Aufgabe 1

Punkten

Gegeben seien stochastisch unabhängige Zufallsvariablen X, Y, Z mit $X \sim \mathrm{bin}(1, \frac{1}{5})$, $Y \sim \mathrm{bin}(1, \frac{1}{5})$ und $Z \sim \mathrm{N}(3, 9)$.
Hierbei bezeichnen $\mathrm{bin}(n, p)$ für $n \in \mathbb{N}$ und $p \in (0, 1)$ die Binomialverteilung mit Zähldichte $f_{n,p} : \{0, \ldots, n\} \longrightarrow [0, \infty)$, definiert durch

$$f_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0, \ldots, n\},$$

sowie $\mathrm{N}(\mu, \sigma^2)$ für $\mu \in \mathbb{R}$ und $\sigma > 0$ die Normalverteilung mit Riemann-Dichte $\varphi_{\mu,\sigma^2} : \mathbb{R} \longrightarrow [0, \infty)$, definiert durch

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

(mit $\exp$ als Exponentialfunktion).

(a) Markieren Sie die **einzige richtige** Aussage für die Verteilung der Zufallsvariablen $X + Y$:

Punkten

- ○ $X + Y \sim \mathrm{N}(2, \frac{2}{5})$
- ○ $X + Y \sim \mathrm{bin}(2, \frac{2}{5})$
- ● $X + Y \sim \mathrm{bin}(2, \frac{1}{5})$ ✔
- ○ $X + Y \sim \mathrm{bin}(1, \frac{1}{5})$
- ○ $X + Y \sim \mathrm{bin}(1, \frac{2}{5})$

(b) Berechnen Sie den Erwartungswert $\mathrm{E}(5Y + Z)$.

Punkten

2 von 2 Punkten

Ergebniseingabe:

(c) Berechnen Sie die Varianz $\mathrm{Var}(Z - 5X)$.

unkten

2 von 2 Punkten

Ergebniseingabe:

(d) Berechnen Sie die Kovarianz $\mathrm{Kov}(-5Y, 10Y)$.

Punkten

3 von 3 Punkten

Ergebniseingabe:

Aufgabe 2

Punkten

Die neu entwickelte Übersetzer-App "Translati" wurden in den bisherigen Testläufen für Übersetzungen von Englisch, Französisch oder Spanisch jeweils in die deutsche Sprache überprüft. Hierbei ergaben sich die folgenden Erfolgsquoten:

(i) Ein zufällig ausgewähltes englisches Wort wurde mit Wahrscheinlichkeit $0{,}9$ fehlerfrei übersetzt.

(ii) Ein zufällig ausgewähltes französisches Wort wurde mit Wahrscheinlichkeit $0{,}6$ fehlerfrei übersetzt.

(iii) Ein zufällig ausgewähltes spanisches Wort wurde mit Wahrscheinlichkeit $0{,}8$ fehlerfrei übersetzt.

In einer ersten Anwendung sollen nun verschiedene Texte mit insgesamt 400 englischen Wörtern, 400 französischen Wörtern und 200 spanischen Wörtern in die deutsche Sprache übersetzt werden. Aus diesen insgesamt 1000 Wörtern wird eines zufällig ausgewählt.

Hinweis: Fassen Sie in Ihrer Lösung die in der Aufgabenstellung enthaltenen relativen Häufigkeiten jeweils als Wahrscheinlichkeiten auf.

(a) Berechnen Sie die Wahrscheinlichkeit dafür, dass das ausgewählte Wort fehlerfrei übersetzt wird.

Punkten

kten

Ergebniseingabe:

(b) Berechnen Sie die Wahrscheinlichkeit dafür, dass das ausgewählte Wort ein englisches oder spanisches ist **und** dass es fehlerfrei übersetzt wird.

nkten

Punkten

Ergebniseingabe:

(c) Berechnen Sie die (bedingte) Wahrscheinlichkeit dafür, dass das ausgewählte Wort ein französisches ist unter der Bedingung, dass das ausgewählte Wort kein spanisches ist.

unkten

Punkten

Ergebniseingabe:

Aufgabe 3 ...unkten

Es seien $X_1,\dots,X_n$ stochastisch unabhängige, identisch verteilte, diskrete Zufallsvariablen mit Werten in $\mathbb{N}_0$. Für $i \in \{1,\dots,n\}$ sei die Zähldichte von X_i gegeben durch

$$P_\vartheta(X_i = k) = (k+1)\,\frac{(\vartheta-1)^k}{\vartheta^{k+2}} \quad \text{für } k \in \mathbb{N}_0$$

mit einem (unbekannten) Parameter $\vartheta \in (1,\infty)$. Dann gilt (und kann ohne eigenen Nachweis verwendet werden):

$$E_\vartheta(X_i) = 2\,(\vartheta-1) \quad \text{für } \vartheta \in (1,\infty) \text{ und } i \in \{1,\dots,n\}.$$

Zu gegebenen Realisationen $x_1,\dots,x_n \in \mathbb{N}_0$ von $X_1,\dots,X_n$ mit arithmetischem Mittelwert $\overline{x} > 0$ bezeichnen $\vartheta \mapsto L(\vartheta \mid x_1,\dots,x_n)$ die zugehörige Likelihood-Funktion, $\vartheta \mapsto l(\vartheta) = \ln\left(L(\vartheta \mid x_1,\dots,x_n)\right)$ die zugehörige Log-Likelihood-Funktion und $\hat{\vartheta}(x_1,\dots,x_n) \in (0,1)$ die zugehörige Maximum-Likelihood-Schätzung, die in der vorliegenden Situation eindeutig bestimmt ist.

(a) Bitte kreuzen Sie die **einzige richtige** Aussage an, die für beliebige $n \in \mathbb{N}$ sowie beliebige $x_1,\dots,x_n \in \mathbb{N}_0$ mit $\overline{x} > 0$ gültig ist. ...Punkten

- [ ] Die Log-Likelihood-Funktion hat die Darstellung

$$l(\vartheta) = \prod_{i=1}^{n} \ln(x_i+1) + n\,\overline{x}\,\ln(\vartheta-1) - n\,(\overline{x}+2)\,\ln(\vartheta)\ , \quad \vartheta \in (1,\infty).$$

- [ ] Die erste Ableitung der Log-Likelihood-Funktion hat die Darstellung ✓

$$l'(\vartheta) = \frac{n\,\overline{x}}{\vartheta-1} - \frac{n\,(\overline{x}+2)}{\vartheta}\ , \quad \vartheta \in (1,\infty).$$

- [x] Die Maximum-Likelihood-Schätzung hat die folgende Darstellung:

$$\hat{\vartheta}(x_1,\dots,x_n) = \overline{x}.$$

- [ ] Die Likelihood-Funktion hat die Darstellung

$$L(\vartheta \mid x_1,\dots,x_n) = (k+1)^n\,(\vartheta-1)^{nk}\,\vartheta^{-n(k+2)}\ , \quad \vartheta \in (1,\infty)\ , \quad k \in \mathbb{N}_0.$$

(b) Bei gegebenem Stichprobenumfang $n = 7$ werde für die Schätzung des Parameters ϑ der Schätzer $\tilde{\vartheta}(X_1,\dots,X_7) = \frac{1}{c}\sum_{i=1}^{7}(X_i+2)$ mit einem $c > 0$ verwendet. ...Punkten

Berechnen Sie das c, für welches $\tilde{\vartheta}(X_1,\dots,X_7)$ erwartungstreu für die Schätzung von ϑ ist.

...kten

Ergebniseingabe:

Aufgabe 4

(a) Gegeben sei der folgende metrische Datensatz:

$$x_1 = 4\,,\ x_2 = 8\,,\ x_3 = 11\,,\ x_4 = 3\,,\ x_5 = 4\,,$$
$$x_6 = 5\,,\ x_7 = 6\,,\ x_8 = 9\,,\ x_9 = 2\,,\ x_{10} = 8\,.$$

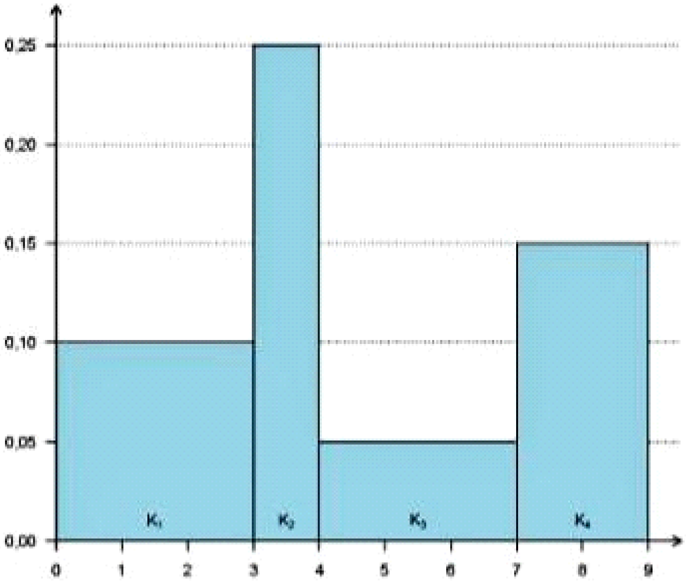
(i) Berechnen Sie das $0{,}15$-Quantil zu $x_1, \ldots, x_{10}$. unkten

Ergebniseingabe:

(ii) Ermitteln Sie den Wert der empirischen Verteilungsfunktion zu $x_1, \ldots, x_{10}$ an der Stelle 5.

Ergebniseingabe:

(b) Gegeben sei ein Datensatz mit 200 Beobachtungen eines metrischen Merkmals, zu dem das folgende Histogramm erstellt wurde.



Markieren Sie diejenige der folgenden Aussagen, die als **einzige falsch** ist: **4 von 4 Punkten**

- Die Klasse K_1 enthält doppelt so viele Beobachtungen wie die Klasse K_3.
- Die relative Klassenhäufigkeit der Klasse K_3 beträgt $0{,}15$.
- Die absolute Klassenhäufigkeit der Klasse K_4 beträgt 60.
- Die Klassen K_2 und K_3 enthalten zusammen 80 Beobachtungen.
- Die relativen Klassenhäufigkeiten der Klassen K_1 und K_4 sind gleich.
- Die Klasse K_2 enthält mehr Beobachtungen als die Klasse K_1. ✔

Aufgabe 4

‾ ınkten

(a) Gegeben sei der folgende metrische Datensatz:

$$x_1 = 4\,,\ x_2 = 8\,,\ x_3 = 11\,,\ x_4 = 3\,,\ x_5 = 4\,,$$
$$x_6 = 8\,,\ x_7 = 6\,,\ x_8 = 9\,,\ x_9 = 2\,,\ x_{10} = 8\,.$$

(i) Berechnen Sie das $0{,}15$-Quantil zu $x_1, \ldots, x_{10}$.

ınkten

ınkten

Ergebniseingabe:

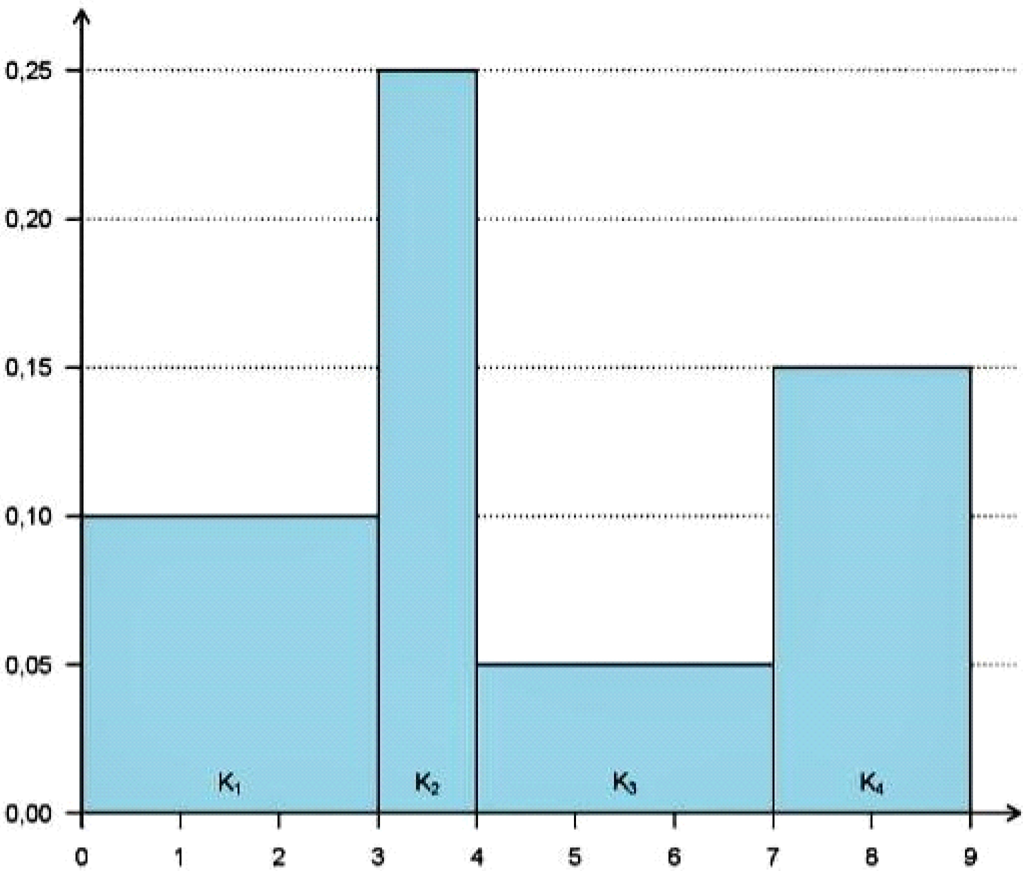
(ii) Ermitteln Sie den Wert der empirischen Verteilungsfunktion zu $x_1, \ldots, x_{10}$ an der Stelle 5.

ınkten

en

Ergebniseingabe:

(b) Gegeben sei ein Datensatz mit 200 Beobachtungen eines metrischen Merkmals, zu dem das folgende Histogramm erstellt wurde:



Markieren Sie diejenige der folgenden Aussagen, die als **einzige falsch** ist:

- Die Klasse K_1 enthält doppelt so viele Beobachtungen wie die Klasse K_3.
- Die relative Klassenhäufigkeit der Klasse K_3 beträgt $0{,}15$.
- Die absolute Klassenhäufigkeit der Klasse K_4 beträgt 60.
- Die Klassen K_2 und K_3 enthalten zusammen 80 Beobachtungen.
- Die relativen Klassenhäufigkeiten der Klassen K_1 und K_4 sind gleich.
- Die Klasse K_2 enthält mehr Beobachtungen als die Klasse K_1. ✓

Aufgabe 5 Punkten

Gegeben seien $n = 10$ Beobachtungspaare $(x_1, y_1), \ldots, (x_{10}, y_{10})$ zweier metrischer Merkmale X und Y, für die das übliche Modell der linearen Regression zugrunde gelegt werde, d.h.

$$Y_i = a + bx_i + \varepsilon_i \quad \text{für} \quad i \in \{1, \ldots, 10\}.$$

Hierbei seien $a, b \in \mathbb{R}$ (unbekannte) Regressionskoeffizienten, und die Fehler $\varepsilon_1, \ldots, \varepsilon_{10}$ seien stochastisch unabhängig und jeweils $\mathrm{N}(0, \sigma^2)$-verteilt mit unbekannter Varianz $\sigma^2 \in (0, \infty)$.

Zu den Beobachtungspaaren werden die folgenden statistischen Kenngrößen ermittelt:

$$\sum_{i=1}^{10} (x_i - \overline{x})^2 = 40 \ , \quad \sum_{i=1}^{10} (y_i - \overline{y})^2 = 4000 \ ,$$

$$\sum_{i=1}^{10} (x_i - \overline{x})(y_i - \overline{y}) = -360 \quad \text{und} \quad \sum_{i=1}^{10} \left(y_i - (\hat{a} + \hat{b}\, x_i)\right)^2 = 760 \ .$$

Hierbei bezeichnen $\overline{x}, \overline{y}$ die arithmetischen Mittelwerte zu $x_1, \ldots, x_{10}$ bzw. $y_1, \ldots, y_{10}$ und $\hat{a}, \hat{b}$ die Kleinste-Quadrate-Schätzungen der Regressionskoeffizienten a und b.

(a) Berechnen Sie mit Hilfe der gegebenen Kenngrößen die Kleinste-Quadrate-Schätzung $\hat{b}$ für den Regressionskoeffizienten b.

nkt

Ergebniseingabe:

(b) Berechnen Sie mit Hilfe der gegebenen Kenngrößen den Bravais-Pearson-Korrelationskoeffizienten r_{xy} zu $(x_1, y_1), \ldots, (x_{10}, y_{10})$.

nkten

Ergebniseingabe:

(c) Berechnen Sie mit Hilfe der gegebenen Kenngrößen die obere Intervallgrenze des unteren einseitigen 95%-Konfidenzintervalls für die unbekannte Varianz σ^2. Benötigte Quantile oder Verteilungsfunktionswerte können Sie gegebenenfalls der unten angehängten Datei entnehmen.

unkten

0 von 3 Punkten

Eingabe der **oberen Grenze** des unteren einseitigen Konfidenzintervalls für σ^2 (Rundung des Zahlenwertes auf **drei Dezimalstellen**):

(d) Berechnen Sie mit Hilfe der gegebenen Kenngrößen die Länge des zweiseitigen 95%-Konfidenzintervalls für den (unbekannten) Regressionskoeffizienten b. Benötigte Quantile oder Verteilungsfunktionswerte können Sie gegebenenfalls der unten angehängten Datei entnehmen.

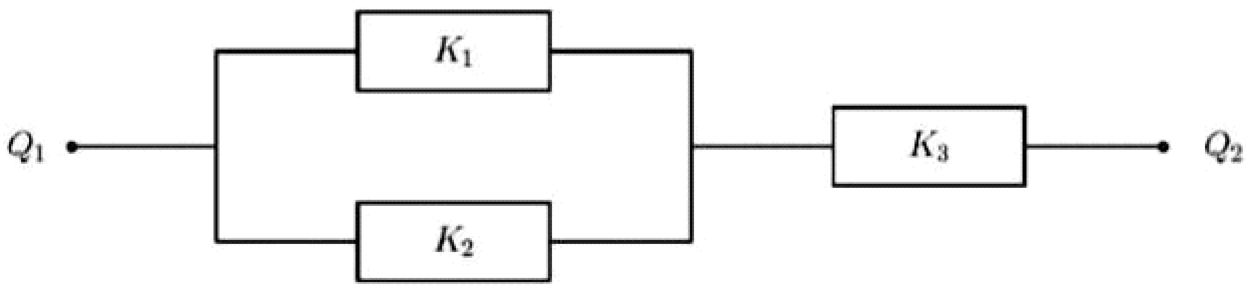
Punkten

Eingabe der **Länge** des zweiseitigen Konfidenzintervalls für b (Rundung des Zahlenwertes auf **drei Dezimalstellen**):

Aufgabe 6

kten

Das folgende System besteht aus drei Komponenten K_1, K_2 und K_3, die unabhängig voneinander jeweils mit Wahrscheinlichkeit $p \in (0,1)$ **intakt** sind. Das gesamte System ist intakt, wenn zwischen den Knotenpunkten Q_1 und Q_2 mindestens eine Verbindung aus intakten Komponenten besteht.



Welche der folgenden Formeln gibt die Wahrscheinlichkeit dafür an, dass das System **nicht intakt** ist? Kreuzen Sie bitte die **einzig richtige** Formel an, die für beliebige $p \in (0,1)$ gültig ist.

Punkten

- $p + p^2 - p^3$
- $1 - p - p^2 + p^3$
- $1 - p + (1-p)^2 - (1-p)^3$
- $2p^2 - p^3$
- $2(1-p)^2 - (1-p)^3$
- $1 - 2p^2 + p^3$

Aufgabe 7 Punkten en

Gegeben seien zwei stochastisch unabhängige, identisch verteilte, stetige Zufallsvariablen X und Y auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$. Hierbei ist die zugehörige Riemann-Dichtefunktion $f^X = f^Y : \mathbb{R} \longrightarrow [0, \infty)$ von X bzw. Y gegeben durch

$$f^X(t) = f^Y(t) = \begin{cases} \dfrac{1}{10\sqrt{t}} & \text{für} \quad t \in [1, 36], \\ 0 & \text{für} \quad t \in \mathbb{R}\setminus[1, 36]. \end{cases}$$

(a) Betrachten Sie zu den gegebenen Voraussetzungen die folgenden drei Aussagen:

(1) Eine Riemann-Dichtefunktion $f^{(X,Y)} : \mathbb{R}^2 \longrightarrow [0, \infty)$ des Zufallsvektors (X, Y) ist gegeben durch

$$f^{(X,Y)}(x, y) = \begin{cases} \dfrac{1}{100\sqrt{x\,y}} & \text{für} \quad (x, y) \in [1, 36]^2, \\ 0 & \text{für} \quad (x, y) \in \mathbb{R}^2\setminus[1, 36]^2. \end{cases}$$

(2) Für $y \in \mathbb{R}$ ist eine bedingte Riemann-Dichtefunktion $f^{X \mid Y=y} : \mathbb{R} \longrightarrow [0, \infty)$ von X unter der Bedingung $Y = y$ gegeben durch

$$f^{X \mid Y=y}(x) = \begin{cases} \dfrac{1}{10\sqrt{x}} & \text{für} \quad x \in [1, 36], \\ 0 & \text{für} \quad x \in \mathbb{R}\setminus[1, 36]. \end{cases}$$

(3) Es gilt:

$$\mathrm{E}(X\,Y) = \mathrm{E}(X)\,\mathrm{E}(Y).$$

Markieren Sie die **einzig richtige** der folgenden Angaben. ıkten

- Die Aussagen (1) und (3) sind jeweils richtig. Aussage (2) ist falsch.
- Alle drei Aussagen sind richtig. ✔
- Die Aussagen (1) und (2) sind jeweils richtig. Aussage (3) ist falsch.
- Nur Aussage (1) ist richtig.
- Die Aussagen (2) und (3) sind jeweils richtig. Aussage (1) ist falsch.

(b) Berechnen Sie den Erwartungswert $\mathrm{E}(2\sqrt{X} - 5)$. Punkten

Punkten

Ergebniseingabe:

(c) Berechnen Sie den Wert $F^X(16)$ der Verteilungsfunktion F^X von X an der Stelle 16. kten

kten

Ergebniseingabe:

(d) Berechnen Sie die Wahrscheinlichkeit $P(X > 16, Y > 16)$. nkten

Ergebniseingabe:

Aufgabe 8

. unkten

In einer Stadt kommt es an einer bestimmten Kreuzung regelmäßig zu Verkehrsstaus. Die Stadtverwaltung ist interessiert an der Gesamtzahl der Verkehrsstaus, die an dieser Kreuzung in den 256 Arbeitstagen eines Jahres vorkommen.

Hierbei kann angenommen werden, dass die Anzahlen der an der Kreuzung auftretenden Verkehrsstaus in den verschiedenen Arbeitstagen durch stochastisch unabhängige, jeweils mit Parameter $\lambda = 9$ Poisson-verteilte Zufallsvariablen $X_1, \ldots, X_{256}$ beschrieben werden können.

Berechnen Sie mit Hilfe des Zentralen Grenzwertsatzes **approximativ** die Wahrscheinlichkeit dafür, dass in den 256 Arbeitstagen insgesamt höchstens 2340 Verkehrsstaus an der Kreuzung vorkommen.

?unkten

?unkten

Hinweis: Angabe des Zahlenwertes gerundet auf **drei** Dezimalstellen. Entnehmen Sie gegebenenfalls benötige Verteilungsfunktionswerte oder Quantile der unten angehängten Datei!

Ergebniseingabe:

Anhang: Verteilungsfunktionswerte und Quantile der Standardnormalverteilung

Anhang öffnen