Entropy Measures in Time Series

Entropy in time-series tells us **how predictable or complex** a signal is. If a signal is **very regular and repetitive**, it has **low entropy**. If a signal is **chaotic and unpredictable**, it has **high entropy**.

Signal Type	Description	Entropy
ECG heartbeat (healthy)	Regular but slightly variable	Medium
Sinus wave	Perfectly regular	Low
Random noise	Completely unpredictable	High

So entropy quantifies the **degree of randomness** or **irregularity** in the signal.

Why We Use Entropy in Time-Series

Entropy measures are used to **quantify complexity** — not just noise or randomness, but structured variability.

💡 Applications

- **Biomedical:** detect heart disease from ECG variability, brain disorder detection from EEG.
- **Finance:** measure market unpredictability (volatility patterns).
- **Climate science:** detect regime shifts (rainfall, temperature irregularity).
- **Sensor monitoring:** detect faults when signals become too regular or too chaotic.

Main Entropy Families in Time-Series

There are many, but the two classic ones are:

1. **Approximate Entropy (ApEn)** – earlier method
2. **Sample Entropy (SampEn)** – improved version

Approximate Entropy (ApEn)

📖 Definition

Approximate Entropy (ApEn) measures the **likelihood that similar patterns of observations** in a time series will remain similar in the future.

It's the **logarithmic probability** that patterns of length **m** that are similar within a tolerance **r** remain similar when one more point is added.

Parameter	Meaning
m	Embedding dimension (length of pattern to compare)
r	Tolerance (usually r = 0.1 × std of signal)
N	Number of data points

Interpretation

ApEn value	Meaning
Low	More predictable, regular (less complex)
High	Less predictable, more complex

Sample Entropy (SampEn)

📖 Definition

Sample Entropy (SampEn) fixes ApEn's bias by **excluding self-matches** and improves stability for smaller datasets.

SampEn = negative natural log of the conditional probability that sequences similar for m points remain similar for m+1 points.

Interpretation

SampEn value	Meaning
Low	Regular and predictable
High	Complex and irregular
Equal to zero	Perfectly predictable pattern

Example Intuition

Signal A: 1, 1, 1, 1, 1, 1

→ Perfectly regular → low entropy (near 0)

Signal B: 1, 2, 3, 2, 3, 4, 3

→ Some variation → medium entropy

Signal C: Random noise

→ Highly unpredictable → high entropy

ApEn vs. SampEn

Feature	Approximate Entropy (ApEn)	Sample Entropy (SampEn)
Self-matching	Included	Excluded
Bias	Biased (especially small N)	Less biased
Consistency	Depends on length	Consistent
Ease of implementation	Simpler	Slightly harder
Preferred for	Long data	Short/Noisy data

Key Terms

Concept	Summary
Entropy	Measures unpredictability/complexity of a time series
ApEn	First algorithm to measure regularity
SampEn	Improved version, unbiased, excludes self-matches
High Entropy	More irregular, complex
Low Entropy	More regular, predictable
Used in	ECG, EEG, weather, finance, machinery monitoring
Related Measures	Shannon, Permutation, Fuzzy, Multiscale Entropy

Forecastability Formula

Forecastability is a measure of **how predictable** a time series is by measuring how concentrated its frequency spectrum.

Forecastability Formula

The forecastability $F(x)$ of a time series $x(t)$ is:

$$F(x) = 1 - \frac{H_s(x)}{\ln(2\pi)}$$

Where:

- $H_s(x)$ = **spectral entropy** of the time series
- $\ln(2\pi)$ = maximum possible entropy for a purely random time series
- $F(x) \in [0, 1]$

Interpretation

Forecastability	Meaning
(F(x) = 0)	Completely unpredictable (white noise)
(F(x) = 1)	Perfectly predictable (pure sine wave)
(0 < F(x) < 1)	Some level of structure + predictability

A **predictable** time series has **dominant frequencies** in its spectrum.

Example:

- A pure sine wave = all energy in one frequency → predictable → low entropy → **high forecastability**
- White noise = energy spread across all frequencies → high entropy → **low forecastability**

Interpretation of Forecastability Values

Type of Signal	Example	Entropy	Forecastability
Pure sinusoid	Sine wave	Low	**High (≈1)**
Biological signal	Heart rate	Medium	Medium
Financial markets	Stock price	High	Low
White noise	Random noise	Max	**0**

So forecastability is basically *“how much useful structure does the time series have for prediction?”*