UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR MATHEMATIK

Theory and Practice of Permutation Entropy

Karsten Keller, Inga Stolz

Institute of Mathematics
University of Lübeck

4th International Electronic Conference on Entropy and Its Applications, 21 November - 1 December 2017

Contents

1. Motivation

- Entropies and entropy-like quantities are playing an increasing role in data analysis, in the contexts of dynamical systems and of stochastic processes.
- Various applications of entropies are given, for example, in the analysis of physiological time series.
- Central classical concepts: Approximate entropy, Sample entropy, and variants.
- Interesting relatively new concepts: **Permutation entropies** and variants based on the ordinal structure of time series and systems behind them.
- Important: **good theoretical understanding** of those concepts and their relationship to other entropies, and their **adequate application** to the analysis of data.

2. Context

Modeling (one observable, simplest setting)

time series $(x_t)_{t=0}^{N-1} = (x_0, x_1, x_2, \ldots, x_{N-1})$ assumed as values measured on the orbit of a dynamical system

- $(\Omega, \mathcal{A}, \mu, T)$ μ-**preserving dynamical system**
- X real-valued random variable on Ω
- **observable** X provides 'outreading' stochastic process $(X \circ T^{\circ t})_{t=0}^{\infty}$ with realizations $(X \circ T^{\circ t}(\omega))_{t=0}^{\infty}$; $\omega \in \Omega$

assumptions:

- **Ergodicity:** realizations represent distribution of X
- no information loss by the measuring process with respect to X ('separation' of orbits), natural[1]

---

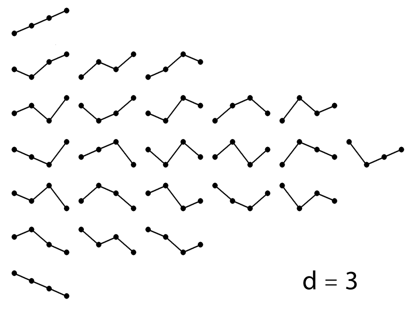
[1] Gutman, Takens embedding theorem with a continuous observable, in: Proceedings of the Erg. Th. Workshops UNC Chapel Hill 2013-2014. De Gruyter

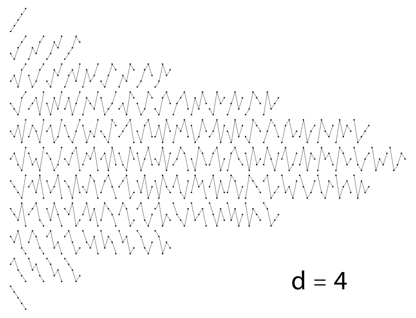
Ordinal patterns

Definition

$(x_0, x_1, \ldots, x_d) \in \mathbb{R}^{d+1}$ has the **(ordinal)** d-**pattern** $\pi = (r_0, r_1, \ldots, r_d) \in \Pi_d$ if

$$x_{r_0} \geq x_{r_1} \geq \ldots \geq x_{r_d}, \text{ and}$$
$$r_{l-1} > r_l \text{ in the case } x_{r_{l-1}} = x_{r_l}.$$





Relative frequencies / probabilities of ordinal d-pattern words

Definition

- **relative frequency of d-patt. word** $(\pi_1, \pi_2, \ldots, \pi_k) \in \Pi_d^k$ in time series $(x_t)_{t=0}^{N-1}$:

$$p_{(\pi_1,\pi_2,\ldots,\pi_k)} := \frac{1}{N-d-k+1} \#\{s \in \{0,\ldots,N-d-k\} \mid \text{for all } i = 1,\ldots,k \; (x_{s+i-1}, x_{s+i} \ldots, x_{s+i-1+d}) \text{ has ordinal pattern } \pi_i\}$$

- **probability of d-pattern word** $(\pi_1, \pi_2, \ldots, \pi_k) \in \Pi_d^k$ for $(\Omega, \mathcal{A}, \mu, T, X)$:

$$P_{(\pi_1,\pi_2,\ldots,\pi_k)} := \mu(\{\omega \in \Omega \mid \text{for all } i = 1,\ldots,k \; X \circ T^{\circ i-1}(\omega), X \circ T^{\circ i}(\omega), \ldots, X \circ T^{\circ d}(\omega) \text{ has ordinal pattern } \pi_i\})$$

- $p_{(\pi_1,\pi_2,\ldots,\pi_k)}$ estimates $P_{(\pi_1,\pi_2,\ldots,\pi_k)}$

3. From KS entropy to permutation entropy

Entropies (I)

Kolmogorov-Sinai entropy (**KS entropy**) is central theoretical complexity measure for dynamical systems, but not easy to determinate and to estimate.

ordinal approach:

- **entropy of** d**-pattern-**k**-words**:

$$H(d,k) := -\sum_{(\pi_1,\ldots,\pi_k)\in\Pi_d^k} P_{(\pi_1,\ldots,\pi_k)} \ln P_{(\pi_1,\ldots,\pi_k)}$$

- **empirical entropy of** d**-pattern-**k**-words**:

$$h(d,k) := -\sum_{(\pi_1,\ldots,\pi_k)\in\Pi_d^k} p_{(\pi_1,\ldots,\pi_k)} \ln p_{(\pi_1,\ldots,\pi_k)}$$

- $h(d,k)$ estimates $H(d,k)$
- $H(d,0) := 0$

Entropies (II)

- **entropy rate of** d**-patterns**:

$$\mathrm{EntroRate}(d) := \lim_{k\to\infty}(\downarrow)\frac{H(d,k)}{k} = \lim_{k\to\infty}(\downarrow)(H(d,k+1)-H(d,k))$$

Theorem

$$\text{KS entropy} = \lim(\uparrow)_{d\to\infty}\mathrm{EntroRate}(d) \leq \text{Permutation entropy}^{2}$$

[2] Antoniouk, K., Maksymenko, Discrete Contin. Dyn. Syst. A 34 (2014); K., Maksymenko,Stolz, Discrete Contin. Dyn. Syst. B 20 (2015);

- only one limit for **Permutation entropy**[3]:

$$\text{Permutation entropy} = \limsup_{d\to\infty}\frac{H(d,1)}{d} \qquad \left(\text{better: } \liminf_{d\to\infty}\frac{H(d,1)}{d}\right)$$

[3] Bandt, Pompe, Phys. Rev. Lett. 88 (2002), X identity

Entropies (III)

Remark

KS entropy = Permutation entropy for a special class of dynamical systems (piecewise monotone interval maps) and X identity[4]

[4] Bandt, G. Keller, Pompe, Nonlinearity 15 (2002)

General ordinal view on KS entropy

relevant entropies:

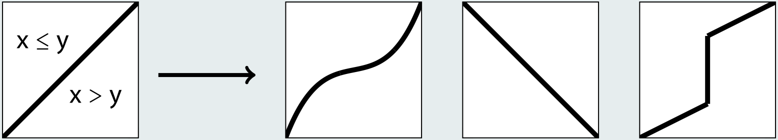
$$\begin{array}{cccccccc}
H(1,1) & H(1,2) & H(1,3) & H(1,4) & H(1,4) & H(1,5) & H(1,6) & \cdots \\
H(2,1) & H(2,2) & H(2,3) & H(2,4) & H(2,4) & H(2,5) & H(2,6) & \cdots \\
H(3,1) & H(3,2) & H(3,3) & H(3,4) & H(3,4) & H(3,5) & H(3,6) & \cdots \\
H(4,1) & H(4,2) & H(4,3) & H(4,4) & H(4,4) & H(4,5) & H(4,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & & &
\end{array}$$

right weighting:

$$\begin{array}{cccccccc}
H(1,1) & H(1,2)/2 & H(1,3)/3 & H(1,4)/4 & H(1,5)/5 & \cdots & \searrow & \mathrm{EntroRate}(1) \\
H(2,1) & H(2,2)/2 & H(2,3)/3 & H(2,4)/4 & H(2,5)/5 & \cdots & \searrow & \mathrm{EntroRate}(2) \\
H(3,1) & H(3,2)/2 & H(3,3)/3 & H(3,4)/4 & H(3,5)/5 & \cdots & \searrow & \mathrm{EntroRate}(3) \\
H(4,1) & H(4,2)/2 & H(4,3)/3 & H(4,4)/4 & H(4,5)/5 & \cdots & \searrow & \mathrm{EntroRate}(4) \\
 & & & & & & & \vdots \\
 & & & & & & & \nearrow \\
 & & & & & & & \text{KS entropy}
\end{array}$$

Generalizations

- generalization to finitely (infinitely) many observables $X_1, X_2, \ldots, X_n(, \ldots)$
- generalization to many non-ergodic cases (ergodic decomposition)
- generalization to new two-dimensionel basic symbolization schemes[5]



[5] Stolz, Keller 2017

4. Conditional viewpoint

‘Conditional’ ordinal view on KS entropy

$$\begin{array}{cccccl}
H(1,2)-H(1,1) & H(1,3)-H(1,2) & H(1,4)-H(1,3) & \cdots & \searrow & \text{EntroRate}(1)\\
H(2,2)-H(2,1) & H(2,3)-H(2,2) & H(2,4)-H(2,3) & \cdots & \searrow & \text{EntroRate}(2)\\
H(3,2)-H(3,1) & H(3,3)-H(3,2) & H(3,4)-H(3,3) & \cdots & \searrow & \text{EntroRate}(3)\\
H(4,2)-H(4,1) & H(4,3)-H(4,2) & H(4,4)-H(4,3) & \cdots & \searrow & \text{EntroRate}(4)\\
 & & & & & \vdots\\
 & & & & & \nearrow\\
 & & & & & \text{KS entropy}
\end{array}$$

- For all natural numbers k it holds

$$H(d,k)-H(d,k-1) \le \frac{1}{k}\sum_{i=1}^{k}\left(H(d,i)-H(d,i-1)\right) = \frac{H(d,k)}{k}$$

Lemma

For each sequence $(k_d)_{d=1}^{\infty}$ of natural numbers it holds

$$\text{KS entropy} \le \liminf_{d\to\infty}\left(H(d,k_d+1)-H(d,k_d)\right) \le \liminf_{d\to\infty}\frac{H(d,k_d)}{k_d}$$

Conditional entropy of ordinal patterns[6]

- special case: $k_d = 1$ for all $d \in \mathbb{N}$:

Definition

conditional entropy of ordinal patterns

$$\liminf_{d\to\infty} H(d,2) - H(d,1)$$

- interpretation: uncertainty of next type of ordinal pattern of high order given one such type
- also interesting:

$$\liminf_{d\to\infty} H(d,k+1) - H(d,k)$$

for $k = 2, 3, 4, \ldots$

[6] Unakafov, K., Physica D 269 (2014), 94-102.

Conditional entropy better than PE?

assuming existence of

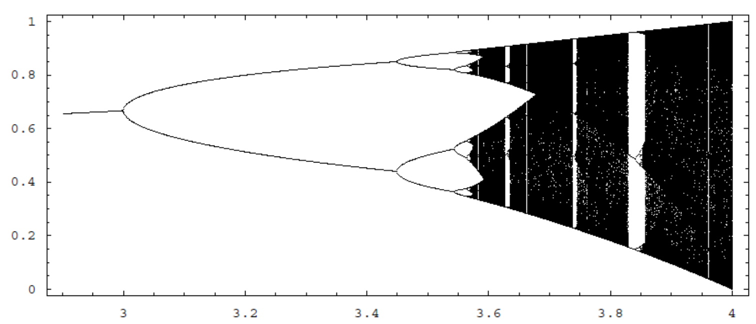
$$\lim_{d\to\infty}(H(d+1,1)-H(d,1)),$$

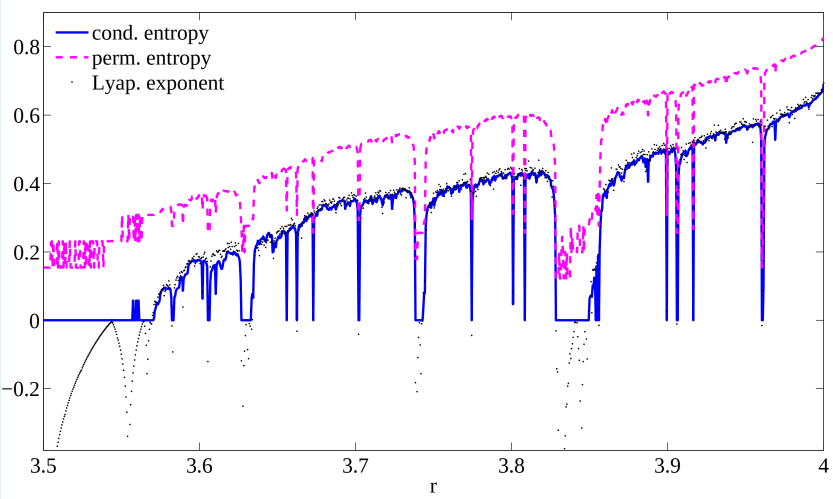
which looks natural, by the Stolz-Cesàro theorem with $H(d,0) := 0$ it holds

$$\begin{aligned}
\textbf{KS entropy} &\leq \liminf_{d\to\infty}(H(d,2)-H(d,1)) \\
&= \textbf{conditional entropy of ordinal patterns} \\
&\leq \lim_{d\to\infty}(H(d+1,1)-H(d,1)) \\
&= \lim_{d\to\infty}(H(d,1)-H(d-1,1)) \\
&= \lim_{d\to\infty}\frac{1}{d}\sum_{i=1}^{d}(H(i,1)-H(i-1,1)) \\
&= \liminf_{d\to\infty}\frac{1}{d}H(d,1) = \textbf{Permutation entropy}
\end{aligned}$$

Logistic family and permutation entropies

- $x \in [0,1] \mapsto rx(1-x)$ for different $r \in [0,4]$
- for almost all $r \in [0,4]$ the KS entropy either coincides with the Lyapunov exponent if it is positive or is equal to zero otherwise (Pesin's formula, 'natural' RB-measure)



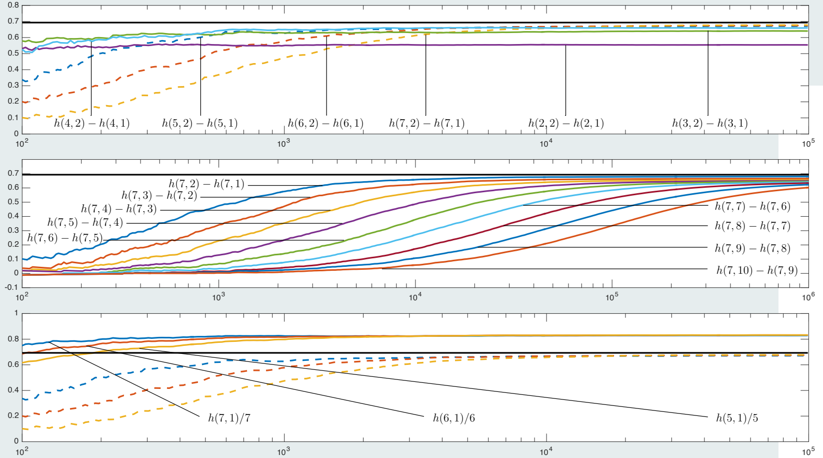


Lyapunov exponent, empirical and **conditional Permutation entropy** (d=9)

5. Practical aspects

Asymptotics versus statistics[7]

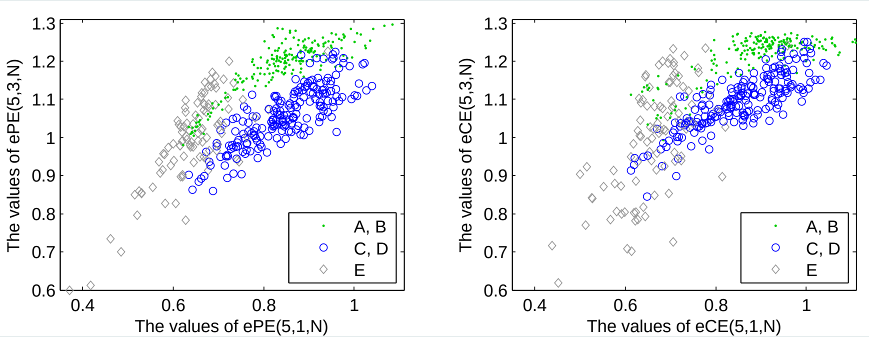
$\Omega = [0,1], T(\omega) = 4\omega(1-\omega), X = \text{id}$



7 Keller, Mangold, Stolz, Werner Entropy 19 (2017)

Combining entropies, delays

example: EEG recordings from the Bonn EEG Database[8]



empirical Permutation entropy versus empirical Permutation entropy/
conditional Permutation entropy versus conditional Permutation entropy,

also other entropies

8 available online: http://epileptologie-bonn.de

Variants of Permutation entropy

Reinventing metric information

- 'Flat' ordinal patterns, i.e. ordinal patterns coming from vectors of low variability, are very sensitive with respect to noise.

⇒ **robust Permutation entropy**[9] excluding 'flat' ordinal patterns

⇒ **weighted Permutation entropy**[10]: weighting ordinal patterns by variance of vectors behind

- theoretically not well understood

Generalization of conditional approach by Armand Eyebe Fouda, Koepf, Jacquir[11]

---

[9] Keller, Unakafov, Unakafova, Entropy 16 (2014)

[10] Fadlallah, Chen, Keil, Prìncipe, Phys. Rev. E 87 (2013)

[11] Communications in Nonlinear Sci. and Num. Simulat. 46 (2017), 103-115

Further aspects

- generally: compromise between theoretical requirements and practical possibilities necessary
- requires better theoretical understanding of measures
- results in standardizations which have to be communicated and commonly used
- standards have to be (automatically) related to kind and size of data
- finding best measures also needs application of machine learning

Thank you!