

Biostatistics 208 HW #1, Due 1/28/14

INSTRUCTIONS

The first part of the assignment is to complete Problems 2.1-2.6 at the end of chapter 2 in the course textbook, *Regression Methods in Biostatistics*. The problems are also posted on the course website. For Problems 2.2 and 2.3, you can control the number of observations using the `set obs` command before you generate the pseudo-random data, as shown in the first lab. The datasets for Problems 2.4-2.6 can be found on the course website.

The second part of the assignment involves a linear regression model with a categorical predictor, the four-level behavioral pattern variable `behpat` in the WCGS dataset used in Problem 2.6.

- Using the `tab`, `sum( )` command, compute average systolic blood pressure (variable `sbp`) according to behavioral pattern. (Note that the `tabstat` command can also be used here.)
- Using the `i.` prefix (to be discussed in the second lecture and lab) for the categorical predictor variable `behpat`, run an unadjusted linear regression with outcome `sbp`.
- Write down the values of the indicator variables required to represent `behpat` in the model, using a table similar to table 4.3 in the textbook.
- Using the `lincom` command, and the appropriate `i.` prefixes to specify the various behavior patterns, obtain the average SBP in the four behavior pattern groups.
- Use the `testparm` command to obtain an F -test of the null hypothesis that mean SBP is the same in all four behavioral pattern groups. Write a brief description of the results.

Please hand in a hard copy of your work by noon on Tuesday, Jan 29, 2013. This can be done in lecture, or given directly to Olivia DeLeon. (Note that hard copies are required for grading, so please don't submit via email).