GABARITO - QUESTÃO 1:

Suponha um modelo de regressão de duas variáveis descrito por $Y_i = \beta_1 + \beta_2 X_i + u_i$ em que os Y_i sejam normal e independentemente distribuídos, com média $\beta_1 + \beta_2 X_i$ e variância σ^2. Tendo em consideração que as funções de densidade individuais são dadas por:

$$f(Y_i) = \frac{1}{\sigma\sqrt{2\pi}} exp\left\{-\frac{1}{2}\frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\}.$$

Pede-se:

(a) Se $Y_1, Y_2, \dots, Y_n$ são conhecidos, ou dados, mas β_1, β_2 e σ^2 não são. Obtenha a função de verossimilhança, denotada por $FV(\beta_1, \beta_2, \sigma^2)$.
(b) Utilizando o método de máxima verossimilhança, obtenha os estimadores de máxima verossimilhança de β_1, β_2 e σ^2.
(c) Sabendo que o estimador de Mínimos Quadrados Ordinários para σ^2 é não-viesado, mostre que o estimador de máxima verossimilhança de σ^2 é viesado e determine a magnitude desse viés.

Resolução

GUJARATI, D.N.; PORTER, D.C. **Econometria básica**. 5.ed. Porto Alegre: AMGH, 2011, p. 124-125.

Suponha que no modelo de duas variáveis $Y_i = \beta_1 + \beta_2 X_i + u_i$ os Y_i sejam normal e independentemente distribuídos, com média $= \beta_1 + \beta_2 X_i$ e variância $= \sigma^2$. (Veja a Equação (4.3.9). Em consequência, a função de densidade de probabilidade conjunta de $Y_1, Y_2, \dots, Y_n$, dadas a média e a variância anteriores, pode ser escrita como

$$f(Y_1, Y_2, \dots, Y_n \mid \beta_1 + \beta_2 X_i, \sigma^2)$$

Mas, tendo em vista a independência dos Y, essa função de densidade de probabilidade conjunta pode ser expressa como um produto de n funções de densidade individuais

$$\begin{aligned} &f(Y_1, Y_2, \dots, Y_n \mid \beta_1 + \beta_2 X_i, \sigma^2) \\ &\quad = f(Y_1 \mid \beta_1 + \beta_2 X_i, \sigma^2) f(Y_2 \mid \beta_1 + \beta_2 X_i, \sigma^2) \cdots f(Y_n \mid \beta_1 + \beta_2 X_i, \sigma^2) \end{aligned} \tag{1}$$

em que

$$f(Y_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\} \tag{2}$$

que é a função de densidade de uma variável com distribuição normal, dadas a média e a variância.

(*Nota:* exp significa *e* elevado à potência da expressão indicada por {}.)

Substituindo a Equação (2) por cada Y_i na Equação (1) obtemos

$$f(Y_1, Y_2, \ldots, Y_n \mid \beta_1 + \beta_2 X_i, \sigma^2) = \frac{1}{\sigma^n \left(\sqrt{2\pi}\right)^n} \exp\left\{-\frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\} \tag{3}$$

Se $Y_1, Y_2, \ldots, Y_n$ são conhecidos ou dados, mas β_1, β_2 e σ^2 não são, a função na Equação (3) é chamada de **função de verossimilhança**, denotada por FV(β_1, β_2 e σ^2), e expressa como[1]

$$\mathrm{FV}(\beta_1, \beta_2, \sigma^2) = \frac{1}{\sigma^n \left(\sqrt{2\pi}\right)^n} \exp\left\{-\frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2}\right\} \tag{4}$$

O **método da máxima verossimilhança**, como o nome indica, consiste em estimar os parâmetros desconhecidos de maneira que a probabilidade de observar os dados Y seja a maior (ou a máxima) possível. Precisamos encontrar o máximo da função na Equação (4). Isso é um exercício direto de cálculo diferencial. Para derivar, é mais fácil expressar a Equação (4) na forma logarítmica, como a seguir.[2] (*Nota:* ln = log natural.)

$$\begin{aligned} \ln \mathrm{FV} &= -n \ln \sigma - \frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2} \\ &= -\frac{n}{2}\ln \sigma^2 - \frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2} \end{aligned} \tag{5}$$

Derivando a Equação (5) parcialmente em relação a β_1, β_2 e σ^2, obtemos

$$\frac{\partial \ln \mathrm{FV}}{\partial \beta_1} = -\frac{1}{\sigma^2}\sum (Y_i - \beta_1 - \beta_2 X_i)(-1) \tag{6}$$

$$\frac{\partial \ln \mathrm{FV}}{\partial \beta_2} = -\frac{1}{\sigma^2}\sum (Y_i - \beta_1 - \beta_2 X_i)(-X_i) \tag{7}$$

$$\frac{\partial \ln \mathrm{FV}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum (Y_i - \beta_1 - \beta_2 X_i)^2 \tag{8}$$

Igualando essas equações a zero (condição de primeira ordem para a otimização) e denotando os estimadores de máxima verossimilhança por $\tilde{\beta}_1, \tilde{\beta}_2$ e $\tilde{\sigma}^2$ obtemos[3]

$$\frac{1}{\tilde{\sigma}^2}\sum (Y_i - \tilde{\beta}_1 - \tilde{\beta}_2 X_i) = 0 \tag{9}$$

$$\frac{1}{\tilde{\sigma}^2}\sum (Y_i - \tilde{\beta}_1 - \tilde{\beta}_2 X_i) X_i = 0 \tag{10}$$

$$-\frac{n}{2\tilde{\sigma}^2} + \frac{1}{2\tilde{\sigma}^4}\sum (Y_i - \tilde{\beta}_1 - \tilde{\beta}_2 X_i)^2 = 0 \tag{11}$$

Após a simplificação, as Equações (9) e (10) ficam como

$$\sum Y_i = n\tilde{\beta}_1 + \tilde{\beta}_2 \sum X_i \tag{12}$$

$$\sum Y_i X_i + \tilde{\beta}_1 \sum X_i + \tilde{\beta}_2 \sum X_i^2 \tag{13}$$

que são exatamente as *equações normais* da teoria dos mínimos quadrados obtidas nas Equações (3.1.4) e (3.1.5). Portanto, os estimadores de máxima verossimilhança, os $\tilde{\beta}$, são idênticos aos estimadores de MQO, os $\hat{\beta}$, dados nas Equações (3.1.6) e (3.1.7). Essa igualdade não é acidental. Examinando a verossimilhança (5), vemos que o último termo entra com sinal negativo. Maximizar a Equação (5) é o mesmo que minimizar esse termo, que é justamente o que faz a abordagem dos mínimos quadrados, como se pode ver na Equação (3.1.2).

Substituindo na Equação (11) os estimadores de máxima verossimilhança (= MQO) e simplificando, obtemos o estimador de máxima verossimilhança para $\tilde{\sigma}^2$ como

$$\begin{aligned}\tilde{\sigma}^2 &= \frac{1}{n}\sum(Y_i - \tilde{\beta}_1 - \tilde{\beta}_2 X_i)^2 \\ &= \frac{1}{n}\sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \\ &= \frac{1}{n}\sum \hat{u}_i^2\end{aligned} \tag{14}$$

Com base na Equação (14) fica óbvio que o estimador de máxima verossimilhança $\tilde{\sigma}^2$ difere do estimador de MQO $\hat{\sigma}^2 = [1/(n-2)]\sum \hat{u}_i^2$, que como já foi demonstrado no Apêndice 3A é um estimador não viesados de σ^2. Assim, o estimador de máxima verossimilhança de σ^2 é viesado. A magnitude desse viés pode ser determinada com facilidade do seguinte modo:

Tomando-se a esperança matemática da Equação (14) de ambos os lados, obtemos

$$\begin{aligned}E(\tilde{\sigma}^2) &= \frac{1}{n}E\left(\sum \hat{u}_i^2\right) \\ &= \left(\frac{n-2}{n}\right)\sigma^2 \qquad \text{usando a Equação (16) da Seção 3A.5 do Apêndice 3A} \\ &= \sigma^2 - \frac{2}{n}\sigma^2\end{aligned} \tag{15}$$

que mostra que $\tilde{\sigma}^2$ é viesado para baixo (isto é, subestima o verdadeiro σ^2) em amostras pequenas. Note que quando n, o tamanho da amostra, aumenta indefinidamente, o segundo termo na Equação (15), o fator de viés,

GABARITO – QUESTÃO 2:

GUJARATI, D.N.; PORTER, D.C. **Econometria básica**. 5.ed. Porto Alegre: AMGH, 2011, Cap.16, p. 590-595.

Os dados em painel são frequentemente analisados usando modelos de efeitos fixos. No entanto, alguns pesquisadores podem optar por estimar modelos pooled sem considerar a presença de efeitos fixos. Com base em Gujarati e Porter (2011):

a) Explique as consequências da estimação de modelos Polled na presença de efeitos não observáveis, invariantes no tempo.
b) Formalize o modelo de Mínimos Quadrados com variáveis dummy para efeitos fixos.
c) Discorra sobre os problemas que podem ser gerados pela estimação de Mínimos Quadrados com a adição de variáveis dummy para efeitos fixos.

Respostas:

Alternativa a:

De acordo com Gujarati e Porter (2011), uma hipótese importante do modelo clássico de regressão linear e que não haja correlação entre os regressores e o termo de erro. Para entender como o termo de erro pode estar correlacionado com os regressores considere que o modelo correto inclui uma variável que mensura a qualidade gerencial ou filosofia gerencial, M_{it}:

$$CT_{it} = \beta_1 + \beta_2 PF_{it} + \beta_3 LF_{it} + \beta_4 M_{it} + u_{it}, \tag{16.3.2}$$

Das variáveis incluídas na Equação (16.3.2), apenas a variável M e invariante no tempo (ou constante no tempo), porque varia entre os indivíduos, mas é constante ao longo do tempo para determinado indivíduo (empresa aérea).

Embora seja invariante em termos de tempo, a variável M não e diretamente observável e, portanto, não podemos medir sua contribuição para a função de custo. Podemos, entretanto, fazer isso indiretamente, se escrevermos a Equação (16.3.2) como:

$$CT_{it} = \beta_1 + \beta_2 PF_{it} + \beta_3 LF_{it} + \alpha_i + u_{it}, \tag{16.3.3}$$

em que α_i, chamado efeito nao observado, ou heterogeneidade, reflete o impacto de M sobre o custo. Note que por simplicidade mostramos apenas o efeito nao observado de M sobre o custo, mas na realidade pode haver mais efeitos não observados, por exemplo, a natureza da propriedade (privada ou pública), se uma empresa e de capital aberto ou fechado, se o CEO e homem ou mulher etc. Embora tais variáveis possam diferir entre os indivíduos (empresas aéreas), provavelmente permanecerão as mesmas para um dado indivíduo sobre todo o período da amostra.

Uma vez que α_i não e diretamente observavel, por que não considerá-la aleatória e incluí-la no termo de erro u_{it}, e considerar o termo de erro $v_{it} = a_i + u_{it}$? Agora escrevemos a Equação (16.3.3) como:

$$CT_{it} = \beta_1 + \beta_2 PF_{it} + \beta_3 LF_{it} + v_{it}, \qquad (16.3.4)$$

Mas se o termo α_i, incluído no termo de erro v_{it} está correlacionado com qualquer um dos regressores na Equação (16.3.4), temos uma violação de uma das hipóteses principais do modelo clássico de regressao linear - ou seja, que o termo de erro não está correlacionado com os regressores. Como sabemos, nessa situação, as estimativas de MQO não são apenas tendenciosas, mas também inconsistentes.

Ha uma possibilidade real de que o α_i não observável esteja correlacionado com um ou mais regressores. Por exemplo, a direção de uma empresa aérea pode ser perspicaz o suficiente para comprar combustivel na bolsa de futuros, a fim de evitar severas flutuações de preço. Isso tera o efeito de reduzir o custo dos serviços de transporte aéreo. Como resultado dessa correlação, pode-se mostrar que $cov\ (v_{it}, v_{is}) = \sigma_u^2;\ t \neq\ s$, que é diferente de zero, e, portanto, a heterogeneidade (não observada) induz a autocorrelação e teremos de prestar atenção nisso.

Alternativa b:

O modelo de mínimos quadrados com variáveis dummy para efeitos fixos (MQVD) conta com a heterogeneidade entre indivíduos:

$$CT_{it} = \beta_1 + \beta_2 PF_{it} + \beta_3 LF_{it} + v_{it}, \qquad (16.3.4)$$

Observe que colocamos o subscrito i no termo de intercepto para sugerir que os interceptos das seis empresas aéreas podem ser diferentes. A diferença pode ser devida a aspectos especiais de cada uma, como estilo gerencial, filosofia gerencial, ou tipo de mercado que cada organização está servindo.

Na literatura específica, o modelo (16.4.1) e conhecido como modelo (de regressão) de efeitos fixos (MEF). O termo “efeitos fixos” deve-se ao fato de que, embora o intercepto possa diferir entre os indivíduos (no caso, seis empresas aéreas), o intercepto de cada indivíduo não varia com o tempo; ele e invariante no tempo. Note que, se tivéssemos de escrever o intercepto como f J >|ele sugeriria que o intercepto de cada indivíduo e variante no tempo. Podemos observar que esse modelo dado na Equação (16.4.1) pressupõe que os coeficientes (angulares) dos regressores não variam entre indivíduos nem com o tempo.

Alternativa c:

A estimação de Mínimos Quadrados com variáveis dummy para efeitos fixo pode gerar vários problemas: Primeiro, a introdução de variáveis dummies demais pode gerar problemas de falta de graus de liberdade. Segundo, a adição de diversas dummies no modelos, tanto individuais quanto interativas ou multiplicativas pode resultar em multicolineariedade, o que pode dificultar a estimação exata de um ou mais parâmetros.

Em terceiro lugar, a estimação de Mínimos Quadrados com variáveis dummy para efeitos fixo pode não ser capaz de identificar o impacto das variáveis que não mudam ao longo do tempo. Suponha que desejemos estimar uma função de salário para um grupo de trabalhadores utilizando dados em painel. Além do salário, uma função de salário pode incluir idade, experiência e educação como variáveis explanatórias. Suponha que também desejemos incluir o gênero, cor e raça como variáveis adicionais do modelo. Uma vez que essas variáveis não mudarão com o tempo para um indivíduo, a abordagem de Mínimos Quadrados com variáveis dummy para efeitos fixo pode não identificar o impacto dessas variáveis invariantes no tempo sobre o salário. Em outras palavras, os interceptos específicos a um sujeito absorvem toda a heterogeneidade que possa existir nas variáveis dependentes e explanatória.

Quarto, supomos que o termo de erro segue as hipóteses clássicas. Porém, existem diversas possibilidades, entre as quais:

1. Podemos considerar que a variância de erro é a mesma para todas as unidades de corte transversal, ou podemos considerar que ela é heterocedástica.
2. Para cada indivíduos, podemos supor que não haja autocorrelação ao longo do tempo.
3. Durante algum tempo, é possível que o termo de erro de um indivíduo esteja correlacionado com o termo de erro de outro. Ou podemos supor que não haja correlação.

Existem ainda outras complicações. Aceitar uma ou mais dessas possibilidades tornará a análise muito mais complicada.

GABARITO – QUESTÃO 3:

WOOLDRIDGE, J.M. **Introdução à Econometria: Uma Abordagem Moderna**. 6.ed. São Paulo: Thomson, 2016, Cap.16.

O seguinte modelo de equações simultâneas impõe a condição de equilíbrio de que a oferta é igual à demanda de moradias em uma cidade:

$$Q_d = \alpha_0 + \alpha_1 P + \alpha_2 Y + u_d$$

$$Q_O = \beta_0 + \beta_1 P + \beta_2 C + u_O$$

Em que:

Q_d é a quantidade demandada de moradias

P é o preço das moradias
Y é a renda média dos consumidores da cidade
u_d é o termo de erro da equação de demanda
Q_O é a quantidade ofertada de moradias
C é o custo médio de construção
u_O é o termo de erro da equação de oferta

Responda:

(a) Suponha que exista disponibilidade de dados sobre $Q, P, Y\ e\ C$. Como os parâmetros das equações de demanda e oferta podem ser estimados pelo método de mínimos quadrados em dois estágios?

No primeiro estágio, deve-se regredir P contra todas as variáveis exógenas (Y e C) e obter os valores de $\hat{P}$.

$$P = \gamma_0 + \gamma_1 Y + \gamma_2 C + v$$

No segundo estágio, utilizar os valores preditos de P ($\hat{P}$) nas equações originais e estimar os parâmetros por MQO ou VI.

Equação de demanda: $Q = \alpha_0 + \alpha_1 \hat{P} + \alpha_2 Y + u_d$

Equação de oferta: $Q = \beta_0 + \beta_1 \hat{P} + \beta_2 C + u_O$

(b) Quais variáveis podem ser usadas como instrumentos para P nas equações de oferta e demanda?

- A equação de <u>demanda</u> está identificada se houver pelo menos uma variável exógena que afete a oferta, mas não a demanda. Nesse caso, C (custo de construção) pode servir como instrumento para P.
- A equação de <u>oferta</u> está identificada se houver pelo menos uma variável exógena que afete a demanda, mas não a oferta. Nesse caso, Y (renda média) pode ser usado como instrumento para P.
- Outra possível resposta é $\hat{P}$, que é uma combinação linear entre C e Y, e também pode ser utilizado como instrumento para P nas equações de oferta e demanda.

Assinaturas do documento

Código para verificação: **1V5XN33U**

Este documento foi assinado digitalmente pelos seguintes signatários nas datas indicadas:

ANALUCIA VIEIRA FANTIN (CPF: 891.XXX.590-XX) em 24/06/2024 às 10:49:50
Emitido por: "SGP-e", emitido em 13/07/2018 - 13:16:58 e válido até 13/07/2118 - 13:16:58.
(Assinatura do sistema)

Para verificar a autenticidade desta cópia, acesse o link https://portal.sgpe.sea.sc.gov.br/portal-externo/conferencia-documento/VURFU0NfMTIwMjJfMDAwMjYwODVfMjYxMjJfMjAyNF8xVjVYTjMzVQ== ou o site https://portal.sgpe.sea.sc.gov.br/portal-externo e informe o processo **UDESC 00026085/2024** e o código **1V5XN33U** ou aponte a câmera para o QR Code presente nesta página para realizar a conferência.