Musterlösung zur zweiten Nachklausur Statistik

TIT/TIM/TIS 19 / Oettinger / 12.2021

Insgesamt erreichbare Punktzahl: 100, 100%: 100 Punkte.

Aufgabe 1

(15 Punkte)

(a) Das arithmetische Mittel wird durch die Ausreißer mit kleinem Betrag geringer, also nach unten verschoben - falsch.

(b) Die Standardabweichung ist als positiver Ast der Wurzel der Varianz definiert (die immer positiv oder Null ist) - richtig.

(c) Die Varianz ist eine Summe quadrierter Größen - sie ist immer größer oder gleich Null, sie ist nur dann Null, wenn die Stichprobe aus identischen Werten besteht - falsch.

(d) Der Modus ist der am häufigsten auftretende Wert, das 100%-Quantil entspricht dem Maximum der Werte - falsch.

(e) Das geometrische Mittel kann bei negativer Veränderungen auch negative Werte annehmen - falsch.

Aufgabe 2

(24 Punkte)
Die arithmetischen Mittel sind $\bar{x} = 5000$ und $\bar{y} = 60000$ EUR. Die benötigten Daten in Tabellenform:

Monat i	1	2	3	4	5	6	(Summe)
Menge x_i	2.000	3.000	6.000	4.000	8.000	7.000	30000
Kosten y_i	30.000	35.000	75.000	55.000	85.000	80.000	360000
$(x_i - \bar{x})$	-300	-200	100	-100	300	200	0
$(y_i - \bar{y})$	-3000	-2500	1500	-500	2500	2000	0
$(x_i - \bar{x})(y_i - \bar{y})$	900000	500000	150000	50000	750000	400000	2750000
$(x_i - \bar{x})^2$	90000	40000	10000	10000	40000	90000	280000
$(y_i - \bar{y})^2$	9000000	6250000	2250000	250000	6250000	4000000	28000000

a) Die Ausgleichsgerade lautet $y = a \cdot x + b$, die Werte für die Steigung und den Achsenabschnitt sind

$$a = \frac{S_{xy}}{S_{xx}} = \frac{sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{2750000}{28000000} = 9,82$$

Achsenabschnitt $b = \bar{y} - m \cdot \bar{x} = 1089,3$

b) Die Gerade beschreibt den Zusammenhang zwischen der produzierten Menge und den Gesamtkosten des Produkts. Die Steigung gibt die Kosten je produzierter Einheit an, der Achsenabschnitt beschreibt Fixkosten, die von der hergestellten Menge unabhängig sind.

Aufgabe 3

(10 Punkte)

Tag	1	2	3	4	5	6	7
km/h (x)	15	16,5	17,5	18	18	20	22

a) Arithmetisches Mittel (in km/h):

$$\bar{x} = \frac{15 + 16,5 + 17,5 + 18 + 18 + 20 + 22}{7} = \frac{127}{7} = 18,1429$$

b) Harmonisches Mittel (in km/h):

$$\bar{x}_H = \frac{7}{\frac{1}{15} + \frac{1}{16,5} + \frac{1}{17,5} + \frac{1}{18} + \frac{1}{18} + \frac{1}{20} + \frac{1}{22}} = 17,9037$$

c) Die Durchschnittsgeschwindigkeit erhält man als Quotient aus der gesamten zurückgelegten Strecke und der dafür benötigten Zeit, also

$$\frac{7 \cdot 10\text{km}}{\frac{10\text{km}}{15\text{km/h}} + \frac{10\text{km}}{16,5\text{km/h}} + \frac{10\text{km}}{17,5\text{km/h}} + \frac{10\text{km}}{18\text{km/h}} + \frac{10\text{km}}{18\text{km/h}} + \frac{10\text{km}}{20\text{km/h}} + \frac{10\text{km}}{22\text{km/h}}}$$

Die Anwendung des harmonischen Mittels ist hier korrekt.

Aufgabe 4

(7 Punkte)
6-stellige ID-Nummern für den neu gegründeten Paketdienst, Berechnung unter der vereinfachenden Annahme, dass auch 000000 als ID-Nummer zulässig sein soll:

(a) Jede beliebige Farbe.

(b) Anzahl A aller möglichen Kombinationen für 6-stellige Nummern:
Jede Stelle kann mit $0 \dots 9$ besetzt werden. $A = 10^6$ (unter der Annahme, dass auch 000000 als ID-Nummer zählt).

(c) Soll keine Nummer mit einer Null beginnen, gibt es für die erste Stelle nur 9 Möglichkeiten ($1 \dots 9$). $A = 9 \cdot 10^5$.

(d) Keine Ziffer soll zweimal vorkommen (Berechnung unter der Annahme, dass eine führende Null erlaubt ist): für die erste Stelle gibt es 10 Möglichkeiten,
für die zweite 9,
für die dritte 8 usw.
$A = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 151200$

Aufgabe 5

(30 Punkte)

(a) Statistische Einheit sind Giovannis Kunden, Merkmale sind Speisen und Getränke. Die Merkmalsausprägungen sind die vom Kunden bestellten Gerichte und Getränke, beispielsweise 'Rotwein' oder 'Pizza'.

(b) Darstellung der Daten im Stab- oder Balkendiagramm. Grafik nicht gefordert!

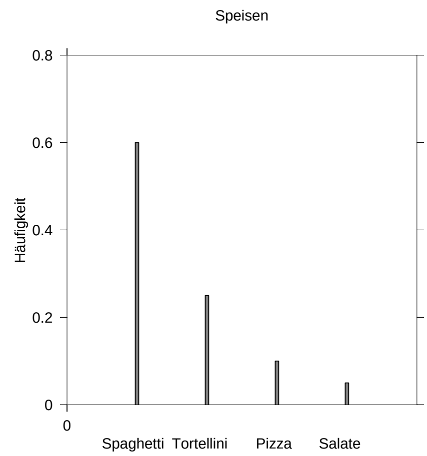

Abbildung 1: Stabdiagramm als Beispiel für die grafische Darstellung

(c) Ereignis A: Kunde bestellt Spaghetti, $P(A) = 0,6$.
Ereignis B: Kunde bestellt Rotwein, $P(B) = 0,7$.
Ereignis C: Kunde bestellt Spaghetti und Rotwein, $P(C) = P(A) \cdot P(B) = 0,6 \cdot 0,7 = 0,42 = 42\%$.
Ereignis D: Kunde bestellt Spaghetti, aber keinen Rotwein, $P(D) = P(A) \cdot (1 - P(B)) = 0,6 \cdot (1 - 0,7) = 0,6 \cdot 0,3 = 0,18 = 18\%$.

(d) Die relative Häufigkeit für eine Spaghettilänge größer als 80cm ist $0,15 + 0,05 = 0,2 = 20\%$.
Das arithmetische Mittel (in cm) muss mit den Klassenmitten gerechnet werden (wegen der Angabe relativer Häufigkeiten wird nicht durch die Gesamtzahl geteilt):

$$\bar{x} = 0,15 \cdot 30 + 25 \cdot 50 + 40 \cdot 70 + 15 \cdot 90 + 5 \cdot 115 = 64,25$$

(e) Der Median (in cm) entfällt auf die 3.Klasse $]60;80]$:

$$\bar{x}_Z = x_3^u + (x_3^o - x_3^u) \cdot \frac{F(\bar{x}_Z) - F(x_3^u)}{F(x_3^o) - F(x_3^u)}$$

$$= 60 + (80 - 60) \cdot \frac{0,5 - 0,4}{0,8 - 0,4} = 65.$$

(f) Da die Klassen unterschiedlich breit sind, wird die Dichte der relativen Häufigkeiten aufgetragen. Daten zur grafischen Darstellung:

Länge (cm)	Klasenbreite (in m) Δ_k	rel. Häufigkeit f_k	Dichte $f_k^* = \frac{f_k}{\Delta_k}$
$(20;40]$	$0,20$	0,15	0,75
$(40;60]$	$0,20$	0,25	1,25
$(60;80]$	$0,20$	0,4	2,0
$(80;100]$	$0,20$	0,15	0,75
$(100;130]$	$0,30$	0,05	0,167

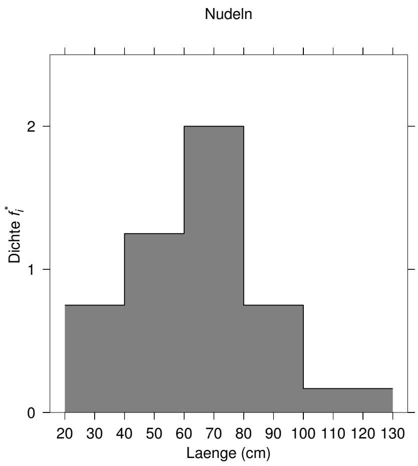

Abbildung 2: Histogramm der Verteilung der Nudellänge

Aufgabe 6

(14 Punkte)
Daten zum Reis:

Jahr	2014	2015	2016	2017	2018	2019	2020
x_i	13	7	12	8	11	5	7

Tabelle 1: Daten zu Reisunfällen in China

Das arithmetische Mittel ist

$$\bar{x} = \frac{1}{7}(13 + 7 + 12 + 8 + 11 + 5 + 7) = 9,$$

der Median ist der zentrale Wert in der geordneten Stichprobe

$$\bar{x}_Z = 8.$$

Die Varianz ist die mittlere qadratische Abweichung vom arithmetischen Mittel

$$s^2 = \frac{1}{7}((5-9)^2 + 2 \cdot (7-9)^2 + (8-9)^2 + (11-9)^2 + (12-9)^2 + ({}^3-9)^2) = 7,71$$

und damit die Standardabweichung

$$s = \sqrt{s^2} = 2,78$$