Otras pruebas con hipótesis compuestas.

→ Consideremos el siguiente caso

$$H_0: \theta \in \Theta_0 \quad vs \quad H_a: \theta \in \Theta_a$$

con $\Theta_a = \Theta \setminus \Theta_0 = \Theta_0^c$

Definición. Cociente de verosimilitud generalizado.
Sea $L(\theta; x_1, \dots, x_n)$ la función de verosimilitud de una m.a. $X_1, \dots, X_n$ con función de densidad o probabilidad conjunta $f(x_1, \dots, x_n; \theta)$. Supongamos $\theta \in \Theta$.
El cociente de verosimilitud generalizado se define como

$$\lambda(x_1, \dots, x_n) = \frac{\sup_{\theta \in \Theta_0} L(\theta; x_1, \dots, x_n)}{\sup_{\theta \in \Theta} L(\theta; x_1, \dots, x_n)}$$

Notar que

i) $$\sup_{\theta \in \Theta} L(\theta; x_1, \dots, x_n) = L(\theta; x_1, \dots, x_n) \Big|_{\theta = \hat{\theta}}$$

donde $\hat{\theta}$ es el EMV sobre todo Θ.

ii) $$\sup_{\theta \in \Theta_0} L(\theta; x_1, \dots, x_n) = L(\theta; x_1, \dots, x_n) \Big|_{\theta = \hat{\hat{\theta}}}$$

donde $\hat{\hat{\theta}}$ es el EMV encontrado sobre Θ_0.

iii) $0 \le \lambda \le 1$, pues el denominador siempre es mayor al numerador.

iv) θ puede ser un vector de parámetros.

v) λ tenderá a ser pequeño cuando H_0 es falsa

vi) $\lambda(x_1,..,x_n)$ es una estadística y tiene su propia distribución, aunque a veces no es fácil identificarla.

vii) En general, se rechaza H_0 si

$$\mathcal{C} = \{ \lambda \le \lambda_0 \}$$ con λ_0 un valor entre 0 y 1

tal que $\sup_{\theta \in \Theta_0} P(\lambda \le \lambda_0 | \theta) \le \alpha$. [Cota al error tipo I]

viii) En ocasiones, el test encontrado en vii) no es bueno, pero en general tiene propiedades aceptables.

ix) En general, este test es usado cuando θ es un vector, es decir, $\theta = (\theta_1, .., \theta_k)$ y cuando

$$H_0: \theta_1 = \theta_1^0, .., \theta_r = \theta_r^0 \quad r \le k \quad vs$$

$$H_a: \text{ alguna } \theta_i \ne \theta_i^0 \text{ para } i = 1,..,r.$$

En este caso, bajo las condiciones de regularidad

$-2 \log \lambda$ tiene una distribución $\chi^2_{(r)}$

cuando H_0 es cierta y n es grande. De donde

Se rechaza H_0 si $-2 \log \lambda > \chi^2_{(r)\, 1-\alpha}$

x) Al usar ix) es importante observar que los grados de libertad se pueden obtener de la siguente forma en muchas ocasiones

Sea K el número de parámetros desconocidos en $\theta \in \Theta$,

Sea K' el número de parámetros desconocidos en $\theta \in \Theta_0$, es decir, en θ cuando H_0 es cierta.

$$\Rightarrow \quad r = K - K'$$

Ejemplo de x)

Sea $X_1, \ldots, X_n$ una m.a de la distribución $N(\mu_1, \sigma^2)$ y sea $Y_1, \ldots, Y_m$ una m.a de la distribución $N(\mu_2, \sigma^2)$, ambas independientes. En este caso hay 3 parámetros desconocidos μ_1, μ_2 y σ^2

Si $H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 \neq \mu_2$

$\Rightarrow$ Bajo H_0 hay sólo dos parámetros desconocidos $\mu = \mu_1 = \mu_2$ y σ^2

$\therefore \quad r = 3 - 2 = 1$. Notar que $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ ya no son i.d. pero dentro de cada una si y son de la fam. exponencial

---

Ejemplos.

Ejemplo.

Sea $X_1, \dots, X_n$ una m.a. de la distribución exponencial (θ)
Encuentre la prueba usando el cociente de verosimilitudes generalizado para:

$$H_0: \theta = \theta_0 \quad vs \quad H_a: \theta \neq \theta_0$$

a) Usando la distribución exacta de una función de $\lambda(x_1, \dots, x_n)$

b) Indique cual sería la región de rechazo (si se puede) asumiendo la distribución asintótica del cociente de verosimilitud generalizado

a) $L(\theta; x_1, \dots, x_n) = \prod_{i=1}^{n} \theta e^{-\theta x_i} \mathbb{I}_{(0,\infty)}(x_i) = \theta^n e^{-\theta \sum x_i} \prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x_i)$

i) $\sup_{\theta \in \Theta_0} L(\theta; x_1, \dots, x_n) = L(\theta_0; x_1, \dots, x_n) = \theta_0^n e^{-\theta_0 \sum x_i} \prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x_i)$

ii) $\sup_{\theta \in \Theta} L(\theta; x_1, \dots, x_n) = L(\hat{\theta}; x_1, \dots, x_n)$ con $\hat{\theta}$ el EMV.

$$\ln L(\theta; x_1, \dots, x_n) = n \ln \theta - \theta \sum x_i + \ln \left(\prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x_i) \right)$$

$$\frac{\partial \ln L}{\partial \theta} = \frac{n}{\theta} - \sum x_i = 0 \iff \hat{\theta} = \frac{1}{\bar{X}}$$

$$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{1}{\theta^2} < 0 \quad \forall \theta > 0. \therefore \hat{\theta} = \frac{1}{\bar{X}} \text{ es el EMV.}$$

$$\Rightarrow \sup_{\theta \in \Theta} L(\theta; x_1, \dots, x_n) = \left(\frac{1}{\bar{x}}\right)^n e^{-\frac{1}{\bar{x}} \sum x_i} \prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x)$$

$$= \left(\frac{1}{\bar{x}}\right)^n e^{-n} \prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x)$$

$$\lambda(x_1,\dots,x_n) = \frac{\theta_0^n e^{-\theta_0 \sum x_i} \prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x)}{\left(\frac{1}{\bar{x}}\right)^n e^{-n} \prod_{i=1}^{n} \mathbb{I}_{(0,\infty)}(x)} = (\bar{x}\theta_0)^n e^{-\theta_0 \sum x_i + n}$$

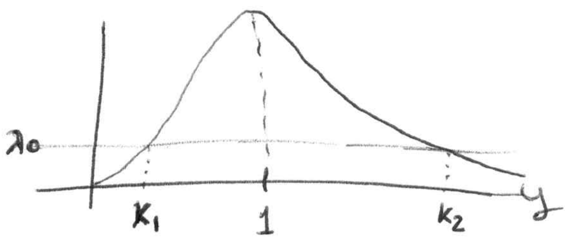
$\Rightarrow$ Se rechaza H_0 si

$$(\bar{x}\theta_0)^n e^{-\theta_0 \sum x_i + n} \leq \lambda_0$$

$$\Leftrightarrow \quad (\bar{x}\theta_0)^n e^{n\left(-\theta_0 \frac{\sum x_i}{n} + 1\right)} \leq \lambda_0$$

Ahora, sea $y = \bar{x}\theta_0$

$\Rightarrow$ Como función $g(y) = y^n e^{-n(y-1)}$ es de la forma



$\Rightarrow$ Si $\bar{x}\theta_0 \leq k_1$ ó

$\bar{x}\theta_0 \geq k_2$

tal que $g(k_1) = g(k_2)$ y

$$\alpha = P[\theta_0 \bar{x} \leq k_1 | \theta_0] + P[\theta_0 \bar{x} \geq k_2 | \theta_0]$$

se cumple que $\lambda(x_1,\dots,x_n) \leq \lambda_0$

La versión asintótica es

Rechazar H_0 si

$$-2\ln\lambda > \chi^2_{r,\,1-\alpha}$$

donde $r = k - k' = 1 - 0$, es decir

$$-2n\left[\ln(\bar{x}\theta_0) - \theta_0\bar{x} + 1\right] > \chi^2_{(1),\,1-\alpha}$$

El uso del p-value

Sin pérdida de generalidad, sea

$$\mathscr{C} = \{(x_1,\ldots,x_n) : W(x_1,\ldots,x_n) < K\}$$

{se podrían considerar $\leq$ $>$ $\geq$}

la región de rechazo de una prueba de hipótesis; donde K es una constante fija y conocida y $W(x_1,\ldots,x_n)$ es una estadística (Depende de la muestra y constantes conocidas).

La regla de decisión para rechazar (o no rechazar) H_0 es:

sean $(x_1,\ldots,x_n) = x$ los valores observados, si $x \in \mathscr{C}$, entonces se rechaza H_0; en otro caso no se rechaza H_0.

En muchos casos a $W(x_1,\ldots,x_n)$ se le conoce como la estadística de la prueba y por lo general se conoce so distribución para un valor fijo de θ.

La probabilidad de cometer el error tipo I se aproxima de forma general como:

$$\alpha = \sup_{\theta \in \Theta_0} P(x \in \mathscr{C} \mid \theta)$$

y se dice que se rechaza o no se rechaza H_0 con un nivel o significancia α.

Una regla de decisión alternativa se basa en el <u>p-value</u>.

Sean $(x_1,\dots,x_n) = x$ los valores observados, entonces el p-value se define como

$$\text{p-value} = \sup_{\theta \in \Theta_0} P\left(W(X_1,\dots,X_n) < w(x_1,\dots,x_n) \mid \theta\right),$$

es decir, es la probabilidad de observar el valor $w(x_1,\dots,x_n)$ o uno más extremo en el sentido de la región de rechazo $\mathcal{C}_\rho$.

La regla es:

Rechazar H_0 si p-value $< \alpha \in (0,1)$

Notemos que las reglas de decisión son equivalentes, pues k en $\mathcal{C}$ es tal que

$$\alpha = \sup_{\theta \in \Theta_0} P(x \in \mathcal{C} \mid \theta) = \sup_{\theta \in \Theta_0} (W(X_1,\dots,X_n < k)$$

De donde

$$\text{p-value} < \alpha = \sup_{\theta \in \Theta_0} P(W(X_1,\dots,X_n) < k \mid \theta)$$

$$\iff$$

$$w(x_1,\dots,x_n) < k$$

$$\iff (x_1,\dots,x_n) \in \mathcal{C}.$$

En la práctica los paquetes estadísticos proporcionan el p-value de varias pruebas de hipótesis ya programadas.

Ejemplo

Sea $X_1, \ldots, X_n$ una m.a de la distribución Normal(μ, σ^2), donde σ^2 es conocida. Suponga que se desea probar

$$H_0: \mu = \mu_0 \quad vs \quad H_a: \mu = \mu_1$$

con μ_0 y μ_1 valores fijos conocidos y tal que $\mu_1 < \mu_0$.

La prueba UMP tiene como región de rechazo para un nivel α a

$$\mathcal{C} = \left\{ (x_1, \ldots, x_n) : \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} < Z_\alpha \right\}$$

El p-value para esta prueba se calcularía como:

Sea $W(X_1, \ldots, X_n) = \dfrac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} \Rightarrow$ Bajo $H_0: \mu = \mu_0$

$W(X_1, \ldots, X_n) \sim N(0,1)$

$\Rightarrow$ para una muestra $x_1, \ldots, x_n$

$$\text{p-value} = P\left(Z < \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} \right) \quad \text{con } Z \sim N(0,1)$$

$$= F\left(\frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} \right)$$

Ver PH_clase 2. R

Algunos test asociados a la distribución normal.

i) Supongamos σ^2 conocida

$$L(\mu; x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \mathbb{I}_{(-\infty,\infty)}(x_i)$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)$$

Sea $\mu_1 > \mu_0$

$$\frac{f(x_1, \dots, x_n; \mu_0)}{f(x_1, \dots, x_n; \mu_1)} = \frac{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_1)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}$$

$$= e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\mu_1)^2\right]}$$

$$= e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i^2-2x_i\mu_0+\mu_0^2) - \sum_{i=1}^{n}(x_i^2-2x_i\mu_1+\mu_1^2)\right]}$$

$$= e^{-\frac{1}{2\sigma^2}\left[-2\sum x_i(\mu_0-\mu_1) + n(\mu_0^2-\mu_1^2)\right]}$$

$$= e^{\frac{\sum x_i}{\sigma^2}(\mu_0-\mu_1)} \cdot M$$

con $M > 0$

Como función de $\sum x_i$ y dado que $\mu_1 > \mu_0$ se tiene que el cociente de verosimilitud es monótono decreciente.

De donde

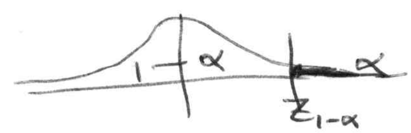
a) $\mathcal{C} = \{ (x_1, \dots, x_n) : T(x_1, \dots, x_n) = \sum X_i > k \}$ es una prueba UMP. de nivel α para $\underline{H_0 : \mu \leq \mu_0}$ vs $\underline{H_a : \mu > \mu_0}$ con k tal que

$$P\left(\sum X_i > k \mid \mu_0 \right) = \alpha .$$

De donde $\sum X_i \sim N(n\mu_0, n\sigma^2)$

$$P\left(\frac{\sum X_i - n\mu_0}{\sqrt{n\sigma^2}} > \frac{k - n\mu_0}{\sqrt{n\sigma^2}} \right) = \alpha$$

$$\Rightarrow \quad \frac{K - n\mu_0}{\sqrt{n\sigma^2}} = Z_{1-\alpha}$$



$$\Rightarrow \quad K = Z_{1-\alpha}\sqrt{n\sigma^2} + n\mu_0$$

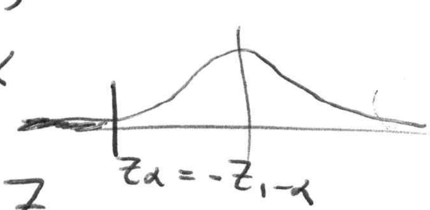
$$\therefore \ \mathcal{C} = \{ (x_1, \dots, x_n) : \sum X_i > Z_{1-\alpha}\sqrt{n\sigma^2} + n\mu_0 \}$$

b) $\mathcal{C} = \{ (x_1, \dots, x_n) : T(X_1, \dots, X_n) = \sum X_i < K \}$ es una prueba UMP de nivel α para $H_0 : \mu \geq \mu_0$ vs $H_a : \mu < \mu_0$ con K tal que

$$P\left(\sum X_i < K \mid \mu_0 \right) = \alpha$$

De donde $\sum X_i \sim N(n\mu_0, n\sigma^2)$

$$\Rightarrow \quad P\left(\frac{\sum X_i - n\mu_0}{\sqrt{n\sigma^2}} < \frac{k - n\mu_0}{\sqrt{n\sigma^2}} \right) = \alpha$$



$$\Rightarrow \quad \frac{k - n\mu_0}{\sqrt{n\sigma^2}} = Z_\alpha = -Z_{1-\alpha}$$

$$\Rightarrow \quad K = -Z_{1-\alpha}\sqrt{n\sigma^2} + n\mu_0$$

$$\therefore\ \mathscr{C} = \{(x_1,\dots,x_n) : \sum X_i < -Z_{1-\alpha}\sqrt{n\sigma^2} + n\mu_0\}$$

c) Para $H_0: \mu=\mu_0$ vs $H_a: \mu\neq\mu_0$

Se puede usar el cociente de verosimilitud generalizado

$$\lambda = \frac{\sup\limits_{\mu\in\Theta_0} L(\mu; x_1,\dots,x_n)}{\sup\limits_{\mu\in\Theta} L(\mu; x_1,\dots,x_n)} = \frac{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu_0)^2} \prod_{i=1}^n \mathbb{I}_{(-\infty,\infty)}(x_i)}{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\bar{x})^2} \prod_{i=1}^n \mathbb{I}_{(-\infty,\infty)}(x_i)}$$

$$= e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n (x_i-\mu_0)^2 - \sum_{i=1}^n (x_i-\bar{x})^2\right]}$$

$$= e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n (x_i^2-2x_i\mu_0+\mu_0^2) - \sum_{i=1}^n (x_i^2-2x_i\bar{x}+\bar{x}^2)\right]}$$

$$= e^{-\frac{1}{2\sigma^2}\left[-2\mu_0\sum x_i + \mu_0^2 n + \bar{x}^2 n\right]}$$

$$= e^{-\frac{n}{2\sigma^2}\left[-2\mu_0\bar{x} + \mu_0^2 + \bar{x}^2\right]}$$

$$= e^{-\frac{n}{2\sigma^2}\left[(\mu_0-\bar{x})^2\right]}$$

Se rechaza H_0 si

$$\lambda \leq \lambda_0$$

$$\Rightarrow \quad e^{-\frac{n}{2\sigma^2}\left[(\mu_0-\bar{x})^2\right]} \leq \lambda_0$$

$$\Rightarrow \quad -\frac{n}{2\sigma^2}\left[(\mu_0-\bar{x})^2\right] \leq \ln\lambda_0$$

$$\Rightarrow \quad \left(\frac{\mu_0-\bar{x}}{\sigma/\sqrt{n}}\right)^2 \geq -2\ln\lambda_0 = k^*$$

De donde buscamos k^* tal que

$$P\left(\left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right)^2 \geq k^* \Big|_{\mu_0}\right)=\alpha, \quad \text{pero} \quad \left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2_{(1)}$$

$$\Rightarrow \quad k^* = \chi^2_{(1)\,1-\alpha}$$

$$\Rightarrow$$

$$\mathcal{C} = \left\{(x_1,\dots,x_n) : \left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right)^2 \geq \chi^2_{(1)\,1-\alpha}\right\}$$

equivalentemente

$$\mathcal{C} = \left\{(x_1,\dots,x_n) : \left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq \left(\chi^2_{(1)\,1-\alpha}\right)^{1/2} = Z_{1-\alpha/2}\right\}$$

$$\therefore \quad \mathcal{C} = \left\{(x_1,\dots,x_n) : \left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq Z_{1-\alpha/2}\right\}$$

Notar que la propiedad asintótica de

$$-2\ln\lambda \sim \chi^2_{(r)}$$

con $r = k - k' = 1 - 0 = 1$

$$\Rightarrow \quad -2\ln\left(e^{-\frac{n}{2\sigma^2}\left[(\mu_0-\bar{x})^2\right]}\right) > \chi^2_{(1)\,1-\alpha}$$

$$\Rightarrow \quad \frac{n}{\sigma^2}\left[(\mu_0-\bar{x})^2\right] > \chi^2_{(1)\,1-\alpha}$$

es <u>equivalente</u> a la prueba basada en la distribución exacta de $\bar{x}$ asumiendo μ_0

ii) asumiendo σ^2 desconocida.

a) $H_0: \mu = \mu_0$ vs $H_a: \mu \neq \mu_0$

Podemos usar el cociente de verosimilitud generalizado

$$\lambda = \frac{\sup_{\{\mu=\mu_0, \sigma^2\}} L(\mu, \sigma^2; x_1, \ldots, x_n)}{\sup_{\{\mu, \sigma^2\}} L(\mu, \sigma^2; x_1, \ldots, x_n)}$$

→ para el numerador, dado $\mu=\mu_0$ el EMV de σ^2 es

$$\hat{\hat{\sigma}}^2 = \frac{\sum_{i=1}^{n} (X_i - \mu_0)^2}{n}$$

→ para el denominador, dado que μ y σ^2 son desconocidos, sus Estimadores Máximo verosímiles son

$$\hat{\mu} = \bar{X} \quad y \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{X})^2}{n}$$

$$\Rightarrow \lambda = \frac{L(\mu_0, \hat{\hat{\sigma}}^2; x_1, \ldots, x_n)}{L(\hat{\mu}, \hat{\sigma}^2; x_1, \ldots, x_n)} = \frac{\frac{1}{(2\pi \hat{\hat{\sigma}}^2)^{n/2}} e^{-\frac{1}{2\hat{\hat{\sigma}}^2} \sum_{i=1}^{n} (x_i - \mu_0)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}{\frac{1}{(2\pi \hat{\sigma}^2)^{n/2}} e^{-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^{n} (x_i - \bar{x})^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}$$

$$= \left(\frac{\hat{\sigma}^2}{\hat{\hat{\sigma}}^2}\right)^{n/2} e^{-\frac{n}{2} + \frac{n}{2}} = \left(\frac{\hat{\sigma}^2}{\hat{\hat{\sigma}}^2}\right)^{\frac{n}{2}}$$

$$= \left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \mu_0)^2}\right)^{n/2}$$

$$= \left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x} + \bar{x} - \mu_0)^2}\right)^{n/2}$$

$$= \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2 + \cancelto{0}{2 \sum_{i=1}^{n} (x_i - \bar{x})(\bar{x} - \mu_0)} + n (\bar{x} - \mu_0)^2} \right]^{n/2}$$

$$= \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2 + n (\bar{x} - \mu_0)^2} \right]^{n/2}$$

$$= \left[\frac{1}{1 + n \frac{(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \right]^{n/2}$$

ahora $\lambda \leq \lambda_0$

$$\Leftrightarrow \quad \lambda^{2/n} \leq \lambda_0^{2/n}$$

$$\Leftrightarrow \quad \lambda_0^{n/2} \leq 1 + n \frac{(\bar{x} - \mu_0)^2}{\sum (x_i - \bar{x})^2}$$

$$\Leftrightarrow \quad \left(\lambda_0^{n/2} - 1\right)(n-1) \leq \frac{n (\bar{x} - \mu_0)^2}{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \frac{\left(\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right)^2}{\frac{\sum (x_i - \bar{x})^2}{\sigma^2 (n-1)}}$$

Sea $K^* = \left(\lambda^{n/2} - 1\right)(n-1) \leq \left(\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right)^2 \Big/ \frac{\sum (x_i - \bar{x})^2}{\sigma^2 (n-1)} = F$

Donde $F \sim F_{1,n-1}$ Bajo $H_0 : \mu = \mu_0$

$\Rightarrow$ Buscamos K^* tal que

$$P(F \geq K^* \mid \mu_0) = \alpha$$

$$\Rightarrow K^* = F^{1,n-1}_{1-\alpha}$$

$$\Rightarrow \mathcal{C} = \left\{ (x_1, \ldots, x_n) : \frac{(\bar{x} - \mu_0)^2}{\frac{\sum (x_i - \bar{x})^2}{n(n-1)}} \geq F_{\substack{1, n-1 \\ 1-\alpha}} \right\}$$

Notar que esto es equivalente a

$$\mathcal{C} = \left\{ (x_1, \ldots, x_n) : \left| \frac{\bar{x} - \mu_0}{\tilde{\sigma}/\sqrt{n}} \right| \geq t_{\substack{n-1 \\ 1-\alpha/2}} \right\}$$

con $\tilde{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$

b) Se puede verificar que cuando σ^2 es desconocida

$$\mathcal{C} = \left\{ (x_1, \ldots, x_n) : \frac{\bar{x} - \mu_0}{\tilde{\sigma}/\sqrt{n}} > t_{\substack{n-1 \\ 1-\alpha}} \right\}$$

es la región de rechazo para

$H_0: \mu \leq \mu_0$ vs $H_a: \mu > \mu_0$

y

$$\mathcal{C} = \left\{ (x_1, \ldots, x_n) : \frac{\bar{x} - \mu_0}{\tilde{\sigma}/\sqrt{n}} < -t_{\substack{n-1 \\ 1-\alpha}} \right\}$$

es la región de rechazo para

$H_0: \mu \geq \mu_0$ vs $H_a: \mu < \mu_0$

iii) Supongamos μ conocida

$$L(\sigma^2; x_1,\dots,x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \mathbb{I}_{(-\infty,\infty)}(x_i)$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)$$

Sean $\sigma_1^2 > \sigma_0^2 > 0$

$$\frac{L(\sigma_0^2; x_1,\dots,x_n)}{L(\sigma_1^2; x_1,\dots,x_n)} = \frac{\frac{1}{(2\pi\sigma_0^2)^{n/2}} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}{\frac{1}{(2\pi\sigma_1^2)^{n/2}} e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i-\mu)^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}$$

$$= \left(\frac{\sigma_1^2}{\sigma_0^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}\left[\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right]}$$

$$= M \cdot e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\left[\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right]}$$

con $M > 0$

$\Rightarrow$ Si $T = \sum_{i=1}^{n}(x_i-\mu)^2$ tenemos que el cociente de verosimilitud es una función monótona decreciente.

A partir de esto se pueden obtener las pruebas uniformemente más potente para

a) $H_0: \sigma^2 \leq \sigma_0^2$ vs $H_a: \sigma^2 > \sigma_0^2$

b) $H_0: \sigma^2 \geq \sigma_0^2$ vs $H_a: \sigma^2 < \sigma_0^2$

→ Cuando μ es conocida y dado que el cociente de verosimilitud es monótono decreciente en términos de $T = \sum_{i=1}^{n}(x_i-\mu)^2$, se puede verificar que la región de rechazo de la prueba UMP con significancia α para

a) $H_0: \sigma^2 \leq \sigma_0^2$ vs $H_a: \sigma^2 > \sigma_0^2$

es

$$\mathcal{C} = \left\{ (x_1,\ldots,x_n) : \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma_0^2} > \chi^2_{(n),\,1-\alpha} \right\}$$

b) $H_0: \sigma^2 \geq \sigma_0^2$ vs $H_a: \sigma^2 < \sigma_0^2$

es

$$\mathcal{C} = \left\{ (x_1,\ldots,x_n) : \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma_0^2} < \chi^2_{(n),\,\alpha} \right\}$$

Para lo anterior se usa que

$$\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma_0^2} \sim \chi^2_{(n)} \quad \text{cuando } X_i \sim N(\mu, \sigma_0^2)$$

→ Cuando μ es desconocida un posible test o prueba con significancia α para

a) $H_0: \sigma^2 \leq \sigma_0^2$ vs $H_a: \sigma^2 > \sigma_0^2$

es

$$\mathcal{C} = \left\{ (x_1,\ldots,x_n) : \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma_0^2} > \chi^2_{(n-1),\,1-\alpha} \right\}$$

b) $H_0: \sigma^2 \geq \sigma_0^2$ vs $H_a: \sigma^2 < \sigma_0^2$

es

$$\mathcal{C} = \left\{ (x_1,\ldots,x_n) : \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma_0^2} < \chi^2_{(n-1),\,\alpha} \right\}$$

Se puede verificar que esto se obtiene a partir del cociente de verosimilitud generalizado.

Para el caso μ conocida o μ desconocida y

$$H_0: \sigma^2 = \sigma_0^2 \quad vs \quad H_a: \sigma^2 \neq \sigma_0^2$$

es posible usar el cociente de verosimilitud generalizado. Supongamos μ desconocida.

$$\lambda = \frac{\sup_{\{\mu, \sigma^2 = \sigma_0^2\}} L(\mu, \sigma^2; x_1, \dots, x_n)}{\sup_{\{\mu, \sigma^2\}} L(\mu, \sigma^2; x_1, \dots, x_n)}$$

→ Para el numerador, dado $\sigma^2 = \sigma_0^2$, el EMV de μ es $\hat{\hat{\mu}} = \overline{X}$.

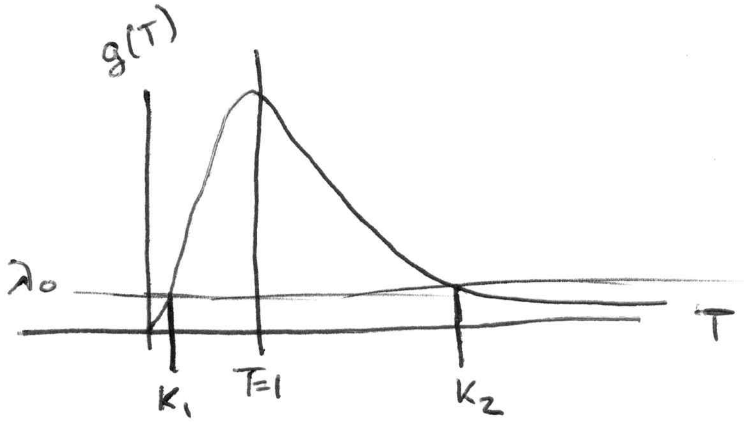
→ Para el denominador los EMV de μ y σ^2 son $\hat{\mu} = \overline{x}$ y $\hat{\sigma}^2 = \frac{\sum (x_i - \overline{x})^2}{n}$

$$\Rightarrow \lambda = \frac{L(\hat{\hat{\mu}}, \sigma_0^2; x_1, \dots, x_n)}{L(\hat{\mu}, \hat{\sigma}^2; x_1, \dots, x_n)} = \frac{\frac{1}{(2\pi\sigma_0^2)^{n/2}} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i - \overline{x})^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}{\frac{1}{(2\pi\hat{\sigma}^2)^{n/2}} e^{-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i - \overline{x})^2} \prod_{i=1}^{n} \mathbb{I}_{(-\infty,\infty)}(x_i)}$$

$$= \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} e^{-\frac{1}{2}\left[\frac{\sum(x_i - \overline{x})^2}{\sigma_0^2} - n\right]} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} e^{-\frac{n}{2}\left[\frac{\hat{\sigma}^2}{\sigma_0^2} - 1\right]}$$

Sea $T = \frac{\hat{\sigma}^2}{\sigma_0^2}$

$$\rightarrow \quad \lambda = (T)^{n/2} e^{-\frac{n}{2}(T-1)} = g(T)$$



$$\lambda \le \lambda_0 \iff T \le k_1 \quad o \quad T \ge k_2$$

Es decir

$$\frac{\hat{\sigma}^2}{\sigma_0^2} \le k_1 \quad o \quad \frac{\hat{\sigma}^2}{\sigma_0^2} \ge k_2$$

O bien

$$\frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \le n k_1 \quad o \quad \frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \ge n k_2$$

$\Rightarrow$ Bajo H_0, $T^* = \frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \sim \chi^2_{(n-1)}$

Por lo que se buscan k_1 y k_2 tal que

$$\alpha = P(T^* \le n k_1) + P(T^* \ge n k_2)$$

y $$g(k_1) = g(k_2)$$

En la práctica se relaja $g(k_1) = g(k_2)$ y sólo se buscan k_1 y k_2 tal que

$$\alpha = P(T^* \le n k_1) + P(T^* \ge n k_2)$$

En cuyo caso, una alternativa es

$$nK_1 = \chi^2_{(n-1)\,\alpha/2} \quad y \quad nK_2 = \chi^2_{(n-1)\,1-\alpha/2}$$

$\Rightarrow$ Se rechaza Ho si

$$\mathcal{C} = \left\{ (x_1,..,x_n) : \frac{\sum (x_i-\bar{x})^2}{\sigma_0^2} \leq \chi^2_{(n-1)\,\alpha/2} \;\text{ ó }\; \frac{\sum (x_i-\bar{x})^2}{\sigma_0^2} \geq \chi^2_{(n-1)\,1-\alpha/2} \right\}$$

Para el caso de μ conocida la región de rechazo para una significancia α y

$$H_0 : \sigma^2 = \sigma_0^2 \quad vs \quad H_a : \sigma^2 \neq \sigma_0^2$$

es

$$\mathcal{C} = \left\{ (x_1,..,x_n) : \frac{\sum (x_i-\mu)^2}{\sigma_0^2} \leq \chi^2_{(n)\,\alpha/2} \;\text{ ó }\; \frac{\sum (x_i-\mu)^2}{\sigma_0^2} \geq \chi^2_{(n)\,1-\alpha/2} \right\}$$

Ahora supongamos que tenemos dos m.a. independientes

$x_1,..,x_n$ con distribución $N(\mu_1, \sigma^2)$ y

$y_1,..,y_m$ con distribución $N(\mu_2, \sigma^2)$

El objetivo es contrastar

$$H_0 : \mu_1 = \mu_2 \quad vs \quad H_a : \mu_1 \neq \mu_2$$

Es decir, si dos poblaciones coinciden en cuanto a la media. Por ejemplo sea X y Y el nivel de colesterol en la sangre de mujeres y hombres mexicanos de 50 años. Se quiere saber si asumiendo que el nivel de colesterol sigue una distribución normal en ambas poblaciones, las medias son iguales, es decir, los niveles de colesterol promedios de los hombres y mujeres son iguales.

Supongamos μ_1, μ_2 y σ^2 <u>desconocidas</u>.

Recordar que

$$L(\mu_1, \mu_2, \sigma^2; x_1, \dots, x_n, y_1, \dots, y_m) = f(x_1, \dots, x_n, y_1, \dots, y_m; \mu_1, \mu_2, \sigma^2)$$

$$\overset{\text{por ser independientes}}{=} f(x_1, \dots, x_n; \mu_1, \mu_2, \sigma^2)\, f(y_1, \dots, y_m; \mu_1, \mu_2, \sigma^2)$$

$$= f(x_1, \dots, x_n; \mu_1, \sigma^2)\, f(y_1, \dots, y_m; \mu_2, \sigma^2)$$

$$= L(\mu_1, \sigma^2; x_1, \dots, x_n)\; L(\mu_2, \sigma^2; y_1, \dots, y_m)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_j-\mu_2)^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_1)^2} \frac{1}{(2\pi\sigma^2)^{m/2}} e^{-\frac{1}{2\sigma^2}\sum_{j=1}^{m}(y_j-\mu_2)^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}+\frac{m}{2}}} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu_1)^2 + \sum_{j=1}^{m}(y_j-\mu_2)^2\right]}$$

Ahora. Bajo H_0 $\mu_1 = \mu_2 = D$

[la hipótesis se puede escribir como

$H_0: \underbrace{\mu_1 - \mu_2 = 0}_{\text{restricción}}, \underbrace{\mu_1, \sigma^2}_{\text{libres}}$ se estiman sólo 2 <u>parametros</u>]

/

Además

$$L(D, \sigma^2, x_1, \ldots, x_n, y_1, \ldots, y_m) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}+\frac{m}{2}}} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-D)^2 + \sum_{j=1}^{m}(y_j-D)^2\right]}$$

ahora hay que sacar los estimadores bajo H_0, es decir $\hat{\hat{D}}$ y $\hat{\hat{\sigma}}^2$.

$$\ln L(D, \sigma^2) = -\left(\frac{n}{2}+\frac{m}{2}\right)\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-D)^2 + \sum_{j=1}^{m}(y_j-D)^2\right]$$

$$\frac{\partial \ln L(D,\sigma^2)}{\partial D} = -\frac{1}{2\sigma^2}\left[-2\sum_{i=1}^{n}(x_i-D) - 2\sum_{j=1}^{m}(y_j-D)\right]$$

$$= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}x_i + \sum_{j=1}^{m}y_j - nD - mD\right]$$

$$= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}x_i + \sum_{j=1}^{m}y_j - D(n+m)\right] \quad \ldots (\mathrm{I})$$

$$\frac{\partial \ln L(D,\sigma^2)}{\partial \sigma^2} = -\frac{\left(\frac{n}{2}+\frac{m}{2}\right)\cancel{2\pi}}{\cancel{2\pi}\sigma^2} + \frac{\left[\sum_{i=1}^{n}(x_i-D)^2 + \sum_{j=1}^{m}(y_j-D)^2\right]}{2\sigma^4} \quad \ldots (\mathrm{II})$$

De I)

$$\frac{\partial \ln L(D,\sigma^2)}{\partial D} = 0 \iff \sum_{i=1}^{n}x_i + \sum_{j=1}^{m}y_j = D(m+n)$$

$$\iff D = \frac{\sum_{i=1}^{n}x_i + \sum_{j=1}^{m}y_j}{m+n}$$

De (II)

$$\frac{\partial \ln L(D,\sigma^2)}{\sigma^2} = 0 \iff \sigma^2 = \frac{\sum_{i=1}^{n}(x_i-D)^2 + \sum_{j=1}^{m}(y_j-D)^2}{n+m}$$

Se puede verificar que son maximos.

$$\hat{\hat{D}} = \frac{\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j}{m+n} = \hat{\hat{\mu}}_1 = \hat{\hat{\mu}}_2 \quad y \quad \hat{\hat{\sigma}}^2 = \frac{\sum_{i=1}^{n}(x_i-\hat{\hat{D}})^2 + \sum_{j=1}^{m}(y_j-\hat{\hat{D}})^2}{n+m}$$

Ahora para el denominador no hay restricción

$$\Rightarrow L(\mu_1,\mu_2,\sigma^2; x_1,\dots,x_n, y_1,\dots,y_m) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}+\frac{m}{2}}} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu_1)^2 + \sum_{j=1}^{m}(y_j-\mu_2)^2\right]}$$

$$\ln L(\mu_1,\mu_2,\sigma^2) = -\left(\frac{n+m}{2}\right)\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu_1)^2 + \sum_{j=1}^{m}(y_j-\mu_2)^2\right]$$

$$\frac{\partial \ln L(\mu_1,\mu_2,\sigma^2)}{\partial \mu_1} = -\frac{1}{2\sigma^2}\left(-2\sum_{i=1}^{n}(x_i-\mu_1)\right) \quad \dots \quad (1)$$

$$\frac{\partial \ln L(\mu_1,\mu_2,\sigma^2)}{\partial \mu_2} = -\frac{1}{2\sigma^2}\left(-2\sum_{j=1}^{m}(y_j-\mu_2)\right) \quad \dots \quad (2)$$

$$\frac{\partial \ln L(\mu_1,\mu_2,\sigma^2)}{\partial \sigma^2} = \frac{-\left(\frac{n+m}{2}\right)}{\cancel{2\pi}\sigma^2}(\cancel{2\pi}) + \frac{1}{2\sigma^4}\left[\sum_{i=1}^{n}(x_i-\mu_1)^2 + \sum_{j=1}^{m}(y_j-\mu_2)^2\right] \quad (3)$$

$$\frac{\partial \ln L}{\partial \mu_1} = 0 \iff \hat{\mu}_1 = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\frac{\partial \ln L}{\partial \mu_2} = 0 \iff \hat{\mu}_2 = \sum_{j=1}^{m} y_j / m$$

$$\frac{\partial \ln L}{\partial \sigma^2} = 0 \iff \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu}_1)^2 + \sum_{j=1}^{m}(y_j-\hat{\mu}_2)^2}{n+m}$$

Se puede verificar que son maximos

$$\therefore \lambda = \frac{\sup_{\{\mu_1,\mu_2,\sigma^2 \in \Theta_0\}} L(\mu_1,\mu_2,\sigma^2)}{\sup_{\{\mu_1,\mu_2,\sigma^2 \in \Theta\}} L(\mu_1,\mu_2,\sigma^2)} = \frac{L(\hat{\hat{\mu}}_1, \hat{\hat{\mu}}_2, \hat{\hat{\sigma}}^2)}{L(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}^2)}$$

$$= \frac{\frac{1}{(2\pi\hat{\hat{\sigma}}^2)^{\frac{n}{2}+\frac{m}{2}}} e^{-\frac{1}{2\hat{\hat{\sigma}}^2}\left[\sum_{i=1}^{n}(x_i-\hat{\hat{\mu}}_1)^2+\sum_{j=1}^{m}(y_j-\hat{\hat{\mu}}_2)^2\right]}}{\left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{\frac{n+m}{2}} e^{-\frac{1}{2\hat{\sigma}^2}\left[\sum_{i=1}^{n}(x_i-\hat{\mu}_1)^2+\sum_{j=1}^{m}(y_j-\hat{\mu}_2)^2\right]}}$$

$$= \left(\frac{\hat{\sigma}^2}{\hat{\hat{\sigma}}^2}\right)^{\frac{n+m}{2}} e^{-\frac{n+m}{2}+\frac{n+m}{2}} = \left(\frac{\hat{\sigma}^2}{\hat{\hat{\sigma}}^2}\right)^{\frac{n+m}{2}}$$

$$= \left(\frac{\sum_{i=1}^{n}(x_i-\hat{\mu}_1)^2+\sum_{j=1}^{m}(y_j-\hat{\mu}_2)^2}{\sum_{i=1}^{n}(x_i-\hat{D})^2+\sum_{j=1}^{m}(y_j-\hat{D})^2}\right)^{\frac{n+m}{2}}$$

Se puede verificar que

$$\lambda = \left(1+\frac{\frac{nm}{m+n}(\hat{\mu}_1-\hat{\mu}_2)^2}{\sum_{i=1}^{n}(x_i-\hat{\mu}_1)^2+\sum_{j=1}^{m}(y_j-\hat{\mu}_2)^2}\right)^{-\frac{n+m}{2}}$$

Además $T = \dfrac{\sqrt{\frac{nm}{n+m}}(\hat{\mu}_1-\hat{\mu}_2)}{\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\hat{\mu}_1)^2+\sum_{j=1}^{m}(y_j-\hat{\mu}_2)^2}{n+m-2}}} \sim t_{n+m-2}$

De donde se obtiene

$$C = \left\{(x_1,\ldots,x_n) : \left|\frac{\sqrt{\frac{nm}{n+m}}(\hat{\mu}_1-\hat{\mu}_2)}{\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\hat{\mu}_1)^2+\sum_{j=1}^{m}(y_j-\hat{\mu}_2)^2}{n+m-2}}}\right| > t^{1-\alpha/2}_{n+m-2}\right\}$$

Supongamos ahora

$X_1, \dots, X_n$ tiene distribución $N(\mu_1, \sigma_1^2)$

$Y_1, \dots, Y_m$ con $N(\mu_2, \sigma_2^2)$, las muestras independientes.

Hay dos objetivos

I) $H_0: \sigma_1^2 = \sigma_2^2$ vs $H_a: \sigma_1^2 \neq \sigma_2^2$

Cuando no se rechaza se puede usar la prueba anterior para

$H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 \neq \mu_2$ con $\underline{\sigma_1^2 = \sigma_2^2 = \sigma^2}$.

II) $H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 \neq \mu_2$.

Se usa cuando se rechaza H_0 en I). En esta prueba es difícil encontrar la distribución exacta, por lo que se usa la propiedad asintótica de $-2 \ln \lambda$, donde λ es el cociente de verosimilitud generalizado.

I. Sea μ_1, μ_2, σ_1^2 y σ_2^2 todos desconocidos

Se busca contrastar

$$H_0: \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_a: \sigma_1^2 \neq \sigma_2^2$$

Aquí

$$L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; x_1, \dots, x_n, y_1, \dots, y_m) = L(\mu_1, \sigma_1^2; x_1, \dots, x_n)\, L(\mu_2, \sigma_2^2; y_1, \dots, y_m)$$

$$= \frac{1}{(2\pi\sigma_1^2)^{n/2}} e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i-\mu_1)^2} \frac{1}{(2\pi\sigma_2^2)^{m/2}} e^{-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n}(y_j-\mu_2)^2}$$

Bajo H_0, $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$\Rightarrow L(\mu_1, \mu_2, \sigma^2; x_1, \dots, x_n, y_1, \dots, y_m) = \frac{1}{(2\pi\sigma^2)^{\frac{n+m}{2}}} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu_1)^2 + \sum_{j=1}^{m}(y_j-\mu_2)^2\right]}$$

y se pueden encontrar los estimadores (ver ejemplo anterior)

$$\hat{\hat{\mu}}_1 = \sum_{i=1}^{n} x_i/n = \bar{x}, \quad \hat{\hat{\mu}}_2 = \sum_{j=1}^{m} y_j/m = \bar{y}, \quad \hat{\hat{\sigma}}^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{j=1}^{m}(y_j-\bar{y})^2}{n+m}$$

Para el caso no restringido se puede ver que cada función de verosimilitud se puede maximizar por separado

$$\Rightarrow \hat{\mu}_1 = \sum_{i=1}^{n} x_i/n = \bar{x}, \quad \hat{\mu}_2 = \sum_{j=1}^{m} y_j/m = \bar{y}, \quad \hat{\sigma}_1^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}, \quad \hat{\sigma}_2^2 = \frac{\sum_{j=1}^{m}(y_j-\bar{y})^2}{m}$$

$$\Rightarrow \lambda = \frac{\sup_{\{\mu_1, \mu_2, \sigma_1^2 = \sigma_2^2\}} L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}{\sup_{\{\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\}} L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)} = \frac{\frac{1}{(2\pi\hat{\hat{\sigma}}^2)^{n/2}} e^{-\frac{1}{2\hat{\hat{\sigma}}^2}\sum_{i=1}^{n}(x_i-\bar{x})^2} \frac{1}{(2\pi\hat{\hat{\sigma}}^2)^{m/2}} e^{-\frac{1}{2\hat{\hat{\sigma}}^2}\sum_{j=1}^{m}(y_j-\bar{y})^2}}{\left(\frac{1}{2\pi\hat{\sigma}_1^2}\right)^{n/2} e^{-\frac{1}{2\hat{\sigma}_1^2}\sum_{i=1}^{n}(x_i-\bar{x})^2} \left(\frac{1}{2\pi\hat{\sigma}_2^2}\right)^{m/2} e^{-\frac{1}{2\hat{\sigma}_2^2}\sum_{j=1}^{m}(y_j-\bar{y})^2}}$$

$$= \left(\frac{\hat{\sigma}_1^2}{\hat{\hat{\sigma}}^2}\right)^{n/2} \left(\frac{\hat{\sigma}_2^2}{\hat{\hat{\sigma}}^2}\right)^{\frac{m}{2}} \frac{e^{-\frac{1}{2\hat{\hat{\sigma}}^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{j=1}^{m}(y_j-\bar{y})^2\right]}}{e^{-\frac{n}{2}} e^{-\frac{m}{2}}}$$

$$= \left(\frac{\hat{\sigma}_1^2}{\hat{\hat{\sigma}}^2}\right)^{n/2} \left(\frac{\hat{\sigma}_2^2}{\hat{\hat{\sigma}}^2}\right)^{\frac{m}{2}} e^{-\frac{m+n}{2} + \frac{n}{2} + \frac{m}{2}}$$

$$= \left(\frac{\hat{\sigma}_1^2}{\hat{\hat{\sigma}}^2}\right)^{n/2} \left(\frac{\hat{\sigma}_2^2}{\hat{\hat{\sigma}}^2}\right)^{m/2}$$

se puede verificar

$$= \frac{(m+n)^{n+m/2}}{m^{m/2}\, n^{n/2}} \frac{\left[\left(\frac{n-1}{m-1}\right) f\right]^{m/2}}{\left(1 + \left(\frac{n-1}{m-1}\right) f\right)^{(m+n)/2}}$$

donde $f = \left[\sum_{i=1}^{n} (x_i-\bar{x})^2 / n-1\right] \Big/ \left[\sum_{j=1}^{m} (y_j-\bar{y})^2 / m-1\right]$

Bajo H_0 $\quad f \sim F_{n-1,\, m-1}$

y se puede verificar

que $\quad \lambda \le \lambda_0$

$\iff \quad f < k_1 \quad$ o $\quad f > k_2$

para determinados k_1 y k_2. Una aproximación es elegir k_1 y k_2 tal que

$P(f < k_1) = \alpha/2 \quad$ y $\quad P(f > k) = \alpha/2$

$\Rightarrow \quad k_1 = F_{n-1,\, m-1,\, \alpha/2} \quad$ y $\quad k_2 = F_{n-1,\, m-1,\, 1-\alpha/2}$

$$\therefore \quad \mathcal{C} = \left\{ (x_1, \ldots, x_n) : \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \right] \Big/ \left(\frac{\sum_{j=1}^{m} (y_j - \bar{y})^2}{m-1} \right) < F_{n-1, m-1, \alpha/2} \quad o \right.$$

$$\left. \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \Big/ \frac{\sum_{j=1}^{m} (y_j - \bar{y})^2}{m-1} \right] > F_{n-1, m-1, 1-\alpha/2} \right\}$$

II. Sea $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ todas desconocidas

$H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 \neq \mu_2$

Aquí nuevamente

$$L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{(2\pi\sigma_1^2)^{n/2}} e^{-\frac{1}{2\sigma_1^2} \sum_{i=1}^{n} (x_i - \mu_1)^2} \cdot \frac{1}{(2\pi\sigma_2^2)^{m/2}} e^{-\frac{1}{2\sigma_2^2} \sum_{j=1}^{m} (y_j - \mu_2)^2}$$

Bajo H_0, $\mu_1 = \mu_2 = D$

$$L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{(2\pi\sigma_1^2)^{n/2}} e^{-\frac{1}{2\sigma_1^2} \sum_{i=1}^{n} (x_i - D)^2} \cdot \frac{1}{(2\pi\sigma_2^2)^{2}} e^{-\frac{1}{2\sigma_2^2} \sum_{j=1}^{m} (y_j - D)^2}$$

$$\ln L = -\frac{n}{2} \ln(2\pi\sigma_1^2) - \frac{n}{2} \ln(2\pi\sigma_2^2) - \frac{1}{2\sigma_1^2} \sum_{i=1}^{n} (x_i - D)^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^{m} (y_j - D)^2$$

$$\frac{\partial \ln L}{\partial D} = -\frac{1}{2\sigma_1^2} \left(-2 \sum_{i=1}^{n} (x_i - D) \right) - \frac{1}{2\sigma_2^2} \left(-2 \sum_{j=1}^{m} (y_j - D) \right)$$

$$= \frac{\sum x_i - nD}{\sigma_1^2} + \frac{\sum y_j - mD}{\sigma_2^2} \quad \text{--- (1)}$$

$$\frac{\partial \ln L}{\partial \sigma_1^2} = -\frac{n}{2} \frac{(2\pi)}{2\pi\sigma_1^2} + \frac{1}{2\sigma_1^4} \sum_{i=1}^{n} (x_i - D)^2 \quad \text{--- (2)}$$

$$\frac{\partial \ln L}{\partial \sigma_2^2} = -\frac{n}{2} \frac{(2\pi)}{2\pi\sigma_2^2} + \frac{1}{2\sigma_2^4} \sum_{j=1}^{m} (y_j - D)^2 \quad \text{--- (3)}$$

Igualando (2) y (3) a cero

$$\hat{\hat{\sigma}}_1^2 = \frac{\sum_{i=1}^{n}(x_i - D)^2}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n} + (\bar{x}-D)^2$$

$$\hat{\hat{\sigma}}_2^2 = \frac{\sum_{j=1}^{m}(y_j - D)^2}{m} = \frac{\sum_{j=1}^{n}(y_j-\bar{y})^2}{m} + (\bar{y}-D)^2$$

Mientras que (1) se puede escribir como

$$\frac{n(\bar{x}-D)}{\hat{\hat{\sigma}}_1^2} + \frac{m(\bar{y}-D)}{\hat{\hat{\sigma}}_2^2} = 0$$

Al sustituir $\hat{\hat{\sigma}}_1^2$ y $\hat{\hat{\sigma}}_2^2$ en

$$L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{(2\pi\hat{\hat{\sigma}}_1^2)^{n/2}} e^{-\frac{1}{2\hat{\hat{\sigma}}_1^2}\sum_{i=1}^{n}(x_i-D)^2} \frac{1}{(2\pi\hat{\hat{\sigma}}_2^2)^{m/2}} e^{-\frac{1}{2\hat{\hat{\sigma}}_2^2}\sum_{j=1}^{m}(y_j-D)^2}$$

$$= \frac{e^{-\frac{1}{2}(n+m)}}{(2\pi)^{\frac{n}{2}+\frac{m}{2}}(\hat{\hat{\sigma}}_1^2)^{n/2}(\hat{\hat{\sigma}}_2^2)^{m/2}}$$

Notar que aun falta resolver (1), se puede hacer numéricamente

En el caso no restringido del denominador

$$L(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2) = \frac{1}{(2\pi\hat{\sigma}_1^2)^{n/2}} \frac{1}{(2\pi\hat{\sigma}_2^2)^{m/2}} e^{-\frac{1}{2}(n+m)}$$

$$\Rightarrow \quad \lambda = \left(\frac{\hat{\sigma}_1^2}{\hat{\hat{\sigma}}_1^2}\right)^{n/2}\left(\frac{\hat{\sigma}_2^2}{\hat{\hat{\sigma}}_2^2}\right)^{m/2}$$

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}, \quad \hat{\sigma}_2^2 = \frac{\sum_{j=1}^{m}(y_j-\bar{y})^2}{m}$$

Se puede usar la propiedad asintótica de

$$-2\ln\lambda = -2\left[\frac{n}{2}\ln\left(\frac{\hat{\sigma}_1^2}{\hat{\hat{\sigma}}_1^2}\right) + \frac{m}{2}\ln\left(\frac{\hat{\sigma}_2^2}{\hat{\hat{\sigma}}_2^2}\right)\right] \sim \chi^2_r$$

aquí $r = \underbrace{k}_{\text{parámetros denominador}} - \underbrace{k'}_{\text{parámetros numerador}} = 4 - 3 = 1$

$\therefore$ Se rechaza H_0 si

$$-n\ln\left(\frac{\hat{\sigma}_1^2}{\hat{\hat{\sigma}}_1^2}\right) - m\ln\left(\frac{\hat{\sigma}_2^2}{\hat{\hat{\sigma}}_2^2}\right) > \chi^2_{(1),\,1-\alpha}$$