

# Inferencia estadística

## Tarea 3

J. Antonio García, jose.ramirez@cimat.mx

23 de noviembre de 2017

1. Sean  $X_1, \dots, X_n \sim \Gamma(\alpha, \beta)$ . Encuentra el estimador de  $\alpha$  y  $\beta$  mediante el método de momentos.

Igualamos el primer momento de la distribución con el primer momento muestral:

$$\begin{aligned}\mu'_1 &= m'_1 \\ \Leftrightarrow \\ \alpha\beta &= \sum_{i=0}^n \frac{X_i}{n} = \bar{X}\end{aligned}\tag{1}$$

Por otro lado, al hacer lo mismo para el segundo momento tenemos: Tenemos que  $\alpha = \frac{\bar{X}}{\beta}$  y sustituyendo este valor en dos obtenemos:

$$\begin{aligned}\mu'_2 &= m'_2 \\ \Leftrightarrow \\ \alpha\beta^2 &= \sum_{i=0}^n \frac{(X_i - \bar{X})^2}{n} = S_n^2\end{aligned}\tag{2}$$

De (1) tenemos que  $\alpha = \frac{\bar{X}}{\beta}$  y sustituyendo este valor en (2) obtenemos:

$$\begin{aligned}\alpha\beta^2 &= \frac{\bar{X}}{n}\beta^2 = \bar{X}\beta = S_n^2 \\ \Rightarrow \beta &= \frac{S_n^2}{\bar{X}} = \frac{\sum_{i=0}^n (x_i - \bar{X})}{\sum_{i=0}^n x_i} \\ \Rightarrow \alpha &= \frac{\bar{X}}{\frac{S_n^2}{\bar{X}}} = \frac{(\sum_{i=0}^n x_i)^2}{n \sum (x_i - \bar{X})^2}\end{aligned}$$

2. Comparación de dos tratamientos. A  $n_1$  personas se les da el tratamiento 1 y a  $n_2$  personas se les da el tratamiento 2. Sea  $X_1$  el número de personas en tratamiento 1 que responden de manera favorable al tratamiento, y  $X_2$  el número correspondiente de personas en tratamiento 2 que responden favorablemente. Asume que  $X_1 \sim \text{Binomial}(n_1, p_1)$  y  $X_2 \sim \text{Binomial}(n_2, p_2)$ .  
Sea  $\psi = p_1 - p_2$ .

a) Encuentra el MLE  $\hat{\psi}$  para  $\psi$ .

Como vimos rápidamente en clase, y como es probado en el libro de Casella G. y Berger R. L. (Statistical Inference) en la pág. 294 (en la sección de máxima verosimilitud), los estimadores de máxima verosimilitud son equivariantes independientemente de la dimensión del espacio de parámetros.

Entonces si consideramos a  $\psi$  como una función de los parámetros  $p_1$  y  $p_2$  tenemos:

$$\begin{aligned}\psi &= \psi(p_1, p_2) = p_1 - p_2 \\ \Rightarrow \hat{\psi} &= \hat{p}_1 - \hat{p}_2 = \frac{n_2 X_1 - n_1 X_2}{n_1 n_2}\end{aligned}$$

b) Encuentra la matriz de información de Fisher  $I(p_1, p_2)$ .

Como los parámetros  $p_1$  y  $p_2$  son independientes su función de distribución conjunta es el producto de sus distribuciones por separado consideremos la función de likelihood con una muestra de tamaño uno.

$$\begin{aligned} L(p_1, p_2) &= L(x; p_1, p_2) = L(x; p_1)L(x; p_2) \\ \Rightarrow \ln L(p_1, p_2) &= \ln L(x; p_1) + \ln L(x; p_2) \end{aligned} \quad (3)$$

Si derivamos lo anterior con respecto a  $p_i$  para obtener los scores  $s(X; p_i)$ :

$$\begin{aligned} s(X; p_i) &= \frac{\partial \ln L(p_1, p_2)}{\partial p_i} = \frac{\partial \ln \binom{n_i}{x} p_i^x (1-p_i)^{n_i-x}}{\partial p_i} \\ &= \frac{\partial}{\partial p_i} ((\ln(n_i!) - \ln((n_i-x)!) - \ln(x!)) + x \ln p_i + (n_i-x) \ln(1-p_i)) \\ &= \frac{\partial}{\partial p_i} (C_i + x \ln p_i + (n_i-x) \ln(1-p_i)) \\ &= \frac{x}{p_i} - \frac{n_i-x}{1-p_i} = \frac{x - n_i p_i}{p_i(1-p_i)} \end{aligned} \quad (4)$$

Por un lado sabemos que la matriz de información de Fisher la construiremos como:

$$I_{(p_1, p_2)} = - \begin{pmatrix} \mathbf{E} \left( \frac{\partial}{\partial p_1} s(X; p_1) \right) & \mathbf{E} \left( \frac{\partial}{\partial p_1} s(X; p_2) \right) \\ \mathbf{E} \left( \frac{\partial}{\partial p_2} s(X; p_1) \right) & \mathbf{E} \left( \frac{\partial}{\partial p_2} s(X; p_2) \right) \end{pmatrix}$$

Si notamos que  $\frac{\partial}{\partial p_j} s(X; p_i) = 0$  si  $i \neq j$  entonces el calculo de  $I(p_1, p_2)$  se reduce a

$$I_{(p_1, p_2)} = - \begin{pmatrix} \mathbf{E} \left( \frac{\partial}{\partial p_1} s(X; p_1) \right) & 0 \\ 0 & \mathbf{E} \left( \frac{\partial}{\partial p_2} s(X; p_2) \right) \end{pmatrix}$$

Finalmente solo resta calcular las matrices de Información de Fisher considerando los parámetros por separado es decir  $\mathbf{E} \left( \frac{\partial^2}{\partial p_j^2} \ln L(X; p_j) \right) = \mathbf{E} \left( \frac{\partial}{\partial p_j} s(X; p_j) \right)$ . Por un resultado visto en clase, en el libro, y como ejercicio en esta tarea sabemos que  $\mathbf{E} \left( \frac{\partial^2}{\partial p_j^2} \ln L(X; p_j) \right) = -\mathbf{E} (s(X; p_j)^2)$ , y lo calculamos :

$$\begin{aligned} \mathbf{E} (s(X; p_i)^2) &= \mathbf{E} \left( \left( \frac{x - n_i p_i}{p_i(1-p_i)} \right)^2 \right) \\ &= \frac{1}{p_i^2(1-p_i)^2} \mathbf{E} (x^2 - 2x n_i p_i + n_i^2 p_i^2) \\ &= \frac{1}{p_i^2(1-p_i)^2} (\mathbf{E}(x^2) - 2\mathbf{E}(x) n_i p_i + n_i^2 p_i^2) \\ &= \frac{1}{p_i^2(1-p_i)^2} (\mathbf{E}(x^2) - 2n_i^2 p_i^2 + n_i^2 p_i^2) \\ &= \frac{1}{p_i^2(1-p_i)^2} (\mathbf{E}(x^2) - n_i^2 p_i^2) \\ &= \frac{1}{p_i^2(1-p_i)^2} (n_i p_i - n_i p_i^2 - n_i^2 p_i^2 + n_i^2 p_i^2) \\ &= \frac{n_i p_i (1-p_i)}{p_i^2(1-p_i)^2} = \frac{n_i}{p_i(1-p_i)} \end{aligned}$$

Donde la penultima igualdad se debe a que  $V(x) = \mathbf{E}(x^2) - \mathbf{E}(x)^2 \rightarrow \mathbf{E}(x^2) = V(X) + \mathbf{E}(x)^2 = n_i p_i (1-p_i) = n_i^2 p_i^2 = n_i(p_i - p_i^2) + n_i^2 p_i^2 = n_i p_i - n_i p_i^2 + n_i^2 p_i^2$

Por lo que la matriz queda:

$$I_{(p_1, p_2)} = - \begin{pmatrix} \mathbf{E} \left( \frac{\partial}{\partial p_1} s(X; p_1) \right) & 0 \\ 0 & \mathbf{E} \left( \frac{\partial}{\partial p_2} s(X; p_2) \right) \end{pmatrix} = \begin{pmatrix} \frac{n_1}{p_1(1-p_1)} & 0 \\ 0 & \frac{n_2}{p_2(1-p_2)} \end{pmatrix}$$

c) Usa el método Delta para encontrar el error estandar asintotico de  $\hat{\psi}$ .

Por el método Delta (de multiparámetros) sabemos que:

$$\hat{se}(\hat{\psi}) = \sqrt{(\hat{\nabla} \psi)^t \hat{J}_n(\hat{\nabla} \psi)}$$

Por un lado  $(\hat{\nabla}\psi)^t = (1, -1)^t$  y por otro:

$$\hat{J}_n = \hat{I}_{(p_1, p_2)}^{-1} = \begin{pmatrix} \frac{n_1}{\hat{p}_1(1-\hat{p}_1)} & 0 \\ 0 & \frac{n_2}{\hat{p}_2(1-\hat{p}_2)} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} & 0 \\ 0 & \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \end{pmatrix}$$

Por lo que podemos escribir:

$$(\hat{\nabla}\psi)^t \hat{J}_n (\hat{\nabla}\psi) = (1, -1)^t \begin{pmatrix} \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} & 0 \\ 0 & \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (1, -1)^t \begin{pmatrix} \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} \\ -\frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \end{pmatrix} = \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}$$

Con lo que obtenemos  $\hat{se}(\hat{\psi}) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

- d) Supon que  $n_1 = n_2 = 200$ ,  $X_1 = 160$  y  $X_2 = 148$ . Encuentra  $\hat{\psi}$ . Encuentra también un intervalo de confianza del 90 % aproximado para  $\psi$  usando el método Delta.

Primero sabemos que  $\hat{\psi} = \hat{p}_1 - \hat{p}_2 = \frac{160}{200} - \frac{148}{200} = \frac{4}{5} - \frac{37}{50} = 0.06$

Además por el método Delta sabemos que  $\frac{\hat{\psi} - \psi}{\hat{se}(\hat{\psi})} \sim N(0, 1)$  por lo que podemos construir un intervalo de confianza como sigue:

$$(\hat{\psi} - Z_{.05}\hat{se}(\hat{\psi}), \hat{\psi} + Z_{.05}\hat{se}(\hat{\psi}))$$

Donde, como con la notación que utilizamos,  $z_{.05}$  es el quinto percentil de una normal estándar y sustituyendo obtenemos el intervalo:

$$(0.06 - (-.6449)(0.04198), 0.06 - (1.6449)(0.04198)) \\ \Rightarrow (-0.009044678, 1.1290447)$$

Como parte de los ejercicios de clase presento los siguientes

3. De las notas ‘clase15.pdf’ en ejercicio de la lamina 46:

La proporción real de familias en cierta ciudad que viven en casa propia es 0.7. Si se escogen al azar 84 familias de esa ciudad y se les pregunta si viven o no en casa propia, ¿Con qué probabilidad podemos asegurar que el valor que se obtendra de la proporción muestral caera entre 0.64 y 0.76 ?

En clase se obtuvo una estimación aproximando la distribución binomial con una Normal estándar, en vista de que el número de muestra es grande, de :

$P(-1.2 < Z < 1.2) = 2 * P(0 < Z < 1 : 2) = 2(0.3849) = 0.7698$  Si utilizamos la corrección por continuidad, como dice el ejercicio, obtenemos:

$$\begin{aligned} P(0.64 \leq \hat{\theta} \leq 0.76) &\approx P(0.64 - \frac{1}{2n} < \hat{\theta} < 0.76 + \frac{1}{2n}) \\ &= P(0.64 - \frac{1}{168} < \hat{\theta} < 0.76 + \frac{1}{168}) \\ &= P(0.6340476 < \hat{\theta} < 0.7659524) \\ &= P(\frac{0.6340476 - .7}{\sqrt{\frac{0.7*0.3}{84}}} < z < P(\frac{0.7659524 - .7}{\sqrt{\frac{0.7*0.3}{84}}}) \\ &= P(-1.319048 < z < 1.319048) \\ &= P(z < 1.319048) - P(z < -1.319048) = 0.8128469 \end{aligned}$$

Qué es mayor a lo estimado sin considerar la corrección por continuidad.

4. De las notas ‘clase15.pdf’ en ejercicio de la lamina 72:

Calcular el sesgo de  $S_n^2$ .

Sabemos que  $s_n^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$ , y en el ejercicio anterior a este en clase mostramos que  $\mathbf{E}(s_n^2) = \frac{n-1}{n}\sigma^2$ , por lo que lo aplicamos a la definición de sesgo para obtener:

$$\begin{aligned} \text{sesgo}(s_n^2) &= \mathbf{E}(s_n^2) - \sigma^2 \\ &= \frac{n-1}{n}\sigma^2 - \sigma^2 = \\ &= \sigma^2 \left( \frac{n-1}{n} - 1 \right) \\ &= \sigma^2 \left( \frac{n-1}{n} - \frac{n}{n} \right) \\ &= \sigma^2 \left( \frac{-1}{n} \right) \end{aligned}$$

5. De las notas ‘clase15.pdf’ en ejercicio de la lamina 78:

Se definió  $Y_{(n)}$  como el n-ésimo el estadístico de orden  $n$  de una m.a.  $X_1, \dots, X_n$  con  $X_1 \sim \text{Unif}(0, \theta)$ , es decir su máximo. El ejercicio consiste en encontrar  $f_{Y_{(n)}}(y)$  y  $V(Y_{(n)})$ .

En este curso mostramos, numerosas veces, que la densidad del máximo de una muestra (aterrizado a nuestro ejemplo) esta dado por:

$$\begin{aligned} f_{Y_{(n)}}(y) &= [F_x(x)]^{n-1} f_x(x) = n \left( \frac{x}{\theta} \right)^{n-1} \left( \frac{1}{\theta} \right) \\ &= \frac{nx^{n-1}}{\theta^n} \end{aligned}$$

Por otro lado  $\mathbf{E}(y) = \int_0^\theta y \frac{n}{\theta^n} y^{n-1} dy = \frac{n}{\theta^n} \int_0^\theta y^n = \frac{n}{\theta^n} \frac{\theta^{n+1}}{n+1} = \theta \frac{n}{n+1}$

$$\begin{aligned} V(Y_{(n)}(y)) &= \mathbf{E}(y^2) - \mathbf{E}(y)^2 \\ &= \int_0^\theta y^2 \frac{ny^{n-1}}{\theta^n} dy - \theta^2 \left( \frac{n}{n+1} \right)^2 \\ &= \frac{n}{\theta^n} \int_0^\theta y^{n+1} dy - \theta^2 \left( \frac{n}{n+1} \right)^2 \\ &= \frac{n}{\theta^n} \frac{\theta^{n+2}}{n+2} - \theta^2 \left( \frac{n}{n+1} \right)^2 \\ &= \theta^2 \left( \frac{n}{n+2} - \frac{n^2}{(n+1)^2} \right) \\ &= n\theta^2 \left( \frac{(n+1)^2 - n(n+2)}{(n+2)(n+1)^2} \right) \\ &= n\theta^2 \left( \frac{1}{(n+2)(n+1)^2} \right) \end{aligned}$$

6. De las notas ‘clase15.pdf’ en ejercicio de la lamina 101, este ejercicio se enumera varias veces en las laminas y aunque también lo mostramos en clase me sente a hacerlo de nuevo:

Demostrar

$$n\mathbf{E} \left[ \left( \frac{\partial}{\partial \theta} \ln f(x) \right)^2 \right] = -n\mathbf{E} \left[ \frac{\partial^2}{\partial \theta^2} \ln f(x) \right]$$

Aquí estan las cuentas:

$$\begin{aligned}
\mathbf{E} \left( \frac{\partial^2}{\partial \theta^2} \ln f(x) \right) &= \mathbf{E} \left( \frac{\partial}{\partial \theta} \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right) \right) \\
&= \mathbf{E} \left( \frac{f(x) \frac{\partial^2 f(x)}{\partial \theta^2} - \left( \frac{\partial f(x)}{\partial \theta} \right)^2}{f(x)^2} \right) \\
&= \mathbf{E} \left( \frac{f(x) \frac{\partial^2 f(x)}{\partial \theta^2} - \left( \frac{\partial f(x)}{\partial \theta} \right)^2}{f(x)^2} \right) \\
&= \mathbf{E} \left( \frac{\frac{\partial^2 f(x)}{\partial \theta^2}}{f(x)} - \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right) \\
&= \mathbf{E} \left( \frac{\frac{\partial^2 f(x)}{\partial \theta^2}}{f(x)} \right) - \mathbf{E} \left[ \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right] \\
&= \int \frac{\frac{\partial^2 f(x)}{\partial \theta^2}}{f(x)} f(x) dx - \mathbf{E} \left[ \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right] \\
&= \int \frac{\partial^2 f(x)}{\partial \theta^2} dx - \mathbf{E} \left[ \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right] \\
&= \frac{\partial^2}{\partial \theta^2} \int f(x) dx - \mathbf{E} \left[ \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right] \\
&= \frac{\partial^2}{\partial \theta^2} 1 - \mathbf{E} \left[ \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right] \\
&= 0 - \mathbf{E} \left[ \left( \frac{\frac{\partial f(x)}{\partial \theta}}{f(x)} \right)^2 \right]
\end{aligned}$$

Donde la cuarta desde atrás igualdad requiere que la función  $f(x)$  cumpla la condición de Leibniz, lo cual es cierto pues  $f(x)$  es una función de densidad.

7. De las notas ‘clase16.pdf’ en ejercicio de la lamina 67, Se pide encontrar los estimadores máximo verosímiles para una m.a. que proviene de una exponencial recorrida dada por:

$$f(x; \lambda, \theta) = \begin{cases} \lambda e^{\lambda(x-\theta)} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

Calculamos la verosimilitud de la muestra

$L(x_1, \dots, x_n; \lambda, \theta) = \prod_{i=1}^n \lambda e^{\lambda(x_i-\theta)} = \lambda^n e^{-\lambda(\sum_{i=1}^n x_i) + n\lambda\theta}$ , si cada una de las observaciones es mayor a  $\theta$  (lo cual a su vez es equivalente a que  $x_{(1)} \geq \theta$ ) y lo importante aquí es mencionar que  $L(x_1, \dots, x_n; \lambda, \theta)$  vale cero si  $x_{(1)} < \theta$ .

Luego, aplicamos la función logaritmo para obtener:

$$\ln L(x_1, \dots, x_n; \lambda, \theta) = n \ln(\lambda) - \lambda \left( \sum_{i=1}^n x_i \right) + n\lambda\theta$$

Y derivando lo anterior e igualando a cero obtenemos  $\lambda_{MV}$

$$\begin{aligned}
0 &= \frac{\partial}{\partial \lambda} \ln L(x_1, \dots, x_n; \lambda, \theta) = \frac{n}{\lambda} - \sum_{i=1}^n x_i + n\theta \\
&\Rightarrow \lambda_{MV} = \frac{n}{\sum_{i=1}^n x_i - n\theta}
\end{aligned}$$

Al derivar con respecto a  $\theta$  obtenmos

$$\frac{\partial}{\partial \theta} \ln L(x_1, \dots, x_n; \lambda, \theta) = n\lambda$$

Como En la expresión anterior no aparece  $\theta$  y como ésta derivada es positiva siempre pues  $\lambda$  lo es, prestamos atención a la condición de que  $x_{(1)} > \theta$  para notar que la verosimilitud se hace cero si  $x \leq \theta$  el estimador  $\theta_{MV} = \min\{x_1, x_2, \dots, x_n\}$ .

Por lo que

$$\theta_{MV} = 17.82$$

Y

$$\lambda_{MV} = \frac{10}{55 - 8 - 10 * 0.64} \approx 0.2015$$

8. De las notas 'clase17.pdf' en ejercicio de la lamina 27, Encuentra en el anterior ejemplo los valores  $z_{\alpha/2}$  utilizados.

Utilicé R para calcular  $z_{.10/2} = -1.644854$  y  $z_{0.5/2} = -1.959964$ .

El intervalo de confianza para un nivel de confianza de 100 % requiere que  $z_{\alpha/2} \rightarrow \infty$ , por lo que el intervalo sería toda la recta real y lo cual no es informativo, yo diría que es equivalente a no estimar el parámetro.

9. De las notas 'clase17.pdf' en el ejercicio de la lamina 43, el Intervalo de confianza que se calculo fue para la media del logaritmo de los datos observados por lo que por su relación de estos con la media y la varianza de la distribución log-normal podemos obtener otro intervalo:

$$e^{\hat{x} + \hat{\sigma}^2/2} \pm z_{\alpha/2} e^{2\hat{x} + \hat{\sigma}^2} (e^{\sigma^2} - 1) = (e^{0.6976 + 0.0552^2/2} \pm z_{10/2} e^{2*0.6976 + 0.0552^2} (e^{0.0552^2} - 1)) = 2.005867 \pm 0.01235353 = (1.993514, 2.018221)$$

10. De las notas 'intervalos\_de\_confianza.pdf' en el ejercicio de la hoja 19 se define el estadístico

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- a) Mostrar que  $\mathbf{E}(s_p^2) = \sigma^2$   
Lo calculamos

$$\begin{aligned} \mathbf{E}(s_p^2) &= \mathbf{E}\left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right) \\ &= \left(\frac{(n_1 - 1)\mathbf{E}(s_1^2) + (n_2 - 1)\mathbf{E}(s_2^2)}{n_1 + n_2 - 2}\right) \\ &= \left(\frac{(n_1 - 1)\sigma^2 + (n_2 - 1)\sigma^2}{n_1 + n_2 - 2}\right) \\ &= \left(\frac{(n_1 - 1 + n_2 - 1)\sigma^2}{n_1 + n_2 - 2}\right) \end{aligned}$$

Donde la antepenúltima igualdad se debe a que cada  $s_i$  es insesgado.

- b) Mostrar que  $\frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2} \sim \chi_{n_1 + n_2 - 2}^2$ .  
Al sustituir el valor de  $s_p^2$  obtenemos:

$$\begin{aligned} \frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2} &= \left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right) \\ &= \frac{(n_1 + n_2 - 2)}{\sigma^2(n_1 + n_2 - 2)} \left( (n_1 - 1) \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1 - 1} + (n_2 - 1) \frac{\sum_{i=1}^{n_2} (x_i - \bar{x})^2}{n_2 - 1} \right) \\ &= \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (x_i - \bar{x})^2}{\sigma^2} \end{aligned}$$

De la expresión anterior recordamos el supuesto de que  $X \sim Norm(0, 1)$  además cada uno de los dos términos anteriores es a su vez una suma de normales estándar al cuadrado por lo que cada una de ellas se distribuye  $\chi_{n_1-1}^2$  y  $\chi_{n_2-1}^2$  respectivamente, este resultado lo vimos en clase, además cada uno de los términos anteriores es independiente del otro por lo que tenemos los supuestos para usar el resultado de que la distribución  $\chi^2$  es cerrada bajo sumas de ellas. Por lo que tenemos que:

$$\frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2} = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (x_i - \bar{x})^2}{\sigma^2} \sim \chi_{n_1-1}^2 + \chi_{n_2-1}^2 = \chi_{n_1+n_2-2}^2$$

c) Mostrar que  $s_p^2$  es de mínima varianza.

Para probar ello primero calculamos la varianza del estadístico: Sabemos que

$$\begin{aligned} \frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2} &\sim \chi_{n_2-1}^2 = \chi_{n_1+n_2-2}^2 \\ \Rightarrow \mathbf{V}\left(\frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2}\right) &= 2(n_1 + n_2 - 2) \end{aligned} \quad (5)$$

Por otro lado:

$$\begin{aligned} \mathbf{V}\left(\frac{(n_1 + n_2 - 2)s_p^2}{\sigma^2}\right) &= \frac{(n_1 + n_2 - 2)^2 \mathbf{V}(s_p^2)}{\sigma^4} \\ \Rightarrow \mathbf{V}(s_p^2) &= \frac{2\sigma^4}{n_1 + n_2 - 2} \end{aligned} \quad (6)$$

A continuación usamos la desigualdad de Cramer-Rao para calcular la varianza mínima que debe tener un estimador para  $\sigma$  de una normal, la primer igualdad es por el ejercicio 7 de esta tarea.

$$\begin{aligned} \mathbf{E}\left[\left(\frac{\partial}{\partial \theta} \ln f(x)\right)^2\right] &= -\mathbf{E}\left[\frac{\partial^2}{\partial \theta^2} \ln f(x)\right] \\ &= -\mathbf{E}\left[\left(\frac{\partial}{\partial \theta} \frac{(x - \mu)^2}{2\sigma^4} - \frac{1}{2\sigma^2}\right)\right] \\ &= -\mathbf{E}\left[\left(\frac{-(x - \mu)^2}{2\sigma^6} + \frac{1}{2\sigma^4}\right)\right] \\ &= -\left[\frac{-2\mathbf{E}[(x - \mu)^2] + \sigma^2}{2\sigma^6}\right] \\ &= -\frac{-2\sigma^2 + \sigma^2}{2\sigma^6} \\ &= -\frac{-1}{2\sigma^4} = \\ &= \frac{1}{2\sigma^4} \end{aligned}$$

Por lo que el estadístico  $s_p^2$  satisface la condición de igualdad en la desigualdad de Cramer-R, la diferencia en las constantes se debe a los grados de libertad pues en una muestra de tamaño  $n_1 + n_2$  se pierden dos grados de libertad al estimar dos parámetros.