Biometrika (2024), **111**, 3, *pp.* 1063–1070 https://doi.org/10.1093/biomet/asae004
Advance Access publication 4 February 2024

On the failure of the bootstrap for Chatterjee's rank correlation

By ZHEXIAO LIN

Department of Statistics, University of California, Berkeley, 367 Evans Hall, Berkeley, California 94720, U.S.A.
zhexiaolin@berkeley.edu

and FANG HAN

Department of Statistics, University of Washington, B-317 Padelford Hall, Seattle, Washington 98195, U.S.A.
fanghan@uw.edu

Summary

While researchers commonly use the bootstrap to quantify the uncertainty of an estimator, it has been noticed that the standard bootstrap, in general, does not work for Chatterjee's rank correlation. In this paper, we provide proof of this issue under an additional independence assumption, and complement our theory with simulation evidence for general settings. Chatterjee's rank correlation thus falls into a category of statistics that are asymptotically normal, but bootstrap inconsistent. Valid inferential methods in this case are Chatterjee's original proposal for testing independence and the analytic asymptotic variance estimator of Lin & Han (2022) for more general purposes.

Some key words: Bootstrap; Rank correlation; Tied data.

1. Introduction

Rank correlation is an essential tool for measuring the association between random variables. Its development is closely linked to the history of statistics as a discipline and has involved many notable figures, including Spearman (1904, 1906), Kendall (1938, 1970), Hoeffding (1940, 1948, 1994), Hodges & Lehmann (1956), Chernoff & Savage (1958), Blum et al. (1961) and Sidak et al. (1999). Unlike other correlation coefficients, a rank correlation relies solely on the rankings of the original data, making it (i) exactly distribution-free when testing independence of continuous random variables; (ii) invariant to marginal monotonic transformations and (iii) robust in the face of outliers and heavy tailedness. Its usefulness is therefore self-explanatory.

Given the remarkable progress made in this area over the past century, it is impressive that, recently, Chatterjee (2021) devised a new rank correlation that is appealing from multiple perspectives. Specifically, consider an independent and identically distributed sample $\{X_i, Y_i\}_{i=1,\ldots,n}$ from a pair of scalars, (X, Y), with joint and marginal distribution functions $F_{X,Y}$ and F_X, F_Y, respectively. Let

$$R_i \equiv \sum_{j=1}^{n} \mathbb{1}(Y_j \leqslant Y_i) \quad \text{and} \quad L_i \equiv \sum_{j=1}^{n} \mathbb{1}(Y_j \geqslant Y_i)$$

be the rank and reversed rank of Y_i with $\mathbb{1}(\cdot)$ representing the indicator function, and let $\{[i], i = 1, \ldots, n\}$ be a rearrangement of $\{1, \ldots, n\}$ such that $X_{[1]} \leqslant \cdots \leqslant X_{[n]}$ with ties broken at random.

Chatterjee (2021) introduced the statistic

$$\xi_n \equiv 1 - \frac{n}{2\sum_{i=1}^{n} L_i(n - L_i)} \sum_{i=1}^{n-1} |R_{[i+1]} - R_{[i]}|, \quad (1)$$

which he showed to be a strongly consistent estimator of Dette–Siburg–Stoimenov's dependence measure (Dette et al., 2013),

$$\xi = \xi(X, Y) \equiv \frac{\int \mathrm{var}[E\{\mathbb{1}(Y \geqslant y) \mid X\}] \, \mathrm{d}F_Y(y)}{\int \mathrm{var}\{\mathbb{1}(Y \geqslant y)\} \, \mathrm{d}F_Y(y)}, \quad (2)$$

as long as Y is not almost surely a constant.

Why is ξ_n in (1) appealing? Chatterjee (2021) outlined three reasons. First, it has a simple form. Second, it has a normal limiting null distribution. Finally, it measures a dependence measure, ξ in (2), that satisfies Rényi's criteria (Rényi, 1959) and Bickel's definition of a measure of functional dependence (Bickel, 2022): ξ is zero if and only if Y is independent of X and one if and only if Y is a measurable function of X. Therefore, ξ_n is a rank correlation that can accurately quantify both independence and functional dependence. This is something that all aforementioned rank correlations, including Spearman's ρ, Kendall's τ, Hoeffding's D, Blum–Kiefer–Rosenblatt's r and Bergsma–Dassios–Yanagimoto's τ^* (Bergsma & Dassios, 2014; Yanagimoto, 1970) rank correlations, fail to achieve.

Because of these appealing properties, Chatterjee's rank correlation has gained significant interest and a wave of research has emerged exploring its applications and extensions. Notable recent works include Cao & Bickel (2020), Deb et al. (2020), Azadkia & Chatterjee (2021), Azadkia et al. (2022), Bickel (2022), Gamboa et al. (2022), Griessenberger et al. (2022), Han & Huang (2022), Huang et al. (2022), Lin & Han (2022, 2023), Shi et al. (2022, 2024), Ansari & Fuchs (2023), Chatterjee & Vidyasagar (2023), Zhang (2023a,b), Auddy et al. (2024), Fuchs (2024), Strothmann et al. (2024) and the 2022 thesis by A. Holma form Umeå University. Additionally, brief surveys on recent progress of Chatterjee's and other rank correlation methods have been conducted by Han (2021) and Chatterjee (2024).

This paper aims to investigate the validity of the standard bootstrap (Efron, 1979, 1981) when applied to fixed and continuous $F_{X,Y}$, with ξ_n taking the form (1) to handle ties in resampled data. We prove, in the simple independence case with $F_{X,Y} = F_X F_Y$, that the standard bootstrap results in an inconsistent estimator of ξ_n's asymptotic variance, and the bootstrap distribution fails to converge to the limiting distribution of $n^{1/2}(\xi_n - \xi)$. Simulations further complement the theory, indicating that the standard bootstrap will likely also fail in general settings with $F_{X,Y} \neq F_X F_Y$, even though $n^{1/2}(\xi_n - \xi)$ still weakly converges to a normal distribution (Lin & Han, 2022). Chatterjee's rank correlation thus falls into a class of statistics that are root-n consistent, asymptotically normal, but bootstrap inconsistent; this class includes Bickel and Freedman's U-statistics (Bickel & Freedman, 1981, § 6), Hodges estimator (Beran, 1982, pp. 213–4) and Abadie and Imbens's matching estimator (Abadie & Imbens, 2008).

There are valid alternatives to using the standard bootstrap for inferring ξ from ξ_n. Chatterjee (2021) derived the limiting null distribution of ξ_n for testing independence between X and Y. Lin & Han (2022) proved conditions under which ξ_n is root-n consistent and asymptotically normal, and proposed an analytic estimator of its asymptotic variance. This paper thus helps to justify the value of these derivations by demonstrating the inconsistency of an otherwise attractive alternative.

2. Main results

2.1. *Set-up*

In this paper we consider the standard model of $(X_1, Y_1), \ldots, (X_n, Y_n)$ to be n independent copies of (X, Y) drawn from a fixed and continuous $F_{X,Y}$ of support in the two-dimensional real space. In this case, with probability one, there is no tie in the observation, and hence ξ_n admits the simpler form

$$\xi_n \equiv 1 - \frac{3}{n^2 - 1} \sum_{i=1}^{n-1} |R_{[i+1]} - R_{[i]}|.$$

To implement the standard bootstrap, consider $(\vec{X}_b, \vec{Y}_b) = \{(X_{b,i}, Y_{b,i})\}_{i=1}^n$ to be the bootstrap sample, of size n and likely embracing ties, by sampling with replacement from $(\vec{X}, \vec{Y}) \equiv \{(X_i, Y_i)\}_{i=1}^n$. Let $\tilde{\xi}_b$ be Chatterjee's rank correlation calculated using the bootstrap sample $(\vec{X}_b, \vec{Y}_b)$. More specifically, write

$$\{[i]_b,\ i = 1, \ldots, n\}$$

to be a rearrangement of $\{1, \ldots, n\}$ such that $X_{b,[1]_b} \leqslant \cdots \leqslant X_{b,[n]_b}$, with ties in $\{X_{b,i}\}_{i=1}^n$ broken in an arbitrary way. Let

$$R_{b,i} \equiv \sum_{j=1}^{n} \mathbb{1}(Y_{b,j} \leqslant Y_{b,i}) \quad \text{and} \quad L_{b,i} \equiv \sum_{j=1}^{n} \mathbb{1}(Y_{b,j} \geqslant Y_{b,i})$$

be the ranks and reversed ranks of the bootstrap sample. The bootstrapped rank correlation $\tilde{\xi}_b$ is then defined to be

$$\tilde{\xi}_b \equiv 1 - \frac{n}{2 \sum_{i=1}^n L_{b,i}(n - L_{b,i})} \sum_{i=1}^{n-1} |R_{b,[i+1]_b} - R_{b,[i]_b}|,$$

namely, substituting (1) into the bootstrap sample.

In this paper, we investigate two commonly used versions of the bootstrap in empirical research. The first version involves centring the bootstrap sample at ξ_n, which is calculated using the original sample. The second version involves centring the bootstrap sample at the mean of the bootstrap distribution $E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$. One can then estimate the asymptotic variance using either

$$nE\{(\tilde{\xi}_b - \xi_n)^2 \mid \vec{X}, \vec{Y}\} \quad \text{or} \quad n\text{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y}),$$

assuming an infinite number of replications for the bootstrap. In addition, it is of interest to evaluate the closeness of the bootstrap distributions of

$$n^{1/2}(\tilde{\xi}_b - \xi_n) \mid \vec{X}, \vec{Y} \quad \text{and} \quad n^{1/2}\{\tilde{\xi}_b - E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\} \mid \vec{X}, \vec{Y}$$

to that of $n^{1/2}(\xi_n - \xi)$.

2.2. *Theory*

Our theory section has to be focused on the simple independence case with $F_{X,Y} = F_X F_Y$, only under which we are able to provide the otherwise formidable calculation of the limits of $nE\{(\tilde{\xi}_b - \xi_n)^2\}$ and $n\text{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$.

The following result of Chatterjee establishes the limiting distribution of ξ_n under independence.

Proposition 1 (Chatterjee, 2021, Theorem 2.1). *Assume that $F_{X,Y} = F_X F_Y$ is fixed and continuous. Then $n^{1/2}\xi_n$ weakly converges to $N(0, 2/5)$.*

We now present the main result of this paper.

THEOREM 1 (BOOTSTRAP INCONSISTENCY). *Assuming the same conditions as in Proposition 1, the following two statements hold.*

(i) *Variance inconsistency:* $nE\{(\tilde{\xi}_b - \xi_n)^2 \mid \vec{X}, \vec{Y}\}$ *and* $n\mathrm{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$ *do not converge to* $2/5$ *in probability.*

(ii) *Distribution inconsistency: there exists a sequence of measurable events* $[\mathcal{E}_i]_{i=1}^{\infty}$, *satisfying*

$$\liminf_{n\to\infty} \mathrm{pr}\{(X_1, Y_1, \ldots, X_n, Y_n) \in \mathcal{E}_n\} > 0,$$

such that the distributions of $n^{1/2}(\tilde{\xi}_b - \xi_n)$ *and* $n^{1/2}\{\tilde{\xi}_b - E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\}$ *do not converge to* $N(0, 2/5)$ *conditional on* $(X_1, Y_1, \ldots, X_n, Y_n) \in \mathcal{E}_n$.

We are intrigued by Theorem 1 and believe that its significance is best appreciated in the context of mathematical statistics history, where the bootstrap method's validity has been a central topic, with establishing/disproving its consistency being particularly imperative.

For an independent and identically distributed sample, bootstrap consistency is often linked to the studied statistic's root-n consistency and asymptotic normality. According to Shao & Tu (1995, p. 128), the conventional wisdom seems to suggest that '[u]sually the consistency of the bootstrap distribution estimator requires some smoothness conditions that are *almost the same as* those required for the asymptotic normality of the given statistic and certain moment conditions'. As a matter of fact, this insight has been partly formalized by Mammen (1991, Theorem 1), who demonstrated, elegantly, that bootstrap consistency is equivalent to asymptotic normality when applied to linear functionals.

Indeed, the majority of theoretical results on bootstrap inconsistency are centred on statistics that do not exhibit a regular pattern of being root-n consistent and asymptotically normal. In this regard, Athreya (1987), Knight (1989) and Hall (1990) focused on the sample mean with a sample drawn from heavy-tailed distributions, Beran & Srivastava (1985) on eigenvalues, Hall et al. (1993) on ranked parameters, and Andrews (2000) and Drton & Williams (2011) on parameters at the boundary. Furthermore, Abrevaya & Huang (2005), Kosorok (2008) and Sen et al. (2010) examined cubic-root-consistent estimators, Bretagnolle (1983) and Arcones & Gine (1992) investigated degenerate U- and V-statistics, and Dümbgen (1993) and Fang & Santos (2019) explored a general class of nonsmooth plug-in estimators.

For statistics that exhibit root-n consistency and asymptotic normality, we categorize the cases where bootstrap inconsistency arises into three groups: (i) those that fail due to a moment condition, e.g., Bickel and Freedman's U-statistics (Bickel & Freedman, 1981, § 6); (ii) those that fail at superefficiency points, e.g., the Hodges and Stein estimators (Beran, 1997; Samworth, 2003) and (iii) those that do not belong to the previous two groups, including, notably, Abadie and Imbens's nearest-neighbour matching estimator of the average treatment effect (Abadie & Imbens, 2008).

Abadie and Imbens's case is particularly relevant to our work on Chatterjee's rank correlation, as both can be perceived as a type of nearest-neighbour graph-based statistics with a fixed number of nearest neighbours. To the best of our knowledge, however, no work has established a general relationship between the inconsistency of the bootstrap and the irregularity of graph-based statistics; this would be an interesting future question for mathematical statisticians.

In § B of the Supplementary Material we give a brief discussion of a weighted bootstrap approach as an alternative to Efron's standard bootstrap. Perhaps as expected, but technically more straightforward to demonstrate, the weighted bootstrap approach also cannot deliver valid inference.

Finally, the issue of bootstrap inconsistency is not a universal problem affecting all rank correlations or rank-based statistics. For example, the bootstrap consistency of Spearman's ρ and Kendall's τ rank correlations can be easily established based on the works of Bickel & Freedman (1981) and Arcones & Gine (1992). On the other hand, the bootstrap inconsistency of Hoeffding's D, Blum–Kiefer–Rosenblatt's r and Bergsma–Dassios–Yanagimoto's τ^* rank correlations under independence between X and Y is caused by the nonnormal convergence of the degenerate U-statistics, but not by the ranking.

2.3. *Simulations*

One might be tempted to speculate that the bootstrap inconsistency observed in Theorem 1 is solely due to the degeneracy property of the null point of independence. While independence does play a crucial role in the cases of Hoeffding's D, Blum–Kiefer–Rosenblatt's r and Bergsma–Dassios–Yanagimoto's τ^* rank correlations, they only become degenerate when X is independent of Y; the case of Chatterjee's rank correlation appears to be different.

As tracking the limits of $nE\{(\tilde{\xi}_b - \xi_n)^2\}$ and $n\text{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$ under dependence between Y and X is technically intimidating, this paper relies on simulations to illustrate this point. To this end, we investigate

(i) (V-LH) the asymptotic variance estimator described by Lin & Han (2022, Theorem 1.2);
(ii) (V-B1) the bootstrap asymptotic variance estimator using $nE\{(\tilde{\xi}_b - \xi_n)^2 \mid \vec{X}, \vec{Y}\}$;
(iii) (V-B2) the bootstrap asymptotic variance estimator using $n\text{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$;
(iv) (D-LH) constructing the confidence interval using the idea described by Lin & Han (2022, Remark 1.4);
(v) (D-HB1) constructing the confidence interval using the hybrid bootstrap (Shao & Tu, 1995, §4.1.5) based on $n^{1/2}(\tilde{\xi}_b - \xi_n) \mid \vec{X}, \vec{Y}$;
(vi) (D-HB2) constructing the confidence interval using the hybrid bootstrap based on $n^{1/2}\{\tilde{\xi}_b - E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\} \mid \vec{X}, \vec{Y}$.

The simulation studies were conducted based on the Gaussian rotation model, where (X, Y) are bivariate Gaussian with mean 0 and the covariance matrix Σ, defined as

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad \text{with } \rho \in (-1, 1).$$

We investigate the performance of different methods for estimating ξ_n's variance and inferring ξ using various sample sizes $n = 1000, 5000, 10\,000$ and population correlations $\rho = 0, 0.3, 0.5, 0.7, 0.9$. For the bootstrap procedure, we adopt a bootstrap size of 5000 and simulate 5000 replications to compute the square roots of the mean squared errors in estimating $n\text{var}(\xi_n)$ of limits 0.4, 0.46, 0.51, 0.47 and 0.24 as ρ changes from 0 to 0.9, as well as the empirical coverage probabilities with the nominal level $\alpha = 0.05$ or 0.1.

Table 1 presents the simulation results, demonstrating that, regardless of the strength of dependence characterized by ρ, the bootstrap methods consistently produce erroneous variance estimators and inaccurate confidence intervals. On the other hand, the method proposed by Lin & Han (2022) performs well for large n.

3. Proof of Theorem 1

Before starting the proof, we first introduce two lemmas.

Lemma 1. *Under the conditions of Proposition 1, we have*

$$\lim_{n\to\infty} E(\tilde{\xi}_b) = \frac{1}{e}.$$

Lemma 2. *Under the conditions of Proposition 1, we have*

$$\limsup_{n\to\infty} nE\{\text{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\} \leqslant \frac{3}{5} - \frac{8}{5}\frac{1}{e^2} \approx 0.3835 < \frac{2}{5}.$$

Table 1. *Variance estimation and empirical coverage probability*

ρ	n	Variance, RMSE			Coverage, $\alpha = 0.05$			Coverage, $\alpha = 0.1$		
		V-LH	V-B1	V-B2	D-LH	D-HB1	D-HB2	D-LH	D-HB1	D-HB2
0	1000	0.18	135.65	0.09	0.90	0.00	0.92	0.85	0.00	0.86
	5000	0.08	676.92	0.09	0.94	0.00	0.91	0.89	0.00	0.85
	10 000	0.06	1353.32	0.09	0.95	0.00	0.92	0.90	0.00	0.85
0.3	1000	0.18	122.19	0.16	0.90	0.00	0.89	0.85	0.00	0.82
	5000	0.07	610.14	0.16	0.94	0.00	0.88	0.89	0.00	0.81
	10 000	0.05	1220.34	0.16	0.95	0.00	0.89	0.90	0.00	0.82
0.5	1000	0.17	98.94	0.23	0.90	0.00	0.84	0.84	0.00	0.76
	5000	0.07	495.33	0.24	0.95	0.00	0.85	0.89	0.00	0.77
	10 000	0.05	990.39	0.24	0.95	0.00	0.85	0.90	0.00	0.77
0.7	1000	0.15	65.81	0.26	0.91	0.00	0.81	0.84	0.00	0.72
	5000	0.06	329.11	0.27	0.95	0.00	0.81	0.89	0.00	0.73
	10 000	0.04	657.98	0.26	0.95	0.00	0.82	0.91	0.00	0.74
0.9	1000	0.12	23.81	0.15	0.82	0.00	0.78	0.76	0.00	0.69
	5000	0.04	119.01	0.14	0.93	0.00	0.78	0.88	0.00	0.69
	10 000	0.03	238.42	0.15	0.94	0.00	0.77	0.89	0.00	0.68

The proofs of Lemma 1 and Lemma 2 are relegated to the Supplementary Material.

Proposition 1, on the other hand, shows that $n^{1/2}\xi_n$ converges in distribution to $N(0, 2/5)$.

Proof of Theorem 1(i). For any $\epsilon > 0$, since $n^{1/2}\xi_n$ is bounded in probability by the above central limit theorem, one can find $C_1 = C_1(\epsilon) > 0$ such that $\mathrm{pr}(|n^{1/2}\xi_n| > C_1) < \epsilon$ for all sufficiently large n.

For any constant $C_2 > 0$,

$$\begin{aligned} E(\tilde{\xi}_b) &= E\{E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\} \\ &= E[E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\mathbb{1}\{E(\tilde{\xi}_b \mid \vec{X}, \vec{Y}) > C_2\}] + E[E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\mathbb{1}\{E(\tilde{\xi}_b \mid \vec{X}, \vec{Y}) \leqslant C_2\}] \\ &\leqslant E[E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\mathbb{1}\{E(\tilde{\xi}_b \mid \vec{X}, \vec{Y}) > C_2\}] + C_2. \end{aligned}$$

By Lemma 1 and the fact that $E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$ is universally bounded for all $\vec{X}, \vec{Y}$, we can take $C_2 < e^{-1}$ and then, for all sufficiently large n, $\mathrm{pr}\{E(\tilde{\xi}_b \mid \vec{X}, \vec{Y}) > C_2\} \geqslant 2\epsilon$ for some $\epsilon > 0$. Then, for all sufficiently large n, with probability at least ϵ, $nE\{(\tilde{\xi}_b - \xi_n)^2 \mid \vec{X}, \vec{Y}\} \geqslant n(C_2^2 - 2C_2C_1/n^{1/2})$. This implies that $nE\{(\tilde{\xi}_b - \xi_n)^2 \mid \vec{X}, \vec{Y}\}$ does not converge in probability to any constant.

If, on the other hand, $n\mathrm{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$ converges to $2/5$ in probability then, by the Portmanteau lemma (van der Vaart, 1998, Lemma 2.2(iv)),

$$\liminf_{n\to\infty} nE\{\mathrm{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\} \geqslant 2/5,$$

which contradicts Lemma 2. Therefore, $n\mathrm{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})$ does not converge to $2/5$ in probability. □

Proof of Theorem 1(ii). We adopt the argument of Abadie & Imbens (2008). For all sufficiently large n, we can establish in a similar way as above that, with probability at least ϵ,

$$n^{1/2}(\tilde{\xi}_b - \xi_n) \geqslant n^{1/2}(C_2 - C_1/n^{1/2}).$$

This shows that $n^{1/2}(\tilde{\xi}_b - \xi_n)$ cannot converge in distribution to $N(0, 2/5)$.

If $n^{1/2}\{\tilde{\xi}_b - E(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\}$ converges in distribution to $N(0, 2/5)$ then

$$\liminf_{n\to\infty} n\mathrm{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y}) \geqslant 2/5$$

by the Portmanteau lemma (van der Vaart, 1998, Lemma 2.2(iv)). If the convergence in distribution holds for almost all sequences $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$, then

$$\liminf_{n\to\infty} nE\{\mathrm{var}(\tilde{\xi}_b \mid \vec{X}, \vec{Y})\} \geqslant 2/5,$$

which contradicts Lemma 2. □

Acknowledgement

The authors would like to thank Sourav Chatterjee for confirming the failure of the bootstrap for his rank correlation and for generously sharing his insights. The authors also thank Peter Bickel for discussing the relation between bootstrap inconsistency and adaptive estimators, Mathias Drton for sharing his work (Drton & Williams, 2011) that led the authors to Beran (1997) and Samworth (2003), and Andres Santos for explaining his findings in Fang & Santos (2019). The authors also benefitted from discussions with Lihua Lei, Bodhisattva Sen and Jon Wellner.

This work was partly motivated by discussions with Mona Azadkia, David Childers, Peng Ding, Andreas Hagemann and Mauricio Olivares. In particular, Peng posed a question regarding the practical relevance of the results presented in Lin & Han (2022) and noted that bootstrap is commonly used in practice. This paper serves as a response to Peng's inquiry.

Supplementary material

The Supplementary Material contains proofs of our results.

References

Abadie, A. & Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators. *Econometrica* **76**, 1537–57.
Abrevaya, J. & Huang, J. (2005). On the bootstrap of the maximum score estimator. *Econometrica* **73**, 1175–204.
Andrews, D. W. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica* **68**, 399–405.
Ansari, J. & Fuchs, S. (2023). A simple extension of Azadkia and Chatterjee's rank correlation to a vector of endogenous variables. *arXiv:* 2212.01621v2.
Arcones, M. A. & Gine, E. (1992). On the bootstrap of U and V statistics. *Ann. Statist.* **20**, 655–74.
Athreya, K. (1987). Bootstrap of the mean in the infinite variance case. *Ann. Statist.* **15**, 724–31.
Auddy, A., Deb, N. & Nandy, S. (2024). Exact detection thresholds for Chatterjee's correlation. *Bernoulli* **30**, 1640–68.
Azadkia, M. & Chatterjee, S. (2021). A simple measure of conditional dependence. *Ann. Statist.* **49**, 3070–102.
Azadkia, M., Taeb, A. & Bühlmann, P. (2022). A fast non-parametric approach for local causal structure learning. *arXiv:* 2111.14969v2.
Beran, R. (1982). Estimated sampling distributions: the bootstrap and competitors. *Ann. Statist.* **10**, 212–25.
Beran, R. (1997). Diagnosing bootstrap success. *Ann. Inst. Statist. Math.* **49**, 1–24.
Beran, R. & Srivastava, M. S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann. Statist.* **13**, 95–115.
Bergsma, W. & Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli* **20**, 1006–28.
Bickel, P. J. (2022). Measures of independence and functional dependence. *arXiv:* 2206.13663v1.
Bickel, P. J. & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9**, 1196–217.
Blum, J. R., Kiefer, J. & Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32**, 485–98.
Bretagnolle, J. (1983). Lois limites du bootstrap de certaines fonctionnelles. *Ann. Inst. H. Poincaré Prob. Statist.* **19**, 281–96.
Cao, S. & Bickel, P. J. (2020). Correlations with tailored extremal properties. *arXiv:* 2008.10177v2.
Chatterjee, S. (2021). A new coefficient of correlation. *J. Am. Statist. Assoc.* **116**, 2009–22.
Chatterjee, S. (2024). A survey of some recent developments in measures of association. In *Probability and Stochastic Processes: A Volume in Honour of Rajeeva L. Karandikar*, Ed. S. Athreya, A. G. Bhatt and B. V. Rao. New York: Springer.
Chatterjee, S. & Vidyasagar, M. (2023). Estimating large causal polytrees from small samples. *arXiv:* 2209.07028v2.

Chernoff, H. & Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29**, 972–94.
Deb, N., Ghosal, P. & Sen, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. *arXiv:* 2010.01768v2.
Dette, H., Siburg, K. F. & Stoimenov, P. A. (2013). A copula-based non-parametric measure of regression dependence. *Scand. J. Statist.* **40**, 21–41.
Drton, M. & Williams, B. (2011). Quantifying the failure of bootstrap likelihood ratio tests. *Biometrika* **98**, 919–34.
Dümbgen, L. (1993). On nondifferentiable functions and the bootstrap. *Prob. Theory Rel. Fields* **95**, 125–40.
Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1–26.
Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Can. J. Statist.* **9**, 139–58.
Fang, Z. & Santos, A. (2019). Inference on directionally differentiable functions. *Rev. Econ. Studies* **86**, 377–412.
Fuchs, S. (2024). Quantifying directed dependence via dimension reduction. *J. Mult. Anal.* **210**, 105266.
Gamboa, F., Gremaud, P., Klein, T. & Lagnoux, A. (2022). Global sensitivity analysis: a new generation of mighty estimators based on rank statistics. *Bernoulli* **28**, 2345–74.
Griessenberger, F., Junker, R. R. & Trutschnig, W. (2022). On a multivariate copula-based dependence measure and its estimation. *Electron. J. Statist.* **16**, 2206–51.
Hall, P. (1990). Asymptotic properties of the bootstrap for heavy-tailed distributions. *Ann. Prob.* **18**, 1342–60.
Hall, P., Härdle, W. & Simar, L. (1993). On the inconsistency of bootstrap distribution estimators. *Comp. Statist. Data Anal.* **16**, 11–8.
Han, F. (2021). On extensions of rank correlation coefficients to multivariate spaces. *Bernoulli News* **28**, 7–11.
Han, F. & Huang, Z. (2022). Azadkia-Chatterjee's correlation coefficient adapts to manifold data. *arXiv:* 2209.11156v1.
Hodges, Jr., J. L. & Lehmann, E. L. (1956). The efficiency of some nonparametric competitors of the *t*-test. *Ann. Math. Statist.* **27**, 324–35.
Hoeffding, W. (1940). Maszstabinvariante Korrelationstheorie. *Schr. Math. Inst. u. Inst. Angew. Math. Univ. Berlin* **5**, 181–233.
Hoeffding, W. (1948). A non-parametric test of independence. *Ann. Math. Statist.* **19**, 546–57.
Hoeffding, W. (1994). Scale-invariant correlation theory. In *The Collected Works of Wassily Hoeffding*, Ed. N. I. Fisher and P. K. Sen, pp. 57–107. New York: Springer.
Huang, Z., Deb, N. & Sen, B. (2022). Kernel partial correlation coefficient—a measure of conditional dependence. *J. Mach. Learn. Res.* **23**, 9699–756.
Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika* **30**, 81–93.
Kendall, M. G. (1970). *Rank Correlation Methods*, 4th ed. London: Griffin.
Knight, K. (1989). On the bootstrap of the sample mean in the infinite variance case. *Ann. Statist.* **15**, 1168–75.
Kosorok, M. R. (2008). Bootstrapping the Grenander estimator. In *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen*, Ed. N. Balakrishnan, E. A. Peña and M. J. Silvapulle, pp. 282–92. Hayward, CA: Institute of Mathematical Statistics .
Lin, Z. & Han, F. (2022). Limit theorems of Chatterjee's rank correlation. *arXiv:* 2204.08031v3.
Lin, Z. & Han, F. (2023). On boosting the power of Chatterjee's rank correlation. *Biometrika* **110**, 283–99.
Mammen, E. (1991). *When Does Bootstrap Work?: Asymptotic Results and Simulations*. New York: Springer.
Rényi, A. (1959). On measures of dependence. *Acta Math. Acad. Sci. Hungar.* **10**, 441–51.
Samworth, R. (2003). A note on methods of restoring consistency to the bootstrap. *Biometrika* **90**, 985–90.
Sen, B., Banerjee, M. & Woodroofe, M. (2010). Inconsistency of bootstrap: the Grenander estimator. *Ann. Statist.* **38**, 1953–77.
Shao, J. & Tu, D. (1995). *The Jackknife and Bootstrap*. New York: Springer.
Shi, H., Drton, M. & Han, F. (2022). On the power of Chatterjee's rank correlation. *Biometrika* **109**, 317–33.
Shi, H., Drton, M. & Han, F. (2024). On Azadkia-Chatterjee's conditional dependence coefficient. *Bernoulli* **30**, 851–77.
Sidak, Z., Sen, P. K. & Hajek, J. (1999). *Theory of Rank Tests*, 2nd ed. New York: Academic Press.
Spearman, C. (1904). The proof and measurement of association between two things. *Am. J. Psychol.* **15**, 72–101.
Spearman, C. (1906). 'Footrule' for measuring correlation. *Br. J. Psychol.* **2**, 89–108.
Strothmann, C., Dette, H. & Siburg, K. F. (2024). Rearranged dependence measures. *Bernoulli* **30**, 1055–78.
van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.
Yanagimoto, T. (1970). On measures of association and a related problem. *Ann. Inst. Statist. Math.* **22**, 57–63.
Zhang, Q. (2023a). On relationships between Chatterjee's and Spearman's correlation coefficients. *arXiv:* 2302.10131v1.
Zhang, Q. (2023b). On the asymptotic null distribution of the symmetrized Chatterjee's correlation coefficient. *Statist. Prob. Lett.* **194**, 109759.

[*Received on* 5 *April* 2023. *Editorial decision on* 10 *January* 2024]