Fall 2019

Estimation of Correlation Confidence Intervals via the Bootstrap: Non-Normal Distributions

John Mart V. DelosReyes
Old Dominion University, johnmart93@gmail.com

ESTIMATION OF CORRELATION CONFIDENCE INTERVALS

VIA THE BOOTSTRAP:

NON-NORMAL DISTRIBUTIONS

by

John Mart V. DelosReyes
B.A., May 2015, Old Dominion University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
December 2019

Approved by:

Miguel A. Padilla (Director)

James M. Henson (Member)

Jing Chen (Member)

ABSTRACT

ESTIMATION OF CORRELATION CONFIDENCE INTERVALS
VIA THE BOOTSTRAP:
NON-NORMAL DISTRIBUTIONS

John Mart V. DelosReyes
Old Dominion University, 2019
Director: Dr. Miguel A. Padilla

Estimating confidence intervals (CIs) for the correlation has been a challenge. The challenge stems from the metamorphic nature of the sampling distribution of the correlation being bound by $-1 \leq \rho \leq 1$. The nonparametric nature of the bootstrap makes it a good option for estimating correlation CIs. However, there have been mixed results about the robustness of bootstrap CIs for the correlation with non-normal data. This had led the literature to suggesting the use of transformation methods to estimate correlation CIs. However, transformation methods carry a risk of the original data being misrepresented. Thus, further investigation of bootstrap CIs for the correlation is necessary to provide pertinent information in choosing a correlation CI.

Here, the coverage probability of non-bootstrap and bootstrap CIs for the correlation are investigated. This was done with a simulation that has condition parity with previous research yet expands upon these conditions. The non-bootstrap CIs investigated were the Fisher z-transformation, Spearman rank-order, and ranked inverse normal (RIN) Transformation. The bootstrap CIs investigated were the percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). All CIs were assessed for 95% coverage probability and the corresponding correlation estimates were assessed with standardized bias. The PB and BCa CIs were the focus of the study and were found to have good coverage probability overall.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Scientific research is conducted to create a deeper understanding of a discipline for its advancement. In the behavioral sciences, this means creating theories that allow for a better understanding of human behavior and potentially making it predictable. A way to find support for the validity of these theories is to demonstrate consistent result replication across research. However, a collaboration of researchers that attempted to replicate 100 scientific studies found a less than 50% replication rate (Collaboration, 2015). Specifically, the replication rate was 36% for studies that used *p*-values and 47% for studies that used effect sizes with confidence intervals (CIs). These results are far from ideal and do not garner much confidence.

There are many potential causes for the low replication rate in the Collaboration (2015) study; however, the authors suggest that publication, selection, and reporting biases are major contributors. In this sense, the results reflect a research climate that incentivizes research that is "statistically significant," which creates a bias on what kind of research is conducted and how it is reported (Ioannidis, 2005). There have been multiple suggestions to address these issues, but they have been met with some contention and other issues. One suggestion is to further lower the *p*-value threshold for declaring statistical significance to curb researchers from gaming the system for specific outputs (Benjamin, 2017). However, technological advancements have made gathering large samples much easier than what was possible in the past. As such, lowering the *p*-value criteria for statistical significance may not be enough. Another approach would be to instead directly address the tools and methods used.

An upfront suggestion from this approach would be to use effect sizes in addition to statistical significance (Cohen, 1990). As shown from the 36% to 47% replication rate jump in the Collaboration (2015) study, this latter approach shows promise. Even so, this suggestion can further be enhanced by including confidence intervals (CIs) for effect sizes to provide precision information.

Along this line of discussion, a popular statistic that can also play the role of an effect size is the correlation. In fact, the correlation was the effect size measure used in the majority of studies in the Collaboration (2015) study. The correlation can take on several forms, but the original Pearson product-moment correlation (henceforth referred to as the correlation) is the most iconic with over 100 years of use (Hald, 2007). Additionally, the correlation is foundational to many other statistics and statistical models like the *t*-test, regression, MANOVA, etc. Given the ubiquity of the correlation, its importance to other statistics, and its use as an effect size, CI research about the correlation is essential. This is especially true if effect sizes are being proposed as a method to help with research replication.

However, CI research about the correlation is limited with contradictory results. Some research shows that the correlation and its CIs are robust to non-normal data and others demonstrate the opposite. The focus on non-normal data is important as there is evidence to suggest that non-normal data is common (Blanca et al., 2013). Furthermore, much of the research is focused on the null hypothesis of when the correlation is zero $(\rho = 0)$. This kind of information is not entirely relevant to situations that call for the use of the correlation as an effect size; i.e., situations where the correlation does not equal zero $(\rho \neq 0)$. Therefore, the research here will address this gap by investigating bootstrap CIs for the correlation when $\rho = 0$, $\rho \neq 0$, and non-normal data. This will also include an exploration of a CI method that has not been previously explored. However, this discussion begins with concept of “statistical significance” and its relation to CIs.

Null Hypothesis Significance Testing

The most prevalent form of hypothesis testing is null hypothesis significance testing (NHST). In fact, the 36% replication rate for *p*-values of studies in the Collaboration (2015) study is based on NHST. Classically and at its simplest, NHST is conducted by setting up two opposing hypotheses: the null (H_0) and alternative (H_A) hypotheses. In NHST, H_0 is assumed to be true (e.g., no effect or no relationship), but data is used to test this claim (Perezgonzalez, 2016). This process requires selecting a criterion (α) that establishes a low probability threshold within H_0 occurring and compares this criterion to the probability of obtaining an effect (*p*-value) from the data. Within this context, α is the probability of rejecting a true H_0 (i.e., the probability of a type I error). If the *p*-value is less than α, then H_0 is rejected in support of H_A (e.g., the effect is statistically significant). As such, the principle idea of statistical significance is to demonstrate that obtaining an observed effect would be highly unlikely if H_0 is true (Fisher, 1929).

Since NHST only establishes the probability of obtaining an effect when assuming that H_0 is true, it restricts *p*-values to only being able to infer the compatibility of an effect to H_0. A common misconception is that the *p*-value is the probability of H_0 being true given an effect (Cohen, 1994). This confusion is further compounded by the misinterpretation that the complement of the *p*-value $(1-p)$ is the probability of H_A being true (Nickerson, 2000). This has led to inaccurate reporting of *p*-values such that there is a belief that a $p < \alpha$ is associated with a greater than $100(1-\alpha)\%$ chance of H_A being true. For example, a $p < .05$ is interpreted as a greater than 95% chance of H_A being true. While these misinterpretations may not be directly caused by NHST, they illustrate the downward spiral of confusion that results from misunderstanding a tool (e.g., NHST in this context).

This is not to say that NHST is limited strictly due to misconceptions or less than noble intentions. Another way to think about NHST is that it only informs researchers how well an effect relates to a hypothesis in terms of probability. For example, say that a test for mean difference has a $p = .03$. If $\alpha = .05$, then the test is statistically significant. However, if $\alpha = .01$, the test is not statistically significant. Thus, the results of NHST can change depending on the α level used to analyze an effect (Wasserstein & Lazar, 2016). Additionally, H_0 is assumed to be equal to a specific value (e.g., $H_0 : \mu_1 - \mu_2 = 0$), and any deviation from this value would demonstrate that H_0 is not true.

Consider a test of mean difference (i.e., the effect) between a group of participants before and after a treatment where H_0 is stated as "There is no mean difference before and after a treatment" (i.e., $\mu_D = 0$) and $\alpha = .05$. The descriptive statistics for this example are: $n = 10$, mean difference is $M_D = 3$, and the variance of the difference is $S^2 = 225$. A related-samples *t*-test of this data would be

$$t = \frac{M_D}{\sqrt{\frac{S_D^2}{n}}} = \frac{3}{\sqrt{\frac{225}{10}}} = 0.6325 \tag{1.1}$$

where the denominator is the standard error (SE). This effect is not statistically significant at $\alpha = .05$ which has $t_{crit} = 2.2622$. If this example were modified by increasing *n* to 100 (i.e., $n = 100$), the new result would be

$$t = \frac{M_D}{\sqrt{\frac{S_D^2}{n}}} = \frac{3}{\sqrt{\frac{225}{100}}} = 2. \tag{1.2}$$

This is statistically significant at $\alpha = .05$ which now has $t_{crit} = 1.9842$. This demonstrates that by increasing the sample size, statistical significance can be obtained by way of decreasing the associated critical value. Note that this occurred despite no change in the mean difference (i.e., the effect) or the variance of the difference. Although presented in an idealistic manner, sample size can impact statistical significance in other ways. For example, increasing the sample size typically decreases variance, which in turn again increases statistical significance. Even so, statistical significance can be obtained by simply increasing the sample size because it decreases the corresponding critical value; a relationship that holds for any form of NHST. This is key as modern technological advancements make obtaining larger sample sizes, and by extension statistical significance, easier than what was possible in the past. Furthermore, the results of statistical significance do not inform in terms of precision and a standard (or common) scale, which makes interpretations of NHST results even more cumbersome. To address the precision limitations of statistical significance, interval estimation should be used about the parameter of interest in NHST.

Interval Estimation and Confidence Intervals

There are two general methods for estimating a parameter: point and interval. A point estimate uses a single value for a parameter estimate. Examples of point estimates are the mean and variance for the difference (e.g., M_D and S_D^2). A point estimate is useful for establishing a best guess for a parameter but does not account for sampling error (or variability) about the parameter estimate. However, an interval estimate consists of a range of possible values that are likely to contain the population parameter. Interval estimates manifest this range by accounting for error to provide precision information about estimating the parameter of interest.

A common application of interval estimation is the confidence interval (CI). Assuming a random sample from a given distribution, a CI is formed by creating an error structure around a point estimate based on the standard error (*SE*) of the corresponding distribution with a confidence level (Hogg, Tanis, & Zimmerman, 2015). A confidence level indicates the consistency of a parameter estimate and is denoted as $100(1-\alpha)\%$ where α represents the probability of making a type I error. In the context of a CI with $\alpha = .05$, a 95% CI indicates an expectation that 95% of all CIs created in the same way will contain the corresponding population parameter.

Besides providing information about the precision of a parameter estimate, CIs also have utility in hypothesis testing. In these cases, statistical significance is met when the parameter specified under H_0 is not within the bounds of the CI. Conversely, having a CI that contains the parameter for H_0 indicates support for H_0.

A CI for a related-samples *t*-test can be constructed as follows:

$$P\left(-t_{\alpha/2} \le \frac{M_D}{\sqrt{\frac{S_D^2}{n}}} \le t_{\alpha/2}\right) = 1-\alpha, \qquad (1.3)$$

$$\left[M_D - t_{\alpha/2}\left(\sqrt{\frac{S_D^2}{n}}\right), M_D + t_{\alpha/2}\left(\sqrt{\frac{S_D^2}{n}}\right)\right], \qquad (1.4)$$

$$M_D \pm t_{\alpha/2}\left(\sqrt{\frac{S_D^2}{n}}\right), \qquad (1.5)$$

and

$$df = n - 1 \tag{1.6}$$

where $t_{\alpha/2}$ refers to the *t*-critical value associated with probability of making a type I error divided by 2. The division of α by 2 is used in this context to create a CI with endpoints around the estimate.

Continuing with the previous two related-samples *t*-test examples. The first example has $n = 10$ with $df = 9$ and the corresponding 95% CI using equation 1.5 is estimated as

$$3 \pm 2.262\left(\sqrt{\frac{225}{10}}\right) \tag{1.7}$$

and $[-7.7296, 13.7296]$. The second example has $n = 100$ with $df = 99$ and the corresponding 95% CI using equation 1.5 is

$$3 \pm 1.984\left(\sqrt{\frac{225}{100}}\right) \tag{1.8}$$

and $[0.0240, 5.9760]$. In the first case, since the CI includes H_0 (i.e., $\mu_D = 0$), the result is not significant. In the second case, the CI does not include H_0 and is significant. In terms of statistical significance, these CIs provide the same information as the original examples. However, because CIs incorporate sampling error through the *SE*, precision information regarding the parameter is provided (μ_D in this case). Nevertheless, the limitation with respect to replication is that CIs are not in a standard (or common) scale because the parameter (or effect) is not in standard scale.

NHST only establishes if a parameter (or effect) is significant with respect to a hypothesis, and a CI provides the same information but adds precision information about the parameter (or effect). Neither lends credence as to whether the effect is meaningful; only that an effect is detected and how well it is estimated. When statistical significance is found under NHST, effects of $1, -10, .001,$ etc. can all hold equal value to the result (Tukey, 1991). This is because there is no sense of scale inherit to NHST due to the binary nature of the result; it only matters that the effect is different from what is stated in H_0. A CI only adds precision information about the effect; i.e., it gives information about how well the parameter is estimated. So while an effect of $.001$ may be statistically significant and may be estimated well, it is difficult to judge if this effect has more real-world weight compared to similar effects from different research (i.e., practically significant). These limitations in clarity demonstrate that other methods should be used that can better demonstrate the utility of an effect. This leads to effect sizes.

Effect Sizes

An effect size is a statistic that measures the magnitude of an effect. The utility of an effect size over NHST is that it explains an effect in terms of a standard (or common) scale rather than the original units of measurement (Cohen, 1988). This means that regardless of the statistical significance of an effect and/or how well it is estimated, the effect can be easily and consistently understood. For example, an effect size that ranges from 0 to 1 can be understood to have greater effect as the effect approaches 1. Furthermore, this process can be extended to compare studies in terms of effect size magnitudes. The utility has been acknowledged and has led to the consideration of reporting effect sizes as a standard (Wilkenson, 1999).

Consider again the previous two mean difference examples. One kind of effect size appropriate here is Cohen's *d* (Cohen, 1988). Cohen suggests the following guidelines for judging the magnitude of *d*: 0.20, 0.50, and 0.80 are small, medium, and large effects, respectively. In this situation, the effect size is

$$d = \left| \frac{M_D}{\sqrt{S_D^2}} \right| = \left| \frac{3}{\sqrt{256}} \right| = 0.1875 \tag{1.9}$$

which is considered a small effect. The key thing to notice is that, regardless of statistical significance and precision, both examples have the same effect size. This occurs because Cohen's *d* only compares the ratio of the effect to error (or variance) without considering sample size. Unlike statistical significance and precision, sample size has little or no impact on effect sizes (Sullivan & Richard, 2012). Therefore, a strict reliance on "statistical significance" does not provide enough information to make proper conclusions about an effect and the example highlights the difference between statistical and practical significance. However, this issue can be alleviated through effect size use.

At this point, it is clear why some researchers suggest using effect sizes in additional to statistical significance (Cohen, 1990). An effect size is less impacted by sample size and puts the effect on a standard (or common) scale. Additionally, they can be used to compare the results from different studies (e.g., meta-analysis). As such, effect sizes help provide a more complete picture for the results found in hypothesis testing. However, effect sizes would have more utility if they had precision information. This can be achieved by forming effect size CIs and would allow researchers to know how well their effects are estimated. Of current contention related to this are CIs about the correlation as CI research regarding the correlation is limited and mixed.

The Pearson Correlation as an Effect Size

As aforementioned, the Pearson product-moment correlation (or correlation for short) is a popular statistic that can also play the role of an effect size. The popularity and importance of the correlation stems from its over 100 years of use and its contribution as a foundational piece to other statistics and statistical models like the *t*-test, regression, MANOVA, etc. (Hald, 2007). Additionally, the correlation was the effect size used in most of the replication studies in the Collaboration (2015) study. Like any effect size, sample size has little to no impact on the correlation and a CI for it would also let researchers know how well the correlation was estimated. However, CI research for the correlation is sparse with varied results. Therefore, the research here will address this gap by evaluating non-bootstrap and bootstrap CIs for the correlation.

The correlation is a method used to measure the linear relationship between two variables. In population settings, the correlation for variables *x* and *y* is defined as

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \tag{1.10}$$

where, and σ_{xy} is the covariance between *x* and *y*, and σ_x and σ_y are the standard deviations for the *x* and *y*, respectively. When dealing with samples, the parameters are replaced with their respective estimates and equation 1.10 becomes

$$r_{xy} = \frac{s_{xy}}{s_x s_y} \tag{1.11}$$

With this structure, it can be understood that the correlation is a ratio between how the *x* and *y* variables vary together. Another way to think of this is how much the variation in *x* is related to variation in *y* and vice versa.

The correlation has a range of $-1 \le \rho \le 1$. This results in a standard (or common) scale where a correlation of 1 indicates a perfect-positive relationship, -1 a perfect-negative relationship, 0 no relationship, and anything in-between as some variation of the previous interpretations. In this context, "perfect-positive" indicates that as one variable increases the other variable will invariably match that increases in the same direction while a "perfect-negative" indicates that as one variable increases the other will invariably match that increase in the opposite direction. Cohen (1988) suggested the following guidelines for judging the magnitude of *r*: ± 0.10, ± 0.30, ± 0.50 are small, medium, and large effects, respectively. The simplicity in implementation and interpretation lends to the correlation's popularity and allows the correlation to serve as a scaffolding piece to other statistics like the coefficient of determination.

The coefficient of determination $\left(r^2\right)$ is another effect size directly related to correlation as it is just the correlation squared and is defined as

$$\rho^2 = \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} \tag{1.12}$$

or

$$r^2 = \frac{s_{xy}^2}{s_x^2 s_y^2} \tag{1.13}$$

depending on whether populations or samples are used, respectively. Like the correlation, r^2 can be thought of as a proportion. In this case, it is the proportion of variance in either variable that is accounted for by the other (Cohen, 1988). This concept can be extended to linear models. For example, r^2 is the amount of variability in the outcome (y) accounted for by a linear regression model. The purpose of making this connection is to note that the correlation is integral to other statistics and that correlation's qualities and interpretations will propagate to those other statistics that are dependent on the correlation.

To summarize, replication of research findings is a concern in scientific research (Collaboration, 2015). In response, one suggestion is to lower the *p*-value threshold for declaring statistical significance (Benjamin, 2017). However, modern technological advancements (e.g., statistical computing, online surveys, etc.) have made gathering large amounts of data easy. Therefore, in addition to reporting statistically significant findings, researchers have also suggested reporting effect size measures as they are minimally impacted by sample size (Cohen, 1990). This suggestion shows promise as effect sizes had a higher replication rate than statistical significance in a collaborative study (Collaboration, 2015). Even so, this suggestion can further be enhanced by including CIs for effect sizes to provide effect size precision information. One common effect size is the correlation (ρ) and was in fact the effect size used in the majority of the studies in the collaborative study. However, CI research about the correlation or its robustness is limited, and in some cases conflicting. Research regarding the robustness of correlation is now discussed.

Robustness of the Correlation *t*-test

Early research regarding the robustness of the correlation primarily focused on the application of the *t*-test for testing $\rho = 0$. An early study of this kind of robustness was done by Blair and Lawson (1982). The authors main contention with previous work was its limitations to familiar distribution shapes that may not be reflective of "actual research contexts" where non-normal data is more common (Kowalski, 1972; Norris & Hjelm, 1961; Pearson & Adyanthaya, 1929). As such, the authors investigated a severely non-normal distribution (i.e., the Bradley distribution with skew = 3 and kurtosis = 17). For distribution shape, the distribution was the same for both variables. This simulation study investigated the impact of (a) distribution shape (normal and Bradley) and (b) sample size $(n = 5, 30, 50, 100)$ on the type I error rate of the correlation *t*-test for testing $\rho = 0$. Results were based on 5,000 simulation replications and $\alpha = .005, .01, .025, \text{and } .05$. The results indicated that the correlation *t*-test type I error rate was accurate under the normal distribution but had inflated type I error rate under the Bradley distribution. Additionally, increasing the sample size did not help and appeared to make the situation worse for the Bradley distribution. Results across α were generally consistent.

In a subsequent study, Edgell and Noon (1984) examined the type I error rate for the correlation *t*-test when testing $\rho = 0$. This was done by investigating (a) distribution shape, (b) distribution pairing, and (c) sample size $(n = 5, 10, 15, 20, 30, 50, 100)$ for $\alpha = .01 \text{ and } .05$. The primary interest of this research was to determine how the correlation *t*-test responds to non-normal distributions. The distributions investigated were:

- Normal (skewness = 0, kurtosis = 0)
- Uniform (skewness = 0, kurtosis = -1.2)
- Exponential (skewness = 2.07, kurtosis = 6.56)
- Cauchy Form 1 (skewness = 21.56, kurtosis = 2171)
- Cauchy Form 2 (skewness = 49.8, kurtosis = 2817.9)

These distributions were then either paired with themselves or with another distribution (i.e., all pairwise combinations) resulting in a total of 15 distribution pairing being explored. Results were based on 10,000 simulation replications. The results of this study showed that the *t*-test for testing $\rho = 0$ is robust at controlling type I error at $\alpha = .05$ and when $n \geq 5$. However, it was not robust for extreme distributions (e.g., Cauchy) at $\alpha = .01$.

Early research on the robustness of the correlation is limited and mixed. According to some of the research, the correlation *t*-test is robust when testing that the population correlation is zero $(\rho = 0)$. However, some of the research indicates that this is not the case when working with severely non-normal data. This is concerning as data typically encountered in research environments is non-normal (Blanca et al., 2013). Furthermore, even if the research showed clear evidence that the correlation *t*-test is robust to distributional assumptions when testing that $\rho = 0$, these results would bear little information about estimating the correlation when $\rho \neq 0$. When the correlation hypotheses are of the form H_0: $\rho = 0$ vs. H_A: $\rho \neq 0$, rejecting H_0 only indicates that $\rho = 0$ is an unlikely event and provides no information about estimating $\rho \neq 0$. As such, if using the correlation as an effect size is of interest, then testing $\rho = 0$ has little utility. It would be of greater interest to investigate the correlation across its range $(-1 \leq \rho \leq 1)$. A way to investigate this is through CIs for the correlation.

Confidence Interval Research About the Correlation

One early attempt to investigate estimating correlation CIs when $\rho \neq 0$ was the Fisher *z*-transformation (Fisher, 1915). The Fisher *z*-transformation is defined as

$$z = \left(\frac{1}{2}\right)\ln\left(\frac{1+r}{1-r}\right). \quad (1.14)$$

This allows for a $100(1-\alpha)\%$ CI to be defined as

$$z \pm z_{\alpha/2} \times SE(z), \quad (1.15)$$

where

$$SE(z) = \frac{1}{\sqrt{n-3}} \quad (1.16)$$

is the estimated standard error of *z*.

Zeller and Levine (1974) evaluated the performance of the Fisher z-transformation CI for the correlation from equations 1.14 to 1.16 under several simulation conditions. The authors investigated the (a) distribution shape, (b) correlation strength $(\rho = 0, .32, .71, .95)$, and (c) sample size $(n = 15, 50, 100)$ for $\alpha = .01$ and $.05$. For distribution shape, the distribution was the same for both variables and the authors investigated the normal, uniform, J, bimodal, and a leptokurtic, but provided no skewness and kurtosis details for these distributions. Results were based on 3,000 simulation replications. One consistent finding was that the estimate of the correlation from equation 1.11 slightly underestimated the true correlation, but this was ultimately negligible when $n > 15$. In addition, the correlation CI from equation 1.15 was shown to be robust to the mild non-normal distributions investigated (e.g., uniform, J, bimodal, and leptokurtic). These results were consistent for $\alpha = .01$ and $.05$.

Additional research on the robustness of the Fisher z-transformation CI for the correlation was conducted by Berry and Mielke (2000). The authors investigated (a) distribution shape, (b) correlation strength $(\rho = 0, .4, .6, .8)$, and (c) sample size $(n = 10, 20, 40, 80)$ for $\alpha = .10$ $.05$, and $.01$. For distribution shape, the distribution was the same for both variables and the authors investigated were the normal, 3 generalized logistic, and 3 symmetric kappa distributions. The generalized logistic distributions were defined by

$$f(x) = \left(\frac{e^{\theta x}}{\theta}\right)^{1/\theta} \left(1 + \frac{e^{\theta x}}{\theta}\right)^{-(\theta+1)/\theta} \tag{1.17}$$

where $\theta = 1, .1, .01$. In this context, $\theta > 1$ results in negative skew and $\theta < 1$ results in positive skew. The symmetric kappa distribution was defined by

$$f(x) = .5\lambda^{-1/\lambda}\left(1+\frac{|x|^{\lambda}}{\lambda}\right)^{-(\lambda+1)/\lambda} \quad (1.18)$$

where $\lambda = 2, 3, 25$. In this context, $\lambda = 2$ represents a distribution similar to a *t* distribution with 2 degree of freedom, $\lambda = 3$ a distribution with heavy tailed distribution, and $\lambda = 25$ is similar to a uniform distribution with thin tails. The authors provided no skewness and kurtosis details for these distributions. Results were based on 1,000,000 simulation replications. The results showed that Fisher z-transformation CIs have appropriate coverage probability when $\alpha = .05$ and $\rho = 0$ for all distribution shapes. However, the coverage probability was consistently underestimated when distribution shapes are non-normal and $\rho \neq 0$. Furthermore, these problems were not remedied with increased sample size but made more severe. These results were consistent for $\alpha = .01 .05$, and $.10$ This discrepancy in the literature about the efficacy of correlation CI methods indicate the need for additional research on the matter.

The issue with correlation CI research is the $-1 \leq \rho \leq 1$ range restriction on the correlation (See Figure 1; Blanca et al., 2013). This range restriction makes estimating the correlation when $\rho \neq 0$ difficult as the correlation becomes more skewed as it approaches ±1. In fact, the only time the correlation is symmetric is when $\rho = 0$. As such, it is apparent that estimating the correlation CIs is difficult when $\rho \neq 0$. However, the bootstrap is a powerful method that can be used for CI estimation of the correlation without making distributional assumptions.

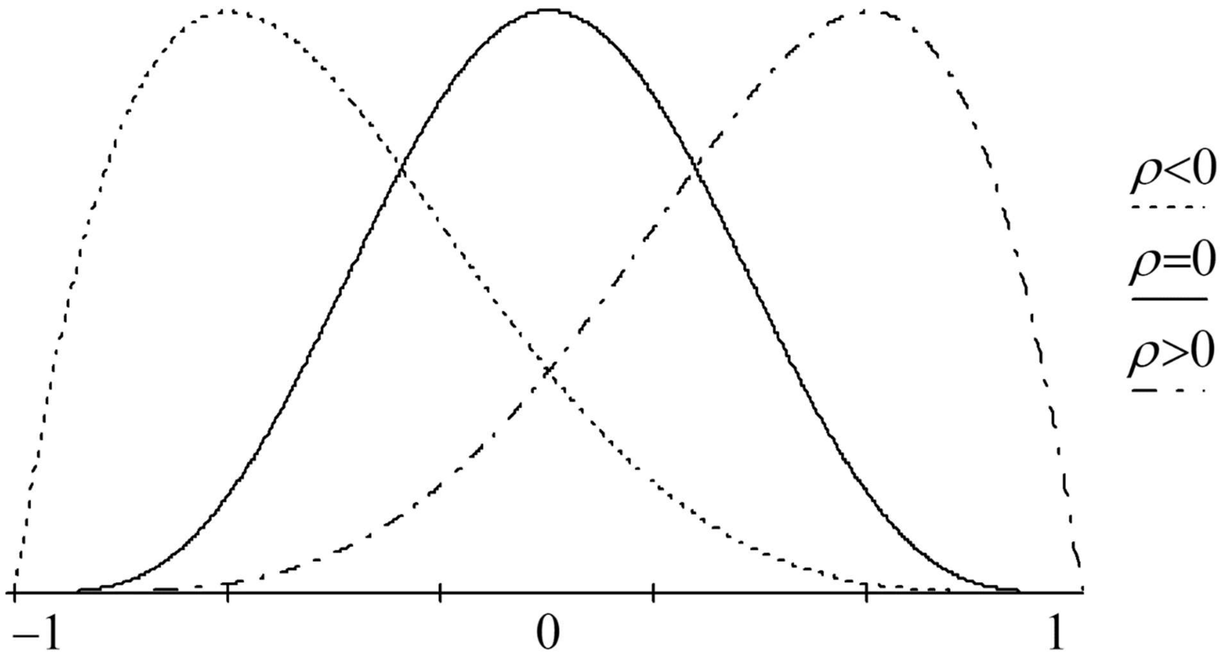

Figure 1. The range restriction effect on the distribution of the correlation.

Bootstrap Method

Bootstrapping is a statistical technique that uses data to create simulations for statistical inference (Efron & Tibshirani, 1993). This method generates the sampling distribution of a parameter estimate through sampling with replacement from a sample. This generated sampling distribution can be used to obtain *SE*s, CIs, and do hypothesis testing. The bootstrap method has two distinct advantages over parametric methods. First, it is a robust alternative used when parametric based inference is in doubt (e.g., the normality assumption is in doubt). Second, bootstrapping is a good alternative when parametric inference is impossible or requires very complicated equations for calculating *SE*s (e.g., the sampling distribution is unknown).

An important detail about the bootstrap method is that it can be used to estimate the sampling distribution of almost any statistic without any prior knowledge of its sampling distribution. To understand this, consider the Central Limit Theorem (CLT) and the mean (μ). The CLT states that regardless of the population distribution of a sample, the sampling distribution of the sample mean (*M*) will approach normality as *n* approaches infinity. In this case, the bootstrap method can simulate the CLT by providing an estimate of the sampling distribution of the statistic (i.e., the empirical sampling distribution or ESD). This also allows for *SE* estimates (i.e., the standard deviation of the sampling distribution) once the ESD is generated. This means that the bootstrap method can be used as a brute force way to estimate the *SE* of almost any statistic.

Consider an example where a sample *x* is obtained of size *n* and the mean is of interest. This sample is defined as $x = (x_1, x_2, ..., x_n)$. The bootstrap is then performed in three steps. First, obtain the b^{th} random sample with replacement from *x*; i.e., $x^{(b)} = \left(x_1^{(b)}, x_2^{(b)}, ..., x_n^{(b)}\right)$. Second, compute and store the b^{th} estimate of the mean $\left(M^{(b)}\right)$ from $x^{(b)}$. Third, compile the stored estimates $M^{(1)}, M^{(2)}, \ldots, M^{(B)}$ to create the ESD of *M* for the $b = 1, 2, ..., B$ bootstrap samples. The bootstrap estimate of the *SE* is defined as

$$SE(M) = \sqrt{\frac{\sum_{i=1}^{B}\left(M^{(b)} - \bar{M}\right)^2}{B-1}} \tag{1.19}$$

where

$$\bar{M} = \frac{1}{B}\sum_{b=1}^{B} M^{(b)} \tag{1.20}$$

is the mean of the ESD. The implication of this example is that this process can be extended to other statistics without knowledge of the corresponding sampling distributions (e.g., the correlation).

As aforementioned, the bootstrap method can be used to estimate the *SE* and that allows for CI estimation. Using the previous example, a CI for the mean can be estimated by

$$M \pm z_{\alpha/2} SE\left(M\right). \tag{1.21}$$

Note the similarities between equations 1.5 and 1.21. A limitation of equation 1.21 is that it requires knowledge of the sampling distribution; in this case, knowledge that the sampling distribution is normal. The bootstrap overcomes this limitation by offering two alternative CIs that do not require this knowledge; the percentile and bias-corrected and accelerated CIs.

Percentile Bootstrap CI

The percentile bootstrap (PB) CI is a distribution-free method for constructing bootstrap CIs based on the percentiles of the ESD. In the context of the previous mean example, the PB CI is defined as

$$\left[M_B^{(\alpha/2)}, M_B^{(1-\alpha/2)}\right] \tag{1.22}$$

where $M_B^{(\alpha/2)}$ and $M_B^{(1-\alpha/2)}$ are the α/2 and $1-\alpha/2$ percentiles from the ESD and α is the probability of type I error. For example, with 1000 bootstrap samples and $\alpha = .05$, 25 and 975 would serve as the percentiles of the ESD. The utility of the PB CI is that it is easy to understand and transformation respecting. This means that for a CI of $\left[M_B^{(\alpha/2)}, M_B^{(1-\alpha/2)}\right]$ for μ, a transformation $t\left(\mu\right)$ will have a corresponding CI of $\left[t\left(M_B^{(\alpha/2)}\right), t\left(M_B^{(1-\alpha/2)}\right)\right]$.

Bootstrap Bias-Corrected and Acceleration CI

Another distribution-free method for creating bootstrap CIs is the bias-corrected and accelerated (BCa) method. This is like the PB CI method but also adjusts for bias and acceleration. In this context, bias refers to the discrepancy between the bootstrap statistics and the corresponding sample statistic and acceleration refers to skew. If bias and acceleration are non-issues, then the PB and the BCa CIs methods yield similar results. This method also benefits from being transformation respecting but is also first and second order accurate. First and second order accurate means for a sample size *n*, the CI will have error that tends to zero at a rate of $1/\sqrt{n}$ and $1/n$, respectively.

Continuing with the mean example, the BCa CI for *M* is defined as

$$\left[M^{(\alpha_1)}, M^{(\alpha_2)}\right] \tag{1.23}$$

where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{(\alpha)}}{1 - \hat{a}\left(\hat{z}_0 + z_{(\alpha)}\right)}\right), \tag{1.24}$$

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{(1-\alpha)}}{1 - \hat{a}\left(\hat{z}_0 + z_{(1-\alpha)}\right)}\right), \tag{1.25}$$

Φ is the standard normal cumulative distribution function, $\hat{z}_0$ is bias, $\hat{a}$ is acceleration, and $z_{(\alpha)}$ and $z_{(1-\alpha)}$ are the percentile cutoffs from the standard normal distribution. Bias is defined as

$$\hat{z}_0 = \Phi^{-1}\left(\frac{1}{B}\sum_{b=1}^{B} I\left(M^{(b)} < M\right)\right), \tag{1.26}$$

where $\Phi^{-1}(.)$ is the inverse standard normal cumulative distribution function, $I(.)$ is the indicator function, and M is the mean of the original data. Acceleration is accounted for by jackknife resampling where data are resampled by removing one observation per resample. Given data $x=(x_1, x_2, ..., x_n)$, the i^{th} jackknife sample for $i = 1, 2, ..., n$ is

$$x_{(i)} = (x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n) \tag{1.27}$$

with the i^{th} data point removed. Acceleration is defined as

$$\hat{a} = \frac{\sum_{i=1}^{n}\left(\bar{M}_{(\cdot)} - M_{(i)}\right)^3}{6\left[\sum_{i=1}^{n}\left(\bar{M}_{(\cdot)} - M_{(i)}\right)^2\right]^{3/2}} \tag{1.28}$$

where

$$\bar{M}_{(\cdot)} = \frac{1}{n}\sum_{i=1}^{n} M_{(i)} \tag{1.29}$$

and $M_{(i)}$ is the mean estimate that excludes the i^{th} data point. Acceleration in this form is simply the skewness multiplied by 1/6 and serves as a skewness correction factor.

An early application of the bootstrap CI for the correlation was presented by Lunneborg (1985). Lunneborg explored the potential of the bootstrap for estimating correlation CIs using SAT verbal and math scores from a pseudorandom sample of 25 college freshman. In this study, the PB CI, based on 500 bootstrap samples, was compared to the Fisher z-transformation CI from equation 1.14 and α was not disclosed. The SAT scores were used because (a) they are real data, (b) the verbal and math scores are known to be bivariate normal, and (c) the verbal-math ρ

is large enough so that useful CIs can be estimated for the small pseudorandom sample $(n = 25)$. The CIs from both methods were similar under bivariate normality.

An early simulation study of the bootstrap CI for the correlation was conducted by Rasmussen (1987). Of interest was the impact of a non-normal distribution on the bootstrap CI for testing $\rho = 0$. This was done by investigating (a) distribution shape (normal and lognormal) and (b) sample size $(n = 5, 15, 30, 60)$ for $\alpha = .01$ and $.05$. For distribution shape, the variable pairings for ρ took the following two forms: normal-normal or normal-lognormal. The PB CI, based on 500 bootstrap samples, was compared to the Fisher z-transformation CI. Results were based on 1,000 simulation replications. The results ran counter to Lunneborg's research (1985) as they showed a lack of parity between the Fisher z-transformation and the PB CI. In this case, the PB CIs demonstrated an overall increase in type I error rate and restrictive CIs compared to the Fisher z-transformation CI under all conditions, including when the variable pairing was normal-normal. Going from $\alpha = .05$ to $\alpha = .01$ further highlights this issue. In addition, Rasmussen noted that the situation did improve as sample size with larger sample sizes but was not able to explore this due to costs in computational power at the time.

In two recent studies, Padilla and Veprinksy (2012, 2014) developed PB and BCa CIs for the deattenuated correlation for $\alpha = .05$. An estimated correlation can become weaker (attenuated) than what may be true in the population due to measurement error (Spearman, 1904). Spearman (1904) developed a correction for this attenuation known as the deattenuated (or disattenuated) correlation (Muchinsky, 1996), but research on this correlation and its corresponding CI is rare. Padilla and Veprinsky (2012, 2014) address this gap by investigating the bootstrap CIs for the deattenuated correlation under four simulation conditions: the (a) distribution shape, (b) strength of the correlation $(\rho = .10, .20, .30, .40, .50)$, (c) reliability of both variables in the correlation $(\rho_{ij} = .50, .60, .70, .80, .90)$, (d) and sample size $(n = 50, 100, 150, 200, 250, 300)$. All bootstrap CIs were based on 2,000 bootstrap samples. For distribution shape, both variables had the same distribution from the following distributions investigated:

- Normal (skewness = 0, kurtosis = 0)
- Uniform (skewness = 0, kurtosis = -1.20)
- Triangular (skewness = 0, kurtosis = -0.60)
- Beta (skewness = -0.85, kurtosis = 0.22)
- Laplace (skewness = 0, kurtosis = 3)
- Pareto (skewness = 2.81, kurtosis = 14.83)

Results were based on 1,000 simulation replications. Overall, the PB and BCa CIs had good coverage under all simulation conditions with negligible differences between the two CIs. Even so, the BCa CI tended to have slightly better coverage than the PB CI. The one exception was that neither CI performed well with the Pareto distribution. However, the Pareto distribution investigated was skewed and highly peaked (kurtosis = 14.83). Such distributions have range restrictions, and it is well known that distributions with range restrictions attenuate the correlation due to less variability.

In a subsequent study, Bishara and Hittner (2017) investigated several CIs for the correlation. Of interest was the impact of various types and combinations of distributions on the correlation CIs. The following correlation CIs were investigated: the 1) Fisher *z*-transformation, 2) Spearman rank-order with Fieller's *SE*, 3) Spearman rank-order with Wright's *SE*, 4) Box-Cox transformation, 5) ranked inverse normal transformation, 6) nonparametric bootstrap, 7) nonparametric bootstrap with asymptotic adjustment (AA), 8) nonparametric bootstrap BCa,

9) observed imposed bootstrap, 10) observed imposed bootstrap with AA, and 11) observed imposed bootstrap with BCa. All bootstrap CIs were based on 9,999 bootstrap samples. The performance of the CIs was investigated through a simulation with the following four conditions: (a) distribution shape, (b) distribution pairing, (d) correlation strength $(\rho = 0, .5)$, and (c) sample size $(n = 10, 20, 40, 80, 160)$ for $\alpha = .05$. The distributions investigated were a result of a combination of population skewness $(\gamma_1 = -4, -3, -2, -1, 0, 1, 2, 3, 4)$ and kurtosis $(\gamma_2 = -1, 0, 2, 4, 6, 8, 10, 20, 30, 40)$ whose feasibility was limited by the lower bound of kurtosis being determined by the squared skewness

$$\gamma_2 \geq \gamma_1^2 - 2. \tag{1.30}$$

This resulted in 46 skewness and kurtosis combinations being investigated. The distribution pairing investigated either had both variables come from the same distribution or had one variable come from a normal distribution and the other from a non-normal distribution. Overall, 920 simulation scenarios were investigated. Results were based on 10,000 simulation replications.

The primary findings were that the RIN followed by the Spearman rank-order with Fieller's *SE* CIs had the best performance when data are non-normal. Of the remaining CI methods, only the observed imposed bootstrap with BCa had good enough performance when data was non-normal. However, it tended to exceed 95% coverage by generating somewhat long CIs. The advantage it has is that it keeps the correlation in the scale of the original variables. This is not the case for the RIN and Spearman rank-order with Fieller's *SE* as both transform the original variables. All the remaining methods did not have good CI coverage when data were non-normal with the Fisher *z*-transformation CI having the least favorable performance, and the situation was made worse by increasing the sample size when $\rho = .5$.

When the variable pairing included a normal distribution, all CIs generally performed better. However, the transformation methods still outperformed the bootstrap methods in this case. The only bootstrap CI methods that was comparable to the transformation methods' performance were the observed imposed bootstrap with AA and observed imposed bootstrap with BCa.

Like research on the robustness of the correlation *t*-test when testing $\rho = 0$, CI research on estimating the correlation when $\rho \neq 0$ is also limited and mixed. According to some research, the Fisher *z*-transformation, PB, and BCa CIs are robust when estimating the correlation. The caveat to the PB and BCa CI research is that the main interest was on the deattenuated correlation (i.e., correlation corrected for attenuation). However, subsequent research indicates that correlation CIs may not be robust when working with severely non-normal distributions. In fact, in such cases only CIs based on transformation methods showed robustness. Therefore, it is logical to investigate the promising non-transformation CIs from previous research to clearly understand under which situations the correlation CI is robust.

Confidence Interval Estimation

Fisher *z*-transformation CI. CIs constructed via the Fisher *z*-transformation (1915) are defined in equations 1.14 – 1.16.

Spearman Rank-Order CI. CIs constructed via the Spearman Rank-Order (1904) have x_{i1} and x_{i2} separately transformed into ascending ranks. From here, the correlation is computed. Confidence intervals are then be constructed by the Fisher *z*-transformation defined by equations

1.14 – 1.16. To draw parity with Bishara and Hittner's work, Fieller's (1957) standard error will be used defined as

$$SE(z) = \frac{1.03}{\sqrt{n-3}}. \tag{1.31}$$

This approach has a robust aspect in that outliers do not heavily affect it. This is because data in ranked form attenuates the differences between any two data points. For example, a sprinter who runs a 100 m dash in 9.58 seconds is faster than another sprinter who runs the 100 m in 12.43 seconds by 2.85 seconds. If this data were ranked, it is known that the first sprinter is faster than the second sprinter but not to what degree.

RIN Transformation CI. In this method, x_{i1} and x_{i2} are separately transformed through Bliss's (1967) rankit transformation defined as

$$f(x) = \phi^{-1}\left(\frac{x_r - .5}{n}\right) \tag{1.32}$$

where ϕ^{1} is the inverse cumulative distribution function and x_r is the ascending rank of each x_i value. The CI for the correlation is then computed through Fisher's z-transformation with equations 11.14 – 1.16. The utility of this process is that it will convert data into an approximately normal distribution.

Bootstrap for the Correlation

The bootstrapped correlation for a pair of variables *x* and *y* can be outlined in three steps. Suppose the observed data is $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ where $\mathbf{x}_i = (x_i, y_i)$ is the pair of variables. First, obtain the b^{th} bootstrap sample with replacement from **X**; i.e., $\mathbf{X}^{(b)} = \left(\mathbf{x}_1^{(b)}, \mathbf{x}_2^{(b)}, \ldots, \mathbf{x}_n^{(b)}\right)$. Second, the b^{th} estimate of the correlation from $\mathbf{X}^{(b)}$ is computed as

$$r_{xy}^{(b)} = \frac{S_{xy}^{(b)}}{S_x^{(b)} S_y^{(b)}} \tag{1.33}$$

and stored. Third, compile the stored estimates $r_{xy}^{(1)}, r_{xy}^{(2)}, ..., r_{xy}^{(B)}$ to create the ESD of r_{xy} for $b = 1, 2, ..., B$ bootstrap samples. The ESD can then be summarized to obtain statistical quantities for inference about r_{xy}.

PB CI. In this method, the percentile bootstrap CI is estimated by obtaining the $\alpha / 2$ and $1 - \alpha / 2$ percentiles from the r_{xy} ESD where α is the significance levels (i.e., probability of Type I error). For example, with $\alpha = .05$ the percentiles are then .025 and .975.

BCa CI. This method follows the same process as the PB CI method apart from how the CI bounds are defined. In this case, the bounds are adjusted according to equations 1.23 – 1.29. This can be interpreted as adjusting the bounds for the bias and skewness (or acceleration) of the r_{xy} ESD.

Highest Probability Density Interval (HPDI). Another method that can be used to form a CI from a distribution but has very limited research is the highest probability density interval (HPDI; Casella and Berger, 2002). If the distribution is unimodal, the HPDI is the narrowest interval that contains the specified probability of the confidence level ($1-\alpha$). In terms of the bootstrap, let $p\left(\hat{\theta} \mid \mathbf{x}\right)$ be the ESD for $\hat{\theta}$ given the data **x**. A $100(1-\alpha)\%$ HPDI for $\hat{\theta}$ is a subset $c \in \Theta$ defined as

$$c = \left\{\hat{\theta} : p\left(\hat{\theta} \mid \mathbf{x}\right) \geq k\right\} \tag{1.34}$$

where k is the largest number such that

$$\int_c p\left(\hat{\theta} \mid \mathbf{x}\right) d\hat{\theta} = 1 - \alpha \tag{1.35}$$

The principal idea is that *k* represents a horizontal line that shifts vertically down through the distribution until its intersections with the distribution capture the region with probability $1-\alpha$. This results in the region being projected upon the *x*-axis as an interval. Another way to think about the HPDI is that it is the narrowest interval because it is optimized based on where the data gather most frequently or is most dense. Additionally, the HPDI has two advantages. First, like the bootstrap CIs, the HPDI is also distribution-free. Second, no region outside of the interval will have higher probability than any region inside the interval. Figure 2 illustrates the difference between the HPDI and percentile-based CIs.

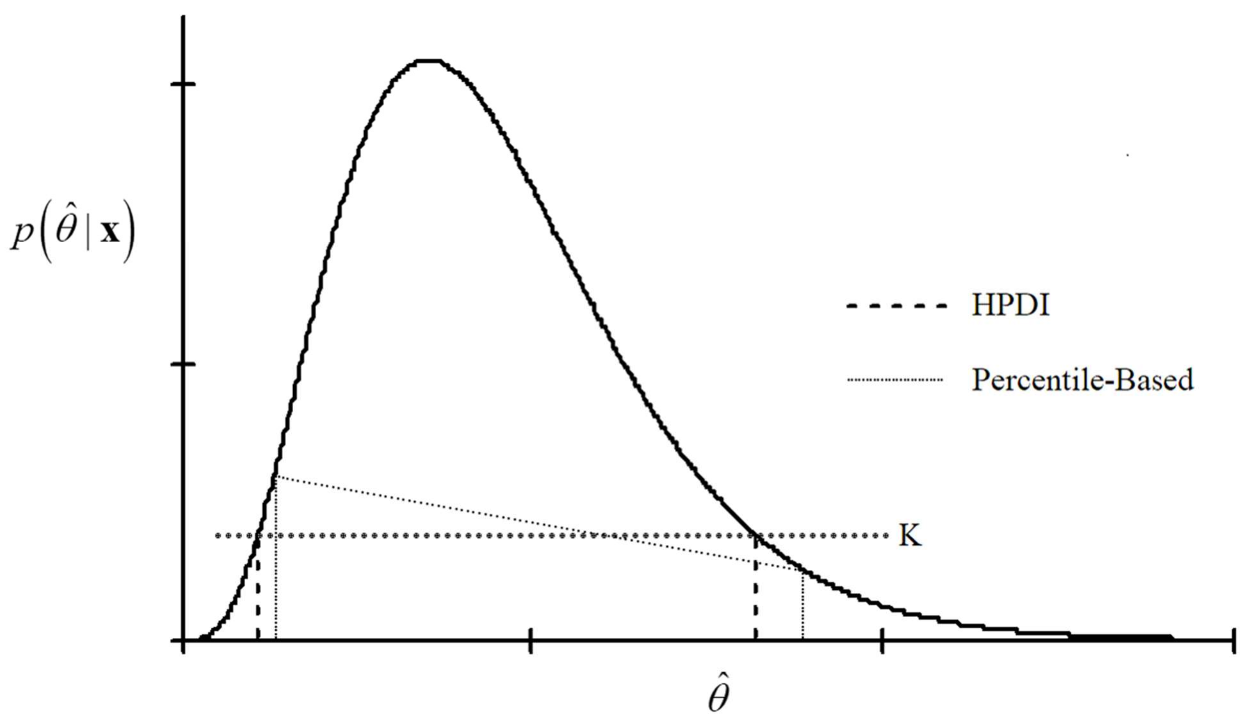

Figure 2. Comparison between HPDI and Percentile-Based CIs

Note that the inclusion of the HPDI is supplemental to the study. The main purpose for including it is was a research question on whether it can be used on the ESD from a parameter of interest.

Rationale

As aforementioned, there is currently a concern in the literature regarding replicability (Collaboration, 2015). One methodological suggestion to help address the issue is to use the correlation as an effect size (Cohen, 1990). Even so, this suggestion can be enhanced by including a correlation CI to provide precision information. However, CI research about the correlation is limited and mixed. Therefore, the aim here is to shed further light and verify the limitations of the bootstrap CI for estimating the correlation. Given that various conditions can impact the performance of CI estimation (e.g., sample size, data distribution, etc.), a simulation study is considered.

This simulation study aims to compare the most promising correlation CIs used in previous research. Two classes of CIs will be investigated: non-bootstrap and bootstrap. The non-bootstrap CIs are the Fisher *z*-transformation, Spearman rank-order with Fieller's *SE*, and RIN. The bootstrap CIs include the PB, BCa, and HPDI. The bootstrap CI methods are the primary focus of the research, but non-bootstrap methods were included for comparison purposes. Additionally, condition parity will be drawn between the previous bootstrap CI research so that results can be directly compared. All CIs were compared in terms of coverage probability. By comparing the performance of these CI estimation methods in different conditions, it is hoped that the findings will benefit applied researchers in choosing the appropriate CI for the correlation, particularly in applications that call for the correlation to be used as an effect size.

Based on the literature review, the following are expected from the correlation CIs:

1. When data for both variables are normally distributed (multivariate normal), all CI estimation method will perform well in terms of coverage probability. This expectation is based on the assumptions of the correlation (Padilla, 2017)
2. When multivariate normality does not hold, the Spearman and RIN CIs will have the best performance in terms of coverage probability; followed by the HPDI, BCa, PB, and Fisher z-transformation CIs. However, the HPDI and BCa CI will perform comparable to one another. This expectation is based on the previous findings from Bishara and Padilla as well as the properties of the HPDI (Bishara & Hittner, 2017; Casella & Berger, 2002; Padilla, 2012, 2014).
3. As the degree of non-normality increases, coverage probability of the Fisher z-transformation and PB CIs will perform worse, while the remaining methods will be robust. This expectation is based on the previous findings from Bishara and Padilla as well as the properties of the HPDI.
4. When the correlation is zero, coverage probability for all CI estimation methods will perform equally well. This expectation is based on previous findings regarding CI estimation methods when the correlation is zero (Berry & Mielke, 2000; Bishara & Hittner, 2017; Zeller & Levine, 1974).
5. As the correlation moves away from zero, coverage probability of the Fisher z-transformation and PB CIs will perform worse, while the remaining methods will be robust. This expectation is based on the previous findings from Bishara and Padilla as well as the properties of the HPDI.

CHAPTER II

METHOD

A Monte Carlo simulation is a suitable option for investigating the performance of a statistical method when analytical methods are not available (Yung & Bentler, 1996). It provides insight into the performance of a statistical method wherein established properties and assumptions may not hold (Bandalos & Leite, 2013). Given that comparisons of the correlation CI estimation methods cannot be achieved through analytical methods, and because the assumptions of the correlation may not be met in applied settings, a simulation study is used to investigate the performance of the correlation CIs.

Data Generation

A Monte Carlo simulation was used to investigate and compare the properties of the correlation CI estimation methods under different simulation conditions. This simulation was structured in a 6 (correlation magnitude) × 11 (sample size) × 23 (distribution pairings) simulation design for a total of 1518 conditions. For each simulation condition, 1,000 data replications were obtained. Data were simulated in three steps as follows. First, generate normal and non-normal data according to Headrick (2002) as follows

$$\begin{bmatrix} x_i \\ x_j \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{bmatrix} \right) \tag{2.1}$$

where

$$x_i = c_{0i} + c_{1i} z_i + c_{2i} z_i^2 + c_{3i} z_i^3 + c_{4i} z_i^4 + c_{5i} z_i^5 , \tag{2.2}$$

$$x_j = c_{0j} + c_{1j} z_j + c_{2j} z_j^2 + c_{3j} z_j^3 + c_{4j} z_j^4 + c_{5j} z_j^5 , \tag{2.3}$$

$$\begin{bmatrix} z_i \\ z_j \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{ij}^* \\ \rho_{ij}^* & 1 \end{bmatrix} \right), \tag{2.4}$$

$c_0, \ldots, c_5$ are constants (See Table 27), and ρ_{ij}^{*} is the intermediate correlation. The idea is to determine the constants and intermediate correlation to obtain the variables in equation 2.1. The unique aspect of this method is that it allows for correlation amongst non-normal data with the use of extra moments. Second, estimate the correlation CIs for each data replication in equation 2.1. Third, determine if the CIs contain the population correlation (ρ). The following simulation conditions were investigated.

Conditions

Sample Size (*n*). Sample size was included because CI estimation is impacted by sample size. The following sample sizes will be investigated: n = 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400 as they will draw parity with the previous research on the correlation CIs (Bishara & Hittner, 2017; Padilla, 2012, 2014). The granularity used here is due to previous research suggesting that bootstrap CIs cease to function if sample size is too small (Rasmussen, 1987).

Correlation Magnitude (ρ). Correlation magnitude was considered because the distribution of the correlation changes when $\rho \neq 0$ and becomes more skewed as it approaches ±1 (see Figure 1). The correlation measures the magnitude of the linear relationship between two continuous variables and can be interpreted as an effect size. Cohen (1988) advises that correlation coefficients equal to .10, .30, and .50 represent small, moderate, and strong correlations, respectively. The present investigation will utilize correlation coefficients ranging from .00 to .50 in increments of .10 to address Cohen's standards but to also gain further insight on a wider range of effects. It is expected that the CIs will perform better the closer the correlation coefficient is to zero.

Distribution Pairings (x_{ij}). The degree of non-normality was the focus in the distribution pairings because the Fisher method assumes multivariate normality while the bootstrap methods do not. Additionally, non-normal data is common in applied settings and there is a lack of consensus in the literature on how to approach the correlation when non-normality occurs. The selected distributions fall into the general categories of symmetric and non-symmetric and were investigated in previous research (Bishara & Hittner, 2017; Padilla, 2012, 2014). The symmetric distributions were as follows: Normal, Triangular, Uniform, and Laplace. The non-symmetric distribution were as follows: Beta $(\alpha = 4, \beta = 1.25)$, Beta $(\alpha = 4, \beta = 1.25)$, Chi-Square $(df = 16)$, Chi-Square $(df = 4)$, Chi-Square $(df = 3)$, Chi-Square $(df = 2)$, Chi-Square $(df = 1)$, and Pareto. Figure 3 shows the graphical representation of these distributions.

Investigation of the distributions considered the pairwise nature of the variables involved in the correlation. This dictated four main types of distribution pairings. In the first pairing, the variables had the same symmetric distribution (e.g., both variables were Uniform). Similarly, in the second pairing, the variables had the same non-symmetric distribution (e.g., both variables were Pareto). In the third pairing, one variable was always normal and other was symmetric (e.g., one variable was Normal and the other Laplace); a normal-normal paring was not included. In the fourth pairing, one variable was always Normal and the other was non-symmetric (e.g., one variable was Normal and the other Pareto).

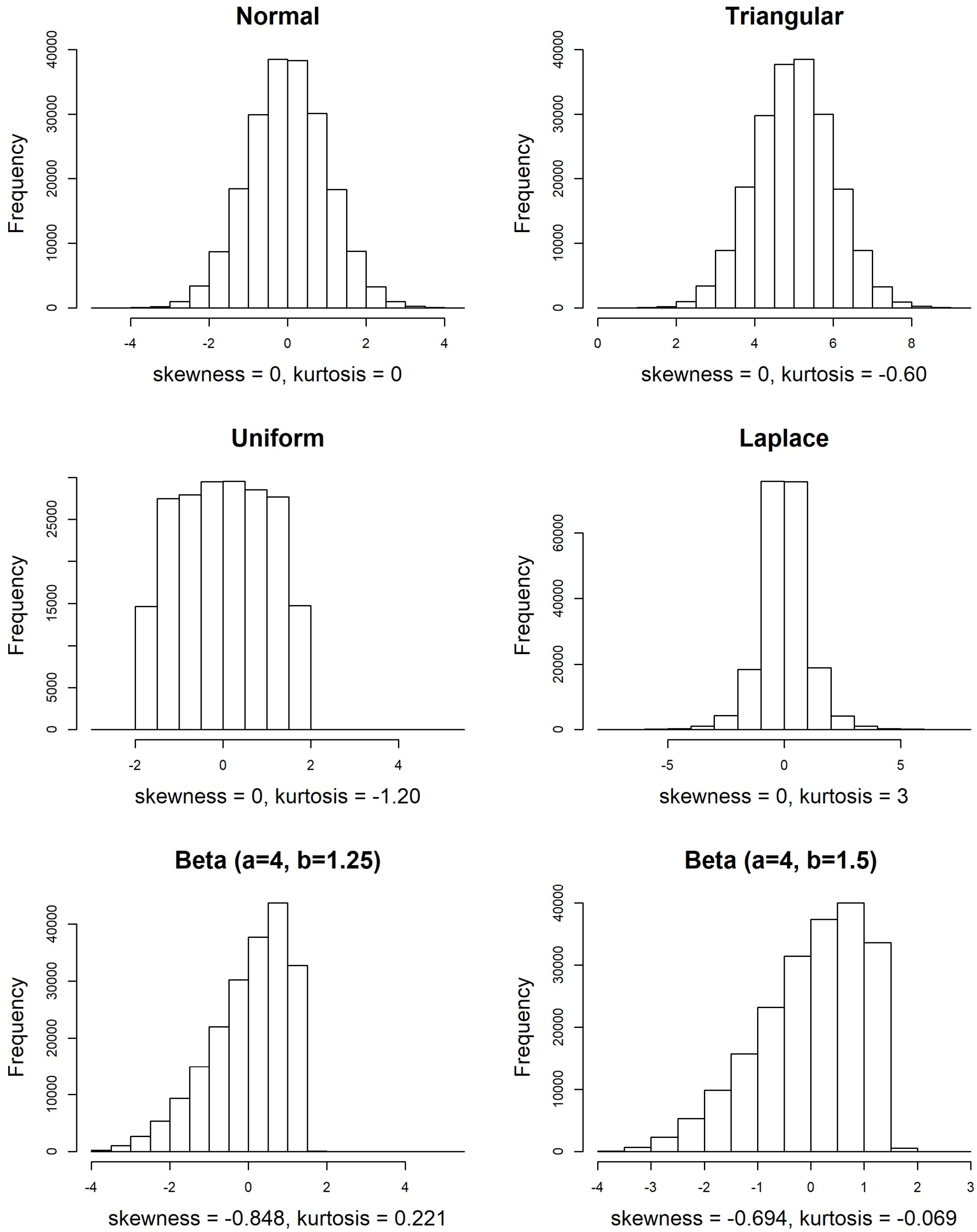

Figure 3. Distributions considered for this study.

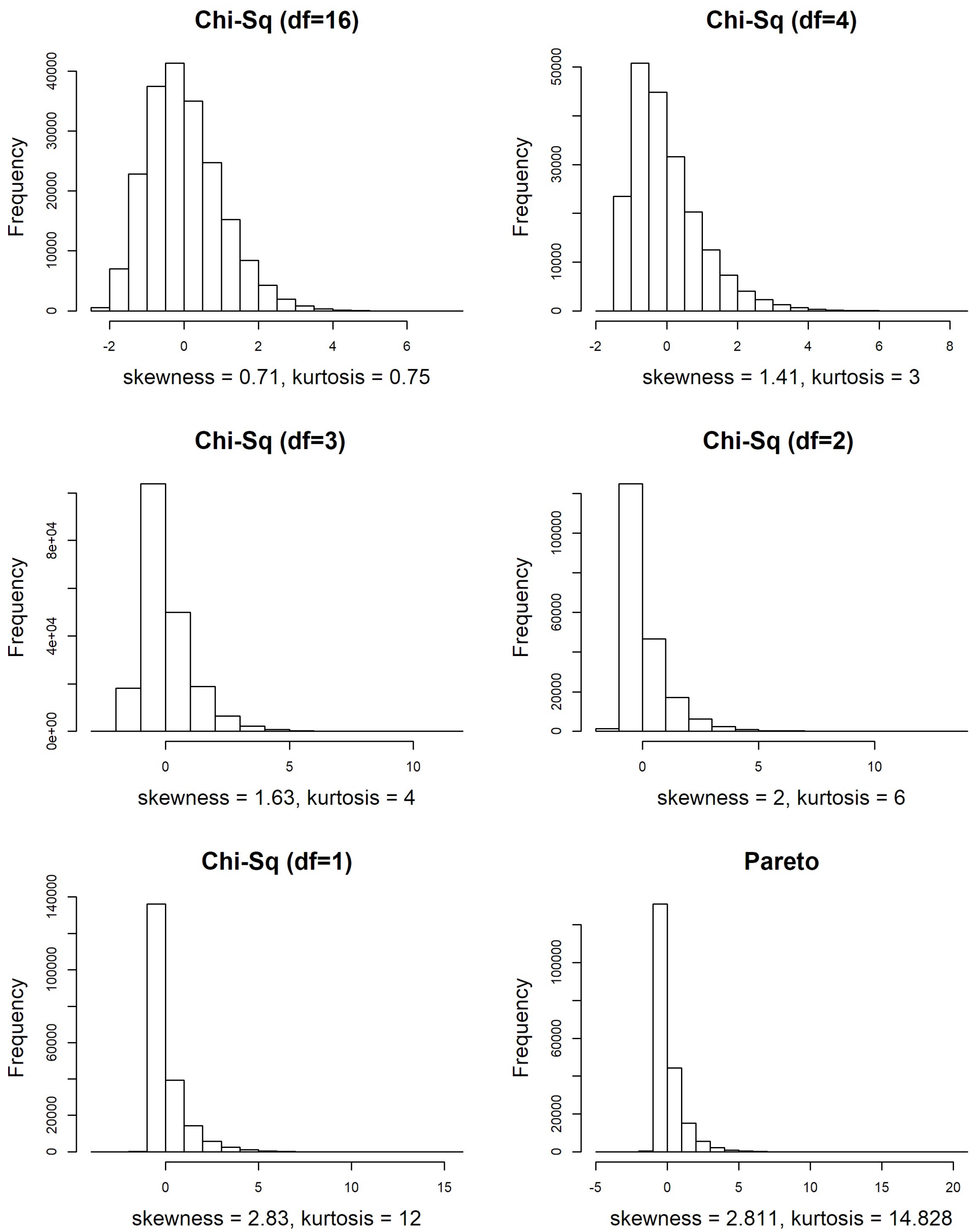

Figure 3 Continued. Distributions considered for this study.

Analysis

The $100(1-\alpha)\%$ CIs were estimated with $\alpha = .05$. CIs that involved using the bootstrap were estimated from a total of 2,000 bootstrap samples. To evaluate the performance of each CI, two criteria were used: coverage probability and bias.

Coverage probability is defined as the proportion of times the estimated CI contain the true population correlation ρ. Bradley's (1978) liberal criterion was used to determine acceptable coverage, which is defined as $1-1.5\alpha \leq 1-\alpha^{*} \leq 1-0.5\alpha$ were α^* is the true probability of Type I error. Therefore, acceptable coverage is given by $[.925, .975]$ for $\alpha = .05$. This criterion has been used in simulation studies (Padilla, 2012, 2014).

Although the focus of the study was CI estimation, point estimates were also reported as a supplement in evaluating the CIs. For each point estimate, standardized bias was computed as

$$\hat{\rho}_{bias} = \frac{\hat{\rho}-\rho}{SE(\hat{\rho})} \tag{2.5}$$

where $\hat{\rho}$ is the point estimate, $SE(\hat{\rho})$ is the standard error of $\hat{\rho}$, and ρ is the population parameter, respectively. Standardized bias allows for the standardized comparison of the Pearson, Spearman, and RIN point estimates. Given that the Pearson point estimates form the ESD of the bootstrap, standardized bias for the bootstrap were not reported. Acceptable bias for a point estimate was defined as $|\hat{\rho}_{bias}| \leq .40$ (Collins et al., 2001).

To further explore the impact of the simulation conditions, on the performance of the CIs, the following models were estimated: logistic regression and 3-way ANOVA. The independent variables for the two models were correlation magnitude, sample size, and distribution pairing. The main effects and two-way interactions were fitted in the logistic regression and ANOVA models. For the logistic regression, the binary dependent variable indicated if the estimated CI covered the population correlation ρ (0 = no; 1 = yes). The Cox-Snell R^2_{CS} (effect size) was used to examine the contribution of each effect in the logistic models. The R^2_{CS} for an effect was computed as the difference between the R^2_{CS} for the full model (i.e., model with all main effects and two-way interactions) and the reduced model (i.e., full model excluding that effect of interest). For ANOVA, the dependent variable was standardized bias, and partial eta squared (η^2) was used to examine the contribution of each effect.

The logistic regression and ANOVA models were conducted for heuristic and descriptive reasons. In simulation work, power and effect size must be considered with care for two reasons. First, simulation work is based on specified population parameters and the estimation methodology must achieve those parameters. To reduce sampling variability and assure the precision of the estimation under investigation, the number of simulation replications need to be large. Regardless of the estimation methodology being investigated, simulation replications start with a minimum of 1,000; i.e., the number of replications for this simulation study. With such large simulation replications, any small deviation will be declared as statistically significant; i.e., simulation work is overpowered. Second, simulation replications and estimation methodology are not human participants or human behavior. As such, using effect size criteria for the behavioral/social sciences to judge the magnitude of an effect size in simulation work must be done with care. Even so, effect sizes (R^2_{CS} and η^2) were used to examine the contribution of an effect in the logistic and ANOVA models in a descriptive manner to help with the overall analysis.

CHAPTER III

RESULTS

Data for the simulation were generated using the *R* statistical package 3.0.2. The 6 (correlation magnitude) × 11 (sample size) × 23 (distribution pairings) simulation design was analyzed using SAS software 9.4. Overall, 1518 simulated conditions for the correlation and its corresponding CIs (i.e., Fisher's z-transformation, Spearman Rank-order, RIN, PB, BCa, and HPDI) were investigated. The Bootstrap CIs used 2000 bootstrap samples. Performance for the correlation CIs and corresponding point estimates were primarily accessed with Bradley's (2015) criterion and standardized bias, respectively. Further analysis of the simulation conditions on the performance of the CIs was done via logistic regression and 3-way ANOVA. The forthcoming sections discuss the results for each condition combination. Coverage probabilities related to the Spearman Rank-order and RIN are not discussed due to consistent coverage near the target of .95 within Bradley's criterion; which was expected.

Coverage Probability

Table 1 presents the Cox-Snell R^2_{CS} (effect size) of the logistic regression for coverage probability for each of non-bootstrap CIs. The Spearman and RIN CIs were not impacted by any of the simulation conditions; the largest $R^2_{CS} = 0.0002$ or .02%. However, the Fisher z-transformation CI was slightly impacted by the simulation conditions. In particular, the Fisher z-transformation CI was slightly impacted by sample size (n) and correlation strength (ρ) main effects; $R^2_{CS} = 0.0099$ and $R^2_{CS} = 0.0091$, respectively. The other effects and interactions did not have as strong impact.

Table 2 presents the Cox-Snell R^2_{CS} (effect size) of the logistic regression for coverage probability for each of bootstrap CIs. In general, all three bootstrap CIs had very small main effects. Even so, they were all most impacted by correlation strength (ρ). This was followed by distribution pairing (d) and sample size (n), respectively, for the PB CI and HPDI. However, for the BCa CI, the order was switched: sample size (n) and distribution pairing (d), respectively.

Affect sizes for statistical performance must be interpreted with caution. The first thing to realize is that although the effect sizes in Tables 1 and 2 appear small, recall that human behavior is not being analyzed. Instead, the performance of statistics is being analyzed. In particular, the performance of coverage probability for CIs. In this context, even a trivially small effect size may reveal unacceptable performance by a statistic. For example, the magnitude of Bradley's (1978) acceptable performance is $.975 - .925 = .05$. Therefore, all effects of the simulation conditions on coverage probability were investigated and presented in Tables $4-26$ and Figures $4-35$. Note that Tables $4-26$ present the interaction of correlation magnitude and sample size for each distribution pairing; individual tables for each simulation condition main effect are not presented.

Coverage for Correlation Magnitude

Figure 4 shows the CI coverage for correlation magnitude. In general, most CIs tended to have acceptable coverage $(\text{i.e., } [.925, .975])$ across the correlations. Even so, there was a clear order of CI coverage performance. The BCa and PB CIs had consistent acceptable coverage that was similar to each other. The Fisher z-transformation CI and HPDI followed, respectively. The Fisher z-transformation CI tended to become worse (wider box plots with more outliers) as the correlation increased. However, the HPDI tended to have the most unacceptable coverage.

Coverage for Sample Size

Figure 5 shows the CI coverage for sample size. Generally, most CIs tended to have acceptable coverage $(\text{i.e., } [.925, .975])$ across the sample sizes. Furthermore, this coverage improved as sample size increased. Still, there was a clear order of CI coverage performance in this case. The Fisher z-transformation CI had consistent acceptable coverage. This was followed by the BCa CI. The PB CI then follows but struggles to maintain adequate coverage when $n \leq 30$. The HPDI had the weakest performance as it does not have adequate coverage until $n \geq 50$. However, all CIs maintain adequate coverage once sample sizes are $n \geq 50$. Additionally, the coverage improves (narrower box plots) for all the CIs as the sample size increases. The exception is the Fisher z-transformation CI, which had more outliers as the sample size increased.

Coverage for Distribution Pairing

Symmetric w/ Symmetric. Figure 6 shows the CI coverage for a symmetric with symmetric variable pairing. In general, most CIs tended to have acceptable coverage $\left(\text{i.e., } [.925, .975]\right)$. However, there was a clear order of CI coverage performance in this case. The Fisher z-transformation CI had consistent acceptable coverage. This was followed by the BCa CI and PB CI coverage, respectively. However, in the Laplace variable pairing, the PB CI coverage outperforms the BCa CI. The HPDI had the weakest performance with more instances of unacceptable coverage.

Non-Symmetric w/ Non-Symmetric. Figure 7 shows the CI coverage for a non-symmetric with non-symmetric variable pairing. In general, CIs had coverage issues with $\chi^2\left(df = 1-4\right)$ and Pareto variable pairings; i.e., when $|\text{skewness}| \geq 1.41$ for both variables. This becomes more profound as skewness increased for the paired variables. The Fisher z-transformation CI was most impacted with more unacceptable coverage (wider box plots). This was followed by the HPDI, BCa CI, and PB CI; respectively.

Most CIs had good coverage with $\text{Beta}(a=4, b=1.25-1.5)$ and $\chi^2(df=16)$ variable pairings; i.e., when $|\text{skewness}| \leq .71$ for both variables. In these cases, the Fisher z-transformation, BCa, and PB CIs had similar coverage. The HPDI consistently had coverage that was behind in these instances.

Symmetric w/ Normal. Figure 8 shows the CI coverage for a symmetric with normal variable pairing. Coverage in this situation was similar to that of the symmetric with symmetric variable pairing (see Figure 6).

Non-Symmetric w/ Normal. Figure 9 shows the CI coverage for a non-symmetric with normal variable pairing. Generally, most CIs tended to have acceptable coverage $(\text{i.e.}, [.925, .975])$. However, there was a clear order of CI coverage performance in this case. The Fisher z-transformation CI had consistent acceptable coverage. The PB and BCa CIs then follow with coverage that tended to be similar to each other. The HPDI had the weakest performance with the most instances of unacceptable coverage.

Coverage for Sample Size by Correlation Magnitude

Sample Size by Correlation Magnitude. Figures 10 and 11 show the CI coverage for sample size by correlation magnitude. In general, most CIs tended to have adequate coverage $(\text{i.e.}, [.925, .975])$. Additionally, coverage improved as the sample size increased. Still, there was a clear order CI coverage performance. The Fisher z-transformation CI had the most consistent acceptable coverage across all sample sizes and correlation magnitudes. However, the Fisher z-transformation CI also had some severe outliers across all sample sizes. This performance was then followed by the BCa CI and PB CI, respectively. However, the PB CI had coverage issues when $n \leq 30$. The HPDI had the weakest performance as it required $n \geq 100$ to maintain adequate coverage. In fact, all CIs maintain adequate coverage once $n \geq 100$ and improve as the sample size increases.

Coverage for Distribution Pairing by Correlation Magnitude

Symmetric w/ Symmetric by Correlation Magnitude. Figures 12 and 13 show the CI coverage for a symmetric with symmetric variable pairing by correlation magnitude. In general, most CIs have tended to have acceptable coverage $\left(\text{i.e.,} \left[.925, .975\right]\right)$ across all correlation magnitudes. However, was a clear order of CI coverage performance. The Fisher z-transformation CI had consistent acceptable coverage across all distribution pairing and correlation magnitudes. This was followed by the BCa and PB CIs, which tended to have similar coverage. However, the PB CI is more robust in the Laplace distribution pairing. The HPDI had the weakest performance as it tended to have more instances of unacceptable coverage.

Non-Symmetric w/ Non-Symmetric by Correlation Magnitude. Figure 14 shows the CI coverage for a non-symmetric with non-symmetric variable pairing by correlation magnitudes of $0-0.2$. Generally, the CIs tended to have to have coverage issues with the $\chi^2\left(df=1\right)$ and Pareto distribution pairing; i.e., the CIs had coverage issues when $\left|\text{skewness}\right| \geq 2.811$ for both variables. In fact, all CIs tended to struggle maintaining acceptable coverage $\left(\text{i.e.,} \left[.925, .975\right]\right)$. In these cases, the PB and BCa CIs have about equal performance and are followed by the HPDI, and Fisher z-transformation CI; respectively. The Fisher z-transformation CI noticeably performs worse as correlation magnitude increases.

Most CIs had good coverage with $\text{Beta}\left(a=4, b=1.25-1.5\right)$ and $\chi^2\left(df=2-4,16\right)$ variable pairings; when $\left|\text{skewness}\right| \leq 1.63$. In this case the Fisher z-transformation CI had the best coverage, followed by the BCa CI, PB CI, and HPDI. Additionally, this performance tended to remain the same across correlation magnitudes. The exception to this was the Fisher z-transformation CI for $\chi^2(df = 2)$ when $\rho = .2$.

Figure 15 shows the CI coverage for a non-symmetric with non-symmetric variable pairing by correlation magnitudes of $0.3 - 0.5$. Generally, the CIs had coverage issues with $\chi^2(df = 1-4)$ and Pareto distribution pairings; i.e., the CIs had coverage issues when $|\text{skewness}| \geq 1.41$ for both variables. In fact, all CIs tended to struggle maintaining acceptable coverage $(\text{i.e.}, [.925, .975])$. In these cases, the PB CI had the best coverage. The BCa CI and HPDI then follow with similar performance to one another. Finally, the Fisher z-transformation CI had the weakest coverage performance. Additionally, the Fisher z-transformation CI had notably worse performance as correlation magnitude increased.

Most CIs had good coverage with $\text{Beta}(a = 4, b = 1.25 - 1.5)$ and $\chi^2(df = 16)$ variable pairings; i.e., had good coverage when $|\text{skewness}| \leq .848$ for both variables. In this case the BCa CI had the best coverage, followed by the PB CI, Fisher z-transformation CI, and HPDI. Additionally, the Fisher z-transformation CI was notably negatively impacted in the $\text{Beta}(a = 4, b = 1.25)$ variable pairing when correlation magnitude increased.

Symmetric w/ Normal by Correlation Magnitude. Figures 16 and 17 show the CI coverage for a symmetric with normal variable pairing by correlation magnitude. In general, most CIs have tended to have acceptable coverage $(\text{i.e.}, [.925, .975])$ across all correlation magnitudes. However, there was a clear order of CI coverage performance. The Fisher z-transformation CI had consistent acceptable coverage across all distribution pairing and correlation magnitudes. This was followed by the BCa and PB CIs, which tended to have similar coverage. The HPDI had the weakest performance as it tended to have more instances of unacceptable coverage.

Non-symmetric w/ Normal by Correlation Magnitude. Figures 18 and 19 show the CI coverage for a non-symmetric with normal variable pairing for correlation magnitudes. In general, most CIs tended to have adequate coverage $\left(\text{i.e.,} \left[.925, .975\right]\right)$. There was a clear order of performance. The Fisher z-transformation CI had the best performance with consistent acceptable coverage. This was followed by the BCa CI and PB CI, respectively. The HPDI had the weakest performance with more instances of unacceptable coverage.

Coverage for Distribution Pairing by Sample Size

Symmetric w/ Symmetric by Sample Size. Figures $20-23$ show the CI coverage for a symmetric with symmetric variable pairing by sample size. In general, the CIs tended to have acceptable coverage $\left(\text{i.e.,} \left[.925, .975\right]\right)$. In addition, coverage improved as the sample size increased. Still, there was a clear order of CI coverage performance. The Fisher z-transformation CI had consistent acceptable coverage across all distribution pairings and sample sizes. The BCa CI had the next best performance but had some coverage difficulty with coverage with the Laplace distribution pairing when $n \leq 100$. The PB CI had the next acceptable coverage and has a dip in performance with the Laplace distribution pairing. The HPDI had the weakest performance with a tendency for unacceptable coverage across all distribution pairings. However, all CIs tended to have acceptable coverage as sample increased. However, all CIs tended to have acceptable coverage as sample size increased.

Non-Symmetric w/ Non-Symmetric by Sample Size. Figures 24 and 25 shows the CI coverage for a non-symmetric with non-symmetric variable pairing by sample sizes $20-150$. In general, the CIs tended to have coverage issues with $\chi^2\left(df = 1-4\right)$ and Pareto variable pairings;

i.e., the CIs had coverage issues when $|\text{skewness}| \geq 1.41$ for both variables. This became more profound as the skewness increased for the paired variables. Even though increasing the sample size improved coverage (made box plots narrower), coverage still tended to be outside of acceptability $(\text{i.e., } [.925, .975])$. The Fisher z-transformation CI is the sole exception in that increasing sample size worsened coverage (wider box plots).

Most CIs had good coverage with $\text{Beta}(a = 4, b = 1.25 - 1.5)$ and $\chi^2(df = 16)$ variable pairings; i.e., had good coverage when $|\text{skewness}| \leq .848$ for both variables. In this case the Fisher z-transformation CI had the best coverage, followed by the BCa CI, PB CI, and HPDI. However, CI coverage improves as sample size increases and all CIs tended to have acceptable coverage when $n \geq 100$.

Figures 26 and 27 show the CI coverage for a non-symmetric with non-symmetric variable pairing by sample sizes of $200 - 400$. In general, the CIs tended to have coverage issues with $\chi^2(df = 1)$ and Pareto variable pairings; i.e., when $|\text{skewness}| \geq 2.811$ for both variables. Here, the PB CI maintains the best coverage. This was followed by the HPDI and BCa CI, respectively. The Fisher z-transformation CI did not have acceptable coverage and increasing sample size only resulted in decreased coverage (wider box plots).

However, the PB CI, BCa CI, and HPDI had good coverage with $\text{Beta}(a = 4, b = 1.25 - 1.5)$ and $\chi^2(df = 2 - 4, 16)$ variable pairings; i.e., when $|\text{skewness}| \leq 2$ for both paired variables. In this case the PB CI had the best performance, followed by the BCa CI, HPDI, and Fisher z-transformation CI. The Fisher z-transformation CI only had good coverage when in $\text{Beta}(a = 4, b = 1.25 - 1.5)$ and $\chi^2(df = 16)$; i.e., when $|\text{skewness}| \leq .71$. When skewness was greater, the fisher z-transformation CI had problems maintaining acceptable coverage regardless of sample size.

Symmetric w/ Normal by Sample Size. Figures $28-31$ shows the CI coverage for a symmetric with normal variable pairing by sample size. In general, the CIs tended to have acceptable coverage $\left(\text{i.e., }[.925, .975]\right)$. In addition, coverage improved as the sample size increased. In this situation, there was a clear order of CI coverage performance. The Fisher z-transformation CI had consistent acceptable coverage across all distribution pairings and sample sizes. For the most part, the BCa CI tended to have consistent acceptable coverage across all distribution pairings and sample sizes. The only exception for the BCa CI was that it had coverage difficulty with the Laplace distribution pairing when $n \le 40$. The PB CI had the next best performance. The HPDI had the weakest overall performance and struggled with coverage when $n \le 50$. Even so, all coverage improved (narrower box plots) with the increase in sample size.

Non-Symmetric w/ Normal by Sample Size. Figures $32-35$ show the CI coverage for a non-symmetric with normal variable pairing by sample size. In general, most of the CIs tended to have acceptable coverage $\left(\text{i.e., }[.925, .975]\right)$. Additionally, coverage improved as the sample size increased. Here, there was a clear order of CI performance. The Fisher z-transformation CI had consistent acceptable coverage across all distribution pairings and sample sizes. The BCa CI follows but has coverage difficulty when one of the variables was $\chi^2\left(df = 1-2\right)$ or Pareto. The PB CI had the next best acceptable coverage. The HPDI had the weakest performance with a tendency for unacceptable coverage across all distribution pairings when $n \le 50$. Even so, CI coverage improved as sample size increased. Likewise, coverage improved (narrower box plot) as sample size increased.

Standardized Bias

Given that CI coverage was impacted by the simulation conditions, parameter bias was also investigated via standardized bias. Table 3 presents the η^2 (effect size) of the ANOVA for standardized bias for each correlation estimate. The Pearson, Spearman, and RIN were not impacted by any of the of the simulation conditions or their interactions; the largest $\eta^2 = .0020$ or .2%. As aforementioned, while these effect sizes are trivial, they may reveal unacceptable performance for a statistic. Therefore, all effects of the simulation conditions on standardized bias were investigated and presented in Figures $36-67$. Tables for standardized bias are not presented as all estimates had acceptable bias.

Standardized Bias for Correlation Magnitude

Figure 36 shows the standardized bias for correlation magnitude. Generally, all estimates had acceptable standardized bias $\left(\text{i.e., } |bias| \leq .40\right)$. However, there was a clear order of performance. The Pearson had the best performance as it had the least variation from the parameter. However, as correlation magnitude increased, the Pearson developed more severe outliers. Following that is the Spearman which varied more from the parameter as correlation magnitude increased. The RIN is the weakest performer as it had the most variation. In general, all methods increased in variation and outliers as the correlation magnitude increased.

Standardized Bias for Sample Size

Figure 37 shows the standardized bias for sample size. Generally, all estimates had acceptable standardized bias $\left(\text{i.e., } |bias| \leq .40\right)$. However, there was a clear order of performance. The Pearson had the best performance as it was closer to the parameter and had the least variation. The Spearman and RIN followed, respectively. Additionally, increasing in sample size resulted in closer estimates to the parameter with reduced variation for standardized bias for all methods (narrower box plots).

Standardized Bias for Distribution Pairing

Symmetric w/ Symmetric. Figure 38 shows the standardized bias for a symmetric with symmetric variable pairing. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, it was evident that Pearson has the best performance as it had the least variation. Following this are the Spearman and RIN, respectively. Additionally, all estimates tended to have a few outliers that underestimated the parameter.

Non-Symmetric w/ Non-Symmetric. Figure 39 shows the standardized bias for a non-symmetric with non-symmetric variable pairing. In general, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, there was a clear order of performance. The Pearson is the most precise around the parameter. Following this are the Spearman and RIN, respectively. Additionally, all estimates tended to have outliers that underestimated the parameter.

Symmetric w/ Normal. Figure 40 shows the standardized bias for a symmetric with normal variable pairing. Standardized bias in this situation was similar to that of symmetric with symmetric variable pairing (see Figure 38).

Non-Symmetric w/ Normal. Figure 41 shows the standardized bias for a non-symmetric with normal variable pairing. Standardized bias in this situation was similar to that of non-symmetric with non-symmetric variable pairing (see Figure 39) with slightly more variation.

Standardized Bias for Sample Size by Correlation Magnitude

Figures 42 and 43 show the standardized bias for sample size by correlation magnitude. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. All estimates tended to become less biased as sample size increased. Additionally, increased sample size tended to result in less variation (narrower box plots). However, increases in correlation magnitude tended to result in more biased estimates.

Standardized Bias for Distribution Pairing by Correlation Magnitude

Symmetric w/ Symmetric by Correlation Magnitude. Figures 44 and 45 show the standardized bias for a symmetric with symmetric variable pairing by correlation magnitude. In generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} |bias| \leq .40\right)$. Even so, a correlation magnitude and sample size had an impact (see Figure 51). In particular, when $\rho = .30 - .50$ and $n = 20$ all the estimates were further from the parameter and had noticeable variation. However, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. In addition, all the estimates got closer to the parameter and had less variation as the sample size increased.

Non-Symmetric w/ Non-Symmetric by Correlation Magnitude. Figures 46 and 47 show the standardized bias for a non-symmetric with non-symmetric variable pairing by correlation magnitude. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} |bias| \leq .40\right)$. Even so, a correlation magnitude had an impact (see Figure 51). In particular, when $\rho = .50$ the Spearman and RIN were further from the parameter and had noticeable variation. However, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Variation among the estimates tended to increase as correlation magnitude increased (wider box plots). Increases in correlation magnitude also resulted in more outliers.

Symmetric w/ Normal by Correlation Magnitude. Figures 48 and 49 show the standardized bias for a symmetric with normal variable pairing by correlation magnitude. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Estimate variation tended to increase as correlation magnitude increased (wider box plots).

Non-symmetric w/ Normal by Correlation Magnitude. Figures 50 and 51 show the standardized bias for a non-symmetric with normal variable pairing by correlation magnitude. In general, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, a correlation magnitude had an impact (see Figure 51). In particular, when $\rho = .50$ the Spearman and RIN were further from the parameter and had noticeable variation. However, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Variation among the estimates tended to increase as correlation magnitude increased (wider box plots). Increases in correlation magnitude also resulted in more outliers.

Standardized Bias for Distribution Pairing by Sample Size

Symmetric w/ Symmetric by Sample Size. Figures $52 - 55$ show the standardized bias for a symmetric with symmetric variable pairing by sample size. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. However, a small sampled had an impact (see Figure 52). In particular, when $n = 20$ all estimates were further from the parameter and had noticeable variation. Even so, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Additionally, increases in sample size resulted in less bias and more precision (narrower box plots).

Non-Symmetric w/ Non-Symmetric by Sample Size. Figures $56-59$ show the standardized bias for a non-symmetric with non-symmetric variable pairing by sample size. In general, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, a small sampled had an impact (see Figure 56). In particular, when $n = 20$ all estimates were further from the parameter and had noticeable variation. However, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Additionally, increases in sample size resulted in less bias and more precision (narrower box plots).

Symmetric w/ Normal by Sample Size. Figures $60-63$ show the standardized bias for a symmetric with normal variable pairing by sample size. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. Even so, a small sampled had an impact (see Figure 60). In particular, when $n = 20$ all estimates were further from the parameter and had noticeable variation. However, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Additionally, increases in sample size resulted in less bias and more precision (narrower box plots).

Non-Symmetric w/ Normal by Sample Size. Figures $64-67$ show the standardized bias for a non-symmetric with normal variable pairing by sample size. Generally, all estimates had acceptable standardized bias $\left(\text{i.e.,} \left|bias\right| \leq .40\right)$. However, a small sample had an impact (see Figure 64). In particular, when $n = 20$ all the estimates were further from the parameter and had noticeable variation. Evens so, there was a clear order of performance. The Pearson had the best performance, followed by the Spearman and RIN. Additionally, increase in sample size tended to result in less bias and less variability (narrower box plots).

Summary of Expected Results

Several important finding can be summarized from the results of the study in conjunction with the study expectations. The major findings are presented below.

First, it was expected that when data for both variables are normally distributed, all CI estimation methods would perform well in terms of coverage probability. Results from the study partially support this expectation. According to the results, coverage probabilities were acceptable for most of the methods for normally distributed data (see Table 1 and Figures 6, $12-13$, $20-23$). The exception was the HPDI, which lingered behind the others.

In fact, the results indicate that only one variable is required to be normally distributed for most CI estimation methods to perform well (see Tables $16-26$ and Figures $8-9$). The exception was when one of the variables was distributed as $\chi^2(df=1)$. Evens so, this result was consistent regardless of the correlation magnitude (see Tables $16-26$ and Figures $16-19$). In terms of sample size, this result was consistent only when the sample size was 50 or more (see Tables $16-26$ and Figures $28-35$). Again, the HPDI lingered behind the other CI estimation methods.

Second, it was expected that when that when multivariate normality does not hold, coverage probability would be best for the BCa CI followed by the PB and Fisher z-transformation CIs, respectively. Results from the study partially support this expectation. According to the results, coverage probabilities were acceptable for most of the methods for non-normally $(\text{i.e., } |\text{skewness}| \leq .71)$ distributed data when both variables were Beta $(a=4, b=1.25-1.5)$ to $\chi^2(df=16)$ (see Tables $8-10$ and Figure 7). Even so, this result was consistent regardless of the correlation magnitude (see Tables $8-10$ and Figures $14-15$). In terms for sample size, this result was consistent only when sample size was 30 or more (See

Tables $8-10$ and Figures $24-27$). The only caveat is that the CI methods were comparable in this situation. As in previous instances, the HPDI seemed to slightly linger behind.

Third, as the degree of non-normality increases, coverage probability of the Fisher z-transformation and PB CIs will perform worse, while the remaining methods will be robust. This expectation was partially supported as all methods were not robust when both variables were $\chi^2(df=1-3)$ or Pareto $(\text{i.e., } |\text{skewness}| \geq 1.63)$. Additionally, it held if the correlation was greater than 0.1 (see Tables $12-15$ and Figures $14-15$). On the other hand, this held for the Fisher z-transformation CI method regardless of sample size. The other CI methods improved with increased sample size; achieving adequate coverage when the sample size was 200 or greater (see Tables $12-15$ and Figures $24-27$).

Fourth, when the correlation is zero, coverage probability for all CI estimation methods will perform well. In general, the results support this expectation (see Tables $4-26$ and Figure 4). However, sample size and the distribution of the variables had a slight impact. Only the BCa and Fisher z-transformation CIs had adequate coverage across all the sample sizes, and the PB CI and HPDI began to have adequate coverage with a sample size of 100 or more (see Tables $4-26$ and Figure 10). The BCa and PB CIs did not have adequate coverage when both variables were $\chi^2(df=1)$ or Pareto $(\text{i.e., } |\text{skewness}| \geq 2.811)$. Interestingly, the Fisher z-transformation CI always had acceptable coverage when the correlation was zero (see Tables $4-26$ and Figures 12, 14, 16, 18).

Fifth, as the correlation moves away from zero, coverage probability of the Fisher z-transformation and PB CIs will perform worse, while the remaining methods will be robust. In general, the results support this expectation with the exception that the Fisher z-transformation CI was also robust (see Tables $4-26$ and Figure 4). However, sample size and the distribution of the variables had an impact. The Fisher z-transformation and BCa CIs had adequate coverage across all the sample sizes, but the PB CI had adequate coverage when the sample size was 50 or more (see Tables $4-26$ and Figures $10-11$). All methods had inadequate coverage when both variables were $\chi^2\left(df=1-3\right)$ or Pareto $\left(\text{i.e., } |\text{skewness}| \geq 0.71\right)$ and get worse as the correlation increases (see Tables $12-15$ and Figures $14-15$). In all other instances, the Fisher z-transformation, BCa, and PB CIs had adequate coverage, respectively (see Tables $4-11$, $16-26$ and Figures $12-13$, $16-19$). It is interesting to note that the Fisher z-transformation CI method tended to be the most robust unless $|\text{skewness}| \geq 0.71$ for both variables. As before, the HPDI tended to linger behind the other methods.

PART IV

DISCUSSION

A core facet of scientific research is to understand and find truth in nature and/or a phenomenon. This is typically done through identification and replication of an effect. However, replication appears to be a concern as a collaboration of researchers achieved a $36\% - 47\%$ replicate rate of 100 studies they attempted to replicate (Collaboration 2015). One way to address this issue would be to change the criteria in which studies are evaluated. One suggestion is to make the criteria for NHST more stringent (i.e., lowering α; Benjamin, 2017). However, this suggestion may not be entirely viable as modern-day technology makes it feasible to gather large sample sizes to overcome the more stringent criteria (e.g., more power by lowering critical value). Another suggestion to address the replication issue is to utilize effect sizes in conjunction with NHST (Cohen, 1990).

The most fundamental tool in scientific research is NHST as it informs researchers of the presence of a potential effect given H_0. Although useful, NHST is ultimately limited as it offers no information regarding the magnitude and scale of the effect. Researchers can address this limitation by also using effect sizes as they contextualize the magnitude of an effect to a standard scale. Using effect sizes in conjunction with NHST allows researchers a simple way to gauge the practical significance of a statistically significant effect. However, NHST and effect sizes can still be enhanced by providing precision information about the effect size.

A way to provide precision information about a statistic is through interval estimation. The most common way to implement interval estimation is through confidence intervals (CIs). CIs provide precision information about a statistic (or effect). As such, combining CIs with effect sizes allows researchers to gauge an effect in a way that is easily understood, comparable to other research, and gives precision information. Effect size CIs can also be used to conduct NHST. In the current study, the focus was on CIs for the correlation because it can be used as an effect size.

The Pearson product-moment correlation (henceforth, the correlation) is a statistic that is simultaneously well-known yet not entirely understood. It is the baseline method used to determine the linear relationship of two variables and is near ubiquitous due to its ease of use, ease of interpretation, and age (Hald, 2007). Early research conducted on the correlation *t*-test revealed that it was robust when $\rho = 0$ (Blair & Lawson, 1982; Edgell & Noon, 1984). This is not surprising as the sampling distribution of the correlation is a *t* distribution when $\rho = 0$ and the *t*-test is known to be robust to non-normality because of the CLT. These early results were interpreted as the correlation being generally robust to non-normality, which includes $\rho \neq 0$. However, this is not necessarily accurate as the sampling distribution of the correlation is not *t* distributed when $\rho \neq 0$ (see Figure 1). The lack of exploration of ρ raises concerns about its robustness when $\rho \neq 0$.

Understanding the correlation when $\rho = 0$ is meaningful but not appropriate for all applications. One of the more prominent exceptions is when using the correlation as an effect size as effects of 0 (i.e., no effect) are generally not desired in research. Older research typically focused on $\rho = 0$ due to complications in simulating the correlation and the sampling distribution of the correlation being non-normal when $\rho \neq 0$ (see Figure 1). However, advancements in both the literature (Headrick, 2002) and computing power have made it feasible to investigate the correlation when $\rho \neq 0$.

Research on correlation CIs was difficult to do in the past due to the metamorphic nature of the sampling distribution of the correlation. However, the bootstrap method can overcome this challenge as it does not require assumptions of the sampling distribution (Efron & Tibshirani, 1993). Generally speaking, a population parameter and how sample estimates differ from this population are not always known (i.e., the true variance is unknown). In the bootstrap, a sample serves as a psuedo-population to the ESD through repeated sampling with replacement. As such, one can infer how well the ESD represents the sample. By extension, this results in being able to infer how well the sample represents the population. Ultimately, the result is that bootrsap provides an aproximation of the sampling distribution of a population parameter from which a CI can be constructed (esimated).

Two pertinent CIs are the PB and BCa CIs. These are theoretically robust as they also do not require any assumptions regarding the distribution of the statistic. Research by Padilla and Veprinsky (2012, 2014) supports this quality as they found the deanenuated correlation PB and BCa CIs to have good coverage in all the distributions they investigated but the most skewed and kurtotic distributions (e.g., Pareto). However, research by Bishara and Hittner (2017) ran counter to those previous findings. In fact, they found that the PB, BCa, and many other bootstrap CIs are not robust to non-normal data and ultimately concluded that ranked transformation methods were the best way to create correlation CIs. This inconsisntency raises issues on how to best construct correlation CIs. This is important to address given the need to create accurate correlation CIs for applied research and verify the theoretical implications of the bootstrap. The present study aimed to address these inconsistencies such that there is a better understanding of the peformance of the bootstrap CIs.

In the present study, a Monte Carlo simulation was used to generate data with the following conditions: sample size, correlation magnitude, and distribution pairing. Distribution pairing considered the pairwise nature of the variables and dictated four main types of distribution pairings: 1) symmetric with symmetric, 2) non-symmetric with non-symmetric, 3) symmetric with normal, and 4) non-symmetric with normal. For each combination of conditions, the following CIs were estimated: Fisher z-transformation, Spearman rank-order with Fieller's *SE*, RIN, PB, BCa, and HPDI. The primary focus of the present study was the bootstrap CIs (e.g., PB, BCa, and HPDI). However, non-bootstrap CIs (e.g., Fisher z-transformation, Spearman, and RIN) were included for comparison purposes. Standardized bias for the point estimates of the respective CIs were also estimated to supplement CI findings.

To begin, the current study had similar findings with the study conducted by Bishara and Hittner (2017) regarding the Spearman and RIN CIs. The two CIs had consistently good coverage probability across all conditions for both studies. However, caution is advised when using the Spearman and RIN CIs as they both use transformations that irreversibly alter the scale of the data. This means that the Spearman and RIN CIs based on the transformed data may not be representative of the original data. Through ranked transformation, these methods attempt to capture the linear monotonicity of the relationship. However, the linear monotonicity may not be necessarily reflective of the of the actual relationship. An example that comes to mind is a relationship where there is a plateau. As such, some reservation should be used when using and interpreting the Spearman and RIN CIs.

The current study also shared similar findings with Bishara and Hittner (2017) regarding the Fisher z-transformation CI. However, those findings were not discussed in detail in their research. Generally, the Fisher z-transformation CI had excellent coverage probability so long as one of the paired variables was normal. If this was not the case, the coverage probability performance broke down. This indicates that the Fisher z-transformation CI is sensitive to non-normality. In the present study, this is most apparent when a distribution has

$|\text{skewness}| \geq 1.41$. The sensitivity to skew is also made apparent by noticing that the correlation sampling distribution becomes more skewed the closer the correlation gets to one (see Figure 1). It is also worth noting that in these such cases, increases in sample size do not alleviate the situation but instead make it worse. It is therefore unwise to use the Fisher z-transformation CI if the correlated variables are non-normal.

The present study also finds more nuance into the qualities of the PB CI. As it stands, the PB CI does have issues maintaining adequate coverage probability when both variables are non-normal and when sample sizes are small. However, this appears to be a non-issue when sample size $n \geq 200$. This quality was not captured in Bishara and Hittner's study as 160 was the largest sample size they investigated. Additionally, the required sample size for adequate coverage probability is lowered to $n \geq 100$ if there is a normal variable. This is also consistent regardless of correlation magnitude. These findings agree with previous research that suggested bootstrap estimations work better with increased sample size (Rasmussen, 1987).

The present study similarly sheds light onto the qualities of the BCa CI. Like the PB CI, the BCa CI has issues maintaining adequate coverage probability when both variables are non-normal and when sample size is small. This becomes a non-issue when sample size $n \geq 200$. Furthermore, the required sample size for adequate coverage probability is lowered to $n \geq 100$ if there is a normal variable. Again, this quality may have not been captured in Bishara and Hittner's (2017) study as 160 was the largest sample size they investigated. It is worth noting that the BCa CI performed better than the PB CI when sample sizes are smaller, but the PB CI outperformed the BCa CI as sample sizes further increase.

Finally, the current study looked at the HPDI as a research question. The HPDI is defined as the smallest interval that contains the probability of the confidence level ($1-\alpha$). This bootstrap CI was not investigated by either Bishara and Hittner (2017) or Padilla and Verprinksy (2012, 2014) but is of interest due to the principals of its implementation being used for credible intervals for Bayesian methods. Furthermore, CI research is generally limited and the hope with investigating the HPDI was to further expand the pool of knowledge regarding potential CIs available to researchers. It was found that the HPDI tended to have problems maintaining acceptable coverage probability but becomes more stable as sample size increased. When considering only the sample size condition, the HPDI had acceptable coverage probability when $n \geq 100$. However, when considering the sample size and distribution pairing conditions together, the HPDI had acceptable coverage when $n \geq 300$. Of note is that when both variables were non-symmetric, the speed of this convergence depended on the skew of the data such that higher skews resulting in slower convergence (i.e., needed a larger sample). Conversely, this convergence accelerated if one of the paired variables was normal (i.e., needed a smaller sample). Nonetheless, the HPDI never outperformed the PB or BCa CIs.

Standardized bias was also investigated to supplement any findings regarding the correlation CIs. As CIs generally center around an estimate, it was of interest to investigate if poor CI coverage was due to bias in the correlation estimates (e.g., Pearson, Spearman, and RIN correlations). Generally, bias was not an issue as every correlation estimate was within $|\hat{\rho}_{bias}| \leq .40$ (Collins et al., 2001). However, this is not to say that the correlation estimates were not influenced by the simulation conditions. The correlation estimates were impacted as the corresponding standardized bias got closer to ±.40 and become more variable as the correlation magnitude increased (see Figure 36). As such, the correlation CIs struggled because of complications in accounting for the variability in sampling distribution of the correlation and not necessarily because of issues with the point estimate. This means that while the correlation was generally estimated well, there were instances in the variability of the estimate that impacted CI coverage.

Despite the findings gathered in the current study, there are still refinements that could be made to further advance the literature. Like the Bishara and Hittner (2017), the current study had a condition where a normal variable was paired with another variable. In the current study, such pairings typically achieved adequate coverage. In future research, it may be of interest to explore if using a non-normal symmetric variable paired with another variable will have adequate coverage. If this is the case, it would afford researchers more flexibility in utilizing bootstrap CIs. It may also be of interest to combine the bootstrap with the Fisher z-transformation. The Fisher z-transformation CI had excellent coverage performance when at least one variable was normal. Additionally, increased sample size tended to improve coverage performance for the bootstrap CIs. Combining these properties of the Fisher z-transformation and bootstrap may yield CIs that have adequate coverage performance even when both paired variables are highly skewed $\left(\text{i.e., } |\text{skewness}| \geq 1.41\right)$.

Further improvements can be made by expanding the pool of conditions explored. Given that the distribution of the correlation becomes more skewed as it gets closer to one (see Figure 1), it may be fruitful to investigate the bootstrap CI coverage when the correlation is greater than .50 $\left(\text{i.e., } \rho > .50\right)$. Also, the bootstrap CIs generally had acceptable coverage when $n = 200$ with improved performance as sample size increased. Therefore, it may be of interest to explore when exactly coverage is maximized (i.e., reaches a point of diminishing returns) and how larger sample sizes impact the correlation when it is greater than .50. As such, further exploration of the bootstrap may yield more promising results.

In summary, the correlation CIs investigated in the current study generally had adequate coverage probability performance but there are some considerations to keep in my mind. The RIN and Spearman CIs both had consistently good performance across all conditions but risk misinterpretation as they involve irreversibly transforming the original data; which is not an issue for the other CIs. The Fisher z-transformation CI had excellent performance when at least one of the paired variables was normal regardless of the sample size investigated. However, the performance of the Fisher z-transformation CI was shown to break down when the paired variables had $|\text{skewness}| \geq 1.41$ and increasing the sample size made the performance worse. The PB and BCa CIs generally had adequate performance when sample sizes were $n \geq 200$, and the HPDI when $n \geq 300$. The sample size needed for adequate performance for the PB CI, BCa CI, and HPDI reduced to $n \geq 100$ each if one of the paired variables was normal. Additionally, the BCa CI has better performance than the PB CI when sample sizes were smaller and the PB CI has better performance than the BCa CI when sample sizes were larger. Given these findings, one can confidently use the Fisher z-transformation CI with $n \geq 20$ in the following two situations: when one of the paired variables is normal or if the paired variables have $|\text{skewness}| \leq 1$. If the paired variables are non-normal $(\text{i.e., } |\text{skewness}| \geq 1)$, the PB and BCa CIs generally performed well with $n \geq 200$, but the PB is recommended as it had better performance for more extreme non-normal paired variables $(\text{i.e., } |\text{skewness}| \geq 2)$.

REFERENCES

Bandalos, D. L., & Leite, W. (2013). Use of Monte Carlo studies in structural equation modeling research. In Hancock, G.R., & Mueller, R.O. (eds.). Structural equation modeling: A second course (2nd ed., p. 625-666). Information Age Publishing.

Benjamin, D. J., et. al. (2017). Redefine statistical significance. *Nature Human Behavior*. doi:10.1038/s41562-017-0189-z

Berry, K. J., & Mielke, P. W. (2000). A monte carlo investigation of the fisher z transformation for normal and nonnormal distributions. *Psychological Reports, 87*(3), 1101-1114. doi:10.2466/pr0.2000.87.3f.1101

Bishara, A. J., & Hittner, J. B. (2017). Confidence intervals for correlations when data are not normal. *Behav Res, 49*, 294-309. doi:10.3758/s13428-016-0702-8

Blair, R. C., & Lawson, S. B. (1982). Another look at the robustness of the product-moment correlation coefficient to population non-normality. *Florida Journal of Educational Research, 24*, 11-15. Retrieved from https://feraonline.org/

Blanca, M. J., Arnau, J., Lopez-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology, 9*(2), 78-84. doi:10.1027/1614-2241/a000057

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology, 31*(2), 144-152. doi:10.1111/j.2044-8317.1978.tb00581.x

Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Australia ; Pacific Grove, CA: Thomson Learning.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*: Lawrence Erlbaum Associates.

Cohen, J. (1990). Things I have learned (So far). *American Psychologist, 45*(12), 1304-1312. doi: 10.1037/0003-066X.45.12.1304

Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6521). doi:10.1126/science.aac4716

Collins, L. M., Schafer, J. L., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330-351. doi:10.1037//1082-989x.6.4.330

Edgell, S. E., & Noon, S. M. (1984). Effect of violation of normality on the t test of the correlation coefficient. *Psychological Methods, 95*(3), 576-583. doi: 10.1037/0033-2909.95.3.576

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Fieller, E. C., Hartley, H. O., & Pearson, E. S. (1957). Tests for rank correlation coefficients: I. *Biometrika, 44*(470-481). doi: 10.2307/2332878

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in the samples from an infinitely large population. *Biometrika, 10*(4), 507-521. doi:10.2307/2331838

Fisher, R. A. (1929). The statistical method in psychical research. *Proceedings of the Society for Psychical Research, 39*, 182-192. Retrieved from https://www.spr.ac.uk/

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*(2), 156-168. doi: 10.1177/2515245919847202

Hald, A. (2007). *A history of parametric statistical inference from Bernoulli to Fisher, 1713-1935*. New York: Springer.

Headrick, T. C. (2002). Fast fifth-order polynomial transforms for generating univariate and multivariate nonnormal distributions. *Computational Statistics & Data Analysis, 40*(4), 685-711. doi:10.1016/S0167-9473(02)00072-5

Hogg, R. V., Tanis, E. A., & Zimmerman, D. L. (2015). *Probability and statistical inference* (Ninth edition. ed.). Boston: Pearson.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *Public Library of Science Medicine, 2*(8). doi:10.1371/journal.pmed.0020124

Kowalski, C. J. (1972). Effects of Non-Normality on Distribution of Sample Product-Moment Correlation Coefficient. *Journal of the Royal Statistical Society Series C-Applied Statistics, 21*(1), 1-12. Retrieved from https://www.rss.org.uk

Lunneborg, C. E. (1985). Estimating the correlation coefficient: The bootstrap approach. *Psychological Bulletin, 98*(1), 209-215. doi:10.1037/0033-2909.98.1.209

Nickerson, R. S. (2000). Null hypothesis testing: A review of an old and continuing controversy. *Psychological Methods, 5*(2), 241-301. doi:10.1037/1082-989X.5.2.241

Norris, R. C., & Hjelm, H. F. (1961). Nonnormality and product moment correlation. *Journal of Experimental Education, 29*(3), 261-270. Retrieved from https://jstor.org

Padilla, M. A. V., A. (2012). Correlation attenuation due to measurement error: A new approach using the bootstrap procedure. *Educational and Psychological Measurement, 72*(5), 827-846. doi:10.1177/0013164412443963

Padilla, M. A. V., A. (2014). Bootstrapped deattenuated correlation: Nonnormal distributions. *Educational and Psychological Measurement, 74*(5), 823-830. doi:10.1177/0013164414531780

Padilla (2017). Fundamental Statistics for the Social, Behavioral, and Health Sciences. San Diego, CA: Cognella.

Pearson, E. S., & Adyanthaya, N. K. (1929). The distribution of frequency constants in small samples from non-normal symmetrical and skew populations. 2nd Paper: The distribution of "Student's" z. *Biometrika, 21*, 259-286. doi:10.2307/2332560

Perezgonzalez. (2016). Fisher, Neyman-Pearspm pr NHST? A tutorial for teaching data testing. *Frontiers in Psychology, 6*(223). doi:10.3389/fpsyg.2015.00223

Rasmussen, J. L. (1987). Estimating the correlation coefficient: Bootstrap and parametric approaches. *Psychological Bulletin, 101*(1), 136-139. doi:10.1037/0033-2909.101.1.136

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15*(1), 72-101. doi:10.2307/1412159

Sullivan, G. M., & Richard, F. (2012). Using effect size - or why the P value is not enough. *Journal of Graduate Medical Education, 4*(3), 279-282. doi:10.4300/JGME-D-12-00156.1

Tukey, J. W. (1991). The philosophy of multiple comparisons. *Statistical Science, 6*(1), 100-116. doi:10.1214/ss/1177011945

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-vales: Context, process, and purpose. *The American Statistician, 7*(2), 129-133. doi:10.1080/00031305.2016.1154108

Wilkenson, L. (1999). Statistical methods in psychology journals. *American Psychologist, 54*(8), 594-604. doi:10.1037/0003-066X.54.8.594

Yung, Y. F., & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), Advanced structural equation modeling: Issues and techniques. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Zeller, R. A., & Levine, Z. H. (1974). The effects of violating the normality assumption underlying r. *Sociological Methods & Research, 2*(4), 511-519. doi:10.1177/004912417400200406

APPENDIX A

TABLES

Table 1
Logistic Model Effect Sizes for Non-Bootstrap Correlations: CI Coverage

Source	df	Fisher: Cox-Snell r^2	Spearman: Cox-Snell r^2	RIN: Cox-Snell r^2
sample size (n)	10	.0099	.0002	.0001
correlation (ρ)	5	.0091	.0001	.0001
distribution (d)	22	.0030	.0002	.0001
$n \times \rho$	50	.0022	.0000	.0001
$n \times d$	220	.0021	.0000	.0001
$\rho \times d$	110	.0001	.0000	.0000

Note. Fisher = Fisher z-transformation; Spearman = Spearman rank-order; RIN = ranked inverse normal transformation.

Table 2
Logistic Model Effect Sizes for Bootstrap Correlation: CI Coverage

Source	df	PB: Cox-Snell r^2	BCa: Cox-Snell r^2	HPDI: Cox-Snell r^2
sample size (n)	10	.0003	.0014	.0006
correlation (ρ)	5	.0015	.0018	.0036
distribution (d)	22	.0012	.0006	.0030
$n \times \rho$	50	.0000	.0001	.0000
$n \times d$	220	.0000	.0000	.0000
$\rho \times d$	110	.0000	.0000	.0000

Note. PB = percentile bootstrap; BCa = bias-corrected and accelerated bootstrap; HPDI = highest probability density interval.

Table 3
ANOVA Effect Sizes for Standardized Correlation Estimate: Standardized Bias

Source	df	Pearson: η^2	Spearman: η^2	RIN: η^2
sample size (n)	10	0.0000	0.0010	0.0020
correlation (ρ)	5	0.0000	0.0000	0.0000
distribution (d)	22	0.0001	0.0001	0.0001
$n \times \rho$	50	0.0000	0.0006	0.0000
$n \times d$	220	0.0001	0.0000	0.0000
$\rho \times d$	110	0.0001	0.0002	0.0003

Note. Spearman = Spearman rank-order; RIN = ranked inverse normal transformation.

Table 4
95% Coverage Probabilities for paired Normal Distributions
(Skewness = 0, Kurtosis = 0)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.943	.938	.950	.948	.950	.959	.956	.945	.945	.959	.949
	RIN	.945	.937	.948	.943	.948	.960	.948	.942	.947	.956	.953
	SPEAR	.950	.940	.950	.957	.953	.961	.965	.945	.957	.966	.954
	PB	**.913**	**.918**	.938	.941	.942	.947	.950	.939	.946	.956	.953
	BCA	.929	**.918**	.941	.940	.974	.948	.951	.938	.944	.953	.951
	HPDI	**.893**	**.899**	**.916**	.926	.931	.940	.945	.931	.937	.951	.948
0.10	FISHER	.951	.960	.949	.946	.965	.952	.946	.956	.954	.955	.957
	RIN	.950	.954	.948	.952	.964	.952	.952	.954	.958	.953	.952
	SPEAR	.957	.959	.953	.958	.962	.958	.947	.962	.956	.955	.957
	PB	.927	.936	**.920**	.941	.949	.943	.948	.952	.950	.954	.954
	BCA	.935	.941	.930	.940	.952	.946	.949	.949	.952	.957	.954
	HPDI	**.915**	**.920**	**.911**	.925	.942	.939	.942	.948	.942	.953	.951
0.20	FISHER	.953	.952	.958	.954	.952	.951	.947	.944	.949	.953	.952
	RIN	.948	.951	.949	.944	.952	.953	.948	.943	.954	.958	.952
	SPEAR	.951	.952	.957	.947	.953	.959	.950	.953	.957	.955	.948
	PB	**.915**	.925	.940	.935	.952	.946	.941	.940	.948	.951	.950
	BCA	.930	.934	.952	.935	.953	.946	.945	.938	.944	.952	.945
	HPDI	**.892**	**.914**	.925	**.919**	.936	.940	.940	.932	.941	.948	.945
0.30	FISHER	.951	.953	.960	.950	.952	.958	.941	.958	.962	.950	.958
	RIN	.955	.960	.960	.946	.948	.959	.947	.955	.964	.953	.965
	SPEAR	.958	.954	.960	.954	.959	.956	.954	.957	.967	.951	.958
	PB	**.916**	.940	.934	.930	.943	.954	.942	.955	.958	.944	.959
	BCA	.930	.942	.946	.938	.943	.956	.943	.960	.959	.949	.959
	HPDI	**.891**	.926	.929	**.917**	.933	.943	.938	.950	.953	.942	.950
0.40	FISHER	.954	.943	.942	.957	.954	.943	.952	.947	.944	.946	.941
	RIN	.938	.950	.936	.962	.950	.943	.955	.943	.945	.948	.938
	SPEAR	.940	.952	.939	.962	.962	.952	.953	.950	.948	.948	.943
	PB	.930	.933	**.920**	.942	.944	.943	.943	.944	.943	.945	.945
	BCA	.945	.939	.927	.948	.943	.941	.944	.943	.941	.945	.941
	HPDI	**.916**	**.914**	**.909**	.940	.940	.937	.941	.933	.936	.942	.935
0.50	FISHER	.955	.946	.941	.949	.954	.940	.954	.949	.959	.942	.947
	RIN	.958	.943	.940	.950	.956	.940	.959	.951	.962	.947	.950
	SPEAR	.951	.953	.937	.941	.947	.937	.955	.954	.949	.948	.945
	PB	.931	.927	**.921**	.936	.947	.934	.953	.946	.961	.941	.942
	BCA	.941	.936	.926	.946	.948	.937	.953	.950	.962	.944	.940
	HPDI	**.911**	**.910**	**.915**	**.919**	.934	**.922**	.950	.940	.953	.943	.940

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 5
95% Coverage Probabilities for Paired Triangular Distributions
(Skewness = 0, Kurtosis = -.06)

ρ	n	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.943	.936	.943	.953	.950	.947	.946	.954	.946	.953	.945
	RIN	.949	.939	.947	.953	.951	.944	.954	.959	.945	.951	.947
	SPEAR	.949	.939	.952	.950	.963	.953	.960	.958	.950	.955	.948
	PB	**.921**	**.911**	.926	.934	.951	.933	.948	.952	.938	.947	.942
	BCA	.935	**.922**	.933	.934	.954	.940	.953	.957	.940	.951	.942
	HPDI	**.910**	**.896**	**.914**	**.918**	.944	.929	.944	.950	.935	.944	.944
0.10	FISHER	.938	.954	.953	.950	.953	.954	.961	.956	.946	.945	.953
	RIN	.938	.953	.955	.952	.958	.957	.962	.960	.946	.942	.957
	SPEAR	.934	.957	.958	.959	.958	.959	.963	.959	.956	.946	.960
	PB	**.908**	.938	.943	.944	.958	.951	.953	.950	.945	.943	.952
	BCA	.925	.946	.951	.950	.960	.952	.956	.955	.942	.946	.951
	HPDI	**.896**	**.922**	.930	.934	.943	.946	.945	.948	.940	.938	.948
0.20	FISHER	.945	.948	.943	.954	.963	.944	.956	.940	.955	.941	.950
	RIN	.943	.953	.943	.963	.962	.948	.961	.942	.950	.942	.953
	SPEAR	.946	.953	.949	.968	.962	.950	.959	.951	.962	.953	.955
	PB	**.919**	.933	.928	.947	.956	.943	.954	.933	.949	.941	.948
	BCA	.932	.939	.935	.952	.962	.944	.959	.942	.947	.941	.945
	HPDI	**.900**	**.920**	**.916**	.935	.949	.939	.953	.933	.950	.939	.952
0.30	FISHER	.949	.949	.947	.941	.953	.951	.951	.955	.948	.941	.949
	RIN	.956	.944	.949	.947	.948	.953	.954	.957	.946	.939	.951
	SPEAR	.959	.954	.954	.951	.958	.955	.959	.962	.951	.947	.956
	PB	.930	.928	.934	.935	.947	.946	.953	.952	.944	.932	.951
	BCA	.936	.937	.934	.937	.952	.952	.953	.954	.944	.933	.953
	HPDI	**.907**	**.911**	.925	**.924**	.943	.939	.947	.951	.939	.931	.948
0.40	FISHER	.953	.948	.948	.955	.944	.955	.950	.953	.950	.960	.953
	RIN	.956	.945	.950	.947	.949	.959	.951	.951	.949	.964	.950
	SPEAR	.960	.947	.949	.958	.944	.954	.949	.953	.957	.959	.945
	PB	.931	**.921**	.931	.946	.939	.953	.947	.946	.940	.955	.947
	BCA	.947	.931	.946	.949	.943	.956	.946	.947	.943	.954	.950
	HPDI	**.905**	**.901**	**.920**	.933	.935	.946	.945	.944	.934	.953	.944
0.50	FISHER	.945	.946	.952	.935	.954	.943	.950	.944	.945	.960	.954
	RIN	.953	.944	.958	.939	.952	.945	.954	.948	.947	.959	.951
	SPEAR	.958	.948	.946	.937	.944	.943	.953	.942	.950	.956	.948
	PB	.937	**.921**	.936	.933	.948	.942	.954	.938	.946	.960	.948
	BCA	.948	.933	.945	.938	.952	.939	.957	.942	.949	.958	.952
	HPDI	**.912**	**.909**	**.920**	**.924**	.936	.937	.950	.930	.940	.956	.948

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 6
95% Coverage Probabilities for Paired Uniform Distributions
(Skewness = 0, Kurtosis = -1.20)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.947	.952	.945	.949	.953	.945	.956	.949	.954	.948	.950
	RIN	.950	.953	.955	.953	.952	.944	.952	.949	.954	.949	.948
	SPEAR	.956	.960	.957	.949	.958	.955	.960	.956	.958	.958	.957
	PB	**.916**	.942	.934	.936	.949	.946	.955	.947	.949	.946	.946
	BCA	.934	.951	.945	.949	.954	.947	.957	.947	.952	.949	.947
	HPDI	**.896**	**.922**	**.921**	**.924**	.937	.940	.951	.949	.950	.945	.942
0.10	FISHER	.935	.951	.957	.941	.951	.951	.947	.953	.951	.938	.950
	RIN	.950	.950	.961	.946	.946	.950	.946	.946	.941	.949	.947
	SPEAR	.947	.951	.960	.949	.957	.961	.953	.955	.957	.943	.959
	PB	**.906**	.928	.942	.931	.951	.953	.946	.951	.951	.934	.949
	BCA	.926	.943	.958	.943	.955	.961	.951	.953	.949	.934	.952
	HPDI	**.880**	**.912**	.930	**.923**	.940	.946	.945	.944	.945	.933	.946
0.20	FISHER	.945	.945	.946	.959	.952	.944	.942	.940	.956	.957	.944
	RIN	.943	.943	.944	.961	.959	.942	.946	.942	.953	.956	.954
	SPEAR	.949	.946	.955	.963	.959	.950	.950	.943	.959	.964	.950
	PB	**.920**	.933	.930	.948	.950	.938	.936	.942	.955	.954	.943
	BCA	.941	.948	.949	.962	.955	.943	.943	.942	.953	.954	.943
	HPDI	**.898**	**.918**	**.916**	.934	.943	.931	.932	.930	.948	.952	.941
0.30	FISHER	.942	.944	.948	.944	.944	.941	.952	.955	.949	.958	.946
	RIN	.949	.957	.960	.948	.943	.952	.957	.953	.948	.963	.956
	SPEAR	.946	.956	.959	.949	.942	.950	.954	.960	.956	.960	.953
	PB	**.913**	.929	.937	.937	.942	.937	.960	.954	.951	.956	.949
	BCA	.945	.949	.954	.947	.949	.940	.958	.957	.952	.961	.947
	HPDI	**.895**	**.920**	.929	.926	.937	.934	.956	.949	.941	.955	.944
0.40	FISHER	.937	.939	.959	.943	.956	.940	.959	.945	.946	.949	.950
	RIN	.956	.949	.955	.949	.954	.946	.956	.956	.959	.955	.952
	SPEAR	.946	.945	.959	.949	.956	.946	.966	.947	.950	.955	.956
	PB	**.916**	**.913**	.952	.938	.950	.939	.957	.943	.950	.947	.952
	BCA	.941	.932	.960	.945	.958	.945	.963	.947	.951	.948	.954
	HPDI	**.899**	**.890**	.940	.928	.940	.925	.957	.938	.948	.947	.949
0.50	FISHER	.932	.945	.953	.940	.934	.958	.946	.951	.948	.949	.943
	RIN	.950	.959	.959	.953	.957	.964	.957	.964	.956	.959	.960
	SPEAR	.944	.953	.955	.945	.943	.964	.956	.956	.951	.947	.946
	PB	**.906**	.931	.937	.934	.929	.957	.950	.950	.947	.949	.948
	BCA	.933	.948	.952	.944	.937	.964	.955	.955	.946	.952	.946
	HPDI	**.886**	.927	**.924**	**.922**	**.924**	.954	.948	.948	.946	.946	.935

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 7
95% Coverage Probabilities for Paired Laplace Distributions
(Skewness = 0, Kurtosis = 3)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.945	.940	.960	.938	.945	.949	.962	.935	.929	.943	.952
	RIN	.952	.941	.955	.934	.954	.947	.961	.936	.942	.950	.949
	SPEAR	.951	.946	.961	.947	.963	.953	.956	.954	.955	.958	.958
	PB	**.911**	**.910**	.937	**.923**	.930	.938	.957	.937	.928	.947	.947
	BCA	.927	**.919**	.932	**.917**	**.917**	.926	.954	.931	**.922**	.943	.944
	HPDI	**.891**	**.887**	**.922**	**.912**	**.920**	.933	.952	.929	**.917**	.940	.945
0.10	FISHER	.938	.946	.947	.953	.944	.961	.942	.940	.949	.952	.960
	RIN	.939	.952	.950	.951	.953	.953	.945	.948	.959	.951	.957
	SPEAR	.942	.957	.960	.958	.967	.955	.955	.958	.956	.961	.958
	PB	**.912**	.939	.926	.932	.926	.947	.941	.942	.951	.952	.958
	BCA	**.916**	.938	.925	**.923**	**.921**	.939	.934	.937	.946	.944	.952
	HPDI	**.879**	**.924**	**.909**	**.921**	**.922**	.944	.934	.930	.945	.945	.951
0.20	FISHER	.949	.942	.948	.947	.951	.950	.947	.950	.955	.939	.945
	RIN	.955	.947	.949	.953	.960	.953	.946	.948	.958	.939	.948
	SPEAR	.956	.956	.953	.953	.962	.963	.953	.948	.958	.946	.951
	PB	**.917**	**.916**	**.924**	**.923**	.939	.939	.939	.941	.952	.940	.942
	BCA	.925	**.917**	**.921**	.925	.936	.934	.931	.938	.945	.933	.945
	HPDI	**.888**	**.897**	**.911**	**.915**	.935	.934	.938	.941	.947	.938	.938
0.30	FISHER	.941	.955	.951	.948	.948	.960	.941	.937	.936	.956	.944
	RIN	.941	.942	.954	.957	.948	.957	.941	.940	.942	.958	.946
	SPEAR	.939	.947	.962	.961	.950	.953	.950	.945	.950	.957	.949
	PB	**.921**	.926	.940	.935	.944	.946	.933	.932	.938	.963	.940
	BCA	**.916**	**.920**	.943	.931	.933	.938	.931	**.920**	.937	.958	.939
	HPDI	**.894**	**.917**	.928	**.924**	.937	.939	**.925**	**.923**	.930	.962	.938
0.40	FISHER	.950	.952	.955	.949	.939	.944	.943	.947	.941	.933	.939
	RIN	.943	.961	.956	.955	.946	.950	.950	.952	.947	.946	.941
	SPEAR	.946	.954	.954	.963	.942	.948	.947	.953	.952	.950	.946
	PB	**.921**	**.919**	.935	.945	.933	.944	.944	.947	.942	.933	.946
	BCA	**.915**	**.923**	.930	.934	.927	.940	.936	.940	.937	.928	.945
	HPDI	**.902**	**.907**	**.921**	.930	**.923**	.937	.942	.939	.935	.930	.940
0.50	FISHER	.948	.934	.934	**.924**	.935	.949	**.915**	.951	.949	.937	.941
	RIN	.960	.939	.946	.938	.945	.964	.932	.962	.960	.951	.951
	SPEAR	.956	.943	.946	.936	.936	.961	.935	.951	.951	.956	.942
	PB	.937	**.919**	.928	**.906**	.938	.950	**.918**	.949	.954	.946	.943
	BCA	.929	**.903**	**.923**	**.905**	.935	.938	**.910**	.944	.950	.939	.939
	HPDI	**.905**	**.896**	**.920**	**.896**	.932	.944	**.909**	.944	.951	.936	.940

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 8
95% Coverage Probabilities for Paired Beta Distributions
(a = 4, b = 1.25, Skewness = -.848, Kurtosis = .221

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.957	.951	.941	.955	.940	.952	.939	.959	.960	.947	.945
	RIN	.954	.949	.945	.958	.942	.956	.940	.951	.965	.953	.946
	SPEAR	.954	.968	.947	.963	.950	.951	.944	.957	.966	.953	.955
	PB	.926	.935	**.920**	.935	.943	.948	.929	.957	.960	.946	.943
	BCA	.945	.945	.929	.935	.940	.953	.932	.959	.959	.946	.943
	HPDI	**.905**	**.911**	**.905**	.925	.928	.940	**.920**	.946	.955	.940	.937
0.10	FISHER	.944	.941	.941	.942	.943	.955	.938	.945	.943	.952	.937
	RIN	.950	.943	.940	.952	.953	.954	.948	.948	.949	.946	.950
	SPEAR	.944	.949	.947	.960	.953	.957	.948	.952	.960	.954	.958
	PB	**.909**	**.911**	.933	.929	.940	.946	.933	.945	.950	.952	.936
	BCA	**.924**	**.922**	.939	.929	.942	.952	.939	.941	.943	.951	.940
	HPDI	**.878**	**.897**	**.920**	**.914**	.936	.944	.930	.934	.938	.947	.931
0.20	FISHER	.942	.946	.949	.952	.940	.931	.953	.951	.934	.953	.941
	RIN	.954	.948	.953	.962	.944	.935	.957	.956	.950	.962	.951
	SPEAR	.960	.949	.955	.966	.952	.953	.954	.962	.953	.959	.953
	PB	.928	.928	.944	.939	.942	.933	.951	.959	.944	.961	.947
	BCA	.942	.938	.942	.948	.939	.933	.951	.959	.947	.962	.950
	HPDI	**.904**	**.910**	.928	.928	.937	.926	.942	.952	.937	.955	.943
0.30	FISHER	**.917**	.930	.931	.942	.932	.927	.929	.941	.943	.933	.935
	RIN	.941	.940	.941	.956	.946	.944	.949	.951	.954	.950	.945
	SPEAR	.947	.960	.948	.955	.948	.949	.954	.952	.952	.955	.952
	PB	**.911**	.930	**.923**	.939	.937	.935	.944	.956	.951	.949	.947
	BCA	**.913**	.935	.928	.945	.937	.932	.945	.953	.948	.943	.946
	HPDI	**.873**	**.911**	**.911**	**.924**	.930	.928	.936	.949	.944	.947	.940
0.40	FISHER	.939	.943	.947	**.918**	.933	.929	.938	.935	.932	.928	.944
	RIN	.958	.947	.957	.945	.948	.950	.954	.946	.942	.940	.960
	SPEAR	.950	.954	.956	.942	.951	.948	.961	.944	.930	.943	.955
	PB	**.918**	.929	.940	**.918**	.938	.940	.949	.947	.944	.935	.950
	BCA	.934	.928	.943	**.918**	.935	.942	.946	.947	.947	.935	.954
	HPDI	**.902**	**.904**	.927	**.905**	.932	.937	.948	.943	.938	.927	.946
0.50	FISHER	.936	.926	**.920**	.933	**.924**	**.923**	.933	.932	.935	.929	.931
	RIN	.951	.959	.949	.952	.957	.948	.942	.951	.941	.944	.943
	SPEAR	.959	.955	.946	.954	.959	.952	.941	.953	.951	.944	.942
	PB	.926	**.919**	.925	.927	.942	.930	.952	.940	.944	.948	.941
	BCA	.937	.929	.925	.926	.942	.933	.949	.937	.945	.948	.944
	HPDI	**.909**	**.907**	**.910**	**.916**	.938	.925	.944	.936	.940	.943	.935

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 9
95% Coverage Probabilities for Paired Beta Distributions
(a = 4, b = 1.5, Skewness = -.694, Kurtosis = -.069)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.941	.950	.950	.940	.948	.957	.943	.945	.932	.943	.948
	RIN	.948	.947	.951	.944	.941	.956	.945	.943	.942	.947	.956
	SPEAR	.951	.950	.960	.953	.952	.956	.957	.944	.953	.950	.959
	PB	**.922**	.932	.943	.942	.941	.951	.938	.937	.933	.938	.944
	BCA	.936	.939	.946	.940	.937	.956	.941	.939	.931	.939	.946
	HPDI	**.900**	**.910**	.927	**.922**	.935	.943	.930	.937	.930	.935	.941
0.10	FISHER	.945	.943	.961	.943	.950	.958	.937	.966	.950	.955	.957
	RIN	.955	.946	.960	.955	.946	.957	.942	.968	.947	.944	.949
	SPEAR	.958	.947	.961	.955	.956	.960	.943	.965	.958	.959	.959
	PB	**.921**	**.915**	.948	.941	.944	.956	.933	.958	.950	.957	.955
	BCA	.936	.927	.952	.940	.951	.956	.935	.959	.953	.958	.956
	HPDI	**.898**	**.899**	.936	.927	.937	.949	.934	.957	.948	.950	.949
0.20	FISHER	.952	.941	.938	.938	.953	.948	.941	.940	.948	.944	.949
	RIN	.953	.950	.940	.946	.956	.957	.944	.950	.957	.945	.955
	SPEAR	.964	.952	.942	.955	.955	.963	.948	.960	.959	.950	.968
	PB	**.924**	.932	.928	.934	.956	.951	.939	.946	.950	.941	.950
	BCA	.931	.940	.931	.931	.947	.951	.937	.945	.953	.940	.950
	HPDI	**.899**	**.916**	**.915**	**.920**	.944	.945	.931	.935	.948	.930	.948
0.30	FISHER	.941	.938	.939	.944	.942	.940	.941	.948	.929	.932	.940
	RIN	.944	.948	.937	.950	.949	.952	.940	.957	.953	.947	.952
	SPEAR	.946	.944	.939	.948	.953	.961	.949	.958	.953	.954	.955
	PB	**.923**	.931	.931	.948	.937	.949	.948	.956	.948	.946	.952
	BCA	.935	.943	.936	.949	.943	.953	.955	.958	.946	.946	.950
	HPDI	**.904**	**.910**	**.917**	.934	.933	.945	.943	.951	.940	.940	.948
0.40	FISHER	.944	.946	.937	.933	.940	.931	.947	.935	.945	.944	.929
	RIN	.946	.955	.949	.955	.951	.952	.958	.940	.945	.957	.950
	SPEAR	.951	.957	.952	.945	.948	.955	.957	.934	.940	.950	.950
	PB	.929	.932	.931	.931	.943	.945	.952	.938	.947	.947	.940
	BCA	.941	.942	.936	.933	.941	.937	.952	.939	.951	.954	.941
	HPDI	**.902**	**.916**	**.914**	**.920**	.937	.930	.942	.933	.940	.943	.937
0.50	FISHER	.938	.943	.945	.935	.938	.929	.949	.940	.942	.935	.931
	RIN	.952	.946	.962	.949	.955	.943	.963	.954	.949	.950	.941
	SPEAR	.950	.946	.954	.950	.953	.938	.954	.950	.955	.947	.943
	PB	.925	.935	.936	.936	.941	.935	.949	.945	.953	.944	.943
	BCA	.935	.938	.932	.941	.939	.935	.951	.948	.952	.944	.946
	HPDI	**.898**	**.911**	**.924**	.926	.938	.929	.945	.946	.944	.939	.940

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 10
95% Coverage Probabilities for Paired Chi-Square Distributions
(df = 16, Skewness = .71, Kurtosis = .75)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.942	.955	.954	.951	.945	.931	.960	.948	.946	.944	.949
	RIN	.942	.957	.952	.952	.948	.935	.957	.951	.953	.947	.949
	SPEAR	.939	.962	.959	.961	.957	.944	.958	.963	.962	.943	.958
	PB	**.895**	.933	.942	.942	.940	.926	.961	.942	.937	.937	.948
	BCA	**.910**	.942	.947	.943	.941	**.924**	.962	.940	.938	.937	.947
	HPDI	**.875**	**.921**	.927	.930	.930	**.920**	.955	.936	.932	.929	.944
0.10	FISHER	.950	.954	.953	.949	.937	.954	.951	.946	.955	.943	.953
	RIN	.951	.953	.954	.957	.945	.958	.954	.955	.952	.957	.952
	SPEAR	.956	.964	.957	.957	.953	.957	.960	.955	.960	.961	.958
	PB	.935	.943	.945	.942	.934	.946	.951	.936	.958	.945	.959
	BCA	.949	.948	.943	.938	.930	.949	.952	.936	.959	.940	.956
	HPDI	**.916**	.925	.936	.931	.929	.944	.949	.934	.948	.939	.954
0.20	FISHER	.949	.952	.943	.947	.950	.937	.951	.932	.941	.941	.942
	RIN	.953	.953	.948	.947	.950	.941	.954	.934	.946	.946	.943
	SPEAR	.962	.960	.949	.956	.958	.958	.949	.950	.958	.953	.953
	PB	.932	.933	.933	.937	.947	.932	.953	.938	.940	.939	.944
	BCA	.942	.938	.931	.937	.941	.929	.956	.938	.940	.938	.938
	HPDI	**.912**	**.923**	**.918**	.925	.945	.931	.946	.928	.939	.928	.940
0.30	FISHER	.939	.955	.941	.944	.942	**.928**	.937	.946	.933	.937	.936
	RIN	.944	.951	.947	.953	.943	.937	.938	.949	.945	.935	.938
	SPEAR	.945	.955	.954	.958	.942	.936	.952	.960	.953	.950	.948
	PB	**.923**	.937	**.924**	.930	.935	**.922**	.936	.945	.934	.938	.940
	BCA	.926	.944	**.924**	.935	.943	.928	.934	.946	.933	.937	.941
	HPDI	**.895**	**.919**	**.904**	.925	.926	**.918**	.930	.935	.932	.938	.937
0.40	FISHER	.940	.950	.935	.960	.949	.935	.948	.945	.943	.946	.948
	RIN	.952	.941	.948	.959	.949	.946	.951	.956	.952	.959	.957
	SPEAR	.951	.945	.953	.957	.951	.953	.960	.953	.956	.954	.954
	PB	.929	.929	**.924**	.943	.951	.950	.958	.942	.954	.953	.954
	BCA	.940	.926	.926	.945	.947	.949	.954	.934	.953	.949	.953
	HPDI	**.903**	**.908**	**.911**	.933	.943	.943	.950	.937	.947	.943	.952
0.50	FISHER	.932	.925	.951	.946	.938	.936	.944	.950	.937	.947	.949
	RIN	.939	.951	.959	.959	.955	.942	.948	.957	.941	.954	.948
	SPEAR	.940	.958	.958	.954	.956	.938	.943	.946	.934	.950	.956
	PB	**.906**	**.917**	.943	.939	.945	.939	.947	.950	.945	.951	.953
	BCA	.933	**.919**	.936	.944	.941	.938	.947	.950	.943	.949	.953
	HPDI	**.880**	**.904**	.927	.929	.937	.938	.945	.945	.940	.950	.953

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 11
95% Coverage Probabilities for Paired Chi-Square Distributions
(df = 4, Skewness = 1.41, Kurtosis = 3)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.941	.944	.949	.944	.951	.938	.949	.953	.942	.944	.953
	RIN	.937	.941	.957	.937	.949	.941	.949	.946	.948	.940	.942
	SPEAR	.946	.939	.958	.936	.956	.956	.948	.954	.959	.949	.946
	PB	.926	.925	.931	**.912**	.935	.938	.948	.942	.942	.942	.947
	BCA	.932	.928	.932	**.919**	.931	.925	.943	.939	.935	.935	.945
	HPDI	**.904**	**.908**	**.916**	**.901**	**.924**	.925	.941	.936	.935	.937	.944
0.10	FISHER	.950	.943	.940	.950	.947	.954	.945	.953	.942	.934	.943
	RIN	.954	.951	.952	.961	.948	.964	.957	.957	.943	.950	.947
	SPEAR	.958	.946	.950	.962	.951	.961	.966	.964	.950	.957	.960
	PB	.927	.925	.932	.935	.946	.957	.945	.955	.951	.942	.943
	BCA	.945	.936	.925	.932	.940	.948	.935	.950	.948	.934	.936
	HPDI	**.907**	**.910**	**.918**	**.917**	.934	.947	.941	.947	.941	.934	.938
0.20	FISHER	.929	.932	.941	.940	.946	.926	.940	.935	**.921**	.933	.942
	RIN	.930	.945	.951	.956	.956	.948	.950	.955	.945	.951	.952
	SPEAR	.934	.950	.959	.966	.954	.962	.956	.959	.950	.958	.953
	PB	**.912**	**.921**	.930	.946	.943	.933	.943	.947	.932	.943	.954
	BCA	**.922**	**.920**	**.919**	.938	.943	**.924**	.941	.947	.925	.940	.955
	HPDI	**.894**	**.903**	**.914**	.927	.935	**.921**	.939	.944	.927	.940	.952
0.30	FISHER	.927	.927	**.923**	.937	.932	**.913**	**.906**	**.909**	**.922**	.933	**.917**
	RIN	.940	.952	.943	.952	.949	.953	.949	.948	.942	.956	.956
	SPEAR	.944	.957	.952	.956	.945	.959	.946	.946	.961	.963	.956
	PB	**.917**	.932	.927	.932	.946	.933	.932	.933	.943	.953	.938
	BCA	.926	.926	**.924**	**.922**	.946	.927	.931	**.922**	.935	.949	.937
	HPDI	**.899**	**.912**	**.905**	**.919**	.942	**.922**	.927	.928	.936	.944	.933
0.40	FISHER	.926	**.917**	**.905**	**.921**	.926	**.917**	**.920**	**.910**	**.910**	**.903**	**.923**
	RIN	.959	.945	.945	.945	.964	.951	.949	.952	.944	.951	.950
	SPEAR	.962	.949	.943	.944	.959	.941	.944	.945	.954	.954	.949
	PB	.926	**.918**	.927	.933	.946	.936	.949	.935	.939	.940	.951
	BCA	**.923**	**.912**	**.909**	**.919**	.933	.927	.937	.934	.936	.929	.947
	HPDI	**.905**	**.891**	**.904**	**.916**	.937	.928	.943	.933	.934	.934	.947
0.50	FISHER	**.924**	**.905**	.907	**.923**	.930	**.896**	**.894**	**.910**	**.892**	**.911**	**.905**
	RIN	.958	.942	.940	.954	.956	.952	.955	.939	.959	.961	.945
	SPEAR	.951	.943	.950	.946	.954	.949	.956	.941	.949	.955	.956
	PB	.938	**.916**	.929	.939	.947	.929	.936	.941	.939	.945	.942
	BCA	.927	**.903**	**.918**	**.924**	.944	.927	**.920**	.933	.932	.939	.938
	HPDI	**.903**	**.899**	**.907**	.928	.941	**.920**	**.923**	.938	.934	.940	.939

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 12
95% Coverage Probabilities for Paired Chi-Square Distributions
(df = 3, Skewness = 1.63, Kurtosis = 4)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.960	.952	.950	.937	.948	.950	.955	.951	.929	.962	.958
	RIN	.950	.953	.940	.936	.955	.948	.954	.953	.935	.956	.946
	SPEAR	.956	.959	.945	.947	.962	.954	.956	.955	.949	.955	.950
	PB	**.915**	**.917**	**.922**	**.924**	.931	.934	.933	.943	.929	.955	.945
	BCA	.944	.933	.934	**.919**	.932	.931	.935	.941	**.919**	.950	.942
	HPDI	**.895**	**.900**	**.910**	**.904**	**.919**	.930	.926	.937	**.923**	.949	.945
0.10	FISHER	.931	.932	.947	.930	.953	.926	.946	.930	.943	.948	.943
	RIN	.946	.940	.958	.938	.954	.946	.961	.949	.949	.956	.944
	SPEAR	.946	.951	.970	.953	.958	.953	.965	.950	.955	.959	.949
	PB	**.913**	**.918**	**.923**	**.919**	.936	.930	.950	.942	.941	.951	.941
	BCA	.925	.928	**.923**	**.920**	.937	**.924**	.942	.937	.944	.948	.942
	HPDI	**.890**	**.902**	**.907**	**.903**	.929	.926	.941	.937	.934	.941	.935
0.20	FISHER	.931	**.907**	.930	**.924**	.925	.925	**.915**	.929	.926	.932	.942
	RIN	.946	.936	.954	.947	.958	.953	.939	.956	.957	.955	.962
	SPEAR	.949	.953	.957	.954	.958	.958	.943	.967	.959	.962	.965
	PB	**.918**	**.911**	.939	.932	.941	.942	.930	.943	.950	.948	.956
	BCA	.935	**.914**	.932	**.919**	.933	.938	**.924**	.937	.943	.941	.949
	HPDI	**.896**	**.895**	**.922**	**.917**	.933	.935	**.919**	.934	.944	.942	.947
0.30	FISHER	.943	.930	**.920**	**.920**	**.907**	**.908**	**.916**	**.900**	**.922**	**.897**	.923
	RIN	.955	.954	.944	.943	.946	.942	.946	.942	.949	.939	.950
	SPEAR	.958	.963	.948	.952	.954	.941	.940	.934	.959	.941	.949
	PB	.925	.931	.931	.932	.925	.929	.947	.935	.950	.935	.941
	BCA	.929	.928	**.920**	**.917**	**.917**	**.919**	.937	.933	.944	.927	.942
	HPDI	**.900**	**.913**	**.902**	**.911**	**.913**	**.924**	.939	.933	.949	.926	.941
0.40	FISHER	**.913**	**.917**	.903	**.912**	**.916**	**.887**	**.912**	**.901**	**.902**	**.899**	**.897**
	RIN	.933	.962	.960	.954	.950	.941	.954	.946	.950	.945	.954
	SPEAR	.937	.961	.955	.953	.946	.946	.946	.953	.948	.944	.953
	PB	**.920**	.927	.935	.933	.935	**.915**	.941	.927	.941	.937	.945
	BCA	**.909**	**.913**	.926	**.922**	.926	**.907**	.930	**.924**	.938	.933	.936
	HPDI	**.896**	**.905**	**.921**	**.919**	.925	**.909**	.935	**.922**	.938	.930	.941
0.50	FISHER	**.893**	**.910**	**.898**	**.897**	**.892**	**.877**	**.899**	**.910**	**.895**	**.872**	**.886**
	RIN	.945	.946	.954	.954	.958	.937	.939	.944	.961	.950	.953
	SPEAR	.943	.937	.952	.949	.957	.950	.936	.944	.950	.943	.953
	PB	**.904**	.934	**.917**	.932	.942	**.909**	.939	.957	.941	.936	.945
	BCA	**.906**	**.922**	**.911**	**.910**	**.917**	**.896**	.932	.945	.936	**.923**	.934
	HPDI	**.874**	**.912**	**.896**	**.907**	.930	**.905**	.930	.948	.934	**.928**	.933

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 13
95% Coverage Probabilities for Paired Chi-Square Distributions
(df = 2, Skewness = 2, Kurtosis = 6)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.939	.954	.944	.953	.956	.955	.950	.961	.959	.952	.951
	RIN	.939	.957	.954	.945	.948	.954	.959	.955	.954	.932	.956
	SPEAR	.945	.967	.962	.953	.957	.958	.965	.959	.967	.946	.953
	PB	**.918**	.930	.927	**.918**	.929	.942	.946	.941	.943	.939	.937
	BCA	.942	.937	.933	**.924**	.934	.941	.939	.948	.941	.938	.937
	HPDI	**.888**	**.903**	**.906**	**.910**	**.917**	.933	.938	.932	.940	.936	**.924**
0.10	FISHER	.943	.947	.937	.940	.935	.931	.925	.931	.935	.939	**.912**
	RIN	.948	.959	.942	.950	.958	.946	.943	.942	.959	.949	.946
	SPEAR	.954	.959	.950	.955	.955	.948	.954	.945	.958	.958	.945
	PB	**.919**	**.915**	.927	.939	.930	.928	.934	.932	.949	.940	.930
	BCA	.934	**.913**	.928	.942	**.922**	**.922**	**.924**	.930	.943	.936	**.924**
	HPDI	**.893**	**.893**	**.906**	.929	**.911**	**.917**	.925	.929	.939	.930	.925
0.20	FISHER	**.916**	.930	.940	**.915**	**.908**	**.911**	.925	**.904**	**.907**	**.900**	**.916**
	RIN	.940	.955	.953	.943	.951	.940	.960	.958	.937	.951	.951
	SPEAR	.938	.952	.964	.960	.959	.956	.958	.960	.941	.962	.964
	PB	**.918**	.931	.932	**.916**	**.921**	.935	.939	.935	.933	.941	.945
	BCA	.925	**.921**	.926	**.906**	**.915**	.929	.930	.932	.932	.935	.943
	HPDI	**.886**	**.916**	**.914**	**.898**	**.910**	**.924**	.931	.928	.926	.929	.940
0.30	FISHER	**.915**	**.903**	**.887**	**.918**	**.887**	**.891**	**.902**	**.893**	**.902**	**.893**	**.884**
	RIN	.940	.932	.943	.949	.959	.949	.942	.954	.954	.939	.950
	SPEAR	.954	.928	.943	.955	.951	.955	.953	.956	.951	.952	.955
	PB	.925	.909	**.922**	.937	.925	.936	.947	.945	.945	.929	.949
	BCA	.925	**.895**	**.905**	**.918**	**.913**	.925	.940	.936	.938	**.922**	.940
	HPDI	**.892**	**.890**	**.903**	**.921**	**.914**	**.917**	.940	.937	.943	**.923**	.943
0.40	FISHER	**.906**	**.893**	**.903**	**.875**	**.892**	**.870**	**.873**	**.881**	**.853**	**.866**	**.902**
	RIN	.952	.942	.958	.936	.953	.943	.942	.953	.946	.948	.955
	SPEAR	.951	.943	.958	.936	.957	.947	.949	.955	.944	.951	.947
	PB	**.920**	**.924**	.928	**.919**	.929	**.920**	.937	.943	**.914**	.926	.951
	BCA	**.923**	**.919**	**.915**	**.896**	**.921**	**.903**	**.918**	.942	**.907**	**.917**	.948
	HPDI	**.891**	**.904**	**.914**	**.895**	**.914**	**.912**	.927	.937	**.912**	**.917**	.945
0.50	FISHER	**.902**	**.881**	**.877**	**.878**	**.864**	**.882**	**.874**	**.881**	**.864**	**.874**	**.853**
	RIN	.945	.937	.948	.957	.949	.951	.963	.945	.942	.944	.953
	SPEAR	.935	.946	.941	.944	.952	.954	.956	.946	.945	.952	.949
	PB	.926	.926	**.924**	.942	.925	.938	.951	.946	.940	.952	.942
	BCA	**.912**	**.904**	**.907**	**.923**	**.902**	**.922**	.929	.933	.933	.938	.927
	HPDI	**.899**	**.901**	**.894**	**.920**	**.910**	.928	.935	.933	.934	.945	.933

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 14
95% Coverage Probabilities for paired Chi-Square Distributions
(df = 1, Skewness = 2.83, Kurtosis = 12)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.943	.943	.947	.956	.957	.951	.961	.960	.962	.945	.950
	RIN	.939	.945	.950	.933	.950	.940	.947	.947	.959	.956	.957
	SPEAR	.946	.963	.961	.948	.955	.946	.947	.958	.958	.968	.963
	PB	**.879**	**.887**	**.896**	**.888**	**.923**	**.910**	.930	.933	**.920**	.932	.926
	BCA	**.916**	**.916**	**.907**	**.896**	**.921**	**.919**	.942	.936	.931	.931	.931
	HPDI	**.856**	**.863**	**.872**	**.863**	**.910**	**.895**	**.922**	**.916**	**.915**	.926	**.917**
0.10	FISHER	**.914**	.927	.943	**.919**	**.913**	**.904**	**.913**	**.913**	**.914**	**.912**	**.912**
	RIN	.947	.950	.951	.949	.950	.944	.944	.948	.959	.959	.937
	SPEAR	.955	.950	.952	.958	.948	.946	.947	.956	.954	.965	.946
	PB	**.907**	**.909**	**.911**	**.912**	**.924**	.930	.929	**.924**	.945	.934	.935
	BCA	.928	**.903**	**.920**	**.908**	**.914**	.932	.926	**.921**	.933	.927	.931
	HPDI	**.880**	**.874**	**.889**	**.885**	**.916**	**.909**	**.919**	**.914**	.935	.926	.929
0.20	FISHER	**.889**	**.901**	**.895**	**.879**	**.880**	**.895**	**.898**	**.897**	**.861**	**.887**	**.884**
	RIN	.948	.951	.942	.953	.946	.948	.955	.953	.951	.955	.946
	SPEAR	.952	.956	.945	.957	.953	.950	.956	.960	.958	.960	.941
	PB	**.906**	**.916**	**.912**	**.911**	**.916**	.925	.943	.935	.928	.944	.948
	BCA	**.916**	**.907**	**.919**	**.898**	**.897**	**.916**	.931	.935	.927	.934	.938
	HPDI	**.876**	**.885**	**.898**	**.886**	**.894**	**.915**	.934	**.923**	**.920**	.938	.938
0.30	FISHER	**.863**	**.879**	**.875**	**.880**	**.859**	**.835**	**.863**	**.847**	**.835**	**.826**	**.821**
	RIN	.952	.961	.953	.940	.950	.953	.957	.955	.943	.957	.961
	SPEAR	.959	.958	.960	.941	.957	.952	.957	.956	.940	.958	.960
	PB	**.919**	**.922**	**.919**	**.919**	.929	.926	.944	.941	.930	.935	.942
	BCA	**.915**	**.894**	**.898**	**.901**	**.897**	**.912**	.927	.932	**.921**	**.921**	.926
	HPDI	**.885**	**.886**	**.899**	**.902**	**.909**	**.908**	.931	.935	**.921**	.928	.932
0.40	FISHER	**.848**	**.851**	**.834**	**.821**	**.829**	**.827**	**.804**	**.807**	**.805**	**.799**	**.814**
	RIN	.953	.959	.943	.948	.953	.949	.952	.962	.959	.956	.945
	SPEAR	.947	.956	.946	.937	.943	.940	.948	.957	.960	.961	.960
	PB	**.922**	.927	**.924**	.926	**.922**	.933	.930	.932	**.921**	.940	**.924**
	BCA	**.901**	**.903**	**.905**	**.893**	**.906**	**.913**	**.904**	**.917**	**.906**	**.916**	**.909**
	HPDI	**.869**	**.896**	**.893**	**.889**	**.913**	**.923**	**.913**	**.919**	**.909**	.927	**.915**
0.50	FISHER	**.822**	**.813**	**.809**	**.826**	**.784**	**.786**	**.791**	**.787**	**.780**	**.783**	**.787**
	RIN	.952	.953	.968	.964	.958	.953	.963	.948	.962	.945	.959
	SPEAR	.948	.945	.956	.956	.950	.941	.951	.945	.951	.945	.954
	PB	**.918**	**.903**	**.921**	.929	.928	.932	.935	.936	.929	.925	.935
	BCA	**.882**	**.879**	**.899**	**.894**	**.903**	**.912**	**.903**	**.917**	**.911**	**.905**	**.920**
	HPDI	**.885**	**.872**	**.891**	**.904**	**.912**	**.909**	**.921**	.925	**.920**	**.918**	**.915**

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 15
95% Coverage Probabilities for paired Pareto Distributions
(Skewness = 2.811, Kurtosis = 14.828)

ρ	n	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.946	.950	.955	.951	.961	.951	.957	.950	.953	.953	.956
	RIN	.950	.947	.950	.949	.949	.948	.946	.945	.948	.947	.953
	SPEAR	.951	.956	.954	.960	.952	.952	.956	.958	.953	.969	.959
	PB	**.907**	**.921**	**.918**	**.903**	**.921**	.932	.926	.933	.933	.926	.927
	BCA	.943	.932	.931	**.910**	.928	.929	.932	.936	.936	.928	.926
	HPDI	**.885**	**.898**	**.901**	**.888**	**.905**	**.923**	**.921**	.925	.929	**.922**	**.923**
0.10	FISHER	**.923**	**.906**	**.919**	.938	**.914**	**.921**	**.906**	.927	**.900**	**.921**	**.910**
	RIN	.946	.938	.950	.949	.936	.942	.925	.956	.945	.954	.951
	SPEAR	.949	.943	.949	.957	.942	.944	.936	.957	.951	.946	.957
	PB	**.904**	**.906**	**.906**	.927	**.914**	**.919**	**.913**	.932	**.920**	.942	.936
	BCA	**.922**	**.910**	**.908**	.930	**.906**	**.913**	**.903**	.928	**.919**	.937	.934
	HPDI	**.873**	**.879**	**.881**	**.913**	**.895**	**.910**	**.904**	**.920**	**.912**	.935	.933
0.20	FISHER	**.913**	.926	**.913**	**.899**	**.903**	**.882**	**.908**	**.891**	**.872**	**.902**	**.879**
	RIN	.939	.952	.964	.951	.938	.944	.957	.960	.935	.954	.946
	SPEAR	.951	.954	.965	.949	.957	.960	.957	.957	.945	.960	.942
	PB	**.913**	.925	.940	**.911**	.935	**.924**	.941	.931	.930	.941	.935
	BCA	**.910**	**.921**	.927	**.902**	**.907**	**.916**	**.923**	.925	**.918**	.933	.925
	HPDI	**.881**	**.898**	**.921**	**.892**	**.917**	**.911**	.929	**.922**	**.919**	.938	.936
0.30	FISHER	**.897**	**.886**	**.883**	**.872**	**.868**	**.861**	**.884**	**.850**	**.879**	**.847**	**.858**
	RIN	.948	.948	.959	.953	.948	.950	.939	.944	.960	.956	.946
	SPEAR	.949	.945	.955	.952	.953	.954	.952	.954	.962	.959	.945
	PB	.925	**.911**	**.918**	**.920**	.927	**.921**	.927	.929	.942	.934	.941
	BCA	**.918**	**.895**	**.900**	**.917**	**.906**	**.901**	**.921**	**.914**	.937	**.916**	.925
	HPDI	**.887**	**.886**	**.896**	**.903**	**.908**	**.907**	**.920**	**.916**	.940	**.921**	.934
0.40	FISHER	**.865**	**.856**	**.879**	**.889**	**.843**	**.829**	**.839**	**.824**	**.827**	**.824**	.931
	RIN	.953	.952	.960	.966	.953	.953	.957	.952	.950	.943	.962
	SPEAR	.955	.952	.954	.959	.949	.953	.953	.947	.948	.950	.963
	PB	**.910**	**.917**	.928	.943	.931	**.916**	.938	.928	.931	.945	.934
	BCA	**.901**	**.904**	**.914**	**.914**	**.902**	**.902**	**.922**	**.907**	**.912**	.931	**.916**
	HPDI	**.873**	**.880**	**.905**	**.917**	**.908**	**.908**	**.920**	**.920**	**.917**	.938	.926
0.50	FISHER	**.868**	**.843**	**.853**	**.820**	**.826**	**.819**	**.816**	**.815**	**.807**	**.813**	**.798**
	RIN	.952	.958	.956	.941	.956	.946	.956	.960	.952	.955	.954
	SPEAR	.950	.957	.936	.939	.958	.946	.941	.946	.942	.953	.954
	PB	.931	**.916**	**.922**	**.920**	.939	.932	.940	.942	.940	.932	.925
	BCA	**.914**	**.890**	**.903**	**.886**	**.909**	**.905**	**.914**	**.920**	**.907**	**.914**	**.905**
	HPDI	**.904**	**.887**	**.895**	**.895**	**.920**	**.921**	.926	.935	.932	**.924**	**.917**

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 16
95% Coverage Probabilities for Normal Distribution Paired with Triangular Distribution (Skewness = 0, Kurtosis = -.06)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.958	.951	.959	.947	.938	.950	.951	.954	.952	.952	.949
	RIN	.960	.951	.966	.946	.937	.947	.956	.952	.953	.952	.953
	SPEAR	.965	.955	.964	.952	.951	.955	.961	.958	.950	.959	.961
	PB	.935	.922	.946	.931	.933	.947	.943	.951	.947	.954	.953
	BCA	.954	.937	.947	.939	.933	.947	.945	.950	.947	.951	.952
	HPDI	**.903**	**.909**	.933	**.918**	.928	.944	.934	.945	.936	.946	.950
0.10	FISHER	.943	.956	.952	.949	.947	.959	.945	.955	.938	.951	.939
	RIN	.942	.954	.955	.945	.945	.957	.948	.953	.940	.955	.941
	SPEAR	.948	.962	.953	.945	.953	.964	.961	.960	.952	.952	.943
	PB	**.914**	.941	.942	.935	.937	.949	.940	.956	.940	.947	.935
	BCA	.932	.943	.950	.945	.943	.955	.943	.957	.939	.946	.939
	HPDI	**.892**	**.922**	.932	**.924**	.932	.950	.934	.948	.937	.939	.931
0.20	FISHER	.957	.956	.943	.939	.955	.956	.955	.950	.959	.949	.948
	RIN	.947	.961	.945	.948	.955	.961	.951	.945	.956	.953	.951
	SPEAR	.947	.955	.952	.940	.962	.971	.958	.950	.955	.958	.956
	PB	**.921**	.938	.937	.918	.947	.951	.954	.946	.952	.952	.948
	BCA	.938	.945	.942	.926	.950	.952	.954	.946	.951	.952	.947
	HPDI	**.893**	**.918**	**.918**	**.902**	.940	.943	.952	.943	.944	.948	.945
0.30	FISHER	.952	.956	.940	.952	.962	.946	.955	.945	.939	.949	.951
	RIN	.945	.946	.941	.948	.956	.946	.956	.947	.937	.953	.946
	SPEAR	.952	.948	.951	.958	.963	.948	.957	.947	.947	.954	.956
	PB	.928	.930	.928	.941	.951	.948	.954	.941	.936	.942	.947
	BCA	.944	.939	.937	.951	.957	.953	.956	.938	.941	.943	.947
	HPDI	**.905**	**.913**	**.914**	.931	.948	.944	.945	.937	.932	.944	.942
0.40	FISHER	.948	.958	.954	.967	.936	.951	.962	.957	.953	.960	.947
	RIN	.950	.955	.955	.964	.942	.948	.959	.951	.946	.963	.952
	SPEAR	.950	.958	.954	.956	.951	.946	.965	.960	.949	.957	.952
	PB	.909	.938	.935	.949	.934	.949	.963	.943	.945	.954	.946
	BCA	.931	.948	.936	.956	.937	.949	.963	.949	.949	.959	.948
	HPDI	**.892**	**.918**	**.918**	.933	.928	.938	.957	.941	.937	.947	.941
0.50	FISHER	.945	.955	.944	.947	.953	.952	.949	.941	.959	.948	.963
	RIN	.949	.954	.943	.946	.945	.939	.947	.944	.957	.947	.958
	SPEAR	.951	.953	.943	.954	.953	.952	.948	.945	.958	.947	.951
	PB	**.916**	.940	.936	.939	.944	.944	.943	.937	.958	.944	.958
	BCA	.929	.942	.945	.942	.946	.942	.947	.939	.958	.943	.960
	HPDI	**.903**	**.923**	**.917**	.930	.944	.936	.943	.930	.950	.938	.950

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 17
95% Coverage Probabilities for Normal Distribution Paired with Uniform Distribution
(Skewness = 0, Kurtosis = -1.2)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.947	.945	.966	.949	.950	.945	.954	.941	.947	.942	.956
	RIN	.946	.945	.960	.953	.947	.952	.953	.936	.952	.936	.954
	SPEAR	.944	.947	.966	.963	.952	.957	.958	.951	.950	.947	.959
	PB	**.907**	.931	.958	.939	.938	.945	.947	.937	.948	.938	.953
	BCA	.929	.949	.962	.952	.940	.949	.954	.942	.951	.937	.956
	HPDI	**.886**	**.921**	.944	.928	.932	.940	.944	**.922**	.942	.934	.948
0.10	FISHER	.954	.940	.964	.948	.945	.953	.958	.959	.938	.957	.962
	RIN	.953	.938	.956	.945	.943	.944	.963	.958	.938	.952	.954
	SPEAR	.952	.945	.965	.955	.956	.953	.963	.963	.948	.960	.963
	PB	.937	**.915**	.945	.937	.939	.947	.958	.957	.935	.953	.959
	BCA	.951	.933	.950	.944	.943	.952	.963	.958	.939	.958	.956
	HPDI	**.918**	**.898**	.932	**.926**	.934	.944	.948	.952	.929	.947	.950
0.20	FISHER	.958	.947	.944	.945	.949	.944	.942	.947	.947	.965	.946
	RIN	.961	.950	.945	.954	.948	.940	.947	.949	.953	.961	.948
	SPEAR	.959	.962	.951	.953	.950	.947	.950	.955	.951	.958	.947
	PB	**.920**	.927	**.924**	.933	.942	.939	.940	.945	.945	.957	.946
	BCA	.937	.941	.938	.946	.943	.939	.943	.947	.950	.956	.947
	HPDI	**.900**	**.913**	**.912**	.926	.937	.931	.938	.939	.943	.957	.940
0.30	FISHER	.944	.957	.956	.959	.956	.960	.956	.945	.964	.945	.944
	RIN	.959	.954	.957	.957	.954	.955	.955	.953	.955	.948	.946
	SPEAR	.960	.957	.959	.961	.958	.955	.961	.942	.961	.948	.945
	PB	.926	.939	.938	.946	.953	.949	.950	.941	.958	.943	.938
	BCA	.943	.949	.949	.951	.953	.957	.958	.943	.954	.949	.936
	HPDI	**.897**	.926	**.920**	.934	.945	.942	.951	.937	.955	.943	.936
0.40	FISHER	.957	.958	.942	.961	.948	.951	.971	.945	.953	.946	.961
	RIN	.960	.960	.946	.959	.952	.948	.971	.943	.954	.943	.960
	SPEAR	.956	.959	.950	.957	.955	.950	.965	.944	.956	.949	.948
	PB	**.920**	.926	.926	.951	.932	.944	.962	.942	.949	.940	.952
	BCA	.942	.944	.941	.960	.944	.947	.963	.944	.951	.946	.951
	HPDI	**.890**	**.918**	**.920**	.932	.931	.932	.952	.938	.946	.936	.947
0.50	FISHER	.953	.959	.954	.961	.951	.958	.963	.947	.960	.951	.952
	RIN	.957	.958	.955	.960	.955	.957	.959	.955	.961	.956	.951
	SPEAR	.951	.964	.946	.959	.944	.947	.952	.950	.952	.945	.946
	PB	**.906**	.941	.935	.948	.941	.948	.957	.939	.951	.945	.938
	BCA	.936	.954	.943	.951	.951	.952	.960	.941	.951	.948	.942
	HPDI	**.880**	.925	**.922**	.937	.932	.943	.955	.938	.942	.942	.933

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 18
95% Coverage Probabilities for Normal Distribution Paired with Laplace Distribution
(Skewness = 0, Kurtosis = 3)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.948	.944	.943	.939	.949	.959	.950	.956	.946	.945	.958
	RIN	.944	.952	.948	.946	.944	.958	.947	.952	.950	.940	.950
	SPEAR	.954	.955	.951	.951	.946	.962	.958	.965	.952	.946	.951
	PB	.926	.925	**.920**	.929	.933	.955	.952	.945	.943	.943	.946
	BCA	.937	.930	.925	.933	.933	.958	.949	.944	.940	.940	.947
	HPDI	**.893**	**.910**	**.904**	**.918**	.925	.948	.946	.939	.939	.938	.942
0.10	FISHER	.942	.946	.950	.954	.950	.955	.952	.940	.937	.951	.954
	RIN	.949	.949	.946	.954	.955	.947	.946	.940	.941	.949	.947
	SPEAR	.957	.953	.953	.952	.967	.957	.951	.950	.950	.959	.954
	PB	**.914**	**.920**	.934	.936	.945	.939	.943	.933	.934	.945	.948
	BCA	**.924**	**.921**	.933	.938	.941	.938	.938	.928	.934	.944	.947
	HPDI	**.894**	**.907**	**.921**	.927	.940	.931	.940	.926	.929	.942	.944
0.20	FISHER	.934	.952	.944	.938	.952	.947	.945	.944	.947	.953	.948
	RIN	.938	.946	.947	.942	.949	.946	.946	.946	.941	.951	.952
	SPEAR	.955	.949	.959	.944	.957	.954	.952	.952	.949	.962	.969
	PB	**.902**	.932	**.922**	.925	.935	.943	.949	.941	.948	.951	.949
	BCA	**.915**	.931	**.920**	**.919**	.937	.944	.943	.940	.944	.945	.947
	HPDI	**.888**	**.912**	**.911**	**.915**	.928	.940	.945	.937	.941	.945	.946
0.30	FISHER	.951	.963	.949	.954	.946	.954	.952	.943	.953	.941	.964
	RIN	.961	.963	.951	.953	.945	.961	.944	.936	.954	.947	.957
	SPEAR	.964	.969	.947	.958	.952	.962	.952	.946	.963	.956	.961
	PB	.935	.943	.928	.938	.936	.939	.938	.941	.951	.935	.955
	BCA	.937	.950	.927	.933	.931	.942	.937	.937	.953	.935	.952
	HPDI	**.908**	.928	**.917**	**.924**	.929	.935	.935	.927	.944	.934	.953
0.40	FISHER	.945	.953	.960	.945	.960	.950	.951	.945	.954	.958	.954
	RIN	.951	.951	.958	.949	.954	.946	.951	.944	.953	.958	.953
	SPEAR	.955	.951	.965	.959	.957	.952	.961	.942	.954	.955	.952
	PB	.925	**.922**	.936	.935	.946	.949	.950	.943	.951	.954	.946
	BCA	.934	.928	.942	.928	.941	.939	.947	.935	.949	.952	.947
	HPDI	**.900**	**.917**	**.923**	**.913**	.935	.940	.944	.937	.944	.949	.943
0.50	FISHER	.944	.959	.943	.955	.956	.968	.955	.959	.957	.954	.954
	RIN	.946	.950	.945	.951	.948	.971	.955	.950	.959	.953	.956
	SPEAR	.942	.953	.943	.953	.948	.959	.943	.951	.959	.951	.942
	PB	**.918**	.929	.925	.938	.939	.960	.941	.948	.955	.947	.950
	BCA	.936	.931	**.912**	.933	.944	.960	.939	.947	.948	.951	.951
	HPDI	**.900**	**.914**	**.904**	.926	.926	.956	.935	.940	.948	.946	.950

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 19

95% Coverage Probabilities for Normal Distribution Paired with Beta Distribution

(a = 4, b = 1.25, Skewness = -.848, Kurtosis = .221)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.960	.942	.942	.953	.951	.949	.952	.956	.952	.946	.940
	RIN	.961	.948	.944	.953	.955	.958	.945	.964	.945	.946	.940
	SPEAR	.963	.944	.957	.952	.957	.962	.962	.958	.955	.953	.948
	PB	.941	**.918**	.928	.939	.944	.949	.950	.952	.946	.940	.936
	BCA	.945	.925	.932	.944	.944	.949	.954	.952	.948	.942	.935
	HPDI	**.909**	**.895**	**.914**	.926	.940	.938	.946	.945	.949	.938	.931
0.10	FISHER	.940	.950	.948	.941	.952	.937	.955	.950	.949	.957	.954
	RIN	.943	.954	.941	.943	.945	.943	.951	.948	.946	.956	.948
	SPEAR	.942	.957	.958	.949	.959	.956	.960	.954	.959	.957	.961
	PB	**.908**	.928	.934	.930	.935	.931	.950	.946	.946	.958	.948
	BCA	.928	.939	.940	.931	.941	.931	.951	.946	.944	.957	.948
	HPDI	**.892**	**.914**	.926	**.919**	.925	**.924**	.943	.941	.939	.955	.945
0.20	FISHER	.956	.953	.951	.948	.941	.948	.953	.951	.950	.958	.952
	RIN	.947	.955	.949	.944	.952	.944	.948	.948	.946	.956	.946
	SPEAR	.947	.965	.956	.950	.949	.959	.955	.957	.960	.960	.948
	PB	**.916**	.938	.936	.937	.932	.942	.949	.944	.940	.957	.947
	BCA	.929	.942	.937	.939	.934	.944	.948	.948	.942	.954	.948
	HPDI	**.899**	**.917**	**.923**	.930	.928	.940	.944	.939	.938	.951	.941
0.30	FISHER	.952	.956	.956	.942	.958	.934	.954	.937	.950	.950	.948
	RIN	.952	.951	.948	.951	.961	.934	.954	.931	.948	.952	.951
	SPEAR	.956	.955	.946	.943	.963	.948	.957	.943	.957	.946	.959
	PB	**.922**	.937	.937	.931	.955	.928	.950	.937	.949	.948	.951
	BCA	.937	.941	.944	.930	.956	.928	.952	.936	.950	.943	.951
	HPDI	**.899**	**.917**	.929	**.917**	.949	**.924**	.947	.928	.946	.943	.944
0.40	FISHER	.955	.950	.959	.958	.956	.953	.951	.946	.954	.951	.954
	RIN	.960	.947	.963	.950	.945	.947	.950	.947	.954	.950	.944
	SPEAR	.963	.955	.963	.950	.947	.954	.943	.951	.957	.961	.951
	PB	.927	.930	.939	.942	.946	.940	.942	.942	.952	.947	.950
	BCA	.942	.940	.940	.944	.947	.941	.946	.941	.951	.948	.948
	HPDI	**.909**	**.912**	**.921**	.930	.936	.932	.942	.941	.950	.947	.945
0.50	FISHER	.947	.952	.944	.963	.946	.958	.954	.948	.952	.956	.957
	RIN	.947	.950	.955	.948	.942	.950	.947	.948	.940	.962	.949
	SPEAR	.941	.951	.951	.943	.933	.948	.945	.947	.948	.944	.945
	PB	**.919**	.928	.939	.945	.938	.953	.945	.935	.939	.947	.952
	BCA	.929	.936	.937	.948	.941	.952	.945	.936	.943	.944	.953
	HPDI	**.894**	**.908**	.928	.943	.930	.940	.935	.931	.934	.941	.946

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 20
95% Coverage Probabilities for Normal Distribution Paired with Beta Distribution
(a = 4, b = 1.5, Skewness = -.694, Kurtosis -.069)

ρ	n	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.944	.941	.948	.951	.958	.951	.941	.950	.952	.939	.952
	RIN	.947	.942	.960	.945	.955	.950	.946	.958	.943	.933	.951
	SPEAR	.952	.946	.956	.949	.966	.958	.946	.961	.959	.947	.960
	PB	**.923**	**.914**	**.923**	.930	.951	.950	.939	.948	.951	.933	.949
	BCA	.929	.928	.940	.931	.955	.949	.938	.951	.951	.931	.950
	HPDI	**.905**	**.904**	**.908**	**.918**	.943	.944	.933	.946	.946	.932	.944
0.10	FISHER	.937	.961	.953	.950	.966	.941	.947	.953	.946	.957	.950
	RIN	.931	.949	.952	.947	.967	.943	.945	.951	.951	.957	.946
	SPEAR	.949	.954	.956	.952	.970	.945	.955	.950	.943	.967	.955
	PB	**.910**	**.920**	.939	.933	.955	.943	.936	.953	.944	.952	.951
	BCA	**.920**	.934	.944	.937	.957	.941	.935	.948	.945	.952	.953
	HPDI	**.878**	**.899**	.926	.925	.949	.938	.931	.943	.945	.945	.944
0.20	FISHER	.952	.947	.949	.953	.947	.941	.956	.940	.942	.940	.958
	RIN	.955	.956	.938	.950	.940	.941	.954	.938	.936	.941	.951
	SPEAR	.960	.958	.949	.963	.944	.945	.962	.948	.946	.946	.958
	PB	**.923**	.927	.927	.937	.939	.936	.945	.943	.945	.934	.955
	BCA	.941	.936	.930	.946	.941	.936	.949	.949	.949	.934	.953
	HPDI	**.897**	**.901**	**.914**	.926	.929	.935	.949	.938	.939	.933	.953
0.30	FISHER	.959	.944	.937	.953	.957	.950	.962	.955	.947	.959	.938
	RIN	.962	.947	.929	.948	.954	.952	.959	.952	.948	.958	.942
	SPEAR	.961	.949	.929	.965	.955	.949	.957	.949	.947	.962	.944
	PB	.927	**.924**	**.912**	.939	.950	.947	.952	.952	.948	.956	.934
	BCA	.938	**.922**	**.919**	.940	.954	.944	.956	.947	.952	.956	.935
	HPDI	**.911**	**.909**	**.896**	.929	.939	.935	.950	.944	.937	.954	.928
0.40	FISHER	.953	.943	.951	.959	.943	.951	.958	.948	.956	.940	.958
	RIN	.955	.942	.955	.957	.943	.951	.958	.951	.954	.950	.946
	SPEAR	.942	.943	.944	.949	.947	.946	.955	.959	.956	.953	.954
	PB	**.922**	**.916**	.932	.947	.941	.948	.951	.945	.954	.935	.956
	BCA	.943	.928	.938	.949	.933	.948	.952	.949	.953	.941	.951
	HPDI	**.896**	**.896**	.925	.929	.934	.942	.940	.938	.949	.930	.941
0.50	FISHER	.956	.953	.947	.953	.960	.956	.947	.950	.936	.949	.956
	RIN	.949	.949	.948	.948	.955	.952	.948	.940	.941	.950	.953
	SPEAR	.955	.950	.950	.948	.953	.944	.949	.940	.931	.946	.960
	PB	**.921**	.937	.929	.936	.947	.949	.940	.939	.934	.941	.950
	BCA	.941	.937	.941	.946	.946	.949	.946	.940	.935	.940	.953
	HPDI	**.898**	**.916**	**.922**	.928	.938	.943	.932	.934	.929	.937	.946

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 21
95% Coverage Probabilities for Normal Distribution Paired with Chi-Square Distribution (df = 16, Skewness = .71, Kurtosis = .75)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.937	.945	.959	.944	.947	.949	.951	.953	.953	.950	.963
	RIN	.945	.943	.969	.956	.951	.946	.956	.956	.945	.946	.958
	SPEAR	.950	.952	.969	.959	.961	.951	.954	.953	.959	.964	.964
	PB	**.919**	.931	.955	.927	.936	.948	.941	.952	.951	.953	.961
	BCA	.935	.938	.951	.927	.937	.948	.939	.950	.951	.953	.960
	HPDI	**.886**	**.919**	.943	**.920**	.933	.943	.931	.944	.947	.943	.958
0.10	FISHER	.937	.945	.953	.937	.948	.946	.957	.959	.947	.953	.961
	RIN	.940	.945	.947	.940	.952	.947	.952	.957	.944	.950	.962
	SPEAR	.945	.940	.952	.949	.967	.951	.962	.962	.953	.956	**.976**
	PB	**.906**	.927	.927	.926	.947	.938	.953	.956	.946	.949	.962
	BCA	**.923**	.931	.928	.928	.946	.944	.955	.952	.949	.946	.960
	HPDI	**.876**	**.904**	**.912**	**.916**	.941	.937	.951	.953	.947	.947	.953
0.20	FISHER	.951	.957	.951	.942	.949	.958	.954	.948	.965	.951	.941
	RIN	.956	.948	.948	.943	.950	.954	.953	.955	.963	.950	.945
	SPEAR	.952	.948	.952	.943	.957	.955	.956	.955	.965	.956	.944
	PB	**.924**	.935	.926	.933	.943	.949	.952	.948	.963	.950	.940
	BCA	.933	.946	.930	.936	.941	.949	.953	.944	.958	.952	.937
	HPDI	**.896**	**.921**	**.918**	**.920**	.938	.942	.946	.933	.955	.943	.934
0.30	FISHER	.950	.956	.948	.946	.952	.943	.949	.944	.961	.945	.952
	RIN	.946	.959	.943	.946	.958	.954	.943	.937	.957	.944	.949
	SPEAR	.955	.961	.949	.945	.957	.959	.953	.937	.969	.953	.950
	PB	**.920**	.930	.933	**.923**	.946	.944	.941	.939	.958	.937	.944
	BCA	.932	.942	.936	.929	.947	.943	.941	.938	.956	.941	.945
	HPDI	**.899**	**.919**	**.910**	**.911**	.940	.935	.936	.934	.950	.933	.941
0.40	FISHER	.948	.950	.944	.941	.954	.940	.957	.957	.952	.951	.948
	RIN	.950	.964	.949	.940	.952	.948	.954	.947	.951	.953	.947
	SPEAR	.951	.958	.954	.942	.953	.951	.948	.955	.957	.958	.950
	PB	**.907**	.930	.932	.933	.942	.943	.944	.947	.946	.951	.939
	BCA	**.923**	.936	.936	.940	.944	.945	.948	.946	.947	.950	.939
	HPDI	**.889**	**.911**	.925	**.923**	.930	.932	.945	.948	.943	.946	.941
0.50	FISHER	.957	.936	.948	.954	.948	.947	.946	.960	.946	.956	.959
	RIN	.957	.936	.943	.959	.949	.944	.948	.959	.944	.953	.955
	SPEAR	.951	.934	.951	.959	.947	.951	.955	.954	.938	.953	.948
	PB	**.923**	.927	.929	.939	.939	.941	.942	.951	.943	.955	.954
	BCA	.934	**.921**	.940	.942	.942	.941	.941	.950	.942	.950	.951
	HPDI	**.902**	**.906**	**.920**	.926	.932	.932	.938	.947	.940	.950	.946

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 22
95% Coverage Probabilities for Normal Distribution Paired with Chi-Square Distribution (df = 4, Skewness = 1.41, Kurtosis = 3)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.943	.946	.947	.939	.950	.943	.953	.956	.942	.948	.946
	RIN	.943	.958	.951	.942	.945	.956	.954	.951	.949	.950	.948
	SPEAR	.958	.957	.956	.952	.950	.963	.962	.958	.949	.963	.950
	PB	**.911**	.928	.938	.931	.933	.943	.942	.950	.944	.940	.941
	BCA	.929	.932	.938	.934	**.924**	.942	.938	.951	.948	.942	.936
	HPDI	**.889**	**.909**	**.924**	**.922**	**.924**	.940	.938	.941	.939	.936	.934
0.10	FISHER	.943	.954	.945	.937	.947	.955	.950	.946	.953	.949	.949
	RIN	.945	.954	.948	.938	.941	.951	.943	.943	.950	.938	.948
	SPEAR	.952	.959	.958	.948	.951	.963	.951	.950	.954	.953	.954
	PB	**.917**	.942	.937	**.924**	.929	.950	.941	.938	.943	.941	.948
	BCA	.929	.950	.939	.926	.926	.946	.939	.940	.936	.940	.944
	HPDI	**.889**	**.923**	**.924**	**.908**	**.918**	.940	.938	.936	.943	.939	.941
0.20	FISHER	.954	.964	.955	.947	.944	.954	.963	.947	.956	.947	.950
	RIN	.967	.963	.950	.943	.949	.957	.965	.952	.960	.943	.948
	SPEAR	.964	.968	.952	.952	.948	.960	.964	.952	.957	.948	.957
	PB	.933	.935	.938	.933	.925	.945	.961	.946	.954	.947	.945
	BCA	.945	.947	.938	.937	.927	.935	.961	.939	.949	.948	.940
	HPDI	**.910**	**.910**	.929	**.918**	**.920**	.930	.954	.937	.947	.945	.943
0.30	FISHER	.937	.959	.943	.936	.954	.952	.957	.959	.950	.959	.950
	RIN	.948	.952	.944	.940	.956	.953	.959	.956	.943	.957	.943
	SPEAR	.950	.949	.947	.947	.959	.950	.958	.958	.946	.962	.948
	PB	**.908**	**.922**	**.923**	**.921**	.942	.943	.955	.953	.944	.953	.951
	BCA	**.919**	.936	**.924**	.926	.943	.939	.953	.953	.940	.956	.950
	HPDI	**.877**	**.910**	**.905**	**.912**	.937	.934	.949	.944	.938	.951	.948
0.40	FISHER	.962	.954	.955	.954	.946	.963	.955	.957	.957	.948	.954
	RIN	.961	.949	.950	.952	.948	.959	.950	.943	.949	.950	.949
	SPEAR	.959	.956	.952	.963	.954	.961	.957	.950	.943	.948	.955
	PB	**.914**	.929	.931	.933	.934	.956	.945	.945	.950	.937	.950
	BCA	.930	.936	.934	.938	.936	.952	.941	.943	.951	.938	.950
	HPDI	**.891**	**.916**	.925	.926	.928	.949	.946	.935	.943	.934	.949
0.50	FISHER	.960	.962	.949	.952	.958	.952	.958	.968	.954	.945	.961
	RIN	.957	.953	.944	.940	.946	.948	.942	.953	.944	.946	.958
	SPEAR	.956	.950	.952	.935	.948	.934	.947	.952	.942	.938	.947
	PB	.926	.931	.929	.925	.940	.940	.944	.960	.943	.935	.949
	BCA	.937	.942	.936	.927	.937	.941	.936	.959	.947	.932	.946
	HPDI	**.905**	**.918**	**.920**	**.905**	.931	.935	.939	.956	.940	.934	.947

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 23
95% Coverage Probabilities for Normal Distribution Paired with Chi-Square Distribution
(df = 3, Skewness = 1.63, Kurtosis = 4)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.932	.950	.946	.955	.956	.942	.940	.942	.945	.964	.938
	RIN	.946	.949	.941	.947	.952	.951	.951	.940	.946	.970	.951
	SPEAR	.947	.963	.950	.958	.951	.951	.956	.944	.955	.971	.957
	PB	**.907**	.933	**.919**	.946	.943	.937	.930	.937	.942	.961	.938
	BCA	.930	.942	**.924**	.950	.945	.931	.927	.933	.942	.959	.937
	HPDI	**.887**	**.917**	**.904**	.934	.937	.933	**.921**	.928	.938	.958	.934
0.10	FISHER	.949	.949	.959	.962	.949	.943	.938	.960	.956	.948	.954
	RIN	.945	.944	.953	.954	.945	.944	.948	.967	.954	.943	.953
	SPEAR	.950	.948	.965	.963	.953	.958	.947	.969	.948	.950	.952
	PB	**.921**	.928	.947	.946	.937	.945	.935	.953	.946	.941	.955
	BCA	.944	.932	.944	.947	.937	.936	.933	.952	.944	.941	.953
	HPDI	**.901**	**.902**	.935	.932	**.923**	.935	.927	.942	.944	.937	.952
0.20	FISHER	.945	.958	.955	.956	.950	.953	.955	.955	.958	.956	.948
	RIN	.958	.951	.940	.946	.949	.958	.954	.950	.953	.957	.939
	SPEAR	.961	.957	.949	.955	.959	.955	.956	.953	.960	.956	.954
	PB	**.919**	.941	927	.933	.945	.950	.945	.948	.951	.950	.948
	BCA	.932	.945	.935	.936	.938	.949	.941	.949	.951	.950	.945
	HPDI	**.905**	.930	**.915**	**.918**	.940	.938	.938	.947	.948	.946	.942
0.30	FISHER	.949	.956	.950	.952	.947	.949	.945	.958	.956	.944	.952
	RIN	.947	.951	.946	.947	.926	.950	.944	.950	.954	.949	.945
	SPEAR	.957	.954	.954	.963	.936	.953	.945	.958	.955	.952	.954
	PB	**.916**	**.917**	.932	.933	**.911**	.935	.939	.948	.948	.940	.943
	BCA	.935	.925	.935	.926	**.913**	.936	.936	.948	.943	.935	.940
	HPDI	**.899**	**.912**	**.918**	.928	**.901**	.933	.932	.943	.943	.933	.942
0.40	FISHER	.952	.951	.942	.951	.952	.948	.968	.954	.959	.949	.947
	RIN	.950	.954	.945	.945	.937	.941	.953	.952	.961	.947	.943
	SPEAR	.948	.969	.948	.943	.948	.937	.959	.958	.957	.959	.954
	PB	**.914**	.926	.928	**.920**	.930	.942	.960	.953	.948	.933	.942
	BCA	.938	.934	.930	.929	.928	.936	.953	.948	.943	.928	.937
	HPDI	**.893**	**.906**	**.913**	**.913**	**.923**	.935	.954	.946	.941	.930	.931
0.50	FISHER	.947	.955	.946	.954	.956	.961	.967	.947	.957	.958	.955
	RIN	.939	.943	.936	.959	.944	.953	.958	.945	.950	.954	.950
	SPEAR	.936	.944	.926	.943	.941	.947	.951	.942	.950	.946	.944
	PB	**.893**	**.923**	**.922**	.929	.932	.936	.950	.935	.949	.950	.942
	BCA	**.924**	**.922**	**.921**	.934	.934	.940	.949	.928	.945	.953	.942
	HPDI	**.869**	**.908**	**.900**	**.920**	.932	.930	.941	.928	.940	.948	.938

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 24
95% Coverage Probabilities for Normal Distribution Paired with Chi-Square Distribution
(df = 2, Skewness = 2, Kurtosis = 6)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.949	.959	.945	.937	.954	.956	.958	.940	.946	.954	.947
	RIN	.952	.951	.943	.940	.955	.961	.958	.951	.935	.962	.950
	SPEAR	.949	.957	.948	.943	.962	.964	.959	.955	.950	.959	.959
	PB	.929	**.920**	.928	**.914**	.932	.947	.951	.941	.937	.943	.949
	BCA	.944	.925	.932	**.915**	.937	.942	.947	.941	.935	.941	.943
	HPDI	**.907**	**.904**	**.914**	**.897**	.930	.942	.946	.933	.934	.939	.940
0.10	FISHER	.931	.943	.934	.943	.958	.952	.952	.951	.950	.948	.951
	RIN	.940	.947	.931	.948	.948	.951	.950	.954	.947	.947	.950
	SPEAR	.942	.954	.949	.957	.960	.960	.948	.960	.963	.954	.958
	PB	**.910**	**.921**	**.920**	.941	.937	.941	.940	.945	.944	.941	.948
	BCA	**.923**	.928	**.920**	.939	.932	.936	.937	.942	.942	.936	.947
	HPDI	**.884**	**.897**	**.895**	.931	**.924**	.933	.936	.940	.939	.937	.942
0.20	FISHER	.935	.949	.956	.947	.943	.952	.957	.936	.946	.956	.943
	RIN	.934	.944	.951	.954	.941	.950	.960	9.45	.951	.952	.953
	SPEAR	.933	.950	.956	.958	.944	.952	.964	.953	.957	.955	.962
	PB	**.903**	.931	.934	.931	**.924**	.942	.946	.933	.939	.946	.951
	BCA	.927	.930	.934	.931	.927	.937	.943	.925	.933	.944	.946
	HPDI	**.881**	**.915**	**.924**	**.922**	**.917**	.938	.942	**.921**	.937	.941	.949
0.30	FISHER	.953	.958	.959	.954	.950	.938	.941	.960	.940	.972	.962
	RIN	.948	.950	.947	.955	.946	.925	.955	.951	.941	.964	.945
	SPEAR	.953	.946	.952	.960	.952	.936	.963	.957	.945	.967	.952
	PB	.929	.933	.936	.932	.946	**.920**	.940	.946	.937	.967	.947
	BCA	.943	.939	.943	.933	.939	**.919**	.930	.951	.940	.964	.954
	HPDI	**.907**	**.914**	**.922**	**.920**	.939	**.916**	.932	.946	.936	.961	.941
0.40	FISHER	.948	.961	.959	.952	.959	.949	.949	.957	.964	.960	.956
	RIN	.953	.948	.959	.947	.942	.946	.935	.952	.957	.957	.956
	SPEAR	.956	.955	.957	.939	.943	.947	.941	.948	.947	.946	.951
	PB	**.921**	.933	.947	**.923**	.941	.937	.932	.946	.950	.956	.940
	BCA	.936	.937	.937	**.919**	.935	.934	.925	.943	.949	.953	.937
	HPDI	**.896**	**.912**	.931	**.907**	.929	.925	.925	.941	.947	.952	.940
0.50	FISHER	.953	.955	.960	.959	.960	.957	.953	.968	.959	.960	.944
	RIN	.950	.955.	.943	.954	.961	.956	.943	.951	.949	.943	.933
	SPEAR	.958	959	.941	.956	.961	.950	.938	.942	.950	.935	.938
	PB	**.916**	.928	.926	.930	.933	.938	.939	.951	.948	.937	.928
	BCA	.925	.931	.927	.925	.936	.939	.934	.951	.948	.931	.927
	HPDI	**.900**	**.906**	**.916**	**.913**	.928	.933	.931	.949	.941	.932	**.922**

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 25
95% Coverage Probabilities for Normal Distribution Paired with Chi-Square Distribution (df = 1, Skewness = 2.83, Kurtosis = 12)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.961	.957	.955	.945	.957	.949	.961	.957	.947	.951	.955
	RIN	.947	.961	.946	.952	.947	.950	.955	.957	.949	.952	.964
	SPEAR	.955	.967	.944	.953	.956	.956	.969	.961	.957	.963	.968
	PB	.934	.941	**.921**	.925	.940	.935	.944	.950	.948	.945	.950
	BCA	.957	.936	**.919**	**.919**	.931	.930	.935	.948	.940	.939	.942
	HPDI	**.913**	**.921**	**.911**	**.911**	.930	.929	.938	.945	.937	.941	.946
0.10	FISHER	.957	.941	.942	.956	.960	.958	.955	.941	.948	.949	.963
	RIN	.952	.931	.941	.942	.966	.957	.947	.949	.946	.958	.945
	SPEAR	.953	.945	.942	.958	.972	.966.	.949	.957	.959	.959	.959
	PB	**.919**	**.917**	**.919**	**.917**	.943	944	.943	.926	.940	.946	.947
	BCA	.935	**.919**	**.923**	**.918**	.938	.932	.932	**.919**	.942	.947	.944
	HPDI	**.892**	**.905**	**.906**	**.903**	.935	.930	.937	.926	.938	.945	.946
0.20	FISHER	.947	.945	.949	.941	.956	.957	.950	.943	.957	.948	.967
	RIN	.952	.954	.947	.950	.951	.957	.949	.950	.950	.953	.950
	SPEAR	.950	.953	.957	.948	.956	.959	.952	.951	.962	.951	.954
	PB	**.915**	**.923**	.936	**.919**	.945	.947	.930	.931	.947	.945	.954
	BCA	.929	**.920**	.932	**.913**	.937	.937	.929	**.921**	.944	.940	.948
	HPDI	**.901**	**.903**	**.920**	**.906**	.939	.943	**.923**	**.921**	.937	.936	.949
0.30	FISHER	.946	.968	.962	.957	.953	.953	.947	.956	.961	.948	.956
	RIN	.942	.968	.963	.952	.948	.953	.948	.956	.949	.954	.95
	SPEAR	.948	.962	.959	.951	.953	.955	.952	.954	.963	.956	.959
	PB	**.918**	.929	.932	.927	.931	.941	.933	.941	.945	.937	.945
	BCA	.943	.930	.928	**.922**	**.920**	.927	.929	.934	.935	.928	.938
	HPDI	**.902**	**.910**	.926	**.914**	**.922**	.936	.927	.936	.938	.932	.946
0.40	FISHER	.958	.960	.957	.950	.966	.955	.949	.964	.970	.959	.961
	RIN	.957	.952	.957	.950	.954	.958	.941	.954	.959	.948	.961
	SPEAR	.954	.950	.952	.943	.959	.953	.942	.947	.958	.953	.959
	PB	.928	**.919**	.932	**.908**	.945	.941	.929	.940	.964	.950	.944
	BCA	.941	**.917**	**.921**	**.904**	.938	.929	**.920**	.934	.955	.946	.941
	HPDI	**.908**	**.900**	**.911**	**.893**	.936	.931	**.922**	.932	.961	.943	.937
0.50	FISHER	.966	.962	.955	.956	.956	.962	.967	.964	.967	.959	.951
	RIN	.961	.965	.954	.963	.955	.957	.959	.958	.961	.959	.954
	SPEAR	.958	.948	.946	.949	.948	.948	.945	.948	.952	.951	.950
	PB	**.924**	.925	**.918**	.933	.937	.929	.952	.950	.952	.943	.934
	BCA	.935	**.918**	.927	**.921**	.928	**.920**	.952	.943	.946	.932	.925
	HPDI	**.910**	**.906**	**.909**	.928	.934	**.923**	.951	.937	.945	.938	.930

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 26
95% Coverage Probabilities for Normal Distribution Paired with Pareto Distribution
(Skewness = 2.811, Kurtosis =14.828)

ρ	*n*	20	30	40	50	100	150	200	250	300	350	400
0	FISHER	.938	.952	.948	.956	.954	.947	.939	.947	.955	.953	.957
	RIN	.945	.956	.945	.954	.958	.953	.957	.960	.957	.960	.961
	SPEAR	.950	.966	.954	.956	.953	.955	.959	.967	.969	.957	.959
	PB	**.923**	.929	.927	.936	.947	.945	.933	.938	.941	.948	.953
	BCA	.937	.941	.929	.940	.940	.937	.935	.936	.939	.944	.949
	HPDI	**.904**	**.914**	**.918**	**.923**	.936	.943	.925	.936	.934	.948	.952
0.10	FISHER	.947	.955	.941	.953	.941	.959	.953	.950	.936	.947	.947
	RIN	.949	.958	.937	.952	.946	.960	.945	.950	.945	.951	.948
	SPEAR	.954	.953	.947	.954	.955	.952	.951	.953	.951	.963	.955
	PB	.925	**.917**	**.907**	.934	.934	.951	.947	.947	.934	.938	.944
	BCA	.931	**.924**	**.914**	.936	.926	.948	.945	.940	.927	.937	.937
	HPDI	**.902**	**.904**	**.894**	**.918**	.927	.947	.942	.940	**.922**	.936	.938
0.20	FISHER	.943	.943	.959	.947	.953	.950	.949	.956	.954	.940	.964
	RIN	.947	.959	.949	.951	.955	.946	.961	.959	.948	.950	.960
	SPEAR	.951	.957	.948	.952	.964	.957	.969	.964	.951	.960	.958
	PB	**.910**	.929	.929	.932	.937	.939	.939	.938	.948	.933	.953
	BCA	**.924**	.930	.925	.927	.932	.930	.930	.931	.941	**.924**	.948
	HPDI	**.888**	**.916**	**.913**	**.919**	.934	.936	.933	.930	.943	.925	.951
0.30	FISHER	.962	.958	.953	.951	.950	.953	.953	.955	.948	.952	.965
	RIN	.958	.945	.950	.945	.948	.949	.947	.961	.953	.949	.947
	SPEAR	.959	.955	.947	.949	.949	.952	.946	.962	.963	.954	.951
	PB	.929	**.922**	.930	**.920**	.928	.939	.948	.945	.943	.945	.952
	BCA	.939	.931	.925	**.922**	**.922**	.929	.941	.938	.935	.938	.950
	HPDI	**.905**	**.899**	**.911**	**.903**	**.923**	.935	.937	.943	.941	.944	.952
0.40	FISHER	.965	.958	.959	.963	.966	.957	.959	.960	.963	.966	.963
	RIN	.960	.955	.955	.949	.958	.961	.963	.948	.947	.954	.943
	SPEAR	.954	.960	.952	.947	.956	.957	.956	.942	.951	.949	.943
	PB	**.920**	**.914**	**.922**	.933	.946	.939	.948	.937	.945	.945	.950
	BCA	.932	**.919**	**.920**	.932	.933	.926	.942	.941	.943	.939	.941
	HPDI	**.910**	**.891**	**.900**	**.922**	.933	.929	.943	.936	.936	.942	.937
0.50	FISHER	.952	.959	.962	.954	.970	.963	.966	.943	.963	.964	.957
	RIN	.941	.946	.959	.958	.960	.963	.957	.951	.959	.956	.939
	SPEAR	.946	.930	.955	.954	.941	.966	.953	.946	.942	.951	.936
	PB	**.902**	**.901**	.931	.928	.940	.942	.943	.927	.947	.939	.929
	BCA	**.913**	**.904**	.931	.929	.939	.935	.939	**.914**	.934	.936	**.922**
	HPDI	**.877**	**.883**	**.916**	**.914**	.929	.927	.936	**.920**	.942	.932	.927

Note. Unacceptable coverage is bolded and outside [.925, .975]. FISHER = Fisher *z*-transformation; RIN = ranked inverse normal transformation; SPEAR = Spearman rank-order with Fieller's *SE* ; PB = percentile bootstrap; BCa = biased-corrected and accelerated bootstrap; HPDI = highest probability density interval. Bootstrap methods based on 2,000 bootstrap samples.

Table 27
Constants for Headrick's (2002) fifth-order polynomial transformation method

Distribution	Skew	Kurtosis	c_0	c_1	c_2	c_3	c_4	c_5
Normal	.000	.000	.000	1.000	.000	.000	.000	.000
Triangular	.000	-.600	.000	1.081	.000	-.029	.000	-.002
Uniform	.000	-1.200	.000	1.347	.000	-.140	.000	.002
Laplace	.000	3.000	.000	0.728	.000	.096	.000	-.002
Beta (a=4, b=1.25)	-.848	.221	.199	1.071	-.229	-.041	.010	.001
Beta (a=4, b=1.5)	-.694	.069	.163	1.089	-.187	-.044	.008	.001
Chi-Square (df=16)	.710	.750	-.117	.976	.117	.004	.000	.000
Chi-Square (df=4)	1.410	3.000	-.228	.901	.232	.015	-.001	.000
Chi-Square (df=3)	1.630	4.000	-.259	.867	.265	.021	-.002	.000
Chi-Square (df=2)	2.000	6.000	-.308	.801	.319	.034	-.004	.000
Chi-Square (df=1)	2.830	12.000	-.398	.621	.417	.068	-.006	.000
Pareto	2.811	14.828	-.346	.712	.347	.028	.000	.004

APPENDIX B

FIGURES

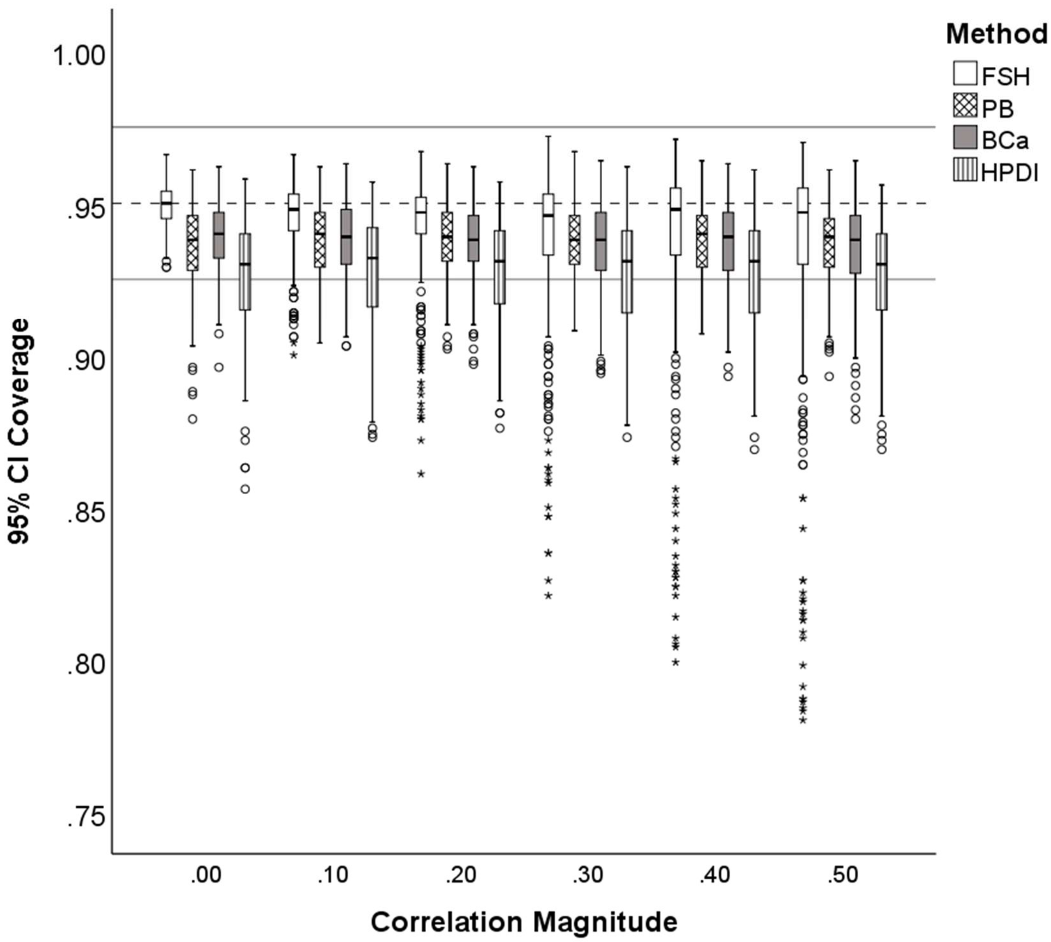

Figure 4. Distribution of 95% CI coverage for correlation magnitude. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

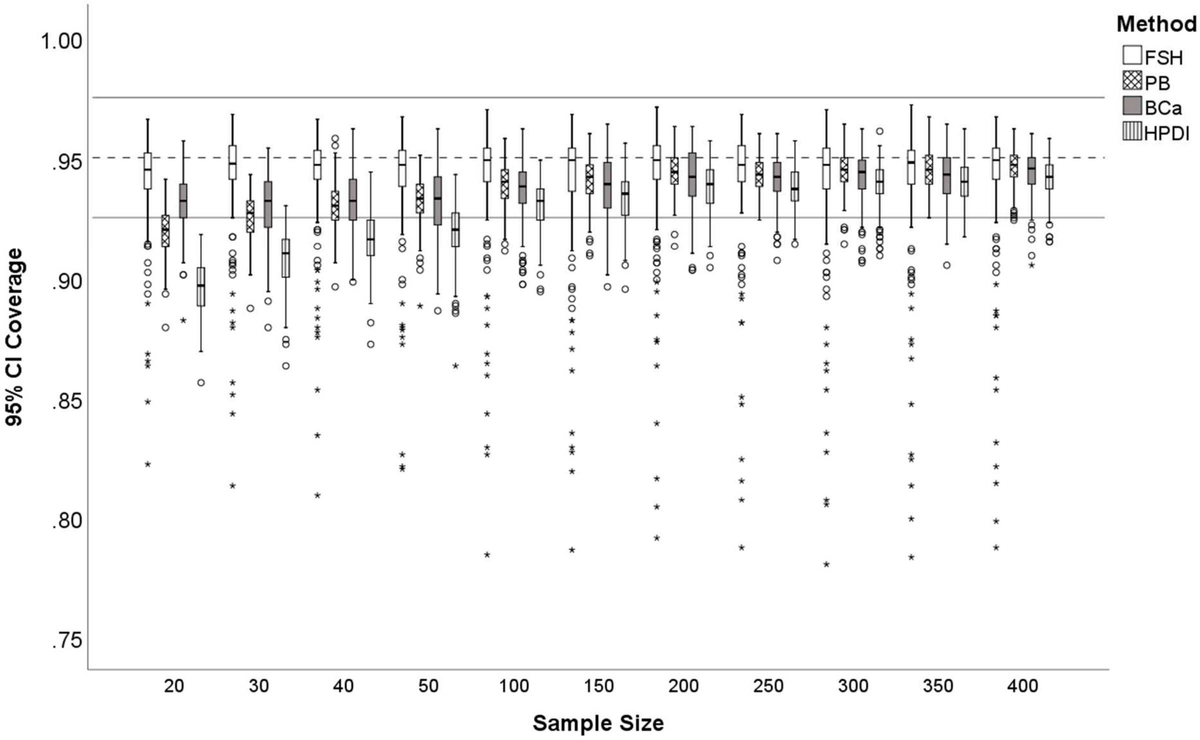

Figure 5. Distribution of 95% CI coverage for sample size. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

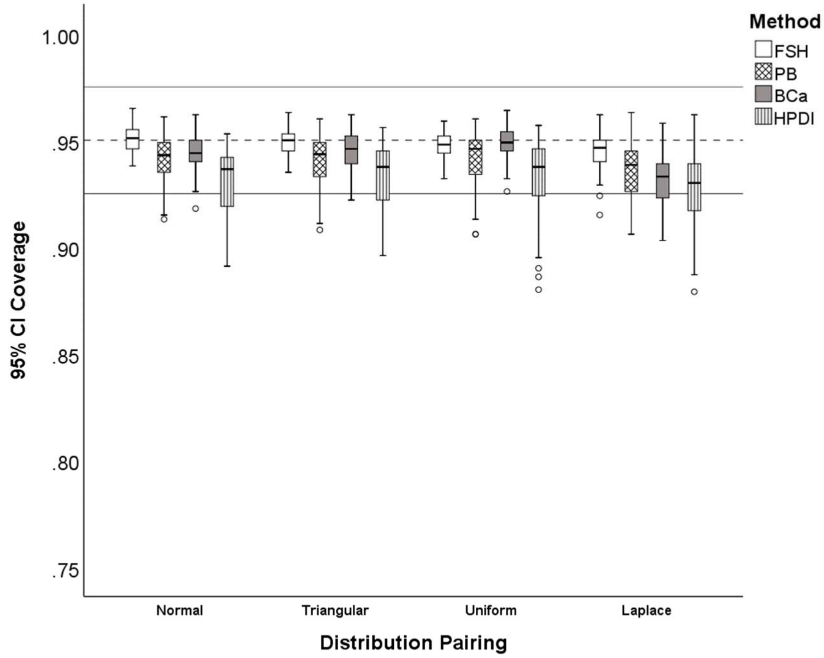

Figure 6. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

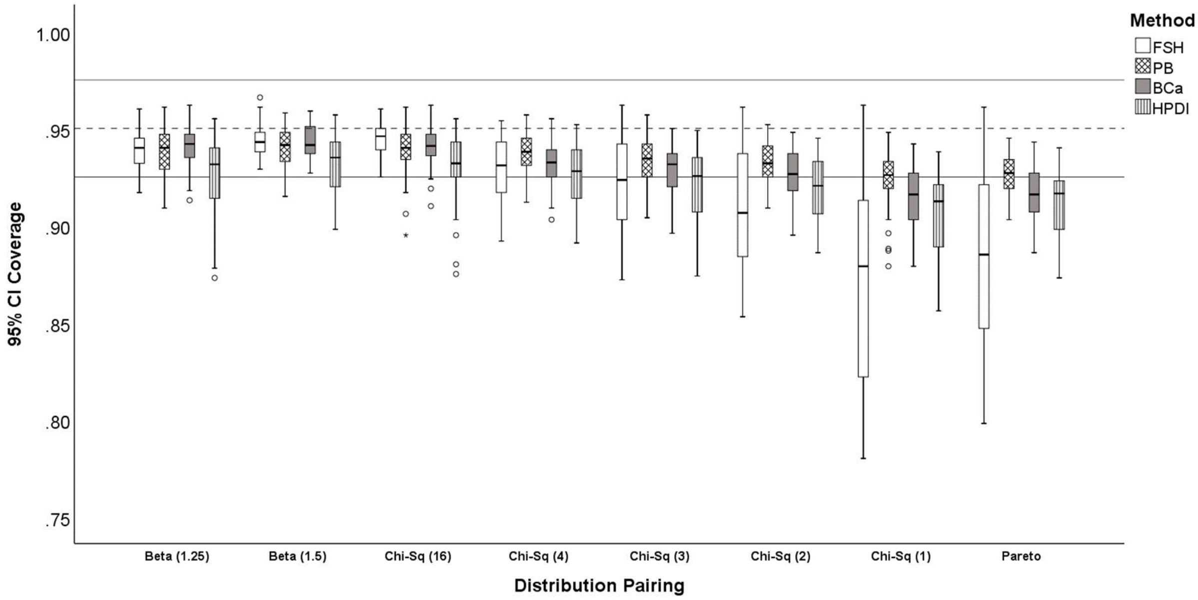

Figure 7. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

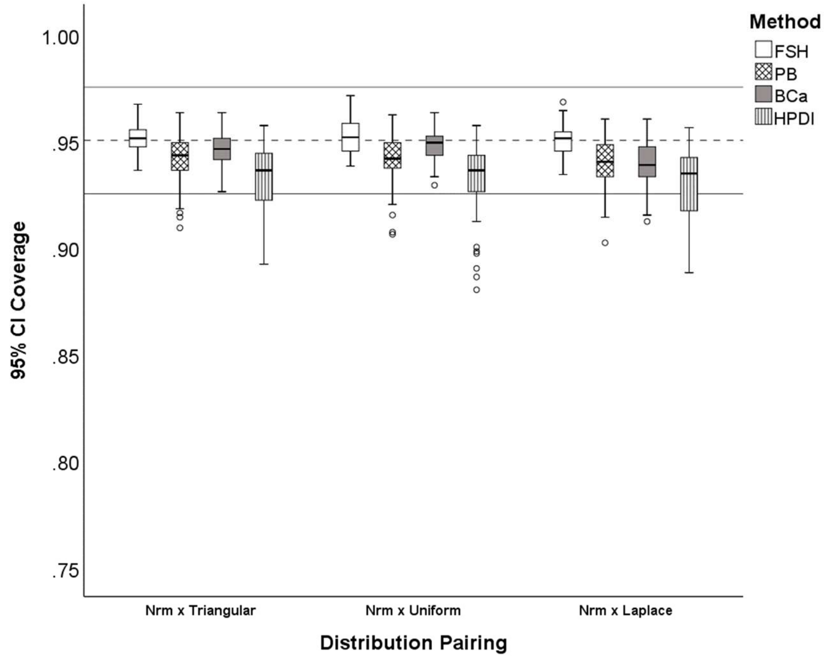

Figure 8. Distribution of 95% CI coverage for symmetric with normal distribution pairings. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

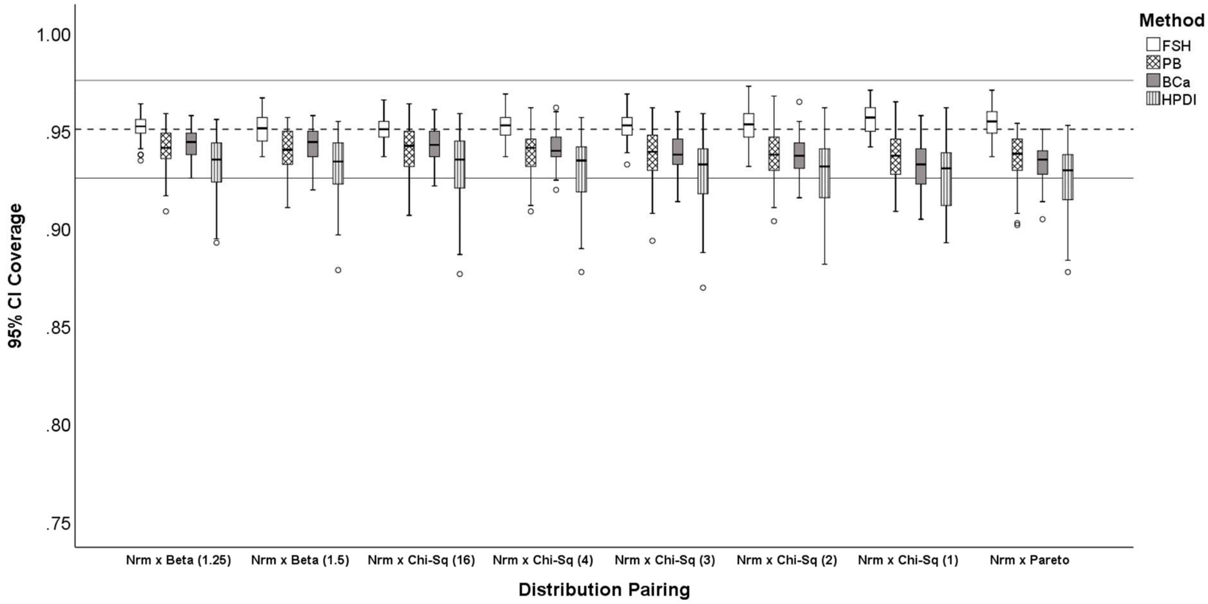

Figure 9. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

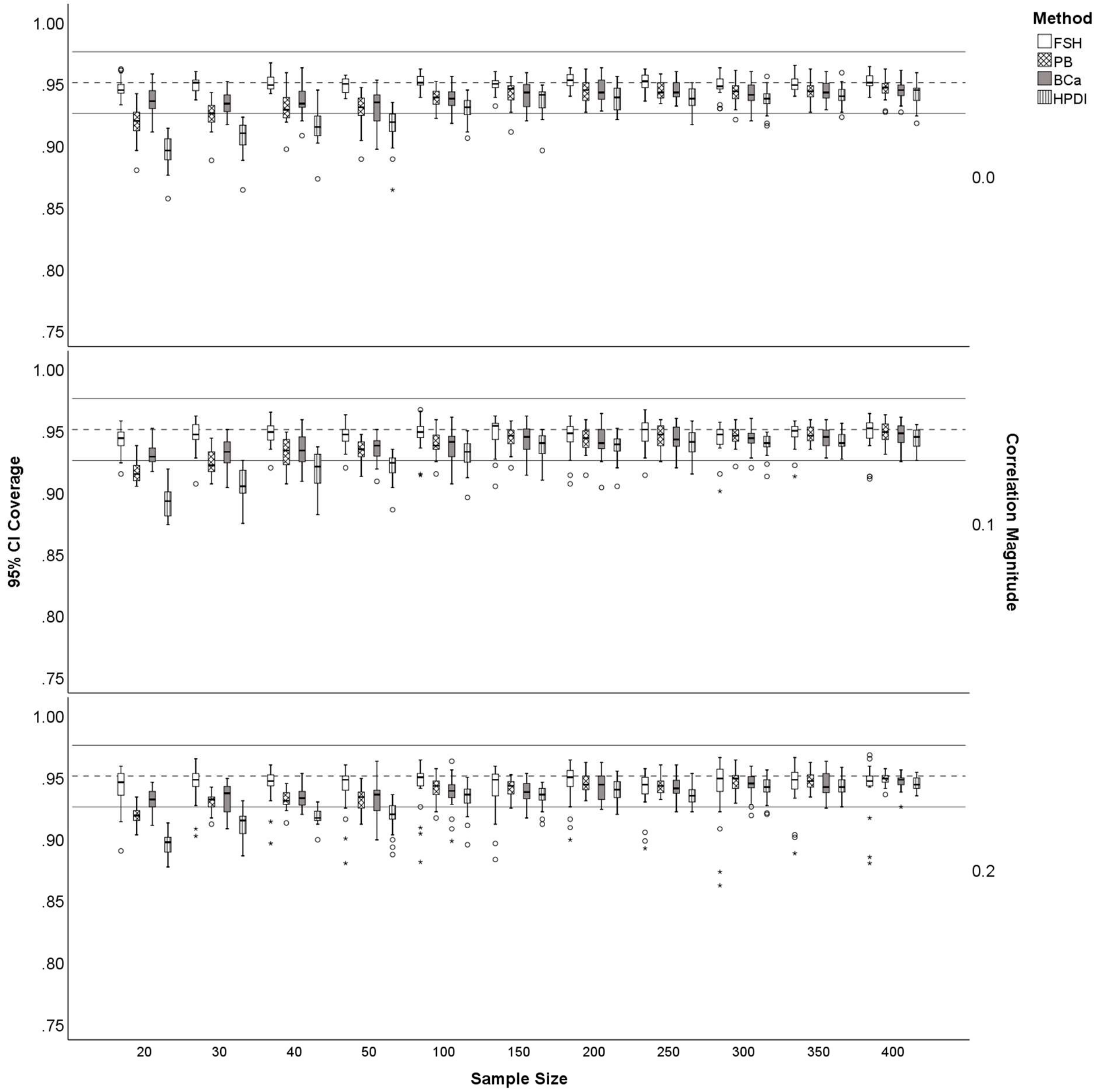

Figure 10. Distribution of 95% CI coverage for sample size by correlation magnitudes of $0-.2$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

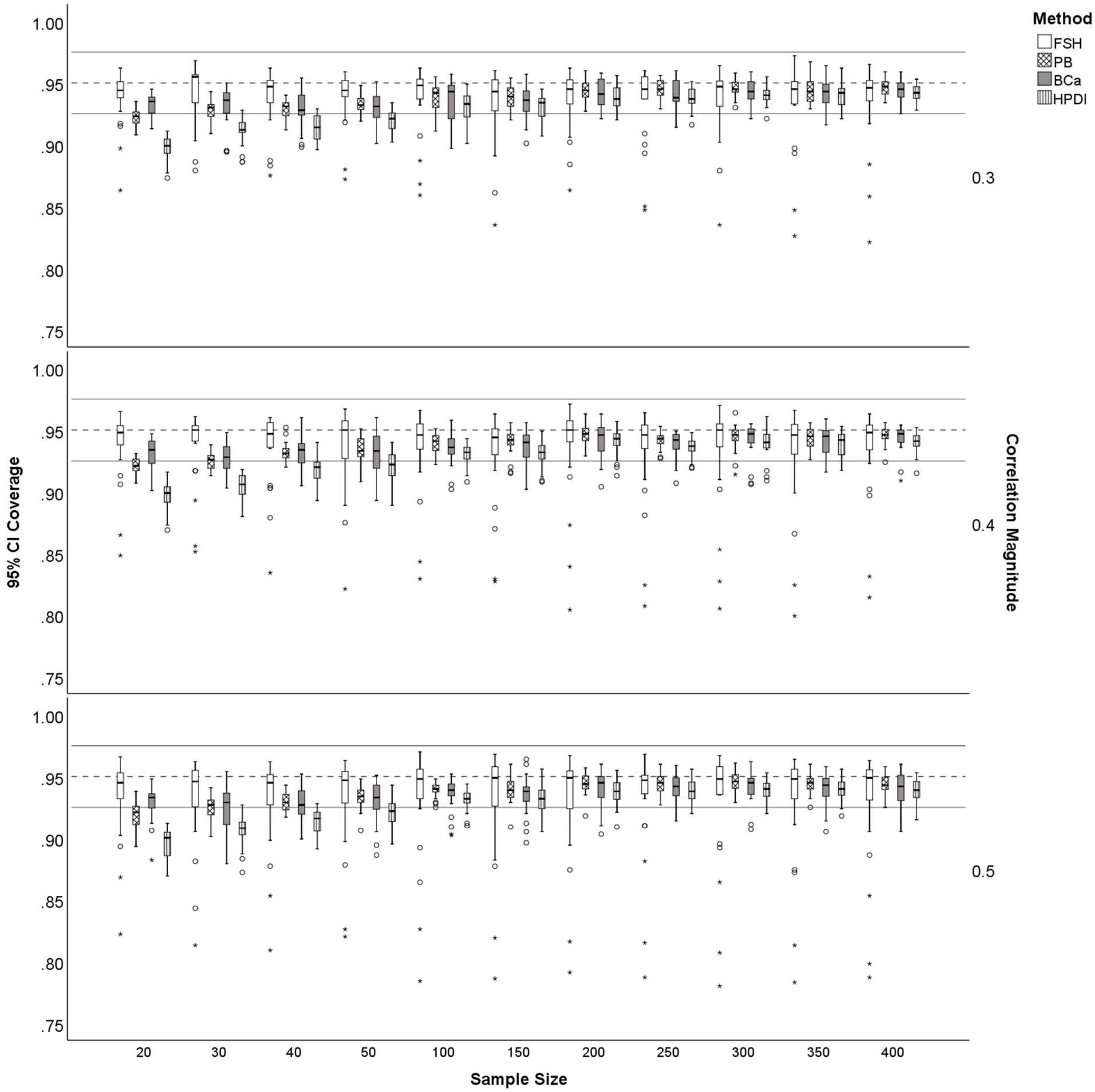

Figure 11. Distribution of 95% CI coverage for sample size by correlation magnitudes of $.3-.5$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

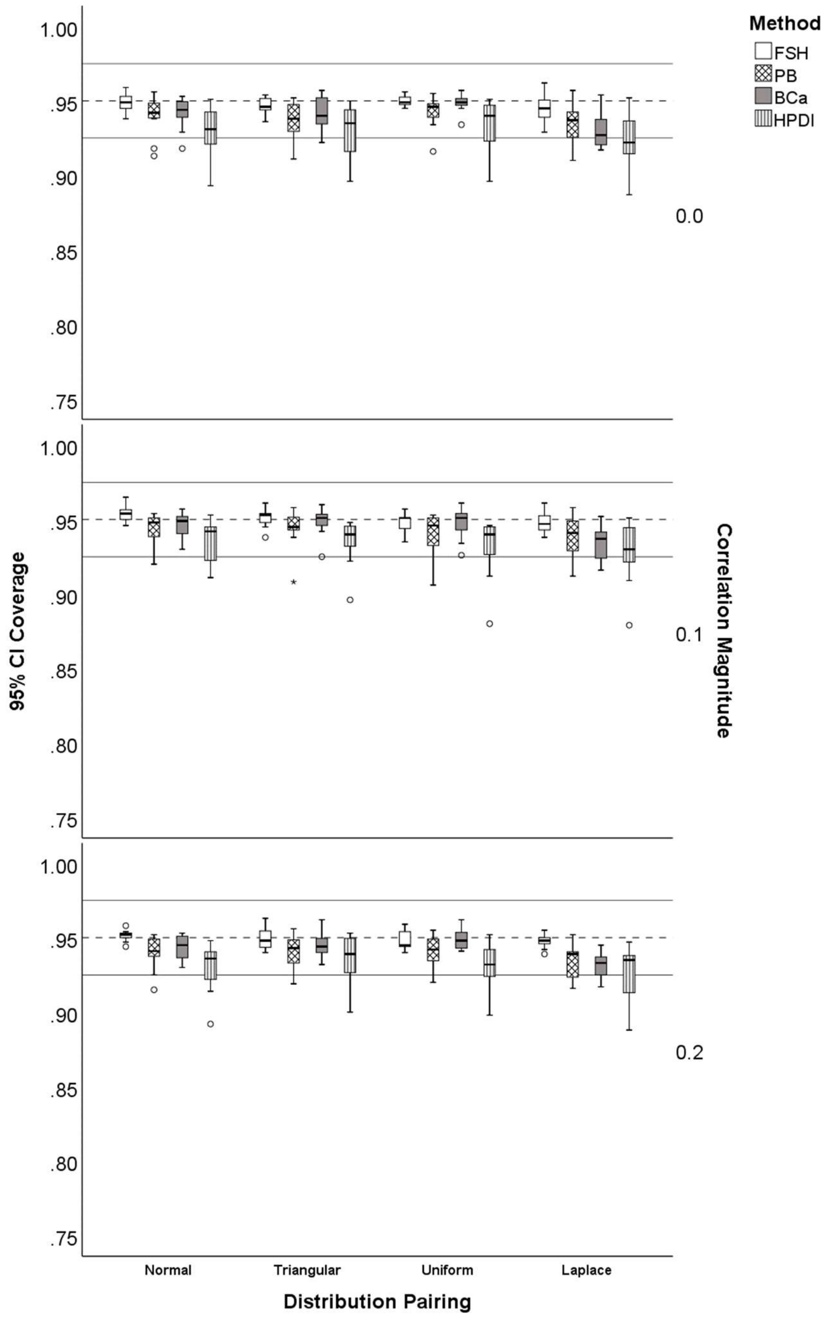

Figure 12. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings by correlation magnitudes of $0-.2$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

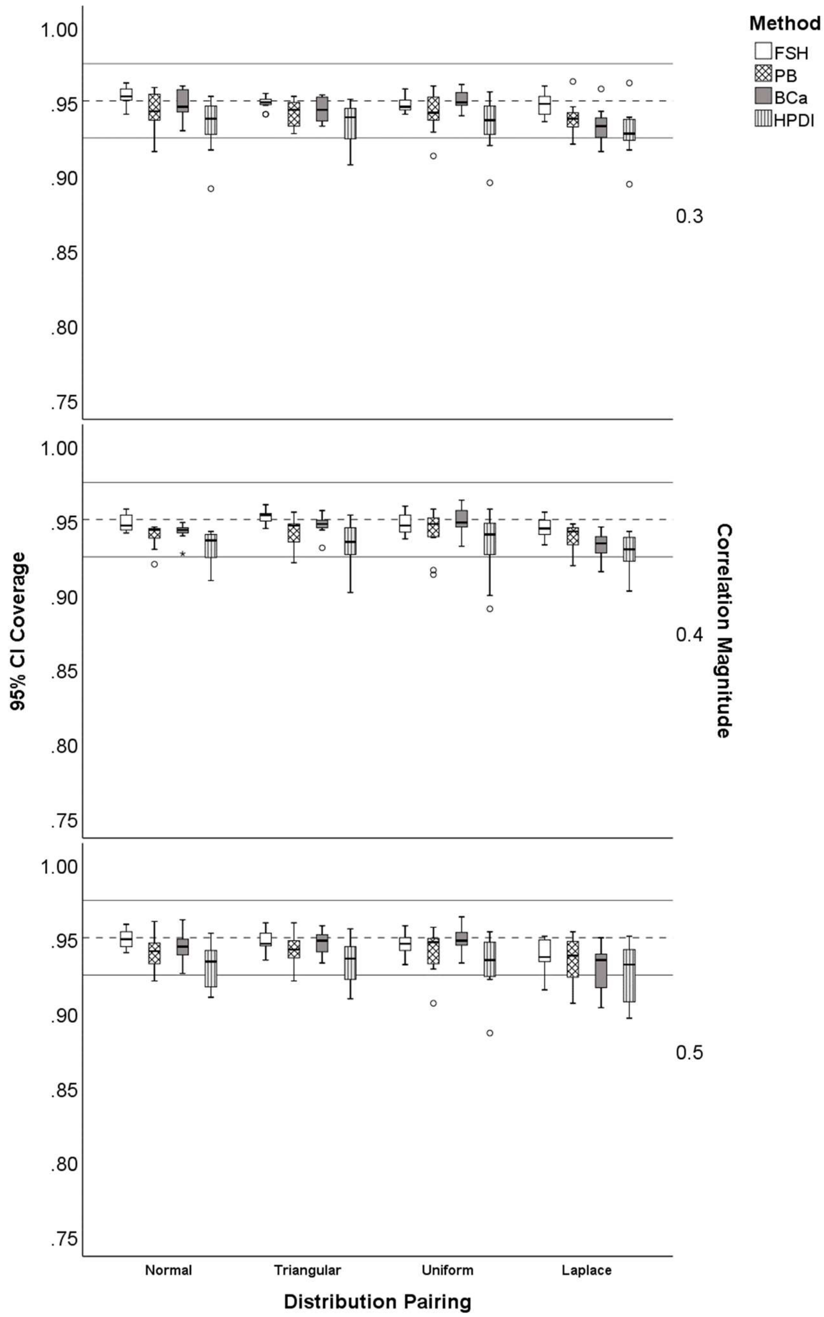

Figure 13. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings by correlation magnitudes of $.3-.5$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $\left[.925, .975\right]$; acceptable coverage.

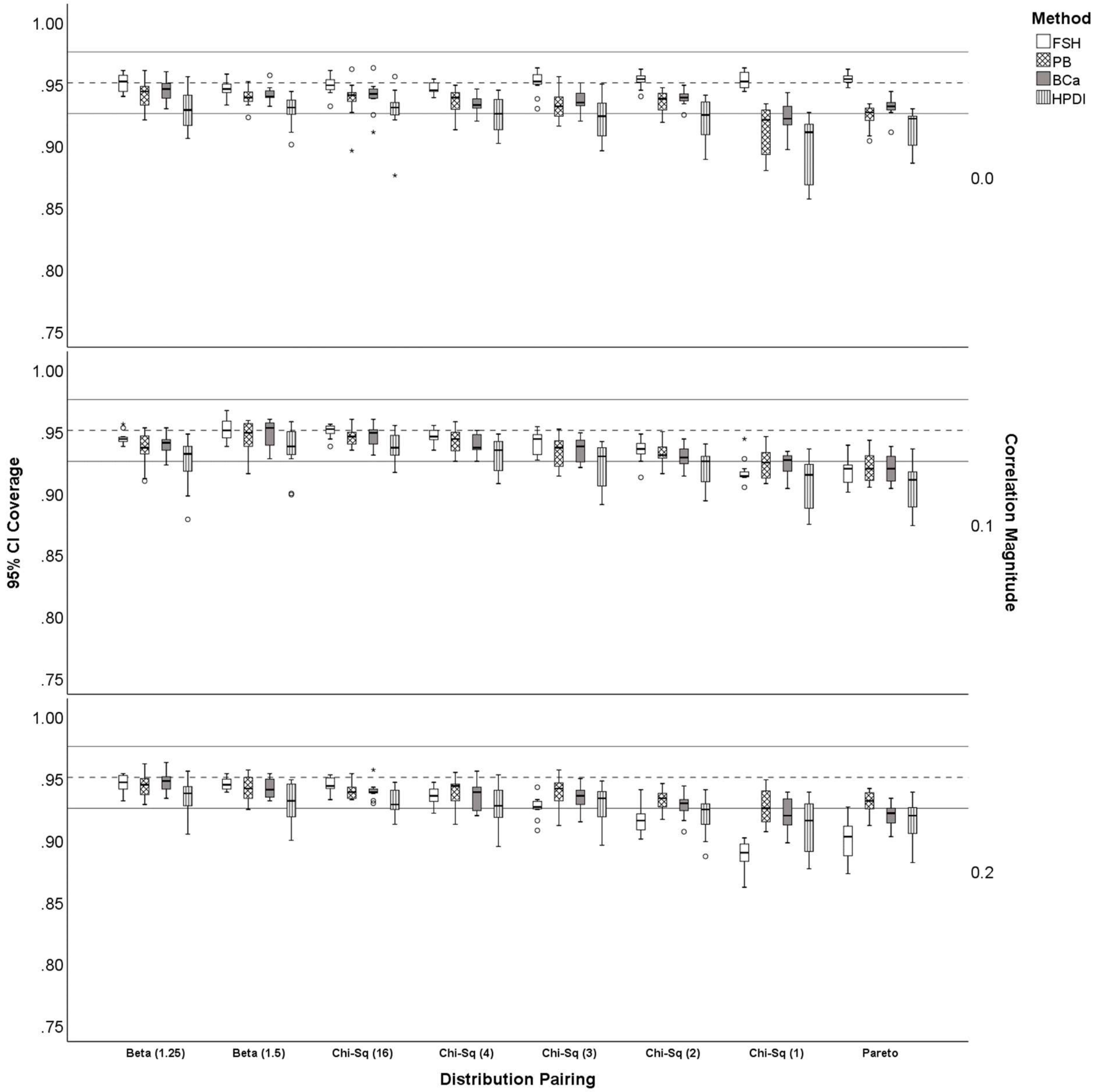

Figure 14. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings by correlation magnitudes of $0-.2$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

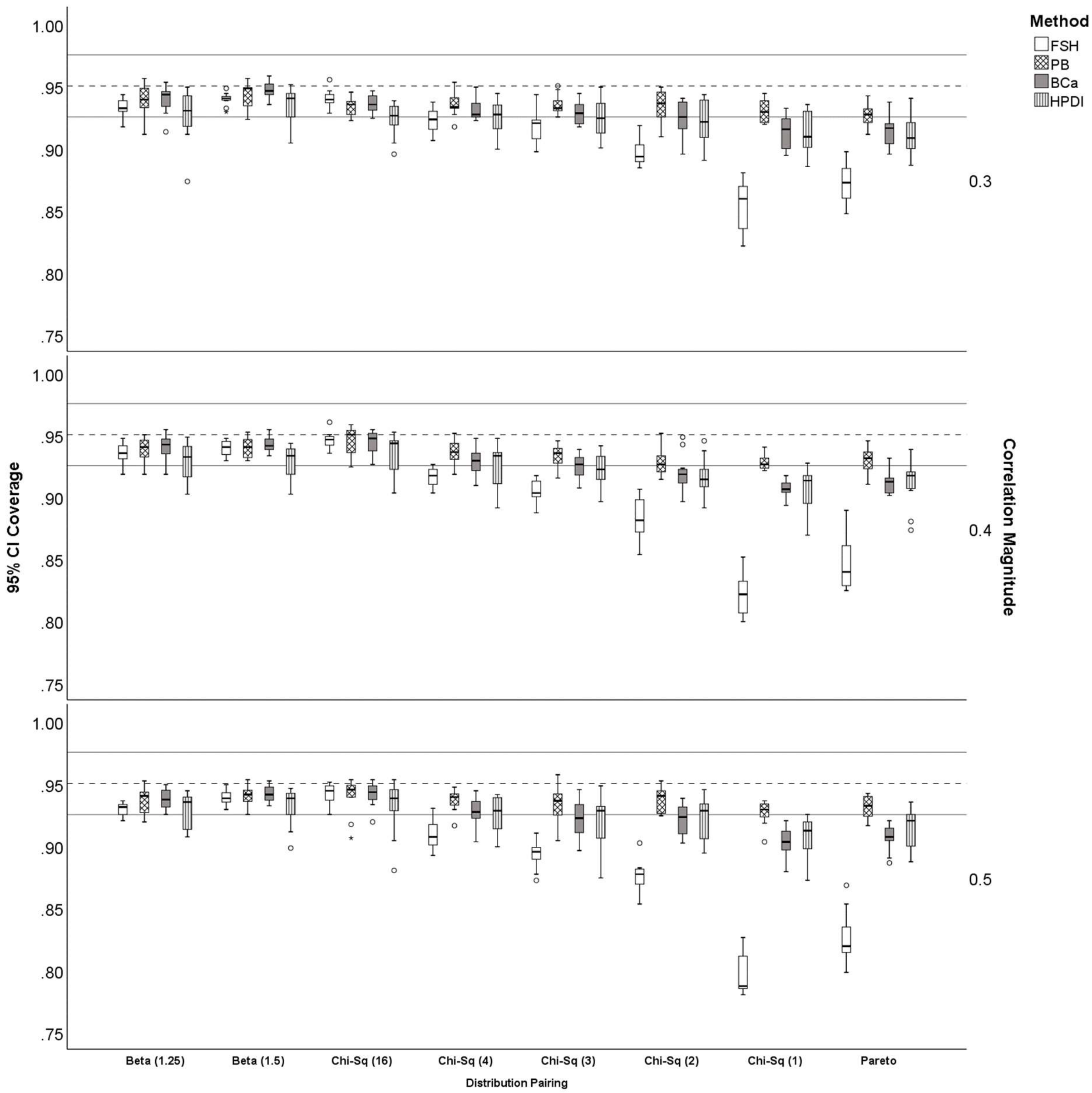

Figure 15. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings by correlation magnitudes of $.3-.5$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

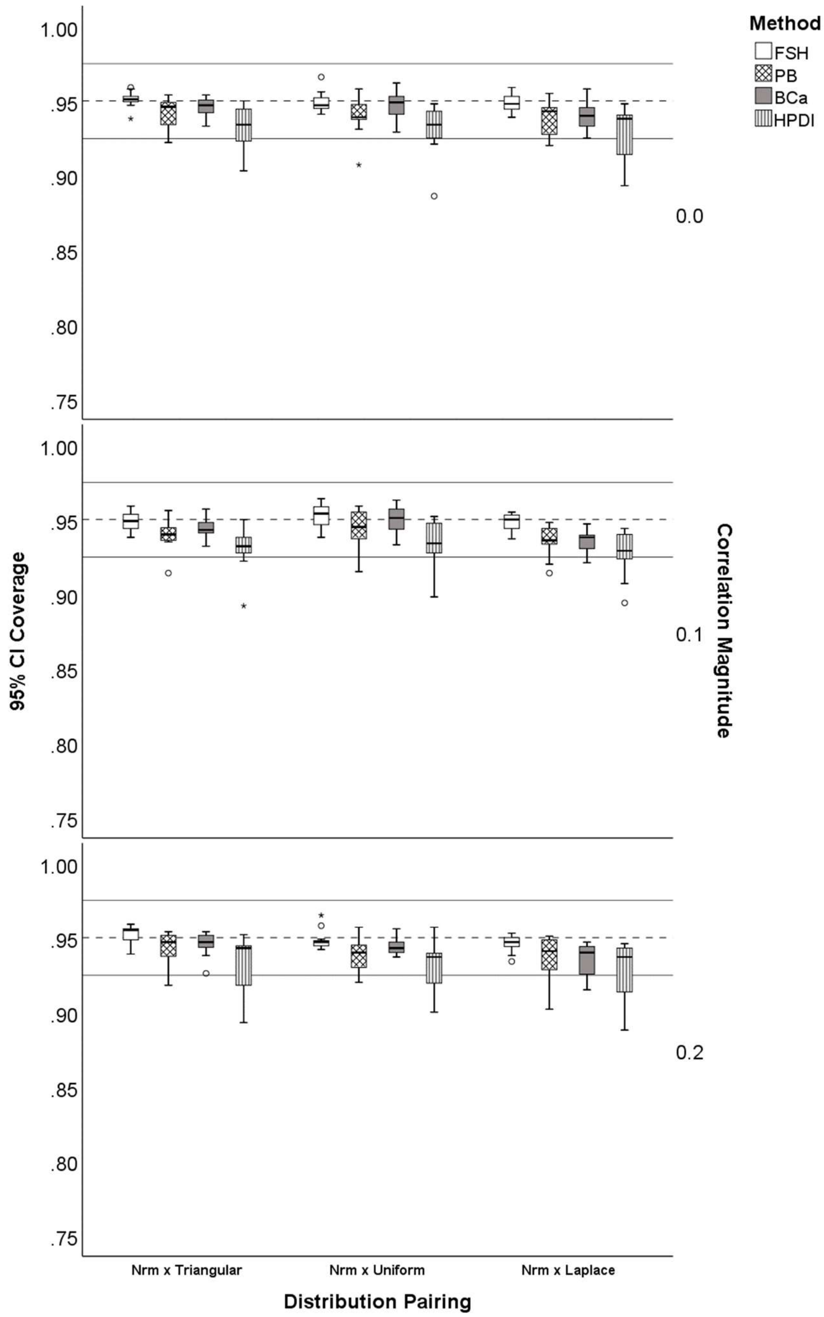

Figure 16. Distribution of 95% CI coverage for symmetric with normal distribution pairings by correlation magnitudes of $0-.2$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

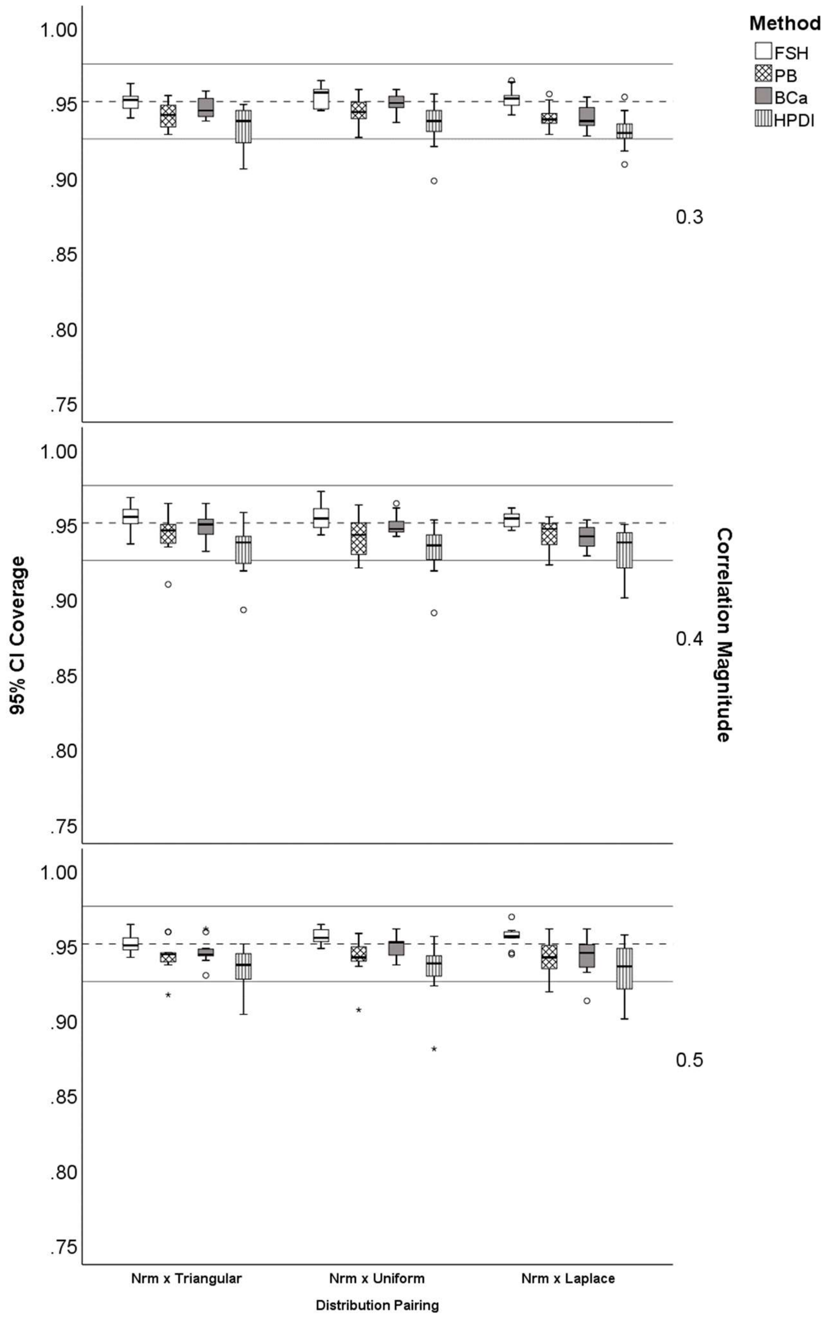

Figure 17. Distribution of 95% CI coverage for symmetric with normal distribution pairings by correlation magnitudes of $.3-.5$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

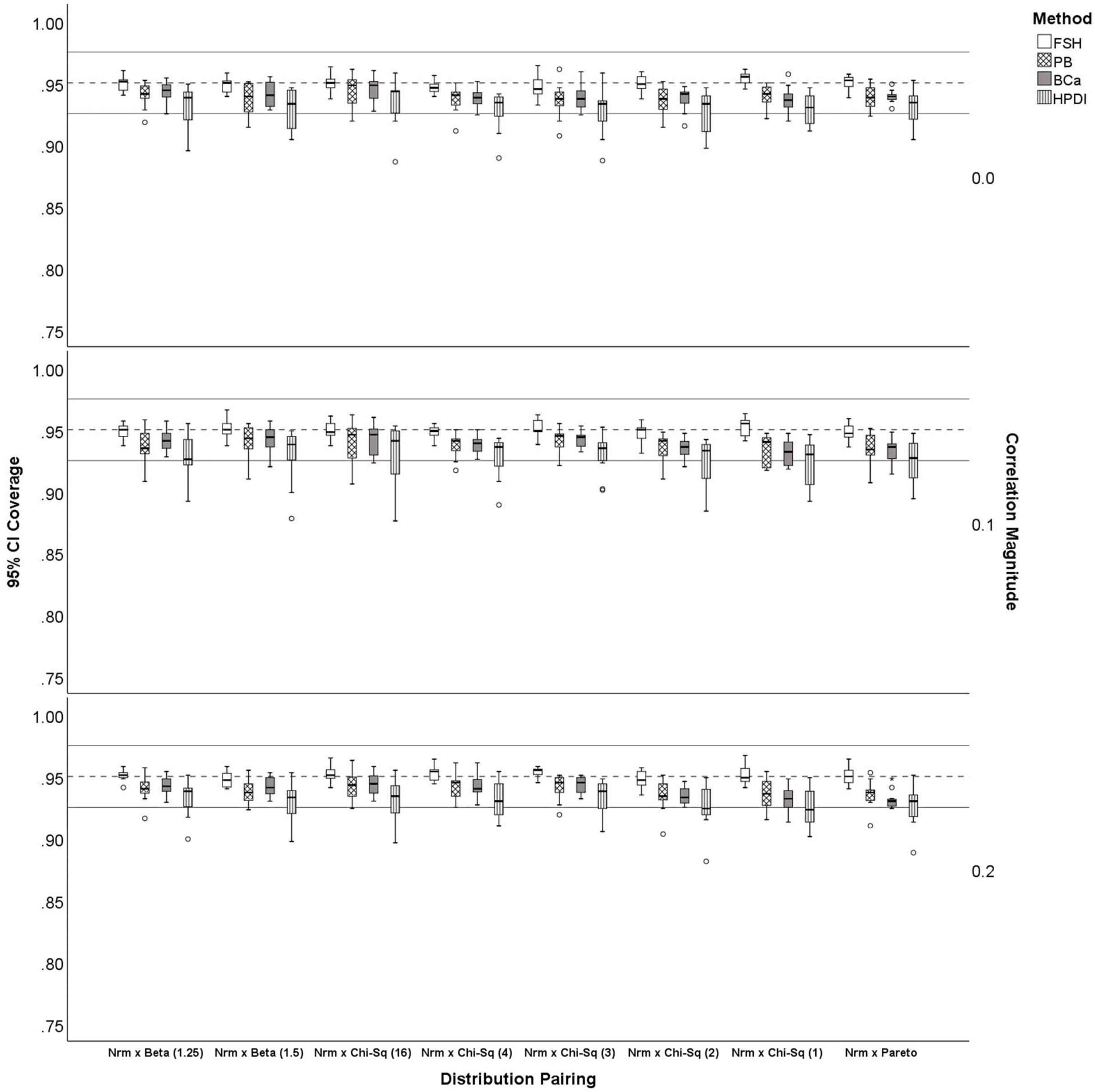

Figure 18. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings by correlation magnitudes of $0-.2$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

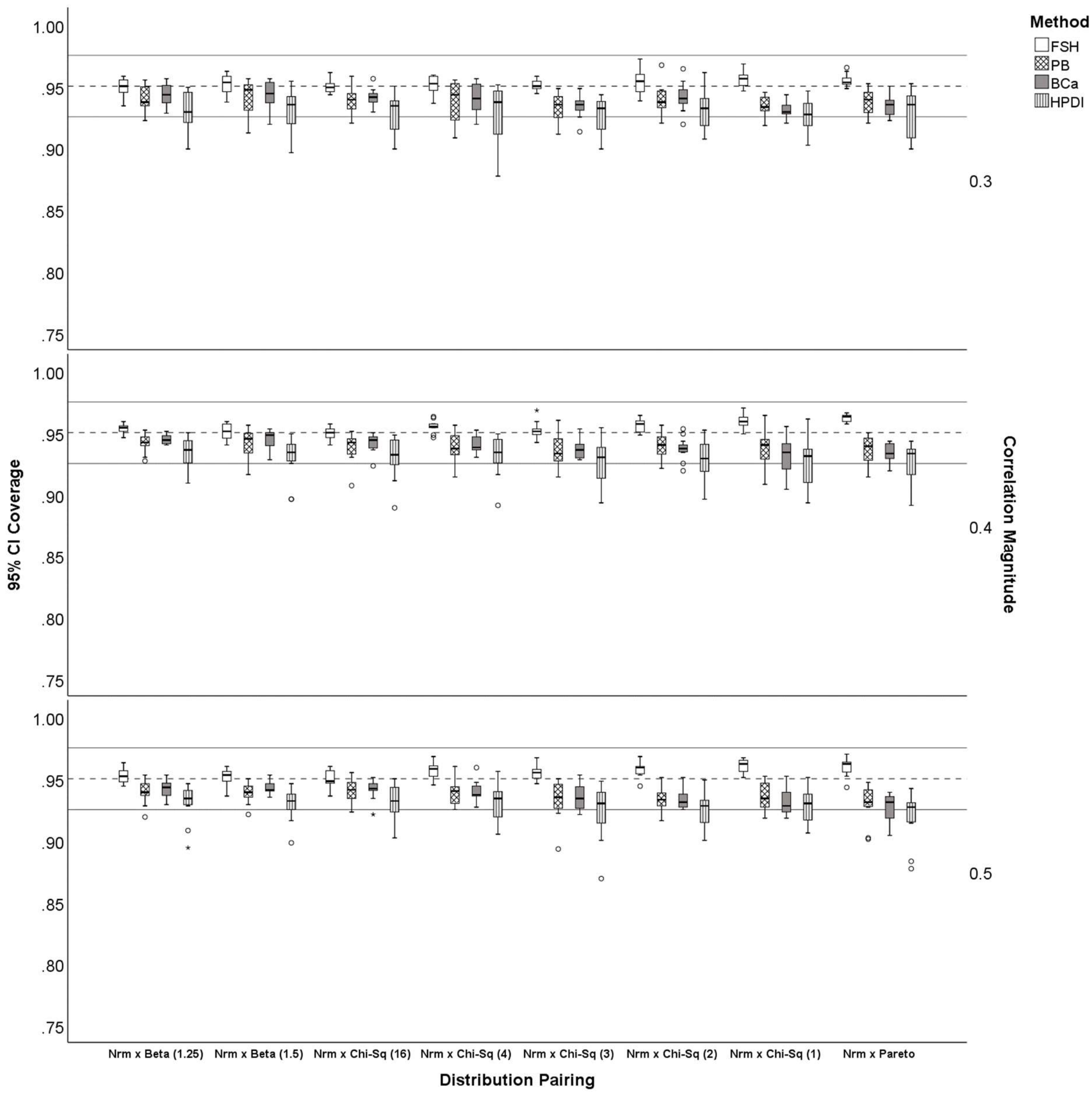

Figure 19. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings by correlation magnitudes of $.3-.5$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

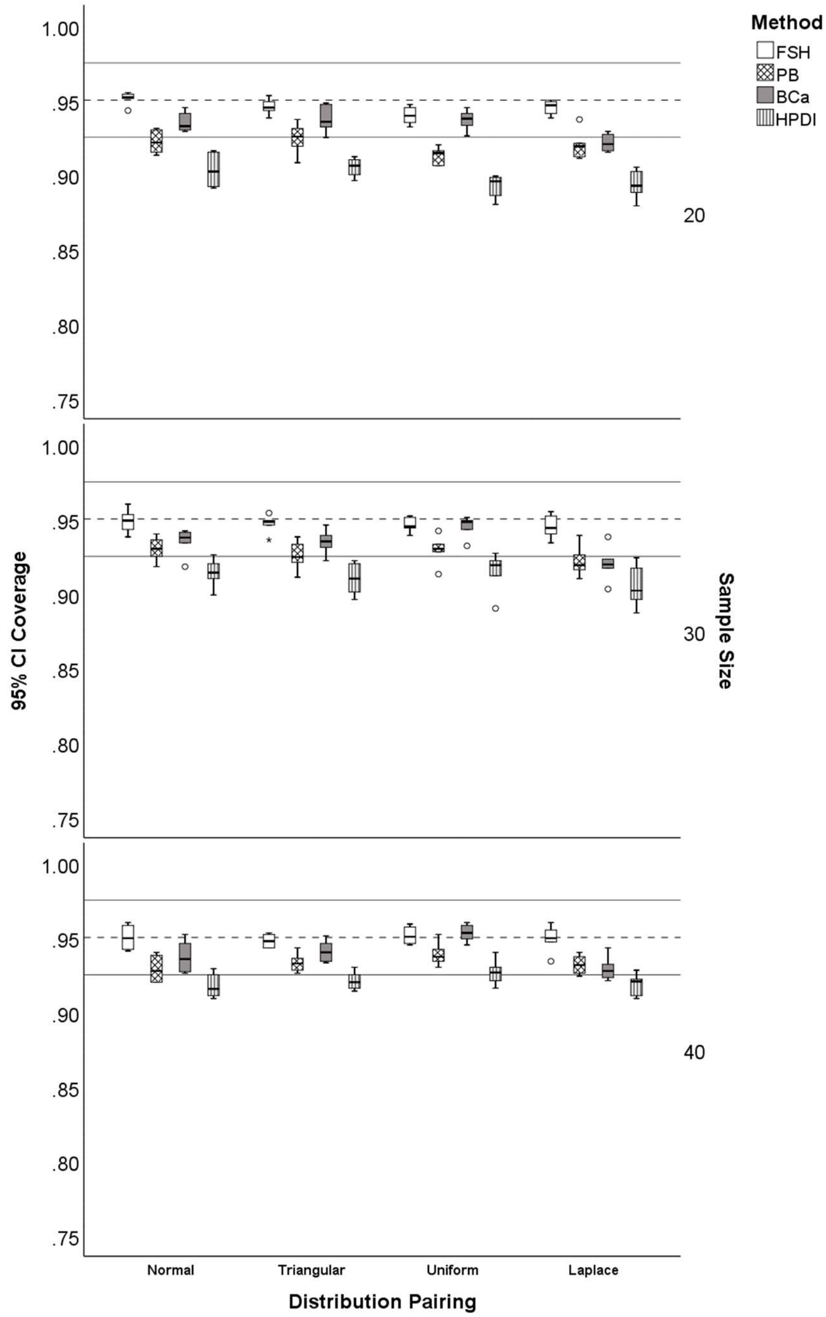

Figure 20. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings by sample size of $20-40$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

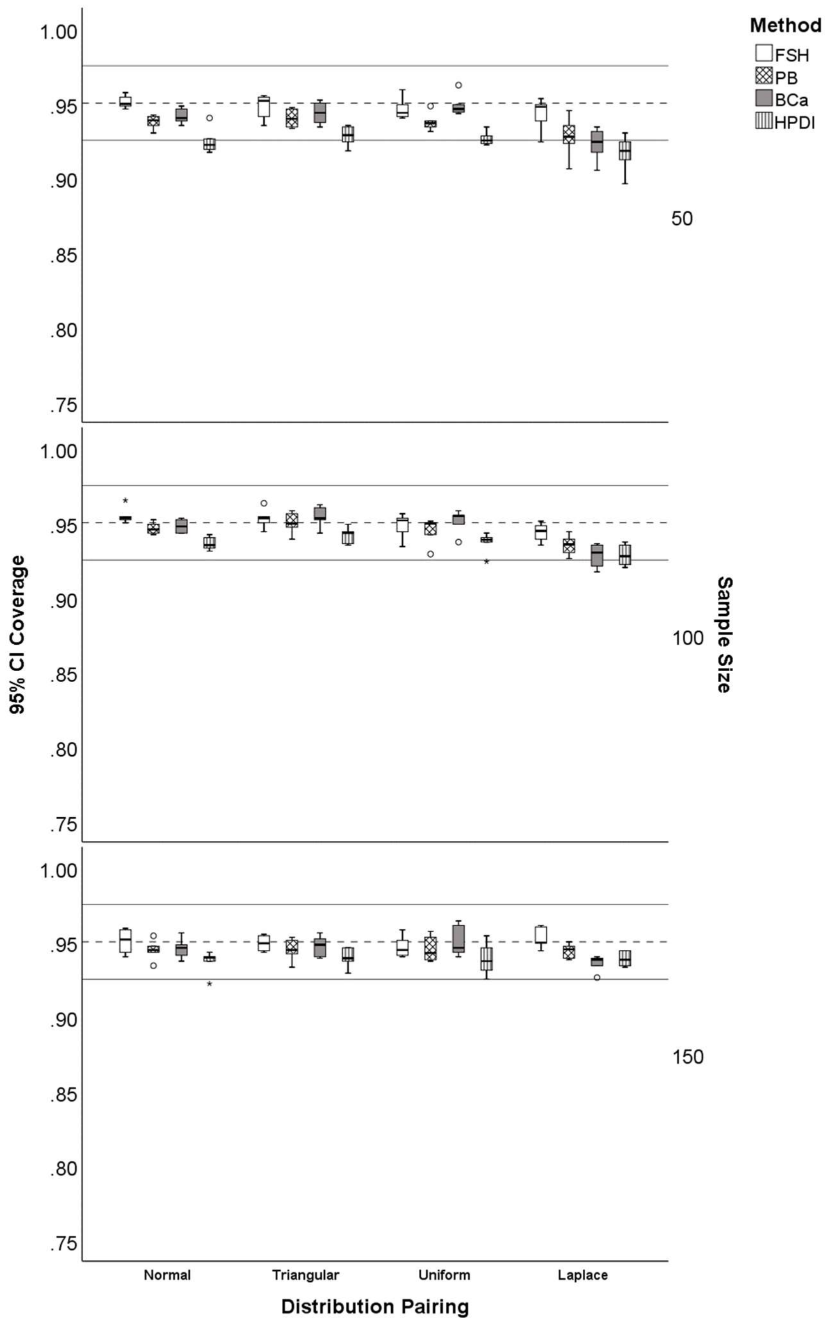

Figure 21. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings by sample size of $50-150$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

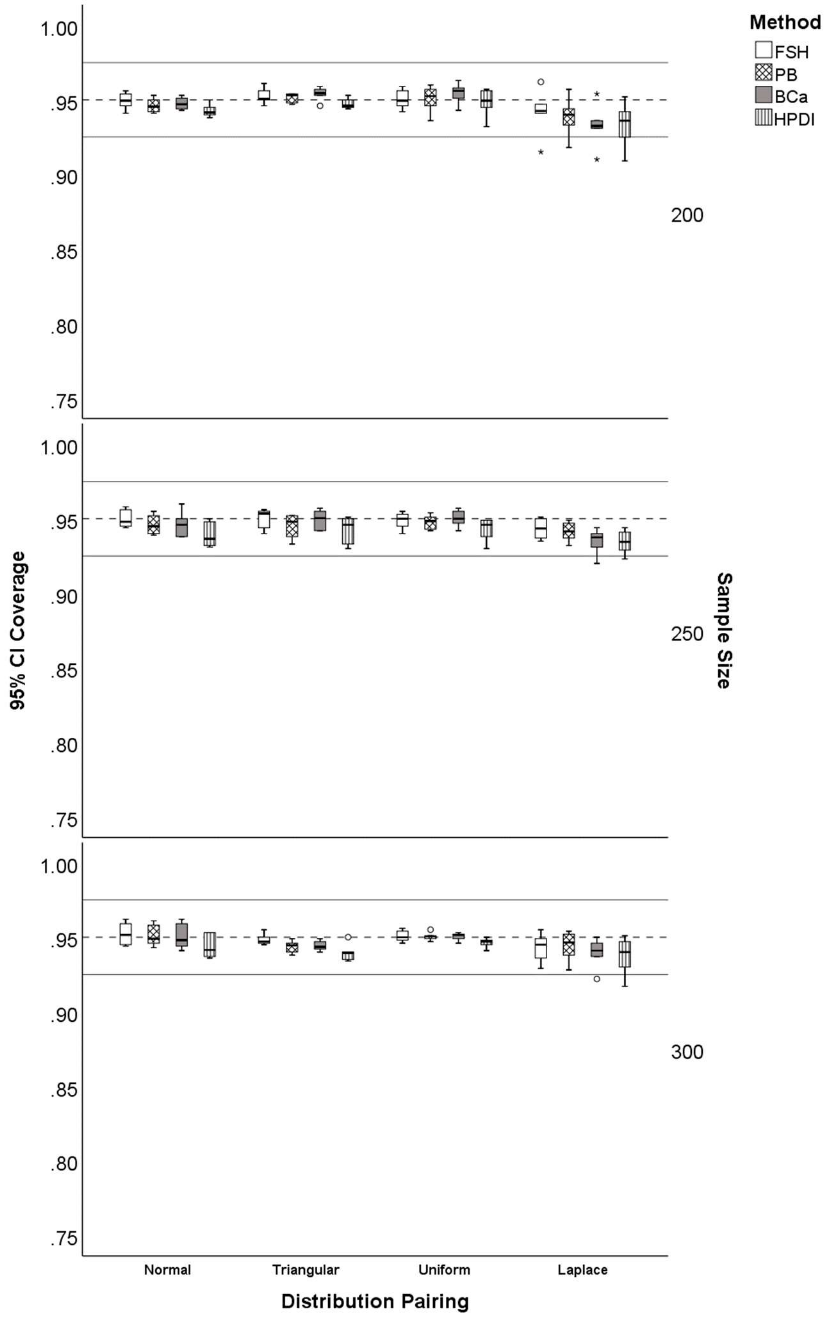

Figure 22. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings by sample size of $200-300$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

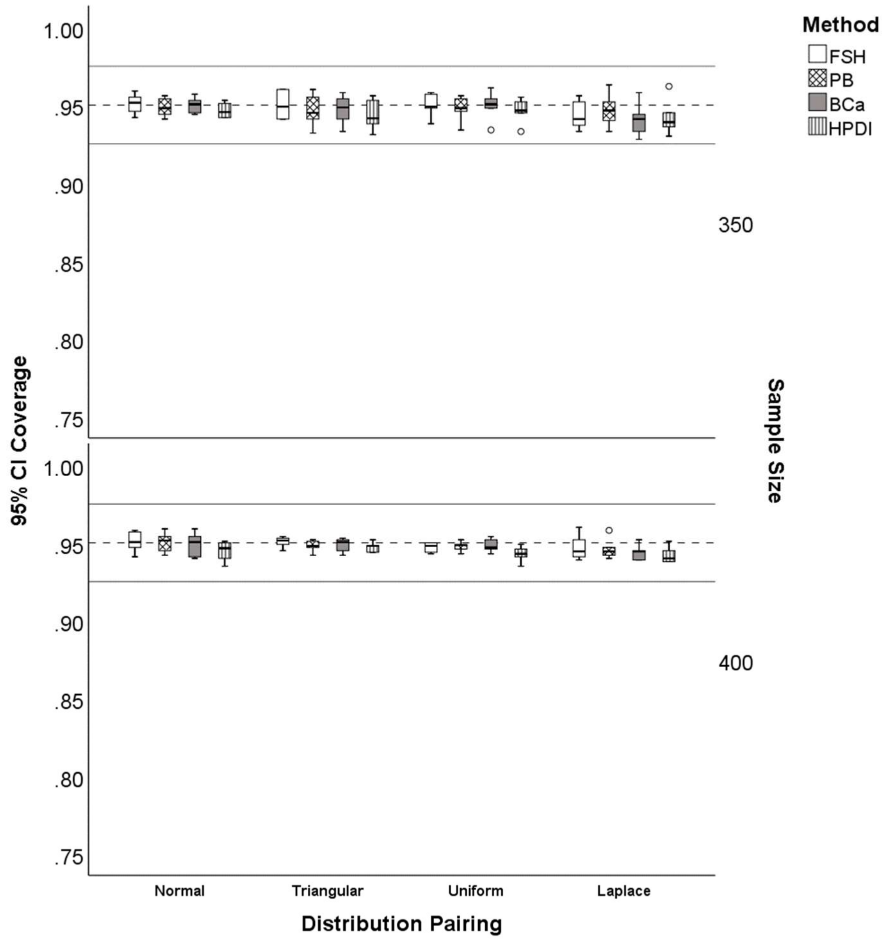

Figure 23. Distribution of 95% CI coverage for symmetric with symmetric distribution pairings by sample size of $350-400$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

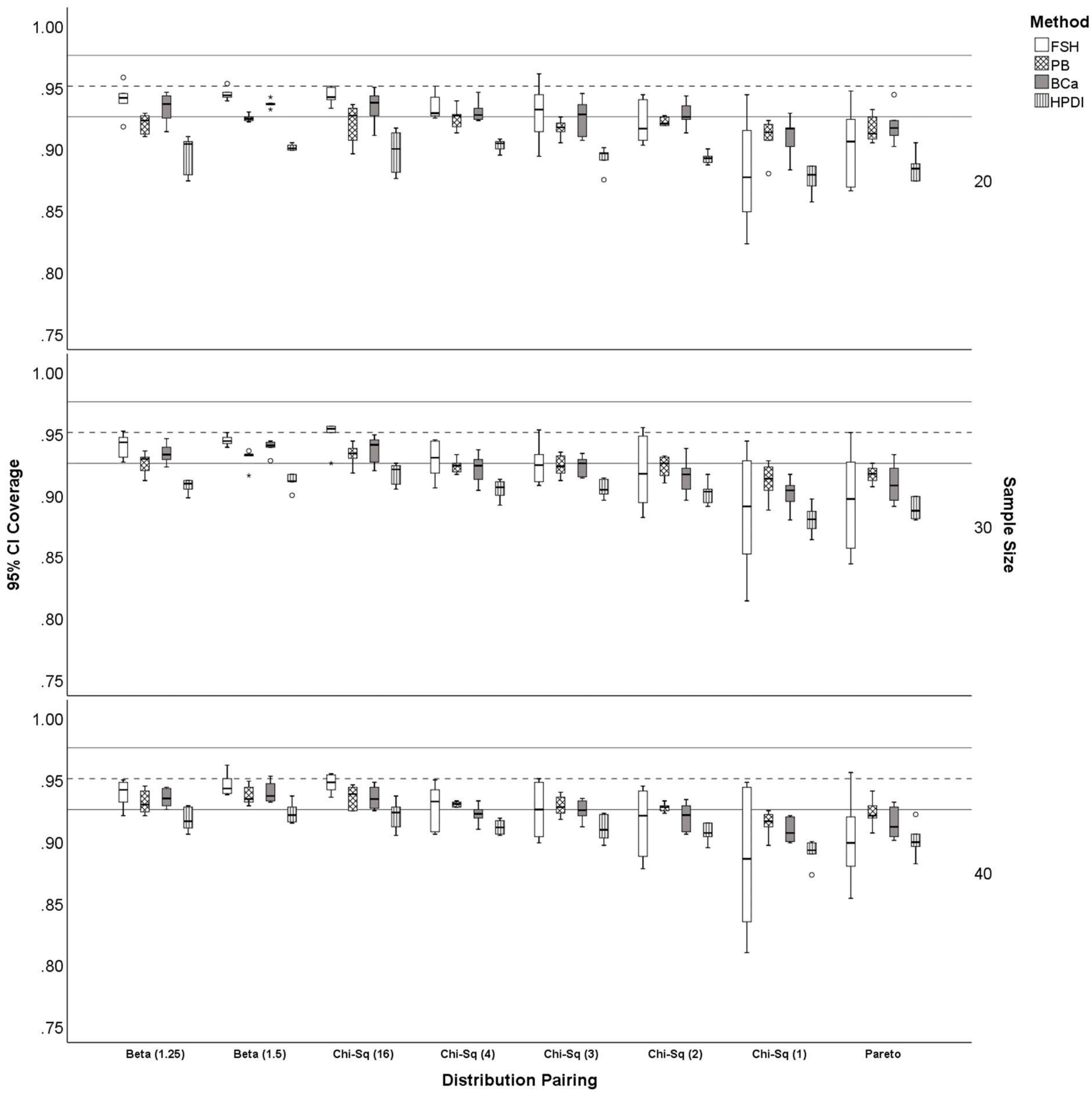

Figure 24. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings by sample size of $20-40$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

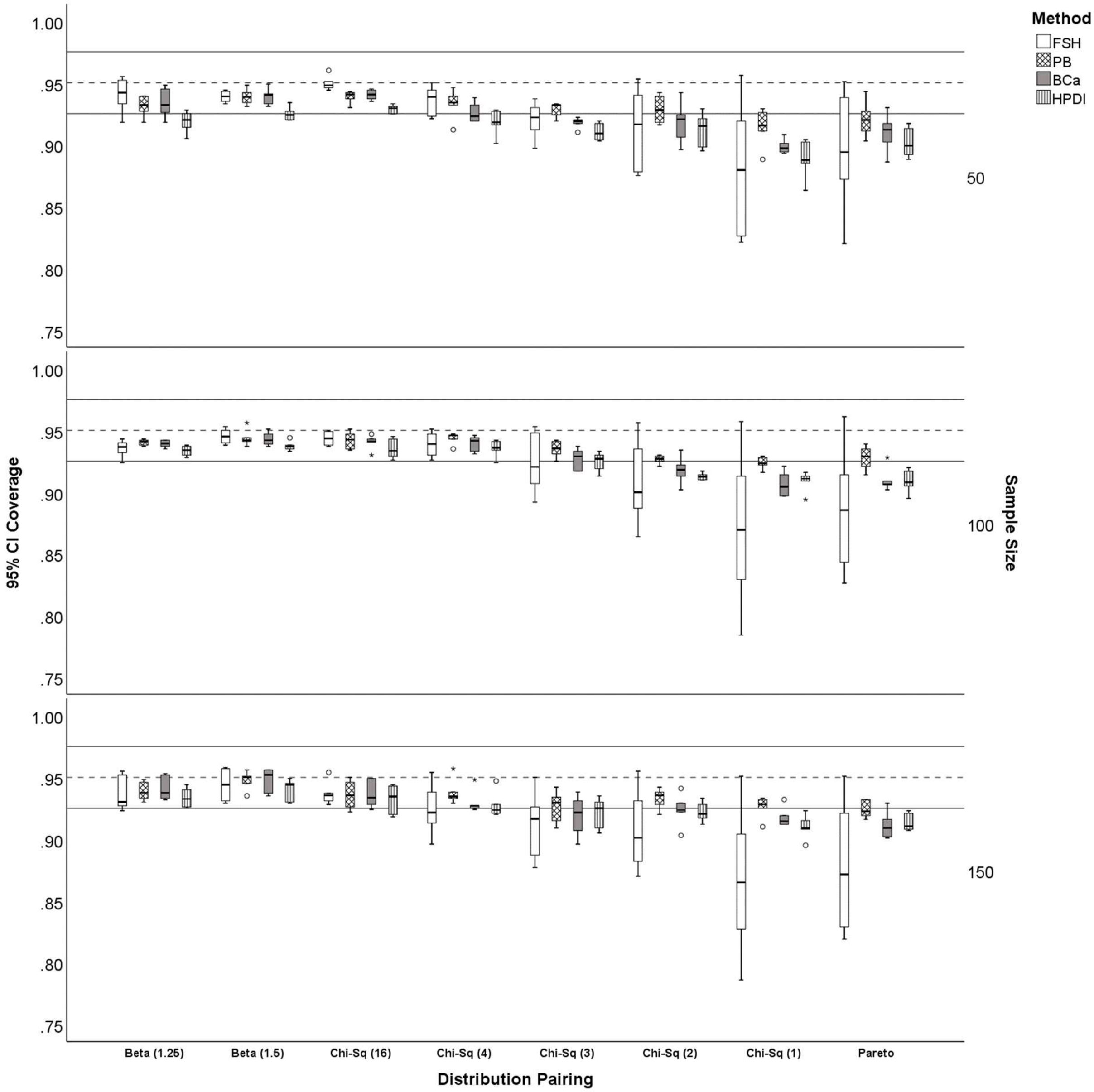

Figure 25. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings by sample size of $50-150$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

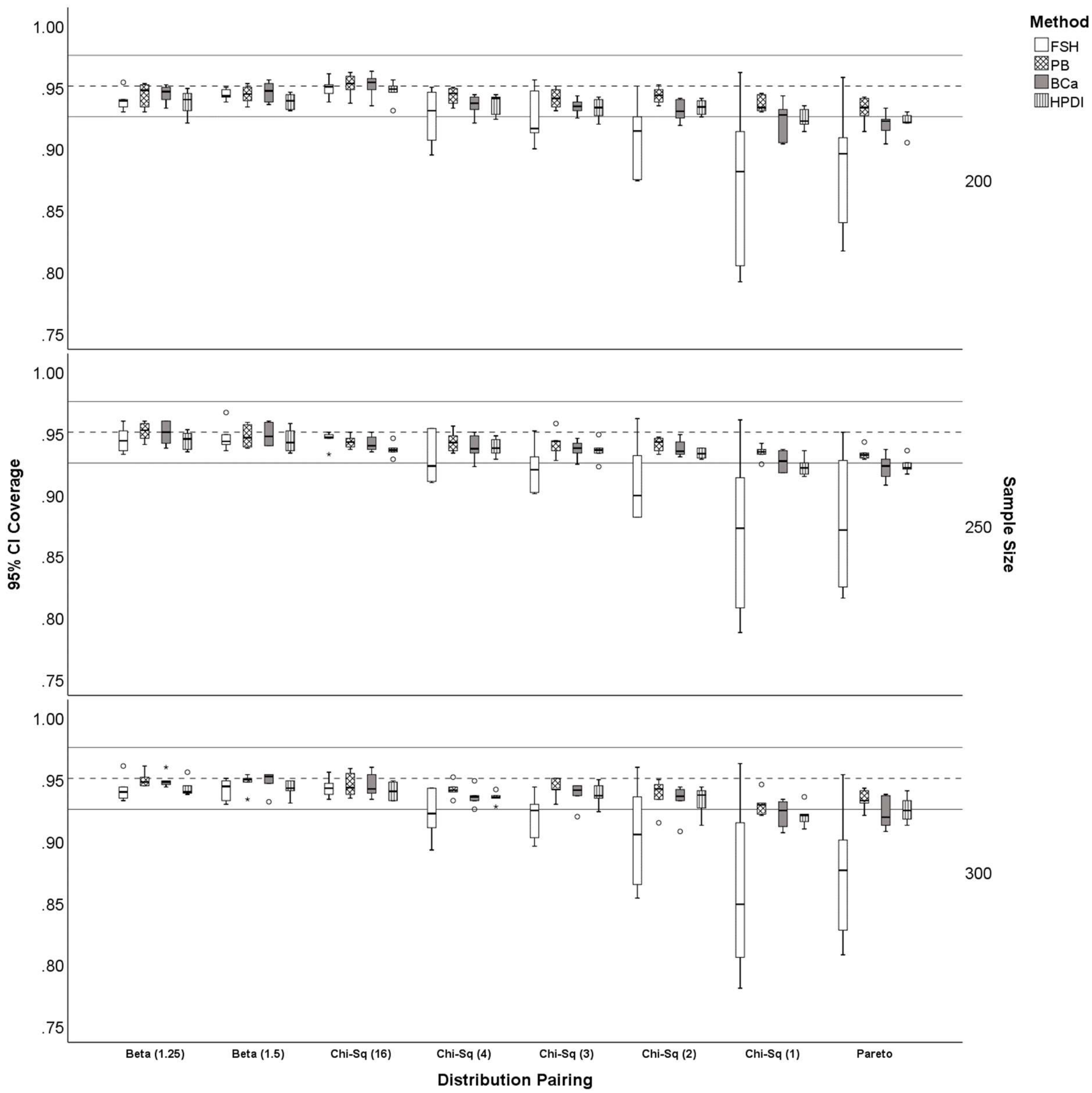

Figure 26. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings by sample size of $200-300$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

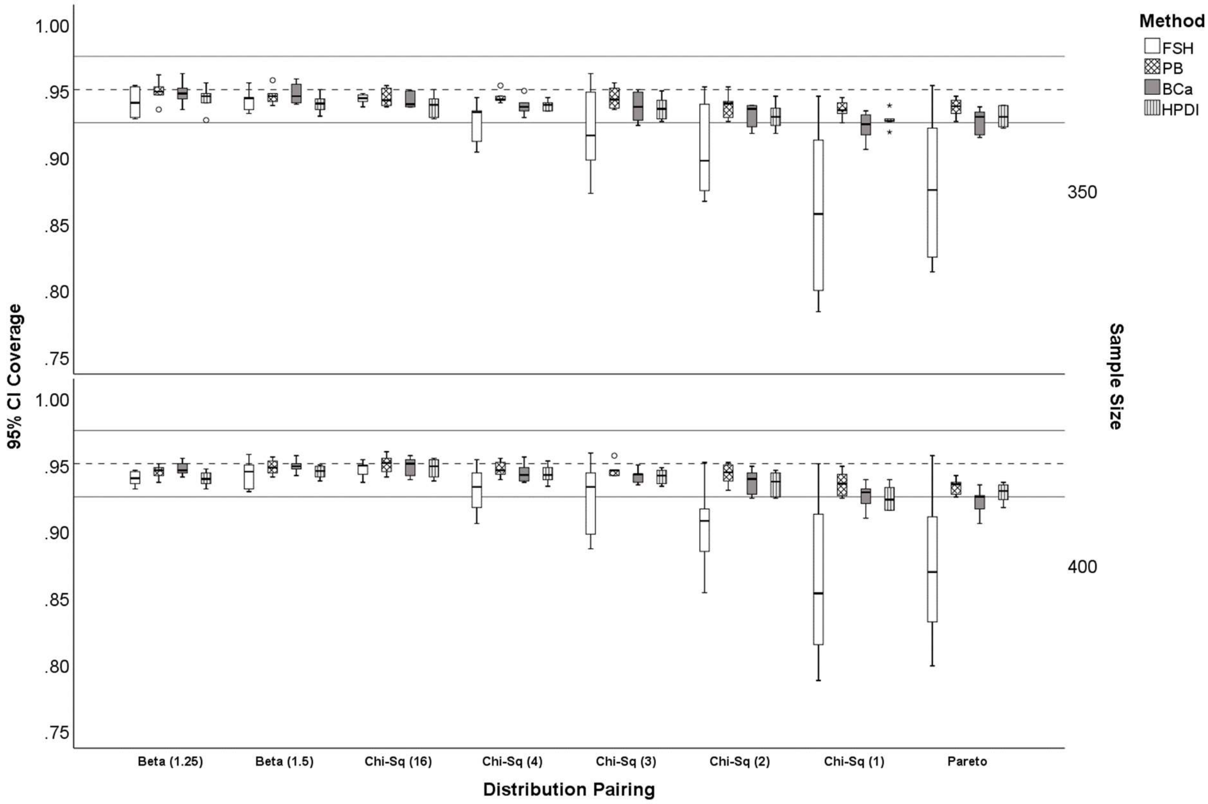

Figure 27. Distribution of 95% CI coverage for non-symmetric with non-symmetric distribution pairings by sample size of $350-400$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

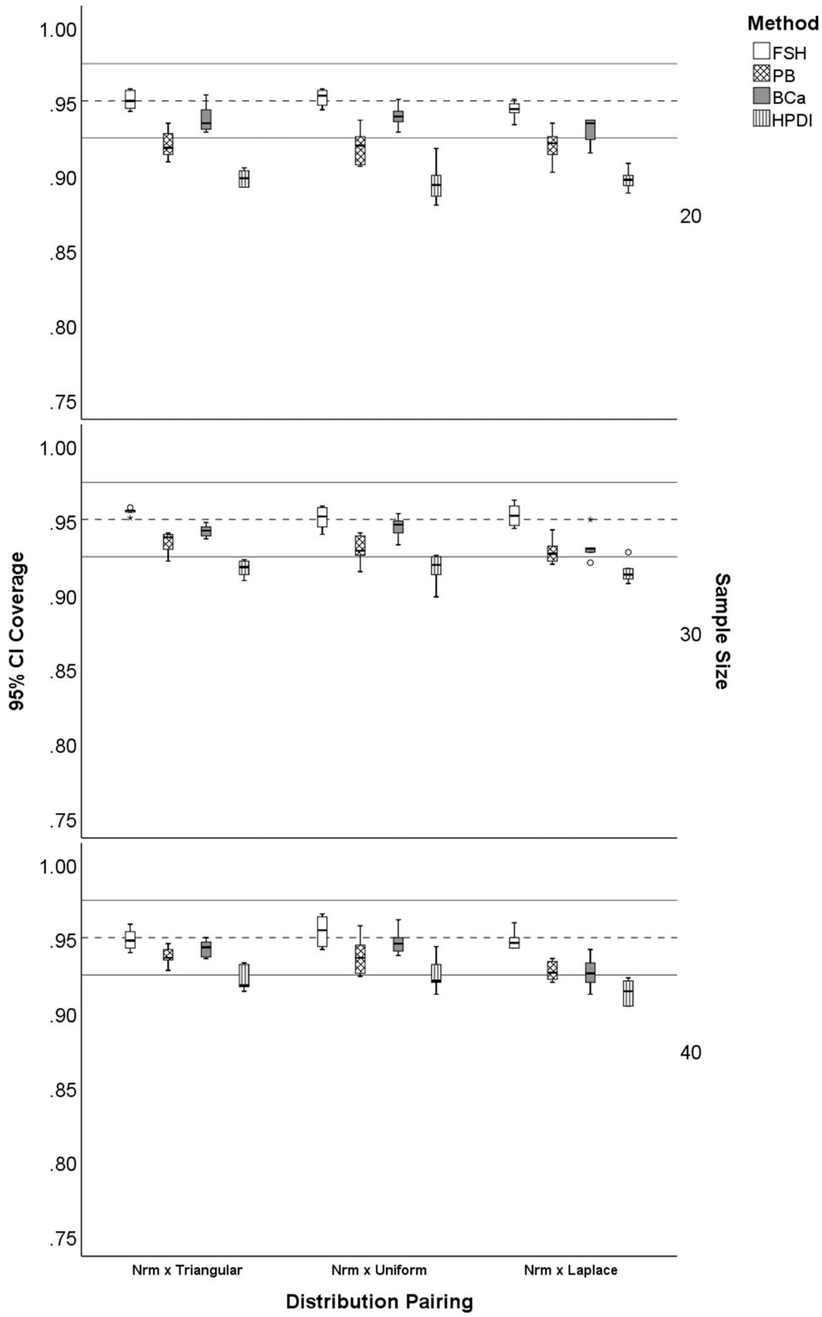

Figure 28. Distribution of 95% CI coverage for symmetric with normal distribution pairings by sample size of $20-40$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

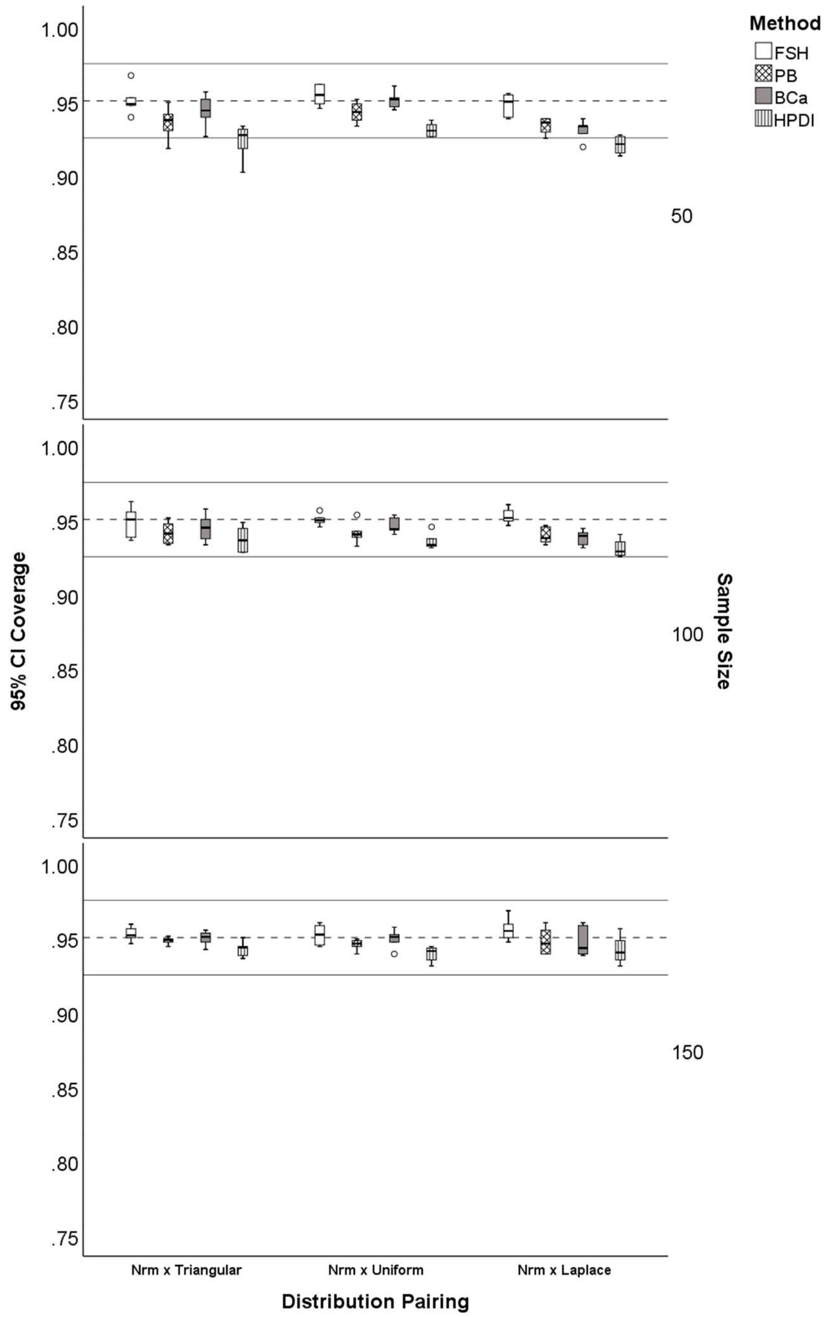

Figure 29. Distribution of 95% CI coverage for symmetric with normal distribution pairings by sample size of $50-150$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

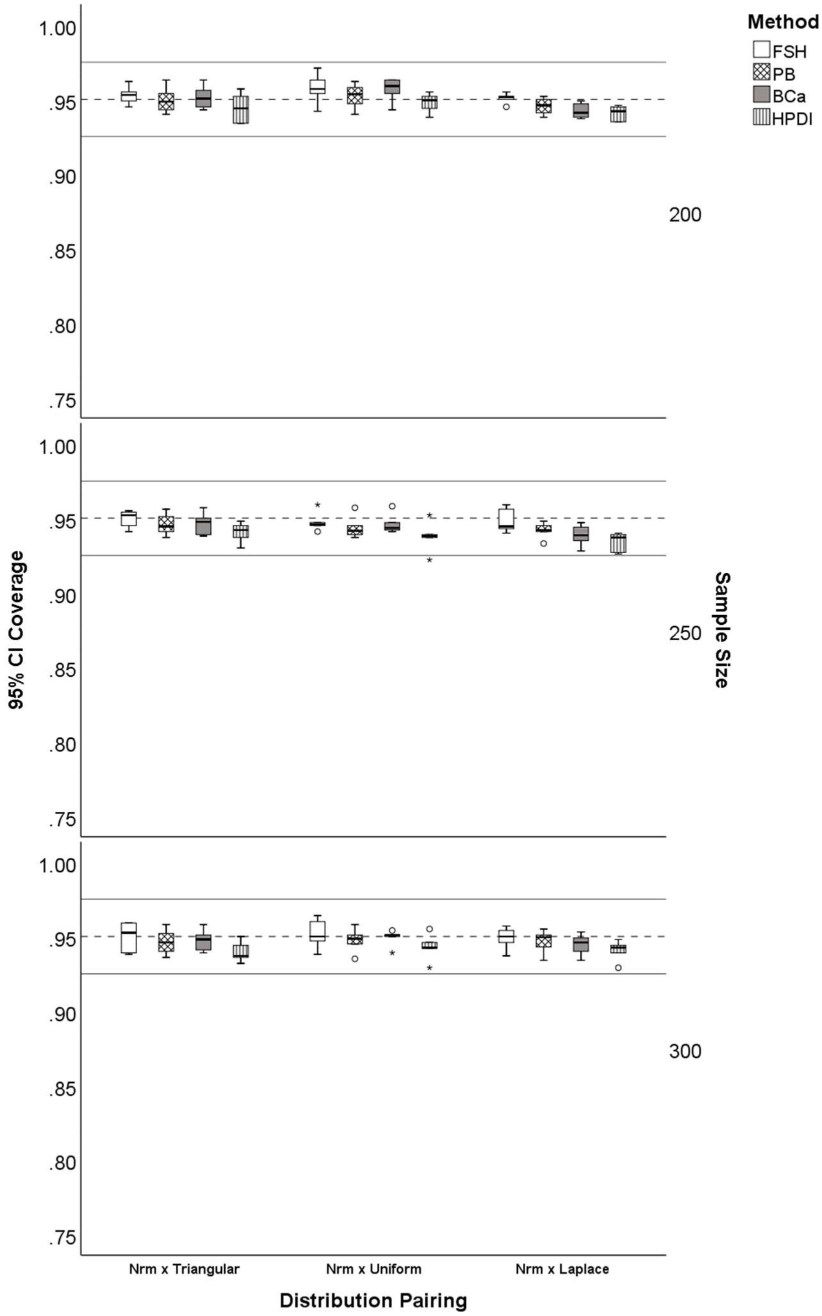

Figure 30. Distribution of 95% CI coverage for symmetric with normal distribution pairings by sample size of $200-300$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

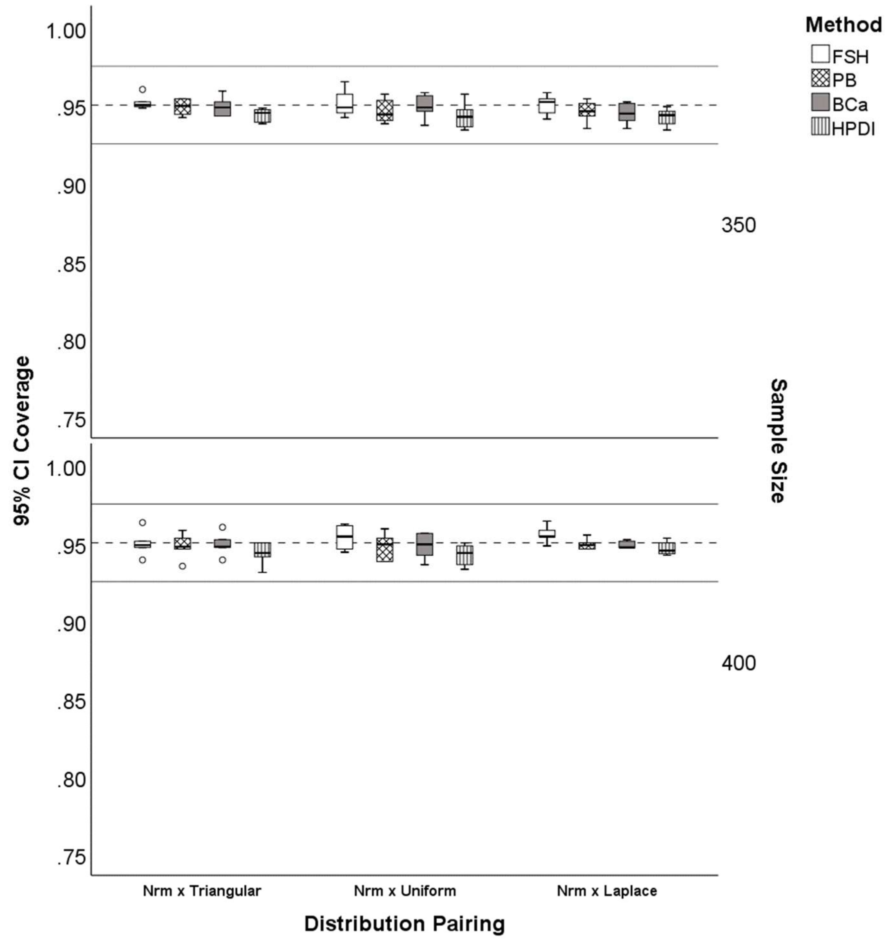

Figure 31. Distribution of 95% CI coverage for symmetric with normal distribution pairings by sample size of $350-400$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

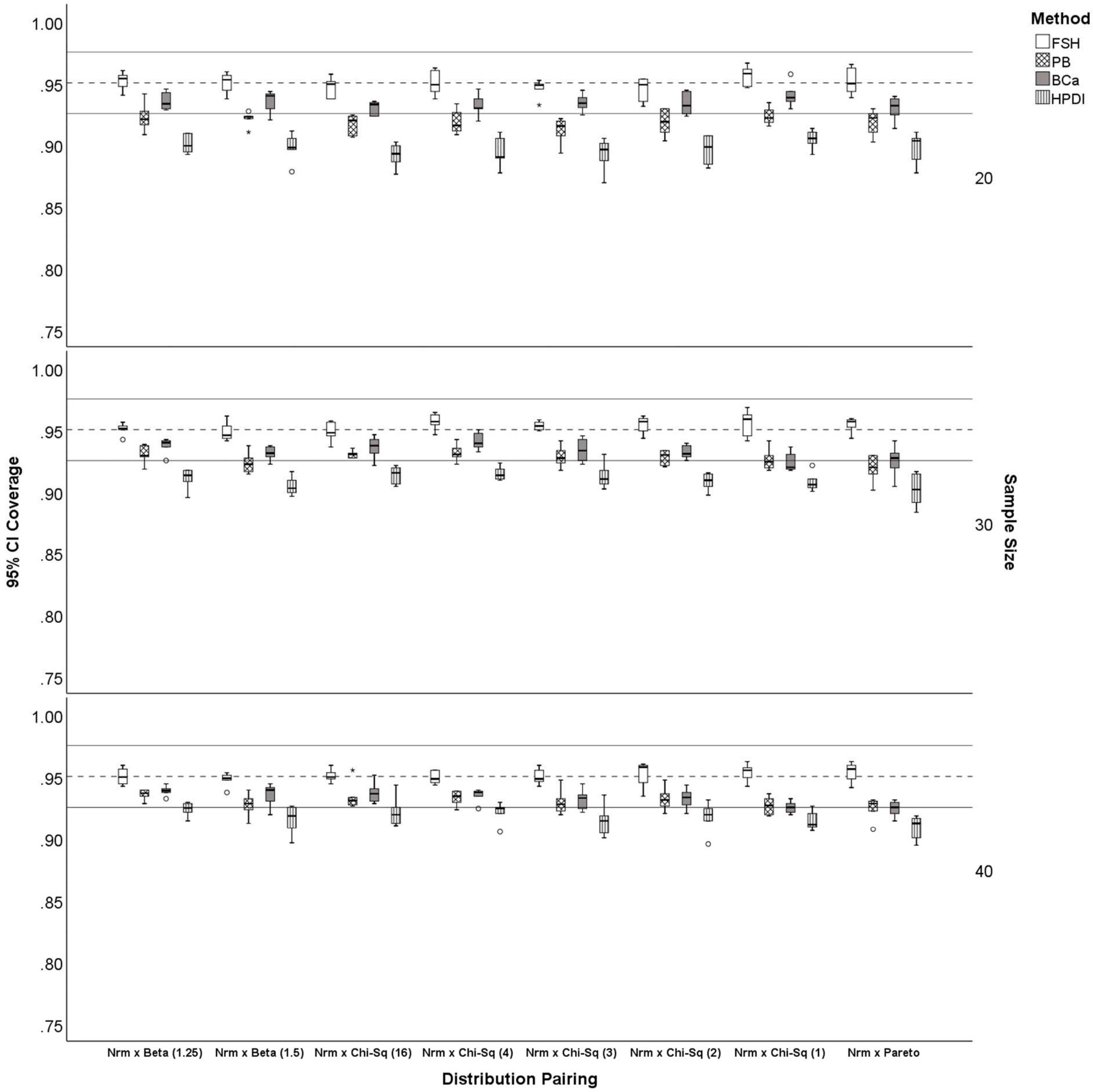

Figure 32. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings by sample size of $20-40$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

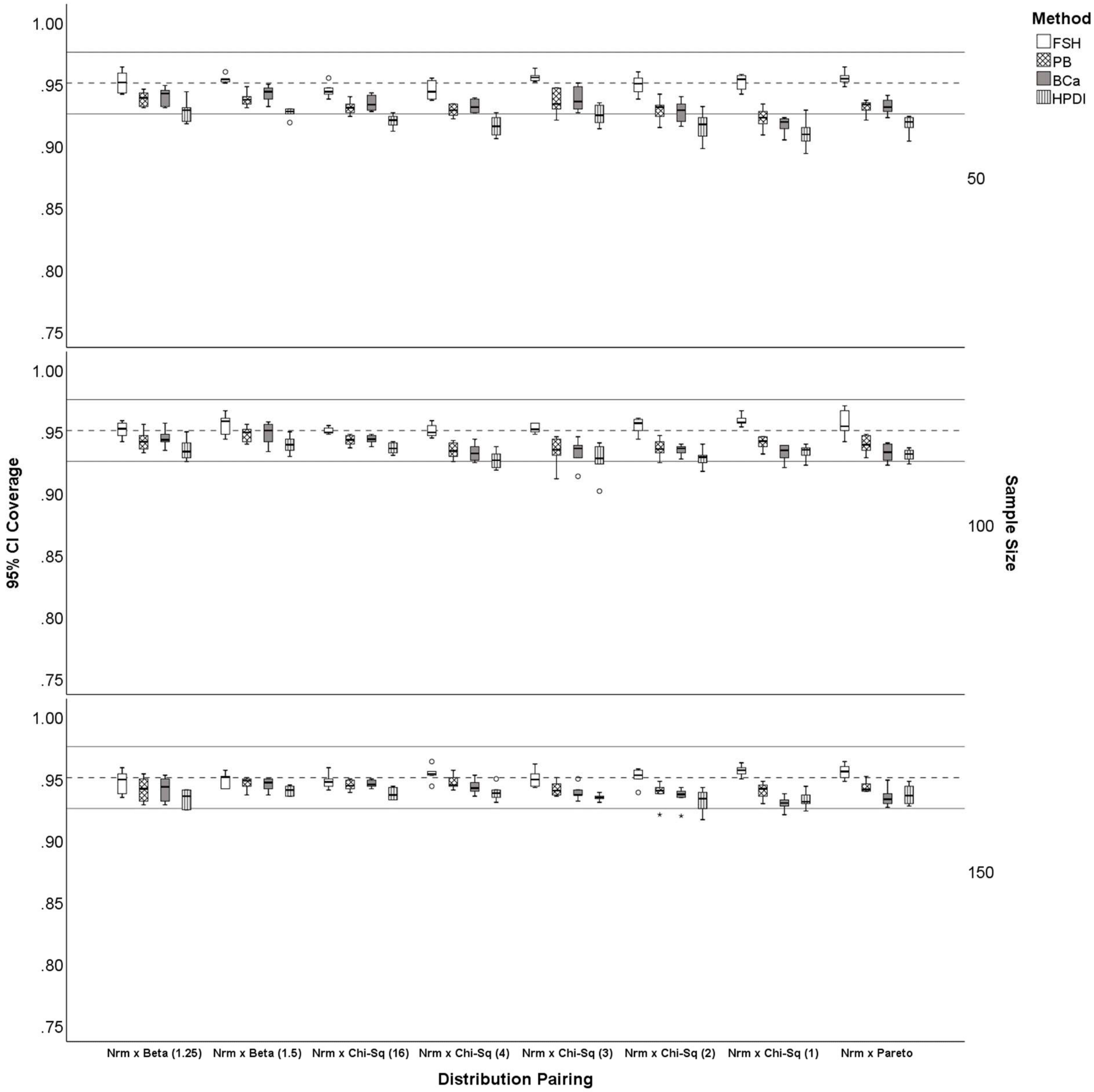

Figure 33. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings by sample size of $50-150$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

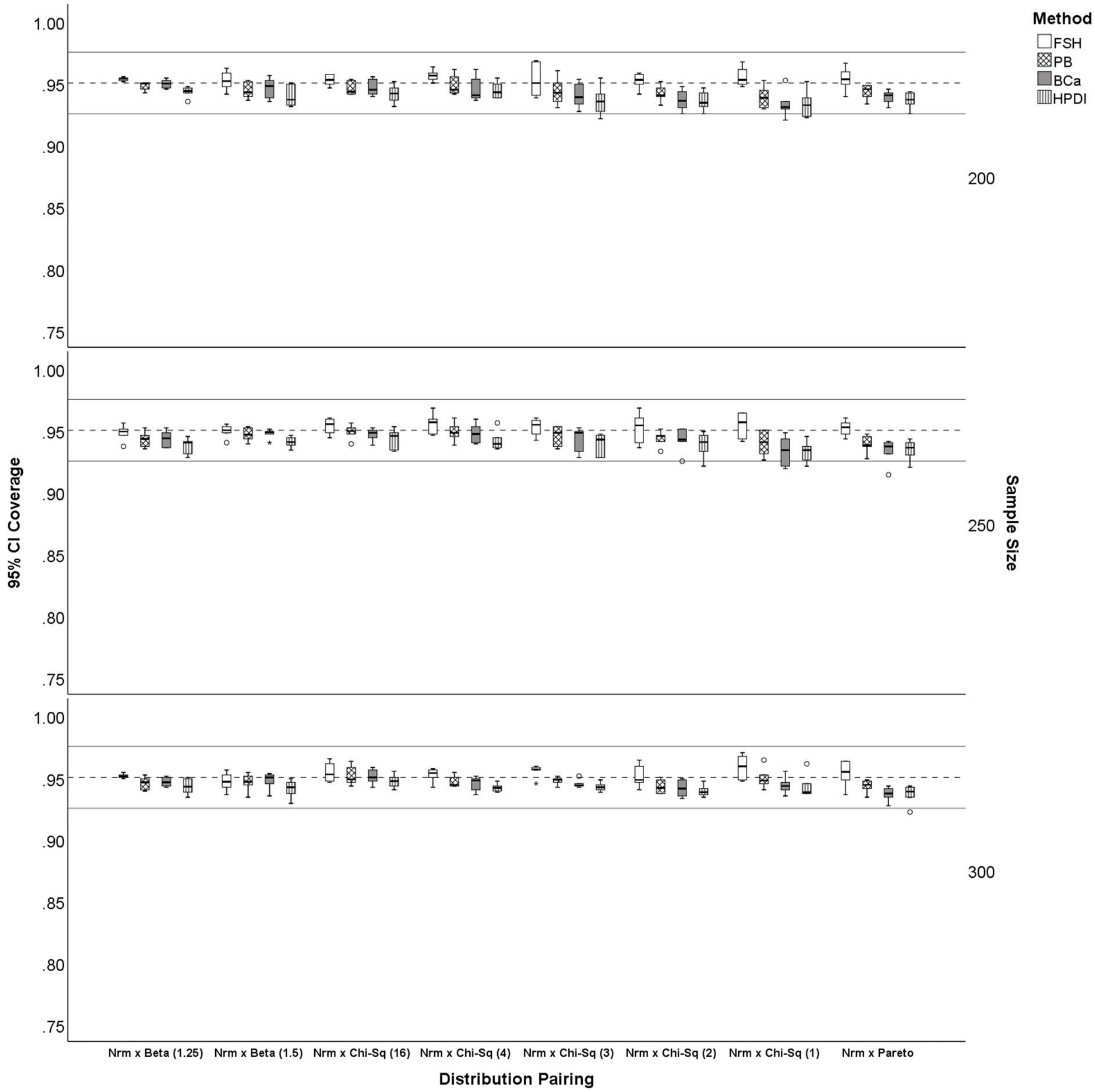

Figure 34. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings by sample size of $200-300$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

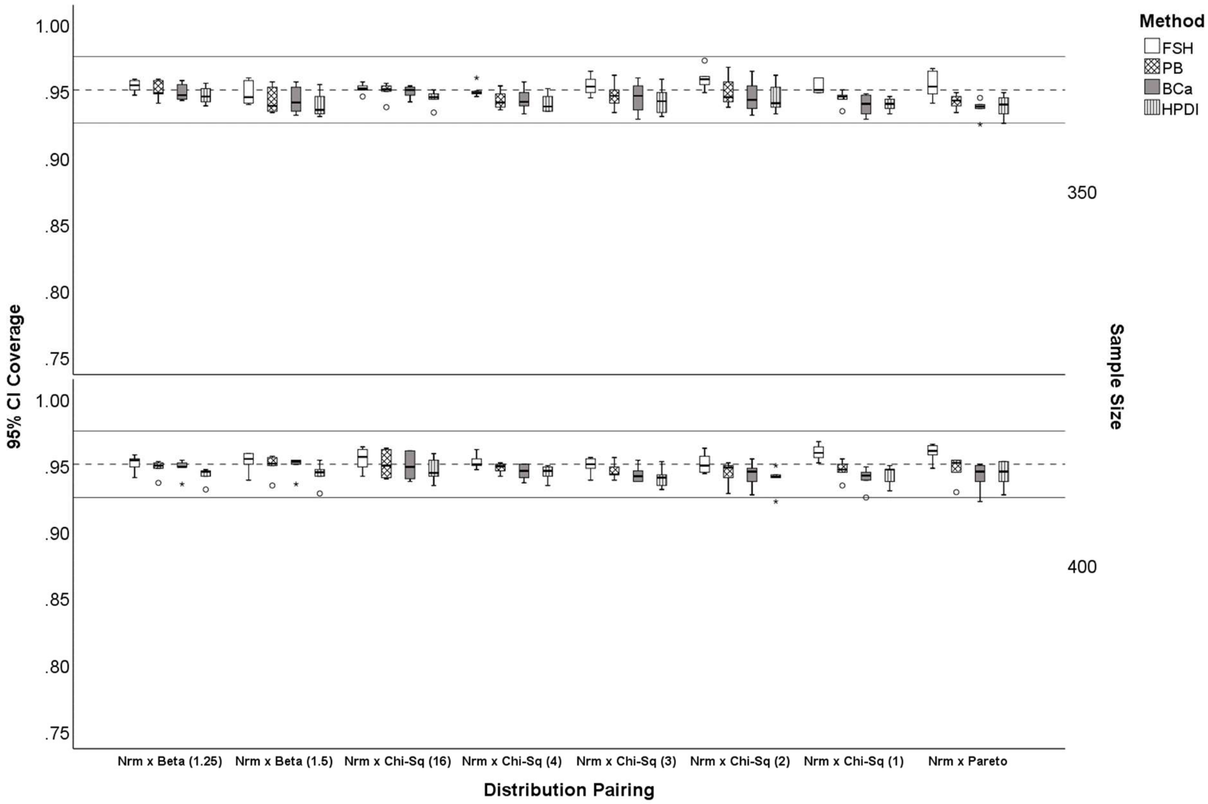

Figure 35. Distribution of 95% CI coverage for non-symmetric with normal distribution pairings by sample size of $350-400$. Fisher *z*-transformation (FSH), percentile bootstrap (PB), bias-corrected and accelerated bootstrap (BCa), and highest probability density interval (HPDI). Bootstrap methods (PB, BCa, HPDI) were based on 2,000 bootstrap samples. The dashed line is at .95 and the solid lines are at $[.925, .975]$; acceptable coverage.

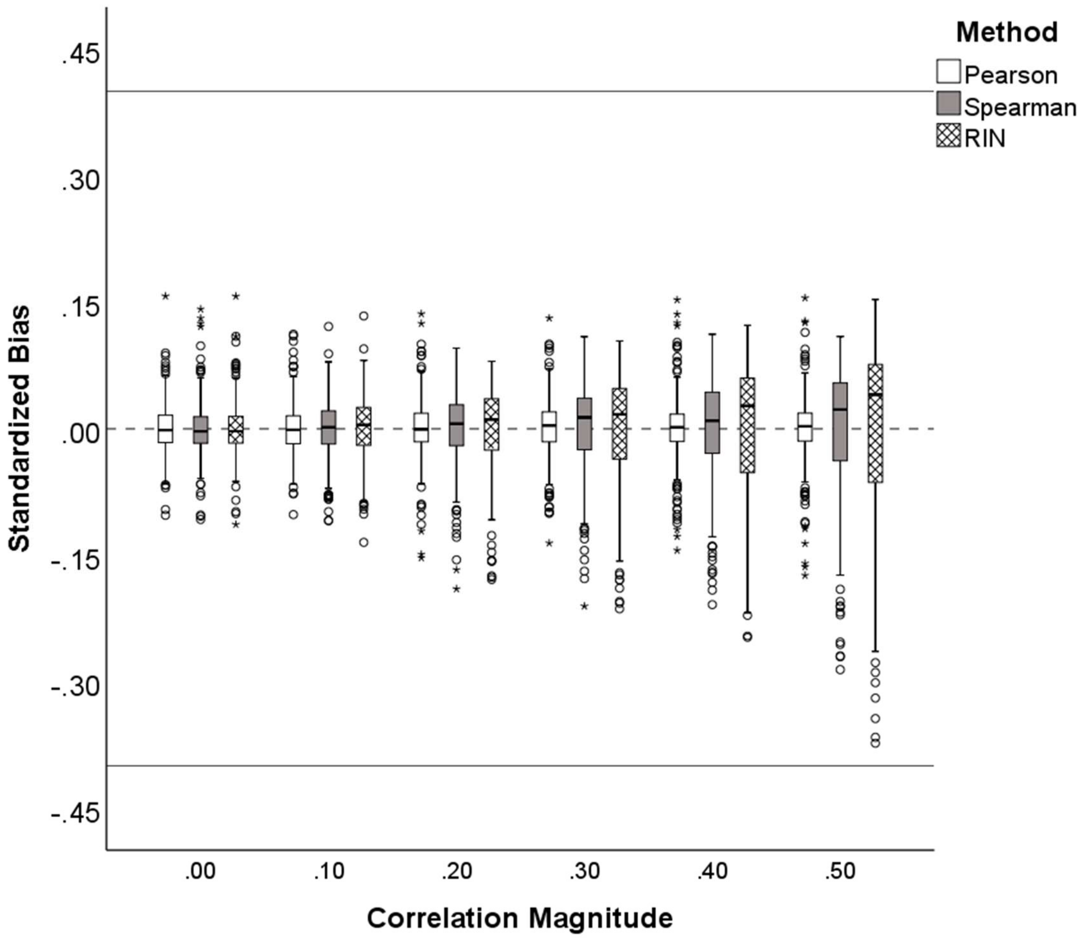

Figure 36. Distribution of standardized bias for correlation magnitude. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

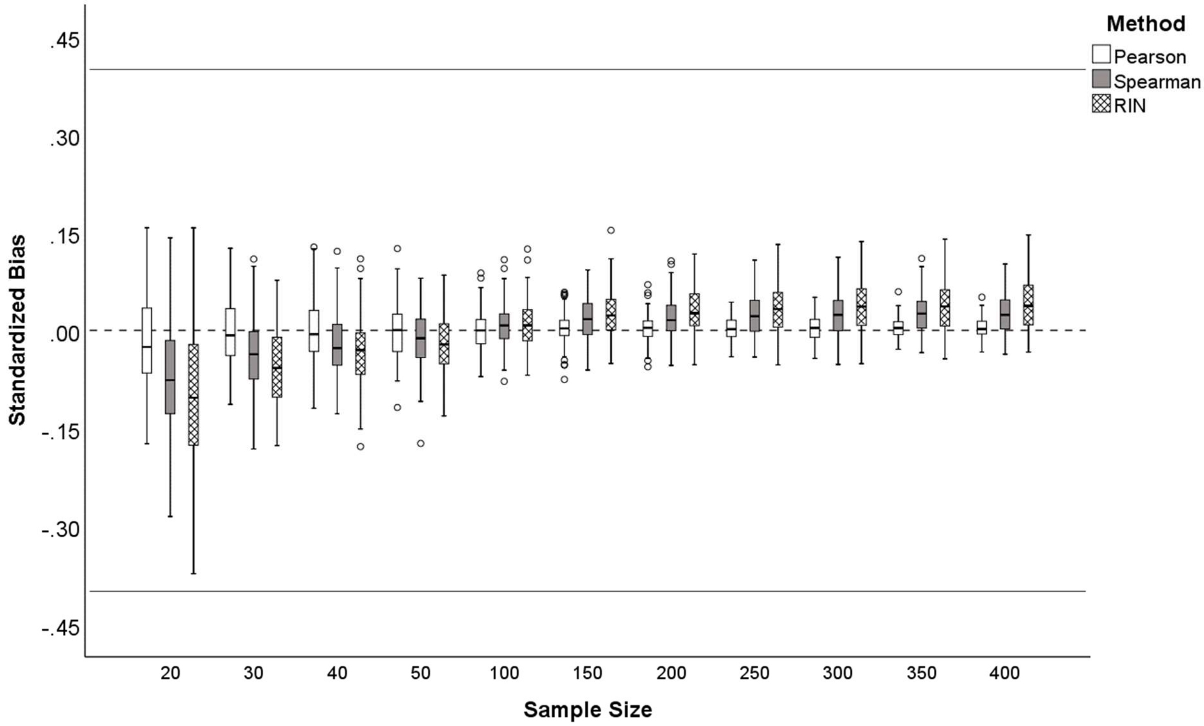

Figure 37. Distribution of standardized bias for sample size. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

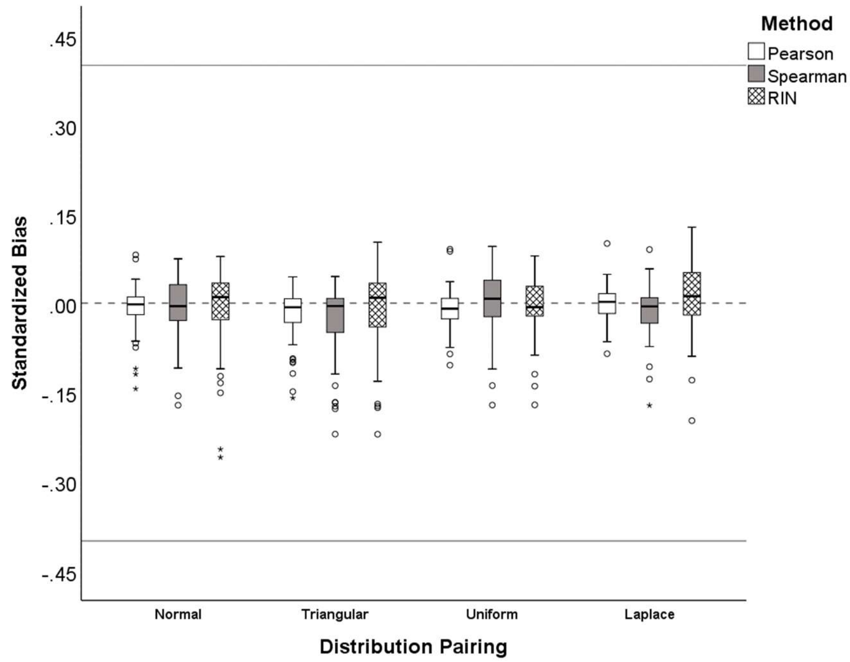

Figure 38. Distribution of standardized bias for symmetric with symmetric distribution pairings. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

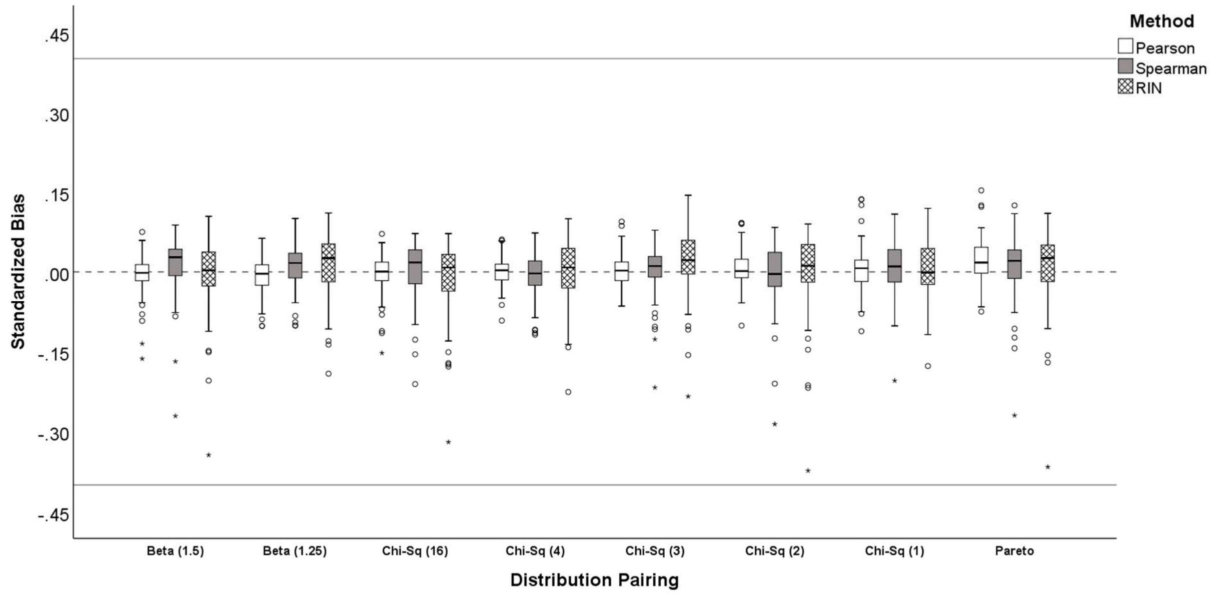

Figure 39. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

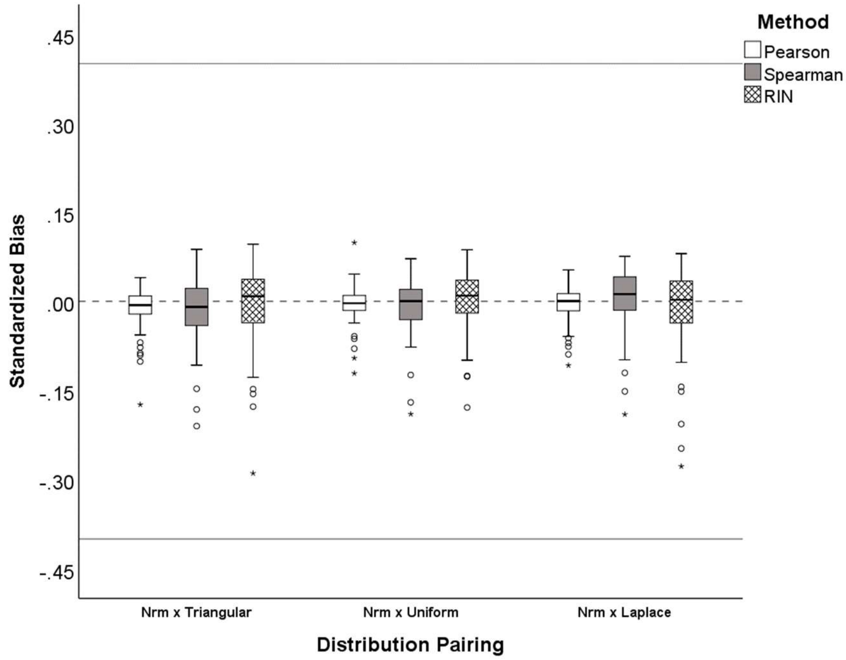

Figure 40. Distribution of standardized bias for symmetric with normal distribution pairings. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

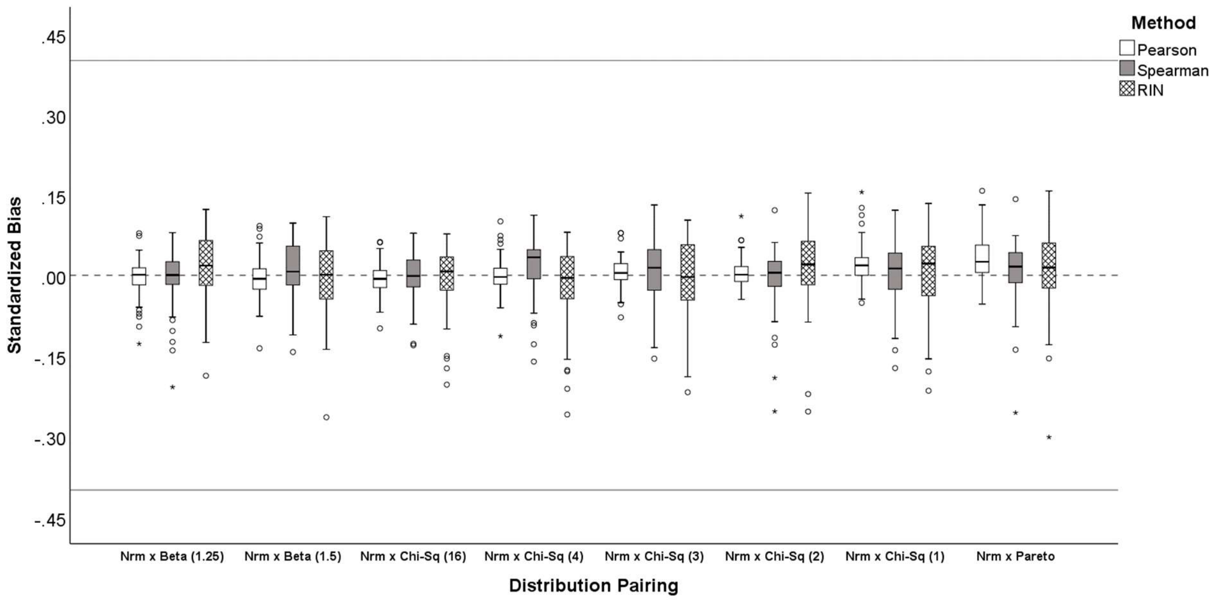

Figure 41. Distribution of standardized bias for non-symmetric with normal distribution pairings. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

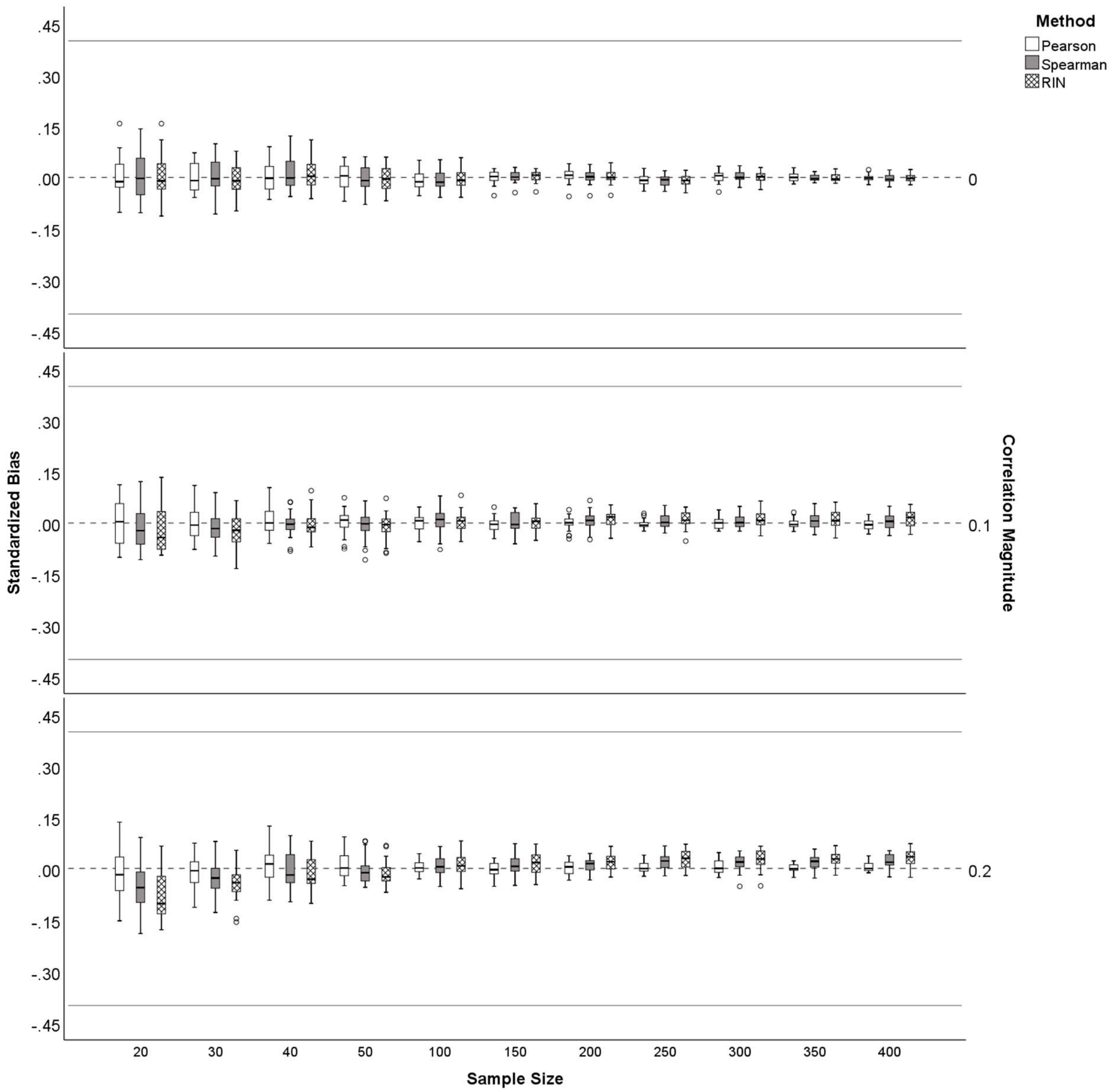

Figure 42. Distribution of standardized bias for sample size by correlation magnitude of $0-.2$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

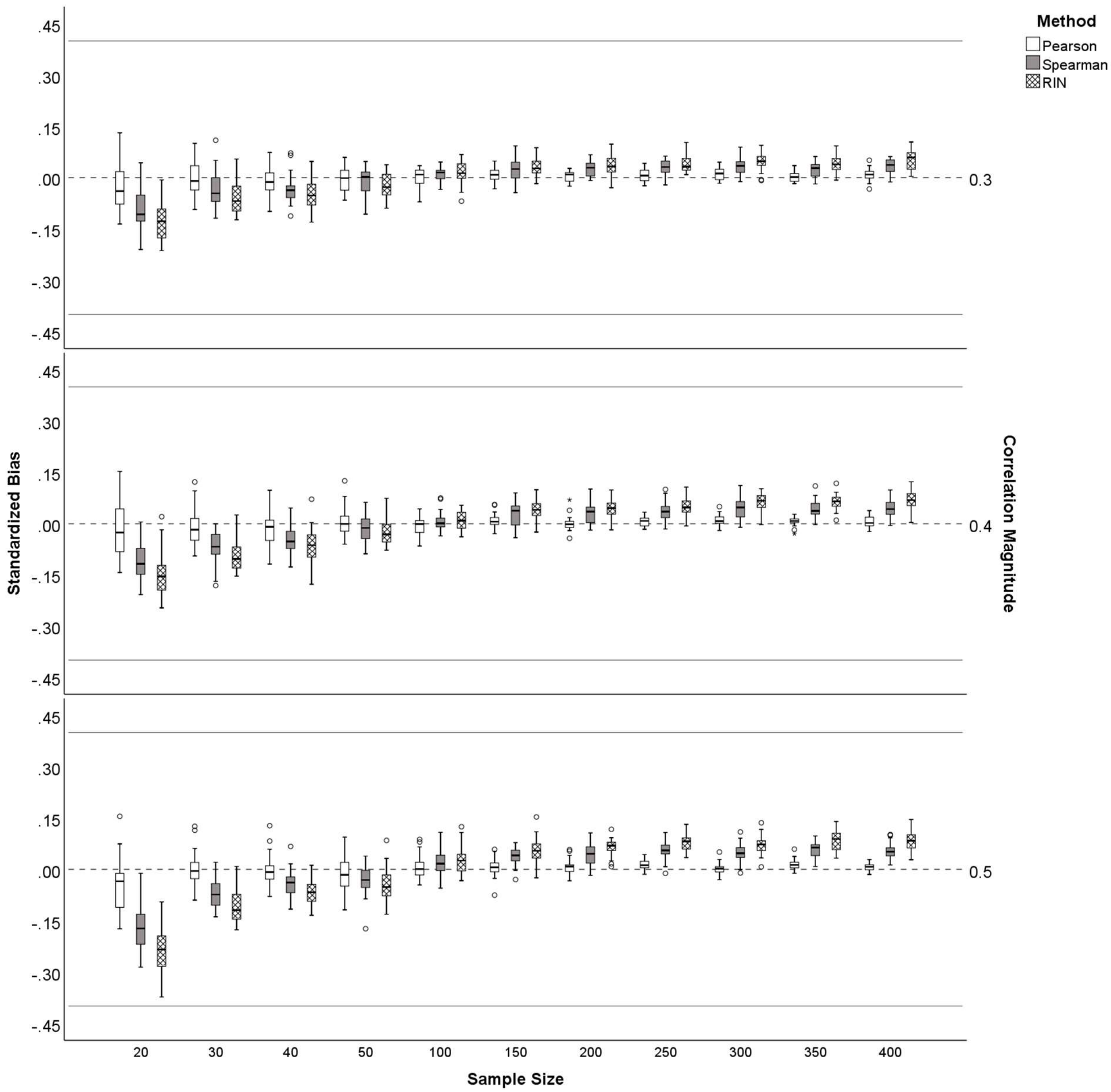

Figure 43. Distribution of standardized bias for sample size by correlation magnitude of $.3-.5$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

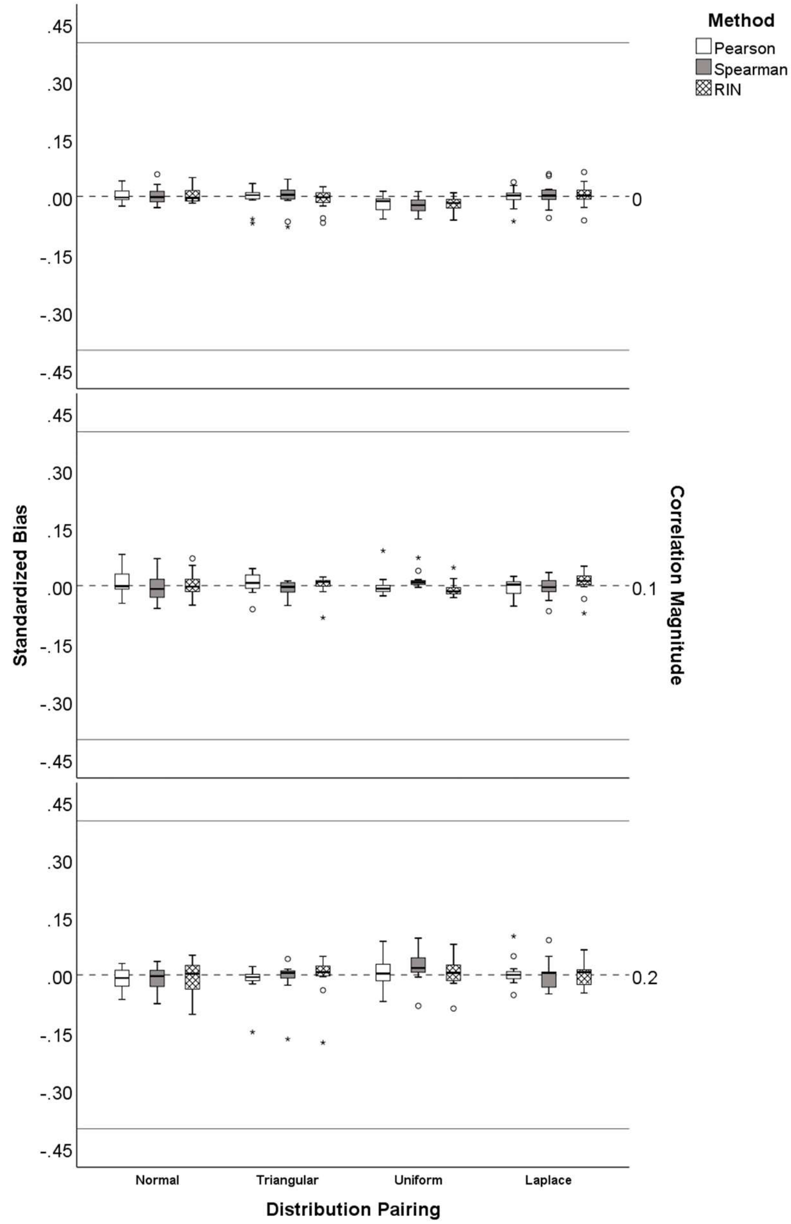

Figure 44. Distribution of standardized bias for symmetric with symmetric distribution pairings by correlation magnitude of $0-.2$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

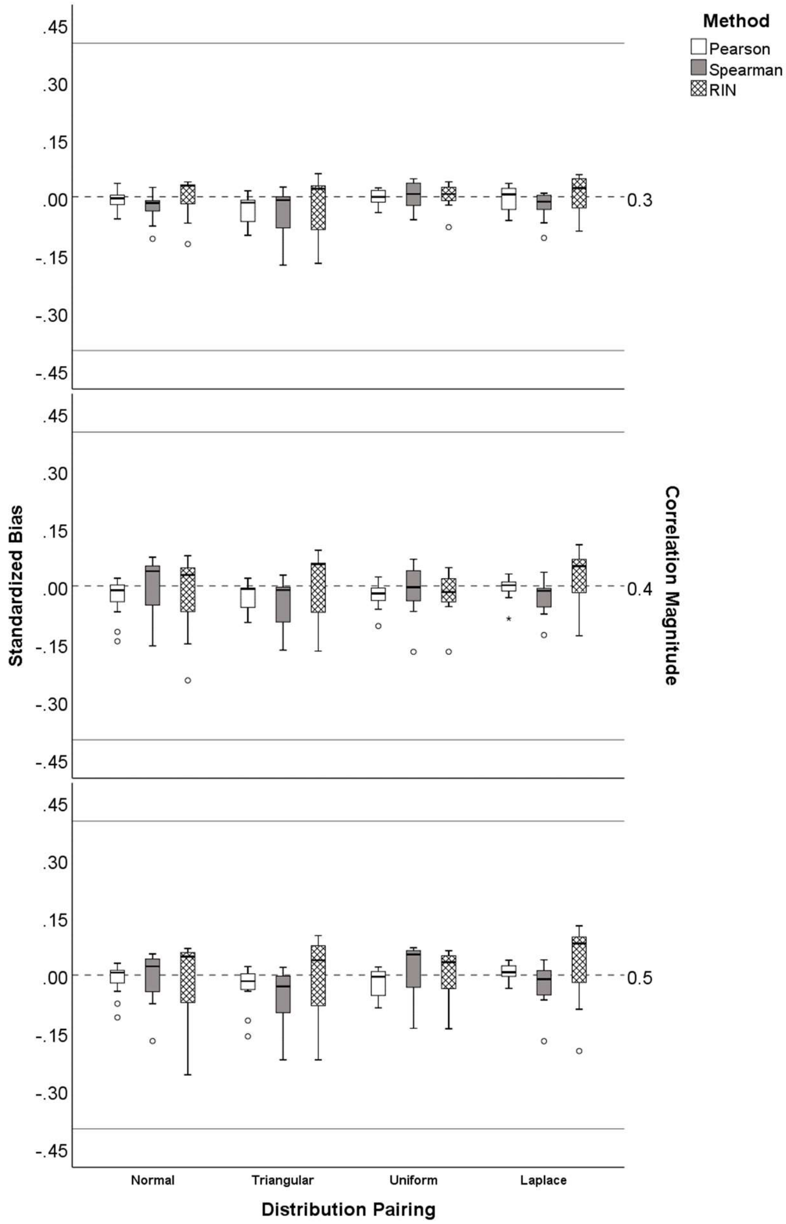

Figure 45. Distribution of standardized bias for symmetric with symmetric distribution pairings by correlation magnitude of $.3-.5$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

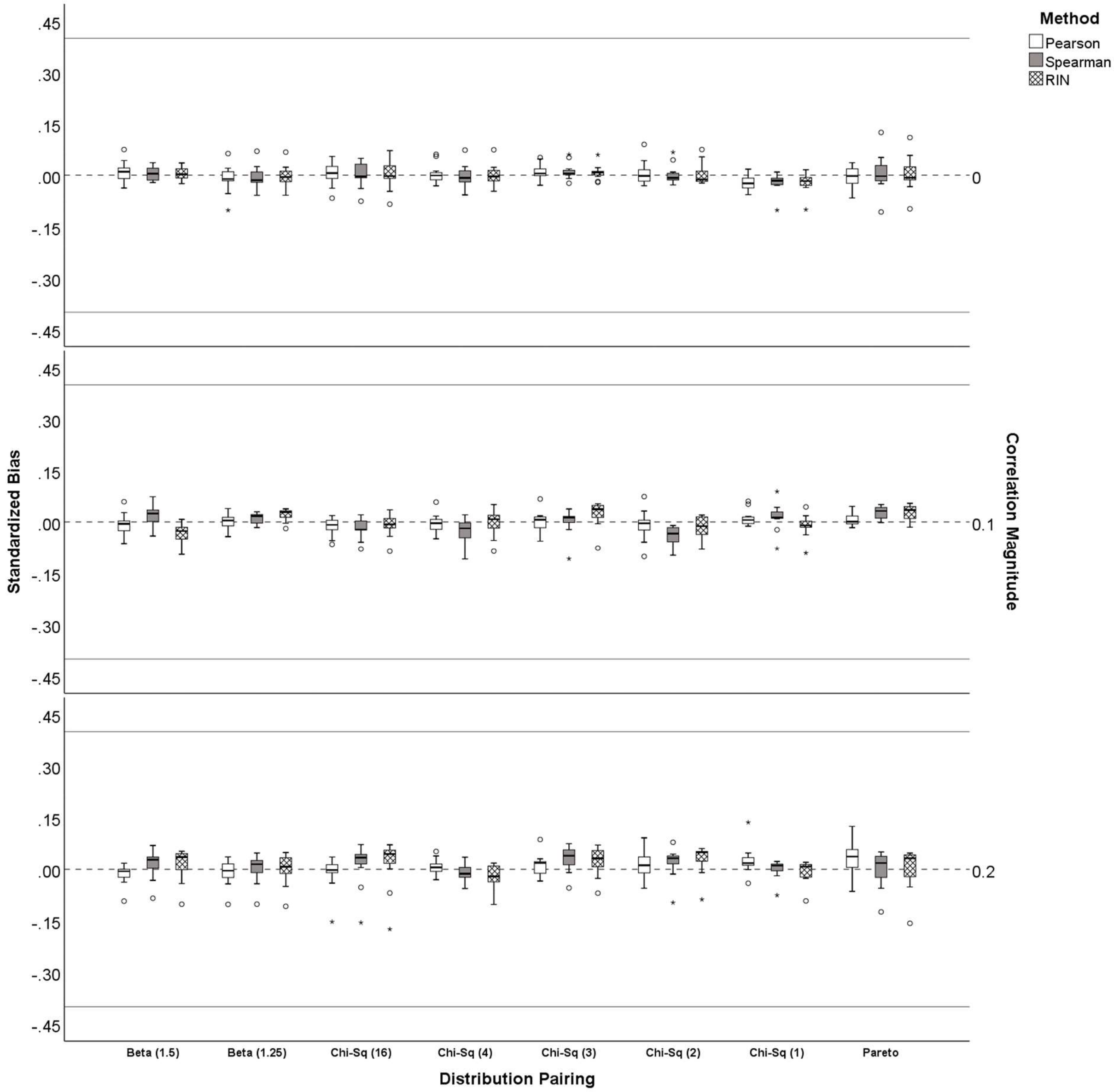

Figure 46. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings by correlation magnitude of $0-.2$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

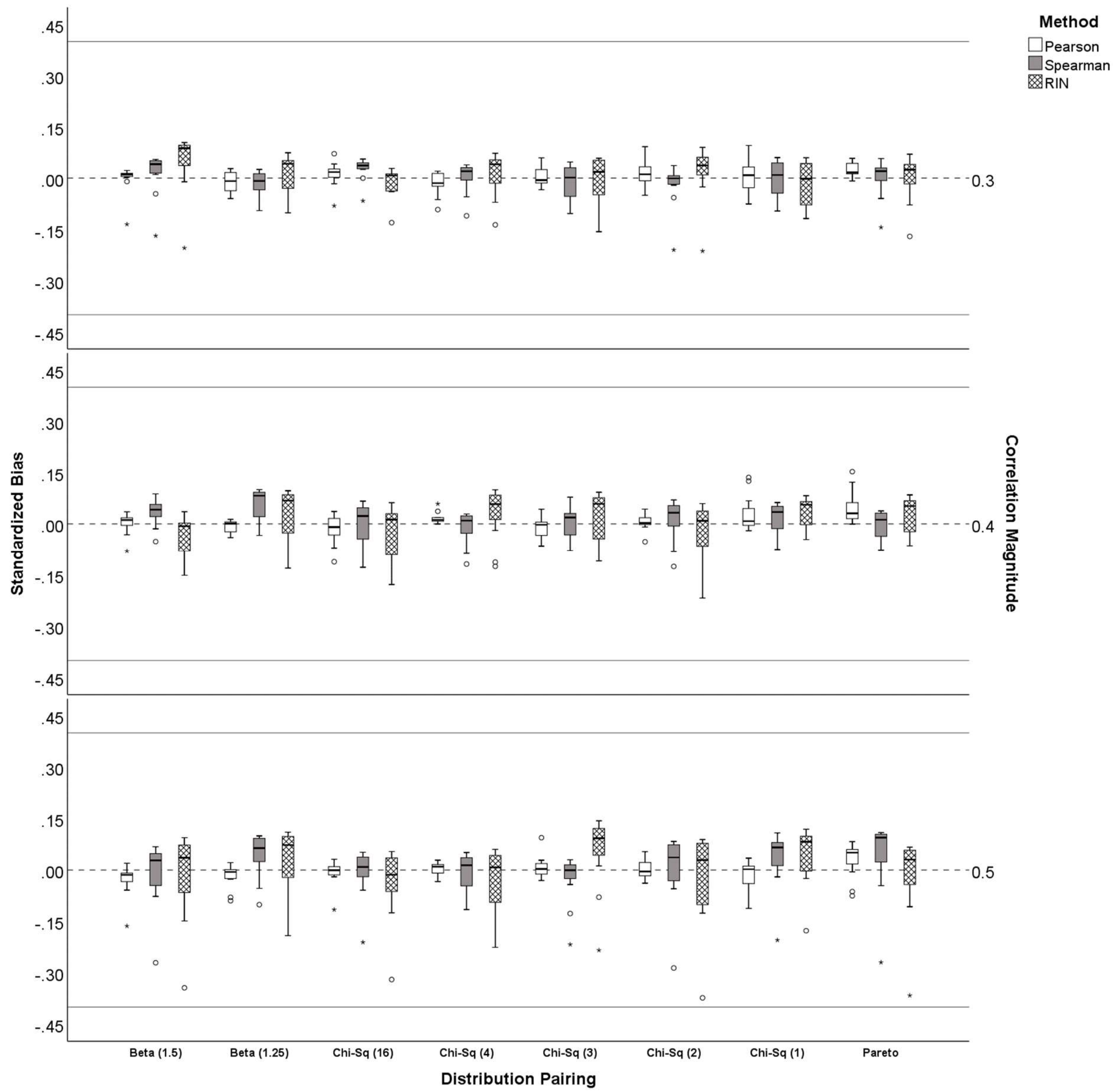

Figure 47. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings by correlation magnitude of $.3-.5$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

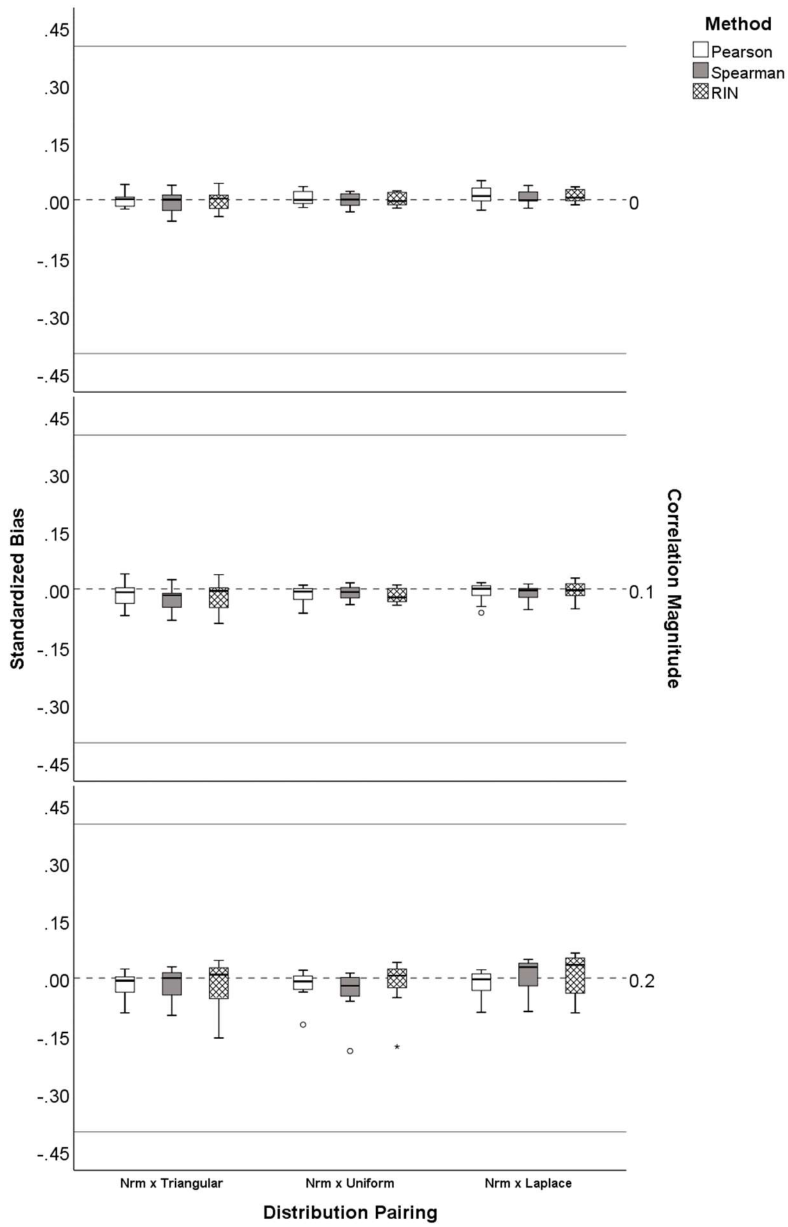

Figure 48. Distribution of standardized bias for symmetric with normal distribution pairings by correlation magnitude of $0-.2$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

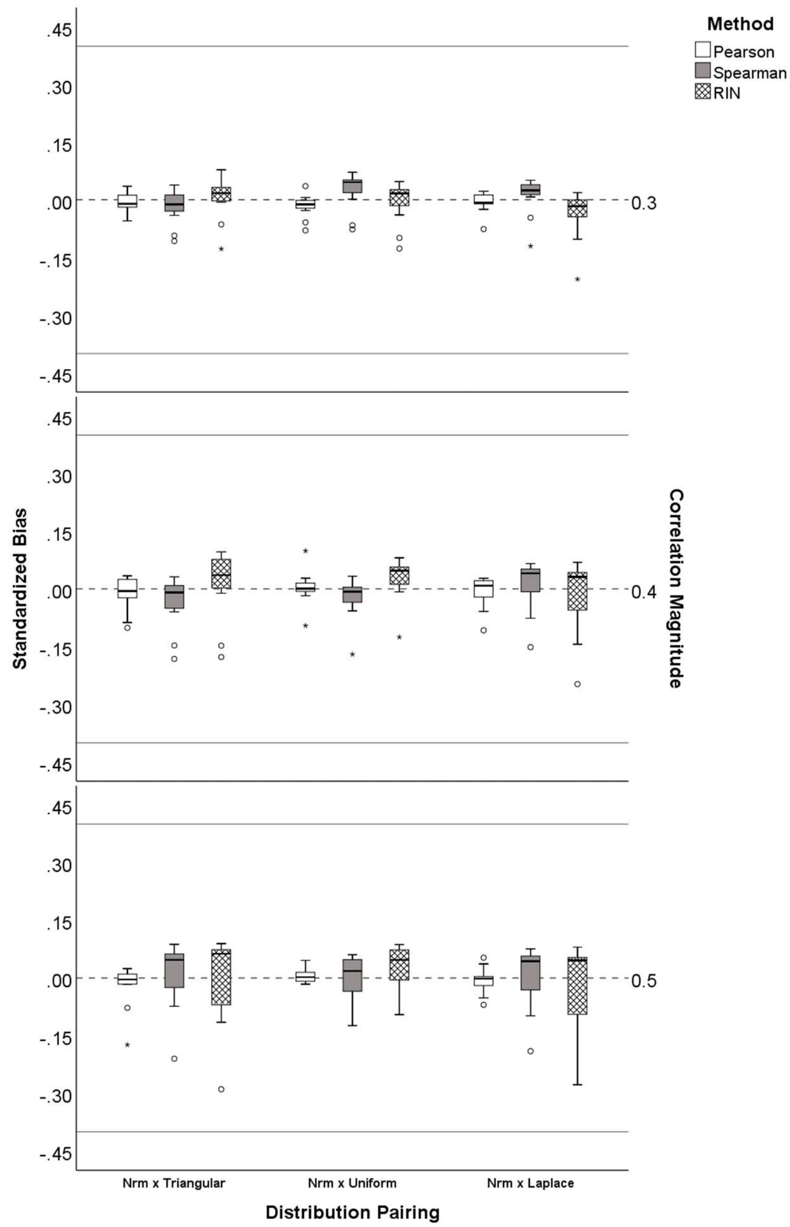

Figure 49. Distribution of standardized bias for symmetric with normal distribution pairings by correlation magnitude of $.3-.5$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

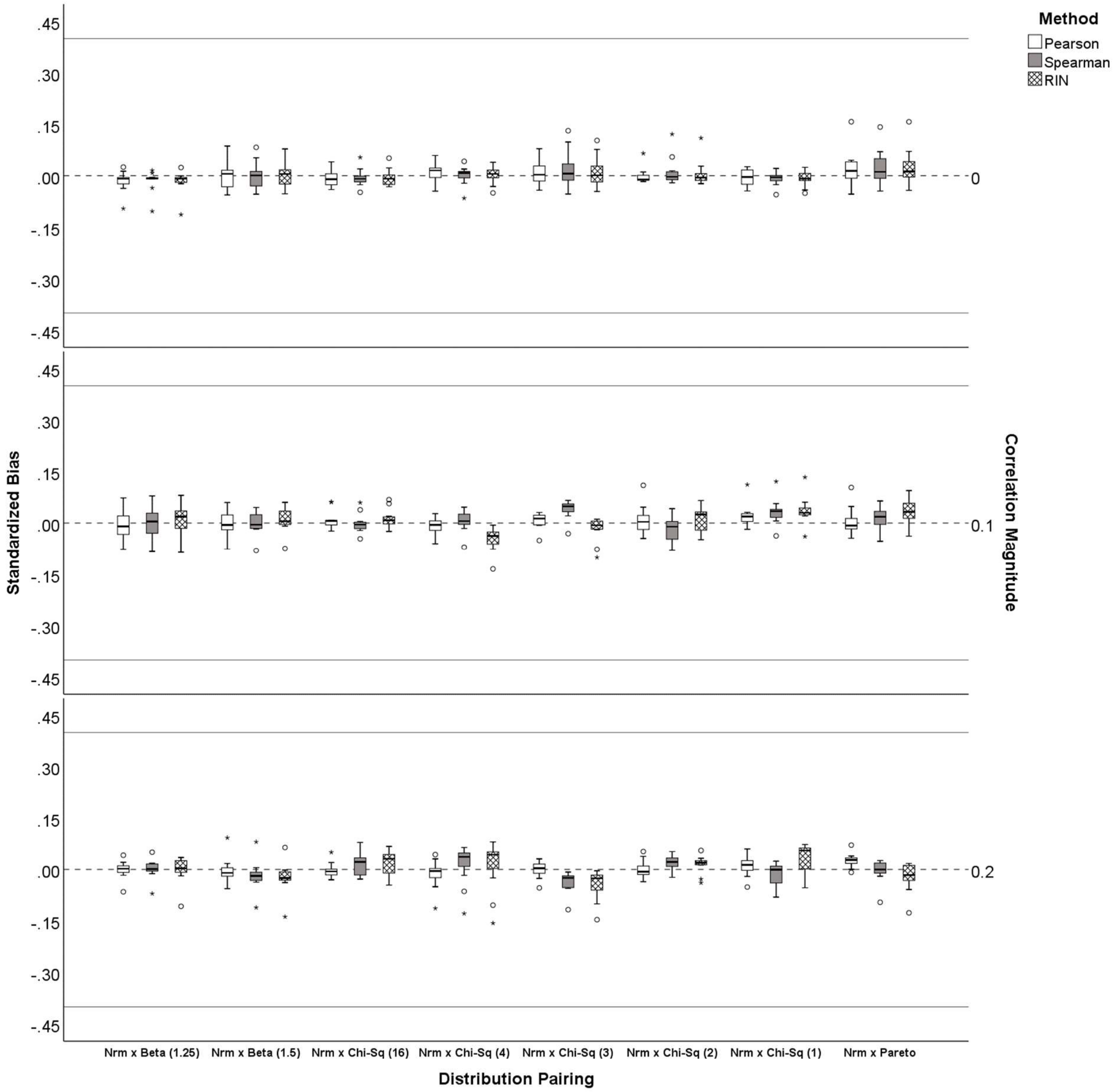

Figure 50. Distribution of standardized bias for non-symmetric with normal distribution pairings by correlation magnitude of $0-.2$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

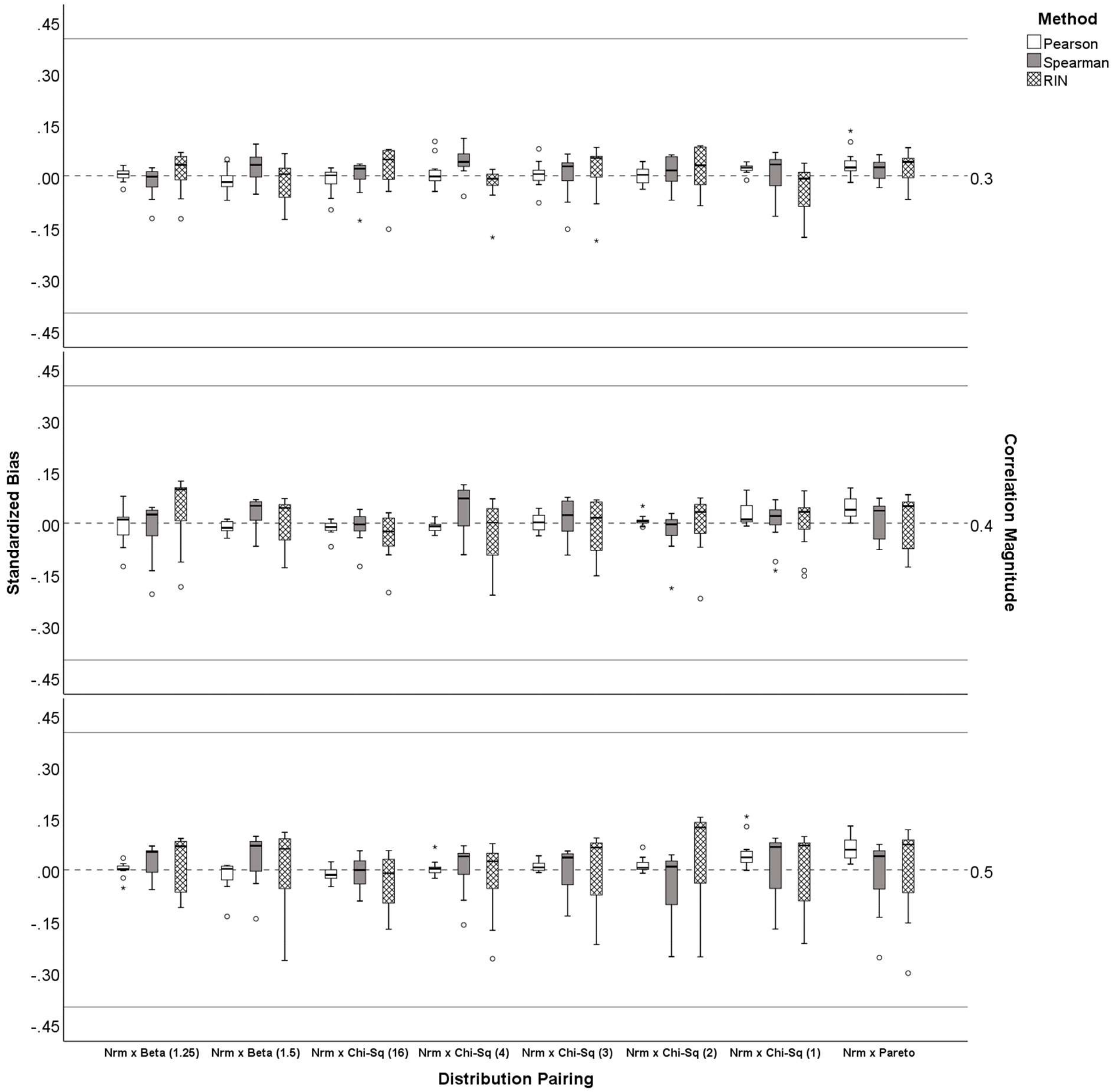

Figure 51. Distribution of standardized bias for non-symmetric with normal distribution pairings by correlation magnitude of $.3-.5$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

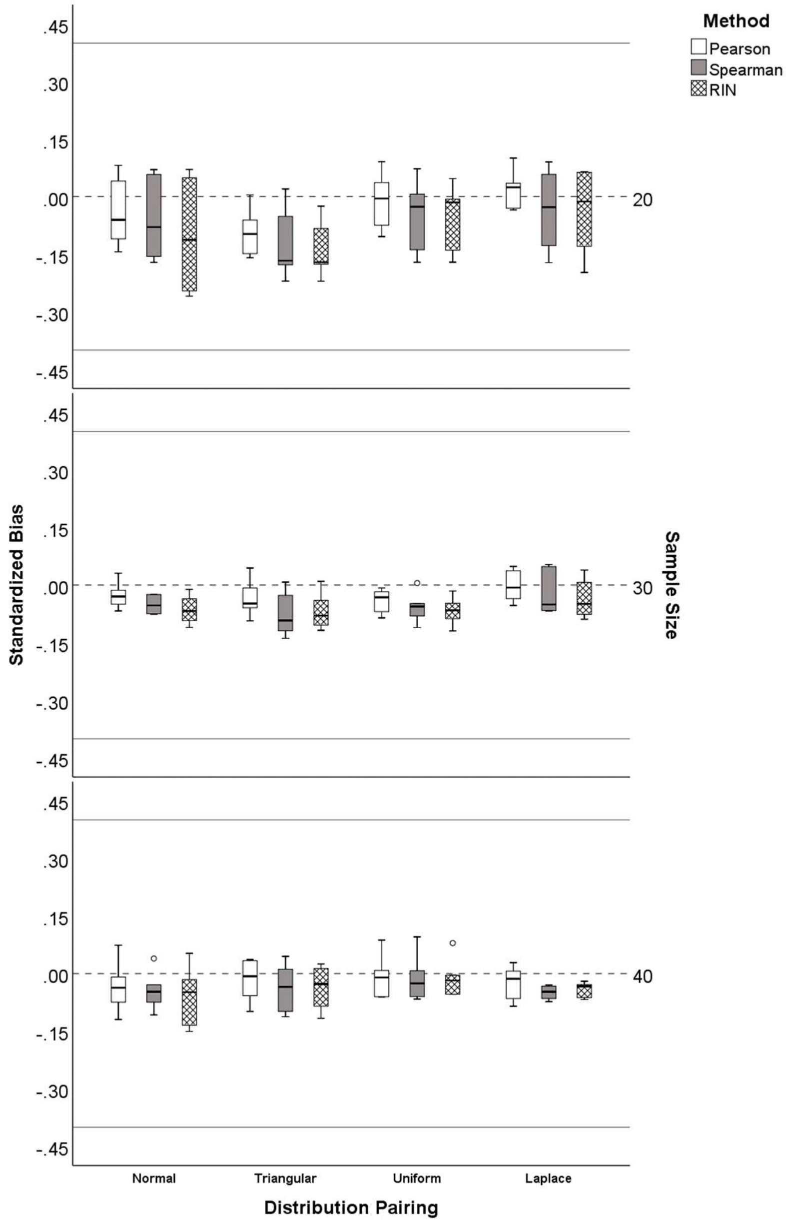

Figure 52. Distribution of standardized bias for symmetric with symmetric distribution pairings by sample size of $20-40$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

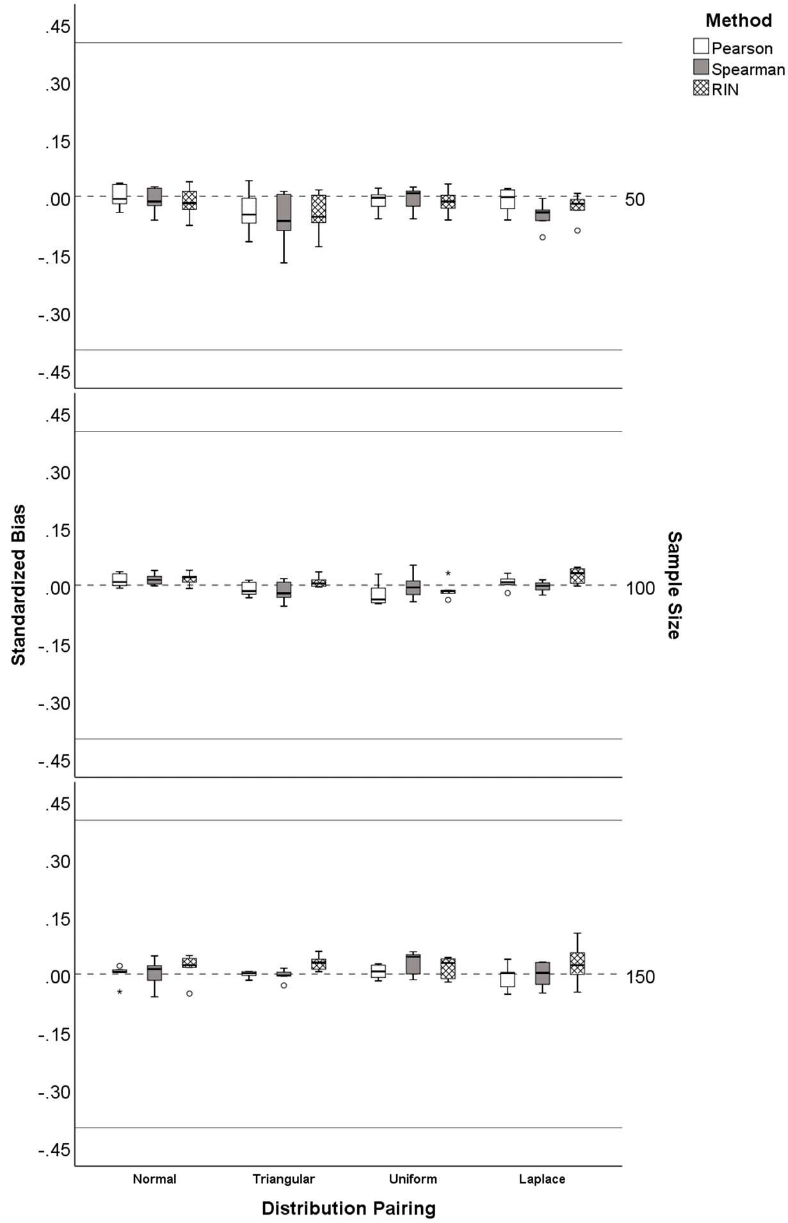

Figure 53. Distribution of standardized bias for symmetric with symmetric distribution pairings by sample size of $50-150$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

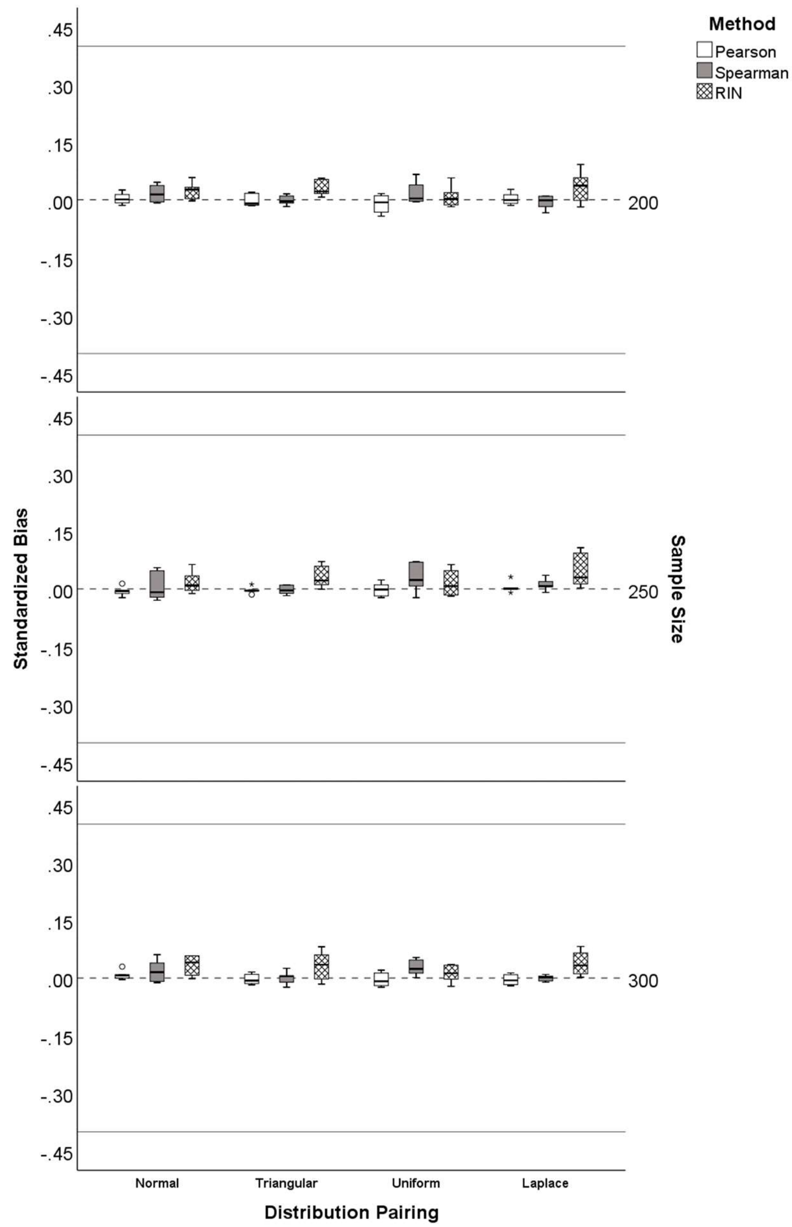

Figure 54. Distribution of standardized bias for symmetric with symmetric distribution pairings by sample size of $200-300$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

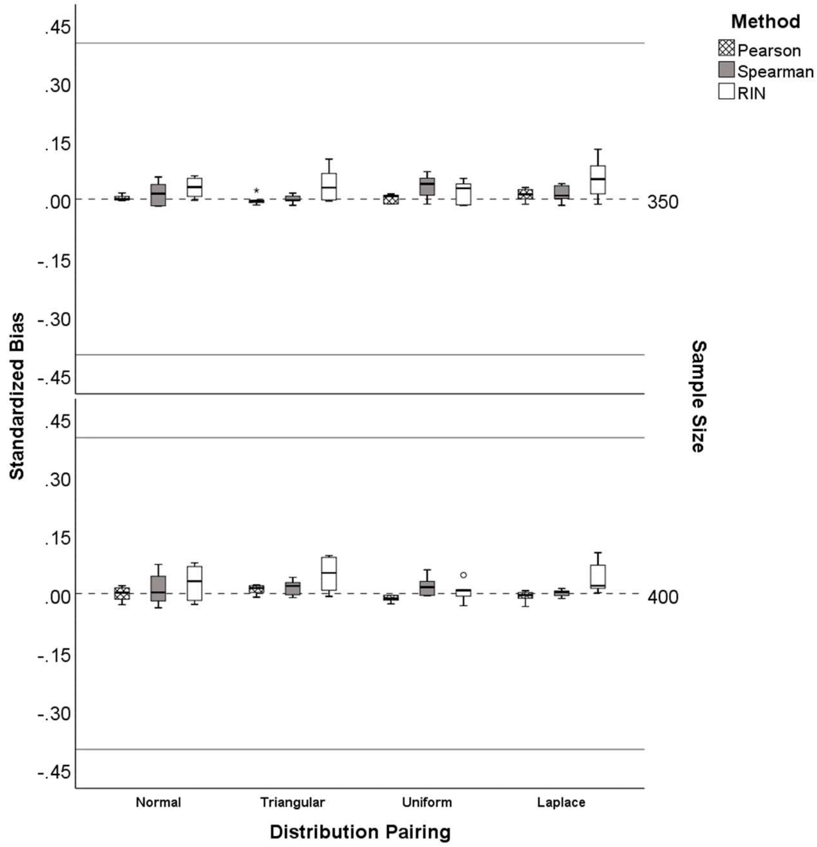

Figure 55. Distribution of standardized bias for symmetric with symmetric distribution pairings by sample size of $350-400$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

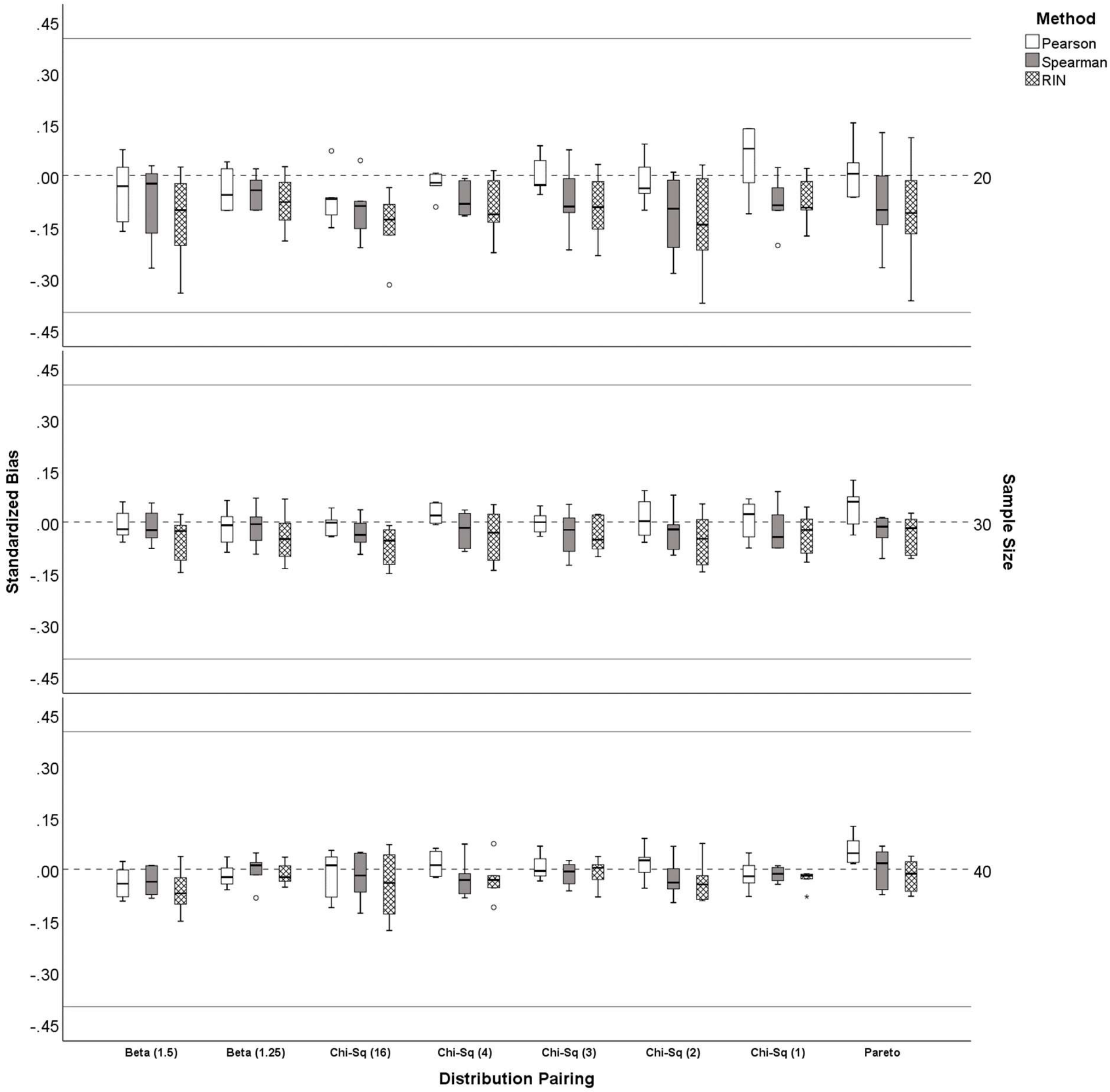

Figure 56. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings by sample size of $20-40$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

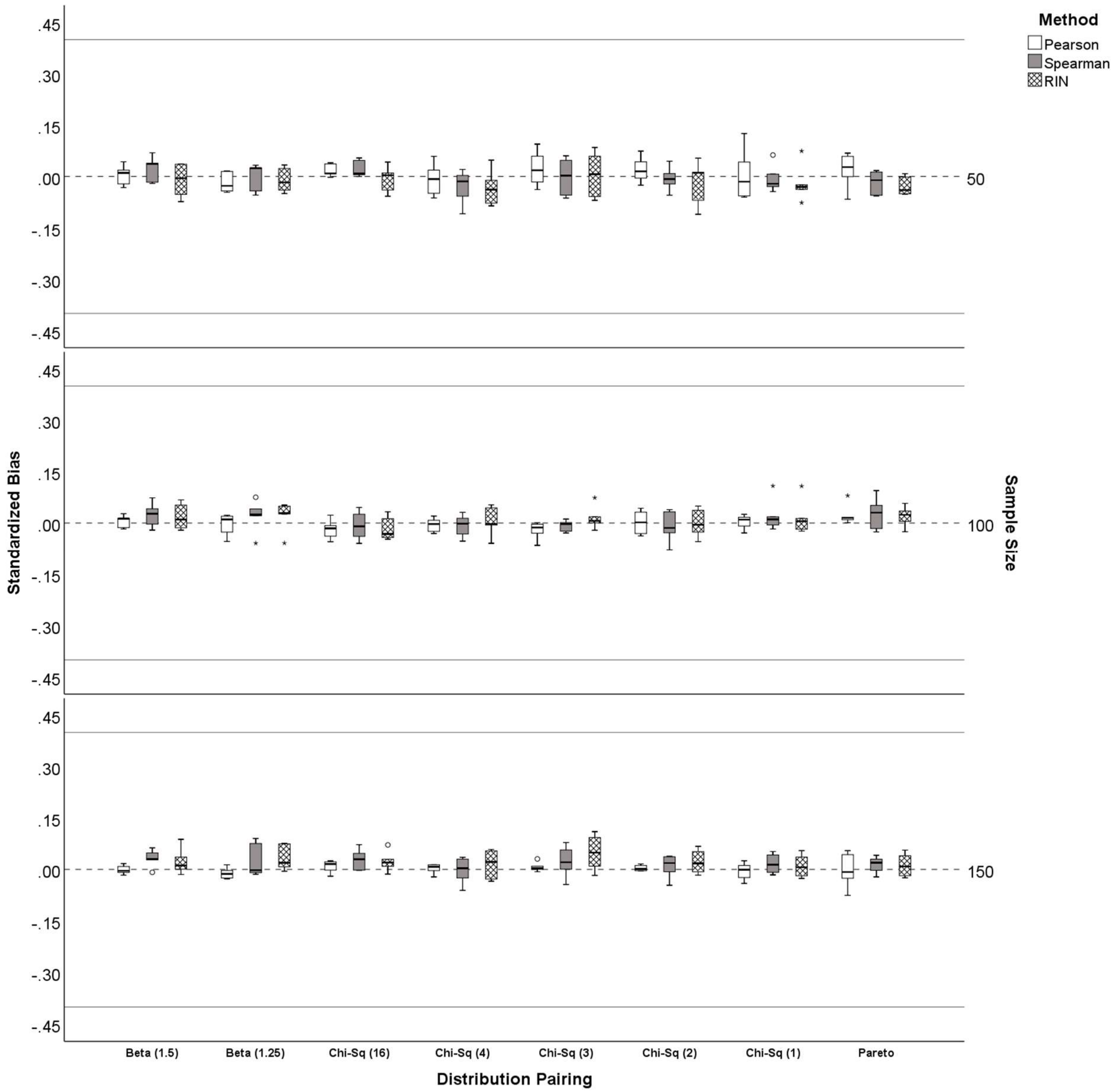

Figure 57. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings by sample size of $50-150$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

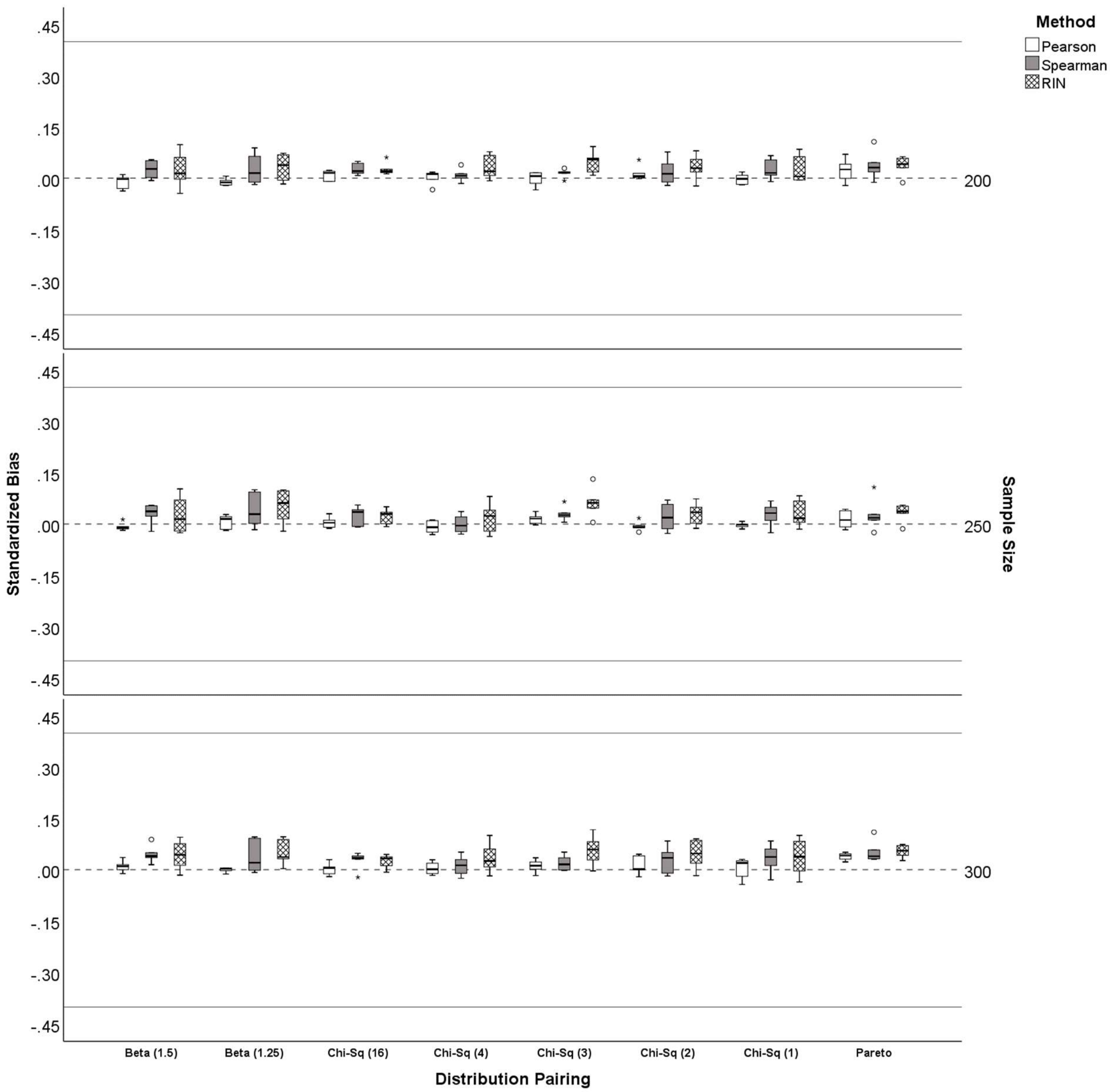

Figure 58. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings by sample size of $200-300$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

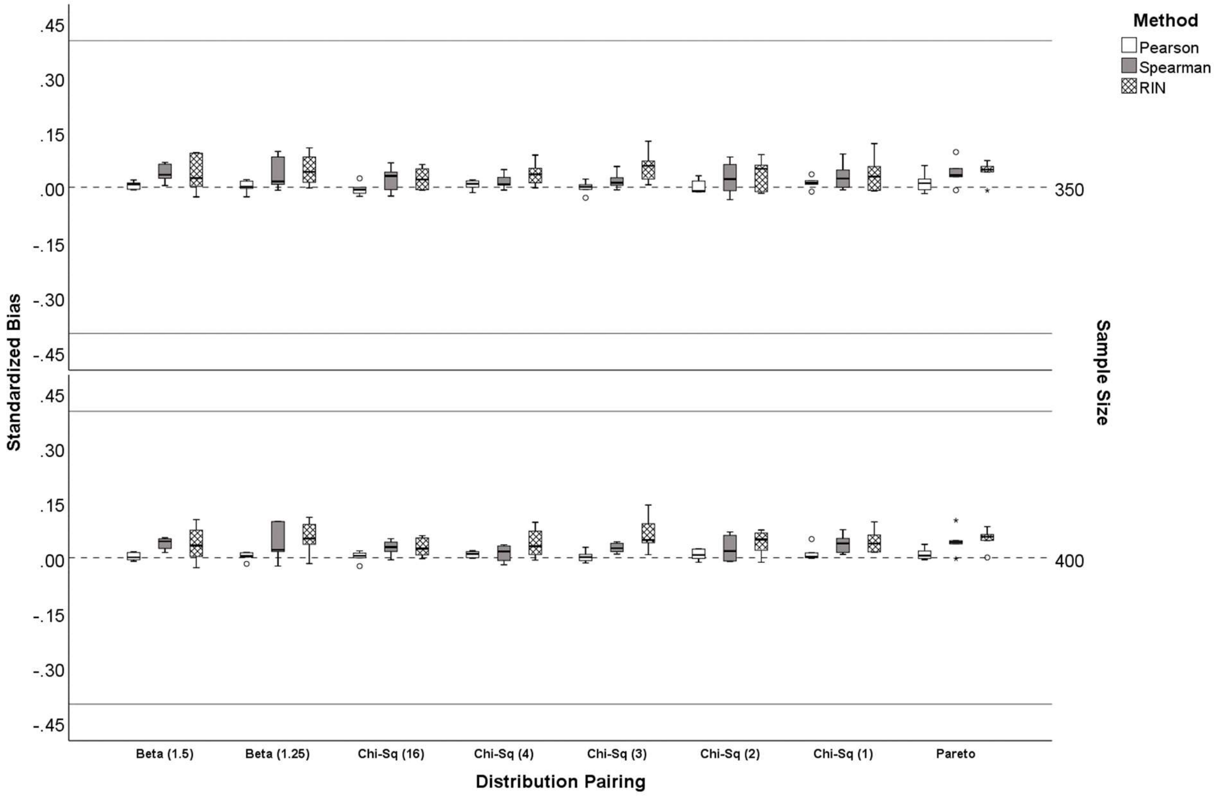

Figure 59. Distribution of standardized bias for non-symmetric with non-symmetric distribution pairings by sample size of $350-400$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

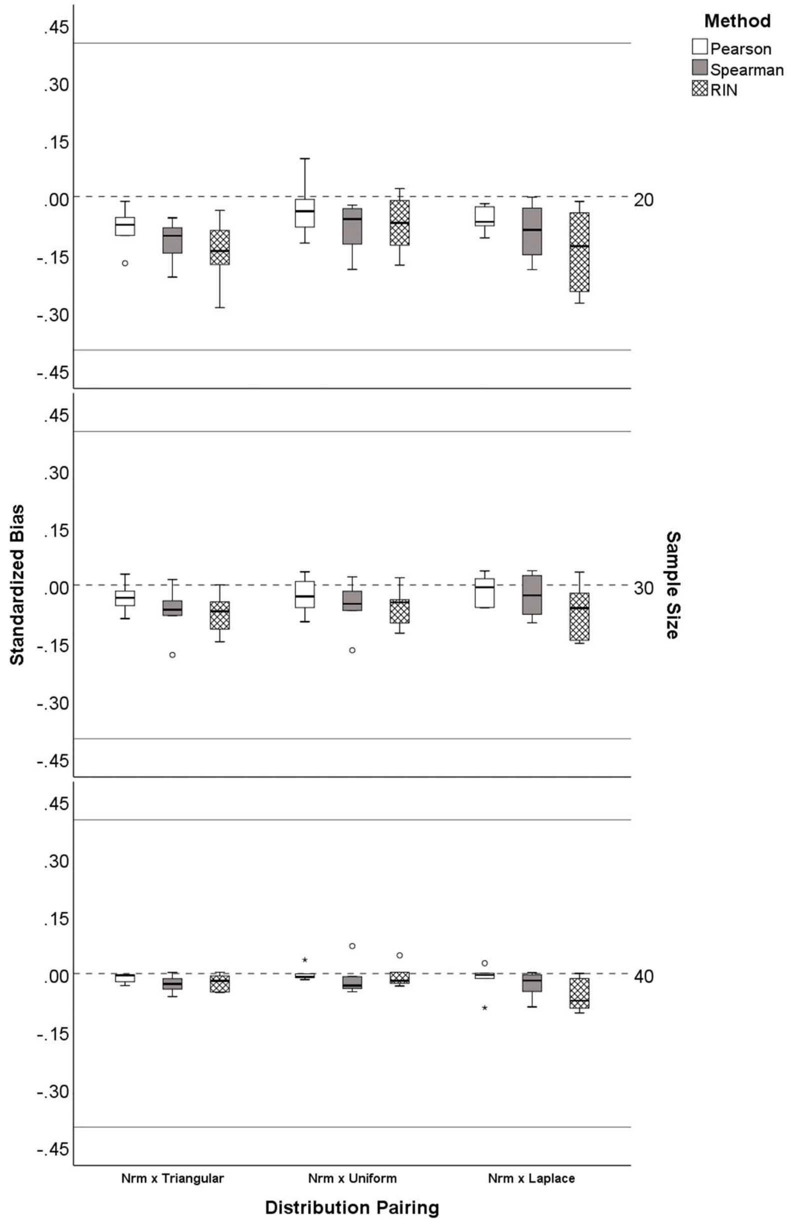

Figure 60. Distribution of standardized bias for symmetric with normal distribution pairings by sample size of $20-40$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

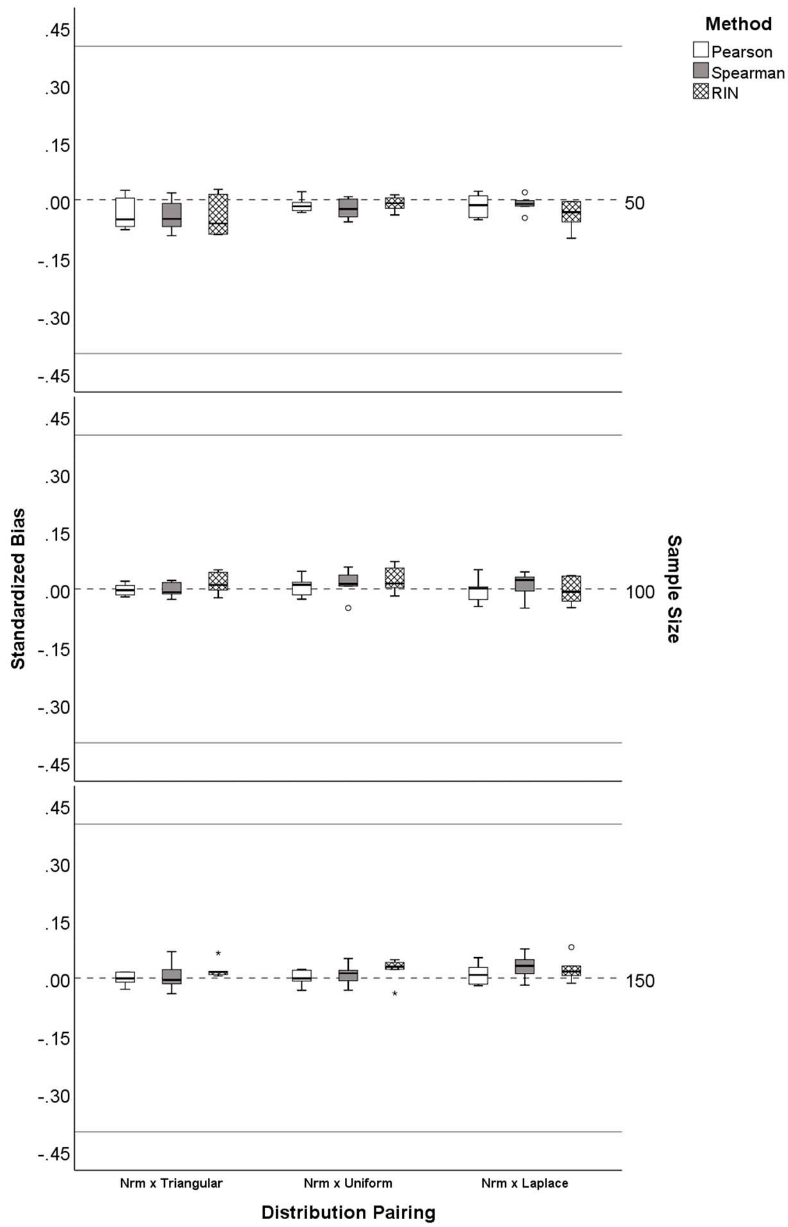

Figure 61. Distribution of standardized bias for symmetric with normal distribution pairings by sample size of $50-150$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

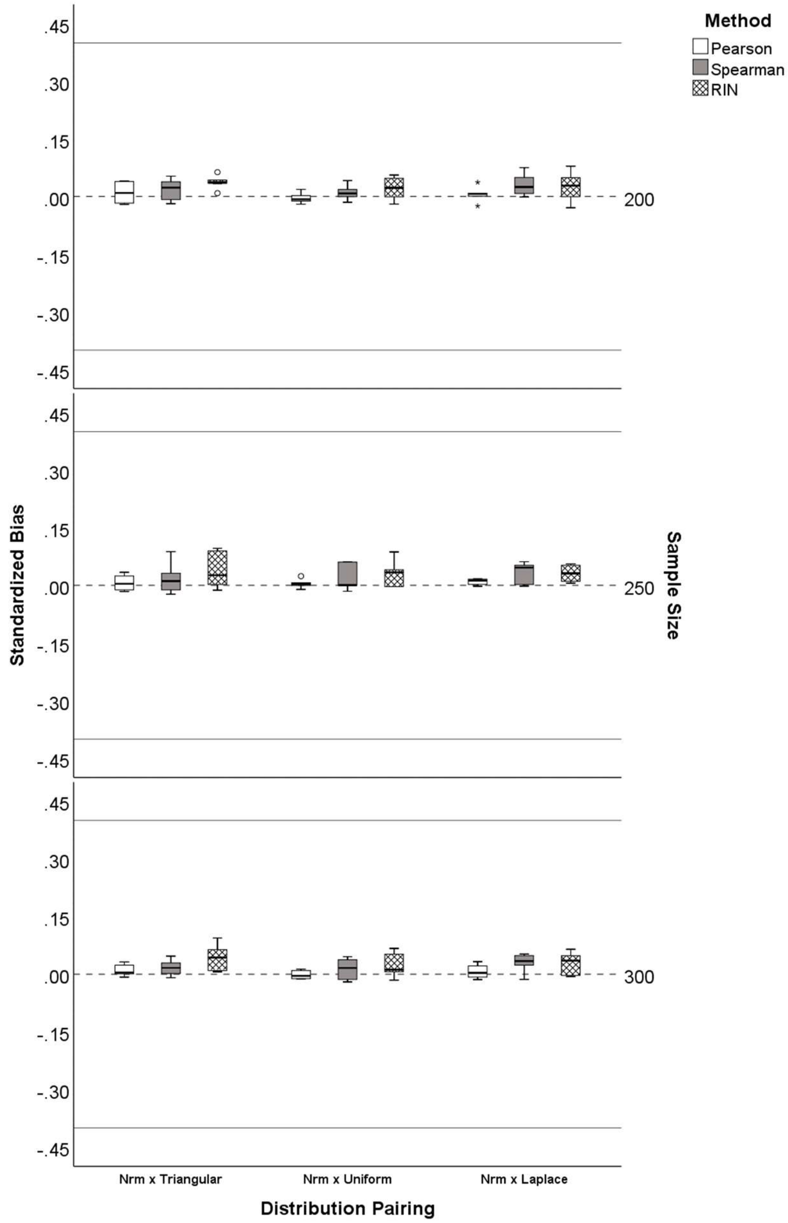

Figure 62. Distribution of standardized bias for symmetric with normal distribution pairings by sample size of $200-300$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

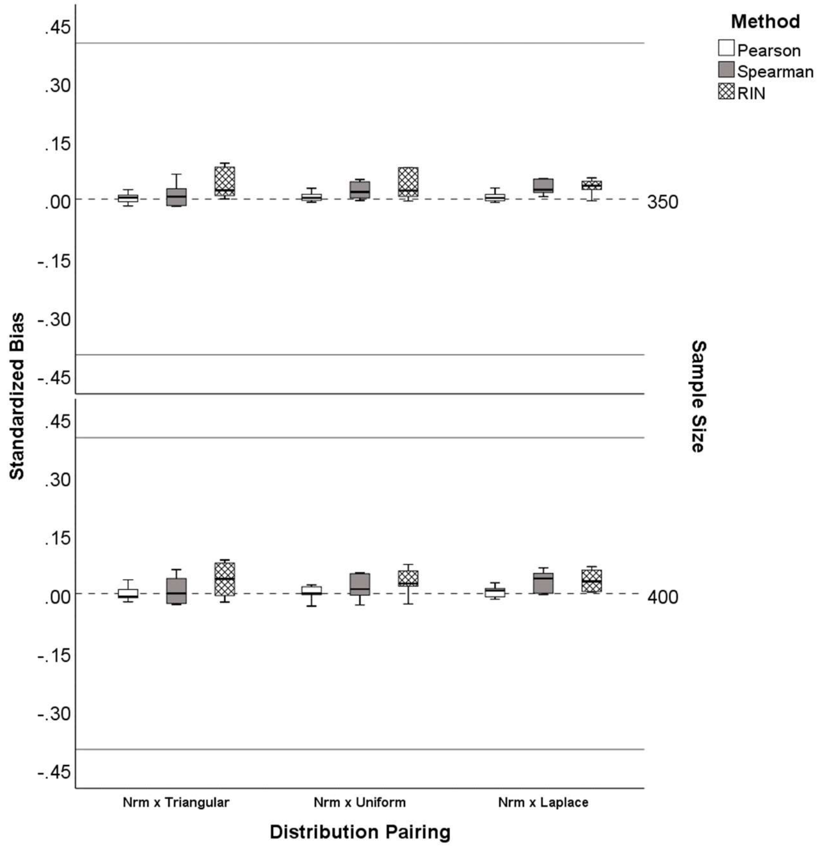

Figure 63. Distribution of standardized bias for symmetric with normal distribution pairings by sample size of $350-400$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

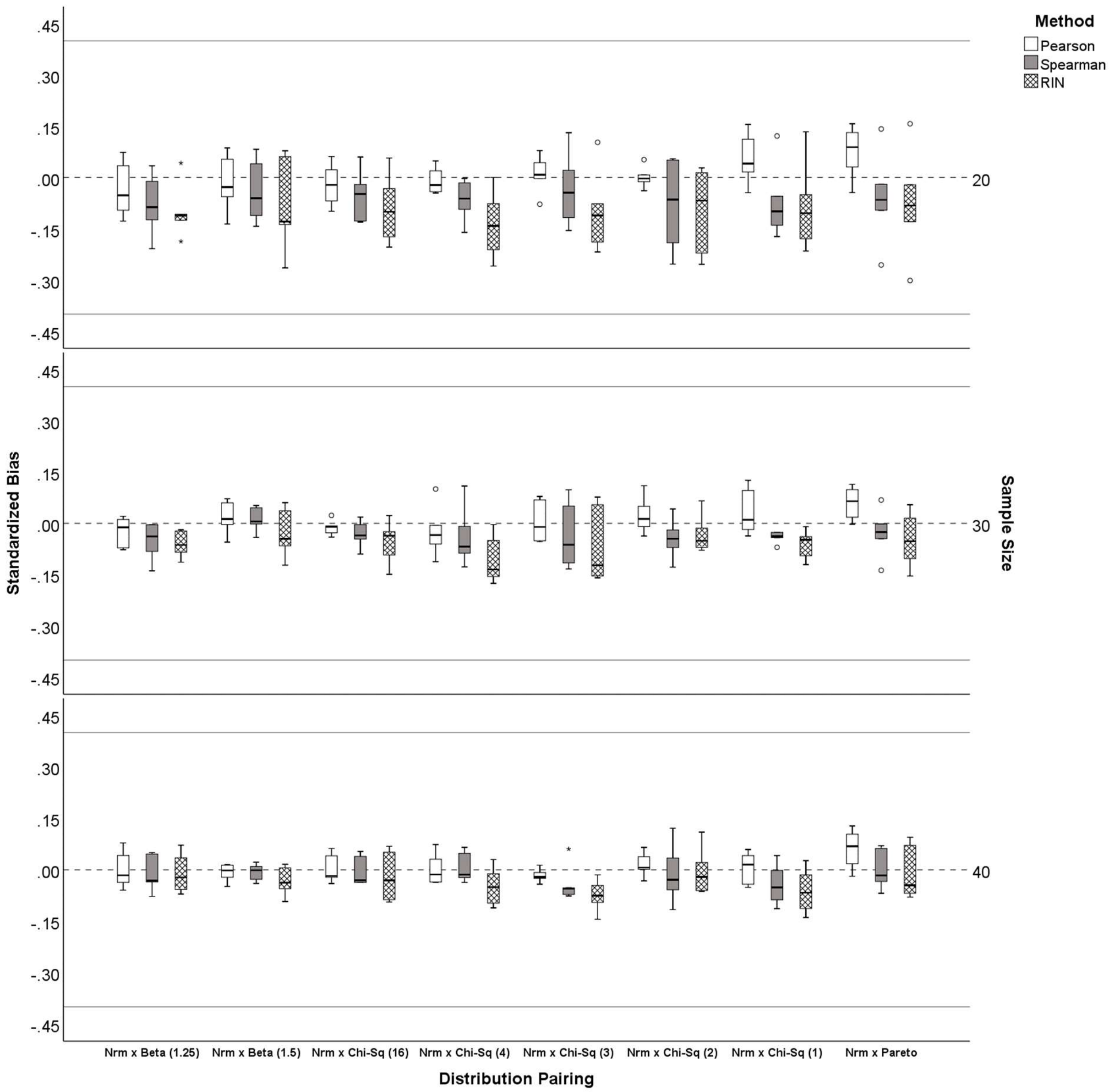

Figure 64. Distribution of standardized bias for non-symmetric with normal distribution pairings by sample size of $20-40$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

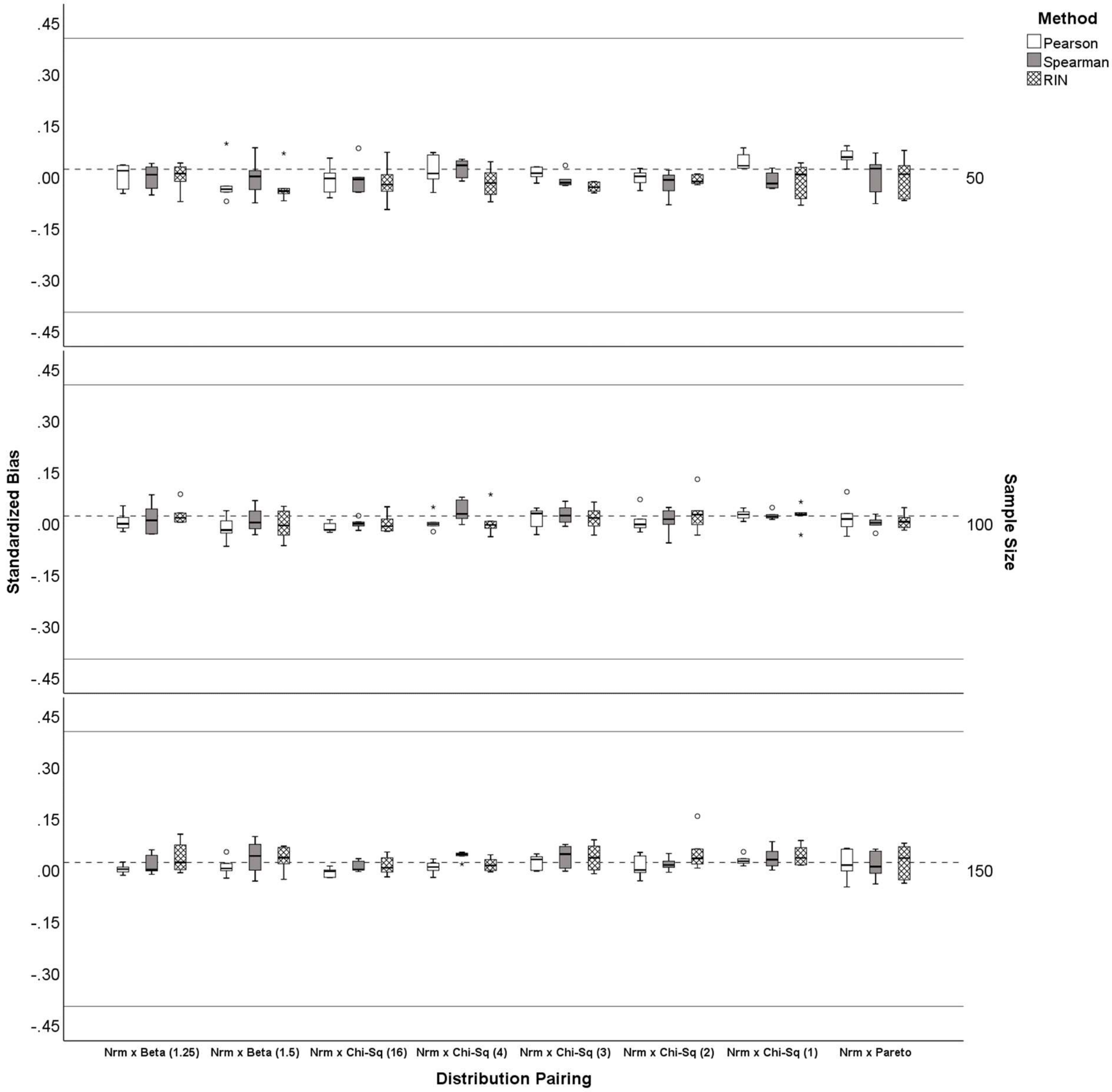

Figure 65. Distribution of standardized bias for non-symmetric with normal distribution pairings by sample size of $50-150$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40,.40]$; acceptable bias.

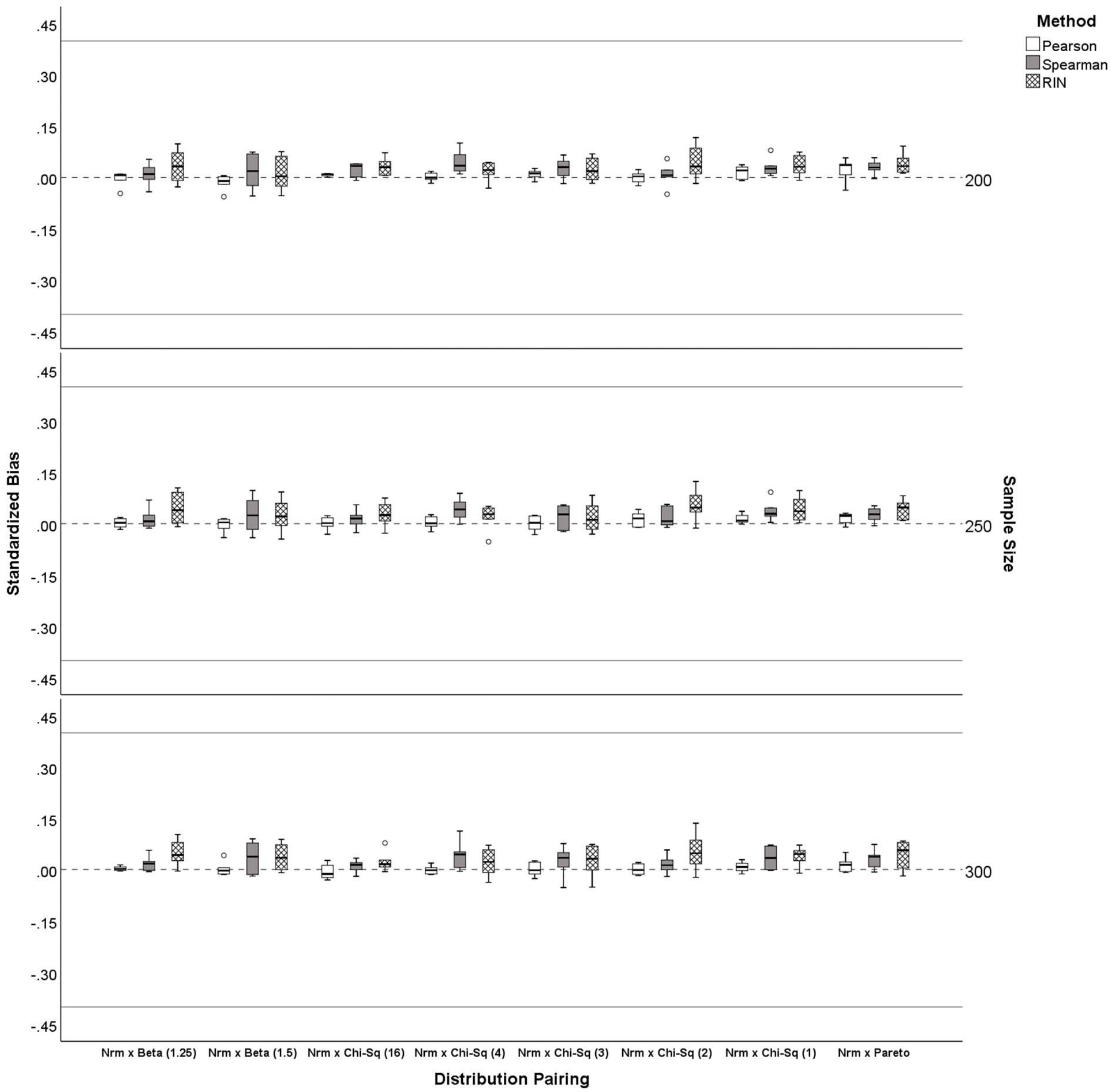

Figure 66. Distribution of standardized bias for non-symmetric with normal distribution pairings by sample size of $200-300$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

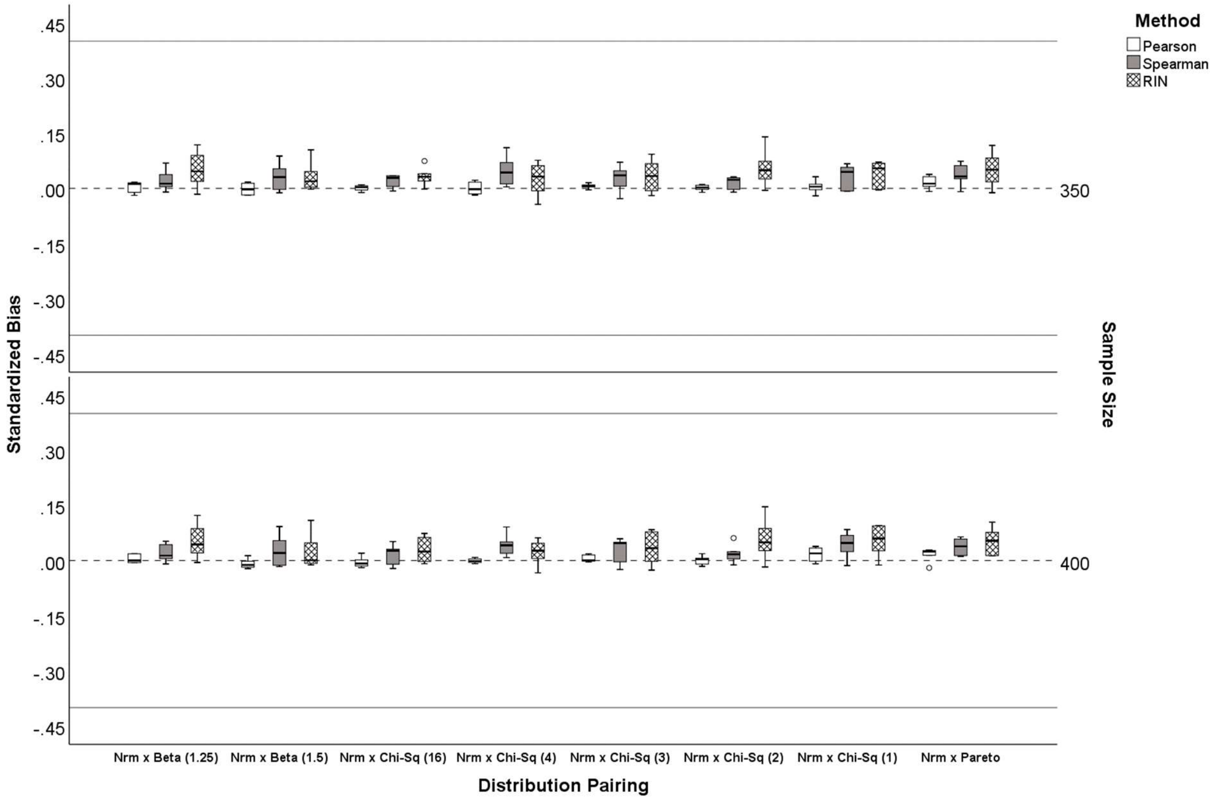

Figure 67. Distribution of standardized bias for non-symmetric with normal distribution pairings by sample size of $350-400$. Ranked inverse normal (RIN). The dashed line is at 0 and the solid lines are at $[-.40, .40]$; acceptable bias.

APPENDIX C

SOURCE CODE

This appendix includes source code for the calculation of the 95% confidence intervals. All code was written in R.

Core Code

```
# Package for Fisher z transformation CI and Spearman CI
library(Desctools);

# Package for ranked inverse normal transformation
library(RNOmni);

# package for highest probability density interval package
library(HDinterval);

# Package for data generation via Headrick's method
library(SimMultiCorrData)
```

```
#Function for bootstrap, percentile bootrstrap IC and bias-
corrected and accelerated CI

boot_fun=function(X,stat,B=1000,prb=.05,qtype=6,rngs=NULL,
stat_out=0,out=0) {

#----------------------------------------------------#
# INPUT ARGUMENTS                                    #
#                                                    #
# X= input data                                      #
# stat= a function of the input data that            #
#       returns a 1xK vector of the stat/parameter   #
#       of interest; K= # of stat/parameter          #
# B= # of bootstrap replicates                       #
# prb= alpha for 100(1-alpha)% confidence interval   #
# qtype= quantile type; see R for details            #
# rngs= random number generator seed                 #
# stat_out= CIs to output                            #
#           0= percentile CI (default)               #
#           1= percentile & BCa CIs                  #
# out= output BxK matrix of Bootstraped              #
#      parameter estimates                           #
#      0= no (default)                               #
#      1= yes                                        #
#----------------------------------------------------#
# OUTPUT                                             #
#                                                    #
# list with CIs & bootstrap parameter estimates      #
#----------------------------------------------------#
```

```
bc_a= function(x,stat,th_,th_0,prb=.05) {

#-----------------------------------------------#
# INPUT ARGUMENTS                               #
#                                               #
# x=  input data                                #
# stat= a function of the input data that       #
#       returns a 1xK vector the stat/parameter #
#       of interest; K= # of stat/parameter     #
# th_= BxK matrix of Bootstraped parameter      #
#      estimates; B= # of bootstrap replicates  #
# th_0= 1xK or Kx1 vector of observed parameter #
#       estimates                               #
# conf= 100(1-alpha)% confifence interval       #
#-----------------------------------------------#
# OUTPUT                                        #
#                                               #
# list with BCa CIs & adjusted alpha            #
#-----------------------------------------------#

   N= dim(x)[1];
   b_rep= dim(th_)[1];
   K= dim(th_)[2];

   if (K!=length(th_0)) {
      stop("th_ & th_0 column dimension do not match");
   }

   th_0= matrix(th_0,byrow=TRUE,b_rep,K);

   # Bias correction factor
   z0= qnorm(apply(th_ < th_0,2,sum)/b_rep);

   jck_th= matrix(NA,N,K);
   for (i in 1:N) {
      jck_th[i,]= stat(x[-i,]);
   }

   L= matrix(apply(jck_th,2,mean),byrow=TRUE,N,K) - jck_th;
   # Acceleration correction factor
   a= apply(L^3,2,sum)/(6*apply(L^2,2,sum)^1.5);

   alph= c(prb,1 + (1 - prb))/2;
   z_a= qnorm(alph);

   adj_alph= matrix(NA,K,length(z_a));
   limit= matrix(NA,K,length(z_a));
```

```
   for (k in 1:K) {
      adj_alph[k,]<-      pnorm(z0[k]      +      (z0[k]+z_a)/(1-
a[k]*(z0[k]+z_a)));
      limit[k,]= quantile(th_[,k], adj_alph[k,], type=qtype);
   }

   name_c= c("bca_lci","bca_uci","adj_a_lci","adj_a_uci");
   limit= cbind(limit,100*adj_alph)
   colnames(limit)= name_c;

   return(list(limit=limit));

}   # End for bc_a

   n= dim(X)[1];

   parm_est= as.matrix(stat(X));
   KK= dim(parm_est)[2];

   b_tht= matrix(NA,B,KK);

   if (!(is.null(rngs))) {
         set.seed(rngs);
   }

   for (b in 1:B) {
      i= sample(1:n,size=n,replace=TRUE);
      b_tht[b,]= stat(X[i,]);
   }

   lo= prb/2;
   up= 1-lo;

   b_mean= colMeans(b_tht);
   b_std= apply(b_tht,2,sd);
   bias= (b_mean - parm_est);
   ratio= abs(bias)/b_std;
   rownames(bias)= "bias";
   rownames(parm_est)= "parm_est";
   rownames(ratio)= "ratio";
   b_mu_ci= apply(b_tht,2,quantile,prob=c(lo,.5,up),type=qtype);
   name_r= c("p_lci","b_med","p_uci");
   rownames(bias)= "bias";
   rownames(parm_est)= "parm_est";
   rownames(ratio)= "ratio";
   rownames(b_mu_ci)= name_r;
```

```
   b_mu_ci=
t(rbind(parm_est,b_mean,bias,b_std,ratio,prb,B,b_mu_ci));

   if (stat_out==1) {
      bca_parm=
bc_a(x=X,stat=stat,th_=b_tht,th_0=parm_est,prb=prb);
      b_mu_ci= cbind(b_mu_ci,bca_parm$limit);
   }

   if (out==0) {
      return( list(b_mu_ci=b_mu_ci) );
   }

   if (out==1) {
      return( list(b_mu_ci=b_mu_ci,b_tht=b_tht) );
   }

}    # End for boot_fun
```

VITA

John Mart V. DelosReyes

Department of Psychology
Mills Godwin Building, Room 250
Old Dominion University
Norfolk, VA 23529

EDUCATION

Ph.D., Health Psychology (current), Old Dominion University, Norfolk, VA
M.S., Health Psychology (current), Old Dominion University, Norfolk, VA
B.S., Psychology (May 2015), Old Dominion University, Norfolk, VA

PUBLICATIONS

Zaharieva, J. N., DelosReyes, J. V., Cullen, K.K., & Padilla, M.A. (2014). *Do published student evaluations of teaching push faculty towards grade inflations and decreased quality of instruction?* Association for Psychological Science Student Research Grant Competition.

CONFERENCES/SYMPOSIUM PRESENTATIONS

Cullen, K.K., DelosReyes, J. V., Zaharieva, J. N., & Padilla, M. A. (2014). *Student satisfaction and faculty evaluations: Does grade-curving have a factor?* Old Dominion University Undergraduate Research Symposium, Norfolk, VA.

DelosReyes, J. V., Cullen, K. K., Zaharieva, J. N., & Padilla, M. A. (2014). *Does course content influence student satisfaction?* Old Dominion University Undergraduate Research Symposium, Norfolk, Va.

DelosReyes, J. V., Manning, S., Sabo, S., Deshpande, A., & Padilla, M. A. (2015). *Developing a measure of psychological aggression: First steps*. Old Dominion University Undergraduate Research Symposium, Norfolk, VA.

DelosReyes, J. V., Manning, S., Sabo, S., Deshpande, A., & Padilla, M. A. (2015). *Developing a measure of psychological aggression: First steps*. Virginia's Collegiate Honors Council Conference, Richmond, VA.

DelosReyes, J.V. (2015) *How to read a scientific article: A primer into basic research literacy.* Psi Chi International Honor Society.

DelosReyes, J. V., & Padilla, M. A. (2019). Correlation confidence intervals robust to non-normality. Association for Psychological Science, Washington, D.C.

DelosReyes, J. V., & Padilla, M. A. (2019). Estimation of the correlation via the bootstrap: Non-normal distributions. William and Mary University, Williamsburg, VA.

TEACHING EXPERIENCE

Teaching Assistant: Quantitative Methods, Old Dominion University (2016 – 2019)

SERVICE

President: Applied Experimental Psychology Student Association (2017 – 2018)
Vice President: Applied Experimental Psychology Student Association (2018 –2019)