Conditional Probability and Contingency Tables

by Sophia

WHAT'S COVERED

This tutorial will cover conditional probability in the context of two-way tables. Our discussion breaks down as follows:

1. Conditional Probability

You can use two-way tables to find conditional probabilities. Recall that conditional probability is the probability that some event (B) occurs given that some other event (A) has already occurred. The probability of B given A is written this way:

$$P(B, \text{given } A) = P(B \mid A)$$

EXAMPLE Suppose that 338 middle school students were asked which was their dominant hand. Here the results are shown in the two-way table:

		Dominant Hand			
		Right	Left	Ambidextrous	
Grade	6th	99	9	2	110
	7th	90	31	0	121
	8th	93	11	3	107
		282	51	5	338

If a student is a sixth grader, what is the probability that he or she is left-handed? To find the answer, isolate

the sixth grade row. This is a question of conditional probability: the probability of a student being left-handed given that the student is a sixth grader, P(L | 6). The formula looks like this:

$$P(L \mid 6) = \frac{P(L \text{ and } 6)}{P(6)} = \frac{\frac{9}{338}}{\frac{110}{338}} = \frac{9\ (Left\ handed\ sixth-graders)}{110\ (All\ sixth-graders)}$$

This formula shows the probability of both left-handed and sixth grade divided by the probability of sixth grade. The probability that a student is left-handed *and* a sixth grader is 9 out of the 338 middle schoolers. The probability that a student is in sixth grade is 110 out of the 338. This reduces to 9 out of 110.

HINT

Notice that you can use the probabilities 9/338 and 110/338, which were both divided by 338, the grand total. Or, you can just use the frequency from the cell for both left-handed and sixth, 9, and from the marginal distribution in the row totals for sixth grade, 110.

TRY IT

Using the table above, find the following probabilities.

Questions	Written as Conditional Probability	Conditional Probability Formula
What is the probability that a seventh-grade student is ambidextrous?	What is the probability that a student is ambidextrous, given they are a seventh grader?	$P(A \mid 7) = \frac{P(A \text{ and } 7)}{P(7)} = \frac{0}{121} = 0$
What is the probability that a student who is right-handed is in eighth grade?	What is the probability that a student is in eight grade, given that he or she is right-handed?	$P(8 \mid R) = \frac{P(8 \text{ and } R)}{P(R)} = \frac{93}{282}$

2. Tables

You can use a two-way table that actually has probabilities in it or relative frequencies.

EXAMPLE A class of 10th graders was asked if they prefer cheese, pepperoni, or sausage pizza. The percentages are shown below: 41% of all of these kids are girls that enjoy cheese pizza, 12% of all of the kids are boys that enjoy pepperoni, etc.

	Cheese	Pepperoni	Sausage	
Boy	0.05	0.12	0.19	0.36
Girl	0.41	0.16	0.07	0.64

0.46	0.28	0.26	1

Given that the student is a boy, what is the probability that he likes cheese?

To find the probability of a student preferring cheese pizza given that he's a boy, you can use the probabilities from the marginal distributions in the column or row totals, instead of the frequencies.

$$P(C \mid B) = \frac{P(C \text{ and } B)}{P(B)} = \frac{0.05}{0.36} = 0.139$$

Therefore, the probability that a student enjoys cheese and is a boy is the 0.05 value from the table. The probability of the student being a boy, in this particular sample, is 36%. This reveals that there's about 14% probability that if you are a boy, you'll prefer cheese pizza.

SUMMARY

Conditional probability is the probability that some event (Event B) follows some other event which has already occurred (Event A). It's calculated by dividing the joint probability-- the probability of A and B--by the probability of the event which has already occurred. This formula works for all events, not just for independent events or mutually exclusive events, and the data for these formulas can be found in two-way tables.

Good luck!

Source: THIS TUTORIAL WAS AUTHORED BY JONATHAN OSTERS FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.