

Homework 4 - due 4/29/13

Please submit your assignment in Word or PDF format by 9:30 am (PDT) to the Biost 513 Assignment Dropbox at <http://canvas.uw.edu/>. You may go to “Assignments” at the left side menus on your Canvas page and then follow the instructions to upload your assignment.

NOTE: Unless explicitly stated, direct computer output is not desired. Typically only part of the computer output is asked for (such as a confidence interval) and then proper interpretation of the statistics is requested.

DATA: The data for these exercises can be found on the class web page:
<http://courses.washington.edu/b513/> in the Homework directory.

Stata.help4: Key Stata commands that are useful for these exercises are described in the text file stata.help4, also available in the homework directory.

PLEASE START YOUR RESPONSE TO EACH QUESTION ON A SEPARATE PAGE AND PUT YOUR NAME ON EACH PAGE.

Interpretation of logistic regression coefficients

- 1) Suppose you code your exposure variable as 1 = exposed, 2 = unexposed and receive the following logistic regression output relating an exposure to development of a disease (assume this is the output of logit)

case	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
exposure	-1.50	etc			
_cons	-0.60	etc			

- a) Provide an interpretation of the coefficient of “exposure”. Specifically, what is the estimated odds ratio for disease given exposure? Does the exposure increase or decrease the risk of disease?
- b) Suppose that the exposure variable had been coded as 0 = unexposed, 1 = exposed. What would the coefficients in the logistic regression have been?

Logistic Regression – Confounding and Effect Modification

- 2) The dataset athens.dat was analyzed in Homework set #2. Define the variables from that analysis as follows:

Y = CHD case = 1 / control = 0 (the outcome variable)
 NEWIRON = 1 if >350mg/month iron / 0 if ≤350mg/month iron
 FEMALE = 1 if female / 0 if male
 AGE = identifier for age group, 1 through 4

- (a) Write a logistic regression model using Y as the response and NEWIRON as the predictor of interest (just write the model, don’t fit it yet).

- (b) Interpret the coefficient of NEWIRON in this model in terms of an odds ratio comparison.
- (c) Use STATA to fit the model in (a) and test whether the NEWIRON coefficient is significantly different from 0.
- (d) Write a logistic regression model that uses NEWIRON, FEMALE, and AGE dummy variables as the predictors.
- (e) Interpret the coefficient of NEWIRON in this model in terms of odds ratio comparisons.
- (f) Interpret the coefficient of AGE(2) (ie. the indicator variable for the third age group) in this model in terms of odds ratio comparisons.
- (g) Interpret the coefficient of AGE(4) (ie. the indicator variable for the fourth age group) in this model in terms of odds ratio comparisons.
- (h) In terms of your logistic regression model coefficients, what would be the odds ratio comparing (NEWIRON=0, FEMALE=0, AGE=4) to (NEWIRON=0, FEMALE=0, AGE=2)?
- (i) Fit logistic regression models with FEMALE and AGE dummy variables (without NEWIRON) and then the model with NEWIRON, FEMALE and AGE dummy variables. Use these models to perform a likelihood ratio test of the coefficient of NEWIRON. State the null hypothesis, test statistic, and interpret the p-value.
- (j) Compare the estimated NEWIRON odds ratio obtained using logistic regression to that obtained via Mantel-Haenszel in Homework set #2. Do these odds ratios have the same interpretation? Justify.
- (k) If your primary interest is in assessing the effect of NEWIRON on (log)odds of CHD, are the AGE variables important variables to include in the model? Justify.
- (l) If your primary interest is in assessing the effect of NEWIRON on (log)odds of CHD, is FEMALE an important variable to include in the model? Justify.
- (m) Is the odds ratio the same for men and women? Formulate a logistic regression model that allows the NEWIRON odds ratio to depend on FEMALE. In terms of the regression coefficients in this model, represent the null hypothesis that the odds ratios are homogeneous across gender. State your conclusions.

Logistic Regression – Testing for Trend – Coming soon...