BNAD 276
Lecture 4
Probability Theory II

Phuong Ho

May 18, 2017

Outline

Outline

Motivating Example

We collected data from 10,000 children. For each child, we ask whether that child is girl or boy. In addition, we also asked whether that child is right-handed or left-handed. Then, we summarize the data in the following table.

	Girl	Boy	Total
RH	4,500	4,275	8,775
LH	500	725	1,225
Total	5,000	5,000	10,000

	Girl	Boy	Total
RH	4500	4275	8775
LH	500	725	1225
Total	5000	5000	10000

Now, suppose that we randomly draw a child from this sample.
Then, that child can be either male or female, and either left-handed or right-handed.
Then, we have basic four possible events, G, B, RH, and LH.

$$\begin{aligned} G &= \text{event that a randomly drawn child is a girl} \\ B &= \text{event that a randomly drawn child is a boy} \\ RH &= \text{event that a randomly drawn child is right-handed} \\ LH &= \text{event that a randomly drawn child is left-handed} \end{aligned}$$

Example Cont'd

Suppose the question is "what are the probabilities of each of those above events?

e.g. What is the probability that event G occurs?

- In other words, what is the probability that the randomly drawn child is a girl?
- Note that we have 10000 children in total. Among 1000 children, 5000 children are girls.
- Thus, $P(G) = \frac{5000}{10000} = \frac{1}{2} = 0.5$

We can do the same thing for another events, B, RH, and LH.

- $P(B) = \frac{5000}{10000} = \frac{1}{2} = 0.5$
- $P(RH) = \frac{8775}{10000} = 0.8775$
- $P(LH) = \frac{1225}{10000} = 0.1225$

Now, the question is

"what is the probability of joint events such as $G \cap RH$, $G \cap LH$, $B \cap RH$, and $B \cap LH$?"

Note that, for example,

$$P(G \cap RH) = P(\text{A child is a girl AND right-handed})$$

- The meaning of $G \cap RH$ is that the event G and RH happens at the same time.
- It means a randomly drawn child is a girl AND right-handed.
- We call this new event **the joint event.** And the probability of this joint event is called **joint probability.**

Note that we have 4,500 children who are girls and left-handed among 10000 children.

Thus,

$$P(G \cap RH) = \frac{4500}{10000} = 0.45$$

We can do the same thing for another joint events, $G \cap LH$, $B \cap RH$, and $B \cap LH$.

- $P(G \cap LH) = \frac{500}{10000} = 0.05$
- $P(B \cap RH) = \frac{4275}{10000} = 0.4275$
- $P(B \cap LH) = \frac{725}{10000} = 0.0725$

Let's summarize the joint probabilities that we've got in the following table.

	Girl	Boy	Total
RH	0.45 $= P(G \cap RH)$	0.4275 $= P(B \cap RH)$	0.8775 $= P(RH)$
LH	0.05 $= P(G \cap LH)$	0.0725 $= P(B \cap LH)$	0.1225 $= P(LH)$
Total	0.5 $= P(G)$	0.5 $= P(B)$	1

- Again, the probability of a joint event (e.g. $G \cap RH$) is called **Joint Probability.**
- The above table that shows all joint probabilities is called **Joint Probability Distribution.**

	Girl	Boy	Total
RH	0.45 $= P(G \cap RH)$	0.4275 $= P(B \cap RH)$	0.8775 $= P(RH)$
LH	0.05 $= P(G \cap LH)$	0.0725 $= P(B \cap LH)$	0.1225 $= P(LH)$
Total	0.5 $= P(G)$	0.5 $= P(B)$	1

- Now, look at the margin of the above table.
- At the right margin, we have $P(RH)$ and $P(LH)$.
- At the bottom margin, we have $P(G)$ and $P(B)$.
- The probabilities at the margin ($P(RH)$, $P(LH)$, $P(G)$, and $P(B)$) are called **Marginal Probabilities**.

Outline

Conditional Probability

- The concept of conditional probability allows us to update the "new probability" when a new piece of information is obtained.
- Suppose we have two relevant events, A and B.
- Then, we heard that event B happens.
- Now the question is what is the probability that event A happens given that event B has already occurred?

Example

- Suppose we are working in a sales management department for a company.
- Suppose we have two events, A and B. A is the event that our company's sales decreases. B is the event that the world economy is collapsed.
- We know that these two events are relevant to each other but not exactly the same or perfectly relevant.
- We want to know the probability that A occurs given that B already occurred.

Conditional Probability

Conditional Probability: Notation

Suppose we have two events A and B. The probability that event A happens when B already occurred is denoted by

$$P(A|B),$$

and we say: probability of A given B.

- Another interpretation: What do we know about the probability of A if we have known B?

Back to the first example in this lecture note

Let's start asking conditional probabilities, e.g. $P(G|LH)$: What is the probability that a randomly drawn child is left-handed GIVEN that the child is a girl?

One can work it out by reasoning as follows:

- There are 5,000 who are girls.
- The fact that our randomly drawn child is a girl is given.
- Now, we know that our randomly drawn child is a girl. We narrow our focus to this group.
- There are 500 left-handed children in the girl group.
- Thus, the probability that a randomly drawn child is left handed given that the child belong to girl group is simply

$$P(LH|G) = \frac{500}{5,000} = \frac{500/10,000}{5,000/10,000} = \frac{0.05}{0.5} = 0.10$$

Motivating the conditional probability formula

Recall the joint probability distribution.

	Girl	Boy	Total
RH	0.45	0.4275	0.8775
LH	0.05	0.0725	0.1225
Total	0.5	0.5	1

We find that

$$P(LH|G) = \frac{0.05}{0.5} = \frac{P(G \cap LH)}{P(G)}$$

It motivates the formula for conditional probability.

Conditional Probability Formula

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

OR

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

Example

The following is the joint probability distribution for 4 events, M, F, P, NP.

	M	F	Total
P	0.24	0.03	0.27
NP	0.56	0.17	0.73
Total	0.80	0.20	1

$$\begin{aligned} M &= \text{event that an employee is male} \\ F &= \text{event that an employee is female} \\ P &= \text{event that an employee is promoted} \\ NP &= \text{event that an employee is not promoted} \end{aligned}$$

Example Cont'd

We want to know if there is a gender-based discrimination in promotion.

	M	F	Total
P	0.24	0.03	0.27
NP	0.56	0.17	0.73
Total	0.80	0.20	1

- At the first glance, it seems that there is the discrimination since the probability that a female employee is promoted is significantly lower than the probability that a male employee is promoted.
- That is, $P(M \cap P) = 0.24 > P(F \cap P) = 0.03$.
- But, this comparison is not quite right.

- We need to compare the probability that an employee is promoted given the employee is male with the probability that an employee is promoted given the employee is female.
- That is, we need to compare the conditional probabilities.
- Comparison between $P(P|M)$ and $P(P|F)$ will give us correct answer for this.
- Recall the formula for the conditional Probability.

$$
\begin{aligned}
P(P|M) &= \frac{P(P \cap M)}{P(M)} \\
P(P|F) &= \frac{P(P \cap F)}{P(F)}
\end{aligned}
$$

	M	F	Total
P	0.24	0.03	0.27
NP	0.56	0.17	0.73
Total	0.80	0.20	1

$$P(P|M) = \frac{P(P \cap M)}{P(M)} = \frac{0.24}{0.80} = 0.3$$

$$P(P|F) = \frac{P(P \cap F)}{P(F)} = \frac{0.03}{0.20} = 0.15$$

- What is our conclusion?

Independent Events

Independent Events

Two events A and B are independent if

$$P(A|B) = P(A)$$

or

$$P(B|A) = P(B)$$

Otherwise, two events are dependent.

- What $P(A|B) = P(A)$ means?
- Intuitively, it means that event B is "irrelevant" to event A.
- $P(A|B)$ is the probability that A occurs when B already happened.
- If $P(A|B) = P(A)$, it mean that probability that A occurs does not change regardless of if B happens or not.

Example

- Consider an example. A is the event that the weather is hot in Tucson and B is the event that a car accident happens in Tucson.
- $P(A)$ is almost close 1.
- $P(A|B)$ is the probability that the weather is hot in Tucson given that an car accident happened in Tucson.
- Is $P(A|B)$ different than $P(A)$?
- No. Regardless of whether a car accident happened or not, the weather is hot in Tucson.

Example

- Consider the other example. A is the event that the weather is hot in Tucson and B is the event that a storm hits west coast.
- Two events, A and B, are definitely dependent.
- $P(A)$ is the probability that the weather is hot. It is close to 1.
- However, $P(A|B)$ is the probability that the weather is hot in Tucson given that a storm hits west coast.
- It seems reasonable to think $P(A|B) < P(A)$.

Multiplication Law

Recall that

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

We have direct implication of the above equation as follows.

Multiplication Law

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

or

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

Multiplication Law for Independent Events

We saw that, if $P(A|B) = P(A)$ or $P(B|A) = P(B)$,

two events, A and B, are independent.

If $P(A|B) = P(A)$ or $P(B|A) = P(B)$, then the multiplication law is simplified as follows.

Multiplication Law for Independent Events

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

or

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

Thus, if $P(A \cap B) = P(A)P(B)$, two events, A and B, are independent.

Outline

Consider a manufacturing company that receives important part from two different suppliers.

Denote the event that a part is from supplier 1 by A1 and the event that a part is from supplier 2 by A2.

Currently, 65% of parts are from supplier 1 and 35% of parts are from supplier 2.

Thus, if we randomly select a part, $P(A1) = 0.65$ and $P(A2) = 0.35$.

Denote the event that a part is Good by G and the event that a part is Bad by B.

The company has historical data about quality of parts from each suppliers. The data tells us

$$\begin{aligned} P(G|A1) &= 0.98 \qquad P(B|A1) = 0.02 \\ P(G|A2) &= 0.94 \qquad P(B|A2) = 0.06. \end{aligned}$$

Suppose that we randomly drawn a part. We use that part and the machine is broken down since the drawn part is bad.

Given the information that the part is bad, what is the probability that the bad part comes from supplier 1?
(In other words, what is $P(A1|B)$?)

Example Cont'd

- Recall that

$$P(A1|B) = \frac{P(A1 \cap B)}{P(B)}.$$

- Let's check what we know.
- We know $P(G|A1)$, $P(B|A1)$, $P(G|A2)$, $P(B|A1)$.
- We also know $P(A1)$ and $P(A2)$.

- What we need is two things: $P(A1 \cap B)$ and $P(B)$.
- Recall the multiplication law

$$P(A1 \cap B) = P(B \cap A1) = P(B|A1)P(A1)$$

- We can calculate $P(A1 \cap B)$ with the multiplication law since we know $P(B|A1)$ and $P(A1)$.
- Thus, $P(A1 \cap B) = P(B|A1)P(A1) = 0.02 \times 0.65$.

- Now only thing need to do is to find $P(B)$.
- Note that a bad part MUST come from either supplier 1 or supplier 2.
- Thus, $P(B) = P(A1 \cap B) + P(A2 \cap B)$.
- In words, probability that a part is bad is the sum of the probability that a part is bad and it comes from supplier 1 and the probability that a part is bad and it comes from supplier 2.
- We already calculate $P(A1 \cap B) = 0.02 \times 0.65$.
- We can also calculate $P(A2 \cap B) = P(B|A2)P(A2) = 0.06 \times 0.35$.

- Now we can calculate $P(A1|B)$.
- Again recall

$$P(A1|B) = \frac{P(A1 \cap B)}{P(B)}$$

- We know $P(A1 \cap B) = P(B|A1)P(A1)$ and $P(B) = P(B|A1)P(A1) + P(B|A2)P(A2)$ now.

Finally,

$$\begin{aligned} P(A1|B) &= \frac{P(A1 \cap B)}{P(B)} = \frac{P(B|A1)P(A1)}{P(B|A1)P(A1) + P(B|A2)P(A2)} \\ &= \frac{0.02 \times 0.65}{0.02 \times 0.65 + 0.06 \times 0.35} = \frac{0.013}{0.034} = 0.38 \end{aligned}$$

Bayes' Theorem (Two-event Case)

Bayes' Theorem (Two-event Case)

$$
\begin{aligned}
P(A1|B) &= \frac{P(B|A1)P(A1)}{P(B|A1)P(A1) + P(B|A2)P(A2)} \\
P(A2|B) &= \frac{P(B|A2)P(A2)}{P(B|A2)P(A2) + P(B|A1)P(A1)}
\end{aligned}
$$

Consider the previous example again.

- Note that before the arrival of new information (before we know that a part is bad), $P(A1) = 0.65$.
- That is, when we randomly select a part, we thought that part comes from supplier 1 with 65% before we know it is bad.
- When the machine is broken with the part we selected, it gives the information that the part we selected is bad.
- Using this information, we update the probability that this part comes from supplier 1.
- Given that the part is bad, the probability that it comes from supplier 1 is 0.38.

Outline

Exercise 1

A survey is conducted to learn about the attitude toward investment for retirement. Respondents are either Male or Female. The question, “Do you think investment for retirement is important?”, is asked to respondents, and respondents answers with only ”Yes” or “NO”. The following joint probability distribution is constructed based on the survey results.

	Male	Female	Total
Y	0.22	0.27	0.49
N	0.28	0.23	0.51
Total	0.50	0.5	1.00

a) What is the probability that a respondent is male and think investment for retirement is important?

b) What is the probability that a respondent is male?

c) What is the probability that a respondent think investment for retirement is important given that the respondent is male?

d) What is the probability that a respondent think investment for retirement is not important given that the responde is male?

Exercise 2

Visa Card USA studied how frequently young consumers, ages 18 to 24, use plastic(debit or credit) cards in making purchases(Associated Press, January 16, 2006). The results of the study provided the following probabilities.

- the probability that a consumer uses a plastic card when making a purchase is .37.
- Given that a consumer uses a plastic card, the probability that the consumer's age is between 18 to 24 is .19 .
- Given that the consumer uses a plastic card, the probability that the consumer's age is more than 24 years old is .81 .

U.S. Census Bureau data show that 14% of the consumer population is 18 to 24 years old. Assume the consumer population considered here is all above 18 years old.

a) Given that the consumer is 18 to 24 years old, what is the probability that the consumer uses plastic card?

b) Given that the consumer is over 24 years old, what is the probability that the consumer uses a plastic card?

c) What is the interpretation of the probabilities shown in parts (a) and (b)?

Exercise 3

A publisher sends advertising materials for an accounting text to 60% of all professors teaching the accounting class. 40% of the professor who received this material decided to use the text book. In addition, 10% of professors who did not receive the material also decided to use the text book.

a) What is the probability that a professor decided to use the text book?

b) What it the probability that a professor received the material given that the professor uses the text book?