

StR.i.H. Albrecht Gündel-vom Hofe

2. Aufgabenblatt zur „Statistik für Biologen“

(Abgabe der H-Aufgaben: Mittwoch, 07.11.2012, in der Großen Übung)

Ü 7. Aufgabe:

Wir betrachten noch einmal die Datenwerte x_i aus der Aufgabe 4 vom 1. Blatt, gegeben durch: 34, 45, 11, 42, 49, 33, 27, 11.

a) Geben Sie die lineare Transformation an, welche die gegebenen Daten x_i standardisiert. Wie lauten die standardisierten Daten x_i^* ? Zeigen Sie an diesen Daten noch einmal konkret, dass gilt: $\overline{x^*} = 0$ sowie $\text{var}(x^*) = 1$.

b) Berechnen Sie den Schiefeparameter $\rho = \frac{1}{n} \sum_{i=1}^n (x_i^*)^3$ für die standardisierten Daten.

H 8. Aufgabe:

Betrachten Sie noch einmal die in Aufgabe 2 gegebenen Daten:

BL	HE	NS	NRW	SH	BB	MV	S
Anteil	16	15	20	20	10	10	19

BL	SA	TH	BW	B	HH	RP	SL
Anteil	14	38	19	19	33	24	19

Ermitteln Sie die zugehörigen standardisierten Daten sowie den Schiefeparameter

$$\rho = \frac{1}{n} \sum_{i=1}^n (x_i^*)^3 \text{ und den Wölbungsparameter } w = \frac{1}{n} \sum_{i=1}^n (x_i^*)^4 .$$

	5,0
--	-----

Ü 9. Aufgabe:

Bei 10 Objekten seien jeweils die Länge X (in cm) und das Gewicht Y (in kg) gemessen worden. Die einzelnen Messwerte sind gegeben durch:

I	1	2	3	4	5	6	7	8	9	10
x_i	1,0	2,0	2,0	4,0	5,0	6,0	9,0	9,0	9,0	13,0
y_i	14,07	15,60	21,92	27,90	30,40	22,25	37,43	38,40	27,95	41,79

Zeichnen Sie die zugehörige Punktwolke und bestimmen Sie die empirische Kovarianz sowie den empirischen Korrelationskoeffizienten dieser Punktwolke.

H 10. Aufgabe:

Bei einer Versuchsreihe muss ein zu reinigendes Medium eine Filterschicht aus kleinen porösen Tonkugeln mit dem Durchmesser D passieren. Die Filterwirkung wird durch die quantitative Variable Y gemessen. Bei $n = 10$ Versuchen ergaben sich folgende Wertepaare:

Y	2,07	2,73	2,52	2,68	2,65	2,30	2,52	1,78	2,37	3,68
D	1	2	3	4	5	6	7	8	9	10

Bestimmen Sie den empirischen Korrelationskoeffizienten zwischen der Filterwirkung Y und der *Größe* der Kugeln. Definieren Sie dabei die *Größe* in folgenden 3 Varianten:

- (a) über den Durchmesser D , (b) über die Oberfläche $O \approx D^2$,
(c) über das Volumen $V \approx D^3$.

	5,0
--	-----

Ü 11. Aufgabe:

Bei einer Untersuchung von Schulkindern wurde die Anzahl der mit Karies befallenen Zähne dokumentiert. Lösen Sie die folgenden Aufgaben unter Zuhilfenahme des Statistiklabors:

1. Lesen Sie den Datensatz „Karies“ ein und stellen Sie ihn in einem Datensatzobjekt dar.
2. Berechnen Sie, wie viele Kinder untersucht wurden.
3. Was ist die kleinste bzw. größte Anzahl an befallenen Zähnen bei einem Kind
4. Stellen Sie die Variable „Karies“ in einem Histogramm dar. Achten Sie auf die Klassenbreite bzw. Klassenanzahl. Erklären Sie, warum dies wichtig ist.
5. Berechnen Sie den Mittelwert, den Median und die Varianz von „Karies“.
6. Warum ist der Mittelwert an dieser Stelle irreführend?
7. Wie viele Kinder haben maximal zwei befallene Zähne? Wie viele haben genau neun?
8. Zeichnen Sie eine Verteilungsfunktion von 'Karies'.

H 12. Aufgabe:

Im Sommersemester 2011 wurde unter den Studenten der Wirtschaftswissenschaften eine Umfrage gemacht.

1. Lesen Sie den Datensatz 'Umfrage2011SoSe' ein und stellen Sie ihn in einem Datensatzobjekt dar.
2. Berechnen Sie mit Hilfe des R-Kalkulators den Mittelwert und den Median der Variablen GEBJAHR.
3. Berechnen Sie das heutige Alter des jüngsten und des ältesten Studierenden.
4. Mit dem Befehl 'm <- (GRO[GESCHL==1])' wird der neuen Variablen 'm' alle Größen zugeordnet, für die 'GESCHL=1', also männlich, gilt. Ordnen Sie einer Variablen 'GROm' und 'GROw' die Größen der weiblichen und männlichen Studierenden zu.
5. Zeichnen Sie mit Hilfe des R-Grafik Wizards zwei Box-Plots (in einem Fenster) für die Variable 'm' und 'w'. Achten Sie auf die Beschriftungen.
6. Interpretieren Sie die Ergebnisse der Grafik.
7. Berechnen Sie den Mittelwert, die Varianz und die Spannweite der neuen Variablen.
8. Erklären Sie, warum die Varianzen und die Spannweite bei den männlichen Studierenden so groß ist.
9. (evtl. Zusatz) Wie kann dies behoben werden?

	10,0
--	------