Kontingenztafeln

Aufgabe 1

Die Teilnehmer der Vorlesung zur deskriptiven Statistik im Wintersemester 2014/2015 wurden zu Beginn des Semesters unter anderem über Auslandsaufenthalte befragt. Folgende (Teil-)Ergebnisse ergaben sich für die 44 Teilnehmer der Umfrage, die diese Frage beantworteten:

27,27 Prozent der Befragten sind weiblich und geben an, während des Studiums ein Auslandssemester machen zu wollen. 63,64 Prozent der männlichen Befragten planen, kein Auslandssemester während des Studiums zu absolvieren. Insgesamt wollen 45,45 Prozent der Befragten ein Auslandssemester machen.

(a) Erstellen Sie die zugehörige Kontingenztabelle der absoluten Häufigkeiten.

(b) Ermitteln Sie für diese Daten, ob und gegebenenfalls inwiefern zwischen dem Geschlecht und einem geplanten Auslandsaufenthalt während des Studiums ein Zusammenhang besteht. Verwenden Sie hierfür eine geeignete Maßzahl mit dem Wertebereich $[0, 1]$ und interpretieren Sie das Ergebnis.

(c) Vergleichen Sie die Chancen für einen geplanten Auslandsaufenthalt zwischen den Geschlechtern mit einer geeigneten Maßzahl.

(d) Wie groß ist der Anteil unter den weiblichen Befragten, ein Auslandssemester absolvieren zu wollen?

Aufgabe 2

Die im Folgenden gegebene Kontingenztafel mit erwarteten relativen Häufigkeiten ist unvollständig. Geben Sie unter der Annahme, dass die beiden Merkmale unabhängig sind, die entsprechenden Werte für die unbekannten Koeffizienten an.

	b_1	b_2	b_3	$\sum$
a_1	0.3	f_{12}	f_{13}	$f_{1\bullet}$
a_2	f_{21}	f_{22}	0.12	0.4
$\sum$	$f_{\bullet 1}$	$f_{\bullet 2}$	$f_{\bullet 3}$	$f_{\bullet\bullet}$

Aufgabe 3

Für 50 Kfz-Werkstätten einer Stadt wurden an einem bestimmten Tag die Merkmale *Anzahl der Beschäftigten (X)* und *Anzahl der reparierten Autos (Y)* erhoben. Die Ergebnisse sind in der folgenden Kontingenztabelle dargestellt:

		Anzahl der reparierten Autos							
		3	5	6	8	10	11	12	15
	2	3	2	0	0	0	0	0	0
Anzahl der	3	1	2	2	0	0	0	0	0
Beschäftigten	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. Geben Sie außerdem die bedingten relativen Häufigkeiten von Y unter der Bedingung $X = 8$ an.

(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 $i = 1, \dots, 5$.

Hinweis: Das bedingte arithmetische Mittel ist ein gewichtetes arithmetisches Mittel. Konkret ist dies der Form

$$\bar{y}_{X=a_i} = \frac{1}{\sum_{j=1}^{8} w_j} \sum_{j=1}^{8} b_j \cdot w_j,$$

wobei b_j die Ausprägungen von Y sind und sich die Gewichte w_j aus den bedingten relativen Häufigkeiten ergeben.

(f) Sind X und Y empirisch unabhängig? Begründen Sie Ihre Entscheidung.

Kontingenztafeln

Aufgabe 1

Die Teilnehmer der Vorlesung zur deskriptiven Statistik im Wintersemester 2014/2015 wurden zu Beginn des Semesters unter anderem über Auslandsaufenthalte befragt. Folgende (Teil-)Ergebnisse ergaben sich für die 44 Teilnehmer der Umfrage, die diese Frage beantworteten:

27,27 Prozent der Befragten sind weiblich und geben an, während des Studiums ein Auslandssemester machen zu wollen. 63,64 Prozent der männlichen Befragten planen, kein Auslandssemester während des Studiums zu absolvieren. Insgesamt wollen 45,45 Prozent der Befragten ein Auslandssemester machen.

(a) Erstellen Sie die zugehörige Kontingenztabelle der absoluten Häufigkeiten.

(b) Ermitteln Sie für diese Daten, ob und gegebenenfalls inwiefern zwischen dem Geschlecht und einem geplanten Auslandsaufenthalt während des Studiums ein Zusammenhang besteht. Verwenden Sie hierfür eine geeignete Maßzahl mit dem Wertebereich $[0, 1]$ und interpretieren Sie das Ergebnis.

(c) Vergleichen Sie die Chancen für einen geplanten Auslandsaufenthalt zwischen den Geschlechtern mit einer geeigneten Maßzahl.

(d) Wie groß ist der Anteil unter den weiblichen Befragten, ein Auslandssemester absolvieren zu wollen?

a)

#	A_h	$\bar{A}_h$	
M_h	h_{11} 8	h_{12} 14	$h_{1\bullet}$ 22
W_h	h_{21} 12	h_{22} 10	$h_{2\bullet}$ 22
	$h_{\bullet 1}$ 20	$h_{\bullet 2}$ 24	44

$h_{\bullet 1} = 0.4545 \cdot 44 = 20$

$h_{21} = 0.2727 \cdot 44 = 12 \Rightarrow h_{11} = h_{\bullet 1} - h_{21} = 20 - 12 = 8$

$\frac{h_{11}}{h_{1\bullet}} = \frac{8}{h_{1\bullet}} \overset{!}{=} 1 - 0.6364 \Rightarrow h_{1\bullet} = 22$

b)

$$\chi^2 = \frac{44 \cdot (8 \cdot 10 - 12 \cdot 14)}{20 \cdot 24 \cdot 22 \cdot 22} = \frac{22}{15} = 1.4\bar{6} \qquad K = \sqrt{\frac{\frac{22}{15}}{44 + \frac{22}{15}}} = \frac{1}{\sqrt{31}} = 0.18$$

$$K^* = \frac{0.18}{\sqrt{\frac{1}{2}}} = \sqrt{\frac{2}{31}} \approx 0.25$$

. Es besteht eine mittelstarke Abhängigkeit zwischen Geschlecht und geplantem Auslandsaufenthalt.

c)

$$\gamma(A_h, \bar{A}_h \mid M_h, W_h) = \frac{\gamma(A_h, \bar{A}_h \mid M_h)}{\gamma(A_h, \bar{A}_h \mid W_h)} = \frac{8 \cdot 10}{12 \cdot 14} = \frac{10}{21} = 0.48 < 1.$$

Die Chancen für den geplanten Auslandsaufenthalt sind bei Männern etwa halb so gross wie bei Frauen.

d)

$$\frac{12}{22} = 0.\overline{54} \approx 54\%$$

Aufgabe 2

Die im Folgenden gegebene Kontingenztafel mit erwarteten relativen Häufigkeiten ist unvollständig. Geben Sie unter der Annahme, dass die beiden Merkmale unabhängig sind, die entsprechenden Werte für die unbekannten Koeffizienten an.

	b_1	b_2	b_3	$\sum$
a_1	0.3	f_{12}	f_{13}	$f_{1\bullet}$
a_2	f_{21}	f_{22}	0.12	0.4
$\sum$	$f_{\bullet 1}$	$f_{\bullet 2}$	$f_{\bullet 3}$	$f_{\bullet\bullet}$

	b_1	b_2	b_3	Σ
a_1	0.3	0.12	0.18	0.6
a_2	0.2	0.08	0.12	0.4
Σ	0.5	0.2	0.3	1

$f_{\bullet\bullet} = 1 \rightarrow f_{1\bullet} = 0.6 \rightarrow f_{\bullet 1} = \frac{f_{11}}{f_{1\bullet}} = \frac{0.3}{0.6} = 0.5 \rightarrow f_{21} = f_{\bullet 1} \cdot f_{2\bullet} = 0.5 \cdot 0.4 = 0.2 \rightarrow f_{22} = f_{2\bullet} - f_{22} - f_{23} = 0.4 - 0.2 - 0.12 = 0.08$

$f_{\bullet 2} = f_{12} + f_{22} = f_{12} + 0.08$, $f_{12} = 0.6 \cdot f_{\bullet 2}$

$f_{\bullet 3} = f_{13} + f_{23} = f_{13} + 0.12$, $f_{13} = 0.6 \cdot f_{\bullet 3}$

$$\begin{cases} f_{\bullet 2} = 0.6 \cdot f_{\bullet 2} + 0.08 \Rightarrow f_{\bullet 2} = \frac{0.08}{0.4} = 0.2 \\ f_{\bullet 3} = 0.6 \cdot f_{\bullet 3} + 0.12 \Rightarrow f_{\bullet 3} = \frac{0.12}{0.4} = 0.3 \end{cases} \Rightarrow f_{12} = 0.12, \; f_{13} = 0.18$$

Aufgabe 3

Für 50 Kfz-Werkstätten einer Stadt wurden an einem bestimmten Tag die Merkmale *Anzahl der Beschäftigten (X)* und *Anzahl der reparierten Autos (Y)* erhoben. Die Ergebnisse sind in der folgenden Kontingenztabelle dargestellt:

		Anzahl der reparierten Autos								
		3	5	6	8	10	11	12	15	
Anzahl der Beschäftigten	2	3 $\frac{1}{2}$	2 $\frac{1}{2}$	0 1	0 1	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$	5
	3	1 $\frac{1}{2}$	2 $\frac{1}{2}$	2 1	0 1	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$	5
	5	1 1	0 1	4 2	4 2	1 1	0 1	0 1	0 1	10
	8	0 2	1 1	4 4	5 4	3 2	5 2	2 2	0 2	20
	10	0 1	0 2	0 2	1 2	1 1	0 1	3 1	5 1	10
		5	5	10	10	5	5	5	5	50

$1 + 1 + 2 + 16 = 20$
$1 + 1 + 0 + 4 = 6$
$1 + 1 + \frac{9}{2} + 1 = 15/2$
$1 + 0 + \frac{1}{2} + 0 = 3/2$
$2 + 2 + \frac{1}{4} + \frac{1}{2} = 19/4$
$2 + 2 + 0 + 2 = 6$
$9 + 1 + 0 + 2 = 12$
$25/2 + \frac{1}{2} + 0 + 2 + 1 = 16$

(a) Bestimmen Sie die Randhäufigkeiten von X und Y. Geben Sie außerdem die bedingten relativen Häufigkeiten von Y unter der Bedingung $X = 8$ an.

b_i	3	5	6	8	10	11	12	15
$f_Y(b_i \mid X=8)$	0	0.05	0.2	0.25	0.15	0.25	0.1	0

(b) Geben Sie unter den Betrieben mit acht Beschäftigten den Anteil derjenigen an, die zehn Autos repariert haben. $f_Y(10 \mid X=8) = 15\%$

(c) Geben Sie den Anteil der Betriebe an, die genau acht Beschäftigte haben und höchstens zehn Autos repariert haben. $f(X=8, Y \leq 10) = f(X=8, Y=3) + f(X=8, Y=5) + f(X=8, Y=6) + f(X=8, Y=8) + f(X=8, Y=10) = \frac{13}{50} = 26\%$

(d) Geben Sie den Anteil der Betriebe an, die höchstens zehn Autos repariert haben. $f_{\cdot 3} + f_{\cdot 5} + \ldots + f_{\cdot 10} = \frac{35}{50} = 70\%$

(e) Berechnen Sie die bedingten arithmetischen Mittel von Y unter der Bedingung $X = a_i$ für alle $i = 1, \ldots, 5$.

Hinweis: Das bedingte arithmetische Mittel ist ein gewichtetes arithmetisches Mittel. Konkret ist dies der Form

$$\bar{y}_{X=a_i} = \frac{1}{\sum_{j=1}^{8} w_j} \sum_{j=1}^{8} b_j \cdot w_j,$$

wobei b_j die Ausprägungen von Y sind und sich die Gewichte w_j aus den bedingten relativen Häufigkeiten ergeben.

(f) Sind X und Y empirisch unabhängig? Begründen Sie Ihre Entscheidung.

e)

$$\bar{y}_{X=2} = \frac{3 \cdot 3 + 2 \cdot 5}{5} = 3.8$$

$$\bar{y}_{X=3} = \frac{1 \cdot 3 + 2 \cdot 5 \cdot 2 \cdot 6}{5} = 5$$

$$\bar{y}_{X=5} = \frac{1 \cdot 3 + 4 \cdot 6 + 4 \cdot 8 + 1 \cdot 10}{10} = 6.9$$

$$\bar{y}_{X=8} = \frac{1 \cdot 5 + 4 \cdot 6 + 5 \cdot 8 + 3 \cdot 10 + 5 \cdot 11 + 2 \cdot 12}{20} = 8.9$$

$$\bar{y}_{X=10} = \frac{3 \cdot 12 + 5 \cdot 15}{10} = 11.1$$

f) Je mehr Mitarbeiter eine Werkstadt hat, desto mehr Autos kann diese reparieren. Das erkennt man u.a. an den steigenden bedingten arithmetischen Mitteln.

$\chi^2 = 73.75$. $K = 0.77$. $K^* = 0.86$ $\rightarrow$ Starker Zusammenhang zwischen X und Y