İstatistik Araştırma Dergisi
Journal of Statistical Research

ISSN: 2791-7614

(2026), 16 (1)
Başvuru / Received :22.02.2026
Kabul / Accepted :20.04.2026

https://dergipark.org.tr/tr/pub/jsstr

ARAŞTIRMA MAKALESİ **RESEARCH ARTICLE**

Central Limit Theorem under Different Continuous Distributions: A Simulation-Based Investigation of Factors Affecting Convergence

Senem BAYBAŞ
Gazi University, Department of Statistics / Research Assistant
senembaybas@gazi.edu.tr
Orcid No: 0009-0003-2249-7572

Abstract

The Central Limit Theorem (CLT) is one of the cornerstones of statistics and probability theory. In this study, the practical validity of the CLT was examined through simulations conducted on different continuous distributions such as Beta, Exponential, Gamma, Log-Normal, and Weibull. While the literature generally addresses the convergence of the sample mean to normality in broad terms, in this study, the rate of convergence and the effect of sample size were analyzed using different metrics, including variance, skewness coefficient, maximum approximate error (MaxError), and the Berry bound based on the Berry-Esseen theorem. The simulation results provide a detailed perspective on the reliability and limitations of the CLT. The history of the CLT dates back to early theoretical studies and, over time, it has been recognized as a tool that allows the distribution of sample means to approach the normal curve even when the underlying population is not normal. This property has made the CLT a fundamental concept in various fields, including political science, computer science, psychology, medical research, and engineering. Despite its widespread use, the practical outcomes of the CLT are often not sufficiently understood in terms of the effects of factors such as sample size and skewness on convergence. In this study, simulation methods were applied to different continuous distributions. Consequently, in addition to classical measures of mean and variance, additional criteria were considered. Thus, the rate at which distributions approach normality and the conditions under which the CLT can be reliably applied were evaluated more comprehensively.

Keywords: Central Limit Theorem, Normal Distribution, Skewness, Simulation, Berry-Essen Theorem

---

Corresponding Author / Sorumlu Yazar: Senem BAYBAŞ, Gazi University, Department of Statistics.

Citation / Atıf: BAYBAŞ S., (2026). Central Limit Theorem under Different Continuous Distributions: A Simulation-Based Investigation of Factors Affecting Convergence. *İstatistik Araştırma Dergisi*, 16 (1), 12-23.

Citation / Atıf: BAYBAŞ S., (2026). Central Limit Theorem under Different Continuous Distributions: A Simulation-Based Investigation of Factors Affecting Convergence. *İstatistik Araştırma Dergisi*, 16 (1), 12-23.

Merkezi Limit Teoremi'nin Farklı Sürekli Dağılımlar Altında İncelenmesi: Yakınsamayı Etkileyen Faktörlerin Simülasyon Temelli Bir Araştırması

Özet

Merkezi Limit Teoremi (MLT), istatistik ve olasılık teorisinin temel taşlarından biridir. Bu çalışmada, MLT'nin pratik geçerliliği; Beta, Üstel, Gamma, Log-Normal ve Weibull gibi farklı sürekli dağılımlar üzerinde gerçekleştirilen simülasyonlar aracılığıyla incelenmiştir. Literatürde genellikle örneklem ortalamasının normalliğe yakınsaması genel çerçevede ele alınırken, bu çalışmada yakınsama hızı ve örneklem büyüklüğünün etkisi; varyans, çarpıklık katsayısı, maksimum yaklaşık hata (MaxError) ve Berry–Esseen teoremine dayalı Berry sınırı gibi farklı ölçütler kullanılarak analiz edilmiştir. Simülasyon sonuçları, MLT'nin güvenilirliği ve sınırlılıkları hakkında ayrıntılı bir bakış açısı sunmaktadır. MLT'nin tarihi erken dönem kuramsal çalışmalara dayanmakta olup, zamanla anakütle dağılımı normal olmasa bile örneklem ortalamalarının dağılımının normal eğriye yaklaştığını gösteren güçlü bir araç olarak kabul edilmiştir. Bu özelliği sayesinde MLT; siyaset bilimi, bilgisayar bilimi, psikoloji, tıp araştırmaları ve mühendislik gibi pek çok alanda temel bir kavram haline gelmiştir. Bununla birlikte, MLT'nin yaygın kullanımına rağmen, örneklem büyüklüğü ve çarpıklık gibi faktörlerin yakınsama üzerindeki etkileri uygulamada çoğu zaman yeterince anlaşılmamaktadır. Bu çalışmada farklı sürekli dağılımlar üzerinde simülasyon yöntemleri uygulanmıştır. Böylece klasik ortalama ve varyans ölçütlerine ek olarak ilave kriterler de dikkate alınmıştır. Sonuç olarak, dağılımların normalliğe hangi hızda yaklaştığı ve MLT'nin hangi koşullar altında güvenle uygulanabileceği daha kapsamlı bir şekilde değerlendirilmiştir.

Anahtar sözcükler: Merkezi Limit Teoremi, Normal Dağılım, Çarpıklık, Simülasyon, Berry–Esseen Teoremi

1. Introduction

The Central Limit Theorem (CLT) is the most fundamental theory in modern statistics (Kwak & Kim, 2017). The history of CLT begins with Pierre-Simon Laplace, who had no concrete theorem and mostly used his approach to CLT as a tool to solve other mathematical problems. Robert Leslie Ellis, Siméon Denis Poisson, Peter Gustav Lejeune Dirichlet, and Augustin Louis Cauchy discussed Laplace's work (Plenar, 2019).

CLT shows that a distribution of means of many samples (rather than individual scores) approximates a normal curve even if the underlying population is not normally distributed. A distribution of means is a distribution consisting of many means calculated from all possible samples of a given size taken from the same population. According to CLT, the shape of the distribution of sample means approximates a normal curve if the distribution of the population of individual scores has a normal shape or if the size of each sample constituting the distribution is at least 30 (Nolan & Heinzen, 2011).

CLT can be used in many real-life fields such as political science, computer science and psychology (Winters, 2022).

Additionally, multiple parametric tests are used to assess the statistical validity of clinical studies performed by medical researchers; however, most researchers are unaware of the value of CLT despite routinely using parametric tests (Kwak & Kim, 2017). Therefore, this study is important for researchers in different branches to understand CLT.

Many studies have been conducted on CLT. Lee et al., derived the exact distribution of the sum for random samples from the uniform distribution and then compared the exact distribution with the approximate normal distribution obtained by CLT. They discussed five existing normality tests based on the difference between the target normal distribution and the empirical distribution to check the similarity between two distributions, namely the Anderson-Darling test, Kolmogorov-Smirnov test, Cramer-von Mises test, Shapiro-Wilk test and Shaprio-Francia test. For comparison purposes, normality tests were applied to simulated data (Lee et al., 2014).

Barri, simulated a population with 10000 data points according to five different distribution types: uniform, platykurtic normal, positively-skewed exponential, negatively-skewed triangular, and bimodal. All populations and sampling distributions with their means and standard deviations are shown in histograms for analysis. He validated the principles of CLT and showed that if the distribution of the population is close to normal, a smaller sample size is needed for CLT to be valid (Barri, 2019).

Zhang et al., explored the misconceptions about CLT among social science researchers and cleared up these misconceptions by explaining the definition and characteristics of CLT in a way that social science researchers can understand. They examined misconceptions about CLT among graduate students and social science researchers by explaining the prerequisites for understanding CLT, presenting a formal definition of CLT, and detailing its implications (Zhang et al., 2023).

In this study, CLT was investigated by simulation method on different continuous probability distributions such as Beta, Exponential, Gamma, Lognormal and Weibull. While the literature mostly evaluates the convergence of the sample mean to normality within a general framework, in this study, criteria such as variance, skewness coefficient, maximum error of approximation (MaxError) and Berry Bound were also taken into account. Thus, not only the validity of CLT but also the speed and sample size at which each distribution approaches normality were evaluated. With simulation, the applicability of CLT under different distributions is discussed more comprehensively. In addition, the Berry-Esseen relationship, which is usually treated theoretically, has been evaluated quantitatively through simulation in this study, and a more concrete and comparative perspective on the applicability of CLT has been presented. This approach goes beyond classical CLT interpretations and provides the opportunity to reveal the speed at which distributions converge to the normal distribution with numerical data.

2. Distributions

In this section, general information about the distributions to be used in the study is given.

2.1. Normal Distribution

The general normal distribution is derived from the standard normal distribution. It is characterized by two parameters, and 2, which, we will shortly see, are in fact the mean and variance (Burkardt, 2014). The Probability Density Function (PDF) for the Normal Distribution is defined in equation 1 (Winters, 2022).

$$f(x;\mu,\sigma^2)=\begin{cases}\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, & -\infty < x < \infty\\ 0, & \text{otherwise}\end{cases} \tag{1}$$

The normal distribution curve is symmetrical and bell-shaped. The score showing the highest frequency is in the middle of the normal distribution curve, and the scores showing the lowest frequency are at the extremes. It can be noted that the two components of normality are skewness and kurtosis. Skewness is related to the symmetric nature of distributions; if the mean of the distribution is to the right of 0 and the scores are stacked to the right, the distribution is skewed to the left (tail long on the left, negatively skewed); if it is to the left, the distribution is skewed to the right (tail long on the right, positively

skewed). Kurtosis is related to the sharpness of the distribution, and distributions can be either sharp (if the kurtosis value is above 0) or nearly flat (if the kurtosis value is below 0, the data density in the tails is high). When skewness and kurtosis are 0, the distribution is considered normal (Uysal & Kılıç, 2022). Figure 1 shows three curves: left-skewed, normal, and right-skewed, respectively (Nolan & Heinzen, 2011).

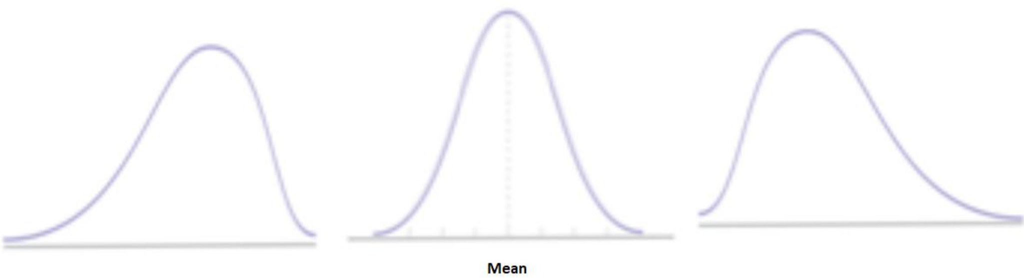

Figure 1. Left-Skewed, Normal, and Right-Skewed Distributions

2.2. Beta Distribution

The beta distribution is a family of continuous probability distributions normalized by two positive parameters, a and β, in the interval [0,1]. PDF for the Beta Distribution is defined in equation 2.

$$f(x;a,\beta) = \begin{cases} \dfrac{x^{a-1}(1-x)^{\beta-1}}{B(a,\beta)}, & a > 0, \beta > 0, 0 \leq x \leq 1 \\ 0, & otherwise \end{cases} \tag{2}$$

The $B(a,\beta)$ expression in the probability density function is defined in equation 3, using Gamma functions (Sarısoy & Gamgam, 2010).

$$B(a,\beta) = \frac{\Gamma(a)\,\Gamma(\beta)}{\Gamma(a+\beta)} \tag{3}$$

2.3. Gamma Distribution

It is a two-parameter continuous probability distribution. One of these parameters is taken as scale parameter: β, shape parameter: a. The exponential distribution is a special case of the gamma distribution, the sum of exponential variables of which is the gamma distribution. PDF for the Gamma Distribution is defined in equation 4 (Yeler, 2019).

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{x^{a}\Gamma(a)} x^{a-1} e^{-x/\beta}, & a > 0, \beta > 0, \Gamma(a) > 0, x > 0 \\ 0, & otherwise \end{cases} \tag{4}$$

2.4. Exponential Distribution

If the probability density of a variable $\boldsymbol{X}$ fits the following definition, it has an exponential distribution and is called an exponential variable. $\boldsymbol{\lambda}$ is the scale parameter. PDF for the Exponential Distribution is defined in equation 5 (Yeler, 2019).

$$f(x;\lambda) = \begin{cases} f(x;\lambda) = \lambda e^{-\lambda x}, & \lambda > 0, \mathrm{x} > 0 \\ 0, & otherwise \end{cases} \tag{5}$$

2.5. Log-Normal Distribution

A random variable X is log-normally distributed if $log(X)$ has a normal distribution. Usually, natural logarithms are used, but other bases would lead to the same family of distributions, with rescaled parameters. PDF for the Log-Normal Distribution is defined in equation 6 (Limpert et al., 2001).

$$f(x;\mu,\sigma)=\begin{cases}\dfrac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & otherwise\end{cases} \tag{6}$$

2.6. Weibull Distribution

It is a three-parameter continuous probability distribution. The parameters $\boldsymbol{a}$, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are known as the shape, scale and location parameters, respectively. PDF of the three-parameter for the Weibull Distribution is defined in equation 7.

$$f(x;\alpha,\beta,\theta)=\begin{cases}\dfrac{a}{\beta}\left(\dfrac{x-\theta}{\beta}\right)^{a-1}e^{-\left(\frac{x-\theta}{\beta}\right)^{a}}, & a>0,\beta>0,x>\theta,-\infty<\theta<\infty\\ 0, & otherwise\end{cases} \tag{7}$$

By a shift transformation, equation (7) can be reduced to the two-parameter version. PDF of the two-parameter for the Weibull Distribution is defined in equation 8 (Teimouri et al., 2013).

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{a}{\beta}\left(\dfrac{x}{\beta}\right)^{a-1}e^{-\left(\frac{x}{\beta}\right)^{a}}, & a>0,\beta>0,x>0\\ 0, & otherwise\end{cases} \tag{8}$$

2.7. Berry-Esseen Theorem

Although some distributions can be approximated by the normal distribution, they may not be useful approximations if they converge too slowly to the normal distribution. The Berry-Esseen Theorem shows how it helps to know when an estimate is good. The Berry-Esseen theorem specifies the rate of convergence to the normal distribution by limiting the maximum error of approximation between the normal and the standardized sample mean $\boldsymbol{F_x}$ distribution. According to the classical Berry-Esseen Theorem, if $\boldsymbol{X_1},\dots,\boldsymbol{X_n}$ are i.i.d. random variables each with a finite third moment, then

$$|F_x(x)-\phi(x)|\le\ a\frac{E|X1-\mu|^3}{\sigma^3\sqrt{n}} \tag{9}$$

Where $\boldsymbol{F_x}$ is the CDF of $\frac{(\bar{x}-\mu)\sqrt{n}}{\sigma}$, $\boldsymbol{\phi}(.)$ is the CDF f the standard normal distribution and $\boldsymbol{a}$ is a constant. The term $\boldsymbol{E|X_1-\mu|^3}$ in equation 9 shows that the random variable's third moment, which often characterizes the variable's skewness, plays an important role in the rate of convergence to the normal distribution (Zhang et al., 2023).

3. Materıals and Methods

3.1. Simulation Study

In this section, Beta, Exponential, Gamma, Lognormal and Weibull distributions are obtained through simulation. For each distribution, 10,000 random numbers were generated. The simulation study was carried out separately for different sample sizes such as $\boldsymbol{n=10,20,30,50,100}$ and $\boldsymbol{1000}$. For each distribution, sample mean, variance, skewness coefficient, maximum error of approximation (MaxError) and Berry-Esseen limit (Berry Bound) were calculated. In addition, the histogram graph of the distribution for each sample size was drawn and the process of convergence to normality was observed visually. Matlab program was used in the simulation study.

3.1.1. Beta Distribution

In the simulation study for the beta distribution, the parameters of the distribution were taken as $\boldsymbol{\alpha}$=5, $\boldsymbol{\beta}$=1. Accordingly, the results obtained according to different sample diameters are shown in Table 1.

Table 1. Values Calculated from Simulation Data for Sample Sizes Different from the Beta (5,1) Distribution

Distribution	Sample Size (n)	Mean	Variance	Coefficient of Skewness	Maximum Approximation Error (MaxError)	Convergence Speed (Berry - Bound)	Comment
Beta Distribution (5,1)	10	0.83336	0.00196	0.38205	0.03433	0.27459	Good convergence (CLT holds)
	20	0.83332	0.00099	0.27964	0.02017	0.19084	Good convergence (CLT holds)
	30	0.83368	0.00066	0.20933	0.02457	0.15400	Good convergence (CLT holds)
	50	0.83338	0.00040	0.16288	0.01892	0.12146	Good convergence (CLT holds)
	100	0.83338	0.00020	0.13277	0.01514	0.08604	Good convergence (CLT holds)
	1000	0.8333	0.00002	0.04548	0.01031	0.02709	Good convergence (CLT holds)

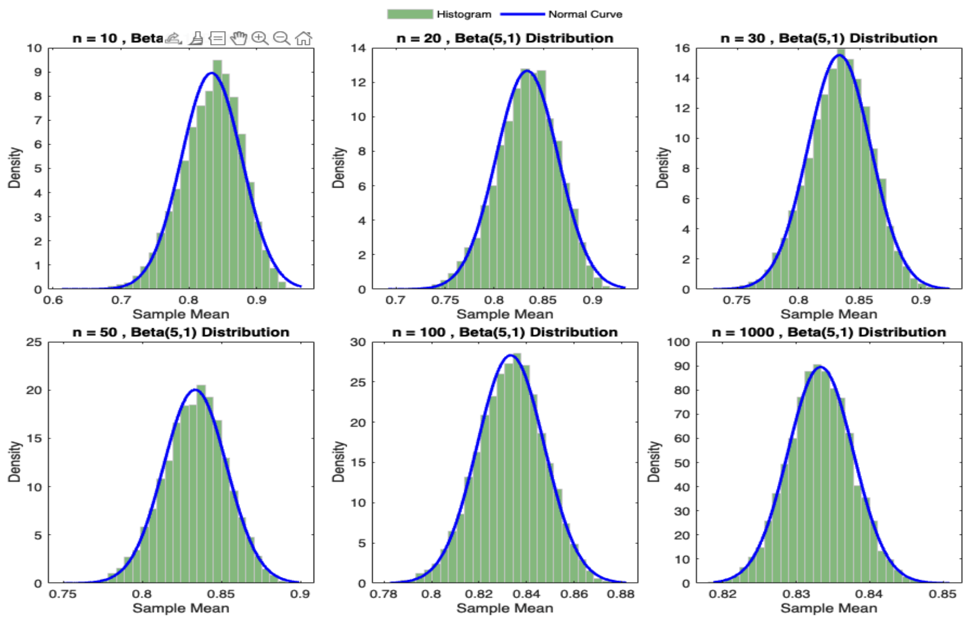

Figure 2. Histograms of Simulation Data Generated for Sample Sizes Different from the Beta (5,1) Distribution

As a result of examining Table 1 and Figure 2, it was observed that the sample mean converged to the theoretical expected value of 0.8333 and the variance became smaller as the sample size increased. Additionally, it was observed that the skewness coefficient decreased as the sample size increased and the distribution became increasingly symmetrical. It has been observed that the beta distribution has a left-skewed structure in small sample sizes with parameter values of $\boldsymbol{\alpha}$=5, $\boldsymbol{\beta}$=1. However, when the sample size is $\boldsymbol{n \geq 50}$, the skewness decreases significantly and the tendency of the distribution to converge to normality becomes stronger. It was observed that as the sample size increases, the maxerror, that is, the difference between the theoretical distribution and the sampling distribution, decreases and this difference remains within the Berry Bound limits. This situation reveals a convergence process that does not exceed the limits of CLT and yields more reliable results as the sample size increases. Additionally, it is observed that the convergence speed decreases significantly when $\boldsymbol{n \geq 50}$ and the MaxError and Berry Bound values decrease rapidly. This shows that convergence is fast when $\boldsymbol{n \geq 50}$ in the Beta distribution. This process reveals that the Beta distribution approaches the normal distribution more rapidly as $\boldsymbol{n \geq 50}$.

3.1.2. Gamma Distribution

In the simulation study for the gamma distribution, the parameters of the distribution were taken as $\boldsymbol{\alpha}$=0.5, $\boldsymbol{\beta}$=1. Accordingly, the results obtained according to different sample diameters are shown in Table 2.

Table 2. Values Calculated from Simulation Data for Different Sample Sizes from the Gamma (0.5,1) Distribution

Distribution	Sample Size (n)	Mean	Variance	Coefficient of Skewness	Maximum Approximation Error (MaxError)	Convergence Speed (Berry - Bound)	Comment
Gamma Distribution (0.5,1)	10	0.50176	0.05077	0.94668	0.05860	0.48144	Good convergence (CLT holds)
	20	0.50099	0.02527	0.66349	0.04232	0.32809	Good convergence (CLT holds)
	30	0.49821	0.01642	0.56192	0.04223	0.26284	Good convergence (CLT holds)
	50	0.49834	0.00991	0.39877	0.03918	0.20136	Good convergence (CLT holds)
	100	0.49968	0.00487	0.25723	0.02012	0.14628	Good convergence (CLT holds)
	1000	0.50000	0.00050	0.10504	0.01115	0.04642	Good convergence (CLT holds)

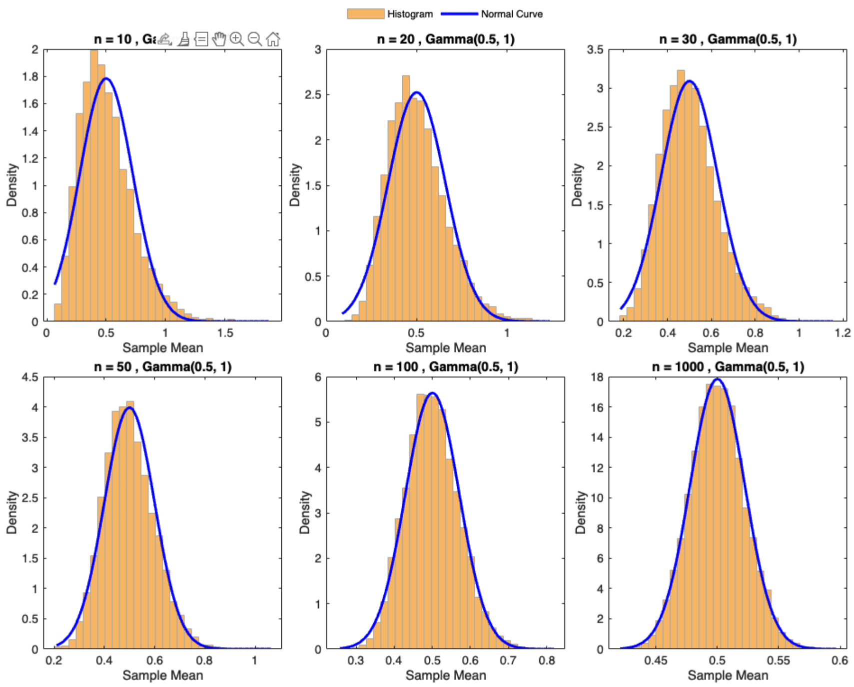

Figure 3. Histograms of Simulation Data Generated for Sample Sizes Different from the Gamma (0.5,1) Distribution

As a result of examining Table 2 and Figure 3, it was observed that the mean converged to the theoretical expected value of 0.5. It was observed that as the sample size increases, the variance becomes smaller. Additionally, it was observed that the skewness coefficient decreased as the sample size increased and the distribution became increasingly symmetrical. It has been observed that the Gamma distribution has a right-skewed structure in small sample sizes with parameter values of $\boldsymbol{\alpha} = \mathbf{0.5}, \boldsymbol{\beta} = \mathbf{1}$. However, when the sample size is $\boldsymbol{n} \geq \mathbf{50}$, the skewness decreases significantly and the tendency of the distribution to converge to normality becomes stronger. It was observed that as the sample size increases, the maxerror, that is, the difference between the theoretical distribution and the sampling distribution, decreases and this difference remains within the Berry Bound limits. This situation reveals a convergence process that does not exceed the limits of CLT and yields more reliable results as the sample size increases. Additionally, the convergence speed decreased significantly when $\boldsymbol{n} \geq \mathbf{50}$ and it was observed that MaxError and Berry Bound values decreased rapidly. This shows that convergence is fast when $\boldsymbol{n} \geq \mathbf{50}$ in the Gamma distribution. This process reveals that the Gamma distribution converges more rapidly to the normal distribution as $\boldsymbol{n} \geq \mathbf{50}$.

3.1.3. Exponential Distribution

In the simulation study for the exponential distribution, the parameter of the distribution was taken as $\boldsymbol{\lambda} = \mathbf{1}$. Accordingly, the results obtained according to different sample diameters are shown in Table 3.

Table 3. Values Calculated from Simulation Data for Sample Sizes Different from the Exponential (1) Distribution

Distribution	Sample Size (n)	Mean	Variance	Coefficient of Skewness	Maximum Approximation Error (MaxError)	Convergence Speed (Berry - Bound)	Comment
Exponential Distribution (1)	10	0.99751	0.10071	0.62584	0.04461	0.35847	Good convergence (CLT holds)
	20	1.00290	0.05035	0.41203	0.02573	0.25313	Good convergence (CLT holds)
	30	1.00150	0.03274	0.40500	0.02628	0.21089	Good convergence (CLT holds)
	50	1.00060	0.02022	0.30443	0.02350	0.16433	Good convergence (CLT holds)
	100	0.99977	0.01016	0.16421	0.01394	0.11482	Good convergence (CLT holds)
	1000	1.00000	0.00099	0.04481	0.00624	0.03641	Good convergence (CLT holds)

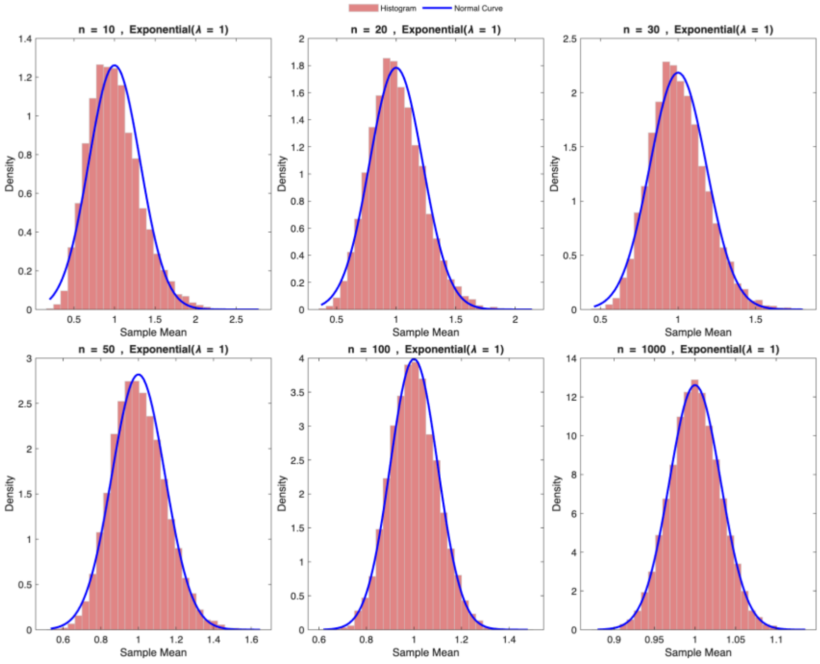

Figure 4. Histograms of Simulation Data Generated for Sample Sizes Different from the Exponential (1) Distribution

As a result of examining Table 3 and Figure 4, it was observed that the mean converged to the theoretical expected value of 1. It was observed that as the sample size increases, the variance becomes smaller. Additionally, it was observed that the skewness coefficient decreased as the sample size increased and the distribution became increasingly symmetrical. It has been observed that the exponential distribution has a right-skewed structure in small sample sizes with a parameter value of λ=1. However, when the sample size is $\boldsymbol{n} \geq \mathbf{50}$, the skewness decreases significantly and the tendency of the distribution to converge to normality becomes stronger. It was observed that as the sample size increases, the maxerror, that is, the difference between the theoretical distribution and the sampling distribution, decreases and this difference remains within the Berry Bound limits. This situation reveals a convergence process that does not exceed the limits of CLT and yields more reliable results as the sample size increases. Additionally, the convergence speed decreased significantly when $\boldsymbol{n} \geq \mathbf{50}$ and it was observed that MaxError and Berry Bound values decreased rapidly. This shows that convergence is fast when $\boldsymbol{n} \geq \mathbf{50}$ in the exponential distribution. This process reveals that the exponential distribution approaches the normal distribution more rapidly as $\boldsymbol{n} \geq \mathbf{50}$.

3.1.4. Log-Normal Distribution

In the simulation study for the Log-Normal distribution, the distribution parameters were taken as $\boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\sigma} = \mathbf{1}$. Accordingly, the results obtained according to different sample diameters are shown in Table 4.

Table 4. Values Calculated from Simulation Data for Sample Sizes Different from the Log-Normal (0,1) Distribution

Distribution	Sample Size (n)	Mean	Variance	Coefficient of Skewness	Maximum Approximation (MaxError)	Convergence Speed Error(Berry - Bound)	Comment
Log-Normal Distribution (0,1)	10	1.64830	0.45884	1.72340	0.08942	0.78349	Good convergence (CLT holds)
	20	1.64660	0.22830	1.26720	0.07431	0.58224	Good convergence (CLT holds)
	30	1.64900	0.15596	1.07160	0.06315	0.52535	Good convergence (CLT holds)
	50	1.65090	0.09162	0.78573	0.04841	0.41498	Good convergence (CLT holds)
	100	1.64660	0.04699	0.62158	0.04502	0.32010	Good convergence (CLT holds)
	1000	1.65000	0.00469	0.19574	0.01099	0.09697	Good convergence (CLT holds)

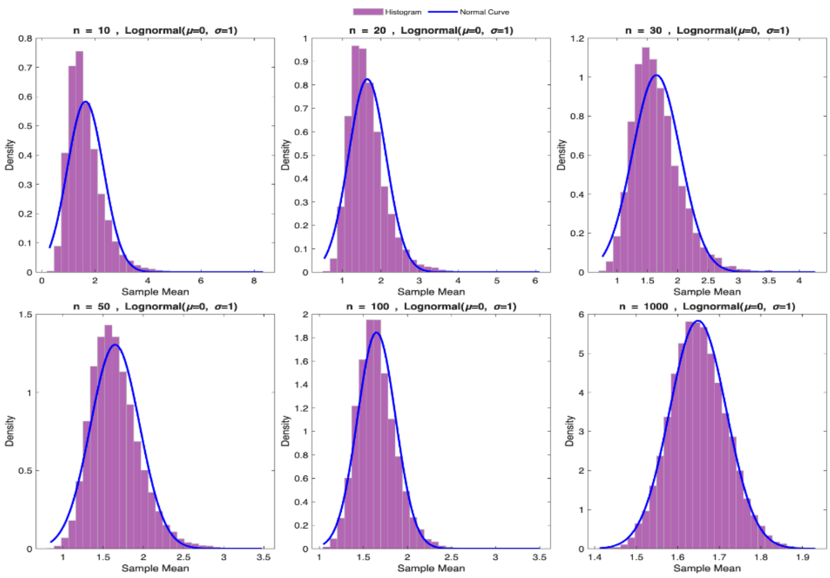

Figure 5. Histograms of Simulation Data Generated for Sample Sizes Different from the Log-Normal (0,1) Distribution

As a result of examining Table 4 and Figure 5, it was observed that the mean converged to the theoretical expected value of 1.65. It was observed that as the sample size increases, the variance becomes smaller. Additionally, it was observed that the skewness coefficient decreased as the sample size increased and the distribution became increasingly symmetrical. It has been observed that the log-normal distribution has a right-skewed structure in small sample sizes at parameter values $\boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\sigma} = \mathbf{1}$. However, when the sample size is n>100, the skewness decreases significantly and the tendency for the distribution to converge to normality becomes stronger. It was observed that as the sample size increases, the maxerror, that is, the difference between the theoretical distribution and the sampling distribution, decreases and this difference remains within the Berry Bound limits. This situation reveals a convergence process that does not exceed the limits of CLT and yields more reliable results as the sample size increases. Additionally, the convergence speed decreased significantly when $\boldsymbol{n} \geq \mathbf{100}$ and it was observed that MaxError and Berry Bound values decreased rapidly. This shows that convergence is fast when $\boldsymbol{n} \geq \mathbf{100}$ in the Log-Normal distribution. This process reveals that the Log-Normal distribution converges more rapidly to the normal distribution as $\boldsymbol{n} \geq \mathbf{100}$.

3.1.5. Weibull Distribution

In the simulation study for the Weibull distribution, the parameters of the distribution were taken as $\boldsymbol{\lambda} = \mathbf{1}, \mathbf{k} = \mathbf{0.5}$. Accordingly, the results obtained according to different sample diameters are shown in Table 5.

Table 5. Values Calculated from Simulation Data for Sample Sizes Different from the Weibull (1,0.5) Distribution

Distribution	Sample Size (n)	Mean	Variance	Coefficient of Skewness	Maximum Approximation Error (MaxError)	Convergence Speed (Berry - Bound)	Comment
Weibull Distribution (1,0.5)	10	1.99400	2.05270	2.07470	0.12231	0.97622	Good convergence (CLT holds)
	20	1.99190	0.99125	1.44870	0.09412	0.67119	Good convergence (CLT holds)
	30	1.99100	0.67221	1.25130	0.08217	0.60628	Good convergence (CLT holds)
	50	1.99970	0.40114	0.93944	0.06021	0.42629	Good convergence (CLT holds)
	100	2.00620	0.20410	0.67862	0.04035	0.30353	Good convergence (CLT holds)
	1000	2.00000	0.01971	0.21638	0.01655	0.10379	Good convergence (CLT holds)

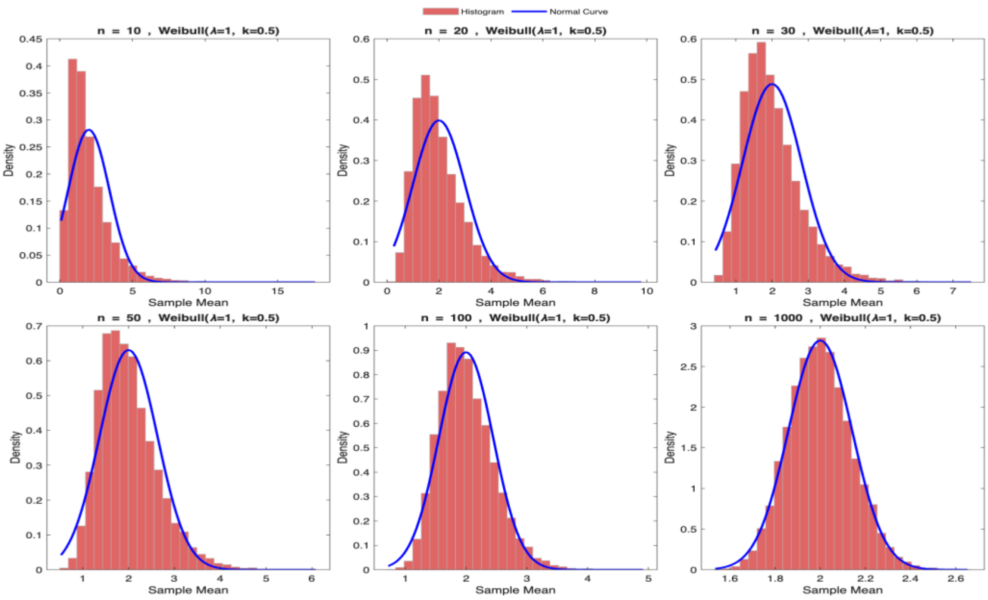

Figure 6. Histograms of Simulation Data Generated for Sample Sizes Different from the Weibull (1,0.5) Distribution

As a result of examining Table 5 and Figure 6, it was observed that the mean converged to the theoretical expected value of 2. It was observed that the variance decreases as the sample size increases. It was also observed that the coefficient of skewness decreases as the sample size increases, and the distribution becomes increasingly symmetrical. It was observed that the weibull distribution had a right-skewed structure in small sample sizes at parameter values λ=1, k=0.5. However, when the sample size is $\boldsymbol{n} \geq \mathbf{100}$, the skewness decreases significantly and the tendency for the distribution to converge to normality becomes stronger. It was observed that as the sample size increases, the maxerror, that is, the difference between the theoretical distribution and the sampling distribution, decreases and this difference remains within the Berry Bound limits. This situation reveals a convergence process that does not exceed the limits of CLT and yields more reliable results as the sample size increases.
Additionally, the convergence speed decreased significantly when $\boldsymbol{n} \geq \mathbf{100}$ and it was observed that MaxError and Berry Bound values decreased rapidly. This shows that convergence is fast when $\boldsymbol{n} \geq \mathbf{100}$ in the Weibull distribution. This process reveals that the Weibull distribution converges more rapidly to the normal distribution as $\boldsymbol{n} \geq \mathbf{100}$.

4. Conclusion

In this study, the effect of sample size on various statistical measures was investigated by simulations on different continuous distributions such as Beta, Exponential, Gamma, Log-Normal and Weibull. Furthermore, by going beyond the classical CLT determinations, we focus not only on the sample mean and variance but also on factors such as skewness of distributions, maximum error of approximation and Berry Bound. Thus, a more in-depth perspective on the convergence process was brought, and the speed at which different distributions approach normality was analyzed. The results show that as the sample size increases, the mean approaches the theoretical value, the variance and skewness coefficient decrease, and accordingly the distribution becomes more symmetrical and reliability increases. When the maximum approximation error (MaxError) and Berry Bound values are taken into account, it is seen that the convergence occurs within the Berry Bound limits in all distributions, therefore the CLT is valid.

In the highly skewed Log-Normal and Weibull distributions, convergence to the normal distribution becomes evident with a sample size of $\boldsymbol{n} \geq \mathbf{100}$, while in the less skewed Beta, Exponential and Gamma distributions, this convergence occurs earlier and faster as of $\boldsymbol{n} \geq \mathbf{50}$. This suggests that CLT should be evaluated not only with sample size but also with other factors.

Although $\boldsymbol{n} \geq \mathbf{30}$ is generally accepted as a threshold value for CLT validity in the literature, there are cases where this value may be insufficient when different criteria are taken into account. In line with these findings, the approach presented in the study reveals that when evaluating the applicability of CLT, not only classical criteria such as mean and variance should be considered, but also additional statistical indicators such as skewness coefficient, Berry Bound value and MaxError should be taken into account. These criteria allow for a more realistic and comparative analysis of the speed at which different distributions converge to normality. In addition, numerical evaluation of Berry Bound values through simulation concretely reveals the practical equivalent of a theoretical result and provides a new approach to assessing the applicability of CLT.

In conclusion, this study not only confirms that increasing sample size plays a critical role in the approximation of distributions to normality, but also makes a significant contribution to the literature by revealing the differences in the convergence speeds of different continuous distributions. This comprehensive approach makes a significant contribution to classical CLT applications and allows for a more accurate interpretation of sample size-dependent convergence processes. In this context, considering the mean, variance, skewness and Berry-Esseen bounds together ensures that more robust and reliable results are obtained in statistical inferences. In future research, expanding these analyses to more diverse distributions and under different parametric conditions can provide significant contributions to the field of theoretical and applied statistics.

Conflict of Interest

The author stated that there are no conflicts of interest regarding the publication of this article.

References

Barri MA. A simulation showing the role of central limit theorem in handling non-normal distributions. *Am J Educ Res*2019; 7(8): 591–598.

Burkardt J. *The truncated normal distribution*. Department of Scientific Computing Website, Florida State University, 2014; 1(35): 58.

Kwak SG, Kim JH. Central limit theorem: the cornerstone of modern statistics. *Korean J Anesthesiol* 2017; 70(2): 144.

Lee CM, Kim SA, Jeong JS. A view on the validity of central limit theorem: an empirical study using random samples from uniform distribution. *Commun Stat Appl Methods* 2014; 21(6): 539–559.

Limpert E, Stahel WA, Abbt M. Log-normal distributions across the sciences: keys and clues. *BioScience* 2001; 51(5): 341–352.

Nolan SA, Heinzen T. *Statistics for the Behavioral Sciences*. New York, NY, USA: Macmillan, 2011.

Plenar F. *A history of the central limit theorem*. Seminar Paper, Technical University of Vienna, 2019.

Sarısoy EE, Gamgam H. Sağa çarpık dağılım ortalamaları için bazı testlerin kullanımı ve karşılaştırmaları. *İstatistik Araştırma Dergisi* 2010; 7(2): 45–57.

Teimouri M, Hoseini SM, Nadarajah S. Comparison of estimation methods for the Weibull distribution. *Statistics* 2013; 47(1): 93–109.

Uysal İ, Kılıç A. Normal dağılım ikilemi. *Anadolu J Educ Sci Int* 2022; 12(1): 220–248.

Winters K. *A new application of the central limit theorem*. Selected Honors Theses, Southeastern University, USA, 2022, p. 158.

Yeler D. *Genelleştirilmiş doğrusal modellerde etkili gözlemlerin incelenmesi ve gamma dağılımı uygulaması*. Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Türkiye, 2019.

Zhang X, Astivia OLO, Kroc E, Zumbo BD. How to think clearly about the central limit theorem. *Psychol Methods* 2023; 28(6): 1427.