Aufgabe 1:

Gegeben sei eine Stichprobe $x_1, \ldots, x_n$ von n unabhängigen Wiederholungen einer eindimensionalen Zufallsvariablen x. Drücken Sie die folgenden häufig verwendeten statistischen Kenngrößen in Matrix- bzw. Vektorschreibweise aus:

(a) Mittelwert

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

(b) Empirische Standardabweichung

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Aufgabe 2:

Gegeben sei eine Stichprobe $\mathbf{x}_1, \ldots, \mathbf{x}_n$ von n unabhängigen Wiederholungen einer p-dimensionalen Zufallsvariablen $\mathbf{x}$.

(a) Definieren Sie die zugehörige Datenmatrix $\mathbf{X}$.

(b) Drücken Sie den Mittelwertsvektor $\bar{\mathbf{x}}$ in Matrixschreibweise aus.

(c) Drücken Sie die empirische Kovarianzmatrix $\mathbf{S}$ in Matrixschreibweise aus.

Aufgabe 3:

Die Matrix

$$\mathbf{H} = \mathbf{I}_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\top \qquad \text{mit} \quad \mathbf{I}_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \quad \text{und} \quad \mathbb{1}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

wird als zentrierende Matrix oder Zentrierungsmatrix bezeichnet. Zeigen Sie, dass gilt:

(a) $\mathbf{H} = \mathbf{H}^\top$

(b) $\mathbf{H} \cdot \mathbf{H} = \mathbf{H}$

und berechnen Sie folgende Kenngrößen

(c) $\mathbf{H} \cdot \mathbf{x}$ mit $\mathbf{x} = (x_1, \ldots, x_n)^\top$ wie in Aufgabe 1

(d) $\mathbf{H} \cdot \mathbf{X}$ mit $\mathbf{X}$ wie in Aufgabe 2(a)