Musterlösung zur Testklausur Statistik

WMS16A (Oettinger, 2017)

Aufgabe 1

(a) Richtig: das arithmetische Mittel reagiert im Gegensatz zum Median empfindlich auf Ausreißer.

(b) Falsch: ein ordinales Merkmal besitzt eine natürliche Rangfolge, ein nominales Merkmal keine.

(c) Richtig: es handelt sich um eine Summe quadrierter Größen.

(d) Falsch: der Modus ist die am häufigsten auftretende Merkmalsausprägung in einer Stichprobe.

(e) Falsch: es handelt sich um ein mittleres Wachstum, es kann auch negativ sein.

Aufgabe 2

Die Tabelle in Form relativer Häufigkeiten:

X (PKW) / Verbrauch Y	1	2	3	4	$\sum$
6,1	0,182	0,2	0,182	0,194	0,185
6,2	0,545	0,6	0,545	0,516	0,546
6,3	0,273	0,2	0,273	0,290	0,269
$\sum$	1	1	1	1	1

(a) Die Spalten der Tabelle relativer Häufigkeiten unterscheiden sich $\Longrightarrow$ die Größen X und Y sind nicht stochastisch unabhängig.

(b) Die bedingte Verteilung $f(x_i|Y = 6,2)$ besteht aus den relativen Häufigkeiten der Zeile mit $Y = 6,2$, also

$$f(x_i|Y = 6,2) = \left\{ \frac{6}{124}; \frac{12}{124}; \frac{90}{124}; \frac{9}{124} \right\}.$$

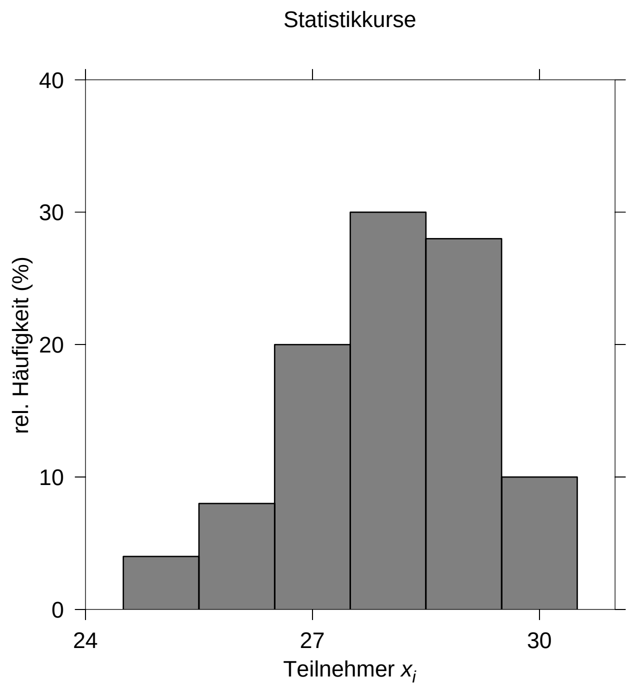
(c) Benötigt wird der Mittelwert des Verbrauchs für PKW 2 (alle Angaben in l):

$$\bar{y}(X=2) = \frac{1}{20}\,(4 \cdot 6,1 + 12 \cdot 6,2 + 4 \cdot 6,3) = 6,2.$$

Die Varianz ist die mittlere quadratische Abweichung vom Mittelwert, also

$$\begin{aligned} s^2(Y|X=2) &= \frac{1}{20}\left(4 \cdot (6,1-6,2)^2 + 12 \cdot (6,2-6,2) + 4 \cdot (6,3-6,2)\right) \\ &= \frac{1}{20} 8 \cdot 0,1^2 = 4 \cdot 10^{-3}. \end{aligned}$$

Aufgabe 3



a)

Abbildung 1: Häufigkeitsverteilung der Teilnehmer

b) arithmetisches Mittel:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} h_i \cdot x_i = \frac{1}{50}(2 \cdot 25 + 4 \cdot 26 + 10 \cdot 27 + 15 \cdot 28 + 14 \cdot 29 + 5 \cdot 30) = 28$$

c) Varianz s^2:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} h_i (x_i - \bar{x})^2 = \frac{1}{5}(2 \cdot (25-28)^2 + 4 \cdot (26-28)^2 + 10 \cdot (27-28)^2 + 15 \cdot (28-28)^2 + 14 \cdot (29-28)^2 +$$

Die Standardabweichung ist die positive Quadratwurzel der Varianz

$$s = +\sqrt{s^2} = +\sqrt{1,56} = 1,249$$

d) Der Modus ist die Merkalsausprägung, die innerhalb der Stichprobe am häufigsten auftritt, also ebenfalls 28. Der Median ist der Wert (x_i aus dem geordneten Vektor der Merkmalsausprägungen)

$$\bar{x}_Z = \frac{x_{25} + x_{26}}{2} = \frac{28 + 28}{2} = 28$$

e) Die Verteilung ist linksschief bzw. rechtssteil.

Aufgabe 4

Es handelt sich beim idealen Wüfel um ein Laplace-Experiment mit einer Wahrscheinlichkeit von $1/6$ für das Auftreten einer bestimmten Augenzahl.

(a) Wahrscheinlichkeit, zweimal hintereinander die zwei zu würfeln: $P = 1/6 \cdot 1/6 = 1/36$.

(b) Für den ersten Würfel sind alle Augenzahlen zulässig, für den zweiten nur die vom ersten Würfel vorgegebene: $P = 1 \cdot 1/6 = 1/6$.

(c) Die Wahrscheinlichkeit dafür, dass die Würfel eine Augensumme zeigen ist 1. Die Wahrscheinlichkeit für unterschiedliche Augenzahlen ist die Wahrscheinlichkeit einer Augensumme minus der Wahrscheinlichkeit dafür, dass beide Würfel dieselbe Augenzahl zeigen, also $P = 1-1/6 = 5/6$.

Aufgabe 5

Die benötigten Daten sind in der folgenden Tabelle zusammengefasst.

Alter (Jahre) ...	Absolute Häufigkeit	$\bar{y}_i$	$s^2_{Y,i}$	Klassen-breite Δ_i	h^*_i	rel. f_i	kumul. F_i	rel. q_i	kumul. Q_i
bis 30	10	2,5	1,8	30	0,37	0,0775	0,0775	0,04	0,04
30 - 40	47	4,5	2,9	10	4,7	0,364	0,4415	0,337	0,377
40 - 50	42	5,3	3,4	10	4,5	0,326	0,768	0,355	0,732
50 -65	30	5,6	3,6	15	0,2	0,233	1	0,268	1

Der Anteil der einzelnen Altersgruppen q_i am Gesamteinkommen kann aus dem mittleren Einkommen $\bar{y}_i$ berechnet werden, indem aus den absoluten Häufigkeiten und dem mittleren Einkommen die Summe des Einkommens in der Altersgruppe berechnet und durch die ebenso bestimmte Gesamtsumme geteilt wird.

(a) Berechnung des arithmetischen Mittels in Jahren: Histogramm:

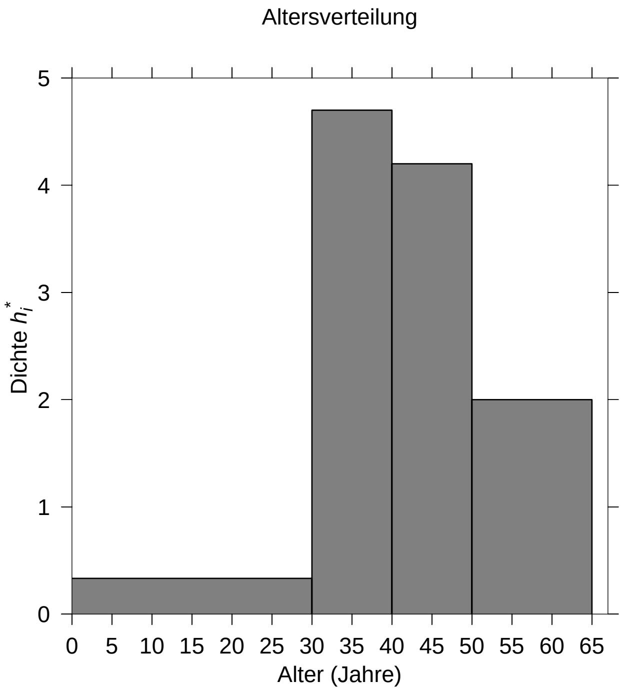


Abbildung 2: Histogramm der Altersverteilung

$$\bar{x} = \frac{1}{129}(15 \cdot 10 + 35 \cdot 47 + 45 \cdot 42 + 57,5 \cdot 30) = 41,94.$$

(b) Berechnung des Medians in Jahren unter Annahme von Gleichverteilung innerhalb der Klassen: der Wert von 0,5 wird in der dritten Klasse ($40 - 50$ Jahre) erreicht.

$$\begin{aligned}\bar{x}_Z &= x_3^u + (x_3^o - x_3^u) \cdot \frac{F(\bar{x}_Z) - F(x_3^u)}{F(x_3^o) - F(x_3^u)} \\ &= 40 + (50 - 40) \cdot \frac{0,5 - 0,4415}{0,768 - 0,4415} = 41,79\end{aligned}$$

(c) Lorenz-Plot:

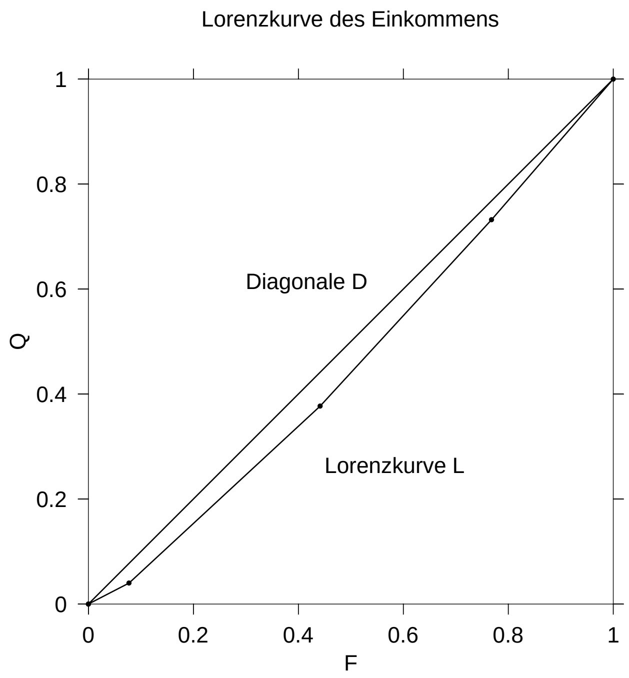


Abbildung 3: Lorenzkurve der Einkommensverteilung

Aufgabe 6

(a) die mittlere erreichte Geschwindigkeit ergibt sich ganz einfach, indem die gesamte zurückgelegte Strecke durch die gesamte benötigte Zeit geteilt wird (Geschwindigkeit in km/h):

$$\bar{v} = \frac{1 \cdot 7 + 5}{1 + 1,25} = 5,\bar{3}$$

(b) Der Mittelwert berechnet sich nach

$$\bar{x} = \frac{1}{20}(5 \cdot 1 + 11 \cdot 2 + 3 \cdot 3 + 4) = 2,0$$

Für die Kandidaten, die nicht bestanden haben, kann keine Note angegeben werden, sie werden zur Berechnung *nicht* herangezogen.

(c) Geometrisches Mittel:

$$\bar{x}_G = \sqrt{1,56 \cdot 1,04} - 1 = 0,274 \text{ oder } 27,4\%$$

(d) Wiederum geometrisches Mittel:

$$\bar{x}_G = \sqrt{2 \cdot 2} - 1 = 2 - 1 = 1 \text{ oder } 100\%,$$

genau das ist ja die Aussage (verdoppelt sich jede Nacht!).

(e)

$$\bar{x}_G = \sqrt{2 \cdot 8} - 1 = 4 - 1 = 3 \text{ oder } 300\%.$$

Aufgabe 7

1. Studierende der DHBW in Ravensburg: Bestandsmasse.
2. Todesfälle in einer Gemeinde: Bewegungsmasse.
3. Maschinenausfälle in einer Werkstatt: Bewegungsmasse.
4. Die Bevölkerung in Ravensburg zum 1.9.2010: Bestandsmasse.