MODELO DE REGRESSÃO HETEROSCEDÁSTICO

Sergio Augusto RODRIGUES[1]
Carlos Alberto Ribeiro DINIZ[2]

- RESUMO: Na teoria clássica de regressão, uma das suposições básicas é a homocesdasticidade dos erros. Quando isso não ocorre, uma alternativa é transformar os dados ou estimar a estrutura heteroscedástica dos mesmos. Este trabalho apresenta o modelo de regressão linear heteroscedástico, bem como uma revisão das alternativas clássicas de estimação e detecção da heteroscedasticidade.

- PALAVRAS-CHAVE: Análise de regressão; heteroscedasticidade; métodos de estimação.

1 Introdução

Na teoria clássica dos modelos de regressão existem suposições básicas para garantir a eficiência e a qualidade de seus estimadores. Quando uma ou mais destas suposições forem violadas é necessário procurar alternativas desejáveis para garantir a eficiência e a qualidade dos estimadores.

Uma das suposições do modelo de regressão é a igualdade das variâncias dos erros. Quando esta não ocorre, diz-se que os erros são heteroscedásticos.

Este trabalho apresenta uma revisão sobre alguns métodos clássicos de estimação dos parâmetros de um modelo de regressão heteroscesdástico, mostrando suas características, os principais testes para detectar variâncias não constantes dos erros e as estruturas heteroscedásticas mais conhecidas na literatura.

[1]Faculdade de Educação São Luis - FESL, CEP 14870-370, Jaboticabal, SP, Brasil. E-mail: *sergio@saoluis.br*
[2]Departamento de Estatística, Universidade Federal de São Carlos - UFSCar, Caixa Postal 676, CEP 13565-905, São Carlos, SP, Brasil. E-mail: *dcad@power.ufscar.br*

2 Modelo de regressão linear heteroscedástico

Quando os erros de um modelo de regressão não têm variância constante pode-se dizer que a suposição de homoscedasticidade não está satisfeita e o modelo pode ser chamado de modelo de regressão heteroscedástico.

Algumas vezes é possível eliminar a heteroscedasticidade por meio de uma transformação na variável resposta (Box e Cox, 1964). Mas, como isto nem sempre é possível, é conveniente considerar uma análise com modelagem explícita da variância.

Um modelo de regressão linear múltipla amostral heteroscedástico é dado por

$$y_i = \mathbf{x_i}'\beta + e_i, \qquad i = 1, 2, ..., n, \tag{1}$$

onde $\mathbf{x}$ é um vetor $k \times 1$ de variáveis regressoras, β é um vetor $k \times 1$ de parâmetros desconhecidos e os $e_i^{'}s$ são os erros do modelo de regressão, os quais são supostamente independentes e normalmente distribuídos com média zero e variância σ_i^2. Duas estruturas de heteroscedasticidade são usualmente consideradas: estrutura aditiva ou estrutura multiplicativa.

2.1 Tipos de heteroscedasticidade

Considerando que os erros tenham distribuição normal com média zero e variância σ_i^2, dois tipos de erros heteroscedásticos bastante populares na literatura (Judge et al., 1985), são dados por:

(a) Estrutura heteroscedástica aditiva: A variância dos erros é proporcional a uma das variáveis regressoras ao quadrado. Neste caso, para um modelo com um termo constante e uma variável regressora, tem-se

$$Var(e_i) = \sigma_i^2 = (\mathbf{z}_i'\delta)^2 = \alpha_0^2(1 + \gamma x_i)^2 \tag{2}$$

onde $\delta = (\alpha_0, \alpha_1)'$, $\gamma = \dfrac{\alpha_1}{\alpha_0}$ e $\mathbf{z}_i' = (1,\ x_i)$. Um modelo com essa estrutura é denominado modelo aditivo.

Nesta estrutura γ representa o grau ou nível da heteroscedasticidade, α_0 um parâmetro de proporcionalidade da variância reparametrizada dos erros e x_i uma das variáveis regressoras do modelo.

(b) Estrutura heteroscedástica multiplicativa: A variância dos erros é proporcional a uma potência desconhecida de uma das variáveis regressoras. Neste caso, para um modelo com uma constante e uma variável regressora, a variância é dada por

$$Var(e_i) = \sigma_i^2 = \sigma^2 x_i^\lambda = \exp(\mathbf{z}_i'\alpha) \tag{3}$$

onde $\alpha = (\log \sigma^2,\ \lambda)'$ e $\mathbf{z}_i' = (1, \log x_i)$.

O parâmetro λ representa o grau ou nível da heteroscedasticidade, x_i é uma das variáveis regressoras do modelo e σ^2 é um parâmetro desconhecido representando a variância do erro nos casos onde $\lambda = 0$.

Essa estrutura foi discutida por Geary (1966), Park (1966), Lancaster (1968), Kmenta (1971) e Harvey (1976); este último denominou-o de modelo multiplicativo. Esse modelo pode ser considerado um caso especial do modelo aditivo.

2.2 Conseqüências da heteroscedasticidade

Quando a suposição de homoscedasticidade ou de independência entre os erros for violada, o modelo de regressão linear clássico pode ser substituído pelo modelo de regressão linear generalizado. As três maiores consequências de se utilizar o estimador de mínimos quadrados ordinários, dado em (6), para os parâmetros do modelo de regressão linear generalizado são: i) as variâncias dos estimadores não são mínimas; ii) embora o estimador de mínimos quadrados ordinários seja não viciado, sua variância é viciada e estimativas intervalares ou testes de hipóteses podem produzir resultados imprecisos ou inválidos. Não é sempre possível detectar se o modelo clássico ou generalizado deva ser empregado, existindo, portanto, um perigo em se fazer inferências falhas usando o estimador de mínimos quadrados ordinários; iii) com a suposição adicional de que os erros são normalmente distribuídos, o estimador de mínimos quadrados ordinários não é o estimador de máxima verossimilhança.

Todas essas conseqüências do uso do estimador de mínimos quadrados ordinários quando os erros são heteroscedásticos sugerem o uso de estimador de mínimos quadrados generalizados. No entanto, o problema de se utilizar esse estimador é que a matriz de variâncias e covariâncias dos erros deve ser conhecida.

2.3 Testes para detectar a heteroscedasticidade

É possível detectar heteroscedasticidade por meio de uma inspeção visual dos resíduos via diagrama de dispersão. Se parecer que os resíduos absolutos possuem a mesma variabilidade em torno de uma linha imaginária, então, provavelmente, não existe heteroscedasticidade. Caso contrário, uma verificaçã

2.3.1 Teste de Goldfeld-Quandt (1965)

Neste teste as n observações são primeiramente ordenadas de acordo com os valores da variável regressora, x, responsável pela presença de heteroscedasticidade. Em seguida é necessário dividir a amostra ordenada em três partes, desconsiderando os c's valores do meio (25% aproximadamente). Com as duas partes restantes, ajusta-se dois modelos de regressão, um utilizando somente os valores da primeira parte (menores valores da variável regressora), e o outro utilizando os valores da segunda parte (maiores valores da variável regressora). Em seguida, verifica-se pelo teste F, se as variâncias dos erros nestas duas regressões são iguais contra a hipótese de que a variância dos erros é maior na segunda parte dos dados, ou seja, no conjunto de dados com grandes valores de x.

É importante notar que, no caso de uma regressão linear múltipla, primeiramente é necessário identificar com qual das variáveis regressoras o resíduo está se relacionando. O teste de Goldfeld-Quandt deve ser realizado utilizando-se cada uma das variáveis regressoras suspeitas de se relacionar com o erro.

Esse teste apresenta algumas limitações, entre elas, exige amostra relativamente grande; pressupõe que a heteroscedasticidade seja decorrente da relação entre o erro aleatório e uma ou algumas das variáveis regressoras. Existem, porém, casos em que a heteroscedasticidade é resultante da utilização de dados agrupados e casos que não apresentam indicações sobre a exata relação existente entre a variável regressora e os resíduos.

2.3.2 Teste de Szroeter (1978)

O teste de homoscedasticidade de Szroeter rejeita a hipótese de homoscedasticidade se a estatística Q,

$$Q = \frac{T(\widetilde{h} - \bar{t})}{\left[2\sum_{t=1}^{T}(t-\bar{t})^2\right]^{1/2}}, \tag{4}$$

for maior que um valor crítico de uma distribuição normal padrão, com $\bar{t} = \sum_{t=1}^{T}\frac{t}{T}$ e $\widetilde{h} = \frac{\sum_{t=1}^{T} t\widehat{e}_t^2}{\sum_{t=1}^{T}\widehat{e}_t^2}$, onde os $t's$ são constantes representando a ordem dos resíduos e os $\widehat{e}_t$'s são os resíduos gerados pelo modelo.

Sob homoscedasticidade, $\widetilde{h}$, que é uma média ponderada dos $t's$ com pesos dados por $\widehat{e}_t^2$, é aproximadamente igual a uma média aritmética simples dos $t's$. Para maiores detalhes sobre a forma deste teste, ver Judge et al. (1985, p.450-453). Griffiths e Surekha (1984) mostraram que o teste de Szroeter é melhor que o teste de Goldfeld-Quandt quando as variâncias dos erros estão ordenadas.

2.3.3 Teste de Glejser (1969)

Esse teste, além da facilidade de uso, permite que se supere a limitação de não poder identificar o tipo de relação existente entre a variável regressora e os resíduos. O teste de Glejser dá indicações sobre o padrão da heteroscedasticidade. Essa informação é útil para a escolha de um método conveniente de estimação do modelo heteroscedástico. Os procedimentos para realização deste teste requer a estimação de um modelo de regressão. Com os resíduos obtidos, estima-se as seguintes regressões: $|\widehat{e}_i| = a_0 + a_1 x_i^h + v_i$, sendo x_i a variável regressora com a qual se supõe que e_i esteja relacionado, v_i o erro aleatório das regressões propostas e h uma potência qualquer, $h = -1, 1, 1/2, 2, ...$ Em seguida, teste a hipótese $H_0 : a_1 = 0$ *versus* $H_1 : a_1 \neq 0$. Se a hipótese nula $H_0 : a_1 = 0$ não for rejeitada para todas as regressões estimadas, então não existe heteroscedasticidade. Se, em pelo menos uma das regressões, a hipótese nula é rejeitada, então existe heteroscedasticidade nos dados.

A limitação mais importante deste teste é que ele só é válido para verificar a heteroscedasticidade decorrente da relação entre o termo aleatório e alguma das variáveis regressoras.

2.3.4 Teste de White (1980)

Este teste examina se a variância dos erros é afetada por alguma das variáveis regressoras, seus quadrados ou seus produtos cruzados. É baseado na idéia de que a variância do estimador de mínimos quadrados ordinários é igual à variância do estimador de mínimos quadrados generalizados e detecta a heteroscedasticidade somente se a variância desigual dos erros afetar a consistência do estimador da matriz de variâncias e covariâncias dado em (8).

White (1980) mostrou que a estatística de teste para detectar a heteroscedasticidade nos dados é igual ao tamanho da amostra, n, multiplicado pela estatística R^2, calculada a partir da regressão ajustada entre os resíduos ao quadrado e o produto cruzado das variáveis regressoras mais a constante. No caso de duas variáveis regressoras, a regressão a ser ajustada seria:

$$\widehat{e}_i^2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2. \tag{5}$$

A estatística de White segue, assintoticamente, distribuição qui-quadrado, χ^2, com o número de graus de liberdade igual ao número de variáveis regressoras do modelo usado para calcular a estatística R^2.

3 Métodos de estimação clássica

Com as suposições do modelo de regressão satisfeitas, o estimador não viciado e de variância mínima para os coeficientes de um modelo de regressão linear múltiplo, e que minimiza a soma de quadrados de resíduos, é o estimador de mínimos quadrados ordinários (EMQO), dado por

$$\widehat{\beta} = \left(X'X\right)^{-1} X'\mathbf{y}. \tag{6}$$

onde X é a matriz do modelo e y é o vetor de respostas. A matriz de variâncias-covariâncias do EMQO é dada por

$$Cov(\widehat{\beta}) = (X'X)^{-1} X'\Omega X (X'X)^{-1}, \tag{7}$$

onde Ω é a matriz de variâncias e covâriancias dos erros.

No caso de homoscedasticidade, um estimador não viciado dessa matriz de variâncias e covariâncias é dado por

$$\widehat{Cov}(\widehat{\beta}) = S^2(X'X)^{-1}, \tag{8}$$

onde

$$S^2 = \frac{\widehat{\mathbf{e}}'\widehat{\mathbf{e}}}{(n-p)} \tag{9}$$

é o estimador da variância dos resíduos.

Este estimador é bastante popular na estatística, pois com as suposições iniciais satisfeitas minimiza a soma de quadrados de resíduos, é não viciado e de variância mínima e se os erros tiverem uma distribuição normal, coincide com o estimador de máxima verossimilhança.

No caso de heteroscedasticidade dos erros, o EMQO continua sendo um estimador não viciado, no entanto, não será um estimador eficiente.

Quando a função que determina a heteroscedasticidade dos erros é completamente conhecida é fácil determinar a matriz de variâncias e covariâncias do EMQO. Quando não se conhece a estrutura de heteroscedasticidade, pode-se utilizar o estimador proposto por White (1980), que é não viciado para grandes amostras e consistente nos casos homoscedástico e heteroscedástico de forma desconhecida. Esse estimador é dado por

$$\widehat{Cov}(\widehat{\beta}) = (X'X)^{-1}X'\widehat{\Omega}X(X'X)^{-1}, \tag{10}$$

onde $\widehat{\Omega} = diag(\widehat{e}_1^2, ..., \widehat{e}_n^2)$, uma matriz diagonal com os quadrados dos resíduos do EMQO na diagonal. No entanto, esse estimador possui vícios altos para pequenas amostras. A fim de diminuir os vícios para pequenas amostras, Cribari Neto ; Silva e Cordeiro (2000) definiram um estimador modificado de White (1980) para a matriz de covariância com erros heteroscedásticos.

Nas situações em que pelo menos uma das suposições da teoria clássica de modelos lineares for violada, uma alternativa seria utilizar o estimador de mínimos quadrados generalizados (EMQG)

$$\widehat{\beta} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}\mathbf{y}. \tag{11}$$

A matriz de variâncias e covariâncias deste estimador é dada por

$$Cov(\widehat{\beta}) = \left(X'\Omega^{-1}X\right)^{-1}. \tag{12}$$

Quando Ω é completamente conhecida é fácil calcular o EMQG e sua matriz de variâncias e covariâncias, mas quando Ω é desconhecida é necessário estimá-la. Quando a função que determina a heteroscedasticidade é conhecida, mas com alguns parâmetros desconhecidos, uma saída seria estimar esta heteroscedasticidade. Harvey (1976) sugere um estimador para o grau da heteroscedasticidade multiplicativa e Harvey (1974) um estimador para o grau da heteroscedasticidade aditiva, considerando independência entre os erros. Uma outra alternativa seria obter uma estimativa global para o grau de heteroscedasticidade e para os coeficientes da regressão, utilizando os algoritmos propostos por Broyden (1970), Fletcher (1970), Goldfarb (1970) e Shanno (1970).

Dependendo da estrutura heteroscedástica dos erros, a maximização da função de máxima verossimilhança pode ser complicada, uma vez que vários parâmetros podem estar presentes. O método Scoring, baseado no processo iterativo de Newton-Raphson, pode ser usado para facilitar a maximização da função de verossimilhança.

Estimando a estrutura heteroscedástica dos erros, e considerando independência entre eles, é possível obter uma estimativa da matriz de variâncias e covariância , $\widehat{\Omega}$, e conseqüentemente, utilizar o EMQG para estimar os coeficientes da regressão.

3.1 Estrutura heteroscedástica aditiva

Neste tipo de estrutura é assumido que os erros são da forma aditiva, ou melhor, $Var(e_i) = \sigma_i^2 = (\mathbf{z}_i'\delta)^2$, $i = 1, ..., n$, onde $\mathbf{z}_i$ é um vetor $p \times 1$ $(p \leq k)$ de um conjunto de variáveis que normalmente são as variáveis regressoras do modelo e δ é um vetor $p \times 1$ de parâmetros da função que forma a variância dos erros.

O caso especial desse tipo de heteroscedasticidade que será considerado neste artigo é dado em (2), ou seja, $Var(e_i) = \alpha_0^2(1 + \gamma x_i)^2$. Neste caso, o parâmetro γ representa o grau ou nível da heteroscedasticidade, x_i é uma das variáveis regressoras do modelo e α_0 um parâmetro de proporcionalidade da variância reparametrizada dos erros. Harvey (1974) propôs um estimador para γ baseado na suposição de que os erros tem uma distribuição normal com média zero e variância σ_i^2.

3.1.1 O estimador de Harvey (1974)

Por suposição, as variáveis padronizadas $\sigma_i^{-1}e_i$ para $i = 1, ...n$, são independentes e identicamente distribuídas com média zero e variância um. Entretanto, o valor esperado do valor absoluto de uma variável padronizada pode ser escrita por $E\left(\sigma_i^{-1}|e_i|\right) = c$, onde c é uma constante igual para todos i's e depende da distribuição de $\sigma_i^{-1}|e_i|$.

Além disso, como $E(|e_i|) = c\sigma_i$ e, se os $e_i's$ são normalmente distribuídos, Harvey mostrou que $c = (2/\pi)^{1/2}$. Destes resultados pode-se escrever

$$|\widehat{e}_i| = c\mathbf{z}_i'\delta + v_i, \tag{13}$$

onde $\widehat{e}_i = y_i - \mathbf{x}_i'\widehat{\beta}$ é o *i-ésimo* resíduo do ajuste do modelo de regressão e v_i é um novo "resíduo", $v_i = |\widehat{e}_i| - E(|e_i|)$.

Aplicando o EMQO no modelo acima, obtém-se um estimador para $c\delta$ e, conseqüentemente, para δ,

$$\widehat{\delta} = c^{-1}\left(\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\sum_{i=1}^{n}\mathbf{z}_i\left|\widehat{e}_i\right|. \tag{14}$$

Harvey (1974) mostrou que $|\widehat{e}_i|$ converge em distribuição para $|e_i|$ e, portanto, assintoticamente, v_i será independentemente distribuída com média zero, pois $E\left(|\widehat{e}_i|\right) = E\left(|e_i|\right)$ quando $n \to \infty$ e $E\left(v_i\right) = E\left(|e_i|\right) - E\left(|e_i|\right) = 0$, e variância dada por

$$\begin{aligned} Var\left(v_i\right) &= Var\left(|\widehat{e}_i|\right) \\ &= E\left(|e_i|^2\right) - \left[E\left(|e_i|\right)\right]^2 = E\left(e_i^2\right) - \left[E\left(|e_i|\right)\right]^2 \\ &= \sigma_i^2 - c^2\sigma_i^2 = \left(\mathbf{z}_i'\delta\right)^2\left(1-c^2\right). \end{aligned}$$

Como v_i não é homoscedástico, o EMQO, $\widehat{\delta}$, não é um estimador eficiente. Entretanto, é um estimador consistente. Nesta situação, o ideal seria utilizar o EMQG, utilizando $\widehat{\delta}$ como estimativa da variância dos erros. Assim, aplicando o EMQG no modelo (13), tem-se

$$\widehat{\widehat{\delta}} = c^{-1}\left(\sum_i\left(\mathbf{z}_i'\widehat{\delta}\right)^{-2}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\sum_i\left(\mathbf{z}_i'\widehat{\delta}\right)^{-2}\mathbf{z}_i\left|\widehat{e}_i\right|, \tag{15}$$

onde $c = (2/\pi)^{1/2}$.

O estimador $\widehat{\widehat{\delta}}$ foi sugerido por Harvey (1974) e também descrito por Judge et al. (1985). A constante c é utilizada porque $E\left(|e_i|\right) \neq 0$ e é necessário assumir que e_i é normalmente distribuído.

A partir disso, o estimador de mínimos quadrados generalizado de β pode ser calculado por

$$\widehat{\beta} = \left(\sum_{i=1}^{n}\left(\mathbf{z}_i'\widehat{\widehat{\delta}}\right)^{-2}\mathbf{x}_i\mathbf{x}_i'\right)^{-1}\sum_{i=1}^{n}\left(\mathbf{z}_i'\widehat{\widehat{\delta}}\right)^{-2}\mathbf{x}_i y_i. \tag{16}$$

3.2 Estrutura heteroscedástica multiplicativa

Nesta estrutura é assumido que os erros são da forma multiplicativa, ou seja, $Var(e_i) = \sigma_i^2 = \exp(\mathbf{z}_i'\alpha) = \sigma^2 x_i^{\lambda}$, $i = 1, ..., n$, onde $\mathbf{z}_i$ é um vetor $p \times 1$ $(p \leq k)$ de um conjunto de variáveis que são, normalmente, mas não necessariamente, as variáveis regressoras do modelo e α é um vetor $p \times 1$ de parâmetros da função que forma a variância dos erros. O primeiro elemento de $\mathbf{z}_i$ sempre será igual a 1, devido ao termo constante do modelo. O parâmetro λ representa o grau ou nível da heteroscedasticidade, σ^2 é um parâmetro desconhecido representando a variância do erro nos casos onde $\lambda = 0$. Quando não é conhecido o verdadeiro valor de λ, Park (1966) propôs um estimador em 2 passos para um modelo geral com um certo grau de heteroscedasticidade da forma multiplicativa. Harvey (1976) propôs um outro estimador, chamado de estimador em 3 passos de Harvey, obtido pelo processo iterativo Scoring.

3.2.1 O estimador em 2 passos de Park

Park (1966) propôs um procedimento de estimação, denominado estimador em 2 passos. O primeiro passo consiste em tomar os logaritmos dos quadrados dos resíduos obtidos no ajuste de um modelo considerando o EMQO. A razão para isto vem do seguinte fato: Uma vez que $\sigma_i^2 = \sigma^2 x_i^{\lambda}$, o logaritmo em ambos os lados produz $\log \sigma_i^2 = \log \sigma^2 + \lambda \log x_i$. Somando $\log \widehat{e}_i^2$ nos dois lados da igualdade, tem-se $\log \widehat{e}_i^2 + \log \sigma_i^2 = \log \sigma^2 + \lambda \log x_i + \log \widehat{e}_i^2$. E, portanto,

$$\begin{aligned} \log \widehat{e}_i^2 &= \log \sigma^2 + \lambda \log x_i + \log \widehat{e}_i^2 - \log \sigma_i^2 \\ &= \log \sigma^2 + \lambda \log x_i + \log \left(\frac{\widehat{e}_i^2}{\sigma_i^2} \right). \end{aligned}$$

Isto é,

$$\log \left(\widehat{e}_i^2 \right) = \mathbf{z}_i' \alpha + w_i. \tag{17}$$

onde $w_i = \log \left(\frac{\widehat{e}_i^2}{\sigma_i^2} \right)$ e $\widehat{e}_i^2$ é o *i-ésimo* resíduo resultado da utilização do EMQO.

Considerando o logaritmo do quadrado do resíduo como uma variável dependente e o logaritmo da variável regressora, que está relacionado com a variância dos erros, como variável regressora, ajusta-se o modelo (17) para estimar λ e $log(\sigma^2)$. A partir dessas estimativas, é possível utilizar o EMQG para estimar os parâmetros do modelo (1).

Utilizando o EMQO para estimar os parâmetros do modelo (17), pode-se dizer que o estimador em 2 passos de Park de α é

$$\widetilde{\alpha} = \left(\sum_{i=1}^{n} \mathbf{z}_i \mathbf{z}_i' \right)^{-1} \sum_{i=1}^{n} \mathbf{z}_i \log \widehat{e}_i^2. \tag{18}$$

Park (1966) mostrou que, assintoticamente, w_i converge em distribuição para uma variável aleatória w_i^* que possui uma distribuição logaritmo de uma variável χ^2 com um grau de liberdade.

Para facilitar o cálculo dos momentos desta distribuição, Abramowitz e Stegum (1965) mostraram que o valor esperado de uma variável, w, distribuída como um logaritmo de uma χ^2 dividida por seus graus de liberdade, v, é igual a $E\left(w/v\right) =$

$\Psi\left(\frac{v}{2}\right) - \log\left(\frac{v}{2}\right)$. $\Psi(s)$ é a função digama definida por $\frac{d\left(\log\Gamma\left(s\right)\right)}{ds} = \frac{\Gamma'(s)}{\Gamma(s)}$ e $\Gamma(s)$ é a função gama.

O segundo momento centrado na média é dado por $\Psi^{(1)}\left(v/2\right)$, pois $\Psi^{(m)}(s)$ denota a *m-ésima* derivada de $\Psi(s)$. Com isso, verifica-se que, para 1 grau de liberdade, a média e a variância são dadas respectivamente por $E\left(w_i^*\right) = \Psi\left(1/2\right) - \log(1/2) = -1,2704$ e $Var(w_i^*) = \Psi^{(1)}(1/2) = 4,9348$.

Supondo que $\widetilde{\alpha}_1$ representa o primeiro elemento do vetor $\widetilde{\alpha}$, e sabendo que $E\left(w_i^*\right) = -1,2704$; pode-se dizer, então, que $\widetilde{\alpha}_1$ é um estimador inconsistente de α_1. Entretanto, os outros $p-1$ elementos de $\widetilde{\alpha}$ são estimadores consistentes de seus correspondentes parâmetros em α. Isto porque, considerando que $n \to \infty$, $w_i^* \sim N(-1,2704;4,9348)$ e conseqüentemente, por ser combinações de normais, $\widehat{\log\left(\widehat{e}_i^2\right)} \sim N(\mathbf{z}_i'\alpha - 1,2704;4,9348)$. Visto que α_1 simplesmente introduz um fator de proporcionalidade na variância dos resíduos, o possível EMQG de β no modelo (1), utilizando o estimador da matriz de variância e covariância dos erros proposto por Park (1966), em geral, será assintoticamente eficiente.

Sabendo que a variância dos resíduos é igual a $4,9348$, então a matriz de variâncias-covariâncias de $\widetilde{\alpha}$ é dada por $Cov\left(\widetilde{\alpha}\right) = 4,9348\left[\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i'\right]^{-1}$.

Assim, considerando $\widetilde{\alpha}_1 + 1,2704$ como o estimador do primeiro elemento do vetor $\widetilde{\alpha}$, pode-se dizer que o estimador corrigido por $1,2704$ é um estimador consistente.

3.2.2 O estimador em 3 passos de Harvey (1976)

Supondo que os erros do modelo apresentado em (1) são independentes e normalmente distribuídos com média zero e variância σ_i^2, o log da função de verossimilhança pode ser escrito como

$$\log\left(L\left(\beta,\alpha\right)\right) = -\frac{n}{2}\log\left(2\pi\right) - \frac{\sum\mathbf{z}_i'\alpha}{2} - \frac{1}{2}\sum_{i=1}^{n}\exp\left(-\mathbf{z}_i'\alpha\right)\left(y_i - \mathbf{x}_i'\beta\right)^2. \qquad (19)$$

O método iterativo mais utilizado para ajustar modelos da forma descrita em (19) é o chamado método de Scoring de Fisher (Harvey, 1976). O método de Scoring, baseado no processo iterativo de Newton-Raphson, pode ser usado para maximizar a função de verossimilhança mais facilmente, e os estimadores obtidos por este processo estarão próximos dos estimadores de máxima verossimilhança.

Sob condições de regularidade, o estimador de máxima verossimilhança de (β,α) segue, para grandes amostras, distribuição normal com média (β,α) e matriz de variâncias e covariâncias igual ao inverso da matriz de informação de Fisher. A matriz de informação é o valor negativo do valor esperado da matriz de derivadas de segunda ordem do logaritmo da função de verossimilhança.

A inversa da matriz de informação de Fisher que, assintoticamente, é a matriz de variâncias e covariâncias dos estimadores de máxima verossimilhança, pode ser escrita por

$$I(\beta,\alpha)^{-1} = \begin{bmatrix} \left(\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\mathbf{x}_i\mathbf{x}_i'\right)^{-1} & 0 \\ 0 & 2\left(\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i'\right)^{-1} \end{bmatrix}. \tag{20}$$

Dado que $E\left[\dfrac{\partial^2 \log L}{\partial\alpha\partial\beta'}\right] = E\left[\dfrac{\partial^2 \log L}{\partial\beta\partial\alpha'}\right] = 0$, a matriz de informação de Fisher é uma matriz bloco diagonal, em que um dos blocos, I_β, corresponde à matriz de informação de β e o outro, I_α, à matriz de informação de α. Isto significa que os parâmetros β e α são globalmente ortogonais e suas estimativas de máxima verossimilhança, $\widehat{\beta}$ e $\widehat{\alpha}$, assintoticamente independentes.

Se $\widehat{\alpha}^{(t)}$ e $\widehat{\beta}^{(t)}$ são as estimativas de α e β respectivamente, obtidos na *t-ésima* iteração do método de Scoring, as equações iterativas para estimar β e α podem, de acordo com Rutemiller e Bowers (1968), ser escritas separadamente da seguinte forma

$$\widehat{\beta}^{(t+1)} = \widehat{\beta}^{(t)} + \left(\sum_{i=1}^{n} e^{-\mathbf{z}_i'\widehat{\alpha}^{(t)}}\mathbf{x}_i\mathbf{x}_i'\right)^{-1}\sum_{i=1}^{n}\mathbf{x}_i e^{-\mathbf{z}_i'\widehat{\alpha}^{(t)}}\left(y_i - \mathbf{x}_i'\widehat{\beta}^{(t)}\right) \tag{21}$$

$$\widehat{\alpha}^{(t+1)} = \widehat{\alpha}^{(t)} + \left(\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\sum_{i=1}^{n}\mathbf{z}_i\left[e^{-\mathbf{z}_i'\widehat{\alpha}^{(t)}}\left(y_i - \mathbf{x}_i'\widehat{\beta}^{(t)}\right)^2 - 1\right]. \tag{22}$$

O valor inicial $\widehat{\beta}^{(0)}$ pode ser igual ao EMQO $(\widehat{\beta})$, e, segundo Rutemiller e Bowers (1968), $\widehat{\alpha}_1^{(0)}$ pode ser igual a $logS^2$, onde $S^2 = n^{-1}\sum\left(y_i - \mathbf{x}_i'\widehat{\beta}\right)^2$ e todos os outros elementos do vetor $\widehat{\alpha}^{(0)}$ serão iguais a zero. Uma outra alternativa pode ser usar $\alpha^{(0)} = \widetilde{\alpha}$, ou seja, o estimador em 2 passos definido por Park (1966).

Se um estimador consistente para α é usado como valor inicial do processo iterativo definido acima, somente uma iteração será necessária para produzir um estimador com as mesmas propriedades assintóticas do estimador de máxima verossimilhança. Harvey (1976) mostrou que o estimador de Park (1966) tem um elemento inconsistente e, por isso, é assintoticamente ineficiente. Por essa razão, Harvey (1976) propôs um procedimento de estimação em 3 passos, o qual possui, assintoticamente, as mesmas propriedades que o estimador de máxima verossimilhança.

Segundo Harvey, se α_1, do estimador de Park (1966), é estimado por $\widetilde{\alpha}_1 + 1,2407$, o procedimento em 2 passos de Park produz um estimador consistente de

α e a expressão (22) pode ser simplificada da seguinte forma

$$\widehat{\alpha} = \widetilde{\alpha} + \begin{pmatrix} 0,2704 \\ 0 \end{pmatrix} + 0,2807 \left(\sum_{i=1}^{n} \mathbf{z}_i \mathbf{z}_i' \right)^{-1} \sum_{i=1}^{n} \mathbf{z}_i \exp\left(-\mathbf{z}_i' \widetilde{\alpha}\right) \widehat{e}_i^2, \tag{23}$$

onde $\widetilde{\alpha} = \left(\sum_{i=1}^{n} \mathbf{z}_i \mathbf{z}_i' \right)^{-1} \sum_{i=1}^{n} \mathbf{z}_i \log\left(\widehat{e}_i^2\right)$.

Este estimador foi chamado por Harvey (1976) de estimador em 3 passos do grau da heteroscedasticidade de um modelo de regressão linear e é um estimador assintoticamente eficiente. Se o EMQG for usado para estimar β, considerando o estimador em 3 passos de Harvey para estimar a matriz de variâncias e covariâncias dos erros, é razoável supor que esse estimador terá propriedades melhores que o correspondente estimador em 2 passos de Park. Neste caso o EMQG de β é dado por

$$\widehat{\beta} = \left(\sum_{i=1}^{n} e^{-\mathbf{z}_i' \widehat{\alpha}} \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \sum_{i=1}^{n} e^{-\mathbf{z}_i' \widehat{\alpha}} \mathbf{x}_i y_i. \tag{24}$$

3.3 O Método de otimização global (Algoritmo BFGS)

Outra forma de se obter estimativas dos parâmetros na situação em que o modelo é heteroscedástico é por meio dos algoritmos iterativos de obtenção de máximos globais de uma função. Segundo Seber e Wild (1988), quando o número de parâmetros do modelo é grande, os algoritmos mais eficientes para maximizar uma função, $h(\theta)$, são os algoritmos de maximização global baseados no clássico método iterativo de Newton em que $\theta^{(t+1)} = \theta^{(t)} + \delta^{(t)}$, $\delta^{(t)} = -H\left(\theta^{(t)}\right)^{-1} g\left(\theta^{(t)}\right)$, onde $g(\theta) = \dfrac{\partial h(\theta)}{\partial \theta}$ é o vetor gradiente e $H(\theta) = \dfrac{\partial^2 h(\theta)}{\partial \theta \partial \theta'}$ é a matriz de derivadas segundas, também chamada de matriz Hessiana. Quando é difícil se obter a expressão analítica de $H(\theta)$, é possível aproximá-la por meio de um algoritmo usando diferenças finitas em $g(\theta)$, assim como $g(\theta)$ pode ser aproximada usando diferenças finitas em $h(\theta)$.

Algumas modificações no método de Newton foram implementadas para melhorar a convergência do algoritmo e em 1970 foi publicado simultaneamente por Broyden (1970), Fletcher (1970), Goldfarb (1970) e Shanno (1970) um algoritmo, conhecido hoje como BFGS, baseado no método iterativo de Quasi-Newton. Este algoritmo é bastante utilizado para encontrar o máximo ou o mínimo global de uma função real de p variáveis reais $\theta = (\theta_1, \theta_2, ..., \theta_p)'$.

4 Discussão

Na teoria clássica de modelos de regressão, uma das suposições básicas para garantir a eficiência dos estimadores padrões convencionais é a suposição de que as variâncias dos erros são iguais, ou seja, os erros são homoscedásticos.

Quando isso não ocorre, utilizar o EMQO do vetor de parâmetros β acarretará conseqüências na eficiência do estimador, pois apesar de ser um estimador não viciado, sua variância não é mínima. Além disso, o estimador de sua variância é um estimador viciado.

As alternativas clássicas para esse problema seriam utilizar o EMQG, estimando a matriz de variâncias e covariâncias a partir do estimador desenvolvido por White (1980) e aperfeiçoado por Cribari Neto; Silva e Cordeiro (2000), para amostras menores; utilizar o método de Scoring, mais precisamente o estimador de Harvey (1976 e 1974), o qual considera a estrutura heteroscedástica na estimação dos parâmetros do modelo; ou ainda, utilizar um método de otimização global, quando a verossimilhança é completamente conhecida.

No entanto, encontrar a solução ideal não é nada fácil devido às várias dificuldades práticas. Nestas situações, estudos mais aprofundados sobre cada método de estimação devem ser considerados a fim de utilizar o método mais apropriado com menor custo computacional para seu cálculo.

RORIGUES, S. A., DINIZ, C. A. R. Heteroscedastic regression model. *Rev. Mat. Estat.*, São Paulo, v.24, n.2, p.133-146, 2006.

▪ *ABSTRACT: One of the basic assumptions in the regression theory is that the errors should have a constant variance. When the violation of this assumption occurs, an option is to transform the data or to reestimate the variance structure of the data. This work introduces the linear heteroscedastic regression model as well as the estimate and detection of heteroscedasticity.*

▪ *KEYWORDS: Heteroscedasticity; estimation.*

Referências

ABRAMOWITZ, M.; STEGUM, I. A. *Handbook of mathematical functions.* New York: Dover, 1965. 1046p.

BOX, G. E. P.; COX, D. R. An analysis of transformation (with discussion). *J. R. Stat. Soc. Ser. B*, London, v. 26, n.2, p.211-252, 1964.

BROYDEN, C. G. The convergence of a class of double-rank minimization algorithms. Parts I and II, *J. Inst. Math. Appl.*, Malden n.6, p.76-90 e p.222-231, 1970.

CRIBARI NETO, F.; SILVA, L. P. F.; CORDEIRO, G. M., Improved Heteroscedasticity-consistent matrix estimators. *Biometrika*, London, v.87, n.4, p.907-918, 2000.

FLETCHER, R. A new approach to variable metric algorithms. *Comput. J.*, Oxford, v.13, p.317-322, 1970.

GEARY, R. C. A note on residual heterovariance and estimation efficiency in regression. *Am. Stat.*, Washington, v.20, n.1 p.30-31, 1966.

GLEJSER, H. A new test for heteroscedasticity. *J. Am. Stat. Assoc.*, New York, v.6, p.316-323, 1969.

GRIFFITHS, W. E.; SUREKHA, K. A Monte Carlo comparison of some Bayesian and sampling theory estimators in two heteroscedastic error models, *Commun. Stat. Simul. Comput.*, New York, v.1, n.13, p.85-105, 1984.

GOLDFARB, D. A family of variable metric methods derived by variational means. *Math. Comp.*, Boston, v.26, p.23-26, 1970.

HARVEY, A. C. Estimation of parameters in a heteroscedastic regression model. In: EUROPEAN MEETING OF THE ECONOMETRIC SOCIETY, 1974, Grenoble. 1974.

HARVEY, A. C. Estimating regression models with multiplicative heteroscedasticity. *Econometrica*, Malden, v.44, p.461-465, 1976.

JUDGE, G. G. et al. *The theory and practice of econometrics.* New York: John Wiley, 1985. 1056p.

KMENTA, J. *Elements of econometrics.* New York: The Macmillan, 1971. 410p.

LANCASTER, T. Grouping estimators on heteroscedastic data. *J. Am. Stat. Assoc.*, New York, v.63, p.182-191, 1968.

PARK, R. Estimation with heteroscedastic error terms. *Econometrica*, Malden, v.34, p.888-896, 1966.

RUTEMILLER, H. C.; BOWERS, D. A. Estimation in a heteroscedastic regression model. *J. Am. Stat. Assoc.*, New York, v.63, p.552-557, 1968.

SEBER, G. A. F.; WILD C. J. *Nonlinear regression.* New York: John Wiley, 1988. 618p.

SHANNO, D, F. Conditioning of quasi-Newton methods for function minimization. *Math. Comp.*, Boston, v.24, p. 647-657, 1970.

SZROETER, J. A class of parametric test for heteroscedasticity in linear econometric models. *Econometrica*, Malden, v.46, p.1311-1327, 1978.

WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, Malden, v.48, p.817-838, 1980.

Recebido em 25.10.2005.

Aprovado após revisão em 19.06.2006.