

### Métodos no paramétricos y de libre distribución

La mayoría de los autores usan ambos términos como sinónimos, aunque otros hacen una distinción.

Bradley (1968): "Un test no paramétrico no testea hipótesis sobre el valor de los parámetros en una cierta distribución, mientras que un test de distribución libre no hace hipótesis sobre la forma precisa de la distribución".

Ejemplos: Consideremos el caso de dos muestras aleatorias independientes:  $X_1, X_2, \dots, X_n$  e  $Y_1, Y_2, \dots, Y_m$ .

a) Test t:  $X_i \sim N(\mu_X, \sigma^2)$ ,  $i=1, \dots, n$  e  $Y_j \sim N(\mu_Y, \sigma^2)$ ,  $j=1, \dots, m$

$$\text{Estadístico del test: } T = \frac{\bar{X} - \bar{Y}}{s} \sqrt{\frac{nm}{n+m}}$$

$$\text{donde } s^2 = \frac{\sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2}{n+m-2}$$

Las hipótesis a testear podrían ser, por ejemplo:

$$H_0: \mu_X = \mu_Y \quad \text{vs} \quad H_1: \mu_X \neq \mu_Y$$

Este es un test paramétrico.

b) Test de Kolmogorov-Smirnov:  $X_i \sim F_X$  e  $Y_j \sim F_Y$ , y las hipótesis a testear son

$$H_0: F_X = F_Y \quad \text{vs} \quad H_1: F_X \neq F_Y$$

El estadístico del test es  $D_{nm} = \|F_n - F_m\| = \sup_x |F_n(x) - F_m(x)|$ , siendo  $F_n$  y  $F_m$  las respectivas funciones de distribución empíricas, es decir

$$F_n(x) = \frac{\text{card}\{X_i / X_i \leq x\}}{n} \quad F_m(y) = \frac{\text{card}\{Y_j / Y_j \leq y\}}{m}$$

Este es un test no paramétrico y de distribución libre.

c) Test de rangos de Mann-Whitney:  $X_i \sim F(x)$  e  $Y_j \sim F(x-\Delta)$ , y las hipótesis a testear son

$$H_0: \Delta = 0 \quad \text{vs} \quad H_1: \Delta \neq 0$$

Se basa en el estadístico:

$$U = \sum_{i=1}^m R(Y_i)$$

siendo  $R(Y_i)$  el rango correspondiente a  $Y_i$  en la muestra total ordenada.

En este caso, la hipótesis es paramétrica pero el test es de distribución libre.

## Repaso de algunos conceptos de inferencia estadística

**Población:** conjunto de individuos, objetos, etc, que deseamos investigar.

**Muestra:** subconjunto de la población.

**Muestra aleatoria:** muestra obtenida a través de un mecanismo tal que todas las muestras posibles tienen igual probabilidad de ser seleccionadas (población finita). En general, puede pensarse como un conjunto de v.a. independientes e idénticamente distribuidas:  $X_1, \dots, X_n$ .

**Estadístico:** función de la muestra  $T : \mathfrak{R}^n \rightarrow \mathfrak{R}$

Ejemplos:

a)  $\bar{X}$

b)  $med(X_1, \dots, X_n) = \tilde{X}$

c)  $min(X_1, \dots, X_n)$

d)  $X^{(k)}$ : k-ésimo estadístico de orden (ocupa el k-ésimo lugar en la muestra ordenada)

e)  $R_j$ : rango de  $X_j$ . Posición de  $X_j$  en la muestra ordenada.

$$R_j = R(X_j) = \text{card} \{ i / X_i \leq X_j \}$$

Si no hay empates,

$$R_j = R(X_j) \Leftrightarrow X^{(R_j)} = X_j$$

f)  $D_j$ : antirango. Indica qué observación ocupa el j-ésimo lugar en la muestra ordenada.

$$X_{D_j} = X^{(j)}$$

Ejemplo: Sean  $X_1 = 93$ ,  $X_2 = 76$  y  $X_3 = 85$ .

$$X^{(1)} = 76 = X_2$$

$$X^{(2)} = 85 = X_3$$

$$X^{(3)} = 93 = X_1$$

Entonces,

$$R_1 = 3$$

$$R_2 = 1$$

$$R_3 = 2$$

$$D_1 = 2$$

$$D_2 = 3$$

$$D_3 = 1$$

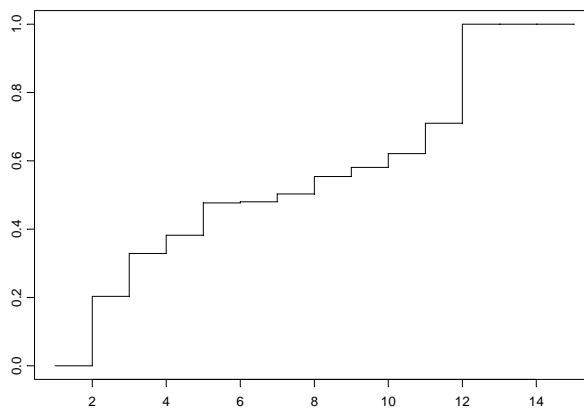
**Estimación:** Un estimador es una función de la muestra (estadístico) que provee un valor aproximado de un parámetro o característica desconocida.

Recordaremos un estimador “particular”.

Función de distribución empírica: Sea  $X_1, \dots, X_n$  una m.a.

$$F_n(x) = \frac{\text{card} \{X_i / X_i \leq x\}}{n}$$

$F_n(x)$  es una función escalera con saltos en los  $X_i$ 's.



Es la función de distribución de una v.a. discreta que toma valores  $X_1, X_2, \dots, X_n$  con probabilidad  $1/n$ . Si llamamos  $W$  a esa variable

$$E(W) = \sum_{i=1}^n X_i / n = \bar{X}$$

Teorema de Glivenko-Cantelli:  $\sup_x |F_n(x) - F(x)| \xrightarrow{pp} 0$ , siendo  $F$  la función de distribución de las  $X_i$ 's.

Parece natural entonces usar características de  $F_n$  para estimar correspondientes características de  $F$ .

Ejemplos:

- a)  $\bar{X}$  para estimar  $\mu = E(X)$
- b)  $Q_p$  (percentil p muestral) para estimar  $x_p$  (percentil p poblacional)

¿Cómo se definen  $x_p$  y  $Q_p$ ?

Sea  $0 < p < 1$ ,  $x_p$  es el valor tal que

$$\begin{cases} P(X \leq x_p) \geq p \\ P(X \geq x_p) \geq 1-p \end{cases} \quad \text{o} \quad \begin{cases} P(X > x_p) \leq 1-p \\ P(X < x_p) \leq p \end{cases}$$

Dada una muestra  $X_1, \dots, X_n$ , el percentil o cuartil  $Q_p$ , es el valor tal que

$$\begin{cases} \frac{\text{card} \{i / X_i > Q_p\}}{n} \leq 1-p \\ \frac{\text{card} \{i / X_i < Q_p\}}{n} \leq p \end{cases}$$

**Intervalos (o regiones) de confianza:** Denotemos  $\vec{X} = (X_1, \dots, X_n)$  donde  $X_i$  tienen distribución  $F(x; \theta)$ ,  $S(\vec{X})$  es una región de confianza para  $\theta$  de nivel  $1 - \alpha$  ( $0 < \alpha < 1$ ) si

$$P_\theta(\theta \in S(\vec{X})) = 1 - \alpha \quad \forall \theta$$

Ejemplos: a)  $X_1, \dots, X_n$ , m.a. de una distribución  $N(\mu, \sigma^2)$

$$\sqrt{n} \frac{\bar{X} - \mu}{s} \sim t_{n-1}$$

y un intervalo de confianza de nivel  $1 - \alpha$  para  $\mu$  es de la forma

$$\left[ \bar{X} - \frac{s}{\sqrt{n}} t_{n-1, \alpha/2}, \bar{X} + \frac{s}{\sqrt{n}} t_{n-1, \alpha/2} \right]$$

b)  $X_1, \dots, X_n$ , m.a. de una distribución exponencial de parámetro  $\lambda$

$$2\lambda \sum_{i=1}^n X_i \sim \chi_{2n}^2$$

y un intervalo de confianza de nivel  $1 - \alpha$  para  $\lambda$  es de la forma

$$\left[ \frac{\chi_{2n, 1-\alpha/2}^2}{2n\bar{X}}, \frac{\chi_{2n, \alpha/2}^2}{2n\bar{X}} \right]$$

Definición: Una familia  $\{S_n(\bar{X})\}$  es una familia de regiones de confianza para  $\theta$  de nivel asintótico  $1 - \alpha$  ( $0 < \alpha < 1$ ) si

$$P_\theta(\theta \in S_n(\bar{X})) \xrightarrow{n \rightarrow \infty} 1 - \alpha \quad \forall \theta$$

Ejemplos: a)  $X_1, \dots, X_n$ , m.a. de una distribución  $Bi(1, p)$ , entonces, por el TCL,

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$$

$$P_\theta \left( \left| \frac{\sum X_i - np}{\sqrt{np(1-p)}} \right| \leq z_{\alpha/2} \right) \cong 1 - \alpha$$

y una región de confianza para  $p$  de nivel asintótico  $1 - \alpha$  es

$$\left\{ p \mid \left| \frac{\sum X_i - np}{\sqrt{np(1-p)}} \right| \leq z_{\alpha/2} \right\}$$

c)  $X_1, \dots, X_n$ , m.a. de una distribución exponencial de parámetro  $\lambda$ , entonces, por el TCL

$$\frac{\sum_{i=1}^n X_i - \frac{n}{\lambda}}{\sqrt{\frac{n}{\lambda^2}}} \xrightarrow{d} N(0,1)$$

y de aquí se puede deducir un intervalo de confianza para  $\lambda$  de nivel asintótico  $1 - \alpha$ .  
(Ejercicio).

**Tests de hipótesis:** Son procedimientos formales para decidir, a partir de una muestra, entre dos opciones o hipótesis.

- Hipótesis nula ( $H_0$ ): Ésta es la hipótesis que se testea y, si los datos muestran fuerte evidencia en su contra, es rechazada.
- Hipótesis alternativa o hipótesis del investigador ( $H_1$  o  $H_a$ )

Para llevar a cabo un test, se selecciona un estadístico  $T$  y una regla de decisión basada en él (región de rechazo).

Ejemplo: Una máquina produce piezas y se considera que funciona adecuadamente si el porcentaje de piezas defectuosas es menor que 5%. Si se supone que cada pieza tiene probabilidad  $p$  de ser defectuosa, se desea testear

$$H_0 : p \leq 0.05 \quad \text{vs} \quad H_1 : p > 0.05$$

Parece razonable rechazar  $H_0$  si hay demasiadas piezas defectuosas en la muestra. Sea  $T$  el número de piezas defectuosas en la muestra. Si se toma una muestra de 10 piezas de la “amplia” producción total, podemos suponer  $T \sim \text{Bi}(10,p)$ .

Si decidimos rechazar  $H_0$  cuando  $T > 2$ , ésta es nuestra región de rechazo o región crítica.

Definición: Una hipótesis es *simple* si la suposición de que es cierta conduce a una sola función de probabilidad. De lo contrario es *compuesta*.

Definición: Supongamos que  $\bar{X} = (X_1, \dots, X_n)$  es un vector aleatorio cuya función de distribución depende de un parámetro  $\theta \in \Theta \subset \mathfrak{R}^p$ . Sean  $\Theta_0$  y  $\Theta_1$  tales que  $\Theta_0 \cap \Theta_1 = \emptyset$ . Un test  $\Phi$  es una función del estadístico, en base a la cual se tomará la decisión entre

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

tal que  $\Phi : T \rightarrow [0,1]$ , siendo  $\Phi(T)$  = probabilidad de rechazar  $H_0$ , para ese valor del estadístico  $T$ .

Si  $\Phi(T)$  toma sólo valores 0 y 1 se trata de un *test no aleatorizado*.

Dado un test no aleatorizado, la región crítica es  $\{t / \Phi(t) = 1\}$ , o sea el conjunto de valores del estadístico para los cuáles se rechaza  $H_0$ .

Tipos de errores:

|          |                  | Realidad     |               |
|----------|------------------|--------------|---------------|
|          |                  | $H_0$ cierta | $H_0$ falsa   |
| Decisión | Rechazo $H_0$    | Error tipo I | OK!!          |
|          | No rechazo $H_0$ | OK!!         | Error tipo II |

Definición: Supongamos que se utiliza el test  $\Phi(T)$  para testear

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

se denomina **función de potencia** del test  $\Phi$  a la función

$$\pi(\theta) = E_{\theta}(\Phi(T)) \quad \forall \theta$$

Observemos que, si el test es no aleatorizado,  $E_{\theta}(\Phi(T)) = P_{\theta}(\Phi(T) = 1)$ , o sea, que  $\pi(\theta)$  es la probabilidad de rechazar  $H_0$  cuando el valor del parámetro es  $\theta$ .

Si  $\theta \in \Theta_0$  (especificado por  $H_0$ ) entonces  $\pi(\theta) = P_{\theta}$ (error tipo I).

Si  $\theta \in \Theta_1$  (especificado por  $H_1$ ) entonces  $\pi(\theta) = 1 - P_{\theta}$ (error tipo II).

Se dice que el test tiene **nivel de significación**  $\alpha$  si

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

**Nivel crítico o p-valor:** es el menor nivel de significación para el que, dado un conjunto de observaciones, el test rechazará  $H_0$ .

Ejemplo: Un fabricante de sistemas de protección contra incendios asegura que el verdadero promedio de temperatura de activación del sistema es 130°F. Al probar una muestra de 9 sistemas se obtiene un promedio muestral de activación de 131.08°F. Si la distribución de las temperaturas de activación es  $N(\mu, (1.5^{\circ}\text{F})^2)$ . ¿Contradicen los datos la afirmación del fabricante a nivel de significación  $\alpha=0.01$ ?

$$H_0 : \mu = 130 \quad \text{vs} \quad H_1 : \mu \neq 130$$

$X_1, \dots, X_9$  m.a.  $X_i \sim N(\mu, (1.5^{\circ}\text{F})^2)$ . El estadístico del test será

$$T = \frac{\bar{X} - 130}{1.5/\sqrt{9}} = \frac{\bar{X} - 130}{0.5}$$

y la zona de rechazo para un nivel de significación  $\alpha=0.01$  es

$$\left| \frac{\bar{X} - 130}{0.5} \right| \geq 2.58$$

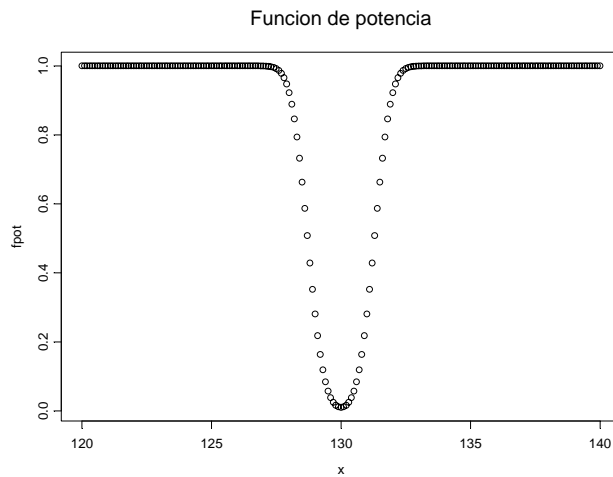
El valor observado de  $T = 2.16$  entonces no se rechaza  $H_0$  a ese nivel.

$$p\text{-valor} = 2 P(T \geq 2.16) = 0.0308$$

Hallemos la función de potencia del test:

$$\pi(\mu) = P_{\mu} \left( \left| \frac{\bar{X} - 130}{0.5} \right| \geq 2.58 \right) = 1 - \left( \left| \frac{\bar{X} - 130}{0.5} \right| \leq 2.58 \right) =$$

$$1 - \Phi \left( 2.58 + \frac{130 - \mu}{0.5} \right) + \Phi \left( -2.58 + \frac{130 - \mu}{0.5} \right)$$



Observemos que  $\pi(130) = 0.01 = \alpha$  y

$$\lim_{\mu \rightarrow \infty} \pi(\mu) = 1$$

$$\lim_{\mu \rightarrow -\infty} \pi(\mu) = 1$$

Es posible hallar  $n$  tal que la potencia sea tan próxima a 1 como se desee para una alternativa fija pues

$$\text{Si } \mu \in \Theta_1 \text{ fijo} \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \pi(\mu) = 1$$

Por ejemplo, con el test anterior,  $\pi(131) = 0.281$ , y por lo tanto la probabilidad de error tipo II es

$$\beta(131) = 1 - \pi(131) = 0.719$$

¿Cuál debe ser el  $n$  para que  $\pi(131) > 0.95$ ?

Para  $n$  variable, la función de potencia del test es:

$$\pi(\mu) = 1 - \Phi\left(2.58 + \frac{130 - \mu}{1.5/\sqrt{n}}\right) + \Phi\left(-2.58 + \frac{130 - \mu}{1.5/\sqrt{n}}\right)$$

Entonces

$$\pi(131) = 1 - \Phi\left(2.58 + \frac{130 - 131}{1.5/\sqrt{n}}\right) + \Phi\left(-2.58 + \frac{130 - 131}{1.5/\sqrt{n}}\right)$$

y para que  $\pi(131) > 0.95$ , debe ser  $n \geq 41$ .



**Algunas propiedades de los tests de hipótesis:**

a) Test uniformemente más potentes: Supongamos que el test  $\Phi$  se utiliza para testear las hipótesis

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

Diremos que  $\Phi$  es UMP de nivel menor o igual que  $\alpha$  si, dado  $\Phi_1$  de nivel menor o igual que  $\alpha$ ,

$$E_\theta(\Phi(\vec{X})) \geq E_\theta(\Phi_1(\vec{X})) \quad \forall \theta \in \Theta_1$$

Si las hipótesis nula y alternativa son simples, el Teorema de Neyman-Pearson provee un test UMP.

Teorema: Sean

$$H_0 : \theta = \theta_1 \quad \text{vs} \quad H_1 : \theta = \theta_2$$

y sea  $X_1, \dots, X_n$  una m.a. de la correspondiente distribución (discreta o continua). Sea  $f(x_i; \theta)$  la función de probabilidad o densidad y

$$f_{\vec{X}}(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

$$\text{El test } \Phi(\vec{X}) = \begin{cases} 1 & \text{si } \frac{f_{\vec{X}}(x_1, \dots, x_n; \theta_2)}{f_{\vec{X}}(x_1, \dots, x_n; \theta_1)} > k \\ v & \text{si } \frac{f_{\vec{X}}(x_1, \dots, x_n; \theta_2)}{f_{\vec{X}}(x_1, \dots, x_n; \theta_1)} = k \\ 0 & \text{si } \frac{f_{\vec{X}}(x_1, \dots, x_n; \theta_2)}{f_{\vec{X}}(x_1, \dots, x_n; \theta_1)} < k \end{cases}$$

con  $k$  y  $v$  tales que  $E_{\theta_1}(\Phi(\vec{X})) = \alpha$ , es UMP de nivel  $\alpha$ .

A partir de este Teorema se deduce que para las familias de cociente de verosimilitud monótono en el estadístico  $T(\vec{X})$ , o sea que para todo  $\theta_1 < \theta_2$  cumplen

a)  $f(\vec{x}, \theta_1) \neq f(\vec{x}, \theta_2)$

b)  $g_{\theta_1, \theta_2}(T(\vec{x})) = \frac{f(\vec{x}, \theta_2)}{f(\vec{x}, \theta_1)}$  es no decreciente en  $S = \{t / t = T(\vec{x}) \text{ con } f(\vec{x}; \theta_1) > 0 \text{ ó } f(\vec{x}; \theta_2) > 0\}$

un test de la forma

$$\Phi(\bar{X}) = \begin{cases} 1 & \text{si } T(\bar{X}) > k \\ v & \text{si } T(\bar{X}) = k \\ 0 & \text{si } T(\bar{X}) < k \end{cases}$$

con  $k$  y  $v$  tales que  $E_{\theta_1}(\Phi(\bar{X})) = \alpha$ , es UMP para las hipótesis

$$H_0 : \theta \leq \theta_1 \quad \text{vs} \quad H_1 : \theta > \theta_1$$

Nota: Para hipótesis bilaterales no se pueden hallar tests UMP sin imponer restricciones.

b) Tests insesgados: Intuitivamente, se desea que la probabilidad de rechazar  $H_0$  cuando es cierta sea menor que la de rechazarla cuando es falsa. Sean

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

y sea  $\Phi$  un test de nivel  $\alpha$ . Diremos que  $\Phi$  es insesgado si

$$E_{\theta}(\Phi(\bar{X})) \geq \alpha \quad \forall \theta \in \Theta_1$$

Todo test UMP es insesgado.

Definición: Un test  $\Phi$  de nivel  $\alpha$  se dice IUMP (insesgado uniformemente más potente) si

- $\sup_{\theta \in \Theta_0} \pi(\theta) = \alpha$
- $\Phi$  es insesgado
- dado  $\Phi_1$  insesgado,  $E_{\theta}(\Phi(\bar{X})) \geq E_{\theta}(\Phi_1(\bar{X}))$ ,  $\forall \theta \in \Theta_1$ .

c) Tests consistentes: este concepto equivale a la idea de convergencia en probabilidad de un estimador y refleja el comportamiento del test para muestras grandes. Una sucesión de tests es consistente para una alternativa dada si la potencia para detectar esa alternativa tiende a 1 cuando  $n$  tiende a infinito, y el nivel de significación se mantiene acotado lejos del 0.

La sucesión de tests  $\{\Phi_n\}$  es consistente en  $\theta_1 \in \Theta_1$  para las hipótesis

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

si

- $\alpha \geq E_\theta(\Phi_n) \geq \gamma > 0$  , para todo  $\theta \in \Theta_0$
- $E_{\theta_1}(\Phi_n) \xrightarrow{n \rightarrow \infty} 1$

Ejemplo: Se desea estudiar si los nacimientos en cierto país tienden a producir más bebés de un sexo. Sean  $X_1, \dots, X_n$  los sexos de los bebés nacidos en cierto período, donde

$$X_i = \begin{cases} 1 & \text{si es varón} \\ 0 & \text{si es mujer} \end{cases}$$

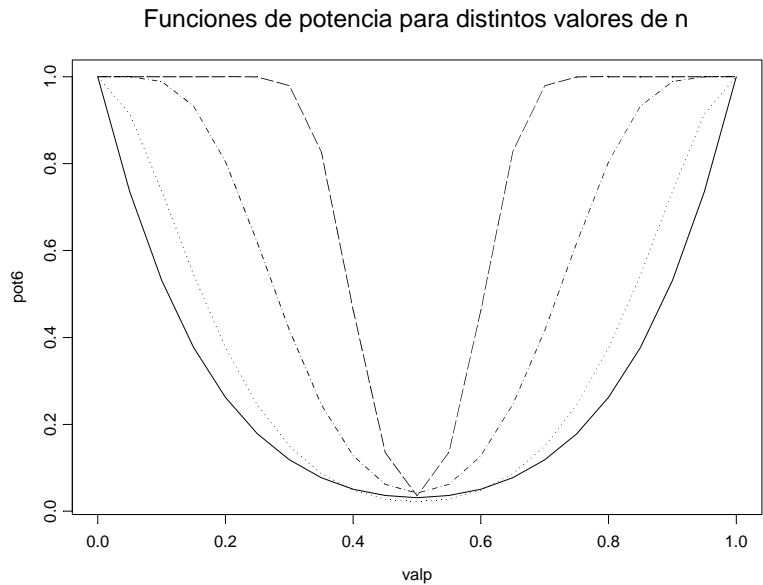
Suponiendo que las  $X_i$  son v.a. independientes y que  $p = P(\text{varón})$  es constante, se desea testear

$$H_0 : p = 0.5 \quad \text{vs} \quad H_1 : p \neq 0.5$$

El estadístico a utilizar es  $T = \sum_{i=1}^n X_i \sim Bi(n, p)$ . La región crítica del test IUMP se elige en forma simétrica y corresponde a valores grandes y chicos de  $T$ .

Supongamos que deseamos obtener tests no aleatorizados de nivel  $\alpha$  menor o igual que 0.05. Se obtiene una sucesión de tests, algunas de cuyas regiones de rechazo se presentan en la siguiente tabla:

| n   | Región crítica            | $\alpha$ |
|-----|---------------------------|----------|
| 5   | ninguna                   | -        |
| 6   | $T=0$ ó $T=6$             | 0.031    |
| 8   | $T=0$ ó $T=8$             | 0.008    |
| 10  | $T \leq 1$ ó $T \geq 9$   | 0.021    |
| 15  | $T \leq 3$ ó $T \geq 12$  | 0.035    |
| 20  | $T \leq 5$ ó $T \geq 15$  | 0.041    |
| 30  | $T \leq 9$ ó $T \geq 21$  | 0.043    |
| 60  | $T \leq 21$ ó $T \geq 39$ | 0.027    |
| 100 | $T \leq 39$ ó $T \geq 61$ | 0.035    |



Se observa que, para cada p fijo, a medida que crece n, la potencia crece hacia 1.

Un test razonable, será consistente para una subclase de alternativas, que se denomina *clase de consistencia* del test.

Notación: Llamando  $\Omega_{nul}$  al conjunto de distribuciones especificadas por  $H_0$  y  $\Omega_{alt}$  al conjunto de distribuciones especificadas por  $H_1$ , las hipótesis del test pueden plantearse como

$$H_0 : G \in \Omega_{nul} \quad \text{vs} \quad H_1 : G \in \Omega_{alt}$$

donde G es la distribución de las observaciones.

Teorema: Sea  $T_n$  el estadístico de un test para las hipótesis

$$H_0 : G \in \Omega_{nul} \quad \text{vs} \quad H_1 : G \in \Omega_{alt}$$

que rechaza  $H_0$  para valores grandes y que satisface

$$T_n \xrightarrow{p} \mu(G) = \begin{cases} \mu_o & \forall G \in \Omega_{nul} \\ > \mu_o & \forall G \in \Omega_c \subseteq \Omega_{alt} \end{cases} \quad (1)$$

Supongamos además que existe  $\sigma_o > 0$  tal que

$$\sqrt{n} \frac{T_n - \mu_o}{\sigma_o} \xrightarrow{d} N(0,1) \quad \forall G \in \Omega_{nul}$$

entonces existe una sucesión de valores críticos  $k_n$  tal que el test

$$\Phi_n = \begin{cases} 1 & \text{si } T_n \geq k_n \\ 0 & \text{si } T_n < k_n \end{cases}$$

es asintóticamente de nivel  $\alpha$  y  $P_G(T_n \geq k_n) \xrightarrow{n \rightarrow \infty} 1, \forall G \in \Omega_c$ , o sea la sucesión de tests es consistente en  $\Omega_c$ .

Dem: Sea  $Z_\alpha$  el percentil de la distribución normal standard que deja un área  $\alpha$  a su derecha y definamos

$$k_n = \mu_o + Z_\alpha \frac{\sigma_o}{\sqrt{n}}$$

Entonces, para toda  $\forall G \in \Omega_{nul}$ ,

$$\alpha_n = P_G(T_n \geq k_n) = P_G\left(\sqrt{n} \frac{T_n - \mu_o}{\sigma_o} \geq Z_\alpha\right) \xrightarrow{n \rightarrow \infty} \alpha$$

y por lo tanto la sucesión de tests  $\Phi_n$  tiene nivel asintótico  $\alpha$ .

Sea ahora  $G \in \Omega_c$  fija, y definamos

$$\varepsilon = \frac{\mu(G) - \mu_o}{2}$$

Por hipótesis,  $\varepsilon > 0$  y, como  $k_n \rightarrow \mu_o$ , para  $n$  suficientemente grande  $k_n < \mu_o + \varepsilon$ .

Como además  $\mu_o = \mu(G) - 2\varepsilon$ , entonces  $k_n < \mu(G) - \varepsilon$ . Entonces

$$|T_n - \mu(G)| < \varepsilon \Rightarrow T_n - \mu(G) > -\varepsilon \Rightarrow T_n > \mu(G) - \varepsilon \Rightarrow T_n \geq k_n$$

y, por lo tanto,  $P(|T_n - \mu(G)| < \varepsilon) \leq P(T_n \geq k_n)$ . Como, por (1), el primer término de esta desigualdad tiende a 1 cuando  $n$  tiende a infinito, el segundo también tiende a 1 y por lo tanto la sucesión de tests  $\Phi_n$  es consistente en  $G \in \Omega_c$ , como queríamos demostrar.

Nota: Cualquier test razonable debería ser consistente y si no lo es para un conjunto lógico de alternativas, debería ser descartado.

Ejemplo:  $X_1, \dots, X_n$  m.a. de una distribución de Cauchy de parámetro  $\theta$ , es decir con densidad

$$f(x; \theta) = \frac{1}{\pi (1 + (x - \theta)^2)}$$

Recordemos que  $\bar{X}$  tiene la misma distribución que  $X_i$ , independientemente de  $n$ . Si, para testear

$$H_0 : \theta = 0 \quad \text{vs} \quad H_1 : \theta > 0$$

la zona de rechazo estuviese dada por  $\bar{X} \geq c$ , la función de potencia del test sería

$$\pi(\theta) = E_\theta(\Phi(\bar{X})) = P_\theta(X_1 \geq c)$$

que no depende de  $n$  y, por lo tanto, no puede tender a 1 para ningún  $\theta > 0$ . Por lo tanto, este test no es consistente.

d) Tests eficientes: este concepto es relativo y permite comparar tests. Intuitivamente, dados dos tests, es más eficiente aquel que requiere un menor tamaño de muestra para alcanzar una potencia dada.

Definición: Sean  $T_n^{(1)}$  y  $T_n^{(2)}$  dos estadísticos basados en una m.a.  $X_1, \dots, X_n$  de  $F(x, \theta)$ , en los que se basan los tests de nivel  $\alpha$ ,  $\Phi_1$  y  $\Phi_2$  respectivamente, para las hipótesis

$$H_0 : \theta = 0 \quad \text{vs} \quad H_1 : \theta > 0$$

Es decir, consideremos los tests:

$$\Phi_1(\bar{X}) = \begin{cases} 1 & \text{si } T_n^{(1)} \geq k_n^{(1)} \\ 0 & \text{si } T_n^{(1)} < k_n^{(1)} \end{cases} \quad \Phi_2(\bar{X}) = \begin{cases} 1 & \text{si } T_n^{(2)} \geq k_n^{(2)} \\ 0 & \text{si } T_n^{(2)} < k_n^{(2)} \end{cases}$$

Fijados  $\theta$  y  $\beta$  ( $\alpha < \beta < 1$ ), sean  $n_1$  y  $n_2$  los tamaños de muestra necesarios para que la potencia de los dos tests en  $\theta$  sea  $\beta$ , o sea

$$P_\theta(T_{n_1}^{(1)} \geq k_n^{(1)}) = \beta \quad P_\theta(T_{n_2}^{(2)} \geq k_n^{(2)}) = \beta$$

Se define la **eficiencia de  $T_n^{(1)}$  relativa a  $T_n^{(2)}$**  como  $n_2/n_1$ .

Ejemplo: Supongamos que se dispone de dos tests de nivel 0.01 para las hipótesis anteriores y que para alcanzar una potencia  $\pi(\theta)=0.14$  en el  $\theta$  que hemos fijado, se requieren  $n_1=75$  observaciones con el primer test y  $n_2=50$  con el segundo, entonces el primer test es menos eficiente que el segundo y la eficiencia relativa es  $50/75=0.67$ .

La eficiencia relativa depende de la alternativa elegida. Preferiríamos una definición que nos permitiese hacer una comparación "global" entre tests.

Eficiencia asintótica relativa o eficiencia de Pitman: Sea  $X_1, \dots, X_n$  m.a. de una distribución absolutamente continua  $F(x; \theta)$  y consideremos las hipótesis

$$H_0 : \theta = 0 \quad \text{vs} \quad H_1 : \theta > 0$$

Sean dos tests de nivel asintótico  $\alpha$  para estas hipótesis

$$\Phi_n^{(1)}(\bar{X}) = \begin{cases} 1 & \text{si } T_n^{(1)} \geq k_n^{(1)} \\ 0 & \text{si } T_n^{(1)} < k_n^{(1)} \end{cases}$$

$$\Phi_n^{(2)}(\bar{X}) = \begin{cases} 1 & \text{si } T_n^{(2)} \geq k_n^{(2)} \\ 0 & \text{si } T_n^{(2)} < k_n^{(2)} \end{cases}$$

$$P_{H_0}(T_n^{(1)} \geq k_n^{(1)}) \xrightarrow{n \rightarrow \infty} \alpha \quad \text{y} \quad P_{H_0}(T_n^{(2)} \geq k_n^{(2)}) \xrightarrow{n \rightarrow \infty} \alpha$$

Sea  $\beta$  fijo ( $\alpha < \beta < 1$ ) y sea  $\{\theta_j\}$  una sucesión de alternativas tales que  $\theta_j \rightarrow 0$  (alternativas contiguas).

Sean  $\{n_j^{(1)}\}$  y  $\{n_j^{(2)}\}$  las correspondientes sucesiones de tamaños de muestra tales que

$$P_{\theta_j}(T_{n_j^{(1)}}^{(1)} \geq k_{n_j^{(1)}}^{(1)}) \xrightarrow{j \rightarrow \infty} \beta$$

$$P_{\theta_j}(T_{n_j^{(2)}}^{(2)} \geq k_{n_j^{(2)}}^{(2)}) \xrightarrow{j \rightarrow \infty} \beta$$

Si  $\lim_{j \rightarrow \infty} \frac{n_j^{(2)}}{n_j^{(1)}} = e_{12}$  existe y es independiente de la sucesión  $\{\theta_j\}$ ,  $\alpha$  y  $\beta$ , entonces  $e_{12}$  se

denomina eficiencia asintótica relativa de  $\Phi^{(1)}$  a  $\Phi^{(2)}$  o eficiencia de Pitman y da una indicación de la razón de los tamaños de muestra requeridos para alcanzar el mismo nivel y la misma potencia para alternativas próximas a la hipótesis nula.

Las siguientes condiciones, debidas a Pitman, facilitan en muchos casos el cálculo de la eficiencia.

Sea  $T_n$  un estadístico para testear las hipótesis:

$$H_0 : \theta = 0 \quad \text{vs} \quad H_1 : \theta > 0$$

y sea  $\Phi_n$  el correspondiente test

$$\Phi_n(\bar{X}) = \begin{cases} 1 & \text{si } T_n \geq k_n \\ 0 & \text{si } T_n < k_n \end{cases}$$

Condiciones:

- 1)  $T_n$  provee un test consistente
- 2) Existen sucesiones  $\{\mu_n(\theta)\}$  y  $\{\sigma_n(\theta)\}$  tales que

$$\frac{T_n - \mu_n(\theta)}{\sigma_n(\theta)} \xrightarrow{d} N(0,1)$$

uniformemente en un entorno de  $\theta = 0$ .

3) Existe  $\frac{d}{d\theta} \mu_n(\theta) \Big|_{\theta=0} = \mu'_n(0)$

4) Si  $\{\theta_n\}$  es una sucesión tal que  $\theta_n \xrightarrow{n \rightarrow \infty} 0$ ,

$$\frac{\sigma_n(\theta_n)}{\sigma_n(0)} \xrightarrow{n \rightarrow \infty} 1 \quad \text{y} \quad \frac{\mu'_n(\theta_n)}{\mu'_n(0)} \xrightarrow{n \rightarrow \infty} 1$$

5)  $\lim_{n \rightarrow \infty} \frac{\mu'_n(0)}{\sqrt{n} \sigma_n(0)} = c > 0$

La cantidad  $c$  se llama eficacia del test basado en  $T_n$ . Para  $n$  grande, mide la velocidad de cambio en unidades standard de la media "asintótica" de  $T_n$  en  $H_0$ . Un test con eficacia relativamente grande, responde rápido a alternativas próximas a 0 y se espera que tenga buenas propiedades de potencia local.

Nota: A menudo el test se construye de manera que

$$T_n \xrightarrow{p} \mu(\theta) \quad \text{o} \quad E_\theta(T_n) \rightarrow \mu(\theta)$$

y  $n \text{Var}(T_n) \rightarrow \sigma^2(\theta)$ , entonces se puede tomar  $\mu_n(\theta) = \mu(\theta)$  para todo  $n$  y  $\sigma_n(\theta) = \sigma(\theta)/\sqrt{n}$ .

En este caso, la condición 2) se convierte en

$$\sqrt{n} \frac{T_n - \mu(\theta)}{\sigma(\theta)} \xrightarrow{d} N(0,1)$$

y la eficacia se calcula a partir de los parámetros asintóticos  $\mu(\theta)$  y  $\sigma(\theta)$ , como

$$c = \frac{\mu'(0)}{\sigma(0)}$$

Ejemplo: Sea  $X_1, \dots, X_n$  una m.a. de una distribución  $N(\theta, \sigma^2)$  y consideremos el test  $t$  para las hipótesis

$$H_0 : \theta = 0 \quad \text{vs} \quad H_1 : \theta > 0$$

El estadístico de este test es  $T_n = \bar{X}$  y se rechaza  $H_0$  si  $T_n \geq k$ . Es fácil verificar las condiciones 1) a 5) anteriores y obtener la eficacia de este test



$$c = 1/\sigma$$

(Lo haremos en la práctica)

Teorema: Sea  $T_n$  un estadístico que satisface las condiciones 1) a 5) de Pitman y provee un test de nivel asintótico  $\alpha$  para testear

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta > 0$$

Sea  $\theta_n = \frac{\theta}{\sqrt{n}}$ , para  $\theta > 0$  fijo. Entonces la potencia asintótica está dada por

$$\lim_{n \rightarrow \infty} P_{\theta_n}(T_n \geq k_n) = 1 - \Phi(Z_\alpha - \theta c)$$

donde  $\lim_{n \rightarrow \infty} P_0(T_n \geq k_n) = 1 - \Phi(Z_\alpha) = \alpha$  y  $c$  es la eficacia del test.

Dem: Pag. 67 Hettmansperger.

Teorema: Sean  $T_n^{(1)}$  y  $T_n^{(2)}$  dos estadísticos que proveen tests para

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta > 0$$

de nivel asintótico  $\alpha$  y que ambos satisfacen las condiciones 1) a 5), entonces la eficiencia asintótica relativa de  $T_n^{(1)}$  a  $T_n^{(2)}$  está dada por

$$e_{12} = \frac{c_1^2}{c_2^2}$$

donde  $c_1$  y  $c_2$  son las eficacias de los tests basados en  $T_n^{(1)}$  y  $T_n^{(2)}$  respectivamente.

Dem: Pag. 68 Hettmansperger.