

Examples: Conditional Probability

Definition: If $P(F) > 0$, then the probability of E given F is defined to be $P(E|F) = \frac{P(E \cap F)}{P(F)}$.

Example 1 A machine produces parts that are either good (90%), slightly defective (2%), or obviously defective (8%). Produced parts get passed through an automatic inspection machine, which is able to detect any part that is obviously defective and discard it. What is the quality of the parts that make it through the inspection machine and get shipped?

Let G (resp., SD , OD) be the event that a randomly chosen shipped part is good (resp., slightly defective, obviously defective). We are told that $P(G) = .90$, $P(SD) = 0.02$, and $P(OD) = 0.08$.

We want to compute the probability that a part is good *given* that it passed the inspection machine (i.e., it is *not* obviously defective), which is

$$P(G|OD^c) = \frac{P(G \cap OD^c)}{P(OD^c)} = \frac{P(G)}{1 - P(OD)} = \frac{.90}{1 - .08} = \frac{90}{92} = .978$$

Example 2 Your neighbor has 2 children. You learn that he has a son, Joe. What is the probability that Joe's sibling is a brother?

The “obvious” answer that Joe's sibling is equally likely to have been born male or female suggests that the probability the other child is a boy is $1/2$. This is not correct!

Consider the experiment of selecting a random family having two children and recording whether they are boys or girls. Then, the sample space is $S = \{BB, BG, GB, GG\}$, where, e.g., outcome “BG” means that the first-born child is a boy and the second-born is a girl. Assuming boys and girls are equally likely to be born, the 4 elements of S are equally likely.

The event, E , that the neighbor has a son is the set $E = \{BB, BG, GB\}$. The event, F , that the neighbor has two boys (i.e., Joe has a brother) is the set $F = \{BB\}$.

We want to compute

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{P(\{BB\})}{P(\{BB, BG, GB\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

[See Example 3m of Ross(8th Ed) (or Example 3l of Ross(7th Ed)), which gives a detailed discussion of how the solution to this type of problem is affected by assumptions we are making in solving it!]

Example 3 Your neighbor has 2 children. He picks one of them at random and comes by your house; he brings a boy named Joe (his son). What is the probability that Joe's sibling is a brother?

This example differs from the above example: There is a mechanism by which the son was selected that gave you the information that your neighbor has a boy (the mechanism was random selection).

Think of it this way: In the first example, we were given the event E that “your neighbor has a son”. Now, we are given the event E' that “your neighbor randomly chose one of his 2 children, and that chosen child is a son”. We note that $E' \subseteq E$, since event E' happening implies that event E happens (if he chose a son at random, then you know for sure that he has a son!). It does not go the other way, though: E' does not imply E (just because he has a son does not mean that he chose that son at random).

We want to compute

$$\begin{aligned} P(F|E') &= \frac{P(F \cap E')}{P(E')} = \frac{P(\{BB\})}{P(E'|\{BB\})P(\{BB\}) + P(E'|\{BG\})P(\{BG\}) + P(E'|\{GB\})P(\{GB\}) + P(E'|\{GG\})P(\{GG\})} \\ &= \frac{1/4}{1(1/4) + (1/2)(1/4) + (1/2)(1/4) + 0(1/4)} = \frac{1/4}{1/2} = \frac{1}{2}. \end{aligned}$$

Multiplication Rule: The “Multiplication Rule” (also known as the “Law of Multiplication”) states that, assuming $P(F) > 0$,

$$P(E \cap F) = P(F) \cdot P(E|F),$$

which is (trivially) just a rewriting of the definition of conditional probability. The more general form is equally easy to prove from the definition:

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1 \cap E_2) \cdots P(E_n|E_1 \cap E_2 \cap \cdots \cap E_{n-1}).$$

[See also Example 2h of Ross(8th Ed) (or Example 2g of Ross(7th Ed)).

Example 4 Suppose that five good fuses and two defective ones have been mixed up. To find the defective fuses, we test them one-by-one, at random and without replacement. What is the probability that we are lucky and find both of the defective fuses in the first two tests?

Let D_1 (resp., D_2) be the event that we find a defective fuse in the first (resp., second) test. We want to compute

$$P(D_1 \cap D_2) = P(D_1)P(D_2|D_1) = \frac{2}{7} \cdot \frac{1}{6} = \frac{1}{21}.$$

Example 5 If six cards are selected at random (without replacement) from a standard deck of 52 cards, what is the probability there will be no pairs? (two cards of the same denomination)

Let E_i be the event that the first i cards have no pair among them. Then we want to compute $P(E_6)$, which is actually the same as $P(E_1 \cap E_2 \cap \cdots \cap E_6)$, since $E_6 \subset E_5 \subset \cdots \subset E_1$, implying that $E_1 \cap E_2 \cap \cdots \cap E_6 = E_6$. We get

$$P(E_1 \cap E_2 \cap \cdots \cap E_6) = P(E_1)P(E_2|E_1) \cdots = \frac{52}{52} \frac{48}{51} \frac{44}{50} \frac{40}{49} \frac{36}{48} \frac{32}{47}.$$

[Alternatively, one can solve the problem directly using counting techniques: Define the sample space to be (equally likely) ordered sequences of 6 cards; then, $|S| = 52 \cdot 51 \cdot 50 \cdots 47$, and the event E_6 has $52 \cdot 48 \cdot 44 \cdots 32$ elements.]

Law of Total Probability: The “Law of Total Probability” (also known as the “Method of Conditioning”) allows one to compute the probability of an event E by conditioning on *cases*, according to a partition of the sample space.

For example, one way to partition S is to break into sets F and F^c , for any event F . This gives us the simplest form of the law of total probability:

$$P(E) = P(E \cap F) + P(E \cap F^c) = P(E|F)P(F) + P(E|F^c)P(F^c).$$

More generally for any partition of S into sets F_1, \dots, F_n ,

$$P(E) = \sum_{i=1}^n P(E|F_i)P(F_i).$$

Example 6 (Parts Inspection) Consider the parts problem again, but now assume that a one-year warranty is given for the parts that are shipped to customers. Suppose that a good part fails within the first year with probability 0.01, while a slightly defective part fails within the first year with probability 0.10. What is the probability that a customer receives a part that fails within the first year and therefore is entitled to a warranty replacement?

From before, we know that $P(G) = \frac{90}{92}$ and $P(SD) = \frac{2}{92}$. Let E be the event that a randomly selected customer’s part fails in the first year. We are told that $P(E|G) = .01$ and $P(E|SD) = 0.10$. We want to compute

$$P(E) = P(E|G)P(G) + P(E|SD)P(SD) = (.01)\frac{90}{92} + (.10)\frac{2}{92} = \frac{11}{920} = 0.012$$

Example 7 Two cards from an ordinary deck of 52 cards are missing. What is the probability that a random card drawn from this deck is a spade?

Let E be the event that the randomly drawn card is a spade. Let F_i be the event that i spades are missing from the 50-card (defective) deck, for $i = 0, 1, 2$.

We want $P(E)$, which we compute by conditioning on how many spades are missing from the original (good) deck:

$$P(E) = P(E|F_0)P(F_0) + P(E|F_1)P(F_1) + P(E|F_2)P(F_2) = \frac{13}{50} \frac{\binom{13}{0}\binom{39}{2}}{\binom{52}{2}} + \frac{12}{50} \frac{\binom{13}{1}\binom{39}{1}}{\binom{52}{2}} + \frac{11}{50} \frac{\binom{13}{2}\binom{39}{0}}{\binom{52}{2}} = \frac{1}{4}$$

[It makes sense that the probability is simply 1/4, since each card is equally likely to be among the lost cards, so why should the chance of getting a spade be changed?]

Bayes Formula. Often, for a given partition of S into sets F_1, \dots, F_n , we want to know the probability that some particular case, F_j , occurs, given that some event E occurs. We can compute this easily using the definition:

$$P(F_j|E) = \frac{P(F_j \cap E)}{P(E)}.$$

Now, using the Multiplication Rule (i.e., the definition of conditional probability), we can rewrite the numerator: $P(F_j \cap E) = P(E|F_j)P(F_j)$. Using the Law of Total Probability, we can rewrite the denominator: $P(E) = \sum_{i=1}^n P(E|F_i)P(F_i)$. Thus, we get

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}.$$

This expression is often called *Bayes Formula*, though I almost always just derive it from scratch with each problem we solve. See Proposition 3.1 of Ross.

Example 8 *Urn 1 contains 5 white balls and 7 black balls. Urn 2 contains 3 whites and 12 black. A fair coin is flipped; if it is Heads, a ball is drawn from Urn 1, and if it is Tails, a ball is drawn from Urn 2. Suppose that this experiment is done and you learn that a white ball was selected. What is the probability that this ball was in fact taken from Urn 2? (i.e., that the coin flip was Tails)*

Let T be the event that the coin flip was Tails. Let W be the event that a white ball is selected. From the given data, we know that $P(W|T) = 3/15$ and that $P(W|T^c) = 5/12$. Since the coin is fair, we know that $P(T) = P(T^c) = 1/2$.

We want to compute $P(T|W)$, which we do using the definition (and the same simple manipulation that results in Bayes Formula):

$$P(T|W) = \frac{P(T \cap W)}{P(W)} = \frac{P(W|T)P(T)}{P(W|T)P(T) + P(W|T^c)P(T^c)} = \frac{(3/15)(1/2)}{(3/15)(1/2) + (5/12)(1/2)} = \frac{12}{37}$$

Example 9 *One half percent of the population has a particular disease. A test is developed for the disease. The test gives a false positive 3% of the time and a false negative 2% of the time. (a). What is the probability that Joe (a random person) tests positive? (b). Joe just got the bad news that the test came back positive; what is the probability that Joe has the disease?*

Let D be the event that Joe has the disease. Let T be the event that Joe's test comes back positive.

We are told that $P(D) = 0.005$, since 1/2% of the population has the disease, and Joe is just an average guy. We are also told that $P(T|D) = .98$, since 2% of the time a person having the disease is missed ("false negative"). We are told that $P(T|D^c) = .03$, since there are 3% false positives.

(a). We want to compute $P(T)$. We do so by conditioning on whether or not Joe has the disease:

$$P(T) = P(T|D)P(D) + P(T|D^c)P(D^c) = (.98)(.005) + (.03)(.995)$$

(b). We want to compute

$$P(D|T) = \frac{P(D \cap T)}{P(T)} = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} = \frac{(.98)(.005)}{(.98)(.005) + (.03)(.995)} \approx .14$$

What a relief! There is only a 14% chance Joe has the disease, even though the test came back positive! The issue here is that the false-positive and false-negative percentages are in fact high, relative to the occurrence of the disease.

You can draw a tree diagram to illustrate the cases, first branching according to whether Joe has the disease or not, then branching according to whether the test is positive or not.

Example 10 *Consider the game of Let's Make a Deal in which there are three doors (numbered 1, 2, 3), one of which has a car behind it and two of which are empty (have "booby prizes"). You initially select Door 1, then, before it is opened, Monty Hall tells you that Door 3 is empty (has a booby prize). You are then given the option to switch your selection from Door 1 to the unopened Door 2. What is the probability that you will win the car if you switch your door selection to Door 2? Also, compute the probability that you will win the car if you do not switch. (What would you do?)*

$S = \{1, 2, 3\}$, where outcome " i " means that the car is behind door i . Let $E = \{\text{Door 3 is empty}\} = \{1, 2\}$. The probability that you win by switching to Door 2, given that he tells you Door 3 is empty is:

$$P(\{2\} | \{1, 2\}) = \frac{P(\{2\} \cap \{1, 2\})}{P(\{1, 2\})} = \frac{1/3}{2/3} = \frac{1}{2}$$

Similarly, the probability that you win by staying with Door 1, given that he tells you Door 3 is empty, is $1/2$. Thus, it does not matter if you switch or stay!

The difference between this situation and that of the Monty Hall problem done in class, in which Monty “shows” an empty door, according to a specific set of guidelines (he must pick an empty one to show, and if he has a choice, he picks at random), is this: the event “Door 3 is empty” is *not* the same event as “Monty shows Door 3 is empty”. (If Monty shows Door 3 is empty, then Door 3 is empty; but the converse is not necessarily true: just because Door 3 is empty does not mean that Monty shows you an empty Door 3.)

Example 11 Consider the game of Let’s Make a Deal in which there are five doors (numbered 1, 2, 3, 4, and 5), one of which has a car behind it and four of which are empty (have “booby prizes”). You initially select Door 1, then, before it is opened, Monty Hall opens two of the other doors that are empty (selecting the two at random if there are three empty doors among $\{2, 3, 4, 5\}$). (We are assuming that Monty Hall knows where the car is and that he selects doors to open only from among those that are empty.) You are then given the option to switch your selection from Door 1 to one of the two remaining closed doors. Given that Monty opens Door 2 and Door 4, what is the probability that you will win the car if you switch your door selection to Door 3? Also, compute the probability that you will win a car if you do not switch. (What would you do?)

$S = \{1, 2, 3, 4, 5\}$, where outcome “ i ” means that the car is behind door i . Let $E = \{\text{Monty shows you doors 2 and 4}\}$. The probability that you win by switching to door 3, given that he shows you doors 2 and 4 is:

$$\begin{aligned} P(\{3\} | E) &= \frac{P(E \cap \{3\})}{P(E)} = \frac{P(E | \{3\})P(\{3\})}{P(E)} \\ &= \frac{P(E | \{3\})P(\{3\})}{P(E | \{1\})P(\{1\}) + P(E | \{2\})P(\{2\}) + P(E | \{3\})P(\{3\}) + P(E | \{4\})P(\{4\}) + P(E | \{5\})P(\{5\})} \\ &= \frac{(1/\binom{3}{2})(1/5)}{(1/\binom{4}{2})(1/5) + 0 + (1/\binom{3}{2})(1/5) + 0 + (1/\binom{3}{2})(1/5)} = \frac{2}{5} \end{aligned}$$

The probability that you win by staying with door 1, given that he shows you doors 2 and 4 is:

$$\begin{aligned} P(\{1\} | E) &= \frac{P(E \cap \{1\})}{P(E)} = \frac{P(E | \{1\})P(\{1\})}{P(E)} \\ &= \frac{P(E | \{1\})P(\{1\})}{P(E | \{1\})P(\{1\}) + P(E | \{2\})P(\{2\}) + P(E | \{3\})P(\{3\}) + P(E | \{4\})P(\{4\}) + P(E | \{5\})P(\{5\})} \\ &= \frac{(1/\binom{4}{2})(1/5)}{(1/\binom{4}{2})(1/5) + 0 + (1/\binom{3}{2})(1/5) + 0 + (1/\binom{3}{2})(1/5)} = \frac{1}{5} \end{aligned}$$

So, I don’t know about you, but I would certainly switch!

Definition: We say that events E and F are *independent* if and only if $P(E \cap F) = P(E) \cdot P(F)$, or, equivalently, $P(E | F) = P(E)$, or, equivalently, $P(F | E) = P(F)$. (See also Section 3.4 of Ross (7th or 8th edition), Examples 4a, 4b, 4c, 4d, 4e, 4f, 4g, and 4h.) Intuitively, two events are independent if knowledge of one event happening does not affect the probability of occurrence of the other event.

Example 12 Let A and B be independent events with $P(A) = \frac{1}{4}$ and $P(A \cup B) = 2P(B) - P(A)$. Find (a). $P(B)$; (b). $P(A|B)$; and (c). $P(B^c|A)$.

We know that $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B)$ (since A, B are independent.)

Thus, we know that $1/4 + P(B) - (1/4)P(B) = 2P(B) - 1/4$, implying that $P(B) = 2/5$. Also, since A and B are independent, $P(A|B) = P(A) = 1/4$. Further, $P(B^c \cap A) + P(B \cap A) = P(A) = 1/4$, so

$$P(B^c|A) = \frac{P(B^c \cap A)}{P(A)} = \frac{1/4 - P(B \cap A)}{1/4} = \frac{1/4 - (2/5)(1/4)}{1/4} = \frac{3}{5}.$$

Example 13 Let three fair coins be tossed. Let $A = \{\text{all heads or all tails}\}$, $B = \{\text{at least two heads}\}$, and $C = \{\text{at most two tails}\}$. Of the pairs of events, (A, B) , (A, C) , and (B, C) , which are independent and which are dependent? (Justify.)

The events can be written explicitly: $A = \{HHH, TTT\}$, $B = \{HHH, HHT, HTH, THH\}$, $C = \{HHH, HHT, HTH, THH, HTT, THT, TTH\}$. $P\{A \cap B\} = 1/8 = (2/8)(4/8) = P\{A\} \cdot P\{B\}$, so A and B are indep. $P\{A \cap C\} = 1/8 \neq (2/8)(7/8) = P\{A\} \cdot P\{C\}$, so A and C are dependent. $P\{B \cap C\} = 4/8 \neq (4/8)(7/8) = P\{B\} \cdot P\{C\}$, so B and C are dependent.

Example 14 (Ross, Section 3.4, Example 4h) Consider independent trials consisting of rolling a pair of fair dice, over and over. What is the probability that a sum of 5 appears before a sum of 7?

Let E be the event that we see a sum of 5 before a sum of 7. We want to compute $P(E)$.

The easy way to solve this is by conditioning on the outcome of the first roll: Let F be the event that the first roll is a 5; let G be the event that the first roll is a 7; let H be the event that the first roll is a sum other than 5 or 7. Then F, G, H partition the sample space (a set of possible cases).

Then, by the law of total probability:

$$P(E) = P(E|F)P(F) + P(E|G)P(G) + P(E|H)P(H).$$

Now, we know that $P(F) = 4/36$, $P(G) = 6/36$, and $P(H) = 26/36$. Also, given that the first roll is a 5, the probability we get a 5 before a 7 is 1: $P(E|F) = 1$. Similarly, given that the first roll is a 7, the probability we get a 5 before a 7 is 0: $P(E|G) = 0$. Now, if the first roll is neither a 5 nor a 7, we can think of the process starting all over again: the chance we get a 5 before a 7 is just like it was ($P(E)$) before we started rolling: $P(E|H) = P(E)$.

Thus,

$$P(E) = P(E|F)P(F) + P(E|G)P(G) + P(E|H)P(H) = 1 \cdot (4/36) + 0 \cdot (6/36) + P(E) \cdot (26/36),$$

which gives us an equation in one unknown ($P(E)$): $P(E) = 4/36 + (26/36)P(E)$, so solving for $P(E)$ we get $P(E) = 2/5$.

Other Textbook Examples: Other Chapter 3 examples to read carefully from Ross (7th edition): 2a–2g, 3a–3g, 3i–3n, 4a–4h; also try doing the Self-Test at the end of the chapter (you can skip problems that are more advanced, such as 16, 23, and 24 in Ross(7th Ed)).