# Conditional Probability

Chapter 2 of Probability and Statistics


Yiping Cheng
ypcheng@bjtu.edu.cn


23-Jul-2018


School of Electronic and Information Engineering
Beijing Jiaotong University

# Chapter Contents

# Concept of Conditional Probability

# Definition of Conditional Probability

Conditional probability is the probability of an event given that (by assumption, presumption, assertion or evidence) another event has occurred.

**Definition**: The conditional probability of $A$ given $B$, which is denoted by $P(A|B)$, is defined as

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

The conditional probability $P(A|B)$ is not defined if $P(B) = 0$.

## Conditional Probability – Examples

**Rolling Dice**. Two dice were rolled and it was observed that the sum $T$ of the two numbers was odd. We shall determine the conditional probability that $T$ was less than 8.

# Conditional Probability – Examples

**Rolling Dice**. Two dice were rolled and it was observed that the sum $T$ of the two numbers was odd. We shall determine the conditional probability that $T$ was less than 8.

Solution: Let $A$ be the event that $T < 8$ and let $B$ be the event that $T$ is odd, then $B$ is the event that $T$ is 3,5,7,9, or 11, and $A \cap B$ is the event that $T$ is 3,5, or 7.

$$P(A \cap B) = P(T = 3) + P(T = 5) + P(T = 7) = \frac{2}{36} + \frac{4}{36} + \frac{6}{36} = \frac{1}{3},$$

$$\begin{aligned} P(B) &= P(T = 3) + P(T = 5) + P(T = 7) + P(T = 9) + P(T = 11) \\ &= \frac{2}{36} + \frac{4}{36} + \frac{6}{36} + \frac{4}{36} + \frac{2}{36} = \frac{1}{2}. \end{aligned}$$

Hence,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{3}.$$

# Conditional Probability – Examples (Continued)

**Rolling Dice Repeatedly**. Two dice are to be rolled repeatedly and the sum $T$ of the two numbers is to be observed for each roll. We shall determine the probability that the value $T = 7$ will be observed before the value $T = 8$ is observed.

# Conditional Probability – Examples (Continued)

**Rolling Dice Repeatedly**. Two dice are to be rolled repeatedly and the sum $T$ of the two numbers is to be observed for each roll. We shall determine the probability that the value $T = 7$ will be observed before the value $T = 8$ is observed.

Solution: Let

- $A$ be the event that $T = 7$ is observed before the value $T = 8$ is observed.
- $B_i$ be the event that $T$ is not 7 nor 8 for the $1, \ldots, (i-1)$th roll, but $T = 7$ or 8 for the $i$th roll.

It is obvious that $B_1, B_2, \ldots$ are disjoint. Let $B = \bigcup_{i=1}^{\infty} B_i$, then $P(B) = 1$.

$$P(A) = P(A \cap B) = \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} P(B_i)P(A|B_i).$$

For each $i$, $P(A|B_i) = \frac{6/36}{(6/36)+(5/36)} = \frac{6}{11}$. Thus

$$P(A) = \frac{6}{11} \sum_{i=1}^{\infty} P(B_i) = \frac{6}{11}.$$

# Conditional Probability – Examples (Continued 2)

**A Clinical Trial**. Consider 150 patients who entered the study after an episode of depression.

Table: Results of the clinical depression study

| Response | Treatment group | | | | Total |
|---|---|---|---|---|---|
| | Imipramine | Lithium | Combination | Placebo | |
| Relapse | 18 | 13 | 22 | 24 | 77 |
| No Relapse | 22 | 25 | 16 | 10 | 73 |
| Total | 40 | 38 | 38 | 34 | 150 |

Compute the conditional probability of relapse given the patient received placebo. Do the same for lithium.

Solution: Let $A$ be the event that the patient had a relapse, $B$ be the event that the patient received the placebo, and $C$ be the event that the patient received lithium.

$$P(B) = \frac{34}{150}, P(A \cap B) = \frac{24}{150}, P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{24}{34}$$

$$P(C) = \frac{38}{150}, P(A \cap C) = \frac{13}{150}, P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{13}{38}.$$

# Multiplication Rule for Conditional Probabilities

**Theorem**

Let $A$ and $B$ be events. If $P(B) > 0$, then

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

If $P(A) > 0$, then

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

## Example

**Selecting Two Balls**. Two balls are to be selected at random, without replacement, from a box containing $r$ red balls and $b$ blue balls. We shall determine the probability that the first ball will be red and the second ball will be blue.

Solution: Let $A$ be the event that the first ball is red, and let $B$ be the event that the second ball is blue.

$$P(A \cap B) = P(A)P(B|A) = \frac{r}{r+b} \cdot \frac{b}{r+b-1}.$$

# Multiplication Rule for Conditional Probabilities (Continued)

**Theorem**

Let $A_1, A_2, \ldots, A_n$ be events. If $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

## Example

**Selecting Four Balls**. Four balls are to be selected at random, without replacement, from a box containing $r$ red balls and $b$ blue balls $(r \geq 2, b \geq 2)$. We shall determine the probability of obtaining the sequence of outcomes red, blue, red, blue.

## Example

**Selecting Four Balls**. Four balls are to be selected at random, without replacement, from a box containing $r$ red balls and $b$ blue balls $(r \geq 2, b \geq 2)$. We shall determine the probability of obtaining the sequence of outcomes red, blue, red, blue.

Solution: Let $R_j$ denote the event that a red ball is obtained on the $j$th draw and let $B_j$ denote the event that a blue ball is obtained on the $j$th draw, then

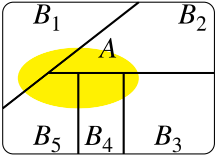
$$\begin{aligned} P(R_1 \cap B_2 \cap R_3 \cap B_4) &= P(R_1)P(B_2|R_1)P(R_3|R_1 \cap B_2)P(B_4|R_1 \cap B_2 \cap R_3) \\ &= \frac{r}{r+b} \cdot \frac{b}{r+b-1} \cdot \frac{r-1}{r+b-2} \cdot \frac{b-1}{r+b-3}. \end{aligned}$$

# Law of Total Probability

By partitioning the sample space into two events $B$ and $B^c$, we have

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

**Definition [Partition].** The events $B_1, \ldots, B_k$ are said to be a partition of the sample space $S$, if they are disjoint and $\bigcup_{i=1}^{k} B_i = S$.




**Theorem [Law of Total Probability].** Suppose that the events $B_1, \ldots, B_k$ form a partition of the sample space $S$ and $P(B_j) > 0$ for $j = 1, \ldots, k$. Then for every event $A$,

$$P(A) = \sum_{j=1}^{k} P(B_j)P(A|B_j).$$

# Law of Total Probability – Examples

**Selecting Bolts**. One box contains 60 long bolts and 40 short bolts, and the other box contains 10 long bolts and 20 short bolts. Now one box is selected at random and a bolt is then selected at random from that box. Determine the probability that this bolt is long.

# Law of Total Probability – Examples

**Selecting Bolts**. One box contains 60 long bolts and 40 short bolts, and the other box contains 10 long bolts and 20 short bolts. Now one box is selected at random and a bolt is then selected at random from that box. Determine the probability that this bolt is long.

Solution: Let $B_1$ be the event that the first box (the one with 60 long and 40 short bolts) is selected, let $B_2$ be the event that the second box (the one with 10 long and 20 short bolts) is selected, and let $A$ be the event that a long bolt is selected.

$$\begin{aligned} P(A) &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) \\ &= \frac{1}{2} \times \frac{60}{100} + \frac{1}{2} \times \frac{10}{30} \\ &= \frac{7}{15}. \end{aligned}$$

# Bayes' Theorem

**Theorem**: Let the events $B_1, \ldots, B_k$ form a partition of the sample space $S$ such that $P(B_j) > 0$ for $j = 1, \ldots, k$, and let $A$ be an event such that $P(A) > 0$. Then, for $i = 1, \ldots, k$,

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{k} P(B_j)P(A|B_j)}.$$

**Proof.**

By the definition of conditional probability, $P(B_i|A) = \frac{P(B_i \cap A)}{P(A)}$ $= \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{k} P(B_j)P(A|B_j)}$, where the last equality for the numerator is according to the multiplication rule, and the last equality for the denominator is according to the law of total probability. $\square$

# Bayes' Theorem – Examples

**Selecting Bolts**. One box contains 60 long bolts and 40 short bolts, and the other box contains 10 long bolts and 20 short bolts. Now one box is selected at random and a bolt is then selected at random from that box. Now we know a long bolt is selected. Given this condition, determine the conditional probability that the first box is selected.

# Bayes' Theorem – Examples

**Selecting Bolts**. One box contains 60 long bolts and 40 short bolts, and the other box contains 10 long bolts and 20 short bolts. Now one box is selected at random and a bolt is then selected at random from that box. Now we know a long bolt is selected. Given this condition, determine the conditional probability that the first box is selected.

Solution: Let $B_1$ be the event that the first box (the one with 60 long and 40 short bolts) is selected, let $B_2$ be the event that the second box (the one with 10 long and 20 short bolts) is selected, and let $A$ be the event that a long bolt is selected.

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2)} = \frac{\frac{1}{2} \times \frac{60}{100}}{\frac{7}{15}} = \frac{9}{14}.$$

# Bayes' Theorem – Examples (Continued)

**Identifying the Source of a Defective Item**. Three machines, $M_1$, $M_2$, and $M_3$, were used for producing a batch of items. The ratio of the batch of product items between the machines was 20:30:50. The defective rates of the product items for the three machines were 0.01,0.02,0.03, respectively. One item is selected at random from the batch and it is found to be defective. We shall determine the probability that this item was produced by machine $M_2$.

# Bayes' Theorem – Examples (Continued)

**Identifying the Source of a Defective Item**. Three machines, $M_1$, $M_2$, and $M_3$, were used for producing a batch of items. The ratio of the batch of product items between the machines was 20:30:50. The defective rates of the product items for the three machines were 0.01,0.02,0.03, respectively. One item is selected at random from the batch and it is found to be defective. We shall determine the probability that this item was produced by machine $M_2$.

Solution: Let $B_i$ be the event that the selected item was produced by $M_i$, $i = 1, 2, 3$, and let $A$ be the event that the selected item is defective. What we need to evaluate is $P(B_2|A)$. From above, we have

$$P(B_1) = 0.2, P(B_2) = 0.3, P(B_3) = 0.5$$
$$P(A|B_1) = 0.01, P(A|B_2) = 0.02, P(A|B_3) = 0.03.$$

$$\text{Thus, } P(B_2|A) = \frac{P(B_2)P(A|B_2)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)}$$
$$= \frac{0.3 \times 0.02}{0.2 \times 0.01 + 0.3 \times 0.02 + 0.5 \times 0.03} = 0.2608....$$

# Independent Events

# Introducing Independence

Event $A$ depends on event $B$ means that the occurrence of $B$ changes the probability of $A$, i.e. $P(A) \neq P(A|B)$. Conversely, $A$ is independent of $B$, if $P(A) = P(A|B)$.

**Example**

**Tossing Coins**. A fair coin is tossed twice. Let event $A = \{H \text{ on second toss}\}$. $B = \{T \text{ on first toss}\}$.

$$S = \{HH, HT, TH, TT\}$$

$$A = \{HH, TH\}, A \cap B = \{TH\}, B = \{TH, TT\}$$

$$P(A) = \frac{2}{4} = \frac{1}{2}, P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{2/4} = \frac{1}{2}. \quad P(A) = P(A|B).$$

The probability of $A$ does not change given the occurrence of $B$. Hence in this case, we say that $A$ is independent of $B$.

# Independence of Two Events

We have two findings regarding the previous described concept of independence.

- Independence is actually symmetric. That is, $A$ independent of $B$ is actually equivalent to $B$ independent of $A$. This is because the equation $P(A) = P(A|B)$ can be rewritten as $P(A) = \frac{P(A\cap B)}{P(B)}$, and this is equivalent to $P(A \cap B) = P(A)P(B)$, assuming $P(B) > 0$.
- Independence can be defined without mentioning conditional probability.

**Definition**: Two events $A$ and $B$ are independent if

$$P(A \cap B) = P(A)P(B).$$

# Independence of Two Events – Examples

**Machine Operation**. Two machines 1 and 2 in a factory are operated independently of each other. Let $A$ be the event that machine 1 will become inoperative during a given 8-hour period, let $B$ be the event that machine 2 will become inoperative during the same period, and suppose that $P(A) = 1/3$ and $P(B) = 1/4$. Determine the probability that at least one of the machines will become inoperative during the given period.

# Independence of Two Events – Examples

**Machine Operation**. Two machines 1 and 2 in a factory are operated independently of each other. Let $A$ be the event that machine 1 will become inoperative during a given 8-hour period, let $B$ be the event that machine 2 will become inoperative during the same period, and suppose that $P(A) = 1/3$ and $P(B) = 1/4$. Determine the probability that at least one of the machines will become inoperative during the given period.
Solution:

$$
\begin{aligned}
P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
&= P(A) + P(B) - P(A) \times P(B) \\
&= \frac{1}{3} + \frac{1}{4} - \frac{1}{3} \times \frac{1}{4} \\
&= \frac{1}{2}.
\end{aligned}
$$

# Independence of Two Events – Examples (Continued)

**Rolling a Die**. A balanced die is rolled. Let $A$ be the event that an even number is obtained, and let $B$ be the event that one of the numbers 1, 2, 3, or 4 is obtained. Show that the events $A$ and $B$ are independent.

Solution:

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

$$P(A \cap B) = P(A) \times P(B).$$

It follows that the events $A$ and $B$ are independent, even though the occurrence of each event depends on the same roll of a die.

# Independence of Complements

**Theorem**: If two events $A$ and $B$ are independent, then

- $A$ and $B^c$ are also independent.
- $A^c$ and $B$ are also independent.
- $A^c$ and $B^c$ are also independent.

**Proof.**

Property 6 of probability (Chapter 1) says that $P(A \cap B^c) = P(A) - P(A \cap B)$. Furthermore, since $A$ and $B$ are independent, $P(A \cap B) = P(A)P(B)$. It now follows that

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A)[1 - P(B)] = P(A)P(B^c).$$

Therefore, the events $A$ and $B^c$ are independent.
The proof for the remaining parts of the theorem is similar to above and is omitted. □

# Independence of Several Events

The definition of independent events can be extended to any number of events. Intuitively, a group of events are independent means that the occurrences or non-occurrences of some of the events do not change the probabilities for any events that depend only on the remaining events.

**Definition**: The $k$ events $A_1, ..., A_k$ are independent (or mutually independent) if, for every subset $A_{i_1}, ..., A_{i_j}$ of $j$ of these events $(j = 2, 3, ..., k)$,

$$P(A_{i_1} \cap ... \cap A_{i_j}) = P(A_{i_1})...P(A_{i_j}).$$

As an example, in order for three events $A$, $B$, and $C$ to be independent, the following four relations must be satisfied:

$$P(AB) = P(A)P(B), P(AC) = P(A)P(C)$$

$$P(BC) = P(B)P(C), P(ABC) = P(A)P(B)P(C).$$

# Independence of Several Events – Example

**Pairwise Independence**. A fair coin is tossed twice so that the sample space $S = \{HH, HT, TH, TT\}$ is simple. Define the following three events:

$$A = \{H \text{ on first toss}\} = \{HH, HT\},$$

$$B = \{H \text{ on second toss}\} = \{HH, TH\},$$

$$C = \{\text{Both tosses the same}\} = \{HH, TT\}.$$

Then $AB = AC = BC = ABC = \{HH\}$. Hence,

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

$$P(AB) = P(AC) = P(BC) = P(ABC) = \frac{1}{4}$$

$$P(AB) = P(A)P(B), P(AC) = P(A)P(C)$$

$$P(BC) = P(B)P(C), P(ABC) \neq P(A)P(B)P(C)$$

These results can be summarized by saying that the events $A$, $B$, and $C$ are pairwise independent, but all three events are not independent.

# Homework for Chapter 2

- P65: 5, 7
- P75: 7, 12
- P85: 4
- P90: 3, 4, 13, 23