Prof. Dr. A. Steland
N. Mause

RWTH Aachen, SS2023

Ausgabe: 06. Juni 2023 Besprechung: 12. Juni 2023

# Einführung in die angewandte Stochastik

## Übungsblatt 7

*Für die Bearbeitung der Aufgaben auf diesem Übungsblatt benötigen Sie die Tabelle mit Funktionswerten der Standardnormalverteilung, die Sie am Ende des Übungsblattes finden können.*

### Aufgabe 27

Seien $X_1, \ldots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert $\mu \in \mathbb{R}$ und Varianz $\sigma^2 > 0$. Außerdem seien $Y_1, \ldots, Y_n$ weitere unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert $\nu \in \mathbb{R}$ und Varianz $\tau^2 > 0$, die zusätzlich von $X_1, X_2, \ldots$ stochastisch unabhängig seien. Alle vorkommenden Zufallsvariablen seien auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ definiert.

Wir möchten in dieser Aufgabe für beliebiges $\epsilon > 0$ den Grenzwert

$$(*) \qquad \lim_{n\to\infty} P\left(\sum_{i=1}^n X_i \leq n(\mu - \nu) + n\epsilon + \sum_{i=1}^n Y_i\right)$$

berechnen. Gehen Sie dazu wie folgt vor:

(a) Definieren Sie Zufallsvariablen $Z_1, \ldots, Z_n$ durch $Z_i = X_i - Y_i$ für $i \in \{1, \ldots, n\}$. Bestimmen Sie $E(Z_i)$ und $\mathrm{Var}(Z_i)$.

(b) Sei nun $S_n = \sum_{i=1}^n Z_i$. Berechnen Sie $E(S_n)$.

(c) Schätzen Sie $(*)$ geeignet nach unten ab und verwenden Sie ein aus der Vorlesung bekanntes Gesetz.

**Lösung:**

a) $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ stoch. unabh.

$\Rightarrow Z_i = X_i - Y_i$ stoch. unabh. $\forall i = 1, \ldots, n$

$\Rightarrow E(Z_i) = E(X_i) - E(Y_i) = \mu - \nu$

$\mathrm{Var}(Z_i) = \mathrm{Var}(X_i - Y_i) \underset{\substack{X_i, Y_i \\ \text{stoch unabh.}}}{=} \mathrm{Var}(X_i) + (-1)^2 \mathrm{Var}(Y_i)$

$= \sigma^2 + \tau^2$

b) $S_n = \sum_{i=1}^n Z_i = \sum_{i=1}^n (X_i - Y_i)$

$$\Rightarrow E(S_n) = \sum_{i=1}^{n} (E(X_i) - E(Y_i)) = \sum_{i=1}^{n} (\mu - \nu) = n(\mu - \nu)$$

c) $\overline{Z}_n = \frac{1}{n} \sum_{i=1}^{n} Z_i = \frac{1}{n} S_n$ mit $E(\overline{Z}_n) = \frac{1}{n} E(S_n) \overset{b)}{=} \mu - \nu$

Für $\varepsilon > 0$ gilt

$$\lim_{n\to\infty} P\left(\sum_{i=1}^{n} X_i \leq n(\mu-\nu) + n\varepsilon + \sum_{i=1}^{n} Y_i\right)$$

$$= \lim_{n\to\infty} P\Big(\underbrace{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i}_{= \sum_{i=1}^{n} Z_i = S_n} \leq \underbrace{n(\mu-\nu)}_{\overset{b)}{=} E(S_n)} + n\varepsilon\Big)$$

$$= \lim_{n\to\infty} P(S_n - E(S_n) \leq n\varepsilon)$$

$$= \lim_{n\to\infty} P\Big(\underbrace{\tfrac{1}{n} S_n}_{= \overline{Z}_n} - \tfrac{1}{n} E(S_n) \leq \varepsilon\Big)$$

$$= \lim_{n\to\infty} P(\overline{Z}_n - (\mu-\nu) \leq \varepsilon)$$

$$= 1 - \lim_{n\to\infty} P(\overline{Z}_n - (\mu-\nu) > \varepsilon)$$

Hier ist $\{\overline{Z}_n - (\mu-\nu) > \varepsilon\} \subset \{|\overline{Z}_n - (\mu-\nu)| > \varepsilon\}$

$\Rightarrow P(\overline{Z}_n - (\mu-\nu) > \varepsilon) \leq P(|\overline{Z}_n - (\mu-\nu)| > \varepsilon)$,
da $P$ monoton ist als W'maß

$\Rightarrow 1 - P(\overline{Z}_n - (\mu-\nu) > \varepsilon) \geq 1 - P(|\overline{Z}_n - (\mu-\nu)| > \varepsilon)$

$$\Rightarrow (*) \geq 1 - \underbrace{\lim_{n\to\infty} P(|\overline{Z}_n - (\mu-\nu)| > \varepsilon)}_{= 0 \text{ nach dem schwachen GGZ}} = 1$$

$P$ ist ein W'maß $\Rightarrow P(A) \leq 1 \ \forall A \subset \Omega$

$$\Rightarrow \lim_{n\to\infty} P\left(\sum_{i=1}^{n} X_i \leq n(\mu-\nu) + n\varepsilon + \sum_{i=1}^{n} Y_i\right) = 1$$

nach dem Sandwich-Theorem

## Aufgabe 28

Im Allgemeinen kann eine (unbekannte) Verteilungsfunktion $F$ durch die empirische Verteilungsfunktion

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}} = \begin{cases} 1, & X_i \leq x \\ 0, & X_i > x \end{cases}, \quad x \in \mathbb{R},$$

auf Basis von $n \in \mathbb{N}$ Zufallsvariablen $X_1, \ldots, X_n$ geschätzt werden.

(a) Im Rahmen einer statistischen Untersuchung wurden die folgenden Beobachtungswerte ermittelt:

$$3.8\,,\ 2.3\,,\ 1.3\,,\ 1.7\,,\ 1.6\,,\ 2.1\,,\ -0.2\,,\ 0.8\,.$$

Diese Daten können aufgefasst werden als Realisationen $x_1, \ldots, x_8$ stochastisch unabhängiger, identisch verteilter Zufallsvariablen $X_1, \ldots, X_8$ mit unbekannter Verteilungsfunktion $F$. Berechnen Sie aus den gegebenen Beobachtungen einen Schätzwert für $F(2)$.

(b) Seien $X_1, \ldots, X_n$ stochastisch unabhängige und identisch verteilte Zufallsvariablen mit Verteilungsfunktion $F$. Bestimmen Sie $E(F_n(x))$ und $\operatorname{Var}(F_n(x))$. Was passiert für $n \to \infty$?

**Lösung:** a) $n = 8$

$$F_8(2) \overset{\text{Def}}{=} \frac{1}{8} \underbrace{\sum_{i=1}^{n} \mathbb{1}_{(-\infty, 2]}(x_i)}_{\substack{\text{Relativer Anteil} \\ \text{der Beobachtungen} \leq 2}} = \frac{5}{8} = 0{,}625$$

b) Sei $Y_i = \mathbb{1}_{\{X_i \leq x\}} = \mathbb{1}_{(-\infty, x]}(X_i) = \begin{cases} 1, & X_i \leq x \\ 0, & X_i > x \end{cases}$

für gegebenes $x \in \mathbb{R}$.

Nach Vor. sind $X_1, \ldots, X_n$ stoch. unabh.

$\Rightarrow Y_1, \ldots, Y_n$ stoch. unabh.

$$P(Y_i = 1) = P(X_i \leq x) \overset{\text{Def}}{=} F(x) =: p$$

$$P(Y_i = 0) = P(X_i > x) = 1 - P(X_i \leq x) = 1 - F(x) = 1 - p$$

für alle $i = 1, \ldots, n$

$\Rightarrow Y_i \sim \operatorname{Ber}(p) = \operatorname{Bin}(1, p)$

$\Rightarrow E(Y_i) = 0 \cdot P(Y_i = 0) + 1 \cdot P(Y_i = 1) = P(Y_i = 1) = p$

$$E(y_i^2) = 0^2 \cdot P(y_i=0) + 1^2 \cdot P(y_i=1) = P(y_i=1) = p$$

$$\Rightarrow Var(y_i) \underset{\text{Verschiebungssatz}}{=} E(y_i^2) - (E(y_i))^2 = p - p^2 = p(1-p)$$

$$= F(x)(1-F(x))$$

$$\Rightarrow E(F_n(x)) = E\left(\frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le x\}}\right) = \frac{1}{n}\sum_{i=1}^{n} E(y_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n} p = F(x) \quad \left[= \frac{1}{n}\sum_{i=1}^{n} F(x) = \frac{1}{n}\cdot n\, F(x)\right]$$

und

$$Var(F_n(x)) = Var\left(\frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le x\}}\right)$$

$$= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} y_i\right) \overset{y_i\text{ stoch.}}{\underset{\text{unabh.}}{=}} \frac{1}{n^2}\sum_{i=1}^{n} Var(y_i)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} F(x)(1-F(x))$$

$$= \frac{F(x)(1-F(x))}{n} \xrightarrow[n\to\infty]{} 0$$

### Aufgabe 29

Es seien $X_1, \dots, X_n$ unabhängige, standard-exponentialverteilte Zufallsvariablen $X_i \sim \mathrm{Exp}(1), i = 1, \dots, n$. Definiere nun $Z_n = \min_{i=1,\dots,n} X_i$. Zeigen Sie, dass $Z_n \xrightarrow{P} 0$ für $n \to \infty$. D.h. zeigen Sie, dass für jedes $\epsilon > 0$ gilt

$$P(|Z_n| > \epsilon) \to 0, \qquad n \to \infty.$$

**Lösung:**

$X_i \sim Exp(1) \Rightarrow f_X(x) = e^{-x} 1\!\!1_{[0,\infty)}$

$\Rightarrow P(X_i \leq 0) = \int_{-\infty}^{0} f_X(x)\,dx = \int_{-\infty}^{0} 0\,dx = 0$

$\Rightarrow P(X_i > 0) = 1 - P(X_i \leq 0) = 1 - 0 = 1 \quad \forall i = 1,..,n$

$\Rightarrow Z_n \overset{Def}{=} \min_{i=1,..,n} X_i > 0 \Rightarrow |Z_n| = Z_n$

Es reicht zu zeigen: $P(Z_n > \varepsilon) \xrightarrow[n\to\infty]{} 0$.

Hier ist $P(Z_n > \varepsilon) \overset{Def\, Z_n}{=} P(\min_{i=1,..,n} X_i > \varepsilon)$

$= P(X_i > \varepsilon \ \forall i = 1,..,n)$

$\overset{X_i \text{ stoch. unabh.}}{=} \prod_{i=1}^{n} P(X_i > \varepsilon)$

$\overset{X_i \text{ identisch verteilt}}{=} \prod_{i=1}^{n} P(X_1 > \varepsilon)$

$= P(X_1 > \varepsilon)^n$

Hier ist

$P(X_1 > \varepsilon) = \int_{\varepsilon}^{\infty} f_{X_1}(x)\,dx = \int_{\varepsilon}^{\infty} e^{-x}\,dx = \left[-e^{-x}\right]_{\varepsilon}^{\infty}$

$= e^{-\varepsilon} \in (0,1)$ für $\varepsilon > 0$

$\Rightarrow P(Z_n > \varepsilon) = e^{-n\varepsilon} \xrightarrow[n\to\infty]{} 0 \quad \forall \varepsilon > 0$

**Aufgabe 30**

Eine Bank betreibt in einer Region insgesamt 200 Geldautomaten, von denen jeder (unabhängig von den übrigen Automaten) mit Wahrscheinlichkeit 0.05 aufgrund einer Störung innerhalb einer Woche mindestens einmal ausfällt. Für die Einrichtung eines ständigen Wartungsdienstes ist die Wahrscheinlichkeit von Interesse, dass die Anzahl der Geldautomaten, die in einer Woche mindestens eine derartige Störung aufweisen, mindestens 5 und höchstens 15 beträgt.

(a) Schätzen Sie die gesuchte Wahrscheinlichkeit mittels der Tschebyscheff–Ungleichung nach unten ab.

(b) Berechnen Sie nun die gesuchte Wahrscheinlichkeit approximativ mit Hilfe des Zentralen Grenzwertsatzes.

**Lösung:** Vorbereitung

Definiere die Zufallsvariablen $X_1, \dots, X_{200}$ durch

$$X_i = \begin{cases} 1, & i\text{-ter Automat fällt mindestens einmal aus} \\ 0, & \text{sonst} \end{cases}$$

Sei $n = 200$ und

$$P(X_i = 1) = 0{,}05 := p$$
$$P(X_i = 0) = 1 - 0{,}05 = 0{,}95 = 1 - p$$

Automaten sind unabh.

$\Rightarrow X_i$ stoch. unabh. $\forall i = 1, \dots, n$

Außerdem sind $X_i \sim Ber(p) = Bin(1, p)$

$$\Rightarrow S_n = S_{200} := \sum_{i=1}^{n} X_i \sim Bin(n, p) = Bin(200, p)$$

da die $X_i$ stoch. unabh. sind.

Weiterhin gilt

$$E(X_i) = p = \frac{1}{20} = 0{,}05$$
$$Var(X_i) = p(1-p) = \frac{1}{20} \cdot \frac{19}{20} = \frac{19}{400}$$
$$E(S_n) = n \cdot p = 200 \cdot \frac{1}{20} = 10$$
$$Var(S_n) = n \cdot p(1-p) = 200 \cdot \frac{1}{20} \cdot \frac{19}{20} = 9{,}5$$

Sei $\overline{X_n} = \frac{1}{n} S_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

$$\Rightarrow E(\overline{X_n}) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{20} = \frac{1}{n} \cdot n \cdot \frac{1}{20} = \frac{1}{20}$$

a) $P(5 \leq S_n \leq 15) = P(-5 \leq S_n - 10 \leq 5)$

$E(S_n) = 10$

$n = 200$

$$= P(|S_n - 10| \leq 5)$$

$$= 1 - P(|S_n - \underbrace{10}_{=E(S_n)}| > 5)$$

$$= 1 - P\left(\left|\frac{1}{n} S_n - \underbrace{\frac{1}{n} E(S_n)}_{=\frac{1}{20}}\right| > \frac{5}{n}\right)$$

$$= 1 - P\left(\left|\overline{X_n} - \frac{1}{20}\right| > \frac{1}{40}\right)$$

Nach der Tschebyscheff-Ungleichung

$$P\left(\left|\overline{X_n} - \underbrace{E(X_1)}_{=\frac{1}{20}}\right| > \frac{1}{40}\right) \leq \frac{Var(X_1)}{n \cdot \frac{1}{40^2}} = \frac{\frac{19}{400}}{200 \cdot \frac{1}{40^2}}$$

$$= \frac{19}{50}$$

$$\Rightarrow P(5 \leq S_n \leq 15) \geq 1 - \frac{19}{50} = \frac{31}{50} = 0{,}62$$

b) Es gilt

$$P(5 \leq S_n \leq 15) = P(S_n \leq 15) - P(S_n \leq 4)$$

$$= P\left(\frac{1}{200} \sum_{i=1}^{200} X_i \leq \frac{15}{200}\right) - P\left(\frac{1}{200} \sum_{i=1}^{200} X_i \leq \frac{4}{200}\right)$$

$$= P\left(\overline{X_n} - \frac{1}{20} \leq \frac{15}{200} - \frac{1}{20}\right) - P\left(\overline{X_n} - \frac{1}{20} \leq \frac{4}{200} - \frac{1}{20}\right)$$

$$= P\left(\sqrt{200} \cdot \frac{\overline{X_n} - \frac{1}{20}}{\sqrt{\frac{19}{400}}} \leq \overbrace{\sqrt{200} \frac{\frac{15}{200} - \frac{1}{20}}{\sqrt{\frac{19}{400}}}}^{=: t_1}\right)$$

$$-P\Big(\underbrace{\sqrt{200}\,\frac{\overline{X}_n-\frac{1}{20}}{\sqrt{\frac{19}{400}}}\le\underbrace{\sqrt{200}\,\frac{\frac{4}{200}-\frac{1}{20}}{\sqrt{\frac{19}{400}}}}_{=:t_2}\Big)}_{\approx\,\Phi(t_2)}$$

ZGWS

$$\approx \Phi(t_1)-\Phi(t_2)$$

$$=\Phi(1{,}62)-\Phi(-1{,}95)$$

$$=\Phi(1{,}62)-(1-\Phi(1{,}95))$$

$$=\Phi(1{,}62)-1+\Phi(1{,}95)$$

$$\overset{\text{Tabelle}}{=} 0{,}9474-1+0{,}9744=0{,}9218$$

$$\boxed{\begin{array}{l}\text{Es gilt}\\ \Phi(-x)=1-\Phi(x)\end{array}}$$