# Measuring Dispersion and Serial Dependence in Ordinal Time Series based on Cumulative Paired $\phi$-Entropy

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

**Christian H. Weiß**

Department of Mathematics & Statistics,

Helmut Schmidt University, Hamburg

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Dispersion & Serial Dependence in Ordinal Time Series

## Introduction

Random variable $X$ **ordinal**

if bounded qualitative range with natural order:

$\mathcal{S} = \{s_0, s_1, \ldots, s_m\}$ with $m \in \mathbb{N} = \{1, 2, \ldots\}$ and $s_0 < \ldots < s_m$.

Realized data: $x_1, \ldots, x_n$ with $n \in \mathbb{N}$,

either from independent and identically distributed (i. i. d.)

replications of $X$ (ordinal **random sample**),

or from stationary ordinal process $(X_t)_{\mathbb{Z}=\{\ldots,-1,0,1,\ldots\}}$

(ordinal **time series**).

(Agresti, 2010; Weiß, 2018)

Variance or mean absolute deviation only for *quantitative* data.

Therefore, several tailor-made measures for ordinal dispersion in the literature, see Kiesl (2003), Kvålseth (2011), and others.

All measures rely on cumulative distribution function (CDF), i. e., on $\boldsymbol{f} = (f_0, \ldots, f_{m-1})^\top$ with $f_i = P(X \leq s_i)$.

They classify any one-point distribution on $\mathcal{S}$ as a scenario of **minimal dispersion** (maximal consensus):

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

**Maximal dispersion** iff extreme two-point distribution, $\boldsymbol{f}_{\text{two}} = (\frac{1}{2}, \ldots, \frac{1}{2})^\top$ (maximal dissent).

Ordinal **dependence measures** defined in analogy to autocorrelation function (ACF) for quantitative time series, namely using 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.

In particular, measures related to **Cohen's** $\kappa$ distinguish positive and negative forms of serial dependence, i. e., extent of (dis)agreement between lagged observations.

(Weiß, 2020)

# Outline



Klein et al. (2016) showed that

many ordinal dispersion measures belong to family

of "cumulative paired $\phi$-entropies", abbreviated as $\mathrm{CPE}_\phi$.

**Task 1:** derive asymptotic distribution of sample version $\widehat{\mathrm{CPE}}_\phi$

for time series data.

**Task 2:** for given $\mathrm{CPE}_\phi$, identify related measure of

serial dependence $\Rightarrow$ novel family of $\kappa_\phi(h)$ measures.

**Task 3:** asymptotics of $\widehat{\kappa}_\phi(h)$ under null of i. i. d. time series.

Check finite-sample performance by simulations,

illustrative real-world data example.

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH
STAT

# The Family of Cumulative Paired $\phi$-Entropies

## Definition & Examples

# Cumulative Paired $\phi$-Entropies



Klein et al. (2016): (normalized) **cumulative paired $\phi$-entropy**

$$\mathrm{CPE}_\phi(\boldsymbol{f}) \;=\; \frac{1}{2\,m\,\phi(1/2)}\sum_{i=0}^{m-1}\Big(\phi(f_i)+\phi(1-f_i)\Big),$$

where **entropy generating function** (EGF) $\phi$ defined on $[0;1]$, satisfies $\phi(0)=\phi(1)=0$, and concave on $[0;1]$.

**Example:** If $q$**-entropies** $\phi_q(z)=1-|2\,z-1|^q$, then $q=2$, i. e., $\phi(z)=z(1-z)$, leads to **index of ordinal variation** (Kiesl, 2003):

$$\mathrm{IOV} \;=\; \tfrac{4}{m}\sum_{i=0}^{m-1} f_i(1-f_i) \;=\; 1-\tfrac{1}{m}\sum_{i=0}^{m-1}(2f_i-1)^2.$$

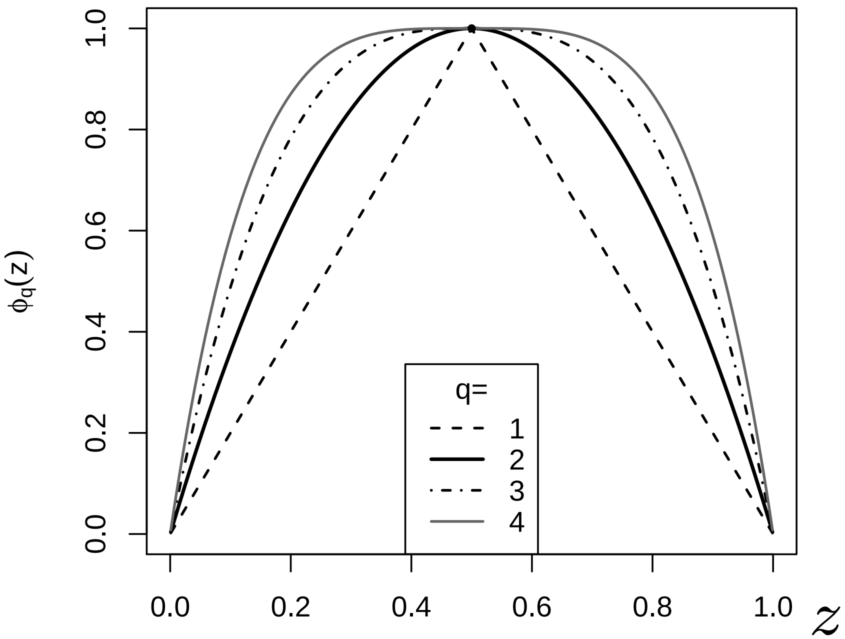
**Example:** If $a$-**entropies** $\phi_a(z) = \frac{z - z^a}{a-1}$, $a > 0$ and $a \neq 1$ (Havrda & Charvát, 1967), then

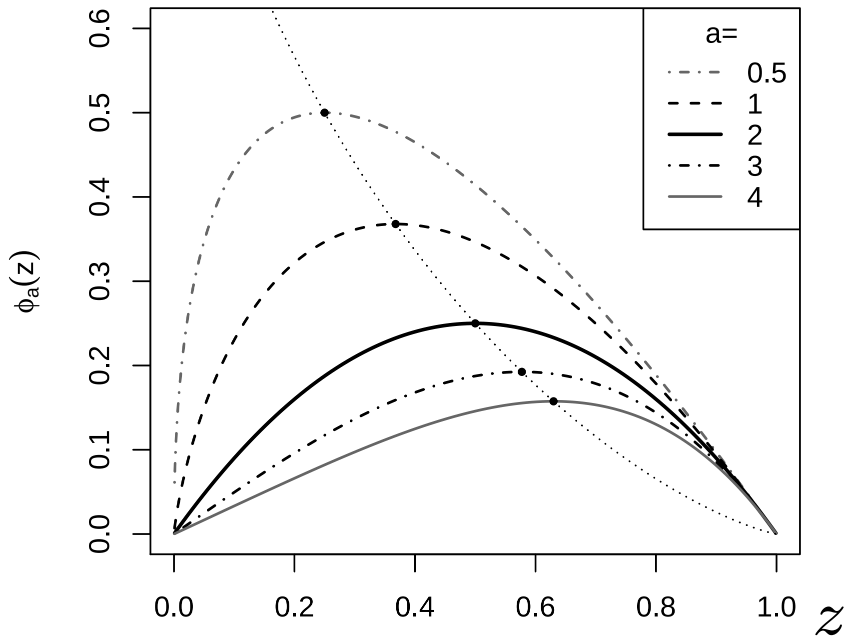
$$\mathrm{CPE}_a = \frac{2^{a-1}}{m\,(2^{a-1}-1)} \sum_{i=0}^{m-1} \Big(1 - f_i^a - (1-f_i)^a\Big).$$

Boundary case $a \to 1$, i. e., $\phi(z) = -z \ln z$, leads to **cumulative paired (Shannon) entropy** (Vogel & Dobbener, 1982):

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

Plot of EGFs $\phi(z)$ against $z$:

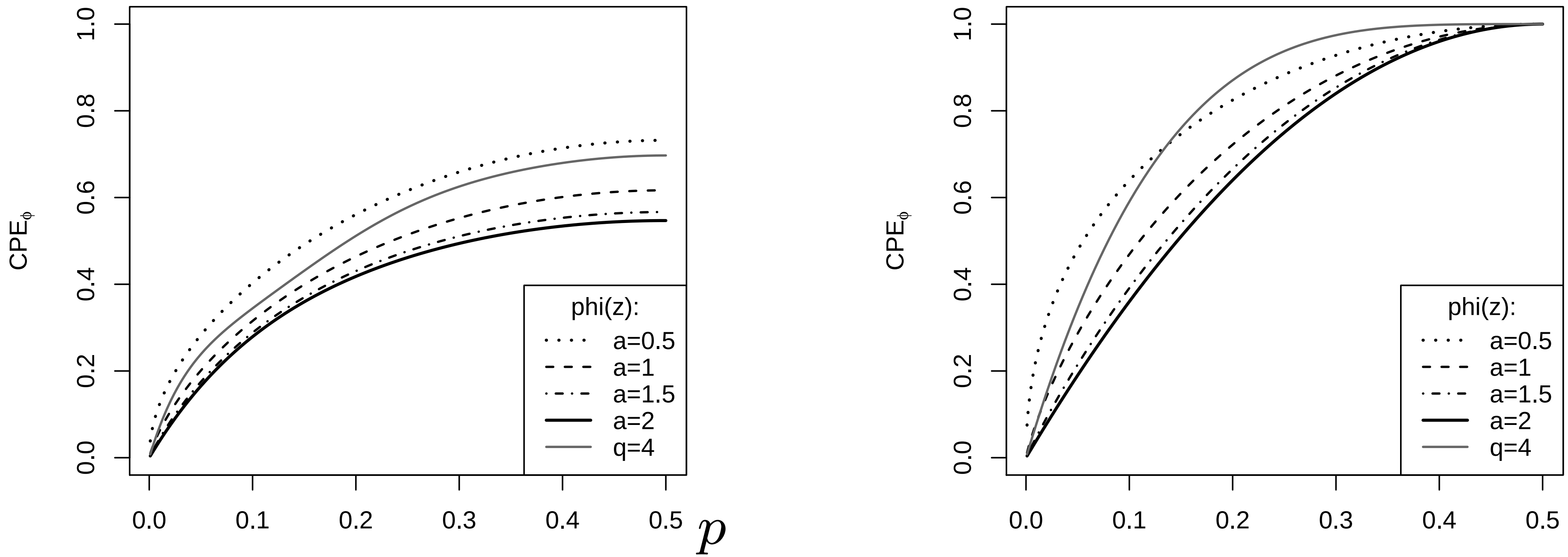



Left: $\phi_q(z) = 1 - |2\,z - 1|^q$;   right: $\phi_a(z) = (z - z^a)/(a - 1)$.

Plot of $\mathrm{CPE}_\phi$ against $p$ for specific cases of

$\phi_a(z) = (z - z^a)/(a-1)$ and $\phi_q(z) = 1 - |2\,z - 1|^q$




Left: Bin$(4, p)$-distribution; right: two-point distr. with $f_0 = p$.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Asymptotic Distribution of Sample CPE$_\phi$

Derivation & Examples

# Asymptotic Distribution of Sample CPE$_\phi$



**Sample version** of CPE$_\phi$:

$\widehat{\text{CPE}}_\phi = \text{CPE}_\phi(\hat{\boldsymbol{f}})$, where sample CDF $\hat{\boldsymbol{f}}$ is vector of cumulative relative frequencies computed from $X_1, \ldots, X_n$.

**Starting point:** If DGP $(X_t)_\mathbb{Z}$ satisfies mixing conditions, e. g., $\alpha$-mixing with exponentially decreasing weights, then Weiß (2020) showed the CLT

$$\sqrt{n}\,\big(\hat{\boldsymbol{f}} - \boldsymbol{f}\big) \xrightarrow{\text{d}} \text{N}(\boldsymbol{0}, \boldsymbol{\Sigma}) \quad \text{with } \boldsymbol{\Sigma} = \big(\sigma_{ij}\big)_{i,j=0,\ldots,m-1}, \quad \text{where}$$

$$\sigma_{ij} \overset{\text{a}}{=} f_{\min\{i,j\}} - f_i\, f_j \;+\; \sum_{h=1}^{\infty} \Big(f_{ij}(h) + f_{ji}(h) - 2\, f_i\, f_j\Big).$$

Finite (co-)variances if $\sum_{h=1}^{\infty} \big(f_{ij}(h) - f_i\, f_j\big) < \infty$ ("short memory").

# Asymptotic Distribution of Sample $\mathrm{CPE}_\phi$



Assuming $\phi$ twice differentiable, then partial derivatives

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

Thus, Hessian $\mathbf{H}=\mathrm{diag}(h_{00},\ldots,h_{m-1,m-1})$ diagonal matrix.

Now, second-order Taylor expansion

$$
\mathrm{CPE}_\phi\big(\widehat{\boldsymbol{f}}\big) \;\approx\; \mathrm{CPE}_\phi(\boldsymbol{f})+\textstyle\sum_{k=0}^{m-1} d_k\,\big(\widehat{f}_k-f_k\big)+\frac{1}{2}\textstyle\sum_{k=0}^{m-1} h_{kk}\,\big(\widehat{f}_k-f_k\big)^2.
$$

Under above mixing assumptions, if $\phi$ twice differentiable,

and if $\mathbf{D}=(d_0,\ldots,d_{m-1})$ does not vanish, it holds that . . .

**Theorem 1:** ... it holds that

$$\sqrt{n}\Big(\mathsf{CPE}_\phi(\hat{\boldsymbol{f}}) - \mathsf{CPE}_\phi(\boldsymbol{f})\Big) \xrightarrow{\mathsf{d}} \mathsf{N}\big(0,\ \sigma_\phi^2\big),$$

$$\sigma_\phi^2 \;=\; \sigma_{\phi,\mathsf{iid}}^2\Big(1 \;+ 2\sum_{h=1}^{\infty} \vartheta_\phi(h)\Big)$$

with $\sigma_{\phi,\mathsf{iid}}^2 \;=\; \sum_{i,j=0}^{m-1} d_i\, d_j\,\big(f_{\min\{i,j\}} - f_i\, f_j\big)$

and $\vartheta_\phi(h) \;=\; \dfrac{\sum_{i,j=0}^{m-1} d_i\, d_j\,\big(f_{ij}(h) - f_i\, f_j\big)}{\sum_{i,j=0}^{m-1} d_i\, d_j\,\big(f_{\min\{i,j\}} - f_i\, f_j\big)}.$

(...)

**Theorem 1:** (continued)

In addition, bias-corrected mean of CPE$_\phi(\widehat{\boldsymbol{f}})$ is

$$E\big[\mathrm{CPE}_\phi(\widehat{\boldsymbol{f}})\big] \approx \mathrm{CPE}_\phi(\boldsymbol{f}) + \tfrac{1}{2\,n}\Big(\sum_{i=0}^{m-1} h_{ii}\, f_i(1-f_i)\Big)\Big(1 + 2\sum_{h=1}^{\infty} \kappa_\phi(h)\Big),$$

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

As second-order derivatives negative due to concavity of $\phi$, CPE$_\phi(\widehat{\boldsymbol{f}})$ exhibits negative bias.

# Asymptotic Distribution of Sample $\text{CPE}_\phi$



Bias of $\text{CPE}_\phi(\widehat{\boldsymbol{f}})$ affected by serial dependence via $\kappa_\phi(h)$,

i. e., $\kappa$-type measure reflecting

extent of (dis)agreement between lagged observations.

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

is weighted type of $\kappa_{\text{ord}}(h)$,

where weights $h_{ii}$ depend on choice of $\phi$.

$\Rightarrow$ novel family of measures of signed serial dependence!

**Examples:** If $\phi(z) = z(1-z)$, then CPE$_\phi(\boldsymbol{f}) =$ IOV,

and as $h_{ii} = -\frac{8}{m}$ constant,

$$\kappa_\phi(h) \;=\; \frac{\sum_{i=0}^{m-1}\big(f_{ii}(h) - f_i^2\big)}{\sum_{i=0}^{m-1} f_i(1-f_i)} \;=\; \kappa_{\text{ord}}(h),$$

i. e., ordinary $\kappa$ (Weiß, 2020). So pair $\big(\text{IOV}, \kappa_{\text{ord}}(h)\big)$.

If $\phi(z) = -z \ln z$, then CPE$_\phi(\boldsymbol{f}) =$ CPE, and

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

Thus, also $\big(\text{CPE}, \kappa^*_{\text{ord}}(h)\big)$ belongs to $\big(\text{CPE}_\phi, \kappa_\phi(h)\big)$-family.

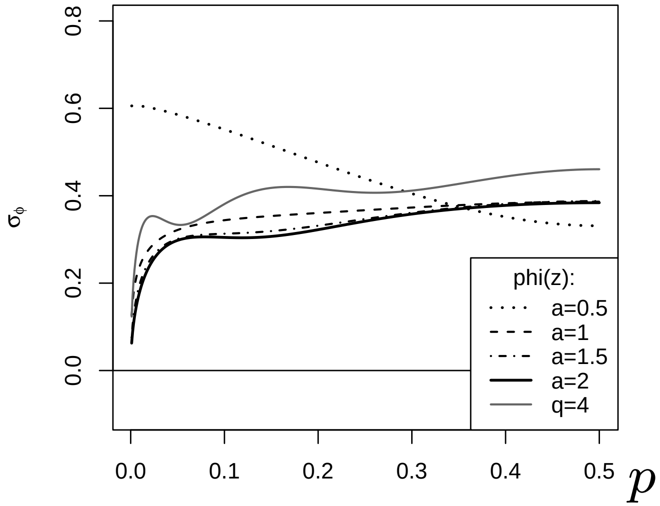
# Asymptotic Distribution of Sample CPE$_\phi$

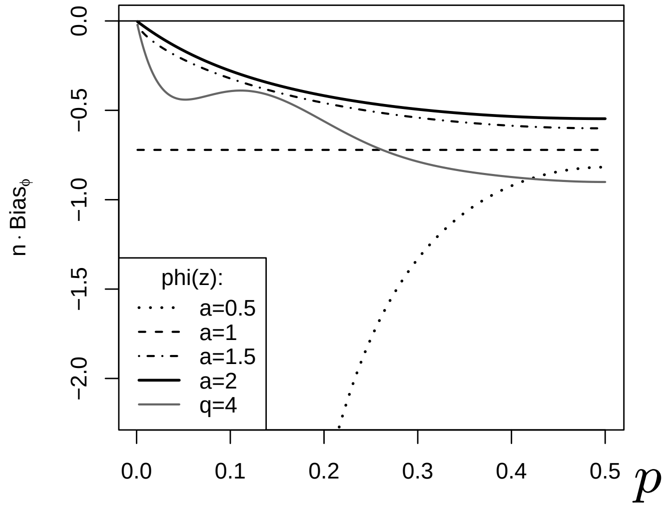


DGP:

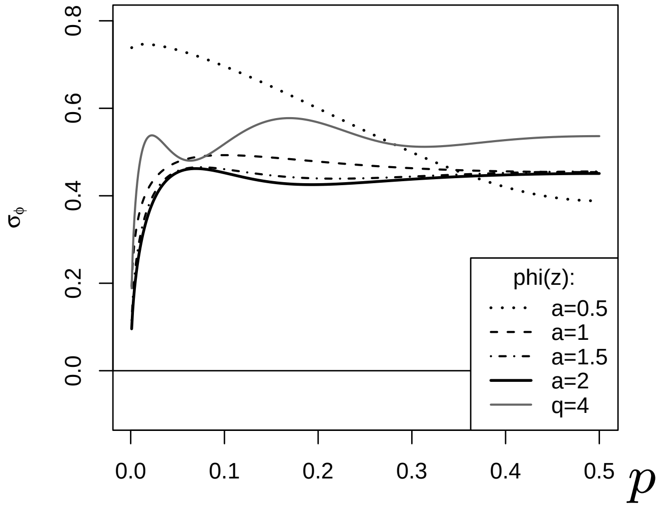
$\sigma_\phi$ | $n\left(E\left[\mathsf{CPE}_\phi\left(\widehat{\boldsymbol{f}}\right)\right]-\mathsf{CPE}_\phi(\boldsymbol{f})\right)$

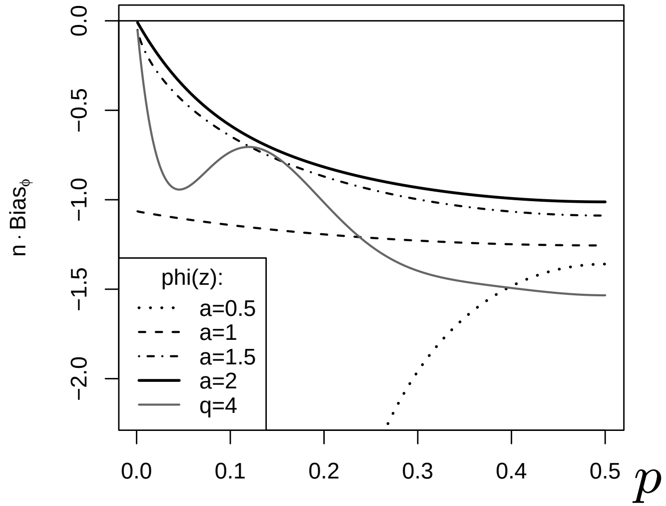
i. i. d. Bin$(4, p)$




BAR(1)
with $\rho = 0.4$



HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Asymptotic Distribution of Sample $\kappa_\phi(h)$

Derivation & Examples

Sample version $\widehat{\kappa}_\phi(h)$ of

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

also depends on bivariate sample CDF $\widehat{f}_{ii}(h)$.

Let $\boldsymbol{f}_{(h)} = \big(f_0,\ldots,f_{m-1},\ f_{00}(h),\ldots,f_{m-1,m-1}(h)\big)^\top$,
denote sample version by $\widehat{\boldsymbol{f}}_{(h)}$.

**CLT** in Weiß (2020):

$$\sqrt{n}\,\big(\widehat{\boldsymbol{f}}_{(h)}-\boldsymbol{f}_{(h)}\big) \;\stackrel{\mathrm{d}}{\to}\; \mathrm{N}\big(\mathbf{0},\boldsymbol{\Sigma}^{(h)}\big) \quad \text{with } \boldsymbol{\Sigma}^{(h)} = \big(\sigma_{i,j}^{(h)}\big)_{i,j=0,\ldots,2m-1}.$$

# Asymptotic Distribution of Sample $\kappa_\phi(h)$



**Aim:** use $\widehat{\kappa}_\phi(h)$ to identify significant dependence in $(X_t)$.

Thus, to test null of independence (i. e., $\kappa_\phi(h) = 0$), asymptotics of $\widehat{\kappa}_\phi(h)$ for i. i. d. $(X_t)$ sufficient.

Simplified $\mathbf{\Sigma}^{(h)}$ in Weiß (2020), with $i, j \in \{0, \ldots, m-1\}$:

$$\sigma_{i,j}^{(h)} = \sigma_{i,j} \;=\; f_{\min\{i,j\}} - f_i\, f_j,$$

$$\sigma_{i,m+j}^{(h)} \;=\; 2\, f_j \big(f_{\min\{i,j\}} - f_i\, f_j\big),$$

$$\sigma_{m+i,m+j}^{(h)} \;=\; \big(f_{\min\{i,j\}} + 3\, f_i\, f_j\big)\big(f_{\min\{i,j\}} - f_i\, f_j\big).$$

Use second-order Taylor expansion for $\widehat{\kappa}_\phi(h)$. Higher-order derivatives of $\phi$ cancel out, still only derivatives of order 2.

## Asymptotic Distribution of Sample $\kappa_\phi(h)$



**Theorem 2:** Under null hypothesis of i. i. d. data,

i. e., if $\kappa_\phi(h) = 0$ for all lags $h \in \mathbb{N}$,

and assuming EGF $\phi$ twice differentiable,

it holds that

$$\sqrt{n}\,\Big(\widehat{\kappa}_\phi(h) - \kappa_\phi(h)\Big) \;\xrightarrow{\text{d}}\; \mathrm{N}\big(0,\ \sigma_\kappa^2\big)$$

$$\text{with } \sigma_\kappa^2 \;=\; \sum_{j,k=0}^{m-1} u_j\, u_k\, \big(f_{\min\{j,k\}} - f_j\, f_k\big)^2,$$

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

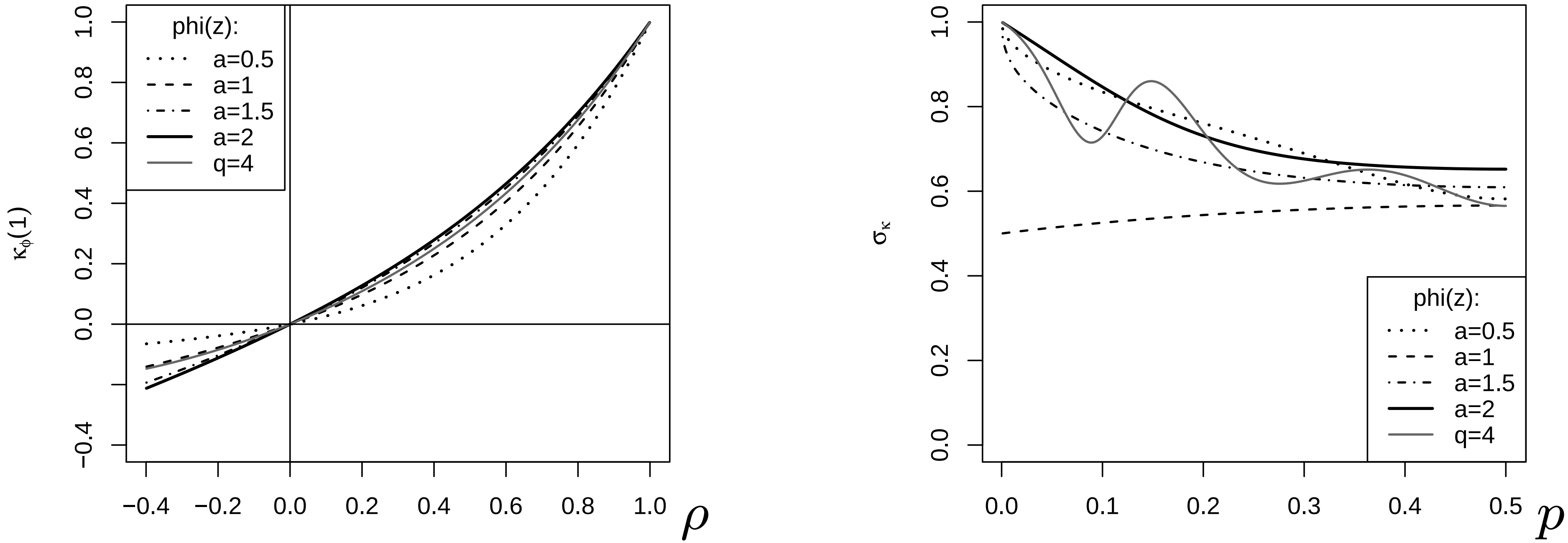
Bias-corrected mean of $\widehat{\kappa}_\phi(h)$ is $\quad E\big[\widehat{\kappa}_\phi(h)\big] \;\approx\; -\frac{1}{n}.$

# Asymptotic Distribution of Sample $\kappa_\phi(h)$



Plots for specific cases of $\phi_a(z) = (z - z^a)/(a-1)$

and $\phi_q(z) = 1 - |2\,z - 1|^q$:




$\kappa_\phi(1)$ against BAR(1)'s dependence parameter $\rho$ with marginal Bin$(4, 0.3)$ (left); $\sigma_\kappa$ against $p$ of marginal Bin$(4, p)$ (right).

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Finite-Sample Performance of Asymptotics

Simulation Results

# Simulation Results



Always $10^4$ replications per scenario,

choices $a = 1, 3/2, 2, 5/2$ as well as $q = 4$,

$X_t = s_{I_t}$ with $I_t \sim \text{Bin}(m, p)$ according to BAR(1) DGP.

**Full simulation results** in Weiß (2022),

few illustrative results in sequel.

# Simulation Results



Simulated vs. asymptotic mean and SE of $\widehat{\mathrm{CPE}}_\phi$,

with $p = 0.3$, $\rho = 0.4$, and sample size $n$.

| | simulated mean | | | | | asymptotic mean | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a=1$ | $a=1.5$ | $a=2$ | $a=2.5$ | $q=4$ | $a=1$ | $a=1.5$ | $a=2$ | $a=2.5$ | $q=4$ |
| 50 | 0.529 | 0.491 | 0.476 | 0.473 | 0.598 | 0.528 | 0.491 | 0.476 | 0.472 | 0.597 |
| 100 | 0.540 | 0.501 | 0.485 | 0.481 | 0.611 | 0.540 | 0.501 | 0.485 | 0.481 | 0.611 |
| 250 | 0.547 | 0.506 | 0.490 | 0.487 | 0.620 | 0.548 | 0.507 | 0.490 | 0.487 | 0.620 |
| 500 | 0.551 | 0.509 | 0.493 | 0.489 | 0.623 | 0.550 | 0.509 | 0.492 | 0.489 | 0.623 |
| 1000 | 0.552 | 0.510 | 0.493 | 0.490 | 0.624 | 0.551 | 0.510 | 0.493 | 0.490 | 0.624 |

| | simulated SE | | | | | asymptotic SE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a=1$ | $a=1.5$ | $a=2$ | $a=2.5$ | $q=4$ | $a=1$ | $a=1.5$ | $a=2$ | $a=2.5$ | $q=4$ |
| 50 | 0.064 | 0.061 | 0.061 | 0.060 | 0.071 | 0.065 | 0.063 | 0.062 | 0.062 | 0.073 |
| 100 | 0.046 | 0.044 | 0.043 | 0.043 | 0.050 | 0.046 | 0.044 | 0.044 | 0.044 | 0.051 |
| 250 | 0.029 | 0.028 | 0.027 | 0.027 | 0.032 | 0.029 | 0.028 | 0.028 | 0.028 | 0.032 |
| 500 | 0.021 | 0.020 | 0.019 | 0.019 | 0.023 | 0.021 | 0.020 | 0.020 | 0.020 | 0.023 |
| 1000 | 0.014 | 0.014 | 0.014 | 0.014 | 0.016 | 0.015 | 0.014 | 0.014 | 0.014 | 0.016 |

Simulated vs. asymptotic mean and SE of $\widehat{\kappa}_\phi(1)$,

i. i. d. Bin$(4, p)$ with $p = 0.3$ and sample size $n$.

| | simulated mean | | | | | asymptotic mean | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ |
| 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 |

| | simulated SE | | | | | asymptotic SE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ | $a = 1$ | $a = 1.5$ | $a = 2$ | $a = 2.5$ | $q = 4$ |
| 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 |

Rejection rate ($\rho = 0$: size; $\rho \neq 0$: power) of $\widehat{\kappa}_{\phi}(1)$-test:

| $n$ | simulated rejection rate $a=1$ | $a=1.5$ | $a=2$ | $a=2.5$ | $q=4$ | simulated rejection rate $a=1$ | $a=1.5$ | $a=2$ | $a=2.5$ | $q=4$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\rho = -0.4$ | | | | | $\rho = 0.4$ | | | | |
| 50 | 0.484 | 0.565 | 0.585 | 0.588 | 0.312 | 0.663 | 0.749 | 0.756 | 0.757 | 0.625 |
| 100 | 0.807 | 0.878 | 0.891 | 0.893 | 0.672 | 0.917 | 0.959 | 0.962 | 0.962 | 0.897 |
| 250 | 0.989 | 0.999 | 1.000 | 1.000 | 0.987 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 |
| 500 | 0.997 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 1000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | $\rho = 0$ | | | | | $\rho = 0.8$ | | | | |
| 50 | *0.061* | *0.056* | *0.054* | *0.054* | *0.061* | 0.997 | 1.000 | 1.000 | 1.000 | 0.981 |
| 100 | *0.056* | *0.060* | *0.058* | *0.057* | *0.057* | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 |
| 250 | *0.047* | *0.048* | *0.046* | *0.046* | *0.049* | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 500 | *0.049* | *0.050* | *0.050* | *0.050* | *0.052* | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 1000 | *0.053* | *0.047* | *0.048* | *0.049* | *0.050* | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Daily Air Quality Level in Shanghai

Data Application

# Data Application



**Daily air-quality** time series $x_1, \ldots, x_n$ in Shanghai for period Dec. 2013–July 2019 ($n = 2\,068$), with levels $s_0 =$ "excellent" to $s_5 =$ "severely polluted" (Liu et al., 2021).

**. . . see plots on next slide!**

Uni-modal shape with mode (=median) in $s_1 =$ "good".

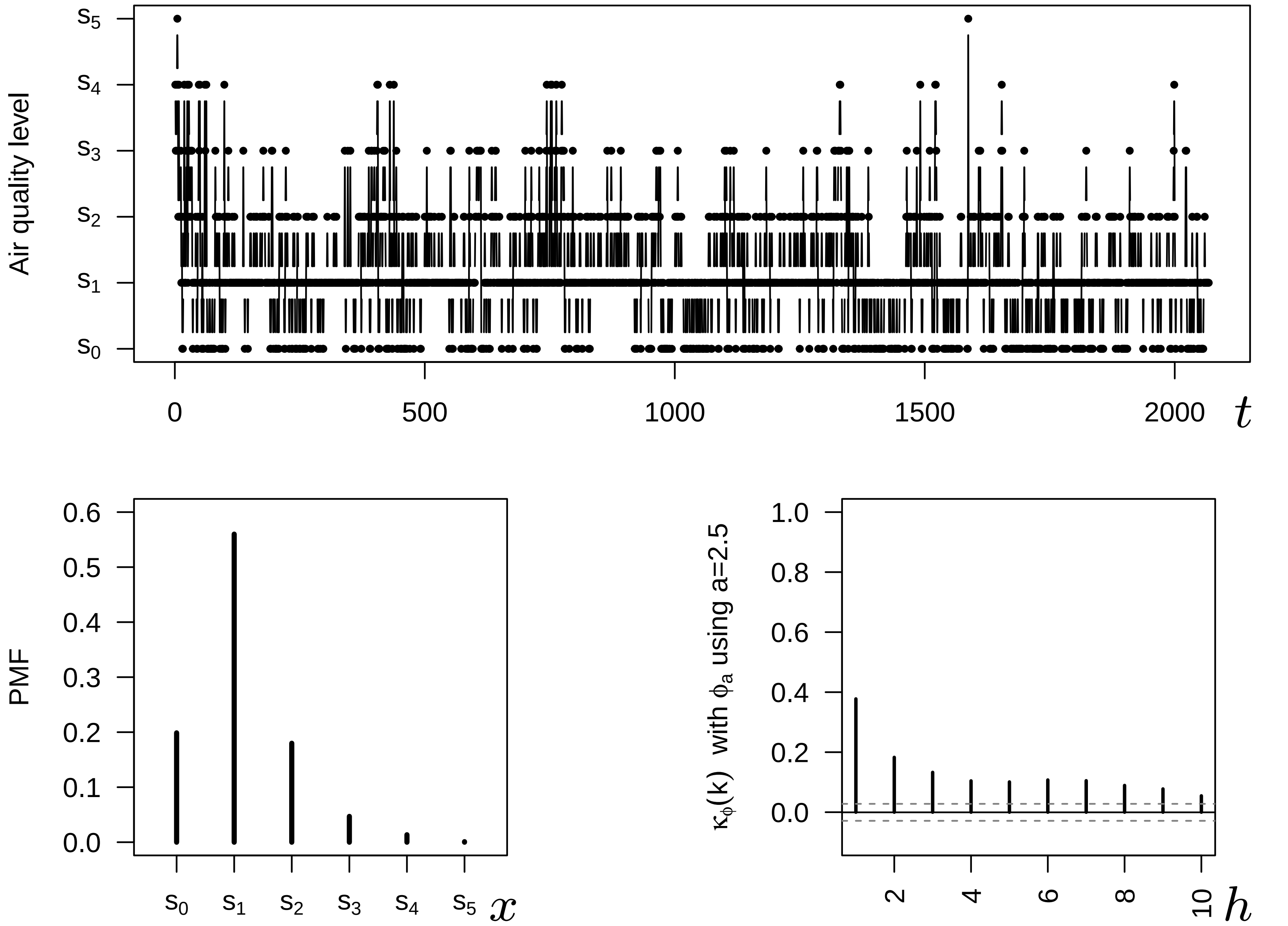
Significant $\widehat{\kappa}_\phi(h)$-values (recommended choice $a = 5/2$).

Medium level of dependence ($\widehat{\kappa}_\phi(1) \approx 0.378$),

quickly decreases with increasing $h$ ($\rightarrow$ AR-type process).

# Data Application






| $\widehat{\mathrm{CPE}}_\phi$ |
|---|
| $a = 1/2$: 0.514 |
| $q = 4$: 0.465 |
| $a = 1$: 0.394 |
| $a = 3/2$: 0.349 |
| $a = 2$: 0.332 |
| $a = 5/2$: 0.328 |

Compute **asymptotics** of $\widehat{\text{CPE}}_\phi$ in example case $a = 5/2$:

Start with **i. i. d.-approximations**

of bias, $\frac{1}{2n}\Big(\sum_{i=0}^{m-1} \widehat{h}_{ii}\, \widehat{f}_i(1-\widehat{f}_i)\Big) \approx -1.54 \cdot 10^{-4}$,

and standard error, $n^{-1/2}\, \widehat{\sigma}_{\phi,\text{iid}} \approx 7.83 \cdot 10^{-3}$.

Liu et al. (2021) use so-called “ZOBPAR model”.

For their model fit,

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

Thus, approximate 95 %-confidence interval (CI) for $\text{CPE}_\phi$

is given by $\approx (0.308; 0.349)$.

# Conclusions



- Asymptotic distribution of $\widehat{\mathrm{CPE}}_\phi$ for ordinal time series data. Useful, e. g., to compute approximate confidence intervals.

- Asymptotic bias of $\widehat{\mathrm{CPE}}_\phi$ implies $\kappa$-type serial dependence measure, i. e., matched pair $\big(\mathrm{CPE}_\phi, \kappa_\phi(h)\big)$ for each EGF $\phi$.

- Asymptotics of sample $\widehat{\kappa}_\phi(h)$ under i. i. d.-null, used to test for significant serial dependence in ordinal time series.

- Choosing $\widehat{\kappa}_\phi(h)$ based on $a$-entropy with $a \geq 2$, such as $a = 5/2$, ensures good finite-sample properties.

# Thank You

# for Your Interest!

**Christian H. Weiß**

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

`weissc@hsu-hh.de`

# Literature

**Weiß (2022) Measuring dispersion and serial dependence in ordinal time series based on the cumulative paired $\phi$-entropy. *Entropy* 24(1), 42.** **($\rightarrow$ open access)**

Agresti (2010) *Analysis of Ordinal Categorical Data*. 2nd edition, Wiley.

Havrda & Charvát (1967) Quantification method . . . *Kybernetika* **3**, 30–35.

Kiesl (2003) *Ordinale Streuungsmaße* . . . (in German). Josef Eul Verlag.

Klein et al. (2016) Cumulative paired $\phi$-entropy. *Entropy* **18**, 248.

Kvålseth (2011) Variation for . . . *Int. Enc. Stat. Sc.*, Springer, 1642–1645.

Liu et al. (2021) Modeling normalcy-dominant . . . *JTSA*, forthcoming.

Vogel & Dobbener (1982) Ein Streuungsmaß . . . *JNS* **197**, 145–158.

Weiß (2018) *An Introduction to Discrete-Valued Time Series*. Wiley.

Weiß (2020) Distance-based analysis of ordinal . . . *JASA* **115**, 1189–1200.