

Homework 8 - Due 6/5/13

Please submit your assignment in PDF format by 9:30 am (PDT) by **Wednesday, June 5, 2013** 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.help8: 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.

Survival Analysis: Model Adequacy and Stratified Analysis

- 1) The file **addicts.dta** on the class web page contains data regarding the time that heroin addicts remain in methadone treatment. These data were analyzed by Caplehorn and Bell (1991) who were interested in factors associated with retaining subjects: "As methadone maintenance is of proven benefit only to those in treatment, retention in treatment is an important measure of the effectiveness of treatment programmes." and "To elucidate the reasons that programmes fail to retain patients, we have studied the relationship between the maximum daily dose and retention in a cohort of addicts." Scientific interest is in whether factors other than dose can be used to identify subjects at high risk for failing to be retained.
 - (a) Create a Cox regression model with clinic as the sole predictor and use `stphptest` to show that the proportional hazards assumption is not satisfied for clinic. How can we control for clinic in a Cox model if clinic does not satisfy the PH assumption?
 - (b) Calculate bivariate summaries for each of the predictor variables and their association with time retained in treatment. Summarize these by creating a table of hazard ratios and 95% CI's for each variable when it is the single predictor in a Cox regression that uses clinic as a stratifying variable. Interpret the coefficient for female gender from the first bivariate Cox regression model.
 - (c) Calculate a Cox regression model using all of the predictors. Summarize the results by creating a single table of regression parameters (or hazard ratios) - use the computer output as directly as possible in order to create this table. (again stratify on clinic). Interpret the coefficient for female gender from this multivariable Cox regression model.
 - (d) What variables appear to be predictive of retention? Simplify the model in (c) by eliminating variables that do not appear to be useful in predicting retention. (Variable selection methods may be acceptable to use here.) Create a table as in (c) for your final model.

- (e) Create log(-log) plots for the variables in your final model (you will need to categorize continuous variables for your plots). Hand these plots in.
- (f) Check the PH assumption for the variables in your final model using `estat phtest` (after fitting your Cox regression model). Also, plot the scaled Schoenfeld residuals versus time for the variables in your final model and hand these plots in. You can use the option, “plot(*var*)” with the `estat phtest` command to obtain these plots. (Please see the Stata.help8 file for additional details.) What do you conclude about the PH assumption for each variable in your model?
- (g) Use martingale residuals to check the functional form for any continuous variables in your final model. (You can obtain martingale residuals when you fit your Cox regression model. (See the Stata.help8 file for details.)
- (h) Describe the assumptions in your Cox regression model and give the model form. Explain what it means to use clinic as a stratifying variable.