VARIÁVEIS "DUMMIES" EM REGRESSÃO; UMA CONSIDERAÇÃO METODOLÓGICA

Pedro Luiz Valls Pereira*

RESUMO

O objetivo deste artigo é apresentar uma formulação alternativa do estudo de "dummies" em modelos lineares. São apresentados três métodos equivalentes de estimação. O teste de Chow (teste de estabilidade dos coeficientes de regressão) é obtido através da estatística de Wald. Estende-se este teste para modelos com erros heteroscedásticos ou correlacionados e para modelos dinâmicos.

ABSTRACT

This article presents an alternative method to deal with dummies in regression models. Three equivalent methods of estimation are presented. The Chow test (test for structural change) is obtained from the Wald statistic. An extension to regression with heteroscedastic or autocorrelated errors as well as dynamic models is presented.

* Pesquisador Assistente do Instituto de Matemática Pura e Aplicada (IMPA). Agradeço a Rubem Klein e Djalma Pessoa e a um "referee" pelos comentários.

1. INTRODUÇÃO

Variáveis "dummies" são amplamente utilizadas em econometria (ver Theil (1971), Johnston (1972), Maddala (1977) entre outros).

Este artigo apresenta um outro tipo de aplicação destas variáveis. Apresenta-se uma outra forma de estimação destes modelos onde a relação entre modelos lineares com "dummies" e análise de con variância é explicitada (para um tratamento semelhante veja Capítulo 6 de Johnston (1972)).

Apresentaremos também o teste de estabilidade dos coeficientes de regressão, devido a Chow (ver Chow (1960) ou Fisher (1970)). Mostraremos que tal teste é válido mesmo para modelos dinâmicos ou mesmo quando as hipóteses de homoscedasticidade ou independência serial não são satisfeitas (o que estende os resultados existentes na literatura).

2. ESTIMAÇÃO DE MODELOS LINEARES COM VARIÁVEIS "DUMMIES"

Variáveis "Dummies" são utilizadas quando os dados são divididos em diversas categorias. O exemplo clássico é "guerra x paz" onde assume-se que o comportamento nos dois regimes é diferente, i.e., parâmetros de regressão mudam nos dois segmentos dos dados.

A utilização de variáveis "dummies" para estimar o termo constante em diferentes subpopulações é um caso especial das aplicações gerais destas variáveis.

Considere o modelo

$$\underline{y} = X\underline{\beta} + \underline{u} \quad (2.1)$$

onde $\underline{y}$ é um vetor $T \times 1$, $X = (\underline{x}_1 : \dots : \underline{x}_n) = \begin{pmatrix} \underline{x}_1' \\ \vdots \\ \underline{x}_T' \end{pmatrix}$ é uma matriz $T \times n$ de variáveis não estocásticas, $\underline{u}$ é $N(0, \sigma^2 I_T)$ e $\underline{\beta}$ um vetor $n \times 1$.

Os dados de (2.1) podem ser de séries temporais ou de "cross-section". Se tais dados não são homogêneos nos períodos de tempo ou de "cross-section", podemos introduzir "dummies" para incorporar tais diferenças, i.e., diferenças entre sub amostras.

Seja então S a amostra total e S_i a i -ésima amostra,

$i = 1, \dots, n_1$ onde S_i tem T_i elementos com $T = \sum_{i=1}^{n_1} T_i$.

Introduz-se então as variáveis "dummies" pela definição

$$q_{ti} = \begin{cases} 1 & t \in S_i \\ 0 & 0 \notin S_i \end{cases}$$

A matriz de "dummies" é dada por

$$D = \begin{matrix} & \left\{ \begin{matrix} T_1 \\ \\ T_2 \\ \\ \vdots \\ \\ T_{n_1} \end{matrix} \right\} & \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & 0 & & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & & \vdots & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} & = [q_1 \dots q_{n_1}] = \begin{bmatrix} q_1' \\ \vdots \\ q_T' \end{bmatrix} \end{matrix} \quad (2.2)$$

onde D é uma matriz $T \times n_1$ com posto n_1 .

Em alguns casos deseja-se medir a variação entre a i -ésima e a k -ésima subamostra (por exemplo variações sazonais veja Phillips & Wickens (1978) exercício 4.6). Neste caso, e assumindo $i = 1$ e $k = n_1$, a matriz D é expressa da seguinte forma alternativa:

$$D^* = \begin{matrix} T_1 \left\{ \begin{bmatrix} 1 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & & 0 \\ \dots & \dots & \dots & \dots \end{bmatrix} \right. \\ T_2 \left\{ \begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & & 0 \\ \dots & \dots & \dots & \dots \end{bmatrix} \right. \\ \vdots \\ T_{n_1-1} \left\{ \begin{bmatrix} 1 & 0 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & & 1 \\ \dots & \dots & \dots & \dots \end{bmatrix} \right. \\ T_{n_1} \left\{ \begin{bmatrix} 1 & -1 & \dots & -1 \\ \vdots & \vdots & & \vdots \\ 1 & -1 & & -1 \end{bmatrix} \right. \end{matrix} = [q_1^*, \dots, q_{n_1}^*] \quad (2.2')$$

O modelo (2.1) pode ser escrito de duas formas

$$y_t = \sum_{i=1}^{n_1} q_{ti} \delta_i + X'_t \beta + u_t$$

$$\Leftrightarrow y_t = \delta_i + X'_t \beta + u_t \text{ se } t \in S_i \text{ } i=1, \dots, n_1 \quad (2.3)$$

onde o termo constante não aparece explicitamente na equação ou

$$y_t = \alpha + \sum_{i=1}^{n_1} q_{ti} \delta_i + X'_t \beta + u_t$$

$$\Leftrightarrow y_t = \alpha + \delta_i + X'_t \beta + u_t \quad \text{se } t \in S_i \quad i=1, \dots, n_1 \quad (2.4)$$

onde o termo constante é explicitado.

Pela definição das "dummies" tem-se as seguintes propriedades

$$(i) \sum_{i=1}^{n_1} q_{it} = 1$$

(ii) (i) $\Rightarrow$ multicolinearidade se (2.4) for utilizado

$$(iii) (ii) \Rightarrow y_t = \alpha + \sum_{i=2}^{n_1} q_{ti}^* \delta_i^* + X'_t \beta + u_t \quad (2.4') (*)$$

$$(iv) D'D = \begin{bmatrix} T_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & T_{n_1} \end{bmatrix} \quad (2.5) \text{ matriz } n_1 \times n_1 \text{ com posto total}$$

$$(v) D^* ' D^* = \begin{bmatrix} T & 0 & 0 & 0 & \dots & 0 \\ 0 & T_1 + T_{n_1} & T_{n_1} & T_{n_1} & \dots & T_{n_1} \\ 0 & T_{n_1} & T_2 + T_{n_1} & T_{n_1} & \dots & T_{n_1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & T_{n_1} & \dots & T_{n_1} & \dots & T_{n_1-1} + T_{n_1} \end{bmatrix} \quad (2.5')$$

(*) Com esta formulação o problema de multicolinearidade é contornado, já que as colunas de D^* , embora incluam o termo constante, são linearmente independentes.

matriz $n_1 \times n_1$ com posto total. É fácil ver que $D(D'D)^{-1}D' = D^*(D^*D^*)^{-1}D^*$, já que D pode ser escrito como

$$D^* \begin{bmatrix} n_1 & +1 \\ \text{---} & \\ \text{---} & \\ \text{---} & \\ j=1 & E_j \end{bmatrix} \quad \text{onde } E_j \text{ são matrizes elementares.}$$

Com a introdução das "dummies" o modelo (2.1) pode ser escrito como

$$y = X_1\beta_1 + X_2\beta_2 + u = X\beta + u \quad (2.6)$$

onde X_1 é a matriz $T \times n_1$ das "dummies" e X_2 a matriz $T \times n_2$ dos regressores onde $n_1 + n_2 = n$.

Deseja-se obter os estimadores de β_1 e β_2 . Abaixo apresentaremos três procedimentos. Todos eles estimam β_2 e como subproduto obtém-se o estimador de β_1 .

Primeiro método: Mínimos Quadrados Ordinários

As equações normais para (2.6) são

$$X_1'Y = X_1'X_1\beta_1 + X_1'X_2\beta_2 \quad (2.7)$$

$$X_2'Y = X_2'X_1\beta_1 + X_2'X_2\beta_2 \quad (2.8)$$

Como o posto $(X_1'X_1) = n_1$, posto $(X_2'X_2) = n_2$ e

$Q_1 = I_T - X_1(X_1'X_1)^{-1}X_1'$ é idempotente com posto $(Q_1) = T - n_1$, se

multiplicarmos (2.7) por $-(X_2'X_1)(X_1'X_1)^{-1}$ e somarmos a (2.8) obtemos

remos

$$X_2'Y - (X_2'X_1)(X_1'X_1)^{-1}X_1'Y = (X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)\beta_2 \quad (2.9)$$

$$\Leftrightarrow X_2'Q_1Y = (X_2'Q_1X_2)\beta_2 \quad (2.10)$$

portanto

$$\hat{\beta}_2 = (X_2' Q_1 X_2)^{-1} X_2' Q_1 \underline{y} \quad (2.11)^*$$

e

$$\hat{\beta}_1 = (X_1' X_1)^{-1} X_1' \underline{y} - (X_1' X_1)^{-1} X_1' X_2 \hat{\beta}_2. \quad (2.12)$$

Observe que aplicando-se diretamente M.Q.O. em (2.6) obtém-se

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} X_1' X_1 & X_1' X_2 \\ X_2' X_1 & X_2' X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1' \underline{y} \\ X_2' \underline{y} \end{pmatrix} \quad (2.13)$$

e usando-se inversa particionada obtém-se (2.11) e (2.12).

Segundo Método: Mínimos Quadrados Generalizados.

Multiplicando-se (2.6) por Q_1 tem-se

$$Q_1 \underline{y} = Q_1 X_1 \beta_1 + Q_1 X_2 \beta_2 + Q_1 \underline{u}. \quad (2.14)$$

É fácil ver que $Q_1 X_1 = 0$ então (2.14) se reduz a

$$Q_1 \underline{y} = Q_1 X_2 \beta_2 + Q_1 \underline{u}. \quad (2.15)$$

Agora como $\underline{u} \sim N(0, \sigma^2 I_T)$ então $Q_1 \underline{u} \sim N(0, \sigma^2 Q_1)$.

Então como Q_1 é positiva definida pode-se usar a decomposição espectral de Q_1 (**) e aplicar M.Q.O. no modelo transformado (que é equivalente a M.Q.G. em (2.15)).

(*) Para que $X_2' Q_1 X_2$ seja inversível necessitamos da hipótese $T - n_1 \geq n_2$.

(**) A decomposição espectral de Q_1 é dada por $H' Q_1 H = \Lambda^2$ onde $H H' = I_{n_1}$ e Λ^2 é diagonal.

Observe que para a aplicação de M.Q.G., a matriz de covariância dos erros tem que ser não singular. Isto não acontece neste caso já que Q_1 tem posto $T-n_1$. Este caso corresponde à seção 6.7 de Theil (1971). Devemos então verificar se as duas condições de posto são satisfeitas (veja Theil (1971)).

1ª condição é satisfeita porque $H'Q_1X_2$ tem posto n_2 onde H é tal que $H'Q_1H = \Lambda^2$ e $HH' = I$.

2ª condição de posto é satisfeita já que $G'X_1X_2 = 0$ onde G' é a matriz de autovetores associados ao autovalor zero de Q_1 i.e., $G'Q_1 = Q_1G' = 0$.

O modelo transformado de (2.15) é dado por

$$\Lambda^{-1}H'Q_1\underline{y} = (\Lambda^{-1}H'Q_1X_2)\underline{\beta}_2 + \Lambda^{-1}H'Q_1\underline{u} \quad (2.16)$$

É fácil ver que se pode aplicar M.Q.O. em (2.16) e obtém-se

$$\begin{aligned} \underline{\beta}_2 &= (X_2'Q_1H\Lambda^{-2}H'Q_1X_2)^{-1}(X_2'Q_1H\Lambda^{-2}H'Q_1)\underline{y} \\ &= (X_2'Q_1Q_1^+Q_1X_2)^{-1}(X_2'Q_1Q_1^+Q_1)\underline{y} \end{aligned} \quad (2.17)$$

onde Q_1^+ é uma pseudo inversa de Q_1 e pelas propriedades de pseudo inversa tem-se $Q_1Q_1^+Q_1 = Q_1$, então (2.17) fica

$$\underline{\tilde{\beta}}_2 = (X_2'Q_1X_2)^{-1}(X_2'Q_1)\underline{y}. \quad (2.18)$$

Analogamente ao que foi feito no primeiro método o estimador de $\underline{\tilde{\beta}}_1$ é dado pela expressão (2.12).

Terceiro Método: Procedimento em Dois Estágios.

Seja

$$\underline{y}^* = Q_1\underline{y} \quad (2.19)$$

$$\underline{x}_2^* = Q_1X_2 \quad (2.20)$$

$$\begin{aligned} \text{Observe que (2.19) } \Leftrightarrow \underline{y}^* &= \underline{y} - X_1 (X_1' X_1)^{-1} X_1' \underline{y}, \\ &= \underline{y} - X_1 \underline{b}_1 \end{aligned} \quad (2.21)$$

i.e., $\underline{y}^*$ é o resíduo de regressão de $\underline{y}$ em X_1 .

Tem-se também

$$\begin{aligned} (2.20) \Leftrightarrow X_2^* &= X_2 - X_1 (X_1' X_1)^{-1} X_1' X_2 \\ &= X_2 - X_1 \underline{B}_2 \end{aligned} \quad (2.22)$$

i.e., X_2^* é o resíduo de regressão de X_2 em X_1 .

Portanto, β_2 pode ser estimado pelo seguinte procedimento em dois estágios

- (1) rodar a regressão de $\underline{y}$ em X_1 e de X_2 em X_1 obtendo os resíduos $\hat{\underline{y}}$ e $\hat{X}_2$
- (2) rodar a regressão de $\hat{\underline{y}}$ em $\hat{X}_2$.

O estimador obtido em (2) será $\tilde{\beta}_2$. Para se obter $\tilde{\beta}_1$ utiliza-se a primeira equação normal i.e.,

$$X_1' \underline{y} = (X_1' X_1) \tilde{\beta}_1 + (X_1' X_2) \tilde{\beta}_2$$

então

$$\tilde{\beta}_1 = (X_1' X_1)^{-1} X_1' \underline{y} + (X_1' X_1)^{-1} X_1' X_2 \tilde{\beta}_2$$

que é exatamente (2.12).

Falta mostrar:

(i) que os resíduos do segundo estágio são os mesmos que os resíduos dos métodos anteriores.

(ii) $\tilde{\beta}_2$ obtido no segundo estágio é equivalente aos procedimentos anteriores.

Mostraremos também que $\hat{\beta}_2$ pode ser calculado usando-se variáveis desviadas da média das variáveis "dummies".

Para mostrar (i)

$$\begin{aligned}\hat{u} &= \underline{y} - \underline{x}_1 \hat{\beta}_1 - \underline{x}_2 \hat{\beta}_2 \\ &= \underline{y} - \underline{x}_1 (\underline{x}'_1 \underline{x}_1)^{-1} \underline{x}'_1 \underline{y} + \underline{x}_1 (\underline{x}'_1 \underline{x}_1)^{-1} \underline{x}'_1 \underline{x}_2 \hat{\beta}_2 - \underline{x}_2 \hat{\beta}_2 \quad (\text{por} \\ (2.11) \text{ e } (2.12)) &= \underline{Q}_1 \underline{y} - \underline{Q}_1 \underline{x}_2 \hat{\beta}_2 \\ &= \underline{y}^* - \underline{x}_2^* \hat{\beta}_2\end{aligned}\quad (2.23)$$

Agora falta mostrar que $\hat{\beta}_2 = \hat{\beta}_2$ para que (2.23) seja o resíduo do segundo estágio.

Tem-se que

$$\begin{aligned}\hat{\beta}_2 &= (\underline{x}_2^* \underline{x}_2^*)^{-1} \underline{x}_2^* \underline{y}^* \\ &= (\underline{x}'_2 \underline{Q}_1 \underline{Q}_1 \underline{x}_2)^{-1} \underline{x}'_2 \underline{Q}_1 \underline{Q}_1 \underline{y} \quad \text{definição de } \underline{x}_2^* \text{ e } \underline{y}^* \\ &= (\underline{x}'_2 \underline{Q}_1 \underline{x}_2)^{-1} \underline{x}'_2 \underline{Q}_1 \underline{y} \quad \underline{Q}_1 \text{ idempotente} \\ &= \hat{\beta}_2 \quad (\text{por 2.18})\end{aligned}$$

Então (2.23) = $\hat{u}^*$.

Agora dado (2.24) tem-se que

$$\underline{x}'_1 \underline{y} = \begin{bmatrix} \underline{q}'_1 & \underline{y} \\ \vdots & \\ \underline{q}'_{n_1} & \underline{y} \end{bmatrix} = \begin{bmatrix} \sum_{t \in S_1} & y_t \\ \vdots & \\ \sum_{t \in S_{n_1}} & y_t \end{bmatrix} \quad (2.24)$$

$$(\underline{x}'_1 \underline{x}_1)^{-1} \underline{x}'_1 \underline{y} = \begin{bmatrix} 1/T_1 & \\ & \ddots \\ & & 1/T_{n_1} \end{bmatrix} \underline{x}'_1 \underline{y} = \begin{bmatrix} \sum_{t \in S_1} & y_t/T_1 \\ \vdots & \\ \sum_{t \in S_{n_1}} & y_t/T_{n_1} \end{bmatrix} \quad (2.25)$$

Observe que (2.25) é o vetor de média em cada subamostra. Tem-se também

$$\underline{y}^* = Q_1 \underline{y} = \underline{y} - X_1 (X_1' X_1)^{-1} X_1' \underline{y} = \begin{bmatrix} y_1 - \frac{\sum y_t / T_1}{S_1} \\ \vdots \\ y_{T_1} - \frac{\sum y_t / T_1}{S_1} \\ \vdots \\ y_{T-T_{n_1}+1} - \frac{\sum y_t / T_{n_1}}{S_{n_1}} \\ \vdots \\ y_T - \frac{\sum y_t / T_{n_1}}{S_{n_1}} \end{bmatrix} \quad \begin{array}{l} \text{Desvios da} \\ \text{média} \end{array} \quad (2.26)$$

Analogamente tem-se

$$\underline{x}_2^* = Q_1 \underline{x}_2 = \underline{x}_2 - X_1 (X_1' X_1)^{-1} X_1' \underline{x}_2 = \begin{bmatrix} x_1^+ - \bar{x}_{T_1}^+ \\ \vdots \\ x_{T_1}^+ - \bar{x}_{T_1}^+ \\ \vdots \\ x_{T-T_{n_1}+1}^+ - \bar{x}_{T_{n_1}}^+ \\ \vdots \\ x_T^+ - \bar{x}_{T_{n_1}}^+ \end{bmatrix} \quad \begin{array}{l} \text{Desvios da} \\ \text{média} \end{array} \quad (2.27)$$

onde $x_2' = (x_1^{+'} : \dots : x_{T_{n_1}}^{+'})$; $\bar{x}_{T_j}^+ = \frac{\sum_{t \in S_j} x_t^+ / T_j}{j} \quad j = 1, \dots, n_1$.

Por (2.26) e (2.27) tem-se que os resíduos do primeiro estágio são as variáveis desviadas da média. Portanto o primeiro estágio do procedimento pode ser calculado facilmente, i.e., dada as observações calcula-se $\underline{y}^*$, $\underline{x}_2^*$ por (2.26) e (2.27) e obtém-se $\underline{\tilde{\beta}}_2$ pelo segundo estágio.

É fácil ver que os estimadores de β_1 obtidos pelos três métodos são os mesmos. Portanto tem-se que os três métodos são equivalentes^(*).

A seguir apresentaremos uma aplicação de variáveis "dummies" para testar estabilidade dos coeficientes.

3. TESTES DE ESTABILIDADE DOS COEFICIENTES DE REGRESSÃO

Analisaremos primeiro o teste de estabilidade para o termo constante, i.e., deseja-se testar se o termo constante é homogêneo entre subamostras. Veremos que o caso geral é análogo ao que apresentaremos a seguir. Para se testar estabilidade do termo constante, deseja-se testar se $\beta_1 = 0$, i.e.

$$H_0: \beta_1 = 0 \quad \text{vs} \quad H_A: \beta_1 \neq 0 \quad (3.1)$$

Sob H_0 o modelo (2.6) toma a forma

$$\underline{y} = \underline{x}_2 \beta_2 + \underline{u} \quad (3.2)$$

e tem-se que

$$\underline{\tilde{\beta}}_2 = (\underline{x}_2' \underline{x}_2)^{-1} \underline{x}_2' \underline{y} \quad (3.3)$$

(*) O terceiro método é usado quando se deseja retirar a sazonalidade das variáveis (ver Seção 6.2 em Johnston (1972) ou exercício 4.7 de Phillips & Wickens (1978)).

e o resíduo é dado por

$$\begin{aligned}\tilde{u}_{H_0} &= \underline{y} - X_2 \tilde{\beta}_2 = (I - X_2 (X_2' X_2)^{-1} X_2') \underline{y} \\ &= Q_2 \underline{y}\end{aligned}\quad (3.4)$$

Os estimadores de $\tilde{\beta}_1$ e $\tilde{\beta}_2$, sob H_1 , foram obtidos na Seção 2. É fácil ver que (2.11) e (2.12) podem ser escritas na seguinte forma

$$\tilde{\beta}_1 = (X_1' Q_2 X_1)^{-1} X_1' Q_2 \underline{y} \quad (3.5)$$

$$\tilde{\beta}_2 = (X_2' X_2)^{-1} X_2' \underline{y} - (X_2' X_2)^{-1} X_2' X_1 \tilde{\beta}_1 \quad (3.6)$$

Então o resíduo é dado por

$$\tilde{u}_{H_A} = \underline{y} - X_1 \tilde{\beta}_1 - X_2 \tilde{\beta}_2 = [Q_2 - Q_2 X_1 (X_1' Q_2 X_1)^{-1} X_1' Q_2] \underline{y} \quad (3.7)$$

Para se testar (3.1) utilizaremos o teste da razão de verossimilhança.

Assumindo-se normalidade nos erros, o teste da razão de verossimilhança tem uma distribuição F. A estatística do teste é dada por

$$\frac{(SQR_{H_0} - SQR_{H_A}) / gl(H_0 - H_A)}{SQR_{H_A} / gl(H_A)} \sim F(gl(H_0 - H_A), gl(H_A)) \quad (3.8) (*)$$

onde $gl(H_A)$ denota graus de liberdade sob H_A , que corresponde ao traço da matriz da forma quadrática SQR_{H_A} (soma dos quadrados dos resíduos em H_A).

Observe que

$$SQR_{H_0} - SQR_{H_A} = \underline{y}' Q_2 X_1 (X_1' Q_2 X_1)^{-1} X_1' Q_2 \underline{y} \quad (3.9) (**)$$

e

(*) Para se obter (3.8) utiliza-se que o numerador e denominador desta expressão são distribuídos como χ^2 independentes.

(**) $gl(H_0 - H_A) = \text{tr}(Q_2 X_1 (X_1' Q_2 X_1)^{-1} X_1' Q_2) = \text{tr}(X_1' Q_2 X_1 (X_1' Q_2 X_1)^{-1}) = \text{tr}(I_{n_1}) = n_1$

e

$$SQR_{H_A} = \underline{y}' (Q_2 - Q_2 X_1 (X_1' Q_2 X_1)^{-1} X_1' Q_2) \underline{y} = \underline{\tilde{u}}_{H_A}' \underline{\tilde{u}}_{H_A} \quad (3.10) (*)$$

$$\text{Portanto (3.8)} \Rightarrow \frac{\underline{y}' Q_2 X_1 (X_1' Q_2 X_1)^{-1} X_1' Q_2 \underline{y}}{n_1 (\underline{\tilde{u}}_{H_A}' \underline{\tilde{u}}_{H_A} / T - n_1 - n_2)} \quad (3.8')$$

Observe que (3.9) pode ser escrita da seguinte forma

$$SQR_{H_0} - SQR_{H_A} = \underline{\tilde{\beta}}_1' (X_1' Q_2 X_1) \underline{\tilde{\beta}}_1 \quad (3.11)$$

usando-se (3.5). Como $\frac{SQR_{H_0} - SQR_{H_A}}{\sigma^2} = \chi_{n_1}^2$ tem-se que

$$\frac{\underline{\tilde{\beta}}_1' (X_1' Q_2 X_1) \underline{\tilde{\beta}}_1}{\sigma^2} = \frac{\underline{\tilde{\beta}}_1' (X_1' Q_2 X_1) \underline{\tilde{\beta}}_1}{\sigma^2} - \chi_{n_1}^2. \quad (3.12)$$

A expressão seria a estatística para o teste (3.1) se σ^2 fosse conhecida.

Observe que o teste (3.1) é um caso particular do teste de restrições lineares nos coeficientes (veja Maddala (1977) p. 457-458) já que $\underline{\beta}_1 = \underline{0} \Leftrightarrow R\underline{\beta} = \underline{r}$ onde $R = [I_{n_1} ; 0]$ é uma matriz $n_1 \times n$, $\underline{\beta} = (\underline{\beta}_1, \underline{\beta}_2)$ é um vetor $n \times 1$ e $\underline{r}$ é um vetor $n_1 \times 1$ que neste caso é um vetor de zeros.

É fácil mostrar que (3.12) pode ser escrita da seguinte forma alternativa

$$K_1 = \frac{(\underline{R}\underline{\tilde{\beta}} - \underline{r})' [R(X'X)^{-1}R']^{-1} (\underline{R}\underline{\tilde{\beta}} - \underline{r})}{\sigma^2} \quad (3.13) (**)$$

já que $(\underline{R}\underline{\tilde{\beta}} - \underline{r}) = \underline{\tilde{\beta}}_1$ e $R(X'X)^{-1}R' = (X_1' Q_2 X_1)^{-1}$.

(*) $gl(H_A) = \text{tr}(Q_2 - Q_2 X_1 (X_1' Q_2 X_1)^{-1} X_1' Q_2) = \text{tr}(Q_2) - gl(H_0 - H_A) = T - n_2 - n_1$.

(**) Observe que $R(X'X)^{-1}R'$ é uma matriz de variância da distribuição de $\underline{\tilde{\beta}}_1$.

A expressão (3.13) é equivalente à (3.12) e portanto tem distribuição $\chi^2_{n_1}$ onde n_1 é o número de restrições nos coeficientes.

Se σ^2 for desconhecido, podemos utilizar um estimador, por exemplo $\hat{s}^2 = \frac{\hat{u}'_H \hat{u}_H}{T - n_1 - n_2}$ nas expressões (3.12) (ou (3.13)).

Neste caso

$$K_2 = \frac{\hat{\beta}'_1 (X'_1 Q_2 X_1) \hat{\beta}_1}{\hat{s}^2} \quad a \quad \chi^2_{n_1} \quad (3.14)$$

onde a significa distribuído assintoticamente.

Observe que (3.8') = K_2/n_1 . Portanto $\frac{K_2}{n_1}$ tem distribuição F em pequenas amostras.

Observe que para se testar (3.1) utilizando-se o teste da razão de verossimilhança, necessitou-se estimar os dois modelos, i.e. restrito (sob H_0) e irrestrito (sob H_A). Note também que o teste obtido é equivalente a $\frac{K_2}{n_1}$, onde só é necessário a estimação do modelo irrestrito (sob H_A). Este teste é conhecido como teste de Wald (veja Buse (1982)). Este último teste será usado nas seções seguintes para se obter o teste (3.1) quando $E(\underline{u}\underline{u}') \neq \sigma^2 I_T$.

3.1 Regressão com Erros Heteroscedásticos ou Correlacionados Serialmente

Para o teste (3.1) assumiu-se que $\underline{u} \sim N(0, \sigma^2 I_T)$. Suponha que $E(\underline{u}\underline{u}') = \sigma^2 \Omega$. O estimador do modelo (2.6) é dado por (usando-se M.Q.G.)

$$\hat{\underline{\beta}} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} \underline{y}. \quad (3.1.1)$$

Observe que o estimador de β_1 é dado por

$$\hat{\beta}_1 = (X'_1 Q_2 X_1)^{-1} X'_1 Q_2 \underline{\tilde{y}} \quad (3.1.2)$$

onde

$$\tilde{x}_i = \Omega^{-1/2} x'_i \quad i = 1, 2$$

$$\tilde{Q}_2 = I - \tilde{x}_2 (\tilde{x}'_2 \tilde{x}_2)^{-1} \tilde{x}_2$$

$$\tilde{y} = \Omega^{-1/2} y.$$

Observe também que (3.1) $\Leftrightarrow R\beta = \beta_1 = 0$ então utilizando-se a estatística de Wald tem-se

$$F^* = \frac{\tilde{\beta}'_1 [R(X'\Omega^{-1}X)^{-1}R']^{-1} \tilde{\beta}_1}{\sigma^2} \quad (3.1.3)$$

Se $\sigma^2\Omega$ for conhecido F^* tem uma distribuição $\chi^2_{n_1}$.

Se σ^2 for desconhecido, devemos usar um estimador sob H_A i.e.

$$\tilde{\sigma}^2 = \frac{\tilde{u}'_{H_A} \Omega^{-1} \tilde{u}_{H_A}}{T - n_1 - n_2} \quad (3.1.4)$$

Neste caso a estatística

$$F^*_1 = \frac{\tilde{\beta}'_1 [R(X'\Omega^{-1}X)^{-1}R']^{-1} \tilde{\beta}_1}{\tilde{\sigma}^2} \quad (3.1.5)$$

tem uma distribuição assintótica $\chi^2_{n_1}$. Além disto $\frac{F^*_1}{n_1}$ tem distribuição $F(n_1, T - n_1 - n_2)$ que é exata para pequenas amostras.

Observe que assumiu-se Ω conhecido em (3.1.4). O caso em que Ω é desconhecido não apresenta maiores dificuldades já que estimando-se Ω consistentemente por $\tilde{\Omega}$ e substituindo-se Ω por $\tilde{\Omega}$ na estatística (3.1.5) a distribuição assintótica de F^*_1 continua sendo uma $\chi^2_{n_1}$. Observe que o resultado exato não se verifica neste caso.

Observe que para se testar estabilidade dos coeficientes i.e.

$$H_0: \beta_2 = 0 \quad \text{vs} \quad H_1: \beta_2 \neq 0$$

o procedimento acima desenvolvido pode ser aplicado, i.e., valem as expressões (3.13), (3.14), (3.1.3) e (3.1.5) com as devidas modificações.

3.2 Modelos Dinâmicos

Suponha agora que X_2 contém variáveis endógenas defasadas, i.e., tem-se o seguinte modelo dinâmico

$$\begin{aligned} \underline{y} &= \underline{X}\underline{\beta} + \underline{u} = \underline{X}_1\underline{\beta}_1 + \underline{X}_2\underline{\beta}_2 + \underline{u} = \underline{X}_1\underline{\beta}_1 + \underline{X}_{21}\underline{\beta}_{21} + \underline{X}_{22}\underline{\beta}_{22} + \underline{u} \\ &= \underline{X}_1^+ \underline{\beta}_1^+ + \underline{X}_2^+ \underline{\beta}_2^+ + \underline{u} \end{aligned} \quad (3.2.1)$$

onde $\underline{X}_1^+ = [\underline{X}_1, \underline{X}_{21}]$ é uma matriz $T \times n_1^+$, $\underline{X}_{21}$ é uma matriz $T \times n_{21}$ de variáveis estritamente exógenas, $\underline{X}_2^+ (= \underline{X}_{22})$ é uma matriz $T \times n_2^+$ de variáveis endógenas defasadas, $n_2 = n_{21} + n_{22} = n_{21} + n_2^+$ e $\underline{u} \sim N(0, \sigma^2 \underline{I}_T)$.

Se, por exemplo, $n_2^+ = 1$, tem-se o seguinte modelo:

$$\underline{y} = \underline{X}_1^+ \underline{\beta}_1^+ + \underline{y}_{-1} \underline{\beta}_2^+ + \underline{u} \quad (3.2.2)$$

onde $\underline{y} = (y_1, \dots, y_T)'$ e $\underline{y}_{-1} = (y_0, \dots, y_{T-1})'$.

Deseja-se testar a estabilidade dos coeficientes das variáveis endógenas defasadas, i.e.,

$$H_0: \underline{\beta}_2^+ = \underline{\bar{\beta}}_2 \quad \text{vs} \quad H_A: \underline{\beta}_2^+ \neq \underline{\bar{\beta}}_2. \quad (3.2.3)$$

A estatística de Wald para este teste depende da distribuição de $\underline{\hat{\beta}}_2^+$, i.e., o estimador de $\underline{\beta}_2^+$ para o modelo irrestrito.

Para se obter a distribuição de $\underline{\hat{\beta}}_2^+$ necessita-se das seguintes hipóteses:

- (i) $p \lim \frac{\underline{X}'\underline{X}}{T} = M$, positiva definida, não singular.
- (ii) $p \lim \frac{\underline{X}'\underline{u}}{T} = 0$.

Particionando M de acordo com a partição de X tem-se

$$M = \begin{bmatrix} M_{11}^+ & M_{12}^+ \\ M_{21}^+ & M_{22}^+ \end{bmatrix} \quad \text{onde}$$

$$M_{11}^+ = p \lim \frac{(X_1^+)' 1}{T} \quad (3.2.4.a)$$

$$M_{12}^+ = p \lim \frac{(X_1^+)' X_2^+}{T} \quad (3.2.4.b)$$

$$M_{21}^+ = p \lim \frac{(X_2^+)' X_1^+}{T} \quad (3.2.4.c)$$

$$M_{22}^+ = p \lim \frac{(X_2^+)' X_2^+}{T} \quad (3.2.4.d)$$

Observe que $M^* = M_{22}^+ - M_{21}^+ (M_{11}^+)^{-1} M_{12}^+$ é não singular já que o determinante, $|M| = |M_{11}^+| |M^*|$.

Observe que $M^* = p \lim \frac{(X_2^+)' Q_1^+ X_2^+}{T}$, utilizando-se (3.2.4.a-d) e a definição de $Q_1^+ = I - X_1^+ (X_1^+)' X_1^+)^{-1} X_1^+$.

Agora como M^* é não singular, existe uma matriz H, que pode ser escolhida na forma triangular inferior, tal que

$$HH' = \frac{(X_2^+)' Q_1^+ X_2^+}{T}. \quad \text{Mais ainda } \bar{H}\bar{H}' = M^* \text{ onde } \bar{H} = p \lim H.$$

Tem-se então a seguinte proposição:

Proposição:

$$t = \sqrt{T} H' (\tilde{\beta}_2^+ - \bar{\beta}_2) \xrightarrow{\mathcal{L}} N(0, \sigma^2 I_{n_2}).$$

Demonstração: Pela hipótese (i) tem-se que

$$\sqrt{T}(\tilde{\beta}_2 - \bar{\beta}_2) \xrightarrow{\mathcal{L}} N(0, \sigma^2 M^{-1}) \text{ então } \sqrt{T}(\tilde{\beta}_2^+ - \bar{\beta}_2) \xrightarrow{\mathcal{L}} N(0, \sigma^2 (M^*)^{-1}).$$

Agora pelo teorema de transformação linear de Cramer (ver Cramer (1946)) tem-se

$$\sqrt{T} H' (\tilde{\beta}_2^+ - \bar{\beta}_2) \stackrel{a}{\sim} N(0, \sigma^2 \bar{H}' (M^*)^{-1} \bar{H}). \quad (3.2.5)$$

Mas

$$\begin{aligned} \bar{H}\bar{H}' &= M^* \Leftrightarrow \bar{H}\bar{H}' (M^*)^{-1} = I_{n_2} \Leftrightarrow \bar{H}' (M^*)^{-1} \bar{H} = (\bar{H})^{-1} \bar{H} \\ &\Leftrightarrow \bar{H}' (M^*)^{-1} \bar{H} = I_{n_2}. \end{aligned}$$

Portanto (3.2.5) é equivalente a

$$\sqrt{T} H' (\tilde{\beta}_2^+ - \bar{\beta}_2) \stackrel{a}{\sim} N(0, \sigma^2 I_{n_2}).$$

Pela proposição I e supondo σ^2 conhecido tem-se que

$$\begin{aligned} \frac{t't}{\sigma^2} &\stackrel{a}{\sim} \chi_{n_2}^2 \Leftrightarrow \frac{\sqrt{T} (\tilde{\beta}_2^+ - \bar{\beta}_2)' \bar{H}\bar{H}' (\tilde{\beta}_2^+ - \bar{\beta}_2) \sqrt{T}}{\sigma^2} \stackrel{a}{\sim} \chi_{n_2}^2 \\ &\Leftrightarrow \frac{(\tilde{\beta}_2^+ - \bar{\beta}_2)' (X_2^{*'} Q_1^* X_2^*) (\tilde{\beta}_2^+ - \bar{\beta}_2)}{\sigma^2} \stackrel{a}{\sim} \chi_{n_2}^2. \end{aligned} \quad (3.2.6.)$$

Se σ^2 é desconhecido, use $\hat{\sigma}^2 = \frac{1}{T-n} \tilde{u}'\tilde{u}$ em (3.2.6)

onde $\tilde{u} = \underline{y} - X\tilde{\beta}$ e $\tilde{\beta}$ é o estimador de $\underline{\beta}$ no modelo irrestrito.

Portanto tem-se a estatística

$$K_3 = \frac{(\tilde{\beta}_2^+ - \bar{\beta}_2)' (X_2^{*'} Q_1^* X_2^*) (\tilde{\beta}_2^+ - \bar{\beta}_2)}{\frac{\tilde{u}'\tilde{u}}{T-n}} \stackrel{a}{\sim} \chi_{n_2}^2 \quad (3.2.7)$$

Observe que a estatística $F = K_3/n_2$ tem distribuição $F(n_2, T - n_1 - n_2)$ somente assintoticamente. Portanto o teste F para peque nas amostras não é válido neste caso.

4. CONCLUSÕES

Uma forma alternativa do estudo de variáveis "dummies" em regressão foi proposta neste artigo.

Três métodos de estimação foram propostos. Mostrou-se que estes métodos eram equivalentes entre si.

Testes de estabilidade dos coeficientes da regressão foram obtidos. Foi possível testar estabilidade do termo constante e/ou dos coeficientes do modelo. Viu-se que, tal teste, é aplicável a modelo com erros correlacionados serialmente ou apresentando heteroscedasticidade. No caso de modelo dinâmico os resultados são válidos assintoticamente.

Foi derivada uma forma alternativa do teste de Chow, que é baseada na estatística de Wald. Estes testes são assintoticamente equivalentes ao teste de Chow. Para se calcular a estatística de Wald, só foi necessário estimar o modelo irrestrito. Esta constitui-se a principal vantagem desta formulação equivalente.

Uma alternativa para o teste de estabilidade é o teste de predição de pós-amostra (ver Harvey (1981) pg. 181-182) que é equivalente ao teste de Chow mas com a vantagem da estatística de Wald (só é preciso a estimação do modelo sob H_A).

REFERÊNCIAS

- BUSE, A. (1982) "Likelihood Ratio, Wald and Lagrange multiplier tests: An expository note". *The American Statistician*, 36, p. 153-157.
- CHOW, G.C. (1960) "Tests for equality between sets of coefficients in two linear regressions", *Econometrica*, 28, p. 591-605.
- CRAMER, H. (1946) *Mathematical methods of Statistics*. Princeton University Press: New Jersey.
- FISHER, F.M. (1970) "Tests of equality between sets of coefficients in two linear regressions: An expository note" *Econometrica*, 38, p. 361-366.
- HARVEY, A.C. (1981) *The Econometric Analysis of Time Series*, Philip Allan: Oxford.
- JOHNSTON, J. (1972) *Econometric methods*, 2nd Ed. McGraw Hill: Tokyo.
- JUDGE, G.G., et. al (1980) *The Theory and Practice of Econometrics*, John Wiley: New York.
- MADDALA, G.S. (1977) *Econometrics*, McGraw Hill: Tokyo.
- PHILIPS, P.C.B. & WICKENS, M.R. (1978) *Exercises in Econometrics*, Vol. I, Philip Allan: Oxford.
- THEIL, H. (1971) *Principles of Econometrics*, John Wiley: New York.