Fachbereich Mathematik
Prof. Dr. J. Lehn
A. Neuenkirch
B. Niese
A. Rößler

TECHNISCHE UNIVERSITÄT DARMSTADT

26./27. April 2006

Einführung in die Mathematische Statistik für WInf, LaB, CE Übung 1

Gruppenübung

G 1 Bei der Messung des Körpergewichts von 20 männlichen Schülern ergaben sich die folgenden Werte [in kg]:

62.5	58.0	57.0	59.5	64.0	54.5	58.5	59.5	57.0	61.5
56.0	61.0	54.5	65.5	62.5	57.0	60.0	56.5	59.5	58.5

a) Erstellen Sie ein Histogramm zu der Klasseneinteilung: $(54.0, 56.0]$, $(56.0, 58.0], \ldots,$ $(64.0, 66.0]$.

b) Skizzieren Sie die empirische Verteilungsfunktion zu der Messreihe.

c) Bestimmen Sie das arithmetische Mittel, den Median, die Spannweite, die empirische Varianz, die empirische Standardabweichung und den Quartilabstand.

d) Veranschaulichen Sie die Struktur der Messreihe mit Hilfe eines Boxplots.

G 2 Bei einer Klausur konnten maximal 120 Punkte erreicht werden. In der folgenden Tabelle ist zu jeder Note diejenige Punktzahl p angegeben, die zum Erhalt dieser Note mindestens erreicht werden mußte.

Note	1.0	1.3	1.7	2.0	2.3	2.7	3.0	3.3	3.7	4.0	5.0
p	120	115	109	102	94	85	75	64	52	40	0

Folgende Punktzahlen wurden von den 25 Teilnehmern erreicht:

41	82	5	84	94	47	73	59	62	20	31	98	76
12	41	52	102	45	70	40	74	51	64	75	60	

a) Bestimmen Sie jeweils im Sinne der Bildung des arithmetischen Mittels die durchschnittliche Punktzahl sowie die Durchschnittsnote.

b) Vergleichen Sie die Note, die man bei durchschnittlicher Punktzahl erhalten würde, mit der Durchschnittsnote.

c) Bestimmen Sie die Mediane zu Noten und Punktzahlen.

d) Vergleichen Sie die Note, die man beim Median der Punktzahlen erhalten würde, mit dem Median der Notenverteilung.

G 3 Gegeben sei eine Messreihe $x_1, \ldots, x_n$ mit arithmetischem Mittel $\overline{x}$ und empirischer Varianz s_x^2.

a) Zeigen Sie: Werden die Messwerte gemäß $y_i = ax_i + b$ für $i = 1, \ldots, n$ linear transformiert, so gilt für das arithmetische Mittel der transformierten Werte $\overline{y} = a\overline{x} + b$ und für die empirische Varianz $s_y^2 = a^2 s_x^2$.

b) In einem Badeort in Florida wurden in einer Juliwoche jeweils morgens zur gleichen Zeit die folgenden Lufttemperaturen in °Fahrenheit gemessen:

79 82 81 80 80 77 80

Berechnen Sie die Durchschnittstemperatur und die empirische Streuung in °Celsius
Hinweis: $[°\mathrm{C}] = 5/9 \cdot ([°\mathrm{F}] - 32)$

Hausübung

H 1 In 40 Haushalten wurde die Anzahl der vorhandenen Elektrogeräte ermittelt. Es ergaben sich die folgenden Werte:

11	20	7	14	13	12	13	9	15	8
12	18	6	10	19	5	8	11	14	11
8	15	12	17	12	6	7	12	9	6
14	12	11	15	14	9	7	15	11	13

a) Bestimmen Sie den Median zu dieser Messreihe.

b) Veranschaulichen Sie die Messreihe mit einem Stabdiagramm.

c) Erstellen Sie die zugehörige empirische Verteilungsfunktion und lesen Sie daraus die empirischen Quantile $x_{0.1}$ und $x_{0.7}$ ab.

d) Erstellen Sie einen Boxplot.

H 2 Bestimmen Sie diejenige lineare Transformation

$$y_i = ax_i + b, \qquad a > 0,$$

durch die sich aus einer Messreihe $x_1, \ldots, x_n$ mit $s_x^2 > 0$ eine transformierte Messreihe $y_1, \ldots, y_n$ ergibt, für die

$$\overline{y} = 0 \qquad \text{und} \qquad s_y^2 = 1$$

gilt (*Standardisierung einer Messreihe*).

H 3 Sei $x_1, \ldots, x_n$ eine Messreihe mit arithmetischem Mittel $\overline{x}$ und dem Median $\widetilde{x}$, $f : \mathbb{R} \to \mathbb{R}$ monoton wachsend und $y_i = f(x_i)$ für $i = 1, \ldots, n$.

a) Zeigen Sie, dass für den Median $\widetilde{y}$ der transformierten Messreihe $y_1, \ldots, y_n$

$$\widetilde{y} = f(\widetilde{x})$$

gilt.

b) Geben Sie den Median der Temperaturen aus Aufgabe G3 in °Celsius an.

c) Gilt für die arithmetischen Mittel $\overline{x}$ und $\overline{y}$ ebenfalls $\overline{y} = f(\overline{x})$?

H 4 Gegeben sei die zweidimensionale Messreihe $(0, c), (1, 1), (2, 0)$ mit $c \in \mathbb{R}$. Bestimmen Sie in Abhängigkeit von c die Regressionsgerade.
Für welches $c \in \mathbb{R}$ liegt der erste Punkt auf der Regressionsgeraden? Fertigen Sie für diesen Fall und für den Fall $c = -1$ jeweils eine Skizze an.

H 5 Im Rahmen von Untersuchungen über zulässige Höchstwerte von Schadstoffen in Futtermitteln wurden Fütterungsversuche an Mastschweinen angestellt. Dem Futter von 13 Schweinen wurden verschiedene Mengen von Bleiacetat $(x_1, \ldots, x_{13})$ beigegeben und nach einem bestimmten Zeitraum wurden dann die Bleikonzentrationen in der Nierenrinde $(y_1, \ldots, y_{13})$ gemessen. Die entsprechenden Werte (in mg Pb pro kg Futter bzw. pro kg Nierenrinde) sind in der nachfolgenden Tabelle dargestellt.

x_i	2.0	3.2	5.4	2.4	8.9	15.8	2.3	6.9	11.8	4.5	2.6	2.9	3.8
y_i	0.12	0.18	0.42	0.18	0.61	1.02	0.14	0.55	0.74	0.24	0.14	0.20	0.20

a) Erstellen Sie ein Punktediagramm und berechnen Sie die empirische Kovarianz sowie die empirische Korrelation zwischen der Bleikonzentration im Futter und in der Nierenrinde.

b) Bestimmen Sie die Regressionsgerade zur Vorhersage des Bleigehalts in der Nierenrinde aus der Bleikonzentration im Futter und tragen Sie diese in das Punktediagramm ein. Wie groß ist die Quadratsumme der Residuen?

c) Welchen Vorhersagewert für den Bleigehalt in der Nierenrinde erhält man bei einer Bleikonzentration des Futters von 10 mg.

H 6 Zeigen Sie, dass die empirische Korrelation bis auf Vorzeichenwechsel invariant gegenüber linearen Transformationen der Daten ist, d.h.:
Sind $(x_1, y_1), \ldots, (x_n, y_n)$ und $(u_1, v_1), \ldots, (u_n, v_n)$ zweidimensionale Messreihen mit

$$u_i = a \cdot x_i + b, \quad v_i = c \cdot y_i + d, \quad i = 1, \ldots, n,$$

für Zahlen $a, b, c, d \in \mathbb{R}$ mit $a \cdot c \neq 0$, so gilt für die jeweiligen empirischen Korrelationen

$$r_{uv} = \begin{cases} r_{xy} & \text{falls } a \cdot c > 0 \\ -r_{xy} & \text{falls } a \cdot c < 0. \end{cases}$$