Probability theory

Lecture 2: Conditional probability and independence

Maksim Zhukovskii

MIPT

Conditional probability

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space, $B \in \mathcal{F}$

Conditional probability

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space, $B \in \mathcal{F}$

If $\mathrm{P}(B) \neq 0$, then $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$.

Conditional probability

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space, $B \in \mathcal{F}$

If $\mathrm{P}(B) \neq 0$, then $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$.
If $\mathrm{P}(B) = 0$, then $\mathrm{P}(A|B) := 0$.

Conditional probability

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space, $B \in \mathcal{F}$

If $\mathrm{P}(B) \neq 0$, then $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$.
If $\mathrm{P}(B) = 0$, then $\mathrm{P}(A|B) := 0$.

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

Conditional probability

Conditional probability measure

$(\Omega, \mathcal{F}, \mathsf{P})$ — probability space,
$B \in \mathcal{F}$, $\mathsf{P}(B) > 0$

Conditional probability measure

$(\Omega, \mathcal{F}, \mathsf{P})$ — probability space,
$B \in \mathcal{F}$, $\mathsf{P}(B) > 0$

$(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$ — probability space:

Conditional probability measure

$(\Omega, \mathcal{F}, \mathsf{P})$ — probability space,
$B \in \mathcal{F}$, $\mathsf{P}(B) > 0$

$(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$ — probability space:

- $\mathsf{P}(\Omega|B) = \frac{\mathsf{P}(B)}{\mathsf{P}(B)} = 1$,

Conditional probability measure

$(\Omega, \mathcal{F}, \mathsf{P})$ — probability space,
$B \in \mathcal{F}$, $\mathsf{P}(B) > 0$

$(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$ — probability space:

- $\mathsf{P}(\Omega|B) = \frac{\mathsf{P}(B)}{\mathsf{P}(B)} = 1$,
- $\mathsf{P}(A_1 \sqcup A_2 \sqcup \ldots |B) =$

Conditional probability measure

$(\Omega, \mathcal{F}, \mathsf{P})$ — probability space,
$B \in \mathcal{F}$, $\mathsf{P}(B) > 0$

$(\Omega, \mathcal{F}, \mathsf{P}(\cdot|B))$ — probability space:

- $\mathsf{P}(\Omega|B) = \frac{\mathsf{P}(B)}{\mathsf{P}(B)} = 1$,
- $\mathsf{P}(A_1 \sqcup A_2 \sqcup \ldots |B) =$
  $= \frac{\mathsf{P}([A_1 \cap B] \sqcup [A_2 \cap B] \sqcup \ldots)}{\mathsf{P}(B)} =$

Conditional probability measure

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space,
$B \in \mathcal{F}$, $\mathrm{P}(B) > 0$

$(\Omega, \mathcal{F}, \mathrm{P}(\cdot|B))$ — probability space:

- $\mathrm{P}(\Omega|B) = \frac{\mathrm{P}(B)}{\mathrm{P}(B)} = 1$,
- $\mathrm{P}(A_1 \sqcup A_2 \sqcup \ldots |B) =$
  $= \frac{\mathrm{P}([A_1 \cap B] \sqcup [A_2 \cap B] \sqcup \ldots)}{\mathrm{P}(B)} =$
  $= \frac{\mathrm{P}(A_1 \cap B) + \mathrm{P}(A_2 \cap B) + \ldots}{\mathrm{P}(B)} =$

Conditional probability measure

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space,
$B \in \mathcal{F}$, $\mathrm{P}(B) > 0$

$(\Omega, \mathcal{F}, \mathrm{P}(\cdot|B))$ — probability space:

- $\mathrm{P}(\Omega|B) = \frac{\mathrm{P}(B)}{\mathrm{P}(B)} = 1$,
- $\mathrm{P}(A_1 \sqcup A_2 \sqcup \ldots |B) =$
  $= \frac{\mathrm{P}([A_1 \cap B] \sqcup [A_2 \cap B] \sqcup \ldots)}{\mathrm{P}(B)} =$
  $= \frac{\mathrm{P}(A_1 \cap B) + \mathrm{P}(A_2 \cap B) + \ldots}{\mathrm{P}(B)} =$
  $= \mathrm{P}(A_1|B) + \mathrm{P}(A_2|B) + \ldots$

Example

Find the probability that a dice shows an even number if it is divisible by 3.

Example

Find the probability that a dice shows an even number if it is divisible by 3.

$$\Omega = \{1, \ldots, 6\},$$
$$A = \{2, 4, 6\},$$
$$B = \{3, 6\}.$$

Example

Find the probability that a dice shows an even number if it is divisible by 3.

$$\begin{aligned}\Omega &= \{1, \dots, 6\},\\ A &= \{2, 4, 6\},\\ B &= \{3, 6\}.\end{aligned}$$

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(\{6\})}{\mathrm{P}(\{3, 6\})} = \frac{1/6}{2/6} = \frac{1}{2}$$

Example

Find the probability that a dice shows an even number if it is divisible by 3.

$$\Omega = \{1, \ldots, 6\},$$
$$A = \{2, 4, 6\},$$
$$B = \{3, 6\}.$$

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(\{6\})}{\mathrm{P}(\{3, 6\})} = \frac{1/6}{2/6} = \frac{1}{2}$$

$\mathrm{P}(A) = \frac{3}{6} = \frac{1}{2}$ independent of B

Law of total probability

$$A,\ B_1, B_2, \ldots \in \mathcal{F}:$$

Law of total probability

$A, \; B_1, B_2, \ldots \in \mathcal{F}$:

- $\Omega = B_1 \sqcup B_2 \sqcup \ldots$

Law of total probability

$A, \; B_1, B_2, \ldots \in \mathcal{F}$:

- $\Omega = B_1 \sqcup B_2 \sqcup \ldots$

The formula

$$\mathrm{P}(A) = \mathrm{P}(A|B_1)\mathrm{P}(B_1) + \mathrm{P}(A|B_2)\mathrm{P}(B_2) + \ldots$$

Law of total probability: the proof

$$\mathrm{P}(A) = \mathrm{P}(A|B_1)\mathrm{P}(B_1) + \mathrm{P}(A|B_2)\mathrm{P}(B_2) + \ldots$$

The first one or the last one?

There are 10 balls in a black box. Two of them are red, and the other eight are blue. Ten children take the balls in turns. Each one take one ball and do not put it back. Winners are those two who get red balls. Who is more likely to be a winner? The first one or the last one?

First solution: law of total probability

Second solution: definition of probability

Bayes' theorem

$A, \; B_1, B_2, \ldots \in \mathcal{F}$:

Bayes' theorem

$A, \; B_1, B_2, \ldots \in \mathfrak{F}$:

- $\mathrm{P}(A) > 0, \; i = 1, 2, \ldots$

Bayes' theorem

$A,\ B_1, B_2, \ldots \in \mathcal{F}$:

- $\mathsf{P}(A) > 0,\ i = 1, 2, \ldots$
- $\Omega = B_1 \sqcup B_2 \sqcup \ldots$

Bayes' theorem

$A,\ B_1, B_2, \ldots \in \mathcal{F}$:

- $\mathrm{P}(A) > 0,\ i = 1, 2, \ldots$
- $\Omega = B_1 \sqcup B_2 \sqcup \ldots$

The formula

$$\mathrm{P}(B_k|A) =$$

$$= \frac{\mathrm{P}(A|B_k)\mathrm{P}(B_k)}{\mathrm{P}(A|B_1)\mathrm{P}(B_1) + \mathrm{P}(A|B_2)\mathrm{P}(B_2) + \ldots}$$

Which condition is most likely?

There are 10 cards in a pack.

Which condition is most likely?

There are 10 cards in a pack.
Three cards are blue, two cards are red, and all the others are green. Alice, Bob and John decide who takes a card from the pack. Alice takes a card with probability 1/2, Bob — 1/3 and John — 1/6.

Which condition is most likely?

There are 10 cards in a pack.
Three cards are blue, two cards are red, and all the others are green. Alice, Bob and John decide who takes a card from the pack. Alice takes a card with probability 1/2, Bob — 1/3 and John — 1/6.
Alice goes to school if she gets a blue card.
Bob goes to school if he gets a red card.
John goes to school if he gets a green card.

Which condition is most likely?

There are 10 cards in a pack.

Three cards are blue, two cards are red, and all the others are green. Alice, Bob and John decide who takes a card from the pack. Alice takes a card with probability $1/2$, Bob — $1/3$ and John — $1/6$.

Alice goes to school if she gets a blue card.
Bob goes to school if he gets a red card.
John goes to school if he gets a green card.

Someone came to school. Find the probabilities that it was Alice, Bob, John.

Which condition is most likely?

Independent events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

Independent events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

- $A, B \in \mathcal{F}$ are independent, if $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$

Independent events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

- $A, B \in \mathcal{F}$ are independent, if $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$
- $A_1, \ldots, A_n \in \mathcal{F}$ are (mutually) independent, if, for any distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$, $\mathrm{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathrm{P}(A_{i_1}) \ldots \mathrm{P}(A_{i_k})$

Pairwise but not mutually

Three faces of a tetrahedra are colored in red, green, blue. And the last one is colored in all three colors. It is rolled. Let R, G, B be the events that the tetrahedra lands on a face with the color red, green, blue respectively.

Pairwise but not mutually

Three faces of a tetrahedra are colored in red, green, blue. And the last one is colored in all three colors. It is rolled. Let R, G, B be the events that the tetrahedra lands on a face with the color red, green, blue respectively.

- R, G, B are pairwise independent,
- R, G, B are not mutually independent

Pairwise but not mutually

Independent sets of events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

Independent sets of events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

- Events in a set $\mathcal{A} \subset \mathcal{F}$ are (mutually) independent,

Independent sets of events

$(\Omega, \mathcal{F}, \mathsf{P})$ — probability space

- Events in a set $\mathcal{A} \subset \mathcal{F}$ are (mutually) independent, if, for any $k \in \mathbb{N}$, and any distinct $A_1, \ldots, A_k \in \mathcal{A}$, $\mathsf{P}(A_1 \cap \ldots \cap A_k) = \mathsf{P}(A_1) \ldots \mathsf{P}(A_k)$.

Independent sets of events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

- Events in a set $\mathcal{A} \subset \mathcal{F}$ are (mutually) independent, if, for any $k \in \mathbb{N}$, and any distinct $A_1, \ldots, A_k \in \mathcal{A}$, $\mathrm{P}(A_1 \cap \ldots \cap A_k) = \mathrm{P}(A_1) \ldots \mathrm{P}(A_k)$.
- $\mathcal{A}_1, \ldots, \mathcal{A}_n \subset \mathcal{F}$ are (mutually) independent,

Independent sets of events

$(\Omega, \mathcal{F}, \mathrm{P})$ — probability space

- Events in a set $\mathcal{A} \subset \mathcal{F}$ are (mutually) independent, if, for any $k \in \mathbb{N}$, and any distinct $A_1, \ldots, A_k \in \mathcal{A}$, $\mathrm{P}(A_1 \cap \ldots \cap A_k) = \mathrm{P}(A_1) \ldots \mathrm{P}(A_k)$.
- $\mathcal{A}_1, \ldots, \mathcal{A}_n \subset \mathcal{F}$ are (mutually) independent, if, for any distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$, every $A_1 \in \mathcal{A}_{i_1}, \ldots, A_k \in \mathcal{A}_{i_k}$ are mutually independent.

An example

Let $A_1, \ldots, A_n$ be independent.

An example

Let $A_1, \ldots, A_n$ be independent.

Then $\mathcal{F}_{A_1}, \ldots, \mathcal{F}_{A_n}$ are independent as well.

Bernoulli scheme: another view

- $A_1, A_2, \ldots$ are independent events,
- $\mathsf{P}(A_1) = \mathsf{P}(A_2) = \ldots$

Bernoulli scheme: another view

- $A_1, A_2, \ldots$ are independent events,
- $\mathrm{P}(A_1) = \mathrm{P}(A_2) = \ldots$

The probability that, say, $A_{i_1}, \ldots, A_{i_k}$ occur, while $A_{j_1}, \ldots, A_{j_m}$ do not, equals $p^k(1-p)^m$.

Random graph: another view

For $1 \leq i < j \leq n$, consider independent events $A_{i,j}$ with $\mathrm{P}(A_{i,j}) = p$.

Random graph: another view

For $1 \leq i < j \leq n$, consider independent events $A_{i,j}$ with $\mathrm{P}(A_{i,j}) = p$.

For every pair of vertices $i < j$ of $\{1, \ldots, n\}$, draw the edge between them if $A_{i,j}$ occurs.

Random graph: another view

For $1 \leq i < j \leq n$, consider independent events $A_{i,j}$ with $\mathrm{P}(A_{i,j}) = p$.

For every pair of vertices $i < j$ of $\{1, \ldots, n\}$, draw the edge between them if $A_{i,j}$ occurs.

The obtained graph is $G(n, p)$.