Journal of Mathematics and Statistics

Research Article

Absolute Measures of Interindividual Variation: A Simulation Study

José Moral de la Rubia

School of Psychology, Universidad Autónoma de Nuevo León, Monterrey, Mexico

Article history
Received: 01-03-2026
Revised: 23-05-2026
Accepted: 09-06-2026

Email:
jose.morald@uanl.edu.mx

Abstract: Four absolute measures and one relative measure of interindividual variability have been defined, together with their point and interval estimation (bootstrap and asymptotic), and an R script has been developed for their computation. However, further simulation studies are required to strengthen their methodological evaluation. This study focused on the three absolute measures with potentially better regularity conditions, Mean Interindividual Differences (*MID*), Interindividual Variance (*IVAR*), and Interindividual Standard Deviation (*ISD*). Its objectives were: (1) To examine whether bootstrap and asymptotic approaches for these statistics show comparable performance, assessed through their standard errors and confidence interval widths; and (2) To determine whether their sampling distributions converge to normality. Using Monte Carlo simulation, 1000 samples were generated for each of 65 conditions (13 sample sizes ×5 distributions). Within each condition, mean Wald-type and bootstrap confidence interval widths, as well as log-transformed asymptotic and bootstrap standard errors, were compared using paired t tests. Equality of variability between asymptotic and bootstrap standard errors was evaluated with the Pitman–Morgan test. Normality of the bootstrap sampling distributions was assessed using the Shapiro–Francia test. The results support asymptotic equivalence between asymptotic and bootstrap standard errors and between Wald-type and Bias-Corrected and accelerated (BCa) confidence interval widths, with faster convergence for *MID* and *ISD* and slower convergence for *IVAR*. The bootstrap sampling distributions of *MID* and *ISD* showed clear convergence to normality, whereas *IVAR* converged more slowly. Overall, the three statistics exhibit regularity properties that justify the application of asymptotic methods. Their use in empirical research is recommended.

Keywords: Monte Carlo Simulation, Bootstrapping; Standard Error, Confidence Interval, Sampling Distribution

Introduction

Measures based on pairwise interindividual differences provide an alternative framework for quantifying variability. Recently, four absolute measures derived from the mean or median of the differences between each sample observation and all remaining observations were introduced, together with corresponding asymptotic and bootstrap inference procedures (Moral-de la Rubia, 2026a). These measures include the Mean Interindividual Difference (*MID*), the Median Interindividual Difference (*MEID*), the Interindividual Variance (*IVAR*), and the Interindividual Standard Deviation (*ISD*).

Despite the availability of inferential procedures for these statistics, their finite-sample behavior and practical performance under varying distributional conditions remain insufficiently understood. In particular, the relative performance of asymptotic and bootstrap approaches, the regularity of the associated sampling distributions, and the practical conditions under which these measures provide stable inference have not yet been systematically evaluated. Accordingly, the present study examines the inferential behavior of these measures through a comprehensive simulation framework focused on sampling variability, confidence interval performance, and convergence to normality under different sample

sizes and probability distributions.

For a sample of size *n*, the $n \times (n - 1) / 2$ positive pairwise differences between data points represent the interindividual differences. The Mean Interindividual Difference (*MID*) is defined as the average of these differences, whereas the Median Interindividual Difference (*MEID*) corresponds to their median. The Interindividual Variance (*IVAR*) is the average of the squared interindividual differences, and the Interindividual Standard Deviation (*ISD*) is the square root of the interindividual variance.

For the four absolute measures of variation based on interindividual differences, asymptotic standard errors have been proposed (Moral-de la Rubia, 2026b). The Mean Interindividual Differences (*MID*) can be formulated as a symmetric second-order U-statistic (Korolyuk and Borovskich, 2013). As a nondegenerate U-statistic, it satisfies Hoeffding's central limit theorem (Hoeffding, 1948), ensuring asymptotic normality under mild regularity conditions and allowing the derivation of an asymptotic standard error. Refer to Equation (1), in which $\hat{h}_1$ corresponds to the linear component, or first-order empirical influence function, of the *MID* statistic when expressed as a degree-2 U-statistic (Korolyuk and Borovskih, 2013):

$$
\begin{aligned}
&x = \{x_1, x_2, \dots, x_n\} \subset X \quad x_{(1)} \le x_{(1)} \le \cdots \le x_{(n)} \\
&\widehat{MID} = \frac{\sum_{i=1}^{n-1}\sum_{i'>i}^{n}\left|x_{(i)} - x_{(i')}\right|}{n(n-1)/2} \\
&\hat{h}_1 = \left\{\hat{h}_1(x_i)\right\}_{i=1}^{n} = \left\{\frac{1}{n-1}\sum_{i\neq i'}|x_i - x_{i'}| - \widehat{MID}\right\}_{i=1}^{n} \\
&ase\left(\widehat{MID}\right) = 2\frac{s_{n-1}(\hat{h}_1)}{\sqrt{n}} = 2 \times \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{h}_1(x_i) - \overline{\hat{h}_1}\right)^2}{n(n-1)}} \\
&\overline{\hat{h}_1} = \frac{\sum_{i=1}^{n}\hat{h}_1(x_i)}{n}
\end{aligned}
\tag{1}
$$

The Median of Interindividual Differences (*MEID*), in turn, corresponds to a symmetric second-order U-quantile (Wendler, 2011). Under continuity assumptions, this class of statistics also satisfies a version of Hoeffding's central limit theorem (Schultzberg and Ankargren, 2022), which permits asymptotic inference. However, the required smoothness and density conditions are stronger than those for nondegenerate U-statistics, potentially limiting its regularity in finite samples.

By contrast, the asymptotic treatment of Interindividual Variance (*IVAR*) and Interindividual Standard Deviation (*ISD*) is algebraically straightforward. It can be shown that *IVAR* equals twice the sample variance and *ISD* equals √2 times the sample standard deviation. Consequently, their asymptotic standard errors are directly obtained from those of the sample variance and standard deviation, appropriately scaled. Since both statistics are smooth functions of sample moments, standard large-sample theory applies, and Wald-type confidence intervals can be constructed in a direct manner. For more details, refer to Equation (2) for *IVAR* and Equation (3) for *ISD*:

$$
\begin{aligned}
&\widehat{IVAR} = \frac{\sum_{i=1}^{n-1}\sum_{i'>i}^{n}\left(x_{(i)} - x_{(i')}\right)^2}{n(n-1)/2} \\
&ase\left(\widehat{IVAR}\right) = 2 \times ase\left(s_{n-1}^2(x)\right) = 2 \times \\
&\sqrt{\frac{m_4(x) - \frac{n-3}{n-1}s_{n-1}^2(x)}{n}} \\
&= 2 \times \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4/n - (n-3)/(n-1)\sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)}{n}}
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
&\widehat{ISD} = \sqrt{\widehat{IVAR}} = \sqrt{\frac{\sum_{i=1}^{n-1}\sum_{i'>i}^{n}\left(x_{(i)} - x_{(i')}\right)^2}{n(n-1)/2}} = \sqrt{2s_{n-1}^2} = \sqrt{2} \times \\
&s_{n-1}(x) \\
&ase\left(\widehat{ISD}\right) = \sqrt{2} \times ase\left(s_{n-1}(x)\right) = \sqrt{2} \times \frac{ase\left(s_{n-1}^2(x)\right)}{4\hat{\sigma}^2(x)} \\
&= \sqrt{2} \times \sqrt{\frac{m_4(x) - \frac{n-3}{n-1}s_{n-1}^4(x)}{4ns_{n-1}^2(x)}} = \sqrt{\frac{m_4(x) - s_{n-1}^4(x)}{2ns_{n-1}^2(x)}} = \frac{ase(\widehat{IVAR})}{2 \times \widehat{ISD}}
\end{aligned}
\tag{3}
$$

Therefore, *MID*, *IVAR*, and *ISD* exhibit clearer and, in general, less restrictive regularity conditions for asymptotic inference than *MEID*, which justifies focusing the present simulation study on these three measures.

The previous study (Moral-de la Rubia, 2026a), which was essentially theoretical, included two illustrative applications intended to demonstrate the computation and interpretation of the proposed measures and their associated inferential procedures. The first used a large random sample of 1,000 observations drawn from a continuous PERT distribution (0, 8, 10), and the second used a medium-sized sample of 100 observations drawn from a discrete binomial distribution (n = 10, p = 0.25). Although these examples suggested similar behavior between asymptotic and bootstrap approaches under specific continuous and discrete distributions, they were not designed as a systematic evaluation of inferential performance across controlled experimental conditions.

In the continuous example, the sampling distributions of the measures were approximately normal, and the asymptotic and bootstrap estimates showed substantial agreement. In the discrete example, *MID* and *ISD* also exhibited stable inferential behavior, whereas *IVAR* showed mild skewness and *MEID* presented instability in bootstrap estimation. These preliminary observations motivated the need for a broader simulation-based assessment involving varying sample sizes and probability distributions.

The *MID* and *ISD* statistics have been used as indicators of intraindividual variability in fields such as differential psychology (Hyun et al., 2022; Kohoutová et al., 2022), evolutionary psychology (Van Geert and Van Dijk, 2002), cognitive science (Shryock, and Weston, 2025), and

population genetics (Chen et al., 2022). However, the present study does not conceptualize these measures as domain-specific indicators. Instead, they are treated as general descriptive measures of variability applicable to quantitative variables and, potentially, to ordinal variables as well (Alabi and Bukola, 2023).

Promoting their use in this broader descriptive sense requires not only reliable and accessible computational tools but also a solid understanding of their statistical behavior. Since an R-based computational tool (R Core Team, 2025) has already been developed for their calculation (Moral-de la Rubia, 2026c) and is available as a Word file in a GitHub repository at https://github.com/josemoraldelarubia579/R_script_for_IV_Measures.git, the remaining step is to determine whether the associated inferential procedures remain stable under realistic sampling and distributional conditions. For these measures to be useful in applied settings, researchers must know whether asymptotic and bootstrap procedures provide comparable results and whether the corresponding sampling distributions adequately approximate normality in finite samples.

In this context, the present study focuses on *MID*, *IVAR*, and *ISD* as the measures with the most favorable regularity conditions for systematic inferential evaluation. Accordingly, the study pursues two objectives: First, to evaluate whether bootstrap and asymptotic approaches for *MID*, *IVAR*, and *ISD* exhibit comparable performance, as assessed through their standard errors and confidence interval widths; and second, to determine whether the bootstrap-generated sampling distributions of these statistics converge to normality under different distributional and sampling conditions.

The present study focuses on *MID*, *IVAR*, and *ISD* because these measures exhibit comparatively clearer and less restrictive regularity conditions for asymptotic inference. The inferential behavior of *MEID*, which depends on stronger regularity assumptions associated with U-quantiles, is examined separately in Moral-de la Rubia (2026b). In addition, the relative interindividual variation measure, whose inferential framework is more complex, is addressed in a complementary study (Moral-de la Rubia, 2026c).

Methods

One thousand random samples were generated for each of 13 sample sizes (30, 40, from 50 to 500 in increments of 50, and 1000) from three continuous probability distributions a symmetric mesokurtic distribution (normal distribution with parameters $\mu = 0$ and $\sigma = 1$), a symmetric platykurtic distribution (arcsine distribution with support (0, 1)), and a skewed leptokurtic distribution (gamma distribution with parameters $\alpha = 10$ and $\beta = 1$) as well as from two discrete probability distributions: a binomial distribution with parameters $n = 10$ and $p = 0.25$, and a Poisson distribution with parameter $\lambda = 5$. A total of 65,000 random samples were generated by fixing a single seed (123) for each of the 65 conditions (13 sample sizes ×5 distributions) to ensure reproducibility.

Within each condition, for 1000 Monte Carlo–generated samples, the bootstrap standard errors and the widths of the 95% bootstrap confidence intervals, using the BCa method (with 1000 drawings), were computed for three measures of variation based on interindividual differences. In addition, the asymptotic standard errors and the widths of the 95% Wald-type confidence intervals were calculated.

Regarding the first objective, for these three statistics, mean widths of the 95% confidence intervals (bootstrap vs. Wald-type) were compared using paired t-tests across the 65 conditions (13 sample sizes ×5 probability distributions). Effect size was estimated using Cohen's paired-samples *d*, defined as $d = t/\sqrt{n}$, where *t* denotes the test statistic and *n* represents the number of paired observations. Values of *d* less than 0.2 indicate a trivial effect size, values between 0.2 and 0.49 indicate a small effect size, values between 0.5 and 0.79 indicate a medium effect size, and values greater than or equal to 0.8 indicate a large effect size (Cohen, 1988).

For these statistics, standard errors obtained using the two estimation methods (bootstrap vs. asymptotic) were compared across the 65 conditions using the *t*-test for equality of variances in correlated (paired) samples proposed by Morgan (1939); Pitman (1939), as implemented in R (Champely, 2018). This test evaluates differences in variability between methods (H_0: $\sigma^2(SE\_\text{asympt}) = \sigma^2(SE\_\text{boot})$). Rejection of H_0 indicates that one estimation method exhibits greater instability, reflected in a larger variance of standard errors across simulation conditions.

In addition, standard errors were log-transformed, and their means were compared using a paired Student's *t* test (H_0: $E[\ln(SE\_\text{asympt})] = E[\ln(SE\_\text{boot})]$, equivalently H_0: $E[\ln(SE\_\text{asympt} / SE\_\text{boot})] = 0$). Rejection of this null hypothesis indicates that one method systematically yields larger standard errors than the other (Raudenbush, 1988). Effect sizes for both tests were estimated using Cohen's paired-samples *d*, computed as $d = t / \sqrt{n}$ (Cohen, 1988).

Regarding the second objective, the goodness of fit to normality of the bootstrap distributions of the three statistics was assessed using the Shapiro Francia W test (Royston, 1993; Shapiro and Francia, 1972). The proportion of samples for which the null hypothesis of normality was retained at a significance level of $\alpha = 0.05$ was calculated for each of the 65 conditions (13 sample sizes ×5 distributions).

Results

Asymptotic vs. Bootstrap SE and Wald vs. BCa CI Widths for MID

The average *MID* estimates were highly stable across sample sizes within each distribution and showed no

monotonic trend (Table 1). For the standard normal distribution, the estimates ranged from 1.1245 to 1.1314, with a mean (*m*) of 1.1282. In the arcsine distribution, values ranged from 0.4041 to 0.4059 (*m* = 0.4053). For the gamma distribution, estimates varied between 3.5034 and 3.5376 (*m* = 3.5239). In the binomial distribution, values ranged from 1.5080 to 1.5191 (*m* = 1.5123), whereas in the Poisson distribution they ranged from 2.4836 to 2.5060 (*m* = 2.4922).

For the normal, gamma, binomial, and Poisson distributions, the Wald-type confidence intervals were consistently narrower than the BCa intervals, with mostly small effect sizes that tended to decrease as sample size increased (Table 2). The reduction in the mean difference was especially evident in the normal distribution, where it decreased from –0.0121 (*n* = 40) to –0.0003 (n = 1000).

A similar pattern was observed for the gamma, binomial, and Poisson distributions, although medium effect sizes were found for some smaller and intermediate sample sizes. In contrast, for the arcsine distribution, the Wald-type intervals were wider than the BCa intervals. In this case, the effect size was large for sample sizes between 30 and 50, moderate for sample sizes of 100 and 150, and became trivial for larger sample sizes, with nonsignificant differences from n = 300 onward.

Table 1: Average Point Estimates of Three Interindividual Variation Statistics

Distrib.	*n*	*MID*	*IVAR*	*ISD*	Distrib.	*n*	*MID*	*IVAR*	*ISD*
Normal	30	1.1248	1.9845	1.3980	Binomial	30	1.5109	3.7451	1.9193
N(0, 1)	40	1.1271	1.9941	1.4031	(*n* = 10,	40	1.5191	3.7846	1.9337
	50	1.1262	1.9896	1.4034	*p* = 0.25)	50	1.5080	3.7368	1.9237
	100	1.1314	2.0085	1.4138		100	1.5085	3.7376	1.9285
	150	1.1289	2.0048	1.4134		150	1.5141	3.7597	1.9359
	200	1.1285	2.0012	1.4129		200	1.5120	3.7496	1.9340
	250	1.1295	2.0037	1.4141		250	1.5112	3.7459	1.9334
	300	1.1283	1.9990	1.4127		300	1.5134	3.7555	1.9364
	350	1.1245	1.9864	1.4084		350	1.5133	3.7512	1.9356
	400	1.1301	2.0063	1.4155		400	1.5123	3.7505	1.9354
	450	1.1305	2.0081	1.4163		450	1.5111	3.7462	1.9345
	500	1.1276	1.9975	1.4126		500	1.5130	3.7506	1.9356
	1000	1.1297	2.0049	1.4156		1000	1.5133	3.7551	1.9373
	m	1.1282	1.9991	1.4108		*m*	1.5123	3.7514	1.9326
Arcsine	30	0.4049	0.2500	0.4989	Poisson	30	2.4896	9.9869	3.1312
≡	40	0.4041	0.2484	0.4976	(λ = 5)	40	2.5060	10.1472	3.1636
Beta	50	0.4059	0.2509	0.5002		50	2.4914	10.0185	3.1479
(α = 0.5,	100	0.4057	0.2506	0.5003		100	2.4836	9.9567	3.1466
β = 0.5)	150	0.4055	0.2503	0.5001		150	2.4959	10.0346	3.1622
	200	0.4057	0.2506	0.5004		200	2.4901	9.9951	3.1571
	250	0.4056	0.2504	0.5003		250	2.4894	9.9936	3.1575
	300	0.4053	0.2500	0.4999		300	2.4948	10.0348	3.1650
	350	0.4054	0.2502	0.5001		350	2.4915	9.9940	3.1590
	400	0.4053	0.2500	0.5000		400	2.4918	10.0093	3.1614
	450	0.4051	0.2497	0.4997		450	2.4890	9.9902	3.1589
	500	0.4052	0.2499	0.4999		500	2.4923	10.0003	3.1605
	1000	0.4052	0.2499	0.4999		1000	2.4928	10.0137	3.1636
	m	0.4053	0.2501	0.4998		*m*	2.4922	10.0135	3.1565
Gamma	30	3.5034	19.8037	4.4003					
(α = 10,	40	3.5376	20.1304	4.4504					
β = 1)	50	3.5257	20.0508	4.4491					
	100	3.5286	20.0416	4.4626					
	150	3.5182	19.9458	4.4567					
	200	3.5290	20.0491	4.4707					
	250	3.5165	19.9221	4.4573					
	300	3.5125	19.8821	4.4543					
	350	3.5298	20.0701	4.4758					
	400	3.5247	19.9828	4.4668					
	450	3.5291	20.0699	4.4766					
	500	3.5284	20.0475	4.4745					
	1000	3.5275	20.0342	4.4744					
	m	3.5239	20.0023	4.4592					

Note: *n* = sample size, *m* = arithmetic mean. Distrib. = probability distribution. Variation measures based on interindividual differences: *MID* = Mean Interindividual Difference, *IVAR* = Interindividual Variance, and *ISD* = Interindividual Standard Deviation. Elaborated by the author

Table 2: Paired Comparison of Wald-type and BCa Confidence Interval Widths for *MID*

	CI_width		Paired t-test				CI_width		Paired t-test		
n	W-type	Bca	*md*	*t*	*p*	*d*	W-type	Bca	*t*	*p*	*d*
	Normal (0, 1)						Arcsine ≡ Beta(α = 0.5, β = 0.5)				
30	0.5689	0.5787	-0.0098	-7.92	<.001	-0.25	0.1136	0.1047	42.81	<.001	1.35
40	0.4927	0.5048	-0.0121	-12.16	<.001	-0.38	0.0938	0.0888	33.96	<.001	1.07
50	0.4409	0.4508	-0.0099	-13.62	<.001	-0.43	0.0807	0.0774	29.72	<.001	0.94
100	0.3144	0.3197	-0.0053	-13.65	<.001	-0.43	0.0535	0.0526	14.46	<.001	0.46
150	0.2592	0.2631	-0.0039	-12.75	<.001	-0.40	0.0428	0.0424	7.71	<.001	0.24
200	0.2237	0.2268	-0.0030	-12.19	<.001	-0.39	0.0367	0.0366	4.25	<.001	0.13
250	0.1996	0.2018	-0.0022	-9.97	<.001	-0.32	0.0325	0.0324	3.38	0.001	0.11
300	0.1820	0.1836	-0.0017	-8.75	<.001	-0.28	0.0296	0.0296	-0.39	0.693	-0.01
350	0.1684	0.1696	-0.0012	-6.85	<.001	-0.22	0.0273	0.0273	0.81	0.418	0.03
400	0.1583	0.1595	-0.0012	-7.49	<.001	-0.24	0.0255	0.0255	-0.34	0.737	-0.01
450	0.1494	0.1505	-0.0012	-7.55	<.001	-0.24	0.0240	0.0240	1.87	0.061	0.06
500	0.1413	0.1423	-0.0010	-7.04	<.001	-0.22	0.0227	0.0227	-1.18	0.240	-0.04
1000	0.1002	0.1005	-0.0003	-3.21	0.001	-0.10	0.0159	0.0160	-2.97	0.003	-0.09
	Gamma (α = 10, β = 1)						Binomial (*n* = 10, *p* = 0.25)				
30	1.8733	1.9287	-0.0553	-10.58	<.001	-0.33	0.7853	0.7923	-4.01	<.001	-0.13
40	1.6331	1.6828	-0.0497	-13.09	<.001	-0.41	0.6846	0.6949	-7.86	<.001	-0.25
50	1.4776	1.5294	-0.0518	-14.12	<.001	-0.45	0.6129	0.6261	-11.16	<.001	-0.35
100	1.0479	1.0761	-0.0282	-17.25	<.001	-0.55	0.4332	0.4383	-9.28	<.001	-0.29
150	0.8578	0.8754	-0.0175	-15.53	<.001	-0.49	0.3538	0.3578	-10.35	<.001	-0.33
200	0.7442	0.7575	-0.0133	-14.78	<.001	-0.47	0.3057	0.3088	-9.39	<.001	-0.30
250	0.6657	0.6752	-0.0095	-13.30	<.001	-0.42	0.2732	0.2750	-6.52	<.001	-0.21
300	0.6088	0.6163	-0.0074	-11.80	<.001	-0.37	0.2498	0.2522	-9.76	<.001	-0.31
350	0.5650	0.5724	-0.0074	-12.19	<.001	-0.39	0.2303	0.2314	-5.26	<.001	-0.17
400	0.5252	0.5306	-0.0054	-9.47	<.001	-0.30	0.2161	0.2180	-9.02	<.001	-0.29
450	0.4993	0.5044	-0.0051	-10.02	<.001	-0.32	0.2039	0.2048	-4.10	<.001	-0.13
500	0.4724	0.4772	-0.0048	-10.06	<.001	-0.32	0.1926	0.1936	-5.62	<.001	-0.18
1000	0.3340	0.3356	-0.0015	-4.81	<.001	-0.15	0.1367	0.1371	-2.98	0.003	-0.09
	Poisson (λ = 5)										
30	1.2942	1.3160	-0.0218	-6.46	<.001	-0.20					
40	1.1410	1.1702	-0.0292	-11.35	<.001	-0.36					
50	1.0186	1.0489	-0.0303	-13.70	<.001	-0.43					
100	0.7206	0.7346	-0.0140	-13.31	<.001	-0.42					
150	0.5888	0.5975	-0.0087	-12.75	<.001	-0.40					
200	0.5101	0.5179	-0.0078	-13.52	<.001	-0.43					
250	0.4567	0.4613	-0.0046	-9.52	<.001	-0.30					
300	0.4181	0.4232	-0.0051	-11.60	<.001	-0.37					
350	0.3845	0.3871	-0.0026	-6.71	<.001	-0.21					
400	0.3613	0.3649	-0.0036	-9.77	<.001	-0.31					
450	0.3409	0.3429	-0.0020	-5.64	<.001	-0.18					
500	0.3217	0.3243	-0.0026	-8.01	<.001	-0.25					
1000	0.2286	0.2293	-0.0006	-2.89	0.004	-0.09					

Note: CI_width = mean width of the Wald-type (W-type) and Bias-Corrected and accelerated (BCa) confidence intervals, *md* = mean difference, *t* = test statistic, *p* = two-tailed p value, and *d* = Cohen's *d* effect size statistic. Degrees of freedom = 999. *MID* = Mean Interindividual Difference. Elaborated by the author

For the normal, gamma, binomial, and Poisson distributions, the bootstrap standard errors were significantly smaller than the asymptotic standard errors, although the differences progressively decreased as sample size increased. Large effect sizes were observed for small samples, medium effects for intermediate sample sizes, and trivial effects for large samples. A similar pattern was found across these three distributions, with the difference becoming nonsignificant at $n = 500$ in the gamma distribution. In contrast, for the arcsine distribution, the asymptotic standard errors were consistently smaller than the bootstrap standard errors across all sample sizes. In this case, the effect size likewise decreased from large in small samples to trivial in large samples (Table 3). According to the Pitman–Morgan test, the variability of the asymptotic standard errors was generally greater than that of the bootstrap standard errors for small sample sizes. For intermediate sample sizes, differences in variability were often nonsignificant, whereas for large sample sizes the variability of the asymptotic standard errors tended to become significantly smaller than that of the bootstrap standard errors, with the magnitude of the difference gradually increasing as sample size increased (Table 3).

Table 3: Comparison of Asymptotic and Bootstrap Standard Errors of *MID* Using Paired t Test of the Mean Log Ratio and the Pitman–Morgan Test

Dist.	*n*	*SE*		Paired t-test for mean_log_ratio				Pitman–Morgan test			
		Asymp	Boot	m_ln	*t*	*p*	*d*	r_{SD}	*t*	*p*	*d*
N	30	0.1451	0.1413	0.0253	31.08	<.001	0.98	0.47	16.98	<.001	0.54
	40	0.1257	0.1231	0.0200	27.24	<.001	0.86	0.35	11.98	<.001	0.38
	50	0.1125	0.1107	0.0152	21.62	<.001	0.68	0.20	6.41	<.001	0.20
	100	0.0802	0.0795	0.0093	13.28	<.001	0.42	0.00	-0.08	0.936	-0.003
	150	0.0661	0.0657	0.0063	8.83	<.001	0.28	-0.08	-2.63	0.009	-0.08
	200	0.0571	0.0569	0.0037	5.21	<.001	0.16	-0.09	-2.89	0.004	-0.09
	250	0.0509	0.0508	0.0035	5.07	<.001	0.16	-0.14	-4.51	<.001	-0.14
	300	0.0464	0.0463	0.0037	5.34	<.001	0.17	-0.11	-3.40	0.001	-0.11
	350	0.043	0.0428	0.0038	5.51	<.001	0.17	-0.16	-5.07	<.001	-0.16
	400	0.0404	0.0403	0.0020	2.79	0.006	0.09	-0.20	-6.50	<.001	-0.21
	450	0.0381	0.0381	0.0016	2.33	<.001	0.07	-0.19	-6.10	<.001	-0.19
	500	0.0360	0.0360	0.0022	3.06	0.018	0.10	-0.24	-7.75	<.001	-0.25
	1000	0.0256	0.0255	0.0020	2.84	0.005	0.09	-0.33	-11.06	<.001	-0.35
AS	30	0.0290	0.0307	-0.0653	-32.21	<.001	-1.02	0.70	31.21	<.001	0.99
	40	0.0239	0.0251	-0.0546	-35.14	<.001	-1.11	0.64	26.62	<.001	0.84
	50	0.0206	0.0215	-0.0485	-36.60	<.001	-1.16	0.56	21.20	<.001	0.67
	100	0.0137	0.0140	-0.0273	-31.86	<.001	-1.01	0.29	9.66	<.001	0.31
	150	0.0109	0.0111	-0.0151	-20.58	<.001	-0.65	0.09	2.85	0.004	0.09
	200	0.0094	0.0095	-0.0151	-20.58	<.001	-0.65	-0.02	-0.56	0.579	-0.02
	250	0.0083	0.0084	-0.0104	-14.15	<.001	-0.45	-0.10	-3.04	0.002	-0.10
	300	0.0076	0.0076	-0.01	-14.16	<.001	-0.45	-0.11	-3.65	<.001	-0.12
	350	0.0070	0.0070	-0.008	-11.41	<.001	-0.36	-0.14	-4.54	<.001	-0.14
	400	0.0065	0.0066	-0.008	-11.05	<.001	-0.35	-0.14	-4.58	<.001	-0.15
	450	0.0061	0.0062	-0.0045	-6.13	<.001	-0.19	-0.16	-5.05	<.001	-0.16
	500	0.0058	0.0058	-0.0049	-6.79	<.001	-0.21	-0.20	-6.35	<.001	-0.20
	1000	0.0041	0.0041	-0.0032	-4.60	<.001	-0.15	-0.31	-10.46	<.001	-0.33
G	30	0.4779	0.4645	0.0267	33.36	<.001	1.06	0.50	18.18	<.001	0.58
	40	0.4166	0.4080	0.0201	27.67	<.001	0.88	0.35	11.74	<.001	0.37
	50	0.3769	0.3697	0.019	26.67	<.001	0.84	0.25	8.14	<.001	0.26
	100	0.2673	0.2651	0.0085	11.86	<.001	0.38	0.00	-0.02	0.984	0.00
	150	0.2188	0.2174	0.0068	9.76	<.001	0.31	-0.03	-0.81	0.418	-0.03
	200	0.1898	0.1889	0.0051	6.98	<.001	0.22	0.02	0.68	0.498	0.02
	250	0.1698	0.1691	0.0048	6.65	<.001	0.21	-0.11	-3.39	0.001	-0.11
	300	0.1553	0.1548	0.0035	5.06	<.001	0.16	-0.12	-3.80	<.001	-0.12
	350	0.1441	0.1438	0.0026	3.55	0.0004	0.11	-0.14	-4.31	<.001	-0.14
	400	0.1340	0.1336	0.003	4.14	<.001	0.13	-0.17	-5.30	<.001	-0.17
	450	0.1274	0.1269	0.0036	5.21	<.001	0.16	-0.10	-3.03	0.003	-0.10
	500	0.1205	0.1204	0.0011	1.65	0.0985	0.05	-0.12	-3.79	0.000	-0.12
	1000	0.0852	0.0851	0.0015	2.21	0.027	0.07	-0.22	-7.02	<.001	-0.22
BN	30	0.2003	0.1958	0.0214	26.73	<.001	0.85	0.45	15.82	<.001	0.50
	40	0.1746	0.1715	0.0173	23.28	<.001	0.74	0.33	10.95	<.001	0.35
	50	0.1563	0.1541	0.0144	19.39	<.001	0.61	0.21	6.85	<.001	0.22
	100	0.1105	0.1095	0.0094	13.51	<.001	0.43	-0.01	-0.24	0.808	-0.01
	150	0.0903	0.0897	0.0064	9.60	<.001	0.30	-0.05	-1.44	0.150	-0.05
	200	0.0780	0.0777	0.0043	6.09	<.001	0.19	-0.12	-3.76	<.001	-0.12
	250	0.0697	0.0694	0.0049	7.00	<.001	0.22	-0.06	-1.97	0.049	-0.06
	300	0.0637	0.0636	0.0029	4.20	<.001	0.13	-0.17	-5.55	<.001	-0.18
	350	0.0587	0.0585	0.0036	5.11	<.001	0.16	-0.15	-4.95	<.001	-0.16
	400	0.0551	0.0551	0.0013	1.96	0.050	0.06	-0.23	-7.38	<.001	-0.23
	450	0.0520	0.0518	0.0039	5.38	<.001	0.17	-0.20	-6.46	<.001	-0.20
	500	0.0491	0.0490	0.0018	2.61	0.009	0.08	-0.17	-5.49	<.001	-0.17
	1000	0.0349	0.0348	0.0019	2.66	0.008	0.08	-0.28	-9.11	<.001	-0.29
Poi	30	0.3301	0.3219	0.0235	30.61	<.001	0.97	0.50	18.17	<.001	0.58
	40	0.2911	0.2853	0.0192	26.09	<.001	0.83	0.37	12.43	<.001	0.39
	50	0.2598	0.2557	0.0153	20.69	<.001	0.65	0.28	9.21	<.001	0.29
	100	0.1838	0.1819	0.0105	14.74	<.001	0.47	0.00	0.08	0.932	0.00
	150	0.1502	0.1491	0.0075	10.91	<.001	0.35	-0.05	-1.57	0.116	-0.05

200	0.1301	0.1295	0.0048	6.64	<.001	0.21	-0.09	-2.83	0.005	-0.09
250	0.1165	0.1159	0.0049	6.94	<.001	0.22	-0.05	-1.49	0.136	-0.05
300	0.1067	0.1063	0.0033	4.77	<.001	0.15	-0.13	-4.25	<.001	-0.13
350	0.0981	0.0977	0.0041	5.94	<.001	0.19	-0.12	-3.69	<.001	-0.12
400	0.0922	0.0920	0.0019	2.78	0.006	0.09	-0.21	-6.71	<.001	-0.21
450	0.0870	0.0867	0.0035	4.87	<.001	0.15	-0.20	-6.34	<.001	-0.20
500	0.0821	0.0820	0.0016	2.36	0.018	0.07	-0.13	-3.98	<.001	-0.13
1000	0.0583	0.0582	0.0021	2.91	0.004	0.09	-0.25	-8.03	<.001	-0.25

Note: Dist. = probability distribution from which 1000 random samples were drawn: N = standard normal ($\mu = 0$, $\sigma^2 = 1$); AS = arcsine or beta ($\alpha = 0.5$, $\beta = 0.5$); G = gamma ($\alpha = 10$, $\beta = 1$); BN = binomial ($n = 10$, $p = 0.25$); Poi = Poisson ($\lambda = 5$); n = sample size. Paired t test: *SE* = asymptotic (Asymp) and bootstrap (Boot) standard errors; m_ln = mean log ratio of asymptotic (numerator) to bootstrap (denominator) standard errors; t = test statistic; p = two-tailed p value under the t distribution with 999 degrees of freedom for H_0: E[ln(SE_asympt / SE_boot)] = 0; d = Cohen's d effect size. Pitman–Morgan test: r_{SD} = Pearson correlation between the sum and the difference of paired data; t = test statistic; p = two-tailed p-value under the t distribution with 998 degrees of freedom for H_0: σ^2(SE_asympt) = σ^2(SE_boot); d = Cohen's d effect size, whose values in most cases coincide with those of r_{SD}. *MID* = Mean Interindividual Difference. Elaborated by the author

Asymptotic vs. Bootstrap SE and Wald vs. BCa CI Widths for IVAR

The average *IVAR* estimates were highly stable across sample sizes within each distribution and showed no monotonic trend (Table 1). For the normal distribution, the estimates ranged from 1.9845 to 2.0085 (m = 1.9991). In the arcsine distribution, they ranged from 0.2484 to 0.2509 (m = 0.2501), whereas in the gamma distribution they varied between 19.8037 and 20.1304 (M = 20.0023). For the binomial and Poisson distributions, the estimates ranged from 3.7368 to 3.7846 (m = 3.7514) and from 9.9567 to 10.1472 (m = 10.0135), respectively.

Across all five distributions, the BCa confidence intervals were significantly wider than the Wald-type intervals (Table 4). The effect size of the interval estimation method decreased progressively as sample size increased, ranging from large effects in small samples to small or trivial effects in large samples. This pattern was consistent across all distributions analyzed.

For the normal, gamma, binomial, and Poisson distributions, the bootstrap standard errors were generally larger than the asymptotic standard errors, although the differences decreased as sample size increased and became nonsignificant for large samples (Table 5). Effect sizes ranged from large or medium in small samples to trivial in large samples. The variability of the bootstrap standard errors also tended to be greater than that of the asymptotic standard errors, although these effects were mostly trivial. A similar pattern was observed across the three distributions.

For the arcsine distribution, the same direction of differences was observed, but with substantially larger effect sizes. Differences between asymptotic and bootstrap standard errors remained significant for nearly all sample sizes, with very large effects in small samples, large effects at n = 100, and medium effects for larger sample sizes. In addition, the variability differences progressively increased with sample size, reaching large effect sizes in the largest samples, a pattern not observed in the other distributions (Table 5).

Asymptotic vs. Bootstrap SE and Wald vs. BCa CI Widths for ISD

The average *ISD* estimates were highly stable across sample sizes within each distribution and generally showed no monotonic trend, although a slight increasing tendency was observed for the binomial distribution (Table 1). For the standard normal distribution, the estimates ranged from 1.3980 to 1.4163 (m = 1.4108). In the arcsine distribution, values ranged from 0.4976 to 0.5004 (m = 0.4998). For the gamma distribution, estimates varied between 4.4003 and 4.4766 (m = 4.4592). In the binomial distribution, values ranged from 1.9193 to 1.9373 (m = 1.9326), whereas in the Poisson distribution they ranged from 3.1312 to 3.1650 (m = 3.1565).

Across all five distributions, the BCa confidence intervals were significantly wider than the Wald-type intervals (Table 6). The effect size of the interval estimation method decreased progressively as sample size increased, ranging from large effects in small samples to trivial or small effects in large samples. For the largest samples (n = 1000), the effects were trivial for the normal, binomial, and Poisson distributions, but remained small for the arcsine and gamma distributions.

For the samples drawn from the standard normal distribution, the bootstrap standard errors were systematically larger than the asymptotic standard errors.

The effect size of the estimation method on the mean standard error decreased progressively as sample size increased, ranging from very large for n = 30 to trivial for $n \geq 350$. Mean differences were not significant for samples of sizes 450 and 1000, whereas they remained significant for the other sample sizes. A similar decreasing trend in effect size was observed for the gamma, binomial, and Poisson distributions. In the binomial distribution, the differences were not significant for samples of sizes 450 and 1000. In the gamma distribution, the differences were additionally nonsignificant for n = 500. In the Poisson distribution, nonsignificant differences were observed for n = 400, 450, and 1000.

Table 4: Paired Comparison of Wald-type and BCa Confidence Interval Widths for *IVAR*

n	CI_width		Paired t-test				CI_width		Paired t-test		
	W-type	Bca	*md*	*t*	*p*	*d*	W-type	Bca	*t*	*p*	*d*
	Normal (0, 1)						Arcsine ≡ Beta (α = 0.5, β = 0.5)				
30	1.8241	2.0856	-0.2615	-37.23	<.001	-1.18	0.1321	0.1390	-39.71	<.001	-1.26
40	1.6131	1.8093	-0.1962	-37.77	<.001	-1.20	0.1131	0.1181	-35.43	<.001	-1.12
50	1.4564	1.6014	-0.1449	-37.40	<.001	-1.18	0.1005	0.1042	-33.19	<.001	-1.05
100	1.0746	1.1406	-0.0660	-32.89	<.001	-1.04	0.0701	0.0718	-24.07	<.001	-0.76
150	0.8941	0.9330	-0.0389	-28.75	<.001	-0.91	0.0570	0.0580	-17.07	<.001	-0.54
200	0.7737	0.7994	-0.0256	-25.14	<.001	-0.80	0.0126	0.0127	-14.67	<.001	-0.46
250	0.6929	0.7119	-0.0190	-23.47	<.001	-0.74	0.0441	0.0446	-11.15	<.001	-0.35
300	0.6320	0.6482	-0.0162	-21.83	<.001	-0.69	0.0402	0.0407	-11.50	<.001	-0.36
350	0.5835	0.5959	-0.0124	-19.53	<.001	-0.62	0.0372	0.0375	-9.21	<.001	-0.29
400	0.5521	0.5624	-0.0104	-17.45	<.001	-0.55	0.0348	0.0350	-7.71	<.001	-0.24
450	0.5219	0.5313	-0.0094	-16.48	<.001	-0.52	0.0328	0.0330	-6.79	<.001	-0.21
500	0.4921	0.4998	-0.0076	-14.92	<.001	-0.47	0.0311	0.0314	-9.14	<.001	-0.29
1000	0.3505	0.3531	-0.0026	-7.55	<.001	-0.24	0.0220	0.0221	-6.24	<.001	-0.20
	Gamma (α = 10, β = 1)						Binomial (*n* = 10, *p* = 0.25)				
30	19.6735	22.8005	-3.1270	-30.47	<.001	-0.96	3.3809	3.8432	-32.70	<.001	-1.03
40	17.6433	20.1097	-2.4664	-29.39	<.001	-0.93	3.0257	3.3981	-31.19	<.001	-0.99
50	16.3173	18.6119	-2.2946	-24.86	<.001	-0.79	2.7294	3.0127	-30.42	<.001	-0.96
100	11.9564	13.0636	-1.1072	-24.29	<.001	-0.77	1.9890	2.1122	-27.64	<.001	-0.87
150	9.8701	10.5691	-0.6991	-27.03	<.001	-0.86	1.6385	1.7108	-27.64	<.001	-0.87
200	8.7172	9.2652	-0.5480	-21.18	<.001	-0.67	1.4212	1.4730	-25.56	<.001	-0.81
250	7.7605	8.1255	-0.3651	-26.79	<.001	-0.85	1.2749	1.3110	-22.95	<.001	-0.73
300	7.1199	7.4229	-0.3030	-17.90	<.001	-0.57	1.1694	1.1999	-22.28	<.001	-0.71
350	6.6541	6.8930	-0.2389	-23.72	<.001	-0.75	1.0764	1.1024	-19.93	<.001	-0.63
400	6.1795	6.3952	-0.2158	-22.89	<.001	-0.72	1.0130	1.0351	-19.47	<.001	-0.62
450	5.9177	6.1064	-0.1887	-20.55	<.001	-0.65	0.9567	0.9757	-18.19	<.001	-0.58
500	5.5887	5.7519	-0.1632	-20.89	<.001	-0.66	0.9040	0.9191	-15.10	<.001	-0.48
1000	3.9648	3.9950	-0.0302	-7.35	<.001	-0.23	0.6443	0.6501	-9.29	<.001	-0.29
	Poisson (λ = 5)										
30	9.3558	10.7243	-0.1885	-31.91	<.001	-1.01					
40	8.5538	9.7349	-0.1455	-28.48	<.001	-0.90					
50	7.6734	8.5870	-0.1292	-26.50	<.001	-0.84					
100	5.6033	6.0226	-0.0689	-24.16	<.001	-0.76					
150	4.6185	4.8722	-0.0454	-28.13	<.001	-0.89					
200	4.0191	4.1980	-0.0366	-26.11	<.001	-0.83					
250	3.6162	3.7513	-0.0266	-24.26	<.001	-0.77					
300	3.3212	3.4286	-0.0226	-26.07	<.001	-0.82					
350	3.0461	3.1379	-0.0182	-21.90	<.001	-0.69					
400	2.8774	2.9543	-0.0170	-20.90	<.001	-0.66					
450	2.7168	2.7840	-0.0133	-20.65	<.001	-0.65					
500	2.5618	2.6119	-0.0122	-16.76	<.001	-0.53					
1000	1.8286	1.8475	-0.0032	-10.36	<.001	-0.33					

Note: *n* = sample size, CI_width = mean width of the Wald-type (W-t) and Bias-Corrected and accelerated (BCa) confidence intervals, *md* = mean difference, *t* = test statistic, *p* = two-tailed p value, and *d* = Cohen's *d* effect size statistic. Degrees of freedom = 999. *IVAR* = Interindividual Variance. Elaborated by the author

Table 5: Comparison of Asymptotic and Bootstrap Standard Errors of *IVAR* Using Paired t Test of the Mean Log Ratio and the Pitman–Morgan Test

Dist.	*n*	*SE*		Paired t-test for mean_log_ratio				Pitman–Morgan test:			
		Asymp	Boot	m_ln	*t*	*p*	*d*	r_{SD}	*t*	*p*	*d*
N	30	0.4653	0.4738	-0.0188	-26.27	<.001	-0.83	-0.12	-3.94	<.001	-0.12
	40	0.4115	0.4173	-0.0141	-19.51	<.001	-0.62	-0.12	-3.67	<.001	-0.12
	50	0.3715	0.3756	-0.0112	-15.34	<.001	-0.49	-0.05	-1.55	0.121	-0.05
	100	0.2741	0.2754	-0.0045	-6.21	<.001	-0.20	-0.07	-2.32	0.020	-0.07
	150	0.2281	0.2288	-0.0026	-3.65	<.001	-0.12	-0.11	-3.48	0.001	-0.11
	200	0.1974	0.1977	-0.0016	-2.18	0.030	-0.07	-0.12	-3.82	<.001	-0.12
	250	0.1768	0.1771	-0.0014	-1.88	0.060	-0.06	-0.08	-2.64	0.009	-0.08
	300	0.1612	0.1615	-0.0012	-1.68	0.093	-0.05	-0.11	-3.47	0.001	-0.11

	350	0.1488	0.1489	-0.0004	-0.54	0.592	-0.02	-0.06	-1.98	0.048	-0.06
	400	0.1408	0.1410	-0.0011	-1.51	0.132	-0.05	-0.09	-3.01	0.003	-0.10
	450	0.1331	0.1332	-0.0002	-0.33	0.740	-0.01	-0.17	-5.33	<.001	-0.17
	500	0.1255	0.1256	-0.0008	-1.49	0.135	-0.05	-0.13	-4.11	<.001	-0.13
	1000	0.0894	0.0894	0.0005	0.67	0.501	0.02	-0.15	-4.88	<.001	-0.15
AS	30	0.0337	0.0356	-0.0541	-61.10	<.001	-1.93	0.01	0.46	0.647	0.01
	40	0.0289	0.0301	-0.0427	-52.70	<.001	-1.67	0.002	0.06	0.951	0.002
	50	0.0256	0.0266	-0.0349	-45.84	<.001	-1.45	-0.08	-2.65	0.008	-0.08
	100	0.0179	0.0182	-0.0199	-28.86	<.001	-0.91	-0.25	-8.09	<.001	-0.26
	150	0.0145	0.0147	-0.0101	-14.32	<.001	-0.45	-0.27	-8.75	<.001	-0.28
	200	0.0494	0.0501	-0.0101	-14.32	<.001	-0.45	-0.34	-11.55	<.001	-0.37
	250	0.0112	0.0113	-0.0072	-10.08	<.001	-0.32	-0.38	-13.16	<.001	-0.42
	300	0.0103	0.0103	-0.0056	-8.19	<.001	-0.26	-0.42	-14.70	<.001	-0.47
	350	0.0095	0.0095	-0.0047	-6.63	<.001	-0.21	-0.45	-16.05	<.001	-0.51
	400	0.0089	0.0089	-0.0041	-5.71	<.001	-0.18	-0.52	-19.32	<.001	-0.61
	450	0.0084	0.0084	-0.0031	-4.41	<.001	-0.14	-0.52	-19.31	<.001	-0.61
	500	0.0079	0.0080	-0.0037	-5.23	<.001	-0.17	-0.56	-21.43	<.001	-0.68
	1000	0.0056	0.0056	-0.0012	-1.60	0.109	-0.05	-0.72	-32.59	<.001	-1.03
G	30	5.0188	5.1020	-0.0181	-24.91	<.001	-0.79	-0.08	-2.56	0.011	-0.08
	40	4.5009	4.5552	-0.013	-16.95	<.001	-0.54	-0.01	-0.40	0.690	-0.01
	50	4.1627	4.2047	-0.0107	-15.12	<.001	-0.48	-0.07	-2.07	0.039	-0.07
	100	3.0501	3.0683	-0.006	-8.46	<.001	-0.27	-0.07	-2.29	0.022	-0.07
	150	2.5179	2.5279	-0.0041	-5.77	<.001	-0.18	0.02	0.75	0.451	0.02
	200	2.2238	2.2287	-0.0019	-2.64	0.008	-0.08	-0.07	-2.36	0.019	-0.07
	250	1.9797	1.9813	-0.0006	-0.77	0.444	-0.02	-0.07	-2.14	0.032	-0.07
	300	1.8163	1.8182	-0.0008	-1.15	0.252	-0.04	-0.06	-2.02	0.044	-0.06
	350	1.6975	1.7006	-0.0014	-2.00	0.046	-0.06	-0.11	-3.65	<.001	-0.12
	400	1.5764	1.5775	-0.0006	-0.82	0.414	-0.03	-0.03	-0.91	0.362	-0.03
	450	1.5096	1.5109	-0.0005	-0.69	0.491	-0.02	-0.12	-3.80	<.001	-0.12
	500	1.4257	1.4275	-0.0009	-1.19	0.235	-0.04	-0.13	-4.22	<.001	-0.13
	1000	1.0114	1.0112	0.0005	0.64	0.520	0.02	-0.08	-2.63	0.009	-0.08
BN	30	0.8625	0.8786	-0.0196	-26.57	<.001	-0.84	-0.08	-2.60	0.009	-0.08
	40	0.7719	0.7834	-0.0153	-21.58	<.001	-0.68	-0.10	-3.06	0.002	-0.10
	50	0.6963	0.7050	-0.0125	-17.11	<.001	-0.54	-0.13	-4.29	<.001	-0.14
	100	0.5074	0.5091	-0.0033	-4.70	<.001	-0.15	-0.08	-2.39	0.017	-0.08
	150	0.4180	0.4193	-0.0029	-3.98	<.001	-0.13	-0.10	-3.06	0.002	-0.10
	200	0.3626	0.3637	-0.0028	-3.91	<.001	-0.12	-0.14	-4.46	<.001	-0.14
	250	0.3252	0.3257	-0.0012	-1.77	0.077	-0.06	-0.10	-3.30	0.001	-0.10
	300	0.2983	0.2989	-0.0016	-2.28	0.023	-0.07	-0.11	-3.41	0.001	-0.11
	350	0.2746	0.2753	-0.0023	-3.28	0.001	-0.10	-0.09	-2.92	0.004	-0.09
	400	0.2584	0.2587	-0.0007	-1.02	0.308	-0.03	-0.12	-3.68	0.000	-0.12
	450	0.2441	0.2444	-0.0010	-1.46	0.145	-0.05	-0.11	-3.49	0.001	-0.11
	500	0.2306	0.2306	0.0002	0.22	0.823	0.01	-0.09	-2.75	0.006	-0.09
	1000	0.1644	0.1646	-0.0009	-1.32	0.188	-0.04	-0.17	-5.35	<.001	-0.17
Poi	30	2.3867	2.4302	-0.0195	-26.40	<.001	-0.84	-0.06	-2.03	0.043	-0.06
	40	2.1821	2.2110	-0.0141	-19.75	<.001	-0.63	-0.07	-2.15	0.032	-0.07
	50	1.9575	1.9794	-0.0116	-15.91	<.001	-0.50	-0.07	-2.24	0.025	-0.07
	100	1.4294	1.4345	-0.0034	-4.71	<.001	-0.15	-0.08	-2.48	0.013	-0.08
	150	1.1782	1.1814	-0.0025	-3.46	0.001	-0.11	-0.06	-2.04	0.041	-0.06
	200	1.0253	1.0283	-0.0027	-3.77	0.000	-0.12	-0.09	-3.01	0.003	-0.10
	250	0.9225	0.9245	-0.0019	-2.66	0.008	-0.08	-0.07	-2.34	0.019	-0.07
	300	0.8473	0.8490	-0.0019	-2.65	0.008	-0.08	-0.08	-2.50	0.013	-0.08
	350	0.7771	0.7790	-0.0023	-3.25	0.001	-0.10	-0.08	-2.43	0.015	-0.08
	400	0.7340	0.7351	-0.0013	-1.75	0.081	-0.06	-0.09	-2.80	0.005	-0.09
	450	0.6931	0.6935	-0.0005	-0.71	0.476	-0.02	-0.08	-2.41	0.016	-0.08
	500	0.6535	0.6538	-0.0002	-0.27	0.789	-0.01	-0.11	-3.34	0.001	-0.11
	1000	0.4665	0.4670	-0.0009	-1.35	0.178	-0.04	-0.15	-4.77	<.001	-0.15

Note: Dist. = probability distribution from which 1000 random samples were drawn: N = standard normal ($\mu = 0$, $\sigma^2 = 1$), AS = arcsine or beta ($\alpha = 0.5$, $\beta = 0.5$), G = gamma ($\alpha = 10$, $\beta = 1$), BN = binomial ($n = 10$, $p = 0.25$), and Poi = Poisson ($\lambda = 5$); n = sample size. Paired t test: *SE* = asymptotic (Asymp) and bootstrap (Boot) standard errors, m_ln = mean log ratio of asymptotic (numerator) to bootstrap (denominator) standard errors, t = test statistic, p = two-tailed p value under the t distribution with 999 degrees of freedom

for H₀: E[ln(SE_asympt / SE_boot)] = 0, and d = Cohen's d effect size. Pitman–Morgan test: r_{SD} = Pearson correlation between the sum and the difference of paired data, t = test statistic, p = two-tailed p value under the t distribution with 998 degrees of freedom for H₀: σ²(SE_asympt) = σ²(SE_boot), and d = Cohen's d effect size, which in most cases takes the same values as r_{SD}. *IVAR* = Interindividual Variance. Elaborated by the author

Table 6: Paired Comparison of Wald-type and BCa Confidence Interval Widths for *ISD*

	CI_width		Paired t-test				CI_width		Paired t-test		
n	W-type	BCa	*md*	*t*	*p*	*d*	W-type	BCa	*t*	*p*	*d*
	Normal (0, 1)						Arcsine ≡ Beta (α = 0.5, β = 0.5)				
30	0.6420	0.6958	-0.0538	-38.91	<.001	-1.23	0.1329	0.1392	-40.17	<.001	-1.27
40	0.5676	0.6065	-0.0389	-37.35	<.001	-1.18	0.1140	0.1186	-35.79	<.001	-1.13
50	0.5138	0.5433	-0.0295	-34.01	<.001	-1.08	0.1007	0.1040	-29.58	<.001	-0.94
100	0.3780	0.3924	-0.0144	-28.60	<.001	-0.90	0.0702	0.0715	-19.70	<.001	-0.62
150	0.3153	0.3241	-0.0089	-23.35	<.001	-0.74	0.0571	0.0580	-16.01	<.001	-0.51
200	0.2732	0.2790	-0.0058	-18.48	<.001	-0.58	0.0126	0.0128	-12.63	<.001	-0.40
250	0.2445	0.2489	-0.0044	-16.78	<.001	-0.53	0.0112	0.0114	-10.66	<.001	-0.34
300	0.2233	0.2273	-0.0041	-16.71	<.001	-0.53	0.0402	0.0407	-11.31	<.001	-0.36
350	0.2069	0.2095	-0.0026	-12.31	<.001	-0.39	0.0372	0.0375	-7.06	<.001	-0.22
400	0.1948	0.1975	-0.0028	-14.20	<.001	-0.45	0.0348	0.0351	-9.91	<.001	-0.31
450	0.1840	0.1862	-0.0022	-11.60	<.001	-0.37	0.0328	0.0331	-7.95	<.001	-0.25
500	0.1740	0.1760	-0.0020	-11.20	<.001	-0.35	0.0311	0.0313	-7.67	<.001	-0.24
1000	0.1238	0.1245	-0.0007	-6.07	<.001	-0.19	0.0220	0.0221	-7.55	<.001	-0.24
	Gamma (α = 10, β = 1)						Binomial (n = 10, p = 0.25)				
30	2.1704	2.3589	-0.1885	-37.56	<.001	-1.19	0.8661	0.9357	-38.02	<.001	-1.20
40	1.9400	2.0856	-0.1455	-32.14	<.001	-1.02	0.7728	0.8253	-32.99	<.001	-1.04
50	1.7980	1.9272	-0.1292	-28.71	<.001	-0.91	0.7023	0.7454	-30.25	<.001	-0.96
100	1.3264	1.3952	-0.0689	-23.27	<.001	-0.74	0.5130	0.5335	-26.44	<.001	-0.84
150	1.1000	1.1454	-0.0454	-23.72	<.001	-0.75	0.4218	0.4338	-20.45	<.001	-0.65
200	0.9698	1.0064	-0.0366	-21.44	<.001	-0.68	0.3665	0.3742	-17.96	<.001	-0.57
250	0.8668	0.8934	-0.0266	-21.50	<.001	-0.68	0.3290	0.3352	-16.00	<.001	-0.51
300	0.7965	0.8191	-0.0226	-18.90	<.001	-0.60	0.3014	0.3065	-16.19	<.001	-0.51
350	0.7411	0.7593	-0.0182	-16.67	<.001	-0.53	0.2777	0.2822	-16.10	<.001	-0.51
400	0.6901	0.7071	-0.0170	-16.35	<.001	-0.52	0.2613	0.2648	-12.73	<.001	-0.40
450	0.6592	0.6726	-0.0133	-16.61	<.001	-0.53	0.2470	0.2501	-11.75	<.001	-0.37
500	0.6231	0.6353	-0.0122	-16.72	<.001	-0.53	0.2333	0.2364	-13.36	<.001	-0.42
1000	0.4425	0.4457	-0.0032	-7.44	<.001	-0.24	0.1662	0.1669	-4.24	<.001	-0.13
	Poisson (λ = 5)										
30	1.4623	1.5842	-0.1219	-37.40	<.001	-1.18					
40	1.3289	1.4255	-0.0967	-31.28	<.001	-0.99					
50	1.2025	1.2818	-0.0793	-30.25	<.001	-0.96					
100	0.8837	0.9245	-0.0408	-24.46	<.001	-0.77					
150	0.7271	0.7511	-0.0240	-20.84	<.001	-0.66					
200	0.6342	0.6513	-0.0171	-20.43	<.001	-0.65					
250	0.5709	0.5844	-0.0136	-17.72	<.001	-0.56					
300	0.5233	0.5341	-0.0107	-17.93	<.001	-0.57					
350	0.4812	0.4904	-0.0092	-17.84	<.001	-0.56					
400	0.4541	0.4617	-0.0076	-15.14	<.001	-0.48					
450	0.4294	0.4354	-0.0059	-12.52	<.001	-0.40					
500	0.4047	0.4111	-0.0065	-14.85	<.001	-0.47					
1000	0.2888	0.2902	-0.0014	-5.02	<.001	-0.16					

Note: n = sample size, CI_width = mean width of the Wald-type (W-type) and bias-corrected and accelerated (BCa) confidence intervals, t = test statistic, p = two-tailed p value, and d = Cohen's d effect size statistic. Degrees of freedom = 999. *ISD* = Interindividual Standard Deviation. Elaborated by the author

By contrast, in the arcsine distribution, all mean differences were statistically significant across the 13 sample sizes considered, although the effect size also decreased progressively from very large to trivial as sample size increased (Table 7).

For the standard normal distribution, the variability of the bootstrap standard errors was greater than that of the asymptotic standard errors, with significant differences for all sample sizes except n = 100. The effect sizes of the estimation method on variability were trivial overall and showed no clear monotonic pattern, although they tended to increase slightly at larger sample sizes, reaching a small

magnitude at $n = 1000$. A similar pattern was observed for the binomial distribution: The variability of the bootstrap errors exceeded that of the asymptotic errors, although the differences were not significant for sample sizes between 50 and 150. Effect sizes were generally trivial, except for the sample of size 1000, where they became small. In the arcsine distribution, the variability of the bootstrap errors was also greater in almost all sample sizes; differences were not significant only for $n = 30$ and 50. In this distribution, the effect sizes increased progressively with sample size, shifting from trivial to moderate and finally to large magnitudes at the highest sample sizes. In the Poisson distribution, the variability of the bootstrap errors was likewise greater than that of the asymptotic errors, although more than half of the comparisons (7 of 13) were not significant. Effect sizes were trivial overall, except for the sample of size 30, where they reached a small magnitude. Finally, in the gamma distribution, most comparisons (9 of 13) were nonsignificant. When significant differences emerged, the variability of the bootstrap standard errors was greater than that of the asymptotic standard errors, with trivial effect sizes except for the sample of size 30, where the effect size was small (Table 7).

Convergence Toward Normality in the Sampling Distributions of Three Interindividual Variation Statistics

The *ISD* statistic showed a similar pattern, with slightly faster convergence toward normality than *MID*. A normality fit in at least 80% of the samples was achieved with sample sizes of 40 in the normal and binomial distributions, 50 in the Poisson distribution, 100 in the gamma distribution, and 200 in the arcsine distribution (Table 8).

Table 7: Comparison of Asymptotic and Bootstrap Standard Errors of *ISD* Using Paired t Test of the Mean Log Ratio and the Pitman–Morgan Test

Dist.	*n*	*SE*		Paired t-test for mean_log_ratio				Pitman–Morgan test			
		Asymp	Boot	m_ln	*t*	*p*	*d*	r_{SD}	*t*	*p*	*d*
N	30	0.1638	0.1707	-0.0427	-55.57	<.001	-1.76	-0.12	-3.80	<.001	-0.12
	40	0.1448	0.1492	-0.0304	-40.88	<.001	-1.29	-0.07	-2.37	0.018	-0.08
	50	0.1311	0.1340	-0.0226	-31.15	<.001	-0.99	-0.07	-2.25	0.025	-0.07
	100	0.0964	0.0975	-0.0106	-14.82	<.001	-0.47	-0.02	-0.48	0.631	-0.02
	150	0.0804	0.0809	-0.006	-8.41	<.001	-0.27	-0.08	-2.67	0.008	-0.08
	200	0.0697	0.0700	-0.0043	-6.07	<.001	-0.19	-0.12	-3.67	<.001	-0.12
	250	0.0624	0.0625	-0.0025	-3.57	<.001	-0.11	-0.15	-4.93	<.001	-0.16
	300	0.0570	0.0572	-0.0044	-6.35	<.001	-0.20	-0.14	-4.31	<.001	-0.14
	350	0.0528	0.0529	-0.0014	-2.05	0.040	-0.06	-0.11	-3.54	<.001	-0.11
	400	0.0497	0.0498	-0.0023	-3.38	0.001	-0.11	-0.13	-4.20	<.001	-0.13
	450	0.0469	0.0470	-0.0012	-1.65	0.100	-0.05	-0.15	-4.84	<.001	-0.15
	500	0.0444	0.0445	-0.0017	-2.45	0.014	-0.08	-0.15	-4.80	<.001	-0.15
	1000	0.0316	0.0316	0.0002	0.2551	0.799	0.01	-0.22	-7.21	<.001	-0.23
AS	30	0.0339	0.0369	-0.0853	-89.67	<.001	-2.84	-0.06	-1.82	0.069	-0.06
	40	0.0291	0.0310	-0.0652	-82.76	<.001	-2.62	-0.07	-2.31	0.021	-0.07
	50	0.0257	0.0271	-0.0526	-66.75	<.001	-2.11	-0.05	-1.57	0.116	-0.05
	100	0.0179	0.0184	-0.0262	-36.92	<.001	-1.17	-0.09	-2.86	0.004	-0.09
	150	0.0146	0.0148	-0.0131	-18.70	<.001	-0.59	-0.18	-5.85	<.001	-0.19
	200	0.0494	0.0500	-0.0131	-18.70	<.001	-0.59	-0.24	-7.91	<.001	-0.25
	250	0.0441	0.0445	-0.01	-13.87	<.001	-0.44	-0.34	-11.34	<.001	-0.36
	300	0.0103	0.0104	-0.0095	-13.53	<.001	-0.43	-0.33	-10.87	<.001	-0.34
	350	0.0095	0.0096	-0.007	-9.88	<.001	-0.31	-0.36	-12.16	<.001	-0.38
	400	0.0089	0.0089	-0.0073	-9.92	<.001	-0.31	-0.39	-13.32	<.001	-0.42
	450	0.0084	0.0084	-0.0051	-7.26	<.001	-0.23	-0.41	-14.16	<.001	-0.45
	500	0.0079	0.0080	-0.0051	-7.37	<.001	-0.23	-0.43	-15.18	<.001	-0.48
	1000	0.0056	0.0056	-0.0026	-3.76	<.001	-0.12	-0.58	-22.36	<.001	-0.71
G	30	0.5537	0.5780	-0.0447	-58.88	<.001	-1.86	-0.247	-8.04	<.001	-0.25
	40	0.4949	0.5088	-0.0291	-37.76	<.001	-1.19	-0.024	-0.76	0.446	-0.02
	50	0.4587	0.4682	-0.0216	-28.41	<.001	-0.90	-0.005	-0.15	0.885	0.00
	100	0.3384	0.3409	-0.0077	-11.07	<.001	-0.35	0.044	1.38	0.169	0.04
	150	0.2806	0.2822	-0.0057	-7.64	<.001	-0.24	0.058	1.84	0.066	0.06
	200	0.2474	0.2482	-0.0036	-5.08	<.001	-0.16	0.049	1.55	0.121	0.05
	250	0.2211	0.2219	-0.0034	-4.71	<.001	-0.15	0.003	0.10	0.924	0.00
	300	0.2032	0.2036	-0.0022	-3.10	0.002	-0.10	0.004	0.14	0.890	0.00
	350	0.1891	0.1894	-0.0019	-2.64	0.008	-0.08	0.005	0.15	0.880	0.00
	400	0.1761	0.1764	-0.0017	-2.32	0.020	-0.07	-0.106	-3.38	0.001	-0.11

	450	0.1682	0.1682	-0.0003	-0.48	0.633	-0.02	0.012	0.36	0.717	0.01
	500	0.1589	0.1590	-0.0004	-0.49	0.626	-0.02	-0.069	-2.20	0.028	-0.07
	1000	0.1129	0.1130	-0.0005	-0.73	0.466	-0.02	-0.150	-4.79	<.001	-0.15
BN	30	0.2209	0.2308	-0.0449	-55.968	<.001	-1.77	-0.19	-6.01	<.001	-0.19
	40	0.1972	0.2031	-0.0304	-40.31	<.001	-1.28	-0.07	-2.08	0.038	-0.07
	50	0.1792	0.1834	-0.0239	-33.42	<.001	-1.06	-0.03	-0.97	0.334	-0.03
	100	0.1309	0.1323	-0.0111	-15.23	<.001	-0.48	-0.04	-1.32	0.189	-0.04
	150	0.1076	0.1083	-0.0067	-9.64	<.001	-0.31	-0.04	-1.32	0.186	-0.04
	200	0.0935	0.0939	-0.0039	-5.53	<.001	-0.17	-0.08	-2.39	0.017	-0.08
	250	0.0839	0.0841	-0.0022	-2.98	0.003	-0.09	-0.11	-3.58	<.001	-0.11
	300	0.0769	0.0772	-0.0037	-5.22	<.001	-0.17	-0.12	-3.90	<.001	-0.12
	350	0.0708	0.0711	-0.003	-4.22	<.001	-0.13	-0.09	-2.78	0.006	-0.09
	400	0.0667	0.0668	-0.0015	-2.10	0.036	-0.07	-0.18	-5.83	<.001	-0.18
	450	0.063	0.0631	-0.0004	-0.61	0.541	-0.02	-0.13	-4.00	<.001	-0.13
	500	0.0595	0.0597	-0.0022	-3.13	0.002	-0.10	-0.17	-5.49	<.001	-0.17
	1000	0.0424	0.0424	0.0001	0.18	0.855	0.01	-0.20	-6.61	<.001	-0.21
Poi	30	0.3731	0.3894	-0.0441	-56.07	<.001	-1.77	-0.24	-7.85	<.001	-0.25
	40	0.3390	0.3485	-0.0290	-38.41	<.001	-1.22	-0.03	-0.82	0.415	-0.03
	50	0.3068	0.3135	-0.0224	-30.64	<.001	-0.97	-0.02	-0.72	0.473	-0.02
	100	0.2254	0.2277	-0.0103	-14.09	<.001	-0.45	-0.03	-1.07	0.285	-0.03
	150	0.1855	0.1865	-0.0055	-7.68	<.001	-0.24	-0.01	-0.22	0.825	-0.01
	200	0.1618	0.1624	-0.0038	-5.39	<.001	-0.17	-0.03	-0.93	0.355	-0.03
	250	0.1456	0.1459	-0.0014	-1.97	0.049	-0.06	-0.07	-2.32	0.021	-0.07
	300	0.1335	0.1340	-0.0034	-4.82	<.001	-0.15	-0.11	-3.46	0.001	-0.11
	350	0.1228	0.1231	-0.0027	-3.87	<.001	-0.12	-0.02	-0.54	0.590	-0.02
	400	0.1158	0.116	-0.0013	-1.82	0.069	-0.06	-0.13	-4.20	<.001	-0.13
	450	0.1095	0.1096	-0.0001	-0.19	0.850	-0.01	-0.04	-1.15	0.250	-0.04
	500	0.1032	0.1035	-0.0021	-2.94	0.003	-0.09	-0.11	-3.62	<.001	-0.11
	1000	0.0737	0.0737	0.0003	0.45	0.652	0.01	-0.15	-4.68	<.001	-0.15

Note: Dist. = probability distribution from which 1000 random samples were drawn: N = standard normal ($\mu = 0$, $\sigma^2 = 1$); AS = arcsine or beta ($\alpha = 0.5$, $\beta = 0.5$); G = gamma ($\alpha = 10$, $\beta = 1$); BN = binomial ($n = 10$, $p = 0.5$); Poi = Poisson ($\lambda = 5$); n = sample size. Paired t test: *SE* = asymptotic (Asymp) and bootstrap (Boot) standard errors; m_ln = mean log ratio of asymptotic (numerator) to bootstrap (denominator) standard errors; t = test statistic; p = two-tailed p value under the t distribution with 999 degrees of freedom for H_0: E[ln(SE_asympt / SE_boot)] = 0; d = Cohen's d effect size. Pitman–Morgan test: r_{SD} (= d) = Pearson correlation between the sum and the difference of paired data, which is algebraically equivalent to Cohen's d effect size; t = test statistic; p = two-tailed p value under the t distribution with 998 degrees of freedom for H_0: σ^2(*SE*_asympt) = σ^2(*SE*_boot). *ISD* = Interindividual Standard Deviation. Elaborated by the author

Table 8: Percentage of samples for which the null hypothesis of normality is NOT rejected at $\alpha = 0.05$ for five distributions using the Shapiro-Francia W' test

Sample size	Normal (0, 1)			Arcsine (0, 1)			Gamma ($\alpha = 10$, $\beta = 1$)		
	MID	*IVAR*	*ISD*	*MID*	*IVAR*	*ISD*	*MID*	*IVAR*	*ISD*
30	71.7	16.3	74.5	1.7	68.2	14.8	65.3	20.4	62.9
40	79.5	17.3	85.9	2.4	78.3	26	71.1	22.6	73.4
50	84.1	18.6	88.9	2.9	84.5	36.1	75.3	20.3	75.8
100	90.1	32.6	89.5	14.4	91.9	65.8	80.9	25.2	83.8
150	90.4	41	91.4	31	93.8	78.8	83.7	34.5	86.2
200	92	50.2	91.2	40.2	93	83.7	85.6	36.7	84.2
250	91	57.7	90.8	56.8	94.3	89.4	88.4	40	85.1
300	91.8	61	92.2	62.9	94.9	90.4	90.6	44.4	85.7
350	93.1	65.8	92.5	68.5	95.4	89.4	89.5	51.1	88
400	93.8	69.5	93.6	68.7	93.9	89.7	91.1	53.1	85.6
450	93.6	71.3	93.1	73.6	94.3	92.3	90.7	53.9	88.9
500	92.3	74.3	94.2	75.9	94.6	91.6	91.1	57.1	87.2
1000	94	83.8	93.4	85.3	94.4	93.6	93.5	74.5	90.4
	Binomial (n =10, p = 0.25)			Poisson ($\lambda = 5$)					
30	66.3	19.4	69.8	66.6	20.8	70.7			
40	77.3	20.8	80.4	72.8	21	78			
50	82.7	22.7	82.6	80.5	21.4	81.9			
100	86.9	34.9	87.6	84.2	31.1	86.1			
150	91.4	45.1	89.2	87.3	37.9	86.9			

200	92.8	54.4	90.7	90.2	47	88
250	92.7	58.7	92.1	92	50.5	88.1
300	93.4	64	92.5	90.5	54.8	91.3
350	95.8	67.2	92.1	93.8	59.2	90.3
400	94.2	70.4	93.4	93.5	62.4	90.1
450	94.1	70.9	92.9	91.6	60.2	91.4
500	94.5	75.5	92.6	94.5	67.9	90.9
1000	95.5	82.8	94.8	94.9	78	92.4

Note: *MID* = Mean Interindividual Difference, *IVAR* = Interindividual Variance, and *ISD* = Interindividual Standard Deviation. Elaborated by the author

The *IVAR* statistic also converged progressively toward normality across the five distributions, although the convergence pattern differed from that observed for *MID* and *ISD*. In this case, convergence was fastest in the arcsine distribution and slowest in the normal, gamma, binomial, and Poisson distributions. At least 80% of the samples drawn from the arcsine distribution showed a normal fit with sample sizes of 50. By contrast, this threshold was reached only with samples of size 1000 in the normal and binomial distributions. Even with n = 1000, the percentages remained below 80% in the Poisson (78%) and gamma (74.5%) distributions, indicating comparatively slower convergence toward normality for *IVAR* in these distributions (Table 8).

Discussion

The first objective was to evaluate the equivalence between the amplitudes of Wald-type and BCa intervals, as well as between asymptotic and bootstrap standard errors of the absolute interindividual variability statistics. The results do not support overall equivalence, although discrepancies depend on the statistic, sample size, and underlying distribution.

With *MID*, the point estimates were stable and very similar across sample sizes within each distribution, with no monotonic trend, suggesting the absence of systematic sample-size bias. Regarding interval width, a general pattern of convergence between methods was observed as sample size increased. In the normal, gamma, binomial, and Poisson distributions, BCa intervals tended to be slightly wider than Wald-type intervals in small samples, a difference that became trivial in large samples. In the arcsine distribution, the opposite pattern was observed in small samples, although also with progressive attenuation. This differential behavior may be related to the U-shaped form of this distribution and the concentration of density at the boundaries of the support. Concerning standard error magnitude, the bootstrap method produced lower average values than the asymptotic method in most distributions, with the relationship reversed in the arcsine distribution. In all cases, effect size decreased as sample size increased. Finally, sample-to-sample variability of standard errors showed a pattern dependent on sampling regime: Greater variability of the asymptotic method in small samples, non-significant differences in intermediate sizes, and in several cases lower asymptotic variability in large samples.

From an applied perspective, these findings suggest that *MID* may provide a practically robust measure of variability across a broad range of distributional conditions. In moderate to large samples, asymptotic and bootstrap procedures yield highly similar results, which supports the use of simpler asymptotic inference when computational efficiency is required. In smaller samples or markedly non-normal distributions, however, bootstrap procedures may still provide a more conservative alternative.

With *IVAR*, there is no equivalence in interval width and error in small samples, where the BCa method systematically produces wider confidence intervals than Wald-type intervals and bootstrap standard errors are systematically larger than asymptotic ones. Practical equivalence is reached in moderate or large samples for approximately regular distributions (normal and gamma), and also for discrete ones (binomial and Poisson), as differences become small or non-significant. The arcsine distribution is a partial exception, since discrepancies persist and bootstrap variability increases with sample size, suggesting greater sensitivity of resampling procedures to strong asymmetry and boundary constraints.

These results have practical implications for applied analyses involving heterogeneous or asymmetric data. Since *IVAR* appears more sensitive to distributional shape, bootstrap procedures may be preferable when sample sizes are limited or when strong skewness and boundary effects are expected. Conversely, asymptotic inference appears sufficiently stable under more regular distributional conditions and moderate-to-large samples.

With *ISD*, there is no strict equivalence in interval width and error in small samples, where the bootstrap produces wider BCa intervals and larger standard errors than the asymptotic method. Nevertheless, practical equivalence exists in moderate or large samples for approximately regular distributions (normal and gamma), and also for discrete ones (binomial and Poisson), as differences become trivial or non-significant. The arcsine distribution is the main exception, since the discrepancy persists even in large sample sizes, especially regarding

standard error variability, suggesting increased bootstrap sensitivity to strong skewness and support restrictions. The practical implications identified for *IVAR* also apply to *ISD*, with the additional advantage that *ISD* is expressed on the same scale as the original variable, whereas *IVAR* is expressed on a quadratic scale.

With respect to the second objective, analyzing convergence toward normality of the interindividual variation statistics, the results indicate that *MID* and *ISD* show clear and consistent convergence toward normality. In approximately 80% of samples, the fit to normality is reached with moderate sample sizes (between 40 and 100) in most distributions analyzed. Convergence is slower in the arcsine distribution, characterized by pronounced non-normality and concentration of mass at the support boundaries.

This pattern suggests that *MID* and *ISD* behave as regular statistics, in the sense that they satisfy functional differentiability conditions that allow application of the delta method. That is, they may be considered differentiable functionals of the population distribution, which ensures asymptotic normality under general conditions.

In the case of *IVAR*, convergence toward normality is more heterogeneous. It is relatively fast under the arcsine distribution, slower in the normal and binomial distributions, and even slower in the Poisson and gamma distributions. This behavior indicates that *IVAR* is more sensitive to population shape than *MID* and *ISD*. Its sampling distribution appears to depend not only on sample size but also on skewness, discreteness, and higher-order moments of the underlying distribution. Although it shows a trend toward normality, its rate of convergence is clearly distribution-dependent.

Beyond their inferential properties, measures based on interindividual differences may provide a complementary perspective to classical dispersion measures. Whereas variance and standard deviation quantify variability through deviations from a central location parameter, *MID* and related statistics quantify variability directly through pairwise differences among observations. This formulation may be especially useful in contexts where interpersonal heterogeneity or dispersion among individuals constitutes the substantive focus of interest. Consequently, these measures may complement traditional variability indices in fields concerned with intraindividual variability, behavioral heterogeneity, psychometrics, or population diversity.

Overall, the results support the practical applicability of *MID*, *IVAR*, and *ISD* as inferential measures of variability under a broad range of sampling conditions, while also highlighting the importance of considering sample size and distributional shape when selecting between asymptotic and bootstrap procedures.

As limitations of the study, it should be noted that only 13 sample sizes were considered. The smallest was 30, which is the minimum generally recommended for applying the bootstrap method (Efron and Narasimhan, 2020), and the largest was 1000, a conventional lower bound for large-sample studies (Muñoz and Christen, 2022). Likewise, only five well-known univariate probability distributions were examined: Three continuous (normal, arcsine, and gamma) and two discrete (binomial and Poisson). All have finite moments and differ in boundedness, skewness, and kurtosis. To reduce the computational burden, 1000 Monte Carlo replications were performed per condition (5 × 13), and 1000 resamples with replacement were generated for the bootstrap procedure in each simulated sample, which corresponds to the minimum commonly recommended (Efron and Narasimhan, 2025; Mokhtar and Sapiri, 2023). Nevertheless, greater stability of the results would likely be achieved with 10,000 Monte Carlo replications and 2,000 bootstrap resamples (Dunn and Shultis, 2022; Rousselet and Wilcox, 2023).

Conclusion

For *MID*, differences between estimation methods in confidence interval width, standard error magnitude, and standard error variability progressively decrease as sample size increases, indicating practical convergence in large samples. In small samples, however, both the magnitude and direction of discrepancies depend on distributional shape, with the arcsine distribution exhibiting the most distinctive pattern. These findings suggest that *MID* may provide a practically stable measure of variability across a broad range of conditions, particularly when moderate or large sample sizes are available.

For *IVAR*, the asymptotic method tends to slightly underestimate uncertainty in small samples. Nevertheless, under regular conditions and with sufficiently large sample sizes, Wald and BCa confidence intervals, as well as asymptotic and bootstrap standard errors, show practical equivalence. Departures from equivalence are driven mainly by sample size and degree of skewness rather than by the discrete or continuous nature of the distribution. Accordingly, bootstrap procedures may be preferable when analyzing highly skewed data or relatively small samples.

A similar pattern is observed for *ISD*. The asymptotic approach slightly underestimates uncertainty in small samples, but under regular conditions and adequate sample sizes, Wald and BCa intervals and both types of standard errors become practically equivalent. Again, discrepancies are primarily associated with sample size and skewness. These results support the practical applicability of *ISD* as a stable inferential measure of variability under a wide range of sampling conditions.

Regarding distributional behavior, the bootstrap-

generated sampling distributions of *MID*, *IVAR*, and *ISD* show convergence toward normality. *MID* and *ISD* exhibit clear and robust convergence, consistent with their properties as regular and differentiable functionals. *IVAR* also converges toward normality, although its rate of convergence is more sensitive to the shape of the population distribution. Overall, the regularity properties required to justify the application of asymptotic methods to these three statistics are empirically supported.

Beyond their inferential properties, measures based on interindividual differences may provide a complementary perspective to classical variability indices. Whereas variance and standard deviation quantify variability through deviations from a central value, *MID* and related statistics quantify variability directly through pairwise differences among observations. This characteristic may be particularly useful in contexts where interpersonal heterogeneity or relative dispersion between individuals is substantively relevant. Consequently, these measures may complement traditional indices of variability in fields such as psychology, behavioral sciences, education, and population studies.

Suggestions

Future research should continue examining the inferential behavior of these statistics through simulation studies involving additional probability distributions, broader distributional irregularities, and empirical samples derived from real-world data. Further studies should also evaluate the robustness of these measures under stronger departures from normality, smaller sample sizes, and more extreme forms of skewness or discreteness.

In addition, the application of these statistics to ordinal variables deserves further investigation, particularly in contexts where pairwise interindividual differences may provide substantively meaningful information about variability or heterogeneity.

Future applied studies should also compare the practical interpretability and usefulness of interindividual variation measures relative to traditional variability indices in empirical settings. Such comparisons may help clarify the specific contexts in which measures based on pairwise differences provide complementary information beyond classical variance-based approaches.

Given their favorable inferential behavior under many conditions examined in the present study, the use of *MID*, *IVAR*, and *ISD* as descriptive indicators of variability in empirical quantitative research is recommended, particularly when the focus of interest involves heterogeneity among individuals or relative distances between observations.

Acknowledgment

The author thanks the reviewers and the editor for their valuable comments and suggestions, which substantially improved the quality of the manuscript.

Funding Information

The study was financed entirely by the author's own resources.

Ethics

Since the study was based on simulated data without specific content and did not involve samples or empirical experimental methods, no ethical issues were incurred. The author declares that there are no conflicts of interest.

References

Alabi, O., & Bukola, T. (2023). Introduction to Descriptive statistics. *Recent Advances in Biostatistics*, 1–12. https://doi.org/10.5772/intechopen.1002475

Champely, S. (2018). Pitman-Morgan test of variance for paired samples. In *PairedData: Paired Data Analysis*. https://www.rdocumentation.org/packages/PairedData/versions/0.9.1/topics/pitman.morgan.test

Chen, L., Zhernakova, D. V., Kurilshikov, A., Andreu-Sánchez, S., Wang, D., Augustijn, H. E., Vich Vila, A., Weersma, R. K., Medema, M. H., Netea, M. G., Kuipers, F., Wijmenga, C., Zhernakova, A., & Fu, J. (2022). Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome. *Nature Medicine*, *28*(11), 2333–2343. https://doi.org/10.1038/s41591-022-02014-8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Dunn, W. L., & Shultis, J. K. (2022). *Exploring Monte Carlo methods* (2nd ed., Vol. 15). Elsevier. https://doi.org/10.1016/C2009-0-16850-2

Efron, B., & Narasimhan, B. (2020). The Automatic Construction of Bootstrap Confidence Intervals. *Journal of Computational and Graphical Statistics*, *29*(3), 608–619. https://doi.org/10.1080/10618600.2020.1714633

Efron, B., & Narasimhan, B. (2025). Packages bcaboot. Bias corrected bootstrap confidence intervals. *Comprehensive R Archive Network (CRAN)*.

Hoeffding, W. (1948). *A Class of Statistics with Asymptotically Normal Distribution*. The Annals of Mathematical Statistics. https://doi.org/10.1214/aoms/1177730196

Hyun, J., Qin, J., Wang, C., Katz, M. J., Pavlovic, J. M., Derby, C. A., & Lipton, R. B. (2022). Reliabilities of Intra-Individual Mean and Intra-Individual Variability of Self-Reported Pain Derived from Ecological Momentary Assessments: Results from the Einstein Aging Study. *The Journal of Pain*, *23*(4), 616–624. https://doi.org/10.1016/j.jpain.2021.10.008

Kohoutová, L., Atlas, L. Y., Büchel, C., Buhle, J. T., Geuter, S., Jepma, M., Koban, L., Krishnan, A., Lee, D. H., Lee, S., Roy, M., Schafer, S. M., Schmidt, L., Wager, T. D., & Woo, C.-W. (2022). Individual variability in brain representations of pain. *Nature Neuroscience*, *25*(6), 749–759. https://doi.org/10.1038/s41593-022-01081-x

Korolyuk, V. S., & Borovskich, Y. V. (2013). *Theory of U-statistics (Mathematics and Its Applications*. *273*.

Mokhtar, S. F., Md Yusof, Z., & Sapiri, H. (2023). Confidence Intervals by Bootstrapping Approach: A Significance Review. *Malaysian Journal of Fundamental and Applied Sciences*, *19*(1), 30–42. https://doi.org/10.11113/mjfas.v19n1.2660

Moral-de la Rubia (2026a). Measures of interindividual variation: Definition and computation. *Open Journal of Statistics, 16*(2), 47-88. https://doi.org/10.4236/ojs.2026.162004

Moral-de la Rubia J. (2026b). A simulation study on statistical properties the Interindividual Coefficient of Variation. *Journal of Mathematical Problems, Equations, and Statistics., 7*(1), 175-185. https://www.doi.org/10.22271/math.2026.v7.i1b.329

Moral-de la Rubia, J. (2026c). On the Regularity and Sampling Behavior of the Median Interindividual Difference. *Journal of Statistics and Applied Sciences*, *13*, 32-45. https://doi.org/10.52693/jsas.1900569

Morgan, W. A. (1939). Test for the significance of the difference between the two variances in a sample from a normal bivariate population. *Biometrika*, *31*(1–2), 13–19. https://doi.org/10.1093/biomet/31.1-2.13

Muñoz, J. D. M., & Christen, J. A. (2022). Criterion to determine the sample size in stochastic simulation processes. *Ingeniería y Universidad*, *26*(9), 1–21. https://doi.org/10.11144/javeriana.iued26.cdss

Pitman, E. J. G. (1939). A note on normal correlation. *Biometrika*, *31*(1/2), 9–12. https://doi.org/10.2307/2334971

R Core Team. (2025). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Raudenbush, S. W. (1988). Estimating Change in Dispersion. *Journal of Educational Statistics*, *13*(2), 148–171. https://doi.org/10.3102/10769986013002148

Rousselet, G., Pernet, C. R., & Wilcox, R. R. (2023). An introduction to the bootstrap: a versatile method to make inferences by using data-driven simulations. *Meta-Psychology*, *7*, 2058. https://doi.org/10.15626/mp.2019.2058

Royston, P. (1993). A Toolkit for Testing for Non-Normality in Complete and Censored Samples [dataset]. In *The Statistician* (Vol. 42, Issue 1, pp. 37–43). https://doi.org/10.2307/2348109

Shapiro, S. S., & Francia, R. S. (1972). An Approximate Analysis of Variance Test for Normality [dataset]. In *Journal of the American Statistical Association* (Vol. 67, Issue 337, pp. 215–216). https://doi.org/10.1080/01621459.1972.10481232

Shryock, I., Condon, D. M., & Weston, S. J. (2025). The stability of intraindividual means and standard deviations of affect. In *Personality Science* (Vol. 6). https://doi.org/10.1177/27000710251318301

Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data [dataset]. In *Infant Behavior and Development* (Vol. 25, Issue 4, pp. 340–374). https://doi.org/10.1016/s0163-6383(02)00140-6

Wendler, M. (2011). Bahadur representation for U-quantiles of dependent data [dataset]. In *Journal of Multivariate Analysis* (Vol. 102, Issue 6, pp. 1064–1079). https://doi.org/10.1016/j.jmva.2011.02.005