Lecture 14: Hypothesis testing, continued

Announcements:

- Reading for this week's subject: pp. 61 - 79 in Vasilj.

Hypothesis testing: example

Hypothesis testing: example

We have a new variety of tomato.

Hypothesis testing: example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety?

Hypothesis testing: example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment.

Hypothesis testing: example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10.

Hypothesis testing: example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10. We know from previous experience that the mean growth rate of the old genotype is 33.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.
4. We decide the criterion for rejecting the null hypothesis.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.
4. We decide the criterion for rejecting the null hypothesis.
5. The test is one of: **two-tailed**, **right-tailed**, or **left-tailed**.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.
4. We decide the criterion for rejecting the null hypothesis.
5. The test is one of: **two-tailed**, **right-tailed**, or **left-tailed**.
6. We take a sample, and calculate our **test statistic** (Z or t for now)

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.
4. We decide the criterion for rejecting the null hypothesis.
5. The test is one of: **two-tailed**, **right-tailed**, or **left-tailed**.
6. We take a sample, and calculate our **test statistic** (Z or t for now)
7. We find if the observed test statistic is in the rejection region (**critical region** or **tail**) of the distribution.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.
4. We decide the criterion for rejecting the null hypothesis.
5. The test is one of: **two-tailed**, **right-tailed**, or **left-tailed**.
6. We take a sample, and calculate our **test statistic** (Z or t for now)
7. We find if the observed test statistic is in the rejection region (**critical region** or **tail**) of the distribution.
8. If the statistic is in the rejection region, we **reject** the null hypothesis and **accept** the alternative hypothesis.

Hypothesis testing: procedure

1. We ask a yes/no question about a **population**.
2. We answer the question yes, and answer the question no, using symbols for the population means.
3. We label one answer the **null hypothesis** and the other answer the **alternative hypothesis**.
4. We decide the criterion for rejecting the null hypothesis.
5. The test is one of: **two-tailed**, **right-tailed**, or **left-tailed**.
6. We take a sample, and calculate our **test statistic** (Z or t for now)
7. We find if the observed test statistic is in the rejection region (**critical region** or **tail**) of the distribution.
8. If the statistic is in the rejection region, we **reject** the null hypothesis and **accept** the alternative hypothesis.
9. If the statistic is not in the rejection region, we **retain** the null hypothesis, and do not accept the alternative hypothesis.

Rules

- If the alternative hypothesis contains the $\neq$ sign, the test is two-tailed.

Rules

- If the alternative hypothesis contains the $\neq$ sign, the test is two-tailed.
- If the alternative hypothesis contains the $<$ sign, the test is left-tailed.

Rules

- If the alternative hypothesis contains the $\neq$ sign, the test is two-tailed.
- If the alternative hypothesis contains the $<$ sign, the test is left-tailed.
- If the alternative hypothesis contains the $>$ sign, the test is right-tailed.

Hypothesis testing: example

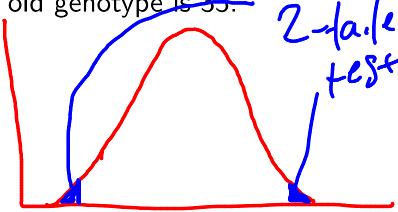
μ_n = resistance of new

μ_o = " " old

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10. We know from previous experience that the mean growth rate of the old genotype is 33.

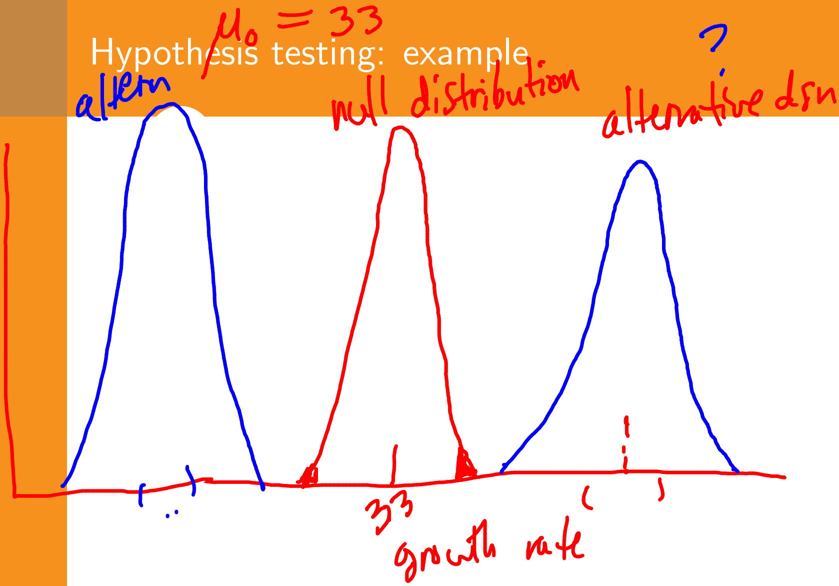
$H_0: \mu_n - \mu_o = 0$

$H_A: \mu_n - \mu_o \neq 0$



2-tailed test

Hypothesis testing: example



Hypothesis testing: example

μ_n new

μ_o old

We have a new variety of tomato. Does it have **lower** resistance to fusarium wilt than the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10. We know from previous experience that the mean growth rate of the old genotype is 33.

Left-tailed test

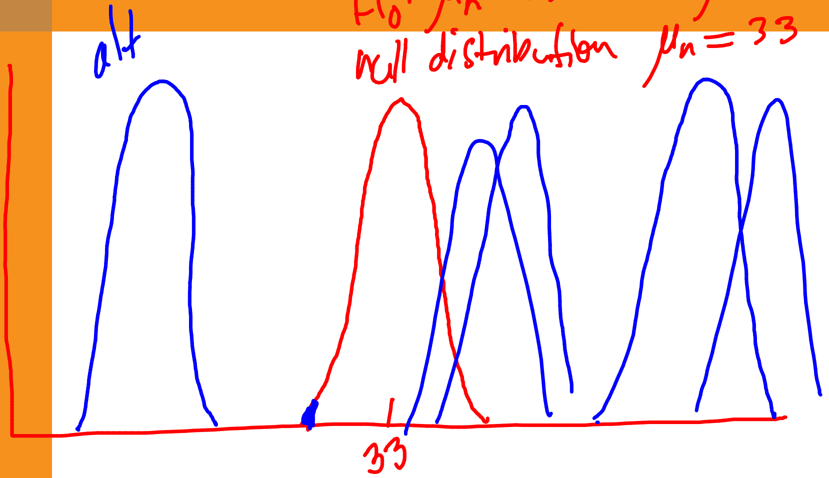
$H_A: \mu_n - \mu_o < 0$

$H_0: \mu_n - \mu_o \geq 0$

Hypothesis testing: example

$H_0: \mu_n - 33 \geq 0 \Rightarrow \mu_n \geq 33$

null distribution $\mu_n = 33$



Hypothesis testing: example

We have a new variety of tomato. Does it have **higher** resistance to fusarium wilt than the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10. We know from previous experience that the mean growth rate of the old genotype is 33.

Hypothesis testing: example

Some terminology

- If we reject the null hypothesis, then we conclude our data are **significant**, or **significantly different** from the null.

Some terminology

- If we reject the null hypothesis, then we conclude our data are **significant**, or **significantly different** from the null.
- The **probability** of the test is the probability of getting the observed test result, or more extreme, than expected by chance, if the null hypothesis were true.

1 - pnorm(z) = P

Some terminology

- If we reject the null hypothesis, then we conclude our data are **significant**, or **significantly different** from the null.
- The **probability** of the test is the probability of getting the observed test result, or more extreme, than expected by chance, if the null hypothesis were true.
- In a two-tailed test, the probability is twice the tail probability. In a one-tailed test, the probability is the one-tail probability.

Some terminology

- If we reject the null hypothesis, then we conclude our data are **significant**, or **significantly different** from the null.
- The **probability** of the test is the probability of getting the observed test result, or more extreme, than expected by chance, if the null hypothesis were true.
- In a two-tailed test, the probability is twice the tail probability. In a one-tailed test, the probability is the one-tail probability.
- This probability is also called the **significance** or the **significance level** of the test.

Some more terminology

- The least significant difference:

Some more terminology

- The least significant difference:

Some more terminology

LSD

- The least significant difference:

$$t = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

Some more terminology

- The least significant difference:

$$t = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$t_{crit} = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

Some more terminology

- The least significant difference:

$$t = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$t_{crit} = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$\text{LSD} =$$

Some more terminology

- The least significant difference:

$$t = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$t_{crit} = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$\text{LSD} = t_{crit}$$

Some more terminology

- The least significant difference:

$$t = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$t_{crit} = \frac{\text{observed} - \text{expected}}{s_{\text{observed}}}$$

$$\text{LSD} = t_{crit} s_{\text{observed}}$$

Hypothesis testing: example

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}} = \sqrt{\frac{S_x^2}{n}}$$

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10. We know from previous experience that the mean growth rate of the old genotype is 33.

$$t = \frac{\bar{x} - \mu_0}{S_{\bar{x}}} = \frac{35-33}{\frac{10}{10}} = 2.$$

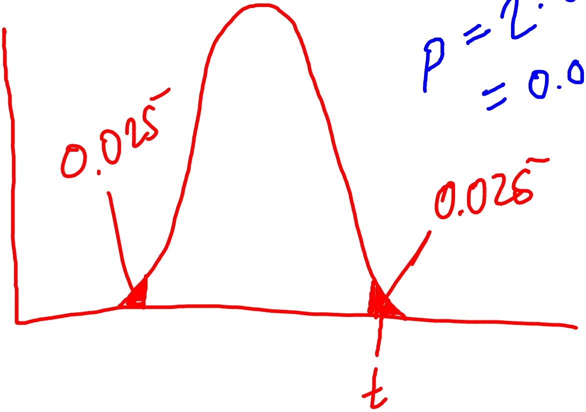
Hypothesis testing: example

$p = P(\text{our test result} \mid H_0 \text{ is true})$

$t = 2 > 1.98 - t_{crit}$

$\therefore$ reject H_0

$p = 2 \cdot 0.0241 = 0.0482$



Hypothesis testing: example

Hypothesis testing: example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants and 100 old plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10 for the new plants, and mean 33 and standard deviation 8 for the old plants.

Hypothesis testing: example

Unequal variance t-test

Calculate separate variances for the two samples:

Unequal variance t-test

Calculate separate variances for the two samples:

$$t =$$

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{}$$

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

What is the degrees of freedom?

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

What is the degrees of freedom? Either use the smallest of $n_1 - 1$ and $n_2 - 1$,

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

What is the degrees of freedom? Either use the smallest of $n_1 - 1$ and $n_2 - 1$, or use R to calculate the following degrees of freedom:

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

What is the degrees of freedom? Either use the smallest of $n_1 - 1$ and $n_2 - 1$, or use R to calculate the following degrees of freedom:

$$\text{d.f.} =$$

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

What is the degrees of freedom? Either use the smallest of $n_1 - 1$ and $n_2 - 1$, or use R to calculate the following degrees of freedom:

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{}$$

Unequal variance t-test

Calculate separate variances for the two samples:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

What is the degrees of freedom? Either use the smallest of $n_1 - 1$ and $n_2 - 1$, or use R to calculate the following degrees of freedom:

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}.$$

Hypothesis testing: example

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} \; =$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance } = \frac{(n_1 - 1)s_i^2 +}{}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance } = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance } = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$s =$

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{(n_1 - 1)s_i^2 +}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{}}$$

Equal variance *t*-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$t =$$

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$t = \frac{\overline{x_1} - \overline{x_2}}{}$$

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s^2/n_1 + s^2/n_2}} =$$

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s^2/n_1 + s^2/n_2}} = \frac{\overline{x_1} - \overline{x_2}}{}$$

Equal variance t-test: use the pooled variance to calculate the sample standard deviation

$$\text{pooled variance} = \frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s = \sqrt{\text{pooled variance}} = \sqrt{\frac{(n_1 - 1)s_i^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s^2/n_1 + s^2/n_2}} = \frac{\overline{x_1} - \overline{x_2}}{s\sqrt{1/n_1 + 1/n_2}}$$

Hypothesis testing: example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants and 100 old plants to fusarium, and found that the mean growth rate was 35 with standard deviation 10 for the new plants, and mean 33 and standard deviation 8 for the old plants.

Hypothesis testing: example

Hypothesis testing: one-sample test of proportion example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants to fusarium, and found that 72 are resistant. We know in advance that 65% of the old plants are resistant.

Hypothesis testing: example

Hypothesis testing: two-sample test of proportions example

We have a new variety of tomato. Does it have **the same** resistance to fusarium wilt as the old variety? To answer this question, we do an experiment. We expose 100 new plants and 100 old plants to fusarium, and found that 72 new are resistant, and 65 old are resistant.

Hypothesis testing: example

Two sample test of proportions

Given x_1 successes in a sample of n_1, and x_2 successes in a sample of n_3. Therefore $\hat{p}_1 = x_1/n_1$, and $\hat{p}_2 = x_2/n_2$.
We define $\overline{p} = \frac{x_1+x_2}{n_1+n_2}$, and $\overline{q} = 1 - \overline{p}$.
Then our test statistic is:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\overline{pq}}{n_1} + \frac{\overline{pq}}{n_2}}}$$