Aufgabe 1 (Empirie)

Aufgabe 1 – Teil A (20P)

Runden Sie die Ergebnisse auf 2 Nachkommastellen!

Statistix interessiert sich sehr für die Geschichte europäischer Stadtentwicklung. In einem alten Buch hat er folgende unvollständige Tabelle gefunden. In der Tabelle wurden Informationen zur Einwohneranzahl von 100 Städten eines europäischen Landes im Jahre 1900 gesammelt. Leider sind große Teile altersbedingt nicht mehr lesbar.
Die Angaben sind in Tsd. Einwohnern.

j	Von g_{j-1} bis g_j	h_j	f_j	$F(x)$	m_j	b_j	$m_j h_j$	$\frac{b_j}{f_j}$
1	[10 -)	39				10		
2	[- 50)			0,65				
3	[-)		0,20		75			
4	[-200)	11						909,09
5	[-]					300		
$\sum$		n =						

1) Wie ist das zugrundeliegende statistische Merkmal skaliert? (1P)
2) Vervollständigen Sie die oben gegebene Tabelle auf ihrem Lösungsbogen! (6P)
3) Welche Annahme muss über die Verteilung der Einwohnerzahl innerhalb der Gruppen getroffen werden, wenn die empirische relative Verteilungsfunktion linear interpoliert wird? (1P)
4) Stellen Sie die empirische relative Verteilungsfunktion unter der in 3) getroffenen Annahme grafisch dar (2P)
5) Berechnen Sie folgende Quantile:
   a. $x_{0,25}$ (1P)
   b. $x_{0,5}$ (1P)
   c. $x_{0,75}$ (1P)
6) Zeichnen Sie den zugehörigen Boxplot! (2P)
7) Berechnen Sie das arithmetische Mittel! (1P)
8) Welche Annahme über die Verteilung der Einwohneranzahl innerhalb der Gruppen wird bei 7) getroffen? (1P)

9) Beurteilen Sie die Schiefe der Verteilung mithilfe der Lageregel! (1P)

10) Erklären Sie, wie der empirische Momentenkoeffizient der Schiefe g_1 beurteilen kann, ob die Verteilung symmetrisch bzw. in welchem Maße diese links- oder rechtssteil ist! (2P)

$$\boxed{g_1 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3}$$

Aufgabe 1 – Teil B (5P)

Runden Sie die Ergebnisse auf zwei Nachkommastellen!

Von seiner Kommilitonin Statistika erhält Statistix folgende Informationen über 5 Städte:

Einwohneranzahl in Tsd. Einwohnern	30	50	20	10	60
Anzahl Kirchen im Stadtgebiet	10	12	12	6	15

1) Berechnen Sie einen <u>geeigneten</u> Korrelationskoeffizient, um die Stärke des Zusammenhangs zwischen den beiden Merkmalen zu bemessen! (3P)

2) Interpretieren Sie das Ergebnis inhaltlich und statistisch exakt! (2P)

Aufgabe 2

Aufgabe 2 – Teil A (Kombinatorik) (15P)

Der Lösungsweg muss klar ersichtlich sein!

Statistix bewirbt sich für einen Studienplatz an verschiedenen Universitäten.
Bei jeder Bewerbung bekommt Statistix genau eine der folgenden Antworten:
„Zusage“ (Z), „Absage“ (A) oder „Warteliste“ (W)

1) Listen Sie die Elemente der Ergebnismenge des Zufallsexperiments auf, wenn Statistix sich bei 2 Universitäten bewirbt! (2P)
2) Wie viele verschiedene Antwortlisten kann Statistix bekommen, wenn er sich bei 9 Universitäten bewirbt? (2P)
   Hinweis: Eine Antwortliste ist eine Liste ohne bestimmte Sortierung von allen Antworten der Universitäten, bei denen sich Statistix bewirbt.
3) Was verändert sich bei 2), wenn die Universitäten nicht unabhängig voneinander antworten? Kurze Begründung! (2P)

Statistix hat von der Universität Bielefeld eine Zusage bekommen und möchte sich dort immatrikulieren.

4) Wie viele verschiedene Worte (ohne dass diese einen Sinn ergeben müssen) können aus den Buchstaben „BIELEFELD“ gebildet werden? (2P)

Nachfolgend die Formeln für die Permutation mit Wiederholung sowie für die Kombination ohne Wiederholung:

$$\boxed{P^W(n,p,q,r,\ldots) = \frac{n!}{p!\cdot q!\cdot r!\cdot \ldots}} \qquad \boxed{K(n,r) = \binom{n}{r} = \frac{n!}{r!\cdot(n-r)!}}$$

5) Erklären Sie, wie sich die Formel der Kombination ohne Wiederholung K(n,r) aus der Permutation mit Wiederholung $P^W(n,p,q)$ ergibt! (2P)
   Gehen Sie dabei darauf ein, was sich bei der Permutation „wiederholt“?

Eine Funktion ordnet jedem Element des Definitionsbereichs genau ein Element des Wertebereichs zu.
Die Funktion ist injektiv, wenn jedes Element des Wertebereichs höchstens ein Urbild im Definitionsbereich hat.
Sei X = {1, 2, 3} der Definitionsbereich und Y = {A, B, C, D} der Wertebereich einer Funktion f von X nach Y.

6) Wie viele
   a) Funktionen, (2P)
   b) injektive Funktionen und (2P)
   c) nicht-injektive Funktionen (1P)

   gibt es?

Aufgabe 2 – Teil B (Ereignisalgebra) (10P)

Der Lösungsweg muss klar ersichtlich sein!
Geben Sie die Ergebnisse in Brüchen an!

Betrachten Sie einen unfairen 6-seitigen Würfel mit den Seiten $S_1, S_2, \ldots, S_6$ mit $\mathbb{P}(S_i) = \frac{i}{21}$.
Pro Wurf ergibt sich immer eindeutig, welche der 6 Seiten des Würfels nach oben zeigt, und dies ergibt die „geworfenen Zahl".
Bei wiederholtem Werfen des Würfels sind die Würfe voneinander unabhängig.

1) Wie groß ist die Wahrscheinlichkeit, eine gerade Zahl zu werfen? (2P)
2) Ist es wahrscheinlicher, unwahrscheinlicher oder gleichwahrscheinlich, bei zwei Würfen zunächst eine 4 und dann eine 6 zu würfeln oder zweimal hintereinander eine 5? (2P)
3) Angenommen, die Augensumme nach zwei Würfen ist gleich 6. Wie groß ist Wahrscheinlichkeit, dass nur gerade Zahlen geworfen wurden? (4P)

Sei $A \subseteq B \subseteq \Omega$.

4) Bestimmen Sie $\mathbb{P}(B\,|\,A)$! (1P)
5) Welche Voraussetzung muss für 4) erfüllt sein? (1P)

Aufgabe 3 (Verteilungsmodelle)

Runden Sie die Ergebnisse auf 4 Nachkommastellen!

Aufgabe 3 – Teil A (7P)

Statistix hilft an 7 Tagen auf einem Festival aus.
In dieser Zeit kommt er immer sehr spät ins Bett, so dass es passieren kann, dass er am nächsten Tag verschläft und zu spät zur Arbeit kommt.
Die Wahrscheinlichkeit, an einem Tag nicht zu verschlafen, beträgt 95 %.
Ob er verschläft oder nicht ist von Tag zu Tag unabhängig voneinander.

1) Wie ist X: „Anzahl der Tage, an denen Statistix während des Festivals verschläft“ verteilt? [Verteilungstyp und -parameter] (1P)
2) Wie groß ist die Wahrscheinlichkeit, dass Statistix während des Festivals
   a. kein einziges Mal (1P)
   b. mehr als 3 Mal (2P)
   c. mindestens 3 Mal, aber weniger als 6 Mal (2P)

   zu verschlafen?
3) Wie oft wird Statistix während des Festivals durchschnittlich verschlafen? (1P)

Aufgabe 3 – Teil B (8P)

Statistix soll am ersten Tag dabei helfen, die Lautsprecher auf den Bühnen zu überwachen.
Aus Erfahrung weiß er, dass durchschnittlich alle 1000 s ein Lautsprecher ausfällt.

1) Wie ist Y: „Wartezeit bis zum nächsten Ausfall eines Lautsprechers in 1000s“ verteilt? [Verteilungstyp und -parameter] (2P)
2) Wie groß ist die Wahrscheinlichkeit, dass
   a. nach höchstens 250 s der nächste Lautsprecher ausfällt? (1P)
   b. nach höchstens 800 s der nächste Lautsprecher ausfällt, wenn in den ersten 400 s kein Lautsprecher ausgefallen ist? (2P)
   c. nach mehr als 400 s, aber höchstens nach 800 s der nächste Lautsprecher ausfällt (2P)
3) Welche Eigenschaft der Verteilung von Y haben Sie bei 2) b) genutzt (bzw. nutzen können)? (1P)

Aufgabe 3 – Teil C (4P)

Die Bühne, an der Statistix zuletzt gearbeitet hat, wird über einen Generator mit Strom versorgt. In regelmäßigen Abständen von 750 Minuten muss der Generator mit neuem Treibstoff befüllt werden. Leider weiß Statistix nicht, wann der Generator das letzte Mal befüllt worden ist, und kann auch den Füllstand nicht einsehen.

1) Wie ist Z: „Dauer bis zur nächsten Treibstoffbefüllung des Generators“ verteilt? (1P)
   [Verteilungstyp und -parameter]
2) Wie groß ist die Wahrscheinlichkeit, dass der Generator
   a. in den nächsten 20 Minuten neu befüllt werden muss? (1P)
   b. in den nächsten 100 Minuten neu befüllt werden muss, nachdem dieser bereits 100 Minuten gelaufen ist (und in dieser Zeit nicht mit Treibstoff befüllt werden musste)? (2P)

Aufgabe 3 – Teil D (6P)

Statistix nimmt an, dass seine Ausgaben auf einem solchen Festival normalverteilt sind mit einem Erwartungswert von $\mu = 133$ € und einer Varianz von $\sigma^2 = 25$ €².

1) Wie groß ist die Wahrscheinlichkeit, dass Statistix mehr als 129 €, aber weniger als 131 € ausgeben wird? (2P)
2) Welchen Betrag höchstens wird er mit einer Wahrscheinlichkeit von 15,87 % ausgeben? (2P)
3) Wie ist die Summe seiner Ausgaben für das Festival in diesem und im nächsten Jahr verteilt, wenn die Ausgaben je Festival unabhängig und identisch verteilt sind? (2P)
   [Verteilungstyp und -parameter]