Introducción

Distribución normal univariante

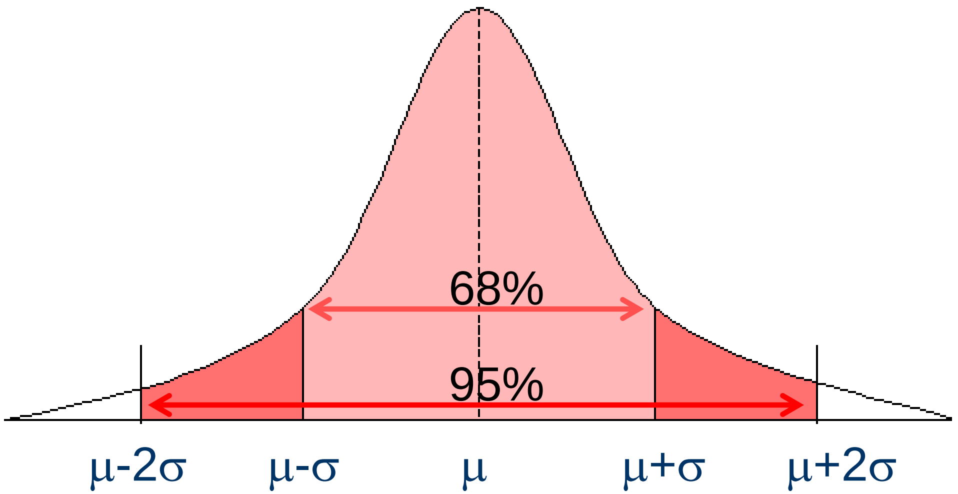
$X \sim N(\mu, \sigma^2)$, donde $\begin{matrix} \mu = E(X) \\ \sigma^2 = V(X) \end{matrix}$

con función de densidad $\quad f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$

$x \in \mathbb{R}$

$\mu \in \mathbb{R}$

$\sigma^2 \geq 0$



Normal bivariante

Normal bivariante

$$X \sim N_2(\mu, \Sigma), \text{ donde } \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}; \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix},$$

con función de densidad

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \exp\left\{ -\frac{1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}} \right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}} \right)^2 - 2\rho_{12} \left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}} \right) \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}} \right) \right] \right\}$$

Normal bivariante

Ejemplo

$$p = 2$$

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}; \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}; \quad \rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$$

Desarrollar $f(x_1, x_2)$

Normal bivariante

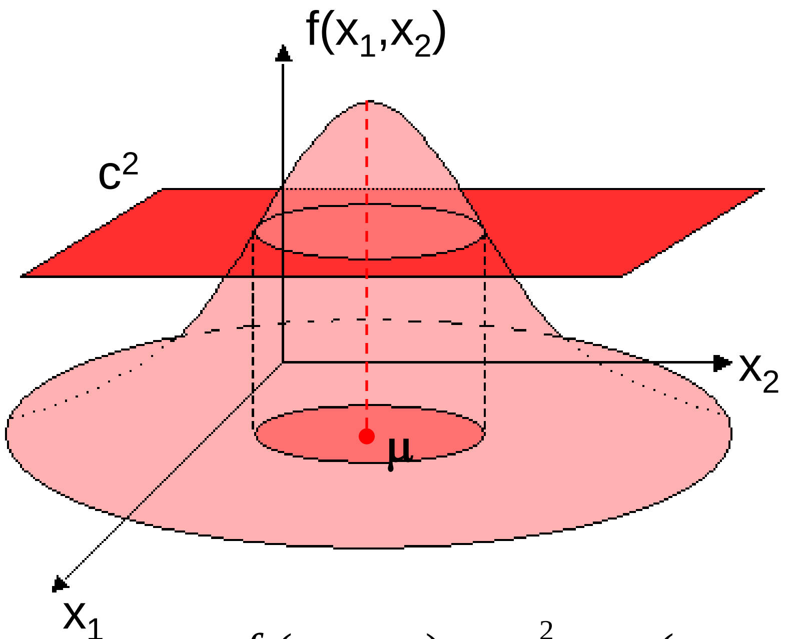
Propiedades

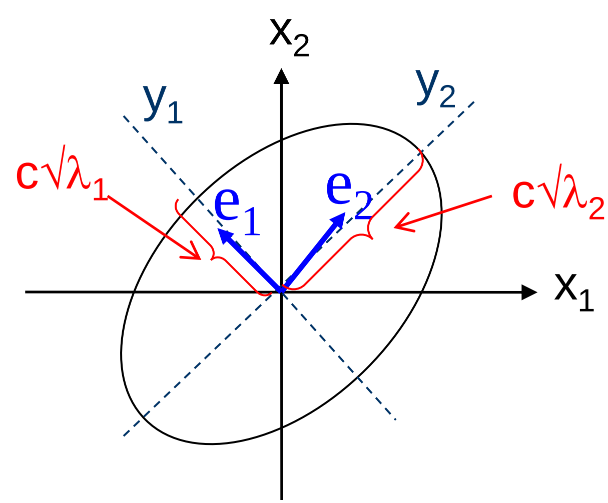
- $\rho_{12} = 0 \Leftrightarrow f(x_1, x_2) = f(x_1) f(x_2) \Leftrightarrow X_1, X_2$ independientes

- (λ, e) autovalor y autovector de $\Sigma \Rightarrow (1/\lambda, e)$ autovalor y autovector de Σ^{-1}

Normal bivariante

Representación gráfica





$$f(x_1, x_2) = c^2 \Leftrightarrow (x-\mu)' \Sigma^{-1} (x-\mu) = c^2$$

λ_1, λ_2 autovalores de Σ
e_1, e_2 autovectores de Σ

Normal bivariante

Ejemplo

Hallar las elipses de densidad constante para

$$x \sim N_2(\mu, \Sigma)$$

$$\sigma_{11} = \sigma_{22} \Rightarrow \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

Normal multivariante

Normal multivariante

$X \sim N_p(\mu, \Sigma)$, donde $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$; Σ definida positiva,

con función de densidad:

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

$x \in \mathbb{R}^p$

Normal multivariante

Propiedades

$$(i)\quad X \sim N_p(\mu, \Sigma)\ ;\ a = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}.\ Entonces\quad :$$

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

$$\forall a \in \mathbb{R}^p,\ a'X \text{ normal} \Rightarrow X \text{ normal } p\text{-variante}$$

Normal multivariante

$$(ii) \quad X \sim N_p(\mu, \Sigma); \quad A_{qxp} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{q1} & a_{q2} & \cdots & a_{qp} \end{pmatrix}$$

$$AX \sim N_q(A\mu, A\Sigma A')$$

$$d = \begin{pmatrix} d_1 \\ \vdots \\ d_p \end{pmatrix}; X + d \sim N_q(\mu + d, \Sigma)$$

Normal multivariante

$$(iii) \qquad X \sim N_p(\mu, \Sigma); \quad X = \begin{pmatrix} X^{(1)} \\ \hline X^{(2)} \end{pmatrix}$$

$$\mu = \begin{pmatrix} \mu^{(1)} \\ \hline \mu^{(2)} \end{pmatrix}; \qquad \Sigma = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right)$$

$$Entonces \quad :$$

$$X^{(1)} \sim N(\mu^{(1)}, \Sigma_{11})$$

$$X^{(2)} \sim N(\mu^{(2)}, \Sigma_{22})$$

Normal multivariante

$$(iv) \quad X^{(1)}, X^{(2)} \ independie \quad ntes \Rightarrow \operatorname{cov}(X^{(1)}, X^{(2)}) = \Sigma_{12} = 0$$

$$\left(\frac{X^{(1)}}{X^{(2)}} \right) \sim N\left(\left(\frac{\mu^{(1)}}{\mu^{(2)}} \right), \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array} \right) \right)$$

$$X^{(1)}, X^{(2)} \ son \ indepe \quad ndientes \Leftrightarrow \Sigma_{12} = 0$$

Normal multivariante

$$(v) \qquad X^{(1)} \sim N_{q1}(\mu_1, \Sigma_1)$$

$$X^{(2)} \sim N_{q2}(\mu_2, \Sigma_2) \;\; independdi \quad entes \;\; \Rightarrow$$

$$\left(\frac{X^{(1)}}{X^{(2)}} \right) \sim N_{q1+q2}\left(\left(\frac{\mu^{(1)}}{\mu^{(2)}} \right), \left(\begin{array}{c|c} \Sigma_{11} & 0 \\ \hline 0 & \Sigma_{22} \end{array} \right) \right)$$

Normal multivariante

$$(vi) \quad Sea \quad \left(\frac{X^{(1)}}{X^{(2)}} \right) \sim N\left(\left(\frac{\mu^{(1)}}{\mu^{(2)}} \right), \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array} \right) \right)$$

$$\left| \Sigma_{22} \right| > 0$$

$$X_1 \mid X_2 = x_2 \sim N(\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$$

Normal multivariante

(vii) Distribución de combinación lineal de normales

$X_1, X_2, \ldots, X_n$ indep.

$$X_i \sim N_p(\mu_{\text{i}}, \Sigma) \;\; ; \; U = c_1 X_1 + \cdots + c_n X_n = \sum_{i=1}^{n} c_i X_i .$$

Entonces $\quad U \sim N_p(\sum_{i=1}^{n} c_i X_i, (\sum_{i=1}^{n} c_i^2) \Sigma).$

Normal multivariante

(viii) Distribución conjunta de normales

Si, además, $V = \sum_{i=1}^{n} b_i X_i$

entonces : $$\begin{pmatrix} U \\ V \end{pmatrix} = N_{2p}\left(\begin{pmatrix} \sum_{i=1}^{n} c_i \mu_i \\ \sum_{i=1}^{n} b_i \mu_i \end{pmatrix}, \begin{pmatrix} (\sum_{i=1}^{n} c_i^2)\Sigma & (b'c)\Sigma \\ (b'c)\Sigma & (\sum_{i=1}^{n} b_i^2)\Sigma \end{pmatrix} \right)$$

Por tanto, U y V son independientes $\Leftrightarrow b'c = 0$

Distribución χ^2

$$\chi_1^2 \equiv Z_1^2, \quad donde \quad Z_1 \equiv N(0,1)$$

$$\sum_{i=1}^{n} Z_{1i}^2 = \chi_n^2,$$

$$donde \quad Z_i \equiv N(0,1) \ independientes \ , i = 1,\ldots,n.$$

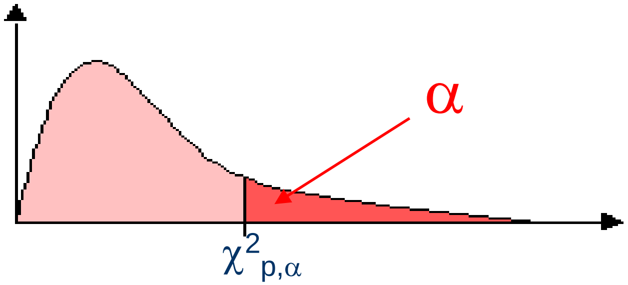
Distribución χ^2

Propiedades

Sea $X \equiv N_p(\mu, \Sigma)$ y $|\Sigma| > 0$. Entonces :

$$(i) \quad (X - \mu)' \Sigma^{-1} (X - \mu) \equiv \chi^2_p$$

$$(ii) \quad P_X\left\{x \in \mathbb{R}^p : (x - \mu)' \Sigma^{-1} (x - \mu) \leq \chi^2_{p,\alpha}\right\} = 1 - \alpha$$



Distribución χ^2

Ejemplo

$$X_1, X_2, X_3, X_4 \equiv N_3(\mu_i, \begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{pmatrix}) \quad indep.$$

$$\mu_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \quad \mu_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \mu_4 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

Distribución χ^2

Ejemplo

$$(i) \quad \sum_{i=1}^{n} c_i X_i \quad con \quad c = \begin{pmatrix} -1 \\ 0 \\ -2 \\ 1 \end{pmatrix}$$

$$(ii) \quad \begin{pmatrix} \sum_{i=1}^{n} c_i X_i \\ \sum_{i=1}^{n} b_i X_i \end{pmatrix} \quad con \quad b = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}$$

$$(iii) \quad \sum_{i=1}^{n} c_i X_i \quad \sum_{i=1}^{n} b_i X_i \quad indep.?$$

$$(iv) \quad Dar \quad una \quad combinación \quad lineal \quad indep.$$

DISTRIBUCION DE WISHART

La distribución de Wishart es la que sigue una matriz aleatoria simétrica definida positiva, generaliza la distribución ji-cuadrado y juega un papel importante en inferencia multivariante. Un ejemplo destacado lo constituye la distribución de la matriz de covarianzas $\mathbf{S}$, calculada a partir de una matriz de datos donde las filas son observaciones normales multivariantes.

Definición

Si las filas de la matriz $\mathbf{Z}_{n\times p}$ son independientes $N_p(0, \Sigma)$ entonces diremos que la matriz $\mathbf{Q} = \mathbf{Z}'\mathbf{Z}$ es Wishart $W_p(\Sigma, n)$, con parámetros Σ y n grados de libertad.

Textos avanzados prueban que cuando Σ es definida positiva y $n \geq p$, la densidad de $\mathbf{Q}$ es

$$f(\mathbf{Q}) = c|\mathbf{Q}|^{(n-p-1)} \exp(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}\mathbf{Q})),$$

siendo

$$c^{-1} = 2^{np/2}\pi^{p(p-1)/4}|\Sigma|^{n/2} \prod_{i=1}^{p} \Gamma(\frac{1}{2}(n+1-i)).$$

PROPÍEDADES

1. Si $\mathbf{Q}_1, \mathbf{Q}_2$ son independientes Wishart $W_p(\Sigma, m), W_p(\Sigma, n)$, entonces la suma $\mathbf{Q}_1 + \mathbf{Q}_2$ es también Wishart $W_p(\Sigma, m+n)$.

2. Si $\mathbf{Q}$ es Wishart $W_p(\Sigma, n)$, y separamos las variables en dos conjuntos y consideramos las particiones correspondientes de las matrices Σ y $\mathbf{Q}$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix},$$

   Entonces $\mathbf{Q}_{11}$ es $W_p(\Sigma_{11}, n)$ y $\mathbf{Q}_{22}$ es $W_p(\Sigma_{22}, n)$.

3. Si $\mathbf{Q}$ es Wishart $W_p(\Sigma, n)$ y $\mathbf{T}$ es una matriz $p \times q$ de constantes, entonces $\mathbf{T}'\mathbf{Q}\mathbf{T}$ es $W_q(\mathbf{T}'\Sigma\mathbf{T}, n)$. En particular, si $\mathbf{t}$ es un vector, entonces

$$\frac{\mathbf{t}'\mathbf{Q}\mathbf{t}}{\mathbf{t}'\Sigma\mathbf{t}} \quad \text{es} \quad \chi_n^2.$$

DISTRIBUCION T^2 DE HOTELLING

Es una generalización multivariante de la distribución t de Student.

Definición

Si $\mathbf{y}$ es $N_p(\mathbf{0}, \mathbf{I})$, $\mathbf{Q}$ es Wishart $W_p(\mathbf{I}, m)$ y además $\mathbf{y}$, $\mathbf{Q}$ son independientes, entonces

$$T^2 = m\mathbf{y}'\mathbf{Q}^{-1}\mathbf{y}$$

sigue la distribución T^2 de Hotelling, que se indica por $T^2(p, m)$.

PROPÍEDADES

1. Si $\mathbf{x}$ es $N_p(\mu,\Sigma)$ independiente de $\mathbf{M}$ que es $W_p(\Sigma, m)$, entonces

$$T^2 = m(\mathbf{x}-\mu)'\mathbf{M}^{-1}(\mathbf{x}-\mu) \sim T^2(p, m).$$

2. T^2 está directamente relacionada con la distribución de Fisher-Snedecor

$$T^2(p, m) \equiv \frac{mp}{m-p+1} F^p_{m-p+1}.$$

3. Si $\overline{\mathbf{x}}, \mathbf{S}$ son el vector de medias y la matriz de covarianzas de la matriz $\mathbf{X}_{n\times p}$ con filas independientes $N_p(\mu, \Sigma)$, entonces

$$(n-1)(\overline{\mathbf{x}}-\mu)'\mathbf{S}^{-1}(\overline{\mathbf{x}}-\mu) \sim T^2(p, n-1),$$

y por lo tanto

$$\frac{n-p}{p}(\overline{\mathbf{x}}-\mu)'\mathbf{S}^{-1}(\overline{\mathbf{x}}-\mu) \sim F^p_{n-p}.$$

4. Si $\overline{\mathbf{x}}, \mathbf{S}_1, \overline{\mathbf{y}}, \mathbf{S}_2$ son el vector de medias y la matriz de covarianzas de las matrices $\mathbf{X}_{n_1\times p}, \mathbf{Y}_{n_2\times p}$, respectivamente, con filas independientes $N_p(\mu, \Sigma)$, y consideramos la estimación conjunta centrada de Σ

$$\widetilde{\mathbf{S}} = (n_1\mathbf{S}_1 + n_2\mathbf{S}_2)/(n_1+n_2-2),$$

entonces

$$T^2 = \frac{n_1 n_2}{n_1+n_2}(\overline{\mathbf{x}}-\overline{\mathbf{y}})'\widetilde{\mathbf{S}}^{-1}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) \sim T^2(p, n_1+n_2-2)$$

y por lo tanto

$$\frac{n_1+n_2-1-p}{(n_1+n_2-2)p} T^2 \sim F^p_{n_1+n_2-1-p}.$$

DISTRIBUCIO DE WILKS

Definición

Si las matrices A, B de orden $p \times p$ son independientes Wishart $W_p(\Sigma, m), W_p(\Sigma, n)$, respectivamente, con $m \geq p$, la distribución del cociente de determinantes

$$\Lambda = \frac{|\mathrm{A}|}{|\mathrm{A}+\mathrm{B}|}$$

es, por definición, la distribución lambda de Wilks, que indicaremos por $\Lambda(p, m, n)$.

PROPÍEDADES

1. $0 \leq \Lambda \leq 1$ y además Λ no depende de Σ. Por lo tanto, podemos estudiarla suponiendo $\Sigma = \mathbf{I}$.

2. Su distribución es equivalente a la del producto de n variables beta independientes:
$$\Lambda(p, m, n) \sim \prod_{i=1}^{n} U_i,$$
donde U_i es beta $B(\frac{1}{2}(m+i-p), \frac{1}{2}p)$. (Obsérvese que debe ser $m \geq p$).

3. Los parámetros se pueden permutar manteniendo la misma distribución. Concretamente:
$$\Lambda(p, m, n) \sim \Lambda(n, m+n-p, p).$$

4. Para valores 1 y 2 de p y n, la distribución de Λ equivale a la F, según las fórmulas
$$\begin{array}{lll} \frac{1-\Lambda}{\Lambda}\frac{m}{n} & \sim F_m^n & (p=1) \\ \frac{1-\Lambda}{\Lambda}\frac{m-p+1}{p} & \sim F_{m-p+1}^p & (n=1) \\ \frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\frac{m-1}{n} & \sim F_{2(m-1)}^{2n} & (p=2) \\ \frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\frac{m-p+1}{p} & F_{2(m-p+1)}^{2p} & (n=2) \end{array} \tag{2.8}$$

5. En general, una transformación de Λ equivale, exacta o asintóticamente, a la distribución F. Si $\Lambda(p, n-q, q)$ es Wilks con n relativamente grande, consideremos
$$F = \frac{ms - 2\lambda}{pq}\frac{1-\Lambda^{1/s}}{\Lambda^{1/s}} \tag{2.9}$$

con $m = n-(p+q+1)/2$, $\lambda = (pq-2)/4$, $s = \sqrt{(p^2q^2-4)/(p^2+q^2-5)}$. Entonces F sigue asintóticamente la distribución F con pq y $(ms-2\lambda)$ g. de lib. (Rao, 1973, p.556).