EPIB 613
November 1, 2012
Exercise #5

Copy the datasets smoke.dta and lowbwt.dta from the course website

Using the dataset smoke.dta:

1. Generate a new variable called bmi for body mass index:
   (BMI=mass(kg)/(height(m))^2). How many women have a BMI greater than 30?

2. Generate a new variable called bmi_high that equals 1 if bmi equals 25
   or greater and 0 if less than 25. What is the mean height for female
   never or former smokers with high BMI?

3. Replace the values of bmi_high so that men are considered as having a
   high BMI only if their value is greater than 28.

4. List the ages, smoking status, and bmi for men now considered to have
   high bmi. Remove all dividers and separators from the results output.

Using the dataset lowbwt.dta:

5. Create a new variable named agegrp3, which contains quintiles of the
   variable age. Label the variable and the categories appropriately.

BONUS. Generate a risk score that classifies a women's risk of giving birth
to a low birthweight baby based on known risk factors. The risk score is a
combination of the following variables: smoking, hypertension, age, and
weight. The following number of points are allocated to the risk score for
each variable:

    Smoking: yes=1 points, no=0 points
    Hypertension: yes=1 point, no=0 points
    Age: <20 or 40+ =2 points, 35-39=1 point, 20-34=0 points
    Weight: >=175lbs=2 points, <175lbs=0 points

Create a new variable called riskcat that groups the risk score so that 0
points=low risk, 1-2 points=moderate risk, 3+points=high risk. What
proportion of white women are in the high-risk category?