Next Item

Let the matrices $B, X \in \mathbb{R}^{3\times 2}$ be given by

$$B = \begin{pmatrix} b_1 & b_2 \\ 3 & 3 \\ 1 & 1 \end{pmatrix}, b_1, b_2 \in \mathbb{R}, \quad X = \begin{pmatrix} 2 & 2 \\ 0 & 1 \\ x_1 & x_2 \end{pmatrix}, x_1, x_2 \in \mathbb{R}.$$

Consider the **two** (normal) linear models

$$\boldsymbol{Y} = B\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1) \qquad \text{and} \qquad \boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

with $\boldsymbol{\beta} = (\beta_1, \beta_2)' \in \mathbb{R}^2$ and $\boldsymbol{\varepsilon} \sim N_3(\mathbf{0}, \sigma^2\Sigma), \sigma^2 > 0, \Sigma \in \mathbb{R}^{3\times 3}_{>0}$.

Give your answers to the tasks below by filling in the blanks. Results that are numerical values should, if necessary, be rounded to two decimals. In case of multiple solutions, please order your solutions from smallest to largest and separate them by " & " (without

$$Y = B\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1) \qquad \text{and} \qquad Y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

with $\boldsymbol{\beta} = (\beta_1, \beta_2)' \in \mathbb{R}^2$ and $\boldsymbol{\varepsilon} \sim N_3(\mathbf{0}, \sigma^2\Sigma), \sigma^2 > 0, \Sigma \in \mathbb{R}^{3\times 3}_{>0}$.

Give your answers to the tasks below by filling in the blanks. Results that are numerical values should, if necessary, be rounded to two decimals. In case of multiple solutions, please order your solutions from smallest to largest and separate them by " & " (without quotations marks but with spaces, e.g.: 3 & 5). If such a value **does not exist**, then type "NA" (without quotation marks) instead into the blank. If the value can be **chosen arbitrarily**, then type "R" (without quotation marks) instead into the blank.

7 of 7 points

1 of 1 point

Let $\Sigma = I_3$ be the identity matrix.

Assume $b_1 = b_2 = b \in \mathbb{R}$. Find the largest set of values for $b \in \mathbb{R}$ such that the LSE (least-squares estimator) $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ in **model (1)** is unique.

[redacted]

1 of 1 point

Assume $x_1 = x_2 = x \in \mathbb{R}$. Find the largest set of values for $x \in \mathbb{R}$ such that the LSE (least-squares estimator) $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ in **model (2)** is unique.

[redacted]

in **model (2)** is unique.

[redacted]

1 of 1 point

Let $b_1 = b$, $b_2 = 2b - 1$ with $b \in \mathbb{R}$. Find the largest set of values for $b \in \mathbb{R}$ such that the LSE (least-squares estimator) $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ in **model (1)** is **not** unique.

[redacted]

1 of 1 point

Let $b_1 = b$ and $b_2 = -2b + 1$ with $b \in \mathbb{R}$. Furthermore, consider the function $g : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$g(\boldsymbol{\beta}) = \begin{pmatrix} \beta_1 - \beta_2 \\ \beta_1 + \beta_2 \end{pmatrix}, \quad \boldsymbol{\beta} \in \mathbb{R}^2.$$

Find the largest set of values for $b \in \mathbb{R}$ such that $g(\boldsymbol{\beta})$ is **not** identifiable in model (1).

[redacted]

1 of 1 point

Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \end{pmatrix}$$

Find the largest set of values for $b \in \mathbb{R}$ such that $g(\boldsymbol{\beta})$ is **not** identifiable in model (1).

[redacted]

1 of 1 point

Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and consider the model

$$A\boldsymbol{Y} = AX\boldsymbol{\beta} + A\boldsymbol{\varepsilon}, \quad (3)$$

which results from model (2) by considering only the first two observations Y_1, Y_2. Assume that

$$\mathrm{Cov}(A\boldsymbol{\varepsilon}) = \mathrm{Cov}\left(\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}\right) = \frac{2}{3} \cdot \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Determine the covariance matrix

$$V = \begin{pmatrix} v_1 & v \\ v & v_2 \end{pmatrix}$$

Determine the covariance matrix

$$V = \begin{pmatrix} v_1 & v \\ v & v_2 \end{pmatrix}$$

of $g(\hat{\boldsymbol{\beta}}) = (\hat{\beta}_1 - \hat{\beta}_2, \hat{\beta}_1 + \hat{\beta}_2)'$, where $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$ is the MLE (maximum likelihood estimator) of $\boldsymbol{\beta}$ in model (3).

$v_1 =$

[redacted]

1 of 1 point

$v_2 =$

[redacted]

1 of 1 point

$v =$

[redacted]

Previous Item
Next Item

Consider the normal linear model

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

with

$$B = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \in \mathbb{R}^2, \qquad \boldsymbol{\varepsilon} \sim N_3(\mathbf{0}, \sigma^2 I_3),\ \sigma^2 > 0.$$

Denote by $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$ the LSE of $\boldsymbol{\beta}$. Find the missing numerical values with a precision of two decimal places.

2.75 of 5 points

Denote by $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$ the LSE of $\boldsymbol{\beta}$. Find the missing numerical values with a precision of two decimal places.

2.75 of 5 points

0.5 points

An upper $(1-\alpha)$-confidence interval for the parameter β_1 is given by

$$I_{\beta_1} = \left[\hat{\beta}_1 - q(\alpha) \cdot \|\boldsymbol{Y} - B\hat{\boldsymbol{\beta}}\| \cdot d, \infty\right)$$

with appropriate choice of $q(\alpha)$ and d; $q(\alpha)$ denotes a quantile of an appropriate distribution; $\|\mathbf{z}\| = \sqrt{\mathbf{z}'\mathbf{z}}$.

For $\alpha = 0{,}05$, determine the values of $q(\alpha)$ and d. [redacted]

$d =$

--- ✕ 0.63

0.5 points [redacted]

$q(\alpha) =$

--- ✕ 6.31

0.75 points

0.75 points

Similarly, an upper $(1-\alpha)$-confidence interval for $\gamma = \beta_1 - 2\beta_2$ is given by

$$I_\gamma = \left[\hat{\gamma} - q^*(\alpha) \cdot ||\boldsymbol{Y} - B\hat{\boldsymbol{\beta}}|| \cdot d^*, \infty\right).$$

For $\alpha = 0{,}1$, determine the values of $q^*(\alpha)$ and d^*.

[redacted]

$d^* =$

--- ✕ 1.1

0.5 [redacted]

$q^*(\alpha) =$

--- ✕ 3.08

0.25 of 0.25 points

Consider the testing problem

$$H_0\colon \beta_2 = 0 \quad \longleftrightarrow \quad H_1\colon \beta_2 \neq 0.$$

Then, there exists an α-level statistical test for H_0 whose decision rule can be formulated as

$$\text{Reject } H_0 \text{ if } \frac{\boldsymbol{Y}' A_0 \boldsymbol{Y}}{\boldsymbol{Y}' A \boldsymbol{Y}} > c(\alpha)$$

0.25 of 0.25 points

Consider the testing problem

$$H_0\colon \beta_2 = 0 \quad \longleftrightarrow \quad H_1\colon \beta_2 \neq 0.$$

Then, there exists an α-level statistical test for H_0 whose decision rule can be formulated as

$$\text{Reject } H_0 \text{ if } \frac{\boldsymbol{Y}' A_0 \boldsymbol{Y}}{\boldsymbol{Y}' A \boldsymbol{Y}} > c(\alpha)$$

for some appropriate **orthogonal projectors** A_0, A, and an appropriately chosen critical value $c(\alpha)$, respectively.

Find the design matrix $B_0 = (x_1, x_2, x_3)'$ associated with the null hypothesis H_0.

The entries x_1, x_2, x_3 of B_0 are

[redacted]

$x_1 =$

1 ✓

0.25 of 0.25 poi[redacted]

$x_2 =$

1 ✓

0.25 of 0.25 points

$x_1 =$

1 ✓ [redacted]

0.25 of 0.25 points

$x_2 =$

1 ✓ [redacted]

0.25 of 0.25 points

$x_3 =$

0 ✓ [redacted]

0.25 of 0.25 points

List the diagonal elements of $A_0 = (a_{ij}^{(0)})_{i,j}$ and $A = (a_{ij})_{i,j}$.

For the diagonal elements $a_{11}^{(0)}, a_{22}^{(0)}, a_{33}^{(0)}$ of A_0 we have

$a_{11}^{(0)} =$

0.17 ✓ [redacted]

0.25 of 0.25 points

$a_{22}^{(0)} =$

0.25 of 0.25 points

List the diagonal elements of $A_0 = (a_{ij}^{(0)})_{i,j}$ and $A = (a_{ij})_{i,j}$.

For the diagonal elements $a_{11}^{(0)}, a_{22}^{(0)}, a_{33}^{(0)}$ of A_0 we have

$a_{11}^{(0)} =$

0.17 ✓ [redacted]

0.25 of 0.25 points

$a_{22}^{(0)} =$

0.17 ✓ [redacted]

0.25 of 0.25 points

$a_{33}^{(0)} =$

0.67 ✓ [redacted]

0.25 of 0.25 points

For the diagonal elements a_{11}, a_{22}, a_{33} of A we have

$a_{11} =$

0.33 ✓ [redacted]

0.67 ✓ [redacted]

0.25 of 0.25 points

For the diagonal elements a_{11}, a_{22}, a_{33} of A we have

$a_{11} =$

0.33 ✓ [redacted]

0.25 of 0.25 points

$a_{22} =$

0.33 ✓ [redacted]

0.25 of 0.25 points

$a_{33} =$

0.33 ✓ [redacted]

0.5 of 0.5 points

For $\alpha = 0{,}01$, find the value of the critical value $c(\alpha)$.

$c(\alpha) =$

4052.181 ✓ [redacted]

Consider measurements $x_1 = 1, x_2 = -1, x_3 = 2, x_4 = 0 \in \mathbb{R}$ and the polynomial regression model

$$Y = \beta_0 + \beta_1 X^3.$$

Give the missing numerical values with a precision of two decimal places. **1.25 of 2 points**

0.25 of 0.25 points

Consider the testing problem

$$H_0 \colon \beta_0 = 0 \quad \longleftrightarrow \quad H_1 \colon \beta_0 \neq 0.$$

Determine the entries of the design matrix $B_0 = (b_1, b_2, b_3, b_4)'$ associated with H_0.

$b_1 =$ [redacted]

0.25 points

$b_2 =$ [redacted]

[redacted]

0.25 points

$b_3 =$

[redacted]

0.25 points

$b_4 =$

[redacted]

1 of 1 point

The decision rule in terms of quantiles of the t-distribution can be formulated as

$$\text{Reject } H_0 \text{ if } \left|\frac{\hat{\beta}_0}{c}\right| > t_{1-\alpha/2}(\mathrm{df}),$$

for an appropriate constant $c \in \mathbb{R}$.

[redacted]

Determine the degrees of freedom df of the t-distribution.

$\mathrm{df} =$

2 ✓

The questions of the second E-Test are based on the tasks of R-Laboratories 4 and 5. Solutions should be given with a precision of 4 digits; so **please round your results to 4 digits**. The names of the data frames, variables etc. are the same as in the corresponding tasks and the solution of these tasks in the RWTHmoodle space. **Notice that the decimal separator is "," (without quotation marks).**

Task 13 (R-Laboratory 4) *Hint*: Please pay attention to the random number generation process. For some questions, the source code has to be changed. When answering the following questions, always carry out the whole Task 13 and remember to change the parameters back to the ones required for the task sheet afterwards.

4 of 5 points

0.5 points

What is the proportion of cases in which the simple model performs better than the correct model according to a comparison of `vec.delta.simple` and `vec.delta.correct`?

38[redacted]

0.5 of 0.5 points

Change the seed in Task 13 (a) (R-Lab 4) to 10. Provide the mean and the standard deviation of both `vec.delta.simple` and `vec.delta.correct`.

mean of `vec.delta.simple`

Change the seed in Task 13 (a) (R-Lab 4) to 10. Provide the mean and the standard deviation of both `vec.delta.simple` and `vec.delta.correct`.

mean of `vec.delta.simple`

[redacted]

0.5 of 0.5 points

standard deviation of `vec.delta.simple`

1.1124 [redacted]

0.5 of 0.5 points

mean of `vec.delta.correct`

[redacted]

0.5 of 0.5 points

standard deviation of `vec.delta.correct`

1.1 [redacted]

0.5 of 0.5 points

Let the seed be set back to 2020. Change Task 13 (b)(i) (R-Lab 4) such that the values of X are generated from a uniform distribution on the interval $[-50, 100]$. Provide the mean and the standard deviation of both standard deviation of `vec.delta.correct`

[redacted]

0.5 of 0.5 points

Let the seed be set back to 2020. Change Task 13 (b)(i) (R-Lab 4) such that the values of X are generated from a uniform distribution on the interval $[-50, 100]$. Provide the mean and the standard deviation of both `vec.delta.simple` and `vec.delta.correct`.

mean of `vec.delta.simple`

2.0167 ✓ [redacted]

0.5 of 0.5 points

standard deviation of `vec.delta.simple`

0.9746 ✓ [redacted]

0.5 of 0.5 points

mean of `vec.delta.correct`

3.8398 ✓ [redacted]

0.5 of 0.5 points

standard deviation of `vec.delta.correct`

[redacted]

0.5 of 0.5 points

mean of `vec.delta.correct`

[redacted]

0.5 of 0.5 points

standard deviation of `vec.delta.correct`

[redacted]

0.5 points

Keep the seed set to 2020 and generate the values of X as in Task 13 (uniform distribution on $[0, 100]$). Change Task 13 (b)(i) (R-Lab 4): generate $N = 1500$ observations. What is the proportion of cases in which the correct model performs better than the simple model according to a comparison of `vec.delta.simple` and `vec.delta.correct` ?

50.6227 ✕ 0.42

[redacted]

The questions of the second E-Test are based on the tasks of R-Laboratories 4 and 5. Solutions should be given with a precision of 4 digits; so **please round your results to4 digits**. The names of the data frames, variables etc. are the same as in the corresponding tasks and the solution of these tasks in the RWTHmoodle space. **Notice that the decimal separator is "," (without quotation marks).**

Task 14 (R-Lab 4) 2 of 2 points

0.5 of 0.5 points

What is the value of the adjusted R-squared of the model in Task 14 (b) (R-Lab 4) ?

0.455 [redacted]

0.5 of 0.5 points

What is the p-value of the Shapiro-Wilk test in Task 14 (c) (R-Lab 4) ?

[redacted]

0.5 of 0.5 points

What is the estimate of parameter b of model (++) (see Task 14 (b) of R-Lab 4) and what is its standard error?

estimate of b

[redacted]

0.5 of 0.5 points

What is the p-value of the Shapiro-Wilk test in Task 14 (c) (R-Lab 4) ?

[redacted]

0.5 of 0.5 points

What is the estimate of parameter b of model (++) (see Task 14 (b) of R-Lab 4) and what is its standard error?

estimate of b

[redacted]

0.5 of 0.5 points

standard error

[redacted]

Previous
Item

The questions of the second E-Test are based on the tasks of R-Laboratories 4 and 5. Solutions should be given with a precision of 4 digits; so **please round your results to 4 digits**. The names of the data frames, variables etc. are the same as in the corresponding tasks and the solution of these tasks in the RWTHmoodle space. **Notice that the decimal separator is "," (without quotation marks).**

Task 17 (R-Lab 5). For this task, the dataset `Solar` introduced in Task 15 is required. **1 of 4 points**

0.5 points

Calculate 80%-confidence intervals for the three differences $\alpha_{i+1} - \alpha_i$ $(i = 1, .., 3)$ of the factor levels $1, \ldots, 4$ of `batch` treated as factor as in Task 17.

lower bound of CI for factor level difference of levels 1 and 2

[redacted]

0.5 points

0.5 points

Calculate 80%-confidence intervals for the three differences $\alpha_{i+1} - \alpha_i \ (i = 1, .., 3)$ of the factor levels $1, \ldots, 4$ of `batch` treated as factor as in Task 17.

lower bound of CI for factor level difference of levels 1 and 2

[redacted]

0.5 points

upper bound of CI for factor level difference of levels 1 and 2

[redacted]

0.5 points

lower bound of CI for factor level difference of levels 2 and 3

[redacted]

0.5 points

upper bound of CI for factor level difference of levels 2 and 3

[redacted]

0.5 points

lower bound of CI for factor level difference of levels 3 and 4

lower bound of CI for factor level difference of levels 2 and 3

[redacted]

0.5 points

upper bound of CI for factor level difference of levels 2 and 3

[redacted]

0.5 points

lower bound of CI for factor level difference of levels 3 and 4

[redacted]

0.5 points

upper bound of CI for factor level difference of levels 3 and 4

[redacted]

1 of 1 point

Fit the model with the `Solar` data using the function `lm()`. What is the p-value of the Shapiro-Wilk test?

[redacted]