11.Proves d'Hipòtesi

Estadística
Grau en Matemàtiques

Josep A. Sanchez
Dept. Estadística i I.O.(UPC)

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Terminologia

- Una **hipòtesi** és, en general una afirmació referent a un **paràmetre poblacional**
- El valor de θ no és conegut, però coneixem l'espai de paràmetres, $\theta \in \Theta$
- Suposem que $\Theta = \Theta_0 \cup \Theta_1$ amb $\Theta_0 \cap \Theta_1 = \emptyset$ és una partició de l'espai de paràmetres.
- Volem decidir si $\theta \in \Theta_0$ o bé, $\theta \in \Theta_1$
- Indiquem:

$$\begin{cases} H_0 : \theta \in \Theta_0 \\ H_1 : \theta \in \Theta_1 \end{cases}$$

- H_0 s'anomena **Hipòtesi nul · la** i H_1 és la **Hipòtesi alternativa**

Hipòtesis Nul·la i Alternativa

- Ambdues hipòtesis són mutuament exclusives. Només una de les dues pot ser certa
- En base a les dades, es tracta de decidir quina hipòtesi rebutgem
- Podem:

$$\text{rebutjar} H_0 \Leftrightarrow \text{acceptar} H_1$$

$$\text{acceptar} H_0 \Leftrightarrow \text{rebutjar} H_1$$

Test d'hipòtesi

Un **test d'hipòtesi** o **prova d'hipòtesi** és un procediment que especifica:

- per quins valors de mostra acceptem H_0 com a veritat
- per quins valors de mostra rebutgem H_0 i acceptem H_1

El subconjunt del espai mostral (valors de $\underset{\sim}{x}$) pels que rebutgem H_0 s'anomena **regió crítica** o **regió de rebuig**. El subconjunt de l'espai mostral pels quals acceptem la H_0 s'anomena **regió d'acceptació**

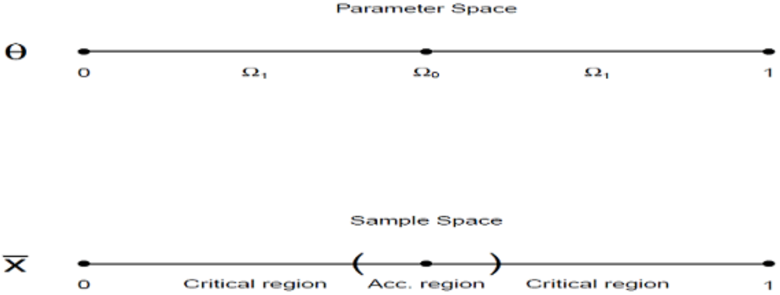
Exemple

Volem testar si una moneda està equilibrada. Llancem la moneda 50 vegades i es pren nota del nombre de cares

$$\begin{cases} H_0 : p = 1/2 \\ H_1 : p \neq 1/2 \end{cases}$$

$$\Theta = [0, 1] \qquad \Theta_0 = \{0.5\} \qquad \Theta_1 = [0, 1] - \{0.5\}$$

$$\Omega = \{0, \dots, 50\}$$



Errors

- La mostra pot portar-nos a creure que la moneda no està equilibrada, quan en realitat sí que ho està. Llavors rebutgem la hipòtesi nul·la quan és certa. Aquesta situació s'anomena **error Tipus I**
- La mostra pot portar-nos a creure que la moneda està equilibrada, quan en realitat no ho està. Llavors no rebutgem la hipòtesi nul·la quan és falsa. Aquesta situació s'anomena **error Tipus II**

Posibilitats

		Test	
		Acceptem H_0	Rebutgem H_0
Realitat	H_0	Correcte	Error Tipus I
	H_1	Error Tipus II	Correcte

α i β

- $\alpha \equiv P(\text{Error tipus I}) = P(\text{Rebutjar } H_0 | H_0 \text{ es certa})$
- $\beta \equiv P(\text{Error tipus II}) = P(\text{Acceptar } H_0 | H_0 \text{ es falsa})$
- Idealment, volem que ambdues probabilitats siguin petites

Exemple

- Un productor de cervesa envasa el seu producte en ampolles de vidre
- Se suposa que les ampolles contenen 500ml del producte
- Posibilitats del procès de producció:
  1. Les ampolles s'omplen, en mitjana, amb 500ml de líquid
  2. Les ampolles s'omplen, en mitjana, amb més o menys de 500ml de líquid
- L'encarregat ha de decidir, a partir d'una mostra d'ampolles, si. . .
  - la producció ha de continuar (en el cas 1)
  - s'ha de parar per reajustar la maquinària (en el cas 2)

En termes estadístics

- Tota la producció d'ampolles és la **població**
- Amb una **mostra aleatòria simple** d'ampolles, volem fer inferència de la població
- Sigui μ la mitjana poblacional del contingut de les ampolles
- Un **test d'hipòtesi** estadístic és una afirmació referent a μ

En termes estadístics

Els estats de la producció es poden formular en forma d'hipòtesis:

- **Hipòtesi Nul · la (H_0)**: Les ampolles s'omplen en mitjana amb 500ml de líquid
- **Hipòtesi Alternativa (H_1)**: Les ampolles no s'omplen en mitjana amb 500ml de líquid (la mitjana és inferior o superior a 500ml)

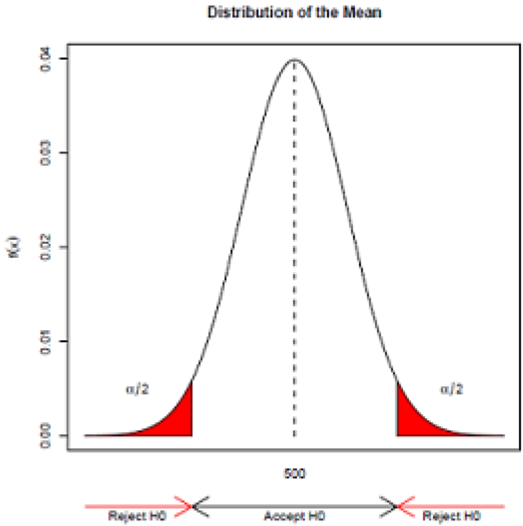
Formalment,

$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu \neq 500 \end{cases}$$

Com prenem la decisió?

- Tenim una m.a.s. de n ampolles i calculem $\bar{X}$, la mitjana mostral
- Si $\bar{X} >> 500$ o $\bar{X} << 500$ rebutgem H_0
- Si $\bar{X} \approx 500$ no rebutgem H_0
- Sabem que $\bar{X} \sim N(\mu, \sigma^2/n)$ (Teorema central del límit)

Com prenem la decisió?



Com prenem la decisió?

- Quines conseqüències té rebutjar la hipòtesi nul · la?
- Quines conseqüències té acceptar la hipòtesi nul · la?
- Consideracions respecte als **costos** de cada decisió ha de determinar el valor d'α
- Un cop s'ha escollit el valor d'α, es determina un **valor crític** per a la mitjana de la mostra (el nostre **estadístic de test** en aquest cas)
- Si el nostre estadístic de test és més gran que el valor crític, rebutgem H_0. En cas contrari no rebutgem H_0 (**Regla de decisió**)

L'estadístic de test

- Hem fet servir la mitjana com a estadístic de test
- Si estandarditzem la mitjana tindrem un estadístic que té distribució N(0,1)
- Amb la distribució N(0,1) podem trobar valors crítics (quantiles d'una probabilitat especificada)
- Si prenem 50 ampolles ($n = 50$) i suposem que coneixem $\sigma^2 = 10^2$, llavors, si el procés està centrat en el nominal ($\mu = 50$)

$$\bar{X} \sim N(500, 10^2/50) = N(500, 2)$$

  Per tant,

$$Z = \frac{\bar{X} - 500}{\sqrt{2}} \sim N(0, 1)$$

- Si fem servir $\alpha = 0.05$, llavors $Z_{1-\alpha/2} = Z_{0.975} = 1.96$ i rebutjaríem H_0 si $|Z| > 1.96$

Example

En l'exemple del procés d'envasat, suposem que s'agafen $n = 75$ ampolles i la seva mitjana mostral dóna $\bar{X} = 495$. Resolem el test:

$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu \neq 500 \end{cases}$$

amb nivell de significació $\alpha = 0.05$. Suposem coneguda la variància del procés $\sigma^2 = 10^2$

$$Z = \frac{495 - 500}{10/\sqrt{75}} = -\frac{\sqrt{75}}{2} = -4.33$$

com

$$|Z| > 1.96$$

rebutgem H_0

- Conclusió pràctica: hi ha evidències estadístiques significatives que permeten afirmar que les ampolles no s'omplen d'acord al valor nominal de 500ml

El p-valor (o valor de probabilitat)

- Ens proporciona un mètode alternatiu per resoldre el test
- Definició: El **p-valor** és la probabilitat d'observar un estadístic de test tan extrem o més que el valor que hem obtingut amb la mostra
- El p-valor quantifica quanta evidència hi ha en la mostra contra la hipòtesi nul · la
- El p-valor es compara amb el nivell de significació prefixat (α)
  - Si el p-valor$< \alpha$, llavors **rebutgem** H_0
  - Si el p-valor$\geq \alpha$, llavors **no rebutgem** H_0

Exemple revisitat

Variable: X: contingut en ml. de les ampolles $\sim N(\mu, \sigma^2 = 10^2)$

Test:

$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu \neq 500 \end{cases}$$

Descriptiva de la mostra: $n = 75 \qquad \bar{X} = 495$

Estadístic de prova: $z = \frac{495-500}{10/\sqrt{75}} = -\frac{\sqrt{75}}{2} = -4.33$

Si $Z \sim N(0,1)$,

$$P(|Z| > z) = P(Z < -4.33) + P(Z \geq 4.33) = 0.00002$$

i com p-value$=0.00002<0.05=\alpha$ rebutgem H_0

Test Z per a la mitjana d'una mostra (σ^2 coneguda)

Variable: $X \sim N(\mu, \sigma^2)$ amb σ^2 coneguda

Resolem el test d'hipòtesi:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

calculant l'estadístic de test

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

i es compara amb la distribució N(0,1)

- Si H_0 és certa, $Z \sim N(0, 1)$ i podem catalogar el valor de l'estadístic obtingut com a plausible (acceptem H_0) o estrany (rebutgem H_0)

Test t per a la mitjana d'una mostra (σ^2 desconeguda)

Variable: $X \sim N(\mu, \sigma^2)$ amb σ^2 desconeguda

Resolem el test d'hipòtesi:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

calculant l'estadístic de test

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

i es compara amb la distribució t_{n-1}

- Si H_0 és certa, $t \sim t_{n-1}$ i podem catalogar el valor de l'estadístic obtingut com a plausible (acceptem H_0) o estrany (rebutgem H_0)

Exemple

En el cas del procès d'envasat, un dia s'obté una mostra de $n = 100$ ampolles, amb una mitjana mostral de 490ml i una desviació estandard mostral de 15ml.

Resolem el test d'hipòtesi:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

calculant l'estadístic de test

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{490 - 500}{15/\sqrt{100}} = -6.67$$

$$P(|t_{99}| > 6.67) < 0.0001$$

Llavors, rebutgem H_0

Procediment per a un test d'hipòtesi

- Escollir un disseny i un estadístic de test
- Formular les hipòtesis, nul · la i alternativa
- Determinar el nivell de significació α
- Especificar la distribució de l'estadístic de test sota la hipòtesi nul · la i establir les assumpcions fetes
- A partir de la mostra, calcular l'estadístic de test i el seu p-valor per determinar la decisió
- Opcionalment, donar un interval de confiança com a suplement del test

Procediment per a la decisió

- Després d'establir les hipòtesis, amb la mostra es calcula l'estadistic de prova. Per la teoria coneixem: la distribució de referència per l'estadístic de prova si H_0 és certa. Per decidir el test, són equivalents els següents procediments:
  1. Determinar la regió crítica pel nivell de significació fixat: si la mostra pertany a la regió crítica, rebutgem H_0, en cas contrari no rebutgem H_0
  2. Calcular el p-valor de l'estadístic: si el p-valor és inferior al valor de significació α prefixat, rebutgem H_0, en cas contrari no rebutgem H_0

Exemple: amb la regió crítica

En el cas del procès d'envasat, un dia s'obté una mostra de $n = 100$ ampolles, amb una mitjana mostral de 490ml i una desviació estandard mostral de 15ml.

Sabem que la distribució de l'estadístic de prova sota H_0 és:

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

Fixat el nivell de significació $\alpha = 0.05$ llavors la regió crítica serà:

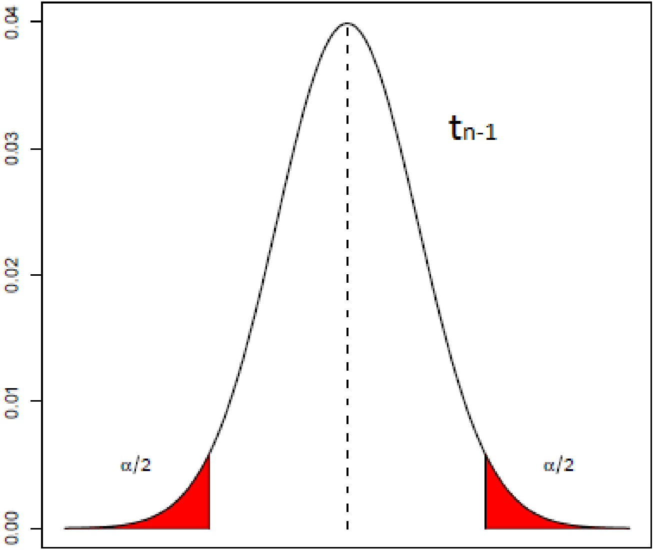
$$C = \{\utilde{X} : |t(\utilde{X})| \geq t_{n-1,0.975}\} = \{\utilde{X} : |t(\utilde{X})| \geq 1.9842\}$$

calculant l'estadístic de test

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{490 - 500}{15/\sqrt{100}} = -6.67$$

i com $|t| > 1.9842$, llavors $\utilde{X} \in C$ i decidim rebutjar H_0

Exemple: amb la regió crítica



Exemple: amb el p-valor

En el cas del procès d'envasat, un dia s'obté una mostra de $n = 100$ ampolles, amb una mitjana mostral de 490ml i una desviació estandard mostral de 15ml.

Sabem que la distribució de l'estadístic de prova sota H_0 és:

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

calculant l'estadístic de test

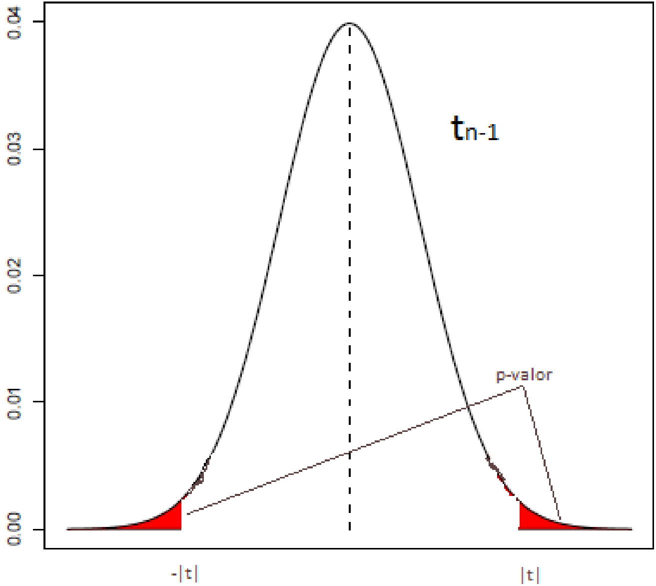
$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{490 - 500}{15/\sqrt{100}} = -6.67$$

la probabilitat d'optenir valors més estranys a la hipòteis nul · la que la que hem trobat és:

$$\text{p-valor} = P(|t_{n-1}| \geq |t|) = P(|t_{n-1}| \geq 6.67) < 0.0001$$

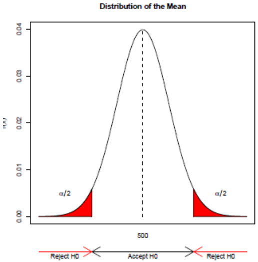
i com p-valor $< \alpha = 0.05$, llavors decidim rebutjar H_0

Exemple: amb el p-valor

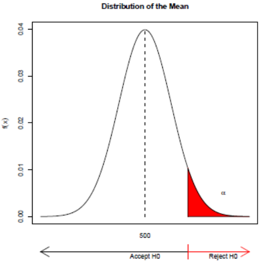


Test bilateral vs.unilateral

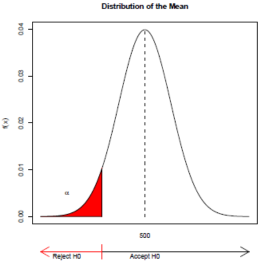
$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu \neq 500 \end{cases}$$



$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu > 500 \end{cases}$$



$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu < 500 \end{cases}$$



Test Unilateral

- Un encarregat del procés d'envasat està preocupat per les queixes dels clients que indiquen que l'empresa omple les ampolles amb menys líquid del que indica l'etiqueta (500ml)
- S'agafa una mostra de $n = 25$ ampolles i es troba que $\bar{X} = 495$ i $S = 15$
- Test:

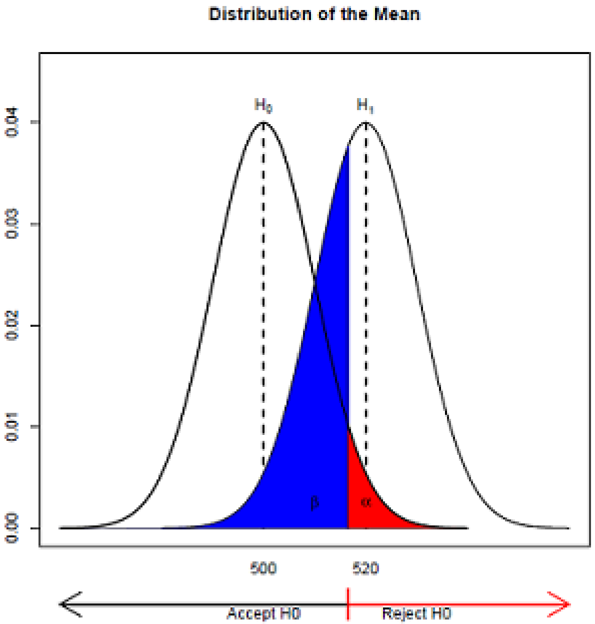
$$\begin{cases} H_0 : \mu = 500 \\ H_1 : \mu < 500 \end{cases}$$

- Estadístic de prova: $t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{495 - 500}{15/\sqrt{25}} = -1.67 \sim_{H_0} t_{n-1}$
    1. Regió crítica: $(\alpha = 0.05)$:
       $C = \{\underset{\sim}{X} : t(\underset{\sim}{X}) \leq t_{24,0.05}\} = \{\underset{\sim}{X} : t(\underset{\sim}{X}) \leq -1.71\}$
    2. p-valor: p-valor $= P(t_{n-1} \leq t) = P(t_{24} \geq -1.67) = 0.0543$

Potència del test

- Es construeix la regió crítica imposant que l'error de Tipus I sigui $\leq \alpha$
- De què depèn el valor de l'error de Tipus II (β)
- En el cas del test per a la mitjana d'una mostra (test Z), si coneixem el valor de μ_1 podem calcular el valor de β

α i β


Distribution of the Mean
H0
H1
0.04
0.03
0.02
0.01
0.00
β
α
500
520
Accept H0
Reject H0

Potència

Definició: La **potència** d'un test és la probabilitat de rebutjar la hipòtesi nul · la

$$\text{Potència} = 1 - \beta$$

- Quins factors afecten la potència del tets per a la mitjana d'una mostra?
    - Nivell de significació (α)
    - Mida mostral (*n*)
    - Veritable valor de μ

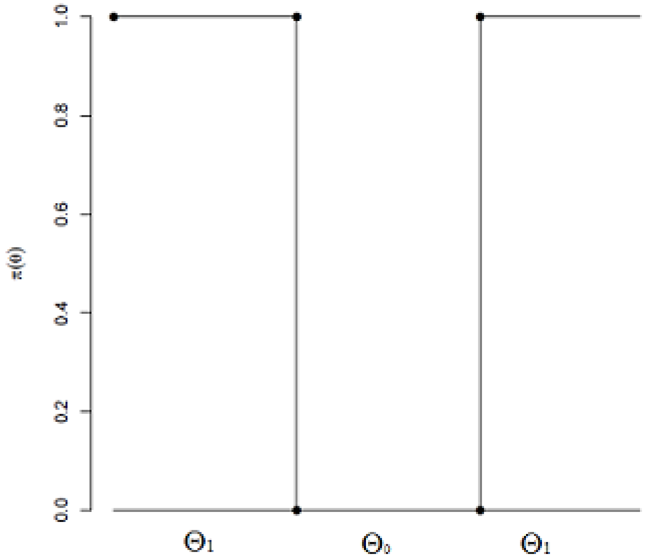
Funció de Potència

Definició: La **funció de potència** d'un test, $\pi(\theta)$, definida a l'espai de paràmetres, dóna la probabilitat de rebutjar H_0 per a cada valor de θ

$$\pi(\theta) = \begin{cases} \alpha = P(\text{Error Tipus I}) & \theta \in \Theta_0 \\ 1 - \beta = 1 - P(\text{Error Tipus II}) & \theta \in \Theta_1 \end{cases}$$

Serveix per comparar procediments de decisió alternatius sobre un mateix test d'hipòtesi

Funció de Potència ideal:



Mida d'un test i nivell de significació

Definició: La **mida d'un test** α és el suprem de la funció de potència en la partició de l'espai de paràmetres corresponent a H_0 (Θ_0), és a dir, la probabilitat màximal d'error de tipus I que podem cometre amb el procediment de decisió per tots els valors de $\theta \in \Theta_0$

$$\alpha = \sup_{\theta \in \Theta_0} \pi(\theta)$$

- Habitualment, s'especifica un límit superior (α_0) per a l'error de Tipus I. Aquest límit s'anomena **nivell de significació** del test. Aquest és un valor α_0 tal que:

$$\sup_{\theta \in \Theta_0} \pi(\theta) \leq \alpha_0$$

- Un cop fixat el valor del nivell de significació, es consideren test de mida $\alpha \leq \alpha_0$

- Habitualment, $\alpha_0 = 0.05$ ó 0.01

Test més potent

Definició: Un test **més potent** de mida α és el que minimitza β, la probabilitat d'error de Tipus II

Exemple de funció de potència

Sigui $X_1, \ldots, X_n$ m.a.s, on $X_i \sim U[0,\theta]$ amb $\theta > 0$

Volem resoldre el test:

$$\begin{cases} H_0 : 3 \leq \theta \leq 4 \\ H_1 : \theta < 3 \text{ ó } \theta > 4 \end{cases}$$

L'estimador ML és $\hat{\theta}_{ML} = X_{(n)}$. Per n gran, el valor de l'estimador estarà a prop del veritable valor de θ (consistència). Escollirem no rebutjar H_0 si el màxim de la mostra està entre 2.9 i 4:

- Regió crítica: $C = \{\underset{\sim}{X} \in [0,\theta]^n : X_{(n)} < 2.9 \text{ o } X_{(n)} > 4\}$
- Funció de potència del test: $\pi(\theta) = P(X_{(n)} < 2.9|\theta) + P(X_{(n)} > 4|\theta)$

$$\pi(\theta) = \begin{cases} 1 + 0 & \theta < 2.9 \\ (2.9/\theta)^n + 0 & 2.9 < \theta < 4 \\ (2.9/\theta)^n + (1 - (4/\theta)^n) & 4 < \theta \end{cases}$$

- La mida d'aquest procediment per n=10 és

$$\alpha = \sup_{3<\theta<4} \pi(\theta) = (2.9/3)^{10} = 0.71$$

Exemple de funció de potència