Einführung in die Wahrscheinlichkeitstheorie und Statistik
Wintersemenster 2022/23
Dozent: Prof. Dr. Rainer Dahlhaus
Assistenz: Marilena Müller

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

# Extrablatt

**Aufgabe 1** (i) Sei $\Omega = \mathbb{N}_0$, $\mathcal{A} = \mathcal{P}(\Omega)$ und $p : \Omega \to [0,1]$ für $p \in (0,1)$, $r \in \mathbb{N}$, $\omega \in \Omega$ gegeben durch

$$p(\omega) = \binom{\omega + r - 1}{\omega} p^r (1-p)^\omega.$$

Zeigen Sie, dass $p$ eine Zähldichte eines Wahrscheinlichkeitsmaßes auf $(\Omega, \mathcal{A})$ ist. Was kann mit Hilfe dieser Zähldichte modelliert werden?
*Zeigen Sie die Regel* $\binom{\alpha+k-1}{k} = (-1)^k \binom{-\alpha}{k}$ *für den verallgemeinerten Binomialkoeffizienten* $\binom{\alpha}{k} := \frac{\alpha\cdot(\alpha-1)\cdot\ldots\cdot(\alpha-(k-1))}{k!}$, *der für* $\alpha \in \mathbb{Z}$ *und* $k \in \mathbb{N}_0$ *definiert ist. Beweisen Sie dann, dass für die binomische Reihe* $(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$ *gilt, wobei* $\alpha \in \mathbb{Z}$, $x \in (-1, 1)$.

(ii) Sei wieder $\Omega = \mathbb{N}_0$ und $\mathbb{P}$ als Wahrscheinlichkeitsverteilung auf $(\Omega, \mathcal{P}(\Omega))$ definiert durch

$$\mathbb{P}(\{n\}) := 2^{-n-1}$$

für $n \in \mathbb{N}_0$. Weiter sei eine Abbildung gegeben durch

$$X : \Omega \to \mathbb{R}, \quad n \mapsto n \mod 3.$$

Bestimmen Sie die induzierte Verteilung $\mathbb{P}^X$ auf $(\Omega^X, \mathcal{P}(\Omega^X)$.

(iii) Nach langjähriger Erfahrung wissen Sie, dass Sie bei einer Runde Skat mit Wahrscheinlichkeit $p = 0.2$ gewinnen. Sie sind nun zu einem Spieleabend eingeladen worden, bei dem ausschließlich Skat gespielt wird. Wie hoch ist die Wahrscheinlichkeit, dass Sie genau beim 30. Spiel zum sechsten Mal gewinnen?

**Aufgabe 2**

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{N}$ eine Zufallsvariable mit Werten in den natürlichen Zahlen.

(a) Es sei $X \sim \mathcal{G}(p)$ geometrisch verteilt mit Parameter $p \in (0,1)$. Zeigen Sie, dass für alle $k, n \in \mathbb{N}$ gilt, dass

$$\mathbb{P}(X = k) = \mathbb{P}(X = n - 1 + k \mid X \geq n). \tag{1}$$

Diese Eigenschaft wird als "Gedächtnislosigkeit" der geometrischen Verteilung bezeichnet.

(b) Zeigen Sie, dass eine Zufallsvariable $X : \Omega \to \mathbb{N}$, die (1) erfüllt, geometrisch verteilt ist mit einem geeigneten Parameter $p$.
*Hinweis: Definieren Sie* $p := \mathbb{P}(X = 1)$, *nutzen Sie (1) mit* $k = 1$ *und leiten Sie so eine rekursive Beziehung der Form* $\mathbb{P}(X \geq n+1) = c \cdot \mathbb{P}(X \geq n)$ *her mit geeignetem* $c \in \mathbb{R}$.

## Aufgabe 3

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}$ eine steig verteilte Zufallsvariable, die einer Pareto-Verteilung folgt, d.h. $X \sim \text{Pareto}(\alpha, x_m)$ mit Parametern $\alpha > 0$ und $x_m > 0$. Die Dichte ist gegeben durch

$$f_X(x) = C_{\alpha,x_m} \cdot x^{-(\alpha+1)} \mathbb{1}_{\{x \geq x_m\}} = \begin{cases} C_{\alpha,x_m} \cdot x^{-(\alpha+1)}, & x \geq x_m, \\ 0, & x < x_m \end{cases}.$$

(a) Bestimmen Sie die Konstante $C_{\alpha,x_m}$ so, dass $f_X(x)$ eine Wahrscheinlichkeitsdichte ist.

(b) Bestimmen Sie die Verteilungsfunktion $F_X(x)$ von $X$.

(c) Welche bekannte Verteilung besitzt die Zufallsvariable $Y$ mit $Y = \log\left(\frac{X}{x_m}\right)$ ?

(d) Sei nun $\alpha = x_m = 1$. Berechnen Sie die Wahrscheinlichkeiten $\mathbb{P}(1 \leq X \leq 2)$ und $\mathbb{P}(X > 2)$.

## Aufgabe 4

Wir untersuchen eine Maschine, die Schrauben herstellt. Uns ist bekannt (und das zweifeln wir nicht an), dass die Länge der Schrauben im Mittel $\mu = 5$ beträgt (Angabe in cm). Der Hersteller behauptet, dass die Standardabweichung der Länge der Schrauben höchstens $\sigma_0 = 0.3$ cm beträgt. Wir vermuten jedoch, dass die Standardabweichung tatsächlich größer ist. Für eine statistische Untersuchung nehmen wir an, dass die Länge der Schrauben normalverteilt ist mit Mittelwert $\mu$ und Standardabweichung $\sigma$. Für $n = 10$ Schrauben beobachten wir folgende Längen (in cm):

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 5.6 | 5.2 | 4.1 | 3.7 | 6.5 | 3.6 | 6.0 | 6.1 | 5.2 | 4.2 |

(a) Sei $n \in \mathbb{N}$ und seien $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ mit $\sigma > 0$. Nutzen Sie das auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ verallgemeinerte Neyman-Pearson-Lemma 6.19, um einen besten Test $\phi^* : \mathbb{R}^n \to \{0, 1\}$ für die einfachen Hypothesen

$$H_0 : \sigma = \sigma_0 \quad \text{gegen} \quad H_1 : \sigma = \sigma_1$$

mit $\sigma_1 > \sigma_0$ zum Niveau $\alpha \in (0, 1)$ anzugeben.
*Hinweis: Bestimmen Sie zunächst die Dichte $f_\sigma$ der gemeinsamen Verteilung $\mathbb{P}_\sigma = \mathbb{P}^{X_1,...,X_n}$ von $X_1, ..., X_n$.*

(b) Zeigen Sie mittels der Technik der monotonen Likelihood-Quotienten, dass $\phi^*$ die Form

$$\phi^*(x) = \begin{cases} 1, & T(x) > c^{**}, \\ 0, & T(x) \leq c^{**} \end{cases}$$

besitzt, wobei $T(x) = \sum_{i=1}^n (x_i - \mu)^2$. Folgern Sie, dass $\phi^*$ sogar gleichmäßig bester Test ist für

$$H_0 : \sigma = \sigma_0 \quad \text{gegen} \quad H_1' : \sigma > \sigma_0.$$

(c) Zeigen Sie, dass $\phi^*$ sogar gleichmäßig bester Test ist für

$$H_0' : \sigma \leq \sigma_0 \quad \text{gegen} \quad H_1' : \sigma > \sigma_0.$$

*Hinweis: Definieren Sie $Z_n := \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2$. Überlegen Sie sich, dass die Verteilung von $Z_n$ nicht mehr von $\sigma$ abhängt. Zeigen Sie dann, dass $\mathbb{P}_\sigma(\phi^* = 1) \leq \alpha$ gilt für $\sigma \leq \sigma_0$.*

(d) Es ist bekannt, dass $Z_n \sim \chi^2_n$, wobei $\chi^2_n$ die Chi-Quadrat-Verteilung mit $n$ Freiheitsgraden bezeichnet. Es bezeichne $\chi^2_{n,1-\alpha}$ das $(1-\alpha)$-Quantil dieser Verteilung. Zeigen Sie, dass $c^{**} = \sigma_0^2 \cdot \chi^2_{n,1-\alpha}$.

(e) Wir drücken nun die Abhängigkeit des Tests $\phi^*$ von $\sigma_0$ explizit aus, indem wir ihn mit $\phi^*_{\sigma_0}$ bezeichnen. Bestimmen Sie den Annahmebereich $A(\sigma_0) := \{x \in \mathbb{R}^n : \phi^*_{\sigma_0}(x) = 0\}$ des Tests und damit ein gleichmäßig bestes $(1-\alpha)$-Konfidenzintervall $S(X)$ für $\sigma$. Für welche falschen Parameter wurde dieses Intervall konstruiert?

(f) Sei nun $\alpha = 0.05$. Werden Sie den Hersteller der Maschine auf Basis unserer Beobachtungen und des Tests $\phi^*$ aus (b) vorwerfen, dass die Angabe der Standardabweichung falsch ist? Geben Sie auf Basis unserer Beobachtungen das 95%-Konfidenzintervall für $\sigma$ aus (e) an.
*Hinweis: Hier sind einige Quantile der $\chi^2_n$-Verteilung:* $\chi^2_{10,0.05} = 3.94$, $\chi^2_{10,0.95} = 18.31$.

**Aufgabe 5**

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und seien $X, Y : \Omega \to \mathbb{R}$ zwei Zufallsvariablen.

(a) Seien $X, Y$ stochastisch unabhängig. Zeigen Sie, dass gilt

$$\begin{aligned}\mathrm{Kov}(XY, Y) &= \mathbb{E}[X] \cdot \mathrm{Var}(Y),\\ \mathrm{Var}(XY) &= \mathrm{Var}(X)\mathrm{Var}(Y) + \mathbb{E}[X]^2\ \mathrm{Var}(Y) + \mathbb{E}[Y]^2\mathrm{Var}(X).\end{aligned}$$

(b) Gelte $\mathrm{Var}(X) = \mathrm{Var}(Y)$. Zeigen Sie, dass

$$\mathrm{Kov}(X+Y, X-Y) = 0.$$

(c) Seien $X, Y \sim \mathrm{Ber}(1,p)$ unabhängig und Bernoulli-verteilt mit Parameter $p \in (0,1)$. Zeigen Sie, dass $X+Y$ und $X-Y$ nicht stochastisch unabhängig sind.

Die Lösungen sind nicht abzugeben und werden dementsprechend nicht korrigiert. Dieses Blatt wird in den Übungen zwischen **Montag, dem 9. Januar 2023** und **Mittwoch, dem 11. Januar 2023** besprochen.

Schöne Feiertage!