Hochschule Düsseldorf
University of Applied Sciences

Fachbereich Wirtschaftswissenschaften
Faculty of Business Studies

Tutorium deskriptive Statistik

SoSe 2025

Die folgenden Aufgaben wurden

Der Klausuraufgabensammlung zu Bamberg/Baur/Krapp: Statistik, Oldenbourg-Verlag, 18. Auflage 2017,

sowie

Klausuren der Vorsemester

entnommen.

Aufgabe 1

Quelle: Bamberg Aufgabe 1

Im Skigebiet A wurden in den letzten zwei Wochen im Jahr 2005 täglich folgende Schneehöhen (in cm) im Tal gemessen:

30, 28, 31, 40, 45, 40, 45, 40, 45, 60, 60, 55, 70, 90

a) Berechnen Sie das arithmetische Mittel, die durchschnittliche Abweichung von $\bar{x}$ sowie die mittlere quadratische Abweichung für obige Daten.

b) Berechnen Sie für eine englische Nachrichtenagentur die unter a) bestimmten Größen in der Maßeinheit Inch (1 Inch = 2;54 cm).

Im Skigebiet B wurden im gleichen Zeitraum die Schneehöhen im Tal (in cm) gemessen. Aus den dort gemessenen Höhen wurden folgende Größen bestimmt:

$$\sum_{i=1}^{14} y_i = 700 \quad ; \quad \sum_{i=1}^{14} y_i^2 = 38250$$

c) Bestimmen Sie die mittlere quadratische Abweichung der Schneehöhen für Skigebiet B.

d) Berechnen Sie das arithmetische Mittel sowie die mittlere quadratische Abweichung für sämtliche (in A und B) gemessenen Schneehöhen.

Lösung

a)

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\bar{x} = \frac{679}{14} = \underline{48{,}5}$$

$$s_{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$$

$$s_{\bar{x}} = \frac{1}{14}\sum_{i=1}^{n} |x_i - 48{,}5| = \frac{185}{14} = \underline{13{,}214}$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$$

$$s^2 = \frac{1}{14}\sum_{i=1}^{n}\left(x_i - 48{,}5\right)^2 \underline{=} \frac{3813{,}5}{14} \underline{=272{,}393}$$

b)

$$\bar{x} = \frac{48{,}5}{2{,}54} = \underline{19{,}09}$$

Lineare Transformation

$$y_i = a + bx_i;\; b = \frac{1}{2{,}54}; a = 0$$

$$s_{\bar{x}} = \frac{13{,}214}{2{,}54} = \underline{5{,}202}$$

$$s_y{}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 = b^2 s_x{}^2$$

$$s^2 = \frac{1}{14}\sum_{i=1}^{n}\left(y_i - 19{,}09\right)^2 = \frac{591{,}09}{14} = \underline{42{,}221}$$

c) $\underline{s^2 = 232{,}143}$

-

Hinweis auf „Verschiebesatz“:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \qquad - \bar{x}^2$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$s^2 = \frac{1}{14}\cdot 38.250 - 50^2 = \underline{\underline{232{,}143}}$$

$$\Uparrow$$

$$\frac{700}{14}$$

d)

$$\bar{x}_{Ges} = \underline{\underline{49{,}25}};\; s^2_{\mathrm{Ges}} = \underline{\underline{258{,}83}}$$

$$\bar{x}_{ges} = \frac{1}{n}\sum_{j=1}^{r} n_j \cdot \bar{x}_j$$

$$\bar{x}_{ges} = \frac{1}{28}(14\cdot 48{,}5 + 14\cdot 50) = \underline{\underline{49{,}25}}$$

$$\Downarrow \qquad\qquad \Downarrow$$

Gebiet A Gebiet B

$$s^2_{ges} = \frac{1}{n}\sum_{j=1}^{r} n_j \cdot s_j^2 + \frac{1}{n}\sum_{j=1}^{r} n_j (\bar{x}_j - \bar{x}_{ges})^2$$

$$s^2_{ges} = \frac{1}{28}(14\cdot 272{,}393 + 14\cdot 232{,}143) + \frac{1}{28}(14\cdot(48{,}5 - 49{,}25)^2 + 14\cdot(50 - 49{,}25)^2)$$

$$s^2_{ges} \underline{= 252{,}83}$$

□Aufgabe 2

Quelle: Bamberg Aufgabe 2

20 Bundesbürger werden bezüglich ihrer Ausgaben für Geflügelprodukte (in € in den letzten zwei Monaten befragt. Unter den 20 Befragten befinden sich 5, die in der letzten Zeit weder Nachrichten gehört noch Zeitung gelesen haben. Diese 5 Befragten geben folgende Werte an:

120, 23, 44, 16, 97

a) Berechnen Sie das arithmetische Mittel, die Spannweite, die Standardabweichung und den Variationskoeffizienten der 5 Uninformierten.

Die mittlere quadratische Abweichung bei allen 20 Befragten beträgt 1.155,5, die innerhalb der Gruppe der 15 Informierten hingegen 396. Das arithmetische Mittel letzterer Gruppe beträgt 12.

b) Berechnen Sie die interne und die externe Varianz.

c) Berechnen Sie das arithmetische Gesamtmittel der Merkmalswerte aller 20 Befragten.

Lösung

a) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$\bar{x}_1 = \frac{300}{5} = \underline{60}$

SP= max x_i - min x_i

i i

$SP_1 = \underline{\underline{120 - 16 = 104}}$

s= $\sqrt{s^2}$

$s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$

$s^2 = \frac{8.530}{5} = \underline{1.706}$

$\underline{s=}\sqrt{1.706} \approx \underline{41,30}$

V= $\frac{s}{\bar{x}}$

$V_1 = \frac{41,3}{60} = \underline{0,688396}$

-

b)

$$\underbrace{s_{ges}^2 = \frac{1}{n}\sum_{j=1}^{r} n_j s_j^2}_{\text{int. Varianz}} + \underbrace{\frac{1}{n}\sum_{j=1}^{r} n_j \left(\bar{x}_j - \bar{x}_{ges}\right)^2}_{\text{ext. Varianz}}$$

$$1.155{,}5 = \frac{1}{20}(5 * 1.706 + 15 * 396) + \text{ext. Varianz}$$

$$1.155{,}5 = 723{,}5 + \text{ext. Varianz} \qquad | - 723{,}5$$

<u>Ext. Varianz = 432</u>

<u>Variation von Aufgabe 2) b:</u>

$$\text{ext. Varianz} = \frac{1}{n}\sum_{j=1}^{n} n_j (\bar{x} - \bar{x}_y)^2$$

$$\underline{\text{ext. Varianz}} = \frac{1}{20}(5 * 36^2 + 15 * 12^2) = \frac{8.640}{20} \underline{= 432}$$

c)

$$\bar{x}_y = \frac{1}{n}\sum_{j=1}^{r} n_j \bar{x}_j$$

$$\bar{x}_y = \frac{1}{20}(5 * 60 + 15 * 12) = \frac{480}{20} = \underline{24}$$

Aufgabe 3

Quelle: Klausur SS 09

An einer Wetterstation wurden in den letzten 14 Tagen folgende Höchsttemperaturen gemessen : (in °C) 18,5; 19,2; 20,0; 17,5; 16,8; 17,0; 19,5; 15,6; 17,1; 18,4; 21,3; 22,1; 24,5; 22,1.

Berechnen Sie (auf 2 Stellen nach dem Komma)

a) Spannweite,
b) Modus,
c) Median,
d) arithmetisches Mittel,
e) Standardabweichung und
f) Variationskoeffizienten

für diese Daten.

Lösung

Musterlösung:

Spannweite 8,9
Modus 22,1
Median 18,85
arithm. Mittel 19,26
Standardabw. 2,42 (bei Anwendung des Verschiebesatzes ~ 2,40)
Variationskoeff 0,13

Aufgabe 4

Quelle: Klausur WS 09/10

Die Verbraucherzentrale hat an einem Stichtag folgende Preise für Butter (jeweils EUR pro 250 Gr.) in 10 Supermärkten in der Innenstadt von Düsseldorf festgestellt:

0,80	1,05	1,15	0,90	0,95	0,89	1,25	1,20	1,02	0,99

Bei 20 weiteren Supermärkten in den Außenbezirken der Stadt wird ein durchschnittlicher Butterpreis (= arithmetisches Mittel) von EUR 1,05 und eine Varianz (= mittlere quadratische Abweichung) von 0,015 ermittelt.

a) Berechnen Sie das arithmetische Mittel (mit 2 Nachkommastellen) und die Varianz (mit 4 Nachkommastellen) der 10 Butterpreise in der Innenstadt.

b) Berechnen Sie das arithmetische Mittel und die Varianz aller 30 Supermärkte.

Lösung

a) Arith. Mittel (10 Superm.): 1,02; Var. (10 Superm.): 0,0189

b) Arith. Mittel (30 Superm.): 1,04; Var. (30 Superm.): 0,0165

Aufgabe 5

Quelle: Bamberg Aufgabe 3

a) Nach 10-maligem Werfen eines Würfels wurden folgende Werte festgestellt:

$$\bar{x} = 3;\ x_{Med} = 3;\ SP = 4;\ h(3) = 0;\ h(5) = 2:$$

Genau die beiden Werte 2 und 4 sind Modalwerte. Wie lautet die Urliste?

b) Mit einem zweiten Würfel wurden folgende Ergebnisse geworfen:

Augenzahl	1	2	3	4	5	6
Anzahl	2	0	2	4	0	4

Bestimmen Sie das arithmetische Mittel und die mittlere quadratische Abweichung dieser Ergebnisse des zweiten Würfels.

c) Wie lautet die gemeinsame mittlere quadratische Abweichung aller Ergebnisse der beiden in a) und b) beschriebenen Würfelvorgänge? (Die mittlere quadratische Abweichung der Resultate aus a) hat den Wert $s^2 = 2{,}2$.)

Lösung

a) (1, 1, 2, 2, 2, 4, 4, 4, 5, 5)

b)

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{k} a_j h(a_j)$$

$$\bar{x} = \frac{1}{12}(1\cdot 2 + 2\cdot 0 + 3\cdot 2 + 4\cdot 4 + 5\cdot 0 + 6\cdot 4) = \frac{48}{12} = \underline{\underline{4}}$$

c) $\bar{x}_{Ges} = \frac{10 \cdot 3 + 12 \cdot 4}{22} = \frac{39}{11} = 3,545$

$$s^2_{Ges} = \frac{1}{22}(10 \cdot 2,2 + 12 \cdot 3) + \frac{1}{22}[10 \cdot (3 - \frac{39}{11})^2 + 12 \cdot (4 - \frac{39}{11})^2] = \underline{\underline{2,884}}$$

□Aufgabe 6

Quelle: Bamberg Aufgabe 4

An einem Lehrstuhl sind sechs Personen beschäftigt, nämlich ein Professor, eine Sekretärin und vier Assistenten. Von diesen Personen wurden im vergangenen Jahr folgende Strecken (gemessen in km) mit dem Fahrrad zurückgelegt:

Professor	Sekretärin	Assistent 1	Assistent 2	Assistent 3	Assistent 4
3.000	0	2.700	3.000	3.000	6.300

Bestimmen Sie zu diesen Daten

a) die durchschnittliche Abweichung vom Median,
b) die mittlere quadratische Abweichung,
c) die Werte der empirischen Verteilungsfunktion an den Stellen 2000 und 3000,
d) alle Knickpunkte der Lorenzkurve.

Lösung

a) Urliste der Größe nach sortieren: .0; 2.700; 3.000; 3.000; 3.000; 6.300 $\Rightarrow x_{med} = 3.000$

$$\Rightarrow \bar{s}_{x_{med}} = \frac{1}{6} \cdot (3.000 + 300 + \cdots + 3.300) = \underline{\underline{1.100}}$$

b)

$$\bar{x} = \frac{1}{6} \cdot (0 + 2.700 + \cdots + 6.300) = 3.000 \Rightarrow s^2 = \frac{1}{6} \cdot (3.000^2 + 300^2 + \cdots + 3.300^2) = \underline{\underline{3.330.000}}$$

c)

$$H(x) = \sum_{a_j \le x} h(a_j) \qquad F(x) = \frac{1}{n} H(x)$$

$$F(x)\begin{cases} \frac{1}{6} & \text{für } x < 2.700 \\ \frac{2}{6} & \text{für } 2.700 \le x < 3.000 \\ \frac{5}{6} & \text{für } 3.000 \le x < 6.300 \\ 1 & \text{für } 6.300 \le x \end{cases}$$

$$\Rightarrow F(2.000) = \underline{\underline{\frac{1}{6}}}; \quad F(3.000) = \underline{\underline{\frac{5}{6}}}$$

d)

K	1	2	3	4	5	6	Σ
xk	0	2700	3000	3000	3000	6300	18000
uk	1/6	2/6	3/6	4/6	5/6	1	
pk	0	3/20	1/6	1/6	1/6	7/20	1
vk	0	3/20	19/60	29/60	13/20	1	

Knickpunkte: $(0,0);\ (\frac{1}{6},0);\ (\frac{1}{3},\frac{3}{20});\ (\frac{5}{6},\frac{13}{20});\ (1,1)$

Aufgabe 7

Quelle: Bamberg Aufgabe 5

Eine Sparkasse hat in einer Stadt fünf Filialen A, …,E. Folgende Tabelle gibt die Anzahl der Beschäftigten in den einzelnen Filialen an:

Filiale	A	B	C	D	E
Beschäftigtenzahl	18	24	63	18	27

Zunächst werden die einzelnen Filialen als Merkmalsträger und die Beschäftigtenzahl als Merkmal X angesehen.

Berechnen Sie mit diesen Daten

a) die Spannweite,
b) die durchschnittliche Abweichung vom Median,
c) die Werte der empirischen Verteilungsfunktion an den Stellen 17, 18 und 19,
d) alle Knickpunkte der Lorenzkurve,
e) den normierten Gini-Koeffizienten.

Lösung

a) SP = $\underline{45}$

b) $x_{med} = 24 \Rightarrow \bar{s}_{Med} = \underline{\underline{10{,}80}}$

c)

$$F(x)\begin{cases} 0 \text{ für } x < 18 \\ \frac{2}{5} \text{ für } 18 \le x < 24 \\ \frac{3}{5} \text{ für } 24 \le x < 27 \\ \frac{4}{5} \text{ für } 27 \le x < 63 \\ 1 \text{ für } 63 \le x \end{cases}$$

$$\Rightarrow F(17) = 0 \;\; F(18) = F(19) = \frac{2}{5}$$

□

d)

$$V_h = \frac{\sum_{i=1}^{h} x_i}{\sum_{i=1}^{n} x_i} \quad U_k = \frac{k}{n}$$

k	1 (A)	2 (D)	3 (B)	4 (E)	5 (C)	Σ
xk	18	18	24	27	63	150
uk	1/5	2/5	3/5	4/5	1	
pk	3/25	3/25	4/25	9/50	21/50	1
vk	3/25	6/25	2/5	29/50	1	

Knickpunkte: (0;0); (2/5;6/25); (3/5;2/5); (4/5;29/50); (1;1)

e)

$$G = \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_i - (n+1)\sum_{i=1}^{n} x_i}{n \cdot \sum_{i=1}^{n} x_i}$$

$$G = \frac{2 \cdot (1 \cdot 18 + 2 \cdot 18 + 3 \cdot 24 + 4 \cdot 27 + 5 \cdot 63) - 6 \cdot 150}{5 \cdot 150} = 0{,}264$$

$$G_* = \frac{n}{n-1} G$$

$$G_* = \frac{5}{4} \cdot 0{,}264 = \underline{\underline{0{,}33}}$$

□Aufgabe 8

Quelle: Bamberg Aufgabe 6

Ein kardinales Merkmal x wurde an n Merkmalsträgern erhoben. Die (zwischen 0 und 600 liegenden) Beobachtungswerte wurden anschließend klassiert. Die Ergebnisse hierzu sind in folgender Tabelle aufgelistet:

Klasse	[0;100)	[100;200)	[200;400)	[400;600)
h_j	10	12	18	10
$\bar{x}_j$	60	150	250	510
s_j^2	800	1.000	10.000	

Dabei bezeichnet $\bar{x}_j$ bzw. s_j^2 das arithmetische Mittel bzw. die mittlere quadratische Abweichung der zur j -ten Klasse gehörenden Beobachtungswerte.

a) Skizzieren Sie das Histogramm.
b) Bestimmen Sie die Knickpunkte der Lorenzkurve zu diesen klassierten Daten.
c) Die gesamte mittlere quadratische Abweichung aller n Beobachtungswerte beträgt $s_{Ges}^2 = 29.000$. Bestimmen Sie den oben fehlenden Wert s_4^2.

Lösung

a) Rechteckhöhe: 10, 12, 9, 5, mit c = 100

b)

u_j	0,2	0,44	0,8	1
$h_j \bar{x}_j$	600	1.800	4.500	5.100
v_j	0,05	0,2	0,575	1

Knickpunkte: (0,0); (0,2;0,05); (0,44;0,2); (0,8;0,575); (1;1)

c)

$$s_{ges}^2 = \frac{1}{n}\sum_{j=1}^{r} n_j s_j^2 + \frac{1}{n}\sum_{j=1}^{r} n_j (\bar{x}_j - \bar{x}_{ges})^2$$

$$29.000 = \frac{1}{50}(10 \cdot 800 + 12 \cdot 1000 + 18 \cdot 10.000 + 10 \cdot s_4^2) +$$
$$\frac{1}{50}(10 \cdot (60-240)^2 + 12(150-240)^2 +$$
$$18 \cdot (350-240)^2 + 10(510-240)^2)$$
$$= \frac{8.000}{50} + \frac{12.000}{50} + \frac{180.000}{50} + \frac{s_4^2}{5} + 23.040$$
$$s_4^2 = \underline{\underline{9.800}}$$

Aufgabe 9

Quelle: Bamberg Aufgabe 7

Im Landkreis Thölen ist Hundefutter der Marken Hundeglück, Kaiser, Schnappi, Snobbydog und Watzen erhältlich. Die Umsatzzahlen der Marken (in T€) für das Jahr 2005 können der nachfolgenden Tabelle entnommen werden.

Marke	Hundeglück	Kaiser	Schnappi	Snobbydog	Watzen
Umsatz	500	300	2.100	900	1.200

a) Bestimmen Sie die Knickpunkte der Lorenzkurve bezüglich des Umsatzanteils der Marken.
b) Berechnen Sie den normierten Gini-Koeffizienten sowie den Herfindahl-Index.

c) Berechnen Sie den Variationskoeffizienten.

Lösung

a)

k	1 (Kaiser)	2 (Hundeg)	3 (Snobby)	4 (Watzen)	5 (Schnap)	Σ
xk	300	500	900	1200	2100	5000
uk	0,2	0,4	0,6	0,8	1	
pk	0,06	0,1	0,18	0,24	0,42	1
vk	0,06	0,16	0,34	0,58	1	

$\Rightarrow$ Knickpunkte: (0;0) (0,2;0,06) (0,4;0,16) (0,6;0,34) (0,8;0,58);(1;1)

$$G = \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_i - (n+1)\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i}$$

b)

$$G = \frac{2 \cdot [1 \cdot 300 + 2 \cdot 500 + 3 \cdot 900 + 4 \cdot 1.200 + 5 \cdot 2.100] - 6 \cdot 5.000}{5 \cdot 5.000}$$

$$G = \underline{\underline{0{,}344}}$$

$$G_* = G \cdot \frac{n}{n-1}$$

$$G_* = 0{,}344 \cdot \frac{5}{4} = \underline{\underline{0{,}43}}$$

$$H = \sum_{i=1}^{n} p_i^2 \qquad ; \qquad \mathrm{p_i} = \frac{x_i}{\sum_{i=1}^{n} x_i}$$

$$H = 0{,}06^2 + 0{,}1^2 + 0{,}18^2 + 0{,}24^2 + 0{,}42^2 = \underline{\underline{0{,}28}}$$

c) $$\mathrm{H} = \frac{1}{n}(V^2+1) \quad 0{,}28 = \frac{1}{5}(V^2+1) \qquad \Rightarrow \mathrm{V} = \underline{\underline{0{,}63}}$$

Aufgabe 10

Quelle: Klausur WS 10/11

Gegeben sei ein Markt mit 6 Unternehmen, die im Jahr 2010 folgende Umsätze erzielten (EUR Mio.): 1, 1, 5, 8, 10,75.

Berechnen Sie (auf 2 Stellen nach dem Komma)

a) den Gini-Koeffizienten sowie

b) den normierten Gini-Koeffizienten und

interpretieren Sie das Ergebnis.

-

Lösung

a)

$$G = \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_i - (n+1)\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i}$$

$$G = \frac{2 \cdot [1 \cdot 1 + 2 \cdot 1 + 3 \cdot 5 + 4 \cdot 8 + 5 \cdot 10 + 6 \cdot 75] - 7 \cdot 100}{6 \cdot 100}$$

$$\mathrm{G} = \frac{2}{3}$$

b)
$$G_* = G \cdot \frac{n}{n-1}$$

$$G_* = \frac{2}{3} \cdot \frac{6}{5} = 0{,}8$$

c) Interpretation: Der Markt weist eine hohe Konzentration auf.

Aufgabe 11

Quelle: Klausur SS 12

Gegeben sei ein Markt mit 7 Unternehmen, die im Jahr 2011 folgende Umsätze erzielten (EUR Mio.): 2, 2, 2, 3, 3, 4, 4.

Berechnen Sie (auf 2 Stellen nach dem Komma)

a) den Gini-Koeffizienten sowie
b) den normierten Gini-Koeffizienten und
c) interpretieren Sie das Ergebnis.

Lösung

a)
$$G = \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_i - (n+1)\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i}$$

$$\mathrm{G} = \frac{2 \cdot [1 \cdot 2 + 2 \cdot 2 + 3 \cdot 2 + 4 \cdot 3 + 5 \cdot 3 + 6 \cdot 4 + 7 \cdot 4] - 8 \cdot 20}{7 \cdot 20}$$

$$\mathrm{G} = \frac{22}{140} = 0{,}16$$

b)
$$G_* = G \cdot \frac{n}{n-1}$$

$$\mathrm{G}_* = \frac{22}{140} \cdot \frac{7}{6} = 0{,}18$$

c) Interpretation: Der Markt weist eine niedrige Konzentration auf.

□

Aufgabe 12

Quelle: Klausur SS 15

Gegeben sei ein Markt mit 6 Unternehmen, die im Jahr 2014 folgende Umsätze erzielten (EUR Mio.): 5,5,5,5,5,5.

Berechnen Sie (auf 3 Stellen nach dem Komma)

a) den Gini-Koeffizienten sowie

b) den normierten Gini-Koeffizienten und

c) interpretieren Sie das Ergebnis.

Lösung

a)

$$G = \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_i - (n+1)\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i}$$

$$G = \frac{2 \cdot [1 \cdot 5 + 2 \cdot 5 + 3 \cdot 5 + 4 \cdot 5 + 5 \cdot 5 + 6 \cdot 5] - 7 \cdot 30}{6 \cdot 30}$$

$$G = 0$$

b)

$$G_* = G \cdot \frac{n}{n-1}$$

$$G_* = 0 \cdot \frac{6}{5} = 0$$

c) Interpretation: Der Markt weist keine Konzentration auf, da alle Merkmalsträger (= Unternehmen) die gleiche Ausprägung (= Umsatz) aufweisen.

□

Aufgabe 13

Ein Betrieb hat im Kalenderjahr 2005 zwölf neue Mitarbeiter eingestellt. Von diesen sind unter anderem folgende Daten bekannt:

Mitarbeiter Nr.	Geschlecht	Ausbildungsdauer (in Jahren)	Abschlussnote
1	männlich	9	4
2	weiblich	10	2
3	weiblich	10	4
4	männlich	11	4
5	weiblich	12	2
6	weiblich	13	2
7	weiblich	14	1
8	männlich	15	3
9	männlich	16	2
10	männlich	17	3
11	weiblich	19	3
12	männlich	22	2

a) Geben Sie die Skalierung der drei Merkmale Geschlecht, Ausbildungsdauer und Abschlussnote an.

b) Ermitteln Sie für jedes der drei Merkmale die folgenden Größen, soweit diese auf Grund des jeweiligen Skalenniveaus Sinnvollerweise berechnet werden können: Modus, Median, arithmetisches Mittel, mittlere quadratische Abweichung und Variationskoeffizient.

c) Betrachten Sie die neuen männlichen und weiblichen Mitarbeiter als disjunkte Gruppen und berechnen Sie die externe und interne Varianz der Ausbildungsdauer.

d) Geben Sie für jedes der zwei Merkmalspaare i) Geschlecht – Abschlussnote und ii) Ausbildungsdauer– Abschlussnote einen statistisch sinnvollen Korrelationskoeffizienten an. (Die Korrelationskoeffizienten müssen nicht berechnet werden.)

Lösung

a) Geschlecht: nominal; Ausbildungsdauer: kardinal; Abschlussnote: ordinal

b) Geschlecht: Modus nicht eindeutig: $x_{Mod} \in \{\text{weiblich; männlich}\}$

$x_{Med} = 13{,}5$ bei Ausbildungdauer

$x_{Med} = 2{,}5$ bei Abschlussnote

Ausbildungsdauer: $x_{Mod}=10$; $x_{Med} \in [13;14]$, $\bar{x} = 14$; $s^2 = 14{,}5$; V=0,27

Abschlussnote: $x_{Mod} = 2$; $x_{Med} = 2{,}5$

c)

$$\frac{1}{n}\sum_{j=1}^{r} n_j s_j^2 = \text{interne Varianz}$$

$$\frac{1}{12}(6\cdot 9,\overline{3} + 6\cdot 17,\overline{6}) = \underline{\underline{13,5}}$$

$$\frac{1}{n}\sum_{j=1}^{r} n_j(\overline{x}_j - \overline{x}_{ges})^2 = \text{ externe Varia}$$

$$\frac{1}{12}(6\cdot(13-14)^2 + 6\cdot(15-14)^2 = \underline{\underline{1}}$$

d) i) Geschlecht – Abschlussnote: Kontingenzkoeffizient; ii) Ausbildungsdauer – Abschlussnote: Kontingenzkoeffizient oder Rangkorrelationskoeffizient von Spearman

□Aufgabe 14

Quelle: Bamberg Aufgabe 9

In Alksburg, dem Austragungsort des ”Brückenfestes“, werden sieben verschiedene (nicht rezeptpflichtige) Präparate gegen Kopfschmerz angeboten. Jedes Präparat basiert auf genau einem der drei alternativen Wirkstoffe A, B oder C, wobei bekannt ist, dass die Substanz C wirksamer ist als die Substanz A und diese wiederum wirksamer als die Substanz B. Auf Grund der großen Nachfrage nach Kopfschmerzmitteln werden in Alksburg derartige Präparate grundsätzlich nur in Packungen zu 50 Tabletten (entspricht Packungsgröße N3) verkauft.

Die nachfolgende Tabelle fasst für jedes Präparat den enthaltenen Wirkstoff, den Verkaufspreis pro Packung sowie die im Jahr 2005 in Alksburg abgesetzte Anzahl Packungen zusammen:

Präparat	Wirkstoff	Preis (in € pro Packung)	Absatz (in Anzahl Packungen)
1	A	8	22.000
2	C	12	18.500
3	C	10	22.000
4	B	9	9.500
5	C	20	9.000
6	A	6	6.000
7	B	5	500

a) Geben Sie die Skalierung der drei Merkmale Wirksamkeit des Wirkstoffes, Preis und Absatz an.

b) Ermitteln Sie für jedes der drei Merkmale aus a) die folgenden Lageparameter, soweit diese auf Grund des jeweiligen Skalenniveaus sinnvollerweise berechnet werden können: Modus, Median und arithmetisches Mittel.

c) Welches Maß ist geeignet zur Quantifizierung des Zusammenhangs zwischen den Merkmalen Preis und Absatz? (Keine Berechnung erforderlich!)

d) Auf Grund der Vermutung, dass hohe Absatzwerte vorwiegend bei niedrigen Preisen auftreten
(und umgekehrt), könnte man eine negative Korrelation dieser beiden Merkmale erwarten. Dennoch hat die Berechnung eines geeigneten, auf [-1;1] normierten Zusammenhangsmaßes den positiven Wert 0;186 ergeben. Erläutern Sie auf Grund obiger Daten (ohne Rechnung!) das Zustandekommen dieses vermeintlichen Widerspruchs.

e) Bestimmen Sie den Wert des normierten Gini-Koeffizienten für den Umsatz im Jahr 2005.

Lösung

a) Wirksamkeit: ordinal; Preis: kardinal; Absatz: kardinal

b) Wirkstoff: $x_{\text{mod}} = C$, $x_{med} = A$

Preis: x_{mod}: jeder Preis; $x_{med} = \underline{9}$; $\bar{x} = \underline{\underline{10}}$

Absatz: $x_{\text{mod}} = \underline{22.000}$; $x_{med} = \underline{9.500}$; $\bar{x} = \underline{12.500}$

c) Sinnvoll: Bravais-Pearson-Korrelationskoeffizient

d) Es liegt eine Scheinkorrelation vor. Die intervenierende Variable Wirkstoff bewirkt, dass Preis und Absatz positiv korreliert sind: Ein wirksameres Medikament wird trotz des höheren Preises stärker nachgefragt als ein weniger wirksames Medikament.

□e) Zuerst wird für jedes Präparat der Umsatz (in T€) ermittelt und die Umsatzzahlen aufsteigend sortiert:

Preis	Absatz	Umsatz (= x_i)	i
8	22.000	176	4
12	18.500	222	7
10	22.000	220	6
9	9.500	85,5	3
20	9.000	180	5
6	600	36	2
5	500	2,5	1
	$\sum$	922	

Einsetzen in (17) i.V.m. (16) liefert:

$$G = \frac{2 \cdot \sum_{i=1}^{n} i \cdot x_i - (n+1)\sum_{i=1}^{n} x_i}{n \cdot \sum_{i=1}^{n} x_i}$$

$$G = \frac{2 \cdot [2{,}5 + 2 \cdot 36 + 3 \cdot 85{,}5 + 4 \cdot 176 + 5 \cdot 180 + 6 \cdot 220 + 7 \cdot 222] - 8 \cdot 922}{7 \cdot 922}$$

$$= 0{,}3474$$

$$G_* = \frac{n}{n-1} G \qquad G_* = \frac{7}{6} \cdot 0{,}3474 = \underline{\underline{0{,}405}}$$

□Aufgabe 15

Quelle: Klausur SS 19

In einer empirischen Untersuchung wurden 170 Frauen nach ihrem eigenen sowie nach dem Beruf ihres Vaters gefragt:

		Beruf des Vaters			
		Arbeiter	Angestellter	Beamter	Selbständiger
Beruf der Tochter	Arbeiterin	20	15	3	7
	Angestellte	15	25	15	5
	Beamtin	5	3	8	4
	Selbständige	10	17	4	14

Bitte prüfen Sie, ob der Beruf der Tochter von dem des Vaters unabhängig ist.

Lösung

Zwei Merkmale X und Y sind unabhängig, wenn für <u>alle</u> i, j (= alle Kombinationen der Ausprägungen von X und Y) gilt:

$$= =$$

Zuerst ermittelt man die Randhäufigkeiten:

		Beruf des Vaters				
		Arbeiter	Angestellter	Beamter	Selbständiger	
Beruf der Tochter	Arbeiterin	20	15	3	7	45
	Angestellte	15	25	15	5	60
	Beamtin	5	3	8	4	20
	Selbständige	10	17	4	14	45
		50	60	30	30	170

$= 20 \neq 13{,}24 = 45x50/170 =$

<u>Ergebnis</u>: Beruf von Vater und Tochter sind nicht unabhängig voneinander.

Aufgabe 16

In einem Fußballturnier mit 16 teilnehmenden Mannschaften wurden für die Vorrunde vier Gruppen, bestehend aus je vier Mannschaften, gebildet. Innerhalb der einzelnen Gruppen spielte "jeder gegen jeden", so dass pro Gruppe sechs und damit insgesamt 24 Spiele in der Vorrunde stattfanden. Im Folgenden ist für jedes dieser (von 1 bis 24 durchnummerierten Spiele) die dabei erzielte Anzahl von Toren registriert.

Spielnummer	1	2	3	4	5	6	7	8	9	10	11	12
Anzahl Tore	0	2	0	1	1	0	2	5	1	1	0	2

Spielnummer	13	14	15	16	17	18	19	20	21	22	23	24
Anzahl Tore	1	2	0	10	1	5	1	0	1	2	2	0

(1) Bestimmen Sie zum Merkmal X = Anzahl der pro Spiel erzielten Tore

a) die Häufigkeitstabelle (mit den absoluten Häufigkeiten),

b) zu jeder auftretenden Ausprägung a_j den Winkel w_j des Kreissektorendiagramms,

c) Modus, Median und arithmetisches Mittel,

d) alle Knickpunkte der Lorenzkurve,

e) den Konzentrationskoeffizienten CR_2 .

(2) Nun sei zusätzlich bekannt, dass unter obigen durchnummerierten Spielen die Spiele mit ungerader
Nummer 1,3,5,…..,23 jeweils nachmittags, die mit gerader Nummer 2,4,6,…..,24 hingegen jeweils abends stattfanden. Berechnen Sie eine geeignete (nicht notwendigerweise normierte) Maßzahl für den Zusammenhang zwischen den Merkmalen

X = Anzahl der pro Spiel erzielten Tore, und

Y = Spieltermin (mit den Ausprägungen nachmittags bzw. abends).

Lösung

(1) a) und b) Mit $w_j = \frac{360}{24} \cdot h(a_j) = 15 \cdot h(a_j)$ folgt:

a_j	0	1	2	5	10
$h(a_j)$	7	8	6	2	1
w_j	105	120	90	30	15

c) $x_{\text{mod}} = 1;\ x_{med} = 1;\ \bar{x} = \frac{1}{24}(0 \cdot 7 + ... + 1 \cdot 10) = \frac{40}{24} = \underline{\underline{1{,}67}}$

d)

x_j	h_i	u_k	v_k
0	7	$7/24$	$0/40$
1	8	$15/24$	$8/40$
2	6	$21/24$	$20/40$
5	2	$23/24$	$30/40$
10	1	1	1

Knickpunkte: (0,0); $(\frac{7}{24},0)$; $(\frac{15}{24},\frac{8}{40})$; $(\frac{21}{24},\frac{20}{40})$; $(\frac{23}{24},\frac{30}{40})$; (1,1)

e)

$$CR_2 = \frac{10+5}{40} = \frac{15}{40} = \underline{\underline{0{,}375}}$$

(2)

Hinweis: Tageszeit ist nominal skaliert, sodass als Maßzahl der Kontingenzkoeffizient zur Anwendung kommt.

h_{ij}	0	1	2	5	10	
nachm.	4	6	2	0	0	12
abends	3	2	4	2	1	12
	7	8	6	2	1	24

$\tilde{h}_{ij}$	0	1	2	5	10
nachm.	3,5	4	3	1	0,5
abends	3,5	4	3	1	0,5

$$\chi^2 = \left(\frac{0{,}5^2}{3{,}5} + \frac{2^2}{4} + \frac{1^2}{3} + \frac{1^2}{1} + \frac{0{,}5^2}{0{,}5}\right) \cdot 2 = 5{,}81 \Rightarrow K = \sqrt{\frac{5{,}81}{29{,}81}} = \underline{\underline{0{,}44}}$$

□Aufgabe 17

Quelle: Bamberg Aufgabe 13

Von einer Firma sind über mehrere Jahre hinweg die Umsätze und die Beschäftigtenzahlen bekannt:

Jahr	1	2	3	4	5	6
Umsatz (in Millionen €)	60	55	57	61	65	62
Anzahl der Beschäftigten	1.000	1.100	960	840	800	700

a) Berechnen Sie den Variationskoeffizienten des Umsatzes.

b) Berechnen Sie den Rangkorrelationskoeffizienten von Spearman zwischen den beiden Merkmalen Umsatz und Beschäftigtenzahl.

c) Berechnen Sie die Regressionsgerade der Beschäftigtenzahl in Abhängigkeit von der Zeit. Mit welcher Anzahl der Beschäftigten ist im Jahr 8 zu rechnen?

Lösung

a)

$$V = \frac{s}{\bar{\bar{x}}}$$

$$V = \frac{3{,}265986}{60} = \underline{\underline{0{,}0544}}$$

b)

$$r_{SP} = 1 - \frac{6\sum_{i=1}^{n}(R_i - R_i^{'})^2}{(n-1)\cdot n\cdot(n+1)} \quad \Rightarrow \quad r_{SP} = 1 - \frac{6\cdot 66}{5\cdot 6\cdot 7} \quad \Rightarrow \quad \underline{\underline{-0{,}8857}}$$

	Jahr	1	2	3	4	5	6	
R_i	*Umsatz*	4	6	5	3	1	2	
$R_i^{'}$	*Beschäftigte*	2	1	3	4	5	6	
$R_i - R_i^{'}$		2	5	2	-1	-4	-4	
$(R_i - R_i^{'})^2$		4	25	4	1	16	16	$\sum 66$

c) Nach Übungsaufgabe 11: $\hat{b} = -72;\ \hat{a} = 1.152 \Rightarrow \hat{y} = 1.152 - 72t;\ \hat{y}(8) = \underline{\underline{576}}$

□Aufgabe 18

Quelle: Klausur SS 10

Bei einer Mathematik-bzw. Physik-Arbeit haben 10 Schüler folgende Noten erhalten:

Schüler	1	2	3	4	5	6	7	8	9	10
Mathematik	1,0	2,3	5,0	1,3	2,0	4,0	3,0	2,7	3,7	1,7
Physik	1,3	3,0	3,7	1,7	2,0	4,3	2,7	3,3	4,0	1,0

a) Bitte berechnen Sie die Korrelation zwischen den beiden Noten,

b) begründen den von Ihnen gewählten Koeffizienten und

c) interpretieren Sie Ihr Ergebnis.

Lösung

a)

Schüler	1	2	3	4	5	6	7	8	9	10
Mathematik	1,0	2,3	5,0	1,3	2,0	4,0	3,0	2,7	3,7	1,7
Physik	1,3	3,0	3,7	1,7	2,0	4,3	2,7	3,3	4,0	1,0

R_i	1	5	10	2	4	9	7	6	8	3
R_i'	2	6	8	3	4	10	5	7	9	1
(Ri - Ri')	-1	-1	2	-1	0	-1	2	-1	-1	2
$(Ri - Ri')^2$	1	1	4	1	0	1	4	1	1	4

rsp = 0,8909091

b) Rangkorrelationskoeffizient von Spearman, da ein ordinal skaliertes Merkmal vorliegt.

c) Rsp = 0,89 deutet auf eine hohe Korrelation/Interdependenz hin: „Wer Mathe kann, kann auch Physik“.

Aufgabe 19

Quelle: Klausur WS 11/12

Bei einer Deutsch- und einer Chemie-Arbeit haben 10 Schüler folgende Noten erhalten:

Schüler	1	2	3	4	5	6	7	8	9	10
Deutsch	1,0	2,7	5,0	1,3	2,0	3,0	4,0	1,7	2,3	3,7
Chemie	3,0	1,3	3,3	1,7	4,3	1,0	2,7	2,3	4,0	2,0

a) Bitte berechnen Sie die Korrelation zwischen den beiden Noten,

b) begründen den von Ihnen gewählten Koeffizienten und

c) interpretieren Sie Ihr Ergebnis.

Lösung

a)

Schüler	1	2	3	4	5	6	7	8	9	10
Deutsch	1,0	2,7	5,0	1,3	2,0	3,0	4,0	1,7	2,3	3,7
Chemie	3,0	1,3	3,3	1,7	4,3	1,0	2,7	2,3	4,0	2,0

Rang	R_i	1	6	10	2	4	7	9	3	5	8
	R_i'	7	2	8	3	10	1	6	5	9	4
	(Ri - Ri')	-6	4	2	-1	-6	6	3	-2	-4	4
	$(Ri - Ri')^2$	36	16	4	1	36	36	9	4	16	16

Summe $(Ri - Ri')^2$ 174

a) rsp = -0,0545

-

-

b) Rangkorrelationskoeffizient von Spearman, da ein ordinal skaliertes Merkmal vorliegt.

c) Rsp = -0,0545 deutet auf (praktisch) keine Korrelation/Interdependenz hin: „Talente in Deutsch und Chemie haben nichts miteinander zu tun.“

Aufgabe 20

Quelle: Klausur SS 09

Ein Unternehmen, hat bei der Herstellung eines Produktes folgende Produktionsmengen und Gesamtkosten registriert:

Periode i	1	2	3	4	5	6	7	8
Menge (Stück)	800	750	600	720	850	630	790	900
Kosten (EUR)	9500	9000	8300	8800	10000	8500	9800	11500

a) Berechnen Sie die Regressionsgrade und interpretieren Sie diese.

b) Bestimmen Sie den Prognosewert der Kosten für die Produktionsmenge von 1.000 Stücken.

Bitte rechnen Sie mit zwei Stellen nach dem Komma.

Lösung

Periode i	Stück			Kosten		
		x-x̄	(x - x̄)2		(y - ȳ)	(x - x̄) * (y - ȳ)
1	800	45	2025	9500	75	3375
2	750	-5	25	9000	-425	2125
3	600	-155	24025	8300	-1125	174375
4	720	-35	1225	8800	-625	21875
5	850	95	9025	10000	575	54625
6	630	-125	15625	8500	-925	115625
7	790	35	1225	9800	375	13125
8	900	145	21025	11500	2075	300875
	755		74200	9425		686000
	x̄			ȳ		

b =	9,25		
â =	2444,81	Summe (x*y)	Summe x²
		57.613.000	4.634.400

a) Regressionsgrade: EUR 2.444,81 + 9,25x. Die Produktion einer zusätzlichen Einheit kostet EUR 9,25.

b) Der Prognosewert für 1.000 Stücke liegt bei EUR 11.690.

Aufgabe 21

Quelle: Klausur WS 09/10

Ein Unternehmen möchte den Zusammenhang zwischen Jahresumsatz und Verkaufsfläche in den Filialen untersuchen. Hierbei wurden für das Jahr 2009 folgende Werte in den 11 Filialen ermittelt:

Filiale i	Verkaufsfläche (in tsd. qm)	Jahresumsatz (in EUR Mio.)
1	0,32	2,96
2	0,94	5,99
3	1,1	6,75
4	1,26	7,15
5	1,05	6,3
6	1,46	8,84
7	0,54	3,54
8	1,29	7,26
9	0,46	3,45
10	1,37	5,95
11	0,85	4,69

Berechnen Sie (mit 3 Stellen nach dem Komma) die Regressionsfunktion und interpretieren Sie diese ökonomisch.

Lösung

xi = Fläche	yi = Umsatz				
xi	yi	x-x̄	(y - ȳ)	(x - x̄) * (y - ȳ)	(x - x̄)2
0,320	2,960	-0,647	-2,756	1,784	0,419
0,940	5,990	-0,027	0,274	-0,007	0,001
1,100	6,750	0,133	1,034	0,137	0,018
1,260	7,150	0,293	1,434	0,420	0,086
1,050	6,300	0,083	0,584	0,048	0,007
1,460	8,840	0,493	3,124	1,539	0,243
0,540	3,540	-0,427	-2,176	0,930	0,183
1,290	7,260	0,323	1,544	0,498	0,104
0,460	3,450	-0,507	-2,266	1,150	0,257
1,370	5,950	0,403	0,234	0,094	0,162
0,850	4,690	-0,117	-1,026	0,120	0,014
arith. Mittel (xi)	arith. Mittel (yi)			Summe (x - x̄) * (y - ȳ)	Summe (x - x̄)2
0,967	5,716			6,7130909	1,4926182
b =		4,498		Summe (x*y)	Summe x²
a =		1,366		67,5352	11,7844

Regressionsfunktion: y = 1,366 + 4,498x. Interpretation: Es ist zu erwarten, dass der Jahresumsatz sich pro 1.000 qm Verkaufsfläche um EUR 4,498 Mio. verändert.

□Aufgabe 22

Quelle: Klausur WS 11/12

Ein Unternehmen, hat bei der Herstellung eines Produktes in den vergangenen 8 Monaten folgende Produktionsmengen und Gesamtkosten registriert:

Monat	Jan	Feb	Mrz	Apr	Mai	Jun	Jul	Aug
Menge (Stück)	600	700	600	740	950	640	810	1000
Kosten (EUR)	9000	9500	8900	8800	9800	8400	9600	12000

a) Berechnen Sie die Regressionsgrade und interpretieren Sie diese.

b) Bestimmen Sie den Prognosewert der Kosten für die Produktionsmenge von 1.200 Stücken.

Bitte rechnen Sie mit zwei Stellen nach dem Komma.

Lösung

Periode i	Stück			Kosten				
		$x-\bar{x}$	$(x - \bar{x})^2$		$(y - \bar{y})$	$(x - \bar{x}) * (y - \bar{y})$	$x*y$	x^2
1	600	-155	24025	9000	-500	77500	5400000	360000
2	700	-55	3025	9500	0	0	6650000	490000
3	600	-155	24025	8900	-600	93000	5340000	360000
4	740	-15	225	8800	-700	10500	6512000	547600
5	950	195	38025	9800	300	58500	9310000	902500
6	640	-115	13225	8400	-1100	126500	5376000	409600
7	810	55	3025	9600	100	5500	7776000	656100
8	1000	245	60025	12000	2500	612500	12000000	1000000
	755		165600	9500		984000	58364000	4725800
	$\bar{x}$			$\bar{y}$				

a) Regressionsgrade: <u>EUR 5.013,77 + 5,94 x</u>.

Interpretation: <u>Die Produktion einer zusätzlichen Einheit kostet EUR 5,94</u>.

b) Der Prognosewert für 1.200 Stücke liegt bei <u>EUR 12.144,20</u>.

□Aufgabe 23

Quelle: Klausur SS 12

Ein Konsumgüterhersteller möchte wissen, welcher Zusammenhang zwischen Preis und Menge für eines seiner Güter besteht. Hierzu wurden in April 2012 in acht, einander vergleichbaren Testmärkten folgende Preis-Mengen-Kombinationen beobachtet:

Testmarkt	1	2	3	4	5	6	7	8
Preis (EUR)	50	48	47	54	55	51	58	45
Menge (Stück)	900	970	970	600	500	830	350	1000

a) Berechnen Sie die Regressionsgrade und interpretieren Sie diese.
b) Bestimmen Sie den Prognosewert für die Absatzmenge in einem Testmarkt bei einem Stückpreis von EUR 53.

Bitte rechnen Sie ohne Nachkommastellen.

Lösung

Testmarkt	Preis	Menge	x*y	x^2
	x	Y		
1	50	900	45000	2500
2	48	970	46560	2304
3	47	970	45590	2209
4	54	600	32400	2916
5	55	500	27500	3025
6	51	830	42330	2601
7	58	350	20300	3364
8	45	1000	45000	2025
	$\bar{x}$	$\bar{y}$	Summe	Summe
	51	765	304680	20944

a) Regressionsgrade: <u>3570 – 55x</u>

b) Die maximale Absatzmenge (bei einem Preis von Null) beträt 3570; pro EUR Preiserhöhung sinkt die abgesetzte Menge um 55 Stück.

c) Der Prognosewert für einen Preis von EUR 53 liegt bei <u>655</u> Stücken.

□Aufgabe 24

Quelle: Bamberg Aufgabe 14

Herr Müller ist Geschäftsführer bei der Alpenhornbahn. Er notiert für jeden Monat in der Saison die mittlere Anzahl der (pro Tag) verkauften Tageskarten für den Kinderlift und die mittlere Schneehöhe (in cm) bei der Bergstation. Seine Notizen der letzten beiden Jahre sind in der folgenden Tabelle wiedergegeben:

Schneehöhe	20	50	100	110	70	30	70	70	100	80

Tageskarten	100 120 180 100 80 30 70 60 100 60

Herr Müller vermutet, dass die Anzahl der verkauften Tageskarten von der Schneehöhe abhängt, und beauftragt Sie, eine entsprechende lineare Regression durchzuführen.

a) Ermitteln Sie die Koeffizienten der entsprechenden Regressionsgeraden.
b) Ermitteln Sie den Determinationskoeffizienten der Regressionsgeraden.
c) Herr Müller geht in diesem Monat von einer mittleren Schneehöhe von 65 cm aus. Prognostizieren Sie die mittlere Anzahl der verkauften Tageskarten auf der Basis der linearen Regression aus Teilaufgabe a).

d) Herr Müller fragt Sie, ob Sie den in c) errechneten Prognosewert für verlässlich halten. Beantworten Sie diese Frage (kurze Begründung).

Lösung

a)

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

$$\hat{b} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\bar{x} = 70$$

$$\bar{y} = 90$$

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 4.400$$

$$\hat{b} = \frac{4.400}{8.000} = \underline{\underline{0,55}}$$

$$\hat{a} = 90 - 0,55 \cdot 70 = \underline{\underline{51,5}}$$

b)

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y} - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

$$R^2 = \frac{2.420}{15.200} = \underline{\underline{0,16}}$$

c) $\hat{y}(65) = 51,5 + 0,55 \cdot 65 = 87,25 \Rightarrow 87$ Tageskarten

d) Die Schneehöhe ist nicht unbedingt der einzige Einflussfaktor. Ferientage (Saison) können beispielsweise einen erheblichen Einfluss auf den Umsatz ausüben. Da R² nur 0;16 beträgt, werden nur 16% der Varianz des Merkmals erklärt. Herr Müller kann sich somit nicht auf die

Prognose verlassen.

□Aufgabe 25

Quelle: Bamberg Aufgabe 15

Zu verschiedenen Zeitpunkten wird der Wasserstand x_t (in cm) der Regnitz gemessen.

t	1	2	3	4	5	6
x_t	100	110	?	?	125	120

Die Messwerte zum Zeitpunkt 3 und 4 sind leider verloren gegangen. Es ist jedoch folgendes bekannt:

$$\sum_{t=1}^{6} x_t = 705 \quad \text{und} \quad \sum_{t=1}^{6} x_t^2 = 83.425$$

Zudem ist bekannt, dass der Messwert zum Zeitpunkt 4 größer ist als der Messwert zum Zeitpunkt 3.

a) Ermitteln Sie die fehlenden Messwerte zum Zeitpunkt 3 und 4.

Falls Sie Teilaufgabe a) nicht lösen können, gehen Sie von einem gemessenen Wasserstand von 100 cm zu beiden Zeitpunkten aus.

b) Prognostizieren Sie auf der Basis der vorliegenden Daten den Wasserstand zum Zeitpunkt t =7 mittels einer linearen Regression.

c) Ermitteln Sie den Determinationskoeffizienten der Regression.

d) Begründen Sie kurz, ob Sie das Vorgehen aus Teilaufgabe b) für sinnvoll erachten.

Lösung

a)
$$x_3 + x_4 = 250$$
$$x_3^2 + x_4^2 = 31.300$$
$$x_3 = 250 - x_4$$
$$(250 - x_4)^2 + x_4^2 = 31.300$$
$$x_4^2 - 500x_4 + 62.500 + x_4^2 = 31.300 \qquad | - 31.300$$
$$2x_4^2 - 500x_4 + 31.200 = 0 \qquad | : 2$$
$$x_4^2 - 250x_4 + 15.600 = 0$$

pq-Formel: $$x_{1,2} = -\frac{p}{2} \pm \sqrt{(\frac{p}{2})^2 - q}$$

$$x_{1,2} = 125 \pm \sqrt{125^2 - 15.600}$$
$$x_{1,2} = 125 \pm \sqrt{25}$$

$x_3 = \underline{\underline{120}}$

$x_4 = \underline{\underline{130}}$

b) $\hat{a} = 102$; $\hat{b} = \frac{31}{7} = 4{,}428571$; $\hat{x}_7 = \frac{31}{7} \cdot 7 + 102 = \underline{\underline{133cm}}$

c) $R^2 = \underline{0{,}584194528}$

d) Der Wasserstand hängt nicht vom Zeitpunkt, sondern von der Niederschlagsmenge ab. Das Vorgehen ist nicht sinnvoll!

Aufgabe 26

Quelle: Bamberg Aufgabe 17

Das Restaurant ”Teuro-Grill“ hat zum Jahreswechsel 2001/2002 alle Preise 1:1 von DM in € umgestellt (obwohl 1 € = 1,95583 DM). Die Speisekarte umfasst 23 Gerichte und 17 Getränke.

a) Geben Sie, falls möglich, folgende Preisindizes für den die 40 Posten der Speisekarte umfassenden Warenkorb zur Berichtsperiode 2002 und zur Basisperiode 2001 an: Laspeyres-Index und Paasche-Index.

Der Student Bernd Breit isst in den Jahren 2001 bis 2003 jeweils 1 Schnitzel pro Woche im ”Teuro-Grill“. Anschließend geht er jeweils in die benachbarte” Trinkbar“, um eine unterschiedliche Anzahl von Bieren zu konsumieren. Folgende Tabelle gibt die Preise und Mengen für die Jahre 2001 bis 2003 wieder:

		2001	2002	2003
Schnitzel	Preis	12 DM	12 Euro	12 Euro
	Menge	1	1	1
Bier	Preis	4,20 DM	2,15 Euro	1,95 Euro
	Menge	3	4	n

Den Indexberechnungen der folgenden Teile b) und c) liegt genau B. Breits Warenkorb der Schnitzel und Biere aus dem Teuro-Grill und der Trinkbar zu Grunde.

b) Berechnen Sie den Paasche- und den Laspeyres-Preisindex von B. Breit zur Basisperiode 2001 und Berichtsperiode 2002.

c) B. beschließt, dass sein Preisindex zur Basisperiode 2002 und zur Berichtsperiode 2003 höchstens 0,95 betragen darf. Welchen von den beiden Indizes von Paasche bzw. Laspeyres sollte er verwenden? Bestimmen Sie die Anzahl n von Bieren, die B. wöchentlich mindestens trinken muss, um sein Ziel zu erreichen.

Lösung

a) Beide Indizes haben den Wert 1,95583.

b)

$$p^P_{2001,2002} = \frac{(1\cdot 12 + 4\cdot 2{,}15)\cdot 1{,}95583}{1\cdot 12 + 4\cdot 4{,}20} = \underline{\underline{1{,}399}}; \qquad p^L_{2001,2002} = \frac{(1\cdot 12 + 3\cdot 2{,}15)\cdot 1{,}95583}{1\cdot 12 + 3\cdot 4{,}20} = \underline{\underline{1{,}467}}$$

c) Paasche, da nur hier n eingeht.

$$0{,}95 \geq p^P_{2002,2003} = \frac{1\cdot 12 + n\cdot 1{,}95}{1\cdot 12 + n\cdot 2{,}15} \Leftrightarrow 11{,}4 + 2{,}0425n \geq 12 + 1{,}95n \Leftrightarrow n \geq \frac{0{,}6}{0{,}0925},$$

also mindestens 7 Biere

□Aufgabe 27

Quelle: Klausur SS 13

Ein Student hat festgestellt, dass sein Verbrauchsverhalten durch die Güter Miete, Brot, Mineralwasser und Kino gut repräsentiert sei. Für die Zeiträume 2010 bis 2012 hat er folgende monatliche Durchschnittspreise und -mengen aufgezeichnet:

	p10(i)	q10(i)	p11(i)	q11(i)	p12(i)	q12(i)
Miete	8 EUR/qm	20 qm	8,5	18	9	21
Brot	2,5 EUR/kg	6 kg	2,7	7	2,8	8
Mineralwasser	0,6 EUR/Liter	40 Liter	0,7	45	0,8	50
Kino	9 EUR/Karte	3 Karten	11	2	12	1

Berechnen Sie (auf 2 Stellen nach dem Komma)

a) die Preisveränderung basierend auf dem Preisindex nach Laspeyres für die Berichtsperiode 2011 gegenüber der Basisperiode 2010 und

b) die Preisveränderung basierend auf dem Preisindex nach Paasche für die Berichtsperiode 2012 gegenüber der Basisperiode 2011.

Lösung

$$\text{a)}\,p^L_{2010,2011} = \frac{8{,}5\cdot 20 + 2{,}7\cdot 6 + 0{,}7\cdot 40 + 11\cdot 3}{8\cdot 20 + 2{,}5\cdot 6 + 0{,}6\cdot 40 + 9\cdot 3} = \underline{\underline{1{,}0938}}$$

$\text{Preisänderung}: 9{,}38\%$

$$\text{b)}\,p^P_{2011,2012} = \frac{9\cdot 21 + 2{,}8\cdot 8 + 0{,}8\cdot 50 + 12\cdot 1}{8{,}5\cdot 21 + 2{,}7\cdot 8 + 0{,}7\cdot 50 + 11\cdot 1} = \underline{\underline{1{,}0703}}$$

$\text{Preisänderung}: 7{,}03\%$

Aufgabe 28

Quelle: Klausur SS 14

Ein Student hat festgestellt, dass sein Verbrauchsverhalten durch die Güter Miete, Brot, Mineralwasser und Kaffee gut repräsentiert sei. Für die Zeiträume 2011 bis 2013 hat er folgende monatliche Durchschnittspreise und -mengen aufgezeichnet:

	p11(i)	q11(i)	p12(i)	q12(i)	p13(i)	q13(i)
Miete	9 EUR/qm	22 qm	9,5	20	10,5	24
Brot	2,7 EUR/kg	8	2,8	9	2,9	7
Mineralwasser	0,7 EUR/Liter	45 Liter	0,8	50	0,9	60
Kaffee	0,5 EUR/Tasse	60 Tassen	0,6	65	0,7	60

Berechnen Sie (auf 2 Stellen nach dem Komma)
die Preisveränderung basierend auf dem Preisindex nach Laspeyres für die Berichtsperiode 2012 gegenüber der Basisperiode 2011 und
die Preisveränderung basierend auf dem Preisindex nach Paasche für die Berichtsperiode 2013 gegenüber der Basisperiode 2012.

Lösung

$$\text{a)}\,p^{L}_{2011,2012}=\frac{9,5\cdot 22+2,8\cdot 8+0,8\cdot 45+0,6\cdot 60}{9\cdot 22+2,7\cdot 8+0,7\cdot 45+0,5\cdot 60}=\underline{\underline{1,0793}}$$

$\text{Preisänderung}: 7,93\%$

$$\text{b)}\,p^{P}_{2012,2013}=\frac{10,5\cdot 24+2,9\cdot 7+0,9\cdot 60+0,7\cdot 60}{9,5\cdot 24+2,8\cdot 7+0,8\cdot 60+0,6\cdot 60}=\underline{\underline{1,1107}}$$

$\text{Preisänderung}: 11,07\%$

$$\text{a)}\,p^{L}_{2010,2011}=\frac{8,5\cdot 20+2,7\cdot 6+0,7\cdot 40+11\cdot 3}{8\cdot 20+2,5\cdot 6+0,6\cdot 40+9\cdot 3}=\underline{\underline{1,0938}}$$

$\text{Preisänderung}: 9,38\%$

$$\text{b)}\,p^{P}_{2011,2012}=\frac{9\cdot 21+2,8\cdot 8+0,8\cdot 50+12\cdot 1}{8,5\cdot 21+2,7\cdot 8+0,7\cdot 50+11\cdot 1}=\underline{\underline{1,0703}}$$

$\text{Preisänderung}: 7,03\%$