

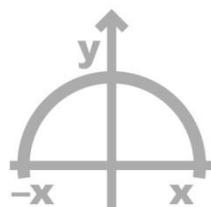
אקונומטריקה



$$\sqrt{2}$$
A grey square icon containing the mathematical expression $\sqrt{2}$, with '1' on both the top-left and bottom-right corners.



$$\{\sqrt{x}\}^2$$
An orange diamond-shaped icon containing the mathematical expression $\{\sqrt{x}\}^2$.



תוכן העניינים

1.	מבוא לקורס	1
7.	אומדי הריבועים הפלטיים	7
21.	מודלים לא ליניאריים	21
25.	מבחני המובהקות וקריאה פלטימ - תוכנת SAS	25
32.	שינוי ייחדות מדידה	32
34.	רגסישה מרובה	34
43.	מבחן ML	43
47.	בעיות ספציפיות	47
48.	תיאוריה מולטיקוליניאריות	48
51.	סיכום ותרגול של בעיות ספציפיות ומולטיקוליניאריות	51
56.	משתנה דמי	56
73.	תיאוריה הפרת ההנחות קלאסיות	73
74.	הטרוסקדסטיות	74
82.	מתאים סדרתי	82
93.	סיכום מתאים סדרתי והטרוקדסטיות	93
94.	מודלים דינמיים	94
100.	סיכום תכונות ארפ	100
(ללא ספר)	WEIVE	
18.	הדרך בקריאה פלטימ של תוכנת WEIVE	18

אקוונומטריקה

פרק 1 - מבוא לקורס

תוכן העניינים

1. כללי

מבוא לקורס:

רקע:

הגדירות וסימוניים:

משתנה אמפירי – תוצאותיו ידועות מראש (למשל: רמת הכנסה, גיל, מס' שנות לימוד במדגם מסוים).

משתנה מקרי – תוצאותיו לא ידועות מראש (כגון תוצאה בהטלה קובייה או בהטלה מטבעה). באקונומטריקה עוסוק בעיקר במשתנים מקרים.

שני סוגי המשתנים יסומנים באותות לועזית עם אינדקס (למשל: Y_t או X).

קבוע – מקבל ערך אחד בלבד (מסומן באותות לועזית ללא אינדקס – למשל a או b).

לכל משתנה מקרי X יש תוחלת המיצגת את מרכז ההתפלגות (μ_X או $E(X)$).

השונות – מייצגת את מידת הפיזור של ההתפלגות (σ^2_X או $V(X)$).

סטיית התקן – היא השורש של השונות (σ_X).

שונות משותפת (covariance) – ממד להתפלגות המשותפת של שני משתנים מקרים ומייצגת את הכיוון של הקשר ביניהם ($Cov(X, Y)$):

$Cov(X, Y) = 0 \Leftrightarrow X, Y$ בלתי מתואמים.

$Cov(X, Y) > 0 \Leftrightarrow$ מתאים חיובי בין המשתנים.

$Cov(X, Y) < 0 \Leftrightarrow$ מתאים שלילי בין המשתנים.

X, Y בלתי תלויים $\Leftrightarrow X, Y$ בלתי מתואמים.

מקדם המתאים של פירסון – ממד לכיוון ולעוצמת הקשר הlieneari בין שני

$$\text{משתנים: } \eta_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y}.$$

$-1 \leq \eta \leq 1$

$\eta = 1$ מתאיםlianeari חיובי מלא בין שני המשתנים.

$\eta = -1$ מתאיםlianeari שלילי מלא בין שני המשתנים.

$\eta = 0$ לא קיים מתאיםlianeari בין שני המשתנים.

אמידה:

פרמטר – ערך המשתנה הנחקר המתאר את כל האוכלוסייה.
סטטיטיסטי/אומד – ערך המשתנה הנחקר המתאר את המדגם.

מדגם	אוכלוסייה
$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$	$E(X) = \mu$
$S_x^2 = \frac{S_{xx}}{n} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$	$V(X) = \sigma^2 = E(X - E(X))^2$
$\frac{S_{xy}}{n} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n}$	$\text{cov}(X, Y) = E(X - E(X))(Y - E(Y))$
$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} \sqrt{S_{yy}}}$	$\eta_{xy} = \frac{\text{cov}(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}}$

נוסחאות וחוקים בסטטיסטיקה:

יהיו X ו- Y משתנים מקרים, ו- a , b קבועים :

חוקי הסיגמה:

$$\cdot \sum_{t=1}^T X_t = X_1 + X_2 + X_3 + \dots + X_T \quad .1$$

$$\cdot \sum_{t=1}^T a = Ta \quad .2 \text{ סכום של קבוע}$$

$$\cdot \sum_{t=1}^T aX_t = a \sum_{t=1}^T X_t \quad .3 \text{ סכום של קבוע כפול משתנה} = \text{ קבוע כפול הסכום}$$

$$\cdot \sum_{t=1}^T (X_t \pm Y)_t = \sum_{t=1}^T X_t \pm \sum_{t=1}^T Y_t \quad .4 \text{ סכום של סכום/הפרש} = \text{ לסכום/הפרש הסכומים}$$

$$\cdot \sum_{t=1}^T X_t^2 \neq \left(\sum_{t=1}^T X_t \right)^2 \quad .5 \text{ יש לשים לב כי :}$$

$$\cdot \sum_{t=1}^T X_t Y_t \neq \sum_{t=1}^T X_t \sum_{t=1}^T Y_t$$

הגדרות ופיתוחים :

1. סכום הסטיות מה ממוצע = 0 : $\sum_{t=1}^T (X_t - \bar{X}) = 0$
2. סכום הסטיות הריבועיות מה ממוצע (МОנה השונות) :

$$\cdot S_{XX} = \sum_{t=1}^T (X_t - \bar{X})^2 = \sum_{t=1}^T X_t^2 - T\bar{X}^2 = \sum_{t=1}^T (X_t - \bar{X})X_t$$

3. מונה של השונות המשותפת :

$$\cdot S_{XY} = \sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y}) = \sum_{t=1}^T X_t Y_t - T\bar{X}\bar{Y} = \sum_{t=1}^T (X_t - \bar{X})Y_t = \sum_{t=1}^T (Y_t - \bar{Y})X_t$$

חוקי התוחלת :

1. תוחלת של קבוע = קבוע : $E(a) = a$

2. תוחלת של סכום/הפרש = לסכום/הפרש התוחלות :

$$\begin{aligned} E(X \pm Y) &= E(X) \pm E(Y) \\ E(\sum(X_i)) &= \sum E(X_i) \end{aligned}$$

3. תוחלת של כפל/חילוק ≠ לכפל/חילוק התוחלות :

$$\begin{aligned} E(X \cdot Y) &\neq E(X) \cdot E(Y) \\ E\left(\frac{X}{Y}\right) &\neq \frac{E(X)}{E(Y)} \\ E(X^2) &\neq [E(X)]^2 \end{aligned}$$

4. השפעת טרנספורמציה ליניארית על התוחלת :

$$\cdot E\left(a / \frac{1}{a} X \pm b\right) = a / \frac{1}{a} \cdot E(X) \pm b$$

חוקי השונות :

1. עבור X ו- Y בלתי תלויים/בלתי מתאימים מתקיים:
שונות של סכום/הפרש = סכום השונות :

$$\begin{aligned} V(X \pm Y) &= V(X) + V(Y) \\ V\sum(X_i) &= \sum V(X_i) \end{aligned}$$

2. עבור X ו- Y תלויים/מתאימים מתקיים :

שונות של סכום/הפרש ≠ סכום השונות :
 $V(X \pm Y) = V(X) + V(Y) \pm 2 \cdot Cov(X, Y)$

$$\begin{aligned} V(a) &= 0 \\ \cdot V(a \pm x) &= V(X) \end{aligned}$$

3. שונות של קבוע = 0 :

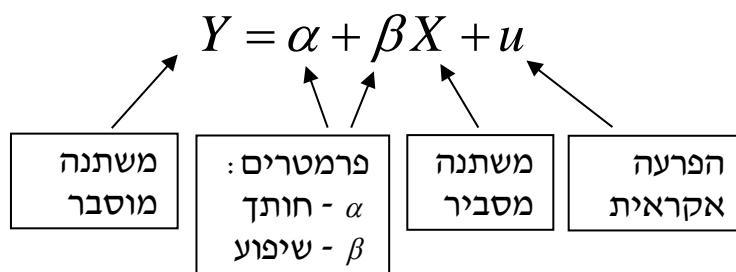
4. השפעת טרנספורמציה ליניארית על השונות : $V(aX + b) = a^2V(X)$

- חוקי התוחלת והשונות מתאימים למשתנים אמפיריים כאלו קבועים (יוצאים מחוץ לתוחלת או לשונות).
- חוקי הסכימה מתאימים למשתנים אמפיריים כמשתנים הנשארים בתוך הסיגמא (רק הקבועים ייצאו מחוץ לסיגמא).

חוקי השונות המשותפת:

1. שונות משותפת בין משתנה קבוע = 0 : $\text{cov}(X, a) = 0$
2. שונות משותפת של משתנים המוכפלים בקבוע : $\text{cov}(aX, bY) = ab \cdot \text{cov}(X, Y)$
3. שונות משותפת של משתנה עם עצמו = שונות המשתנה : $\text{cov}(X, X) = V(X)$

המודל האקונומטרי:



1. במודל : $Y_t = \alpha + \beta X_t + u_t$, α ו- β הם מספרים קבועים אך לא ידועים. אנו יכולים להעריך אותם ולקבל אומדיים (תהליך קבלת האומדיים נקרא אמידה).
2. $\hat{\alpha}$ הוא האומד ל- α ו- $\hat{\beta}$ הוא האומד ל- β .
3. אומדי ריבועים פחותים (אר"פ) הם אומדיים שהושבו בשיטת הריבועים הפחותים. מסומנים בד"כ ע"י 'קובע' - $\hat{\beta}$. אומדיים אחרים מסומנים בד"כ ע"י 'תלטלי' - $\tilde{\beta}$.

4. בעוד α ו- β הם קבועים, $\hat{\alpha}$ ו- $\hat{\beta}$ הם משתנים מקרים כיון שבכל מדגם מתקבלים $\hat{\alpha}$ ו- $\hat{\beta}$ אחרים.

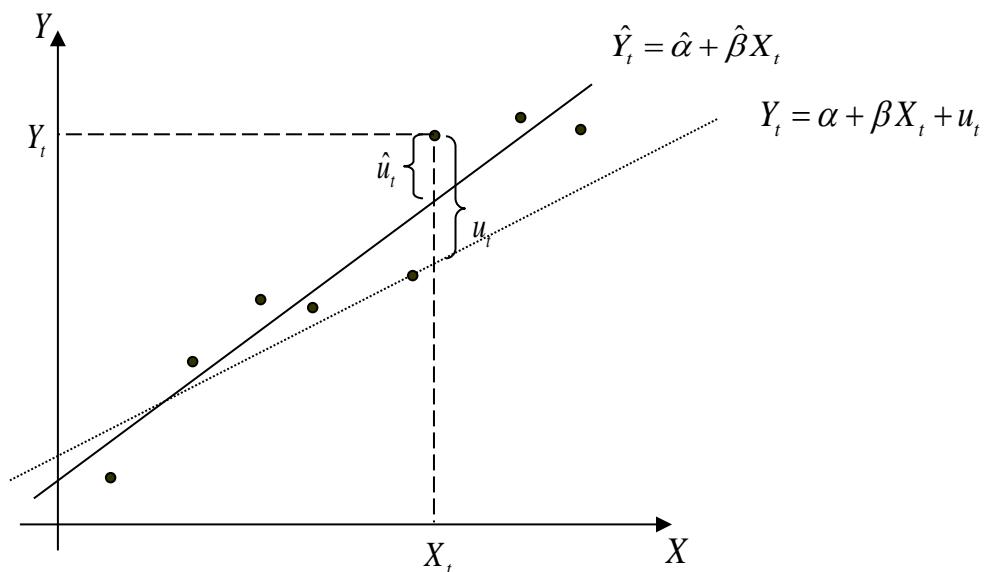
5. את α ו- β ו- u_t לא ניתן לדעת (אלא רק לאמוד מנתוני המדגם) – הקו האמיתי באוכי לא ידוע.

6. אפשר לדעת את \hat{u}_t , שהיא הסטייה מקו הרגרסיה במדגם :

- עבור X_t , הערך הצפוי של המשתנה הקשור (\hat{Y}_t) המתתקבל לפי הרגרסיה

$$\text{הוא : } \hat{Y}_t = \hat{\alpha} + \hat{\beta}X_t.$$

- הסטייה של התצפית (\hat{Y}_t) מהערך הצפוי לפי הרגרסיה (Y_t) היא :



— קו הרגרסיה הנאמד (במדגם)
..... קו הרגרסיה האמיתי באוכלוסייה
• תצפית בודדת

שאלות:**(1)** הבא נוכח את הזיהויות הבאות:

$$\cdot \sum_{t=1}^T (X_t - \bar{X})^2 = \sum_{t=1}^T X_t^2 - T\bar{X}^2 . \text{ א}$$

$$\cdot \sum_{t=1}^T (X_t - \bar{X})^2 = \sum_{t=1}^T (X_t - \bar{X})X_t . \text{ ב}$$

$$\cdot \sum (X_t - \bar{X}) = 0 . \text{ ג}$$

$$\cdot \sum \frac{(X_t - \bar{X})^2}{\sum (X_t - \bar{X})X_t} = 1 . \text{ ד}$$

$$\cdot \sum (\hat{\alpha} + \hat{\beta}x_i)(x_i - \bar{x}) = \hat{\beta}(x_i - \bar{x})^2 . \text{ ה}$$

$$\cdot \sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y}) = \sum_{t=1}^T X_t Y_t - T\bar{X}\bar{Y} . \text{ ו}$$

$$\cdot \sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y}) = \sum_{t=1}^T (X_t - \bar{X})Y_t . \text{ ז}$$

(2) בטא באמצעות $\text{cov}(x, y), \text{var}(x), \text{var}(y)$ והקבועים a ו- b את הביטויים

הבאים:

$$\cdot \text{Var}(ax) . \text{ א}$$

$$\cdot \text{Var}(x+y) . \text{ ב}$$

$$\cdot \text{Var}(ax+b) . \text{ ג}$$

$$\cdot \text{Cov}(x, ay) . \text{ ד}$$

$$\cdot \text{Cov}(x+a, y+b) . \text{ ה}$$

ו. מקדם המתאים בין x ל- y .**תשובות סופיות:****(1)** הוכחה.

$$\cdot a \text{cov}(x, y) . \text{ ז} \quad \cdot a^2 \text{var}(x) . \text{ ג} \quad \cdot \text{var}(x) + \text{var}(y) + 2\text{cov}(x, y) . \text{ ב} \quad \cdot a^2 \text{var}(x) . \text{ א} \quad \text{ (2)}$$

$$\cdot r = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)} \cdot \sqrt{\text{var}(y)}} . \text{ ו} \quad \cdot \text{cov}(x, y) . \text{ ה}$$

אקוונומטריקה

פרק 2 - אומדי הריבועים הפחותים

תוכן העניינים

- 7 1. כללי

אומדי הריבועים הפחותים:

רקע:

שיטת האמידה של α ושל β לקבלת אומדים $\hat{\alpha}$ ו- $\hat{\beta}$ – Ordinary Least Squares (OLS) שיביאו למינימום את סכום ריבועי טעויות האמידה :

$$\min_{\hat{\alpha}\hat{\beta}} \sum \hat{u}_t^2 = \min_{\hat{\alpha}\hat{\beta}} \sum (y_t - \hat{y}_t)^2 = \min_{\hat{\alpha}\hat{\beta}} \sum [y_t - (\hat{\alpha} + \hat{\beta}X_t)]^2 = ?$$

מתוך גזירת הפונקציה זו מתקבלים האומדים $\hat{\alpha}$ ו- $\hat{\beta}$.

מודל רק עם חותך	מודל ללא חותך	מודל עם חותך ושיפוע	
$Y_t = \alpha + u_t$	$Y_t = \beta X_t + u_t$	$Y_t = \alpha + \beta X_t + u_t$ $\hat{\beta} = \frac{S_{XY}}{S_{XX}} = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^T (X_t - \bar{X})^2}$ $= \frac{\sum_{t=1}^T (X_t - \bar{X})Y_t}{\sum_{t=1}^T (X_t - \bar{X})^2}$ $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$	חישוב האומדים
$E(\hat{\alpha}) = \alpha$	$E(\hat{\beta}) = \beta$	$E(\hat{\beta}) = \beta$ $E(\hat{\alpha}) = \alpha$	תוחלת האומדים
$V(\hat{\alpha}) = \frac{\sigma_u^2}{T}$	$V(\hat{\beta}) = \frac{\sigma_u^2}{\sum_{i=1}^T X_t^2}$	$V(\hat{\beta}) = \frac{\sigma_u^2}{S_{XX}}$ $V(\hat{\alpha}) = \sigma_u^2 \left(\frac{1}{T} + \frac{\bar{X}^2}{S_{XX}} \right)$	שונות האומדים

"המשוואות הנורמליות" מתקבלות בתחילת הגזירה של פונקציית הריבועים

הפחותים וחייבות להתקיים על מנת שהפונקציה התקיים ($\sum_i \hat{u}_i^2 = \min$) :

עבור המודל הקליני (עם חותך) :

$$\text{בגזרה של } \alpha : \sum_i \hat{u}_i = 0$$

$$\text{בגזרה של } \beta : \sum_i \hat{u}_i \cdot x_i = 0$$

עבור מודל ללא חותך :

$$\text{בגזרת } \beta \text{ בלבד} : \sum_i \hat{u}_i \cdot x_i = 0$$

מן המשוואות הנורמליות נובעות :

1. התכונות הגיאומטריות :

$$\text{א. } \sum_i \hat{u}_i = 0$$

$$\text{ב. } \sum_i x_i \hat{u}_i = 0$$

- ברגסיה לא שיפוע מתקיימת רק התכונה הגיאומטרית הראשונה.
- ברגסיה לא חותן מתקיימת רק התכונה הגיאומטרית השנייה.

2. התכונות האלגבריות :

$$\text{א. } \text{cov}(x_i, \hat{u}_i) = 0$$

$$\text{ב. } \text{cov}(\hat{y}_i, \hat{u}_i) = 0$$

$$\text{ג. } \bar{y} = \hat{\alpha} + \hat{\beta} \bar{x} = \bar{\hat{y}}$$

- התכונות האלגבריות תקפות עבור קו הרוגסיה הקליני (עם חותך ושיפוע) במדגם בלבד.

הנחהות הקלאסיות של מודל הרוגסיה:

1. קיים קשר ליניארי בין המשתנה המוסבר למשתנה המסביר.

$$\text{2. } X \text{ איננו קבוע} : S_{XX} = \sum_{t=1}^T (X_t - \bar{X})^2 \neq 0$$

3. תוחלת ההפרעה האקראית היא אפס לכל תצפית: $E(u_t) = 0$ לכל t .

4. X_t אינם משתנים מקרים \Leftarrow ניתן להוציא אותם מחוץ לתוחלת ולשונו \Leftarrow

$$\text{cov}(X_t, u_t) = 0$$

5. הומוסקדיות: שונות ההפרעה האקראית קבועה לכל תצפית:

$$V(u_t) = \sigma_u^2$$

$$\text{6. } u_t \text{ ב''ת: } \text{cov}(u_t, u_s) = 0 \text{ לכל } t \neq s.$$

7. ההפרעות האקראיות מתפלגות נורמלית: $N(u_t) \approx N$.

תכונות האומדים:

אומדי הריבועים הפחותים הם לינאריים, חסרי הטיה, יעילים ועקבים.

1. לינאריות:

אר"פ ניתנים להציג כטרנספורמציה לינארית של \hat{Y}_t .

כדי ש- $\hat{\beta}$ למשל, יהיה אומד לינארי צריך להתקיים:

$$\hat{\beta} = \frac{\sum X_t \cdot Y_t}{\sum X_t^2}$$

כדי להביא את האומד לצורה: $\tilde{\beta} = \sum w_i \cdot y_i$ נעזר בשוויון:

אומד זה ניתן להציג בצורה הבאה:

$$\hat{\beta} = \sum \frac{X_t}{\sum X_t^2} Y_t = \sum W_t \cdot Y_t$$

$$W_t = \frac{X_t}{\sum X_t^2}$$

לפיכך מדובר באומד לינארי.

• שימוש לב Ci:

W_t אסור שיכלול את \hat{Y}_t .

\hat{Y}_t אסור שייהי במבנה או בשורש/חזקה (אליא אם כי במודל הנתון הוא מצוי בשורש/חזקה).

2. חסר הטיה:

אומד $\hat{\theta}$ מסויים יהיה אח"ה לפרמטר θ אותו הוא אומד באוכלוסייה אם מתקיים: $E(\hat{\theta}) = \theta$.

כיצד יודעים אם אומד הוא חסר הטיה?

1. בשלב הראשון יש לבצע עבודה הכנה – מבטאים את האומד באמצעות הפרמטר האמייתי – מציבים במקום ה- $\hat{\theta}$ את המודל ופתחים אלגברית.

- יש לזכור כי:

u_t מהווים משתנים מקרים \leftarrow נשארים בתוך התוחלת, השונות y_t וה- \sum .

x_t איננו משתנה מקרי (על פי הנחה מס' 4) \leftarrow יוצא מחוץ לתוחלת ולשונות אך נשאר בתוך ה- $\sum \frac{\alpha}{\beta}$ קבועים \leftarrow יוצאים מחוץ לתוחלת, לשונות ול- \sum .

2. בשלב השני מפעילים תוחלת על האומד המפותח ואם התוחלת שווה לפרמטר האמייתי אז האומד חסר הטיה.

- חסר הטיה מחייב את התקיימותן של הנחות (3) $E(u_t) = 0$ לכל t

ו- (4) $\text{cov}(X_t, u_t) = 0$.

3. **יעילות:**
 יעילות פירושה השונות הקטנה ביותר. ככל שהשונות של האומד קטנה יותר, כך יש הסתברות גבוהה יותר שהוא יהיה קרוב לפרמטר האמייתי באוכלוסייה אותו הוא אומד.
 $\hat{\theta}_1$ יקרא אומדיעיל יותר מ- $\hat{\theta}_2$ אם מתקיים שהשונות שלו קטנה יותר: $V(\hat{\theta}_1) < V(\hat{\theta}_2)$.

משפט גאוס מרקוב – אר"יפ הם בעלי השונות הנמוכה ביותר בקבוצה שלהם (קבוצת האומדים הליינריים חסרי ההטיה), והם נקראים: B.L.U.E. (Best Linear Unbiased Estimation).

כיצד מחשבים שונות של אומד? חייבות להתקיים הנחות (4) $\text{cov}(X_t, u_t) = 0$ לכל t ו-(6) $\text{cov}(u_t, u_s) = 0$ לכל $s \neq t$. אם הן מתקיימות, מחשבים את השונות של האיברים המכילים את u_t מהפיתוח הקודם (לפי כללי הסיגמא והשונות).

4. עיקיות:

כל שהמדגם יגדל כך יתקרב האומד לערך האמתי של הפרמטר. אם נגדיל את המדגם לאינסוף תצפויות ונחשב את האומד, הוא יהיה שווה

$$\text{לפרמטר האמתי באוכטוסייה: } (\hat{\theta} \rightarrow \theta, T \rightarrow \infty)$$

תנאי הכרחי לעיקיות:

האומד חייב להיות פונקציה של גודל המדגם. במקרים אחרים, האומד צריך להיות מושפע מגודל המדגם. ברגע שהאומד עונה על תנאי זה הוא יהיה עיקיב. אומד המוחש במדגם סופי בהגדרה לא יוכל להיות עיקיב לפרמטר באוכטוסייה.

סיכום: השלבים להוכחת התוכנות:

1. הוכחת ליניאריות.
2. הכנת האומד \leftarrow להציב במקום Y את המודל האמתי.
במודל עם חותך: $Y_t = \alpha + \beta X_t + u_t$
במודל ללא חותך: $Y_t = \beta X_t + u_t$
3. פיתוח האלגברה.
4. חישוב תוחלת, שונות, עיקיות.
 - ליניאריות מהויה תנאי הכרחי לחוסר הטיה.
 - ליניאריות וחוסר הטיה מהויה תנאי הכרחי לבחינת הייעילות של האומד לפי משפט גאוס-מרקוב.
 - עיקיות איננה תלואה בתוכנות האחרות, אלא רק בהיותו של האומד פונקציה של גודל המדגם (לא מוחש על מדגם סופי). כך שאומד לא חייב להיות ליניארי או חסר הטיה כדי להיות עיקיב.
 - העיקיות משפיעה על הייעילות של האומד. עבור אומדים התלויים בגודל המדגם: ככל שגודל המדגם גדול יותר כך שונות האומד קטנה והאומד יהיה יעיל יותר לפרמטר באוכ'.

שאלות:

גזרת אוף:

- 1) כלכלן החליט לאמוד את המודל: $y_i = \alpha + \beta x_i + u_i$
- נסחו את בעיית ה-OLS.
 - מצאו את תנאי סדר ראשון של בעיית ה-OLS (המשוואות הנורמלליות).
 - מצאו נוסחה לקבלת האומדים: $\hat{\alpha}, \hat{\beta}$.
 - הוכחו כי קו הרוגרסיה עובר דרך נקודת הממוצעים (\bar{X}, \bar{Y}) .
 - בහנחה והיינו בוחרים אומד אחר ל- β שאינו אומד הריבועים הפחותים, מה היה יחס הביטויים: $\sum e_i^2$ ו- $\sum e_i$ של אומד זה ביחס לאומד הריבועים הפחותים?
- 2) כלכלן החליט לאמוד את המודל: $y_i = \beta x_i + u_i$
- נסחו את בעיית ה-OLS.
 - מצאו את תנאי סדר ראשון של בעיית ה-OLS.
 - מצאו נוסחה לקבלת $\hat{\beta}$.
 - הוכחו כי קו הרוגרסיה אינו עובר דרך נקודת הממוצעים (\bar{X}, \bar{Y}) .
 - מהו התנאי שבו אומד הריבועים הפחותים שמצאתם בסעיף ג' יהיה זהה לנוסחה של אומד הריבועים הפחותים שנמצא בשאלת הקודמת (במודל עם חותך)?
- 3) חוקר רצה לחקר האם ציוני IQ משפיעים על הציון באקונומטריקה ולכן אסף תוצאות מ-5 סטודנטים:

SCORE	IQ	e_i
80	100	1
75	110	-1
80	110	1
90	103	2
85	102	-3

איזה מבין המודלים הבאים נAMD?

- $scôre_i = \hat{\alpha} + \hat{\beta} \cdot IQ_i$
- $scôre_i = \hat{\beta} \cdot IQ_i$
- $scôre_i = \hat{\alpha}$
- $scôre_i = \bar{y}$

4) נבדק הקשר שבין שכר לשעה שעבוד מסויים מרוויח אצל מעסיק מסוים (X) לבין כמות העובדים שਮועסקים אצל אותו מעסיק (Y) (הניחו שכר שווה בין העובדים אצל אותו המעסיק).

לשם כך נדגמו 10 מעסיקים באופן מקרי ונתקבלו התוצאות הבאות:

$$\bar{x} = 35$$

$$\bar{y} = 5.8$$

$$\sum_{i=1}^{10} x_i^2 = 19,100$$

$$\sum_{i=1}^{10} y_i^2 = 440$$

$$\sum_{i=1}^{10} x_i y_i = 2858.85$$

מהי תחזית כמות העובדים המועסקים אצל מעסיק מסוים המשתכנים 25 לשעה?

5) כלכלן החליט לאמוד את המודל: $y_i = \alpha + u_i$.

א. נסחו את בעיית ה-OLS.

ב. מצאו את תנאי סדר ראשון של בעיית ה-OLS.

ג. מצאו נוסחה לקבלת $\hat{\alpha}$.

6) חוקר רצה לבדוק את המודל: $y_i = \hat{\beta}x_i + u_i$ כאשר המשתנה תלוי הוא הציון במקשו והב"ת הוא ציוני IQ. לשם כך אסף תצפיות של 5 סטודנטים:

SCORE	IQ	ציון חיזי	e_i
80	100		
90	110		
95	110		
92			5
	102		3

מאמידת הרוגרסיה התקבל כי: $\hat{\beta} = 0.85$

השלם את התאים הריאקים בטבלה.

הנחות המודל:

7) שכר של עובדים מנובא על ידי השכלהם במודל הבא : $w_i = \alpha + \beta \cdot s_i + u_i$.

א. כתבו את הנחות הקלאסיות בmonoח המשתנים של המודל הנתון וסבירו אותן.

ב. התייחסו לכל אחת מהטענות הבאות וקבעו האם היא :

הנחה קלאסית / משווה נורמללית (או תוצאה הנובעת ממשווה נורמללית) / אף אחד מהשניים :

$$\text{cov}(s_i, u_i) = 0 \quad .i$$

$$E(u_i) = 0 \quad .ii$$

$$\text{cov}(u_i, u_j) = 0 \quad .iii$$

$$\bar{e} = 0 \quad .iv$$

$$\bar{w} = \hat{\bar{w}} \quad .v$$

$$\sum u_i = 0 \quad .vi$$

$$V(u_i) = \sigma_i \quad .vii$$

$$S_s^2 \neq 0 \quad .viii$$

$$\text{cov}(s, e) = 0 \quad .ix$$

$$\text{cov}(\hat{y}_i, e) = 0 \quad .x$$

8) חוקר מעוניין לאמוד את ההשפעה של נוכחות בתרגולים על הציון בקורס

אקוונומטריה. לשם כך אמד את המשווה : $score = \alpha + \beta attendance + u$

הועלתה הטענה כי מודל זה אינו מקיים את הנחה מס' 4 של אי תלות בין

המשתנה הבית לטיעויות . חוויה דעתך על טענה זו.

לייניאריות:

$$9) \text{ האם האומד : } \hat{\beta} = \frac{\sum_{t=1}^T (X_t - \bar{X})Y_t}{\sum_{t=1}^T (X_t - \bar{X})^2} \text{ הוא לייניארי?}$$

$$10) \text{ האם האומד : } \tilde{\beta} = \frac{\sum_{t=1}^T Y_t^3 \sum_{t=1}^T X_t (Z_t + Y_t)}{\sum_{t=1}^T X_t^2} \text{ הוא לייניארי?}$$

11) כלכלן החליט לאמוד את המודל: $y_i = \alpha + \beta x_i + u_i$ מי מהאומדים הבאים הוא ליניארי ומהן המשקولات:

$$\text{א. } \tilde{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$\text{ב. } \tilde{\beta} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\text{ג. } \tilde{\beta} = \sum \left(\frac{y_i}{x_i} \right)^2$$

$$\text{ד. } \tilde{\beta} = \sum \frac{Y_i}{n}$$

$$\text{ה. } \tilde{\beta} = \sum \frac{X_t}{Y_i}$$

$$\text{ו. } \tilde{\beta} = \frac{\sum x_i y_i}{\sum (x_i - \bar{x})^2}$$

$$\text{ז. } \tilde{\beta} = \frac{\bar{Y}}{\bar{X}}$$

חומר הטיה:

$$\text{12) נתון האומד הבא: } \tilde{\beta} = \frac{\sum_{t=1}^T X_t Y_t}{\sum_{t=1}^T X_t^2}$$

האם האומד הנ"ל הוא חסר הטיה?

- א. בדוק במודל עם חותך.
- ב. בדוק במודל ללא חותך.

13) נתון המודל הבא: $y_i = \alpha + \beta x_i + u_i$ נתו בנוסף כי האומד ל- β הינו ליניארי וחסר הטיה.

איזה מן הטענות מתקיימת בהכרח:

$$\text{א. } \sum w_i x_i = 1$$

$$\text{ב. } \sum w_i x_i = 0$$

$$\text{ג. } \sum w_i = 0$$

יעילות ועקבות:

$$14) \text{ ככלון חzie את האומד הבא עבור } \beta : \tilde{\beta} = \frac{y_9 - y_5 + y_2}{x_9 - x_5 + x_2}$$

- א. בדוק האם האומד חסר הטיה עבור המודל הקלסי.
 ב. האם תשתנה תשובה אם מדובר באומד ללא חותך?
 ג. חשב את שונות האומד עבור מודל ללא חותך.

תרגול ממבחןים:

15) נתון המודל: $T = 100$, $Y_t = \alpha + \beta X_t + u_t$, כאשר מתקיימות כל ההנחהות הקלסיות.

$$\text{נתון האומד : } \tilde{\beta} = \frac{\sum_{t=51}^{100} Y_t - \sum_{t=1}^{50} Y_t}{\sum_{t=51}^{100} X_t - \sum_{t=1}^{50} X_t}$$

- א. האומד $\tilde{\beta}$ הינו אומד חסר הטיה ל- $-\beta$.
 ב. האומד $\tilde{\beta}$ הינו אומד עקיב ל- $-\beta$.
 ג. האומד $\tilde{\beta}$ הינו אומד לינארי ל- $-\beta$.
 ד. האומד $\tilde{\beta}$ הינו אומד יעיל ל- $-\beta$.
 ה. השונות האמיתית של $\tilde{\beta}$ היא?

16) נתון המודל: $Y_t = \beta X_t + u_t$, כאשר כל ההנחהות הקלסיות מתקיימות.
 (יש לשים לב המודל ללא חותך).

$$\text{נתון האומד : } \tilde{\beta} = \frac{\sum Y_t}{\sum X_t}$$

- א. האומד $\tilde{\beta}$ הינו אומד מוטה ל- $-\beta$: נכון / לא נכון / אי אפשר לדעת
 ב. על סמך משפט גאוס-מרקוב ניתן להסיק כי $\tilde{\beta}$ איננו אומד יעיל יותר
 נכון / לא נכון / לא ניתן לדעת מאומד הריבועים הפחותים:
 ג. מהי השונות האמיתית של $\tilde{\beta}$?

17) נתון המודל: $Y_t = \beta X_t + u_t$, כאשר כל ההנחות הקלאסיות מתקיימות.
(יש לשים לב המודל ללא חותם).

$$\text{נתון האומד: } \tilde{\beta} = \frac{\sum X_t Y_t}{\sum (X_t - \bar{X})^2}$$

א. מהי התוחלת של $\tilde{\beta}$?

ncou / לא ncou / אי אפשר לדעת ב. $E(\tilde{\beta}) < \beta$.

ג. על סמך משפט גאוס-מרקוב ניתן להסיק כי אומד הריבועים הפחותים הינו אומד יעיל יותר מ- $\tilde{\beta}$.
ncou / לא ncou / לא ניתן לדעת

$$\text{ד. מהי השונות האמיתית של האומד: } \frac{\sum X_t Y_t}{\sum X_t^2}?$$

18) בכל השאלהות ההנחות הקלאסיות מתקיימות.
האומדים הם אר"יפ, והמודל הוא: $Y_t = \alpha + \beta X_t + u_t$.

ncou / לא ncou / אי אפשר לדעת א. $E(Y_t) = E(\hat{Y}_t)$

ncou / לא ncou / אי אפשר לדעת ב. $\sum_{t=1}^T (X_t - \bar{X}) \bar{Y} \neq 0$

ג. אמידת המודל בשיטות הריבועים הפחותים תיתן את

ncou / לא ncou / אי אפשר לדעת ה. התוצאה: $\sum_{t=1}^T u_t = 0$

ד. אם נתון ש $r_{xy} = 0.57$, אז $\hat{\beta}$:

i. הוא בהכרח שלילי.

ii. הוא בהכרח חיובי.

iii. הוא בהכרח שווה לאפס.

iv. לא ניתן לקבוע את סימנו על סמך הנתונים הקיימים.

ה. סמן את הטענה הנכונה בהכרח:

$$\text{i. } \sum_{t=1}^T (X_t - \bar{Y}) \hat{u}_t = 0$$

$$\text{ii. } S_{xx} = \sum_{t=1}^T X_t^2 - (T \bar{X})^2$$

$$\text{iii. } \sum_{t=1}^T X_t u_t = 0$$

iv. אף אחת מהטענות הנילאיינה נכונה בהכרח.

- ו. אומדי הריבועים הפחותים אינם חסרי הטיה, אם נתון שהשונות של n נכון / לא נכון / אי אפשר לדעת אינה קבועה.
- ז. אומד חסר הטיה הוא אינו בהכרח נכון / לא נכון / אי אפשר לדעת גם אומד עקיף.

תשובות סופיות:

$$\hat{\alpha} \Leftarrow \sum_{i=1}^n \hat{u}_i = 0, \hat{\beta} \Leftarrow \sum_{i=1}^n \hat{u}_i x_i = 0 \quad \text{ב.} \quad \min_{\hat{\alpha}, \hat{\beta}} \sum \left[y_t - (\hat{\alpha} + \hat{\beta} x_t) \right]^2 \quad \text{א.} \quad (1)$$

ה. ראה סרטון. ד. הוכחה. . $\frac{S_{xy}}{S_{xx}} = \hat{\beta}$, $\bar{y} - \hat{\beta} \bar{x} = \hat{\alpha}$. ג.

$$\hat{\beta} \Leftarrow \sum_{i=1}^n \hat{u}_i x_i = 0 \quad \text{ב.} \quad \min_{\hat{\beta}} \sum \left[y_t - (\hat{\beta} x_t) \right]^2 \quad \text{א.} \quad (2)$$

ה. ראה סרטון. ד. הוכחה. . $\frac{\sum y_i x_i}{\sum x_i^2} = \hat{\beta}$. ג.

א'. ג.

$$\hat{\alpha} = \bar{y} \quad \text{ג.} \quad \sum_{i=1}^n \hat{u}_i = 0 \quad \text{ב.} \quad \min_{\hat{\alpha}} \sum \left[y_t - (\hat{\alpha}) \right]^2 \quad \text{א.} \quad (5)$$

ה. ראה סרטון. ד. הוכחה. . $\hat{y}_i = 4.59$. (4)

(6)

SCORE	IQ	ציון חזוי	e_i
80	100	85	-5
90	110	93.5	-3.5
95	110	93.5	1.5
92	114	97	5
89.7	102	86.7	3

- (7) א. ראה סרטון.
ב.נ. הנחה, לא בהכרח מתקיים.
ii. הנחה, לא בהכרח מתקיים.
iii. הנחה, לא בהכרח מתקיים.
iv. נוצר מהתשוויה, מתקיים בהכרח.
v. נוצר מהתשוויה, מתקיים בהכרח.
vi. אף אחד מהשניים.
vii. אף אחד מהשניים.
viii. הנחה, לא בהכרח מתקיים.
ix. נוצר מהתשוויה, מתקיים בהכרח.
x. נוצר מהתשוויה, מתקיים בהכרח.
- (8) ראה סרטון.
(9) כן.
(10) לא.

. $W_i = \frac{(x_i - \bar{x})}{\sum(x_i - \bar{x})^2}$ ב. ליניארי, . $W_i = \frac{x_i}{\sum x_i^2}$ א. ליניארי,

. $W_i = \frac{1}{n}$ ד. ליניארי, ג. לא ליניארי.

. $W_i = \frac{x_i}{\sum(x_i - \bar{x})^2}$ ח. לא ליניארי.

. $W_i = \frac{1}{\sum x_i}$ ז. ליניארי,

ב. כן.

א. לא.

כ. או-ו-ג.

. $\sigma_u^2 \frac{1}{(x_9 - x_5 + x_2)^2}$ ג. חסר הטיה. ב. מوطה.

ד. לא נכון. ג. נכון. א. נכון.

ב. לא נכון.

ג. נכון.

ד. לא נכון.

. $V(\tilde{\beta}) = \frac{100\sigma_u^2}{\left(\sum_{t=51}^{100} X_t - \sum_{t=1}^{50} X_t\right)^2}$

. $V(\tilde{\beta}) = \frac{T\sigma_u^2}{\left(\sum X_t\right)^2}$ ג. נכון. ב. נכון. א. לא נכון.

. $E(\tilde{\beta}) = \frac{\beta \sum X_t^2}{\sum(X_t - \bar{X})^2}$ א. נכון. ב. לא נכון. ג. לא נכון.

. $\frac{\sigma_u^2}{\sum X_t^2}$ ז.

ד. נכון. ב. לא נכון. ג. לא נכון. א. נכון.

ו. לא נכון. ז. נכון. ח. או.

אקוונומטריקה

פרק 3 - מודלים לא ליניארים

תוכן העניינים

- 21 1. כללי

מודלים לא ליניארים:

רקע:

הגמישות $\left(\frac{\partial Y}{\partial X} \cdot \frac{X}{Y} \right)$ בכמה % ישנה ב- X אם נגדיל את Y ב- 1% ?	השינוי השولي $\left(\frac{\partial Y}{\partial X} \right)$ בכמה ישנה Y אם נגדיל את X ביחידה?	משמעות ה- β	המודל
$\frac{\beta X}{Y}$	β	השינוי השولي אם נגדיל את X ביחידה Y ישנה ב- β יחידות	لينיארי: $Y = \alpha + \beta X + u$
βX	βY	שיעור ההשינוי השולי אם נגדיל את X ביחידה Y ישנה ב- $100 \cdot \beta \%$	חצי לוגריטמי: $\ln Y = \alpha + \beta X + u$ $(Y = e^{\alpha + \beta X + u})$
β	$\frac{\beta Y}{X}$	הגמישות אם נגדיל את X ב- 1% Y ישנה ב- $\beta\%$	לוגריטמי כפול: $\ln Y = \alpha + \beta \ln X + u$ $(Y = e^\alpha \cdot X^\beta \cdot e^u)$
$\frac{\beta}{Y}$	$\frac{\beta}{X}$	אין משמעות כלכלית אם נגדיל את X ב- 1% Y ישנה ב- β	לוג ליניארי: $Y = \alpha + \beta \ln X + u$ $(e^y = e^\alpha \cdot X^\beta \cdot e^u)$

- המשתנה שישי בו LN השינוי בו יהיה באחוזים.

תזכורת של חוקי לוגים:

$$\begin{aligned} LN(e^x) &= X \\ LN(X^y) &= Y \cdot LN(X) \end{aligned}$$

$$LN(X \cdot Y) = LN(X) + LN(Y)$$

$$LN\left(\frac{X}{Y}\right) = LN(X) - LN(Y)$$

שאלות:

1) על מנת לאמד את התשואה להשכלה בישראל בשנים 1948-1990 נאמדו המודלים הבאים :

$$\cdot MWAGE_t = 139.547 + 118.628 \cdot SCL_t \quad .1$$

$$\cdot MWAGE_t = -1445.08 + 1239.60 \cdot LN(SCL)_t \quad .2$$

$$\cdot LN(MWAGE)_t = 5.244 + 0.778 \cdot LN(SCL)_t \quad .3$$

$$\cdot LN(MWAGE)_t = 6.292 + 0.070 \cdot SCL_t \quad .4$$

א. הסבירו את המשמעות של β בכל אחד מהמודלים.

ב. חשבו את הגמישות בנקודות הממוצעים : (12.311,1600.01) עבור כל אחד מהמודלים.

2) נתונים תוצאות האמידה של המודלים הבאים :

$$\hat{Y} = e^{4.5} \cdot X^{0.05} \quad .1$$

$$\hat{Y} = e^{4.5+0.05X} \quad .2$$

$$\hat{Y} = 4.5 + \frac{0.05}{X} \quad .3$$

$$\hat{Y} = \frac{1}{1 + e^{4.5+0.05X}} \quad .4$$

א. כתבו את המודלים בצורה ליניארית בעזרת טרנספורמציה מתאימה.

ב. עבור כל אחד מהמודלים ערכו תחזית נקודתית עבור $X = 6$.

3) נתונים המודלים הבאים עבור התוצר במשק :

$$\cdot Q_i = AK_i^{\beta_1} e^{u_i} \quad .1$$

$$\cdot Q_i = Ae^{\beta_1 L_i + u_i} \quad .2$$

$$\cdot Q_i = A + K_i^{\beta_1} + e^{u_i} \quad .3$$

$$\cdot Q_i = A + \frac{\beta_1}{L_i} + u_i \quad .4$$

$$\cdot Q_i = A + \beta_1 \sqrt{K_i} + u_i \quad .5$$

$$\cdot Q_i = e^{A + \beta_1 K_i + u_i} \quad .6$$

$$\cdot Q_i = A \left(\frac{K_i}{2} + 7 \right)^{\beta_1} e^{u_i} \quad .7$$

$$\cdot Q_i = A + \beta_1 L_i + u_i \quad .8$$

$$\cdot Q_i = A + \beta_1 \left(\frac{K_i}{L_i} \right) + u_i \quad .9$$

כאשר :

- Q - הוצאות צריכה על מוצר מסוים על ידי פרט מסוים.
- A - הוצאות צריכה על המוצר בהינתן רמת הכנסה אפסית.
- K - הכנסת הפרט.
- L - שנות לימוד.

- א. מי מהמודלים הבאים ניתן לאמידה בשיטת OLS?
- ב. מי מבין המודלים שלא ניתנים לאמידה בשיטת OLS ניתן להביא למודל ליניארי בפרמטרים ועל כן לאמוד את הפרמטרים שלו?
- ג. עברו כל אחד מהמודלים קבעו מיהו המשנה המושבר ומיהו המסbir במשוואת הרגרסיה הליניארית.
- ד. עוקמת אנגל מתארת את גמישות הצריכה של הפרט מוצר מסוים ביחס להכנסתו. איזה מהמודלים מתאים כדי לתאר את עוקמת אנגל?

$$4) \text{ נתון המודל הבא : } Q_i = \frac{A}{K_i^{\beta_i}} e^{u_i} .$$

- א. האם ניתן לאמוד את המודל בשיטת OLS?
- ב. מה המשועה שצריך לאמוד על מנת לקבל את הפרמטרים למודל זה (כלומר כיצד הופכים את המודל ליניארי בפרמטרים)?
- ג. נאמד המודל הבא : $\ln(Q_i) = \alpha_0 + \alpha_1 \ln(K_i) + u_i$, והתקבלו התוצאות הבאות : $\hat{\alpha}_0 = 3$, $\hat{\alpha}_1 = 0.8$
מהם האומדנים עברו : ? A , β_1

- 5) נתון כי הקשר באוכטוסייה בין X ל- Y נתון על ידי המודל הבא : $u + \ln Y = \alpha + \beta \ln X$. נתון גם כי עברו המודל הניל כל ההנחהות הקלאסיות מתקיימות.

$$\text{כלכלן הציע את האומד הבא עבור } \beta : \tilde{\beta} = \frac{\sum_{t=1}^T (\ln X_t - \bar{\ln X}) \ln Y_t}{\sum_{t=1}^T (\ln X_t - \bar{\ln X})^2}$$

- א. האם האומד ליניארי?
- ב. האם האומד חסר הטיה?
- ג. האם האומד blue?
- ד. מהי שונותו?

תשובות סופיות:

- (1) א. 1. השינוי השולי. 2. אין משמעות כלכלית. 3. גמישות.
4. שיעור השינוי השולי.

ב. $0.861 \cdot 4 = 0.778 \cdot 3 = 0.77 \cdot 2 = 0.912$

. $\ln(Y) = 4.5 + 0.05X$. 2 . $\ln(Y) = 4.5 + 0.05 \cdot \ln(X)$. 1. (2)

$$\cdot \ln\left(\frac{1-\hat{Y}}{\hat{Y}}\right) = 4.5 + 0.05X \cdot 4$$

3. אין צורך.

ב. $0.00816 \cdot 4 = 4.50833 \cdot 3 = 121.51 \cdot 2 = 98.45$

(3) א. מודלים: 8, 5, 4, 1-7.

ב. מודלים: 6, 2, 1, 7-1.

ג. 1. מסביר: $\ln(Q_i)$, מושבר: $\ln(K_i)$

2. מסביר: L_i , מושבר: $\frac{1}{L_i}$

3. אינם ליניארי.

4. מסביר: Q_i , מושבר: K_i

5. מסביר: $\sqrt{K_i}$, מושבר: Q_i

6. מסביר: K_i , מושבר: L_i

7. מסביר: L_i , מושבר: $\ln(Q_i)$

8. מסביר: L_i , מושבר: K_i

9. מסביר: Q_i , מושבר: $\frac{K_i}{L_i}$

ד. מודלים: 1, 7-1.

(4) א. לא. $\ln(Q_i) = \ln(A) - \beta_1 \ln(K_i) + u_i$

ג. $\beta_1 = -0.8$, $A = 20$

. $V(\hat{\beta}) = \frac{\sigma_u^2}{SS \ln x}$. 7 . ב. כן. ג. כן.

אקוונומטריקה

פרק 4 - מבחני המובהקות וקריאת פלטים - תוכנת SAS

תוכן העניינים

- 25 1. כללי

מבחני המובהקות וקריאת פלטיהם – תוכנת SAS

רקע:

פלט ניתוח שונות (Analysis of Variance)

Analysis of Variance					
Source	DF	Sum of	Mean	F Value	Prob>F
		Squares	Square		
Model	k	RSS	$RSS/k = MSR$	$F = \frac{MSR}{MSE}$	PF
	+	+			
Error	$T - k - 1$	ESS	$ESS/T - k - 1 = MSE$		
C Total	$\overline{T - 1}$	\overline{TSS}			
<hr/>					
Root MSE		$\sqrt{MSE} = s_u$	R-square	$R^2 = \frac{RSS}{TSS}$	
Dep Mean		\bar{Y}	Adj R-sq	$\bar{R}^2 = 1 - \frac{ESS}{TSS} \cdot \frac{T - 1}{T - k - 1}$	
C.V.		$\frac{s_u}{\bar{Y}} \cdot 100$			

פלט מקדמי הרוגרסיה (Parameter Estimates)

Parameter Estimates					
Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	$\hat{\alpha}$	$s_{\hat{\alpha}}$	$\frac{\hat{\alpha}}{s_{\hat{\alpha}}} = t_{(\hat{\alpha}=0)}$	$Pt_{\hat{\alpha}}$
X	1	$\hat{\beta}$	$s_{\hat{\beta}}$	$\frac{\hat{\beta}}{s_{\hat{\beta}}} = t_{(\hat{\beta}=0)}$	$Pt_{\hat{\beta}}$

פלט ה – Covariance of Estimates

פלט שמתאר את השונות המשותפת (covariance) של האומדנים- $\hat{\alpha}$ ו- $\hat{\beta}$:

Covariance of Estimates		
COVB	INTERCEP	X
INTERCEP	$s_{\hat{\alpha}}^2$	$\text{cov}(\hat{\alpha}, \hat{\beta})$
X	$\text{cov}(\hat{\alpha}, \hat{\beta})$	$s_{\hat{\beta}}^2$

ערכית תחזית וקריאת פלטים (תוכנת SPSS):

אמידה נקודתית :

אמידה נקודתית עבור X_0 מסוים (תחזית).

מחושבת על פי הרגסיה במדגם : $\hat{Y}_0 = \hat{\alpha} + \hat{\beta} \cdot X_0$.

אמידת מרוחק ל- $E(Y)$:

אמידת התחזית באוכלוסייה עבור X_0 מסוים. נחשב רוח בר סמך לערך ממוצע של Y

באוכי עבור X_0 מסוים ($E(Y)$) ברמת סמך $1-\alpha$.

$$\hat{Y} \pm t_{n-2; 1-\frac{\alpha}{2}} \hat{\sigma}_u \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum(X_i - \bar{X})^2}}$$

נוסחת הרב"ס :

$$\hat{\sigma}_u = MSE = \frac{SSE}{n-2}, \quad \sum(X_i - \bar{X})^2 = S_{xx} = (n-1)S_x^2$$

$$\text{רישום הרב"ס} : p(\underline{\quad} \leq E(Y) \leq \underline{\quad}) = 1-\alpha$$

אמידת מרוחק ל- Y :

אמידת ערך בודד של Y באוכלוסייה עבור X_0 מסוים. נחשב רוח בר סמך לערך בודד

של Y באוכי עבור X_0 מסוים (Y_0) ברמת סמך $1-\alpha$.

$$\hat{Y} \pm t_{n-2; 1-\frac{\alpha}{2}} \hat{\sigma}_u \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum(X_i - \bar{X})^2}}$$

נוסחת הרב"ס :

$$\text{רישום הרב"ס} : p(\underline{\quad} \leq Y \leq \underline{\quad}) = 1-\alpha$$

- רב"ס לערך בודד יהיה רחב יותר מאשר רב"ס לערך ממוצע משומש שטויות התקן בראשון גדולה מאשר באחרון.

שאלות:**פלט ניתוח שונות:**

- 1) חוקר רצה לבחון את השפעת ההכנסה (*INCOME*) על גובה המס (*TAX*)
 $.TAX_t = \alpha + \beta \cdot INCOME_t + u_t$ לפי המודל:
 לשם כך אסף נתונים מ-51 מדינות. להלן התוצאות:

Model: MODEL1**Dependent Variable: TAX****Analysis of Variance**

Source	DF	Sum of	Mean	F Value	Prob>F
		Squares	Square		
Model	1	2046.89694	2046.89694	8798.672	0.0001
Error	49	11.39922	0.23264		
C Total	50	2058.29615			

Root MSE	0.48232	R-square	0.9945
Dep Mean	5.4242	Adj R-sq	0.9943
C.V.	8.88711		

בדקו את ההשערה כי המודל מובהק ברמת מובהקות של 0.05.

פלט מקדמי הרוגסיה:

- 2) בהמשך לדוגמא הקודמת – בדיקת השפעת ההכנסה על גודל המס, התקבלו גם התוצאות הבאות:

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	-0.086912	0.08953904	-0.971	0.3365
INCOME	1	0.152232	0.0016229	93.801	0.0001

- א. אמדנו את המודל: $.TAX = \alpha + \beta \cdot INCOME + U$: מהי המשמעות הכלכלית של β ?
- ב. האם המודל מובהק? בדקו על סמך הפלט הנ"ל ברמת מובהקות של 0.05.
- ג. מהי רמת המובהקות הקטנה ביותר, עבורה עדין תידחה השערת האפס מסעיף ב'?

- ד. בדקו את ההשערה כי ככל שההכנסה עולה כך עולה גם המס (שיוף β חיובי) ברמת מובהקות של 0.01.
- ה. בנו רוח-סמן ברמת סמן של 95% עבור β .
- ו. בדקו את ההשערה שתוספת של מיליארד \$ להכנסה תגדיל את המס ב-2 מיליארד \$, ברמת מובהקות של 0.05.

- שימוש לב Ci :

במודל עם משתנה מסביר אחד בלבד קיימת זהות בין מבחן F למובהקות המודל בין מבחן t למובהקות ה- β :

$$F_{(1,T-2;1-\alpha)} = t_{(T-2;1-\frac{\alpha}{2})}^2$$

$$F = t_{\hat{\beta}}^2$$

כלומר: כל החלטה המתקבלת במבחן אחד חייבת להיות זהה להחלטה המתקבלת במבחן השני.

פלט שוניות משותפות:

(3) נתון פلت האמידה של המודל: $Y_t = \alpha + \beta X_t + u_t$, שלצורך אמידתו נאספו 240 תצפיות:

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	
		Estimate	Error	Parameter=0	Prob> T
INTERCEP	1	5.25	0.25	21	0.0000
X	1	0.96	0.12	8	0.0000

Covariance of Estimates

	INTERCEP	X
INTERCEP	0.0625	-0.003
X	-0.003	0.0144

יש לבדוק את ההשערה: $H_0: \alpha = 5\beta$.

שאלה מס' 4:

(4) חוקר רצה לבדוק את השפעת הותק בעבודה (*SALARY*) על השכר (*EXP*) לפי המודל: $\ln(SALARY_t) = \alpha + \beta \cdot EXP_t + u_t$. הוא אסף 403 תצפיות, ואמד את הפרמטרים בתוכנת SAS. להלן חלקים מהפלט ויש להשלימו:

Analysis of Variance

Source	DF	Sum of		F Value	Prob>F
		Squares	Mean Square		
Model	---	---	5.68015	---	---
Error	---	205.22539	---		
C Total	---	---			

Root MSE	---	R-square	---
Dep Mean	7.14247	Adj R-sq	0.0245
C.V.	10.01602		

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	---	---	---	---
EXP	1	-0.008740	---	---	0.0009

Covariance of Estimates

COVB	INTERCEP	EXP
INTERCEP	0.0047463101	---
EXP	-0.000154685	6.882844 E-6

- נתון נספָּה: $EXP = 22$
- א. קיים קשר חיובי מובהק בין ותק ללוג השכר. נכון / לא נכון
- ב. שיעור התשואה בשכר לשנת ותק הוא?
- ג. תחזית לוג השכר עבור אדם בעל 10 שנות ותק היא?

ביצוע תוצאות:

5) במדגם של 30 דירות מושכrotein לסטודנטים ברדיוס של עד 2 ק"מ מסביב. למקרה נחקק הקשר בין שכר דירה למספר הסטודנטים הנרים בדירה.
להלן התוצאות:

Descriptive Statistics

	Mean	Std. Deviation	N
שכר הדירה	1386.7667	509.46027	30
מספר הסטודנטים	3.0000	1.31306	30

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.602 ^a	.362	.339	414.05503

a. Predictors: (Constant), number of students

b. Dependent Variable: rent

ANOVA^b

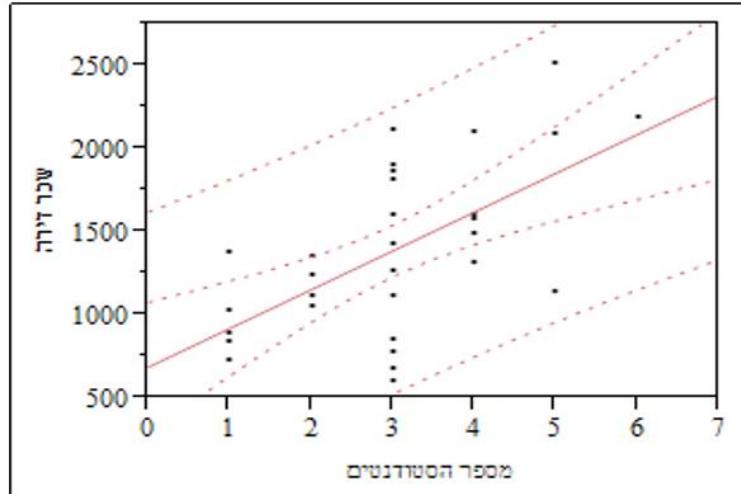
Model	Sum of Squares	Df	Mean Square	F	Sig.
1 Regression	2726579.520	1	2726579.520	15.904	.000 ^a
Residual	4800363.847	28	171441.566		
Total	7526943.367	29			

a. Predictors: (Constant), number of students

b. Dependent Variable: rent

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
1 (Constant)	686.207	191.244		3.588	.001
מספר הסטודנטים	233.520	58.556	.602	3.988	.000



- א. חשב אומדן נקודתי לשכר הדירה אותו ישלמו סטודנטים החולקים את הדירה עם שותף אחד בלבד.
- ב. אמודד את שכר הדירה הממוצע שישלמו סטודנטים החולקים את הדירה עם שותף אחד בלבד, ברמת בטיחון של .95%.
- ג. אמודד את שכר הדירה שישלם סטודנט יחיד החולק את הדירה עם שותף אחד בלבד, ברמת בטיחון של .95%.

תשובות סופיות:

- (1) יש עדות לכך.
- (2) א. ראה סרטון.
- ב. יש עדות לכך.
- ג. $P_{t_{\hat{\beta}}} = 0.0001$.
- ה. $P(0.1488 \leq \beta \leq 0.1554) = 0.95$.
- ד. יש עדות לכך.
- ו. יש עדות לכך.
- (3) אין עדות לכך.
- (4) א. לא נכון.
- .7.24735. ג. -0.87% .
- ב. $p(957.4 \leq \mu_{Y_{X=2}} \leq 1349.08) = 0.95$.
- ב. $.1153.247$.
- ג. $p(282.94 \leq Y_{X=2} \leq 2023.55) = 0.95$.

אקוונומטריקה

פרק 5 - שינוי ייחדות מדידה

תוכן העניינים

- | | |
|----------|---------------|
| 32 | 1. כללי |
|----------|---------------|

שינוי ייחידות מדידה:

רקע:

טרנספורמציה ליניארית: הוספה/חיסכונה של קבוע ו/או הכפלת/חילוק של קבוע של אחד או שני המשתנים (התלווי והבלתי).

- טרנספורמציה ליניארית של המשתנים לא תשפיע על: $t_{\hat{\beta}}, F, R^2$ ו- PF .

השינויים מסוכמים בטבלה הבאה:

$S\hat{\alpha}$	$S_{\hat{\beta}}$	$\hat{\alpha}'$	$\hat{\beta}'$	
$s_{\hat{\alpha}'} \neq s_{\hat{\alpha}}$	$s_{\hat{\beta}'} = s_{\hat{\beta}}$	$\hat{\alpha}' = \hat{\alpha} - \hat{\beta}d$	$\hat{\beta}' = \hat{\beta}$	הוספת קבוע ל- X : $Y = \alpha' + \beta'(X + d) + v$
$s_{\hat{\alpha}'} = s_{\hat{\alpha}}$	$s_{\hat{\beta}'} = s_{\hat{\beta}}$	$\hat{\alpha}' = \hat{\alpha} + d$	$\hat{\beta}' = \hat{\beta}$	הוספת קבוע ל- Y : $Y + d = \alpha' + \beta'X + v$
$s_{\hat{\alpha}'} = s_{\hat{\alpha}}$	$s_{\hat{\beta}'} = \frac{s_{\hat{\beta}}}{d}$	$\hat{\alpha}' = \hat{\alpha}$	$\hat{\beta}' = \frac{\hat{\beta}}{d}$	הכפלת X פי קבוע: $Y = \alpha' + \beta'(dX) + v$
$s_{\hat{\alpha}'} = ds_{\hat{\alpha}}$	$s_{\hat{\beta}'} = ds_{\hat{\beta}}$	$\hat{\alpha}' = d\hat{\alpha}$	$\hat{\beta}' = d\hat{\beta}$	הכפלת Y פי קבוע: $dY = \alpha' + \beta'X + v$

• $t_{(\hat{\beta}'=0)} = t_{(\hat{\beta}=0)}$ תמיד.

• רק בהכפלות $t_{(\hat{\alpha}'=0)} = t_{(\hat{\alpha}=0)}$.

שאלות:

(1) חוקר ביקש לאמוד את הקשר בין שכר ב-₪ (*MWAGE*) לבין שנות לימוד (*SCL*) במאזעות 2 מודלים שונים.

להלן תוצאות האמידה:

$$\text{א. } MWAGE_t = 139.54 + 118.62 \cdot SCL_t$$

$$\text{ב. } MWAGE_t = -1445.08 + 1239.60 \cdot LN(SCL)_t$$

חשבו מחדש את מקדמי הרגרסיה וסטטיסטי המבחן *F* בכל אחד מהמודלים כתוצאה:

1. התברר כי נעשו טעות בחישוב מספר שנות הלימוד, ויש צורך להוסיף 20% למשתנה המקורי.

2. התברר כי הקשר בין שכר לשנות לימוד הוא ריבועי ולכן יש צורך להעלות את המשנה המקורי של מספר שנות הלימוד בריבוע.

(2) בהמשך לנוטוני השאלה לדוגמא מהפרק החמישי:

חוקר טען כי יש לבדוק את הקשר בין שכר לוותק ע"י שימוש בשכר נטו (*NET*) ולא בשכר ברוטו (*SALARY*). (קיים שיעור מס קבוע של 20%).

$$\text{המודל הוא: } \ln(NET_t) = \alpha' + \beta' \cdot EXP_t + \nu_t$$

מה יהיו ערכי האומדים, סטיות התקן שלהם וטיב ההתאמה באמידת מודל זה?

תשובות סופיות:

(1) א. $\hat{\beta}' = 98.85$, $\hat{\alpha}' = 139.59$, סטטיסטי *F* לא משתנה.

ב. $\hat{\beta}' = 1239.6$, $\hat{\alpha}' = -1671$, סטטיסטי *F* לא משתנה.

2. לא ניתן לדעת.

$\hat{\beta}' = \hat{\beta} = -0.00874$, $\hat{\alpha}' = 7.11161$, $S_{\hat{\beta}'} = S_{\hat{\beta}} = 0.0026235$, $S_{\hat{\alpha}'} = S_{\hat{\alpha}} = 0.0688935$ (2)

$$R^2 = 0.0269$$

אקוונומטריקה

פרק 6 - רגרסיה מרובה

תוכן העניינים

34 1. רגרסיה מרובה

גרסיה מרובה:

רקע:

מבחן T ו-F:

כאשר יש יותר משתנה מסביר אחד, מדובר ברגסיה מרובה.
המודל הקלاسي: $Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_k X_{kt} + u_t$

- קבוע α יש אחד.
- מספר ה- β טות כמספר המשתנים הב'ית במודל.

מבחן F לモביהקוט המודל :

$$\begin{aligned} H_0 &= \beta_1 = \beta_2 = \dots = \beta_k = 0 \\ \text{השערות: } H_1 &: \text{OTHERWISE} \end{aligned}$$

סטטיסטי המבחן F וככל ההכרעה :

$$F = \frac{\frac{RSS}{k}}{\frac{ESS}{T-k-1}} = \frac{\frac{R^2}{k}}{\frac{1-R^2}{T-k-1}} > F(k, T-k-1; 1-\alpha)$$

מבחן t לモביהקוט ה- β טות :

מבחן לבדיקת מוביהקוט β ספציפית :

$$\begin{aligned} H_0 &= \beta_1 = 0 \\ \text{השערות: } H_1 &: \beta_1 \neq 0 \end{aligned}$$

סטטיסטי המבחן t וככל ההכרעה :

$$\left| t_{\hat{\beta}_i} \right| = \left| \frac{\hat{\beta}_i}{S_{\hat{\beta}_i}} \right| > t_{(T-k-1; 1-\alpha/2)}$$

השוואה בין מודלים – \bar{R}^2 וחוק חיטובסקי:

בכדי להחליט האם כדאי לנו להוסיף למודל משתנה ב'ית מסויים':
 השוו את פרופורציית השונות הקשורות המתוקנת \bar{R}^2 בין המודל ללא המשתנה
 הקשור למועדן עם המשתנה הקשור שהואנו.

- ניתן להשתמש גם באומד המוטה - R^2 להשוואה בין מודלים אם מתקימים
 שני התנאים הבאים:
 1. מספר המשתנים זהה.
 2. המשתנה הקשור זהה.

לפי חוק חיטובסקי – בהוספת משתנה מסביר אחד בלבד למודל ה- \bar{R}^2 עולה אך
 ורק אם: $|t_{\hat{\beta}}| > 1$.

כאשר: $|t_{\hat{\beta}}| < 1$ אז \bar{R}^2 ירד בהוספת המשתנה והוא גם לא יהיה רלוונטי למודל (МОВАХ).

כאשר: $|t_{\hat{\beta}}| > 2$ אז \bar{R}^2 עולה והמשתנה שהוסף יהיה גם מובייח.

כאשר: $1 < |t_{\hat{\beta}}| < 2$ אז ה- \bar{R}^2 עולה אך יש לבדוק את רלוונטיות המשתנה שהוסף למודל
 על פי מבחן t .

שאלות:

מבחן T ו F-ו:

1) נאמד המודל : $Y_t = \alpha + \beta_x X_t + \beta_z Z_t + \beta_w W_t + \beta_s S_t + u_t$ והתקבלו התוצאות הבאות :

Analysis of Variance

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

Root MSE	-----	R-square	-----
Dep Mean	178.6645	Adj R-sq	0.999022
C.V.	0.988075		

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	5.067731	0.456604	11.09874	0.0000
X	1	-----	0.042711	22.84485	0.0000
Z	1	3.005385	0.008679	346.2721	0.0000
W	1	-5.029101	0.073149	-----	0.0000
S	1	8.974106	0.029075	308.6485	0.0000

- א. השלים את הנתונים החסרים בפלט.
- ב. האם המודל מובהק? בדקו ברמת מובהקות של 0.05.
- ג. האם משתנה W רלוונטי למודל? בדקו ברמת מובהקות של 0.01.

השוואה בין מודלים:

- (2) במודל לניבוי ההכנסה על פי שנות לימוד וותק במקום העבודה, התקבל: $\bar{R}^2 = 0.266$. הוסף המשתנה היקף המשרחה. ב מבחון לモובחיקות המשתנה הנוסף התקבל: $t_{\hat{\beta}} = 0.456$. האם ערך \bar{R}^2 יעליה/יריד/לא ישנה בהוספה המשתנה הנוסף למודל?

 מבחון Wald ו-T מורכב:

- (3) נאמד המודל: $Y_t = \alpha + \beta_x X_t + \beta_z Z_t + \beta_w W_t + \beta_s S_t + u_t$ והתקבלו התוצאות הבאות:

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Prob>F
Model	4	646169.84	161542.46	51835.84	0.0000
Error	199	620.1683	3.1164236		
C Total	203	646790.01			
Root MSE		1.7653395		R-square	0.999041
Dep Mean		178.6645		Adj R-sq	0.999022
C.V.		0.988075			

Parameter Estimates

Variable	DF	Parameter	T for H0:			
			Estimate	Error	Parameter=0	Prob> T
INTERCEP	1	5.067731	0.456604	11.09874	0.0000	
X	1	0.975736	0.042711	22.84485	0.0000	
Z	1	3.005385	0.008679	346.2721	0.0000	
W	1	-5.029101	0.073149	-68.75141	0.0000	
S	1	8.974106	0.029075	308.6485	0.0000	

הוועלה ההשערה כי ההשפעה על Y של משתנה S היא פי 3 מזו של משתנה Z, וכן כי החוווץ הוא 5.

- א. מהי השערת האפס?
ב. מהו המודל המוגבל שאותו צריך לאמוד?

להלן אמידת המודל המוגבל:

Analysis of Variance

Source	DF	Sum of	Mean	F Value	Prob>F
		Squares	Square		
Model	2	646166.01	323083.01		
Error	201	623.9983	3.104469		
C Total	203	646790.01			
Root MSE		1.7619504		R-square	0.999035
Dep Mean		173.6645		Adj R-sq	0.999026
C.V.		1.0145714			

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
X	1	0.978491	0.036399	26.88240	0.0000
Z+3S	1	2.999995	0.003669	817.6080	0.0000
W	1	-5.043109	0.071218	-70.81249	0.0000

- ג. חשב את הסטטיסטי של WALD.
- ד. כמה דרגות חופש יש במונה וכמה במכנה?
- ה. האם דוחים או מקבלים את השערת האפס?

על מנת לאמוד את פונקציית התצורך נאספו נתונים על 42 משקי בית בשנת 4)

$$\text{. } C_t = \alpha + \beta_1 \cdot W_t + \beta_2 \cdot P_t + u_t \text{ ונאמדה המשווה הבאה :}$$

להלן תוצאות האמידה של המשווה הניל:

Analysis of Variance

Source	DF	Sum of Squares	Mean		
			Square	F Value	Prob>F
Model	---	-----	-----	-----	-----
Error	---	-----	52968		
C Total	---	-----			

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0:	
				Parameter=0	Prob> T
INTERCEP	1	-107.226	-----	-----	-----
W	1	0.743	-----		
P	1	0.561	-----		

Covariance of Estimates

COV	INTERCEP	W	P
INTERCEP	-----	-----	-----
W	-----	0.0046	-0.0090
P	-----	-0.0090	0.016

על מנת לבדוק את ההשערה שהנתיחה השולית לצרוך מתוך ההכנסה זהה

לנתיחה השולית לצרוך מתוך ההון, נאמדה גם המשווה הבאה :

$$\text{. } C_t = \alpha + \beta_1 \cdot Y_t + u_t \text{, כאשר : } Y_t = \text{סה"כ ההכנסה של משק בית t.}$$

התΚבל : $ESS = 0.4566$

בדקו את ההשערה בשתי דרכים.

תרגיל מסכם:

- 5) חוקר אמד את התצרוכת של 500 משקי בית כפונקציה של הכנסה שלהם לפי המשוואה : $EXPENSE_t = \alpha + \beta \cdot INCOME_t + u_t$
- התצרוכת של משק הבית ה- t -י באלפי שקליםים . $EXPENSE_t$
- הכנסה של משק הבית ה- t -י באלפי שקליםים . $INCOME_t$
- ההפרעות האקראיות מקיימות את כל ההנחה הקלאסיות התקבל הפלט הבא :

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Prob>F
Model	1	2013.105	2013.105	6495.745	0.0000
Error	498	154.3358	0.3099112		
C Total	499	2167.441			
Root MSE		0.556697	R-square	0.928794	
Dep Mean		3.990208	Adj R-sq	0.928651	
C.V.		13.95157			

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	0.041995	0.054951	0.764236	0.4451
INCOME	1	0.713503	0.008853	80.59618	0.0000

- א. מהו Pvalue לבדיקה מובהקות המודל ע"י מבחן F?
- ב. מהו אחוז השינוי בתצרוכת המוסבר ע"י הכנסה?
- ג. מהו אומדן לתצרוכת ההתחלתיות של משק בית?
- ד. האם אומדן זה מובהק?
- ה. על עוזר מחקר הטיל החוקר לבדוק את השערה כי על כל 1000 ש"נ נוספים בהכנסה צורך הפרט 700 ש"נ, נגד השערה כי הוא צורך יותר מ-700 ש"נ. נסח את השערת האפס ואת השערה האלטרנטיבית.
- ו. מהו הסטטיסטי t לבדיקה השערה?
- ז. מהו הסטטיסטי $WALD$ לבדיקה השערה?
- ח. התברר כי הייתה טעות בנתונים, וכי יש להוציא 1000 ש"נ לתצרוכת של כל משק בית :
- ט. ההוספה תגדיל את האומד $\hat{\alpha}$: נכון / לא נכון / לא ניתן לדעת.

- ii. בעקבות ההוספה האומד ל- α נכון / לא נכון / לא ניתן לדעת. יהיה מובחן :
- iii. ההוספה תנסה את האומד ל- β : נכון / לא נכון / לא ניתן לדעת.
- iv. ההוספה תנסה את R^2 : נכון / לא נכון / לא ניתן לדעת.

החוקר טען כי יש להווסף לפונקציית התצורך גם את השפעת העושר העוזר של משק בית מורכב מתוכניות החסכון שלו (*SAVINGS*) ומニアרות הערך שיש לו (*NE*). שתי סדרות הנתונים הן באלפי שקלים.

החוקר אמד את המשוואה :

$$EXPENSE_t = \alpha + \beta_1 \cdot INCOME_t + \beta_2 \cdot SAVINGS_t + \beta_3 \cdot NE_t + u_t$$

וקיבל כי סכום ריבועי הסטיות של הטיעויות הוא 121.

ט. מהי השערת האפס לבדיקת הטענה של החוקר (שהמודל החדש נכון ולא המקורי)?

י. מהו הסטטיסטי *WALD* לבדיקת ההשערה?

החוקר רצה לבדוק את ההשערה כי הנש"ץ מתוך הכנסה שווה ל-6.0 וכי השפעת ניארות הערך על התצורך היא פי 2 מהשפעת תוכניות החסכון.

יא. מהי השערת האפס לבדיקה זו?

יב. המודל המוגבל לבדיקת ההשערה יהיה מהצורה :

$$\text{בטא את } Z_t, \text{ ו- } W_t \text{ באמצעות המשתנים המקוריים.}$$

תשובות סופיות:

. א. (1)

Analysis of Variance

Source	DF	Sum of	Mean	F Value	Prob>F
		Squares	Square		
Model	4	646169.84	161542.46	51835.84	0.0000
Error	199	620.1683	3.1164236		
C Total	203	646790.01			
Root MSE		1.7653395	R-square	0.999041	
Dep Mean		178.6645	Adj R-sq	0.999022	
C.V.		0.988075			

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	5.067731	0.456604	11.09874	0.0000
X	1	0.975736	0.042711	22.84485	0.0000
Z	1	3.005385	0.008679	346.2721	0.0000
W	1	-5.029101	0.073149	-68.75141	0.0000
S	1	8.974106	0.029075	308.6485	0.0000

ב. יש עדות לכך. ג. יש עדות לכך.
ירד. (2)

. $Y_t = \beta_x X_t + \beta_z (Z_t + 3S_t) + \beta_w W_t + u_t$. א. $H_0: \alpha = 5, \beta_s = 3\beta_z$ (3)

ג. מונח: 2, מכנה: -199. ד. $WALD_{stat} = 0.6145$.

ה. מקבלים.

בדיקה ע"י מבחן WALD ו-t: אין עדות לכך. (4)

א. לא. ב. $\hat{\alpha} = 0.04195$ ג. 92% . $PF = 0.000$ (5)

. $WALD_{stat} = 2.505$. $t_{\hat{\beta}} = 1.583$. $H_0: \beta = 0.70, H_1: \beta > 0.70$ ה. $\beta = 0.70$

ח. נכוון. ii. נכוון. iii. לא נכוון. iv. לא נכוון.

. $WALD_{stat} = 68.32$. $H_0: \beta_2 = \beta_3 = 0, H_1: \text{OTHERWISE}$ ט. $H_0: \beta_2 = \beta_3 = 0, H_1: \beta_1 = 0.6$

. $W_t = SAVINGS_t + 2 \cdot NE_t, Z_t = EXPENCE_t - 0.6 \cdot INCOME_t$ ב.

אקוונומטריקה

פרק 7 - מבחן LM

תוכן העניינים

- 43 1. כללי

מבחן LM:

רקע:

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

לדוגמא :

נניח שיש לנו מודל הכלול 4 משתנים מסבירים (UNRESTRICTED) :

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

לגביה השניים הראשונים אנו בטוחים כי הם רלוונטיים וחיבאים להופיע במודל. לגבי השניים האחרונים אנחנו לא בטוחים.

$$\begin{aligned} H_0 &: \beta_3 = \beta_4 = 0 \\ H_1 &: \text{OTHERWISE} \end{aligned}$$

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 הרגرسיבית : $\hat{Y}_t - \hat{u}_t$.

כעת אומדים את רגרסיבית העזר שבנה מנסים לנבא את הסטייה מקו הרגסיבית עבור כל תצפית : $\omega_t = \delta_0 + \delta_1 x_{1t} + \delta_2 x_{2t} + \delta_3 x_{3t} + \delta_4 x_{4t}$.

חישוב הסטטיסטי : $(R^2 \text{ של רגרסיבית העזר } * \text{ מספר התצפיות}) \cdot T$.

כלל הכלעה : אם $LM_{stat} > \chi_m^2$ נדחה את H_0 (m = מס' ההגבלות ב- H_0).

- שימוש לב Ci :

עבור המשתנים הנוספים למודל – כל המדדים (הבטות, ערכי t וה- $Pvalu$)

ברגרסיבית העזר שווים לאלו של הרגסיבית הלא מוגבלת.

עבור המשתנים הקיימים במודל – המדדים אינם שווים בין שתי הרגסיביות.

שאלות:

(1) נניח מודל הכלול 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		Mean	
		Squares	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	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
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		F Value	Prob>F
		Squares	Mean Square		
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	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
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

רגרסיות עזר

Dependent variable :RES

Analysis of Variance

Source	DF	Sum of		F Value	Prob>F
		Squares	Mean Square		
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	Standard	T for H0:	
		Estimate	Error	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

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

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

$$\hat{u}_t = \hat{\delta}_0 + \hat{\delta}_1 x_{1t} + \hat{\delta}_2 x_{2t} + \hat{\delta}_3 x_{3t} + \hat{\delta}_4 x_{4t} + \hat{o}_t : \text{עזרה}$$

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

ה. שחזרו בעזרת המשוואה האחורונה (רגression העזר) את $WALD_{stat}$.

תשובות סופיות:

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

$$\cdot pt_{\hat{\beta}_3} = pt_{\hat{\beta}_4} = 0.00$$

ב. עזר. $U=R+R_{\text{עזרה}}$

$$\cdot ESS_U = ESS_{\text{עזרה}}$$

$$\cdot ESS_R = TSS_{\text{עזרה}}$$

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

$$\cdot LM_{stat} = 43.489 . v$$

$$\cdot WALD_{stat} = 26.962 . vi$$

אקוונומטריקה

פרק 8 - בעיות ספציפיקציה

תוכן העניינים

- 47 1. תיאוריה

בעיות ספציפיקציה:

רקע:

טעויות ספציפיקציה הן טעויות בניסוח משווהת הרגרסיבית.

1. הוספת משתנה לא רלוונטי :

למשל, המודל האמתי: $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$.

המודל הנאמד (הטועתי): $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \varepsilon$.

אם קיבל את H_0 בבחן t לモבוקות β_3 נסיק כי המשתנה איננו רלוונטי ונאמד את המודל מחדש ללא המשתנה השלישי. אולם, גם אם לא נוכל לאמוד מחדש, הימצאותו של משתנה שאיננו רלוונטי במודל הרגRESSEDIVE אינה פוגמת ברלוונטיות של המשתנים האחרים במודל ולא בתכונות החיוניות ל מבחני המובוקות שלהם.

2. השמטת משתנה רלוונטי :

למשל, המודל האמתי: $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$.

המודל הנאמד (הטועתי): $Y_t = \alpha + \beta_1 x_{1t} + \varepsilon$.

בහיעדר x_2 , בדיקות ההשערות לפרמטרים של המודל הטועתי אינן תקפות:

אומד לשינוי הפרמטרים	אומד ל- α	אומד ל- β_1	
מויטה אלא אם: $\bar{x}_2 = 0$	חוור הטיה		$S_{12} = 0$
מויטה (כלפי מעלה)	מויטה <u>כיוון החטיה:</u> חיובי: S_{12} ו- β_2 שווי סימן שלילי: S_{12} ו- β_2 מנוגדי סימן		$S_{12} \neq 0$

אקוונומטריקה

פרק 9 - תיאוריה מולטיקוליניארית

תוכן העניינים

1. כללי

48

מולטיקוליניאריות:

רקע:

מולטיקוליניאריות היא תופעה סטטיסטית בעיתית המתיחסת למתאם בין המשתנים המסבירים במודל.

נבחין בין מולטיקוליניאריות מלאה לחלקית.

מולטיקוליניאריות מלאה:

מתאים מלא בין המשתנים המסבירים במודל.

הדבר קורה כאשר משתנה מסביר אחד הוא קומבינציה ליניארית מלאה של המשתנה המסביר השני: $x_1 = a + bx_2$ (x_1 הוא קומבינציה ליניארית מלאה של x_2)
מכאן ש: $r_{12} = 1$.

- שימוש לב Ci מזובר בטרנספורמציה ליניארית ולא בטרנספורמציה אחרת (למשל: $x_2^2 = x_1$), אז בהכרח: $r_{12} \neq 1$.

במצב של מולטיקוליניאריות מלאה אין כל השפעה של המשתנה האחד מעבר לשני.
מדוע זה בעיתתי?

כיוון שלא ניתן לאמוד את המודל שכן אר"פ אינס מוגדרים.
פתרון: הורדת אחד המשתנים ומידת המשווה מוחדש בפועל.

מולטיקוליניאריות חלקית:

כאשר יש מתאם גבוה מאוד בין משתנים מסבירים במודל (אך לא מושלם) עלולה להיווצר בעיה של מולטיקוליניאריות חלקית.

מכיוון שיש מתאם גבוה בין המשתנים הב"ית לא נוכל לבדוק באופן מלא את ההשפעה המדויקת של כל אחד מהם על ציוני המשתנה התלווי.
כל אחד מהמשתנים הב"ית "יגזולי" מן ההשפעה הייחודית שיש למשתנה הב"ית השני על המשתנה התלווי, כך שבסתו של דבר, למרות שהמודל עם שני המשתנים הב"ית יהיה מובhawk, התרומה הייחודית של כל משתנה ב"ית לניבוי התלווי לא תהיה מובhawkת.

זיהוי מולטיקוליניאריות חלקית :

1. כאשר קיימת סטייה בין התוצאה ב מבחן F לモובקהות המודל (המודל מובחן) לבין מבחני t לモובקהות השיפועים (אף אחד מן השיפועים אינו מובחן).

הסתירה נוצרת כתוצאה מהגדלת השונות של כל אחד מהSHIPועים בשל המתאם הגובה בין הב"ית, באופן שלא מאפשר לדוחות את השערת האפס

$$\text{لمובקהות השיפועים : } t = \frac{\hat{\beta}_1}{S_{\hat{\beta}_1}}, S_{\hat{\beta}_1}^2 = \frac{MSE}{SSX_1(1 - r_{12})}$$

2. רגישות לSPECIFICITY – הורדת משתנה ב"ית שאינו מובחן תהפהך משתנים ב"ית אחרים במודל לMOובחים. אם אין בעיה של מולטיקוליניאריות, הורדת משתנים ב"ית שאינם רלוונטיים מהמודל, לא אמורה להשפיע על MOובקהותם של המשתנים הב"ית האחרים.

3. סימנים הפוכים – כאשר השיפועים של המשתנים הב"ית מקבלים סימנים הפוכים מכיוון ההשפעה שלהם על המשתנה התלו. אם למשל, x_1 משפיע חיובית על y ואילו x_2 משפיע שלילית על y אבל הם יופיעו במשווה הרגרסיה עם סימנים הפוכים ($\hat{\beta}_1$ שלילית ואילו $\hat{\beta}_2$ חיובית), יש לחשוד שקיימת בעיה.

השלכות של מולטיקוליניאריות חלקית :

מולטיקוליניאריות חלקית איננה פוגעת בתכונות של ARIF (הם נותרים LINEARיים, חסרי הטיה, יעילים ועקיבים) ולא באומד השונות של האומדים (שנותר חסר הטיה כך שבדיקת השערות תוך שימוש באומדים הללו תהיה תקפה (זאת בניגוד למולטיקוליניאיות מלאה).

במונח זהה, בעיה של מולטיקוליניאריות חלקית דומה לבעה של הוספת משתנה ב"ית שאינו רלוונטי.

פתרונות למולטיקוליניאריות חלקית :

1. ברוב המקרים נסקול להוריד את אחד המשתנים. יחד עם זאת, כאשר המובקהות של המשתנים היא גבולית: $t_{\hat{\beta}} < 1 < t_{\hat{\beta}_2}$, ניתן את שניהם בתוך המודל כיון שבסץ הכל יש עלייה ב- $AdjR^2$ (לפי חוק CHIOTOBISKI).
2. ניתן לעיתים לאחד את שני המשתנים למשנה אחד.

שלבי בדירת ההשערות:

1. מבצעים מבחן F לבדיקה מובהקות המודל.
2. במידה והמודל מובהדק, מבצעים מבחן t למובהקות כל אחד מהSHIPועים.
3. ביצוע מבחן WALD לבדיקה כל השיפועים שלא יצאו מובהקים:
 - א. אם מקבלים את H_0 : אין סתירה בין מבחן WALD ל מבחני t - אין בעיה של מולטיקוליניאריות חלקית, נוריד את קבוצת המשתנים הלא רלוונטיים מהמודל.
 - ב. אם דוחים את H_0 : יש סתירה בין מבחן WALD ל מבחני t - קיימת בעיה של מולטיקוליניאריות חלקית, יש להוריד מן המודל כל פעם משתנה אחד ולבצע מבחן WALD בפועל, עד ש모זהים את המשתנה / משתנים שיש להוריד מהמודל.

אקוונומטריקה

פרק 10 - סיכום ותרגול של בעיות ספציפיקציה ומולטיקוליניאריות

תוכן העניינים

- 51 1. כללי

סיכום ותרגול של בעיות ספציפיקציה ומולטיקוליניאריות:

רקע:

פתרונות	השלכות					זיהוי	הגדרה	הבעיה															
*הורדת המשנה	<p>ניתן לבצע בדיקת השערות</p> <p>אר"פ $(\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2)$ חסרי הטיה</p> <p>$(S_{\hat{\alpha}}^2, S_{\hat{\beta}_1}^2, S_{\hat{\beta}_2}^2)$ אומדי השונות</p> <p>חסרי הטיה</p>					H_0 קבלת t במבחן לモבוקות β_2	המודל האמיתי: $Y_t = \alpha + \beta_1 x_{1t} + \varepsilon$ המודל הנאמד (הטועני): $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$	הוספת משתנה לא רלוונטי															
הוספת המשתנה	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 5px;">אומד לשונות הפרמטרים</td> <td style="padding: 5px;">α</td> <td style="padding: 5px;">β_1</td> <td style="padding: 5px;">אומד ל-x_2</td> <td style="padding: 5px; text-align: right;">בහיעדר: x_2</td> </tr> <tr> <td style="padding: 5px;">מוטה חיובית</td> <td style="padding: 5px;">מוטה: אלא אם: $\bar{x}_2 = 0$</td> <td style="padding: 5px;">חסר הטיה</td> <td style="padding: 5px;">$S_{12} = 0$</td> <td style="padding: 5px;"></td> </tr> <tr> <td style="padding: 5px;">מוטה חיובית</td> <td style="padding: 5px;">מוטה חיובית: $\beta_2 > S_{12}$</td> <td style="padding: 5px;">מוטה שילית: $\beta_2 < S_{12}$</td> <td style="padding: 5px;">$S_{12} \neq 0$</td> <td style="padding: 5px;"></td> </tr> </table> <p style="text-align: center; margin-top: 10px;">לא ניתן לבצע בדיקת השערות</p>					אומד לשונות הפרמטרים	α	β_1	אומד ל- x_2	בහיעדר: x_2	מוטה חיובית	מוטה: אלא אם: $\bar{x}_2 = 0$	חסר הטיה	$S_{12} = 0$		מוטה חיובית	מוטה חיובית: $\beta_2 > S_{12}$	מוטה שילית: $\beta_2 < S_{12}$	$S_{12} \neq 0$		H_0 דחיית t במבחן לモבוקות β_2	המודל האמיתי: $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$ המודל הנאמד (הטועני): $Y_t = \alpha + \beta_1 x_{1t} + \varepsilon$	השמטה משתנה רלוונטי
אומד לשונות הפרמטרים	α	β_1	אומד ל- x_2	בහיעדר: x_2																			
מוטה חיובית	מוטה: אלא אם: $\bar{x}_2 = 0$	חסר הטיה	$S_{12} = 0$																				
מוטה חיובית	מוטה חיובית: $\beta_2 > S_{12}$	מוטה שילית: $\beta_2 < S_{12}$	$S_{12} \neq 0$																				
הורדת אחד המשתנים	<p>לא ניתן לבצע בדיקת השערות</p> <p>אר"פ $(\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2)$ בלתי מוגדרים.</p>					$x_1 = a + bx_2$ $r_{12} = 1$: אם:	מתאים מלא בין המשתנים המסבירים במודל $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$ $r_{12} = \pm 1$ כאשר:	מולטיקוליניאריות מלאה															
**הורדת אחד המשתנים או איחוד	<p>ניתן לבצע בדיקת השערות</p> <p>אין פגיעה בתוכנות אר"פ ושונוות</p>					א. סתירה בין mbחן F ל-t ב. רגישות לספציפיקציה ג. סימנים הפוכיים	מתאים חזק בין המשתנים המסבירים במודל $Y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$ $0.7 < r_{12} < 1$ כאשר:	מולטיקוליניאריות חלוקת															

- * במידה ומובהקות גבולית ($t_{\hat{\beta}} < 2$) נסקול להשאיר משתנה לא רלוונטי כי מעלה את $AdjR^2$ (חוק חיטובסקי).
- * במידה ומובהקותם גבולית ($t_{\hat{\beta}} < 2$) נסקול להשאיר את שניהם בשל העלייה ב- $AdjR^2$ (חוק חיטובסקי).

שאלות:

1) להלן מודל של שכר W_t , כפונקציה של שנות לימוד S_t :

$$. W_t = \alpha + \beta \cdot S_t + u_t . \quad 1$$

להלן מודל של שכר W_t , כפונקציה של שנות לימוד S_t ושל גיל A_t :

$$. W_t = \alpha + \beta_1 \cdot S_t + \beta_2 \cdot A_t + v_t . \quad 2$$

כל האומדדים חיוביים ומובהקים וקיים קשר שלילי בין גיל להשכלה.

א. $\hat{\beta}_1$ במשווהה (1) הוא :

i. אומד חסר הטיה.

ii. אומד מוטה שלילית.

iii. אומד מוטה חיובית.

iv. אומד מוטה, אך לא ניתן לדעת את כיוון ההטיה.

ב. ניתן להשתמש בבחן χ^2 לבדיקת מובהקות

נכוע/לא נכון/לא ניתן לדעת השיפוע במשווהה (1).

ג. בנוסף למשתנים במשווהה השנייה, הchlilit החוקר להוסיף גם את משתנה הוותק, EXP_t . מכיוון שלא היו בידו נתונים על הוותק, הchlilit החוקר להערכו עבור כל עובד על ידי הגיל של העובד פחות 24 שנים (מתוך ההנחה שהחכים המוצעאים מתחלים בגיל זה לערך).

להלן משווהה מס' 3 :

$$. W_t = \alpha + \beta_1 \cdot S_t + \beta_2 \cdot A_t + \beta_3 \cdot EXP_t + w_t . \quad 3$$

חווה דעתך על המשווהה השלישית.

2) נתונות ארבע משוואות הרגרסיה הבאות (כאשר הסטיות במודל האמתי מקיימות את הנחות הרגרסיה הקלאליסיות) :

$$. \sum \hat{V}_t^2 = \sum (X_{2t} - \bar{X}_{2t})^2 . \text{ כאשר התקבל : } X_{2t} = \lambda + \delta \cdot X_{1t} + V_t . \quad 1$$

$$. Y_t = \alpha + \beta_1 \cdot X_{1t} + \beta_2 \cdot X_{2t} + U_t . \quad 2$$

(10.3) (19.8)

$$. Y_t = \alpha + \beta_1 \cdot X_{1t} + \beta_2 \cdot X_{2t} + \beta_3 \cdot X_{3t} + W_t . \quad 3$$

(9.9) (17.3) (0.37)

$$. Y_t = \alpha + \beta_1 \cdot X_{1t} + \sum_t \beta_2 \cdot X_{2t} + \beta_3 \cdot X_{3t} + \beta_4 \cdot X_{4t} + Z_t . \quad 4$$

(6.3)

(המספרים בסוגרים הם ערכי t של אומדני המקדמים).

לגביה הטענות הבאות, קבעו לגבי כל טענה אם היא נכונה או לא, והסבירו :

א. האומד של β_1 במשווהה (2) הינו חסר הטיה, אך אומד השונות של β_1 מוטה.

ב. האומד של β_1 במשווהה (3) הינו חסר הטיה, אך אומד השונות של β_1 מוטה.

- ג. האומד של β_1 במשווהה (4) הינו חסר הטיה, אך אומד השונות של β_1 מוטה.
- ד. האומדן $\hat{\beta}_1$ במשווהה (4) זהה ל- $\hat{\beta}_1$ במשווהה (2).
- ה. השונות התיאורטיבית של האומדן $\hat{\beta}_1$ במשווהה (4) זהה לשונות התיאורטיבית של $\hat{\beta}_1$ במשווהה (2), אך אומדני השונות שונים.
- ו. האומד ל- α במשווהה (4) הינו חסר הטיה.
- ז. האומד ל- α במשווהה (3) הינו חסר הטיה.
- ח. R^2 של משווהה (2) גדול מ- R^2 של משווהה (3).
- ט. \bar{R}^2 של משווהה (2) גדול מ- \bar{R}^2 של משווהה (3).

3) נתון המודל : $Y_t = \alpha + \beta_1 \cdot X_{1t} + \beta_2 \cdot X_{2t} + U_t$.
חוו דעתכם על הטענות הבאות (כל סעיף עומד בפני עצמו):

A. בהנחה כי מתקיים: $R^2 = 0.92$ $Y_t = \alpha + \beta_1 \cdot X_{1t} + \beta_2 \cdot X_{2t} + U_t$ (0.5) (0.3)

הערכתים בסוגרים הם ערכי t.
למובחיקות הבוטות יש טעות במודל
כימודל מובחיק והמקדים לא:
נכון/לא נכון/לא ניתן לדעת

B. בהנחה כי מתקיים: $1 = X_{1t} - 2X_{2t}$ לא ניתן לאמוד
את המודל בשיטת הריבועים הפחותים:
נכון/לא נכון/לא ניתן לדעת

C. בהנחה כי מתקיים: $X_{2t}^2 = X_{1t}$ לא ניתן לאמוד
את המודל בשיטת הריבועים הפחותים:
נכון/לא נכון/לא ניתן לדעת

D. הוכיחו תשובהיכם לסעיפים א' ו-ב'.

E. בהנחה כי מתקיים: $r_{12} = 0.98$.

F. לא ניתן לאמוד את המודל בשיטת
הריבועים הפחותים:
נכון/לא נכון/לא ניתן לדעת.

G. איזו בעיה עלולה להיווצר במודל ומהן השלכותיה.

H. בהנחה שהמודל יצא מובחיק אולם הבוטות אינן מובחיקות וערכי t
למובחיקות הבוטות הן כدلמן: $t_{\hat{\beta}_1} = 1.31$, $t_{\hat{\beta}_2} = 1.45$, מה יהיה הפתרון

הטוב ביותר, לדעתכם, לבעה במודל (אליה התיחסתם בסעיף ii)?

1. להוריד את x_1 .

2. להוריד את x_2 .

3. להוריד את שני המשתנים.

4. להותיר את שני המשתנים.

תשובות סופיות:

- (1) א. ii. ג. קיימות בעיות מולטיקוליניאריות מלאה.
 ב. לא נכון. ה. נכון.
- (2) א. לא נכון. ג. נכון. ד. נכון. ה. נכון.
 ו. לא נכון. ז. נכון. ט. נכון.
- (3) א. לא נכון. ד. הוכחה. ג. נכון. ה. לא נכון.
 ii. מולטיקוליניאריות חלקית. iii. 4.

אקוונומטריקה

פרק 11 - משתנה דמי

תוכן העניינים

1. כללי

56

משתנה דמי:

רקע:

הכנסת משתנים ב"ת איקוטיים למודל הרגressive.

למשל, נתונה משווהת הרגressive: $W_t = \alpha + \beta \cdot S_t$.

W_t = השכר (התלווי).

S_t = שנות לימוד (הבטח) שניהם כמותיים.

נניח שאנו סבורים שגם משתנה המגדר (משתנה איקוטי) משפיע על השכר.

כדי להכין למשווהת הרגressive יש להגדיר משתני דמי (dummy variable):

נדיר משתנה D שיקבל את הערך 0 אם מדובר בא"יישה" ואת הערך 1 אם מדובר ב"גבר".

ניתן להכין את המשתנה הדמי למודל בשלושה אופנים שונים:

1. משתנה דמי לחותך – המגדר משפיע על השכר התחלתי בלבד.
2. משתנה דמי לשיפור – המגדר משפיע על התוספת לשכר בגין שנות הלימוד.
3. משתנה דמי לכל הפונקציה – המגדר משפיע גם על החותך וגם על השיפור.

משתנה דמי לחותך:

המין משפיע על השכר התחלתי בלבד.

המודל: $W_t = \alpha_0 + \alpha_1 D + \beta \cdot S_t + u_t$ החותך מייצג כאן את השכר התחלתי.

שכר התחלתי של אישה: α_0 .

שכר התחלתי של גבר: $\alpha_0 + \alpha_1$.

הבדל בשכר בין נשים וגברים: α_1 (הפרש בין החותכים).

בדיקות השערות על משתנה הדמי: מבחן t לモבוקות הפרש החותכים: $H_0: \alpha_1 = 0$.

- השיפור מייצג את התוספת בשכר כפונקציה של מס' שנות הלימוד והוא זהה עבור נשים ובברים.

פונקציית רgressie המכילה משתנים איקוטיים בלבד:

המגדר הוא המשתנה היחיד במשווהת: $W_t = \alpha_0 + \alpha_1 D + u_t$

החותך מייצג כאן את השכר הממוצע עבור כל קטgorיה:

שכר הממוצע של אישה: α_0 .

שכר הממוצע של גבר: $\alpha_0 + \alpha_1$.

הבדל בשכר הממוצע בין נשים וגברים: α_1 (הפרש בין החותכים).

בדיקות השערות על משתנה הדמי: מבחן t : $H_0: \alpha_1 = 0$ (מבחן זהה למבחן t להבדל בין ממוצעים).

משתנה דמי לשיפור:

- . $W_t = \alpha + \beta_0 S_t + \beta_1 DS_t + u_t$ המגדר משפייע על התוספת לשכר בגין שנות הלימוד : β_1 השיפור מיציג בכך את התוספת לשכר בגין שנות לימוד.
- אצל אישה : התוספת לשכר בגין שנות לימוד : β_0 .
- אצל גבר : התוספת לשכר בגין שנות לימוד : $\beta_1 + \beta_0$.
- הבדל בין גברים לנשים בתוספת לשכר בגין שנות הלימוד : β_1 (הפרש השיפורים).
- בדיקות השערות על משתנה הדמי : מבחון t לモבפקות הפרש השיפורים : $H_0 : \beta_1 = 0$.
- החותך, המיציג את השכר ההתחלתי, יהיה זהה עבור גברים ונשים.

משתנה דמי לכל הפונקציה:

- המין משפייע גם על החותך וגם על השיפור – גם על השכר ההתחלתי וגם על התוספת לשכר ההתחלתי בגין שנות הלימוד.
- המודל : $W_t = \alpha_0 + \alpha_1 D + \beta_0 S_t + \beta_1 DS_t + u_t$ השכר ההתחלתי של אישה : α_0 .
- השכר ההתחלתי של גבר : $\alpha_0 + \alpha_1$.
- הבדל בשכר ההתחלתי בין המינים : α_1 (הבדל בחותכים).
- אצל אישה : התוספת לשכר בגין שנות הלימוד : β_0 .
- אצל גבר : התוספת לשכר בגין שנות הלימוד : $\beta_1 + \beta_0$.
- הבדל בין המינים בתוספת לשכר בגין שנות הלימוד : β_1 (הבדל בשיפורים).

2 דרכי לבדיקה האם יש השפעה למשתנה האיקומי :

1. בדיקת השערות למשתני הדמי :
באמצעות מבחון WALD יש לבדוק : $H_0 : \alpha_1 = \beta_1 = 0$ לפחות אחד הפרמטרים שונה מ-0 : H_1 .
אם דוחים את השערת האפס, יש לבצע מבחני t עבור כל אחד מהפרמטרים בפרט : $H_0 : \beta_1 = 0$ ו- $H_0 : \alpha_1 = 0$.
2. מבחון CHOW :
דרך נוספת לבדיקה ההבדל בין הקטגוריות ללא יצירת משתני דמי :
חלוקת המדגם לפי הקטגוריות של המשתנה האיקומי.
مدגם של גברים (T_f) ושל נשים (T_m).
עבור כל קבוצה לאמוד משוואות רגסיביות לניבוי שכר על ידי שנות לימוד :
נשים : $W_t = \alpha_f + \beta_f X_t + u_t$
גברים : $W_t = \alpha_m + \beta_m X_t + u_t$
השערות : $H_0 : \alpha_f = \alpha_m ; \beta_f = \beta_m$

לבדיקה ההשערה נשתמש ב מבחן CHOW (הזהה ל מבחן WALD) :
 המודל המוגבל (R) לא לוקח בחשבון את השפעת המגדר ולכן יכולות את המדגם המאוחד :

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

המודל הלא מוגבל (U) כולל את שני חלקיו המדגמים :

$$ESS_U = ESS_f + ESS_m$$

$$DF_U = DF_f + DF_m$$

$$CHOW_{stat} = \frac{\frac{ESS_R - (ESS_f + ESS_m)}{DF_R - (DF_f + DF_m)}}{\frac{ESS_f + ESS_m}{DF_f + DF_m}} = WALD_{stat}$$

למרות התוצאות הזהות בשתי הדרכים, שיטת משני הדמי עדיפה :

1. אם דחינו את H_0 ב מבחן CHOW נתקשה לברר את מקור הבדל שנמצא.
2. בהערכת שתי רגרסיות נפרדות אנו בודקים הבדל בכל הפונקציה ואילו שיטת משני הדמי מאפשרת לבדוק הבדל רק בחותך או רק בשיפוע.

סיכום ביניים:

משנה דמי לכל הפונקציה	משנה דמי לשיפוע	משנה דמי לחותך	
$Y_t = \alpha_0 + \alpha_1 D + \beta_0 X_t + \beta_1 DX_t + u_t$	$Y_t = \alpha + \beta_0 X_t + \beta_1 DX_t + u_t$	$Y_t = \alpha_0 + \alpha_1 D + \beta \cdot X_t + u_t$	המודל
קיים הבדל בין הקטגוריות במשוואת הרגרסיה כולה (בחותך ובשיפוע).	קיים הבדל בין הקטגוריות בתוספת ל-Y בגון X (בשיפוע).	קיים הבדל בין הקטגוריות ב-Y ההתחלתי (בחותך).	ההשערה במילימט
בחן WALD להפרש בין הפונקציות (חוותcis והשיפועים) : $H_0: \alpha_1 = \beta_1 = 0$ **ניתן לבדוק את ההשערה בזאת הבדל בין הפונקציות גם בבחן CHOW. אם דוחים את H_0 יש לברר את מקור הבדל באמצעות בחן t (אפשרי רק ב- $H_0: \alpha_1 = 0$: (WALD $H_0: \beta_1 = 0$)	בחן t להפרש החותcis : השיפועים : $H_0: \beta_1 = 0$	בחן t להפרש החותcis : $H_0: \alpha_1 = 0$	בדיקה ההשערה

משתני דמי אס המשנה האיקוטי יכול לקבל יותר משני ערכיים:

כאשר המשנה האיקוטי כולל יותר משני ערכיים/קטגוריות נגידיר מס' משתני דמי כמספר הקטגוריות פחות אחד.

למשל, את המשנה האיקוטי של עונות השנה הכולל 4 ערכיים : אביב, קיץ, סתיו, חורף נציג באמצעות 3 משתני דמי :

- D_1 מקבל את הערך 1 אם מדובר באביב ו-0 אחרת.
- D_2 מקבל את הערך 1 אם מדובר בקיץ ו-0 אחרת.
- D_3 מקבל את הערך 1 אם מדובר בסתיו ו-0 אחרת.

אם מדובר בחורף אז כל משתני הדמי מקבלו את הערך 0 ולכן החורף היא קבועה היחס. נניח שאנו רוצים לבדוק עונתיות במחירים הירקות :

 $V_t = \text{מדד מחירים הירקות.}$
 $p_t = \text{מדד המחרירים לצרכן.}$

1. משתני דמי לחותך :

הטענה : יש הבדל בין עונות השנה במדד התחלתי של הירקות.

המודל : $u_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \beta \cdot P_t$.

כל עליה של יחידה אחת במדד המחרירים לצרכן תעלה את מחירים הירקות ב- β . למחיר זה יתווסף α_0 בחורף, $\alpha_1 + \alpha_0$ באביב, $\alpha_2 + \alpha_0$ בקיץ ו- $\alpha_3 + \alpha_0$ בסתיו.

ניתן לראות כי : α_0 - החותך בקטgorיה שהושמטה, $\alpha_1 + \alpha_0$ - החותך בקטgorיה .

בדיקות השערות :

השערות : $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$
 $H_1 : \text{OTHERWISE}$

הבחן הסטטיסטי – מבחן WALD :

$$V_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \beta \cdot P_t + u_t : (\text{U})$$

$$V_t = \alpha + \beta \cdot P_t + u_t : (\text{R})$$

- שימושו לב שחותך במשווה המוגבלת איינו α שכן המשנה המסביר של עונות השנה ירד.

אם נדחה את H_0 ב מבחון הסטטיסטי של הסעיף הקודם, יש לבדוק מה מקור ההבדל בין החותכים על ידי מבחן t :

1. אם יש הבדל במחיר ההתחלתי של הירקות בין האביב לחורף:

$$H_0: \alpha_1 = 0$$

2. אם יש הבדל במחיר ההתחלתי של הירקות בין הקיץ לחורף:

$$H_0: \alpha_2 = 0$$

3. אם יש הבדל במחיר ההתחלתי של הירקות בין הסתיו לחורף:

$$H_0: \alpha_3 = 0$$

2. משתני דמי לשיפוע:

הטענה: יש הבדל בין עונות השנה בתוספת מחיר הירקות בגין המחיר לצרכן.

$$\text{המודל: } V_t = \alpha + \beta_0 P_t + \beta_1 (D_{1i} P_t) + \beta_2 (D_{2i} P_t) + \beta_3 (D_{3i} P_t) + u_t$$

המחיר ההתחלתי של הירקות שווה בין עונות השנה (α) אולם כל עלייה

של ייחידה אחת במידת המחיר לצרכן מעלה את מחירי הירקות

$$\text{ב: } \beta_0 + \beta_1 + \beta_2 + \beta_3 \text{ בקיץ, } \beta_0 + \beta_3 \text{ בסתיו.}$$

ניתן לראות כי $\beta_0 + \beta_1 + \beta_2 + \beta_3$ השיפוע בקטגוריה שהושמטה:

הSHIPOU בקטgorיה.

בדיקות השערות:

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_1: \text{OTHERWISE}$$

הבחן הסטטיסטי – מבחן WALD:

$$(U) : V_t = \alpha + \beta_0 P_t + \beta_1 (D_{1i} P_t) + \beta_2 (D_{2i} P_t) + \beta_3 (D_{3i} P_t) + u_t$$

$$(R) : V_t = \alpha + \beta \cdot P_t + u_t$$

- שימושו לב שהSHIPOU במשוואה המוגבלת איננו $\beta_0 + \beta_1 + \beta_2 + \beta_3$ שכן המשתנה המסביר של עונות השנה ירד.

אם נדחה את H_0 ב מבחון הסטטיסטי של הסעיף הקודם, יש לבדוק מה מקור ההבדל בין השיפועים על ידי מבחן t .

3. **משתני דמי לכל הפונקציה:**
הטענה: יש הבדל בין עונות השנה בפונקציית הרגרסיה לניבוי מחיר הירקות באמצעות המחיר לצרכן.
המודל:

$$V_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \beta_0 P_t + \beta_1 (D_{1t} P_t) + \beta_2 (D_{2t} P_t) + \beta_3 (D_{3t} P_t) + u_t$$

בדיקה השערות:

$$\text{השערות: } H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \beta_1 = \beta_2 = \beta_3 = 0$$

הבחן הסטטיסטי - מבחון WALD :

(U)

$$. V_t = \alpha_0 + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \beta_0 P_t + \beta_1 (D_{1t} P_t) + \beta_2 (D_{2t} P_t) + \beta_3 (D_{3t} P_t) + u_t$$

$$. V_t = \alpha + \beta \cdot P_t + u_t : (R)$$

אם דוחים את H_0 , יש לבדוק במבחן WALD האם ההבדל הוא בין החותכים

$$\text{או בין השיפועים: } H_0 : \beta_1 = \beta_2 = \beta_3 = 0 , H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$$

אם דוחים את H_0 יש להמשיך לבדוק באמצעות מבחני t :

$$. H_0 : \beta_j = 0 , H_0 : \alpha_j = 0$$

משתני דמי עברו שני משתנים איקוטיים:

לדוגמא – שני משתנים איקוטיים המשפיעים על פונקציית השכר:
מגדר (אישה, גבר) וגובה (לבן, שגור).

נגידר משתנה דמי G שיקבל 1 אם מדובר בגבר ו-0 אחרת (אישה).

נגידר משתנה דמי R שיקבל 1 אם מדובר לבן ו-0 אחרת (שגור).

נבודוק כיצד מגדר וגובה משפיעים על השכר ההתחלתי (החותך), כאשר השכר תלוי גם בשנות לימוד (S_t).

1. הבדל בחותך ללא אינטראקציה:

$$\text{המודל: } W_t = \alpha_0 + \alpha_1 G + \alpha_2 R + \beta \cdot S_t + u_t$$

במודל זה – אין השפעה משלבת של מגדר וגובה על השכר ההתחלתי.

ניתן לבדוק השערות על כל אחד מהמשתנים הב'ית האיקוטיים בנפרד:

$$1. \text{ הבדל בשכר ההתחלתי בין גברים לנשים: } H_0 : \alpha_1 = 0$$

$$2. \text{ הבדל בשכר ההתחלתי בין שגורים לבנים: } H_0 : \alpha_2 = 0$$

2. הבדל בחותך עם אינטראקציה :

$$\text{המודל : } W_t = \alpha_0 + \alpha_1 G + \alpha_2 R + \alpha_3 G \cdot R + \beta \cdot S_t + u_t$$

במודל זה הטענה היא כי קיימת, בנוסף להשפעה של מגדר וגזע בנפרד על השכר, גם השפעה מושלבת (אינטראקציה) של מגדר וגזע על השכר ההתחלתי.

במודל זה, לעומת זאת, נוספת ההשערה לבדיקת השפעת האינטראקציה בין מגדר לגזע על השכר ההתחלתי :

$$\text{. } H_0 : \alpha_3 = 0 \quad .3$$

3. דרך נוספת לייצירת מודל עם אינטראקציה :

הגדרת משתני דמי המיצגים שילוב בין המשתנים האיכוטיים גזע ומגדר באופן הבא :

D_1 יקבל 1 אם מדובר בגבר לבן ו-0 אחרת.

D_2 יקבל 1 אם מדובר בגבר שחור ו-0 אחרת.

D_3 יקבל 1 אם מדובר באשה לבנה ו-0 אחרת.

הנשים השחורות מהוות כאן את קבוצת הייחוס.

$$\text{המודל : } W_t = \gamma_0 + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 D_3 + \delta \cdot S_t + u_t$$

עזר בטבלה כדין לנ Sach את ההשערות לבדיקת האינטראקציה :

הפרש	אישה	גבר	
$\gamma_1 - \gamma_3$	$\gamma_0 + \gamma_3$	$\gamma_0 + \gamma_1$	לבן
γ_2	γ_0	$\gamma_0 + \gamma_2$	שחור
	γ_3	$\gamma_1 - \gamma_2$	הפרש

ההשערות לבדיקת קיום האינטראקציה : $H_0 : \gamma_1 - \gamma_3 = \gamma_2$ או $H_0 : \gamma_1 - \gamma_2 = \gamma_3$

התוצאות שיתקבלו כאן יהיו כמפורט להלן להתוצאות שהתקבלו בדרך

$$\text{הקודמת : } \begin{aligned} WALD &= t^2 \\ PF &= Pt \end{aligned}$$

שאלות:**משתנה דמי לחותך:**

1) על בסיס מדגם של 50 איש העובדים בחברה מסוימת התקבלו התוצאות הבאות :

$$W_t = 5500 + 1043 \cdot D + 119 \cdot S_t$$

(S.E) (24) (56) (134)

המספרים בסוגרים הם טיעיות התקן של מבחני המובהקות לפרמטרים.

א. מהו השכר התחלתי של גבר בעל 12 שנות לימוד?

ב. מה ההבדל בשכר התחלתי בין גברים לנשים?

ג. האם הבדל זה מובהק באוכלוסייה?

ד. בדקו את הטענה כי השכר התחלתי של גברים גבוהה ביותר מ-500 ש"מ
מזה של נשים.

ה. בדקו את הטענה שהשכר התחלתי של נשים נמוך ב-600 ש"מ מזה של גברים.

פונקציית רגרסיה המכילה משתנים איכוטיים בלבד:

2) על אותו המדגם של 50 איש העובדים בחברה מסוימת ביקש החוקר לבדוק

אם יש הבדל בשכר הממוצע בין גברים לנשים.

$$תוצאות האמידה: D \cdot W_t = 5200 + 1120 \cdot S_t$$

$$\text{נתון: } S_{\hat{\alpha}_1} = 63$$

בדקו האם קיים הבדל מובהק בשכר הממוצע בין נשים וגברים?

משתנה דמי לשיפוע:

3) על בסיס אותו מדגם, ביקש החוקר לדעת האם קיים הבדל מובהק בין גברים

לנשים בתוספת לשכר בגין שנות הלימוד.

תוצאות האמידה נתונות להלן :

$$W_t = 5000 + 110 \cdot S_t + 120 \cdot D \cdot S_t + u_t$$

(68) (23) (25)

בדוק את ההשערה.

משתנה דמי לכל פונקציה:

4) חוקר רצה לבדוק את הטענה שסוג הכביש משפיע על מס' תאונות הדרכים בקטעי כביש ביןעירוניים, בהינתן נפח התנועה. החוקר בדק האם הפונקציה של מס' התאונות בהינתן נפח התנועה, שונה בין כבישים מהירים לבין כבישים שאינם מהירים. לשם כך אמד החוקר את ארבע המשוואות הבאות:

$$1. \text{ כבישים מהירים בלבד. } NUM_t = \gamma_1 + \delta_1 \cdot AVG D_t + \varepsilon_{1t}$$

$$2. \text{ כבישים לא מהירים בלבד. } NUM_t = \gamma_2 + \delta_2 \cdot AVG D_t + \varepsilon_{2t}$$

$$3. \text{ שני סוגי הכביש (כל המדגמים). } NUM_t = \gamma_3 + \delta_3 \cdot AVG D_t + \varepsilon_{3t}$$

$$4. \text{ כאשר: } NUM_t = \alpha + \beta_1 \cdot TYPE_t + \beta_2 \cdot AVG D_t + \beta_3 \cdot (AVGD \cdot TYPE)_t + U_t$$

NUM_t - מס' תאונות הדרכים הקטלניות בקטע כביש t בשנה.

$AVGD_t$ - נפח התנועה בקטע כביש t ליום באלפים.

$TYPE_t$ - משתנה דמי המקבל את הערך 1 כאשר הכביש מהיר, ו-0 כאשר הכביש לא מהיר.

תוצאות אמידת המשוואות מופיעות בהמשך השאלה.

א. בדקו את טענת החוקר בשתי דרכים שונים. ציינו איזה מן המשוואות רלוונטיות עבור כל דרך.

ב. חשבו את הערכיהם המספריים עבור אומדי משווה (4).

ג. מהו האומדן הנקודתי למס' התאונות בכביש מהיר כאשר נפח התנועה עומד על ארבעת מכוניות ליום בקטע הכביש האמור?

הoulתת הטענה כי המקדם להשפעה של נפח התנועה בדרכים מהירות הינו כפול מזה שבדרכים לא- מהירות.

ד. מהי השערת האפס לבדיקת הטענה (במונחי משווה (4))?

ה. מהי הרגרסיה "תחת H_0 " ל מבחן WALD?

משוואת (1) - כבישים מהיריים בלבד:

The REG Procedure

Model: MODEL1

Dependent Variable: num num

Number of Observations Read 344

Number of Observations Used 344

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Pr > F
Model	1	4700.81174	4700.81174	89.12	<.0001
Error	342	18039	52.74684		
Corrected Total	343	22740			

Root MSE	7.26270	R-Square	0.2067
Dependent Mean	5.10465	Adj R-Sq	0.2044
Coeff Var	142.27617		

Parameter Estimates

Variable	DF	Parameter		Standard	
		Estimate	Error	t Value	Pr > t
Intercept	1	1.55289	0.54303	2.86	0.0045
avgd	1	0.02098	0.00222	9.44	<.0001

משוואת (2) - כבישים לא מהירים בלבד:

The REG Procedure

Model: MODEL1

Dependent Variable: num num

Number of Observations Read 410

Number of Observations Used 410

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Pr > F
Model	1	971.99073	971.99073	145.83	<.0001
Error	408	2719.34830	6.66507		
Corrected Total	409	3691.33902			
		Root MSE		2.58168	R-Square 0.2633
		Dependent Mean		1.38780	Adj R-Sq 0.2615
		Coeff Var		186.02612	

Parameter Estimates

Variable	DF	Parameter		Standard	
		Estimate	Error	t Value	Pr > t
Intercept	1	0.14978	0.16360	0.92	0.3605
avgd	1	0.02877	0.00238	12.08	<.0001

משוואה (3) - שני סוגי הכਬיש (כל המינים):

The REG Procedure

Model: MODEL1

Dependent Variable: num num

Number of Observations Read 754

Number of Observations Used 754

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Pr > F
Model	1	8052.00804	8052.00804	288.84	<.0001
Error	752	20964	27.87730		
Corrected Total	753	29016			

Root MSE	5.27990	R-Square	0.2775
Dependent Mean	3.08355	Adj R-Sq	0.2765
Coeff Var	171.22758		

Parameter Estimates

Variable	DF	Parameter		Standard	
		Estimate	Error	t Value	Pr > t
Intercept	1	0.73903	0.23665	3.12	0.0019
avgd	1	0.02330	0.00137	17.00	<.0001

משוואות :(4)

The REG Procedure

Model: MODEL1

Dependent Variable: num num

Number of Observations Read 754

Number of Observations Used 754

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Pr > F
Model	3	8256.966	2752.322	99.44	<.0001
Error	750	20759	27.678		
Corrected Total	753	29016			

Root MSE	5.26102	R-Square	0.2846
Dependent Mean	3.08355	Adj R-Sq	0.2817
Coeff Var	170.61553		

Parameter Estimates

Variable	DF	Parameter		Standard	
		Estimate	Error	t Value	Pr > t
Intercept	1	0.14978	0.33340	0.45	0.6534
type	1				0.0067
avgd	1				<.0001
avgdtype	1				0.1283

משתנה איכוטי עם יותר משתני קטגוריות:

(5) ענה על הסעיפים הבאים :

- א. הoulטה הטענה כי יש הבדל במחיר ההתחלתי בין האביב לקיץ.
- מהי השערת האפס לבדיקת הטענה?
 - פרטו שני מבחנים סטטיסטיים בעזרתם ניתן לבדוק את הטענה.
- ב. הoulטה הטענה כי יש רק שתי עונות המשפיעות על מחיר הירקות ההתחלתי : קיץ + אביב, חורף + סתיו.
- מהי השערת האפס לבדיקת הטענה?
 - מהו המבחן הסטטיסטי המתאים? פרטו.

משתנה דמי עבור שני משתנים איכוטיים:(6) חוקר בדק השפעות של השכלה, גזע (שחור, לבן) וניסיון (EXP) על לוג השכר ($\ln(Y)$) במדגם בן 306 תצפיות :

$$\ln(Y)_t = \alpha_0 + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \beta_1 EXP_t + \beta_2 EXP_t^2 + u_t$$

$\ln(Y)$ - לוג השכר.
 EXP - שנות ניסיון.

- מקבל את הערך 1 עבור שחורים בעלי השכלה גבוהה (ו-0 אחרת).
 D_1
- מקבל את הערך 1 עבור שחורים בעלי השכלה נמוכה (ו-0 אחרת).
 D_2
- מקבל את הערך 1 עבור לבנים בעלי השכלה גבוהה (ו-0 אחרת).
 D_3
- תוצאות אמידת משוואות הרגסיבית מוצגות בפלט להלן :

Analysis of Variance

Source	DF	Sum of		Mean	
		Squares	Square	F Value	Pr > F
Model	5	-----	-----	-----	-----
Error	300	140	-----		
Corrected Total	305	210			

Root MSE	-----	R-Square	-----
Dependent Mean	-----	Adj R-Sq	-----
Coeff Var	-----		

Parameter Estimates						
		Parameter	Standard			
Variable	DF	Estimate	Error	t Value	Pr > t	
Intercept	1	-----	-----	60.84	0.00	
D1	1	-----	-----	-3.20	0.00	
D2	1	-----	-----	-5.56	0.00	
D3	1	-----	-----	7.23	0.00	
EXP	1	-----	-----	8.11	0.00	
EXP ²	1	-----	-----	-7.45	0.00	

- א. לפי המשוואת הניסיון זהה עבור שחורים ולבנים : נכון/לא נכון / לא ניתן לדעת.
- ב. בדוק את הטענה כי בקרב אנשים בעלי השכלה נמוכה אין השפעה לגזע.
- ג. בדוק את הטענה כי אין השפעות השכלה בקרב לבנים.
- ד. מהי השערת האפס לבדיקת הטענה כי אין אינטראקציה בין גזע לשכלה?
- ה. לבדיקת ההשערה של הסעיף הקודם בוצע מבחן WALD.

הרגסיסיה המוגבלת תחת השערת האפס הינה :

$$Z_0 = \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + \gamma_3 Z_3 + \gamma_4 Z_4 + \varepsilon$$

מהם ה-Z'ים ?

- ו. בדוק את ההשערה אם ידוע שבמודל המוגבל $R^2 = 0.33$.
- ז. החוקר החליט לאמוד במקומות את המשוואת המקורית את המשוואת :
- $$\ln(Y)_t = \lambda_0 + \lambda_1 S + \lambda_2 E + \lambda_3 (S \cdot E) + \delta_1 EXP + \delta_2 EXP^2 + \omega_t$$

כאשר :

- S מקבל את הערך 1 עבור שחורים ו-0 אחריו (לבנים).
- E מקבל את הערך 1 עבור השכלה גבוהה ו-0 אחריו (השכלה נמוכה).
- מה הקשר בין המקדים של שני המודלים ?
- ח. אם יאמוד החוקר את המשוואת :

$$\ln(Y)_t = \lambda_0 + \lambda_1 S + \lambda_2 E + \delta_1 EXP + \delta_2 EXP^2 + \omega_t$$

спциficzija של השמota משטנה רלוונטי (היעזר בסעיפים ד', ו' ו-ז').

7) חוקרת בדקה השפעות השכלה, מגדר וניסיון על הכנסה מעבודה לפי המשוואה הבאה :

$$\ln(MWAGE) = \alpha_0 + \alpha_1 \cdot S + \alpha_2 \cdot E + \alpha_3 \cdot (S \cdot E) + \beta_0 \cdot EXP + \beta_1 \cdot (EXP \cdot S) + \beta_2 \cdot (EXP \cdot S \cdot E) + U$$

כאשר :

S משתנה דמי : 1 = עבור נשים, 0 = גברים.

E משתנה דמי : 1 = עbor השכלה גבוהה ($scl > 12$), 0 = השכלה נמוכה.

א. רשמו את הפונקציה לחישוב :

- i. תחזית לוג השכר עבור גבר בעל השכלה נמוכה ו- 10 שנים ניסיון.
- ii. תחזית לוג השכר ההתחלתי עבור נשים משכילות.
- iii. לאחר כמה שנים ניסיון ישתווה השכר של נשים משכילות לזה של גברים משכילים?

ב. רשמו את השערות האפס המתאימות לבדיקת הטענות הבאות :

- i. אין השפעה של מגדר והשכלה על השכר.
- ii. השפעת ההשכלה אינה תלולה במגדר.
- iii. אין השפעות השכלה אצל גברים.
- iv. אין הבדל בשיעורי התשואה לניסיון, בקרב הנשים.

תשובות סופיות:

(1) א. $W_t = 7971$
 ב. יש עדות לכך.
 ג. כן.
 ד. יש עדות לכך.

(2) יש עדות לכך.

(3) יש עדות לכך.

(4) (4) א. יש עדות לכך, מבחן CHOW 1-3, 2-1, 3-1, מבחן DWLS 3-4.

$$\hat{\alpha} = 0.14978, \hat{\beta}_1 = 1.40311, \hat{\beta}_2 = 0.002877, \hat{\beta}_3 = -0.008$$

$$\begin{aligned} H_0: \beta_2 + \beta_3 &= 2 \cdot \beta_2 \\ \text{. } H_0: \beta_3 &= \beta_2 \end{aligned}$$

$$\text{. } NUM_t = \alpha + \beta_1 \cdot TYPE_t + \beta_3 \cdot (AVGD_t + AVGD \cdot TYPE)_t + U_t$$

(5) א. $H_0: \alpha_1 = \alpha_2, \alpha_3 = 0$ ב. t-WALD .ii
 ג. $H_0: \alpha_1 = \alpha_2$.i. נ. WALD .ii
 ד. WALD .ii

(6) א. נכון. ב. יש עדות לכך. ג. יש עדות לכך.

$$\text{. } H_0: \alpha_3 = \alpha_1 - \alpha_2 \text{ או } H_0: \alpha_2 = \alpha_1 - \alpha_3$$

$$Z_0 = \ln(Y)_t, Z_1 = D_1 + D_3, Z_2 = D_2 - D_3, Z_3 = EXP_t, Z_4 = EXP^2_t$$

ו. אין עדות לכך. ז. לא.
 ח. לא.

$$\text{. } \hat{ln}(MWAGE) = \hat{\alpha}_0 + \hat{\beta}_0 \cdot 10$$

$$\text{. } EXP_t = \frac{-(\alpha_1 + \alpha_3)}{\beta_1 + \beta_3}$$

$$\text{. } \hat{ln}(MWAGE) = \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3$$

ב. $H_0: \alpha_3 = \beta_3 = 0$.ii
 ג. $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \beta_1 = \beta_2 = \beta_3 = 0$.i.

ד. $H_0: \beta_2 + \beta_3 = 0$.iv
 א. $H_0: \alpha_2 = \beta_2 = 0$.iii

אקוונומטריקה

פרק 12 - תיאוריה הפרת ההנחהות קלאסיות

תוכן העניינים

1. כללי

- 73

הפרת הנחות הקלאסיות:

רקע:

ארבעת הנושאים הבאים עוסקים במצב של הפרת אחת הנחות הקלאסיות הדרושות לאמידת הפרמטרים בשיטת OLS :

- הטווסקדיוטיות (הפרת הנחה מס' 5) – שונות קבועה ויחידה לאורך קו הרגרסיה : $\sigma^2 = V(u_t)$.
- מתאם סידרתי (הפרת הנחה מס' 6) – אי תלות בין הטעויות : $\text{cov}(u_t, u_s) = 0$.
- מודלים דינמיים ו-משוואות סימולטניות (הפרת הנחה מס' 4) – אי תלות בין המשתנים הבית' לטעויות : $\text{cov}(x, u) = 0$.

בכל אחד מן הנושאים נלמד :

- מהן ההשלכות של הפרת הנחות הללו על אומדי הריבועים הפחותים.
- מהם המבחנים הסטטיסטיים המשמשים לזיהוי קיומה של ההפרה.
- כיצד נתקן את משווהת הרגרסיה כך שניתן יהיה לאמוד את הפרמטרים בשיטת OLS.

אקוונומטריקה

פרק 13 - הטרוסקסטיות

תוכן העניינים

1. כללי

74

הטרוסקדייטיות:

רקע:

הטרוסקדייטיות הוא מצב שבו מופרת הנחת ההומוסקדייטיות, הגורסת כי שונות הטועיות היא אותה שנותן עבור כל תצפית ותצפית: $\sigma^2 = \text{Var}(u_t)$, כלומר התצפיות מפוזרות באופן אחיד סביר כו הרגסית. במצב של הטרוסקדייטיות שנותן הטועיות של כל תצפית היא שונה: $\text{Var}(u_t) = \sigma^2$.

ההשלכות של הטרוסקדייטיות על אומדי OLS:

בחינתן הטרוסקדייטיות מופרת תכונת הייעילות של אומדי הריבועים הפחותים שכן בכך לחשב שנותן יעלה של האומדים השתמשנו בהנחה של שנותן קבועה.

מבחנים לזיהוי הטרוסקדייטיות:

החיש לקיומה של בעיית הטרוסקדייטיות בתנאים צריך להתעורר כאשר אנו בוחנים את גרפ' השאריות – באיזה אופן השונות של הטועיות משתנה בין תצפית לתצפית. שיטות לזיהוי הטרוסקדייטיות: מבחן GQ (Goldfeld-Quandt) ובבחן White. מבחן GQ מניח כי במקום שנותן אחת אחת של הטועיות לכל התצפיות, קיימות שתי שנותן שונות בלבד. ואילו מבחן White מניח כי לכל תצפית ותצפית שונות של טועיות.

1. מבחן GQ :

הנחה העומדת בבסיס מבחן זה היא כי קיימות שתי שנותן של טועיות.

ביצוע המבחן :

- מחלקים את המדגם לשני חלקים:

1. החלק שבו אנו חושדים שיש שנותן גבוהה יותר.

2. החלק שבו אנו חושדים שיש שנותן נמוכה יותר.

מקובל להשRITE מס' תצפיות (בין 1/6 ל-3/1) במרכז המדגם.

- אומדים כל אחד מחלוקת נפרד ומתקבלים את ה-ESS של כל חלק.

- מחשבים את הסטטיסטי: $F_{stat} = \frac{\text{ESS}_1/T_1 - K - 1}{\text{ESS}_2/T_2 - K - 1}$ (תמיד השנותן הגבוהה חלקו הקטנה).

- סטטיסטי זה מתפלג: $F_{(\alpha; T_1-K-1, T_2-K-1)}$.
- כלל ההכרעה: אם $F_{stat} > F_C$ או דוחים את H_0 .

$$\begin{aligned} H_0: \sigma_1^2 &= \sigma_2^2 \\ H_1: \sigma_1^2 &> \sigma_2^2 \end{aligned}$$

2. מבן White: הנחה העומדת בבסיס מבחון זה כי לכל תצפית ותצפית שוונות של טעויות. הביטוי המתמטי של הנחה זו היא היotta של השוונות פונקציה ליניארית של כל המשתנים המסבירים, ריבועיהם והאיברים הצלבים:

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

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

האומד ל- σ_t^2 הוא $\hat{\sigma}_t^2$.
הבחן הוא מבחן LM:

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

פתרונות בעיית ההטרוסקסדסטיטוות – ריבועיםՓחות משוקללים (WLS):

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

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

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

$$\begin{aligned} Y_t &= \alpha + \beta X_t + u_t \\ \cdot Y_t W_t &= \alpha \cdot W_t + \beta (X_t W_t) + u_t W_t \end{aligned}$$

המשוואת המקורית:

בצורתה המפורשת המשוואת החדשה נראה כך:

$$\cdot \frac{Y_t}{\sqrt{Z_t}} = \alpha \cdot \frac{1}{\sqrt{Z_t}} + \beta \cdot \frac{X_t}{\sqrt{Z_t}} + \frac{u_t}{\sqrt{Z_t}}$$

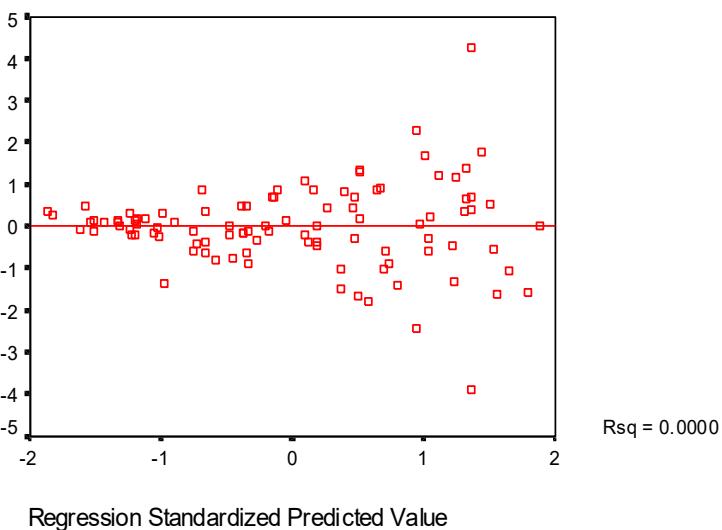
שאלות:

מבחן GQ:

- 1) נאמד הקשר שבין הכנסה לתשוכת: $. Y_t = \alpha + \beta X_t + u_t$
גרף השאריות של הרוגרסיה הנ"ל נתון להלן:

Scatterplot

Dependent Variable: Y



מגרף זה אנו למדים כי השונות איננה אחידה סביר קו הרוגרסיה אלא תלויות ברמת הכנסה – זהו מצב של הטורוסקסטטיות.
בכדי לבצע מבחן GQ:

- התוצאות של משתנה הכנסה סודרו מהגדול לפחות והມדגם חולק לשלוש קבוצות שוות.
- רוגרסיה נפרדת הוראה על השלישי הראשון ועל השלישי האחרון.

התוצאות של אמידת הקשר בין הכנסה לצרכים מוצג בפלטים 1 ו-2 בהתאם:

משואה (1)

The REG Procedure

Model: MODEL1

Dependent Variable: y

Number of Observations Read	16
Number of Observations Used	16

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model					
Error	14	166452.9			
Corrected Total					

משואה (2)

The REG Procedure

Model: MODEL1

Dependent Variable: y

Number of Observations Read	16
Number of Observations Used	16

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model					
Error	14	2934638			
Corrected Total					

האם יש עדות לקיום הטרוסקסדסטיות נתונים? (בצעו את המבחן המתאים: רשמו השערות, חשבו סטטיסטי מבחן, רשמו כלל הכרעה והגיעו למסקנה).

על מנת לבדוק את פונקציית הייצור בענף מסויים נאספו נתונים על 150 FIRMOOT. נסמן:

Q - תפוקה שנתית באלפי שקלים.

L - מספר עובדים.

המודל הנאמד: $L = \alpha + \beta \cdot \ln(Q)$.

חוקר חש שההפרעה המקראית אינה הומוסקסדסטית. לשם כך הוא מין את התצפויות בסדר עולה של מספר העובדים, השמיט $\frac{1}{3}$ מהתצפויות האמצעיות

והרייך שתי וגרסאות נפרדות עם מספר שווה של תצפויות:

ברגרסיה הכוללת את הערכים הנומוכים יחסית של תשומת העבודה הוא

קיבל : $R^2 = 0.403$, $ESS = 279.3$

ברגרסיה הכללת את הערכאים הגבוהים יחסית של תשומת העובדה הוא

קיבל : $R^2 = 0.238$, $ESS = 493.8$

האם יש עדות לקיום הטרוסקדיוטיות בנתונים? (בצעו את המבחן המתאים : רשמו השערות, חשבו סטטיסטי מבחן, רשמו כלל הכרעה והגיעו למסקנה).

מבחן WHITE:

(3) על אותו הקשר שבען הכנסה לתצרוכת : $Y_t = \alpha + \beta X_t + u_t$ שבו מצב של הטרוסקדיוטיות. שאנו חושדים על פי גרפ השאריות כי קיימים בו מצב של הטרוסקדיוטיות. בצד י写下 בוצע את מבחן WHITE :

- נחשב את השאריות של הרגרסיה : $\hat{u}_t = Y_t - \hat{Y}_t$
- נעלם את השאריות בריבוע : \hat{u}_t^2
- נאמוד את המשווהה : $\hat{u}_t^2 = \alpha_0 + \alpha_1 X_t + \beta_1 X_t^2 + v_t$

תוצאות האמידה מוצגות להלן :

```
The REG Procedure
Model: MODEL1
Dependent Variable: RES2
Number of Observations Read      48
Number of Observations Used      48

Analysis of Variance
Source        DF     Sum of Squares      Mean Square      F Value      Pr > F
Model
Error
Corrected Total

Root MSE      Dependent Mean      R-Square      0.390763
                          Adj R-Sq
```

בוצעו מבחן White (השערות, סטטיסטי המבחן, כלל הכרעה ומסקנה).

(4) חוקר מנהח כי מכירות של חנותו הנו פונקציה של שיטתה, דמי שכירות והאפשרות של מכירת עיתונים.

נסמן :

- מכירות חודשיות (לפ). Y_SALES

- שטח החנות (מ"ר). $X1_SQUARES$

- דמי שכירות (\$). $X2_RENT$

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

החוקר חשב כי קיימת בעיה של הטראנסדסטיות בנתונים.

החוקר ביצע מבחן ליזיהו הטראנסדסטיות שתוצאותיו נתונות להלן :

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

- . הרגסיה המופיעה בפלט לעיל נועדה לבדיקת :
 על ידי מבחן : _____ .
 המשתנה התלו依 הינו : _____ .
 המשתנים הב"ית : _____ .
 ההשערות הינו : _____ .
 גודל הסטטיסטי למבחן הינו (רשמו תוצאה מספרית) : _____ .
 המסקנה המתתקבלת היא : _____ .

שיטת WLS:**5)** נתון המודל :

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

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

- a. מהי הביעה שנוצרת באמידת משווהה (1)?
 b. מהן תכונות אומדי הריבועים הפחותים של משווהה (1)?

כדי לפטור את הביעה שנוצרה, נאמדת המשווהה הבאה :

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

ג. מהו W שבזורתו ניתן לאמוד את α ו- β בצורה עיליה?ד. מהו האומד הייעיל של σ^2 ?ה. האם ניתן להשווות בין המודלים על בסיס R^2 ? אם לא, האם ניתן להחליט בכל זאת איזה מודל טוב יותר?

ו. חווו דעתכם על הטענות הבאות, ונמקו :

i. אם נתון כי : $Z_t = a + b \cdot \bar{X}$, התחשבות לסעיפים א' ו-ב' נשארות ללא שינוי.ii. המשווהה הנורמללית : $0 = \sum \hat{\varepsilon}_t$ (כאשר : $\varepsilon_t = U_t \cdot W_t$) היא אחת המשווהות הנורמלליות לאמידת משווהה (2).**6)** ענו על השאלה הקודמת, כאשר נתון כי : $. VAR(U_t) = \sigma^2 \cdot X_t^2$ **7)** נתון המודל : $Y_t = \alpha + \beta X_t + u_t$. וקיים מוגם של 100 תצפיות כאשר נתוןci : $V(u_t) = \sigma_t^2 = \begin{cases} X_t \sigma^2 & \Leftrightarrow t \geq 50 \\ X_t^2 \sigma^2 & \Leftrightarrow t \leq 50 \end{cases}$

א. במשווהה מס' 1 יש בעיה של _____.

ב. אמידת משווהה (1) תניב אומדים

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

ג. פתרו את הביעה הקיימת במשווהה (1) ייתכן על ידי אמידת המשווהה

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

כאשר : $W_t = _____$ 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

תשובות סופיות:

.2.48) יש עדות לכך, השערות:
 $H_0: \sigma_1^2 = \sigma_2^2$, חישוב סטטיסטי: 17.62, כלל הכרעה: 8.

.1.69) יש עדות לכך, השערות:
 $H_0: \sigma_1^2 = \sigma_2^2$, חישוב סטטיסטי: 1.77, כלל הכרעה: 1.

.3) יש עדות לכך, השערות:
 $H_0: \alpha_1 = \beta_1 = 0$, חישוב סטטיסטי: 18.75
 $H_1: \text{OTHERWISE}$

כלל הכרעה: 5.991
 קיום הטרנסקסטיות בנתונים.
 .(LM)WHITE

.RES²

X1_SQUARE , X1_SQUARE² , X1_SQUARE * X2_RENT ,

. X2_RENT , X2_RENT² , X2_RENT * PAPERS , PAPERS

. LM_{stat} = 1.73

אין עדות לכך.

. W_t = $\frac{1}{\sqrt{Z_t^2}} = Z_t$ ג. ב. ראו סרטון.
 א. הטרנסקסטיות. 5

ה. לא, המודל השני. ו. לא נכון. ii. לא נכון.
 . $\sigma^2 = \frac{ESS}{T-K}$ ד.

. W_t = $\frac{1}{X_t}$ ג. ב. ראו סרטון.
 א. הטרנסקסטיות. 6

ה. ראו סרטון. ג. ראו סרטון.
 . S² = $\frac{ESS}{T-k-1}$ ד.

ב. נכון. א. הטרנסקסטיות. 7

. W_t = $\frac{1}{\sqrt{X_t^2}}$ t ≤ 50 , W_t = $\frac{1}{\sqrt{X_t}}$ t ≥ 50
 ד. לא. ג.

אקוНОמטריקה

פרק 14 - מתאם סדרתי

תוכן העניינים

- | | |
|----------|---------------|
| 82 | 1. כללי |
|----------|---------------|

מתאים סדרתי:

רקע:

מתאים סדרתי עוסקת במצב שבו מופרת ההנחה (מש' 6) של אי תלות בין הטיעוות: $\text{cov}(u_s, u_t) = 0$ וኖצרת תלות סטטיסטית בין הטיעוות במודל: $0 \neq \text{cov}(u_s, u_t)$. תלות כזו בין הטיעוות קיימת בכך כלל כאשר הנתונים הנאספים הם נתוני סדרות עיתיות ולא נתוני חתך בהם עסוקנו עד כה. בנתוני סדרות עיתיות, לאחר ומדובר באותו הפרט הנמדד בזמןנים שונים סביר שהטיעוות בניובי שלו תהיה תלויות אחת בשנית.

השלכות על אומדי הריבועים הפחותים (OLS):

מבין התכונות של אר"פ (lieniarיות, חוסר הטיה, עקיבות ויעילות) היחידה שמופרת כאשר קיים מתאם סידרתי היא: **תכונת הייעילות**. משום שתכונת הייעילות היא היחידה מבין תכונות אר"פ התלויה להוכחתה בקיומה של הנחת אי תלות בין הטיעוות. משום הפגיעה בתכונת הייעילות, בדיקת ההשערות לא תהיה תקפה.

- שימוש לב: כי במידה וקיים מתאם סידרתי חיובי בין הטיעוות ולמשתנים יש מגמת זמן (X עולה או יורדת עם הזמן) אומד השונות (ESS) יהיה מוטה כלפי מטה ואז נקבל: R^2 , F ו- t מוטים כלפי מעלה.

מבנה המתאים הסדרתי:

מתאים סדרתי מסדר ראשון:
 ההנחה היא כי יש מתאם בין הטיעוות למרחק אחד, כלומר u_t תלוי ישירות רק ב- u_{t-1} : $\text{cov}(u_t, u_{t-1}) \neq 0$.

את המתאים בין הטיעוות מסדר ראשון ניתן לנתח באופן הבא: $u_t = \rho \cdot u_{t-1} + \varepsilon_t$ כך ש:

1. $0 \neq \rho$ (כי אם $0 = \rho$ אין מתאים סידרתי).
2. $-1 < \rho < 1$ (כי אם חורג מ-1 הטעות הולכת ונגדלה עם הזמן).
3. ρ חיובי פירושו מתאים סדרתי חיובי ואילו ρ שלילי פירושו מתאים סדרתי שלילי (לא נפוץ).

ε_t מקיים את ההנחות הקלasicות מאחר ומהויה סטייה מקרית לחלוטין .4.

$$E(\varepsilon_t) = 0$$

$$\text{. } V(\varepsilon_t) = \sigma_\varepsilon^2 \quad : \text{ (בניגוד ל- } u_t \text{) כך ש:}$$

$$\text{cov}(\varepsilon_t, \varepsilon_{t-s}) = 0$$

המודל יכולشتwei משוואות-המשווהה העיקרית והגדרת המתאים הסדרתי (mse) :

$$\begin{aligned} Y_t &= \alpha + \beta X_t + u_t \\ \text{ריאשוון: } u_t &= \rho \cdot u_{t-1} + \varepsilon_t \end{aligned}$$

מלבד α ו- β נרצה לאמוד גם את ρ .

מתאים סדרתי מסדר שני :

$$u_t = \rho_1 \cdot u_{t-1} + \rho_2 \cdot u_{t-2} + \varepsilon_t$$

הן u_{t-1} והן u_{t-2} משפיעים ישירות על u_t .

מתאים סדרתי מסדר P :

$$u_t = \rho_1 \cdot u_{t-1} + \rho_2 \cdot u_{t-2} + \dots + \rho_P \cdot u_{t-P} + \varepsilon_t$$

וושפע מתקופות שונות בעבר .

תכונות המתאים הסדרתי:

מכיוון שככל טעות בזמן מסויים מתואמת עם הטעות הסמוכה לה בזמן :

המתאים של u_t הולך ופוחת עם הזמן :

בנוסף לכך, התוחלת, השונות והשונות המשותפת של הטעויות :

$$E(u_t) = 0$$

$$V(u_t) = \sigma_u^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$COV(u_t, u_{t-s}) = \rho^s \sigma_{u_t}^2 = \rho^s \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

מבחנים לזיהוי מתאם סדרתי:

מבחן DW (דרבין ווטסון) לקיים מתאם סדרתי מסדר ראשון :

- נניח תחילת כי אין מתאם סדרתי ונאמוד את המשווה הראשית בשיטת OLS.
- חלק מתוצאות האמידה נקבל ציון DW (יכול לקבל ערכים בין 0 ל-4 בלבד).
- נתבונן בטבלה DW ולפי K = מס' המשתנים הב"ת במודל ו-T = מס' התצפיות בדוגמנו נשלוף שני ערכים : d_L ו- d_U .
- חלק את הטווח שבין 0 ל-4 באופן הבא :

$$0 - \rho > 0 - d_L - d_U - \rho = 0 - 4 - d_U - 4 - d_L - \rho < 0 - 4$$

- נראה היכן נופל ציון ה-DW שהתקבל לחלק מתוצאות האמידה.
- ניתן לדעת אם יש מתאם ואיזה סוג של מתאם רק אם ציון ה-DW ייפול בחלוקת המודגשים באדום.

$$\begin{aligned} \text{השערות: } H_0: \rho = 0 \\ H_1: \rho > 0, \rho < 0 \end{aligned}$$

חישוב הסטטיסטי : $DW_{stat} \cong 2 \cdot (1 - \hat{\rho})$
 אם אנו מקבלים ציון $\hat{\rho}$ ניתן להציב בנוסחה ולקבל DW_{stat} .

מבחן DW יש שתי בעיות עיקריות :

1. מתאים רק למתאם סדרתי מסדר ראשון.
2. יש אזורים "מתאים" בטווח בהם לא ניתן לדעת האם יש מתאם סדרתי.

בנוסף לכך על מספר תנאים להתקיים כדי שאפשר יהיה להשתמש במבחן DW :

1. הרגרסיה כוללת חותך.
2. ה-Xים קבועים ולא משתנים.
3. אין משתנים מסבירים שהם פיגור של המשתנה המוסף.
4. אין תכיפות חסרות באמצע.
5. אם קיימים מתאם סדרתי מסדר ראשון אז הוא מהצורה : $\varepsilon_t = \rho \cdot \varepsilon_{t-1} + u_t$.

מבחן LM :
לעומת מבחן DW מבחן LM מתאים גם לבחינת קיומו של מתאם סדרתי מסדריים

$$\begin{aligned} Y_t &= \alpha + \beta \cdot X_t + u_t \\ u_t &= \rho \cdot u_{t-1} + \varepsilon_t \end{aligned}$$

המודל הלא מוגבל :

$$(U) \quad Y_t = \alpha + \beta \cdot X_t + \rho \cdot u_{t-1} + \varepsilon_t$$

ניתן להתייחס לבחינת קיומו של מתאם סדרתי כהוספת משתנה מסביר : \hat{u}_t .
השלבים לביצוע המבחן :

- נאמדות את המודל המקורי ונחשב \hat{u}_t ו- \hat{u}_{t-1} .
- נאמדות את רגרסיות העזר : $\hat{u}_t / x_1, x_2, \dots, x_k; \hat{u}_{t-1}$.
- נחישב סטטיסטי LM : $LM_{stat} = T \cdot R^2$.
- נדחה את H_0 כאשר $LM_{stat} > \chi_m^2$ כאשר $m =$ סדר המתאם הסדרתי.
אם נדחה את H_0 נדע את סימנו של המתאם הסדרתי לפי המקדם של \hat{u}_{t-1}
ברגרסיות העזר ששויה ל- $\hat{\rho}$.
שימוש לב Ci אם נרצה לבדוק מתאים סדרתיים מסדריים גבוהים יותר :

$$\begin{aligned} H_0: \rho_1 &= \rho_2 = \dots = \rho_s = 0 \\ H_1: &OTHERWISE \end{aligned}$$

$$\hat{u}_t / x_1, x_2, \dots, x_k; \hat{u}_{t-1}, \hat{u}_{t-2}, \dots, \hat{u}_{t-s}$$

פתרונות בעיית המתאם הסדרתי – רגרסיות הפרשיים (שיטת קווקון-אורקט) :

ניצור משווהה שהיא קומבינציה ליניארית של המשווהה המקורית שבאה לא יהיה מתאם סדרתי ולכן ניתן יהיה לאמוד אותה בשיטת הריבועים הפחותים, האומדדים יהיו יעילים ונitin ייה לבצע בדיקת השערות.

משווהה (1) : המודל בזמן t : $Y_t = \alpha + \beta X_t + u_t$

משווהה (2) : המודל בזמן $-t$ מוכפל ב- ρ : $\rho \cdot Y_{t-1} = \rho \cdot \alpha + \rho \cdot \beta X_{t-1} + \rho \cdot u_t$

החסרת משווהה (2) ממשווהה (1) :

$$Y_t - \rho Y_{t-1} = \alpha(1 - \rho) + \beta(X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$$

כדי לאמוד את הפרמטרים של רגרסיות ההפרשיים נגדיר :

$$Y_t^* = Y_t - \rho Y_{t-1}$$

$$\alpha^* = \alpha(1 - \rho)$$

$$\beta^* = \beta$$

$$X_t^* = X_t - \rho X_{t-1}$$

$$\varepsilon_t = u_t - \rho u_{t-1}$$

כך "נטרלנו" את המתאם הסדרתי : $\varepsilon_t = u_t - \rho \cdot u_{t-1}$, $u_t = \rho \cdot u_{t-1} + \varepsilon_t \iff$

מתקנים את כל הנקודות הקלסיות ולכון שוניות הרגסיה וכן שוניות הפרמטרים הנאמדים לא תהיה תלולה במקדם המתאים הסדרתי.
 המשוואה "המתוקנת" אותה נאמד: $Y^* = \alpha^* + \beta^* X_t^* + \varepsilon_t^*$.
 לאחר אמידת משווה זו ניתן לחץ את האומדים של הפרמטרים המקוריים: $\hat{\beta}$, $\hat{\alpha}$.
 מאחר ש- ε אינו ידוע יש צורך לאמוד אותו.

אמידת ε בשיטת קווקן אורקוט:
 שיטת קווקן אורקוט לאמידת ε היא שיטה איטרטיבית – מבוססת על חזרות של תהליך מסוים עד להתקנסות.
 התהליך הממוחשב נקרא אוטורגסיה (AUTOREGRESION) מסדר ראשון, שני, שלישי וכו' (תלוי בסדר המתאים הסדרתי). התקון למתאים הסדרתי יתבצע על ידי הרצת רגסיה עם משתנה AR(1) (אוטו רגסיה מסדר ראשון), AR(2) ו-AR(1) (אם מניחים קיום אוטורגסיה מסדר שני) וכו'. אם משתנה AR מובחק זו אינדיקציה שפתרנו את הבעיה של המודל המקורי.

שאלות:**תכונות המתאים הסדרתי:**

- 1) נתון מתאם סדרתי מסדר ראשון: $u_t = \rho \cdot u_{t-1} + \varepsilon_t$ ונתון כי $\rho = 0.9$ וכי $V(\varepsilon_t) = \sigma^2 = 1$. מצאו את:
 א. המתאים בין u_t ל- u_{t-4} .
 ב. המתאים בין u_t ל- u_{t-4} . הסבר את ההבדל בין המתאיםים (סעיף א' ו-ב').
 ג. השונות σ_u^2 .
 ד. חזרו על סעיפים א' עד ג' בעבר $\rho = 0.4$. הסבירו את ההבדל בין התוצאות.

 מבחון DW:

- 2) חוקר רצה לאמוד את מחיר סגירה של מנתה כפונקציה של הזמן שעובר:
 $CLOSE_t = \alpha + \beta \cdot TIME_t + u_t$
 כאשר:
 $CLOSE_t$ = מחיר סגירה של מנתה ב-\$ ביום t .
 $TIME_t$ = משתנה זמן שמקבל את הערכים: ..., 1, 2, 3, ...
 תוצאות האמידה שהתקבלו:

Dependent variable: CLOSE

Analysis of Variance						
F Value	Prob>F	Mean Square	Sum of Squares	DF	Source	
55.78		-----	-----	1	Model	
		-----	-----	151	Error	
		-----	-----	152	C Total	
0.181		R-square	---- Root MSE			
-----		Adj R-sq	---- Dep Mean			
			---- C.V.			
Parameter Estimates						
T for H0:		Standard	Parameter			
Parameter=0	Prob> T	Error	Estimate	DF	Variable	
91.047	0.0000	0.0148	1.3474	1	INTERCEP	
0.0000	-7.468	0.0001	-0.00075	1	TIME	
Durbin-Watson D	0.150					

האם קיים מתאם סדרתי?

TABLE 12 Cutoff Points for the Distribution of the Durbin-Watson Test Statistic

Let d_α be the number such that $P(d < d_\alpha) = \alpha$, where the random variable d has the distribution of the Durbin-Watson statistic under the null hypothesis of no autocorrelation in the regression errors. For probabilities $\alpha = .05$ and $\alpha = .01$, the tables show, for numbers of independent variables, K , values d_L and d_U such that $d_L \leq d_\alpha \leq d_U$ for numbers n of observations.

n	$\alpha = .05$									
	K									
	1		2		3		4		5	
	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	1.82	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
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
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

3) המודל הינו :

$$\begin{aligned} Y_t &= \alpha + \beta X_t + \varepsilon_t \\ u_t &= \rho u_{t-1} + \varepsilon_t \end{aligned}$$

 u_t - הפרעה מקראית קלאסית.

וידוע כי : $T = 100$

$$u_t = 0.9u_{t-1}$$

$$d_L = 1.57$$

$$d_U = 1.65$$

האם קיימים מתאימים סדרתיים ברמת מובהקות של 5%?

מבחן LM:

4) עבר הדוגמא הקודמת – ניבוי מחיר סגירה של מנתה כפונקציה של הזמן :

$$CLOSE_t = \alpha + \beta \cdot TIME + u_t$$

נבחן את קיומו של מתאם סדרתי מסדר ראשון באמצעות מבחן LM.

$$u_t = \gamma_0 + \gamma_1 \cdot TIME_t + \rho \cdot u_{t-1} + \varepsilon_t$$

נתוצאות האמידה שהתקבלו :

Dependent variable: RES

Analysis of Variance					
F Value	Prob>F	Mean Square	Sum of Squares	DF	Source
		-----	-----	2	Model
		-----	-----	150	Error
		-----	-----	152	C Total

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

Parameter Estimates

Parameter	Estimate	DF	Variable
	-0.00096	1	INTERCEP
	1.16331E-05	1	TIME
	0.927172	1	RES1

- א. האם קיימים מתאימים סדרתיים?
- ב. מהו ערכו של המתאים הסדרתי הנאמד?
- ג. מהו כיוונו של המתאים הסדרתי באוכטולסיה?

תיקון המתאים הסדרתי:

- 5) סטודנט הניח כי במודל: $Y_t = \alpha + \beta X_t + u_t$ קיימים מתאימים סדרתיים מסדר ראשון בשאריות כך שמתקיים: $u_t = 0.7u_{t-1} + \varepsilon_t$ ולכן במקומם לאמוד את המודל המקורי אמד את המודל: $Y_t - 0.7Y_{t-1} = \alpha(1-0.7) + \beta(X_t - 0.7X_{t-1}) + u_t$. הטענה טען כי במודל החדש לא קיימים מתאימים סדרתיים. נוכונה / לא נוכונה / לא ניתן לדעת טענת הסטודנט:
- 6) נשיק עם הדוגמא של ניבוי מחיר סגירה של מנתה כפונקציה של הזמן: $CLOSE_t = \alpha + \beta \cdot TIME + u_t$. נניח כי קיימים מתאימים סדרתיים מסדר ראשון נתונים. תוצאות האמידה בשיטת אוטורגרסיה מסדר ראשון מוצגות להלן:

Parameter Estimates					
Prob> T	T for H0: Parameter=0	Standard Error	Parameter Estimate	DF	Variable
0.000			1.333	1	INTERCEP
Durbin-Watson D	2.235		-0.0006	1	TIME
			0.927	1	(1)AR

- א. בדקו האם נפתרה בעיית המתאים הסדרתי.
 ב. מהי המשועואה לאמידת מחיר הסגירה הצפוי ביום המסחר הבא?

תרגול מסכם:

- 7) נאמד הקשר שבין הכנסתה לתקופת ינואר 1994 עד דצמבר 1997 ($T=48$).
 המודל הינו: $C_t = \alpha + \beta Y_t + u_t$.
 בניסיון לבדוק האם מתקיים קשר מסווג הבא: $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \varepsilon_t$.
 נאמדת המשועואה הבא: $\hat{u}_t = \gamma_1 \hat{u}_{t-1} + \gamma_2 \hat{u}_{t-2} + \gamma_3 \hat{u}_{t-3} + \gamma_4 Y_t + \omega_t$.
 תוצאות האמידה מוצגות להלן:

Depended Variable: RES

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob> T
INTERCEP	1	-54.709	710.85	-0.076	0.939
RESID1	1	0.705	0.152	4.631	0.000
RESID2	1	-0.0066	0.188	-0.035	0.972
RESID3	1	-0.337	0.167	-2.012	0.051
Y	1	0.0027	0.032	0.085	0.932
Durbin-Watson D		1.954			

- נתון בנוסף כי: $R^2 = 0.479$.
- א. הרגרסיה המומפיה בפלט לעיל נועדה לבדיקת:
על ידי מבחן: _____.
- ההשערות הינו: _____.
- גודל הסטטיסטי לבחן הינו (רשמו תוצאה מספרית): _____.
- המסקנה המתקבלת היא: _____.

בהתaha כי קיים מתאם סדרתי מסדר שלישי בנתונים נאמד מחדש הקשר שבין ההכנסה לצורכת בהתאם לשיטתם של קווקן ואורקוט.
תוצאות האמידה מוצגות להלן:

Depended Variable:C

Parameter Estimates						
	Error	:T for H0	Standard Parameter=0	Prob> T Estimate	DF	Variable
	1128.38	-1.864	0.069	-2103.53	1	INTERCEP
	0.0510	14.086	0.000	0.71944	1	Y
AR(1)	1	0.7070	0.1525	4.634	0.000	
AR(2)	1	-0.0064	0.1889	-0.034	0.972	
AR(3)	1	-0.3282	0.1669	-1.966	0.0562	
Durbin-Watson D		1.954				

ב. רשמו את המשווה המתוקנת המשמשת לעריכת תחזיות.

- 8) חוקר רצה לאמוד את עקומת הביקוש לטיסות לאירופה.
לשנותו נתונים שבועיים לאורך 3 שנים (52 שבועות).
נסמן:
 Y_t - מספר כרטיסי הטיסה לאירופה שנמכרו בשבוע t.
 p_t - מחיר ממוצע ב-\$ של הכרטיסים שנמכרו בשבוע t.
 החוקר אמד את המודל: $Y_t = e^{\alpha} \cdot P_t^{\beta_1} \cdot P_{t-1}^{\beta_2}$.
 וקיים לאחר הטרנספורמציה הלוגריתמית: $R^2 = 0.81$.
 לבדיקת ההשערה כי קיים מתאם סדרתי בנתונים מסדר ראשון הוא חישב את
 \hat{u}_t ולאחר מכן חישב את הרגרסיה: $u_t = \gamma_0 + \gamma_1 \ln P_t + \gamma_2 \ln P_{t-1} + \gamma_3 u_{t-1}$.
 מקדם ההסבר המרובה ברגסיה זו הוא 0.282.
 א. נסח את ההשערה ובחנו אותה בר"מ של 0.05.
 החוקר מניח שיש מתאם סדרתי מסדר ראשון.
 לאחר תיקון Cochrane-Orcutt התקבל: $\hat{u}_t = 0.2 + 0.2 \ln P_t + 0.4 \ln P_{t-1}$, $\ln Y_t = 7.3 - 0.2 \ln P_t + 0.4 \ln P_{t-1}$.
 הניחו שהשבוע ושבוע שעבר מחיר ממוצע של כרטיס היה \$500.
 השבוע נמכרו 6,185 כרטיסים. בשבוע הבא צפוי מחיר של \$400.
 ב. כמה כרטיסים יימכרו?
 החוקר גם מנסה לקבוע האם בנתונים אלה קיים מתאם מסדר שני.
 ג. רשמו את המשווה הנוסף שעליו לאמוד.
 במשווה הנוסף התקבל מרובה השווה ל-12.0.
 ד. מהי המסקנה בר"מ של 0.05?

תשובות סופיות:

. $\sigma_u^2 = 5.263$. **ג.** $r_{(u_t, u_{t-4})} = 0.6561$. **ב.** $r_{(u_t, u_{t-1})} = 0.9$. **א.** **(1)**

. $r_{(u_t, u_{t-1})} = 0.4$, $r_{(u_t, u_{t-4})} = 0.0256$, $\sigma_u^2 = 1.19$. **ד.** **(2)** יש עדות לכך.

(3) יש עדות לכך.

. $\hat{\rho} = 0.927$. **ב.** $\hat{\rho} = 0.927$. **א.** יש עדות לכך. **(4)**

(5) לא נכונה.

(6) ראו סרטון.

(7) **א.** קיומו של מתאם סדרתי מסדר שלישי בתנויים.

$$. LM$$

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$$

$$. H_1 : \text{OTHERWISE}$$

$$. LM_{stat} = 22.99$$

יש עדות לכך.

. $\hat{C}_t = -2103.53 + 0.719 \cdot Y_t + 0.707 \cdot \hat{u}_{t-1} - 0.0064 \cdot \hat{u}_{t-2} - 0.328 \cdot \hat{u}_{t-3}$. **ב.**

. $Y_{t+1} = 5,568$. **ב.** $H_0 : \rho = 0$, **א.** $H_1 : \rho \neq 0$, יש עדות לכך. **(8)**

. $u_t = \gamma_0 + \gamma_1 \ln P_t + \gamma_2 \ln P_{t-1} + \gamma_3 u_{t-1} + \gamma_4 u_{t-2} + \omega_t$. **ג.** אין עדות לכך.

אקוונומטריקה

פרק 15 - סיכום בהתאם סדרתי והטרוקדסטיות

תוכן העניינים

1. כללי

93

סיכום מתאם סדרתי והטרוקדסטטיות:

רקע:

הטרוקדסטטיות	מתאם סדרתי	
$Y_t = \alpha + \beta X_t + u_t$: למשל:		המשוואת העיקרית של המודל
$V(u_t) = \sigma^2$	$\text{cov}(u_t, u_{t-s}) = 0$	הנחה הקלאסית המופרת
$V(u_t) = \sigma^2$	$\text{cov}(u_t, u_{t-s}) \neq 0$	המצב לאחר ההפרה
$V(u_t) = W_t \sigma^2$	$u_t = \rho u_{t-1} + \varepsilon_t$	המשוואת המאפיינת את ההפרה
מתקבלים אומדים חסרי הטיה ועקבים, אך תוכנות הייעילות נגעה.		מה קורה אם אומדים ב-OLS
מבחן GQ מבחן White	מבחן DW מבחן LM	זיהוי הבעיה
שיטת WLS	שיטת קווקרן – אורוקוט (רגסיטיות ההפירשימים) הכנסת משתנה מושבר בפיגור (מודל דינמי)	פתרון הבעיה

אקוונומטריקה

פרק 16 - מודלים דינמיים

תוכן העניינים

- 94 1. כללי

מודלים דינמיים:

רקע:

מודל דינמי הוא מודל שיש בו משתנה הקשור בפיגור, כלומר Y היום מושפע מ- Y של אתמול: $Y_t = \alpha + \beta_0 X_t + \beta_1 Y_{t-1} + u_t$.

המכפילים הדינמיים:

שלוש סוגים של השפעות בהקשר של המודל הדינמי (מכפילים):

1. מכפל לטוח קצר (מיידי):

איך X היום משפייע על Y היום: $\frac{\partial Y_t}{\partial X_t}$

2. מכפל ביןים מסדר j (מכפל דינמי):

איך X לפני j תקופות משפייע על Y היום: $\frac{\partial Y_t}{\partial X_{t-j}}$

3. מכפל טוח ארוך (מצב עמיד):

איך X משפייע על Y לאחר P תקופות: $\frac{\partial Y^*}{\partial X^*}$

כאשר X ו- Y נוטרים קבועים על פני הזמן (מצב עמיד):

$$Y_t = Y_{t-1} = \dots = Y_{t-p} = Y^*$$

$$X_t = X_{t-1} = \dots = X_{t-p} = X^*$$

הקשר בין מתאם סדרתי למודלים דינמיים

המתאים הסדרתי נובע מהשמטה של דינמיות מבנית במודל. המודל המקורי היה צריך להיות מודל דינמי אך נAMD בטיעות מודל סטטיסטי. הדינמיות Taboa אז לידי ביטוי בטיעויות, כלומר מותאם הסדרתי.

ורגסית ההפרשים, המהווה פתרון למתאים הסדרתי, היא למעשה מודל דינמי.

לסיכום:

בכדי לפתור את בעיית המתאים הסדרתי יש לאמוד מלבד את המשתנה הקשור בזמן t גם את המשתנה המוסבר והסביר בזמן $t-1$.

המשתנה בפיגור Y_{t-1} נועד לפתור מותאם סדרתי מסדר ראשון, Y_{t-2} משמש לפתרת מותאם סדרתי מסדר שני וכך הלאה.

בכדי לבדוק קיומו של מותאם סדרתי במודל דינמי לא יוכל לבצע מבחן DW אלא רק מבחן LM.

השלכות על אר"פ של משתנה מושבר בפיגור כמשתנה מסביר:

בניגוד למשתנה מסביר רגיל (X), Y_{t-1} הינו משתנה מקרי.
 שימוש כך אר"פ ברגסיה הכללת משתנים כאלה הם מוטים (להזיכרכם בהוכחת חוסר הטיה של האומדים השתמשו בהנחה מס' 4 הגורסת כי המשתנים המסבירים אינם משתנים מקרים).
 בנוסף לכך העקבות של האומדים תלויות בקיום מתאם סדרתי:

$$\hat{\beta} \rightarrow \beta + \frac{COV(Y_{t-1}, u_t)}{V(Y_{t-1})}$$

אם אין מתאם סדרתי: $COV(Y_{t-1}, u_t) = 0 \Leftarrow$ האומד עקיב.

אם יש מתאם סדרתי: $COV(Y_{t-1}, u_t) \neq 0 \Leftarrow$ האומד אינו עקיב.

לסיכום – ההשלכות על אר"פ:

1. האומדים מוטים ולכן ניתן לבצע בדיקת השערות רק במקרים גדולים ($T > 30$).
2. אם אין מתאם סדרתי \Leftarrow האומדים עקיבים ויעילים (ניתן לבצע בדיקת השערות במקרים גדולים).
 אם יש מתאם סדרתי \Leftarrow האומדים אינם עקיבים ואינם יעילים (לא ניתן לבצע בדיקת השערות גם במקרים גדולים).

שאלות:**חישוב מכפלים:**

1) חשבו את שלושת סוגי המכפלים של המודלים הדינמיים הבאים :

- א. $Y_t = \alpha + \beta_1 X_{t-1} + \beta_2 Y_{t-1} + u_t$.
 ב. $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 Y_{t-1} + u_t$.

תרגול מסכם:

2) המודל הבא הורץ ב-SAS עם מדגם בעל 100 תצפיות : $Y_t = \alpha + \beta X_t + u_t$.

א. מהפלט עולה $DW = 0.195$ לפיכך :

i. לא קיים מתאם סדרתי.

ii. קיים מתאם סדרתי והוא : _____.

iii. לא ניתן לקבוע אם המתאם הסדרתי מובהק.

ב. לפי תשובה לך לסעיף א' חוויה דעתך על תכונות האומדים :

i. מוטים נכון/ לא נכון/ לא ניתן לדעת

ii. ליניאריים נכון/ לא נכון/ לא ניתן לדעת

iii. עילים נכון/ לא נכון/ לא ניתן לדעת

iv. עקיבים נכון/ לא נכון/ לא ניתן לדעת

ג. אמידה של איזו משווהה תפטור באופן מלא את הבעה שנוצרה במודל :

$$. Y_t = \alpha + \beta_1 X_t + \beta_2 X_{t-1} + u_t . i$$

$$. Y_t = \alpha + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 X_{t-1} + u_t . ii$$

$$. Y_t = \alpha + \beta_1 X_t + \beta_2 Y_{t-1} + u_t . iii$$

ד. בדוק את ההשערה כי לפי מודל (3) השפעת X על Y הולכת ופוחתת עם הזמן.

מצורף החלק הרלוונטי מהפלט :

Parameter Estimates

Variable	DF	Parameter	Standard	T for H0:	Prob> T
		Estimate	Error	Parameter=0	
INTERCEP	1	0.42	0.06	7.00	0.000
X	1	0.25	0.03	8.33	0.000
Y1	1	0.85	0.05	17.00	0.000

ה. מהו המכפיל הדינמי בתקופה $t-8$?

(3) הקשר בין כמות הכספי לבין רמת האינפלציה במשק נאמד בסדרה עתית על ידי המשוואה הבאה :

$$1. M_t = \alpha + \beta \cdot P_t + U_t$$

כאשר :

M_t - כמות הכספי במשק בחודש t .

P_t - מדד המוצרים לצרכן במשק בחודש t .

משוואה (1) נאמדה בפלט מס' 1.

א. לפי מבחון על הסטטיסטי DW , נראה כי $-U_t$:

i. לא ניתן לחשב את הסטטיסטי DW מהנתונים הקיימים.

ii. קיימים מתאימים סדרתי שלילי.

iii. קיימים מתאימים סדרתי חיובי.

iv. לא קיימים מתאימים סדרתי.

v. לא ניתן לקבוע אם המתאים הסדרתי מובהק.

ב. סמננו את התשובה הנכונה בהכרח :

i. האומדיים ליניאריים חסרי הטיה, עקיבים אך לא יעילים.

ii. האומדיים ליניאריים חסרי הטיה, עקיבים ויעילים.

iii. האומדיים מוטים אך עקיבים.

iv. האומדיים חסרי הטיה, אך לא עקיבים.

v. כל התשובות אינן נכונות.

ג. אומד השונות מوطה ובדיקת השערות

נכונה/לא נכון/אי אפשר לדעת לא תקפה :

חוקר טען כי הבעיה שנוצרה במשוואה (1) טיפול ע"י אמידת המשוואה הבאה :

$$2. M_t = \alpha_1 + \beta_1 \cdot P_t + \beta_2 \cdot M_{t-1} + \varepsilon_t$$

כאשר :

M_{t-1} - כמות הכספי בשנה הקודמת.

משוואה (2) נאמדת בפלט מס' 2, כמו כן נאמדת על ידי החוקר המשוואה המופיעה בפלט מס' 3.

ד. הרגרסיה המופיעה בפלט מס' 3 נועדה לבדיקת :

במשוואה : _____.

על ידי מבחון : _____.

גודל הסטטיסטי لمבחן הינו (רשום תוצאה מספרית) : _____.

ה. לאור תשובהך לסעיף ד' טענת החוקר : נכון/לא נכון/אי אפשר לדעת

ו. האומד ל- β_1 במשוואה (2) הוא מوطה

אך עקיב : _____.

ז. ניתן להשתמש בסטטיסטי DW לבדיקת

מתאימים סדרתי במשוואה (2) : נכון/לא נכון/אי אפשר לדעת.

ח. חשבו את המכפיל הדינמי לשינוי P בתקופה 1- t .

ט. בדקו את הטענה כי המכפיל הדינמי לשינוי P בתקופה $t-1$

הינו 90% מהמכפיל המיידי בטוחה הקצר.

ו. רשמו את השערת האפס עבור הטענה כי המכפיל בט"א שווה ל-1.

מהו המבחן הסטטיסטי המתאים לבחינת ההשערה?

פלט מס' 1 – משואה (1):

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	44.03828	44.03828	5145.80	<.0001
Error	103	0.88148	0.00856		
Corrected Total	104	44.91976			
Root MSE		0.09251	R-Square	0.9804	
Dependent Mean		8.53854	Adj R-Sq	0.9802	
Coeff Var		1.08344			
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	1.49372	0.09862	15.15	<.0001
p	1	1.69267	0.02360	71.73	<.0001
Durbin-Watson D		0.208			

פלט מס' 2 – משואה (2):

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	42.19946	21.09973	15988.1	<.0001
Error	101	0.13329	0.00132		
Corrected Total	103	42.33275			
Root MSE		0.03633	R-Square	0.9969	
Dependent Mean		8.55393	Adj R-Sq	0.9968	
Coeff Var		0.42469			
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	0.40374	0.06790	5.95	<.0001
m1	1	0.81811	0.03857	21.21	<.0001
p	1	0.28127	0.06633	4.24	<.0001

$$M1 = M_{t-1}$$

פלט מס' 3:

Dependent Variable: res Residual $\hat{RES} = \hat{\varepsilon}_t = (\text{a}) \text{ נקיון ו/צ}$

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00032896	0.00010965	0.08	0.9695
Error	99	0.13173	0.00133		
Corrected Total	102	0.13206			
Root MSE		0.03648	R-Square	0.0025	
Dependent Mean		0.00033853	Adj R-Sq	-0.0277	
Coeff Var		10775			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	0.03298	0.08192	0.40	0.6881
p		1	0.02640	0.07979	0.33	0.7415
m1		1	-0.01672	0.04705	-0.36	0.7230
res1		1	-0.01304	0.11039	-0.12	0.9062

$$\hat{RES1} = \hat{\varepsilon}_{t-1}$$

תשובות סופיות:

1) א. מכפיל מיידי: $\frac{\partial Y_t}{\partial X_{t-j}} = \beta_1 \beta_2^{j-1}$, מכפיל טוח ביןים: $\frac{\partial Y_t}{\partial X_t} = 0$, מכפיל של הטוח הארוך:

$$\frac{\partial Y^*}{\partial X^*} = \beta_1 \cdot \frac{1}{1 - \beta_2}$$

ב. מכפיל מיידי: $\frac{\partial Y_t}{\partial X_{t-j}} = \beta_2^{j-1} (\beta_1 + \beta_0 \beta_2)$, מכפיל טוח ביןים: $\frac{\partial Y_t}{\partial X_t} = \beta_0$

מכפיל של הטוח הארוך:

2) א. ii, חיובי. ב. i. לא נכון. ג. iii. לא נכון.

ד. יש עדות לכך. א. iv. נכון.

$$\frac{\partial Y_t}{\partial X_{t-j}} = 0.068$$

3) א. iii. נכון. ב. i. לא נכון. ג. ii. לא נכון.

ד. קיומ מתחם סדרתי במשוואת (2).

מבחן LM.

$$LM_{stat} = 0.2575$$

4) $\frac{\partial M_t}{\partial P_{t-i}} = \beta_1 \beta_2^i$ א. לא נכון. ב. נכון. ג. נכון.

ט. יש עדות לכך. י. WALD, $H_0: \beta_1 = 1 - \beta_2$.

אקוונומטריקה

פרק 17 - סיכום תכונות ארפ

תוכן העניינים

100.....
1. כללי.....

סיכום תכונות ארכ'פ:

רקע:

אם המודל שאותו אנו אומדים מנסה לנכון את הקשר המתמטי בין המשתנים וכ כולל את כל המשתנים הרלוונטיים להסביר התופעה ורק אותם ולא קיים קשר חזק או מלא בין המשתנים הבלטי תלויים במודל וכל ההנחהות הקלאסיות מתקיימות, אז האומדים שאותם נקבל הם א.ח.ה, הם BLUE (יעילים) והם עקיפים.

פגיעה בתכונות ארכ'פ	כל הנקודות הקלאסיות מתיקיות	אין קשר חזק או מלא בין הב'ית	לא כולל משתנים שלא אינם רלוונטיים	כולל את כל המשתנים הרלוונטיים	הבעיה במודל
אין פגיעה	✓	✓	✓	✓	אין בעיה
פגיעה בכל התכונות - *מוטים, לא יעילים ולא עקיפים	✓	✓	✓	✗	השמדת משתנים רלוונטיים
אין פגיעה	✓	✓	✗	✓	הוספת משתנים לא רלוונטיים
אין פגיעה	✓	✗	✓	✓	מולטיקוליניאריות חיליקת
ארכ'פ אינים מוגדרים	✓	✗	✓	✓	מולטיקוליניאריות מלאה
יעילות	$V(u_t) = \sigma^2$	✓	✓	✓	הטרוסקડסטיטות
יעילות	$\text{cov}(u_t, u_s) \neq 0$	✓	✓	✓	מתאים סדרתי
מוטים הייעילות והעקבות תלויים בקיים מתאם סדרתי	$\text{cov}(x, u) \neq 0$	✓	✓	✓	מודלים דינמיים
**אומדים מוטים אך עקיפים. הייעילות תלואה בשיטה.	$\text{cov}(x, u) \neq 0$	✓	✓	✓	משוואות סימולטניות

* אלא אם אין מתאם בין המשתנה הב'ית שהושמט לאלו המוצאים במודל.

* בהנחה שהמשוואות מזוהות.

אקוНОמטריקה

פרק 18 - הדרכה בקריאת פלטימ של תוכנת VIEW

תוכן העניינים

1. כללי

(ללא ספר)