

StR.i.H. Albrecht Gündel-vom Hofe

3. Aufgabenblatt zur „Statistik für Biologen“

(Abgabe der H-Aufgaben: Mittwoch, 14.11.2012, in der Großen Übung)

Ü 13. Aufgabe:

Der Ernteertrag Y einer Pflanze hängt unter anderem vom Wassergehalt X des Bodens ab. Dabei ist Y minimal, wenn der Boden zu trocken oder zu feucht ist. Optimal ist er bei mittleren Werten $X \approx x_{opt}$. Dann ist die Korrelation $\rho(X, Y)$ abhängig vom Wertebereich, in dem X gemessen wird. Gilt nun:

a) $\rho(X, Y) > 0$, falls $X \leq x_{opt}$, und $\rho(X, Y) < 0$, falls $X \geq x_{opt}$

oder gilt:

b) $\rho(X, Y) < 0$, falls $X \leq x_{opt}$, und $\rho(X, Y) > 0$, falls $X \geq x_{opt}$

oder gilt weder (a) noch (b)?

H 14. Aufgabe:

Beantworten Sie die folgenden Fragen und begründen Sie jeweils Ihre Antwort.

- a) Bei einer Verpackungsmaschine seien das Nettogewicht N des Füllgutes und das Gewicht T der Verpackung voneinander unabhängig. Das Bruttogewicht B ist die Summe von beiden; also gilt: $B = N + T$. Sind B und N nun unkorreliert, positiv korreliert oder negativ korreliert?
- b) Bei einer Abfüllmaschine werden Ölsardinen in Öl in Dosen verpackt. Es sei S das Gewicht der Sardinen und O das Gewicht des Öls in einer Dose. Sind dann S und O unkorreliert oder positiv korreliert oder sogar negativ korreliert? Wie groß ist die Korrelation zwischen S und O , wenn das Gesamtgewicht $S + O$ genau 100 Gramm beträgt?

| | |
|--|-----|
| | 6,0 |
|--|-----|

Ü 15. Aufgabe:

Es seien die folgenden 6 Werte x_i als Merkmalsausprägungen zu X gegeben:

| | | | | | | |
|-------------------------|--------|--------|--------|-------|--------|--------|
| I | 1 | 2 | 3 | 4 | 5 | 6 |
| x_i | -1.188 | -1.354 | -1.854 | 0.146 | -0.354 | -0.521 |

- a) Die zugehörigen Werte y_i des Merkmals Y seien definiert durch die Gleichung $y_i = x_i + 0.85433$ ($i = 1, \dots, 6$). Wie groß ist die Korrelation $r(\mathbf{x}, \mathbf{y})$?
- b) Nun sei definiert $u_i = \frac{1}{x_i}$ sowie $v_i = \frac{1}{y_i}$ ($i = 1, \dots, 6$). Wie groß ist nun die Korrelation $r(\mathbf{u}, \mathbf{v})$ der transformierten Daten u_i und v_i ?

H 16. Aufgabe:

Gegeben sei die folgende (fiktive) Tabelle mit Daten aus 10 verschiedenen Ländern. Dabei steht F_i für die Fläche des Landes in km^2 , B_i für die Anzahl (in Tausend) der im letzten

Jahr dort geborenen Babys, S_i für die Anzahl der Störche in dem entsprechenden Land sowie W_i für die Gesamtgröße der Wasserfläche in km^2 in den einzelnen Ländern. Weiterhin sind mit S_F, B_F und S_W entsprechende Quoten mal 100 (siehe dazu speziell die folgenden Aufgabenteile) miterfasst.

| Land | F_i | B_i | S_i | W_i | S_F | B_F | S_W |
|------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 8624 | 370 | 213 | 157 | 2.47 | 4.3 | 135 |
| 2 | 9936 | 210 | 48 | 150 | 0.48 | 2.1 | 32 |
| 3 | 2093 | 323 | 100 | 190 | 4.78 | 15.4 | 53 |
| 4 | 3150 | 306 | 152 | 185 | 4.83 | 9.7 | 82 |
| 5 | 4584 | 373 | 146 | 177 | 3.18 | 8.1 | 82 |
| 6 | 4294 | 556 | 95 | 179 | 2.21 | 13.0 | 53 |
| 7 | 15570 | 520 | 85 | 122 | 0.55 | 3.3 | 69 |
| 8 | 9260 | 300 | 149 | 154 | 1.61 | 3.2 | 97 |
| 9 | 2377 | 580 | 149 | 288 | 6.27 | 24.4 | 52 |
| 10 | 12149 | 287 | 192 | 139 | 1.58 | 2.4 | 138 |

- Bestimmen Sie die Korrelationen $r(F, B)$, $r(F, S)$ sowie $r(B, S)$.
- Beziehen Sie dann die jeweilige Anzahl der Babys und der Störche auf die zur Verfügung stehende Fläche. Dabei bezeichne (siehe Tabelle) $(S_F)_i = \frac{S_i}{F_i}$ die Anzahl der Störche pro km^2 und $(B_F)_i = \frac{B_i}{F_i}$ die Anzahl der Babys pro km^2 . Bestimmen Sie die Korrelation $r(S_F, B_F)$.
- Beziehen Sie nun die Anzahl der Störche nicht auf die Größe des Landes, sondern auf die Größe der nahrungsspendenden Wasserfläche W . Dann bezeichne $(S_W)_i = \frac{S_i}{W_i}$ die Anzahl der Störche pro km^2 Wasserfläche. Bestimmen Sie die Korrelation $r(S_W, B_W)$.

Was lässt sich aus den berechneten Korrelationen lernen?

| | |
|--|-----|
| | 6,0 |
|--|-----|

Ü 17. Aufgabe:

Diese Aufgabe ist unter Zuhilfenahme des Statistiklabors zu bearbeiten und nimmt Bezug auf den Datensatz 'Umfrage2011SoSe' vom 2. Aufgabenblatt.

- Lesen Sie den Datensatz 'Umfrage2011SoSe' ein und stellen Sie es in einem Datensatzobjekt dar.
- Berechnen Sie den empirischen Korrelationskoeffizienten der Variablen 'GRO' und 'GEW'. Wie kann das Ergebnis interpretiert werden?
- Es soll der empirische Korrelationskoeffizient anhand der Formel aus dem Skript über die Kovarianz und die Standardabweichungen von 'GRO' und 'GEW' berechnet werden.

Definieren Sie hierfür die Werte 'covber', 'sdgro' und 'sdgew' und berechnen Sie den empirischen Korrelationskoeffizienten. Stimmt er mit Ihrem Ergebnis aus b) überein?

4. Zeichnen Sie ein Streudiagramm für die Variable 'GEW' und 'GRO'.
5. Vergleichen Sie die Ergebnisse des Streudiagramms mit ihren berechneten Werten.

H 18. Aufgabe:

Diese Aufgabe ist eine Weiterführung der Aufgabe Ü 11 unter Zuhilfenahme des Statistiklabors unter Rückgriff auf den Datensatz 'Umfrage2011SoSe' – speziell auf die Variablen 'FACHSEM' und 'KLAUSUR'.

1. Bestimmen Sie die durchschnittliche Anzahl an Fachsemestern sowie der bereits bestandenen Klausuren.
2. Berechnen Sie, wie viele Studierende mehr als 5 Klausuren bestanden haben. Wie viel Prozent aller Studierenden sind dies?
3. Erstellen Sie die empirische Verteilungsfunktion zu 'KLAUSUR'.
4. Deckt sich Ihr Ergebnis aus 2) mit der Darstellung?
5. Bestimmen Sie den empirischen Korrelationskoeffizienten der beiden Variablen. Was kann man daraus interpretieren?
6. Erstellen Sie ein Streudiagramm zu den beiden Variablen.

| | |
|--|-----|
| | 8,0 |
|--|-----|