

A 1 (5 Punkte)

Gegeben sind die Werte einer zweidimensionalen Stichprobe

| | | | | | |
|-------|---|---|----|---|---|
| x_i | 3 | 7 | 11 | 4 | 5 |
| y_i | 4 | 5 | 9 | 6 | 7 |

- a) Untersuchen Sie die Korrelation zwischen x_i und y_i . Berechnen Sie dazu die Zahlenwerte für:
- S_x^2 , d. h. die Varianz der x -Werte
 - S_y^2 , d. h. die Varianz der y -Werte
 - S_{xy} , d. h. die Kovarianz der Stichprobe
 - r_{xy} , d. h. den Korrelationskoeffizienten
- b) Bestimmen Sie die Regressionsgeraden von y bzgl. x .
- c) Wie lautet die Prognose für y , wenn x den Wert 5 annimmt?

A 2 (5 Punkte)

In einer Zeitschrift wird eine Studie zitiert, nach der Männer mit weniger Haaren mehr verdienen. Die Gleichung der Regressionsgerade für die Größe Y := 'Monatliches Nettogehalt in Euro' in Abhängigkeit von der Größe X := 'Anzahl Haare pro cm^2 Kopfhaut' wird mit $g(x) = -1.8x + 1980$ angegeben. Das durchschnittliche Nettogehalt der Studienteilnehmer in Euro betrug $\bar{y} = 1500$. Als weitere Angaben finden sich die beiden Varianzen $s_x^2 = 1630$ und $s_y^2 = 5820$.

- a) Kann man auf Grundlage der Daten davon ausgehen, dass zwischen den Größen 'Anzahl der Haare' und 'Gehalt' ein linearer Zusammenhang besteht?
- b) Wie viele Haare hatten die Studienteilnehmer durchschnittlich pro cm^2 Kopfhaut?
- c) Die Zeitschrift empfiehlt sich eine Glatze rasieren zu lassen um seine Gehaltschancen zu erhöhen. Was halten Sie von diesem Vorschlag (Begründung!)?

A 3 (10 Punkte)

Bitte untersuchen Sie mit Hilfe von R den Zusammenhang zwischen Körpergröße und Schuhgröße im Datensatz `Umfrage.RData`.

- a) Erstellen Sie ein Streudiagramm (*Größe* auf der x -Achse, *Schuhgröße* auf der y -Achse) (Funktion `plot` in R).
- b) Führen Sie eine lineare Regression für die beiden Größen durch (Funktion `lm` in R) und interpretieren Sie die Parameter der linearen Regression (`summary`).
- c) Fügen Sie dem Streudiagramm aus a) eine Regressionsgerade hinzu (`abline`) und interpretieren Sie die Befunde.