Notas de aula para o curso de Econometria I

Nota 7: propriedades assintóticas do MCRL e teste do multiplicador de Lagrange

Thiago Fonseca Morello

fonseca.morello@ufabc.edu.br

sala 301, Bloco Delta, SBC

1 Propriedades assintóticas do MCRL

1.1 Introdução

Há três propriedades do MCRL bastante úteis na prática, sendo as duas primeiras ausência de viés e eficiência, esta última estabelecida pelo teorema de Gauss-Markov. A terceira propriedade diz respeito a funções de distribuição de probabilidades (FDs) das estatísticas empregadas nos testes de hipóteses para significância individual, para significância conjunta e para restrições lineares. Nisso também deve ser incluída a estatística tomada por base para a construção de intervalos de confiança. A terceira propriedade é a de que todas as estatísticas mencionadas têm FDs conhecidas, t de Student e F de Snedecor.

Uma quarta propriedade ainda não estudada é a de consistência que diz respeito ao comportamento do valor esperado do estimador de MQO em uma amostra de tamanho arbitrariamente grande, infinito. O estudo do comportamento de estimadores em amostras de tamanho infinito é objeto de um componente específico da teoria estatística, denominada por teoria assintótica. Essa nota de aula introduz alguns dos principais resultados atingidos em tal teoria para algumas propriedades dos estimadores de MQO.

No caso específico da propriedade de consistência, é desejável que a expansão ilimitada da amostra, i.e., a incorporação de conteúdo informacional arbitrariamente grande aos estimadores, se traduza em maior acurácia dos mesmos. Mais do que isso, na atualidade, em pesquisa empírica com econometria, predomina a avaliação de que inconsistência é sinônimo de inacurácia intolerável e a recomendação é a de que estimadores inconsistentes sejam descartados. Este tema é tratado na segunda subseção deste tópico.

Uma segunda razão para o interesse nas propriedades assintóticas dos estimadores de MQO, i.e., propriedades em amostras de tamanho ilimitado, infinito, diz respeito à validade da terceira propriedade supramencionada. Esta depende da hipótese de que a FD das perturbações é normal (hipótese MCRL7), o que pode não ser condizente com os dados. Basta examinar histogramas para os resíduos e realizar testes para a hipótese nula de normalidade. Caso atestada a violação de MCRL7, os testes de hipóteses e a estimação por intervalo estariam se apoiando em estatísticas que não tem FDs t de Student ou F de Snedecor, mas sim outras FDs desconhecidas. Resultados equivocados seriam gerados, portanto, ao serem tomados por base valores críticos ou p-valores obtidos das FDs tradicionais. Este último problema pode ser mitigado, ainda que não corrigido plenamente, com base no resultado de que as estatísticas utilizadas em testes de hipóteses e intervalos de confiança têm FDs equivalentes às tradicionais em amostras de tamanho infinito, com pequenas diferenças nos parâmetros (especialmente quanto à variância). É possível, pois, entender que os testes e intervalos de confiança possuem, em amostra finita, distribuições aproximadamente equivalentes às FDs válidas para amostra infinita. Ou seja, com base na teoria assintótica, torna-se possível realizar testes de hipóteses e calcular intervalos de confiança aproximados. Este ponto é brevemente discutido na subseção 1.4 abaixo.

Conforme esclarece Wooldridge, a propriedade de ausência de viés é exata pois vale para qualquer tamanho amostral[1]. Ela é chamada de propriedade em “amostra finita” ou, no jargão, em “amostra pequena”. Já as propriedades assintóticas são aproximadas e valem apenas com o tamanho amostral tendendo ao infinito, o que não diz respeito a um tamanho amostral específico. É utilizado o termo “amostra grande” para se referir à situação assintótica, mas não há, em geral, um limiar a partir do qual é possível dizer se uma amostra é suficientemente grande ou não para que as propriedades assintóticas sejam exatamente (e não aproximadamente) válidas.

1.2 Consistência (convergência em probabilidade)

O conceito de consistência se apoia na definição de convergência em probabilidade. Esta, por sua vez, se refere a uma estatística cujo valor varia com o tamanho da amostra, tal como é o caso, por exemplo, da média amostral, da variância amostral e do estimador de MQO para os parâmetros da FRP,$\hat{\beta}_{MQO}$. A estatística genérica pode ser denotada por $\hat{\theta}_N$, em que N é o tamanho da amostra e cabe recordar que toda estatística, sendo uma função de variáveis aleatórias (dados) é também uma variável aleatória (VA).

Definição: convergência em probabilidade

Uma sequencia de variáveis aleatórias $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N$ *converge em probabilidade para o valor* θ_L *se, qualquer que seja* $\varepsilon > 0$, $\lim_{N \to \infty} P(¿\hat{\theta}_N - \theta_L \vee ¿\varepsilon) = 0$.

É preciso interpretar a definição. Ela está dizendo que, com o aumento irrestrito do tamanho da amostra, o valor da estatística se torna de tal maneira próximo do valor-limite, θ_L, que a diferença entre ambas chega a ser desprezível. A medida desta diferença é probabilística, trata-se da probabilidade de que ela seja relevante, superior a um valor ε qualquer, não nulo – tal ε é um limiar de materialidade, ou seja a referência para dizermos se um número é ou não desprezível. Dizer que é nula a probabilidade de que a diferença entre $\hat{\theta}_N$ e θ_L seja relevante é uma maneira de expressar, com rigor estatístico, a afirmação de que a diferença em questão é desprezível, irrelevante.

[1] A propriedade de eficiência, que atribui ao estimador de MQO o status de MELNV, também é uma propriedade exata quando se refere a amostra finita. Mas é também comum tomar por base a versão assintótica da propriedade de eficiência.

Antes de passar à definição de consistência, cabe apresentar um enunciado geral para a lei fraca dos grandes números, atribuído ao matemático Alexandr Khinchin (Teorema D5, Greene).

Lei fraca dos grandes números de Khinchin

Seja $X_1,...,X_N$ uma amostra aleatória de tamanho N de uma variável X, tal que $E[X_i]<\infty$, finita, e $E[X_i]=E[X_j]$, i,j=1,...,N.

$$\lim_{N\to\infty} P\left(\left|\frac{1}{N}\sum_{i=1}^{N} X_i - E[X_i]\right| > \varepsilon\right) = 0$$

Ou seja, para uma amostra arbitrariamente grande, a média amostral converge em probabilidade para a expectativa. Uma diferença relevante entre as duas ocorre com probabilidade zero assim que o tamanho da amostra aumenta de maneira irrestrita. Notar que a expressão anterior é equivalente à definição de convergência em probabilidade, com $\hat{\theta}_N$ tendo sido substituído por $\frac{1}{N}\sum_{i=1}^{N} X_i$ e θ por $E[X_i]$. Após a lei fraca de Khinchin, a definição de consistência a seguir se torna mais clara.

Consistência

O estimador de MQO, $\hat{\beta}_{MQO,N}$, é consistente quando converge em probabilidade para o valor populacional do vetor de parâmetros da FPR. Formalmente:

$$\lim_{N\to\infty} P\left(\left|\hat{\beta}_{MQO,N} - \beta\right| > \varepsilon\right) = 0$$

O estimador de MQO é consistente, pois, quando seu valor se torna arbitrariamente próximo do valor populacional do vetor de parâmetros, com a expansão ilimitada do tamanho do conteúdo informacional incorporado. Uma maneira sucinta de denotar o limite de probabilidade da definição de consistência é por meio do operador "plim", de modo que a definição passa a:

O estimador de mínimos quadrados ordinários, $\hat{\beta}_{MQO}$, é dito consistente se e somente se $plim(\hat{\beta}_{MQO}) = \beta$.

Com base na noção de plim, é possível verificar se, sob as hipóteses do MCRL, o estimador $\hat{\beta}_{MQO}$ é de fato consistente, ou seja, se a definição acima se aplica. Para isso será utilizada a notação matricial das notas de aula 4 e 5.

$$plim\left(\hat{\beta}_{MQO}\right) = plim\left(\beta + \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i u_i\right) = \beta + plim\left(\left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i u_i\right)$$

Nessa primeira passagem, utilizou-se uma propriedade do operador plim, o fato de que o limite de probabilidade de uma constante é a própria constante. Fica nítido que há consistência apenas se o plim remanescente é nulo. Aplica-se, agora, uma segunda propriedade, segundo a qual o plim de um produto é o produto dos plims dos fatores:

$$plim\left(\hat{\beta}_{MQO}\right)=\beta+plim\left(\left(\sum_{i=1}^{N}x_i x_i^{'}\right)^{-1}\sum_{i=1}^{N}x_i u_i\right)=\beta+\left(plim\left[\left(\sum_{i=1}^{N}x_i x_i^{'}\right)^{-1}\right]plim\left[\sum_{i=1}^{N}x_i u_i\right]\right)(1)$$

Para tornar mais claros quais são os limites de probabilidade de cada fator, é útil multiplicar a segunda expressão à esquerda por N/N, obtendo-se:

$$plim\left(\hat{\beta}_{MQO}\right)=\beta+\frac{N}{N}\left(plim\left[\left(\sum_{i=1}^{N}x_i x_i^{'}\right)^{-1}\right]plim\left[\sum_{i=1}^{N}x_i u_i\right]\right)$$

$$plim\left(\hat{\beta}_{MQO}\right)=\beta+\left(plim\left[\left(\frac{1}{N}\sum_{i=1}^{N}x_i x_i^{'}\right)^{-1}\right]plim\left[\frac{1}{N}\sum_{i=1}^{N}x_i u_i\right]\right)$$

Neste ponto serão introduzidos dois fatos, sem a devida demonstração pois está fora do escopo da disciplina (ver Greene, 2002, seção 5.2). A noção intuitiva para os fatos é que, geralmente, o plim da média amostral de uma quantidade é a expectativa da quantidade. Os fatos são:

$$plim\left[\left(\sum_{i=1}^{N}x_i x_i^{'}\right)^{-1}\right]=E\left[x_i x_i{'}\right]^{-1}=Q^{-1}(2)$$

$$plim\left[\sum_{i=1}^{N}x_i u_i\right]=E\left[x_i u_i\right](3)$$

Sendo Q uma matriz não-nula, finita e invertível. Aplicando a hipótese MCRL4 (exogeneidade), tem-se que E[ui|X] = 0, i=1,...,N. O que implica: $E[u_i x_i] = E[E[u_i|X].x_i] = E[0.x_i] = 0$, i=1,...,N (conforme visto na nota de aula 3, seção 3.2.2). Deste modo, pois, $plim\left[\sum_{i=1}^{N}x_i u_i\right]$ = 0 (3’). Incorporando (2) e (3’) em (1):

$$plim\left(\hat{\beta}_{MQO}\right)=\beta+\left(Q^{-1}X\,0\right)=\beta$$

Sendo $Q^{-1}X\,0$ o produto de uma matriz não-nula e finita pelo escalar zero, o que resulta em uma matriz de zeros.

Deve-se assinalar que não apenas a hipótese MCRL4 é necessária à consistência, mas também as hipóteses MCRL1 a MCRL3. Em segundo lugar, um corolário da primeira afirmação que merece ser assinalado é o de que a violação da hipótese de exogeneidade ($E[u_i|X]=0$, i=1,...,N) implica na perda da propriedade de consistência.

1.3 Viés assintótico de estimador alternativo

Cabe apresentar o cálculo do tamanho da inconsistência (ou do viés assintótico) de um estimador de MQO na situação em que MCRL4 é violada, seguindo os passos de Wooldridge. Conforme seção anterior, o tamanho da inconsistência, no caso geral, é:

$$plim\left(\hat{\beta}_{MQO}\right)-\beta=E\left[x_i u_i\right]E\left[x_i x_i'\right]^{-1}=cov\left(x,u\right)V\left[x_i\right]^{-1}$$

A última passagem incorporando a notação utilizada por Wooldrige.

Seja assumindo que a FRP verdadeira é $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u$, i=1,...,N. Contudo, por erro, assume-se $y_i = \beta_0 + \beta_1 x_{1i} + v$, i=1,...,N, $v = \beta_2 x_{2i} + u$, de modo que a segunda variável independente é captada pelo termo de perturbação. Neste caso, refazendo o raciocínio a seção anterior e considerando-se que se trata de uma regressão simples e não múltipla, chega-se a:

$$plim\left(\hat{\beta}_{MQO,1}\right)-\beta_1=cov\left(x_1,v\right)V\left[x_{1i}\right]^{-1}=cov\left(x_1,\beta_2 x_2+u\right)V\left[x_{1i}\right]^{-1}(4)$$

É possível demonstrar que cov(x+y,z) = cov(x,z) + cov(y,z)[2]. Ou seja, $cov\left(x_1,\beta_2 x_2+u\right)=cov\left(x_1,\beta_2 x_2\right)+cov\left(x_1,u\right)=\beta_2 cov\left(x_1,x_2\right)+0(5)$. A última passagem leva em conta uma propriedade do operador covariância e a hipótese de que E[u|x1,x2] = 0 e, pois, cov(u,x2) = 0. Incorporando (5) em (4):

$$plim\left(\hat{\beta}_{MQO,1}\right)-\beta_1=\beta_2 cov\left(x_1,x_2\right)V\left[x_{1i}\right]^{-1}(4')$$

Tem-se, pois, uma inconsistência não-nula caso as duas variáveis independentes forem estatisticamente relacionadas, i.e., apresentem covariância não-nula. O que revela as duas condições que implicam em estimador inconsistente de MQO, sendo elas:

1. Uma variável independente relevante é omitida do modelo, compondo, pois, o termo de perturbação;
2. A variável independente omitida possui covariância não-nula com pelo menos uma variável independente incluída.

Como resultado, há covariância não-nula entre o termo de perturbação e a variável independente incluída. Cabe também notar que o sinal (positivo ou negativo) da inconsistência é igual ao sinal da covariância entre as variáveis.

Notar, pois, que omitir variável relevante não torna os estimadores de MQO inconsistentes caso tal variável seja não-correlacionada com as variáveis independentes incluídas.

1.4 Convergência em distribuição

Mesmo que o comportamento dos resíduos refute a hipótese de normalidade das perturbações, os teoremas a seguir garantem que em amostras suficientemente grandes,

[2] E[(x+y-E[x+y])(z-E[z])] = E[(x-E[x]+y-E[y])(z-E[z])] = E[(x-E[x])(z-E[z]) + (y-E[y])(z-E[z])] = E[(x-E[x])(z-E[z])] + E[(y-E[y])(z-E[z])].

as estatísticas utilizadas na inferência do MCRL têm FDs desprezivelmente distintas das tradicionalmente tomadas por base.

Teorema 1: sob as hipóteses MCRL1-MCRL6, $\sqrt{N}\left(\hat{\beta}_{MQOk}-\beta_k\right) \to N\left(0,\sigma^2 E\left[x_i x_i{}'\right]^{-1}\right) e\, t=\frac{\left(\hat{\beta}_{MQOk}-\beta_k\right)}{\sqrt{\widehat{V\left(\hat{\beta}_{MQOk}\right)}}} \to^{D} t_{N-(K+1)}$ *(ver Wooldridge, teorema 5.2 e equação 5.8)*

Teorema 2: sob as hipóteses MCRL1-MCRL6, $\hat{F}=\frac{\left(R_{IR}^2-R_R^2\right)/s}{\left(1-R_{IR}^2\right)/\left(N-(K+1)\right)} \to^{D} F_{s,N-(K+1)}$

O símbolo "→D" indica a convergência da função de distribuição da estatística do lado esquerdo para a função de distribuição do lado direito, convergência essa que ocorre com o aumento ilimitado do tamanho da amostra.

É preciso assinalar que o teorema 2 engloba a estatística do teste de significância global como caso particular.

As demonstrações de ambos os teoremas fogem ao escopo de um curso de graduação. Cabe, pois, estabelecer que em ambos os casos o ponto de partida é o Teorema do Limite Central, enunciado a seguir, o qual estabelece a convergência em termos da função de distribuição de probabilidades – ou, simplesmente, "convergência em distribuição" - para uma soma de variáveis aleatórias.

Teorema do Limite Central[3]

Seja $X_1,X_2,\ldots,X_N$ uma amostra aleatória de tamanho N em que $E[X_i] = E[X_j]$, i, j = 1,...,N, $V[Xi] = \sigma^2 < \infty$. É sabido que $V(\sum_{i=1}^{N} X_i) = \sigma\sqrt{N}$. Então:

$$\frac{\sum_{i=1}^{N} X_i - NE\left[X_i\right]}{\sigma\sqrt{N}} \to^{D} N(0,1)$$

Em que N(0,1) é a FD normal padrão (notar que (i) $\sum_{i=1}^{N} X_i$ é a estatística total na amostra e, (ii) na população, a estatística total é $NE[X_i]$).

Sob o teorema do limite central e as hipóteses MCRL1-4, tem-se:

$$\hat{\beta}_{MQO} \to^{D} N\left(\beta, plim\left[\left(X'X\right)^{-1} V\right]\right)$$

Sendo V a matriz de variância-covariância do termo de perturbação. Adotando-se, adicionalmente, as hipóteses MCRL5-6, tem-se que $V = \sigma^2 I_{K+1}$, sendo I_{K+1} a matriz

[3] Transcrito de Magalhães, M.N., Probabilidade e Variáveis Aleatórias, 2006, Edusp.

identidade (quadrada) com K+1 linhas (e K+1 colunas). Em que σ^2 é a variância do termo de perturbação. Ou seja, sob MCRL1-6:

$$\hat{\beta}_{MQO} \to^{D} N\left(\beta, \sigma^2 \, plim\left[\left(X' X\right)^{-1}\right]\right)$$

Ou, analogamente, utilizando a notação de Wooldridge para ao estimador do j-ésimo coeficiente, tem-se:

$$\sqrt{N}\left(\hat{\beta}_j - \beta_j\right) \to^{D} N\left(0, \frac{\sigma^2}{a_j^2}\right)$$

Em que $a_j^2 = plim\,¿$ é a variância assintótica $\sqrt{N}\left(\hat{\beta}_j - \beta_j\right)$,sendo $\hat{r}_{ij}^2$ o i-ésimo resíduo da regressão da j-ésima explicativa contra as demais explicativas.

Um último elemento referido no capítulo 5 de Wooldridge é o estimador da variância do estimador do j-ésimo coeficiente, sendo ele:

$$\widehat{Var}\left(\hat{\beta}_j\right) = \frac{\hat{\sigma}^2}{SQT_j\left(1 - R_j^2\right)}$$

Em que $\hat{\sigma}^2$ é o estimador da variância do termo de perturbação, SQTj é a soma dos quadrados dos desvios da j-ésima explicativa em relação à própria média, e R_j^2 é o coeficiente de determinação da regressão da j-ésima explicativa contra as demais explicativas. Segundo o argumento do autor, o estimador acima converge para zero com o aumento ilimitado da amostra, o que está de acordo com a definição de consistência. Basta notar que $plim\left(\hat{\beta}_{MQO}\right) = \beta$ significa que todos os estimadores, que são variáveis aleatórias, convergem para constantes, o que apenas pode ocorrer com a variância dos estimadores convergindo para zero. Como o estimador da variância acima é consistente, ele também tem de convergir para zero. Deve ser também mencionado que $\hat{\sigma}^2 = SQR/\left(N - K - 1\right)$ é um estimador consistente para a variância do termo de perturbação (a última sendo $E[u^2]$).

2 O teste do multiplicador de Lagrange

O teste de restrições de exclusão pode ser realizado a partir de uma estatística alternativa, denominada multiplicador de Lagrange. O resultado obtido coincide, assintoticamente, i.e., em amostra suficientemente grande, ao atingido com a estatística F de Snedecor.

O nome do teste se deve a que a estatística é, formalmente, obtida a partir da maximização de uma função de máxima verossimilhança associada ao modelo restrito, i.e., sujeito à hipótese nula – os detalhes não serão expostos por estarem fora do escopo da disciplina[4]. Na versão apresentada a seguir, considera-se uma forma alternativa de calcular a estatística.

<u>Primeiro passo: definição do teste</u>

As hipóteses do teste de restrições de exclusão referente às explicativas com índice k a k + s são:

$$H_0: \beta_{k+1} = 0 \text{ e } \beta_{k+2} = 0 \text{ e ... e } \beta_{k+s} = 0 \text{ vs } H_1: \beta_{k+1} \neq 0 \text{ ou } \beta_{k+2} \neq 0 \text{ ou...ou } \beta_{k+s} \neq 0$$

<u>Segundo passo: estatística do teste, distribuição e cálculo do valor observado</u>

A estatística do teste é $LM = N.R^2\_Ru$, em que N é o tamanho amostral e R^2_Ru é o R2 da regressão em que a variável dependente é o resíduo da regressão irrestrita e todas as variáveis independentes são utilizadas. O que significa que é preciso realizar um procedimento de dois estágios para obter a estatística:

1. Estimar a regressão restrita, a qual exclui as variáveis-objeto do teste e obter os resíduos;
2. Estimar a regressão dos resíduos obtidos no passo anterior contra todas as variáveis explicativas.

(Notar, pois, que a regressão irrestrita não é necessária).

O fundamento do teste está em que se as variáveis-alvo são de fato irrelevantes, então elas não podem estar contidas no termo de perturbação da regressão restrita. Cabe recordar que tal termo contém todos os fatores relevantes para explicar a variável dependente omitidos da regressão restrita. Uma maneira de verificar, pois, se tais fatores de fato não são captados pelo termo de perturbação da regressão restrita é verificando se a relação entre os resíduos da regressão restrita e as variáveis-alvo é nula. O que é feito no segundo estágio. Como afirma Wooldridge, é necessário, por "razões técnicas e irrelevantes" incluir também no segundo estágio também as variáveis explicativas contidas na regressão restrita. Uma medida do grau de relação entre resíduos da regressão restrita e as explicativas-alvo do teste é o coeficiente de determinação da regressão de segundo estágio (este sendo a parcela da variação dos resíduos explicada pelas variáveis-alvo).

A estatística LM tem distribuição qui-quadrado com um número de graus de liberdade equivalente ao número de variáveis-alvo do teste, i.e., $LM \sim \chi^2(s)$.

<u>Terceiro passo: região crítica e p-valor</u>

A região crítica é dada por $RC[5\%] = \{\chi_0,\infty\}$, tal que $P(\chi_s > \chi_0) = 0.05$.

O p-valor é calculado de maneira coerente, considerando-se apenas a cauda superior, $P(\chi_s > LM)$.

[4] Os interessados encontrarão os detalhes em Wooldridge, "Econometric Analysis of Cross Section and Panel Data", seção 12.6.2, e Greene, "Econometric Analysis", apêndice E, p.940.

Quarto passo: decisão

O critério de decisão padrão se aplica, i.e., a hipótese de irrelevância populacional do conjunto de variáveis deve ser rejeitado sempre que o valor observado da estatística pertencer à região crítica ou o p-valor for inferior ao nível de significância do teste (5%).

A decisão tomada no teste LM deve ser equivalente à embasada no teste F e o p-valor de ambos deve ser aproximadamente igual em amostras grandes.

3 Eficiência assintótica

Por eficiência assintótica se entende que, com o aumento ilimitado da amostra, a variância do estimador de MQO converge para o menor valor possível. Ou, de maneira mais rigorosa, a variância assintótica do estimador de MQO, entendida com a variância-limite, é a menor possível entre todos os estimadores de um determinado grupo. A demonstração rigorosa foge ao escopo da disciplina e Wooldridge apresenta apenas uma discussão simplificada. Cabe explicar a ideia central de tal discussão. Eficiência é uma propriedade relativa, i.e., estabelecida com base na comparação do estimador-alvo com outros estimadores do mesmo grupo. Por exemplo, pode-se comparar o estimador de MQO, $\hat{\beta}_{MQO\square}$ com um estimador linear (nos parâmetros) genérico alternativo, $\tilde{\beta}_{MQO\square}$. É exatamente o que Wooldridge faz na demonstração simplificada, tomando um estimador cuja fórmula contém, no numerador, ao invés da covariância entre x e y, a covariância entre uma função de x e y. No caso, considera-se uma regressão simples e foca-se o estimador do coeficiente, $\hat{\beta}_{MQO1}$. No segundo passo, o autor demonstra que o estimador alternativo possui variância assintótica no mínimo equivalente à do estimador de MQO, com base em uma identidade estatística denominada desigualdade de Cauchy-Schwartz. Na demonstração, Wooldridge considera como variância assintótica a variância da distribuição normal para a qual a função $\sqrt{N}\left(\hat{\beta}_{MQO1}-\beta_1\right)$ do estimador de MQO converge assintoticamente. Essa é talvez a definição mais comum de variância assintótica.