Chapter 3

Conditional Probability and Independence

Yingbo Li

Clemson University

MATH 4000 / 6000

Recap

For event E in a sample space S with equally likely outcomes,

$$P(E) = \frac{\#(E)}{\#(S)}$$

Question

In the game of bridge, the entire deck of 52 cards is dealt out to 4 players. What is the probability that one of the players receives all face cards?

$$\#(\text{Player A gets all face cards}) = 40$$

$$\#(\text{one player gets all face cards}) = 40 + 40 + 40 + 40 = 160$$

$$P(\text{one player gets all face cards}) = \frac{160}{\binom{52}{13}}$$

Conditional probability: motivation

The probability of getting a one when rolling a 6-sided die is usually assumed to be $1/6$

Suppose you were given the extra information that the die roll was an odd number (hence 1, 3 or 5)

Conditional on this new information, the probability of a one is now $1/3$

Conditional probability

Definition

Given two events A and B with $P(B) > 0$, the conditional probability of A given B has occurred is defined as

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

- When B is the sample space: $P(A \mid B) = P(A)$
- Intuition: in a sample space with equally likely outcomes,

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

Example

Consider the die roll example: $B = \{1 \text{ or } 3 \text{ or } 5\}, A = \{1\}$

$$\begin{aligned}
P(\text{get 1 given that roll is odd}) =& P(A \mid B) \\
=& \frac{P(A \cap B)}{P(B)} \\
=& \frac{P(\text{get 1})}{P(\text{get 1 or 3 or 5})} \\
=& \frac{1/6}{3/6} = \frac{1}{3}
\end{aligned}$$

Example: roll two fair 6-sided dice. Find the probability of the sum of two rolls is 7, given the condition that the first roll is 4.

Denote two events $A = \{\text{Sum is 7}\}$ and $B = \{\text{First roll is 4}\}$.
Find $P(A \mid B)$. Formulas to use?

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

In context,

$$P(A \cap B) = P\left(\{(4,3)\}\right) = 1/36, \quad P(B) = 1/6$$

$$P(A \mid B) = \frac{1/36}{1/6} = \frac{1}{6}$$

A survey asked if whether voters who are familiar with the DREAM act support or oppose it.

- 32% of the respondents are Democrats,
- 51% of the respondents support the DREAM act, and
- 21% of the respondents are Democrats and support the DREAM act.

If we randomly select a respondent who supports the DREAM act, what is the probability that s/he is a Democrat?

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

$$\begin{aligned} P(\text{support}) &= 0.51 \\ P(\text{Democrat and support}) &= 0.21 \\ P(\text{Democrat} \mid \text{support}) &= \frac{0.21}{0.51} = 0.41 \end{aligned}$$

Question

At an apartment complex, 58% of the units have a washer and dryer, 32% have double parking, and 20% have both washer & dryer and double parking.

1. What percent of apartments have neither double parking nor washer and dryer?
2. A unit with double parking just became available at this apartment complex, what is the probability that it also has washer and dryer?

$$P(\text{w\&d} \cup \text{dbl prk}) \\ = 0.58 + 0.32 - 0.20 \\ = 0.70$$

$$P(\text{neither w\&d nor dbl prk}) \\ = 1 - 0.70 = 0.30$$

$$P(\text{w\&d} \mid \text{dbl parking}) \\ = \frac{P(\text{w\&d} \cap \text{dbl prk})}{P(\text{double prk})} \\ = \frac{0.20}{0.32} = 0.625$$

Propositions of conditional probability

1. $P(A \mid A) = 1$

2. $P(A^c \mid A) = 0$

3. $P(A^c \mid B) = 1 - P(A \mid B)$

Multiplication rule

- ☞ Dy the definition of conditional probability, the *joint probability* of A and B is

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

  - Usually, $P(A)$ and $P(B)$ are called *marginal probabilities*.

- ☞ Generalize to n events: chaining of probabilities

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1, A_2) \cdots P(A_n \mid A_1, \ldots, A_{n-1})$$

A box obtains 8 white balls and 4 black ones. 3 balls are drawn at random without replacement.

1. What's the probability of getting $W_1B_2B_3$?
2. What's the probability of getting black in draw 2?

1. Multiplication rule

$$P(W_1B_2B_3) = P(W_1)P(B_2 \mid W_1)P(B_3 \mid W_1B_2) = \frac{8}{12}\cdot\frac{4}{11}\cdot\frac{3}{10}$$

2. If we want to consider draw 1, which can be W_1 or B_1 (disjoint!)

$$P(B_2) = P(W_1B_2) + P(B_1B_2) = \frac{8}{12}\cdot\frac{4}{11} + \frac{4}{12}\cdot\frac{3}{11} = \frac{44}{132} = \frac{4}{12}$$

Actually we do not need to consider draw 1. The marginal probability

$$P(B_2) = \frac{4}{12}$$

Law of total probability

For events $A_1, \ldots, A_n$ are **disjoint**, and

$$\bigcup_{i=1}^{n} A_i = S,$$

then for any event B,

☞ law of total probability

$$P(B) = P(B \mid A_1)P(A_1) + \ldots + P(B \mid A_n)P(A_n)$$

- Such collection of sets $A_1, \ldots, A_n$ is also called a partition of sample space.

Example: a firm is considering a drug-testing program for its employees, but before it begins it wants to know the scope of the problem if any exists. Realizing the sensitivity of this issue, the personnel director decides to use a randomized response survey. It is felt that respondents are more likely to be honest when such forms are used. Each employee is asked to flip a fair coin. If the coin comes up heads, answer the question "Do you carpool to work?". If the coin comes up tails, answer the question "Have you used illegal drugs within the last month?". Assume that all employees answer the survey honestly. Out of 8000 responses, 1420 answered "YES". Suppose the firm knows that 35% of its employees carpool to work. What is the probability that an employee chosen at random used illegal drugs within the last month?

Let events $A_1 = \{\text{head}\}$, $A_2 = \{\text{tail}\}$, $B = \{\text{answer YES}\}$.
Since $\{A_1, A_2\}$ is a partition, so we can use the law to total probability

$$\begin{aligned} P(B) =& P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) \\ P(\text{YES}) =& P(\text{carpool})P(\text{head}) + P(\text{drug})P(\text{tail}) \\ \frac{1420}{8000} =& 0.35 \times 0.5 + P(\text{drug}) \times 0.5 \\ P(\text{drug}) =& 0.005 \end{aligned}$$

Recap

- Marginal probability: $P(A), P(B)$
- Joint probability: $P(A \cap B)$
- Conditional probability

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

- Multiplication rule:

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

- Law of total probability: for a partition $\{A_1, A_2, \ldots, A_n\}$ of S,

$$P(B) = \sum_{j=1}^{n} P(B \mid A_j) P(A_j)$$

Question

Which is the correct notation for the following probability?
"At a coffee shop you overhear a recent college graduate discussing that she doesn't believe that online courses provide the same educational value as one taken in person. What's the probability that she has taken an online course before?"

(a) *P(took online course | not valuable)*
(b) P(not valuable | took online course)
(c) P(took online course and not valuable)
(d) P(valuable | didn't take online course)

Question

My neighbor has two children. I know one of them is a son (i.e. at least one boy). What is the probability that she has two boys?

$$P(\text{both boys} \mid \text{at least one boy}) = \frac{1/4}{3/4} = \frac{1}{3}$$

Monty Hall problem: suppose you are on a game show, and you are given the choice of three doors: Behind one door is a car; behind the others, goats.

You pick a door, say No. 1, and the host, who knows whats behind the doors, opens another door, say No. 3, which he knows has a goat. He then asks you, “Do you want to switch to door No. 2?” Is it to your advantage to switch your choice?

Event A: You first pick the correct door. Event B: You win the car.
If you switch, by the law of total probability

$$\begin{aligned} P(B) &= P(B \mid A)P(A) + P(B \mid A^c)P(A^c) \\ &= 0 \times 1/3 + 1 \times 2/3 = 2/3 \end{aligned}$$

If you don't switch, by the law of total probability

$$\begin{aligned} P(B) &= P(B \mid A)P(A) + P(B \mid A^c)P(A^c) \\ &= 1 \times 1/3 + 0 \times 2/3 = 1/3 \end{aligned}$$

Good to switch.

Bayes theorem (also called Bayes rule)

Suppose events $A_1, \ldots, A_n$ are disjoint, and $\bigcup_{i=1}^{n} A_i = S$, with $P(A_i) > 0$, $i = 1, 2, \ldots, n$. Then for any event B with $P(B) > 0$,

$$\begin{aligned} P(A_i \mid B) = &\frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)P(A_j)}, \quad i = 1, \ldots, n \\ &= \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)} \end{aligned}$$

- $P(A_i)$ is often called *prior probability*
- $P(A_i \mid B)$ is called *posterior probability*.

Example: Clemson students

- Suppose Clemson has only three possible majors: math, engineering, and econ. And 25% of the students are in math, 55% are in engineering, and the rest are in econ (with no double majors).
- Campus folklore say that 90% of math majors want to go on dates, compared with 50% of engineers and 10% of econ majors.
- Your room-mate meets someone at the TD's and gets a date. What is the probability that she is dating a math major?
- Note that the majors define a finite partition, and the campus folklore gives us the conditional probabilities $\Pr(B \mid A_i)$.
- The point of Bayes' rule is to reverse the conditioning to get $\Pr(A_i \mid B)$.

Example: Suppose Clemson has only three possible majors: math, engineering, and econ. And 25% of the students are in math, 55% are in engineering, and the rest are in econ (with no double majors). Campus folklore say that 90% of math majors want to go on dates, compared with 50% of engineers and 10% of econ majors. Find $P(\text{math} \mid \text{date})$.

Events $A_1 = \{\text{math}\}, A_2 = \{\text{eng}\}, A_3 = \{\text{econ}\}$ define a partition.
Let $B = \{\text{date}\}$. In order to calculate $P(A_1 \mid B)$, use the Bayes theorem:

$$
\begin{aligned}
P(A_1 \mid B) =& \frac{P(B \mid A_1)P(A_1)}{P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + P(B \mid A_3)P(A_3)} \\
=& \frac{0.9 \times 0.25}{0.9 \times 0.25 + 0.5 \times 0.55 + 0.1 \times 0.20} = 0.43
\end{aligned}
$$

Example: diagnostic tests

Lupus is a medical phenomenon where antibodies that are supposed to attack foreign cells to prevent infections instead see plasma proteins as foreign bodies, leading to a high risk of blood clotting. It is believed that 2% of the population suffer from this disease.

The test for lupus is very accurate if the person actually has lupus, however is very inaccurate if the person does not. More specifically, the test is 98% accurate if a person actually has the disease. The test is 74% accurate if a person does not have the disease.

There is a line from the Fox television show *House*, often used after a patient tests positive for lupus: “It’s never lupus.” Do you think there is truth to this statement?

- Event $A = \{\text{lupus}\}, \quad P(A) = 0.02, \quad P(A^c) = 1 - P(A) = 0.98$
- Event $B = \{\text{positive}\}, \quad P(B \mid A) = 0.98$
- $P(\text{positive} \mid \text{no lupus}) = P(B \mid A^c) = 1 - 0.74 = 0.26$

$$\begin{aligned} P(\text{lupus} \mid \text{positive}) &= P(A \mid B) \\ &= \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A^c)P(A^c)} \\ &= \frac{0.98 \times 0.02}{0.98 \times 0.02 + 0.26 \times 0.98} \\ &= 0.0714 \end{aligned}$$

Even when a patient tests positive for lupus, there is only a 7.14% chance that he/she actually has lupus. This is because the background rate of lupus is quite low. House may be right.

http://www.ted.com/talks/peter_donnelly_shows_how_stats_fool_juries.html/ From about 11:00

Recap

Bayes theorem

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)P(A_j)}, \quad i = 1, \ldots, n$$

Question

Suppose in the Monty Hall problem, I tossed a fair coin to decide whether to stay or switch. Say I won the car, what's the probability that I switched?

Let $A_1 = \{\text{switch}\}, A_2 = \{\text{switch}\}, B = \{\text{win}\}$, using Bayes theorem

$$P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2)}$$

$$= \frac{\frac{2}{3} \cdot \frac{1}{2}}{\frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2}} = \frac{2}{3}$$

Product rule for independent events

Definition

Events A and B are independent if

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

- If A and B are independent, then

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

Knowing B doesn't affect the odds of A.

Example: you toss a coin twice, what is the probability of getting two tails in a row?

- Method 1: sample space $S = \{HH, TT, HT, TH\}$.
  Since one out of four possible outcomes match this definition, the probability is $\frac{1}{4}$.
  Inefficient way if the number of trials was much higher.
- Method 2: use independence.

$$P(\text{T on the first toss}) \times P(\text{T on the second toss}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Example: roll two fair 6-sided dice. Set

- $A = \{\text{Sum is } 7\}$
- $B = \{\text{First roll is } 5\}$
- $C = \{\text{Maximum roll is } 5\}$

Are A and B independent? How about B and C?

$$A \cap B = \{(5,2)\}, P(A \cap B) = 1/36$$

$$P(A) = 6/36 = 1/6, P(B) = 1/6$$

$$P(A \cap B) = P(A) \times P(B) \Longrightarrow A \text{ and } B \text{ are independent.}$$

$$B \cap C = \{(5,1),(5,2),(5,3),(5,4),(5,5)\}, P(B \cap C) = 5/36$$

$$P(B) = 1/6, P(C) = 9/36$$

$$P(B \cap C) \neq P(B) \times P(C) \Longrightarrow B \text{ and } C \text{ are dependent.}$$

Properties of independence

If A and B are independent, then

- B and A are independent.

- A and B^c are independent. And so are A^c and B^c.

- Independent is not disjoint.

Mutually independent

- Three events A, B, C are called *mutually independent* if
  1. $$P(A \cap B) = P(A)P(B)$$
     $$P(B \cap C) = P(B)P(C)$$
     $$P(A \cap C) = P(A)P(C)$$
  2. $$P(A \cap B \cap C) = P(A)P(B)P(C)$$
- If only (1) holds but not (2), then A, B, C are called *pairwise independent*.

Example: pairwise indpt but not mutually indpt

Roll two fair 6-sided dice. Set

- $A = \{\text{Sum is 7}\}$
- $B = \{\text{First roll is 4}\}$
- $C = \{\text{Second roll is 3}\}$

$$\underset{1/36}{P(A \cap B)} \overset{?}{=} \underset{1/6}{P(A)} \times \underset{1/6}{P(B)}$$

$$\underset{1/36}{P(B \cap C)} \overset{?}{=} \underset{1/6}{P(B)} \times \underset{1/6}{P(C)}$$

A, B, C are pairwise indpt

$$\underset{1/36}{P(A \cap C)} \overset{?}{=} \underset{1/6}{P(A)} \times \underset{1/6}{P(C)}$$

$$\underset{1/36}{P(A \cap B \cap C)} \overset{?}{=} \underset{1/6}{P(A)} \times \underset{1/6}{P(B)} \times \underset{1/6}{P(C)}$$

NOT mutually indpt

☞ Events $A_1, A_2, \ldots, A_n$ are called *mutually independent* (or just independent) if for any $1 \leq r \leq n$ of them

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_r})$$

☞ The key to compute the probability of independent events is to just multiply the probability of the individual events.

Question

A fair coin is tossed 5 times. Compute $P(5 \text{ tails})$.

$$P(T_1\ T_2\ T_3\ T_4\ T_5) = P(T_1)P(T_2)P(T_3)P(T_4)P(T_5) = \frac{1}{2^5}$$

Example: two fair dice are rolled independently until a sum of 5 or 7 is obtained. What are the probability the trials end with a sums of 5?

$E_n = \{n^{\text{th}}$ trial is 5, and the first $n-1$ trials do not have a sum of 5 or 7$\}$

$$P(\{\text{Game ends in 5}\}) = P(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n)$$

Let S_i denote the i^{th} sum.

$$P(E_n) = P(\{S_1 \neq 5, 7\}) \cdots P(\{S_{n-1} \neq 5, 7\}) P(\{S_n = 5\})$$

$$P(\{S_i = 5\}) = \frac{4}{36}, P(\{S_i = 7\}) = \frac{6}{36}, P(\{S_i \neq 5, 7\}) = \frac{13}{18}$$

$$P(\{\text{Game ends in 5}\}) = \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} \frac{1}{9} = \left(\frac{1}{1 - \frac{13}{18}}\right) \frac{1}{9} = 0.4$$