PROBABILITY

Randomness

Random: -True Random means no predictive Power.

When We can't predict if it is true Randomness.

Example: - Flipping a Coin.

Event & Sample space

In a random experiment

Set of all possible outcomes→ Sample space S S={H,T}

Set of desired possible outcome→Event E E={H}

Event E is the Subset of Sample Space S.

Example:-

Tossing a dice., getting even no.

S=6# (1,2,3,4,5,6)

Probability of event

E=3# getting even no.(2,4,6)

E= count(E)/count(S)

Probability=E/S=3/6=1/2.

When toss a coin :-

Probability of Head=0.5

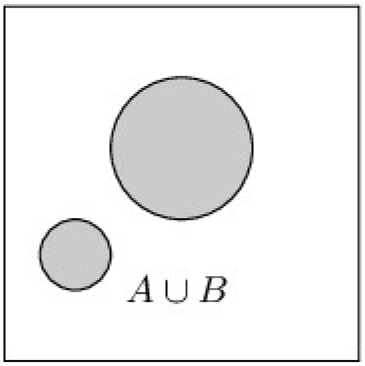
Probability of Tail= 0.5

Union and Intersection

Union

The union of two EVENTS that contains all of the elements that are in at least one of the two sets. The union is written as AUB it is **"A or B"**.

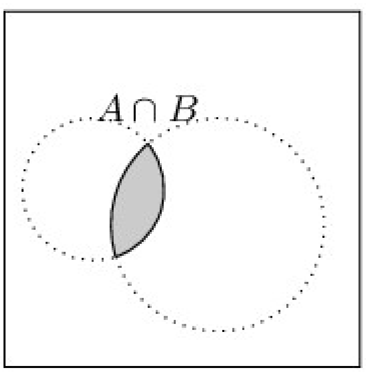
AUB occurs if the event A occurs or the Event B occurs.



Intersection

The intersection of two sets is a new set that contains all of the elements that are in both sets. The intersection is written as A∩B it is **"A and B"**.

A∩B occurs means Event A and Event B both occurs.



here we removing P(A∩B) once because

1) P(A∩B) will be null means no common elements

2) If there is common elements P(A),P(B) each of them have P(A∩B), so P(A)+P(B) will have P(A∩B) 2 times, so we need to remove P(A∩B) once to avoid duplicate

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

$$P(A\cap B)=P(A)+P(B)-P(A\cup B)$$

•In order to get common part we have to remove P(AUB) from P(A)+P(B)

Mutually exclusive

Two events A & B are called mutually exclusive if

$A \cap B = \varnothing$ null event → Event A and B can't occur together

Example:- flipping a coin , Head and Tail can't occurs together

$P(H \cap T) = \varnothing$ null event

$H \cup T$ it is **"H or T"**.

$P(H \cup T) = P(H) + P(T)$ coz $P(H \cap T) = \varnothing$

Compliment EVENT: -

Compliment of A →Probability(Not A) = 1- P(A) $P(A)+P(A')=1$

Q.A box contains 20 cards, numbered from 1 to 20. A card is drawn from the box at random. Find the probability that the number on the card drawn is (i) even (ii) prime and (iii) multiple of 3.

Ans:-

Even number = 2,4,6,8,10,12,14,16,18,20

Prime Number = 2,3,5,7,11,13,17,19

Multiple of 3 = 3,6,9,12,15,18

(i)Probability of getting an Even Number

$$= \frac{\text{Number of Even Numbers}}{\text{Total Numbers}} = \frac{10}{20}$$

(ii)Probability of getting a prime number

$$= \frac{\text{Number of prime Numbers}}{\text{Total Numbers}} = \frac{8}{20}$$

(iii)Probability of getting a multiple of 3

$$= \frac{\text{Number of mu;ltiples of 3}}{\text{Total numbers}} = \frac{6}{20}$$

EVENTS

(1) Dependent Events or conditional

(2) Independent Events.

Dependent Event: - A card is chosen out of 52 cards

(no replacement allowed).

4/52 → probability first card is Queen.

4/51→ Probability next card is Jack.

P(Queen and Jack)→ 4/52 x 4/51=16/2652=4/663. P(A and B)=P(A)*P(B)

P(Queen OR Jack)→4/52 + 4/51. P(A U B)=P(A)+P(B)

Independent Event: - A card is chosen out of 52 cards (replacement allowed).

4/52 → probability first card is Queen.

4/52→ Probability next card is Jack.

P(Queen and Jack)→ 4/52 x 4/52.

P(Queen OR Jack)→4/52 + 4/52.

Addition Theorem of Probability

Or Theorem of Total probability: -

Addition Theorem: - If two events A and B are mutually exclusive, the probability of occurrence of either A or B is the sum of the individual probability of A and B symbolically.

P(A or B)=P(A)+P(B)

The theorem can be extended to three or more mutually exclusive events.

P(A or B or C)=P(A)+P(B)+P(C).

Example: - A bag contains 30 balls numbered 1 to 30. One ball is drawn at random, find the probability that the number of the ball will be the multiple of 5 or 9.

Sol. :- Number of multiple of 5 (Event A)=(5,10,15,20,25 ,30)=6

Number of multiple of 9(Event B)=(9,18,27)=3

Sample space=Total no. of Event=30

we can see both events dont have common elements so they are mutually exclusive.

P(A)=6/30

P(B)=3/30

P(A or B)=P(A)+P(B)

6/30 + 3/30 =9/30 =3/10.

(Here the Events are mutually exclusive)

Multiplication Theorem of Probability

The is theorem of probability states that if two events A and B are independent the probability the probability that both will occur is equal to the product of their individual probabilities.

P(A and B)= P(A) x P(B)

The theorem can be extended to three or more independent events.

P(A and B and C)=P(A)x P(B)x P(C).

Example: - A bag contains 5 white and 3 black balls. Two Balls are drawn at random one after another without replacement. Find the probability that

(a)Both balls are black. P(B and B)=3/8*2/7=3/28

(b)One black and one white. P(B and W)=3/8*5/7= 0.267

Answer: -

(a) P(A) =Probability of drawing a black ball in the first attempt=3/(3+5)=3/8.
P(B)=Probability of drawing again a black ball in the second attempt=2/(2+5)=2/7.
P(A and B)=P(A) x P(B)= 3/8 x 2/7 =3/28.

(b) ----------- try yourself

Conditional probability or dependent probability: -

If we are told that event B has occurred, what is the probability that Event A will also occur.?

P(A/B)→ Probability of A , when B already occurred.

$$P(A/B) = P(A\cap B)/P[B]$$

P(spam/if there is bad word) =P(mail is spam and there is bad word)/P(bad word)

Q. If P(A)=7/13, P(B)=9/13, P(A∩B)=4/13.

What is P(A/B)?

P(A/B)→Probability of A when B has already occurred.

Ans:-P(A/B)=P(A∩B)/P(B)= (4/13)/(9/13)=4/9.

Q.A family has two children , what is the probability that both the children are boys. Given that at least one of them is a boy.

Ans:- Recall the Formula

P(A/B)→ Probability of A , when B already occurred.

$P(A/B) = P(A \cap B)/P[B]$

E→ both the children are boy.

F→at least one of them is a boy.

Sample space={bb,bg,gb,gg}

$P(E)=1/4$

$P(F)=3/4$

$(E \cap F)=\{bb\}$

$P(E \cap F)=1/4$

P(E/F)→ probability that both children are boy , given that one child is boy.

$P(E/F)= P(E \cap F)/P(F)=(1/4)/3/4)=1/3$.

Q. A New movie released, for a married couple the probability that husband will watch the movie is 70%.The probability that wife watch the movie is 65 %. The probability that both will watch the movie is 60%.

If the husband is watching the movie,

what is the probability that wife is also watching the movie?

Ans:- Try yourself

$$P(A/B) = P(A \cap B)/P[B]$$

P(H)=.70
P(W)=.65
P(H and W)=.60
P(W/H)=P(W and H)/P(H)
.60/.70=.8571

Q. In a school there are 1000 students, out of which 430 are girls, it is known that out of 430 girls 10% girls are studying in class 12 .

A student chosen randomly from the school,

what is the probability that chosen one is a student of class 12?.

It is given that the chosen student is a girl. P(S of 12th/G) =P(S of 12th and G)/P(G)
43/1000/(430/1000) =.1

Ans:- Try your self

$P(A/B) = P(A \cap B)/P[B]$

Q.

owner	Probability Have pet animal	Probability Don't have Pet animal	total
male	0.41	0.08	0.49
female	0.45	0.06	0.51
total	0.86	0.14	1

What is the probability that the randomly selected person is a male?

Given that the selected person has a Pet animal?

Ans:- Recall the Formula

P(A/B)→ Probability of A , when B already occurred.

$P(A/B) = P(A \cap B)/P[B]$

P(M/P)=P(M and P)/P(P)
= .41/.86

$P(M)=0.49$

$P(M \cap \text{Pet-own})=0.41$

$P(M/PO) = P(M \cap \text{Pet-own})/P(\text{Pet-own})$

=0.41/0.86=0.4777

P(A/B)=P(A and B)/P(B)
P(B/A)=P(B and B)/P(A)
P(A/B)*P(B)=P(B/A)*P(A)
P(A/B)=P(B/A)*P(A)/P(B)

Bayes Theorem

Named after Thomas Bayes.

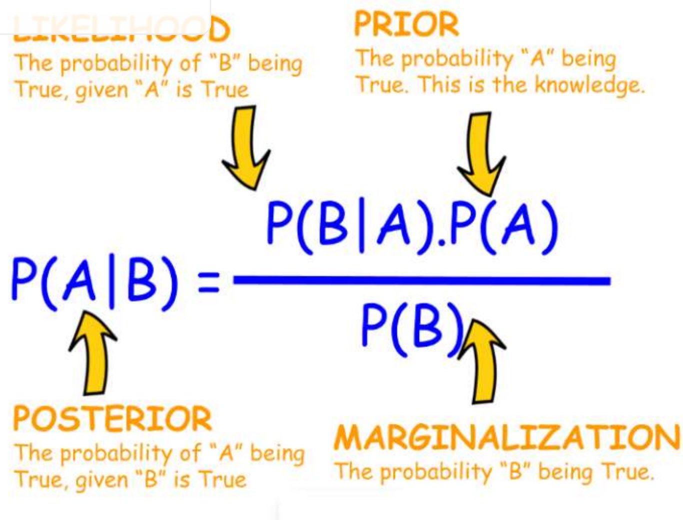
Bayes theorem is actually an extension of conditional probability. It is represented as:

P(A|B) = P(B|A) * P(A)/P(B)

Here, The conditional probability answers the probability of occurrence of A when B has already occurred. In that case, the Bayes theorem answers using prior beliefs and comes to a posterior conclusion.

Describes the probability of an event based on the prior knowledge of conditions that might be related to the event.

The conditional Probability is known as hypothesis. The Hypothesis is calculated through previous evidence or knowledge.



Deriving Bayes Equation from conditional Probability:-

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

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

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

Solving above two equations we get Bayes Theorem.

One key to understanding the essence of Bayes' theorem is to recognize that we are dealing with sequential events, whereby new additional information is obtained for a subsequent event, and that new information is used to revise the probability of the initial event.

In this context, the terms prior probability and posterior probability are commonly used.

Definitions

A prior probability is an initial probability value originally obtained before any additional information is obtained.

A posterior probability is a probability value that has been revised by using additional information that is later obtained.

Example 1

What is the probability of a patient having liver disease if they are alcoholic?

P(d/a)=(p(a/d)*p(d))/p(a)

Given data(**Prior Information): -**

(1)As per earlier records of the clinic, it states that 10% of the patient's entering the clinic are suffering from liver disease. p(d)=.1

(2)Earlier records of the clinic showed that 5% of the patients entering the clinic are alcoholic. p(a)=.05

(3)Earlier records of the clinic showed, 7% out of the patient's that are diagnosed with liver disease, are alcoholics. P(a/d)=.07

This defines the B|A: probability of a patient being alcoholic, given that they have a liver disease is 7%.

What is the probability that Patient Being Alcoholic, chances that he is having a liver disease?

Ans:- (.07*.1)/.05=0.14

P(A)=Probability that Patient having liver disease =0.10

P(B)= Probability that Patient is alcoholic =0.05

$P(A|B)$= Probability that Patient having liver disease, it is known that he is alcoholic =?

$P(B|A)$= Probability that Patient is alcoholic having liver disease=0.07.

As, per Bayes theorem formula,

$$P(A|B) = (P(B|A) * P(A))/ P(B)$$

$$P(A|B) = (0.07 * 0.1)/0.05 = 0.14$$

Therefore, for a patient being alcoholic, the chances of having a liver disease is 0.14 (14%).

This is a large increase from 10% suggested by past data.

Example 2

In a particular pain clinic , 10% of patients are prescribed narcotic pain killers. Overall 5% of clinic patients are addicted to narcotics (including pain killers and illegal substances)

Out of all the people prescribed pain killers 8% are addicted.

If a patient is addicted what is the probability that he is prescribed pain killer?

ANS:-

TRY YOUR SELF

P(pres/a)=(p(a/pres)*p(pres))/p(a)
p(a)=.05
p(pres)=0.1
p(a/pres)=.08

(.08*.01)/.05

Example 3: You are planning a picnic today, but the morning is cloudy. 50% of all rainy days start off cloudy. But cloudy mornings are common (about 40% of days start cloudy). Also this is usually a dry month (only 3 of 30 days tend to be rainy, or 10%). What is the chance of rain during the day?

Sol:-

Try yourself.

p(r/c)=p(c/r)*p(r)/p(c)
=(.5*.1)/.4

Example 4

What is the probability of a patient having liver disease if they are alcoholic?

Given data(**Prior Information): -**

(1)As per earlier records of the clinic, it states that 10% of the patient's entering the clinic are suffering from liver disease.

(2)Earlier records of the clinic showed that 5% of the patients entering the clinic are alcoholic.

(3)Earlier records of the clinic showed, 7% out of the patient's that are diagnosed with liver disease, are alcoholics.

This defines the $B|A$: probability of a patient being alcoholic, given that they have a liver disease is 7%.

What is the probability that Patient Being Alcoholic, chances that he is having a liver disease?

Ans:

$P(A)$=Probability that Patient having liver disease =0.10

$P(B)$= Probability that Patient is alcoholic =0.05

$P(A|B)$= Probability that Patient having liver disease, it is known that he is alcoholic =?

$P(B|A)$= Probability that Patient is alcoholic having liver disease=0.07.

As, per Bayes theorem formula,

$$P(A|B) = (P(B|A) * P(A))/ P(B)$$

$$P(A|B) = (0.07 * 0.1)/0.05 = 0.14$$

Therefore, for a patient being alcoholic, the chances of having a liver disease is 0.14 (14%).

This is a large increase from 10% suggested by past data.

Besides statistics, the Bayes' theorem is also used in various disciplines, with medicine and pharmacology as the most notable examples. In addition, the theorem is commonly employed in different fields of finance. modeling the risk of lending money to borrowers or forecasting the probability of the success of an investment.

Bayes theorem also used for SPAM Filters.

Bayesian Spam Filtering

it's used to filter spam. The **event** in this case is that the message is spam. The **test** for spam is that the message contains some flagged words (like "Money, Lottery" or "you have won"). Here's the equation set up (from Wikipedia), read as "The probability a message is spam given that it contains certain flagged words":

$$\Pr(\text{spam}|\text{words}) = \frac{\Pr(\text{words}|\text{spam})\Pr(\text{spam})}{\Pr(\text{words})}$$

Example 5.

The probability that the word MONEY word appears in an email, given that the email is spam is 8%.

Probability that an email can be a spam =20%

Probability that Money can appear in an email=2.4%

Find the probability that the email is spam, given that Money word is in the email?

Solution:-

M→Event of Money, S→ Even of SPAM

P(Money/Spam)=0.08,   P(Spam)=0.2 , P(Money)=0.024

But We are interested to know the probability that the email is spam, given that Money appears in the email.

P(Spam/Money)=?

P(Spam/Money)=P(Money/Spam)x P(Spam)/P(Money)

=0.08x0.2/0.024=0.67=67%.

Please Note:- There are several forms of Bay's theorem out there, and they are all equivalent (they are written in slightly different ways)

Here we have three conditions:-

$$P(A/D) = \frac{P(A) * P(D/A)}{P(A)*P(D/A) + P(B) * P(D/B) + P(C) * P(D/C)}$$

Example 6

Black box manufactures for aircraft

A → 75% production → defect 4%.

B → 15% production → defect 6%.

C → 10% production → defect 8%.

A defective Black box is randomly chosen, what is the probability that it was manufactured by company A?

Ans:-

Prior Information

P(A)=75%=0.75

P(B)=15%=0.15

P(C)=10%=0.10

P(D/A)=4%=0.04

P(D/B)=6%=0.06

P(D/C)=6%=0.08

P(A/D)=?

$$P(A/D) = \frac{P(A) * P(D/A)}{P(A)*P(D/A) + P(B) * P(D/B) + P(C) * P(D/C)}$$

=0.75*0.04 / (O.75*0.04)+(0.15*0.06)+(0.10+0.08)

=0.03/0.03+0.009+0.008

=0.03/0.047 = 30/47

=0.6382

Example 7

A factory produces an item using three machines A,B & C which account for 20%, 30% and 50% of its output respectively.

Defective Item manufactured by machine A, B & C are 5 % , 3 % and 1% respectively. If a randomly selected item is defective, what is the probability is was produced by machine C?

Ans:- Try yourself.

Given that item is defective, find the probability of it being Machine C product?

P(C/D)=?

$$P(C/D) = \frac{P(C) * P(D/C)}{P(A)*P(D/A) + P(B) * P(D/B) + P(C) * P(D/C)}$$

Please Note:- One another forms of Bay's theorem , related to True Positive, True negative condition.

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not } B)\,P(\text{not } B)}$$

$$P(B|A) = \frac{P(A|B)\,.\,P(B)}{P(A)} \qquad P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not } B)\,P(\text{not } B)}$$

Please note :-

Regarding TP,TN,FP,FN :-

P(test positive/no disease)+ P(test negative/no disease)=1

P(test positive/ disease)+ P(test negative/ disease)=1

Example 8.

1% of people have genetic defect.

90% of test correctly detect the Genetic defect (True Positive).

9.6% of the test are False Positive.

If a person gets a positive result, what is the probability that actually the person has the genetic defect?

Ans:-

Assume

A=Positive test report

B=Person having Genetic Disorder.

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not }B)\,P(\text{not }B)}$$

P(A/B)=Probability of finding the Positive report when the person having the Genetic disorder =90%.

P(B)=Probability the person has genetic defect=1%

P(A/~B)=Probability with positive test result person has not genetic defect=9.6%

P(~B)=1-P(B)=100-01=99% =Probability that Person has not genetic defect=99%

P(B/A)=probability that actually the person has the genetic defect when the test showing the Positive Report.

P(B/A)=?

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not } B)\,P(\text{not } B)}$$

P(B/A)=(0.9x0.01)/[(0.9x0.01)+(0.096x0.99)]

=0.09/[0.09+0.09504]

=0.09/0.18504

=0.4863

=48.63 % chance that in spite of having a positive test report person is patient of Genetic disorder.

Q 9 . Epidemiologists claim that the probability of breast cancer among Caucasian women in their mid-50 is 0.005. An established test identified people who had breast cancer and were healthy. A new Mammography test in clinical trials has a probability of 0.85 for detecting for detecting cancer correctly. In women without breast cancer, it has a chance of 0.925 for a negative result. If a 55 year old Caucasian woman tests positive for breast cancer, what is the probability that she,in fact,has breast cancer?

Solution:-

Assume :-

A=Positive test report

B=Person having Disease.

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not } B)\,P(\text{not } B)}$$

P(A/B)=Probability of finding the Positive report when the person having the Disease =0.85

P(B)=Probability the person has Disease=0.005

P(~A/~B)=Probability with negative test result person has not Disease= 0.925

Important Note:-

P(test positive/no disease)+ P(test negative/no disease)=1

P(test positive/ disease)+ P(test negative/ disease)=1

P(~A/~B) + P(A/~B)=1

P(A/~B)=1- P(~A/~B)=1-0.925=0.075.

P(not B)=P(~B)=1-P(B)=1-0.005=0.995

P(B/A)=probability that actually the person has the Disease when the test showing the Positive Report.

P(B/A)=?

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not } B)\,P(\text{not } B)}$$

P(B/A)=0.85*0.005/[(0.85*0.005)+(0.075*0.995)]

=0.00425/[0.00425+0.074625] = 0.00425/0.078875

=0.05388

Q10. 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 20% of the mails marked as spam. Assuming 0.1% 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.

Ans:-

Try yourself

Q 11 :- 1% of the population has certain disease. If an infected person is tested, then there is a 95% chance that the test is positive. If the person is not infected, then there is a 2% chance that the test gives an erroneous positive result (“False Positive”).

Given that a person tests positive, what are the chances that she/he has the disease?

Answer:-

The probability that a person has a disease given that the predicted result is Positive.

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|\text{not } B)\,P(\text{not } B)}$$

Try yourself…