

Klausur Statistik für B_MInf(v110), II(v103), IAW(154)

Klausurdatum: 11.8.08, 8:30, Bearbeitungszeit: 60 Minuten

Erlaubte Hilfsmittel:

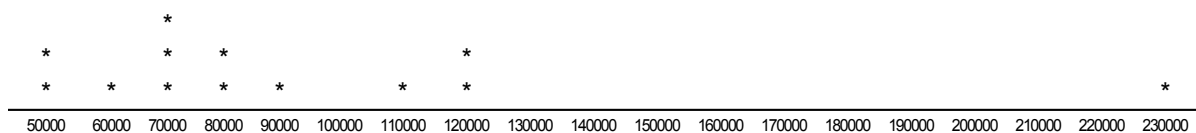
Taschenrechner, Zeichengeräte, Tabellenwerk und Formeln aus dem Handoutverzeichnis.

Aufgabe 1:

Sie fragen auf der Mönckebergstraße in Hamburg 13 Männer, die Krawatte, Anzug und geputzte Schuhe tragen (MKA_gS) nach dem aktuellen Jahreseinkommen. Weil das auf den Euro genau niemand weiß, lassen Sie sich das Resultat als den nächstliegenden glatten 10 000er-Euro-Betrag angeben. Sie erhalten folgende Urliste:

MKA_gS1: 90000, MKA_gS2: 50000, MKA_gS3: 110000, MKA_gS4: 120000, MKA_gS5: 80000, MKA_gS6: 70000, MKA_gS7: 80000, MKA_gS8: 50000, MKA_gS9: 60000, MKA_gS10: 70000, MKA_gS11: 230000, MKA_gS12: 70000, MKA_gS13: 120000

a) Fertigen Sie einen Dot Plot und eine Tabelle der Häufigkeit, der relativen Häufigkeit und der kumulierten relativen Häufigkeit der Verteilung an. **(4P)**



Einkommen in 10 TSD	Häufigk	rel. Häuf.	kum. rel. Häuf.
5	2	0,15	0,15
6	1	0,08	0,23
7	3	0,23	0,46
8	2	0,15	0,62
9	1	0,08	0,69
10	0	0,00	0,69
11	1	0,08	0,77
12	2	0,15	0,92
13	0	0,00	0,92
14	0	0,00	0,92
15	0	0,00	0,92
16	0	0,00	0,92
17	0	0,00	0,92
18	0	0,00	0,92
19	0	0,00	0,92
20	0	0,00	0,92
21	0	0,00	0,92
22	0	0,00	0,92
23	1	0,08	1,00
13		1	

b) Ein Herrenausstatter hat Interesse an Ihren Daten und möchte sein Sortiment auf MKAgS-Kunden maßschneidern. Welcher Lageparameter (arithmetisches Mittel, Median oder Modalwert) des Jahreseinkommens ist Ihrer Meinung nach am geeignetsten und warum?(1P)

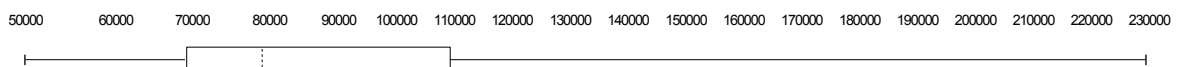
Median, weil der 230000 Ausreißer im arithmetischen Mittel eine falsche Lage vortäuschen würde. (Der Modalwert wäre auch geeignet.)

c) Geben Sie arithmetisches Mittel, Median und Modalwert des Einkommens Ihrer MKAgS-Stichprobe an.(2P)

arithm. Mittel: 92 300, Median: 80 000, Modalwert: 70 000

d) Fertigen Sie eine „Five Point Summary“ der MKAgS-Einkommen an und zeichnen Sie sie als Box-Whisker Plot.(3P)

Min: 50 000, 25er Perzentil: 70 000, Median: 80 000, 75er Perzentil: 110 000, Max: 230 000



e) Nennen Sie zwei Ihnen bekannte Streuparameter und ermitteln Sie die Zahlenwerte für die angegebenen Einkommensdaten.(2P)

*Varianz : 2 269 000 000 €²,
Standardabweichung : 47 600 €*

Einkommen in 10 TSD	Häufigk	# * Ek	#*(x- mitt)^2
5	2	10	35,8
6	1	6	10,44
7	3	21	14,93
8	2	16	3,03
9	1	9	0,05
10	0	0	0
11	1	11	3,13
12	2	24	15,34
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0
21	0	0	0
22	0	0	0
23	1	23	189,59
	13	120	272,31

Mittel 9,23
 Varianz: 22,69 in 10⁸ Euro²
 Std.abw. 4,76 in 10⁴ Euro

Aufgabe 2 (4 P):

Jemand behauptet Männer mit Krawatten, Anzügen und geputzten Schuhen, die auf der Mönckebergstraße herumlaufen, verdienen durchschnittlich 130000,-€ . Testen Sie diese Hypothese anhand Ihrer Stichprobe (nehmen Sie dabei an: $\alpha = 0,05$).

Test bei unbekannter Standardabw. -> Student-T Verteilung verwenden bei $s_1 = 47600$ Euro, $s_{13} = 47600/\sqrt{13} = 13202$ Euro, $\mu = 92300$ Euro.

Standardisieren: $t = (130000 - 92300)/13202 = 2,8556$

Zweiseitiger Test für $\alpha = 0,05$ -> T-Tabelle für 12 DOF, welcher t-Wert für $p = 0,025$?

-> $t_{\text{schwelle}} = 2,1788$

Ist geringer als 2,8556, die wahre Abweichung ist also größer als die für Beibehaltung der Hypothese gerade noch zu tolerierende Abweichung.

Wir müssen die Hypothese verwerfen!

Aufgabe 3:

Im 17. Jahrhundert lebt der junge Adelige Siegfried Freiherr von Süßholz. Weil er schöne Uniformen liebt, wird er Leutnant bei den kaiserlichen Kürassieren. Er ist gutaussehend, liebt die Literatur, die Kunst, edle Pferde und vor allem schöne Frauen. Militärische Übungen, Fechten und Schießen verabscheut er.

Aus dieser Kombination ergibt sich für ihn ein Problem. Wegen seiner notorischen Schürzenjägeri besteht ständig das Risiko, von gehörnten Ehemännern und gekränkten Verlobten zum Duell gefordert zu werden. Leider ist für einen solchen Fall wegen mangelnder Fähigkeiten seinerseits die Wahrscheinlichkeit, ein Duell siegreich zu überstehen nur etwa 0,2. Neben der Schande zieht eine Niederlage(es gibt kein Unentschieden beim Duell) immer auch den Tod oder (selten) nur eine so schwere Verletzung nach sich, dass damit auch die Schürzenjägeri endgültig beendet ist.

Durchschnittlich alle vier Monate macht er eine neue Eroberung in der Damenwelt. Er ist nicht dumm und meidet deshalb Damen, deren offizielle Verehrer besonders kampfeswillig und misstrauisch sind. Dadurch erreicht er, dass bei einer neuen Eroberung die Wahrscheinlichkeit, entdeckt und zum Duell gefordert zu werden nur 0,47 beträgt.

a) (2P) Wie groß ist die Wahrscheinlichkeit, dass Freiherr von Süßholz bei diesem Lebenswandel 3 Jahre lebendig oder zumindest unverletzt zu übersteht ?

Für eine Eroberung: $p(\text{kein Duell}) = (1-0,47) = 0,53$

$p(\text{Duell gewonnen} | \text{Duell}) = 0,47 * 0,2 = 0,094$

$p(\text{überlebt}) = 0,53 + 0,094 = 0,624$

oder $1 - p(\text{Duell verloren} | \text{Duell}) = 1 - 0,47 * 0,8 = 0,624$

3 Jahre oder 9 Eroberungen unverletzt = $0,624^9 = 0,01434$

b) (2P) Wenn er trotz lausiger Fähigkeiten aus purem Glück seine ersten 5 Duelle gewänne, eilte ihm ein Ruf voraus, der potenzielle Duellgegner derartig abschrecken würde, dass er fortan ungestraft den Damen nachstellen könnte. Wie groß ist die Wahrscheinlichkeit, dass er diesen paradisischen Zustand irgendwann erreicht ?

$p(5 \text{ Duelle, alle gewonnen}) = 0,2^5 = 0,00032$

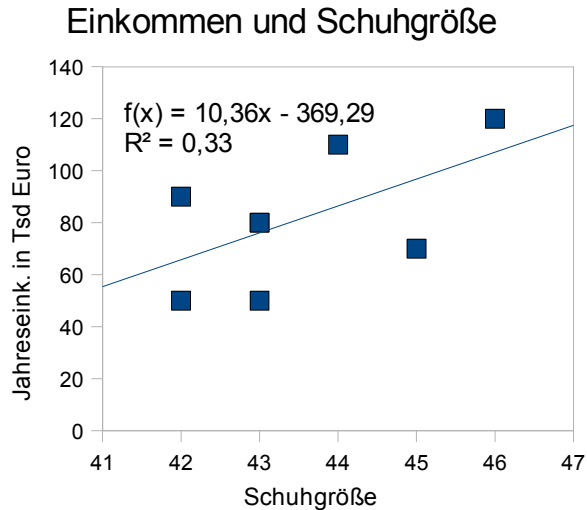
Aufgabe 4:

Sie fragen auf der Mönckebergstraße in Hamburg 8 Männer, die Krawatte, Anzug und geputzte Schuhe tragen (MKA_gS) nach dem aktuellen Jahreseinkommen und der Schuhgröße:

MKA_gS1: (90000, 42) MKA_gS2: (50000, 43), MKA_gS3: (110000, 44), MKA_gS4: (120000, 46) MKA_gS5: (80000, 43), MKA_gS6: (70000, 45), MKA_gS7: (80000, 43), MKA_gS8: (50000, 42)

a) (2P) Zeichnen Sie einen sogenannten Scatter-Plot der Daten (mit richtigen Achsenbezeichnungen).

Schuhgr.	EK in Tsd
42	90
43	50
44	110
46	120
43	80
45	70
43	80
42	50



b) (4P) Berechnen Sie über Varianz und Kovarianz die Konstanten a und b für die Lineare Regression und den Korrelationskoeffizienten. Geben Sie außerdem die konkrete Form der ermittelten Regressionsgerade (mit physikalischen Einheiten) an.

Erw.Eink. = 10,36 Tsd€/Schuhgr * x - 369,29 Tsd€

Schuhgr.	EK in Tsd	yvar	xvar	kovar
42	90	76,56	2,25	-13,13
43	50	976,56	0,25	15,63
44	110	826,56	0,25	14,38
46	120	1501,56	6,25	96,88
43	80	1,56	0,25	0,63
45	70	126,56	2,25	-16,88
43	80	1,56	0,25	0,63
42	50	976,56	2,25	46,88

mitt		4487,5	14,00	145,00
43,5	81,25			

xVarianz	2,00000	b:	10,36
kovarianz	20,71	a:	-369,29
yVarianz	641,07	r:	0,58
		r ² :	0,33

c) (2P) Kommentieren Sie den Wert des Korrelationskoeffizienten. Besteht eine starke Korrelation? Wie ändert sich der Korrelationskoeffizient, wenn Sie das Einkommen statt in Euro in DM angeben würden ? (Hinweis 1,95 DM ≈ 1€)

Keine Überzeugende Korrelation, aber nicht "NICHTS". In DM gerechnet würde der gleiche Korrelationskoeffizient herauskommen, da er dimensionslos ist. (a und b hätten natürlich andere Zahlenwerte und andere Einheiten.)