

Article

Measuring Dispersion and Serial Dependence in Ordinal Time Series Based on the Cumulative Paired ϕ -Entropy

Christian H. Weiß 

Department of Mathematics and Statistics, Helmut Schmidt University, 22043 Hamburg, Germany; weissc@hsu-hh.de; Tel.: +49-40-6541-2779

Abstract: The family of cumulative paired ϕ -entropies offers a wide variety of ordinal dispersion measures, covering many well-known dispersion measures as a special case. After a comprehensive analysis of this family of entropies, we consider the corresponding sample versions and derive their asymptotic distributions for stationary ordinal time series data. Based on an investigation of their asymptotic bias, we propose a family of signed serial dependence measures, which can be understood as weighted types of Cohen's κ , with the weights being related to the actual choice of ϕ . Again, the asymptotic distribution of the corresponding sample κ_ϕ is derived and applied to test for serial dependence in ordinal time series. Using numerical computations and simulations, the practical relevance of the dispersion and dependence measures is investigated. We conclude with an environmental data example, where the novel ϕ -entropy-related measures are applied to an ordinal time series on the daily level of air quality.

Keywords: Cohen's κ ; dispersion; entropy; ordinal time series; serial dependence



Citation: Weiß, C.H. Measuring Dispersion and Serial Dependence in Ordinal Time Series Based on the Cumulative Paired ϕ -Entropy. *Entropy* **2022**, *24*, 42. <https://doi.org/10.3390/e24010042>

Academic Editor: Luca Faes

Received: 1 December 2021

Accepted: 23 December 2021

Published: 26 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

During the last years, ordinal data in general [1] and ordinal time series in particular [2] received a great amount of interest in research and applications. Here, a random variable X is said to be ordinal if X has a bounded qualitative range exhibiting a natural order among the categories. We denote the range as $\mathcal{S} = \{s_0, s_1, \dots, s_m\}$ with some $m \in \mathbb{N} = \{1, 2, \dots\}$, and we assume that $s_0 < \dots < s_m$. The realized data are denoted as x_1, \dots, x_n with $n \in \mathbb{N}$. They are assumed to stem either from independent and identically distributed (i. i. d.) replications of X (then, we refer to x_1, \dots, x_n as an ordinal random sample), or from a stationary ordinal stochastic process $(X_t)_{\mathbb{Z}=\{\dots, -1, 0, 1, \dots\}}$ (then, x_1, \dots, x_n are said to be an ordinal time series).

In what follows, we take up several recent works on measures of dispersion and serial dependence in ordinal (time series) data. Regarding ordinal dispersion, the well-known measures such as variance or mean absolute deviation cannot be used as the data are not quantitative. Therefore, several tailor-made measures for ordinal dispersion have been developed and investigated in the literature, see, among others, [3–15]. The unique feature of all these measures is that they rely on the cumulative distribution function (CDF) of X , i.e., on $\mathbf{f} = (f_0, \dots, f_{m-1})^\top$ with $f_i = P(X \leq s_i)$ for $i = 0, \dots, m$ (f_m is omitted in \mathbf{f} as it necessarily equals one). They classify any one-point distribution on \mathcal{S} as a scenario of minimal dispersion, i.e., if all probability mass concentrates on one category from \mathcal{S} (maximal consensus):

$$\mathbf{f}_{\text{one}} \in \left\{ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \right\}.$$

By contrast, maximal dispersion is achieved exactly in the case of the extreme two-point distribution (polarized distribution), $\mathbf{f}_{\text{two}} = (\frac{1}{2}, \dots, \frac{1}{2})^\top$, where we have 50% proba-

bility mass in each of the outermost categories (maximal dissent). Further details on ordinal dispersion measures are presented in Section 2 below.

Building upon earlier works by Klein [16], Yager [17], it was recently shown by Klein & Doll [18], Klein et al. [19] that the aforementioned ordinal dispersion measures can be subsumed under the family of so-called “cumulative paired ϕ -entropies” (see Section 2), abbreviated as CPE_ϕ , which constitutes the starting point of the present article. Our first main task is to derive the asymptotic distribution of the corresponding sample version, \widehat{CPE}_ϕ , for both i. i. d. and time series data, and to check the finite sample performance of the resulting approximate distribution, see Sections 3 and 5.

In the recent paper by Weiß [20] on the asymptotics of some well-known dispersion measures for nominal data (i.e., qualitative data without a natural ordering), it turned out that the corresponding dispersion measures—if these are applied to time series data—are related to specific measures of serial dependence. Therefore, our second main task is to explore for a similar relation in the ordinal case, if considering the CPE_ϕ -family for measuring dispersion. Ordinal dependence measures can be defined in analogy to the popular autocorrelation function (ACF) for quantitative time series, namely by using the lagged bivariate CDF $f_{ij}(h) = P(X_t \leq s_i, X_{t-h} \leq s_j)$ for time lags $h \in \mathbb{N}$ as their base [14]. For the novel family of $\kappa_\phi(h)$ measures, which cover the existing ordinal Cohen’s κ [14,15] as a special case, we derive the asymptotics under the null hypothesis of i. i. d. time series data, see Section 4. This result is used in Section 5 to test for significant serial dependence, in analogy to the application of the sample ACF to quantitative time series. In Section 6, we discuss an illustrative real-world data example from an environmental application, namely regarding the daily level of air quality. The article concludes in Section 7.

2. The Family of Cumulative Paired ϕ -Entropies

Klein & Doll [18], Klein et al. [19] proposed and investigated a family of cumulative paired ϕ -entropies. Although their main focus was on continuously distributed random variables, they also referred to the ordinal case and pointed out that many well-known ordinal dispersion measures are included in this family. Here, we exclusively concentrate on the ordinal case as introduced in Section 1, and we define the (normalized) cumulative paired ϕ -entropy as (see Section 2.3 in Klein et al. [19])

$$CPE_\phi(f) = \frac{1}{2m\phi(1/2)} \sum_{i=0}^{m-1} (\phi(f_i) + \phi(1 - f_i)). \tag{1}$$

The entropy generating function (EGF) ϕ is defined on $[0; 1]$, it satisfies $\phi(0) = \phi(1) = 0$, and it is assumed to be concave on $[0; 1]$. Later in Section 3, when deriving the asymptotic distribution of the sample counterpart $\widehat{CPE}_\phi = CPE_\phi(\hat{f})$, we shall also require that ϕ is (twice) differentiable. As pointed out in Sections 2.3 and 3.1 of Klein et al. [19], several well-known measures of ordinal dispersion can be expressed by (1) with an appropriate choice of ϕ .

- *Leik’s ordinal variation* [11] corresponds to the choice $\phi(z) = \min\{z, 1 - z\}$ (which is not differentiable in $z = 1/2$) because of the equality $|2z - 1| = 1 - 2 \min\{z, 1 - z\}$:

$$LOV = 1 - \frac{1}{m} \sum_{i=0}^{m-1} |2f_i - 1| = \frac{2}{m} \sum_{i=0}^{m-1} \min\{f_i, 1 - f_i\}. \tag{2}$$

- An analogous argument applies to the whole family of ordinal variation measures, OV_q with $q \geq 1$ [3,9,10,13]. Choosing $\phi_q(z) = 1 - (1 - 2 \min\{z, 1 - z\})^q = 1 - |2z - 1|^q$ with $\phi_q(1/2) = 1$, we have the relation

$$OV_q = 1 - \left(\frac{1}{m} \sum_{i=0}^{m-1} |2f_i - 1|^q\right)^{1/q} = 1 - \left(1 - CPE_\phi(f)\right)^{1/q}. \tag{3}$$

Note that $q = 1$ leads to the LOV, and the case $q = 2$ is known as the coefficient of ordinal variation, $COV = OV_2$ [4,8].

- Related to the previous case OV_q with $q = 2$, $CPE_\phi(f)$ becomes the widely-used index of ordinal variation [3,7,8] if choosing $\phi(z) = z(1 - z)$:

$$IOV = \frac{4}{m} \sum_{i=0}^{m-1} f_i(1 - f_i) = 1 - \frac{1}{m} \sum_{i=0}^{m-1} (2f_i - 1)^2. \tag{4}$$

- The cumulative paired (Shannon) entropy [12] corresponds to the choice $\phi(z) = -z \ln z$ (with the convention $0 \ln 0 = 0$):

$$CPE = \frac{-1}{m \ln 2} \sum_{i=0}^{m-1} (f_i \ln f_i + (1 - f_i) \ln(1 - f_i)). \tag{5}$$

- $\phi(z) = -z \ln z$ can be embedded into the family of a -entropies [21,22],

$$\phi_a(z) = \frac{z - z^a}{a - 1} \quad \text{with } a > 0 \text{ and } a \neq 1, \tag{6}$$

as the boundary case $a \rightarrow 1$. Plugging-in (6) into (1), one obtains

$$CPE_a = \frac{2^{a-1}}{m(2^{a-1}-1)} \sum_{i=0}^{m-1} (1 - f_i^a - (1 - f_i)^a). \tag{7}$$

Note that both $a = 2$ and $a = 3$ in (7) lead to the IOV (4).

The EGFs ϕ involved in (2)–(4) are symmetric around $1/2$, i.e., they satisfy $\phi(z) = \phi(1 - z)$. This is also illustrated by Figure 1, where the cases $q = 2$ (left) and $a = 2$ (right; both in bold black) agree with each other except a scaling factor. The plotted EGFs $\phi_q(z)$ are maximal in $1/2$ with $\phi_q(1/2) = 1$. The EGF $\phi_a(z)$ is maximal in $a^{1/(1-a)}$ with $\phi_a(a^{1/(1-a)}) = a^{a/(1-a)}$ for $a \neq 1$, whereas in the boundary case $a \rightarrow 1$, $\phi_a(z)$ takes its maximal value $1/e$ at $1/e$. However, since CPE_ϕ in (1) is normalized, it is not necessary to care for a possible rescaling of $\phi_a(z)$ if computing CPE_ϕ .

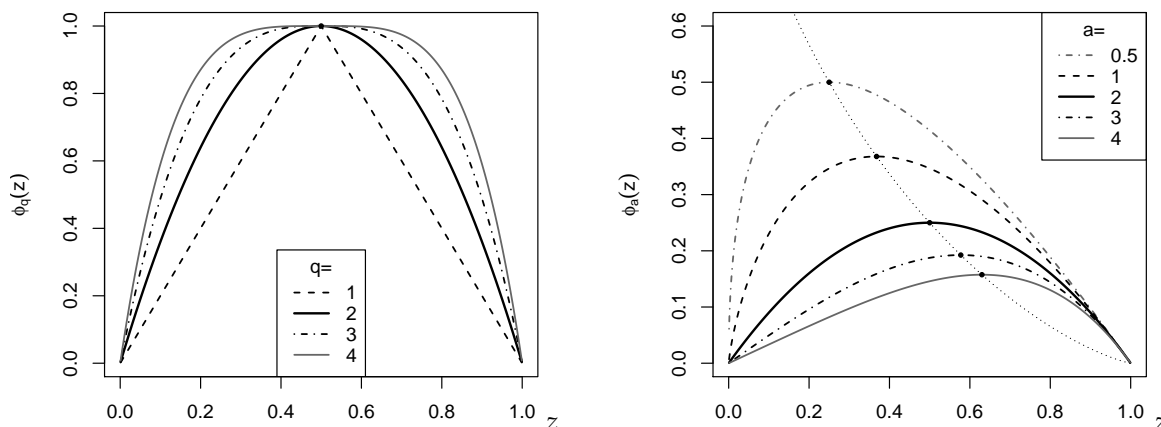


Figure 1. Plot of EGFs $\phi(z)$ against z . (Left): $\phi_q(z) = 1 - |2z - 1|^q$; (right): $\phi_a(z) = (z - z^a)/(a - 1)$.

Remark 1. The dotted curve in the right part of Figure 1, which connects the maxima of $\phi_a(z)$ for different a , is computed by using the Lambert W function (also referred to as the product logarithm). This can be seen as follows: $\phi_a(z)$ is maximal in $z_0 = a^{1/(1-a)}$ with $\phi_a(z_0) = z_0^a$. It holds that

$$a \cdot z_0^a = a^{1+a/(1-a)} = z_0.$$

Using that $z_0^a = \exp(a \ln z_0)$, this implies that

$$(a \ln z_0) \cdot \exp(a \ln z_0) = z_0 \ln z_0.$$

The Lambert W function is defined to solve the equation $w \exp w = z$ as $w = W(z)$, so we get

$$a \ln z_0 = W(z_0 \ln z_0) \quad \Leftrightarrow \quad a = W(z_0 \ln z_0) / \ln z_0.$$

Thus, since $\phi_a(z_0) = a^{a/(1-a)} = a^{1/(1-a)} / a = z_0 / a$, the dotted curve in Figure 1 follows the function $z \ln z / W(z \ln z)$. More precisely, since $z \ln z$ is minimal in $z = 1/e$ with minimal value $-1/e$, we have to use the principal branch $W(z) = W_0(z)$ for $z \leq 1/e$, and the secondary branch $W(z) = W_{-1}(z)$ for $z > 1/e$.

Let us conclude this section with some examples of CPE_ϕ measures (see Figure 2). For all examples, we set $m = 4$, i.e., we have five ordinal categories like, for example, in the case of a simple Likert scale. In the left part of Figure 2, f was computed according to the binomial distribution $\text{Bin}(4, p)$, which has maximal dispersion for $p = 0.5$. This is also recognized by any of the plotted measures CPE_ϕ , with their maximal dispersion values varying around 0.6. This medium level of dispersion is plausible because $\text{Bin}(4, 0.5)$ is far away from the extreme two-point distribution. The right part of Figure 2, by contrast, shows the CPE_ϕ for the two-point distribution with $f_0 = p$ ($= f_1 = \dots = f_{m-1}$). So $p = 0$ corresponds to a one-point distribution in s_m (minimal dispersion), and $p = 0.5$ to the extreme two-point distribution (maximal dispersion). Accordingly, all measures reach their extreme values 0 and 1, respectively. It is interesting to note that outside these extreme cases, the dispersion measures judge the actual dispersion level quite differently; see the related discussion in Kvålseth [10], Weiß [13].

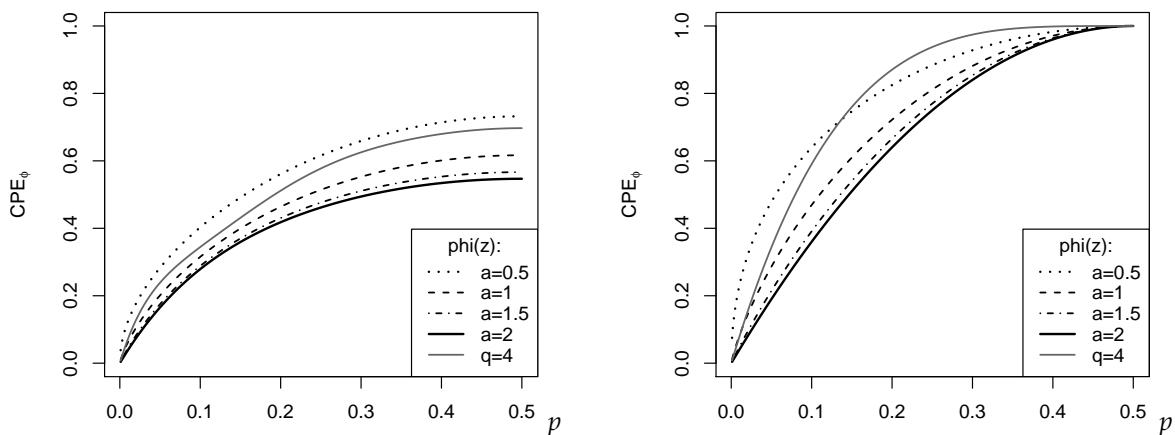


Figure 2. Plot of CPE_ϕ against p for specific cases of $\phi_a(z) = (z - z^a)/(a - 1)$ and $\phi_q(z) = 1 - |2z - 1|^q$. (Left): binomial distribution $\text{Bin}(4, p)$; (Right): two-point distribution with $f_0 = p$.

3. Asymptotic Distribution of Sample CPE_ϕ

From now on, we focus on the sample version of CPE_ϕ from (1), i.e., on $\widehat{CPE}_\phi = CPE_\phi(\hat{f})$, where \hat{f} denotes the vector of cumulative relative frequencies computed from X_1, \dots, X_n . To derive the asymptotic distribution of \widehat{CPE}_ϕ , which is to be used as an approximation to the true distribution of CPE_ϕ , we recall the following results from Weiß [14]. Provided that the data-generating process (DGP) satisfies appropriate

mixing conditions, e.g., α -mixing with exponentially decreasing weights (which includes the i. i. d.-case), it holds that

$$\begin{aligned} \sqrt{n}(\hat{f} - f) &\xrightarrow{d} N(\mathbf{0}, \Sigma) \quad \text{with } \Sigma = (\sigma_{ij})_{i,j=0,\dots,m-1}, \quad \text{where} \\ \sigma_{ij} &\stackrel{a}{=} f_{\min\{i,j\}} - f_i f_j + \sum_{h=1}^{\infty} (f_{ij}(h) + f_{ji}(h) - 2 f_i f_j). \end{aligned} \tag{8}$$

For an analogous result in the presence of missing data, see Theorem 1 in Weiß [15]. In (8), finite (co)variances are ensured if we require that $\sum_{h=1}^{\infty} (f_{ij}(h) - f_i f_j) < \infty$ holds for all i, j (“short memory”). In particular, all sums $\sum_{h=1}^{\infty} (f_{ij}(h) + f_{ji}(h) - 2 f_i f_j)$ vanish in the i. i. d.-case. Otherwise, they account for the serial dependence of the DGP. This can be seen by considering the trace of Σ , which equals

$$\sum_{i=0}^{m-1} \left(f_i(1 - f_i) + 2 \sum_{h=1}^{\infty} (f_{ii}(h) - f_i^2) \right) = \sum_{i=0}^{m-1} f_i(1 - f_i) \left(1 + 2 \sum_{h=1}^{\infty} \kappa_{\text{ord}}(h) \right).$$

Here, the term $\sum_{i=0}^{m-1} f_i(1 - f_i)$ agrees with the IOV in (4) except the normalizing factor $\frac{4}{m}$, i.e., it corresponds to CPE_{ϕ} with $\phi(z) = z(1 - z)$. The term $\kappa_{\text{ord}}(h)$, in turn, is the ordinal Cohen’s κ [14] defined by

$$\kappa_{\text{ord}}(h) = \frac{\sum_{i=0}^{m-1} (f_{ii}(h) - f_i^2)}{\sum_{i=0}^{m-1} f_i(1 - f_i)}. \tag{9}$$

It is a measure of signed serial dependence, which evaluates the extent of (dis)agreement between X_t and X_{t-h} by positive (negative) values.

Based on Taylor expansions of $\widehat{\text{CPE}}_{\phi} = \text{CPE}_{\phi}(\hat{f})$ in f , we shall now study its asymptotic behavior. To establish asymptotic normality and to derive the asymptotic variance of $\widehat{\text{CPE}}_{\phi}$, we need ϕ to be differentiable. For an asymptotic bias correction, which relies on a second-order Taylor expansion, ϕ even has to be twice differentiable (then, the concavity of ϕ implies that $\phi''(z) < 0$).

Remark 2. From the examples discussed in Section 2, the EGF corresponding to the LOV (i.e., ϕ_q with $q = 1$) is not differentiable (in $z = 1/2$). ϕ_q is only once differentiable for $1 < q < 2$, while $q \geq 2$ ensures ϕ_q to be at least twice differentiable; see Example 1 below. Accordingly, in these cases, it is not possible to establish asymptotic normality ($q = 1$) or an asymptotic bias correction ($1 < q < 2$), respectively. In fact, Weiß [13] was faced with the same problem when studying the asymptotics of the sample OV_q , and a solution to it was not possible. In simulations, he showed that even modified asymptotics (using a folded-normal distribution) did not lead to an acceptable approximation quality. We shall therefore exclude such cases from our further discussion. If, in applications, the cases $q = 1$ or $1 < q < 2$ appear to be relevant, a bootstrap approach might be an option to gain insight into the sample distribution of $\widehat{\text{CPE}}_{\phi}$.

Assuming ϕ to be (twice) differentiable, the partial derivatives of $\text{CPE}_{\phi}(f)$ according to (1) are

$$\begin{aligned} \frac{\partial}{\partial f_k} \text{CPE}_{\phi}(f) &= \frac{1}{2m\phi(1/2)} \left(\phi'(f_k) - \phi'(1 - f_k) \right) =: d_k, \\ \frac{\partial^2}{\partial^2 f_k} \text{CPE}_{\phi}(f) &= \frac{1}{2m\phi(1/2)} \left(\phi''(f_k) + \phi''(1 - f_k) \right) =: h_{kk}, \\ \frac{\partial^2}{\partial f_k \partial f_l} \text{CPE}_{\phi}(f) &= 0 \quad \text{for } k \neq l. \end{aligned} \tag{10}$$

We denote the gradient (Jacobian) of $\text{CPE}_{\phi}(f)$ by $\mathbf{D} = (d_0, \dots, d_{m-1})$, and the Hessian equals $\mathbf{H} = \text{diag}(h_{00}, \dots, h_{m-1,m-1})$. Note that if ϕ is symmetric around $1/2$, i.e., if $\phi(z) = \phi(1 - z)$, then $d_k = \phi'(f_k) / (m\phi(1/2))$ and $h_{kk} = \phi''(f_k) / (m\phi(1/2))$.

Example 1. Let us compute the derivatives required in (10) for the EGF examples presented in Section 2.

For $\phi_a(z) = \frac{z-z^a}{a-1}$, the constant factor becomes $\frac{1}{2m\phi_a(1/2)} = \frac{2^{a-1}(a-1)}{m(2^{a-1}-1)}$, and the derivatives are $\phi'_a(z) = \frac{1-az^{a-1}}{a-1}$ and $\phi''_a(z) = -az^{a-2}$. Here, $\phi'_a(z)$ exists in the boundary value $z = 0$ only if $a > 1$, and $\phi''_a(z)$ if $a \geq 2$. Important special cases are

$$\phi(z) = -z \ln z \Rightarrow \frac{1}{2m\phi(1/2)} = \frac{1}{m \ln 2}, \quad \phi'(z) = -1 - \ln z, \quad \phi''(z) = -1/z \quad \text{for } a \rightarrow 1,$$

$$\phi(z) = z(1-z) \Rightarrow \frac{1}{2m\phi(1/2)} = \frac{2}{m}, \quad \phi'(z) = 1-2z, \quad \phi''(z) = -2 \quad \text{for } a = 2,$$

and for $a = 3$,

$$\phi(z) = \frac{1}{2}z(1-z^2) \Rightarrow \frac{1}{2m\phi(1/2)} = \frac{8}{3m}, \quad \phi'(z) = \frac{1}{2}(1-3z^2), \quad \phi''(z) = -3z.$$

Note that both $a = 2, 3$ lead to the same expressions for d_k, h_{kk} in (10); see Table 1. This is in accordance with the equivalence of these cases as discussed after (7).

For $\phi_q(z) = 1 - (1 - 2 \min\{z, 1-z\})^q = 1 - |2z-1|^q$ with $\frac{1}{2m\phi_q(1/2)} = \frac{1}{2m}$, the derivatives are expressed using the sign function, $\text{sgn}(\cdot)$, which is not continuous at 0. Note that the following relations hold:

$$\frac{d}{dx} |x| = \text{sgn}(x) \text{ for } x \neq 0, \quad x = \text{sgn}(x) \cdot |x|, \quad |x| = \text{sgn}(x) \cdot x.$$

For $q \geq 2$, it then follows by applying the chain rule and the product rule that

$$\begin{aligned} \phi'_q(z) &= -2q|2z-1|^{q-1} \text{sgn}(2z-1) = -2q(2z-1)|2z-1|^{q-2}, \\ \phi''_q(z) &= -4q|2z-1|^{q-2} - 4q(q-2)(2z-1)|2z-1|^{q-3} \text{sgn}(2z-1) \\ &= -4q(q-1)|2z-1|^{q-2}. \end{aligned}$$

Note that these derivatives are continuous in $z = 1/2$ for $q \geq 2$, using that $0^0 = 1$. The final expressions for (10) are summarized in Table 1.

Table 1. Partial derivatives (10) for EGFs discussed in Example 1.

| EGF $\phi(z)$ | d_k | h_{kk} |
|-------------------------------|--|--|
| $(z - z^a)/(a - 1)$ | $\frac{2^{a-1}a}{m(2^{a-1}-1)} \left((1-f_k)^{a-1} - f_k^{a-1} \right)$ | $-\frac{2^{a-1}a(a-1)}{m(2^{a-1}-1)} \left(f_k^{a-2} + (1-f_k)^{a-2} \right)$ |
| $-z \ln z$ | $\frac{1}{m \ln 2} \left(\ln(1-f_k) - \ln f_k \right)$ | $\frac{-1}{m \ln 2} \left(1/f_k + 1/(1-f_k) \right)$ |
| $z(1-z), \frac{1}{2}z(1-z^2)$ | $\frac{4}{m}(1-2f_k)$ | $-\frac{8}{m}$ |
| $1 - 2z-1 ^q$ | $-\frac{2q}{m}(2f_k-1) 2f_k-1 ^{q-2}$ | $-\frac{4q(q-1)}{m} 2f_k-1 ^{q-2}$ |

Using (10), the second-order Taylor expansion equals

$$\text{CPE}_\phi(\hat{f}) \approx \text{CPE}_\phi(f) + \sum_{k=0}^{m-1} d_k (\hat{f}_k - f_k) + \frac{1}{2} \sum_{k=0}^{m-1} h_{kk} (\hat{f}_k - f_k)^2. \tag{11}$$

According to (8), the linear term in (11) is asymptotically normally distributed (“Delta method”), provided that \mathbf{D} does not vanish (see Remark 3 below). Then, we conclude from (8) that

$$\sqrt{n} \left(\text{CPE}_\phi(\hat{f}) - \text{CPE}_\phi(f) \right) \xrightarrow{d} \mathbf{N}(0, \mathbf{D} \Sigma \mathbf{D}^\top), \quad \mathbf{D} \Sigma \mathbf{D}^\top = \sum_{i,j=0}^{m-1} d_i d_j \sigma_{ij}. \tag{12}$$

The approximate bias correction relies on the quadratic term in (11), because $E[\hat{f}_k - f_k] = 0$. Using (8), we conclude that

$$n E[\text{CPE}_\phi(\hat{f}) - \text{CPE}_\phi(f)] \approx \frac{1}{2} \sum_{i=0}^{m-1} h_{ii} \sigma_{ii}. \tag{13}$$

Let us summarize the results implied by (12) and (13) in the following theorem.

Theorem 1. Under the mixing assumptions stated before (8), assuming the EGF ϕ to be twice differentiable, and recalling the d_k, h_{kk} from (10) where \mathbf{D} must not vanish, it holds that

$$\sqrt{n} \left(\text{CPE}_\phi(\hat{f}) - \text{CPE}_\phi(f) \right) \xrightarrow{d} \text{N}(0, \sigma_\phi^2), \quad \sigma_\phi^2 = \sigma_{\phi,\text{iid}}^2 \left(1 + 2 \sum_{h=1}^{\infty} \vartheta_\phi(h) \right)$$

$$\text{with } \sigma_{\phi,\text{iid}}^2 = \sum_{i,j=0}^{m-1} d_i d_j (f_{\min\{i,j\}} - f_i f_j) \quad \text{and} \quad \vartheta_\phi(h) = \frac{\sum_{i,j=0}^{m-1} d_i d_j (f_{ij}(h) - f_i f_j)}{\sum_{i,j=0}^{m-1} d_i d_j (f_{\min\{i,j\}} - f_i f_j)}.$$

In addition, the bias-corrected mean of $\text{CPE}_\phi(\hat{f})$ is

$$E[\text{CPE}_\phi(\hat{f})] \approx \text{CPE}_\phi(f) + \frac{1}{2n} \left(\sum_{i=0}^{m-1} h_{ii} f_i (1 - f_i) \right) \left(1 + 2 \sum_{h=1}^{\infty} \kappa_\phi(h) \right),$$

$$\text{where } \kappa_\phi(h) = \frac{\sum_{i=0}^{m-1} h_{ii} (f_{ii}(h) - f_i^2)}{\sum_{i=0}^{m-1} h_{ii} f_i (1 - f_i)}.$$

Note that the second-order derivatives are negative due to the concavity of ϕ , so $\text{CPE}_\phi(\hat{f})$ exhibits a negative bias. $\vartheta_\phi(h)$ expresses the effect of serial dependence on σ_ϕ^2 . For i. i. d. ordinal data, $\vartheta_\phi(h) = 0$, so Theorem 1 simplifies considerably in this case, namely to $\sigma_\phi^2 = \sigma_{\phi,\text{iid}}^2$. The bias of $\text{CPE}_\phi(\hat{f})$ is affected by serial dependence via $\kappa_\phi(h)$, which is a κ -type measure reflecting the extent of (dis)agreement between lagged observations, recall (9). More precisely, $\kappa_\phi(h)$ can be interpreted as weighted type of $\kappa_{\text{ord}}(h)$, where the weights h_{ii} depend on the particular choice of ϕ . It thus provides a novel family of measures of signed serial dependence, the asymptotics of which are analyzed in Section 4 below.

Example 2. In the special case $\phi(z) = z(1 - z)$ (as well as for $\phi(z) = \frac{1}{2} z (1 - z^2)$), which corresponds to the IOV in (4), $h_{ii} = -\frac{8}{m}$ is constant (see Table 1). Thus, $\kappa_\phi(h) = \kappa_{\text{ord}}(h)$ in this case (see (9)). Furthermore, the first factor of the bias in Theorem 1 becomes

$$\frac{1}{2n} \sum_{i=0}^{m-1} h_{ii} f_i (1 - f_i) = -\frac{1}{n} \cdot \frac{4}{m} \sum_{i=0}^{m-1} f_i (1 - f_i) = -\frac{1}{n} \text{IOV}.$$

Hence, the bias is determined by both the serial dependence and the dispersion of the process.

As another simple example, consider the choice $\phi(z) = -z \ln z$ for the CPE in (5). Then, using $h_{ii} = \frac{-1}{m \ln 2} \left(\frac{1}{f_i} + \frac{1}{1-f_i} \right) = \frac{-1}{m \ln 2} \frac{1}{f_i(1-f_i)}$ from Table 1, we get

$$\frac{1}{2n} \sum_{i=0}^{m-1} h_{ii} f_i (1 - f_i) = \frac{1}{2n} \frac{-1}{m \ln 2} \sum_{i=0}^{m-1} 1 = \frac{-1}{(2 \ln 2) n}.$$

Thus, we have a unique i. i. d.-bias, independent of the actual marginal CDF f . Under serial dependence, we get

$$\kappa_\phi(h) = \frac{\sum_{i=0}^{m-1} h_{ii} (f_{ii}(h) - f_i^2)}{\sum_{i=0}^{m-1} h_{ii} f_i (1 - f_i)} = \frac{1}{m} \sum_{i=0}^{m-1} \frac{f_{ii}(h) - f_i^2}{f_i(1 - f_i)} =: \kappa_{\text{ord}}^*(h). \tag{14}$$

So, besides the pair $(\text{IOV}, \kappa_{\text{ord}}(h))$, $(\text{CPE}, \kappa_{\text{ord}}^*(h))$ also belongs to the $(\text{CPE}_\phi, \kappa_\phi(h))$ -family.

A few examples are plotted in Figure 3, where the DGP $X_t = s_{I_t}$ assumes the rank counts I_t to have $\text{Bin}(4, p)$ -marginals. In the top panel, the DGP is i. i. d., whereas (I_t) follows a so-called first-order binomial autoregressive (BAR(1)) model with dependence parameter $\rho = 0.4$ ([2] Section 3.3)

in the lower panel, i.e., the DGP exhibits a medium level of positive dependence. While the resulting dependence structure is investigated in more detail in Section 4, Figure 3 considers the asymptotic standard error (SE) σ_ϕ and bias $n(E[CPE_\phi(\hat{f})] - CPE_\phi(f))$ according to Theorem 1. Obviously, an increase of serial dependence causes an increase of both SE and bias. While most measures have a rather stable SE for varying p (except for extremely small p , where we are close to a one-point distribution), the EGF ϕ_a with $a = 1/2$ varies a lot. In particular, the bias takes rather extreme values with decreasing p for this case, which can be explained by the strongly negative exponents at $f_k, 1 - f_k$ in h_{kk} from Table 1. Thus, choices $a < 1$ seem not advisable for practice. The boundary case $a = 1$ has a constant bias for an i. i. d. DGP. For ϕ_q with $q = 4$, we note an oscillating behavior of both bias and SE. The lowest bias and SE are achieved for the cases ϕ_a with $a > 1$.

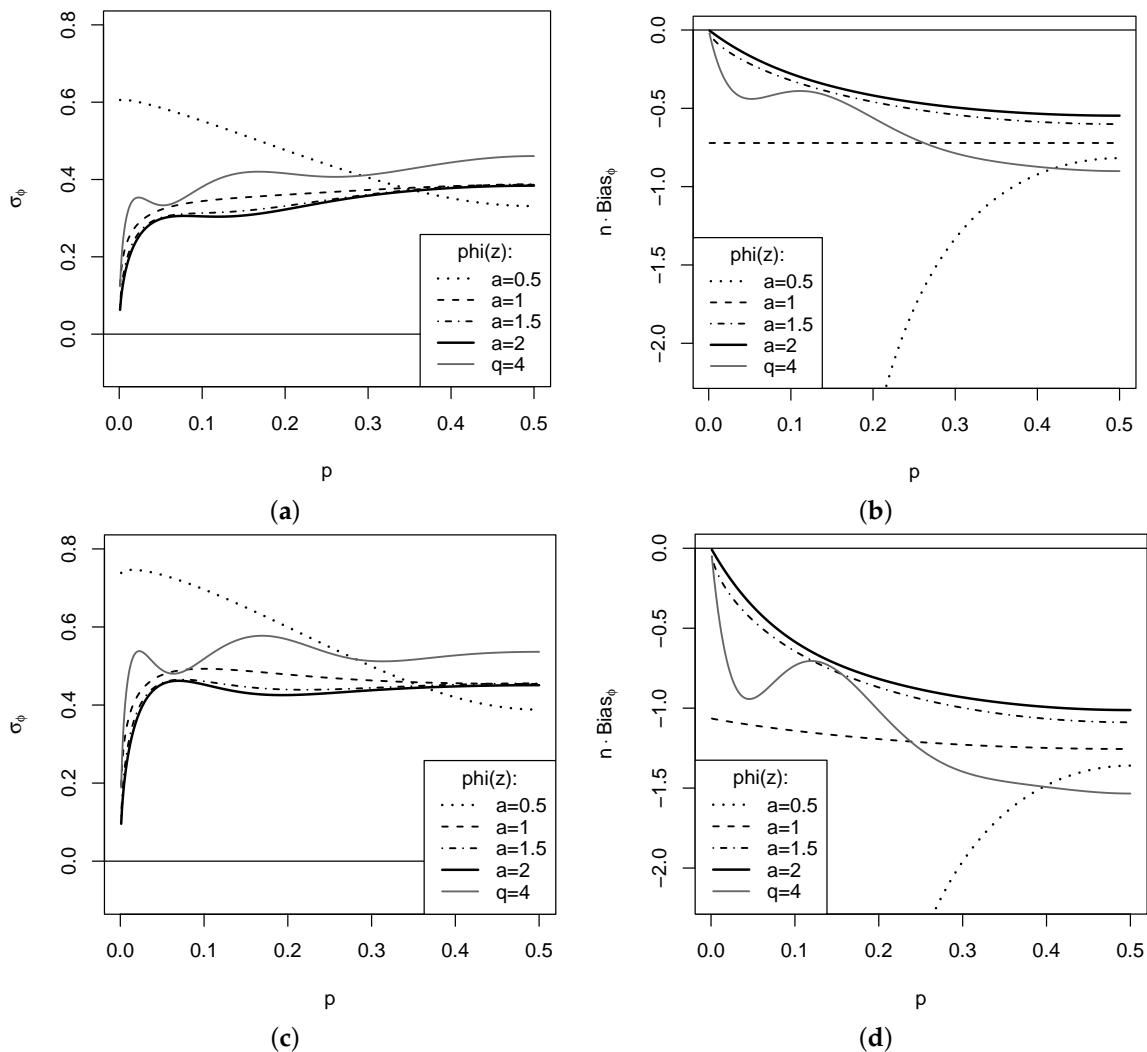


Figure 3. Plots for specific cases of $\phi_a(z) = (z - z^a)/(a - 1)$ and $\phi_q(z) = 1 - |2z - 1|^q$: σ_ϕ (left) and $n(E[CPE_\phi(\hat{f})] - CPE_\phi(f))$ (right) against p of marginal Bin(4, p), where i. i. d. DGP in (a,b), and BAR(1) DGP with dependence parameter $\rho = 0.4$ in (c,d).

The newly obtained measure $\kappa_{ord}^*(h)$ from (14) constitutes a counterpart to the nominal measures $\kappa^*(h), \kappa^*(h)$ in Weiß [20]. It is worth pointing out that the latter measures were derived from the nominal entropy and extropy, respectively, while the CPE in (5) can be interpreted as a combination of cumulative entropy and extropy. It has to be noted that $\kappa_{ord}^*(h)$ also shares a disadvantage with $\kappa^*(h)$: if only one of the f_i equals 0 or 1, we suffer from a division by 0 in (14). For $\kappa_{ord}(h)$ according to (9), by contrast, a division by 0 only happens in the (deterministic) case of a one-point distribution. As a workaround when computing $\kappa_{ord}^*(h)$, it is recommended to replace the affected summands in (14) by 0.

Remark 3. If $f = f_{\text{two}}$, then all $d_k = 0$ in (10). Therefore, the linear term in (11) vanishes. In fact, for any two-point distribution on s_0 and s_m , we necessarily have $f_0 = \dots = f_{m-1}$ and $\hat{f}_0 = \dots = \hat{f}_{m-1}$. Therefore, $\text{CPE}_\phi(\hat{f})$ reduces to $\text{CPE}_\phi(\hat{f}) = \frac{1}{2\phi(1/2)} (\phi(\hat{f}_0) + \phi(1 - \hat{f}_0))$, and the quadratic term in (11) to $\frac{m}{2} h_{00} (\hat{f}_0 - f_0)^2$. Hence, in this special case,

$$n \left(\text{CPE}_\phi(\hat{f}) - \text{CPE}_\phi(f) \right) \overset{a}{\sim} \frac{m}{2} h_{00} \sigma_{00} \cdot \chi_1^2.$$

For example, plugging-in $h_{00} = -\frac{8}{m}$ for $\phi(z) = z(1 - z)$ corresponding to the IOV in (4), we obtain the result in Remark 7.1.2 in Weiß [14].

4. Asymptotic Distribution of Sample $\kappa_\phi(h)$

The bias equation in Theorem 1 gives rise to a novel family of serial dependence measures for ordinal time series, namely

$$\kappa_\phi(h) = \frac{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1 - f_i)) (f_{ii}(h) - f_i^2)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1 - f_i)) f_i(1 - f_i)} \tag{15}$$

for a given EGF ϕ . Some examples are plotted in the left part of Figure 4, where the DGP $X_t = s_{I_t}$ assumes that the rank counts I_t follow the BAR(1) model with marginal distribution $\text{Bin}(4, 0.3)$ and dependence parameter ρ ; recall the discussion of Figure 3. So, the rank counts $(I_t)_\mathbb{Z}$ have the first-order ACF ρ , whereas the plotted $\kappa_\phi(1)$ have absolute value $\leq |\rho|$.

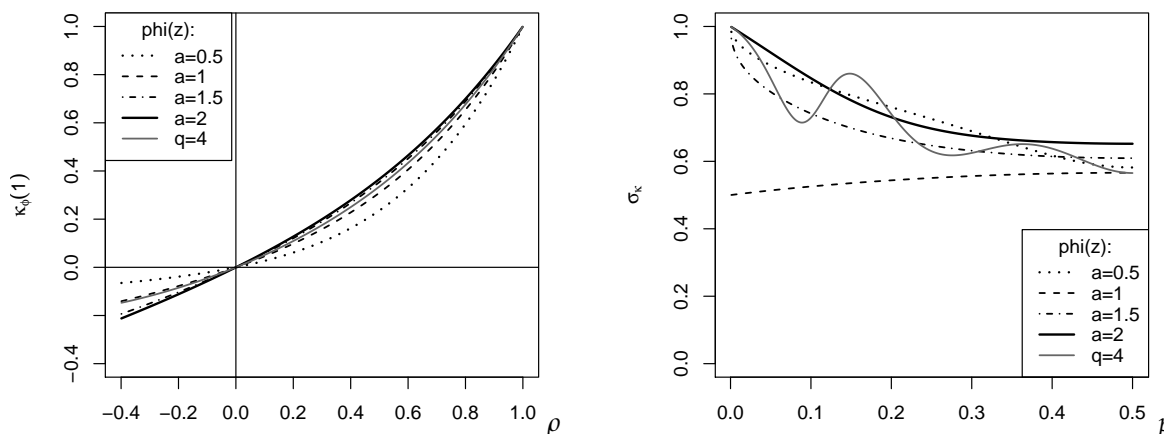


Figure 4. Plots for specific cases of $\phi_a(z) = (z - z^a)/(a - 1)$ and $\phi_q(z) = 1 - |2z - 1|^q$: $\kappa_\phi(1)$ against BAR(1)'s dependence parameter ρ with marginal $\text{Bin}(4, 0.3)$ (left); σ_κ against p of marginal $\text{Bin}(4, 0.3)$ (right).

In practice, the sample version of this measure, $\hat{\kappa}_\phi(h)$, is particularly important, where the cumulative (bivariate) probabilities are replaced by the corresponding relative frequencies. For uncovering significant deviations from the null hypothesis of serial independence (then, $\kappa_\phi(h) = 0$), the asymptotic distribution of $\hat{\kappa}_\phi(h)$ under the null of i.i.d. time series data is required. For its derivation, we proceed in an analogous way as in Section 3. As the starting point, we have to extend the asymptotics of the marginal sample CDF in (8) by also considering the bivariate sample CDF $\hat{f}_{ii}(h)$. Let $f_{(h)} = (f_0, \dots, f_{m-1}, f_{00}(h), \dots, f_{m-1, m-1}(h))^\top$, and denote its sample version by $\hat{f}_{(h)}$. Then, under the same mixing conditions as in Section 3, Weiß [14] established the asymptotic normality

$$\sqrt{n} (\hat{f}_{(h)} - f_{(h)}) \overset{d}{\rightarrow} N(\mathbf{0}, \Sigma^{(h)}) \quad \text{with } \Sigma^{(h)} = (\sigma_{i,j}^{(h)})_{i,j=0,\dots,2m-1} \tag{16}$$

and he derived general expressions for the asymptotic (co)variances $\sigma_{i,j}^{(h)}$. Analogous results for the case of missing data can be found in Supplement S.3 of Weiß [15]. For the present task, the asymptotics of the i. i. d.-case are sufficient. Then, $f_{ii}(h) = f_i^2$ for all $i = 0, \dots, m - 1$, and the covariances in (16) are given by

$$\begin{aligned} \sigma_{i,j}^{(h)} = \sigma_{i,j} &= f_{\min\{i,j\}} - f_i f_j \quad (\text{see (8)}), \\ \sigma_{i,m+j}^{(h)} &= 2 f_j (f_{\min\{i,j\}} - f_i f_j), \quad \text{for } i, j \in \{0, \dots, m - 1\}, \\ \sigma_{m+i,m+j}^{(h)} &= (f_{\min\{i,j\}} + 3 f_i f_j) (f_{\min\{i,j\}} - f_i f_j) \end{aligned} \tag{17}$$

see Weiß [14] (as well as p. 8 in Supplement S.3 of Weiß [15] if being concerned with missing data).

Next, we derive the asymptotics of $\hat{\kappa}_\phi(h)$ under the i. i. d.-null, and this requires to derive the second-order Taylor expansion for $\hat{\kappa}_\phi(h)$; details are postponed to Appendix A.1. As higher-order derivatives of ϕ , which are initially used while deriving a bias correction of $\hat{\kappa}_\phi(h)$, cancel out, the final result still relies on derivatives of ϕ up to order 2 only.

Theorem 2. *Under the null hypothesis of i. i. d. data, i.e., if $\kappa_\phi(h) = 0$ for all lags $h \in \mathbb{N}$, and assuming the EGF ϕ to be twice differentiable, it holds that*

$$\begin{aligned} \sqrt{n} \left(\hat{\kappa}_\phi(h) - \kappa_\phi(h) \right) &\xrightarrow{d} N(0, \sigma_\kappa^2) \quad \text{with } \sigma_\kappa^2 = \sum_{j,k=0}^{m-1} u_j u_k (f_{\min\{j,k\}} - f_j f_k)^2, \\ \text{where } u_j &= \frac{\phi''(f_j) + \phi''(1 - f_j)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1 - f_i)) f_i(1 - f_i)}. \end{aligned}$$

In addition, the bias-corrected mean of $\hat{\kappa}_\phi(h)$ is $E[\hat{\kappa}_\phi(h)] \approx -\frac{1}{n}$.

Note that we have a unique bias correction for any of the measures $\hat{\kappa}_\phi(h)$, independent of the choice of the EGF ϕ . Thus, for application in practice, it remains to compute the asymptotic variance σ_κ^2 in Theorem 2. This only requires knowledge about $\phi''(z)$ to evaluate the u_j , but not about higher-order derivatives of the EGF ϕ (see Example 3 for illustration). Further examples are plotted in the right part of Figure 4, where σ_κ was computed for the marginal distribution $\text{Bin}(4, p)$. The oscillating behavior of σ_κ for $\phi_q(z)$ with $q = 4$ is quite striking. It is also interesting to note that among the plotted κ -measures, the novel $\hat{\kappa}_{\text{ord}}^*(h)$ (case $a = 1$) has the lowest variance.

Example 3. *While we have a unique bias correction for $\hat{\kappa}_\phi(h)$, the asymptotic variance σ_κ^2 according to Theorem 2 differs for different choices of the EGF ϕ , as the involved u_j depend on $\phi''(z)$. For example,*

- for $\phi_a(z) = \frac{z-z^a}{a-1}$, we have $\phi_a''(z) = -a z^{a-2}$ according to Example 1,
- while for $\phi_q(z) = 1 - |2z - 1|^q$ with $q \geq 2$, we have $\phi_q''(z) = -4q(q - 1) |2z - 1|^{q-2}$.

Important special cases are:

- For $\phi(z) = z(1 - z)$, i.e., for the basic $\kappa_{\text{ord}}(h)$ according to (9), we have

$$\phi''(z) = -2, \quad \text{so } u_j = \left(\sum_{i=0}^{m-1} f_i(1 - f_i) \right)^{-1}.$$

Thus, we get

$$\sigma_\kappa^2 = \frac{\sum_{j,k=0}^{m-1} (f_{\min\{j,k\}} - f_j f_k)^2}{\left(\sum_{i=0}^{m-1} f_i(1 - f_i) \right)^2};$$

see Theorem 7.2.1 in Weiß [14].

- For $\phi(z) = -z \ln z$, i.e., for the novel $\kappa_{\text{ord}}^*(h)$ according to (14), we have

$$\phi''(z) = -\frac{1}{z}, \quad \text{so } u_j = \frac{1}{m} \frac{1}{f_j(1-f_j)}.$$

Thus, we get

$$\sigma_{\kappa}^2 = \frac{1}{m^2} \sum_{j,k=0}^{m-1} \frac{(f_{\min\{j,k\}} - f_j f_k)^2}{f_j(1-f_j) f_k(1-f_k)} = \frac{1}{m} + \frac{2}{m^2} \sum_{j < k} \frac{f_j(1-f_k)}{f_k(1-f_j)}.$$

For any other choice of $\phi_a(z)$ and $\phi_q(z)$, σ_{κ}^2 is easily computed using the aforementioned expressions for $\phi_a''(z)$ and $\phi_q''(z)$ from Example 1. Since the obtained closed-form formulae do not much simplify, further details are omitted here.

5. Simulation Results

In what follows, we discuss results from a simulation study, being tabulated in Appendix B, where 10^4 replications per scenario were used throughout. In view of our previous findings, achieved when discussing the asymptotics plotted in Figures 3 and 4, we do not further consider the choice $a = 1/2 < 1$ for the EGF ϕ_a , but we use $a = 5/2 > 2$ instead. The latter choice, in turn, was not presented before as the resulting asymptotic curves could hardly be distinguished from the case $a = 2$. So, altogether, $a = 1, 3/2, 2, 5/2$ as well as $q = 4$ were taken into account for simulations. The ordinal data were generated via binomial rank counts, $X_t = s_{I_t}$ with $I_t \sim \text{Bin}(m, p)$, which either exhibit serial dependence caused by a BAR(1) DGP with dependence parameter ρ , or which are i. i. d. (corresponding to $\rho = 0$).

Let us start with the ordinal dispersion measures $\widehat{\text{CPE}}_{\phi}$. Table A1 presents the simulated means (upper part) and SEs (lower part) for the case of i. i. d. ordinal data, and these are compared to the asymptotic values obtained from Theorem 1. Generally, we have an excellent agreement between simulated and asymptotic values, i.e., the derived asymptotic approximation to the true distribution of $\widehat{\text{CPE}}_{\phi}$ works well in practice. This is even more remarkable as also the low sample size $n = 50$ is included. There is a somewhat larger deviation only for the mean in the case $a = 1$, i.e., for the CPE (5), if $n \leq 100$ and $p = 0.1$. In this specific case, the simulated sample distribution might be quite close to a one-point distribution, which might cause computational issues for (5); recall that the convention $0 \ln 0 = 0$ has to be used. However, as the approximation quality is good throughout, a pivotal argument for the choice of ϕ in practice might be that the least SEs are observed if using ϕ_a with $a = 3/2, 2, 5/2$.

Table A2 considers exactly the same marginals as before, but now in the presence of additional serial dependence ($\rho = 0.4$). Comparing Tables A1 and A2, it becomes clear that the additional dependence causes increased bias and SE. However, and this is the crucial point for practice, the asymptotic approximations from Theorem 1 work as well as they do in the i. i. d.-case. If there are visible deviations at all, then these happen again mainly for $p = 0.1$ and low sample sizes. Overall, we have an excellent approximation quality throughout, but with least SEs again for $a = 3/2, 2, 5/2$.

While the CPE_{ϕ} -type dispersion measures and their asymptotics perform well, essentially for any choice of ϕ , the gap becomes wider when looking at the serial dependence measure $\hat{\kappa}_{\phi}(h)$. The asymptotics in Theorem 2 refer to the i. i. d.-case, which is used as the null hypothesis (H_0) if testing for significant serial dependence. Thus, let us start by investigating again the mean and SE of $\hat{\kappa}_{\phi}(1)$ for i. i. d. data (same DGPs as in Table A1); see the results in Table A3. For the asymptotic mean, we have the unique approximation $-1/n$, and this works well except for $p = 0.1$ and low sample sizes. In particular, for ϕ_a with $a = 1$, i.e., for $\hat{\kappa}_{\text{ord}}^*(1)$, we get notable deviations. The reason is given by the computation of $\hat{\kappa}_{\text{ord}}^*(1)$, where division by zero might happen (in the simulations, this was circumvented by replacing a zero by 10^{-6}). In a weakened form, we observe a similar issue for the case $a = 3/2$; generally, we are faced with the zero problem if $a < 2$ because of the second-order

derivatives of $\phi_a(z)$. Analogous deviations are observed for the SEs. Here, generally, the simulated SEs tend to be larger than the asymptotic ones. As a consequence, if using the asymptotic SEs for calculating the critical values when testing H_0 , we expect a tendency to oversizing.

If looking at the simulated rejection rates in Table A4, first at the size values ($\rho = 0$) being printed in italic font, we indeed see sizes being slightly larger than the nominal 5%-level, as long as $n \leq 100$. For larger sample sizes, by contrast, the $\hat{\kappa}_\phi(1)$ -test satisfies the given size requirement quite precisely. Here, we computed the critical values by plugging-in the respective sample CDF \hat{f} into the formula for σ_κ^2 . In Table A4, power values for $\rho \neq 0$ are also shown. Note that for a BAR(1) process, ρ can take any positive value in $(0; 1)$, but the attainable range of negative values is bounded from below by $\max\{-\frac{1-p}{p}, -\frac{p}{1-p}\}$ [2]. Thus, only $\rho = -0.4, -0.2$ are considered in Table A4. Generally, the $\hat{\kappa}_\phi(1)$ -tests are powerful regarding both positive and negative dependencies, but the actual power performance differs a lot for different ϕ . The worst power is observed for ϕ_q with $q = 4$, followed by ϕ_a with $a = 1$. Regarding the remaining ϕ_a -cases, $a = 3/2$ does slightly worse than $a = 2, 5/2$, and we often have a slight advantage for $a = 5/2$, especially for negative dependencies.

To sum up, while the whole $(CPE_\phi, \kappa_\phi(h))$ -family is theoretically appealing, and while there are hardly any noteworthy problems with the sample dispersion measures \widehat{CPE}_ϕ , the performance of $\hat{\kappa}_\phi(h)$ clearly depends on the choice of ϕ . It is recommended to use the family of a -entropies (6), and there, $a \geq 2$ is preferable. The measure $\hat{\kappa}_{\text{ord}}^*(1)$ from (14), for example, although theoretically appealing as a combination of entropy and extropy, has a relatively bad finite-sample performance. The probably most well-known pair, $(\text{IOV}, \kappa_{\text{ord}}(h))$, has a good performance, although there appears to be a slight advantage if choosing a somewhat larger than 2, such as $a = 5/2$ (recall that $a = 3$ leads back to the case $a = 2$).

6. Data Application

Ordinal time series are observed in quite diverse application areas. Economic examples include time series on credit ratings [14] or on fear states at the stock market [20], and a climatological example is the level of cloudiness of the sky [23]. Health-related examples are time series of electroencephalographic (EEG) sleep states [24], the pain severity of migraine attacks, and the level of perceived stress [15]. In this section, we are concerned with an environmental application, namely the level of air quality. Different definitions of air quality have been reported in the literature. In Chen & Chiu [25], the air quality index (AQI) is used for expressing the daily air quality, with levels ranging from $s_0 = \text{“good”}$ to $s_5 = \text{“hazardous”}$. Another case study is reported by Liu et al. [26], who use the classification of the Chinese government, which again distinguishes $m + 1 = 6$ levels, but now ranging from $s_0 = \text{“excellent”}$ to $s_5 = \text{“severely polluted”}$. The latter article investigates daily time series from thirty Chinese cities for the period December 2013–July 2019, i.e., the sample size equals $n = 2068$. In what follows, we use one of the time series studied by Liu et al. [26], namely the daily air quality levels x_1, \dots, x_n in Shanghai, for illustrating our novel results about cumulative paired ϕ -entropies.

The considered time series is plotted in the top panel of Figure 5. The bottom left graph shows the sample version of the probability mass function (PMF) $P(X = s_i)$, i.e., the relative frequencies of the categories. It exhibits a unimodal shape with mode (=median) in $s_1 = \text{“good”}$. The serial dependence structure is analyzed in the bottom right graph, where $\hat{\kappa}_\phi(h)$ with a -entropy having $a = 5/2$ is used, as this is the recommended choice according to Section 5. All of the plotted $\hat{\kappa}_\phi(h)$ -values are significantly different from 0 at the 5%-level, where the critical values (plotted as dashed lines in Figure 5) are computed as $\{-0.029, 0.028\}$ according to Theorem 2 (and by plugging-in the sample CDF). We recognize a medium level of dependence ($\hat{\kappa}_\phi(1) \approx 0.378$), which quickly decreases with increasing time lag h , similar to an AR-type process.

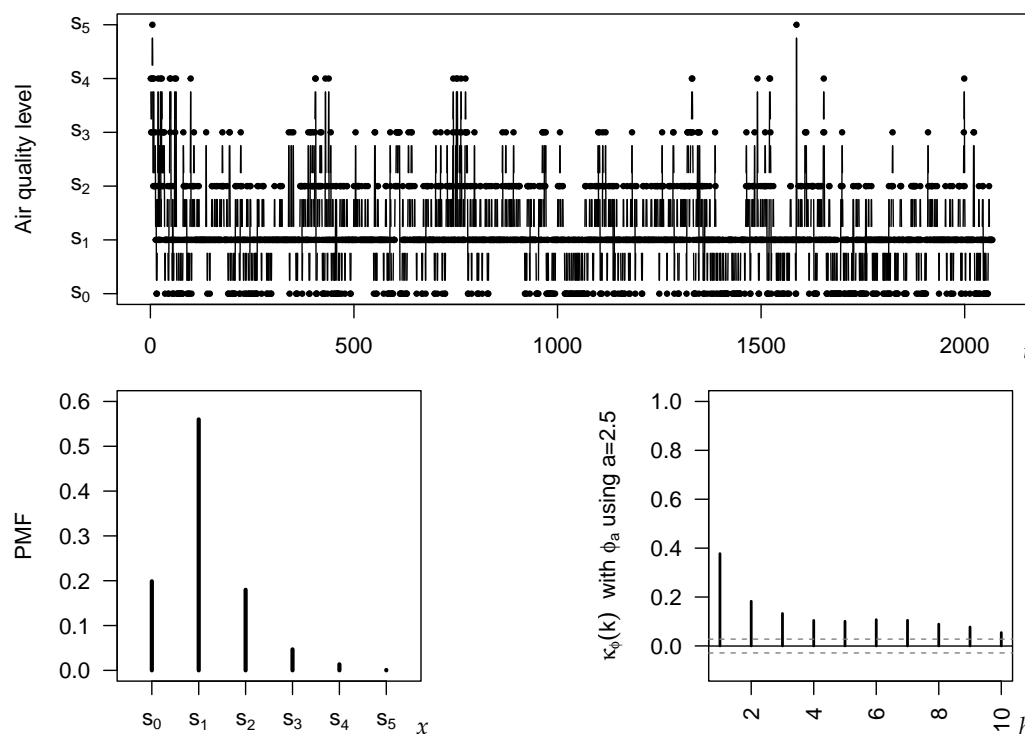


Figure 5. Daily air quality level in Shanghai: plot of time series (x_t) in top panel; plot of sample PMF (left) and $\hat{\kappa}_\phi(h)$ (right) in bottom panel; $\hat{\kappa}_\phi(h)$ uses a -entropy with $a = 5/2$.

Let us now have a closer look at the dispersion properties of the Shanghai series. The different choices of the CPE_ϕ -measure considered so far provide slightly different results regarding the extent of dispersion. In accordance with Figure 2, the largest point estimates are computed for $a = 1/2$ (0.514) and $q = 4$ (0.465), followed by $a = 1$ with 0.394, whereas $a = 3/2$ (0.349), $a = 2$ (0.332), and $a = 5/2$ (0.328) lead to similar but clearly lower values. Comparing the sample PMF in Figure 5 to the extreme scenarios of a one-point and an extreme two-point distribution, i.e., the lower ones among the above dispersion values seem to be more realistic here.

The novel asymptotics of Theorem 1 allow to judge the estimation uncertainty for the above point estimates. To keep the discussion simple, let us focus again on the case $a = 5/2$. In the first step, we compute the i. i. d.-approximations of bias and SE, $\frac{1}{2n} (\sum_{i=0}^{m-1} h_{ii} f_i (1 - f_i))$ and $n^{-1/2} \sigma_{\phi, \text{iid}}$, respectively. By plugging-in the sample CDF, these are obtained as $-1.54 \cdot 10^{-4}$ and $7.83 \cdot 10^{-3}$, respectively. However, these i. i. d.-results are misleading in the present example as the data exhibit significant serial dependence (recall Figure 5). As we know from Theorem 1, the bias has to be increased by the factor $1 + 2 \sum_{h=1}^\infty \kappa_\phi(h)$, and the SE by $(1 + 2 \sum_{h=1}^\infty \vartheta_\phi(h))^{1/2}$. These factors shall now be computed based on the so-called “ZOBPAR model” proposed by Liu et al. [26], which constitutes a rank-count approach, $X_t = s_{I_t}$. In view of the AR(1)-like dependence structure and the high frequency for s_1 , namely 0.560, the conditional distribution of $I_t | I_{t-1}, \dots$ is assumed to be a truncated Poisson distribution, truncated to the range $\{0, \dots, 5\}$, with time-varying Poisson parameter $\lambda_t = 0.3489 + 0.7594 I_{t-1}$ and additional one-inflation parameter 0.3463 ([26] Table III). For this model fit, we compute

$$1 + 2 \sum_{h=1}^\infty \kappa_\phi(h) \approx 1.907, \quad \sqrt{1 + 2 \sum_{h=1}^\infty \vartheta_\phi(h)} \approx 1.343.$$

Thus, an approximate 95 %-confidence interval (CI) for CPE_ϕ is given by $\approx (0.308; 0.349)$. CIs for the remaining CPE_ϕ -measures are computed analogously, leading to $(0.487; 0.544)$

for $a = 1/2$, to $(0.372; 0.418)$ for $a = 1$, to $(0.328; 0.371)$ for $a = 3/2$, to $(0.312; 0.353)$ for $a = 2$, and to $(0.439; 0.492)$ for $q = 4$.

7. Conclusions

In this article, we considered the family of cumulative paired ϕ -entropies. For each appropriate choice of the EGF ϕ , an ordinal dispersion measure $CPE_\phi(f)$ is implied. For example, particular choices from the families of a -entropies or q -entropies, respectively, lead to well-known dispersion measures from the literature. The first main contribution of this work was the derivation of the asymptotic distribution of the sample version \widehat{CPE}_ϕ for ordinal time series data. These asymptotics can be used to approximate the true distribution of \widehat{CPE}_ϕ , e.g., to compute approximate confidence intervals. Simulations showed that these asymptotics lead to an excellent finite-sample performance. Based on the obtained expression for the asymptotic bias of \widehat{CPE}_ϕ , we recognized that each EGF ϕ also implies a κ -type serial dependence measures, i.e., altogether, we have a matched pair $(CPE_\phi, \kappa_\phi(h))$ for each EGF ϕ . Again, we analyzed the asymptotics of the sample version $\hat{\kappa}_\phi(h)$, and these can be utilized for testing for significant serial dependence in the given ordinal time series. This time, however, the finite-sample performance clearly depends on the choice of ϕ . Choosing $\hat{\kappa}_\phi(h)$ based on an a -entropy with $a \geq 2$, such as $a = 5/2$, ensures good finite-sample properties. The practical application of the measures $(CPE_\phi, \kappa_\phi(h))$ and their asymptotics was demonstrated with an ordinal time series on the daily level of air quality in Shanghai.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data discussed in Section 6 were taken from Liu et al. [26] and are available at <https://doi.org/10.1111/jtsa.12625> (accessed on 22 December 2021).

Acknowledgments: The author thanks the two referees for their useful comments on an earlier draft of this article.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Proofs

Appendix A.1. Proof of Theorem 2

First, we derive the second-order Taylor expansion for $\hat{\kappa}_\phi(h)$. For this purpose, define $g : (0; 1)^{2^m} \rightarrow \mathbb{R}$ by mapping $\mathbf{y} = (y_0, \dots, y_{2^m-1})$ onto

$$g(\mathbf{y}) = \frac{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) (y_{m+i} - y_i^2)}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)}.$$

So, according to (15), $g(\mathbf{f}_{(h)}) = \kappa_\phi(h)$ and $g(\hat{\mathbf{f}}_{(h)}) = \hat{\kappa}_\phi(h)$. In fact, because of the i. i. d.-assumption, we even have $g(\mathbf{f}_{(h)}) = 0$.

For the intended Taylor expansion of $\hat{\kappa}_\phi(h)$, we have to compute all partial derivatives of $g(\mathbf{y})$ up to order 2, and to evaluate these derivatives in $\mathbf{y} = \mathbf{f}_{(h)}$, using that $f_{ii}(h) = f_i^2$ for all $i = 0, \dots, m - 1$ under the i. i. d.-assumption, and thus $g(\mathbf{f}_{(h)}) = 0$. We use the notations

$$u_j := \frac{\phi''(f_j) + \phi''(1 - f_j)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1 - f_i)) f_i(1 - f_i)}, \quad v_j := \frac{\phi'''(f_j) - \phi'''(1 - f_j)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1 - f_i)) f_i(1 - f_i)}. \tag{A1}$$

Note that if ϕ is symmetric around $1/2$, i.e., if $\phi(z) = \phi(1 - z)$, then u_j, v_j simplify because of $\phi''(f_j) + \phi''(1 - f_j) = 2\phi''(f_j)$ and $\phi'''(f_j) - \phi'''(1 - f_j) = 2\phi'''(f_j)$. Using (A1), we get for $0 \leq j, k \leq m - 1$ that

$$\begin{aligned}
 d_j^{(\kappa)} &:= \frac{\partial}{\partial y_j} g(\mathbf{f}(h)) = -2f_j u_j, \\
 d_{m+j}^{(\kappa)} &:= \frac{\partial}{\partial y_{m+j}} g(\mathbf{f}(h)) = u_j, \\
 h_{j,k}^{(\kappa)} &:= \frac{\partial^2}{\partial y_j \partial y_k} g(\mathbf{f}(h)) = 2f_j f_k ((1 - f_k) u_j v_k + (1 - f_j) u_k v_j) \\
 &\quad + 2(f_j + f_k - 4f_j f_k) u_j u_k - 2\delta_{j,k} (u_j + 2f_j v_j), \tag{A2} \\
 h_{j,m+k}^{(\kappa)} &:= \frac{\partial^2}{\partial y_j \partial y_{m+k}} g(\mathbf{f}(h)) = (\delta_{j,k} - f_j(1 - f_j) u_k) v_j - (1 - 2f_j) u_j u_k, \\
 h_{m+j,m+k}^{(\kappa)} &:= \frac{\partial^2}{\partial y_{m+j} \partial y_{m+k}} g(\mathbf{f}(h)) = 0.
 \end{aligned}$$

The proof of Equation (A2) is provided by Appendix A.2. We denote the gradient (Jacobian) of $g(\mathbf{f}(h))$ by $\mathbf{D}_\kappa = (d_0^{(\kappa)}, \dots, d_{2m-1}^{(\kappa)})$, and the Hessian equals $\mathbf{H}_\kappa = \text{diag}(h_{00}^{(\kappa)}, \dots, h_{2m-1,2m-1}^{(\kappa)})$.

Example A1. Let us pick up Example 1 and continue with the derivatives of the EGFs, as required for evaluating (A1) and (A2).

For $\phi_a(z) = \frac{z-z^a}{a-1}$, we have $\phi_a''(z) = -a z^{a-2}$, and thus $\phi_a'''(z) = -a(a-2) z^{a-3}$. Specific examples are

- $a \rightarrow 1$ corresponding to the EGF $\phi(z) = -z \ln z$, then

$$\phi''(z) = -\frac{1}{z}, \phi'''(z) = \frac{1}{z^2}, \quad \text{so } u_j = \frac{1}{m} \frac{1}{f_j(1-f_j)}, v_j = \frac{1}{m} \frac{2f_j-1}{f_j^2(1-f_j)^2};$$

- $a = 2$ corresponding to the EGF $\phi(z) = z(1 - z)$, then

$$\phi''(z) = -2, \phi'''(z) = 0, \quad \text{so } u_j = (\sum_{i=0}^{m-1} f_i(1 - f_i))^{-1}, v_j = 0;$$

- $a = 3$ corresponding to the EGF $\phi(z) = \frac{1}{2} z(1 - z^2)$, then

$$\phi''(z) = -3z, \phi'''(z) = -3, \quad \text{so } u_j, v_j \text{ are as before.}$$

For the equivalence of the $a = 2, 3$, recall (7).

For $\phi_q(z) = 1 - (1 - 2 \min\{z, 1 - z\})^q = 1 - |2z - 1|^q$ with $q > 3$, we get

$$\phi_q''(z) = -4q(q-1) |2z - 1|^{q-2}, \phi_q'''(z) = -8q(q-1)(q-2) |2z - 1|^{q-3} \text{sgn}(2z - 1).$$

Recall that (A1) simplifies because of the symmetry of $\phi_q(z)$ around $1/2$.

Using (A2), the second-order Taylor expansion of $\hat{\kappa}_\phi(h)$ equals

$$\hat{\kappa}_\phi(h) \approx \kappa_\phi(h) + \mathbf{D}_\kappa (\hat{\mathbf{f}}(h) - \mathbf{f}(h)) + \frac{1}{2} (\hat{\mathbf{f}}(h) - \mathbf{f}(h))^\top \mathbf{H}_\kappa (\hat{\mathbf{f}}(h) - \mathbf{f}(h)). \tag{A3}$$

In analogy to Section 3, we now conclude that

$$\begin{aligned}
 \sqrt{n} (\hat{\kappa}_\phi(h) - \kappa_\phi(h)) &\xrightarrow{d} \mathbf{N}(0, \sigma_\kappa^2) \quad \text{with } \sigma_\kappa^2 = \mathbf{D}_\kappa \boldsymbol{\Sigma}^{(h)} \mathbf{D}_\kappa^\top, \\
 \text{where } \sigma_\kappa^2 &= \sum_{j,k=0}^{m-1} u_j u_k (f_{\min\{j,k\}} - f_j f_k)^2.
 \end{aligned} \tag{A4}$$

The proof of Equation (A4) is provided by Appendix A.3. σ_κ^2 is computed explicitly by plugging-in (A1).

An approximate bias correction is obtained from (A3) (see the analogous arguments in Section 3). Using that $h_{m+j,m+k}^{(\kappa)} = 0$, we get that

$$E[\hat{\kappa}_\phi(h)] \approx 0 + \frac{1}{2n} \sum_{j,k=0}^{m-1} (h_{j,k}^{(\kappa)} \sigma_{j,k} + 2h_{j,m+k}^{(\kappa)} \sigma_{j,m+k}^{(h)}) = -\frac{1}{n}, \tag{A5}$$

see the proof in Appendix A.4. Using (A4) and (A5), the derivation of Theorem 2 is complete.

Appendix A.2. Proof of Equation (A2)

For

$$g(\mathbf{y}) = \frac{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) (y_{m+i} - y_i^2)}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)},$$

we compute all partial derivatives up to order 2. For doing this, we have to require that the EGF ϕ is even four times differentiable. Then,

$$\frac{\partial}{\partial y_{m+j}} g(\mathbf{y}) = \frac{\phi''(y_j) + \phi''(1 - y_j)}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)} \quad \text{and} \quad \frac{\partial^2}{\partial y_{m+j} \partial y_{m+k}} g(\mathbf{y}) = 0 \quad \text{for all } 0 \leq j, k \leq m - 1.$$

Next, using the product and chain rule,

$$\frac{\partial}{\partial y_j} g(\mathbf{y}) = \frac{\frac{\partial}{\partial y_j} (\phi''(y_j) + \phi''(1 - y_j)) (y_{m+j} - y_j^2)}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)} - \frac{g(\mathbf{y}) \frac{\partial}{\partial y_j} (\phi''(y_j) + \phi''(1 - y_j)) y_j(1 - y_j)}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)}.$$

Here,

$$\begin{aligned} &\frac{\partial}{\partial y_j} (\phi''(y_j) + \phi''(1 - y_j)) (y_{m+j} - y_j^2) \\ &= (\phi'''(y_j) - \phi'''(1 - y_j)) (y_{m+j} - y_j^2) - 2y_j (\phi''(y_j) + \phi''(1 - y_j)), \end{aligned}$$

and

$$\begin{aligned} &\frac{\partial}{\partial y_j} (\phi''(y_j) + \phi''(1 - y_j)) y_j(1 - y_j) \\ &= (\phi'''(y_j) - \phi'''(1 - y_j)) y_j(1 - y_j) + (1 - 2y_j) (\phi''(y_j) + \phi''(1 - y_j)). \end{aligned}$$

Thus,

$$\begin{aligned} \frac{\partial}{\partial y_j} g(\mathbf{y}) &= \left(\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i) \right)^{-1} \\ &\quad \cdot \left((\phi'''(y_j) - \phi'''(1 - y_j)) (y_{m+j} - y_j^2 - g(\mathbf{y}) y_j(1 - y_j)) \right. \\ &\quad \left. - (\phi''(y_j) + \phi''(1 - y_j)) (2y_j + g(\mathbf{y}) (1 - 2y_j)) \right). \end{aligned}$$

Consequently,

$$\begin{aligned} \frac{\partial^2}{\partial y_j \partial y_{m+k}} g(\mathbf{y}) &= \frac{\delta_{j,k} (\phi'''(y_j) - \phi'''(1 - y_j))}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)} - \frac{\frac{\partial}{\partial y_{m+k}} g(\mathbf{y})}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1 - y_i)) y_i(1 - y_i)} \\ &\quad \cdot \left((\phi'''(y_j) - \phi'''(1 - y_j)) y_j(1 - y_j) + (1 - 2y_j) (\phi''(y_j) + \phi''(1 - y_j)) \right). \end{aligned}$$

Finally,

$$\frac{\partial^2}{\partial y_j \partial y_k} g(\mathbf{y}) = - \frac{\frac{\partial}{\partial y_j} g(\mathbf{y}) \frac{\partial}{\partial y_k} (\phi''(y_k) + \phi''(1-y_k)) y_k(1-y_k)}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1-y_i)) y_i(1-y_i)} + \left(\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1-y_i)) y_i(1-y_i) \right)^{-1} \cdot \frac{\partial}{\partial y_k} \left((\phi'''(y_j) - \phi'''(1-y_j)) (y_{m+j} - y_j^2 - g(\mathbf{y}) y_j(1-y_j)) - (\phi''(y_j) + \phi''(1-y_j)) (2y_j + g(\mathbf{y}) (1-2y_j)) \right) \Bigg\} =: A.$$

Here, we have for $k \neq j$ that

$$A = - \left(y_j(1-y_j) (\phi'''(y_j) - \phi'''(1-y_j)) + (1-2y_j) (\phi''(y_j) + \phi''(1-y_j)) \right) \frac{\partial}{\partial y_k} g(\mathbf{y}).$$

For $k = j$, we have that

$$\begin{aligned} A &= \frac{\partial}{\partial y_j} (\phi'''(y_j) - \phi'''(1-y_j)) (y_{m+j} - y_j^2 - g(\mathbf{y}) y_j(1-y_j)) \\ &\quad - \frac{\partial}{\partial y_j} (\phi''(y_j) + \phi''(1-y_j)) (2y_j + g(\mathbf{y}) (1-2y_j)) \\ &= (\phi^{(4)}(y_j) + \phi^{(4)}(1-y_j)) (y_{m+j} - y_j^2 - g(\mathbf{y}) y_j(1-y_j)) \\ &\quad - (\phi'''(y_j) - \phi'''(1-y_j)) y_j(1-y_j) \frac{\partial}{\partial y_j} g(\mathbf{y}) \\ &\quad - 2 (\phi'''(y_j) - \phi'''(1-y_j)) (2y_j + g(\mathbf{y}) (1-2y_j)) \\ &\quad - (\phi''(y_j) + \phi''(1-y_j)) (2-2g(\mathbf{y}) + (1-2y_j) \frac{\partial}{\partial y_j} g(\mathbf{y})). \end{aligned}$$

So, altogether,

$$\begin{aligned} A &= -\delta_{jk} \left(2 (\phi''(y_j) + \phi''(1-y_j)) (1-g(\mathbf{y})) + 2 (\phi'''(y_j) - \phi'''(1-y_j)) (2y_j + g(\mathbf{y}) (1-2y_j)) \right. \\ &\quad \left. - (\phi^{(4)}(y_j) + \phi^{(4)}(1-y_j)) (y_{m+j} - y_j^2 - g(\mathbf{y}) y_j(1-y_j)) \right) \\ &\quad - \left(y_j(1-y_j) (\phi'''(y_j) - \phi'''(1-y_j)) + (1-2y_j) (\phi''(y_j) + \phi''(1-y_j)) \right) \frac{\partial}{\partial y_k} g(\mathbf{y}). \end{aligned}$$

Hence,

$$\begin{aligned} \frac{\partial^2}{\partial y_j \partial y_k} g(\mathbf{y}) &= - \frac{\left(y_k(1-y_k) (\phi'''(y_k) - \phi'''(1-y_k)) + (1-2y_k) (\phi''(y_k) + \phi''(1-y_k)) \right) \frac{\partial}{\partial y_j} g(\mathbf{y})}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1-y_i)) y_i(1-y_i)} \\ &\quad - \frac{\left(y_j(1-y_j) (\phi'''(y_j) - \phi'''(1-y_j)) + (1-2y_j) (\phi''(y_j) + \phi''(1-y_j)) \right) \frac{\partial}{\partial y_k} g(\mathbf{y})}{\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1-y_i)) y_i(1-y_i)} \\ &\quad - \delta_{jk} \left(\sum_{i=0}^{m-1} (\phi''(y_i) + \phi''(1-y_i)) y_i(1-y_i) \right)^{-1} \\ &\quad \cdot \left(2 (\phi''(y_j) + \phi''(1-y_j)) (1-g(\mathbf{y})) + 2 (\phi'''(y_j) - \phi'''(1-y_j)) (2y_j + g(\mathbf{y}) (1-2y_j)) \right. \\ &\quad \left. - (\phi^{(4)}(y_j) + \phi^{(4)}(1-y_j)) (y_{m+j} - y_j^2 - g(\mathbf{y}) y_j(1-y_j)) \right). \end{aligned}$$

For the required Taylor expansion, we have to evaluate all these derivatives in $\mathbf{y} = \mathbf{f}_{(h)}$, using that $f_{ii}(h) = f_i^2$ for all $i = 0, \dots, m-1$, and thus $g(\mathbf{f}_{(h)}) = 0$. We use the notations introduced in (A1), i.e.,

$$u_j := \frac{\phi''(f_j) + \phi''(1-f_j)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)}, \quad v_j := \frac{\phi'''(f_j) - \phi'''(1-f_j)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)}.$$

Then, we get for $0 \leq j \leq m - 1$ that

$$d_j^{(\kappa)} := \frac{\partial}{\partial y_j} g(\mathbf{f}(h)) = \frac{0 - 2 f_j (\phi''(f_j) + \phi''(1-f_j))}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} = -2 f_j u_j,$$

$$d_{m+j}^{(\kappa)} := \frac{\partial}{\partial y_{m+j}} g(\mathbf{f}(h)) = u_j.$$

For the second-order derivatives, we get for $0 \leq j, k \leq m - 1$ that

$$\begin{aligned} h_{j,k}^{(\kappa)} &:= \frac{\partial^2}{\partial y_j \partial y_k} g(\mathbf{f}(h)) \\ &= \frac{\left(f_k(1-f_k) (\phi'''(f_k) - \phi'''(1-f_k)) + (1-2 f_k) (\phi''(f_k) + \phi''(1-f_k)) \right) d_j^{(\kappa)}}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} \\ &\quad - \frac{\left(f_j(1-f_j) (\phi'''(f_j) - \phi'''(1-f_j)) + (1-2 f_j) (\phi''(f_j) + \phi''(1-f_j)) \right) d_k^{(\kappa)}}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} \\ &\quad - \frac{\delta_{jk} 2 \left((\phi''(f_j) + \phi''(1-f_j)) + 2 f_j (\phi'''(f_j) - \phi'''(1-f_j)) \right)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} \\ &= -f_k(1-f_k) v_k d_j^{(\kappa)} - (1-2 f_k) u_k d_j^{(\kappa)} \\ &\quad - f_j(1-f_j) v_j d_k^{(\kappa)} - (1-2 f_j) u_j d_k^{(\kappa)} - 2 \delta_{jk} u_j - 4 \delta_{jk} f_j v_j \\ &= -f_k(1-f_k) (-2 f_j u_j) v_k - (1-2 f_k) (-2 f_j u_j) u_k \\ &\quad - f_j(1-f_j) (-2 f_k u_k) v_j - (1-2 f_j) (-2 f_k u_k) u_j - 2 \delta_{jk} (u_j + 2 f_j v_j) \\ &= 2 f_j f_k (1-f_k) u_j v_k + 2 f_j f_k (1-f_j) u_k v_j \\ &\quad + 2 f_j (1-2 f_k) u_j u_k + 2 f_k (1-2 f_j) u_j u_k - 2 \delta_{jk} (u_j + 2 f_j v_j) \\ &= 2 f_j f_k ((1-f_k) u_j v_k + (1-f_j) u_k v_j) + 2 (f_j + f_k - 4 f_j f_k) u_j u_k - 2 \delta_{jk} (u_j + 2 f_j v_j). \end{aligned}$$

Similarly,

$$\begin{aligned} h_{j,m+k}^{(\kappa)} &:= \frac{\partial^2}{\partial y_j \partial y_{m+k}} g(\mathbf{f}(h)) = \frac{\delta_{j,k} (\phi'''(f_j) - \phi'''(1-f_j))}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} \\ &\quad - \frac{d_{m+k}^{(\kappa)} (\phi'''(f_j) - \phi'''(1-f_j)) f_j(1-f_j)}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} - \frac{d_{m+k}^{(\kappa)} (1-2 f_j) (\phi''(f_j) + \phi''(1-f_j))}{\sum_{i=0}^{m-1} (\phi''(f_i) + \phi''(1-f_i)) f_i(1-f_i)} \\ &= (\delta_{j,k} - d_{m+k}^{(\kappa)} f_j(1-f_j)) v_j - d_{m+k}^{(\kappa)} (1-2 f_j) u_j \\ &= (\delta_{j,k} - f_j(1-f_j) u_k) v_j - (1-2 f_j) u_j u_k. \end{aligned}$$

Finally,

$$h_{m+j,m+k}^{(\kappa)} := \frac{\partial^2}{\partial y_{m+j} \partial y_{m+k}} g(\mathbf{f}(h)) = 0.$$

This completes the proof of (A2).

Appendix A.3. Proof of Equation (A4)

$\sigma_\kappa^2 = \mathbf{D}_\kappa \Sigma^{(h)} \mathbf{D}_\kappa^\top$ is computed as follows:

$$\begin{aligned} \sigma_\kappa^2 &= \sum_{j,k=0}^{m-1} (d_j^{(\kappa)} d_k^{(\kappa)} \sigma_{j,k} + 2 d_j^{(\kappa)} d_{m+k}^{(\kappa)} \sigma_{j,m+k}^{(h)} + d_{m+j}^{(\kappa)} d_{m+k}^{(\kappa)} \sigma_{m+j,m+k}^{(h)}) \\ &\stackrel{(A2)}{=} \sum_{j,k=0}^{m-1} (4 f_j f_k u_j u_k \sigma_{j,k} - 4 f_j u_j u_k \sigma_{j,m+k}^{(h)} + u_j u_k \sigma_{m+j,m+k}^{(h)}) \\ &\stackrel{(17)}{=} \sum_{j,k=0}^{m-1} \left(-4 f_j f_k u_j u_k (f_{\min\{j,k\}} - f_j f_k) + u_j u_k (f_{\min\{j,k\}} + 3 f_j f_k) (f_{\min\{j,k\}} - f_j f_k) \right) \\ &= \sum_{j,k=0}^{m-1} u_j u_k (f_{\min\{j,k\}} - f_j f_k)^2. \end{aligned}$$

This completes the proof of (A4).

Appendix A.4. Proof of Equation (A5)

Using that $h_{m+j,m+k}^{(\kappa)} = 0$, we get that

$$\begin{aligned}
 E[\hat{\kappa}_\phi(h)] &\approx 0 + \frac{1}{2n} \sum_{j,k=0}^{m-1} (h_{j,k}^{(\kappa)} \sigma_{j,k} + 2h_{j,m+k}^{(\kappa)} \sigma_{j,m+k}^{(h)}) \\
 &\stackrel{(17)}{=} \frac{1}{2n} \sum_{j,k=0}^{m-1} \sigma_{j,k} (h_{j,k}^{(\kappa)} + 4f_k h_{j,m+k}^{(\kappa)}) \\
 &\stackrel{(A2)}{=} \frac{1}{2n} \sum_{j,k=0}^{m-1} \sigma_{j,k} \left(2f_j f_k ((1-f_k) u_j v_k + (1-f_j) u_k v_j) \right. \\
 &\quad \left. + 2(f_j + f_k - 4f_j f_k) u_j u_k - 2\delta_{j,k} (u_j + 2f_j v_j) \right. \\
 &\quad \left. + 4f_k (\delta_{j,k} - f_j(1-f_j) u_k) v_j - 4f_k (1-2f_j) u_j u_k \right) \\
 &= \frac{1}{2n} \sum_{j,k=0}^{m-1} \sigma_{j,k} \left(4f_j f_k (1-f_j) u_k v_j + 4f_k (1-2f_j) u_j u_k - 2\delta_{j,k} u_j - 4\delta_{j,k} f_j v_j \right. \\
 &\quad \left. + 4\delta_{j,k} f_j v_j - 4f_k f_j (1-f_j) u_k v_j - 4f_k (1-2f_j) u_j u_k \right) \\
 &= \frac{1}{2n} \sum_{j,k=0}^{m-1} \sigma_{j,k} (0 - 2\delta_{j,k} u_j) = -\frac{1}{n} \sum_{j=0}^{m-1} u_j \sigma_{j,j} = -\frac{1}{n},
 \end{aligned}$$

where the last equality follows from (A1) by using $\sigma_{j,j} = f_j(1-f_j)$ according to (17). This completes the proof of (A5).

Appendix B. Tables

Table A1. Simulated vs. asymptotic mean and SE of \widehat{CPE}_ϕ for specific cases of $\phi_a(z) = (z - z^a)/(a - 1)$ and $\phi_q(z) = 1 - |2z - 1|^q$, where i. i. d. rank counts $I_t \sim \text{Bin}(4, p)$ and sample size n .

| p | n | Simulated Mean | | | | | Asymptotic Mean | | | | |
|-----|------|----------------|-----------|---------|-----------|---------|-----------------|-----------|---------|-----------|---------|
| | | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ |
| 0.1 | 50 | 0.305 | 0.282 | 0.273 | 0.272 | 0.336 | 0.301 | 0.282 | 0.273 | 0.272 | 0.337 |
| | 100 | 0.310 | 0.285 | 0.276 | 0.274 | 0.341 | 0.308 | 0.285 | 0.276 | 0.274 | 0.341 |
| | 250 | 0.313 | 0.287 | 0.278 | 0.276 | 0.343 | 0.312 | 0.287 | 0.278 | 0.276 | 0.343 |
| | 500 | 0.314 | 0.288 | 0.278 | 0.277 | 0.344 | 0.314 | 0.288 | 0.278 | 0.277 | 0.344 |
| | 1000 | 0.314 | 0.288 | 0.279 | 0.277 | 0.344 | 0.315 | 0.288 | 0.279 | 0.277 | 0.344 |
| 0.3 | 50 | 0.538 | 0.500 | 0.484 | 0.481 | 0.610 | 0.538 | 0.500 | 0.484 | 0.481 | 0.610 |
| | 100 | 0.546 | 0.506 | 0.490 | 0.486 | 0.618 | 0.545 | 0.505 | 0.489 | 0.486 | 0.617 |
| | 250 | 0.550 | 0.508 | 0.492 | 0.488 | 0.622 | 0.550 | 0.509 | 0.492 | 0.489 | 0.622 |
| | 500 | 0.551 | 0.509 | 0.493 | 0.489 | 0.624 | 0.551 | 0.510 | 0.493 | 0.490 | 0.624 |
| | 1000 | 0.552 | 0.510 | 0.494 | 0.490 | 0.625 | 0.552 | 0.510 | 0.494 | 0.490 | 0.625 |
| 0.5 | 50 | 0.601 | 0.554 | 0.536 | 0.532 | 0.679 | 0.602 | 0.555 | 0.536 | 0.532 | 0.679 |
| | 100 | 0.609 | 0.560 | 0.541 | 0.537 | 0.687 | 0.609 | 0.561 | 0.541 | 0.537 | 0.688 |
| | 250 | 0.613 | 0.564 | 0.544 | 0.540 | 0.693 | 0.614 | 0.564 | 0.545 | 0.540 | 0.693 |
| | 500 | 0.616 | 0.566 | 0.546 | 0.542 | 0.696 | 0.615 | 0.565 | 0.546 | 0.542 | 0.695 |
| | 1000 | 0.616 | 0.566 | 0.546 | 0.542 | 0.696 | 0.616 | 0.566 | 0.546 | 0.542 | 0.696 |
| 0.1 | 50 | 0.048 | 0.044 | 0.043 | 0.043 | 0.053 | 0.049 | 0.044 | 0.043 | 0.043 | 0.054 |
| | 100 | 0.034 | 0.031 | 0.030 | 0.030 | 0.037 | 0.034 | 0.031 | 0.030 | 0.030 | 0.038 |
| | 250 | 0.022 | 0.020 | 0.019 | 0.019 | 0.024 | 0.022 | 0.020 | 0.019 | 0.019 | 0.024 |
| | 500 | 0.015 | 0.014 | 0.014 | 0.014 | 0.017 | 0.015 | 0.014 | 0.014 | 0.014 | 0.017 |
| | 1000 | 0.011 | 0.010 | 0.010 | 0.009 | 0.012 | 0.011 | 0.010 | 0.010 | 0.010 | 0.012 |
| 0.3 | 50 | 0.053 | 0.051 | 0.051 | 0.050 | 0.059 | 0.053 | 0.051 | 0.051 | 0.051 | 0.058 |
| | 100 | 0.037 | 0.036 | 0.036 | 0.036 | 0.041 | 0.037 | 0.036 | 0.036 | 0.036 | 0.041 |
| | 250 | 0.024 | 0.023 | 0.023 | 0.023 | 0.026 | 0.024 | 0.023 | 0.023 | 0.023 | 0.026 |
| | 500 | 0.017 | 0.016 | 0.016 | 0.016 | 0.018 | 0.017 | 0.016 | 0.016 | 0.016 | 0.018 |
| | 1000 | 0.012 | 0.011 | 0.011 | 0.011 | 0.013 | 0.012 | 0.011 | 0.011 | 0.011 | 0.013 |

Table A3. Cont.

| p | n | Simulated Mean | | | | | Asymptotic Mean | | | | |
|-----|------|----------------|-----------|---------|-----------|---------|-----------------|-----------|---------|-----------|---------|
| | | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ |
| 0.3 | 50 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 |
| | 100 | -0.010 | -0.010 | -0.010 | -0.010 | -0.011 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 |
| | 250 | -0.003 | -0.004 | -0.004 | -0.004 | -0.003 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 |
| | 500 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 |
| | 1000 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 |
| 0.5 | 50 | -0.020 | -0.021 | -0.021 | -0.021 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 | -0.020 |
| | 100 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 |
| | 250 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 |
| | 500 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 | -0.002 |
| | 1000 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 |
| 0.1 | 50 | 0.124 | 0.124 | 0.130 | 0.132 | 0.134 | 0.074 | 0.105 | 0.120 | 0.122 | 0.103 |
| | 100 | 0.134 | 0.082 | 0.088 | 0.089 | 0.085 | 0.053 | 0.074 | 0.085 | 0.086 | 0.073 |
| | 250 | 0.086 | 0.050 | 0.055 | 0.056 | 0.050 | 0.033 | 0.047 | 0.054 | 0.055 | 0.046 |
| | 500 | 0.045 | 0.034 | 0.038 | 0.039 | 0.034 | 0.023 | 0.033 | 0.038 | 0.039 | 0.033 |
| | 1000 | 0.022 | 0.024 | 0.027 | 0.027 | 0.023 | 0.017 | 0.023 | 0.027 | 0.027 | 0.023 |
| 0.3 | 50 | 0.098 | 0.097 | 0.101 | 0.101 | 0.101 | 0.079 | 0.089 | 0.096 | 0.097 | 0.088 |
| | 100 | 0.066 | 0.067 | 0.070 | 0.071 | 0.067 | 0.056 | 0.063 | 0.068 | 0.068 | 0.062 |
| | 250 | 0.040 | 0.040 | 0.043 | 0.043 | 0.040 | 0.035 | 0.040 | 0.043 | 0.043 | 0.040 |
| | 500 | 0.026 | 0.028 | 0.030 | 0.031 | 0.028 | 0.025 | 0.028 | 0.030 | 0.031 | 0.028 |
| | 1000 | 0.018 | 0.020 | 0.021 | 0.021 | 0.020 | 0.018 | 0.020 | 0.021 | 0.022 | 0.020 |
| 0.5 | 50 | 0.084 | 0.092 | 0.098 | 0.099 | 0.089 | 0.080 | 0.086 | 0.092 | 0.094 | 0.080 |
| | 100 | 0.058 | 0.063 | 0.067 | 0.068 | 0.059 | 0.057 | 0.061 | 0.065 | 0.066 | 0.057 |
| | 250 | 0.036 | 0.039 | 0.042 | 0.043 | 0.037 | 0.036 | 0.039 | 0.041 | 0.042 | 0.036 |
| | 500 | 0.025 | 0.027 | 0.029 | 0.030 | 0.025 | 0.025 | 0.027 | 0.029 | 0.030 | 0.025 |
| | 1000 | 0.018 | 0.019 | 0.021 | 0.021 | 0.018 | 0.018 | 0.019 | 0.021 | 0.021 | 0.018 |

Table A4. Simulated rejection rate ($\rho = 0$: size; $\rho \neq 0$: power) of $\hat{\kappa}_\phi(1)$ -test at 5%-level ($H_0: \rho = 0$, i.e., i. i. d.-case) for specific cases of $\phi_a(z) = (z - z^a)/(a - 1)$ and $\phi_q(z) = 1 - |2z - 1|^q$, where $\text{BAR}(1)$ rank counts $I_t \sim \text{Bin}(4, 0.3)$ and sample size n .

| p | n | Simulated Rejection Rate | | | | | Simulated Rejection Rate | | | | |
|------|-------|--------------------------|-----------|---------|-----------|---------------|--------------------------|-----------|---------|-----------|--------------|
| | | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ |
| | | | | | | $\rho = 0.2$ | | | | | |
| | | | | | | 50 | 0.227 | 0.265 | 0.263 | 0.262 | 0.222 |
| | | | | | | 100 | 0.367 | 0.447 | 0.451 | 0.450 | 0.370 |
| | | | | | | 250 | 0.721 | 0.830 | 0.830 | 0.829 | 0.723 |
| | | | | | | 500 | 0.950 | 0.982 | 0.982 | 0.981 | 0.947 |
| | | | | | | 1000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 |
| | | | | | | $\rho = -0.4$ | | | | | $\rho = 0.4$ |
| 50 | 0.484 | 0.565 | 0.585 | 0.588 | 0.312 | 50 | 0.663 | 0.749 | 0.756 | 0.757 | 0.625 |
| 100 | 0.807 | 0.878 | 0.891 | 0.893 | 0.672 | 100 | 0.917 | 0.959 | 0.962 | 0.962 | 0.897 |
| 250 | 0.989 | 0.999 | 1.000 | 1.000 | 0.987 | 250 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 |
| 500 | 0.997 | 1.000 | 1.000 | 1.000 | 1.000 | 500 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 1000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | | | | | $\rho = -0.2$ | | | | | $\rho = 0.6$ |
| 50 | 0.151 | 0.188 | 0.195 | 0.196 | 0.094 | 50 | 0.956 | 0.979 | 0.981 | 0.981 | 0.923 |
| 100 | 0.298 | 0.365 | 0.374 | 0.375 | 0.231 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| 250 | 0.643 | 0.744 | 0.751 | 0.751 | 0.584 | 250 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 500 | 0.925 | 0.965 | 0.966 | 0.965 | 0.900 | 500 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 1000 | 0.994 | 1.000 | 1.000 | 1.000 | 0.998 | 1000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Table A4. Cont.

| <i>p</i> | <i>n</i> | Simulated Rejection Rate | | | | | <i>q</i> = 4 | Simulated Rejection Rate | | | | |
|----------|----------|--------------------------|----------------|--------------|----------------|--------------|--------------|--------------------------|--------------|----------------|--------------|--|
| | | <i>a</i> = 1 | <i>a</i> = 1.5 | <i>a</i> = 2 | <i>a</i> = 2.5 | <i>a</i> = 1 | | <i>a</i> = 1.5 | <i>a</i> = 2 | <i>a</i> = 2.5 | <i>q</i> = 4 | |
| | | $\rho = 0$ | | | | | | $\rho = 0.8$ | | | | |
| 50 | 0.061 | 0.056 | 0.054 | 0.054 | 0.061 | 50 | 0.997 | 1.000 | 1.000 | 1.000 | 0.981 | |
| 100 | 0.056 | 0.060 | 0.058 | 0.057 | 0.057 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | |
| 250 | 0.047 | 0.048 | 0.046 | 0.046 | 0.049 | 250 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | |
| 500 | 0.049 | 0.050 | 0.050 | 0.050 | 0.052 | 500 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | |
| 1000 | 0.053 | 0.047 | 0.048 | 0.049 | 0.050 | 1000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | |

References

- Agresti, A. *Analysis of Ordinal Categorical Data*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010.
- Weiß, C.H. *An Introduction to Discrete-Valued Time Series*; John Wiley & Sons, Inc.: Chichester, UK, 2018.
- Blair, J.; Lacy, M.G. Measures of variation for ordinal data as functions of the cumulative distribution. *Percept. Mot. Ski.* **1996**, *82*, 411–418. [[CrossRef](#)]
- Blair, J.; Lacy, M.G. Statistics of ordinal variation. *Sociol. Methods Res.* **2000**, *28*, 251–280. [[CrossRef](#)]
- Gadrich, T.; Bashkansky, E. ORDANOVA: analysis of ordinal variation. *J. Stat. Plan. Inference* **2012**, *142*, 3174–3188. [[CrossRef](#)]
- Gadrich, T.; Bashkansky, E.; Zitikis, R. Assessing variation: A unifying approach for all scales of measurement. *Qual. Quant.* **2015**, *49*, 1145–1167. [[CrossRef](#)]
- Kiesl, H. *Ordinale Streuungsmaße—Theoretische Fundierung und statistische Anwendungen*; Josef Eul Verlag: Lohmar, Cologne, 2003. (In German).
- Kvålseth, T.O. Coefficients of variation for nominal and ordinal categorical data. *Percept. Mot. Ski.* **1995**, *80*, 843–847. [[CrossRef](#)]
- Kvålseth, T.O. Variation for categorical variables. In *International Encyclopedia of Statistical Science*; Lovric, M., ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 1642–1645.
- Kvålseth, T.O. The lambda distribution and its applications to categorical summary measures. *Adv. Appl. Stat.* **2011**, *24*, 83–106.
- Leik, R.K. A measure of ordinal consensus. *Pac. Sociol. Rev.* **1966**, *9*, 85–90. [[CrossRef](#)]
- Vogel, F.; Dobbener, R. Ein Streuungsmaß für comparative Merkmale (In German). *Jahrbücher für Natl. und Stat.* **1982**, *197*, 145–158.
- Weiß, C.H. On some measures of ordinal variation. *J. Appl. Stat.* **2019**, *46*, 2905–2926. [[CrossRef](#)]
- Weiß, C.H. Distance-based analysis of ordinal data and ordinal time series. *J. Am. Stat. Assoc.* **2020**, *115*, 1189–1200. [[CrossRef](#)]
- Weiß, C.H. Analyzing categorical time series in the presence of missing observations. *Stat. Med.* **2021**, *40*, 4675–4690. [[CrossRef](#)] [[PubMed](#)]
- Klein, I. Rangordnungsstatistiken als Verteilungsmaßzahlen für ordinalskalierte Merkmale: I. Streuungsmessung. *Rangordnungsstatistiken als Verteilungsmaßzahlen für ordinalskalierte Merkmale: I. Streuungsmessung*; Diskussionspapier No. 27/1999; Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Statistik und Ökonometrie: Nuremberg, Germany, 1999; Volume 27.
- Yager, R.R. Dissonance: A measure of variability for ordinal random variables. *Int. J. Uncertainty, Fuzziness -Knowl.-Based Syst.* **2001**, *9*, 39–53. [[CrossRef](#)]
- Klein, I.; Doll, M. (Generalized) maximum cumulative direct, residual, and paired Φ entropy approach. *Entropy* **2020**, *22*, 91. [[CrossRef](#)] [[PubMed](#)]
- Klein, I.; Mangold, B.; Doll, M. Cumulative paired ϕ -entropy. *Entropy* **2016**, *18*, 248. [[CrossRef](#)]
- Weiß, C.H. Measures of dispersion and serial dependence in categorical time series. *Econometrics* **2019**, *7*, 17. [[CrossRef](#)]
- Havrda, J.; Charvát, F. Quantification method of classification processes: concept of structural α -entropy. *Kybernetika* **1967**, *3*, 30–35.
- Rao, C.R. Convexity properties of entropy functions and analysis of diversity. *IMS Lect. Notes—Monogr. Ser.* **1984**, *5*, 68–77.
- Weiß, C.H. Regime-switching discrete ARMA models for categorical time series. *Entropy* **2020**, *22*, 458. [[CrossRef](#)] [[PubMed](#)]
- Stoffer, D.S.; Tyler, D.E.; Wendt, D.A. The spectral envelope and its applications. *Stat. Sci.* **2000**, *15*, 224–253. [[CrossRef](#)]
- Chen, C.W.S.; Chiu, L.M. Ordinal time series forecasting of the air quality index. *Entropy* **2021**, *23*, 1167. [[CrossRef](#)] [[PubMed](#)]
- Liu, M.; Zhu, F.; Zhu, K. Modeling normalcy-dominant ordinal time series: An application to air quality level. *J. Time Series Anal.* **2021**, forthcoming.