Distance-based Analysis of Ordinal Time Series

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

Distance-based Analysis of Ordinal Time Series

Motivation & Outline

We analyze **ordinal time series** $X_1, \ldots, X_n$ having ordered categorical range $\mathcal{S} = \{s_0, \ldots, s_m\}$ with $s_0 < \ldots < s_m$.

Ordinal r.v. X closely related to **rank-count** variable I with bounded range $\{0, \ldots, m\}$ defined by $X = s_I$.

Count process $(I_t)_\mathbb{Z}$ generates ordinal $(X_t)_\mathbb{Z}$ by $X_t = s_{I_t}$.

$(X_t)_\mathbb{Z}$ inherits distributional properties from $(I_t)_\mathbb{Z}$ like stationarity, mixing, etc. (Weiß, 2018). In particular,

$$P(X_t = s_k) = P(I_t = k),\; P(X_t = s_k, X_{t-h} = s_l) = P(I_t = k, I_{t-h} = l).$$

Rank-count duality related to **latent-variable appr.** (Agresti, 2010), but neither continuous-valued nor unobservable.

Here: **distance functions** on ordinal range $\mathcal{S}$, measures of marginal or serial properties based on these distances:

- unified approach with obvious interpretation,
- great flexibility with respect to tailor-made distances,
- universal applicability, also beyond ordinal categories.

Distance-based approaches common in statistical analysis, e. g., Székely et al. (2007): serial dependence in real-valued processes as distances between characteristic functions, or Kvålseth (2011): ordinal variation via distance between CDFs.

Present work: distances defined on $\mathcal{S}$ itself, not between CDFs or other stochastic properties.

Outline:

- ordinal distances and their properties;
- distance-based analytic tools for marginal properties of ordinal t.s. like location, dispersion, or skewness;
- serial dependence of ordinal time series;
- sample versions and their asymptotics;
- application to time series of monthly credit ratings.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Ordinal Distances

Definition & Properties

Definition:

$d : \mathcal{S} \times \mathcal{S} \to [0; \infty)$ **distance measure** on $\mathcal{S}$ if for all $x, y \in \mathcal{S}$:

(D1) If $x = y$, then $d(x, y) = 0$. (positive semi-definite)

(D2) $d(x, y) = d(y, x)$. (symmetry)

Distance d is **pseudo-metric** if it even satisfies

(D3) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in \mathcal{S}$.

(triangular inequality)

Pseudo-metric d is **metric** if it even satisfies

(D1') $x = y$ iff $d(x, y) = 0$ for all $x, y \in \mathcal{S}$. (positive definite)

Possible Properties of Ordinal Distances:

Let $d: \mathcal{S}\times\mathcal{S} \to [0;\infty)$ be distance measure on *ordinal* range $\mathcal{S} = \{s_0,\ldots,s_m\}$. It may satisfy:

(O1) $d(s_0, s_m) \ = \ \max_{x,y\in\mathcal{S}} d(x,y)$. (maximization)

(O2) d is compatible with the ordering if

$$x<y<z \qquad \text{implies that} \qquad d(x,z) \ > \ d(x,y),\, d(y,z).$$

(O3) d is additive if for given $d_1,\ldots,d_m>0$, it holds that

$$d(s_i, s_{i+k}) = d_{i+1} + \ldots + d_{i+k} \quad \forall_{i=0,\ldots,m-1;\ k=1,\ldots,m-i}.$$

(O4) $d(s_i, s_j) \ = \ d(s_{m-i}, s_{m-j}) \quad \forall_{0\le i<j\le m}$. (centrosymmetry)

We have (O3) $\Rightarrow$ (O2) $\Rightarrow$ (O1).

(O4) relevant for possible symmetry of ordinal distribution.

Examples:

Hamming distance $d_{\mathsf{H}}(s_k, s_l) \;:=\; 1 - \delta_{k,l}$

metric with (O1), (O4), but not (O2).

Ordinal block distance $d_{\mathsf{o},1}(s_k, s_l) \;:=\; |k-l|$,

so number of categories between s_k and s_l; all properties above.

Squared Euclidean distance $d_{\mathsf{o},2}(s_k, s_l) \;:=\; (k-l)^2$

(quadratic penalty) no metric (not (D3)), not (O3).

Tailor-made distances possible!

Choice of weights for dissimilarities depends on application.

Successive ordinal states not necessarily "equally-spaced":

- At HSU, grades are $1, 1.3, \ldots, 3.7, 4$ (pass), and $4.3, 5$ (fail). $d(4, 4.3)$ largest among successive differences?!
- Taleb & Limam (2002): quality of porcelain products, "standard" (S), "second choice" (SC), "third choice" (TC) or "chipped" (C). $d(\text{C}, \text{TC}) > d(\text{TC}, \text{SC})$?!

HELMUT SCHMIDT
UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Distance-based Analysis of Ordinal Random Variables

Approaches & Examples

Proposal: quantify properties of ordinal r.v.

(location, dispersion, skewness) by **expected distances**.

Since distance between r.v. non-negative real number,

expectation thereof well defined.

Universal and flexible approach,

covers already existing measures of marginal properties.

Notations:

- $\boldsymbol{p} = (p_0, \ldots, p_m)^\top$, where $p_i = P(X = s_i)$.

  So $\boldsymbol{p}$ summarizes PMF.

- CDF: $f_k = \sum_{l=0}^{k} p_l = P(X \leq s_k)$ and $\boldsymbol{f} = (f_0, \ldots, f_{m-1})^\top$.

Location:

Location as $x_{\text{loc},d} \;:=\; \arg\min_{x\in\mathcal{S}} E\big[d(X,x)\big]$.

Then $x_{\text{loc},d} \in \mathcal{S}$.

Hamming distance $\mapsto$ mode(s),

block distance $\mapsto$ median.

Note that $d_{\text{o},2} \;\not\mapsto$ mean $E[I]$, because

$E\big[d_{\text{o},2}(X,s_i)\big] \;=\; E\big[(I-i)^2\big]$ minimized for $i \in \{0,\ldots,m\}$.

So $d_{\text{o},2}$ leads to $s_{\text{round}(E[I])}$.

Dispersion (variability) of X as expected distance between independent copies X_1, X_2 of X, i. e., $E\big[d(X_1, X_2)\big]$
→ **diversity coefficient (DIVC)** by Rao (1982). So

$$\mathrm{disp}_d(X) \;:=\; E\big[d(X_1, X_2)\big] \;=\; \textstyle\sum_{i,j=0}^{m} d(s_i, s_j)\, p_i\, p_j \;=\; \boldsymbol{p}^\top \mathbf{D}\, \boldsymbol{p},$$

If d positive definite, then $\mathrm{disp}_d = 0$ iff one-point distribution.
If property (O1) holds, $\mathrm{disp}_d \leq d(s_0, s_m)$.

Extreme scenarios for ordinal data:
minimal for one-point distribution $\boldsymbol{p}_{\mathrm{one}}$ (maximal consensus),
maximal for extreme two-point d. $\boldsymbol{p}_{\mathrm{two}} = (0.5, 0, \ldots, 0, 0.5)^\top$
(polarized distribution, maximal dissent), see Kiesl (2003).

Examples:

- Block distance leads to **Gini's mean difference**.

  In ordinal case,

$$\text{disp}_{d_{o,1}} \;=\; \textstyle\sum_{k,l=0}^{m} |k-l|\, p_k p_l \;=\; 2\,\textstyle\sum_{k=0}^{m-1} f_k(1-f_k).$$

  To be normalized by factor $2/d_{o,1}(s_0, s_m) = 2/m$, then **index of ordinal variation** (Kvålseth, 1995; Kiesl, 2003).

- Squared Euclidean distance leads to **variance**:

$$\text{disp}_{d_{o,2}} \;=\; E\big[(I_1 - I_2)^2\big] \;=\; 2\,V[I],$$

  normalized version given by $\frac{2}{m^2}\,\text{disp}_{d_{o,2}} = \frac{4}{m^2}\,V[I]$.

Symmetry (brief sketch!)

If X ordinal r.v. with $P(X = s_i) = p_i$,

then X^{r} **reflected copy** of X if $P(X = s_i) = p_{m-i}$.

Symmetry reasonably defined only if d **centrosymmetric**.

Symmetric if X and X^{r} are identically distributed, i. e.,

if $p_i = p_{m-i}$ or $f_i = 1 - f_{m-1-i}$ for all $i = 0, \ldots, m$.

$$\mathrm{asym}_d(X) \;=\; E\big[d(X, X^{\mathrm{r}})\big] - \mathrm{disp}_d(X) \;=\; \textstyle\sum_{i,j=0}^{m} d(s_i, s_j)\, p_i (p_{m-j} - p_j).$$

Example:

$$\mathrm{asym}_{d_{o,1}} \;=\; \textstyle\sum_{j=0}^{m-1} (1 - f_j - f_{m-j-1})^2.$$

Also more refined distinction possible,
skewness (Klein & Doll, 2018).

Maximal positive (negative) skewness iff $p_0 = 1$ ($p_m = 1$).

Signed **skewness measure:**

$$\text{skew}_d(X) \;=\; E\big[d(X, s_m)\big] - E\big[d(X, s_0)\big].$$

Example:

$$\text{skew}_{d_{o,1}} \;=\; 2\sum_{i=0}^{m-1} f_i \;-m.$$

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Distance-based Analysis of Ordinal Time Series

Serial Dependence

Now $x_1,\ldots,x_n$ time series from stationary ordinal process $(X_t)_{\mathbb{Z}}$.

DGP assumed to satisfy appropriate mixing condition

such that application of CLT possible,

e. g., α-mixing with exponentially decreasing weights.

We use bivariate probabilities (for $i,j=0,\ldots,m-1$)

$$f_{ij}^{(h)} \;=\; P(X_t\le s_i, X_{t-h}\le s_j) \;=\; P(I_t\le i, I_{t-h}\le j)$$

and corresponding sample frequencies $\widehat{f}_{ij}{}^{(h)}$.

We assume that $(X_t)_{\mathbb{Z}}$ some kind of "short memory":

$\sum_{h=1}^{\infty} \left|f_{ij}^{(h)} - f_i\,f_j\right| < \infty$ has to hold for all $i,j=0,\ldots,m-1$.

Measure of **serial dependence** at lag h:

$$\kappa_d(h) \;=\; \frac{\mathrm{disp}_d(X) - E\big[d(X_t, X_{t-h})\big]}{\mathrm{disp}_d(X)} \qquad \text{for } h \geq 1.$$

Type of **weighted Cohen's** κ (Cohen, 1968).

$\mathrm{disp}_d(X)$ as $E\big[d(X_t, X_{t-h})\big]$ under null of independence, i. e.,

$\kappa_d(h)$ relative deviation of DIVCs for dependent and indep. r.v.

Examples:

$$\kappa_{d_{\mathrm{H}}}(h) \;=\; \frac{\sum_{i=0}^{m}\, p_{ii}^{(h)} - s_2(\boldsymbol{p})}{1 - s_2(\boldsymbol{p})},$$

so serial Cohen's $\kappa(h)$ for nominal t.s. (Weiß & Göb, 2008).

Examples:

$\kappa_{d_{o,2}}(h) \;=\; \rho_I(h)$, i. e., lag-$h$ ACF of rank counts $(I_t)_{\mathbb{Z}}$.

Particularly relevant is $d = d_{o,1}$:

$$\kappa_{d_{o,1}}(h) \;=\; \frac{\sum_{i=0}^{m-1}\big(f_{ii}^{(h)} - f_i^2\big)}{\sum_{i=0}^{m-1} f_i(1-f_i)} \quad\in\; \Big[\frac{-\sum_{i=0}^{m-1} f_i^2}{\sum_{i=0}^{m-1} f_i(1-f_i)};\, 1\Big].$$

Strongest negative dependence if all ${f_{ii}}^{(h)} = 0$,

i. e., $X_{t-h} \leq s_i$ necessarily followed by $X_t > s_i$.

Maximal positive dependence

if $X_{t-h} \leq s_i$ necessarily followed by $X_t \leq s_i$.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Distance-based Analysis of Ordinal Time Series

Sample measures

Sample versions easily obtained by replacing

expected distances by sample means of distances, e. g.,

sample version of $x_{\mathrm{loc},d}$ as $\quad \widehat{x}_{\mathrm{loc},d} := \arg\min_x \frac{1}{n} \sum_{t=1}^{n} d(x_t, x)$.

DIVC-related measures lead to $\quad \widehat{\mathrm{disp}}_d \; = \; \widehat{\boldsymbol{p}}^\top \, \mathbf{D} \, \widehat{\boldsymbol{p}}$,

$\widehat{\mathrm{asym}}_d \; = \; \widehat{\boldsymbol{p}}^\top \, (\mathbf{J} - \mathbf{I}) \mathbf{D} \, \widehat{\boldsymbol{p}}, \quad \widehat{\mathrm{skew}}_d \; = \; \sum_{i=0}^{m} \Big(d(s_i, s_m) - d(s_i, s_0) \Big) \, \widehat{p}_i.$

For asymptotics, we focus on ordinal block distance $d_{\mathrm{o},1}$:

$$\widehat{\mathrm{disp}}_{d_{\mathrm{o},1}} \; = \; 2 \sum_{i=0}^{m-1} \widehat{f}_i (1 - \widehat{f}_i),$$

$$\widehat{\mathrm{asym}}_{d_{\mathrm{o},1}} \; = \; \sum_{j=0}^{m-1} (1 - \widehat{f}_j - \widehat{f}_{m-j-1})^2,$$

$$\widehat{\mathrm{skew}}_{d_{\mathrm{o},1}} \; = \; 2 \sum_{i=0}^{m-1} \widehat{f}_i \; - m, \quad \widehat{\kappa}_{d_{\mathrm{o},1}}(h) \; = \; \frac{\sum_{i=0}^{m-1} \left(\widehat{f}_{ii}^{(h)} - \widehat{f}_i^2 \right)}{\sum_{i=0}^{m-1} \widehat{f}_i (1 - \widehat{f}_i)}.$$

Theorem:

(i) If $(f_0, \ldots, f_{m-1}) \neq (\frac{1}{2}, \ldots, \frac{1}{2})$ (extreme two-point),

then $\widehat{\text{disp}}_{d_{\text{o},1}}$ is asymptotically normally distributed with mean

$$(1 - \tfrac{1}{n})\, \text{disp}_{d_{\text{o},1}} \; - \; \tfrac{4}{n} \textstyle\sum_{h=1}^{n-1}(1 - \tfrac{h}{n}) \textstyle\sum_{i=0}^{m-1} \big(f_{ii}^{(h)} - f_i^2\big)$$

and variance

$$\tfrac{4}{n} \textstyle\sum_{i,j=0}^{m-1}(1 - 2f_i)(1 - 2f_j) \big(f_{\min\{i,j\}} - f_i\, f_j\big)$$
$$+ \tfrac{8}{n} \textstyle\sum_{h=1}^{n-1}(1 - \tfrac{h}{n}) \textstyle\sum_{i,j=0}^{m-1}(1 - 2f_i)(1 - 2f_j) \big(f_{ij}^{(h)} - f_i\, f_j\big).$$

Theorem:

(ii) If distribution of X_i not symmetric, then $\widehat{\text{asym}}_{d_{o,1}}$ asymptotically normally distributed with mean

$$\text{asym}_{d_{o,1}} + \frac{1}{n}\,\text{disp}_{d_{o,1}} + \frac{2}{n}\sum_{j=0}^{m-1}\big(f_{\min\{m-j-1,j\}} - f_j\, f_{m-j-1}\big)$$
$$+ \frac{4}{n}\sum_{h=1}^{n-1}(1-\tfrac{h}{n})\sum_{j=0}^{m-1}\big(f_{jj}^{(h)} - f_j^2 + f_{j,m-j-1}^{(h)} - f_j\, f_{m-j-1}\big)$$

and variance

$$\frac{16}{n}\sum_{i,j=0}^{m-1}(1-f_i-f_{m-i-1})(1-f_j-f_{m-j-1})\big(f_{\min\{i,j\}} - f_i\, f_j\big) +$$
$$\frac{32}{n}\sum_{h=1}^{n-1}(1-\tfrac{h}{n})\sum_{i,j=0}^{m-1}(1-f_i-f_{m-i-1})(1-f_j-f_{m-j-1})\big(f_{ij}^{(h)} - f_i\, f_j\big).$$

Theorem:

(iii) $\widehat{\text{skew}}_{d_{\text{o},1}}$ asympt. normal with mean $\text{skew}_{d_{\text{o},1}}$ and variance

$$\frac{4}{n}\sum_{i,j=0}^{m-1}\big(f_{\min\{i,j\}}-f_i\,f_j\big)+\frac{8}{n}\sum_{h=1}^{n-1}\left(1-\frac{h}{n}\right)\sum_{i,j=0}^{m-1}\big(f_{ij}^{(h)}-f_i\,f_j\big).$$

Modifications necessary for **boundary cases**,

leads to asymptotic quadratic-form distributions.

Finite-sample performance confirmed with simulations.

Testing for significant dependence:

Let $X_1, \ldots, X_n$ be time series from i. i. d. ordinal process $(X_t)_{\mathbb{Z}}$, let $h \geq 1$.

Then asymptotic approximation for $\widehat{\kappa}_{d_{\mathrm{o},1}}(h)$ is normal distribution with mean $-1/n$ and variance

$$\frac{1}{n}\,\frac{4}{\mathrm{disp}^2_{d_{\mathrm{o},1}}}\,\sum_{k,l=0}^{m-1}\big(f_{\min\{k,l\}} - f_k\, f_l\big)^2.$$

Finite-sample performance confirmed with simulations.

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH STAT

Application: Credit Ratings

EU countries

S&P ratings for 28 EU countries, taken from Trading Economics platform `tradingeconomics.com`.

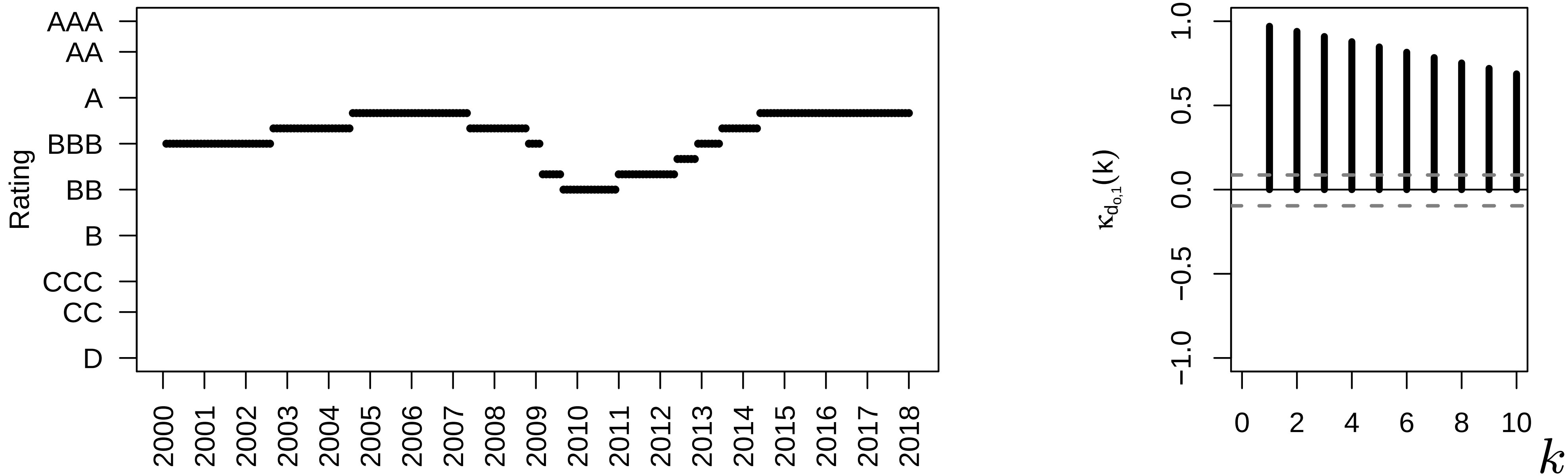
Possible outcomes from 'D' (worst) to 'AAA' (best), commonly refined to $m+1 = 23$ states $s_0, \ldots, s_{22}$ given by 'D', 'SD', 'R', 'CC', 'CCC−', 'CCC', 'CCC+', . . . , 'AA+', 'AAA'.

Motivated by Trading Economics rating, we treat $s_0, \ldots, s_{22}$ as equidistant, so $d_{o,1}$.

Unique time range for all EU countries: $n = 216$ months from January 2000 to December 2017.

Monthly S&P credit rating of Latvia:

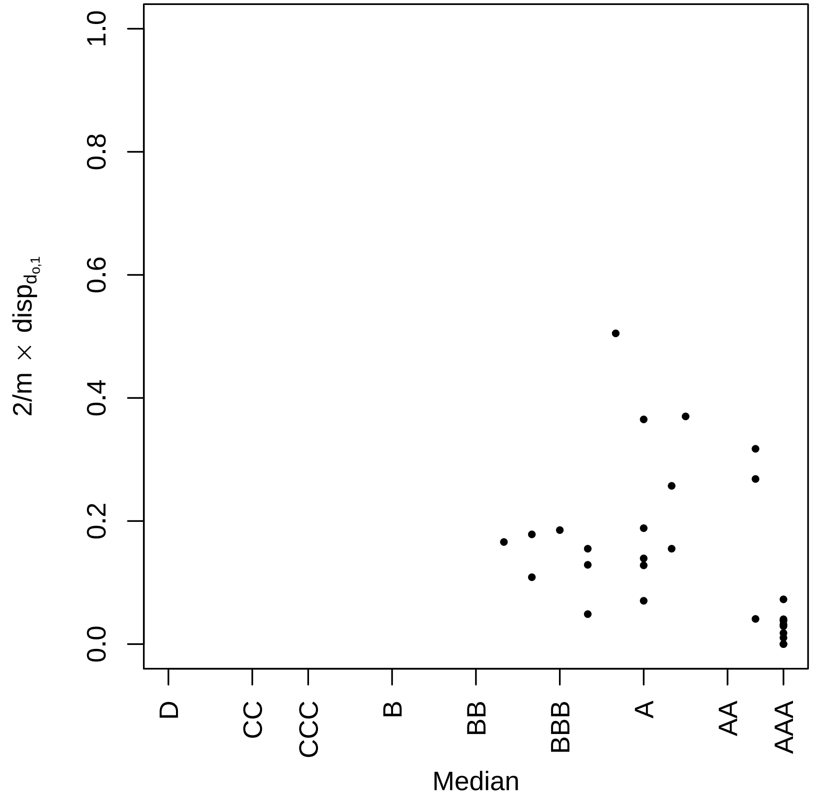
time series plot (left), $\widehat{\kappa}_{d_{\mathrm{O},1}}(k)$ (right).

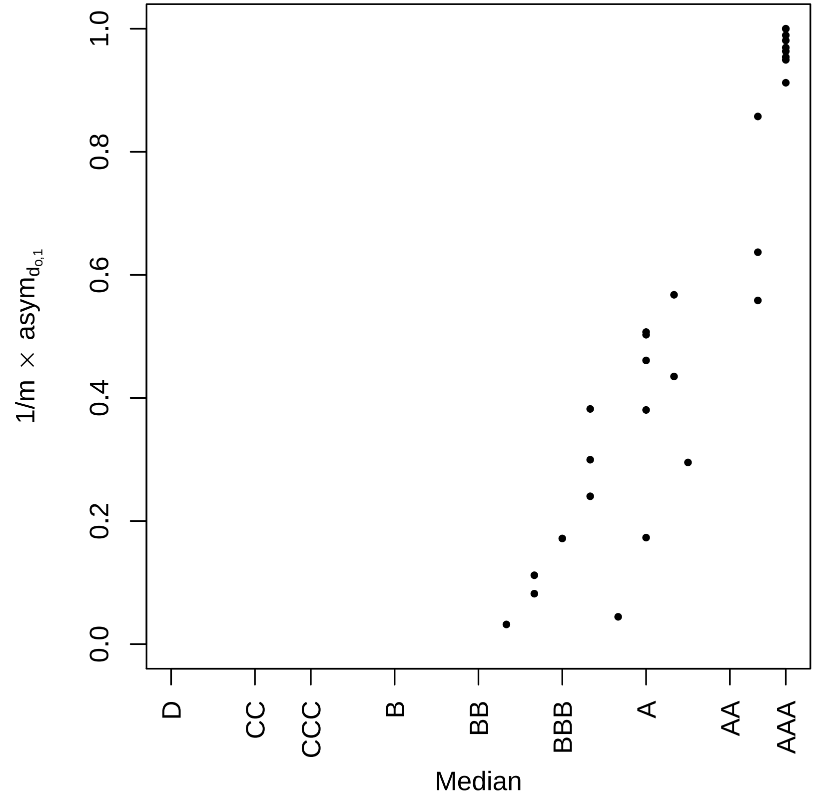


Norm. dispersion 0.155, asymmetry 0.240 and skewness −0.318.

Application: Credit Ratings

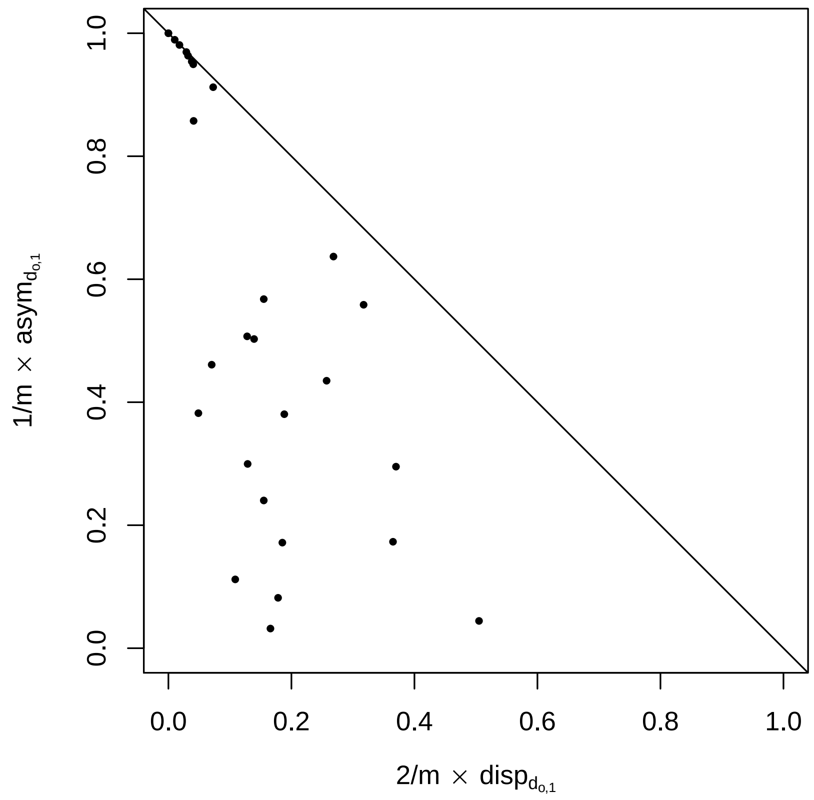
Marginal properties of S&P time series of 28 EU countries:

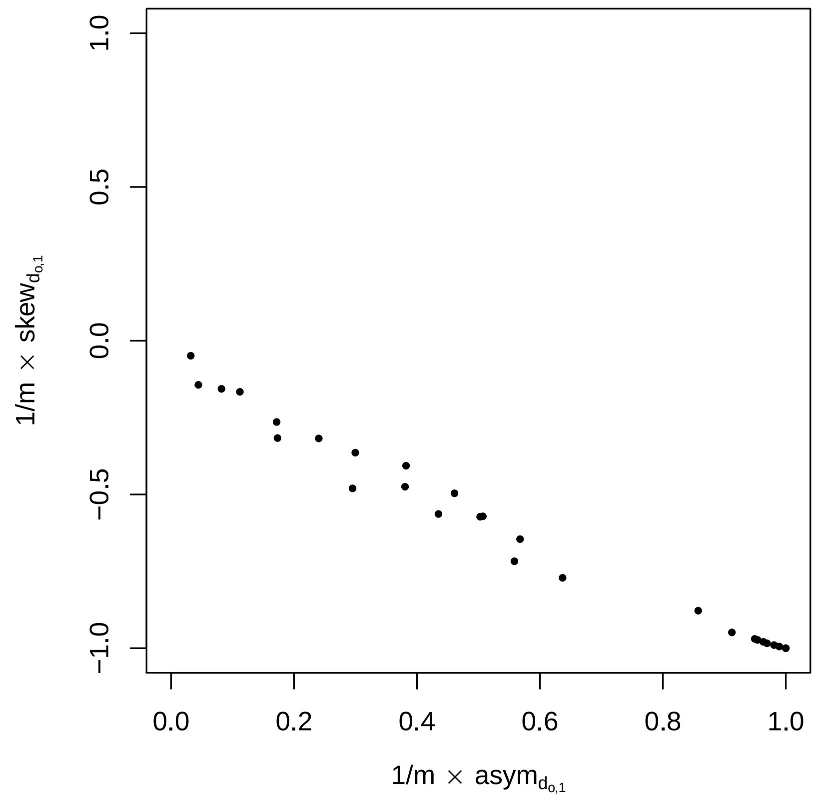




Zero dispersion for Germany and Luxembourg (s_{22} = AAA).

Marginal properties of S&P time series of 28 EU countries:





Inequality $\frac{1}{m}\,\text{asym}_{d_{\text{o},1}} \leq 1 - \text{IOV}$.

Would also be reasonable to consider reduced range with $\tilde{m}+1=10$ states $\tilde{s}_0,\ldots,\tilde{s}_9$ given by 'D', 'R', 'CC', 'CCC', 'B', 'BB', 'BBB', 'A', 'AA', 'AAA'.

To make results comparable, equidistance not appropriate.

Following (additive) distance matrix used for reduced range:

$$\mathbf{D} = \big(d(\tilde{s}_i,\tilde{s}_j)\big)_{i,j=0,\ldots,9} = \begin{pmatrix} 0 & 2 & 3 & 5 & 8 & 11 & 14 & 17 & 20 & 22 \\ 2 & 0 & 1 & 3 & 6 & 9 & 12 & 15 & 18 & 20 \\ 3 & 1 & 0 & 2 & 5 & 8 & 11 & 14 & 17 & 19 \\ 5 & 3 & 2 & 0 & 3 & 6 & 9 & 12 & 15 & 17 \\ 8 & 6 & 5 & 3 & 0 & 3 & 6 & 9 & 12 & 14 \\ 11 & 9 & 8 & 6 & 3 & 0 & 3 & 6 & 9 & 11 \\ 14 & 12 & 11 & 9 & 6 & 3 & 0 & 3 & 6 & 8 \\ 17 & 15 & 14 & 12 & 9 & 6 & 3 & 0 & 3 & 5 \\ 20 & 18 & 17 & 15 & 12 & 9 & 6 & 3 & 0 & 2 \\ 22 & 20 & 19 & 17 & 14 & 11 & 8 & 5 & 2 & 0 \end{pmatrix}.$$

Not centrosymmetric (e. g., $d(\tilde{s}_0,\tilde{s}_2)=3\neq 5=d(\tilde{s}_9,\tilde{s}_7)$).

Location, dispersion and serial dependence computed w.r.t. d:

$\widehat{x}_{\text{loc},d} = \arg\min_x \frac{1}{n} \sum_{t=1}^{n} d(x_t, x)$ instead of median $\widehat{x}_{\text{loc},d_{\text{o},1}}$,

similarly $\widehat{\text{disp}}_d = \widehat{\boldsymbol{p}}^\top \mathbf{D}\, \widehat{\boldsymbol{p}}$ (normalizing factor $2/d(\tilde{s}_0, \tilde{s}_9) = 1/11$)

and $\widehat{\kappa}_d(h) = \big(\widehat{\text{disp}}_d - \frac{1}{n-h} \sum_{t=h+1}^{n} d(x_t, x_{t-h})\big)\big/\widehat{\text{disp}}_d$.

Crucial question: how do location, dispersion and serial dependence change if using reduced range instead of full one? (...)

Location and dispersion robust against this change in range.

Serial dependence stronger affected, mainly larger values:

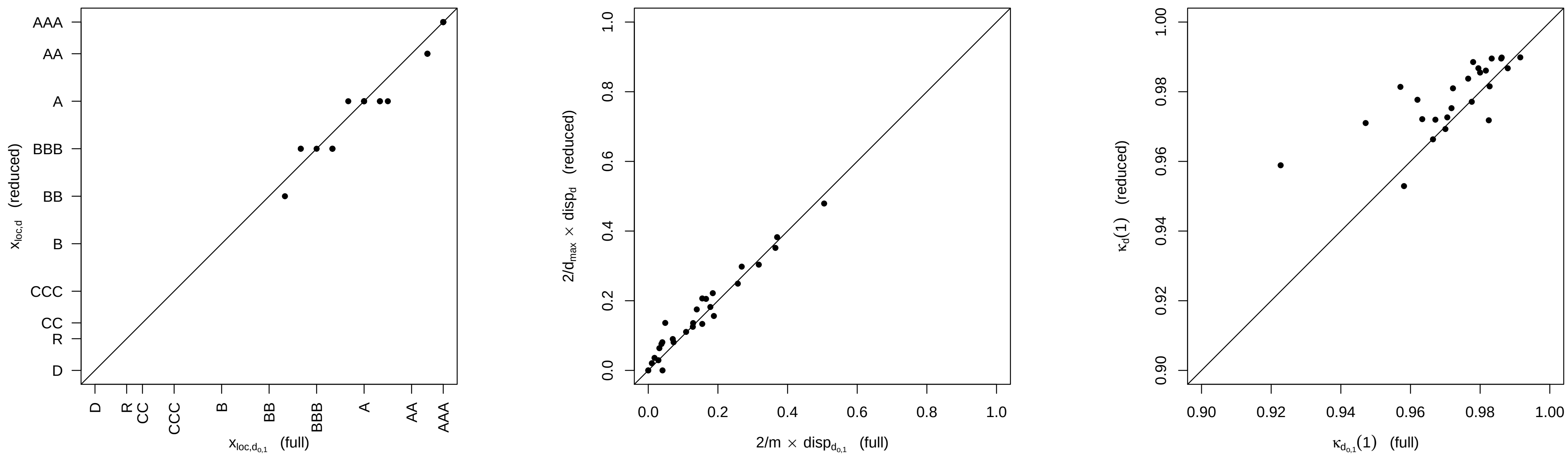
reduction of states causes less variation,

i. e., longer runs and thus increased positive dependence.

Application: Credit Ratings

MATH
STAT

Comparison of properties of S&P time series of 28 EU countries if using either full range or reduced range.



Not centrosymmetric, so dispersion of original t.s. and reflected copies thereof might differ. Latvia: 0.207 vs. 0.165.

Conclusions

- Location, dispersion, symmetry and serial dependence for ordinal time series based on expected distances.
  Well interpretable, can deal with many types of distance function, asymptotics available.

- Distance-based approach also for r.v. and processes having other types of range, e. g., compositional data.
  Thus allows for novel analytic tools for such data types, and offers opportunity to treat different data types with unique approach and unique interpretation.

Thank You

for Your Interest!

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

Literature

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

Cohen (1968) Weighted kappa: ... *Psychol. Bull.* **70**, 213–220.

Efron (1979) Bootstrap methods: another look at the jackknife. *Ann. Stat.* **7**, 1–26.

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

Klein & Doll (2018) Tests on asymmetry ... *FAU Discussion Papers* 03/2018.

Kvålseth (1995) Coefficients of variation ... *Perc. Motor Skills* **80**, 843–847.

Kvålseth (2011) Variation for categ... In *Int. Encyc. Stat. Science*, Springer, 1642–1645.

Lomnicki (1952) The standard error of Gini's mean diff... *Ann. Math. Stat.* **23**, 635–637.

Nair (1936) The standard error of Gini's mean difference. *Biometrika* **28**, 428–436.

Rao (1982) Diversity and dissimilarity coefficients ... *Theo. Popul. Biol.* **21**, 24–43.

Székely et al. (2007) Measuring and testing dependence ... *Ann. Stat.* **35**, 2769–2794.

Taleb & Limam (2002) On fuzzy and probabilistic control charts. *IJPR* **40**, 2849–2863.

Weiß (2013) Serial dependence of NDARMA processes. *CSDA* **68**, 213–238.

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

Weiß & Göb (2008) Measuring serial dependence ... *AStA* **92**, 71–89.