

Section 3.2 Conditional Probabilities

EXAMPLES:

1. Consider rolling a die. We know that $S = \{1, 2, 3, 4, 5, 6\}$. Let E be the event that the outcome is in $\{2, 3, 6\}$. We know that $P(E) = 1/2$. Now suppose I tell you that the roll was an even number. Given this knowledge, what is the probability that E occurred?

Solution: Let $F = \{2, 4, 6\}$. Since we know that each of 2, 4, 6 is equally likely, the probability of E now appears to be

$$\text{Cond. Prob. of 2} + \text{Cond. Prob. of 3} + \text{Cond. Prob. of 6} = \frac{1}{3} + 0 + \frac{1}{3} = \frac{2}{3}$$

2. Now consider rolling an unfair die. We know that $S = \{1, 2, 3, 4, 5, 6\}$, but assume that $P(\{i\}) = 0.1$ for $i \in \{1, 2, 3, 4, 5\}$ and $P(\{6\}) = 0.5$. Let E be the event that the outcome is in $\{2, 3, 6\}$. We know that

$$P(E) = P(\{2\}) + P(\{3\}) + P(\{6\}) = 0.7$$

Now suppose I tell you that the roll was an even number. Given this knowledge, what is the probability that E occurred?

Solution: Let $F = \{2, 4, 6\}$. Clearly, the outcome $\{3\}$ did not occur, so that has zero probability now. What about $\{2\}$ and $\{6\}$? We know that $P(F) = 0.7$ and that $P(\{2\}) = 0.1$, $P(\{4\}) = 0.1$, $P(\{6\}) = 0.5$. Thus, it seems like

$$\text{Cond. Prob. of 2} + \text{Cond. Prob. of 3} + \text{Cond. Prob. of 6} = \frac{0.1}{0.7} + 0 + \frac{0.5}{0.7} = \frac{6}{7}$$

How did we intuitively compute that? For E to occur given that F already has, the outcome *must* be in EF (so you intuitively ignored the outcome “3” in E). Because we know that $F \subset S$ has occurred, we make F our new, or *reduced* sample space. Next, the probability of EF occurring given that F has should be the probability of EF relative to the probability of F : i.e. the probability of $\{2, 6\}$ relative to $\{2, 4, 6\}$ (consider Venn diagrams). That is,

DEFINITION: Let $E, F \subset S$ with $P(F) > 0$, then

$$P(E|F) = \frac{P(EF)}{P(F)}$$

In the first example above, $EF = \{2, 6\}$ and so

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{P(\{2, 6\})}{P(\{2, 4, 6\})} = \frac{2/6}{3/6} = \frac{2}{3}$$

The above is taken to be the definition of conditional probability in all cases.

EXAMPLES:

1. The probability that a person lives to 80 is 0.64. The probability that a person lives until 90 is 0.51. What is the probability that a person will live to 90, given that that person is 80?

1. The probability that a person lives to 80 is 0.64. The probability that a person lives until 90 is 0.51. What is the probability that a person will live to 90, given that that person is 80?

Solution: Let A be the event of living until 90 and B be the event of living until 80. Note that $A \subset AB$. Then,

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{0.51}{0.64} \approx 0.797$$

2. A coin is flipped twice. Assuming all points of $S = \{(H, H), (T, H), (H, T), (T, T)\}$ are equally likely, what is the conditional probability that both flips are heads, given that

(a) the first flip is a head?

(b) at least one flip lands on heads?

Solution: Let E be the event that both flips are heads. Let A be the event that the first flip is a head. Then

$$P(E|A) = \frac{P(EA)}{P(A)} = \frac{P(\{(H, H)\})}{P(\{(H, T), (H, H)\})} = \frac{1/4}{2/4} = \frac{1}{2}$$

Next, let B be the event that at least one flip is a head. Then

$$P(E|B) = \frac{P(EB)}{P(B)} = \frac{P(\{(H, H)\})}{P(\{(H, T), (H, H), (T, H)\})} = \frac{1/4}{3/4} = \frac{1}{3}$$

This seems counter-intuitive as “the other could be heads or tails.” So it seems that answer should be $1/2$. However, the information “at least one coin landed on heads” only tells us that (T, T) didn’t happen. The remaining three options are equally likely.

3. In the Senate, there are 53 liberals and 47 conservatives. A committee is composed randomly of 9 senators. We know that the committee does not contain more than 3 conservative senators. What is the probability that there are exactly 7 liberal senators?

3. In the Senate, there are 53 liberals and 47 conservatives. A committee is composed randomly of 9 senators. We know that the committee does not contain more than 3 conservative senators. What is the probability that there are exactly 7 liberal senators?

Solution: Let A be the event of exactly 7 liberal senators. Let B be the event of 0, 1, 2, or 3 conservative senators. We want $P(A|B) = P(A \cap B)/P(B)$. Firstly,

$$P(B) = \sum_{i=0}^3 \frac{\binom{47}{i} \binom{53}{9-i}}{\binom{100}{9}} = 0.307519$$

Second, $A \cap B = A$ (7 liberal and 2 conservative senators). Thus,

$$P(A \cap B) = P(A) = \frac{\binom{53}{7} \binom{47}{2}}{\binom{100}{9}} = 0.087596$$

Therefore, the probability of interest is

$$\frac{0.087596}{0.307519} = 0.28485$$

Multiplication Rule

We have

$$P(E|F) = \frac{P(EF)}{P(F)}$$

Therefore,

$$\boxed{P(EF) = P(F)P(E|F)}$$

Thus, the probability of E and F occurring is equal to the probability of F occurring times the probability that E occurs given that F has.

EXAMPLE: Suppose an urn contains 8 red balls and 4 white balls. We draw two balls without replacement. If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

EXAMPLE: Suppose an urn contains 8 red balls and 4 white balls. We draw two balls without replacement. If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

Solution: Let R_1 and R_2 be the events that the first and second ball are red, respectively. Then we want $P(R_1R_2)$. Note that we easily have that $P(R_1) = 8/12$. Also, $P(R_2|R_1) = 7/11$ because the *reduced* sample space R_1 consists of 7 red balls and 4 white balls, and we are asking for the probability of choosing a red ball. Therefore,

$$P(R_1R_2) = P(R_1)P(R_2|R_1) = \frac{8}{12} \cdot \frac{7}{11} = \frac{56}{132} = \frac{14}{33}$$

Note, could also just do straightaway as $\binom{8}{2} / \binom{12}{2}$.

The *multiplication rule* can be generalized.

THEOREM: If $P(E_1E_2 \dots E_n) > 0$, then

$$\boxed{P(E_1E_2 \dots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2) \dots P(E_n|E_1E_2 \dots E_{n-1})}$$

Proof:

$$\begin{aligned} & P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3) \dots P(E_n|E_1E_2 \dots E_{n-1}) \\ &= P(E_1E_2)P(E_3|E_1E_2)P(E_4|E_1E_2E_3) \dots P(E_n|E_1E_2 \dots E_{n-1}) \\ &= P(E_1E_2E_3)P(E_4|E_1E_2E_3) \dots P(E_n|E_1E_2 \dots E_{n-1}) \\ &\vdots \\ &= P(E_1E_2 \dots E_n) \blacksquare \end{aligned}$$

EXAMPLE: An urn contains 8 red balls and 4 white balls. Suppose that we draw 4 balls without replacement. What is the probability that the sequence draws was $RRWW$?

EXAMPLE: An urn contains 8 red balls and 4 white balls. Suppose that we draw 4 balls without replacement. What is the probability that the sequence draws was $RRWW$?

Solution: Let E_1 be the event that the first ball was red, E_2 be the event that the second ball was red, E_3 be the event that the third ball was white, E_4 be the event that the fourth ball was white. We want $P(E_1E_2E_3E_4)$. We have

$$P(E_1E_2E_3E_4) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3) = \frac{8}{12} \cdot \frac{7}{11} \cdot \frac{4}{10} \cdot \frac{3}{9} \approx 0.057$$

Note that we could also have solved using generalized counting principle.

Section 3.3 Bayes's Formula

THEOREM (Law of total probability): Let E and F be events, then

$$\boxed{P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)}$$

Proof: We have

$$P(E|F)P(F) + P(E|F^c)P(F^c) = P(EF) + P(EF^c) = P(EF \cup EF^c) = P(E)$$

where the second to last equality followed from mutual exclusiveness. ■

EXAMPLES:

1. In a hospital, 35% of those with high blood pressure have had strokes and 20% of those without high blood pressure have had strokes. If 40% of the patients have high blood pressure, what percent of the patients have had strokes?

Solution: Let A be the event of a random patient having had a stroke. Let B be the event that a person had high blood pressure. We want $P(A)$.

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) = 0.35 \cdot 0.4 + 0.2 \cdot 0.6 = 0.14 + 0.12 = 0.26$$

2. In answering a question on a multiple-choice test, a student either knows the answer or guesses. Suppose the student knows the answer with probability $2/3$, and guesses with probability $1/3$. Assume that a student who guesses at the answer will be correct with probability $1/5$ (i.e. five questions). What is the conditional probability that a student knows the answer to a question, given that he or she answered correctly?

2. In answering a question on a multiple-choice test, a student either knows the answer or guesses. Suppose the student knows the answer with probability $2/3$, and guesses with probability $1/3$. Assume that a student who guesses at the answer will be correct with probability $1/5$ (i.e. five questions). What is the conditional probability that a student knows the answer to a question, given that he or she answered correctly?

Solution: Let C and K be the events that the student answers correctly, and actually knew the answer, respectively. We have

$$P(K) = 2/3$$

$$P(K^c) = 1/3$$

$$P(C|K) = 1$$

$$P(C|K^c) = 1/5$$

Then

$$\begin{aligned} P(K|C) &= \frac{P(KC)}{P(C)} = \frac{P(C|K)P(K)}{P(C)} = \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K^c)P(K^c)} \\ &= \frac{1 \cdot 2/3}{1 \cdot 2/3 + (1/5) \cdot (1/3)} = 10/11 \end{aligned}$$

3. A blood test is 95% effective in detecting a disease in someone that has it. It also gives a “false positive” 1 percent of the time. If 0.5% of the population has the disease, what is the probability that a person has the disease given that the test result is positive?

Solution: Let D be the event that a person has the disease. Let E be event that the test is positive. We are interested in $P(D|E)$. We know that

$$P(D) = 0.005$$

$$P(D^c) = 0.995$$

$$P(E|D) = 0.95$$

$$P(E|D^c) = 0.01$$

Therefore,

$$P(D|E) = \frac{P(DE)}{P(E)} = \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)} = \frac{0.95 \cdot 0.005}{0.95 \cdot 0.005 + 0.01 \cdot 0.995} \approx 0.323$$

Why? Think of a population of 200 people, where one person has disease. That person will be found positive with probability 0.95. Of other 199 people, approximately $199 \cdot 0.01 = 1.99$ will be found to have the disease. So, on some average, $0.95 + 1.99 \approx 3$ people would be found to have disease. So more false positives than actual positives.

We can generalize the above ideas.

DEFINITION: Let $\{F_1, \dots, F_n\}$ be a set of nonempty subsets of the sample space S . If the sets F_i are mutually exclusive and $\cup F_i = S$ then the set $\{F_1, \dots, F_n\}$ is called a partition of S .

THEOREM (Law of total probability): Let $\{F_1, \dots, F_n\}$ be a partition of S with $P(F_i) > 0$. Then for any E we have

$$P(E) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + \dots + P(E|F_n)P(F_n) = \sum_{i=1}^n P(E|F_i)P(F_i)$$

Proof:

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

THEOREM (Bayes's formula): For a partition $\{F_1, \dots, F_n\}$ and a set E we have

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

Terminology: F_i 's are the *hypotheses*, $P(F_i)$ is the *prior* probability of F_i , and $P(F_i|E)$ is called the *posterior* probability of F_i given E . That is, how your probability should be changed with new evidence.

EXAMPLES:

1. An army has 4 sharpshooters. The probabilities of each hitting a target at a given distance are 0.2, 0.3, 0.15, and 0.7, respectively. The first three shooters are each chosen with probability 0.2, while the fourth is chosen with probability 0.4. Suppose that the target is hit. What is the probability the shooter was shooter number 4.

Solution: Let E be the event that the target was hit. Let F_i be the event that the i th shooter was chosen. Note that our *prior* is $P(F_4) = 0.4$. Then, with our new information

$$P(F_4|E) = \frac{P(E|F_4)P(F_4)}{\sum P(E|F_i)P(F_i)} = \frac{0.7 \cdot 0.4}{0.2 \cdot 0.2 + 0.3 \cdot 0.2 + 0.15 \cdot 0.2 + 0.7 \cdot 0.4} \approx 0.683$$

2. Suppose we have 10 coins such that if the i th coin is flipped, heads will appear with probability $i/10$, $i = 1, \dots, 10$. When one of the coins is randomly selected and flipped, it shows heads. What is the conditional probability that it was the fifth coin?

2. Suppose we have 10 coins such that if the i th coin is flipped, heads will appear with probability $i/10$, $i = 1, \dots, 10$. When one of the coins is randomly selected and flipped, it shows heads. What is the conditional probability that it was the fifth coin?

Solution: Let B_i be the event that the i th coin was chosen. Let A be the event that the flip showed heads. We are interested in $P(B_5|A)$. I will solve for $P(B_i|A)$. We know that $P(B_i) = 1/10$ and $P(A|B_i) = i/10$. We want $P(B_i|A)$. We have

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{k=1}^{10} P(A|B_k)P(B_k)} = \frac{(i/10)(1/10)}{\sum_{k=1}^{10} (k/10)(1/10)} = \frac{i}{\sum_{k=1}^{10} k} = \frac{i}{55}$$

Thus, $P(B_5|A) = 5/55 = 1/11$.

Section 3.4 Independent Events

We think of A being independent from B if $P(A|B) = P(A)$. But using that

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

So, if A is independent of B , then B is independent of A . Symmetric.

DEFINITION: Two events, A and B , are called independent if

$$\boxed{P(AB) = P(A)P(B)}$$

Two sets are dependent if they are not independent.

EXAMPLES:

1. A card is drawn at random from a deck of 52. Let S be the event that it is a spade, and A be the event of an ace. Are these events independent?

Solution: There are 52 possible outcomes. We have

$$P(S) = 13/52, \quad P(A) = 4/52, \quad P(A \cap S) = 1/52$$

Finally, $P(S)P(A) = 13 \cdot 4/52^2 = 1/52$. So yes.

2. Consider choosing randomly from $\{1, 2, 3\}$. We perform the choice twice. The sample space for this experiment is $S = \{(i, j)|i, j \in \{1, 2, 3\}\}$. Let E be the event of a 1 in the first choice. Thus, $E = \{(1, 1), (1, 2), (1, 3)\}$. Let F be the event that the sum total is 3 and let G be the event that the sum total is 4. Then $F = \{(1, 2), (2, 1)\}$ and $G = \{(1, 3), (2, 2), (3, 1)\}$. Assuming all are equiprobable, we have

$$P(E) = 1/3, \quad P(F) = 2/9, \quad P(G) = 1/3, \quad P(FE) = 1/9, \quad P(GE) = 1/9$$

So, E and G are independent whereas E and F are not. Why? Intuitively, when considering whether or not you will sum to 4, the first roll tells you nothing as no matter what you get, the probability will be $1/3$ that the second roll will be such that the sum will be 4. Not so with a sum of 3. Getting a 1 on first role makes a sum of 3 possible whereas getting a 3 (for example) makes it already impossible. Thus information is there.

THEOREM: If A and B are independent, then A and B^c are independent as well.

Proof: We have $P(A) = P(AB) + P(AB^c)$, therefore

$$P(AB^c) = P(A) - P(AB) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c) \blacksquare$$

COROLLARY: If A and B are independent, then A^c and B^c are independent.

REMARK: If A and B are mutually exclusive and $P(A) > 0$ and $P(B) > 0$, then they are dependent (counter to intuition?), because $P(A \cap B) = 0$.

QUESTION: Suppose that E is independent from F and E is independent from G . Is E necessarily independent from FG ?

EXAMPLE: Two fair dice are thrown. Let E be the event the sum is 7. Let F denote the event that the first die is a 3, and let G denote the event that the second die is 3. We have that

$$P(E) = 1/6, \quad P(F) = 1/6, \quad P(EF) = P\{(3, 4)\} = \frac{1}{36}, \quad P(EG) = P\{(4, 3)\} = \frac{1}{36}$$

Thus, E and F are independent. So are E and G . However,

$$P(E|FG) = 0 \neq P(E) = \frac{1}{6}$$

DEFINITION: Three events E , F and G are independent if

$$P(EFG) = P(E)P(F)P(G)$$

and

$$P(EF) = P(E)P(F), \quad P(EG) = P(E)P(G), \quad P(FG) = P(F)P(G)$$

Note that if E , F and G are independent, then E is independent from any set formed from F and G . For example,

$$\begin{aligned} P(E(F \cup G)) &= P(EF \cup EG) = P(EF) + P(EG) - P(EFG) \\ &= P(E)P(F) + P(E)P(G) - P(E)P(F)P(G) \\ &= P(E)[P(F) + P(G) - P(FG)] = P(E)P(F \cup G) \end{aligned}$$

DEFINITION: Multiple events, E_1, \dots, E_n are independent if all subsets of those sets satisfy

$$\boxed{P(E_{i_1}E_{i_2} \dots E_{i_k}) = P(E_{i_1}) \dots P(E_{i_k})}$$

DEFINITION: Multiple events, E_1, \dots, E_n are called pairwise-independent if

$$\boxed{P(E_iE_j) = P(E_i)P(E_j)}$$

for any two.

TERMINOLOGY: When an experiment consists of repeated subexperiments (like tossing a coin), we call each subexperiment a *trial*. If we assume that the outcome of any subset of the trials has no effect on the outcomes of the other trials, we say the trials are independent.

EXAMPLE: An infinite sequence of independent trials are to be performed. Each trial results in a success with probability p and a failure with probability $1 - p$. What is the probability that

1. at least 1 success occurs in the first n trials?
2. exactly k successes occur in the first n trials?

EXAMPLE: An infinite sequence of independent trials are to be performed. Each trial results in a success with probability p and a failure with probability $1 - p$. What is the probability that

1. at least 1 success occurs in the first n trials?
2. exactly k successes occur in the first n trials?

Solution: Let E be the event of at least one success. Then E^c is the probability of no successes in the first n trials. This is easier to compute. Let E_i be the event of failure on the i th event. Then

$$P(E^c) = P(E_1 E_2 \dots E_n) = P(E_1)P(E_2) \dots P(E_n) = (1 - p)^n$$

Therefore, $P(E) = 1 - (1 - p)^n$.

Now consider the second question. The probability for any particular sequence of k successes and $n - k$ failures happening is $p^k(1 - p)^{n - k}$. There are $\binom{n}{k}$ ways to have k successes, therefore the probability of exactly k successes is

$$P\{\text{exactly } k \text{ successes}\} = \binom{n}{k} p^k (1 - p)^{n - k}$$

Section 3.5 $P(\cdot|F)$ is a Probability

For any $F \subset S$ with $P(F) > 0$, $P(\cdot|F)$ satisfies the axioms of being a probability measure on S .

THEOREM: Suppose that $P(F) > 0$. Then

1. For any $E \subset S$, we have $0 \leq P(E|F) \leq 1$.
2. $P(S|F) = 1$.
3. If E_i , $i = 1, 2, \dots$, are mutually exclusive events, then

$$P\left(\bigcup_{i=1}^{\infty} E_i|F\right) = \sum_{i=1}^{\infty} P(E_i|F)$$

Therefore, if we let $Q(E) = P(E|F)$, $Q(E)$ is a probability function on the events of S and we have all results we used to. For example,

$$Q(E_1 \cup E_2) = Q(E_1) + Q(E_2) - Q(E_1 E_2)$$

is equivalent to

$$P(E_1 \cup E_2|F) = P(E_1|F) + P(E_2|F) - P(E_1 E_2|F)$$

which is certainly less obvious. Also, for E_1, E_2 we have

$$Q(E_1) = Q(E_1|E_2)Q(E_2) + Q(E_1|E_2^c)Q(E_2^c)$$

But,

$$Q(E_1|E_2) = \frac{Q(E_1 E_2)}{Q(E_2)} = \frac{P(E_1 E_2|F)}{P(E_2|F)} = \frac{P(E_1 E_2 F)}{P(F)} / \frac{P(E_2 F)}{P(F)} = P(E_1|E_2 F)$$

Thus, above is equivalent to

$$P(E_1|F) = P(E_1|E_2 F)P(E_2|F) + P(E_1|E_2^c F)P(E_2^c|F)$$