Lezione 12 – Hypothesis Testing on a single population

Second Extra Points Test: Mon 26 May

Statistic Hypothesis

Partially known distribution, θ is unknown.

Random sample $X_1, \dots, X_n \overset{iid}{\sim} f_X(-, \theta)$, θ unknown parameter

observed sample $(x_1, \dots, x_n) \in \mathbb{R}^n$

We do not aim estimate features of the population (not estimating θ), we are testing hypothesis on the unknown parameter θ.

We are going to look at procedures to see if our data is consistent with the statement.

Example

Manufacturer claims batteries last more than N hours: how do we check his claim?

Suppose we extract n batteries and measure their lifetime.

We can compute $\bar{x}_n$, if this value is much smaller than N, we can conclude that we distrust the manufacturer.

Example

How do we check a coin is fair?

We flip n times and register the frequency of HEADS, if this is not too far from 0,5, we conclude the coin is fair.

General procedure to check if hypothesis is consistent with our data.

Definition:

A statistic hypothesis is a statement about the unknown parameters of one or more populations.

We talk about hypothesis since we don't know if the statement is true or not, since in general we cannot definitely show the veracity.

The truth or falsity of a particular hypothesis can never be known for certain. It is possible for a procedure to reach a wrong conclusion.
Also because our data is a realisation of random variables.

Example

We are interested in the burning rate of a solid propellant used to power aircrew escape systems.

$X \sim N(\theta, (2{,}5)^2)$

↳ Suppose known for now.

We are interested in deciding whether the mean rate is 50 cm/s

Formally: $H_0: \theta = 50$ $H_1: \theta \neq 50$

Remark:

Hypotheses are always statements about the population or distribution under study, not a statement about the observed sample/data.

NEVER write $H_0: \bar{x}_n = 50$

Definition

If we determine that the observed sample and hypothesis H_0 are consistent we say H_0 is accepted, while if they aren't consistent we say that H_0 is rejected.

REMARK: For this course, the rejection of H_0 is the acceptance of H_1.

$H_0 \rightarrow$ null hypothesis
$H_1 \rightarrow$ alternative hypothesis.

A simple hypothesis is a hypothesis which completely specifies the distribution.

The hypothesis is called composite if we are not completely defining θ (i.e. $\theta \leq 50$).

Definition:

A predecure leading to a decision on a hypothesis is called a test of hypotheses.

<u>Decision Rule</u>

↳ We build a subset C of $\mathbb{R}^n$ such that: ← observed sample space

If the observed sample $(x_1, \ldots, x_n) \in C$, we conclude that H_0 is rejected,

If $(x_1, \ldots, x_n) \notin C$ we conclude H_0 is <u>not rejected</u>

The set C is called the critical region or rejection region.

Our conclusion could be wrong, since our data is the observation of a random sample.

Type I error: rejecting H_0, when it's true

Type II error: rejecting H_1, when it is true.

Truth

		H_0	H_1
Accept	H_0	Ok	Type II error
	H_1	Type I error	Ok

It would be nice to keep both error probabilities very small: this is not possible though!

$\rightarrow$ can control

Type I probability $\quad P_{H_0}(\text{reject } H_0) = P_{H_0}(X \in C) \leq \alpha$ for a small α

Type II probability $\quad P_{H_1}(\text{rejecting } H_1) = P_{H_1}(\text{accepting } H_0) = P_{H_1}(X \notin C)$

$\hookrightarrow$ cannot control $\qquad$ $\rightarrow$ we can't control both.

Remark: The hypothesis (are) NOT symmetric in the test; since I'm able to control only Type I error probability, I'll choose (as H_0) the hypothesis which gives the worst consequences when wrong.

Definition:

α is called the significance level or size of the test

Power of a test

Power = probability of rejecting H_0, when H_1 is true, the probability of correctly rejecting H_1.

Equivalent: 1 - probability of incurring into Type II error

$\hookrightarrow$ shown on slides.

General Procedure to Build a Test and a Critical Region.

1. Find estimator $\rightarrow T(X)$ for Θ
2. Define $C \subseteq \mathbb{R}^n$ such that we reject H_0, when $T(x)$ is far from Θ_0

$(x_1, \dots, x_{10})$ realisation of $X_1, \dots, X_{10} \overset{iid}{\sim} N(\theta, (2,5)^2)$

$H_0: \theta = 50$

↳ Average burning rate

$H_1: \theta \neq 50$

known

$X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma_0^2) \rightarrow (x_1, \dots, x_n)$

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

Test for the mean of a gaussian population with known variance

1) Estimator for μ $\rightarrow \bar{X}_n$

2) I need to know the distribution of $\bar{X}_n$ under H_0

$$\bar{X}_n \overset{H_0}{\sim} N\left(\mu_0, \frac{\sigma_0^2}{n}\right)$$

If H_0 is true $\Rightarrow X_1, \dots, X_n \overset{iid}{\sim} N(\mu_0, \sigma_0^2)$

↳ 50 in this case

Test statistic :

$$\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma_0^2 / n}} \overset{H_0}{\sim} N(0,1)$$

→ Distribution of Test Statistic

How to construct Test Region

If H_0 true $\Rightarrow |\bar{X}_n - \mu_0|$ small else if H_1 = true $\Rightarrow |\bar{X}_n - \mu_0|$ is large

So we define C is such a way that

$$C: |\bar{X}_n - \mu_0| \geq k$$

↳ since we are defining the rejection region.

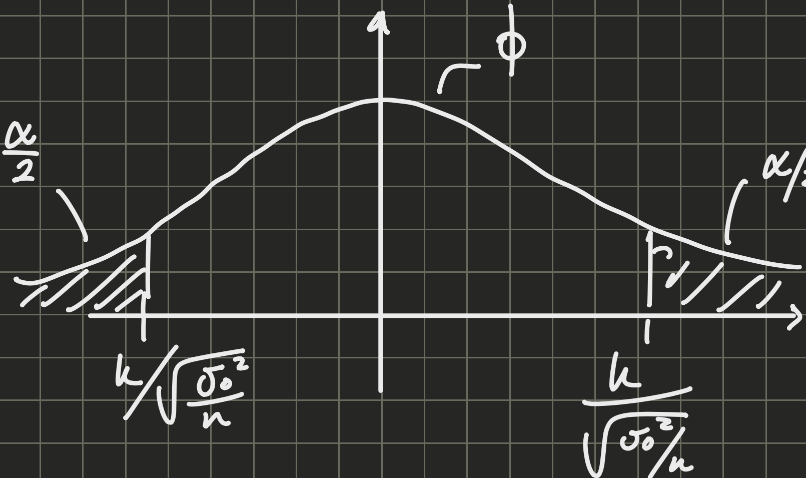
We fix k, such that $P(\text{Type I error}) = \alpha$

↳ size of test

$$\alpha = P_{H_0}\big((X_1, \dots, X_n) \in C\big) = P_{\underbrace{\mu = \mu_0}_{H_0 \text{ is true}}}\big(|\bar{X}_n - \mu_0| \geq k\big)$$

⇒ We can find k as a quantile of the Gaussian distribution

$$k: \quad \alpha = P_{\mu=\mu_0}\left(\underbrace{\frac{|\bar{X}_n - \mu_0|}{\sqrt{\sigma_0^2/n}}}_{\sim N(0,1)} \geq \frac{k}{\sqrt{\sigma_0^2/k}} \right)$$



$$\frac{k}{\sqrt{\frac{\sigma_0^2}{n}}} = z_{1-\frac{\alpha}{2}} \Rightarrow k = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_0^2}{n}}$$

$(x_1, \dots, x_n)$

$$C: |\bar{x}_n - \mu_0| \geq z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_0^2}{n}}$$

$\Rightarrow$ Reject H_0 at significant level α $\iff$ $\bar{x}_n \le \mu_0 - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}}$ or $\bar{x}_n \ge \mu_0 + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}}$

If α decreases, $z_{1-\frac{\alpha}{2}} \uparrow \Rightarrow P(H_0) \uparrow = P(H_1) \downarrow$

General procedure written on slides.

<u>Example</u>

$n=10$ $\qquad X_1, \dots, X_{10} \overset{iid}{\sim} N(\theta, (2,5)^2)$

$(x_1, \dots, x_{10})$ $\qquad x_i \in \mathbb{R}^+$ cm/s

$\bar{x}_{10} = 51,4$ cm/s

$H_0 : \theta = 50$ $\qquad H_1 : \theta \neq 50$

If $\alpha = 5\%$ reject $H_0 \iff \dfrac{|\bar{x}_{10} - 50|}{\underbrace{\frac{2.5}{\sqrt{10}}}_{1,7705}} \ge z_{.975} = 1,96$

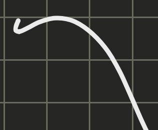
Inequality not satisfied $\Rightarrow$ H_0 not rejected.

$\boxed{Z\text{-test}} \to$ This type of test

z because we use a standard normal as a pivot.

Reject H_0 at significant level $\alpha \iff \dfrac{|\bar{x}_n - \mu_0|}{\sqrt{\frac{\sigma_0^2}{n}}} \ge z_{1-\frac{\alpha}{2}}$

$z_0 := \dfrac{\bar{x}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}$



$\to$ Notation on formulary.

Reject H_0 at significant level α $\iff$ $|z_0| \geq z_{1-\frac{\alpha}{2}}$

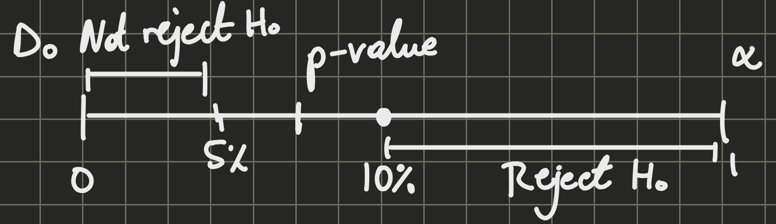
If $\alpha = 10\%$?

Reject H_0 at significant level 10% $\iff$ $1{,}7709 = |z_0| \geq z_{0.95} = 1{,}645$

This is true $\Rightarrow$ reject H_0 at significance level 10%

p-values

If there is a value of α for which we reject H_0, we will reject H_0 at any α greater than that.



p-value = smallest value of significance level which with our data leads to a rejection of H_0.

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

and $X_i \overset{iid}{\sim} N(\mu, \sigma_0^2)$

$\hookrightarrow 0$ always indicates that it is known.

Reject H_0 at significant level α $\iff |z_0| \geq z_{1-\frac{\alpha}{2}} \iff \phi(|z_0|) \geq \underbrace{\phi(z_{1-\frac{\alpha}{2}})}_{1-\frac{\alpha}{2}}$

$$\Rightarrow \alpha \geq \underbrace{2\left[1-\phi(|z_0|)\right]}_{p\text{-value}}$$

$$p\text{-value} = 2\left(1-\phi(|z_0|)\right]$$

↳ observed value of test statistic

Suppose p value $\approx 0 < 0{,}00001$

$\Rightarrow$ very strong evidence against H_0,
because we can reject H_0 and the probability of Type I error is very small.

p-values	Conclusion
if $\alpha < 0{,}001$	Strong evidence against H_0
~	~
~	~
~	~
~	~

Remark: the p-value is not the probability H_0 is false.

How C changes if the hypothesis changes to a unilateral hypothesis

$$\left.\begin{array}{lll} H_0 : \mu = \mu_0 & \text{vs.} & H_1 : \mu > \mu_0 \\ H_0 : \mu \leq \mu_0 & \text{vs.} & H_1 : \mu > \mu_0 \end{array}\right\} \text{Same result}$$

Reject H_0 at significant level $\alpha \iff$ $\bar{x}_n$

(Also slide with p-value)