November 22, 2012
Morten Frydenberg

POSTGRADUATE COURSE IN
LINEAR AND LOGISTIC REGRESSION
Day 4

Consider the dataset: "`obese.dta`". In this exercise, focus is on high systolic blood pressure (`sbp`) and possible explanatory variables Body Mass Index (`bmi`), age, and `sex`.
We will here define high systolic blood pressure as SBP >140 mmHg.
I suggest that you use age=50 years and BMI==25 kg/m^2 as reference values.

1. Consider the logistic regression model with response high SBP and age, sex and BMI as explanatory variables – the model should include an interaction between sex and age.
   Write down the model for the log odds of having high SBP.
   Write down the model for the risk of having high SBP.

2. Fit the model.
   Explain all the coefficients in the model.
   Explain all the hypotheses tested in the output.

3. Plot the log odds for high SBP as a function of age for four 'persons' a man and a woman with BMI=25 kg/m^2 and BMI=30 kg/m^2.
   Plot the risk (probability) of high SBP as a function of age for four 'persons' a man and a woman with BMI=25 kg/m^2 and BMI=30 kg/m^2.

4. What is the OR comparing a 40 year old man and woman adjusted for BMI?
   What is the OR comparing a 60 year old man and woman adjusted for BMI?
   What is the risk (probability) of high SBP for a 40 year old man with BMI=27?

5. Plot the risk (probability) of high SBP as a function of BMI for a 40 year old man and a 40 year old woman?

6. Discuss, based on what you have seen, whether or not it is reasonable to assume no effect modification between age and sex.

7. In the model above BMI is introduced as one continuous variable (linear on the logit scale) with no interaction with sex or age. Fit two or more models in order to check the validity of these assumptions.
   Write a summary of your findings.
   What is your conclusion?