

S2-P3-Statistik-LÖS-6

September 25, 2024

0.0.1 Aufgabe 6

Gegeben sind die Daten, welche in der Datei «daten_p1_6.csv» vorliegen. Die beiden Merkmale sind durch das Trennzeichen «;» voneinander getrennt.

- Lies diese Daten ein und lass dir die ersten fünf Datensätze anzeigen, damit du dir ein Bild über die Struktur der Daten und die Merkmale machen kannst.
- Berechne die beiden Koeffizienten der Regressions- respektive Ausgleichsgeraden.
- Wie gross ist der Zusammenhang (Korrelation) zwischen den beiden Merkmalen? Berechne dazu den Korrelationskoeffizienten und interpretiere das Resultat.

```
[3]: import scipy.stats as sci
import numpy as np
import pandas as pd

# a) Einlesen der Daten mit den Parametern:
#   1) sep:      Trennzeichen zwischen Spalten
#   2) decimal:  Dezimalzeichen
#   3) index_col: soll ein Index beim Importieren erstellt werden oder nicht
daten = pd.read_csv('Daten/daten_p1_6.csv', sep = ';', decimal = '.', index_col=False)

# Temperatur und Lebensdauer in Variablen X und Y übernehmen
X = daten['Temperatur_C']
Y = daten['Lebensdauer_h']

# b) Grad des Polynoms ist 1, also eine Gerade
koeffizienten = np.polyfit(X, Y, 1)      # die Koeffizienten (a, b) ermitteln
print('Koeffizienten =', koeffizienten)

# c) Korrelationskoeffizient
korr_koeff = sci.pearsonr(X, Y)
print("Korrelationskoeffizient:", "%.4f" % korr_koeff[0])
```

Koeffizienten = [-0.53866905 202.06601748]

Korrelationskoeffizient: -0.7963

- Da der Korrelationskoeffizient negativ ist, können wir erkennen, dass die Lebensdauer mit erhöhter Temperatur abnimmt.