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: $X_1, X_2, \ldots, X_n$ e $Y_1, Y_2, \ldots, Y_m$.

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

Estadístico del test: $T = \dfrac{\overline{X} - \overline{Y}}{s} \sqrt{\dfrac{nm}{n+m}}$

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

Las hipótesis a testear podrían ser, por ejemplo:

$H_o: \mu_X = \mu_Y$ vs $H_1: \mu_X \neq \mu_Y$

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

$H_o: F_X = F_Y$ vs $H_1: F_X \neq F_Y$

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

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

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

$H_o: \Delta = 0$ vs $H_1: \Delta \neq 0$

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: $X_1,..., X_n$.

Estadístico: función de la muestra $T : \Re^n \to \Re$

Ejemplos:

a) $\overline{X}$

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

c) $min(X_1,..., X_n)$

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.

$$R_j = R(X_j) = card\left\{ i / X_i \le X_j \right\}$$

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.

$$X^{(1)} = 76 = X_2 \qquad X^{(2)} = 85 = X_3 \qquad X^{(3)} = 93 = X_1$$

Entonces,

$$R_1 = 3 \qquad R_2 = 1 \qquad R_3 = 2$$
$$D_1 = 2 \qquad D_2 = 3 \qquad D_3 = 1$$

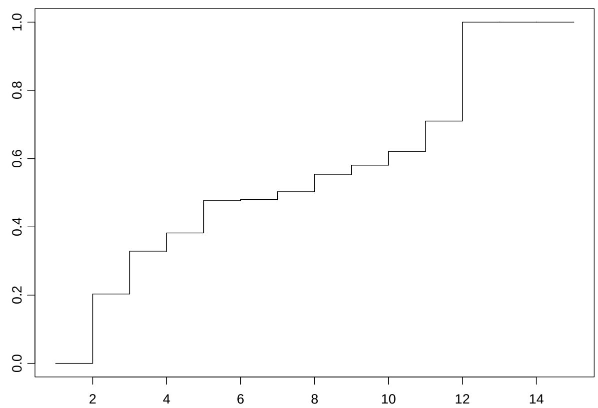
Estimación: Un estimador es una función de la muestra (estadístico) que provee un valor aproximado de un parámetro o característica desconocida.

Recordaremos un estimador "particular".

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

$$F_n(x) = \frac{card\,\{X_i \,/\, X_i \le x\}}{n}$$

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



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

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

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) $\overline{X}$ para estimar μ=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

$$\begin{cases} P(X \le x_p) \ge p \\ P(X \ge x_p) \ge 1-p \end{cases} \qquad \text{o} \qquad \begin{cases} P(X > x_p) \le 1-p \\ P(X < x_p) \le p \end{cases}$$

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

$$\begin{cases} \dfrac{card\ \{i / X_i > Q_p\}}{n} \le 1-p \\ \dfrac{card\ \{i / X_i < Q_p\}}{n} \le p \end{cases}$$

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

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

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

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

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

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

b) $X_1,..., X_n$, m.a. de una distribución exponencial de parámetro λ

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

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

$$\left[\frac{\chi^2{}_{2n,1-\alpha/2}}{2n\overline{X}}, \frac{\chi^2{}_{2n,\alpha/2}}{2n\overline{X}}\right]$$

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

$$P_\theta(\theta \in S_n(\vec{X})) \xrightarrow[n\to\infty]{} 1-\alpha \qquad \forall\,\theta$$

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

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

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

y una región de confianza para p de nivel asintótico 1-α es

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

c) $X_1,..., X_n$, m.a. de una distribución exponencial de parámetro λ, 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 λ de nivel asintótico 1-α. (*Ejercicio*).

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

- Hipótesis nula (H_o): 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

$$H_o : p \leq 0.05 \qquad vs \qquad H_1 : p > 0.05$$

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* ~ 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 $\vec{X} = (X_1,..., X_n)$ es un vector aleatorio cuya función de distribución depende de un parámetro $\theta \in \Theta \subset \Re^p$. Sean Θ_0 y Θ_1 tales que $\Theta_0 \cap \Theta_1 = \emptyset$. Un test Φ es una función del estadístico, en base a la cual se tomará la decisión entre

$$H_o : \theta \in \Theta_0 \qquad \text{vs} \qquad H_1 : \theta \in \Theta_1$$

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

Si Φ(*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_o.

Tipos de errores:

		Realidad	
		H_o cierta	**H_o falsa**
Decisión	**Rechazo Ho**	Error tipo I	OK!!
	No rechazo Ho	OK!!	Error tipo II

Definición: Supongamos que se utiliza el test Φ(T) para testear

$$H_o : \theta \in \Theta_o \qquad \text{vs} \qquad H_1 : \theta \in \Theta_1$$

se denomina **función de potencia** del test Φ a la función

$$\pi(\theta) = E_\theta(\Phi(T)) \qquad \forall\, \theta$$

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_o cuando el valor del parámetro es θ.
Si $\theta \in \Theta_o$ (especificado por H_o) entonces $\pi(\theta) = P_\theta(\text{error tipo I})$.

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

Se dice que el test tiene **nivel de significación** α si

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

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

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 $(\mu,(1.5°F)^2)$. ¿Contradicen los datos la afirmación del fabricante a nivel de significación α=0.01?

$$H_o : \mu = 130 \qquad vs \qquad H_1 : \mu \neq 130$$

$X_1,...,X_9$ m.a. $X_i \sim$ N $(\mu,(1.5°F)^2)$. El estadístico del test será

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

y la zona de rechazo para un nivel de significación α=0.01 es

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

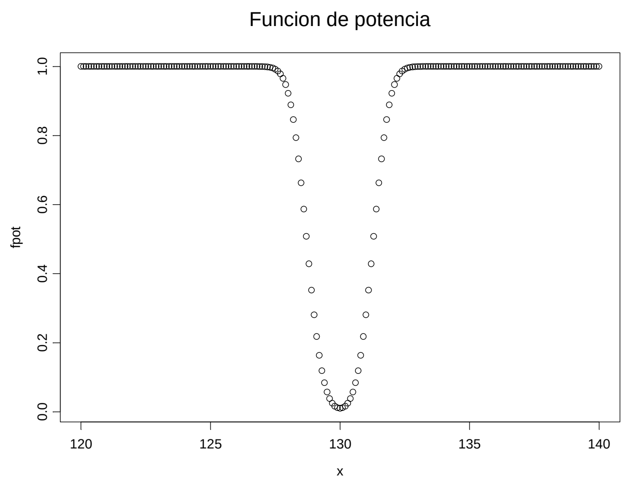
El valor observado de T = 2.16 entonces no se rechaza H_o a ese nivel.

$$p - valor = 2\, P(T \geq 2.16) = 0.0308$$

Hallemos la función de potencia del test:

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

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



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

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

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

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

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) <u>Test uniformemente más potentes</u>: Supongamos que el test Φ se utiliza para testear las hipótesis

$$H_o : \theta \in \Theta_o \qquad \text{vs} \qquad H_1 : \theta \in \Theta_1$$

Diremos que Φ es UMP de nivel menor o igual que α si, dado Φ_1 de nivel menor o igual que α,

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

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

<u>Teorema</u>: Sean

$$H_o : \theta = \theta_1 \qquad \text{vs} \qquad H_1 : \theta = \theta_2$$

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

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

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

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

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) $g_{\theta_1,\theta_2}(T(\vec{x})) = \dfrac{f(\vec{x},\theta_2)}{f(\vec{x},\theta_1)}$ 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

$$\Phi(\vec{X}) = \begin{cases} 1 & \text{si} \quad T(\vec{X}) > \mathrm{k} \\ \nu & \text{si} \quad T(\vec{X}) = \mathrm{k} \\ 0 & \text{si} \quad T(\vec{X}) < \mathrm{k} \end{cases}$$

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

$$H_o : \theta \leq \theta_1 \qquad \text{vs} \qquad H_1 : \theta > \theta_1$$

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_o cuando es cierta sea menor que la de rechazarla cuando es falsa. Sean

$$H_o : \theta \in \Theta_o \qquad \text{vs} \qquad H_1 : \theta \in \Theta_1$$

y sea Φ un test de nivel α. Diremos que Φ es insesgado si

$$E_\theta(\Phi(\vec{X})) \geq \alpha \qquad \forall\, \theta \in \Theta_1$$

Todo test UMP es insesgado.

Definición: Un test Φ de nivel α se dice IUMP (insesgado uniformemente más potente) si

- $\sup_{\theta \in \Theta_o} \pi(\theta) = \alpha$
- Φ es insesgado
- dado Φ_1 insesgado, $E_\theta(\Phi(\vec{X})) \geq E_\theta(\Phi_1(\vec{X}))$, $\forall\ \theta \in \Theta_1$.

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

$$H_o : \theta \in \Theta_o \qquad \text{vs} \qquad H_1 : \theta \in \Theta_1$$

si

- $\alpha \geq \mathrm{E}_\theta(\Phi_n) \geq \gamma > 0$, para todo $\theta \in \Theta_o$
- $\mathrm{E}_{\theta_1}(\Phi_n) \xrightarrow[n\to\infty]{} 1$

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

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

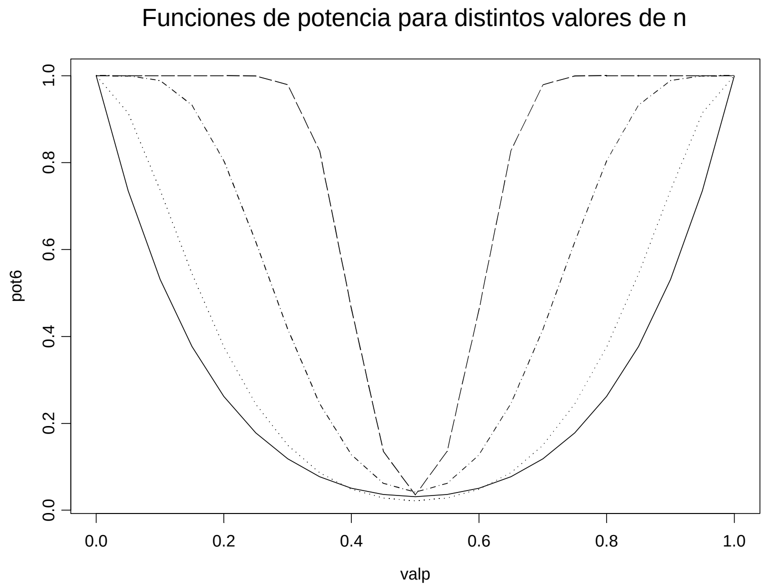
Suponiendo que las X_i son v.a. independientes y que p = P(varón) es constante, se desea testear

$$H_o : p = 0.5 \qquad vs \qquad H_1 : p \neq 0.5$$

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 α 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**	α
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 Ω_{nul} al conjunto de distribuciones especificadas por H_o y Ω_{alt} al conjunto de distribuciones especificadas por H_1, las hipótesis del test pueden plantearse como

$$H_o : G \in \Omega_{nul} \qquad \text{vs} \qquad H_1 : G \in \Omega_{alt}$$

donde *G* es la distribución de las observaciones.

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

$$H_o : G \in \Omega_{nul} \qquad \text{vs} \qquad H_1 : G \in \Omega_{alt}$$

que rechaza H_o para valores grandes y que satisface

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

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

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

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 α y $P_G(T_n \geq k_n) \xrightarrow[n\to\infty]{} 1$, $\forall$ $G \in \Omega_c$, o sea la sucesión de tests es consistente en Ω_c.

Dem: Sea Z_α el percentil de la distribución normal standard que deja un área α a su derecha y definamos

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

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

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

y por lo tanto la sucesión de tests Φ_n tiene nivel asintótico α.

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

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

Por hipótesis, $\varepsilon > 0$ y, como $k_n \to \mu_o$, 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, $P\big(|T_n - \mu(G)| < \varepsilon\big) \leq P(T_n \geq k_n)$. 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 Φ_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: $X_1,\ldots,X_n$ m.a. de una distribución de Cauchy de parámetro θ, 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

$$H_o: \theta = 0 \qquad vs \qquad H_1: \theta > 0$$

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

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

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) <u>Tests eficientes</u>: 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.

<u>Definición:</u> Sean $T_n^{(1)}$ y $T_n^{(2)}$ dos estadísticos basados en una m.a. $X_1,...,X_n$ de $F(x,\theta)$, en los que se basan los tests de nivel α, Φ_1 y Φ_2 respectivamente, para las hipótesis

$$H_o: \theta = 0 \qquad vs \qquad H_1: \theta > 0$$

Es decir, consideremos los tests:

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

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

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

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

<u>Ejemplo</u>: 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 θ 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.

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

$$H_o: \theta = 0 \qquad vs \qquad H_1: \theta > 0$$

Sean dos tests de nivel asintótico α para estas hipótesis

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

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

$$P_{H_o}(T_n^{(1)} \geq k_n^{(1)}) \xrightarrow[n\to\infty]{} \alpha \quad \text{y} \quad P_{H_o}(T_n^{(2)} \geq k_n^{(2)}) \xrightarrow[n\to\infty]{} \alpha$$

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

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

$$P_{\theta_j}\left(T^{(1)}_{n_j^{(1)}} \geq k^{(1)}_{n_j^{(1)}} \right) \xrightarrow[j\to\infty]{} \beta$$

$$P_{\theta_j}\left(T^{(2)}_{n_j^{(2)}} \geq k^{(2)}_{n_j^{(2)}} \right) \xrightarrow[j\to\infty]{} \beta$$

Si $\lim\limits_{j\to\infty} \dfrac{n_j^{(2)}}{n_j^{(1)}} = e_{12}$ existe y es independiente de la sucesión $\{\theta_j\}$, α y β, 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:

$$H_o : \theta = 0 \qquad vs \qquad H_1 : \theta > 0$$

y sea Φ_n el correspondiente test

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

Condiciones:

1) T_n provee un test consistente

2) Existen sucesiones $\{\mu_n(\theta)\}$ y $\{\sigma_n(\theta)\}$ tales que

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

uniformemente en un entorno de θ = 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 $\theta_n \xrightarrow[n\to\infty]{} 0$,

$$\frac{\sigma_n(\theta_n)}{\sigma_n(0)} \xrightarrow[n\to\infty]{} 1 \qquad \text{y} \qquad \frac{\mu_n'(\theta_n)}{\mu_n'(0)} \xrightarrow[n\to\infty]{} 1$$

5) $\lim\limits_{n\to\infty} \frac{\mu_n'(0)}{\sqrt{n}\,\sigma_n(0)} = c > 0$

La cantidad c se llama <u>eficacia</u> 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_o. Un test con eficacia relativamente grande, responde rápido a alternativas próximas a 0 y se espera que tenga buenas propiedades de potencia local.

<u>Nota</u>: A menudo el test se construye de manera que

$$T_n \xrightarrow{p} \mu(\theta) \qquad \text{o} \qquad E_\theta(T_n) \to \mu(\theta)$$

y $n\,Var(T_n) \to \sigma^2(\theta)$, 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

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

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

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

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

$$H_o: \theta = 0 \qquad vs \qquad H_1: \theta > 0$$

El estadístico de este test es $T_n = \overline{X}$ y se rechaza H_o 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 α para testear

$$H_o : \theta = 0 \qquad vs \qquad H_1 : \theta > 0$$

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

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

donde $\lim_{n\to\infty} P_0(T_n \geq k_n) = 1 - \Phi(Z_\alpha) = \alpha$ 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

$$H_o : \theta = 0 \qquad vs \qquad H_1 : \theta > 0$$

de nivel asintótico α 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.