Open Access Repository
www.ssoar.info

# Measures of Dispersion and Serial Dependence in Categorical Time Series

Weiß, Christian H.

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

**Empfohlene Zitierung / Suggested Citation:**
Weiß, C. H. (2019). Measures of Dispersion and Serial Dependence in Categorical Time Series. *Econometrics*, *7*(2), 1-23. https://doi.org/10.3390/econometrics7020017

**Nutzungsbedingungen:**
*Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:*
*https://creativecommons.org/licenses/by/4.0/deed.de*

**Terms of use:**
*This document is made available under a CC BY Licence (Attribution). For more Information see:*
*https://creativecommons.org/licenses/by/4.0*

gesis
Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-71349-0

*econometrics*

MDPI

*Article*

# Measures of Dispersion and Serial Dependence in Categorical Time Series

**Christian H. Weiß**

Department of Mathematics and Statistics, Helmut Schmidt University, 22043 Hamburg, Germany; weissc@hsu-hh.de

Received: 21 December 2018; Accepted: 17 April 2019; Published: 22 April 2019

**Abstract:** The analysis and modeling of categorical time series requires quantifying the extent of dispersion and serial dependence. The dispersion of categorical data is commonly measured by Gini index or entropy, but also the recently proposed extropy measure can be used for this purpose. Regarding signed serial dependence in categorical time series, we consider three types of $\kappa$-measures. By analyzing bias properties, it is shown that always one of the $\kappa$-measures is related to one of the above-mentioned dispersion measures. For doing statistical inference based on the sample versions of these dispersion and dependence measures, knowledge on their distribution is required. Therefore, we study the asymptotic distributions and bias corrections of the considered dispersion and dependence measures, and we investigate the finite-sample performance of the resulting asymptotic approximations with simulations. The application of the measures is illustrated with real-data examples from politics, economics and biology.

**Keywords:** Cohen's $\kappa$; extropy; nominal variation; signed serial dependence; asymptotic distribution

## 1. Introduction

In many applications, the available data are not of quantitative nature (e.g., real numbers or counts) but consist of observations from a given finite set of categories. In the present article, we are concerned with data about political goals in Germany, fear states in the stock market, and phrases in a bird's song. For stochastic modeling, we use a *categorical* random variable $X$, i.e. a qualitative random variable taking one of a finite number of categories, e.g. $m+1$ categories with some $m \in \mathbb{N}$. If these categories are unordered, $X$ is said to be a *nominal* random variable, whereas an *ordinal* random variable requires a natural order of the categories (Agresti 2002). To simplify notations, we always assume the possible outcomes to be arranged in a certain order (either lexicographical or natural order), i.e. we denote the range (state space) as $\mathcal{S} = \{s_0, s_1, \ldots, s_m\}$. The stochastic properties of $X$ can be determined based on the vector of marginal probabilities by $\boldsymbol{p} = (p_0, \ldots, p_m)^\top \in [0;1]^{m+1}$, where $p_i = P(X = s_i)$ (probability mass function, PMF). We abbreviate $s_k(\boldsymbol{p}) := \sum_{j=0}^m p_j^k$ for $k \in \mathbb{N}$, where $s_1(\boldsymbol{p}) = 1$ has to hold. The subscripts "$0, 1, \ldots, m$" are used for $\mathcal{S}$ and $\boldsymbol{p}$ to emphasize that only $m$ of the probabilities can be freely chosen because of the constraint $p_0 = 1 - p_1 - \ldots - p_m$.

Well-established dispersion measures for quantitative data, such as variance or inter quartile range, cannot be applied to qualitative data. For a categorical random variable $X$, one commonly defines dispersion with respect to the uncertainty in predicting the outcome of $X$ (Kvålseth 2011b; Rao 1982; Weiß and Göb 2008). This uncertainty is maximal for a uniform distribution $\boldsymbol{p}_{\text{uni}} = (\frac{1}{m+1}, \ldots, \frac{1}{m+1})^\top$ on $\mathcal{S}$ (a reasonable prediction is impossible if all states are equally probable, thus maximal dispersion), whereas it is minimal for a one-point distribution $\boldsymbol{p}_{\text{one}}$ (i.e., all probability mass concentrates on one category, so a perfect prediction is possible). Obviously, categorical dispersion is just the opposite concept to the concentration of a categorical distribution. To measure the dispersion of the categorical random variable $X$, the most common approach is to use either the (normalized) *Gini index* (also *index of qualitative variation, IQV*) (Kvålseth 1995; Rao 1982) defined as

$$\nu_{\mathrm{G}} \;=\; \tfrac{m+1}{m}\,\big(1-s_2(\boldsymbol{p})\big), \tag{1}$$

or the (normalized) *entropy* (Blyth 1959; Shannon 1948) given by

$$\nu_{\mathrm{En}} \;=\; \tfrac{-1}{\ln(m+1)}\,\textstyle\sum_{i=0}^{m} p_i\,\ln p_i \qquad \text{with } 0\cdot\ln 0 := 0. \tag{2}$$

Both measures are minimized by a one-point distribution $\boldsymbol{p}_{\mathrm{one}}$ and maximized by the uniform distribution $\boldsymbol{p}_{\mathrm{uni}}$ on $\mathcal{S}$. While *nominal* dispersion is always expressed with respect to these extreme cases, it has to be mentioned that there is an alternative scenario of maximal *ordinal* variation, namely the extreme two-point distribution; however, this is not further considered here.

If considering a (stationary) categorical process $(X_t)_{\mathbb{Z}}$ instead of a single random variable, then not only marginal properties are relevant but also information about the serial dependence structure (Weiß 2018). The (signed) autocorrelation function (ACF), as it is commonly applied in case of real-valued processes, cannot be used for categorical data. However, one may use a type of *Cohen's* $\kappa$ instead (Cohen 1960). A $\kappa$-measure of signed serial dependence in categorical time series is given by (see Weiß 2011, 2013; Weiß and Göb 2008);

$$\kappa(h) \;=\; \sum_{i=0}^{m}\frac{p_{ii}(h)-p_i^2}{1-s_2(\boldsymbol{p})} \;\in\; \Big[\frac{-s_2(\boldsymbol{p})}{1-s_2(\boldsymbol{p})};1\Big] \qquad \text{for lags } h\in\mathbb{N}. \tag{3}$$

Equation (3) is based on the lagged bivariate probabilities $p_{ij}(h) = P(X_t = i, X_{t-h} = j)$ for $i,j = 0,\ldots,m$. $\kappa(h) = 0$ for serial independence at lag $h$, and the strongest degree of positive (negative) dependence is indicated if all $p_{ii}(h) = p_i$ ($p_{ii}(h) = 0$), i.e. if the event $X_{t-h} = s_i$ is necessarily followed by $X_t = s_i$ ($X_t \neq s_i$).

Motivated by a mobility index discussed by Shorrocks (1978), a simplified type of $\kappa$-measure, referred to as the *modified* $\kappa$, was defined by Weiß (2011, 2013):

$$\kappa^*(h) \;=\; \frac{1}{m}\sum_{i=0}^{m}\frac{p_{ii}(h)-p_i^2}{p_i} \;\in\; \Big[-\frac{1}{m};1\Big] \qquad \text{for lags } h\in\mathbb{N}. \tag{4}$$

Except the fact that the lower bound of the range differs from the one in Equation (3) (note that this lower bound is free of distributional parameters), we have the same properties as stated before for $\kappa(h)$. The computation of $\kappa^*(h)$ is simplified compared to the one of $\kappa(h)$ and, in particular, its sample version $\hat{\kappa}^*(h)$ has a more simple asymptotic normal distribution, see Section 5 for details. Unfortunately, $\kappa^*(h)$ is not defined if only one of the $p_i$ equals 0, whereas $\kappa(h)$ is well defined for any marginal distribution not being a one-point distribution. This issue may happen quite frequently for the sample version $\hat{\kappa}^*(h)$ if the given time series is short (a possible circumvention is to replace all summands with $p_i = 0$ by 0). For this reason, $\kappa^*(h), \hat{\kappa}^*(h)$ appear to be of limited use for practice as a way of quantifying signed serial dependence. It should be noted that a similar "zero problem" happens with the entropy $\nu_{\mathrm{En}}$ in Equation (2), and, actually, we work out a further relation between $\nu_{\mathrm{En}}$ and $\hat{\kappa}^*(h)$ below.

In the recent work by Lad et al. (2015), *extropy* was introduced as a complementary dual to the entropy. Its normalized version is given by

$$\nu_{\mathrm{Ex}} \;=\; \tfrac{-1}{m\,\ln\left(\frac{m+1}{m}\right)}\,\textstyle\sum_{i=0}^{m}\,(1-p_i)\,\ln(1-p_i). \tag{5}$$

Here, the zero problem obviously only happens if one of the $p_i$ equals 1 (i.e., in the case of a one-point distribution). Similar to the Gini index in Equation (1) and the entropy in Equation (2), the extropy takes its minimal (maximal) value 0 (1) for $\boldsymbol{p} = \boldsymbol{p}_{\mathrm{one}}$ ($\boldsymbol{p} = \boldsymbol{p}_{\mathrm{uni}}$), thus also Equation (5) constitutes a normalized measure of nominal variation. In Section 2, we analyze its properties in comparison to Gini index and entropy. In particular, we focus on the respective sample versions $\hat{\nu}_{\mathrm{Ex}}$, $\hat{\nu}_{\mathrm{G}}$ and $\hat{\nu}_{\mathrm{En}}$ (see Section 3). To be able to do statistical inference based on $\hat{\nu}_{\mathrm{Ex}}$, $\hat{\nu}_{\mathrm{G}}$ and $\hat{\nu}_{\mathrm{En}}$, knowledge about their distribution is required. Up to now, only the asymptotic distribution of $\hat{\nu}_{\mathrm{G}}$ and (to some part) of $\hat{\nu}_{\mathrm{En}}$ has been derived; in Section 3, comprehensive results for all considered dispersion measures are provided. These asymptotic distributions are then used as approximations to the true sample distributions of $\hat{\nu}_{\mathrm{Ex}}$, $\hat{\nu}_{\mathrm{G}}$ and $\hat{\nu}_{\mathrm{En}}$, which is further investigated with simulations and a real application (see Section 4).

The second part of this paper is dedicated to the analysis of serial dependence. As a novel competitor to the measures in Equations (3) and (4), a new type of modified $\kappa$ is proposed, namely

$$\kappa^{\star}(h) \;=\; \sum_{i=0}^{m} \frac{p_{ii}(h)-p_i^2}{1-p_i} \;\in\; \Big[\sum_{i=0}^{m} \frac{-p_i^2}{1-p_i};1\Big] \qquad \text{for lags } h\in\mathbb{N}. \tag{6}$$

Again, this constitutes a measure of signed serial dependence, which shares the before-mentioned (in)dependence properties with $\kappa(h), \kappa^*(h)$. However, in contrast to $\kappa^*(h)$, the newly proposed $\kappa^{\star}(h)$ does not have a division-by-zero problem: except for the case of a one-point distribution, $\kappa^{\star}(h)$ is well defined. Note that, in Section 3.2, it turns out that $\kappa^{\star}(h)$ is related to $\nu_{\mathrm{Ex}}$ in some sense, e.g. $\kappa(h)$ is related to $\nu_{\mathrm{G}}$ and $\kappa^*(h)$ to $\nu_{\mathrm{En}}$. In Section 5, we analyze the sample version of $\kappa^{\star}(h)$ in comparison to those of $\kappa(h), \kappa^*(h)$, and we derive its asymptotic distribution under the null hypothesis of serial independence. This allows us to test for significant dependence in categorical time series. The performance of this $\hat{\kappa}^{\star}$-test, in comparison to those based on $\hat{\kappa}, \hat{\kappa}^*$, is analyzed in Section 6, where also two further real applications are presented. Finally, we conclude in Section 7.

## 2. Extropy, Entropy and Gini Index

As extropy, entropy and Gini index all serve for the same task, it is interesting to know their relations and differences. An important practical issue is the "0 ln 0"-problem, as mentioned above, which never occurs for the Gini index, only occurs in the case of a (deterministic) one-point distribution for the extropy, and always occurs for the entropy if only one $p_i = 0$. Lad et al. (2015) further compared the non-normalized versions of extropy and entropy, and they showed that the first is never smaller than the latter. Actually, using the inequality $\ln(1+x) > x/(1+x/2)$ for $x > 0$ from Love (1980), it follows that

$$-\textstyle\sum_{i=0}^{m}(1-p_i)\,\ln(1-p_i) \;\geq\; -\sum_{i=0}^{m} p_i\,\ln p_i \;\geq\; 1-s_2(\boldsymbol{p}), \tag{7}$$

(see Appendix B.1 for further details).

Things change, however, if considering the normalized versions $\nu_{\mathrm{Ex}}$, $\nu_{\mathrm{En}}$ and $\nu_{\mathrm{G}}$. For illustration, assume an underlying Lambda distribution $L_m(\lambda)$ with $\lambda\in(0;1)$ defined by the probability vector $\boldsymbol{p}_{m;\lambda} = (1-\lambda+\frac{\lambda}{m+1}, \frac{\lambda}{m+1}, \ldots, \frac{\lambda}{m+1})^{\top}$ (Kvålseth 2011a). Note that $\lambda\to 0$ leads to a one-point distribution, whereas $\lambda\to 1$ leads to the uniform distribution; actually, $L_m(\lambda)$ can be understood as a mixture of these boundary cases. For $L_m(\lambda)$, the Gini index satisfies $\nu_{\mathrm{G}} = \lambda(2-\lambda)$ for all $m\in\mathbb{N}$ (see Kvålseth (2011a)). In addition, the extropy $\nu_{\mathrm{Ex}}$ has rather stable values for varying $m$ (see Figure 1a), whereas the entropy values in Figure 1b change greatly. This complicates the interpretation of the actual level of normalized entropy.

Finally, the example $m = 10$ plotted in Figure 1c shows that, in contrast to Equation (7), there is no fixed order between the normalized entropy $\nu_{\mathrm{En}}$ and Gini index $\nu_{\mathrm{G}}$. In this and many further numerical experiments, however, it could be observed that the inequalities $\nu_{\mathrm{Ex}} \geq \nu_{\mathrm{En}}$ and $\nu_{\mathrm{Ex}} \geq \nu_{\mathrm{G}}$ hold. These inequalities are formulated as a general conjecture here.

From now on, we turn towards the sample versions $\hat{\nu}_{\mathrm{Ex}}, \hat{\nu}_{\mathrm{En}}, \hat{\nu}_{\mathrm{G}}$ of $\nu_{\mathrm{Ex}}, \nu_{\mathrm{En}}, \nu_{\mathrm{G}}$. These are obtained by replacing the probabilities $p_i, \boldsymbol{p}$ by the respective estimates $\hat{p}_i, \hat{\boldsymbol{p}}$, which are computed as relative frequencies from the given sample data $x_1,\ldots,x_n$. As detailed in Section 3, $x_1,\ldots,x_n$ are assumed as time series data, but we also consider the case of independent and identically distributed (i.i.d.) data.

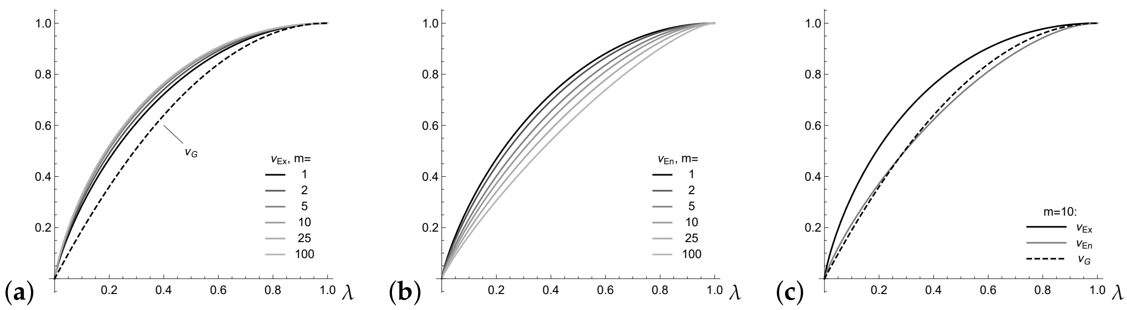


**Figure 1.** Normalized dispersion measures for Lambda distribution $L_m(\lambda)$ against $\lambda$: $\nu_{Ex}, \nu_G$ in (**a**), $\nu_{En}$ in (**b**), comparison for $m = 10$ in (**c**).

## 3. Distribution of Sample Dispersion Measures

To be able to derive the asymptotic distribution of statistics computed from $X_1, \ldots, X_n$, Weiß (2013) assumed that the nominal process is $\phi$-mixing with exponentially decreasing weights such that the CLT on p. 200 in Billingsley (1999) is applicable. This condition is not only satisfied in the i.i.d.-case, but also for, among others, the so-called NDARMA models as introduced by Jacobs and Lewis (1983) (see Appendix A.1 for details). Then, Weiß (2013) derived the asymptotic distribution of $\sqrt{n}\,(\hat{\boldsymbol{p}} - \boldsymbol{p})$, which is the normal distribution $\mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=0,\ldots,m}$ given by

$$\sigma_{ij} \;=\; p_j\,(\delta_{i,j} - p_i) \;+\; \textstyle\sum_{h=1}^{\infty} \big(p_{ij}(h) + p_{ji}(h) - 2\,p_i\,p_j\big). \tag{8}$$

Using this result, the asymptotic properties of $\hat{\nu}_G$, $\hat{\nu}_{En}$ and $\hat{\nu}_{Ex}$ can be derived, as shown in Appendix B.2. The following subsections present and compare these properties in detail.

### *3.1. Asymptotic Normality*

As shown in Appendix B.2, provided that $\boldsymbol{p} \neq \boldsymbol{p}_{one}, \boldsymbol{p}_{uni}$, all variation measures $\hat{\nu}_G$, $\hat{\nu}_{En}$ and $\hat{\nu}_{Ex}$ are asymptotically normally distributed. More precisely, $\sqrt{n}\,(\hat{\nu}_G - \nu_G)$ is asymptotically normally distributed with variance

$$\begin{aligned} \sigma_G^2 \;&=\; 4\,\big(\tfrac{m+1}{m}\big)^2\, \textstyle\sum_{i,j=0}^{m} p_i p_j\, \sigma_{ij} \\ &\overset{\text{Equation (8)}}{=}\; 4\,\big(\tfrac{m+1}{m}\big)^2\, \big(s_3(\boldsymbol{p}) - s_2^2(\boldsymbol{p})\big)\, \Bigg(1 \;+\; 2\,\sum_{h=1}^{\infty} \underbrace{\sum_{i,j=0}^{m} \frac{\big(p_{ij}(h) - p_i\,p_j\big)\, p_i p_j}{s_3(\boldsymbol{p}) - s_2^2(\boldsymbol{p})}}_{=:\vartheta(h)}\Bigg), \end{aligned} \tag{9}$$

a result already known from Weiß (2013). Here, $\vartheta(h)$ might be understood as a measure of serial dependence, in analogy to the measures in Equations (3), (4) and (6). In particular, $\vartheta(h) = 0$ in the i.i.d.-case, and $\vartheta(h) = \kappa(h)$ for NDARMA processes (Appendix A.1).

Analogously, $\sqrt{n}\,(\hat{\nu}_{En} - \nu_{En})$ is asymptotically normally distributed with variance

$$\begin{aligned} \sigma_{En}^2 \;&=\; \textstyle\sum_{i,j=0}^{m} \frac{1+\ln p_i}{\ln(m+1)}\, \frac{1+\ln p_j}{\ln(m+1)}\, \sigma_{ij} \\ &\overset{\text{Equation (8)}}{=}\; \frac{1}{(\ln(m+1))^2}\, \Big( \textstyle\sum_{i=0}^{m} p_i\,(\ln p_i)^2 - \big(\sum_{i=0}^{m} p_i\, \ln p_i\big)^2 \Big) \\ &\qquad \cdot \Bigg(1 \;+\; 2\,\sum_{h=1}^{\infty} \underbrace{\frac{\sum_{i,j=0}^{m} (1+\ln p_i)(1+\ln p_j)\,\big(p_{ij}(h) - p_i\,p_j\big)}{\sum_{i=0}^{m} p_i\,(\ln p_i)^2 - \big(\sum_{i=0}^{m} p_i\, \ln p_i\big)^2}}_{=:\vartheta^*(h)}\Bigg). \end{aligned} \tag{10}$$

In the i.i.d.-case, where the last factor becomes 1 (cf. Appendix A.1), this result was given by Blyth (1959), whereas the general expression in Equation (10) can be found in the work of Weiß (2013).

A novel result follows for the extropy, where $\sqrt{n}\,(\hat{\nu}_{\mathrm{Ex}} - \nu_{\mathrm{Ex}})$ is asymptotically normally distributed with variance

$$\begin{aligned} \sigma^2_{\mathrm{Ex}} &= \textstyle\sum_{i,j=0}^{m} \frac{1+\ln(1-p_i)}{m\,\ln(\frac{m+1}{m})}\,\frac{1+\ln(1-p_j)}{m\,\ln(\frac{m+1}{m})}\,\sigma_{ij} \\ &\overset{\text{Equation (8)}}{=} \frac{1}{\big(m\,\ln(\frac{m+1}{m})\big)^2}\,\Big(\textstyle\sum_{i=0}^{m} p_i\,\big(\ln(1-p_i)\big)^2 - \big(\sum_{i=0}^{m} p_i\,\ln(1-p_i)\big)^2\Big) \\ &\qquad \cdot\Big(1 + 2\sum_{h=1}^{\infty} \underbrace{\frac{\sum_{i,j=0}^{m} \big(1+\ln(1-p_i)\big)\big(1+\ln(1-p_j)\big)\,\big(p_{ij}(h) - p_i\,p_j\big)}{\sum_{i=0}^{m} p_i\,\big(\ln(1-p_i)\big)^2 - \big(\sum_{i=0}^{m} p_i\,\ln(1-p_i)\big)^2}}_{=:\vartheta^{\star}(h)}\Big). \end{aligned} \tag{11}$$

Again, the last factor becomes 1 in the i.i.d.-case as $\vartheta^{\star}(h) = 0$, and $\vartheta^{\star}(h) = \kappa(h)$ for NDARMA processes (Appendix A.1).

In Equations (9)–(11), the notations $\vartheta(h)$, $\vartheta^{*}(h)$ and $\vartheta^{\star}(h)$ have been introduced (see the respective expressions covered by the curly bracket) to highlight the similar structure of the asymptotic variances, and to locate the effect of serial dependence. Actually, one might use $\vartheta(h)$, $\vartheta^{*}(h)$ and $\vartheta^{\star}(h)$ as measures of serial dependence in categorical time series, although their definition is probably too complex for practical use. In Section 3.2, when analyzing the bias of $\hat{\nu}_{\mathrm{G}}$, $\hat{\nu}_{\mathrm{En}}$ and $\hat{\nu}_{\mathrm{Ex}}$, analogous relations to the $\kappa$-measures defined in Section 1 are established.

### 3.2. Asymptotic Bias

In Appendix B.2, we express the variation measures $\hat{\nu}_{\mathrm{G}}$, $\hat{\nu}_{\mathrm{En}}$ and $\hat{\nu}_{\mathrm{Ex}}$ as centered quadratic polynomials (at least approximately), and subsequently derive a bias formula. For the sample Gini index, it follows that

$$E[\hat{\nu}_{\mathrm{G}}] \approx \nu_{\mathrm{G}} - \tfrac{1}{n}\,\tfrac{m+1}{m}\,\textstyle\sum_{i=0}^{m}\sigma_{ii} \overset{\text{Equation (8)}}{=} \nu_{\mathrm{G}}\,\Big(1 - \tfrac{1}{n}\,\big(1 + 2\textstyle\sum_{h=1}^{\infty}\kappa(h)\big)\Big). \tag{12}$$

This formula was also derived by Weiß (2013), and it leads to the exact corrective factor $(1-\frac{1}{n})$ in the i.i.d.-case. For $\hat{\nu}_{\mathrm{En}}, \hat{\nu}_{\mathrm{Ex}}$, such bias formulae do not exist yet. However, from our derivations in Appendix B.2, we newly obtain that

$$E[\hat{\nu}_{\mathrm{En}}] \approx \nu_{\mathrm{En}} - \tfrac{1}{2n}\,\textstyle\sum_{i=0}^{m}\frac{p_i^{-1}}{\ln(m+1)}\,\sigma_{ii} \overset{\text{Equation (8)}}{=} \nu_{\mathrm{En}} - \tfrac{1}{2n}\,\tfrac{m}{\ln(m+1)}\,\big(1 + 2\textstyle\sum_{h=1}^{\infty}\kappa^{*}(h)\big). \tag{13}$$

In the i.i.d.-case, the last factor reduces to 1, and $\kappa^{*}(h) = \kappa(h)$ for NDARMA processes (Appendix A.1). Comparing Equations (12) and (13), we see that the effect of serial dependence on the bias is always expressed in terms of a $\kappa$-measure, using the ordinary $\kappa$ (Equation (3)) for the Gini index, and the modified $\kappa$ (Equation (4)) for the entropy. Concerning the extropy, it turns out that the newly proposed $\kappa$-measure from Equation (6) takes this role:

$$E[\hat{\nu}_{\mathrm{Ex}}] \approx \nu_{\mathrm{Ex}} - \tfrac{1}{2n}\,\textstyle\sum_{i=0}^{m}\frac{(1-p_i)^{-1}}{m\,\ln(\frac{m+1}{m})}\,\sigma_{ii} \overset{\text{Equation (8)}}{=} \nu_{\mathrm{Ex}} - \tfrac{1}{2n}\,\frac{1}{m\,\ln(\frac{m+1}{m})}\,\big(1 + 2\textstyle\sum_{h=1}^{\infty}\kappa^{\star}(h)\big). \tag{14}$$

In the i.i.d.-case, the last factor again reduces to 1, and also $\kappa^{\star}(h) = \kappa(h)$ holds for NDARMA processes (Appendix A.1). Altogether, Equations (12)–(14) show a unique structure regarding the effect of serial dependence. Furthermore, the computed bias corrections imply the relations $\nu_{\mathrm{G}} \leftrightarrow \kappa$, $\nu_{\mathrm{En}} \leftrightarrow \kappa^{*}$ and $\nu_{\mathrm{Ex}} \leftrightarrow \kappa^{\star}$. The sample versions of $\kappa, \kappa^{*}, \kappa^{\star}$ are analyzed later in Section 5.

### 3.3. Asymptotic Properties for Uniform Distribution

The asymptotic normality established in Section 3.1 certainly does not apply to the deterministic case $\boldsymbol{p} = \boldsymbol{p}_{\mathrm{one}}$, but we also have to exclude the boundary case of a uniform distribution $\boldsymbol{p}_{\mathrm{uni}}$. As shown in Appendix B.2, the asymptotic distribution of $\hat{\nu}_{\rm G}$, $\hat{\nu}_{\rm En}$ and $\hat{\nu}_{\rm Ex}$ in the uniform case is not a normal distribution but a quadratic-form one. All three statistics can be related to the Pearson's $\chi^2$-statistic:

$$\left.\begin{array}{rcl} n\,m\,(1-\hat{\nu}_{\rm G}) & = & \\ 2n\,\ln(m+1)\cdot(1-\hat{\nu}_{\rm En}) & \approx & \\ 2n\,m^2\,\ln\left(\frac{m+1}{m}\right)\cdot(1-\hat{\nu}_{\rm Ex}) & \approx & \end{array}\right\}\ n\sum_{j=0}^{m}\frac{(\hat{p}_j-\frac{1}{m+1})^2}{\frac{1}{m+1}}. \tag{15}$$

The actual asymptotic distribution can now be derived by applying Theorem 3.1 in Tan (1977) to the asymptotic result in Equation (8), which requires computing the eigenvalues of $\mathbf{\Sigma}$. In special cases, however, one is not faced with a general quadratic-form distribution but with a $\chi^2_m$-distribution; this happens for NDARMA processes and certainly in the i.i.d.-case (see Weiß (2013)). Then, defining $c = 1 + 2\sum_{h=1}^{\infty}\kappa(h)$, it holds that Equation (15) asymptotically follows $c$ times a $\chi^2_m$-distribution (with $c = 1$ in the i.i.d.-case).

## 4. Simulations and Applications

Section 4.1 presents some simulation results regarding the quality of the asymptotic approximations for $\hat{\nu}_{\rm G}$, $\hat{\nu}_{\rm En}$ and $\hat{\nu}_{\rm Ex}$ as derived in Section 3. Section 4.2 then applies these measures within a longitudinal study about the most important goals in politics in Germany.

### *4.1. Finite-Sample Performance of Dispersion Measures*

In applications, the normal and $\chi^2$-distributions derived in Section 3 were used as an approximation to the true distribution of $\hat{\nu}_{\rm G}$, $\hat{\nu}_{\rm En}$ and $\hat{\nu}_{\rm Ex}$. Therefore, the finite-sample performance of these approximations had to be analyzed, which was done by simulation (with 10,000 replications per scenario). In the tables provided by Appendix C, the simulated means (Table A1) and standard deviations (Table A2) for $\hat{\nu}_{\rm G}$, $\hat{\nu}_{\rm En}$ and $\hat{\nu}_{\rm Ex}$ are reported and compared to the respective asymptotic approximations. Then, a common application scenario was considered, the computation of two-sided 95% confidence intervals (CIs) for $\hat{\nu}_{\rm G}$, $\hat{\nu}_{\rm En}$ and $\hat{\nu}_{\rm Ex}$. Since the true parameter values are not known in practice, one has to plug-in estimated parameters into the formulae for mean and variance given in Section 3. The simulated coverage rates reported in Table A3 refer to such plug-in CIs. For these simulations, we either used an i.i.d. data-generating process (DGP) or an NDARMA DGP (see Appendix A.1): a DMA(1) process with $\varphi_1 = 0.25$ or a DAR(1) process with $\phi_1 = 0.40$. These were combined with the marginal distributions ($m = 3$) summarized in Table 1: $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ were used before by Weiß (2011, 2013), $\boldsymbol{p}_3$ by Kvålseth (2011b), and $\boldsymbol{p}_4$ to $\boldsymbol{p}_6$ are Lambda distributions $L_m(\lambda)$ with $\lambda \in \{0.25, 0.50, 0.75\}$ (see Kvålseth 2011a).

**Table 1.** Marginal distributions considered in Section 4.1 together with the corresponding dispersion values.

| PMF | $\nu_{\rm G}$ | $\nu_{\rm En}$ | $\nu_{\rm Ex}$ |
|---|---|---|---|
| $\boldsymbol{p}_1 = (0.2, 0.2, 0.25, 0.35)^\top$ | 0.980 | 0.979 | 0.988 |
| $\boldsymbol{p}_2 = (0.05, 0.1, 0.15, 0.7)^\top$ | 0.633 | 0.660 | 0.745 |
| $\boldsymbol{p}_3 = (0.2, 0.15, 0.05, 0.6)^\top$ | 0.767 | 0.767 | 0.848 |
| $\boldsymbol{p}_4 = (0.8125, 0.0625, 0.0625, 0.0625)^\top$ | 0.438 | 0.497 | 0.574 |
| $\boldsymbol{p}_5 = (0.625, 0.125, 0.125, 0.125)^\top$ | 0.750 | 0.774 | 0.832 |
| $\boldsymbol{p}_6 = (0.4375, 0.1875, 0.1875, 0.1875)^\top$ | 0.938 | 0.940 | 0.961 |

Finally, we used the results in Section 3.3 to test the null hypothesis (level 5%) of the uniform distribution $L_3(1)$ ($\nu_{\rm G} = \nu_{\rm En} = \nu_{\rm Ex} = 1$) based on $\hat{\nu}_{\rm G}$, $\hat{\nu}_{\rm En}$ and $\hat{\nu}_{\rm Ex}$. The considered alternatives were $L_m(\lambda)$ with $\lambda \in \{0.98, 0.96, 0.94, 0.92, 0.90\}$, and the DGPs were as presented above. The simulated rejection rates are summarized in Table A4. Note that the required factor $c = 1 + 2\sum_{h=1}^{\infty}\kappa(h)$ equald $c = (1+\phi_1)/(1-\phi_1)$ with $\phi_1 = \kappa(1)$ in the DAR(1) case, and $c = 1 + 2\,\kappa(1) = 1 + 2\,\varphi_1(1-\varphi_1)$ in the DMA(1) case, and was thus easy to estimate from the given time series data.

Let us now investigate the simulation results. Comparing the simulated mean values in Table A1 with the true dispersion values in Table 1, we realized a considerable negative bias for small sample sizes, which became even larger with increasing serial dependence. Fortunately, this bias is explained very well by the asymptotic bias correction, in any of the considered scenarios. With some limitations, this conclusion also applies to the standard deviations reported in Table A2; however, for sample size $n = 100$ and increasing serial dependence, the discrepancy between asymptotic and simulated values increased. As a result, the coverage rates in Table A3 performed rather poorly for sample size 100; thus, reliable CIs generally required a sample of size at least 250. It should also be noted that Gini index and extropy performed very similarly and often slightly worse than the entropy. Finally, the rejection rates in Table A4 concerning the tests for uniformity showed similar sizes (columns "1.00"; slightly above 0.05 for $n = 100$) but little different power values: best for extropy and worst for entropy.

### 4.2. Application: Goals in Politics

The monitoring of public mood and political attitudes over time is important for decision makers as well as for social scientists. Since 1980, the German General Social Survey ("ALLBUS") is carried out by the "GESIS—Leibniz Institute for the Social Sciences" in every second year (exception: there was an additional survey in 1991 after the German reunification). In the years before and including 1990, the survey was done only in Western Germany, but in all Germany for the years 1991 and later. In what follows, we consider the cumulative report for 1980–2016 in GESIS—Leibniz Institute for the Social Sciences (2018), and there the question "If you had to choose between these different goals, which one would seem to you personally to be the most important?". The four possible (nominal) answers are

- $s_0$: "To maintain law and order in this country";
- $s_1$: "To give citizens more influence on government decisions";
- $s_2$: "To fight rising prices"; and
- $s_3$: "To protect the right of freedom of speech".

The sample sizes of this longitudinal study varied between 2795 and 3754, and are thus sufficiently large for the asymptotics derived in Section 3. If just looking at the mode as a summary measure (location), there is not much change in time: from 1980 to 2000 and again in 2016, the majority of respondents considered $s_0$ (law and order) as the most important goal in politics, whereas $s_1$ (influence) was judged most important between 2002 and 2014.

Much more fluctuations are visible if looking at the dispersion measures in Figure 2a. Although the absolute values of the measures differ, the general shapes of the graphs are quite similar. Usually, the dispersion measures take rather large values ($\geq 0.90$ in most of the years), which shows that any of the possible goals is considered as being most important by a large part of the population. On the other hand, the different goals never have the same popularity. Even in 2008, where all measures give a value very close to 1, the corresponding uniformity tests lead to a clear rejection of the null of a uniform distribution ($p$-values approximately 0 throughout).

Let us now analyze the development of the importance of the political goals in some more detail, by looking at the extropy for illustration. Figure 2b shows the approximate 95%-CIs in time (a bias correction does not have a visible effect because of the large sample sizes). There are phases where successive CIs overlap, and these are interrupted by breaks in the dispersion behavior. Such breaks happen, e.g., in 1984 (possibly related to a change of government in Germany in 1982/83), in 2002 (perhaps related to 9/11), or in 2008 and 2010 (Lehman bankruptcy and economic crisis). These changes in dispersion go along with reallocations of probability masses, as can be seen from Figure 2c. From the frequency curves of $s_0$ and $s_1$, we can see the above-mentioned change in mode, where the switch back to $s_0$ (law and order) might be caused by the refugee crisis. In addition, the curve for $s_2$ (fight rising prices) helps for explanation, as it shows that $s_2$ is important for the respondents in the beginning of the 1980s (where Germany suffered from very high inflation) and in 2008 (economic crisis), but not otherwise. Thus, altogether, the dispersion measures together with their approximate CIs give a very good summary of the distributional changes over time.

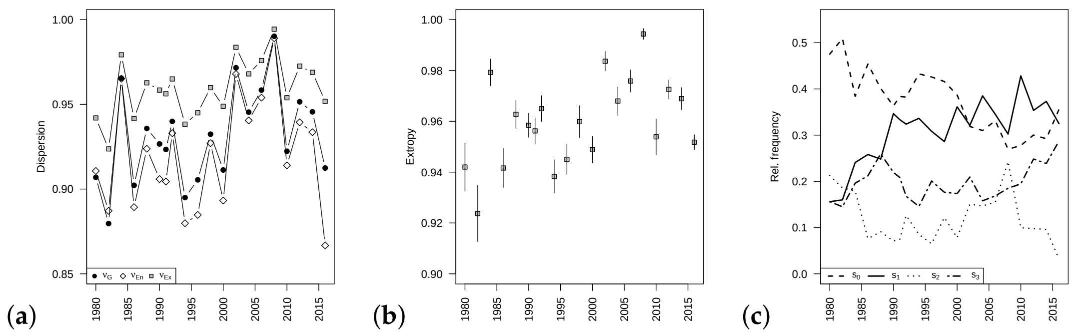


**Figure 2.** ALLBUS data from Section 4.2: (**a**) dispersion measures; (**b**) extropy with 95% CIs; and (**c**) relative frequencies; plotted against year of survey.

## 5. Measures of Signed Serial Dependence

After having discussed the analysis of marginal properties of a categorical time series, we now turn to the analysis of serial dependencies. In Section 1, two known measures of signed serial dependence, Cohen's $\kappa(h)$ in Equation (3) and a modification of it, $\kappa^*(h)$ in Equation (4), are briefly surveyed, and, in Section 3.2, we realize a connection to $\nu_{\mathrm{G}}$ and $\nu_{\mathrm{En}}$, respectively. Motivated by a zero problem with $\kappa^*(h)$, a new type of modified $\kappa$ is proposed in Equation (6), the measure $\kappa^\star(h)$, and this turns out to be related to $\nu_{\mathrm{Ex}}$.

If replacing the (bivariate) probabilities in Equations (3), (4) and (6) by the respective (bivariate) relative frequencies computed from $x_1, \ldots, x_n$, we end up with sample versions of these dependence measures. Knowledge of their asymptotic distribution is particularly relevant for the i.i.d.-case, because this allows us to test for significant serial dependence in the given time series. As shown by Weiß (2011, 2013), $\hat{\kappa}(h)$ then has an asymptotic normal distribution, and it holds approximately that

$$E\big[\hat{\kappa}(h)\big] \;\approx\; -\frac{1}{n}, \qquad V\big[\hat{\kappa}(h)\big] \;\approx\; \frac{1}{n}\left(1 \;-\; \frac{1+2s_3(\boldsymbol{p})-3s_2(\boldsymbol{p})}{(1-s_2(\boldsymbol{p}))^2}\right). \tag{16}$$

The sample version of $\kappa^*(h)$ has a more simple asymptotic normal distribution with Weiß (2011, 2013)

$$E\big[\hat{\kappa}^*(h)\big] \;\approx\; -\frac{1}{n}, \qquad V\big[\hat{\kappa}^*(h)\big] \;\approx\; \frac{1}{m\,n}, \tag{17}$$

but it suffers from the before-mentioned zero problem, especially for short time series.

Thus, it remains to derive the asymptotics of the novel $\hat{\kappa}^\star(h)$ under the null of an i.i.d. sample $X_1, \ldots, X_n$. The starting point is an extension of the limiting result in Equation (8). Under appropriate mixing assumptions (see Section 3), Weiß (2013) derived the joint asymptotic distribution of *all* univariate and equal-bivariate relative frequencies, i.e. of all $\sqrt{n}\,(\hat{p}_i - p_i)$ and $\sqrt{n}\,(\hat{p}_{jj}(h) - p_{jj}(h))$, which is the $2\,(m+1)$-dimensional normal distribution $\mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}^{(h)})$. The covariance matrix $\boldsymbol{\Sigma}^{(h)}$ consists of four blocks with entries

$$\begin{aligned}
\sigma_{i,j} \;&=\; p_j\,(\delta_{i,j}-p_i) \;+\; \textstyle\sum_{k=1}^{\infty}\,(p_{ij}(k)+p_{ji}(k)-2\,p_i\,p_j),\\
\sigma_{i,m+1+j}^{(h)} \;&=\; 2(\delta_{i,j}-p_i)\,p_{jj}(h) \;+\; \textstyle\sum_{k=1}^{\infty}\,\big(p_{ijj}(k,h)-p_i p_{jj}(h)\big)\\
&\quad+\; \textstyle\sum_{k=h+1}^{\infty}\,\big(p_{jji}(h,k-h)-p_i p_{jj}(h)\big)\\
&\quad+\; \textstyle\sum_{k=1}^{h-1}\,\big(p_{jij}(k,h-k)-p_i p_{jj}(h)\big) \;=\; \sigma_{m+1+j,i}^{(h)},\\
\sigma_{m+1+i,m+1+j}^{(h)} \;&=\; \big(\delta_{i,j}-p_{jj}(h)\big)\,p_{ii}(h) \;+\; 2\big(\delta_{i,j}\,p_{iijj}(h,h)-p_{ii}(h)p_{jj}(h)\big)\\
&\quad+\; \textstyle\sum_{k=1}^{h-1}\,\big(p_{jiji}(k,h-k,k)+p_{ijij}(k,h-k,k)-2p_{ii}(h)p_{jj}(h)\big)\\
&\quad+\; \textstyle\sum_{k=h+1}^{\infty}\,\big(p_{jjii}(h,k-h,h)+p_{iijj}(h,k-h,h)-2p_{ii}(h)p_{jj}(h)\big),
\end{aligned} \tag{18}$$

where always $i, j = 0, \ldots, m$, and where

$$p_{abc}(k,l) = P(X_t = a, X_{t-k} = b, X_{t-k-l} = c),$$
$$p_{abcd}(k,l,m) = P(X_t = a, X_{t-k} = b, X_{t-k-l} = c, X_{t-k-l-m} = d).$$

This rather complex general result simplifies greatly special cases such as an NDARMA-DGP (Weiß 2013) and, in particular, for an i.i.d. DGP:

$$\begin{aligned} \sigma_{i,j} &= p_j\,(\delta_{i,j} - p_i), \\ \sigma_{i,m+1+j} &= 2\,(\delta_{i,j} - p_i)\,p_j^2 \;=\; \sigma_{m+1+j,i}, \\ \sigma_{m+1+i,m+1+j} &= \delta_{i,j}\,p_i^2\,(1 + 2\,p_i) \;-\; 3\,p_i^2 p_j^2. \end{aligned} \tag{19}$$

Now, the asymptotic properties of $\hat{\kappa}^\star(h)$ can be derived, as done in Appendix B.3. $\sqrt{n}\,\big(\hat{\kappa}^\star(h) - \kappa^\star(h)\big)$ is asymptotically normally distributed, and mean and variance can be approximated by plugging Equation (18) into

$$\begin{aligned} E\big[\hat{\kappa}^\star(h)\big] &\approx \kappa^\star(h) + \frac{1}{n}\sum_{j=0}^{m} \frac{(1-p_j)\,\sigma^{(h)}_{j,m+1+j} - (1-p_{jj}(h))\,\sigma_{jj}}{(1-p_j)^3}, \\ V\big[\hat{\kappa}^\star(h)\big] &\approx \frac{1}{n}\sum_{i,j=0}^{m} \frac{p_{ii}(h) - 2p_i + p_i^2}{(1-p_i)^2}\,\frac{p_{jj}(h) - 2p_j + p_j^2}{(1-p_j)^2}\,\sigma_{i,j} \\ &\quad + \sum_{i,j=0}^{m} \frac{\sigma^{(h)}_{m+1+i,m+1+j}}{(1-p_i)(1-p_j)} + 2\sum_{i,j=0}^{m} \frac{p_{ii}(h) - 2p_i + p_i^2}{(1-p_i)^2}\,\frac{\sigma^{(h)}_{i,m+1+j}}{1-p_j}. \end{aligned} \tag{20}$$

In the i.i.d.-case, we simply have

$$E\big[\hat{\kappa}^\star(h)\big] \approx -\frac{1}{n}, \qquad V\big[\hat{\kappa}^\star(h)\big] \approx \frac{1}{n}\left(s_2(\boldsymbol{p}) - \sum_{i=0}^{m}\Big(\frac{p_i^2}{1-p_i}\Big)^2 + \Big(\sum_{i=0}^{m}\frac{p_i^2}{1-p_i}\Big)^2\right). \tag{21}$$

Comparing Equations (16), (17) and (21), we see that all three measures have the same asymptotic bias $-1/n$, but their asymptotic variances generally differ. An exception to the latter statement is obtained in the case of a uniform distribution, then also the asymptotic variances coincide (see Appendix B.3).

## 6. Simulations and Applications

Section 6.1 presents some simulation results, where the quality of the asymptotic approximations for $\hat{\kappa}(h), \hat{\kappa}^*(h), \hat{\kappa}^\star(h)$ according to Section 5 is investigated, as well as the power if testing against different types of serial dependence. Two real-data examples are discussed in Sections 6.2 and 6.3, first an ordinal time series with rather strong positive dependence, then a nominal time series exhibiting negative dependencies.

### *6.1. Finite-Sample Performance of Serial Dependence Measures*

In analogy to Section 4.1, we compared the finite-sample performance of the normal approximations in Equations (16), (17) and (21) via simulations[1]. As the power scenarios, we not only included NDARMA models but also the NegMarkov model described in Appendix A.2. These

[1] The "zero problem" for $\kappa^*(h)$ described after Equation (4) happened mainly for $n = 100$ and for distributions with low dispersion such as $\boldsymbol{p}_2$ to $\boldsymbol{p}_4$, in about 0.5% of the i.i.d. simulation runs. It increased with positive dependence, to about 2% for the DAR(1) simulation runs. This problem was circumvented by replacing all affected summands by 0.

models were combined with the marginal distributions $\boldsymbol{p}_1$ to $\boldsymbol{p}_6$ in Table 1 plus $\boldsymbol{p}_7 = \boldsymbol{p}_{\text{uni}}$; for the NegMarkov model, it was not possible to have the marginals $\boldsymbol{p}_2$ and $\boldsymbol{p}_4$ with very low dispersion. The full simulation results are available from the author upon request, but excerpts thereof are shown in the sequel to illustrate the main findings. As before, all tables are collected in Appendix C.

First, we discuss the distributional properties of $\hat{\kappa}(h), \hat{\kappa}^*(h), \hat{\kappa}^\star(h)$ under the null of an i.i.d. DGP. The unique mean approximation $-1/n$ worked very well without exceptions. The quality of the standard deviations' approximations is investigated in Table A5. Generally, the actual marginal dispersion was of great influence. For large dispersion (e.g., $\boldsymbol{p}_6, \boldsymbol{p}_7$ in Table A5), we had a very good agreement between asymptotic and simulated standard deviation, where deviations were typically not larger than $\pm 0.001$. For low dispersion (e.g., $\boldsymbol{p}_5$ in Table A5), we found some discrepancy for $n = 100$ and increasing $h$: the asymptotic approximation resulted in lower values for $\hat{\kappa}(h), \hat{\kappa}^\star(h)$ (discrepancy up to 0.004), and in larger values for $\hat{\kappa}^*(h)$ (discrepancy up to 0.002). Consequently, if testing for serial independence, we expect the size for $\hat{\kappa}^*(h)$ to be smaller than the nominal 5%-level, and to be larger for $\hat{\kappa}(h), \hat{\kappa}^\star(h)$. This was roughly confirmed by the results in Table A6 (and by further simulations), with the smallest size values for $\hat{\kappa}^*(h)$ (might be smaller by up to 0.01) and the largest for $\hat{\kappa}^\star(h)$. The sizes of $\hat{\kappa}(h)$ tended to be smaller than 0.05 for $h = 1$ (discrepancies $\leq 0.003$), while those of $\hat{\kappa}^*(h)$ were always rather close to 5%.

A more complex picture was observed with regard to the power of $\hat{\kappa}(h), \hat{\kappa}^*(h), \hat{\kappa}^\star(h)$ (see Table A6). For positive dependence (DMA(1) and DAR(1)), $\hat{\kappa}(h), \hat{\kappa}^\star(h)$ performed best if being close to a marginal uniform distribution, whereas $\hat{\kappa}^*(h)$ had superior power for lower marginal dispersion levels (and $\hat{\kappa}(h)$ performed second-best). For negative dependence (NegMarkov), in contrast, $\hat{\kappa}^\star(h)$ was the optimal choice, and $\hat{\kappa}^*(h)$ might have a rather poor power, especially for low dispersion. Thus, while $\hat{\kappa}(h), \hat{\kappa}^\star(h)$ both showed a more-or-less balanced performance with respect to positive and negative dependence, we had a sharp contrast for $\hat{\kappa}^*(h)$.

### *6.2. Application: Fear Index*

The Volatility Index (VIX) serves as a benchmark for U.S. stock market volatility, and increasing VIX values are interpreted as indications of greater fear in the market (Hancock 2012). Hancock (2012) distinguished between the $m + 1 = 13$ ordinal fear states given in Table 2. From the historical closing rates of the VIX offered by the website https://finance.yahoo.com/, a time series of daily fear states was computed for the $n = 4287$ trading days in the period 1990–2006 (before the beginning of the financial crisis). The obtained time series is plotted in Figure 3.

**Table 2.** Definition of fear states for Volatility Index (VIX) according to Hancock (2012).

| State and Explanation | | VIX | State and Explanation | | VIX |
|---|---|---|---|---|---|
| $s_0$ | extreme complacency | [0; 10) | $s_7$ | extremely high anx. | [40; 45) |
| $s_1$ | very low anx. = high compl. | [10; 15) | $s_8$ | near panic | [45; 50) |
| $s_2$ | low anx. = moderate compl. | [15; 20) | $s_9$ | moderate panic | [50; 55) |
| $s_3$ | moderate anx. = low compl. | [20; 25) | $s_{10}$ | panic | [55; 60) |
| $s_4$ | moderately high anxiety | [25; 30) | $s_{11}$ | intense panic | [60; 65) |
| $s_5$ | high anxiety | [30; 35) | $s_{12}$ | extreme panic | [65; 100] |
| $s_6$ | very high anxiety | [35; 40) | | | |

As shown in the plots in the top panel of Figure 3, the states $s_9$–$s_{12}$ are never observed during 1990–2006, thus we have zero frequencies affecting the computation of $\hat{\nu}_{\text{En}}$ and $\hat{\kappa}^*(h)$. The marginal distribution itself deviates visibly from a uniform distribution, thus it is reasonable that the dispersion measures $\hat{\nu}_{\text{G}}$, $\hat{\nu}_{\text{En}}$ and $\hat{\nu}_{\text{Ex}}$ are clearly below 1. Actually, the PMF mainly concentrates on the low to moderate fear states, high anxiety (or more) only happened for few of the trading days. Even more important is to investigate the development of these fear states over time. While negative serial dependence would indicate a permanent fluctuation between the states, positive dependence would imply some kind of inertia regarding the respective states. The serial dependence structure was analyzed, as shown in the bottom panel of Figure 3, where the critical values for level 5% (dashed lines) were computed according to the asymptotics in Section 5. All measures indicated significantly positive dependence, thus the U.S. stock market has a tendency to stay in a state once attained. However, $\hat{\kappa}^*(h)$ (Figure 3, center) produced notably smaller values than $\hat{\kappa}(h), \hat{\kappa}^\star(h)$ (Figure 3, left and right). Considering that the time series plot with its long runs of single states implies a rather strong positive dependence, the values produced by $\hat{\kappa}^*(h)$ did not appear to be that plausible. Thus, $\hat{\kappa}(h), \hat{\kappa}^\star(h)$, which resulted in very similar values, appeared to be better interpretable in the given example. Note that the discrepancy between $\hat{\kappa}^*(h)$ and $\hat{\kappa}(h), \hat{\kappa}^\star(h)$ went along with the discrepancy between $\hat{\nu}_{\text{En}}$ and $\hat{\nu}_{\text{G}}, \hat{\nu}_{\text{Ex}}$.

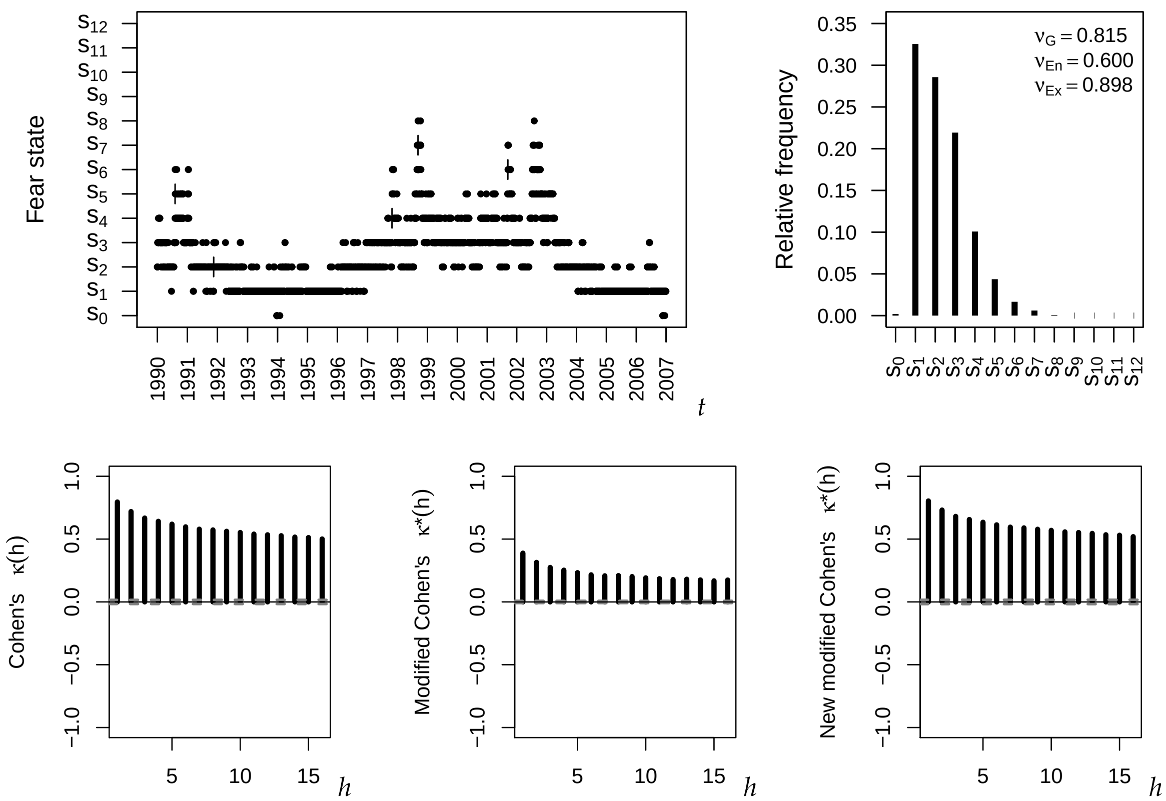


**Figure 3.** Fear states time series plot and PMF (**top**); and plots of $\hat{\kappa}(h), \hat{\kappa}^*(h), \hat{\kappa}^\star(h)$ (**bottom**).

### *6.3. Application: Wood Pewee*

Animal behavior studies are an integral part of research in biology and psychology. The time series example to be studied in the present section dates back to one of the pioneers of ethology, to Wallace Craig. The data were originally presented by Craig (1943) and further analyzed (among others) in Chapter 6 of the work by Weiß (2018). They constitute a nominal time series of length $n = 1327$, where the three states $s_0, s_1, s_2$ express the different phrases in the morning twilight song of the Wood Pewee ("pee-ah-wee", "pee-oh" and "ah-di-dee", respectively). Since the range of a nominal time series lacks a natural order, a time series plot is not possible. Thus, Figure 4 shows a rate evolution graph as a substitute, where the cumulative frequencies of the individual states are plotted against time $t$. From the roughly linear increase, we conclude on a stable behavior of the time series (Weiß 2018).

The dispersion measures $\hat{\nu}_{\text{G}}$, $\hat{\nu}_{\text{En}}$ and $\hat{\nu}_{\text{Ex}}$ all led to values between 0.90 and 0.95, indicating that all three phrases are frequently used (but not equally often) within the morning twilight song of the Wood Pewee. This is confirmed by the PMF plot in Figure 4, where we found some preference for the phrase $s_0$. From the serial dependence plots in the bottom panel of Figure 4, a quite complex serial dependence structure becomes visible, with both positive and negative dependence values and with a periodic pattern. Positive values happen for even lags $h$ (and particularly large values for multiples of 4), and negative values for odd lags. The positive values indicate a tendency for repeating a phrase, and such repetitions seem to be particularly likely after every fourth phrase. Negative values, in contrast, indicate a change of the phrase, e.g., it will rarely happen that the same phrase is presented twice in a row. While $\hat{\kappa}(h), \hat{\kappa}^\star(h)$ gave a very clear (and similar) picture of the rhythmic structure, it was again $\hat{\kappa}^*(h)$ that caused some implausible values, e.g., the non-significant value at lag 2. Thus, both data examples indicate that $\hat{\kappa}^*(h)$ should be used with caution in practice (also because of the zero problem). A decision between $\hat{\kappa}(h)$ and $\hat{\kappa}^\star(h)$ is more difficult; $\hat{\kappa}(h)$ is well established and slightly advantageous for uncovering positive dependencies, whereas $\hat{\kappa}^\star(h)$ is computationally simpler and shows a very good performance regarding negative dependencies.

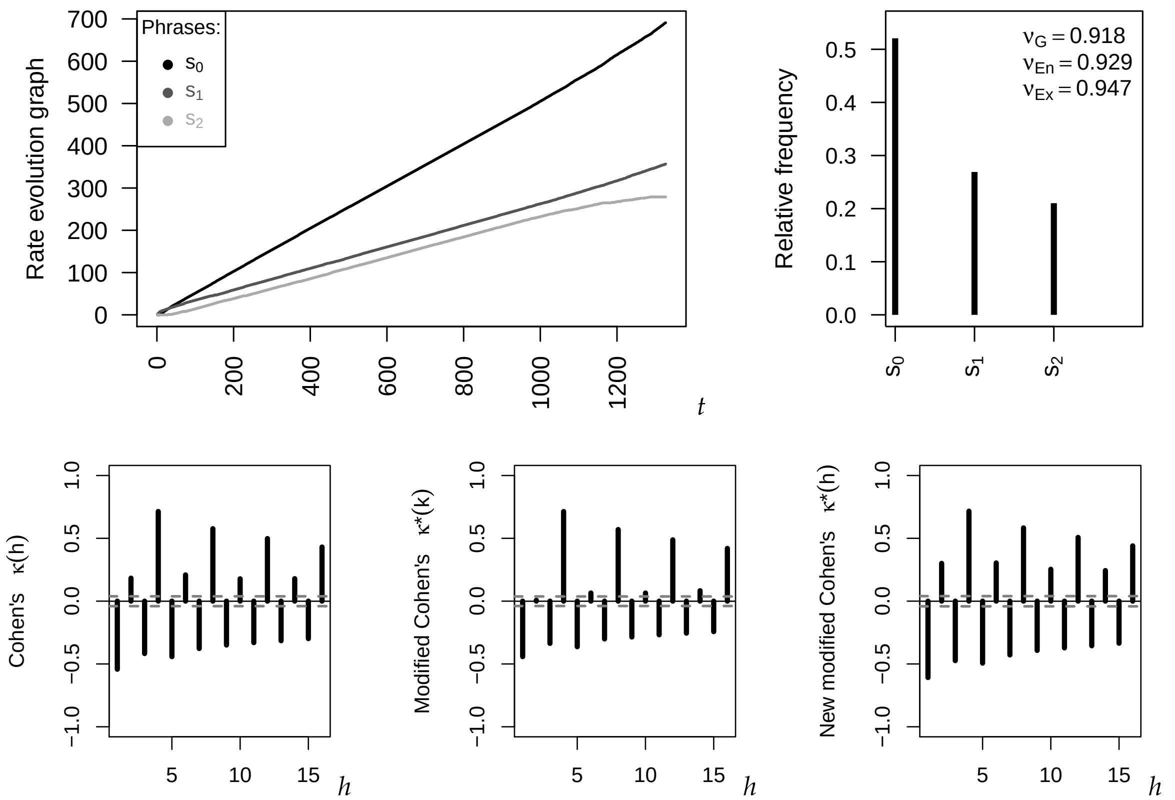


**Figure 4.** Wood Pewee time series: rate evolution graph and PMF plot (**top**); and plots of $\hat{\kappa}(h), \hat{\kappa}^*(h), \hat{\kappa}^\star(h)$ (**bottom**).

## 7. Conclusions

This work discusses approaches for measuring dispersion and serial dependence in categorical time series. Asymptotic properties of the novel extropy measure for categorical dispersion are derived and compared to those of Gini index and entropy. Simulations showed that all three measures performed quite well, with slightly better coverage rates for the entropy but computational advantages for Gini index and extropy. The extropy was most reliable if testing the null hypothesis of a uniform distribution. The application and interpretation of these measures was illustrated with a longitudinal study about the most important political goals in Germany.

The analysis of the asymptotic bias of Gini index, entropy and extropy uncovered a relation between these three measures and three types of $\kappa$-measures for signed serial dependence. While two of these measures, namely $\kappa(h)$ and $\kappa^*(h)$, have already been discussed in the literature, the "$\kappa$-counterpart" to the extropy turned out to be a new type of modified Cohen's $\kappa$, denoted by $\kappa^\star(h)$. The asymptotics of $\hat{\kappa}^\star(h)$ were investigated and utilized for testing for serial dependence. A simulation study as well as two real-data examples (time series of fear states and song of the Wood Pewee) showed that $\hat{\kappa}^*(h)$ has several drawbacks, while both $\hat{\kappa}(h)$ and $\hat{\kappa}^\star(h)$ work very well in practice. The advantages of $\hat{\kappa}^\star(h)$ are computational simplicity and a superior performance regarding negative dependencies.

**Funding:** This research received no external funding.

**Acknowledgments:** The author thanks the editors and the three referees for their useful comments on an earlier draft of this article.

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

## Appendix A. Some Models for Categorical Processes

This appendix provides a brief summary of those models for categorical processes that are used for simulations and illustrative computations in this article. More background on these and further models for categorical processes can be found in the book by Weiß (2018).

### *Appendix A.1. NDARMA Models for Categorical Processes*

The *NDARMA model* ("new" discrete autoregressive moving-average model) was proposed by Jacobs and Lewis (1983), and its definition might be given as follows (Weiß and Göb 2008):

Let $(X_t)_{\mathbb{Z}}$ and $(\epsilon_t)_{\mathbb{Z}}$ be categorical processes with state space $\mathcal{S}$, where $(\epsilon_t)_{\mathbb{Z}}$ is i.i.d. with marginal distribution $\boldsymbol{p}$, and where $\epsilon_t$ is independent of $(X_s)_{s<t}$. Let

$$(\alpha_{t,1},\ldots,\alpha_{t,\mathrm{p}},\beta_{t,0},\ldots,\beta_{t,\mathrm{q}}) \quad\sim\ \mathrm{MULT}(1;\ \phi_1,\ldots,\phi_{\mathrm{p}},\varphi_0,\ldots,\varphi_{\mathrm{q}})$$

be i.i.d. multinomial random vectors, which are independent of $(\epsilon_t)_{\mathbb{Z}}$ and of $(X_s)_{s<t}$. Then, $(X_t)_{\mathbb{Z}}$ is said to be an NDARMA(p, q) process (and the cases $\mathrm{q}=0$ and $\mathrm{p}=0$ are referred to as a DAR(p) process and DMA(q) process, respectively) if it follows the recursion

$$X_t\ =\ \alpha_{t,1}\cdot X_{t-1}+\ldots+\alpha_{t,\mathrm{p}}\cdot X_{t-\mathrm{p}}\ +\ \beta_{t,0}\cdot\epsilon_t+\ldots+\beta_{t,\mathrm{q}}\cdot\epsilon_{t-\mathrm{q}}. \tag{A1}$$

(Here, if the state space $\mathcal{S}$ is not numerically coded, we assume $0\cdot s=0$, $1\cdot s=s$ and $s+0=s$ for each $s\in\mathcal{S}$.)

NDARMA processes have several attractive properties, e.g. $X_t$ and $\epsilon_t$ have the same stationary marginal distribution: $P(X_t=s_i)\ =\ p_i\ =\ P(\epsilon_t=s_i)$ for all $i\in\mathcal{S}$. Their serial dependence structure is characterized by a set of Yule–Walker-type equations for the serial dependence measure Cohen's $\kappa$ from Equation (3) (Weiß and Göb 2008):

$$\kappa(h)\ =\ \textstyle\sum_{j=1}^{\mathrm{p}}\phi_j\,\kappa(|h-j|)\ +\ \sum_{i=0}^{\mathrm{q}-h}\varphi_{i+h}\,r(i)\qquad\text{for } h\geq 1, \tag{A2}$$

where the $r(i)$ satisfy $r(i)\ =\ \sum_{j=\max\{0,i-\mathrm{p}\}}^{i-1}\phi_{i-j}\cdot r(j)\ +\ \varphi_i\,\mathbb{1}(0\leq i\leq \mathrm{q})$. It should be noted that NDARMA processes satisfy $\kappa(h)\geq 0$, i.e. they can only handle positive serial dependence. Another important property is that the bivariate distributions at lag $h$ are $p_{i|j}(h)\ =\ p_i\ +\ \kappa(h)\,(\delta_{i,j}-p_i)$. This implies that $p_{ij}(h)-p_i\,p_j\ =\ \kappa(h)\,(\delta_{i,j}-p_i)\,p_j$ and, as a consequence, that all of the serial dependence measures mentioned in this work coincide for NDARMA processes: $\vartheta(h)=\vartheta^*(h)=\vartheta^\star(h)=\kappa(h)=\kappa^*(h)=\kappa^\star(h)$.

Finally, Weiß (2013) showed that an NDARMA process is $\phi$-mixing with exponentially decreasing weights such that the CLT on p. 200 in Billingsley (1999) is applicable.

### *Appendix A.2. Markov Chains for Categorical Processes*

A discrete-valued Markov process $(X_t)_{\mathbb{Z}}$ is characterized by a "memory of length $\mathrm{p}\in\mathbb{N}$", in the sense that

$$P(X_t=x_t\mid X_{t-1}=x_{t-1},\ldots)\ =\ P(X_t=x_t\mid X_{t-1}=x_{t-1},\ldots,X_{t-\mathrm{p}}=x_{t-\mathrm{p}})$$

has to hold for all $x_i \in \mathcal{S}$. In the case p $= 1$, $(X_t)_{\mathbb{Z}}$ is commonly called a *Markov chain*. If the transition probabilities $P(X_t = i \mid X_{t-1} = j)$ of the Markov chain do not change with time $t$, i.e., if $P(X_t = i \mid X_{t-1} = j) = p_{i|j}$ for all $t \in \mathbb{N}$, it is said to be homogeneous (analogously for higher-order Markov processes). An example of a parsimoniously parametrized homogeneous Markov chain (Markov process) is the DAR(1) process (DAR(p) process) according to Appendix A.1, which always exhibits positive serial dependence.

A parsimoniously parametrized Markov model with negative serial dependence was proposed by Weiß (2011), the "Negative Markov model" (*NegMarkov model*). For a given probability vector $\boldsymbol{\pi} \in (0;1)^{m+1}$ and some $\alpha \in (0;1]$, its transition probabilities are defined by

$$p_{i|j} \;=\; \begin{cases} \alpha\,\pi_j & \text{if } i = j, \\ \beta_j\,\pi_i & \text{if } i \neq j, \end{cases} \qquad \text{where } \beta_j \;=\; \frac{1-\alpha\,\pi_j}{1-\pi_j} \;\geq 1.$$

The resulting ergodic Markov chain has the stationary marginal distribution

$$p_j \;=\; \frac{\pi_j/\beta_j}{\sum_{i=0}^m \pi_i/\beta_i} \qquad \text{for } j = 0,\ldots,m.$$

As an example, if $\boldsymbol{\pi} = \boldsymbol{p}_{\text{uni}}$, then the $\beta_j$ become $1 + (1-\alpha)/m$ such that $\boldsymbol{p}$ is also a uniform distribution. However, the conditional distribution given $X_{t-1} = j$ is not uniform:

$$p_{i|j} \;=\; \begin{cases} \frac{\alpha}{m+1} & \text{if } i = j, \\ \left(1 + \frac{1-\alpha}{m}\right)\frac{1}{m+1} & \text{if } i \neq j. \end{cases}$$

## Appendix B. Proofs

### *Appendix B.1. Derivation of the inequality in Equation (7)*

The first inequality in Equation (7) was shown by Lad et al. (2015). Using the inequality $\ln(1+x) > x/(1+x/2)$ for $x > 0$ from Love (1980), it follows for $p_i \in (0;1)$ that

$$-\ln(1-p_i) \;=\; \ln\left(1 + \tfrac{p_i}{1-p_i}\right) \;>\; \tfrac{2p_i}{2-p_i}, \qquad -\ln p_i \;=\; \ln\left(1 + \tfrac{1-p_i}{p_i}\right) \;>\; 2\,\tfrac{1-p_i}{1+p_i}.$$

Consequently, we have

$$-\textstyle\sum_{i=0}^m (1-p_i)\,\ln(1-p_i) \;\geq\; \sum_{i=0}^m p_i(1-p_i)\,\frac{2}{2-p_i} \;\geq\; 1 - s_2(\boldsymbol{p}),$$

as well as

$$-\textstyle\sum_{i=0}^m p_i\,\ln p_i \;\geq\; \sum_{i=0}^m p_i(1-p_i)\,\frac{2}{1+p_i} \;\geq\; 1 - s_2(\boldsymbol{p}).$$

### *Appendix B.2. Derivations for Sample Dispersion Measures*

For studying the asymptotic properties of the sample Gini index according to Equation (1), it is important to know that $\hat{\nu}_G$ can be *exactly* rewritten as the centered quadratic polynomial

$$\hat{\nu}_G \;=\; \nu_G \;-\; 2\,\tfrac{m+1}{m}\,\textstyle\sum_{i=0}^m p_i\,(\hat{p}_i - p_i) \;-\; \tfrac{m+1}{m}\,\sum_{i=0}^m (\hat{p}_i - p_i)^2. \tag{A3}$$

Since $E[\hat{\boldsymbol{p}}] = \boldsymbol{p}$ holds exactly, this representation immediately implies an *exact* way of bias computation, $n\,\big(E[\hat{\nu}_G] - \nu_G\big) = \frac{m+1}{m}\,\sum_{i=0}^m V\big[\sqrt{n}\,(\hat{p}_i - p_i)\big]$, where $V\big[\sqrt{n}\,(\hat{p}_i - p_i)\big]$ is approximately given by $\sigma_{ii}$ according to Equation (8). Furthermore, provided that $\boldsymbol{p} \neq \boldsymbol{p}_{\text{uni}}$, the resulting linear approximation together with the asymptotic normality of $\sqrt{n}\,(\hat{\boldsymbol{p}} - \boldsymbol{p})$ can be used to derive the asymptotic result (Equation (9)) (Delta method). Here, $\boldsymbol{p} = \boldsymbol{p}_{\text{uni}}$ has to be excluded, because then the linear term in Equation (A3) vanishes. Hence, in the boundary case $\boldsymbol{p} = \boldsymbol{p}_{\text{uni}}$, we end up with an asymptotic quadratic-form distribution instead of a normal one. Actually, it is easily seen that

$n\,m\,(1-\hat{\nu}_{\mathrm{G}})$ then coincides with the Pearson's $\chi^2$-statistic with respect to $\boldsymbol{p}_{\mathrm{uni}}$; see Section 4 in Weiß (2013) for the asymptotics.

For the entropy in Equation (2), an exact polynomial representation such as in Equation (A3) does not exist, thus we have to use a Taylor approximation instead:

$$\hat{\nu}_{\mathrm{En}} \;\approx\; \nu_{\mathrm{En}} \;-\; \textstyle\sum_{j=0}^{m} \frac{1+\ln p_j}{\ln(m+1)}\,(\hat{p}_j-p_j) \;-\; \frac{1}{2}\,\textstyle\sum_{j=0}^{m} \frac{p_j^{-1}}{\ln(m+1)}\,(\hat{p}_j-p_j)^2. \tag{A4}$$

However, then, one can proceed as before. Thus, an approximate bias formula follows from

$$n\,\big(E[\hat{\nu}_{\mathrm{En}}]-\nu_{\mathrm{En}}\big) \;\approx\; -\tfrac{1}{2}\,\textstyle\sum_{j=0}^{m} \frac{p_j^{-1}}{\ln(m+1)}\,V\big[\sqrt{n}\,(\hat{p}_j-p_j)\big] \;\approx\; -\tfrac{1}{2}\,\textstyle\sum_{j=0}^{m} \frac{p_j^{-1}}{\ln(m+1)}\,\sigma_{jj}.$$

For $\boldsymbol{p}\neq\boldsymbol{p}_{\mathrm{uni}}$, we can use the linear approximation implied by Equation (A4) to conclude on the asymptotic normality of $\sqrt{n}\,(\hat{\nu}_{\mathrm{En}}-\nu_{\mathrm{En}})$ with variance

$$\sigma_{\mathrm{En}}^2 \;=\; \sum_{i,j=0}^{m} \left(\frac{1+\ln p_i}{\ln(m+1)}\right)\left(\frac{1+\ln p_j}{\ln(m+1)}\right)\sigma_{ij},$$

also see the results in (Blyth 1959; Weiß 2013). Note that, in the i.i.d.-case, where $\sigma_{ij}=p_j\,(\delta_{i,j}-p_i)$, one computes

$$\begin{aligned}
&\textstyle\sum_{i,j=0}^{m}(1+\ln p_i)(1+\ln p_j)\,p_j\,(\delta_{i,j}-p_i)\\
&\quad=\; \textstyle\sum_{i=0}^{m}(1+\ln p_i)^2\,p_i \;-\; \big(\textstyle\sum_{i=0}^{m}(1+\ln p_i)\,p_i\big)^2\\
&\quad=\; 1+2\,\textstyle\sum_{i=0}^{m} p_i\,\ln p_i+\textstyle\sum_{i=0}^{m} p_i\,(\ln p_i)^2 \;-\; \big(1+\textstyle\sum_{i=0}^{m} p_i\,\ln p_i\big)^2\\
&\quad=\; \textstyle\sum_{i=0}^{m} p_i\,(\ln p_i)^2 \;-\; \big(\textstyle\sum_{i=0}^{m} p_i\,\ln p_i\big)^2.
\end{aligned}$$

Finally, in the boundary case $\boldsymbol{p}=\boldsymbol{p}_{\mathrm{uni}}$, again the linear term vanishes such that

$$2n\,\ln(m+1)\cdot(1-\hat{\nu}_{\mathrm{En}}) \;\approx\; n\,\sum_{j=0}^{m}\frac{(\hat{p}_j-\frac{1}{m+1})^2}{\frac{1}{m+1}}$$

equals the Pearson's $\chi^2$-statistic.

Finally, we do analogous derivations concerning the extropy in Equation (5). Starting with the Taylor approximation

$$\hat{\nu}_{\mathrm{Ex}} \;\approx\; \nu_{\mathrm{Ex}} \;+\; \textstyle\sum_{j=0}^{m} \frac{1+\ln(1-p_j)}{m\,\ln(\frac{m+1}{m})}\,(\hat{p}_j-p_j) \;-\; \frac{1}{2}\,\textstyle\sum_{j=0}^{m} \frac{(1-p_j)^{-1}}{m\,\ln(\frac{m+1}{m})}\,(\hat{p}_j-p_j)^2, \tag{A5}$$

it follows that

$$n\,\big(E[\hat{\nu}_{\mathrm{Ex}}]-\nu_{\mathrm{Ex}}\big) \;\approx\; -\tfrac{1}{2}\,\textstyle\sum_{j=0}^{m} \frac{(1-p_j)^{-1}}{m\,\ln(\frac{m+1}{m})}\,\sigma_{jj}.$$

For $\boldsymbol{p}\neq\boldsymbol{p}_{\mathrm{uni}}$, we can use the linear approximation implied by Equation (A5) to conclude on the asymptotic normality of $\sqrt{n}\,(\hat{\nu}_{\mathrm{Ex}}-\nu_{\mathrm{Ex}})$ with variance

$$\sigma_{\mathrm{Ex}}^2 \;=\; \sum_{i,j=0}^{m} \frac{1+\ln(1-p_i)}{m\,\ln(\frac{m+1}{m})}\,\frac{1+\ln(1-p_j)}{m\,\ln(\frac{m+1}{m})}\,\sigma_{ij}.$$

Note that, in the i.i.d.-case, where $\sigma_{ij}=p_j\,(\delta_{i,j}-p_i)$, one computes

$$\begin{aligned}
&\textstyle\sum_{i,j=0}^{m}\big(1+\ln(1-p_i)\big)\big(1+\ln(1-p_j)\big)\,p_j\,(\delta_{i,j}-p_i)\\
&\quad=\; \textstyle\sum_{i=0}^{m} p_i\,\big(\ln(1-p_i)\big)^2 \;-\; \big(\textstyle\sum_{i=0}^{m} p_i\,\ln(1-p_i)\big)^2
\end{aligned}$$

as before. Finally, in the boundary case $\boldsymbol{p} = \boldsymbol{p}_{\text{uni}}$, again the linear term vanishes such that

$$2n\, m^2\, \ln\left(\tfrac{m+1}{m}\right) \cdot (1 - \hat{\nu}_{\text{Ex}}) \;\approx\; n \sum_{j=0}^{m} \frac{(\hat{p}_j - \frac{1}{m+1})^2}{\frac{1}{m+1}}$$

equals the Pearson's $\chi^2$-statistic.

*Appendix B.3. Derivations for Measures of Signed Serial Dependence*

We partition $\boldsymbol{x} \in (0;1)^{2(m+1)}$ as $\boldsymbol{x} = (x_0, \ldots, x_m,\ x_{m+1+0}, \ldots, x_{m+1+m})^\top$, and we define

$$f(\boldsymbol{x}) \;=\; \sum_{j=0}^{m} \frac{x_{m+1+j} - x_j^2}{1 - x_j}.$$

Then,

$$\tfrac{\partial}{\partial x_j}\, f(\boldsymbol{x}) \;=\; \frac{x_{m+1+j} - 2x_j + x_j^2}{(1-x_j)^2}, \qquad \tfrac{\partial}{\partial x_{m+1+j}}\, f(\boldsymbol{x}) \;=\; \frac{1}{1-x_j},$$

and

$$\tfrac{\partial^2}{\partial x_j^2}\, f(\boldsymbol{x}) \;=\; \frac{-2(1 - x_{m+1+j})}{(1-x_j)^3}, \qquad \tfrac{\partial^2}{\partial x_j\, \partial x_{m+1+j}}\, f(\boldsymbol{x}) \;=\; \frac{1}{(1-x_j)^2},$$

all other second-order derivatives equal 0. Thus, a second-order Taylor approximation of $\hat{\kappa}^\star(h) = f(\ldots, \hat{p}_j, \ldots, \hat{p}_{jj}(h), \ldots)$ is given by

$$\begin{aligned}
\hat{\kappa}^\star(h) \;\approx\; & \kappa^\star(h) + \sum_{j=0}^{m} \frac{p_{jj}(h) - 2p_j + p_j^2}{(1-p_j)^2}\, (\hat{p}_j - p_j) + \sum_{j=0}^{m} \frac{1}{1-p_j}\, \big(\hat{p}_{jj}(h) - p_{jj}(h)\big) \\
& - \sum_{j=0}^{m} \frac{1 - p_{jj}(h)}{(1-p_j)^3}\, (\hat{p}_j - p_j)^2 + \sum_{j=0}^{m} \frac{1}{(1-p_j)^2}\, (\hat{p}_j - p_j)\, \big(\hat{p}_{jj}(h) - p_{jj}(h)\big).
\end{aligned}$$

Hence, using Equation (18), it follows that

$$\begin{aligned}
n\,\big(E\big[\hat{\kappa}^\star(h)\big] - \kappa^\star(h)\big) \;\approx\; & \sum_{j=0}^{m} \frac{1}{(1-p_j)^2}\, \sigma^{(h)}_{j,m+1+j} - \sum_{j=0}^{m} \frac{1 - p_{jj}(h)}{(1-p_j)^3}\, \sigma_{jj} \\
=\; & \sum_{j=0}^{m} \frac{(1-p_j)\, \sigma^{(h)}_{j,m+1+j} - \big(1 - p_{jj}(h)\big)\, \sigma_{jj}}{(1-p_j)^3}.
\end{aligned}$$

Furthermore, the Delta method implies that $\sqrt{n}\,\big(\hat{\kappa}^\star(h) - \kappa^\star(h)\big) \sim \mathrm{N}(0, \sigma^2)$ with

$$\begin{aligned}
\sigma^2 \;=\; & \sum_{i,j=0}^{m} \frac{p_{ii}(h) - 2p_i + p_i^2}{(1-p_i)^2}\, \frac{p_{jj}(h) - 2p_j + p_j^2}{(1-p_j)^2}\, \sigma_{i,j} \\
& + \sum_{i,j=0}^{m} \frac{\sigma^{(h)}_{m+1+i,m+1+j}}{(1-p_i)(1-p_j)} + 2 \sum_{i,j=0}^{m} \frac{p_{ii}(h) - 2p_i + p_i^2}{(1-p_i)^2}\, \frac{\sigma^{(h)}_{i,m+1+j}}{1-p_j}.
\end{aligned}$$

Note that, under the null of an i.i.d. DGP, we have the simplifications

$$\frac{p_{jj}(h) - 2p_j + p_j^2}{(1-p_j)^2} \;=\; \frac{-2p_j}{1-p_j}, \qquad \frac{1 - p_{jj}(h)}{(1-p_j)^3} \;=\; \frac{1+p_j}{(1-p_j)^2}.$$

Thus, the second-order Taylor approximation of $\hat{\kappa}^\star(h)$ then simplifies to

$$\begin{aligned}\hat{\kappa}^\star(h) \;&\approx\; \kappa^\star(h) \,+\, \sum_{j=0}^{m} \frac{-2p_j}{1-p_j}\,(\hat{p}_j - p_j) \,+\, \sum_{j=0}^{m} \frac{1}{1-p_j}\,(\hat{p}_{jj}(h) - p_j^2) \\ &\quad -\, \sum_{j=0}^{m} \frac{1+p_j}{(1-p_j)^2}\,(\hat{p}_j - p_j)^2 \,+\, \sum_{j=0}^{m} \frac{1}{(1-p_j)^2}\,(\hat{p}_j - p_j)\,(\hat{p}_{jj}(h) - p_j^2).\end{aligned}$$

Furthermore,

$$\sigma_{i,j} \;=\; p_j\,(\delta_{i,j} - p_i), \quad \sigma_{i,m+1+j} \;=\; 2p_j^2\,(\delta_{i,j} - p_i), \quad \sigma_{m+1+i,m+1+j} \;=\; \delta_{i,j}\,p_i^2(1+2p_i) - 3p_i^2 p_j^2,$$

according to Equation (19). Thus, for an i.i.d. DGP, it follows that

$$\begin{aligned} n\left(E\big[\hat{\kappa}^\star(h)\big] - \kappa^\star(h)\right) \;=\; n\,E\big[\hat{\kappa}^\star(h)\big] \;&\approx\; \sum_{j=0}^{m} \frac{\sigma_{j,m+1+j} \;-\; (1+p_j)\,\sigma_{jj}}{(1-p_j)^2} \\ &=\; \sum_{j=0}^{m} \frac{2p_j^2\,(1-p_j) \;-\; (1+p_j)\,p_j(1-p_j)}{(1-p_j)^2} \\ &=\; \sum_{j=0}^{m} \frac{-p_j(1-p_j)\,(1+p_j-2p_j)}{(1-p_j)^2} \;=\; -1. \end{aligned}$$

Furthermore,

$$\begin{aligned} \sigma^2 \;&=\; \sum_{i,j=0}^{m} \frac{-2p_i}{1-p_i}\,\frac{-2p_j}{1-p_j}\,p_j\,(\delta_{i,j} - p_i) \,+\, 2\sum_{i,j=0}^{m} \frac{-2p_i}{1-p_i}\,\frac{1}{1-p_j}\,2p_j^2\,(\delta_{i,j} - p_i) \\ &\quad +\, \sum_{i,j=0}^{m} \frac{1}{1-p_i}\,\frac{1}{1-p_j}\,\big(\delta_{i,j}\,p_i^2(1+2p_i) - 3p_i^2 p_j^2\big) \\ &=\; 4\sum_{i=0}^{m} \frac{p_i^3}{(1-p_i)^2} \,-\, 8\sum_{i=0}^{m} \frac{p_i^3}{(1-p_i)^2} \,+\, \sum_{i=0}^{m} \frac{p_i^2(1+2p_i)}{(1-p_i)^2} \\ &\quad -\, 4\sum_{i,j=0}^{m} \frac{p_i^2}{1-p_i}\,\frac{p_j^2}{1-p_j} \,+\, 8\sum_{i,j=0}^{m} \frac{p_i^2}{1-p_i}\,\frac{p_j^2}{1-p_j} \,-\, 3\sum_{i,j=0}^{m} \frac{p_i^2}{1-p_i}\,\frac{p_j^2}{1-p_j} \\ &=\; \sum_{i=0}^{m} \frac{p_i^2(1-2p_i)}{(1-p_i)^2} \,+\, \left(\sum_{i=0}^{m} \frac{p_i^2}{1-p_i}\right)^2. \end{aligned}$$

This leads to Equation (21).

In the special case of a uniform distribution, i.e. where all $p_i = \frac{1}{m+1}$, the asymptotic variances according to Equations (16), (17) and (21) coincide. Then, we have in Equation (16) that

$$1 \,-\, \tfrac{1+2s_3(\boldsymbol{p})-3s_2(\boldsymbol{p})}{(1-s_2(\boldsymbol{p}))^2} \;=\; 1 \,-\, \tfrac{1+\frac{2}{(m+1)^2}-\frac{3}{m+1}}{(1-\frac{1}{m+1})^2} \;=\; 1 \,-\, \tfrac{(m+1)^2+2-3\,(m+1)}{m^2} \;=\; 1 \,-\, \tfrac{m^2-m}{m^2} \;=\; \tfrac{1}{m},$$

which corresponds to Equation (17). However, the same expression also follows in Equation (21):

$$s_2(\boldsymbol{p}) \,-\, \textstyle\sum_{i=0}^{m}\left(\frac{p_i^2}{1-p_i}\right)^2 \,+\, \left(\sum_{i=0}^{m}\frac{p_i^2}{1-p_i}\right)^2 \;=\; \frac{1}{m+1} \,-\, \frac{1}{m^2\,(m+1)} \,+\, \frac{1}{m^2} \;=\; \frac{1}{m}.$$

## Appendix C. Tables

**Table A1.** Asymptotic vs. simulated mean (M.-a vs. M.-s) of $\hat{\nu}_{\mathrm{G}}$, $\hat{\nu}_{\mathrm{En}}$, $\hat{\nu}_{\mathrm{Ex}}$ for DGPs i.i.d., DMA(1) with $\varphi_1 = 0.25$, and DAR(1) with $\phi_1 = 0.40$.

| | **DGP** | **i.i.d.** | | **DMA(1), 0.25** | | **DAR(1), 0.40** | | **i.i.d.** | | **DMA(1), 0.25** | | **DAR(1), 0.40** | | **i.i.d.** | | **DMA(1), 0.25** | | **DAR(1), 0.40** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PMF** | $n$ | **M$_{\mathrm{G}}$-a** | **M$_{\mathrm{G}}$-s** | **M$_{\mathrm{G}}$-a** | **M$_{\mathrm{G}}$-s** | **M$_{\mathrm{G}}$-a** | **M$_{\mathrm{G}}$-s** | **M$_{\mathrm{En}}$-a** | **M$_{\mathrm{En}}$-s** | **M$_{\mathrm{En}}$-a** | **M$_{\mathrm{En}}$-s** | **M$_{\mathrm{En}}$-a** | **M$_{\mathrm{En}}$-s** | **M$_{\mathrm{Ex}}$-a** | **M$_{\mathrm{Ex}}$-s** | **M$_{\mathrm{Ex}}$-a** | **M$_{\mathrm{Ex}}$-s** | **M$_{\mathrm{Ex}}$-a** | **M$_{\mathrm{Ex}}$-s** |
| $p_1$ | 100 | 0.970 | 0.970 | 0.967 | 0.967 | 0.957 | 0.957 | 0.969 | 0.968 | 0.965 | 0.964 | 0.954 | 0.954 | 0.982 | 0.982 | 0.980 | 0.980 | 0.975 | 0.975 |
| | 250 | 0.976 | 0.976 | 0.975 | 0.975 | 0.971 | 0.971 | 0.975 | 0.975 | 0.973 | 0.973 | 0.969 | 0.969 | 0.986 | 0.986 | 0.985 | 0.985 | 0.983 | 0.983 |
| | 500 | 0.978 | 0.978 | 0.977 | 0.977 | 0.975 | 0.975 | 0.977 | 0.977 | 0.976 | 0.976 | 0.974 | 0.974 | 0.987 | 0.987 | 0.987 | 0.987 | 0.985 | 0.985 |
| | 1000 | 0.979 | 0.979 | 0.979 | 0.979 | 0.978 | 0.978 | 0.978 | 0.978 | 0.978 | 0.978 | 0.977 | 0.977 | 0.988 | 0.988 | 0.987 | 0.987 | 0.987 | 0.987 |
| $p_2$ | 100 | 0.627 | 0.627 | 0.625 | 0.625 | 0.619 | 0.619 | 0.649 | 0.649 | 0.645 | 0.644 | 0.634 | 0.634 | 0.739 | 0.739 | 0.737 | 0.737 | 0.731 | 0.731 |
| | 250 | 0.631 | 0.631 | 0.630 | 0.630 | 0.627 | 0.627 | 0.655 | 0.655 | 0.654 | 0.654 | 0.649 | 0.649 | 0.743 | 0.743 | 0.742 | 0.742 | 0.739 | 0.739 |
| | 500 | 0.632 | 0.632 | 0.632 | 0.631 | 0.630 | 0.630 | 0.657 | 0.657 | 0.657 | 0.656 | 0.654 | 0.654 | 0.744 | 0.744 | 0.743 | 0.743 | 0.742 | 0.742 |
| | 1000 | 0.633 | 0.633 | 0.632 | 0.633 | 0.632 | 0.632 | 0.658 | 0.658 | 0.658 | 0.658 | 0.657 | 0.657 | 0.744 | 0.744 | 0.744 | 0.744 | 0.744 | 0.743 |
| $p_3$ | 100 | 0.759 | 0.759 | 0.756 | 0.756 | 0.749 | 0.749 | 0.756 | 0.755 | 0.752 | 0.751 | 0.741 | 0.741 | 0.842 | 0.842 | 0.840 | 0.840 | 0.835 | 0.835 |
| | 250 | 0.764 | 0.764 | 0.762 | 0.762 | 0.760 | 0.760 | 0.762 | 0.762 | 0.761 | 0.761 | 0.757 | 0.757 | 0.846 | 0.846 | 0.845 | 0.845 | 0.843 | 0.843 |
| | 500 | 0.765 | 0.765 | 0.765 | 0.765 | 0.763 | 0.763 | 0.764 | 0.764 | 0.764 | 0.764 | 0.762 | 0.762 | 0.847 | 0.847 | 0.846 | 0.847 | 0.845 | 0.845 |
| | 1000 | 0.766 | 0.766 | 0.766 | 0.766 | 0.765 | 0.765 | 0.766 | 0.765 | 0.765 | 0.765 | 0.764 | 0.764 | 0.847 | 0.847 | 0.847 | 0.847 | 0.847 | 0.847 |
| $p_4$ | 100 | 0.433 | 0.433 | 0.431 | 0.431 | 0.427 | 0.428 | 0.486 | 0.486 | 0.482 | 0.481 | 0.471 | 0.471 | 0.568 | 0.568 | 0.566 | 0.566 | 0.560 | 0.561 |
| | 250 | 0.436 | 0.436 | 0.435 | 0.435 | 0.433 | 0.434 | 0.492 | 0.492 | 0.491 | 0.491 | 0.487 | 0.487 | 0.572 | 0.572 | 0.571 | 0.571 | 0.569 | 0.569 |
| | 500 | 0.437 | 0.437 | 0.436 | 0.436 | 0.435 | 0.436 | 0.495 | 0.495 | 0.494 | 0.494 | 0.492 | 0.492 | 0.573 | 0.573 | 0.572 | 0.572 | 0.571 | 0.571 |
| | 1000 | 0.437 | 0.437 | 0.437 | 0.437 | 0.436 | 0.436 | 0.496 | 0.496 | 0.495 | 0.495 | 0.494 | 0.494 | 0.573 | 0.573 | 0.573 | 0.573 | 0.573 | 0.573 |
| $p_5$ | 100 | 0.743 | 0.743 | 0.740 | 0.740 | 0.733 | 0.733 | 0.764 | 0.764 | 0.760 | 0.759 | 0.749 | 0.749 | 0.827 | 0.827 | 0.824 | 0.824 | 0.819 | 0.819 |
| | 250 | 0.747 | 0.747 | 0.746 | 0.746 | 0.743 | 0.743 | 0.770 | 0.770 | 0.768 | 0.769 | 0.764 | 0.765 | 0.830 | 0.830 | 0.829 | 0.829 | 0.827 | 0.827 |
| | 500 | 0.749 | 0.749 | 0.748 | 0.748 | 0.747 | 0.747 | 0.772 | 0.772 | 0.771 | 0.772 | 0.769 | 0.769 | 0.831 | 0.831 | 0.831 | 0.831 | 0.830 | 0.830 |
| | 1000 | 0.749 | 0.749 | 0.749 | 0.749 | 0.748 | 0.748 | 0.773 | 0.773 | 0.773 | 0.773 | 0.772 | 0.772 | 0.832 | 0.832 | 0.832 | 0.832 | 0.831 | 0.831 |
| $p_6$ | 100 | 0.928 | 0.928 | 0.925 | 0.925 | 0.916 | 0.916 | 0.929 | 0.929 | 0.925 | 0.925 | 0.915 | 0.915 | 0.956 | 0.956 | 0.953 | 0.954 | 0.948 | 0.948 |
| | 250 | 0.934 | 0.934 | 0.932 | 0.932 | 0.929 | 0.929 | 0.936 | 0.936 | 0.934 | 0.934 | 0.930 | 0.930 | 0.959 | 0.959 | 0.958 | 0.958 | 0.956 | 0.956 |
| | 500 | 0.936 | 0.936 | 0.935 | 0.935 | 0.933 | 0.933 | 0.938 | 0.938 | 0.937 | 0.937 | 0.935 | 0.935 | 0.960 | 0.960 | 0.960 | 0.960 | 0.959 | 0.959 |
| | 1000 | 0.937 | 0.937 | 0.936 | 0.936 | 0.935 | 0.935 | 0.939 | 0.939 | 0.939 | 0.939 | 0.938 | 0.938 | 0.961 | 0.961 | 0.961 | 0.961 | 0.960 | 0.960 |

**Table A2.** Asymptotic vs. simulated standard deviation (S.-a vs. S.-s) of $\hat{\nu}_{G}$, $\hat{\nu}_{En}$, $\hat{\nu}_{Ex}$ for DGPs i.i.d., DMA(1) with $\varphi_1 = 0.25$, and DAR(1) with $\phi_1 = 0.40$.

| | **DGP** | **i.i.d.** | | **DMA(1), 0.25** | | **DAR(1), 0.40** | | **i.i.d.** | | **DMA(1), 0.25** | | **DAR(1), 0.40** | | **i.i.d.** | | **DMA(1), 0.25** | | **DAR(1), 0.40** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PMF** | $n$ | $\mathbf{S_G}$**-a** | $\mathbf{S_G}$**-s** | $\mathbf{S_G}$**-a** | $\mathbf{S_G}$**-s** | $\mathbf{S_G}$**-a** | $\mathbf{S_G}$**-s** | $\mathbf{S_{En}}$**-a** | $\mathbf{S_{En}}$**-s** | $\mathbf{S_{En}}$**-a** | $\mathbf{S_{En}}$**-s** | $\mathbf{S_{En}}$**-a** | $\mathbf{S_{En}}$**-s** | $\mathbf{S_{Ex}}$**-a** | $\mathbf{S_{Ex}}$**-s** | $\mathbf{S_{Ex}}$**-a** | $\mathbf{S_{Ex}}$**-s** | $\mathbf{S_{Ex}}$**-a** | $\mathbf{S_{Ex}}$**-s** |
| $\boldsymbol{p}_1$ | 100 | 0.017 | 0.019 | 0.020 | 0.023 | 0.027 | 0.032 | 0.017 | 0.020 | 0.020 | 0.024 | 0.027 | 0.033 | 0.011 | 0.012 | 0.012 | 0.014 | 0.016 | 0.020 |
| | 250 | 0.011 | 0.011 | 0.013 | 0.014 | 0.017 | 0.018 | 0.011 | 0.012 | 0.013 | 0.014 | 0.017 | 0.019 | 0.007 | 0.007 | 0.008 | 0.008 | 0.010 | 0.011 |
| | 500 | 0.008 | 0.008 | 0.009 | 0.009 | 0.012 | 0.012 | 0.008 | 0.008 | 0.009 | 0.009 | 0.012 | 0.013 | 0.005 | 0.005 | 0.006 | 0.006 | 0.007 | 0.008 |
| | 1000 | 0.006 | 0.006 | 0.006 | 0.007 | 0.008 | 0.009 | 0.006 | 0.006 | 0.006 | 0.007 | 0.008 | 0.009 | 0.003 | 0.003 | 0.004 | 0.004 | 0.005 | 0.005 |
| $\boldsymbol{p}_2$ | 100 | 0.071 | 0.071 | 0.084 | 0.083 | 0.109 | 0.106 | 0.063 | 0.064 | 0.074 | 0.075 | 0.097 | 0.096 | 0.057 | 0.058 | 0.067 | 0.068 | 0.088 | 0.088 |
| | 250 | 0.045 | 0.045 | 0.053 | 0.053 | 0.069 | 0.068 | 0.040 | 0.040 | 0.047 | 0.047 | 0.061 | 0.061 | 0.036 | 0.036 | 0.043 | 0.043 | 0.055 | 0.055 |
| | 500 | 0.032 | 0.032 | 0.037 | 0.037 | 0.049 | 0.049 | 0.028 | 0.029 | 0.033 | 0.033 | 0.043 | 0.043 | 0.026 | 0.026 | 0.030 | 0.030 | 0.039 | 0.039 |
| | 1000 | 0.023 | 0.023 | 0.027 | 0.027 | 0.035 | 0.035 | 0.020 | 0.020 | 0.024 | 0.024 | 0.031 | 0.031 | 0.018 | 0.018 | 0.021 | 0.021 | 0.028 | 0.028 |
| $\boldsymbol{p}_3$ | 100 | 0.058 | 0.058 | 0.068 | 0.068 | 0.088 | 0.086 | 0.053 | 0.054 | 0.062 | 0.063 | 0.081 | 0.081 | 0.042 | 0.043 | 0.049 | 0.050 | 0.064 | 0.065 |
| | 250 | 0.037 | 0.037 | 0.043 | 0.043 | 0.056 | 0.055 | 0.033 | 0.034 | 0.039 | 0.039 | 0.051 | 0.051 | 0.027 | 0.027 | 0.031 | 0.031 | 0.041 | 0.041 |
| | 500 | 0.026 | 0.026 | 0.030 | 0.030 | 0.039 | 0.039 | 0.024 | 0.024 | 0.028 | 0.028 | 0.036 | 0.036 | 0.019 | 0.019 | 0.022 | 0.022 | 0.029 | 0.029 |
| | 1000 | 0.018 | 0.018 | 0.021 | 0.021 | 0.028 | 0.028 | 0.017 | 0.017 | 0.020 | 0.020 | 0.025 | 0.025 | 0.013 | 0.013 | 0.016 | 0.016 | 0.020 | 0.020 |
| $\boldsymbol{p}_4$ | 100 | 0.078 | 0.077 | 0.092 | 0.090 | 0.119 | 0.115 | 0.072 | 0.073 | 0.085 | 0.085 | 0.110 | 0.109 | 0.073 | 0.073 | 0.085 | 0.086 | 0.111 | 0.110 |
| | 250 | 0.049 | 0.049 | 0.058 | 0.058 | 0.075 | 0.075 | 0.046 | 0.046 | 0.054 | 0.054 | 0.070 | 0.070 | 0.046 | 0.046 | 0.054 | 0.054 | 0.070 | 0.071 |
| | 500 | 0.035 | 0.035 | 0.041 | 0.041 | 0.053 | 0.053 | 0.032 | 0.032 | 0.038 | 0.038 | 0.049 | 0.049 | 0.033 | 0.033 | 0.038 | 0.038 | 0.050 | 0.050 |
| | 1000 | 0.025 | 0.025 | 0.029 | 0.029 | 0.038 | 0.038 | 0.023 | 0.023 | 0.027 | 0.027 | 0.035 | 0.035 | 0.023 | 0.023 | 0.027 | 0.027 | 0.035 | 0.035 |
| $\boldsymbol{p}_5$ | 100 | 0.065 | 0.064 | 0.076 | 0.075 | 0.099 | 0.096 | 0.056 | 0.057 | 0.066 | 0.067 | 0.086 | 0.086 | 0.048 | 0.048 | 0.056 | 0.056 | 0.073 | 0.073 |
| | 250 | 0.041 | 0.041 | 0.048 | 0.048 | 0.062 | 0.062 | 0.036 | 0.036 | 0.042 | 0.042 | 0.054 | 0.054 | 0.030 | 0.030 | 0.035 | 0.035 | 0.046 | 0.046 |
| | 500 | 0.029 | 0.029 | 0.034 | 0.034 | 0.044 | 0.044 | 0.025 | 0.025 | 0.029 | 0.029 | 0.038 | 0.038 | 0.021 | 0.021 | 0.025 | 0.025 | 0.032 | 0.033 |
| | 1000 | 0.020 | 0.020 | 0.024 | 0.024 | 0.031 | 0.031 | 0.018 | 0.018 | 0.021 | 0.021 | 0.027 | 0.027 | 0.015 | 0.015 | 0.018 | 0.018 | 0.023 | 0.023 |
| $\boldsymbol{p}_6$ | 100 | 0.033 | 0.034 | 0.039 | 0.040 | 0.051 | 0.052 | 0.030 | 0.032 | 0.036 | 0.037 | 0.046 | 0.050 | 0.021 | 0.022 | 0.025 | 0.026 | 0.032 | 0.034 |
| | 250 | 0.021 | 0.021 | 0.025 | 0.025 | 0.032 | 0.032 | 0.019 | 0.019 | 0.022 | 0.023 | 0.029 | 0.030 | 0.013 | 0.013 | 0.016 | 0.016 | 0.020 | 0.021 |
| | 500 | 0.015 | 0.015 | 0.017 | 0.017 | 0.023 | 0.023 | 0.014 | 0.014 | 0.016 | 0.016 | 0.021 | 0.021 | 0.009 | 0.010 | 0.011 | 0.011 | 0.014 | 0.015 |
| | 1000 | 0.010 | 0.011 | 0.012 | 0.012 | 0.016 | 0.016 | 0.010 | 0.010 | 0.011 | 0.011 | 0.015 | 0.015 | 0.007 | 0.007 | 0.008 | 0.008 | 0.010 | 0.010 |

**Table A3.** Simulated coverage rate for 95% CIs of $\hat{\nu}_{\mathrm{G}}$, $\hat{\nu}_{\mathrm{En}}$, $\hat{\nu}_{\mathrm{Ex}}$ for DGPs i.i.d., DMA(1) with $\varphi_1 = 0.25$, and DAR(1) with $\phi_1 = 0.40$.

| | | **Coverage $\hat{\nu}_{\mathrm{G}}$, DGP:** | | | **Coverage $\hat{\nu}_{\mathrm{En}}$, DGP:** | | | **Coverage $\hat{\nu}_{\mathrm{Ex}}$, DGP:** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **PMF** | $n$ | **i.i.d.** | **DMA(1), 0.25** | **DAR(1), 0.40** | **i.i.d.** | **DMA(1), 0.25** | **DAR(1), 0.40** | **i.i.d.** | **DMA(1), 0.25** | **DAR(1), 0.40** |
| $p_1$ | 100 | 0.895 | 0.893 | 0.899 | 0.904 | 0.901 | 0.905 | 0.892 | 0.890 | 0.897 |
| | 250 | 0.914 | 0.909 | 0.898 | 0.920 | 0.915 | 0.905 | 0.912 | 0.907 | 0.896 |
| | 500 | 0.931 | 0.922 | 0.912 | 0.935 | 0.927 | 0.918 | 0.930 | 0.921 | 0.910 |
| | 1000 | 0.938 | 0.935 | 0.926 | 0.940 | 0.937 | 0.930 | 0.938 | 0.934 | 0.925 |
| $p_2$ | 100 | 0.936 | 0.926 | 0.908 | 0.937 | 0.929 | 0.915 | 0.931 | 0.928 | 0.911 |
| | 250 | 0.944 | 0.941 | 0.933 | 0.945 | 0.942 | 0.937 | 0.944 | 0.941 | 0.935 |
| | 500 | 0.947 | 0.946 | 0.941 | 0.947 | 0.946 | 0.943 | 0.947 | 0.946 | 0.941 |
| | 1000 | 0.949 | 0.947 | 0.944 | 0.949 | 0.948 | 0.946 | 0.949 | 0.947 | 0.945 |
| $p_3$ | 100 | 0.931 | 0.920 | 0.904 | 0.936 | 0.926 | 0.918 | 0.930 | 0.919 | 0.901 |
| | 250 | 0.941 | 0.938 | 0.930 | 0.943 | 0.943 | 0.937 | 0.940 | 0.937 | 0.929 |
| | 500 | 0.945 | 0.945 | 0.940 | 0.946 | 0.946 | 0.943 | 0.944 | 0.944 | 0.939 |
| | 1000 | 0.948 | 0.947 | 0.945 | 0.948 | 0.948 | 0.947 | 0.948 | 0.947 | 0.945 |
| $p_4$ | 100 | 0.941 | 0.925 | 0.904 | 0.938 | 0.925 | 0.905 | 0.938 | 0.931 | 0.913 |
| | 250 | 0.947 | 0.939 | 0.929 | 0.943 | 0.940 | 0.931 | 0.948 | 0.942 | 0.934 |
| | 500 | 0.949 | 0.946 | 0.941 | 0.948 | 0.946 | 0.942 | 0.950 | 0.947 | 0.943 |
| | 1000 | 0.952 | 0.945 | 0.946 | 0.949 | 0.946 | 0.947 | 0.949 | 0.946 | 0.947 |
| $p_5$ | 100 | 0.933 | 0.922 | 0.904 | 0.938 | 0.928 | 0.915 | 0.939 | 0.923 | 0.906 |
| | 250 | 0.945 | 0.939 | 0.931 | 0.947 | 0.942 | 0.935 | 0.944 | 0.940 | 0.931 |
| | 500 | 0.946 | 0.945 | 0.941 | 0.947 | 0.946 | 0.943 | 0.947 | 0.945 | 0.941 |
| | 1000 | 0.948 | 0.947 | 0.945 | 0.949 | 0.948 | 0.946 | 0.949 | 0.947 | 0.945 |
| $p_6$ | 100 | 0.909 | 0.894 | 0.884 | 0.920 | 0.908 | 0.901 | 0.906 | 0.890 | 0.877 |
| | 250 | 0.933 | 0.923 | 0.908 | 0.936 | 0.929 | 0.919 | 0.931 | 0.921 | 0.905 |
| | 500 | 0.938 | 0.935 | 0.927 | 0.940 | 0.938 | 0.933 | 0.937 | 0.934 | 0.925 |
| | 1000 | 0.943 | 0.941 | 0.939 | 0.945 | 0.943 | 0.942 | 0.943 | 0.941 | 0.938 |

**Table A4.** Rejection rate if testing null hypothesis of uniform distribution ($m = 3$) on 5%-level based on $\hat{v}_{\mathrm{G}}$, $\hat{v}_{\mathrm{En}}$, $\hat{v}_{\mathrm{Ex}}$. DGPs i.i.d., DMA(1) with $\varphi_1 = 0.25$, and DAR(1) with $\phi_1 = 0.40$, with marginal distribution $L_m(\lambda)$.

| | **$\lambda$** | **1.00** | **0.98** | **0.96** | **0.94** | **0.92** | **0.90** | **1.00** | **0.98** | **0.96** | **0.94** | **0.92** | **0.90** | **1.00** | **0.98** | **0.96** | **0.94** | **0.92** | **0.90** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **DGP** | **$n$** | **Rejection rate $\hat{v}_{\mathrm{G}}$** | | | | | | **Rejection rate $\hat{v}_{\mathrm{En}}$** | | | | | | **Rejection rate $\hat{v}_{\mathrm{Ex}}$** | | | | | |
| i.i.d. | 100 | 0.049 | 0.056 | 0.081 | 0.122 | 0.183 | 0.271 | 0.050 | 0.057 | 0.082 | 0.120 | 0.178 | 0.262 | 0.050 | 0.058 | 0.084 | 0.127 | 0.190 | 0.281 |
| | 250 | 0.052 | 0.068 | 0.132 | 0.250 | 0.421 | 0.605 | 0.051 | 0.068 | 0.129 | 0.243 | 0.407 | 0.590 | 0.051 | 0.068 | 0.132 | 0.252 | 0.423 | 0.608 |
| | 500 | 0.050 | 0.088 | 0.223 | 0.465 | 0.723 | 0.899 | 0.050 | 0.087 | 0.219 | 0.455 | 0.712 | 0.893 | 0.050 | 0.088 | 0.225 | 0.468 | 0.727 | 0.901 |
| | 1000 | 0.051 | 0.132 | 0.420 | 0.782 | 0.961 | 0.997 | 0.051 | 0.132 | 0.414 | 0.775 | 0.959 | 0.997 | 0.051 | 0.133 | 0.423 | 0.785 | 0.962 | 0.997 |
| DMA(1), 0.25 | 100 | 0.054 | 0.060 | 0.076 | 0.109 | 0.150 | 0.213 | 0.057 | 0.063 | 0.079 | 0.109 | 0.150 | 0.210 | 0.055 | 0.062 | 0.078 | 0.112 | 0.155 | 0.219 |
| | 250 | 0.052 | 0.065 | 0.108 | 0.194 | 0.317 | 0.467 | 0.052 | 0.067 | 0.108 | 0.190 | 0.309 | 0.455 | 0.052 | 0.066 | 0.110 | 0.197 | 0.322 | 0.474 |
| | 500 | 0.051 | 0.077 | 0.172 | 0.350 | 0.575 | 0.778 | 0.052 | 0.077 | 0.169 | 0.343 | 0.564 | 0.769 | 0.051 | 0.078 | 0.174 | 0.354 | 0.579 | 0.782 |
| | 1000 | 0.049 | 0.106 | 0.315 | 0.634 | 0.879 | 0.978 | 0.050 | 0.106 | 0.310 | 0.625 | 0.873 | 0.976 | 0.050 | 0.107 | 0.317 | 0.638 | 0.881 | 0.978 |
| DAR(1), 0.4 | 100 | 0.056 | 0.059 | 0.067 | 0.088 | 0.112 | 0.146 | 0.059 | 0.062 | 0.071 | 0.091 | 0.115 | 0.147 | 0.059 | 0.062 | 0.071 | 0.092 | 0.118 | 0.153 |
| | 250 | 0.053 | 0.060 | 0.085 | 0.132 | 0.199 | 0.294 | 0.054 | 0.062 | 0.085 | 0.131 | 0.195 | 0.286 | 0.054 | 0.062 | 0.087 | 0.135 | 0.205 | 0.301 |
| | 500 | 0.050 | 0.066 | 0.120 | 0.218 | 0.366 | 0.535 | 0.051 | 0.067 | 0.119 | 0.213 | 0.356 | 0.522 | 0.051 | 0.067 | 0.121 | 0.221 | 0.371 | 0.542 |
| | 1000 | 0.050 | 0.083 | 0.199 | 0.404 | 0.647 | 0.846 | 0.050 | 0.083 | 0.195 | 0.396 | 0.636 | 0.838 | 0.050 | 0.084 | 0.200 | 0.408 | 0.651 | 0.850 |

**Table A5.** Asymptotic vs. simulated standard deviation (S.-a vs. S.-s) of $\hat{\kappa}(h)$, $\hat{\kappa}^*(h)$, $\hat{\kappa}^\star(h)$ for i.i.d. DGPs with $m = 3$.

| | | $S_{\hat{\kappa}}$-a | $S_{\hat{\kappa}}$-s, $h =$ | | | $S_{\hat{\kappa}^*}$-a | $S_{\hat{\kappa}^*}$-s, $h =$ | | | $S_{\hat{\kappa}^\star}$-a | $S_{\hat{\kappa}^\star}$-s, $h =$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PMF** | $n$ | | **1** | **2** | **3** | | **1** | **2** | **3** | | **1** | **2** | **3** |
| $p_5$ | 100 | 0.064 | 0.064 | 0.065 | 0.066 | 0.058 | 0.056 | 0.057 | 0.057 | 0.074 | 0.075 | 0.077 | 0.078 |
| | 250 | 0.040 | 0.041 | 0.041 | 0.041 | 0.037 | 0.036 | 0.036 | 0.036 | 0.047 | 0.047 | 0.048 | 0.048 |
| | 500 | 0.029 | 0.029 | 0.029 | 0.029 | 0.026 | 0.026 | 0.026 | 0.026 | 0.033 | 0.033 | 0.033 | 0.034 |
| | 1000 | 0.020 | 0.020 | 0.020 | 0.020 | 0.018 | 0.018 | 0.018 | 0.018 | 0.023 | 0.023 | 0.024 | 0.024 |
| $p_6$ | 100 | 0.060 | 0.060 | 0.061 | 0.061 | 0.058 | 0.057 | 0.057 | 0.058 | 0.063 | 0.064 | 0.064 | 0.065 |
| | 250 | 0.038 | 0.038 | 0.038 | 0.038 | 0.037 | 0.036 | 0.036 | 0.037 | 0.040 | 0.040 | 0.040 | 0.040 |
| | 500 | 0.027 | 0.027 | 0.027 | 0.027 | 0.026 | 0.026 | 0.026 | 0.026 | 0.028 | 0.028 | 0.028 | 0.028 |
| | 1000 | 0.019 | 0.019 | 0.019 | 0.019 | 0.018 | 0.018 | 0.018 | 0.018 | 0.020 | 0.020 | 0.020 | 0.020 |
| $p_7$ | 100 | 0.058 | 0.058 | 0.059 | 0.059 | 0.058 | 0.057 | 0.058 | 0.058 | 0.058 | 0.058 | 0.059 | 0.059 |
| | 250 | 0.037 | 0.036 | 0.037 | 0.037 | 0.037 | 0.036 | 0.037 | 0.036 | 0.037 | 0.037 | 0.037 | 0.037 |
| | 500 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 | 0.026 |
| | 1000 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 |

**Table A6.** Rejection rate (RR) if testing null hypothesis of i.i.d. data on 5%-level based on $\hat{\kappa}(1)$, $\hat{\kappa}^*(1)$, $\hat{\kappa}^\star(1)$. DGPs i.i.d., DMA(1) with $\varphi_1 = 0.15$, DAR(1) with $\phi_1 = 0.15$, and NegMarkov with $\alpha = 0.75$.

| | **DGP** | **i.i.d., RR for** | | | **DMA(1), 0.15, RR for** | | | **DAR(1), 0.15, RR for** | | | **NMark, 0.75, RR for** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PMF** | $n$ | $\hat{\kappa}(1)$ | $\hat{\kappa}^*(1)$ | $\hat{\kappa}^\star(1)$ | $\hat{\kappa}(1)$ | $\hat{\kappa}^*(1)$ | $\hat{\kappa}^\star(1)$ | $\hat{\kappa}(1)$ | $\hat{\kappa}^*(1)$ | $\hat{\kappa}^\star(1)$ | $\hat{\kappa}(1)$ | $\hat{\kappa}^*(1)$ | $\hat{\kappa}^\star(1)$ |
| $p_5$ | 100 | 0.047 | 0.042 | 0.051 | 0.484 | 0.537 | 0.400 | 0.594 | 0.634 | 0.507 | 0.478 | 0.235 | 0.521 |
| | 250 | 0.049 | 0.048 | 0.049 | 0.851 | 0.896 | 0.756 | 0.925 | 0.947 | 0.861 | 0.886 | 0.654 | 0.908 |
| | 500 | 0.049 | 0.049 | 0.050 | 0.988 | 0.994 | 0.961 | 0.997 | 0.999 | 0.989 | 0.995 | 0.936 | 0.997 |
| | 1000 | 0.049 | 0.048 | 0.049 | 1.000 | 1.000 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 1.000 |
| $p_6$ | 100 | 0.048 | 0.047 | 0.049 | 0.540 | 0.559 | 0.509 | 0.662 | 0.673 | 0.631 | 0.338 | 0.268 | 0.353 |
| | 250 | 0.050 | 0.049 | 0.049 | 0.903 | 0.916 | 0.878 | 0.960 | 0.966 | 0.947 | 0.725 | 0.636 | 0.735 |
| | 500 | 0.050 | 0.049 | 0.051 | 0.996 | 0.997 | 0.992 | 0.999 | 1.000 | 0.999 | 0.958 | 0.920 | 0.961 |
| | 1000 | 0.049 | 0.050 | 0.050 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 1.000 |
| $p_7$ | 100 | 0.048 | 0.047 | 0.048 | 0.577 | 0.569 | 0.573 | 0.699 | 0.688 | 0.697 | 0.275 | 0.269 | 0.276 |
| | 250 | 0.047 | 0.048 | 0.048 | 0.925 | 0.924 | 0.924 | 0.973 | 0.972 | 0.973 | 0.640 | 0.636 | 0.641 |
| | 500 | 0.049 | 0.049 | 0.049 | 0.998 | 0.998 | 0.998 | 1.000 | 1.000 | 1.000 | 0.918 | 0.916 | 0.918 |
| | 1000 | 0.050 | 0.050 | 0.050 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 0.998 | 0.998 |

## References


Agresti, Alan. 2002. *Categorical Data Analysis*, 2nd ed. Hoboken: John Wiley & Sons, Inc.

Billingsley, Patrick. 1999. *Convergence of Probability Measures*, 2nd ed. New York: John Wiley & Sons, Inc.

Blyth, Colin R. 1959. Note on estimating information. *Annals of Mathematical Statistics* 30: 71–79. [CrossRef]

Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20: 37–46. [CrossRef]

Craig, Wallace. 1943. *The Song of the Wood Pewee, Myiochanes virens Linnaeus: A Study of Bird Music*. New York State Museum Bulletin No. 334. Albany: University of the State of New York.

GESIS—Leibniz Institute for the Social Sciences. 2018. *ALLBUS 1980–2016 (German General Social Survey)*. ZA4586 Data File (Version 1.0.0). Cologne: GESIS Data Archive. (In German)

Hancock, Gwendolyn D'Anne. 2012. VIX and VIX futures pricing algorithms: Cultivating understanding. *Modern Economy* 3: 284–94.

Jacobs, Patricia A., and Peter A. W. Lewis. 1983. Stationary discrete autoregressive-moving average time series generated by mixtures. *Journal of Time Series Analysis* 4: 19–36. [CrossRef]

Kvålseth, Tarald O. 1995. Coefficients of variation for nominal and ordinal categorical data. *Perceptual and Motor Skills* 80: 843–47. [CrossRef]

Kvålseth, Tarald O. 2011a. The lambda distribution and its applications to categorical summary measures. *Advances and Applications in Statistics* 24: 83–106.

Kvålseth, Tarald O. 2011b. Variation for categorical variables. In *International Encyclopedia of Statistical Science*. Edited by Miodrag Lovric. Berlin: Springer, pp. 1642–45.

Lad, Frank, Giuseppe Sanfilippo, and Gianna Agro. 2015. Extropy: Complementary dual of entropy. *Statistical Science* 30: 40–58. [CrossRef]

Love, Eric Russell. 1980. Some logarithm inequalities. *Mathematical Gazette* 64: 55–57. [CrossRef]

Rao, C. Radhakrishna. 1982. Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology* 21: 24–43. [CrossRef]

Shannon, Claude Elwood. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423, 623–56. [CrossRef]

Shorrocks, Anthony F. 1978. The measurement of mobility. *Econometrica* 46: 1013–24. [CrossRef]

Tan, Wai-Yuan. 1977. On the distribution of quadratic forms in normal random variables. *Canadian Journal of Statististics* 5: 241–50. [CrossRef]

Weiß, Christian H. 2011. Empirical measures of signed serial dependence in categorical time series. *Journal of Statistical Computation and Simulation* 81: 411–29. [CrossRef]

Weiß, Christian H. 2013. Serial dependence of NDARMA processes. *Computational Statistics and Data Analysis* 68: 213–38. [CrossRef]

Weiß, Christian H. 2018. *An Introduction to Discrete-Valued Time Series*. Chichester: John Wiley & Sons, Inc.

Weiß, Christian H., and Rainer Göb. 2008. Measuring serial dependence in categorical time series. *AStA Advances in Statistical Analysis* 92: 71–89. [CrossRef]

© 2019 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 (http://creativecommons.org/licenses/by/4.0/).