STA 2023

Module 11B
Hypothesis Tests for
One Population Mean

Learning Objectives

Upon completing this module, you should be able to

1. Define the terms associated with hypothesis testing.
2. Choose the null and alternative hypotheses for a hypothesis test.
3. Explain the logic behind hypothesis testing.
4. Identify the test statistic, rejection region, non-rejection region, and critical value(s) for a hypothesis test.
5. Define and apply the concepts of Type I and Type II errors.
6. State and interpret the possible conclusions for a hypothesis test.

Learning Objectives

7. Obtain the critical value(s) for a specified significance level.
8. Perform a hypothesis test for a population mean when the population standard deviation is known.
9. Obtain the P-value of a hypothesis test.
10. State and apply the steps for performing a hypothesis test, using the critical-value approach to hypothesis testing.
11. State and apply the steps for performing a hypothesis test, using the P-value approach to hypothesis testing.
12. Perform a hypothesis test for a population mean when the population standard deviation is unknown.

How a Statistic is Used?

In the first part of Module 11, we looked at methods for obtaining the confidence intervals for one population mean.

We have learned that a confidence interval for a population mean, is based on a sample mean (a statistic).

In this second part of Module 11, we are going to look at how a statistic (sample mean) is used to make decisions about hypothesized values of a parameter (population mean.)

Quick Review: Decision and Hypothesis Test

One of the commonly used methods for making decision is to perform a hypothesis test

A hypothesis test involves two hypotheses: the null hypothesis and the alternative hypothesis.

"Testing a hypothesis" is like "testing a claim."

What are Hypotheses?

- Hypotheses are working models that we adopt temporarily.
- Our starting hypothesis is called the null hypothesis.
- The null hypothesis, that we denote by H_0, specifies a population model parameter of interest and proposes a value for that parameter.
- We usually write down the null hypothesis in the form H_0: *parameter* = *hypothesized* value.
- The alternative hypothesis, which we denote by H_A, contains the value of the parameter that we consider plausible when we reject the null hypothesis.

Testing Hypotheses

- The first step in defining the null and alternative hypotheses is to determine which parameter is being tested. A parameter describes a population. Examples are the population mean, the population standard deviation and a population proportion.
- The next step is to define the null hypothesis, specifies a population model parameter of interest and proposes a value for that parameter.
  - We might have, for example, H_0: *parameter* = 0.20
  - We want to compare our data to what we would expect, given that H_0 is true.
- Note that the null hypothesis is a statement of equality (with the equal sign.)

A Trial as a Hypothesis Test

- Think about the logic of jury trials:
  - To prove someone is guilty, we start by *assuming* they are innocent.
  - We retain that hypothesis until the facts make it unlikely beyond a reasonable doubt.
  - Then, and only then, we reject the hypothesis of innocence and declare the person guilty.

A Trial as a Hypothesis Test (cont.)

- The same logic used in jury trials is used in statistical tests of hypotheses:
  - We begin by assuming that a hypothesis is true.
  - Next we consider whether the data are consistent with the hypothesis.
  - If they are, all we can do is retain the hypothesis we started with. If they are not, then like a jury, we ask whether they are unlikely beyond a reasonable doubt.

Quick Review (cont.)
What are Null Hypothesis, Alternative Hypothesis and Hypothesis Test?

Null and Alternative Hypotheses; Hypothesis Test

Null hypothesis: A hypothesis to be tested. We use the symbol H_0 to represent the null hypothesis.

Alternative hypothesis: A hypothesis to be considered as an alternative to the null hypothesis. We use the symbol H_a to represent the alternative hypothesis.

Hypothesis test: The problem in a hypothesis test is to decide whether the null hypothesis should be rejected in favor of the alternative hypothesis.

What are Type I Error and Type II Error?

Type I and Type II Errors

Type I error: Rejecting the null hypothesis when it is in fact true.

Type II error: Not rejecting the null hypothesis when it is in fact false.

The probability of making type I error, that is, of rejecting a true null hypothesis, is called the significance level of a hypothesis test.

What is the Power of a Hypothesis Test?

Power

The **power** of a hypothesis test is the probability of not making a Type II error, that is, the probability of rejecting a false null hypothesis. We have

$$\text{Power} = 1 - P(\text{Type II error}) = 1 - \beta.$$

What to do with an “Innocent” defendant?

- If the evidence is not strong enough to reject the presumption of innocent, the jury returns with a verdict of “not guilty.”
  - The jury does not say that the defendant is innocent.
  - All it says is that there is not enough evidence to convict, to reject innocence.
  - The defendant may, in fact, be innocent, but the jury has no way to be sure.

What to Do with an “Innocent” Defendant ?(cont.)

- Said statistically, we will *fail to reject* the null hypothesis.
  - We never declare the null hypothesis to be true, because we simply do not know whether it's true or not.
  - Sometimes in this case we say that the *null hypothesis has been retained*.

What to Do with an “Innocent” Defendant ?(cont.)

- In a trial, the burden of proof is on the prosecution.
- In a hypothesis test, the burden of proof is on the unusual claim.
- The null hypothesis is the ordinary state of affairs, so it's the alternative to the null hypothesis that we consider unusual (and for which we must marshal evidence).

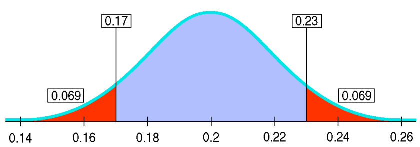
What are the Three Possible Alternative Hypotheses?

There are three possible alternative hypotheses:

- H_A: *parameter < hypothesized value*
- H_A: *parameter ≠ hypothesized value*
- H_A: *parameter > hypothesized value*

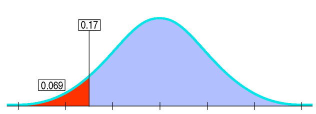
What is a Two-Sided Alternative Hypothesis?

- H_A: *parameter ≠ value* is known as a two-sided alternative or two-tailed alternative because we are equally interested in deviations on either side of the null hypothesis value.
- For two-sided alternatives, the P-value is the probability of deviating in *either* direction from the null hypothesis value.

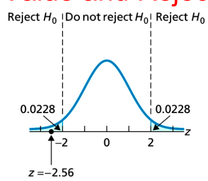


What is One-Sided Alternative Hypothesis?

- The other two alternative hypotheses are called one-sided alternatives or one-tailed alternative
- A one-sided alternative focuses on deviations from the null hypothesis value in only one direction.
- Thus, the P-value for one-sided alternatives is the probability of deviating *only in the direction of the alternative* away from the null hypothesis value.



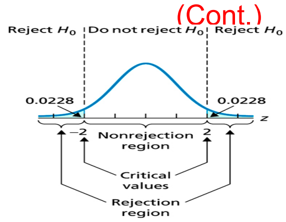
Critical Value and Rejection Region



The rejection region (shaded area) is based on the chosen significance level, which determines the critical value -2 and 2.

When the value of the test-statistic (z = -2.56) falls in this rejection region, we reject the null hypothesis.

Critical Value and Rejection Region (Cont.)



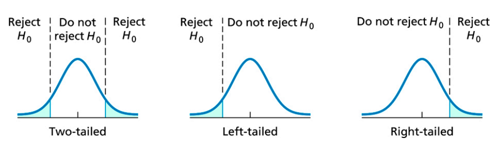
How to obtain the critical value -2 in this case (by TI-84+)?

Press the following keys: [2nd] [DISTR]

Then perform: invNorm(0.0228)

The critical value(s) can be obtained from the standard-normal table or technology, based on the chosen significance level. In this case, the significance level is 4.56%. Since this is a two-tailed alternative/two-tailed test, we divide the significance level by two (4.56/2 = 2.28% or 0.0228) for each tail.

Where is the Rejection Region?



For a two-tailed or two-sided test, the rejection region is on both the left and right. For a left-tailed test, the rejection region is on the left. For a right-tailed test, the rejection region is on the right.

Note that the rejection region is always at the tail(s).

One-Mean z-Test (Critical-value Approach)

One-Mean *z*-Test (Critical-Value Approach)

Purpose To perform a hypothesis test for a population mean, μ

Assumptions

1. Simple random sample
2. Normal population or large sample
3. σ known

STEP 1 **The null hypothesis is H_0: $\mu = \mu_0$, and the alternative hypothesis is**

$$H_a: \mu \neq \mu_0 \text{ (Two tailed)} \quad \text{or} \quad H_a: \mu < \mu_0 \text{ (Left tailed)} \quad \text{or} \quad H_a: \mu > \mu_0 \text{ (Right tailed)}$$

STEP 2 **Decide on the significance level, α.**

STEP 3 **Compute the value of the test statistic**

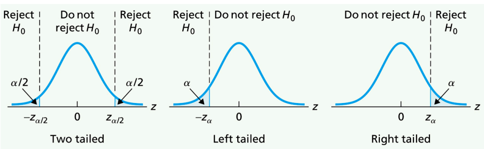
$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.$$

STEP 4 **The critical value(s) are**

$$\pm z_{\alpha/2} \text{ (Two tailed)} \quad \text{or} \quad -z_{\alpha} \text{ (Left tailed)} \quad \text{or} \quad z_{\alpha} \text{ (Right tailed)}$$

Use Table II to find the critical value(s).

One-Mean z-Test (Critical-value Approach)



STEP 5 **If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.**

STEP 6 **Interpret the results of the hypothesis test.**

The hypothesis test is exact for normal populations and is approximately correct for large samples from nonnormal populations.

How to obtain the P-value of a Hypothesis Test?

***P*-Value**

To obtain the ***P*-value** of a hypothesis test, we assume that the null hypothesis is true and compute the probability of observing a value of the test statistic as extreme as or more extreme than that observed. By *extreme* we mean "far from what we would expect to observe if the null hypothesis is true." We use the letter P to denote the P-value.

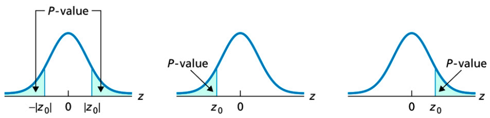
What is P-Value, again?

- The statistical twist is that we can quantify our level of doubt.
  - We can use the model proposed by our hypothesis to calculate the probability that the event we've witnessed could happen.
  - That's just the probability we're looking for—it quantifies exactly how surprised we are to see our results.
  - This probability is called a P-value.

What is P-Value, again? (cont.)

- When the data are consistent with the model from the null hypothesis, the P-value is high and we are unable to reject the null hypothesis.
  - In that case, we have to "retain" the null hypothesis we started with.
  - We can't claim to have proved it; instead we "*fail to reject the null hypothesis*" when the data are consistent with the null hypothesis model and in line with what we would expect from natural sampling variability.
- If the P-value is low enough, we'll "*reject the null hypothesis*," since what we observed would be very unlikely were the null model true.

Where is the P-value?



(a) Two-tailed (b) Left-tailed (c) Right-tailed

Note that P-value is the probability at the tail(s). If the P-value is small, we reject the null hypothesis.

P-Values and Decisions: What to Tell About a Hypothesis Test

- How small should the P-value be in order for you to reject the null hypothesis?
- It turns out that our decision criterion is context-dependent.
  - When we're screening for a disease and want to be sure we treat all those who are sick, we may be willing to reject the null hypothesis of no disease with a fairly large P-value.
  - A longstanding hypothesis, believed by many to be true, needs stronger evidence (and a correspondingly small P-value) to reject it.
- Another factor in choosing a P-value is the importance of the issue being tested.

P-Values and Decisions (cont.)

- Your conclusion about any null hypothesis should be accompanied by the P-value of the test.
  - If possible, it should also include a confidence interval for the parameter of interest.
- Don't just declare the null hypothesis rejected or not rejected.
  - Report the P-value to show the strength of the evidence against the hypothesis.
  - This will let each reader decide whether or not to reject the null hypothesis.

How to Use P-value as Evidence against the null hypothesis?

P-value	Evidence against H_0
$P > 0.10$	Weak or none
$0.05 < P \leq 0.10$	Moderate
$0.01 < P \leq 0.05$	Strong
$P \leq 0.01$	Very strong

One-Mean z-Test (P-value Approach)

One-Mean z-Test (P-Value Approach)

Purpose To perform a hypothesis test for a population mean, μ

Assumptions
1. Simple random sample
2. Normal population or large sample
3. σ known

STEP 1 **The null hypothesis is H_0: $\mu = \mu_0$, and the alternative hypothesis is**

$$H_a: \mu \neq \mu_0 \text{ (Two tailed)} \quad \text{or} \quad H_a: \mu < \mu_0 \text{ (Left tailed)} \quad \text{or} \quad H_a: \mu > \mu_0 \text{ (Right tailed)}$$

STEP 2 **Decide on the significance level, α.**

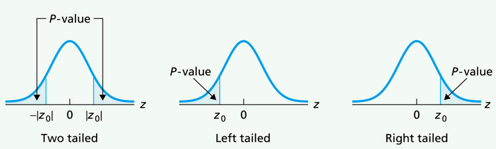
STEP 3 **Compute the value of the test statistic**

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

and denote that value z_0.

One-Mean z-Test (P-value Approach)

STEP 4 **Use Table II to obtain the P-value.**



STEP 5 **If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.**

STEP 6 **Interpret the results of the hypothesis test.**

The hypothesis test is exact for normal populations and is approximately correct for large samples from nonnormal populations.

Here is a Comparison of Critical-Value and P-Value Approach

CRITICAL-VALUE APPROACH	or	*P*-VALUE APPROACH
STEP 1 State the null and alternative hypotheses.		**STEP 1** State the null and alternative hypotheses.
STEP 2 Decide on the significance level, α.		**STEP 2** Decide on the significance level, α.
STEP 3 Compute the value of the test statistic.		**STEP 3** Compute the value of the test statistic.
STEP 4 Determine the critical value(s).		**STEP 4** Determine the P-value.
STEP 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.		**STEP 5** If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.
STEP 6 Interpret the result of the hypothesis test.		**STEP 6** Interpret the result of the hypothesis test.

How to conduct a One-Mean t-Test?

One-Mean *t*-Test

Purpose To perform a hypothesis test for a population mean, μ

Assumptions

1. Simple random sample
2. Normal population or large sample
3. σ unknown

STEP 1 The null hypothesis is H_0: $\mu = \mu_0$, and the alternative hypothesis is

$$H_a: \mu \neq \mu_0 \text{ (Two tailed)} \quad \text{or} \quad H_a: \mu < \mu_0 \text{ (Left tailed)} \quad \text{or} \quad H_a: \mu > \mu_0 \text{ (Right tailed)}$$

STEP 2 Decide on the significance level, α.

STEP 3 Compute the value of the test statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

and denote that value t_0.

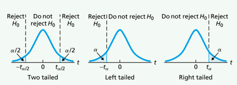
How to Conduct a One-Mean t-Test

CRITICAL-VALUE APPROACH or ***P*-VALUE APPROACH**

STEP 4 The critical value(s) are

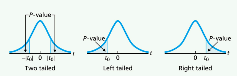
$$\pm t_{\alpha/2} \text{ (Two tailed)} \quad \text{or} \quad -t_\alpha \text{ (Left tailed)} \quad \text{or} \quad t_\alpha \text{ (Right tailed)}$$

with df $= n - 1$. Use Table IV to find the critical value(s).



STEP 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

STEP 4 The t-statistic has df $= n - 1$. Use Table IV to estimate the P-value, or obtain it exactly by using technology.



STEP 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

STEP 6 Interpret the results of the hypothesis test.

The hypothesis test is exact for normal populations and is approximately correct for large samples from nonnormal populations.

What Can Go Wrong?

- Don't base your null hypothesis on what you see in the data.
  - *Think* about the situation you are investigating and develop your null hypothesis appropriately.
- Don't base your alternative hypothesis on the data, either.
  - Again, you need to *Think* about the situation.

What Can Go Wrong? (cont.)

- Don't make your null hypothesis what you want to show to be true.
  - You can reject the null hypothesis, but you can never "accept" or "prove" the null.
- Don't forget to check the conditions.
  - We need randomization, independence, and a sample that is large enough to justify the use of the Normal model.

What have we learned?

We have learned to:

1. Define the terms associated with hypothesis testing.
2. Choose the null and alternative hypotheses for a hypothesis test.
3. Explain the logic behind hypothesis testing.
4. Identify the test statistic, rejection region, non-rejection region, and critical value(s) for a hypothesis test.
5. Define and apply the concepts of Type I and Type II errors.
6. State and interpret the possible conclusions for a hypothesis test.

What have we learned? (Cont.)

7. Obtain the critical value(s) for a specified significance level.
8. Perform a hypothesis test for a population mean when the population standard deviation is known.
9. Obtain the P-value of a hypothesis test.
10. State and apply the steps for performing a hypothesis test, using the critical-value approach to hypothesis testing.
11. State and apply the steps for performing a hypothesis test, using the P-value approach to hypothesis testing.
12. Perform a hypothesis test for a population mean when the population standard deviation is unknown.

Credit

Some of these slides have been adapted/modified in part/whole from the slides of the following textbooks.

- Weiss, Neil A., Introductory Statistics, 8th Edition
- Bock, David E., Stats: Data and Models, 3rd Edition