MAT-269: Distribuciones T^2 de Hotelling, beta multivariada y Λ de Wilks

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Definición 1 (Distribución T^2 de Hotelling)

Sean $\boldsymbol{z} \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$ y $\boldsymbol{U} \sim \mathsf{W}_p(m, \boldsymbol{I}_p)$ independientes. La distribución de la variable aleatoria

$$T^2 = \boldsymbol{z}^\top \big(\boldsymbol{U}/m\big)^{-1} \boldsymbol{z} = m\boldsymbol{z}^\top \boldsymbol{U}^{-1} \boldsymbol{z},$$

se denomina T^2 de Hotelling con p y m grados de libertad, y escribimos:

$$T^2 = m\boldsymbol{z}^\top \boldsymbol{U}^{-1} \boldsymbol{z} \sim \mathsf{T}^2(p, m).$$

Resultado 1

Si $\boldsymbol{x}$ y $\boldsymbol{W}$ son independientemente distribuídos $\mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\mathsf{W}_p(m, \boldsymbol{\Sigma})$, respectivamente. Entonces,

$$m(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{W}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \sim \mathsf{T}^2(p, m).$$

Demostración:

Basta notar que $\boldsymbol{z} = \boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$, donde $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$, y

$$\boldsymbol{U} = \boldsymbol{B}^{-1}\boldsymbol{W}\boldsymbol{B}^{-\top} \sim \mathsf{W}_p(m, \boldsymbol{B}^{-1}\boldsymbol{\Sigma}\boldsymbol{B}^{-\top}) \stackrel{\mathsf{d}}{=} \mathsf{W}_p(m, \boldsymbol{I}).$$

Es decir,

$$\begin{aligned} m\boldsymbol{z}^\top \boldsymbol{U}^{-1}\boldsymbol{z} &= m(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{B}^{-\top}\boldsymbol{B}^\top \boldsymbol{W}^{-1}\boldsymbol{B}\boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \\ &= m(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{W}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}), \end{aligned}$$

de donde sigue el resultado.

Distribución T^2 de Hotelling

Observación:

Sabemos que $\overline{\boldsymbol{x}}$ y $\boldsymbol{S}$ son independientes y $\overline{\boldsymbol{x}} \sim \mathsf{N}_p(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma})$, mientras que

$$\boldsymbol{Q} = \boldsymbol{X}^\top \boldsymbol{C}\boldsymbol{X} \sim \mathsf{W}_p(n-1, \boldsymbol{\Sigma}).$$

De ahí que

$$T^2 = n(\overline{\boldsymbol{x}} - \boldsymbol{\mu})^\top \boldsymbol{S}^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu}) \sim \mathsf{T}^2(p, n-1).$$

Es fácil notar que $\sqrt{n}\boldsymbol{B}^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu}) \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$ y $\boldsymbol{B}^{-1}\boldsymbol{Q}\boldsymbol{B}^{-\top} \sim \mathsf{W}_p(n-1, \boldsymbol{\Sigma})$, luego

$$\begin{aligned} T^2 &= \{\sqrt{n}\boldsymbol{B}^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu})\}^\top \big(\boldsymbol{B}^{-1}\boldsymbol{Q}\boldsymbol{B}^{-\top}/(n-1)\big)^{-1}\{\sqrt{n}\boldsymbol{B}^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu})\} \\ &= n(\overline{\boldsymbol{x}} - \boldsymbol{\mu})^\top \boldsymbol{S}^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu}), \end{aligned}$$

pues $\boldsymbol{S} = \frac{1}{n-1}\boldsymbol{Q}$, es decir $(n-1)\boldsymbol{Q}^{-1} = \boldsymbol{S}^{-1}$. Alternativamente podemos considerar $\boldsymbol{S}_* = \frac{n-1}{n}\boldsymbol{S}$. Entonces

$$T^2 = (n-1)(\overline{\boldsymbol{x}} - \boldsymbol{\mu})^\top \boldsymbol{S}_*^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu}) \sim \mathsf{T}^2(p, n-1).$$

Resultado 2

Considere $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ variables aletorias independientes desde $\mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\boldsymbol{y}_i = \boldsymbol{A}\boldsymbol{x}_i + \boldsymbol{b}$, para $i = 1, \ldots, n$, donde $\boldsymbol{A} \in \mathbb{R}^{q\times p}$ con $\operatorname{rg}(\boldsymbol{A}) = q$. Si $\overline{\boldsymbol{x}}$ y $\boldsymbol{S}_X$ son la media muestral y la matriz de covarianza muestral, respectivamente. Tenemos

$$\overline{\boldsymbol{y}} = \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{b} \sim \mathsf{N}_q\Big(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \frac{1}{n}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top\Big)$$

$$(n-1)\boldsymbol{S}_Y = (n-1)\boldsymbol{A}\boldsymbol{S}_X\boldsymbol{A}^\top \sim \mathsf{W}_p(n-1, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top).$$

De este modo,

$$n(\boldsymbol{A}\overline{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{\mu})^\top(\boldsymbol{A}\boldsymbol{S}_X\boldsymbol{A}^\top)^{-1}(\boldsymbol{A}\overline{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{\mu}) \sim \mathsf{T}^2(q, n-1).$$

Resultado 3

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

Demostración:

Considere $T^2 = m\boldsymbol{z}^\top \boldsymbol{U}^{-1}\boldsymbol{z}$ y escriba

$$T^2 = m\,\frac{\boldsymbol{z}^\top \boldsymbol{U}^{-1}\boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{z}}\,\boldsymbol{z}^\top \boldsymbol{z},$$

dado que $\boldsymbol{U}$ es independiente de $\boldsymbol{z}$. Tenemos que $\boldsymbol{z}^\top \boldsymbol{z}/\boldsymbol{z}^\top \boldsymbol{U}^{-1}\boldsymbol{z}$ dado $\boldsymbol{z}$ tiene distribución[1] $\chi^2(m-p+1)$. Además $\boldsymbol{z}^\top \boldsymbol{z} \sim \chi^2(p)$. De ahí que

$$\begin{aligned} T^2 &= m\,\frac{\boldsymbol{z}^\top \boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{z}/\boldsymbol{z}^\top \boldsymbol{U}^{-1}\boldsymbol{z}} \stackrel{\mathsf{d}}{=} \frac{mp}{m-p+1}\,\frac{\chi^2(p)/p}{\chi^2(m-p+1)/(m-p+1)} \\ &= \frac{mp}{m-p+1} F(p, m-p+1). \end{aligned}$$

[1] ver Resultado 10 en Slides 4.

Definición 2

Sean $\boldsymbol{A} \sim \mathsf{W}_p(n_1, \boldsymbol{\Sigma})$ y $\boldsymbol{B} \sim \mathsf{W}_p(n_2, \boldsymbol{\Sigma})$ independientes, con $n_1 > p - 1$ y $n_2 > p - 1$. Considere $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{T}^\top \boldsymbol{T}$ donde $\boldsymbol{T}$ es matriz triangular superior $p \times p$ con elementos diagonales positivos y $\boldsymbol{U}$ matriz simétrica $p \times p$ definida como $\boldsymbol{A} = \boldsymbol{T}^\top \boldsymbol{U} \boldsymbol{T}$. Entonces $\boldsymbol{A} + \boldsymbol{B}$ y $\boldsymbol{U}$ son independientes, con $\boldsymbol{A} + \boldsymbol{B} \sim \mathsf{W}_p(n_1 + n_2, \boldsymbol{\Sigma})$ y la densidad de $\boldsymbol{U}$ es

$$f(\boldsymbol{U}) = \frac{\Gamma_p(\frac{n_1+n_2}{2})}{\Gamma_p(\frac{n_1}{2})\Gamma_p(\frac{n_2}{2})} (\det \boldsymbol{U})^{(n_1-p-1)/2} \det(\boldsymbol{I}_p - \boldsymbol{U})^{(n_2-p-1)/2},$$

para $0 < \boldsymbol{U} < \boldsymbol{I}_p$.[2] En cuyo caso escribimos

$$\boldsymbol{U} \sim \mathsf{Beta}_p\Big(\frac{n_1}{2}, \frac{n_2}{2}\Big).$$

Evidentemente, también se tiene que $\boldsymbol{I} - \boldsymbol{U} \sim \mathsf{Beta}_p(n_2/2, n_1/2)$.

[2] Esto significa que $\boldsymbol{U} > 0$ y $\boldsymbol{I}_p - \boldsymbol{U} > 0$.

Considere $\boldsymbol{H} \sim \mathsf{W}_p(n_1, \boldsymbol{\Sigma})$ y $\boldsymbol{E} \sim \mathsf{W}_p(n_2, \boldsymbol{\Sigma})$ con $n_1 + n_2 \geq p$ y sea

$$\boldsymbol{H}\boldsymbol{E}^{-1}, \qquad \text{y} \qquad \boldsymbol{H}(\boldsymbol{E}+\boldsymbol{H})^{-1}.$$

Estas matrices no son simétricas y no llevan a densidades de interés. Sin embargo, $\boldsymbol{E}$ y $\boldsymbol{H}$ y de ahí que $\boldsymbol{H}+\boldsymbol{E}$ son positivas definidas con probabilidad 1. Defina

$$\boldsymbol{R} = \boldsymbol{E}^{-1/2}\boldsymbol{H}\boldsymbol{E}^{-1/2}, \qquad \boldsymbol{U} = (\boldsymbol{E}+\boldsymbol{H})^{-1/2}\boldsymbol{H}(\boldsymbol{E}+\boldsymbol{H})^{-1/2}.$$

Ahora

$$\begin{aligned}\boldsymbol{I} - \boldsymbol{U} &= (\boldsymbol{E}+\boldsymbol{H})^{-1/2}(\boldsymbol{E}+\boldsymbol{H}-\boldsymbol{H})(\boldsymbol{E}+\boldsymbol{H})^{-1/2} \\ &= (\boldsymbol{E}+\boldsymbol{H})^{-1/2}\boldsymbol{E}(\boldsymbol{E}+\boldsymbol{H})^{-1/2}.\end{aligned}$$

En este caso $\boldsymbol{R}$ tiene función de densidad[3]

$$f(\boldsymbol{R}) = \frac{1}{B_p(n_1/2, n_2/2)} (\det \boldsymbol{R})^{(n_1-p-1)/2} \det(\boldsymbol{I}+\boldsymbol{R})^{-(n_1+n_2)/2},$$

para $\boldsymbol{R} > 0$.

[3] $B_p(a,b) = \Gamma_p(a)\Gamma_p(b)/\Gamma_p(a+b)$.

Definición 3

Considere $\boldsymbol{A} \sim \mathsf{W}_p(m, \boldsymbol{I})$ y $\boldsymbol{B} \sim \mathsf{W}_p(n, \boldsymbol{I})$ independientes con $m \geq p$. Entonces se dice que

$$\Lambda = \frac{|\boldsymbol{A}|}{|\boldsymbol{A} + \boldsymbol{B}|} = |\boldsymbol{I} + \boldsymbol{A}^{-1}\boldsymbol{B}|,$$

tiene distribución Λ de Wilks con parámetros p, m y n y escribimos

$$\Lambda = |\boldsymbol{I} + \boldsymbol{A}^{-1}\boldsymbol{B}| \sim \Lambda(p, m, n)$$

Resultado 4

Tenemos

$$\Lambda(p, m, n) \stackrel{\mathsf{d}}{=} \prod_{i=1}^{n} u_i,$$

donde $u_1, \ldots, u_n$ son variables independientes y $u_i \sim \text{Beta}((m + i - p)/2, p/2)$, para $i = 1, \ldots, n$.

Para probar este resultado, considere los siguientes Lemas.

Lema 1

Sea $\boldsymbol{z} \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$ y $\boldsymbol{M} \sim \mathsf{W}_p(m, \boldsymbol{I})$. Entonces $\boldsymbol{z}^\top \boldsymbol{M}^{-1} \boldsymbol{z}$ y $\boldsymbol{M} + \boldsymbol{z}\boldsymbol{z}^\top$ son independientes.

Lema 2

Considere $\boldsymbol{z} \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$ y $\boldsymbol{M} \sim \mathsf{W}_p(m, \boldsymbol{I})$. Entonces

$$\frac{|\boldsymbol{M}|}{|\boldsymbol{M} + \boldsymbol{z}\boldsymbol{z}^\top|} \sim \mathsf{Beta}\Big(\frac{m-p+1}{2}, \frac{p}{2}\Big).$$

Demostración:

En efecto,

$$\begin{aligned}
\frac{|\boldsymbol{M}|}{|\boldsymbol{M} + \boldsymbol{z}\boldsymbol{z}^\top|} &= \frac{|\boldsymbol{M}|}{|\boldsymbol{M}(\boldsymbol{I} + \boldsymbol{M}^{-1}\boldsymbol{z}\boldsymbol{z}^\top)|} = \frac{|\boldsymbol{M}|}{|\boldsymbol{M}||\boldsymbol{I} + \boldsymbol{M}^{-1}\boldsymbol{z}\boldsymbol{z}^\top|} \\
&= \frac{1}{1 + \boldsymbol{z}^\top\boldsymbol{M}^{-1}\boldsymbol{z}} = \frac{m}{m + m\boldsymbol{z}^\top\boldsymbol{M}^{-1}\boldsymbol{z}} = \frac{m}{m + T^2},
\end{aligned}$$

como

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

el resultado sigue.

Demostración del Resultado 4:

Sea $\boldsymbol{Z} \sim \mathsf{N}_{n,p}(\boldsymbol{0}, \boldsymbol{I})$ y $\boldsymbol{B} = \boldsymbol{Z}^\top \boldsymbol{Z}$. Ahora sea $\boldsymbol{Z}_i$ una matriz $i \times p$ conformada por las primeras i filas de $\boldsymbol{Z}$, y sea

$$\boldsymbol{M}_i = \boldsymbol{A} + \boldsymbol{Z}_i^\top \boldsymbol{Z}_i, \qquad i = 1, \dots, n.$$

Note que $\boldsymbol{M}_0 = \boldsymbol{A}$, $\boldsymbol{M}_n = \boldsymbol{A} + \boldsymbol{B}$, y además

$$\boldsymbol{M}_i = \boldsymbol{M}_{i-1} + \boldsymbol{z}_i \boldsymbol{z}_i^\top.$$

Ahora, sea

$$\Lambda(p, m, n) = \frac{|\boldsymbol{A}|}{|\boldsymbol{A} + \boldsymbol{B}|} = \frac{|\boldsymbol{M}_0|}{|\boldsymbol{M}_n|} = \frac{|\boldsymbol{M}_0|}{|\boldsymbol{M}_1|} \frac{|\boldsymbol{M}_1|}{|\boldsymbol{M}_2|} \cdots \frac{|\boldsymbol{M}_{n-1}|}{|\boldsymbol{M}_n|} = u_1\, u_2 \cdots u_n,$$

con $u_i = |\boldsymbol{M}_{i-1}|/|\boldsymbol{M}_i|$, para $i = 1, \dots, n$. Ahora, usando el Lema 2

$$u_i = \frac{|\boldsymbol{M}_{i-1}|}{|\boldsymbol{M}_{i-1} + \boldsymbol{z}_i \boldsymbol{z}_i^\top|} = \frac{1}{1 + \boldsymbol{z}_i^\top \boldsymbol{M}_{i-1}^{-1} \boldsymbol{z}_i} \sim \mathrm{Beta}\Big(\frac{m + i - p}{2}, \frac{p}{2}\Big).$$

Para notar la indepencia, recuerde que (por el Lema 1) $\boldsymbol{M}_i$ es independiente de

$$1 + \boldsymbol{z}_i^\top \boldsymbol{M}_{i-1}^{-1} \boldsymbol{z}_i = \frac{|\boldsymbol{M}_i|}{|\boldsymbol{M}_{i-1}|} = u_i^{-1}.$$

Además, u_i es independiente de $\boldsymbol{z}_{i+1}, \dots, \boldsymbol{z}_n$ y

$$\boldsymbol{M}_{i+l} = \boldsymbol{M}_i + \sum_{k=1}^{l} \boldsymbol{z}_{i+k} \boldsymbol{z}_{i+k}^\top,$$

de lo que sigue que u_i es también independiente de $\boldsymbol{M}_{i+1}, \boldsymbol{M}_{i+2}, \dots, \boldsymbol{M}_n$ y de ahí que independiente de $u_{i+1}, \dots, u_n$. Por tanto el resultado sigue.

Desde la relación entre la distribución Beta con la F, sigue que:

$$\frac{1-\Lambda(p,m,1)}{\Lambda(p,m,1)} \sim \frac{p}{m-p-1}\,F(p,m-p-1),$$

$$\frac{1-\Lambda(1,m,n)}{\Lambda(1,m,n)} \sim \frac{n}{m}\,F(n,m),$$

$$\frac{1-\sqrt{\Lambda(p,m,2)}}{\sqrt{\Lambda(p,m,2)}} \sim \frac{p}{m-p+1}\,F(2p,2(m-p+1)),$$

$$\frac{1-\sqrt{\Lambda(2,m,n)}}{\sqrt{\Lambda(2,m,n)}} \sim \frac{n}{m-1}\,F(2n,2(m-1)).$$

Para otros valores de n y p (siempre que m sea grande), podemos usar la aproximación de Bartlett:

$$-[m-\tfrac{1}{2}(p-n+1)]\log\Lambda(p,m,n) \xrightarrow{\mathsf{D}} \chi^2(np).$$