

Using contingency tables in Conditional Probability

Problem: According to a survey of 50 new-car dealers, there were 20 that had been in business 10 years or longer, and 30 had been in business for less than 10 years. Among the dealers that had been in business for 10 years or longer, 16 of them gave good service under warranty; and among the dealers who had been in business less than 10 years, 10 of them gave good service under warranty.

- What is the probability that a dealer who has been in business for 10 years or longer will give good service under warranty?
- What is the probability that a dealer who has been in business for less than 10 years will give good service under warranty?
- If we select a new-car dealer at random from these 50 dealers, what is the probability that they will give good service under warranty?
- If we select a new-car dealer at random from these 50 dealers and find that they give good service under warranty. what is the probability that they have been in business for 10 years or longer?

Here I use the notation:

G = the dealer gives good service under warranty

T = the dealer has been in business 10 years or more

The data from the 50 new-car dealers surveyed can be summarized in a contingency table as follows:

	G	G'	totals
T	16 $ T \cap G $	4 $ T \cap G' $	20 $ T $
T'	10 $ T' \cap G $	20 $ T' \cap G' $	30 $ T' $
totals	26 $ G $	24 $ G' $	50 $ S $

The main body of the table (the part inside the double lines, excluding the “totals” rows and columns) can be viewed as a Venn diagram, with each rectangle representing an intersection as noted above. The numbers then represent the number of outcomes (dealerships) in each of the intersections.

The event G is represented by the entire first column; G' is represented by the entire second column; T is represented by the entire first row; and T' is represented by the entire second row. The whole main body of the table represents the sample space S .

Notation: to indicate the number of elements in a set A , you may use any one of the following notations: $\#A$ or $|A|$ or $n(A)$

Problem: A company rents cars from three rental companies, Company 1, Company 2, and Company 3. 60% of the cars it rents come from Company 1, 30% come from Company 2, and 10% come from Company 3. 9% of the cars from Company 1 need oil changes, 20% of the cars from Company 2 need oil changes, and 6% of the cars from Company 3 need oil changes.

If we pick one of the rented cars at random, what is the probability that it will need an oil change?

If we pick a car at random and it needs an oil change, what is the probability that it came from Company 2?

Here I use the notation:

A = “The car needs an oil change”

B_1 = “The car came from Company 1”

B_2 = “The car came from Company 2”

B_3 = “The car came from Company 3”

Using the probabilities given in the problem:

	B_1	B_2	B_3	totals
A	$(0.09)(0.6) = 0.054$ $P(A \cap B_1)$	$(0.20)(0.3) = 0.06$ $P(A \cap B_2)$	$(0.06)(0.10) = 0.006$ $P(A \cap B_3)$	0.12 $P(A)$
A'	$(0.91)(0.6) = 0.546$ $P(A' \cap B_1)$	$(0.30)(0.8) = 0.24$ $P(A' \cap B_2)$	$(0.94)(0.10) = 0.094$ $P(A' \cap B_3)$	0.88 $P(A')$
totals	0.6 $P(B_1)$	0.3 $P(B_2)$	0.1 $P(B_3)$	1 $P(S)$

It can sometimes be more helpful to use a numerical table, similar to that we would have if we were given data as in the problem on page 1. To make up such a table, assume you have chosen a random sample from the population, large enough that all the numbers we compute below come out to be whole numbers. Since none of the probabilities in the table above have more than three decimal places, we will choose the sample size as 1000, so assume 1000 cars randomly selected from all the cars rented by this firm. Then we expect that in this sample the number of cars in each category to be as follows: (or nearly so!)

	B_1	B_2	B_3	totals
A	54 $ A \cap B_1 $	60 $ A \cap B_2 $	6 $ A \cap B_3 $	120 $ A $
A'	546 $ A' \cap B_1 $	240 $ A' \cap B_2 $	94 $ A' \cap B_3 $	880 $ A' $
totals	600 $ B_1 $	300 $ B_2 $	100 $ B_3 $	1000 $ S $

The entries in this table are found by multiplying the probabilities in the previous table by the sample size, here 1000.