

## 14. Tutoriumsblatt: Regressionsanalyse und Wiederholung, Teil 2

### Aufgabe 1<sup>1</sup>: (Regression)

Ein Medikament zur Behandlung von Depressionen steht im Verdacht, als Nebenwirkung das Reaktionsvermögen zu reduzieren. In einer Klinik wurde deshalb eine Studie durchgeführt, an der zehn zufällig ausgewählte Patienten teilnahmen, die das Präparat in verschiedenen Dosierungen verabreicht bekamen. Es ergeben sich die folgenden Werte für die Dosierung  $X$  in mg und die dazugehörige Reaktionszeit  $Y$  in Sekunden.

<b>Dosierung</b> ( $x_i$ )	1	5	3	8	2	2	10	8	7	4
<b>Reaktionszeit</b> ( $y_i$ )	1	6	1	6	3	2	8	5	6	2

- Was sagt das Streudiagramm über den Zusammenhang von  $X$  und  $Y$  aus?
- Berechnen Sie einen geeigneten Korrelationskoeffizienten. Was sagt dieser Wert über den Zusammenhang von  $X$  und  $Y$  aus?
- Passen Sie eine Gerade an die beobachteten Datenpunkte an. Berechnen Sie hierfür die KQ-Schätzer, interpretieren Sie diese, und skizzieren Sie die sich ergebende Regressionsgerade, indem Sie auch die Werte der Schätzer eintragen.
- Beurteilen Sie die Güte Ihrer Anpassung!
- Ein Patient wird mit einer Dosis von 5.5 mg des Medikaments behandelt. Welche Reaktionszeit prognostizieren Sie?

Hinweis:

$$\sum_{i=1}^{10} x_i = 50, \sum_{i=1}^{10} x_i^2 = 336, \sum_{i=1}^{10} y_i = 40, \sum_{i=1}^{10} y_i^2 = 216, \sum_{i=1}^{10} x_i y_i = 262$$

### Aufgabe 2: (Regression)

Die abhängige Variable  $Y$  sei das Geld in Euro, das die Befragten monatlich für Schuhe ausgeben. Geschlecht (0: männlich, 1: weiblich) und das monatliche Nettoeinkommen in Euro seien die unabhängigen Variablen.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	7.427758	0.432091	17.19	<2e-16
Geschlecht	9.375962	0.307495	30.49	<2e-16
Nettoeinkommen	0.113609	0.007685	14.78	<2e-16

Interpretieren Sie die Parameterschätzer des Regressionsoutputs!

<sup>1</sup>angelehnt an: Fahrmeir et al. (2005<sup>4</sup>): "Arbeitsbuch Statistik", Springer

**Aufgabe 3:** (Wiederholung, Teil 2)

Zur Wiederholung der ersten Kapitel, siehe auch Tutorium 9 (Wiederholungsquiz, Teil 1).

- a) Von 100 Personen wurden die Merkmale Geschlecht ( $X$ ) und Zahlungsart ( $Y$ ) erhoben. Ergänzen Sie die folgende Kontingenztafel, indem Sie auch die Randhäufigkeiten berechnen.

$X \setminus Y$	bar	Kreditkarte	
Männer	30	25	
Frauen		17	

- Wie viel Prozent der Befragten sind männlich und haben in bar bezahlt?
  - Wie viel Prozent der Frauen haben mit Kreditkarte bezahlt?
  - Sind die beiden Merkmale  $X$  und  $Y$  unabhängig?
- b) Diskutieren Sie kurz das Zusammenhangsmaß für nominalskalierte Merkmale, welches die Grundlage für einige andere Maße bildet. Schlagen Sie jeweils ein Maß vor, welches herangezogen werden könnte/sollte, wenn...
- ... den zu vergleichenden Kontingenztafeln ein unterschiedlicher Stichprobenumfang und eine unterschiedliche Spaltenanzahl zugrundeliegt
  - ... insbesondere die Richtung des Zusammenhangs von Interesse ist (im Vierfelder-Fall)
- c) Erläutern Sie die Idee der PRE-Maße und nennen Sie ein Maß, welches auf dieser Idee beruht.
- d) Was ist der wesentliche Unterschied zwischen dem Korrelationskoeffizienten nach Spearman und dem nach Bravais-Pearson hinsichtlich der Interpretation des Zusammenhangs?