

Biostat 510
Homework 8
Due Thursday, April 9, 2009

State the alpha level that you are using for any statistical tests carried out for this homework. Don't forget to include the test statistic, df, and p-value for the results of any statistical tests that you report. Paste all commands for this homework in a syntax window and include this syntax as the first part of your write-up.

The questions in this homework use the Afifi data set (afifi.sav) which you created for homework 7, or you can download this SPSS data set from my web page (<http://www.umich.edu/~kwelch>).

1. Create new variables and value labels.
 - a) DIED = 1 if the person died, 0 if the person lived. This is the variable you will use for logistic regression models.
 - b) DIEDCAT = 1 if the person died, 2 if the person lived. This is the variable you will use for cross-tabs.
 - c) SHOCKDUM=1 if the person was in shock (shoktype=3,4,5,6,7), and SHOCKDUM=0 if the person was not in shock (shoktype=2). This variable already exists in the Afifi.sav dataset on the web, but you may need to create it if you're using your own version of the dataset.
 - d) SHOCKCAT = 1 if the person was in shock (shoktype=3,4,5,6,7), and SHOCKCAT= 2 if the person was not in shock (shoktype=2). This is the variable you will use for cross-tabs.
 - e) URINECAT1=1 if the person had any urinary output at time 1, and 0 if not.
 - f) URINECAT2=1 if the person had any urinary output at time 2, and 0 if not.
 - g) Create value labels for the variables DIEDCAT, SHOCKCAT, SEX, and SHOKTYPE.
 - h) Get descriptive statistics for all numeric variables in your dataset. Check the sample sizes for each variable to be sure that they match between the original variables and the new variables that you created.
 - i) Include the descriptive statistics table in your write-up.

2. Create a cross-tab between SHOCKCAT and DIEDCAT. Remember, SHOCKCAT is the risk factor, and DIEDCAT is the outcome, get the counts (they are included by default), the expected counts, and the row percentages, a Chi-square test and the relative risk and odds ratio.
 - a) What percentage of those patients who were in shock died? What percentage of those patients not in shock died?
 - b) Carry out a chi-square test of whether being in shock is independent of dying, and report the results of your test. (Use the Pearson chi-square test results).
 - c) What is the risk ratio of dying, for a patient who was in shock compared to a patient who was not in shock? Include the risk ratio and the 95% confidence interval calculated by SPSS.

- d) What is the Odds Ratio of dying, for a patient who was in shock compared to a patient who was not in shock? Please include the odds ratio and the 95% confidence interval calculated by SPSS.
 - e) Include the output from this cross-tab in your write-up.
3. Fit a logistic regression model, with DIED (the 0,1 variable) as the dependent variable, and SHOCKDUM (the 0,1 variable) as the predictor.
 - a) How many observations are included in the logistic regression model? What is the pseudo R-square for this model? (Report the Nagelkerke R-square.)
 - b) Report the odds ratio of dying for those in shock compared to those not in shock, along with the 95% CI from this logistic regression model. Do these values agree with the Odds Ratio and 95% CI from the cross-tab in Question 2 above?
 - c) Is there a significant relationship between being in shock and dying? Report the Wald test chi-square, df, and p-value for this test.
 - d) Include the following tables from the Logistic Regression output in your write-up: Case Processing Summary, Omnibus Test of Model Coefficients, Model Summary, and Variables in the Equation.
 4. Create a cross-tab between SHOKTYPE (the variable with 6 categories) and DIEDCAT, get a chi-square test of independence, row percents and expected values for this table.
 - a) Is there a significant relationship between SHOKTYPE and whether the patient died?
 - b) Include the output from this cross-tab in your write-up.
 5. Fit a logistic regression model with DIED as the dependent variable and SHOKTYPE as the Categorical predictor.
 - a) You will need to include SHOKTYPE as a categorical predictor. Use SHOKTYPE=2 (the first category for this variable) as the reference category.
 - b) Check the dummy variable coding reported by SPSS to be sure that the first category is indeed the reference category
 - c) Which categories of SHOKTYPE have significantly different probabilities of dying than the non-shock group (SHOKTYPE=2)?
 - d) What is the Nagelkerke R-square for this model?
 - e) Is there a significant overall relationship between SHOKTYPE and DIED? Report the overall test for SHOKTYPE from the table of Results for variables in the model.
 - f) Include the output from this logistic regression in your write-up (the same output tables as for question 3, plus the table of Categorical Variables Codings) .
 6. Fit a logistic regression model with DIED as the dependent variable and systolic blood pressure at time 1 (SBP1) as the predictor.
 - a) What is the Nagelkerke R-square for this model?
 - b) What is the Odds Ratio for SBP1?
 - i. Interpret this Odds ratio in words
 - c) Include the Model Summary table and the Variables in the Equation tables in your write-up.

7. Fit a logistic regression model with DIED as the dependent variable and both SHOKTYPE and SBP1 as predictors. Be sure SHOKTYPE is included as a categorical predictor and that the first category is the reference.
 - a) Is there a significant relationship between SHOKTYPE and the probability of dying in the model when it is added alone? (Check the test for SHOKTYPE in the table labeled "Variables not in the Equation". This test should have 5 degrees of freedom).
 - b) Is there a significant relationship between SBP1 and the probability of dying in the model when SBP1 is added alone? (Again, check the test for SBP1 in the table labeled "Variables not in the Equation").
 - c) Is there a significant relationship between SHOKTYPE and the probability of dying in this new model, after controlling for SBP1? (look at the Wald test for SHOKTYPE overall; this test will be in the table labeled "Variables in the Equation", and will have 5 degrees of freedom).
 - d) Is there a significant relationship between SBP1 and the probability of dying in this new model, after controlling for SHOKTYPE? (Look at the Wald test for SBP1 in the table "Variables in the Equation").
 - e) Include the tables: Case Processing Summary, Categorical Variables Codings, Variables not in the Equation, Omnibus Test of Model Coefficients, Model Summary and Variables in the Equation, in the write-up.

8. Create a cross-tab with URINECAT1 as the row variable and URINECAT2 as the column variable. Include only the total percentages for this table, and not the row percents and column percents.
 - a) What percentage of patients had any urinary output at time 1?
 - b) What percentage of patients had any urinary output at time 2?
 - c) Carry out a McNemar's test of symmetry to compare the percentage of patients with any urinary output at time 1 to the percentage of patients with any urinary output at time 2.
 - i. Report the results of your test.
 - ii. What do you conclude from this test? Please explain your results.
 - d) Include the cross-tab for this question in your write-up.

9. Extra Credit: Create a bar chart showing the Mean of DIED for each level of SHOKTYPE. Include this bar chart in your write-up (don't put error bars on this chart). Interpret this bar chart.

Include the printout of your SPSS commands as the first part of your write-up. You will be graded on your SPSS commands, the output and its appearance, and your interpretation.