Guillermo Ribero

Hoja de fórmulas:

También ∃ el vector de variables X

$\mathbb{X}$ ← Matriz de variables (Columnas va y filas obs)

Media de la muestra: $\bar{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$ (va → j; Filas, Col)

Diagonal de la matriz S. ← **Varianza muestral:** $S_k^2 = S_{kk} = \frac{1}{n}\sum_{j=1}^{n}(X_{jk}-\bar{X}_k)^2$, $k=1,2,\dots,p$

(Su raíz cuadrada → desviación estandar)

$$\begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{p1} & \cdots & X_{pp} \end{bmatrix}$$

Covarianza muestral: $S_{ik} = \frac{1}{n}\sum_{j=1}^{n}(X_{ji}-\bar{X}_i)(X_{jk}-\bar{X}_k)$ $i,k = 1,2,\dots,p$

↳ $S_{ik} = S_{ki}$ pues S es simétrica.

Coeficiente de Correlación: de la muestra

$$r_{ki} = r_{i,k} = \frac{S_{ik}}{\sqrt{S_{ii}}\cdot\sqrt{S_{kk}}} = \frac{\sum_{j=1}^{n}(X_{ij}-\bar{X}_i)(X_{jk}-\bar{X}_k)}{\sqrt{\sum_{j=1}^{n}(X_{ji}-\bar{X}_i)^2}\sqrt{\sum_{j=1}^{n}(X_{jk}-\bar{X}_k)^2}}$$

→ Dejaría usar n o n-1. (Puede servir en alguna demostración)

↳ Lo que estandariza.

$= S_{12} = S_{21} = S_{ki}$ → Si $k=i$ es varianza muestral.

→ Correla entre 2 vars.
+ Ambas crecen o decrecen
− Van pa lados dif.
0 no se correla.

- $-1 \le r \le 1$

$$r_{11} = 1 = \frac{S_{ii}}{\sqrt{S_{ii}^2}} = 1.$$

- r es que tanto se relacionan i,k linealmente.
  - $r>0$ Tiende a que los valores i,k se comporten igual. ej: uno grande el otro también, uno pequeño el otro también. ← Algo como esto
  - $r=0$ i,k no se relacionan linealmente. ← Si no es 0, indica la dirección de la rela lineal.
  - $r<0$ Tiende a que un valor del par i,k sea mayor a su promedio dado que el otro es menor al promedio. (Poderoso pues sabiendo que uno es menor puedo intuir sobre el otro valor de la pareja i,k).
- El valor de $r_{i,k}$ se mantiene incluso si los valores de i pasan a ser $Y_{ji} = aX_{ji}+b$, y los de k $Y_{jk} = cX_{jk}+d$. Siempre y cuando a,c tengan el mismo signo.

r_{ik} es invariante si las medidas de i e k cambian de forma lineal:
$Y_{ji} = aX_{ji}+b$ y $Y_{jk} = cX_{jk}+d$

$S_{kk} = \frac{1}{n}\sum(X_{jk}-\bar{X}_k)^2$

Es simétrica ← **Matriz de Varianzas y Covarianzas:** $S = \begin{pmatrix} & & S_{ik} \\ & \ddots & \\ S_{ik} & & \end{pmatrix}$

Media muestral: $\bar{\mathbb{X}} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix}$

= "estandarizada"

Son simétricas → **Matriz de Correlaciones muestrales:** $R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix}$

Distancia euclidiana: $P=(x_1, x_2, \dots, x_p)$, origen $O=(0,0,\dots,0)$

$d(O,P)=\sqrt{x_1^2+x_2^2+\dots+x_p^2}$ y sea $Q=(y_1, y_2, \dots, y_p)$

$d(P,Q)=\sqrt{(x_1-y_1)^2+\dots+(x_p-y_p)^2}$, $d^2(O,P)=x_1^2+x_2^2+\dots+x_p^2$ → Puntos a una distancia constante cuadrada del origen

Distancia estadística:

Asumimos ind. ← media 0.

$$\frac{V_{ai}-0}{\sqrt{S_{ii}}}$$

$$x_1^* = \frac{x_1}{\sqrt{S_{11}}}, \quad x_2^* = \frac{x_2}{\sqrt{S_{22}}}$$

$P=(x_1, x_2)$ $Q=(y_1, y_2)$ → Cualesquiera puntos

$$d(O,P)=\sqrt{(x_1^*)^2+(x_2^*)^2}=\sqrt{\frac{x_1}{\sqrt{S_{11}}}+\frac{x_2}{\sqrt{S_{22}}}}$$

$$d(P,Q)=\sqrt{\frac{(x_1-y_1)^2}{S_{11}}+\frac{(x_2-y_2)^2}{S_{22}}}$$

Se generaliza hasta p variables. (Q punto fijo con coord. ind.)

$$d(P,Q)=\sqrt{\frac{(x_1-y_1)^2}{S_{11}}+\dots+\frac{(x_p-y_p)^2}{S_{pp}}}$$

Todos los ptos con dist. cuadrada:

$$C^2=\frac{x_1^2}{S_{11}}+\frac{x_2^2}{S_{22}}$$

Una elipse centrada en el origen

Nueva dist. estadística: $d(O,P)=\sqrt{\frac{\tilde{x}_1^2}{\tilde{S}_{11}}+\frac{\tilde{x}_2^2}{\tilde{S}_{22}}}$

→ Varianzas calculadas con $\tilde{x}_1$ y $\tilde{x}_2$.

$\tilde{x}_1 = x_1\cos(\theta)+x_2\operatorname{sen}(\theta)$ $\tilde{x}_2 = -x_1\operatorname{sen}(\theta)+x_2\cos(\theta)$

Asumimos correlación no nula

$$d=\sqrt{a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2}$$

→ Asume correla. no nula. aquí es 0.

Ai's funciones de θ.

$$d(P,Q)=\sqrt{a_{11}(x_1-y_1)^2+\dots+a_{pp}(x_p-y_p)^2+2a_{12}(x_1-y_1)(x_2-y_2)+\dots+2a_{p-1,p}(x_{p-1}-y_{p-1})\cdot(x_p-y_p)}$$

→ $x'Ax$

Distancia válida:

Distancia de P a Q y R un pto intermedio debe satisfacer:

- $d(P,Q)=d(Q,P)$
- $d(P,Q)>0$ si $P\neq Q$
- $d(P,Q)=0$ si $P=Q$
- $d(P,Q)\leq d(P,R)+d(R,Q)$ (Desigualdad triangular)

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

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

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

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \\ a_{p1} & & a_{pp} \end{pmatrix}$$

Repaso lineal:

Longitud vector:

$$L_x = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{x \cdot x^T}$$

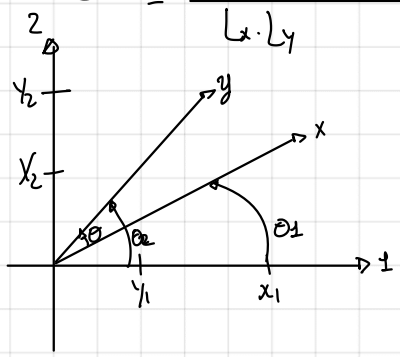
$$L_{cx} = \sqrt{c^2 x_1^2 + \dots + c^2 x_2^2} = \underset{c \in \mathbb{R}}{|c|} \sqrt{x_1^2 + \dots + x_n^2} = |c| L_x$$

Si $c = L_x^{-1}$ cX es el vector unitario en dirección de X.

Ángulo entre 2 vectores:

X, Y Son Perpendiculares Si $X' \cdot Y = 0$.

$$\cos\theta = \frac{X \cdot Y}{L_x \cdot L_y} = \frac{X \cdot Y}{\sqrt{(X'X)(Y' \cdot Y)}}$$



$$\cos\theta_1 = \frac{X_1}{L_x}, \quad \cos\theta_2 = \frac{Y_1}{L_y}$$

$$\sin\theta_1 = \frac{X_2}{L_x}, \quad \sin\theta_2 = \frac{Y_2}{L_y}$$

$$\cos\theta = \cos(\theta_2 - \theta_1) = \cos\theta_2 \cos\theta_1 + \sin\theta_2 \sin\theta_1$$

Independencia lineal: $X_1 \dots X_k$ son li

X, Y Son LD Si $\exists$ $C_1, C_2 \neq 0$ tq $C_1 X + C_2 Y = 0$.

El uno Combo lineal del otro.

$$C_1 X_1 + \dots C_n X_n = 0 \rightarrow \underline{\text{Solo Cumple con } C_1 = \dots = C_n = 0}.$$

Proyecciones:

Sean X, Y vectores. La proyección ortogonal de X sobre Y se calcula como:

L_y^{-1} tiene long $= 1$.

$$P_y(X) = \left(\frac{X' \cdot Y}{Y' \cdot Y}\right) Y = \frac{X' \cdot Y}{L_y} \cdot Y$$

Longitud de la proyección $L_{P_y(x)} = |\cos(\theta)| \cdot L_y$

$AB \neq BA$

- Matriz cuadrada: mismo num filas y columnas.
- A simétrica si $A = A^T$ (Generalmente $AB \neq BA$)
  ó $a_{ij} = a_{ji}$ $\forall i, j$.
- I_k identidad $k \times k$. $I_k \cdot A = A I_k = A$

A es invertible si todos los vectores columna son LI.

- Una matriz cuadrada es invertible si $\exists$ B tq: $AB = BA = I_k$ $(B = A^{-1})$

Matrices ortogonales (Cuadrada)

$QQ' = Q'Q = I$

$Q' = Q^{-1}$

Vectores y valores propios:

Sea A una matriz cuadrada, decimos que λ es un valor propio de A con vector propio.

$x \neq 0$ Si $\underbrace{Ax}_{\text{col}} = \lambda x$

Sea A_{kk} simétrica, ent A tiene k pares de valores y vectores propios.

$\lambda_1 e_1, \dots, \lambda_2 e_n$

$e_1, \dots, e_n \to$ Ortonormales. (entre ellos)

Vecs propios

Sea A una matriz cuadrada simétrica

Descomp. Espectral $\Rightarrow A = \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \dots + \lambda_k \cdot e_k e_k^T$

Matrices no neg ≥ 0 Forma Cuadrática

$x'Ax > 0$ Def Positiva

Forma Cuadrática:

Sea A una matriz cuadrada simétrica $k \times k$, sea x un vector k, $x_1, x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$

La forma cuadrática asociada a A es $x'Ax = a_{11}x_1^2 + \dots + a_{ij}x_i \cdot x_j + \dots a_{k-1,k} x_k \cdot x_{k-1}$ $\hookrightarrow c^2$

- Si $\forall x \in \mathbb{R}^n$ $0 \leq x'Ax$ ent A se dice no negativa.
  Si A es def no negativa.

  $x'Ax = 0$ Si $x = 0$.

- A definida positiva

  1) A def positiva si: todos los valores propios son positivos.

  2) A def no neg si: Todos los valores propios son no neg.

$P = [e_1, e_2, \dots, e_k]$

$PP' = PP' = I$

$A^{-1} = P \Lambda^{-1} P'$

$= \sum_{i=1}^{n} \frac{1}{\lambda_i} \cdot e_i \cdot e_i'$

Sea $A = \sum_{i=1}^{k} \lambda_i e_i \cdot e_i'$

$= P \Lambda P'$

Matriz diagonal.

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix} \quad \lambda_i > 0$$

Raíz Cuadrada de una matriz:

$$A^{1/2} = \sum_{i=1}^{k} \sqrt{\lambda_i}\, e_i e_i' = P \Lambda^{1/2} P'$$

$\sqrt{\lambda_i}$ i-ésimo

1. $(A^{1/2})^T = A^{1/2}$ (Simétrica).
2. $A^{1/2} \cdot A^{1/2} = A$
3. $(A^{1/2})^{-1} = \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}} e_i \cdot e_i' = P \Lambda^{-1/2} P'$
4. $A^{1/2} \cdot A^{-1/2} = A^{-1/2} \cdot A^{1/2} = I = I$ y $A^{-1/2} \cdot A^{-1/2} = A^{-1}$

$A^{-1/2} = (A^{1/2})^{-1}$

Vectores aleatorios:

Vec/Mat elem son va, $E(\text{vec}_a/M_a) = E(\%\text{ de sus elems})$

Elem vec/Mat

Elem matriz

$$E(X_{ij}) = \begin{cases} \int_{-\infty}^{\infty} x_{ij} f_{ij}(x_{ij})\, dx_{ij}, & X_{ij} \text{ cont con densidad } f_{ij}(x_{ij}) \\ \sum P_{ij}(x_{ij}) \cdot x_{ij}, & X_{ij} \text{ disc con PMF } P_{ij}(x_{ij}) \end{cases}$$

Valor esperado Sumas y Productos matrices:

Sean $\mathbb{X}$, $\mathbb{Y}$ matrices (vectores) aleatoreas del mismo tamaño.
y A,B matrices (vecs) constantes (%elem):

1) $E(\mathbb{X}+\mathbb{Y}) = E(\mathbb{X}) + E(\mathbb{Y})$

2) $E(A\mathbb{X}B) = AE(\mathbb{X})B$

Vecs media y Ma de Cov:

Sean μ_i y σ_i^2 la media y varianza de X_i (marginales)

$$\mu_i = E(X_i) = \begin{cases} \int_{-\infty}^{\infty} x_i \cdot f_i(x_i)\, dx_i & \text{Pdf} \\ \sum x_i \cdot P_i(x_i) \end{cases} \qquad \sigma_i^2 = E((X_i - \mu_i)^2) = \begin{cases} \int_{-\infty}^{\infty} (x_i-\mu_i)^2 \cdot f_i(x_i)\, dx_i \\ \sum (x_i-\mu_i)^2 \cdot P_i(x_i) \end{cases}$$

$\sigma_i^2 = \sigma_{ii}$ $\quad$ $i=k$ es varianza.

Covarianza entre X_i, X_k $\quad i,k = 1,\dots,p$

$$\sigma_{ik} = Cov(X_i, X_k) = E((X_i-\mu_i)(X_k-\mu_k))$$

$$= \begin{cases} \int_{-\infty}^{\infty} (x_i-\mu_i)(x_k-\mu_k) \cdot f_{ik}(x_i, x_k)\, dx_i\, dx_k \\ \sum (x_i-\mu_i)(x_k-\mu_k) \cdot P_{ik}(x_i, x_k) \end{cases}$$

$\rightarrow$ PMF Conjunta

Obs: Sea X un vec. Su dist es la conjunta de las entradas:

$$f_{X_1 \dots X_p}(x_1 \dots x_p) = f_X(x) \rightarrow \text{Casi Siempre Pdf multivariada}$$

Independencia estadística P-variables:

Si son ind $\Downarrow$ Cov = 0.

- X_i, X_k son est. ind si: $F_{ik}(x_i, x_k) = F_i(x_i) \cdot F_k(x_k)$
  $\hookrightarrow f_{ik}(x_i, x_k) = f_i(x_i) \cdot f_k(x_k)$
- Gen $\rightarrow$ $X_1 \dots X_p$ son estind. Si $f_{X_1 \dots X_p}(x_1, \dots, x_p) = \prod_{i=1}^{p} f_i(x_i)$, $\forall x_1, \dots x_p$

$\mu_x = E(X)$ (Vec Media Poblacional)

Notación: Sea X vecc$_{p\times 1}$

Denotamos $\mu_X = \mu = E(X)$ (Vec medias Poblacional)

Matriz de varianzas y covarianzas. $\leftarrow$ $\underset{p\times p}{\Sigma} = E((X-\mu)(X-\mu)^T) \rightarrow$ Elementos son:

$\hookrightarrow$ Simétrica

- $\sigma_{ii} = \sigma_i^2$ diag. Principal (varianzas)
- $\sigma_{ik} = \sigma_{ki}$ Covarianzas $\frac{p(p-1)}{2}$

$$E(\underset{p\times 1}{X}) = \mu_X = \mu = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} \rightarrow$$ Vector de medias Poblacional.

$$\Sigma_{p\times p} = E\left(\begin{pmatrix} X_1-\mu_1 \\ \vdots \\ X_p-\mu_p \end{pmatrix} \cdot \left(X_1-\mu_1, \dots, X_p-\mu_p\right)\right)$$

$$Cov(X) \leftarrow \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix} \rightarrow$$ Matriz de varianza-cov. Poblacional

Mide asociación lineal entre vars. $\leftarrow$ **Def:** Matriz de Correlación Poblacional (Simétrica):

$$\underset{p\times p}{\rho} = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1p} \\ \vdots & \ddots & \vdots \\ \rho_{1p} & \cdots & \rho_{pp} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & \rho_{1p} \\ \vdots & \ddots & \vdots \\ \rho_{p1} & & 1 \end{pmatrix}, \quad \rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}}$$

$\sigma_{ik} \leftarrow$ Cov

Matriz desviación estandar:

0's Pues no nos interesa la Covarianza $\leftarrow$

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

$$\boxed{\begin{aligned} &1)\ V^{1/2}\rho V^{1/2} = \Sigma_{p\times p} \\ &2)\ \rho = (V^{1/2})^{-1}_{p\times p}\,\Sigma_{p\times p}\,(V^{1/2})^{-1}_{p\times p} \end{aligned}}$$

Los autovals y vecs se encuentran así:

1. $\det(A-\lambda I)$ da ecua característica
2. $\det(A-\lambda I) = 0 \rightarrow$ Polinomio Característico.
   $\hookrightarrow$ Las raíces son los vals propios.
3. Para el vector propio: $(A-\lambda I)v = 0$ (Se resuelve usando c/λ_i)

$\det(A_{n\times n})$ Σ elem diag. Principal

$\underbrace{(-1)^{i+j}}_{\text{Cofactor.}} \cdot \underbrace{\det(\text{menor del elem})}_{\text{Se hace recursivo.}}$

$i+j \rightarrow$ Pose elem

variables a despejar.

autoval.

Particiones de vectores y matrices:

$$X_{p\times 1} = \begin{pmatrix} X_1 \\ \vdots \\ X_q \\ \hline X_{q+1} \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \underset{q\times 1}{X_1} \\ \hline \underset{(p-q)\times 1}{X_2} \end{pmatrix}, \quad \mu = E(X) = \begin{pmatrix} E(X_1) \\ \hline E(X_2) \end{pmatrix} = \begin{pmatrix} \mu^{(1)} \\ \hline \mu^{(2)} \end{pmatrix}$$

Matriz de Covarianzas entre las Componentes de $X^{(1)}$ y $X^{(2)}$.

(Nota: no es necesariamente simétrica ni cuadrada)

$$\Sigma_{(1)(2)} = E\left((X^{(1)} - \mu^{(1)})(X^{(2)} - \mu^{(2)})^T \right)$$

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

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

$$\Sigma_{21} = \Sigma_{12}^T$$

$$\Sigma_{p\times p} = \begin{array}{c|c} & \begin{array}{cc} 1\cdots q & q\cdots p \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ q \\ \hline q \\ \vdots \\ p \end{array} & \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right) \end{array} = \left(\begin{array}{c|c} \begin{matrix} \sigma_{11} & \cdots & \sigma_{1q} \\ \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} \end{matrix} & \Sigma_{12} \\ \hline \Sigma_{21} & \begin{matrix} \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,p} \\ \vdots & \ddots & \vdots \\ \sigma_{p,q+1} & \cdots & \sigma_{pp} \end{matrix} \end{array}\right)$$

La covarianza de $X^{(1)}$ es Σ_{11} y la de $X^{(2)}$ es Σ_{22}.

Combinaciones lineales de vecas:

Recordeis: $Y = aX + b$, $a, b \in \mathbb{R}$ X vca

$E(Y) = aE(X) + b$, $Var(Y) = a^2 Var(X)$

$a, b, c \in \mathbb{R}$, X, Y, Z vca's

$Cov(aX, bY) = ab\,Cov(X,Y)$ / $Cov(X, Y+c) = Cov(X,Y)$ / $Cov(X, Y+Z) = Cov(X,Y) + Cov(X,Z)$

Sea $c = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}$, $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ X_1, X_2 vca

$$E(c'X) = c'E(X) = c'\mu_X \;/\; Var(c'X) = c'\Sigma c = (a\ b) \begin{pmatrix} Var(X_1) & Cov(X_1X_2) \\ Cov(X_1X_2) & Var(X_2) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

Generalizando a p-variables tenemos:

$$c = \begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix}_{p\times 1} \quad X = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}_{p\times 1}$$

$$c'X = c_1X_1 + c_2X_2 + \dots + c_pX_p$$

$$E(c'X) = c' \cdot E(X) = c'\mu_X$$

$$Var(c'X) = \underset{1\times p}{c'} \underset{p\times p}{\Sigma} \underset{p\times 1}{c}$$

Generalización a combolin de P-vars:

$$Z_1 = c_{11}X_1 + \dots + c_{1p}X_p$$
$$\vdots$$
$$Z_q = c_{q1}X_1 + \dots + c_{qp}X_p$$

ent $\underset{q\times 1}{Z} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_q \end{pmatrix}$, $\underset{p\times 1}{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$, $\underset{q\times p}{C} = \begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{q1} & \cdots & c_{qp} \end{pmatrix}$

Luego $$\underset{q\times 1}{Z} = \underset{q\times p}{C}\,\underset{p\times 1}{X} = \underset{q\times p}{\begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{q1} & \cdots & c_{qp} \end{pmatrix}} \underset{p\times 1}{\begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}}$$

$$\underset{q\times 1}{\mu_Z} = E(Z) = E(CX) = CE(X) = \underset{q\times p}{C}\,\underset{p\times 1}{\mu_X}$$

$$\underset{q\times q}{\Sigma_Z} = \underset{q\times p}{C}\,\underset{p\times p}{\Sigma_X}\,\underset{p\times q}{C'}$$

Muestreo y Geometría:

Asumimos ma:

$$\underset{n\times p}{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} x_1' \\ \vdots \\ x_n \end{pmatrix}$$

ma ← X

n mediciones (obs) en p-variables.

obs → n, vars → p

x_1' → Vectores fila $(1\times p)$ características o n vecs en $\mathbb{R}^p$ para observación en P variables.

x_n → última obs

$\bar{X}$ es como el centro de gravedad (como alrededor de que estén las obs.) de la "nube" de puntos.

También podemos ver X como:

o P-vectores en el espacio n-dimensional

$$\underset{n\times p}{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} = (Y_1 | \cdots | Y_p)$$

Y_i es un vector columna de tamaño $n\times 1$, 1 variable y n observaciones $i = 1,\dots,p$

2 formas de graficar o agrupar.

Vector 1_n: $1_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \Big\} n$

Obs: El vec. 1 forma un ángulo igual con todos los ejes.

El vector $\left(\frac{1}{\sqrt{n}}\right)1$ es el vector unitario en dirección de 1_n.

Si $y_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$ La proyección ortogonal de Y_i sobre $\frac{1}{\sqrt{n}}1$ es:

$$\left[y_i'\left(\frac{1}{\sqrt{n}}\right)1\right]\left(\frac{1}{\sqrt{n}}1\right) = \left(\frac{x_{1i}+\cdots+x_{ni}}{n}\right)1$$

$$= \bar{x}_i 1.$$

Vector desviación:

$d_i = Y_i - \bar{x}_i 1$ es perpendicular a $\bar{x}1$

$$d_i = Y_i - \bar{x}_i 1 = \begin{pmatrix} x_{1i}-\bar{x}_i \\ \vdots \\ x_{ni}-\bar{x}_i \end{pmatrix}$$

→ Cada elemento de d_i son las desviaciones de la media. $(x_{1i}-\bar{x}_i)$

Propiedades Vec. desviación:

Norma

$\text{Length}(d_i) = \sqrt{d_i' \cdot d_i}$ $\qquad L(d_i d_k) = \sqrt{(d_i d_k)(d_i d_k)}$

$L^2 d_i = d_i' d_i = \sum_{j=1}^{n} (x_{ji} - \bar{x}_i)^2$ $\qquad L^2(d_i - d_k) = (d_i d_k)(d_i d_k)'$

$= \sum (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_i)$

$\frac{L^2 d_i}{n} = S_{ii}$ (Varianza) $\qquad L^2 d_i = n S_{ii}$ $\qquad \frac{L^2(d_i d_k)}{n} \Rightarrow$ Correla. (Creo)

Valores grandes d_i representan más variabilidad en la i-ésima var.

$d_i \cdot d_k = L d_i \cdot L d_k \cdot \cos(\theta_{ik})$ (Producto Vect) → Preguntar Yepes

$\sum_{j=1}^{n} (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k) = n \cdot Cov(X_i, X_k)$?

$$\cos(\theta_{ik}) = \frac{S_{ik}}{\sqrt{S_{ii}}\sqrt{S_{kk}}} = r_{ik}$$ (Ángulo entre 2 vecs. desv)

Cos del ángulo entre d_i y d_k es el coef de Correlación muestral r_{ik}

Ma y Val. esperado de la media muestral y de la matriz de covarianza muestral:

Sup. que no hemos recogido datos

$$X_{n \times p} = \begin{pmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$ → Agrupamos Por obs

Los vecs $X_1, \ldots, X_n$ son aleatorios, ind e identicamente dist. Si la densidad conjunta (dF) es: $f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n) = f_X(x)$

Observación: Son ind pero las variables no necesariamente lo son.

donde $f(X_j) = f(X_{j1}) \cdot \ldots \cdot f(X_{jp})$ (Tmb ind en % obs)

$f(X_j) = f(X_{j1}, \ldots, X_{jp})$

Obs: • Las medidas de las p-variables en una misma obs, no son necesariamente ind.

• Las medidas asociadas a obs diferentes deben ser ind.

Sup: $\overset{ma}{\mathbb{X}}_{n\times p} = \begin{pmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{pmatrix}$ y $\mathbb{Y}_k = \begin{pmatrix} X_{1k} \\ \vdots \\ X_{nk} \end{pmatrix} \rightarrow$ Observaciones de la variable k en la ma $\mathbb{X}$.

La distribución de $\mathbb{Y}_k$ depende de:

$$f(Y|k) = f(X_{1k}, \ldots, X_{nk}) = f_k(x_{1k}) \cdot \ldots \cdot f_k(x_{nk})$$

Distribuciones Muestrales de $\bar{X}$ y S_n:

Sean $X_1 \ldots X_n$ vecas (ma multivariada $\Rightarrow$ mav) tomadas de una Población con medio Poblacional μ, $(E(X_i) = \mu)$ y Covarianza Poblacional

$\hookrightarrow \Sigma \; (Var(X_i) = \Sigma)$ $i = 1, \ldots, n$

ent:

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

$Cov(\bar{X}) = \frac{1}{n}\Sigma$

2) $E(S_n) = \frac{n-1}{n}\Sigma \rightarrow$ Dem de Tarea

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

Estimador insesgado de Σ pues S_n es la misma matriz pero de la muestra.

Sesgo de S_n:

$$B(S_n) = E(S_n) - \Sigma$$
$$= \frac{n-1}{n}\Sigma - \Sigma$$
$$= \Sigma\left(\frac{n-1}{n} - 1\right) \frac{\cancel{n} - 1 - \cancel{n}}{n}$$
$$= -\frac{\Sigma}{n} \rightarrow \text{Es sesgado.}$$

Dem: $\bar{X} = \left(\frac{1}{n}\right)(X_1 + \ldots + X_n)$

$$E(\bar{X}) = \frac{1}{n}\underbrace{E(X_1)}_{\mu} + \frac{1}{n}\underbrace{E(X_2)}_{\mu} + \ldots + \frac{1}{n}\underbrace{E(X_n)}_{\mu}$$
$$= \mu$$

$E((\bar{X} - \mu)(\bar{X} - \mu)')$

Varianza muestral gen. y Normal multivariada:

Matriz de var y Cov de la muestra. ← $S_n = \frac{1}{n} \sum_j (X_{ji} - \bar{X}_j)^2$

Matriz de var y Cov de la ma. insesgada ← $S = \frac{1}{n-1} \sum_j (X_{ji} - \bar{X}_j)^2$

Varianza muestral generalizada $= |S|$ ← det.

- 0 S_{ii} at least 1 vector desviación yace en el hiperplano formado por todas las combo lineales de los vectores restantes.
  Ed es 0 si al menos una columna de la matriz S es L.D. de otra.

$\underbrace{n \leq p}_{\text{Un número mayor o igual de variables a las obs.}}$ $X_{n \times p} = \begin{array}{c} n \left\{ \begin{bmatrix} \quad \end{bmatrix} \right. \\ \underbrace{\qquad\qquad}_{p} \end{array}$

Var. gen. determinada por $|R|$: → Notación

- La var muestral gen. es afectada por la variabilidad de las mediciones en una sola variable.
- Ej: Si S_{ii} es muy grande (o pequeño) geométricamente el vector de desviación correspondiente

  $d_i = (Y_i - \bar{X}_i 1)$ Será muy largo (o muy corto) y afectará el volumen

- Para solucionar lo anterior escalamos todos los vectores (d_i) de tal forma que tengan la misma longitud.
  ↳ Reemplazar la obs d_i por $\frac{(X_{ji} - \bar{X}_i)}{\sqrt{S_{ii}}}$ (Estandarizar)

Ent. la matriz de varianzas y covarianzas es R → La matriz de correlación (Muestral, est)

$|R|$ = Var. muestral. gen. de las variables estandarizadas

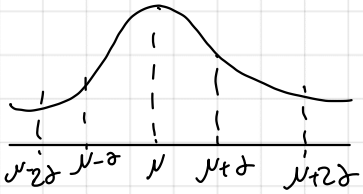
Varianza total de la muestra $= S_{11} + S_{22} + \ldots + S_{pp}$
(Suma de varianzas)

Normal multivariada:

Recordemos: Pdf va normal con media μ y var σ^2 es:

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

Ley empírica:



Potencia de una matriz:

$$A^n = \begin{pmatrix} a^{n-1} & b^{n-1} \\ c^{n-1} & d^{n-1} \end{pmatrix}$$

Denotada $N(\mu, \sigma^2)$, X normal con media μ y var σ^2

$X \sim N(\mu, \sigma^2)$

$$\frac{(X-\mu)^2}{\sigma^2} \rightarrow \text{Dist. estadística entre } X \text{ y } \mu.$$

$$= (X-\mu)(\sigma^2)^{-1}(X-\mu)$$

Generalizando $\vec{X}$ (→ una fila de $\mathbb{X}$) vector de obs de $\mathbb{X}$. (1 obs Pvars) (← Matriz)

dist. estadística de $\mathbb{X}$ a μ

Da un número. ← $\underset{1\times p}{(\mathbb{X}-\mu)'} \underset{p\times p}{\Sigma^{-1}} \underset{p\times 1}{(\mathbb{X}-\mu)}$ $\quad \mathbb{X} \rightarrow$ Vector $p\times 1$

Ahora generalizando para obtener PDF de la normal multivariada:

$$f(\mathbb{X}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \cdot \exp\left(-(\mathbb{X}-\mu)\Sigma^{-1}(\mathbb{X}-\mu)/2\right)$$

Vector de observaciones ← $\mathbb{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ de un vector aleatorio $\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$ → Va. → Agrupar por va. las obs. de c/u.

Notación: Normal multivariada. (p-variada)

$$\mathbb{X} \sim \underset{\text{vars}}{N_p}(\underset{p\times 1}{\mu}, \underset{p\times p}{\Sigma})$$

(→ vector, → Matriz)

Obs: Si $p=2$ y $\rho_{12}=0$ → ($Cov(X_1, X_2)=0$ ⇓ Ind X_1, X_2) Como hay normalidad si la pueda implicar

Ecl Para N_p tenemos: Ind sii $Cov=0$

Ma def Positiva > 0
- Autovals (All) > 0
- $z' M z > 0$

Teo: Si Σ es def Positiva (Σ^{-1} existe) ent 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 $\frac{1}{\lambda}$. Ademas Σ^{-1} es def positiva.

Obs:

1) $\underbrace{(X-\mu)'\Sigma^{-1}(X-\mu)}_{\text{Suma de normales estandar al cuadrado estandarizadas}} \leq \chi^2_p(\alpha)$ con Proba $1-\alpha$. → Chi cuadrada

Est. $\Sigma^{-1} = \frac{1}{\sigma_{ii}}$

$\frac{X-\mu}{\sqrt{\sigma^2}}$

Prop 1: Si X es un vecc normal multivariado

$X \sim N_p(\mu, \Sigma)$ ent

Si $a \in \mathbb{R}^p$ (vec const.), $a'X = a_1X_1 + a_2X_2 + \dots + a_pX_p$

$\sim N_1(a'\mu, a'\Sigma a)$

Prop 2: Si $a'X \sim N_1(a'\mu, a'\Sigma a)$, $\forall a \in \mathbb{R}$ (Número; univariada)

ent $X \sim N_p(\mu, \Sigma)$

Recordemos que $a = (1, 0, 0, 0, \dots, 0_p)$. El 1 cambia pues se usa para seleccionar 1 variable.

Prop 3: $X \sim N_p(\mu, \Sigma)$

Sea $A_{q \times p} \in \mathbb{R}^{q \times p}$ (Matriz Const)

$A_{q \times p} X_{p \times 1} \sim N_q(A\mu, A\Sigma A')$ ← q variada.

Prop 4: $X \sim N_p(\mu, \Sigma)$ $d \in \mathbb{R}^p$ (vector desviación)

$X + d \sim N_p(\mu + d, \Sigma)$ → Var. no Cambia (Pues $Var(Const) = 0$)

Prop 5: Todas las particiones de vecs normales resultan en vecs normales:

$X \sim N_p(\mu, \Sigma)$ $\quad X = \begin{pmatrix} X_{(1)} \\ \hline X_{(2)} \end{pmatrix} \begin{matrix} \} q \\ \} p-q \end{matrix}$ $\quad \mu = \begin{pmatrix} \mu_{(1)} \\ \hline \mu_{(2)} \end{pmatrix}$

$\Sigma = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right)$

$X_{(1)} \sim N_q(\mu_{(1)}, \Sigma_{11})$

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

Prop 6: 1) Si $X_{(1)}$ ($q_1 \times 1$), $X_{(2)}$ ($q_2 \times 1$) norm. multi. ind ent $\underbrace{Cov(X_{(1)}, X_{(2)})}_{\text{Matriz } q_1 \times q_2} = 0$ → Matriz 0 → Pues Cov de c/par de elementos de (1) y (2) es 0.

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

$X_{(1)}$ y $X_{(2)}$ son ind sii $\underbrace{\Sigma_{12 \text{ or } 21}}_{(Cov)^s} = 0$ (lo Permite el que sean normales)

3) $X_{(1)}$ ($q_1 \times 1$), $X_{(2)}$ ($q_2 \times 2$) normales ind Sii $X = \left(\begin{array}{c} X_{(1)} \\ \hline X_{(2)} \end{array}\right) \sim N_{q_1+q_2}\left(\begin{array}{c} \mu_{(1)} \\ \hline \mu_{(2)} \end{array}, \left(\begin{array}{c|c} \Sigma_{11} & 0 \\ \hline 0 & \Sigma_{22} \end{array}\right)\right)$

Prop 7): Sea $X \sim N_p(\mu, \Sigma)$, $|\Sigma| > 0$ → det

Dist estadística de X a μ.

Ent $(X-\mu)'(\Sigma^{-1})(X-\mu) \sim \chi^2_p$ ← chi Squared.

La distancia de los valores a la media tiene dist. $\chi^2_{p\,el}$ ← Chisquared

Ed la probabilidad de que tenga esa distribución. ← La probabilidad del evento: $\{(X-\mu)'\Sigma^{-1}(X-\mu) \leq \chi^2_p(\alpha)\}$ es $1-\alpha$

Muestreo normal multivariada:

Sup que $X_1 \dots X_n$ son una ma de una Población normal multivariada con media μ y cov Σ

Pdf ← Densidad conjunta de $X_1 \dots X_n$

Son ind. ← $$\prod_{j=1}^{n}\left(\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \cdot \exp\left(-\underset{1\times p}{(X-\mu)'} \cdot \underset{p\times p}{\Sigma^{-1}} \cdot \underset{p\times 1}{(X-\mu)}/2\right)\right)$$

Función de verosimilitud: (función de μ y Σ)

dadas las obs $x_1, \dots, x_n$: $L(\mu, \Sigma)$

Mayor Valor → Mejor estimador. → mejor explican datos.

Método de máximo-verosimilitud:

Utilizar como estimaciones de parámetros poblacionales desconocidos los valores que maximizan $L(\mu, \Sigma)$

Teo: Sean $X_1, \dots, X_n$ una ma de una pobla normal con media μ y Cov Σ, ent los MLE (est máx vero) son:

$$\hat{\mu} = \bar{X}, \quad \hat{\Sigma} = \frac{1}{n}\left(\sum_{j=1}^{n}\underset{p\times 1}{(X_j - \bar{X})}\underset{1\times p}{(X_j - \bar{X})'}\right)$$

→ X_j obs j tiene p vars. Pues agrupo por obs.

$$= S_n = \frac{n-1}{n} \cdot S$$

Obs: 1) $\hat{\mu}$, $\hat{\Sigma}$ son vecas. (¿No es $\hat{\Sigma}$ una matriz?) antes de tomar datos.

2) $$L(\hat{\mu}, \hat{\Sigma}) = \underbrace{\frac{1}{(2\pi)^{p/2}} \cdot e^{-np/2}}_{\text{Const.}} \cdot \frac{1}{|\hat{\Sigma}|^{n/2}}$$

máx vero.

$\hat{\Sigma} = \frac{n-1}{n} \cdot S$ ← Parte constante.

Var.. → Sale tras operar Π.

$$= \text{Constante} \cdot |S|^{-n/2}$$

La varianza generalizada determina el máximo de la función de verosimilitud.

15/02/2022

Estimador de max-verosimilitud:

$\mu = \bar{X}$ $\quad \Sigma = S_n$ (est.) → Para matriz varianzas y covarianzas.

Distribución muestral $\bar{X}, S$

Si $X_1, \dots, X_n$ es una m.a. de Pobl. normal con media μ y cov Σ la distribución de $\bar{X}$ y de S se puede determinar completamente

1) Si $p=1$ $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

2) Caso general
$\bar{X} \sim N_p(\mu, \frac{1}{n}\Sigma)$

Si $p=1$
$(n-1)S^2 = \sum_{j=1}^{n} (X_j - \bar{X})^2$ $\quad$ $(n-1)S^2$ se distribuye como $\sigma^2(Z_1^2 + \dots + Z_{n-1}^2)$ → $N(0,1)$

Distribución $\chi^2_{\underbrace{n-1}_{gl}}$ → Suma de normales est^2

En el caso general, la distr. de $(n-1)S_{p\times p}$ (Matriz) es la dist. Wishart con $n-1$ gl → gen de χ^2 a matrices → Buscar que características tiene

$W_m(\cdot|\Sigma)$ es la distr de $\sum_{i=1}^{m} Z_i Z_i'$ → $N(0,\Sigma)$ (Normal estandar Sandra)

3) $\bar{X}$ y S (→ matriz) son ind

Propiedades asíntoticas: (LNG... etc)

1). $\bar{X}$ (← vector) converge en proba a μ. ($\bar{X}$ est consistente de μ) → Poco o nulo sesgo

Pues n grande ← 2) S y S_n convergen en probabilidad a Σ
(S, S_n son est consistentes de Σ)
S → $\frac{1}{n-1}$ insesgado, S_n → sesgado

Teo Lím Central:

Sean $X_1, \dots, X_n$ obs independientes de una población con media μ y cov Σ.

$\sqrt{n}(\bar{X} - \mu)$ es aprox $N(0, \Sigma)$ (valor esp, varianza)

Si $n >> p$ ($n-p$ es grande)

$$\bar{X} \approx N(\mu, \tfrac{1}{n}\Sigma)$$

Sin importar la distribución de los X_i

Obs: ∘ Usar S en lugar de Σ no afectará los resultados seriamente. (Pues es insesgado).

∘ $n-p$ Debe ser grande

Inferencia Sobre μ (Medias):

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$$ → Media de X var.

Vamos a usar teoría aprendida previamente y la adaptaremos a p-variables.

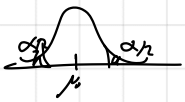
Ej: $\mu = \begin{pmatrix} \mu_1 \to 3 \\ \mu_2 \to 2 \\ \mu_3 \to 1 \end{pmatrix}$ → Podemos hacer muchas o una conjunta pues tiene más potencia (Proba de rechazar la nula correctamente).

Recordemos

P.H. respecto a una media poblacional μ_0

$H_0: \mu = \mu_0 \qquad H_a: \mu \neq \mu_0$

Tip: rechazar H_0 when true.

nivel α  y RR

Si $X_1 \dots X_n$ es una ma de una pobla. normal ent el estadístico de prueba es.

$$t = \frac{(\bar{X} - \mu_0)}{\frac{S}{\sqrt{n}}} \qquad S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})$$

$$\bar{X} = \frac{1}{n}\sum(X_i)$$

Si T muy pos o muy neg rechazamos pues la dist es muy grande.

← Si H_0 true, t tiene una distr. t-Student con $n-1$ gl

Rechazamos H_0 si $t \in RR$ i.e. $|t| > t_{\alpha/2}$ ← 2 colas

$t > t_\alpha$ ← Cola Sup

$t < t_\alpha$ ← Cola inf.

Si h_0 no se rechaza μ_0 es un valor posible de μ.

Rechazamos para $|t|$ grande y eso es equivalente a rechazar si t^2 es grande

$$t^2 = n(\bar{X} - \mu_0)(S^2)^{-1}(\bar{X} - \mu_0) \quad \text{¿(Dist estadística cuadrada)?}$$

Ej si la dist entre $\bar{X}$ y μ_0 es grande, rechazamos H_0.

Dada una muestra, se rechaza H_0 si:

$$n(\bar{X} - \mu_0)(S^2)^{-1}(\bar{X} - \mu_0) > t^2_{n-1}(\alpha/2)$$ usando $\alpha/2$

Recordar que no rechazar H_0 a un nivel α es equivalente a que el valor de μ_0 esté en un Int. Conf. con confianza $1-\alpha$.

Valores μ_0 en el I. C. son aquellos para los cuales no se rechaza $\underline{H_0: \mu = \mu_0}$.

Los límites de los int. de confianza son v.a. (Pues dependen de la muestra).

Probabilidad de que el I.C. contenga a μ es $1-\alpha$, luego si tomo muchas muestras $1-\alpha$ (100%) tendrán a μ.

Generalizando a p variables:

$$T^2 = (\bar{X} - \mu_0)' \left(\frac{1}{n} \overset{\text{Matriz}}{S}\right)^{-1} (\bar{X} - \mu_0)$$

$$= n(\bar{X} - \mu_0)' (S^{-1}) (\bar{X} - \mu_0)$$

donde $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ $\quad S = \frac{1}{n-1}\sum (X_j - \bar{X})(X_j - \bar{X})'$

$$\mu_0 = \begin{pmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{p0} \end{pmatrix}$$

Obs: 1) T^2 se llama T^2 de Hotelling.

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

3) Si T^2 es muy grande, $\bar{X}$ estará lejos de μ_0. Luego se rechaza H_0.

χ^2 para varianza.

4) En este caso, $\underline{\text{si } H_0 \text{ es cierta}}$

$$T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,n-p}$$

distribución f (Para comparar varianzas) con p, n-p gl

$$\text{Ent } \alpha = P\Big(\underbrace{T^2 > \frac{(n-1)p}{n-p} \cdot F_{p,n-p}(\alpha)}_{\text{Rechazar } H_0} \;\Big|\; H_0 \text{ cierta}\Big).$$

dado que

$$= P\left(n(\bar{X} - \mu_0)' (S^{-1}) (\bar{X} - \mu_0) > \triangle \;\middle|\; H_0 \text{ cierta}\right)$$

Con datos

Si tenemos una PH de la forma

$H_0: \mu = \mu_0$, $H_a: \mu \neq \mu_0$ con nivel α

rechazamos si

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

Obs: T^2 es invariante bajo cambios de unidades de las mediciones de X

$$Y = \underset{p\times p}{\overset{\text{fijo}}{C}}\ \underset{p\times 1}{\overset{\text{Var.}}{X}} + \underset{p\times 1}{\overset{\text{cons}}{d}}$$

Pruebas más potentes ← Método de razón de Verosimilitud:

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

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

Asumiendo $H_0: \mu = \mu_0$

Verosimilitud normal multivariada ← $$L(\mu_0,\Sigma) = \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left(-\frac{1}{2}\sum (X_j - \mu_0)' \Sigma^{-1} (X_j - \mu_0)\right)$$

Valor fijo

← matriz.

Pues el resto es constante

La varianza generalizada determina el max de la verosimilitud → Σ

Podemos variar Σ para encontrar valores más posibles, con las obs se tiene:

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

→ Variable

Calcular matriz var y cov asumiendo H_0 True. ← $$\hat{\Sigma}_0 = \frac{1}{n}\sum_{j=1}^{n} (X_j - \mu_0)'(X_j - \mu_0)$$

Para determinar si μ_0 es un valor posible de μ_1 se comparan los máximos $\underbrace{L(\mu_0,\Sigma)}_{\text{Asumiendo } H_0 \text{ true}}$ y $\underbrace{L(\mu,\Sigma)}_{\text{Sin asumir}}$

Máx sobre todos los posibles Σ (μ fijo)

$$\Lambda = \frac{\max_{\Sigma} L(\mu_0,\Sigma)}{\underbrace{\max_{\mu,\Sigma} L(\mu,\Sigma)}_{\text{Max sobre todos los posibles } \mu \text{ y } \Sigma}} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{n/2}$$

Cte $(|S|)^{-n/2}$

Lambda de Wilk

(Si Λ pequeño $H_0: \mu = \mu_0$ es poco probable y H_0 se rechaza)

(Si μ_0 está muy lejos de μ, $|\hat{\Sigma}_0|$ va a ser mucho más grande que $|\hat{\Sigma}|$)

Más concretamente,

H_0 se rechaza si $\Lambda < C_\alpha$

↓ α percentil de la distribución Λ