Martes 26 de octubre de 2021.

MÉTODO DE LOS MOMENTOS PARA ESTIMAR PARÁMETROS.

Este método da por hecho que los momentos muestrales son una buena aproximación de los momentos poblacionales, por lo tanto se igualan ambos momentos y de ahí se despeja el estimador del parámetro de interés.

$\mu_k = E(X^k)$ k-ésimo momento poblacional de la v.a X.

$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ k-ésimo momento muestral de la v.a X.

$\boxed{\mu_k = m_k}$ y de aquí se despeja el estimador.

Al estimador del parámetro θ obtenido por el método de los momentos se le pone una tilde de la forma $\tilde{\theta}$.

Ejemplo 9: Sea $X \sim N(x;\mu,\sigma^2)$ estima $\underline{\theta} = (\mu,\sigma^2)$ en base a una muestra aleatoria de tamaño n usando el método de los momentos.

Solución

Tenemos dos parámetros a estimar por lo tanto necesitamos los dos primeros momentos muestrales y los dos primeros momentos poblacionales.
Se igualan los primeros momentos; $\tilde{\mu} = \bar{x}$
Se igualan también los segundos momentos

$V(\bar{X}) = E(\bar{X}^2) - \mu^2$
$E(\bar{X}^2) = V(\bar{X}) + \mu^2$
$= \sigma^2 + \mu^2$

$$\mu_2 = E(X^2) = \sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = m_2 \quad \text{entonces } \boxed{\tilde{\mu} = \bar{x}}$$

$$\tilde{\sigma}^2 + \tilde{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\tilde{\sigma}^2 + (\bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

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

$$\boxed{\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

¿Qué pero le ponemos a este estimador?
R.- No es insesgado
(El estimador insesgado para σ^2 es $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Ejemplo 10: Determina los estimadores por momentos para los parámetros $\underbrace{\alpha \text{ y } \beta}_{\text{2 estimadores}}$ de la distribución gamma, en base a una m.a de tamaño n.

∴ Necesitamos 2 momentos poblacionales y 2 momentos muestrales

Solución

Igualando los dos primeros momentos poblacionales con los muestrales tenemos un sistema de dos ecuaciones con dos incógnitas a resolver.

$$\left.\begin{aligned} \mu_1 &= \frac{\tilde{\alpha}}{\tilde{\beta}} = \bar{X} = m_1 \\ \mu_2 &= \frac{\tilde{\alpha}}{\tilde{\beta}^2} + \frac{\tilde{\alpha}^2}{\tilde{\beta}^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = m_2 \end{aligned}\right\} \text{sistema de ecs. 2x2}$$

Despejando $\tilde{\alpha}$ de la primera ecuación y sustituyendo en la segunda se tiene:

si $\frac{\tilde{\alpha}}{\tilde{\beta}} = \bar{x}$ $\quad \therefore \left(\frac{\tilde{\alpha}}{\tilde{\beta}}\right)\frac{1}{\tilde{\beta}} + \left(\frac{\tilde{\alpha}}{\tilde{\beta}}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$$\tilde{\alpha} = \tilde{\beta}\bar{X}$$

$$\frac{\tilde{\beta}\bar{X}}{\tilde{\beta}^2} + \frac{\tilde{\beta}^2\bar{X}^2}{\tilde{\beta}^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\frac{\bar{X}}{\tilde{\beta}} + \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\frac{\bar{X}}{\tilde{\beta}} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - \bar{X}\right)^2$$

$\frac{\bar{x}}{\tilde{\beta}} + \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$\frac{\bar{x}}{\tilde{\beta}} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$

$\frac{1}{\tilde{\beta}} = \frac{\frac{1}{n}\left[\sum x_i^2 - n\bar{x}^2\right]}{\bar{x}}$

$\tilde{\beta} = \frac{\bar{x}}{\frac{1}{n}\left[\sum x_i^2 - n\bar{x}^2\right]} = \frac{n\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

Finalmente, los estimadores por el método de los momentos para los parámetros de la distribución Gamma son: $\tilde{\beta} = \frac{n\bar{X}}{\sum_{i=1}^{n}\left(X_i^2 - \bar{X}\right)^2}$ y $\tilde{\alpha} = \tilde{\beta}\bar{X} = \frac{n\bar{X}^2}{\sum_{i=1}^{n}\left(X_i^2 - \bar{X}\right)^2}$

Ejemplo 11: Sea X una v.a binomial con parámetros n y p desconocidos, encuentra los estimadores por el método de los momentos para tales parámetros en base a una muestra aleatoria de tamaño k.

Solución

$$\mu_1 = np = \bar{X} = m_1$$
$$\mu_2 = np(1-p) + n^2p^2 = \frac{1}{k}\sum_{i=1}^{k} X_i^2 = m_2$$

$$\bar{X}(1-\tilde{p}) + \bar{X}^2 = \frac{1}{k}\sum_{i=1}^{k} X_i^2$$

$$\bar{X} - \tilde{p}\bar{X} + \bar{X}^2 = \frac{1}{k}\sum_{i=1}^{k} X_i^2$$

$$\bar{X} - \tilde{p}\bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i^2 - \bar{X}^2 = \frac{1}{k}\sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2$$

$$\tilde{p}\bar{X} = \bar{X} - \frac{1}{k}\sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2$$

$$\boxed{\tilde{p} = 1 - \frac{1}{k\bar{X}}\sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2}$$

$$\boxed{\tilde{n} = \frac{\bar{X}}{\tilde{p}} = \frac{k\bar{X}^2}{k\bar{X} - \sum_{i=1}^{k}\left(X_i - \bar{X}\right)^2}}$$

Ejemplo 12: Sea una población con distribución Geométrica con parámetro p. Encuentra el estimador para el parámetro usando el método de los momentos y en base a una muestra aleatoria de tamaño n.

Solución
$\mu_1 = \frac{1}{\tilde{p}} = \bar{X} = m_1 \quad \therefore \quad \tilde{p} = \frac{1}{\bar{X}}$

MÉTODO DE MÁXIMA VEROSIMILITUD.

Definición 5: La función de verosimilitud de n variables aleatorias se define como su densidad conjunta $f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)$. En particular si $X_1,X_2,\ldots,X_n$ es una muestra aleatoria de la densidad $f(x;\theta)$ entonces la función de verosimilitud de la muestra está dada por

$$L(\theta) = f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta)$$
$$\boxed{L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)}$$

Definición 6: Se dice que $\hat{\theta}$ es un estimador máximo verosímil si maximiza la función de verosimilitud.

Recuerda que para maximizar se deriva parcialmente respecto al parámetro de interés, se iguala a cero y se despeja el valor que hace posible la ecuación.
Por facilidad de cálculo y dado que las funciones $L(\theta)$ y $\ln L(\theta)$ se maximizan en el mismo punto, se acostumbra encontrar al estimador máximo verosímil resolviendo la ecuación $\frac{\partial}{\partial\theta}\ln L(\theta) = 0$

Método:
1) Escribir la función de verosimilitud. $L(\theta)$
[Escribir los 2 primeros términos y el último]
2) Sacar el logaritmo natural. $\ln(L(\theta))$
3) Derivar parcialmente respecto al parámetro o parámetros de interés $\underbrace{\text{igualar a cero}}_{\text{poner su gorro "^" a los parámetros}}$ y despejar al estimador de interés o resolver el sistema para encontrar a los estimadores de interés.

Cuando se tiene un vector de parámetros a estimar por ejemplo $\underline{\theta} = (\theta_1,\theta_2,\ldots,\theta_k)$ la manera de proceder es resolver las k ecuaciones siguientes:

$$\frac{\partial}{\partial\theta_1}\ln L(\underline{\theta}) = \frac{\partial}{\partial\theta_1}\ln L(\theta_1,\theta_2,\ldots,\theta_k) = 0$$
$$\frac{\partial}{\partial\theta_2}\ln L(\underline{\theta}) = \frac{\partial}{\partial\theta_2}\ln L(\theta_1,\theta_2,\ldots,\theta_k) = 0$$
$$\vdots$$
$$\frac{\partial}{\partial\theta_k}\ln L(\underline{\theta}) = \frac{\partial}{\partial\theta_k}\ln L(\theta_1,\theta_2,\ldots,\theta_k) = 0$$

Ejemplo 13: Encuentra el estimador máximo verosímil de p en base a una m.a de tamaño n de una población Bernoulli. $\mu = p$
Solución

$$f_X(x;p) = p^x(1-p)^{1-x} \text{ si } x = 0,\ 1$$

i) La función de verosimilitud de una m.a de tamaño n

$$L(p) = \prod_{i=1}^{n} f(x_i,p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

$= p^{x_1}(1-p)^{1-x_1}\, p^{x_2}(1-p)^{1-x_2} \cdots p^{x_n}(1-p)^{1-x_n}$
$= p^{x_1+x_2+\cdots+x_n}(1-p)^{1-x_1+1-x_2+\cdots+1-x_n}$

ii) Sacar el logaritmo natural de la función de verosimilitud.

$$\ln L(p) = \left(\sum_{i=1}^{n} x_i\right)\ln(p) + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-p)$$

iii) Sacar la derivada parcial respecto a p.

$\frac{\partial}{\partial p}\ln L(p) = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p}$ ***iv)*** Igualar a cero y resolver la ecuación $\frac{\partial}{\partial p}\ln L(p) = 0$. En el momento en que se iguala a cero se le pone un "gorrito" al parámetro para indicar que se trata de su estimador máximo verosímil.

$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} - \frac{n - \sum_{i=1}^{n} x_i}{1-\hat{p}} = 0$$
$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} = \frac{n - \sum_{i=1}^{n} x_i}{1-\hat{p}}$$
$$\frac{1-\hat{p}}{\hat{p}} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$
$$\frac{1}{\hat{p}} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1$$
$$\frac{1}{\hat{p}} = \frac{n}{\sum_{i=1}^{n} x_i} \therefore \hat{p} = \bar{x}$$

Ejemplo 14: Sea $X_1,X_2,\ldots,X_n$ una m.a de una densidad Normal con media μ y varianza σ^2 encuentra el estimador máximo verosímil de $\underline{\theta} = (\mu,\sigma^2)$.

Ejemplo 14: Sea $X_1,X_2,\ldots,X_n$ una m.a de una densidad Normal con media μ y varianza σ^2 encuentra el estimador máximo verosímil de $\underline{\theta} = (\mu,\sigma^2)$.

Solución

$$f_X(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \text{ si } x > 0$$

i) La función de verosimilitud de una m.a de tamaño n es:

$$L(\mu,\sigma^2) = \prod_{i=1}^{n} f(x_i,\mu,\sigma^2) = \prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$= \left(\sigma\sqrt{2\pi}\right)^{-n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

ii) Sacar el logaritmo natural de la función de verosimilitud.

$$\ln L(\mu,\sigma^2) = -n\ln\sigma - n\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

iii) Sacar las derivadas parciales respecto a μ y respecto a σ^2.

$$\frac{\partial}{\partial\mu}\ln L(\mu,\sigma^2) = \frac{2}{2\sigma^2}\sum_{i=1}^{n}2(x_i-\mu) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)$$
$$\frac{\partial}{\partial\sigma^2}\ln L(\mu,\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

iv) Igualar a cero y resolver el sistema para encontrar los estimadores máximo verosímiles de los parámetros de interés.

$$0 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} x_i - n\hat{\mu}\right] \quad \therefore \quad \hat{\mu} = \bar{x}$$

$$0 = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$
$$n = \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad \therefore \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Observa que el estimador máximo verosímil para la media poblacional de la distribución Normal es la media muestral que además es un estimador insesgado, óptimo y consistente. Y que el estimador máximo verosímil para la varianza poblacional de una distribución Normal es un estimador sesgado pero consistente.

Definición 7: Supón que $X_1,X_2,\ldots,X_n$ es una muestra aleatoria de tamaño n en orden de aparición, si ordenamos esta muestra de menor a mayor tenemos lo que se conoce como estadísticas de orden y se denota $X_{(1)},X_{(2)},\ldots,X_{(n)}$ tales que $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$.

Ejemplo 15: Se toma una muestra aleatoria de 5 estudiantes y se registra su calificación en el segundo parcial de probabilidad y estadística, siendo las calificaciones en orden de aparición $x_1 = 10$, $x_2 = 5.3$, $x_3 = 8$, $x_4 = 6.7$, $x_5 = 1$, las estadísticas de orden correspondientes son $x_{(1)} = 1$, $x_{(2)} = 5.3$, $x_{(3)} = 6.7$, $x_{(4)} = 8$, $x_{(5)} = 10$.

Ejemplo 16: Se toma una muestra aleatoria de tamaño n de una población distribuida uniformemente en el intervalo $(0,a)$ ¿Cuál es el estimador máximo verosímil para a?

Solución

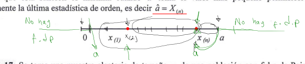
i) La función de verosimilitud está dada por

$$L(a) = \prod_{i=1}^{n} f(x_i,a) = \prod_{i=1}^{n}\frac{1}{a} = \frac{1}{a^n}$$

Queremos maximizar esta función.
Será máxima cuando a^n sea mínima o equivalentemente cuando a sea mínima $(n>0)$
Aprovechando la muestra

ii) Encontrar el valor de a que maximiza la función de verosimilitud.

Observa que esta función es máxima cuando a toma el valor más pequeño permisible que es precisamente la última estadística de orden, es decir $\hat{a} = X_{(n)}$



Ejemplo 17: Se toma una muestra aleatoria de tamaño n de una población con f.d.p de Poisson con media λ.
a) Encuentra el estimador máximo verosímil para λ.
b) Encuentra el valor esperado y la varianza del estimador.
c) ¿Es un estimador consistente?

$X \sim P(x,\lambda)$
$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

Solución
i) La función de verosimilitud está dada por

$$L(\lambda) = \prod_{i=1}^{n}\frac{\lambda^{x_i}e^{-\lambda}}{x_i!} = e^{-\lambda n}\frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

$$\ln L(\lambda) = -\lambda n + \sum_{i=1}^{n} x_i\ln(\lambda) - \ln\prod_{i=1}^{n} x_i!$$

$$\frac{\partial}{\partial\lambda}\ln L(\lambda) = -n + \sum_{i=1}^{n}\frac{x_i}{\lambda}$$

i) $L(\lambda) = \prod_{i=1}^{n}\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$
$= e^{-\lambda}\frac{\lambda^{x_1}}{x_1!}e^{-\lambda}\frac{\lambda^{x_2}}{x_2!}\cdots e^{-\lambda}\frac{\lambda^{x_n}}{x_n!}$
$= \left(e^{-\lambda n}\right)\left(\lambda^{\sum_{i=1}^{n} x_i}\right)\left(\prod_{i=1}^{n}\frac{1}{x_i!}\right)$

2) $\ln(L(\lambda)) = -\lambda n + \left(\sum_{i=1}^{n} x_i\right)\ln(\lambda) + \ln\left(\prod\frac{1}{x_i!}\right)$
$\frac{\partial}{\partial\lambda}\ln(L(\lambda)) = -n + \sum_{i=1}^{n} x_i\left(\frac{1}{\lambda}\right) = 0$

ii) Igualando a cero y resolviendo se encuentra el estimador máximo verosímil.

$$n = \sum_{i=1}^{n}\frac{x_i}{\lambda} \quad \therefore \quad \hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$\frac{1}{\hat{\lambda}}\sum_{i=1}^{n} x_i = n \rightarrow \hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X} = \hat{\lambda}$

b) Valor esperado y varianza del estimador.

$$E(\bar{X}) = \lambda \text{ y } V(\bar{X}) = \frac{\lambda}{n}$$

c) $\lim_{n\to\infty} V(\bar{X}) = \lim_{n\to\infty}\frac{\lambda}{n} = 0$ entonces sí es consistente.

$E(\bar{X}) = \lambda$ y $V(\bar{X}) = \frac{\lambda}{n}$
¿Es consistente? $\lim_{n\to\infty} V(\hat{\lambda}) = \lim_{n\to\infty}\frac{\lambda}{n} = 0 \rightarrow$ sí es consistente.

Ejemplo 18: Si $X_1,X_2,\ldots,X_n$ es una muestra aleatoria de una población con f.d.p $f_X(x;\theta) = \frac{r}{\theta}x^{r-1}e^{-x^r/\theta}$ con $\theta > 0$ y $x > 0$ con r constante y positiva. Encuentra el estimador máximo verosímil para θ.

$f_X(x) = \frac{r}{\theta}x^{r-1}\exp\left(-\frac{x^r}{\theta}\right)$

Valor: +1 sobre el 3er parcial.
Revisar 1) Castillo
2) Trinidad
3) Mendoza

Solución

i) La función de verosimilitud está dada por

$$L(\theta) = \prod_{i=1}^{n}\frac{r}{\theta}x_i^{r-1}e^{-x_i^r/\theta} = \frac{r^n}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i^r}\prod_{i=1}^{n} x_i^{r-1}$$

ii) Sacar el logaritmo natural de la función de verosimilitud.

$$\ln L(\theta) = \ln\left(\frac{r^n}{\theta^n}\right) - \frac{1}{\theta}\sum_{i=1}^{n} x_i^r + \sum_{i=1}^{n}\ln(x_i^{r-1}) \quad \ln\left(\prod_{i=1}^{n} x_i^{r-1}\right)$$
$$= n\ln(r) - n\ln(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i^r + (r-1)\sum_{i=1}^{n}\ln(x_i) + \ln\left(\prod_{i=1}^{n} x_i^{r-1}\right)$$
$$\frac{\partial}{\partial\theta}\ln L(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i^r$$

iii) Igualando a cero y resolviendo se tiene que

$\frac{n}{\hat{\theta}} = \frac{1}{\hat{\theta}^2}\sum_{i=1}^{n} x_i^r$ despejando al estimador del parámetro se tiene que $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i^r$ r-ésimo momento muestral

$\frac{1}{n}\sum_{i=1}^{n} X_i$ media muestral

$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2$ varianza muestral

$X_{(n)} - X_{(1)}$ rango

$E(X^k) = \begin{cases}\sum_x x^k P(X = x) \\ \int_{-\infty}^{\infty} x^k f_X(x)\,dx\end{cases}$

$m_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$

Ejemplo 19: Supón que $X_1,X_2,\ldots,X_m$ es una m.a de la producción por acre de la variedad A de trigo, la cual tiene una distribución Normal con media μ_1 y varianza σ^2 y que $Y_1,Y_2,\ldots,Y_n$ es una m.a de la producción por acre de la variedad B de trigo, la cual se distribuye como una Normal con media μ_2 y varianza σ^2. Si $X \perp Y$ encuentra el estimador máximo verosímil para la varianza común σ^2 si se desconocen las medias poblacionales.
Solución

$$X \sim N(x;\mu_1,\sigma^2) \text{ y } Y \sim N(y;\mu_2,\sigma^2)$$

Sabemos que el estimador máximo verosímil para la media poblacional está dado por la media muestral es decir $\hat{\mu}_1 = \bar{X}$ y $\hat{\mu}_2 = \bar{Y}$

La función de verosimilitud para la muestra conjunta está dada por

$$L(\mu_1,\mu_2,\sigma^2) = f(X_1,X_2,\ldots,X_m,Y_1,Y_2,\ldots,Y_n)$$
$$= \prod_{i=1}^{m} f(x_i;\mu_1,\sigma^2)\prod_{i=1}^{n} f(y_i;\mu_2,\sigma^2)$$
$$= \prod_{i=1}^{m}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2}\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(y_i-\mu_2)^2}$$
$$= \prod_{i=1}^{m}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2}\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(y_i-\mu_2)^2}$$
$$= (2\pi)^{-\frac{m+n}{2}}(\sigma^2)^{-\frac{m+n}{2}}\exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(X_i-\mu_1)^2 + \sum_{i=1}^{n}(Y_i-\mu_2)^2\right]\right)$$

Sacar el logaritmo natural a la función de máximo verosímilitud

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

Derivando parcialmente respecto a la varianza

$$\frac{\partial}{\partial\sigma^2}\ln L(\mu_1,\mu_2,\sigma^2) = -\frac{m+n}{2\sigma^2} + -\frac{1}{2\sigma^4}\left[\sum_{i=1}^{m}(X_i-\mu_1)^2 + \sum_{i=1}^{n}(Y_i-\mu_2)^2\right]$$

Igualando a cero y despejando se tiene que el estimador máximo verosímil para la varianza común es:

$$\hat{\sigma}^2 = \frac{1}{m+n}\left[\sum_{i=1}^{m}(X_i-\bar{X})^2 + \sum_{i=1}^{n}(Y_i-\bar{Y})^2\right]$$