Testing for Serial Dependence by Using Ordinal Patterns

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

Christian H. Weiß

Department of Mathematics & Statistics,
Helmut Schmidt University, Hamburg

MATH STAT

Alexander Schnurr

Department of Mathematics,
University of Siegen

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Ordinal Patterns in Continuously-distributed Real-valued Time Series

Introduction

Ordinal pattern (OP) introduced by Bandt & Pompe (2002).

Basic idea: map segments

$\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ of length m from

continuously distrib., real-valued process $(X_t)_{t\in\mathbb{Z}=\{\ldots,-1,0,1,\ldots\}}$

onto permutations from symmetric group S_m of order m,

where selected $\pi_t \in S_m = \{\pi^{[1]}, \ldots, \pi^{[m!]}\}$ expresses

order among values in $\boldsymbol{X}_t$ in certain way: (...).

Rank representation, see Berger et al. (2019):

Entries of $\pi = (r_1, \ldots, r_m) \in S_m$ express

ranks within $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$, i. e.,

$$r_k < r_l \qquad \Leftrightarrow \qquad x_k < x_l \quad \text{or} \quad (x_k = x_l \text{ and } k < l)$$

for all $k, l \in \{1, \ldots, m\}$. Here, "$x_k = x_l$" if ties within $\boldsymbol{x}$.

Example: $(1.2, -0.7, 3.4, 1.9) \quad \mapsto \quad (2, 1, 4, 3)$,

$(1.2, -0.7, 3.4, -0.7) \quad \mapsto \quad (3, 1, 4, 2)$.

Marginal distribution of OP series (π_t) provides insights into **serial dependence structure** of original process (X_t).

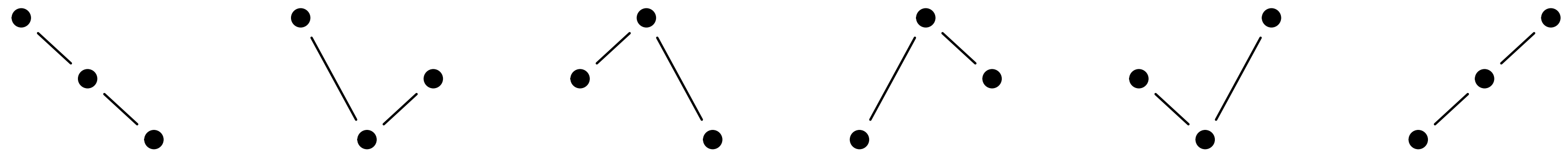
Focus on $m!$-dimensional **pmf vector** $\boldsymbol{p}$ (or sample pmf $\widehat{\boldsymbol{p}}$), with kth component being $p_k = P(\pi_t = \pi^{[k]})$.

Here, order m of OPs (thus dimension $m!$) chosen by user.

However, range of π_t quickly increases with m as $|S_m| = m!$, so estimation $\widehat{\boldsymbol{p}}$ of $\boldsymbol{p}$ quickly difficult in practice.

If $m = 2$: only downward OP $(2,1)$ and upward OP $(1,2)$.
Convenient choice is $m = 3$ (Bandt, 2019):

$(3,2,1)$ $(3,1,2)$ $(2,3,1)$ $(1,3,2)$ $(2,1,3)$ $(1,2,3)$



HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Testing for Serial Dependence in Continuously-distributed Real-valued Time Series

Approaches & Asymptotics

Let (X_t) be *continuously distributed* real-valued process, independent and identically distributed (i. i. d.) *under null*. Probability of ties $= 0$, so ties at most rarely in data.

Following **properties** crucial for dependence tests:

1. OPs invariant w.r.t. strictly monotonically increasing transformations of X_t. Thus, OPs do not depend on actual marginal distribution of $(X_t)_{\mathbb{N}}$ (**distribution-free** approach).
2. $(X_t)_{\mathbb{N}}$ is i. i. d. under null ($\rightarrow$ exchangeability). Thus, π_t discrete uniform on S_m, i. e., $P\big(\pi_t = \pi\big) = 1/m!$ for each $\pi \in S_m$ (**no parameter estimation** required).

OP-test statistics built upon $\widehat{\boldsymbol{p}}$ computed from $\pi_1, \ldots, \pi_n$ with $n \in \mathbb{N} = \{1, 2, \ldots\}$, where π_t from $\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ for $t = 1, 2, \ldots, n$ (moving-window approach).

First, asymptotics of $\sqrt{n}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}_0\right)$ under i. i. d.-null required, where $\boldsymbol{p}_0 = (1/m!, \ldots, 1/m!)$.

Note: moving-window for (π_t) causes $(m-1)$-dependence!

Elsinger (2010), Weiß (2022): $\sqrt{n}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}_0\right) \;\to\; \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_m)$, where $\boldsymbol{\Sigma}_m$ computed explicitly by combinatorial arguments.

Then, distribution of OP-test statistics by Taylor approximations ("Delta method").

Asymptotics $\sqrt{n}\,\big(\widehat{\boldsymbol{p}} - \boldsymbol{p}_0\big) \;\to\; \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_m)$

$\boldsymbol{\Sigma}_m = (\sigma_{ij})_{i,j=1,\ldots,m!}$ given by

$$\sigma_{ij} \;=\; 1/m!\,\big(\delta_{ij} - 1/m!\big) \;+\; \sum_{h=1}^{m-1} \Big(p_{ij}(h) + p_{ji}(h) - 2/m!^2\Big).$$

Case $m = 3$**:**

$$\boldsymbol{\Sigma}_3 \;=\; \frac{1}{360} \begin{pmatrix} 46 & -23 & -23 & 7 & 7 & -14 \\ -23 & 28 & 10 & -2 & -20 & 7 \\ -23 & 10 & 28 & -20 & -2 & 7 \\ 7 & -2 & -20 & 28 & 10 & -23 \\ 7 & -20 & -2 & 10 & 28 & -23 \\ -14 & 7 & 7 & -23 & -23 & 46 \end{pmatrix}.$$

Possible OP-test statistics, see Bandt (2019), Weiß (2022):

- entropy $H(\widehat{\boldsymbol{p}}) = -\sum_{k=1}^{m!} \widehat{p}_k \ln \widehat{p}_k$,
- distance to white noise $\Delta^2(\widehat{\boldsymbol{p}}) = \sum_{k=1}^{m!} \big(\widehat{p}_k - 1/m!\big)^2$,
- extropy $H_{\text{ex}}(\widehat{\boldsymbol{p}}) = -\sum_{k=1}^{m!} \big(1-p_k\big) \ln \big(1-p_k\big)$.

Theorem: If $(X_t)_{\mathbb{Z}}$ i. i. d., then

$$n\,\Delta^2(\widehat{\boldsymbol{p}}), \quad -n\,\tfrac{2}{m!}\Big(H(\widehat{\boldsymbol{p}}) - \ln m!\Big),$$

and $$-\,2n\,(1-\tfrac{1}{m!})\Big(H_{\text{ex}}(\widehat{\boldsymbol{p}}) - (m!-1)\,\ln\big(\tfrac{m!}{m!-1}\big)\Big)$$

asymptotically distributed like $Q_m := \sum_{i=1}^{l} \lambda_i\, \chi^2_{r_i}$, where

$\lambda_1,\ldots,\lambda_l$ distinct eigenvalues of $\boldsymbol{\Sigma}_m$ (multiplicities $r_1,\ldots,r_l$).

Corollary: If $(X_t)_{\mathbb{Z}}$ i.i.d. and $m = 3$, then

$$n\,\Delta^2\big(\widehat{\boldsymbol{p}}\big), \quad -\tfrac{n}{3}\Big(H\big(\widehat{\boldsymbol{p}}\big) - \ln 6\Big), \quad -\tfrac{5n}{3}\Big(H_{\text{ex}}\big(\widehat{\boldsymbol{p}}\big) - 5\,\ln\big(\tfrac{6}{5}\big)\Big)$$

asymptotically distributed like

$$\tfrac{1}{12}(2+\sqrt{2})\cdot\chi_1^2 \;+\; \tfrac{2}{15}\cdot\chi_1^2 \;+\; \tfrac{1}{10}\cdot\chi_1^2 \;+\; \tfrac{1}{12}(2-\sqrt{2})\cdot\chi_1^2.$$

Asymptotic mean $\frac{17}{30}$ and variance $\frac{2}{9}$.

Furthermore, (...)

Corollary: (...) the statistics (proposed by Bandt, 2019)

up-down balance: $\widehat{\beta} = \widehat{p}_6 - \widehat{p}_1$,

persistence: $\widehat{\tau} = \widehat{p}_6 + \widehat{p}_1 - \frac{1}{3}$,

rotational asymmetry: $\widehat{\gamma} = \widehat{p}_5 + \widehat{p}_3 - \widehat{p}_4 - \widehat{p}_2$,

up-down scaling: $\widehat{\delta} = \widehat{p}_4 + \widehat{p}_5 - \widehat{p}_3 - \widehat{p}_2$,

satisfy

$$\sqrt{n}\,\widehat{\beta} \stackrel{a}{\sim} \mathrm{N}\big(0,\ 1/3\big), \qquad \sqrt{n}\,\widehat{\tau} \stackrel{a}{\sim} \mathrm{N}\big(0,\ 8/45\big),$$

$$\sqrt{n}\,\widehat{\gamma} \stackrel{a}{\sim} \mathrm{N}\big(0,\ 2/5\big), \qquad \sqrt{n}\,\widehat{\delta} \stackrel{a}{\sim} \mathrm{N}\big(0,\ 2/3\big).$$

Recall: $(3,2,1)$ $(3,1,2)$ $(2,3,1)$ $(1,3,2)$ $(2,1,3)$ $(1,2,3)$

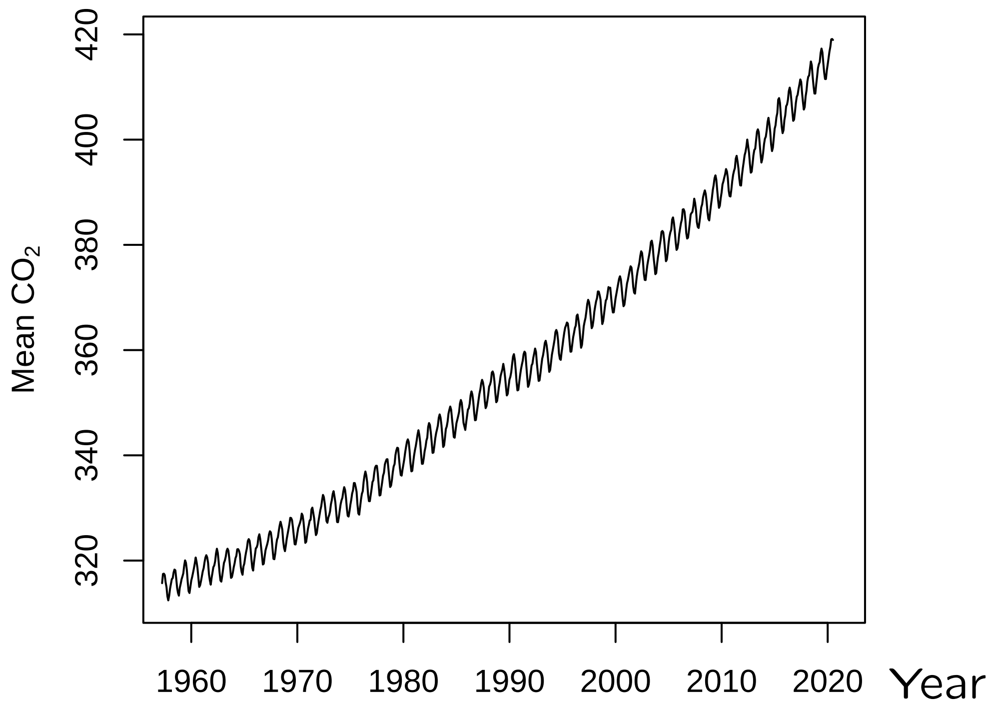
Application: Tests of i. i. d.-null based on previous statistics and corresponding asymptotics (distribution-free!).

Allowing for delays $d = 1, 2, \ldots$,
$H_{\text{ex}}\big(\widehat{\boldsymbol{p}}^{(d)}\big)$, $\widehat{\beta}^{(d)}$, $\widehat{\tau}^{(d)}$, etc. are counterparts to autocorrelation function (ACF) $\widehat{\rho}(h)$ with lags $h = 1, 2, \ldots$

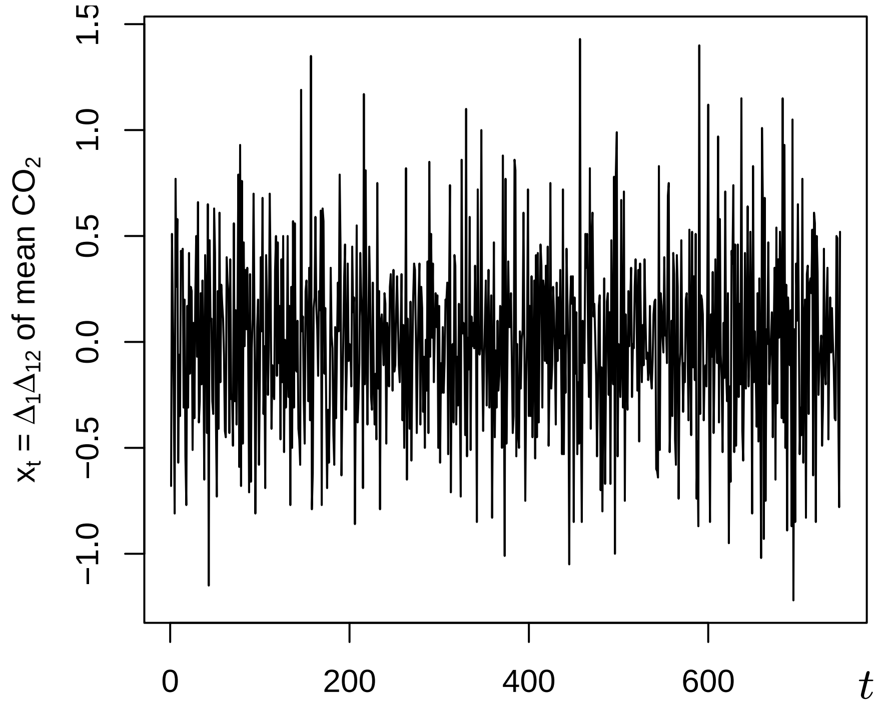
Simulation study in Weiß (2022),
ACF superior for linear dependence (ARMA processes), but OP-tests often superior for non-linear dependence.
OP-tests also robust against outliers (ranks!).
$\widehat{\tau}^{(d)}$ excels if also ACF reasonable,
$H_{\text{ex}}\big(\widehat{\boldsymbol{p}}^{(d)}\big)$ quite universally applicable.

Data example: atmospheric CO_2 at Mauna Loa Observatory on Hawaii (monthly means; mole fractions in ppm).

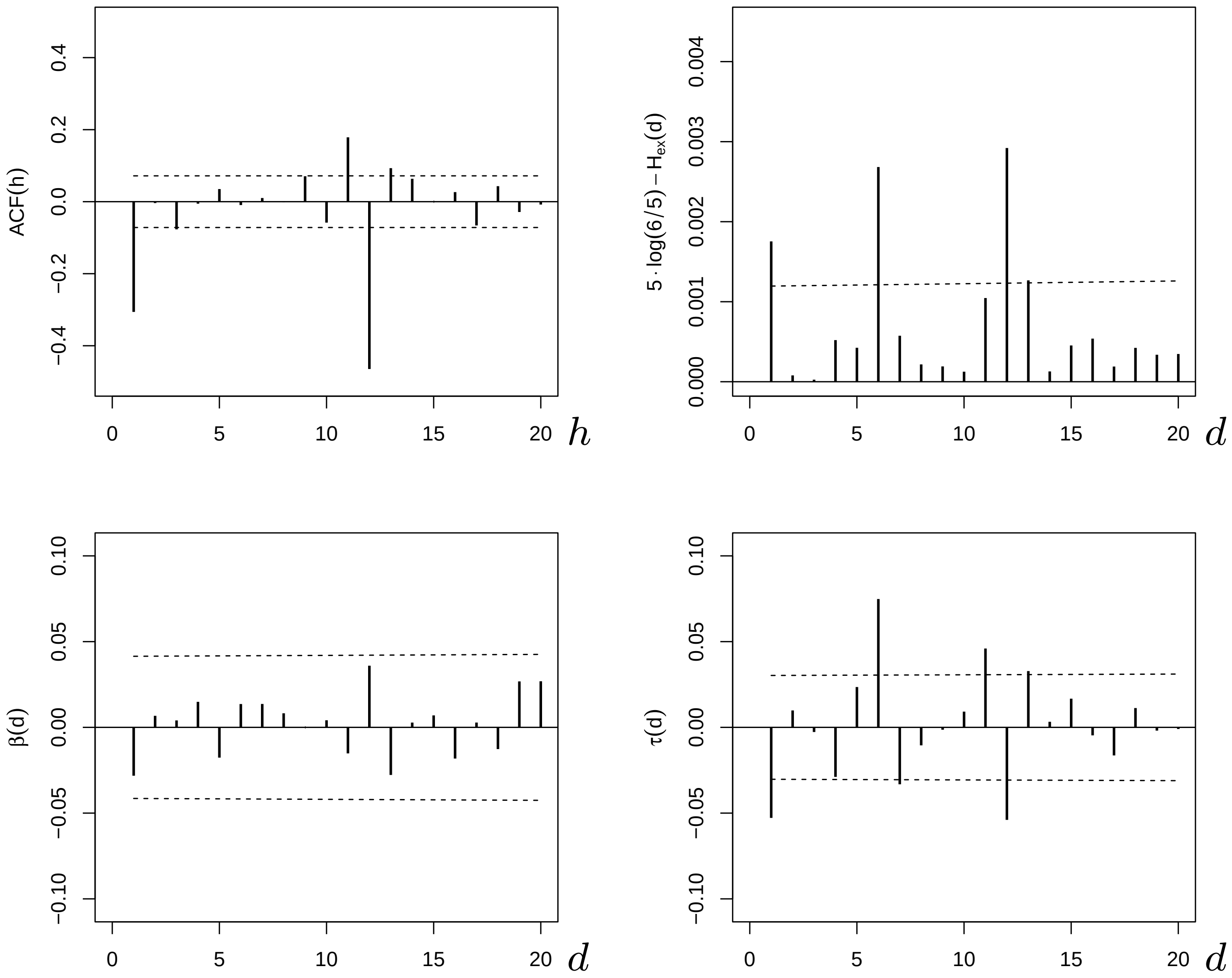
Original data:



Differenced data:



Dependence measures applied to differenced data:



Half-year dependence not recognized by ACF,

but detected by $H_{\text{ex}}\big(\widehat{\boldsymbol{p}}^{(d)}\big)$ and $\widehat{\tau}^{(d)}$.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Testing for Serial Dependence in Univariate Discrete-valued Time Series

(jointly with A. Schnurr)

Challenges & Approaches

Previous OP-tests depend critically on assumption of *continuously distributed* process (X_t), ensures distribution-free tests with unique null distribution.

In many applications, however, **discrete-valued processes** (see Weiß, 2018), say (Y_t).

If Y_t at least ordinal scale (or even quantitative, e. g., count process), then OPs could still be computed as before.

But frequent ties, π_t not uniformly distributed anymore.

In fact, see below, vector $\boldsymbol{p}$ strongly depends on actual distribution of X_t (and its parametrization).

So how use **OPs for discrete-valued processes?**

First solution: add continuously distributed noise to Y_t.

For example, if Y_t integer range from $\mathbb{Z}$,

then Weiß & Schnurr (2023) suggest **uniform noise** (U_t):

(U_t) i. i. d. uniform on $(0;1)$ and independent of (Y_t),

compute OPs from (X_t) defined as $X_t = Y_t + U_t$.

Then, $Y_s < Y_t$ implies $X_s < X_t$, i. e.,

strict orders in (Y_t) preserved, only ties randomly removed.

(X_t) continuously distributed, previous methods applicable.

However, ties in discrete-valued (Y_t) contain valuable

information about serial dependence, lost after adding noise.

Second solution: Like in Bian et al. (2012), Unakafova & Keller (2013), Schnurr & Fischer (2022), consider **generalized OPs** (GOPs) computed from (Y_t) directly.
GOPs complement above "strict patterns" by "tied patterns" (i. e., having at least one duplicate rank):

$(y_1, \ldots, y_m)$ has GOP $(i_1, \ldots, i_m) \in C_m$ if

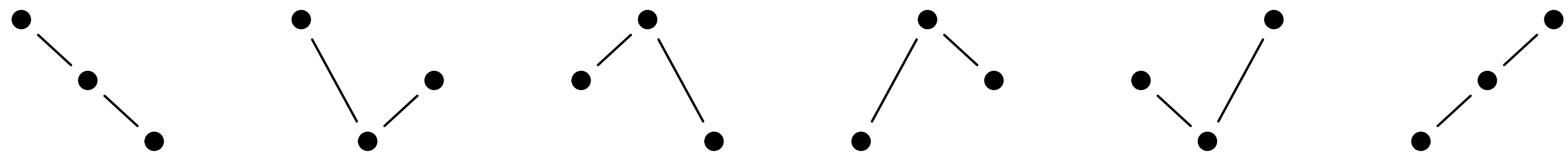
$$i_k < i_l \;\Leftrightarrow\; y_k < y_l \qquad \text{and} \qquad i_k = i_l \;\Leftrightarrow\; y_k = y_l,$$

for all $k, l \in \{1, \ldots, m\}$. Here, C_m set of mth-order Cayley permutations, its cardinality is mth ordered Bell number (Fubini number), $b(m)$, see Unakafova & Keller (2013).

Examples: $b(2)=3$ and $C_2=\big\{(2,1),(1,2)\big\}\cup\big\{(1,1)\big\}$.

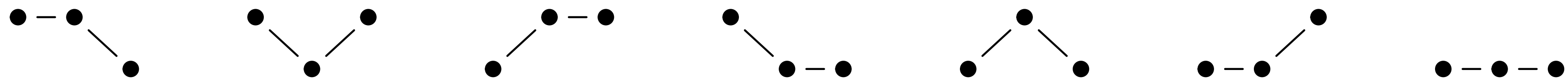
For $m=3$, $b(3)=13$ GOPs by complementing

$(3,2,1)$ $(3,1,2)$ $(2,3,1)$ $(1,3,2)$ $(2,1,3)$ $(1,2,3)$



by

$(2,2,1)$ $(2,1,2)$ $(1,2,2)$ $(2,1,1)$ $(1,2,1)$ $(1,1,2)$ $(1,1,1)$



$\Rightarrow$ GOPs make use of information contained in ties.

Remark: For $m = 3$, $b(3) = 13$ GOPs,

so in short time series, maybe some GOPs never observed.

$\Rightarrow$ Weiß & Schnurr (2023) suggest to form groups of GOPs:

Group 1 (increasing GOPs): $G_1 = \{(1,2,3),(1,2,2),(1,1,2)\}$;

Group 2 (decreasing GOPs): $G_2 = \{(3,2,1),(2,2,1),(2,1,1)\}$;

Group 3 (non-monotone GOPs): $G_3 = C_3 \setminus (G_1 \cup G_2)$.

Test statistics based on $\hat{\mathbf{p}}, \mathbf{p}_0$

for either full set of GOPs or grouped GOPs.

Weiß & Schnurr (2023) use different distances $d(\hat{\mathbf{p}}, \mathbf{p}_0)$.

Distribution of GOPs: Let $P\big(\pi_t = \pi\big) = \mathrm{p}(\pi)$,

denote PMF $p(y) = P(Y_t = y)$ and CDF $f(y) = P(Y_t \leq y)$.

Proposition: Let $(Y_t)_{\mathbb{Z}}$ be i. i. d., let $m = 2$, then

$$\begin{aligned} \mathrm{p}(1,1) &= \sum_y p(y)^2 = E\big[p(Y)\big], \\ \mathrm{p}(1,2) &= \mathrm{p}(2,1) = \tfrac{1}{2}\big(1 - \mathrm{p}(1,1)\big) \\ &= \sum_y p(y)\,\big(1 - f(y)\big) = E\big[1 - f(Y)\big]. \end{aligned}$$

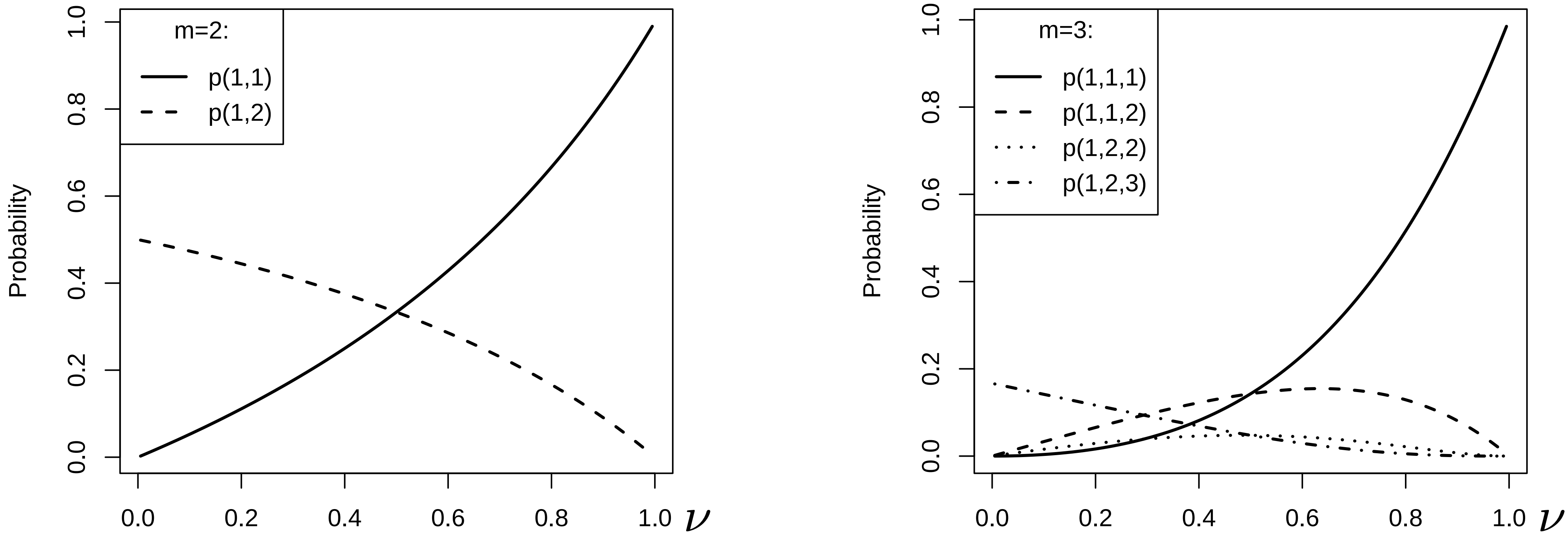
Proposition: Let $(Y_t)_{\mathbb{Z}}$ be i. i. d., let $m = 3$, then

$$\begin{aligned}
\mathrm{p}(1,1,1) &= \sum_y p(y)^3 = E\big[p(Y)^2\big],\\
\mathrm{p}(1,2,2) &= \mathrm{p}(2,1,2) = \mathrm{p}(2,2,1)\\
&= \sum_y f(y-1)\,p(y)^2 = E\big[f(Y-1)\,p(Y)\big],\\
\mathrm{p}(1,1,2) &= \mathrm{p}(1,2,1) = \mathrm{p}(2,1,1)\\
&= \sum_y p(y)^2\,\big(1-f(y)\big) = E\big[p(Y)\,\big(1-f(Y)\big)\big],
\end{aligned}$$

and all strict patterns have unique probability

$$\mathrm{p}(1,2,3) = \sum_y f(y-1)\,p(y)\,\big(1-f(y)\big) = E\big[f(Y-1)\,\big(1-f(Y)\big)\big].$$

Example: Probabilities of GOPs for i. i. d. Geom(ν)-counts:



$\Rightarrow$ GOP distribution depends on actual distribution of Y_t, so GOP-based test statistics of parametric nature.

Asymptotics for $m = 2$: see Weiß & Schnurr (2023).

We circumvent parametric nature by

Scheme for bootstrap implementation:

Let $Y_1, \ldots, Y_n$ be a stationary discrete-valued time series.

1. (a) Compute sample pmf from $Y_1, \ldots, Y_n$, compute corresponding null distribution $\mathbf{p}_0$ by Propositions.
   (b) Compute estimator $\hat{\mathbf{p}}$ from GOPs derived from $Y_1, \ldots, Y_n$.
   (c) Test statistic: distance $d(\hat{\mathbf{p}}, \mathbf{p}_0)$.

2. (...)

Scheme for bootstrap implementation:

Let $Y_1, \ldots, Y_n$ be a stationary discrete-valued time series.

1. (. . .)
2. Apply Efron bootstrap to $Y_1, \ldots, Y_n$:
   (a) Generate B i. i. d. time series $Y^*_{b,1}, \ldots, Y^*_{b,n}$, $b = 1, \ldots, B$.
   (b) Compute $\hat{\mathbf{p}}^*_b$ from GOPs from $Y^*_{b,1}, \ldots, Y^*_{b,n}$, $b = 1, \ldots, B$.
   (c) Test statistics (distances) $d(\hat{\mathbf{p}}^*_1, \mathbf{p}_0), \ldots, d(\hat{\mathbf{p}}^*_B, \mathbf{p}_0)$.

   $\Rightarrow$ approximate sample distribution of $d(\hat{\mathbf{p}}, \mathbf{p}_0)$ *under null*.

3. Compute $(1 - \alpha)$-quantile of $d(\hat{\mathbf{p}}^*_1, \mathbf{p}_0), \ldots, d(\hat{\mathbf{p}}^*_B, \mathbf{p}_0)$,

   use as critical value for $d(\hat{\mathbf{p}}, \mathbf{p}_0)$ to test i. i. d.-null.

Simulation study in Weiß & Schnurr (2023):

- GOP-tests hold level quite accurately (grouped GOPs) or tend to be conservative (all GOPs).
- Much better power than if using noisy OPs.
- Better power than ACF for nonlinear DGPs, if contamination by additive outliers, or if remarkable sample paths (exhibiting zero inflation, long runs of counts, or cascades of decaying counts).

Weiß & Schnurr (2023): application to **hydrologigal data** and **infectious disease data**.

Conclusions & Outlook

(G)OPs are well-interpretable, robust, and flexibly adapted to different types of dependence.
If data continuously distributed, we get non-parametric tests.

Work in progress:

In Weiß & Testik (2023), sequential testing of independence in continuously distributed process (X_t) via non-parametric OP-based EWMA control charts.
Currently: monitoring of *discrete-valued* processes, parametric EWMA control charts based on GOPs.

Thank You

for Your Interest!

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

Literature

Weiß (2022) Non-parametric tests for serial dependence in time series based on asymptotic . . . *Chaos* 32(9), 093107.

Weiß & Schnurr (2023) Generalized ordinal patterns in discrete-valued time series: . . . *Journal of Nonparametric Statistics*, in press.

Bandt (2019) Small order patterns . . . *Entropy* **21**, 613.

Bandt & Pompe (2002) Permutation entropy . . . *Phys Rev L* **88**, 174102.

Berger et al. (2019) Teaching ordinal patterns . . . *Entropy* **21**, 1023.

Bian et al. (2012) Modified permutation . . . *Phys Rev E* **85**, 021906.

Elsinger (2010) Independence tests . . . *Working paper* **165**, Öst. Nat.bank.

Schnurr & Fischer (2022) Generalized OPs . . . *CSDA* **171**, 107472.

Unakafova & Keller (2013) Efficiently meas. . . *Entropy* **15**, 4392–4415.

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

Weiß & Testik (2023) Non-param. control ch. . . *Technomet* **65**, 340–350.