STAT 515 fa 2023 Lec 3

Conditional probability, independence, and Bayes' rule

Karl Gregory

Conditional probability

Two events can be related to one another such that the occurrence of one event makes the occurrence of the other event more or less likely.

Definition: Conditional probability

The *conditional probability* of the event A given the event B is the probability that the event A occurs given that B occurs. We denote this as $P(A|B)$ and call it "the probability of A given B". It is equal to

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

Example. On the day of a test, select a student at random and let A be the event that he or she scores 90% or higher.

1. Now consider only students who completed the last homework assignment. One would expect the probability of the event A to be higher when sampling a student from this group than when sampling from the whole class.

2. Now consider only students who *did not* complete the last homework assignment. One would expect the probability of the event A to be *lower* when sampling a student from this group than when sampling from the whole class.

Consider selecting a student at random from the whole class and letting B be the event that the selected student completed the last homework assignment. Then, corresponding to 1 and 2 above:

1. The conditional probability $P(A|B)$ is the probability that a student selected at random from the group of students who completed the last homework assignment will score 90% or higher on the test. We expect this to be higher than $P(A)$, which is the probability that a student drawn from the whole class will score 90% or higher on the test.

2. The conditional probability $P(A|B^c)$ is the probability that a student selected at random from the group of students who *did not* complete the last homework assignment will score 90% or higher on the test. We expect this to be lower than $P(A)$, which is the probability that a student drawn from the whole class will score 90% or higher on the test.

Suppose that two-thirds of the students in the class completed the last homework assignment and that one half of the class completed the last homework assignment and scored 90% or higher on the test. Then

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/2}{2/3} = 0.75.$$

Example. Suppose friend 1 of yours shows up to parties with probability 0.4 and friends 1 and 2 of yours both show up to a party with probability 0.35. Then the probability that friend 2 shows up to a party given that friend 1 shows up is

$$\begin{aligned} P(\text{friend 2 shows up}|\text{friend 1 shows up}) &= \frac{P(\text{friends 1 and 2 both show up})}{P(\text{friend 1 shows up})} \\ &= \frac{0.35}{0.4} \\ &= 0.875. \end{aligned}$$

Example. Suppose that 80,250 spectators are in attendance at the next Gamecocks vs Aggies football game, and that one lucky winner will be selected at random to receive free "big gulps" for life from the concession stand. Suppose that 20,000 of the spectators are students, and that there are 10,000 Aggies present, 200 of which are students.

We can begin making a table of the spectators based on the provided information as follows

	Gamecock	Aggie	total
student		200	20,000
non-student			
total		10,000	80,250

Filling in the table, we get

	Gamecock	Aggie	total
student	19,800	200	20,000
non-student	50,450	9,800	60,250
total	70,250	10,000	80,250

1. Let G be the event that a Gamecock is chosen. Since each person is equally likely to be chosen, $P(G) = 70{,}250/80{,}250 = 0.8754$.

2. Let S be the event that a student is chosen. In the same way, $P(S) = 20{,}000/80{,}250 = 0.2492$

3. The event G^c is the event that an Aggie is chosen. We have $P(G^c) = 1 - P(G) = 0.1246$.

4. The event that an Aggie student is chosen is $G^c \cap S$. This has probability $P(G^c \cap S) = 200/80{,}250 = 0.0025$.

5. Given that a student is chosen, what is the probability that he or she will be a Gamecock? The answer is
$$P(G|S) = \frac{P(G \cap S)}{P(S)},$$
so we need $P(G \cap S) = 19{,}800/80{,}250$ as well as $P(S) = 20{,}000/80{,}250$. The answer is thus
$$P(G|S) = P(G \cap S) \times \frac{1}{P(S)} = \frac{19{,}800}{80{,}250} \times \frac{80{,}250}{20{,}000} = \frac{19{,}800}{20{,}000} = 0.99.$$

6. Given that an Aggie is chosen, what is the probability that he or she will be a student? The answer is
$$P(S|G^c) = \frac{P(G^c \cap S)}{P(G^c)} = \frac{200}{10{,}000} = 0.02.$$

A direct consequence of the definition of the conditional probability is the following.

Probability rule: Multiplicative rule of probability

For any two events A and B,
$$P(A \cap B) = P(A|B)P(B) \quad \text{or equivalently} \quad P(A \cap B) = P(B|A)P(A).$$

Example. Suppose you go on a safari in South Africa and you are interested in which animals are present at a certain riverbank, namely the giraffe (G), the wildebeest (W), and the crocodile (C). Suppose that on a given Safari

$$P(W) = 0.40$$

$$P(C) = 0.60$$

$$P(G) = 0.20$$

$$P(C|W) = 0.775$$

$$P(C|G) = 0.65$$

$$P(G \cap W) = 0.06$$

$$P(G \cap W \cap C) = 0.01$$

Then we can fill in all the spaces on a Venn Diagram. We will need

$$P(W \cap C) = P(C|W)P(W) = 0.775 \times 0.40 = 0.31$$

$$P(G \cap C) = P(C|G)P(G) = 0.65 \times 0.20 = 0.13$$

Independent events

Events which have nothing to do with other are *independent*. Formally:

Definition: Independence

The following definitions of *independence* are equivalent:

- Two events A and B are *independent* if $P(A|B) = P(A)$.
- Two events A and B are *independent* if $P(B|A) = P(B)$.
- Two events A and B are *independent* if $P(A \cap B) = P(A)P(B)$.

Exercise. Flip a coin twice and record the flips. Let A_1 be the event that flip one is heads and A_2 be the event that flip two is heads. Our intuition tells us that the two flips should be independent; the outcome of the first flip has no bearing on the second flip, so we should have $P(A_2|A_1) = P(A_2) = 1/2$. So, what is the probability that both flips are heads?

Answer:

$$P(A_1 \cap A_2) = P(A_1)P(A_2) = 1/2 \times 1/2 = 1/4.$$

Example. Let A be the event that your bike gets a flat tire on a given day and let B be the event that you have forgotten to bring a spare tube. Suppose that $P(A) = 0.02$ and $P(B) = 0.10$ and $P(A \cap B) = 0.002$. Then the events are independent, since $P(A \cap B) = P(A)P(B)$.

Example. Suppose you send out a survey to 10 randomly selected people and suppose that each person will complete the survey with probability 0.20. Define the events

$$\begin{array}{ll} R_1: & \text{1st person responds} \\ \vdots & \vdots \\ R_{10}: & \text{10th person responds} \end{array}$$

1. What is the probability that everyone completes the survey?

$$\begin{aligned} P(\text{everyone completes survey}) &= P(R_1 \cap R_2 \cap \cdots \cap R_{10}) \\ &= P(R_1) \times P(R_2) \times \cdots \times P(R_{10}) \\ &= 0.20^{10} \\ &= 0.0000001024 = 1.024 \times 10^{-7} \end{aligned}$$

2. What is the probability that no one completes the survey?

$$\begin{aligned} P(\text{no one completes survey}) &= P(R_1^c \cap R_2^c \cap \cdots \cap R_{10}^c) \\ &= P(R_1^c) \times P(R_2^c) \times \cdots \times P(R_{10}^c) \\ &= 0.80^{10} \\ &= 0.1074 \end{aligned}$$

This last example illustrates the rule that for K independent events $A_1, \ldots, A_K$,

$$P\left(\bigcap_{k=1}^{K} A_k\right) = \prod_{k=1}^{K} P(A_k),$$

where

$$\bigcap_{k=1}^{K} A_k = A_1 \cap A_2 \cap \cdots \cap A_K$$

and

$$\prod_{k=1}^{K} P(A_k) = P(A_1) \times P(A_2) \times \ldots P(A_K).$$

Figure 1: Thomas Bayes (1701 – 1761)

Bayes' rule

Bayes' rule is a useful re-expression of the definition of conditional probability. Before stating Bayes' rule, note that for any events A and B, we have

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

Draw a Venn Diagram to convince yourself! We can also write

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

using the definition of conditional probability. This allows us to write

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}, \tag{1}$$

which is an instance of Bayes' rule. More generally, Bayes' rule is the following:

Probability rule: Bayes' rule

For an event A and a set of mutually exclusive and exhaustive events $B_1, \ldots, B_K$ such that $P(B_1) + \cdots + P(B_K) = 1$, we have

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A|B_1)P(B_1) + \cdots + P(A|B_K)P(B_K)}.$$

Remark 1. *We get the instance of Bayes' rule from (1) by letting $K = 2$ and setting $B_1 = B$ and $B_2 = B^c$. Also, the word "exhaustive" means that the union of the events $B_1, \ldots, B_K$ is equal to the whole sample space; that is, one of the events* must *happen. If $K = 2$ and we set $B_1 = B$ and $B_2 = B^c$, then B and B^c comprise a set of mutually exclusive and exhaustive events.*

Draw a Venn Diagram where A is a circle in the middle and the $B_1, \ldots, B_K$ are cells covering the whole square to see that

$$P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_K)P(B_K).$$

Exercise. The National Cancer Institute estimates that 3.65% of women in their sixties develop breast cancer. Suppose that if a woman has breast cancer, a mammogram will detect it 85% of the time and if a woman does not have breast cancer, the mammogram still results in a positive test for breast cancer 5% of the time.

1. What is the probability that a woman has breast cancer given that her mammogram is positive for breast cancer?

   Answer: We want to find $P(\text{Cancer}|\text{Pos})$. We know from Bayes' rule that

$$\begin{aligned} P(\text{Cancer}|\text{Pos}) &= \frac{P(\text{Pos}|\text{Cancer})P(\text{Cancer})}{P(\text{Pos}|\text{Cancer})P(\text{Cancer}) + P(\text{Pos}|\text{No cancer})P(\text{No cancer})} \\ &= \frac{0.85 \times 0.0365}{0.85 \times 0.0365 + 0.05 \times (1 - 0.0365)} \\ &= 0.319 \end{aligned}$$

2. What is the probability that a randomly selected woman tests positive for breast cancer?

   Answer: We have computed this probability already; it is

$$\begin{aligned} P(\text{Pos}) &= P(\text{Pos} \cap \text{Cancer}) + P(\text{Pos} \cap \text{No cancer}) \\ &= P(\text{Pos}|\text{Cancer})P(\text{Cancer}) + P(\text{Pos}|\text{No cancer})P(\text{No cancer}) \\ &= 0.85 \times 0.0365 + 0.05 \times (1 - 0.0365) \\ &= 0.0792. \end{aligned}$$

Exercise. Suppose that 1.0% of all Craigslist advertisements for housing sublets are posted by scammers. Suppose 90% of scammers request that payment be sent to a P.O. box instead of a residential address, and only 20% of non-scammers, that is legitimate advertisers, request that payment be sent to a P.O. box instead of a residential address.

1. What is the probability that a randomly selected Craigslist advertiser for a housing sublet will request that payment be sent to a P.O. box?

Answer: We have

$$\begin{aligned} P(\text{P.O}) &= P(\text{P.O.} \cap \text{scam}) + P(\text{P.O.} \cap \text{not a scam}) \\ &= P(\text{P.O.}|\text{scam})P(\text{scam}) + P(\text{P.O.}|\text{not a scam})P(\text{not a scam}) \\ &= 0.90 \times 0.01 + 0.20 \times (1 - 0.01) \\ &= 0.207. \end{aligned}$$

2. If you make an inquiry into a housing sublet post and the advertiser requests that payment be sent to a P.O. box, what is the probability that it is a scam?

   Answer: We have

$$\begin{aligned} P(\text{scam}|\text{P.O.}) &= \frac{P(\text{scam} \cap \text{P.O.})}{P(\text{P.O.})} \\ &= \frac{P(\text{P.O.}|\text{scam})P(\text{scam})}{P(\text{P.O.})} \\ &= \frac{0.90 \times 0.01}{0.207} \\ &= 0.0435, \end{aligned}$$

   where we have used Bayes' formula:

$$P(\text{scam}|\text{P.O.}) = \frac{P(\text{P.O.}|\text{scam})P(\text{scam})}{P(\text{P.O.}|\text{scam})P(\text{scam}) + P(\text{P.O.}|\text{not a scam})P(\text{not a scam})}.$$

Exercise. Let's play a game. There are three bags, each containing four bills as follows:

Bag 1:	3 five dollar bills and 1 twenty dollar bill
Bag 2:	2 five dollar bills and 2 twenty dollar bills
Bag 3:	1 five dollar bill and 3 twenty dollar bills

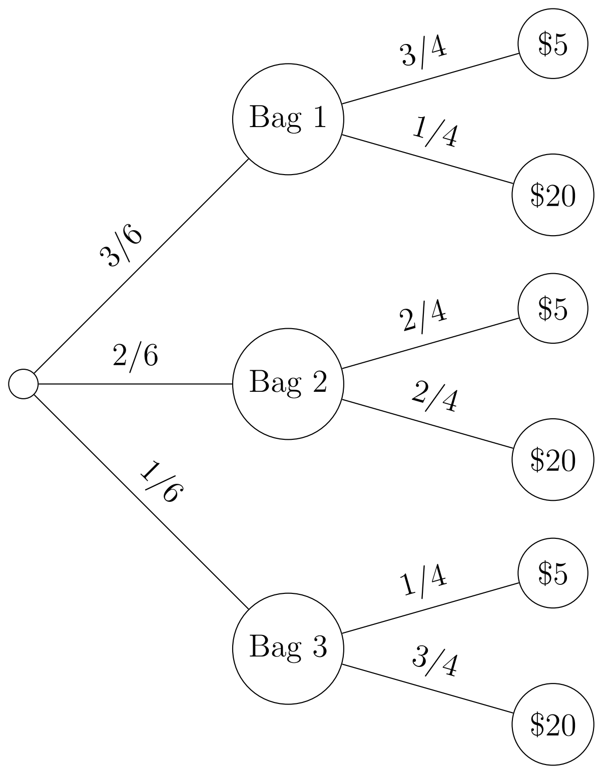
You roll a die and then draw a bill from Bag 1 if the roll is 1,2, or 3, from Bag 2 if the roll is 3 or 4, and from Bag 3 if the roll is 6.

1. What is the probability that you play the game and get twenty dollars?

   Answer: We have

$$\begin{aligned} P(\$20) &= P(\$20 \cap \text{Bag 1}) + P(\$20 \cap \text{Bag 2}) + P(\$20 \cap \text{Bag 3}) \\ &= P(\$20|\text{Bag 1})P(\text{Bag 1}) + P(\$20|\text{Bag 2})P(\text{Bag 2}) + P(\$20|\text{Bag 3})P(\text{Bag 3}) \\ &= (1/4)(3/6) + (2/4)(2/6) + (3/4)(1/6) \\ &= 5/12. \end{aligned}$$

This is an example in which it can be useful to draw what is called a tree diagram like the one below:



This diagram nicely summarizes all the information we are given. To find the probability of an outcome, we trace the tree diagram starting from the beginning until we reach that outcome, and multiplying the probabilities we pass along the way. For example, $P(\$20 \cap \text{Bag } 1) = (3/6)(1/4)$.

2. Given that a player got twenty dollars, what is the probability that the player drew from Bag 1?

Answer: We have

$$\begin{aligned} P(\text{Bag } 1|\$20) &= \frac{P(\$20|\text{Bag } 1)P(\text{Bag } 1)}{P(\$20|\text{Bag } 1)P(\text{Bag } 1) + P(\$20|\text{Bag } 2)P(\text{Bag } 2) + P(\$20|\text{Bag } 3)P(\text{Bag } 3)} \\ &= \frac{(1/4)(3/6)}{5/12} \\ &= 3/10. \end{aligned}$$