Bachelor in de wiskunde, Master of Statistics and Data Science
Prof. dr. Stefan Van Aelst

Statistical Inference and Data Analysis

Name: **Student No.**: January 2023

INSTRUCTIONS

- This is a closed book examination. You only need some pencils and a calculator. This is a written examination, so write down your complete answers and clearly indicate to which question each answer belongs. Only one answer is allowed for each question.
- The examination last maximum **3 hours**. You have to wait at least 1 hour before you can leave the room.
- Any use of communication devices is strictly forbidden and will be considered as a fraudulent activity.
- **Hand in** your examination before leaving the room. Make sure that your **name** is clearly written on **each page**. Fill in your **name and student number** on this page at the **start** of the examination.
- This examination counts for 14 out of 20 points. There are 3 questions. Each question counts for approximately one third of the examination grade.

QUESTIONS

1. Consider a random sample $X_1, \dots, X_n$ from the statistical model $(\mathbb{N}, 2^{\mathbb{N}}, \{\text{Geom}(p); p \in]0,1[\})$. For $X \sim \text{Geom}(p)$, the geometric distribution with success probability p, it holds that

$$P(X = k) = p(1-p)^k \qquad \text{for } k = 0, 1, 2, \dots,$$

so $F_X(k) = 1 - (1-p)^{k+1}$, $E[X] = \frac{1-p}{p}$, $\text{Var}(X) = \frac{1-p}{p^2}$ and $\phi_X(t) = \frac{p}{1-(1-p)e^{it}}$.

(a) Calculate the MLE for p.

(b) Use the properties of the MLE to determine its asymptotic distribution

(c) Construct an estimator for the asymptotic variance of the MLE and find its asymptotic distribution.

(d) Consider the hypothesis test problem

$$H_0 : p = 1/2$$
$$H_1 : p \neq 1/2$$

Construct the Wald and score tests for this problem. Are they different or not?

2. Consider $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and split $\boldsymbol{X}, \boldsymbol{\mu}, \boldsymbol{\Sigma}$ as follows

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \begin{matrix} \}q \\ \}p-q \end{matrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

The dependence between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ is contained in the matrix $\boldsymbol{\Sigma}_{12}$ (or equivalently $\boldsymbol{\Sigma}_{21}$). Evaluating the association between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ based on $\boldsymbol{\Sigma}_{12}$ is difficult. Instead, we can measure the maximal possible correlation between $\boldsymbol{a}^\top \boldsymbol{X}_1$ and $\boldsymbol{b}^\top \boldsymbol{X}_2$ over all $\boldsymbol{a} \neq \boldsymbol{0} \in \mathbb{R}^q$ and $\boldsymbol{b} \neq \boldsymbol{0} \in \mathbb{R}^{p-q}$

(a) Show that

$$\mathrm{Cor}(\boldsymbol{a}^\top \boldsymbol{X}_1, \boldsymbol{b}^\top \boldsymbol{X}_2) = \frac{\boldsymbol{a}^\top \boldsymbol{\Sigma}_{12} \boldsymbol{b}}{(\boldsymbol{a}^\top \boldsymbol{\Sigma}_{11} \boldsymbol{a})^{1/2} (\boldsymbol{b}^\top \boldsymbol{\Sigma}_{22} \boldsymbol{b})^{1/2}}$$

(b) Use appropriate transformations to show that

$$\max_{\boldsymbol{a} \neq \boldsymbol{0}, \boldsymbol{b} \neq \boldsymbol{0}} \mathrm{Cor}(\boldsymbol{a}^\top \boldsymbol{X}_1, \boldsymbol{b}^\top \boldsymbol{X}_2) = \max_{\boldsymbol{u} \neq \boldsymbol{0}, \boldsymbol{v} \neq \boldsymbol{0}} \frac{(\boldsymbol{u}^\top \boldsymbol{M} \boldsymbol{v})^2}{(\boldsymbol{u}^\top \boldsymbol{u})(\boldsymbol{v}^\top \boldsymbol{v})}$$

with $\boldsymbol{M} = \boldsymbol{\Sigma}_{11}^{-1/2} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1/2}$.

(c) Use the Cauchy-Schwarz inequality to show that

$$\max_{\boldsymbol{u} \neq \boldsymbol{0}, \boldsymbol{v} \neq \boldsymbol{0}} \frac{(\boldsymbol{u}^\top \boldsymbol{M} \boldsymbol{v})^2}{(\boldsymbol{u}^\top \boldsymbol{u})(\boldsymbol{v}^\top \boldsymbol{v})} \leq \max_{\boldsymbol{v} \neq \boldsymbol{0}} \frac{\boldsymbol{v}^\top \boldsymbol{M}^\top \boldsymbol{M} \boldsymbol{v}}{\boldsymbol{v}^\top \boldsymbol{v}}$$

with equality if $\boldsymbol{u} = c\boldsymbol{M}\boldsymbol{v}$ for some constant $c \neq 0$.

(d) Show that

$$\max_{\boldsymbol{a} \neq \boldsymbol{0}, \boldsymbol{b} \neq \boldsymbol{0}} \mathrm{Cor}(\boldsymbol{a}^\top \boldsymbol{X}_1, \boldsymbol{b}^\top \boldsymbol{X}_2) = \sqrt{\rho_1}$$

with ρ_1 the largest eigenvalue of the matrix $\boldsymbol{\Sigma}_{22}^{-1/2} \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1/2}$ and is reached for $\boldsymbol{b} = \boldsymbol{\Sigma}_{22}^{-1/2} \boldsymbol{e}_1$ and $\boldsymbol{a} = \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \boldsymbol{b}$ with $\boldsymbol{e}_1$ the eigenvector corresponding to ρ_1.

(e) Now, consider a random sample $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ distributed as $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Construct a likelihood ratio test for the null hypothesis of independence between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, i.e. $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$ based on the random sample.

3. We consider a dataset with a representative sample of 310 General Motor (GM) cars from the year 2015. The goal of this study is to estimate retail prices of future cars. The data provides information on the following variables:

Variable name	Description
Price	Retail price
Mileage	Number of miles the car has been driving
Make	Manufacturer of the car (Buick, Cadillac or Pontiac)

The following model was fit

```
Call:
lm(formula = log(Price) ~ Mileage + Make + Mileage * Make, data = cars)

Residuals:
     Min       1Q   Median       3Q      Max
-0.38369 -0.12319 -0.01977  0.10720  0.53910

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)            1.008e+01  6.687e-02 150.728  < 2e-16 ***
```

```
Mileage              -7.099e-06  3.102e-06  -2.289    0.0228 *
MakeCadillac          6.469e-01  8.357e-02   7.740 1.48e-13 ***
MakePontiac          -1.819e-01  7.814e-02  -2.328    0.0205 *
Mileage:MakeCadillac  1.578e-07  3.921e-06   0.040    0.9679
Mileage:MakePontiac   2.019e-06  3.653e-06   0.553    0.5808
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1911 on 304 degrees of freedom
F-statistic: 191.5 on 5 and 304 DF,  p-value: < 2.2e-16
```

The fitted model is shown in Figure 1

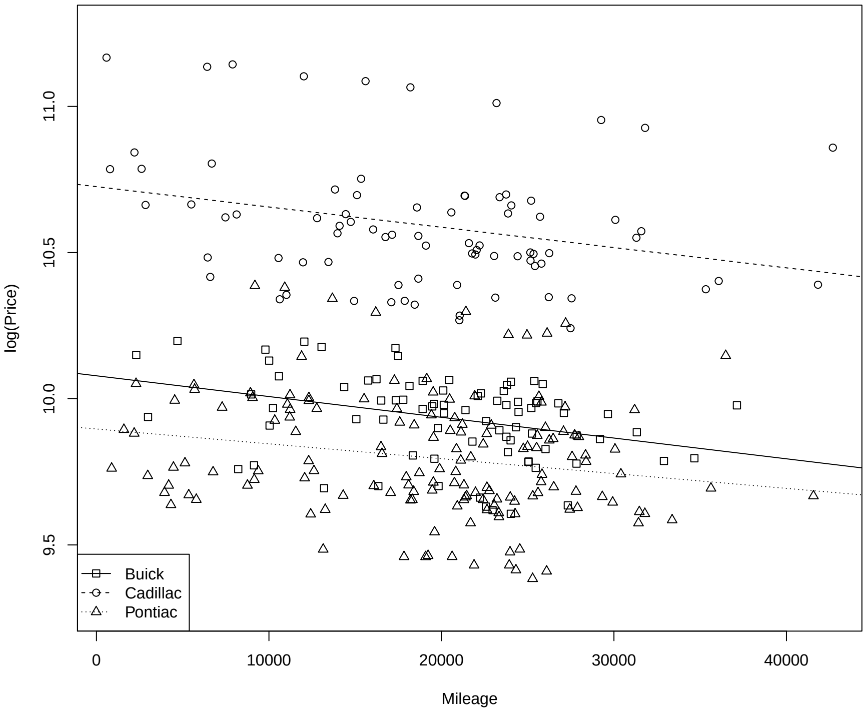

Figure 1: GM cars data: estimated regression lines of the model above

(a) Give the expression for the three regression lines shown in Figure 1.

(b) Given that the variance of the response (log(Price)) equals 46.08, calculate the coefficient of determination (R^2) and the adjusted coefficient of determination (R^2_a) for this model.

(c) Explain in words what the coefficient of determination is.

(d) Consider the following anova table

```
Analysis of Variance Table

Model 1: log(Price) ~ Mileage + Make
Model 2: log(Price) ~ Mileage + Make + Mileage * Make
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    306 11.121
2    304 11.102  2  0.018448      ?  0.777
```

Explain which hypothesis test is performed here and calculate the value of the F-statistic that is missing in the table. What is your conclusion based on this result?

(e) Let $\boldsymbol{X}$ denote the design matrix of the linear model. Given the matrix

$$(\boldsymbol{X}^t\boldsymbol{X})^{-1} = \begin{pmatrix} 0.12 & -0.00 & -0.12 & -0.12 & 0.00 & 0.00 \\ -0.00 & 0.00 & 0.00 & 0.00 & -0.00 & -0.00 \\ -0.12 & 0.00 & 0.19 & 0.12 & -0.00 & -0.00 \\ -0.12 & 0.00 & 0.12 & 0.17 & -0.00 & -0.00 \\ 0.00 & -0.00 & -0.00 & -0.00 & 0.00 & 0.00 \\ 0.00 & -0.00 & -0.00 & -0.00 & 0.00 & 0.00 \end{pmatrix}$$

construct simultaneous confidence intervals for β_2 and β_3. Motivate your method of choice and give the expressions for the confidence intervals (Fill in the available values in the formulas, but you do not have to do any calculations).

(f) Which assumptions need to hold for the above inference results to be valid? Explain for each of your assumptions how they can be verified based on the plots in Figure 2 and formulate your conclusion.

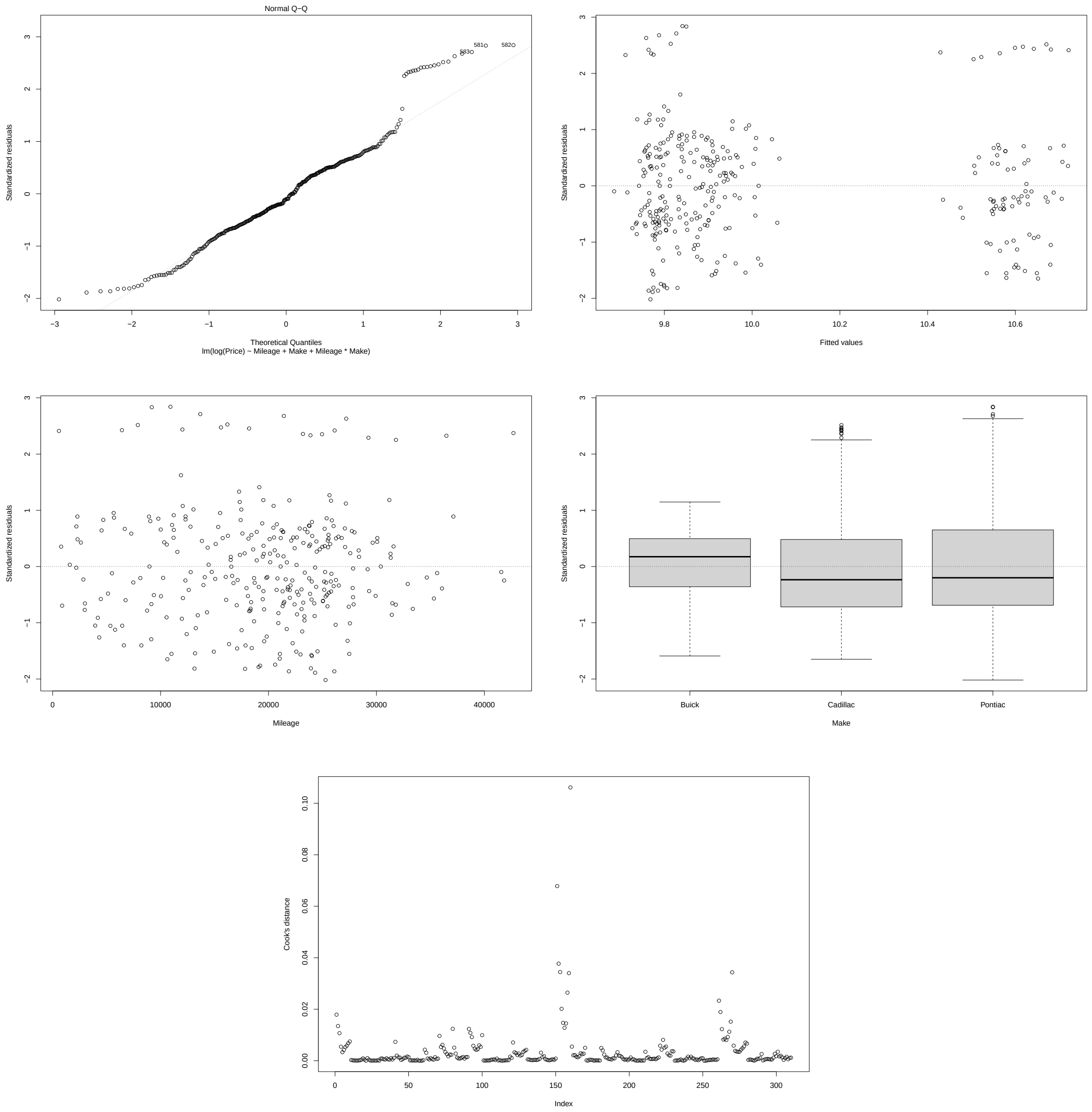

Figure 2: GM cars data: diagnostic plots

Formulas

- If T_n is an **asymptotically normal (distributed)** estimator for $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^k$:

$$\sqrt{n}(T_n - \boldsymbol{\theta}) \xrightarrow{D} \mathrm{N}_k(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \text{ if } n \to \infty$$

  and the function $g : \mathbb{R}^k \to \mathbb{R}$ is differentiable at $\boldsymbol{\theta}$ with gradient $\nabla g(\boldsymbol{\theta}) = (\frac{\partial g}{\partial t_1}\big|_{\boldsymbol{t}=\boldsymbol{\theta}}, \dots, \frac{\partial g}{\partial t_k}\big|_{\boldsymbol{t}=\boldsymbol{\theta}})^\top \neq \mathbf{0}$, then

$$\sqrt{n}(g(T_n) - g(\boldsymbol{\theta})) \xrightarrow{D} \mathrm{N}(0, \nabla g(\boldsymbol{\theta})^\top \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \nabla g(\boldsymbol{\theta})) \quad \text{if} \quad n \to \infty$$

- Hypothesis test $H_0 : R(\boldsymbol{\theta}) = 0$ vs $H_1 : R(\boldsymbol{\theta}) \neq 0$ with nonzero gradient $\nabla R(\hat{\boldsymbol{\theta}})$. $\hat{\boldsymbol{\theta}}$ is the MLE in the full model and $\hat{\boldsymbol{\theta}}_0$ the MLE in the restricted model under H_0, then
  - LR test

$$\text{Under } H_0 : D_n = 2\ln L(\hat{\boldsymbol{\theta}}; \boldsymbol{X}) - 2\ln L(\hat{\boldsymbol{\theta}}_0; \boldsymbol{X}) \xrightarrow{D} \chi^2_l \text{ if } n \to \infty$$

  - Wald test

$$\text{Under } H_0 : \; W_n = n(R(\hat{\boldsymbol{\theta}}))^\top [\nabla R(\hat{\boldsymbol{\theta}})^\top \mathbf{I}(\hat{\boldsymbol{\theta}})^{-1} \nabla R(\hat{\boldsymbol{\theta}})]^{-1} R(\hat{\boldsymbol{\theta}}) \xrightarrow{D} \chi^2_l \text{ if } n \to \infty$$

  - Score test

$$\text{Under } H_0 : R_n = \frac{1}{n}(S_n(\hat{\boldsymbol{\theta}}_0; X_1, \dots, X_n))^\top \mathbf{I}(\hat{\boldsymbol{\theta}}_0)^{-1} S_n(\hat{\boldsymbol{\theta}}_0; X_1, \dots, X_n) \xrightarrow{D} \chi^2_l \text{ if } n \to \infty$$

- Bayesian hypothesis test of $H_0 : \boldsymbol{\theta} \in \Theta_0 \subset \Theta$ vs $H_1 : \boldsymbol{\theta} \in \Theta_1 = \Theta \setminus \Theta_0$ is based on the posterior odds

$$\frac{\mathrm{P}(H_0|\boldsymbol{X} = \boldsymbol{x})}{\mathrm{P}(H_1|\boldsymbol{X} = \boldsymbol{x})} = \frac{\int_{\Theta_0} L(\boldsymbol{\theta}; \boldsymbol{x})\pi_0(\boldsymbol{\theta})\, d\boldsymbol{\theta}}{\int_{\Theta_1} L(\boldsymbol{\theta}; \boldsymbol{x})\pi_1(\boldsymbol{\theta})\, d\boldsymbol{\theta}} \times \frac{p_0}{p_1} = \text{Bayes factor} \times \text{prior odds}$$

- Multivariate normal distribution $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
  - Characteristic function: $\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \exp(i\boldsymbol{t}^\top \boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t})$
  - Conditional distribution: $\mathrm{P}_{\boldsymbol{X}_2|\boldsymbol{X}_1=\boldsymbol{x}_1}$ is a $(p-q)$-dimensional normal distribution with

$$\begin{aligned} \mathrm{E}[\boldsymbol{X}_2|\boldsymbol{X}_1 = \boldsymbol{x}_1] &= \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1) \\ \mathrm{Cov}[\boldsymbol{X}_2|\boldsymbol{X}_1 = \boldsymbol{x}_1] &= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \end{aligned}$$

  - Unbiased estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$:
    - $\overline{\boldsymbol{X}}_n \sim \mathrm{N}_p\left(\boldsymbol{\mu}, \frac{\boldsymbol{\Sigma}}{n}\right)$
    - $\boldsymbol{C}_n = \frac{1}{n-1}\sum_{i=1}^n (\boldsymbol{X}_i - \overline{\boldsymbol{X}}_n)(\boldsymbol{X}_i - \overline{\boldsymbol{X}}_n)^\top$ with $(n-1)\boldsymbol{C}_n = \boldsymbol{W}_n \sim W_p(\boldsymbol{\Sigma}, n-1)$
    - $\overline{\boldsymbol{X}}_n$ and $\boldsymbol{C}_n$ are independent
  - $T^2 = n(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu})^\top \boldsymbol{C}_n^{-1}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}) \sim T^2(p, n-1) = \frac{(n-1)p}{n-p} F_{p,n-p}$
  - $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ vs $H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$:

$$D_n = \mathrm{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}_n\right) - \ln\det\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}_n\right) - \frac{np}{2} \xrightarrow{D} \chi^2_{p(p+1)/2} \text{ if } n \to \infty$$

- Two sample Hotelling statistic

$$\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1} (\overline{\boldsymbol{X}}_{1n_1} - \overline{\boldsymbol{X}}_{2n_2} - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2))^\top \boldsymbol{C}_P^{-1} (\overline{\boldsymbol{X}}_{1n_1} - \overline{\boldsymbol{X}}_{2n_2} - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)) \sim \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{p, n_1+n_2-p-1}$$

  with $\boldsymbol{C}_P = \frac{(n_1-1)\boldsymbol{C}_{1n_1} + (n_2-1)\boldsymbol{C}_{2n_2}}{(n_1+n_2-2)}$

- $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ vs $H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$:

$$D_n = -2 \ln \Lambda_n = n \ln \det \left(\frac{n-2}{n} \boldsymbol{C}_P \right) - \sum_{j=1}^{2} n_j \ln \det \left(\frac{n_j - 1}{n_j} \boldsymbol{C}_{j n_j} \right) \xrightarrow{D} \chi^2_{p(p+1)/2}$$

- MANOVA: $H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_g$ vs $H_1 :$ not H_0:

$$D_n = n \ln \det(\boldsymbol{I}_p + \boldsymbol{W}_P^{-1} \boldsymbol{B}_P) \xrightarrow{D} \chi^2_{p(g-1)} \text{ if } n \to \infty$$

- $H_0 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_g$ vs $H_1 :$ not H_0:

$$D_n = -2 \ln \Lambda_n = n \ln \det \left(\frac{\boldsymbol{W}_P}{n} \right) - \sum_{l=1}^{g} n_l \ln \det \left(\frac{\boldsymbol{W}_{l\, n_l}}{n_l} \right) \xrightarrow{D} \chi^2_{(g-1)p(p+1)/2}$$

- Least squares for linear regression
  - $\hat{\boldsymbol{\beta}} \sim \mathrm{N}_{p+1} \left(\boldsymbol{\beta}, \sigma^2 \left(\tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{X}} \right)^{-1} \right)$
  - $S^2 = \frac{1}{n-p-1} (\boldsymbol{Y} - \tilde{\boldsymbol{X}} \hat{\boldsymbol{\beta}})^\top (\boldsymbol{Y} - \tilde{\boldsymbol{X}} \hat{\boldsymbol{\beta}})$ with $\frac{(n-p-1)S^2}{\sigma^2} \sim \chi^2_{n-p-1}$
  - $\hat{\boldsymbol{\beta}}$ and S^2 are independent
  - Linear hypothesis test $H_0 : \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0}$ vs $H_1 : \boldsymbol{C}\boldsymbol{\beta} \neq \boldsymbol{0}$ with $\boldsymbol{C} \in \mathbb{R}^{r \times (p+1)}$. $\hat{\boldsymbol{\beta}}$ is the estimate in the full model and $\hat{\boldsymbol{\beta}}_0$ is the estimate in the restricted model under H_0, then

$$F = \frac{(\mathrm{RSS}(\hat{\boldsymbol{\beta}}_0) - \mathrm{RSS}(\hat{\boldsymbol{\beta}}))/r}{\mathrm{RSS}(\hat{\boldsymbol{\beta}})/(n-p-1)} \sim F_{r,n-p-1}$$

  - Cook's distance: $D_i = \frac{(\hat{\boldsymbol{\beta}}_{-i} - \hat{\boldsymbol{\beta}})^\top \tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{X}} (\hat{\boldsymbol{\beta}}_{-i} - \hat{\boldsymbol{\beta}})}{(p+1)S^2}$
  - Mallow's C_p: $C_q = \frac{\mathrm{RSS}(\hat{\boldsymbol{\beta}}_1)}{S^2} + 2(q+1) - n$
  - $\mathrm{AIC} = \frac{\mathrm{RSS}(\hat{\boldsymbol{\beta}}_1)}{\sigma^2} + 2(q+1)$
  - $\mathrm{BIC} = \frac{\mathrm{RSS}(\hat{\boldsymbol{\beta}}_1)}{\sigma^2} + (q+1) \ln n$
- Principal components: $\boldsymbol{Y}_j = (\tilde{\boldsymbol{X}} - \boldsymbol{1}\overline{\boldsymbol{X}}_n^\top) \boldsymbol{e}_j \qquad j = 1, \ldots, p$
- Fisher LDA: discriminant directions are eigenvectors of $\boldsymbol{W}_P^{-1} \boldsymbol{B}_P$ corresponding to the largest eigenvalues
- Bayesian discriminant analysis: Assign $\boldsymbol{x}$ to $\min_{1 \leq k \leq g} d^2_{\boldsymbol{C}_P}(\boldsymbol{x}, \overline{\boldsymbol{X}}_{k\, n_k}) - 2 \ln \pi_k$
- EM algorithm
  1. E-step: Calculate $Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}_c) = \mathrm{E}[\ln L(\boldsymbol{\theta}, \tilde{\boldsymbol{Z}}) | \tilde{\boldsymbol{X}}, \hat{\boldsymbol{\theta}}_c]$
  2. M-step: $\hat{\boldsymbol{\theta}}_n$ is solution of $\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}_c)$

Matrix algebra

- Spectral decomposition of a symmetric square matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top = \sum_{j=1}^{p} \lambda_j \boldsymbol{e}_j \boldsymbol{e}_j^\top$$

  with $\boldsymbol{P} := [\boldsymbol{e}_1 \boldsymbol{e}_2 \ldots \boldsymbol{e}_p]$ an orthogonal matrix and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$

- For a p-dimensional positive definite matrix $\boldsymbol{A}$ and $\boldsymbol{d} \in \mathbb{R}^p$ it holds that

$$\max_{\boldsymbol{u} \neq \boldsymbol{0}} \frac{(\boldsymbol{u}^\top \boldsymbol{d})^2}{\boldsymbol{u}^\top \boldsymbol{A} \boldsymbol{u}} = \boldsymbol{d}^\top \boldsymbol{A}^{-1} \boldsymbol{d}$$

  with the maximum attained at $\boldsymbol{u} = c\boldsymbol{A}^{-1}\boldsymbol{d}$ for any constant $c \neq 0$