

3 Multivariate deskriptive Statistik

zugehörige Seiten in Fahrmeir et al. (2007): Kap. 3.1 - 3.4

Aufgabe 27

Für die Kfz-Werkstätten einer Stadt wurden an einem bestimmten Tag die folgenden statistischen Merkmale erhoben

X : „Anzahl der Beschäftigten“ und Y : „Anzahl der reparierten Autos“.

Die Ergebnisse sind in folgender Kontingenztabelle dargestellt

Anzahl der Beschäftigten	Anzahl der reparierten Autos							
	3	5	6	8	10	11	12	15
2	3	2	0	0	0	0	0	0
3	1	2	2	0	0	0	0	0
5	1	0	4	4	1	0	0	0
8	0	1	4	5	3	5	2	0
10	0	0	0	1	1	0	3	5

- (a) Bestimmen Sie die Randhäufigkeiten von X und Y sowie die bedingten relativen Häufigkeiten von Y unter der Bedingung $X = 8$.
- (b) Geben Sie unter den Betrieben mit acht Beschäftigten den Anteil derjenigen an, die zehn Autos repariert haben.
- (c) Geben Sie den Anteil der Betriebe an, die genau acht Beschäftigte haben und höchstens zehn Autos repariert haben.
- (d) Geben Sie den Anteil der Betriebe an, die höchstens zehn Autos repariert haben.
- (e) Berechnen Sie die bedingten arithmetischen Mittel von Y unter der Bedingung $X = a_i$ für alle $i = 1, \dots, 5$.
- (f) Sind X und Y deskriptiv unabhängig? Begründen Sie Ihre Entscheidung.

Aufgabe 28

Der χ^2 -Koeffizient für n Beobachtungen zweier nominalskalierten Merkmale $X \in \{a_1, \dots, a_k\}$ und $Y \in \{b_1, \dots, b_\ell\}$ ist definiert durch

$$\chi^2 := \sum_{i=1}^k \sum_{j=1}^{\ell} \frac{(h_{ij} - e_{ij})^2}{e_{ij}} = n \left(\sum_{i=1}^k \sum_{j=1}^{\ell} \frac{h_{ij}^2}{h_{i\bullet} h_{\bullet j}} - 1 \right)$$

mit

$$e_{ij} := \frac{h_{i\bullet} h_{\bullet j}}{n}.$$

Zeigen Sie, dass $\chi^2 \in [0, n(\min\{k, \ell\} - 1)]$ gilt.

Aufgabe 29

Zehn befragte Passanten gaben ihre Einstellung zum lokalen Fußballverein (F) und zum dortigen Torschützenkönig der letzten Saison (T) anhand einer Skala von -4 bis +4 an:

Person	1	2	3	4	5	6	7	8	9	10
Einstellung zu F	1	-1	0	-2	2	4	3	1	-4	-3
Einstellung zu T	2	-2	2	0	3	4	2	-4	-4	0

- (a) Ermitteln Sie mit Hilfe einer geeigneten Maßzahl, ob und gegebenenfalls inwiefern ein deskriptiver Zusammenhang zwischen der Einstellung der Befragten zum Fußballverein und zum Torkönig besteht.
- (b) Ändert sich der Zusammenhang, wenn man nur die Personen mit positiver Einstellung zum lokalen Fußballverein berücksichtigt?

Aufgabe 30

Zeigen Sie, dass falls keine Bindungen vorliegen, für den Rangkorrelationskoeffizient nach Spearmans für zwei mindestens ordinalskalierte Merkmale X und Y die rechentechnisch günstigste Version

$$r_{SP} = 1 - \frac{6 \sum_{i=1}^n d_i^2}{(n^2 - 1)n},$$

mit den Rangdifferenzen $d_i = rg(x_i) - rg(y_i)$, $i = 1, \dots, n$ verwendet werden kann.

Aufgabe 31* (6 Punkte)

Die im folgenden gegebene Kontingenztafel mit relativen Häufigkeiten ist unvollständig. Vervollständigen Sie die Tabelle unter der Annahme, dass die beiden Merkmale „statistisch unabhängig“ sind.

	b_1	b_2	b_3	
a_1	0.3			
a_2			0.12	0.4

Aufgabe 32* (10 Punkte)

100 Personen wurden befragt, ob sie rauchen und ob bei ihnen eine chronische Bronchitis diagnostiziert worden sei. Die Befragung ergab folgende (Teil-)Ergebnisse:

10 Prozent der Befragten waren Raucher und hatten eine chronische Bronchitis. Ein Drittel der an chronischer Bronchitis erkrankten Personen raucht nicht. 60 Prozent der Raucher sind nicht an chronischer Bronchitis erkrankt.

- (a) Ermitteln Sie für diese Stichprobe, ob und gegebenenfalls inwiefern zwischen Rauchen und dem Auftreten von chronischer Bronchitis ein deskriptiver Zusammenhang besteht. Verwenden Sie hierfür eine geeignete Maßzahl mit dem Wertebereich $[0, 1]$.
- (b) Vergleichen Sie das Risiko, an chronischer Bronchitis zu erkranken zwischen der Gruppe der Raucher und der Gruppe der Nichtraucher.
- (c) Laden Sie sich unter <http://www.stat.uni-muenchen.de/service/datenarchiv/> den Datensatz zur Studie zu Chronischer Bronchitis und Staubkonzentration herunter. Lesen Sie die Daten mit dem Befehl `read.table` in R ein. Führen Sie für diesen Datensatz dieselben Analysen wie in Teil a) und b) durch. (Hinweis: Betrachten Sie dabei nur die Variablen `cbr` und `smoking`!)