Univariate Datensätze

Aufgabe 1

Vor einer Bürgermeisterwahl, bei der fünf Kandidaten (A bis E) zur Auswahl stehen, wurden 160 Wahlberechtigte nach ihrer Wahlabsicht befragt. Die Befragung lieferte folgende Tabelle:

Kandidat	A	B	C	D	E
Stimmen	12	40	60	20	28

a) Bestimmen Sie die relativen Häufigkeiten und zeichnen Sie das dazugehörige Stabdiagramm.

b) Zeichnen Sie das dazugehörige Kreisdiagramm.

Aufgabe 2

Bei einer Untersuchung wurde die Körpergröße in cm von 20 Personen bestimmt. Dabei ergaben sich folgende Werte:

172	164	160	162	173	180	158	185	158	192
171	181	162	184	177	175	177	174	151	177

a) Stellen Sie die relativen Häufigkeiten in einem Histogramm dar. Teilen Sie dafür die Stichprobe in die fünf Klassen $[150;160); [160;170); [170;180); [180;190); [190;200)$ ein.

b) Stellen Sie die relativen Häufigkeiten in einem Histogramm dar. Teilen Sie dafür die Stichprobe in die vier Klassen $[140;160); [160;165); [165;190); [190;200)$ ein.

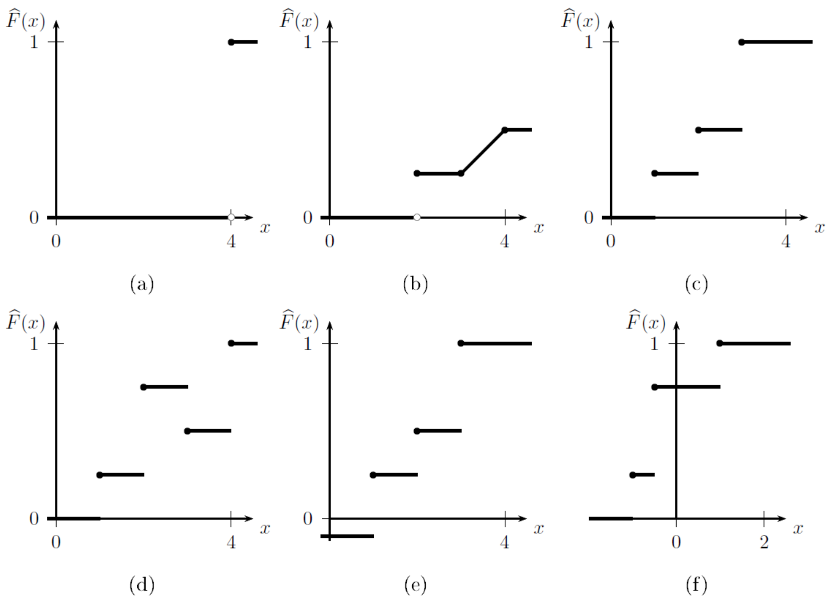
Aufgabe 3

Eine Stichprobe vom Umfang $n = 8$ lieferte folgende Ergebnisse:

$$26 \quad 23 \quad 16 \quad 32 \quad 23 \quad 18 \quad 27 \quad 18$$

a) Zeichnen Sie die empirische Verteilungsfunktion.

b) Bei welchen der folgenden Abbildungen handelt es sich nicht um eine empirische Verteilungsfunktion?



Beschreibung univariater Datensätze

Aufgabe 1

Gegeben seien die Merkmalsausprägungen x_i mit:

i	1	2	3	4	5	6	7	8	9
x_i	3,75	4,43	5,5	3,5	3	3	6,5	6,25	2,5

a) Bestimmen Sie den Modus, den Median und das arithmetische Mittel.

b) Bestimmen Sie die Varianz und die Standardabweichung.

c) Geben Sie an, ob die empirische Verteilung linkssteil, rechtssteil oder symmetrisch ist. Begründen Sie kurz Ihre Antwort.

d) Wie ändern sich der Modus, der Median, das arithmetische Mittel und die Varianz, wenn die x_i alle um 20% steigen?

e) Wie ändern sich der Modus, der Median und das arithmetische Mittel, wenn sich x_7 verdoppelt? Eine Tendenz genügt als Antwort.

f) Gegeben sei eine zweite Stichprobe vom Umfang $n = 10$ mit $\bar{y} = 50$ und $\Sigma(y_i - \bar{y})^2 = 250$. Welche Stichprobe weist die größere Streuung auf?

Aufgabe 2

Gegeben seien die Merkmalsausprägungen x_i mit:

i	1	2	3	4	5	6	7	8	9	10
x_i	5	2	4	9	8	4	2	6	4	8

Zusätzlich seien bekannt:

$$\widehat{M}_3 = \frac{1}{10}\sum_{i=1}^{10}(x_i - \bar{x})^3 = 2,856 \qquad \widehat{M}_4 = \frac{1}{10}\sum_{i=1}^{10}(x_i - \bar{x})^4 = 54,7792$$

a) Bestimmen Sie den Modus und den Median.

b) Bestimmen Sie die rechnerisch die Schiefe und die Wölbung.

Aufgabe 3

Gegeben seien nachfolgende Beobachtungswerte:

5,5	6	8	8,5	10,6	12,5	13,2
13,7	15	15,2	16,6	16,8	18,9	19,4

a) Bestimmen Sie das arithmetische Mittel $\bar{x}$ aus den Beobachtungswerten der Urliste.

b) Bilden Sie die Klassen über 5 bis 10, über 10 bis 15, über 15 bis 20 und berechnen Sie die Durchschnitte $\bar{x}_k$ pro Klasse und daraus erneut das arithmetische Mittel $\bar{x}$.

c) Bestimmen Sie mit Hilfe der Klassenmitten eine Näherung $\bar{x}'$ für das arithmetische Mittel und vergleichen Sie die bisher ermittelten Ergebnisse.

Aufgabe 4

Der Absatz eines Gutes ist in nachfolgender Tabelle gegeben. Berechnen Sie die mittlere jährliche Absatzsteigerung.

Jahr i	1	2	3	4	5	6	7	8	9	10	11
Absatz x_i	9	20	40	75	90	140	244	410	705	1036	1718

Aufgabe 5

Gegeben seien nachfolgende 17 Werte:

0 0 1 2 2 3 3 3 3 4 5 6 6 7 8 9 15

a) Zeichen Sie den dazugehörigen Box-Plot.

b) Treffen Sie eine Aussage zur Schiefe. Begründen Sie Ihre Aussage.

Beschreibung bivariater Datensätze: Abhängigkeitsmaße

Aufgabe 1

Bei einer Verkehrskontrolle wurden 10 Personen erfasst. Für die beiden Merkmale Alter und Alkoholgehalt ergab sich folgende Tabelle:

Alter	20	21	20	22	20	22	21	20	21	21
Promille	$0,2$	$0,1$	0	$0,2$	$0,1$	$0,1$	$0,2$	0	$0,1$	0

a) Stellen Sie in einer Kontingenztabelle die absoluten und relativen Häufigkeiten dar.

b) Bestimmen Sie den χ^2-Koeffizienten und den normierten Kontingenzkoeffizienten und interpretieren Sie Ihr Ergebnis.

c) Wie ändern sich die beiden Koeffizienten, wenn sich der Stichprobenumfang verdoppelt, aber die relativen Häufigkeiten gleich bleiben? Ändert eine Umordnung der Spalten und/oder Zeilen das Ergebnis?

Aufgabe 2

Gegeben sei nachfolgende Tabelle:

x_i	16	19	17	15	20	19	20
y_i	100	150	120	90	150	140	160

a) Zeichnen Sie das zugehörige Streudiagramm.

b) Bestimmen Sie den Korrelationskoeffizienten von Bravais-Pearson. Interpretieren Sie Ihr Ergebnis.

Aufgabe 3

Gegeben sei nachfolgende Tabelle:

x_i	-2	-1	0	1	2
y_i	4	1	0	1	4

a) Zeichnen Sie das zugehörige Streudiagramm.

b) Bestimmen Sie den Korrelationskoeffizient von Bravais-Pearson. Nehmen Sie zur Aussage, es bestehe kein Zusammenhang, Stellung.

Aufgabe 4

Es wurden sechs Abiturienten befragt, welche Note sie bei ihrer Prüfung erzielt haben ((g)ut, (b)efriedigend, (a)usreichend, (m)angelhaft). Zusätzlich wurden sie befragt, ob sie eine vorherige Probeklausur bestanden haben ((j)a, (n)ein). Dabei ergaben sich folgende Antworten:

(b, n)   (b, j)   (b, j)   (a, n)   (m, n)   (g, j)

Kann ein Zusammenhang zwischen der erzielten Note und dem Abschneiden bei der Probeklausur festgestellt werden?

Das lineare Regressionsmodell

Aufgabe 1

In einem Schwellenland wurde eine Studie zum Zusammenhang zwischen dem Einkommen der Eltern und dem Geburtsgewicht des Kindes durchgeführt. Dabei wurden das monatliche Einkommen x_i in 1000 GE und das Geburtsgewicht y_i in Pfund betrachtet:

x_i	2,7	1,9	3,1	3,9	4	3,4	2,1	2,9
y_i	5	6	9	8	7	6	7	8

a) Stellen Sie eine sinnvolle Regressionsbeziehung auf.

b) Zeichnen Sie das zugehörige Streudiagramm.

c) Bestimmen Sie die Regressionsgerade.

d) Das Einkommen einer Familie beträgt 3000 GE. Welches Gewicht wird dann prognostiziert?

e) Ist die gewählte Regression geeignet? Nutzen Sie das Bestimmtheitsmaß für Ihre Stellungnahme.

Aufgabe 2

Gegeben sei folgende Regressionsgerade:

$$\widehat{schlaf} = 3586,4 - 0,151 \cdot arbeit$$

Dabei ist *schlaf* die Zeit in Minuten, die pro Woche zum Schlafen genutzt wird. *arbeit* gibt die Minuten, die pro Woche gearbeitet werden, an.

a) Interpretieren Sie den Achsenabschnitt.

b) Angenommen *arbeit* steigt um 2 Stunden. Wie wirkt sich diese Erhöhung auf *schlaf* aus? Handelt es sich um einen großen Effekt?

Aufgabe 3

Gegeben sei folgende Tabelle:

P	11	9	12	13	14	16	17	30	28	28	42	55
Q	67	55	56	58	51	57	46	38	39	69	36	35

Dabei ist P der beobachtete Preis und Q die nachgefragte Menge eines Gutes in den USA zwischen den Jahren 1958 und 1969.

a) Bestimmen Sie die zugehörige Regressionsgerade und unterstellen Sie, dass P der Regressor ist. Interpretieren Sie außerdem die Parameter der Regressionsgerade.

b) Führen Sie eine log-log-Regression durch. Interpretieren Sie außerdem den Steigungsparameter der Regressionsgerade.

Zufallsvariablen und spezielle Verteilungen

Aufgabe 1

Entscheiden Sie für die nachfolgenden Funktionen, ob diese die Eigenschaften einer Wahrscheinlichkeits-/Dichte- bzw. Verteilungsfunktion besitzen.

$$g_1(x) = \begin{cases} \frac{x}{2} & \text{für } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

$$g_2(x) = e^x \quad \text{für } x \in \mathbb{R}$$

$$g_3(x) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{1}{9}x^2 & \text{für } 0 \leq x \leq 3 \\ 1 & \text{für } x > 3 \end{cases}$$

$$g_4(x) = \begin{cases} \frac{x}{28} & \text{für } x \in \{1,2,3,4,5,6,7\} \\ 0 & \text{sonst} \end{cases}$$

$$g_5(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

$$g_6(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ 1 & \text{für } x > 0 \end{cases}$$

Aufgabe 2

Gegeben sei eine Zufallsvariable X mit nachfolgender Dichtefunktion:

$$f_X(x) := \begin{cases} 0,1x & \text{für } 0 \leq x \leq 4 \\ 2 - 0,4x & \text{für } 4 < x \leq 5 \\ 0 & \text{sonst} \end{cases}$$

a) Bestimmen Sie die zugehörige Verteilungsfunktion.

b) Bestimmen Sie den Erwartungswert und die Varianz.

Aufgabe 3

In einer Urne befinden sich fünf weiße und drei schwarze Kugeln. Es werden fünf Kugeln mit Zurücklegen gezogen.

a) Durch welche Ihnen bekannte Verteilung könnte das Zufallsexperiment beschrieben werden? Welche Parameter besitzt diese Verteilung? Bestimmen Sie deren Werte für das angegebene Zufallsexperiment.

b) Bestimmen Sie den Erwartungswert und die Varianz.

c) Wie groß sind die Wahrscheinlichkeiten 0, 1, 2, 3, 4 oder 5 schwarze Kugeln zu ziehen?

d) Wie groß ist die Wahrscheinlichkeit, mindestens drei weiße Kugeln zu ziehen?

e) Wie groß ist die Wahrscheinlichkeit, mindestens eine, aber höchstens zwei schwarze Kugeln zu ziehen?

Aufgabe 4

Es sei $X \sim \mathcal{N}(35, 144)$. Bestimmen Sie mittels geeigneter Tabelle nachfolgende Wahrscheinlichkeiten.

a) $P(X \leq 41)$, $P(X \leq 50)$, $P(X \leq 23)$

b) $P(X > 41)$, $P(X > 20)$

c) $P(23 < X \leq 41)$, $P(38 < X \leq 50)$, $P(20 < X \leq 23)$

Zufallsvariablen und spezielle Verteilungen

Aufgabe 5

Gegeben sei nachfolgende Verteilungsfunktion:

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 0,05x^2 & \text{für } 0 \leq x \leq 4 \\ -0,2x^2 + 2x - 4 & \text{für } 4 < x \leq 5 \\ 1 & \text{für } x > 5 \end{cases}$$

Bestimmen Sie das 10% und das 95%-Quantil.

Parameterschätzung

Aufgabe 1

Für einen interessierenden Parameter θ sind zwei Schätzer $\hat{\theta}_1$ und $\hat{\theta}_2$ mit folgenden Eigenschaften gegeben:

$$\mathrm{E}(\hat{\theta}_1) = \theta - 3 \qquad \mathrm{E}(\hat{\theta}_2) = \theta - 2 \qquad \mathrm{Var}(\hat{\theta}_1) = 4 \qquad \mathrm{Var}(\hat{\theta}_2) = 9$$

$\hat{\theta}_1$ hat demnach im Vergleich zu $\hat{\theta}_2$ zwar die kleinere Varianz, allerdings den größeren Bias. Welchen der beiden Schätzer würden Sie bevorzugen? Begründen Sie kurz Ihre Antwort.

Aufgabe 2

Bei einer Untersuchung der Belastbarkeit $X \sim \mathcal{N}(\mu; \sigma^2)$ von Spielwaren ergab sich bei einer unabhängigen Stichprobe vom Umfang 10:

$$\bar{x} = 34{,}55 \qquad\qquad s^2 = 14{,}0539$$

Bestimmen Sie ein nach oben offenes 95%-Konfidenzintervall für die durchschnittliche Belastbarkeit.

Aufgabe 3

Die Zufallsvariable $X :=$ „Effektiv-Füllung von 5l-Ketchup-Eimern“ eines Abfüllers sei approximativ normalverteilt. Ein Abnehmer erhält eine zufällig zusammengestellte Lieferung von 20 Eimern, die eine durchschnittliche Füllung von $\bar{x} = 4{,}92$ Litern aufweisen. Unterstellen Sie eine Standardabweichung von $\sigma = 0{,}16$.

a) Bestimmen Sie ein zentrales 90%-Konfidenzintervall für die mittlere Füllung.

b) Welcher Stichprobenumfang ist mindestens erforderlich, damit ein zentrales 90%-Konfidenzintervall für μ höchstens halb so breit wie das unter a) bestimmte Intervall ist?

Aufgabe 4

Für eine bestimmte Glühbirnenart kann die Lebensdauer im Dauerbetrieb aufgefasst werden als eine (näherungsweise) normalverteilte Zufallsvariable X (gemessen in Tagen).

Eine unabhängige Zufallsstichprobe von 20 Birnen ergab:

$$\bar{x} = 132 \text{ Tage} \qquad\qquad \sum_{i=1}^{20} x_i^2 = 349020 \text{ Tage}^2$$

a) Bestimmen Sie erwartungstreue Schätzer für den Erwartungswert und die Varianz von X.

b) Bestimmen Sie ein 90%-Konfidenzintervall mit kleinstmöglicher Untergrenze für die Standardabweichung von X.

c) Geben Sie die Wahrscheinlichkeit an, dass die „wahre“ Standardabweichung im berechneten Konfidenzintervall liegt. Wie ändert sich Ihre Aussage, wenn Sie ein 90%-Konfidenzintervall bestimmen würden, dass nicht die kleinstmögliche Untergrenze besitzt?

Testen von Hypothesen

Aufgabe 1

Eine Firma verschickt Tee in Holzkisten mit jeweils 10 Teepackungen. Das Gewicht in kg einer Teepackung sei $X \sim \mathcal{N}(6; 0{,}0036)$. Einer großen Lieferung werden 16 Teepackungen entnommen. Dabei ergab sich ein durchschnittliches Gewicht von 5,95 kg.

a) Bestimmen Sie das 95%-Konfidenzintervall für das durchschnittliche Gewicht einer einzelnen Teepackung.

b) Sie vermuten, dass im Mittel nicht der Sollwert abgefüllt wird. Treffen Sie eine Testentscheidung auf dem 5%-Signifikanzniveau und führen Sie gegebenenfalls einen geeigneten Test durch.

Aufgabe 2

Ein Unternehmen füllt seine Produkte in 0,5l-Flaschen ab. Die Abfüllmaschine arbeitet jedoch nicht exakt. Die Füllmenge X in Litern kann dabei als normalverteilt angesehen werden. Zusätzlich sei aus der Vergangenheit die Standardabweichung $\sigma = 0{,}1$ bekannt. Eine unabhängige Stichprobe vom Umfang 81 ergab:

$$\sum_{i=1}^{81} x_i = 41{,}3$$

Nehmen Sie an, ein Mitarbeiter ist beauftragt worden zu überprüfen, ob die Füllmenge von der Vorgabe abweicht. Allerdings hat sich der Chef noch nicht für ein Signifikanzniveau entschieden. Welche Größe könnte der Mitarbeiter trotzdem bestimmen und seinem Chef übermitteln? Bestimmen Sie diese Größe.

Aufgabe 3

Ein Unternehmen bezieht seit langem von einem bestimmten Lieferanten einen Massenartikel, wobei der Ausschussanteil 5% beträgt. Ein Konkurrenzangebot verspricht bei gleichem Preis einen Ausschussanteil unter 5%. Unter 100 zufällig ausgewählten Artikeln des konkurrierenden Anbieters waren 2 Ausschussteile.

a) Können Sie schließen, dass der Ausschussanteil des Konkurrenzanbieters unter 5% liegt. Formulieren Sie einen statistischen Test als Entscheidungsgrundlage, wobei Sie von der Unabhängigkeit der Artikel ausgehen können. Wie ist auf dem 5%-Signifikanzniveau zu entscheiden?

b) Berechnen Sie den p-Wert des Tests aus Aufgabenteil a). Welches Signifikanzniveau müsste man vorgeben, damit H_0 gerade noch abgelehnt wird?

Spezielle Testprobleme

Aufgabe 1

Es besteht die Vermutung, dass die handelsüblichen Sahnebecher einer bestimmten Marke, die mit dem Aufdruck 200ml verkauft werden, im Mittel mit weniger Inhalt gefüllt sind.

Um dies zu untersuchen, wird eine Zufallsstichprobe von 120 Bechern gezogen. Diese ergibt für die Füllmenge den Mittelwert $\bar{x} = 194$ml und die Standardabweichung $s = 16$ml.

a) Schätzen Sie den Erwartungswert und die Varianz erwartungstreu.

b) Die durch den Aufdruck gemachte Behauptung soll überprüft werden. Sie möchten dabei eine zu geringe Füllmenge nachweisen. Führen Sie hierfür einen geeigneten Test auf dem 1%-Signifikanzniveau durch.

Aufgabe 2

In einer Stichprobe von 21 neu emittierten Unternehmensanleihen mit einem Rating von AAA wies die Laufzeit eine empirische Varianz von 40 auf. In einer anderen Stichprobe von 13 neu emittierten Unternehmensanleihen mit einem Rating von CCC betrug die Varianz der Laufzeit nur 12.

Prüfen Sie, ob die Varianz der Laufzeit der AAA-Papiere signifikant größer als die der CCC-Papiere ist. Treffen Sie für Ihren Test eine sinnvolle Verteilungsvoraussetzung und führen Sie diesen auf dem 2,5%-Signifikanzniveau durch.

Aufgabe 3

Ein Marktforschungsinstitut wurde von einem Einzelhändler beauftragt, das Kaufverhalten in einem Selbstbedienungsmarkt zu untersuchen. Dabei werden zwei Produkte A und B an zwei aufeinander folgenden Wochen je einmal oben im Regal bzw. unten im Regal platziert. Es ergaben sich folgende Verkaufszahlen:

Anordnung	A	B	Σ
1. Woche (A oben; B unten)	146	124	270
2. Woche (B oben; A unten)	84	118	202
Σ	230	242	472

Prüfen Sie mit einem geeigneten Test auf dem 5%-Signifikanzniveau, ob das Kaufverhalten von der Platzierung im Regal abhängt.

Aufgabe 4

In einem Unternehmen wird ein bestimmtes Produkt an drei Produktionsstätten A, B und C hergestellt.

Aus der laufenden Produktion wurden zufällig 800 Produkte ausgewählt. Anhand des Kontrollzettels wird festgestellt, dass 336 bzw. 288 davon an der Produktionsstätte A bzw. B hergestellt wurden. Von den 800 Produkten sind 120 fehlerhaft. 53 bzw. 44 dieser fehlerhaften Produkte stammen von der Produktionsstätte A bzw. B.

Prüfen Sie mit einer Irrtumswahrscheinlichkeit von 0,05 die Hypothese, dass die Fehlerhaftigkeit unabhängig von der Produktionsstätte ist.

Lineare Einfachregression

Aufgabe 1

Zur Schätzung einer Kostenfunktion verwendet ein Unternehmer die Daten von 10 Monaten für die monatliche Produktionsmenge M (in 10000 Stück) und die Herstellungskosten K pro Monat (in 10000 Euro). Aus den Datenpaaren (m_i, k_i) für die Monate $i = 1, \ldots, 10$ ergeben sich folgende Summenwerte:

$$\sum_{i=1}^{10} m_i = 55 \qquad \sum_{i=1}^{10} m_i^2 = 385 \qquad \sum_{i=1}^{10} k_i = 720 \qquad \sum_{i=1}^{10} k_i^2 = 53458 \qquad \sum_{i=1}^{10} m_i k_i = 4323$$

a) Geben Sie ein geeignetes Regressionsmodell an.

b) Schätzen Sie die unbekannten Parameter mit der KQ-Methode und interpretieren Sie den Steigungskoeffizienten.

c) Bestimmen Sie das symmetrische 95%-Konfidenzintervall für den Steigungskoeffizieten.

Aufgabe 2

Der Inhaber eines Eiscafes untersucht die Abhängigkeit des Eisumsatzes im Straßenverkauf V (in 1000 Euro) von der Mittagstemperatur T (in Grad Celsius). Aus seiner Erhebung an 15 Tagen ergaben sich folgende Werte:

$$\frac{1}{15}\sum_{i=1}^{15} v_i = 5{,}3 \qquad \frac{1}{15}\sum_{i=1}^{15} (v_i - \bar{v})^2 = 3{,}24 \qquad \frac{1}{15}\sum_{i=1}^{15} t_i = 22 \qquad \frac{1}{15}\sum_{i=1}^{15} t_i^2 = 494{,}24 \qquad \frac{1}{15}\sum_{i=1}^{15} v_i t_i = 121{,}78$$

a) Geben Sie ein geeignetes Regressionsmodell an.

b) Schätzen Sie die unbekannten Parameter mit der KQ-Methode und interpretieren Sie den Steigungskoeffizienten.

c) Kann auf dem 1%-Signifikanzniveau geschlossen werden, dass der Eisumsatz mit steigenden Temperaturen zunimmt?

Aufgabe 3

Bei einer Verkehrsuntersuchung wurde für 25 Verkehrsbezirke der Individualverkehr erfasst. Als Beschreibung für die Struktur des Verkehrs wurden die Merkmale

X : „Einwohnerzahl in 1000“ und Y : „Zahl der PKW pro Stunde in 1000“

für die 25 Verkehrsbezirke ermittelt. Dabei ergaben sich folgende Werte:

$$\sum_{i=1}^{25} y_i = 79{,}75 \qquad \sum_{i=1}^{25} x_i = 100 \qquad \sum_{i=1}^{25} y_i^2 = 277{,}825 \qquad \sum_{i=1}^{25} x_i^2 = 475 \qquad \sum_{i=1}^{25} y_i x_i = 345{,}5$$

a) Schätzen Sie die unbekannten Parameter mit der KQ-Methode und unterstellen Sie dabei, dass die Zahl der PKW von der Einwohnerzahl abhängt.

b) Testen Sie die Nullhypothese, dass die Steigung der Regressionsgeraden auf dem 10%-Signifikanzniveau höchstens 0,3 beträgt.

Multiple lineare Regression

Aufgabe 1

Mit den Daten von 177 Mietwohnungen einer Stadt wurde versucht, die Determinanten des Mietzinses empirisch zu ermitteln. Mittels KQ-Methode wurde folgendes Modell geschätzt:

$$MIETE = \beta_0 + \beta_1 NWF + \beta_2 Alter + \beta_3 DZENT + \varepsilon$$

Dabei bedeuten die Variablen MIETE: Mietpreis in Euro, NWF: Nettowohnfläche in qm, ALTER: Alter der Wohnung in Jahren, DZENT: Distanz zum Stadtzentrum in km.
EViews liefert folgende Regressionsergebnisse:

Dependent Variable: MIETE
Method: Least Squares
Date: 12/11/02 Time: 08:00
Sample: 1 177
Included observations: 177

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C		144.7024	6.335762	0.0000
NWF	11.40628		7.947048	0.0000
ALTER	-6.469049	0.937492		0.0000
DZENT		20.54634	-3.433036	
R-squared		Mean dependent var	1303.944	
Adjusted R-squared		S.D. dependent var	526.3415	
S.E. of regression				
Sum squared resid	26928912			

a) Bestimmen Sie die fehlenden Werte des Outputs. Dabei ist die Summe der quadrierten Abweichungen der abhängigen Variablen MIETE von ihrem Mittelwert $48.758.231,3776$.

b) Beurteilen Sie die Signifikanz von $\widehat{\beta}_0$ und $\widehat{\beta}_1$ und interpretieren Sie sie inhaltlich.

c) Bilden Sie das 90%-Konfidenzintervall für das Alter der Wohnung.

d) Testen Sie, ob die Distanz zum Stadtzentrum auf dem 1%-Signifikanzniveau einen signifikanten Einfluss auf den Mietzins hat.

e) Testen Sie die Aussage, dass der Quadratmeter Wohnfläche mehr als 10 Euro kostet ($\alpha = 0{,}025$).

Aufgabe 2

Gegeben sei folgende Stichprobe mit 5 Beobachtungen für x_1, x_2 und y:

x_1	1,2	3	4,5	5,8	7,2
x_2	0,6	0,75	0,8	0,9	1,4
y	2,6	1,6	4	3	4,9

Bestimmen Sie die Parameter des Modells (mit Intercept):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Nutzen Sie für die Bestimmung:

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} 3,0448 & 0,4025 & -5,1593 \\ 0,4025 & 0,2264 & -1,5565 \\ -5,1593 & -1,5565 & 13,3871 \end{pmatrix}$$

Aufgabe 3

Für eine multiple Regression mit $n = 14$ ergab sich folgender Schätzer für den Vektor $\hat{\boldsymbol{\beta}}$:

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} -1,0014 \\ -0,2347 \\ -0,001 \end{pmatrix}$$

Für das Bestimmtheitsmaß wurde zusätzlich ein Wert von $0,475$ errechnet. Prüfen Sie auf dem $2,5\%$-Signifikanzniveau, ob mindestens ein Regressor zur Erklärung der Zielvariablen beiträgt.