

Biostat 510 Winter 2005

SPSS Homework #9

Due Tuesday, April 5, 2005

Be sure to include your SPSS commands with your homework. Provide appropriate output for each question (don't include output for questions that require syntax only).

- 1) Read in all variables for the Afifi data from Afifi.dat so that each observation has a time1 and a time2 value for each variable. Be sure to include decimal points appropriately in your data list statement. Your data set should have 113 observations.
- a) Create **value labels** for the variables SEX, SURVIVE, and SHOKTYPE

<u>Sex Value</u>	<u>Label</u>
2	Male
3	Female

<u>Survive</u>	<u>Label</u>
1	Survived
3	Died

SHOKTYPE	
<u>Value</u>	<u>Label</u>
2	Non-shock
3	Hypovolemic
4	Cardiogenic
5	Bacterial
6	Neurogenic
7	Other

Recode new variables and assign value labels (for b through d below). Use Transform...Recode...into different variables.

- b) Recode Shoktype into SHOCK, as shown below. Assign value labels.
- 1=Shock
 - 2=No Shock
- c) Recode Survive into DIED. Assign value labels to this new variable.
- 1=Died
 - 2=Lived
- d) Recode Urine1 and Urine2, into UR1 and UR2 and assign value labels, as shown below:
- 1=No urinary output
 - 2=Some urinary output.
- e) Create a new variable called SHOCKDUM that is coded as:
- 1=In shock
 - 0=Not in shock
- f) Create a new variable called DIEDDUM that is coded as:
- 1=Died
 - 0=Lived

- g) Get frequency tabulations for Sex, Survive, Shocktype, and all of the new variables created in this problem. Check your new variables by getting oneway frequencies for them.
 - h) Get descriptive statistics on all variables in the data set. How many patients have complete data on all variables? Make sure your new variables have the same number of cases as the original variables.
 - i) Include these frequency tabulations and descriptive statistics in the output that you hand in.
- 2) Crosstab with a chi-square test for independence..
- a) Get a crosstab of Shoktype (row variable) and Survive (column variable). Be sure to get the frequencies, row percents, and a chi-square test of independence for these two variables.
 - b) What percent of patients in each type of shock survived? Note: in the dialog box, row percents can be chosen after clicking on the Cells button and then selecting the appropriate percents—be sure observed counts is still selected.
 - c) Carry out a chi-square test of independence between shoktype and survive. Please write out the chi-square test statistic, the degrees of freedom and the p-value. What do you conclude about the relationship between shock type and survival from this test?
 - d) Include this crosstab in your output.
- 3) Crosstab for a 2x2 table, with chi-square test, relative risk and odds ratio
- a) Get a crosstab of SHOCK (the 1, 2 coding) as the row variable and DIED (the 1,2 coding) as the column variable, with row percents.
 - b) Get a chi-square test of independence, the relative risk and the odds ratio.
 - c) What do you conclude about the relationship between these two variables? (Be sure to include the test statistic, degrees of freedom and p-value).
 - d) Report the relative risk and the odds ratio, their 95% confidence intervals, and their p-values. Compare and interpret both of these measures.
- 4) Crosstab for related proportions
- a) Get a crosstab of UR1 as the row variable and UR2 as the column variable, with **table percents.**
 - b) What percent of patients had some urinary output at time 1? What percent of patients had urinary output at time 2?
 - c) Test the null hypothesis that the proportion of patients with at least some urinary output is the same for time1 and time2.
 - d) What statistical test is appropriate here? (Remember, these are correlated proportions, so a Pearson chi-square is not appropriate).What do you conclude based on this test?
 - e) Include your crosstab and the statistical test results in your output.
- 5) Regression analysis with continuous and dummy variables as predictors.
- a) Run a regression model with MAP2 as the dependent variable and the predictors: shockdum (the 1,0 dummy variable), MAP1, SBP1, body surface area at time 1, and urinary output at time1. Include all of these predictors in the model at once.
 - b) Select the options to obtain collinearity diagnostics, plots of residuals vs. predicted values (zresid and zpred would be OK), Histogram of residuals and normal p-p plot of residuals.
 - c) What is the overall significance of the model? The model R-square? The adjusted R-square?

- d) Which variables are significant in this model? Interpret the results for each **significant** predictor variable.
 - e) Check the collinearity diagnostics for the model, including tolerance, VIF, and the condition index for the model. What do you conclude about collinearity for this model? Is it a problem? If it is, which variables are collinear?
 - f) Examine the plot of the residuals vs. predicted values for this model. What does this plot tell you about the adequacy of your model?
 - g) Examine the histogram of the residuals from this model. How symmetric do the residuals appear to be?
 - h) Examine the normal p-p plot of residuals. How normal do the residuals appear to be?
 - i) Turn in the regression output and plots from this regression analysis..
- 6) Logistic Regression with a categorical predictor. (Analyze...Regression...Binary Logistic)
- a) Run a logistic regression in which DIEDDUM (the 0,1 variable) is the dependent variable and SHOKTYPE (the original categorical variable with all the different categories of shock) is the predictor.
 - b) Be sure to include shoktype as a **categorical** variable in the model, with the first level of shoktype (shoktype=2, non-shock) as the reference category. (Click on categorical, in the "Define Categorical Variables" window, enter SHOKTYPE as a categorical variable. Select Reference Category "First" and click on the "Change" button. Click on "Continue").
 - c) What is the Nagelkerke R-square for this model? (Note: this is the same as the Max. Re-scaled R-square from Proc Logistic in SAS)
 - d) Which levels of shoktype are significant in the model? Discuss the meaning of the odds ratios for **each level** of shoktype.
 - e) Include computer output for this logistic regression in your homework.
- 7) Logistic Regression with categorical and continuous predictors.
- a) Run a logistic regression in which DIEDDUM is the dependent variable and SHOCKDUM (the 0,1 variable), BSA1, and SBP1 are the predictor variables.
 - b) What is the overall likelihood ratio test for this model? What is the Nagelkerke R-square?
 - c) Interpret the parameter estimates and odds ratios for each predictor.
 - d) Include computer output for this logistic regression in your homework.
- 8) Save your data file as Afifi.sav and save your syntax as homework9.sps. Include the syntax with your homework.