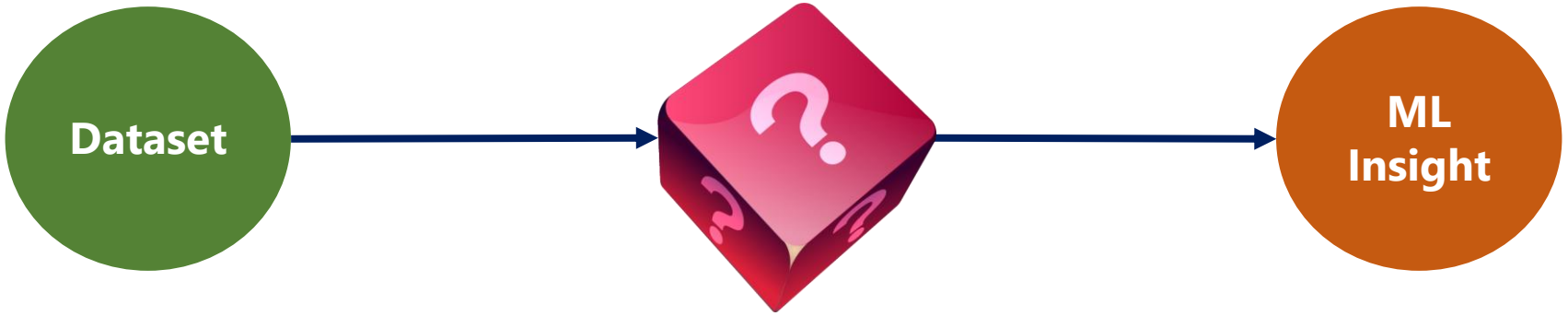
365 DataScience

PROBABILITY FOR STATISTICS AND DATA SCIENCE

Introduction to Probability: Cheat Sheet

Probability Formula | Sample Space | Expected Values | Complements

Words of welcome



You are here because you want to comprehend the basics of probability before you can dive into the world of statistics and machine learning. Understanding the driving forces behind key statistical features is crucial to reaching your goal of mastering data science. This way you will be able to extract important insight when analysing data through supervised machine learning methods like regressions, but also fathom the outputs unsupervised or assisted ML give you.

Bayesian Inference is a key component heavily used in many fields of mathematics to succinctly express complicated statements. Through Bayesian Notation we can convey the relationships between elements, sets and events. Understanding these new concepts will aid you in interpreting the mathematical intuition behind sophisticated data analytics methods.

Distributions are the main way we lie to classify sets of data. If a dataset complies with certain characteristics, we can usually attribute the likelihood of its values to a specific distribution. Since many of these distributions have elegant relationships between certain outcomes and their probabilities of occurring, knowing key features of our data is extremely convenient and useful.

What is probability?

Probability is **the likelihood of an event occurring**. This event can be pretty much anything – getting heads, rolling a 4 or even bench pressing 225lbs. We measure probability with numeric values between 0 and 1, because we like to *compare* the relative likelihood of events. Observe the general probability formula.

$$P(X)=\frac{Preferred\ outcomes}{Sample\ Space}$$

Probability Formula:

- The Probability of event X occurring equals the *number* of preferred outcomes over the *number* of outcomes in the sample space.
- Preferred outcomes are the outcomes we want to occur or the outcomes we are interested in. We also call refer to such outcomes as "Favorable".
- Sample space refers to all possible outcomes that can occur. Its "size" indicates the amount of elements in it.

If two events are independent:

The probability of them occurring simultaneously equals the product of them occurring on their own.

P(A♥) = P(A) . P(♥)

365 DataScience

Expected Values

Trial – Observing an event occur and recording the outcome.

Experiment – A collection of one or multiple trials.

Experimental Probability – The probability we assign an event, based on an experiment we conduct.

Expected value – the specific outcome we expect to occur when we run an experiment.

Example: Trial

Flipping a coin and recording the outcome.

Example: Experiment

Flipping a coin 20 times and recording the 20 individual outcomes.

In this instance, the **experimental probability** for getting heads would equal the number of heads we record over the course of the 20 outcomes, over 20 (the total number of trials).

The **expected value** can be numerical, Boolean, categorical or other, depending on the type of the event we are interested in. For instance, the expected value of the trial would be the more likely of the two outcomes, whereas the expected value of the experiment will be the number of time we expect to get either heads or tails after the 20 trials.

Expected value for **categorical** variables.

$$E(X) = n \times p$$

Expected value for **numeric** variables.

$$E(X) = \sum_{i=1}^{n} x_i \times p_i$$

Probability Frequency Distribution

What is a probability frequency distribution?:
A collection of the probabilities for each possible outcome of an event.

Why do we need frequency distributions?:
We need the probability frequency distribution to try and predict future events when the expected value is unattainable.

What is a frequency?:
Frequency is the number of times a given value or outcome appears in the sample space.

What is a frequency distribution table?:
The frequency distribution **table** is a table matching each distinct outcome in the sample space to its associated frequency.

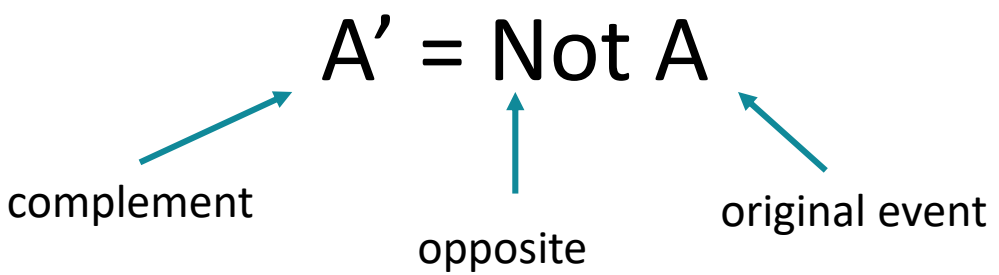
How do we obtain the probability frequency distribution from the frequency distribution table?:
By dividing every frequency by the size of the sample space. (Think about the "favoured over all" formula.)

Sum	Frequency	Probability
2	1	1/36
3	2	1/18
4	3	1/12
5	4	1/9
6	5	5/36
7	6	1/6
8	5	5/36
9	4	1/9
10	3	1/12
11	2	1/18
12	1	1/36

365 DataScience

Complements

The complement of an event is **everything** an event is **not.** We denote the complement of an event with an apostrophe.



Characteristics of complements:

- Can never occur simultaneously.
- Add up to the sample space. **(A + A’ = Sample space)**
- Their probabilities add up to 1. **(P(A) + P(A’) = 1)**
- The complement of a complement is the original event. **((A’)’ = A)**

Example:

- Assume event A represents drawing a spade, so P(A) = 0.25.
- Then, A’ represents **not** drawing a spade, so drawing a club, a diamond or a heart. P(A’) = 1 – P(A), so P(A’) = 0.75.

365 DataScience