TEMA 1 Introducción

1.1 Conceptos básicos de la inferencia estadística

Def: (Muestra aleatoria)

X v.a.

La contaminación por cobre. Variable, pues en cada parte puede haber una contaminación distinta.

Luego una muestra aleatoria de X es un conjunto $X_1, \dots, X_n$ i.i.d.

En la práctica tenemos una observación de $X_1, \dots, X_n$, $x_1, \dots, x_n$, que es una realización muestral.

Observación

Nos podemos hacer varias preguntas $\begin{cases} ¿E(X) = \mu? \to \text{Cont. media de cobre. Est. puntual} \\ ¿IC(\mu; 0'95)? \to \text{Estimación por Int. de conf. o orquillas} \\ ¿H_0 : \mu = 20? \to \text{Contraste de hipótesis.} \end{cases}$

Con estas herramientas debemos ~~responder~~ responder a las preguntas que nos formulan. Usando la realización muestral.

Cuando no solo estudiamos una única variable si no que queremos estudiar más por ejemplo X_1, X_2, X_3

$$\text{Co} \begin{cases} X_1 \equiv \text{cant. cobre} \\ X_2 \equiv \text{cant. hierro} \\ X_3 \equiv \text{Cant. Pb restos} \end{cases} \Rightarrow \underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

→ Ahora me planteo el problema con el vector $\underline{X}$ y hago un estudio de las 3 variables a la vez.

Se puede hacer un estudio de cada variable de forma individual. Pero esto presenta algunos inconvenientes:

1. Se pierde toda la relación que puedan existir entre las variables. No es exactamente que se pierda, sino, que con el estudio individual no se tiene en cuenta.
2. En los contrastes de hipótesis es acumula la ~~propa~~ probabilidad de error de tipo 1.

$$\boxed{\begin{aligned} &① \; S_c^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2 \\ &② \; S^2 = \frac{(n-1)}{n} \sum (X_i - \bar{X})^2 \\ &③ \; \bar{X} = \frac{1}{n} \sum X_i \end{aligned}}$$

1.2. Dos muestras aleatorias relacionadas

$$(P) \begin{cases} X = \text{cant. per cobre en la superficie} \\ Y = \text{cant. per cobre en 10 cm.} \end{cases} \Big\} \rightarrow \text{Tengo una muestra de la forma } \begin{pmatrix} X \\ Y \end{pmatrix}$$

Nos podemos plantear el contraste de hip. $H_0 : \mu_X = \mu_Y$.

$$\rightarrow D := X - Y \rightarrow H_0 : \underset{\mu_X - \mu_Y = 0}{\mu_D = 0} \iff H_0 : \mu_Y = \mu_X$$

↳ Me puedo plantear el problema de 1 variable y aplicar lo mismo que en el caso 1-dimensional.

Recordemos...

Sobre una m.a $X_1, \dots, X_m$ (i.i.d. a X) existían 2 estadísticos fundamentales:

- $\bar{X} = \frac{1}{m} \sum_{k=1}^{m} X_k$
- $S_c^2 = \frac{1}{m-1} \sum_{k=1}^{m} (X_k - \bar{X})^2 = \frac{m}{m-1} S^2 \left(= \frac{1}{m} \sum X_i^2 - \bar{X}^2 \right)$

$\bar{X}$ y S_c^2 son estimadores insesgados de μ y σ^2 respectivamente, pues $E(\bar{X}) = \mu$ y $E(S_c^2) = \sigma^2$ $\left(E(S^2) = \frac{\sigma^2}{m} \Rightarrow E(mS^2) = \sigma^2 \right)$

1.3 Características poblacionales y muestras aleatorias

Caso p-dimensional $\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix} = [X_1, \dots, X_p]'$

1.3.1 Características (Parámetros) poblacionales

⇝ Vector de medias $E(\underline{X}) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} = \underline{\mu}$

⇝ Matriz de varianzas y covarianzas (MVC)

$$\sigma_{jk} = cov(X_j, X_k) = E((X_j - \mu_j)(X_k - \mu_k))$$

$$\sigma_j^2 = \sigma_{jj} = Cov(X_j, X_j) = Var(X_j)$$

$$\Sigma = cov(\underline{X}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} = [\sigma_{jk}], \ j,k = 1, \dots, p.$$

(MVC es simétrica pues $Cov(X, Y) = Cov(Y, X)$)

Mat. de correlaciones

$$\rho_{jk} = \frac{\sigma_{jk}}{\sigma_j \sigma_k}, \quad \rho_{jj} = 1;$$

- R :: Mat. de correlaciones $= (\rho_{jk})_{j,k=1,\dots,p}$.
- $R = \Delta^{-1} \Sigma \Delta^{-1}$ con $\Delta \equiv diag[\sigma_1, \dots, \sigma_p]$.

Nota

① $V(X) = E[(X-\mu)^2] = E(X^2) - \mu^2$

② $Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)] = E(XY) - E(X)E(Y) = Cov(Y,X)$

③ Coeficiente de correlación lineal entre X e Y

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

↳ Propiedades

- $\rho \in [-1, 1]$
- $\rho^2 \in [0,1]$
- El signo de ρ depende de $Cov(X,Y)$
- $\rho^2 = 0 \iff X, Y$ son incorreladas $(Cov(X,Y)=0)$
  - X, Y ind $\Rightarrow X, Y$ incorreladas
  - X, Y ind $\iff X, Y$ incorreladas → Normal Multivariante.
- $\rho^2 = 1 \iff \exists a, b \mid Y = a + bX$

El coeficiente de correlación lineal mide cuanta relación lineal existe entre X e Y.

④ R es simétrica

⑤ Elementos básicos de un vector X
$\begin{cases} \text{Vector de Medias} \\ \text{Mat. V.C} \\ \text{Mat. de correlaciones.} \end{cases}$

1.3.2 Vector de variables estandarizadas

$$X(s) = \begin{bmatrix} \frac{X_1 - \mu_1}{\sigma_1} \\ \vdots \\ \frac{X_p - \mu_p}{\sigma_p} \end{bmatrix} = \Delta^{-1}(X - \mu)$$

$$X \begin{cases} E(X) = \mu \\ V(X) = \sigma^2 \end{cases} \rightsquigarrow Y = \frac{X-\mu}{\sigma}$$

$$\Rightarrow Y \begin{cases} E(Y) = 0 \\ V(Y) = \frac{V(X-\mu)}{\sigma^2} = \frac{V(X)}{\sigma^2} = 1 \end{cases}$$

Las técnicas estadísticas dependen mucho de la unidad con la que se miden los experimentos. Luego el procedimiento de estandarización es muy usado en la práctica para poder obtener variables a-dimensionales

La matriz de correlaciones no se ve afectada por la estandarización pero la M.V.C. sí.

$$\boxed{R = \Sigma_{(s)} = R_{(s)}}$$

1.3.3 Muestra aleatoria de $\underline{X} \sim (\underline{\mu}, \Sigma)$

~> $\underline{X}_1, \dots, \underline{X}_n$ iid a $\underline{X}$.

- Elementos muestrales

$$\underline{X}_i = \begin{bmatrix} X_{i1} \\ \vdots \\ X_{ip} \end{bmatrix} \quad \underline{X}_i^t = (X_{i1}, \dots, X_{ip}), \; i = 1, \dots, n.$$

~> Matriz (aleatoria) de datos muestrales

Importante →

$$\boxed{\underset{n \times p}{\underline{X}}} = \begin{bmatrix} \underline{X}_1^t \\ \vdots \\ \underline{X}_n^t \end{bmatrix} = \begin{bmatrix} X_{11} & \dots & X_{1p} \\ \vdots & & \vdots \\ X_{n1} & \dots & X_{np} \end{bmatrix}$$

Datos correspondientes a la variable 1 ... Datos corresp. a la variable p.

→ Matriz de datos
- Tantas filas como elementos tenga en la muestra
- Tantas columnas como variables

- X_{ij} es independiente de X_{kt}, para $i \neq k$
- X_{ij} no es independiente de X_{it}

Nota

~> Dos elementos de una fila no tienen porqué ser independientes pues son componentes del vector base $\underline{X}$.

~> Dos elementos de distintas ~~columnas~~ filas son indep. al ser $\underline{X}_1, \dots, \underline{X}_n$ una m.a.

~> La columna i es una m.a. de X_i

~> Y la matriz aleatoria es una M.a. de $\underline{X}$

1.4. Estadísticos muestrales

1.4.1 Vector de medias

Importante →

$$\bar{\underline{X}} = \frac{1}{n} \sum_{i=1}^{n} \underline{X}_{ij} = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^{n} X_{i1} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^{n} X_{ip} \end{bmatrix} = \begin{bmatrix} \bar{X}_1 \\ \vdots \\ \bar{X}_p \end{bmatrix}$$

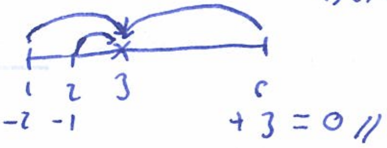
→ $\boxed{\text{Buen estimador de } \mu}$

Propiedades

$1, 2, 6 \quad \bar{X} = 3$

$$\sum_{i=1}^{n} (\underline{X}_i - \bar{\underline{X}}) = \underline{0}$$

~> Esto quiere decir que está en el medio



1.4.2 Matriz de suma de cuadrados y productos cruzados (s.c.p.c)

~> Suma de cuadrados

$$S_{jj} = \sum_{i=1}^{n} (X_{ij} - \bar{X}_j)^2 \quad 1 \leq j \leq p$$

1.4.4 Matriz de covarianzas muestrales

1. (Cuasi-) varianza muestral de la j-ésima componente

$$\hat{\sigma}_j^2 = \hat{\sigma}_{jj} = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij}-\bar{X}_j) = \frac{1}{n-1}S_{jj} \quad 1 \leq j \leq p.$$

2. (Cuasi-) covarianza muestral entre las componentes j y k.

$$\hat{\sigma}_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij}-\bar{X}_j)(X_{ik}-\bar{X}_k) = \frac{1}{n-1}S_{jk}$$

$1 \leq j \neq k \leq p$.

3. MATRIZ DE VARIANZAS Y COVARIANZAS MUESTRALES.

$$\hat{\Sigma} = (\hat{\sigma}_{jk})_{j,k=1,\dots,p}.$$

* $\underset{p\times p}{\hat{\Sigma}}$ es simétrica semidefinida positiva (con prob. 1)

* $$\boxed{\hat{\Sigma} = \frac{1}{n-1}S.}$$

1.4.5 Matriz de correlaciones muestrales

→ Matriz de correlaciones poblacionales y correlaciones poblacionales.

$$\rho_{jk} = \frac{\sigma_{jk}}{\sigma_j\sigma_k} \quad ; \quad \rho_{jj} = 1.$$

$$R = (\rho_{jk})_{1\leq j,k\leq p} = \Delta^{-1}\Sigma\Delta^{-1}.$$

↳ Correlaciones muestrales

$$\hat{\rho}_{jk} = \frac{\hat{\sigma}_{jk}}{\hat{\sigma}_j\hat{\sigma}_k} \quad ; \quad \hat{\rho}_{jj} = 1.$$

$$\hat{R} = \hat{\Delta}^{-1}\hat{\Sigma}\hat{\Delta}^{-1} \quad ; \quad \hat{\Delta} = \text{diag}\{\hat{\sigma}_1,\dots,\hat{\sigma}_p\}.$$

~> Suma de productos cuadrados

$$\boxed{S_{jk} = \sum_{i=1}^{n} (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)} \quad i \leq j \neq k \leq p. \quad \boxed{\to \text{Buen estimador de } \Sigma}$$

~> Matriz s.c.p.c.

$$\boxed{S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix} = [S_{jk}]}$$

$$S = \sum_{i=1}^{n} (\underline{X}_i - \underline{\bar{X}})(\underline{X}_i - \underline{\bar{X}}^t) = \sum_{i=1}^{n} \underline{X}_i \underline{X}_i^t - n \underline{\bar{X}}\,\underline{\bar{X}}^t =$$

$$S = \tilde{\mathbb{X}}^t \tilde{\mathbb{X}} = \mathbb{X}^t H \mathbb{X}$$

Nota

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \ldots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \quad (X_i - \bar{X})(Y_i - \bar{Y}) \quad // \quad \boxed{\text{Cov. muestral} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{n}}$$

- Cov- muestral

$$\frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) = \frac{1}{n}\sum_{n=1}^{n} X_i Y_i - \bar{X}\bar{Y}$$
$$= \overline{XY} - \bar{X}\bar{Y}$$

- Coef. de correlación lineal Muestral

$$\frac{\overline{XY} - \bar{X}\bar{Y}}{S_{cX}^2 S_{cY}^2} \quad // \quad \rho = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

1.4.3 Matriz de datos centrados

$$\boxed{\tilde{\mathbb{X}}} = \begin{bmatrix} X_{11} - \bar{X}_1 & X_{12} - \bar{X}_2 & \cdots & X_{1p} - \bar{X}_p \\ X_{21} - \bar{X}_1 & X_{22} - \bar{X}_2 & \cdots & X_{2p} - \bar{X}_p \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} - \bar{X}_1 & X_{n2} - \bar{X}_2 & \cdots & X_{np} - \bar{X}_p \end{bmatrix} = \begin{bmatrix} \underline{X}_1^t - \underline{\bar{X}}^t \\ \underline{X}_2^t - \underline{\bar{X}}^t \\ \vdots \\ \underline{X}_n^t - \underline{\bar{X}}^t \end{bmatrix}$$

Importante.

① $$\tilde{\mathbb{X}} = \mathbb{X} - 1_n \underline{\bar{X}}^t = \mathbb{X} - \frac{1}{n}\underbrace{1_n 1_n^t}_{E_n}\mathbb{X} = \mathbb{X} - \frac{1}{n}E_n \mathbb{X} = \underbrace{\left[I_n - \frac{1}{n}E_n\right]}_{H}\mathbb{X}$$

$$\Rightarrow \tilde{\mathbb{X}} = H\mathbb{X} \ ; \quad H \text{ es simétrica e idempotente } (H^2 = H)$$

1.4.6 Elementos estandarizados

① Vector de variables estandarizadas.

$$\underline{X}_{(S)} = \begin{bmatrix} \frac{X_1 - \mu_1}{\sigma_1} \\ \vdots \\ \frac{X_p - \mu_p}{\sigma_p} \end{bmatrix} = \Delta^{-1}(\underline{X} - \underline{\mu})$$

$\Delta^{-1} = \text{diag}(1/\sigma_1, \dots, 1/\sigma_p)$

② Vector de Datos estandarizados

$$\underline{X}_{i(S)} = \begin{bmatrix} \frac{X_{i1} - \overline{X}_1}{\hat{\sigma}_1} \\ \vdots \\ \frac{X_{ip} - \overline{X}_p}{\hat{\sigma}_p} \end{bmatrix} = \hat{\Delta}^{-1}(\underline{X}_i - \overline{\underline{X}})$$

③ Matriz de datos estandarizados

$$\boxed{\mathbb{X}_{(S)}} = \begin{bmatrix} \frac{X_{11} - \overline{X}_1}{\hat{\sigma}_1} & \frac{X_{12} - \overline{X}_2}{\hat{\sigma}_2} & \cdots & \frac{X_{1p} - \overline{X}_p}{\hat{\sigma}_p} \\ \vdots & & & \\ \frac{X_{m1} - \overline{X}_1}{\hat{\sigma}_1} & \frac{X_{m2} - \overline{X}_2}{\hat{\sigma}_2} & \cdots & \frac{X_{mp} - \overline{X}_p}{\hat{\sigma}_p} \end{bmatrix} = \begin{bmatrix} \underline{X}_1^t - \overline{\underline{X}}^t \\ \vdots \\ \underline{X}_m^t - \overline{\underline{X}}^t \end{bmatrix} \hat{\Delta}^{-1} = \boxed{\tilde{\mathbb{X}} \hat{\Delta}^{-1}}$$

$$\boxed{\hat{R} = \hat{\Sigma}_{(S)} = \hat{R}_{(S)}}$$

Esto quiere decir que la matriz de correlaciones muestrales es lo mismo que la Mat. var.cov. muestral estandarizada.

1.5 Transformaciones lineales

Dada $A_{m \times p}$ y sea $\underline{Y}_{m \times 1} = \underset{m \times p}{A} \, \underset{p \times 1}{\underline{X}}$ ⇝ muestra transformada $\underline{Y}_1, \dots, \underline{Y}_m$ con $\underline{Y}_i = A \underline{X}_i$

① $\overline{\underline{Y}} = A \overline{\underline{X}}$

② $S_{\underline{Y}} = A S_{\underline{X}} A^t$

$\hat{\Sigma}_{\underline{Y}} = A \hat{\Sigma}_{\underline{X}} A^t$

Prueba

1. $$\frac{1}{m}\sum_{i=1}^{m} \underline{Y}_{ij} = \frac{1}{m}\sum_{i=1}^{m} A\,\underline{X}_{ij} = \frac{1}{m} A \sum_{i=1}^{m} \underline{X}_{ij} = A\begin{pmatrix} \bar{\underline{X}}_1 \\ \vdots \\ \bar{\underline{X}}_p \end{pmatrix} = A\bar{\underline{X}}$$

2. $$S_{\underline{Y}} = \sum_{i=1}^{m} (\underline{Y}_i - \bar{\underline{Y}})(\underline{Y}_i - \bar{\underline{Y}})^t$$
$$= \sum_{i=1}^{m} (A\underline{X}_i - A\bar{\underline{X}})(A\underline{X}_i - A\bar{\underline{X}})^t$$
$$= A \sum (\underline{X}_i - \bar{\underline{X}})(\underline{X}_i - \bar{\underline{X}})^t A^t = A\hat{\Sigma}_x A^t$$

Análogo para $\hat{\Sigma}_{\underline{Y}} = A\hat{\Sigma}_{\underline{X}} A^t$ ya que $\hat{\Sigma} = \frac{1}{m-1} S$ $\square$

1.6 Matrices Aleatorias

Def: • Sea $\{Z_{ij}, i=1,\dots,m; j=1,\dots,g\}$ un conjunto de v.a.

$\Rightarrow$ Matriz aleatoria $\rightarrow$ $$\mathbf{Z} = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1g} \\ \vdots & \vdots & & \\ Z_{m1} & Z_{m2} & \cdots & Z_{mg} \end{bmatrix} = [Z_{ij}]$$

1.6.1 Operador Esperanza. Propiedades

Def: (Matriz esperanza o matriz de medias)

$$E[\mathbf{Z}] = \begin{bmatrix} E[Z_{11}] & \cdots & E[Z_{1g}] \\ \vdots & & \\ E[Z_{m1}] & \cdots & E[{mg}] \end{bmatrix} = [E(Z_{ij})]$$

Lema 1.6.1.1

Sea $\mathbf{Z}$ e $\mathbf{Y}$ dos matrices aleatorias de la misma dimensión. Sean A, B y C matrices constantes de las dimensiones adecuadas.

1. $E[A\mathbf{Z}B + C] = A\,E[\mathbf{Z}]\,B + C.$

2. $E[\mathbf{Z} + \mathbf{Y}] = E[\mathbf{Z}] + E[\mathbf{Y}].$

Lema 1.6.1.2

Sea $\underline{X} \sim (\underline{\mu}, \Sigma)$

1. $\Sigma = E((\underline{X} - E(\underline{X}))(\underline{X} - E(\underline{X}))^t) = E[\underline{X}\underline{X}^t] - \underline{\mu}\underline{\mu}^t$

Prueba

$\underbrace{(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^t}$

Matriz $p \times p$ $\sim$ su elemento $(j,k) \to (X_j - \mu_j)(X_k - \mu_k)$

$$\Rightarrow E((\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^t) = [E(j,k)]$$

$$E(j,k) = E((X_j - \mu_j)(X_k - \mu_k)) = \sigma_{jk} \quad \boxtimes$$

$$E[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^t] = E[\underline{X}\underline{X}^t - \underline{X}\underline{\mu}^t - \underline{\mu}\underline{X}^t + \underline{\mu}\underline{\mu}^t]$$

$$= E(\underline{X}\underline{X}^t) - \underbrace{\cancel{E(\underline{X})}}_{\mu}\mu^t - \mu(\underbrace{\cancel{E(\underline{X}^t)}}_{\mu^t}) + \mu\mu = \underbrace{E(\underline{X}\underline{X}^t) - \underline{\mu}\underline{\mu}^t}.$$

$V(X) = E(X^2) - \mu^2.$

2. $E(A\underline{X} + \underline{b}) = A E[\underline{X}] + \underline{b}$

3. $Cov(A\underline{X} + \underline{b}) = \Sigma_{A\underline{X}+\underline{b}} = A \Sigma A^t$

Prueba

$A = [a_{ij}]$, $\underline{b} = (b_i)$ $\qquad \underline{Y} = A\underline{X} + \underline{b}$

$$E \quad \Sigma_{\underline{Y}} = E((\underbrace{\underline{Y}}_{A\underline{X}+\underline{b}} - \underbrace{\mu_Y}_{A\mu_x + b})(\underline{Y} - \mu_Y)^t) = E[A(\underline{X} - \mu_X)(\underline{X} - \mu_X)^t A^t] =$$

$$= A \Sigma_{X} A^t.$$

4. En particular para $\underline{c} \in \mathbb{R}^p$

$$E(\underline{c}^t \underline{X}) = \underline{c}^t \underline{\mu}$$

$$Var(\underline{c}^t \underline{X}) = \underline{c}^t \Sigma \underline{c}$$

5. Σ es semidefinida positiva. $(\Sigma \geq 0)$

Prueba

Dado cualquier $c \sim \underline{c}^t \Sigma \underline{c} \geq 0$ como $\underline{c}^t \Sigma \underline{c} = \underbrace{Var[\underline{c}^t \underline{X}]}\geq 0$

$\Rightarrow \Sigma$ es ~~real~~ semidefinida positiva.

Esto es una cantidad positiva pues es la var. de una v.a.

Def: (Mat. de cov.)

Sean $\underline{X}$ e $\underline{Y}$ dos vectores aleatorios. Se define la matriz de covarianza entre $\underline{X}$ e $\underline{Y}$:

$$Cov(\underline{X},\underline{Y}) = [Cov(X_i, Y_j)]$$

Lema 1.6.1.3

① Si $\underline{X}$ e $\underline{Y}$ son independientes $\Rightarrow Cov(\underline{X},\underline{Y}) = \mathbb{O}_{p\times q}$

② $Cov(\underline{X}) = Cov(\underline{X},\underline{X}) = \Sigma_{\underline{X}}$

③ $Cov(\underline{X},\underline{Y}) = E[(\underline{X}-E(\underline{X}))(\underline{Y}-E(\underline{Y}))^t]$

④ $Cov(\underline{X},\underline{Y}) = Cov(\underline{Y},\underline{X})^t$

Prueba

$\underline{X}+\underline{Y}=\underline{Z}$ $\quad E_{\underline{Z}} \doteq E((\underline{Z}-\mu_Z)(Z-\mu_Z)^t) = E[(\underline{X}-\mu_x+\underline{Y}-\mu_y)(\underline{X}\cdots)^t]$

$$= \Sigma_{\underline{X}} + \underbrace{E[(\underline{X}-\mu_x)(\underline{Y}-\mu_y)^t]}_{Cov(\underline{X},\underline{Y})} + \underbrace{E[(\underline{Y}-\mu_y)(\underline{X}-\mu_{\underline{X}})]}_{Cov(\underline{Y},\underline{X})} + \Sigma_y$$

⑤ $Cov(A\underline{X}+\underline{c}, B\underline{Y}+\underline{d}) = A\,Cov(\underline{X},\underline{Y})\,B^t$.

⑥ $Cov(\underline{X}_1+\underline{X}_2,\underline{Y}) = Cov(\underline{X}_1,\underline{Y}) + Cov(\underline{X}_2,\underline{Y})$

⑦ Si $p=q \Rightarrow \Sigma_{\underline{X}+\underline{Y}} = \Sigma_{\underline{X}} + Cov(\underline{X},\underline{Y}) + Cov(\underline{Y},\underline{X}) + \Sigma_Y$

⑧ $Cov\left(\sum_{i=1}^{n}\underline{X}_i\right) = \sum_{i=1}^{n}\Sigma_{\underline{X}_i} + \sum_{i\neq j} Cov(\underline{X}_i,\underline{X}_j)$

1.7 Propiedades de los estadísticos muestrales

Proposición 1.7.1

Sea $\underline{X}_1, \dots, \underline{X}_n$ una m.a. de una v.a. $\underline{X}$ p-dim. $\underline{X} \sim (\mu, \Sigma)$

Se verifica:

1. $E(\overline{\underline{X}}) = \mu$
2. $Cov(\overline{\underline{X}}) = \frac{1}{n}\Sigma$
3. $E(\hat{\Sigma}) = E\left[\frac{1}{n-1} S\right] = \Sigma$

→ El elemento j de $\overline{\underline{X}}$ es un estimador insesgado de μ_j.

→ El elemento (j,k) de $\hat{\Sigma}$ es un estimador insesgado de σ_{jk}.

Demostración

1. $\overline{\underline{X}} = \frac{1}{n}\sum_{i=1}^{n} \underline{X}_i \to E(\overline{\underline{X}}) = E\left(\frac{1}{n}\sum_{i=1}^{n} \underline{X}_i\right) = \frac{1}{n}\sum E(\underline{X}_i) = \mu.$

   v.a. i.i.d. a $\underline{X}$

2. $\Sigma_{\overline{\underline{X}}} = cov\left(\frac{1}{n}\cdot\sum_{i=1}^{n} \underline{X}_i\right) = \frac{1}{n^2} cov\left(\sum_{i=1}^{n} \underline{X}_i\right) =$

   $= \frac{1}{n^2}\left\{\sum_{i=1}^{n} \Sigma_{\underline{X}_i} + \underbrace{\sum_{i\neq j} cov(\underline{X}_i, \underline{X}_j)}_{0}\right\} = \frac{1}{n}\Sigma_{\underline{X}}$

   0 por ser una m.a. ya que son ind.

3. $S = \underbrace{\sum_{i=1}^{n} \underline{X}_i \underline{X}_i^t}_{} \underbrace{- n\overline{\underline{X}}\,\overline{\underline{X}}^t}_{}$

$$E(S) = \underbrace{n\,E[\underline{X}\,\underline{X}^t]}_{} - \underbrace{n\,E(\overline{\underline{X}}\,\overline{\underline{X}}^t)}_{} = n\Sigma_{\underline{X}} - n\Sigma_{\overline{\underline{X}}} =$$

$$n\left(\Sigma_{\underline{X}} + \cancel{\mu\mu^t} - \Sigma_{\overline{\underline{X}}} - \cancel{E(\overline{\underline{X}})E(\overline{\underline{X}})^t}\right) \quad E(\overline{\underline{X}}) = \mu,\ E(\overline{\underline{X}})^t = \mu^t$$

$$= n\Sigma_{\underline{X}} - n\frac{1}{n}\Sigma_{\underline{X}} = \Sigma_{\underline{X}}(n-1)$$

$$E[S] = \Sigma_{\underline{X}}(n-1) \leadsto E[\hat{\Sigma}] = E\left[\frac{1}{n-1}S\right] = \Sigma_{\underline{X}} \qquad \emptyset$$

Def:

1. Var. muest. gen: $|\hat{\Sigma}| = \left|\frac{1}{n-1}S\right|$
2. Var. muest. total: $tr(\hat{\Sigma}) = \frac{1}{n-1} tr(S)$
3. Var. muest. efectiva: $|\hat{\Sigma}|^{1/p} = \frac{1}{n-1}|S|^{1/p}$.

TEMA 2 Inferencia en poblaciones normales

2.1 Distribución normal.

2.1.1. Caso unidimensional $X \sim N(\mu, \sigma^2)$

① $f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}(X-\mu)^2\right)$; $-\infty < x < \infty$

② $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$; $E(X)=\mu$. $V(X)=\sigma^2$.

③ $aX+b \sim N(a\mu+b; a^2\sigma^2)$

④ X, Y ind. $\Rightarrow X+Y \sim N(\mu_x+\mu_y; \sigma_x^2+\sigma_y^2)$

2.1.2 Caso multidimensional

① $f(\underline{X}) = f(x_1,\dots,x_p) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\underline{X}-\underline{\mu})^t\Sigma^{-1}(\underline{X}-\underline{\mu})\right\}$

$-\infty < x_i < \infty$; $i=1,\dots,p$.

② $Z = \underset{p\times p}{\Sigma^{-1/2}} (\underset{p\times 1}{\underline{X}} - \underset{p\times 1}{\underline{\mu}}) \sim N_p(0,\Sigma) \rightarrow E[\underline{X}]=\underline{\mu}$; $V(\underline{X})=\Sigma$.

③ $\underset{m\times p}{A}\,\underline{X} + \underset{m\times 1}{b} \sim N_m(A\underline{\mu}+b ; A\Sigma A^t)$

④ $\underline{X}$ y $\underline{Y}$ i.d $\Rightarrow \underline{X}+\underline{Y} \sim N_p(\underline{\mu}_x+\underline{\mu}_y ; \Sigma_x+\Sigma_y)$

2.2. Dist. Chi-cuadrado y Dist. de Wishart.

→ La distribución de Wishart será en el caso multidimensional lo que es χ^2_n en el caso unidimensional.

2.2.1 Dist. Chi-cuadrado

① $X_1,\dots,X_n$ m.a. $N(0,\sigma^2) \Rightarrow \sum_{i=1}^{n} \frac{X_i^2}{\sigma^2} \sim \chi^2_n$; $\sum_{i=1}^{n} X_i^2 \sim \sigma^2\chi^2_n$

② $f_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-1/2\,x} \sim X \sim \chi^2_n \equiv G\left(\frac{n}{2},\frac{1}{2}\right)$ $\begin{cases} E(X)=n. \\ V(X)=2n. \end{cases}$

③ Reproductiva.

2.2.2. Dist. de Wishart

Sea $X_1,\dots,X_n$ una m.a. de una $N_p(\boxed{0},\Sigma)$ y sea $X=[X_1,\dots,X_n]'$

$\Rightarrow \sum_{i=1}^{n} X_i X_i^t \sim W_p(n;\Sigma)$ → Distribución de Wishart p-dim. de n grados de libertad y con matriz de escala Σ.

$W \sim W_p \Rightarrow E[W]=E\left[\sum_{i=1}^{n} X_i X_i^t\right] = \sum_{i=1}^{n} E[X_i X_i^t] \overset{i.i.d}{=} nE\,\underline{X}\,\underline{X}^t = \Sigma n$

$\Sigma = E[(\underline{X}-\underline{\mu})(\underline{X}-\underline{\mu})^t] = E[\underline{X}\,\underline{X}^t]$ pues $\boxed{\underline{\mu}=0}$

2.3. Distr. t-Student y T^2 de Hotelling

2.3.1 D. t-Student

$$\left.\begin{array}{l} Z \sim N(0,1) \\ V \sim \chi^2_m \\ \text{independientes} \end{array}\right\} \quad T = \frac{Z}{\sqrt{V/m}} \sim t_m$$

equivalentemente
$$\left[\begin{array}{l} T^2 = m\,\dfrac{Z^2/1}{V/m} \sim F_{1,m} \\ Z^2 \sim \chi^2_1 \text{ (pues } Z \sim N(0,1)) \end{array}\right]$$

2.3.2 D. T^2 de Hotelling $\boxed{T^2 \sim T^2_{p,m}}$

Sean $\underline{d} \sim N_p(\underline{0} ; I_p)$ y $W \sim W_p(m, I_p)$ independientes.

Se denomina T^2 de Hotelling a la distribución del estadístico

$$T^2 = m \underset{1\times p}{\underline{d}^t} \underset{p\times p}{W^{-1}} \underset{p\times 1}{\underline{d}} \rightarrow 1\times 1 \text{ es un escalar. Variable unidimensional.}$$

Nota

① $T^2_{p,m}$ es una distribución unidimensional

② La T^2 Hotelling salvo parámetros es una F-de Snedecor.

Lema 2.3.1

Sea $\underline{X} \sim N_p(\mu, \Sigma)$ y $W \sim W_p(m, \Sigma)$ independientes.

Entonces:

$$T^2 = m(\underline{X} - \underline{\mu})^t W^{-1} (\underline{X} - \underline{\mu}) \sim T^2_{p,m}.$$

Prueba

$$\underline{X} \sim N_p(\mu ; \Sigma) \Rightarrow \Sigma^{-1/2}(\underline{X} - \mu) \sim N_p(\underline{0}, I_p)$$

$$W (\sim W_p(m ; \Sigma)) = \sum_{i=1}^{m} \underline{X}_i \underline{X}_i^t \; (\underline{X}_i \sim N(0 ; \Sigma)) \Rightarrow \Sigma^{-1/2} W \Sigma^{-1/2}$$

$$= \sum_{i=1}^{m} (\Sigma^{-1/2} \underline{X}_i)(\Sigma^{-1/2} \underline{X}_i)^t \text{ (ambas siguen } N_p(\underline{0}, I_p)) \sim W_p(m ; I)$$

Entra dentro de esta forma y no traspuesta pues es una matriz simétrica.

$$\Rightarrow m(\underline{X} - \mu)^t \cancel{\Sigma^{-1/2}} (\cancel{\Sigma^{-1/2}} W \cancel{\Sigma^{-1/2}})^{-1} \cancel{\Sigma^{-1/2}} (\underline{X} - \mu)$$

$$= m(\underline{X} - \mu)' W^{-1} (\underline{X} - \underline{\mu}) \sim T^2_{p,m}$$

Lema 2.3.2

Sea $T^2 \sim T^2_{p,m}$. Se verifica

$$\frac{m-p+1}{mp} T^2 \sim F_{p,m-p+1} \quad \text{equivalentemente} \quad \left[T^2 \sim \frac{mp}{m-p+1} F_{p;m-p+1} \right]$$

2.4 Estimación de parámetros

Sea $\underline{X}_1, \dots, \underline{X}_m$ una muestra aleatoria de una población $N_p(\underline{\mu}, \Sigma)$

1. Vector media muestral $\overline{\underline{X}} = \frac{1}{m} \sum_{i=1}^{m} \underline{X}_i$

2. Mat. sum. cuad. y prod. cruzados.

$$S = \sum_{i=1}^{m} (\underline{X}_i - \overline{\underline{X}})(\underline{X}_i - \overline{\underline{X}})^t$$

3. Mat. var. y cov. muestrales.

$$\hat{\Sigma} = \frac{1}{m-1} S.$$

Teorema 2.4.1

1. $\overline{\underline{X}} \sim N_p(\underline{\mu}; \frac{1}{m}\Sigma)$
   a) $\overline{\underline{X}}$ es un est. insesgado de $\underline{\mu}$.
2. $(m-1)\hat{\Sigma} = S \sim W_p(m-1, \Sigma)$
   a) $\hat{\Sigma}$ es un est. insesgado de Σ.
3. $\overline{\underline{X}}$ y $\hat{\Sigma}$ son independientes
4. $\overline{\underline{X}}$ y $\tilde{\Sigma} = \frac{1}{m} S$ son estimadores en max. verosimilitud de $\underline{\mu}$ y Σ.