t-tests

ENTMLGY 6707 Entomological Techniques and Data Analysis

Learning objectives

1) Identify the information required to conduct a *t*-test

2) Distinguish between types of *t*-tests

3) Interpret the outcome of *t*-tests

Hypothesis Testing

State the hypotheses:
Null Hypothesis (H_0)
Alternative Hypothesis (H_a)

↓

Data collection/sampling

↓

Calculate the test statistic

↓

Calculate the P-value

↓

Make a decision

t-tests

Used to determine if the means of two groups have statistically clear differences OR used to determine if a mean differs from a specified value (e.g., population mean or 0).

Null Hypothesis (H_0): The mean is equal to a specified value $H_0: \mu = c$

Alternative Hypothesis (H_a): The mean is different from the specified value $H_a: \mu \neq c$

Null Hypothesis (H_0): The means of the two groups are equal $H_0: \mu_1 = \mu_2$

Alternative Hypothesis (H_a): The means of the two groups are not equal $H_a: \mu_1 \neq \mu_2$

Exercise

Discuss the null and alternative hypotheses for your research/experiment?

t-tests

Used to determine if the means of two groups have statistically clear differences OR used to determine if a mean differs from a specified value (e.g., population mean or 0).

Assumptions:

1.
2.
3.
4.

Activity: Based on what you have learned so far about experimental design, sampling, and distributions, list assumptions of the standard t-test

t-tests

Used to determine if the means of two groups have statistically clear differences OR used to determine if a mean differs from a specified value (e.g., population mean or 0).

<u>Assumptions</u>:

1. Response (aka dependent) variable (the one we are comparing between groups) is continuous
2.
3.
4.

Activity: Based on what you have learned so far about experimental design, sampling, and distributions, list assumptions of the standard t-test

t-tests

Used to determine if the means of two groups have statistically clear differences OR used to determine if a mean differs from a specified value (e.g., population mean or 0).

Assumptions:

1. Response (aka dependent) variable (the one we are comparing between groups) is continuous
2. Observations are independent (typically meaning they comprise a random sample)
3.
4.

Activity: Based on what you have learned so far about experimental design, sampling, and distributions, list assumptions of the standard t-test

t-tests

Used to determine if the means of two groups have statistically clear differences OR used to determine if a mean differs from a specified value (e.g., population mean or 0).

<u>Assumptions</u>:

1. Response (aka dependent) variable (the one we are comparing between groups) is continuous
2. Observations are independent (typically meaning they comprise a random sample)
3. Response variable is normally distributed
4.

Activity: Based on what you have learned so far about experimental design, sampling, and distributions, list assumptions of the standard t-test

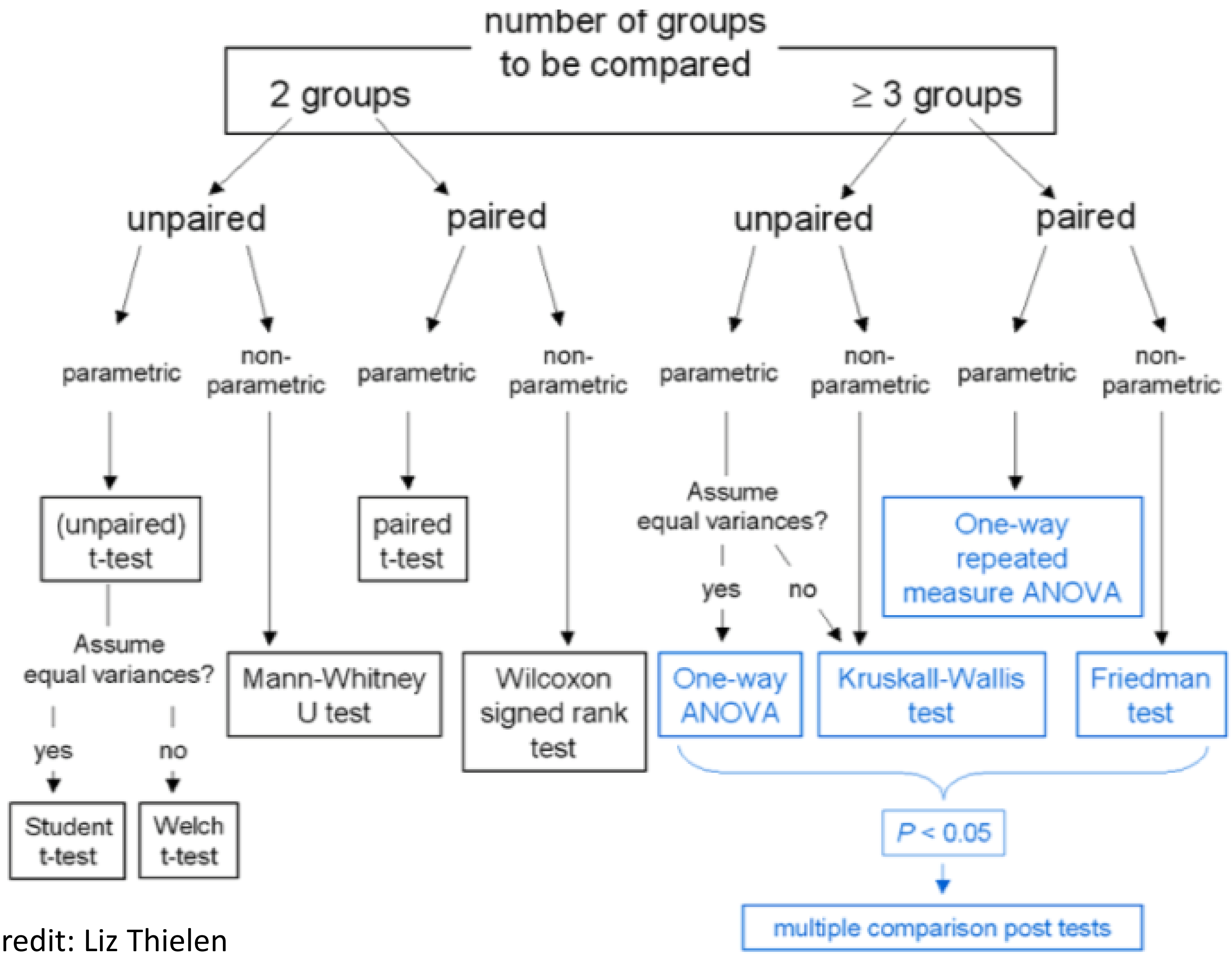
t-tests

Used to determine if the means of two groups have statistically clear differences OR used to determine if a mean differs from a specified value (e.g., population mean or 0).

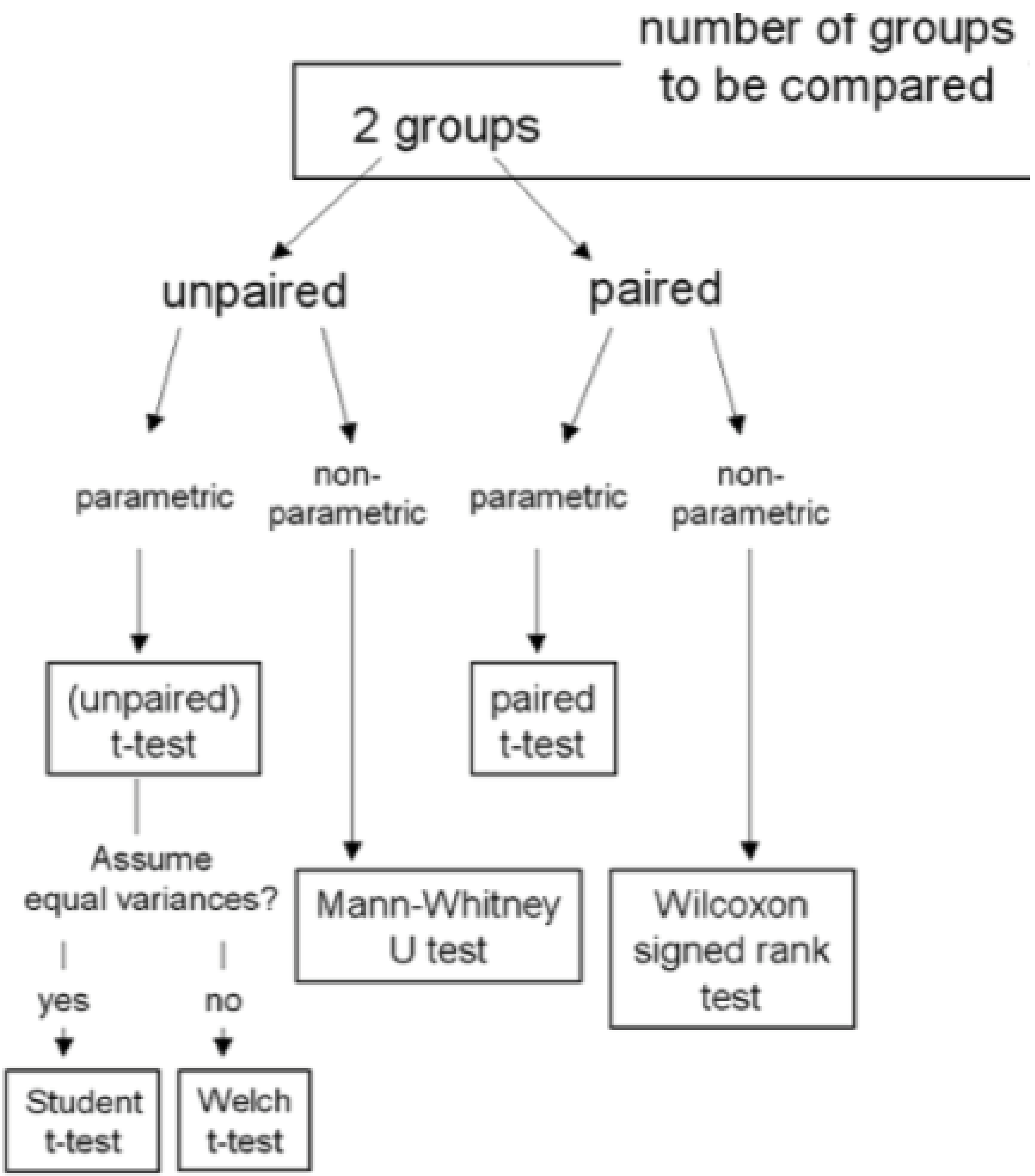
Assumptions:

1. Response (aka dependent) variable (the one we are comparing between groups) is continuous
2. Observations are independent (typically meaning they comprise a random sample)
3. Response variable is normally distributed
4. Variances of the response variable are equal across groups (= homogeneity of variances)

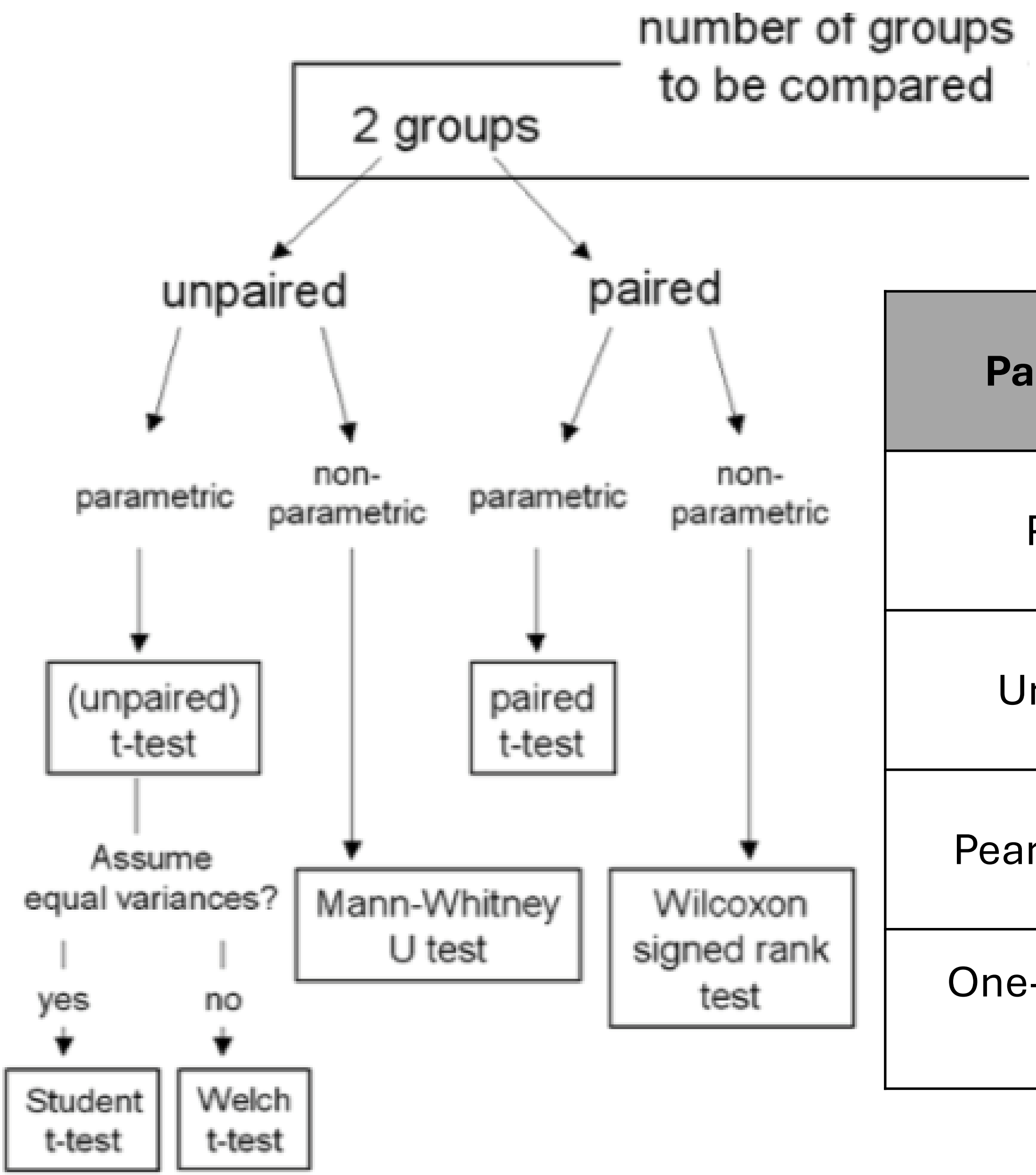
Activity: Based on what you have learned so far about experimental design, sampling, and distributions, list assumptions of the standard t-test



Credit: Liz Thielen



Credit: Liz Thielen

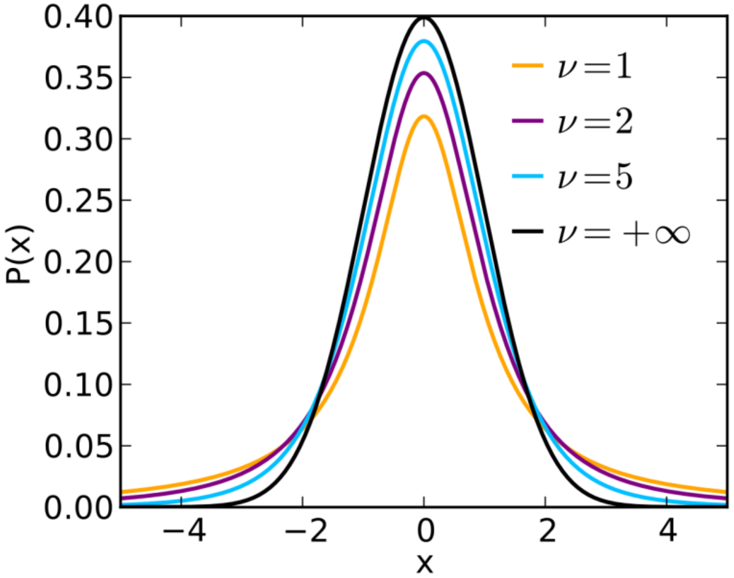


Credit: Liz Thielen

Parametric Test	Non-Parametric Test
Paired t-test	Wilcoxon signed–rank test
Unpaired t-test	Mann-Whitney U test
Pearson correlation	Spearman correlation
One-way Analysis of Variance	Kruskal-Wallis test

William Sealy Gosset

- 1876 – 1937
- Head Experimental Brewer of Guinness
- Developed Student's t-distribution
- Argued – in the literature and probably at the pub – with Ronald Fisher
- Had awesome mustache

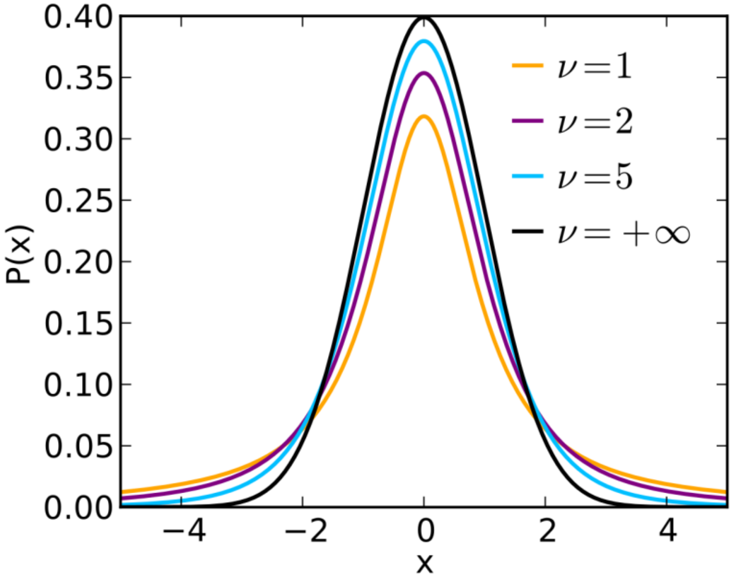


$$f(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \cdot \frac{1}{(1+t^2/r)^{(r+1)/2}}$$

Note: *r* and *t* in the above equation are indicated by *v* and *x* on the figure

William Sealy Gosset

- 1876 – 1937
- Head Experimental Brewer of Guinness
- Developed Student's t-distribution
- Argued – in the literature and probably at the pub – with Ronald Fisher
- Had awesome mustache



$$f(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \cdot \frac{1}{(1+t^2/r)^{(r+1)/2}}$$

Note: *r* and *t* in the above equation are indicated by *v* and *x* on the figure

One sample t-test

(includes paired t-tests)

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

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

$$\mathrm{d.f.} = n - 1$$

Two sample t-test

These equations are for when your two groups have unequal sample sizes and variances

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{\Delta}}}$$

$$s_{\bar{\Delta}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

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

One sample t-test

(includes paired t-tests)

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

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

$$\mathrm{d.f.} = n - 1$$

Two sample t-test

These equations are for when your two groups have unequal sample sizes and variances

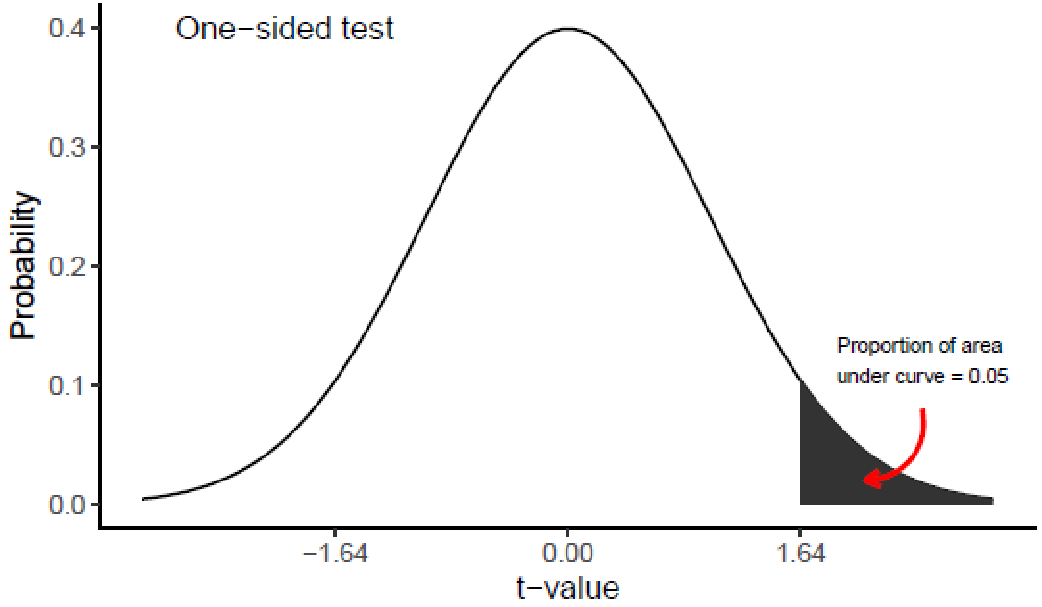
$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{\Delta}}}$$

$$s_{\bar{\Delta}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

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

$H_0: \mu \leq c$

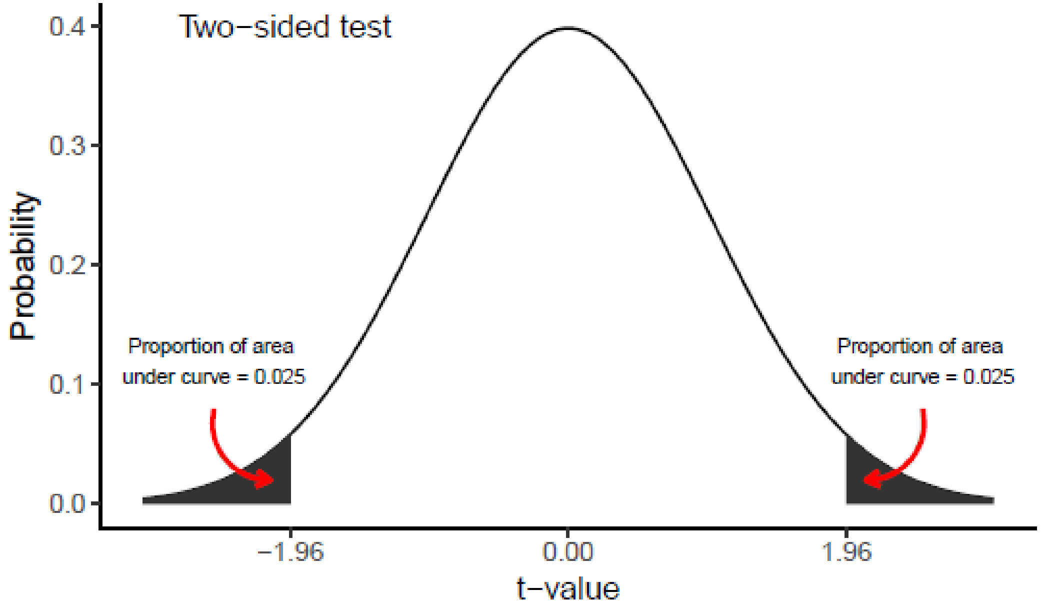
$H_a: \mu > c$



$p \leq 0.05$

$H_0: \mu = c$

$H_a: \mu \neq c$



One-sample t-test

A company claims that its candy weighs 20 grams. You randomly sample 31 candies, and the average weight is 21.5 grams with a sample standard deviation of 2 grams. At a significance level of 0.05, can you conclude that the average weight of the candy differs from the company's claim?

Null Hypothesis (H_0): The population mean candy weight is 20 grams ($\mu = 20$)

Alternative Hypothesis (H_1): The population mean candy weight is not 20 grams ($\mu \neq 20$)

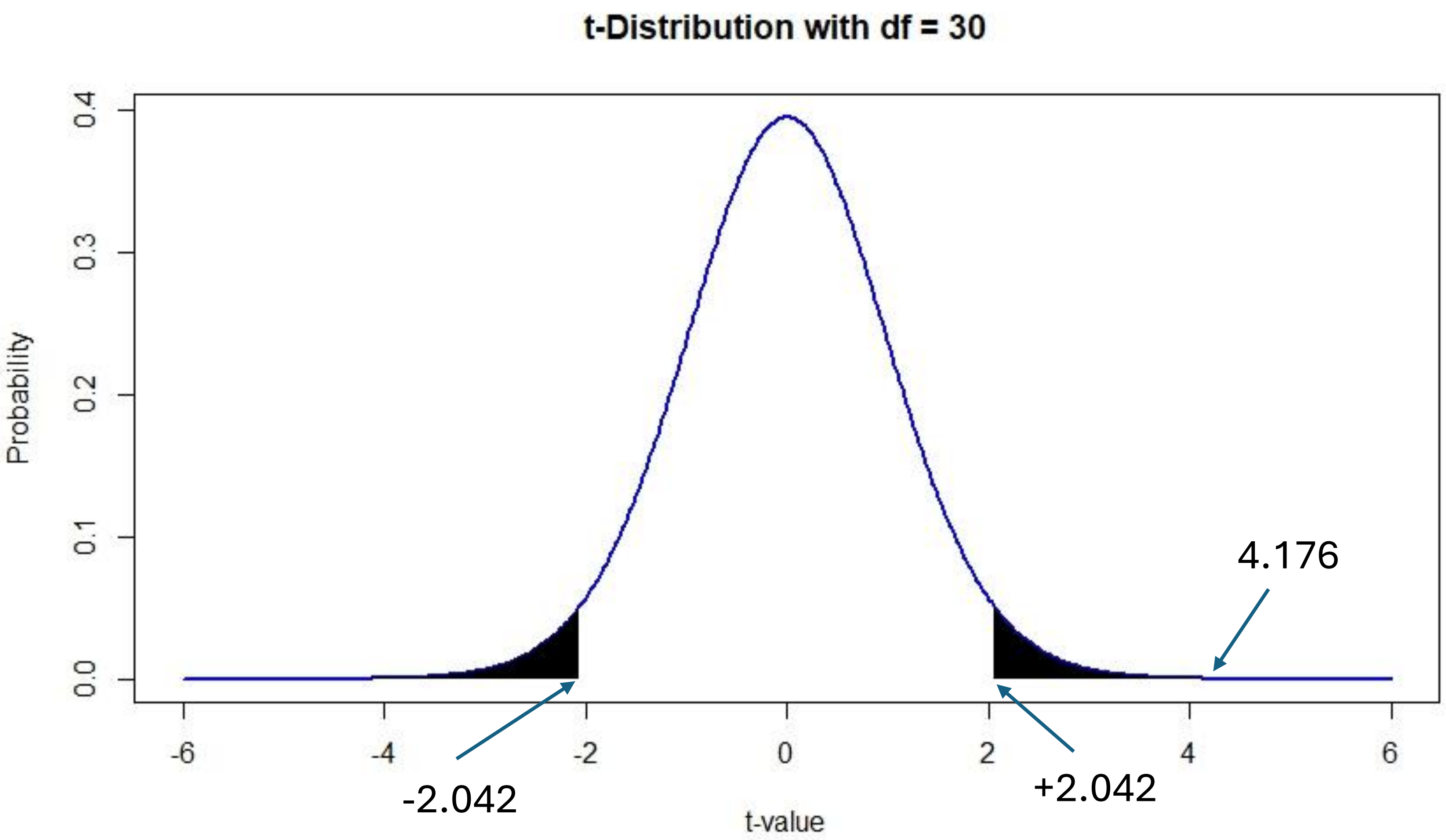
df = n – 1 = 31 – 1 = 30

Using a t-table for a two-tailed test with df = 30 and α = 0.05, the critical t-value is approximately ±2.042.

$t = (21.5 - 20) / (2 / \sqrt{31})$
$t \approx 4.176$

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

The calculated t-statistic (4.176) is greater than the critical t-value (2.042).



Since the t-statistic (4.176) falls in the rejection region (greater than 2.042), we reject the null hypothesis.

Conclusion: At the 0.05 significance level, there is enough evidence to suggest that the average weight of the candy is significantly different from the 20 grams claimed by the company.