

'voyeuristic'의 검색결과 - 네이버 사진 | A Data Science Career Survival Guide | by Chris Kenwri... | Comma before "where": Rules and Examples | Kurs: Business Process Intelligence (VO) [21ss-12.251... | Learner | Orbit | Dynexite

E-Test 3 | DYNEXITE | 414760 | 16:39:54

Let a family of distributions be given by their pdfs (probability density functions) defined for $\alpha > 0, \gamma > 0$ as

$$f(x; \alpha, \gamma) = \sqrt{\frac{\gamma}{2\pi x^2}} \exp\left(-\frac{\gamma(x-\alpha)^2}{2\alpha^2 x}\right), \quad x > 0. \quad (1)$$

For fixed (known) $\gamma > 0$, $f_\gamma(x; \alpha) = f(x; \alpha, \gamma)$ defines a subfamily of the exponential dispersion family (EDF) of distributions with

$$c(x, \phi) = \frac{1}{2} \left(\ln(\gamma) - \ln(2\pi x^2) - \frac{\gamma}{x} \right).$$

Find the missing numerical values. **For all numerical results the exact values have to be provided without any rounding.** 3 von 6 Punkten

(a) Let $\gamma = 2$ and $X \sim f_2(\cdot; \alpha)$. 1 von 1 Punkt
 (i) Determine the values of the natural parameter θ and the dispersion $a(\phi)$, when $\alpha = 1$.
 $\theta =$ -0.5 ✓

(ii) Calculate the expectation $E(X)$, when $\alpha = 1$. 1 Punkt
 $E(X) =$ -1 ✗ 1 ↗

(iii) Calculate the variance $\text{Var}(X)$, when $\alpha = 1$. 1 von 1 Punkt

1 2 3 4 5 6

ÜBERSICHT | EINSICHT VERLASSEN

'voyeuristic'의 검색결과 - 네이버 사진 | A Data Science Career Survival Guide | by Chris Kenwri... | Comma before "where": Rules and Examples | Kurs: Business Process Intelligence (VO) [21ss-12.251... | Learner | Orbit | Dynexite

E-Test 3 | DYNEXITE | 414760 | 16:39:41

$f(x; \alpha, \gamma) = \sqrt{\frac{\gamma}{2\pi x^2}} \exp\left(-\frac{\gamma(x - \alpha)^2}{2\alpha^2 x}\right), \quad x > 0. \quad (1)$

For fixed (known) $\gamma > 0$, $f_\gamma(x; \alpha) = f(x; \alpha, \gamma)$ defines a subfamily of the exponential dispersion family (EDF) of distributions with

$$c(x, \phi) = \frac{1}{2} \left(\ln(\gamma) - \ln(2\pi x^2) - \frac{\gamma}{x} \right).$$

Find the missing numerical values. For all numerical results the exact values have to be provided without any rounding. 3 von 6 Punkten

(a) Let $\gamma = 2$ and $X \sim f_2(\cdot; \alpha)$. 1 von 1 Punkt

(i) Determine the values of the natural parameter θ and the dispersion $a(\phi)$, when $\alpha = 1$.

$\theta =$ -0.5 ✓

$a(\phi) =$ 0.5 ✓

(ii) Calculate the expectation $E(X)$, when $\alpha = 1$. 1 Punkt

$E(X) =$ -1 ✗ 1 ↗

(iii) Calculate the variance $\text{Var}(X)$, when $\alpha = 1$. 1 von 1 Punkt

$\text{Var}(X) =$ 0.5 ✓

ÜBERSICHT | 1 | 2 | 3 | 4 | 5 | 6 | EINSICHT VERLASSEN

E-Test 3

DYNEXITE

1 von 1 Punkt

(iii) Calculate the variance $\text{Var}(X)$, when $\alpha = 1$.

$\text{Var}(X) =$

0.5 ✓

1 Punkt

(b) Further assume that Y is a binary response variable and let

$\pi(x) = P(Y = 1 | X = x).$

Suppose that

$(X|Y = j) \sim f(\cdot | \alpha_j, \gamma_j),$

that is, conditionally on $Y = j$ the explanatory variable X has pdf (1) with parameters $\alpha_j, \gamma_j > 0, j \in \{0, 1\}$, and consider the model

$\text{logit}(\pi(x)) = \log\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x^{-1} + \beta_2 x.$

Assume that $\alpha_0 = \gamma_0 = 2$ and $\alpha_1 = \gamma_1 = 1$. Calculate the values of β_1 and β_2 .

Hint: Use Bayes' Theorem applied to probability distributions: For random variables X_1, X_2 with pdfs or pmfs f^{X_1}, f^{X_2} , respectively, it holds

$$f^{X_1|X_2=x_2}(x_1) = \frac{f^{X_2|X_1=x_1}(x_2)f^{X_1}(x_1)}{f^{X_2}(x_2)} I_{\text{supp}(X_2)}(x_2), \quad x_1 \in \mathbb{R},$$

where $f^{X_i|X_j}, i \neq j, i, j \in \{1, 2\}$ is the conditional probability density or mass function of X_i given X_j and I_A is the indicator function on a set A .

$\beta_1 =$

1 ✕ -0.5 ↗

1 Punkt

$\beta_2 =$

0 ✕ 0.25 ↗

1 2 3 4 5 6

EINSICHT VERLASSEN

dynexite.rwth-aachen.de

A Data Science Career Survival Guide | by Chris Kenwri... Comma before "where": Rules and Examples Kurs: Business Process Intelligence (VO) [21ss-12.251... Learner | Orbit | Dynexite

E-Test 3 414760 16:39:28

Consider a GLM $E(\mathbf{Y}) = g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$, where g is the canonical link, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ and

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Further assume that the responses Y_1, Y_2, Y_3 are jointly independent with $Y_i \sim \mathcal{P}(\mu_i)$, where $\mathcal{P}(\mu_i)$ denotes the Poisson distribution with parameters $\mu_i = 2i^2, i \in \{1, 2, 3\}$.

Determine the expected Fisher information matrix

$$\mathcal{I}_F = \begin{pmatrix} i_1 & i \\ i & i_2 \end{pmatrix}$$

with respect to the parameter vector $\boldsymbol{\beta}$.

Calculate the missing numerical values. **For all numerical results the exact values have to be provided without any rounding.** 3 von 3 Punkten

$i_1 =$ 1 von 1 Punkt 20 ✓

$i_2 =$ 1 von 1 Punkt 26 ✓

$i =$ 1 von 1 Punkt -18 ✓

ÜBERSICHT EINSICHT VERLASSEN

Let $(X_i)_i \stackrel{iid}{\sim} \mathcal{P}(\mu_1)$, $(Y_i)_i \stackrel{iid}{\sim} \mathcal{P}(\mu_2)$ be two jointly independent sequences of Poisson random counts with parameter μ_i , $i \in \{1, 2\}$, respectively. Further, define

$$\hat{\mu}_{1(n)} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \hat{\mu}_{2(n)} = \frac{1}{n} \sum_{i=1}^n Y_i, \quad n \in \mathbb{N}.$$

Derive the asymptotic variance $\sigma^2 > 0$, say, of

$$Z_n = \sqrt{n}(\hat{\mu}_{1(n)} + \hat{\mu}_{2(n)}^2)$$

as $n \rightarrow \infty$ for $\mu_1 = 1$ and $\mu_2 = 2$.

Calculate the missing numerical value. The numerical result has to be given as an exact values without any rounding.

0 von 4 Punkten

$\sigma^2 =$

4 Punkte
3 ✘ 33 🎁

Vorherige
← Aufgabe

Nächste
→ Aufgabe

Nachkorrekturantrag anlegen?

Bitte beachte, dass dieses Dokument zu einem Teil deiner Prüfungsakte wird!

- ⚠ Sei höflich und freundlich.
- Beschreibe deine Begründung so präzise wie möglich.

A Data Science Career Survival Guide | by Chris Kenwri... Comma before "where": Rules and Examples Kurs: Business Process Intelligence (VO) [21ss-12.251... Learner | Orbit | Dynexite

E-Test 3 414760 16:39:15

Task 24 2 von 3 Punkten

0.5 von 0.5 Punkten
Which of the three models of task 24(b) and 24(c) has the best fit in terms of AIC? Type in "1" for the logistic regression model, "2" for the probit and "3" for the cloglog regression model (without quotation marks). Provide the corresponding AIC value for this model (**requested precision: 1 digit**)
best model
 ✓

0.5 von 0.5 Punkten
AIC value
 ✓

1 von 1 Punkt
Transform the variable **sex** into a factor variable. Fit a cloglog regression model that predicts **transport** using **time**, **sex** and their interaction. What is the percentage of correct classified observations for this model? (**requested precision: 4 digits**)
 ✓

0.5 Punkte
What are the values for AIC and BIC of the model above? (**requested precision: 1 digit**)
Value AIC
 ✗ ↗

0.5 Punkte
Value BIC
 ✗ ↗

ÜBERSICHT 1 2 3 4 5 6 EINSICHT VERLassen

A screenshot of a web-based assignment interface for Dynexite. The top navigation bar shows tabs for various courses and a user profile. The main content area is titled "Task 25".

Task 25

At the beginning of your code, set the seed to 2021. Load the dataset *credits.csv* from RWTHmoodle into your workspace. **Make sure that your working directory contains an unmodified version of *credits.csv*. In particular, we recommend downloading a fresh version of *credits.csv* from the RWTHmoodle space.**

3.5 von 5.5 Punkten

1 von 1 Punkt

In the *credits* dataset, the variable **guaran** stands for whether the possible credits recipient is "No further guarantor of debtor" (**guaran** = 1), "co applicant" (**guaran** = 2) or "guarantor" (**guaran** = 3). Introduce a new variable **guaran2** with two categories where **guaran2** = 1 for **guaran** = 1 and **guaran** = 2 and **guaran2** = 2 otherwise. For how many cases in the data set *credits* it holds that **guaran2**=1? (requested precision: whole number)

948 ✓

0.5 von 0.5 Punkten

Split randomly the data in training (65 % of the data) and testing data (the remaining 35 % of the data). Transform **account**, **behavior**, **rate**, **finance**, **furthered**, **home**, **job**, **pers** into factor variables. Fit the following logistic regression model on the training data:

```
repayment ~ time+age+account+behavior+savings+rate+guaran2+finance+furthcred+home+job+pers
```

In the following, we will refer to this model as **model b**.

What are the AIC and BIC values for this model? (requested precision: 1 digit)

AIC value

663.1 ✓

BIC value

0.5 Punkte

779.5 ✗ 774.4 ↗

1 von 1 Punkt

What is the AUC (area under the curve) of **model b**? (requested precision: 3 digits)

0.814 ✓

Navigation: ÜBERSICHT (Overview) | EIN SICHT VERLASSEN (Leave view)

Page navigation: 1 2 3 4 5 6

A screenshot of a web-based test interface from Dynexite. The top navigation bar shows the URL dynexite.rwth-aachen.de. The page title is "E-Test 3". On the right, there are user statistics: 414760 points and the time 16:39:03.

The main content area displays a series of six questions:

- What is the AUC (area under the curve) of *model b*? (requested precision: 3 digits)**
1 von 1 Punkt
0.814 ✓
- Fit the null model corresponding to *model b*. What is the null deviance for this model? (requested precision: 4 digits)**
1 Punkt
797.5 ✗ 797.4835 ↕
- Select the best model nested in *model b* in terms of AIC using a backward stepwise selection algorithm. In the following, we will refer to this selected model as *model e*. What is the median of the estimated probabilities for on time credit repayment for the training sample, based on this model? (requested precision: 4 digits)**
0.5 Punkte
0.776 ✗ 0.7573 ↕
- Select the best model nested in *model b* in terms of BIC using a backward stepwise selection algorithm. In the following, we will refer to this selected model als *model f*. Test *model f* and *model e* on the test data. What are the values for predicted residual sum of squares (PRESS) for both models ? (requested precision: 4 digits)**
0.5 von 0.5 Punkten
PRESS for *model e*
62.6048 ✓
- PRESS for *model f***
0.5 von 0.5 Punkten
61.4974 ✓

At the bottom, there are navigation buttons: "Vorherige Aufgabe" (Previous Task) and "Nächste Aufgabe" (Next Task). Below the tasks, a horizontal navigation bar shows the current task number 5 highlighted in blue, with other numbers 1, 2, 3, 4, 6, and a "ÜBERSICHT" (Overview) button.

A screenshot of a web-based E-Test interface on the Dynexite platform. The page title is "Task 26" with a total of 2 von 3.5 Punkten available. The task consists of five questions, each worth 0.5 Punkte.

Question 1: Fit a probit regression model that predicts **Good**, using **Dist** as explanatory variable. In the following, we will refer to this model as **model 1**. What is the deviance of this model? (**requested precision: 4 digits**)

Inputs: -795 (incorrect), 794.9805 (correct)

Question 2: What is the value of the predicted probability for a good kick when the distance is 47 yards for **model 1**? (**requested precision: 4 digits**)

Inputs: 0.6243 (incorrect), 0.6942 (correct)

Question 3: What is the 90 % profile confidence interval for the intercept of **model 1**? (**requested precision: 4 digits**)

Inputs: lower bound of CI: 0.3634 (incorrect), 3.3634 (correct)

Inputs: upper bound of CI: 4.2432 (correct)

Question 4: Fit a probit model that predicts **Good**, using the attributes **Dist**, **Blk.**, **Pressure**, **Roof.type**, **Altitude** and **Field**. Select the model with smallest AIC using a backward stepwise selection algorithm. In the following, we will refer to this model as **model 2**. What is the number of parameters included in the selected model beyond the intercept? (**requested precision: whole number**)

Inputs: 2 (correct)

Question 5: What are the percentages of correct classified observations for **model 1** and **model 2** when the threshold probability is 0.5? (**requested precision: 4 digits**)

The navigation bar at the bottom shows tabs for "ÜBERSICHT" (Overview) and "EINSICHT VERLASSEN" (Leave Insights). The page header includes links to various resources like "A Data Science Career Survival Guide" and "Kurs: Business Process Intelligence (VO) [21ss-12.251...]."

A Data Science Career Survival Guide | by Chris Kenwri... Comma before "where": Rules and Examples Kurs: Business Process Intelligence (VO) [21ss-12.251... Learner | Orbit | Dynexite

E-Test 3 414760 16:38:52

DYNEXITE

What is the value of the predicted probability for a good kick when the distance is 47 yards for *model 1*? (requested precision: 4 digits) 0.5 Punkte
0.6243 ✗ 0.6942 ↗

What is the 90 % profile confidence interval for the intercept of *model 1*? (requested precision: 4 digits) 0.5 Punkte
lower bound of CI 0.3634 ✗ 3.3634 ↗

upper bound of CI 0.5 von 0.5 Punkten
4.2432 ✓

Fit a probit model that predicts *Good*. using the attributes *Dist*, *Blk.*, *Pressure*, *Roof.type*, *Altitude* and *Field*. Select the model with smallest AIC using a backward stepwise selection algorithm. In the following, we will refer to this model as *model 2*. What is the number of parameters included in the selected model beyond the intercept? (requested precision: whole number) 0.5 von 0.5 Punkten
2 ✓

What are the percentages of correct classified observations for *model 1* and *model 2* when the threshold probability is 0.5? (requested precision: 4 digits) 0.5 von 0.5 Punkten
percentage correct classified observations *model 1* 0.8232 ✓

percentage correct classified observations *model 2* 0.5 von 0.5 Punkten
0.8409 ✓

ÜBERSICHT 1 2 3 4 5 6 EINSICHT VERLASSEN

A screenshot of a web-based test interface for Dynexite. The top navigation bar shows the URL dynexite.rwth-aachen.de. The page title is "E-Test 3". On the right, there are user statistics: 414760 points and the time 16:38:46.

The main content area contains several questions:

- What is the 90 % profile confidence interval for the intercept of *model 1?* (requested precision: 4 digits)**
lower bound of CI
0.3634 X 3.3634 P
- upper bound of CI**
4.2432 V
- Fit a probit model that predicts Good. using the attributes Dist, Blk., Pressure, Roof.type, Altitude and Field. Select the model with smallest AIC using a backward stepwise selection algorithm. In the following, we will refer to this model as *model 2*. What is the number of parameters included in the selected model beyond the intercept? (requested precision: whole number)**
2 V
- What are the percentages of correct classified observations for *model 1* and *model 2* when the threshold probability is 0.5? (requested precision: 4 digits)**
percentage correct classified observations *model 1*
0.8232 V
- percentage correct classified observations *model 2***
0.8409 V

At the bottom left, there is a "Vorherige Aufgabe" button with a left arrow icon. At the bottom center, there is a navigation bar with numbered buttons 1 through 6, where 6 is highlighted in blue. At the bottom right, there is a "EINSICHT VERLASSEN" button with a shield icon.