MAT-468: Modelo no lineal con efectos mixtos

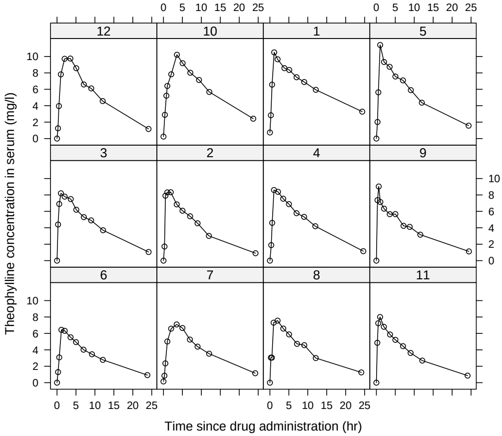
Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Cinética de Teofilina (Boeckmann, Sheiner y Beal, 1984)

Dosis orales de droga anti-asmática Teofilina son suministradas a doce individuos, luego las concentraciones en la sangre (mg/L) son medidas en 11 instantes en un periodo de 25 horas.



Cinética de Teofilina (Boeckmann, Sheiner y Beal, 1984)

Observaciones:

- Estudio de laboratorio (farmacocinética) con pocos individuos y varias observaciones balanceadas.
- **Modelo mecánico:** modelo de un compartimento (primer orden) para expresar la concentración de Teofilina $C(t)$ en un instante t después de una dosis inicial Δ

$$C(t) = \frac{\Delta k_a}{V(k_a - Cl/V)}\Big\{\exp\Big(-\frac{Cl}{V}t\Big) - \exp(-k_a t)\Big\},$$

donde k_a es la tasa de absorción, V representa el volumen de distribución y Cl es el Clearance

(es usual reparametrizar en términos de la tasa de eliminación $K = Cl/V$).

Objetivo:

- Determinar, basado en los perfiles observados, valores promedios de (k_a, V, Cl) para realizar inferencias en los individuos de la población.
- Diseñar regímenes de dosificación para mantener las concentraciones de la droga en los niveles deseados.

Etapa 1: Modelo sujeto-específico

Considere:

$$Y_{ij} = f(\boldsymbol{x}_{ij}, \boldsymbol{\phi}_i) + \epsilon_{ij}, \qquad i = 1, \dots, n; j = 1, \dots, m_i, \tag{1}$$

donde Y_{ij} es la j-ésima observacion del i-ésimo individuo, f es una función que gobierna el comportamiento individual, $\boldsymbol{\phi}_i$ es el vector de parámetros sujeto específico y $\boldsymbol{x}_{ij}$ representa un vector de covariables.

En muchas aplicaciones $\boldsymbol{x}_{ij} = (t_{ij}, \boldsymbol{u}_i^T)^T$ donde t_{ij} es el "tiempo" y $\boldsymbol{u}_i$ representa condiciones adicionales. (en nuestro *ejemplo de juguete*, $\boldsymbol{u}_i = \Delta_i$).

Etapa 2: Modelo poblacional

Sea

$$\boldsymbol{\phi}_i = \boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i), \qquad i = 1, \dots, n, \tag{2}$$

donde $\boldsymbol{d}$ es una función q-dimensional que depende del vector de efectos fijos $\boldsymbol{\beta}$, el vector de efectos aleatorios $\boldsymbol{b}_i$ y covariables $\boldsymbol{a}_i$ (asociadas a los atributos del individuo).

Caso especial:

$$\boldsymbol{\phi}_i = \boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{B}_i\boldsymbol{b}_i, \qquad i = 1, \dots, n.$$

En nuestro ejemplo (Modelo poblacional):

Sea $\boldsymbol{\phi}_i = (\log k_{ai}, \log V_i, \log Cl_i)^T$, $\boldsymbol{b}_i = (b_{i1}, b_{i2}, b_{i3})^T$ y considere

$$k_{ai} = \exp(\beta_1 + b_{i1}), \quad V_i = \exp(\beta_2 + b_{i2}), \quad Cl_i = \exp(\beta_3 + b_{i3}).$$

Este modelo fuerza positividad de los parámetros farmacocinéticos.

Sea $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{im_i})^T$, $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{im_i})^T$ y

$$\boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{\phi}_i) = (f(\boldsymbol{x}_{i1}, \boldsymbol{\phi}_i), \ldots, f(\boldsymbol{x}_{in_i}, \boldsymbol{\phi}_i))^T.$$

Supuesto distribucional:

Considere

$$\boldsymbol{Y}_i|\boldsymbol{\phi}_i \overset{\text{ind}}{\sim} N_{m_i}(\boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{\phi}_i), \sigma^2 \boldsymbol{I}_{m_i}), \qquad \text{y}$$

$$\boldsymbol{\phi}_i = \boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{b}_i, \qquad \boldsymbol{b}_i \overset{\text{ind}}{\sim} N_q(\boldsymbol{0}, \boldsymbol{\Psi}), \tag{3}$$

para $i = 1, \ldots, n$, con $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\boldsymbol{\theta})$

Modelo no lineal con efectos mixtos (Vonesh y Carter, 1992)

Vonesh y Carter (1992) desarrollaron un modelo mixto no lineal, pero lineal en los efectos aleatorios

$$\boldsymbol{Y}_i = \boldsymbol{f}(\boldsymbol{z}_i, \boldsymbol{\beta}) + \boldsymbol{Z}_i(\boldsymbol{\beta})\boldsymbol{b}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, n,$$

donde $\boldsymbol{Z}_i$ es una matriz de rango completo de funciones conocidas que depende de los efectos fijos $\boldsymbol{\beta}$. Además se asume que

$$\boldsymbol{b}_i \overset{\text{ind}}{\sim} N(\boldsymbol{0}, \boldsymbol{\Psi}), \qquad \boldsymbol{\epsilon}_i \overset{\text{ind}}{\sim} N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \qquad \boldsymbol{b}_i \perp \boldsymbol{\epsilon}_i$$

Nota:

Este enfoque se concentra principalmente en inferencias relacionadas a los efectos fijos.

Estimación máximo verosímil en (3) se basa en la distribución marginal para la respuesta observada $\boldsymbol{Y}$

$$f(\boldsymbol{y};\boldsymbol{\beta},\sigma^2,\boldsymbol{\theta}) = \int_{\mathbb{R}^q} f(\boldsymbol{y}|\boldsymbol{b};\boldsymbol{\beta},\sigma^2) f(\boldsymbol{b};\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{b}.$$

Note que,

$$\begin{aligned} L(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) &= \frac{|\boldsymbol{\Delta}|^n}{(2\pi\sigma^2)^{(N+nq)/2}} \\ &\quad \times \prod_{i=1}^n \int_{\mathbb{R}^q} \exp\Big\{-\frac{1}{2\sigma^2}(\|\boldsymbol{Y}_i - \boldsymbol{f}_i(\boldsymbol{z}_i,\boldsymbol{b}_i)\|^2 + \|\boldsymbol{\Delta}\boldsymbol{b}_i\|^2)\Big\}\,\mathrm{d}\boldsymbol{b}_i, \end{aligned}$$

en general esta integral no tiene solución explícita cuando $\boldsymbol{f}$ es no lineal en $\boldsymbol{b}_i$.

Modelo no lineal con efectos mixtos

Algunos procedimientos de estimación:

- Aproximación de Laplace (Wolfinger, 1993; Vonesh, 1996; Pinheiro y Bates, 1995).
- Procedimientos de linealización (Lindstrom y Bates, 1990; Vonesh y Carter, 1992).
- Integración numérica (Pinheiro y Bates, 1995).
- Algoritmos MCEM (Walker, 1996) y SAEM (Kuhn y Lavielle, 2005).

EX UMBRA IN SOLEM

Note que

$$L(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2;\boldsymbol{y}) = \prod_{i=1}^{n} L_i(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2;\boldsymbol{y}_i),$$

donde

$$L_i(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2;\boldsymbol{y}_i) = \int_{\mathbb{R}^q} (2\pi\sigma^2)^{-(n_i+q)/2} |\boldsymbol{\Delta}| \exp\{-g(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\boldsymbol{b}_i)/2\sigma^2\}\,\mathrm{d}\boldsymbol{b}_i$$

con

$$g(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\boldsymbol{b}_i) = \|\boldsymbol{Y}_i - \boldsymbol{f}_i(\boldsymbol{z}_i,\boldsymbol{b}_i)\|^2 + \|\boldsymbol{\Delta}\boldsymbol{b}_i\|^2.$$

Usando la aproximación de Laplace en NLME

Sea

$$\widehat{\boldsymbol{b}}_i = \widehat{\boldsymbol{b}}_i(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i) = \arg\min_{\theta} g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i),$$

$$\dot{g}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i) = \frac{\partial g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i},$$

$$\ddot{g}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i) = \frac{\partial^2 g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i \partial \boldsymbol{b}_i^T},$$

y considere la expansión de Taylor de segundo orden

$$g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i) \approx g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \widehat{\boldsymbol{b}}_i) + \frac{1}{2}(\boldsymbol{b}_i - \widehat{\boldsymbol{b}}_i)^T \ddot{g}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \widehat{\boldsymbol{b}}_i)(\boldsymbol{b}_i - \widehat{\boldsymbol{b}}_i).$$

Recuerde que $\dot{g}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \widehat{\boldsymbol{b}}_i) = \boldsymbol{0}$.

Usando la aproximación de Laplace en NLME

La aproximación de Laplace es definida como

$$L(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}}|\boldsymbol{\Delta}|^n \exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^n g(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\widehat{\boldsymbol{b}}_i)\Big\}$$
$$\times\prod_{i=1}^n\int_{\mathbb{R}^q}(2\pi\sigma^2)^{\frac{q}{2}}\exp\Big\{-\frac{1}{2\sigma^2}(\boldsymbol{b}_i-\widehat{\boldsymbol{b}}_i)^T\ddot{g}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\widehat{\boldsymbol{b}}_i)(\boldsymbol{b}_i-\widehat{\boldsymbol{b}}_i)\Big\}\,\mathrm{d}\boldsymbol{b}_i$$

esto es,

$$L(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}}|\boldsymbol{\Delta}|^n \exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^n g(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\widehat{\boldsymbol{b}}_i)\Big\}$$
$$\times\prod_{i=1}^n|\ddot{g}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\widehat{\boldsymbol{b}}_i)|^{-\frac{1}{2}}.$$

Note que

$$\dot{g}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i) = -2\frac{\partial \boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i}(\boldsymbol{Y}_i - \boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{b}_i)) + 2\boldsymbol{\Delta}^T\boldsymbol{\Delta}\boldsymbol{b}_i$$

y podemos utilizar la siguiente iteración para obtener $\widehat{\boldsymbol{b}}_i$:

$$\widehat{\boldsymbol{b}}_i^{(r+1)} = (\boldsymbol{\Delta}^T\boldsymbol{\Delta})^{-1}\frac{\partial \boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i}\Big|_{b_i = \widehat{b}_i^{(r)}}(\boldsymbol{Y}_i - \boldsymbol{f}_i(\boldsymbol{z}_i, \widehat{\boldsymbol{b}}_i^{(r)})).$$

Usando la aproximación de Laplace en NLME

Finalmente, considerando la siguiente aproximación para la matriz Hessiana

$$\ddot{g}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\boldsymbol{b}_i) \approx \boldsymbol{G}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i) = \frac{\partial \boldsymbol{f}_i(\boldsymbol{z}_i,\boldsymbol{b}_i)}{\partial \boldsymbol{b}_i^T}\Big|_{b_i=\widehat{b}_i} \frac{\partial \boldsymbol{f}_i(\boldsymbol{z}_i,\boldsymbol{b}_i)}{\partial \boldsymbol{b}_i}\Big|_{b_i=\widehat{b}_i} + \boldsymbol{\Delta}^T\boldsymbol{\Delta}.$$

Tenemos que la aproximación Laplace para la log-verosimilitud está dada por

$$\begin{aligned}\ell_{\mathsf{Lapl}}(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) &= -\frac{N}{2}\log(2\pi\sigma^2) + n\log|\boldsymbol{\Delta}| \\ &\quad - \frac{1}{2}\Big\{\sum_{i=1}^n \log|\boldsymbol{G}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i)| + \sum_{i=1}^n g(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\widehat{\boldsymbol{b}}_i)/\sigma^2\Big\}.\end{aligned}$$

Usando la aproximación de Laplace en NLME

Note que $\widehat{\boldsymbol{b}}_i$ no depende de σ^2, para $\boldsymbol{\beta}$ y $\boldsymbol{\theta}$ fijados el MLE de σ^2 es

$$\widehat{\sigma}^2 = \widehat{\sigma}^2(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}) = \frac{1}{N} \sum_{i=1}^{n} g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \widehat{\boldsymbol{b}}_i).$$

Perfilando $\ell_{\mathsf{Lapl}}(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2)$ en $\widehat{\sigma}^2$, obtenemos

$$\ell^*_{\mathsf{Lapl}}(\boldsymbol{\beta}, \boldsymbol{\theta}) = -\frac{N}{2}[1 + \log(2\pi\widehat{\sigma}^2)] + n \log|\boldsymbol{\Delta}| - \frac{1}{2} \sum_{i=1}^{n} \log|\boldsymbol{G}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)|.$$

Suponga el modelo en dos etapas

$$\boldsymbol{Y}_i = \boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{\phi}_i) + \boldsymbol{\epsilon}_i, \qquad \boldsymbol{\epsilon}_i \overset{\text{ind}}{\sim} N_{m_i}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}),$$

y

$$\boldsymbol{\phi}_i = \boldsymbol{A}_i \boldsymbol{\beta} + \boldsymbol{b}_i, \qquad \boldsymbol{b}_i \overset{\text{ind}}{\sim} N_q(\boldsymbol{0}, \boldsymbol{\Psi}),$$

para $i = 1, \dots, n$, con $\boldsymbol{\epsilon}_i \perp \boldsymbol{b}_i$ y $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\boldsymbol{\lambda})$.

Sea $\boldsymbol{\eta}_i(\boldsymbol{\phi}_i) = \boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{\phi}_i)$ y considere una aproximación de Taylor de primer orden de $\boldsymbol{\eta}_i$ en torno de $\boldsymbol{b}_i^*$,

$$\boldsymbol{\eta}_i(\boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{b}_i) \approx \boldsymbol{\eta}_i(\boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{b}_i^*) + \frac{\partial \boldsymbol{\eta}_i(\boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i^T}\Big|_{b_i = b_i^*} (\boldsymbol{b}_i - \boldsymbol{b}_i^*),$$

y defina[1]

$$\boldsymbol{Z}_i(\boldsymbol{\beta}, \boldsymbol{b}_i^*) = \frac{\partial \boldsymbol{\eta}_i(\boldsymbol{\phi}_i)}{\partial \boldsymbol{b}_i^T}\Big|_{b_i = b_i^*}.$$

[1] en ocasiones sólo escribiremos $\boldsymbol{Z}_i^*(\boldsymbol{\beta})$.

Aproximación de primer orden

Para la elección de $\boldsymbol{b}_i^*$ en la expansión de Taylor, se ha propuesto:

- El valor promedio de los $\boldsymbol{b}_i$s, es decir, $\mathrm{E}(\boldsymbol{b}_i) = \mathbf{0}$ (FO).
- Alguna estimación para $\boldsymbol{b}_i$, digamos la esperanza condicional de los efectos aleatorios $\widehat{\boldsymbol{b}}_i$ (FOCE).

Otra alternativa:

- Expandir $\boldsymbol{\eta}(\boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{b}_i)$ en torno de ambos $\widehat{\boldsymbol{\beta}}$ y $\widehat{\boldsymbol{b}}_i$ (Aproximación LME[2])

[2] implementada en biblioteca nlme.

Para cada una de las alternativas anteriores, se aproxima la distribución condicional de $\boldsymbol{Y}_i$ como

$$\boldsymbol{Y}_i|\boldsymbol{b}_i \mathrel{\dot{\sim}} N_{m_i}(\boldsymbol{\eta}_i(\boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{b}_i^*) - \boldsymbol{Z}_i^*\boldsymbol{b}_i^* + \boldsymbol{Z}_i^*\boldsymbol{b}_i, \sigma^2\boldsymbol{I}),$$

además $\boldsymbol{b}_i \sim N_q(\boldsymbol{0}, \boldsymbol{\Psi})$.

De este modo, tenemos la siguiente aproximación para la distribución marginal

$$\boldsymbol{Y}_i \mathrel{\dot{\sim}} N_{m_i}(\boldsymbol{\eta}_i(\boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{b}_i^*) - \boldsymbol{Z}_i^*\boldsymbol{b}_i^*, \sigma^2\boldsymbol{V}_i^*),$$

donde

$$\boldsymbol{V}_i^* = \boldsymbol{Z}_i^*\boldsymbol{D}\boldsymbol{Z}_i^{*T} + \boldsymbol{I}, \qquad \text{con } \boldsymbol{\Psi} = \sigma^2\boldsymbol{D}.$$

Para el modelo poblacional

$$\boldsymbol{\phi}_i = \boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i), \qquad i = 1, \dots, n,$$

también podemos considerar una aproximación de primer orden,

$$\begin{aligned}\boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i) &\approx \boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i^*) + \frac{\partial \boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i^T}\Big|_{b_i = b_i^*} (\boldsymbol{b}_i - \boldsymbol{b}_i^*) \\ &= \boldsymbol{g}_i^*(\boldsymbol{a}_i, \boldsymbol{\beta}) + \boldsymbol{B}_i^*(\boldsymbol{\beta})(\boldsymbol{b}_i - \boldsymbol{b}_i^*)\end{aligned}$$

con

$$\boldsymbol{g}_i^*(\boldsymbol{a}_i, \boldsymbol{\beta}) = \boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i^*), \qquad \boldsymbol{B}_i^*(\boldsymbol{\beta}) = \frac{\partial \boldsymbol{d}(\boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i^T}\Big|_{b_i = b_i^*}$$

Linealización FO

Cuando consideramos $\boldsymbol{b}_i^* = \mathbf{0}$ $(= \mathsf{E}(\boldsymbol{b}_i))$ tenemos el modelo jerárquico

$$\begin{aligned} \boldsymbol{Y}_i | \boldsymbol{b}_i &\dot{\sim} N_{m_i}(\boldsymbol{\eta}_i(\boldsymbol{\beta}) + \boldsymbol{Z}_i(\boldsymbol{\beta})\boldsymbol{b}_i, \sigma^2 \boldsymbol{I}) \\ \boldsymbol{b}_i &\sim N_q(\mathbf{0}, \boldsymbol{\Psi}), \end{aligned}$$

con distribución marginal[3].

$$\boldsymbol{Y}_i \,\dot{\sim}\, N_{m_i}(\boldsymbol{\eta}_i(\boldsymbol{\beta}), \boldsymbol{Z}_i(\boldsymbol{\beta})\boldsymbol{\Psi}\boldsymbol{Z}_i^T(\boldsymbol{\beta}) + \sigma^2 \boldsymbol{I}),$$

donde $\boldsymbol{Z}_i(\boldsymbol{\beta}) = \boldsymbol{Z}_i(\boldsymbol{\beta}, \mathbf{0})$.

La aproximación FO de la log-verosimilitud marginal está dada por

$$\begin{aligned} \ell_{\mathsf{FO}}(\boldsymbol{\theta}) = -\frac{1}{2}\Big\{ N \log \sigma^2 + \sum_{i=1}^n \log |\boldsymbol{Z}_i(\boldsymbol{\beta})\boldsymbol{D}\boldsymbol{Z}_i^T(\boldsymbol{\beta}) + \boldsymbol{I}| \Big\} \\ - \frac{1}{2\sigma^2} \sum_{i=1}^n (\boldsymbol{Y}_i - \boldsymbol{\eta}_i(\boldsymbol{\beta}))^T (\boldsymbol{Z}_i(\boldsymbol{\beta})\boldsymbol{D}\boldsymbol{Z}_i^T(\boldsymbol{\beta}) + \boldsymbol{I})^{-1} (\boldsymbol{Y}_i - \boldsymbol{\eta}_i(\boldsymbol{\beta})). \end{aligned}$$

y utilizamos lme para realizar la estimación de parámetros.

[3] compare con modelo de Vonesh y Carter (1992)

Linealización FOCE

Considerando $\boldsymbol{b}_i^* = \widehat{\boldsymbol{b}}_i$ $(= \mathsf{E}(\boldsymbol{b}_i|\boldsymbol{Y}_i))$ tenemos el modelo jerárquico

$$\begin{aligned}
\boldsymbol{Y}_i|\boldsymbol{b}_i &\,\dot{\sim}\, N_{m_i}(\boldsymbol{\eta}_i(\boldsymbol{\beta}, \widehat{\boldsymbol{b}}_i) - \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\widehat{\boldsymbol{b}}_i + \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\boldsymbol{b}_i, \sigma^2 \boldsymbol{I}),\\
\boldsymbol{b}_i &\sim N_q(\boldsymbol{0}, \boldsymbol{\Psi}),
\end{aligned}$$

en cuyo caso la distribución marginal asume la forma

$$\boldsymbol{Y}_i \,\dot{\sim}\, N_{m_i}(\boldsymbol{\eta}_i(\boldsymbol{\beta}, \widehat{\boldsymbol{b}}_i) - \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\widehat{\boldsymbol{b}}_i, \sigma^2 \widehat{\boldsymbol{V}}_i(\boldsymbol{\beta})),$$

donde

$$\widehat{\boldsymbol{V}}_i(\boldsymbol{\beta}) = \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\boldsymbol{D}\widehat{\boldsymbol{Z}}_i^T(\boldsymbol{\beta}) + \boldsymbol{I},$$

con $\widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta}) = \boldsymbol{Z}_i(\boldsymbol{\beta}, \widehat{\boldsymbol{b}}_i)$.

Linealización FOCE

Sea

$$\widehat{\boldsymbol{w}}_i(\boldsymbol{\beta}) = \boldsymbol{Y}_i - \boldsymbol{\eta}_i(\boldsymbol{\beta}, \widehat{\boldsymbol{b}}_i) + \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\widehat{\boldsymbol{b}}_i,$$

luego, la aproximación FOCE para la log-verosimilitud marginal asume la forma

$$\begin{aligned}\ell_{\mathsf{FOCE}}(\boldsymbol{\theta}) &= -\frac{1}{2}\Big\{N\log\sigma^2 + \sum_{i=1}^n \log|\widehat{\boldsymbol{Z}}(\boldsymbol{\beta})\boldsymbol{D}\widehat{\boldsymbol{Z}}_i^T(\boldsymbol{\beta}) + \boldsymbol{I}|\Big\} \\ &\quad - \frac{1}{2\sigma^2}\sum_{i=1}^n (\widehat{\boldsymbol{w}}_i - \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\boldsymbol{b}_i)^T(\widehat{\boldsymbol{Z}}(\boldsymbol{\beta})\boldsymbol{D}\widehat{\boldsymbol{Z}}_i^T(\boldsymbol{\beta}) + \boldsymbol{I})^{-1}(\widehat{\boldsymbol{w}}_i - \widehat{\boldsymbol{Z}}_i(\boldsymbol{\beta})\boldsymbol{b}_i).\end{aligned}$$

y utilizamos lme para realizar la estimación de parámetros.

Considere la siguiente aproximación de primer orden en torno de $\widehat{\boldsymbol{\beta}}$ y $\widehat{\boldsymbol{b}}_i$,

$$\boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i) \approx \boldsymbol{\eta}_i(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}_i) + \left(\frac{\partial \boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)}{\partial \boldsymbol{\beta}^T}, \frac{\partial \boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i^T}\right)\Big|_{\beta=\widehat{\beta}, b_i=\widehat{b}_i} \begin{pmatrix} \boldsymbol{\beta} - \widehat{\boldsymbol{\beta}} \\ \boldsymbol{b}_i - \widehat{\boldsymbol{b}}_i \end{pmatrix}$$

defina

$$\widehat{\boldsymbol{X}}_i = \frac{\partial \boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)}{\partial \boldsymbol{\beta}^T}\Big|_{\beta=\widehat{\beta}, b_i=\widehat{b}_i}, \qquad \widehat{\boldsymbol{Z}}_i = \frac{\partial \boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i^T}\Big|_{\beta=\widehat{\beta}, b_i=\widehat{b}_i},$$

de este modo, tenemos

$$\boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i) \approx \boldsymbol{\eta}_i(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}_i) + \widehat{\boldsymbol{X}}_i(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) + \widehat{\boldsymbol{Z}}_i(\boldsymbol{b}_i - \widehat{\boldsymbol{b}}_i).$$

Usando la aproximación LME, tenemos el modelo jerárquico

$$\begin{aligned}\boldsymbol{Y}_i|\boldsymbol{b}_i &\stackrel{\cdot}{\sim} N_{m_i}(\boldsymbol{\eta}_i(\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{b}}_i) - \widehat{\boldsymbol{X}}_i\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{Z}}_i\widehat{\boldsymbol{b}}_i + \widehat{\boldsymbol{X}}_i\boldsymbol{\beta} + \widehat{\boldsymbol{Z}}_i\boldsymbol{b}_i, \sigma^2\boldsymbol{I}),\\ \boldsymbol{b}_i &\sim N_q(\boldsymbol{0},\boldsymbol{\Psi}),\end{aligned}$$

con distribución marginal (aproximada)

$$\boldsymbol{Y}_i \stackrel{\cdot}{\sim} N_{m_i}(\boldsymbol{\eta}_i(\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{b}}_i) - \widehat{\boldsymbol{X}}_i\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{Z}}_i\widehat{\boldsymbol{b}}_i + \widehat{\boldsymbol{X}}_i\boldsymbol{\beta}, \sigma^2\widehat{\boldsymbol{V}}_i),$$

donde

$$\widehat{\boldsymbol{V}}_i = \widehat{\boldsymbol{Z}}_i\boldsymbol{D}\widehat{\boldsymbol{Z}}_i^T + \boldsymbol{I}.$$

Etapa PNLS

$\widehat{\boldsymbol{\beta}}$ y $\widehat{\boldsymbol{b}}_i$ son obtenidos por medio de resolver un problema de mínimos cuadrados no lineal penalizado (PNLS)

$$\min_{\beta, b_i} \sum_{i=1}^{n} \Big\{ \|\boldsymbol{Y}_i - \boldsymbol{\eta}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)\|^2 + \boldsymbol{b}_i \boldsymbol{D}^{-1} \boldsymbol{b}_i \Big\},$$

y defina

$$\widehat{\boldsymbol{w}}_i = \boldsymbol{Y}_i - \boldsymbol{\eta}_i(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}_i) + \widehat{\boldsymbol{X}}_i \widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{Z}}_i \widehat{\boldsymbol{b}}_i,$$

para $i = 1, \dots, n$.

Etapa LME

La aproximación LME para la log-verosimilitud marginal está dada por

$$\ell_{\mathsf{LME}}(\boldsymbol{\theta}) = -\frac{1}{2}\Big\{N\log\sigma^2 + \sum_{i=1}^{n}\log|\widehat{\boldsymbol{Z}}_i\boldsymbol{D}\widehat{\boldsymbol{Z}}_i^T + \boldsymbol{I}|\Big\}$$

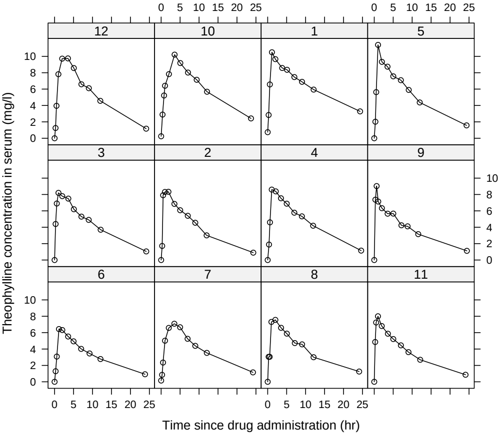
$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\widehat{\boldsymbol{w}}_i - \widehat{\boldsymbol{X}}_i\boldsymbol{\beta})^T(\widehat{\boldsymbol{Z}}_i\boldsymbol{D}\widehat{\boldsymbol{Z}}_i^T + \boldsymbol{I})^{-1}(\widehat{\boldsymbol{w}}_i - \widehat{\boldsymbol{X}}_i\boldsymbol{\beta}).$$

y se utiliza lme para estimar de parámetros de escala $\boldsymbol{D}$ y σ^2.

Observaciones:

- Opcionalmente también es posible actualizar las estimaciones $\widehat{\boldsymbol{\beta}}$ y $\widehat{\boldsymbol{b}}_i$ en la etapa LME (Wolfinger, 1993).
- La estrategia anterior converge a los mismos valores que los obtenidos por el algoritmo alternante (PNLS-LME), aunque más lentamente.
- Wolfinger (1993) mostró que la aproximación LME para REML es equivalente a la aproximación de Laplace asumiendo una priori plana para $\boldsymbol{\beta}$.

Dosis orales de droga anti-asmática Teofilina son suministradas a doce individuos, luego las concentraciones en la sangre (mg/L) son medidas en 11 instantes en un periodo de 25 horas.



Modelo:

$$Y_{ij} = \frac{\Delta_i K_i k_{ai}}{Cl_i(k_{ai} - K_i)}\{\exp(-K_i t_{ij}) - \exp(-k_{ai} t_{ij})\}, \tag{4}$$

donde Δ_i representa la dosis inicial, k_{ai} y K son las constantes de absorción y eliminación, respectivamente y Cl_i es el Clearance, con

$$k_{ai} = \exp(\beta_1 + b_{i1}), \quad K_i = \exp(\beta_2 + b_{i2}), \quad Cl_i = \exp(\beta_3 + b_{i3}).$$

donde $\boldsymbol{b}_i = (b_{1i}, b_{2i}, b_{3i})^T \overset{\text{ind}}{\sim} N_3(\mathbf{0}, \boldsymbol{\Psi})$.

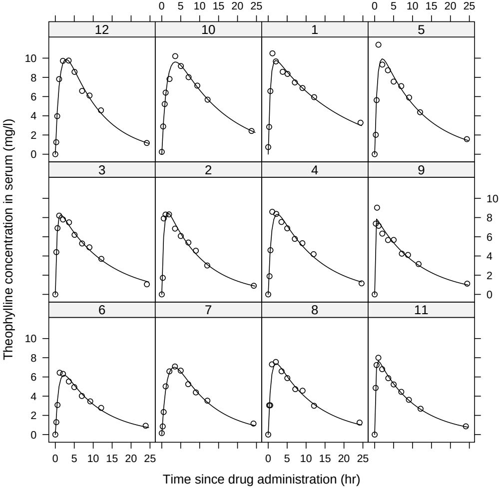
Cinética de Teofilina (ajuste individual)

Ajustando el modelo sin considerar los efectos aleatorios (usando función `nlsList`), tenemos

i	$\log \widehat{K}_i$	$\log \widehat{k}_{a_i}$	$\log \widehat{Cl}_i$
1	-2.920	0.575	-3.916
2	-2.286	0.664	-3.106
3	-2.508	0.898	-3.230
4	-2.436	0.158	-3.286
5	-2.425	0.386	-3.133
6	-2.307	0.152	-2.973
7	-2.280	-0.386	-2.964
8	-2.386	0.319	-3.069
9	-2.446	2.182	-3.421
10	-2.604	-0.363	-3.428
11	-2.322	1.348	-2.860
12	-2.248	-0.183	-3.170

además $\widehat{\sigma}^2 = 0.490$.

Cinética de Teofilina



Realizando el ajuste del modelo en (4) (usando función nlme), tenemos

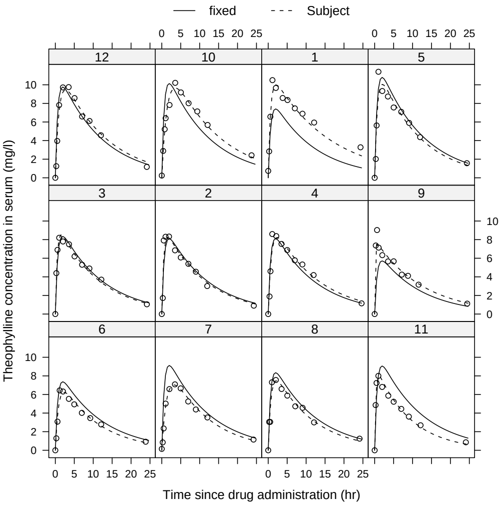
$$\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)^T = (0.451, -2.433, -3.214)^T, \qquad \widehat{\sigma}^2 = 0.465,$$

$$\widehat{\boldsymbol{C}} = \begin{pmatrix} 1.000 & 0.012 & 0.995 \\ 0.012 & 1.000 & -.089 \\ 0.995 & -.089 & 1.000 \end{pmatrix}, \qquad \widehat{\boldsymbol{V}} = \operatorname{diag}(0.131, 0.638, 0.251),$$

i	$\widehat{b_{1i}}$	$\widehat{b_{2i}}$	$\widehat{b_{3i}}$	i	$\widehat{b_{1i}}$	$\widehat{b_{2i}}$	$\widehat{b_{3i}}$
1	-0.289	0.038	-0.554	7	0.097	-0.777	0.217
2	0.023	0.284	0.034	8	0.078	-0.163	0.155
3	0.029	0.331	0.042	9	-0.113	1.501	-0.276
4	-0.055	-0.242	-0.095	10	-0.106	-0.831	-0.169
5	0.053	-0.098	0.105	11	0.220	0.699	0.393
6	0.108	-0.295	0.219	12	-0.046	-0.447	-0.070

donde $\widehat{\boldsymbol{\Psi}} = \widehat{\boldsymbol{V}}\widehat{\boldsymbol{C}}\widehat{\boldsymbol{V}}$. Además $\ell_{\mathsf{LME}}(\widehat{\boldsymbol{\theta}}) = -173.321$ y $AIC = 366.641$.

Cinética de Teofilina



Cinética de Teofilina

Considerando que $\boldsymbol{b}_i \sim N_3(\mathbf{0}, \boldsymbol{\Psi})$ con $\boldsymbol{\Psi} = \mathrm{diag}(\psi_1, \psi_2, \psi_3)$, obtenemos

$$\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)^T = (0.465, -2.455, -3.227)^T, \qquad \widehat{\sigma}^2 = 0.503,$$
$$\widehat{\boldsymbol{\Psi}} = \mathrm{diag}(\widehat{\psi}_1, \widehat{\psi}_2, \widehat{\psi}_3) = \mathrm{diag}(0.000, 0.644, 0.167),$$

además $\ell_{\mathsf{LME}}(\widehat{\boldsymbol{\theta}}) = -177.024$ y $AIC = 368.047$.

Por otro lado si consideramos el modelo poblacional

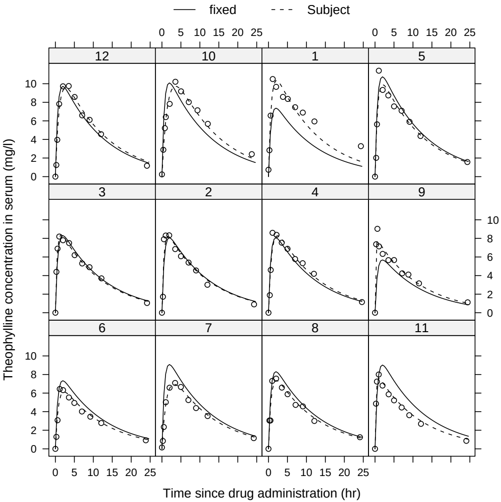
$$\log k_{a_i} = \beta_1 + b_{1i}, \qquad \log Cl_i = \beta_3 + b_{2i}$$

con $\boldsymbol{b}_i \sim N_2(\mathbf{0}, \boldsymbol{\Psi})$ talque $\boldsymbol{\Psi} = \mathrm{diag}(\psi_1, \psi_2)$, obtenemos

$$\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3)^T = (0.466, -2.455, -3.227)^T, \qquad \widehat{\sigma}^2 = 0.503,$$
$$\widehat{\boldsymbol{\Psi}} = \mathrm{diag}(\widehat{\psi}_1, \widehat{\psi}_2) = \mathrm{diag}(0.644, 0.167),$$

además $\ell_{\mathsf{LME}}(\widehat{\boldsymbol{\theta}}) = -177.021$ y $AIC = 366.043$.

Cinética de Teofilina



Cuadratura Gaussiana

Considere $h(t)$ función regular (uni-dimensional) y

$$g(t) = h(t)(2\pi\sigma^2)^{-1/2} \exp\Big\{ -\tfrac{1}{2}\big(\tfrac{t-\mu}{\sigma}\big)^2\Big\}.$$

Cuadratura gaussiana utiliza la siguiente aproximación

$$\int_{\mathbb{R}} g(t)\,\mathrm{d}t \approx \sum_{j=1}^{M} m_j\, g(z_j),$$

donde

$$m_j = w_j \exp(t_j^2)\sqrt{2}\sigma, \qquad z_j = \mu + \sqrt{2}\sigma t_j.$$

Para implementar este enfoque, tablas de t_j, w_j y $w_j \exp(t_j^2)$ son requeridas (Naylor y Smith, 1982).

- Aproxima integrales con relación a un kernel fijado por un promedio ponderado del integrando evaluado en abscisas predeterminadas.
- Los pesos y abscisas usadas en la cuadratura Gaussiana para los kernels usuales pueden ser obtenidos desde tablas (Abramowitz y Stegun, 1964) o por un algoritmo propuesto por Golub (1973).
- Es bien conocido que las reglas para cuadratura Gaussiana en el caso de integrales múltiples es numéricamente complejo (curse of dimensionality).

Función de verosimilitud

Se desea aproximar la función de verosimilitud

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) = \prod_{i=1}^{n} \int_{\mathbb{R}^q} (2\pi\sigma^2)^{-(n_i+q)/2} |\boldsymbol{\Delta}| \exp\{-g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i)/2\sigma^2\} \, \mathrm{d}\boldsymbol{b}_i$$

donde

$$g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i) = \|\boldsymbol{Y}_i - \boldsymbol{f}_i(\boldsymbol{z}_i, \boldsymbol{b}_i)\|^2 + \|\boldsymbol{\Delta}\boldsymbol{b}_i\|^2.$$

Sea

$$\widehat{\boldsymbol{b}}_i = \widehat{\boldsymbol{b}}_i(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i) = \arg\min_{\theta} g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i),$$

$$\boldsymbol{G}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i) = \frac{\partial^2 g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_i)}{\partial \boldsymbol{b}_i \partial \boldsymbol{b}_i^T}\Big|_{b_i = \widehat{b}_i}.$$

Por Slide 13, tenemos que el integrando es (salvo una constante) aproximadamente igual a la densidad[4].

$$N_q(\widehat{\boldsymbol{b}}_i, \sigma^2 \boldsymbol{G}^{-1}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)).$$

[4] que es una elección natural para la distribución de importancia

Obtenemos una observación desde $N_q(\widehat{\boldsymbol{b}}_i, \sigma^2 \boldsymbol{G}^{-1}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i))$, tomando $\boldsymbol{z}^* \sim N(\boldsymbol{0}, \boldsymbol{I})$ y calculando

$$\boldsymbol{b}_i^* = \widehat{\boldsymbol{b}}_i + \sigma \boldsymbol{G}^{-1/2}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)\boldsymbol{z}^*,$$

donde $\boldsymbol{G}^{1/2}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)$ es el factor Cholesky de $\boldsymbol{G}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)$.

Repetimos el procedimiento anterior M veces (número de muestras de importancia).

La aproximación importance sampling de la log-verosimilitud marginal está dada por

$$\ell_{\mathsf{IS}}(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) = -\frac{1}{2}\Big\{N \log(2\pi\sigma^2) + n \log |\boldsymbol{D}| + \sum_{i=1}^n \log |\boldsymbol{G}(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2)|\Big\}$$
$$+ \sum_{i=1}^n \log\Big\{\sum_{j=1}^M \exp[-g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \boldsymbol{b}_{ij}^*)/2\sigma^2 + \|\boldsymbol{z}_j^*\|^2/2]/M\Big\}.$$

En este caso no es posible obtener expresiones en forma cerrada del MLE de σ^2 para $\boldsymbol{\beta}$ y $\boldsymbol{\theta}$ fijos.

Idea:

Aprovechar la estructura del integrando en nlme para transformar el problema en la aplicación sucesiva de reglas de cuadratura Gaussiana unidimensionales.

Sean z_j^*, w_j, $j = 1, \dots, M$, las abscisas y los pesos para la cuadratura Gaussiana (unidimensional) con M puntos basados en un kernel $N(0,1)$. Entonces

$$\begin{aligned}
&\int_{\mathbb{R}^q} (2\pi\sigma^2)^{-q/2} |\boldsymbol{D}|^{-1/2} \exp\{-\|\boldsymbol{y}_i - \boldsymbol{f}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)\|^2/2\sigma^2\} \exp(-\boldsymbol{b}_i^T \boldsymbol{D}^{-1} \boldsymbol{b}_i/2\sigma^2)\, \mathrm{d}\boldsymbol{b}_i \\
&\quad = \int_{\mathbb{R}^q} (2\pi)^{-q/2} \exp\{-\|\boldsymbol{y}_i - \boldsymbol{f}_i(\boldsymbol{\beta}, \sigma \boldsymbol{D}^{T/2} \boldsymbol{z}^*)\|^2/2\sigma^2\} \exp(-\|\boldsymbol{z}^*\|^2/2)\, \mathrm{d}\boldsymbol{z}^* \\
&\quad \approx \sum_{j_1=1}^{M} \cdots \sum_{j_q=1}^{M} \exp\{-\|\boldsymbol{y}_i - \boldsymbol{f}_i(\boldsymbol{\beta}, \sigma \boldsymbol{D}^{T/2} \boldsymbol{z}^*_{j_1,\dots,j_q})\|^2/2\sigma^2\} \prod_{k=1}^{q} w_{j_k}
\end{aligned}$$

Luego, la aproximación para la función de log-verosimilitud usando cuadratura Gaussiana está dada por

$$\begin{aligned}\ell_{\mathsf{GQ}}(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) &= -\frac{N}{2}\log(2\pi\sigma^2) \\ &\quad + \sum_{i=1}^{n}\log\Big\{\sum_{\boldsymbol{j}}^{M}\exp(-\|\boldsymbol{y}_i - \boldsymbol{f}_i(\boldsymbol{\beta},\sigma\boldsymbol{D}^{T/2}\boldsymbol{z}_{\boldsymbol{j}}^{*})\|^2/2\sigma^2)\prod_{k=1}^{q} w_{j_k}\Big\}\end{aligned}$$

donde $\boldsymbol{j} = (j_1,\dots,j_q)^T$.

Aproximación por cuadratura Gaussiana adaptativa

Observaciones:

- Cuadratura Gaussiana corresponde a una versión determinista de integración Monte Carlo, donde las muestras de $\boldsymbol{b}_i$ son generadas desde $N_q(\mathbf{0}, \sigma^2 \boldsymbol{D})$.
- Cuadratura Gaussiana adaptativa (AGQ) es el homólogo determinista de Importance sampling.
- En AGQ, la grilla de abscisas en la escala $\boldsymbol{b}_i$ está centrada en $\widehat{\boldsymbol{b}}_i$ y $\boldsymbol{G}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)$ se utiliza para escalar $\boldsymbol{z}^*$.

La cuadratura Gaussiana adaptativa está dada por

$$
\begin{aligned}
&\int_{\mathbb{R}^q} (2\pi\sigma^2)^{-q/2} |\boldsymbol{D}|^{-1/2} \exp\{-\|\boldsymbol{y}_i - \boldsymbol{f}_i(\boldsymbol{\beta}, \boldsymbol{b}_i)\|^2/2\sigma^2\} \exp(-\boldsymbol{b}_i^T \boldsymbol{D}^{-1} \boldsymbol{b}_i/2\sigma^2)\,\mathrm{d}\boldsymbol{b}_i \\
&\qquad = \int_{\mathbb{R}^q} (2\pi)^{-q/2} |\boldsymbol{G}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)\boldsymbol{D}|^{-1/2} \\
&\qquad \times \exp\{-g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \widehat{\boldsymbol{b}}_i + \sigma \boldsymbol{G}^{-1/2}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)\boldsymbol{z}^*)/2\sigma^2 + \|\boldsymbol{z}^*\|^2/2\} \exp(-\|\boldsymbol{z}^*\|^2/2)\,\mathrm{d}\boldsymbol{z}^* \\
&\qquad \approx \sum_{j_1=1}^{M} \cdots \sum_{j_q=1}^{M} \exp\{-g(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i, \widehat{\boldsymbol{b}}_i + \sigma \boldsymbol{G}^{-1/2}(\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{y}_i)\boldsymbol{z}^*_{j_1,\dots,j_q})/2\sigma^2 \\
&\qquad + \|\boldsymbol{z}^*_{j_1,\dots,j_q}\|^2/2\} \prod_{k=1}^{q} w_{j_k}.
\end{aligned}
$$

La aproximación de la log-verosimilitud asume la forma

$$\ell_{\mathsf{AGQ}}(\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) = -\frac{1}{2}\Big\{N\log(2\pi\sigma^2) + n\log|\boldsymbol{D}| + \sum_{i=1}^{n}\log|\boldsymbol{G}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i)|\Big\}$$
$$+\sum_{i=1}^{n}\log\Big\{\sum_{\boldsymbol{j}}^{M}\exp\{-g(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i,\widehat{\boldsymbol{b}}_i+\sigma\boldsymbol{G}^{-1/2}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{y}_i)\boldsymbol{z}_{\boldsymbol{j}}^*)/2\sigma^2+\|\boldsymbol{z}_{\boldsymbol{j}}^*\|^2\}\prod_{k=1}^{q}w_{j_k}\Big\}$$

donde $\boldsymbol{j}=(j_1,\dots,j_q)^T$.

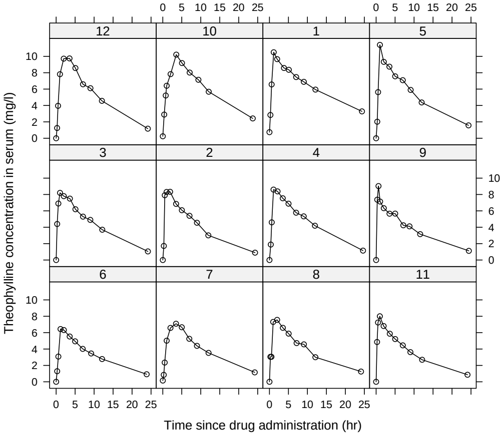
Aproximación por cuadratura Gaussiana adaptativa

Observaciones:

- AGQ usa abscisas y pesos fijos, mientras que Importance sampling las determina por medio de simulación.
- Cuando $M = 1$ cuadratura Gaussiana adaptativa reduce a la aproximación Laplace.
- AGQ es exacto cuando $\boldsymbol{f}$ es lineal en $\boldsymbol{b}_i$, esto no es verdad para cuadratura Gaussiana.

Cinética de Teofilina

Dosis orales de droga anti-asmática Teofilina son suministradas a doce individuos, luego las concentraciones en la sangre (mg/L) son medidas en 11 instantes en un periodo de 25 horas.



Modelo:

$$Y_{ij} = \frac{\Delta_i K k_{ai}}{Cl_i(k_{ai} - K)}\{\exp(-Kt_{ij}) - \exp(-k_{ai}t_{ij})\},$$

donde Δ_i representa la dosis inicial, k_{ai} y K son las constantes de absorción y eliminación, respectivamente y Cl_i es el Clearance, con

$$Cl_i = \exp(\beta_1 + b_{i1}), \quad k_{ai} = \exp(\beta_2 + b_{i2}), \quad K = \exp(\beta_3).$$

donde $\boldsymbol{b}_i = (b_{1i}, b_{2i})^T \overset{\text{ind}}{\sim} N_2(\mathbf{0}, \mathbf{\Psi})$.

Resultados de estimación:

Aproximación	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\log\widehat{\sigma}^2$	ℓ
LME	-3.22719	0.46548	-2.45464	-0.68660	-177.0237
Laplace	-3.22946	0.46876	-2.46432	-0.68658	-176.9995
Imp. sampling$_{1000}$	-3.22682	0.47614	-2.45851	-0.68747	-177.7689
Gaussiana$_5$	-3.30411	0.50046	-2.48743	-0.48395	-182.4680
Gaussiana$_{10}$	-3.23814	0.59525	-2.46872	-0.70276	-176.1008
Gaussiana$_{100}$	-3.22684	0.47947	-2.45893	-0.68539	-177.7293
Gaussiana Adap.$_5$	-3.22503	0.47566	-2.45788	-0.68677	-177.7499
Gaussiana Adap.$_{10}$	-3.22705	0.47377	-2.45942	-0.68533	-177.7473

Cinética de Teofilina (Pinheiro y Bates, 1995)

Número de evaluaciones funcionales hasta convergencia:

Aproximación	Evaluaciones funcionales
LME	1.512
Laplace	7.683
Gaussiana Adap.$_5$	30.020
Gaussiana Adap.$_{10}$	96.784
Gaussiana$_5$	47.700
Gaussiana$_{10}$	318.000
Gaussiana$_{100}$	10.200.000
Imp. sampling$_{1000}$	11.211.284