Revista Colombiana de Estadística
No.29 y No. 30

MEDIDAS DE NO LINEALIDAD EN REGRESION NO LINEAL

HUGO GRISALES ROMERO

Facultad Nacional de Salud Pública
Universidad de Antioquia

RESUMEN Usualmente la extensión de las inferencias de los modelos de regresión lineal al caso no lineal se han hecho obviando el grado de no linealidad y la importancia de la bondad de ajuste en muestras pequeñas. En este artículo se analizan las medidas de Bates y Watts (1980) y Hougaard (1985) para la validación de la extensión inferencial confrontando esta última con la de Box (1971). Se desarrolló un software (RENOL: Regresión No Lineal) para el hallazgo de estas medidas y con su uso se evaluaron dos conjuntos de datos utilizados en sendas investigaciones exponiéndose las conclusiones pertinentes.

1. INTRODUCCIÓN

La utilización de los modelos lineales como un patrón de medida del comportamiento no lineal se suele considerar natural ya que el comportamiento de los modelos no lineales en estimación con muestras grandes se aproxima al de un modelo lineal.

Para un modelo de Regresión Lineal los estimadores de los parámetros son insesgados, normalmente distribuidos y tienen la varianza mínima posible en la clase de estimadores conocidos como Uniformemente Insesgados de Mínima Varianza, siempre y cuando los errores sean independientes e idénticamente distribuidos de manera normal con media cero y varianza σ^2. En regresión no lineal, en muestras pequeñas, los estimadores mínimo-cuadráticos de los parámetros no son insesgados, no son normalmente distribuidos y no tienen mínima varianza aunque los errores sean $N(0, \sigma^2)$. Sobre esto último Ratkowsky (1983) muestra una serie de ejemplos ilustrativos.

En el presente artículo se presentan medidas que permiten determinar analíticamente, en muestras pequeñas, si las aproximaciones asintóticas de los modelos no lineales de regresión son válidas para fines inferenciales. Yáñez (1988), enfatizó la importancia de estos problemas utilizando la medida del sesgo de Box (1971) para los datos de consumo de energía eléctrica en dos ciudades colombianas. Aquí se incluirán las medidas de Bates y Watts (1980) y de Hougaard para cuantificar la no linealidad, las cuales permitan mejorar el análisis ya que el sesgo de Box (1971), como medida única para medir no linealidad, presenta problemas en su uso; con ello se espera contribuir a la enseñanza de la regresión no lineal, alertando sobre sus dificultades y mostrando sus recientes desarrollos. Para tal efecto se implementó un programa en Fortran que calcula además del sesgo de Box, otras medidas de singular importancia tales como las medidas de curvatura de Bates & Watts y la medida de asimetría de Hougaard. El programa mencionado que se implementó, con el nombre de RENOL (Regresión no lineal), se ensayó con 36 modelos, entre ellos los sugeridos por Bates & Watts (1988).

El programa **RENOL** está disponible para quien lo solicite y se espera que sea de utilidad a la comunidad estadística nacional.

2. Medidas De Curvatura De Bates & Watts

El desarrollo formal de las medidas anteriormente citadas, se precisará a continuación siguiendo los lineamientos teóricos de Seber (1988).

Sea $y = \eta + \epsilon$ donde $\eta \in \Omega$, siendo Ω un subconjunto del espacio muestral, el cual usualmente es $\mathbb{R}^n$. Ω puede ser descrito en términos de un parámetro $\psi \in \Phi_p$, donde Φ_p es un conjunto parametrizable localmente por p coordenadas, y una función modelo $f(\psi)$; así $\Omega = \{\eta : \eta = f(\psi), \psi \in \Phi_p\}$ donde $[f(\psi)]_i = f_i(\psi) = f(x_i, \psi)$, x_i es la fila i-ésima de la matriz de diseño X.

Suponga que ψ está suficientemente cerca de $\hat{\psi}$; entonces tenemos la siguiente aproximación cuadrática de Taylor:

$$\begin{aligned} \eta - \hat{\eta} &= f(\psi) - f(\hat{\psi}) \\ &\approx \hat{F}.(\psi - \hat{\psi}) + 1/2(\psi - \hat{\psi})^t \hat{F}..(\psi - \hat{\psi}) \\ &= \hat{F}.\delta + \frac{1}{2}\delta^t \hat{F}..\delta \end{aligned} \tag{2.0.1}$$

donde $\delta = \psi - \hat{\psi}$, $\hat{F}.$ denota la matriz de primeras derivadas dada por

$$\hat{F}. = [(\partial f_i(\psi)/\partial \psi_j]_{\psi = \hat{\psi}}$$

y $\hat{F}..$ denota el arreglo de segundas derivadas dado por

$$\hat{F}.. = [(\partial^2 f_i(\psi_r)/\partial \psi_r \partial \psi_s]_{\psi = \hat{\psi}} = [(\hat{f}_{rs})].$$

Si se ignora el término cuadrático en (2.0.1), tenemos la aproximación lineal para ψ en una vecindad de $\hat{\psi}$,

$$\eta - \hat{\eta} \approx \hat{F}.(\psi - \hat{\psi}) \tag{2.0.2}$$

Por lo tanto una región de confianza al $100(1-\alpha)\%$ para ψ es el conjunto Ψ en el plano tangente tal que:

$$||\eta - \hat{\eta}||^2 \approx ||F.(\psi - \hat{\psi})|| \leq ps^2 F_\alpha \tag{2.0.3}$$

donde $F_\alpha = F^{\alpha}_{p,n-p}$ es el percentil $1-\alpha$ de la distribución $F(p, n-p)$ y $s^2 = ||y - \hat{\eta}||^2/(n-p)$, siendo $||\cdot||$ la norma usual.

De (2.0.3) se observa que η está contenida en una esfera con centro $\hat{\eta}$ y radio $(ps^2F_\alpha)^{1/2}$. Es apropiado introducir en esta parte el parámetro $\rho = s\sqrt{p}$ llamado el radio estándar, así que $\rho\sqrt{F_\alpha}$ es el radio de la esfera arriba mencionada . Nótese que

$$(\psi - \hat{\psi})^t \hat{F}.^t \hat{F}.(\psi - \hat{\psi}) \leq \rho^2 F_\alpha$$

es un elipsoide con centro en $\hat{\psi}$.

La validez de la aproximación al plano tangente (2.0.2) dependerá de la magnitud del término cuadrático $\delta^t \hat{F}..\delta$ en (2.0.1). Al hacer esta comparación se halla que es útil dividir el término cuadrático, (un vector n dimensional), en dos componentes ortogonales; las respectivas proyecciones al plano tangente $\hat{F}^T$ y normal al plano tangente $\hat{F}^N$. Esta descomposición puede llevarse a cabo utilizando la matriz de proyección,

$$\hat{P}_F = \hat{F}.(\hat{F}.^t\hat{F}.)^{-1}\hat{F}.^t$$

Las dos componentes están dadas por: $\hat{F}.. = \hat{F}^T.. + \hat{F}^N..$, donde $\hat{F}..^T = [(\hat{f}^T_{rs})]$, $\hat{F}..^N = [(\hat{f}^N_{rs})]$ con $\hat{f}^T_{rs} = \hat{P}_F \hat{f}_{rs}$ y $\hat{f}^N_{rs} = (I_n - \hat{P}_F)\hat{f}_{rs}$ Los conceptos de no linealidad intrínseca y de efectos de los parámetros, medidas básicas para medir no linealidad, se precisarán en los numerales siguientes.

2.1. Curvatura Intrinseca

De la discusión previa, Bates & Watts definen la siguiente medida para comparar cada componente cuadrática con el término lineal, considerando la proyección al plano normal:

$$K_\delta^N = ||\delta^t \hat{F}..^N \delta|| \,/\, ||\hat{F}.\delta||^2 \tag{2.1.1}$$

La razón exhibida en (2.1.1) se conoce como curvatura intrínseca, en la dirección δ. Puesto que $f_{rs} = f_{rs}^T + f_{rs}^N$, se tiene

$$||\delta^t F..\delta||^2 = ||\delta^t F..^T \delta||^2 + ||\delta^t F..^N \delta||^2. \tag{2.1.2}$$

La no linealidad intrínseca (**IN**) mide la curvatura del lugar geométrico (conjunto de soluciones) en un espacio muestral, donde el lugar representa todas las posibles soluciones al problema de estimación. La solución de mínimos cuadrados es el punto sobre el lugar geométrico más cercano al vector respuesta.

Para un modelo de regresión lineal, **IN** es cero ya que el lugar geométrico es recto (una línea, un plano, un hiperplano). Para un modelo de regresión no lineal, el lugar geométrico es curvado con **IN** midiendo el grado de Curvatura. Afortunadamente, como lo hallaron Bates y Watts (1980) y Ratkowsky (1983) y estudios subsiguientes de éste último autor, por lo general IN es pequeño para casi todos los modelos de interés práctico. Esto significa que si un modelo es alejado de lo lineal, la mayor contribución a la no linealidad se debe a la parametrización.

2.2 Efectos De Los Parametros

Considerando la proyección al plano tangente , de la discusión previa se tiene la siguiente medida que sirve, también, para comparar cada componente cuadrática con el término lineal, Bates & Watts (1980), La razón exhibida en (2.2.1) se conoce como la curvatura de efectos de los parámetros denotada por PE. Esta es una medida de la ausencia de paralelismo y de la desigualdad de espaciamiento de las lineas de los parámetros sobre el lugar geométrico en la solución de mínimos cuadrados.

Si IN es pequeña podrá obtenerse una reparametrización que haga PE tan pequeña como se quiera. Una guía práctica para decidir si IN o PE son pequeñas, es comparada con los valores $1/\sqrt{F_\alpha}$ y $1/2\sqrt{F}$ respectivamente

3. Medidas De Sesgo Y Asimetria

Cook et al, citado por Ratkowsky (1990), mostraron que las aproximaciones a los sesgos propuestos por Cox y Snell; Box; Clarke, Amari y Hougaard son similares para el modelo de regresión lineal bajo las consideraciones que se hacen.

La fórmula expuesta por Box (1971) para calcular el sesgo en los estimadores mínimo-cuadráticos de los parámetros en los modelos de regresión no lineal, está dada por,

$$\text{Sesgo}(\hat{\psi}) = -\sigma^2/2[\sum_{u=1}^{n} F_u F_u^T\}^{-1} \sum_{t=1}^{n} F_t \, tr[\{\sum F_u F_u^T\}^{-1} H_t]$$

donde $F_t(= F_u)$ es el vector p-dimensional de primeras derivadas de $f(X_t, \psi)$ y $\mathbf{H}_t$ es la matriz $p \times p$ de segundas derivadas con respecto a cada uno de los parámetros de ψ, evaluadas en $\mathbf{X}_t$

El porcentaje de sesgo del estimativo mínimo cuadrático ordinario es una medida útil puesto que un valor absoluto mayor al 1% parece ser una buena regla práctica como indicativo del comportamiento no lineal (Cálculo del sesgo de J.M Box). Sin embargo, se debe enfatizar que el sesgo de Box al ser una cantidad dependiente de la localización, puede conducir a inferencias equivocadas cuando los estimadores estén cercanos a cero. Por lo anterior, es mejor utilizar la medida directa de asimetría de Hougaard, presentada por Ratkowsky (1983), la cual se describe a continuación.

Para calcular la medida de asimetría de Hougaard, se utiliza el término L^{ik} que denota un elemento de

$$\mathbf{L} = [\mathbf{J}^T(\hat{\psi})\mathbf{J}(\hat{\psi})]^{-1},$$

donde $\mathbf{J}(\hat{\psi})$ es la matriz jacobiana $n \times p$ (Ratkowsky,1983, Apéndice 2.A, para una detallada definición de esta matriz) con elemento típico $\mathbf{J}_{mj}$, evaluado en $\hat{\psi}$. Sea $s^2 = RSS(\hat{\psi})/(n-p)$ el estimador de la varianza de los residuales calculados con la suma de cuadrados de los residuales en $\hat{\psi}$ y sea $\mathbf{H}(\hat{\psi})$ la el arreglo $n \times p \times p$ de segundas derivadas con respecto a los parámetros, con elemento típico $\mathbf{H}_{mkl}$, evaluado en $\hat{\psi}$. Si $\mathbf{W}_{jkl}$ es el número definido por $\mathbf{W}_{jkl} = \sum \mathbf{J}_{mj}\mathbf{H}_{mkl}$, un estimador del tercer momento está dado por

$$E[\hat{\psi}_i - E(\hat{\psi}_i)]^3 = -(s^2)^2 \sum_{j,kl=1}^{p} \mathbf{L}^{ij}\mathbf{L}^{ik}\mathbf{L}^{il}(\mathbf{W}_{jkl} + \mathbf{W}_{kjl} + \mathbf{W}_{lkj}).$$

El tercer momento puede ser convenientemente estandarizado usando el elemento apropiado de la matriz de covarianza asintótica que produce

$$g_{1i} = E[\hat{\psi}_i - E(\hat{\psi}_i)]^3/(s^2\mathbf{L}^{ii})^{3/2},$$

la cual genera una medida directa de la asimetría $\hat{\psi}_i$. Dado que la medida g_1 es familiar en estadística y dada la cercanía entre el grado de comportamiento no lineal del estimador y el grado de no normalidad en la distribución muestral del estimador.

$\hat{\psi}_i$, ponderado por g_{1i}, es aproximadamente lineal o contiene una considerable no linealidad. Así, es posible decir que si $g_{1i} < 0.1$, el estimador $\hat{\psi}_i$ de el parámetro ψ_i tiene un comportamiento aproximadamente lineal en su comportamiento y si $0.1 < g_{1i} < 0.25$, el estimador es razonablemente lineal. Si $g_{1i} \geq 0.25$, la asimetría es aparente y $g_{1i} > 1$ indica un considerable comportamiento no lineal.

4. Ejemplos

EJEMPLO 1: Como primera ilustración se considera la información presentada por González (1994) en un trabajo dedicado al análisis del crecimiento de Prioria Copaifera.g en condiciones naturales. Se quiere determinar si el modelo

$$Y = P_1 X_1^{1.157} + (P_2 + P_3 X_2 + P_4 X_3) X_1 + \epsilon_i \ ,$$

considerado por el autor es no lineal o intrínsecamente no lineal, siendo

Y: crecimiento anual del árbol

X_1: diámetro del árbol

X_2: posición de la copa

X_3: índice de competencia

Se consideraron como estimativos iniciales de los parámetros los siguientes

$$g_0^{(0)} = -0.063801; \qquad g_1^{(0)} = 0.15413;$$
$$g_2^{(0)} = -0.0080734; \qquad g_3^{(0)} = -0.0050511.$$

En la Tabla 1 se presenta un resumen de los resultados obtenidos mediante el programa RENOL.

ESTIMATIVOS PARAMETROS	SESGO	%SESGO	HOUGAARD	IN	PE
-0.581E-1	0	0	0	0	0
0.138E-1	0	0	0	0	0
-0.574E-2	0	0	0	0	0
-0.505E-2	0	0	0	0	0

Tabla 1: Estimadores y medidas de no linealidad para

el incremento medio anual.

El modelo ajustado toma la forma

$$Y = -0.0581084X_1^{1.157} + (0.138085 - 0.00574896X_2 - 0.00505418X_3)X_1 \quad (4.0.1)$$

El porcentaje de sesgo para los cuatro parámetros estimados no supera el 1%. Desde este punto de vista podría afirmarse que el modelo se comporta linealmente. Así mismo el tercer momento asintótico de Hougaard para los cuatro parámetros , al no exceder a 0.1, corrobora la afirmación previa.

En cuanto a las medidas de no linealidad (**IN** y **PE**), al ser menores que $1/2\sqrt{F}$, reafirman el hecho de que el modelo se comporta linealmente.

En conclusión, las inferencias usuales de regresión lineal se pueden extender con este modelo, lo cual era de esperarse ya que el modelo elegido es realmente de regresión polinomial lineal.

EJEMPLO 2: El segundo ejemplo ilustrativo es tomado de un estudio de Vélez et al (1987) para analizar la demanda residencial de energía eléctrica en dos ciudades colombianas.

El modelo para el caso de Medellín es :

$$Q_t = 2.8\{\alpha C_t^{\beta/Dt} Y_t^{\tau}\}^{Dt/(Dt-\beta)} \quad (4.0.2)$$

donde $t = 1970, \ldots, 1983$; $N = 14$ datos

Q_t: Consumo del suscriptor medio.

C_t: Intercepto de la función de oferta cuando el precio es uno.

D_t: Elasticidad de la oferta respecto al precio.

Y_t: Ingreso per cápita.

α: Efecto de las preferencias y necesidades de los suscriptores sobre la demanda.

β: Elasticidad de la demanda con respecto al precio.

τ: Elasticidad de la demanda con respecto al ingreso.

ϵ: Error del modelo.

En Yáñez (1988) se analiza el modelo de Vélez et al (1987) donde se dice que "el ajuste fue satisfactorio razonando por analogía al caso lineal". El modelo se utilizó para hacer inferencias basadas en la interpretación de los estimativos de los parámetros y para proyectar demanda de energía eléctrica por subscriptor medio desde 1984 hasta el 2005. Yáñez (1988) utilizó la medida del sesgo de Box como indicativa del grado de no linealidad del modelo; con ella se llegaron a resultados inconsistentes con respecto a la conclusión anterior, resultados que permitían aseverar que el modelo se comportaba no linealmente.

Mediante el programa RENOL se obtuvieron los resultados que se exhiben en la Tabla 2.

ESTIMATIVOS DE LOS PARAMETROS	%SESGO	HOUGAARD	IN	PE
0.166E+3	21.895	2.018		
0.258E-1	0.061	2.720	0.051	123.25
0.306E+0	0.016	-0.030		

Tabla 2. Estimadores y medidas de no linealidad para la

demanda residencial de energia eléctrica

al considerar medida de curvatura máxima puede observarse que **IN** = 0.0151. Dado que $F(P;N-P;0.05) = F(3;11;0.05) = 3.59$, y **IN** $< 1/2\sqrt{F} = 0.263$, la no linealidad intrínseca es adecuadamente baja. Análogamente $PE >>> 1/2\sqrt{F} = 0.263$, con $PE = 123.2522$ mostrando una no linealidad de los efectos de los parámetros notablemente alta. Como puede observarse el modelo se comporta no linealmente. Las inferencias sobre el modelo (4.0.2) no pueden extenderse de las inferencias en regresión lineal.

Desde el punto de vista estadístico se concluye que el modelo debe ser revisado ya que tal como está es altamente no lineal y las inferencias respecto a los parámetros y a la predicción no tienen ninguna validez estadística a pesar de que los diagnósticos realizados sobre el modelo eran " buenos ", utilizando criterios análogos a los de regresión lineal.

Para complementar el análisis de este ejemplo podría pensarse en hacer una reparametrización y con ella ajustar los datos. Pero ¿cuál reparametrización podría hacerse ?. Esta es la pregunta natural. El valor inconsistente de PE para este modelo, no así el de IN, sugiere en primera instancia pensar en una reparametrización del modelo.

5. Comentarios Finales

5.1 Se puede recomendar a **IN**, **PE** y el estadístico de Hougaard como las medidas básicas de no linealidad que deben analizarse con el fin de considerar la extensión de las inferencias de la regresión lineal con respecto a un modelo intrínsecamente no lineal.

5.2 Si simultáneamente la no linealidad de los efectos de los parámetros es despreciable, más válidos serán los test estadísticos cuya justificación reposa en los supuestos de linealidad. Como en un modelo de regresión lineal los parámetros son lineales, se tiene que las segundas derivadas son todas cero, lo cual implica que PE igual a cero. En los modelos de regresión no lineal PE crece a medida que el modelo se desvía de un modelo lineal.

5.3 Podría suceder que al ajustar un grupo de datos a un modelo de regresión no lineal tanto IN como PE o alguna de éstas dos medidas estuvieran estrechamente cercanas al valor frontera , $1/\sqrt{F}$ o $1/2\sqrt{F}$, cantidad que permite decidir cuando podrían extenderse o no las inferencias del caso lineal al no lineal. Sería interesante hacer un estudio de simulación para detectar si en la frontera de éstos valores es apresurado hacer la extensión inferencial.

5.4 En la Revista Colombiana de Estadística, Número 27, se hace una exposición de los fundamentos teóricos de la Regresión No Lineal con un ejemplo donde el tamaño de la muestra es pequeño. Sin embargo no se menciona la importancia de las medidas de Bates and Watts como guía de extensión de las inferencias del campo lineal al campo no lineal. Si esto se ignora, se pueden cometer errores graves en la extensión inferencial.

BIBLIOGRAFÍA

BATES, D. M. y WATTS, D. G. (1980), *Relative Curvature Measures of Nonlinearity*, J. R. Statist. Soc., Ser. B 42, 1-25.

BOX, M. J. (1971), *Bias in Nonlinear Estimation*, J. R. Statist. Soc., Ser. B 33, 171- 201.

CASTELLANOS, R. (1991), *Regresión No Lineal*, Trabajo de grado presentado para optar al título de Matemático. Universidad de Antioquia, Facultad de Ciencias Exactas y Naturales, pág:136-149.

DRAPER, N.R. y SMITH, H. (1981), *Applied Regression Analysis*, Willey, New York.

Gonzalez, H. (1994), *Análisis del crecimiento de Prioria Copaifera.g en condiciones naturales por medio de modelos matemáticos*, Tesis de Magister, Posgrado en Silvicultura y Manejo de Bosques, Universidad Nacional, Medellín.

NETTER, WASSERMAN y KUTNER (1985), *Applied Linear Statistical Models*, Segunda Edición.

RATKOWSKY, D.A. (1983), *Nonlinear regression modeling*, Mercel Dekker, New York.

RATKOWSKY, D.A. (1990), *Handbook of nonlinear regression models.*, Mercel Dekker, New York.

SEBER, G.A.F. (1989), *Nonlinear regression*, Willey, New York:.

VELEZ, C.E, BOTERO J., YAÑEZ S. (1987), Memorias del VII Encuentro Latinoamericano de la Econometric Society, *La demanda residencial de energía eléctrica en dos ciudades colombianas 1970-1983*, Sao Paulo Brasil.

YAÑEZ C., SERGIO. (1988), *Inferencia en Regresión No Lineal*, Revista Colombiana de Estadística, 17-18, pp 75-88.