



XLVIII Coloquio Argentino de Estadística

VI JORNADA DE EDUCACIÓN ESTADÍSTICA "MARTHA DE ALIAGA"

27 al 30 oct 2020

Poster:

Uso de métodos diagnósticos en modelos mixtos

Cecilia Rapelli, María del Carmen García



Esta obra está bajo una
Licencia Creative Commons
Atribución-NoComercial 4.0
Internacional



FACULTAD
DE CIENCIAS
ECONÓMICAS



Universidad
Nacional
de Córdoba



USO DE MÉTODOS DIAGNÓSTICOS EN MODELOS MIXTOS

CECILIA RAPELLI, MARÍA DEL CARMEN GARCIA

Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística. Facultad de Ciencias Económicas y Estadística. Universidad Nacional de Rosario

Resumen

Los modelos lineales mixtos constituyen una poderosa herramienta inferencial que se utiliza para el análisis de datos longitudinales. La complejidad de estos modelos, por la presencia de efectos aleatorios, hace que, para la elección de un modelo adecuado para los datos, se requiera el uso de herramientas de diagnóstico para evaluar la existencia de alguna deficiencia en el modelo estimado. El propósito de este trabajo es presentar y comparar diferentes enfoques para identificar valores y unidades atípicos y/o influyentes y elegir un modelo "óptimo". Se ilustra su uso para modelar el nivel de ácido láctico en pacientes con HIV. La aplicación sugiere que los tres métodos presentados se complementan.

Los datos

Se realizó un estudio para investigar el efecto de un tratamiento antirretroviral sobre los valores de ácido láctico en sangre. Se incluyeron 45 pacientes con serología positiva para el HIV. De éstos, 30 comenzaron un tratamiento antirretroviral 3 meses antes del comienzo del estudio y 15 no realizan tratamiento antirretroviral. Se registraron los valores de ácido láctico en la primera consulta (visita 1) y luego en los controles, a los 3, 6 y 9 meses de la visita inicial (visitas 2, 3 y 4).

Distancia de Cook clásica (DC)

Es la estadística más utilizada. Su eficacia es limitada por estar orientada a la unidad. Puede no ser conveniente para detectar unidades influyentes en vista de la posición relativa de las observaciones dentro y entre los sujetos.

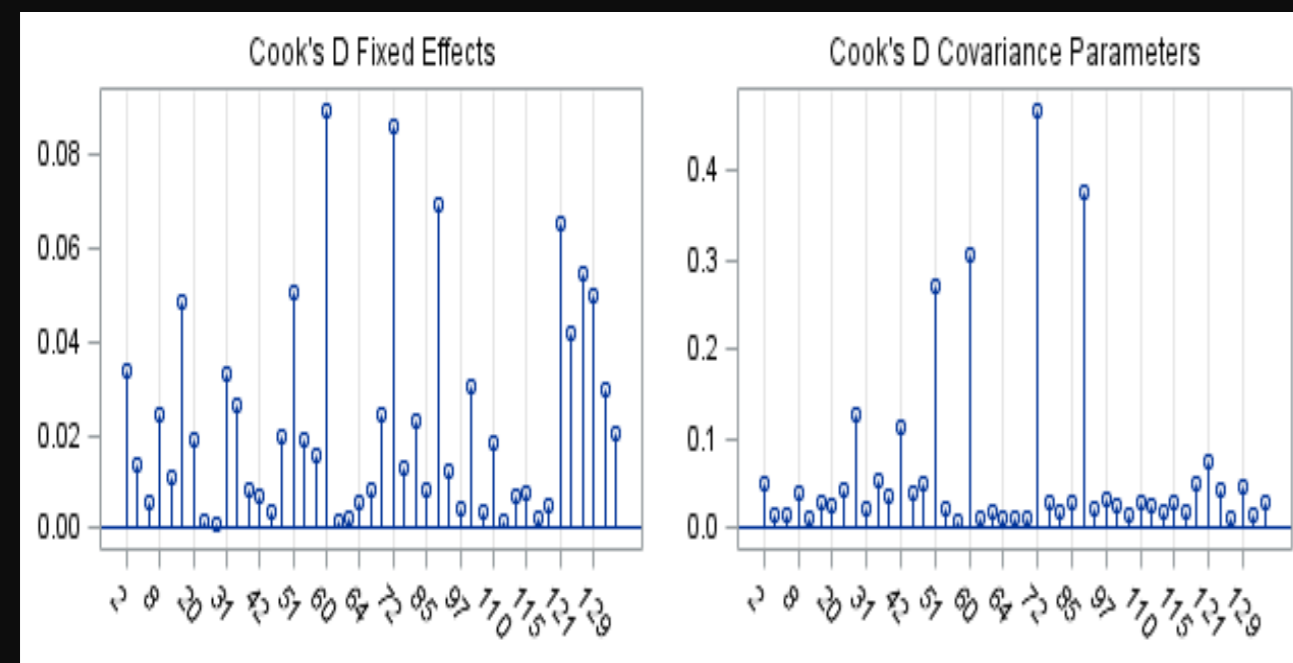


Gráfico TRSS

El gráfico de las sumas de cuadrados de los residuos estudentizados (TRSS) mide simultáneamente la desviación entre la media específica de una unidad y la media poblacional (tipo L) y la distancia entre una trayectoria individual y su media específica (tipo S).

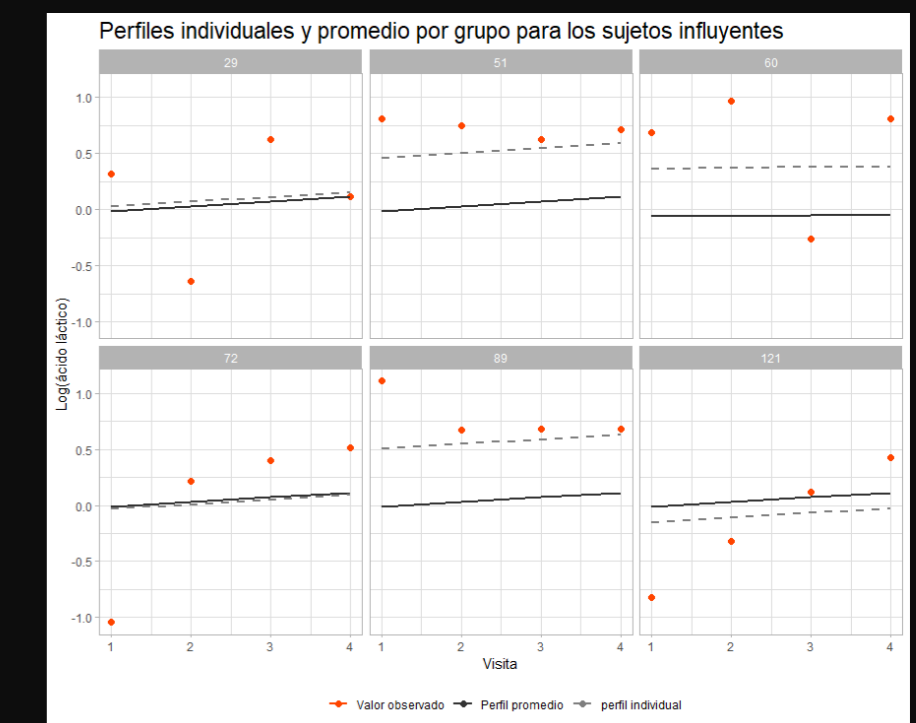
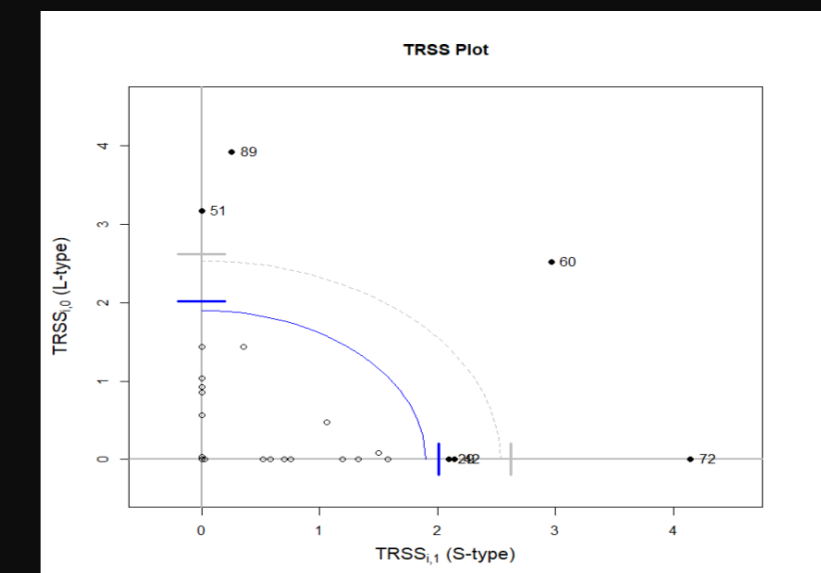


Tabla 1. Individuos influyentes detectados según herramienta

Unidad	Distancia Cook (DC)	Distancia Cook condicional (DCc)	Gráfico TRSS
51	COV		TIPO L
60	EF COV		TIPO L y S
72	EF COV	DC1 en1 DC2 en1	TIPO S
89	EF COV		TIPO L
29		DC2 en2	TIPO S
42			TIPO S
121		DC2 en1	

Consideraciones finales

- El perfil observado del paciente 29 se desvía de la trayectoria lineal de los restantes y es detectado por el gráfico TRSS. Como el perfil individual estimado es muy cercano al promedio, la DC no lo detecta.
- El individuo 72 es detectado por todas las estadísticas. La DC la detecta como la más influyente y el gráfico TRSS la señala como tipo S lo que indica que su estructura de correlación es diferente a las demás.
- Los pacientes 51 y 89 son tipo L en el gráfico TRSS causado por el gran efecto para predecir su ordenada al origen. La ausencia de estos pacientes produciría serios cambios en la estimación de los parámetros fijos y de covariancia.
- El paciente 121 es sólo detectado por la DCc, que considera el valor en la visita 1 como discordante. La DC no la detecta porque no es todo el perfil influyente sino solo una obs.
- El gráfico TRSS y la DC detectaron al paciente 60 como influyente, tanto sobre los parámetros fijos como los de covariancia. La DCc no lo detecta puesto que el perfil completo es discordante y no una observación en particular.

Distancia de Cook condicional (DCc)

Basada en la observación. Se descompone en tres medidas $Dc_{ij} = DC1_{ij} + DC2_{ij} + DC3_{ij}$

- Influencia sobre los efectos fijos
- Cambios en los efectos específicos de la unidad
- Covariación entre un cambio en el perfil promedio y un cambio en el perfil individual relativo al perfil medio

