Übungsklausur zu „Statistik I für Statistiker, Mathematiker und Informatiker“

Aufgabe 1

Die folgende Tabelle zeigt für die 16 deutschen Bundesländer die Anzahl der Geburten je 1000 Einwohner (G) im Jahr 2002. (Quelle: DER SPIEGEL 44/2003).

BB	ST	MV	TH	SN	SL	HB	BE	RP	NW	HE	HH	SH	BY	BW	NI
6.8	6.9	7.1	7.1	7.2	7.4	8.3	8.5	8.6	9.0	9.1	9.1	9.2	9.2	9.4	9.5

Aus den angegebenen Daten lässt sich $\sum_{i=1}^{16} G_i = 132.40$ ermitteln.

(a) Entscheiden Sie für das betrachtete Merkmal G, ob es stetiger oder diskreter Natur ist und welches Skalenniveau es besitzt.

(b) Ermitteln Sie arithmetisches Mittel, Modus und Median von G. Was lässt sich anhand dieser Parameter über die Gestalt der Verteilung der Anzahl von Geburten je 1000 Einwohner in den 16 Bundesländern sagen (symmetrisch, rechts-, linkssteil)?

(c) Laut dem original Zeitungsartikel wurden im Jahr 2002 in Deutschland durchschnittlich 8.7 Geburten je 1000 Einwohner gezählt. Erläutern Sie kurz, warum das in (b) berechnete arithmetische Mittel von diesem Wert abweicht. Wie müsste man in (b) vorgehen, um eine Übereinstimmung zwischen beiden Werten zu erhalten?

(d) Erstellen Sie für die Daten ein Histogramm der relativen Häufigkeiten mit den Klassen $(6, 7], (7, 8], (8, 9], (9, 10]$. Begründen Sie dabei kurz allgemein, in welchem Spezialfall ein Histogramm mit dem einfachen Säulendiagramm identisch ist.

Aufgabe 2

Für das Alter (X) und den Händlerverkaufspreis (Y) gebrauchter PKW eines bestimmten Typs liegen folgende Informationen vor: Die Kovarianz zwischen Alter und Verkaufspreis beträgt -5.4; die Varianz des Verkaufspreises ist 4. Durch eine lineare Abhängigkeit vom Alter werden 81% der Variation in den Verkaufspreisen erklärt.
Berechnen Sie die Standardabweichung des Alters.

Aufgabe 3

Ein vermisstes Flugzeug kann mit gleichen Wahrscheinlichkeiten in jeder von drei Regionen abgestürzt sein. Sei $1 - \beta_i$ die Wahrscheinlichkeit, dass das Flugzeug in Region i, $i = 1, 2, 3$ gefunden wird, falls es dort abgestürzt ist, wobei β_i die Wahrscheinlichkeit bezeichnet, mit der das Flugzeug übersehen wird, $i = 1, 2, 3$.

(a) Wie groß ist die Wahrscheinlichkeit, dass sich das Flugzeug in Region i, $i = 1, 2, 3$ befindet, wenn die Suche in Region 1 erfolglos war?

(b) Stellt das Ereignis E : „Die Suche in Region 1 war erfolglos“ einen Informationsgewinn für die weitere Suche dar?

Aufgabe 4

Die Telefonzentrale einer Feuerwache empfängt in einer Stunde durchschnittlich 0.5 Alarmmeldungen.

(a) Wie groß ist die Wahrscheinlichkeit, daß während der sechsstündigen Dienstzeit einer Feuerwehrmannschaft

- kein Alarm,
- mindestens dreimal Alarm,
- höchstens siebenmal Alarm gegeben wird?

(b) Wie groß ist die Wahrscheinlichkeit, daß die Feuerwehrmannschaft

- innerhalb der ersten Dienststunde den ersten Alarm bekommt?
- länger als zwei Stunden auf den ersten Alarm warten muß?
- ausgerechnet in der letzten Dienststunde zum ersten Alarm „ausrücken" muß, nachdem in der gesamten fünfstündigen Dienstzeit zuvor kein Alarm gekommen ist?

(c) Der Oberbrandmeister erklärt seiner Feuerwehrmannschaft, daß mit 95%iger Wahrscheinlichkeit der erste Alarm noch in der Dienstzeit dieser Mannschaft fallen wird. Hat er recht? Muß die Mannschaft also tatsächlich weniger als 6 Stunden warten, um mit 95%iger Wahrscheinlichkeit den ersten Alarm zu bekommen?

Aufgabe 5

Sei X eine stetige Zufallsgröße, deren Dichte

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{sonst} \end{cases}$$

von einem unbekannten Parameter λ abhängt.

(a) Zeigen Sie, dass $f(x)$ für alle $\lambda > 0$ tatsächlich eine Dichtefunktion ist.

(b) Verifizieren Sie, dass der Erwartungswert von X durch $E(X) = \frac{1}{\lambda}$ bestimmt ist.
(*Hinweis:* Es gilt

$$\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx, \quad a, b \in \mathbb{R} \cup \{-\infty, \infty\},$$

wobei $u(x)$ und $v(x)$ stetige Ableitungen besitzen müssen.)

Aufgabe 6

Es werde zufällig eine Zahl X aus $\{1, \ldots, 5\}$ und dann zufällig eine weitere Zahl Y aus $\{1, \ldots, X\}$ ausgewählt.

(a) Berechnen Sie die gemeinsame Wahrscheinlichkeitsfunktion $P(X = x, Y = y)$.

(b) Berechnen Sie die bedingte Wahrscheinlichkeitsfunktion von X gegeben $Y = j$, $(j = 1, \ldots, 5)$

(c) Sind X und Y unabhängig?

Zusatz-Aufgabe 7

Der χ^2-Koeffizient für n Beobachtungen zweier nominalskalierter Merkmale $X \in \{a_1, \ldots, a_k\}$ und $Y \in \{b_1, \ldots, b_\ell\}$ ist definiert durch

$$\chi^2 := \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(h_{ij} - e_{ij})^2}{e_{ij}} = n \left(\sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{h_{ij}^2}{h_{i\bullet} h_{\bullet j}} - 1 \right)$$

mit

$$e_{ij} := \frac{h_{i\bullet} h_{\bullet j}}{n}.$$

Zeigen Sie, dass $\chi^2 \in [0, n(\min\{k, \ell\} - 1)]$ gilt.