ESTADÍSTICA MULTIVARIANTE

UGR, GRADO EN MATEMÁTICAS

Curso Académico 2023-2024

José Miguel Angulo Ibáñez (*jmangulo@ugr.es*)

Departamento de Estadística e Investigación Operativa
Universidad de Granada

- **TEMA 1. Distribución Normal Multivariante**

  Aspectos generales sobre vectores aleatorios

Esperanza y covarianza ('operadores')

- Sea $\mathbf{X} = (X_1, \ldots, X_p)'$ un vector aleatorio
- Se define el *vector de medias* de $\mathbf{X}$ como

$$\boldsymbol{\mu}_{\mathbf{X}} := E[\mathbf{X}] := \begin{pmatrix} E[X_1] \\ \vdots \\ E[X_p] \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$$

(siempre que existan las esperanzas unidimensionales)

- PROPIEDAD DE LINEALIDAD: Sea $\mathbf{Y} = (Y_1, \ldots, Y_q)'$ definido como

$$\mathbf{Y} = B\mathbf{X} + \mathbf{b}, \quad \text{con} \quad \begin{array}{l} B \text{ matriz } q \times p \text{ (cte.)} \\ \mathbf{b} \text{ vector } q \times 1 \text{ (cte.)} \end{array}$$

Entonces,

$$\boldsymbol{\mu}_{\mathbf{Y}} = BE[\mathbf{X}] + \mathbf{b} = B\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}$$ (☐ Probar)

(El resultado se extiene convenientemente al caso de transformaciones lineales de matrices aleatorias) (☐ Probar)

Esperanza y covarianza ('operadores') (cont.)

- Se define la *matriz de covarianzas* de $\mathbf{X}$ como

$$\Sigma_{\mathbf{X}} = \mathrm{Cov}(\mathbf{X}) := E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'] = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix},$$

  con

$$\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] \qquad (= \sigma_{ji})$$

  (siempre que existan las esperanzas unidimensionales).

  En particular,

$$\sigma_{ii} = E[(X_i - \mu_i)^2] = \mathrm{Var}(X_i) \qquad (= \sigma_i^2, \text{ notación})$$

- PROPIEDADES:

  (1) $\Sigma_{\mathbf{X}}$ es simétrica

  (2) Los elementos de la diagonal de $\Sigma_{\mathbf{X}}$ son no negativos

  (3) La clase de matrices de covarianzas (dim. $p \times p$) coincide con la clase de matrices simétricas definidas no negativas (dim. $p \times p$) ($\square$)

En relación con la propiedad **(3)**, para cualquier matriz de covarianzas, Σ, se distinguen los casos siguientes:

A. $\boxed{\Sigma \text{ definida positiva}}$

(notación: $\Sigma > 0$)

En este caso, Σ es no singular, con $|\Sigma| > 0$, y $\exists \Sigma^{-1}$

'Normalización' y *distancia de Mahalanobis*

Dado un vector aleatorio p-dimensional $\mathbf{X} \sim (\boldsymbol{\mu}, \Sigma)$ con $\Sigma > 0$, para cualquier elección de una matriz C de dimensión $p \times p$ tal que $\Sigma = CC'$ se obtiene una 'normalización' (en origen y escala multidimensionales) del vector mediante la transformación

$$\mathbf{Z} = C^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

En efecto, se tiene que $\mathbf{Z} \sim (\mathbf{0}_p, I_{p\times p})$ (□ Probar)

Esperanza y covarianza ('operadores') (cont.)

Se define la *distancia de Mahalanobis* de $\mathbf{X} \sim (\boldsymbol{\mu}, \Sigma)$ (con $\Sigma > 0$) con respecto a su vector de medias $\boldsymbol{\mu}$ como

$$\Delta(\mathbf{X}, \boldsymbol{\mu}) := \left\{ (\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \right\}^{\frac{1}{2}}$$

Interpretación y observaciones:

- Se comprueba fácilmente que $\Delta(\mathbf{X}, \boldsymbol{\mu}) = \|\mathbf{Z}\|_p$, para cualquier normalización $\mathbf{Z}$ de $\mathbf{X}$ ($\|\cdot\|_p$ denota la norma euclídea en el espacio $\mathbb{R}^p$) (□ Probar)
- $\Delta(\mathbf{X}, \boldsymbol{\mu})$ es una variable aleatoria, cumpliéndose que

  $$E[\Delta^2(\mathbf{X}, \boldsymbol{\mu})] = p$$ (□ Probar)

- La ecuación

  $$\Delta(\mathbf{x}, \boldsymbol{\mu}) = k,$$

  para $k \geq 0$ (cte.) y $\mathbf{x} \in \mathbb{R}^p$, define un hiperelipsoide en $\mathbb{R}^p$, de tal modo que los puntos transformados por normalización se corresponden con la esfera euclídea p-dimensional de radio k con centro en el origen $\mathbf{0}$ (□ Probar)

B. $\boxed{\Sigma \text{ semidefinida positiva}}$

(notación: $\Sigma \geq 0$ indicará, en general, 'definida no negativa')

En este caso, Σ es singular, es decir, $|\Sigma| = 0$ y $\nexists \Sigma^{-1}$

Por tanto, no se puede realizar una normalización en todo el espacio $\mathbb{R}^p$, ni definir la distancia de Mahalanobis de $\mathbf{X} \sim (\boldsymbol{\mu}, \Sigma)$ con respecto a su vector de medias $\boldsymbol{\mu}$ a nivel p-dimensional

Observaciones:

- Se tendrá que, siendo rango$(\Sigma) = r < p$, se puede escribir

$$\Sigma = CC', \quad \text{con } C \text{ matriz } p \times r \text{ de rango } r$$

- Como consecuencia: Con probabilidad 1, las componentes del vector aleatorio $\mathbf{X} = (X_1, \ldots, X_p)'$ cumplirán (al menos) una relación de dependencia lineal del tipo

$$\boldsymbol{\alpha}'\mathbf{X} = k, \quad \text{con } \boldsymbol{\alpha} \neq \mathbf{0}$$ (□ Probar)

(es decir, la variabilidad de $\mathbf{X}$ se sitúa ($P_{\mathbf{X}}$-c.s.) en un hiperplano afín en $\mathbb{R}^p$)

Esperanza y covarianza ('operadores') (cont.)

- TRANSFORMACIONES LINEALES: Sean

$$\begin{array}{rcll} \mathbf{X} & \sim & (\boldsymbol{\mu}_{\mathbf{X}}, \Sigma_{\mathbf{X}}) & \text{vec.a. } p\text{-dimensional} \\ \mathbf{Y} & = & B\mathbf{X} + \mathbf{b} & \text{vec.a. } q\text{-dimensional, con} \\ & & B & \text{matriz } q \times p \text{ (cte.)} \\ & & \mathbf{b} & \text{vector } q \times 1 \text{ (cte.)} \end{array}$$

  Entonces, se tiene que

$$\mathbf{Y} \sim (\boldsymbol{\mu}_{\mathbf{Y}}, \Sigma_{\mathbf{Y}}) = (B\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}, B\Sigma_{\mathbf{X}}B')$$

  (□ Probar)

- ALGUNAS MEDIDAS GLOBALES DE VARIACIÓN:
  - $|\Sigma| = \prod_{j=1}^{p} \lambda_j$ (determinante)
  - $\mathrm{tr}(\Sigma) = \sum_{j=1}^{p} \sigma_j{}^2 = \sum_{j=1}^{p} \lambda_j$ (traza)

    ($\lambda_1, \ldots, \lambda_p$, autovalores de Σ)

Función característica

- [Recordatorio] Dado un vector aleatorio $\mathbf{X} = (X_1, \ldots, X_p)'$, se define su *función característica* como

$$\begin{aligned}\phi_{\mathbf{X}}(\mathbf{t}) &= E\left[e^{i\mathbf{t}'\mathbf{X}}\right] \\ &= \int_{\Omega} e^{i\mathbf{t}'\mathbf{X}(\omega)} P(d\omega) = \int_{\mathbb{R}^p} e^{i\mathbf{t}'\mathbf{x}} P_{\mathbf{X}}(d\mathbf{x}),\end{aligned}$$

  $\forall \mathbf{t} = (t_1, \ldots, t_p)' \in \mathbb{R}^p$

- TEOREMA: Sea $\mathbf{X} = (X_1, \ldots, X_p)'$ un vector aleatorio. Se tiene que la distribución (multivariante) de $\mathbf{X}$ queda unívocamente determinada por el conjunto de todas las distribuciones (univariantes) de variables aleatorias de la forma

$$\boldsymbol{\alpha}'\mathbf{X}, \quad \forall \boldsymbol{\alpha} \in \mathbb{R}^p$$

  (□ Probar)

Sea $\mathbf{X} = (X_1, \ldots, X_p)'$ un vector aleatorio con función característica $\phi_{\mathbf{X}}(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^p$

- <u>MOMENTOS</u>

  Se define el *momento (no centrado)* p-dimensional de orden $(r_1, \ldots, r_p)$ de $\mathbf{X}$ como

$$\mu^{1\ldots p}_{r_1\ldots r_p} = E\left[X_1^{r_1} \cdots X_p^{r_p}\right]$$

  Los momentos pueden obtenerse a partir de la función característica, derivándose de su expansión de Taylor (respecto al origen):

$$\begin{aligned} \phi(\mathbf{t}) &= E[e^{i\mathbf{t}'\mathbf{x}}] = E\left[\sum_{r=0}^{\infty} \frac{1}{r!}(i\mathbf{t}'\mathbf{X})^r\right] \\ &= \sum_{r=0}^{\infty} \sum_{r_1+\cdots+r_p=r} \mu^{1\ldots p}_{r_1\ldots r_p} \frac{(it_1)^{r_1}\cdots(it_p)^{r_p}}{r_1!\cdots r_p!} \end{aligned}$$

En particular, se tiene el siguiente resultado:

- TEOREMA: Si $E\left[|X_1^{m_1}|\cdots|X_p|^{m_p}\right] < \infty$, entonces la función característica de $\mathbf{X}$ es $(m_1, \ldots, m_p)$ veces continuamente diferenciable, y
$$\left.\frac{\partial^m}{\partial_{t_1}^{m_1}\cdots\partial_{t_p}^{m_p}}\phi(\mathbf{t})\right|_{\mathbf{t}=\mathbf{0}} = i^m E\left[X_1^{m_1}\cdots X_p^{m_p}\right] = i^m \mu_{m_1\ldots m_p}^{1\ldots p}$$
$(m = m_1 + \cdots + m_p)$

Sea $\mathbf{X} = (X_1, \ldots, X_p)'$ un vector aleatorio con función característica $\phi_{\mathbf{X}}(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^p$. Consideremos la función

$$\log \phi(\mathbf{t})$$

- <u>CUMULANTES</u>

  Se definen los *cumulantes* p-dimensionales de orden $(r_1, \ldots, r_p)$ de $\mathbf{X}$ como los coeficientes de la correspondiente expansión (respecto al origen),

$$\log \phi(\mathbf{t}) = \sum_{r=0}^{\infty} \sum_{r_1+\cdots+r_p=r} \kappa_{r_1\ldots r_p}^{1\ldots p} \frac{(it_1)^{r_1} \cdots (it_p)^{r_p}}{r_1! \cdots r_p!}$$

- TEOREMA: Sea $\mathbf{X} = (X_1, \ldots, X_p)'$ un vector aleatorio con función de densidad $f_{\mathbf{X}}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^p$, positiva sobre $S \subseteq \mathbb{R}^p$ y continua.

  Sea $\mathbf{Y} = (Y_1, \ldots, Y_p)'$ un vector aleatorio con
  $$\mathbf{Y} = \mathbf{g}(\mathbf{X}) = (g_1(\mathbf{X}), \ldots, g_p(\mathbf{X}))',$$
  siendo la restricción
  $$\mathbf{g} = (g_1, \ldots, g_p)' : S \longrightarrow T \equiv \mathbf{g}(S) \subseteq \mathbb{R}^p$$
  una aplicación biyectiva, y sea $\mathbf{g}^{-1} =: \mathbf{h} = (h_1, \ldots, h_p)'$. Supongamos que existen las derivadas parciales
  $$\frac{\partial h_i(\mathbf{y})}{\partial y_j}, \qquad (i, j = 1, \ldots, p)$$
  y son continuas sobre T.

  Entonces, la función de densidad $f_{\mathbf{Y}}(\mathbf{y})$, $\mathbf{y} \in \mathbb{R}^p$, del vector aleatorio $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ viene dada por
  $$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{g}^{-1}(\mathbf{y})) \cdot \mathrm{abs}(J_{\mathbf{g}^{-1}}(\mathbf{y}))$$
  siendo $J_{\mathbf{g}^{-1}}(\mathbf{y}) = \left[J_{\mathbf{g}}(\mathbf{g}^{-1}(\mathbf{y}))\right]^{-1}$ el determinante de la matriz jacobiana de la transformación $\mathbf{g}^{-1}$

- CASO LINEAL: Sean

  $\mathbf{X}$ vec.a. p-dimensional

  $\mathbf{Y}$ vec.a. p-dimensional, definido por la transformación

  $$\mathbf{Y} = \mathbf{g}(\mathbf{X}) := B\mathbf{X} + \mathbf{b}, \quad \text{con}$$

  B matriz $p \times p$ (cte.) no singular

  $\mathbf{b}$ vector $p \times 1$ (cte.)

  Entonces, en este caso, se tiene que

  $$\mathbf{X} = B^{-1}(\mathbf{Y} - \mathbf{b})$$

  $$J_{\mathbf{g}^{-1}}(\cdot) \equiv |B^{-1}| = |B|^{-1}$$

  $$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(B^{-1}(\mathbf{y} - \mathbf{b})) \cdot \text{abs}(|B|^{-1})$$