Modelo Lineal General (GLM)
vs
Modelo Lineal Mixto (LMM)

Valentín Pando Fernández

Departamento de Estadística e Investigación Operativa
Universidad de Valladolid

DISTRIBUCIÓN

NORMAL

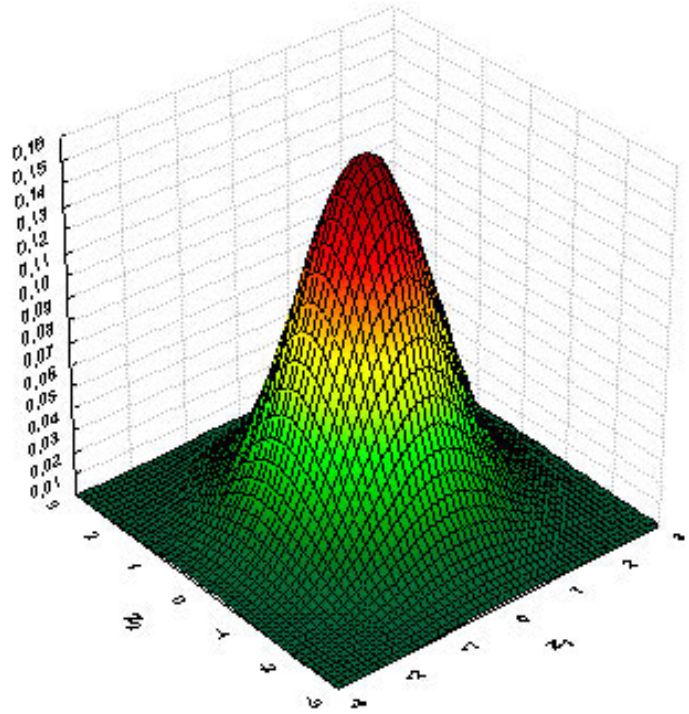
n-DIMENSIONAL

DISTRIBUCIÓN NORMAL ESTÁNDAR I

Dado vector aleatorio $\mathbf{z} = (z_1, z_2, ..., z_n)' \in \mathbb{R}^n$ diremos que $\mathbf{z}$ tiene una distribución normal n-dimensional estándar si su función de densidad es

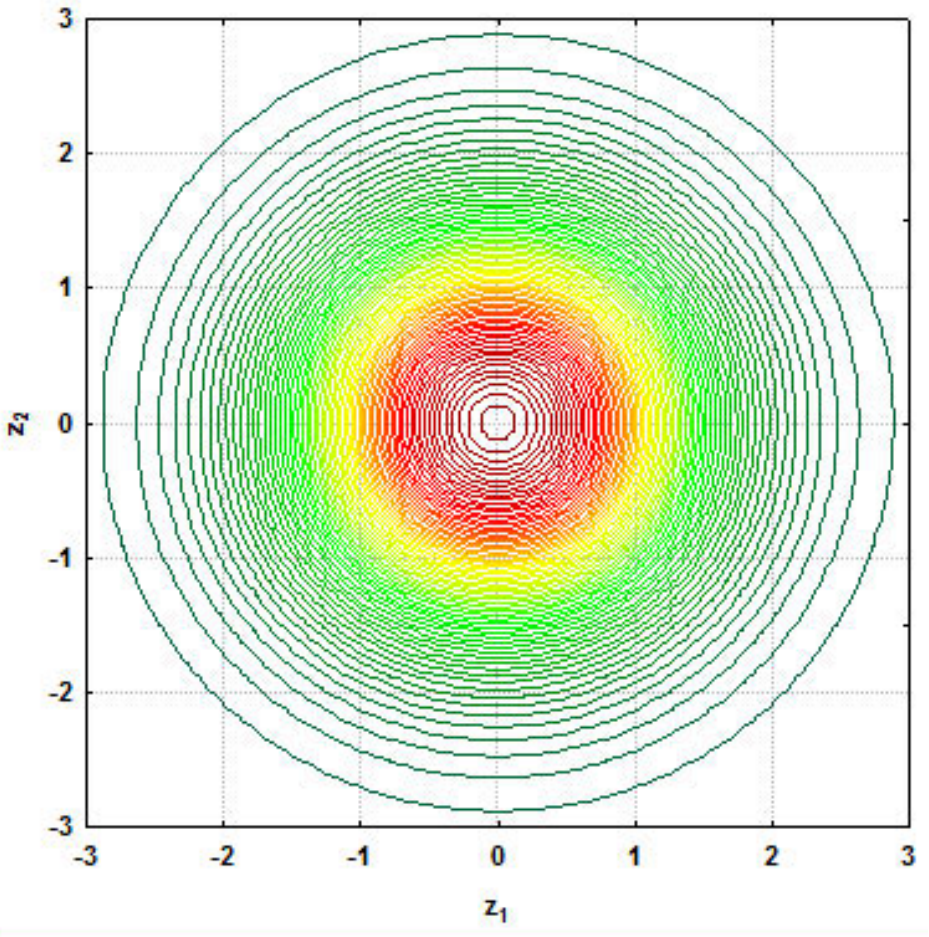
$$f(\mathbf{z}) = \frac{\exp\left[-0.5\mathbf{z}'\mathbf{z}\right]}{(2\pi)^{n/2}} = \frac{\exp\left[-0.5\sum_{i=1}^{n} z_i^2\right]}{(2\pi)^{n/2}}$$

Gráficamente:



DISTRIBUCIÓN NORMAL ESTÁNDAR II

Las curvas de nivel son:



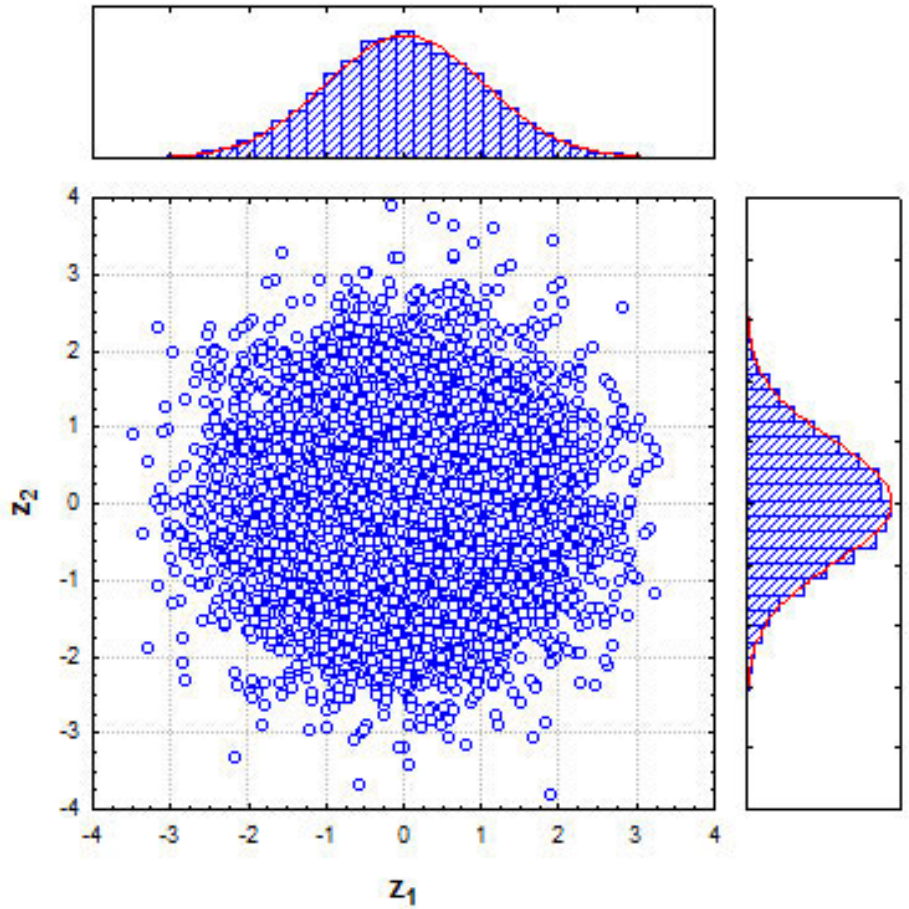
Se verifica: $E(\mathbf{z}) = \mathbf{0}_n$ y $Var(\mathbf{z}) = I_n$, con I_n matriz identidad de orden n

Escribiremos: $\mathbf{z} \rightsquigarrow N_n(\mathbf{0}_n, I_n)$

DISTRIBUCIÓN NORMAL ESTÁNDAR III

$$\mathbf{z} \rightsquigarrow N_n(\mathbf{0}_n, I_n) \Leftrightarrow z_i \rightsquigarrow N(0,1) \quad \forall i = 1, 2, ..., n \text{ e independientes}$$

El siguiente gráfico representa una muestra aleatoria simple de una distribución $\mathbf{z} \rightsquigarrow N_2(\mathbf{0}_2, I_2)$ con 10000 observaciones:



DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\mu = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \mu$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\mu, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que **y** tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de **y** es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\mu = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \mu$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\mu, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que **y** tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de **y** es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)
  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$, existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$ (raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$ (raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL I

- Si $\mathbf{y} = (y_1, y_2, ..., y_n)' \in \mathbb{R}^n$ diremos que $\mathbf{y}$ tiene distribución normal n-dimensional si se verifican las siguientes condiciones equivalentes:
  - Si $\mathbf{u} \in \mathbb{R}^n$, con $\mathbf{u} \neq \mathbf{0_n}$, $\mathbf{u}'\mathbf{y}$ tiene distribución de probabilidad normal (Cualquier combinación lineal de las componentes de $\mathbf{y}$ es normal)
  - Existe $\boldsymbol{\mu} \in \mathbb{R}^n$ y $A \in M_{n\times m}$ tal que $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z}$ con $\mathbf{z} \rightsquigarrow N_m(\mathbf{0}_m, I_m)$ y $m \leq n$
- Si $rg(A) = n$ es una distribución normal n-dimensional no degenerada

  Si $rg(A) < n$ es una distribución normal n-dimensional degenerada
- Si $\boldsymbol{\mu} = \mathbf{0}_n$ y $A = I_n$ entonces tenemos distribución normal estándar
- Se verifica que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $Var(\mathbf{y}) = \Sigma = AA' \in M_{n\times n}$. Por ello usaremos la notación $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$
- Si $rg(A) = n$ la distribución es no degenerada y se verifica que: $|\Sigma| \neq 0$,existe Σ^{-1}, existe una matriz $\Sigma^{1/2}$ con $\Sigma = \left(\Sigma^{1/2}\right)' \Sigma^{1/2}$ y existe $\Sigma^{-1/2}$ con $\Sigma^{-1/2}\left(\Sigma^{-1/2}\right)' = \Sigma^{-1}$(raíz de Cholesky)

  Observar que en este caso $A = \left(\Sigma^{1/2}\right)'$

DISTRIBUCIÓN NORMAL GENERAL II

- Si $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA'$ y $T \in M_{t\times n}$ no nula, entonces $T\mathbf{y} = T\boldsymbol{\mu} + TA\mathbf{z} \in \mathbb{R}^t$ tiene distribución normal con
$$T\mathbf{y} \rightsquigarrow N_t(T\boldsymbol{\mu}, T\Sigma T')$$
- Si y_i =componente i-ésima del vector $\mathbf{y}$, μ_i =componente i-ésima del vector $\boldsymbol{\mu}$ y $T = (0, ..0, 1^{(i)}, 0, .., 0)$, entonces $y_i = T\mathbf{y} \rightsquigarrow N(\mu_i, \sigma_i^2)$ siendo σ_i^2 el elemento diagonal i-ésimo de la matriz Σ
- Si $\mathbf{y} \rightsquigarrow N_n(\mu, \Sigma)$ no degenerada, entonces $\mathbf{z} = \left(\Sigma^{-1/2}\right)'(\mathbf{y} - \boldsymbol{\mu})$ y la función de densidad de $\mathbf{y}$ es:
$$f(\mathbf{y}) = \frac{\exp\left[-0.5(\mathbf{y} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{y} - \boldsymbol{\mu})\right]}{(2\pi)^{n/2}|\Sigma|^{1/2}}$$
- $(\mathbf{y} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{y} - \boldsymbol{\mu})$ es el cuadrado de la distancia de Mahalanobis entre $\mathbf{y}$ y $\boldsymbol{\mu}$. Si $\Sigma = I_n$ tenemos la distancia euclídea habitual
- Si el vector normal es degenerado entonces $|\Sigma| = 0$ y no tenemos función de densidad para el vector

DISTRIBUCIÓN NORMAL GENERAL II

- Si $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA'$ y $T \in M_{t\times n}$ no nula, entonces $T\mathbf{y} = T\boldsymbol{\mu} + TA\mathbf{z} \in \mathbb{R}^t$ tiene distribución normal con
$$T\mathbf{y} \rightsquigarrow N_t(T\boldsymbol{\mu}, T\Sigma T')$$
- Si y_i =componente i-ésima del vector $\mathbf{y}$, μ_i =componente i-ésima del vector $\boldsymbol{\mu}$ y $T = (0, ..0, 1^{(i)}, 0, .., 0)$, entonces $y_i = T\mathbf{y} \rightsquigarrow N(\mu_i, \sigma_i^2)$ siendo σ_i^2 el elemento diagonal i-ésimo de la matriz Σ
- Si $\mathbf{y} \rightsquigarrow N_n(\mu, \Sigma)$ no degenerada, entonces $\mathbf{z} = \left(\Sigma^{-1/2}\right)'(\mathbf{y} - \mu)$ y la función de densidad de $\mathbf{y}$ es:
$$f(\mathbf{y}) = \frac{\exp\left[-0.5(\mathbf{y} - \mu)'\Sigma^{-1}(\mathbf{y} - \mu)\right]}{(2\pi)^{n/2}|\Sigma|^{1/2}}$$
- $(\mathbf{y} - \mu)'\Sigma^{-1}(\mathbf{y} - \mu)$ es el cuadrado de la distancia de Mahalanobis entre $\mathbf{y}$ y μ. Si $\Sigma = I_n$ tenemos la distancia euclídea habitual
- Si el vector normal es degenerado entonces $|\Sigma| = 0$ y no tenemos función de densidad para el vector

DISTRIBUCIÓN NORMAL GENERAL II

- Si $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA'$ y $T \in M_{t\times n}$ no nula, entonces $T\mathbf{y} = T\boldsymbol{\mu} + TA\mathbf{z} \in \mathbb{R}^t$ tiene distribución normal con
$$T\mathbf{y} \rightsquigarrow N_t(T\boldsymbol{\mu}, T\Sigma T')$$
- Si y_i =componente i-ésima del vector $\mathbf{y}$, μ_i =componente i-ésima del vector $\boldsymbol{\mu}$ y $T = (0, ..0, 1^{(i)}, 0, .., 0)$, entonces $y_i = T\mathbf{y} \rightsquigarrow N(\mu_i, \sigma_i^2)$ siendo σ_i^2 el elemento diagonal i-ésimo de la matriz Σ
- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada, entonces $\mathbf{z} = \left(\Sigma^{-1/2}\right)'(\mathbf{y} - \boldsymbol{\mu})$ y la función de densidad de $\mathbf{y}$ es:
$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y} - \boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y} - \boldsymbol{\mu})\right]}{(2\pi)^{n/2}\,|\Sigma|^{1/2}}$$
- $(\mathbf{y} - \mu)'\,\Sigma^{-1}\,(\mathbf{y} - \mu)$ es el cuadrado de la distancia de Mahalanobis entre $\mathbf{y}$ y μ. Si $\Sigma = I_n$ tenemos la distancia euclídea habitual
- Si el vector normal es degenerado entonces $|\Sigma| = 0$ y no tenemos función de densidad para el vector

DISTRIBUCIÓN NORMAL GENERAL II

- Si $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA'$ y $T \in M_{t\times n}$ no nula, entonces $T\mathbf{y} = T\boldsymbol{\mu} + TA\mathbf{z} \in \mathbb{R}^t$ tiene distribución normal con
$$T\mathbf{y} \rightsquigarrow N_t(T\boldsymbol{\mu}, T\Sigma T')$$
- Si y_i =componente i-ésima del vector $\mathbf{y}$, μ_i =componente i-ésima del vector $\boldsymbol{\mu}$ y $T = (0, ..0, 1^{(i)}, 0, .., 0)$, entonces $y_i = T\mathbf{y} \rightsquigarrow N(\mu_i, \sigma_i^2)$ siendo σ_i^2 el elemento diagonal i-ésimo de la matriz Σ
- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada, entonces $\mathbf{z} = \left(\Sigma^{-1/2}\right)'(\mathbf{y} - \boldsymbol{\mu})$ y la función de densidad de $\mathbf{y}$ es:
$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y} - \boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y} - \boldsymbol{\mu})\right]}{(2\pi)^{n/2}\,|\Sigma|^{1/2}}$$
- $(\mathbf{y} - \boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y} - \boldsymbol{\mu})$ es el cuadrado de la distancia de Mahalanobis entre $\mathbf{y}$ y $\boldsymbol{\mu}$. Si $\Sigma = I_n$ tenemos la distancia euclídea habitual
- Si el vector normal es degenerado entonces $|\Sigma| = 0$ y no tenemos función de densidad para el vector

DISTRIBUCIÓN NORMAL GENERAL II

- Si $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{z} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA'$ y $T \in M_{t\times n}$ no nula, entonces $T\mathbf{y} = T\boldsymbol{\mu} + TA\mathbf{z} \in \mathbb{R}^t$ tiene distribución normal con
$$T\mathbf{y} \rightsquigarrow N_t(T\boldsymbol{\mu}, T\Sigma T')$$
- Si y_i =componente i-ésima del vector $\mathbf{y}$, μ_i =componente i-ésima del vector $\boldsymbol{\mu}$ y $T = (0, ..0, 1^{(i)}, 0, .., 0)$, entonces $y_i = T\mathbf{y} \rightsquigarrow N(\mu_i, \sigma_i^2)$ siendo σ_i^2 el elemento diagonal i-ésimo de la matriz Σ
- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada, entonces $\mathbf{z} = \left(\Sigma^{-1/2}\right)'(\mathbf{y} - \boldsymbol{\mu})$ y la función de densidad de $\mathbf{y}$ es:
$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y} - \boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y} - \boldsymbol{\mu})\right]}{(2\pi)^{n/2}\,|\Sigma|^{1/2}}$$
- $(\mathbf{y} - \boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y} - \boldsymbol{\mu})$ es el cuadrado de la distancia de Mahalanobis entre $\mathbf{y}$ y $\boldsymbol{\mu}$. Si $\Sigma = I_n$ tenemos la distancia euclídea habitual
- Si el vector normal es degenerado entonces $|\Sigma| = 0$ y no tenemos función de densidad para el vector

DISTRIBUCIÓN NORMAL NO DEGENERADA I

Sea $\mathbf{y} = (y_1, y_2)' = \boldsymbol{\mu} + A\mathbf{z}$ con $\boldsymbol{\mu} = (10, 15)'$, $A = \begin{pmatrix} 2 & 0 \\ 3/2 & 3\sqrt{3}/2 \end{pmatrix}$ y $\mathbf{z} \rightsquigarrow N_2(\mathbf{0}_2, I_2)$. Entonces $\mathbf{y} \rightsquigarrow N_2(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA' = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$

En este caso $|\Sigma| = 27$ y $\Sigma^{-1} = \frac{1}{27}\begin{pmatrix} 9 & -3 \\ -3 & 4 \end{pmatrix}$

Por tanto, la función de densidad del vector $\mathbf{y}$ es

$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y}-\boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y}-\boldsymbol{\mu})\right]}{2\pi\,|\Sigma|^{1/2}}$$

$$= \frac{\exp\left[-\frac{0.5}{27}\left[9\,(y_1-10)^2 + 4\,(y_2-15)^2 - 6\,(y_1-10)\,(y_2-15)\right]\right]}{2\pi\sqrt{27}}$$

La representación gráfica tridimensional y con curvas de nivel de esta función es:

DISTRIBUCIÓN NORMAL NO DEGENERADA I

Sea $\mathbf{y} = (y_1, y_2)' = \boldsymbol{\mu} + A\mathbf{z}$ con $\boldsymbol{\mu} = (10, 15)'$, $A = \begin{pmatrix} 2 & 0 \\ 3/2 & 3\sqrt{3}/2 \end{pmatrix}$ y

$\mathbf{z} \rightsquigarrow N_2(\mathbf{0}_2, I_2)$. Entonces $\mathbf{y} \rightsquigarrow N_2(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA' = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$

En este caso $|\Sigma| = 27$ y $\Sigma^{-1} = \frac{1}{27}\begin{pmatrix} 9 & -3 \\ -3 & 4 \end{pmatrix}$

Por tanto, la función de densidad del vector $\mathbf{y}$ es

$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y}-\boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y}-\boldsymbol{\mu})\right]}{2\pi\,|\Sigma|^{1/2}}$$

$$= \frac{\exp\left[-\frac{0.5}{27}\left[9\,(y_1-10)^2 + 4\,(y_2-15)^2 - 6\,(y_1-10)\,(y_2-15)\right]\right]}{2\pi\sqrt{27}}$$

La representación gráfica tridimensional y con curvas de nivel de esta función es:

DISTRIBUCIÓN NORMAL NO DEGENERADA I

Sea $\mathbf{y} = (y_1, y_2)' = \boldsymbol{\mu} + A\mathbf{z}$ con $\boldsymbol{\mu} = (10, 15)'$, $A = \begin{pmatrix} 2 & 0 \\ 3/2 & 3\sqrt{3}/2 \end{pmatrix}$ y

$\mathbf{z} \rightsquigarrow N_2(\mathbf{0}_2, I_2)$. Entonces $\mathbf{y} \rightsquigarrow N_2(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA' = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$

En este caso $|\Sigma| = 27$ y $\Sigma^{-1} = \frac{1}{27}\begin{pmatrix} 9 & -3 \\ -3 & 4 \end{pmatrix}$

Por tanto, la función de densidad del vector $\mathbf{y}$ es

$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y}-\boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y}-\boldsymbol{\mu})\right]}{2\pi\,|\Sigma|^{1/2}}$$

$$= \frac{\exp\left[-\frac{0.5}{27}\left[9\,(y_1-10)^2 + 4\,(y_2-15)^2 - 6\,(y_1-10)\,(y_2-15)\right]\right]}{2\pi\sqrt{27}}$$

La representación gráfica tridimensional y con curvas de nivel de esta función es:

DISTRIBUCIÓN NORMAL NO DEGENERADA I

Sea $\mathbf{y} = (y_1, y_2)' = \boldsymbol{\mu} + A\mathbf{z}$ con $\boldsymbol{\mu} = (10, 15)'$, $A = \begin{pmatrix} 2 & 0 \\ 3/2 & 3\sqrt{3}/2 \end{pmatrix}$ y

$\mathbf{z} \rightsquigarrow N_2(\mathbf{0}_2, I_2)$. Entonces $\mathbf{y} \rightsquigarrow N_2(\boldsymbol{\mu}, \Sigma)$ con $\Sigma = AA' = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$

En este caso $|\Sigma| = 27$ y $\Sigma^{-1} = \frac{1}{27}\begin{pmatrix} 9 & -3 \\ -3 & 4 \end{pmatrix}$

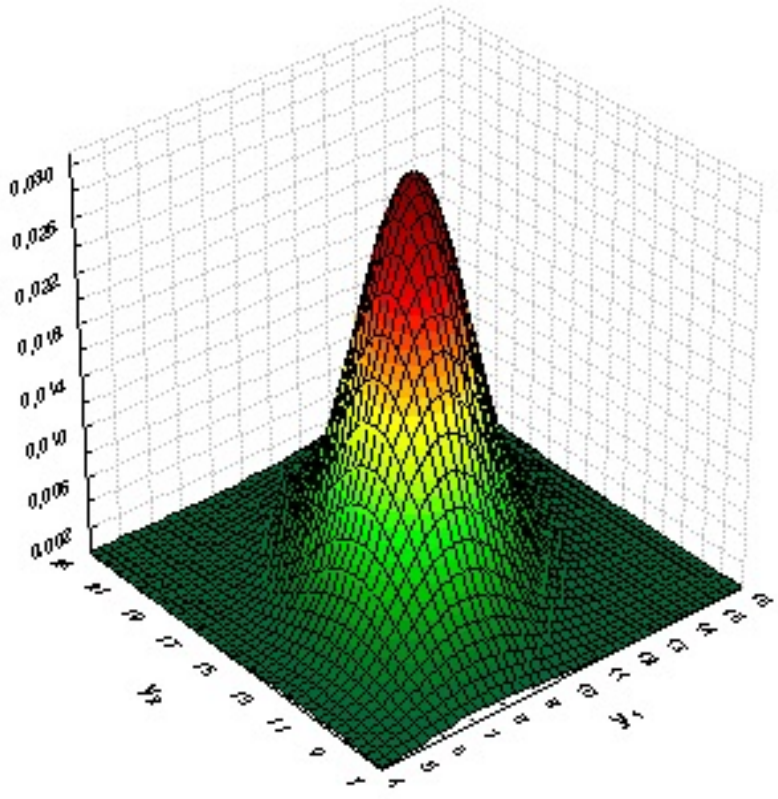
Por tanto, la función de densidad del vector $\mathbf{y}$ es

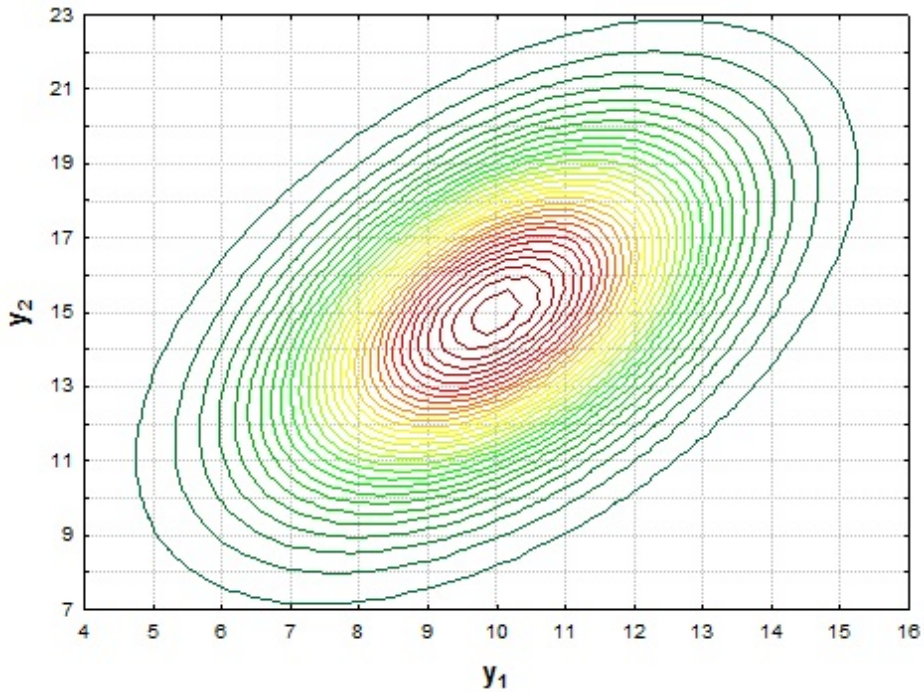
$$f(\mathbf{y}) = \frac{\exp\left[-0.5\,(\mathbf{y}-\boldsymbol{\mu})'\,\Sigma^{-1}\,(\mathbf{y}-\boldsymbol{\mu})\right]}{2\pi\,|\Sigma|^{1/2}}$$

$$= \frac{\exp\left[-\frac{0.5}{27}\left[9\,(y_1-10)^2 + 4\,(y_2-15)^2 - 6\,(y_1-10)\,(y_2-15)\right]\right]}{2\pi\sqrt{27}}$$

La representación gráfica tridimensional y con curvas de nivel de esta función es:

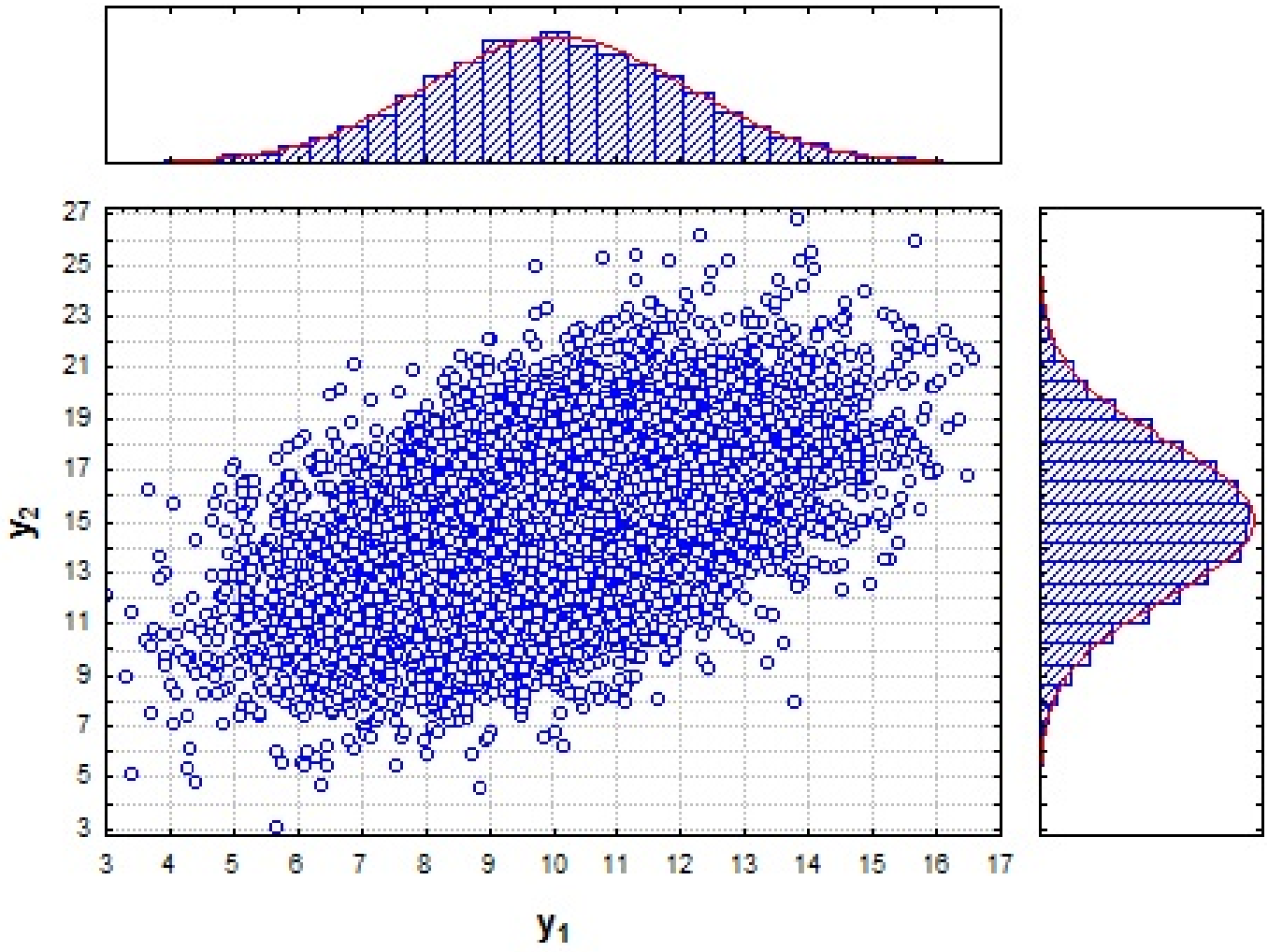
DISTRIBUCIÓN NORMAL NO DEGENERADA II





DISTRIBUCIÓN NORMAL NO DEGENERADA III

El siguiente gráfico representa una muestra aleatoria simple con 10000 observaciones para este vector aleatorio no degenerado $\mathbf{y} \rightsquigarrow N_2(\boldsymbol{\mu}, \Sigma)$



DISTRIBUCIÓN NORMAL DEGENERADA I

Sea $\mathbf{y}=(y_1,y_2,y_3)'=\boldsymbol{\mu}+A\mathbf{z}$ con $\boldsymbol{\mu}=(10,15,13)'$, $A=\begin{pmatrix} 2 & 0 & 1 \\ 3/2 & 3\sqrt{3}/2 & 0 \\ 4 & 0 & 2 \end{pmatrix}$

y $\mathbf{z} \rightsquigarrow N_3(\mathbf{0}_3, I_3)$

En este caso $\Sigma = AA' = \begin{pmatrix} 5 & 3 & 10 \\ 3 & 9 & 6 \\ 10 & 6 & 20 \end{pmatrix}$ y $|\Sigma| = 0$

Por tanto $\mathbf{y} \rightsquigarrow N_3(\mu, \Sigma)$ degenerada

Observar que $rg(A) = rg(\Sigma) = 2$ e $y_3 = 13 + 4z_1 + 2z_3 = -7 + 2y_1$. Por tanto $\mathbf{y}=(y_1,y_2,y_3)'=(y_1,y_2,-7+2y_1)'$ y el vector $\mathbf{y}$ queda definido por el

vector 2-dimensional $(y_1,y_2)' \rightsquigarrow N_2\left((10,15)', \begin{pmatrix} 5 & 3 \\ 3 & 9 \end{pmatrix}\right)$

También podríamos usar el vector $(y_2,y_3)'$ pero no el vector $(y_1,y_3)'$

DISTRIBUCIÓN NORMAL DEGENERADA I

Sea $\mathbf{y}=(y_1,y_2,y_3)'=\boldsymbol{\mu}+A\mathbf{z}$ con $\boldsymbol{\mu}=(10,15,13)'$, $A=\begin{pmatrix} 2 & 0 & 1 \\ 3/2 & 3\sqrt{3}/2 & 0 \\ 4 & 0 & 2 \end{pmatrix}$

y $\mathbf{z} \rightsquigarrow N_3(\mathbf{0}_3, I_3)$

En este caso $\Sigma = AA' = \begin{pmatrix} 5 & 3 & 10 \\ 3 & 9 & 6 \\ 10 & 6 & 20 \end{pmatrix}$ y $|\Sigma| = 0$

Por tanto $\mathbf{y} \rightsquigarrow N_3(\boldsymbol{\mu}, \Sigma)$ degenerada

Observar que $rg(A) = rg(\Sigma) = 2$ e $y_3 = 13 + 4z_1 + 2z_3 = -7 + 2y_1$. Por tanto $\mathbf{y}=(y_1,y_2,y_3)'=(y_1,y_2,-7+2y_1)'$ y el vector $\mathbf{y}$ queda definido por el

vector 2-dimensional $(y_1,y_2)' \rightsquigarrow N_2\left((10,15)', \begin{pmatrix} 5 & 3 \\ 3 & 9 \end{pmatrix}\right)$

También podríamos usar el vector $(y_2,y_3)'$ pero no el vector $(y_1,y_3)'$

DISTRIBUCIÓN NORMAL DEGENERADA I

Sea $\mathbf{y}=(y_1,y_2,y_3)'=\boldsymbol{\mu}+A\mathbf{z}$ con $\boldsymbol{\mu}=(10,15,13)'$, $A=\begin{pmatrix} 2 & 0 & 1 \\ 3/2 & 3\sqrt{3}/2 & 0 \\ 4 & 0 & 2 \end{pmatrix}$

y $\mathbf{z} \rightsquigarrow N_3(\mathbf{0}_3, I_3)$

En este caso $\Sigma = AA' = \begin{pmatrix} 5 & 3 & 10 \\ 3 & 9 & 6 \\ 10 & 6 & 20 \end{pmatrix}$ y $|\Sigma| = 0$

Por tanto $\mathbf{y} \rightsquigarrow N_3(\boldsymbol{\mu}, \Sigma)$ degenerada

Observar que $rg(A) = rg(\Sigma) = 2$ e $y_3 = 13 + 4z_1 + 2z_3 = -7 + 2y_1$. Por tanto $\mathbf{y}=(y_1,y_2,y_3)'=(y_1,y_2,-7+2y_1)'$ y el vector $\mathbf{y}$ queda definido por el vector 2-dimensional $(y_1,y_2)' \rightsquigarrow N_2\left((10,15)', \begin{pmatrix} 5 & 3 \\ 3 & 9 \end{pmatrix}\right)$

También podríamos usar el vector $(y_2,y_3)'$ pero no el vector $(y_1,y_3)'$

DISTRIBUCIÓN NORMAL DEGENERADA I

Sea $\mathbf{y}=(y_1,y_2,y_3)'=\boldsymbol{\mu}+A\mathbf{z}$ con $\boldsymbol{\mu}=(10,15,13)'$, $A=\begin{pmatrix} 2 & 0 & 1 \\ 3/2 & 3\sqrt{3}/2 & 0 \\ 4 & 0 & 2 \end{pmatrix}$

y $\mathbf{z} \rightsquigarrow N_3(\mathbf{0}_3, I_3)$

En este caso $\Sigma = AA' = \begin{pmatrix} 5 & 3 & 10 \\ 3 & 9 & 6 \\ 10 & 6 & 20 \end{pmatrix}$ y $|\Sigma| = 0$

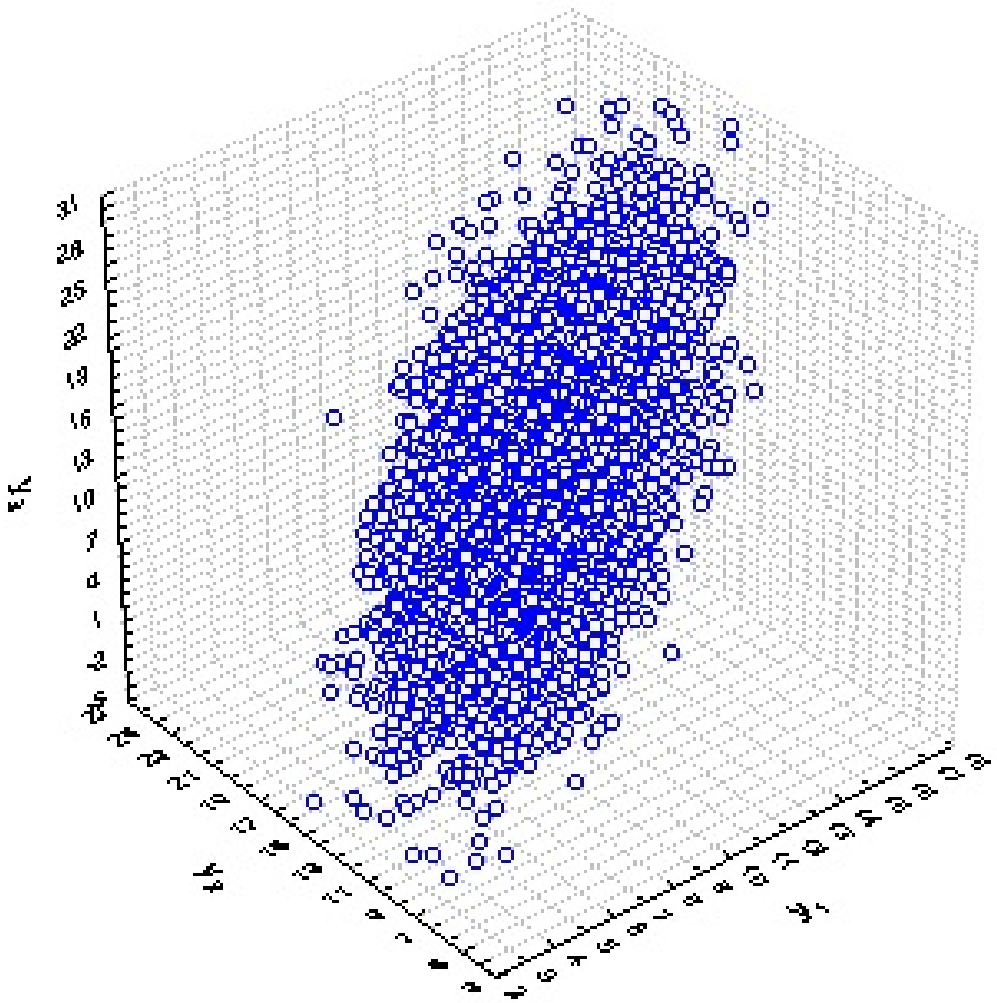
Por tanto $\mathbf{y} \rightsquigarrow N_3(\boldsymbol{\mu}, \Sigma)$ degenerada

Observar que $rg(A) = rg(\Sigma) = 2$ e $y_3 = 13 + 4z_1 + 2z_3 = -7 + 2y_1$. Por tanto $\mathbf{y}=(y_1,y_2,y_3)'=(y_1,y_2,-7+2y_1)'$ y el vector $\mathbf{y}$ queda definido por el vector 2-dimensional $(y_1,y_2)' \rightsquigarrow N_2\left((10,15)', \begin{pmatrix} 5 & 3 \\ 3 & 9 \end{pmatrix}\right)$

También podríamos usar el vector $(y_2,y_3)'$ pero no el vector $(y_1,y_3)'$

DISTRIBUCIÓN NORMAL DEGENERADA II

El siguiente gráfico representa una muestra aleatoria simple con 10000 observaciones para este vector aleatorio degenerado $\mathbf{y} \rightsquigarrow N_3(\boldsymbol{\mu}, \Sigma)$



PROPIEDADES DE LA DISTRIBUCIÓN NORMAL I

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$, $P_1 \in M_{p_1 \times n}$, $P_2 \in M_{p_2 \times n}$ y $p_1, p_2 \leq n$ entonces:

$$\begin{pmatrix} P_1\mathbf{y} \\ P_2\mathbf{y} \end{pmatrix} \rightsquigarrow N_{p_1+p_2}\left(\begin{pmatrix} P_1\boldsymbol{\mu} \\ P_2\boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} P_1\Sigma P_1' & P_1\Sigma P_2' \\ P_2\Sigma P_1' & P_2\Sigma P_2' \end{pmatrix}\right)$$

- Sean $\mathbf{y} \rightsquigarrow N_n(\mu, \Sigma)$ no degenerada, $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^n$ y $P_1, P_2 \in M_{n \times n}$ matrices simétricas. Entonces:
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\boldsymbol{u}_2'\mathbf{y}$ son independientes si y sólo si $\boldsymbol{u}_1'\Sigma\boldsymbol{u}_2 = 0$
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\mathbf{y}'P_1\mathbf{y}$ son independientes si y sólo si $P_1\Sigma\boldsymbol{u}_1 = \mathbf{0}_n$
  - $P_1\mathbf{y}$ y $P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$
  - $\mathbf{y}'P_1\mathbf{y}$ y $\mathbf{y}'P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada entonces $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_n$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n \times n}$ simétrica con $rg(P) = p$. Entonces: $(\mathbf{y} - \boldsymbol{\mu})' P (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$ si y sólo si $P\Sigma$ es idempotente

  (es decir $(P\Sigma)(P\Sigma) = P\Sigma$)

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL I

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$, $P_1 \in M_{p_1 \times n}$, $P_2 \in M_{p_2 \times n}$ y $p_1, p_2 \leq n$ entonces:

$$\begin{pmatrix} P_1\mathbf{y} \\ P_2\mathbf{y} \end{pmatrix} \rightsquigarrow N_{p_1+p_2}\left(\begin{pmatrix} P_1\boldsymbol{\mu} \\ P_2\boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} P_1\Sigma P_1' & P_1\Sigma P_2' \\ P_2\Sigma P_1' & P_2\Sigma P_2' \end{pmatrix}\right)$$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada, $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^n$ y $P_1, P_2 \in M_{n \times n}$ matrices simétricas. Entonces:
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\boldsymbol{u}_2'\mathbf{y}$ son independientes si y sólo si $\boldsymbol{u}_1'\Sigma\boldsymbol{u}_2 = 0$
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\mathbf{y}'P_1\mathbf{y}$ son independientes si y sólo si $P_1\Sigma\boldsymbol{u}_1 = \mathbf{0}_n$
  - $P_1\mathbf{y}$ y $P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$
  - $\mathbf{y}'P_1\mathbf{y}$ y $\mathbf{y}'P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada entonces $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_n$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n \times n}$ simétrica con $rg(P) = p$. Entonces: $(\mathbf{y} - \boldsymbol{\mu})' P (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$ si y sólo si $P\Sigma$ es idempotente

  (es decir $(P\Sigma)(P\Sigma) = P\Sigma$)

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL I

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$, $P_1 \in M_{p_1 \times n}$, $P_2 \in M_{p_2 \times n}$ y $p_1, p_2 \leq n$ entonces:

$$\begin{pmatrix} P_1\mathbf{y} \\ P_2\mathbf{y} \end{pmatrix} \rightsquigarrow N_{p_1+p_2}\left(\begin{pmatrix} P_1\boldsymbol{\mu} \\ P_2\boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} P_1\Sigma P_1' & P_1\Sigma P_2' \\ P_2\Sigma P_1' & P_2\Sigma P_2' \end{pmatrix}\right)$$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada, $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^n$ y $P_1, P_2 \in M_{n \times n}$ matrices simétricas. Entonces:
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\boldsymbol{u}_2'\mathbf{y}$ son independientes si y sólo si $\boldsymbol{u}_1'\Sigma\boldsymbol{u}_2 = 0$
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\mathbf{y}'P_1\mathbf{y}$ son independientes si y sólo si $P_1\Sigma\boldsymbol{u}_1 = \mathbf{0}_n$
  - $P_1\mathbf{y}$ y $P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$
  - $\mathbf{y}'P_1\mathbf{y}$ y $\mathbf{y}'P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada entonces $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_n$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n \times n}$ simétrica con $rg(P) = p$. Entonces: $(\mathbf{y} - \boldsymbol{\mu})' P (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$ si y sólo si $P\Sigma$ es idempotente

  (es decir $(P\Sigma)(P\Sigma) = P\Sigma$)

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL I

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$, $P_1 \in M_{p_1 \times n}$, $P_2 \in M_{p_2 \times n}$ y $p_1, p_2 \leq n$ entonces:

$$\begin{pmatrix} P_1\mathbf{y} \\ P_2\mathbf{y} \end{pmatrix} \rightsquigarrow N_{p_1+p_2}\left(\begin{pmatrix} P_1\boldsymbol{\mu} \\ P_2\boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} P_1\Sigma P_1' & P_1\Sigma P_2' \\ P_2\Sigma P_1' & P_2\Sigma P_2' \end{pmatrix}\right)$$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada, $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^n$ y $P_1, P_2 \in M_{n \times n}$ matrices simétricas. Entonces:
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\boldsymbol{u}_2'\mathbf{y}$ son independientes si y sólo si $\boldsymbol{u}_1'\Sigma\boldsymbol{u}_2 = 0$
  - $\boldsymbol{u}_1'\mathbf{y}$ y $\mathbf{y}'P_1\mathbf{y}$ son independientes si y sólo si $P_1\Sigma\boldsymbol{u}_1 = \mathbf{0}_n$
  - $P_1\mathbf{y}$ y $P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$
  - $\mathbf{y}'P_1\mathbf{y}$ y $\mathbf{y}'P_2\mathbf{y}$ son independientes si y sólo si $P_1\Sigma P_2 = \mathbf{0}_{n \times n}$

- Si $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada entonces $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_n$

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n \times n}$ simétrica con $rg(P) = p$. Entonces: $(\mathbf{y} - \boldsymbol{\mu})' P (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$ si y sólo si $P\Sigma$ es idempotente

  (es decir $(P\Sigma)(P\Sigma) = P\Sigma$)

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL II

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n\times n}$ idempotente y simétrica con $rg(P) = p$. Entonces:

$$(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$$

Si además $P\left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} = \mathbf{0}_n$ entonces:

$$\mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} \rightsquigarrow \chi^2_p$$

porque se verifica:

- $\boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} = \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right)' \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right) = 0$
- $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' (\mathbf{y} - \boldsymbol{\mu}) = \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y}$
  $-\mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} - \boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} + \boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}$
  $= \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} - \mathbf{y}' \Sigma^{-1/2} \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right) - \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right)' \left(\Sigma^{-1/2}\right)' \mathbf{y}$
  $= \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y}$

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL II

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n\times n}$ idempotente y simétrica con $rg(P) = p$. Entonces:

$$(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$$

Si además $P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} = \mathbf{0}_n$ entonces:

$$\mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} \rightsquigarrow \chi^2_p$$

porque se verifica:

- $\boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} = \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right)' \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right) = 0$
- $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' (\mathbf{y} - \boldsymbol{\mu}) = \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y}$
  $-\mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} - \boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} + \boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}$
  $= \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} - \mathbf{y}' \Sigma^{-1/2} \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right) - \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right)' \left(\Sigma^{-1/2}\right)' \mathbf{y}$
  $= \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y}$

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL II

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P \in M_{n\times n}$ idempotente y simétrica con $rg(P) = p$. Entonces:

$$(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' (\mathbf{y} - \boldsymbol{\mu}) \rightsquigarrow \chi^2_p$$

Si además $P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} = \mathbf{0}_n$ entonces:

$$\mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} \rightsquigarrow \chi^2_p$$

porque se verifica:

- $\boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} = \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right)' \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right) = 0$
- $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' (\mathbf{y} - \boldsymbol{\mu}) = \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y}$
  $-\mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu} - \boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} + \boldsymbol{\mu}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}$
  $= \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y} - \mathbf{y}' \Sigma^{-1/2} \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right) - \left(P \left(\Sigma^{-1/2}\right)' \boldsymbol{\mu}\right)' \left(\Sigma^{-1/2}\right)' \mathbf{y}$
  $= \mathbf{y}' \Sigma^{-1/2} P \left(\Sigma^{-1/2}\right)' \mathbf{y}$

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL III

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P_1, P_2 \in M_{n\times n}$ simétricas e idempotentes con $rg(P_1)=p_1$, $rg(P_2)=p_2$ y $P_1P_2=\mathbf{0}_{n\times n}$. Entonces:

$$\left(\frac{p_2}{p_1}\right)\left(\frac{(\mathbf{y}-\boldsymbol{\mu})'\Sigma^{-1/2}P_1\left(\Sigma^{-1/2}\right)'(\mathbf{y}-\boldsymbol{\mu})}{(\mathbf{y}-\boldsymbol{\mu})'\Sigma^{-1/2}P_2\left(\Sigma^{-1/2}\right)'(\mathbf{y}-\boldsymbol{\mu})}\right) \rightsquigarrow F_{p_1,p_2}$$

Si además $P_1\left(\Sigma^{-1/2}\right)'\boldsymbol{\mu} = P_2\left(\Sigma^{-1/2}\right)'\boldsymbol{\mu} = \mathbf{0}_n$ entonces:

$$\left(\frac{p_2}{p_1}\right)\left(\frac{\mathbf{y}'\Sigma^{-1/2}P_1\left(\Sigma^{-1/2}\right)'\mathbf{y}}{\mathbf{y}'\Sigma^{-1/2}P_2\left(\Sigma^{-1/2}\right)'\mathbf{y}}\right) \rightsquigarrow F_{p_1,p_2}$$

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL III

- Sean $\mathbf{y} \rightsquigarrow N_n(\boldsymbol{\mu}, \Sigma)$ no degenerada y $P_1, P_2 \in M_{n\times n}$ simétricas e idempotentes con $rg(P_1)=p_1$, $rg(P_2)=p_2$ y $P_1P_2=\mathbf{0}_{n\times n}$. Entonces:

$$\left(\frac{p_2}{p_1}\right)\left(\frac{(\mathbf{y}-\boldsymbol{\mu})'\Sigma^{-1/2}P_1\left(\Sigma^{-1/2}\right)'(\mathbf{y}-\boldsymbol{\mu})}{(\mathbf{y}-\boldsymbol{\mu})'\Sigma^{-1/2}P_2\left(\Sigma^{-1/2}\right)'(\mathbf{y}-\boldsymbol{\mu})}\right) \rightsquigarrow F_{p_1,p_2}$$

Si además $P_1\left(\Sigma^{-1/2}\right)'\boldsymbol{\mu} = P_2\left(\Sigma^{-1/2}\right)'\boldsymbol{\mu} = \mathbf{0}_n$ entonces:

$$\left(\frac{p_2}{p_1}\right)\left(\frac{\mathbf{y}'\Sigma^{-1/2}P_1\left(\Sigma^{-1/2}\right)'\mathbf{y}}{\mathbf{y}'\Sigma^{-1/2}P_2\left(\Sigma^{-1/2}\right)'\mathbf{y}}\right) \rightsquigarrow F_{p_1,p_2}$$

MODELO

LINEAL

GENERAL

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) I

- $y =$ variable respuesta aleatoria, numérica y continua
- $x_1, x_2, ..., x_k =$ variables independientes continuas (regresores) o categóricas (factores)
- OBJETIVO: Estimar el valor esperado de y para ciertos valores fijos $x_1, x_2, ..., x_k$ de las variables independientes
- Hipótesis: $E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$

  $y/x_1, x_2, ..., x_k \rightsquigarrow N(\mu, \sigma^2)$
- Necesitamos estimar los parámetros $\beta_0, \beta_1, ..., \beta_k, \sigma^2$
- Sean $(y_i, x_{i1}, x_{i2}, ..., x_{ik})$ con $i = 1, 2, ..., n$ observaciones independientes de la variable respuesta y ($x_{i1}, x_{i2}, ..., x_{ik}$ pueden ser observados o fijados por el experimentador)
- Por tanto: $y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon_i$

  con $\varepsilon_i \rightsquigarrow N(0, \sigma^2)$ y $Cov(\varepsilon_i, \varepsilon_{i'}) = 0$ para todo i, i'

Formulación del Modelo Lineal General (GLM) II

Matricialmente, si $p = k + 1$, tenemos

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ con } \begin{cases} \mathbf{y} \in \mathbb{R}^n, \boldsymbol{\varepsilon} \in \mathbb{R}^n \\ X \in M_{n\times p},\ \boldsymbol{\beta} \in \mathbb{R}^p \end{cases}$$

siendo

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdot & \cdot & x_{1k} \\ 1 & x_{21} & x_{22} & \cdot & \cdot & x_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & \cdot & \cdot & x_{nk} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{pmatrix}}_{\boldsymbol{\varepsilon}}$$

con $\boldsymbol{\varepsilon} \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right)$ y por tanto $\mathbf{y}/X \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 I_n\right)$

Siempre suponemos que $n > p$ y $rg(X) = p$

Formulación del Modelo Lineal General (GLM) II

Matricialmente, si $p = k + 1$, tenemos

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ con } \begin{cases} \mathbf{y} \in \mathbb{R}^n, \boldsymbol{\varepsilon} \in \mathbb{R}^n \\ X \in M_{n\times p},\ \boldsymbol{\beta} \in \mathbb{R}^p \end{cases}$$

$$\text{siendo } \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdot & \cdot & x_{1k} \\ 1 & x_{21} & x_{22} & \cdot & \cdot & x_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & \cdot & \cdot & x_{nk} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{pmatrix}}_{\boldsymbol{\varepsilon}}$$

con $\boldsymbol{\varepsilon} \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right)$ y por tanto $\mathbf{y}/X \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 I_n\right)$

Siempre suponemos que $n > p$ y $rg(X) = p$

Solución del modelo GLM

- $\boldsymbol{\varepsilon}=\mathbf{y}-X\boldsymbol{\beta} \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\boldsymbol{\varepsilon}=(\sigma^{-1})(\mathbf{y}-X\boldsymbol{\beta}) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ debería ser mínimo
- Si definimos $S(\boldsymbol{\beta})=(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ el problema es: $\min_{\boldsymbol{\beta}\in\mathbb{R}^p} S(\boldsymbol{\beta})$
- En forma escalar, $S(\boldsymbol{\beta})=\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{k}\beta_j x_{ij}\right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\boldsymbol{\beta})=-2X'(\mathbf{y}-X\boldsymbol{\beta})\in\mathbb{R}^p$
- Y para la derivada segunda: $Hess(\boldsymbol{\beta})=2X'X\in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y}-X\boldsymbol{\beta})=\mathbf{0}_n$ o equivalentemente $(X'X)\boldsymbol{\beta}=X'\mathbf{y}$
- Por ser $rg(X)=p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\boldsymbol{\beta}}=(X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\boldsymbol{\varepsilon}=\mathbf{y}-X\boldsymbol{\beta} \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\boldsymbol{\varepsilon}=(\sigma^{-1})(\mathbf{y}-X\boldsymbol{\beta}) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ debería ser mínimo
- Si definimos $S(\boldsymbol{\beta})=(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ el problema es: $\min_{\boldsymbol{\beta}\in\mathbb{R}^p} S(\boldsymbol{\beta})$
- En forma escalar, $S(\boldsymbol{\beta})=\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{k}\beta_j x_{ij}\right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\boldsymbol{\beta})=-2X'(\mathbf{y}-X\boldsymbol{\beta})\in\mathbb{R}^p$
- Y para la derivada segunda: $Hess(\boldsymbol{\beta})=2X'X\in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y}-X\boldsymbol{\beta})=\mathbf{0}_n$ o equivalentemente $(X'X)\boldsymbol{\beta}=X'\mathbf{y}$
- Por ser $rg(X)=p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\boldsymbol{\beta}}=(X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\boldsymbol{\varepsilon}=\mathbf{y}-X\boldsymbol{\beta} \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\boldsymbol{\varepsilon}=(\sigma^{-1})(\mathbf{y}-X\boldsymbol{\beta}) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ debería ser mínimo
- Si definimos $S(\boldsymbol{\beta})=(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ el problema es: $\min_{\boldsymbol{\beta}\in\mathbb{R}^p} S(\boldsymbol{\beta})$
- En forma escalar, $S(\boldsymbol{\beta})=\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{k}\beta_j x_{ij}\right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\boldsymbol{\beta})=-2X'(\mathbf{y}-X\boldsymbol{\beta})\in\mathbb{R}^p$
- Y para la derivada segunda: $Hess(\boldsymbol{\beta})=2X'X\in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y}-X\boldsymbol{\beta})=\mathbf{0}_n$ o equivalentemente $(X'X)\boldsymbol{\beta}=X'\mathbf{y}$
- Por ser $rg(X)=p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\boldsymbol{\beta}}=(X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\boldsymbol{\varepsilon}=\mathbf{y}-X\boldsymbol{\beta} \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\boldsymbol{\varepsilon}=(\sigma^{-1})(\mathbf{y}-X\boldsymbol{\beta}) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ debería ser mínimo
- Si definimos $S(\boldsymbol{\beta})=(\mathbf{y}-X\boldsymbol{\beta})'(\mathbf{y}-X\boldsymbol{\beta})$ el problema es: $\min_{\boldsymbol{\beta}\in\mathbb{R}^p} S(\boldsymbol{\beta})$
- En forma escalar, $S(\boldsymbol{\beta})=\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{k}\beta_j x_{ij}\right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\boldsymbol{\beta})=-2X'(\mathbf{y}-X\boldsymbol{\beta})\in\mathbb{R}^p$
- Y para la derivada segunda: $Hess(\boldsymbol{\beta})=2X'X\in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y}-X\boldsymbol{\beta})=\mathbf{0}_n$ o equivalentemente $(X'X)\boldsymbol{\beta}=X'\mathbf{y}$
- Por ser $rg(X)=p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\boldsymbol{\beta}}=(X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\varepsilon = \mathbf{y} - X\beta \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\varepsilon = (\sigma^{-1})(\mathbf{y} - X\beta) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ debería ser mínimo
- Si definimos $S(\beta) = (\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ el problema es: $\min_{\beta \in \mathbb{R}^p} S(\beta)$
- En forma escalar, $S(\beta) = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\beta) = -2X'(\mathbf{y} - X\beta) \in \mathbb{R}^p$
- Y para la derivada segunda: $Hess(\beta) = 2X'X \in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y} - X\beta) = \mathbf{0}_n$ o equivalentemente $(X'X)\beta = X'\mathbf{y}$
- Por ser $rg(X) = p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\beta} = (X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\varepsilon = \mathbf{y} - X\beta \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\varepsilon = (\sigma^{-1})(\mathbf{y} - X\beta) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ debería ser mínimo
- Si definimos $S(\beta) = (\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ el problema es: $\min_{\beta \in \mathbb{R}^p} S(\beta)$
- En forma escalar, $S(\beta) = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\beta) = -2X'(\mathbf{y} - X\beta) \in \mathbb{R}^p$
- Y para la derivada segunda: $Hess(\beta) = 2X'X \in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y} - X\beta) = \mathbf{0}_n$ o equivalentemente $(X'X)\beta = X'\mathbf{y}$
- Por ser $rg(X) = p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\beta} = (X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\varepsilon = \mathbf{y} - X\beta \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\varepsilon = (\sigma^{-1})(\mathbf{y} - X\beta) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ debería ser mínimo
- Si definimos $S(\beta) = (\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ el problema es: $\min_{\beta \in \mathbb{R}^p} S(\beta)$
- En forma escalar, $S(\beta) = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\beta) = -2X'(\mathbf{y} - X\beta) \in \mathbb{R}^p$
- Y para la derivada segunda: $Hess(\beta) = 2X'X \in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y} - X\beta) = \mathbf{0}_n$ o equivalentemente $(X'X)\beta = X'\mathbf{y}$
- Por ser $rg(X) = p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\beta} = (X'X)^{-1}X'\mathbf{y}$$

Solución del modelo GLM

- $\varepsilon = \mathbf{y} - X\beta \rightsquigarrow N_n\left(\mathbf{0}_n, \sigma^2 I_n\right) \Rightarrow (\sigma^{-1})\varepsilon = (\sigma^{-1})(\mathbf{y} - X\beta) \rightsquigarrow N_n\left(\mathbf{0}_n, I_n\right)$
- Por tanto, para cada σ fijo, $(\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ debería ser mínimo
- Si definimos $S(\beta) = (\mathbf{y} - X\beta)'(\mathbf{y} - X\beta)$ el problema es: $\min_{\beta \in \mathbb{R}^p} S(\beta)$
- En forma escalar, $S(\beta) = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2$, es decir, tenemos el problema de mínimos cuadrados
- Para la derivada primera tenemos: $\nabla S(\beta) = -2X'(\mathbf{y} - X\beta) \in \mathbb{R}^p$
- Y para la derivada segunda: $Hess(\beta) = 2X'X \in M_{pxp}$
- Por tanto la solución del problema se obtiene resolviendo el sistema $X'(\mathbf{y} - X\beta) = \mathbf{0}_n$ o equivalentemente $(X'X)\beta = X'\mathbf{y}$
- Por ser $rg(X) = p$ la matriz $X'X$ tiene inversa y la solución única es:

$$\widehat{\beta} = (X'X)^{-1}X'\mathbf{y}$$

Estimadores y distribuciones de probabilidad

- El vector de parámetros estimados por el modelo es

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1} X'\mathbf{y} \rightsquigarrow N_p\left(\boldsymbol{\beta}, \sigma^2 (X'X)^{-1}\right)$$

- El vector de valores predichos por el modelo es

$$\widehat{\mathbf{y}} = X\widehat{\beta} = X (X'X)^{-1} X'\mathbf{y} = H\mathbf{y} \rightsquigarrow N_n (X\beta, \sigma^2 H)$$

siendo $H = X (X'X)^{-1} X' \in M_{n\times n}$ la conocida matriz hat

- La matriz H verifica que $HH = H$ y $(I_n - H)(I_n - H) = I_n - H$
- El vector de errores del modelo es

$$\mathbf{e} = \mathbf{y} - X\widehat{\beta} = (I_n - H)\mathbf{y} \rightsquigarrow N_n\left(\mathbf{0}, \sigma^2(I_n - H)\right)$$

- Además se verifica que $\mathbf{e}'\mathbf{e} = S(\widehat{\beta}) = \mathbf{y}' (I_n - H)\, \mathbf{y} \rightsquigarrow \sigma^2 \chi^2_{n-p}$
- Por tanto podemos estimar σ^2 como: $\widehat{\sigma^2} = \dfrac{\mathbf{e}'\mathbf{e}}{n-p} = \dfrac{\mathbf{y}' (I_n - H)\, \mathbf{y}}{n-p}$

Estimadores y distribuciones de probabilidad

- El vector de parámetros estimados por el modelo es

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1} X'\mathbf{y} \rightsquigarrow N_p\left(\boldsymbol{\beta}, \sigma^2 (X'X)^{-1}\right)$$

- El vector de valores predichos por el modelo es

$$\widehat{\mathbf{y}} = X\widehat{\boldsymbol{\beta}} = X (X'X)^{-1} X'\mathbf{y} = H\mathbf{y} \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 H\right)$$

siendo $H = X (X'X)^{-1} X' \in M_{n\times n}$ la conocida matriz hat

- La matriz H verifica que $HH = H$ y $(I_n - H)(I_n - H) = I_n - H$
- El vector de errores del modelo es

$$\mathbf{e} = \mathbf{y} - X\widehat{\boldsymbol{\beta}} = (I_n - H)\mathbf{y} \rightsquigarrow N_n\left(\mathbf{0}, \sigma^2(I_n - H)\right)$$

- Además se verifica que $\mathbf{e}'\mathbf{e} = S(\widehat{\boldsymbol{\beta}}) = \mathbf{y}'(I_n - H)\mathbf{y} \rightsquigarrow \sigma^2 \chi^2_{n-p}$
- Por tanto podemos estimar σ^2 como: $\widehat{\sigma^2} = \dfrac{\mathbf{e}'\mathbf{e}}{n-p} = \dfrac{\mathbf{y}'(I_n - H)\mathbf{y}}{n-p}$

Estimadores y distribuciones de probabilidad

- El vector de parámetros estimados por el modelo es

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1} X'\mathbf{y} \rightsquigarrow N_p\left(\boldsymbol{\beta}, \sigma^2 (X'X)^{-1}\right)$$

- El vector de valores predichos por el modelo es

$$\widehat{\mathbf{y}} = X\widehat{\boldsymbol{\beta}} = X (X'X)^{-1} X'\mathbf{y} = H\mathbf{y} \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 H\right)$$

siendo $H = X (X'X)^{-1} X' \in M_{n\times n}$ la conocida matriz hat

- La matriz H verifica que $HH = H$ y $(I_n - H)(I_n - H) = I_n - H$
- El vector de errores del modelo es

$$\mathbf{e} = \mathbf{y} - X\widehat{\boldsymbol{\beta}} = (I_n - H)\mathbf{y} \rightsquigarrow N_n\left(\mathbf{0}, \sigma^2(I_n - H)\right)$$

- Además se verifica que $\mathbf{e}'\mathbf{e} = S(\widehat{\boldsymbol{\beta}}) = \mathbf{y}'\left(I_n - H\right)\mathbf{y} \rightsquigarrow \sigma^2\chi^2_{n-p}$
- Por tanto podemos estimar σ^2 como: $\widehat{\sigma^2} = \dfrac{\mathbf{e}'\mathbf{e}}{n-p} = \dfrac{\mathbf{y}'\left(I_n - H\right)\mathbf{y}}{n-p}$

Estimadores y distribuciones de probabilidad

- El vector de parámetros estimados por el modelo es

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1} X'\mathbf{y} \rightsquigarrow N_p\left(\boldsymbol{\beta}, \sigma^2 (X'X)^{-1}\right)$$

- El vector de valores predichos por el modelo es

$$\widehat{\mathbf{y}} = X\widehat{\boldsymbol{\beta}} = X (X'X)^{-1} X'\mathbf{y} = H\mathbf{y} \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 H\right)$$

siendo $H = X (X'X)^{-1} X' \in M_{n\times n}$ la conocida matriz hat

- La matriz H verifica que $HH = H$ y $(I_n - H)(I_n - H) = I_n - H$

- El vector de errores del modelo es

$$\mathbf{e} = \mathbf{y} - X\widehat{\boldsymbol{\beta}} = (I_n - H)\mathbf{y} \rightsquigarrow N_n\left(\mathbf{0}, \sigma^2(I_n - H)\right)$$

- Además se verifica que $\mathbf{e}'\mathbf{e} = S(\widehat{\boldsymbol{\beta}}) = \mathbf{y}'\left(I_n - H\right)\mathbf{y} \rightsquigarrow \sigma^2 \chi^2_{n-p}$

- Por tanto podemos estimar σ^2 como: $\widehat{\sigma^2} = \dfrac{\mathbf{e}'\mathbf{e}}{n-p} = \dfrac{\mathbf{y}'\left(I_n - H\right)\mathbf{y}}{n-p}$

Estimadores y distribuciones de probabilidad

- El vector de parámetros estimados por el modelo es

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1} X'\mathbf{y} \rightsquigarrow N_p\left(\boldsymbol{\beta}, \sigma^2 (X'X)^{-1}\right)$$

- El vector de valores predichos por el modelo es

$$\widehat{\mathbf{y}} = X\widehat{\boldsymbol{\beta}} = X (X'X)^{-1} X'\mathbf{y} = H\mathbf{y} \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 H\right)$$

siendo $H = X (X'X)^{-1} X' \in M_{n\times n}$ la conocida matriz hat

- La matriz H verifica que $HH = H$ y $(I_n - H)(I_n - H) = I_n - H$
- El vector de errores del modelo es

$$\mathbf{e} = \mathbf{y} - X\widehat{\boldsymbol{\beta}} = (I_n - H)\mathbf{y} \rightsquigarrow N_n\left(\mathbf{0}, \sigma^2(I_n - H)\right)$$

- Además se verifica que $\mathbf{e}'\mathbf{e} = S(\widehat{\boldsymbol{\beta}}) = \mathbf{y}' (I_n - H)\, \mathbf{y} \rightsquigarrow \sigma^2 \chi^2_{n-p}$
- Por tanto podemos estimar σ^2 como: $\widehat{\sigma^2} = \dfrac{\mathbf{e}'\mathbf{e}}{n-p} = \dfrac{\mathbf{y}' (I_n - H)\, \mathbf{y}}{n-p}$

Estimadores y distribuciones de probabilidad

- El vector de parámetros estimados por el modelo es

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1} X'\mathbf{y} \rightsquigarrow N_p\left(\boldsymbol{\beta}, \sigma^2 (X'X)^{-1}\right)$$

- El vector de valores predichos por el modelo es

$$\widehat{\mathbf{y}} = X\widehat{\boldsymbol{\beta}} = X (X'X)^{-1} X'\mathbf{y} = H\mathbf{y} \rightsquigarrow N_n\left(X\boldsymbol{\beta}, \sigma^2 H\right)$$

siendo $H = X (X'X)^{-1} X' \in M_{n\times n}$ la conocida matriz hat

- La matriz H verifica que $HH = H$ y $(I_n - H)(I_n - H) = I_n - H$
- El vector de errores del modelo es

$$\mathbf{e} = \mathbf{y} - X\widehat{\boldsymbol{\beta}} = (I_n - H)\mathbf{y} \rightsquigarrow N_n\left(\mathbf{0}, \sigma^2(I_n - H)\right)$$

- Además se verifica que $\mathbf{e}'\mathbf{e} = S(\widehat{\boldsymbol{\beta}}) = \mathbf{y}'\left(I_n - H\right)\mathbf{y} \rightsquigarrow \sigma^2\chi^2_{n-p}$
- Por tanto podemos estimar σ^2 como: $\widehat{\sigma^2} = \dfrac{\mathbf{e}'\mathbf{e}}{n-p} = \dfrac{\mathbf{y}'\left(I_n - H\right)\mathbf{y}}{n-p}$

Test global del modelo y calidad del ajuste

- Descomposición de la varianza:

$$\underbrace{\boldsymbol{y}'\boldsymbol{y} - n\overline{y}^2}_{\text{SST}} = \underbrace{(\boldsymbol{y}'H\boldsymbol{y} - n\overline{y}^2)}_{\text{SSR}} + \underbrace{\boldsymbol{y}'(I_n - H)\boldsymbol{y}}_{\text{SSE}} = (\widehat{\mathbf{y}}'\widehat{\mathbf{y}} - n\overline{y}^2) + \mathbf{e}'\mathbf{e}$$

siendo $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$

- Test global del modelo: $H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0 \Rightarrow \dfrac{(n-p)SSR}{(p-1)SSE} \rightsquigarrow F_{p-1,n-p}$
- Coeficiente de determinación: $R^2 = \dfrac{SSR}{SST} = 1 - \dfrac{SSE}{SST} \in [0,1]$
- Coeficiente de determinación ajustado: $\overline{R^2} = 1 - \dfrac{SSE/(n-p)}{SST/(n-1)}$
- Error estándar de estimación: $\widehat{\sigma} = \sqrt{\dfrac{SSE}{n-p}} = \sqrt{MSE}$
- Coeficiente de variación del modelo: $CV = \dfrac{\widehat{\sigma}}{\overline{y}}$ (porcentaje)

Test global del modelo y calidad del ajuste

- Descomposición de la varianza:
$$\underbrace{\boldsymbol{y'y} - n\overline{y}^2}_{\text{SST}} = \underbrace{(\boldsymbol{y'Hy} - n\overline{y}^2)}_{\text{SSR}} + \underbrace{\boldsymbol{y'}(I_n - H)\boldsymbol{y}}_{\text{SSE}} = (\widehat{\mathbf{y}}'\widehat{\mathbf{y}} - n\overline{y}^2) + \mathbf{e}'\mathbf{e}$$
siendo $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Test global del modelo: $H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0 \Rightarrow \frac{(n-p)SSR}{(p-1)SSE} \rightsquigarrow F_{p-1,n-p}$
- Coeficiente de determinación: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \in [0,1]$
- Coeficiente de determinación ajustado: $\overline{R^2} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}$
- Error estándar de estimación: $\widehat{\sigma} = \sqrt{\frac{SSE}{n-p}} = \sqrt{MSE}$
- Coeficiente de variación del modelo: $CV = \frac{\widehat{\sigma}}{\overline{y}}$ (porcentaje)

Test global del modelo y calidad del ajuste

- Descomposición de la varianza:
$$\underbrace{\boldsymbol{y}'\boldsymbol{y} - n\overline{y}^2}_{\text{SST}} = \underbrace{(\boldsymbol{y}'H\boldsymbol{y} - n\overline{y}^2)}_{\text{SSR}} + \underbrace{\boldsymbol{y}'(I_n - H)\boldsymbol{y}}_{\text{SSE}} = (\widehat{\mathbf{y}}'\widehat{\mathbf{y}} - n\overline{y}^2) + \mathbf{e}'\mathbf{e}$$
siendo $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Test global del modelo: $H_0{:}\beta_1{=}\beta_2{=}\ldots{=}\beta_k{=}0 \Rightarrow \frac{(n-p)SSR}{(p-1)SSE} \rightsquigarrow F_{p-1,n-p}$
- Coeficiente de determinación: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \in [0,1]$
- Coeficiente de determinación ajustado: $\overline{R^2} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}$
- Error estándar de estimación: $\widehat{\sigma} = \sqrt{\frac{SSE}{n-p}} = \sqrt{MSE}$
- Coeficiente de variación del modelo: $CV = \frac{\widehat{\sigma}}{\overline{y}}$ (porcentaje)

Test global del modelo y calidad del ajuste

- Descomposición de la varianza:
$$\underbrace{\boldsymbol{y'y} - n\overline{y}^2}_{\text{SST}} = \underbrace{(\boldsymbol{y'Hy} - n\overline{y}^2)}_{\text{SSR}} + \underbrace{\boldsymbol{y'}(I_n - H)\boldsymbol{y}}_{\text{SSE}} = (\widehat{\mathbf{y}}'\widehat{\mathbf{y}} - n\overline{y}^2) + \mathbf{e}'\mathbf{e}$$
siendo $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Test global del modelo: $H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0 \Rightarrow \frac{(n-p)SSR}{(p-1)SSE} \rightsquigarrow F_{p-1,n-p}$
- Coeficiente de determinación: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \in [0,1]$
- Coeficiente de determinación ajustado: $\overline{R^2} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}$
- Error estándar de estimación: $\widehat{\sigma} = \sqrt{\frac{SSE}{n-p}} = \sqrt{MSE}$
- Coeficiente de variación del modelo: $CV = \frac{\widehat{\sigma}}{\overline{y}}$ (porcentaje)

Test global del modelo y calidad del ajuste

- Descomposición de la varianza:
$$\underbrace{\mathbf{y'y} - n\overline{y}^2}_{\text{SST}} = \underbrace{(\mathbf{y'Hy} - n\overline{y}^2)}_{\text{SSR}} + \underbrace{\mathbf{y'}(I_n - H)\mathbf{y}}_{\text{SSE}} = (\widehat{\mathbf{y}}'\widehat{\mathbf{y}} - n\overline{y}^2) + \mathbf{e'e}$$
siendo $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Test global del modelo: $H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0 \Rightarrow \frac{(n-p)SSR}{(p-1)SSE} \rightsquigarrow F_{p-1,n-p}$
- Coeficiente de determinación: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \in [0,1]$
- Coeficiente de determinación ajustado: $\overline{R^2} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}$
- Error estándar de estimación: $\widehat{\sigma} = \sqrt{\frac{SSE}{n-p}} = \sqrt{MSE}$
- Coeficiente de variación del modelo: $CV = \frac{\widehat{\sigma}}{\overline{y}}$ (porcentaje)

Test global del modelo y calidad del ajuste

- Descomposición de la varianza:
$$\underbrace{\mathbf{y'y} - n\overline{y}^2}_{\text{SST}} = \underbrace{(\mathbf{y'Hy} - n\overline{y}^2)}_{\text{SSR}} + \underbrace{\mathbf{y'}(I_n - H)\mathbf{y}}_{\text{SSE}} = (\widehat{\mathbf{y}}'\widehat{\mathbf{y}} - n\overline{y}^2) + \mathbf{e'e}$$
siendo $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Test global del modelo: $H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0 \Rightarrow \frac{(n-p)SSR}{(p-1)SSE} \rightsquigarrow F_{p-1,n-p}$
- Coeficiente de determinación: $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \in [0,1]$
- Coeficiente de determinación ajustado: $\overline{R^2} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}$
- Error estándar de estimación: $\widehat{\sigma} = \sqrt{\frac{SSE}{n-p}} = \sqrt{MSE}$
- Coeficiente de variación del modelo: $CV = \frac{\widehat{\sigma}}{\overline{y}}$ (porcentaje)

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test de hipótesis e intervalos de confianza de parámetros

- c_{jj}=elemento diagonal j-ésimo de $C = (X'X)^{-1}$ con $j = 1, 2, ..., p$
- h_{ii}=elemento diagonal i-ésimo de $H = XCX'$ con $i = 1, 2, ..., n$
- $\frac{\widehat{\beta}_{j-1} - \beta_{j-1}}{\widehat{\sigma}\sqrt{c_{jj}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\beta_{j-1} = 0$
- $\widehat{\beta}_{j-1} \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{c_{jj}} \Rightarrow$ intervalo de confianza para β_{j-1}
- $\frac{\widehat{y}_i - \mu_i}{\widehat{\sigma}\sqrt{h_{ii}}} \rightsquigarrow t_{n-p} \Rightarrow$ test de hipótesis para $\mu_i = E(y/x_{i1}, x_{i2}, ..., x_{ik}) = 0$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{h_{ii}} \Rightarrow$ intervalo de confianza para $E(y/x_{i1}, x_{i2}, ..., x_{ik})$
- $\widehat{y}_i \pm t_{n-p;\alpha/2}\widehat{\sigma}\sqrt{1 + h_{ii}} \Rightarrow$ intervalo de predicción para una nueva observación de $y/x_{i1}, x_{i2}, ..., x_{ik}$
- $\left(\frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;1-\alpha/2}}, \frac{\mathbf{e}'\mathbf{e}}{\chi^2_{n-p;\alpha/2}} \right) \Rightarrow$ intervalo de confianza para σ^2

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \frac{\hat{\beta}' L' (LCL')^{-1} L\hat{\beta}}{\nu \hat{\sigma}^2} \rightsquigarrow F_{\nu, n-p}$ y tenemos test de hipótesis para H_0
- Los test F de una tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\hat{\sigma}\sqrt{LCL'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L\beta$ es $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L_1\beta - L_2\beta$ es $(L_1 - L_2)\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \frac{\hat{\beta}' L' (LCL')^{-1} L\hat{\beta}}{\nu \hat{\sigma}^2} \rightsquigarrow F_{\nu, n-p}$ y tenemos test de hipótesis para H_0
- Los test F de una tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\hat{\sigma}\sqrt{LCL'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L\beta$ es $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L_1\beta - L_2\beta$ es $(L_1 - L_2)\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \frac{\hat{\beta}' L' (LCL')^{-1} L\hat{\beta}}{\nu \hat{\sigma}^2} \rightsquigarrow F_{\nu, n-p}$ y tenemos test de hipótesis para H_0
- Los test F de una tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\hat{\sigma}\sqrt{LCL'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L\beta$ es $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L_1\beta - L_2\beta$ es $(L_1 - L_2)\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \frac{\hat{\beta}' L' (LCL')^{-1} L\hat{\beta}}{\nu \hat{\sigma}^2} \rightsquigarrow F_{\nu, n-p}$ y tenemos test de hipótesis para H_0
- Los test F de una tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\hat{\sigma}\sqrt{LCL'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L\beta$ es $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L_1\beta - L_2\beta$ es $(L_1 - L_2)\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \dfrac{\hat{\beta}' L' (LCL')^{-1} L\hat{\beta}}{\nu \hat{\sigma}^2} \rightsquigarrow F_{\nu, n-p}$ y tenemos test de hipótesis para H_0
- Los test F de una tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \dfrac{L\hat{\beta} - L\beta}{\hat{\sigma}\sqrt{LCL'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L\beta$ es $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$
- Si $L_1, L_2 \in M_{1xp}$, $t = \dfrac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L_1\beta - L_2\beta$ es $(L_1 - L_2)\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \frac{\hat{\beta}' L' (LCL')^{-1} L\hat{\beta}}{\nu \hat{\sigma}^2} \rightsquigarrow F_{\nu, n-p}$ y tenemos test de hipótesis para H_0
- Los test F de una tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\hat{\sigma}\sqrt{LCL'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L\beta$ es $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}} \rightsquigarrow t_{n-p}$ y un intervalo de confianza para $L_1\beta - L_2\beta$ es $(L_1 - L_2)\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{(L_1 - L_2)C(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{n-p;\alpha/2}\hat{\sigma}\sqrt{LCL'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \frac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \dfrac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \frac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \frac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \dfrac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \dfrac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Validación del modelo GLM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales e_i o mejor aún con los residuales studentizados $r_i = \dfrac{e_i}{\widehat{\sigma}\sqrt{1-h_{ii}}}$
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i)$ y (x_{ij}, r_i) para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: nube de dispersión (i, r_i) y test de Durbin-Watson
- Posibles residuales atípicos (outliers): aquellos con $|r_i| > 2$ (o mejor 3)
- Posibles puntos influyentes: aquellos con residual press studentizado mayor que 2 (o mejor 3), D-Cook's$>4/n$ o diagonal hat $h_{ii}>2p/n$
- Existencia de multicolinealidad: número de condición> 1000 o presencia de variables con $VIF > 10$ (o tolerancia< 0.1)
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$

Software para el modelo GLM

- Si usamos el software R disponemos del paquete lm
- Si usamos el software SAS tenemos el procedimiento proc glm, ó también proc reg si sólo hay regresores
- Sintaxis básica general de proc glm:

ods pdf file="..";ods graphics on; *(opcional)*
PROC GLM options;
    CLASS factores;
    MODEL dependent=independents / options;
    CONTRAST 'label' [effect values].../options;
    ESTIMATE 'name' effect values.../options;
    LSMEANS effects/options; (**sólo factores incluidos en CLASS**)
    OUTPUT OUT=SAS data set keyword=names...;
    TEST H=effects E=effect/options;
RUN;
ods graphics off;ods pdf close;*(si usamos los dos primeros comandos)*
QUIT;

Software para el modelo GLM

- Si usamos el software R disponemos del paquete lm
- Si usamos el software SAS tenemos el procedimiento proc glm, ó también proc reg si sólo hay regresores
- Sintaxis básica general de proc glm:

```
ods pdf file="...";ods graphics on; (opcional)
PROC GLM options;
    CLASS factores;
    MODEL dependent=independents / options;
    CONTRAST 'label' [effect values].../options;
    ESTIMATE 'name' effect values.../options;
    LSMEANS effects/options; (sólo factores incluidos en CLASS)
    OUTPUT OUT=SAS data set keyword=names...;
    TEST H=effects E=effect/options;
RUN;
ods graphics off;ods pdf close;(si usamos los dos primeros comandos)
QUIT;
```

Software para el modelo GLM

- Si usamos el software R disponemos del paquete lm
- Si usamos el software SAS tenemos el procedimiento proc glm, ó también proc reg si sólo hay regresores
- Sintaxis básica general de proc glm:

ods pdf file="...";ods graphics on; *(opcional)*
PROC GLM options;
    CLASS factores;
    MODEL dependent=independents / options;
    CONTRAST 'label' [effect values].../options;
    ESTIMATE 'name' effect values.../options;
    LSMEANS effects/options; (**sólo factores incluidos en CLASS**)
    OUTPUT OUT=SAS data set keyword=names...;
    TEST H=effects E=effect/options;
RUN;
ods graphics off;ods pdf close;*(si usamos los dos primeros comandos)*
QUIT;

MODELO

LINEAL

MIXTO

Formulación del Modelo Lineal Mixto (LMM) I

- Al igual que en el modelo GLM, suponemos que

$$E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$$

- Las observaciones de y pueden estar acompañadas de ciertos efectos aleatorios γ_l, con $l = 1, ..., q$ y $E(\gamma_l) = 0$, además del error aleatorio ε
- Consideramos ciertas observaciones $(y_i, x_{i1}, ..., x_{ik}, z_{i1}, ..., z_{iq},)$ con $i = 1, 2, ..., n$, donde z_{il} son variables dummy de tipo 0-1 (presencia o no de cada efecto aleatorio en cada observación) o numéricas (con coeficientes aleatorios cuya varianza queremos estimar)
- Entonces tendremos $y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{q} \gamma_l z_{il} + \varepsilon_i$
- Suponemos que γ_l y ε_i tienen todos distribución normal con $E(\gamma_l) = E(\varepsilon_i) = 0$, y además γ_l y ε_i son siempre independientes
- Sin embargo, no suponemos independencia y homogeneidad de varianzas entre los valores γ_l ni entre los valores ε_i

Formulación del Modelo Lineal Mixto (LMM) I

- Al igual que en el modelo GLM, suponemos que

$$E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$$

- Las observaciones de y pueden estar acompañadas de ciertos efectos aleatorios γ_l, con $l = 1, ..., q$ y $E(\gamma_l) = 0$, además del error aleatorio ε
- Consideramos ciertas observaciones $(y_i, x_{i1}, ..., x_{ik}, z_{i1}, ..., z_{iq},)$ con $i = 1, 2, ..., n$, donde z_{il} son variables dummy de tipo 0-1 (presencia o no de cada efecto aleatorio en cada observación) o numéricas (con coeficientes aleatorios cuya varianza queremos estimar)
- Entonces tendremos $y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{q} \gamma_l z_{il} + \varepsilon_i$
- Suponemos que γ_l y ε_i tienen todos distribución normal con $E(\gamma_l) = E(\varepsilon_i) = 0$, y además γ_l y ε_i son siempre independientes
- Sin embargo, no suponemos independencia y homogeneidad de varianzas entre los valores γ_l ni entre los valores ε_i

Formulación del Modelo Lineal Mixto (LMM) I

- Al igual que en el modelo GLM, suponemos que

$$E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$$

- Las observaciones de y pueden estar acompañadas de ciertos efectos aleatorios γ_l, con $l = 1, ..., q$ y $E(\gamma_l) = 0$, además del error aleatorio ε
- Consideramos ciertas observaciones $(y_i, x_{i1}, ..., x_{ik}, z_{i1}, ..., z_{iq},)$ con $i = 1, 2, ..., n$, donde z_{il} son variables dummy de tipo 0-1 (presencia o no de cada efecto aleatorio en cada observación) o numéricas (con coeficientes aleatorios cuya varianza queremos estimar)
- Entonces tendremos $y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{q} \gamma_l z_{il} + \varepsilon_i$
- Suponemos que γ_l y ε_i tienen todos distribución normal con $E(\gamma_l) = E(\varepsilon_i) = 0$, y además γ_l y ε_i son siempre independientes
- Sin embargo, no suponemos independencia y homogeneidad de varianzas entre los valores γ_l ni entre los valores ε_i

Formulación del Modelo Lineal Mixto (LMM) I

- Al igual que en el modelo GLM, suponemos que

$$E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$$

- Las observaciones de y pueden estar acompañadas de ciertos efectos aleatorios γ_l, con $l = 1, ..., q$ y $E(\gamma_l) = 0$, además del error aleatorio ε
- Consideramos ciertas observaciones $(y_i, x_{i1}, ..., x_{ik}, z_{i1}, ..., z_{iq},)$ con $i = 1, 2, ..., n$, donde z_{il} son variables dummy de tipo 0-1 (presencia o no de cada efecto aleatorio en cada observación) o numéricas (con coeficientes aleatorios cuya varianza queremos estimar)
- Entonces tendremos $y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{q} \gamma_l z_{il} + \varepsilon_i$
- Suponemos que γ_l y ε_i tienen todos distribución normal con $E(\gamma_l) = E(\varepsilon_i) = 0$, y además γ_l y ε_i son siempre independientes
- Sin embargo, no suponemos independencia y homogeneidad de varianzas entre los valores γ_l ni entre los valores ε_i

Formulación del Modelo Lineal Mixto (LMM) I

- Al igual que en el modelo GLM, suponemos que

$$E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$$

- Las observaciones de y pueden estar acompañadas de ciertos efectos aleatorios γ_l, con $l = 1, ..., q$ y $E(\gamma_l) = 0$, además del error aleatorio ε
- Consideramos ciertas observaciones $(y_i, x_{i1}, ..., x_{ik}, z_{i1}, ..., z_{iq},)$ con $i = 1, 2, ..., n$, donde z_{il} son variables dummy de tipo 0-1 (presencia o no de cada efecto aleatorio en cada observación) o numéricas (con coeficientes aleatorios cuya varianza queremos estimar)
- Entonces tendremos $y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{q} \gamma_l z_{il} + \varepsilon_i$
- Suponemos que γ_l y ε_i tienen todos distribución normal con $E(\gamma_l) = E(\varepsilon_i) = 0$, y además γ_l y ε_i son siempre independientes
- Sin embargo, no suponemos independencia y homogeneidad de varianzas entre los valores γ_l ni entre los valores ε_i

Formulación del Modelo Lineal Mixto (LMM) I

- Al igual que en el modelo GLM, suponemos que

$$E(y/x_1, x_2, ..., x_k) = \mu = \beta_0 + \sum_{j=1}^{k} \beta_j x_j$$

- Las observaciones de y pueden estar acompañadas de ciertos efectos aleatorios γ_l, con $l = 1, ..., q$ y $E(\gamma_l) = 0$, además del error aleatorio ε
- Consideramos ciertas observaciones $(y_i, x_{i1}, ..., x_{ik}, z_{i1}, ..., z_{iq},)$ con $i = 1, 2, ..., n$, donde z_{il} son variables dummy de tipo 0-1 (presencia o no de cada efecto aleatorio en cada observación) o numéricas (con coeficientes aleatorios cuya varianza queremos estimar)
- Entonces tendremos $y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{q} \gamma_l z_{il} + \varepsilon_i$
- Suponemos que γ_l y ε_i tienen todos distribución normal con $E(\gamma_l) = E(\varepsilon_i) = 0$, y además γ_l y ε_i son siempre independientes
- Sin embargo, no suponemos independencia y homogeneidad de varianzas entre los valores γ_l ni entre los valores ε_i

Formulación del Modelo Lineal Mixto (LMM) II

Matricialmente, si $p = k + 1$, tenemos:

$$\mathbf{y} = X\boldsymbol{\beta} + Z\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \text{ con } \begin{cases} \mathbf{y}, \boldsymbol{\varepsilon} \in \mathbb{R}^n \\ X \in M_{n\times p}, \boldsymbol{\beta} \in \mathbb{R}^p \\ Z \in M_{n\times q}, \boldsymbol{\gamma} \in \mathbb{R}^q \end{cases}$$

siendo X la matriz de diseño para las variables independientes, con $rg(X) = p$, y Z la matriz de diseño para los efectos aleatorios del modelo, con $rg(Z) = q$

Además, las hipótesis para los vectores aleatorios $\boldsymbol{\gamma}$ y $\boldsymbol{\varepsilon}$ son:

$$\begin{cases} \boldsymbol{\gamma} \rightsquigarrow N_q\left(\mathbf{0}_q, G(\boldsymbol{\theta_1})\right) \\ \boldsymbol{\varepsilon} \rightsquigarrow N_n\left(\mathbf{0}_n, R(\boldsymbol{\theta_2})\right) \\ \boldsymbol{\gamma}, \boldsymbol{\varepsilon} \text{ independientes} \end{cases}$$

siendo $G(\boldsymbol{\theta_1}) \in M_{q\times q}$, $R(\boldsymbol{\theta_2}) \in M_{n\times n}$ ciertas matrices de varianzas-covarianzas dependientes de ciertos vectores de parámetros de los efectos aleatorios $\boldsymbol{\theta_1} \in \mathbb{R}^{t_1}$ y $\boldsymbol{\theta_2} \in \mathbb{R}^{t_2}$

Formulación del Modelo Lineal Mixto (LMM) II

Matricialmente, si $p = k + 1$, tenemos:

$$\mathbf{y} = X\beta + Z\gamma + \varepsilon \text{ con } \begin{cases} \mathbf{y}, \varepsilon \in \mathbb{R}^n \\ X \in M_{n\times p}, \beta \in \mathbb{R}^p \\ Z \in M_{n\times q}, \gamma \in \mathbb{R}^q \end{cases}$$

siendo X la matriz de diseño para las variables independientes, con $rg(X) = p$, y Z la matriz de diseño para los efectos aleatorios del modelo, con $rg(Z) = q$

Además, las hipótesis para los vectores aleatorios γ y ε son:

$$\begin{cases} \gamma \rightsquigarrow N_q\left(\mathbf{0}_q, G(\boldsymbol{\theta_1})\right) \\ \varepsilon \rightsquigarrow N_n\left(\mathbf{0}_n, R(\boldsymbol{\theta_2})\right) \\ \gamma, \varepsilon \text{ independientes} \end{cases}$$

siendo $G(\boldsymbol{\theta_1}) \in M_{q\times q}$, $R(\boldsymbol{\theta_2}) \in M_{n\times n}$ ciertas matrices de varianzas-covarianzas dependientes de ciertos vectores de parámetros de los efectos aleatorios $\boldsymbol{\theta_1} \in \mathbb{R}^{t_1}$ y $\boldsymbol{\theta_2} \in \mathbb{R}^{t_2}$

Formulación del Modelo Lineal Mixto (LMM) III

- Con esta formulación e hipótesis tenemos:

$$\begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, \begin{pmatrix} G(\boldsymbol{\theta_1}) & \mathbf{0}_{qxn} \\ \mathbf{0}_{nxq} & R(\boldsymbol{\theta_2}) \end{pmatrix}\right)$$

- Si definimos el vector aleatorio $\varepsilon_{\boldsymbol{m}} = Z\gamma + \varepsilon \in \mathbb{R}^n$, podemos escribir el modelo como $\mathbf{y} = X\beta + \varepsilon_{\boldsymbol{m}}$ con la hipótesis $\varepsilon_{\boldsymbol{m}} \rightsquigarrow N_n(\mathbf{0}_n, V(\theta))$, siendo $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \in \mathbb{R}^t$, con $t = t_1 + t_2$, y $V(\theta) = ZG(\boldsymbol{\theta_1})Z' + R(\boldsymbol{\theta_2})$
- Esta formulación del modelo se denomina modelo marginal asociado al modelo lineal mixto. Si $V(\theta) = \sigma^2 I_n$ el modelo marginal coincide con el modelo GLM
- El modelo marginal es una generalización del modelo GLM con una matriz de varianzas-covarianzas para los errores $V(\theta)$, que no exige independencia y homogeneidad de varianzas de los errores

Formulación del Modelo Lineal Mixto (LMM) III

- Con esta formulación e hipótesis tenemos:

$$\begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, \begin{pmatrix} G(\boldsymbol{\theta_1}) & \mathbf{0}_{qxn} \\ \mathbf{0}_{nxq} & R(\boldsymbol{\theta_2}) \end{pmatrix}\right)$$

- Si definimos el vector aleatorio $\boldsymbol{\varepsilon_m} = Z\gamma + \varepsilon \in \mathbb{R}^n$, podemos escribir el modelo como $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon_m}$ con la hipótesis $\boldsymbol{\varepsilon_m} \rightsquigarrow N_n(\mathbf{0}_n, V(\boldsymbol{\theta}))$, siendo $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta_1} \\ \boldsymbol{\theta_2} \end{pmatrix} \in \mathbb{R}^t$, con $t = t_1 + t_2$, y $V(\boldsymbol{\theta}) = ZG(\boldsymbol{\theta_1})Z' + R(\boldsymbol{\theta_2})$
- Esta formulación del modelo se denomina modelo marginal asociado al modelo lineal mixto. Si $V(\theta) = \sigma^2 I_n$ el modelo marginal coincide con el modelo GLM
- El modelo marginal es una generalización del modelo GLM con una matriz de varianzas-covarianzas para los errores $V(\theta)$, que no exige independencia y homogeneidad de varianzas de los errores

Formulación del Modelo Lineal Mixto (LMM) III

- Con esta formulación e hipótesis tenemos:

$$\begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, \begin{pmatrix} G(\boldsymbol{\theta_1}) & \mathbf{0}_{qxn} \\ \mathbf{0}_{nxq} & R(\boldsymbol{\theta_2}) \end{pmatrix}\right)$$

- Si definimos el vector aleatorio $\boldsymbol{\varepsilon_m} = Z\gamma + \varepsilon \in \mathbb{R}^n$, podemos escribir el modelo como $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon_m}$ con la hipótesis $\boldsymbol{\varepsilon_m} \rightsquigarrow N_n(\mathbf{0}_n, V(\boldsymbol{\theta}))$, siendo $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta_1} \\ \boldsymbol{\theta_2} \end{pmatrix} \in \mathbb{R}^t$, con $t = t_1 + t_2$, y $V(\boldsymbol{\theta}) = ZG(\boldsymbol{\theta_1})Z' + R(\boldsymbol{\theta_2})$
- Esta formulación del modelo se denomina modelo marginal asociado al modelo lineal mixto. Si $V(\boldsymbol{\theta}) = \sigma^2 I_n$ el modelo marginal coincide con el modelo GLM
- El modelo marginal es una generalización del modelo GLM con una matriz de varianzas-covarianzas para los errores $V(\theta)$, que no exige independencia y homogeneidad de varianzas de los errores

Formulación del Modelo Lineal Mixto (LMM) III

- Con esta formulación e hipótesis tenemos:

$$\begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, \begin{pmatrix} G(\boldsymbol{\theta_1}) & \mathbf{0}_{qxn} \\ \mathbf{0}_{nxq} & R(\boldsymbol{\theta_2}) \end{pmatrix}\right)$$

- Si definimos el vector aleatorio $\boldsymbol{\varepsilon_m} = Z\gamma + \varepsilon \in \mathbb{R}^n$, podemos escribir el modelo como $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon_m}$ con la hipótesis $\boldsymbol{\varepsilon_m} \rightsquigarrow N_n(\mathbf{0}_n, V(\boldsymbol{\theta}))$, siendo $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta_1} \\ \boldsymbol{\theta_2} \end{pmatrix} \in \mathbb{R}^t$, con $t = t_1 + t_2$, y $V(\boldsymbol{\theta}) = ZG(\boldsymbol{\theta_1})Z' + R(\boldsymbol{\theta_2})$
- Esta formulación del modelo se denomina modelo marginal asociado al modelo lineal mixto. Si $V(\boldsymbol{\theta}) = \sigma^2 I_n$ el modelo marginal coincide con el modelo GLM
- El modelo marginal es una generalización del modelo GLM con una matriz de varianzas-covarianzas para los errores $V(\boldsymbol{\theta})$, que no exige independencia y homogeneidad de varianzas de los errores

Formulación del Modelo Lineal Mixto (LMM) IV

- Observar que, para el vector respuesta $\mathbf{y}$, tenemos que

$$\mathbf{y}/X, Z \rightsquigarrow N_n\left(X\boldsymbol{\beta}, V\left(\boldsymbol{\theta}\right)\right) \text{(modelo marginal)}$$

mientras que, para $\boldsymbol{\gamma}$ fijo, se tendrá

$$\mathbf{y}/X, Z, \boldsymbol{\gamma} \rightsquigarrow N_n\left(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}, R\left(\boldsymbol{\theta_2}\right)\right) \text{(modelo condicional)}$$

- Si no hay efectos aleatorios en el modelo, entonces $Z = 0$ y estos dos modelos coinciden. Pero para obtener el modelo GLM necesitamos además que $R\left(\boldsymbol{\theta}_2\right) = \sigma^2 I_n$
- Por tanto, podemos tener modelos LMM sin efectos aleatorios
- Existen matrices $G(\boldsymbol{\theta}_1)^{1/2}$, $R(\boldsymbol{\theta}_2)^{1/2}$ y $V\left(\boldsymbol{\theta}\right)^{1/2}$ verificando $\left(G(\boldsymbol{\theta}_1)^{1/2}\right)' G(\boldsymbol{\theta}_1)^{1/2} = G(\boldsymbol{\theta}_1)$, $\left(R(\boldsymbol{\theta}_2)^{1/2}\right)' R(\boldsymbol{\theta}_2)^{1/2} = R(\boldsymbol{\theta}_2)$ y $\left(V\left(\boldsymbol{\theta}\right)^{1/2}\right)' V\left(\boldsymbol{\theta}\right)^{1/2} = V\left(\boldsymbol{\theta}\right)$ (Cholesky root matrix)
- También existen sus inversas $G(\boldsymbol{\theta}_1)^{-1/2}$, $R(\boldsymbol{\theta}_2)^{-1/2}$ y $V\left(\boldsymbol{\theta}\right)^{-1/2}$

Formulación del Modelo Lineal Mixto (LMM) IV

- Observar que, para el vector respuesta $\mathbf{y}$, tenemos que

$$\mathbf{y}/X, Z \rightsquigarrow N_n\left(X\boldsymbol{\beta}, V\left(\boldsymbol{\theta}\right)\right) \text{(modelo marginal)}$$

mientras que, para $\boldsymbol{\gamma}$ fijo, se tendrá

$$\mathbf{y}/X, Z, \boldsymbol{\gamma} \rightsquigarrow N_n\left(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}, R\left(\boldsymbol{\theta_2}\right)\right) \text{(modelo condicional)}$$

- Si no hay efectos aleatorios en el modelo, entonces $Z = 0$ y estos dos modelos coinciden. Pero para obtener el modelo GLM necesitamos además que $R\left(\boldsymbol{\theta_2}\right) = \sigma^2 I_n$
- Por tanto, podemos tener modelos LMM sin efectos aleatorios
- Existen matrices $G(\boldsymbol{\theta_1})^{1/2}$, $R(\boldsymbol{\theta_2})^{1/2}$ y $V\left(\boldsymbol{\theta}\right)^{1/2}$ verificando $\left(G(\boldsymbol{\theta_1})^{1/2}\right)' G(\boldsymbol{\theta_1})^{1/2} = G(\boldsymbol{\theta_1})$, $\left(R(\boldsymbol{\theta_2})^{1/2}\right)' R(\boldsymbol{\theta_2})^{1/2} = R(\boldsymbol{\theta_2})$ y $\left(V\left(\boldsymbol{\theta}\right)^{1/2}\right)' V\left(\boldsymbol{\theta}\right)^{1/2} = V\left(\boldsymbol{\theta}\right)$ (Cholesky root matrix)
- También existen sus inversas $G(\boldsymbol{\theta_1})^{-1/2}$, $R(\boldsymbol{\theta_2})^{-1/2}$ y $V\left(\boldsymbol{\theta}\right)^{-1/2}$

Formulación del Modelo Lineal Mixto (LMM) IV

- Observar que, para el vector respuesta $\mathbf{y}$, tenemos que

$$\mathbf{y}/X, Z \rightsquigarrow N_n\left(X\boldsymbol{\beta}, V\left(\boldsymbol{\theta}\right)\right) \text{(modelo marginal)}$$

mientras que, para $\boldsymbol{\gamma}$ fijo, se tendrá

$$\mathbf{y}/X, Z, \boldsymbol{\gamma} \rightsquigarrow N_n\left(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}, R\left(\boldsymbol{\theta_2}\right)\right) \text{(modelo condicional)}$$

- Si no hay efectos aleatorios en el modelo, entonces $Z = 0$ y estos dos modelos coinciden. Pero para obtener el modelo GLM necesitamos además que $R\left(\boldsymbol{\theta_2}\right) = \sigma^2 I_n$
- Por tanto, podemos tener modelos LMM sin efectos aleatorios
- Existen matrices $G(\boldsymbol{\theta_1})^{1/2}$, $R(\boldsymbol{\theta_2})^{1/2}$ y $V\left(\boldsymbol{\theta}\right)^{1/2}$ verificando $\left(G(\boldsymbol{\theta_1})^{1/2}\right)' G(\boldsymbol{\theta_1})^{1/2} = G(\boldsymbol{\theta_1})$, $\left(R(\boldsymbol{\theta_2})^{1/2}\right)' R(\boldsymbol{\theta_2})^{1/2} = R(\boldsymbol{\theta_2})$ y $\left(V\left(\boldsymbol{\theta}\right)^{1/2}\right)' V\left(\boldsymbol{\theta}\right)^{1/2} = V\left(\boldsymbol{\theta}\right)$ (Cholesky root matrix)
- También existen sus inversas $G(\boldsymbol{\theta_1})^{-1/2}$, $R(\boldsymbol{\theta_2})^{-1/2}$ y $V\left(\boldsymbol{\theta}\right)^{-1/2}$

Formulación del Modelo Lineal Mixto (LMM) IV

- Observar que, para el vector respuesta $\mathbf{y}$, tenemos que

$$\mathbf{y}/X, Z \rightsquigarrow N_n\left(X\boldsymbol{\beta}, V\left(\boldsymbol{\theta}\right)\right) \text{(modelo marginal)}$$

mientras que, para $\boldsymbol{\gamma}$ fijo, se tendrá

$$\mathbf{y}/X, Z, \boldsymbol{\gamma} \rightsquigarrow N_n\left(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}, R\left(\boldsymbol{\theta_2}\right)\right) \text{(modelo condicional)}$$

- Si no hay efectos aleatorios en el modelo, entonces $Z = 0$ y estos dos modelos coinciden. Pero para obtener el modelo GLM necesitamos además que $R\left(\boldsymbol{\theta_2}\right) = \sigma^2 I_n$
- Por tanto, podemos tener modelos LMM sin efectos aleatorios
- Existen matrices $G(\boldsymbol{\theta_1})^{1/2}$, $R(\boldsymbol{\theta_2})^{1/2}$ y $V\left(\boldsymbol{\theta}\right)^{1/2}$ verificando $\left(G(\boldsymbol{\theta_1})^{1/2}\right)' G(\boldsymbol{\theta_1})^{1/2} = G(\boldsymbol{\theta_1})$, $\left(R(\boldsymbol{\theta_2})^{1/2}\right)' R(\boldsymbol{\theta_2})^{1/2} = R(\boldsymbol{\theta_2})$ y $\left(V\left(\boldsymbol{\theta}\right)^{1/2}\right)' V\left(\boldsymbol{\theta}\right)^{1/2} = V\left(\boldsymbol{\theta}\right)$ (Cholesky root matrix)
- También existen sus inversas $G(\boldsymbol{\theta_1})^{-1/2}$, $R(\boldsymbol{\theta_2})^{-1/2}$ y $V\left(\boldsymbol{\theta}\right)^{-1/2}$

Formulación del Modelo Lineal Mixto (LMM) IV

- Observar que, para el vector respuesta $\mathbf{y}$, tenemos que

$$\mathbf{y}/X, Z \rightsquigarrow N_n\left(X\boldsymbol{\beta}, V\left(\boldsymbol{\theta}\right)\right) \text{(modelo marginal)}$$

mientras que, para $\boldsymbol{\gamma}$ fijo, se tendrá

$$\mathbf{y}/X, Z, \boldsymbol{\gamma} \rightsquigarrow N_n\left(X\boldsymbol{\beta} + Z\boldsymbol{\gamma}, R\left(\boldsymbol{\theta_2}\right)\right) \text{(modelo condicional)}$$

- Si no hay efectos aleatorios en el modelo, entonces $Z = 0$ y estos dos modelos coinciden. Pero para obtener el modelo GLM necesitamos además que $R\left(\boldsymbol{\theta_2}\right) = \sigma^2 I_n$
- Por tanto, podemos tener modelos LMM sin efectos aleatorios
- Existen matrices $G(\boldsymbol{\theta_1})^{1/2}$, $R(\boldsymbol{\theta_2})^{1/2}$ y $V\left(\boldsymbol{\theta}\right)^{1/2}$ verificando $\left(G(\boldsymbol{\theta_1})^{1/2}\right)' G(\boldsymbol{\theta_1})^{1/2} = G(\boldsymbol{\theta_1})$, $\left(R(\boldsymbol{\theta_2})^{1/2}\right)' R(\boldsymbol{\theta_2})^{1/2} = R(\boldsymbol{\theta_2})$ y $\left(V\left(\boldsymbol{\theta}\right)^{1/2}\right)' V\left(\boldsymbol{\theta}\right)^{1/2} = V\left(\boldsymbol{\theta}\right)$ (Cholesky root matrix)
- También existen sus inversas $G(\boldsymbol{\theta_1})^{-1/2}$, $R(\boldsymbol{\theta_2})^{-1/2}$ y $V\left(\boldsymbol{\theta}\right)^{-1/2}$

Solución del modelo LMM I

- Si $\gamma^* = G(\boldsymbol{\theta_1})^{-1/2}\gamma$ y $\varepsilon^* = R(\boldsymbol{\theta_2})^{-1/2}\varepsilon$ tenemos que

$$\begin{pmatrix}\gamma^* \\ \varepsilon^*\end{pmatrix} = \begin{pmatrix}G(\boldsymbol{\theta_1})^{-1/2}\gamma \\ R(\boldsymbol{\theta_2})^{-1/2}\varepsilon\end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, I_{q+n}\right)$$

  y por tanto, para cada $\boldsymbol{\theta}$ fijo, $(\gamma^*)'\gamma^* + (\varepsilon^*)'\varepsilon^*$ debería ser mínimo

- Si $S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = \gamma' G(\theta_1)^{-1}\gamma + (\mathbf{y} - X\beta - Z\gamma)' R(\theta_2)^{-1}(\mathbf{y} - X\beta - Z\gamma)$,

  para cada θ fijo, el problema es: $\min_{\beta\in\mathbb{R}^p, \gamma\in\mathbb{R}^q} S\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$

- Si derivamos dos veces tenemos:

$$\nabla S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = 2\begin{pmatrix} -X' R(\theta_2)^{-1}(\mathbf{y} - X\beta - Z\gamma) \\ G(\theta_1)^{-1}\gamma - Z' R(\theta_2)^{-1}(\mathbf{y} - X\beta - Z\gamma)\end{pmatrix} \in \mathbb{R}^{p+q}$$

$$Hess\,(\theta) = 2\begin{pmatrix} X' R(\theta_2)^{-1} X & X' R(\theta_2)^{-1} Z \\ Z' R(\theta_2)^{-1} X & G(\theta_1)^{-1} + Z' R(\theta_2)^{-1} Z\end{pmatrix} \in M_{(p+q)\times(p+q)}$$

Solución del modelo LMM I

- Si $\gamma^* = G(\boldsymbol{\theta_1})^{-1/2}\gamma$ y $\varepsilon^* = R(\boldsymbol{\theta_2})^{-1/2}\varepsilon$ tenemos que

$$\begin{pmatrix}\gamma^* \\ \varepsilon^*\end{pmatrix} = \begin{pmatrix}G(\boldsymbol{\theta_1})^{-1/2}\gamma \\ R(\boldsymbol{\theta_2})^{-1/2}\varepsilon\end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, I_{q+n}\right)$$

  y por tanto, para cada $\boldsymbol{\theta}$ fijo, $(\gamma^*)'\gamma^* + (\varepsilon^*)'\varepsilon^*$ debería ser mínimo

- Si $S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = \gamma' G(\boldsymbol{\theta_1})^{-1}\gamma + (\mathbf{y} - X\beta - Z\gamma)' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma)$,

  para cada $\boldsymbol{\theta}$ fijo, el problema es: $\min_{\beta\in\mathbb{R}^p, \gamma\in\mathbb{R}^q} S\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$

- Si derivamos dos veces tenemos:

$$\nabla S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = 2\begin{pmatrix} -X' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma) \\ G(\boldsymbol{\theta_1})^{-1}\gamma - Z' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma)\end{pmatrix} \in \mathbb{R}^{p+q}$$

$$Hess\,(\boldsymbol{\theta}) = 2\begin{pmatrix} X' R(\boldsymbol{\theta_2})^{-1} X & X' R(\boldsymbol{\theta_2})^{-1} Z \\ Z' R(\boldsymbol{\theta_2})^{-1} X & G(\boldsymbol{\theta_1})^{-1} + Z' R(\boldsymbol{\theta_2})^{-1} Z\end{pmatrix} \in M_{(p+q)\times(p+q)}$$

Solución del modelo LMM I

- Si $\gamma^* = G(\boldsymbol{\theta_1})^{-1/2}\gamma$ y $\varepsilon^* = R(\boldsymbol{\theta_2})^{-1/2}\varepsilon$ tenemos que

$$\begin{pmatrix}\gamma^* \\ \varepsilon^*\end{pmatrix} = \begin{pmatrix}G(\boldsymbol{\theta_1})^{-1/2}\gamma \\ R(\boldsymbol{\theta_2})^{-1/2}\varepsilon\end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, I_{q+n}\right)$$

  y por tanto, para cada $\boldsymbol{\theta}$ fijo, $(\gamma^*)'\gamma^* + (\varepsilon^*)'\varepsilon^*$ debería ser mínimo

- Si $S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = \gamma' G(\boldsymbol{\theta_1})^{-1}\gamma + (\mathbf{y} - X\beta - Z\gamma)' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma)$,

  para cada $\boldsymbol{\theta}$ fijo, el problema es: $\min_{\beta\in\mathbb{R}^p, \gamma\in\mathbb{R}^q} S\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$

- Si derivamos dos veces tenemos:

$$\nabla S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = 2\begin{pmatrix} -X' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma) \\ G(\boldsymbol{\theta_1})^{-1}\gamma - Z' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma)\end{pmatrix} \in \mathbb{R}^{p+q}$$

$$Hess\,(\boldsymbol{\theta}) = 2\begin{pmatrix} X' R(\boldsymbol{\theta_2})^{-1} X & X' R(\boldsymbol{\theta_2})^{-1} Z \\ Z' R(\boldsymbol{\theta_2})^{-1} X & G(\boldsymbol{\theta_1})^{-1} + Z' R(\boldsymbol{\theta_2})^{-1} Z\end{pmatrix} \in M_{(p+q)\times(p+q)}$$

Solución del modelo LMM I

- Si $\gamma^* = G(\boldsymbol{\theta_1})^{-1/2}\gamma$ y $\varepsilon^* = R(\boldsymbol{\theta_2})^{-1/2}\varepsilon$ tenemos que

$$\begin{pmatrix}\gamma^* \\ \varepsilon^*\end{pmatrix} = \begin{pmatrix}G(\boldsymbol{\theta_1})^{-1/2}\gamma \\ R(\boldsymbol{\theta_2})^{-1/2}\varepsilon\end{pmatrix} \rightsquigarrow N_{q+n}\left(\mathbf{0}_{q+n}, I_{q+n}\right)$$

  y por tanto, para cada $\boldsymbol{\theta}$ fijo, $(\gamma^*)'\gamma^* + (\varepsilon^*)'\varepsilon^*$ debería ser mínimo

- Si $S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = \gamma' G(\boldsymbol{\theta_1})^{-1}\gamma + (\mathbf{y} - X\beta - Z\gamma)' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma)$,

  para cada $\boldsymbol{\theta}$ fijo, el problema es: $\min_{\beta\in\mathbb{R}^p, \gamma\in\mathbb{R}^q} S\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$

- Si derivamos dos veces tenemos:

$$\nabla S\begin{pmatrix}\beta \\ \gamma\end{pmatrix} = 2\begin{pmatrix} -X' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma) \\ G(\boldsymbol{\theta_1})^{-1}\gamma - Z' R(\boldsymbol{\theta_2})^{-1}(\mathbf{y} - X\beta - Z\gamma)\end{pmatrix} \in \mathbb{R}^{p+q}$$

$$Hess\,(\boldsymbol{\theta}) = 2\begin{pmatrix} X' R(\boldsymbol{\theta_2})^{-1} X & X' R(\boldsymbol{\theta_2})^{-1} Z \\ Z' R(\boldsymbol{\theta_2})^{-1} X & G(\boldsymbol{\theta_1})^{-1} + Z' R(\boldsymbol{\theta_2})^{-1} Z\end{pmatrix} \in M_{(p+q)x(p+q)}$$

Solución del modelo LMM II

- Por tanto, $\nabla S\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 0 \Longleftrightarrow Hess\left(\boldsymbol{\theta}\right)\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 2\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$
- Si $Hess\left(\theta\right)$ tiene inversa la solución para cada θ fijo es:

$$\begin{pmatrix}\hat{\beta}_\theta\\ \hat{\gamma}_\theta\end{pmatrix} = 2\left(Hess\left(\theta\right)\right)^{-1}\begin{pmatrix}X'R(\theta_2)^{-1}\mathbf{y}\\ Z'R(\theta_2)^{-1}\mathbf{y}\end{pmatrix}$$

- Si hacemos $C\left(\theta\right) = \left(X'V\left(\theta\right)^{-1}X\right)^{-1}$ la solución es:

$$\begin{pmatrix}\hat{\beta}_\theta\\ \hat{\gamma}_\theta\end{pmatrix} = \begin{pmatrix}C\left(\theta\right)X'V\left(\theta\right)^{-1}\mathbf{y}\\ G(\theta_1)Z'V\left(\theta\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_\theta\right)\end{pmatrix}$$

- Además, el valor mínimo buscado es:

$$\begin{aligned} S\begin{pmatrix}\hat{\beta}_\theta\\ \hat{\gamma}_\theta\end{pmatrix} &= \left(\mathbf{y}-X\hat{\beta}_\theta\right)'V\left(\theta\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_\theta\right)\\ &= \mathbf{y}'\left(V\left(\theta\right)^{-1}-V\left(\theta\right)^{-1}XC\left(\theta\right)X'V\left(\theta\right)^{-1}\right)\mathbf{y}\end{aligned}$$

Solución del modelo LMM II

- Por tanto, $\nabla S\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 0 \Longleftrightarrow Hess\left(\boldsymbol{\theta}\right)\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 2\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$
- Si $Hess\left(\boldsymbol{\theta}\right)$ tiene inversa la solución para cada $\boldsymbol{\theta}$ fijo es:

$$\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} = 2\left(Hess\left(\boldsymbol{\theta}\right)\right)^{-1}\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$$

- Si hacemos $C\left(\boldsymbol{\theta}\right) = \left(X'V\left(\boldsymbol{\theta}\right)^{-1}X\right)^{-1}$ la solución es:

$$\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} = \begin{pmatrix}C\left(\boldsymbol{\theta}\right)X'V\left(\boldsymbol{\theta}\right)^{-1}\mathbf{y}\\ G(\boldsymbol{\theta_1})Z'V\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)\end{pmatrix}$$

- Además, el valor mínimo buscado es:

$$\begin{aligned} S\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} &= \left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)'V\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)\\ &= \mathbf{y}'\left(V\left(\boldsymbol{\theta}\right)^{-1}-V\left(\boldsymbol{\theta}\right)^{-1}XC\left(\boldsymbol{\theta}\right)X'V\left(\boldsymbol{\theta}\right)^{-1}\right)\mathbf{y}\end{aligned}$$

Solución del modelo LMM II

- Por tanto, $\nabla S\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 0 \Longleftrightarrow Hess\left(\boldsymbol{\theta}\right)\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 2\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$
- Si $Hess\left(\boldsymbol{\theta}\right)$ tiene inversa la solución para cada $\boldsymbol{\theta}$ fijo es:

$$\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} = 2\left(Hess\left(\boldsymbol{\theta}\right)\right)^{-1}\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$$

- Si hacemos $C\left(\boldsymbol{\theta}\right) = \left(X'V\left(\boldsymbol{\theta}\right)^{-1}X\right)^{-1}$ la solución es:

$$\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} = \begin{pmatrix}C\left(\boldsymbol{\theta}\right)X'V\left(\boldsymbol{\theta}\right)^{-1}\mathbf{y}\\ G(\boldsymbol{\theta_1})Z'V\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)\end{pmatrix}$$

- Además, el valor mínimo buscado es:

$$\begin{aligned} S\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} &= \left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)'V\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)\\ &= \mathbf{y}'\left(V\left(\boldsymbol{\theta}\right)^{-1}-V\left(\boldsymbol{\theta}\right)^{-1}XC\left(\boldsymbol{\theta}\right)X'V\left(\boldsymbol{\theta}\right)^{-1}\right)\mathbf{y}\end{aligned}$$

Solución del modelo LMM II

- Por tanto, $\nabla S\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 0 \Longleftrightarrow Hess\left(\boldsymbol{\theta}\right)\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 2\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$
- Si $Hess\left(\boldsymbol{\theta}\right)$ tiene inversa la solución para cada $\boldsymbol{\theta}$ fijo es:

$$\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} = 2\left(Hess\left(\boldsymbol{\theta}\right)\right)^{-1}\begin{pmatrix}X'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\\ Z'R(\boldsymbol{\theta_2})^{-1}\mathbf{y}\end{pmatrix}$$

- Si hacemos $C\left(\boldsymbol{\theta}\right) = \left(X'V\left(\boldsymbol{\theta}\right)^{-1}X\right)^{-1}$ la solución es:

$$\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} = \begin{pmatrix}C\left(\boldsymbol{\theta}\right)X'V\left(\boldsymbol{\theta}\right)^{-1}\mathbf{y}\\ G(\boldsymbol{\theta_1})Z'V\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)\end{pmatrix}$$

- Además, el valor mínimo buscado es:

$$\begin{aligned} S\begin{pmatrix}\hat{\beta}_{\boldsymbol{\theta}}\\ \hat{\gamma}_{\boldsymbol{\theta}}\end{pmatrix} &= \left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)'V\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{y}-X\hat{\beta}_{\boldsymbol{\theta}}\right)\\ &= \mathbf{y}'\left(V\left(\boldsymbol{\theta}\right)^{-1}-V\left(\boldsymbol{\theta}\right)^{-1}XC\left(\boldsymbol{\theta}\right)X'V\left(\boldsymbol{\theta}\right)^{-1}\right)\mathbf{y}\end{aligned}$$

Solución del modelo LMM III

- Con esta solución obtenida, si $L(\boldsymbol{\theta})$ es la función de verosimilitud y $\Lambda(\boldsymbol{\theta})=-2ln(L(\boldsymbol{\theta}))$ se verifica: $\Lambda(\boldsymbol{\theta})=n\ln(2\pi)+\ln|V(\boldsymbol{\theta})|+S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$
- Habitualmente, existen tres formas de estimar el vector θ:
  - Minimizando $S\begin{pmatrix}\hat{\beta}_{\theta}\\ \hat{\gamma}_{\theta}\end{pmatrix}$: estimador insesgado cuadrático de mínima varianza (MIVQUE0, minimum variance quadratic unbiased estimation)
  - Maximizando $L(\theta)$, es decir, minimizando $S\begin{pmatrix}\hat{\beta}_{\theta}\\ \hat{\gamma}_{\theta}\end{pmatrix}+\ln|V(\theta)|$: estimador de máxima verosimilitud (ML, maximum likehood estimation)
  - Minimizando la función $S\begin{pmatrix}\hat{\beta}_{\theta}\\ \hat{\gamma}_{\theta}\end{pmatrix}+\ln|V(\theta)|-\ln|C(\theta)|$: estimador de máxima verosimilitud restringida (REML, restricted maximum likehood estimation)

Solución del modelo LMM III

- Con esta solución obtenida, si $L(\boldsymbol{\theta})$ es la función de verosimilitud y $\Lambda(\boldsymbol{\theta})=-2ln(L(\boldsymbol{\theta}))$ se verifica: $\Lambda(\boldsymbol{\theta})=n\ln(2\pi)+\ln|V(\boldsymbol{\theta})|+S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$
- Habitualmente, existen tres formas de estimar el vector $\boldsymbol{\theta}$:
  - Minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$: estimador insesgado cuadrático de mínima varianza (MIVQUE0, minimum variance quadratic unbiased estimation)
  - Maximizando $L(\boldsymbol{\theta})$, es decir, minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|$: estimador de máxima verosimilitud (ML, maximum likehood estimation)
  - Minimizando la función $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|-\ln|C(\boldsymbol{\theta})|$: estimador de máxima verosimilitud restringida (REML, restricted maximum likehood estimation)

Solución del modelo LMM III

- Con esta solución obtenida, si $L(\boldsymbol{\theta})$ es la función de verosimilitud y $\Lambda(\boldsymbol{\theta})=-2ln(L(\boldsymbol{\theta}))$ se verifica: $\Lambda(\boldsymbol{\theta})=n\ln(2\pi)+\ln|V(\boldsymbol{\theta})|+S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$
- Habitualmente, existen tres formas de estimar el vector $\boldsymbol{\theta}$:
  - Minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$: estimador insesgado cuadrático de mínima varianza (MIVQUE0, minimum variance quadratic unbiased estimation)
  - Maximizando $L(\boldsymbol{\theta})$, es decir, minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|$: estimador de máxima verosimilitud (ML, maximum likehood estimation)
  - Minimizando la función $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|-\ln|C(\boldsymbol{\theta})|$: estimador de máxima verosimilitud restringida (REML, restricted maximum likehood estimation)

Solución del modelo LMM III

- Con esta solución obtenida, si $L(\boldsymbol{\theta})$ es la función de verosimilitud y $\Lambda(\boldsymbol{\theta})=-2ln(L(\boldsymbol{\theta}))$ se verifica: $\Lambda(\boldsymbol{\theta})=n\ln(2\pi)+\ln|V(\boldsymbol{\theta})|+S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$
- Habitualmente, existen tres formas de estimar el vector $\boldsymbol{\theta}$:
  - Minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$: estimador insesgado cuadrático de mínima varianza (MIVQUE0, minimum variance quadratic unbiased estimation)
  - Maximizando $L(\boldsymbol{\theta})$, es decir, minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|$: estimador de máxima verosimilitud (ML, maximum likehood estimation)
  - Minimizando la función $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|-\ln|C(\boldsymbol{\theta})|$: estimador de máxima verosimilitud restringida (REML, restricted maximum likehood estimation)

Solución del modelo LMM III

- Con esta solución obtenida, si $L(\boldsymbol{\theta})$ es la función de verosimilitud y $\Lambda(\boldsymbol{\theta})=-2ln(L(\boldsymbol{\theta}))$ se verifica: $\Lambda(\boldsymbol{\theta})=n\ln(2\pi)+\ln|V(\boldsymbol{\theta})|+S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$
- Habitualmente, existen tres formas de estimar el vector $\boldsymbol{\theta}$:
  - Minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}$: estimador insesgado cuadrático de mínima varianza (MIVQUE0, minimum variance quadratic unbiased estimation)
  - Maximizando $L(\boldsymbol{\theta})$, es decir, minimizando $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|$: estimador de máxima verosimilitud (ML, maximum likehood estimation)
  - Minimizando la función $S\begin{pmatrix}\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}\\ \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}\end{pmatrix}+\ln|V(\boldsymbol{\theta})|-\ln|C(\boldsymbol{\theta})|$: estimador de máxima verosimilitud restringida (REML, restricted maximum likehood estimation)

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\theta}} = C(\hat{\theta})X'V(\hat{\theta})^{-1}\mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\theta}} = G(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\theta})^{-1}(\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\theta}) = XC(\hat{\theta})X'V(\hat{\theta})^{-1}$ y $P(\hat{\theta}) = ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\theta})^{-1} + R(\hat{\boldsymbol{\theta}}_2)V(\hat{\theta})^{-1}H(\hat{\theta})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\theta})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\theta})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\theta})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\theta})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}}) X' V(\hat{\boldsymbol{\theta}})^{-1} \mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1) Z' V(\hat{\boldsymbol{\theta}})^{-1} (\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}}) X' V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1) Z' V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2) V(\hat{\boldsymbol{\theta}})^{-1} H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}}) X' V(\hat{\boldsymbol{\theta}})^{-1} \mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1) Z' V(\hat{\boldsymbol{\theta}})^{-1} (\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}}) X' V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1) Z' V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2) V(\hat{\boldsymbol{\theta}})^{-1} H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}(\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2)V(\hat{\boldsymbol{\theta}})^{-1}H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}(\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2)V(\hat{\boldsymbol{\theta}})^{-1}H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}(\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2)V(\hat{\boldsymbol{\theta}})^{-1}H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}(\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2)V(\hat{\boldsymbol{\theta}})^{-1}H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Solución del modelo LMM IV

- Sea $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_1 \\ \hat{\boldsymbol{\theta}}_2 \end{pmatrix}$ el estimador obtenido para el vector $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix}$
- La solución del modelo es entonces:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}} = C(\hat{\boldsymbol{\theta}}) X' V(\hat{\boldsymbol{\theta}})^{-1} \mathbf{y}$$
$$\hat{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}} = G(\hat{\boldsymbol{\theta}}_1) Z' V(\hat{\boldsymbol{\theta}})^{-1} (\mathbf{y} - X\hat{\boldsymbol{\beta}})$$

- Si consideramos las matrices $H(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}}) X' V(\hat{\boldsymbol{\theta}})^{-1}$ y $P(\hat{\boldsymbol{\theta}}) = ZG(\hat{\boldsymbol{\theta}}_1) Z' V(\hat{\boldsymbol{\theta}})^{-1} + R(\hat{\boldsymbol{\theta}}_2) V(\hat{\boldsymbol{\theta}})^{-1} H(\hat{\boldsymbol{\theta}})$
- El vector de valores predichos es: $\hat{\mathbf{y}} = X\hat{\boldsymbol{\beta}} + Z\hat{\boldsymbol{\gamma}} = P(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores es: $\mathbf{e} = \mathbf{y} - X\hat{\boldsymbol{\beta}} - Z\hat{\boldsymbol{\gamma}} = \left(I_n - P(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- El vector de valores predichos marginales es: $\hat{\mathbf{y}}_m = X\hat{\boldsymbol{\beta}} = H(\hat{\boldsymbol{\theta}})\mathbf{y}$
- El vector de errores marginales es: $\mathbf{e}_m = \mathbf{y} - X\hat{\boldsymbol{\beta}} = \left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$
- Los vectores $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{y}}$, $\mathbf{e}$, $\hat{\mathbf{y}}_m$ y $\mathbf{e}_m$ tienen todos distribución normal porque $\mathbf{y}/X, Z \rightsquigarrow N_n(X\boldsymbol{\beta}, V(\boldsymbol{\theta}))$

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\theta)$ cuanto menor sea $\Lambda(\hat{\theta}) = -2ln(L((\hat{\theta}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\theta}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\theta}) + d\,\ln(n^*)$ con d igual que antes y $n^*=$ sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\theta}) + \frac{2dn}{n-d-1}$
- Si $Z{=}0$ tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\theta){=}R(\theta_2){=}\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\theta}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\theta}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\theta}) + d\,\ln(n^*)$ con d igual que antes y $n^*=$ sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\theta}) + \dfrac{2dn}{n-d-1}$
- Si $Z=0$ tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\theta)=R(\theta_2)=\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\theta}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\boldsymbol{\theta}}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\boldsymbol{\theta}}) + d\,\ln(n^*)$ con d igual que antes y n^*= sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\boldsymbol{\theta}}) + \dfrac{2dn}{n - d - 1}$
- Si Z=0 tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\boldsymbol{\theta})$=$R(\boldsymbol{\theta_2})$=$\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\boldsymbol{\theta}}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\boldsymbol{\theta}}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\boldsymbol{\theta}}) + d\,\ln(n^*)$ con d igual que antes y n^*= sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\theta}) + \dfrac{2dn}{n-d-1}$
- Si Z=0 tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\theta)$=$R(\theta_2)$=$\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\theta}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\boldsymbol{\theta}}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\boldsymbol{\theta}}) + d\,\ln(n^*)$ con d igual que antes y n^*= sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\boldsymbol{\theta}}) + \dfrac{2dn}{n-d-1}$
- Si Z=0 tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\boldsymbol{\theta})$=$R(\boldsymbol{\theta_2})$=$\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\boldsymbol{\theta}}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\boldsymbol{\theta}}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\boldsymbol{\theta}}) + d\,\ln(n^*)$ con d igual que antes y n^*= sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\boldsymbol{\theta}}) + \dfrac{2dn}{n-d-1}$
- Si Z=0 tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\boldsymbol{\theta})$=$R(\boldsymbol{\theta}_2)$=$\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\boldsymbol{\theta}}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\boldsymbol{\theta}}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\boldsymbol{\theta}}) + d\,\ln(n^*)$ con d igual que antes y n^*= sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\boldsymbol{\theta}}) + \dfrac{2dn}{n-d-1}$
- Si Z=0 tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\boldsymbol{\theta})$=$R(\boldsymbol{\theta_2})$=$\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\boldsymbol{\theta}}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Adecuación del modelo LMM

- No disponemos de un test global ni de coeficiente de determinación
- Mayor verosimilitud $L(\boldsymbol{\theta})$ cuanto menor sea $\Lambda(\hat{\boldsymbol{\theta}}) = -2ln(L((\hat{\boldsymbol{\theta}}))$
- Se usa el Criterio de Información de Akaike: $AIC = \Lambda(\hat{\boldsymbol{\theta}}) + 2d$ con $d = t$ si usamos REML o $d = p + t$ si usamos ML
- También el Criterio de Información de Bayes: $BIC = \Lambda(\hat{\boldsymbol{\theta}}) + d\,\ln(n^*)$ con d igual que antes y $n^*=$ sujetos efectivos (depende del modelo)
- Mejor modelos y estructuras de varianza con menor AIC y menor BIC
- También AIC Modificado: $AICC = \Lambda(\hat{\boldsymbol{\theta}}) + \dfrac{2dn}{n-d-1}$
- Si $Z=0$ tenemos Test de la Razón de Verosimilitud del Modelo Nulo con $V(\boldsymbol{\theta})=R(\boldsymbol{\theta_2})=\sigma^2 I_n$, usando el estadístico $\Lambda(\hat{\boldsymbol{\theta}}) - \Lambda(\hat{\sigma}^2) \rightsquigarrow \chi^2_{t-1}$
- Se puede utilizar el Coeficiente de Determinación ($pseudo{-}R^2$) entre observados ($\mathbf{y}$) y predichos ($\hat{\mathbf{y}}$) (o predichos marginales $\hat{\mathbf{y}}_m$)

Test e intervalos de confianza de parámetros y predicciones

- Intervalo de confianza para β_{j-1} con $j=1,...,p$: $\hat{\beta}_{j-1} \pm t_{df;\alpha/2}\sqrt{c_{jj}}$ siendo c_{jj} el elemento diagonal j-ésimo de la matriz $C(\hat{\boldsymbol{\theta}})$ y df calculado en cada caso según el modelo (intervalo aproximado)
- Si $H_0 : \beta_{j-1} = 0$ es cierta, $\frac{\hat{\beta}_{j-1}}{\sqrt{c_{jj}}} \rightsquigarrow t_{df}$ y tenemos un test para H_0
- Intervalo de confianza para y_i con $i=1,...,n$: $\hat{y}_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $P(\hat{\boldsymbol{\theta}})V(\hat{\boldsymbol{\theta}})P'(\hat{\boldsymbol{\theta}})$
- Intervalo de confianza para $(y_m)_i$ con $i=1,...,n$: $(\hat{y}_m)_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $XC(\hat{\boldsymbol{\theta}})X'$
- Intervalo de confianza de Wald para θ_τ con $\tau=1,...,t$: $\hat{\theta}_\tau \pm z_{\alpha/2}\sqrt{d_{\tau\tau}}$ siendo $d_{\tau\tau}$ el elemento diagonal τ-ésimo de la inversa de la matriz de información de Fisher (matriz hessiana de $\Lambda(\hat{\boldsymbol{\theta}})$)

Test e intervalos de confianza de parámetros y predicciones

- Intervalo de confianza para β_{j-1} con $j=1,...,p$: $\hat{\beta}_{j-1} \pm t_{df;\alpha/2}\sqrt{c_{jj}}$ siendo c_{jj} el elemento diagonal j-ésimo de la matriz $C(\hat{\boldsymbol{\theta}})$ y df calculado en cada caso según el modelo (intervalo aproximado)
- Si $H_0 : \beta_{j-1} = 0$ es cierta, $\dfrac{\hat{\beta}_{j-1}}{\sqrt{c_{jj}}} \rightsquigarrow t_{df}$ y tenemos un test para H_0
- Intervalo de confianza para y_i con $i=1,...,n$: $\hat{y}_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $P(\hat{\boldsymbol{\theta}})V(\hat{\boldsymbol{\theta}})P'(\hat{\boldsymbol{\theta}})$
- Intervalo de confianza para $(y_m)_i$ con $i=1,...,n$: $(\hat{y}_m)_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $XC(\hat{\boldsymbol{\theta}})X'$
- Intervalo de confianza de Wald para θ_τ con $\tau=1,...,t$: $\hat{\theta}_\tau \pm z_{\alpha/2}\sqrt{d_{\tau\tau}}$ siendo $d_{\tau\tau}$ el elemento diagonal τ-ésimo de la inversa de la matriz de información de Fisher (matriz hessiana de $\Lambda(\hat{\boldsymbol{\theta}})$)

Test e intervalos de confianza de parámetros y predicciones

- Intervalo de confianza para β_{j-1} con $j=1,...,p$: $\hat{\beta}_{j-1} \pm t_{df;\alpha/2}\sqrt{c_{jj}}$ siendo c_{jj} el elemento diagonal j-ésimo de la matriz $C(\hat{\boldsymbol{\theta}})$ y df calculado en cada caso según el modelo (intervalo aproximado)
- Si $H_0 : \beta_{j-1} = 0$ es cierta, $\frac{\hat{\beta}_{j-1}}{\sqrt{c_{jj}}} \rightsquigarrow t_{df}$ y tenemos un test para H_0
- Intervalo de confianza para y_i con $i=1,...,n$: $\hat{y}_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $P(\hat{\boldsymbol{\theta}})V(\hat{\boldsymbol{\theta}})P'(\hat{\boldsymbol{\theta}})$
- Intervalo de confianza para $(y_m)_i$ con $i=1,...,n$: $(\hat{y}_m)_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $XC(\hat{\boldsymbol{\theta}})X'$
- Intervalo de confianza de Wald para θ_τ con $\tau=1,...,t$: $\hat{\theta}_\tau \pm z_{\alpha/2}\sqrt{d_{\tau\tau}}$ siendo $d_{\tau\tau}$ el elemento diagonal τ-ésimo de la inversa de la matriz de información de Fisher (matriz hessiana de $\Lambda(\hat{\theta})$)

Test e intervalos de confianza de parámetros y predicciones

- Intervalo de confianza para β_{j-1} con $j=1,...,p$: $\hat{\beta}_{j-1} \pm t_{df;\alpha/2}\sqrt{c_{jj}}$ siendo c_{jj} el elemento diagonal j-ésimo de la matriz $C(\hat{\boldsymbol{\theta}})$ y df calculado en cada caso según el modelo (intervalo aproximado)
- Si $H_0 : \beta_{j-1} = 0$ es cierta, $\frac{\hat{\beta}_{j-1}}{\sqrt{c_{jj}}} \rightsquigarrow t_{df}$ y tenemos un test para H_0
- Intervalo de confianza para y_i con $i=1,...,n$: $\hat{y}_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $P(\hat{\boldsymbol{\theta}})V(\hat{\boldsymbol{\theta}})P'(\hat{\boldsymbol{\theta}})$
- Intervalo de confianza para $(y_m)_i$ con $i=1,...,n$: $(\hat{y}_m)_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $XC(\hat{\boldsymbol{\theta}})X'$
- Intervalo de confianza de Wald para θ_τ con $\tau=1,...,t$: $\hat{\theta}_\tau \pm z_{\alpha/2}\sqrt{d_{\tau\tau}}$ siendo $d_{\tau\tau}$ el elemento diagonal τ-ésimo de la inversa de la matriz de información de Fisher (matriz hessiana de $\Lambda(\hat{\theta})$)

Test e intervalos de confianza de parámetros y predicciones

- Intervalo de confianza para β_{j-1} con $j=1,...,p$: $\hat{\beta}_{j-1} \pm t_{df;\alpha/2}\sqrt{c_{jj}}$ siendo c_{jj} el elemento diagonal j-ésimo de la matriz $C(\hat{\boldsymbol{\theta}})$ y df calculado en cada caso según el modelo (intervalo aproximado)
- Si $H_0 : \beta_{j-1} = 0$ es cierta, $\frac{\hat{\beta}_{j-1}}{\sqrt{c_{jj}}} \rightsquigarrow t_{df}$ y tenemos un test para H_0
- Intervalo de confianza para y_i con $i=1,...,n$: $\hat{y}_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $P(\hat{\boldsymbol{\theta}})V(\hat{\boldsymbol{\theta}})P'(\hat{\boldsymbol{\theta}})$
- Intervalo de confianza para $(y_m)_i$ con $i=1,...,n$: $(\hat{y}_m)_i \pm t_{df;\alpha/2}\sqrt{d_{ii}}$ siendo d_{ii} el elemento diagonal i-ésimo de la matriz $XC(\hat{\boldsymbol{\theta}})X'$
- Intervalo de confianza de Wald para θ_τ con $\tau=1,...,t$: $\hat{\theta}_\tau \pm z_{\alpha/2}\sqrt{d_{\tau\tau}}$ siendo $d_{\tau\tau}$ el elemento diagonal τ-ésimo de la inversa de la matriz de información de Fisher (matriz hessiana de $\Lambda(\hat{\boldsymbol{\theta}})$)

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu xp}$ y $rg(L) = \nu < p$, entonces $F = \nu^{-1}\hat{\beta}' L' (LC(\hat{\boldsymbol{\theta}})L')^{-1} L\hat{\beta} \rightsquigarrow F_{\nu,df}$ con df calculado en cada caso según el modelo, y tenemos test de hipótesis para H_0
- Los test F en la tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\sqrt{LC(\hat{\boldsymbol{\theta}})L'}} \rightsquigarrow t_{df}$, y entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $L\beta$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}} \rightsquigarrow t_{df}$ y para $L_1\beta - L_2\beta$ el intervalo de confianza es $(L_1 - L_2)\hat{\beta} \pm t_{df;\alpha/2}\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu xp}$ y $rg(L) = \nu < p$, entonces $F = \nu^{-1}\hat{\beta}' L' (LC(\hat{\boldsymbol{\theta}})L')^{-1} L\hat{\beta} \rightsquigarrow F_{\nu,df}$ con df calculado en cada caso según el modelo, y tenemos test de hipótesis para H_0
- Los test F en la tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\sqrt{LC(\hat{\boldsymbol{\theta}})L'}} \rightsquigarrow t_{df}$, y entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $L\beta$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}} \rightsquigarrow t_{df}$ y para $L_1\beta - L_2\beta$ el intervalo de confianza es $(L_1 - L_2)\hat{\beta} \pm t_{df;\alpha/2}\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu xp}$ y $rg(L) = \nu < p$, entonces $F = \nu^{-1}\hat{\beta}' L' (LC(\hat{\boldsymbol{\theta}})L')^{-1} L\hat{\beta} \rightsquigarrow F_{\nu,df}$ con df calculado en cada caso según el modelo, y tenemos test de hipótesis para H_0
- Los test F en la tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\sqrt{LC(\hat{\boldsymbol{\theta}})L'}} \rightsquigarrow t_{df}$, y entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $L\beta$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}} \rightsquigarrow t_{df}$ y para $L_1\beta - L_2\beta$ el intervalo de confianza es $(L_1 - L_2)\hat{\beta} \pm t_{df;\alpha/2}\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu xp}$ y $rg(L) = \nu < p$, entonces $F = \nu^{-1}\hat{\beta}' L' (LC(\hat{\boldsymbol{\theta}})L')^{-1} L\hat{\beta} \rightsquigarrow F_{\nu,df}$ con df calculado en cada caso según el modelo, y tenemos test de hipótesis para H_0
- Los test F en la tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\sqrt{LC(\hat{\boldsymbol{\theta}})L'}} \rightsquigarrow t_{df}$, y entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $L\beta$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}} \rightsquigarrow t_{df}$ y para $L_1\beta - L_2\beta$ el intervalo de confianza es $(L_1 - L_2)\hat{\beta} \pm t_{df;\alpha/2}\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \nu^{-1} \hat{\beta}' L' (LC(\hat{\boldsymbol{\theta}})L')^{-1} L\hat{\beta} \rightsquigarrow F_{\nu, df}$ con df calculado en cada caso según el modelo, y tenemos test de hipótesis para H_0
- Los test F en la tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\sqrt{LC(\hat{\boldsymbol{\theta}})L'}} \rightsquigarrow t_{df}$, y entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $L\beta$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}} \rightsquigarrow t_{df}$ y para $L_1\beta - L_2\beta$ el intervalo de confianza es $(L_1 - L_2)\hat{\beta} \pm t_{df;\alpha/2}\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Test F, intervalos de confianza y Ls-medias

- En general, si $H_0: L\beta = 0$ con $L \in M_{\nu x p}$ y $rg(L) = \nu < p$, entonces $F = \nu^{-1} \hat{\beta}' L' (LC(\hat{\boldsymbol{\theta}})L')^{-1} L\hat{\beta} \rightsquigarrow F_{\nu, df}$ con df calculado en cada caso según el modelo, y tenemos test de hipótesis para H_0
- Los test F en la tabla ANOVA son casos particulares con ciertas L
- Si $L \in M_{1xp}$, $t = \frac{L\hat{\beta} - L\beta}{\sqrt{LC(\hat{\boldsymbol{\theta}})L'}} \rightsquigarrow t_{df}$, y entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $L\beta$
- Si $L_1, L_2 \in M_{1xp}$, $t = \frac{(L_1 - L_2)\hat{\beta} - (L_1 - L_2)\beta}{\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}} \rightsquigarrow t_{df}$ y para $L_1\beta - L_2\beta$ el intervalo de confianza es $(L_1 - L_2)\hat{\beta} \pm t_{df;\alpha/2}\sqrt{(L_1 - L_2)C(\hat{\boldsymbol{\theta}})(L_1 - L_2)'}$
- Los test e intervalos de confianza para Ls-medias y sus diferencias en un ANOVA son casos particulares con ciertas matrices L_1 y L_2
- Si $L = (1, x_1, ..., x_k) \in M_{1xp}$ entonces $L\hat{\beta} \pm t_{df;\alpha/2}\sqrt{LC(\hat{\boldsymbol{\theta}})L'}$ es un intervalo de confianza para $E(y/x_1, x_2, ..., x_k)$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\theta})\left(V(\hat{\theta}) - Q(\hat{\theta})\right)K'(\hat{\theta})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\theta}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\theta}) - Q(\hat{\theta})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\theta})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\theta})^{-1/2}\mathbf{e}_m = V(\hat{\theta})^{-1/2}\left(I_n - H(\hat{\theta})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\theta})\left(V(\hat{\theta}) - Q(\hat{\theta})\right)K'(\hat{\theta})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\theta}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\theta}) - Q(\hat{\theta})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\theta})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\theta})^{-1/2}\mathbf{e}_m = V(\hat{\theta})^{-1/2}\left(I_n - H(\hat{\theta})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\boldsymbol{\theta}})\left(V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})\right)K'(\hat{\boldsymbol{\theta}})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\boldsymbol{\theta}}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\boldsymbol{\theta}})^{-1/2}\mathbf{e}_m = V(\hat{\boldsymbol{\theta}})^{-1/2}\left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\boldsymbol{\theta}})\left(V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})\right)K'(\hat{\boldsymbol{\theta}})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\boldsymbol{\theta}}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\boldsymbol{\theta}})^{-1/2}\mathbf{e}_m = V(\hat{\boldsymbol{\theta}})^{-1/2}\left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\boldsymbol{\theta}})\left(V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})\right)K'(\hat{\boldsymbol{\theta}})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\boldsymbol{\theta}}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\boldsymbol{\theta}})^{-1/2}\mathbf{e}_m = V(\hat{\boldsymbol{\theta}})^{-1/2}\left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\boldsymbol{\theta}})\left(V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})\right)K'(\hat{\boldsymbol{\theta}})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\boldsymbol{\theta}}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\boldsymbol{\theta}})^{-1/2}\mathbf{e}_m = V(\hat{\boldsymbol{\theta}})^{-1/2}\left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\boldsymbol{\theta}})\left(V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})\right)K'(\hat{\boldsymbol{\theta}})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\boldsymbol{\theta}}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\boldsymbol{\theta}})^{-1/2}\mathbf{e}_m = V(\hat{\boldsymbol{\theta}})^{-1/2}\left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$

Tipos de residuales en el modelo LMM

- Sean las matrices $Q(\hat{\boldsymbol{\theta}}) = XC(\hat{\boldsymbol{\theta}})X'$ y $K(\hat{\boldsymbol{\theta}}) = I_n - ZG(\hat{\boldsymbol{\theta}}_1)Z'V(\hat{\boldsymbol{\theta}})^{-1}$
- Residuales condicionales $r_i = e_i$
  - Studentizados: $r_i^{student} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $K(\hat{\boldsymbol{\theta}})\left(V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})\right)K'(\hat{\boldsymbol{\theta}})$
  - Pearson: $r_i^{pearson} = \dfrac{r_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $R(\hat{\boldsymbol{\theta}}_2)$
- Residuales marginales $(r_m)_i = (e_m)_i$
  - Studentizados: $(r_m^{student})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}})$
  - Pearson: $(r_m^{pearson})_i = \dfrac{(r_m)_i}{\sqrt{d_{ii}}}$ con d_{ii}=elemento diagonal i-ésimo de $V(\hat{\boldsymbol{\theta}})$
  - Escalados: $(r_{m(c)})_i$. Son las componentes del vector $\mathbf{e}_{m(c)} = V(\hat{\boldsymbol{\theta}})^{-1/2}\mathbf{e}_m = V(\hat{\boldsymbol{\theta}})^{-1/2}\left(I_n - H(\hat{\boldsymbol{\theta}})\right)\mathbf{y}$

Validación del modelo LMM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales condicionales studentizados $r_i^{student}$ (también con los marginales $(r_m^{student})$)
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i^{student})$ y $(x_{ij}, r_i^{student})$ para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: coeficiente de autocorrelación de residuales escalados $(r_{m(c)})_i$ y nube de dispersión $(i, (r_{m(c)})_i)$
- Posibles puntos influyentes: aquellos con D de Cook's $> 4/n$, diagonal hat $h_{ii} > 2p/n$ o residual studentizado mayor que 2 (o mejor 3)
- Existencia de multicolinealidad: analizarlo previamente utilizando un modelo GLM
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$ y su recta de regresión para evaluar $pseudo-R^2$

Validación del modelo LMM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales condicionales studentizados $r_i^{student}$ (también con los marginales $(r_m^{student})$)
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i^{student})$ y $(x_{ij}, r_i^{student})$ para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: coeficiente de autocorrelación de residuales escalados $(r_{m(c)})_i$ y nube de dispersión $(i, (r_{m(c)})_i)$
- Posibles puntos influyentes: aquellos con D de Cook's $> 4/n$, diagonal hat $h_{ii} > 2p/n$ o residual studentizado mayor que 2 (o mejor 3)
- Existencia de multicolinealidad: analizarlo previamente utilizando un modelo GLM
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$ y su recta de regresión para evaluar $pseudo-R^2$

Validación del modelo LMM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales condicionales studentizados $r_i^{student}$ (también con los marginales $(r_m^{student})$)
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i^{student})$ y $(x_{ij}, r_i^{student})$ para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: coeficiente de autocorrelación de residuales escalados $(r_{m(c)})_i$ y nube de dispersión $(i, (r_{m(c)})_i)$
- Posibles puntos influyentes: aquellos con D de Cook's $> 4/n$, diagonal hat $h_{ii} > 2p/n$ o residual studentizado mayor que 2 (o mejor 3)
- Existencia de multicolinealidad: analizarlo previamente utilizando un modelo GLM
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$ y su recta de regresión para evaluar $pseudo-R^2$

Validación del modelo LMM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales condicionales studentizados $r_i^{student}$ (también con los marginales $(r_m^{student})$)
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i^{student})$ y $(x_{ij}, r_i^{student})$ para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: coeficiente de autocorrelación de residuales escalados $(r_{m(c)})_i$ y nube de dispersión $(i, (r_{m(c)})_i)$
- Posibles puntos influyentes: aquellos con D de Cook's $> 4/n$, diagonal hat $h_{ii} > 2p/n$ o residual studentizado mayor que 2 (o mejor 3)
- Existencia de multicolinealidad: analizarlo previamente utilizando un modelo GLM
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$ y su recta de regresión para evaluar $pseudo-R^2$

Validación del modelo LMM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales condicionales studentizados $r_i^{student}$ (también con los marginales $(r_m^{student})$)
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i^{student})$ y $(x_{ij}, r_i^{student})$ para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: coeficiente de autocorrelación de residuales escalados $(r_{m(c)})_i$ y nube de dispersión $(i, (r_{m(c)})_i)$
- Posibles puntos influyentes: aquellos con D de Cook's $> 4/n$, diagonal hat $h_{ii} > 2p/n$ o residual studentizado mayor que 2 (o mejor 3)
- Existencia de multicolinealidad: analizarlo previamente utilizando un modelo GLM
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$ y su recta de regresión para evaluar $pseudo-R^2$

Validación del modelo LMM

- Normalidad de residuales: plot de probabilidad normal y test de Kolmogorov con los residuales condicionales studentizados $r_i^{student}$ (también con los marginales $(r_m^{student})$)
- Homegeneidad de varianzas: nubes de dispersión de $(\widehat{y}_i, r_i^{student})$ y $(x_{ij}, r_i^{student})$ para $i = 1, ..., n$ y $j = 1, ..., k$
- Independencia de residuales: coeficiente de autocorrelación de residuales escalados $(r_{m(c)})_i$ y nube de dispersión $(i, (r_{m(c)})_i)$
- Posibles puntos influyentes: aquellos con D de Cook's $> 4/n$, diagonal hat $h_{ii} > 2p/n$ o residual studentizado mayor que 2 (o mejor 3)
- Existencia de multicolinealidad: analizarlo previamente utilizando un modelo GLM
- Falta de adecuación del modelo: nube de dispersión $(y_i, \widehat{y}_i)$ y su recta de regresión para evaluar $pseudo-R^2$

Software para el modelo LMM I

- Si usamos el software R disponemos del paquete lme4
- Si usamos el software SAS tenemos el procedimiento proc mixed
- Sintaxis básica general de proc mixed:

```
ods pdf file="...";ods graphics on; (opcional)
PROC MIXED method=REML,ML,MIVQUE0 options;
    CLASS factores;
    MODEL dependent=fixed effects / options;
    RANDOM random-effects / type=vc group=variable options;
    REPEATED effects / subject=var type=cs group=var options;
    PARMS value list/options;
    CONTRAST 'label' [effect values].../options;
    ESTIMATE 'name' effect values.../options;
    LSMEANS effects/options;
RUN;
ods graphics off;ods pdf close;(si usamos los dos primeros comandos)
QUIT;
```

Software para el modelo LMM I

- Si usamos el software R disponemos del paquete lme4
- Si usamos el software SAS tenemos el procedimiento proc mixed
- Sintaxis básica general de proc mixed:

```
ods pdf file="...";ods graphics on; (opcional)
PROC MIXED method=REML,ML,MIVQUE0 options;
    CLASS factores;
    MODEL dependent=fixed effects / options;
    RANDOM random-effects / type=vc group=variable options;
    REPEATED effects / subject=var type=cs group=var options;
    PARMS value list/options;
    CONTRAST 'label' [effect values].../options;
    ESTIMATE 'name' effect values.../options;
    LSMEANS effects/options;
RUN;
ods graphics off;ods pdf close;(si usamos los dos primeros comandos)
QUIT;
```

Software para el modelo LMM I

- Si usamos el software R disponemos del paquete lme4
- Si usamos el software SAS tenemos el procedimiento proc mixed
- Sintaxis básica general de proc mixed:

```
ods pdf file="...";ods graphics on; (opcional)
PROC MIXED method=REML,ML,MIVQUE0 options;
    CLASS factores;
    MODEL dependent=fixed effects / options;
    RANDOM random-effects / type=vc group=variable options;
    REPEATED effects / subject=var type=cs group=var options;
    PARMS value list/options;
    CONTRAST 'label' [effect values].../options;
    ESTIMATE 'name' effect values.../options;
    LSMEANS effects/options;
RUN;
ods graphics off;ods pdf close;(si usamos los dos primeros comandos)
QUIT;
```

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Software para el modelo LMM II

Observaciones importantes si usamos PROC MIXED de SAS:

- Los efectos aleatorios no deben incluirse en el comando MODEL
- Puede no haber factores de medidas repetidas en REPEATED o no haber efectos aleatorios en RANDOM
- Sólo puede haber uno o dos factores de medidas repetidas
- Si hay factores de medidas repetidas, OJO CON EL ORDEN DE LOS DATOS EN EL FICHERO
- No hay comando OUTPUT OUT, se deben usar las opciones OUTP o OUTPM dentro del comando MODEL despues de /, habitualmente con la opción RESIDUALS
- En el comando LSMEANS, sólo factores incluidos en CLASS

Principales estructuras de varianza con SAS I

Variance Components (VC) (Componentes de Varianza)

Estructura por defecto para RANDOM y REPEATED

Si tenemos t_1 factores aleatorios y $n_1, .., n_{t_1}$ son los niveles de cada factor, con $n_1 + .. + n_{t_1} = q$, entonces $\boldsymbol{\theta_1} = \left(\sigma_1^2, \sigma_2^2, .., \sigma_{t_1}^2\right)' \in \mathbb{R}^{t_1}$ y

$$G\left(\boldsymbol{\theta_1}\right) = \begin{pmatrix} \sigma_1^2 I_{n_1} & 0_{n_1 x n_2} & . & 0_{n_1 x n_{t_1}} \\ 0_{n_2 x n_1} & \sigma_2^2 I_{n_2} & . & 0_{n_2 x n_{t_1}} \\ . & . & . & . \\ 0_{n_{t_1} x n_1} & 0_{n_{t_1} x n_2} & . & \sigma_{t_1}^2 I_{n_{t_1}} \end{pmatrix} \in M_{qxq}$$

Cada factor tiene una varianza distinta

Los factores aleatorios y los niveles dentro de cada factor son independientes

Principales estructuras de varianza con SAS II

Compound Simmetry (CS) (Simetría Compuesta)
La más usada en REPEATED. Si hay s sujetos con r medidas repetidas de cada sujeto ($n = rs$), la matriz de varianzas-covarianzas en cada sujeto es:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma^2+\sigma_1 & \sigma_1 & . & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & . & \sigma_1 \\ . & . & . & . \\ \sigma_1 & \sigma_1 & . & \sigma^2+\sigma_1 \end{pmatrix} = \sigma^2 I_r + \sigma_1 1_{rxr} \in M_{rxr}$$

La matriz $R(\theta_2)$ se define entonces como:

$$R(\boldsymbol{\theta_2}) = \begin{pmatrix} R_1(\boldsymbol{\theta_2}) & 0_{rxr} & . & 0_{rxr} \\ 0_{rxr} & R_1(\boldsymbol{\theta_2}) & . & . \\ . & . & . & 0_{rxr} \\ 0_{rxr} & 0_{rxr} & . & R_1(\boldsymbol{\theta_2}) \end{pmatrix} = R_1(\boldsymbol{\theta_2}) \otimes I_s \in M_{nxn}$$

Si $\sigma_1 = 0$ coincide con **VC**, y si $Z = 0$ tenemos $V(\theta) = \sigma^2 I_n$ (modelo GLM). Por ello la opción **VC** es el valor por defecto en REPEATED

Principales estructuras de varianza con SAS II

Compound Simmetry (CS) (Simetría Compuesta)
La más usada en REPEATED. Si hay s sujetos con r medidas repetidas de cada sujeto ($n = rs$), la matriz de varianzas-covarianzas en cada sujeto es:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma^2+\sigma_1 & \sigma_1 & . & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & . & \sigma_1 \\ . & . & . & . \\ \sigma_1 & \sigma_1 & . & \sigma^2+\sigma_1 \end{pmatrix} = \sigma^2 I_r + \sigma_1 1_{rxr} \in M_{rxr}$$

La matriz $R(\boldsymbol{\theta_2})$ se define entonces como:

$$R(\boldsymbol{\theta_2}) = \begin{pmatrix} R_1(\boldsymbol{\theta_2}) & 0_{rxr} & . & 0_{rxr} \\ 0_{rxr} & R_1(\boldsymbol{\theta_2}) & . & . \\ . & . & . & 0_{rxr} \\ 0_{rxr} & 0_{rxr} & . & R_1(\boldsymbol{\theta_2}) \end{pmatrix} = R_1(\boldsymbol{\theta_2}) \otimes I_s \in M_{nxn}$$

Si $\sigma_1 = 0$ coincide con **VC**, y si $Z = 0$ tenemos $V(\theta) = \sigma^2 I_n$ (modelo GLM). Por ello la opción **VC** es el valor por defecto en REPEATED

Principales estructuras de varianza con SAS II

Compound Simmetry (CS) (Simetría Compuesta)
La más usada en REPEATED. Si hay s sujetos con r medidas repetidas de cada sujeto ($n = rs$), la matriz de varianzas-covarianzas en cada sujeto es:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma^2 + \sigma_1 & \sigma_1 & . & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & . & \sigma_1 \\ . & . & . & . \\ \sigma_1 & \sigma_1 & . & \sigma^2 + \sigma_1 \end{pmatrix} = \sigma^2 I_r + \sigma_1 1_{rxr} \in M_{rxr}$$

La matriz $R(\boldsymbol{\theta_2})$ se define entonces como:

$$R(\boldsymbol{\theta_2}) = \begin{pmatrix} R_1(\boldsymbol{\theta_2}) & 0_{rxr} & . & 0_{rxr} \\ 0_{rxr} & R_1(\boldsymbol{\theta_2}) & . & . \\ . & . & . & 0_{rxr} \\ 0_{rxr} & 0_{rxr} & . & R_1(\boldsymbol{\theta_2}) \end{pmatrix} = R_1(\boldsymbol{\theta_2}) \otimes I_s \in M_{nxn}$$

Si $\sigma_1 = 0$ coincide con **VC**, y si $Z = 0$ tenemos $V(\boldsymbol{\theta}) = \sigma^2 I_n$ (modelo GLM). Por ello la opción **VC** es el valor por defecto en REPEATED

Principales estructuras de varianza con SAS III

Heterogeneous Compound Simmetry (CSH)
(Simetría Compuesta Heterogénea)

Similar a la anterior con distintas varianzas y usando el coeficiente de correlación ρ en vez de la covarianza σ_1

El número de parámetros es $r + 1$

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho & . & \sigma_r\sigma_1\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & . & \sigma_r\sigma_2\rho \\ . & . & . & . \\ \sigma_r\sigma_1\rho & \sigma_r\sigma_2\rho & . & \sigma_r^2 \end{pmatrix} \in M_{rxr}$$

Igual que antes, la matriz $R(\theta_2)$ se define como:

$$R(\theta_2) = R_1(\theta_2) \otimes I_s \in M_{nxn}$$

Principales estructuras de varianza con SAS III

Heterogeneous Compound Simmetry (CSH)
(Simetría Compuesta Heterogénea)

Similar a la anterior con distintas varianzas y usando el coeficiente de correlación ρ en vez de la covarianza σ_1

El número de parámetros es $r+1$

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_2\sigma_1\rho & . & \sigma_r\sigma_1\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & . & \sigma_r\sigma_2\rho \\ . & . & . & . \\ \sigma_r\sigma_1\rho & \sigma_r\sigma_2\rho & . & \sigma_r^2 \end{pmatrix} \in M_{rxr}$$

Igual que antes, la matriz $R(\boldsymbol{\theta_2})$ se define como:

$$R(\boldsymbol{\theta_2}) = R_1(\boldsymbol{\theta_2}) \otimes I_s \in M_{nxn}$$

Principales estructuras de varianza con SAS IV

Unstructured (UN) (No estructurada)

Es la estructura más general, con varianzas y covarianzas cualesquiera. Si hay t niveles en el factor el número de parámetros es $\frac{t(t+1)}{2}$. Por ejemplo para $t = 4$ es:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

Unstructured Correlations (UNR)

Igual que la anterior utilizando correlaciones en vez de covarianzas

$$R_1(\theta_2) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_1\sigma_2\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & \sigma_2\sigma_4\rho_{42} \\ \sigma_1\sigma_3\rho_{31} & \sigma_2\sigma_3\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_1\sigma_4\rho_{41} & \sigma_2\sigma_4\rho_{42} & \sigma_3\sigma_4\rho_{43} & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS IV

Unstructured (UN) (No estructurada)

Es la estructura más general, con varianzas y covarianzas cualesquiera. Si hay t niveles en el factor el número de parámetros es $\frac{t(t+1)}{2}$. Por ejemplo para $t = 4$ es:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

Unstructured Correlations (UNR)

Igual que la anterior utilizando correlaciones en vez de covarianzas

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_1\sigma_2\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & \sigma_2\sigma_4\rho_{42} \\ \sigma_1\sigma_3\rho_{31} & \sigma_2\sigma_3\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_1\sigma_4\rho_{41} & \sigma_2\sigma_4\rho_{42} & \sigma_3\sigma_4\rho_{43} & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS V

Banded Main Diagonal UN(1)

Varianzas distintas para cada nivel e independencia entre niveles

Tiene r parámetros. Si $r = 4$ la expresión sería:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}$$

Banded Unstructured UN(2) o Unstructured Correlation UNR(2)

Tienen $2r - 1$ parámetros. Por ejemplo, para **UNR(2)** y $r = 4$ sería:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & 0 & 0 \\ \sigma_1\sigma_2\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & 0 \\ 0 & \sigma_2\sigma_3\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ 0 & 0 & \sigma_3\sigma_4\rho_{43} & \sigma_4^2 \end{pmatrix}$$

En general, **UN(s)** o **UNR(s)** con $s < r$ y $\frac{s}{2}(2r - s + 1)$ parámetros

Principales estructuras de varianza con SAS V

Banded Main Diagonal UN(1)

Varianzas distintas para cada nivel e independencia entre niveles

Tiene r parámetros. Si $r = 4$ la expresión sería:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}$$

Banded Unstructured UN(2) o Unstructured Correlation UNR(2)

Tienen $2r - 1$ parámetros. Por ejemplo, para **UNR(2)** y $r = 4$ sería:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & 0 & 0 \\ \sigma_1\sigma_2\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & 0 \\ 0 & \sigma_2\sigma_3\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ 0 & 0 & \sigma_3\sigma_4\rho_{43} & \sigma_4^2 \end{pmatrix}$$

En general, **UN(s)** o **UNR(s)** con $s < r$ y $\frac{s}{2}(2r - s + 1)$ parámetros

Principales estructuras de varianza con SAS V

Banded Main Diagonal UN(1)

Varianzas distintas para cada nivel e independencia entre niveles

Tiene r parámetros. Si $r = 4$ la expresión sería:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}$$

Banded Unstructured UN(2) o Unstructured Correlation UNR(2)

Tienen $2r - 1$ parámetros. Por ejemplo, para **UNR(2)** y $r = 4$ sería:

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & 0 & 0 \\ \sigma_1\sigma_2\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & 0 \\ 0 & \sigma_2\sigma_3\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ 0 & 0 & \sigma_3\sigma_4\rho_{43} & \sigma_4^2 \end{pmatrix}$$

En general, **UN(s)** o **UNR(s)** con $s < r$ y $\frac{s}{2}(2r - s + 1)$ parámetros

Principales estructuras de varianza con SAS VI

Autoregressive AR(1) (Autoregresiva)

Sólo dos parámetros. Con la misma notación que en simetría compuesta, $R(\boldsymbol{\theta_2}) = R_1(\boldsymbol{\theta_2}) \otimes I_s \in M_{nxn}$ con $\boldsymbol{\theta_2} = (\sigma^2, \rho)'$ y para $r = 4$

$$R_1(\boldsymbol{\theta_2}) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

Heterogeneous Autoregressive ARH(1)

Igual con distintas varianzas y $r+1$ parámetros

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS VI

Autoregressive AR(1) (Autoregresiva)

Sólo dos parámetros. Con la misma notación que en simetría compuesta, $R(\boldsymbol{\theta_2}) = R_1(\boldsymbol{\theta_2}) \otimes I_s \in M_{nxn}$ con $\boldsymbol{\theta_2} = (\sigma^2, \rho)'$ y para $r = 4$

$$R_1(\boldsymbol{\theta_2}) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

Heterogeneous Autoregressive ARH(1)

Igual con distintas varianzas y $r + 1$ parámetros

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS VII

Toeplitz (TOEP)

Con la misma notación y r parámetros

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Heterogeneus Toeplitz (TOEPH)

Con la misma notación y $2r-1$ parámetros

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma_2^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma_3^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS VII

Toeplitz (TOEP)

Con la misma notación y r parámetros

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Heterogeneus Toeplitz (TOEPH)

Con la misma notación y $2r-1$ parámetros

$$R_1(\boldsymbol{\theta_2}) = \begin{pmatrix} \sigma_1^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma_2^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma_3^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS VII

Banded Toeplitz TOEP(s)

Con la misma notación y s parámetros. Para $r = 4$ y $s = 3$

$$R_1(\boldsymbol{\theta_2}) = TOEP(3) = \begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 & 0 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ 0 & \sigma_2 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Heterogeneous Banded Toeplitz TOEPH(s)

Igual con varianzas distintas y $s + r - 1$ parámetros. Para $r = 4$ y $s = 3$

$$R_1(\theta_2) = TOEPH(3) = \begin{pmatrix} \sigma_1^2 & \sigma_1 & \sigma_2 & 0 \\ \sigma_1 & \sigma_2^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma_3^2 & \sigma_1 \\ 0 & \sigma_2 & \sigma_1 & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS VII

Banded Toeplitz TOEP(s)

Con la misma notación y s parámetros. Para $r = 4$ y $s = 3$

$$R_1(\boldsymbol{\theta_2}) = TOEP(3) = \begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 & 0 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ 0 & \sigma_2 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Heterogeneous Banded Toeplitz TOEPH(s)

Igual con varianzas distintas y $s + r - 1$ parámetros. Para $r = 4$ y $s = 3$

$$R_1(\boldsymbol{\theta_2}) = TOEPH(3) = \begin{pmatrix} \sigma_1^2 & \sigma_1 & \sigma_2 & 0 \\ \sigma_1 & \sigma_2^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma_3^2 & \sigma_1 \\ 0 & \sigma_2 & \sigma_1 & \sigma_4^2 \end{pmatrix}$$

Principales estructuras de varianza con SAS VIII

Direct Product

Si hay dos factores de medidas repetidas, el primero con r_1 niveles y el segundo con r_2 niveles, podemos combinar estructuras con el producto directo de matrices. SAS contempla estas tres que ilustraremos para el caso $r_1 = 2$ y $r_2 = 3$:

UN@AR(1). Con $\frac{r_1(r_1+1)}{2} + 1$ parámetros:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \otimes \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} = \left(\begin{array}{ccc|ccc} \sigma_1^2 & \sigma_1^2\rho & \sigma_1^2\rho^2 & \sigma_{21} & \sigma_{21}\rho & \sigma_{21}\rho^2 \\ \sigma_1^2\rho & \sigma_1^2 & \sigma_1^2\rho & \sigma_{21}\rho & \sigma_{21} & \sigma_{21}\rho \\ \sigma_1^2\rho^2 & \sigma_1^2\rho & \sigma_1^2 & \sigma_{21}\rho^2 & \sigma_{21}\rho & \sigma_{21} \\ \hline \sigma_{21} & \sigma_{21}\rho & \sigma_{21}\rho^2 & \sigma_2^2 & \sigma_2^2\rho & \sigma_2^2\rho^2 \\ \sigma_{21}\rho & \sigma_{21} & \sigma_{21}\rho & \sigma_2^2\rho & \sigma_2^2 & \sigma_2^2\rho \\ \sigma_{21}\rho^2 & \sigma_{21}\rho & \sigma_{21} & \sigma_2^2\rho^2 & \sigma_2^2\rho & \sigma_2^2 \end{array}\right)$$

Para $r_1 = 2$ y $r_2 = 3$ tendríamos 4 parámetros

Principales estructuras de varianza con SAS VIII

Direct Product

Si hay dos factores de medidas repetidas, el primero con r_1 niveles y el segundo con r_2 niveles, podemos combinar estructuras con el producto directo de matrices. SAS contempla estas tres que ilustraremos para el caso $r_1 = 2$ y $r_2 = 3$:

UN@AR(1). Con $\frac{r_1(r_1+1)}{2} + 1$ parámetros:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \otimes \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} = \left(\begin{array}{ccc|ccc} \sigma_1^2 & \sigma_1^2\rho & \sigma_1^2\rho^2 & \sigma_{21} & \sigma_{21}\rho & \sigma_{21}\rho^2 \\ \sigma_1^2\rho & \sigma_1^2 & \sigma_1^2\rho & \sigma_{21}\rho & \sigma_{21} & \sigma_{21}\rho \\ \sigma_1^2\rho^2 & \sigma_1^2\rho & \sigma_1^2 & \sigma_{21}\rho^2 & \sigma_{21}\rho & \sigma_{21} \\ \hline \sigma_{21} & \sigma_{21}\rho & \sigma_{21}\rho^2 & \sigma_2^2 & \sigma_2^2\rho & \sigma_2^2\rho^2 \\ \sigma_{21}\rho & \sigma_{21} & \sigma_{21}\rho & \sigma_2^2\rho & \sigma_2^2 & \sigma_2^2\rho \\ \sigma_{21}\rho^2 & \sigma_{21}\rho & \sigma_{21} & \sigma_2^2\rho^2 & \sigma_2^2\rho & \sigma_2^2 \end{array}\right)$$

Para $r_1 = 2$ y $r_2 = 3$ tendríamos 4 parámetros

Principales estructuras de varianza con SAS IX

UN@CS. Con $\frac{r_1(r_1+1)}{2}+1$ parámetros:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1-\sigma^2 & 1-\sigma^2 \\ 1-\sigma^2 & 1 & 1-\sigma^2 \\ 1-\sigma^2 & 1-\sigma^2 & 1 \end{pmatrix} =$$

$$= \left(\begin{array}{ccc|ccc} \sigma_1^2 & \sigma_1^2(1-\sigma^2) & \sigma_1^2(1-\sigma^2) & \sigma_{21} & \sigma_{21}(1-\sigma^2) & \sigma_{21}(1-\sigma^2) \\ \sigma_1^2(1-\sigma^2) & \sigma_1^2 & \sigma_1^2(1-\sigma^2) & \sigma_{21}(1-\sigma^2) & \sigma_{21} & \sigma_{21}(1-\sigma^2) \\ \sigma_1^2(1-\sigma^2) & \sigma_1^2(1-\sigma^2) & \sigma_1^2 & \sigma_{21}(1-\sigma^2) & \sigma_{21}(1-\sigma^2) & \sigma_{21} \\ \hline \sigma_{21} & \sigma_{21}(1-\sigma^2) & \sigma_{21}(1-\sigma^2) & \sigma_2^2 & \sigma_2^2(1-\sigma^2) & \sigma_2^2(1-\sigma^2) \\ \sigma_{21}(1-\sigma^2) & \sigma_{21} & \sigma_{21}(1-\sigma^2) & \sigma_{21}\rho & \sigma_2^2 & \sigma_2^2(1-\sigma^2) \\ \sigma_{21}(1-\sigma^2) & \sigma_{21}(1-\sigma^2) & \sigma_{21} & \sigma_2^2(1-\sigma^2) & \sigma_2^2(1-\sigma^2) & \sigma_2^2 \end{array}\right)$$

con la condición $0 \leq \sigma^2 \leq 1$

Para $r_1 = 2$ y $r_2 = 3$ tendríamos 4 parámetros

Principales estructuras de varianza con SAS X

UN@UN. Con $\frac{r_1(r_1+1)}{2} + \frac{r_2(r_2+1)}{2}$ parámetros:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \otimes \begin{pmatrix} \omega_1^2 & \omega_{21} & \omega_{31} \\ \omega_{21} & \omega_2^2 & \omega_{32} \\ \omega_{31} & \omega_{32} & \omega_3^2 \end{pmatrix} =$$

$$= \left(\begin{array}{ccc|ccc} \sigma_1^2\omega_1^2 & \sigma_1^2\omega_{21} & \sigma_1^2\omega_{31} & \sigma_{21}\omega_1^2 & \sigma_{21}\omega_{21} & \sigma_{21}\omega_{31} \\ \sigma_1^2\omega_{21} & \sigma_1^2\omega_2^2 & \sigma_1^2\omega_{32} & \sigma_{21}\omega_{21} & \sigma_{21}\omega_2^2 & \sigma_{21}\omega_{32} \\ \sigma_1^2\omega_{31} & \sigma_1^2\omega_{32} & \sigma_1^2\omega_3^2 & \sigma_{21}\omega_{31} & \sigma_{21}\omega_{32} & \sigma_{21}\omega_3^2 \\ \hline \sigma_{21}\omega_1^2 & \sigma_{21}\omega_{21} & \sigma_{21}\omega_{31} & \sigma_2^2\omega_1^2 & \sigma_2^2\omega_{21} & \sigma_2^2\omega_{31} \\ \sigma_{21}\omega_{21} & \sigma_{21}\omega_2^2 & \sigma_{21}\omega_{32} & \sigma_{21}\omega_{21} & \sigma_2^2\omega_2^2 & \sigma_2^2\omega_{32} \\ \sigma_{21}\omega_{31} & \sigma_{21}\omega_{32} & \sigma_{21}\omega_3^2 & \sigma_2^2\omega_{31} & \sigma_2^2\omega_{32} & \sigma_2^2\omega_3^2 \end{array}\right)$$

Para $r_1 = 2$ y $r_2 = 3$ tendríamos 9 parámetros

¿Cuándo debemos usar LMM y no GLM?

- Regresiones o Análisis de Varianza donde hayamos detectado heterogeneidad de varianzas o falta de independencia de residuales
- Regresiones con datos correlados donde existe correlación temporal o espacial en los datos observados
- Diseños con estructura de split-plot o split-split-plot porque tienen más de una varianza que estimar
- Regresiones o Análisis de Varianza con medidas repetidas (varianza inter-sujetos, varianza intra-sujetos y posible falta de independencia entre observaciones de un mismo sujeto)
- En general, podemos usar siempre LMM. Si hacemos $Z = 0$ y $R = \sigma^2 I_n$ tenemos GLM. Con SAS bastaría usar PROC MIXED sin los comandos RANDOM y REPEATED

¿Cuándo debemos usar LMM y no GLM?

- Regresiones o Análisis de Varianza donde hayamos detectado heterogeneidad de varianzas o falta de independencia de residuales
- Regresiones con datos correlados donde existe correlación temporal o espacial en los datos observados
- Diseños con estructura de split-plot o split-split-plot porque tienen más de una varianza que estimar
- Regresiones o Análisis de Varianza con medidas repetidas (varianza inter-sujetos, varianza intra-sujetos y posible falta de independencia entre observaciones de un mismo sujeto)
- En general, podemos usar siempre LMM. Si hacemos $Z = 0$ y $R = \sigma^2 I_n$ tenemos GLM. Con SAS bastaría usar PROC MIXED sin los comandos RANDOM y REPEATED

¿Cuándo debemos usar LMM y no GLM?

- Regresiones o Análisis de Varianza donde hayamos detectado heterogeneidad de varianzas o falta de independencia de residuales
- Regresiones con datos correlados donde existe correlación temporal o espacial en los datos observados
- Diseños con estructura de split-plot o split-split-plot porque tienen más de una varianza que estimar
- Regresiones o Análisis de Varianza con medidas repetidas (varianza inter-sujetos, varianza intra-sujetos y posible falta de independencia entre observaciones de un mismo sujeto)
- En general, podemos usar siempre LMM. Si hacemos $Z = 0$ y $R = \sigma^2 I_n$ tenemos GLM. Con SAS bastaría usar PROC MIXED sin los comandos RANDOM y REPEATED

¿Cuándo debemos usar LMM y no GLM?

- Regresiones o Análisis de Varianza donde hayamos detectado heterogeneidad de varianzas o falta de independencia de residuales
- Regresiones con datos correlados donde existe correlación temporal o espacial en los datos observados
- Diseños con estructura de split-plot o split-split-plot porque tienen más de una varianza que estimar
- Regresiones o Análisis de Varianza con medidas repetidas (varianza inter-sujetos, varianza intra-sujetos y posible falta de independencia entre observaciones de un mismo sujeto)
- En general, podemos usar siempre LMM. Si hacemos $Z = 0$ y $R = \sigma^2 I_n$ tenemos GLM. Con SAS bastaría usar PROC MIXED sin los comandos RANDOM y REPEATED

¿Cuándo debemos usar LMM y no GLM?

- Regresiones o Análisis de Varianza donde hayamos detectado heterogeneidad de varianzas o falta de independencia de residuales
- Regresiones con datos correlados donde existe correlación temporal o espacial en los datos observados
- Diseños con estructura de split-plot o split-split-plot porque tienen más de una varianza que estimar
- Regresiones o Análisis de Varianza con medidas repetidas (varianza inter-sujetos, varianza intra-sujetos y posible falta de independencia entre observaciones de un mismo sujeto)
- En general, podemos usar siempre LMM. Si hacemos $Z = 0$ y $R = \sigma^2 I_n$ tenemos GLM. Con SAS bastaría usar PROC MIXED sin los comandos RANDOM y REPEATED

Bibliografía

- Jiang J. (2007). *Linear and Generalized Linear Mixed Models and Their Applications*. Springer Series in Statistics. Springer. New York, USA
- McCulloch C.E., Searle S.R. (2001). *Generalized, Linear and Mixed Models*. Wiley-Interscience. New York
- Muller K.E., Fetterman B.A. (2002). *Regression and ANOVA. An Integrated Approach Using SAS® Software*. SAS Institute Inc. and John Wiley and Sons Inc. North Carolina, USA
- SAS Institute, The MIXED procedure 2005. *SAS/STAT User's Guide*. SAS On-Line Documentation, Cary, NC
- Searle S.R., Casella G., McCulloch C.E. (1992). *Variance Components*. John Wiley & Sons. New York
- Verbeke G., Molenberghs G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, Berlin.