Statistisk Modellering ST523, ST813
Exercise Session 4

All exercises should be prepared BEFORE the exercise session.

4 Exercises

Exercise 4.1

Distribution of quadratic forms and extensions to Cochrans theorem

Let $\mathbf{y}$ be an p-dimensional random vector such that $\mathbf{y} \sim N(\mu, \mathbf{V})$ for some $\mu \in \mathbb{R}^p$ and $\mathbf{V}$ a covariance matrix of rank p. Show that the following statements hold:

1. $(\mathbf{y} - \mu)^T \mathbf{V}^{-1}(\mathbf{y} - \mu) \sim \chi^2_p$, and
2. $\mathbf{y}^T \mathbf{V}^{-1} \mathbf{y} \sim \chi^2_{p,\lambda}$ where $\lambda = \mu^T \mathbf{V}^{-1} \mu$

   Hereby, $\chi^2_{p,\lambda}$ denotes the non-central χ^2-distribution with p degrees of freedom and non-centrality parameter $\lambda \geq 0$, i.e. a distribution that arises from the sum $\sum_{i=1}^p X_i^2$, where $X_1, \ldots, X_p$ are independent r.v.s such that $X_i \sim N(\mu_i, 1)$ and $\lambda = \sum_{i=1}^p \mu_i^2$.

Exercise 4.2

ML-estimator for linear model with Laplace distribution (Exercise 2.2 from Agresti)

In the linear model $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$, suppose that the errors are iid Laplace distributed with density $f(x) = \frac{1}{2b} \exp -\frac{|x|}{b}$ for $x \in \mathbb{R}$.

Show that the ML estimate minimizes $\sum_i |y_i - \hat{y}_i|$ where $\hat{y}_i$ denote the fitted values.

Exercise 4.3

Multivariate normal distribution

1. Let $\mathbf{X} = (X_1, \ldots, X_d)^T$ be a (multivariate) normally distributed random vector. Show that:

   $X_1, \ldots, X_d$ independent $\Leftrightarrow X_1, \ldots, X_d$ pairwise uncorrelated, i.e. $Cov(X_i, X_j) = 0, i \neq j$

   Note: Use the fact, that a multivariate normal distribution $N(\mu, \mathbf{V})$ is uniquely identified by its mean vector $\mu \in \mathbb{R}^d$ and variance matrix $V \in \mathbb{R}^{d \times d}$.

2. Let X and Y iid $\sim N(0,1)$. Show that $X + Y$ and $X - Y$ are independent.

Exercise 4.4

Change of units

Consider simple linear regression $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, $i = 1, \ldots n$. Does the line from the usual least squares fit depend on the measurement scale(s) of the variables? Formulate and prove the corresponding invariance property of the LS fit.

Hint: The measurement scale of a variable refers to whether e.g. a length is measured in units of mm, cm, m, in, ft, etc.