STATISTICS II

Lisbon School of Economics & Management
Universidade de Lisboa

Bachelor's degrees in Economics, Finance and Management

2nd year/2nd Semester
2025/2026

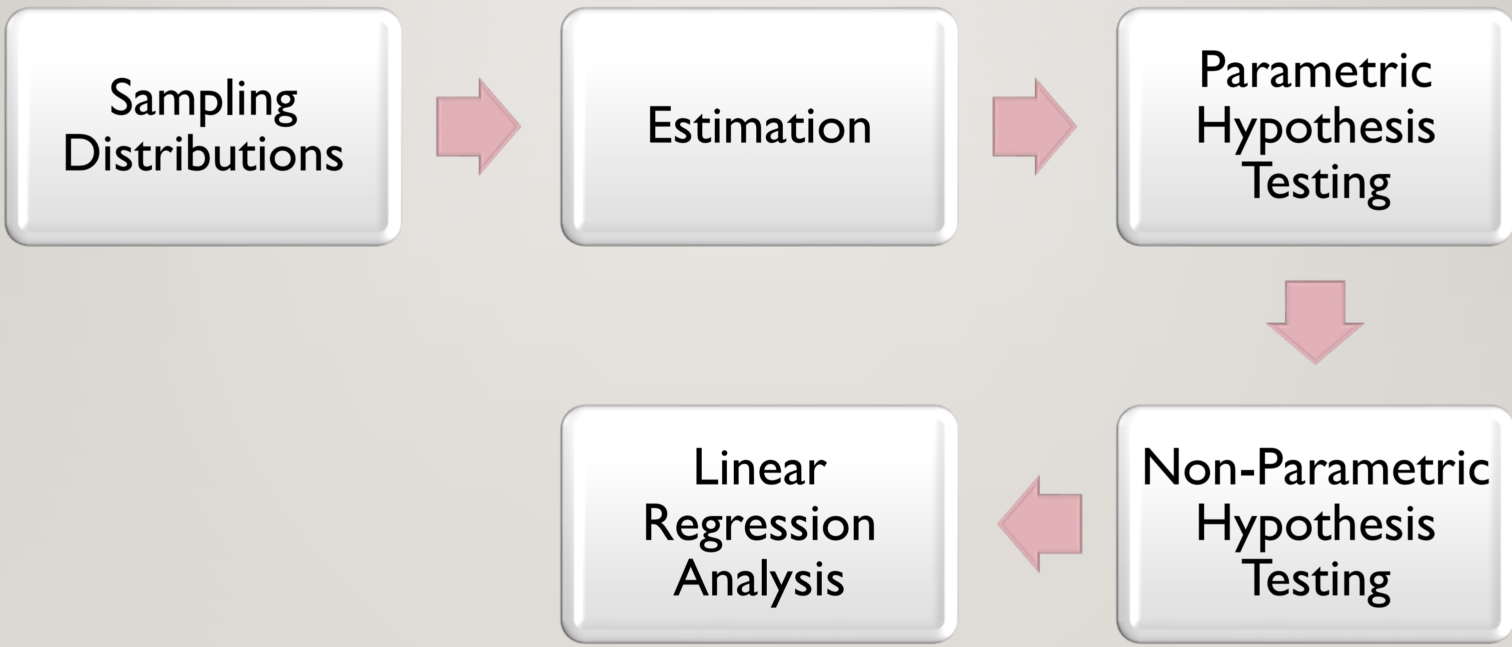
CONTACT

Professor: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

PROGRAM


Sampling Distributions
Estimation
Parametric Hypothesis Testing
Non-Parametric Hypothesis Testing
Linear Regression Analysis

HOMEWORK OF LECTURE 15: QUESTIONS AND SOLUTIONS

EXERCISE 9.33

9.33 Of a random sample of 199 auditors, 104 indicated some measure of agreement with this statement: *Cash flow is an important indication of profitability.* Test at the 10% significance level against a two-sided alternative the null hypothesis that one-half of the members of this population would agree with this statement. Also find and interpret the *p*-value of this test.

Newbold et al (2013)

EXERCISE 9.33: SOLUTION

Answer:

n×$\hat{p}$×(1- $\hat{p}$) = 199 × 0.5226 × (1-0.5226) = 49.65> 5

Note:
Since n×$\hat{p}$×(1- $\hat{p}$) > 5, the normal approximation is valid and the one-sample proportion test can be applied.

Given

- Sample size: $n = 199$
- Number agreeing: $x = 104$
- Sample proportion:

$$\hat{p} = \frac{104}{199} \approx 0.5226$$

Two-Tailed Test

- Null hypothesis: $H_0 : P = 0.5$
- Alternative hypothesis: $H_1 : P \neq 0.5$ (**two-sided**)
- Significance level: $\alpha = 0.10$

Step 1: Test statistic

For large samples, the **Z-test for a proportion** is:

$$Z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

Where $P_0 = 0.5$.

Standard error:

$$SE = \sqrt{\frac{0.5 \cdot 0.5}{199}} = \sqrt{\frac{0.25}{199}} \approx \sqrt{0.001256} \approx 0.0354$$

Test statistic:

$$Z_0 = \frac{0.5226 - 0.5}{0.0354} \approx \frac{0.0226}{0.0354} \approx 0.64$$

EXERCISE 9.33: SOLUTION

Answer:

Decision rule: Reject H_0 if the test statistic z_0 is in the rejection region (RR) or if p-value < α.

Step 2: Critical values

Two-sided test at $\alpha = 0.10$:

RR =] -∞; -1.645]U[1.645; +∞[

$$z_{1-\alpha/2} = Z_{0.95} = 1.645$$

Decision rule:

$$\text{Reject } H_0 \text{ if } |Z_0| > 1.645$$

Step 3: Decision

$z_0 = 0.64 \in RR$ ⟺ $|Z_0| = 0.64 < 1.645$

$\Rightarrow$ Do not reject H_0.

Step 4: p-value

Two-sided p-value:

P-value = 2 × P(T > 0.64) ~ 0.522

From standard normal table:

$$P(Z > 0.64) \approx 0.2611$$

$$p\text{-value} = 2 \cdot 0.2611 \approx 0.522$$

Conclusion

- p-value = 0.522 > 0.10 → do not reject H_0.
- There is **no significant evidence** at the 10% level that the proportion of auditors agreeing differs from 50%.

INFERENCE FOR A POPULATION PROPORTION

Note: The statistic used for a **confidence interval** and for a **hypothesis test** for a population proportion are **different**.

Confidence Interval for a Population Proportion

$$\frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

- Uses the **sample proportion** $\hat{p}$ to estimate the standard error.

Confidence Interval for a Population Proportion

A $100(1-\alpha)\%$ confidence interval for the population proportion p is:

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- $\hat{p}$: sample proportion
- n: sample size
- $z_{\alpha/2}$: critical value from the standard normal distribution

Hypothesis Test for a Population Proportion

The test statistic is:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- p_0: proportion under the null hypothesis

Key difference:
Confidence interval: standard error based on $\hat{p}$.
Hypothesis test: standard error based on p_0.

POPULATION PROPORTIONS: WHY ARE THE STATISTICS DIFFERENT FOR CONFIDENCE INTERVALS AND HYPOTHESIS TESTS?

Note: The difference comes from the **role played by the unknown population proportion** p in each procedure.

Confidence Interval

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- The true proportion p is **unknown**.
- The standard error is therefore **estimated** using the sample proportion $\hat{p}$.
- The goal is **estimation**, not decision-making.

Hypothesis Test

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- Under the null hypothesis, the population proportion is **fixed at** p_0.
- The standard error is computed using this **assumed value**.
- The goal is to **test a claim** about p.

Key Takeaway:
Confidence interval: estimates variability using $\widehat{p}$.
Hypothesis test: evaluates evidence assuming p = p_0.
The statistics differ because **estimation** and **hypothesis testing** answer different statistical questions.

LECTURE 16: TESTS OF THE VARIANCE OF A NORMAL DISTRIBUTION

TESTS OF THE VARIANCE OF A NORMAL DISTRIBUTION

- Goal: Test hypotheses about the population variance, σ^2 (e.g., $H_0 : \sigma^2 = \sigma_0^2$)
  - If the population is normally distributed,

$$Q = \frac{(n-1)s^2}{\sigma^2}$$

has a chi-square distribution with $(n-1)$ degrees of freedom

Two-Tailed Test

$H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 \neq \sigma_0^2$

Newbold et al (2013)

TESTS OF THE VARIANCE OF A NORMAL DISTRIBUTION

The test statistic for hypothesis tests about one population variance is

$$Q = \frac{(n-1)s^2}{\sigma_0^2}$$

Newbold et al (2013)

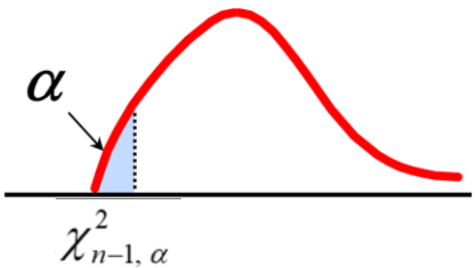
TESTS OF THE VARIANCE OF A NORMAL DISTRIBUTION: DECISION USING THE RR

Population variance

Left-Tailed Test

$H_0 : \sigma^2 \geq \sigma_0^2$

$H_1 : \sigma^2 < \sigma_0^2$



Reject H_0 if

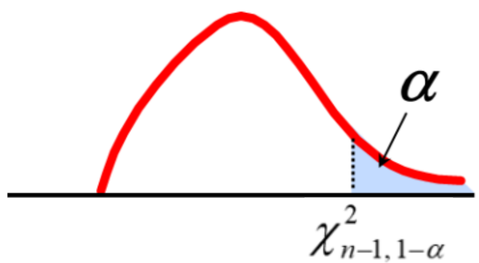
q $< \chi^2_{n-1,\,\alpha}$

RR =]-∞; $\chi^2_{n-1,\alpha}$]

Right-Tailed Test

$H_0 : \sigma^2 \leq \sigma_0^2$

$H_1 : \sigma^2 > \sigma_0^2$



Reject H_0 if

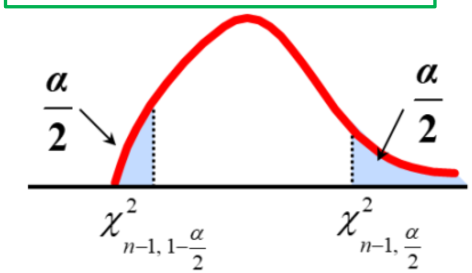
q $> \chi^2_{n-1,\,1-\alpha}$

RR = [$\chi^2_{n-1,1-\alpha}$; +∞[

Two-Tailed Test

$H_0 : \sigma^2 = \sigma_0^2$

$H_1 : \sigma^2 \neq \sigma_0^2$



Reject H_0 if

q $> \chi^2_{n-1,\,1-\frac{\alpha}{2}}$

or

q $< \chi^2_{n-1,\,\frac{\alpha}{2}}$

RR =] -∞; $\chi^2_{n-1,\alpha/2}$]U[$\chi^2_{n-1,1-\alpha/2}$; +∞[

t al (2013)

TESTS OF THE VARIANCE OF A NORMAL DISTRIBUTION: DECISION USING THE P-VALUE

Test Statistic

$$Q = \frac{(n-1)s^2}{\sigma_0^2}, \qquad Q \sim \chi^2_{n-1}$$

Left-Tailed Test

- Hypotheses:

$$H_0 : \sigma^2 \geq \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 < \sigma_0^2$$

- p-value:

$$p\text{-value} = P(Q \leq q_0)$$

Right-Tailed Test

- Hypotheses:

$$H_0 : \sigma^2 \leq \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 > \sigma_0^2$$

- p-value:

$$p\text{-value} = P(Q \geq q_0)$$

Two-Tailed Test

- Hypotheses:

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

- p-value:

$$p\text{-value} = 2 \times \min\left\{P(Q \leq q_0),\ P(Q \geq q_0)\right\}$$

Note:
Q denotes the **random test statistic**, while q_0 is its **observed value**.
For a two-tailed test, the p-value is twice the smaller tail probability.

EXERCISE 9.49

9.49 Plastic sheets produced by a machine are periodically monitored for possible fluctuations in thickness. If the true variance in thicknesses exceeds 2.25 square millimeters, there is cause for concern about product quality. Thickness measurements for a random sample of 10 sheets produced in a particular shift were taken, giving the following results (in millimeters):

226 226 232 227 225 228 225 228 229 230

a. Find the sample variance.
b. Test, at the 5% significance level, the null hypothesis that the population variance is at most 2.25.

Newbold et al (2013)

EXERCISE 9.49 A): SOLUTION

Answer:

Given

- Sample size: $n = 10$
- Sample measurements (in mm): 226, 226, 232, 227, 225, 228, 225, 228, 229, 230
- Significance level: $\alpha = 0.05$
- Null hypothesis: $H_0 : \sigma^2 \leq 2.25$
- Alternative hypothesis: $H_1 : \sigma^2 > 2.25$ (**right-tailed test** for variance)

Right-Tailed Test

We will use the **chi-squared test for variance**:

$$Q = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1} \text{ under } H_0$$

Sample Variance

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

- x_i = each observation
- $\bar{x}$ = sample mean
- n = sample size

$$\bar{x} = \frac{2276}{10} = 227.6$$

$$S^2 = \frac{46.40}{n-1} = \frac{46.40}{9} \approx 5.1556$$

Sample variance: $S^2 \approx 5.156\,\text{mm}^2$

EXERCISE 9.49 B): SOLUTION

Answer:

b) Test for variance

$$q_0 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{9 \cdot 5.156}{2.25} \approx \frac{46.404}{2.25} \approx 20.62$$

Degrees of freedom: $df = n - 1 = 9$

Step 2: Critical value

Right-tailed test at $\alpha = 0.05$:

RR = [16.919; +∞[

$\chi^2_{0.95,9} \approx 16.919$

Decision Using the RR

Step 3: Decision

$$q_0 = 20.62 > 16.919$$

$\Rightarrow$ Reject H_0 at the 5% significance level.

Conclusion

There is **evidence that the population variance exceeds 2.25 mm²**, so the machine may indeed be producing **excessive variability in sheet thickness**.

HOMEWORK OF LECTURE 16: QUESTIONS

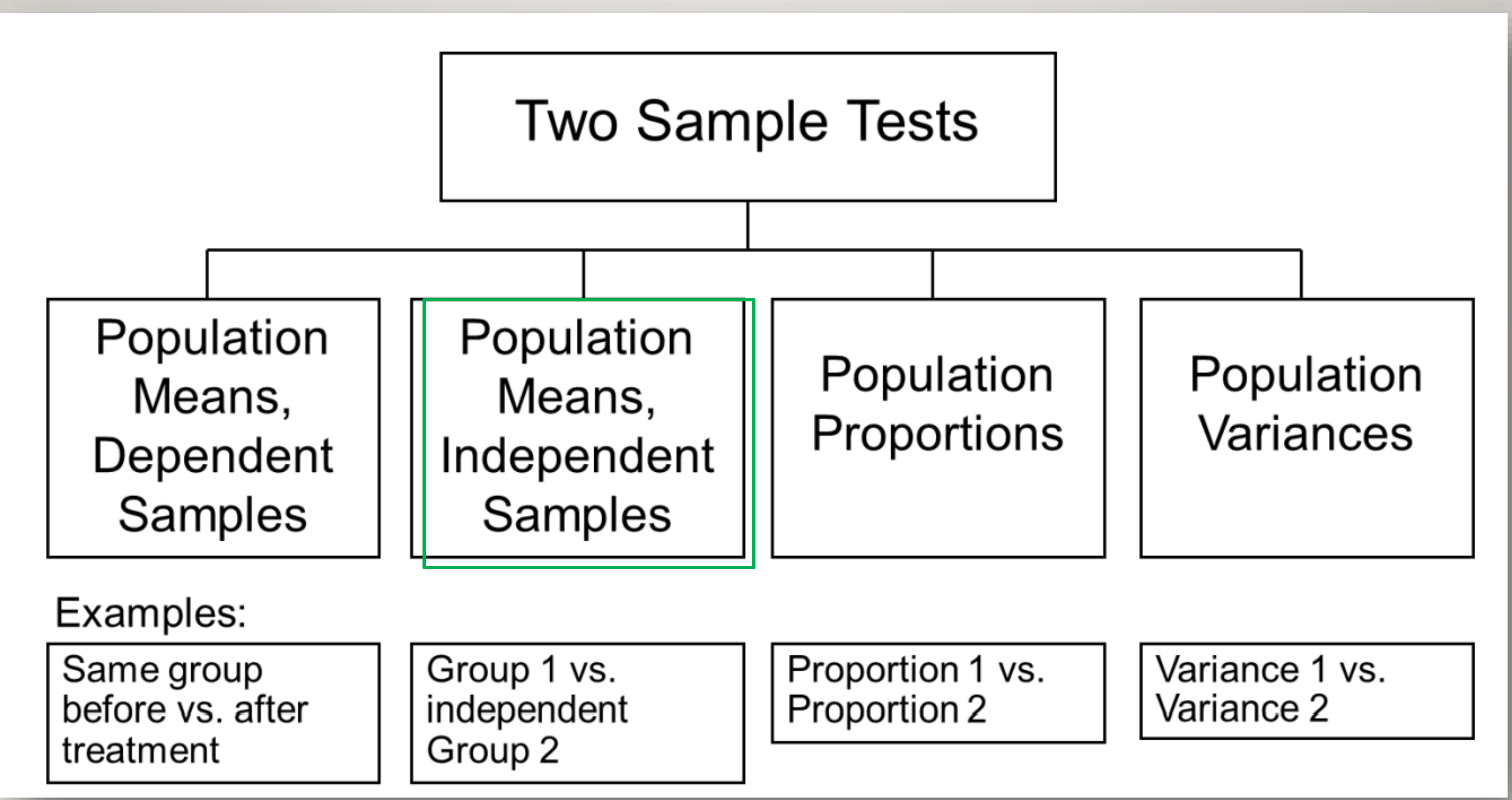
EXERCISE 9.51

9.51 A company produces electric devices operated by a thermostatic control. The standard deviation of the temperature at which these controls actually operate should not exceed 2.0°F. For a random sample of 20 of these controls, the sample standard deviation of operating temperatures was 2.36°F. Stating any assumptions you need to make, test, at the 5% level, the null hypothesis that the population standard deviation is 2.0 against the alternative that it is larger.

Newbold et al (2013)

LECTURE 16: TESTS OF THE DIFFERENCE BETWEEN TWO NORMAL POPULATION MEANS

TWO SAMPLE TESTS



Newbold et al (2013)

INDEPENDENT SAMPLES

Population means, independent samples

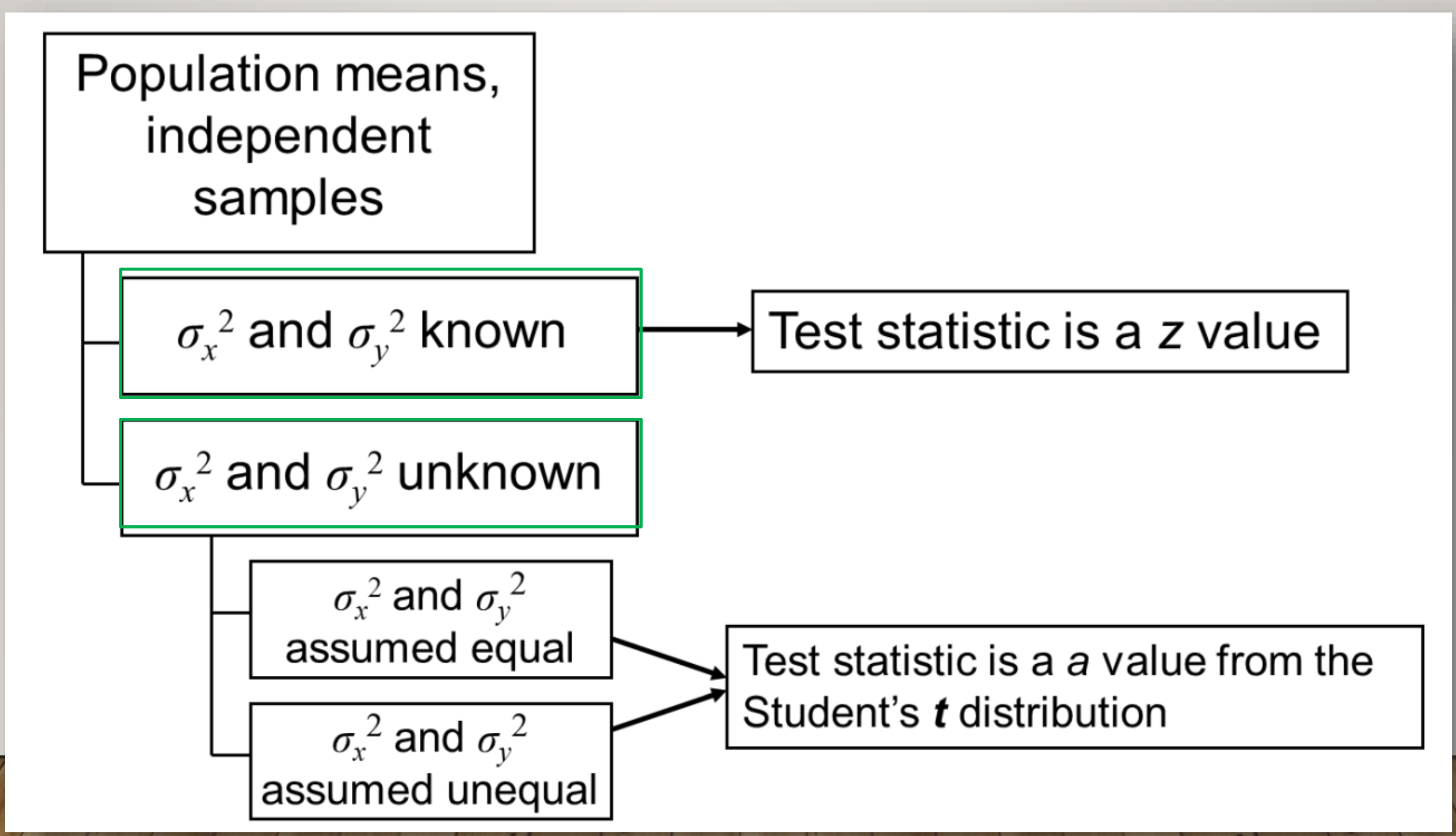
Tests of the Difference Between Two Normal Population Means: Dependent Samples

Goal: Form a confidence interval for the difference between two population means, $\mu_x - \mu_y$

- Different populations
  - Unrelated
  - Independent
    - Sample selected from one population has no effect on the sample selected from the other population
  - Normally distributed

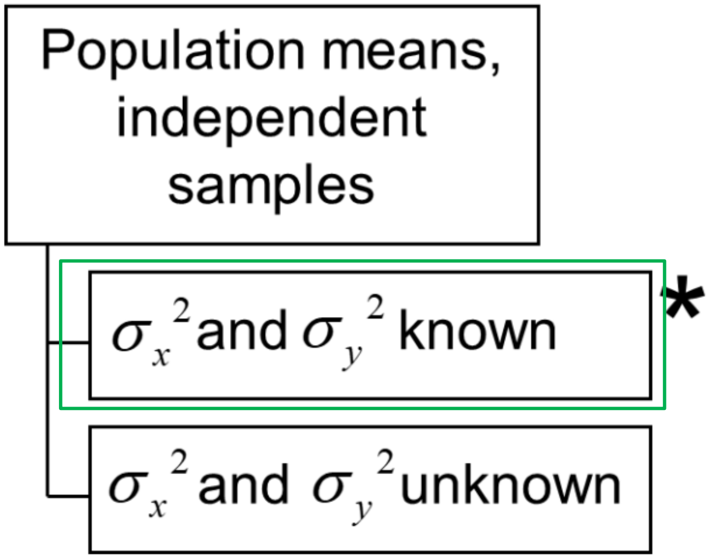
Newbold et al (2013)

DIFFERENCE BETWEEN TWO MEANS



Newbold et al (2013)

TESTS FOR THE DIFFERENCE BETWEEN TWO NORMAL POPULATION MEANS (KNOWN VARIANCES)



Assumptions:

- Samples are randomly and independently drawn
- both population distributions are normal
- Population variances are known

Newbold et al (2013)

TESTS FOR THE DIFFERENCE BETWEEN TWO NORMAL POPULATION MEANS (KNOWN VARIANCES)

Population means, independent samples

- σ_x^2 and σ_y^2 known *
- σ_x^2 and σ_y^2 unknown

Formula given in the formula sheet

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

When σ_x^2 and σ_y^2 are known and both populations are normal, the variance of $\bar{X} - \bar{Y}$ is

$$\sigma^2_{\bar{X}-\bar{Y}} = \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}$$

…and the random variable

$$Z = \frac{(\bar{x} - \bar{y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma^2_x}{n_X} + \frac{\sigma^2_y}{n_Y}}}$$

has a standard normal distribution

Newbold et al (2013)

TESTS FOR THE DIFFERENCE BETWEEN TWO NORMAL POPULATION MEANS (KNOWN VARIANCES)

Population means, independent samples

- σ_x^2 and σ_y^2 known *
- σ_x^2 and σ_y^2 unknown

Two-Tailed Test

$$H_0 : \mu_x - \mu_y = 0$$

The test statistic for $\mu_x - \mu_y$ is:

$$z = \frac{(\bar{x} - \bar{y})}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$

Newbold et al (2013)

HYPOTHESIS TESTS FOR TWO POPULATION MEANS

Two Population Means, Independent Samples

Left-Tailed Test	Right-Tailed Test	Two-Tailed Test
$H_0: \mu_x \geq \mu_y$	$H_0: \mu_x \leq \mu_y$	$H_0: \mu_x = \mu_y$
$H_1: \mu_x < \mu_y$	$H_1: \mu_x > \mu_y$	$H_1: \mu_x \neq \mu_y$
i.e.,	i.e.,	i.e.,
$H_0: \mu_x - \mu_y \geq 0$	$H_0: \mu_x - \mu_y \leq 0$	$H_0: \mu_x - \mu_y = 0$
$H_1: \mu_x - \mu_y < 0$	$H_1: \mu_x - \mu_y > 0$	$H_1: \mu_x - \mu_y \neq 0$

Newbold et al (2013)

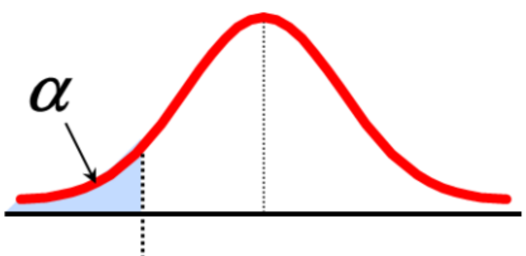
TESTS FOR THE DIFFERENCE BETWEEN TWO NORMAL POPULATION MEANS (KNOWN VARIANCES): DECISION USING THE RR

Two Population Means, Independent Samples, Variances Known

Left-Tailed Test

$H_0 : \mu_x - \mu_y \geq 0$

$H_1 : \mu_x - \mu_y < 0$



z_α

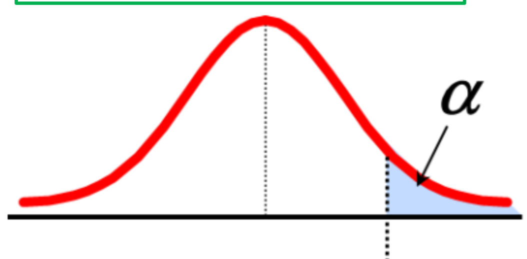
Reject H_0 if $z < z_\alpha$

RR = $]-\infty; z_\alpha]$

Right-Tailed Test

$H_0 : \mu_x - \mu_y \leq 0$

$H_1 : \mu_x - \mu_y > 0$



$z_{1-\alpha}$

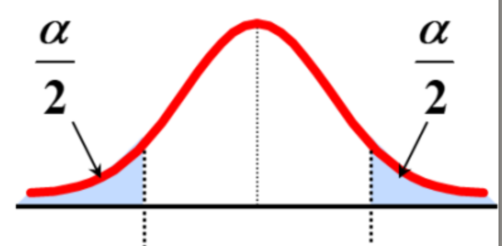
Reject H_0 if $z > z_{1-\alpha}$

RR = $[z_{1-\alpha}; +\infty[$

Two-Tailed Test

$H_0 : \mu_x - \mu_y = 0$

$H_1 : \mu_x - \mu_y \neq 0$



$z_{\alpha/2}$ $z_{1-\alpha/2}$

Reject H_0 if $z < z_{\alpha/2}$

or $z > z_{1-\alpha/2}$

RR = $]-\infty; -z_{1-\alpha/2}] \cup [z_{1-\alpha/2}; +\infty[$

Newbold et al (2013)

EXERCISE 10.6

10.6 You have been asked to determine if two different production processes have different mean numbers of units produced per hour. Process 1 has a mean defined as μ_1 and process 2 has a mean defined as μ_2. The null and alternative hypotheses are as follows:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

Use a random sample of 25 observations from process 1 and 28 observations from process 2 and the known variance for process 1 equal to 900 and the known variance for process 2 equal to 1,600. Can you reject the null hypothesis using a probability of Type I error $\alpha = 0.05$ in each case?

a. The process means are 50 and 60.

Newbold et al (2013)

THANKS!

Questions?