

Musterlösung zur Klausur Statistik

WMS13A

Oettinger 06.2014

Zeit: 60Min. / 100 Punkte

Aufgabe 1

- (a) Jede Antwort ist zulässig
- (b) Statistische Einheit sind Giovanni's Kunden, Merkmale sind Speisen und Getränke. Die Merkmalsausprägungen sind die vom Kunden bestellten Gerichte und Getränke, beispielsweise 'Rotwein' oder 'Pizza'.
- (c) Darstellung der Daten im Stab- oder Balkendiagramm. Grafik nicht gefordert!

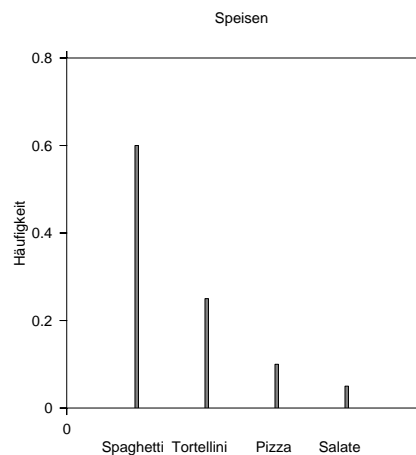


Abbildung 1: Stabdiagramm als Beispiel für die grafische Darstellung

- (d) 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, $P(D) = P(A) \cdot (1 - P(B)) = 0,6 \cdot (1 - 0,7) = 0,6 \cdot 0,3 = 0,18 = 18\%$.

- (e) 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):

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

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

$$\begin{aligned} \bar{x}_Z &= x_3^u + (x_3^o - x_3^u) \cdot \frac{F(\bar{x}_Z) - F(x_3^u)}{F(x_3^o) - F(x_3^u)} \\ &= 60 + (80 - 60) \cdot \frac{0,5 - 0,4}{0,8 - 0,4} = 65. \end{aligned}$$

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

| Länge (cm) | Klassenbreite (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 |

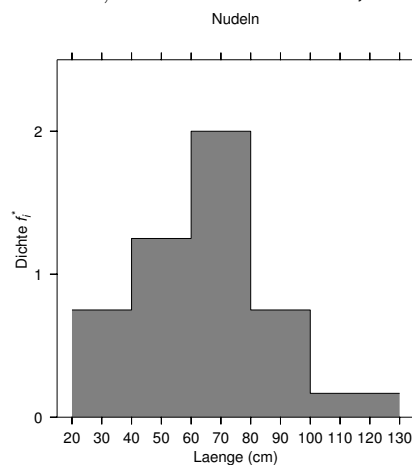


Abbildung 2: Histogramm der Verteilung der Nudellänge

Aufgabe 2

- a) Geometrisches Mittel:

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

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

Aufgabe 3

- a) Kreuztabelle, bei der das Merkmal 'Passagierklasse' die Spaltenvariable darstellt:

| | 1.Klasse | 2. Klasse | 3.Klasse | Randspalte |
|------------|----------|-----------|----------|------------|
| gerettet n | 202 | 125 | 180 | 507 |
| % | 60 | 43,9 | 25 | 37,8 |
| vermisst n | 135 | 160 | 541 | 836 |
| % | 40 | 56,1 | 75 | 62,2 |
| Summe n | 337 | 285 | 721 | 1343 |
| % | 100 | 100 | 100 | 100 |

- b) Die Spalten der relativen Häufigkeiten unterscheiden sich deutlich, was auf einen Zusammenhang zwischen Passagierklasse und Überlebenschance hinweist. Die Überlebenschance in einer besseren Passagierklasse scheint höher zu sein.

Aufgabe 4

- a) die Körpergröße von Studentinnen ist ein stetiges kardinales Merkmal
- b) Die Farbe der Krawatten von Dozenten ist ein nominales Merkmal
- c) Die Länge dieser Krawatten ist ein stetiges und kardinales Merkmal
- d) Die Qualität von Statistikvorlesungen mit den angegebenen Ausprägungen ist ein ordinales Merkmal (Rangfolge).

Aufgabe 5

Welche der folgenden Aussagen sind richtig?

- (a) Hat das arithmetische Mittel einen größeren Wert als der Median, so deutet dies auf Ausreißer zu höheren Werten hin. Richtig, das arithmetische Mittel reagiert auf Ausreißer.
- (b) Ein ordinales Merkmal besitzt im Gegensatz zu einem nominalen Merkmal keine natürliche Rangfolge der Ausprägungen. Falsch, ein ordinales Merkmal besitzt eine Rangfolge, ein nominales nicht.
- (c) Die mittlere absolute Abweichung kann nur nur positive Werte annehmen. Richtig, es handelt sich um eine Summe positiver Größen.
- (d) Der Modus einer Stichprobe kardinaler Daten ist immer positiv. Falsch, die Merkmalswerte können negativ sein.
- (e) Der Modus einer Stichprobe nominaler Daten ist immer positiv. Falsch, nominale Daten lassen sich nicht als Zahl ausdrücken, sie können also auch nicht positiv sein.