

Aufgabe 33

Gegeben sind zwei Datensätze jeweils vom Umfang $n = 5$ (siehe Beispiel aus der Vorlesung):

ohne Bindungen						mit Bindungen					
X	0	0.3	0.5	0.4	1	X	0	0.3	0.5	0.3	1
Y	0	0	0	1	1	Y	0	0	0	1	1

Zeichnen Sie die zugehörigen ROC-Kurven.

Aufgabe 34

In der sog. Sonntagsfrage wird regelmäßig folgende Frage gestellt: *“Welche Partei würden Sie wählen, wenn am nächsten Sonntag Bundestagswahl wäre?”* Für den Befragungszeitraum 11.1. bis 24.1.1995 ergab sich dabei folgende Kontingenztafel:

	CDU/CSU	SPD	FDP	Grüne	Rest	
Männer	144	153	17	26	95	435
Frauen	200	145	30	50	71	496
	344	298	47	76	166	931

Bei der Untersuchung des Zusammenhangs zwischen Geschlecht und Parteipräferenz betrachte man die sich ergebenden (2×2) -Tabellen, wenn man die Union CDU/CSU jeweils lediglich einer anderen Partei gegenüberstellt. Bestimmen und interpretieren Sie jeweils Odds-Ratio, den χ^2 -Koeffizienten und den Kontingenzkoeffizienten.

Aufgabe 35

Kendall's Tau: Die Noten in Mathematik und Physik einiger Schüler sind in folgender Tabelle gegeben:

Schüler	Mathematiknote (X)	Physiknote (Y)
1	5	5
2	4	4
3	1	3
4	2	2
5	3	1
6	4	3
7	2	3

Berechnen Sie Kendall's Tau für diese Stichprobe. Was lässt sich über den Zusammenhang von Mathematik- und Physiknote der Schüler aussagen?

[Hilfe: Kendall's Tau bei Bindungen:

$$\tau = \frac{N_c - N_d}{\sqrt{(N_c + N_d + T_x)(N_c + N_d + T_y)}} \quad]$$

4 Regression

zugehörige Seiten in Fahrmeir et al. (2007): Kap. 3.6

Aufgabe 36

Für die zehn umsatzstärksten Unternehmen Deutschlands ergaben sich 1995 folgende Umsätze Y (in Mrd. DM) und Beschäftigungszahlen X (in Tsd.):

Unternehmen	Umsatz (Y)	Beschäftigte (X)
1	103.54	311.0
2	88.76	373.0
3	88.12	242.4
4	72.37	125.2
5	65.50	135.1
6	52.17	161.6
7	49.40	106.6
8	46.14	115.8
9	44.58	142.9
10	41.93	83.8

Die dahinterstehenden Firmen waren Daimler-Benz, Siemens, Volkswagen, VEBA, RWE, Hoechst, BASF, BMW, Bayer und VIAG.

- (a) Zeichnen Sie ein Streudiagramm für die 10 Unternehmen und interpretieren Sie dieses.
- (b) Man bestimme den Bravais-Pearson- und den Spearmanschen Korrelationskoeffizienten.
- (c) Wie ändern sich die Korrelationskoeffizienten, wenn man in absoluten DM-Umsätzen und absoluten Beschäftigungszahlen rechnet? Wie wenn man die Umsätze in Euro umrechnet?
- (d) Man berechne die Regressionskoeffizienten für eine lineare Einfachregression von Y auf X . Man zeichne die entsprechende Ausgleichsgerade in das Streudiagramm aus (a). Man gebe das Bestimmtheitsmaß an.

Aufgabe 37* (16 Punkte)

Die folgende Tabelle zeigt die Ausprägungen zweier „Indikatoren der Wissensgesellschaft“ für verschiedene Länder (In Anlehnung an die Quelle: Deutschland in Zahlen, Ausgabe 2005).

Land	#Forscher im Jahr 2002 (je 1000 Beschäftigte)	#Triadepatente ¹ im Jahr 2002 (je 1 Mio Beschäftigte)
Belgien	7.8	215
Deutschland	6.8	280
Dänemark	9.4	356
Spanien	5.2	38
Frankreich	8.2	167
Irland	4.8	127
Portugal	3.7	8
GB	6.2	201
Japan	9.7	346
USA	8.1	241

- (a) Welchen Zusammenhang erwarten Sie zwischen den beiden Variablen „#Forscher im Jahr 2002“ und „#Triadepatente im Jahr 2002“?
- (b) Zeichnen Sie für die Daten ein Streudiagramm mit geeignet gewählter X-Variable. Was sagt es über den Zusammenhang aus?
- (c) Ermitteln Sie die Varianzen der beiden Merkmale sowie die Kovarianz.
- (d) Bestimmen und interpretieren Sie nun den Bravais-Pearson-Korrelationskoeffizienten sowie den Rangkorrelationskoeffizienten nach Spearman.
- (e) Wie ändern sich die Koeffizienten, wenn man die Anzahl der Triadepatente je 1000 Beschäftigte gemessen hätte?
- (f) Berechnen Sie die Parameter einer linearen Einfachregression mit „#Triadepatente im Jahr 2002 (je 1 Mio Beschäftigte)“ als abhängiger Variable. Beurteilen Sie die Güte des Modells.
- (g) Zeichnen Sie die Regressionsgerade in das Streudiagramm aus (b).

Sie können für alle Berechnungen die Software R zu Hilfe nehmen!

(Hinweis: Benutzen Sie dazu die Funktionen `var`, `cor` und `lm`!)

¹Patente, die zusätzlich zum Inland in mindestens zwei Auslandsmächten der Triaderegion USA, Europa, Japan angemeldet werden.