Klausur: **Statistik**

Jürgen Meisel

Zugelassene Hilfsmittel: nicht progr. Taschenrechner
Bearbeitungszeit: **60 Minuten**

Anmerkung zur Bearbeitung:
Die Klausur besteht aus insgesamt 6 Aufgaben. Sie müssen nur 5 davon bearbeiten. Bitte wählen Sie 5 Aufgaben aus und streichen Sie die nicht bearbeitete Aufgabe auf dem Aufgabenblatt.

Aufgabe 1: Mittelwerte und Streumaße

Aufgabe 2: Preisindizes und Korrelation

Aufgabe 3: Konzentrationsprozesse

Aufgabe 4: Lineare Regression

Aufgabe 5: Grundlagen Wahr'keit, Satz von Bayes und Binomialverteilung

Aufgabe 6: Normalverteilung

Teilbereich Statistik

1.) Mittelwerte und Streumaße

20	

Die Firma MediCare stellt medizinische Geräte her. In der folgenden Tabelle sind die Ausgaben für die Garantieleistungen in den vergangenen Jahren dargestellt:

Jahr	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Garantieleistung (in Tausend Euro)	18	15	16	12	11	13	11	9	4	8

a) Ermitteln Sie das arithmetische Mittel der Garantieleistung.

Lösung:
Mittelwert:

$$\mu = \frac{18+15+16+12+11+13+11+9+4+8}{10} = \frac{117}{10} = 11{,}7$$

b) Berechnen Sie die Varianz und die Standardabweichung der Garantieleistung.

Lösung:

Varianz: $$\sigma^2 = \frac{1.521}{10} - 11{,}7^2 = 15{,}21 \quad \rightarrow \quad \sigma = 3{,}9$$

c) Wie viele Werte liegen innerhalb des Bereichs $\bar{x} \pm s$?
Geben Sie dieses Ergebnis auch als Prozentanteil aller Werte an.

Anmerkung: s ist die Standardabweichung.

Lösung:
Im Intervall [7,8 ; 15,6] liegen insgesamt 7 Werte;
d.h. 70 % der Werte sind im σ-Intervall.

d) Bestimmen Sie nun noch den Median der Garantieleistung, die Quartile q_1 und q_3 und berechnen Sie den Abstand zwischen den Quartilen.

Lösung:

Median: $\overline{x_{0,5}} = \frac{1}{2}\left(x_5 + x_6\right) \rightarrow \overline{x_{0,5}} = \overline{x_M} = 11,5$

Quartile 1: $\overline{x_{0,25}} = x_{[n \cdot p]+1} \rightarrow \overline{x_{0,25}} = x_{[10 \cdot 0,25]+1} = x_3 = 9$

Quartile 3: $\overline{x_{0,75}} = x_{[n \cdot p]+1} \rightarrow \overline{x_{0,75}} = x_{[10 \cdot 0,75]+1} = x_8 = 15$

Der Quartilabstand beträgt 6.

2.) Preisindizes und Korrelation

20	

Familie Winzig leistet sich leistet sich folgende Produkte und Ausgaben.
Folgende Übersicht wurde erstellt:

	Preis 2008	Umsatz 2008	Preis 2012	Verbrauch 2012
Chips	1,00 € / Tüte	50,00 €	2,50 € / Tüte	70 Tüten
Erdnüsse	7,00 € / kg	42,00 €	6,00 € / kg	8 kg
Bier	2,50 € / ltr.	550,00 €	4,50 € / ltr.	200 ltr.

a) Berechnen Sie den Preisindex nach Laspeyres.

Lösung:

$$L_P = \frac{\sum p_{1i} \cdot q_{0i}}{\sum p_{0i} \cdot q_{0i}} \rightarrow \frac{2,5 \cdot 50 + 6,00 \cdot 6 + 4,50 \cdot 220}{50 + 42 + 550} = \frac{1.151}{642} = 1,7928$$

b) Wie groß ist der durchschnittliche jährliche Preisanstieg in %, wenn der Preisindex nach Laspeyres zugrunde gelegt wird?

Lösung: $g = \sqrt[4]{1,7928} = 1,1571 \rightarrow 15,71[\%]$

c) Ein Landwirt möchte feststellen, ob ein Zusammenhang zwischen Blütebeginn und Erntebeginn von hellen Süßkirschen besteht. Im Jahre 2012 machte er an 5 Bäumen folgende Beobachtungen:

Baum	**Blütebeginn**	**Erntebeginn**
Apfel	28.04.	12.06.
Birne	29.04.	07.07.
Mirabelle	01.05.	27.06.
Pfirsich	02.05.	03.07.
Zitrone	03.05.	10.07.

Berechnen Sie den geeigneten Korrelationskoeffizienten.

Lösung:

Baum	**Blütebeginn**	**Erntebeginn**	**Rang Blüte**	**Rang Ernte**	**Rang-Dif.**	**Rang-Qudr.**
Apfel	28.04.	12.06.	1	1	0	0
Birne	29.04.	07.07.	2	4	2	4
Mirabelle	01.05.	27.06.	3	2	1	1
Pfirsich	02.05.	03.07.	4	3	1	1
Zitrone	03.05.	10.07.	5	5	0	0

$$r_S = 1 - \frac{6\sum_{i=1}^{n}(d_i)^2}{n(n^2-1)} \quad \Rightarrow \quad r_S = 1 - \frac{6\cdot 6}{5\cdot 24} = 0{,}7$$

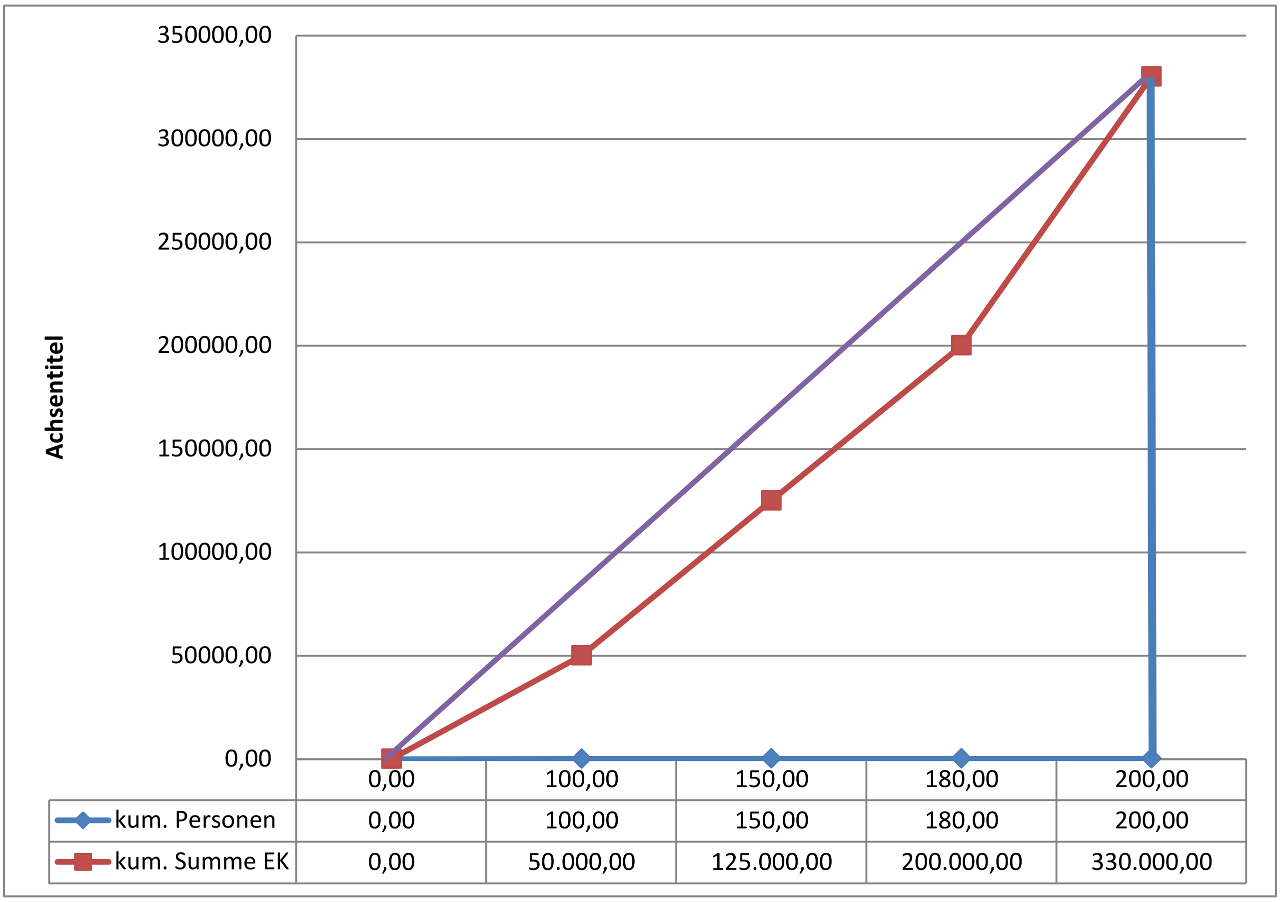
3.) Konzentrationsprozesse

20	

Gegeben sei folgende Einkommensverteilung:

Einkommensklasse	**Anzahl Personen**	**Kum. Personen**	**Klassen-mitte**	**Einkommens-summe**	**Kum. EK-Summe**
0 - 1.000	100	100	500	50.000	50.000
1.000 - 2.000	50	150	1.500	75.000	125.000
2.000 - 3.000	30	180	2.500	75.000	200.000
3.000 - 10.000	20	200	6.500	130.000	330.000

a) Bestimmen Sie die Lorenzkurve und den Gini-Koeffizient.



	0,00	100,00	150,00	180,00	200,00
kum. Personen	0,00	100,00	150,00	180,00	200,00
kum. Summe EK	0,00	50.000,00	125.000,00	200.000,00	330.000,00

$$Gini = \frac{\text{Konzentrationsfläche K}}{\text{maximale Konzentrationsfläche } \mathrm{K}_{\max}}$$

$$Gini = \frac{15.950.000}{33.000.000} = 0,48333$$

	A	B	C	D	E	F	G
1				Gini-Koeffizient mit relativen Anteilen			
2		Personen	kum. Personen	Klassenmitte EK	Summe EK	kum. Summe EK	Flächen
3			0,00			0,00	
4		100	100,00	500,00	50.000,00	50.000,00	2.500.000,00
5		50	150,00	1.500,00	75.000,00	125.000,00	4.375.000,00
6		30	180,00	2.500,00	75.000,00	200.000,00	4.875.000,00
7		20	200,00	6.500,00	130.000,00	330.000,00	5.300.000,00
8							
9							
10							
11							
12							
13							
14							
15			Berechnung mit absoluten Werten				
16		Breite	kum. Summe EK	Flächen			
17		100	50.000,00	2.500.000,00			
18		50	125.000,00	4.375.000,00			
19		30	200.000,00	4.875.000,00			
20		20	330.000,00	5.300.000,00			
21			Summe	17.050.000,00			
22							
23			Summe:	17.050.000,00		Fläche unterhalb der Lorenzkurve	
24							
25				33.000.000,00		Fläche unterhalb der Gleichverteilungsgeraden	
26							
27				15.950.000,00		Fläche zwischen Lorenzkurve und Gleichverteilungsgeraden	
28							
29			Gini:	0,483333333			

b) Ermitteln Sie den Anteil der 25 % reichsten Einkommensbezieher am Gesamteinkommen.

Lösung:

Die 25 % reichsten Einkommensbezieher vereinigen insgesamt 205.000,00 € auf sich.

Das bedeutet, dass 25 % der Einkommensbezieher 62,12 % des Gesamteinkommens erhalten.

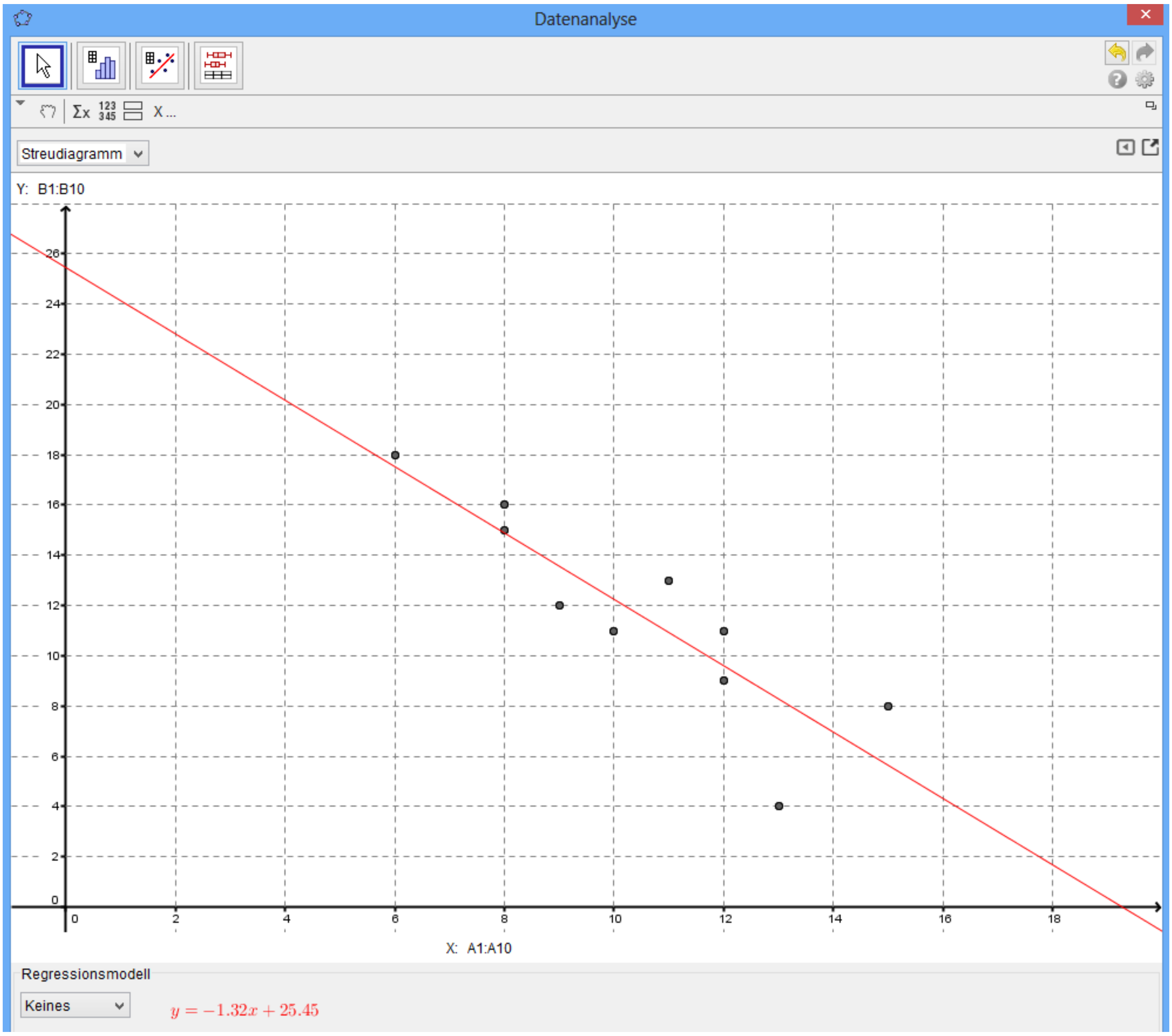
20	

4.) Lineare Regression und Korrelation

Die Firma MediCare möchte die Ausgaben für Qualität in Bezug zu den Ausgaben für Garantieleistungen untersuchen. Die notwendigen Daten sind der folgenden Tabelle zu entnehmen:

Jahr	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Summe
Ausgaben für Qualität	6	8	8	9	10	11	12	12	13	15	104
Garantieleistung (in Tausend Euro)	18	15	16	12	11	13	11	9	4	8	117
$x*y$	108	120	128	108	110	143	132	108	52	120	1.129
x^2	36	64	64	81	100	121	144	144	169	225	1.148

a) Zeichnen Sie die Datenpunkte in ein Koordinatensystem.



b) Ermitteln Sie die Regressionsgerade $y = b_0 + b_1 x$.

Ansatz: $y = b_0 + b_1 x$

$$b_1 = \frac{\sigma_{xy}}{\sigma_{x^2}} = \frac{\sum_{i=1}^{n}(x_i y_i) - n \cdot \mu_x \cdot \mu_y}{\sum_{i=1}^{n}(x_i^2) - n \cdot \mu_x^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i y_i) - \mu_x \cdot \mu_y}{\frac{1}{n}\sum_{i=1}^{n}(x_i^2) - \mu_x^2}$$

$$b_0 = \mu_y - b_1 \cdot \mu_x$$

$$\left.\begin{aligned} b_1 &= \frac{1.129 - 10 \cdot 10,4 \cdot 11,7}{1.148 - 10 \cdot 108,16} = \frac{-87,8}{66,4} = -1,32 \\ b_0 &= 11,7 - (-1,32) \cdot 10,4 = 25,43 \end{aligned}\right\} y = 25,43 - 1,32x$$

c) Welche Ausgaben für Qualität sind zu tätigen, damit die Garantieleistungen bei ca. 2.500,00 € liegen?

Lösung: $2,5 = 25,43 - 1,32x \rightarrow x = 17,37$

Es müssen ca. 17.370,00 € für die Qualität investiert werden, damit man eine Garantieleistung von 2.500,00 € erwarten darf.

d) Erörtern Sie zwei mögliche Konsequenzen der Geschäftsführung von MediCare aufgrund der unter b) berechneten Regressionsgeraden.

Lösung:

(1) Mehr Investition in die Produktionsfaktoren,

(2) Mehr Investition in Forschung und Entwicklung,

damit die Qualität der Produkte steigt.

(3) Prozessoptimierung

Teilbereich Stochastik

5.) Grundlagen Wahr'keit, Satz von Bayes und Binomialverteilung

20	

Ein Taxiunternehmer beschäftigt zwei Fahrer. Der Chef fährt 50 %, Fahrer A 30 % und Fahrer B 20 % der Fahrten. Bei 8 % der Fahrten verursacht der Chef, bei 10 % der Fahrer A und bei 12 % der Fahrer B einen Schaden.

a) Mit welcher (totalen) Wahrscheinlichkeit fahren die drei ihre Touren ohne Schäden?

Lösung:

$$P(schadensfrei) = 0,5 \cdot 0,92 + 0,3 \cdot 0,90 + 0,2 \cdot 0,88$$
$$P(schadensfrei) = 0,906 = 90,6\,\%$$

b) Es wurde an einem der Fahrzeuge ein Schaden festgestellt.
Mit welcher Wahrscheinlichkeit hat ihn Chef persönlich verursacht?

Lösung:

$$P_{Schaden}(Chef) = \frac{0,5 \cdot 0,08}{0,094} = 0,4255 = 42,55\,\%$$

c) Wie groß darf der Schadensprozentwert bei Fahrer B höchstens sein, wenn die Gesamtwahrscheinlichkeit für einen Schaden auf einer Tour auf höchstens 8 % sinken soll?

Lösung:

$$Ansatz: \quad 0,5 \cdot 0,08 + 0,3 \cdot 0,1 + 0,2 \cdot x = 0,08$$
$$\rightarrow \quad 0,5 \cdot 0,08 + 0,3 \cdot 0,1 + 0,2 \cdot x = 0,08$$
$$\rightarrow \quad 0,07 + 0,2 \cdot x = 0,08 \quad \rightarrow \quad 0,2 \cdot x = 0,01 \quad \rightarrow \quad x = 0,05 = 5\,\%$$

6.) Normalverteilung

20	

Bei Transporten der Firma EiEiEi werden im Normalfall 6 % der Eier beschädigt.
Ein Geschäft bekommt eine Lieferung von 1.500 Eiern.

a) Zeigen Sie, dass der Erwartungswert $\mu = 90$ und die Varianz $\sigma^2 = 84{,}6$ betragen.

Lösung:

$$\mu = 1.500 \cdot 0{,}06 = 90 \quad und$$
$$\sigma^2 = 1.500 \cdot 0{,}06 \cdot 0{,}94 = 84{,}6$$

Wie groß ist die Wahrscheinlichkeit, dass

b) mindestens 100 Eier beschädigt sind?

Lösung:

$$P(X \geq 100) = 1 - P(X < 100) = 1 - \Phi(1{,}09) = 1 - 0{,}862 = 0{,}138$$

$$Umrechnung: \quad z = \frac{100-90}{\sqrt{84{,}6}} = 1{,}09$$

c) höchstens 90 Eier beschädigt sind?

Lösung:

$$P(X \leq 90) = 1 - \Phi(1{,}09) = 0{,}5 \qquad mit\ Umrechnung: \quad z = \frac{90-90}{\sqrt{84{,}6}} = 0$$

d) Wie viele Eier dürfen höchstens beschädigt sein, wenn bei einer Lieferung mit einer Wahrscheinlichkeit von mind. 97 % insgesamt 1.450 Eier einwandfrei angeliefert werden sollen?

Lösung:

Erwartungswert und Varianz müssen nun von n Eiern berechnet werden, da die Gesamtmenge unbekannt ist:

$$\Phi(z) = 0{,}97 \xrightarrow{Tabelle} z = 1{,}89$$

$$\rightarrow 1{,}89 = \frac{1.450 - n \cdot 0{,}94}{\sqrt{n \cdot 0{,}94 \cdot 0{,}06}} \rightarrow n = 1.523{,}91 \approx 1.524$$

$$\rightarrow 1.524 - 1.450 = 74$$

Es dürfen von insgesamt 1.524 Eiern höchstens 74 Eier beschädigt sein, damit man mit einer Wahrscheinlichkeit von 97 % eine Lieferung von 1.450 unversehrten Eiern bekommt.