

EXTENSIÓN DEL MODELO SUR PARA DATOS LONGITUDINALES

Giaquinta, Sofia¹; Cuesta, Cristina¹; Rapelli, Cecilia¹.

¹Instituto de Investigaciones Teóricas y Aplicadas en Estadística, FCEyE, UNR

sofiagiaquinta01@gmail.com; cbcuesta@gmail.com; cecirapelli@gmail.com

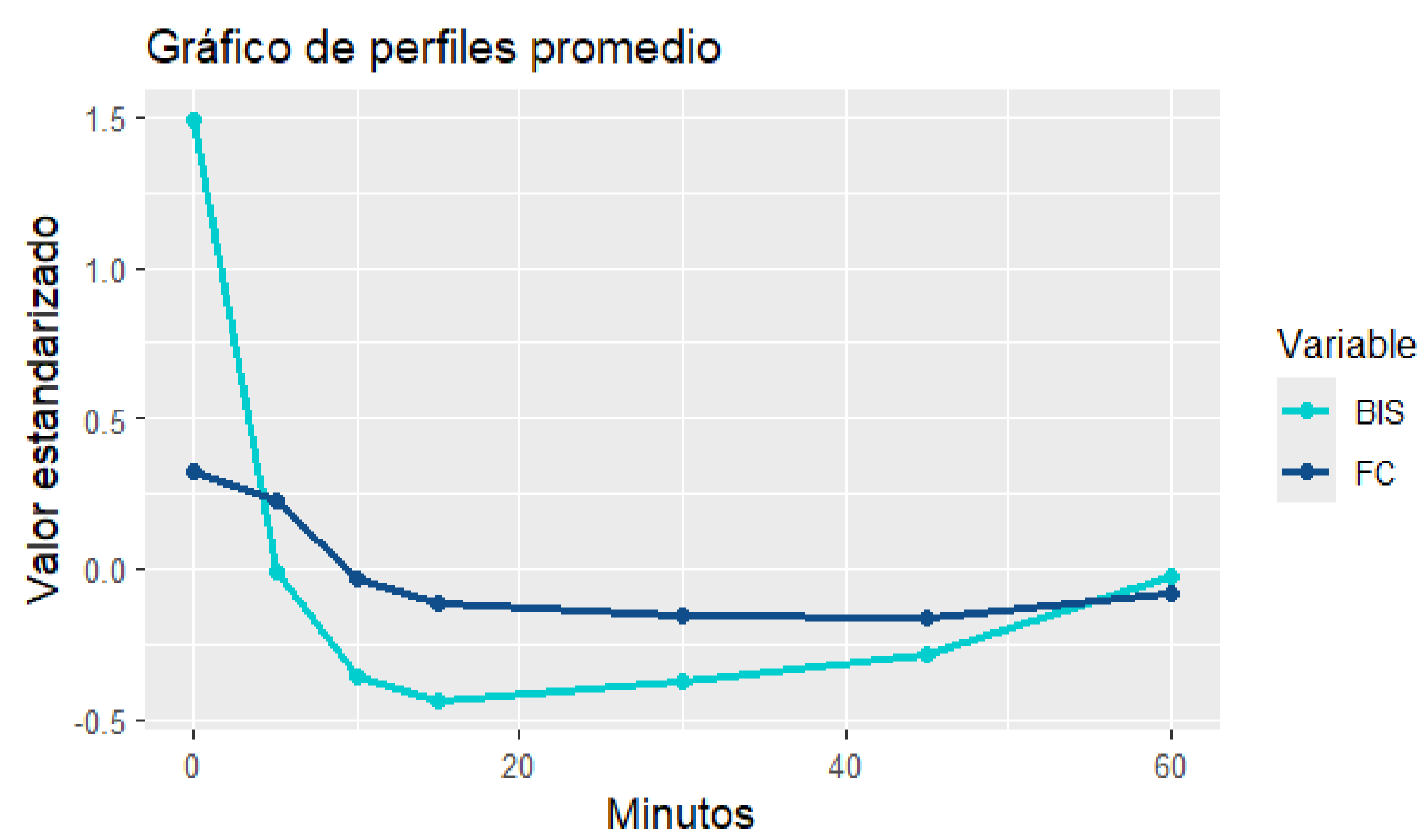
Palabras claves: Modelos SUR, Respuesta multivariada, Modelos longitudinales.

INTRODUCCIÓN

En diversas áreas de investigación es usual que, para cada unidad estudiada, se registren varias variables de interés en más de una ocasión sucesiva. Esto conlleva dos desafíos para su análisis, por un lado tener en cuenta la respuesta simultánea de varias variables y por otro considerar la posible correlación entre observaciones de una misma unidad a través del tiempo. Los métodos estadísticos para modelar estas situaciones son escasos. Entre las posibilidades se consideran los modelos mixtos y los modelos lineales para medidas repetidas. Ambas opciones requieren de una especificación a priori de las matrices de variancias y covariancias de las observaciones y su estimación puede resultar dificultosa. En este trabajo se presenta la adaptación de otro tipo de modelos denominados SUR (Seemingly Unrelated Regressions), originalmente desarrollados para el abordaje de respuestas múltiples en una única ocasión, al caso de datos longitudinales.

DATOS

Se piensa que la Frecuencia Cardíaca e Índice Biespectral de una persona luego de dormirse tiene una alta correlación. Se le miden estas dos variables a 57 pacientes antes de dormirse y después de 5, 10, 15, 30, 45 y 60 minutos de dormirse.



METODOLOGÍA

$$y_{ij}^1 = \beta_0^1 + \beta_1^1 \text{Tiempo}_{ij} + \beta_2^1 (\text{Tiempo}_{ij} - 15)_+ + \beta_3^1 \text{Edad} + \varepsilon_{ij}^1$$

$$y_{ij}^2 = \beta_0^2 + \beta_1^2 \text{Tiempo}_{ij} + \beta_2^2 (\text{Tiempo}_{ij} - 10)_+ + \varepsilon_{ij}^2$$

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \begin{bmatrix} X^1 & \mathbf{0} \\ \mathbf{0} & X^2 \end{bmatrix} \begin{bmatrix} \beta^1 \\ \beta^2 \end{bmatrix} + \begin{bmatrix} \varepsilon^1 \\ \varepsilon^2 \end{bmatrix}$$

Ecuación 1= Frecuencia cardiaca
Ecuación 2= Índice Biespectral

$$E(\varepsilon\varepsilon' | X^1, X^2) = \Omega = \Sigma \otimes (I_N \otimes T)$$

$$\Sigma = \begin{bmatrix} \sigma_{11}^s & \sigma_{12}^s \\ \sigma_{21}^s & \sigma_{22}^s \end{bmatrix} \quad T = \begin{bmatrix} \sigma_{11}^t & \sigma_{12}^t & \sigma_{13}^t & \sigma_{14}^t & \sigma_{15}^t & \sigma_{16}^t & \sigma_{17}^t \\ \sigma_{21}^t & \sigma_{22}^t & \sigma_{23}^t & \sigma_{24}^t & \sigma_{25}^t & \sigma_{26}^t & \sigma_{27}^t \\ \sigma_{31}^t & \sigma_{32}^t & \sigma_{33}^t & \sigma_{34}^t & \sigma_{35}^t & \sigma_{36}^t & \sigma_{37}^t \\ \sigma_{41}^t & \sigma_{42}^t & \sigma_{43}^t & \sigma_{44}^t & \sigma_{45}^t & \sigma_{46}^t & \sigma_{47}^t \\ \sigma_{51}^t & \sigma_{52}^t & \sigma_{53}^t & \sigma_{54}^t & \sigma_{55}^t & \sigma_{56}^t & \sigma_{57}^t \\ \sigma_{61}^t & \sigma_{62}^t & \sigma_{63}^t & \sigma_{64}^t & \sigma_{65}^t & \sigma_{66}^t & \sigma_{67}^t \\ \sigma_{71}^t & \sigma_{72}^t & \sigma_{73}^t & \sigma_{74}^t & \sigma_{75}^t & \sigma_{76}^t & \sigma_{77}^t \end{bmatrix}$$

ESTIMACION

Primera etapa:

$$\hat{\beta}_{MCO}^1 = (X^1 X^1)^{-1} X^1 y^1$$

$$\hat{\beta}_{MCO}^2 = (X^2 X^2)^{-1} X^2 y^2$$

$$\hat{\Omega} = \hat{\Sigma} \otimes (I_N \otimes \hat{T}) \longrightarrow$$

Segunda etapa:

$$\hat{\beta}_{MCGF} = (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} y$$

ESTRATEGIAS DE ESTIMACION DE Σ Y T

SUR residuos promedios:

$\hat{\Sigma}$ = Matriz de covariancia entre los residuos promedio de cada individuo de la ecuación 1 y 2

\hat{T} = Promedio de la matrices de covariancia en el tiempo de las ecuaciones 1 y 2

SUR residuos individuales:

$\hat{\Sigma}$ = Matriz de covariancia entre los residuos de cada individuo en cada tiempo de la ecuación 1 y 2

\hat{T} = Matriz de covariancia en el tiempo de las observaciones de los 57 individuos en ambas ecuaciones

RESULTADOS

Se estimaron los modelos SUR, con las dos opciones de $\hat{\Omega}$ y los modelos individuales para las variables Frecuencia cardiaca e Índice Biespectral y se compararon sus parámetros estimados y sus sumas de cuadrados del error. Para esto se utilizo el software R.

Parámetros	SUR Residuos promedio		SUR Residuos acumulados		Modelos individuales	
	Estimación	EE	Estimación	EE	Estimación	EE
β_0^1	0.967**	0.194	0.950**	0.231	0.845**	0.309
β_1^1	-0.033**	0.006	-0.029**	0.007	-0.023**	0.008
β_2^1	0.035**	0.007	0.030**	0.008	0.024**	0.009
β_3^1 (edad)	-0.016**	0.004	-0.016**	0.005	-0.015*	0.007
β_0^2	1.406**	0.045	1.374**	0.066	1.518**	0.023
β_1^2	-0.149**	0.006	-0.159**	0.008	-0.142**	0.008
β_2^2	0.152**	0.006	0.164**	0.009	0.140**	0.009
SCE	FC	BIS	FC	BIS	FC	BIS
	380.89	290.41	380.66	272.78	380.64	320.7

CONCLUSIONES

Todos los parámetros resultan significativos
Estimaciones más precisas con los modelos SUR que con los modelos individuales
Mayor ganancia en el ajuste para la variable BIS