Review Test Submission: Problem Set #2

User	Kohei Nishitani
Course	2024GSP_ANA_510_02 Statistical Modeling
Test	Problem Set #2
Started	4/5/24 12:58 PM
Submitted	4/6/24 1:05 PM
Due Date	4/7/24 11:59 PM
Status	Completed
Attempt Score	75 out of 100 points
Time Elapsed	24 hours, 6 minutes
Instructions	You should answer all questions first, either using paper and pencil or another computer program such as gretl. Then, enter your answers in the online assignment. I have setup this assignment so you will have three chances to take it, one to get the questions and one to really enter your answers. Be sure to keep track of your work and answers. If for any reason I have to reset your assignment it will wipe out all the work you did before! There are a variety of types of questions. You should select the best choice or choices. If you are entering a numeric value you should round your answer to two decimal places unless there are other specific instructions for a specific question. Not all questions are worth the same number of points, i.e. some questions are worth more points than others. If you have any questions – ask! There is additional information in the short description below.

Question 1 5 out of 5 points

How many observations are there in the contract/road construction dataset?

Question 2 0 out of 5 points

How many independent variables are in the contract data/road construction dataset?

Question 3 5 out of 5 points

Fill in the exact variable name of the dependent variable in the contract data/road construction dataset. Be careful, this field is case sensitive!

Question 4 5 out of 5 points

The possible values for the dependent variable, BIDStatusy are 0 or 1.

Question 5 5 out of 5 points

If BIDStatusy is coded with the number 1 that indicates the bid is a fixed bid. If BIDStatusy is coded with the number 0 that indicates the bid is a competitive bid.

Question 6 0 out of 5 points

The reported mean value of BIDStatusy, 0.38710, indicates that (select the best choice to complete this sentence).

Question 7 5 out of 5 points

The output of the logit model built for the contract data/road construction dataset indicates that the model is able to accurately predict _____ % of all cases (or observations). Note that the answer is in a percentage from the model output.

Question 8 5 out of 5 points

The logit model for the contract data/road construction dataset results in _____ Type II errors.

Question 9 5 out of 5 points

A Type Ii error for the contract data/road construction logit model means that (select the best choice below to complete this sentence).

Question 10 5 out of 5 points

For the contract data/road construction logit model as reported in the textbook Example 9.5, the (Hosmer and Lemeshow Goodness-of-Fit) P-value equal to 0.2324 which is greater than the 0.05 level of significance indicates that _____ (select the choice below that best completes this sentence). Note that I have redirected you to the textbook output for this question because we do not typically compute the Hosmer and Lemeshow Goodness-of-Fit test.

Question 11 5 out of 5 points

The number of observations in the cokeVsPepsi dataset is _____.

Question 12 0 out of 5 points

The number of variables in the cokeVsPepsi dataset is _____.

Question 13 5 out of 5 points

For the Coke-Pepsi dataset, the three models produce virtually identical results.

Question 14 5 out of 5 points

Although the probit and logit models built for the coke-pepsi dataset are different, both models result in the same error rates. (Hint, look for the confusion tables in the model output!)

Question 15 5 out of 5 points

The dependent variable, coke, in the cokeVsPepsi dataset means that _____ (select the choice below that best completes this sentence).

Question 16 0 out of 5 points

The number of correctly predicted "cases" by either the probit or logit models is _____. (Enter the number of correct cases this time, NOT the percent.)

Question 17 0 out of 5 points

The number of "cases" or observations that were predicted to have coke chosen when pepsi should have been chosen is _____. (Hint, look at the confusion tables in the model output!)

Question 18 5 out of 5 points

The expectation is that if a store has a coke display then coke is more likely chosen. If a store has a pepsi display then pepsi is more likely chosen. All the models; linear, probit, and logit, show that the coefficients for the variables disp_coke and disp_pepsi have opposite signs. This indicates that _____. (Select the choice below that best completes this sentence.)

Question 19 5 out of 5 points

Both probit and logit models can be used to predict "choice". This is due to their ability to predict "discrete" or binary outcomes. (Note, that this does NOT include multinomial models which involve more than 2 choices, e.g. either/or problems.)

Question 20 5 out of 5 points

To keep choice probability, "p," within the interval [0, 1], both the probit and logit models are based on a nonlinear "S-curve" or S-shaped relationship between "x" and "p". However, where the probit function is related to the standard normal probability distribution, the logit function is related to the sigmoid function.

Saturday, April 6, 2024 1:05:15 PM EDT

← OK