Mock Exam 2: Inf Estad Oto 2021

ITAM
Departamento de Estadística

Inferencia Estadística– Examen de Prueba 2
Parte Abierta

1. Suponga que $X_1, ..., X_n$ es una m.a. con $f_{X_i}(x_i|\theta) = \theta x_i^{\theta-1}\mathbb{I}_{(0,1)}(x_i)$, $\theta > 0$. Se desea probar $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ con $\theta_1 > \theta_0$

   (a) Mediante el lema de Neyman-Pearson, demuestre que la región de rechazo correspondiente la prueba uniformemente más potente es de la forma $RR = \left\{ -\sum_i^n ln(X_i) \leq \lambda \right\}$.

   (b) Demuestre que para una prueba de tamaño α, el valor crítico de la región de rechazo del inciso anterior es $\lambda = \frac{\chi^2_{2n,(1-\alpha)}}{2\theta_0}$. $\Big[$Hint: Demuestra que $-ln(X_i) \sim Exp(\frac{1}{\theta})$ y recuerda $T \sim Gamma(\alpha, \beta)$ entonces $\frac{2T}{\beta} \sim \chi^2_{2n}$ $\Big]$

a) La prueba uniformemente más potente es H_0: $\theta = \theta_0$ vs. H_1^* : $\theta = \theta_1$ para $\theta_0 < \theta_1$. Del Lema de Neyman-Pearson se sabe que $RR = \left\{ \frac{L(\theta_0)}{L(\theta_1)} \leq k \right\}$ para $k > 0$.

$L(\theta) = \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1}$, entonces $\frac{L(\theta_0)}{L(\theta_1)} = \left(\frac{\theta_0}{\theta_1} \right)^n \left(\prod_{i=1}^n x_i \right)^{\theta_0-\theta_1} \leq k \Leftrightarrow$

$(\theta_0 - \theta_1)\sum_{i=1}^n \ln(x_i) \leq k\left(\frac{\theta_1}{\theta_0}\right)^n$, y como $\theta_0 - \theta_1 < 0$, $\sum_{i=1}^n \ln(x_i) \geq \frac{k}{\theta_0 - \theta_1}\left(\frac{\theta_1}{\theta_0}\right)^n = -\lambda$, por lo tanto $RR = \left\{ -\sum_{i=1}^n \ln(X_i) \leq \lambda \right\}$.

b)Para la distribución de muestreo: $X_i \in (0, 1) \Rightarrow Y_i = g(X_i) = -\ln(X_i) > 0$ y
$g^{-1}(X_i) = e^{-X_i}$,

entonces para $y_i > 0$ $f_{Y_i}(y_i) = f_{X_i}\left(g^{-1}(y_i)\right)\left|\frac{d}{dy_i} g^{-1}(y_i)\right| = \theta\left(e^{-y_i}\right)^{\theta-1}\left|-e^{-y_i}\right| = \theta\, e^{-\theta\, y_i}$, es

decir, $-\ln(X_i) \sim Exp(\frac{1}{\theta}) \Rightarrow S = -\sum_{i=1}^n \ln(X_i) \sim Gamma(n, \frac{1}{\theta}) \Rightarrow$

$W = \frac{2S}{\frac{1}{\theta}} = 2\theta\, S \sim \chi^2(2n)$.

Para que la prueba sea de tamaño α:

$$\alpha = P[S \leq \lambda | \theta = \theta_0] = P[2\theta_0 S \leq 2\theta_0 \lambda] = P[W \leq 2\theta_0 \lambda],$$

entonces $2\theta_0 \lambda = p_\alpha = \chi^2_{2n,1-\alpha}$, por lo tanto $\lambda = \frac{\chi^2_{2n,1-\alpha}}{2\theta_0}$.

2. Suponga que $X_1, ..., X_n$ es una m.a. de una población cuya densidad es una $Bernoulli(\theta)$ i.e $p(x|\theta) = \theta^x(1-\theta)^{1-x}$.
   (a) Verifique que el EMV para θ es $\hat{\theta} = \sum_{i=1}^{n} \frac{X_i}{n}$
   (b) Demuestre que el EMV $\hat{\theta}$ del inciso anterior es un estimador eficiente de θ.[Hint: Use la Cota Inferior de Crao Ramer$= \frac{[g'(\theta)]^2}{I_{\tilde{X}}}$ donde $I_{\tilde{X}}$ es la información de Fisher]
   (c) Con base en las propiedades de los EMV's para muestras grandes, encuentre un IC$(1-\alpha)x100\,\%$ para la varianza de la distribución Bernoulli.

Solución (problema 2 examen tipo A, problema 3 examen tipo B) Recordar Bernoulli (θ)

$p(x|\theta) = \theta^x (1-\theta)^{1-x}$

$E(X) = \theta$

$V(X) = \theta(1-\theta)$

a) $L(\underline{X}, \theta) = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i} \Rightarrow$

$l(\theta) = ln(L(\underline{X}, \theta)) = \sum_{i=1}^{n} x_i (\ln(\theta)) - (n - \sum_{i=1}^{n} x_i)\ln(1-\theta)$ y

$$\frac{dl(\theta)}{d\theta} = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{(n - \sum_{i=1}^{n} x_i)}{1-\theta} \Rightarrow \frac{dl(\theta)}{d\theta} = 0 \Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} \text{ y } \frac{d^2 l(\theta)}{d\theta^2} < 0$$

b) $ln(P(X=x)) = xln(\theta) + (1-x)ln(1-\theta) \Rightarrow \frac{dlnP(X=x)}{d\theta} = \frac{x}{\theta} + \frac{1-x}{1-\theta}$ → Tenemos que $g(\theta) = \theta$

$$I_X = E\left[\frac{dlnP(X=x)}{d\theta}\right]^2 = E\left[\frac{x}{\theta} + \frac{1-x}{1-\theta}\right]^2 = \frac{\cancel{E[x-\theta]^2}^{\,V(X)}}{\theta^2(1-\theta)^2} = \frac{1}{\theta(1-\theta)} = I_{X_i}$$

$\Rightarrow [g'(\theta)]^2 = 1$

$\Rightarrow I_{\tilde{X}}(\theta) = \frac{n}{\theta(1-\theta)}$

$E\left[\left(\frac{d}{\delta\theta} \ln(g(t;\theta))\right)^2\right]$

$$C.I.C.R = \frac{1}{\frac{n}{\theta(1-\theta)}} = \frac{\theta(1-\theta)}{n} = Var(\hat{\theta}), por\ lo\ tanto\ \hat{\theta}\ es\ eficiente$$

c) El estimador de MV de $E(X) = \theta$ es $\hat{\theta} = \sum_{i=1}^{n} X_i/n$ y el de $VAR(X) = \theta(1-\theta)$ es

$$\hat{Var}(X) = \hat{g}(\theta) = g(\hat{\theta}) = \hat{\theta} - \hat{\theta}^2.$$

Entonces, se sabe que, aproximadamente

$$Z = \frac{\hat{\theta}(1-\hat{\theta}) - Var(X)}{\sqrt{\hat{\theta}(1-\hat{\theta})(1-2\hat{\theta})^2/n}} \approx N(0,1)$$

así que

$$P\left(-z_{\alpha/2} < \frac{\hat{\theta}(1-\hat{\theta}) - Var(X)}{\sqrt{\hat{\theta}(1-\hat{\theta})(1-2\hat{\theta})^2/n}} < z_{\alpha/2}\right) = 1-\alpha$$

de donde

$$P\left(\hat{\theta}(1-\hat{\theta}) - z_{\alpha/2}\sqrt{\hat{\theta}(1-\hat{\theta})(1-2\hat{\theta})^2/n} < Var(X) < \hat{\theta}(1-\hat{\theta}) + z_{\alpha/2}\sqrt{\hat{\theta}(1-\hat{\theta})(1-2\hat{\theta})^2/n}\right) = 1-\alpha$$

y el intervalo resultante es

$$\hat{\theta}(1-\hat{\theta}) \pm z_{\alpha/2}\sqrt{\hat{\theta}(1-\hat{\theta})(1-2\hat{\theta})^2/n}.$$

3. Suponga que $X_1, ..., X_n$ es una m.a. de una $Exp(\theta)$.

(a) Encuentre un IC de 90 % para θ, cuando $n = 9$

(b) Sea $\hat{\theta}$ el EMV de θ. Encuentre una función $g()$ tal que $\sqrt{n}[g(\hat{\theta}) - g(\theta)]$ siga una distribución $N(0, 1)$. Usando este hecho, encuentre un IC al 90 % para θ, cuando $n \to \infty$

Remember if $X \sim Gamma(\alpha, \beta)$
$kX \sim Gamma(\alpha, \frac{\beta}{k})$

(a) $\underline{n=9}$ We know $\sum X_i$ is a suff stat for θ

$\Rightarrow \sum X_i \sim Gamma(n, \theta) \Rightarrow 2\theta \sum X_i \sim Gamma(n, \theta/2\theta) \sim Gamma(n, \frac{1}{2}) \sim \chi^2_{2n}$

Now
$.90 = 1-\alpha \Rightarrow \alpha = \underline{.10}$

$\Rightarrow \mathbb{P}\{ \chi^2_{(2n, \alpha/2)} \leq 2\theta \sum X_i \leq \chi^2_{(2n, 1-\alpha/2)} \} = 1-\alpha$

$\Rightarrow \mathbb{P}\{ \frac{\chi^2_{(2n, \alpha/2)}}{2\sum X_i} \leq \theta \leq \frac{\chi^2_{(2n, 1-\alpha/2)}}{2\sum X_i} \} = \mathbb{P}\{ \frac{\chi^2_{(2n, \alpha/2)}}{2n\bar{X}} \leq \theta \leq \frac{\chi^2_{(2n, 1-\alpha/2)}}{2n\bar{X}} \}$

Given then $\bar{X} = \frac{1}{n}\sum X_i \Rightarrow n\bar{X} = \sum X_i$

$= 1-\alpha$

$\therefore$ my 90% CI for θ is $\left[\frac{\chi^2_{(2n, \alpha/2)}}{2n\bar{X}}, \frac{\chi^2_{(2n, 1-\alpha/2)}}{2n\bar{X}} \right]$

$n=9, \alpha = .1$

$= \left[\frac{\chi^2_{(18, .05)}}{18\bar{X}}, \frac{\chi^2_{(18, .95)}}{18\bar{X}} \right]$

(b) Let's find $\hat{\theta}$

$L(\theta) = \prod_1^n \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum_1^n x_i}$

$\ell(\theta) = n\ln(\theta) - \theta \sum_1^n x_i \Rightarrow \ell'(\theta) = \frac{n}{\theta} - \sum x_i = 0$ $(*)$

$\Leftrightarrow \hat{\theta}_{ML} = \frac{n}{\sum_1^n x_i} = \frac{1}{\bar{X}}$. Now, we know that $\underbrace{\frac{(g(\hat{\theta}) - g(\theta))}{1/\sqrt{n}} \sim N(0,1)}_{(*)}$

Since $\hat{\theta}_{ML} = 1/\bar{X} \Rightarrow g(\hat{\theta}) = \frac{\theta}{1/\bar{X}} = \bar{X}\theta$

$X_i \overset{iid}{\sim} Exp(\theta)$

$\Rightarrow E[g(\hat{\theta})] = E[\theta\bar{X}] = \theta E[\bar{X}] = \theta \frac{1}{\theta} = 1$

$V[g(\hat{\theta})] = V(\theta\bar{X}) = \theta^2 V(\bar{X}) = \theta^2 \frac{1}{\theta^2 n} = \frac{1}{n}$

$\therefore g(s) = \frac{\theta}{s}$ $\underline{\underline{\text{works}}}$

Now, $\frac{g(\hat{\theta}) - g(\theta)}{\frac{1}{\sqrt{n}}} = \frac{\bar{X}\theta - 1}{\frac{1}{\sqrt{n}}} \sim N(0,1)$

$\mathbb{P}\{ -z_{\alpha/2} \leq \sqrt{n}(\bar{X}\theta - 1) \leq z_{\alpha/2} \} = \mathbb{P}\{ -\frac{z_{\alpha/2}}{\sqrt{n}} \leq \bar{X}\theta - 1 \leq \frac{z_{\alpha/2}}{\sqrt{n}} \}$

$= \mathbb{P}\{ \frac{1 - z_{\alpha/2}/\sqrt{n}}{\bar{X}} \leq \theta \leq \frac{1 + z_{\alpha/2}/\sqrt{n}}{\bar{X}} \} = 1-\alpha$

$\Rightarrow \left[\frac{1 - z_{.05}/3}{\bar{X}}, \frac{1 + z_{.05}/3}{\bar{X}} \right]$ is a 90% CI for θ $\because n=9$.

4. Sean $X_1, ..., X_n$ es una m.a. de una $N(\mu, 1)$ y queremos probar $H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$. Si $\alpha = P(errortipoI)$ y la región de rechazo está dada por

$$RR = \left\{ x | \bar{x} < \mu_0 + \frac{\Phi^{-1}(\alpha)}{\sqrt{n}} \right\}$$

(a) Encuentre la función potencia.

(b) Demuestre que la función potencia es decreciente en μ

(a) $\pi_{\psi^*}(X) = E[\psi^*(X)] = P\{ \bar{X} \leq \mu_0 + \frac{\Phi^{-1}(\alpha)}{\sqrt{n}} \}$

Normalize Z_α

$$\Rightarrow P\left\{ \frac{\bar{X} - \mu}{\frac{1}{\sqrt{n}}} \leq \left[\mu_0 + \frac{\Phi^{-1}(\alpha)}{\sqrt{n}} - \mu \right] \sqrt{n} \right\}$$

$$= P\{ Z_\alpha \leq \mu_0\sqrt{n} + \Phi^{-1}(\alpha) - \mu\sqrt{n}] = \Phi(\mu_0\sqrt{n} + \underbrace{\Phi^{-1}(\alpha)}_{z} - \mu\sqrt{n})$$

$$= \Phi(\underbrace{\mu_0\sqrt{n} + z - \mu\sqrt{n}}_{g(\mu)})$$

(b) Since $\Phi(g(\mu))$ is an increasing function in $g(\mu)$

$$\Rightarrow \frac{\partial g(\mu)}{\partial \mu} = -\sqrt{n} < 0 \quad \because \quad n \geq 0$$

$\therefore$ From composition of functions $\Phi(g(\mu))$ is decreasing in μ.

5. Sean $X_1, ..., X_{25}$ es una m.a. de una $U(1,\theta)$ con $\theta > 1$. Encuentre la prueba uniformemente más poderosa para $\alpha = 0{,}05$ para probar $H_0 : 1 < \theta \leq 3$ vs $H_1 : \theta > 3$

Take $\begin{cases} H_0' : \theta_0 = 3 \\ H_1' : \theta_1 > 3 \end{cases}$ Let's use N-P where $f_{X_i}(x) = \frac{1}{\theta - 1}$

$F_{X_i}(x) = \frac{x_i - 1}{\theta - 1}$

$$L(\theta|x) = \prod_{i=1}^{n} \frac{1}{\theta-1} = \left(\frac{1}{\theta-1}\right)^n$$

$$\Rightarrow \frac{L(\theta_1)}{L(\theta_0)} = \left[\frac{\frac{1}{\theta_1-1}}{\frac{1}{\theta_0-1}}\right]^n \frac{\mathbb{1}_{(1,\theta_1)}(X_{(n)})}{\mathbb{1}_{(1,\theta_0)}(X_{(n)})} \Rightarrow \begin{cases} 1 \ ; \ X_{(n)} \leq \theta_0 < \theta_1 \\ \infty \ ; \ \theta_0 < X_{(n)} < \theta_1 \\ \nexists \ ; \ \theta_0 < \theta_1 < X_{(n)} \end{cases}$$

$X_{(n)}$ increasing

$$= \left(\frac{\theta_0 - 1}{\theta_1 - 1}\right)^n \frac{\mathbb{1}_{(1,\theta_1)} X_{(n)}}{\mathbb{1}_{(1,\theta_0)} X_{(n)}}$$

As $X_{(n)}$ increases the ratio is an increasing funtion in $T(\underset{\sim}{X}) = X_{(n)}$

$$\therefore \varphi(X) = \begin{cases} 1 & \text{if } X_{(n)} \geq k \\ 0 & \text{if } X_{(n)} < k \end{cases}$$

Remember

$F_{X_{(n)}} = [F(x)]^n$

$f_{X_{(n)}}(x) = n f(x) [F(x)]^{n-1}$

Now

$$\alpha = \underset{H_0}{\mathbb{P}}\{X_{(n)} \geq k\} = 1 - \mathbb{P}\{X_{(n)} \leq k\} = 1 - F_{X_{(n)}}(k)$$

$U(a-b)$

$\Rightarrow F(X) \sim U(0,1)$

$$= 1 - \left[\frac{k-1}{\theta_0 - 1}\right]^n = .05 \Rightarrow .95 = \frac{(k-1)^n}{(\theta_0-1)^n}$$

$$\Rightarrow k^* = 2^{-1/25}\left(\theta_0 + \sqrt[25]{20} - 1\right)$$

$$\therefore \varphi^*(X) = \begin{cases} 1 \ , & X_{(n)} \geq k^* \\ 0 \ , & X_{(n)} < k^* \end{cases}$$ is the UMPT for H_0' v. H_1'

Finally, we know $X_{(n)} \sim Beta(n,1)$ is an increasing function in θ

$$\Rightarrow \sup_{\theta \in \Theta}\{E[\varphi^*(x)]\} = \sup_{1<\theta\leq 3}\{\underbrace{E[\varphi^*(x)]}_{\pi_{\varphi^*}(\theta)}\} = E_{\underset{\to \theta_0}{3}}[\varphi^*(x)]$$

$$\Rightarrow E_\theta[\varphi^*(x)] \leq \overbrace{E_{\underset{\theta_0}{3}}[\varphi^*(x)]}^{\alpha} \therefore \forall \ 1 < \theta \leq 3 \ \ E_\theta[\varphi^*(x)] \leq \alpha$$

$\therefore \varphi^*(X)$ is the UMPT for H_0 v H_1

6. Sea $X_1, ..., X_n$ es una m.a. con $f(x;\theta)\frac{1}{2\theta}e^{|x|/\theta}$ donde $x \in \mathbb{R}$; $\theta > 0$. Encuentre un IC $(1-\alpha)x100\,\%$ para θ.

Since $f(x;\theta) \in$ Exp Fam $\Rightarrow |X|$ is a suff stat

$$Y = \frac{|X|}{\theta} \Rightarrow F_Y(y) = \mathbb{P}\left\{\frac{|X|}{\theta} \leq y\right\} = \mathbb{P}\{|X| \leq \theta y\} = \mathbb{P}\{-\theta y \leq X \leq \theta y\}$$

$$= F_X(\theta y) - F_X(-\theta y) \Rightarrow f_Y(y) = \theta f_X(\theta y) + \theta f_X(-\theta y)$$

$$= \theta\left[\frac{1}{2\theta} e^{-|\theta y|/\theta} + \frac{1}{2\theta} e^{-|-\theta y|/\theta}\right]$$

$$= \frac{1}{2}\left[e^{-\cancel{\theta} y/\cancel{\theta}} + e^{-\cancel{\theta} y/\cancel{\theta}}\right] = e^{-y} \Rightarrow Y \sim Exp(\theta = 1) \in \text{Exp Fam}$$

$\Rightarrow$ By Factorization Theorem $\sum_1^n Y_i$ is a suff stat $\sim Gamma(n, 1)$

$$\Rightarrow Q(y;\theta) = 2\sum_1^n Y_i \sim \chi^2_{2n} = \frac{2}{\theta}\sum_1^n |X_i|$$

$$\Rightarrow \mathbb{P}\left\{\chi^2_{(2n,1-\alpha/2)} \leq \frac{2\sum_1^n |X_i|}{\theta} \leq \chi^2_{(2n,\alpha/2)}\right\} = \mathbb{P}\left\{\frac{\chi^2_{(2n,1-\alpha/2)}}{2\sum_1^n |X_i|} \leq \frac{1}{\theta} \leq \frac{\chi^2_{(2n,\alpha/2)}}{2\sum_1^n |X_i|}\right\}$$

$$= \mathbb{P}\left\{\frac{2\sum_1^n |X_i|}{\chi^2_{(2n,\alpha/2)}} \leq \theta \leq \frac{2\sum_i^n |X_i|}{\chi^2_{(2n,1-\alpha/2)}}\right\} = 1-\alpha$$

$\therefore \left[\frac{2\sum_1^n |X_i|}{\chi^2_{(2n,\alpha/2)}}, \frac{2\sum_1^n |X_i|}{\chi^2_{(2n,1-\alpha/2)}}\right]$ is my $100(1-\alpha)\%$ CI for θ.