

Comparing Proportions Using Relative Risk

We've previously learned that in a Chi-Squared Homogeneity Test the null hypothesis suggests the proportions between two (or more) groups are the same. Rather than comparing their difference, an alternative method for comparing proportions focuses on their **ratio**.

When an outcome / response variable is dichotomous (has two options), the ratio of the probabilities is called the **relative risk**.

Getting Warmed-up – Thinking About Percentages

1. Can a person's weekly wage increase by more than 100%? If so, what would this entail? Explain.
2. Can a person's weekly wage decrease by more than 100%? If so, what would this entail? Explain.

Example 1 – Malaria Vaccine

An article that appeared in the journal Lancet in May of 2021 (Dattoo, et al.) described a study that investigated a potential vaccine that might protect children against malaria. Researchers recruited children between the ages of 5 and 17 months in Burkina Faso, a country in western Africa, as participants. The children were randomly assigned into one of three groups: one group received a large dose of the potential vaccine, another group received a small dose, and a third group received a placebo. Researchers observed the children for the next 18 months, keeping track of whether or not the child developed malaria. Researchers hoped, of course, that children who received a vaccine would be less likely to develop malaria than children who received a placebo. Results are shown in the following table of counts:

	High Dose	Low Dose	Placebo	Total
Developed Malaria	38	43	105	186
Did Not Develop Malaria	108	103	42	253
Total	146	146	147	439

3. Identify the observational units, explanatory variable, and response variable. Classify each variable as categorical or numerical.

4. Is this an observational study or a randomized experiment? Explain how you can tell.
5. For each of the three groups, calculate the proportion of children who developed malaria. (Show how you calculate these, and report your answers with **three** decimal places of accuracy.)
6. Convert the proportions you calculated in #3 to percentages.
7. Use these percentages in a sentence to explain what they represent in the context of this study.

Key Question: Does this mean that children who received the placebo were $71.4\% - 26.0\% = 45.4\%$ more likely to develop malaria than children who received the high dose?

8. How should we interpret the difference of 45.4%?
9. How would we calculate the *percentage difference* in the proportion who developed malaria, comparing the placebo to the high dose group?

A Small Detour

Let's suppose (hypothetically) President Armstrong's IQ is 80 and Dr. Theobald's is 140.

10. By what percentage difference is Dr. Theobald's IQ score larger than President Armstrong's?
11. How would you write a general formula for the percentage difference?
12. Express the value you calculated in #10 in a sentence.

Back At It – Malaria Vaccine

13. Calculate the percentage difference in the proportion of children who developed malaria, comparing the low dose group to the high dose group.
14. Write a sentence to interpret this value in context.
15. Calculate the percentage difference in the proportion of children who developed malaria, comparing the placebo to the high dose group.
16. Write a sentence to interpret this value in context.

17. Calculate the *ratio* of the proportions who develop malaria between the low dose and high dose groups.

18. Write a sentence to interpret this value in context

Relative Risk

The number you calculated in 15 should look similar to what you saw in #11. You found that children who received the low dose of the malaria vaccine were **13.2%** more likely to develop malaria than those who received the high dose. Then, you found that children who received the low dose were **1.132** times more likely to develop malaria than those who didn't.

There is a formula for how these quantities are related:

$$\text{percentage difference} = (\text{ratio of proportions} - 1) \times 100\%$$

$$\text{percentage difference} = (\text{relative risk} - 1) \times 100\%$$

Suppose you know the risk of getting malaria for children who received the placebo is 2.744 times more likely for children who received the placebo compared to children who received the high dose.

19. Calculate the percentage difference in the likelihood of developing malaria between these two groups.

Another Detour

Recall, Dr. Theobold's IQ score is (hypothetically) 75% greater than President Armstrong's IQ score.

20. Does that mean President Armstrong's IQ is 75% lower than Dr. Theobold's? Why or why not?

21. Complete the sentence:

President Armstrong's IQ is _____ % lower than Dr. Theobold's.

Changing the Baseline Malaria Group

22. Complete the sentence:

The percentage who developed malaria among those who received the high dose is _____ %
smaller than the percentage who developed malaria who received the placebo.

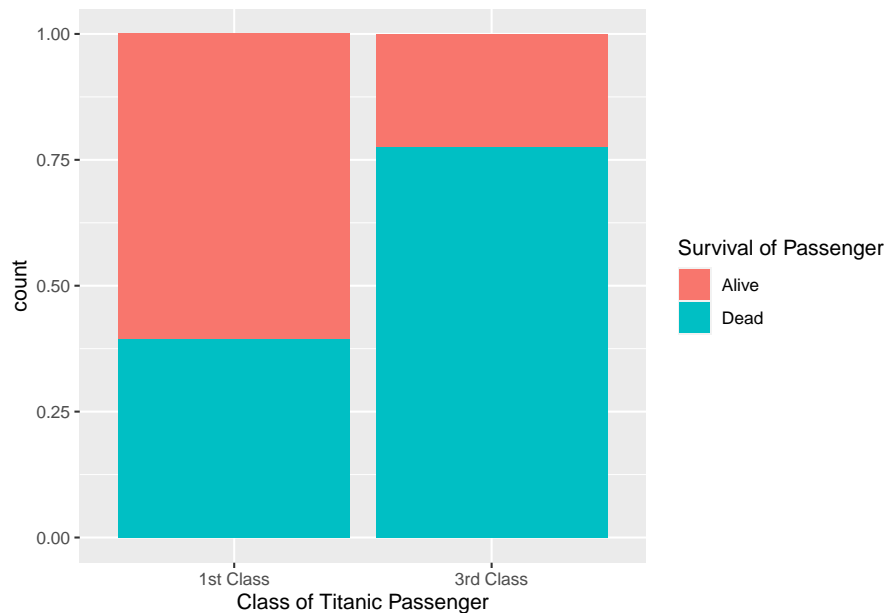
23. Write a different sentence that uses this value in context.

Titanic Survivors

A complete data set exists listing all those aboard HMS Titanic and includes related facts about each person including age, how much they paid for their ticket, which boat they survived in (if they survived), and their job if they were crew members. Stories, biographies and pictures can be found on the site: www.encyclopedia-titanica.org/. Did all passengers aboard the Titanic have the same chance of survival? Was the risk of death higher among 3rd class passengers compared to 1st class passengers?

Data Exploration

Below is a filled bar plot, visualizing the relationship



24. Based on the plot, does it appear that there is a relationship between a passenger's probability of survival and their class?

Next, we'll look at a table of the observed counts for the deaths of 1st class passengers and 3rd class passengers.

Survived	1st Class	3rd Class	Total
Alive	166	147	313
Dead	108	509	617
Total	274	656	930

25. Calculate the conditional proportion of 1st class passengers that died.

26. Calculate the conditional proportion of 3rd class passengers that died.
27. Calculate the percentage difference in the proportion of passengers who died, comparing the 3rd class passengers to the 1st class passengers.
28. Write a sentence to interpret this value in context.
29. Calculate the relative risk of dying, comparing the 3rd class passengers to the 1st class passengers.
30. Write a sentence to interpret this value in context.