

Modelos de regresión beta mixtos

Olga Cecilia Usuga Manco

Departamento de Ingeniería Industrial
Facultad de Ingeniería
Universidad de Antioquia

Septiembre 23 de 2013

Modelos de regresión beta mixtos

Olga Cecilia Usuga Manco

Departamento de Ingeniería Industrial
Facultad de Ingeniería
Universidad de Antioquia

Septiembre 23 de 2013

Resumen

1 Introducción

2 Modelos de regresión beta con efectos aleatorios

Modelo de regresión beta con intercepto aleatorio normal

Modelos de regresión beta con intercepto aleatorio no normal

Modelos de regresión beta con intercepto y pendiente aleatoria

3 Selección del modelo y análisis de residuos

4 Aplicación

5 Paquete BLMM

6 Conclusiones

1 Introducción

2 Modelos de regresión beta con efectos aleatorios

3 Selección del modelo y análisis de residuos

4 Aplicación

5 Paquete BLMM

6 Conclusiones

Motivación

Análisis de datos longitudinales.

- Naturaleza de la variable respuesta.
- Naturaleza de las observaciones repetidas.
- Flexibilidad en la suposición de normalidad de la distribución de los efectos aleatorios.

Estudio oftalmológico

- Individuos: 29 pacientes.
- Variable respuesta: Porcentaje de gas presente en los ojos del paciente i en el tiempo j , con $i = 1, \dots, 29$ y $j = 1, \dots, n_i$.
- Tiempo: Dias después de la cirugía (3, 4, \dots , 15).
- Variable regresora: Concentración de gas intraocular (15 %, 20 %, 25 %).

Gráfico de perfil

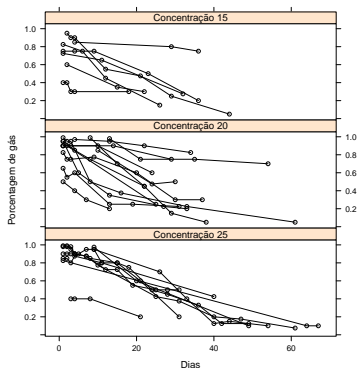


Figura 1: Gráfico de perfil del porcentaje de gas presente en los ojos de los pacientes según los niveles de concentración de gas.

Modelo de regresión beta

- Paolino (2001)
- Kieschnick y McCullough (2003)
- Ferrari y Cribari-Neto (2004)
- Rigby y Stasinopoulos (2005)
- Simas, Barreto-Souza y Rocha (2010)
- Ospina y Ferrari (2012)

Distribución beta

Parametrización de la distribución beta propuesta por Ferrari y Cribari-Neto (2004).

La función de densidad de probabilidad de la distribución beta denotada por $Be(\mu, \phi)$ es definida como

$$f(y | \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1 \quad (1)$$

con $0 < \mu < 1$ y $\phi > 0$.

La media y la varianza de Y están dadas por

$$\begin{aligned} E(Y) &= \mu, \\ \text{Var}(Y) &= \frac{\mu(1-\mu)}{1+\phi}. \end{aligned}$$

Distribución beta

La función de densidad de probabilidad de la distribución beta, denotada por $Be(\mu, \sigma)$, se define como

$$f(y | \mu, \sigma) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}, \quad 0 < y < 1 \quad (2)$$

donde

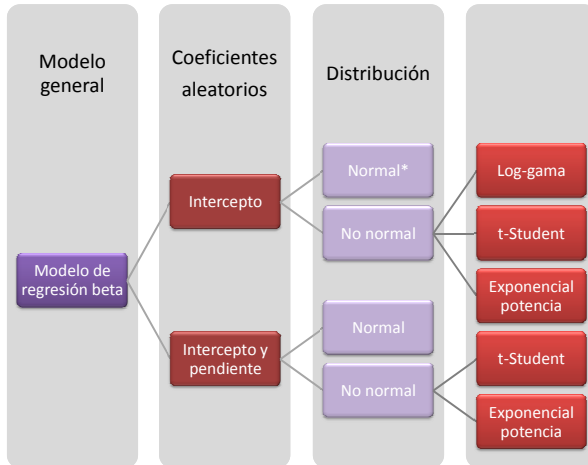
$$\alpha = \frac{\mu(1 - \sigma^2)}{\sigma^2},$$
$$\beta = \frac{(1 - \mu)(1 - \sigma^2)}{\sigma^2},$$

con $0 < \mu < 1$ y $0 < \sigma < 1$.

La media y la varianza de Y están dadas por

$$\begin{aligned} E(Y) &= \mu, \\ \text{Var}(Y) &= \sigma^2 \mu(1 - \mu). \end{aligned}$$

- 1 Introducción
- 2 Modelos de regresión beta con efectos aleatorios
- 3 Selección del modelo y análisis de residuos
- 4 Aplicación
- 5 Paquete BLMM
- 6 Conclusiones



1 Introducción

2 Modelos de regresión beta con efectos aleatorios

Modelo de regresión beta con intercepto aleatorio normal

Modelos de regresión beta con intercepto aleatorio no normal

Modelos de regresión beta con intercepto y pendiente aleatoria

3 Selección del modelo y análisis de residuos

4 Aplicación

5 Paquete BLMM

6 Conclusiones

Modelo de regresión beta con intercepto aleatorio normal

En los modelos de regresión beta con interceptos aleatorios normales se asume que

- La distribución condicional de y_{ij} dados los efectos aleatorios sigue una distribución beta con función de densidad de probabilidad dada en (2),
- dados los efectos aleatorios, los $y_{i1}, y_{i2}, \dots, y_{in_i}$ son independientes, y
- los efectos aleatorios γ_{i1} e γ_{i2} son independientes e idénticamente distribuidos con distribuciones $N(0, \lambda_1)$ y $N(0, \lambda_2)$, respectivamente.

Modelo de regresión beta con intercepto aleatorio normal

La formulación del modelo es

$$\begin{aligned}y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\stackrel{\text{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ \gamma_{i1} &\stackrel{\text{i.i.d}}{\sim} N(0, \lambda_1), \\ \gamma_{i2} &\stackrel{\text{i.i.d}}{\sim} N(0, \lambda_2),\end{aligned}$$

$$\begin{aligned}g_1(\mu_{ij}) = \eta_{ij1} &= \mathbf{x}_{ij1}^T \boldsymbol{\beta}_1 + \gamma_{i1}, \\ g_2(\sigma_{ij}) = \eta_{ij2} &= \mathbf{x}_{ij2}^T \boldsymbol{\beta}_2 + \gamma_{i2}.\end{aligned}$$

Máxima verosimilitud aproximada

Asumiendo que los N individuos son independientes la función de verosimilitud toma la forma

$$L(\boldsymbol{\theta}) = \prod_{i=1}^N \int \int_{\mathbb{R}^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \boldsymbol{\beta}_1, \boldsymbol{\beta}_2) \cdot f(\gamma_{i1}; \lambda_1) f(\gamma_{i2}; \lambda_2) d\gamma_{i1} d\gamma_{i2},$$

Las integrales de $L(\boldsymbol{\theta})$ son evaluadas por medio de cuadratura de Gauss-Hermite.

Logaritmo de la función de verosimilitud aproximada

$$\ell(\boldsymbol{\theta}) \cong \sum_{i=1}^N \log \left(\sum_{k_1=1}^{Q_1} \sum_{k_2=1}^{Q_2} \prod_{j=1}^{n_i} f(y_{ij} \mid \sqrt{2}\lambda_1 z_{k_1}, \sqrt{2}\lambda_2 z_{k_2}; \boldsymbol{\beta}_1, \boldsymbol{\beta}_2) \frac{w_{k_1} w_{k_2}}{\pi} \right).$$

Predicción de los efectos aleatorios

Mejor predictor empírico de Bayes

$$\tilde{\gamma}_{i1} = \frac{\int \int_{\mathbb{R}^2} \gamma_{i1} \prod_{j=1}^{ni} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\beta}_1, \hat{\beta}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) d\gamma_{i1} d\gamma_{i2}}{\int \int_{\mathbb{R}^2} \prod_{j=1}^{ni} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\beta}_1, \hat{\beta}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) d\gamma_{i1} d\gamma_{i2}},$$

$$\tilde{\gamma}_{i2} = \frac{\int \int_{\mathbb{R}^2} \gamma_{i2} \prod_{j=1}^{ni} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\beta}_1, \hat{\beta}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) d\gamma_{i1} d\gamma_{i2}}{\int \int_{\mathbb{R}^2} \prod_{j=1}^{ni} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\beta}_1, \hat{\beta}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) d\gamma_{i1} d\gamma_{i2}},$$

1 Introducción

2 Modelos de regresión beta con efectos aleatorios

Modelo de regresión beta con intercepto aleatorio normal

Modelos de regresión beta con intercepto aleatorio no normal

Modelos de regresión beta con intercepto y pendiente aleatoria

3 Selección del modelo y análisis de residuos

4 Aplicación

5 Paquete BLMM

6 Conclusiones

Modelo de regresión beta con intercepto aleatorio no normal

En los modelos de regresión beta con interceptos aleatorios no normales se asume que

- La distribución condicional de y_{ij} dados los efectos aleatorios sigue una distribución beta con función de densidad de probabilidad dada en (2),
- dados los efectos aleatorios, los $y_{i1}, y_{i2}, \dots, y_{in_i}$ son independientes, y
- los efectos aleatorios γ_{i1} y γ_{i2} son independientes e idénticamente distribuidos con funciones densidades de probabilidad $f(\gamma_{i1}; \lambda_1)$ y $f(\gamma_{i2}; \lambda_2)$, respectivamente.

Modelo de regresión beta con intercepto aleatorio no normal

La formulación del modelo es

$$\begin{aligned}y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\stackrel{\text{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ \gamma_{i1} &\stackrel{\text{i.i.d}}{\sim} f(\gamma_{i1}; \boldsymbol{\lambda}_1), \\ \gamma_{i2} &\stackrel{\text{i.i.d}}{\sim} f(\gamma_{i2}; \boldsymbol{\lambda}_2),\end{aligned}\tag{3}$$

$$\begin{aligned}g_1(\mu_{ij}) = \eta_{ij1} &= \mathbf{x}_{ij1}^T \boldsymbol{\beta}_1 + \gamma_{i1}, \\ g_2(\sigma_{ij}) = \eta_{ij2} &= \mathbf{x}_{ij2}^T \boldsymbol{\beta}_2 + \gamma_{i2},\end{aligned}\tag{4}$$

Distribuciones de los interceptos aleatorios

- Distribución log-gama
- Distribución t-Student
- Distribución exponencial potencia

Maxima verosimilitud aproximada

El método de estimación usado es máxima verosimilitud aproximada y esta basado en el trabajo de Liu y Yu (2008).

Función de verosimilitud

$$L(\theta) = \prod_{i=1}^N \int_{\mathbb{R}^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot f(\gamma_{i1}; \lambda_1) f(\gamma_{i2}; \lambda_2) d\gamma_{i1} d\gamma_{i2}$$

Función de verosimilitud reformulada

$$L(\theta) = \prod_{i=1}^N \int_{\mathbb{R}^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot \frac{f(\gamma_{i1}; \lambda_1)}{\phi(\gamma_{i1})} \frac{f(\gamma_{i2}; \lambda_2)}{\phi(\gamma_{i2})} \phi(\gamma_{i1}) \phi(\gamma_{i2}) d\gamma_{i1} d\gamma_{i2}$$

Función de log-verosimilitud aproximada

$$\ell(\theta) \cong \sum_{i=1}^N \log \left(\sum_{k_1=1}^{Q_1} \sum_{k_2=1}^{Q_2} \prod_{j=1}^{n_i} f(y_{ij} \mid z_{k_1}, z_{k_2}; \beta_1, \beta_2) \cdot \frac{f(z_{k_1}; \lambda_1)}{\phi(z_{k_1})} \frac{f(z_{k_2}; \lambda_2)}{\phi(z_{k_2})} w_{k_1} w_{k_2} \right)$$

1 Introducción

2 Modelos de regresión beta con efectos aleatorios

Modelo de regresión beta con intercepto aleatorio normal

Modelos de regresión beta con intercepto aleatorio no normal

Modelos de regresión beta con intercepto y pendiente aleatoria

3 Selección del modelo y análisis de residuos

4 Aplicación

5 Paquete BLMM

6 Conclusiones

Modelos de regresión beta con intercepto y pendiente aleatoria

En los modelos de regresión beta con intercepto y pendiente aleatoria se asume que

- La distribución condicional de y_{ij} dados los efectos aleatorios sigue una distribución beta con función de densidad de probabilidad dada en (2),
- dados los efectos aleatorios, los $y_{i1}, y_{i2}, \dots, y_{in_i}$ son independientes, y
- los efectos aleatorios γ_{i1} y γ_{i2} son independientes e idénticamente distribuidos con distribuciones elípticas bivariadas $El_2(\mathbf{0}, \Sigma_1)$ y $El_2(\mathbf{0}, \Sigma_2)$, respectivamente.

Modelos de regresión beta con intercepto y pendiente aleatoria

La formulación del modelo es

$$\begin{aligned}y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\stackrel{\text{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ \gamma_{i1} &\stackrel{\text{i.i.d}}{\sim} El_2(\mathbf{0}, \Sigma_1), \\ \gamma_{i2} &\stackrel{\text{i.i.d}}{\sim} El_2(\mathbf{0}, \Sigma_2),\end{aligned}\tag{5}$$

$$\begin{aligned}g_1(\mu_{ij}) = \eta_{ij1} &= \mathbf{x}_{ij1}^T \boldsymbol{\beta}_1 + \mathbf{z}_{ij1}^T \gamma_{i1}, \\ g_2(\sigma_{ij}) = \eta_{ij2} &= \mathbf{x}_{ij2}^T \boldsymbol{\beta}_2 + \mathbf{z}_{ij2}^T \gamma_{i2},\end{aligned}\tag{6}$$

Estructuras de varianza-covarianza

Cuadro 2: Estructuras de varianza-covarianza

Estructura	Ejemplo	Número de parámetros
S	$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$	1
CV	$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$	2
AR(1)	$\begin{pmatrix} \lambda & \rho\lambda \\ \rho\lambda & \lambda \end{pmatrix}$	2
SCH	$\begin{pmatrix} \lambda_1 & \rho\sqrt{\lambda_1}\sqrt{\lambda_2} \\ \rho\sqrt{\lambda_1}\sqrt{\lambda_2} & \lambda_2 \end{pmatrix}$	3

Distribuciones de los efectos aleatorios

- Distribución normal
- Distribución t-Student
- Distribución exponencial potencia

Máxima verosimilitud aproximada

Función de verosimilitud

$$L(\theta) = \prod_{i=1}^N \int_{\mathbb{R}^4} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot f(\gamma_{i1}; \Sigma_1) f(\gamma_{i2}; \Sigma_2) d\gamma_{i1} d\gamma_{i2}.$$

Función de verosimilitud reformulada

$$L(\theta) = \prod_{i=1}^N \int_{\mathbb{R}^4} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot \frac{f(\gamma_{i1}; \Sigma_1)}{\phi(\gamma_{i1})} \frac{f(\gamma_{i2}; \Sigma_2)}{\phi(\gamma_{i2})} \phi(\gamma_{i1}) \phi(\gamma_{i2}) d\gamma_{i1} d\gamma_{i2},$$

Logaritmo de la función de verosimilitud

$$\ell(\theta) = \sum_{i=1}^N \log \left(\int_{\mathbb{R}^4} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot \frac{f(\gamma_{i1}; \Sigma_1)}{\phi(\gamma_{i1})} \frac{f(\gamma_{i2}; \Sigma_2)}{\phi(\gamma_{i2})} \phi(\gamma_{i1}) \phi(\gamma_{i2}) d\gamma_{i1} d\gamma_{i2} \right).$$

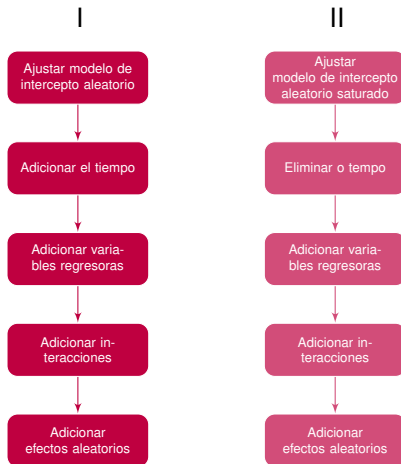
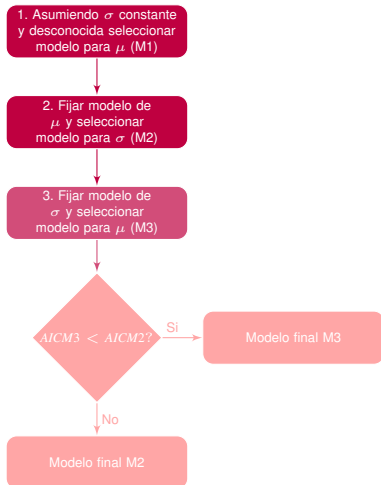
- 1 Introducción
- 2 Modelos de regresión beta con efectos aleatorios
- 3 Selección del modelo y análisis de residuos**
- 4 Aplicación
- 5 Paquete BLMM
- 6 Conclusiones

Selección del modelo

El método de selección propuesto de los modelos de regresión beta con efectos aleatorios para datos longitudinales esta basado en los métodos de selección de los modelos

- Aditivos generalizados para posición, escala y forma, Stasinopoulos, et.al (2012)
- Modelos lineales mixtos para datos longitudinales, Ryyo(2010)

Selección del modelo



Análisis de residuos

- Residuo cuantil aleatorizado
- Residuo condicional
- Residuo marginal
- Residuo aleatorio

Análisis de residuos

- Residuo cuantil aleatorizado propuesto por Dunn y Smyth (1996)

$$r_{qij} = \Phi^{-1} (F(y_{ij}; \hat{\mu}_{ij}, \hat{\sigma}_{ij})) ,$$

donde $\hat{\mu}_{ij} = g_1^{-1}(\mathbf{x}_{ij1}^T \hat{\beta}_1 + \hat{\gamma}_{i1})$ y $\hat{\sigma}_{ij} = g_2^{-1}(\mathbf{x}_{ij2}^T \hat{\beta}_2 + \hat{\gamma}_{i2})$.

- Residuo condicional

$$r_{cij} = \frac{y_{ij} - \hat{E}(y_{ij} \mid \gamma_{i1}, \gamma_{i2})}{\sqrt{\widehat{\text{Var}}(y_{ij} \mid \gamma_{i1}, \gamma_{i2})}},$$

donde $\hat{E}(y_{ij} \mid \gamma_{i1}, \gamma_{i2}) = \hat{\mu}_{ij}$ y $\widehat{\text{Var}}(y_{ij} \mid \gamma_{i1}, \gamma_{i2}) = \hat{\sigma}_{ij}^2 \hat{\mu}_{ij}(1 - \hat{\mu}_{ij})$.

Análisis de residuos

- Residuo marginal

$$r_{mij} = \frac{y_{ij} - \hat{E}(y_{ij})}{\sqrt{\widehat{\text{Var}}(y_{ij})}},$$

donde

$$\begin{aligned}\widehat{\mathbf{E}}(\mathbf{y}_{ij}) &= \mathbf{E}(\hat{\mu}_{ij}), \\ \widehat{\mathbf{Var}}(\mathbf{y}_{ij}) &= \mathbf{E}(\hat{\mu}_{ij}^2) - (\mathbf{E}(\hat{\mu}_{ij}))^2 + \mathbf{E}(\hat{\sigma}_{ij}^2) \cdot \mathbf{E}(\hat{\mu}_{ij}(1 - \hat{\mu}_{ij})).\end{aligned}$$

$$\begin{aligned} E(y_{ij}) &= 1 - E\left(\frac{1}{1 + ae^{\gamma_{i1}}}\right), \\ \text{Var}(y_{ij}) &= E\left(\frac{1}{(1 + ae^{\gamma_{i1}})^2}\right) - \left[E\left(\frac{1}{1 + ae^{\gamma_{i1}}}\right)\right]^2 + ab^2 E\left(\frac{e^{2\gamma_{i2}}}{(1 + be^{\gamma_{i2}})^2}\right) E\left(\frac{e^{\gamma_{i1}}}{(1 + ae^{\gamma_{i1}})^2}\right). \end{aligned}$$

Análisis de residuos

- Residuo aleatorio

$$r_{a_{\mu}} = \frac{\hat{\gamma}_{i1}}{\sqrt{\widehat{\text{Var}}(\gamma_{i1})}},$$

$$r_{a_{\sigma}} = \frac{\hat{\gamma}_{i2}}{\sqrt{\widehat{\text{Var}}(\gamma_{i2})}}.$$

Estudios de simulación

- 1 Evaluar el desempeño del proceso de estimación.
- 2 Estudiar el comportamiento de las distribuciones empíricas de los residuos propuestos.

Estudios de simulación

Los datos fueron generados de acuerdo con el modelo

$$\begin{aligned}y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\stackrel{\text{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ \gamma_{i1} &\stackrel{\text{i.i.d}}{\sim} f(\gamma_{i1}; \boldsymbol{\lambda}_1), \\ \gamma_{i2} &\stackrel{\text{i.i.d}}{\sim} f(\gamma_{i2}; \boldsymbol{\lambda}_2),\end{aligned}$$

$$\begin{aligned}\text{logit}(\mu_{ij}) &= \beta_{11} + \beta_{21}x_{ij} + \beta_{31}t_{ij} + \gamma_{i1}, \\ \text{logit}(\sigma_{ij}) &= \beta_{12} + \beta_{22}x_{ij} + \beta_{32}t_{ij} + \gamma_{i2},\end{aligned}$$

Estudios de simulación

- $\beta_1 = \beta_2 = (-0, 15; 0, 15; -0, 15)^T$
- $N = 20, 40, 60, 100, 150$ y $n_i = 3, 5, 8, 12$.
- $\lambda_1 = \lambda_2 = 0, 5; 1, 0; 1, 5$.
- $\lambda_1 = \lambda_2 = 0, 22; 0, 70; 1, 22$.
- $\nu = 3, 4$.
- $\lambda = 0, 44; 0, 89; 1, 33$ y $\nu = 1, 5$.

Estudios de simulación

El desempeño del proceso de estimación fue evaluado por medio de la raíz del error cuadrático medio (REQM) propuesta por Wissel (2009) y definida como

$$REQM = (\text{traza}(\Sigma(\hat{\theta})) + (\hat{\theta} - \theta)^T(\hat{\theta} - \theta))^{1/2}.$$

Valores de la REQM próximos de cero indican buen desempeño en el proceso de estimación.

Estudios de simulación

Los datos fueron generados de acuerdo con el modelo

$$\begin{aligned} y_{ij} | \gamma_{i1} &\overset{\text{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ \gamma_{i1} &\overset{\text{i.i.d}}{\sim} N_2(\mathbf{0}, \Sigma_1), \end{aligned} \quad (7)$$

con

$$\text{logit}(\mu_{ij}) = (\beta_{11} + \gamma_{i11}) + (\beta_{21} + \gamma_{i21})t_{ij1}, \quad (8)$$

- $\beta_1 = (-0, 3; 0, 6)^T$, $\lambda_1 = (0, 5; 0, 3)^T$ y $\rho_1 = -0, 5$.
- $\beta_{12} = -2, 0$.
- Número de individuos $N = 15, 20, 30$.
- Número de observaciones por individuo $n_i = 4, 6, 8, 10$.

Estudo de simulação

El proceso de estimación mejora su desempeño en la medida que

- La cantidad de información por individuo aumenta.
- La cantidad de individuos en el estudio longitudinal aumenta.
- La variabilidad de la distribución de los interceptos aleatorios asociados a μ y σ disminuye.
- Las estructuras de varianza-covarianza son simples.

Distribución empírica de los residuos

- La distribución empírica del residuo cuantil aleatorizado sigue una distribución aproximadamente normal.
- Las distribuciones empíricas de los residuos condicional, marginal y aleatorios presentan asimetría.
- Se recomienda el residuo cuantil aleatorizado para realizar análisis de residuos de los modelos propuestos.

- 1 Introducción
- 2 Modelos de regresión beta con efectos aleatorios
- 3 Selección del modelo y análisis de residuos
- 4 Aplicación**
- 5 Paquete BLMM
- 6 Conclusiones

Estudio oftalmológico

- Individuos: 29 pacientes.
- Variable respuesta: y_{ij} : Porcentaje de gas presente en los ojos del paciente i en el tiempo j , con $i = 1, \dots, 29$ y $j = 1, \dots, n_i$.
- Tiempo: Dias después de la cirugía (3, 4, \dots , 15).
- Variable regresora: Concentración estándar de gas

$$x_{ij} = \frac{C_{ij} - 20}{5} = \begin{cases} -1 & \text{si la concentración de gas es 15 \%} \\ 0 & \text{si la concentración de gas es 20 \%} \\ 1 & \text{si la concentración de gas es 25 \%.} \end{cases}$$

Estudio oftalmológico

El modelo ajustado a los datos fue

$$\begin{aligned}y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\overset{\text{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ \gamma_{i1} &\overset{\text{i.i.d}}{\sim} N(0, \lambda_1^2), \\ \gamma_{i2} &\overset{\text{i.i.d}}{\sim} N(0, \lambda_2^2),\end{aligned}$$

con

$$\begin{aligned}\text{logit}(\mu_{ij}) &= \beta_{11} + \beta_{31} \log^2(t_{ij}) + \beta_{41} x_{ij} + \gamma_{i1}, \\ \text{logit}(\sigma_{ij}) &= \beta_{12} + \gamma_{i2}.\end{aligned}$$

Estudio oftalmológico

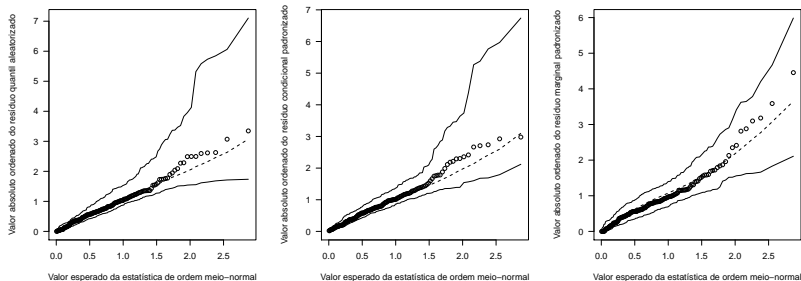


Figura 2: Gráficos de probabilidad medio-normal del residuo cuantil aleatorizado, condicional y marginal del modelo beta normal ajustado a los datos del porcentaje de gas presente en los ojos de los pacientes.

- 1 Introducción
- 2 Modelos de regresión beta con efectos aleatorios
- 3 Selección del modelo y análisis de residuos
- 4 Aplicación
- 5 Paquete BLMM**
- 6 Conclusiones

Paquete BLMM

El objetivo principal del paquete BLMM es ajustar modelos de regresión beta con efectos aleatorios.

- `blmmML`
- `print.blmmML`
- `summary.blmmML`
- GHQ (pruning)
- `re.prediction`
- `residuals.blmmML`
- `halfnorm`

Función blmmML

```
blmmML(formula, sigma.formula, data,  
  param = "BE_DISP",  
  link.mu = "logit", link.sigma = "logit",  
  re.dist.mu = "NORMAL",  
  re.dist.sigma = "NORMAL",  
  varcov = "CSH", method = "GHQ",  
  n.points = 10,  
  optimizer = "nlminb",  
  optim.method = "Nelder-Mead",  
  control = list(), A = NULL,  
  transf.par.disp = FALSE)
```

- 1 Introducción
- 2 Modelos de regresión beta con efectos aleatorios
- 3 Selección del modelo y análisis de residuos
- 4 Aplicación
- 5 Paquete BLMM
- 6 Conclusiones**

Conclusões

- Los modelos con intercepto aleatorio son versátiles debido a su flexibilidad para capturar comportamientos normales, no normales y asimétricos que pueden ocurrir al nivel de los efectos aleatorios.
- La inclusión de intercepto y pendiente aleatoria en la estructura de regresión del modelo enriquece la clase de modelos de regresión beta con efectos aleatorios dado que la correlación entre observaciones del mismo individuo es acomodada de forma sistemática por medio de las estructuras de varianza-covarianza.
- La metodología presentada es una herramienta completa y flexible para el análisis de datos longitudinales y correlacionados restringidos a un intervalo con heterogeneidad entre individuos en los coeficientes de regresión como el intercepto y la pendiente aleatoria.

Principais referencias bibliográficas

- Dunn, P. K. y Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5(3), 236-244.
- Ferrari, S.L.P. y Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Statistical Computation and Simulation*, 31(7), 799-815.
- Kieschnick, R. y McCullough, B.D. (2003). Regression analysis of variates observed on (0,1): percentages, proportions, and fractions. *Statistical Modelling*, 3(3), 193-213.
- Liu, L. y Yu, Z. (2008). A likelihood reformulation method in non-normal random effects models. *Statistics in medicine*, 27(16), 3105-3124.

Principais referencias bibliográficas

- Meyers, S.M., Ambler, J.S., Tan, M., Werner, J.C. y Huang, S.S. (1992). Variation of perfluoropropane disappearance after vitrectomy. *Retina*, 12, 359-363.
- Ospina, R. y Ferrari, S.L.P. (2012). A general class of zero-or-one inflated beta regression models. *Computational Statistics and Data Analysis*, 56(6), 1609-1623.
- Paolino, P. (2001). Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables. *Political Analysis*, 9(4), 325-346.
- Rigby, R.A. e Stasinopoulos, D.M. (2005). Generalized additive models for location, scale and shape. *Applied Statistical*, 54(3), 507-554.

Principais referencias bibliográficas

- Ryoo, J.H. (2010). Model Selection with the Linear Mixed Effects Model for Longitudinal Data. Tese de Doutorado, University of Minnesota, USA.
- Simas, A.B., Barreto-Souza, W. y Rocha, A.V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics and Data Analysis*, 54(2), 348-366.
- Stasinopoulos, M., Rigby, B. y Akantziliotou, C. (2012). Instructions on how to use the gamlss package in R, second edição, 2012. URL <http://www.R-project.org/>.
- Wissel, J. (2009). *A new biased estimator for multivariate regression models with highly collinear variables*. Tese de Doutorado, Institut für Mathematik, Universität Würzburg, Germany.

Modelos de regresión beta mixtos

Olga Cecilia Usuga Manco

Departamento de Ingeniería Industrial
Facultad de Ingeniería
Universidad de Antioquia

Septiembre 23 de 2013