

Testy diagnostyczne

Natalia Nehrebecka
Stanisław Cichocki

Wykład 10

Plan wykładu

1. Testy diagnostyczne
2. Testowanie prawidłowości formy funkcyjnej modelu
3. Testowanie normalności składników losowych
4. Testowanie stabilności parametrów
5. Testowanie heteroskedastyczności

Plan wykładu

1. Testy diagnostyczne
2. Testowanie prawidłowości formy funkcyjnej modelu
3. Testowanie normalności składników losowych
4. Testowanie stabilności parametrów
5. Testowanie heteroskedastyczności

Testy diagnostyczne

- Służą do weryfikacji założeń KMRL
- Sprawdzenie założeń KMRL jest ważne \longrightarrow na nich opierają się własności estymatorów MNK
- Jeśli któreś z założeń nie jest spełnione \longrightarrow należy zastanowić się nad przeformułowaniem modelu lub zastosować bardziej zaawansowane narzędzia ekonometryczne
- Testy są stosowane po wyestymowaniu modelu

Testy diagnostyczne

- W praktyce do testowania jednego założenia KMRL używa się często kilku testów
- Czasami różne testy zastosowane do testowania tej samej hipotezy zerowej dają sprzeczne wnioski

Każdy test należy rozpatrywać następująco:

- ▶ Nazwa testu
- ▶ Hipotezy
- ▶ Jakie założenie KMRL nie jest spełnione przy odrzuceniu H_0
- ▶ Jakie są skutki niespełnienia założenia KMRL
- ▶ W jaki sposób można rozwiązać problemy zasygnalizowane przez wynik testów?

Plan wykładu

1. Testy diagnostyczne
2. Testowanie prawidłowości formy funkcyjnej modelu
3. Testowanie normalności składników losowych
4. Testowanie stabilności parametrów
5. Testowanie heteroskedastyczności

Testowanie prawidłowości formy funkcyjnej

Test RESET (*Regression Specification Error Test*):

$$H_0 : y_i = x_i\beta + \varepsilon_i \quad - \text{liniowa postać modelu}$$

$$H_1 : y_i = f(x_i\beta) + \varepsilon_i \quad - \text{nieliniowa postać modelu}$$

gdzie $f(\bullet)$ jest nieliniowa

Procedura przeprowadzenia testu RESET

Krok 1: estymujemy współczynniki regresji w modelu:

$$y_i = x_i\beta + \varepsilon_i$$

Krok 2: uzyskujemy wartości dopasowane:

$$\hat{y}_i = x_i b$$

Krok 3: przeprowadzamy regresję pomocniczą:

$$y_i = x_i\beta + \alpha_1 \hat{y}_i^2 + \dots + \alpha_p \hat{y}_i^{p+1} + u_i$$

Testowanie prawidłowości formy funkcyjnej

Krok 4: testujemy łączną nieistotność zmiennych $\hat{y}_i^2, \dots, \hat{y}_i^{p+1}$

- ▶ **Hipoteza zerowa** oznacza łączną nieistotność tych zmiennych, implikuje poprawność formy funkcyjnej przyjętej przez nas w regresji wyjściowej.
- ▶ za pomocą *testu F* testujemy H_0 :

$$H_0 : \alpha_1 = \dots = \alpha_p = 0$$

- Statystyka testowa: F-Snedecora (**w dużych próbach**)
- Statystyka krytyczna: $F(p, N-K-p)$

Testowanie prawidłowości formy funkcyjnej

▶ Druga forma:

$$e_i = x_i\beta + \alpha_1\hat{y}_i^2 + \dots + \alpha_p\hat{y}_i^{p+1} + \eta_i$$

- Test $LM = NR^2 \xrightarrow{D} \chi^2_p$

Jakie założenie KMRL nie jest spełnione przy odrzuceniu H_0 ?

- ▶ Związek pomiędzy zmienną zależną a zmiennymi niezależnymi opisany jest równaniem:

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_K x_{Ki} + \varepsilon_i \quad i = 1, 2, 3, \dots, N$$

Jakie są skutki niespełnienia założenia KMRL

Odrzucenie hipotezy zerowej o poprawności przyjętej formy funkcyjnej

- podważa interpretacje ekonomiczną modelu (*interpretacja oszacowanych parametrów*)
- niemożliwe udowodnienie własności estymatora MNK (*nieobciążoność czy efektywność estymatora MNK*)

W jaki sposób można rozwiązać problemy zasygnalizowane przez wynik testu?

Możemy próbować poprawić formę funkcyjną modelu wprowadzając do modelu:

- interakcje między zmiennymi,
- dokonać przekształceń zmiennych,
- zastosować model wielomianowy, schodkowy lub krzywej łamanej.

Przykład

- ▶ Analiza prestiżu wykonywanego zawodu (*siops*)
 - a) wiek (*age*)
 - b) płeć (*sex*, 1 Mężczyzna, 2 kobieta)
 - c) wykształcenie respondenta (*dp* - podstawowe, *ds* – średnie, *dw* –wyższe)
 - d) miejsce zamieszkania respondenta (*size*)
 - e) prestiż wykonywanego przez ojca zawodu (*pasiops* całkowite od 1 do 100.)
 - f) prestiż wykonywanego przez matkę zawodu (*masiops* całkowite od 1 do 100.)
 - g) wykształcenie ojca (*padeq*)
 - h) wykształcenie matki (*madeq*)

Przykład

```
. xi: regress siops age pasiops masiops i.sex ds dw  
i.sex          _Isex_1-2          (naturally coded; _Isex_1 omitted)
```

Source	SS	df	MS	Number of obs = 1232		
Model	50595.8547	6	8432.64244	F(6, 1225)	=	106.59
Residual	96913.6778	1225	79.1132064	Prob > F	=	0.0000
Total	147509.532	1231	119.829027	R-squared	=	0.3430
				Adj R-squared	=	0.3398
				Root MSE	=	8.8946

siops	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0627069	.0192075	3.26	0.001	.0250236	.1003903
pasiops	.08292	.0304436	2.72	0.007	.0231926	.1426474
masiops	.0960088	.0317817	3.02	0.003	.0336562	.1583614
_Isex_2	-1.066076	.5158061	-2.07	0.039	-2.078037	-.0541146
ds	5.08828	.6766924	7.52	0.000	3.760676	6.415885
dw	13.55178	.7347977	18.44	0.000	12.11018	14.99338
_cons	23.4708	1.798279	13.05	0.000	19.94275	26.99885

```
. /*test Ramsey'a RESET na poprawność formy funkcyjnej*/  
. ovtest
```

```
Ramsey RESET test using powers of the fitted values of siops  
Ho: model has no omitted variables  
F(3, 1222) = 1.15  
Prob > F = 0.3264
```

OVTEST: w regresji pomocniczej testowana jest łączna nieistotność wartości dopasowanych podniesionych do 2,3 i 4 potęgi

Pytania teoretyczne

1. Do czego służą testy diagnostyczne?
2. Za pomocą jakiego testu testujemy prawidłowość formy funkcyjnej? Jakiemu założeniu KMRL odpowiada H_0 w tym teście? Jaka jest hipoteza alternatywna w tym teście?

Dziękuję za uwagę