots$ eine unendliche Folge von unabhängigen, identisch verteilten Zufallsvariablen, wobei $X_1$ diskret ist mit

$$\mathbb{P}[X_1 = -1] = \mathbb{P}[X_1 = 0] = \frac{1}{4} \quad \text{und} \quad \mathbb{P}[X_1 = 1] = \frac{1}{2}.$$

Für jedes $n \in \mathbb{N} \setminus \{0\}$ betrachten wir die Partialsumme

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

(a) Bestimmen Sie den Erwartungswert und die Varianz von $S_n$.

(b) Zeigen Sie, dass

$$\mathbb{P}\big[S_n \geq n/2\big] \leq \mathbb{P}\big[|S_n - n/4| \geq n/4\big].$$

Schlussfolgern Sie, dass

$$\mathbb{P}[S_n \geq n/2] \to 0 \quad \text{für } n \to \infty.$$

(c) Bestimmen Sie die Wahrscheinlichkeit

$$\mathbb{P}\left[\lim_{n\to\infty} S_n/n = 1/4\right].$$

Wenn du Feedback zur Serie hast, schreibe bitte in das Forum (oder eine Mail an Jakob Heiss).

[2] Für den Namen „Chebychev“ sind mehrere unterschiedliche Schreibweisen gebräuchlich: z.B. „Tschebyschow“, „Tschebyscheff“ oder „Tschebyschev“. Daher ist auch die Chebychev Ungleichung unter mehreren Schreibweisen bekannt.