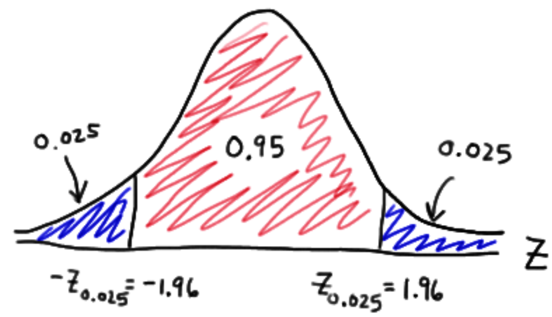
Hypothesis Testing and Confidence Intervals

Understand Z-Test, P-Value, Significance Level, Sampling Distribution, Central Limit Theorem, Confidence Level, and Confidence Intervals



Overview

In this post, I will use hypothesis testing and confidence intervals to answer the questions in an example and summarize the relationship between them. During the process, I will also introduce other important statistical concepts that I have applied.

Introduction

Statistical knowledge is significant for people who want to work in data science. In my first quarter at graduate program, I took the course "Probability Theory and Introductory Statistics" to learn basic statistics for data analysis. For instance, in this class I learned when analyzing data, inferential statistics is the process of using random sample data to make inferences about the larger population. The most common methodologies in inferential statistics are *hypothesis testing, confidence intervals,* and *regression*. I'll illustrate the first two concepts in the following example to show what I have learned!

Example

A researcher claims that the average wind speed in a certain city is 8 miles per hour. A random sample of 36 days has an average wind speed of 8.2 miles per hour. The standard deviation of the population is 0.6 miles per hour.

Question 1

At $\alpha = 0.05$, is there enough evidence to reject the claim?

Let's check assumptions before we state the hypotheses.

√ The sample is a random sample.

√ Either $n \geq 30$ or the population is normally distributed when $n < 30$ (n=sample size).

$$H_0: \mu = 8$$

$$H_1: \mu \neq 8$$

H_0 (Null hypothesis): Specifies that the mean will remain unchanged.

H_1 (Alternative hypothesis): States that the mean will be different.

Because of the symbol "≠", this is a two-tailed test. Also, since σ (standard deviation of the population) is known, the Z-test is used instead of the t-test. To solve question 1, we can use a Z-test or P-value method.

Z-test Method

Step 1: Find the critical values

A critical value is a point on the test distribution that splits the graph into sections to determine whether to reject the null hypothesis. According to Critical Values Calculator, when $\alpha = 0.05$ and the test is two-tailed ($\frac{\alpha}{2} = 0.025$), the critical values are +1.96 and -1.96. That is, $Z_{0.025}$ equals 1.96.

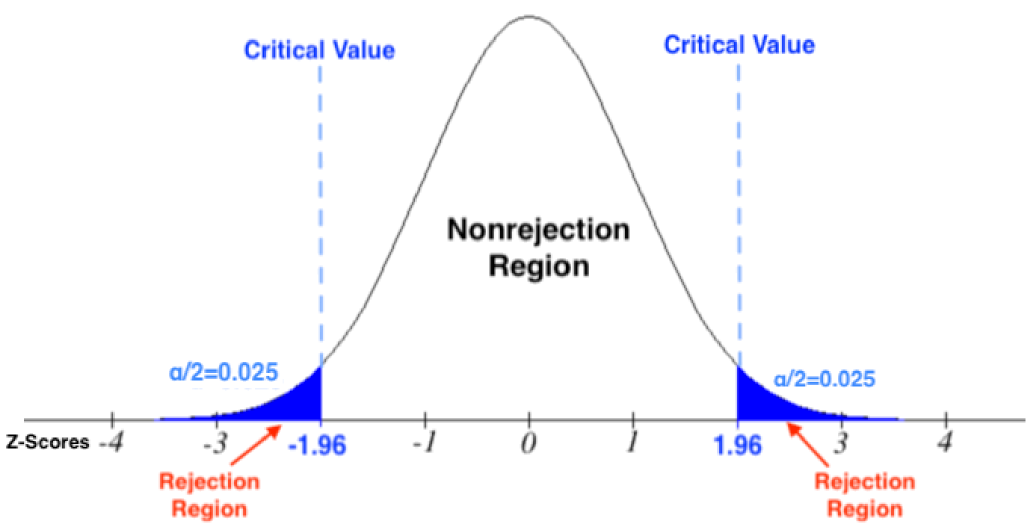


Figure 1. Critical Value.

Step 2: Compute the test value

The test value is computed to compare with the critical values. From the result, we can tell there are 2 standard errors between the sample mean and the population mean.

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{8.2-8}{0.6/\sqrt{36}} = 2$$

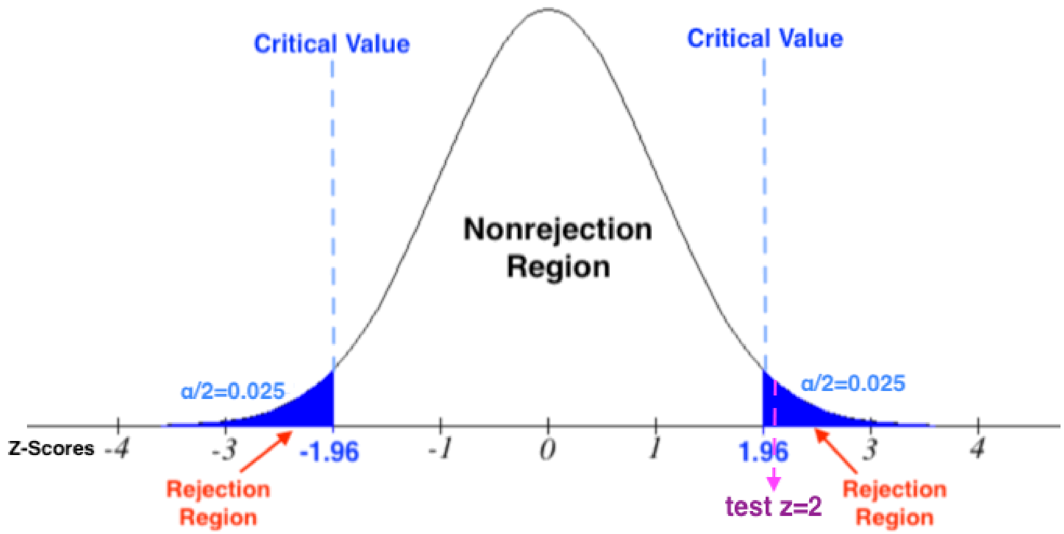


Figure 2. Test Value and Critical Value.

Step 3: Summarize the results

Figure 2 shows that the Z-test value (2) is greater than the critical value (1.96), falling in the critical region. Therefore, at the α = 0.05 level, there is enough evidence to reject the null hypothesis and to support the claim that the average wind speed in a certain city is not equal to 8 miles per hour.

P-value Method

Before we start using the P-value method, let's explain the difference between the significance level and the P-value. The significance level (α) is used to refer to a pre-chosen probability and the P-value is used to represent a probability that has been calculated after a given study.

Significance Level (α): the probability of rejecting the null hypothesis when it is true. That is, in question 1, we chose a significance level of 0.05, indicating a 5% risk of concluding that a difference exists when there is no actual difference.

P-value: the probability of a particular sample statistic or a more extreme sample statistic occurring if the null hypothesis is true. For instance, a small P-value means that the observed result is highly unlikely if the null hypothesis were true. Therefore, we may conclude that the null hypothesis is unlikely to be true and reject it.

Step 1: Compute the test value

From step 2 of the Z-test method, we computed that the test value is 2.

Step 2: Find the P-value

To compare with the significance level, we have to use test value Z=2 to find the corresponding area under the normal distribution using Z-Score Calculator and then subtract this value for the area from 1.0000.

$$1.0000 - 0.9772 = 0.0228$$

Since this is a two-tailed test, the area of 0.0228 must be doubled to get the P-value.

$$2 \times 0.0228 = 0.0456$$

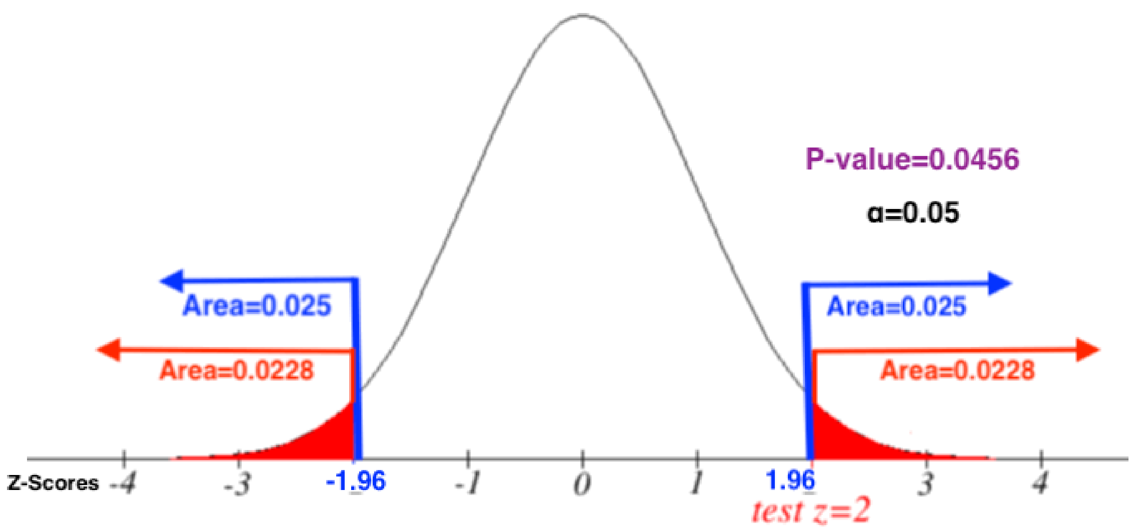


Figure 3. P-Value and Level of Significance.

Step 3: Summarize the results

Since the P-value (0.0456) is less than 0.05 level of significance, once again, there is sufficient evidence to reject the null hypothesis and accept the alternative hypothesis.

Note: In practice, we use one sample data + additional knowledge to find the parameters of a sampling distribution. For instance, based on the Central Limit Theorem, the sampling distribution of the sample mean will follow a normal distribution. Therefore, to test mean, we can draw normal distribution and locate sample data on it.

Sampling Distribution: The distribution of all possible values of a statistic for a given sample size.

Central Limit Theorem: If sample size ≥ 30 or population is normally distributed, the sampling distribution of the mean is approximately or exactly normal.

However, sometimes the Central Limit Theorem can't be applied. For example, when the sample size is small or σ is unknown. In these cases, to test mean, T-distribution would be used as the sampling distribution.

Question 2

Regardless of the research claim, based on this random sample data, what is the confidence interval for the population mean at a 95% confidence level?

A confidence interval is a specific interval estimate of a parameter determined by using data obtained from a sample and by using the specific confidence level of the estimate. According to question 2, the confidence level is 95%. Since the confidence level is equal to 1- α, the corresponding significance level is also 0.05 (1-0.95=0.05), the same as question 1.

Level of Confidence = Confidence Level = 1- α	95%
Level of Significance = Significance Level = α	0.05

Again, let's check assumptions before we find the confidence interval.

√ The sample is a random sample.

√ Either $n \geq 30$ or the population is normally distributed when $n < 30$ (n=sample size).

Step 1: Find the critical values

For a 95% confidence interval,

$$\alpha = 1-0.95 = 0.05$$

$$Z^{\alpha}/_2 = Z^{0.05}/_2 = Z_{0.025} = 1.96$$

Step 2: Compute confidence interval

$Z^{\alpha}/_2(\frac{\sigma}{\sqrt{n}})$ is called the margin of error, indicating the degree of error in results received from random sampling surveys. In this sample, the sample size equals 36, the sample mean equals 8.2, and the standard deviation of the population is 0.6.

Margin of error = Critical value * Standard Error (SE).

$$\bar{x} - Z^{\alpha}/_2(\frac{\sigma}{\sqrt{n}}) < \mu < \bar{x} + Z^{\alpha}/_2(\frac{\sigma}{\sqrt{n}})$$

$$8.2 - 1.96(\frac{0.6}{\sqrt{36}}) < \mu < 8.2 + 1.96(\frac{0.6}{\sqrt{36}})$$

$$8.004 < \mu < 8.396$$

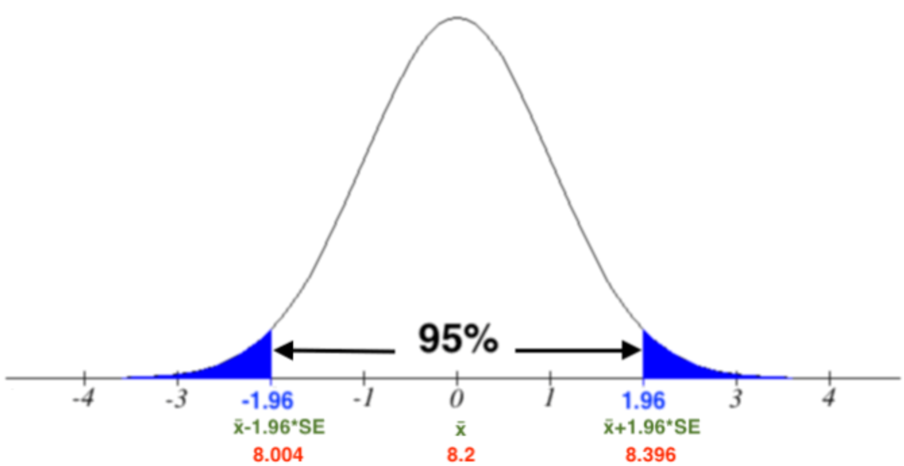


Figure 4. Confidence Interval.

Step 3: Summarize the results

Based on a sample of 36 days, we can be 95% confident that the mean number of the wind speed in a certain city (population mean) is between 8.004 and 8.396.

(Strictly speaking, 95% confident interval for the mean indicates that if repeated samples were taken and the 95% confidence interval was computed for each sample, 95% of the intervals would contain the population mean.)

Summary

In question 1, we know that there are 2 standard errors between the sample mean and the population mean. Then, in question 2, we calculated that the confidence interval is the sample mean ±1.96 standard errors. That is, the 95% confidence interval is constructed to be [8.004,8.396]. However, since the value specified by the null hypothesis in question 1 (H_0: $\mu = 8$) is not contained within the 95% confidence interval, it is not a reasonable estimate of the population mean. Thus, we would expect to have a P-value less than 0.05 and to reject the null hypothesis at the 0.05 level. Not surprisingly, our expectation is just the same as the result we obtained in question 1. In conclusion, when the confidence intervals and P-values are generated by the same hypothesis test and used an equivalent confidence level (significance level), the two approaches would agree.

References

Book: Elementary Statistics 10th Edition

P Values: https://www.statsdirect.com/help/basics/p_values.htm

How to Correctly Interpret P Values: https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-correctly-interpret-p-values

Confidence Intervals & Hypothesis Testing
https://online.stat.psu.edu/stat200/lesson/6/6.6

How Hypothesis Tests Work: Confidence Intervals and Confidence Levels
https://statisticsbyjim.com/hypothesis-testing/hypothesis-tests-confidence-intervals-levels/

Graph generator: http://www.imathas.com/stattools/norm.html

z-Score Calculator: https://www.zscorecalculator.com/

Critical Values Calculator:
https://www.socscistatistics.com/tests/criticalvalues/default.aspx