

5.4: Two-Way Tables (2 of 5)

Learning Objectives

- Calculate marginal, joint, and conditional percentages and interpret them as probability estimates.

In the previous section, we used the information in a two-table to examine the relationship between two categorical variables. Our goal was to answer the big question: *Are the variables related?*

In this section, we continue to work with two-way tables, but we ask a different set of questions.

Example

Community College Enrollment

The following table summarizes the full-time enrollment at a community college located in a West Coast city. There are a total of 12,000 full-time students enrolled at the college. The two categorical variables here are *gender* and *program*. The programs include academic and vocational programs at the college. Assume that a student can enroll in only one program.

	Arts-Sci	Bus-Econ	Info Tech	Health Science	Graphics Design	Culinary Arts	Row Totals
Female	4,660	435	494	421	105	83	6,198
Male	4,334	490	564	223	97	94	5,802
Column Totals	8,994	925	1,058	644	202	177	12,000

Let's consider a few preliminary questions to get familiar with this new data set.

1. What proportion of the total number of students are male students?

Answer

$$\frac{\text{number of male students}}{\text{total number of students}} = \frac{5,802}{12,000} = \mathbf{0.4835} \text{ (or } 48.35\%)$$

2. What proportion of the total number of students are Bus-Econ students?

Answer

$$\frac{\text{number of Bus-Econ students}}{\text{total number of students}} = \frac{925}{12,000} = \mathbf{0.077} \text{ (or } 7.7\%)$$

Note that to calculate this proportion, we used two numbers in the margin that relate to just one of the categorical variables (program). This calculation is therefore called a **marginal proportion**.

Note: This proportion does not help us determine if gender is related to program because it involves only one of the variables.

Now consider the following question:

If we choose one student at random from among all 12,000 students at the college, how likely is it that this student will be in the Bus-Econ program?

From our previous calculation, we know that only about 8% (7.7%) of the students at the college are in the Bus-Econ program. That's a fairly low number, so it is not very likely that our random student will be a Bus-Econ student.

One way to state our conclusion is to say:

There is about an **8% chance** of picking a Bus-Econ major.

This means that if we selected 100 students at random, we would expect on average that 8 of them would be in the Bus-Econ program.

Here is another way to state this conclusion:

There is about an **0.08 probability** of picking a Bus-Econ major.

Because this probability is exactly the same as the marginal proportion we calculated earlier, we call it a **marginal probability**.

Note: P for Probability

It is customary to use the capital letter P to stand for probability. So instead of writing “The probability that a student is in Bus-Econ program equals 0.08,” we can write $P(\text{student is in Bus-Econ}) = 0.08$.

The following table is used for the next Learn By Doing and Did I Get This? activities.

	Arts-Sci	Bus-Econ	Info Tech	Health Science	Graphics Design	Culinary Arts	Row Totals
Female	4,660	435	494	421	105	83	6,198
Male	4,334	490	564	223	97	94	5,802
Column Totals	8,994	925	1,058	644	202	177	12,000

Learn By Doing

<https://assessments.lumenlearning.co...sessments/3536>

<https://assessments.lumenlearning.co...sessments/3537>

Example

Conditional Probability

Here is the same community college enrollment data.

	Arts-Sci	Bus-Econ	Info Tech	Health Science	Graphics Design	Culinary Arts	Row Totals
Female	4,660	435	494	421	105	83	6,198
Male	4,334	490	564	223	97	94	5,802
Column Totals	8,994	925	1,058	644	202	177	12,000

Here is our first question:

If we select a female student at random, what is the probability that she is in the Health Sciences program?

Answer Of the 6,198 female students at the college, **421** are enrolled in Health Sciences. (Find these numbers in the table.) The probability we are looking for is:

$$\frac{\text{421}}{\text{6,198}} \approx \text{0.07}$$

Therefore, the probability that a female student is in the Health Sciences program is approximately 0.07.

Focus on Language

We need to pause here and be very careful about the language we use in describing this situation.

Note that we *start* with a female student and *then* ask what is the probability that this female student is in the Health Sciences department.

In this case, our *starting point is that the student is a female*. This information sets the conditions for calculating the probability. Once the condition (*student is female*) is set, we focus on the female student population. In terms of the two-way table, it means that the only numbers we will be using are in the Female row: 421 and 6,198.

What Is a Conditional Probability?

The probability we calculated earlier is an example of a **conditional probability**. In general, a conditional probability is one that is based on a given condition. Here the *given condition* is that the student is female.

Here is the notation we use for a conditional probability:

- Original question: *If we select a female student at random, what is the probability that she is in the Health Sciences program?*
- Notation: $P(\text{student is in Health Sciences} \textbf{ given that student is female})$.
- We also write this as $P(\text{Health Sciences} \textbf{ given female})$.

An even shorter way of writing this is to use a vertical bar | in place of *given*: $P(\text{Health Sciences} \mid \text{female})$.

The following table is used for the next Learn By Doing and Did I Get This? activities.

	Arts-Sci	Bus-Econ	Info Tech	Health Science	Graphics Design	Culinary Arts	Row Totals
Female	4,660	435	494	421	105	83	6,198
Male	4,334	490	564	223	97	94	5,802
Column Totals	8,994	925	1,058	644	202	177	12,000

Learn By Doing

<https://assessments.lumenlearning.co...sessments/3538>

<https://assessments.lumenlearning.co...sessments/3539>

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** *CC BY: Attribution*

This page titled [5.4: Two-Way Tables \(2 of 5\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Bill Pelz](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.