Name ___________________________ Date _________________ Class __________________

WISE Sampling Distribution of the Mean Tutorial

Exercise 1: How accurate is a sample mean?

Overview

A friend of yours developed a scale to measure Life Satisfaction. For the population of American adults the scale has a normal distribution with a mean of 500 and a standard deviation of 100. She plans to measure Life Satisfaction for various groups of people, and she is very interested in knowing how accurate her sample means are likely to be depending upon the size of her samples. She will come to you for advice. This exercise is designed to help you prepare for her questions.

At this point you may choose to either review *z*-scores and the normal distribution or skip this review and continue on below.

- Because the Life Satisfaction Scale is normally distributed in the population, *z* scores will also be normally distributed. This is extremely important to the probability calculations we will be making throughout this tutorial. If the distribution of interest is far from being normal, these calculations would not be valid!
- According to the Central Limit Theorem (CLT), although the population distribution may be non-normal, its sampling distribution of the mean approaches normal with increasing sample size. To explore more in detail the effects of the non-normal population distribution and sample size on the sampling distribution of the mean, complete the CLT tutorial.

Let us examine drawing a single Life Satisfaction score from the population.

Q1. What is the probability that a randomly selected American adult has a Life Satisfaction score within 30 points of the population mean (i.e., in the range 470 and 530)?

Use Figure 1 to make an approximation. A good way to make an approximation here is to estimate what percent of the total area under the curve falls between 470 and 530. Remember, the total area is 1.0 (or 100%).

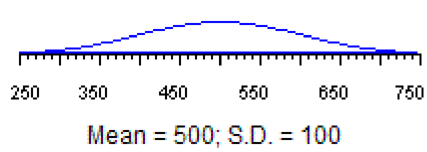

Figure 1. Population distribution of the Life Satisfaction Scale

Probability of a score from 470 to 530 is about ______. (Make your best estimate.)

Recall that the *z*-score reflects how many standard deviations a score differs from the mean. You may use a table of probabilities for the standardized normal distribution or use the *p*-*z* converter to convert from a *z*-score to probability. For examples of how to use this calculator, consult the review of *z*-scores.

Now, before going on, solve for the exact answer ______.

The answer is about 24%. The exact answer and detailed calculations are in the answer section.

Sampling distribution of the mean: *N*=100

You may now briefly review the Central Limit Theorem or skip this review and continue on below.

Our researcher friend plans to draw a sample of 100 people to take the Life Satisfaction scale. To simulate this in the applet, select the following options: **Normal, *N*=100, Show sample data** (no other 'show' options for now). If you have questions about any of the options, please consult the instructions for using the applet.

Click on **Draw a sample**, and enter the value displayed for **Last mean =** in the first space below. Then click on **Draw a sample** nine more times and record each mean below.

How many of your 10 sample means fell outside of the range 470 to 530?

Click on **Show obtained means** to see a display of all ten means.

Notice how tightly clustered the obtained means are compared to the individual scores.

Click **Draw 100 samples** to see the means for 100 different samples, each of size *N*=100.

Click **Show sampling distribution of the mean** to see how closely the observed sample means match the actual distribution of all possible means for samples of size *N*=100.

Now in terms of their means and variability, describe in the box below how this sampling distribution of the mean for *N*=100 compares to the distribution of scores in the population (click **Show population**).

Q2. What is the probability that a randomly selected sample of *N*=100 American adults has a mean Life Satisfaction score within 30 points of the population mean (i.e., between 470 and 530)?

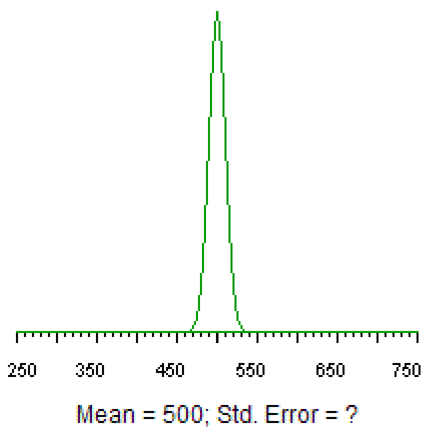

Figure 2. Sampling distribution of the mean for the Life Satisfaction Scale, *N*=100.

First, make an estimate of the answer using Figure 2.

You can also estimate the answer by counting the number of sample means out of 100 that fall within the range 470 to 530. It may be easier to see if you turn off **Show sampling distribution of the mean**. You don't need to get an exact count, but we see that when we draw a sample with *N*=100, it is unlikely that the sample mean is in error by more than 30 points as an estimate of the mean Life Satisfaction for the population (i.e., it is unlikely to find a sample mean that deviates as much as 30 points from the mean). Were any of your 100 sample means 30 or more points away from the population mean?

Now solve for the exact answer. You may use a table of probabilities for the standardized normal distribution or use the *p*-*z* converter to convert from a *z*-score to probability.

The answer is over 99%. The exact answer and detailed calculations are in the answer section.

Sampling distribution of the mean: *N*=25

Our researcher friend wishes to know how accurate the sample mean is likely to be if she samples 25 people. To simulate this, in the applet select the following options: **Normal, *N*=25, Show sample data** (no other 'show' options).

Click on **Draw a sample**, and enter the value shown for **Last mean** in the first space below. Then click on **Draw a sample** nine more times and record each mean below.

How many of your 10 sample means fell outside of the range 470 to 530?

Click on **Show obtained means** to see all ten means displayed. Notice how tightly clustered the obtained means are compared to the individual scores.

Click **Draw 100 samples** to see the means for 100 different samples, each of size *N*=25.

Click **Show sampling distribution of the mean** to see how closely the distribution of 100 observed sample means matches the actual distribution of possible means of size *N*=25.

In the box below describe how this sampling distribution of the mean (for *N*=25) compares to the sampling distribution of the mean for *N*=100. Be sure to consider the variability of sample means (i.e., the standard error of the sampling distribution of the means). [Hint: Select **Show sampling distribution of the mean** and compare the sampling distributions for *N*=100 and *N*=25.]

Q3. What is the probability that a randomly selected sample of *N*=25 American adults has a mean Life Satisfaction score within 30 points of the population mean?

First, estimate the answer by examining your ten sample means, the displays of 100 sample means with *N*=25 for each mean, and the sampling distribution of the mean. What is your rough estimate based on these observations?

Now solve for the exact answer. You may use a table of probabilities for the standardized normal distribution or use the *p*-*z* converter to convert from a *z*-score to probability.

The exact answer is about 87%. A detailed solution is in the answer section, but try it on your own before consulting the solution.

Sampling distribution of the mean: *N*=5

Your researcher friend has considered using sample sizes of only five people. She will ask you to explain the advantages and disadvantages of this plan. To begin, select the following options: **Normal, *N*=5, Show sample data** (no other 'show' options).

Click on **Draw a sample**, and record the value shown for **Last mean** in the first space below. Then click on **Draw a sample** nine more times and record each mean below.

How many of your 10 sample means fell outside of the range 470 to 530?

Click on **Show obtained means** to see all ten means displayed**.** Notice how tightly clustered the obtained means are compared to the individual scores, and to the distributions of means when *N*=100.

Click **Draw 100 samples** to see the means for 100 different samples, each of size *N*=5.

Click **Show sampling distribution of the mean** to see how closely the observed sample means match the actual distribution of possible means of size *N*=5.

In the box below describe how this sampling distribution of the mean (for *N*=5) compares to the sampling distribution of the mean for *N*=100.

Q4. What is the probability that a randomly selected sample of *N*=5 American adults has a mean Life Satisfaction score within 30 points of the population mean?

You can estimate the answer by examining your ten sample means and the displays of 100 sample means with *N*=5 for each mean. What is your estimate based on these observations?

The approximate answer is about 50%. What is the exact answer?

You may use a table of probabilities for the standardized normal distribution or use the *p*-*z* converter to convert from a *z*-score to probability.

A detailed solution is in the answer section, but try it on your own before consulting the answer.

In Conclusion

In responding to the concluding exercise questions below, you may refer to your responses on the previous pages regarding the approximate and exact probabilities that sample means fall within 30 points of the population mean and the number of obtained sample means that fell out of this range depending upon the sample size.

Q5. Complete the following table and then comment on the relationship between sample size and the expected accuracy of a sample mean:

Sample size	Probability that a sample mean differs from the population mean by more than 30 points
100	0.2%*
25	%
5	%

***In Q2 we determined that there is a 99.8% chance that the mean of a sample of *N*=100 will be within 30 points of the population mean. Thus, the probability that a mean for a sample with *N*=100 will differ from the population mean by more than 30 points is 100% - 99.8%, or 0.2%. You can use your findings in Q3 and Q4 to calculate the values for samples of *N*=25 and *N*=5.**

Your researcher friend says "we know that for any population, the best estimate of the mean is the sample mean -- therefore, it shouldn't matter what size sample I use, right? Since that is the case, I'll use a sample of *N*=5 as this will save a good deal of time and money." What do you tell your friend? In your answer in the box below, include information from the questions you have just completed.