Probability and Stochastic Processes
Lecture 6: Conditional Probability and Independence

Prof. Washim Uddin Mondal
Department of Electrical Engineering
Indian Institute of Technology Kanpur

Conditional Probability

Definition

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and A, B be $\mathcal{F}$-measurable sets. If $\mathbf{P}(B) > 0$, then the conditional probability of A given B is defined as:

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

Frequentist Interpretation

If, out of n trials, event B occurs n_B times and A, B simultaneously occur n_{AB} times, then

$$\mathbf{P}(A|B) \approx \frac{n_{AB}}{n_B} = \frac{(n_{AB}/n)}{n_B/n} \approx \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

Conditional Probability

Example

Two dice are thrown such that all outcomes are equally likely. Find the conditional probability that the number appearing on the first die is 5 given that the sum of the numbers on the two dice is 10.

Solution

Let A be the event that the number appearing on the first die is 5 and B denote the event that the sum of the numbers on the two dice is 10. Note that

$$A = \{(5,1),(5,2),\cdots,(5,6)\}, \qquad B = \{(4,6),(5,5),(6,4)\}$$

Clearly, $\mathbf{P}(B) = 3/36$, and $\mathbf{P}(A \cap B) = \mathbf{P}(\{5,5\}) = 1/36$. Hence, $\mathbf{P}(A|B) = 1/3$.

Conditional Probability

Example

The probability that a certain machine works for at least t months without wearing down is given by e^{-2t}, $t \geq 0$. Find the probability that a 4 month-old machine will continue to work for at least 2 additional months without breaking down.

Solution

Let A_t denote the event that the machine works for at least t months without breaking down. The desired probability is given below.

$$\mathbf{P}(A_6|A_4) = \frac{\mathbf{P}(A_4 \cap A_6)}{\mathbf{P}(A_4)} = \frac{\mathbf{P}(A_6)}{\mathbf{P}(A_4)} = \frac{e^{-12}}{e^{-8}} = e^{-4}$$

Conditional Probability

Notes

- If $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $B \in \mathcal{F}$ be such that $\mathbf{P}(B) > 0$, then one can define a function $\mathbf{P}(\cdot|B) : \mathcal{F} \to [0,1]$ as follows.

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}, \ \forall A \in \mathcal{F}$$

- Since $0 \leq \mathbf{P}(A \cap B) \leq \mathbf{P}(B)$, the function $\mathbf{P}(\cdot|B)$ is well-defined.

Proposition 6.1

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and $B \in \mathcal{F}$. If $\mathbf{P}(B) > 0$, then $(\Omega, \mathcal{F}, \mathbf{P}(\cdot|B))$ is also a probability space.

Conditional Probability

Proof of Proposition 6.1

(a) $\mathbf{P}(\Omega|B) = \mathbf{P}(\Omega \cap B)/\mathbf{P}(B) = 1.$

(b) Let $\{A_i\}_{i\in I}$ be a countable collection of disjoint $\mathcal{F}$-measurable sets. Clearly, $\{A_i \cap B\}_{i\in I}$ is also a countable collection of disjoint $\mathcal{F}$-measurable sets.

$$\begin{aligned}
\mathbf{P}\left(\cup_{i\in I} A_i | B\right) &= \frac{\mathbf{P}((\cup_{i\in I} A_i) \cap B)}{\mathbf{P}(B)} \\
&= \frac{\mathbf{P}(\cup_{i\in I}(A_i \cap B))}{\mathbf{P}(B)} = \frac{\sum_{i\in I} \mathbf{P}(A_i \cap B)}{\mathbf{P}(B)} = \sum_{i\in I} \mathbf{P}(A_i|B)
\end{aligned}$$

Properties of Conditional Probability Measure

Proposition 6.2

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and A, B, C be $\mathcal{F}$-measurable. If $\mathbf{P}(C) > 0$, then

(a) $\mathbf{P}(\phi|C) = 0$.

(b) $\mathbf{P}(A^c|C) = 1 - \mathbf{P}(A|C)$.

(c) $\mathbf{P}(A^c \cap B|C) = \mathbf{P}(B|C) - \mathbf{P}(A \cap B|C)$.

(d) $\mathbf{P}(A \cup B|C) = \mathbf{P}(A|C) + \mathbf{P}(B|C) - \mathbf{P}(A \cap B|C)$.

(e) $\mathbf{P}(A \cap B|C) \leq \mathbf{P}(B|C)$.

(f) If $A \subset B$, then $\mathbf{P}(A|C) \leq \mathbf{P}(B|C)$.

Independence of Events

Definition

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and A, B be $\mathcal{F}$-measurable. The events A and B are called independent if $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

Notes

(a) If $\mathbf{P}(B) > 0$, and $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$, then $\mathbf{P}(A|B) = \mathbf{P}(A)$.

(b) Three $\mathcal{F}$-measurable events A, B, C are called mutually independent if

  (1) $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$

  (2) $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$

  (3) $\mathbf{P}(B \cap C) = \mathbf{P}(B)\mathbf{P}(C)$

  (4) $\mathbf{P}(C \cap A) = \mathbf{P}(C)\mathbf{P}(A)$

Independence of Events

Example

Two dice are thrown such that all outcomes are equally likely. Let A denote the event that the sum of the numbers on the two dice is 6 and B denote the event that an even number appears on the first die. Check if A and B are independent events.

Solution

- Note that $A = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$. Therefore, $\mathbf{P}(A) = 5/36$.
- Moreover, $B = \{(2,a),(4,a),(6,a) \mid a \in \{1,\cdots,6\}\}$. Hence, $\mathbf{P}(B) = 1/2$.
- Clearly, $\mathbf{P}(A \cap B) = \mathbf{P}(\{(2,4),(4,2)\}) = 2/36$.
- In conclusion, A and B are not independent.

Independence of Events

Proposition 6.3

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and $A, B \in \mathcal{F}$. If A, B are independent, then

(a) A^c and B are also independent,

(b) A^c and B^c are also independent.

Proof of Proposition 6.3(a)

$$\begin{aligned}\mathbf{P}(A^c \cap B) = \mathbf{P}(B) - \mathbf{P}(A \cap B) &= \mathbf{P}(B) - \mathbf{P}(A)\mathbf{P}(B) \\ &= \mathbf{P}(B)[1 - \mathbf{P}(A)] = \mathbf{P}(B)\mathbf{P}(A^c)\end{aligned}$$

Law of Total Probability

Proposition 6.4

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and A be $\mathcal{F}$-measurable. If $\{B_i\}_{i \in I}$ is a countable collection of disjoint $\mathcal{F}$-measurable sets such that $\cup_{i \in I} B_i = \Omega$ (i.e., B_i's partition Ω), then

$$\mathbf{P}(A) = \sum_{i \in I} \mathbf{P}(A | B_i) \mathbf{P}(B_i)$$

Proof of Proposition 6.4

- Note that $\mathbf{P}(A) = \mathbf{P}(A \cap \Omega) = \mathbf{P}(A \cap (\cup_{i \in I} B_i)) = \mathbf{P}(\cup_{i \in I}(A \cap B_i))$.
- Since $\{B_i\}_{i \in I}$ are disjoint, $\{A \cap B_i\}_{i \in I}$ are disjoint as well.
- Since I is countable, $\mathbf{P}(A) = \sum_{i \in I} \mathbf{P}(A \cap B_i) = \sum_{i \in I} \mathbf{P}(A | B_i) \mathbf{P}(B_i)$.

Law of Total Probability

Example

There are three boxes. Box 1 contains 4 red and 2 black balls, box 2 contains 3 red and 3 black balls, and box 3 contains 2 red and 4 black balls. If one picks up a ball at random from a uniformly chosen box, what is the probability that the chosen ball is red?

Solution

Let B_i be the event that the ith box is chosen, $i \in \{1, 2, 3\}$, and R be the event that a red ball is chosen. We have $\mathbf{P}(B_1) = \mathbf{P}(B_2) = \mathbf{P}(B_3) = 1/3$, $\mathbf{P}(R|B_1) = 2/3$, $\mathbf{P}(R|B_2) = 1/2$, and $\mathbf{P}(R|B_3) = 1/3$. Hence,

$$\mathbf{P}(R) = \sum_{i=1}^{3} \mathbf{P}(R|B_i)\mathbf{P}(B_i) = 0.5$$

Bayes' Theorem

Proposition 6.5

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and A be $\mathcal{F}$-measurable. If $\{B_i\}_{i\in I}$ is a countable collection of disjoint $\mathcal{F}$-measurable sets such that $\cup_{i\in I} B_i = \Omega$, and $\mathbf{P}(A) > 0$ then

$$\mathbf{P}(B_i|A) = \frac{\mathbf{P}(A|B_i)\mathbf{P}(B_i)}{\sum_{j\in I}\mathbf{P}(A|B_j)\mathbf{P}(B_j)}$$

Notes

(a) A : observation, B_i : hypothesis, $\mathbf{P}(A|B_i)$: likelihood of A given B_i, $\mathbf{P}(B_i)$: prior probability, $\mathbf{P}(B_i|A)$: posterior probability.

(b) Bayes' Theorem updates belief about B_i using observation A.

Bayes' Theorem

Example

Consider a noisy binary channel. If the transmitter transmits the bits 0 and 1, they are received at the receiver error-free with probabilities 0.8 and 0.6, respectively. Assume that the transmitter is equally likely to send any of the bits. If the receiver receives 1, what is the probability that 0 was transmitted?

Solution

Let A be the event that 1 is received, while B_0, B_1 respectively denote the events that 0 and 1 are transmitted. We have $\mathbf{P}(B_0) = \mathbf{P}(B_1) = 0.5$, $\mathbf{P}(A|B_0) = 0.2$, and $\mathbf{P}(A|B_1) = 0.6$.

$$\mathbf{P}(B_0|A) = \frac{\mathbf{P}(A|B_0)\mathbf{P}(B_0)}{\mathbf{P}(A|B_0)\mathbf{P}(B_0) + \mathbf{P}(A|B_1)\mathbf{P}(B_1)} = 0.25$$