Learning Goal: Use a two-way table to estimate probability and risk.

Introduction to probability:

Previously, we used a two-way table to examine the association between two categorical variables. In this activity we will do the same calculations through a new lens; the lens is probability.

Later in the course we will study probability in more detail because it is the heart of statistical inference, but this is a natural place to begin to use probability language. We use the word probability to mean "likelihood" or "chance." For our purposes, probability describes the chance of a specific outcome if we are randomly selecting from an individual from a group.

When we ask questions such as "who is more likely to use math tutoring services, STEM or non-STEM majors," we are using probability language. Similarly, the question, "what percentage of the STEM majors use math tutoring" is the same as asking, "what is the probability that a STEM student uses the college's math tutoring services?" We do the same calculations to answer each of these questions.

We use the word *risk* when the probability describes a negative outcome. For example, we talk about the probability of winning the lottery and the risk of being struck by lightning.

1) This data comes from a Massachusetts study.

	Smoker Yes	Smoker No	Total
Low Birth Weight Yes	30	29	59
Low Birth Weight No	44	86	130
Total	74	115	189

a) For a smoker, what is the risk of having a low weight baby?

b) For a non-smoker, what is the risk of having a low weight baby?

c) How much does smoking increase the risk of having a low weight baby? What is the percent increase?

$$percent\ increase\ when\ smoking = \frac{increase\ in\ risk}{risk\ when\ not\ smoking} =$$

Smoking increases the risk of having a low weight baby by _____%.

d) Other ways to compare risks:

- Calculate the percentage point difference by subtracting percentages:

  The risk associated with low birth weight is _____ percentage points higher for smokers.

- Calculate the relative risk by dividing the percentages:

  Smokers are _____ times more likely to have a low weight baby.

2) This table summarizes the results of a survey of firefighters in New York.

	No alcohol problems	Moderate to severe alcohol problems	Totals
Participated in 9/11 rescue	793	309	1,102
Did not participate in 9/11	441	110	551
Totals	1,234	419	1,653

a) Are the firefighters who participated in the 9/11 rescue at greater risk of alcohol problems? Support your answer.

b) How much does participating in the 9/11 rescue increase the risk of alcohol problems? What is the percent increase?

c) What is the relative risk of alcohol problems? Write a sentence to explain the relative risk.

3) StatCrunchU is a fictitious university of 46,000 students. In StatCrunch we can select a random sample of StatCrunchU students. Use this data to answer the questions for StatCrunchU students. Support your answers with appropriate calculations; include the ratios and explanations so that someone else can follow your reasoning.

	Student Loan no	Student Loan yes	Total
Job	17	65	82
No Job	77	41	118
Total	94	106	200

a) What is the probability that a StatCrunchU student has a student loan?

b) At StatCrunchU are students with jobs more likely to have student loans than students without jobs?

c) Are the following statements accurate based on this sample of StatCrunchU students? Support your answer.

Students with a job are more than twice as likely to have a student loan than students without a job.