tarea teorica, Juan Pablo Cabeza

P1: Consideremos un vector bidimensional con sus dos comp. ortogonales, independientes y con distribución normal.

Sea $U=(U_1,\dots,U_n)$ con $U_n \sim N(0,\sigma^2)$

$W=(W_1,\dots,W_n)$ con $W_n \sim N(0,\sigma^2)$

en base a estas MAS, se obtiene $X=(X_1,\dots,X_n)$ dado por

$$X_i = \sqrt{U_i^2+W_i^2}$$

obs: $U_i \sim N(0,\sigma^2)$
$\Rightarrow U_i^2 \sim \chi_i^2(1)$
de ahí su distribución

a) Función densidad de X_i, $i=1,\dots,n$.

Notemos que, si $X_i = \sqrt{U_i^2+W_i^2}$, entonces,

$$F_{X_i}(z) = \mathbb{P}(X_i \leq z) = \mathbb{P}\left(\sqrt{U_i^2+W_i^2} \leq z\right) = \mathbb{P}\left(U_i^2+W_i^2 \leq z^2\right)$$

$$= \iint_{x^2+y^2 \leq z^2} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-(x^2+y^2)}{2\sigma^2}\right) dx\,dy$$

tomando $x=r\cos\theta$, $y=r\sin\theta$, $r>0$, $\theta\in[0,2\pi)$, se sigue que del teorema de cambio de variable,

$$F_{X_i}(z) = \int_0^{2\pi}\int_0^{z} \frac{1}{2\pi\sigma^2} \cdot \exp\left(\frac{-r^2}{2\sigma^2}\right) r\,dr\,d\theta$$

$$= \int_0^{z} \frac{1}{\sigma^2} \exp\left(\frac{-r^2}{2\sigma^2}\right) r\,dr = \int_0^{z} \frac{1}{\sigma^2}\left(-\frac{\sigma^2}{r}\frac{d}{dr}\exp\left(\frac{-r^2}{2\sigma^2}\right)\right) r\,dr$$

$$= -\left(\exp\left(\frac{-z^2}{2\sigma^2}\right) - 1\right) = 1-\exp\left(\frac{-z^2}{2\sigma^2}\right)$$

Escaneado con CamScanner

Por lo tanto, $F_{X_i}(z^2) = 1 - e^{-\frac{z^2}{2\sigma^2}}$, luego,

$$\Rightarrow F_X(z) = 1 - \exp\left(-\frac{z}{2\sigma^2}\right)$$

$$\Rightarrow f(z) = \frac{d}{dz}\left(1 - e^{-\frac{z}{2\sigma^2}}\right) = -e^{\frac{-z}{2\sigma^2}}\left(-\frac{1}{2\sigma^2}\right)$$

es decir,

$$f(z) = \begin{cases} \frac{1}{2\sigma^2} e^{\frac{-z}{2\sigma^2}} & , 0 < z \\ 0 & , eoc \end{cases}$$

P1.

b) Calculemos el estimador máxima verosimilitud ($\hat{\sigma}^2_{EMV}$) de σ^2.

Sea $\mathcal{L}(z_1, \ldots, z_n | \sigma^2) = \prod_{i=1}^{n} f(z_i | \sigma^2)$

$$= \prod_{i=1}^{n}\left(\frac{1}{2\sigma^2}\exp\left(-\frac{z_i}{2\sigma^2}\right)\right) = \frac{1}{(2\sigma^2)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} z_i\right)$$

tomando log,

$$\ln(\mathcal{L}(z_1, \ldots, z_n | \sigma^2)) = -n\ln(2\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} z_i$$

dado que buscamos $\hat{\sigma}^2_{EMV}$, debemos maximizar la función,

$$\frac{\partial}{\partial\sigma^2}\ln(\mathcal{L}(z_1, \ldots, z_n | \sigma^2)) = 0$$

$$\Rightarrow -n\frac{1}{\cancel{2}\sigma^2}\cdot \cancel{2} - \frac{1}{2}\left(\frac{-1}{(\sigma^2)^2}\cdot\sum_{i=1}^{n} z_i\right) = 0$$

$$\Rightarrow \frac{n}{\sigma^2} = \frac{1}{2}\sum_{i=1}^{n} z_i \cdot \frac{1}{(\sigma^2)^2} \Rightarrow n = \frac{1}{2\sigma^2}\sum_{i=1}^{n} z_i \Rightarrow \frac{2n}{\sum_{i=1}^{n} z_i} = \frac{1}{\sigma^2}$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n} z_i}{2n} \Rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n} z_i\right)}$$

Escaneado con CamScanner

P1.c) Veamos que $\hat{\sigma}^2_{EMV}$ es insesgado para σ^2 y calculemos la varianza de $\hat{\sigma}^2_{EMV}$.

Notemos que, usando la función gamma, (solo para ahorrar cálculos, también se puede integrar por partes)

$$\mathbb{E}(z) = \int_0^\infty \frac{z}{2\sigma^2} \exp\left(-\frac{z}{2\sigma^2}\right) dz$$

$$= 2\sigma^2 \int_0^\infty \left(\frac{z}{2\sigma^2}\right)^{2-1} \exp\left(\frac{-z}{2\sigma^2}\right) \frac{dz}{2\sigma^2} = 2\sigma^2 \Gamma(2) = 2\sigma^2$$

Luego, $\mathbb{E}(\hat{\sigma}^2) = \mathbb{E}\left(\frac{1}{2n}(z_1 + z_2 + \cdots + z_n)\right)$ de la linealidad de $\mathbb{E}$

$$= \frac{1}{2n}\left(\mathbb{E}(z_1) + \cdots + \mathbb{E}(z_n)\right)$$

$$= \frac{1}{2n}\left(2\sigma^2 + \cdots + 2\sigma^2\right) = \frac{2n\sigma^2}{2n} = \sigma^2$$

Por lo tanto $\hat{\sigma}^2$ es insesgado.

Por otra parte,

$$\mathbb{E}(z^2) = \int_0^\infty \frac{z^2}{2\sigma^2} \exp\left(\frac{-z}{2\sigma^2}\right) dz = (2\sigma^2)^2 \int_0^\infty \left(\frac{z}{2\sigma^2}\right)^{3-1} \exp\left(\frac{-z}{2\sigma^2}\right) \frac{dz}{2\sigma^2}$$

$$= 4(\sigma^2)^2 \Gamma(3) = 4\sigma^4 \, 2! = 8\sigma^4$$

Por lo tanto,

$$V(z) = \mathbb{E}(z^2) - (\mathbb{E}(z))^2 = 8\sigma^4 - (2\sigma^2)^2 = 8\sigma^4 - 4\sigma^4 = 4\sigma^4$$

Luego, $V(\hat{\sigma}^2) = V\left(\frac{z_1 + \cdots + z_n}{2n}\right) = \frac{1}{4n^2}\left(V(z_1) + \cdots + V(z_n)\right)$

$$= \frac{1}{4n^2}\left(4\sigma^4 + \cdots + 4\sigma^4\right) = \frac{4n\sigma^4}{4n^2} = \frac{\sigma^4}{n}$$

Escaneado con CamScanner

P1

(Bonus) Notemos que,

$$\frac{\partial^2}{\partial^2(\sigma^2)} \ln\left(L(z_1, \dots, z_n \mid \sigma^2)\right) = -n\left(\frac{-1}{\sigma^2}\right) + \frac{1}{2}\left(\frac{-2}{(\sigma^2)^3}\right)\sum_{i=1}^{n} z_i$$

entonces, $E_\theta\left[\frac{\partial^2(\ln L(z_1, \dots, z_n \mid \sigma^2))}{\partial^2(\sigma^2)}\right] = E\left[\frac{n}{(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum_{i=1}^{n} z_i\right]$

$$= \frac{n}{(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\left(E(z_1) + \dots + E(z_n)\right) = \frac{n}{(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\left(2\sigma^2 + \dots + 2\sigma^2\right)$$

$$= \frac{n}{(\sigma^2)^2} - \frac{2\sigma^2 n}{(\sigma^2)^3} = \frac{n}{(\sigma^2)^2} - \frac{2n}{(\sigma^2)^2} = -\frac{n}{(\sigma^2)^2}$$

Luego $I(\sigma^2) = -E\left(\frac{\partial^2 \log L(\sim)}{\partial^2 \sigma^2}\right) = \frac{n}{(\sigma^2)^2}$ //

Ahora, calculemos $I(\sigma^2)^{-1}$. Note que $y = \frac{n}{(\sigma^2)^2} \Rightarrow (\sigma^2)^2 = \frac{n}{y}$ / $\sqrt{\ }$

$\Rightarrow \sigma^2 = \sqrt{\frac{n}{y}}$

Por lo tanto $I(\sigma^2)^{-1} = \sqrt{\frac{n}{\sigma^2}}$.

Se sigue que, $\sqrt{n}\left(\sigma^{2(n)}_{EMV} - \sigma^2\right) \sim N\left(0, \sqrt{\frac{n}{\sigma^2}}\right)$

Finalmente, $\left(\sigma^{2(n)}_{EMV} - \sigma^2\right) \sim N\left(0, \frac{1}{\sigma}\right)$ //

Escaneado con CamScanner

P2. b) Sea $y = a_0 + a_1 x_1 + a_2 x_2 + \varepsilon$, con $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$ y $D = \{x_1^{(i)}, x_2^{(i)}, y^{(i)}\}_{i=1}^n$ el conj. de datos de entrenamiento.

Notemos que la densidad conjunta de $\varepsilon_i \sim N(0, \sigma^2)$ es

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right) = \frac{1}{\sqrt{(2\pi\sigma^2)^n}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n \varepsilon_i^2\right)$$

Determinemos el estimador máxima verosimilitud, es decir,

$$L(a_i, (y, x_i)) = \frac{1}{\sqrt[n]{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n \left(y - (a_0 + a_1 x_1 + a_2 x_2)\right)^2\right)$$

$$\Rightarrow \ln L(a_i, (y, x_i)) = -\frac{n}{2}\left(\ln(2\pi) + \ln \sigma^2\right) - \frac{1}{2\sigma_M^2}\sum_{i=1}^n \left(y - (a_0 + a_1 x_1 + a_2 x_2)\right)^2$$

Por tanto, maximizar la función de verosimilitud es equivalente a minimizar,

$$\sum_{i=1}^n \left(y - (a_0 + a_1 x_1 + a_2 x_2)\right)^2$$

es decir, encontrar el estimador de máxima verosimilitud de la log, es equivalente a minimizar el error cuadrático, P2.d) Considere el funcional

$$J_{reg}(y,\theta) = \frac{1}{2}\| y - \bar{X}\theta\|_2^2 + \frac{1}{2}\rho\|\theta\|_2^2$$

Recordemos que si $U \subset \mathbb{R}^n$ es un subconj. abto y $f: U \to \mathbb{R}$ es diferenciable entonces ∇f es una función de U a $\mathbb{R}^n$ tal que,

$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - \nabla f(x)\cdot h|}{\|h\|} = 0$$

Notemos que,

$$J_{reg}(y,\theta) = \frac{1}{2}(y - \bar{X}\theta)^T(y - \bar{X}\theta) + \frac{1}{2}\rho\theta^T\theta$$

$$= \frac{1}{2}\left(y^Ty - y^T\bar{X}\theta - \theta^T\bar{X}^Ty + \theta^T\bar{X}^T\bar{X}\theta + \rho\theta^T\theta\right)$$

De sigue que,

$$J_{reg}(y,\theta+h) = \frac{1}{2}\Big(y^Ty - y^T\bar{X}(\theta+h) - (\theta+h)^T\bar{X}^Ty + (\theta+h)^T\bar{X}^T\bar{X}(\theta+h) + \rho(\theta+h)^T(\theta+h)\Big)$$

$$= J_{reg}(y,\theta) + \frac{1}{2}\Big(-y^T\bar{X}h - h^T\bar{X}^Ty + \theta^T\bar{X}^T\bar{X}h + h^T\bar{X}^T\bar{X}h + h^T\bar{X}^T\bar{X}\theta + \rho(\theta^Th + h^T\theta + h^Th)\Big)$$

$$\Rightarrow J_{reg}(y,\theta+h) - J_{reg}(y,\theta) = \frac{1}{2}\Big(-y^T\bar{X} - (\bar{X}^Ty)^T + \theta^T\bar{X}^T\bar{X} + h^T\bar{X}^T\bar{X} + (\bar{X}^T\bar{X}\theta)^T + \rho\theta^T + \rho\theta^T + \rho h^T\Big)\cdot h$$

De donde concluimos que el gradiente es,

$$\nabla_\theta J_{reg}(y,\theta) = \theta^T\bar{X}^T\bar{X} - y^T\bar{X} + \rho\theta^T$$

$$= (\theta^T\bar{X}^T - y^T)\bar{X} + \rho\theta^T$$

Escaneado con CamScanner

P2.d) Sea $y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \varepsilon$, $\varepsilon \sim N(0, \sigma^2)$

Determinemos $\mathbb{P}(y \mid x_1, x_2)$

Dado que $\varepsilon_i \sim N(0, \sigma^2)$, entonces la densidad es

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)$$

lo cual implica que,

$$\mathbb{P}(y \mid x_1, x_2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - (\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2))^2}{2\sigma^2}\right)$$

Escaneado con CamScanner

P3) La estimación de cada científico es $N(\mu, \sigma_i), i \in (1, \dots 7)$ entonces

$$f(A, \dots, G) = \frac{1}{\sqrt[7]{2\pi}} \frac{1}{\sigma_A \cdots \sigma_G} \exp\left(-\frac{1}{2\sigma_A^2}(-27{,}000-\mu)^2\right)$$

$$\vdots$$

$$\exp\left(-\frac{1}{2\sigma_G^2}(-10{,}056-\mu)^2\right)$$

$$\frac{\partial \ln}{\partial \mu}(P(A, \dots, G; \mu, \sigma_i)) = \frac{1}{\sigma^2}\left(24{,}24300 - 7\mu\right)$$

luego sumando, obtenemos

$$\mu \simeq 3{,}4632$$

Escaneado con CamScanner