Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2019

EYP 2405/EYP 2114
Métodos Estadísticos / Inferencia Estadística
Clase de Ejercicios 3

1. Considere $X_1, \ldots, X_n$ i.i.d $\text{Exp}(\lambda_1)$ independiente de $Y_1, \ldots, Y_n$ i.i.d. $\text{Exp}(\lambda_2)$.

   a) Derive el TRV para $H_0 : \lambda_1 = \lambda_2$ vs $H_1 : \lambda_1 \neq \lambda_2$.

   b) Muestre que el test derivado en a) puede expresarse en términos del estadístico $T = \sum_{i=1}^n X_i / (\sum_{i=1}^n X_i + \sum_{i=1}^n Y_i)$.

   c) Encuentre la distribución de T bajo H_0 y explique cómo utilizar este resultado para obtener explícitamente el valor para el cual el TRV rechaza H_0.

2. Considere una muestra aleatoria (i.i.d.) $X_1, \ldots, X_n$ de la distribución $N(0, \sigma^2)$. En base a dicha muestra, se quiere realizar el test de las hipótesis $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 = \sigma_1^2$.

   a) Muestre que el test UMP de nivel $0 < \alpha < 1$ consiste en rechazar H_0 si $\sum_{i=1}^n X_i^2 > c$.

   b) Para el caso en que $\sigma_0^2 = 1$ y $\sigma_1^2 = 100$, encuentre explícitamente c.

   c) Calcule la potencia de este test y compárela con la que se obtiene cuando se considera $H_1' : \sigma^2 = 4$

   d) Si se observa $\sum_{i=1}^2 X_i^2 = 6$, calcule el valor-p asociado al test.

3. Considere $Y \sim f(x \mid \theta)$ con

$$f(x \mid \theta) = [2(1-\theta)x + \theta] I\{0 \leq x \leq 1\}, \quad 0 \leq \theta \leq 2$$

   a) Suponga que se quieren evaluar las hipótesis $H_0 : \theta = 0$ vs $H_1 : \theta = 2$. Encuentre el test UMP de tamaño $0 < \alpha < 1$ para este caso y calcule su potencia.

   b) Suponga ahora que interesa el test $H_0 : \theta \geq 1$ vs $H_1 : \theta < 1$. Considere un procedimiento de test que rechaza si $X \geq 0.9$. Calcule la función potencia y tamaño respectivo del test.

4. Considere una muestra $X_1, \ldots, X_n$ con densidad común

$$f(x \mid \theta) = \theta x^{\theta - 1} I\{0 \leq x \leq 1\}, \quad \theta > 0$$

   a) Encuentre la forma de la región de rechazo para el test UMP de las hipótesis $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ para $\theta_0 > 0$.

   b) Considerando un nivel $0 < \alpha < 1$, exprese la región de rechazo del test en a) en términos de percentiles de una distribución conocida, y calcule la función de potencia del test resultante. **Hint:** ¿cuál e sla distribución de $Y_i = 2\theta \log(1/X_i)$?

Clase de ejercicios 3

1) Como las muestras son independientes se tiene que:

$$f(\underset{\sim}{x},\underset{\sim}{y} \mid \lambda_1,\lambda_2) = f(\underset{\sim}{x}\mid\lambda_1)\, f(\underset{\sim}{x}\mid\lambda_2)$$

$$= \left[\prod_{i=1}^{n} \lambda_1 \exp\{-\lambda_1 x_i\}\right]\left[\prod_{i=1}^{n} \lambda_2 \exp\{-\lambda_2 y_i\}\right]$$

$$= (\lambda_1\lambda_2)^n \exp\{-\lambda_1 \Sigma x_i\} \exp\{-\lambda_2 \Sigma y_i\}$$

$$\lambda(\underset{\sim}{x},\underset{\sim}{y}) = \frac{\sup\limits_{\lambda_1,\lambda_2\in\Theta_0} L(\lambda_1,\lambda_2 \mid \underset{\sim}{x},\underset{\sim}{y})}{\sup\limits_{\lambda_1,\lambda_2\in\Theta} L(\lambda_1,\lambda_2 \mid \underset{\sim}{x},\underset{\sim}{y})} = \frac{L(\hat{\lambda}_0 \mid \underset{\sim}{x},\underset{\sim}{y})}{L(\hat{\lambda}_1,\hat{\lambda}_2 \mid \underset{\sim}{x},\underset{\sim}{y})}$$

en donde, bajo H_0, $\lambda_1=\lambda_2=\lambda_0$ y $\hat{\lambda}_0$ es el EMV, y $\hat{\lambda}_1, \hat{\lambda}_2$ son los EMV no restringidos.

- calculemos primero $\hat{\lambda}_0$.

$$L(\lambda_0 \mid \underset{\sim}{x},\underset{\sim}{y}) = \lambda_0^{2n} \exp\{-\lambda_0(\Sigma x_i + \Sigma y_i)\}$$

$$\ell(\lambda_0 \mid \underset{\sim}{x},\underset{\sim}{y}) = 2n \log\lambda_0 - \lambda_0(\Sigma x_i + \Sigma y_i)$$

$$\ell'(\lambda_0 \mid \underset{\sim}{x},\underset{\sim}{y}) = 0 \iff \frac{2n}{\lambda_0} - (\Sigma x_i + \Sigma y_i) = 0$$

$$\iff \frac{2n}{\lambda_0} = \Sigma x_i + \Sigma y_i$$

$$\iff \hat{\lambda}_0 = \frac{2n}{\Sigma x_i + \Sigma y_i}$$

$$\ell''(\lambda_0 \mid \underset{\sim}{x},\underset{\sim}{y}) = -\frac{2n}{\lambda_0^2}\bigg|_{\lambda_0=\hat{\lambda}_0} < 0 \implies \hat{\lambda}_0 \text{ es máximo.}$$

- calculemos ahora $\hat{\lambda}_1$ y $\hat{\lambda}_2$.

$$\ell(\lambda_1,\lambda_2 \mid \underset{\sim}{x},\underset{\sim}{y}) = n\log\lambda_1 + n\log\lambda_2 - \lambda_1\Sigma x_i - \lambda_2 \Sigma y_i$$

$$\frac{\partial \ell(\;)}{\partial \lambda_1} = 0 \iff \frac{n}{\lambda_1} - \Sigma x_i = 0 \iff \hat{\lambda}_1 = \frac{n}{\Sigma x_i}$$

$$\frac{\partial \ell(\;)}{\partial \lambda_2} = 0 \iff \frac{n}{\lambda_2} - \Sigma y_i = 0 \iff \hat{\lambda}_2 = \frac{n}{\Sigma y_i}$$

②

entonces

$$\lambda(\underset{\sim}{x},\underset{\sim}{y}) = \frac{\left(\frac{2n}{\sum x_i+\sum y_i}\right)^{2n} \cancel{\exp\left\{-\left(\frac{2n}{\sum x_i+\sum y_i}\right)(\sum x_i+\sum y_i)\right\}}}{\left(\frac{n}{\sum x_i}\right)^{n}\left(\frac{n}{\sum y_i}\right)^{n} \cancel{\exp\left\{-\left(\frac{n}{\sum x_i}\right)\sum x_i-\left(\frac{n}{\sum y_i}\right)\sum y_i\right\}}}$$

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

y la región de rechazo es $R=\{(\underset{\sim}{x},\underset{\sim}{y})\in(\mathbb{R}^+)^2 : \lambda(\underset{\sim}{x},\underset{\sim}{y})\le c\}$

$0\le c\le 1$

b) si $T=\dfrac{\sum x_i}{\sum x_i+\sum y_i}$, entonces

$$\lambda(\underset{\sim}{x},\underset{\sim}{y}) = (2n)^{2n}\left(\frac{1}{\sum x_i+\sum y_i}\right)^{n}\left(\frac{1}{\sum x_i+\sum y_i}\right)^{n}\left[\left(\frac{\sum x_i}{n}\right)^{n}\left(\frac{\sum y_i}{n}\right)^{n}\right]$$

$$= \frac{(2n)^{2n}}{n^n n^n}\left(\frac{\sum x_i}{\sum x_i+\sum y_i}\right)^{n}\left(\frac{\sum y_i}{\sum x_i+\sum y_i}\right)^{n}$$

$$= \frac{(2n)^{2n}}{n^{2n}}\left[T^n(1-T)^n\right]$$

$$= \left(\frac{2n}{n}\right)^{2n}\left[T^n(1-T)^n\right]$$

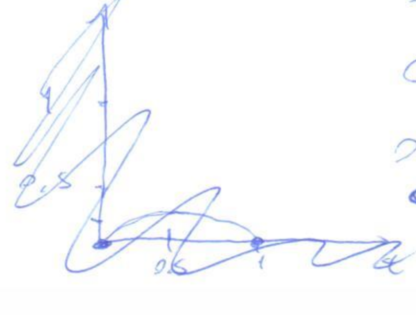
$$= 4^n\left[T^n(1-T)^n\right] \qquad T(1-T)\le c$$

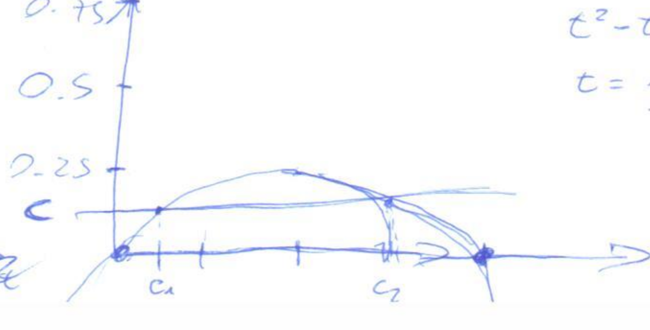
$$= \left[4T(1-T)\right]^n$$

con $R=\{t\in(0,1) ; \lambda(t)\le c\}$, $0\le c\le 1$ $\qquad 1-4c>0$

$\frac{1}{4}>c$

$f(t)=t(1-t)$





$t-t^2-c=0$

$t^2-t+c=0$

$t=\dfrac{1\pm\sqrt{1-4c}}{2}$

$\lambda(t)<c_1$ ó

$\lambda_2>c_2$

b) Bajo H_0: $\lambda_1 = \lambda_2 = \lambda \Rightarrow$

$$\sum x_i \sim \text{Gamma}(n, \lambda)$$
$$\sum y_i \sim \text{Gamma}(n, \lambda)$$
$$\sum x_i + \sum y_i \sim \text{Gamma}(2n, \lambda)$$
$$\Rightarrow T = \frac{\sum x_i}{\sum x_i + \sum y_i} \sim \beta\text{eta}(n, 2n)$$

Para un test de nivel α, tenemos que

$$P_{\theta_0}(R) \leq \alpha \iff P_{\theta_0}(T \leq c_1) + P(T \geq c_2) \leq \alpha$$

Tomamos $P_{\theta_0}(T \leq c_1) = P_{\theta_0}(T \geq c_2) = \frac{\alpha}{2}$, entonces

$$P_{\theta_0}(T \leq c_1) = \frac{\alpha}{2} \Rightarrow c_1 = \gamma^{-1}\left(\frac{\alpha}{2}\right)$$ con $\gamma(\cdot)$ la acumulada de una Beta$(n, 2n)$

$$P_{\theta_0}(T \geq c_2) = \frac{\alpha}{2} \iff 1 - \frac{\alpha}{2} = P_{\theta_0}(T \leq c_2)$$

$$c_2 = \gamma^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Luego el test rechaza para

$$T \leq \gamma^{-1}\left(\frac{\alpha}{2}\right) \text{ ó } T \geq \gamma^{-1}\left(1 - \frac{\alpha}{2}\right)$$ a nivel α.

(2) se tiene que

$$f(\underset{\sim}{x}|\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum x_i^2\right\}$$

Además,

$$\frac{f(\underset{\sim}{x}|\theta_1)}{f(\underset{\sim}{x}|\theta_0)} = \frac{f(\underset{\sim}{x}|100)}{f(\underset{\sim}{x}|1)} = \frac{(200\pi)^{-n/2}\exp\left\{-\frac{1}{200}\sum x_i^2\right\}}{(2\pi)^{-n/2}\exp\left\{-\frac{1}{2}\sum x_i^2\right\}}$$

$$= (100)^{-n/2}\exp\left\{\sum x_i^2\left(\frac{1}{2}-\frac{1}{200}\right)\right\}$$

$$= (100)^{-n/2}\exp\left\{\frac{99}{200}\sum x_i^2\right\}$$

Ahora $\frac{f(\underset{\sim}{x}|\theta_1)}{f(\underset{\sim}{x}|\theta_0)} > k \iff \exp\left\{\frac{99}{200}\sum x_i^2\right\} > k'$

$\iff \frac{99}{200}\sum x_i^2 > k''$

$\iff \sum x_i^2 > k'''$ i)

además

ii) $\alpha = P_{\theta_0}(\underset{\sim}{x}\in R) \iff \alpha = P(\sum x_i^2 > k''' | \sigma^2=1)$

y como bajo H_0, $\sum x_i^2 \sim \chi^2_{(n)}$, entonces

$$1-\alpha = P(\sum x_i^2 \le k''') \Rightarrow k''' = \chi^2_{(n,1-\alpha)}$$

Luego, por Teorema de Neyman-Pearson (se cumplen i) y ii)) el test es UMP

b) Listo!

~~c) Potencia del test es por def $\beta(\theta) = P_\theta(\underset{\sim}{x}\in R)$, $\theta\in\Theta$~~

~~$\Theta = \{\sigma_0^2\}\cup\{\sigma_1^2\}$ $\beta(\theta_0) = P_{\theta_0}(\underset{\sim}{x}\in R)$, $\theta_0\in\Theta_0$~~

~~$\beta(\theta_1) = P_{\theta_1}(\underset{\sim}{x}\in R)$, $\theta_1\in\Theta_1$~~

~~$P_{\sigma^2}(\sum x_i^2 > c) = 1 - P_{\sigma^2}(\sum x_i^2 \le c)$~~

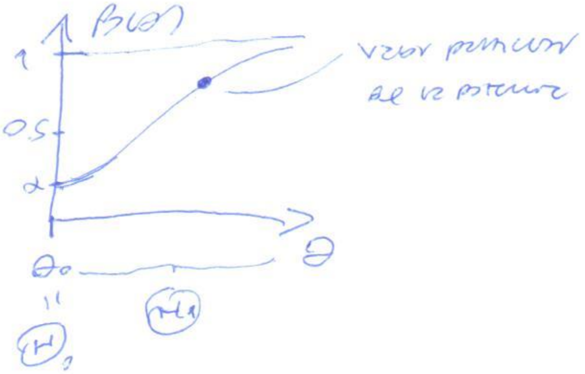
~~$\alpha = 1 - P_{\sigma^2=1}(W\le c)$ $W\sim\chi^2_{(n)}$~~

~~$P_{\sigma^2=1}(W\le c) = 1-\alpha$~~

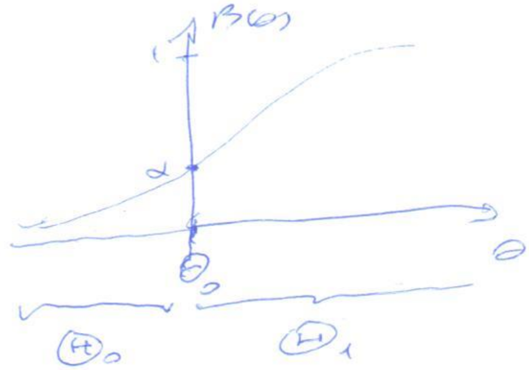
c) La potencia de un test es la probabilidad de rechazar H_0 cuando esta es falsa ($P(\underset{\sim}{X} \in R \mid H_1)$). ← cuidado con esta frase!

La función potencia de un test es $\beta(\theta) = P_\theta(\underset{\sim}{X} \in R)$ $\theta \in \Theta$

$H_0: \theta = \theta_0$
$H_1: \theta > \theta_0$



$H_0: \theta \le \theta_0$
$H_1: \theta > \theta_0$



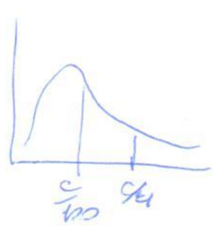
En este caso $\Theta_1 = \{100\}$, entonces $\left(\frac{X_i}{10} \sim N(0,1) \to \frac{\sum X_i^2}{100} \sim \chi^2_{(n)} \right)$

$$P\left(\sum X_i^2 > c \mid \sigma^2 = 100\right) = P\left(\frac{\sum X_i^2}{100} > \frac{c}{100}\right)$$
$$= 1 - P\left(W \le \frac{c}{100}\right) \quad W \sim \chi^2_{(n)}$$

Si $H_1': \sigma^2 = 4$, entonces la potencia del test es

$$P\left(\sum X_i^2 > c \mid \sigma^2 = 4\right) = 1 - P\left(W \le \frac{c}{4}\right)$$

Como $\frac{c}{4} > \frac{c}{100}$ entonces $P\left(W \le \frac{c}{4}\right) > P\left(W \le \frac{c}{100}\right)$



y entonces $1 - P\left(W \le \frac{c}{4}\right) < 1 - P\left(W \le \frac{c}{100}\right)$

→ la potencia es menor para el caso H_1'

d) $P\left(\sum_{i=1}^{2} X_i^2 > 6 \mid \sigma^2 = 1\right) = P(W > 6)$ $W \sim \chi^2_{(2)}$ $\chi^2_{(2)} = Ga\left(\frac{2}{2}, 2\right) = \exp(2)$

$$= 1 - P(W \le 6)$$
$$= 1 - \left[1 - e^{-6/2}\right]$$
$$= e^{-3} \simeq 0.0498$$

3) El test UMP tiene región de rechazo

$$\frac{f(x|2)}{f(x|0)} > k \iff \frac{-2x+2}{2x} = \frac{1-x}{x} > k \iff 1-x > xk$$
$$\iff 1 > (k+1)x$$
$$\iff X < c$$

y $\alpha = P(X<c|\theta=0) = \int_0^c 2x\,dx = x^2\Big|_0^c = c^2 \Rightarrow c = \sqrt{\alpha}$

luego el test UMP de nivel α rechaza si $X < \sqrt{\alpha}$.

La potencia es

$$P(Y<\sqrt{\alpha}|\theta=2) = \int_0^{\sqrt{\alpha}} 2(1-y)\,dy = \int_0^{\sqrt{\alpha}} 2\,dy - \int_0^{\sqrt{\alpha}} 2y$$

$$= 2\sqrt{\alpha} - \alpha.$$

b) Si $R=\{x: x \geq 0.9\}$ entonces el tamaño del test es

$$\sup_{\theta\in\Theta_0} \underbrace{P(X\geq 0.9|\theta)}_{\beta(\theta)} = \sup_{1\leq\theta\leq 2} \int_{0.9}^1 \{2(1-\theta)y+\theta\}\,dy$$
$$= \quad " \quad \Big[(1-\theta)y^2+\theta y\Big]_{0.9}^1$$
$$= \quad " \quad \Big[1-\theta+\theta-(1-\theta)(0.9)^2-0.9\theta\Big]$$
$$= \quad " \quad \Big[1-0.81+0.81\theta-0.9\theta\Big]$$
$$= \sup_{1\leq\theta\leq 2} \underline{\Big[0.19-0.09\theta\Big]}$$

Si $\theta=1 \rightarrow \beta(\theta)=0.10$
Si $\theta=2 \rightarrow \beta(\theta)=0.01$

luego $\sup_{\theta\in\Theta_0} \beta(\theta) = 0.10 \equiv \alpha$ es el tamaño del test

La función potencia es $\beta(\theta) = P(Y\leq 0.9|\theta) = 0.19-0.09\theta, \quad \theta<1$

(4) $f(\underset{\sim}{x}|\theta) = \prod_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n \left[\prod_{i=1}^{n} x_i\right]^{\theta-1}$

a) Dados $0<\theta_1<\theta_2$, se tiene

$$\frac{f(\underset{\sim}{x}|\theta_2)}{f(\underset{\sim}{x}|\theta_1)} = \frac{\theta_2^n \left(\prod_{i=1}^{n} x_i\right)^{\theta_2-1}}{\theta_1^n \left(\prod_{i=1}^{n} x_i\right)^{\theta_1-1}} = \left(\frac{\theta_2}{\theta_1}\right)^n \left(\prod x_i\right)^{\theta_2-\theta_1}$$

que es creciente en $T = \prod x_i$. Luego el modelo tiene RVM y por teorema de K-R, el test UMP para $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0$ tiene

$$R = \{ \underset{\sim}{x} \in \mathcal{X} : T > c \}$$

al nivel $\alpha = P_{\theta_0}(T > t_0)$.

b) Primero, notar que

$$F_X(x) = \begin{cases} 0 & ; x<0 \\ x^\theta & , 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

Luego, si $Y_i = 2\theta \log\left(\frac{1}{x_i}\right)$, entonces

$$\begin{aligned} P(Y_i \leq t) &= P\left(2\theta \log\left(\tfrac{1}{x_i}\right) \leq t\right) \\ &= P\left(\log\left(\tfrac{1}{x_i}\right) \leq \tfrac{t}{2\theta}\right) \\ &= P\left(\tfrac{1}{x_i} \leq \exp\left\{\tfrac{t}{2\theta}\right\}\right) \\ &= 1 - P\left(x_i \leq \exp\left\{-\tfrac{t}{2\theta}\right\}\right) \\ &= 1 - \exp\left\{-\tfrac{t}{2\theta}\right\}^\theta \\ &= 1 - \exp\left\{-\tfrac{t}{2}\right\} \implies Y_i \sim \exp(2) \equiv \text{Gamma}(1,2) \\ &\qquad\qquad \equiv \text{Gamma}\left(\tfrac{2}{2},2\right) \\ &\qquad\qquad \equiv \chi^2_{(2)} \end{aligned}$$

Luego, dado $0<\alpha<1$ se tiene:

$$\alpha = P_{\theta_0}\left(\prod x_i > c\right) = P_{\theta_0}\left(\frac{1}{c} > \prod \frac{1}{x_i}\right)$$

$$= P_{\theta_0}\left(\log \frac{1}{c} > \sum \log \frac{1}{x_i}\right)$$

$$= P_{\theta_0}\left(2\theta_0 \log \frac{1}{c} > \sum 2\theta_0 \log \frac{1}{x_i}\right)$$

y como bajo $\theta = \theta_0$ tenemos que

$$\sum 2\theta_0 \log\left(\frac{1}{x_i}\right) \sim \chi^2_{(2n)},$$

se tiene que $2\theta_0 \log\left(\frac{1}{c}\right) = \chi^2_{(2n,\alpha)}$

$$\Rightarrow \quad c' = \exp\left\{\frac{-\chi^2_{(2n,\alpha)}}{2\theta_0}\right\}$$

$$y \quad R = \{\underline{x} \in \mathcal{X} : \prod x_i > c'\}$$

Finalmente, si $\theta > 0$,

$$\beta(\theta) = P_\theta\left(\prod x_i > c'\right) = P_\theta\left(\frac{1}{c'} > \prod\left(\frac{1}{x_i}\right)\right)$$

$$= P_\theta\left(2\theta \log \frac{1}{c'} > 2\theta \sum \log \frac{1}{x_i}\right)$$

$$= P\left(W < 2\theta \log\left(\frac{1}{c'}\right)\right)$$

con $W \sim \chi^2_{(2n)}$