t-tests

STAT 587 (Engineering)
Iowa State University

September 30, 2020

Statistical hypothesis testing

A hypothesis test consists of two hypotheses:

- null hypothesis (H_0) and
- an alternative hypothesis (H_A)

which make a claim about parameters in a model

Statistical hypothesis testing

A hypothesis test consists of two hypotheses:

- null hypothesis (H_0) and
- an alternative hypothesis (H_A)

which make a claim about parameters in a model and a decision to either

- reject the null hypothesis or
- fail to reject the null hypothesis.

t-tests

If $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$, then typical hypotheses about the mean are

$$H_0 : \mu = \mu_0$$

t-tests

If $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$, then typical hypotheses about the mean are

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

t-tests

If $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$, then typical hypotheses about the mean are

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

or

$$H_0 : \mu \leq \mu_0 \qquad \text{versus} \qquad H_A : \mu > \mu_0$$

t-tests

If $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$, then typical hypotheses about the mean are

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

or

$$H_0 : \mu \leq \mu_0 \qquad \text{versus} \qquad H_A : \mu > \mu_0$$

or

$$H_0 : \mu \geq \mu_0 \qquad \text{versus} \qquad H_A : \mu < \mu_0$$

t-tests

If $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$, then typical hypotheses about the mean are

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

or

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu > \mu_0$$

or

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu < \mu_0$$

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.

The as or more extreme region is determined by the alternative hypothesis.

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.
The as or more extreme region is determined by the alternative hypothesis.

$$H_A : \mu < \mu_0$$

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.
The as or more extreme region is determined by the alternative hypothesis.

$$H_A : \mu < \mu_0 \implies T \leq t$$

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.
The as or more extreme region is determined by the alternative hypothesis.

$$H_A : \mu < \mu_0 \implies T \leq t$$

or

$$H_A : \mu > \mu_0$$

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.
The as or more extreme region is determined by the alternative hypothesis.

$$H_A : \mu < \mu_0 \implies T \leq t$$

or

$$H_A : \mu > \mu_0 \implies T \geq t$$

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.
The as or more extreme region is determined by the alternative hypothesis.

$$H_A : \mu < \mu_0 \implies T \leq t$$

or

$$H_A : \mu > \mu_0 \implies T \geq t$$

or

$$H_A : \mu \neq \mu_0$$

t-statistic

Then

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$$

has a t_{n-1} distribution when H_0 is true.
The as or more extreme region is determined by the alternative hypothesis.

$$H_A : \mu < \mu_0 \implies T \leq t$$

or

$$H_A : \mu > \mu_0 \implies T \geq t$$

or

$$H_A : \mu \neq \mu_0 \implies |T| \geq |t|$$

where $T \sim t_{n-1}$.

Example data

Suppose we assume $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$ with $H_0 : \mu = 3$

Example data

Suppose we assume $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$ with $H_0 : \mu = 3$ and we observe

$$n = 6,\ \overline{y} = 6.3, \text{ and } s = 4.1.$$

Example data

Suppose we assume $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$ with $H_0 : \mu = 3$ and we observe

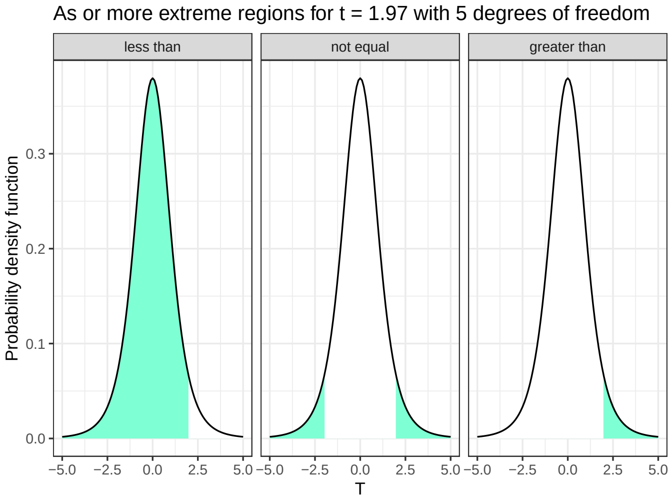
$$n = 6,\ \overline{y} = 6.3,\ \text{and } s = 4.1.$$

Then we can calculate

$$t = 1.97$$

which has a t_5 distribution if the null hypothesis is true.

as or more extreme regions



R Calculation

$$H_A : \mu < 3$$

R Calculation

$$H_A : \mu < 3$$

```
t.test(y, mu = mu0, alternative = "less")$p.value

[1] 0.9461974
```

R Calculation

$$H_A : \mu < 3$$

```
t.test(y, mu = mu0, alternative = "less")$p.value

[1] 0.9461974
```

$$H_A : \mu > 3$$

R Calculation

$$H_A : \mu < 3$$

```
t.test(y, mu = mu0, alternative = "less")$p.value

[1] 0.9461974
```

$$H_A : \mu > 3$$

```
t.test(y, mu = mu0, alternative = "greater")$p.value

[1] 0.05380256
```

R Calculation

$$H_A : \mu < 3$$

```r
t.test(y, mu = mu0, alternative = "less")$p.value
```

```
[1] 0.9461974
```

$$H_A : \mu > 3$$

```r
t.test(y, mu = mu0, alternative = "greater")$p.value
```

```
[1] 0.05380256
```

$$H_A : \mu \neq 3$$

R Calculation

$$H_A : \mu < 3$$

```
t.test(y, mu = mu0, alternative = "less")$p.value

[1] 0.9461974
```

$$H_A : \mu > 3$$

```
t.test(y, mu = mu0, alternative = "greater")$p.value

[1] 0.05380256
```

$$H_A : \mu \neq 3$$

```
t.test(y, mu = mu0, alternative = "two.sided")$p.value

[1] 0.1076051
```

Interpretation

The null hypothesis is a model.

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model.

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model. So, possibly

- the Y_i are not independent

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model. So, possibly

- the Y_i are not independent or
- they don't have a common σ^2

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model. So, possibly

- the Y_i are not independent or
- they don't have a common σ^2 or
- they aren't normally distributed

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model.
So, possibly

- the Y_i are not independent or
- they don't have a common σ^2 or
- they aren't normally distributed or
- $\mu \neq \mu_0$

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model. So, possibly

- the Y_i are not independent or
- they don't have a common σ^2 or
- they aren't normally distributed or
- $\mu \neq \mu_0$ or
- you got unlucky.

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model.
So, possibly

- the Y_i are not independent or
- they don't have a common σ^2 or
- they aren't normally distributed or
- $\mu \neq \mu_0$ or
- you got unlucky.

If you fail to reject H_0,

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y_i \stackrel{ind}{\sim} N(\mu_0, \sigma^2)$$

if we reject H_0, then we are saying the data are incompatible with this model. So, possibly

- the Y_i are not independent or
- they don't have a common σ^2 or
- they aren't normally distributed or
- $\mu \neq \mu_0$ or
- you got unlucky.

If you fail to reject H_0,then there is insufficient evidence to say that the data are incompatible with the null model.

Quality control example

An I-beam manufacturing facility has a design specification for I-beam thickness of 12 millimeters.

Quality control example

An I-beam manufacturing facility has a design specification for I-beam thickness of 12 millimeters. During manufacturing a random sample of I-beams are taken from the line and their thickness is measured.

Quality control example

An I-beam manufacturing facility has a design specification for I-beam thickness of 12 millimeters. During manufacturing a random sample of I-beams are taken from the line and their thickness is measured.

```
y
```

```
 [1] 12.04 11.98 11.97 12.12 11.90 12.05 12.14 12.13 12.18 12.23 12.03 12.03
```

Quality control example

An I-beam manufacturing facility has a design specification for I-beam thickness of 12 millimeters. During manufacturing a random sample of I-beams are taken from the line and their thickness is measured.

```
y

 [1] 12.04 11.98 11.97 12.12 11.90 12.05 12.14 12.13 12.18 12.23 12.03 12.03
```

```
t.test(y, mu = 12)


One Sample t-test

data:  y
t = 2.4213, df = 11, p-value = 0.03393
alternative hypothesis: true mean is not equal to 12
95 percent confidence interval:
 12.00607 12.12727
sample estimates:
mean of x
 12.06667
```

Quality control example

An I-beam manufacturing facility has a design specification for I-beam thickness of 12 millimeters. During manufacturing a random sample of I-beams are taken from the line and their thickness is measured.

```
y
```

```
 [1] 12.04 11.98 11.97 12.12 11.90 12.05 12.14 12.13 12.18 12.23 12.03 12.03
```

```
t.test(y, mu = 12)
```

```

One Sample t-test

data:  y
t = 2.4213, df = 11, p-value = 0.03393
alternative hypothesis: true mean is not equal to 12
95 percent confidence interval:
 12.00607 12.12727
sample estimates:
mean of x
 12.06667
```

The small p-value suggests the data may be incompatible with the model $Y_i \stackrel{ind}{\sim} N(12, \sigma^2)$.

Summary

- t-test, $Y_i \overset{ind}{\sim} N(\mu, \sigma^2)$:

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

Summary

- t-test, $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$:

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

- Use p-values to determine whether to
  - reject the null hypothesis or
  - fail to reject the null hypothesis.

Summary

- t-test, $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$:

$$H_0 : \mu = \mu_0 \qquad \text{versus} \qquad H_A : \mu \neq \mu_0$$

- Use p-values to determine whether to
  - reject the null hypothesis or
  - fail to reject the null hypothesis.
- More assessment is required to determine if other model assumptions hold.