Statistics & Probability – Given data how can we use them.

Both deal with uncertainty, RANDOMNESS.

Probability: It deals with the prediction of likelihood of future events.

- logically self contained
- Has one correct ans

Statistics: Have Past data and doing analysis currently. Involves the analysis of the frequency of past events

- works on experimental data
- No single correct ans.

Datasets: Some may increase noise (irrelevant features) helps understand patterns, relations, trends

Statistics – Classification (Given features, understand data)

- Descriptive: summarizing and presenting data in concise manner.

  Ex: Mean, Median, mode

- Inferential: Involves using sample data to draw inferences, predictions for larger population.

  Ex: Confidence intervals / Drug Testing*

Q. what is sample data, larger population?

Ex: Voting in a state

Population: Entire group we want to draw conclusions from.

Sample: A specific group is chosen by random sampling

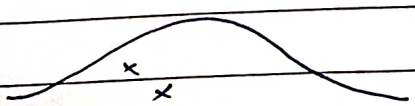
Population ⇄ Sample (Sampling →, ← Inference)

Independent Identical Distributed

↳ mutually independent

↳ None of the datapoints should be related.

↳ identically distributed (⇒ comes from same prob distrib)



when dataset is being prepared samples must be IID

Probability models:

Discrete Probability Models: In discrete probality models we compute prob by adding all corresponding outcomes.

- Bernoulli Model: Binary outcomes (yes or no)
- Binomial Model: Describes the # of successes in ind trials
- Poisson Model: # of events in a fixed time/space interval.
  Ex: like no. of flights in 30 min duration

[redacted] predictions are based of independent events.

Continuous Probability Models:

- A random var X can take on any value (is continuous).
- Uniform Distribution: where all values in interval are equally likely.
- Normal / Gaussian Distribution: symmetrically distributed (Bell curve)

Joint Probability
The likelihood of
- More than one event is occuring at the same time.

Conditions: ① X & Y must happen at same time
② X & Y must be independent of each other.

Q. what will happen to joint prob of two dependent events
↳ leads to Conditional Probability

Given an event A has already occurred, what is the prob that B will take place.

∴ Joint Probability of two dependent events

$$P(A \cap B) = P(A) \cdot P(B|A)$$

Bayes Theorem

Joint Probability of two dependent events:

$P(A \cap B) = P(A)\,P(B|A)$ $\quad$ $P(B \cap A) = P(B) \cdot P(A|B)$

$\Rightarrow P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$ $\quad$ $\Rightarrow P(A|B) = P(A) \cdot P(B|A) / P(B)$

Probability Distribution.

learns how data is distributed

(we can generate new samples. Image classification - Dogs & Cats

- A probability distribution is a mathematical function that describes the likelihood of various outcomes in a random experiment.

Probability Mass Function - Complete distribution

Ex: Rolling 1000 times (↑ iterations ⇒ equally likely)

Prob of each outcome

at larger iterations, it becomes almost equally likely.

$$E(x) = \sum x_i \cdot P_i$$

$$PMF = f(x, p) = \begin{cases} p & x = 0 \\ q = 1-p & x = 1 \end{cases}$$

most used

Prob density function (Distribution over a range)

- of a cont random var, the set of possible outcomes is an uncountably ∞ range.

$$P(a \le X \le b) = \int_a^b f(x) \cdot dx$$

Gaussian Distribution → Bell shape

$$PDF = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

$$PMF = \int PDF \cdot dx$$

Normal distribution: (68-95-99.7 rule)

~~Data~~ Describes the approximate distribution of data in a normal distribution.

1σ - 68.26%

2σ - 95.44%

3σ - 99.72%

rest 0.28% outliers: ones that create noise in the dataset.

Var less - Bell curve more spiked

variance more - flattened curve.