Forma Canónica del Modelo Ω

Dada una base ortonormal de $\mathcal{V}_r = \mathcal{V}_{\mathbf{X}}$, digamos $\{\boldsymbol{\alpha}_1, \ldots \boldsymbol{\alpha}_r\}$, sabemos que podemos extenderla a una base ortonormal de $\Re^n$: $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \ldots, \boldsymbol{\alpha}_n\}$.

Por lo tanto,

$$\mathbf{y} \in \Re^n : \mathbf{y} = \sum_{j=1}^{n} z_j \boldsymbol{\alpha}_j$$

y tenemos que

$$\boldsymbol{\alpha}_i' \mathbf{y} = \sum_{j=1}^{n} z_j \boldsymbol{\alpha}_i' \boldsymbol{\alpha}_j = z_i \boldsymbol{\alpha}_i' \boldsymbol{\alpha}_i = z_i \quad \forall i = 1, \ldots, n$$

Luego, si definimos a $\mathbf{T}$ como la matriz que tiene filas $\boldsymbol{\alpha}_i'$, entonces

$$\mathbf{z} = \mathbf{T}\mathbf{y}$$

Observemos que

$$E(z_i) = \begin{cases} \boldsymbol{\alpha}_i' \boldsymbol{\eta} = \xi_i & \text{si } 1 \leq i \leq r \\ 0 & \text{si } r+1 \leq i \leq n \end{cases}$$

$$\Sigma_{\mathbf{z}} = \mathbf{T}\Sigma_{\mathbf{y}}\mathbf{T}' = \sigma^2 \mathbf{I}$$

Por lo tanto, ahora podemos reescribir a Ω como

$$\boxed{\begin{array}{l} \Omega : \\ \qquad E(z_i) = \begin{cases} \xi_i & \text{si } 1 \leq i \leq r \\ 0 & \text{si } r+1 \leq i \leq n \end{cases} \\ \qquad \Sigma_{\mathbf{z}} = \sigma^2 \mathbf{I} \end{array}}$$

donde $\boldsymbol{\xi}$ y σ^2 son parámetros desconocidos.

En términos de esta forma caónica es sencillo demostrar que

$$\boxed{s^2 = \frac{\|\mathbf{Y} - \widehat{\mathbf{Y}}\|^2}{n-r} == \frac{\|\mathbf{Y} - \widehat{\boldsymbol{\eta}}\|^2}{n-r}}$$

es un estimador insesgado de σ^2. Sólo habíamos demostrado hasta ahora el caso de rango completo.

Distribución Normal Multivariada

Definición 1: Se dice que un vector $\mathbf{X}$, $k-$dimensional tiene distribución normal multivariada $N_k(\boldsymbol{\mu}, \mathbf{Q})$ donde $\boldsymbol{\mu}$ es un vector $k-$dimensional, $\mathbf{Q}$ una matriz de $k \times k$ definida positiva, si su densidad es de la forma

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^k |\mathbf{Q}|^{1/2}} \, e^{-\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{Q}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2}}$$

donde $|\mathbf{Q}|$ indica determinante de $\mathbf{Q}$.

Por ejemplo, si X_i son k v.a. independientes tales que $X_i \sim N(\mu_i, \sigma_i^2)$, entonces el vector $\mathbf{X}' = (X_1, \ldots, X_k)$ tiene densidad

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^k \, \Pi_{j=1}^k (\sigma_i^2)^{1/2}} \, e^{-1/2\{\Sigma_{i=1}^k (x_i-\mu_i)^2/\sigma_i^2\}}$$

Luego, resulta que $\mathbf{X}$ es $N_k(\boldsymbol{\mu}, \mathbf{Q})$ donde $\boldsymbol{\mu}' = (\mu_1, \ldots, \mu_k)$ y

$$\mathbf{Q} = diag(\sigma_1^2, \ldots, \sigma_k^2) = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_k^2 \end{pmatrix}$$

Más aún, en el caso en que las k v.a. X_i son todas $N(0,1)$, $\mathbf{X}$ es $N(\mathbf{0_k}, \mathbf{I_k})$ donde $\mathbf{0}'_\mathbf{k} = (\mathbf{0}, \ldots, \mathbf{0}) \in \Re^\mathbf{k}$ y $\mathbf{I}_k$ es la matriz identidad de $k \times k$.

Recordemos el enunciado del Teorema de Cambio de Variable:

Sean $\mathbf{x}$ es un vector aleatorio con densidad f y $\mathbf{y} = g(\mathbf{x})$, tal que $g^{-1}(\mathbf{y}) = \mathbf{x}$. Supongamos que en un abierto $\mathcal{G}$ existen las derivadas parciales $\dfrac{\partial x_i}{\partial y_j}$ y sea $J = det \left\{ \dfrac{\partial x_i}{\partial y_j} \right\}$, entonces

$$f_\mathbf{Y}(\mathbf{y}) = f_\mathbf{X}(g^{-1}(\mathbf{y}))|J|$$

Teorema N1: Si $\mathbf{X}$ es un vector aleatorio $k-$dimensional con distribución $N_k(\boldsymbol{\mu}, \mathbf{Q})$, $\mathbf{A}$ es una matriz no singular de $k \times k$ y $\mathbf{b}$ un vector $k-$dimensional, entonces

$$\mathbf{Y} = \mathbf{AX} + \mathbf{b} \quad \text{es} \quad N_k(\mathbf{A}\boldsymbol{\mu} + \mathbf{b},\ \mathbf{AQA}')$$

Teorema N2:

i) Un vector aleatorio $k-$dimensional $\mathbf{X}$ es $N_k(\boldsymbol{\mu}, \mathbf{Q})$ si y sólo si $\mathbf{X} = \mathbf{B}\mathbf{Y} + \boldsymbol{\mu}$, donde $\mathbf{Y}$ es $N_k(\mathbf{0_k}, \mathbf{I_k})$ y $\mathbf{B}$ es una matriz de $k \times k$ no singular tal que $\mathbf{B}\mathbf{B}' = \mathbf{Q}$.

ii) Si $\mathbf{X}$ es $N_k(\boldsymbol{\mu}, \mathbf{Q})$ entonces

$$E(\mathbf{X}) = \boldsymbol{\mu} \quad \text{y} \quad \Sigma_{\mathbf{X}} = \mathbf{Q}$$

Teorema N3: Sea $\mathbf{X}$ un vector aleatorio $k-$dimensional $N_k(\boldsymbol{\mu}, \mathbf{Q})$ y $\mathbf{A}$ una matriz de $h \times k$ con rango h, luego si $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ entonces

$$\mathbf{Y} \sim N_h(A\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\mathbf{Q}\mathbf{A}')$$

Teorema N4: Sea $\mathbf{X}' = (X_1, \ldots, X_k)$ un vector $k-$dimensional con distribución normal multivariada, luego la distribución marginal de cualquier subconjunto de componentes tiene distribución normal multivariada. En particular cada componente es normal.

Demostración: Sea $\mathbf{X}^* = (X_{k_1}, \ldots, X_{k_h})$, $k_1 < k_2 < \ldots < k_h$, luego se tiene que $\mathbf{X}^* = A\mathbf{X}$, donde A es la matriz de $h \times k$ dada por:

$$a_{ij} = \begin{cases} 1 & \text{si } j = k_i \\ 0 & \text{si } j \neq k_i \end{cases}$$

$1 \leq i \leq h$, $1 \leq j \leq k$.

Es fácil ver que A es una matriz de rango h.

Teorema N5: Si $\mathbf{X}$ es un vector $k-$dimensional con distribución $N_k(\boldsymbol{\mu}, \mathbf{Q})$, luego

$$(\mathbf{X} - \boldsymbol{\mu})'\mathbf{Q}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2_k$$

.

Demostración: Por lo ya visto, resulta que $\mathbf{X} = \mathbf{B}\mathbf{Y} + \boldsymbol{\mu}$ donde $\mathbf{Y}$ es $N(\mathbf{0_k}, \mathbf{I_k})$

$$\mathbf{Y} = \mathbf{B}^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

y además

$$\mathbf{B}\mathbf{B}' = \mathbf{Q}$$

Luego

$$\mathbf{Y}\mathbf{Y}' = (\mathbf{X} - \boldsymbol{\mu})'\mathbf{B}'^{-1}\mathbf{B}^{-1}(\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{X} - \boldsymbol{\mu})'\mathbf{Q}^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

El teorema resulta del hecho que

$$\mathbf{Y}'\mathbf{Y} = \sum_{i=1}^{k} Y_i^2$$

tiene distribución χ^2_k, ya que las Y_i son variables aleatorias independientes con distribución $N(0, 1)$.

Teorema N6: Si $\mathbf{X}$ es un vector $k-$dimensional con distribución $N_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ y $\mathbf{P}$ una matriz simétrica e idempotente de rango r, entonces

$$\frac{(\mathbf{X}-\boldsymbol{\mu})'\mathbf{P}(\mathbf{X}-\boldsymbol{\mu})}{\sigma^2} \sim \chi^2_r$$

.

Tests y Regiones de Confianza

Hasta ahora hemos trabajado sólo con las hipótesis Ω. Sin embargo para deducir tests y regiones de confianza con nivel exacto será necesario que hagamos un supuesto adicional: **normalidad conjunta de los errores**

Supondremos que las y_i's se distribuyen conjuntamente según una normal multivariada.

Podremos deducir:

- intervalos de confianza de nivel exacto para funciones paramétricas estimables
- tests de nivel exacto para hipótesis que involucran a los parámetros
- conjuntos o regiones de confianza para la estimación simultánea de más de una función paramétrica estimable.

Nuestro nuevo modelo será:

$$\boxed{\Omega : \quad \mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}) \quad rg(\mathbf{X}) = r \quad \boldsymbol{\beta} \in \Re^p}$$

Observemos que en este caso suponer que $\Sigma_{\mathbf{Y}} = \sigma^2\mathbf{I}$ es equivalente a asumir que las $y_i, 1 \leq i \leq n$, son independientes.

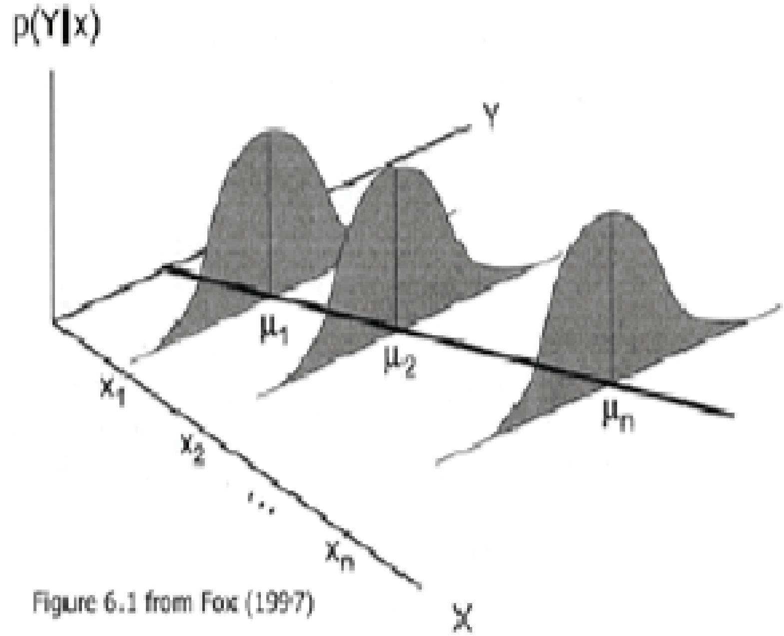

Figure 6.1 from Fox (1997)

Bajo estas condiciones se obtiene el siguiente resultado:

Teorema: Supongamos que se tiene el modelo

$$\Omega : \quad \mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}) \quad \mathrm{rg}(\mathbf{X}) = p \quad \boldsymbol{\beta} \in \Re^p \,.$$

Luego, $\widehat{\boldsymbol{\beta}}$ y s^2 son funciones de estadísticos suficientes y completos y por lo tanto, $\widehat{\boldsymbol{\beta}}$ y s^2 son estimadores IMVU de $\boldsymbol{\beta}$ y σ^2, respectivamente.

Si nuestro modelo es

$$E(\mathbf{Y}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p$$

podríamos tener interés en testear hipótesis como las que siguen:

$$\begin{aligned} H_o : \beta_j = 0 &\ \text{ vs. } \ H_1 : \beta_j \neq 0 \\ H_o : \beta_1 - \beta_2 = 0 &\ \text{ vs. } \ H_1 : \beta_1 - \beta_2 \neq 0 \\ H_o : \beta_1 = \beta_2 = \ldots = \beta_p = 0 &\ \text{ vs. } \ H_1 : \text{existe } j : \beta_j \neq 0 \end{aligned}$$

Todas estas hipótesis son de la forma $\mathbf{c}'\boldsymbol{\beta} = 0$ o $\mathbf{C}\boldsymbol{\beta} = \mathbf{0}$.

Supongamos que tenemos q funciones estimables $\psi_1, \psi_2, \ldots, \psi_q$ donde:

$$\psi_i = \sum_{j=1}^{p} c_{ij}\beta_j \quad 1 \leq i \leq q$$

Por ser estimables, por el Teorema de Gauss–Markov tenemos que

$$\widehat{\psi_i} = \sum_{j=1}^{n} a_{ij}^* y_j \quad 1 \leq i \leq q\,,$$

donde $\mathbf{a}_i^* \in \mathcal{V}_r \subset \Re^n$; de manera que

$$\begin{aligned} \Psi &= \mathbf{C}\boldsymbol{\beta} \qquad \mathbf{C} \in \Re^{q\times p} \\ \widehat{\Psi} &= \mathbf{A}^*\mathbf{Y} \qquad \mathbf{A}^* \in \Re^{q\times n} \end{aligned}$$

Más aún, sabemos que

$$\begin{aligned} \widehat{\Psi} &= \mathbf{C}\widehat{\boldsymbol{\beta}} \\ \Sigma_{\widehat{\Psi}} &= \sigma^2 \mathbf{A}^*\mathbf{A}^{*'} \end{aligned}$$

Estimamos a σ^2 por

$$s^2 = \frac{\|\mathbf{Y} - \widehat{\mathbf{Y}}\|^2}{n-r}$$

Bajo estas nuevas hipótesis obtenemos el siguiente resultado:

Teorema: Supongamos que se cumple Ω, es decir $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, $\mathrm{rg}(\mathbf{X}) = r$, $\boldsymbol{\beta} \in \Re^p$ y que además que $\psi_1, \psi_2, \ldots, \psi_q$ son q funciones estimables l.i., de manera que $\mathrm{rg}(\mathbf{C}) = q$. Entonces,

i) $\widehat{\Psi} \sim N_q(\Psi, \Sigma_{\widehat{\Psi}})$ (o lo que es igual $N_q(\Psi, \sigma^2\mathbf{A}^*\mathbf{A}^{*\prime})$)

ii) $\widehat{\Psi}$ y $\dfrac{s^2(n-r)}{\sigma^2}$ son independientes

iii) $\dfrac{(n-r)s^2}{\sigma^2} \sim \chi^2_{n-r}$

En el caso de rango completo, es decir cuando $r = p$, obtenemos el siguiente resultado:

Teorema: Supongamos que se cumple Ω, es decir $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, $\mathrm{rg}(\mathbf{X}) = p$, $\boldsymbol{\beta} \in \Re^p$. Entonces,

i) $\widehat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$

ii) $\dfrac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\mathbf{X}'\mathbf{X})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{\sigma^2} \sim \chi^2_p$

iii) $\widehat{\boldsymbol{\beta}}$ y $\dfrac{(n-p)s^2}{\sigma^2}$ son independientes

iv) $\dfrac{(n-p)s^2}{\sigma^2} \sim \chi^2_{n-p}$

Estos resultados nos permiten deducir intervalos de confianza o tests para cada uno de los coeficientes del modelo lineal: