Biostat II: Lab 4, Practice with confidence intervals and hypothesis tests
Date: 25 April 2008

1. The data in the following tables is a 30% sample from the 2003 Biostatistics class at JHSPH.

   Summaries for 30% sample of males

Variable	Observations	Mean	Std. Dev.	Min	Max
Age	34	29.2	4.13	22	38
Weight	34	165.6	26.2	120	250
Height	34	70.3	4.25	60	86
BMI	34	23.6	3.14	15.2	30.5

   Summaries for 30% sample of females

Variable	Observations	Mean	Std. Dev.	Min	Max
Age	73	29.2	6.97	21	55
Weight	73	133.8	22.7	95	198
Height	73	64.8	3.54	54	74
BMI	73	22.4	3.21	16.1	33.0

   Suppose that the average BMI (kg/m^2) of all females in the 2002 Biostatistics class was 25 kg/m^2. We conjecture that females in the 2003 class (from which the above 30% sample was taken) should have the same mean BMI. For the moment, assume that the true standard deviation of the sample mean for samples of size 73 is $\frac{3.21}{\sqrt{72}} = 0.38$ kg/m^2. Calculate the probability that the sample mean (for samples of size 73) is:

   (a) < 24.62
   (b) < 23.86
   (c) > 25.38
   (d) > 26.14
   (e) What is the probability of observing the sample mean of 22.44 kg/m^2 or a more extreme value?

2. Construct by hand a 95% confidence interval for the true BMI value for females in the class population based on the above sample.

3. Construct by hand a 95% confidence interval for the true BMI value for males in the class population based on the above sample. How does this confidence interval compare to the one you constructed for females?

4. The output below provides summary measures for the entire 2003 class. How do the actual population mean BMI values compare to the estimates and the 95% confidence intervals that you constructed? Comment.

   Summaries for all males

Variable	Observations	Mean	Std. Dev.	Min	Max
Age	93	30.4	6.80	21	50
Weight	93	168.4	28.8	83	270
Height	93	69.9	3.44	56	77
BMI	93	24.2	3.65	11.3	37.1

   Summaries for all females

Variable	Observations	Mean	Std. Dev.	Min	Max
Age	219	28.8	6.68	21	56
Weight	219	131.6	22.2	87	230
Height	219	64.4	3.75	46	76
BMI	219	22.4	3.91	14.0	41.6

5. By hand, test the null hypothesis that men and women have the same average BMI values.

   (a) Look at the available data and decide whether you think it is reasonable to assume variance is equal across males and females.

   (b) Find the test statistic and corresponding p-value. Write a sentence interpreting the p-value. *Hint: In R, you can find the* $P(t_d < x)$ *for a t distribution with d degrees of freedom using the command* `pt(x, d)`.

   (c) Construct the 95% confidence interval for difference in mean.

   (d) How do parts (b) and (c) compare?