

Introduction to Statistics

Nels Grevstad

Metropolitan State University of Denver
ngrevsta@msudenver.edu

October 13, 2019

Nels Grevstad

Topics

- 1 Introduction to Hypothesis Testing
- 2 Hypothesis Test for μ when σ is Known

Nels Grevstad

Objectives

Objectives:

- Distinguish between the null and alternative hypotheses.
- State the role of the level of significance in hypothesis testing, and describe how the choice of a significance level can affect the conclusion of a hypothesis test.
- Interpret the p-value of a hypothesis test.
- Explain what's meant by a "statistically significant" study result.
- Carry out a one-mean z test for a population mean μ when the population standard deviation σ is known.

Nels Grevstad

Introduction to Hypothesis Testing (9.1, 9.3)

Null and Alternative Hypotheses

- A statistical **hypothesis** is a **claim** about the (unknown) value of a population parameter (e.g. μ).
- A **hypothesis test** is a statistical procedure for **deciding** between **two** competing **hypotheses**.

Nels Grevstad

Notes

Notes

Notes

Notes

- The two **hypotheses** play different roles:
 - The **null hypothesis** (H_0) is the hypothesis we seek to **discredit**, but to which we give the **benefit of the doubt**.
 - The **alternative hypothesis** (H_a) is the hypothesis we seek to **substantiate**.

- The conclusion of any hypothesis test will be to either **reject** or **fail to reject** H_0 .
- The decision will be based on a **test statistic**, its associated **p-value**, and a **decision criterion** involving a **level of significance**.

Steps for Carrying Out a Hypothesis Test:

1. State H_0 and H_a .
2. Choose a level of significance α .
3. Calculate the test statistic from the data.
4. Determine the p-value.
5. State the conclusion using the decision criterion (two slides ahead).
6. Interpret the result of the hypothesis test.

- The **p-value** is a **probability** that answers the question:
"If H_0 was true, what's the chance we'd get a test statistic value that's as contradictory to H_0 (and consistent with H_a) as the one we got?"

Smaller p-values indicate **stronger evidence against H_0** in favor of H_a .
- The **level of significance**, denoted α , is a threshold used in the **decision criterion**, which states (see next slide):

Decision Criterion:

Reject H_0 if p-value $< \alpha$

Fail to Reject H_0 if p-value $\geq \alpha$

where α is the value chosen for the **level of significance**.

- The most common choices for α are **0.01**, **0.05**, and **0.10**.
- Using a **smaller** value for α means we require **stronger evidence** against H_0 before we're willing to reject H_0 .

- When we reject H_0 , we say the result is ***statistically significant***.

A ***statistically significant*** result is one that **can't be explained by mere chance** (sampling error).

Hypothesis Test for μ when σ is Known (9.3, 9.4)

Introduction to the One-Mean Z Test

- The ***one-mean z test*** for a population mean μ is used when we have a **random sample** from a **population** whose **mean** and **standard deviation** are μ and σ , and

σ is ***known***
 μ is ***unknown***.

- We'll see how to use the sample to decide if μ is different from some **claimed value μ_0** .

- We'll be seeking to **discredit** the claim that μ is equal to μ_0 , so the **null hypothesis** is:

Null Hypothesis:

$$H_0 : \mu = \mu_0$$

Notes

Notes

Notes

Notes

- The **alternative hypothesis** will depend on what we're trying to **substantiate**:

Alternative Hypothesis: The alternative hypothesis will be one of

1. $H_a : \mu > \mu_0$ (**one-sided, upper-tailed test**)
2. $H_a : \mu < \mu_0$ (**one-sided, lower-tailed test**)
3. $H_a : \mu \neq \mu_0$ (**two-sided, two-tailed test**)

depending on what we're trying to substantiate in our study.

Exercise

Hemoglobin is a protein in red blood cells that carries oxygen from the lungs to body tissues. People with less than **12** grams of hemoglobin per deciliter of blood (g/dl) are said to be anemic.

A health official in Jordan suspects that the mean μ for all children in that country is **less than 12**. To test his claim, he measures the hemoglobin a random sample of $n = 50$ children.

The **null hypothesis** is

$$H_0 : \mu = 12$$

Which of the following three **alternative hypotheses** would he test?

- $H_a : \mu > 12$
- $H_a : \mu < 12$
- $H_a : \mu \neq 12$

Would he be performing a **lower-tailed, upper-tailed, or two-tailed test**?

Exercise

The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures students' study habits and attitudes toward school. The mean SSHA score for all college students is **115**.

A teacher suspects that the true mean μ for *older* students is **higher than 115**. To test her claim, she gives the SSHA to a random sample of $n = 25$ students who are at least 30 years old.

The **null hypothesis** is

$$H_0 : \mu = 115$$

Notes

Notes

Notes

Notes

Which of the following three **alternative hypotheses** would she test?

- $H_a : \mu > 115$
- $H_a : \mu < 115$
- $H_a : \mu \neq 115$

Would she be performing a **lower-tailed, upper-tailed, or two-tailed test**?

Exercise

The diameter of a spindle in a small motor is supposed to be 5 mm. If the spindle is either too small or too large, the motor will not work properly.

The manufacturer measures the diameters in a random sample of $n = 10$ spindles to determine whether the true mean diameter μ is **any different** from the target value 5 mm.

The **null hypothesis** is

$$H_0 : \mu = 5$$

Which of the following three **alternative hypotheses** would they test?

- $H_a : \mu > 5$
- $H_a : \mu < 5$
- $H_a : \mu \neq 5$

Would they be performing a **lower-tailed, upper-tailed, or two-tailed test**?

- The **test statistic** for the **one-mean z test for μ** is

One-Mean Z Test Statistic:

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

When

$$H_0 : \mu = \mu_0$$

is true, the **sampling distribution** of the test statistic z is a **standard normal** distribution.

(The **one-mean z test** is valid if either the sample is from a **normal** population or the sample size n is **large** ($n \geq 30$).)

Notes

Notes

Notes

Notes

- z measures how many **standard errors** \bar{x} is **away from** the claimed value μ_0 .
- \bar{x} is an **estimator** of the true (unknown) population mean μ , so we'd expect $\bar{x} \approx \mu$, and thus ...
 1. If $\mu = \mu_0$, we'd expect $\bar{x} \approx \mu_0$ and z to be approximately **zero**.
 2. If $\mu > \mu_0$, we'd expect $\bar{x} > \mu_0$ and z to be **positive**.
 3. If $\mu < \mu_0$, we'd expect $\bar{x} < \mu_0$ and z to be **negative**.

1. Values of z **close to zero** provide almost **no evidence against the null hypothesis**
 $H_0 : \mu = \mu_0$.

2. **Positive** values of z provide **evidence against the null hypothesis in favor of**
 $H_a : \mu > \mu_0$.
3. **Negative** values of z provide **evidence against the null hypothesis in favor of**
 $H_a : \mu < \mu_0$.
4. **Positive and negative** values of z provide **evidence against the null hypothesis in favor of**
 $H_a : \mu \neq \mu_0$.

- The ***p-value*** is the probability that just by chance (under H_0) we'd get a test statistic value as far from zero, in the direction predicted by H_a , as the observed value.

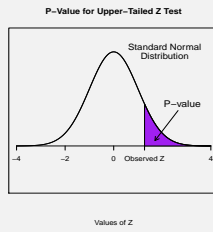
Notes

Notes

Notes

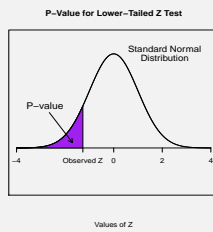
Notes

1. **P-value** = Area to the **right** of the observed z if the alternative hypothesis is $H_a : \mu > \mu_0$.



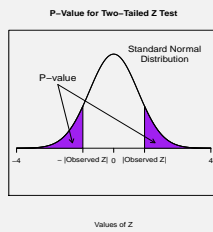
Nels Grevstad

2. **P-value** = Area to the **left** of the observed z if the alternative hypothesis is $H_a : \mu < \mu_0$.



Nels Grevstad

3. **P-value** = Area to the **left** of $-|z|$ and **right** of $|z|$ if the alternative hypothesis is $H_a : \mu \neq \mu_0$.



Nels Grevstad

Example

Recall that a health official in Jordan suspects that the mean hemoglobin level μ for all children in that country is **less than 12**. To test his claim, he measures the hemoglobin a random sample of $n = 50$ children.

He'll test the **hypotheses**

$$H_0 : \mu = 12$$

$$H_a : \mu < 12$$

Nels Grevstad

Notes

Notes

Notes

Notes

In the **sample**, the mean hemoglobin level is

$$\bar{x} = 11.7.$$

Suppose the population standard deviation is **known** to be $\sigma = 2$.

Then the observed **test statistic** is

$$\begin{aligned} z &= \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \\ &= \frac{11.7 - 12}{2/\sqrt{50}} \\ &= -1.06. \end{aligned}$$

Notes

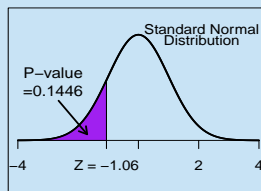
Thus the **sample mean** hemoglobin level, $\bar{x} = 11.7$, is **1.06 standard errors below 12**.

The **p-value** is the **probability** that we'd get this result by chance if the **population mean μ** was **12**.

From the **left tail** of the **standard normal** distribution, the **p-value is 0.1446** (see the next slide).

Notes

P-Value for Lower-Tailed Z Test



Values of Z

Notes

Thus we'd get a result like the one we got **14.46%** of the time **even if the population mean μ** was **12**.

Using a **level of significance $\alpha = 0.05$** , the **decision criterion** is

- Reject H_0 if p-value < 0.05 .
- Fail to reject H_0 if p-value ≥ 0.05 .

Because $0.1446 \geq 0.05$, we **fail to reject H_0** .

There's **no statistically significant evidence** that the population mean hemoglobin level μ is less than 12.

The result he got (by taking a random sample) can be explained by chance variation (sampling error).

Notes

Exercise

A concern of employers is time spent surfing the Internet and emailing friends during work hours.

An article in the *San Luis Obispo Tribune* (Aug. 3, 2016) ran under the headline "Who Goofs Off 2 Hours a Day? Most Workers, Survey Says".

The CEO of a company **wants to decide** whether the **true (unknown) population mean** amount of wasted time per workday for her employees is **less than** the reported **120** minutes.

A random sample of $n = 10$ employees was asked about daily wasted time at work.

(They were guaranteed anonymity to obtain truthful answers!)

Here are the data:

108 110 117 122 111 124 113 107 105 128

The sample mean is

$$\bar{x} = 114.5.$$

Suppose we know that in the company's **employee population**, wasted time follows a **normal** distribution with **known standard deviation** $\sigma = 9.0$ (but **unknown mean** μ).

Do the data provide **statistically significant** evidence, at the $\alpha = 0.05$ level, that the population mean wasted time μ for this company is **less than 120** minutes?

- a) State the **hypotheses** that should be tested.
- b) Calculate the **test statistic**.
- c) Determine the **p-value**.

Hint: You should get **-1.93**.

Hint: You should get **0.0268**.

- d) State the **conclusion** using a **level of significance** $\alpha = 0.05$, in which case the **decision criterion** is

Reject H_0 if p-value < 0.05 .

Fail to reject H_0 if p-value ≥ 0.05 .

- e) **Interpret** the result: Is there **statistically significant** evidence that the population mean wasted time μ for this company is **less than 120** minutes?
- f) If instead we had used a **significance level** $\alpha = 0.01$, would the conclusion have been different? Explain

Notes

Notes

Notes

Notes
