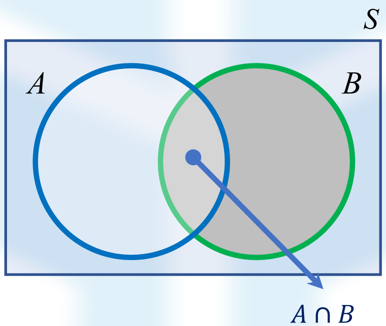
Machine Learning

Review of Probability



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

Dr. *Ali Valinejad*

valinejad.ir
valinejad@umz.ac.ir

What Is Probability?

❖ Definition:

- *Probability* is the study of chance or the likelihood of an event happening. Directly or indirectly, probability plays a role in all activities.
- *Probability* is a measurement of the chance that some event is likely to happen.

For example, we may say that there is a 70% chance of rain today (because most of the days we have observed were rainy days. However, in mathematics, we would require a more accurate way of measuring probability).

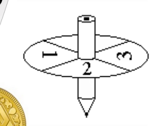
Terminology

❖ Experiment:

- An ***experiment*** or ***trial*** is any procedure that can be infinitely repeated and has a well-defined set of possible *outcomes*, known as the sample space.
- An ***experiment*** is said to be random if it has more than one possible *outcome*, and deterministic if it has only one.

Examples:

Throwing a die, Rotating a spinner and Tossing a coin.

❖ Outcome:

- An ***outcome*** is a possible result of a random *experiment*.

Examples:

Throwing a die has six possible outcomes 1, 2, 3, 4, 5 and 6.
Tossing a coin has two possible outcomes, *head* or *tail*.

❖ Event:

An ***event*** is a *single* result of an experiment. An ***event*** is a set of outcomes to which a probability is assigned.

Examples:

One possible event in throwing *a die* is "rolling a number less than 3".

❖ Sample space:

A ***sample space*** is the set of all possible *outcomes* in the *experiment*. It is usually denoted by the letter S. Sample space can be written using the set notation, { }.

Examples:

Sample space of throwing a die: S={1, 2, 3, 4, 5, 6}.
Sample space of tossing a coin: S={*head, tail*}

Terminology

Experimental Probability vs. Theoretical Probability

Experimental Probability is found by repeating an experiment and observing the outcomes.

$$P(\text{event}) = \frac{\text{number of times event occurs}}{\text{total number of trials}}$$

Example:

A coin is tossed 10 times:
A head is recorded 7 times and a tail 3 times.

$$P(\text{head}) = \frac{7}{10}$$

$$P(\text{tail}) = \frac{3}{10}$$

Theoretical Probability is what is expected to happen based on mathematics

$$P(\text{event}) = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}}$$

Example:

A coin is tossed.

$$P(\text{head}) = \frac{1}{2}$$

$$P(\text{tail}) = \frac{1}{2}$$

Terminology

❖ How To Find The Experimental Probability Of An Event?

Step 1: Conduct an experiment and record the number of times the event occurs and the number of times the activity is performed.

Step 2: Divide the two numbers to obtain the Experimental Probability.

❖ What Is The Experimental Probability Formula?

The formula for theoretical probability of an event is

$$P(event) = \frac{Number\ of\ times\ event\ occurs}{Total\ number\ trials}$$

Example:

A bag contains 10 red marbles, 8 blue marbles and 2 yellow marbles. Find the experimental probability of getting a blue marble.

Solution:

1. Take a marble from the bag.
2. Record the color and return the marble.
3. Repeat a few times (maybe 10 times).
4. Count the number of times a blue marble was picked (Suppose it is 6)
5. The experimental probability of getting a blue marble from the bag is $\frac{6}{10} = 0.6$

Terminology

❖ How To Find The Theoretical Probability Of An Event?

The Theoretical Probability of an event is the number of ways the event can occur (favorable outcomes) divided by the number of total outcomes.

❖ What Is The Theoretical Probability Formula?

The formula for theoretical probability of an event is

$$P(event) = \frac{Number\ of\ favorable\ outcomes}{Number\ of\ total\ outcomes}$$

Example:

A bag contains 10 red marbles, 8 blue marbles and 2 yellow marbles. Find the theoretical probability of getting a blue marble.

Solution:

1. There are 8 blue marbles. Therefore, the number of favorable outcomes = 8
2. There are a total of 20 marbles. Therefore, the number of total outcomes = 20

$$P(event) = \frac{Number\ of\ favorable\ outcomes}{Number\ of\ total\ outcomes} = \frac{8}{20} = \frac{4}{10} = 0.4$$

Terminology

❖ What Is a Random Variable?

- A ***random variable*** is a *variable* whose value is unknown or a function that assigns values to each of an experiment's outcomes.
- Random variables are often designated by letters and can be classified as ***discrete***, which are variables that have specific values, or ***continuous***, which are variables that can have any values within a continuous range.

Examples:

A typical example of a random variable is the *outcome of a coin toss*.
If random variable, Y, is ***the number of heads we get from tossing two coins***, then
Y could be 0, 1, or 2.
This means that we could have no heads, one head, or both heads on a two-coin toss.

However, the two coins land in four different ways: TT, HT, TH, and HH. Therefore,
the P(Y=0) = 1/4=0.25.
since we have one chance of getting no heads (i.e., two tails [TT] when the coins are tossed).

Similarly, the probability of getting two heads (HH) is also ¼=0.25.

Notice that getting one head has a likelihood of occurring twice: in HT and TH. In this case, P (Y=1) = 2/4 = 0.5.

Terminology

Independent Events

The outcome of one event does not affect the outcome of the other.

If A and B are independent events then the probability of both occurring is

$$P(A \text{ and } B) = P(A) \times P(B)$$

Dependent Events

The outcome of one event affects the outcome of the other.

If A and B are dependent events then the probability of both occurring is

$$P(A \text{ and } B) = P(A) \times P(B | A)$$

Probability of B given A

Terminology

Independent events

Example:

A coin is tossed twice, what is the probability of getting a head after a tail, $P(head\ after\ a\ tail)$.

Solution:

Since the first toss *doesn't affect* the second toss, the events are independent. So

$$P(first\ toss = tail) = \frac{1}{2} \quad \text{and} \quad P(second\ toss = head) = \frac{1}{2}$$

The probability of tossing a head after tossing a tail is

$$P(tail, head\ after\ tail) = P(tail) \times P(head\ after\ tail) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

H: head

T: Tail

S={HH,HT,TH,TT} ➔ $P(tail, head\ after\ tail) = P(TH) = \frac{1}{4}$

Terminology

Dependent events

Example:

A bag contains 6 red, 5 blue and 4 yellow marbles. Two marbles are drawn, but the first marble drawn is not replaced. Find $P(blue\ after\ red)$

Solution:

There are a total of 15 marbles which 6 of them are red. So $P(red) = \frac{6}{15}$.

The result of the first draw *affected* the probability of the second draw.

In second drawn, there are a total of 14 marbles left which 5 of them are blue marbles. So

$$P(blue\ after\ red) = \frac{5}{14}$$

The probability of drawing a red marble and then a blue marble is

$$P(red, then\ belue) = P(red) \times P(blue\ after\ red) = \frac{6}{15} \times \frac{5}{14} = \frac{1}{7}$$

Terminology

❖ What Is Joint Probabilities:

The ***joint probability*** is the probability of some different events occurring at the ***same time***.

❖ What Is Marginal Probabilities:

Assume two random variables X and Y. The probability of one event in the presence of all (or a subset of) outcomes of the other random variable is called the ***marginal probability*** or the ***marginal distribution***.

$$P(X=\mathrm{d}) = \sum_{y \in S_Y} P(X = d, Y = y)$$

S_Y represents all the possible values of the random variable Y.

Examples:

x_2=Type

x_1=**Color**	Sports	SUV	
Red	0.2	0.4	0.6
Yellow	0.1	0.3	0.4
	0.3	0.7	1.0

Joint Probabilities:

$P(x_1\text{=Red}, x_2\text{=Sport}) = 0.2$

$P(x_1\text{=Yellow}, x_2\text{=SUV}) = 0.3$

$P(x_1\text{=Red}, x_2\text{=SUV}) = 0.4$

$P(x_1\text{=Yellow}, x_2\text{=Sport}) = 0.1$

Marginal Probabilities:

$P(x_1\text{=Red}) = P(x_1\text{=Red}, x_2\text{=Sport}) + P(x_1\text{=Red}, x_2\text{=}SUV) = 0.2 + 0.4 = 0.6$

$P(x_1\text{=Yellow}) = 0.1 + 0.3 = 0.4$

$P(x_2\text{=SUV}) = 0.4 + 0.3 = 0.7$

$P(x_2\text{=Sport}) = 0.2 + 0.1 = 0.3$

Terminology

❖ What Is Conditional Probability?

The probability of an event occurring given that another event has already occurred is called a conditional probability.

Recall:

When two events, A and B, are *dependent*, the probability of both occurring is:

$P(A\ and\ B) = P(A) \times P(B\ given\ A)$

or $P(A\ and\ B) = P(A) \times P(B|A)$

Conditional Probability Formula:

$$\underbrace{P(B|A)}_{\text{Probability of } B\ given\ A} = \frac{\overbrace{P(A, B)}^{\text{Probability of } A\ and\ B}}{\underbrace{P(A)}_{\text{Probability of } A}}$$

Example:

Sara took two tests. The probability of her passing both tests is 0.6. The probability of her passing the first test is 0.8. What is the probability of her passing the second test given that she has passed the first test?

Solution:

$$P(second|first) = \frac{P(second\ and\ first)}{P(first)} = \frac{0.6}{0.8} = 0.75$$

Terminology

Conditional Probability Formula:

$$P(B=b|A=a)=\frac{P(A=a\,,B=b)}{P(A=a)}=\frac{P(A=a\,,B=b)}{\sum_{t\in S_B}P(A=a,B=t)}$$

Examples:

x_2=**Type**

x_1=**Color**	Sports	SUV	
Red	0.2	0.4	0.6
Yellow	0.1	0.3	0.4
	0.3	0.7	1.0

Joint Probabilities:

$P(x_1\text{=Red}, x_2\text{=Sport}) = 0.2$

$P(x_1\text{=Yellow}, x_2\text{=SUV}) = 0.3$

$P(x_1\text{=Red}, x_2\text{=SUV}) = 0.4$

$P(x_1\text{=Yellow}, x_2\text{=Sport}) = 0.1$

Marginal Probabilities:

$P(x_1\text{=Red}) = 0.6$

$P(x_1\text{=Yellow}) = 0.4$

$P(x_2\text{=SUV}) = 0.7$

$P(x_2\text{=Sport}) = 0.3$

$$P(x_1\text{=Red} \mid x_2\text{=Sport}) = \frac{P(x_1\text{=Red}\,,\,x_2\text{=Sport})}{\sum_{t\in\{\text{Sport,Suv}\}}P(x_1\text{=Red}\,,\,x_2\text{=}t)} = \frac{P(x_1\text{=Red}\,,\,x_2\text{=Sport})}{P(x_1\text{=Red}, x_2\text{=Suv})+P(x_1\text{=Red}, x_2\text{=Sport})} = \frac{0.2}{0.4+0.2} = 0.\overline{3}$$

References

1- http://www.pitt.edu/~super4/34011-35001/34301.ppt.
2- https://www.mathsteacher.com.au/year10/ch05_probability/01_long_run/prob.htm#trial
3- https://www.onlinemathlearning.com/dependent-events.html
4- https://www.onlinemathlearning.com/theoretical-probability.html
5- https://slideplayer.com/slide/7467928/
6- https://www.onlinemathlearning.com/samples-in-probability.html
7- https://www.investopedia.com/terms/r/random-variable.asp
8- https://tinyheero.github.io/2016/03/20/basic-prob.html
9- http://pages.stat.wisc.edu/~ifischer/Intro_Stat/Lecture_Notes/4_-_Classical_Probability_Distributions/4.1_-_Discrete_Models.pdf