

[Return to Cover Page](#)

LESSON 10 - CENTRAL LIMIT THEOREM

The purpose of this lesson is to illustrate the concepts involved in the Central Limit Theorem. In particular, we illustrate through repeated sampling, how to estimate the mean and standard deviation of a distribution of sample means. We will also see that the distribution of sample means is close to being a normal distribution if the samples are large enough.

Start Minitab and retrieve worksheet DJSS5.MTW that you created in Lesson 1. Delete everything below the date/time stamp and type your name, Lesson 10, and Example as usual.

Use the procedures from Lesson 5 and Lesson 6 to display the count, mean, and standard deviation of the variable DJC in the session window.

Create 10 more samples with replacement of size 5 as you did in Lesson 1. They should be stored in C7 through C16 and label them Sample6 through Sample15. Now click **Stat > Basic Statistics > Store Descriptive Statistics**. Select Sample1 through Sample15 into the "Variables:" box. Click on "Statistics" and make sure that "Mean" is the only box that is checked. Now click "OK" and "OK". You will see the means of the 15 samples stored in columns C17 through C31, and Minitab has labeled them Mean1 through Mean15. Now click on **Data > Stack > Columns**. In the dialog box that opens, select Mean1 through Mean15 into the "Stack the following columns:" box, click the "Column of current worksheet:" button, and type C32 into the related input box. When you click "OK", the means will be copied into column C32. Label this column "Means5". Now use the procedure from Lessons 5 and 6 to display the count, mean, and standard deviation of Means5 in the session window.

Now close this worksheet, and retrieve DJSS30 that you created in Lesson 5. Create 10 more samples with replacement of size 30 and repeat the procedures above to compute the mean and standard deviation of the means of these 15 samples. Call the column of means in C32 "Means30" in this case.

After deleting extraneous instructions, your session window should now look like the figure to the right. Your numbers will be different since you will have produced different random samples, but they should be similar.

First, we observe that mean of the sample means from samples of size 30 is quite a bit closer to the population mean than is the mean of the sample means from samples of size 5. The Central Limit Theorem tells us that the standard deviation for the means of samples of size 5 should be the population standard deviation divided by the square root of five. Hence, $899.7/\sqrt{5} = 402.4$. The experimental value of 365.6 is not exactly right, but close. For samples of size 30, the standard deviation of the means should be the population standard

Descriptive Statistics: DJC

Variable	Total Count	Mean	StDev
DJC	518	9706.6	899.7

Descriptive Statistics: Means5

Variable	Total Count	Mean	StDev
Means5	15	9729.0	365.6

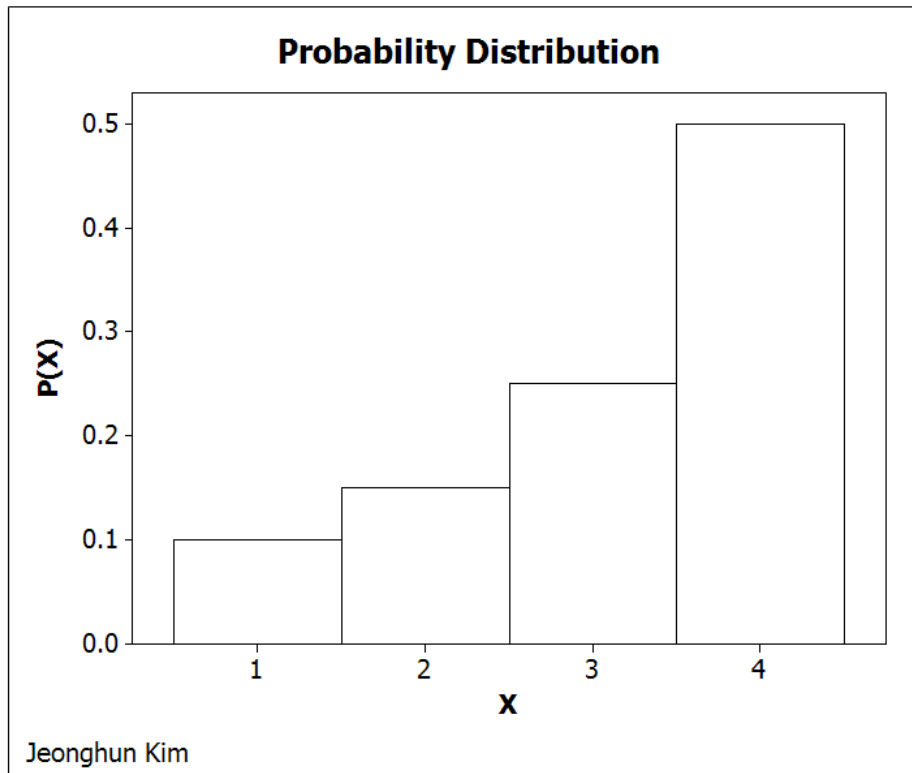
Descriptive Statistics: Means30

Variable	Total Count	Mean	StDev
Means30	15	9711.0	160.5

deviation divided by the square root of 30. Hence $899.7/\sqrt{30} = 164.3$ The experimental value of 160.5 is pretty close!

NORMALITY OF SAMPLE MEANS

Close your data window from the previous example and open the worksheet ProbDist.mtw that you saved from Lesson 8. First make a bar graph of this discrete distribution. (See pages 39 and 40 of Lesson 9.) The graph is shown below, and it certainly does not look normal.

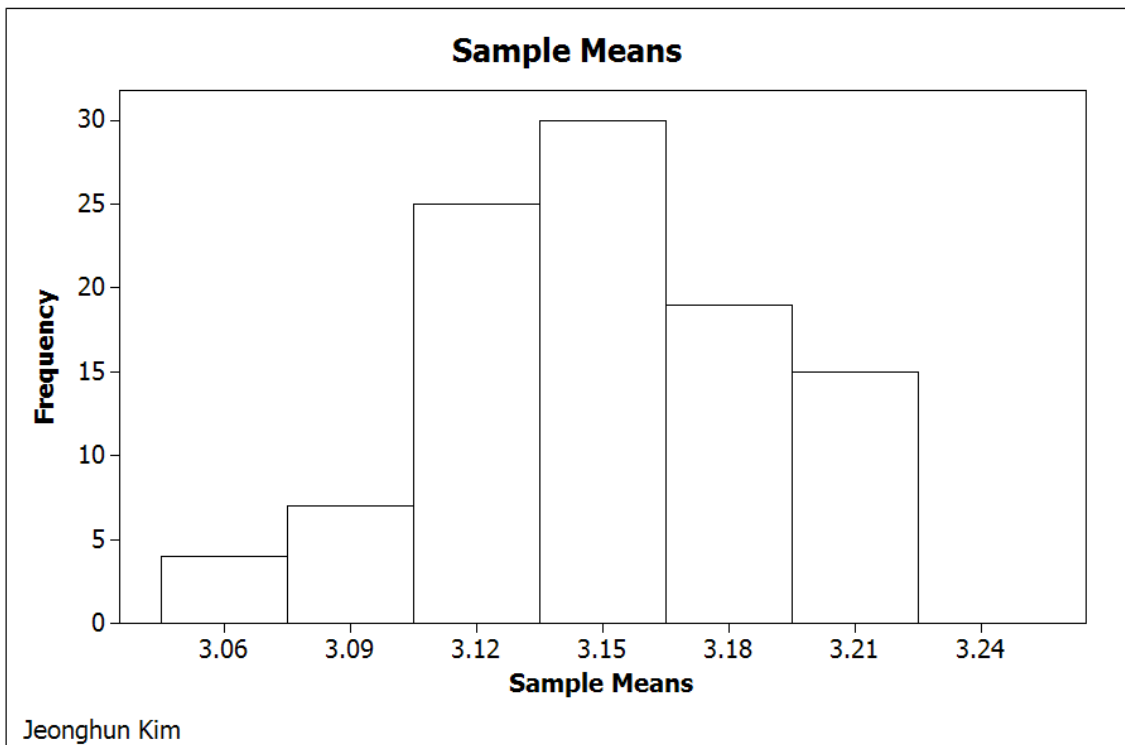


Now click on Calc > Random Data > Discrete. Type 1000 in the "Number of rows of data to generate" box and C3-C102 in the "Store in column(s):" box. Select C1 X into the "Values in:" box and P(X) into the "Probabilities in:" box. Now click "OK". You now have 100 samples of size 1000 in columns C3 through C102. Find the means of these 100 samples, (they will be stored in C103 through C202) and stack them in C203 as we did in the example above. Give C203 the label "Sample Means." Now use our basic statistics function to find the mean, and variance of the sample means. They are shown in the figure below, but your may not match exactly.

Descriptive Statistics: Sample Means		
Variable	Mean	Variance
Sample Means	3.1504	0.00131

In Lesson 8 we computed the mean and variance of this distribution to be 3.15 and 1.0275 respectively, so we should expect the mean and variance for the sample means to be 3.15 and $1.0275/1000 = 0.0010275$ respectively. We can see that the experimental values for the mean and variance are quite close to what we would expect.

Finally, we will draw a histogram of the sample means. We will use 7 bins but leave the labels in the middle rather than at the cut points.



We can see that from a distribution that was skewed left we have obtained a probability distribution of sample means that is approximately normal by repeated sampling. The Central Limit Theorem tells us that the distribution of the means will get closer to normal the larger the sample size.

MINITAB ASSIGNMENT 10

See instructions on page 8.

1. Use worksheet EISS8.MTW that you created in Minitab Assignment 1 and EISS40.MTW that you created in Minitab Assignment 5 to recreate an experiment like the one in this lesson. Compare the mean of the means for samples of size 8 and size 40 to the population mean. Compare the standard deviation of the means for samples of size 8 and size 40 to the standard deviation that is predicted by the Central Limit Theorem. Type your observations in the session window.
2. Consider Problem 35 on page 203. You should have this distribution saved as Quiz.MTW from Lesson 8. Do each of the following. Note that all of your answers, except part (a) and the graphs, will appear in the session window. In some cases you will need to type the answers into the session window.
 - (a) Create 100 samples of size 1000, find the means of the samples, and stack the sample means in a column labeled "Sample Means."
 - (b) Find the mean and variance of Sample Means.
 - (c) How do your answers to part (b) compare to the expected values? (You would have computed the mean and variance for this distribution as part of Minitab Assignment 8.)
 - (d) Create a bar graph of this distribution. Be sure to set the space between bars to zero as was done in the example above so that it looks like a probability distribution.
 - (e) Create a histogram with 7 classes of Sample Means.
 - (f) Does your graph from part (e) look closer to being normal than the one from part (d)?

[Return to Cover Page](#)