מודלים ושיטות חיזוי

פרק 10 - מבחן LM

תוכן העניינים

1. כללי 1

מבחן LM:

רקע:

במבחן כופלי לגרנג' (LM) אנו בודקים האם משתנה או משתנים מסבירים מסוימים רלוונטיים למודל.

לדוגמא:
נניח שיש לנו מודל הכולל 4 משתנים מסבירים (UNRESTRICTED):
$Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t}$.
לגבי השניים הראשונים אנו בטוחים כי הם רלוונטיים וחייבים להופיע במודל. לגבי השניים האחרונים אנחנו לא בטוחים.
השערות:
$$\begin{matrix} H_0 : \beta_3 = \beta_4 = 0 \\ H_1 : OTHERWISE \end{matrix}$$.

המודל המוגבל (RESTRICTED): $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t}$.
במבחן LM אומדים את המודל המוגבל ומקבלים עבור כל תצפית את הסטייה מקו הרגרסיה: $Y_t - \hat{Y}_t = \hat{u}_t$.
כעת אומדים את רגרסיית העזר שבה מנסים לנבא את הסטייה מקו הרגרסיה עבור כל תצפית: $\hat{u}_t = \delta_0 + \delta_1 x_{1t} + \delta_2 x_{2t} + \delta_3 x_{3t} + \delta_4 x_{4t} + \omega_t$.
חישוב הסטטיסטי: (R^2 של רגרסיית העזר * מספר התצפיות) $LM_{stat} = R^2 \cdot T$.
כלל הכרעה: אם $LM_{stat} > \chi^2_m$ נדחה את H_0 (m = מס' ההגבלות ב-H_0).

- שימו לב כי:
  עבור המשתנים הנוספים למודל – כל המדדים (הבטות, ערכי t וה-$Pvalu$) ברגרסיית העזר שווים לאלו של הרגרסיה הלא מוגבלת.
  עבור המשתנים הקיימים במודל – המדדים אינם שווים בין שתי הרגרסיות.

שאלות:

1) נניח מודל הכולל 4 משתנים מסבירים (UNRESTRICTED):

$$Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} .$$

UNRESTRICTED

Dependent variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	------	646169.84	----------	----------	0.0000
Error	------	620.17			
C Total	203	646790.01			

Root MSE	------	R-square	-----
Dep Mean	----	Adj R-sq	-----
C.V.	-----		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	5.067731	0.456604	11.09874	0.0000
X1	1	0.975726	0.042711	22.84485	0.0000
X2	1	3.005385	0.008679	346.2721	0.0000
X3	1	-5.029101	0.073149	-68.75146	0.0000
X4	1	8.974106	0.029075	308.6485	0.0000

RESTRICTED

Dependent variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	------	646001.81			
Error	------	788.2			
C Total	203	646790.01			

Root MSE	---------	R-square	-----
Dep Mean	---------	Adj R-sq	-----
C.V.	---------		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	7.067731	0.656604	10.76406	0.0000
X1	1	26.36455	0.756627	34.84485	0.0000
X2	1	29.58626	0.076993	384.2721	0.0000

רגרסיית עזר

Dependent variable :RES

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	------	646169.84	--------	-------	0.0000
Error	------	620.17			
C Total	203	646790.01			

Root MSE	---------	R-square	0.213
Dep Mean	---------	Adj R-sq	-------
C.V.	---------		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	5.9608892	0.776604	7.675584	0.0000
X1	1	1.2077723	0.978845	1.233875	0.8455
X2	1	0.4840697	0.886754	0.545889	0.9976
X3	1	-5.029101	0.073149	-68.75146	0.0000
X4	1	8.974106	0.029075	308.6485	0.0000

א. בדוק את הטענה כי לפחות אחד מן המשתנים הנוספים רלוונטי למודל בשתי דרכים.

ב. איזה מן המשתנים הנוספים רלוונטי למודל?

ג. הסבירו את הקשרים בין שלוש המשוואות: R, U ועזר ואת הקשר בין מבחן WALD ומבחן LM.

U: $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} + \hat{\beta}_3 x_{3t} + \hat{\beta}_4 x_{4t} + \hat{v}_t$.

R: $\hat{Y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 x_{1t} + \hat{\alpha}_2 x_{2t} + \hat{u}_t$.

עזר: $\hat{u}_t = \hat{\delta}_0 + \hat{\delta}_1 x_{1t} + \hat{\delta}_2 x_{2t} + \hat{\delta}_3 x_{3t} + \hat{\delta}_4 x_{4t} + \hat{\omega}_t$.

ד. שחזרו בעזרת שתי המשוואות הראשונות (U ו-R) את LM_{stat}.

ה. שחזרו בעזרת המשוואה האחרונה (רגרסיית העזר) את $WALD_{stat}$.

תשובות סופיות:

1) א. מבחן LM ומבחן WALD, יש עדות לכך.

ב. $pt_{\hat{\beta}_3} = pt_{\hat{\beta}_4} = 0.00$.

ג. i. עזר+R=U.

ii. $ESS_U = ESS_y$.

iii. $ESS_R = TSS_y$.

iv. $R^2 = 1 - \frac{ESS}{TSS} = 1 - \frac{ESS_U}{ESS_R} = \frac{ESS_R - ESS_U}{ESS_R}$.

ד. $LM_{stat} = 43.489$.

ה. $WALD_{stat} = 26.962$.