### Classic versus Bayesian approach

Statistics can be split in two separate approaches: 1) Classic or frequentist 2) Bayesian. Frequentist statistics tries to eliminate uncertainty by providing estimates while Bayesian statistics allows the update of prior beliefs based on new information available. Classic versus Bayesian approach has been a long debate in statistics. Bayesian statistics seems to be closer to AI and machine learning. Research showed that human brain behaves like a Bayesian predictor, updating its beliefs and forecasts based on new information received. Therefore studying Bayesian statistics is crucial for the in depth understanding of machine learning and AI. Bayesian statistics is based on Bayes’ formula.

*Bayesian statistics allows us to update our subjective beliefs in light of new data or evidence.*

Frequentist statistics -> probabilities are the long-run frequency of random events in repeated trials

Bayesian statistics -> the subjective prior beliefs are updated based on new data or evidence

### Bayesian framework in examples

### Probability of an event occurring is the main concept in Bayesian framework. The probability can take values from 0 to 1.

### Example 1

*Probability = 0 -> No chance of an event occurring*

*Probability = 1 -> Absolutely certain of an event occurring*

### Example 2

### Bayesian statistics in finance:

*Some finance experts consider that certain large companies are impossible to go bankrupt. They associate a zero probability to this event. Lehman Brothers, Barings Bank and Long-Term Capital Management failure show that large company bankruptcy is a rare event but not impossible. Therefore we can use Bayesian statistics to associate a low probability to this event rather than consider it zero.*

### Bayes’ Rule

– *prior*

The strength of belief without considering information

– *posterior*

The updated belief considering information

– *likelihood*

The probability of seeing data as generated by a model with parameter

– *evidence*

The probability of the data by summing all the possible values of weighted by how strongly we believe in those particular values of .

### Coin-Flipping

Coin-flipping is a well-known example in statistics for showing random independent outcomes.

|  |  |
| --- | --- |
| Bayes’ Rule components | Coin Flipping description |
| *Prior* | Our prior view regarding the probability of how fair the coin is |
| *Posterior* | The updated view of the fairness of the coin after seeing n1 heads out of n2 flips |
| *Likelihood* | The probability of seeing a number of heads in a particular number of flips |
| *Evidence* | The probability of seeing a certain sequence of flips for all possibilities of our belief in the coin’s fairness |

### Bayesian Approach concepts

|  |  |
| --- | --- |
| Concept | Explanation |
| *Assumptions* | * Coin flipping has two outcomes * Flips are random and independent events * Coin fairness is stationary (quantified by parameter) |
| *Prior beliefs* | * The prior beliefs regarding coin fairness are quantified using a probability distribution * Beta distribution is used in coin flipping case |
| *Experimental data* | * Coin-flips will be carried in order to obtain data * The number of heads will be quantified considering the total number flips * The probability of obtaining heads will be quantified considering given fairness of the coin () * Likelihood functions will be considered in this case |
| *Posterior beliefs* | * Posterior belief regarding the fairness of the coin is constructed using the prior belief, the likelihood function and Bayes’ rule * The prior beliefs are coupled with the observed data and are updated * Are also known as conjugate priors * Beta distribution is considered posterior |
| *Inference* | * Once we have a posterior belief we can estimate the coin’s fairness * Predict the probability of heads on the next flip * See how results depend upon different choices of prior beliefs |