

MOD500 Decision Analysis with Artificial Intelligence Support

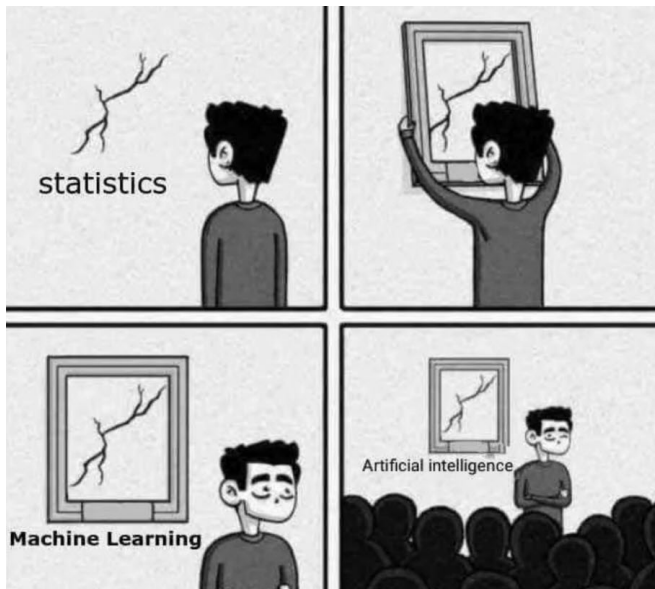
Enrico Riccardi¹

Department of Energy Resources, University of Stavanger (UiS).¹

Oct 17, 2024



A bias recap from last classes



1 Metrics and distributions

2 Information Theory

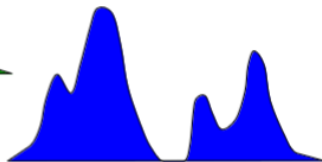
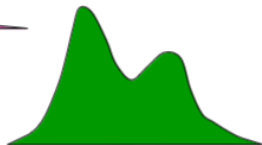
3 Comparing information

Probability distribution

A metric is needed

Amount of uncertainty

Number of Modes: unimodal, bimodal, polymodal



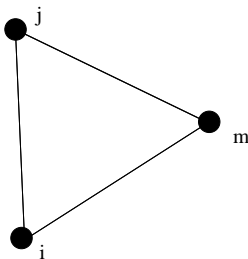
Metrics

Triangle inequality

Considering a vectors in an N-dimensional space, to be a **distance** it must satisfy the **triangle inequality**:

$$d_{ij} + d_{im} \geq d_{jm}$$

If also $d_{jj} = 0$, if $i = j$ and $d_{jj} - d_{ii} = 0$, then we call it a *metric*.



Common metrics:

- Counting
- Euclidean

$$d_{ij}^{(E)} = \left[\sum_{k=1}^N (x_{ik} - x_{jk})^2 \right]^{\frac{1}{2}}$$

Categorical variables do not necessarily satisfy these relations.

Common metrics:

- Counting
- Euclidean

$$d_{ij}^{(E)} = \left[\sum_{k=1}^N (x_{ik} - x_{jk})^2 \right]^{\frac{1}{2}}$$

Categorical variables do not necessarily satisfy these relations.

Task assignment

Code a discrete probability distribution in Python

Calculate the Mean and Standard Deviation

How to get an experiment out of this distribution?

1 Metrics and distributions

2 Information Theory

3 Comparing information

How to relate a distribution to information?

First, the origins:

- Daddy: Claude Shannon (1940)
- His initial work has been done on signal transmission.
- It uses Entropy as key measurement of information uncertainty.

How to relate a distribution to information?

First, the origins:

- Daddy: Claude Shannon (1940)
- His initial work has been done on signal transmission.
- It uses Entropy as key measurement of information uncertainty.

How to relate a distribution to information?

First, the origins:

- Daddy: Claude Shannon (1940)
- His initial work has been done on signal transmission.
- It uses Entropy as key measurement of information uncertainty.

How to relate a distribution to information?

First, the origins:

- Daddy: Claude Shannon (1940)
- His initial work has been done on signal transmission.
- It uses Entropy as key measurement of information uncertainty.

How Entropy can alter a decisions?

It is an interface between data and decisions.

A question to sum up the idea

Does more data bring value?

It has permitted the advances of several fields:

cryptography, neurobiology, signal processing, linguistics,
bioinformatics, statistical physics, black holes, quantum computing,
information retrieval, intelligence gathering, plagiarism detection,
pattern recognition, anomaly detection, etc

zip files, phones, internet!

How Entropy can alter a decisions?

It is an interface between data and decisions.

A question to sum up the idea

Does more data bring value?

It has permitted the advances of several fields:

cryptography, neurobiology, signal processing, linguistics,
bioinformatics, statistical physics, black holes, quantum computing,
information retrieval, intelligence gathering, plagiarism detection,
pattern recognition, anomaly detection, etc

zip files, phones, internet!

How Entropy can alter a decisions?

It is an interface between data and decisions.

A question to sum up the idea

Does more data bring value?

It has permitted the advances of several fields:

cryptography, neurobiology, signal processing, linguistics,
bioinformatics, statistical physics, black holes, quantum computing,
information retrieval, intelligence gathering, plagiarism detection,
pattern recognition, anomaly detection, etc

zip files, phones, internet!

How Entropy can alter a decisions?

It is an interface between data and decisions.

A question to sum up the idea

Does more data bring value?

It has permitted the advances of several fields:

cryptography, neurobiology, signal processing, linguistics,
bioinformatics, statistical physics, black holes, quantum computing,
information retrieval, intelligence gathering, plagiarism detection,
pattern recognition, anomaly detection, etc

zip files, phones, internet!

Entropy of an information source

$$H_X = - \sum_i (p_i) \log(p_i)$$

H_X of a discrete random variable X is a measure of the amount of uncertainty associated with the value of X when only its distribution is known

What is p_i ?

It is a numerical descriptions of how likely an event is to occur

Do not mix the concepts!

Assigned probability and computed probability are different

Entropy of an information source

$$H_X = - \sum_i (p_i) \log(p_i)$$

H_X of a discrete random variable X is a measure of the amount of uncertainty associated with the value of X when only its distribution is known

What is p_i ?

It is a numerical descriptions of how likely an event is to occur

Do not mix the concepts!

Assigned probability and computed probability are different

Entropy of an information source

$$H_X = - \sum_i (p_i) \log(p_i)$$

H_X of a discrete random variable X is a measure of the amount of uncertainty associated with the value of X when only its distribution is known

What is p_i ?

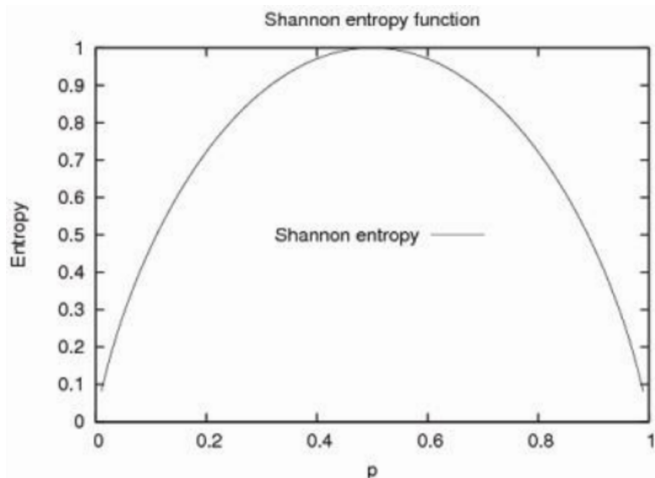
It is a numerical descriptions of how likely an event is to occur

Do not mix the concepts!

Assigned probability and computed probability are different

Shannon's entropy shape

$$H_X = - \sum_i (p_i) \log(p_i)$$



Task assignment [2]

Code a discrete probability distribution in Python

Calculate the Mean and Standard Deviation

How to get an experiment out of this distribution?

Calculate Shannon's Entropy as a function of the number of experiments

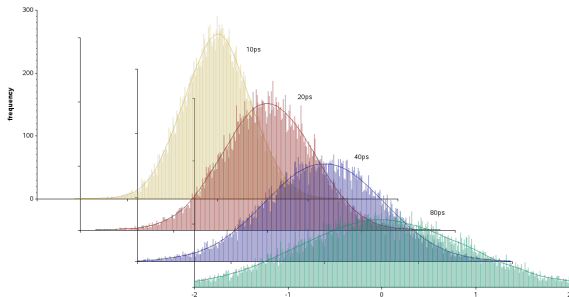
1 Metrics and distributions

2 Information Theory

3 Comparing information

Comparing information

A distribution can be the sum of multiple distributions



Kullback–Leibler divergence (information gain)

How to compare multiple information sources?

$$D_{KL} = \sum_i (p_i) \log\left(\frac{p_i}{q_i}\right)$$

- It is a divergence, as it is asymmetric
- It is NOT a metric

Kullback–Leibler divergence (information gain)

How to compare multiple information sources?

$$D_{KL} = \sum_i (p_i) \log\left(\frac{p_i}{q_i}\right)$$

- It is a divergence, as it is asymmetric
- It is NOT a metric

Download a discrete probability distribution

Calculate Kullback–Leibler divergence

Find the distribution that minimise K-L divergence