8.1.5: Additional Information on Hypothesis Tests

- In a hypothesis test problem, you may see words such as "the level of significance is 1%." The "1%" is the preconceived or preset α.
- The statistician setting up the hypothesis test selects the value of α to use before collecting the sample data.
- If no level of significance is given, a common standard to use is $\alpha = 0.05$.
- When you calculate the p-value and draw the picture, the p-value is the area in the left tail, the right tail, or split evenly between the two tails. For this reason, we call the hypothesis test left, right, or two tailed.
- The alternative hypothesis, H_a, tells you if the test is left, right, or two-tailed. It is the key to conducting the appropriate test.
- H_a never has a symbol that contains an equal sign.
- Thinking about the meaning of the p-value: A data analyst (and anyone else) should have more confidence that he made the correct decision to reject the null hypothesis with a smaller p-value (for example, 0.001 as opposed to 0.04) even if using the 0.05 level for alpha. Similarly, for a large p-value such as 0.4, as opposed to a p-value of 0.056 ($\alpha = 0.05$ is less than either number), a data analyst should have more confidence that she made the correct decision in not rejecting the null hypothesis. This makes the data analyst use judgment rather than mindlessly applying rules.

The following examples illustrate a left-, right-, and two-tailed test.

Example 8.1.5.1

$H_0 : \mu = 5, H_a : \mu < 5$

Test of a single population mean. H_a tells you the test is left-tailed. The picture of the p-value is as follows:

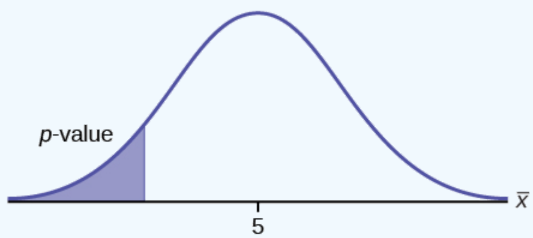

Figure 8.1.5.1

Exercise 8.1.5.1

$H_0 : \mu = 10, H_a : \mu < 10$

Assume the p-value is 0.0935. What type of test is this? Draw the picture of the p-value.

Answer

left-tailed test

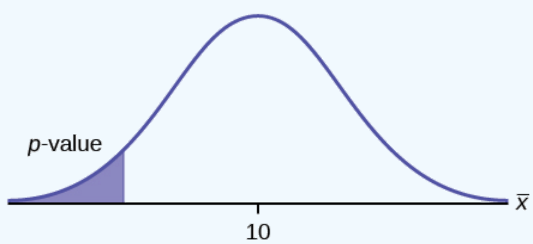

Figure 8.1.5.2

Example 8.1.5.2

$H_0 : \mu \leq 0.2, H_a : \mu < 0.2$

This is a test of a single population proportion. H_a tells you the test is **right-tailed**. The picture of the p-value is as follows:

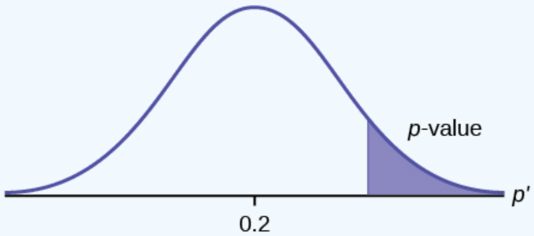

Figure 8.1.5.3

Exercise 8.1.5.2

$H_0 : \mu \leq 1, H_a : \mu > 1$

Assume the p-value is 0.1243. What type of test is this? Draw the picture of the p-value.

Answer

right-tailed test

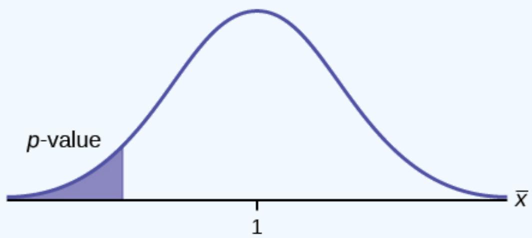

Figure 8.1.5.4

Example 8.1.5.3

$H_0 : \mu = 50, H_a : \mu \neq 50$

This is a test of a single population mean. H_a tells you the test is **two-tailed**. The picture of the p-value is as follows.

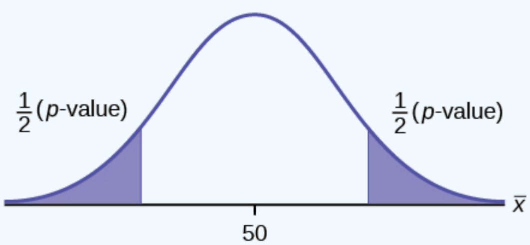

Figure 8.1.5.5

Exercise 8.1.5.3

$H_0 : \mu = 0.5, H_a : \mu \neq 0.5$

Assume the p-value is 0.2564. What type of test is this? Draw the picture of the p-value.

Answer

two-tailed test

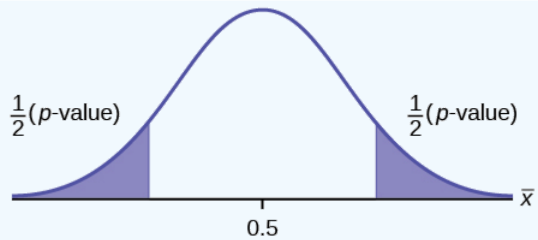

Figure 8.1.5.6

The **hypothesis test** itself has an established process. This can be summarized as follows:

1. Determine H_0 and H_a. Remember, they are contradictory.
2. Determine the random variable.
3. Determine the distribution for the test.
4. Draw a graph, calculate the test statistic, and use the test statistic to calculate the $p\text{-value}$. (A *z*-score and a *t*-score are examples of test statistics.)
5. Compare the preconceived α with the *p*-value, make a decision (reject or do not reject H_0), and write a clear conclusion using English sentences.

Notice that in performing the hypothesis test, you use α and not β. β is needed to help determine the sample size of the data that is used in calculating the $p\text{-value}$. Remember that the quantity $1-\beta$ is called the **Power of the Test**. A high power is desirable. If the power is too low, statisticians typically increase the sample size while keeping α the same.If the power is low, the null hypothesis might not be rejected when it should be.

Exercise 9.6.8

Assume $H_0 : \mu = 9$ and $H_a : \mu < 9$. Is this a left-tailed, right-tailed, or two-tailed test?

Answer

This is a left-tailed test.

Exercise 9.6.9

Assume $H_0 : \mu \leq 6$ and $H_a : \mu > 6$. Is this a left-tailed, right-tailed, or two-tailed test?

Exercise 9.6.10

Assume $H_0 : p = 0.25$ and $H_a : p \neq 0.25$. Is this a left-tailed, right-tailed, or two-tailed test?

Answer

This is a two-tailed test.

Exercise 9.6.11

Draw the general graph of a left-tailed test.

Exercise 9.6.12

Draw the graph of a two-tailed test.

Answer

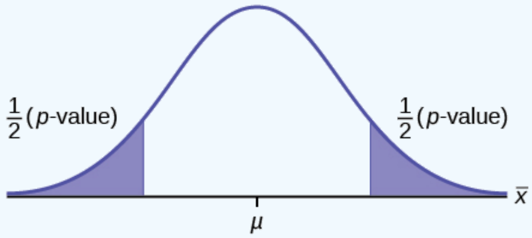

Figure 8.1.5.16

Exercise 9.6.13

A bottle of water is labeled as containing 16 fluid ounces of water. You believe it is less than that. What type of test would you use?

Exercise 9.6.14

Your friend claims that his mean golf score is 63. You want to show that it is higher than that. What type of test would you use?

Answer

a right-tailed test

Exercise 9.6.15

A bathroom scale claims to be able to identify correctly any weight within a pound. You think that it cannot be that accurate. What type of test would you use?

Exercise 9.6.16

You flip a coin and record whether it shows heads or tails. You know the probability of getting heads is 50%, but you think it is less for this particular coin. What type of test would you use?

Answer

a left-tailed test

Exercise 9.6.17

If the alternative hypothesis has a not equals ($\neq$) symbol, you know to use which type of test?

Exercise 9.6.18

Assume the null hypothesis states that the mean is at least 18. Is this a left-tailed, right-tailed, or two-tailed test?

Answer

This is a left-tailed test.

Exercise 9.6.19

Assume the null hypothesis states that the mean is at most 12. Is this a left-tailed, right-tailed, or two-tailed test?

Exercise 9.6.20

Assume the null hypothesis states that the mean is equal to 88. The alternative hypothesis states that the mean is not equal to 88. Is this a left-tailed, right-tailed, or two-tailed test?

Answer

This is a two-tailed test.

References

1. Data from Amit Schitai. Director of Instructional Technology and Distance Learning. LBCC.
2. Data from *Bloomberg Businessweek*. Available online at http://www.businessweek.com/news/2011- 09-15/nyc-smoking-rate-falls-to-record-low-of-14-bloomberg-says.html.
3. Data from energy.gov. Available online at http://energy.gov (accessed June 27. 2013).
4. Data from Gallup®. Available online at www.gallup.com (accessed June 27, 2013).
5. Data from *Growing by Degrees* by Allen and Seaman.
6. Data from La Leche League International. Available online at www.lalecheleague.org/Law/BAFeb01.html.
7. Data from the American Automobile Association. Available online at www.aaa.com (accessed June 27, 2013).
8. Data from the American Library Association. Available online at www.ala.org (accessed June 27, 2013).
9. Data from the Bureau of Labor Statistics. Available online at http://www.bls.gov/oes/current/oes291111.htm.
10. Data from the Centers for Disease Control and Prevention. Available online at www.cdc.gov (accessed June 27, 2013)
11. Data from the U.S. Census Bureau, available online at quickfacts.census.gov/qfd/states/00000.html (accessed June 27, 2013).
12. Data from the United States Census Bureau. Available online at www.census.gov/hhes/socdemo/language/.
13. Data from Toastmasters International. Available online at http://toastmasters.org/artisan/deta...eID=429&Page=1.
14. Data from Weather Underground. Available online at www.wunderground.com (accessed June 27, 2013).
15. Federal Bureau of Investigations. "Uniform Crime Reports and Index of Crime in Daviess in the State of Kentucky enforced by Daviess County from 1985 to 2005." Available online at http://www.disastercenter.com/kentucky/crime/3868.htm (accessed June 27, 2013).
16. "Foothill-De Anza Community College District." De Anza College, Winter 2006. Available online at research.fhda.edu/factbook/DA...t_da_2006w.pdf.
17. Johansen, C., J. Boice, Jr., J. McLaughlin, J. Olsen. "Cellular Telephones and Cancer—a Nationwide Cohort Study in Denmark." Institute of Cancer Epidemiology and the Danish Cancer Society, 93(3):203-7. Available online at http://www.ncbi.nlm.nih.gov/pubmed/11158188 (accessed June 27, 2013).
18. Rape, Abuse & Incest National Network. "How often does sexual assault occur?" RAINN, 2009. Available online at www.rainn.org/get-information...sexual-assault (accessed June 27, 2013).

Glossary

Central Limit Theorem

Given a random variable (RV) with known mean μ and known standard deviation σ. We are sampling with size n and we are interested in two new RVs - the sample mean, $\bar{X}$, and the sample sum, $\sum X$. If the size n of the sample is sufficiently large, then $\bar{X} - N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ and $\sum X - N\left(n\mu, \sqrt{n}\sigma\right)$. If the size *n* of the sample is sufficiently large, then the distribution of the sample means and the distribution of the sample sums will approximate a normal distribution regardless of the shape of the population. The mean of the sample means will equal the population mean and the mean of the sample sums will equal n times the population mean. The standard deviation of the distribution of the sample means, $\frac{\sigma}{\sqrt{n}}$, is called the standard error of the mean.

Contributors and Attributions

-

Barbara Illowsky and Susan Dean (De Anza College) with many other contributing authors. Content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at http://cnx.org/contents/30189442-699...b91b9de@18.114.