Central Limit Theorem workshop using TLCshiny

The TLCshiny is a web application created by using the shiny R package. This app can be used to explore the effect of the sample size n on the distribution of the statistics $\bar{X}$.

The app considers five populations gamma (default), normal, uniform, beta, and a mix of two normals. The user selects a population (source distribution), and the app will take 1000 samples of size n, then the app creates a histogram for the sample means $\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_{1000}$, finally, the app creates a QQplot to explore the distribution of the sample means visually.

The following figure illustrates the first visualization of the app. The gray box on the left side corresponds to the inputs (population, parameters, and n). The right side of the app contains the results: histogram, limit distribution, QQplot, and p-value of the Shapiro-Wilk test.

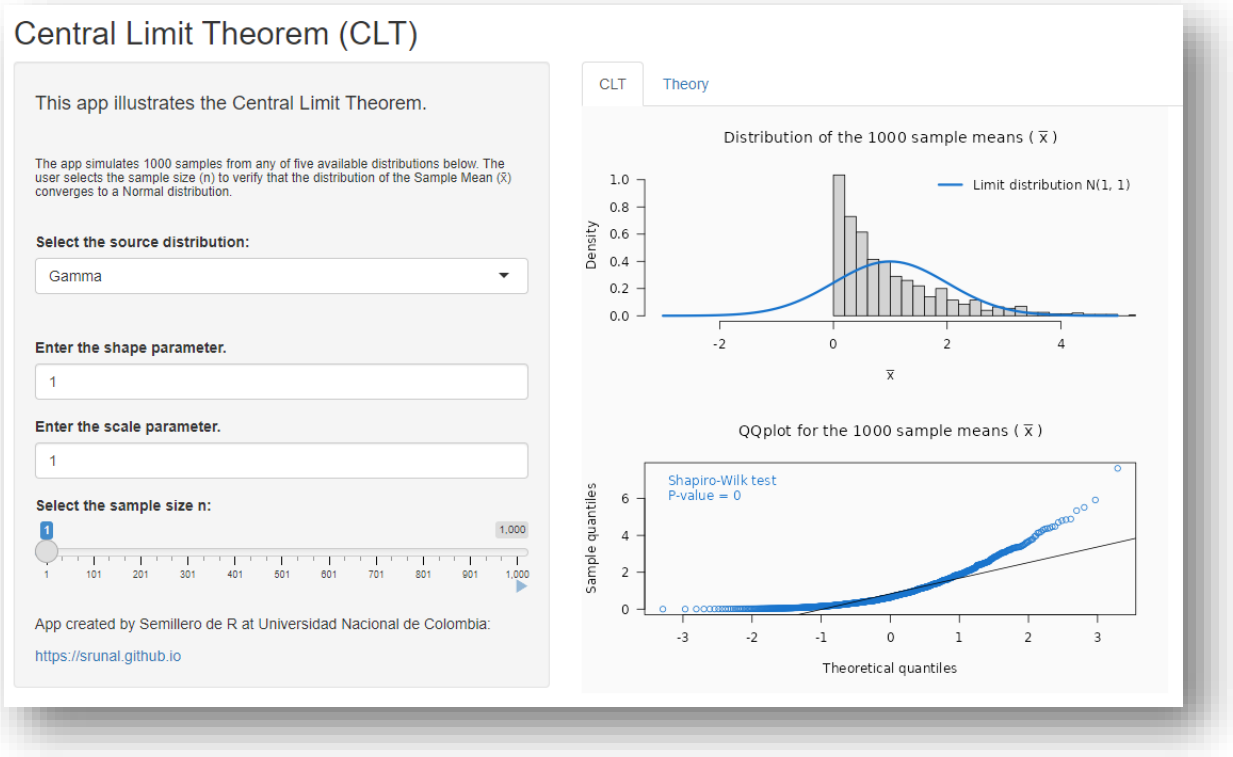


Figure 1. Screenshot of the TLCshiny with the default values.

Please follow the steps of this workshop to interact with the TLCshiny and to ensure an understanding of the concepts.

1. Visit the URL http://gauss.medellin.unal.edu.co:3838/fhernanb/tlc/ or scan the QR to open the TLCshiny.

2. Using the default values of the app, Gamma(1, 1), complete the gaps on the second and third columns of the next table to explore the effect of n on the p-value and decision. The table contains some values as a hint.

n	P-value of the Shapiro-Wilk test	Using $\alpha = 0.05$, could be rejected or not rejected H_0: the sample of $\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_{1000}$ follow from a Normal distribution?
1	0	
5		Reject
20		
45	3×10^{-4}	
50		
53		No reject
60		
80	0.2956	
100		
700		No reject
900		

Table 1. Effect of n when the samples are taken from Gamma(1, 1).

What can you conclude from the last table? _____

_____

3. Every time that sample size is $n = 1$, the histogram of the $\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_{1000}$ corresponds to the approximate density curve of the population. Complete the following figure with a sketch of the histogram fixing $n = 1$ for each population. Use the default values for the parameter distribution.

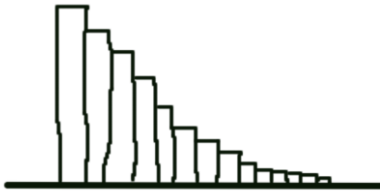
Normal(10, 3)	Uniform(0, 1)	Gamma(1, 1)
		
Beta(1.5, 0.9)	Mix of two normal(-4,6,1,1.3, 0.3)	

Table 1. Screenshot of the TLCshiny with the default values.

4. The current objective is to explore the minimum sample size n when the Shapiro-Wilk test does not reject the null hypothesis of normality for the first time. Select the correct distribution and parameters for each case, then click the blue triangular button to activate the animation for sample size n. The

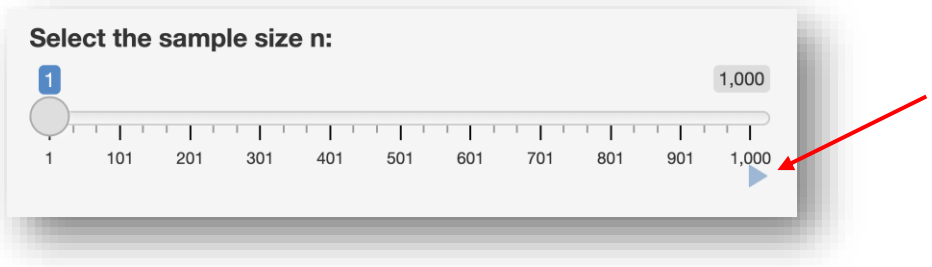


Figure 2. Screenshot of the slider input for sample size. Right bottom, there is a blue triangular button to activate the animation.

Initiate the animation in $n = 1$, and press the blue triangular button Be aware the first time the p-value exceeds the value 0.05 to stop the animation. Next, complete the gaps on the table below with the results of the minimum n to not reject the null hypothesis.

Source distribution	Parameters	Minimum n value to not reject the null hypothesis of normality for the first time	Limit distribution observed in the first figure of the App
Normal	$\mu = 3, \sigma = 8$		$N(3, 64)$
Normal	$\mu = -7, \sigma = 2$	1	
Uniform	$a = -5, b = 3$		$N(-1, 1.07)$
Uniform	$a = 7, b = 15$		
Gamma	$shape = 4, scale = 6$	24	
Gamma	$shape = 6, scale = 20$		
Beta	$a = 2.8, b = 0.9$		
Beta	$a = 3.5, b = 15$		
Mix of two normal	$\mu_1 = -5, \mu_2 = 4, \sigma_1 = 1, \sigma_2 = 2, \rho = 0.7$		
Mix of two normal	$\mu_1 = -6, \mu_2 = 6, \sigma_1 = 3, \sigma_2 = 3, \rho = 0.5$	5	

Table 2. Minimum n value to not reject the null hypothesis of normality for the first time.

What can you conclude from the last table? _____

_____

_____