

Biostat 510
Homework 9
Due Thursday, April 6, 2006

For the SAS part of this homework, use the permanent SAS data set, afifi.sas7bdat that you created for homework 3, or you can access it from my web page. **Be sure to use the permanent version of the SAS data file – and not just read in the raw data file. If you create the permanent SAS data set from scratch, as for homework 3, close SAS, and reopen it, and use SAS syntax to use the permanent file.**

For the SPSS part of the homework, you will be using the SPSS data file afifi.sav, which you can also get from my web page, or you can create it using SPSS syntax, and save it as an SPSS data set and open it to use it in the homework.

SAS and SPSS:

1. Frequency tabulations and crosstabs.
 - a. Get oneway frequency tabulations of DIED, SHOCK, and SHOKTYPE (your variable may be spelled as SHOCKTYPE). Be sure that DIED is coded as 0 and 1, with 1 indicating that the patient died and 0 indicating that the patient did not die. Also check to be sure that SHOCK is coded as 0 and 1, with 1 indicating that the patient was in shock when admitted to the hospital, and 0 indicating that they were not in shock.
 - b. Get a crosstab of SHOCK as the row variable vs. DIED as the column variable, along with a chi-square test of independence (do not get relative risk for this 2x2 table, because it is not set up properly for this statistic).
 - c. Get a crosstab of SHOKTYPE as the row variable vs. DIED as the column variable, along with a chi-square test of independence.
2. Logistic regression with a continuous predictor
 - a. Carry out a logistic regression to predict DIED using SBP1 as a predictor in SAS and SPSS. (Be sure to use the descending option in SAS for this and all problems in this homework).
 - b. In SPSS, go to Analyze...Regression...Binary Logistic. Choose DIED as the dependent variable and SBP1 as the covariate. Select the "Options" button at the bottom, and click on "CI for exp(B)" to get a 95% confidence interval for the parameter estimates.
 - c. Get the pseudo R-square statistic for both SAS and SPSS (use the max rescaled r-square in SAS and Nagelkerke r-square in SPSS), You need to specify the rsquare option in SAS to get this statistic. It is automatic in SPSS.
 - d. *For extra credit in SAS, get the ods output graph of the predicted probability vs. SBP1 value, as shown in the handout.
3. Logistic regression with a binary predictor
 - a. Carry out a logistic regression to predict DIED using SHOCK as the predictor in SAS and SPSS.

- b. Get the pseudo R-square statistic for both SAS and SPSS.
 - c. Do not set up SHOCK as a categorical predictor. Because it is already a dummy variable, it can simply be put into the model.
4. Logistic regression with a categorical predictor
- a. Carry out a logistic regression to predict DIED using SHOKTYPE (which may be spelled as SHOCKTYPE in your data set) as the predictor in SAS and SPSS. Use SHOKTYPE=2 (non-shock) as the reference category.
 - b. Use the Class statement to specify SHOKTYPE as a categorical predictor in SAS (use the ref="2" option in SAS to have non-shock be the reference category).
 - c. In SPSS, select SHOKTYPE as the covariate, and then click on the "Categorical..." button. In the window that opens, select SHOKTYPE and put it into the "Categorical Covariates" box on the right, Click on "First" as the reference category, and then click on "Change" to make your choice take effect, then click on "Continue" to return to the logistic regression setup window.
 - d. Get the model pseudo r-square in both SAS and SPSS.
5. Logistic regression with a categorical and continuous predictor.
- a. Carry out a logistic regression to predict DIED using SHOKTYPE and SPB1 as the predictors in SAS and SPSS. Use SHOKTYPE=2 (non-shock) as the reference category for SHOKTYPE, as before.
 - b. Get the model pseudo r-square in both SAS and SPSS.
6. Answer the following questions about your analyses.
- a. Frequencies and crosstabs:
 - i. What number and percent of patients died? lived?
 - ii. What number and percent of patients who were in shock died? What number and percent of patients who were not in shock died? What is the value of the Pearson chi-square test for SHOCK vs. DIED, what is the p-value and degrees of freedom?
 - iii. What number and percent of patients in each category of SHOKTYPE died? What is the value of the Pearson chi-square test for SHOKTYPE vs DIED? Give the chi-square statistic, df and p-value.
 - iv. Do your SAS and SPSS results agree? If not, fix your data and rerun these items.
 - b. Logistic regression with a continuous predictor:
 - i. What is the overall test of significance for SBP1 as a predictor of DIED? (Report the likelihood ratio chi-square test statistic, the df and p-value).
 - ii. What is the coefficient for SBP1? Please give an interpretation of this coefficient in terms of the change in the log-odds of dying for a one-unit increase in SBP1.
 - iii. What is the odds ratio for SBP1? Note that the odds ratio is less than one, which means that the odds of dying go *down* as SBP1 increases.

- iv. What is the pseudo r-square for this model (report the max-rescaled r-square in SAS and the Nagelkerke r-square in SPSS).
 - v. Do your SAS and SPSS results agree?
- c. Logistic regression with a binary predictor:
- i. What is the overall test of significance for SHOCK? Report the likelihood ratio chi-square test, df and p-value.
 - ii. What is the coefficient for SHOCK? Please interpret this coefficient in terms of a change in the log-odds of dying for a person who is in SHOCK vs. a person who is not in SHOCK.
 - iii. What is the odds ratio for SHOCK? Please interpret this coefficient in terms of a change in the odds of dying for person who is in SHOCK vs. a person who is not in SHOCK.
 - iv. What is the pseudo r-square for this model?
 - v. Do your SAS and SPSS results agree?
 - vi. Compare the Wald chi-square test for SHOCK in the logistic regression with the chi-square test for independence in the crosstabs of SHOCK vs. DIED in SAS. Do they agree? Compare the likelihood ratio test from the logistic regression to the likelihood ratio chi-square from the crosstabs in SAS. Do they agree?
- d. Logistic regression with a categorical predictor:
- i. What is the overall test of significance for SHOKTYPE? Report the likelihood ratio chi-square test, df and p-value.
 - ii. What is the coefficient for each type of shock, vs. non-shock. Interpret these coefficients in terms of the log-odds of dying for a person who is in each type of shock vs. a person who is not in shock.
 - iii. What is the odds ratio for each type of shock? Please interpret these coefficients in terms of a change in the odds of dying for a person who is in each type of shock vs. a person who is not in shock.
 - iv. What is the pseudo r-square for this model?
 - v. Do your SAS and SPSS results agree?
 - vi. Compare the Wald chi-square test for SHOKTYPE in the logistic regression with the chi-square test for independence in the crosstabs of SHOKTYPE vs. DIED in SAS. Do they agree? Compare the likelihood ratio test from the logistic regression to the likelihood ratio chi-square from the crosstabs in SAS. Do they agree?
- e. Logistic regression with a categorical and continuous predictor.
- i. What is the overall test of significance for this model? Report the likelihood ratio chi-square test, df and p-value.
 - ii. What is the overall significance of SHOKTYPE in this model? Report the chi-square test, df and p-value.
 - iii. What is the test of significance of SBP1 in this model? Report the chi-square test, df and p-value.