Hypothesis testing

I Principle :

The question with hypothesis testing is different from the estimation problem. We are not trying to find an approximation of an unknown parameter or trying to provide the user with a confidence interval but we want to decide whether an hypothesis is true or not.

The main difference lies in the hypothesis which assumes some properties of the random variable that we expect to confirm or not thanks to the observations.

Let Θ be the set of parameters such that the statistical model has a parametric distribution is $\{P_\theta, \theta \in \Theta\}$. Let Θ_0 and Θ_1 s.t $\Theta_0 \cap \Theta_1 = \emptyset$ and $\Theta_0 \sqcup \Theta_1 \subseteq \Theta$

Let $H_0 : \{\theta \in \Theta_0\}$ and $H_1 : \{\theta \in \Theta_1\}$ be two hypothesis

def (pure test) A pure test of H_0 against H_1 is a measurable function φ s.t.

- If $\varphi(z) = 0$, we accept H_0 called the null hypothesis (or equivalently we reject H_1 called the alternative hypothesis)
- If $\varphi(z) = 1$, we reject H_0 (or equivalently we accept H_1)

The sets $\mathcal{R} = \{z \in \mathcal{Z} : \varphi(z) = 1\}$ is the reject region of the test φ and $\mathcal{A} = \{z \in \mathcal{Z} : \varphi(z) = 0\}$ is the acceptation region.

Example: Let $Z \sim \mathcal{N}(\theta, 1)$, we only observe one sample of Z

Let $\lambda > 0$, $\Theta_0 = \{0\}$ and $\Theta_1 = \{\lambda\}$

One can propose a very intuitive test : we accept H_0 if the observation is closer to 0 than to λ. This leads us to $\mathcal{R} = \{z \in \mathbb{R} : z > \frac{\lambda}{2}\}$ and

$\mathcal{A} = \{z \in \mathbb{R} : z < \frac{\lambda}{2}\}$

In this case $\varphi(z) = \mathbb{1}_{\{z > \frac{\lambda}{2}\}}$

def: When the set Θ_0 is a singleton (ex $\Theta_0 = \{0\}$) the hypothesis is called simple, otherwise it is called composite.

Example: (survey) The aim of this survey is to know if candidate A is the winner.

We ask randomly people in the street if they voted for A or B. The random variable which models this process is a Bernoulli of parameter θ. The test is therefore to know if $\theta \leq \theta_0 = \frac{1}{2}$ (we assume that $0 \leftrightarrow A$ and $1 \leftrightarrow B$) where θ is therefore the probability of voting for B.

$\Theta_0 = \{ \theta \in [0,1] : \theta \leq \theta_0 \}$ and $\Theta_1 = \{ \theta \in [0,1] : \theta > \theta_0 \}$

We have access to $X_1, \dots, X_n$ samples. The question is not: find $\hat{\theta}_n$ which is close to θ or find $[a_n, b_n]$ a confidence interval st $\theta \in [a_n, b_n]$ with a certain probability But tell if $\theta \leq \theta_0 = \frac{1}{2}$ with an evaluation of the reliability of our answer.

In this example, we can choose the following statistic: for some $t \geq 0$

$$\varphi(X_1, \dots, X_n) = \mathbb{1}_{\{\sum_{i=1}^n X_i \geq t\}} \quad \text{where } X_i = \begin{cases} 1 & \text{if } i \text{ voted for B and} \\ 0 & \text{otherwise.} \end{cases}$$

So the test is $Z = (X_1, \dots, X_n)$: $\varphi(Z) = \mathbb{1}_{\{\sum_{i=1}^n X_i \geq t\}}$

The rejection region is the region for which at least t persons voted for B

$$\mathcal{R} = \{ (x_1, \dots, x_n) \in \{0,1\}^n : \sum_{i=1}^n x_i \geq t \}$$

If the rejection region writes $\mathcal{R} = \{ z \in \mathcal{Z} : T(z) \geq c \}$ where T is a statistic and c a threshold then T is called hypothesis testing statistic and c is a critical value.

As one can see, there are 2 errors that can arise:

① rejecting H_0 although H_0 is true
② Accepting H_0 although H_0 is false

they are not symmetric. Imagine if $H_0 = \{\text{innocent}\}$ it sounds "better" to have the lowest ① even if it creates errors of type ② ... → better in France ...

Now that we can build tests, one have to choose a "good" one. To this purpose we have to introduce the notion of power.

def: The power function of a test φ is given by

$$\beta_\varphi : \begin{array}{l} \Theta \longrightarrow [0,1] \\ \theta \longmapsto \mathbb{P}_\theta(Z \in \mathcal{R}) = \mathbb{P}_\theta(\varphi(Z) = 1) \end{array}$$

We can now use this function to quantify the 2 types of errors.

Sol: the first type risk is given by $\bar{\alpha}(\varphi) = \sup_{\theta \in \Theta_0} \beta_\varphi(\theta) = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta[Z \in \mathcal{R}] = \sup_{\theta \in \Theta_0} \mathbb{E}_\theta[\varphi(Z)]$

the second type risk is given by $1 - \bar{\beta}(\varphi) = 1 - \inf_{\theta \in \Theta_1} \beta_\varphi(\theta)$

$$= \sup_{\theta \in \Theta_1} \mathbb{E}_\theta[1 - \varphi(Z)] = \sup_{\theta \in \Theta_1} \mathbb{P}_\theta[Z \notin \mathcal{R}]$$

Given these definitions, the first type risk measures the maximum probability of rejecting H_0 although H_0 was true. The second type risk measures the maximum probability of accepting H_0 although it was false.

Example: (following)(i) $Z \sim \mathcal{N}(\theta, 1)$. Note $\Theta_0 = \{\theta \leq 0\}$; $\Theta_1 = \{\theta > 0\}$

Let $c_1 > c_2$ and $\varphi_1(z) = 1_{\{z > c_1\}}$; $\varphi_2(z) = 1_{\{z > c_2\}}$

the hypothesis testing statistic is here $T(z) = z$

Associated to φ_1 and φ_2 we have 2 rejection regions

$$\mathcal{R}_1 = \{z \in \mathbb{R} : z > c_1\} \quad \text{and} \quad \mathcal{R}_2 = \{z \in \mathbb{R} : z > c_2\}$$

Because $c_1 > c_2$ $\quad \mathcal{R}_1 \subsetneq \mathcal{R}_2$

therefore $\beta_{\varphi_1}(\theta) = \mathbb{P}_\theta(z \in \mathcal{R}_1) < \mathbb{P}_\theta(z \in \mathcal{R}_2) = \beta_{\varphi_2}(\theta)$

which yields $\bar{\alpha}(\varphi_1) < \bar{\alpha}(\varphi_2)$

Moreover $\mathbb{P}_\theta(z \in \mathcal{R}_1) = \mathbb{P}_\theta(z > c_1)$

$$= 1 - \mathbb{P}_\theta(z \leq c_1)$$

$$= 1 - \Phi(c_1 - \theta)$$ as $z \sim \mathcal{N}(\theta, 1)$ with Φ the cumulative distribution function of $\mathcal{N}(0,1)$

This gives $\bar{\alpha}(\varphi_1) = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(z \in \mathcal{R}_1) = \sup_{\theta \leq 0} (1 - \Phi(c_1 - \theta))$

$$= 1 - \inf_{\theta \leq 0} \Phi(c_1 - \theta) = 1 - \Phi(c_1)$$

$$\bar{\beta}(\varphi_1) = \inf_{\theta \in \Theta_1} \mathbb{P}_\theta(z \in \mathcal{R}_1) = \inf_{\theta > 0} (1 - \Phi(c_1 - \theta))$$

$$= 1 - \sup_{\theta > 0} \Phi(c_1 - \theta) = 1 - \Phi(c_1)$$

and the second type error is $\Phi(c_1)$

Comparing φ_1 and φ_2: $\begin{cases} \bar{\alpha}(\varphi_1) < \bar{\alpha}(\varphi_2) \\ 1 - \bar{\beta}(\varphi_1) > 1 - \bar{\beta}(\varphi_2) \end{cases}$

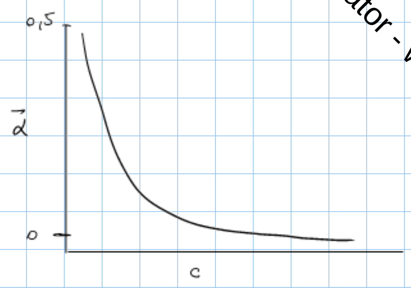
This shows that it can be difficult to choose between φ_1 and φ_2

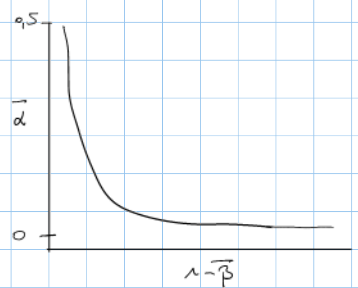
ii) Now with the same example and $\Theta_0 = \{0\}$, $\Theta_1 = \{\vartheta\}$

Taking as above $T(z) = z$ and a critical value c (so that $\varphi(z) = 1_{\{T(z) > c\}}$)

$\bar{\alpha}(\varphi) = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(z \in \mathcal{R}) = \mathbb{P}_{\theta=0}(z > c) = 1 - \mathbb{P}_{\theta=0}(z \leq c) = 1 - \Phi(c)$

and $1 - \bar{\beta}(\varphi) = 1 - \inf_{\theta \in \Theta_1} \mathbb{P}_\theta(z > c) = 1 - \mathbb{P}_\vartheta(z > c) = \mathbb{P}_\vartheta(z \leq c) = \Phi(c - \vartheta)$





$\vartheta = 2$

<u>Example</u> (survey following) $\Theta_0 : \{\theta \leq \vartheta_0\}$

$T(X_1, \dots, X_n) = T(z) = \sum_{i=1}^n X_i = A_n$

As $(X_1, \dots X_n)$ are iid $\mathcal{B}(\vartheta)$, $A_n \sim \mathcal{B}in(\theta, n)$

ie $\mathbb{P}_\theta(A_n = k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$

We want to compute $\bar{\alpha}$ and $1 - \bar{\beta}$ for a test with T as statistic and a critical value c

Let us first prove that $f(\theta) = \mathbb{P}_\theta(A_n \geq c)$ for a given c is a non decreasing function.

Let $U_1, \dots, U_n$ iid $\mathcal{U}([0,1])$ and $0 \leq \theta \leq \vartheta \leq 1$ let $c \in \mathbb{R}$.

$\forall\ i \in \{1, \dots, n\}$ let $X_i = 1_{\{U_i \leq \theta\}}$, $Y_i = 1_{\{U_i \leq \vartheta\}}$

this implies that $X_i \leq Y_i$ and $\sum_{i=1}^n X_i \sim \mathcal{B}in(\theta, n)$; $\sum_{i=1}^n Y_i \sim \mathcal{B}in(\vartheta, n)$

$\sum_{i=1}^n X_i \leq \sum_{i=1}^n Y_i$

therefore $\mathbb{P}_\theta(A_n > c) = \mathbb{P}\left(\sum_{i=1}^n X_i > c\right) \leq \mathbb{P}\left(\sum_{i=1}^n Y_i > c\right) = \mathbb{P}_\vartheta(A_n > c)$

this property enables to compute $\sup_{\theta \leq \vartheta_0} \mathbb{P}_\theta(A_n > c) = \mathbb{P}_{\vartheta_0}(A_n > c)$

which leads to $\bar{\alpha}(\varphi) = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(A_n > c) = \mathbb{P}_{\vartheta_0}(A_n > c)$

$$1 - \bar{\beta}(\varphi) = \sup_{\theta \in \Theta_1} \mathbb{P}_\theta(A_n \leq c)$$
$$= 1 - \inf_{\theta > \vartheta_0} \mathbb{P}_\theta(A_n > c) = 1 - \mathbb{P}_{\vartheta_0}(A_n > c)$$
$$= \mathbb{P}_{\vartheta_0}(A_n \leq c)$$

To construct a test which first type risk is below α, one has to choose a critical value c_α s.t. $P_{\theta_0}(A_n > c_\alpha) \leq \alpha$ (in this example)

This reads: to get a 1st type risk α one has to reduce the rejection region

On the other hand to get a low 2nd type risk, on has to increase the rejection region.

These 2 objectives are contradictory.

This conclusion is unfortunately always true so that one has to find a tradeoff between the 2 risks.

In order to do so, Neyman - Pearson proposed to remove the symmetry of the 2 hypothesis and therefore the 2 errors.

Mathematically this leads to choose a $\alpha \in (0,1)$ small and construct a family of tests Φ st $\forall \varphi \in \Phi$: $\bar{\alpha}(\varphi) \leq \alpha$ this reads φ is of level α.

Now that $\bar{\alpha}(\varphi)$ is bounded, we minimize $1-\bar{\beta}(\varphi)$ over Φ

Example (following)

$$\mathcal{R} = \{z \in \mathbb{R} : z > c\} \quad \text{and} \quad \bar{\alpha}(\varphi) = 1 - \Phi(c)$$

so $\bar{\alpha}(\varphi) \leq \alpha \iff 1 - \Phi(c_\alpha) \leq \alpha \iff \Phi(c_\alpha) \geq 1-\alpha$

that is to say $c_\alpha = z_{1-\alpha}$ the $(1-\alpha)$ quantile of $\mathcal{N}(0,1)$

Note that in this case $1 - \bar{\beta}(\varphi) = \Phi(c_\alpha) \geq 1-\alpha$. One choose $c_\alpha = z_{1-\alpha}$ to have equality.

Remark: If now $H_0 : \theta = \nu$ and $H_1 : \theta = 0$

$$\bar{\alpha}(\varphi) = \sup_{\theta \in \Theta} P_\theta(Z \in \mathcal{R}) = P_\nu(Z \in \mathcal{R}')$$

$$\mathcal{R}' = \{z : z \leq c\} \quad \text{so} \quad \bar{\alpha}(\varphi) = P_\nu(\{z \leq c\}) = \Phi(c - \nu)$$

$$\bar{\alpha}(\varphi) \leq \alpha \iff \Phi(c-\nu) \leq \alpha \iff c - \nu = z_\alpha \iff c = \nu + z_\alpha$$

Noticing that $z_\alpha = -z_{1-\alpha}$ we get: $\mathcal{R}' = \{z \in \mathbb{R} : z \leq \nu - z_{1-\alpha}\}$

In order to understand the disymmetry between H_0 and H_1 we will consider this example in the case $0 \leq \nu \leq z_{1-\alpha}$

De même que dans un procès aux assises, où le principe de présomption d'innocence du prévenu conduit l'avocat général à devoir étayer ses accusations de manière (quasi) irréfutable, le principe de présomption sur H0 conduitàminimiserenprioritéleniveaudutestenimposantqu'ilnedépassepas une valeur fixée. Puis, le test est construit de telle sorte que son erreur de 2ème espèce soit minimale. Cette démarche en deux temps porte le nom de principe de Neyman.

In this setting, one can see 3 different cases:

(i). $\mathcal{R}_0 = \{ z < ? - z_{1-\alpha} \}$ → both hypothesis decide $\theta = 0$

(ii). $\mathcal{R}_1 = \{ z > z_{1-\alpha} \}$ → both hypothesis decide $\theta = ?$

(iii). $\mathcal{R}_2 = \{ ? - z_{1-\alpha} \le z \le z_{1-\alpha} \}$ → the decision depends on the choice of H_0.

these regions correspond to 3 equivalent cases where we were looking for confidence intervals $\hat{I}_\alpha$ of level $1-\alpha$ for θ, knowing that $\theta \in \{0, ?\}$

(i). When $z \in \mathcal{R}_0$, the confidence interval only contains 0

(ii). When $z \in \mathcal{R}_1$, the confidence interval only contains $?$

(iii). when $z \in \mathcal{R}_2$, both 0 and $?$ belong to I_α.

If you are considering the problem with a confidence interval point of view then the last region $\mathcal{R}_2$ would suggest that both values are relevant. this means that just looking at the data is not sufficient to make a decision.

However, looking at this problem with a test point of view requires to introduce some prior information: what is put in H_0. the decision in the Neyman-Pearson's framework is then to choose this prior knowledge that is to say accept H_0.

This example enables to understand the general principle which says that it is always easier to accept the null hypothesis H_0 rather than the alternative one.

In practice therefore the crucial question when constructing a hypothesis test is to choose H_0.

<u>Examples</u>: . For a clinical trial in order to test the efficiency of a new drug, H_0 will be that the drug has <u>no</u> effect → If you are a customer of a compagny, you may not have the same H_0 ...

• Testing the security of a place subject to natural risks, it is better to reject a place although it was <u>not</u> risky that accepting a place of high risk. therefore H_0 is this place is risky.

One can wonder now whether there exists one (at least) best test φ among the family of test of level α. this question again got an answer thanks to Neyman-Pearson.

def: (Neyman's principle): let $\alpha \in (0,1)$ a risk level. the test φ^* is optimal (also called uniformly more powerful UMP) to test $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta_1$ if φ^* satisfies: $\bar{\alpha}(\varphi^*) \leq \alpha$ and $\forall \varphi$ s.t $\bar{\alpha}(\varphi) \leq \alpha$ then

$\forall \theta \in \Theta_1$ $\quad \beta_\theta(\varphi) \leq \beta_\theta(\varphi^*)$ $\quad$ (ie $\mathbb{P}_\theta(Z \in \mathcal{R}_\varphi) \leq \mathbb{P}_\theta(Z \in \mathcal{R}_{\varphi^*})$)

In the case of 2 simple hypothesis, one knows how to get φ^*:

Let $\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \{\theta_1\}$ then:

theorem: (Neyman-Pearson's lemma) In this particular setting, there exist an optimal test φ^* and one can construct explicitly φ^*

II P-value:

What we have seen in the previous section leads to a binary decision: Accept or Reject H_0.

We introduce here a different notion; the p-value, which enables to quantify the level of uncertainty.

Def: (p-value) Consider a statistical model where we have an observation Z (which can be a n-tuple!)

We assume that $Z \sim \mathbb{P}_\theta$ for some $\theta \in \Theta$.

Let $\{\varphi_\alpha, \alpha \in [0,1]\}$ a family of pure tests such that φ_α is of level α

$$\left(\text{ie} \quad \forall \alpha \in (0,1) \quad \bar{\alpha}(\varphi_\alpha) = \sup_{\theta \in \Theta_0} \beta_{\varphi_\alpha}(\theta) = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(Z \in \mathcal{R}_{\varphi_\alpha}) \leq \alpha \right.$$
$$\left. \left(= \sup_{\theta \in \Theta_0} \mathbb{E}_\theta[\varphi_\alpha(Z)] \right) \right)$$

such that $\forall\ 0 \leq \alpha \leq \alpha' < 1$, $\mathcal{R}_{\varphi_\alpha} \subset \mathcal{R}_{\varphi_{\alpha'}}$

the p-value of this family is $\hat{\alpha}(Z) = \inf \{\alpha \in [0,1], Z \in \mathcal{R}_{\varphi_\alpha}\}$

Interpretation: the p-value of the test is $\hat{\alpha}$ st H_0 is always rejected for any $\alpha \geq \hat{\alpha}$ and accepted for $\alpha < \hat{\alpha}$

Example: (following) $Z \sim \mathcal{N}(\theta, 1)$; $H_0 : \theta = 0$; $H_1 : \theta = ?$

$\mathcal{R} = \{Z \in \mathbb{R} : Z > z_{1-\alpha}\}$

therefore, if we observe z, H_0 will be rejected for all α st $z > z_{1-\alpha}$

and accepted for α st $z \leq z_{1-\alpha}$

this writes $\hat{\alpha}(Z)$ satisfies $z = z_{1-\hat{\alpha}(Z)}$ ie $\Phi(z) = 1-\hat{\alpha}(Z)$

$$\Leftrightarrow \hat{\alpha}(Z) = 1-\Phi(z)$$

In practice, accepting or rejecting H_0 is subject to caution (Remember the example where the 2 values were possible and a test would choose H_0). This is highly dependant on α the level of the test, in particular, when α tends to the critical value where things change from H_0 to H_1. However this critical point is very important. This is the p-value.

To interpret the p-value of a test one can say:

p-value	suspicion of rejection		
<0,01	very High suspicion against H_0		
0.01 - 0.05	High	–	–
0.05 - 0,1	weak	–	–
>0,1	no	–	–

← this does not mean you accept H_0 but rather you do not reject it.

There may be two reasons for a p-value to be high:

- H_0 is true
- H_0 is false BUT the test is not powerful (a lot of false positives) and the second type risk is high.

The p-value has an important property: its distribution (under the null Hypothesis) is uniform as soon as P_θ is continuous.

Proposition: Consider a statistical model with an observation Z st $Z \sim P_\theta$, $\theta \in \Theta$

Let H_0 be of the form $H_0 : \theta \in \Theta_0$

Assume that the rejection regions satisfy $\forall \alpha \leq \alpha'$ $\mathcal{R}_\alpha \subset \mathcal{R}_{\alpha'}$

(i) If for all $\alpha \in]0,1[$ $\sup_{\theta \in \Theta_0} P_\theta(Z \in \mathcal{R}_\alpha) \leq \alpha$

then $\hat{\alpha}(Z)$ satisfies: $\forall \theta \in \Theta_0$ $P_\theta(\hat{\alpha}(Z) \leq u) \leq u$ $\forall\ 0 \leq u \leq 1$

(ii) If $\forall \theta \in \Theta_0$ $P_\theta(Z \in \mathcal{R}_\alpha) = \alpha$ $\forall \alpha \in]0,1[$

then $P_\theta(\hat{\alpha}(Z) \leq u) = u$ $\forall\ 0 \leq u \leq 1$

The question that arises now is: Can we compute the p-value in practice? The answer is yes if the test is given by a statistic T and a critical value c.

Proposition: Assume that $\forall \alpha \in [0,1]$, $\varphi_\alpha(Z) = \mathbb{1}_{\{T(Z) \geq c_\alpha\}}$ and $\forall \alpha \in]0,1[$

$\bar{\alpha}(\varphi_\alpha) = \alpha$

Assume that c_α is a non decreasing function on $[0,1]$ then

$$\hat{\alpha}(Z) = \inf\left\{ \alpha \in]0,1[\ : T(Z) \geq c_\alpha \right\}$$

Example: let $(X_1, \dots, X_n)$ iid $\sim \mathcal{N}(\theta, 1)$ with $\theta \in \mathbb{R}$

let $H_0 : \theta \in \mathbb{R}^-$

$\forall \theta \in \mathbb{R}$ $\quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \sim \mathcal{N}(\theta ; \frac{1}{n})$

Proof: $X_i \sim \mathcal{N}(\theta, 1)$ iid $\Rightarrow \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\theta ; n) \Rightarrow \frac{1}{n} \sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\frac{n\theta}{n} ; \frac{n}{n^2}\right)$

Let $\alpha \in [0,1]$ and $\varphi_\alpha(X_1, \dots, X_n) = \mathbb{1}_{\{\bar{X}_n \geq \frac{z_{1-\alpha}}{\sqrt{n}}\}}$

$$\forall \theta \in \mathbb{R}^- : \quad P_\theta\left(\bar{X}_n \geq \frac{z_{1-\alpha}}{\sqrt{n}}\right) = P_\theta\left(\bar{X}_n - \theta \geq \frac{z_{1-\alpha}}{\sqrt{n}} - \theta\right)$$

$$= P_\theta\left(\underbrace{\sqrt{n}(\bar{X}_n - \theta)}_{\sim \mathcal{N}(0,1)\ \forall \theta} \geq z_{1-\alpha} - \sqrt{n}\theta\right)$$

$$= 1 - \Phi(z_{1-\alpha} - \underset{\theta \leq 0}{\sqrt{n}\theta}) \leq 1 - \Phi(z_{1-\alpha}) = \alpha$$

this yields $\quad \sup_{\theta \in \Theta_0} P_\theta\left(\bar{X}_n \geq \frac{z_{1-\alpha}}{\sqrt{n}}\right) = \alpha$

For all $0 \leq \alpha \leq \alpha' < 1$ $\quad z_{1-\alpha} \geq z_{1-\alpha'} \Rightarrow \mathcal{R}_\alpha \subset \mathcal{R}_{\alpha'}$

thus
$$\hat{\alpha}(X_1, \dots, X_n) = \inf\left\{\alpha \in [0,1] : \bar{X}_n \geq \frac{z_{1-\alpha}}{\sqrt{n}}\right\}$$
$$= \inf\left\{\alpha \in [0,1] : \Phi(\sqrt{n}\bar{X}_n) \geq 1 - \alpha\right\}$$
$$= \inf\left\{\alpha \in [0,1] : \alpha \geq 1 - \Phi(\sqrt{n}\bar{X}_n)\right\} = 1 - \Phi(\sqrt{n}\bar{X}_n)$$

III Optimality

In this section, we will extend the first notion of the Neyman-Pearson's principle to better understand and sometimes construct optimal tests.

1) Uniformly more powerful tests

We recall our notations: Let $Z \sim P_\theta$ an observation, $\theta \in \Theta$. Let $(\alpha, \beta) \in]0, 1/2[^2$ and Θ_0 and Θ_1 2 disjoint subsets of Θ.

Let $H_0 : \theta \in \Theta_0$ and $H_1 : \theta \in \Theta_1$

The critical function φ of a test H_0 against H_1 is a measurable function of Z in $\{0,1\}$. The power function is given by:

$$\beta_\varphi(\theta) = E_\theta[\varphi(Z)] = P_\theta(\varphi(Z)=1) = P_\theta(Z \in \mathcal{R}_\varphi)$$

In this section, we restrict our framework to tests of level $\alpha \in [0,1]$: i.e:

$$\sup_{\theta \in \Theta_0} \beta_\varphi(\theta) \leq \alpha$$

We will try to find (when possible) the UMP test φ^* of level α i.e:

$$\sup_{\theta \in \Theta_0} \beta_{\varphi^*}(\theta) \leq \alpha$$

$$\text{and} \quad \forall \varphi \text{ st } \sup_{\theta \in \Theta_0} \beta_\varphi(\theta) \leq \alpha, \quad \forall \theta \in \Theta_1 \quad \beta_{\varphi^*}(\theta) \geq \beta_\varphi(\theta)$$

To be able to construct optimal tests, we will need to make an hypothesis:

the 1^{st} type risk $\bar{\alpha}(\varphi)$ has to be equal to α

$$\left(\bar{\alpha}(\varphi) = \sup_{\theta \in \Theta_0} \beta_\varphi(\theta)\right)$$

We will see that this may not be possible, in particular when dealing with discret random variable. This requires to introduce a more general family of tests: the randomised testings.

Principle: Let $\varphi : \mathcal{Z} \to [0,1]$ a measurable function

Given an observation $Z \in \mathcal{Z}$, we sample a Bernoulli r.v of parameter $\varphi(Z)$.

The null Hypothesis H_0 is chosen if the sample is 0 otherwise H_1 is chosen

Because of the introduction of this Bernoulli r.v the decision is random and this leads to what is called randomised testings

Now the test is not binary ($\varphi = 0$ or 1) but we have 3 cases:

- $\mathcal{Z}_1 = \{z \in \mathcal{Z} : \varphi(z) = 1\}$: Reject H_0
- $\mathcal{Z}_0 = \{z \in \mathcal{Z} : \varphi(z) = 0\}$: Accept H_0
- $\mathcal{Z}_p = \{z \in \mathcal{Z} : \varphi(z) \in]0,1[\}$: the decision depends on the sample (and not only on the observations) $\mathcal{B}(\varphi(z))$

Now that we have defined a randomised testing, one can define as before for pure tests the power function.

def: • the power function of a randomised testing of critical function φ is

$$\beta_\varphi(\theta) = E_\theta[\varphi(Z)] \in [0,1]$$

• the 1st type risk is $\bar{\alpha}(\varphi) = \sup_{\theta \in \Theta_0} \beta_\varphi(\theta)$

• A test is of level $\alpha \in [0,1]$ if $\bar{\alpha}(\varphi) \leq \alpha$

We denote $K_\alpha(\Theta_0)$ the set of tests of level α for the null hypothesis $\theta \in \Theta_0$.

Example: Let $Z \sim \mathcal{P}(\theta)$, $\theta > 0$ (Poisson distribution)

$H_0 : \theta = 1$ $\quad H_1 : \theta > 1$ and we want a level of $\alpha = 0,05$

• A pure test requires to find t_α as big as possible s.t

$$\mathbb{P}_{\theta=1}(Z > t_\alpha) \leq \alpha$$

→ Problem: a Poisson distribution produces integers and

$\mathbb{P}_1(Z > 1) \simeq 0,08$ and $\mathbb{P}_1(Z > 2) \simeq 0,019$

this leads to the choice of $t_\alpha = 2$

The level is $0,05$ but the risk of 1s type is $0,019$.

• Let us now consider a randomised testing:

Let $\gamma \in [0,1]$ and $\varphi_\gamma(Z) = \begin{cases} 1 & \text{if } Z > 2 \quad (\text{ie } Z \geq 3) \\ \gamma & \text{if } Z = 2 \\ 0 & \text{if } Z < 2 \quad (\text{ie } Z \leq 1) \end{cases}$

$\bar{\alpha}(\varphi_\gamma) = E_{\theta=1}[\varphi_\gamma(Z)] = \mathbb{P}_1(Z > 2) + \gamma\, \mathbb{P}_1(Z = 2)$

Choosing $\gamma = \dfrac{\alpha - \mathbb{P}_1(Z>2)}{\mathbb{P}_1(Z=2)}$ implies $\bar{\alpha}(\varphi_\gamma) = \alpha$

For $\alpha = 0,05$: $\gamma \simeq 0,168$.

def: (UMP test) A test with critical function φ is UMP of level α if

$$\varphi \in K_\alpha(\Theta_0) \text{ and } \forall \theta \in \Theta_1 \quad \beta_\varphi(\theta) = \inf_{\psi \in K_\alpha(\Theta_0)} \beta_\psi(\theta)$$

2) Neyman – Pearson's theorem:

this theorem enables to construct UMP tests in some particular cases (although important)

def: Let P_0 and P_1 two probability distributions. We denote by $x(z) = \dfrac{P_1(z)}{P_0(z)}$ the likelihood ratio where we have identified P_0 with its pdf.

Theorem: $\forall \alpha \in [0,1]$, $\exists\, C_\alpha > 0$ and $\gamma_\alpha \in [0,1]$ st

$$\varphi^*(z) = \begin{cases} 1 & \text{if } r(z) > C_\alpha \\ \gamma_\alpha & \text{if } r(z) = C_\alpha \\ 0 & \text{if } r(z) < C_\alpha \end{cases}$$

where $r(z) = \dfrac{P_1(z)}{P_0(z)}$ and we test $H_0: \theta = \theta_0$ and $H_1: \theta = \theta_1$

with $P_{\theta=\theta_0} = P_0$ and $P_{\theta=\theta_1} = P_1$

And φ^* satisfies $E_{\theta=\theta_0}[\varphi^*(z)] = \int \varphi^*(z)\, p_0(z)\, dz = \alpha$

the test associated with the critical function φ^* is UMP of level α and

its power is $\geq \alpha$: $E_{\theta=\theta_1}[\varphi^*(z)] = \int \varphi^*(z)\, p_1(z)\, dz \geq \alpha$

Moreover, if φ^{**} is UMP(α) then for almost every $z \in Z$

$$\varphi^{**}(z) = \begin{cases} 1 & \text{if } r(z) > C_\alpha \\ 0 & \text{if } r(z) < C_\alpha \end{cases}$$

Remark: if $P_{\theta=0}(r(z) = c) = 0 \quad \forall c \geq 0$, one can choose $\gamma = 0$ and come back to pure test.

Example: Let $p_i(z) = \dfrac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\dfrac{1}{2\sigma_i^2}(z-\mu_i)^2\right)$ for $i=1$ and $i=0$.
H_1 H_0

$$r(z) = \frac{\sigma_0}{\sigma_1} \exp\left(-\frac{1}{2\sigma_1^2}(z-\mu_1)^2 + \frac{1}{2\sigma_0^2}(z-\mu_0)^2\right)$$

the critical region is given by: $z \in Z$ st.

$$-\frac{1}{2\sigma_1^2}(z-\mu_1)^2 + \frac{1}{2\sigma_0^2}(z-\mu_0)^2 \geq \log C_\alpha + \frac{1}{2}\log\frac{\sigma_1^2}{\sigma_0^2}$$

Plot the log likelihood ratio for $(\mu_0, \sigma_0^2) = (-1, 1)$ and $(\mu_1, \sigma_1^2) = (1, 0.5)$
and $(\mu_0, \sigma_0^2) = (-1, 1)$ and $(\mu_1, \sigma_1^2) = (1, 1)$

En exercice proposer l'exemple 5.8 p 106 d'Eric
— 5.9 p 107 —
— 5.10 p 107 —

In the case of composite hypothesis the issue of constructing a UMP(α) test is more complex and requires to restrict ourself to exponential families.

Voir si on développe la suite ...

En particulier quid des vitesses de séparation ?