

## Klausur „Stochastik“

Bitte diese Felder in Druckschrift ausfüllen

Name:		Vorname:	
Matrikelnr.:		Studiengang:	

### Wichtige Hinweise:

- **Es sind keine eigenen Unterlagen, Handys, Taschenrechner u.ä. zugelassen!**
- Dieses Deckblatt ist vollständig ausgefüllt zusammen mit den Lösungen abzugeben. Jedes abgegebene Blatt ist zudem mit Namen und Matrikelnummer zu versehen.
- Bitte legen Sie den Studentenausweis und einen amtlichen Lichtbildausweis gut sichtbar neben Ihren Platz!
- Abgabe bis spätestens 11.00 Uhr.

**Viel Erfolg!**

---

Diese Felder NICHT ausfüllen:

Aufgabe	1	2	3	4			Summe	Note
Punkte								

## Werte zur Standardnormalverteilung

$\Phi(0, 0) = 0,500$	$\Phi(1, 0) = 0,841$	$\Phi(2, 0) = 0,977$
$\Phi(0, 1) = 0,539$	$\Phi(1, 1) = 0,864$	$\Phi(2, 1) = 0,982$
$\Phi(0, 2) = 0,579$	$\Phi(1, 2) = 0,884$	$\Phi(2, 2) = 0,986$
$\Phi(0, 3) = 0,617$	$\Phi(1, 3) = 0,903$	$\Phi(2, 3) = 0,989$
$\Phi(0, 4) = 0,655$	$\Phi(1, 4) = 0,919$	$\Phi(2, 4) = 0,991$
$\Phi(0, 5) = 0,691$	$\Phi(1, 5) = 0,933$	$\Phi(2, 5) = 0,993$
$\Phi(0, 6) = 0,725$	$\Phi(1, 6) = 0,945$	$\Phi(2, 6) = 0,995$
$\Phi(0, 7) = 0,758$	$\Phi(1, 7) = 0,955$	$\Phi(2, 7) = 0,996$
$\Phi(0, 8) = 0,788$	$\Phi(1, 8) = 0,964$	$\Phi(2, 8) = 0,997$
$\Phi(0, 9) = 0,815$	$\Phi(1, 9) = 0,971$	$\Phi(2, 9) = 0,998$

**1. (Zufallsvariablen, Unabhängigkeit, zentraler Grenzwertsatz)**

[30 Punkte]

a) Geben Sie eine kurze, aber vollständige Definition der folgenden Begriffe:

- (i) Wahrscheinlichkeitsraum,
- (ii) reellwertige Zufallsvariable  $X$ ,
- (iii) Unabhängigkeit von Zufallsvariablen  $X$  und  $Y$  mit Werten in einer abzählbaren Menge  $S$ .

b) Sei  $U$  eine auf  $\Omega = \{1, 2, 3, 4, 5, 6\}$  gleichverteilte Zufallsvariable, und seien

$$X := \begin{cases} 1 & \text{falls } U \text{ ungerade,} \\ 0 & \text{sonst,} \end{cases} \quad \text{und} \quad Y := U + X.$$

- (i) Bestimmen Sie die Verteilungen von  $X$  und  $Y$ .
  - (ii) Bestimmen Sie die gemeinsame Verteilung von  $X$  und  $Y$ .
  - (iii) Sind  $X$  und  $Y$  unabhängig?
  - (iv) Berechnen Sie die Varianz von  $X + Y$ .
- c) Sei  $T$  eine reellwertige Zufallsvariable auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ . Zeigen Sie

$$P[T \geq 1] \leq e^{-1} E[e^T].$$

d) Formulieren Sie den zentralen Grenzwertsatz (mit Voraussetzungen, aber ohne Beweis).

**Lösung:**

- a) (i) Ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$  besteht aus einer nicht-leeren Menge  $\Omega$ , einer  $\sigma$ -Algebra  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ , und einer Wahrscheinlichkeitsverteilung  $\mathbb{P}$  auf  $(\Omega, \mathcal{A})$ . Dabei hat eine  $\sigma$ -Algebra  $\mathcal{A}$  die folgenden Eigenschaften:

1.  $\Omega \in \mathcal{A}$
2. Für alle  $A \in \mathcal{A}$  ist  $A^c \in \mathcal{A}$ .
3. Für alle  $A_1, A_2, \dots \in \mathcal{A}$  ist  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ .

Eine Wahrscheinlichkeitsverteilung auf  $(\Omega, \mathcal{A})$  ist eine Abbildung  $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$  mit den folgenden Eigenschaften:

1.  $\mathbb{P}[\Omega] = 1$
2. Für alle  $A_1, A_2, \dots$  mit  $A_i \cap A_j = \emptyset$  für alle  $i \neq j$  gilt

$$\mathbb{P} \left[ \bigcup_{n \in \mathbb{N}} A_n \right] = \sum_{n=1}^{\infty} \mathbb{P}[A_n]$$

- (ii) Sei  $(\Omega, \mathcal{A}, \mathbb{P})$  ein Wahrscheinlichkeitsraum. Eine Zufallsvariable mit Werten in  $\mathbb{R}$  ist eine Funktion  $X : \Omega \rightarrow \mathbb{R}$  mit der Eigenschaft, dass

$$X^{-1}((-\infty, c]) = \{X \leq c\} = \{\omega \in \Omega : X(\omega) \leq c\} \in \mathcal{A}$$

für alle  $c \in \mathbb{R}$  gilt. Hier bezeichnet  $\mathcal{B}(\mathbb{R})$  die Borelsche  $\sigma$ -Algebra auf  $\mathbb{R}$ . Dies ist die kleinste  $\sigma$ -Algebra, die alle halboffenen Intervalle  $(-\infty, c]$  für  $c \in \mathbb{R}$  enthält.

- (iii) Die Zufallsvariablen  $X, Y$  heißen unabhängig, wenn für alle  $s, t \in S$  gilt, dass

$$\mathbb{P}[X = s, Y = t] = \mathbb{P}[X = s]\mathbb{P}[Y = t]$$

- b) (i) Es gilt

$$\mathbb{P}[X = 1] = \mathbb{P}[U \text{ ungerade}] = \mathbb{P}[U \in \{1, 3, 5\}] = \frac{|\{1, 3, 5\}|}{|\{1, 2, 3, 4, 5, 6\}|} = \frac{3}{6} = \frac{1}{2}$$

Analog ist  $\mathbb{P}[X = 0] = \frac{1}{2}$ . Für  $y \in \{1, \dots, 6\}$  gilt

$$\begin{aligned} \mathbb{P}[Y = y] &= \mathbb{P}[U + X = y] \\ &= \underbrace{\mathbb{P}[U = y, U \text{ gerade}]}_{=\frac{1}{6} \text{ für } y=2,4,6} + \underbrace{\mathbb{P}[U + 1 = y, U \text{ ungerade}]}_{=\frac{1}{6} \text{ für } y=2,4,6} \\ &= \begin{cases} \frac{1}{3} & \text{falls } y = 2, 4, 6 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

- (ii) Sei  $x \in \{0, 1\}$  dann ist

$$\mathbb{P}[X = x, Y = y] = \begin{cases} \frac{1}{6} & \text{falls } y = 2, 4, 6 \\ 0 & \text{sonst} \end{cases}$$

(iii) Für  $x \in \{0, 1\}$  und  $y \in \{2, 4, 6\}$  ist

$$\mathbb{P}[X = x, Y = y] = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = \mathbb{P}[X = x]\mathbb{P}[Y = y]$$

Für  $x \in \{0, 1\}$  und  $y \in \{1, 3, 5\}$  gilt hingegen

$$\mathbb{P}[X = x, Y = y] = 0 = \frac{1}{2} \cdot 0 = \mathbb{P}[X = x]\mathbb{P}[Y = y]$$

Somit sind  $X$  und  $Y$  unabhängig.

(iv) Da  $X$  und  $Y$  unabhängig sind folgt, dass sie ebenso unkorreliert sind. Dies nutzen wir zum Berechnen der Varianz. Es gilt

$$\begin{aligned} \text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ &= \frac{1}{2} - \frac{1}{4} + \frac{1}{3}(4 + 16 + 36) - \left(\frac{2 + 4 + 6}{3}\right)^2 \\ &= \frac{1}{4} + \frac{8}{3} = \frac{35}{12} \end{aligned}$$

c) Durch Betrachten der Fälle  $T \geq 1$  und  $T < 1$  sehen wir, dass

$$\mathbb{1}_{\{T \geq 1\}} e^1 \leq e^T.$$

Auf Grund der Monotonie und Linearität des Erwartungswertes folgt, dass

$$\mathbb{E}[\mathbb{1}_{\{T \geq 1\}}] e^1 \leq \mathbb{E}[e^T], \quad \text{also} \quad \mathbb{P}[T \geq 1] \leq e^{-1} \mathbb{E}[e^T].$$

d) (**Zentraler Grenzwertsatz**) Seien  $X_1, X_2, \dots$  unabhängige, identisch verteilte, reellwertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$  mit Erwartungswert  $m \in \mathbb{R}$  und endlicher Varianz  $\sigma^2 > 0$ , und sei

$$S_n = X_1 + \dots + X_n$$

Dann konvergieren die Verteilungen der standardisierten Summen  $\tilde{S}_n := \frac{S_n - np}{\sqrt{n}}$  für  $n \rightarrow \infty$  schwach gegen  $N(0, \sigma^2)$ , d.h. für alle  $a, b \in \mathbb{R}$  mit  $a \leq b$  gilt

$$\mathbb{P}\left[a \leq \frac{S_n - np}{\sqrt{n}} \leq b\right] \xrightarrow{n \rightarrow \infty} \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx.$$

## 2. (Verteilungsfunktionen und Dichten)

[30 Punkte]

- a) Sei  $X : \Omega \rightarrow \mathbb{R}$  eine reellwertige Zufallsvariable auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ .
- (i) Wie ist die Verteilungsfunktion von  $X$  definiert?
  - (ii) Welche Eigenschaften hat die Verteilungsfunktion einer reellwertigen Zufallsvariable?
- b) Sind die Verteilungen der folgenden Zufallsvariablen diskret, absolutstetig, oder weder diskret noch absolutstetig? (ohne Beweis). Skizzieren Sie die Graphen der Verteilungsfunktionen, sowie, im absolutstetigen Fall, die Graphen der Dichten (mit Beschriftung der Koordinatenachsen):
- (i)  $X \sim N(m, v)$  mit  $m = -3$  und  $v = 9$ ,
  - (ii)  $T \sim \text{Exp}(\lambda)$  mit  $\lambda = 1/5$ ,
  - (iii)  $Y = \max(U, 0)$ , wobei  $U$  gleichverteilt auf dem Intervall  $(-1, 1)$ ,
  - (iv)  $V = U^2$ .
- c) Sei  $c \in (0, \infty)$ . Die Verteilung der Zufallsvariable  $Z$  sei absolutstetig mit Dichtefunktion

$$f(x) = \begin{cases} c \cdot x^2 & \text{für } x \in [0, 1], \\ c & \text{für } x \in (1, 2], \\ 0 & \text{sonst.} \end{cases}$$

- (i) Zeigen Sie, dass die Konstante  $c$  den Wert  $3/4$  hat.
- (ii) Berechnen Sie die Wahrscheinlichkeit  $P[Z > 1]$ .
- (iii) Berechnen Sie den Median von  $Z$ .

### Lösung:

- a) (i) Die Verteilungsfunktion  $F_X : \mathbb{R} \rightarrow [0, 1]$  einer reellwertigen Zufallsvariable  $X : \Omega \rightarrow \mathbb{R}$  bzw. der Wahrscheinlichkeitsverteilung  $\mu_X$  auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  ist definiert als

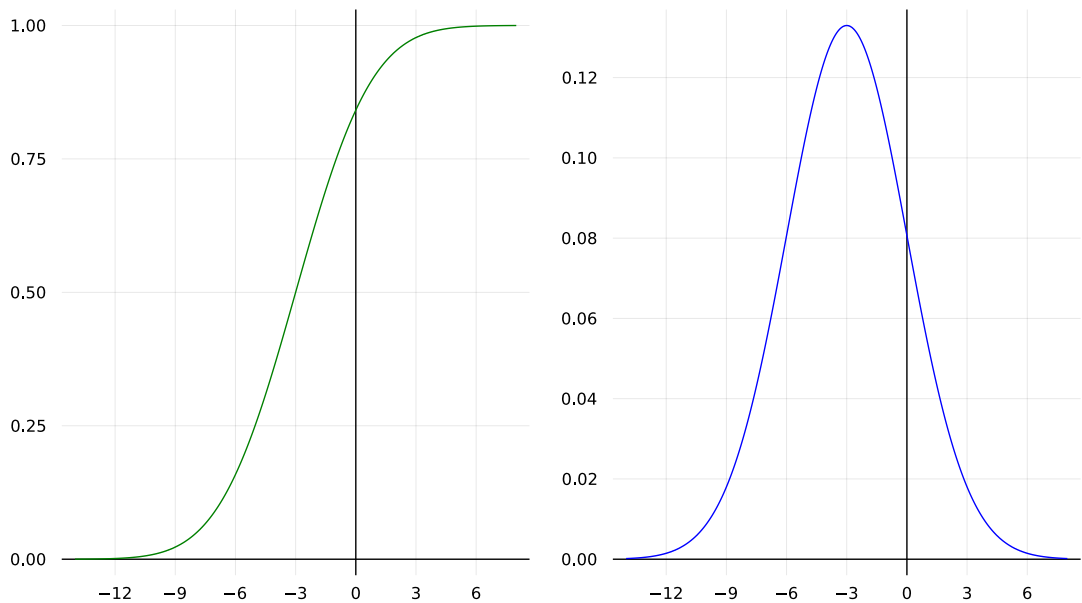
$$F_X(c) := \mu_X[(-\infty, c]] = \mathbb{P}[X \leq c]$$

für alle  $c \in \mathbb{R}$ .

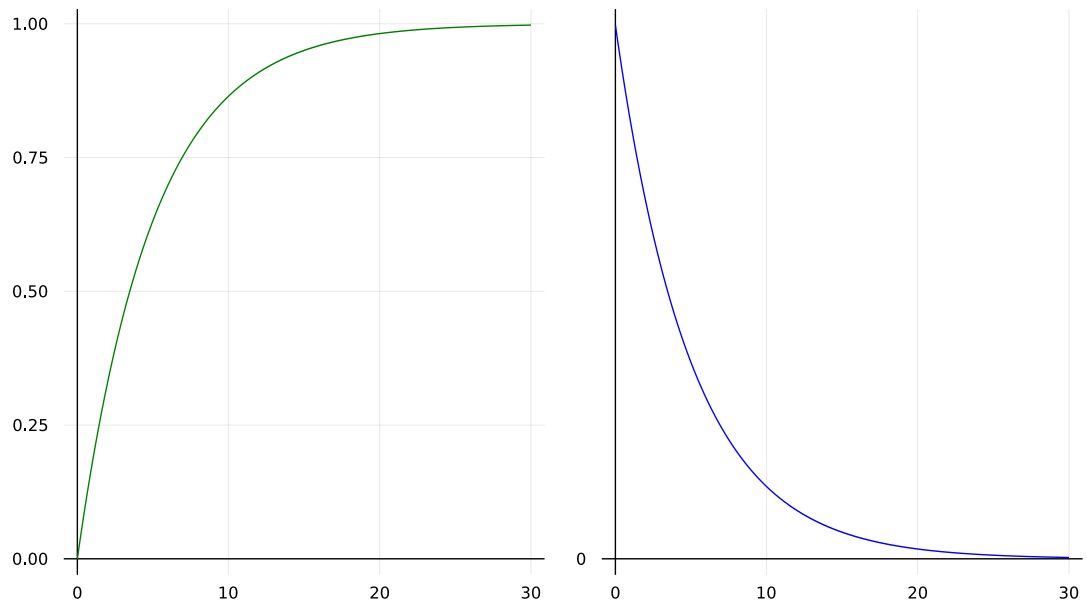
- (ii) Für die Verteilungsfunktion  $F_X : \mathbb{R} \rightarrow [0, 1]$  einer reellwertigen Zufallsvariable  $X$  gilt

1.  $F_X$  ist *monoton wachsend*
2.  $\lim_{c \rightarrow -\infty} F_X(c) = 0$  und  $\lim_{c \rightarrow \infty} F_X(c) = 1$
3.  $F_X$  ist *rechtsstetig*, d.h.  $F_X(c) = \lim_{y \searrow c} F_X(y)$  für alle  $c \in \mathbb{R}$
4.  $F_X(c) = \lim_{y \searrow c} F_X(y) + \mu_X[\{c\}]$ . Insbesondere ist  $F_X$  genau dann stetig bei  $c$ , wenn  $\mu_X[\{c\}] = 0$  gilt.

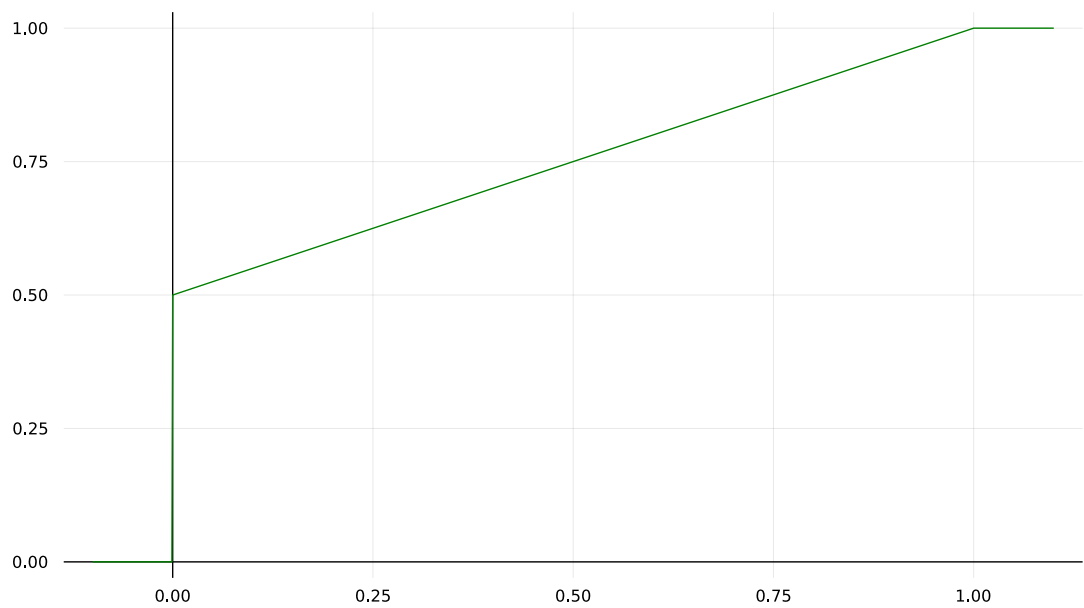
- b) (i) Die Verteilung von  $X \sim N(-3, 9)$  ist absolutstetig.



(ii) Die Verteilung von  $T \sim \text{Exp}(1/5)$  ist absolutstetig.



(ii) Die Verteilung von  $Y = \max(U, 0)$  mit  $U \sim \text{Unif}(-1, 1)$  ist weder absolutstetig noch diskret.



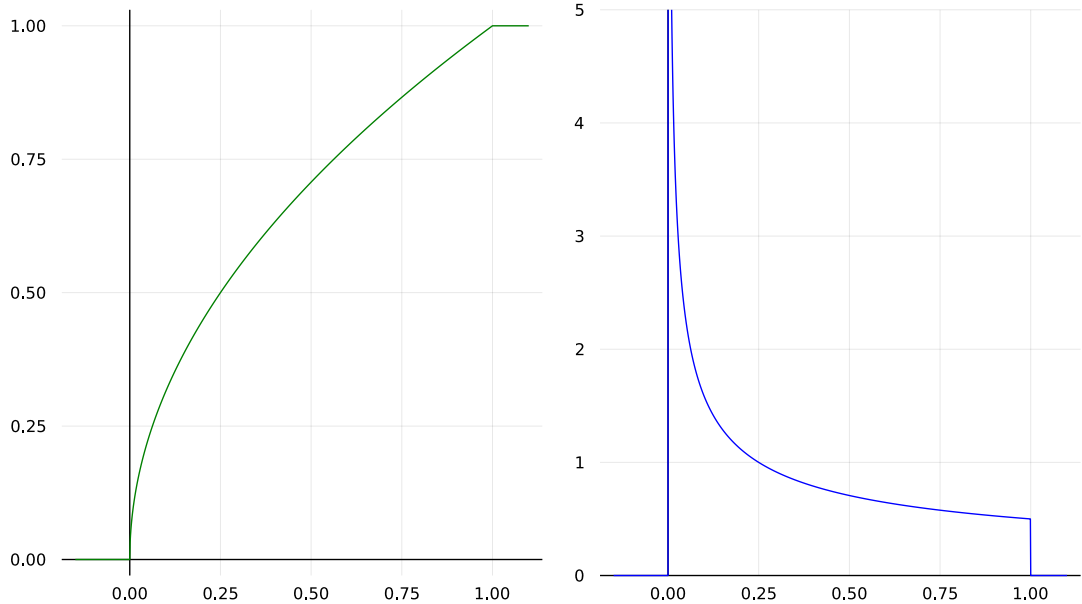


(iv) Die Verteilung von  $V = U^2$  ist absolutstetig mit Verteilungsfunktion

$$F_V(c) = \mathbb{P}[U^2 \leq c] = \mathbb{P}[U \leq \sqrt{c}] = \begin{cases} 0 & \text{falls } c < 0, \\ \sqrt{c} & \text{falls } 0 \leq c \leq 1, \\ 1 & \text{sonst,} \end{cases}$$

und Dichte

$$f_V(y) = F'_V(y) = \frac{1}{2\sqrt{y}} \mathbf{1}_{(0,1)}(y).$$



c) (i) Es gilt

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^1 cx^2 dx + \int_1^2 c dx = \frac{c}{3} + c = \frac{4}{3}c$$

Da  $f$  die Dichte einer Zufallsvariable ist muss diese zu eins integrieren. Also gilt  $\frac{4}{3}c = 1$ , das heißt  $c = \frac{3}{4}$ .

(ii) Wir berechnen

$$\mathbb{P}[Z > 1] = \int_1^{\infty} f(x) dx = \int_1^2 \frac{3}{4} dx = \frac{3}{4}$$

(iii) Aus (ii) folgt, dass  $\mathbb{P}[Z \leq 1] = \frac{1}{4} < \frac{1}{2}$ . Daher liegt der Median  $m$  von  $Z$  zwischen 1 und 2. Für  $m \in [1, 2]$  gilt

$$\mathbb{P}[Z \geq m] = \int_m^{\infty} f(x) dx = \int_m^2 \frac{3}{4} dx = \frac{3}{4}(2 - m).$$

Der Median ist das eindeutige  $m \in (1, 2)$ , für das diese Wahrscheinlichkeit gleich  $1/2$  ist. Also ist  $m = 4/3$ .

### 3. (Stochastische Modelle)

[25 Punkte]

- a) Sie bringen Ihr Fahrrad zur Reparatur in eine Fahrradwerkstatt. Die Werkstatt hat zehn Mitarbeiter, von denen acht fest angestellt, und zwei Aushilfskräfte sind. Sie können davon ausgehen, dass die Reparatur von jedem dieser Mitarbeiter mit derselben Wahrscheinlichkeit übernommen wird. Nach der Wartung durch einen der fest angestellten Mitarbeiter ist im Durchschnitt eines von 20 Fahrrädern noch defekt, nach der Wartung durch eine Aushilfskraft ist eines von 5 Fahrrädern noch defekt. Beim Abholen kontrollieren Sie das Fahrrad. Dabei erkennen und reklamieren Sie einen Defekt in 9 von 10 Fällen.
- (i) Formulieren Sie ein geeignetes stochastisches Modell.
  - (ii) Berechnen Sie die bedingte Wahrscheinlichkeit, dass ihr Fahrrad noch defekt ist, gegeben dass Sie es nicht reklamieren.
- b) Ein Geiger-Müller-Zählrohr  $Z$  und eine radioaktive Quelle  $Q$  seien so positioniert, dass ein Teilchen welches von  $Q$  emittiert wird, von  $Z$  mit Wahrscheinlichkeit  $10^{-4}$  registriert wird. Während der Beobachtungszeit emittiert  $Q$  genau 30.000 Teilchen.
- (i) Formulieren Sie ein geeignetes Modell.
  - (ii) Berechnen Sie approximativ die Wahrscheinlichkeit dafür, dass  $Z$  höchstens zwei Teilchen registriert.

**Lösung:**

- a) (i) Sei  $(\Omega, \mathcal{A}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $X, Y, Z : \Omega \rightarrow \{0, 1\}$  drei Zufallsvariablen. Dabei entspreche das Ereignis  $\{X = 1\}$ , dass das Fahrrad durch einen fest angestellten Mitarbeiter bearbeitet wurde und das Ereignis  $\{X = 0\}$  der Bearbeitung des Rads durch eine Aushilfskraft. Weiter sei  $\{Y = 1\}$  das Ereignis, dass das Fahrrad repariert werden konnte und  $\{Y = 0\}$ , dass es auch nach der Reparatur noch defekt ist. Zuletzt bezeichne  $\{Z = 0\}$  das Ausbleiben einer Reklamation und  $\{Z = 1\}$  dem Erkennen und Reklamieren eines Defektes. Dann ist die Verteilung von  $X$  gegeben durch

$$\mathbb{P}[X = 0] = \frac{2}{10} = 0,2 \quad \text{und} \quad \mathbb{P}[X = 1] = \frac{8}{10} = 0,8.$$

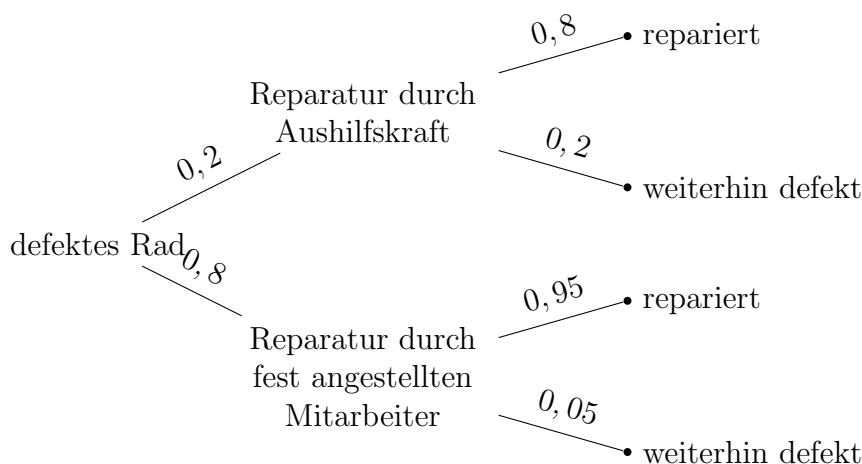
und die bedingte Verteilung von  $Y$  gegeben  $X$  ist gegeben durch

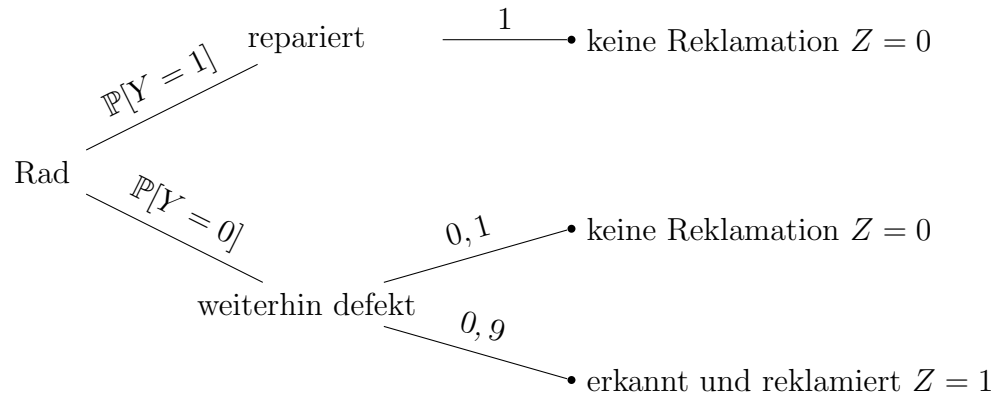
$$\begin{aligned} \mathbb{P}[Y = 0|X = 0] &= \frac{1}{5} = 0,2 \quad \text{und} \quad \mathbb{P}[Y = 0|X = 1] = \frac{1}{20} = 0,05, \\ \mathbb{P}[Y = 1|X = 0] &= \frac{4}{5} = 0,8 \quad \text{und} \quad \mathbb{P}[Y = 1|X = 1] = \frac{19}{20} = 0,95. \end{aligned}$$

weiter ist die bedingte Verteilung von  $Z$  gegeben  $Y$  gegeben durch

$$\begin{aligned} \mathbb{P}[Z = 0|Y = 0] &= \frac{1}{10} = 0,1 \quad \text{und} \quad \mathbb{P}[Z = 0|Y = 1] = 1, \\ \mathbb{P}[Z = 1|Y = 0] &= \frac{9}{10} = 0,9 \quad \text{und} \quad \mathbb{P}[Z = 1|Y = 1] = 0. \end{aligned}$$

Die folgenden Pfaddiagramme veranschaulichten das beschriebene mehrstufige Modell:





- (ii) Das Ereignis, dass das Fahrrad noch defekt ist, gegeben dass der Fehler nicht reklamiert wurde ist gegeben durch  $Y = 0$  bedingt auf  $Z = 0$ . Dazu berechnen wir zunächst die Wahrscheinlichkeit ein defektes Fahrrad zurückzubekommen aus der gegebenen bedingten Verteilung,

$$\begin{aligned} \mathbb{P}[Y = 0] &= \mathbb{P}[Y = 0|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = 0|X = 1]\mathbb{P}[X = 1] \\ &= \frac{1}{5} \cdot \frac{2}{10} + \frac{1}{20} \cdot \frac{8}{10} \\ &= \frac{2}{25} = \frac{8}{100}, \end{aligned}$$

und damit  $\mathbb{P}[Y = 1] = \frac{92}{100}$ . Also folgt mit der Bayesschen Regel

$$\begin{aligned} \mathbb{P}[Y = 0|Z = 0] &= \frac{P[Z = 0|Y = 0] \cdot \mathbb{P}[Y = 0]}{P[Z = 0|Y = 0] \cdot \mathbb{P}[Y = 0] + P[Z = 0|Y = 1] \cdot \mathbb{P}[Y = 1]} \\ &= \frac{\frac{1}{10} \cdot \frac{8}{100}}{\frac{1}{10} \cdot \frac{8}{100} + 1 \cdot \frac{92}{100}} = \frac{8}{8 + 920} = \frac{8}{928} = \frac{1}{116}. \end{aligned}$$

- b) (i) Sei  $(\Omega, \mathcal{A}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $Z : \Omega \rightarrow \{0, 1, \dots, 30.000\}$  eine Zufallsvariable, die die Anzahl der registrierten Teilchen während des Beobachtungszeitraums wiedergibt. Dann ist  $Z \sim \text{Bin}(30.000, 10^{-4})$ .
- (ii) Da die Ereignisse mit einer Wahrscheinlichkeit von  $\frac{1}{10.000}$  sehr selten auftreten, kann die Binomialverteilung durch die Poissonverteilung mit Parameter  $\lambda = np = 30.000 \cdot 10^{-4} = 3$  approximiert werden. Demnach ist

$$\begin{aligned} \mathbb{P}[Z \leq 2] &= \mathbb{P}[Z = 0] + \mathbb{P}[Z = 1] + \mathbb{P}[Z = 2] \\ &\approx \frac{3^0}{0!}e^{-3} + \frac{3^1}{1!}e^{-3} + \frac{3^2}{2!}e^{-3} \\ &= e^{-3} \left( 1 + 3 + \frac{9}{2} \right) \\ &= \frac{17}{2}e^{-3} \approx 0,423. \end{aligned}$$

#### 4. (Konfidenzintervalle)

[20 Punkte]

Bei einer Wahlumfrage unter 121 zufällig ausgewählten Personen gaben 50 an, für die Kandidatin A zu stimmen.

- a) Geben Sie ein statistisches Modell und einen Schätzer  $\hat{p}$  für den tatsächlichen Stimmenanteil  $p$  an.
- b) Ist der Schätzer erwartungstreu?
- c) Bestimmen Sie  $\epsilon > 0$  so, dass das Intervall  $[\hat{p} - \epsilon, \hat{p} + \epsilon]$  ein approximatives Konfidenzintervall zum Konfidenzniveau 95% für den unbekanntem Parameter  $p$  ist.
- d) Bestimmen Sie  $\delta > 0$  so, dass das Intervall  $[0, \hat{p} + \delta]$  ein approximatives Konfidenzintervall für  $p$  zum Konfidenzniveau 95% ist.
- e) Kann die Kandidatin noch hoffen, die absolute Mehrheit der Stimmen zu erhalten?

**Lösung:**

a) Seien  $X_1, \dots, X_{121}$  unter  $\mathbb{P}_p$  unabhängige und Bernoulli( $p$ )-verteilte Zufallsvariablen, wobei  $X_i = 1$  dafür steht, dass der  $i$ -te Wähler für die Kandidatin A stimmen wird, und sei  $S_n = X_1 + \dots + X_n$  die Anzahl der Stimmen für die Kandidatin A in unserer Stichprobe. In unserem Modell ist  $S_n$  also Bin ( $n, p$ )-verteilt mit  $n = 121$  und unbekannter Wahrscheinlichkeit  $p$ . Ein Schätzer für den tatsächlichen Stimmenanteil  $p$  ist  $\hat{p} := \frac{S_n}{n}$ .

b) Dieser ist erwartungstreu, denn

$$\mathbb{E}[\hat{p}] = \frac{\mathbb{E}[S_n]}{n} = \frac{\mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]}{n} = \frac{np}{n} = p$$

c) Für ein  $\varepsilon > 0$  approximieren wir

$$\begin{aligned} \mathbb{P}_p [|\hat{p} - p| \leq \varepsilon] &= \mathbb{P}_p \left[ \left| \frac{S_n}{n} - p \right| \leq \varepsilon \right] \\ &= \mathbb{P}_p \left[ \left| \frac{S_n - np}{\sqrt{np(1-p)}} \right| \leq \frac{\varepsilon n}{\sqrt{np(1-p)}} \right] \\ &\approx 2 \left( \Phi \left( \frac{\varepsilon \sqrt{n}}{\sqrt{p(1-p)}} \right) - \frac{1}{2} \right) \\ &\geq 2\Phi(2\varepsilon\sqrt{n}) - 1 \end{aligned}$$

wobei wir im letzten Schritt genutzt haben, dass  $p(1-p) \leq \frac{1}{4}$  für alle  $p \in [0, 1]$ . Nun wählen wir  $\varepsilon > 0$  sodass

$$2\Phi(2\varepsilon\sqrt{n}) - 1 \geq 0,95,$$

Nach der Tabelle der Verteilungsfunktion ist dies erfüllt, wenn  $2\varepsilon\sqrt{n} \geq 2$  gilt, also für

$$\varepsilon \geq \frac{2}{2\sqrt{n}} = \frac{2}{22} = \frac{1}{11}$$

Wir erhalten also das approximative Konfidenzintervall

$$[\hat{p} - \varepsilon, \hat{p} + \varepsilon] = \left[ \frac{50}{121} - \frac{1}{11}, \frac{50}{121} + \frac{1}{11} \right] = \left[ \frac{39}{121}, \frac{61}{121} \right].$$

d) Für ein  $\delta > 0$  approximieren wir

$$\begin{aligned} \mathbb{P}_p [p \leq \hat{p} + \delta] &= \mathbb{P}_p \left[ \frac{S_n}{n} - p \geq -\delta \right] \\ &= \mathbb{P}_p \left[ \frac{S_n - np}{\sqrt{np(1-p)}} \geq -\frac{\delta n}{\sqrt{np(1-p)}} \right] \\ &\approx 1 - \Phi \left( -\frac{\delta \sqrt{n}}{\sqrt{p(1-p)}} \right) = \Phi \left( \frac{\delta \sqrt{n}}{\sqrt{p(1-p)}} \right) \\ &\geq \Phi(2\delta\sqrt{n}), \end{aligned}$$

wobei wir im letzten Schritt wieder genutzt haben, dass  $p(1-p) \leq \frac{1}{4}$  für alle  $p \in [0, 1]$ . Nun wählen wir  $\delta > 0$  so, dass  $\Phi(2\delta\sqrt{n}) \geq 0,95$ . Nach Tabelle ist dies erfüllt, falls  $2\delta\sqrt{n} \geq 1,7$ , also für

$$\delta \geq \frac{1,7}{2\sqrt{n}} = \frac{0,85}{11}.$$

Wählen wir entsprechend  $\delta = \frac{0,85}{11}$ , dann erhalten wir das Konfidenzintervall

$$[0, \hat{p} + \delta] = \left[ 0, \frac{50}{121} + \frac{9,35}{121} \right] = \left[ 0, \frac{59,35}{121} \right].$$

- e) Der Wert  $1/2$  liegt im dem beidseitigen Konfidenzintervall aus c), aber nicht in dem einseitigen Konfidenzintervall aus d). Da wir nur entscheiden wollen, ob der tatsächliche Stimmenanteil der Kandidatin größer als  $1/2$  ist, ist das einseitige Konfidenzintervall genauer. Basierend darauf würden wir die Hypothese einer absoluten Mehrheit zum Konfidenzniveau 95% verwerfen. Dabei haben wir natürlich viele andere Unsicherheiten (Meinungsänderungen, Zusammensetzung der Stichprobe,...) außer acht gelassen.