Contents

1 Combinatorics

yeah, somebody should add it here :/.

2 Probability

2.1 Introduction

Probability is a fundamental concept in statistics and machine learning. Understanding the basics of probability is crucial for mastering data science, as it enables the extraction of important insights from data. Bayesian Inference is a key component heavily used in many fields of mathematics to succinctly express complicated statements. Through Bayesian Notation, relationships between elements, sets, and events can be conveyed. This understanding can aid in interpreting the mathematical intuition behind sophisticated data analytics methods.
Distributions are the main way to classify sets of data. If a dataset complies with certain characteristics, the likelihood of its values can be attributed to a specific distribution. Many of these distributions have elegant relationships between certain outcomes and their probabilities of occurring, making it extremely convenient and useful to know key features of our data.
Overall, understanding the basics of probability and distributions is crucial for anyone looking to dive into the world of statistics and machine learning, as it enables the extraction of meaningful insights from data.

2.1.1 What is probability?

Probability is the measure of the likelihood of an event occurring, and it can range from 0 to 1. The probability formula states that 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, while the sample space refers to all possible outcomes that can occur.

$$P(X) = \frac{\text{preferred outcomes}}{\text{sample space}}$$

If two events are independent, the probability of them occurring simultaneously equals the product of them occurring on their own. This concept is useful in many areas, including statistics and machine learning, where probabilities are used to make predictions and inferences.

2.1.2 expected values

Trail - Ovserving an event occur and recording the outcome.
Experiment - A collectoin of one or multiple trails.
Experimental Probability - The probability we assign an event, based on a experiment we conduct.
Expected Value - The specific outcome we expect to occur when we run an experiment.

Example: Trail
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 trails).
The **expected value** can be numerical, Boolearn, categorical or other depending on the type of the events we are interested in. For instance, the expected value of the trail would be the more likely of the two outcomes, whereas the expected value of the experiment will be the number of time we expect get either heads or tails after 20 trails.
Expected value for **categorical** variables.

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

Expected value for **numeric** variables.

$$E(X) = \sum_{i=1}^{n} x_1 \cdot p_i$$

2.1.3 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	$\frac{1}{36}$
3	2	$\frac{1}{18}$
4	3	$\frac{1}{12}$
5	4	$\frac{1}{9}$
6	5	$\frac{5}{36}$
7	6	$\frac{1}{6}$
8	5	$\frac{5}{36}$
9	4	$\frac{1}{9}$
10	3	$\frac{1}{12}$
11	2	$\frac{1}{18}$
12	1	$\frac{1}{36}$

2.1.4 Complements

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

$$A\prime = \text{not}A$$

Where $A\prime$ is an complement, not the opposite and A The original event.
characteristics of complements:

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

Example:

- Assume event A represents drawing a spade, so P(A) = 0.25.
- Then, $A\prime$ represents **not** drawing a spade, so drawing a club, a diamond or a heart. $P(A\prime) = 1 - P(A),\ \text{so}P(A\prime) = 0.75$.

2.2 Bayesian thingies

A **set** is a collection of elements, which hold vertain values. Additionally, every event has a set of outcomes that satisfy it. The **null-set** (or empty set), denoted $\emptyset$, is an set which contain no values.
An element is denoted in lower-case, e.g. x.
A set is written with upper-case, e.g. A.
To make sense out of it, it can be used the following ways:

Notation	**Interpretation**	**Example**
$x \in A$	Element x is part of set A	$2 \in$ All even numbers
$A \ni x$	Set A contains element x	All even numbers $\ni 2$
$x \notin A$	Element x is not part of set A	$1 \notin$ All even numbers
$\forall x :$	For all x such that...	$\forall x : x \in$ All even numbers
$A \subseteq B$	Set A is a subset of set B	Even numbers $\subseteq$ Integers

Table 1: Notations and interpretations for set theory

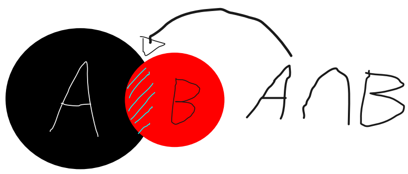
Remember every set has at least two subsets:

- $A \subseteq A$
- $\emptyset \subseteq A$

There are three different types of multible events for now, one can imagine them as circles which are:

1. Not touching at all.
2. Intersect (partially overlap).
3. One completely overlaps another.

Intersection We denote the intersection of two sets with the "intersect" sign, which resembles an upside down capital letter U: $A \cap B$.



Union The union is two or more events expresses the set of outcome that satify at least one of the events. Graphically, is the area that includes both sets. Noted: $A \cup B$, $A \cup B = A + B - A \cap B$.

Mutually exclusive Sets Sets with no overlapping elements are called mutually exclusive. Graphically, their circles never touch.
If $A \cap B = \emptyset$, then the two sets are mutually exclusive. Remember: All complements are mutually exclusive, but not all mutually exclusive sets are complements.
Example: Dogs and cats are mutually exclusive sets, since no species is simultaneously a feline and a canine, but the two are not complements, since there exist other types of animals as well.

2.2.1 Independent and Dependent Events

If the likelihood of event A occurring (P(A)) is affected event B occurring, then we say that A and B are **dependent** events. Alternatively, if it isn't – the two events are **independent**.
We express the probability of event A occurring, given event B has occurred the following way **P(A|B)**. We call this the conditional probability.
Independent:

- All the probabilities we have examined so far.
- The outcome of A does not depend on the outcome of B.
- $P(A|B) = P(A)$
- **Example:**
- A -> Hearts
- B -> Jacks

Dependent:

- New concept
- The outcome of A depends on the outcome of B.
- $P(A|B) \neq P(A)$

- **Example:**
- A -> Hearts
- B -> Red

2.2.2 Conditional Probability

For any two events A and B, such that the likelihood of B occuring is greater than 0 (P(B)>0), the conditional probability formula states the following. Also writable as:

$$\forall A, B \quad P(B) > 0 : \quad P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Intuition behind the formula:

- Only interest in the outcome where B is satisfied.
- Oly the elements in the intersectino would satisfly A as well.
- Paralll to the "favored over all" formula.

  Intersectino = "preferred outcomes"

  B = "sample space"

Remember:

- Unlike the union or the intersection, changing the order of A and B in the conditional probability alters its meaning.
- P(A|B) is not the same as P(B|A), even if P(A|B)=P(B|A) numerically.
- the two conditional probabilities posess **different meanings** even if they have equal values.

2.2.3 Law of total probability

The **law of total probability** dictates that for any set A, which is a union of many mutually exclusive sets $B_1, B_2, ..., B_n$ its probability equals the following sum.

$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \cdots + P(A|B_n) \cdot P(B_n)$$

Meaning of the opperators:
$P(A|B_1)$= Conditional probability of A, given B_1 has occurred.
$P(B_1)$ = Probability of B_1 occurring.
$P(A|B_2)$= Conditional of A, given B_2 has occurred.
$P(B_2)$ = Probability of B_2 occuring.
Intution behind the formula:

- Since P(A) is the union of mutually exclusive sets, so it equals to the sum of the individual sets.
- The **intersection** of a union and one of its subsets is the entire subset.
- We can rewrite the conditional probability formula

  $P(A|B) = \frac{P(A \cap B)}{P(B)}$ to get $P(A \cap B) \cdot P(B)$.
- Another way to express the law of probability is:

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_n)$$

2.2.4 Additive Law

The additive law calculates the probability of the union based on the probability of the individual sets it accounts for.

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

Meaning of the operators:
$P(A \cup B)$= Probability of the union.
$P(A \cap B)$= Probability of the intersection.
Intution behind the formula:

- Recall the formula for finding the size of the uunion using the size of the intersction:
  $A \cup B = A + B - A \cap B$
- The probability of each oen is simply its size over the size of the sample space.
- this holds true for any events A and B.

2.2.5 The Multiplication Rule

The multiplication rule calculates the probability of the intersection based on the conditional probability.

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

meaning of the operators:
$P(A \cap B)$= Probability of the intersection
$P(A|B)$= Conditional Probability
$P(B)$= Probability of event B
Intuition behind the formula:

- We can multiply both sides of the conditional probability formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$ by $P(B)$ to get $P(A \cap B) = P(A|B) \cdot P(B)$.
- If event B occurs in 40% of the time ($P(B) = 0.4$) and event A occurs in 50% of the time of the time mB occurs ($P(A|B) = 0.5$), then they would simantaneously occur 20% of the time ($P(A|B) \cdot P(B) = 0,5 \cdot 0,4 = 0,2$).

2.2.6 Bayes' Law

Bayes' Law helps us understand the relationship between two events by computing the different conditional probabilities. We also call it Bayes' Rue or Bayes' Theorem.

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

Intuition behind the formula

- According to the multiplication rule $P(A \cap B) = P(A|B) \cdot P(B)$, so $P(B \cap A) = P(B|A) \cdot P(A)$.
- Since $P(A \cap B) = P(B \cap A)$, we plug in $P(B|A) \cdot P(A)$ for $P(A \cap B)$ in the conditional probability formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Bayes' Law is often used in medical or business analysis to determine which of two symption affects the other one more.