Statistik 1, Test 3

Aufgabe 1

(a) Der Mittelwert der Länge ist

$$\begin{aligned}\mu_l &= \frac{5.1+4.9+4.7+4.6+5.0+5.4+4.6+5.0+4.4+4.9}{10} \\ &= 4.86\end{aligned}$$

(b) Der Mittelwert der Breite ist

$$\begin{aligned}\mu_b &= \frac{3.5+3.0+3.2+3.1+3.6+3.9+3.4+3.4+2.9+3.1}{10} \\ &= 3.31\end{aligned}$$

(c) Die Stichprobenvarianz der Länge ist

$$\begin{aligned}s_l^2 &= \frac{(5.1-4.86)^2+(4.9-4.86)^2+\cdots+(4.9-4.86)^2}{9} \\ &= 0.0849\end{aligned}$$

(d) Die Stichprobenvarianz der Breite ist

$$\begin{aligned}s_b^2 &= \frac{(3.5-3.31)^2+(3.0-3.31)^2+\cdots+(3.1-3.31)^2}{9} \\ &= 0.0943\end{aligned}$$

(e) Die Stichprobenkovarianz ist dann

$$\begin{aligned}s_{lb} &= \frac{(5.1-4.86)(3.5-3.31)+(4.9-4.86)(3.0-3.31)+\cdots+(4.9-4.86)(3.1-3.31)}{9} \\ &= 0.0704\end{aligned}$$

(f) Der Korrelationskoeffizient ist dann

$$\begin{aligned}r &= \frac{s_{lb}}{\sqrt{s_l^2}\cdot\sqrt{s_b^2}} \\ &= \frac{0.0704}{\sqrt{0.0849}\cdot\sqrt{0.0943}} \\ &= 0.7868\end{aligned}$$

$\Rightarrow$ Das lässt auf einen mittleren, linearen, positiven Zusammenhang schließen.

Aufgabe 2

(a) Zuerst müssen wir den Daten einen Rang zuordnen. Für die Michelin-Sterne gilt

x_i	2	2	3	3	3	3	3	4	4	4	4	4
$R(x_i)$	1.5	1.5	5	5	5	5	5	10	10	10	10	10

Für die Kreativität

y_i	0	0	0	0	0	1	1	0	0	1	1	1
$R(y_i)$	4	4	4	4	4	10	10	4	4	10	10	10

Der durchschnittliche Rang ist damit

$$\begin{aligned}\bar{R} &= \frac{1.5 + 1.5 + 5 + \cdots + 10 + 4 + 4 + \cdots + 10}{24} \\ &= 6.5\end{aligned}$$

(b) Der Korrelationskoeffizient nach Spearman ist dann

$$\begin{aligned}R &= \frac{1.5 \cdot 4 + 1.5 \cdot 4 + \cdots + 10 \cdot 4 + 10 \cdot 10 - 12 \cdot 6.5^2}{\sqrt{1.5^2 + 1.5^2 + \cdots + 10^2 - 12 \cdot 6.5^2} \cdot \sqrt{4^2 + 4^2 + \cdots + 10^2 - 12 \cdot 6.5^2}} \\ &= 0.396\end{aligned}$$

Es handelt sich um einen mittleren monotonen positiven Zusammenhang.

Aufgabe 3

(a) Die Kontingenztafel der absoluten Häufigkeiten ist

	männlich	weiblich	Σ
Informatik	15	5	20
Wiwi	35	35	70
Geisteswissenschaften	0	10	10
Σ	50	50	100

(b) Die Kontingenztafel der relativen Häufigkeiten ist

	männlich	weiblich	Σ
Informatik	0.15	0.05	0.2
Wiwi	0.35	0.35	0.7
Geisteswissenschaften	0	0.1	0.1
Σ	0.5	0.5	1

(c) χ^2 ist dann

$$\begin{aligned}\chi^2 &= 100 \left(\frac{15^2}{20 \cdot 50} + \frac{5^2}{20 \cdot 50} + \frac{35^2}{70 \cdot 50} + \frac{35^2}{70 \cdot 50} + \frac{0^2}{10 \cdot 50} + \frac{10^2}{10 \cdot 50} - 1 \right) \\ &= 15\end{aligned}$$

(d) Für C_{max} gilt

$$\begin{aligned} C_{max} &= \sqrt{\frac{2-1}{2}} \\ &= 0.7071 \end{aligned}$$

(e) Für C gilt

$$\begin{aligned} C &= \sqrt{\frac{15}{15+100}} \\ &= 0.3612 \end{aligned}$$

(f) Für C_{Korr} gilt

$$\begin{aligned} C_{Korr} &= \frac{0.3612}{0.7071} \\ &= 0.5108 \end{aligned}$$

Es besteht also ein mittlerer Zusammenhang.

Aufgabe 4

(a) Alter und Haarfarbe: Kontingenzkoeffizient

(b) Alter und Einkommen: Bravis-Pearson, Spearman, Kontingenzkoeffizient

(c) Abiturnote und Bachelornote: Spearman, Kontingenzkoeffizient

(d) Preis ein Kunstwerkes und Names des Künstlers: Kontingenzkoeffizient

(e) Körpergröße und Körpergewicht: Bravis-Pearson, Spearman, Kontingenzkoeffizient

(f) Zufriedenheit des Kunden und Rabatt: Spearman, Kontingenzkoeffizient

Aufgabe 5

Die Aussagen sind

- Wenn der Korrelationskoeffizient nach Pearson zwischen zwei Stichproben klein ist, so sind die Merkmale unabhängig voneinander. FALSCH, die Merkmale könnten z.B. auch quadratisch voneinander abhängen.
- $R_{XY} = \mathrm{Cor}(R(X), R(Y))$. RICHTIG
- Die Zuordnung $x \mapsto R(x)$ ist eineindeutig. FALSCH, da man aus $R(x)$ nicht auf x schließen kann.
- Der Korrelationskoeffizient von Spearman liegt immer im Intervall $[-1, 1]$. RICHTIG
- Genau dann wenn X und Y unabhängig sind, folgt $C_{Korr} = 0$. RICHTIG