INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

Comparative Performance Between Spearman's Rho and Pearson's R Correlation Coefficient in Measuring Statistical Relationships at Different Population Distribution

Gilbert Alerta

Department of Science and Technology, Philippine Science High School, Caraga Region Campus, Butuan City

DOI: https://doi.org/10.47772/IJRISS.2026.100800250

Received: 21 August 2026; Accepted: 26 August 2026; Published: 01 September 2026

ABSTRACT

In many correlational studies, Pearson's correlation coefficient (r) is commonly used to assess bivariate distributions. However, this statistic is inappropriately applied when variables are highly skewed, whereas Spearman's Rho (ρ_s) is better suited. This study used a Monte Carlo simulation to compare the accuracy and stability of Pearson's r and Spearman's rho across three population distributions that differed in skewness, kurtosis, and the presence of outliers (normal, uniform, exponential), at sample sizes ranging from n = 10 to n = 1,000, and at population associations of $\rho_s = 0$ (no relationship) and $\rho_s = .50$ (moderate monotonic relationship). Bivariate samples were generated using a Gaussian copula so that the rank-order dependence structure was held constant while the marginal distributions were manipulated. Each condition was replicated 3,000 times. Results showed that under the symmetric distributions (normal, uniform), Pearson's r was slightly more efficient (lower sampling variability) than Spearman's rho, consistent with classical asymptotic relative efficiency theory. Under the skewed exponential distribution, this pattern reversed: Spearman's rho showed consistently lower variability and, at larger sample sizes, lower root-mean-square error than Pearson's r, with the variance ratio favoring Spearman's rho by as much as 1.6:1 at n = 1,000. Both coefficients were approximately unbiased under the null hypothesis of no association, regardless of distributional shape. These findings indicate that no single coefficient is uniformly superior: Pearson's r is preferable for approximately symmetric, light-tailed data, whereas Spearman's rho offers a more robust and often more efficient estimate of association when data are skewed.

Keywords: Pearson's correlation, Spearman's correlation, bivariate distribution, Monte Carlo simulation

INTRODUCTION

Correlation analysis is among the most frequently applied statistical procedures for quantifying the strength and direction of association between two continuous or ordinal variables. Of the many correlation indices available, Pearson's product-moment correlation coefficient (r) and Spearman's rank-order correlation coefficient (rho, ρ_s) are by far the most commonly reported in empirical research across psychology, education, health sciences, and business (Dazong & Terrang). Pearson's r estimates the degree of linear association between two variables and is derived under the assumption that the variables are bivariate normally distributed, measured on interval or ratio scales, and free of substantial outliers. Hurley (2025) noted that Spearman's rho, in contrast, is computed on the ranked data and estimates the degree of monotonic, not necessarily linear, association, making it a distribution-free, nonparametric alternative that does not require the assumption of normality.

Despite the conceptual distinction between these two coefficients, applied researchers often inconsistently select one over the other, sometimes reporting Pearson's r by default even when distributional assumptions are clearly violated, or applying Spearman's rho unnecessarily when the data are already well-behaved. This inconsistency is problematic because the relative statistical performance of the two coefficients, in terms of bias, sampling variability, and accuracy in recovering the true population relationship, depends heavily on the shape of the underlying sampling distribution and on sample size (Bocianowski et. al., 2024). Classical theory establishes that when data are bivariate normal, Pearson's r is the maximum-likelihood estimator of the population correlation and is more statistically efficient than Spearman's rho, with an asymptotic relative efficiency of approximately 91.2% for rho relative to r (Sathyanarayana & Mohanasundaram, 2025). However, this efficiency advantage of Pearson's r is known to erode, and even reverse, once the normality assumption is violated.

A growing body of simulation research has demonstrated that non-normal data, particularly skewed data, data with heavy tails, or data containing outliers, can substantially inflate the variability and bias of Pearson's r, while Spearman's rho remains comparatively stable under the same conditions. Ventura-León et al., (2023) conducted an extensive simulation across sample sizes ranging from 5 to 1,000 and found that although Pearson's r and Spearman's rho performed almost identically under normality, Spearman's rho showed lower variability than Pearson's r once the variables exhibited high kurtosis, and reduced the standard deviation of the estimate by roughly 20% relative to Pearson's r in heavy-tailed empirical survey data. Similarly, Zrelak (2026) mentioned that Pearson's r becomes increasingly biased and its associated significance tests increasingly invalid as non-normality increases, and that rank-based alternatives such as Spearman's rho largely eliminate this inflation. Lo & Christian (2026) further showed, using both analytical and Monte Carlo methods, that Spearman's rho is considerably more robust than Pearson's r to outlier contamination in otherwise normal data. Conversely, Ji (2023) cautioned that rank-based coefficients do not automatically outperform Pearson's r in every non-normal scenario, particularly for hypothesis testing of the null of no association, where the two approaches often perform similarly.

Taken together, the existing literature suggests that the comparative performance of Pearson's r and Spearman's rho is not fixed but is conditional on the sampling distribution from which the data arise. What remains useful to document in a compact, easily replicable design is a systematic comparison across a small, well-differentiated set of distributional conditions, spanning a symmetric mesokurtic baseline, a symmetric bounded (platykurtic) distribution, and a right-skewed distribution, evaluated jointly against sample size and against both a null and a non-null population relationship. This gap motivates the present study.

Accordingly, this study used a controlled Monte Carlo simulation to address the following research questions: (1) How do Pearson's r and Spearman's rho compare in terms of bias relative to the population correlation across the normal, uniform, and exponential sampling distributions? (2) How do the two coefficients compare in terms of sampling variability (standard deviation) and overall accuracy (root-mean-square error) across these distributions and across a range of sample sizes? (3) Under which of these distributional conditions does one coefficient demonstrate a clear performance advantage over the other? It was hypothesized that Pearson's r would outperform Spearman's rho under the symmetric normal and uniform distributions, whereas Spearman's rho would outperform Pearson's r under the right-skewed exponential distribution.

METHODS

This study employed a quantitative, simulation-based comparative design. Rather than relying on a single empirical dataset, a Monte Carlo simulation approach was used so that the population parameters generating the data were known exactly, allowing bias, variability, and accuracy to be evaluated against a true reference value rather than an estimated one. This design follows the tutorial logic used in prior methodological comparisons of correlation coefficients (Jopkins et al., 2024; Ceritoğlu, 2026).

Bivariate samples were generated using a Gaussian copula procedure. For each condition, a pair of standard normal latent variables (Z1, Z2) was drawn from a bivariate normal distribution with a specified latent correlation. These latent variables were transformed to uniform marginals via the standard normal cumulative distribution function and then transformed a second time into the target marginal distribution using its inverse cumulative distribution function (quantile function). Because rank-order (Spearman) correlation is invariant to any monotonic transformation of the marginals, this procedure guaranteed that the population Spearman correlation, ρ_s, was known exactly and held constant across all marginal distributions, while the resulting population Pearson correlation varied naturally as a function of the marginal shape, mirroring the way non-normality distorts the linear correlation in real data without altering the underlying monotonic dependence.

Three marginal distributions were examined, selected to represent a minimal but well-differentiated continuum of distributional shapes commonly encountered in applied research: (a) Normal, a symmetric, mesokurtic baseline consistent with the assumptions underlying Pearson's r; (b) Uniform, a symmetric, bounded, platykurtic

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

distribution with thinner tails than the normal; and (c) Exponential, a right-skewed distribution with a lower bound at zero, representative of the moderately skewed data (e.g., reaction times, income, wait times) frequently encountered in applied research.

Two population associations were examined: $\rho_s = 0$, representing the null hypothesis of no monotonic relationship, and $\rho_s = .50$, representing a moderate monotonic relationship typical of many applied findings. Six sample sizes were crossed with each distribution and population association: n = 10, 30, 50, 100, 300, and 1,000, spanning small, exploratory-study sample sizes through large-sample conditions. For each of the resulting 36 design cells (3 distributions × 2 population associations × 6 sample sizes), 3,000 independent samples were generated, and both Pearson's r and Spearman's rho were computed for each replicate, yielding 108,000 simulated samples in total. The population Pearson correlation for each distribution and association level was estimated from a separate very large reference sample (N = 2,000,000) generated under the same copula procedure, providing a stable benchmark against which the sample estimates were evaluated.

Three criteria were used to evaluate and compare the performance of the two coefficients, consistent with standard Monte Carlo evaluation practice (Soltani & Imani, 2024; Diaz et al., 2022). Bias, calculated as the mean of the sample estimates across replications minus the population Pearson correlation for that condition, indicating systematic over- or under-estimation of the true relationship. Sampling variability, calculated as the standard deviation of the sample estimates across replications, indicating the precision or stability of the estimator.

In addition, Root-mean-square error (RMSE), calculated as the square root of the average squared deviation of the sample estimates from the population value, jointly captures both bias and variability into a single index of overall estimation accuracy. In addition, a variance ratio was computed for each condition; values greater than 1 indicate that Spearman's rho was the more efficient (less variable) estimator, whereas values less than 1 indicate that Pearson's r was more efficient.

RESULTS AND DISCUSSIONS

Table 1 summarizes the mean estimate, standard deviation, root-mean-square error (RMSE), and variance ratio for Pearson's r and Spearman's rho across the three distributions at three representative sample sizes (n = 30, 100, and 1,000), under the moderate population association condition ($\rho_s = .50$). Figure 1 displays the full pattern of sampling variability (standard deviation) across all six sample sizes for each distribution.

Under the Normal distribution, Pearson's r closely tracked the population value at every sample size and showed lower variability than Spearman's rho (variance ratio ≈ 0.86–0.90 across sample sizes), meaning Pearson's r was modestly more efficient, consistent with its status as the maximum-likelihood estimator under bivariate normality. A similar, though smaller, advantage for Pearson's r was observed under the Uniform distribution (variance ratio ≈ 0.94–0.97). Both coefficients were essentially unbiased in these two conditions once sample size reached $n \geq 100$.

Table 1 Bias, Variability, and Accuracy of Pearson's r and Spearman's Rho Across Distributions and Sample Sizes (Population $\rho_s = .50$)

Distribution	n	Population Pearson r	Mean rp (SD)	Mean rs (SD)	RMSE rp	RMSE rs	Variance Ratio (rp/rs)
Normal	30	.519	.515 (.141)	.488 (.149)	.141	.152	0.90
Normal	100	.519	.517 (.074)	.497 (.079)	.074	.082	0.87
Normal	1000	.519	.517 (.023)	.499 (.025)	.023	.032	0.86
Uniform	30	.500	.496 (.143)	.485 (.147)	.143	.148	0.94
Uniform	100	.500	.497 (.075)	.494 (.077)	.075	.077	0.96
Uniform	1000	.500	.500 (.024)	.500 (.024)	.024	.024	0.97
Exponential	30	.470	.468 (.177)	.487 (.148)	.177	.149	1.42

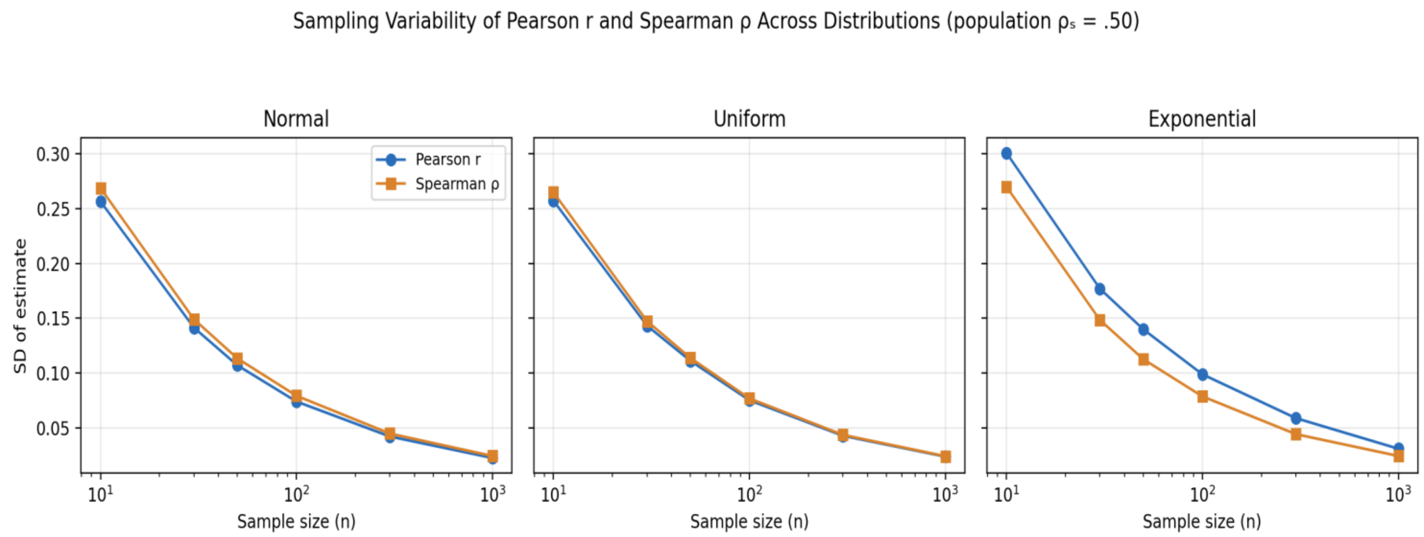
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

Distribution	n	Population Pearson r	Mean rp (SD)	Mean rs (SD)	RMSE rp	RMSE rs	Variance Ratio (rp/rs)
Exponential	100	.470	.471 (.099)	.495 (.079)	.099	.083	1.57
Exponential	1000	.470	.472 (.031)	.499 (.025)	.031	.038	1.61

Note: Values in parentheses under "Mean rp (SD)" and "Mean rs (SD)" are the standard deviations of the 3,000 sample estimates for that condition. RMSE = root-mean-square error relative to the population Pearson correlation. Variance ratio = Var(rp) / Var(rs); values above 1.00 favor Spearman's rho, values below 1.00 favor Pearson's r.

Under the right-skewed Exponential distribution, this pattern reversed: Spearman's rho showed lower variability than Pearson's r at every sample size, with the variance ratio increasing from 1.42 at n = 30 to a peak of 1.61 at n = 1,000. Pearson's r also tended to track the population Pearson value slightly more closely on average than Spearman's rho tracked it, because Spearman's rho converges toward the (numerically larger) population rank correlation rather than the population Pearson correlation; nonetheless, Spearman's rho showed the lower RMSE of the two coefficients at the two largest sample sizes examined (n = 100 and n = 1,000), reflecting its markedly smaller sampling variability under skew.

Figure 1 Sampling Variability (SD) of Pearson's r and Spearman's Rho by Distribution and Sample Size



Note: SD = standard deviation of the sampling distribution across 3,000 replications per cell; population ρ_s = .50. The x-axis is displayed on a logarithmic scale.

The purpose of this study was to compare the statistical performance of Pearson's r and Spearman's rho across three sampling distributions that varied in skewness, kurtosis, and susceptibility to outliers. Consistent with the study's hypothesis and with classical statistical theory, Pearson's r outperformed Spearman's rho in the sense of lower sampling variability, under the normal and uniform distributions, both of which are symmetric. This finding replicates the well-established result that Pearson's r is the more statistically efficient estimator when its parametric assumptions are satisfied (Harary, 2024), and it aligns closely with the simulation findings reported by Okoye & Hosseini (2024), who likewise found that Spearman's rho was somewhat more variable than Pearson's r for normally distributed variables, particularly at moderate-to-strong correlations.

However, this advantage for Pearson's r was not universal. Under the right-skewed exponential distribution, Spearman's rho consistently demonstrated lower sampling variability than Pearson's r, and this advantage grew rather than diminished as sample size increased, reaching a variance ratio of roughly 1.6 to 1 at n = 1,000. This pattern is directionally consistent with, though more modest in magnitude than, prior findings involving more extreme forms of non-normality; for example, Alshangiti et al. (2025) demonstrated the superior robustness of Spearman's rho under outlier contamination, and Rahaman (2026) showed that non-normality inflates the bias and error of Pearson's r while rank-based alternatives largely resist this inflation. The present results extend this literature by showing that even a comparatively moderate departure from symmetry, the exponential distribution, which lacks the extreme kurtosis of heavier-tailed or outlier-contaminated distributions, is sufficient to reverse the efficiency advantage that favors Pearson's r under normality.

This simulation study demonstrated that the comparative statistical performance of Pearson's r and Spearman's rho is conditional on the shape of the sampling distribution from which the data arise, even within a compact set of three distributions. Pearson's r retains a modest efficiency advantage under the normal and uniform distributions, consistent with classical statistical theory. However, Spearman's rho is the more efficient estimator once data depart from symmetry, as demonstrated under the right-skewed exponential distribution, an advantage that widens, rather than narrows, with increasing sample size. These results reinforce the recommendation that researchers examine the distributional properties of their data, particularly skewness, before defaulting to Pearson's r, and that Spearman's rho should be considered a genuinely competitive general-purpose measure of association whenever symmetry cannot be confidently assumed.

REFERENCES

1. Alshangiti, A. M., Gemeay, A. M., Bakr, M. E., Yousuf, A. M., El-Qurashi, M. A., Balogun, O. S., ... & Hammad, A. T. (2025). New robust estimator for handling outliers and multicollinearity in gamma regression model with application to breast cancer data. Scientific Reports, 15(1), 38436.
2. Bocianowski, J., Wrońska-Pilarek, D., Krysztofiak-Kaniewska, A., Matusiak, K., & Wiatrowska, B. (2024). Comparison of Pearson's and Spearman's correlation coefficients for selected traits of Pinus sylvestris L. Biometrical Letters, 61(2), 115-135.
3. Ceritoğlu, F. (2026). Systematic Undercoverage in Bootstrap Confidence Intervals for Xi Correlation Coefficient: A Simulation Study. Mathematics, 14(12), 2136.
4. Dazong, R. H., & Terrang, A. U. (2026). Investigating the Relationship between Pearson's and Spearman Rank Correlation Mathematically and through Simulation. KolaDaisi University Journal of Applied Sciences, 3, 29-34.
5. Díaz, H., Teixeira, A. P., & Soares, C. G. (2022). Application of Monte Carlo and Fuzzy Analytic Hierarchy Processes for ranking floating wind farm locations. Ocean Engineering, 245, 110453.
6. Harary, M. (2024). Efficient algorithms for the sensitivities of the Pearson correlation coefficient and its statistical significance to online data. arXiv preprint arXiv:2405.14686.
7. Hopkins, V., Kagalwala, A., Philips, A. Q., Pickup, M., & Whitten, G. D. (2024). How do we know what we know? Learning from Monte Carlo simulations. The Journal of Politics, 86(1), 36-53.
8. Hurley, L. (2025). Completing and studentising Spearman's correlation in the presence of ties. arXiv preprint arXiv:2512.23993.
9. Ji, X. (2023). Multivariate extreme wind loads: copula-based analysis. Journal of Engineering Mechanics, 149(1), 04022082.
10. Lo, A. A. P., & Christian, C. (2026). Behavioural Predictors of Forward Head Posture Risk: A Correlation, Machine Learning, and Clustering Analysis. Journal of Applied Informatics and Computing, 10(1), 398-405.
11. Okoye, K., & Hosseini, S. (2024). Correlation tests in R: pearson cor, kendall's tau, and spearman's Rho. In R programming: Statistical data analysis in research (pp. 247-277). Singapore: Springer Nature Singapore.
12. Rahaman, A. (2026). Comprehensive classification of statistical methods: A hierarchical taxonomy from foundations to research frontiers. Available at SSRN 6604119.
13. Sathyanarayana, S., & Mohanasundaram, T. (2025). Correlation methodologies revisited: From classical statistics to contemporary residual diagnostics. Asian Journal of Economics, Business and Accounting, 25(11), 86-103.
14. Soltani, A., & Imani, M. A. (2024). Overcoming implementation barriers to renewable energy in developing nations: A case study of Iran using MCDM techniques and Monte Carlo simulation. Results in Engineering, 24, 103213.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

15. Ventura-León, J., Peña-Calero, B. N., & Burga-León, A. (2023). The effect of normality and outliers on bivariate correlation coefficients in psychology: A Monte Carlo simulation. The Journal of General Psychology, 150(4), 405-422.
16. Zrelak, P. A. (2026). Inferential Statistics and the Pitfalls of Nonrandomized Sampling in Nursing Research. Nursing Research, 75(1), 63-69.

R-CODES AND OUTPUTS

```
> set.seed(20260821)
> generate_pair <- function(n, rho_s, dist, rng_extra = NULL)
+ {
+   rho_latent <- 2 * sin(rho_s * pi / 6)
+   Sigma <- matrix(c(1, rho_latent, rho_latent, 1), nrow = 2)
+   L <- chol(Sigma)                      # Cholesky factor for correlated normals
+   Z <- matrix(rnorm(2 * n), ncol = 2) %*% L
+   U <- pnorm(Z)                         # transform to uniform margins
+
+   if (dist == "Normal") {
+     x <- qnorm(U[, 1]); y <- qnorm(U[, 2])
+   } else if (dist == "Uniform") {
+     x <- U[, 1]; y <- U[, 2]
+   } else if (dist == "Exponential") {
+     x <- qexp(U[, 1]); y <- qexp(U[, 2])
+   } else {
+     stop("Unknown distribution: ", dist)
+   }
+   data.frame(x = x, y = y)
+ }
> get_population_pearson <- function(rho_s, dist, N = 2e6)
+ {
+   d <- generate_pair(N, rho_s, dist)
+   cor(d$x, d$y, method = "pearson")
+ }
> distributions   <- c("Normal", "Uniform", "Exponential")
> sample_sizes    <- c(10, 30, 50, 100, 300, 1000)
> target_rho_s    <- c(0, 0.5)
> R               <- 3000
> results <- data.frame()
>
> for (rho_s in target_rho_s) {
```

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

```
+  for (dist in distributions) {
+
+    pop_r <- get_population_pearson(rho_s, dist)
+
+    for (n in sample_sizes) {
+
+      rp <- numeric(R)
+      rs <- numeric(R)
+
+      for (i in 1:R) {
+        d <- generate_pair(n, rho_s, dist)
+        rp[i] <- cor(d$x, d$y, method = "pearson")
+        rs[i] <- cor(d$x, d$y, method = "spearman")
+      }
+
+      bias_rp <- mean(rp) - pop_r
+      bias_rs <- mean(rs) - pop_r
+      sd_rp   <- sd(rp)
+      sd_rs   <- sd(rs)
+      rmse_rp <- sqrt(mean((rp - pop_r)^2))
+      rmse_rs <- sqrt(mean((rs - pop_r)^2))
+      var_ratio <- (sd_rp^2) / (sd_rs^2)
+
+      results <- rbind(results, data.frame(
+        dist = dist, target_rho_s = rho_s, n = n,
+        pop_pearson_r = pop_r,
+        mean_rp = mean(rp), mean_rs = mean(rs),
+        bias_rp = bias_rp, bias_rs = bias_rs,
+        sd_rp = sd_rp, sd_rs = sd_rs,
+        rmse_rp = rmse_rp, rmse_rs = rmse_rs,
+        var_ratio_rp_over_rs = var_ratio
+      ))
+
+      cat(sprintf("dist=%-11s rho_s=%.1f n=%4d | pop_r=%.3f  rp=%.3f(%.3f) rs=%.3f(%.3f)  VR=%.2f\n",
+                  dist, rho_s, n, pop_r, mean(rp), sd_rp, mean(rs), sd_rs, var_ratio))
+    }
+  }
+ }
```

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

```
dist=Normal      rho_s=0.0 n=  10 | pop_r=-0.001  rp=0.008(0.341) rs=0.008(0.340)  VR=1.01
dist=Normal      rho_s=0.0 n=  30 | pop_r=-0.001  rp=0.004(0.184) rs=0.003(0.182)  VR=1.02
dist=Normal      rho_s=0.0 n=  50 | pop_r=-0.001  rp=0.002(0.144) rs=0.002(0.144)  VR=0.99
dist=Normal      rho_s=0.0 n= 100 | pop_r=-0.001  rp=-0.000(0.102) rs=-0.000(0.101)  VR=1.01
dist=Normal      rho_s=0.0 n= 300 | pop_r=-0.001  rp=0.002(0.057) rs=0.002(0.057)  VR=0.99
dist=Normal      rho_s=0.0 n=1000 | pop_r=-0.001  rp=-0.001(0.032) rs=-0.001(0.032)  VR=0.99
dist=Uniform     rho_s=0.0 n=  10 | pop_r=0.001  rp=0.005(0.341) rs=0.001(0.338)  VR=1.02
dist=Uniform     rho_s=0.0 n=  30 | pop_r=0.001  rp=0.003(0.188) rs=0.003(0.187)  VR=1.01
dist=Uniform     rho_s=0.0 n=  50 | pop_r=0.001  rp=0.005(0.141) rs=0.005(0.141)  VR=1.00
dist=Uniform     rho_s=0.0 n= 100 | pop_r=0.001  rp=0.001(0.100) rs=0.001(0.100)  VR=1.00
dist=Uniform     rho_s=0.0 n= 300 | pop_r=0.001  rp=0.001(0.058) rs=0.001(0.058)  VR=1.00
dist=Uniform     rho_s=0.0 n=1000 | pop_r=0.001  rp=0.000(0.031) rs=0.000(0.031)  VR=1.00
dist=Exponential rho_s=0.0 n=  10 | pop_r=-0.000  rp=0.005(0.339) rs=0.005(0.335)  VR=1.03
dist=Exponential rho_s=0.0 n=  30 | pop_r=-0.000  rp=-0.000(0.188) rs=0.002(0.188)  VR=0.99
dist=Exponential rho_s=0.0 n=  50 | pop_r=-0.000  rp=-0.004(0.142) rs=-0.001(0.145)  VR=0.96
dist=Exponential rho_s=0.0 n= 100 | pop_r=-0.000  rp=0.001(0.101) rs=0.000(0.101)  VR=0.99
dist=Exponential rho_s=0.0 n= 300 | pop_r=-0.000  rp=-0.002(0.059) rs=-0.002(0.058)  VR=1.03
dist=Exponential rho_s=0.0 n=1000 | pop_r=-0.000  rp=0.001(0.032) rs=0.001(0.031)  VR=1.03
dist=Normal      rho_s=0.5 n=  10 | pop_r=0.518  rp=0.493(0.269) rs=0.452(0.280)  VR=0.92
dist=Normal      rho_s=0.5 n=  30 | pop_r=0.518  rp=0.513(0.137) rs=0.486(0.146)  VR=0.88
dist=Normal      rho_s=0.5 n=  50 | pop_r=0.518  rp=0.514(0.107) rs=0.493(0.113)  VR=0.89
dist=Normal      rho_s=0.5 n= 100 | pop_r=0.518  rp=0.517(0.074) rs=0.497(0.079)  VR=0.87
dist=Normal      rho_s=0.5 n= 300 | pop_r=0.518  rp=0.515(0.042) rs=0.497(0.045)  VR=0.88
dist=Normal      rho_s=0.5 n=1000 | pop_r=0.518  rp=0.518(0.023) rs=0.500(0.024)  VR=0.88
dist=Uniform     rho_s=0.5 n=  10 | pop_r=0.500  rp=0.483(0.271) rs=0.456(0.279)  VR=0.94
dist=Uniform     rho_s=0.5 n=  30 | pop_r=0.500  rp=0.496(0.147) rs=0.487(0.150)  VR=0.96
dist=Uniform     rho_s=0.5 n=  50 | pop_r=0.500  rp=0.498(0.109) rs=0.492(0.111)  VR=0.96
dist=Uniform     rho_s=0.5 n= 100 | pop_r=0.500  rp=0.498(0.078) rs=0.495(0.079)  VR=0.96
dist=Uniform     rho_s=0.5 n= 300 | pop_r=0.500  rp=0.499(0.045) rs=0.498(0.046)  VR=0.96
dist=Uniform     rho_s=0.5 n=1000 | pop_r=0.500  rp=0.500(0.024) rs=0.500(0.024)  VR=0.97
dist=Exponential rho_s=0.5 n=  10 | pop_r=0.472  rp=0.461(0.297) rs=0.454(0.272)  VR=1.19
dist=Exponential rho_s=0.5 n=  30 | pop_r=0.472  rp=0.467(0.178) rs=0.487(0.148)  VR=1.45
dist=Exponential rho_s=0.5 n=  50 | pop_r=0.472  rp=0.473(0.140) rs=0.492(0.112)  VR=1.58
dist=Exponential rho_s=0.5 n= 100 | pop_r=0.472  rp=0.472(0.100) rs=0.496(0.078)  VR=1.63
dist=Exponential rho_s=0.5 n= 300 | pop_r=0.472  rp=0.470(0.059) rs=0.499(0.046)  VR=1.65
dist=Exponential rho_s=0.5 n=1000 | pop_r=0.472  rp=0.470(0.032) rs=0.499(0.024)  VR=1.76
> library(readxl)
> dataset <- read_excel(NULL)
```

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

```
Error: `path` must be a string
> write.csv(results, "results_R.csv", row.names = FALSE)
> print(subset(results, target_rho_s == 0.5,
+            select = c(dist, n, pop_pearson_r, mean_rp, sd_rp,
+                       mean_rs, sd_rs, rmse_rp, rmse_rs, var_ratio_rp_over_rs)))
          dist    n pop_pearson_r   mean_rp      sd_rp   mean_rs      sd_rs
19      Normal   10     0.5178741 0.4928210 0.26858754 0.4524768 0.28012072
20      Normal   30     0.5178741 0.5127304 0.13660573 0.4861267 0.14578231
21      Normal   50     0.5178741 0.5144608 0.10679001 0.4928539 0.11304251
22      Normal  100     0.5178741 0.5173306 0.07416449 0.4970499 0.07929274
23      Normal  300     0.5178741 0.5153307 0.04203946 0.4973246 0.04473782
24      Normal 1000     0.5178741 0.5178342 0.02280855 0.5002408 0.02426178
25     Uniform   10     0.4996806 0.4826918 0.27111021 0.4561778 0.27940897
26     Uniform   30     0.4996806 0.4960508 0.14743209 0.4867298 0.15021392
27     Uniform   50     0.4996806 0.4979409 0.10865623 0.4917135 0.11084923
28     Uniform  100     0.4996806 0.4983834 0.07773258 0.4952911 0.07928792
29     Uniform  300     0.4996806 0.4990410 0.04537621 0.4979712 0.04624188
30     Uniform 1000     0.4996806 0.5000548 0.02372505 0.4996924 0.02408900
31 Exponential   10     0.4715453 0.4608476 0.29726367 0.4535636 0.27237043
32 Exponential   30     0.4715453 0.4674033 0.17799096 0.4866980 0.14769723
33 Exponential   50     0.4715453 0.4727792 0.14037769 0.4923203 0.11174378
34 Exponential  100     0.4715453 0.4723364 0.09989249 0.4957204 0.07834747
35 Exponential  300     0.4715453 0.4702942 0.05851069 0.4985183 0.04550741
36 Exponential 1000     0.4715453 0.4701994 0.03194653 0.4987094 0.02409249
      rmse_rp    rmse_rs var_ratio_rp_over_rs
19 0.26970888 0.28760785            0.9193508
20 0.13667978 0.14917539            0.8780680
21 0.10682676 0.11575992            0.8924373
22 0.07415412 0.08196884            0.8748330
23 0.04210934 0.04922488            0.8830080
24 0.02280478 0.02998950            0.8837913
25 0.27159688 0.28272928            0.9414799
26 0.14745220 0.15074622            0.9633047
27 0.10865205 0.11111675            0.9608240
28 0.07773044 0.07939614            0.9611519
29 0.04537316 0.04626576            0.9629097
30 0.02372404 0.02408499            0.9700108
31 0.29740658 0.27291805            1.1911426
32 0.17800949 0.14844799            1.4522828
33 0.14035971 0.11364026            1.5781544
34 0.09987898 0.08197997            1.6256076
35 0.05851431 0.05289402            1.6531269
36 0.03196955 0.03630624            1.7582647
```