Revista de la Facultad de Ciencias Universidad Nacional de Colombia, Sede Medellín
V 1 N°2 diciembre 2012 • ISSN 0121-747X • Artículo Investigación
Páginas 69 - 82

COMPARACIÓN DE ALGUNOS R^2 COMO MEDIDAS DE BONDAD DE AJUSTE EN MODELOS LINEALES MIXTOS

COMPARISON OF SOME R^2 AS MEASURES OF GOODNESS OF FIT IN LINEAR MIXED MODELS

DIANA GUZMÁN [1], *JUAN SALAZAR*[2], *JUAN CORREA*[2], *CAMILO ZULUAICA*[3]

Recibido 22-10-12, aceptado 12-12-12, versión final 28-12-12

Artículo Investigación

RESUMEN: En este artículo se proponen dos estadísticos R^2 para evaluar la bondad de ajuste en modelos lineales mixtos con datos longitudinales. Se llevó a cabo un estudio de simulación para evaluar y comparar los estadísticos R^2 propuestos junto con otros estadísticos R^2 reportados en la literatura. Finalmente, se aplican los distintos R^2 a datos reales. Se llega a la conclusión que los estadísticos R^2 propuestos tiene una interpretación intuitiva, rango bien definido y son medidas que se pueden interpretar en un sentido absoluto sin hacer referencia a un modelo nulo. Sin embargo, el estudio de simulación evidencia que al igual que los otros estadísticos, los estadísticos R^2 propuestos, en algunas ocasiones, son incapaces de discriminar cuando las variables eliminadas del modelo son importantes.

PALABRAS CLAVE: Bondad de ajuste, coeficiente de determinación, datos longitudinales, modelos lineales mixtos.

ABSTRACT: A simulation study was conducted to evaluate and compare the R^2 statistics proposed together with other R^2 statistics reported in the literature. Finally, the studied R^2 statistics are illustrated using real data. In conclusion, the two proposed R^2 statistics have the following properties: intuitive interpretation, well defined range and they can be interpreted in an absolute scale without doing reference to a null model. However, the simulation study evidences that, as well as other R^2 statistics, the two proposed R^2 statistics sometimes are not able to discriminate when the removed variables from the model are important.

KEYWORDS: Coefficient of determination, goodness of fit, linear mixed models, longitudinal data.

[1]Magíster en Ciencias-Estadística, Escuela de Estadística, Universidad Nacional de Colombia, Medellín, Colombia, email: dsguzman@unal.edu.co
[2]Profesor Asociado, Escuela de Estadística, Universidad Nacional de Colombia, Medellín, Colombia.
[3]Matemático, Escuela de Matemáticas, Universidad Nacional de Colombia, Medellín, Colombia

1 INTRODUCCIÓN

Los modelos lineales con efectos mixtos son eficientes para analizar datos longitudinales y medidas repetidas, Liu et al. (2008) y Vonesh et al. (1996). Una gran número de investigadores han enfocado sus estudios en la modelación de la varianza, la estimación de los parámetros, las pruebas de significancia y las pruebas de bondad de ajuste. En este último aspecto, se han desarrollado varias medidas tales como el criterio de información de Akaike (AIC), el criterio de información bayesiano (BIC) y la prueba de razón de verosimilitud (LRT), comúnmente utilizadas en modelos lineales mixtos y que se encuentran disponibles en la mayoría de los paquetes estadísticos. Sin embargo, su uso se limita a la comparación de modelos con el fin de seleccionar algunos de ellos, a diferencia de los estadísticos R^2 tradicionales que, además de utilizarse como criterios de selección, ayudan a medir la proporción de la variación explicada de una manera intuitiva.

Como el coeficiente de determinación, el R^2 del modelo lineal clásico tiene una interpretación simple e intuitiva y además es fácil de calcular, hay un gran interés en desarrollar medidas de este tipo para modelos lineales mixtos que sean extensiones naturales de los R^2 de los modelos tradicionales y que posean las propiedades básicas de los R^2 que han sido enunciadas por Kvalseth (1985). El autor propone ocho criterios para evaluar los estadísticos R^2:

1. Los R^2 deben tener una interpretación razonable y útil como una medida de bondad de ajuste (Goodnes of fit, GOF siglas en inglés).

2. Los R^2 deben ser independientes de las unidades de medida.

3. El rango de valores potencial debe estar bien definido con criterios correspondientes para perfecto ajuste o total falta de ajuste.

4. Los R^2 deben ser lo suficientemente generales como para ser aplicados a cualquier tipo de modelo.

5. Los R^2 no deben estar confinados a una técnica específica de ajuste del modelo especificado. Es decir, sin importar que técnica de estimación del modelo se utilice debe ser factible calcular los estadísticos R^2.

6. Los R^2 se deben poder comparar cuando se apliquen al mismo conjunto de datos pero con diferentes modelos ajustados (comparación entre modelos).

7. Los valores relativos de los R^2 deben ser en general comparables con otros criterios razonables de bondad de ajuste (comparación entre criterios de bondad de ajuste).

8. Los residuales positivos y negativos deberían tener igual ponderación.

Recientemente, Orelien y Edwards (2007) reportaron resultados de simulación para evaluar la habilidad de varios estadísticos R^2 propuestos en la literatura en la elección del modelo más parsimonioso en el marco de modelos lineales mixtos, es decir, un modelo que captura un buen porcentaje de la información con pocas covariables. Los autores comparan el valor de los estadísticos R^2 de un modelo completo con los R^2 de otros modelos en donde se han removido covariables de efectos fijos. También comparan los R^2 del modelo completo con los de un modelo sobreajustado. En ese estudio, todos los modelos tienen los mismos efectos aleatorios.

Como una extensión del trabajo propuesto por Orelien y Edwards (2007), en este estudio se realiza un análisis de simulación para evaluar el desempeño de algunos estadísticos R^2 condicionales propuestos en la literatura, aplicados a modelos lineales mixtos, ajustando modelos en los cuales no se asume que las covariables de efectos aleatorios están bien especificadas. Además, se propone dos estadísticos R^2 y se evalúa el desempeño de los mismos como medidas de bondad de ajuste en un sentido absoluto, para el caso de los modelos lineales mixtos aplicados a datos longitudinales con respuesta continua, mediante diferentes escenarios de simulación donde se generan conjuntos de datos desbalanceados.

Este artículo está organizado de la siguiente forma: la sección 2 presenta la especificación matemática del modelo; en la sección 3 se presentan cinco estadísticos R^2 seleccionados de la literatura; en la sección 4 se propone dos estadísticos R^2 junto con sus propiedades; la sección 5 contiene una descripción y los resultados de un estudio de simulación; en la sección 6 se estiman los siete estadísticos R^2 para evaluar el ajuste de diferentes modelos ajustados a un conjunto de datos referente a un estudio longitudinal que analiza el crecimiento de 27 niños de ocho a catorce años de edad y finalmente, en la sección 7 se expone las conclusiones derivadas de los resultados obtenidos.

2 EL MODELO LINEAL MIXTO (MLM)

Para modelar datos con medidas repetidas, Laird y Ware (1982) suponen que se tienen m sujetos y que cada sujeto i ha sido observado n_i veces, por tanto, se tienen disponibles $N=\sum_{i=1}^{m} n_i$ observaciones en total. Sea $\boldsymbol{Y}_i$ el vector de respuestas $n_i \times 1$ del *i-ésimo* sujeto que satisface el siguiente modelo:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i, \tag{1}$$

donde $\mathbf{b}_i$ es el vector k-dimensional que contiene los efectos aleatorios del *i-ésimo* individuo, $\boldsymbol{\beta}$ es el vector de parámetros $p \times 1$ correspondiente a los efectos fijos, $\mathbf{X}_i$ es la matriz de diseño $n_i \times p$ de los efectos fijos específica para el *i-ésimo* sujeto, $\mathbf{Z}_i$ es una matriz de diseño $n_i \times k$ asociada a los efectos aleatorios y $\boldsymbol{e}_i$ es el vector de errores intra-sujeto. Se asume que $\mathbf{e}_i \sim N(0, \mathbf{R}_i)$ donde $\mathbf{R}_i$ es la matriz de covarianzas intra-individual de dimensión $n_i \times n_i$. Además, se supone que el vector de efectos aleatorios $\mathbf{b}_i$ proviene de una distribución normal con media cero y matriz de dispersión $\mathbf{D}_{k\times k}$ y se asume que los $\mathbf{b}_i$, $i = 1, 2, ..., m$ son independientes entre sí y del vector de errores $\mathbf{e}_i$. Con estos supuestos se tiene que, $\mathbf{Y}_i|\mathbf{b}_i \sim N(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i,\ \mathbf{R}_i)$ y $\mathbf{Y}_i \sim N(\mathbf{X}_i\boldsymbol{\beta},\ \mathbf{V}_i)$, donde $\mathbf{V}_i = Cov[\mathbf{Y}_i] = \mathbf{R}_i + \mathbf{Z}_i\mathbf{D}\mathbf{Z'}_i$.

3 ESTADÍSTICOS R^2 PARA MODELOS LINEALES MIXTOS

3.1 Estadístico R_1^2

Vonesh y Chinchilli (1997) propusieron un estadístico R^2 para evaluar bondad de ajuste en un MLM generalizado, el cual se denota R_1^2. Asumiendo $R_i = \sigma^2 I_{n_i}$, el estadístico R_1^2 se define como:

$$R_1^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)'(y_i - \hat{y}_i)}{\sum_{i=1}^m (y_i - \bar{y}1_{n_i})'(y_i - \bar{y}1_{n_i})}, \tag{2}$$

con m el número de individuos, y_i el vector de valores observados del i-ésimo sujeto; $\hat{y}$ el vector de valores predichos, $\bar{y}$ la media de los elementos de todos los valores observados y 1_{n_i} un vector $n_i \times 1$ de unos. El R_1^2 coincide con el tradicional estadístico R^2 para modelos lineales cuando $m = 1$, por lo tanto R_1^2 es fácil de interpretar. Sin embargo, R_1^2 no considera explícitamente las componentes aleatorias del modelo.

3.2 Estadístico R_c

Motivado por el coeficiente de correlación de concordancia ρ_c, Vonesh et al. (1996) proponen

$$R_c = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)'(y_i - \hat{y}_i)}{\sum_{i=1}^m (y_i - \bar{y}1_{n_i})'(y_i - \bar{y}1_{n_i}) + \sum_{i=1}^m (\hat{y}_i - \hat{y}1_{n_i})'(\hat{y}_i - \hat{y}1_{n_i}) + N(\bar{y} - \hat{y})^2}, \tag{3}$$

como una medida de bondad de ajuste para modelos no lineales con efectos mixtos generalizados, donde N es el número total de observaciones; m es el número de individuos, y_i es el vector de valores observados del i-ésimo sujeto; $\hat{y}$ es el vector de valores predichos, $\bar{y}$ es la media de los elementos de todos los valores observados y 1_{n_i} es un vector $n_i \times 1$ de unos.

3.3 Estadístico $\hat{\Omega}^2$

Xu (2003) propuso el estadístico $\hat{\Omega}^2$ que sirve para explicar la variación en el modelo lineal. Está dado por:

$$\hat{\Omega}^2 = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}, \tag{4}$$

Este estadístico tiene como objetivo estimar $\Omega^2 = 1 - \frac{V(y_{ij}|X,b)}{V(y_{ij}\text{bajo un modelo nulo})}$, donde $j \in \{1, 2, \ldots, n_i\}$ y y_{ij} es el j-ésimo elemento de y_i, $\hat{\sigma}^2$ es el estimador de σ^2 del modelo ajustado (o modelo interés), $\hat{\sigma}_0^2$ es el estimador de la varianza de los residuales de un modelo nulo de la siguiente forma:

$$\mathbf{Y}_i = \mathbf{1}_{n_i}\boldsymbol{\beta}_o + \mathbf{1}_{n_i}\mathbf{b}_{oi} + \epsilon_i, \tag{5}$$

donde $\boldsymbol{\beta}_o$ es un parámetro fijo desconocido, $\boldsymbol{b}_{oi}$ es un coeficiente aleatorio desconocido que tiene distribución normal con media cero y ϵ_i es un error aleatorio intra-sujeto.

3.4 Estadístico R_2^2

El estadístico R_2^2 propuesto por Xu (2003), que al igual que $\hat{\Omega}^2$, mide la proporción de la variación explicada por las covariables en el modelo, se obtiene mediante la expresión:

$$R_2^2 = 1 - \frac{RSS}{RSS_0}, \tag{6}$$

donde RSS es la suma de los residuales al cuadrado del modelo (1) y RSS_0 es la suma de los residuales al cuadrado de un modelo nulo dado en (5).

3.5 Estadístico $\hat{\rho}^2$

Xu (2003) propuso el estadístico $\hat{\rho}^2$ como estimador de ρ^2 que mide la proporción de la aleatoriedad explicada de una variable aleatoria Y como una transformación monótona de su entropía, $\exp\left[-2I(\theta)\right]$, donde $I(\theta) = E[\log p(y;\theta)]$ es la esperanza de la log-verosimilitud bajo un modelo lineal mixto que tiene la forma dada en la ecuación (1), la aleatoriedad residual está definida como $D(y_{ij}|X,b)$=$\exp\left(-2E[\log p(y_{ij}|X,b)]\right)$. La proporción de la aleatoriedad explicada está dada por:

$$\rho^2 = 1 - \frac{D(y_{ij}|X,b)}{D\left(y_{ij}|b_o^*\right)}, \tag{7}$$

donde $D(y_{ij}|b_o^*)$ es la esperanza de la log-verosimilitud del modelo nulo dado en la ecuación (5).

$$\hat{\rho}^2 = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2} \exp\left(\frac{RSS}{\hat{\sigma}^2} - \frac{RRS_0}{\hat{\sigma}_0^2}\right), \tag{8}$$

donde RSS es la suma de los residuales al cuadrado del modelo (1) y RSS_0 es la suma de los residuales al cuadrado de un modelo nulo dado en (5).

3.6 Estadísticos R^2 marginales y condicionales

Vonesh et al. (1996) propusieron los conceptos de estadístico R^2 marginal y condicional. Para calcular la versión condicional del estadístico R^2, se utiliza el valor ajustado $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\mathbf{b}}$, no obstante, para la versión marginal del estadístico R^2, se utiliza el valor ajustado $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$. La versión condicional cuenta con ambos efectos fijos y aleatorios para medir la bondad de ajuste y el poder de predicción, mientras que la versión marginal sólo cuenta con la parte de efectos fijos, es decir, la media del modelo.

4 ESTADÍSTICOS R^2 PROPUESTOS: R^2_{DG} y R^2_{DGP}

En esta sección se proponen dos estadísticos R^2 para evaluar la bondad de ajuste de los modelos lineales mixtos, el R^2_{DG} y el R^2_{DGP}. El R^2_{DG} da igual peso a cada residual mientras que el R^2_{DGP} pondera los residuales de forma tal que da mayor peso a los residuales de los individuos con mayor cantidad de observaciones. Intuitivamente, lo anterior se fundamenta en el hecho de que el individuo con mayor cantidad de datos debe ser el que aporta más información al modelo. Ambos estadísticos se pueden utilizar para datos balanceados, es decir, cuando el número de observaciones es el mismo para cada individuo, o desbalanceados. Sin embargo, los estadísticos coinciden cuando los datos son balanceados. En este sentido el R^2_{DGP} es una generalización de R^2_{DG}. Para el cálculo de estos estadísticos los valores predichos se obtienen a partir de la siguiente relación:

$$\hat{\mathbf{Y}}_i = \mathbf{X}_i\hat{\beta} + \mathbf{Z}_i\hat{\mathbf{b}}_i. \tag{9}$$

Para modelos lineales con efectos mixtos ajustados a conjuntos de datos con medidas repetidas, se proponen los estadísticos R^2_{DGP} y R^2_{DG} dados por:

$$R^2_{DG} = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)'(y_i - \hat{y}_i)}{\sum_{i=1}^{m}(y_i - \bar{y}_i 1_{n_i})'(y_i - \bar{y}_i 1_{n_i})} \tag{10}$$

$$R^2_{DGP} = 1 - \frac{m}{N}\frac{\sum_{i=1}^{m} n_i(y_i - \hat{y}_i)'(y_i - \hat{y}_i)}{\sum_{i=1}^{m}(y_i - \bar{y}_i 1_{n_i})'(y_i - \bar{y}_i 1_{n_i})}, \tag{11}$$

donde m es el número de individuos; y_i, $\hat{y}_i$, $\bar{y}_i$ son los vectores de valores observados, valores ajustados o predichos y la media de los valores observados del individuo i; $i = 1, 2, \ldots, m$, respectivamente. N es el total de observaciones; n_i es el número de observaciones del individuo i, 1_{n_i} es un vector de unos de tamaño $n_i \times 1$.

4.1 Justificación

Los estadísticos R^2 buscan estimar el porcentaje de la variación que es explicada por el modelo. En general son de la forma,

$$1 - \frac{\text{Variabilidad no explicada por el modelo}}{\text{Variabilidad total de la variable de interés}}.$$

Los estadísticos definidos en las ecuaciones (9) y (10) estiman la variabilidad no explicada por el modelo en forma natural, mediante la suma de los cuadrados de los residuales de los modelos condicionales. Se debe tener presente que el modelo lineal con efectos mixtos ajusta una recta a cada individuo y por tanto a cada individuo le corresponde una cantidad de variabilidad que no es explicada por el modelo, así, la suma de dichas cantidades es la variabilidad no explicada por el modelo. Para estimar la variabilidad total se supone que las observaciones de cualquier par de individuos diferentes son independientes entre sí, un individuo no tiene relación con ningún otro.

Así, se estima la variabilidad total como la suma de las variabilidades totales de cada individuo, donde la variabilidad total de un individuo es la suma de cuadrados de la diferencia entre las observaciones de dicho individuo y su respectiva media. Parece razonable estimar la variabilidad como la suma de las variabilidades individuales bajo el supuesto de independencia entre los individuos.

4.2 Propiedades

Los estadísticos R^2, propuestos en las ecuaciones (9) y (10), poseen las siguientes propiedades.

1. Son adimensionales.
2. Fáciles de interpretar como la cantidad de variación explicada por el modelo ajustado.
3. Son menores o iguales a uno.
4. Alcanzan la cota superior, uno, cuando hay perfecto ajuste.
5. Un valor de cero o negativo indica completa falta de ajuste.
6. Si los datos son balanceados entonces $R^2_{DGP}=R^2_{DG}$.
7. El estadístico R^2_{DG} es menor o igual que el estadístico R^2_1.

5 ESTUDIO DE SIMULACIÓN

En esta sección se presenta la metodología utilizada para evaluar el desempeño de los estadísticos R^2 definidos en la sección 3 y 4, además se muestran los resultados de dicha simulación.

5.1 Metodología

Se genera un conjunto datos desbalanceados a partir de un modelo lineal mixto controlando los siguientes parámetros: tamaño de muestra, el vector de parámetros fijos del modelo, la distribución del efecto aleatorio y del error aleatorio y la covarianza de los efectos aleatorios. Se generó una variable respuesta longitudinal continua para 100 individuos y siete repeticiones para cada uno, con tres términos de efectos fijos: Una covariable de tiempo y dos variables dicotómicas. Las covariables de efecto aleatorio consisten de un término de intercepto aleatorio y un término lineal para el tiempo. Luego, de este conjunto, se extrajeron algunos valores en forma aleatoria manteniendo como mínimo una observación por individuo, de forma tal que el número de repeticiones por individuo no sea fijo, formando así el conjunto de datos desbalanceados.

Se simularon tres conjuntos de datos asumiendo que la covarianza de los efectos aleatorios no tiene una estructura determinada y la covarianza de los errores intrasujeto es $\mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i}$. Los tres conjuntos difieren en los valores de correlación intrasujeto, los respectivos valores de σ^2 fueron:

$14, 58, 220$, estos valores fueron seleccionados para que la correlación intrasujeto estuviera alrededor de 0.8, 0.5, 0.2, de tal forma que se pueda evaluar el desempeño de los estadísticos R^2 cuando el grado de asociación entre las observaciones de un mismo individuo es baja, media y alta. Los parámetros utilizados en la simulación se presentan a continuación:

$$\mathbf{X} = [\mathbf{1}_7, (9{,}9.5{,}9.75{,}10{,}10.5{,}10.75{,}11)', 1_7 I_1, 1_7 I_2];\ \mathbf{Z} = [\mathbf{1}_7, (9{,}9.5{,}9.75{,}10{,}10.5{,}10.75{,}11)']$$

donde $\mathbf{X}$ es una matriz de diseño que contiene las covariables asociadas a los efectos fijos, $\mathbf{Z}$ es la matriz de diseño que contiene las covariables para los efectos aleatorios y $\mathbf{1}_7$ es un vector de unos 7×1, $\mathbf{I}_1$ e $\mathbf{I}_2$ son funciones indicadoras que toman el valor 0 ó 1, donde 1 indica que el individuo posee una característica determinada y 0 lo contrario. $\mathbf{1}_7$ en $\mathbf{X}$ y $\mathbf{Z}$ es una columna de unos necesaria para incluir en el modelo el intercepto fijo y el intercepto aleatorio respectivamente, $\mathbf{1}_7\mathbf{I}_1$ y $\mathbf{1}_7\mathbf{I}_2$ corresponden a dos variables dicotómicas, que según sea el caso se tendrá, para $k = 1$ o $k = 2$, $\mathbf{1}_7\mathbf{I}_k = \mathbf{0}_7$ cuando el individuo no posee la característica, ó, $\mathbf{1}_7\mathbf{I}_k = \mathbf{1}_7$ cuando el individuo posee la característica. $\mathbf{1}_7\mathbf{I}_1$ y $\mathbf{1}_7\mathbf{I}_2$ se generaron en forma aleatoria para cada individuo. Se tomó

$$\beta = \begin{pmatrix} 16 \\ 9 \\ 17 \\ 17 \end{pmatrix} \text{ y } D = \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Los valores de los parámetros para dos de los efectos fijos (las variables dicotómicas) tienen el mismo valor (17) con el objetivo que tenga igual peso dentro de los modelos ajustados. Para la matriz de varianzas y covarianzas de los efectos aleatorios, los valores fueron seleccionados con el fin de que la correlación entre el intercepto y la pendiente fuera de 0.5.

Tabla 1: Descripción de los modelos ajustados

σ^2	**Modelo**	**Efectos Fijos**	**Efectos Aleatorios**
14	*Modelo Verdadero*	*Efectos Fijos*	*Intercepto + Pendiente*
	Modelo 1	*Efectos fijos +*	*Sólo intercepto*
	Modelo 2	*una covariable*	*Intercepto + Pendiente*
	Modelo 3	*independiente*	*Intecepto + Pendiente +*
			una variable independiente
58	*Modelo Verdadero*	*Efectos Fijos*	*Intercepto + Pendiente*
	Modelo 1	*Efectos Fijos +*	*Sólo intercepto*
	Modelo 2	*una covariable*	*Intercepto + Pendiente*
	Modelo 3	*independiente*	*Intecepto + Pendiente +*
			una variable independiente
220	*Modelo Verdadero*	*Efectos Fijos*	*Intercepto + Pendiente*
	Modelo 1	*Efectos Fijos +*	*Sólo intercepto*
	Modelo 2	*una covariable*	*Intecepto + Pendiente*
	Modelo 3	*independiente*	*Intecepto + Pendiente +*
			una variable independiente

Tabla 2: Valores Ω^2 para el modelo verdadero

β	var(Tiempo)	σ^2	γ	Ω^2
		14	0.8	0.75
9	0.51	58	0.5	0.42
		220	0.2	0.16

Se probaron tres escenarios de simulación, uno para cada valor de σ^2. Para cada conjunto de datos simulados, se ajustó el modelo verdadero que consiste del tiempo y las dos variables dicotómicas como efectos fijos y del intercepto y el tiempo como efectos aleatorios. Aparte del modelo verdadero se ajustaron tres diferentes modelos. En cada uno de ellos se tomaron los mismos efectos fijos: El tiempo, las dos variables dicotómicas y una covariable no relacionada con la variable respuesta (variable generada a partir de una distribución uniforme en forma independiente de la variable respuesta); en tanto que los efectos aleatorios varían de modelo a modelo. En el primer modelo se ajustaron los efectos fijos ya mencionados y un intercepto aleatorio solamente; en el segundo modelo se agregó el tiempo como efecto aleatorio con respecto al primer modelo; y por último, en el tercer modelo se agregó la covariable no relacionada con la variable respuesta como efecto aleatorio con respecto al segundo modelo. Los modelos que se ajustaron para cada valor de σ^2 se muestran en la tabla 1.

En comparación con el modelo verdadero, el modelo 1 es un modelo reducido en donde se elimina el tiempo en los efectos aleatorios, similarmente, en comparación con el modelo verdadero, el modelo 2 y el modelo 3 son modelos sobre ajustados. En el modelo 2 se adiciona una variable no relacionada con la respuesta como efecto fijo y en el modelo 3 se adiciona, tanto para los efectos fijos como para los efectos aleatorios, una variable no relacionada con la respuesta. Para ajustar cada uno de los modelos se utilizó el sofware estadístico R (2009) versión 2.9.2. Se simularon 1000 muestras dentro de cada conjunto de datos (cada valor de σ^2). Para cada muestra se ajustaron los diferentes modelos usando verosimilitud restringida para estimar los parámetros. Para cada modelo se calcularon los estadísticos R^2 descritos en las secciones 3 y 4 en su versión condicional.

Con el propósito de estimar la previsibilidad de las covariables en el modelo se utilizó la fórmula propuesta por Xu (2003) para calcular Ω^2 de la siguiente manera:

$$\Omega^2 = 1 - \frac{\sigma^2}{(\beta^2_{tiempo} + \gamma^2)var(tiempo) + \sigma^2}, \tag{12}$$

donde γ es el coeficiente de correlación intrasujeto y *var(tiempo)* es la varianza de la variable tiempo. A partir de los valores dados anteriormente se obtienen los siguientes valores para Ω^2 para el modelo verdadero:

5.2 Resultados

A partir de los modelos ajustados se obtuvieron estimaciones de los siete estadísticos R^2 definidos en las secciones 3 y 4. La tabla 3 muestra la media, el mínimo y el máximo de los valores obtenidos de los estadísticos R^2 ajustados para cada valor de σ^2.

Los estadísticos propuestos por Xu (2003) y los estadísticos propuestos en este trabajo están cercanos al valor de Ω^2, el porcentaje de previsibilidad de las covariables definido en la ecuación (11). Tal como se mencionó anteriormente los estadísticos R^2_{DG} y R^2_{DGP} no son iguales, además se observa que los estadísticos R^2 condicionales, son invariantes o presentan leves cambios al evaluarlos entre los diferentes modelos cuando se modifican los efectos aleatorios y por tanto, los estadísticos R^2 no permiten elegir el verdadero modelo y conducen a la elección de un modelo mal especificado.

Las simulaciones también muestran que los estadísticos R^2 condicionales analizados no son adecuados como medidas de bondad de ajuste cuando se trabaja con modelos lineales mixtos, esto se debe principalmente a que los estadísticos R^2 en los diferentes modelos ajustados no presentan mayores variaciones en relación al modelo verdadero y en consecuencia no logran identificar cuando las covariables son importantes. Lo anterior porque se supone que una propiedad deseada en un R^2 es que pueda disminuir su valor cuando se eliminan una o más covariables importantes para el modelo y la disminución en el valor debe ser proporcional a la cantidad de la variación de la variable respuesta que es explicada por la variable eliminada. Por otra parte, al adicionar una covariable independiente (no relacionada) de la variable respuesta se espera que el valor del estadístico R^2 no aumente.

Cabe señalar que a medida que la variabilidad intraindividual (σ^2) aumenta, los valores de los estadísticos R^2 disminuyen. El hecho de que los estadísticos R^2 estudiados no logren discriminar entre un modelo verdadero y uno mal especificado se debe básicamente, a que en el ajuste de un efecto aleatorio se incluyen muchos parámetros en el modelo, uno por cada individuo, por tanto es difícil medir su efecto. Teniendo en cuenta que la correlación intrasujeto dificulta esta labor, al existir correlación entre los efectos aleatorios. Por tanto, al incluir el intercepto aleatorio en un modelo, este contiene información de los otros efectos aleatorios importantes para el modelo, así la inclusión de un segundo efecto aleatorio no es tan significativa como se pudiera esperar, su efecto se ve reducido y es más difícil determinar si el segundo efecto aleatorio es importante para el modelo. Similarmente, la inclusión de los dos primeros efectos aleatorios contiene información de los demás efectos aleatorios importantes para el modelo, lo que reduce aún más el efecto de la inclusión del próximo efecto aleatorio; y así ocurre sucesivamente. Entonces el ajuste de los primeros efectos aleatorios puede incrementar tanto el valor del R^2 que al incluir los demás no se alcanza a notar un cambio significativo en el R^2 lo que conduce a descartar efectos aleatorios importantes para el modelo.

Tabla 3: Medias, máximos y mínimos de los estadísticos R^2 estudiados para datos desbalanceados

σ^2	Modelo	Ω^2	R_1^2	R_c	$\hat{\Omega}^2$	R_2^2	$\hat{\rho}^2$	R_{DG}^2	R_{DGP}^2
14	***Modelo Verdadero.*** Efectos fijos: Tiempo y dos variables dicotómicas. Efectos aleatorios: Intercepto y el tiempo como Pendiente aleatoria.	0.75	0.95(0.93,0.97)	0.98(0.97,0.98)	0.75(0.7,0.8)	0.75(0.7,0.81)	0.75(0.7,0.81)	0.75(0.7,0.81)	0.75(0.69,0.8)
	Modelo 1. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto.		0.95(0.93,0.97)	0.98(0.96,0.98)	0.75(0.7,0.79)	0.75(0.7,0.79)	0.75(0.7,0.79)	0.75(0.7,0.79)	0.74(0.69,0.78)
	Modelo 2. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto y el tiempo como Pendiente.		0.95(0.93,0.97)	0.98(0.97,0.98)	0.75(0.7,0.79)	0.75(0.7,0.79)	0.75(0.7,0.79)	0.75(0.7,0.79)	0.75(0.69,0.79)
	Modelo 3. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto, el tiempo como Pendiente aleatoria más una covariable independiente.		0.95(0.93,0.97)	0.98(0.97,0.99)	0.75(0.7,0.8)	0.76(0.7,0.81)	0.76(0.7,0.81)	0.76(0.7,0.81)	0.75(0.69,0.81)
58	***Modelo Verdadero.*** Efectos fijos: Tiempo y dos variables dicotómicas. Efectos aleatorios: Intercepto y el tiempo como Pendiente aleatoria.	0.42	0.83(0.77,0.88)	0.91(0.87,0.94)	0.42(0.32,0.52)	0.42(0.32,0.53)	0.42(0.32,0.53)	0.42(0.31,0.53)	0.41(0.3,0.52)
	Modelo 1. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto		0.83(0.76,0.88)	0.91(0.86,0.94)	0.42(0.32,0.51)	0.41(0.31,0.51)	0.41(0.31,0.51)	0.41(0.31,0.51)	0.4(0.3,0.5)
	Modelo 2. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto y el tiempo como Pendiente aleatoria.		0.83(0.77,0.88)	0.91(0.87,0.94)	0.42(0.32,0.52)	0.42(0.32,0.52)	0.42(0.32,0.52)	0.42(0.31,0.51)	0.4(0.3,0.51)
	Modelo 3. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto, el tiempo como Pendiente aleatoria más una covariable independiente.		0.84(0.78,0.89)	0.91(0.88,0.94)	0.43(0.32,0.54)	0.44(0.32,0.56)	0.44(0.32,0.56)	0.43(0.31,0.55)	0.42(0.3,0.55)
220	***Modelo Verdadero.*** Efectos fijos: Tiempo y dos variables dicotómicas. Efectos aleatorios: Intercepto y el tiempo como Pendiente aleatoria.	0.16	0.57(0.45,0.69)	0.73(0.62,0.82)	0.17(0.09,0.27)	0.15(0.06,0.27)	0.15(0.07,0.27)	0.13(0.05,0.26)	0.12(0.02,0.24)
	Modelo 1. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto.		0.57(0.46,0.66)	0.72(0.63,0.8)	0.16(0.08,0.24)	0.14(0.05,0.22)	0.14(0.05,0.23)	0.12(0.03,0.2)	0.1(0,0.18)
	Modelo 2. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto y el tiempo como Pendiente aleatoria.		0.57(0.46,0.66)	0.73(0.63,0.8)	0.17(0.09,0.25)	0.15(0.06,0.26)	0.15(0.06,0.26)	0.13(0.04,0.24)	0.11(0.01,0.21)
	Modelo 3. Efectos fijos: Tiempo, dos variables dicotómicas y una covariable independiente Efectos aleatorios: Intercepto , el tiempo como Pendiente aleatoria más una covariable independiente.		0.59(0.47,0.69)	0.74(0.64,0.82)	0.18(0.1,0.3)	0.17(0.07,0.32)	0.17(0.08,0.32)	0.16(0.05,0.31)	0.14(0.02,0.3)

Los estadísticos R^2 fueron calculados usando en el numerador $\hat{\mathbf{Y}}_i = X_i\hat{\beta} + Z_i\hat{b}_i$

Los modelos son especificados en la tabla 1.

Tabla 4: Estadísticos R^2 condicionales evaluados en los datos desbalanceados de crecimiento Potthoff y Roy (1964)

modelo	efectos fijos	efectos aleatorios	R_1^2	R_c	$\hat{\Omega}^2$	R_2^2	$\hat{\rho}^2$	R_{DG}^2	R_{DGP}^2
1	Edad + género +edad*género	Intercepto+ edad	0.8681996	0.9294506	0.6806271	0.7403232	0.7253565	0.7202088	0.6947732
2	Edad + género +edad*género	Intercepto	0.8227952	0.902784	0.6211161	0.6508662	0.6443861	0.6238224	0.5896244
3	Edad + género	Intercepto + Edad	0.8694504	0.9301668	0.6770288	0.7427875	0.7259786	0.7228639	0.6976697
4	Edad + género	Intercepto	0.8087004	0.8942337	0.5976335	0.6230962	0.6176716	0.5939013	0.5569833
5	Edad	Intercepto + Edad	0.868738	0.929759	0.6770043	0.741384	0.7250083	0.7213517	0.69602
6	Edad	Intercepto	0.810769	0.8954969	0.5975584	0.6271718	0.6207672	0.5982927	0.5617738
7	Intercepto	Intercepto + Edad	0.8924633	0.9431763	0.6776356	0.7881281	0.7555531	0.7717166	0.7509636

Los estadísticos R^2 fueron calculados usando en el numerador $\hat{\mathbf{Y}}_i = X_i\hat{\beta} + Z_i\hat{b}_i$.

6 Aplicación: Datos de crecimiento dental

Para aplicar los estadísticos R^2 a datos reales se utilizan datos de crecimiento dental que fueron introducidos por Potthoff y Roy (1985); estos contiene medidas del crecimiento de 27 niños: 16 hombres y 11 mujeres. A cada niño con rayos X se le registró la distancia desde el centro de la pituitaria hasta la fisura del maxilar en las edades de 8, 10 ,11 y 14 años. La base de datos contiene 27 individuos (niños) con cuatro repeticiones cada uno, para un total de 108 observaciones. Los estadísticos R^2 se calcularon con los siguientes modelos:

1. Modelo 1. Consiste de las covariables *edad*, *género* y la interacción *edad* $\times$ *género* como efectos fijos, además posee un intercepto aleatorio y la covariable *edad* como efecto aleatorio.

2. Modelo 2. Los efectos fijos están dados por las covariables *edad*, *género* y la interacción *edad* $\times$ *género*, en tanto que como efecto aleatorio se ajusta un intercepto aleatorio.

3. Modelo 3. Se ajustan efectos fijos para la *edad* y el *género*, y como efectos aleatorios el intercepto y la pendiente medida a través de la *edad*.

4. Modelo 4. Se ajustan efectos fijos para la *edad* y el *género* junto con un intercepto aleatorio.

5. Modelo 5. Este modelo explica la distancia desde el centro de la pituitaria hasta la fisura del maxilar a través de la covariable *edad* como efecto fijo, un intercepto aleatorio y la *edad* como efecto aleatorio.

6. Modelo 6. Este modelo explica la distancia desde el centro de la pituitaria hasta la fisura del maxilar a través de la covariable *edad* como efecto fijo y un intercepto aleatorio.

7. Modelo 7. Posee solamente un efecto fijo, el intercepto, un intercepto aleatorio y la covariable *edad* como efecto aleatorio.

En la tabla 4 se muestran los resultados de los estadísticos R^2 condicionales. Todos los estadísticos R^2 presentaron diferencias pequeñas al ser calculados en los diferentes modelos desde el modelo uno hasta el modelo siete. Los estadísticos R_1^2 y R_c toman valores más altos que los otros estadísticos. Los resultados obtenidos son consistentes con el estudio de simulación.

7 CONCLUSIONES

Cuando se evalúa el desempeño de los estadísticos R^2 como medidas de bondad de ajuste y como criterios de selección del mejor modelo que explique adecuadamente la variabilidad presente se obtuvieron los siguientes resultados:

A través de las simulaciones mostradas en la sección 5, queda claro que el desempeño de los estadísticos condicionales propuestos por Xu (2003), los propuestos por Vonesh et al. (1996) y Vonesh y Chinchilli (1997) y los propuestos en este trabajo no presentaron cambios en los valores medios al ser evaluados en un modelo subajustado o sobreajustado en relación al modelo verdadero cuando se adicionan o se quitan covariables de efectos aleatorios. Lo que pone en tela de juicio su utilidad como criterios de selección y medidas de bondad de ajuste en modelos lineales mixtos con respuesta continua donde están mal especificadas las covariables de efecto aleatorio en el contexto de los datos longitudinales.

Orelien y Edwards (2007) obtuvieron resultados similares. Ellos evalúan la utilidad de estos estadísticos asumiendo que los efectos aleatorios están bien especificados y ajustan modelos adicionando y quitando covariables de efectos fijos.

Si bien todos los estadísticos R^2 no cambian de modelo a modelo en relación al verdadero, condición no deseable, se resalta el hecho que los estadísticos propuestos, en este trabajo, poseen las propiedades principales que fueron descritas por Kvalseth (1985) para medidas tipo R^2, es decir, poseen buenas propiedades estadísticas. Por ejemplo: son fáciles de interpretar, intuitivamente dan el porcentaje de la variación explicada por el modelo; tienen por cota superior el uno que se alcanza cuando hay perfecto ajuste, si toman valores negativos o cero se tiene total falta de ajuste.

Los estadísticos propuestos por Xu (2003) y los estadísticos R_{DG}^2 y R_{DGP}^2 son cercanos al verdadero valor de Ω^2, el coeficiente de previsibilidad, que mide la proporción de la variación explicada por las covariables en un modelo. Es decir, estos estadísticos son buenos estimadores de la proporción de la variabilidad de la variable respuesta explicada por el modelo.

Referencias

Kvalseth, T. O. (1985), Cautionary note about R^2. *American Statistical Association*, **39**, 279-285.

Laird, N. M.; Ware, J. H. (1982), Random - effects models for longitudinal data. *Biometrics*, **38**, 963-974.

Little, R. J. A.; Rubin, B. D. (1987), *Statistical Analysis with Missing Data. John Wiley* & *Sons, New York.*

Liu, H.; Zheng, Y.; Shen, Jie. (2008), Goodness-of-fit measures of R^2 for repeated measures mixed effect models. *Journal of Applied Statistics*, **35**, 1081-1092.

Orelien, J. G.; Edwards L. J. (2007), Fixed-effect variable selection in linear mixed models using R^2 statistics. *Computational Statistics* & *Data Analysis*, **52**, 1896-1907.

Potthoff, R. F.; Roy, S. N. (1964), A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **85**, 313-326.

R Development Core Team (2009), *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0

Vonesh, E. F.; Chinchilli, V. M.; Pu, K. (1996), Goodness-of-fit in generalized nonlinear mixed-effects models. *Biometrics*, **52**, 572-587.

Vonesh, E. F.; Chinchilli, V. M. (1997), Goodness-of-fit in generalized nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York, 419-424.

Xu, Ronghui. (2003), Measuring explained variation in linear mixed effects models. *Statistics In Medicine*, **22**, 3527-3541.