Math 160/263 – Assignment #3

Title: Sampling Distributions of the Mean

Worksheet: None

The purpose of this assignment is to study the sampling distribution of the mean. You will learn how to:

- Compute the mean and standard deviation of this distribution
- Compute probabilities from its distribution and
- Simulate a sampling distribution.

This assignment uses material from chapter 10 in the Minitab Manual.

The physical fitness of a patient is often measured by the patient's maximum oxygen uptake, recorded in milliliters of oxygen per kilogram of body weight. The mean maximum oxygen uptake for cardiac patients who regularly participate in sports or exercise programs was found to be 24.1 ml/kg, with a standard deviation, σ, of 6.30 ml/kg. (*Adapted Physical Activity Quarterly*, October, 1997). We assume that the distribution of the maximum oxygen uptake for these patients is Normal.

Calculating Probabilities with Minitab

1. Calculate the probability that the maximum oxygen uptake in a randomly selected patient is between 21 and 28 ml/kg.
Put 21 and 28 in a column and use **Calc** ⇒ **Probability Distributions** ⇒ **Normal** and select "Cumulative probability" to calculate the cumulative probabilities for the two endpoints and subtract them using Minitab's calculator.

The Sampling Distribution of the Mean

2. Consider a simple random sample of 20 such cardiac patients.

a. Using what you know about sampling distributions from class, find the mean and standard deviation of the sampling distribution of the mean maximum oxygen uptake for all such samples of 20, and its standard deviation.

b. Find the probability that this mean of a sample of 20 lies between 21 and 28 ml/kg. Use the same commands as in problem #1, but with the correct mean and standard deviation.

c. Explain the difference between the answer to this question and the answer to #1.

3. Now create a simulated version of the Sampling Distribution of the Mean for samples of size twenty. Use **Calculate** ⇒ **Random Data** ⇒ **Normal** to calculate 100 samples of 20 patients each from the distribution for mean oxygen uptake. You will need 100 rows of 20 columns(C1-C20) each(use mean 24.1 and standard deviation 6.3). (this will overwrite the entries made in problem #1, so be sure the results of problem #1 are in the Report-Pad). Once this is done, we change our perspective and consider each row as representing one sample of size twenty.

4. Use **Calc** ⇒ **Row Statistics** to compute the means of the rows. Put C1-C20 in "Input variables". Store the means in C21. This represents a sample of size 100 from the Sampling Distribution of the Mean of samples of size twenty.

5. Create a histogram of C21, using **Graph** ⇒ **Histogram** ⇒ **Simple**. In addition, calculate its mean and standard deviation with **Stat** ⇒ **Basic Statistics** ⇒ **Display Descriptive Statistics**. You will need to select **Statistics** to choose what Minitab will calculate.
Are these calculated statistics as expected from the class discussion of the Sampling Distribution of the Mean? Explain.

Looking at the Central Limit Theorem with Minitab

According to the University of Arizona Fact Book, 2004-2005, 15.8% of all students enrolled in the University of Arizona in the Fall Semester 2004 are Hispanic. Consider a Simple Random Sample of 30 students. Then X, the number of Hispanic Students in the sample, has a Binomial Distribution with parameters $n = 30$ and $p = 0.158$.

6. Open a new worksheet. Use the **File** ⇒ **New** ⇒ **Minitab Worksheet**.

7. Use the **Calculate** ⇒ **Random Data** ⇒ **Binomial** to calculate 500 samples (500 rows) of the distribution $B(30, 0.158)$ in first forty columns(C1-C40) in your worksheet. Use **Calc** ⇒ **Row Statistics** command to compute the row means for the first ten columns of simulated data, and store these in C41.The rows of the first ten columns represent samples of size ten from the sampling distribution.

8. Now do the same with the first 20 columns, and put the results in the C42.

9. Repeat the same procedure with all 40 columns and put the results in C43. What do the results of #7 and #8 represent?

10. Use the **Graph** ⇒ **Boxplots** ⇒ **Multiple Y's Simple** command to create graphical descriptions of the observations in the first row of simulated data (C1), and the three means of different sample sizes (C41-C43).

11. What happens to the mean and standard deviation of the distributions of the sampling distribution of the mean as sample size increases? Verify your statements by computing the means and standard deviations of the first column of data (C1), and the means from different sample sizes(C41-C43). Use **Stat** ⇒ **Basic Statistics** ⇒ **Display Descriptive Statistics** What are your conclusions? Now save and edit your report for submission.