AUTOCORRELAÇÃO

PROF. DR. VASCONCELOS REIS WAKIM

Autocorrelação

- A autocorrelação serial dos resíduos é outro problema importante que recai sobre o Modelo Clássico de Regressão Linear (MCRL) e que pode tornar os betas estimados em inconsistentes e não eficientes.

- As hipóteses do MCRL não podem ser violadas

- 3 tipos dados: Cross-Section, Painel e Séries Temporais

Autocorrelação

- Para cada tipo de dados → pode haver problemas diferentes

- No que tange ao problema de autocorrelação serial dos resíduos este irá recair basicamente sobre os dados de séries temporais

Autocorrelação

- No caso de cross-section → a base de dados é aleatório (unidades de análises diferentes) → o erro de uma ud pode estar relacionado com o de outra ud de análise;

- Neste caso, Gujarati e Porter (2011) mencionam que se chama de **correlação espacial**;

- Pois refere-se a correlação espacial e não de tempo

Autocorrelação

- Veremos então como assuntos em autocorrelação:

  1) Natureza da autocorrelação;

  2) Consequências;

  3) Detecção; e

  4) Correção

Natureza da Autocorrelação

- Como acontece com os estimadores na presença de heterocedasticidade, quando existe o problema da autocorrelação, os betas estimados serão lineares, não tendenciosos e o termo de erro será iid (independente e identicamente distribuídos), no entanto, não apresentará variância constante e/ou mínima ($Var\,(u_i) = \sigma_i^2$), logo os estimadores serão ineficientes.

$$Cov(u_i, u_j | X_i X_j) = 0 \;\; \forall\, i \neq j$$

$$Cov(u_i, u_j | X_i X_j) \neq 0 \; \forall\, i \neq j$$

Demonstrando que E(u) é não autocorrelacionado

$$E(uu') = \begin{vmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{vmatrix} x \begin{vmatrix} u_1 & u_2 & u_3 & \dots & u_n \end{vmatrix}$$

$$E(uu') = \begin{vmatrix} u_1^2 & u_1 u_2 & u_1 u_3 & \dots & u_1 u_n \\ u_2 u_1 & u_2^2 & u_2 u_3 & \dots & u_2 u_n \\ u_3 u_1 & u_3 u_2 & u_3^2 & \dots & u_3 u_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_n u_1 & u_n u_2 & \cdots & \cdots & u_n^2 \end{vmatrix}$$

Demonstrando que E(u) é não autocorrelacionado

$$E(uu') = \begin{vmatrix} E(u_1^2) & E(u_1u_2) & E(u_1u_3) & \dots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & E(u_2u_3) & \dots & E(u_2u_n) \\ E(u_3u_1) & E(u_3u_2) & E(u_3^2) & \dots & E(u_3u_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & \cdots & E(u_n^2) \end{vmatrix}$$

Matriz de Variância e Covariância do Termo de Erro

$$E(uu') = \begin{vmatrix} \sigma^2 & Cov(u_1u_2) & Cov(u_1u_3) & \dots & Cov(u_1u_n) \\ Cov(u_2u_1) & \sigma^2 & Cov(u_2u_3) & \dots & Cov(u_2u_n) \\ Cov(u_3u_1) & Cov(u_3u_2) & \sigma^2 & \dots & Cov(u_3u_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Cov(u_nu_1) & Cov(u_nu_2) & \cdots & \cdots & \sigma^2 \end{vmatrix}$$

Demonstrando que E(u) é não autocorrelacionado

$$E(uu') = \begin{vmatrix} \frac{\sigma^2}{\sigma^2} & \frac{Cov(u_1u_2)}{\sigma^2} & \frac{Cov(u_1u_3)}{\sigma^2} & \cdots & \frac{Cov(u_1u_n)}{\sigma^2} \\ \frac{Cov(u_2u_1)}{\sigma^2} & \frac{\sigma^2}{\sigma^2} & \frac{Cov(u_2u_3)}{\sigma^2} & \cdots & \frac{Cov(u_2u_n)}{\sigma^2} \\ \frac{Cov(u_3u_1)}{\sigma^2} & \frac{Cov(u_3u_2)}{\sigma^2} & \frac{\sigma^2}{\sigma^2} & \cdots & \frac{Cov(u_3u_n)}{\sigma^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{Cov(u_nu_1)}{\sigma^2} & \frac{Cov(u_nu_2)}{\sigma^2} & \cdots & \cdots & \frac{\sigma^2}{\sigma^2} \end{vmatrix}$$

Demonstrando que E(u) é não autocorrelacionado

$$\Omega = \begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \cdots & \rho_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \cdots & 1 \end{vmatrix}$$

$$E(uu') = \sigma^2 \Omega,$$

COM ESTA DEMONSTRAÇÃO, É POSSÍVEL PERCEBER QUE O TERMO DE ERRO NÃO É AUTOCORRELACIONADO

Origem da Autocorrelação

Origem da Autocorrelação

- Qual é a origem da autocorrelação? → resposta básica → nos próprios dados

- Mas existem algumas situações mais específicas que podem ajudar a entender a origem da autocorrelação:

Origem da Autocorrelação

- Inércia da série temporal (lentidão)

- Viés de especificação de modelo (exclusão ou omissão de variável importante no modelo);

$$Y_i = \gamma_o + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \gamma_4 X_4 + \gamma_5 X_5 + \vartheta_i \longrightarrow \text{Modelo original}$$

$$Y_i = \gamma_o + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_4 X_4 + u_i \longrightarrow \text{Modelo estimado}$$

$$u_i = \gamma_3 X_3 + \gamma_5 X_5 + u_t \longrightarrow Cov(u_i, u_j | X_i X_j) \neq 0$$

Origem da Autocorrelação

- Forma funcional incorreta;
- Modelos em Diferenças

$$\begin{cases} Y_t = \beta_1 + \beta_2 X_t + u_t & (1) \\ Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1} & (2) \end{cases}$$

$$\Delta Y_t = 0 + \beta_2 X_t + X_{t-1} + u_t - u_{t-1}$$

$$\Delta Y_t = \beta_2 \Delta X_t + \Delta u_t$$

$$\Delta u_t = u_t - u_{t-1}$$

WAKIM

Origem da autocorrelação

- É importante frisar que na presença da autocorrelação os betas estimados serão lineares, não tendenciosos, com distribuição normal, no entanto, apresentaram a característica de ineficiência, pois não possuirão variância mínima.

Detecção da Autocorrelação

Detecção da autocorrelação

- A autocorrelação prova → betas estimados ineficientes
- Métodos Informais e Formais
- Informais → forma gráfica
- Formais → testes de detecção

Detecção da autocorrelação

- **Durbin-Watson (DW)** - foi o primeiro teste formal desenvolvido para testar a existência da autocorrelação utilizando os resíduos do MQO. Tsay (2005) comenta que o Teste de DW é muito utilizado, pois este considera uma defasagem da correlação serial.

- Mils (2019) comenta que a autocorrelação ocorre entre os valores atuais (X_t), com os valores futuros (X_{t+1}), ou mesmo com os valores defasados (X_{t-k}), sendo que k varia de 1 a n.

Detecção da autocorrelação

- Matematicamente:

$$d = \frac{\sum_{t=2}^{T}(u_t - u_{t-1})^2}{\sum_{t=1}^{T} u_t^2}$$

- Pressupostos do teste:
  - A regressão tem que ter intercepto;
  - Os X's tem que ser fixo;
  - Ui tem que ter distribuição normal;
  - O termo de erro é gerado por um processo AR(1)
  - A regressão não pode incluir o Y como variável explicativa defasada Y_{t-1}
  - Não pode faltar observações

Como apresenta limitações, não é muito utilizado!

Detecção da autocorrelação

- **Breusch-Godfrey (BG)**: é considerado um teste mais geral e mais indicado para a detecção da autocorrelação.

- Este teste parte de um modelo geral → $Y_t = \beta_1 + \beta_2 X_t + u_i$

- Sendo que u_i → um vetor gerado por um processo AR(p)

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + \delta_t$$

δ_t ➡ É considerado um ruído branco → **estacionário**

Detecção da autocorrelação

- **Rotina para estimação do teste BG manualmente:**
  - Estimar a regressão principal;
  - Obter os resíduos e estimar uma regressão auxiliar tendo o erro estimado como Y
  - $\hat{u}_t = \alpha_1 + \alpha_2 X_t + \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + \gamma_t$
  - Deve-se inserir todas as variáveis do modelo básico acrescentando o termo de erro defasado e obtenha o R^2 da regressão auxiliar;

  - A hipótese nula (H_0)do teste = $\rho_1 = \rho_2 = \cdots = \rho_p = 0$
  - Se a amostra é considerada grande → quando tende ao infinito

$(n-p)R^2 \approx \chi_p^2$ Se $\chi_p^2\ caclulado > \chi_p^2\ tabelado$ rejeita-se H_0

H_0 → ausência de autocorrelação
H_1 → presença de autocorrelação

Detecção da autocorrelação

- **Rotina para estimação do teste BG no Stata**
- Estimar modelo básico → **reg Y X_1 X_2 X_3 X_4**
- Obter o resíduo ($\hat{u}_t$) → **predict u, r**
- Defasar o $\hat{u}_t$ em t-1 períodos
  - **gen lagU_{t-1}= u[_n-1]**
  - **gen lagU_{t-2}= u[_n-2]**
  - **gen lagU_{t-3}= u[_n-3]**
  - **gen lagU_{t-4}= u[_n-4]**

Detecção da autocorrelação

- **Rotina para estimação do teste BG no Stata**
- Estimar regressão auxiliar
- **reg u X_1 X_2 X_3 X_4 lagU_{t-1} lagU_{t-2} lagU_{t-3} lagU_{t-4}**
- Obter o R^2 da regressão auxiliar
- $BG = (n-p)R^2$
- H_0 – ausência
- H_1 - presença

$$BG > \chi^2_{tab} - rejeita\, H_0$$

$$BG < \chi^2_{tab} - aceita\, H_0$$

Se p-value → significativo → rejeita H_0

Comando no Stata → estat bgodfrey, lag(n) small
N = nº de defasagens

Correção da Cutocorrelação

Correção da Autocorrelação

- Relembrando → na presença de autocorrelação os betas são ineficientes

- Tem-se 2 situações para correção da autocorrelação: ρ conhecido ou ρ desconhecido

- Modelo Geral: $Y_t = \beta_1 + \beta_2 X_t + u_t$

Correção da Autocorrelação

- ρ conhecido → $Y_t = \beta_1 + \beta_2 X_t + u_t$, considerando que u_t segue um processo de geração AR(1)

- Se a equação $Y_t = \beta_1 + \beta_2 X_t + u_t$ for verdadeira no tempo t, ela também será verdadeira no t-1

- $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$

Correção da Autocorrelação

- $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$ → multiplicando ambos os lados por ρ → obtêm-se

- $\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1}$ → subtraindo esta expressão de $Y_t = \beta_1 + \beta_2 X_t + u_t$

Correção da Autocorrelação

- $\begin{cases} \rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1} \\ Y_t = \beta_1 + \beta_2 X_t + u_t \end{cases}$

- $\rho Y_{t-1} - Y_t = \beta_1(1-\rho) + \beta_2(1+\rho X_{t-1}) + \xi_t$

- Sendo que: $\xi_t = u_t - \rho u_{t-1}$

Correção da Autocorrelação

- Logo: $\rho Y_{t-1} - Y_t = \beta_1(1-\rho) + \beta_2(X_t - \rho X_{t-1}) + \xi_t$ pode ser rescrita como:
- A 1º diferença pode ser escrita como:
- $Y_t^* = \beta_1^* + \beta_2^* X_t^* + \theta_t^*$
- $Y_t^* = \rho Y_{t-1} - Y_t$
- $\beta_1^* = \beta_1(1-\rho)$
- $\beta_2^* = \beta_2$
- $X_t^* = X_t - \rho X_{t-1}$

Desta forma, estima-se o MQO normalmente e os betas estimados serão **BLUE**.

No entanto, a 1ª diferença pode levar a perda de 1 observação de Y e X, para contornar, usa-se a correção de Prais-Winsten

$Y_1\sqrt{1-\rho^2}$ $X_1\sqrt{1-\rho^2}$

Correção da Autocorrelação

- ρ desconhecido → ρ raramente é conhecido
- Método da 1ª diferença
- $\begin{cases} Y_t = \beta_1 + \beta_2 X_t + u_t \\ Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1} \end{cases}$
- $\Delta Y_t = \beta_2 \Delta X_t + \Delta u_t$ → a 1ª ≠ não tem intercepto → passa pela origem a reta

Correção da Autocorrelação

- Método de **Newey-West** → corrige os erros padrões → contorna heterocedasticidade e autocorrelação simultaneamente.

- **No stata:** newey Y X_1 X_2 X_3 $X_{4,}$ lag(0) → é equivalente ao uso do **vce(robust)**

- **Newey-West** propõe uma matriz de correção para a Var-cov($\hat{\beta}$) que corrige heterocedasticidade e autocorrelação simultaneamente.

$$Var(\hat{\beta}) = (X'X)^{-1}X'\sigma^2\Omega(X'X)^{-1}$$

fim