Fórmulas Corte 1 AED

Media

$$\bar{x}_k = \frac{1}{n}\sum_{j=1}^{n} x_{jk}, \quad k = 1,\dots,p \qquad X = \begin{array}{c} \begin{array}{cc} Var_1 & Var_p \end{array} \\ \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \end{array}$$

Varianza y covarianza

$$S = \begin{bmatrix} s_{11} & \cdots & \\ s_{12} & s_{22} & \\ \vdots & & \ddots \; s_{pp} \end{bmatrix}$$

Varianzas

Covarianzas

$$s_k^2 = \frac{1}{n}\sum_{j=1}^{n}(x_{jk} - \bar{x}_k)^2$$

Varianza

$$s_k^2 = \boxed{s_{kk}} \rightarrow \text{Varianza nueva notación.}$$

$$\Rightarrow \sqrt{s_k^2} = \sqrt{s_{kk}} = s_k$$

Covarianza

Coeficiente de correlación

$$r_{ik} = \frac{s_{ik}}{\sqrt{s_{ii}} \cdot \sqrt{s_{kk}}} = \frac{\sum (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)}{\sqrt{\sum (x_{ji} - \bar{x}_i)^2} \cdot \sqrt{\sum (x_{jk} - \bar{x}_k)^2}}$$

desviación estándar de (i, k) normalizada

Distancia

$$d(Q,P) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \rightarrow \text{Distancia euclideana en espacio}$$

la versión estadística de la distancia:

$$x_i^* = \frac{x_i}{\sqrt{s_{ii}}}, \quad d(Q,P) = \sqrt{\sum_{i=1}^{n} (y_i^* - x_i^*)^2} \longrightarrow \text{Distancia estadística}$$

* Se pueden hacer rotaciones a los puntos para quitar correlaciones y calcular la distancia de los puntos así:

$$\text{Matriz de rotación} \leftarrow \begin{bmatrix} \begin{pmatrix} \cos(\theta) & \operatorname{sen}(\theta) \\ -\operatorname{sen}(\theta) & \cos(\theta) \end{pmatrix} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} \rightarrow \text{puntos rotados}$$

$$\underbrace{d(Q,P)}_{} = \sqrt{a_{11}(y_1-x_1)^2 + 2a_{12}(y_1-x_1)(y_2-x_2) + a_{22}(x_2-y_2)^2}$$

Por simetría de la matriz A

→ Luego que existe una matriz que relaciona $\tilde{x}$ con x, se puede calcular la distancia en función de x como unos coeficientes.

→ Los a_{ij} se calculan con base a θ.

$$a_{11} = \frac{\cos^2(\theta)}{\cos^2(\theta)s_{11} + 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{22}} + \frac{\sin^2(\theta)}{\cos^2(\theta)s_{22} - 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{11}}$$

$$a_{22} = \frac{\sin^2(\theta)}{\cos^2(\theta)s_{11} + 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{22}} + \frac{\cos^2(\theta)}{\cos^2(\theta)s_{22} - 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{11}}$$

and

$$a_{12} = \frac{\cos(\theta)\sin(\theta)}{\cos^2(\theta)s_{11} + 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{22}} - \frac{\sin(\theta)\cos(\theta)}{\cos^2(\theta)s_{22} - 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{11}}$$

el caso de distancia N-dimensional:

$$d(Q,P) = \sqrt{\left[a_{11}(y_1-x_1)^2 + a_{22}(y_2-x_2)^2 + \dots + a_{pp}(y_p-x_p)^2 + 2a_{12}(y_1-x_1)(y_2-x_2) + 2a_{13}(y_1-x_1)(y_3-x_3) + \dots + 2a_{p-1,p}(y_{p-1}-x_{p-1})(y_p-x_p)\right]}$$

$$A = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{p1} \\ a_{12} & a_{22} & \cdots & a_{p2} \\ \vdots & \vdots & \ddots & \\ a_{1p} & a_{2p} & & a_{pp} \end{bmatrix}$$

columna fila

⇝ los coeficientes de A siempre producen distancias cuadradas no negativas.

Propiedades de la distancia:

- $D(P,Q) = D(Q,P)$
- $D(P,Q) > 0$ sii $P \neq Q$
- $D(P,Q) = 0$ sii $P = Q$
- $D(P,Q) \leq D(P,R) + D(R,Q)$

Angulo entre 2 vectores

$$\cos(\theta) = \frac{x'y}{\|x\| \cdot \|y\|}$$

Proyecciones

Proyección ortogonal de $\vec{x}$ sobre $\vec{y}$

- $P_y(\vec{x}) = \left(\frac{\vec{x}'\vec{y}}{\vec{y}'\vec{y}}\right)\vec{y}$

longitud de la proyección: $|\cos(\theta)\, y'|$

Vectores y Valores propios

- Sea A una matriz cuadrada λ es el valor propio de A con vector propio $x \neq 0$ si:

$$Ax = \lambda x$$

- Sea $A_{k \times k}$ simétrica cuadrada entonces tiene k eigen values y k eigen vectors.

$$\underbrace{\lambda_1 e_1, \dots, \lambda_1 e_1}_{\text{son ortonormales}}$$

Descomposición espectral

- Sea $A_{k \times k}$ simétrica su descomposición espectral es:

$$A = \lambda_1 \underset{(k\times 1)}{e_1} \underset{(1\times k)}{e_1^T} + \dots + \lambda_k \underset{(k\times 1)}{e_k} \underset{(1\times k)}{e_k^T}$$

Descomposición espectral

$$A = \sum_{i=i}^{k} \lambda_i e_i e_i'$$

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_k} \end{pmatrix}$$

Sea $P = \begin{pmatrix} | & | & & | \\ e_1 & e_2 & \dots & e_k \\ | & | & & | \end{pmatrix}$, $A = P \Lambda P'$, $A^{-1} = P \Lambda^{-1} P'$

Forma cuadrática y Matriz definida positiva

- $x'Ax$ } forma cuadrática
- Si $\forall x \in \mathbb{R}^n$, $0 \leq x'Ax \Rightarrow A$ es def no negativa
- Si A es def no negativa y $x'Ax = 0 \Leftrightarrow x = 0$, entonces A es def. positiva.

1) A es def. positiva $\Leftrightarrow$ todos los valores propios son positivos
2) A son def. no neg $\Leftrightarrow$ todos los valores propios son no negativos

Matriz raiz cuadrada de A

Sea $A_{k \times k}$ simétrica, la matriz raiz cuadrada de A, $A^{1/2}$ es:

$$P \underbrace{\Lambda^{1/2}}_{\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_k} \end{pmatrix}} P' = A^{1/2}$$

$A^{1/2}$ satisface:

- $(A^{1/2})^T = A^{1/2}$
- $A^{1/2} A^{1/2} = A$
- $(A^{1/2})^{-1} = P \Lambda^{-1/2} P'$
- $A^{1/2} A^{-1/2} = I_{k \times k}$
- $A^{-1/2} A^{-1/2} = A^{-1}$

Vectores y Matrices aleatorias

$\underbrace{X, Y}_{\text{Matrices de vect. aleatorios}}$

- el valor esperado de una matriz aleatoria es la matriz de valores esperados de cada entrada.
- $E(X+Y) = E(X) + E(Y)$

  $E(\underbrace{A}_{\text{Matriz cte}} X \underbrace{B}_{\text{Matriz cte}}) = A\,E(X)\,B$

- Sea X vector aleatorio. Su distribución se puede escribir mediante la distribución conjunta de sus entradas:

$$f_{x_1,\dots,x_p}(x_1,\dots,x_p) = \underbrace{f_X(x)}_{\text{Muchas veces es la Normal Multivariada}}$$

Si $x_1, \dots, x_p$ son independientes:

$$f_{x_1,\dots,x_p}(x_1,\dots,x_p) = \prod_{i=1}^{p} f_i(x_i) \quad \forall\, x_i, \dots, x_p$$

Recuerde

$$X_i, X_k \text{ indep} \Rightarrow Cov(X_i, X_k) = 0$$

Matriz Covarianzas y Varianzas

$$\Sigma = E\left((\underbrace{X}_{\text{vector aleatorio}} - \underbrace{\mu}_{E(X)})(X-\mu)^T\right)$$

- p varianzas en la diagonal (σ_{ii})
- $\frac{p(p-1)}{2}$ covarianzas diferentes (σ_{ij})

$$\Sigma = E\left(\begin{pmatrix} x_1-\mu_1 \\ \vdots \\ x_p-\mu_p \end{pmatrix}(x_1-\mu_1, \dots, x_p-\mu_p)\right) = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{1p} & \cdots & \sigma_{pp} \end{pmatrix}$$

Matriz correlación poblacional

$$\rho = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1p} \\ \vdots & \ddots & \vdots \\ \rho_{1p} & \cdots & \rho_{pp} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & \rho_{1p} \\ \vdots & \ddots & \vdots \\ \rho_{1p} & \cdots & 1 \end{pmatrix}$$

$$\rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}}$$ → Mide la asociación lineal que hay entre las variables

Vectorialmente:

$$V^{1/2} = \begin{pmatrix} \sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}} \end{pmatrix}$$

matriz de desviación estándar

1) $V^{1/2} \rho V^{1/2} = \Sigma$ → mat varianzas y covarianzas

2) $\rho = (V^{1/2})^{-1} \Sigma (V^{1/2})^{-1}$

Particiones de vectores y matrices

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_q \\ \hline X_{q+1} \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} X^{(1)} \\ \hline X^{(2)} \end{pmatrix}, \quad \mu = E(X) = \begin{pmatrix} E(X^{(1)}) \\ \hline E(X^{(2)}) \end{pmatrix} = \begin{pmatrix} \mu^{(1)} \\ \hline \mu^{(2)} \end{pmatrix}$$

$$\Sigma_{(1)(2)} = E\left((X^{(1)} - \mu^{(1)}) \cdot (X^{(2)} - \mu^{(2)})^T \right) = E\left(\begin{pmatrix} X_1 - \mu_1 \\ \vdots \\ X_q - \mu_q \end{pmatrix} (X_{q+1} - \mu_{q+1}, \dots, X_p - \mu_p) \right)$$

$$= \begin{pmatrix} \sigma_{1,q+1} & \cdots & \sigma_{1,p} \\ \vdots & \ddots & \\ \sigma_{q,q+1} & & \sigma_{qp} \end{pmatrix}$$

Matriz de cov. y var. entre componentes $X^{(1)}$ y $X^{(2)}$

→ No necesariamente es simétrica o cuadrada

μ y Σ para combinaciones lineales

- $E(aX+b) = aE(X) + b$
- $Var(aX+b) = a^2 Var(X)$
- $a, b, c \in \mathbb{R}$ X, Y, Z v.a's

→ $Cov(aX, bY) = ab\, Cov(X,Y)$

→ $Cov(X, Y+c) = Cov(X,Y)$

→ $Cov(X, Y+Z) = Cov(X,Y) + Cov(X+Z)$

Vectorialmente:

Sea $c = \begin{pmatrix} a \\ b \end{pmatrix}$, $\quad c'X = [a \ b] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = aX_1 + bX_2$

$$E(c'X) = aE(X_1) + bE(X_2) = a\mu_1 + b\mu_2 = \boxed{c'\mu}$$

$$Var(c'X) = a^2 Var(X_1) + b^2 Var(X_2) + 2ab\, Cov(X_1, X_2)$$

$$= a^2\sigma_{11} + b^2\sigma_{22} + 2ab\sigma_{12}$$

$$\boxed{= c'\Sigma c}$$

Generalizado:

$$Z = \underbrace{\begin{bmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{q1} & \cdots & c_{qp} \end{bmatrix}}_{(q \times p)} \underbrace{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}}_{(p \times 1)} = CX, \quad E(Z) = C\mu_X, \quad Cov(Z) = C\Sigma_X C'$$

Muestreo y geometría

Notación:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ x_{21} & \cdots & x_{2p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} X_1' \\ X_2' \\ \vdots \\ X_p' \end{pmatrix}$$

X_1' → 1ra observación multivariada

X_p' → n-ésima observación multivariada

$$= \left(\underbrace{Y_1}_{\text{vector de mediciones de la variable 1}} \mid Y_2 \mid \cdots \mid \underbrace{Y_p}_{\text{vector de mediciones de la variable } p} \right)$$

1_n → vector de unos n dimensional (tiene mil usos).

$\frac{1}{\sqrt{n}} \cdot 1_n$ → vector de unos n dimensional normalizado.

Recuerde:

Proyección ortogonal de $\vec{x}$ sobre $\vec{y}$

- $P_y(\vec{x}) = \left(\frac{\vec{x}'\vec{y}}{\vec{y}'\vec{y}} \right) \vec{y}$

Proyectamos $Y_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$ sobre $\frac{1}{\sqrt{n}} \cdot 1_n$

$$\left(\frac{Y_i' \cancel{\frac{1}{\sqrt{n}}} \cdot 1_n}{\cancel{\left(\frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}}\right)} 1_n' 1_n} \right) \cancel{\frac{1}{\sqrt{n}}} 1_n = \frac{Y_i'}{n} \vec{1}_n \vec{1}_n$$

$$= \frac{(x_{1i}, \ldots, x_{ni})}{n} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= \left(\frac{x_{1i} + \ldots + x_{ni}}{n} \right) \vec{1}_n$$

$$= \bar{x}_i \vec{1}_n$$

la proyección ortogonal de un vector sobre $\vec{1}_n$ es su media por el vector $\vec{1}_n$

el vector d_i:

- $d_i = Y_i - \bar{X}_i 1_n = \begin{pmatrix} x_{1i} - \bar{x}_i \\ \vdots \\ x_{ni} - \bar{x}_i \end{pmatrix}$
- $L_{d_i} = \sqrt{d_i' d_i}$, $L^2_{d_i} = d_i' d_i = \sum_{j=1}^{n} (x_{ji} - x_i)^2 = n s_{ii}$

$$\frac{L^2_{d_i}}{n} = s_{ii}$$

- $d_i' d_k = \sum (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$

$$= L_{d_i} \cdot L_{d_k} \cdot \cos(\theta_{ik}) \Rightarrow \cos(\theta_{ik}) = \frac{s_{ik}}{\sqrt{s_{ii}}\sqrt{s_{kk}}} = r_{ik}$$

Distribuciones muestrales de $\bar{X}$ y S_n

Sean x_i vectores aleatorios tomados de una población con media poblacional μ $(E(x_i) = \mu)$ y cov. poblacional Σ $(var(x_i) = \Sigma)$

entonces:

1) $E(\bar{X}) = \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$

2) $E(S_n) = \frac{n-1}{n} \Sigma$

$E\left(\underbrace{\frac{n}{n-1} \cdot S_n}_{\text{estimador insesgado de } \Sigma}\right) = \Sigma$

3) $\overbrace{S}^{\text{Insesgado}} = \frac{1}{n-1} \cdot \sum_i (x_{ij} - \bar{x}_i)^2$

$\underbrace{S_n}_{\text{sesgado}} = \frac{1}{n} \cdot \sum_i (x_{ij} - \bar{x}_i)^2$

Varianza muestral generalizada $= \underbrace{|S|}$

- es cero si todas las columnas no son linealmente independientes
- tiene el problema de que si una variable tiene mucha variabilidad $|S|$ se ve muy afectada por esto.

$d_i = \vec{y}_i - \bar{x}_i(\vec{1}_n)$, es más grande si hay mucha variabilidad.

entonces si se escala todo: $\frac{(x_{ij} - \bar{x}_i)}{\sqrt{s_{ii}}}$ se puede eliminar el problema de la variabilidad y volver un valor sin unidad de medida a la matriz S. esto es la matriz R (de correlación).

$|R|$ = **Varianza muestral generalizada de las variables estandarizadas.**

Varianza total de la muestra $= \sum_{i=1}^{p} s_{ii}$

Recuerde la **PDF** de **x normal**:

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(x-\mu)/2\sigma^2} \quad x \in \mathbb{R}$$

$\frac{(x-\mu)^2}{\sigma^2}$ → Distancia estadística entre x y μ

Distancia estadística vectorial

$$(\vec{x} - \vec{\mu})' \Sigma^{-1} (\vec{x} - \vec{\mu})$$

PDF normal multivariada

$$f_x(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \cdot e^{-(\vec{x}-\vec{\mu})' \Sigma^{-1} (\vec{x}-\vec{\mu})/2}$$

Normal multivariada

$$\mathbb{X} \sim N_p(\underbrace{\vec{\mu}}_{\text{vector}}, \underbrace{\Sigma}_{\text{matriz}}),$$ Si $\mathbb{X}$ es normal multivariada, $p=2$ y $cov(x_1, x_2) = \rho_{12} = 0 \Rightarrow x_1$ y x_2 son indep.

Teorema

Si Σ def positiva (Σ^{-1} existe) entonces si e es un vector propio de Σ con valor propio asociado λ. e es un vector propio de Σ^{-1} con valor propio asociado $\frac{1}{\lambda}$. Además Σ^{-1} es def. positiva:

$$1) \quad (\vec{x} - \vec{\mu})' (\Sigma^{-1})(\vec{X} - \vec{\mu}) \leq \overbrace{\chi^2_p(\alpha)}^{\text{chi-cuadrada}}$$

Proposiciones Normal

Si $\mathbb{X}$ es un vector aleatorio normal multivariado.

1) $\mathbb{X} \sim N_p(\mu, \Sigma) \Rightarrow$ si $a \in \mathbb{R}^p$, $a'\mathbb{X} = \sum a_i x_i$, $a'\mathbb{X} \sim N(a'\mu, a'\Sigma a')$

2) Si $a'\mathbb{X} \sim N(a'\mu, a'\Sigma a)$, $\forall a \in \mathbb{R}^p \Rightarrow \mathbb{X} \sim N_p(\mu, \Sigma)$

3) $\mathbb{X} \sim N_p(\mu, \Sigma)$, sea $A_{q \times p} \in \mathbb{R}^{q \times p} \Rightarrow A\mathbb{X} \sim N_p(A\mu, A\Sigma A)$

4) $\mathbb{X} \sim N_p(\mu, \Sigma)$, $d \in \mathbb{R}^p \Rightarrow \mathbb{X} + d \sim N_p(\mu + d, \Sigma)$

5) todas las particiones de vectores aleatorios normales resultan en vect. aleatorios normales

$$\mathbb{X} \sim N_p(\mu, \Sigma), \quad \mathbb{X} = \begin{pmatrix} \mathbb{X}_{(1)} \\ \hline \mathbb{X}_{(2)} \end{pmatrix} \begin{matrix} \}q \\ \}p-q \end{matrix}, \quad \mu = \begin{pmatrix} \mu_{(1)} \\ \hline \mu_{(2)} \end{pmatrix}, \quad \Sigma = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right)$$

$$\mathbb{X}_{(1)} \sim N_q(\mu_{(1)}, \Sigma_{11})$$

$$\mathbb{X}_{(2)} \sim N_{p-q}(\mu_{(2)}, \Sigma_{22})$$

6) ⊛ Si $\mathbb{X}_{(1)}^{(q_1 \times 1)}$, $\mathbb{X}_{(2)}^{(q_2 \times 1)}$ normales multivariadas e independientes $\Rightarrow \underbrace{cov(\mathbb{X}_{(1)}, \mathbb{X}_{(2)})}_{\text{matriz } q_1 \times q_2} = 0$

⊛ Si $\mathbb{X} = \begin{pmatrix} \mathbb{X}_{(1)} \\ \hline \mathbb{X}_{(2)} \end{pmatrix} \sim N_{q_1+q_2}\left(\begin{pmatrix} \mu_{(1)} \\ \hline \mu_{(2)} \end{pmatrix}, \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right) \right)$

⊛ $\mathbb{X}_{(1)}$ y $\mathbb{X}_{(2)}$ indep $\Leftrightarrow \Sigma_{12} = 0$ (covarianza cero implica independencia sii las variables son normales)

7) Sea $\mathbb{X} \sim N_p(\mu, \Sigma)$, $|\Sigma| > 0$ entonces:

$$(\mathbb{X} - \mu)'(\Sigma^{-1})(\mathbb{X} - \mu) \sim \chi^2_p$$

la distancia de los valores a la media sigue distribución χ^2.

$$P\left(\left\{(\mathbb{X} - \mu)'(\Sigma^{-1})(\mathbb{X} - \mu) \leq \chi^2_p(\alpha)\right\}\right) = 1 - \alpha$$

Muestreo de la normal multivariada

Suponga que $\mathbb{X}_1, \dots, \mathbb{X}_n$ son una muestra aleatoria de una población normal multivariada con media μ y cov Σ.

Densidad conjunta de $\mathbb{X}_1, \dots, \mathbb{X}_n$

función de Verosimilitud dadas $\mathbb{X}_1, \dots, \mathbb{X}_n$

$$L(\mu, \Sigma) = \prod_{j=1}^{n} \left(\frac{1}{(2\pi)^{p/2} \cdot |\Sigma|^{1/2}} \cdot e^{-\frac{1}{2}\left[(\mathbb{X}_j - \mu)' \Sigma^{-1} (\mathbb{X}_j - \mu)\right]} \right)$$

Método de máxima Verosimilitud → Utilizar como estimadores de parámetros los valores que maximizan $L(\mu, \Sigma)$

los MLE μ y cov. Σ para $\mathbb{X}_1, \dots, \mathbb{X}_n$ muestras aleatorias normales con media son:

$$\hat{\mu} = \bar{\mathbb{X}}$$

$$\hat{\Sigma} = \frac{1}{n}\left(\sum_{j=1}^{n} (\mathbb{X}_j - \bar{\mathbb{X}})(\mathbb{X}_j - \bar{\mathbb{X}})'\right) = \frac{n-1}{n} S_n$$

$$L(\hat{\mu}, \hat{\Sigma}) = \frac{1}{(2\pi)^{p/2}} \cdot e^{-\frac{np}{2}} \times \frac{1}{\underbrace{|\hat{\Sigma}|}_{\left(\frac{n-1}{n} \cdot S_n\right)}^{n/2}}$$

Inferencia sobre Vector de medias

Sean $\mathbb{X}_1, \dots, \mathbb{X}_n$ una muestra aleatoria de población normal la distribución de $\bar{\mathbb{X}}$:

$$\boxed{\bar{\mathbb{X}} \sim N_p\left(\mu, \frac{1}{n}\Sigma\right)}$$

1) Si $p = 1$ $(n-1)S^2 = \sum_{j=1}^{n} (\mathbb{X}_j - \bar{\mathbb{X}})^2$ tiene distribución χ^2_{n-1}, $(n-1)S^2$ se distribuye como $\sigma^2 (\underbrace{Z_1}_{N(0,1)}^2 + \dots + Z_{n-1})^2 \approx \chi^2_{n-1}$ g.l.

2)

Distribución Wishart → la distribución de $(n-1)\overset{\text{matriz}}{S}$ es wishart con $n-1$ g.l.

$W_n(\cdot \mid \Sigma)$ es la distribución de $\sum_{i=1}^{n} \underbrace{Z_i Z_i'}_{N(0,\Sigma)}$

3) $\bar{X}$ y $\overset{\text{matriz}}{S}$ son independientes

Teorema del límite central

Sean $X_1, \dots X_n$ observaciones independientes de una población con media μ y covarianza Σ. $\sqrt{n}(\bar{X}-\mu)$ es aproximadamente $N(0, \Sigma)$

si $n \gg p$ ($n-p$ grande)

$$\bar{X} \approx N\left(\mu, \frac{1}{n}\Sigma\right)$$

usar S en lugar de Σ no es grave.

Inferencia sobre μ

$$\boxed{t} = \frac{(\bar{X} - \mu_0)}{\frac{S}{\sqrt{n}}}$$

t-student

$$S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

$$\bar{x} = \frac{1}{n} \sum x_i$$

Rechazamos H_0 si $t \in RR$ por ejemplo:
- $|t| > t_{\alpha/2}$
- $t > t_\alpha$
- $t < -t_\alpha$

rechazar para $|t|$ grande es equivalente a rechazar para t^2 grande

T^2 de Hotelling $\boxed{T^2} = n \cdot (\bar{X} - \mu_0)' (S^{-1}) (\bar{X} - \mu_0)$

Vectorialmente

matriz

$$\begin{pmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{p0} \end{pmatrix}$$

$$\bar{X} = \frac{1}{n} \sum_{j=1}^{n} X_j \qquad S = \frac{1}{n-1} \sum \overbrace{(X_j - \bar{X})(X_j - \bar{X})'}^{\text{vectores}}$$

1) T^2 se llama T^2 de Hotelling

2) $\frac{1}{n} S$ es la estimación de la covarianza de $\bar{X}$

3) Si H_0 es cierta $T^2 \sim \frac{(n-1)}{(n-p)} F_{p, n-p}$ → distribución F con $p, n-p$ g.l.

Aplicando:

Prueba de hipótesis $H_0: \mu = \mu_0$, $H_a: \mu \neq \mu_0$, nivel α

$$T^2 = n\underbrace{(\bar{X} - \mu_0)'}_{\substack{\text{Normal} \\ N_p(0,\Sigma)}}\underbrace{(S^{-1})}_{\substack{\text{Wishart} \\ (p,\, n-1)}}\underbrace{(\bar{X} - \mu_0)}_{\substack{\text{Normal} \\ N_p(0,\Sigma)}} \geq \underbrace{\frac{(n-1)p}{(n-p)} \cdot F_{p,n-p}(\alpha/2)}_{F_{\underbrace{p,n-p}_{g.l.}}}$$

4) T^2 es invariante bajo cambios de unidades de las medidas de X

$$Y = \underbrace{C_{p\times p}}_{cte}\, X_{p\times 1} + d_{p\times 1}, \quad T^2 = n(\bar{Y} - \mu_{Y_0})'(S_Y^{-1})(\bar{Y} - \mu_{Y_0})$$
$$= n(\bar{X} - \mu_0)(S^{-1})(\bar{X} - \mu_0)$$

Método de razón de verosimilitud (pruebas potentes)

$$\max_{\mu,\Sigma} L(\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\hat{\Sigma}|^{n/2}} \cdot e^{-\frac{np}{2}}$$

$$\hat{\Sigma} = \frac{1}{n} \sum (X_j - \bar{X})(X_j - \bar{X})' \qquad \hat{\mu} = \bar{X}$$

Asumiendo $H_0: \mu = \mu_0$

$$L(\mu_0, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{n/2}} \cdot \exp\left(-\frac{1}{2} \sum_{j=1}^{n} (X_j - \mu_0)'(\Sigma^{-1})(X_j - \mu_0)\right)$$

Dado μ_0 fijo variamos Σ para encontrar

$$\max_{\Sigma} L(\mu_0, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} \cdot |\hat{\Sigma}_0|^{n/2}} \cdot e^{-\frac{np}{2}}$$

$$\hat{\Sigma}_0 = \frac{1}{n} \cdot \sum_{j=1}^{n} (X_j - \mu_0)(X_j - \mu_0)'$$

$$\underbrace{\Lambda}_{\text{lambda de wilk}} = \frac{\max_{\Sigma} L(\mu_0, \Sigma)}{\max_{\mu,\Sigma} L(\mu, \Sigma)} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{n/2}$$

Si Λ pequeño entonces H_0 es poco probable y H_0 se rechaza

H_0 se rechaza si $\Lambda < \underbrace{C_\alpha}_{\alpha \text{ percentil de la distribución de } \Lambda}$