50.034 - Introduction to Probability and Statistics

Week 2 – Lecture 3

January–May Term, 2019

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Lecture 3 Assigned Readings: 2.1, 2.2, 2.3

Outline of Lecture

- Conditional probability
- Independent events
- Bayes' theorem
- Prior and posterior probabilities

Conditional probability

Most of the probabilities we encounter in our lives are actually conditional probabilities:

- Weather forecast (e.g. 40% probability of rain) is *conditional* on factors such as radar readings, or if it is monsoon season.
- The probability of getting into an accident during a Grab ride is *conditional* on factors such as driving records of driver, trip route, traffic conditions, etc.
- The probability that your newly launched mobile app will hit 100k downloads after a month is *conditional* on the size of your potential customer base, marketing strategy, etc.

Conditional probability

The conditional probability of an event A given that an event B has occurred, is denoted by $\Pr(A|B)$, where B is the conditioning event.

Example: Live Announcement of Results of Lucky Draw Contest

- Suppose there are 100 participants, including you.
- Every participant has a lucky draw ticket with a number from $001, 002, \ldots, 100$. Suppose your ticket number is 093.
- Every ticket has an equal chance of being selected.

Let A be the event that the winning ticket number is 093.
Question: What is $\Pr(A)$?

Let B be the event that the announcer confirms the winning ticket number is > 90. Suppose event B has occured.
Question: What is your chance of winning, given the new info from the announcer? In other words, what is $\Pr(A|B)$?

Example 1

A helicopter company has two factories. This year, the company has produced 200 helicopters in total. The info about the products is given in the table.

	D (defective)	D^c (not defective)
F_1 (factory 1)	3	77
F_2 (factory 2)	2	118

1. Given that a helicopter has been found defective, what is the probability $\Pr(F_2|D)$ that it came from factory 2?
2. What is the relationship between $\Pr(D)$, $\Pr(F_2 \cap D)$ and $\Pr(F_2|D)$?

Example 1

According to the table, there are 5 defective helicopters, two of which are from factory 2. Therefore,

$$\Pr(F_2|D) = \frac{2}{5} = \frac{2/200}{5/200}.$$

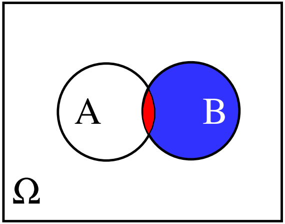
According to the table, we can also calculate

$$\begin{aligned} \Pr(D) &= \frac{5}{200} \\ \Pr(F_2 \cap D) &= \frac{2}{200} \end{aligned} \qquad \Rightarrow \qquad \Pr(F_2|D) = \frac{\Pr(F_2 \cap D)}{\Pr(D)}.$$

Definition of conditional probability

For any two events A and B with $\Pr(B) > 0$, the conditional probability of A given that B has occurred is defined by

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



$\Pr(A|B)$ is the ratio of the red and blue regions.

For convenience, we usually say "conditional probability of A given B" to mean $\Pr(A|B)$.

Interpretation of conditional probability

The conditional probability is expressed as a ratio of unconditional probabilities.

Given that B has occurred:

- The relevant outcomes are no longer all possible outcomes in the sample space, but only those outcomes that are contained in B.
- A has occurred if and only if one of the outcomes in $A \cap B$ has occurred.

Thus, the conditional probability of A given B is proportional to $\Pr(A \cap B)$.

Example 2

Doctors decide whether a person is sick or not by performing a blood test. The test has two outcomes: positive (implying sick) and negative (implying healthy). The joint probabilities of test result and health status are given in the table

	Healthy	Sick
Negative	0.72	0.005
Positive	0.18	0.095

1. What is the probability that a person is sick?
2. If a person is known to be sick, what is the probability that his/her blood test result is negative?
3. If a person is known to be healthy, what is the probability that his/her blood test is positive?

Example 2

1. What is the probability that a person is sick?

$$\begin{aligned}\Pr(\text{sick}) &= \Pr(\text{sick and negative}) + \Pr(\text{sick and positive})\\ &= 0.005 + 0.095 = 0.1\end{aligned}$$

2. If a person is known to be sick, what is probability that his/her blood test result is negative?

$$\begin{aligned}\Pr(\text{negative}|\text{sick}) &= \Pr(\text{negative and sick}) / \Pr(\text{sick})\\ &= 0.005/0.1 = 0.05\end{aligned}$$

3. If a person is known to be healthy, what is the probability that his/her blood test is positive?

$$\begin{aligned}\Pr(\text{positive}|\text{healthy}) &= \Pr(\text{positive and healthy}) / \Pr(\text{healthy})\\ &= 0.18/0.9 = 0.02\end{aligned}$$

Intuition for Independent Events

Coin toss experiment: Suppose that a fair coin is tossed twice.

- 4 possible outcomes: HH, HT, TH, TT (probability $\frac{1}{4}$ each).

Let A be the event $A = \{\text{H on second toss}\}$. What is $\Pr(A)$?

- We check that $A = \{\text{HH, TH}\}$, so $\Pr(A) = \frac{2}{4} = \frac{1}{2}$.

Suppose we know that the first toss was T. What is the condition probability of A given the event $B = \{\text{T on first toss}\}$?

- By the definition of conditional probability,
$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(\{\text{TH}\})}{\Pr(\{\text{TH, TT}\})} = \frac{1/4}{2/4} = \frac{1}{2}.$$

Consequence: The probability that A occurs does not change, even after we have learned that B has occurred.

- Intuitively, the outcome of the first toss does not affect the outcome of the second toss.
- So intuitively, A and B are independent events.

Definition of Independence

Two events A and B are called independent if

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

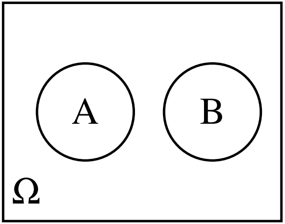
Two events A and B are called dependent if

$$\Pr(A \cap B) \neq \Pr(A)\Pr(B).$$

Important Remark: Disjoint doesn't imply independence!

Let A and B be two disjoint events, whose probabilities are proportional to their areas in the diagram.

- Disjoint $\Rightarrow \Pr(A \cap B) = 0$.
- BUT: $\Pr(A)\Pr(B) > 0$.



Thus, A and B are not independent, even though they are disjoint.

Independence and conditional probability

Fact: If A and B are independent events, and $\Pr(B) > 0$, then $\Pr(A|B) = \Pr(A)$.

Proof: By definitions of conditional probability and independence,

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

Fact: Similarly, if A and B are independent events, and $\Pr(A) > 0$, then $\Pr(B|A) = \Pr(B)$.

Proof: Again, by the same definitions,

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

What about dependent events?

- If A, B are dependent, and $\Pr(B) > 0$, then $\Pr(A|B) \neq \Pr(A)$.
- If A, B are dependent, and $\Pr(A) > 0$, then $\Pr(B|A) \neq \Pr(B)$.

Interpretation of independent events

Suppose A and B are events, such that $\Pr(A) > 0$ and $\Pr(B) > 0$.
Question: When are A and B independent events?

From the previous slide (assuming $\Pr(A) > 0$ and $\Pr(B) > 0$),

- A and B are independent if and only if $\Pr(A|B) = \Pr(A)$.
- A and B are independent if and only if $\Pr(B|A) = \Pr(B)$.

In other words, independence means that the probability of whether one event (A or B) occurs or not, is NOT affected even if we know whether the other event (B or A) has occurred or not.

Independence of several events

Events $A_1, A_2, \ldots, A_n$ are mutually independent if for every subset of indices $\{i_1, i_2, \ldots, i_k\}$ (for $k = 2, 3, \ldots, n$),

$$\Pr(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \Pr(A_{i_1}) \Pr(A_{i_2}) \cdots \Pr(A_{i_k}).$$

In other words, the events $A_1, \ldots, A_n$ are mutually independent if the probability of the intersection of **any subset** of the n events is equal to the product of the individual probabilities.

Question

Events A_1, A_2 and A_3 are pairwise independent. Are they mutually independent? Not necessarily.

Example: A fair coin is tossed twice. Let A be the event of head on the first toss, B be the event of head on the second toss, and C the event that exactly one head occurs in two tosses.

$\Pr(A) = \Pr(B) = \Pr(C) = 0.5$. A and B are clearly independent. To see that A and C are independent, we note that $\Pr(C|A) = 0.5$. However,

$$\Pr(A \cap B \cap C) = 0 \neq \Pr(A)\Pr(B)\Pr(C).$$

The law of total probability

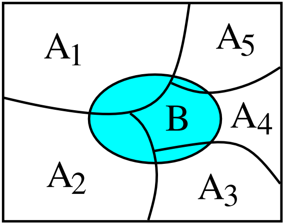
Let $A_1, A_2, \ldots, A_k$ be events in some sample space Ω.

- The events $A_1, A_2, \ldots, A_k$ are called exhaustive if at least one A_i must occur, i.e. $A_1 \cup A_2 \cup \cdots \cup A_k = \Omega$.
- The events $A_1, A_2, \ldots, A_k$ are called mutually exclusive if no two distinct A_i, A_j can occur at the same time, i.e. $A_i \cap A_j = \varnothing$ for all $i \neq j$.

Let $A_1, A_2, \ldots, A_k$ be **mutually exclusive** and **exhaustive** events. Then for any event B, the **law of total probability** states that

$$\Pr(B) = \sum_{i=1}^{k} \Pr(B|A_i)P(A_i)$$

(See Sec. 2.1.4 in textbook for a proof.)



Bayes' theorem

Let $A_1, A_2, \ldots, A_k$ be **mutually exclusive** and **exhaustive** events. Let B be an event such that $\Pr(B) > 0$. Then **Bayes' theorem** states that for every $j = 1, \ldots, k$,

$$\Pr(A_j|B) = \frac{\Pr(B|A_j)\Pr(A_j)}{\Pr(B)} = \frac{\Pr(B|A_j)\Pr(A_j)}{\sum_{i=1}^{k}\Pr(B|A_i)\Pr(A_i)}$$

Proof: By the definition of conditional probability,

$$\Pr(A_j|B) = \frac{\Pr(A_j \cap B)}{\Pr(B)}.$$

Using the definition of conditional probability again, the numerator $\Pr(A_j \cap B)$ can be replaced by

$$\Pr(A_j \cap B) = \Pr(B|A_j)\Pr(A_j).$$

By the law of total probability, the denominator $\Pr(B)$ can be replaced by

$$\Pr(B) = \sum_{i=1}^{k}\Pr(B|A_i)P(A_i).$$

Example 3

A mobile phone manufacturer subcontracts mobile LCD screen production to 3 factories F_1, F_2 and F_3. Based on historical record, these factories have a defective product rate of 0.1%, 0.2% and 0.3% respectively. Last year, factories F_1, F_2, F_3 produced 4000, 16000, 4000 LCD screens respectively for the manufacturer.

If a defective LCD screen is found in a randomly selected product from last year, what is the probability that this defective LCD screen is produced by F_1, F_2 and F_3 respectively?

Example 3: Solution

Define the following events:

- A_1 = "selected LCD screen is produced by F_1".
- A_2 = "selected LCD screen is produced by F_2".
- A_3 = "selected LCD screen is produced by F_3".
- B = "selected LCD screen is defective".

By the law of total probability,

$$\begin{aligned}\Pr(B) &= \Pr(B|A_1)P(A_1) + \Pr(B|A_2)P(A_2) + \Pr(B|A_3)P(A_3) \\ &= \frac{0.001 \cdot 4000 + 0.002 \cdot 16000 + 0.003 \cdot 4000}{24000} \\ &= 0.002.\end{aligned}$$

Example 3: Solution (continued)

By Bayes' theorem, the probability that the defective LCD screen is produced by F_j is

$$\Pr(A_i|B) = \frac{\Pr(B|A_i)\Pr(A_i)}{\Pr(B)}.$$

Therefore:

$$\Pr(A_1|B) = \frac{\Pr(B|A_1)\Pr(A_1)}{\Pr(B)} = \frac{0.001 \cdot \frac{4000}{24000}}{0.002} = \frac{1}{12}.$$

$$\Pr(A_2|B) = \frac{\Pr(B|A_2)\Pr(A_2)}{\Pr(B)} = \frac{0.002 \cdot \frac{16000}{24000}}{0.002} = \frac{2}{3}.$$

$$\Pr(A_3|B) = \frac{\Pr(B|A_3)\Pr(A_3)}{\Pr(B)} = \frac{0.003 \cdot \frac{4000}{24000}}{0.002} = \frac{1}{4}.$$

Prior and posterior probabilities

Fair coin versus biased coin: Your friend has two coins, a fair coin, and a biased coin that always gives heads. He randomly selects one of the coins, and asks if the selected coin is fair.

- Let A be the event "selected coin is fair".
- Your *prior* guess: $\Pr(A) = 0.5$. Without more information, you have no reason to favour A (coin is fair) or A^c (coin is biased).
- You toss the coin 10 times and record all 10 outcomes.

Suppose the event $B =$ "all heads for 10 tosses" occurs.

Event B would strongly suggest that the selected coin is NOT fair. Should you update your guess, given that B has occurred?

In other words, what should $\Pr(A|B)$ be?

Gathering experimental evidence to check your prior guess is common practice. In such a scenario, $\Pr(A)$ is called the prior probability, and $\Pr(A|B)$ is called the posterior probability.

Bayes' theorem reinterpreted

Bayes' theorem allows us to compute the posterior probability from some prior probability.

Reinterpretation of Bayes' theorem:

Let $A_1, A_2, \ldots, A_k$ be **mutually exclusive** and **exhaustive** events with **prior probabilities** $\Pr(A_i)$ (for $i = 1, 2, \ldots, k$). Suppose B is an event with probability $\Pr(B) > 0$. Then for every $j = 1, \ldots, k$, the **posterior probability** of A_j given that B has occurred is

$$\Pr(A_j|B) = \frac{\Pr(B|A_j)\Pr(A_j)}{\Pr(B)} = \frac{\Pr(B|A_j)\Pr(A_j)}{\sum_{i=1}^{k}\Pr(B|A_i)\Pr(A_i)}.$$

Remarks:

- Typically, the conditional probabilities $\Pr(B|A_1), \ldots, \Pr(B|A_k)$ are easy to determine directly.
- Since the prior probabilities $\Pr(A_1), \ldots, \Pr(A_k)$ are assumed to be known, we can then use Bayes' theorem to compute the posterior probability $\Pr(A_j|B)$ (for each $j = 1, \ldots, k$).

An initial introduction to the Bayesian philosophy

The Bayesian philosophy is based on Bayes' theorem. The main idea of this philosophy is that the probability of a random event can be **updated** with new evidence, as follows:

- The event of interest (your initial hypothesis, e.g. "selected coin is fair") is assigned a prior probability.
- As we gather experimental evidence, we update our guess on whether the hypothesis is true with the posterior probability.
- If A is the event of interest, and B is the event representing experimental evidence, then $\Pr(A)$ is the prior probability, and $\Pr(A|B)$ is the posterior probability.
- The posterior probability $\Pr(A|B)$ can then be computed using Bayes' theorem.

Fair coin versus biased coin revisited

Event A = "selected coin is fair". Event A^c = "selected coin always gives heads". Event B = "all heads for 10 tosses".

- Events A and A^c are mutually exclusive and exhaustive.
- Prior probabilities: $\Pr(A) = 0.5$, $\Pr(A^c) = 0.5$.

We can compute the following conditional probabilities:

$$\begin{aligned}
\Pr(B|A) &= \text{Probability that selected coin gives all heads for 10} \\
&\quad\ \text{tosses, given that selected coin is fair} \\
&= \tfrac{1}{2^{10}} \approx 0.0009766. \\
\Pr(B|A^c) &= \text{Probability that selected coin gives all heads for 10} \\
&\quad\ \text{tosses, given that selected coin always gives heads} \\
&= 1.
\end{aligned}$$

Therefore by Bayes' theorem, the **posterior probability** is

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A^c)\Pr(A) + \Pr(B|A^c)\Pr(A^c)} \approx 0.001949.$$

Summary

- Conditional probability
- Independent events
- Bayes' theorem
- Prior and posterior probabilities

Reminder:
Make-up Lecture 4 is held this Friday, 2:30pm-4pm, at LT5.