Seminar 3

Conditional probability in classic probability

Let's review classical probability: $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ is the set of equally likely outcomes. Let $B \subset \Omega$ be some non-empty event. Then the probability of event A conditioned on event B is by definition

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

We can expand it as follows:

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

Problem 1

Two dice were rolled and the sum of results is more than 6. Find the probability that the first die result is less or equal to 3.

What are the events A and B here?

- Sum of results is more than 6 is event B (we condition on it).
- The first die results is less or equal to 3 is event A (our target).

$$\mathbb{P}(A|B) = \frac{|A \cap B|}{|B|} = \frac{1+2+3}{6+5+4+\ldots+1} = \frac{6}{21} = \frac{2}{7}$$

Recap of axiomatic definition of probability

A probability space is the following tuple: $(\Omega, \mathcal{F}, \mathbb{P})$.

- **Sample space** Ω
- **Set of events** $\mathcal{F}$
- **Probability measure** $\mathbb{P}$

Set of events is $\mathcal{F} \subset 2^{\Omega}$, such that

1. $\Omega \in \mathcal{F}$
2. If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$ (closed under complement operation)
3. If $A_1, A_2, \ldots \in \mathcal{F}$, then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$ (closed under countable union operation)

Probability measure is $\mathbb{P} : \mathcal{F} \to \mathbb{R}_+$, such that

1. $\mathbb{P}(\Omega) = 1$
2. If $A_1, A_2, \ldots \in \mathcal{F}$ and $A_i \cap A_j = \varnothing$ for $i \neq j$, then $\mathbb{P}\left(\bigcup_{k=0}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$ (σ-additivity)

Conditional probability in axiomatic probability

In axiomatic probability, we will use the same formula as before, and say it's a definition.

We are working with probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we are given event $B \in \mathcal{F}$ such that $\mathbb{P}(B) > 0$. Then the probability of any event $A \in \mathcal{F}$ conditioned on event B is by definition:

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

Let's say our sample space is now B. We can prove that:

- $\mathcal{F}_{\mathcal{B}} = \{\mathcal{A} \cap \mathcal{B} : \mathcal{A} \in \mathcal{F}\}$ is a σ-algebra for B
- $\mathbb{P}_B = \mathbb{P}(A|B)$ is a porbability measure

This means that $(B, \mathcal{F}_{\mathcal{B}}, \mathbb{P}_{\mathcal{B}})$ is a porbability space and it is called **conditional probability space** given B.

Conditional probability in axiomatic probability

- If $A \subset B$, then
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} \geqslant P(A)$$
- If $A \supset B$, then
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

Probability of intersection of events

We can multiply both sides by $\mathbb{P}(B)$ to obtain the probability of the intersection of two events:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) \stackrel{?}{=} \mathbb{P}(B|A)\mathbb{P}(A)$$

Applying this formula n times for events $A_1, \ldots, A_n$ such that $\mathbb{P}(A_2 \cap \ldots \cap A_n) > 0$, we obtain:

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbb{P}\left(A_1 | \bigcap_{k=2}^{n} A_k\right)\mathbb{P}\left(\bigcap_{k=2}^{n} A_k\right) = \ldots = \left(\prod_{m=1}^{n-1} \mathbb{P}\left(A_m | \bigcap_{k=m+1}^{n} A_k\right)\right)\mathbb{P}(A_n)$$

Law of total probability

Let $\{B_n\}$ be a countable set of events, such that

- $\mathbb{P}(B_n) > 0, \forall n$
- $B_i \cap B_j = \varnothing, i \neq j$

Then for any event $A \subset \bigcup_n B_n$ holds the law of total probability:

$$\mathbb{P}(A) = \mathbb{P}\left(A \cap \bigcup_n B_n\right) = \mathbb{P}\left(\bigcup_n (A \cap B_n)\right) = \sum_n \mathbb{P}(A \cap B_n) = \sum_n \mathbb{P}(A|B_n)\mathbb{P}(B_n)$$

Problem 2

A hat contains 100 coins, where 99 are fair but one is double-headed (always landing Heads). A coin is chosen uniformly at random. The chosen coin is flipped 7 times, and it lands Heads all 7 times. Given this information, what is the probability of such event?

Solution

Let A be the event of observing heads 7 times. Let B_1 be the event that we picked a fair coin and B_2 be the event that we picked the double-headed coin.

Do we satisfy the requirements for the law of total probability:

- $\mathbb{P}(B_n) > 0, \forall n$
- $B_i \cap B_j = \varnothing, i \neq j$
- $A \subset \bigcup_n B_n$

Then, $\mathbb{P}(B_1) = 0.99$ and $\mathbb{P}(B_2) = 0.01$. Next, let's compute the probabilities $\mathbb{P}(A|B_1)$ and $\mathbb{P}(A|B_2)$.

For a fair coin $\mathbb{P}(A|B_1) = 0.5^7$, and for double-headed coin $\mathbb{P}(A|B_2) = 1$.

Next, let's apply the law of total probability:

$$P(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) = 0.5^7 \times 0.99 + 1 \times 0.01 = 0.017734375$$

Bayes rule

Let $\{B_n\}$ be a countable set of events, such that

- $\mathbb{P}(B_n) > 0, \forall n$
- $B_i \cap B_j = \varnothing, i \neq j$
- $A \subset \bigcup_n B_n$
- $\mathbb{P}(A) > 0$

Then holds the Bayes rule:

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

Problem 3

A hat contains 100 coins, where 99 are fair but one is double-headed (always landing Heads). A coin is chosen uniformly at random. The chosen coin is flipped 7 times, and it lands Heads all 7 times. Given this information, what is the probability that the chosen coin is double-headed?

Solution

We have estimated $\mathbb{P}(A)$ previously and showed that it is greater than zero. Now we only need to apply the Bayes rule:

$$\mathbb{P}(B_2|A) = \frac{\mathbb{P}(A|B_2)\mathbb{P}(B_2)}{\mathbb{P}(A)} = \frac{1 \times 0.01}{0.5^7 \times 0.99 + 0.01 \times 1} = \frac{0.01}{0.017734375} = 0.563876652$$

Problem 4

According to the CDC (Centers for Disease Control and Prevention), men who smoke are 23 times more likely to develop lung cancer than men who don't smoke. Also according to the CDC, 21.6% of men in the U.S. smoke. What is the probability that a man in the U.S. is a smoker, given that he develops lung cancer?

Solution

Let A the event that a man developes a lung cancer. Let B_1 be the event that a man is a smoker and B_2 be the event that a man is a non-smoker.

Then, we need to estimate $\mathbb{P}(B_1|A)$. We will be using the Bayes rule:

$$\mathbb{P}(B_1|A) = \frac{\mathbb{P}(A|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2)}$$

We will need the following:

- $\mathbb{P}(B_1) = 0.216$ from the formulation, $\mathbb{P}(B_2) = 1 - 0.216$
- Denote $\mathbb{P}(A|B_2) = x$, then $\mathbb{P}(A|B_1) = 23x$

We obtain

$$\mathbb{P}(B_1|A) = \frac{23x \times 0.216}{23x \times 0.216 + x \times (1 - 0.216)} = 0.8636995827538247$$

Problem 5

On the game show Let's Make a Deal, hosted by Monty Hall, a contestant chooses one of three closed doors, two of which have a goat behind them and one of which has a car. Monty, who knows where the car is, then opens one of the two remaining doors. The door he opens always has a goat behind it (he never reveals the car!). If he has a choice, then he picks a door at random with equal probabilities. Monty then offers the contestant the option of switching to the other unopened door. If the contestant's goal is to get the car, should she switch doors?

Solution

Let's find the probability of finding a car if the contestant does not switch his choice.

Let's label

- the door that the contestant originally picked as door 1
- the door that Monty opens as door 2
- the probabilities of a car being behind a certain door k as C_k
- the probability of Monty opening a certain door k as M_k

Then,

$$\mathbb{P}(C_1|M_2) = \frac{\mathbb{P}(M_2|C_1)\mathbb{P}(C_1)}{\mathbb{P}(M_2)} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

This means that the contestant should switch his choice.

Independence of events

Consider event B, such that $0 < \mathbb{P}(B) < 1$. We will say that events A and B are **independent** if $\mathbb{P}(A|B) = \mathbb{P}(A|\overline{B})$, i.e.

$$\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \cap \overline{B})}{\mathbb{P}(\overline{B})} = \frac{\mathbb{P}(A) - \mathbb{P}(A \cap B)}{1 - \mathbb{P}(B)}$$

Cross-multiplication gives:

$$\mathbb{P}(A \cap B) - \mathbb{P}(A \cap B)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A \cap B)\mathbb{P}(B)$$

Tidying up, we obtain

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

Properties of independence

- If $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$, this event is independent of any event including itself
- If A and B are independent, their negations are independent as well:
$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(\overline{B})$$
- If there is a causal relation between the events, e.g. $A \subset B$ and $0 < \mathbb{P}(A) < 1$, they are dependent in probabilistic sense as well: $\mathbb{P}(A \cap B) = \mathbb{P}(A) \neq \mathbb{P}(A)\mathbb{P}(B)$
- The converse to the above is **generally** not true, and probabilistic independence does not yield causal relation

Pairwise and mutual independence

Let $\mathcal{A} \subset \mathcal{F}$ be a family of events.

These events are called **pairwise independent** if for any $i \neq j$ holds:

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

These events are called **mutually independent** if for any finite set $A_1, \ldots, A_n \in \mathcal{A}$ holds:

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \ldots \mathbb{P}(A_n)$$

Problem 6

Three dice are rolled. Let

- A_1 be the event that die 1 and die 2 have the same result
- A_2 be the event that die 2 and die 3 have the same result
- A_3 be the event that die 3 and die 1 have the same result

Are the events $\mathcal{A} = \{A_1, A_2, A_3\}$ pairwise independent? Are they mutually independent?

Solution

The probabilities of the events are all equal $\mathbb{P}(A_i) = \frac{1}{6}$.

1. $\mathbb{P}(A_1 \cap A_2) = \left(\frac{1}{6}\right)^2 = \mathbb{P}(A_1)\mathbb{P}(A_2)$
2. $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \left(\frac{1}{6}\right)^2 \neq \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)$ because of symmetry

Problem 7

Let $\mathbb{P}(A) > 0$ and events $B, C \subset A$ are independent in the conditional probability space given A. Are B and C independent in the original probability space?

Solution

We know that $\mathbb{P}_A(B \cap C) = \mathbb{P}_A(B)\mathbb{P}_A(C)$. The definition of conditional probability measure is:

$$\mathbb{P}_A(B \cap C) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(A)}$$

Since $B, C \subset A$, we can get rid of intersections with A:

$$\frac{\mathbb{P}(B \cap C)}{\mathbb{P}(A)}$$

On the other hand, we have (and again getting rid of intersections with A)

$$\mathbb{P}_A(B)\mathbb{P}_A(C) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B)\mathbb{P}(C)}{(\mathbb{P}(A))^2}$$

We can see that

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

So they are not independent in original probability space.