MOD500 Decision Analysis with Artificial Intelligence Support

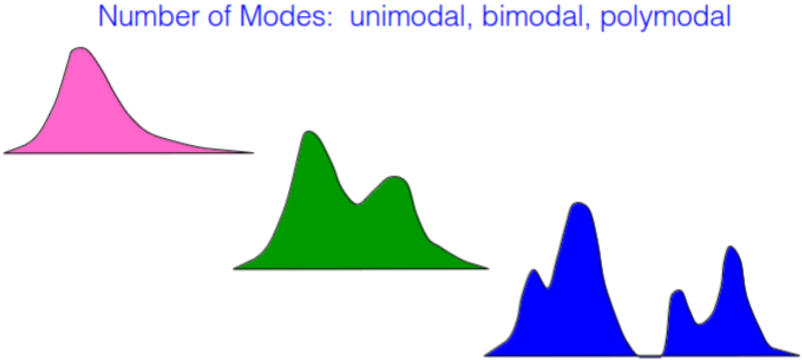
Enrico Riccardi[1]

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

Oct 17, 2024

Probability distribution

Amount of uncertainty



Task assignment

Code a discrete probability distribution in Python

Calculate the Mean and Standard Deviation

How to get an experiment out of this distribution?

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 uncertainity.

Why is it important/userfull ?

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

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

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

Task assignment

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

Kullback–Leibler divergence (information gain)

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