I.1. (2P) Welche Aussage ist wahr?

Für zwei standardisierte Stichproben $x_1^*, \dots x_n^*$ und $y_1^*, \dots, y_n^*$ gilt immer:

A. $\bar{x}^* = 1$, $V_{x^*} = 0$
B. $\rho_{x^* y^*} = \mathrm{Cov}_{x^* y^*}$
C. x_i^* und y_i^* haben die gleiche Verteilung
D. $\bar{x}^* = \bar{y}^*$ und $\mathrm{Cov}_{x^* y^*} = 1$

I.2. (2P) Für zwei Stichproben $x_1, \dots, x_n$ und $y_1, \dots, y_n$ gelte $\rho_{xy} = 1$ und $V_x > 0, V_y > 0$. Welche Aussage ist richtig?

A. Die zugehörige Regressionsgerade hat Steigung 1.
B. Die zugehörige Regressionsgerade geht durch den Ursprung.
C. Es gilt $C_{xy} = \rho_{xy}$.
D. Es gibt $a, b \in \mathbb{R}$ so dass $x_i = a y_i + b$ für alle $i = 1, \dots, n$ gilt.

I.3. (2P) Gegeben seien zwei Ereignisse A und B mit $P(A) = 0.1$ und $P(B) = 0.2$. Was lässt sich allgemein über $P(A \cap B)$ sagen?

A. $P(A \cap B) = 0$
B. Es ist keine Aussage möglich.
C. $P(A \cap B) = 0.02$
D. $P(A \cap B) = -0.02$

I.4. (2P) Es sei V das Ereignis, dass sich eine Person vegan ernährt und M das Ereignis, dass eine Person eine Mangelerscheinung hat. Welche Antwort sagt, dass die Wahrscheinlichkeit für eine Magelerscheinung bei einer vegan lebenden Person größer ist, als eine Magelerscheinung bei einer nicht-veganen Person?
A. $P(M|V) > P(M|V^c)$
B. $P(V|M) > P(V^c|M)$
C. $P(M|V) < P(M|V^c)$
D. $P(V \cap M) > P(V^c \cap M)$

I.5. (2P) Die Wahrscheinlichkeit, dass bei der Herstellung eines Werkstücks ein bestimmter Materialfehler auftritt ist 0.05. Die Monatsproduktion umfasst 400 Werkstücke.

Dann ist die zufällige Anzahl der fehlerhaften Werkstücke

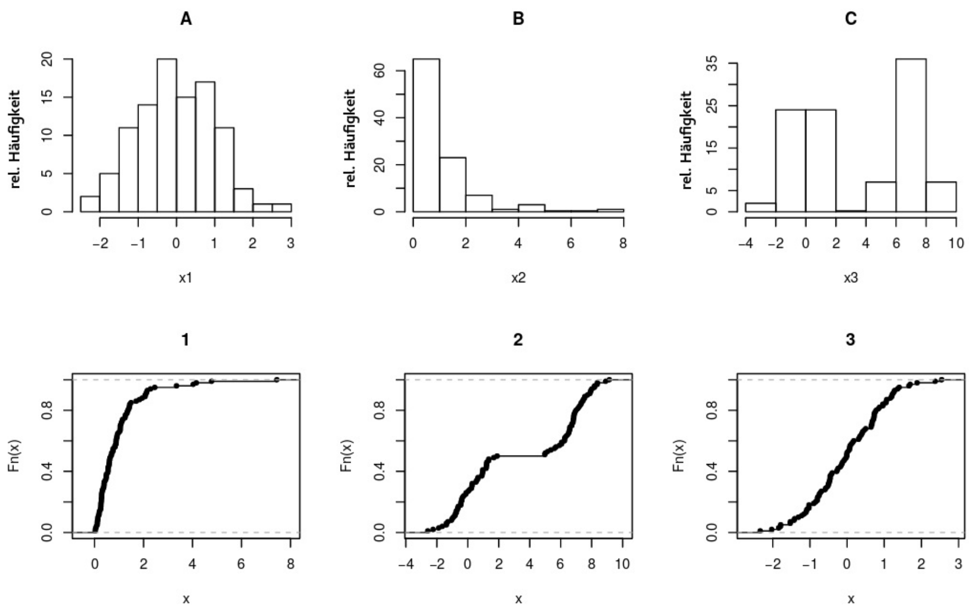
A. binomialverteilt mit Parametern $n = 400$ und $p = 0.05$
B. geometrisch verteilt mit Parameter $q = 0.05$
C. geometrisch verteilt mit Parameter $q = 0.95$
D. binomialverteilt mit Parametern $n = 400$ und $p = 0.95$

I.6. (2P) Welche der folgenden Aussagen ist richtig? A, B und C seien drei unterschiedliche Ereignisse eines Zufallsexperiments. Weiterhin gelte $A \cap B = C$ sowie $P(C) \neq 0$. Dann gilt immer:

A. $P(C) = P(B|A) \cdot P(B)$
B. $P(B) \leq P(C)$
C. $P(C) = P(A|B) \cdot P(B)$
D. $P(C) = P(A) + P(B) - P(A \cap B)$.

I.7. (2P) Die untenstehende Abbildung zeigt 3 Histogramme (A,B,C) und 3 empirische Verteilungsfunktionen (1,2,3). Ordne jedem Histogramm seine empirische Verteilungsfunktion zu.

A. 1A,2B,3C B. 2A,3B,1C C. 3A,1B,2C D. 1A,2C,3B



I.8. (2P) Gegeben seien die Zufallsvariablen X und Y von denen bekannt sei, dass
$Y = 0.5X + 1$, $\mathbb{E}(X) = 3$ und $Var(Y) = 4$.
Welche Aussage ist richtig?
A. $\mathbb{E}(Y) = 1.5$
B. $Var(X) = 16$
C. $Var(X) = 6$
D. keine der anderen Antwortmöglichkeiten stimmt.

I.9. (2P) Sie haben einen Weihnachtsbaum von dem sich 15 Zweige gut eignen um eine Kugel daran aufzuhängen. Sie besitzen 10 identische rote Kugeln. Wie viele Möglichkeiten haben Sie dann den Baum zu schmücken?
A. 10^{15} B. $\binom{15}{10}$ C. $\frac{15!}{5!}$ D. $\binom{24}{10}$

I.10. (2P) Jede Einzelmessung habe eine Standardabweichung von 2. Wie groß ist die Standardabweichung des Mittelwerts von 16 unabhängigen Messungen?
A. 0.125 B. 2 C. 32 D. 0.5

TEIL II: Rechenaufgaben

Bemerkungen:

- **Alle Ergebnisse und Arbeitsschritte sind zu begründen.**
- **Ergebnisse ohne Lösungswege werden nicht gewertet.**

II.1. (13P) In der Gemeinde Unterleuten gibt es zwei Postdienstleister. 70% der Briefe werden mit A-Post und 30% der Briefe mit B-Post verschickt. Aus Erfahrung wissen wir, dass ein A-Post-Brief mit einer Wahrscheinlichkeit von 3.2% **nicht** rechtzeitig beim Empfänger ankommt. Im Falle der B-Post liegt diese Wahrscheinlichkeit bei 1.5%.

Wir definieren folgende Ereignisse:

- A: „Brief wird mit A-Post versandt.“
- B: „Brief wird mit B-Post versandt.“
- R: „Brief trifft rechtzeitig beim Empfänger ein.“
- R^C: „Brief trifft nicht rechtzeitig beim Empfänger ein.“

(a) (2P) Geben Sie die im Aufgabentext gegebenen Wahrscheinlichkeiten an, z.B. $P(A) = ...$

(b) (2P) Sie senden einen wichtigen Brief per A-Post. Wie groß ist die Wahrscheinlichkeit, dass dieser Brief rechtzeitig ankommt? (Ersatzresultat: 0.95).

(c) (3P) Wie groß ist die Wahrscheinlichkeit, dass ein beliebiger Brief **nicht** rechtzeitig ankommt? (Ersatzresultat: 0.04)

(d) (3P) Sie haben einen Brief verspätet erhalten. Wie groß ist die Wahrscheinlichkeit, dass dieser Brief mit der B-Post versandt wurde?

(e) (3P) Sind die Ereignisse B und R^C unabhängig?

II.2. (18P)Sie betreiben ein Theater mit 20 Sitzplätzen. Aus Erfahrung wissen Sie, dass die Wahrscheinlichkeit dafür, dass ein verkauftes Ticket ungenutzt bleibt, 0.05 beträgt. Sie gehen weiterhin davon aus, dass das Erscheinen eines Gastes unabhängig vom Erscheinen jedes anderen Gastes ist.
Sei X die Zufallsgröße, welche die Anzahl der nicht erschienenen Gäste bei n verkauften Tickets angibt.

(a) (2P) Welche Verteilung besitzt X? Was sind die zugehörigen Parameter?

(b) (2P) Wie viele Gäste erscheinen im Mittel nicht zu einer Aufführung?

(c) (3P) Sie spekulieren auf höhere Einnahmen und verkaufen 21 Tickts, obwohl es nur 20 Plätze gibt. Wie groß ist die Wahrscheinlichkeit, dass Sie zu wenige Sitzplätze haben?

(d) (5P) Sie haben 21 Tickets verkauft und wissen bereits, dass 19 Gäste sicher kommen. Wie groß ist dann die Wahrscheinlichkeit, dass es zu wenige Sitzplätze gibt?

(e) (4P) Ein Kollege besitzt ein Theater mit 195 Sitzplätzen und ist angetan von Ihren Rechnungen. Er fragt Sie nach der Wahrscheinlichkeit, dass es bei ihm zu wenige Sitzplätze gibt, wenn er 200 Tickets verkauft. Benutzen Sie den Poissonschen Grenzwertsatz um die gesuchte Wahrscheinlichkeit zu bestimmen.

II.3. (10P) Die Graf Zahl GmbH verlegt Statistiklehrbücher. Für die Produktion des neuesten Titels „Datenfreie Statistik“ hat die Firma die Wahl zwischen den beiden Standorten Leipzig und Dresden. Die jährliche Produktionskapazität hängt jedoch von vielen zufälligen Faktoren ab und kann nicht genau vorhergesagt werden.
Es bezeichne L die Zufallsvariable „produzierte Stückzahl in Leipzig“ und D die Zufallsvariable „produzierte Stückzahl in Dresden“. Die beiden folgenden Tabellen geben die Wahrscheinlichkeiten für die jährliche Produktion an beiden Standorten an.

l	2	3	4	5
$P(L = l)$	0.4	0.3	0.2	0.1

Tabelle 1: Jährliche Produktion in Leipzig (in 1000 Stück)

d	2	3	4	5
$P(D = d)$	0.1	0.4	0.4	0.1

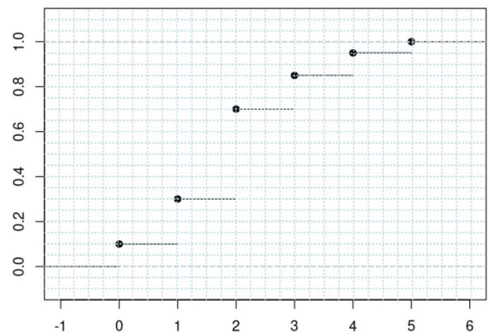
Tabelle 2: Jährliche Produktion in Dresden (in 1000 Stück)

(a) (3P) Bestimmen Sie die Wahrscheinlichkeit, dass
   i. mehr als 3000 Bücher in Leipzig produziert werden,
   ii. mindestens 4000 Bücher in Dresden produziert werden,
   iii. zwischen 3400 und 4500 Bücher in Leipzig produziert werden.

(b) (4P) Bestimmen Sie für beide Standorte die erwartete Anzahl an produzierten Büchern; für welchen Standort entscheiden Sie sich, wenn Sie die erwartete Anzahl an produzierten Büchern maximieren wollen?

(c) (3P) Die Kosten für jedes produzierte Buch sind für die beiden Standorte unterschiedlich. In Leipzig beträgt der Nettogewinn pro Buch 4 Euro, in Dresden ist der Nettogewinn dagegen um 25% niedriger. Für welchen Standort entscheiden Sie sich, wenn Sie den erwarteten jährlichen Gewinn maximieren wollen? (Mit Begründung!)

II.4. (15P) Sie erhalten Daten von 200 Befragten zum täglichen Kaffekonsum x_i (in Anzahl an Tassen) in Form der empirischen Verteilungsfunktion:



(a) (1P) Wie groß ist der Anteil der Befragten, der mehr als eine Tasse am Tag trinkt?

(b) (2P) Wie viele Befragte trinken mehr als eine und weniger als 4 Tassen Kaffe am Tag?

(c) (4P) Zeichnen Sie einen Boxplot und geben Sie die eingezeichneten Werte an.

(d) (3P) Bestimmen Sie den empirischen Mittelwert und die Stichprobenvarianz.
Sie dürfen $\sum_{i=1}^{200}(x_i - \bar{x})^2 = 318$ benutzen.

(e) (3P) Geben Sie das Intervall um den Mittelwert an, welches gemäß der Tschebyscheff Ungleichung mindestens 50% der Daten enthält.

(f) (2P) Sie erfahren, dass auch die Anzahl der verzehrten Kuchenstücke y_i pro Woche erfragt wurde und die Kovarianz zwischen täglichem Kaffekonsum und verzehrten Kuchenstücken und der Mittelwert mit $C_{xy} = 1.34$ und $\bar{y} = 2.25$ berechnet wurde. Bestimmen Sie die Funktionsvorschrift der zugehörigen Regressionsgerade.