

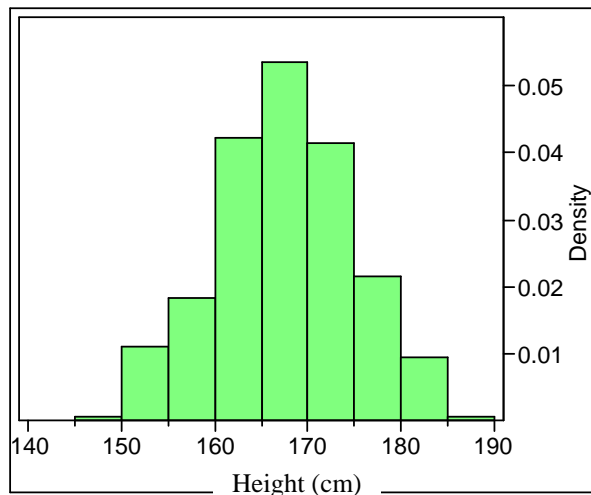
## Statistics 101 – Laboratory 11

### Distribution of the sample mean $\bar{y}$ .

This lab looks at the sample mean  $\bar{y}$  by sampling from a population of 250 females who took an introductory statistics class. The characteristic we are interested in is the average or mean height of female students. We can look at the variation in a sample statistic by taking many samples (called repeated sampling) from the population and looking at the distribution of the sample statistics obtained.

1. Refer to the table titled “Heights of all females in the population.” This table contains a listing of the heights, in centimeters (cm), of the population members. Rather than list the names of the population members, this table numbers them by rows numbered (00, 01, . . . , 09, 10, . . . , 24) and columns numbered (0, 1, 2, . . . , 9). For example, student 037 (Row 03 and Column 7) is 165 cm tall.
  - a) Use the random number table to select a simple random sample of 10 students from this population. Write the student numbers and heights on the answer sheet. Calculate the sample mean height. Note: You can make more efficient use of 3 digit random numbers doing the following. For numbers between 000 and 249 go directly to the table of heights. For numbers between 250 and 499 subtract 250 and then go to the table of heights. For numbers between 500 and 749 subtract 500 and then go to the table of heights. For numbers between 750 and 999 subtract 750 and then go to the table heights.
  - b) Take three other random samples of size 10 from the population. This requires different random numbers than the ones used in a). Record the student numbers and heights on the answer sheet and calculate the sample mean for each sample.
  - c) How different are the four sample means? What causes the differences?

This population has a population mean height  $\mu = \mathbf{166.6\text{ cm}}$  and a population standard deviation  $\sigma = \mathbf{7.6\text{ cm}}$ . The distribution of heights is given below.



We are interested in the distribution of the sample mean,  $\bar{y}$  for random samples of size  $n = 10$ .

- d) Describe the shape of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 10$ .
- e) What is the mean of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 10$ ?
- f) What is the standard deviation of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 10$ ?

Combine your first two samples together to create a random sample of size  $n = 20$ .

- g) Calculate the sample mean for this larger random sample,  $n = 20$ .
  - h) Describe the shape of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 20$ .
  - i) What is the mean of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 20$ ?
  - j) What is the standard deviation of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 20$ ?
  - k) Which should be a better estimate of the population mean, the sample mean for a random sample of size  $n = 10$  or size  $n = 20$ ? Support your answer statistically.
2. For this problem you will use the random samples you collected earlier in the lab. Each member of the group should take one of the random samples and
- a) Calculate the sample mean and sample standard deviation. Use your calculator to do these calculations.
  - b) Construct a 95% confidence interval for the population mean height of female students.
  - c) How many of your confidence intervals contain the population mean  $\mu = 166.6$  cm?
  - d) Put this on the white board. Once all the groups in lab have put their numbers of intervals that contain the population mean on the board, compute a relative frequency of capturing the population mean for your lab section. You can continue to work on other parts of this lab while waiting for all the groups to put their numbers on the board.
  - e) What should be the long run relative frequency of capturing the population mean?
  - f) Suppose you were to combine the first two random samples of size  $n = 10$  to create a random sample of size  $n = 20$ . Would the confidence interval for the population mean height based on the sample of size  $n = 20$  be wider, about the same width or narrower than a confidence interval for the population mean height based on the sample of size  $n = 10$ ? Explain your answer briefly. Note: You do not, and should not, compute the confidence interval for the sample of size  $n = 20$ .

## Stat 101 – Laboratory 11 Group Answer Sheet

Names of Group Members: \_\_\_\_\_, \_\_\_\_\_  
 \_\_\_\_\_, \_\_\_\_\_

1.

Sample 1		Sample 2		Sample 3		Sample 4	
Number	Height	Number	Height	Number	Height	Number	Height
Sample Mean		Sample Mean		Sample Mean		Sample Mean	

- c) How different are the four sample means? What causes the differences?
  
- d) Describe the shape of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 10$ .
  
- e) What is the mean of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 10$ ?

- f) What is the standard deviation of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 10$  ?
- g) Calculate the sample mean for this larger random sample  $n = 20$  .
- h) Describe the shape of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 20$  .
- i) What is the mean of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 20$  ?
- j) What is the standard deviation of the theoretical distribution of the sample mean,  $\bar{y}$  for samples of size  $n = 20$  ?
- k) Which should be a better estimate of the population mean, the sample mean for a random sample of size  $n = 10$  or size  $n = 20$  ? Support your answer statistically.

2. Each member of the group should take one of the random samples and

			95% Confidence Interval for $\mu$	
	Sample Mean, $\bar{y}$	Sample Std. Dev., $s$	$\bar{y} - t^* \left( \frac{s}{\sqrt{n}} \right)$	$\bar{y} + t^* \left( \frac{s}{\sqrt{n}} \right)$
Sample 1				
Sample 2				
Sample 3				
Sample 4				

- How many of the intervals contain  $\mu = \mathbf{166.6\text{ cm}}$  ?
- The relative frequency of capturing the population mean for your lab section.
- What should be the long run relative frequency of capturing the population mean?
- Suppose you were to combine the first two random samples of size  $n = \mathbf{10}$  to create a random sample of size  $n = \mathbf{20}$ , would the confidence interval for the population mean height based on the sample of size  $n = \mathbf{20}$  be wider, about the same width or narrower than a confidence interval for the population mean height based on the sample of size  $n = \mathbf{10}$  ? Explain your answer briefly.