CPH 636 — Spring 2009 — Dr. Charnigo

Written Assignment 1

Written Assignment 1 is due on Monday 02 February at the end of lecture, but *I request that you do task 1 before the lecture on Monday 26 January.* You are encouraged to work in groups of two or three, though you may work individually if you prefer. *If you work in groups of two or three, please be sure that each group member is able to run the SAS macros, as each group member must be able to use the SAS macros for the noncollaborative final project.*

[20] 1. To prepare for an activity in Lecture 2, visit {http://www.humanmetrics.com/cgi-win/JTypes2.asp} and take the personality test. Record whether you are classified as an introvert ("I") or an extrovert ("E"), intuitive ("N") or sensing ("S"), thinking ("T") or feeling ("F"), judging ("J") or perceiving ("P").

[10] 2. Download the macros and data sets from Fernandez's web site. Detailed instructions are available at {SASMacroInstr.txt} on the CPH 636 web site. You do not need to hand anything in for this task; completion of task 4 below will be *prima facie* evidence that you have completed this task successfully.

[10] 3. Apply the EXCELSAS macro to create a SAS data set from {Diabetes.xls}, which may be found on the CPH 636 web site. Detailed instructions are available at {SASMacroInstr.txt} on the CPH 636 web site. Again, you do not need to hand anything in for this task; completion of task 4 will be *prima facie* evidence that you have completed this task successfully.

[20] 4. Apply the RANSPLIT macro to the SAS data set that you created in task 3. Divide it into three subsets: training (50%), validation (25%), and testing (25%). Detailed instructions are available at {SASMacroInstr.txt} on the CPH 636 web site. Please hand in a printout of the first ten observations in the training subset.

[20] 5. Using the SAS data set that you created in task 3, run a logistic regression with DIAB as the response variable and with NPREG and GLU as explanatory variables. The syntax in point 9 of "Use the EXCELSAS macro." in {SASMacroInstr.txt} can be used verbatim. Do not worry if you have not studied logistic regression before; we will talk about it later this semester. Please record the point estimates of the odds ratios for NPREG and GLU.

[20] 6. Using the training subset that you created in task 4, run a logistic regression with DIAB as the response variable and with NPREG and GLU as explanatory variables. The syntax in point 9 of "Use the EXCELSAS macro." in {SASMacroInstr.txt} can be used verbatim except that you will need to change "RCLib.Diabetes" to reference your training subset. Please record the point estimates of the odds ratios for NPREG and GLU. Are they the same as the point estimates you obtained in task 5? Did you expect them to be? Why or why not?