Minitab 17 Exercise Module 3 – Normal & Lognormal Distribution
(Rev 12/15)

In this session you will review normal and lognormal distribution concepts and learn for these in Minitab, to

- Determine cumulative probabilities,
- Determine inverse cumulative probabilities
- Generate random data and sort data
- Calculate descriptive statistics

Step 1. The problem set

Complete the following problems prior to starting up Minitab. Include a sketch of the normal curve with the area of interest identified. You will be checking your answers using Minitab.

A firm that manufactures grape juice has a machine that automatically fills bottles. The mean of the process is assumed to be the machine's setting. The process variation (standard deviation) is 1.2 oz. (Assume that the process has a normal distribution.)

1. Customers get unhappy if the actual level is less than 36 oz but do not mind if it is greater than 36 oz. If you set the machine at 37 oz. what % of the time would the bottle contain less than 36 oz.?

2. a. The bottle will actually hold 40 oz. If you set the machine to 38, what percent of the time will the bottles overflow?

   b. If 10 bottles from this process are filled, what is the probability that at least one will have overflowed? (Basic probability problem. Refer back to prior lectures as needed.)

3. With the machine set at 38 oz., how big would the bottle have to be not to overflow 99.95% of the time?

Step 2. Minitab Cumulative probability

Open Minitab to a new project.

Problem 1) You were asked to find the % of time the bottle would contain less than 36oz, given a mean of 37 and a standard deviation of 1.2.

Choose: **Calc>Probability Distributions>Normal**
Click **Cumulative probability**
Mean: **37**
Standard Deviation: **1.2**
Click **Input constant: 36**
Click **OK**.

The response appears in the Session box. Check your answer. If you did not get it correct recalculate it and/or double check your Minitab inputs.

Problem 2a)
Follow the same process as problem 1) changing the inputs accordingly.
(To get back to the same dialog box that you previously used, you can click  on the menu.

Remember then to subtract the session answer from 1 to compare to your answer.

Step 3. Inverse Cumulative probability

Problem 3). This time you want to know the constant, x, that will give you the cumulative probability of 99.95% when the mean is 38 and standard deviation is 1.2.

Choose: **Calc>Probability Distributions>Normal**
Click **Inverse cumulative probability**
Mean: **38**
Standard Deviation: **1.2**
Click **Input constant: .9995**
Click **OK**.

The response appears in the Session box. Check your answer.

Step 4 Review of normal concepts

In the above problems, we assumed that the mean of the process was the actual machine setting, and the standard deviation was 1.2.

In all likelihood, we would take a sample to estimate the mean and standard deviation.

Minitab will allow us to generate a random sample from given distribution.

Choose: **Calc>Random Data>Normal**
Number of rows to generate: **100**
Store in column: **C1**
Mean: **37**
Standard Deviation: **1.2**
Click **OK**.

Now let's run some statistics on the data.

Choose: **Stat>Basic Statistics>Display Descriptive Statistics**

Variable: **C1**
Click **Statistics**
Check: **Mean, Standard deviation, Minimum, Maximum, Range**
Click **OK**.
Click **Graphs**
Check: Histogram **of data with normal curve**.
Click **OK**. Click **OK**.

In the Session window, the checked statistics will appear.

On the Histogram, the mean and standard deviation appear along with the associated curve

Enter your sample statistics here: Mean = $\bar{x}$ = _____
Standard Deviation = s = _____

If you did not know the true population parameters, μ and σ, you would likely estimate it with the sample statistics $\bar{x}$ and s. When we estimate population parameters, typically the ^ is put over the symbols.

Recall that natural process limits are considered to be $\mu \pm 3\sigma$. Using your sample statistics, what would you estimate the natural process limits to be?

Natural process limits: $\hat{\mu} - 3\hat{\sigma}$ _____
$\hat{\mu} + 3\hat{\sigma}$ _____

Recall that 99.73 % of observations will fall within $\pm 3\sigma$ and approximately 95% within $\pm 2\sigma$. Now we will sort the data to see how many actually fell outside the limits.

Choose: **Data>Sort**
Columns to sort by: C1 *under Column and leave Increasing under Order*
Columns to sort: All Columns *If there were more than one column of data we might need a different choice.*
Storage location for current: **In the original columns** *use dropdown to select*
Click **OK**.

Look at the sorted data in the worksheet.

How many values fell outside the natural process limits? _____

What percent fell out of the natural process limits? _____

How does this compare with what you expected?

What are the $\pm 2\sigma$ boundaries? $\hat{\mu} - 2\hat{\sigma}$ _____
$\hat{\mu} + 2\hat{\sigma}$ _____

How many values fell below? _____ above? _____

What was the total number out of the $\pm 2\sigma$ boundaries? _____

What was the percent out of the $\pm 2\sigma$ boundaries? _____

How does this compare with what you expected?

Step 5. Lognormal Problem – transforming data

A study is performed on the latency period of chicken pox. The latency period is the time between exposure to another person with chicken pox and when symptoms first become apparent. The data for the study is found in ChickenPox.mtw. Initial analysis of the data needs to be performed to determine the distribution to use for analysis of the data.

Histogram of the data to view the general shape to get an indication of what type of distribution to use for analysis.

Generate a histogram of the data.

Does it look like it might be lognormal? _____

For further review, you decide to take the natural log of the observations and create a histogram those values to see if they look like a normal distribution.

Calc/Calculator

Store result in : C2
Use drop down to select: Natural Log (log base e)
Then double click on C1 and it will fill in the formula for you.
Click OK.

You are hoping it is normal data so you decide to generate a Histogram With Fit for the transformed data

Graph/Histogram/With fit

Graph variable: C2
Data view/Distribution –
Check Fit distribution
Use dropdown box to select normal (should be the default)
Click Ok
Click OK

Does the histogram look reasonably normal? _____

What is the mean of the distribution? _____ Standard deviation? _____

Use this mean and standard deviation for your parameters of the lognormal function in the next step.

Step 6. Work Lognormal problem by hand.

A study is performed on the latency period of chicken pox. Based on the study the latency period is assumed to follow a lognormal distribution with parameters μ = _____ and σ = _____.

What is the probability that a person exposed to the chicken pox virus will begin to show symptoms of chicken pox after 15 days?

What is the expected value of the latency period for chicken pox?

Step 7. Lognormal Calculations in Minitab.

We now want to check your calculations in Minitab.

What is the probability that a person exposed to the chicken pox virus will begin to show symptoms of chicken pox after 15 days?

Calc/Probability Distribution/Lognormal Distribution
Check cumulative probability
Location: enter mean
Scale: enter standard deviation
Threshold: leave at 0
Check Input constant: enter 15
OK

The answer will appear in the session window. The problem says "after fifteen days", so you need to subtract from 1.

$P(X>15) = 1-P(X<15) =$ _____

Check your work and revise your previous calculations if necessary.

For the expected value (using a calculator) the answer is 4.6876. (I have not found a way to do this in Minitab). Check your work and revise your previous calculations if necessary.

Step 8. Generate one more graph.

Generate a Histogram with Fitted Line for the original data. Be sure to change the type to lognormal under Data Views.

What are the location and scale parameters on the graph? Loc _____ Scale _____

How do these compare with the mean and standard deviation of the associated normal curve?