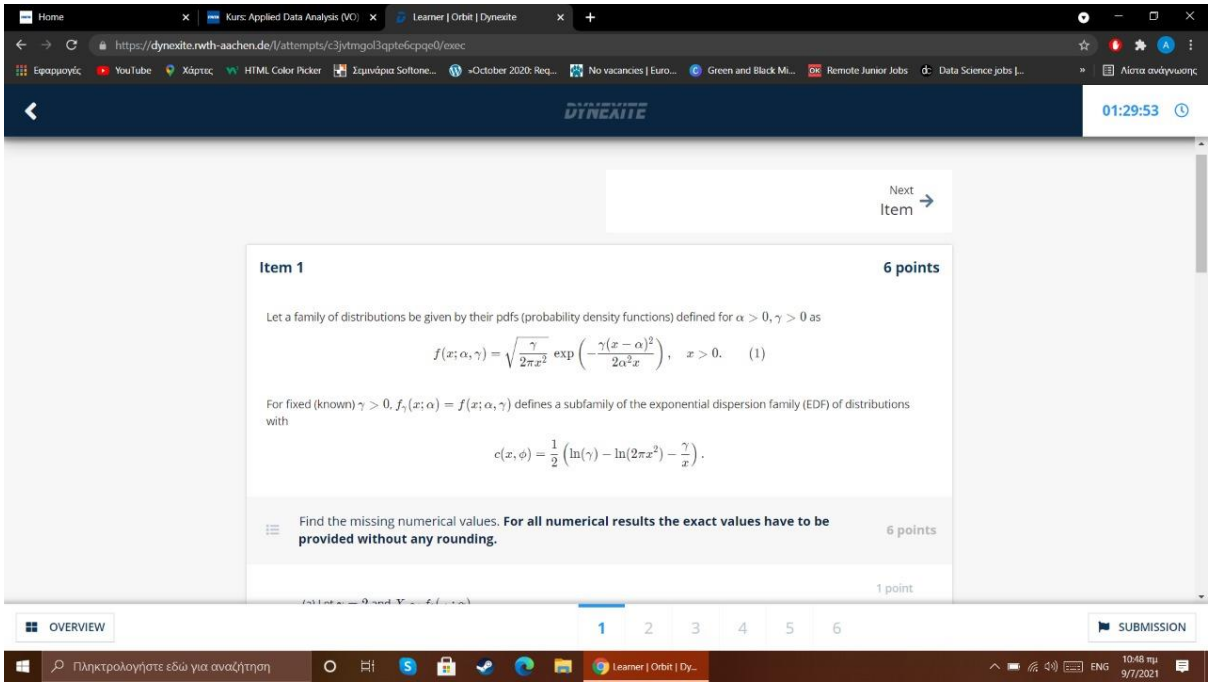
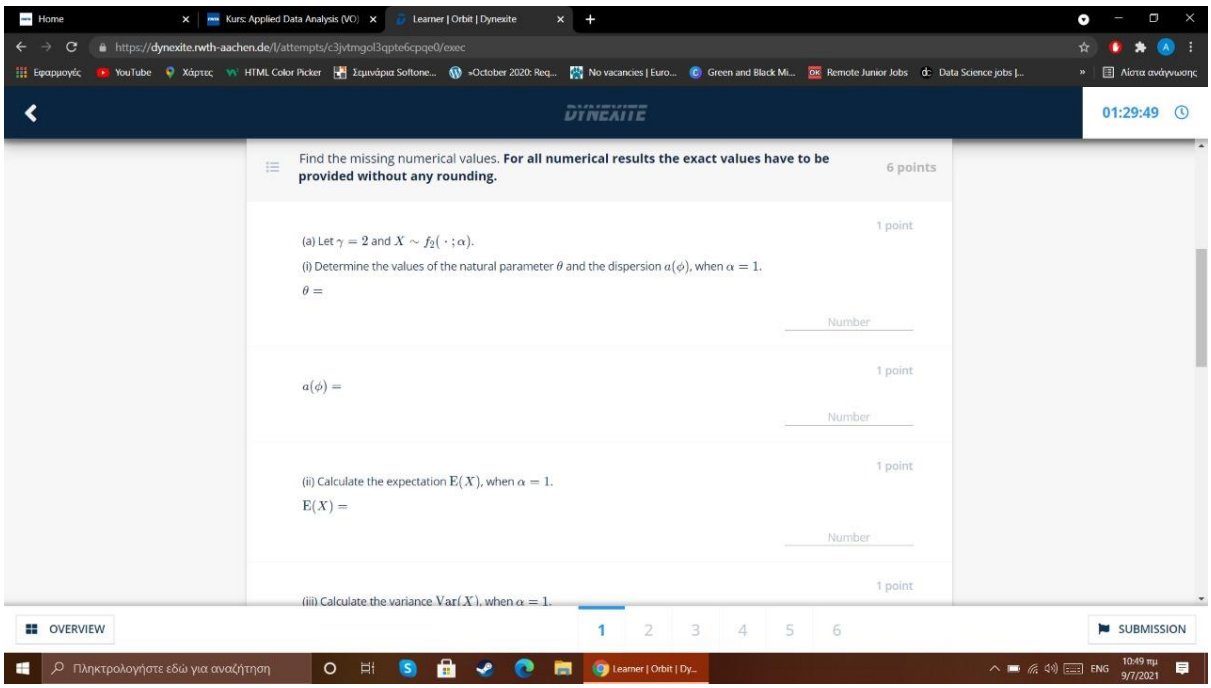
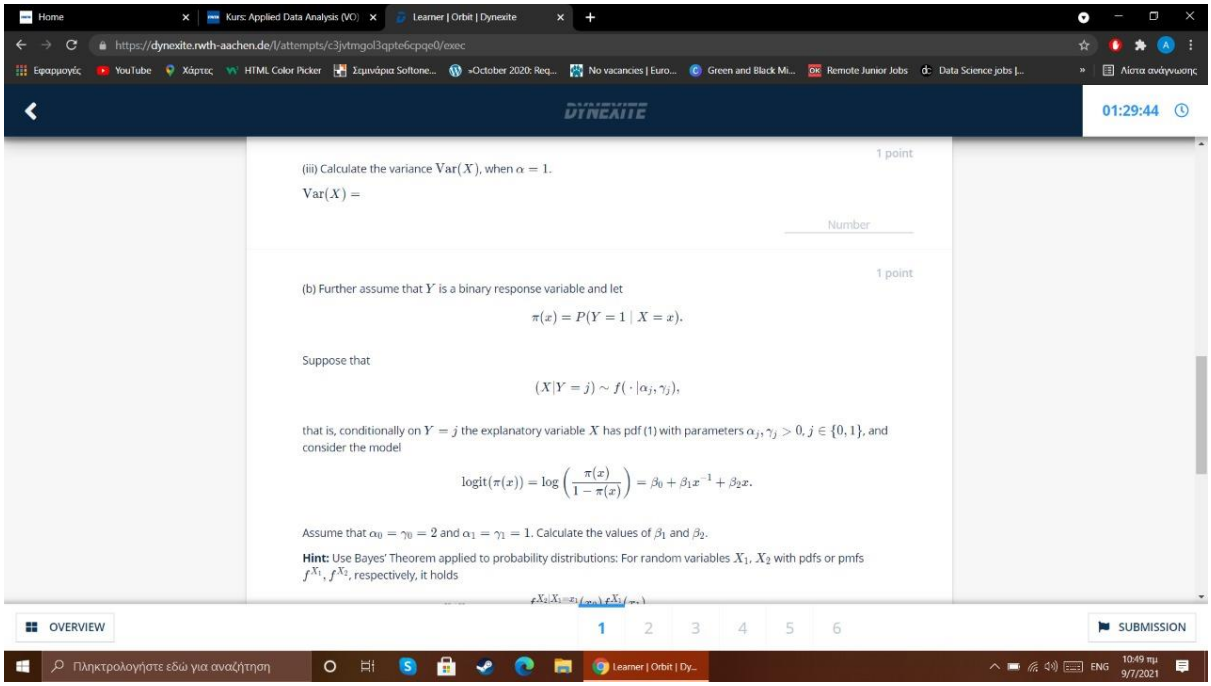
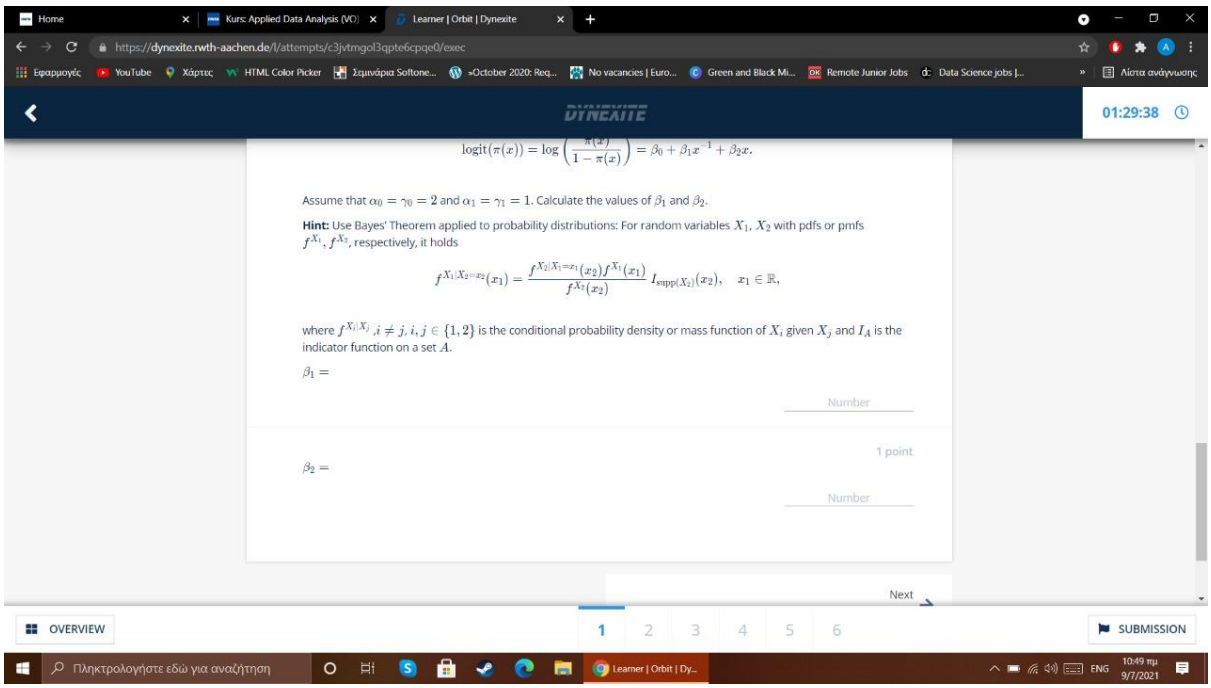
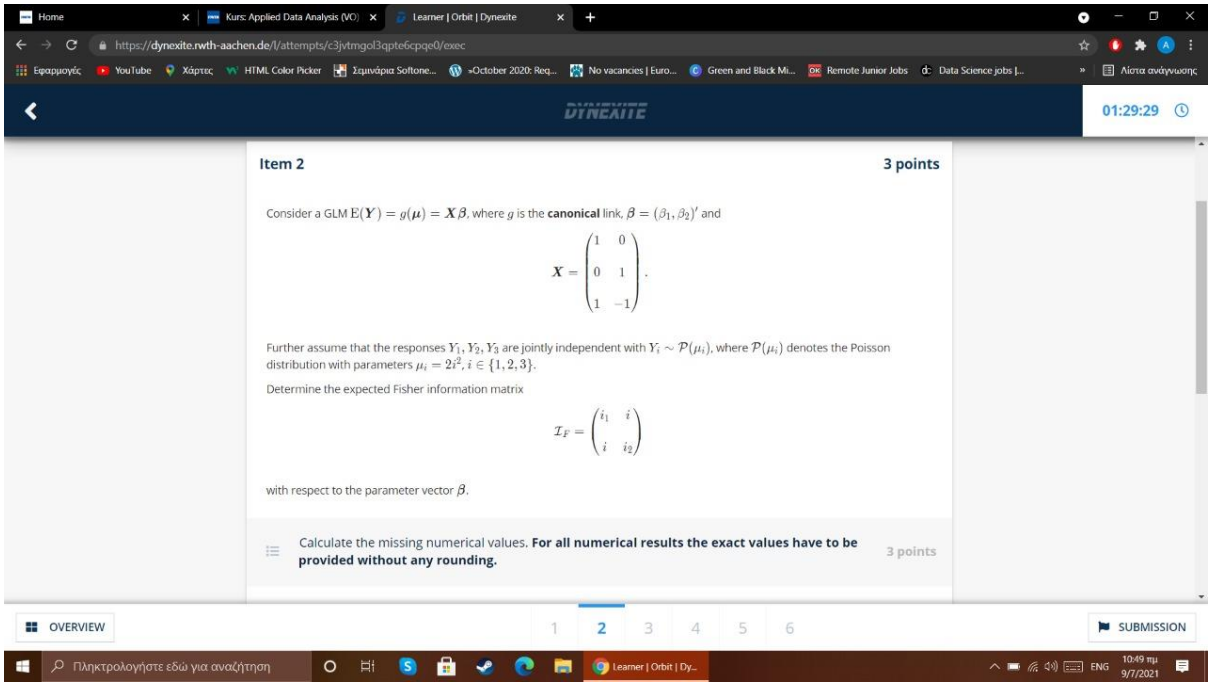
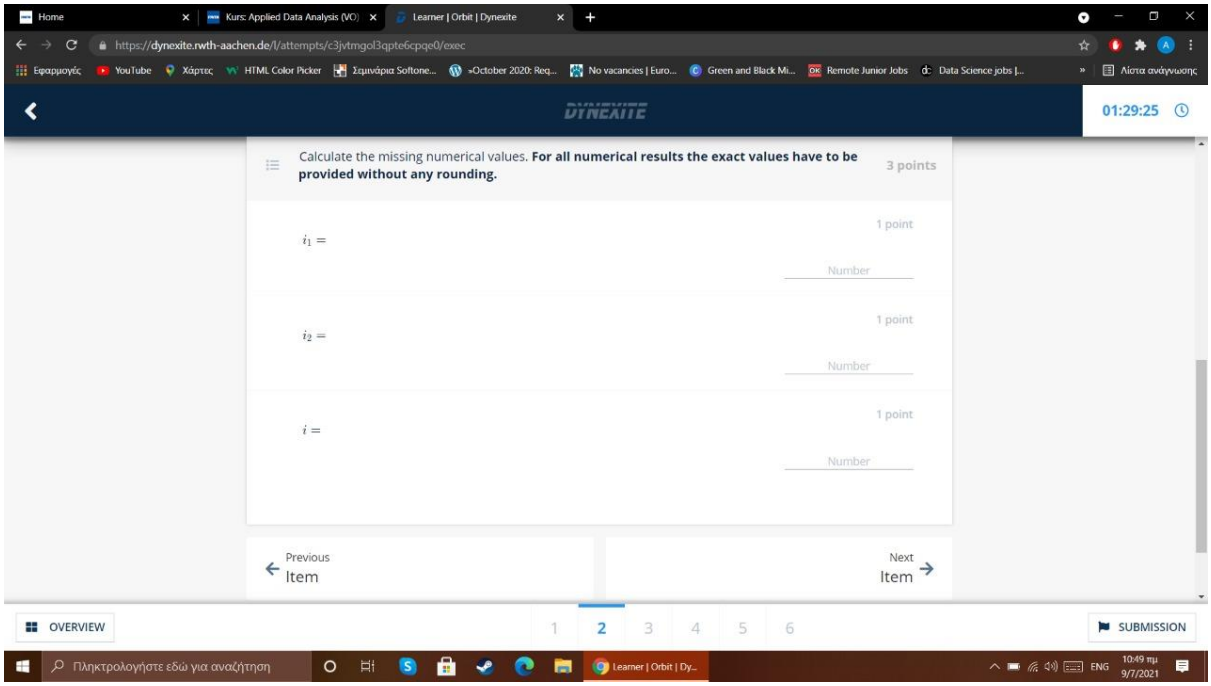

Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
DYNEXITE
01:29:53
Next Item
Item 1
6 points
Let a family of distributions be given by their pdfs (probability density functions) defined for $\alpha > 0, \gamma > 0$ as
$$f(x;\alpha,\gamma) = \sqrt{\frac{\gamma}{2\pi x^2}}\,\exp\left(-\frac{\gamma(x-\alpha)^2}{2\alpha^2 x}\right), \quad x > 0. \qquad (1)$$
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.**
6 points
1 point
OVERVIEW
1 2 3 4 5 6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:48 πμ
9/7/2021


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
DYNEXITE
01:29:49
Find the missing numerical values. **For all numerical results the exact values have to be provided without any rounding.**
6 points
1 point
(a) Let $\gamma = 2$ and $X \sim f_2(\,\cdot\,;\alpha)$.
(i) Determine the values of the natural parameter θ and the dispersion $a(\phi)$, when $\alpha = 1$.
$\theta =$
Number
1 point
$a(\phi) =$
Number
1 point
(ii) Calculate the expectation $\mathrm{E}(X)$, when $\alpha = 1$.
$\mathrm{E}(X) =$
Number
1 point
(iii) Calculate the variance $\mathrm{Var}(X)$, when $\alpha = 1$.
OVERVIEW
1 2 3 4 5 6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:49 πμ
9/7/2021


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
DYNEXITE
01:29:44
1 point
(iii) Calculate the variance $\mathrm{Var}(X)$, when $\alpha = 1$.
$\mathrm{Var}(X) =$
Number
1 point
(b) Further assume that Y is a binary response variable and let
$$\pi(x) = P(Y = 1 \mid X = x).$$
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
$$\mathrm{logit}(\pi(x)) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x^{-1} + \beta_2 x.$$
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
OVERVIEW
1 2 3 4 5 6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:49 πμ
9/7/2021


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
DYNEXITE
01:29:38
$$\mathrm{logit}(\pi(x)) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x^{-1} + \beta_2 x.$$
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
$$f^{X_1|X_2=x_2}(x_1) = \frac{f^{X_2|X_1=x_1}(x_2) f^{X_1}(x_1)}{f^{X_2}(x_2)} I_{\mathrm{supp}(X_2)}(x_2), \quad x_1 \in \mathbb{R},$$
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.
$\beta_1 =$
Number
1 point
$\beta_2 =$
Number
Next
OVERVIEW
1 2 3 4 5 6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:49 πμ
9/7/2021


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:29:29
Item 2
3 points
Consider a GLM $\mathrm{E}(\boldsymbol{Y}) = g(\boldsymbol{\mu}) = \boldsymbol{X}\boldsymbol{\beta}$, where g is the **canonical** link, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ and
$$\boldsymbol{X} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}.$$
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 points
OVERVIEW
1
2
3
4
5
6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:49 πμ
9/7/2021


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:29:25
Calculate the missing numerical values. **For all numerical results the exact values have to be provided without any rounding.**
3 points
$i_1 =$
1 point
Number
$i_2 =$
1 point
Number
$i =$
1 point
Number
Previous Item
Next Item
OVERVIEW
1
2
3
4
5
6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:49 πμ
9/7/2021

Item 3

4 points

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

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

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

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

as $n \to \infty$ 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.** 4 points

4 points

$\sigma^2 =$ Number

Home | Kurs: Applied Data Analysis (VO) | Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
Εφαρμογές YouTube Χάρτες HTML Color Picker Σεμινάρια Softone... »October 2020: Req... No vacancies | Euro... Green and Black Mi... Remote Junior Jobs Data Science jobs |... Λίστα ανάγνωσης

Item 4

3 points

The questions of the third E-Test are based on the tasks of R-Laboratory 8 and 9. **Please provide numbers in the requested precision within each question. The use of different precision is evaluated as wrong.**

Task 24 3 points

0.5 points

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

Number

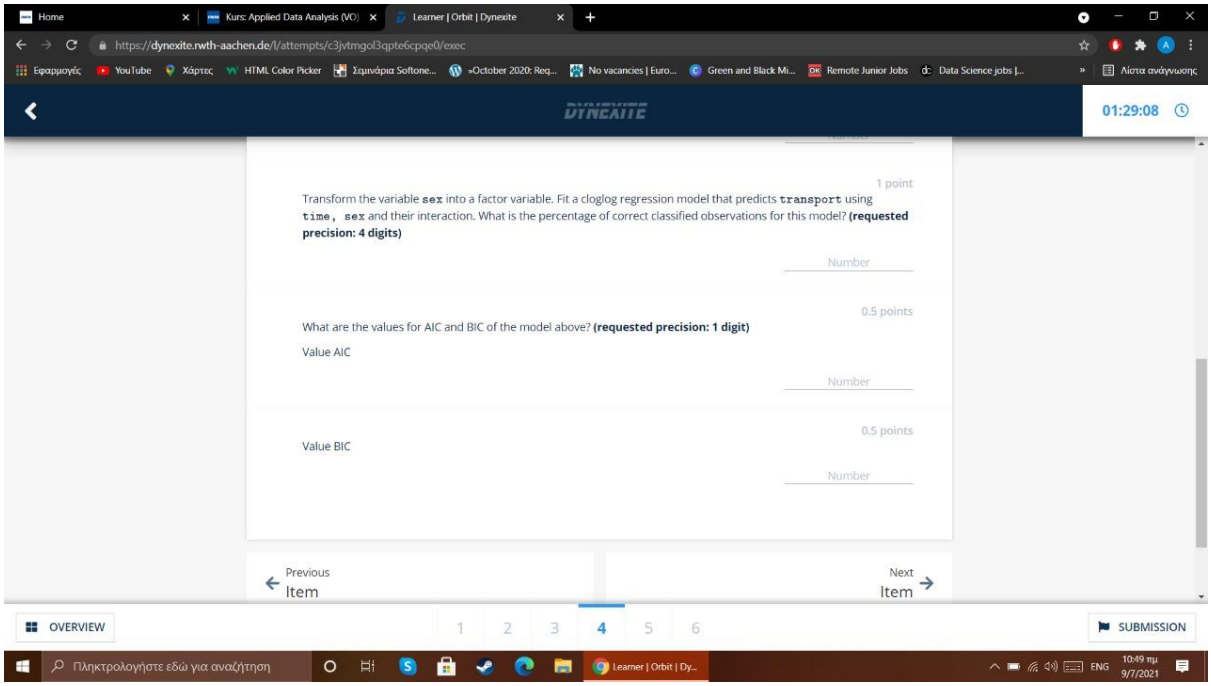
0.5 points

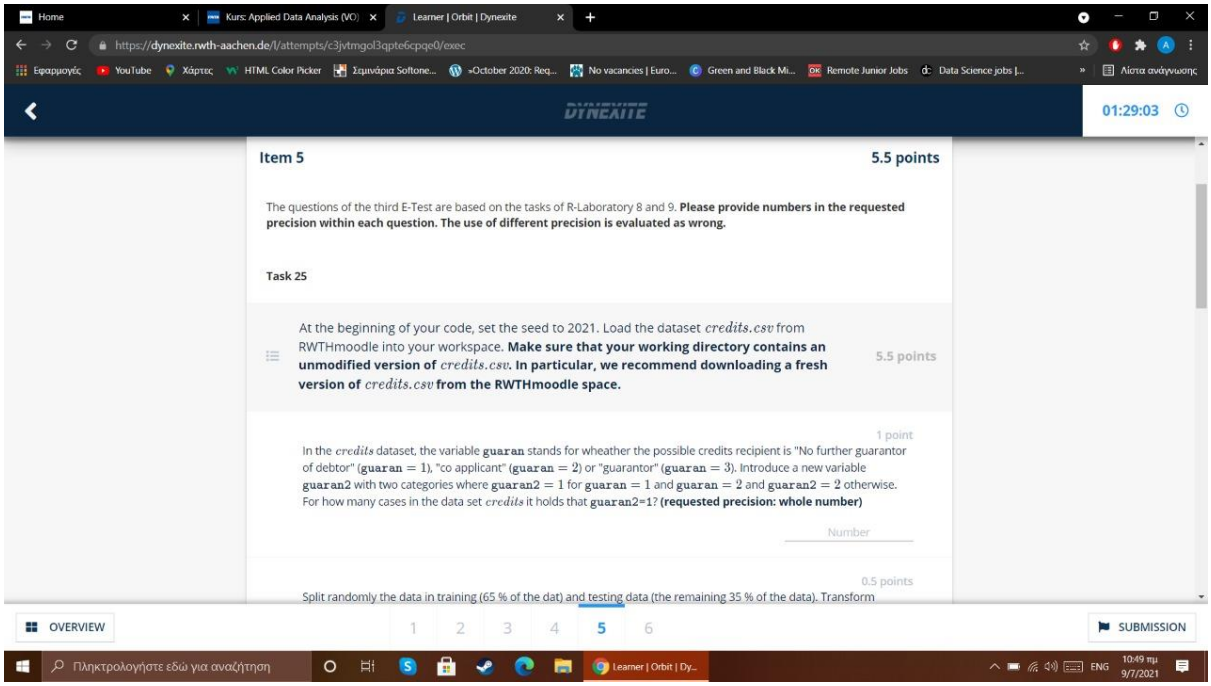
AIC value

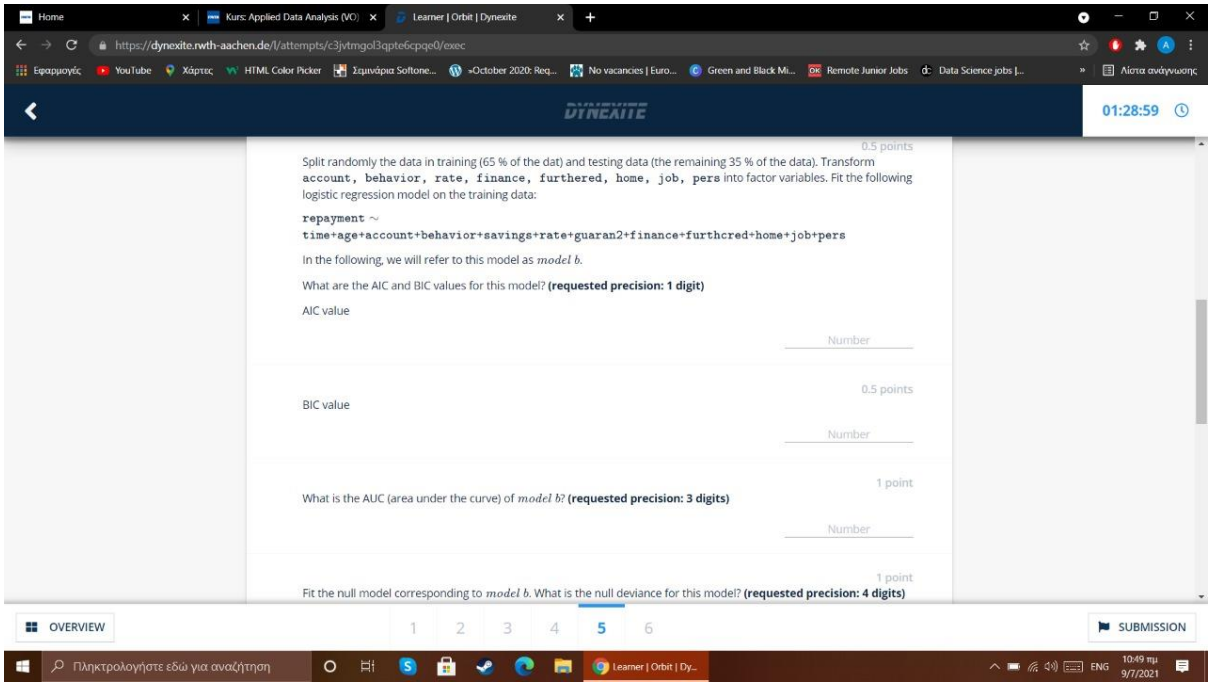
Number

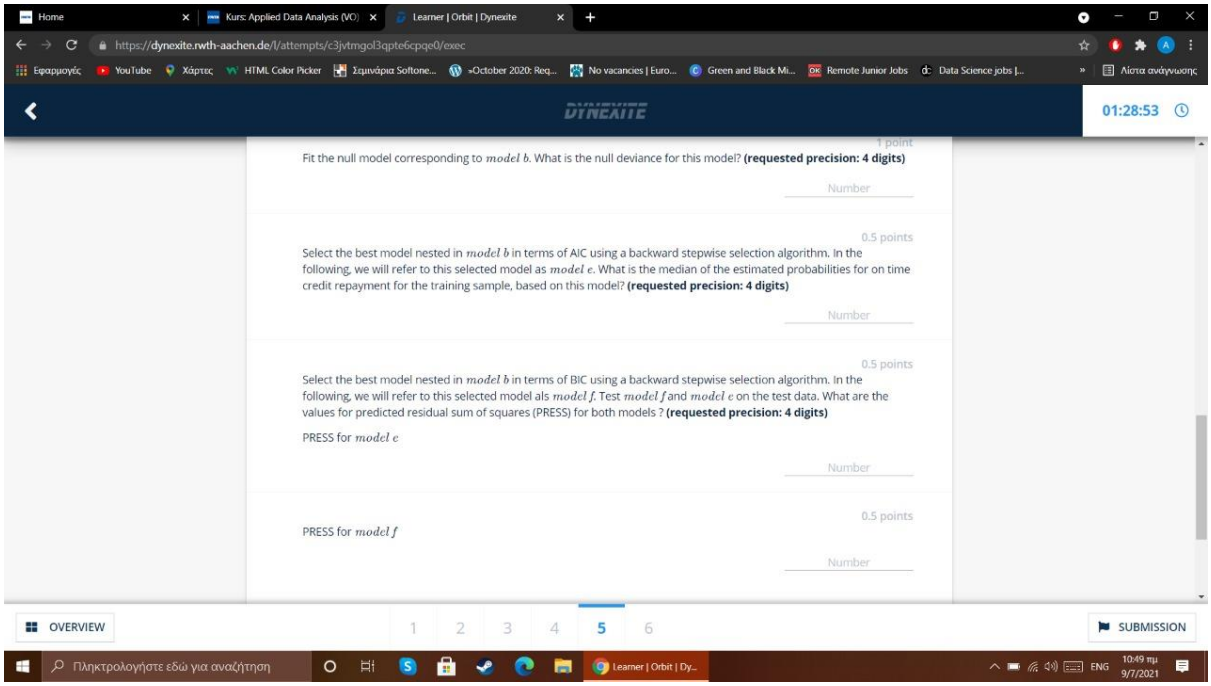
1 point

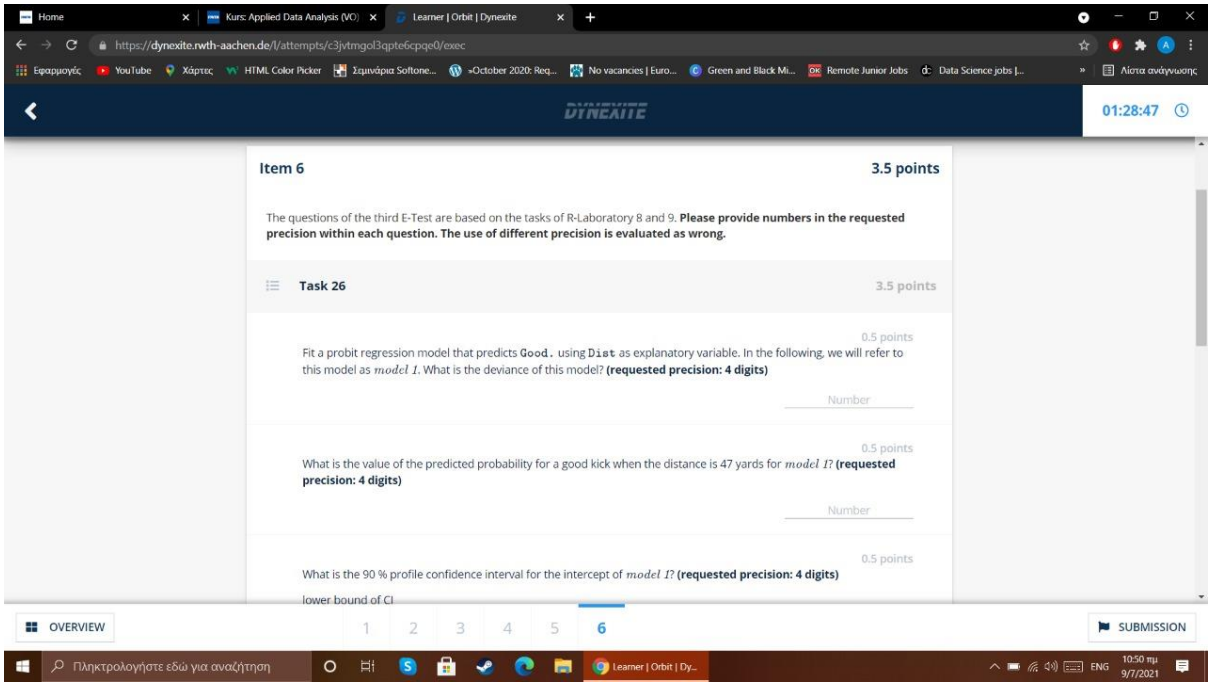
Transform the variable `sex` into a factor variable. Fit a cloglog regression model that predicts `transport` using `time, sex` and their interaction. What is the percentage of correct classified observations for this model? **(requested**

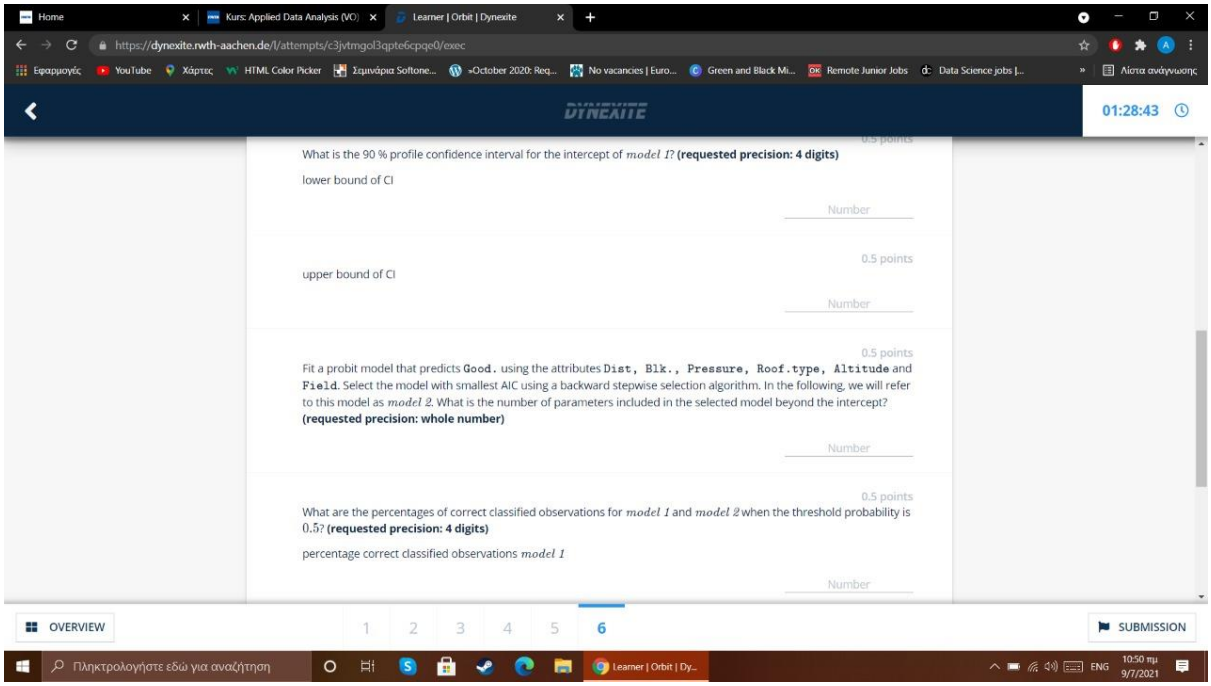

Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:29:08
1 point
Transform the variable sex into a factor variable. Fit a cloglog regression model that predicts transport using time, sex and their interaction. What is the percentage of correct classified observations for this model? (requested precision: 4 digits)
Number
0.5 points
What are the values for AIC and BIC of the model above? (requested precision: 1 digit)
Value AIC
Number
0.5 points
Value BIC
Number
Previous Item
Next Item
OVERVIEW
1 2 3 4 5 6
SUBMISSION

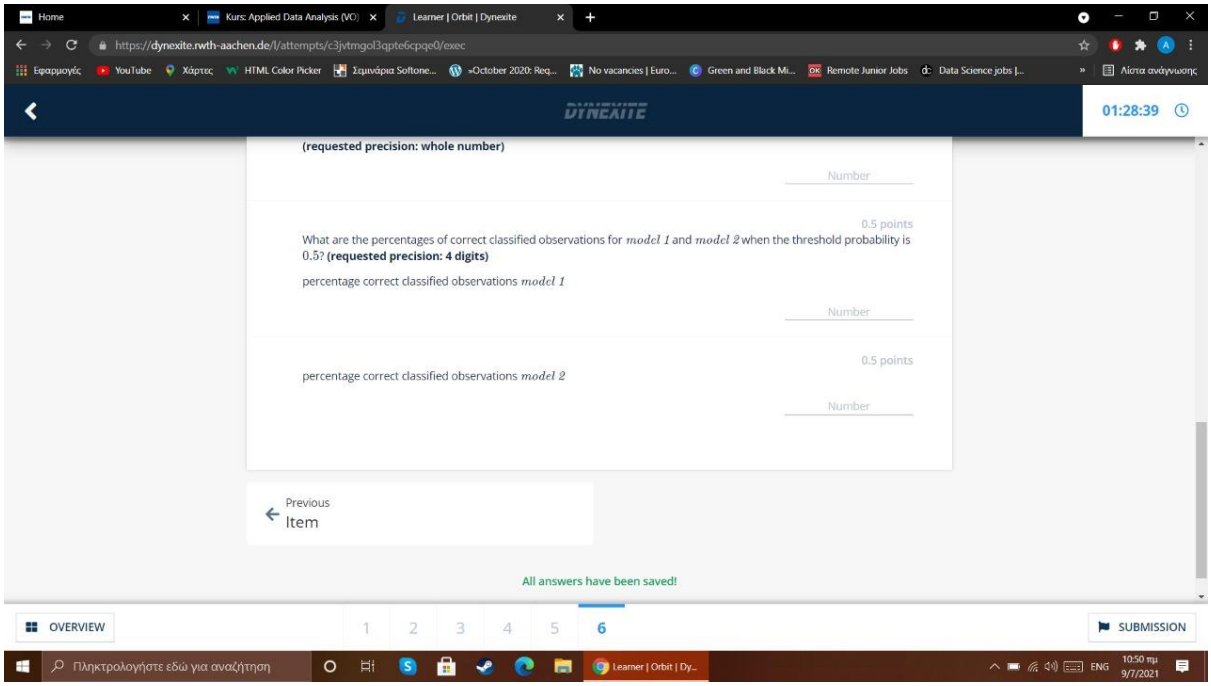

Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:29:03
Item 5
5.5 points
The questions of the third E-Test are based on the tasks of R-Laboratory 8 and 9. Please provide numbers in the requested precision within each question. The use of different precision is evaluated as wrong.
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.
5.5 points
1 point
In the credits dataset, the variable guaran stands for wheather 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)
Number
0.5 points
Split randomly the data in training (65 % of the dat) and testing data (the remaining 35 % of the data). Transform
OVERVIEW
1 2 3 4 5 6
SUBMISSION


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:28:59
0.5 points
Split randomly the data in training (65 % of the dat) and testing data (the remaining 35 % of the data). Transform account, behavior, rate, 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
Number
0.5 points
BIC value
Number
1 point
What is the AUC (area under the curve) of model b? (requested precision: 3 digits)
Number
1 point
Fit the null model corresponding to model b. What is the null deviance for this model? (requested precision: 4 digits)
OVERVIEW
1 2 3 4 5 6
SUBMISSION


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:28:53
1 point
Fit the null model corresponding to model b. What is the null deviance for this model? (requested precision: 4 digits)
Number
0.5 points
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)
Number
0.5 points
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)
PRESS for model e
Number
0.5 points
PRESS for model f
Number
OVERVIEW
1 2 3 4 5 6
SUBMISSION


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
DYNEXITE
01:28:47
Item 6
3.5 points
The questions of the third E-Test are based on the tasks of R-Laboratory 8 and 9. Please provide numbers in the requested precision within each question. The use of different precision is evaluated as wrong.
Task 26
3.5 points
0.5 points
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)
Number
0.5 points
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)
Number
0.5 points
What is the 90 % profile confidence interval for the intercept of model 1? (requested precision: 4 digits)
lower bound of CI
OVERVIEW
1 2 3 4 5 6
SUBMISSION


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
DYNEXITE
01:28:43
What is the 90 % profile confidence interval for the intercept of model 1? (requested precision: 4 digits)
lower bound of CI
Number
0.5 points
upper bound of CI
Number
0.5 points
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)
Number
0.5 points
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
Number
OVERVIEW
1 2 3 4 5 6
SUBMISSION


Home
Kurs: Applied Data Analysis (VO)
Learner | Orbit | Dynexite
https://dynexite.rwth-aachen.de/l/attempts/c3jvtmgol3qpte6cpqe0/exec
01:28:39
(requested precision: whole number)
Number
0.5 points
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
Number
0.5 points
percentage correct classified observations model 2
Number
Previous Item
All answers have been saved!
OVERVIEW
1
2
3
4
5
6
SUBMISSION
Πληκτρολογήστε εδώ για αναζήτηση
Learner | Orbit | Dy...
ENG
10:50 πμ
9/7/2021