

4.6: Conditional Probability

The probability of event B happening, given that event A already happened, is called the **conditional probability**. The conditional probability of B, given A is written as $P(B | A)$, and is read as “the probability of B given A happened first.” We can use the General Multiplication Rule when two events are dependent.

Definition: General Multiplication Rule

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B|A)P(A \cap B) \\ &= P(A) \cdot P(B|A) \end{aligned}$$

A bag contains 10 colored marbles: 7 red and 3 blue. A random experiment consists of drawing a marble from the bag, then drawing another marble without replacement (without putting the first marble back in the bag). Find the probability of drawing a red marble on the first draw (event R_1), and drawing another red marble on the second draw (event R_2).

Solution

Drawing a red marble on the first draw and drawing a red marble on the second draw are dependent events because we do not place the marble back in the bag. The probability of drawing a red marble on the first draw is $P(R_1) = \frac{7}{10}$, but on the second draw, the probability of drawing a red marble given that a red marble was drawn on the first draw is $P(R_2|R_1) = \frac{6}{9}$.

Thus, by the general multiplication rule, $P(R_1 \text{ and } R_2) = P(R_1) \cdot P(R_2|R_1) = (\frac{7}{10}) (\frac{6}{9}) = 0.4667$.

A bag contains 10 colored marbles: 7 red and 3 blue. A random experiment consists of drawing a marble from the bag, then drawing another marble without replacement. Create the tree diagram for this experiment and compute the probabilities of each outcome.

Solution

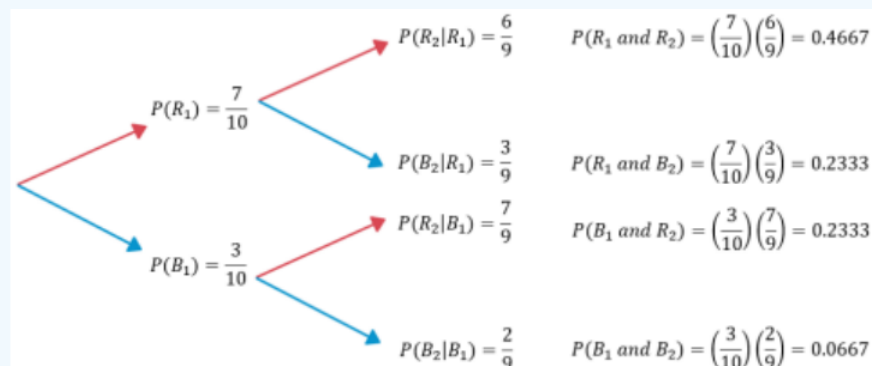


Figure 4-15

If we were to multiply the probabilities as we move from left to right up each set of tree branches as shown in Figure 4-15, we get the intersections. For example, by the general multiplication rule, $P(R_1 \text{ and } R_2) = P(R_1) \cdot P(R_2|R_1) = (\frac{7}{10}) (\frac{6}{9}) = 0.4667$.

Put the four intersection values into a contingency table and total the rows and columns. The table will help solve probability questions of other events.

	R_2	B_2	Total
R_1	0.4667	0.2333	0.7
B_1	0.2333	0.0667	0.3
Total	0.7	0.3	1

The grand total should add up to 1 since we have 100% of the sample space.

$$\text{Conditional Probability Rule: } P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ or } P(B|A) = \frac{P(A \cap B)}{P(A)}$$

The following table shows the utility contract granted for a specific year. One contractor is randomly chosen.

Corporation	Government	Individual	Total	United States	0.45	0.007	0.08	0.537	Foreign	0.41	0.003	0.05	0.463	Total	0.86	0.01
0.13 1																

- Compute the probability the contractor is from the United States and is a corporation.
- Compute the probability the contractor is from the United States given that they are a corporation.
- If the contractor is from a foreign country, what is the probability that it is from a government?
- Are the events a “contractor is an individual” independent of a “contractor from the United States?”

Solution

a) For the intersection in the contingency tables use where the row and column meet. $P(\text{U.S.} \cap \text{Corp}) = 0.45$.

$$b) P(\text{U.S.}|\text{Corp}) = \frac{P(\text{U.S.} \cap \text{Corp})}{P(\text{Corp})} = \frac{0.45}{0.86} = 0.5233.$$

$$c) P(\text{Gov}|\text{Foreign}) = \frac{P(\text{Gov} \cap \text{Foreign})}{P(\text{Foreign})} = \frac{0.003}{0.463} = 0.0065.$$

d) Do not assume independence between two variables in a contingency table since the data may show relationships that you didn't know were there.

Use the definition of independent events. If the two events are independent then we would have $P(\text{Individual} \cap \text{U.S.}) = P(\text{Individual}) \cdot P(\text{U.S.})$. First find the intersection using where the row and column meet to get $P(\text{Individual} \cap \text{U.S.}) = 0.08$. Then use the row and column totals to find $P(\text{Individual}) \cdot P(\text{U.S.}) = 0.13 \cdot 0.537 = 0.0698$. Since $P(\text{Individual} \cap \text{U.S.}) \neq P(\text{Individual}) \cdot P(\text{U.S.})$ these two events are dependent.

A random sample of 500 people was taken from the 2010 United States Census. Their marital status and race were recorded in the following contingency table. A person is randomly chosen, find the following.

Race	Marital Status	American	Indian	Black	Asian	White	Two Major Races	Total	Divorced	0	6	1	30	1	38	Married	1	25	23	156
4 209 Single 2 33 21 155 11 222 Widowed 0 7 2 22 0 31 Total 3 71 47 363 16 500																				

- $P(\text{Single and Asian})$
- $P(\text{Single} | \text{Asian})$
- Given that a person is single what is the probability their race is Asian?

Solution

a) The intersection for a contingency table is found by simply finding where the row or intersection meets. There are 21 single Asians, therefore the $P(\text{Single} \cap \text{Asian}) = P(\text{Single and Asian}) = 21/500 = 0.042$. Do not multiply the row total times the column total since there is no indication that these are independent events.

b) In words we are trying to find the probability that the person is single given that we already know that their race is Asian. Using the conditional probability formula, we get $P(\text{Single} | \text{Asian}) = \frac{P(\text{Single} \cap \text{Asian})}{P(\text{Asian})} = \frac{21}{47} = 0.4468$.

c) This seems similar to the last question, however the part we know is that the person is single, but we do not know their race. In symbols we want to find the $P(\text{Asian} | \text{Single}) = \frac{P(\text{Asian} \cap \text{Single})}{P(\text{Single})} = \frac{21}{222} = 0.0946$.

Keep in mind that $P(A | B) \neq P(B | A)$ since we would divide by a different total in the equation.

A blood test correctly detects a certain disease 95% of the time (positive result), and correctly detects no disease present 90% of the time (negative result). It is estimated that 25% of the population have the disease. A person takes the blood test and they get a positive result. What is the probability that they have the disease?

Solution

Let D = Having the Disease, D^C = Not having the disease, $+$ is a positive result, and $-$ is a negative result. We are given in the problem the following: $P(+ | D) = 0.95$, $P(- | D^C) = 0.90$, $P(D) = 0.25$. We want to find $P(D|+) = \frac{P(D \cap +)}{P(+)}$.

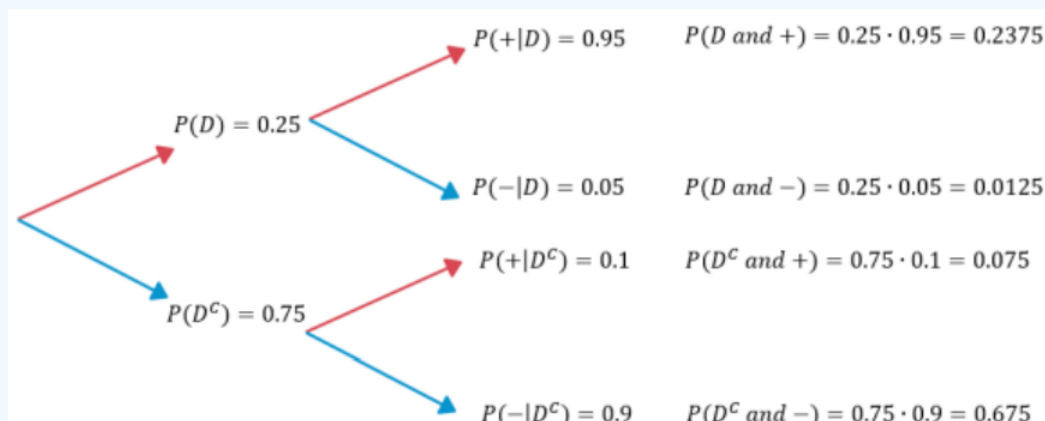


Figure 4-16

When you multiply up each pair of tree branches from left to right as shown in Figure 4-16, you are finding the intersection of the events. Place the multiplied values into a table. Note that the 0.2375 is not our answer. This is the people who have the disease and tested positive, but does not take into consideration the false positives. Since we know that the result was positive, we only divide by the proportion of positive results.

	D^C	D	Total
+	0.075	0.2375	0.3125
-	0.675	0.0125	0.6875
Total	0.75	0.25	1

$$P(D|+) = \frac{P(D \cap +)}{P(+)} = \frac{0.2375}{0.3125} = 0.76 \quad (4.6.1)$$

There is a 76% chance that they have the disease given that they tested positive. Many of the more difficult probability problems can be set up in a table, which makes the probabilities easier to find.