Testes Complementares na Análise de Resíduos

Prof Silvia Nagib Elian
Sala 215 - Bloco A

Instituto de Matemática e Estatística
Universidade de São Paulo

Agenda

1. Teste do Sinal
2. Teste da Aleatoriedade
3. Teste de Durbin-Watson
4. Teste de homocedasticidade de Goldfeld e Quandt
5. Referências adicionais

Teste do Sinal

H_0: $\mathrm{P}(+) = \mathrm{P}(-) = 1/2$ $\quad$ $+$: resíduo positivo
H_a: $\mathrm{P}(+) \neq \mathrm{P}(-)$ $\quad$ $-$: resíduo negativo

T $=$ número de resíduos positivos
n$'$ $=$ n - número de resíduos nulos

RC $= \{0, 1, \ldots, t, n'\text{-}t, \ldots, n'\}$

$$P(W \leqslant t) \leqslant \frac{\alpha}{2} \qquad W \sim \text{Binomial}(n', 1/2)$$

Para $n > 30$, usar a aproximação da Binomial pela normal.

Teste da Aleatoriedade

H_0: Há aleatoriedade

n_1 = número de resíduos positivos
n_2 = número de resíduos negativos
u = número de sequências de resíduos de mesmo sinal

$$n_1 = 2 \quad n_2 = 4 \quad u = 2 \quad + + - - - -$$

$$n_1 = 2 \quad n_2 = 4 \quad u = 4 \quad + - + - - -$$

i) Se $n_1 \leqslant 10$ e $n_2 \leqslant 10$, usar o teste exato

x	3	4
5	2 ⊢⊣ 7	
6		

$2 \leqslant u \leqslant 7$ aceita-se H_0

Tabela: Bussab e Severo, página 12.

ii) Caso contrário, usar a aproximação normal

$$\mu = \frac{2n_1 n_2}{n_1 + n_2} + 1 \qquad \sigma^2 = \frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

Sob H_0, $Z = \frac{u-\mu+1/2}{\sigma} \approx N(0,1)$

Teste de Durbin-Watson para detectar auto-correlação entre resíduos

A principal fonte de auto-correlação é tomar observações sequencialmente no tempo.

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + \epsilon_t, \ \ t = 1, \ldots, n.$$

$E(\epsilon_t \epsilon_{t+s}) \neq 0$ para $s > 0 \Rightarrow \epsilon_t$ e ϵ_{t+s} são v.a. correlacionadas.

$\text{Var}(\boldsymbol{\epsilon}) = \sigma^2 V, \quad V$ não diagonal.

Gráfico de resíduos (na ordem do tempo) são úteis para detectar autocorrelação.

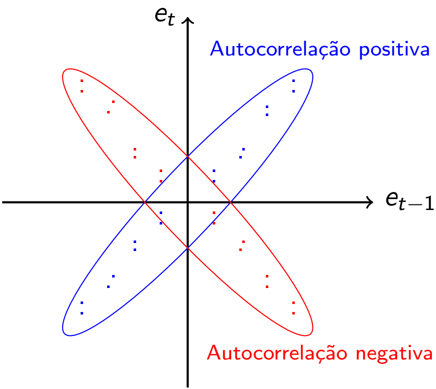
Autocorrelação positiva: resíduos de sinais idênticos ocorrem em "clusters".

$$+ + + + - - - - + + + + - - - -$$

Autocorrelação negativa: resíduos alternam sinais rapidamente.

$$+ - + - + - + - + - + - + - + -$$

Para verificar existência de autocorrelação serial de 1ª ordem (s=1)



$(e_1,\, e_2)$
$(e_2,\, e_3)$
$\vdots$
$(e_{n-1},\, e_n)$

$\downarrow$

r

$$r = \hat{\rho}(\epsilon_{t-1}, \epsilon_t) = \frac{\sum_{t=2}^{n} e_{t-1} e_t}{\sqrt{\sum_{t=1}^{n-1} e_t^2 \sum_{t=2}^{n} e_t^2}}$$

Frequentemente, supõe-se que:

- $\epsilon_t = \rho\epsilon_{t-1} + v_t$
- $E(v_t) = 0$
- $\text{Var}(v_t) = \sigma^2$
- $E(v_t v_s) = 0, \quad t \neq s$

$$\Downarrow$$

Modelo auto Regressivo de Ordem 1, AR(1)

Verifica-se que $\rho(\epsilon_{t-1}, \epsilon_t) = \rho$.

O teste de Durbin-Watson é apropriado para detectar auto correlação neste caso.

Teste de Durbin-Watson

$$\epsilon_t = \rho\epsilon_{t-1} + v_t$$

$H_0 : \rho = 0$
$H_a : \rho \neq 0.$

$$d = \frac{\displaystyle\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=}^{n} e_t^2}$$

Tabelas 3.2, 3.3 e 3.4 Draper e Smith 2ª ed, $\alpha = 0,05$, $\alpha = 0,025$, $\alpha = 0,01$ pág 164, 165, 166.

Teste de Durbin-Watson

1. $H_a : \rho > 0$
   $d < d_L \Rightarrow$ Rejeitar H_0
   $d > d_U \Rightarrow$ Não rejeitar H_0
   $d_L \leqslant d \leqslant d_U \Rightarrow$ Teste é inconclusivo
2. $H_a : \rho < 0$
   $4 - d < d_L \Rightarrow$ Rejeitar H_0
   $4 - d > d_U \Rightarrow$ Não rejeitar H_0
   $d_L \leqslant 4 - d \leqslant d_U \Rightarrow$ Teste é inconclusivo

- A validade do teste depende da validade da condição $\epsilon_i \sim N(0, \sigma^2), \forall i$
- Existe um teste modificado para o caso em que o resultado é inconclusivo. (Draper e Smith, pág. 167)

Teste de homocedasticidade de Goldfeld e Quandt

1. Ordenamos as observações de acordo com a magnitude de X_j (com a qual supomos que σ^2 está associado).
2. Selecionamos um número arbitrário c de observações centrais que omitimos da análise.
3. Ajustamos retas de regressão a cada um dos grupos de $\frac{n-c}{2}$ observações. Sejam

   S_1 - SSE do grupo com menores valores de X_j

   S_2 - SSE do grupo com maiores valores de X_j

   $$F^* = \frac{S_2/(\frac{n-c}{2} - 2)}{S_1/(\frac{n-c}{2} - 2)} = \frac{S_2}{S_1}$$

   Rejeita-se a hipótese de homocedasticidade para

   $$F^* \geqslant F_c\left(\frac{n-c}{2} - 2, \frac{n-c}{2} - 2\right)$$

   c grande $\Rightarrow$ o teste é pouco poderoso.

   c pequeno $\Rightarrow$ as variabilidades residuais estarão mais próximas e o ganho em poder pode ser obscurecido.

Teste de Glejser

- Construímos a regressão de Y em função das variáveis independentes e obtemos os resíduos e_i
- Obtemos a regressão de $|e_i|$ em função da variável X_j com a qual imaginamos que σ^2 está associado.
  Sugestões:
  - $|\hat{e}| = \hat{\beta}_0 + \hat{\beta}_1 x_j$
  - $|\hat{e}| = \hat{\beta}_0 + \hat{\beta}_1 \frac{1}{x_j}$
  - $|\hat{e}| = \hat{\beta}_0 + \hat{\beta}_1 \sqrt{x_j}$

  Testamos a significância de β_0 e β_1 ($\beta_1 \neq 0$ indica heterocedasticidade.)

  A vantagem deste teste é que, se σ^2 não é constante, fornece uma relação entre σ^2 e X_j.

Tabela com valores significantes de d_L e d_U, $\alpha = 0,05$.

	k=1		k=2		k=3		k=4		k=5	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

Tabela: Fonte: Durbin e Watson, "Testing for serial correlation in least squares regression II", Biometrika, 38, 1951, 159-178.

Referências adicionais

Livro Wooldridge - Introductory Econometrics, 3ª ed.

- Teste para independência dos erros:
  - Para checar autocorrelação serial dos erros com o modelo AR(q),
    - Pág. 422.
- Outros testes de heterocedasticidade:
  - Teste de White
    - Pág. 282.