60080079 Introduction to Statistical Methods
Semester 2 2023-2024
Handout 4

The Law of Large Numbers

Each column of the SPSS file **LLN.sav** contains 5,000 sample means from a particular distribution. Each column represents one particular sample size, namely, Columns 1 through 6 represent n = 10, 20, 50, 100, 1,000 and 10,000.

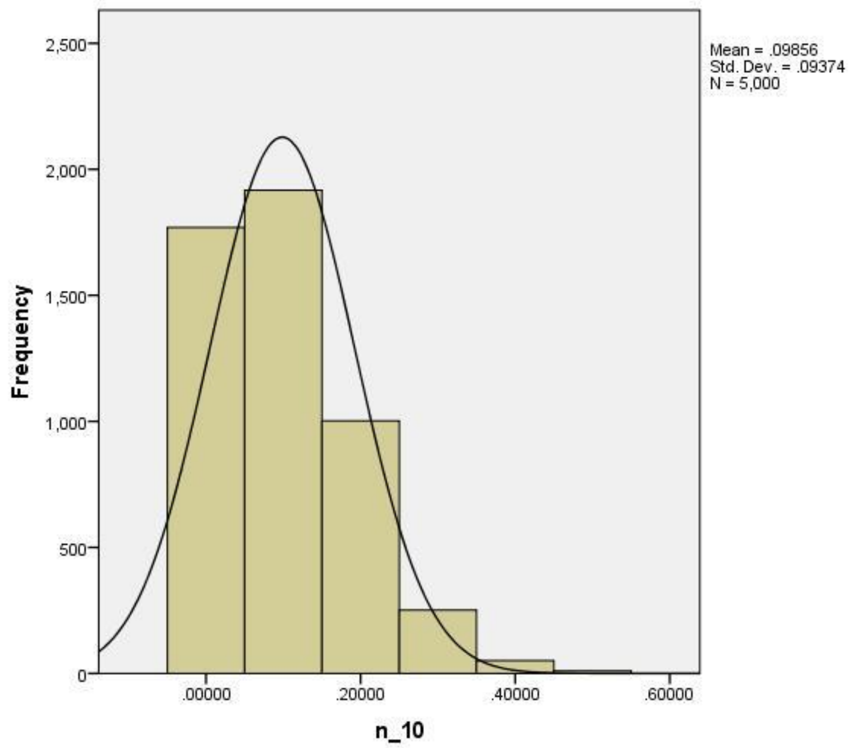
1. Obtain the histogram of each of the six variables, as in, do

   Graphs → Legacy Dialogs → Histogram

   Each histogram will give us an approximate distribution of the sample means (or sampling distribution of the mean) based on a particular sample size. (The approximation would be better if we use more observations, say, 100,000 instead of 5,000.).

   To better understand the shape of the sampling distribution of the mean, check the **Display normal curve** option just below the **Variable** box before hitting **OK**.

   For example, you should get the following histogram for **n_10**:



2. Based on the six histograms, complete the following statement.

As the sample size gets larger, the sampling distribution of the mean becomes **more** normal?

3. Use **Analyze → Descriptive Statistics → Explore** to fill in the missing values in the table below.

	Variable (Sample Size)					
Statistic	n_10	n_20	n_50	n_100	n_1K	n_10K
Mean	0.0986	0.0989	0.0995	0.1008	0.0998	0.1000
Minimum	0.0000	**0.0000**	0.0000	0.0100	0.0610	0.0872
Maximum	**0.5000**	0.3500	**0.3000**	0.2200	**0.1360**	0.1104
Range	0.5000	0.3500	0.3000	**0.2100**	0.0750	0.0232

Note: All the samples were drawn from a population that has a mean of **0.10**.

4. Based on the completed table above, complete the following statements:

Regardless of the sample size, the mean of the sampling distribution (of the mean) is approximately equal to **0.10**.

As the sample size gets larger, the smallest of the means gets **larger**, whereas the largest of the means gets **smaller**.

As a consequence, the range gets **smaller** as the sample size gets larger.

This indicates that for n = 10, the sample mean can be off from the true mean by as much as **0.2500**.?

In contrast, when n = 10,000, the sample mean can never be off from the true mean by more than **0.1000 – 0.0872 = 0. 0128**.?

Remember that your answers above are based on a finite sample of 5,000 observations. Your answers can change when you change the number of samples, or when you obtain a different sample of 5,000 observations.

In general, this exercise shows us that when the sample mean is computed using larger and larger sample sizes, we are guaranteed that it would be closer the true mean.