Central Limit Theorem Simulation Study: A Monte Carlo Analysis of Normal Approximation

Jialiang Wu (alfredwu)

December 7, 2025

1 Introduction

The Central Limit Theorem (CLT) is one of the foundational results in probability. It guarantees that the standardized sample mean converges to a normal distribution as the sample size increases. A common empirical rule states that a sample size of $n \geq 30$ is "large enough" for the CLT to hold. However, this rule often fails for highly skewed, heavy-tailed, or discrete distributions. In such cases, the rate of convergence to normality can be extremely slow, leading to inaccurate inference when normal-based methods are used.

This project aims to empirically investigate how quickly the standardized sample mean approaches normality under several representative non-normal distributions. The goal is to evaluate when the CLT approximation becomes reliable, and to assess whether the common $n \geq 30$ guideline is adequate across distributions with different structural characteristics.

2 Statistical Methodology

To capture different deviations from normality, I study four distributions representing distinct structural challenges for CLT convergence:

- **Bernoulli(p)**: a discrete, bounded distribution. For extreme probabilities (e.g., $p = 0.1$ or 0.9), the distribution is highly skewed, making it a strong test case for discreteness and asymmetry.
- **Exponential(λ)**: a continuous, strictly right-skewed distribution with a fixed shape and moderate heavy tail. Because scaling does not affect its shape, it isolates the effect of skewness alone on CLT convergence.
- **Lognormal($\mu = 0, \sigma$)**: a continuous distribution with strong right skewness and heavy tails. Larger values of σ amplify skewness and kurtosis, making the Lognormal distribution representative of extremely slow convergence cases.
- **Student's t_ν**: a symmetric but heavy-tailed distribution. Smaller degrees of freedom ν induce substantial kurtosis, allowing me to study heavy-tailed effects independent of skewness.

These four distributions collectively cover discreteness, pure skewness, heavy tails, and combinations thereof, enabling a comprehensive assessment of factors influencing CLT behavior.

For each distribution, I generate a large pseudo-population and repeatedly draw bootstrap samples across a range of sample sizes. To better capture the early stages of CLT convergence, the grid is denser for small sample sizes and sparser for larger ones: $n \in \{2, 3, 4, 5, 8, 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500\}$. For each sample, the standardized sample mean is computed, and its empirical distribution is compared to the standard normal based on Monte Carlo replications (e.g., $B = 2000$). This design allows me to track how normality emerges as n increases.

To evaluate convergence comprehensively, I compute seven metrics for each distribution–sample-size combination, according to Table 1

3 Results

3.1 Bernoulli Distributions

For Bernoulli distributions, the rate of convergence to normality depends strongly on the success probability p. As shown in Figures 1–2, when p is close to 0.5, the distribution is symmetric with moderate variance, and the standardized mean converges rapidly. In contrast, when $p \in \{0.1, 0.9\}$, the distribution is highly skewed, and the CLT operates much more slowly. This pattern appears consistently across all seven normality metrics. The KS statistic and the Shapiro–Wilk statistic both show that samples from $p = 0.1$ and $p = 0.9$ remain far from normal at $n = 30$, while the cases $p = 0.3, 0.5, 0.7$ display substantially better fit. Similarly, the 95% confidence-interval coverage stabilizes near the nominal level for moderate values of p, but is still unstable for the more extreme parameters. QQ-plot R^2, tail-probability error, skewness, and kurtosis also indicate slow convergence at $p = 0.1$ and $p = 0.9$, with persistent deviations from symmetry and Gaussian tail behavior. Overall, the results demonstrate that $n = 30$ is insufficient for the CLT to yield an approximately normal standardized mean under highly imbalanced Bernoulli distributions, whereas moderate values of p show much faster convergence.

3.2 Exponential Distribution

For the exponential distribution, the standardized sample mean converges to normality much faster than in the highly skewed Bernoulli cases, which are shown in Figures 3–4. Both the KS statistic and the Shapiro–Wilk statistic show that by $n = 30$, the distribution of the standardized mean is already close to Gaussian, although further improvement continues as n increases. The nominal 95% confidence interval coverage is reasonably accurate at $n = 30$, but exhibits noticeable oscillation for larger n, reflecting the sensitivity of exponential samples to occasional large observations arising from the heavy right tail. The QQ-plot R^2 statistic is near one by $n = 30$, and the kurtosis of the standardized mean stabilizes around the Gaussian value of three. In contrast, skewness decays more slowly and remains above zero at $n = 30$. Tail-probability error decreases substantially with increasing n, reaching its minimum around $n \approx 200$ before fluctuating slightly due to occasional tail-driven variability. As expected, changing the rate parameter λ affects only the scale of the original distribution and therefore has no substantive impact on the convergence behavior. All seven normality metrics exhibit nearly identical patterns across $\lambda = 1$ and $\lambda = 3$. Overall, the exponential distribution demonstrates relatively fast convergence to normality.

3.3 Lognormal Distribution

For the lognormal distribution, the degree of right-skewness is controlled by the scale parameter σ, and the rate of convergence to normality decreases rapidly as σ increases. These can be viewed in Figures 5–6, When $\sigma = 0.5$, the standardized mean is already close to Gaussian by $n = 30$, with the KS statistic, Shapiro–Wilk statistic, QQ-plot R^2, and skewness all indicating near-normal behavior, although several metrics continue to improve slightly as n grows. In contrast, for $\sigma = 1$, convergence is much slower: normality metrics remain far from their Gaussian targets at $n = 30$, and substantial improvement requires sample sizes in the hundreds. The case $\sigma = 1.5$ is even more extreme; the distribution is so heavily skewed that even at $n = 500$, the standardized mean exhibits severe deviations from normality. The 95% CI coverage reflects this pattern as well, with $\sigma = 0.5$ showing approximate validity at $n = 30$, while $\sigma = 1$ and $\sigma = 1.5$ display large and persistent instability. Tail-probability error also highlights the difficulty of achieving normality: even at $n = 500$, the larger σ cases exhibit substantial discrepancies relative to the Gaussian tail. In contrast, kurtosis stabilizes quickly for all σ, as expected from the rapid attenuation of higher-order moments under averaging. Overall, the lognormal distribution demonstrates extremely slow CLT convergence when σ is large, and only the mildly skewed case $\sigma = 0.5$ attains approximate normality at $n \approx 30$.

3.4 t-Distribution

For the Student-t distribution, the convergence behavior under the CLT depends strongly on the degrees of freedom. The cases df $= 1$ and df $= 2$ are special: the former has no finite mean, and the latter has infinite variance, so the standardized sample mean converges extremely slowly or can't approach convergent. This is reflected in all the normality metrics, which can be viewed in Figures 7–8, where skewness and kurtosis do not approach their Gaussian targets until sample sizes exceed $n \approx 200$. For degrees of freedom df $= 3, 4, 5$, however, the variance is finite, and the CLT operates in the usual manner. Across these values of df, all seven normality metrics indicate rapid convergence. The KS statistic, Shapiro–Wilk statistic, QQ-plot R^2, and tail-probability error all show near-Gaussian behavior by $n = 30$. The 95% confidence-interval coverage is also close to the nominal level, with only slight oscillation in the case df $= 5$, where the distribution retains slightly heavier tails. Skewness and kurtosis both stabilize quickly as well, consistent with the symmetry of the t distribution. Overall, for df ≥ 3, the standardized mean converges to normality much faster than in the lognormal or exponential settings, and even modest sample sizes yield behavior very close to the Gaussian.

4 Discussion and Conclusion

This study evaluates the rate at which the standardized sample mean approaches normality under four representative families of non-normal distributions. Across all distributions, the results show that the common rule, $n \geq 30$, works well only in some situations. The adequacy of $n = 30$ depends essentially on the underlying distribution's skewness, tail heaviness, discreteness, and moment existence.

According to the results, it is showed that the common rule, $n \geq 30$, works well only in some situations. For symmetric or mildly skewed distributions—such as Bernoulli with moderate p, Exponential, and t-distributions with $\nu \geq 3$—the CLT approximation becomes reasonably accurate by $n = 30$. In these cases, most normality metrics are close to their target values, and the standardized mean behaves much like a Gaussian.

However, for distributions that are highly skewed or heavy-tailed, convergence is much slower. Log-normal distributions with large σ and t-distributions with very small degrees of freedom require samples in the hundreds before the standardized mean looks approximately normal, and some tail-related metrics remain unstable even at $n = 500$. These findings highlight that skewness and heavy tails are the main factors that delay CLT convergence.

Putting the results together, the study shows that the four distribution families form a comparatively clear ordering of convergence difficulty. The ranking highlights that skewness + heavy tails in combination produce the slowest convergence overall. By contrast, symmetry + finite variance tends to guarantee fast convergence, even in heavy-tailed situations.

5 Appendix

5.1 A.1 Summary of Normality Metrics

Table 1 summarizes the seven metrics used to evaluate the convergence of the standardized sample mean.

Table 1: Summary of normality metrics

Metric	**Definition**	**Purpose**
KS statistic	$D = \sup_z \lvert F_n(z) - \Phi(z) \rvert$	CDF deviation
95% CI coverage	$\mathbf{1}\{\mu \in [\bar{X} \pm 1.96\sigma/\sqrt{n}]\}$	Interval accuracy
QQ-plot R^2	R^2 from normal QQ-line fit	Quantile alignment
Tail error	$\lvert P(Z > 2) - 0.0228 \rvert$	Extreme tail fit
Skewness	$\gamma_1 = m_3/m_2^{3/2}$	Asymmetry
Excess kurtosis	$\gamma_2 = m_4/m_2^2 - 3$	Tail heaviness
Shapiro–Wilk W	$\frac{(\sum a_i x_{(i)})^2}{\sum (x_i - \bar{x})^2}$	Normality test

5.2 A.2 Bernoulli(p) Results

Figures 1–2 present all seven metrics for the Bernoulli distributions.

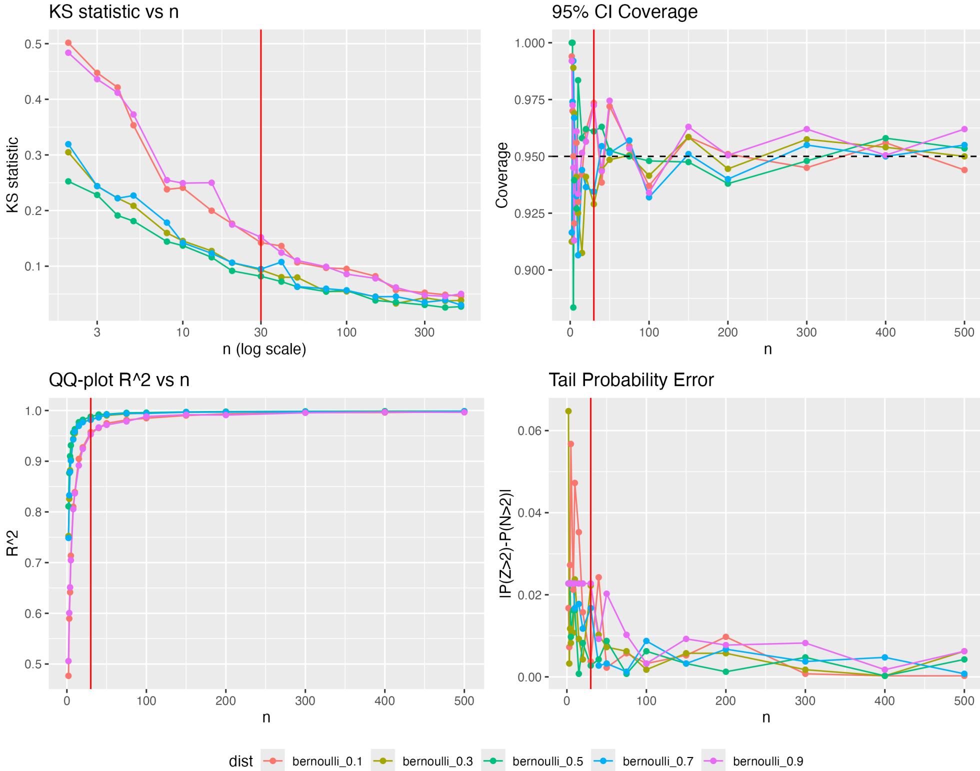

Figure 1: KS, CI coverage, QQ-plot R^2, and tail error for Bernoulli(p).

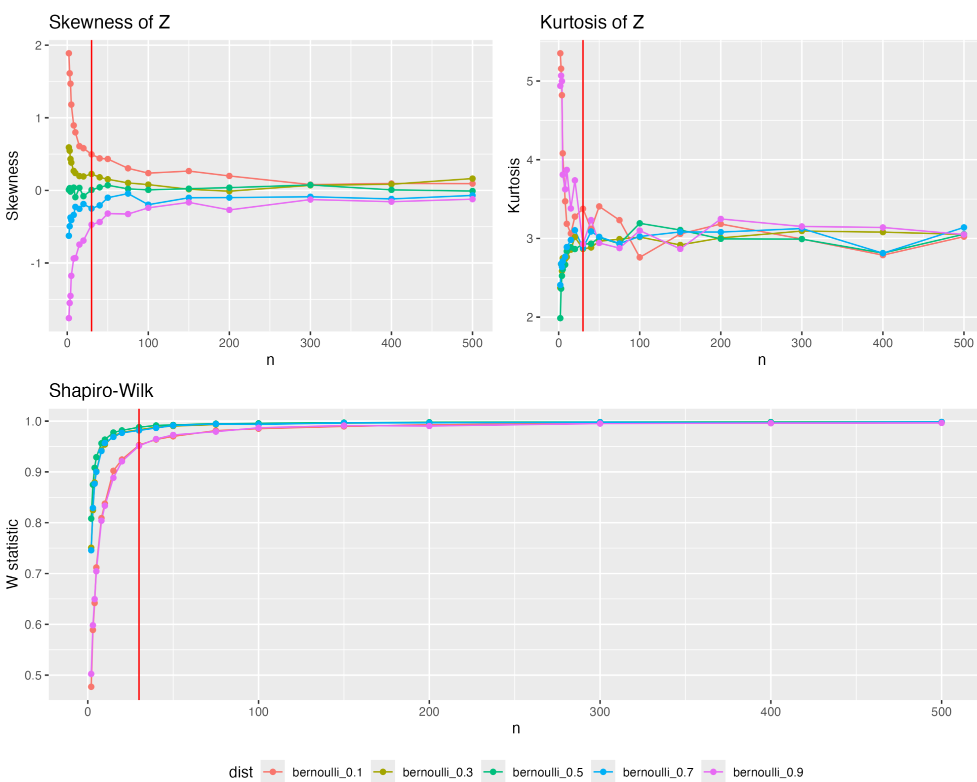

Figure 2: Skewness, kurtosis, and Shapiro–Wilk statistic for Bernoulli(p).

5.3 A.3 Exponential(λ) Results

Figures 3–4 present all seven metrics for the exponential distributions.

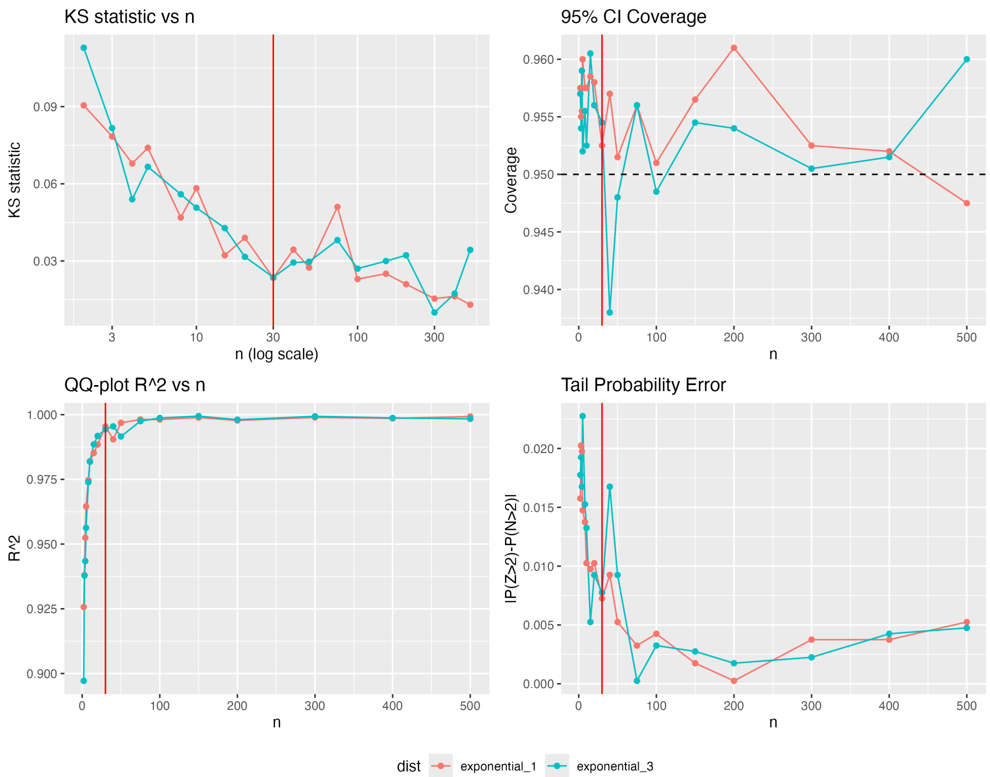

Figure 3: KS, CI coverage, QQ-plot R^2, and tail error for Exponential(λ).

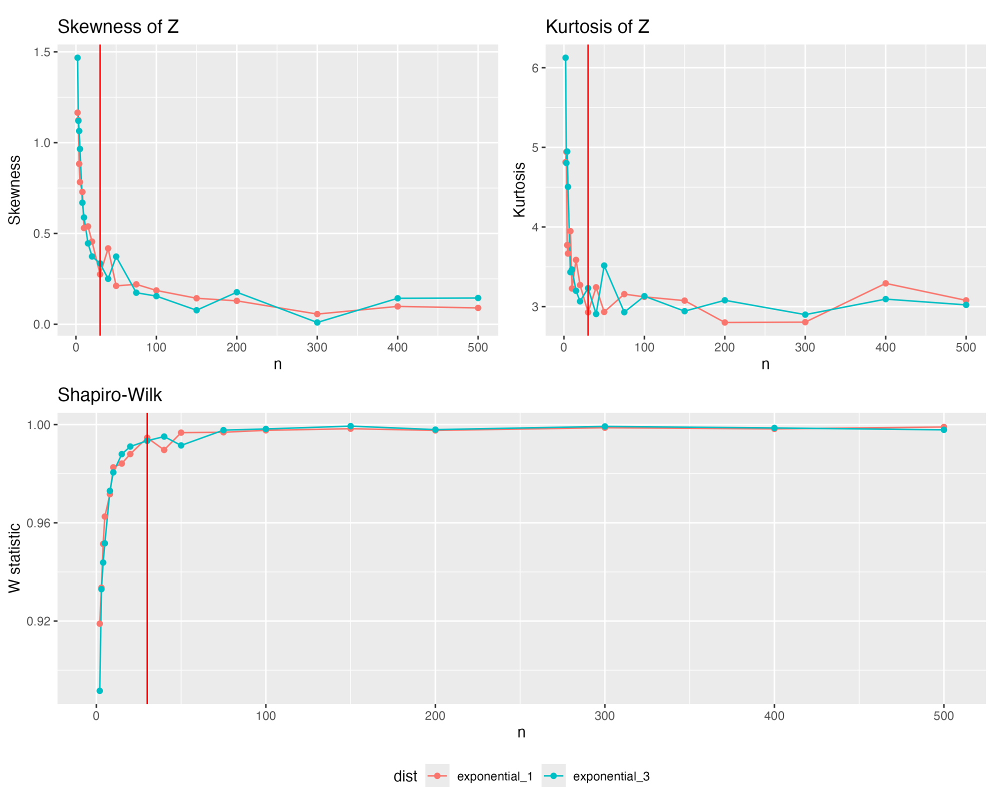

Figure 4: Skewness, kurtosis, and Shapiro–Wilk statistic for Exponential(λ).

5.4 A.4 lognormal($\mu = 0,\ \sigma$) Results

Figures 5–6 present all seven metrics for the Bernoulli distributions.

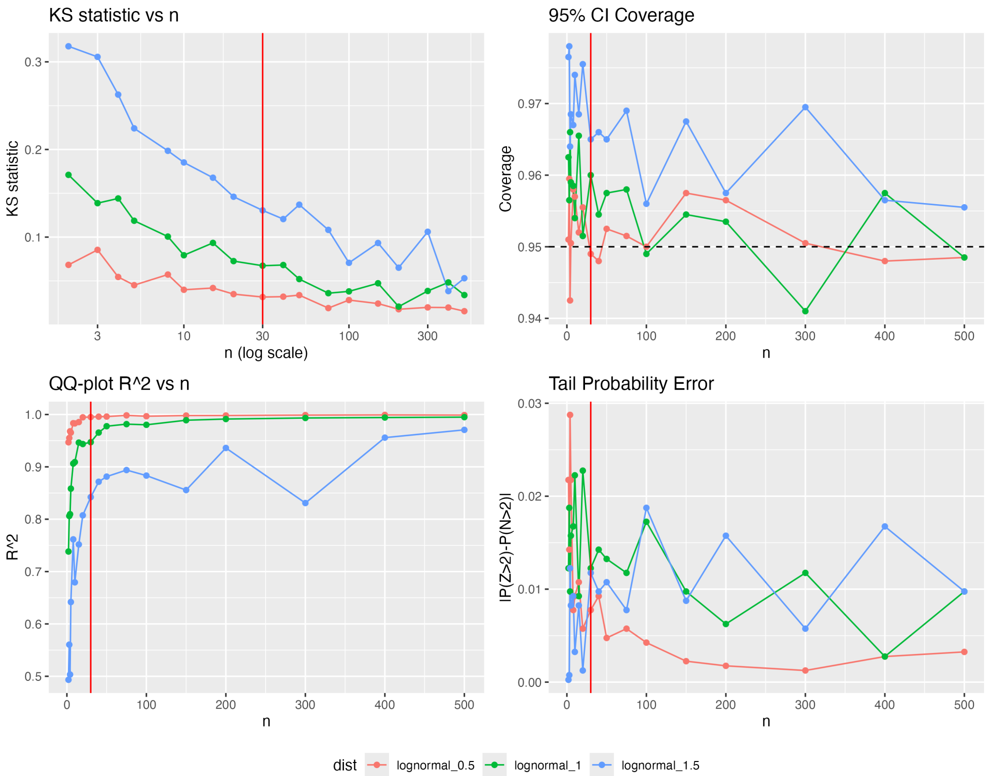

Figure 5: KS, CI coverage, QQ-plot R^2, and tail error for lognormal($\mu = 0,\ \sigma$).

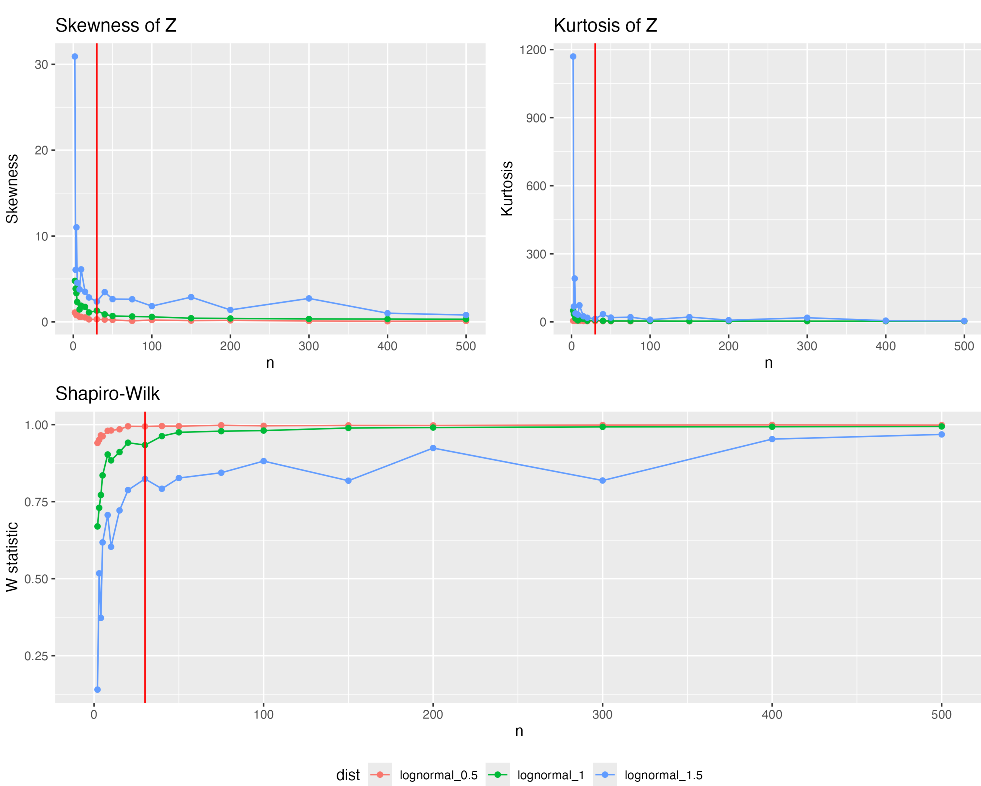

Figure 6: Skewness, kurtosis, and Shapiro–Wilk statistic for lognormal($\mu = 0,\ \sigma$)).

5.5 A.5 t distribution(df) Results

Figures 7–8 present all seven metrics for the Bernoulli distributions.

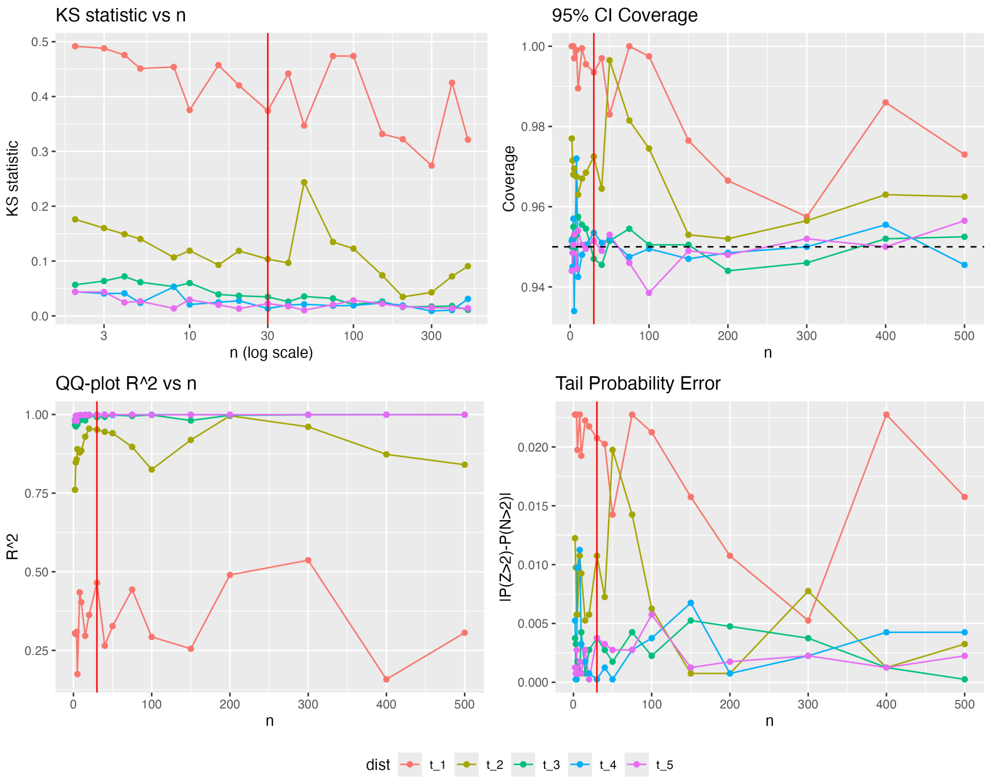

Figure 7: KS, CI coverage, QQ-plot R^2, and tail error for t distribution(df).

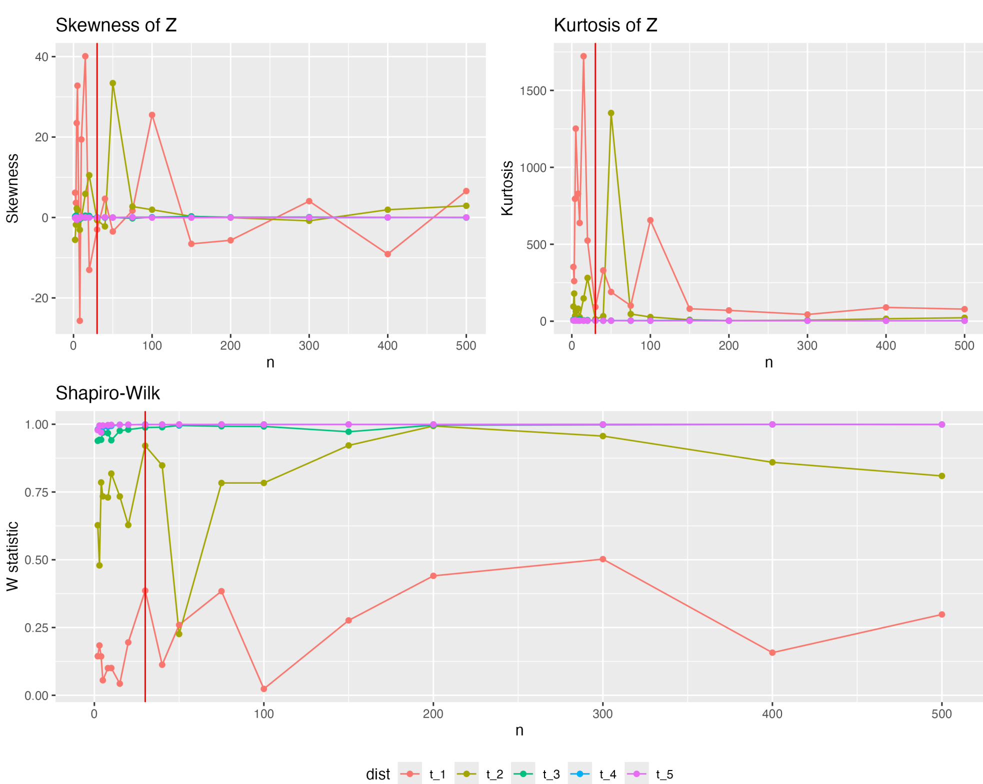

Figure 8: Skewness, kurtosis, and Shapiro–Wilk statistic for t distribution(df).