

Statistik und Wahrscheinlichkeitsrechnung (WS 2020/21)

Aufgabenblatt 3

zu bearbeiten bis: 22.11.2020 (23:59 Uhr). Abgabe in StudIP-Ordner "Abgabe Blatt 03".

Aufgabe 3.1 (Python: Autos)

Bearbeiten Sie das Notebook `01_cars.ipynb`. Notebooks laden Sie ebenso wie PDFs in den Abgabeordner hoch. Beachten Sie das übliche Benamungsschema.

Aufgabe 3.2 (Kovarianz)

Gegeben ist die bivariate Stichprobe:

$$(x_1, y_1), \dots, (x_5, y_5) = (0, 2), (5, 8), (3, 9), (-2, -1), (4, 2)$$

- a) Berechnen Sie die Kovarianzmatrix.
- b) Berechnen Sie die Korrelation zwischen den beiden Variablen x und y . Würden Sie die beiden Variablen als stark korreliert bezeichnen?
- c) Wie verändern sich Mittelwert und Kovarianzmatrix, wenn wir die Daten verschieben, d.h.

$$(x'_i, y'_i) := (x_i, y_i) + (-1, 2) \quad \text{für alle } i = 1, \dots, n$$

- d) Wie verändert sich (nur) die Kovarianzmatrix, wenn wir die Daten in x-Richtung strecken, d.h.

$$x'_i := 2 \cdot x_i \quad \text{für alle } i = 1, \dots, n$$

Begründen Sie bei (c) und (d) Ihre Antwort mit den in der Vorlesung vorgestellten Rechenregeln.

Aufgabe 3.3 (Lineare Regression)

Gegeben sei die zweidimensionale Stichprobe $(x_1, y_1), \dots, (x_4, y_4) = (-1, 1), (0, -2), (2, 2), (3, 6)$.

- a) Nähern Sie die Stichprobe mit dem Modell $\mathcal{M}_{a,b}(x) = ax + b$ an.
- b) Nähern Sie die Stichprobe mit dem Modell $\mathcal{M}_c(x) = x^2 + c$ an.
- c) Welches der beiden Modelle nähert die Stichprobe besser an? Begründen Sie.