7 One-Sample Hypothesis Tests

MTH 3240 Environmental Statistics

Spring 2020

Objectives

Objectives:

- Explain the meanings of the terms hypothesis, test statistic, level of significance, p-value, statistical significance.
- Carry out a one-sample t test for a population mean using the rejection region and p-value approaches.

Introduction

- A statistical ***hypothesis*** is a claim about the value(s) of one or more population parameters.

Introduction

- A statistical ***hypothesis*** is a claim about the value(s) of one or more population parameters.
- A ***hypothesis test*** uses data to **decide** between two ***hypotheses***.

Introduction

- A statistical ***hypothesis*** is a claim about the value(s) of one or more population parameters.
- A ***hypothesis test*** uses data to **decide** between two ***hypotheses***.

Introduction

- A statistical ***hypothesis*** is a claim about the value(s) of one or more population parameters.
- A ***hypothesis test*** uses data to **decide** between two ***hypotheses***.
- The ***alternative hypothesis***, H_a, is the claim that's of primary interest.

Introduction

- A statistical ***hypothesis*** is a claim about the value(s) of one or more population parameters.
- A ***hypothesis test*** uses data to **decide** between two ***hypotheses***.
- The ***alternative hypothesis***, H_a, is the claim that's of primary interest.

  The ***null hypothesis***, H_0, is the claim that's not of interest.

- The **decision** will be either **reject H_0** or **fail to reject H_0**, depending on whether the data provide ample evidence to reject it.

- The **decision** will be either **reject** H_0 or **fail to reject** H_0, depending on whether the data provide ample evidence to reject it.
- The decision will be based on the value of a ***test statistic*** and a ***decision rule***.

- The **decision** will be either **reject** H_0 or **fail to reject** H_0, depending on whether the data provide ample evidence to reject it.
- The decision will be based on the value of a ***test statistic*** and a ***decision rule***.
- **"Failing to reject"** H_0 only means there was insufficient evidence to reject it, *not* that there was sufficient evidence to accept it.

Example

According to the U.S. Forest Service, to be classified as ***old growth***, the **mean tree diameter** μ in a Douglas-fir tree stand should be **at least 32** inches.

Example

According to the U.S. Forest Service, to be classified as ***old growth***, the **mean tree diameter** μ in a Douglas-fir tree stand should be **at least 32** inches.

A logging company **claims** that μ is **less than 32** inches. The government **claims** it's **at least 32**.

Before the government will allow logging of the forest, it requires **convincing evidence** that μ is **less than 32.**

Before the government will allow logging of the forest, it requires **convincing evidence** that μ is **less than 32.**

The company is allowed to take a **random sample** of trees and use the mean diameter $\bar{\mathbf{X}}$ to justify its claim **...**

Before the government will allow logging of the forest, it requires **convincing evidence** that μ is **less than 32.**

The company is allowed to take a **random sample** of trees and use the mean diameter $\bar{X}$ to justify its claim **...**

... but it must convince the government that $\bar{X}$ didn't fall **below 32** just as a result of **sampling variation** (chance).

The logging company, which is conducting the study, is **interested** in the claim that μ is **less than 32** inches.

The logging company, which is conducting the study, is **interested** in the claim that μ is **less than 32** inches.

This is the **alternative hypothesis**.

The logging company, which is conducting the study, is **interested** in the claim that μ is **less than 32** inches.

This is the **alternative hypothesis**.

The **null hypothesis**, therefore, is that μ is **32 inches or larger**.

The logging company, which is conducting the study, is **interested** in the claim that μ is **less than 32** inches.

This is the **alternative hypothesis**.

The **null hypothesis**, therefore, is that μ is **32 inches or larger**.

These are stated as

$$\begin{aligned} H_0 : \mu &\geq 32 \\ H_a : \mu &< 32 \end{aligned}$$

Evidence that μ is **less than 32** would come in the form of an $\bar{X}$ value **less than 32**.

Evidence that μ is **less than 32** would come in the form of an $\bar{X}$ value **less than 32**.

If that happens, a **hypothesis test** can be used to decide whether it's just due to **sampling variation**.

Rather than basing the decision on $\bar{X}$, it will be easier if we first ***standardize*** it, and use the ***test statistic***

$$t = \frac{\bar{X} - 32}{S_{\bar{X}}},$$

Rather than basing the decision on $\bar{X}$, it will be easier if we first ***standardize*** it, and use the ***test statistic***

$$t = \frac{\bar{X} - 32}{S_{\bar{X}}},$$

where $S_{\bar{X}}$ is the (estimated) **standard error** of $\bar{X}$,

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}.$$

- Often a set of hypotheses such as

$$\begin{aligned} H_0 : \mu &\geq 32 \\ H_a : \mu &< 32 \end{aligned}$$

is stated as

$$\begin{aligned} H_0 : \mu &= 32 \\ H_a : \mu &< 32 \end{aligned}$$

(the idea being that if the data provide enough evidence to reject $H_0 : \mu = 32$ in favor of $H_a : \mu < 32$, they provide *at least enough* evidence to reject $H_0 : \mu \geq 32$).

- There are two methods of forming the **decision rule**:
  - The ***rejection region approach***
  - The ***p-value approach***

- There are two methods of forming the **decision rule**:
    - The ***rejection region approach***
    - The ***p-value approach***

  In either case, $\boldsymbol{H_0}$ is **rejected** if the observed **test statistic** value is one that would **rarely occur just by sampling variation (chance)** if H_0 was true.

- There are two methods of forming the **decision rule**:
  - The ***rejection region approach***
  - The ***p-value approach***

  In either case, $\boldsymbol{H_0}$ is **rejected** if the observed **test statistic** value is one that would **rarely occur just by sampling variation (chance)** if H_0 was true.

  Both involve **comparing** the **observed test statistic** value to the **sampling distribution** the test statistic would follow **if $\boldsymbol{H_0}$ was true.**

- There are two methods of forming the **decision rule**:
  - The ***rejection region approach***
  - The ***p-value approach***

  In either case, $\boldsymbol{H_0}$ is **rejected** if the observed **test statistic** value is one that would **rarely occur just by sampling variation (chance)** if H_0 was true.

  Both involve **comparing** the **observed test statistic** value to the **sampling distribution** the test statistic would follow **if $\boldsymbol{H_0}$ was true.**

  We call this the ***null distribution***.

- In every hypothesis test, a ***level of significance*** α must be chosen.

- In every hypothesis test, a ***level of significance*** α must be chosen.

  It determines **how strong the evidence** against H_0 **needs to be** before we're willing to reject that hypothesis.

- In every hypothesis test, a ***level of significance*** α must be chosen.

  It determines **how strong the evidence** against H_0 **needs to be** before we're willing to reject that hypothesis.

  Smaller α values require stronger evidence.

- In every hypothesis test, a ***level of significance*** α must be chosen.

  It determines **how strong the evidence** against H_0 **needs to be** before we're willing to reject that hypothesis.

  Smaller α values require stronger evidence.

  The most common choices for α are **0.01**, **0.05**, and **0.10**.

Decision Rule for Rejection Region Approach:

Reject H_0 if the observed test statistic value falls in the rejection region.

Fail to reject H_0 if it doesn't fall in the rejection region.

- Using $\alpha = 0.05$, the ***rejection region*** consists of values in the **extreme 5%** of the **null distribution**, in the direction specified by H_a.

Decision Rule for Rejection Region Approach:

Reject H_0 if the observed test statistic value falls in the rejection region.

Fail to reject H_0 if it doesn't fall in the rejection region.

- Using $\alpha = 0.05$, the ***rejection region*** consists of values in the **extreme 5%** of the **null distribution**, in the direction specified by H_a.

  (More generally, it's the extreme **100α%** of the distribution).

Decision Rule for P-Value Approach:

Reject $\boldsymbol{H_0}$ if the p-value is less than 0.05 (or more generally, less than α).

Fail to reject $\boldsymbol{H_0}$ if the p-value isn't less than 0.05 (α).

- The ***p-value*** is the **probability** (under H_0) of getting a test statistic value **as extreme** (in the direction specified by H_a) as the **observed value**.

- **Smaller p-values** provide **more compelling evidence against H_0.**

- The **rejection region** and **p-value** approaches **always** lead to the **same conclusion**.

- The **rejection region** and **p-value** approaches **always** lead to the **same conclusion.**

  This is because (as we'll see later) the **test statistic** will fall in the **rejection region** if (and only if) the **p-value** is **less than α.**

- For a given choice of α, **if H_0 is rejected**, we say the result is ***statistically significant*** at the α **level.**

- For a given choice of α, **if H_0 is rejected**, we say the result is ***statistically significant*** at the α **level.**

  A **statistically significant** result is one for which **sampling variation (chance) can be ruled out** as being the sole explanation for the evidence supporting H_a.

Steps in Performing a Hypothesis Test

1. Identify the population parameter(s) of interest.
2. State the null and alternative hypotheses.
3. Choose an appropriate test procedure and check any assumptions (e.g. normality of the data).
4. Choose a level of significance α.
5. Compute the test statistic.
6. Find the p-value **or** determine if the test statistic falls in the rejection region.
7. State the conclusion (”reject” or ”fail to reject” H_0).

The One-Sample *t* Test

- The ***one-sample*** t ***test*** is a hypothesis test for an **(unknown) population mean** μ**.**

The One-Sample t Test

- The ***one-sample*** t ***test*** is a hypothesis test for an **(unknown) population mean** μ.
- The **null hypothesis** is that μ is equal to some ***claimed value*** μ_0.

Null Hypothesis:

$$H_0 : \mu = \mu_0.$$

- The **alternative hypothesis** is one of the following.

Alternative Hypothesis:

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

depending on what we're trying to verify using the data.

- The **alternative hypothesis** is one of the following.

Alternative Hypothesis:

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

depending on what we're trying to verify using the data.

One-Sample t Test Statistic:

$$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}},$$

where

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}.$$

One-Sample t Test Statistic:

$$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}},$$

where

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}.$$

- t indicates how many **standard errors** $\bar{X}$ is **away from** $\boldsymbol{\mu_0}$, and in what direction (positive or negative).

- $\bar{X}$ is an estimate of μ, so ...

- $\bar{X}$ is an estimate of μ, so ...
  - **If H_0 was true**, ...

- $\bar{X}$ is an estimate of μ, so ...
  - **If H_0 was true**, ...

    ... we'd expect $\bar{X}$ to be close μ_0.

- $\bar{X}$ is an estimate of μ, so ...
  - **If H_0 was true**, ...

    ... we'd expect $\bar{X}$ to be close μ_0.
  - But **if H_a was true**, ...

- $\bar{X}$ is an estimate of μ, so ...
  - **If H_0 was true**, ...

    ... we'd expect $\bar{X}$ to be close μ_0.
  - But **if H_a was true**, ...

    ... we'd expect $\bar{X}$ to differ from μ_0 in the direction specified by H_a

- $\bar{X}$ is an estimate of μ, so ...
  - **If H_0 was true**, ...

    ... we'd expect $\bar{X}$ to be close μ_0.
  - But **if H_a was true**, ...

    ... we'd expect $\bar{X}$ to differ from μ_0 in the direction specified by H_a
- Thus ...

- $\bar{X}$ is an estimate of μ, so ...
  - **If H_0 was true**, ...

    ... we'd expect $\bar{X}$ to be close μ_0.
  - But **if H_a was true**, ...

    ... we'd expect $\bar{X}$ to differ from μ_0 in the direction specified by H_a
- Thus ...
  1. t will be approximately **zero** (most likely) if H_0 is true.
  2. It will **differ from zero** (most likely) in the direction specified by H_a if $H_a$is true.

1. *Large positive* values of t provide evidence in favor of $H_a : \mu > \mu_0$.
2. *Large negative* values of t provide evidence in favor of $H_a : \mu < \mu_0$.
3. *Both large positive and large negative* values of t provide evidence in favor of $H_a : \mu \neq \mu_0$.

- Suppose we have a random sample from a population.

- Suppose we have a random sample from a population.

  If either

  1. The population is normal, or
  2. The sample size n is large,

- Suppose we have a random sample from a population.

  If either

  1. The population is normal, or
  2. The sample size n is large,

  the **null distribution** is as follows.

- Suppose we have a random sample from a population.

  If either

  1. The population is normal, or
  2. The sample size n is large,

  the **null distribution** is as follows.

Sampling Distribution of the Test Statistic Under H_0:
If t is the one-sample t test statistic, then when

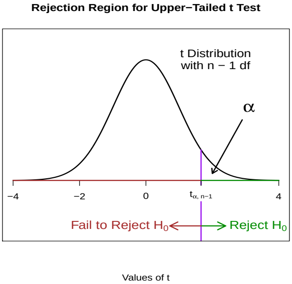
$$H_0 : \mu = \mu_0$$

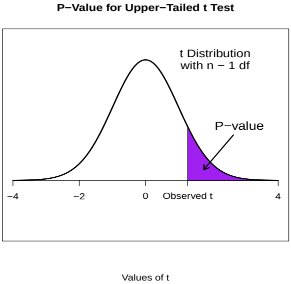
is true,

$$t \sim t(n-1).$$

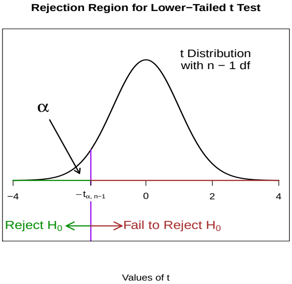
- **P-values** and **rejection regions** are obtained from the appropriate tail(s) of the $t(n-1)$ **distribution**, as shown on the next slides.

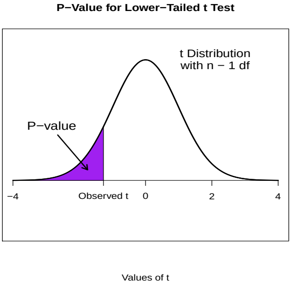
1. $H_a : \mu > \mu_0$ (Upper-Tailed Test)



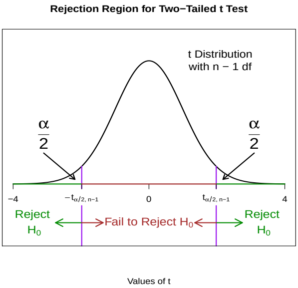


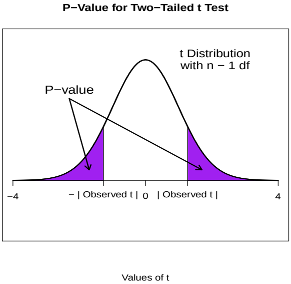
2. $H_a : \mu < \mu_0$ (Lower-Tailed Test)





3. $H_a : \mu \neq \mu_0$ (Two-Tailed Test)





One-Sample t Test for μ

Assumptions: The data $x_1, x_2, \ldots, x_n$ are a random sample from a population and either the population is *normal* or n is *large*.

Null hypothesis: $H_0 : \mu = \mu_0$.

Test statistic value: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$.

Decision rule: Reject H_0 if p-value $< \alpha$ or t is in rejection region.

One-Sample t Test for μ

Alternative hypothesis	P-value = area under t distribution with $n-1$ d.f.:	Rejection region = t values such that:*
$H_a : \mu > \mu_0$	to the right of t	$t > t_{\alpha,n-1}$
$H_a : \mu < \mu_0$	to the left of t	$t < -t_{\alpha,n-1}$
$H_a : \mu \neq \mu_0$	to the left of $-\|t\|$ and right of $\|t\|$	$t > t_{\alpha/2,n-1}$ or $t < -t_{\alpha/2,n-1}$

* $t_{\alpha,n-1}$ is the $100(1-\alpha)$th percentile of the t distribution with $n-1$ d.f.

Exercise

In the previous example, the logging company was interested in testing hypotheses that can be stated as

$$\begin{aligned} H_0 : \mu &= 32 \\ H_a : \mu &< 32 \end{aligned}$$

where μ is the true (unknown) **population mean** tree diameter.

Suppose that in a random sample of $n = 100$ trees, the **sample mean** and **standard deviation** of the diameters are

$$\begin{aligned} \bar{X} &= \mathbf{30.3} \\ S &= \mathbf{8.16}. \end{aligned}$$

Suppose that in a random sample of $n = 100$ trees, the **sample mean** and **standard deviation** of the diameters are

$$\begin{aligned} \bar{X} &= \mathbf{30.3} \\ S &= \mathbf{8.16}. \end{aligned}$$

Thus the (estimated) **standard error** of $\bar{X}$ is

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{8.16}{\sqrt{100}} = \mathbf{0.82}.$$

Suppose also a histogram indicates the data are from a **normal population**, so the **one-sample t test** is appropriate.

Suppose also a histogram indicates the data are from a **normal population**, so the **one-sample t test** is appropriate.

(It would be appropriate even if the normality assumption wasn't met because n is large.)

Suppose also a histogram indicates the data are from a **normal population**, so the **one-sample t test** is appropriate.

(It would be appropriate even if the normality assumption wasn't met because n is large.)

The **observed value** of the **test statistic** is

$$t \;=\; \frac{\bar{X} - \mu_0}{S_{\bar{X}}} \;=\; \frac{30.3 - 32}{0.82} \;=\; \mathbf{-2.07}.$$

a) For the **rejection region approach**, using level of significance $\alpha = \mathbf{0.05}$, the **decision rule** is

Reject H_0 if $t < -t_{0.05,99}$
Fail to reject H_0 if $t \geq -t_{0.05,99}$

a) For the **rejection region approach**, using level of significance $\alpha = \mathbf{0.05}$, the **decision rule** is

Reject H_0 if $t < -t_{0.05,99}$
Fail to reject H_0 if $t \geq -t_{0.05,99}$

Carry out the test using the **rejection region approach.**

b) For the **p-value approach**, using $\boldsymbol{\alpha = 0.05}$ again, the **decision rule** is

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

b) For the **p-value approach**, using $\boldsymbol{\alpha = 0.05}$ again, the **decision rule** is

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

Carry out the test again using the **p-value approach.**