Aufgabe 1

2 Punkte

Consider $X_1, \ldots, X_n$, $n \in \mathbb{N}$, independently and identically distributed random variables with probability density and cumulative distribution function, respectively, given by

$$f(x;\alpha,\gamma) = \begin{cases} \alpha\gamma^{\alpha}x^{-(\alpha+1)}, & \gamma \leq x, \\ 0, & \text{else}, \end{cases}, \qquad F(x;\alpha,\beta) = \begin{cases} 1-\gamma^{\alpha}x^{-\alpha}, & \gamma \leq x \\ 0, & \text{else} \end{cases} \tag{1}$$

with real parameters $\alpha > 1, \gamma > 0$; that is, the (full) likelihood of the parameters for a sample $\boldsymbol{x} = (x_1, \ldots, x_n)$ is given by

$$L(\alpha,\gamma;\boldsymbol{x}) = \begin{cases} \alpha^n\gamma^{n\alpha}\prod_{i=1}^{n} x_i^{-(\alpha+1)}, & \gamma \leq x_{(1)}, \\ 0, & \text{else}, \end{cases}$$

where $x_{(1)} = \min\{x_1, \ldots, x_n\}$ denotes the minimum of the observed sample values.

Consider the profile likelihood approach, where γ is the nuisance parameter, while α is the parameter of interest. Denote by $\hat{\gamma}(\alpha)$ the nuisance parameter estimate of γ for fixed α.

Suppose we have observed the sample

$$x_1 = 6, x_2 = 4, x_3 = 3$$

of size $n = 3$.

Find the missing numerical values with a precision of two decimals. **2 Punkte**

1 Punkt

(a) Calculate $\hat{\gamma}(\alpha)$ and give the value of $\hat{\gamma}(\alpha)$ for $\alpha = 2$.

3

1 Punkt

(b) Calculate the profile maximum likelihood estimate for the parameter α based on the observed sample given above (with a precision of two decimals).

$\hat{\alpha}^{profile} =$

$$L(\alpha, \gamma; \mathbf{x}) = \begin{cases} & \\ 0, & \text{else,} \end{cases}$$

where $x_{(1)} = \min\{x_1, \ldots, x_n\}$ denotes the minimum of the observed sample values.

Consider the profile likelihood approach, where γ is the nuisance parameter, while α is the parameter of interest. Denote by $\hat{\gamma}(\alpha)$ the nuisance parameter estimate of γ for fixed α.

Suppose we have observed the sample

$$x_1 = 6, x_2 = 4, x_3 = 3$$

of size $n = 3$.

Find the missing numerical values with a precision of two decimals. **2 Punkte**

1 Punkt

(a) Calculate $\hat{\gamma}(\alpha)$ and give the value of $\hat{\gamma}(\alpha)$ for $\alpha = 2$.

3

1 Punkt

(b) Calculate the profile maximum likelihood estimate for the parameter α based on the observed sample given above (with a precision of two decimals).

$\hat{\alpha}^{profile} =$

3,06

Nächste Aufgabe

Alle Antworten wurden gespeichert!

Vorherige Aufgabe
Nächste Aufgabe

Aufgabe 2

3 Punkte

Let $Y \mid X = k \sim \mathcal{P}(k)$ be a Poisson distributed random variable conditional on the value of X. Further suppose that

$$P(X = k) = (1-p)^k p, \quad k \in \mathbb{N}_0,$$

for $p \in (0,1)$ with $\mathrm{E}(X) = \frac{1-p}{p}$ and $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

Find the missing numerical values with a precision of two decimals. **3 Punkte**

1 Punkt

(a) Derive the expectation and variance of Y for $p = \frac{1}{3}$.

$\mathrm{E}(Y) =$

2

1 Punkt

$\mathrm{Var}(Y) =$

8

1 Punkt

(b) Denote by $p_{\min}$ the smallest $p \in (0,1)$, such that $P(Y = 0) \geq \frac{1}{3}$. Find $p_{\min}$ with a precision of two decimals.

Hint: Use the law of total probability.

$p_{\min} =$

0,24

$$P(X = k) = (1-p)^k p, \quad k \in \mathbb{N}_0,$$

for $p \in (0,1)$ with $\mathrm{E}(X) = \frac{1-p}{p}$ and $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

Find the missing numerical values with a precision of two decimals. **3 Punkte**

1 Punkt

(a) Derive the expectation and variance of Y for $p = \frac{1}{3}$.

$\mathrm{E}(Y) =$

2

1 Punkt

$\mathrm{Var}(Y) =$

8

1 Punkt

(b) Denote by $p_{\min}$ the smallest $p \in (0,1)$, such that $P(Y = 0) \geq \frac{1}{3}$. Find $p_{\min}$ with a precision of two decimals.

Hint: Use the law of total probability.

$p_{\min} =$

0,24

Vorherige Aufgabe

Nächste Aufgabe

Alle Antworten wurden gespeichert!

Vorherige Aufgabe
Nächste Aufgabe

Aufgabe 3

5 Punkte

Consider a linear model with design matrix

$$X = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{3}} \\ -1 & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

and parameter vector $\boldsymbol{\beta} = (\beta_1, \beta_2)'$. Denote the lasso and ridge regression estimators as solutions to the objective function

$$\min_{\beta \in \mathbb{R}^2} \left\{ \frac{1}{2} ||\boldsymbol{y} - X\boldsymbol{\beta}||^2 + \lambda ||\boldsymbol{\beta}||_q^q \right\}$$

for $q = 1$ or $q = 2$, respectively, as $\hat{\boldsymbol{\beta}}^{lasso}$ and $\hat{\boldsymbol{\beta}}^{ridge}$. Suppose we have observed

$$\boldsymbol{y} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Hint: The following taks have to be solved theoretically (see related theoretical exercise). In particular, *do not* use R, since the objective function in `glmnet` - and thus the tuning paramater λ - are scaled differently than above.

Find the missing numerical values with a precision of two decimals. **5 Punkte**

1 Punkt

(a) Find the smallest value $\lambda_{\min}$, say, for the tuning parameter λ, such that *at least* one of the lasso estimates $\hat{\beta}_i^{lasso}$, $i \in \{1, 2\}$, equals exactly zero (with a precision of two decimal places).

Find the missing numerical values with a precision of two decimals. 5 Punkte

1 Punkt

(a) Find the smallest value $\lambda_{\min}$, say, for the tuning parameter λ, such that $at\ least$ one of the lasso estimates $\hat{\beta}_i^{lasso}$, $i \in \{1, 2\}$, equals exactly zero (with a precision of two decimal places).

$\lambda_{\min} =$

0,5

1 Punkt

(b) Calculate the corresponding ridge regression estimates for the value of $\lambda_{\min}$ derived in (a) with a precision two decimal places.

(i)The corresponding ridge regression estimate of the first component is

$\hat{\beta}_1^{ridge} =$

0,33

1 Punkt

(ii)The corresponding ridge regression estimate of the second component is

$\hat{\beta}_2^{ridge} =$

-0,58

1 Punkt

(c) Suppose the underlying model is normal, that is

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

with $\boldsymbol{\varepsilon} \sim N_2(\mathbf{0}, 4 \cdot I_2)$, where I_2 denotes the 2-dimensional identity matrix.

Consider the mean squared error (MSE) of the ridge estimator of the first component β_1 of $\boldsymbol{\beta}$,

$$\mathrm{MSE}(\hat{\beta}_1^{ridge}) = \mathrm{bias}(\hat{\beta}_1^{ridge})^2 + \mathrm{Var}(\hat{\beta}_1^{ridge}).$$

Let $\lambda = 1$ and denote by

$$(\beta_1^{lower}, \beta_1^{upper})$$

1 Punkt

(c) Suppose the underlying model is normal, that is

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

with $\boldsymbol{\varepsilon} \sim N_2(\mathbf{0}, 4 \cdot I_2)$, where I_2 denotes the 2-dimensional identity matrix.

Consider the mean squared error (MSE) of the ridge estimator of the first component β_1 of $\boldsymbol{\beta}$,

$$\mathrm{MSE}(\hat{\beta}_1^{ridge}) = \mathrm{bias}(\hat{\beta}_1^{ridge})^2 + \mathrm{Var}(\hat{\beta}_1^{ridge}).$$

Let $\lambda = 1$ and denote by

$$(\beta_1^{lower}, \beta_1^{upper})$$

the interval of values for β_1, for which

$$\mathrm{MSE}(\hat{\beta}_1^{ridge}) < \mathrm{Var}(\hat{\beta}_1^{LS}),$$

i.e., the MSE of $\hat{\beta}_1^{ridge}$ is smaller than the variance of the least squares estimator $\hat{\beta}_1^{LS}$ of β_1. Find this interval (with a precision of two decimal places).

Hint: Consult Exercise 27(a) and Theorem I.4.25 of the lecture to derive the distribution of the estimator $\hat{\beta}_1^{LS}$.

$\beta_1^{lower} =$

-3

1 Punkt

$\beta_1^{upper} =$

3

Vorherige Aufgabe
Nächste Aufgabe
ÜBERSICHT
1 2 3 4 5
ABGABE

Vorherige Aufgabe
Nächste Aufgabe

Aufgabe 4

8 Punkte

Please provide numbers in the requested precision within each question. The use of different precision is evaluated as wrong!

Load the dataset `Boston` from the package `MASS` into your `R` workspace. Use `?Boston` to get information about the dataset. Let n denote the number of observations included in this dataset.

(a) Set the seed to 2021. Split the dataset randomly into a training and a test dataset where the training dataset contains $n - 100$ observations. Fit a linear model based on the training data where `medv` is the response variable and `crim, zn, indus, nox, rm, age, dis, tax, ptratio, black, lstat` are the explanatory variables. What is the mean value of the resulting fitted values?

1 Punkt

1 Punkt

mean value **(requested precision: 4 digits)**

22,3111

(b) Fit the penalized regression model with Lasso with tuning parameter $\lambda = 0.5$. What are the resulting estimates for the intercept and the coefficient of `dis`?

2 Punkte

1 Punkt

estimate of intercept **(requested precision: 4 digits)**

10,1197

(b) Fit the penalized regression model with Lasso with tuning parameter $\lambda = 0.5$. What are the resulting estimates for the intercept and the coefficient of `dis`? **2 Punkte**

estimate of intercept **(requested precision: 4 digits)** — 1 Punkt

10,1197

estimate of coefficient of `dis` **(requested precision: 4 digits)** — 1 Punkt

-0,0615

(c) Fit the penalized regression model with Lasso using cross validation (CV). What are the values for the minimum of λ (denoted by λ_{min}) and λ_{1-SE}? **2 Punkte**

value for λ_{min} **(requested precision: 4 digits)** — 1 Punkt

0,0227

value for λ_{1-SE} **(requested precision: 4 digits)** — 1 Punkt

1,0293

(d) Fit the penalized regression model with Lasso with tuning parameter λ_{1-SE} from (c). What is the minimum value of the resulting estimated coefficients ? **1 Punkt**

minimum value of resulting estimated coefficients **(requested precision: 4 digits)** — 1 Punkt

-0,5996

(e) Fit the penalized regression model with Lasso with tuning parameter λ_{min} from (c). What is the minimum value of the resulting estimated coefficients ? **1 Punkt**

(c) Fit the penalized regression model with Lasso using cross validation (CV). What are the values for the minimum of λ (denoted by λ_{min}) and λ_{1-SE}? **2 Punkte**

1 Punkt

value for λ_{min} **(requested precision: 4 digits)**

0,0227

1 Punkt

value for λ_{1-SE} **(requested precision: 4 digits)**

1,0293

(d) Fit the penalized regression model with Lasso with tuning parameter λ_{1-SE} from (c). What is the minimum value of the resulting estimated coefficients ? **1 Punkt**

1 Punkt

minimum value of resulting estimated coefficients **(requested precision: 4 digits)**

-0,5996

(e) Fit the penalized regression model with Lasso with tuning parameter λ_{min} from (c). What is the minimum value of the resulting estimated coefficients ? **1 Punkt**

1 Punkt

minimum value of resulting estimated coefficients **(requested precision: 4 digits)**

-10,6274

(f) Test the models fitted in (d) and (e) on the test dataset and compute the values for predicted residual sum of squares (PRESS). Which model has the lowest value of PRESS? Type in "1" for the model fitted in (d) and type in "2" for the model fitted in (e) (without quotation marks). **1 Punkt**

1 Punkt

model with lowest value of PRESS

(d) Fit the penalized regression model with Lasso with tuning parameter λ_{1-SE} from (c). What is the minimum value of the resulting estimated coefficients ? 1 Punkt

1 Punkt

minimum value of resulting estimated coefficients **(requested precision: 4 digits)**

-0,5996

(e) Fit the penalized regression model with Lasso with tuning parameter λ_{min} from (c). What is the minimum value of the resulting estimated coefficients ? 1 Punkt

1 Punkt

minimum value of resulting estimated coefficients **(requested precision: 4 digits)**

-10,6274

(f) Test the models fitted in (d) and (e) on the test dataset and compute the values for predicted residual sum of squares (PRESS). Which model has the lowest value of PRESS? Type in "1" for the model fitted in (d) and type in "2" for the model fitted in (e) (without quotation marks). 1 Punkt

1 Punkt

model with lowest value of PRESS

2

Vorherige Aufgabe
Nächste Aufgabe

Alle Antworten wurden gespeichert!

Vorherige
Aufgabe

Aufgabe 5

7 Punkte

Please provide numbers in the requested precision within each question. The use of different precision is evaluated as wrong!

Load the dataset `UCBAdmissions` from the package `datasets` into your R workspace.

(a) Use `?UCBAdmissions` to get information about the dataset. Create a flat contingency table and add the marginal sums. What are the mean values for the row and the column marginals? **2 Punkte**

mean of row marginals **(requested precision: 1 digit)** — 1 Punkt

1131,5

mean of column marginals **(requested precision: 1 digit)** — 1 Punkt

754,3

(b) Transform the flat contingency table from (a) into a data frame. Fit the saturated model that predicts the count variable `Freq`. What is the value for AIC of the saturated model? **1 Punkt**

AIC **(requested precision: 1 digit)** — 1 Punkt

207,1

(b) Transform the flat contingency table from (a) into a data frame. Fit the saturated model that predicts the count variable **`Freq`**. What is the value for AIC of the saturated model? 1 Punkt

1 Punkt

AIC **(requested precision: 1 digit)**

207,1

(c) Starting with the saturated model of (b), use a backward selection algorithm to select the best nested model based on AIC. What is the value of the null deviance for this selected model? 1 Punkt

1 Punkt

null deviance **(requested precision: whole numbers)**

2650

(d) Fit the model of no three way interactions that predicts the count variable **`Freq`**. What is the value for λ_i^X for X given by Gender and i chosen to be the category "Female"? 1 Punkt

1 Punkt

λ_i^X **(requested precision: 4 digits)**

0,6819

(e) What is the mean value for the pearsonian residuals for the model of no three way interactions of (d)? 1 Punkt

1 Punkt

mean value **(requested precision: 3 digits)**

-0,035

(f) What is the mean value for the deviance residuals for the model of no three way interactions of (d)? 1 Punkt

(d) Fit the model of no three way interactions that predicts the count variable **Freq**. What is the value for λ_i^X for X given by Gender and i chosen to be the category "Female"? 1 Punkt

1 Punkt

λ_i^X **(requested precision: 4 digits)**

0,6819

(e) What is the mean value for the pearsonian residuals for the model of no three way interactions of (d)? 1 Punkt

1 Punkt

mean value **(requested precision: 3 digits)**

-0,035

(f) What is the mean value for the deviance residuals for the model of no three way interactions of (d)? 1 Punkt

1 Punkt

mean value **(requested precision: 3 digits)**

-1,205

Vorherige Aufgabe

Alle Antworten wurden gespeichert!