Probability Theory

Olivér Kiss

Mathematics and Informatics Pre-session for Business Analytics

2023

Topics

- Random Experiments
- Probability distributions

Random experiments

Probability theory is based on random experiments, that are experiments whose outcome cannot be determined in advance. Examples:

- Tossing a coin (heads/tails)
- The number we get when casting a dice twice
- The number of left-handers in class
- The number of service disruptions on a metro line in a month
- The monthly rainfall in September in Vienna

We can still, to some extent, analyze these random experiments. We will build a proper framework for it.

Sample space

Although we don't know the outcome, we often know the set of possible outcomes. This is called the sample space, and is denoted by Ω. Examples:

- Tossing a coin: $\Omega = \{H, T\}$
- Cast a dice twice: $\Omega = \{(1,1),(1,2)\dots(6,6)\}$
- The number of left-handers in class: $\Omega = \{1,2,3,4\dots N\}$
- The number of service disruptions on metro line M3 in a month: $\Omega = \mathbb{Z}_0^+$
- The monthly rainfall in September in Budapest: $\Omega = \mathbb{R}_0^+$

Events

An event $\mathcal{A}$ is a pre-defined set of events. If the outcome of the random experiment falls into this set, then event $\mathcal{A}$ occurs. Examples:

- Sum of two dice is 10: $\mathcal{A} = \{(6,4),(4,6),(5,5)\}$
- The number of left-handers in class is even: $\mathcal{A} = \{2,4,6\ldots\}$
- The monthly rainfall in September in Budapest is less than 15 mm: $\mathcal{A} = [0,15)$

After observing the outcome of the random experiment, we can always decide whether an event occured or not.

Events

Since events are sets, they can be combined like sets. Let's have two events, $\mathcal{A}$ and $\mathcal{B}$.

- $\mathcal{A} \cup \mathcal{B}$ occurs if $\mathcal{A}$ or $\mathcal{B}$ or both occur
- $\mathcal{A} \cap \mathcal{B}$ occurs if both $\mathcal{A}$ and $\mathcal{B}$ occur
- $\mathcal{A}^c$ occurs if $\mathcal{A}$ does not occur

Quick terminology: If $\mathcal{A} \cap \mathcal{B} = \emptyset$, then $\mathcal{A}$ and $\mathcal{B}$ are disjoint events.

Probability

Probability is a number between 0 and 1 telling us how likely it is that certain events happen. Probabilities have to satisfy:

- $\mathbb{P}(\mathcal{A}) \geq 0$
- $\mathbb{P}(\Omega) = 1$
- For disjoint events: $\mathbb{P}\left(\cup_i \mathcal{A}_i\right) = \sum_i \mathbb{P}(\mathcal{A}_i)$

It is not necessarily easy to find probabilities. It is the easiest if the sample space is finite. Then if we can calculate the probabilities of elementary events, we can combine them to get the probabilities of events.

Probability

Example: We toss a coin three times. What is the probability of getting Tails twice?

- We can define the sample space:

$$\Omega = \{(H, H, H), (H, H, T), (H, T, H), (T, H, H),\\ (H, T, T), (T, H, T), (T, T, H), (T, T, T)\}$$

- These elementary events are equally likely and disjoint.
- Since $\mathbb{P}(\Omega) = 1$, and for disjoint events $\mathbb{P}\left(\cup_i \mathcal{A}_i\right) = \sum_i \mathbb{P}(\mathcal{A}_i)$, we know that with 8 elementary events each has a probability of $\frac{1}{8}$.
- The event of getting Tails twice is: $\mathcal{A} = \{(H, T, T), (T, H, T), (T, T, H)\}$.
- Since these are disjoint events:
$\mathbb{P}(\mathcal{A}) = \mathbb{P}((H, T, T)) + \mathbb{P}((T, H, T)) + \mathbb{P}((T, T, H)) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$

Solve the following problems

Define the sample space for the following random experiments and calculate the probability of an event:

- We cast a dice twice. What is the probability of having 6 at least once?
- We toss a coin four times. What is the probability of getting four heads or four tails?

Counting

- It is common in the previous examples that we had a sample space with equally likely elementary events, and counted the number of elementary events in event $\mathcal{A}$.
- This counting is not necessarily easy either.
- You have 10 multiple choice questions with 4 options each. How many ways are there to fill this test?
- Imagine a car race with 20 cars. How many combinations of 1st, 2nd and 3rd places are there?
- The CEU library has 200 000 books, you want to borrow 5. How many options do you have?
- You throw 5 dice simultaneously. How many different results can you have?
- Luckily these mostly fall into a few simple categories.

The Urn

- A common example in probability theory is the urn with n different balls, out of which you select k. There are two important aspects:
  - Do you replace the balls?
  - Do you care about their order?
- Most counting problems are equivalent to these urn problems, and thus you will be able to calculate the number of elementary events easily.

Drawing without replacement, ordered

- Also called a permutation of size k from set $\{1, 2, 3 \dots n\}$
- Equivalent to the problem of the car race. No one can take both the 1st and the 2nd place, and order definitely matters.
- The first place can be taken by n balls/racers
- Once the first place is taken, the second can be taken by $n - 1$ balls/racers etc.
- Thus if you select k balls, the number of permutations is:

$$\#P_k^n = n(n-1)(n-2)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

Drawing without replacement, unordered

- Also called a combination of size k from set $\{1, 2, 3 \ldots n\}$
- Equivalent to the problem of the library. You can't borrow the same book twice, and there is no ordering of borrowed books.
- Let's start from the permutation case. Imagine that order actually mattered. If we borrow the same k books, they can be borrowed in $k!$ orders.
- Now since order doesn't matter, P_k^n counted all combinations $k!$ times
- Thus the number of combinations is:

$$\# C_k^n = \frac{\# P_k^n}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

Drawing with replacement, ordered

- Equivalent to the problem of the multiple choice test. You can choose A, B, C or D for multiple questions, and order definitely matters.
- You have n options at each question
- You answer k questions
- For the first question you have n options. Then for the second you have n options again. Therefore for 2 questions you have n^2 answer options. For 3, you have n^3.
- For k questions you have altogether n^k options.

Drawing with replacement, unordered

- This is the equivalent of the dice throwing example. You can't distinguish the dice and you can easily end up with multiple 1s, 2s, 3s etc.
- The easiest way to understand this case is that you basically have 5 balls (throws) that you want to distribute between 6 urns (numbers).
- You need $n-1$ separators and k balls, out of which the balls will occupy k positions.
- Altogether there are:

$$\binom{k+n-1}{k}$$

options

Solve the following problems

- If we throw 3 dice simultaneously, what is the probability of ending up with three 6s?
- If you have 20 CDs and you want to select 3 of them, how many options do you have?
- In a multiple choice test with 20 questions and 4 options a student answers randomly. What is the probability of a perfect score?

Conditional probability

- How do probabilities change when we know some event $\mathcal{B} \subset \Omega$ has occurred?
- If we are looking for the probability of $\mathcal{A}$ also occuring, we should find the relative probability of $\mathcal{A} \cap \mathcal{B}$ compared to $\mathcal{B}$
- This is the so-called conditional probability, the probability of $\mathcal{A}$ happening given that we know that $\mathcal{B}$ occured

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

Solve the following problems

- What is the probability of throwing two 3s with two dice, if we know that the first throw is a 3?
- What is the probability of throwing two 3s with two dice, if we know that the first throw is a 2?
- Look up the Monty Hall problem at home!

Law of total probability

- Assume that you can partition the sample space to n disjoint events $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$.
- Also assume that you know the conditional probability of $\mathcal{A}$ given each partition, thus you know $\mathbb{P}(\mathcal{A}|\mathcal{B}_1), \mathbb{P}(\mathcal{A}|\mathcal{B}_2), \ldots, \mathbb{P}(\mathcal{A}|\mathcal{B}_n)$
- You also know the probability of each $\mathcal{B}_i$ event
- Then

$$\mathbb{P}(\mathcal{A}) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B}_i)\mathbb{P}(\mathcal{A}|\mathcal{B}_i)$$

Example

You know that a firm produces the same product in 3 factories. 20% of the production comes from the 1st factory, 45% from the second, and the remaining 35% from the third. You also know that 10% of the products in the 1st factory are faulty. This rate is 15% in the second factory and 5% in the 3rd factory. What is the probability of a randomly selected product being faulty?

$$\mathbb{P}(\mathcal{A}) = 0.2 \cdot 0.1 + 0.45 \cdot 0.15 + 0.35 \cdot 0.05 = 0.105$$

Bayes' rule

- Assume that you can partition the sample space to n disjoint events $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$.
- Also assume that you know the conditional probability of $\mathcal{A}$ given each partition, thus you know $\mathbb{P}(\mathcal{A}|\mathcal{B}_1), \mathbb{P}(\mathcal{A}|\mathcal{B}_2), \ldots, \mathbb{P}(\mathcal{A}|\mathcal{B}_n)$
- You also know the probability of each $\mathcal{B}_i$ event
- Then

$$\mathbb{P}(\mathcal{B}_j|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A}|\mathcal{B}_j)\mathbb{P}(\mathcal{B}_j)}{\sum_{i=1}^{n} \mathbb{P}(\mathcal{B}_i)\mathbb{P}(\mathcal{A}|\mathcal{B}_i)}$$

Example

You know that a firm produces the same product in 3 factories. 20% of the production comes from the 1st factory, 45% from the second, and the remaining 35% from the third. You also know that 10% of the products in the 1st factory are faulty. This rate is 15% in the second factory and 5% in the 3rd factory. What is the probability that a randomly selected faulty product came from the first factory?

$$\mathbb{P}(\mathcal{B}_\infty|\mathcal{A}) = \frac{0.2 \cdot 0.1}{0.2 \cdot 0.1 + 0.45 \cdot 0.15 + 0.35 \cdot 0.05} = \frac{0.02}{0.105}$$

Solve the following problems

We know that in 2012 there were altogether 166960 criminal cases registered in Central Hungary. Out of these, 118600 were registered in Budapest, the rest in the Budapest Metropolitan Area. In Budapest, 3769 of these registered cases were crimes against the person. The same number for the BMA is 2059.

- What is the probability that a randomly selected criminal case in Central Hungary is a crime against the person?
- What is the probability that a crime against the person in Central Hungary was registered in Budapest?

Random variables

If the outcome of a random experiment is a number, it is called a random variable. Random variables can be

- Discrete: if the sample space is finite or countably infinite. Examples:
  - The result of a throw with a fair dice
  - The number of days until a species goes extinct
- Continuous: If the sample space is uncountable. Examples:
  - The exact quantity of water in a bottle
  - The height of a randomly selected individual in the classroom

Discrete random variables - pmf

Discrete random variables have a probability mass function, which shows the probability of each elementary event in the sample space. For example, let's draw the pmf of:

- The result of a throw with a fair dice
- The sum of two dice

The pmf is basically a function that associates each elementary event with it's probability. If we denote our random variable by x, then in the first case the pmf is $\mathbb{P}(x) = \frac{1}{6} \quad \forall x \in \{1, 2, 3, 4, 5, 6\}$

Discrete random variables - cdf

Discrete random variables have a cumulative density function, which shows the probability that the outcome of an experiment is smaller than or equal to X. For example, let's draw the cdf of:

- The result of a throw with a fair dice
- The sum of two dice

The cdf basically shows the following for a random variable x:

$$F(X) = \mathbb{P}(x < X)$$

Discrete random variables - Expected value

Discrete random variables have an expected value, which is by the Law of Large Numbers the average of the results obtained from a large number of trials. It can be calculated by:

$$E(x) = \sum x\mathbb{P}(x)$$

Example: The expected value of a throw with a fair dice:

$$E(x) = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3.5$$

Discrete random variables - Specific types

There are a few types of discrete random variables that we use quite often:

- Uniform random variable
- Binomial random variable
- Geometric random variable
- Bernoulli random variable

Let's check these one by one.

Uniform random variable

This distribution is characterized by equally likely elementary events. For example: throw a fair dice.

- The p.m.f. is

$$\mathbb{P}(x) = \frac{1}{n}$$

  where n is the number of elementary events.

- The expected value is

$$E(x) = \frac{\sum x}{n}$$

Bernoulli random variable

This random variable takes the value of 1 with probability p, and takes the value of 0 with probability $1 - p$.

- The p.m.f. is
$$\mathbb{P}(x) = p^x(1-p)^{1-x}$$
- The expected value is
$$E(x) = p$$

Binomial random variable

Assume that you run a random experiment n times, and each time the probability of an event happening is p. Then the probability that this event happens exactly x times out of n experiments follows a binomial distribution.

- The p.m.f. is

$$\mathbb{P}(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- The expected value is

$$E(x) = np$$

Geometric random variable

Assume that you run a random experiment where the probability of an event happening is p. You repeat this experiment until the event happens. Then the probability of repeating this experiment exactly x times follows a geometric distribution.

- The p.m.f. is

$$\mathbb{P}(x) = (1-p)^{x-1}p$$

- The expected value is

$$E(x) = \frac{1}{p}$$

Solve the following problems

1. Assume that you keep throwing a dice until you get a 6. How many times do you have to throw on average?
2. Assume that you keep throwing a dice until you get a 6. What is the probability that you have to throw less than 3 times?
3. Assume that 5% of oranges at a certain store are rotten inside. You buy 20 oranges. What is the probability of having exactly 5 rotten oranges?
4. Assume that 5% of oranges at a certain store are rotten inside. You buy 20 oranges. What is the probability of having less than 3 rotten oranges?