



1. Define Experiment, Sample space, Outcome and Event.
2. What is probability and explain different types of probability?
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%
4. Define Bayes theorem and write the formulae.
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.,  $P(\text{Free} | \text{Spam}) = 0.30$ . 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.

1. Experiment :

It is an action that can be repeated infinite no of times and has a well-defined set of possible outcomes.

ex : rolling a die ...

Random Experiment :

Experiment we conduct where

the output is uncertain.

ex : tossing a coin ....

Sample space :

It has all the possible outcomes

for the experiment

ex : rolling a die  $S = \{1, 2, 3, 4, 5, 6\}$

Event :

It is the set of favorable outcomes of an experiment

It <sup>is</sup> subset of sample space.

ex : rolling a even number the set of outcomes  $\{2, 4, 6\}$ .

Outcome :

It is a single result of an experiment

ex : rolling a die is a number b/w 1 and 6, inclusive

2. probability

It is simply how likely something is to happen

$$P(E) = \frac{n(\text{favourable cases})}{n(\text{total cases})}$$

different types of probability :

1. Theoretical probability
2. Experimental probability
3. Axiomatic probability

3.

Given that

$$\begin{aligned} p(\text{older people} \mid \text{loan defaulters}) &= 1.4\% \\ &= 0.014 \end{aligned}$$

$$p(\text{loan default}) = 0.184$$

$$\begin{aligned} p(\text{older people}) &= 0.8\% \\ &= 0.008 \end{aligned}$$

$$p(\text{loan default} \mid \text{older people}) = ?$$

from Bayes theorem

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$$P(\text{loan default} / \text{older people}) = \frac{0.014 \times 0.184}{0.008} = \underline{0.322}$$

4. Bayes theorem :

It can be derived for events and random variables separately using the definition of conditional probability

Bayes theorem :

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad \text{where } P(B) \neq 0$$

$P(A|B)$  is the probability of condition when event A is occurring while event B is already occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{where } P(B) \neq 0$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \quad \text{where } P(A) \neq 0$$

so the joint probability  $P(A \cap B)$  of both events A and B being

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

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

$$\therefore P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad \text{where } P(B) \neq 0$$

5.

Given that

$$p(\text{free} | \text{spam}) = 0.30$$

$$p(\text{free} | \text{not spam}) = 1\% = 0.01$$

$$p(\text{spam}) = 50\% = 0.50$$

$$p(\text{not} | \text{spam}) = 1 - p(\text{spam})$$

$$= 1 - 0.50$$

$$= 0.50$$

$$p(\text{Free}) = p(\text{free} | \text{spam}) * p(\text{spam}) + p(\text{free} | \text{not spam}) * p(\text{not spam})$$

$$p(\text{free}) = 0.30 * 0.50 + 0.01 * 0.50$$

$$= 0.165$$

from Bayes theorem

$$p(\text{spam} | \text{free}) = \frac{p(\text{free} | \text{spam}) * p(\text{spam})}{p(\text{free})}$$

$$= \frac{0.30 * 0.50}{0.165}$$

$$= \underline{\underline{0.909}}$$