

1. Define Experiment, Sample space, Outcome and Event.

* **Experimen**t- An experiment with a known set of pssible outcomes is called a experiment.
* **Sample space**-The set of all possible outcomes.
* **Outcome**- A result of a random experiment.
* **Even**t-  A subset of the sample space

1. What is probability and explain different types of probability?

* **Probability**- It is used to measure the Uncertainity

There are two types of probability-

1. **Theoretical Probability**-

The theoretical probability is defined as the ratio of the number of favourable outcomes to the number of possible outcomes.

2. **Empirical Probability-**

Empirical probability is a type of experimental probability that depends on past data

3. In loan defaulters older people make up only 1.4%. Now the probability that someone defaults on a loan is 0.184, Find the probability of default on loan knowing that he is an old person. Older people make up only 0.8%

Sol-

p(yes)=0.184

p(old person)=0.8% =0.008

p(older person| yes)= 1.4%=0.14

p(yes | older person)=p(yes) x p(olderperson | yes) / p(older person)

= 0.184x0.14 / 0.08

= 0.322

4. Define Bayes theorem and write the formulae.

* **Bayes Theorem**-

Bayes' Theorem states that the conditional probability of an event, based on the occurrence of another event, is equal to the likelihood of the second event given the first event multiplied by the probability of the first event.

P(A|B)= [P(B|A). P(A)] / P(B)

5. Solve the below problem using Bayes theorem:

Spam Assassin works by having users train the system. It looks for patterns in the words in emails marked as spam by the user.

For example, it may have learned that the word “free” appears in 30% of the mails marked as spam, i.e., P(Free | Spam) = 0.30. Assuming 1% of non-spam mail includes the word “free” and 50% of all mails received by the user are spam, find the probability that a mail is spam if the word “free” appears in it.

**Sol-**

P(spam)=50% =0.50

P(free|spam)=30% =0.30

P(free|not sapm)=1%=0.01

P(not spam)=5%=0.50

P(spam|free)=P(free|spam) x P(spam)/ P(free|spam) x p(spam)+ P(free|notspam) x P(spam)

=0.30 x 0.50 / (0.30 x 0.50) + (0.01 x 0.50)

= 0.967