JOURNAL OF STATISTICAL MODELING & ANALYTICS (JOSMA)
(ISSN: 2180-3102)
UNIVERSITI MALAYA

Sample Size Requirements for the Central Limit Theorem for Skewed Distributions: A simulation study

Sinha Aziz[1], HM Nayem[1] & Dr. B M Golam Kibria[2*]

[1]*Graduate Student, Department of Mathematics and Statistics, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA*

[2]*Professor, Department of Mathematics and Statistics, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA*

**Corresponding author: kibriag@fiu.edu*

RESEARCH ARTICLE

Received 23 July 2025
Accepted 18 Nov 2025
Published 26 Dec 2025

Abstract

The Central Limit Theorem (CLT) plays a foundational role in statistical inference, often serving as the rationale for assuming a normal approximation of the sample mean. Yet, the pace at which this assumption becomes valid is influenced by the shape of the parent distribution, especially its skewness. This research quantifies the minimum number of observations required for the mean of samples drawn from skewed, non-normal distributions specifically Gamma, Poisson, Binomial, and Beta to achieve a satisfactory normal approximation. We implemented a Monte Carlo simulation and applied both the Shapiro-Wilk and Kolmogorov-Smirnov tests to assess the adequacy of the normal approximation. Results indicate a nonlinear association between the degree of skewness and the sample size required for acceptable normal approximation. For distributions with mild asymmetry (|skewness| < 0.5), 20 samples often suffice, whereas more heavily skewed distributions (|skewness| ≥ 2.5) may necessitate sample sizes beyond 100. These findings call into question the blanket use of the "$n \geq 30$" heuristic and suggest more tailored guidelines are necessary for accurate inference. A graphical overview summarizes these results across the examined distributional families, offering clear guidance for applied researchers working with non-normal data.

Keywords: Binomial, Beta, Gamma, Normal, Poisson, Sample Size, Skewness

1. Introduction

The Central Limit Theorem (CLT) is one of the cornerstones of probability theory and statistical inference. It posits that, for a sufficiently large sample size, the distribution of the sample mean tends to achieve a normal approximation regardless of the shape of the original population distribution, provided the population has a finite mean and variance (Casella & Berger, 2002; Rice, 2006). This foundational theorem underpins many parametric inferential methods ranging from the construction of confidence intervals to hypothesis testing. The CLT enables researchers to draw reliable conclusions from sample data even when the underlying distribution is not normal, thereby making it an indispensable tool in fields such as economics, psychology, public health, and engineering (Agresti & Finlay, 2009; Hogg et al., 2013).

However, while the CLT offers powerful theoretical assurances, the rate at which convergence to the normal approximation occurs varies significantly based on the characteristics of the parent distribution. Specifically, the convergence is influenced by the degree of skewness, kurtosis, and the presence of outliers (Lumley et al., 2002; Hogg et al., 2013; Thode, 2002). Distributions with high skewness or leptokurtic behavior often require much larger sample sizes for the sample mean to a satisfactory normal approximation compared to symmetric or light-tailed distributions (Micceri, 1989; Kim, 2013). While the CLT provides asymptotic guarantees, it does not specify the rate of convergence, nor does it prescribe the precise sample size at which the normal approximation becomes sufficiently accurate for practical applications (Blanca et al., 2013; Kim & Park, 2019). This presents a significant challenge for applied researchers who must decide whether parametric methods based on normal approximation are valid for their data.

A commonly cited rule of thumb suggests that a sample size of 30 or more is sufficient for the CLT to hold and for the sample mean to achieve an adequate normal approximation (Chang et al., 2006; Kwak & Kim, 2017; Field, 2013). This heuristic is often referenced in textbooks and classroom instruction as a guideline for employing parametric tests. However, growing empirical evidence suggests that this threshold is not universally applicable. For example, Kim (2013) showed that the degree of skewness and kurtosis directly impacts the validity of the normal approximation, and that sample sizes far exceeding 30 may be necessary in the presence of substantial asymmetry. Similarly, Blanca et al. (2013) examined non-normal data in psychological research and reported that commonly used sample sizes may not always suffice to justify parametric tests. These findings call into question the validity of universally applying the $n \geq 30$ rule and emphasize the need for a distribution-specific understanding of the CLT's normal approximation behavior.

Several prior simulation studies and empirical investigations have explored how quickly the CLT yields a reliable normal approximation under different conditions. Ahad et al. (2011) assessed the CLT's validity across Gamma, Exponential, and Chi-square distributions and demonstrated that larger sample sizes ($n > 100$) were needed for distributions with pronounced skewness. Similarly, Koh and Ahad (2020) focused on the Beta and Exponential distributions and found that moderate skewness levels could delay convergence even beyond $n = 100$. Their findings corroborate earlier results by Micceri (1989), who analyzed real-world datasets and concluded that data in education and psychology often deviate substantially from normality, rendering CLT-based normal approximations unreliable at conventional sample sizes. In addition, Zhang et al. (2015) studied the Poisson distribution and noted that while approximation may occur at small sample sizes under high rate parameters ($\lambda > 10$), it deteriorates significantly at lower values ($\lambda < 5$).

The Poisson and Binomial distributions are of particular practical interest due to their prevalence in count data scenarios such as epidemiological studies, defect counts in manufacturing, and response rates in surveys. Although they are discrete in nature, these distributions are frequently assumed to achieve a normal approximation in large samples. However, convergence behavior is highly parameter-dependent. Lumley et al. (2002) observed that the Poisson distribution with small rate parameters shows slower convergence, while the Binomial distribution behaves similarly when the probability of success is close to 0 or 1. The implications are clear: for discrete distributions, the sample size required for a satisfactory normal approximation is not just a function of n but also the distribution's shape, variance, and tail behavior (Elliott & Woodward, 2007; Sawilowsky & Blair, 1992).

A key problem highlighted in the literature is the lack of generalizable, empirically grounded guidelines for determining the minimum sample size at which the CLT yields a sufficiently accurate normal approximation across various distributions and levels of skewness. While some studies suggest heuristics based on specific parameter ranges, there is no universally accepted framework for identifying when the sample mean achieves the normal approximation for skewed or kurtotic distributions. Furthermore, most standard textbooks and guidelines do not account for the distributional differences or provide tools to empirically test the adequacy of the normal approximation in the sampling distribution (Field, 2013; Gravetter & Wallnau, 2012). Even when convergence occurs, the speed and quality of the normal approximation can vary enough to materially affect statistical inference-particularly for small to moderate sample sizes (Razali & Wah, 2011). Given the prevalence of non-normal data in applied settings, this lack of clarity can result in misguided model choices, invalid test statistics, and misleading confidence intervals.

The motivation for this study arises from this gap in applied statistical practice. While the theoretical CLT is well understood, its real-world use frequently involves assumptions that lack empirical validation. Existing rules of thumb (e.g., $n \geq 30$) are overly simplistic and fail to accommodate the distributional diversity encountered in real-world data. Researchers frequently rely on parametric methods without validating the normal approximation for their sample means potentially invalidating results. Moreover, previous studies on this topic, though informative, are often narrow in scope focusing on a limited number of distributions or specific parameter values. There is a need for a comprehensive, simulation-based approach that can examine the convergence behavior of sample means across a variety of skewed distributions and under a range of sample sizes. Formal normality tests such as Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling can provide diagnostic insights into the suitability of normal approximation in each case (Thode, 2002; Razali & Wah, 2011). Such a framework can equip researchers with distribution-specific guidelines that enhance the robustness of statistical inferences in small or moderately sized samples.

This study aims to provide a systematic, simulation-based investigation into the minimum sample size required for the sampling distribution of the mean achieve a satisfactory normal approximation across a range of skewed distributions. Focusing on the Gamma, Beta, Poisson, and Binomial distributions, it examines how varying levels of skewness influence the convergence behavior of the sample mean. Using graphical diagnostics and formal normality tests-specifically the Shapiro-Wilk and Kolmogorov-Smirnov tests, the study evaluates the adequacy of normal approximation under different sample sizes and distributional parameters. The goal is to develop empirically grounded, distribution-specific guidelines that help researchers determine whether parametric methods based on normality assumptions are appropriate, particularly in small to moderate sample size settings.

While earlier investigations (e.g., Micceri, 1989; Ahad et al., 2011; Koh & Ahad, 2020) evaluated the CLT's validity for a few continuous distributions such as Gamma, Exponential, or Chi-square, this study expands the scope by incorporating both continuous (Gamma, Beta) and discrete (Poisson, Binomial) distributions within a single, unified simulation framework. Each distribution is examined under multiple parameter combinations specifically chosen to represent low, moderate, and high skewness levels, thereby enabling a systematic cross-distribution comparison of convergence behavior. The analysis covers a wider range of sample sizes ($n = 5$ to 200) than prior studies and applies two normality criteria (Shapiro-Wilk and Kolmogorov-Smirnov tests) to determine the minimum sample size at which the CLT holds. This design not only generalizes earlier findings but also provides practically interpretable thresholds that guide researchers on when normal-based inference is reliable for skewed or discrete data. By empirically mapping how skewness affects convergence to normality, the study bridges a key applied gap in statistics education and research methodology offering distribution-specific, data-driven guidance rather than relying on the oversimplified "$n \geq 30$" heuristic.

2. Materials and Methods

This study aims to identify the minimum sample sizes required for the sample mean to exhibit a satisfactory normal approximation across distributions with varying skewness. Specifically, we focus on the Poisson, Exponential, Gamma, Binomial, and Beta distributions. The assessment of the normal approximation was carried out using two widely adopted statistical tests: the Shapiro-Wilk test and the Kolmogorov-Smirnov (KS) test**.**

2.1 Distributions and Parameters

To investigate the relationship between skewness and the sample size necessary for the CLT to hold, we selected four commonly used non-normal distributions Gamma, Poisson, Binomial, and Beta. These distributions span a broad range of skewness values and are frequently used in practical statistical applications.

Each distribution's parameters were carefully adjusted to create varying levels of skewness. The corresponding probability distributions were used to simulate random samples. The following Table 2.1 outlines each distribution, their respective functional forms, parameters, domains, and theoretical skewness expressions:

Table 2.1: Probability Distributions: PDFs/PMFs, Parameters, Support, and Skewness

Distribution	PDF / PMF	Parameters	Support	Skewness
Gamma	$f(x) = \frac{1}{[\Gamma(k)\theta^k]} x^{k-1} e^{\frac{-x}{\theta}}$	*Shape:* $k > 0$ *Scale:* $\theta > 0$	$x > 0$	$\frac{2}{\sqrt{k}}$
Binomial	$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$	$n \in \mathbb{N}$, $0 < p < 1$	$x = 0,1,\dots,n$	$\frac{(1 - 2p)}{\sqrt{[np(1-p)]}}$
Poisson	$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$	$\lambda > 0$	$x = 0,1,2,\dots$	$\frac{1}{\sqrt{\lambda}}$
Beta	$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ where, $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$	$\alpha, \beta > 0$	$0 < x < 1$	$\frac{[2(\beta-\alpha)\sqrt{(\alpha+\beta+1)}]}{\sqrt{[(\alpha+\beta+2)\sqrt{(\alpha\beta)}]}}$

2.2 Simulation study

A Monte Carlo simulation has been conducted in R to empirically assess how skewness influences the sample size required for the sampling distribution of the mean to approach normality. Multiple configurations of each distribution were used, with parameters selected to produce varying skewness.

2.2.1 Simulation Technique

The simulation design followed a fully crossed, Monte-Carlo framework in which each distribution parameter condition was paired with an ascending grid of sample sizes to trace the point at which the CLT provides an adequate normal approximation. For every parameter set-defined by a specific combination of distribution family, shape (skewness) and scale parameters-independent random samples of size n were generated, increasing in steps of five. Specifically, the sample size n increased in uniform increments of five, beginning at $n = 5$ and continuing through $n = 200$; that is, $n = \{5,10,15,\dots,200\}$. This fixed-step progression was chosen to balance computational efficiency with resolution in detecting convergence trends. At each n, 3000 replications were drawn to ensure stable estimates of the sampling distribution of the mean; this replication count is commonly recommended for Monte-Carlo accuracy in CLT studies (Micceri, 1989; Ahad et al., 2011). The mean of each replicate was computed, yielding a distribution of 3000 sample means per n. This entire loop was repeated for every parameter configuration across the Gamma, Beta, Poisson and Binomial families, thereby producing a comprehensive, distribution-specific map of convergence behavior.

In this simulation, different parameter combinations were selected for each probability distribution to generate datasets with varying degrees of skewness. Specifically, parameters were systematically adjusted within the Gamma, Beta, Poisson, and Binomial families to represent low, moderate, and high skewness conditions. This approach ensured that the simulation captured a broad spectrum of distributional shapes, allowing a comprehensive assessment of how increasing asymmetry influences the minimum sample size required for the sampling distribution of the mean to achieve normal approximation.

To evaluate whether the resulting distribution of sample means was consistent with normality, we employed two complementary goodness-of-fit tests that are widely cited for their sensitivity to different departures from normality:

1. Shapiro-Wilk test-renowned for high power against alternatives driven by skewness or kurtosis (Shapiro & Wilk, 1965; Razali & Wah, 2011).
2. Kolmogorov-Smirnov (K-S) test applied to standardized sample means, using the standard normal cumulative distribution as the reference. The K-S statistic is particularly sensitive to global shape discrepancies (Massey, 1951; Thode, 2002).

Shapiro-Wilk test

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where $x_{(1)} \leq \cdots \leq x_{(n)}$ are the order statistics of the sample, $\bar{x}$ is the sample mean, and a_i are fixed weights derived from the expected values and covariance matrix of order statistics from a standard normal distribution (Shapiro & Wilk, 1965). A smaller W indicates stronger evidence against normality.

Kolmogorov-Smirnov statistic

$$D = \max_{1 \leq i \leq n} |F_n(x_i) - F_0(x_i)|$$

where $F_n(\cdot)$ is the empirical cumulative distribution function (ECDF) of the standardized sample means, and $F_0(\cdot)$ is the standard normal CDF (Kolmogorov, 1933; Smirnov, 1948). Larger D values signal greater divergence from normality.

For both tests: H_0 : The sampling distribution of the mean is normal vs. H_1 : The distribution is not normal. With significance level $\alpha = 0.05$, we fail to reject H_0 when the test p-value exceeds 0.05. A particular sample size n was deemed sufficient for the CLT to hold only if both Shapiro-Wilk and K-S tests failed to reject H_0. The simulation for a given parameter configuration terminated at the first n satisfying this dual acceptance criterion or, if none satisfied it, at $n = 200$. The smallest n that met the dual-test criterion was recorded as the minimum sample size required for normality under that distribution-parameter setting. These minima were then collated into summary tables, organized by distribution family and indexed by skewness level, to facilitate direct comparison of convergence behavior across conditions. Such empirically derived thresholds provide practitioners with distribution-specific guidance on when parametric inference predicated on the CLT is warranted (Blanca et al., 2013; Kim & Park, 2019).

2.2.2 Analytic Strategy

The analysis is conducted based on two normality tests: the Shapiro-Wilk (SW) and Kolmogorov-Smirnov (KS) tests, with a significance level of 0.05. The primary focus is on how skewness influences the sample size required for the Central Limit Theorem (CLT) to hold.

3. Results Discussion

This section presents the simulation results for estimating the minimum sample size required to approximate the Gamma, Binomial, Poisson, and Beta distributions. Simulation was used because theoretical comparison is challenging, and later, real-life data was also employed to validate and compare the simulation outcomes.

3.1 Results for Gamma Distribution

The scale parameter of Gamma distribution which is θ equal to 1 is set as fixed. The generated sample data are tested with two normality test methods as set in R.

Table 3.1: Normality Tests for Gamma Distribution, $\alpha = 0.05$

Skewness	Minimum Sample Size	Shapiro p-value	KS p-value	Shape (k)
0.6667	15	0.8637	0.8564	9.0000
0.6984	20	0.0929	0.7181	8.2000
0.7071	20	0.4245	0.9872	8.0000
0.7845	25	0.1758	0.9098	6.5000
0.8305	25	0.6805	0.9912	5.8000
0.8944	30	0.1133	0.3844	5.0000
1.1547	45	0.1158	0.6805	3.0000
1.2172	30	0.0773	0.8642	2.7000
1.2403	35	0.2263	0.6868	2.6000
1.4142	35	0.1640	0.6933	2.0000
1.4907	65	0.0670	0.1214	1.8000
1.6330	60	0.0874	0.8088	1.5000
2.0000	75	0.1414	0.2891	1.0000
2.1082	75	0.1137	0.9591	0.9000
2.2361	85	0.2055	0.4730	0.8000

Table 3.1 presents the simulation results for the Gamma distribution, where the scale parameter was fixed at 1, and the shape parameter (k) was varied to manipulate the degree of skewness. The analysis examines how skewness affects the minimum sample size required for the sampling distribution of the mean to approximate normality, as assessed by both the Shapiro–Wilk (SW) and Kolmogorov–Smirnov (KS) tests at a significance level of 0.05. The results clearly indicate a strong and systematic relationship between the degree of skewness in the underlying distribution and the rate of convergence to normality. Specifically, as the shape parameter decreases (i.e., as skewness increases), the minimum sample size required to pass both normality tests increases substantially.

For low levels of skewness ($\leq$ 0.7), the sampling distribution of the mean approaches normality with relatively small sample sizes (typically 15-20 observations). However, for moderate skewness levels (approximately 1.0 to 1.5), sample sizes in the range of 30 to 60 are needed to meet the normality criteria. In cases of high skewness (≥ 2.0), the required sample size rises sharply, with at least 75 to 85 observations necessary for the sample mean to be considered approximately normal. These results illustrate that there is no fixed or universally applicable minimum sample size at which the CLT guarantees normality of the sample mean; rather, the appropriate sample size is distribution-specific and must account for underlying skewness.

These findings are consistent with previous studies that have examined the effect of skewness on the applicability of the CLT. Ahad et al. (2011) reported that the Gamma distribution required progressively larger sample sizes for the sample mean to converge to normality as skewness increased, with required values exceeding n = 100 for highly skewed parameter configurations. Similarly, Koh and Ahad (2020) found through simulation that Gamma and Exponential distributions with higher skewness demanded substantially greater sample sizes for the sampling distribution of the mean to pass normality tests. The present results reaffirm such observations and highlight a crucial practical implication: commonly cited rules of thumb such as "$n \geq 30$" can be misleading and inadequate when applied to non-normal or highly skewed data. Instead, skewness should be explicitly evaluated when assessing whether the assumption of normality for the sample mean is tenable.

3.2 Results for Poisson Distribution

Table 3.2: Normality Tests for Poisson Distribution, $\alpha = 0.05$

Skewness	Minimum Sample Size	Shapiro p-value	KS p-value	λ
0.3162	10	0.0805	0.1837	10.0000
0.3558	10	0.0515	0.1187	7.9000
0.3536	15	0.3514	0.2975	8.0000
0.4472	15	0.0751	0.0806	5.0000
0.5241	15	0.0767	0.1175	3.6400
0.5455	25	0.0698	0.0706	3.3600
0.6325	30	0.0853	0.0688	2.5000
0.7670	40	0.0624	0.1005	1.7000
0.7906	40	0.1065	0.1074	1.6000
0.8165	40	0.0833	0.0807	1.5000
1.1180	50	0.0894	0.0776	0.8000
1.0000	85	0.0826	0.1143	1.0000
1.2910	90	0.0949	0.0773	0.6000
1.4142	85	0.0684	0.0804	0.5000
1.1952	105	0.0751	0.0926	0.7000

Table 3.2 presents the simulation results for the Poisson distribution, focusing on how increasing skewness, governed by the rate parameter λ, influences the minimum sample size required for the sampling distribution of the mean to approximate normality. The Poisson distribution's skewness is inversely related to its mean; specifically, skewness is defined as Skewness $= \lambda^{-0.5}$, implying that lower values of λ produce greater asymmetry (Lumley et al., 2002; Zhang et al., 2015). This makes the Poisson distribution an ideal candidate for examining the CLT under varying conditions of discrete skewed data. As expected under the CLT, the rate of convergence to normality improves with increasing λ, which corresponds to decreasing skewness. For low-skewed Poisson distributions-specifically, when $\lambda \geq 5$, corresponding to skewness values between approximately 0.32 and 0.45, the simulation results show that relatively small sample sizes ($n = 10 - 15$) are sufficient for the sample mean to pass both

normality tests. This indicates that, in these conditions, the sample mean exhibits near-normal behavior even in small samples.

However, as skewness increases-driven by decreasing λ values-the sample size required for normality grows significantly. For moderately skewed distributions (i.e., $0.5 \leq$ skewness ≤ 0.8, corresponding to λ between 1.5 and 3.5), the required sample size rises to approximately 25-40 to achieve convergence as measured by both the SW and KS tests. This trend becomes more pronounced in the highly skewed region, where skewness exceeds 1.0. For example, when $\lambda = 0.5$, the skewness reaches approximately 1.4142, and the simulation reveals that at least 85 observations are needed for the sampling distribution of the mean to be accepted as approximately normal. Similarly, for skewness levels in the range of $1.2 - 1.3$ (e.g., $\lambda = 0.6 - 0.7$), the required sample size remains high-between 85 and 105-to meet the dual normality criteria.

These findings confirm that the convergence of the sample mean to normality under the CLT is substantially slower for Poisson distributions with low λ (i.e., high skewness). This result aligns with prior literature demonstrating that discrete distributions with strong asymmetry require large sample sizes for the CLT approximation to hold (Lumley et al., 2002; Zhang et al., 2015; Ahad et al., 2011). Notably, these results further challenge the validity of the commonly used heuristic that "$n \geq 30$" is sufficient for the CLT, especially when applied indiscriminately across distributions. In practical terms, this suggests that researchers and practitioners must explicitly account for the skewness inherent in their data when determining appropriate sample sizes for parametric inference. Simply relying on traditional rules of thumb without reference to distributional properties may result in misleading conclusions, particularly when working with count data or rare events, which often follow highly skewed Poisson processes.

3.3 Results for Binomial Distribution

Table 3.3: Normality Tests for Binomial Distribution, $\alpha = 0.05$

Skewness	Minimum Sample Size	Shapiro p-value	KS p-value	Size	*p*
-0.2739	10	0.0521	0.0746	30	0.8000
-0.0592	10	0.0832	0.0890	30	0.5800
-0.2869	15	0.0977	0.1428	20	0.7700
-0.2723	15	0.2517	0.1980	20	0.7600
-0.1320	15	0.3411	0.4105	30	0.6700
0.2739	15	0.0977	0.1854	30	0.2000
0.3354	15	0.1052	0.1561	20	0.2000
0.1952	20	0.0985	0.0585	20	0.3000
-0.2760	25	0.0971	0.2093	10	0.7000
-0.1291	30	0.4001	0.3429	10	0.6000
0.1291	30	0.0520	0.1202	10	0.4000
0.2760	30	0.0526	0.1132	10	0.3000
0.4743	60	0.2701	0.1827	10	0.2000
0.8433	65	0.1688	0.1124	10	0.1000
-0.7396	75	0.3051	0.1900	10	0.8800

Table 3.3 summarizes the simulation results for the Binomial distribution, demonstrating how varying degrees of skewness governed by the number of trials (size) and the probability of success (p) influence the minimum sample size required for the sampling distribution of the mean to approximate normality. The skewness of the Binomial distribution is a function of both the shape and asymmetry of the

distribution. As such, the skewness becomes positive when $p < 0.5$ and negative when $p > 0.5$, with the distribution being symmetric when $p = 0.5$ (Zhang et al., 2015; Rice, 2006). This unique property allows for a balanced exploration of how both positive and negative skewness affect the behavior of the CLT in discrete bounded distributions.

The simulation results reveal that the rate of convergence to normality is relatively fast for symmetrical or mildly skewed binomial distributions. For instance, when the skewness remains within the range of ± 0.3, which typically occurs when p is near 0.5 and n is moderate (e.g., $n = 10, p = 0.4$ or 0.6), sample sizes as small as 10-15 were sufficient for the sampling distribution of the mean to pass both the Shapiro-Wilk and Kolmogorov-Smirnov normality tests at the 0.05 significance level. This aligns with previous findings that discrete symmetric distributions such as the binomial with $p = 0.5$ exhibit rapid convergence of the sample mean to a normal distribution, even at modest sample sizes (Zhang et al., 2015; Lumley et al., 2002).

In cases of moderate skewness (e.g., around ± 0.47, such as with $p = 0.2$ or 0.8 and $n = 10$), larger sample sizes of approximately 50 to 60 were needed to achieve normality. For highly skewed binomial distributions, such as those with $p = 0.1$ or 0.88 and a small number of trials ($n = 10$), the skewness exceeds ± 0.7, and the results indicate that at least 65 to 75 observations are required for the sampling distribution of the mean to conform to the normal model. Importantly, the simulation reveals a near symmetrical pattern in convergence behavior: positively and negatively skewed parameter configurations exhibit similar sample size thresholds, suggesting that the magnitude rather than the direction of skewness primarily governs the convergence rate.

These results reinforce the principle that the CLT applies most efficiently to binomial distributions that are symmetric or only mildly skewed, allowing for reliable normal approximations even in small samples. However, when the distribution is highly skewed due to extreme probability values or small trial counts, substantially larger samples are necessary to ensure that parametric methods relying on normality assumptions are valid. The findings highlight the limitations of applying universal sample size heuristics, such as "$n \geq 30$" across all binomial configurations. Instead, practitioners must carefully evaluate the skewness introduced by their choice of p and n and adjust their sample size expectations accordingly to uphold statistical rigor in inference.

3.4 Results for Beta Distribution

Table 3.4: Normality Tests for Beta Distribution, $\alpha = 0.05$

Skewness	Minimum Sample Size	Shapiro p-value	KS p-value	α	β
-0.0820	5	0.5561	0.7088	4.0000	3.5000
0.3911	5	0.0731	0.2727	4.0000	8.5000
-0.4340	5	0.1063	0.7144	5.0000	2.5000
-0.3098	5	0.3754	0.9668	5.0000	3.0000
0.4825	10	0.1116	0.8090	3.0000	7.0000
1.0000	15	0.0582	0.4532	0.5000	1.5000
-0.5657	20	0.6681	0.9901	2.0000	1.0000
1.3613	25	0.0601	0.6460	0.7300	4.0000
-1.2257	35	0.1911	0.8172	2.0000	0.5100
-1.9660	50	0.0788	0.4087	3.5000	0.4000
1.5149	65	0.1007	0.4540	0.7200	5.0000
2.8645	105	0.4022	0.9330	0.3000	7.0000
-3.1252	105	0.0569	0.9619	4.0000	0.2000
3.1252	110	0.0563	0.9091	0.2000	4.0000
2.6389	150	0.0906	0.9444	0.3000	5.0000

Table 3.4 presents the simulation results for the Beta distribution, which is widely recognized for its flexibility in modeling a range of distributional shapes due to the influence of its two shape parameters, α and β. Depending on the combination of these parameters, the Beta distribution can be symmetric, positively skewed, or negatively skewed. This simulation investigates how varying degrees of skewness impact the minimum sample size required for the sampling distribution of the mean to approximate normality, based on the Shapiro-Wilk and Kolmogorov-Smirnov tests applied at the 0.05 significance level.

For Beta distributions that are approximately symmetric or only mildly skewed with skewness values between –0.4 and +0.4 the mean demonstrates rapid convergence to normality. In these cases, very small sample sizes, typically between 5 and 10, were sufficient for the sample mean to pass both normality tests. As skewness increases, whether in the positive or negative direction, the sample size required to achieve a normal approximation increases substantially. For distributions with moderate skewness, around ±1.0 to ±1.5, sample sizes between 15 and 65 were necessary for the mean to approximate a normal distribution.

In cases of extreme skewness, where skewness values exceed ±2.5, the minimum required sample size rises considerably. For instance, in Beta distributions where the skewness was around ±3.1252, a sample size of 105 to 110 or more was needed for the sampling distribution of the mean to pass the normality tests. These results indicate a clear relationship between the magnitude of skewness and the rate at which the CLT holds for the mean, even for bounded distributions such as the Beta.

The simulation also shows that the direction of skewness does not substantially alter the required sample size. Distributions with equivalent skewness magnitude but opposite direction (e.g., $\alpha = 4, \beta = 0.2$ versus $\alpha = 0.2, \beta = 4$) demonstrated similar convergence behavior in terms of required sample size. This suggests that the extent of asymmetry, rather than its direction, is the more critical factor in determining how quickly the sampling distribution of the mean approximates normality.

These results emphasize the importance of considering distributional shape when determining sample size for inferential statistics. While smaller samples may be sufficient for symmetric or lightly skewed data, substantially larger samples are needed for highly skewed distributions to justify the use of parametric methods that rely on normality assumptions. In applied settings where the Beta distribution is often used such as in proportion data, rates, and bounded outcomes such sample size considerations are critical for ensuring valid statistical inference.

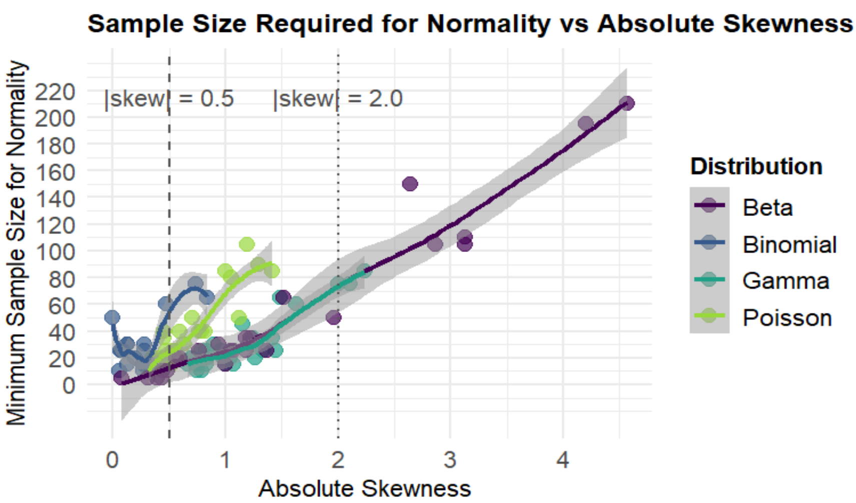

Figure 1. Visualization of Skewness Effect on Required Sample Size

Figure 1 provides a comprehensive visual summary of the relationship between absolute skewness and the minimum sample size required for the sampling distribution of the mean to approximate normality. The figure synthesizes simulation results across four distinct probability distributions Beta, Binomial, Gamma, and Poisson each representing different combinations of distributional shape, support, and discreteness. The x-axis displays the absolute value of skewness, allowing for a unified analysis regardless of the direction of asymmetry, while the y-axis represents the minimum sample size at which both the Shapiro-Wilk and Kolmogorov-Smirnov tests fail to reject the null hypothesis of normality at the 0.05 significance level.

Each data point corresponds to a specific parameter configuration from the simulation, and points are color-coded by distribution to facilitate interpretation. Smoothed trend lines are included for each distributional family to capture the general pattern of how required sample size changes with increasing skewness.

The figure reveals a clear positive nonlinear relationship between skewness and the sample size required for satisfactory normal approximation. As the absolute skewness increases, the minimum sample size needed for the sample mean to achieve the normal approximation rises substantially, supporting the theoretical expectation that convergence under the CLT slows in the presence of asymmetry. Among the distributions, the Gamma distribution exhibits a gradual and monotonic increase in required sample size, following a smooth, predictable curve. In contrast, the Beta distribution (red) displays a sharper and steeper increase, particularly for cases with extreme skewness (i.e., > 2.0), where the required sample size exceeds the upper bound of the simulation range. These cases are visually capped at $y = 210$ to indicate that convergence had not occurred even at the highest simulated sample size.

The Poisson and Binomial distributions exhibit more irregular and nonlinear patterns, which may be attributed to their discrete nature and the interaction between their defining parameters (e.g., probability and size in Binomial; rate in Poisson). Despite this variability, both distributions show a general upward trend, with lower-skewed configurations (e.g., skewness < 0.5) achieving the normal approximation with relatively small sample sizes (5-20). This reinforces the understanding that the CLT applies more quickly in near-symmetric settings.

4. Application

To complement the findings from the simulation study, an analysis was conducted using a real-world dataset to assess the implications of skewness on the applicability of the CLT in practical settings. The dataset selected for this purpose is the, publicly available on Kaggle (Arun Jangir & Willian Oliveira, 2023). This dataset contains records for 1,338 individuals residing in the United States and includes a variety of demographic and behavioral attributes, such as age, body mass index (BMI), smoking status, region, sex, and most notably, medical insurance charges the variable of primary interest in this analysis.

The charges variable, which reflects each individual's annual medical expenses, is known to be highly right-skewed, making it a suitable candidate for evaluating the robustness of the CLT in the context of skewed empirical data. Preliminary exploratory analysis, including a histogram and summary statistics, confirmed a pronounced skew in the distribution. The range of charges spans from approximately $1,122 to over $63,000, with a strong concentration of observations in the lower range and a long right tail.

To quantify the asymmetry, sample skewness was calculated and found to be 1.51, indicating a substantial departure from symmetry. This level of skewness falls within the range identified in the simulation study as requiring moderate to large sample sizes for the sample mean to exhibit a satisfactory normal approximation. Therefore, this dataset provides a meaningful real-world illustration of the simulation findings, allowing for empirical validation of the relationship between skewness and the convergence behavior of the sample mean.

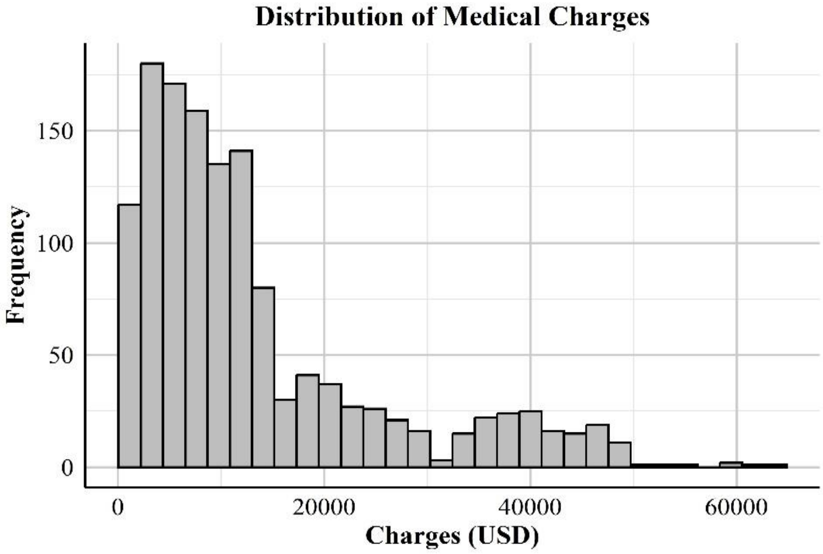

Figure 2. Histogram of Medical Charges

To empirically determine the minimum sample size required for the sample mean of annual medical charges to achieve normal approximation, a Monte Carlo simulation was conducted using bootstrap resampling techniques. Specifically, for each sample size $n = 10, 20, \dots, 300$, a total of 1,000 bootstrap samples were drawn with replacement from the original dataset, and the sample mean was computed for each replicate. The resulting distribution of sample means for each value of n was subjected to formal normality tests, namely the Shapiro-Wilk and Kolmogorov-Smirnov tests. A significant level of α=0.05 was employed, and the proportion of simulations that failed to reject the null hypothesis of normal approximation was recorded at each sample size.

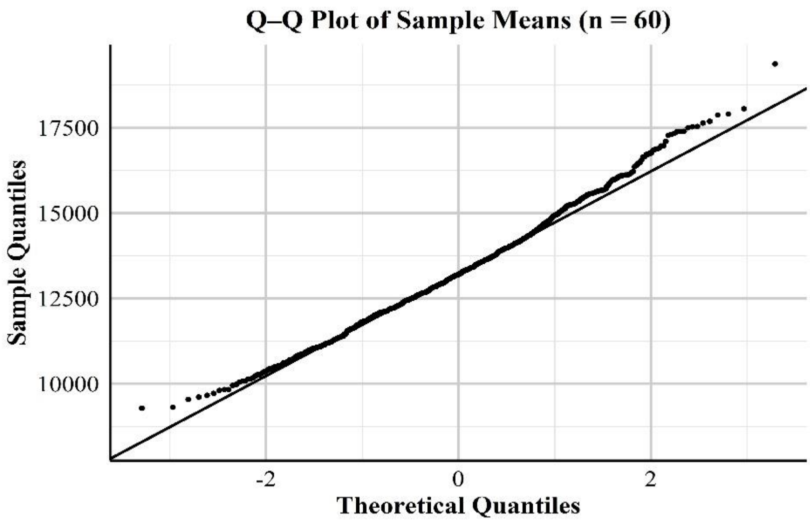

Figure 3. Normal Q-Q Plot

The results revealed that when the sample size reached $n = 60$, both normality tests consistently failed to reject the null hypothesis, indicating that the sampling distribution of the mean provided a sufficiently accurate normal approximation. Despite the strong right skewness (Figure 2) of the original charges variable (skewness = 1.51), these findings suggest that the CLT begins to take practical effect at moderate sample sizes in real-world data scenarios. Additional visual diagnostics at the $n = 60$ threshold supported these statistical results: the Q–Q plot (Figure 3) of the sample means exhibited near-linear alignment with the theoretical normal distribution, and the corresponding histogram demonstrated a bell-shaped curve characteristic of the normal approximation (Figure 4). The real-world figures (histogram and Q–Q plot in Figures 3 and 4) were generated from the same set of bootstrap resamples used for the normality tests at the selected sample size ($n = 60$), ensuring consistency between the statistical evaluation and visual diagnostics, and they were independent of the earlier parametric simulation runs.

The sample size requirement observed in this empirical setting closely mirrors the simulation findings for Gamma distributions with comparable levels of skewness, further validating the broader conclusion that higher skewness necessitates larger samples for the sample mean to conform to the normal approximation. However, unlike controlled simulations, real-world datasets often contain additional complexities such as outliers, heteroscedasticity, and data irregularities, which may impact convergence behavior. Thus, conducting empirical resampling-based assessments remains a critical step in determining the appropriateness of parametric inference in applied research contexts involving skewed data.

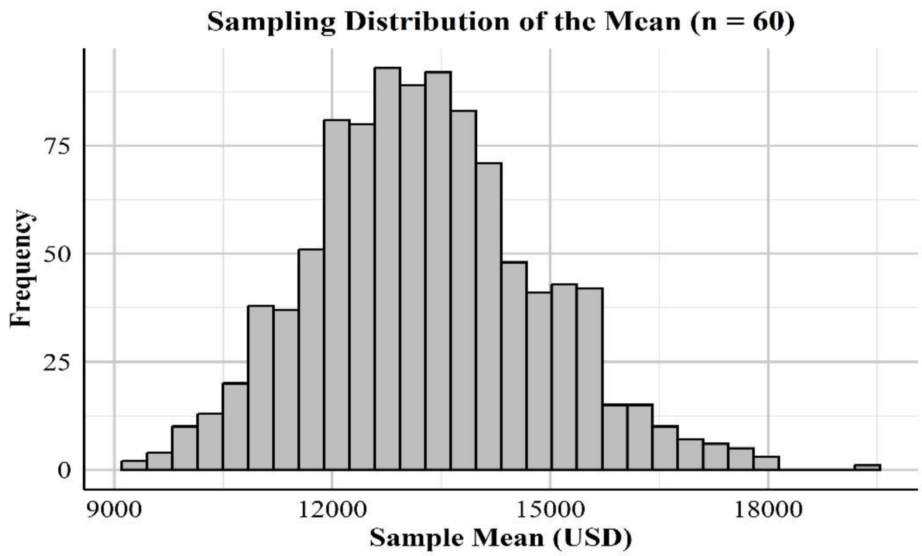

Figure 4. Sampling Distribution

5. Conclusions and Future Research

5.1 Conclusions

This study systematically examined the relationship between skewness in a population distribution and the minimum sample size required for the sampling distribution of the mean to achieve a satisfactory normal approximation, as prescribed by the Central Limit Theorem (CLT). Using extensive Monte Carlo simulations, we evaluated four commonly encountered skewed distributions: Gamma, Poisson, Binomial, and Beta, assessed convergence to the normal approximation through two established tests: Shapiro–Wilk and Kolmogorov–Smirnov, each conducted at a 5% significance level.

The results demonstrated a clear and consistent trend: as skewness increases, a larger sample size is required for the sample mean to achieve a reliable normal approximation. However, the rate of increase and threshold varied by distributional form. For example, in the Gamma distribution, only 15 samples were needed under low skewness, but more than 75 were required when skewness exceeded 2.0. The Poisson distribution showed similarly sharp increases, with sample size requirements rising from 10 to over 100 as skewness grew. In the Binomial distribution, symmetric and mildly skewed configurations required relatively small samples (10-15), but higher skewness values (±0.7 or more) necessitated 65-75 observations. The Beta distribution, due to its extreme flexibility in shape, required over 100 observations when skewness surpassed ±3.0, indicating a slower convergence to the normal approximation than the other distributions.

These findings were further supported by an analysis of a real-world dataset the Healthcare Insurance dataset which contains highly skewed medical expense data. Bootstrap-based simulation revealed that the sample mean of charges achieved a satisfactory normal approximation at a sample size of $n = 60$, confirming the simulation trends observed for distributions with similar skewness levels. This empirical validation reinforces the practical significance of accounting for skewness when applying the CLT to real-world problems.

In sum, this study highlights that rules of thumb, such as assuming $n \geq 30$ is sufficient for a normal approximation, may be overly simplistic and potentially misleading when applied to skewed distributions. The simulation results, combined with real-world validation, suggest that researchers and practitioners must assess distributional characteristics, especially skewness, when planning studies or making statistical inferences that rely on the normal approximation of the sample mean. By doing so, one can ensure more accurate conclusions and avoid violations of underlying assumptions that could compromise statistical validity.

5.2 Future Research

Building on these findings, future research may investigate how dependencies among observations, outliers, or heteroscedasticity influence convergence to the normal approximation**.** Additionally, exploring the performance of other estimators such as medians, trimmed means, or variance estimates under similar conditions could offer further insight into robust statistical inference. Expanding the analysis to include non-iid data, transformation-based approaches, or Bayesian frameworks may also enhance our understanding of inference under skewness. Future studies might also explore transformation techniques (e.g., Box–Cox) or nonparametric methods as alternative pathways when normal approximation fail, providing flexibility for distributions that resist the normal approximation. Ultimately, the development of automated, data-driven diagnostic tools to assess normal approximation requirements could help guide sample size decisions in practice, ensuring that the Central Limit Theorem is applied judiciously and effectively in real-world settings.

Acknowledgements

The authors are grateful to the anonymous referees for their constructive comments and suggestions, which certainly helped to improve the quality and presentation of the paper.

6. References

Arun Jangir, and willian oliveira. (2023). Healthcare Insurance [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/6678394

Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences* (4th ed.). Pearson Education.

Ahad, N. A., Rohani, M. F., & Abidin, A. Z. (2011). The Validity of Central Limit Theorem in non-normal Distributions. *Proceedings of the International Conference on Statistics in Science, Business and Engineering* (ICSSBE), 1–5.

Ahad, N. A., Rohani, M. F., & Abidin, A. Z. (2011). The validity of Central Limit Theorem in non normal Distributions. *Proceedings of the International Conference on Statistics in Science, Business and Engineering.*

Blanca, M. J., Arnau, J., López Montiel, D., Bono, R., & Bendayan, R. (2013). Non normal data: Is ANOVA Still a Valid Option? *Psicothema*, 25(2): 306–312.

Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). Duxbury Press.

Chang, L., & Zhang, Y. (2006). Central Limit Theorem and Convergence to Normality. *Journal of Modern Applied Statistical Methods*, 5(1): 140–144. https://doi.org/10.22237/jmasm/1146456000

Elliott, A. C., & Woodward, W. A. (2007). *Statistical Analysis Quick Reference Guidebook: With SPSS Examples.* Sage Publications.

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics* (4th ed.). Sage Publications.

Gravetter, F. J., & Wallnau, L. B. (2012). *Essentials of Statistics for the Behavioral Sciences* (7th ed.). Cengage Learning.

Hogg, R. V., McKean, J., & Craig, A. T. (2013). *Introduction to Mathematical Statistics* (7th ed.). Pearson Education.

Kim, H. Y. (2013). Statistical Notes for Clinical Researchers: Assessing Normal Distribution (2) using Skewness and Kurtosis. *Restorative Dentistry & Endodontics*, 38(1): 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Kim, T. K., & Park, J. H. (2019). More about the Central Limit Theorem and Why it Matters in Medical Research. *Korean Journal of Anesthesiology*, 72(4): 324–330. https://doi.org/10.4097/kja.19168

Koh, K. L., & Ahad, N. A. (2020). Assessing the Validity of the Central Limit Theorem for Skewed Distributions. *Malaysian Journal of Mathematical Sciences*, 14(1): 53–68.

Kolmogorov, A. N. (1933). Sulla Determinazione Empirica di Una Legge di Distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4: 83–91.

Kwak, S. G., & Kim, J. H. (2017). Central Limit Theorem: The cornerstone of Modern Statistics. *Journal of Rheumatic Diseases*, 24(2): 83–88. https://doi.org/10.4078/jrd.2017.24.2.83

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health*, 23: 151–169. https://doi.org/10.1146/annurev.publhealth.23.100901.140546

Massey, F. J. Jr. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46: 68–78.

Micceri, T. (1989). The Unicorn, the Normal Curve, and other Improbable Creatures. *Psychological Bulletin*, 105(1): 156–166. https://doi.org/10.1037/0033-2909.105.1.156

Razali, N. M., & Wah, Y. B. (2011). Power Comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling Tests. *Journal of Statistical Modeling and Analytics*, 2(1): 21–33.

Rice, J. A. (2006). *Mathematical Statistics and Data Analysis* (3rd ed.). Duxbury Press.

Sawilowsky, S. S., & Blair, R. C. (1992). A More Realistic Look at the Robustness and Type II Error Properties of the T-Test to Departures from Population Normality. *Psychological Bulletin*, 111(2): 352–360. https://doi.org/10.1037/0033-2909.111.2.352

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3–4): 591–611.

Smirnov, N. V. (1948). Table for Estimating the Goodness of Fit of Empirical Distributions. *Annals of Mathematical Statistics*, 19: 279–281.

Thode, H. (2002). *Testing for Normality*. Marcel Dekker. Inc. New York, 99-123.

Zhang, C., Liu, Y., & Liu, H. (2015). Simulation Study on the Normality of Sample Means from Poisson and Binomial Distributions. *Journal of Applied Mathematics and Physics*, 3(1): 52–58. https://doi.org/10.4236/jamp.2015.31008