TOPIC 5: HIPOTHESIS TESTING

5.1. KEY ISSUES

- A hypothesis test is a statistical method allowing to choose between two alternatives regarding the behaviour of the random variable under study.
- When the goal consists in giving values to the unknown parameters in the population, we will use parametric tests (topic 5).
- When the hypothesis concern other aspects of the variable, such as the probability distribution or sample features (independence and homogeneity) we will employ nonparametric tests (topic 6).

- General approach:

  - Let ξ be a random variable with probability distribution $f(x;\theta)$ θ being an unknown parameter.
  - Then we formulate a null hypothesis H_0 over the value of the parameter, for instance, H_0: $\theta = \theta_0$ against an alternative hypothesis H_1: $\theta \neq \theta_0$
  - Later on we take a s.r.s. and based on the point estimator θ^* we elaborate a decision rule in order to accept or reject H_0.

- ❖ The hypothesis are established by the researcher following some criterion (look at point 1 in section 5.2)
- ❖ Then data are collected and hypothesis are tested with the sample with the purpose of choosing one of them.
- ❖ In the decision making process four situations may arise depending on which hypothesis is true or false:

	States of nature (reality)	
Decision	H_0 true	H_0 false
Accept H_0 (Reject H_1)	Correct decision 1-α (Confidence level)	Type II Error $\beta = P(Accept\ H_0/H_0\ false)$
Reject H_0 (Accept H_1)	Type I Error $\alpha = P(Reject\ H_0/\ H_0\ true)$ (Significance level)	Correct decision 1-β (Power of the contrast)

- ❖ We would like α and β be as small as possible. However, for a given sample, a reduction in α involves an increase in β and viceversa, except when enlarging the sample size *n* at the same time.
- ❖ The choice between controlling α or β depends on the cost arising from incurring in type I error or type II error, respectively
- ❖ Nevertheless, we usually prefer to fix α in a low level and choose, among all possible tests for that α, that one with the smallest β (hence, with the highest power)

TYPE OF HYPOTHESIS

- Hypothesis are conjectures made about the value of some unknown parameter of a given population.
- There are two kind of hypothesis: single and composite
  a. Single hypothesis specify a unique value for the parameter. In these cases the density function of the variable is completely determined.
  b. Composite hypothesis assign a range of values to the unknown parameter.
- Under another point of view, the null hypothesis must be differentiated from the alternative one.
  a. The null hypothesis H_0 is established first and will be kept true unless data gathered from a s.r.s. shows strong evidence against that. However, to accept the null hypothesis doesn't mean it is correct, just that data do not offer strong evidence to reject it.
  b. The alternative hypothesis H_1 is that one accepted when the null hypothesis H_0 is rejected.

❖ Possible ways to establish a hypothesis testing:

I	$H_0: \theta = \theta_0$ $H_1: \theta \neq \theta_0$		
II	$H_0: \theta = \theta_0$ $H_1: \theta < \theta_0$	or	$H_0: \theta \geq \theta_0$ $H_1: \theta < \theta_0$
III	$H_0: \theta = \theta_0$ $H_1: \theta > \theta_0$	or	$H_0: \theta \leq \theta_0$ $H_1: \theta > \theta_0$

❖ If the possible direction of the alternative hypothesis is known, then a one-tailed test will be formulated (cases II and III), otherwise a two-tailed test will apply (case I)

5.2. METHODOLOGY

To carry out a hypothesis testing we must follow some steps:

1.- To formulate both, the null hypothesis H_0 and the alternative hypothesis H_1

General rule: The researcher chooses as alternative hypothesis H_1 what he or she considers true, this is, that one including the real value for the parameter.

Particular situations:

- In a risky context, the conjecture provoking the highest risk if it is rejected being true must appear as the null hypothesis H_0. The reason is that we control that risk through α.
- Sometimes the researcher believes the parameter being either higher or lower than certain value. Then that conjecture appears as the alternative hypothesis H_1. In this situation, the researcher thinks H_1 is true in the sense that it incorporates the value of the unknown parameter.
- Often the researcher tests a single value for the unknown parameter (with or without information). In such a case, that conjecture appears in the null hypothesis H_0.

In any context, the researcher will accept (not reject) the null hypothesis H_0 unless data provide strong evidence against it, hence accepting the alternative hypothesis H_1.

2.- <u>To determine the test statistic</u>

- It is the statistic to be used in order to accept or reject the null hypothesis H_0.
- Such statistic must incorporate the discrepancy existing between the sample (represented by the point estimator θ^*) and the null hypothesis H_0 (represented by the value θ_0 taken by the parameter in that hypothesis):

$$d(\theta_0; \theta^*) \equiv d$$

- The test statistic must follow a known probability distribution when H_0 is true.

3.- <u>To establish the level of significance α and determine the critical region</u>

$$\alpha = P(Reject\ H_0\ /\ H_0\ true)$$

- Critical region is the interval or intervals in R in which the discrepancy between the sample and H_0 is large enough as to reject H_0
- When will the discrepancy be large enough?

  Answer: when such discrepancy has a low probability to happen if H_0 is true

  That probability is the level of signficance and is represented by α:

$$P(d > d_c\ /\ H_0\ true) = \alpha$$

  d_c is called *critical value*

α is established by the researcher, usually 5% or 1%

- Those intervals in R being complementary to the critical region are called acceptance region.

  The position of both regions depending on the hypothesis (one-tail or two-tails)

4.- <u>To select a s.r.s., to calculate the value of the test statistic and make a decision</u>

In this stage data are collected and the estimate of the test statistic, <u>which we will denote as d_0</u> is obtained. Then a decision must be made:

a. If that estimate (d_0) falls inside the acceptance region, then the discrepancy between the sample and the null hypotehsis H_0 is not considered large enough and H_0 is accepted (not rejected)
b. On the contrary if such estimate (d_0) falls inside the critical región, then the discrepancy is considered large enough and H_0 is rejected (H_1 is accepted)

OTHER CONCEPTS

<u>p-value</u>

- The determination of the critical region through the significant level α has some limitations

  By one side, the decision making depends strongly on the level of α. Hence, the researcher may reject the null hypothesis H_0 for

$\alpha = 5\%$ but accept the same H_0 for $\alpha = 2\%$.

By another side, the result of the test does not inform of the degree of strength used to reject H_0.

The procedure referred to as *p-value* allows to overcome these difficulties.

❖ The *p-value* is the probability of obtaining a discrepancy at least as extreme as that one observed in the sample when the null hypothesis H_0 is true.

The discrepancy observed in the sample being d_0 the estimate of the test statistic. Therefore:

$$p - value = P(d \geq d_0 \;/\; H_0)$$

The lower the *p-value*, the smaller the probability to find a discrepancy as extreme as that one observed under the null hypothesis and, therefore, the lower the likelihood of H_0 (Peña, D.).

The *p-value* can also be interpreted as the sampling significant level, meaning that it is the probability of rejecting the null hypothesis H_0 when the null hypothesis is true supplied by the sample

❖ Hence, the decision being:

a. If *p-value* $< \alpha$ reject H_0

b. If *p-value* $\geq \alpha$ accept H_0

Power of a test

- When H_0 is rejected in a hypothesis test, the probability α of type I error or the *p-value* are known, both being of a small amount.
- However when H_0 is accepted it may be true or not, the latter leading to make a type II error with probability β.

  β is the probability of accepting H_0 when it is false. This is a problem because the parameter is unknown, then β is not unique but depending on the infinite values the parameter may take under the alternative hypothesis H_1:

$$\beta = P(Accept\ H_0\ /\ H_0\ false) = P(Accept\ H_0\ /\ H_1\ true)$$

- Varying the actual value of the parameter in H_1 a function called Operating Caracteristic (OC) curve is developed:

$$\beta(\theta) = P(Accept\ H_0 /\ \theta)$$

  Nevertheless it is customary to use its complementary, the power function:

$$Power(\theta) = 1 - \beta(\theta) = P(Reject\ H_0 /\ \theta)$$

- The power of a test does not directly affect the decision of either accepting or rejecting the null hypothesis, provided that the decision depends only on α (Newbold P. et al.)
- However, for a given α, the power of a test grows when the sample size n increases

- ❖ By another side, having two tests for a given α based on two different discrepancies, the test with the smallest β will be preferred. In order words we will prefer the most powerful test (Peña, D.)

- ❖ Finally, as indicated above, to accept H_0 provokes uncertainty because β cannot be calculated and this explains why the test is intended to reject H_0.

Confidence Interval Estimation and Hypothesis Testing

- ❖ In general it can be proved that the $\gamma=1-\alpha$ confidence interval estimation includes all the θ_0 values of the unknown parameter for which the null hypothesis H_0 would be accepted with a significance level of α

- ❖ As an example let's consider the case of the mean from a N(μ;σ) with σ unknown:

H_0: $\mu=\mu_0$
H_1: $\mu\neq\mu_0$

$$If\ H_0\ is\ true\ \rightarrow d=\frac{\bar{x}-\mu_0}{s_1/\sqrt{n}}\sim t_{n-1}$$

$$With\ Acceptance\ Region\ \left|\frac{\bar{x}-\mu_0}{s_1/\sqrt{n}}\right|\leq t_{n-1,\alpha/2}$$

$$and\ probability\ P\left(-t_{n-1,\alpha/2}\leq\frac{\bar{x}-\mu_0}{s_1/\sqrt{n}}\leq +t_{n-1,\alpha/2}\right)=1-\alpha$$
$$=\gamma$$

Solving for μ_0: $$\mu_0\epsilon\left[\bar{x}\mp t_{n-1,\alpha/2}\frac{s_1}{\sqrt{n}}\right]_\gamma$$

5.3. HYPOTHESIS TESTING FOR ONE POPULATION

5.3.1. Tests for the mean. Normal Populations

Test	Test Statistic d	AR	CR	p-value
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$ σ^2 known	$\frac{(\bar{x}-\mu_0)\sqrt{n}}{\sigma} \sim Z$	$\frac{\lvert\bar{x}-\mu_0\rvert\sqrt{n}}{\sigma} \leq z_{\alpha/2}$	$\frac{\lvert\bar{x}-\mu_0\rvert\sqrt{n}}{\sigma} > z_{\alpha/2}$	$P(\lvert Z\rvert \geq \lvert d_0\rvert) =$ $= 2 \cdot P(Z \geq \lvert d_0\rvert)$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$ σ^2 known	idem		$\frac{(\bar{x}-\mu_0)\sqrt{n}}{\sigma} > z_{\alpha}$	$P(Z \geq d_0)$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$ σ^2 known	idem		$\frac{(\bar{x}-\mu_0)\sqrt{n}}{\sigma} < -z_{\alpha}$	$P(Z \leq d_0)$

$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$ σ^2 unknown	$\frac{(\bar{x}-\mu_0)\sqrt{n}}{s_1} \sim t_{n-1}$	$\left\lvert\frac{(\bar{x}-\mu_0)\sqrt{n}}{s_1}\right\rvert \leq t_{n-1,\alpha/2}$	$\left\lvert\frac{(\bar{x}-\mu_0)\sqrt{n}}{s_1}\right\rvert > t_{n-1,\alpha/2}$	$P(\lvert t_{n-1}\rvert \geq \lvert d_0\rvert) =$ $= 2 \cdot P(t_{n-1} \geq \lvert d_0\rvert)$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$ σ^2 unknown	idem		$\frac{(\bar{x}-\mu_0)\sqrt{n}}{s_1} > t_{n-1,\alpha}$	$P(t_{n-1} \geq d_0)$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$ σ^2 unknown	idem		$\frac{(\bar{x}-\mu_0)\sqrt{n}}{s_1} < -t_{n-1,\alpha}$	$P(t_{n-1} \leq d_0)$

Note 1: Z means a N(0;1)

Note 2: Remember that when n is large enough the Student's t can be approximated by a N(0;1)

Note 3: If the population distribution is unknown but n is large enough the N(0;1) would be used instead of the Student's t in the last three situations

5.3.2. Tests for the population proportion (n large enough)

Test	Test Statistic d	AR	CR	p-value
$H_0: p = p_0$ $H_1: p \neq p_0$	$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} \sim Z$	$\dfrac{\left\lvert \hat{p} - p_0 \right\rvert}{\sqrt{\dfrac{p_0 q_0}{n}}} \leq z_{\alpha/2}$	$\dfrac{\left\lvert \hat{p} - p_0 \right\rvert}{\sqrt{\dfrac{p_0 q_0}{n}}} > z_{\alpha/2}$	$P(\lvert Z \rvert \geq \lvert d_0 \rvert) =$ $= 2 \cdot P(Z \geq \lvert d_0 \rvert)$
$H_0: p = p_0$ $H_1: p > p_0$	idem		$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} > z_{\alpha}$	$P(Z \geq d_0)$
$H_0: p = p_0$ $H_1: p < p_0$	idem		$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} < -z_{\alpha}$	$P(Z \leq d_0)$

5.3.3. Tests for the variance. Normal Populations.

Test	Test Statistic d	AR	CR	p-value
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$\frac{(n-1)s_1^2}{\sigma_0^2} \sim \chi^2_{n-1}$	$\chi^2_{n-1,1-\alpha/2} \leq \frac{(n-1)s_1^2}{\sigma_0^2} \leq \chi^2_{n-1,\alpha/2}$	$\frac{(n-1)s_1^2}{\sigma_0^2} < \chi^2_{n-1,1-\alpha/2}$ *o* $\frac{(n-1)s_1^2}{\sigma_0^2} > \chi^2_{n-1,\alpha/2}$	$2 \cdot min[P(\chi^2_{n-1} \geq d_0),$ $P(\chi^2_{n-1} \leq d_0)]$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$	idem		$\frac{(n-1)s_1^2}{\sigma_0^2} > \chi^2_{n-1,\alpha}$	$P(\chi^2_{n-1} \geq d_0)$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$	idem		$\frac{(n-1)s_1^2}{\sigma_0^2} < \chi^2_{n-1,1-\alpha}$	$P(\chi^2_{n-1} \leq d_0)$

5.4. HYPOTHESIS TESTING FOR TWO POPULATIONS

5.4.1. Tests for the differences of population means. Independent Samples. Normal Populations

Test	Test Statistic d	AR	CR	p-value
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y \neq D$ σ_X and σ_Y are known	$\dfrac{\bar{x} - \bar{y} - D}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} \sim Z$	$\dfrac{\lvert\bar{x} - \bar{y} - D\rvert}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} \leq +z_{\alpha/2}$	$\dfrac{\lvert\bar{x} - \bar{y} - D\rvert}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} > +z_{\alpha/2}$	$P(\lvert Z\rvert \geq \lvert d_0\rvert) =$ $= 2 \cdot P(Z \geq \lvert d_0\rvert)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y > D$ σ_X and σ_Y are known	idem		$\dfrac{\bar{x} - \bar{y} - D}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} > z_\alpha$	$P(Z \geq d_0)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y < D$ σ_X and σ_Y are known	idem		$\dfrac{\bar{x} - \bar{y} - D}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} < -z_\alpha$	$P(Z \leq d_0)$

Test	Test Statistic d	AR	CR	p-value
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y \neq D$ $\sigma_X = \sigma_Y = \sigma$ unknown	$\dfrac{(\bar{x} - \bar{y}) - D}{s^* \sqrt{\dfrac{1}{n} + \dfrac{1}{m}}} \sim t_{n+m-2}$ $s^* = \sqrt{\dfrac{(n-1)s_{1x}^2 + (m-1)s_{1y}^2}{n+m-2}}$	$\lvert d \rvert \leq t_{n+m-2,\alpha/2}$	$\lvert d \rvert > t_{n+m-2,\alpha/2}$	$P(\lvert t_{n+m-2} \rvert \geq \lvert d_0 \rvert) =$ $= 2 \cdot P(t_{n+m-2} \geq \lvert d_0 \rvert)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y > D$ $\sigma_X = \sigma_Y = \sigma$ unknown	idem		$d > t_{n+m-2,\alpha}$	$P(t_{n+m-2} \geq d_0)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y < D$ $\sigma_X = \sigma_Y = \sigma$ unknown	idem		$d < -t_{n+m-2,\alpha}$	$P(t_{n+m-2} \leq d_0)$

Test	Test Statistic d	AR	CR	p-value
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y \neq D$ σ_X and σ_Y unknown n and m large	$\dfrac{(\bar{x}-\bar{y})-D}{\sqrt{\dfrac{s_{1x}^2}{n}+\dfrac{s_{1y}^2}{m}}} \sim Z$	$\lvert d\rvert \leq +z_{\alpha/2}$	$\lvert d\rvert > +z_{\alpha/2}$	$P(\lvert Z\rvert \geq \lvert d_0\rvert) =$ $= 2 \cdot P(Z \geq \lvert d_0\rvert)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y > D$ σ_X and σ_Y unknown n and m large	idem		$d > z_\alpha$	$P(Z \geq d_0)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y < D$ σ_X and σ_Y unknown n and m large	idem		$d < -z_\alpha$	$P(Z \leq d_0)$

$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y \neq D$ σ_X and σ_Y unknown n and m small	$\dfrac{(\bar{x}-\bar{y})-D}{\sqrt{\dfrac{s_{1x}^2}{n}+\dfrac{s_{1y}^2}{m}}} \sim t_k$ k = n+m-2-g $g = \dfrac{\left[(m-1)\dfrac{s_{1x}^2}{n}-(n-1)\dfrac{s_{1y}^2}{m}\right]^2}{(m-1)\dfrac{s_{1x}^4}{n^2}-(n-1)\dfrac{s_{1y}^4}{m^2}}$	$\lvert d \rvert \leq t_{k,\alpha/2}$	$\lvert d \rvert > t_{k,\alpha/2}$	$P(\lvert t_k \rvert \geq \lvert d_0 \rvert) =$ $= 2 \cdot P(t_k \geq \lvert d_0 \rvert)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y > D$ σ_X and σ_Y unknown n and m small	idem		$d > t_{k,\alpha}$	$P(t_k \geq d_0)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y < D$ σ_X and σ_Y unknown n and m small	idem		$d < -t_{k,\alpha}$	$P(t_k \leq d_0)$

5.4.2. Tests for the difference of population means. *Dependent Samples*. Normal Populations

Test	Test Statistic d	AR	CR	p-value
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y \neq D$	$\dfrac{(\bar{x}-\bar{y})-D}{s_{1d}/\sqrt{n}} \sim t_{n-1}$ S_{1d} is the bias-corrected sample standard deviation of the differences in the samples	$\lvert d\rvert \leq +t_{n-1,\alpha/2}$	$\lvert d\rvert > +t_{n-1,\alpha/2}$	$P(\lvert t_{n-1}\rvert \geq \lvert d_0\rvert) =$ $= 2 \cdot P(t_{n-1} \geq \lvert d_0\rvert)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y > D$	idem		$d > t_{n-1,\alpha}$	$P(t_{n-1} \geq d_0)$
$H_0: \mu_X - \mu_Y = D$ $H_1: \mu_X - \mu_Y < D$	idem		$d < -t_{n-1,\alpha}$	$P(t_{n-1} \leq d_0)$

5.4.3. Tests for the difference of proportions. Large independent samples.

Test	Test Statistic d	AR	CR	p-value
$H_0: p_X - p_Y = p_0$ $H_1: p_X - p_Y \neq p_0$	$\dfrac{(\widehat{p_x} - \widehat{p_y}) - p_0}{\sqrt{\dfrac{\widehat{p_x}\widehat{q_x}}{n} + \dfrac{\widehat{p_y}\widehat{q_y}}{m}}} \sim Z$	$\lvert d \rvert \leq +z_{\alpha/2}$	$\lvert d \rvert > +z_{\alpha/2}$	$P(\lvert Z\rvert \geq \lvert d_0\rvert) =$ $= 2 \cdot P(Z \geq \lvert d_0\rvert)$
$H_0: p_X - p_Y = p_0$ $H_1: p_X - p_Y > p_0$	idem		$d > z_\alpha$	$P(Z \geq d_0)$
$H_0: p_X - p_Y = p_0$ $H_1: p_X - p_Y < p_0$	idem		$d < -z_\alpha$	$P(Z \leq d_0)$

5.4.4. Tests for the quotient of variances. Normal Populations. Independent Samples.

Test	Test Statistic d	AR	CR	p-value
$H_0: \sigma_X^2 = \sigma_Y^2$ $H_0: \sigma_X^2 \neq \sigma_Y^2$	$\frac{S_{1X}^2}{S_{1Y}^2} \sim F_{n-1,m-1}$	$F_{n-1,m-1,1-\alpha/2} \leq d \leq F_{n-1,m-1,\alpha/2}$ where $F_{n-1,m-1,1-\alpha/2} = \frac{1}{F_{m-1,n-1,\alpha/2}}$	$d > F_{n-1,m-1,\frac{\alpha}{2}}$ or $d < F_{n-1,m-1,1-\frac{\alpha}{2}}$	$2 \cdot mín\left[P\left(F_{n-1,m-1,\alpha/2} \geq d_0\right),\ P\left(F_{n-1,m-1,\alpha/2} \leq d_0\right)\right]$
$H_0: \sigma_X^2 = \sigma_Y^2$ $H_0: \sigma_X^2 > \sigma_Y^2$			$d > F_{n-1,m-1,\alpha}$	$P\left(F_{n-1,m-1,\alpha/2} \geq d_0\right)$
$H_0: \sigma_X^2 = \sigma_Y^2$ $H_0: \sigma_X^2 < \sigma_Y^2$			$d < F_{n-1,m-1,1-\alpha}$	$P\left(F_{n-1,m-1,\alpha/2} \leq d_0\right)$