Applied Data Analysis

Exercise Sheet 2

Exercise 7

(a) Let $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p$ and $\Sigma \in \mathbb{R}^{p\times p}_{>0}$ (i.e. Σ positive definite).
Further, let $\emptyset \neq K, L \subseteq \{1,\ldots,p\}$ with $K \cap L = \emptyset$, $k := |K|$, $l := |L|$ and let $\Sigma_{K,L} := \operatorname{Cov}(\boldsymbol{X}_K, \boldsymbol{X}_L)$, $\Sigma_{KK|L} := \Sigma_{K,K} - \Sigma_{K,L}\,\Sigma_{L,L}^{-1}\,\Sigma_{K,L}'$.
Show (1) of Theorem I.2.8, i.e., the conditional distribution of $\boldsymbol{X}_K$ given $\boldsymbol{X}_L = \boldsymbol{x}_L$ for $\boldsymbol{x}_L \in \mathbb{R}^l$ is given by

$$P^{\boldsymbol{X}_K \,|\, \boldsymbol{X}_L = \boldsymbol{x}_L} \;=\; N_k\big(\boldsymbol{\mu}_K + \Sigma_{K,L}\,\Sigma_{L,L}^{-1}\,(\boldsymbol{x}_L - \boldsymbol{\mu}_L)\,,\, \Sigma_{KK|L}\big)$$

(with P denoting the underlying probability distribution).

(b) Show that $\boldsymbol{X}_K$ and $\boldsymbol{X}_L$ are stochastically independent if and only if the following equation holds for each $\boldsymbol{x}_L \in \mathbb{R}^l$:

$$\operatorname{E}(\boldsymbol{X}_K \,|\, \boldsymbol{X}_L = \boldsymbol{x}_L) \;=\; \operatorname{E}(\boldsymbol{X}_K)\,.$$

(c) Let $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, $\rho \in (-1, 1)$ and $\boldsymbol{X} = (X_1, X_2)' \sim N_2(\boldsymbol{\mu}, \Sigma)$ with

$$\boldsymbol{\mu} := \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \quad \text{and} \quad \Sigma := \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\,.$$

For $t \in \mathbb{R}$, derive the conditional distribution of X_2 given $X_1 = t$.

Hint to (a): Argue that, without loss of generality, one may assume $K \cup L = \{1,\ldots,p\}$ and therefore

$$\boldsymbol{X} = \begin{pmatrix}\boldsymbol{X}_K\\ \boldsymbol{X}_L\end{pmatrix}\,,\quad \boldsymbol{\mu} = \begin{pmatrix}\boldsymbol{\mu}_K\\ \boldsymbol{\mu}_L\end{pmatrix}\,,\quad \Sigma = \begin{pmatrix}\Sigma_{K,K} & \Sigma_{K,L}\\ \Sigma_{L,K} & \Sigma_{L,L}\end{pmatrix}\,,$$

i.e. $K = \{1,\ldots,k\}$, $L = \{k+1,\ldots,p\}$. Then, apply the results of Exercise 4 to factorize the density of $\boldsymbol{X} = \boldsymbol{X}_{K\cup L}$ into a marginal and a conditional part.

Exercise 8

Let $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, where Σ is an orthogonal projector (according to Definition I.1.8) and $\boldsymbol{\mu} \in \operatorname{Im}(\Sigma)$. Show:

$$\boldsymbol{X}'\boldsymbol{X} \;\sim\; \chi^2\Big(\operatorname{rank}(\Sigma)\,,\, \frac{1}{2}\,\boldsymbol{\mu}'\boldsymbol{\mu}\Big)\,.$$

Hint: Apply Theorem I.3.5.

Exercise 9

Let $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \sigma^2 I_p)$ with $\sigma > 0$ and let $A, B \in \mathbb{R}^{p\times p}$, where A is symmetric and $B\,A = 0_{p\times p}$. Show the following statements:

(a) $\boldsymbol{X}'\,A\,\boldsymbol{X}$ and $B\,\boldsymbol{X}$ are stochastically independent.

(b) If furthermore, B is also symmetric, then $\boldsymbol{X}'\,A\,\boldsymbol{X}$ and $\boldsymbol{X}'\,B\,\boldsymbol{X}$ are stochastically independent.

Solange die beiden BX und AX stochastisch unabhängig sind,
jegliche lineare Transformation von den beiden sind eben stochastisch unabhängig!
wenn wir annehmen, es gibt funtion f(.) sodass f(AX) -> X'AA+AX und g(.) g(BX)-> IpBX(Identity)
, dann sind f(AX) und g(BX) wiederum stochastisch unabhängig, weil die beiden BX und AX stoch. unabh sind
dasselbe Prinzip betrifft (b)

Exercise 10

Consider the linear model

$$\boldsymbol{Y} \;=\; B\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

of Definition I.4.2 with $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\,I_n$, where $\sigma > 0$. Furthermore, assume that $\det(B'B) > 0$. Then, the (estimated) residuals are defined by

$$\hat{\boldsymbol{\varepsilon}} \;:=\; (I_n - B\,(B'B)^{-1}\,B')\,\boldsymbol{Y}\;.$$

Calculate

(a) $\mathrm{E}(\hat{\boldsymbol{\varepsilon}})$,

(b) $\mathrm{Cov}(\hat{\boldsymbol{\varepsilon}})$.