Universidad de la República | **Estadística Multivariada Computacional**
Facultad de Ingeniería - IMERL | **Primer Semestre 2021**

Práctico 1

1. Recordamos que si $\mathbf{x}, \mathbf{y}$ son vectores aleatorios en $\mathbb{R}^p$

   - $\mathbb{E}(\mathbf{x}) = \mu_x$
   - $Var(\mathbf{x}) = \Sigma_{x,x} = \mathbb{E}\big((\mathbf{x} - \mu_x)(\mathbf{x} - \mu_x)'\big)$
   - $Cov(\mathbf{x}, \mathbf{y}) = \Sigma_{x,y} = \mathbb{E}\big((\mathbf{x} - \mu_x)(\mathbf{y} - \mu_y)'\big)$ siendo $\mathbf{x} \sim (\mu_x, \Sigma_{xx})$ e $\mathbf{y} \sim (\mu_y, \Sigma_{yy})$

   Sean $\mathbf{x}$ e $\mathbf{y}$ vectores aleatorios, A y B matrices y c un vector fijo (no aleatorio) real. Pruebe que:

   a) $\mathbb{E}(A\mathbf{x}) = A\mathbb{E}(\mathbf{x})$

   b) $Cov(A\mathbf{x}, B\mathbf{y}) = ACov(\mathbf{x}, \mathbf{y})B'$

   c) $Var(A\mathbf{x}) = AVar(\mathbf{x})A'$

   d) $Cov(\mathbf{x}, \mathbf{y}) = \mathbb{E}(\mathbf{x}\mathbf{y}') - \mathbb{E}(\mathbf{x})\mathbb{E}(\mathbf{y})'$

   e) $Var(\mathbf{x} - c) = Var(\mathbf{x})$

   f) Si $\mathbf{x} \sim (\mu, \Sigma)$ entonces $\mathbb{E}(\mathbf{x}'A\mathbf{x}) = tr(A\Sigma) + \mu' A\mu$.

2. Sea $\mathbf{x_1}, \ldots, \mathbf{x_n}$ es una muestra aleatoria simple proveniente de una variable aleatoria $\mathbf{x} \sim (\mu, \Sigma)$, y $\overline{\mathbf{x}}_\mathbf{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x_i}$ el promedio muestral. Pruebe que:

   a) $\mathbb{E}(\overline{\mathbf{x}}_\mathbf{n}) = \mu$ y $Var(\overline{\mathbf{x}}_n) = \frac{1}{n}\Sigma$

   b) si $S = \dfrac{1}{n}\sum_{i=1}^{n}(\mathbf{x_i} - \overline{\mathbf{x}}_n)(\mathbf{x_i} - \overline{\mathbf{x}}_n)'$ (matriz de varianzas y covarianzas muestrales), entonces $\mathbb{E}(S) = \dfrac{n-1}{n}\Sigma$. De un estimador insesgado de Σ a partir de S.

3. En este ejercicio, se sugiere usar R para hacer las cuentas. Se dispone de 3 indicadores económicos x_1, x_2, x_3 que se miden en 4 paises, con los resultados siguientes:

x_1	x_2	x_3
2	3	-1
1	5	-2
2	2	1
2	3	1

   a) Calcule el vector de medias, la matriz de varianzas y covarianzas, la varianza generalizada, la matriz de correlación y los valores y vectores propios de dicha matriz.

   b) Se construyen los nuevos indicadores

$$y_1 = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3 \qquad y_2 = x_1 - 0{,}5x_2 - 0{,}5x_3$$

   Calcule el vector de medias para $\mathbf{y} = (y_1, y_2)'$, su matriz de varianzas y covarianzas y su matriz de correlación

4. Demuestre que si $\mathbf{Y} = \mathbf{X}\mathbf{A}$ donde $\mathbf{Y} \in \mathcal{M}_{n\times m}$ y $\mathbf{X} \in \mathcal{M}_{n\times p}$, las matrices de covarianzas de $\mathbf{Y}$ y de $\mathbf{X}$ están relacionadas por $\mathbf{S}_y = \mathbf{A}'\mathbf{S}_x\mathbf{A}$.

5. Demuestre que para un conjunto de datos $\dfrac{1}{np}\sum\limits_{i=1}^{n}(\mathbf{x_i} - \overline{\mathbf{x}}_\mathbf{n})'S^{-1}(\mathbf{x_i} - \overline{\mathbf{x}}_\mathbf{n}) = 1$.

6. Pruebe que si $\mathbf{y} \sim Mult\big(n, \mathbf{p}\big)$ con $\mathbf{p} = (p_1, \ldots, p_G)'$ entonces

$$\mathbb{E}(\mathbf{y}) = n\mathbf{p} \quad \text{y} \quad Var(\mathbf{y}) = n(diag(\mathbf{p}) - \mathbf{p}\mathbf{p}')$$

7. Se considera la función de densidad dada por

$$f(x, y) = \begin{cases} Kx & \text{si } 0 < x < 1,\ 0 < y < 1 - x \\ 0 & \text{en otro caso} \end{cases}$$

Halle K, las funciones de densidad marginales, las densidades condicionadas, el vector de medias y la matriz de varianzas-covarianzas y el coeficiente de correlación de la variables aleatorias.

8. Se extraen sucesivamente, sin reemplazamiento, dos bolas de una urna que contiene tres blancas y dos negras. En relación a este experimento, se definen las siguientes variables:

$$X = \begin{cases} 0 & \text{si la primera bola es blanca} \\ 1 & \text{si la primera bola es negra} \end{cases} \qquad Y = \begin{cases} 0 & \text{si la segunda bola es blanca} \\ 1 & \text{si la segunda bola es negra} \end{cases}$$

   a) Calcular $\mathbb{E}(X|Y)$

   b) Descomponer la varianza de la variable X teniendo en cuenta el condicionamiento a Y.

9. Sea (X, Y) un vector aleatorio con funcón de densidad $f(x, y) = 2$ si $0 < x < y < 1$

   a) Calcular $\mathbb{E}(X^3|Y)$

   b) Descomponer la varianza de la variable Y teniendo en cuenta el condicionamiento a X.