Musterlösung zur Klausur Statistik

TI21

Oettinger 03.2023

Zeit: 90Min.

Insgesamt erreichbare Punktzahl: 106, 100%: 100 Punkte.

Aufgabe 1

(15 Punkte)

(a) Nominale Merkmale besitzen keine natürliche Rangfolge, eine Symmetrie kann nicht definiert werden - falsch.

(b) Der Median ist der Wert in der Mitte der Stichprobe, er entspricht dem 50%-Quantil - falsch.

(c) Der Median kann je nach Symmetrie einer Häufigkeitsverteilung im Vergleich zum arithmetischen Mittel größer, gleich oder kleiner sein - falsch.

(d) Die Varianz kann nur positive Werte annehmen - richtig, sie ist eine Summe quadrierter Größen.

(e) Das arithmetische Mittel kann natürlich größer als der Median sein - die Stichprobe ist dann meist linkssteil - richtig.

Aufgabe 2

(14 Punkte) Daten zur Gewichtsverteilung mit Häufigkeitsdichte:

i	Klasse $(x_i^u; x_i^o]$	h_i	Mitte	Breite	Dichte h_i^*	rel. Häufigkeit f_i	kumuliert F_i
1	(985; 995]	15	990	10	1,5	0,3	0,3
2	(995; 1000]	5	997,5	5	1	0,1	0,4
3	(1000; 1005]	20	1002,5	5	4	0,4	0,8
4	(1005; 1020]	10	1012,5	15	0,666	0,2	1

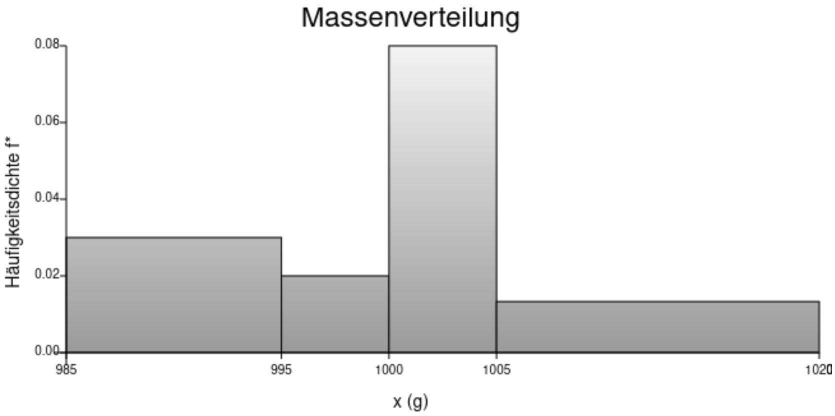

Abbildung 1: Histogramm zum Gewicht des Quallengelees

Die Näherung für das arithmetische Mittel (in Gramm) über die Klassenmitten ist

$$\bar{x} \approx \frac{1}{50}(15 \cdot 990 + 5 \cdot 997,5 + 20 \cdot 1002,5 + 10 \cdot 1012,5) = 1000,25.$$

Die Hälfte der untersuchten Pakete (25) wird in der dritten Klasse erreicht ($i = 3$), der Median kann ebenfalls geschätzt werden (in Gramm)

$$\bar{x_Z} \approx x_3^u + (x_3^o - x_3^u)\frac{0,5 - F(x_3^u)}{F(x_3^o) - F(x_3^u)} = 1000 + 5 \cdot \frac{0,5 - 0,4}{0,8 - 0,4} = 1001,25$$

Als typischer Wert sind beide geeignet - in den meisten Fällen gibt der Median die Verhältnisse der Verteilung aber etwas besser wieder.

Aufgabe 3

(27 Punkte)
Die arithmetischen Mittel sind $\bar{x} = 5000$ und $\bar{y} = 60000$, die benötigten Daten sind

Monat i	1	2	3	4	5	6	(Summe)
Menge x_i	2.000	3.000	6.000	4.000	8.000	7.000	
Kosten y_i (EUR)	30.000	35.000	75.000	55.000	85.000	80.000	
$(x_i - \bar{x})$	-300	-200	100	-100	300	200	
$(y_i - \bar{y})$	-3000	-2500	1500	-500	2500	2000	
$(x_i - \bar{x})(y_i - \bar{y})$	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 $y = m \cdot x + b$, die Werte für die Steigung und den Achsenabschnitt sind

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

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.

c) Der Bravais-Pearson-Koeffizient ist

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{2750000}{\sqrt{280000}\sqrt{28000000}} = 0,98$$

Aufgabe 4

(27 Punkte)
Die benötigten Daten zur Aufgabe:

Zahl der Tabellen x_i	Tage früher	Tage jetzt	kumuliert früher	kumuliert jetzt
1	60	5	60	5
2	160	10	220	15
3	110	25	330	40
4	0	20	330	60
5	60	0	390	60
6	50	0	440	60
8	0	40	440	100

a) Die Aussage ist: Lohnt sich der Droide (ist die Produktivität angestiegen?).

b) Benötigt werden das Arithmetische Mittel und der Median:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{m} h_i x_i$$

$$\bar{x} = \frac{1}{440}\left(60 + 2 \cdot 160 + 3 \cdot 110 + 4 \cdot 0 + 5 \cdot 60 + 6 \cdot 50 + 8 \cdot 0\right) = 2,977 \text{ ohne Droide,}$$

$$\bar{x} = \frac{1}{100}\left(5 + 2 \cdot 10 + 3 \cdot 25 + 4 \cdot 20 + 5 \cdot 0 + 6 \cdot 0 + 8 \cdot 40\right) = 5 \text{ mit Droide.}$$

Der Median lässt sich aus den Daten in der Tabelle ablesen, für die Daten ohne den Droiden

$$\bar{x}_Z = \frac{x_{220} + x_{221}}{2} = 2,5,$$

nach der Einführung des Droiden

$$\bar{x}_Z = \frac{x_{50} + x_{51}}{2} = 4.$$

Die mittlere absolute Abweichung vom arithmetischen Mittel ist

$$\begin{aligned} d_{\bar{x}} &= \frac{1}{n}\sum_{i=1}^{m} h_i (x_i - \bar{x}), \\ &= 1,25 \text{ ohne Droide bzw.} \\ &= 2,4 \text{ mit Droide.} \end{aligned}$$

Die mittlere absolute Abweichung vom Median ist

$$\begin{aligned} d_{\bar{x}_Z} &= \frac{1}{n}\sum_{i=1}^{m} h_i (x_i - \bar{x}_Z), \\ &= 1,25 \text{ ohne Droide bzw.} \\ &= 2,2 \text{ mit Droide.} \end{aligned}$$

Aufgabe 5

(14 Punkte)
Geeignete Mittelwerte.

1. Wenn Ted eine mittlere Geschwindigkeit von 60 km/h fahren will, benötigt er für die insgesamt 8 km Weg eine Zeit von 8/60 h = 8 min. Da er aber für den Rückweg von 4km bereits eine Zeit von 4/30 h = 8min einplant, kann er die geplante Durchschnittsgeschwindigkeit nicht erreichen.

2. Geometrisches Mittel: das mittlere Wachstum ist

$$\bar{x}_G = \sqrt[3]{(1+0,1)\cdot(1+0,15)\cdot(1-0,0005)} - 1 = 8,13\%.$$

   Ist die Größe der Kultur zu Beginn N_0, beträgt sie nach 5 Tagen

$$N_5 = (1+\bar{q})^5 N_0 = (1,0813)^5 N_0 = 1,48 N_0,$$

   das entspricht einem Wachstum von von 48%.

3. Insgesamt befragte Personen: 100 + 1000 = 1100. Für die Abschaffung sind 60 + 380 = 440. Also sind 440/1100 = 40% dafür.

Aufgabe 6

(8 Punkte)
Nominale/ordinale/kardinale Merkmale:

(a) Körpergröße: kardinal (Zahl)

(b) Farbe: nominal (keine Rangfolge)

(c) Krawattenlänge: kardinal (Zahl)

(d) Qualität von Vorlesungen: ordinal (keine Zahl, aber mit Rangfolge)