Statistics & Probability

DeepLearning.AI

Independent Event: Throwing coins one after another. The outcome doesn't depend on each other.
So, they are independent event.

Dependent Event: On chess, 11 th move depends on the 10 th move.
So, that is dependent event.

For independent event, $P(A \cap B) = P(A) \cdot P(B)$

Question: You tossed a coin five times. What is the probability of landing heads five time?

$$P(\text{Head}) = \left(\frac{1}{2}\right)^5$$
$$= \frac{1}{32}$$

Question: You have 2 dice. What is the probability that both of the dice will 6,6 after 1 throw?

Combination of 2 dice $= (6)^2 = 36$

So, $Pr(6,6) = \frac{1}{36}$

[Both of the events are independent]

Question: If you have 10 dice, what is the probability of getting 10 sixes?

$$Pr(6,6,6,6,6,6,6,6,6,6) = \left(\frac{1}{6}\right)^{10}$$

$$=$$

Conditional probability: What is the probability of something given that something is already happened?

Example, What is the probability of landing on heads twice given that the first one was head after landing.

$$P(HH \mid 1st \text{ is } H) = ?$$

	2nd H	2nd T
1st H	HH	HT
1st T	TH	TT

→ Initial

As for one time ~~time~~ the coin tossed and Heads come,

So, ~~Number~~ Sample space will reduce to →

	H	T
H	HH	HT

Already Head appeared

So, posibilities are either (H,H) or (H,T)

So probability $(H,H) = \frac{1}{2}$

Question: What is the probability of landing on heads twice given that first one was tails

	H	T
H	HH	HT
T	TH	TT

→ From this sample event we can see that there is no chance of landing on heads twice if given that first one is tail.

Because for tail at first (T,H) and (T,T) this posibilities available

So, $P(HH \mid 1st \text{ is tail}) = 0$

For Dependent event, $P(A \cap B) = P(A) \cdot P(B \mid A)$

Question: What is the probability of two dice that the sum = 10. Given that the first dice = 6

Here, $P(B \mid A) = \frac{1}{6}$

[P(A) = sum 10
P(B) = First dice 6]

[(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)]

↳ First one 6

$\cancel{P(A \cap B) = P(A) \cdot P(B \mid A)}$

$= \frac{3}{36} \cdot \frac{1}{36}$

$= \frac{1}{12} \cdot \frac{1}{6} = \frac{1}{84}$

Question: Among 100 students 40% play soccer. The students who play soccer among them 80% were running shoes. Show the probability of students play soccer and were running shoes?

40% percent student play soccer, mean $P(\text{soccer}) = 0.4$

Among them, 80% wear running shoes, means $P(\text{Running Shoes} \mid \text{Soccer})$

$\cancel{0.4 \times 80\%} = 80\%$

$\cancel{0.4 \times 0.8} = 0.8$

$=$

$\therefore$ P(soccer and Running Shoe)

$\Rightarrow P(S \cap R) = P(S) \times P(R|S)$

$= 0.4 \times 0.8$

$= 0.32 = 32\%$

Question: The probability of a kid wears running shoes when they don't like soccer = 50%. So what's the probability of not playing soccer and wear running shoe?

40% play soccer, Means 60% don't play soccer

$\therefore P(NS) = 0.6$

Among them 50% were running shoes; $P(R \mid NS) = 0.5$

$$P(NS \cap R) = P(NS) \cdot P(R|NS)$$
$$= 0.6 \times 0.5$$
$$= 0.3$$

Bayes Theorem

Imagine a scenario. There is a rare diseases going on and you are going to get tested for it. You go to a doctor who said he has a very effective test which is 99% accurate most of the time. You did the test and tested positive. Now calculate the probability that you have the diseases given the fact that you tested positive.

Population = 1000000 (1 M)

Illness effect 1 person in every 1000 people

Test effectiveness = 99% (accuracy)

99% accuracy actually means two thing →

1) Out of 100 sick people, 99 people tested sick
   1 people tested Healthy
   (although he is sick)

2) Out of 100 healthy people, 99 people tested healthy
   1 people tested sick
   (Although he is healthy)

You went to the doctor and tested sick. The question is are you really sick or not. More specificly,

What is the probability that you are sick, given that you are tested sick?

Answer:

Among 1000000 people, 1 out of every 10000 people are sick

Means total 100 people are sick from 1000000 people

Healthy people = 1000000 − 100 = 999900

Sick people = 100

These are actual number

But in terms of diagnosis, test accuracy is 99%

~~Means, from 100 sick people 99 people are sick~~

~~∴ 1 " = $\frac{99}{100}$ " " "~~

~~∴ 1000000 " " $\frac{99 \times 1000000}{100}$ are sick~~

~~= 990000~~

Means,

From 999900 people who are healthy

Among them, 9999 people diagnosed sick

(Although they are healthy)

From, 100 sick people, 1 people diagnosed healthy
(Although he is sick)

Now we have to find the probability of people who are sick given that they are diagnosed sick.

$$P(\text{sick} \mid \text{diagnosed Sick}) = \frac{99}{99+9999}$$

Actually healthy but tested sick

Actually sick and tested sick

$$= 0.0098$$

So, You are diagnosed sick. But the probability that you are actually sick is < ~~100~~ 1%.

- 1M People
  - sick = 100
    - tested sick | sick = 99% → 99 people
    - tested healthy | sick = 1% → 1 people
  - healthy = 999,900
    - tested sick | healthy ~~99~~ .1% → 9999 people
    - tested healthy | healthy = 99% → 989901 people

$P(\text{sick}) = 0.01\% = P(A)$

$P(\text{not sick}) = 99.99\% = P(A')$

Let's say → A: Sick

B: Diagnosed Sick

$P(\text{diagnosed sick} \mid \text{sick}) = 99\% \quad P(B|A)$

$P(\text{diagnosed sick} \mid \text{not sick}) = 1\% \quad P(B|A')$

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

$$P(\text{sick} \mid \text{Diagnosed Sick}) = \frac{P(\text{sick}) \cdot P(\text{diagnosed sick} \mid \text{sick})}{P(\text{sick}) \cdot P(\text{diagnosed sick} \mid \text{sick}) + P(\text{not sick}) \cdot P(\text{diagnosed sick} \mid \text{not sick})}$$

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

$$= \frac{0.0001 \times 0.99}{\cancel{0.0001} \; (0.0001 \times 0.99) + (0.9999 \times 0.01)}$$

$$= P(A|B) = 0.0098$$

Bayes Theorem - Spam Example:

We have a dataset of 100 emails in which 20 emails are spam. So we want to build a classifier. The classifier says everything is a spam with a 20% probability. Now we will consider some spam word to check. We will ~~the~~ find the word "Lottery" in the spam and ham emails. Suppose, 14 out of 20 Spam email contain "lottery" word. And from the ham emails, there are 10 emails out of 80, which contains the word "lottery".

Now what is the probability that an email containing the word "lottery" is a spam? P(spam | lottery)

A = spam
B = lottery

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

We will notice only the emails that contains word lottery. Other emails doesn't matter.

24 emails contain the word lottery.

$$P(\text{spam} | \text{Lottery}) = \frac{P(\text{spam and Lottery})}{P(\text{Lottery})}$$

$$= \frac{14}{24}$$

$$= \frac{7}{12}$$

$$= 0.583$$

According to Bayer' theorem,

Spam : A
Lottery : B

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

$$P(\text{Spam}|\text{Lottery}) = \frac{P(\text{spam}) \cdot P(\text{Lottery}|\text{Spam})}{P(\text{Spam}) \cdot P(\text{Lottery}|\text{Spam}) + P(\text{Not spam}) \cdot P(\text{Lottery}|\text{Not spam})}$$

From the story context, $P(\text{spam}) = \frac{20}{100}$

$= 0.2$

$P(\text{not spam}) = 1 - 0.2 = 0.8$

$P(\text{Lottery}|\text{Spam}) = \frac{14}{20}$

$= 0.7$

Because,
Spam emails = 20
from spam emails
14 are lottery

$P(\text{Lottery}|\text{Not spam}) = \frac{10}{80}$

$= 0.125$

Because
Not spam email = 80
from not spam emails,
10 are lottery.

$$\therefore P(\text{Spam}|\text{Lottery}) = \frac{0.2 \times 0.7}{(0.2 \times 0.7) + (0.8 \times 0.125)}$$

$= 0.583$

Also, Bayer Theorem can be written as $\rightarrow \frac{P(B|A) \cdot P(A)}{P(B)}$

Given $P(B) != 0$

Prior	Event	Posterior
$P(spam) = \frac{20}{100}$	Email contain "lottery"	$P(spam \mid lottery) = 0.583$

In case of the dice example,

where you have to find the probability of getting the sum 10 after throwing a ~~dice~~ second dice, given that 1st one is 6

Prior	Event	Posterior
$P(sum=10) = \frac{3}{36}$	First dice = 6	$P(sum=10 \mid first\ dice=6) = \frac{1}{36}$

In case of the coin example,

where you have to find the probability of getting two heads, given that first coin is head.

Prior	Event	Posterior
$P(HH) = \frac{1}{4}$	first coin = H	$P(HH \mid first = H) = \frac{1}{2}$

Problem Question: A Factory producing markers have three machines

Factory

- M1 (20 out of 100) → Manufacturing defect → 5%
- M2 (30 out of 100) → Manufacturing defect → 3%
- M3 (50 out of ~~300~~ 100) → Manufacturing defect → 1%

So, M_1 Makes 20% of Markers and has 5% defect

M_2 " 30% " " " 3% "

M_3 " 50% " " " 1% "

Question: I have a batch of markers. I randomely pick one marker and found it defective. Tell the probability it has been manufactured by M_3?

Here,

$$P(M_1) = \frac{20}{100} = \frac{1}{5}, \quad P(M_2) = \frac{30}{100} = \frac{3}{10}, \quad P(M_3) = \frac{50}{100} = \frac{1}{2}$$

$$P(D|M_1) = \frac{5}{100} = \frac{1}{20}, \quad P(D|M_2) = \frac{3}{100}, \quad P(D|M_3) = \frac{1}{100} \quad [D = \text{Defect}]$$

We have to find $P(M_3|D)$?

According to Bayes theorem,

$$P(M_3|D) = \frac{P(D|M_3) \cdot P(M_3)}{P(D)}$$

$$= \frac{1/100 \times \frac{1}{2}}{P(D)}$$

We can see

$$P(D) = P(D \cap M_1) + P(D \cap M_2) + P(D \cap M_3)$$

Now,

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

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

$$\Rightarrow P(D \cap M_1) = P(D|M_1) \times P(M_1)$$

$$= \frac{1}{20} \times \frac{1}{5}$$

$$= \frac{1}{100}$$

$$P(D \cap M_2) = P(D|M_2) \times P(M_2).$$

$$= \frac{3}{100} \times \frac{3}{10}$$

$$= \frac{9}{1000}$$

$$P(D \cap M_3) = P(D|M_3) \times P(M_3)$$

$$= \frac{1}{100} \times \frac{1}{2}$$

$$= \frac{1}{200}$$

$$\therefore P(D) = \frac{1}{100} + \frac{9}{1000} + \frac{1}{200}$$

$$= \frac{10+9+5}{1000}$$

$$= \frac{\cancel{24}}{\cancel{1000}} = \frac{3}{175}$$

$$\therefore P(M_3|D) = \frac{\frac{1}{100} \times \frac{1}{2}}{\frac{3}{175}} = \frac{1}{200} \times \frac{175}{3} = \frac{\cancel{175}}{\cancel{600}} = \frac{7}{24}$$

$$= \frac{7}{24}$$

Naive Bayes Classifier:

CSK Cricke data

Toss	Venue	Outlook	Result
Won	Mumbai	Overcast	Won
Lost	Chennai	Sunny	Won
Won	Kolkata	Sunny	Won
Lost	Mumbai	Sunny	Lost
Won	Chennai	Overcast	Lost
Won	Kolkata	Overcast	Lost
Won	Mumbai	Sunny	Won

Predict that, for {lost, Mumbai, sunny} CSK would win or lost?

Now ~~wat~~ what Naive Bayes will do is, it will calculate the probabilities →

$$P(W \mid \{lost, Mumbai, Sunny\})$$

and $P(L \mid \{lost, Mumbai, Sunny\})$ [W = Win, L = Lose]

Whose probability will be greater, that would be ~~so~~ the prediction.

From, Bayes theorem we know,

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

$$\therefore P(W \mid \{lost, Mumbai, Sunny\}) = \frac{\overset{P(\{l,M,s\} \mid W)}{\cancel{P(W \mid \{l,M,s\})}} \cdot P(W)}{P(\{l,M,s\})}$$

Now if we look back the table,

We will find → $P(W \mid \{l,M,s\}) = 0$

But that will make the whole prediction 0.

That's why we will change the formula a bit.

$$\cancel{P(W \mid \{l,M,s\})} = \frac{\cancel{P(W \mid l) \cdot P(W \mid M) \cdot P(W \mid s) \cdot P(W)}}{\cancel{P(W \mid l) \cdot P(W \mid M) \cdot P(W \mid s)} \quad \cancel{P(\{l,M,s\})}}$$

$$P(W|\{l,M,S\}) = \frac{P(l|W)\cdot P(M|W)\cdot P(S|W)\cdot P(W)}{P(\{l,M,S\})}$$

$$= \frac{\frac{1}{4}\cdot\frac{2}{4}\cdot\frac{3}{4}\cdot\frac{4}{7}}{\frac{1}{7}}$$

$$= 0.375$$

$$\therefore P(L|\{\text{lost, Mumbai, Sunny}\}) = \frac{P(\{l,M,S\}|L)\cdot P(L)}{P(\{l,M,S\})}$$

$$= \frac{1/\cancel{7}\cdot\frac{3}{7}}{1/\cancel{7}}$$

$$= \frac{3}{7} = 0.428 > P(W|\{l,M,S\})$$

So, CSK will lose the match in $\left\{\begin{array}{l}\text{toss = lost}\\ \text{Venue : Mumbai}\\ \text{Weather = Sunny}\end{array}\right.$ $\rightarrow$ this condition