Inferencia I

UNIVERSITAT DE BARCELONA

MULTIVARIATE ANALYSIS
MESIO (15-16)

PROF. SERGI CIVIT
PROF. MIQUEL SALICRÚ

Función de verosimilitud

Sea $x_1, x_2,\ldots, x_n$ una muestra aleatoria representativa de la variable $X \approx N_p(\mu, \Sigma)$. Ya que $x_1, x_2,\ldots, x_n$ son independientes, la función de densidad conjunta es el producto de marginales:

$$L(x_1,\ldots,x_n) = \prod_{i=1}^{n}\left\{\frac{1}{|\Sigma|^{1/2}(2\pi)^{p/2}} e^{-\frac{1}{2}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)}\right\}$$

$$= (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left\{-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)\right\}$$

Interés práctico

Estimación de parámetros. $\hat{\mu}$ y $\hat{\Sigma}$ maximizan la función de verosimilitud: son solución del sistema de ecuaciones

$$\frac{\partial L(x_1,\ldots,x_n)}{\partial \mu}=0 \quad , \quad \frac{\partial L(x_1,\ldots,x_n)}{\partial \Sigma}=0$$

Test de hipótesis. $H_0:\theta\in\Theta_0$ se rechaza en favor de $H_1:\theta\in\Theta$ si

$$\Lambda=\frac{\max_{\theta\in\Theta_0} L(x_1,\ldots,x_n,\theta)}{\max_{\theta\in\Theta} L(x_1,\ldots,x_n,\theta}<c$$

Asintoticamente: $-2\ln\Lambda\approx\chi^2_{\nu-\nu_0}$

Equivalencias en la función de verosimilitud

$$\ln L(x_1,\ldots,x_n)=\frac{-np}{2}\ln(2\pi)-\frac{n}{2}\ln|\Sigma|\boxed{-\tfrac{1}{2}\underline{\underline{\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)}}}$$

Igualdad 1.

$$\ln L(x_1,\ldots,x_n)=\frac{-np}{2}\ln(2\pi)-\frac{n}{2}\ln|\Sigma|-\frac{n}{2}\underline{\underline{tr\left[\Sigma^{-1}\cdot S_0\right]}}$$

Igualdad 2

$$\ln L(x_1,\ldots,x_n)=\frac{-np}{2}\ln(2\pi)-\frac{n}{2}\ln|\Sigma|-\frac{n}{2}\underline{\underline{tr\left[\Sigma^{-1}\cdot S\right]}}-\frac{n}{2}\underline{\underline{(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu)}}$$

Estimación de parámetros 1/2

$$\ln L(x_1,\dots,x_n) = \frac{-np}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{n}{2}\mathrm{tr}(\Sigma^{-1}\cdot S) - \frac{n}{2}(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu)$$

1.- $\dfrac{\partial L(x_1,\dots,x_n)}{\partial\mu} = 0 \Leftrightarrow \underline{\underline{\dfrac{\partial}{\partial\mu}\left(-\dfrac{n}{2}(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu)\right)}} = 0 \Leftrightarrow \dfrac{n}{2}\cdot 2\Sigma^{-1}(\mu-\bar{X}) = 0 \Leftrightarrow \underline{\underline{\hat{\mu} = \bar{X}}}$

2.- $\dfrac{\partial L(x_1,\dots,x_n)}{\partial\Sigma} = 0 \Leftrightarrow \underline{\underline{\dfrac{\partial}{\partial\Sigma}\left(-\dfrac{n}{2}\ln|\Sigma| - \dfrac{n}{2}\mathrm{tr}(\Sigma^{-1}\cdot S) - \dfrac{n}{2}(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu)\right)}} = 0$

$$-\frac{n}{2}\cdot(2\Sigma^{-1} - \mathrm{diag}(\Sigma^{-1})) - \frac{n}{2}\cdot(-2\Sigma^{-1}S\Sigma^{-1} + \mathrm{diag}(\Sigma^{-1}S\Sigma^{-1}))$$

$$\underbrace{-\frac{n}{2}(-2\Sigma^{-1}((\bar{X}-\mu)(\bar{X}-\mu)')\Sigma^{-1} + \mathrm{diag}(\Sigma^{-1}((\bar{X}-\mu)(\bar{X}-\mu)')\Sigma^{-1}))}_{=0\ (\hat{\mu}=\bar{X})} = 0$$

Estimación de parámetros 2/2

$$-\frac{n}{2}\cdot(2\Sigma^{-1} - \mathrm{diag}(\Sigma^{-1})) - \frac{n}{2}\cdot(-2\Sigma^{-1}S\Sigma^{-1} + \mathrm{diag}(\Sigma^{-1}S\Sigma^{-1})) = 0$$

$$\Leftrightarrow\ 2\cdot(\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1}) - \mathrm{diag}(\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1}) = 0$$

$$\Leftrightarrow\ \Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1} = 0\ \Leftrightarrow\ \underline{\underline{\hat{\Sigma} = S}}$$

Test de la media

$$\left.\begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array}\right\}$$

$$\Lambda = \frac{\max_{\theta \in \Theta_0} L(x_1, \dots, x_n, \theta)}{\max_{\theta \in \Theta} L(x_1, \dots, x_n, \theta)} < c$$

Test de la media 1/4

$$\max L_{H_0}(x_1, \dots, x_n, \theta)$$

H_0: $$\frac{\ln L(x_1, \dots, x_n)}{\partial \Sigma} = 0 \Leftrightarrow \frac{\frac{-np}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{n}{2}\mathrm{tr}(\Sigma^{-1} \cdot S_0)}{\partial \Sigma} = 0$$

$$\Leftrightarrow -\frac{n}{2} \cdot (2\Sigma^{-1} - \mathrm{diag}(\Sigma^{-1})) - \frac{n}{2}(-2\Sigma^{-1} S_0 \Sigma^{-1} + \mathrm{diag}(\Sigma^{-1} S_0 \Sigma^{-1})) = 0$$

$$\Leftrightarrow \Sigma^{-1} - \Sigma^{-1} S_0 \Sigma^{-1} = 0 \Leftrightarrow \hat{\Sigma} = S_0$$

$$\max L_{H_0}(x_1, \dots, x_n) = (2\pi)^{-np/2} |S_0|^{-n/2} \exp\{-\frac{n}{2}\mathrm{tr} S_0^{-1} S_0\} = (2\pi)^{-np/2} |S_0|^{-n/2} e^{-np/2}$$

Test de la media 2/4

$$\max L_{H_0}(x_1,\ldots,x_n,\theta)$$

H_1: $\hat{\mu}=\bar{X}$ y $\hat{\Sigma}=S$

$$L_{H_1}(x_1,\ldots,x_n)=(2\pi)^{-np/2}|S|^{-n/2}\exp\{-\frac{n}{2}trS^{-1}S-\frac{n}{2}(\bar{X}-\bar{X})'\Sigma^{-1}(\bar{X}-\bar{X})'\}$$

$$\underline{\underline{\max L_{H_1}(x_1,\ldots,x_n)=(2\pi)^{-np/2}|S|^{-n/2}e^{-np/2}}}$$

$$\Lambda=\frac{\max_{\theta\in\Theta_0}L(x_1,\ldots,x_n,\theta)}{\max_{\theta\in\Theta}L(x_1,\ldots,x_n,\theta)}=\left(\frac{|S_0|}{|S|}\right)^{-n/2}$$

Test de media 3/4

$$\Lambda=\underline{\underline{\left(\frac{|S_0|}{|S|}\right)}}^{-n/2}=\left(\frac{|S+(\bar{X}-\mu_0)(\bar{X}-\mu_0)'|}{|S|}\right)^{-n/2}$$

$$=|S^{-1}S+S^{-1}(\bar{X}-\mu_0)(\bar{X}-\mu_0)'|^{-n/2}$$

$$=1+(\bar{X}-\mu_0)'S^{-1}(\bar{X}-\mu_0)=1+\underline{\underline{\frac{T^2}{n-1}}}$$

siendo

$$T^2=(n-1)(\bar{X}-\mu_0)'S_n^{-1}(\bar{X}-\mu_0)$$

Test de la media 4/4

Estadístico de contraste (transformación monótona del estadístico Λ para generalizar el caso univariante)

$$\underline{\underline{T^2}} = n\cdot(\bar{X}-\mu_0)' S_{n-1}^{-1} (\bar{X}-\mu_0)$$

$$= \underbrace{n^{1/2}(\bar{X}-\mu_0)'}_{N_p(0,\Sigma)} \cdot \underbrace{S_{n-1}^{-1}}_{W_{p,n-1}(\Sigma)} \cdot \underbrace{n^{1/2}(\bar{X}-\mu_0)}_{N_p(0,\Sigma)} \approx \underline{\underline{\frac{(n-1)p}{n-p} F_{p,n-p}}}$$

Criterio de decisión. El p valor asociado a la ley F (Fisher-Snedecor) se obtiene igual que en el caso univariante

p-dimensional confidence region

Let $\boldsymbol{\theta}$ be a vector of unknown population parameters such that $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. The $100(1-\alpha)\%$ Confidence Region determined by data

$$\mathbf{X}' = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$$

is denoted $R(\mathbf{X})$, is the region satisfying

$$P[R(\mathbf{X}) \in \boldsymbol{\theta}] = 1-\alpha$$

Confidence region about mean

The corresponding confidence region in p-dimensional multivariate space:

$$P\left[n(\bar{\mathbf{x}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}) \leq \frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha)\right] = 1-\alpha$$

The $100(1-\alpha)\%$ confidence region for the mean vector $\boldsymbol{\mu}$ of a p-dimensional normal distribution is the ellipsoid determined by all possible points ? that satisfy,

$$n(\bar{\mathbf{x}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}) \leq \frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha)$$

Test con datos apareados

Con muestras de tamaño n, X e Y apareadas, se considera la variable diferencia D=Y-X. El test de medias, se reduce al test de una población (la población de diferencias)

X	Y	D=Y-X
x_1	y_1	$d_1 = y_1 - x_1$
x_2	y_2	$d_2 = y_2 - x_2$
x_n	y_n	$d_n = y_n - x_n$

$$\left.\begin{array}{l} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{array}\right\} \Leftrightarrow \left.\begin{array}{l} H_0: d = d_0 \\ H_1: d \neq d_0 \end{array}\right\}$$

En este caso, se utiliza el estadístico T^2-Hotelling

$$\underline{\underline{T^2}} = n \cdot (\bar{d}-d_0)' S_d^{-1} (\bar{d}-d_0) \approx \underline{\underline{\frac{(n-1)p}{n-p} F_{p,n-p}}}$$

con criterio de decisión acorde a la ley F (Fisher-Snedecor)

Test de la matriz de varianzas-covarianzas

$$\left.\begin{array}{l} H_0: \Sigma = \Sigma_0 \\ H_1: \Sigma \neq \Sigma_0 \end{array}\right\}$$

$\Sigma_0 = \sigma^2 I$ con σ desconocido

$\Sigma_0 = I \ (\sigma = 1)$

esfericidad

$$\Sigma_0 = \sigma^2 \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \dots & 1 \end{pmatrix}$$

igual varianza y correlación

$$\Sigma_0 = \begin{pmatrix} \Sigma_{11} & 0 & \dots & 0 \\ 0 & \Sigma_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \Sigma_{kk} \end{pmatrix}$$

independencia por cajas

Adherencia a una matriz

$$\left.\begin{array}{l} H_0: \Sigma = \Sigma_0 \\ H_1: \Sigma \neq \Sigma_0 \end{array}\right\}$$

H_0: $\max_{\theta\in\Theta_0} L(x_1,\dots,x_n,\theta)$ se consigue cuando $\hat{\mu} = \overline{X}$, $\Sigma = \Sigma_0$

H_1: $\max_{\theta\in\Theta} L(x_1,\dots,x_n,\theta)$ se consigue cuando $\hat{\mu} = \overline{X}$, $\hat{\Sigma} = S$

Estadístico,

$$-2\ln\Lambda = -2\ln\frac{\max_{\theta\in\Theta_0} L(x_1,\dots,x_n,\theta)}{\max_{\theta\in\Theta} L(x_1,\dots,x_n,\theta)} = -\frac{4}{n}\ln\frac{|S|}{|\Sigma_0|} \approx \chi^2_{p(p+1)/2}$$

Criterio de decisión: igual al caso univariante

Otros test

H_0: Esfericidad

H_0: Igual varianza y correlación

H_0: Independencia por cajas

Paso 1. Estimar parámetros que maximizan

$$L_{\theta\in\Theta_0}(x_1,\ldots,x_n,\theta) \quad y \quad L_{\theta\in\Theta}(x_1,\ldots,x_n,\theta)$$

Paso 2. Obtener estadístico

$$\Lambda = \frac{\max_{\theta\in\Theta_0} L(x_1,\ldots,x_n,\theta)}{\max_{\theta\in\Theta} L(x_1,\ldots,x_n,\theta}$$

Paso 3. Obtener la distribución exacta o asintótica del estadístico

Paso 4. Establecer el criterio de decisión

Distribuciones multivariantes

Definición y propiedades básicas

Distribución de Wishart 1/3

Z_{nxp} matriz con filas independientes que se distribuyen según una distribución normal centrada ($Z_i \sim N_p(0,\Sigma)$)

$$Q=Z'Z \sim W_p(\Sigma,n)$$

Función de densidad ($\Sigma>0$)

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

n=1: $Z=(Z_1,\ldots,Z_p)'$, Q se reduce a la suma de cuadrados de normales independientes (generalización de ji-cuadrado)

$$Q=(Z_1,\ldots,Z_p)'(Z_1,\ldots,Z_p)=Z_1^2+\ldots+Z_p^2$$

Distribución de Wishart 2/3

Propiedad 1. $Q_1 \sim W_p(\Sigma, m)$, $Q_2 \sim W_p(\Sigma, n)$, independientes

$$Q_1+Q_2 \sim W_p(\Sigma, m+n)$$

Propiedad 2. $Q \sim W_p(\Sigma, n)$ y se consideran las distribuciones por cajas

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

entonces: $Q_1 \sim Wp1(\Sigma_{11}, n)$ y $Q_2 \sim Wp2(\Sigma_{22}, n)$

Propiedad 3. $Q \sim W_p(\Sigma, n)$ y $T \sim M_{pxq}$, entoces

$$T'QT \sim W_q(T'\Sigma T, n)$$

Distribución de Wishart 3/3

Propiedad 4. $nS \sim W_p(\Sigma, n\text{-}1)$,

$$nS = X'PX = X'P^2X = (XP)'(XP) \approx W_p(\Sigma, n-1)$$

siendo $P = I - \frac{1}{n}11'$

Distribución de T² Hotelling

$Y \sim N_p(0, I)$ y $Q \sim W_p(I, n)$, entonces

$$T^2 = n \cdot Y' \cdot Q^{-1} \cdot Y \approx T^2(p,n)$$

Propiedad 1. $X \sim N_p(\mu, \Sigma)$ y $Q \sim W_p(\Sigma, n)$, entonces

$$T^2 = n \cdot (X-\mu)' \cdot Q^{-1} \cdot (X-\mu) \approx T^2(p,n)$$

Propiedad 2.

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

Distribución Λ de Wilks

Definición. A y B independientes, $A \sim W_p(\Sigma, m)$ y $B \sim W_p(\Sigma, n)$ con $m>p$,

$$\Lambda = \frac{|A|}{|A|+|B|} \approx \Lambda(p,m,n)$$

Propiedad 1. $0 \le \Lambda \le 1$

Propiedad 2. Relación $\Lambda(p,n-q,q)$ y F

$$\frac{ms-2\lambda}{pq}\frac{1-\Lambda^{1/s}}{\Lambda^{1/s}} \xrightarrow{L_{n\to\infty}} F_{pq,ms-2\lambda}$$

$m=n-(p+q+1)/2$, $\lambda=(pq-2)/4$, $s=((p^2q^2-4)/(p^2+q^2-5))^{1/2}$

Propiedad 3. Relación $\Lambda(p,m,n)$ y χ^2

$$\left(\frac{p-n+1}{2}-m\right)\cdot\log\Lambda \xrightarrow{L_{m\to\infty}} \chi^2_{np}$$

Referencias

Anderson, T.W. (2003). An introduction to multivariate statistical analysis. Wiley.

Johnson, R.A. y Wichern, D.W. (1982). Applied multivariate statistical analysis. Prentice-Hall.

Mardia, K.V., Kent, J.T. y Bibby, J.M. (1979). Multivariate analysis. Academic Press.

Muirhead R.J. (1982). Aspects of Multivariate Statistical Theory. Wiley Series in Probability and Statistics.

Peña D (2002) Analisis de datos multivariantes. Mc Graw Hill

Seber, G.A.F. (1984). Multivariate observations. John Wiley & Sons