

אקונומטריקה ב

פרק 5 - הטרסקדסטיות

תוכן העניינים

1. השלכות על אומדי OLS.....1
2. מבחנים לזיהוי הטרסקדסטיות.....2
3. פתרון בעיית ההטרסקדסטיות WLS.....6

השלכות על אומדי OLS:

רקע:

הטרנסקדסטיות הוא מצב שבו מופרת הנחת ההומוסקדסטיות, הגורסת כי שונות הטעויות היא אותה שונות עבור כל תצפית ותצפית: $V(u_t) = \sigma^2$ לכל t , כלומר התצפיות מפוזרות באופן אחיד סביב קו הרגרסיה. במצב של הטרנסקדסטיות שונות הטעויות של כל תצפית היא שונה: $V(u_t) = \sigma_t^2$.

בהינתן הטרנסקדסטיות מופרת תכונת היעילות של אומדי הריבועים הפחותים שכן בכדי לחשב שונות יעילה של האומדים השתמשנו בהנחה של שונות קבועה.

מבחנים לזיהוי הטרוסקדסטיות:

רקע:

החשד לקיומה של בעיית הטרוסקדסטיות בנתונים צריך להתעורר כאשר אנו בוחנים את גרף השאריות – באיזה אופן השונות של הטעויות משתנה בין תצפית לתצפית. שיטות לזיהוי הטרוסקדסטיות: מבחן GQ (Goldfeld-Qauandt) ומבחן White. מבחן GQ מניח כי במקום שונות אחת אחידה של הטעויות לכל התצפיות, קיימות שתי שונות שונות בלבד. ואילו מבחן White מניח כי לכל תצפית ותצפית שונות שונה של טעויות.

1. מבחן GQ:

ההנחה העומדת בבסיס מבחן זה היא כי קיימות שתי שונות שונות של טעויות.
ביצוע המבחן:

- מחלקים את המדגם לשני חלקים:
- 1. החלק שבו אנו חושדים שיש שונות גבוהה יותר.
- 2. החלק שבו אנו חושדים שיש שונות נמוכה יותר.
- מקובל להשמיט מס' תצפיות (בין 1/6 ל-1/3) במרכז המדגם.
- אומדים כל אחד מהחלקים בנפרד ומקבלים את ה-ESS של כל חלק.
- מחשבים את הסטטיסטי: $F_{stat} = \frac{ESS_1 / T_1 - K - 1}{ESS_2 / T_2 - K - 1}$ (תמיד השונות הגבוהה חלקי הקטנה).
- סטטיסטי זה מתפלג: $F_{(\alpha; T_1 - K - 1, T_2 - K - 1)}$.
- כלל ההכרעה: אם $F_{stat} > F_C$ אז דוחים את H_0 .
- ההשערות: $H_0: \sigma_1^2 = \sigma_2^2$
 $H_1: \sigma_1^2 > \sigma_2^2$

2. מבחן White:

ההנחה העומדת בבסיס מבחן זה היא כי לכל תצפית ותצפית שונות שונה של טעויות. הביטוי המתמטי של הנחה זו היא היותה של השונות פונקציה ליניארית של כל המשתנים המסבירים, ריבועיהם והאיברים הצולבים:

$$\sigma_t^2 = f(x_j, x_j^2, x_j x_j)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_k x_k + \beta_1 x_k^2 + \dots + \beta_k x_k^2 + \gamma_{12} x_1 x_2 + \gamma_{13} x_1 x_3 \dots$$

האומד ל- σ_t^2 הוא $\hat{u}_t^2$.

המבחן הוא מבחן LM:

- אומדים את המודל המקורי ומקבלים את הסטיות מקו הרגרסיה $\hat{u}_t$ (המכונה בתוכנה הסטטיסטית RES-SAS).
- אומדים את $\hat{u}_t^2$ כפונקציה ליניארית של כל המשתנים המסבירים, ריבועיהם והאיברים הצולבים: $\hat{u}_t^2 / x_j, x_j^2, x_j x_j$: זוהי רגרסיית העזר.
- נחשב את סטטיסטי LM : $LM_{stat} = T_y \cdot R_y^2$.
- אם $LM_{stat} > \chi_m^2$, כאשר $m =$ מס' המשתנים ברגרסיית העזר.
 השערות: $H_0 : \alpha_j = \beta_j = \gamma_{jj} = 0$
 $H_1 : OTHERWISE$

שאלות:

1) חוקר מניח כי מכירות של חנות הן פונקציה של שיטחה, דמי שכירות והאפשרות של מכירת עיתונים.

נסמן:

Y_SALES - מכירות חודשיות (נח).

$X1_SQUARES$ - שטח החנות (מ"ר).

$X2_RENT$ - דמי שכירות (\$).

$PAPERS$ - משתנה איכותי המקבל 1 אם החנות מוכרת גם עיתונים ו-0 אם לא.

החוקר חשד כי קיימת בעיה של הטרוסקדסטיות בנתונים.

החוקר ביצע מבחן לזיהוי הטרוסקדסטיות שתוצאותיו נתונות להלן:

:Dependent Variable					
20	Number of Observations Read				
20	Number of Observations Used				
Analysis of Variance					
< Pr	F Value	Mean Square	Sum of Squares	DF	Source
----	-----	-----	-----	---	Model
		-----	-----	---	Error
			----	---	Corrected Total
0.086942	R-Square	-----			Root MSE
-----	Adj R-Sq	-----			Dependent Mean
Parameter Estimates					
< Pr	t Value	Standard Error	Parameter Estimate	DF	Variable
----	-----	----	-----	1	Intercept
----	-----	-----	-----	1	X1_SQUARE
----	-----	-----	-----	1	X1_SQUARE^2
----	-----	-----	-----	1	X1_SQUARE*X2_RENT
----	-----	-----	-----	1	X2_RENT
----	-----	-----	-----	1	X2_RENT^2
----	-----	-----	-----	1	X2_RENT*PAPERS
--	-----	-----	-----	1	PAPERS

- הרגרסיה המופיעה בפלט לעיל נועדה לבדיקת: _____.
- על ידי מבחן: _____.
- המשתנה התלוי הינו: _____.
- המשתנים הב"ת: _____.
- ההשערות הינן: _____.
- גודל הסטטיסטי למבחן הינו (רשמו תוצאה מספרית): _____.
- המסקנה המתקבלת היא: _____.

תשובות סופיות:

1) קיום הטרוסקדסטיות בנתונים.

$(LM)_{WHITE}$.

RES^2 .

$X1_SQUARE$, $X1_SQUARE^2$, $X1_SQUARE * X2_RENT$,

$X2_RENT$, $X2_RENT^2$, $X2_RENT * PAPERS$, $PAPERS$

$LM_{stat} = 1.73$.

אין עדות לכך.

פתרון בעיית ההטרוסקדסטיות WLS:

רקע:

נניח שאנו רוצים לאמוד את המודל: $Y_t = \alpha + \beta X_t + u_t$ וידוע כי לכל קריזה שונות אחרת. ההנחה שעומדת בבסיס שיטת ה-WLS היא כי השונות המשתנה כוללת בתוכה מרכיב קבוע ומרכיב משתנה: $\sigma_t^2 = Z_t \cdot \sigma^2$ את המרכיב המשתנה בשונות (Z_t) יש לנטרל. לשם כך ניצור משתנה חדש W_t שיהווה השורש ההופכי

$$W_t = \frac{1}{\sqrt{Z_t}} \quad \text{לאותו מרכיב משתנה:}$$

נכפיל כל תצפית במשתנה החדש W_t וניצור משוואה שהיא קומבינציה ליניארית של

$$Y_t = \alpha + \beta X_t + u_t$$

המשוואה המקורית: $Y_t W_t = \alpha \cdot W_t + \beta (X_t W_t) + u_t W_t$

$$\frac{Y_t}{\sqrt{Z_t}} = \alpha \cdot \frac{1}{\sqrt{Z_t}} + \beta \cdot \frac{X_t}{\sqrt{Z_t}} + \frac{u_t}{\sqrt{Z_t}} \quad \text{בצורתה המפורשת המשוואה החדשה נראית כך:}$$

שאלות:

(1) נתון המודל:

$$1. Y_t = \alpha + \beta \cdot X_t + U_t.$$

ונתון כי: $VAR(U_t) = \frac{\sigma^2}{Z_t^2}$ (משתנה ידוע).

א. מהי הבעיה שנוצרת באמידת משוואה (1)?

ב. מהן תכונות אומדי הריבועים הפחותים של משוואה (1)?

כדי לפתור את הבעיה שנוצרה, נאמדה המשוואה הבאה:

$$2. Y_t \cdot W_t = \alpha \cdot W_t + \beta \cdot (X_t \cdot W_t) + U_t \cdot W_t.$$

ג. מהו W_t שבעזרתו ניתן לאמוד את α ו- β בצורה יעילה?ד. מהו האומד היעיל של σ^2 ?ה. האם ניתן להשוות בין המודלים על בסיס R^2 ? אם לא, האם ניתן להחליט בכל זאת איזה מודל טוב יותר?

ו. חוו דעתכם על הטענות הבאות, ונמקו:

i. אם נתון כי: $Z_t = a + b \cdot \bar{X}$, התשובות לסעיפים א' ו-ב' נשארות ללא שינוי.

ii. המשוואה הנורמאלית: $\sum \hat{\varepsilon}_t = 0$ (כאשר: $\varepsilon_t = U_t \cdot W_t$) היא אחת המשוואות הנורמאליות לאמידת משוואה (2).

(2) ענו על השאלה הקודמת, כאשר נתון כי: $VAR(U_t) = \sigma^2 \cdot X_t^2$.(3) נתון המודל: $Y_t = \alpha + \beta X_t + u_t$. וקיים מדגם של 100 תצפיות כאשר נתון

כי: $V(u_t) = \sigma_t^2 = \begin{cases} X_t \sigma^2 & \Leftrightarrow t \geq 50 \\ X_t^2 \sigma^2 & \Leftrightarrow t \leq 50 \end{cases}$ (שאר ההנחות הקלאסיות מתקיימות).

א. במשוואה מס' 1 יש בעיה של: _____.

ב. אמידת משוואה (1) תניב אומדים

בלתי מוטים ועקיבים: נכון/ לא נכון/ לא ניתן לדעת

ג. פתרון הבעיה הקיימת במשוואה (1) ייתכן על ידי אמידת המשוואה

הבאה: $Y_t \cdot W_t = \alpha \cdot W_t + \beta \cdot (X_t \cdot W_t) + u_t$ כאשר: $W_t = \underline{\hspace{2cm}}$

ד. אם נתון כי: $V(u_t) = \sigma_t^2 = \begin{cases} 3\sigma^2 & \Leftrightarrow t \geq 50 \\ \sigma^2 & \Leftrightarrow t \leq 50 \end{cases}$

האם ישתנו תשובותיכם לסעיפים א' ו-ב': כן/לא/לא ניתן לדעת

(4) להלן משוואה המציגה את הקשר בין שנות השכלה X והכנסה Y :

$$Y_t = \alpha + \beta_1 X_t + \beta_2 X_t^2 + u_t$$

כמו כן ידוע שמתקיים הקשר: $\text{var}(u_t) = 64X_t^2$

הריצו מספר רגרסיות וקיבלו את התוצאות הבאות:

$$1. \quad \frac{\hat{Y}_t}{X_t} = 21 + 8 \frac{1}{X_t - 2X_t}$$

$$2. \quad \frac{\hat{Y}_t}{X_t^2} = 13 \frac{1}{X_t} + 3 \frac{1}{X_t^2} - 4X_t$$

$$3. \quad \frac{\hat{Y}_t}{X_t} = 2 - 20 \frac{1}{X_t^2} + 7X_t$$

$$4. \quad \frac{\hat{Y}_t}{8} = 7 + 6 \frac{X_t}{8} - 4 \frac{X_t^2}{8}$$

אומדני ה-OLS עבור α ו- β המבוססים על רגרסיה המקיימת את כל ההנחות הקלאסיות הינם:

$$א. \quad \hat{\alpha} = 8; \hat{\beta}_1 = 21$$

$$ב. \quad \hat{\alpha} = 3; \hat{\beta}_1 = 13$$

$$ג. \quad \hat{\alpha} = 20; \hat{\beta}_1 = 7$$

$$ד. \quad \hat{\alpha} = 7; \hat{\beta}_1 = 6$$

תשובות סופיות:

- (1) א. הטרוסקדסטיות. ב. ראו סרטון. ג. $W_t = \frac{1}{\sqrt{\frac{1}{Z_t^2}}} = Z_t$. ד. $\sigma^2 = \frac{ESS}{T-K}$. ה. לא, המודל השני. ו. i. לא נכון. ii. לא נכון.
- (2) א. הטרוסקדסטיות. ב. ראו סרטון. ג. $W_t = \frac{1}{X_t}$. ד. $S^2 = \frac{ESS}{T-k-1}$. ה. ראו סרטון. ו. ראו סרטון. ז. נכון.
- (3) א. הטרוסקדסטיות. ב. נכון. ג. $W_t = \frac{1}{\sqrt{X_t^2}} \quad t \leq 50, W_t = \frac{1}{\sqrt{X_t}} \quad t \geq 50$. ד. לא.
- (4) א'.