Permutation Complexity Measures for Time Series

Daryl DeFord and Kate Moore

Dartmouth College
Department of Mathematics

SIAM Annual Meeting
Pittsburgh, PA
July 11, 2017
Abstract

Permutation entropy has become a standard tool in time-series analysis that exploits the temporal properties of these data sets. Many current applications use an approach based on Shannon entropy, which implicitly assumes an underlying uniform distribution on patterns. In this paper, we consider several additional null models for time series data and determine the corresponding permutation distributions. This allows us to compare real-world data to more complex generative processes. Additionally, building on recent results of Martinez, we define a measure of complexity that allows us to characterize when a random walk is an appropriate model for a time series.
Outline

1. Introduction
2. Background and Motivation
3. Null Models
4. Random Walk Metrics
5. Walks on $S_n$
6. Conclusion
Permutation Patterns

\[
\begin{pmatrix}
3 & 6 & 7 & 9 & 4 & 8 & 1 & 2 & 5 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
\end{pmatrix}
\]
Permutation Patterns (123)

\[
\begin{pmatrix}
3 & 6 & 7 & 9 & 4 & 8 & 1 & 2 & 5 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{pmatrix}
\]
Permutation Patterns (123)

\[
\begin{pmatrix}
3 & 6 & 7 & 9 & 4 & 8 & 1 & 2 & 5 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{pmatrix}
\]
Permutation Patterns (231)

\[
\begin{pmatrix}
3 & 6 & 7 & 9 & 4 & 8 & 1 & 2 & 5 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{pmatrix}
\]
Permutation Patterns (312)

\[
\begin{pmatrix}
3 & 6 & 7 & 9 & 4 & 8 & 1 & 2 & 5 \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{pmatrix}
\]
Given a function $f : [0, 1] \rightarrow [0, 1]$ and a point $x \in [0, 1]$, consider the behavior of $\{x, f(x), f(f(x)), f(f(f(x))), \ldots\}$. 

Example

Let $f(x) = 4x(1-x)$ and $x_0 = 0.2$. Then, the list of values is:

$[0.2, 0.64, 0.92, 0.28, 0.82, 0.58, 0.97, 0.11, 0.40, \ldots]$. 

Given a function $f : [0, 1] \rightarrow [0, 1]$ and a point $x \in [0, 1]$, consider the behavior of \{ $x, f(x), f(f(x)), f(f(f(x))), \ldots$ \}.

Example

Let $f(x) = 4x(1 - x)$ and $x_0 = .2$. Then, the list of values is:

$[0.20, 0.64, 0.92, 0.28, 0.82, 0.58, 0.97, 0.11, 0.40, \ldots]$. 
Iterated Example

\[ 4x(1-x) \]

\[ x \]
Iterated Example
Iterated Example (12)
Iterated Example (231)
Iterated Example (2413)
Iterated Example (2413)
Iterated Example (35142)
Iterated Example (35142)
Iterated Example (351426)
Iterated Example (351426)
Iterated Example (4625371)
Iterated Example (4625371)
Iterated Example (57264813)
Iterated Example (268375914)
Forbidden Patterns
Forbidden Patterns

Definition (Topological Entropy)

\[ TE = \lim_{n \to \infty} \frac{\log(|\text{Allow}(f)|)}{n - 1} \]
Forbidden Patterns

Definition (Topological Entropy)

\[ TE = \lim_{n \to \infty} \frac{\log(|\text{Allow}(f)|)}{n - 1} \]


Simple Time Series
Simple Time Series

![Graph showing a simple time series with a linear relationship between time and value.]

- **Value** vs. **#occurrences**
- The graph on the left illustrates a linear increase in value with respect to time. The right graph shows the number of occurrences for each time series value.

---

Dartmouth Logo
Complex Time Series
Complex Time Series
Time Series Complexity

Definition (Permutation Entropy)

\[ PE(\{X_i\}) = -\frac{1}{\log(N!)} \sum_{\pi \in S_n} p_\pi \log(p_\pi) \]
# Time Series Complexity

## Definition (Permutation Entropy)

\[
PE\left(\{X_i\}\right) = -\frac{1}{\log(N!)} \sum_{\pi \in S_n} p_\pi \log(p_\pi)
\]

## Definition (Uniform KL Divergence)

\[
D_{KL}\left(\{X_i\}||\text{uniform}\right) = \sum_{\pi \in S_n} p_\pi \log \left(\frac{p_\pi}{\frac{1}{n!}}\right)
\]
Time Series Complexity

**Definition (Permutation Entropy)**

\[ PE(\{X_i\}) = -\frac{1}{\log(N!)} \sum_{\pi \in S_n} p_\pi \log(p_\pi) \]

**Definition (Uniform KL Divergence)**

\[ D_{KL}(\{X_i\}||\text{uniform}) = \sum_{\pi \in S_n} p_\pi \log \left( \frac{p_\pi}{\frac{1}{n!}} \right) \]

The uniformity assumption is already associated to a null model.

**Lemma**

Let \( \{X_i\} \) be a set of I.I.D. random variables, then \((\pi, t)\) and \((\tau, s)\) are independent random variables if and only if \(|t - s| > n\).
The uniformity assumption is already associated to a null model.

**Lemma**

Let \( \{X_i\} \) be a set of I.I.D. random variables, then \((\pi, t)\) and \((\tau, s)\) are independent random variables if and only if \(|t - s| > n\).

**Lemma**

Let \( \{X_i\} \) be a set of I.I.D. random variables, then \( P(\pi) = \frac{1}{n!} \) for all \( \pi \in S_n \).
Stock Data (Closing Prices)
Stock Data (n=3)
Stock Data (n=4)
Stock Data (n=5)
Stock Data (n=6)
Random Walk Models

**Definition (Random Walk)**

Let \( \{X_i\} \) be a set of I.I.D. random variables and define \( \{Z_i\} \) by

\[
Z_j = \sum_{i=0}^{j} X_j.
\]
Time Series Entropy
Null Models

Random Walk Models

Definition (Random Walk)
Let \( \{X_i\} \) be a set of I.I.D. random variables and define \( \{Z_i\} \) by
\[
Z_j = \sum_{i=0}^{j} X_j.
\]

Theorem
If \( \{Z_i\} \) are defined as above then either 123 \ldots n or \( n(n-1)(n-2)\ldots1 \) occurs with the highest probability.
Random Walk Models

Definition (Random Walk)
Let \( \{X_i\} \) be a set of I.I.D. random variables and define \( \{Z_i\} \) by
\[
Z_j = \sum_{i=0}^{j} X_j.
\]

Theorem
If \( \{Z_i\} \) are defined as above then either \( 123 \ldots n \) or \( n(n-1)(n-2)\ldots1 \)
occurs with the highest probability.

Corollary
If \( \{Z_i\} \) are defined as above and \( n \geq 3 \) then the expected distribution of
permutations is not uniform.
Uniform Walks

Definition

Let the \( \{X_i\} \) be defined as uniform random variables over \( [b - 1, b] \) where \( 0 < b < 1 \) and define \( \{Z_i\} \) with \( Z_j = \sum_{i=0}^{j} X_i \).
Definition

Let the \( \{X_i\} \) be defined as uniform random variables over \([b - 1, b]\) where \(0 < b < 1\) and define \( \{Z_i\} \) with \( Z_j = \sum_{i=0}^{j} X_i \).
**Expected Distributions**

**Theorem**

Let \( \{ Z_i \} \) be a random walk as defined above, then the expected distributions of permutations for \( S_3 \) and \( S_4 \) are characterized by the following:

\[
\begin{align*}
P(123) &= b^2 \\
P(132) + P(231) &= (1 - b)b \\
P(213) + P(312) &= (1 - b)b \\
P(1234) &= b^3 \\
P(4321) &= (1 - b)^3 \\
P(1243) + P(1342) + P(2341) &= (1 - b)b^2 \\
P(1432) + P(2431) + P(3421) &= (1 - b)^2b \\
P(2134) + P(3124) + P(4123) &= (1 - b)b^2 \\
P(3214) + P(4213) + P(4312) &= (1 - b)^2b \\
P(1324) + P(1423) + P(2314) + P(2413) + P(3412) &= (1 - b)b^2 \\
P(4231) + P(3241) + P(4132) + P(3142) + P(2143) &= (1 - b)^2b
\end{align*}
\]
Hyperplanes

Example

In order for the pattern 1342 to appear in the time series we need the following inequalities to hold:

- $X_1 > 0$
- $X_2 > 0$
- $X_3 < 0$
- $X_3 > X_2$
- $X_3 < X_1 + X_2$
**Hyperplanes**

**Example**

In order for the pattern 1342 to appear in the time series we need the following inequalities to hold:

- $X_1 > 0$
- $X_2 > 0$
- $X_3 < 0$
- $X_3 > X_2$
- $X_3 < X_1 + X_2$

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>$P(\pi)$</th>
<th>$\pi$</th>
<th>$P(\pi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>$b^2$</td>
<td>132</td>
<td>$(b - 1) - \frac{3}{2}(b - 1)^2$</td>
</tr>
<tr>
<td>213</td>
<td>$(b - 1) - \frac{3}{2}(b - 1)^2$</td>
<td>231</td>
<td>$\frac{1}{2}(b - 1)^2$</td>
</tr>
<tr>
<td>312</td>
<td>$\frac{1}{2}(b - 1)^2$</td>
<td>321</td>
<td>$(b - 1)^2$</td>
</tr>
</tbody>
</table>
Integration Regions
Stock Comparison
Equivalence Classes

**Theorem (Martinez 2015)**

Let \( \pi, \tau \in S_n \). The \( \mathbb{P}(\pi) = \mathbb{P}(\tau) \) for every probability distribution on the \( \{X_i\} \) if and only if \( \pi \) and \( \tau \) are equivalent under repeated application of the reverse complement operation on bounded cylindrical blocks.
Equivalence Classes

**Theorem (Martinez 2015)**

Let $\pi, \tau \in S_n$. The $P(\pi) = P(\tau)$ for every probability distribution on the $\{X_i\}$ if and only if $\pi$ and $\tau$ are equivalent under repeated application of the reverse complement operation on bounded cylindrical blocks.

\[
\begin{align*}
\{123\} & \{132, 213\} \{231, 312\} \{321\} \\
\{1234\} & \{1243, 2134\} \{1324\} \{1342, 3124\} \{1423, 2314\} \{1432, 2143, 3214\} \\
& \{2341, 3412, 4123\} \{2413\} \{2431, 4213\} \{4231\} \{3142\} \{3241, 4132\} \{3421, 4312\} \{4321\} \\
\{12345\} & \{14325\} \{21354\} \{21453\} \{25314\} \{41352\} \{45312\} \{52341\} \{54321\} \\
& \{53142\} \{42531\} \{53241\} \{52431\} \\
\end{align*}
\]
Related Metrics

**Definition**

To measure how closely the distribution matches the conditions of Martinez, we compute

\[
g_n(T) = \sum_{\Lambda_i \subseteq S_n} \sum_{\pi \in \Lambda_i} p_\pi |p_\pi - \mu_i|,
\]

or “equivalently”,

\[
h_n(T) = \sum_{\Lambda_i \subseteq S_n} \sum_{\pi \in \Lambda_i} p_\pi \log \left( \frac{p_\pi}{\mu_i} \right).
\]
Related Metrics

**Definition**

To measure how closely the distribution matches the conditions of Martinez we compute

\[ g_n(T) = \sum_{\Lambda_i \subseteq S_n} \sum_{\pi \in \Lambda_i} p_\pi |p_\pi - \mu_i|, \]

or “equivalently”,

\[ h_n(T) = \sum_{\Lambda_i \subseteq S_n} \sum_{\pi \in \Lambda_i} p_\pi \log \left( \frac{p_\pi}{\mu_i} \right). \]

**Corollary**

If \( Z^N = \{Z_i\}_{i=1}^N \) is a random walk, then

\[ \lim_{N \to \infty} g_n(Z^N) = 0 \quad \text{and} \quad \lim_{N \to \infty} h_n(Z^N) = 0. \]
Examples

- Uniform
- Exxon
- UPS
- Boeing
- Intel
- CCE
- Sin(n)
- Heart Rate

\[ g_0(T) \]

\[ 5 \cdot 10^{-2} \]

\[ 0.1 \]

\[ 0.15 \]
(a) The permutation graph for $n = 3$ whose edges, $\pi \rightarrow \tau$, are weighted with probability $P_X(\pi \rightarrow \tau)$. 
Lemma

Let $\pi, \tau \in S_n$. Then,

$$P(\tau \rightarrow \pi) = \frac{P(\tau \wedge \pi)}{P(\tau)}$$
Let $\pi, \tau \in S_n$. Then, 

$$P(\tau \rightarrow \pi) = \frac{P(\tau \wedge \pi)}{P(\tau)}$$

Lemma

Let $\{Z_i\}$ be a random walk as above. Then, the transition probability between permutations are not the uniform distribution.
CCE Random Walk
Further Reading


Code (Try it yourself...)

https://math.dartmouth.edu/~ddeford/time_series.html
That’s all...

Thank You!