Computational Neuroscience - Lecture 3

Kugiumtzis Dimitris

Department of Electrical and Computer Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, Greece
e-mail: dkugiu@auth.gr  http://users.auth.gr/dkugiu

22 May 2018
Outline

1. Single-channel EEG analysis - Introduction
2. Features of signal morphology
3. Features from linear analysis
4. Features from nonlinear analysis
Single-channel EEG analysis - Introduction
1. Single-channel EEG analysis - Introduction
2. Features of signal morphology
Outline

1. Single-channel EEG analysis - Introduction
2. Features of signal morphology
3. Features from linear analysis
1. Single-channel EEG analysis - Introduction
2. Features of signal morphology
3. Features from linear analysis
4. Features from nonlinear analysis
Single-channel EEG analysis - Introduction

Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI) ⇒ univariate time series analysis: study the static and dynamic properties of the time series (signal), model the time series.

Three main perspectives:
1. Characteristics of signal morphology, e.g. identify bursts.
2. Stochastic process (linear analysis), e.g.
   \[ X_t = \phi_0 + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t \]
3. Nonlinear dynamics and chaos (nonlinear analysis), e.g.
   \[ S_t = f(S_{t-1}), \quad S_t \in \mathbb{R}^d, \quad X_t = h(S_t) + \epsilon_t \]
Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI)
⇒
univariate time series analysis: study the static and dynamic properties of the time series (signal), model the time series.

Three main perspectives:
1. Characteristics of signal morphology, e.g. identify bursts.
2. Stochastic process (linear analysis), e.g. $X_t = \phi_0 + \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + \epsilon_t$
3. Nonlinear dynamics and chaos (nonlinear analysis), e.g. $S_t = f(S_{t-1}), S_t \in \mathbb{R}^d, X_t = h(S_t) + \epsilon_t$
Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI) ⇒
Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI) ⇒ univariate time series analysis: study the static and dynamic properties of the time series (signal), model the time series.
Single-channel EEG analysis - Introduction

Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI) ⇒ univariate time series analysis: study the static and dynamic properties of the time series (signal), model the time series.

Three main perspectives:
1. Characteristics of signal morphology, e.g. identify bursts.
Single-channel EEG analysis - Introduction

Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI) ⇒ univariate time series analysis: study the static and dynamic properties of the time series (signal), model the time series.

Three main perspectives:

1. Characteristics of signal morphology, e.g. identify bursts.

2. Stochastic process (linear analysis), e.g.

\[ X_t = \phi_0 + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t \]

Kugiumtzis Dimitris Computational Neuroscience - Lecture 3
Single-channel EEG analysis - Introduction

Region of Interest (ROI) of the brain

Eyeball judgement:
Doctor or EEGer may diagnose abnormalities by visual inspection of EEG signals (in ROI).

Alternatively:
Quantify automatically information from EEG signal (in ROI) ⇒

univariate time series analysis: study the static and dynamic properties of the time series (signal), model the time series.

Three main perspectives:

1. Characteristics of signal morphology, e.g. identify bursts.

2. stochastic process (linear analysis), e.g.
   \[ X_t = \phi_0 + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t \]

3. nonlinear dynamics and chaos (nonlinear analysis), e.g.
   \[ S_t = f(S_{t-1}), S_t \in \mathbb{R}^d, \]
   \[ X_t = h(S_t) + \epsilon_t \]
Clinical applications:

- predicting epileptic seizures
- classifying sleep stages
- measuring depth of anesthesia
- detection and monitoring of brain injury
- detecting abnormal brain states
- others ...
Clinical applications:
- predicting epileptic seizures
- classifying sleep stages
- measuring depth of anesthesia
- detection and monitoring of brain injury
- detecting abnormal brain states
- others ...

Example:
\(\alpha\)-wave (8-13 Hz) is reduced in children and in the elderly, and in patients with dementia, schizophrenia, stroke, and epilepsy
Features of signal morphology

Any feature $f_i$ is computed on sliding windows $i$ (overlapped or non-overlapped) across the EEG signal (of length $n_w$).
Any feature $f_i$ is computed on \textit{sliding windows $i$} (overlapped or non-overlapped) across the EEG signal (of length $n_w$). A feature can be \textit{normalized}, $f_i := f_i / \sum_{j=1}^{n} f_j$, where $n$ number of windows.

Features of signal morphology:
Indices (statistics) that capture some characteristic of the shape of the signal.
- Descriptive statistics: center (mean, median), dispersion (variance, SD, interquartile range), higher moments (skewness, kurtosis).
Features of signal morphology

Any feature $f_i$ is computed on **sliding windows** $i$ (overlapped or non-overlapped) across the EEG signal (of length $n_w$). A feature can be normalized, $f_i := f_i / \sum_{j=1}^{n} f_j$, where $n$ number of windows.

If $f_i$ is measured in several (many) channels, $f_{i,j}$, for channels $j = 1, \ldots, K$: take average at each time window, e.g. $f_i := \frac{1}{K} \sum_{j=1}^{K} f_{i,j}$. 

Indices (statistics) that capture some characteristic of the shape of the signal.

- **Descriptive statistics**: center (mean, median)
- **Dispersion**: variance, SD, interquartile range
- **Higher moments**: skewness, kurtosis
Any feature $f_i$ is computed on sliding windows $i$ (overlapped or non-overlapped) across the EEG signal (of length $n_w$). A feature can be normalized, $f_i := f_i / \sum_{j=1}^{n} f_j$, where $n$ number of windows.

If $f_i$ is measured in several (many) channels, $f_{i,j}$, for channels $j = 1, \ldots, K$: take average at each time window, e.g. $f_i := \frac{1}{K} \sum_{j=1}^{K} f_{i,j}$.

**Features of signal morphology:**
Indices (statistics) that capture some characteristic of the shape of the signal.
Features of signal morphology

Any feature $f_i$ is computed on \textit{sliding windows} $i$ (overlapped or non-overlapped) across the EEG signal (of length $n_w$). A feature can be \textbf{normalized}, $f_i := f_i / \sum_{j=1}^{n} f_j$, where $n$ number of windows.

If $f_i$ is measured in several (many) \textbf{channels}, $f_{i,j}$, for channels $j = 1, \ldots, K$: take \textit{average} at each time window, e.g. $f_i := \frac{1}{K} \sum_{j=1}^{K} f_{i,j}$.

\textbf{Features of signal morphology:}
Indices (statistics) that capture some characteristic of the shape of the signal.

\textbf{Descriptive statistics:}
- center (mean, median)
- dispersion (variance, SD, interquartile range)
- higher moments (skewness, kurtosis)
Hjorth parameters:

- activity = variance = total power, \( \text{Var}(x(t)) = \sum_{i=0}^{n_f} P_{XX}(f_i) \).

- mobility = \( \sqrt{\frac{\text{Var}(dx(t)dt)}{\text{Var}(x(t))}} \), estimates the SD of \( P_{XX}(f) \) (along the frequency axis).

- complexity = \( \frac{\text{mobility}(dx(t)dt)}{\text{mobility}(x(t))} \), estimates similarity of the signal to a pure sine wave (deviation of complexity from one).

Hjorth parameters:

- activity = variance = total power, $\text{Var}(x(t)) = \sum_{i=0}^{n_f} P_{XX}(f_i)$.
- mobility = $\sqrt{\text{Var} \left( \frac{dx(t)}{dt} \right) / \text{Var}(x(t))}$,
  estimates the SD of $P_{XX}(f)$ (along the frequency axis).

Hjorth parameters:

- activity = variance = total power, \( \text{Var}(x(t)) = \sum_{i=0}^{nf} P_{XX}(f_i) \).
- mobility = \( \sqrt{\text{Var} \left( \frac{dx(t)}{dt} \right) / \text{Var}(x(t))} \),
  estimates the SD of \( P_{XX}(f) \) (along the frequency axis)
- complexity = mobility \( \left( \frac{dx(t)}{dt} \right) / \text{mobility}(x(t)) \),
  estimates similarity of the signal to a pure sine wave
  (deviation of complexity from one).
Hjorth parameters:

- **activity** = variance = total power, \( \text{Var}(x(t)) = \sum_{i=0}^{n_f} P_{XX}(f_i) \).
- **mobility** = \( \sqrt{\text{Var} \left( \frac{dx(t)}{dt} \right) / \text{Var}(x(t))} \),
  estimates the SD of \( P_{XX}(f) \) (along the frequency axis)
- **complexity** = mobility \( \left( \frac{dx(t)}{dt} \right) / \text{mobility}(x(t)) \),
  estimates similarity of the signal to a pure sine wave (deviation of complexity from one).

Line Length: \( L(i) = \sum_{t=1}^{n_w-1} |x(t+1) - x(t)| \).

Normalized line length: \( L(i) := \frac{L(i)}{\sum_{j=1}^{n} L(j)} \).

Median over all channels: \( L(i) := \text{median} L(i,j) \), \( j \) denotes channel.
**Line Length:** \( L(i) = \sum_{t=1}^{n_w-1} |x(t+1) - x(t)|. \)

**Normalized line length:** \( L(i) := L(i) / \sum_{j=1}^{n} L(j). \)

**Median over all channels:** \( L(i) := \text{median}L(i,j), \) \( j \) denotes channel.

**Nonlinear energy:** \( \text{NLE}(i) = \sum_{t=2}^{n_w-1} |x(t)^2 - x(t-1)x(t+1)|. \)
... and many other features of signal morphology:

Fisher information (from normalized spectrum of singular values)
Katz’s fractal dimension (from line length and largest distance from start)

Petrosian Fractal Dimension (PFD) (from number of sign changes in the signal derivative)

Hurst exponent (the Rescaled Range statistics (R/S))
Detrended Fluctuation Analysis (DFA) (long range correlation measure)
Features of linear analysis - frequency domain

Features from **power spectrum** $P_{XX}(f_k)$, $f_k = 0, \ldots, f_s/2,$
Features from power spectrum $P_{XX}(f_k)$, $f_k = 0, \ldots, f_s/2$.

Power in bands:
$\delta[0, 4]$Hz, $\theta[4, 8]$Hz, $\alpha[8, 13]$Hz, $\beta[13, 30]$Hz, $\gamma > 30$Hz,
Features from power spectrum $P_{XX}(f_k)$, $f_k = 0, \ldots, f_s/2$.

Power in bands:
$\delta[0, 4]Hz$, $\theta[4, 8]Hz$, $\alpha[8, 13]Hz$, $\beta[13, 30]Hz$, $\gamma > 30Hz$,

e.g. for $\alpha$-band power: $P_{XX}(\alpha) = \sum_{f_k=8}^{f_k=13} P_{XX}(f_k)$
or relative $\alpha$-band power: $RP_{XX}(\alpha) = P_{XX}(\alpha)/\sum_{i=0}^{n_f} P_{XX}(f_i)$
Features from power spectrum $P_{XX}(f_k), f_k = 0, \ldots, f_s/2$.

Power in bands:
$\delta[0, 4]Hz, \theta[4, 8]Hz, \alpha[8, 13]Hz, \beta[13, 30]Hz, \gamma > 30Hz$,
e.g. for $\alpha$-band power: $P_{XX}(\alpha) = \sum_{f_k=8}^{f_k=13} P_{XX}(f_k)$
or relative $\alpha$-band power: $RP_{XX}(\alpha) = P_{XX}(\alpha)/\sum_{i=0}^{nf} P_{XX}(f_i)$

In the same way the bands are defined on wavelets.
Features from power spectrum $P_{XX}(f_k), f_k = 0, \ldots, f_s/2,$

Power in bands:
$\delta[0, 4]Hz, \theta[4, 8]Hz, \alpha[8, 13]Hz, \beta[13, 30]Hz, \gamma > 30Hz,$

e.g. for $\alpha$-band power: $P_{XX}(\alpha) = \sum_{f_k=8}^{f_k=13} P_{XX}(f_k)$
or relative $\alpha$-band power: $RP_{XX}(\alpha) = P_{XX}(\alpha)/\sum_{i=0}^{nf} P_{XX}(f_i)$

In the same way the bands are defined on wavelets.

Spectral edge frequency (SEF), e.g. SEF(90) is the frequency at 90% of total power, $f^{90}: \sum_{f=0}^{f^{90}} P_{XX}(f)/\sum_{f=0}^{f_s/2} P_{XX}(f) = 0.90$

median frequency: SEF(50)
Example:
5 datasets of 100 single-channel EEG of $N = 4096$ ($f_s = 173.61$ Hz)
A: scalp EEG, healthy eyes open    B: scalp EEG, healthy eyes closed
C, D: intracranial EEG, interictal period    E: intracranial EEG, ictal period
Example:
5 datasets of 100 single-channel EEG of $N = 4096\ (f_s = 173.61\ Hz)$
A: scalp EEG, healthy eyes open  B: scalp EEG, healthy eyes closed
C, D: intracranial EEG, interictal period  E: intracranial EEG, ictal period

Features of linear analysis - time domain

Autocorrelation: \( r_X(\tau) = \frac{c_X(\tau)}{s_X^2} = \frac{\sum_{t=1}^{n-\tau} (x_t - \bar{x})(x_{t+\tau} - \bar{x})}{\sum_{t=1}^{n-\tau} (x_t - \bar{x})^2} \)
Autocorrelation: $r_X(\tau) = \frac{c_X(\tau)}{s_X^2} = \frac{\sum_{t=1}^{n-\tau} (x_t-\bar{x})(x_{t+\tau}-\bar{x})}{\sum_{t=1}^{n-\tau} (x_t-\bar{x})^2}$

Assuming the linear stochastic process
$X_t = \phi_0 + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t$

Autoregressive model:
$\hat{X}_t = \hat{\phi}_0 + \hat{\phi}_1 x_{t-1} + \ldots + \hat{\phi}_p x_{t-p}$
Autocorrelation: \( r_X(\tau) = \frac{c_X(\tau)}{s_X^2} = \frac{\sum_{t=1}^{n-\tau} (x_t - \bar{x})(x_{t+\tau} - \bar{x})}{\sum_{t=1}^{n-\tau} (x_t - \bar{x})^2} \)

Assuming the linear stochastic process
\( X_t = \phi_0 + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t \)

Autoregressive model:
\( \hat{x}_t = \hat{\phi}_0 + \hat{\phi}_1 x_{t-1} + \ldots + \hat{\phi}_p x_{t-p} \)

Mean square error of fit: \( \text{MSE} = \frac{1}{n-p} \sum_{t=p+1}^{n} (x_t - \hat{x}_t)^2 \)
Features of nonlinear analysis - 1

Extension of autocorrelation to linear and nonlinear correlation

Entropy: information from each sample of $X$

$$H(X) = \sum x p_X(x) \log p_X(x)$$

Mutual information: information for $Y$ knowing $X$ and vice versa

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = \sum x, y p_{XY}(x, y) \log p_{XY}(x, y)$$

$$p_X(x)$$

$$p_Y(y)$$

For $X \rightarrow X_t$ and $Y \rightarrow X_{t+\tau}$, delayed mutual information

$$I_X(\tau) = I(X_t, X_{t+\tau}) = \sum x_t, x_{t+\tau} p_{X_t X_{t+\tau}}(x_t, x_{t+\tau}) \log p_{X_t X_{t+\tau}}(x_t, x_{t+\tau})$$

$$p_{X_t}(x_t)$$

$$p_{X_{t+\tau}}(x_{t+\tau})$$

To compute $I_X(\tau)$ make a partition of $\{x_t\}_{n_t=1}$, a partition of $\{y_t\}_{n_t=1}$ and compute probabilities for each cell from the relative frequency.
Features of nonlinear analysis - 1

Extension of autocorrelation to linear and nonlinear correlation

**Entropy**: information from each sample of $X$ (assume proper discretization of $X$)

$$H(X) = \sum_x p_X(x) \log p_X(x)$$
Features of nonlinear analysis - 1

Extension of autocorrelation to linear and nonlinear correlation

**Entropy**: information from each sample of $X$ (assume proper discretization of $X$)

$$H(X) = \sum_x p_X(x) \log p_X(x)$$

**Mutual information**: information for $Y$ knowing $X$ and vice versa

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}$$

Kugiumtzis Dimitris Computational Neuroscience - Lecture 3
Features of nonlinear analysis - 1

Extension of autocorrelation to linear and nonlinear correlation

**Entropy**: information from each sample of $X$ (assume proper discretization of $X$)

$$H(X) = \sum_x p_X(x) \log p_X(x)$$

**Mutual information**: information for $Y$ knowing $X$ and vice versa

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}$$

For $X \rightarrow X_t$ and $Y \rightarrow X_{t+\tau}$, **Delayed mutual information**:

$$I_X(\tau) = I(X_t, X_{t+\tau}) = \sum_{x_t,x_{t+\tau}} p_{X_tX_{t+\tau}}(x_t,x_{t+\tau}) \log \frac{p_{X_tX_{t+\tau}}(x_t,x_{t+\tau})}{p_{X_t}(x_t)p_{X_{t+\tau}}(x_{t+\tau})}$$
Features of nonlinear analysis - 1

Extension of autocorrelation to linear and nonlinear correlation

**Entropy**: information from each sample of $X$ (assume proper discretization of $X$)

$$H(X) = \sum_x p_X(x) \log p_X(x)$$

**Mutual information**: information for $Y$ knowing $X$ and vice versa

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = \sum_{x,y} p_{XY}(x, y) \log \frac{p_{XY}(x, y)}{p_X(x)p_Y(y)}$$

For $X \rightarrow X_t$ and $Y \rightarrow X_{t+\tau}$, **Delayed mutual information**:

$$I_X(\tau) = I(X_t, X_{t+\tau}) = \sum_{x_t,x_{t+\tau}} p_{X_tX_{t+\tau}}(x_t, x_{t+\tau}) \log \frac{p_{X_tX_{t+\tau}}(x_t, x_{t+\tau})}{p_{X_t}(x_t)p_{X_{t+\tau}}(x_{t+\tau})}$$

To compute $I_X(\tau)$ make a partition of $\{x_t\}_{t=1}^n$, a partition of $\{y_t\}_{t=1}^n$ and compute probabilities for each cell from the relative frequency.
Features of nonlinear analysis - 2

Extension of linear autoregressive models to nonlinear models

All models require the state space reconstruction, points $x_t \in \mathbb{R}^m$ from scalar $x_t$.

Delay embedding:
$$x_t = [x_t, x_{t-\tau}, ..., x_{t-(m-1)\tau}]$$

Nonlinear model:
$$x_{t+1} = f(x_t)$$

1. parametric models, e.g. polynomial autoregressive models
2. semilocal (black-box) models, e.g. neural networks
3. local models, e.g. nearest neighbor models.

For the prediction of $x_{t+1}$ having $x_1, x_2, ..., x_t$.

Find the $k$ nearest neighbors of $x_t$, $\{x_t^{(1)}, ..., x_t^{(k)}\}$.

Fit a linear autoregressive model on the $k$ neighbors.

$$x_{t+i} = \phi_0 + \phi_1 x_{t+i} + \cdots + \phi_m x_{t+i-(m-1)\tau} + \epsilon_{t+i}$$

Predict $\hat{x}_{t+1} = \hat{\phi}_0 + \hat{\phi}_1 x_{t+1} + \cdots + \hat{\phi}_m x_{t+1-(m-1)\tau}$. 
Extension of linear autoregressive models to **nonlinear models**

All models require the **state space reconstruction**, points $x_t \in \mathbb{R}^m$ from scalar $x_t$.

**Delay embedding:**

$x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]$. 
Features of nonlinear analysis - 2

Extension of linear autoregressive models to nonlinear models

All models require the state space reconstruction, points \( x_t \in \mathbb{R}^m \) from scalar \( x_t \).

Delay embedding:

\[
x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}].
\]

Nonlinear model: \( x_{t+1} = f(x_t) \)

1. parametric models, e.g. polynomial autoregressive models
2. semilocal (black-box) models, e.g. neural networks
3. local models, e.g. nearest neighbor models.
Features of nonlinear analysis - 2

Extension of linear autoregressive models to nonlinear models

All models require the state space reconstruction, points \( x_t \in \mathbb{R}^m \) from scalar \( x_t \).

**Delay embedding:**

\[
x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}].
\]

Nonlinear model: \( x_{t+1} = f(x_t) \)

1. **Parametric models**, e.g. polynomial autoregressive models
2. **Semilocal (black-box) models**, e.g. neural networks
3. **Local models**, e.g. nearest neighbor models.

For the prediction of \( x_{t+1} \) having \( x_1, x_2, \ldots, x_t \).

- Find the \( k \) nearest neighbors of \( x_t \), \( \{x_{t(1)}, \ldots, x_{t(k)}\} \)
- Fit a linear autoregressive model on the \( k \) neighbors.

\[
x_{t(i)+1} = \phi_0 + \phi_1 x_{t(i)} + \cdots + \phi_m x_{t(i)-(m-1)\tau} + \epsilon_{t+1}
\]

- Predict \( \hat{x}_{t+1} = \hat{\phi}_0 + \hat{\phi}_1 x_t + \cdots + \hat{\phi}_m x_{t-(m-1)\tau} \)
Entropy: Estimates of the entropy of 
\( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \):

\[
H(X) = \sum_x p_x(x) \log p_x(x)
\]
Entropy: Estimates of the entropy of 
\[ x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \]:
\[ H(X) = \sum_x p_X(x) \log p_X(x) \]

- **Approximate entropy**, ApEn (uses a bandwidth in the estimation of \( p_X(x) \))
Entropy: Estimates of the entropy of
\( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \):

\[
H(X) = \sum_x p_X(x) \log p_X(x)
\]

- **Approximate entropy**, ApEn (uses a bandwidth in the estimation of \( p_X(x) \))
- **Sample entropy** (similar to ApEn)
Entropy: Estimates of the entropy of \( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \):

\[
H(X) = \sum_x p_X(x) \log p_X(x)
\]

- **Approximate entropy**, ApEn (uses a bandwidth in the estimation of \( p_X(x) \))
- **Sample entropy** (similar to ApEn)
- **Permutation entropy** (entropy on ranks of the components in \( x_t \))
Entropy: Estimates of the entropy of 
\( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \):

\[ H(X) = \sum_x p_X(x) \log p_X(x) \]

- **Approximate entropy**, ApEn (uses a bandwidth in the estimation of \( p_X(x) \))
- **Sample entropy** (similar to ApEn)
- **Permutation entropy** (entropy on ranks of the components in \( x_t \))
- **Spectral entropy** (entropy on \( P_{XX}(f) \))

Others, e.g. fuzzy entropy, multiscale entropy.
Entropy: Estimates of the entropy of
\( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \):
\[
H(X) = \sum_x p_X(x) \log p_X(x)
\]
- Approximate entropy, ApEn (uses a bandwidth in the estimation of \( p_X(x) \))
- Sample entropy (similar to ApEn)
- Permutation entropy (entropy on ranks of the components in \( x_t \))
- Spectral entropy (entropy on \( P_{XX}(f) \))
- Others, e.g. fuzzy entropy, multiscale entropy.
Dimension and Complexity:

- Correlation dimension, estimates the fractal dimension of \( x_t = [x_t, x_{t-\tau}, ..., x_{t-(m-1)\tau}] \) (assuming chaos):
  \[ d(x_i, x_j) \]: distance of two points in \( \mathbb{R}^m \)
  Scaling of probability of distance with \( r \),
  \[ p(d(x_i, x_j) < r) \propto r^\nu \]
  e.g. \( \nu \): the dimension, estimated by the slope of log \( p(r) \) vs log \( r \).

- Higuchi dimension (distance is defined in a different way).
- Lempel-Ziv algorithmic complexity (turning the signal to a series of symbols by discretization and searching for new symbol patterns in the series).

Dimension and Complexity:

- **correlation dimension**, estimates the fractal dimension of
  \[ x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \] (assuming chaos):

  \[ d(x_i, x_j): \text{distance of two points in } \mathbb{R}^m \]

  scaling of probability of distance with \( r, p(d(x_i, x_j) < r) \propto r^\nu, \]

  e.g.

  \( \nu: \text{the dimension, estimated by the slope of } \log p(r) \text{ vs } \log r. \)
Dimension and Complexity:

- **correlation dimension**, estimates the fractal dimension of 
  \[ x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \] (assuming chaos):
  \[ d(x_i, x_j) \]: distance of two points in \( \mathbb{R}^m \)

  scaling of probability of distance with \( r \),
  \[ p(d(x_i, x_j) < r) \propto r^\nu, \]
  e.g.
  \( \nu \): the dimension, estimated by the slope of \( \log p(r) \) vs \( \log r \).

- **Higuchi dimension** (distance is defined in a different way).
Dimension and Complexity:

- **correlation dimension**, estimates the fractal dimension of \( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \) (assuming chaos):
  \[
d(x_i, x_j) = \text{distance of two points in } \mathbb{R}^m\]
  scaling of probability of distance with \( r \), \( p(d(x_i, x_j) < r) \propto r^{\nu} \), e.g.
  \( \nu \): the dimension, estimated by the slope of \( \log p(r) \) vs \( \log r \).

- **Higuchi dimension** (distance is defined in a different way).

- **Lempel-Ziv algorithmic complexity** (turning the signal to a series of symbols by discretization and searching for new symbol patterns in the series).
Dimension and Complexity:

- **correlation dimension**, estimates the fractal dimension of 
  \( x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}] \) (assuming chaos):
  
  \[ d(x_i, x_j) \]: distance of two points in \( \mathbb{R}^m \)
  
  scaling of probability of distance with \( r \), 
  \[ p(d(x_i, x_j) < r) \propto r^{\nu} \],
  
  e.g. \( \nu \): the dimension, estimated by the slope of \( \log p(r) \) vs \( \log r \).

- **Higuchi dimension** (distance is defined in a different way).

- **Lempel-Ziv algorithmic complexity** (turning the signal to a series of symbols by discretization and searching for new symbol patterns in the series).

**Source:** [Kugiumtzis et al, Int. J of Bioelectromagnetism, 2007]

[Kugiumtzis and Tsimpiris, J. Stat. Softw., 2010], MATS module in Matlab
Example: 5 datasets of healthy extracranial EEG (A,B) and intracranial interictal (C,D) and ictal (E) EEG

[Bao et al, Comput Intell Neurosci, 2011]