Hypothesis testing : recap

Definition 4.1. A **hypothesis** is a statement about a parameter (or parameters) of interest.

H_0 : null hypothesis, default position – gives us distribution of test statistic

H_1 : alternative hypothesis

Example: $x_1, x_2, \ldots, x_n \sim N(\theta, 1)$

unknown

$H_0: \theta = 0$

$H_1: \theta \neq 0$

$H_0: \theta \leq 2$

$H_1: \theta > 2$

> 'Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis.'

Result	Decisions and statements we can make regarding the	
	Null hypothesis H_0	**Alternative hypothesis** H_1
$p < \alpha$	Reject H_0	Data supports H_1
$p \not< \alpha$	Fail to reject H_0	–

Summary: making decisions and statements about hypotheses

- We **specify** our null and alternative hypotheses **in advance** of the experiment.
- Our decision about the null hypothesis is based on **data from an experiment**.
- We **never 'accept'** a hypothesis as being true after analysing the data.
- We either **reject** or **fail to reject** the null hypothesis H_0.
- Failing to reject H_0 does not mean it is 'accepted'; the result of the experiment is **inconclusive**.
- If H_0 is rejected, we may say that **the data supports** the alternative hypothesis H_1.

Specify a significance threshold α in advance

Common values : $\alpha = 0.05$, $\alpha = 0.1$

$\alpha = 0.01$, $\alpha = 0.0001$

Example

- $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$, $n = 100$

  θ unknown

z	$P(Z < z)$
1.281	0.900
1.645	0.950
1.960	0.975
2.326	0.990
2.576	0.995

- $H_0: \theta = 0$

  $H_1: \theta \neq 0$

- Use $\alpha = 0.05$

- Observe $x_1, x_2, \ldots, x_{100}$; $\bar{x} = 1.5$

$F_Z(1.5) = 0.9932$

$$\bar{X} \sim N\left(\theta, \frac{1}{100}\right)$$

$$Z = \frac{\theta - \bar{X}}{\sigma/\sqrt{n}} = \frac{\theta - \bar{X}}{1/10}$$

$H_0: \theta = 0$

observe $\bar{x} = 1.5$ $\quad z = \frac{0 - 1.5}{1/10} = -15$

Suppose $\bar{x} = -0.26$

$$z = \frac{0 - (-0.26)}{1/10} = 2.6$$

$$p = 1 - F_Z(z)$$
$$= 1 - F_Z(2.6)$$
$$< 1 - 0.995$$
$$< 0.005$$

$F_Z(2.6) > 0.995$

$-F_Z(2.6) < -0.995$

$1 - F_Z(2.6) < 1 - 0.995$

< 0.005

p-value : Score between 0 and 1
which gives an assessment of
how well the data follow
the assumptions of the null hypothesis

- values close to 0 mean
data are extreme in relation to
null hypothesis.

what if instead of $\bar{x} = -0.26$

$\bar{x} = +0.26$

$p = 1 - F_Z(z)$ (typically)

$$z = \frac{\theta - \bar{x}}{\sigma/\sqrt{n}} = \frac{0 - 2.6}{1/10}$$

$$= -2.6$$

$$F_Z(-2.6) \approx 0.005$$

$$p = 1 - F_Z(-2.6) \approx 0.995$$

Example

- $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$, $n = 100$

  θ unknown

- $H_0 : \theta = 0$

  $H_1 : \theta \neq 0$

- Use $\alpha = 0.05$

z	$P(Z < z)$
1.281	0.900
1.645	0.950
1.960	0.975
2.326	0.990
2.576	0.995

- Observe $x_1, x_2, \ldots, x_{100}$; $\bar{x} = 1.5$

$F_Z(1.5) = 0.9932$

$$\bar{X} \sim N\left(\theta, \frac{1}{100}\right)$$

$$Z = \frac{\theta - \bar{X}}{\sigma/\sqrt{n}} = \frac{\theta - \bar{X}}{1/10}$$

$H_0 : \theta = 0$

observe $\bar{x} = 1.5$ $\quad z = \frac{0 - 1.5}{1/10} = -15$

Suppose $\bar{x} = -0.26$

$$z = \frac{0 - (-0.26)}{1/10} = 2.6$$

$$p = 1 - F_Z(z)$$

$$= 1 - F_Z(2.6)$$

$$< 1 - 0.995$$

$$< 0.005$$

$F_Z(2.6) > 0.995$

$-F_Z(2.6) < -0.995$

$1 - F_Z(2.6) < 1 - 0.995$

< 0.005

p-value : Score between 0 and 1
which gives an assessment of
how well the data follow
the assumptions of the null hypothesis

- values close to 0 mean
data are extreme in relation to
null hypothesis.

what if instead of $\bar{x} = -0.26$

$\bar{x} = +0.26$

$p = 1 - F_Z(z)$ (typically)

$$z = \frac{\theta - \bar{x}}{\sigma/\sqrt{n}} = \frac{0 - 2.6}{1/10}$$

$$= -2.6$$

$F_Z(-2.6) \approx 0.005$

$p = 1 - F_Z(-2.6) \approx 0.995$

Distribution of a p-value

Theorem 6.5.1

Let X be a continuous random variable with cumulative distribution function (cdf) F_X, and let $Y = F_X(X)$.

Then $Y \sim U(0,1)$

uniform distribution on interval $[0,1]$.

Proof: For simplicity assume F_X^{-1} (inverse) exists

for any $x \in \mathbb{R}$

$0 \leq F_X(x) \leq 1$; Recall $Y = F_X(X)$

$\therefore P(Y<0) = 0 = P(Y>1)$

$P(Y \leq 1) = 1$

pick any $y \in (0,1)$ then

$$P(Y \leq y) = P(F_X(X) \leq y)$$

$F_X(x) = P(X \le x) \qquad = P(X \le F_X^{-1}(y))$

$$= F_X(F_X^{-1}(y))$$

$$\Rightarrow P(Y \le y) = y$$

cdf of Y:

$$F_Y(y) = P(Y \le y) = \begin{cases} 0 & \text{if } y \le 0 \\ y & \text{if } 0 < y < 1 \\ 1 & \text{if } y \ge 1 \end{cases}$$

This is the cdf of $U(0,1)$

$$\Rightarrow \quad Y \sim U(0,1)$$

p-value is usually defined:

$$p = 1 - F_T(T) \qquad T: \text{test statistic}$$

or $p = F_T(T) \quad \sim U(0,1) \qquad V \sim U(0,1)$

$1 - V \sim U(0,1)$

Under H_0: T follows distribution with cdf F_T

Transformation:

"two sided" → "one sided" p-value

such that p close to 0 → extreme

p not close to 0 → NOT extreme

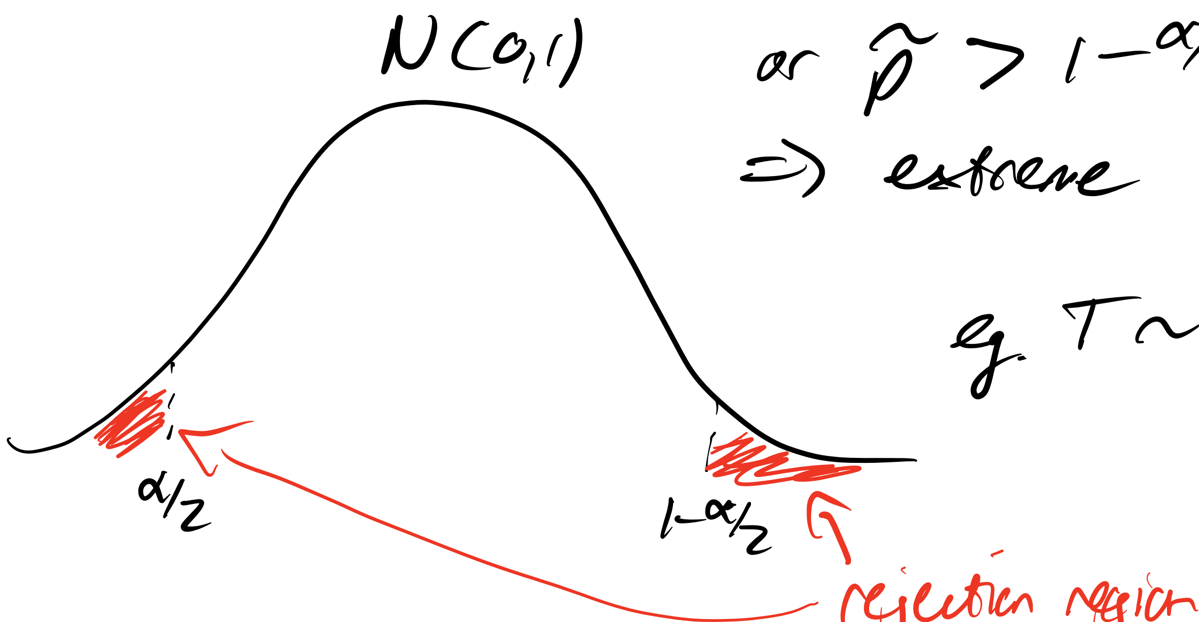
Start by alway defining

$\tilde{p} = 1 - F_T(T)$: T test statistic

for some $\alpha \in (0,1)$ if $\tilde{p} < \alpha/2$

or $\tilde{p} > 1 - \alpha/2$

$\Rightarrow$ extreme



e.g. $T \sim N(0,1)$

Proposition 4.4.2: Define $p = 1 - 2|\tilde{p} - \tfrac{1}{2}|$

then if $\tilde{p} < \alpha/2$ or $\tilde{p} > 1-\alpha/2$

$\Rightarrow p < \alpha$

$$\tilde{p} \sim U(0,1)$$

$$\tilde{p} - 1/2 \sim U(-1/2, 1/2)$$

$$|\tilde{p} - 1/2| \sim U(0, 1/2)$$

$$-2|\tilde{p} - 1/2| \sim U(-1, 0)$$

$$-2|\tilde{p} - 1/2| + 1 \sim U(0,1)$$

Type I and Type II errors

Definition 4.3. If the null hypothesis has been rejected, when in fact the null hypothesis is true, then we say a **Type I** error has occurred.

Type I error : α

Definition 4.4. If the null hypothesis fails to be rejected, when in fact the null hypothesis is false, then we say a **Type II** error has occurred.

Type II error : β

Definition 4.5. The probability of correctly rejecting the null hypothesis, when in fact the null hypothesis is false, is defined as the **power** of the test, and is computed as $\mathrm{P}\left(\text{Reject } H_0 \,|\, H_0 \text{ is false}\right) = 1 - \beta$, where β is the probability of a Type II error occurring.

All of these quantities are summarised in the following table.

		Given that the null hypothesis H_0 is	
		True	False
Decision	Reject H_0	**Type I Error** $\mathrm{P}\left(\text{Reject } H_0 \,\vert\, H_0 \text{ is true}\right) = \alpha$	Correct decision: **Power** $\mathrm{P}\left(\text{Reject } H_0 \,\vert\, H_0 \text{ is false}\right) = 1 - \beta$
	Fail to reject H_0	Correct decision $\mathrm{P}\left(\text{Fail to reject } H_0 \,\vert\, H_0 \text{ is true}\right) = 1 - \alpha$	**Type II Error** $\mathrm{P}\left(\text{Fail to reject } H_0 \,\vert\, H_0 \text{ is false}\right) = \beta$