Chapter Seven

Hypothesis Testing with \( z \) Tests

Note to Instructor

This chapter is critical to an understanding of hypothesis testing, which students will use frequently in the coming chapters. Some of the most essential topics covered in this chapter include the steps involved in hypothesis testing. To help students understand these steps, be sure to use activities in the textbook and the discussion questions provided here. You may also provide groups of students with the six steps written on note cards, and have the students work in their groups to put the steps in the proper order.

Outline of Resources

I. The \( z \) Table
   - Discussion Question 7-1 (p. 60)
   - Discussion Question 7-2 (p. 60)
   - Classroom Activity 7-1: Determining Percentiles (p. 61)
   - Discussion Question 7-3 (p. 61)
   - Discussion Question 7-4 (p. 61)
   - Classroom Activity 7-2: Make It Your Own: Statistics Test Scores (p. 62)

II. The Assumptions and the Steps of Hypothesis Testing
   - Discussion Question 7-5 (p. 62)
   - Discussion Question 7-6 (p. 63)
   - Discussion Question 7-7 (p. 63)
   - Classroom Activity 7-3: Make It Your Own: Short-Term Memory (p. 64)

III. An Example of the \( z \) Test
   - Discussion Question 7-8 (p. 65)
   - Classroom Activity 7-4: Research in the Media (p. 65)
   - Additional Readings (p. 66)
   - Online Resources (p. 66)

IV. Next Steps: Cleaning Data
   - Discussion Question 7-9 (p. 65)
   - Classroom Activity 7-5: Cleaning Data (p. 66)

V. Handouts
   - Handout 7-1: Determining Percentiles (p. 67)
I. The $z$ Table

1. After converting a raw score into a $z$ score, we can use the $z$ table to look up a given $z$ score and find the percentage of scores between the mean and that $z$ score.

2. The $z$ scores in the table are all positive because this normal curve is symmetrical. As a result, calculating the percentage between the mean and a given positive $z$ score is identical to calculating the percentage between the mean and the negative version of that $z$ score.

> Discussion Question 7-1

*Why are there no negative values in the $z$ table?*

Your students’ answers should include:

- There are no negative values in the $z$ table because the percentage of scores between the mean and a positive score is equivalent to the percentage of scores between the mean and a negative $z$ score, so inclusion of the negative values would be redundant.

3. If we are given the population mean and standard deviation, we can determine the percentile score of one individual and the percentage of scores below and above the individual’s score. We can also find scores at least as extreme as our individual’s score in both directions.

4. With a drawing of the bell-shaped curve, we can then label the midpoint with a $z$ score of 0. We can also convert raw scores of interest to $z$ scores and add this $z$ score to our bell curve. We should then shade the portion of the curve for which we are trying to calculate percentages. We can also look up the $z$ score on the table to determine its percentage. Lastly, we can use addition or subtraction to calculate the percentage of our shaded portion.

> Discussion Question 7-2

*If you were given an individual’s raw score and it was above the mean, what steps would you take to determine that individual’s percentile?*

Your students’ answers should include:

- If you were given an individual’s raw score and it was above the mean and you wanted to find that individual’s percentile, you would first convert the raw score to a $z$ score and add 50% to that $z$ score, which represents the percentage of scores below the mean.
Classroom Activity 7-1
Determining Percentiles

For this activity, students will determine their percentile for a particular variable.

- You could have students complete a questionnaire and use their score from the questionnaire for this activity, or for the purpose of this activity, you could record students’ heights (be sure to record the entire heights in inches only). Whether you decide to use a questionnaire or heights, make sure you choose a neutral topic.
- Once you collect the information, calculate the mean and standard deviation.
- Then, divide the class into small groups, and have each person in the group determine their percentile for the variable.

Use Handout 7-1, found at the end of this chapter, to structure this activity.

5. In addition, we can use the bell curve to determine a raw score from a z score.

6. To do this, we can draw a picture of the curve including the percentage with which we are starting. After drawing a vertical line at this percentage and the mean, we can shade the area of the curve between the percentage and the mean that is of interest. We can use addition or subtraction as needed to calculate the percentage between the mean and the z score of interest. We then look up the percentage on the z table (adding a negative sign if it is below the mean) and convert the z score to a raw score.

7. When converting the z score to a raw score, we want to use the formula: 
   \[ X = z(\sigma) + \mu. \]

> Discussion Question 7-3

If you were given an individual’s z score and it was less than zero, what steps would you take to determine that individual’s raw score?

Your students’ answers should include:

- If you were given an individual’s z score and it was less than zero, and you wanted to find that individual’s raw score, you would first subtract the percentage between the mean and the z score of interest. Next, you would look up the percentage on the z table. Finally, you would convert the z score to a raw score using the formula: \( X = z(\sigma) + \mu. \)

8. The same process can be conducted on a sample mean instead of a score. In this case, we must use the mean and the standard error of the distribution of scores. The mean will be the same, but we need to divide the standard deviation by the square root of the sample size.

> Discussion Question 7-4

How is the process of determining percentiles different if you were given a sample mean rather than an individual raw score?
Your students’ answers should include:

- The process of determining percentiles is different for a given sample mean rather than an individual raw score because, for a sample mean, you use the mean and standard error of the distribution of means instead of the mean and the standard deviation of the distribution of scores.

Classroom Activity 7-2
Make It Your Own: Statistics Test Scores

Use your students’ latest statistics test results as the basis for this exercise.

- Calculate the class mean and standard deviation as the population mean and standard deviation.

- Assuming each student knows their own score, have them calculate their \( z \) score and percentile rank.

II. The Assumptions and the Steps of Hypothesis Testing

1. In the field of statistics, **assumptions** are the characteristics that we ideally require the population (from which we are sampling) to have so that we can make accurate inferences.

2. **Parametric tests** are statistical analyses based on a set of assumptions about the population whereas **nonparametric tests** are statistical analyses that are *not* based on a set of assumptions about the population.

3. In hypothesis testing, we assume that the dependent variable is assessed using a scale measure.

4. Our second assumption in hypothesis testing is that the participants are randomly selected.

5. Our final assumption is that the distribution of the population of interest is approximately normal.

> Discussion Question 7-5

*What three assumptions do we make in hypothesis testing?*

Your students’ answers should include:

- Three assumptions that you make in hypothesis testing are (1) that the dependent variable is measured on an interval scale, (2) that participants in the sample have been selected randomly, and (3) that the underlying population distribution is well approximated by a normal curve.

6. Sometimes in hypothesis testing our assumptions are not met, but our statistics are **robust** against violations of the assumptions. Robust hypothesis tests are those that produce fairly accurate results even when the data suggest that the population might not meet some of the assumptions.

7. The three statistical assumptions represent the ideal conditions of research, but not meeting the assumptions doesn’t necessarily invalidate the research.
Discussion Question 7-6
If one of your assumptions is violated, should you abandon the research? Why or why not?
Your students’ answers should include:
- If one of your assumptions is violated, you are not required to abandon your research.
- This is because the inferential statistics that you use are often robust against violations of these assumptions and can still produce comparative results.

8. The first step in hypothesis testing is to identify the populations to be compared, the comparison distribution, the appropriate test, and its assumptions. We want to know that it is okay to proceed with a particular statistical analysis.

9. The second step is to state the null and research hypotheses.

10. Our third step is to explicitly state the relevant characteristics of the comparison distribution or the distribution based on the null hypothesis.

11. Our fourth step is to determine critical values, or cutoffs, indicating how extreme our data must be to reject the null hypothesis. Critical values are the test statistic values beyond which we will reject the null hypothesis. In most cases, we determine two cutoffs, one for extreme samples below the mean and one for extreme cases above the mean. In general, the critical values are the most extreme 5% of the distribution curve, which becomes 2.5% on either end. The critical region refers to the area in the tails of the distribution in which we reject the null hypothesis if our test statistic falls there. The probabilities used to determine the critical values in hypothesis testing are called p levels (or alphas).

12. Our fifth step is to calculate our test statistic.

13. Lastly, we decide whether to reject or fail to reject the null hypothesis. If the test statistic falls in the critical region, we know that it is in the most extreme 5% of possible test statistics only if the hypothesis is true.

Discussion Question 7-7
What the six steps of hypothesis testing?
Your students’ answers should include:
The six steps of hypothesis testing are:
- In step one, state the population represented by each of the samples, list the comparison distribution, state the appropriate hypothesis test, and check the assumptions for that hypothesis test.
- In step two, explicitly state the null and alternative hypothesis.
- In step three, determine the appropriate mean and the measure of spread for the comparison distribution to clarify the distribution represented by the null hypothesis.
- In step four, determine the critical values, or cutoffs, that indicate the points beyond which you will reject the null hypothesis.
In step five, calculate the test statistic and compare the test statistic to the obtained critical value.

In step six, decide whether to reject or fail to reject the null hypothesis.

Classroom Activity 7-3
Make It Your Own: Short-Term Memory

Use Miller's classic paper as the basis for this class demonstration. See Miller, G. A. (March 1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. To view or purchase this article, go to your local library or visit the American Psychological Association online at http://www.apa.org. Use the 7 ± 2 as the basis for a z score distribution.

Tell the students: “We are going to conduct a memory test. Everyone take out a piece of paper. Please listen carefully. I am going to tell you some words and, when I am done, I will ask you to try to write down all the words you can recall.” Read the list of words. (See Transparency Master 7-1 found at the end of this chapter. You may display the transparency on an overhead projector by photocopying it onto acetate, or you may use PowerPoint by scanning the transparency master into your computer.)

Tell the students: “Please write down all the words you can remember. The order of the words is not important.”

Have the students calculate their individual z scores using the formula given in the text: \( z = \frac{\text{score} - 7}{2} \) adapted for this task.

Have the students graph their scores.

Next, have the students determine their percentile scores.

With your students, complete the six steps of hypothesis testing, including calculating the z test. (See Transparency Master 7-2 found at the end of this chapter. Again, you may display the transparency on an overhead projector by photocopying it onto acetate, or you may use PowerPoint by scanning the transparency master into your computer.)

14. If we reject the null hypothesis because the pattern in the data differs from what we would expect by chance, we say that our results are statistically significant.

III. An Example of the z Test

1. A one-tailed test is a hypothesis test in which the research hypothesis is directional, positing either a decrease or an increase in the dependent variable, but not both, as a result of the independent variable. One-tailed tests are rarely used because the researcher must be absolutely certain that the effect cannot go in the other direction.

2. In contrast, a two-tailed test is a hypothesis test in which the research hypothesis does not indicate a direction of difference or change in the dependent variable, but merely indicates that there will be a difference. It is more conservative to begin with using a two-tailed test.
> **Discussion Question 7-8**

*What is the difference between a one-tailed and two-tailed test? Why is a two-tailed test preferred over a one-tailed test?*

Your students’ answers should include:

- The difference between a one-tailed and a two-tailed test is that in a one-tailed test the hypothesis is directional in that you would expect to see either a decrease or an increase on the dependent variable, compared to a two-tailed test in which you would hypothesize a difference without specifying an increase or decrease on the dependent variable.

**Classroom Activity 7-4**

Research in the Media

For this assignment, students will search for articles in the news about research (or you could provide them with the articles). Using these articles, students will determine in small groups whether the original hypothesis of the researchers was a one-tailed or two-tailed test. Handout 7-2, found at the end of this chapter, is designed to help your students with this process.

3. When we use a sample mean, rather than a single score in our test, we must use the standard error of the mean instead of the population standard deviation of the scores.

**IV. Next Steps: Cleaning Data**

1. “Dirty data” refers to missing data, misleading data, and outliers.
2. Sometimes it is a good idea to throw out dirty data. This is especially the case when data loss is widespread, systematic, affects most of the participants in a particular experimental condition, or affects only a few individuals but almost all of their data.
3. If there is occasional, nonsystematic loss of data, researchers can choose to use the modal or mean score for that variable, the modal or mean score from that participant’s other responses (if there are similar ones), or a random number that falls within the range of possible numbers.

> **Discussion Question 7-9**

*When is it best to throw out dirty data? When should we keep our dirty data? Why?*

Your students’ answers should include:

- It is best to throw out dirty data when there is a significant amount of missing data or if there are systematic patterns to the missing data that can lead to misleading conclusions.

4. We can use a similar solution for misleading data depending on the type.
5. Outliers may be misleading data. $z$ scores can be used to identify outliers so that we can specify exactly how different one data point is from all the other data points.
Classroom Activity 7-5
Cleaning Data

Many researchers have an abundance of dirty data. If you are one of those researchers, bring to class an example of dirty data from your work. This could be either in the form of a completed questionnaire, a discussion of particular research scenarios, or computerized datasets. Alternatively, if you do not have access to dirty data, create some using different types of dirty data (systematic as well as nonsystematic missing data). Have students analyze the data and determine why it should be kept or thrown out. Use Handout 7-3, found at the end of this chapter, to aid in this activity.

Additional Readings


Reviews the pros and cons of hypothesis testing and cautions that there is still value in descriptive statistics. This paper looks at the pros and cons of hypothesis testing.


Baseball is a game that is analyzed on the basis of statistics: batting averages, pitching statistics, and so on. If you are a baseball enthusiast, then this book may be very helpful in bringing the real world into your classroom. This book is a collection of case studies applying probabilistic thinking to the game of baseball. Section Three (Comparing Batches and Standardization) is particularly relevant to this chapter.

Online Resources

The Merlot Multimedia Educational Resource for Learning and Online Teaching has an extensive collection of demonstrations with explanations for most statistical topics.

http://www.merlot.org/merlot/materials.htm?category=2595
HANDOUT 7-1: DETERMINING PERCENTILES

Directions: For this assignment, your instructor should have given you the mean and standard deviation for the variable of interest for the entire class. With this information, answer the following questions regarding your percentile.

1. Describe the steps you would take to calculate your percentile for this variable.

2. What is the $z$ score for your score for this variable?

3. What is your percentile on this variable?

4. What percentage of scores is above your score and below your score?
HANDOUT 7-2: RESEARCH IN THE MEDIA

Directions: For this assignment, you need to gather newspaper articles discussing different types of scientific research. Answer each of the questions below based on these articles.

1. Summarize the research described in each of the articles in the space below.

2. For each article, based on the hypotheses of the researchers, do you think that they used a one-tailed or two-tailed test? Why?

3. Did you find more one-tailed tests or two-tailed tests? Why do you think you observed more of one type of test?
HANDOUT 7-3: CLEANING DATA

Directions: For this activity, your instructor will give you examples of dirty data. Use this handout to interpret this data.

1. In the space below, describe the data that your instructor presented to you. What does the data measure?

2. Should any of the data be thrown out? Which data, and why?

3. Should any of the data be kept? Which data, and why?

4. If there are data that we can keep, how should we modify the data so that we can use them in our analyses?
Word List

Station  Consumer
Director  Interview
Investment  Passenger
Album  Engine
Facility  Resource
Monitor  Telescope
Cabinet  Feature
Driveway  Volunteer
Passenger  District
Vessel  Mission
The Six Steps of Hypothesis Testing

Step 1. Populations, distributions, and assumptions.
   The populations are our class and all those who have ever completed Miller’s task.
   The comparison distribution is the distribution of means.
   The hypothesis test is the \( z \) test because we know the population mean (7) and population standard deviation (2).

Step 2. State the null and research hypotheses.
   \[ H_0 \]
   \[ H_1 \]

Step 3. Determine the characteristics of the comparison distributions.
   Here we can assume that the population mean is 7 and the population standard deviation is 2.

Step 4. Determine the critical values, or cutoffs.
   By behavioral science convention, \( p < 0.05 \): that is, 2.5% of the comparison distribution at the bottom of the distribution and 2.5% of the distribution at the top of the distribution. (See Figure 7-10, p. 181, in your textbook.) The critical values of the \( z \) score are \(-1.96\) and \(1.96\).

Step 5. Calculate the test statistics.
   \[ z = (7 - \text{class mean})/2 \]

Step 6. Make a decision.
   Compare the test statistic with the critical values of \( z \). Is it beyond the cutoff?