אקונומטריקה למדעי הנתונים

פרק 34 - מבחן לדוגמה מס' 2

תוכן העניינים

1. כללי 1

מבחן לדוגמא מס' 2:

שאלות:

לשם חישובים הנח כי ערך t הינו 2 וערך F הינו 4.

1) הנח כי הקשר באוכ' בין X ל-Y נתון על ידי המשוואה הבאה: $Y_t = \beta \cdot X_t + U_t$, כאשר כל ההנחות הקלאסיות מתקיימות.

נתון האומד: $\tilde{\beta} = \frac{\sum Y_t}{S_{XX}}$.

א. האם האומד ליניארי?

ב. האם האומד חסר הטיה?

ג. אומד זה יעיל פחות מאומד הריבועים הפחותים: נכון/לא נכון/אי אפשר לדעת

ד. האם אומד זה הוא blue?

ה. אומד $\tilde{\beta}$ מוגדר רק כאשר: $S_X^2 \neq 0$: נכון/לא נכון/אי אפשר לדעת

ו. חשבו את השונות של β עבור מודל שבו $\alpha \neq 0$.

ז. שונות האומד (שחושבה בסעיף הקודם) הינה גדולה משונות המודל הנתון: נכון/לא נכון/אי אפשר לדעת

2) על סמך מדגם של 60 משפחות שלכל אחת 3 ילדים נאמדו המשוואות הבאות:

1. $\hat{y}_i = 15 + 0.7x_{1i} + 0.35x_{2i} + 0.20x_{3i}$ $R^2 = 0.85$.
2. $y_i = 2 + 0.1z_i$ $R^2 = 0.25$.
3. $z_i = x_{1i} - x_{2i} + 2x_{3i}$.

כאשר y_i הינן הוצאות משק הבית על חינוך הילדים ואילו x_{ji} הינו גילו של הילד j.

א. ההשערה שניתן לבדוק באמצעות המשוואות הנתונות הינה:

i. $H_0 : \beta_1 = \beta_2; \beta_1 = 2\beta_3$.

ii. $H_0 : \beta_1 = -\beta_2 = 2\beta_3$.

iii. $H_0 : \beta_2 = -\beta_1; \beta_3 = 2\beta_1$.

iv. לא ניתן לדעת.

ב. סטטיסטי המבחן שניתן לבדוק באמצעות המשוואות הנתונות שווה בקירוב ל:

i. 56.

ii. 57.

iii. 112.

iv. 74.66.

3) כלכלן הציע את המודלים הבאים:

1. $y_i = \beta_1 \ln(x_i) + \beta_2 \ln(0.5x_i) + u_i$.
2. $y_i = \beta_1 \ln(x_i) + \beta_2 \ln(x_i^{0.5}) + u_i$.

האם ניתן לאמוד את המודלים בשיטת OLS?

א. אין בעיה לאמוד את שני המודלים.
ב. לא ניתן לאמוד את המודל הראשון בלבד.
ג. לא ניתן לאמוד את המודל השני בלבד.
ד. לא ניתן לאמוד את שני המודלים.

4) כלכלן אמד את המודל הבא: $y_i = \alpha_0 + \alpha_1 \ln(x_i) + u_i$ וקיבל את האומדנים:
$\hat{\alpha}_0 = 10$ ו- $\hat{\alpha}_1 = 6$.

על אותו המדגם אמד חברו את המודל הבא: $y_i = \beta_0 + \beta_1 \ln(x_i^2) + u_i$. מכאן ש:

א. $\hat{\beta}_0 = 5$ ו- $\hat{\beta}_1 = 3$.
ב. $\hat{\beta}_0 = 10$ ו- $\hat{\beta}_1 = 3$.
ג. $\hat{\beta}_0 = 5$ ו- $\hat{\beta}_1 = 6$.
ד. כל התשובות שגויות.

5) על סמך מדגם של 95 תצפיות נאמד המודל הבא: $y_i = 2 + 0.5x_{1i} + 0.3x_{2i}$ $R^2 = 0.73$.
(1) (2)

הערכים שבסוגריים הם סטיות התקן של המקדמים.

א. בדוק האם המודל מובהק.
ב. בדוק האם מקדמי השיפוע מובהקים.
ג. מה תוכל להסיק מסעיפים א' ו-ב'?

6) על סמך מדגם של 52 תצפיות נאמדו המשוואות הבאות:

1. $\hat{y}_i = 4 + 0.1x_{1i} + 0.8x_{2i}$ $R^2 = 0.84$.
2. $\hat{y}_i = 2 + 0.8x_{1i}$ $R^2 = 0.7$.
3. $\hat{y}_i = 7 + 0.23x_{2i}$ $R^2 = 0.25$.
4. $\hat{y}_i = 3 + 0.23z_i$ $R^2 = 0.55$.

כאשר x_{1i} ו- x_{2i} הם השכלת הבעל והאישה בהתאמה במשפחה i ו- y_i הכנסת משק בית i.

כמו כן נתון כי: $z_i = x_{1i} + 2.2x_{2i}$.

א. בדוק את ההשערה כי להשכלה אין השפעה על הכנסות המשפחה.
ב. איזה השערה ניתן לבדוק תוך שימוש במשוואות (1) ו-(4)? בדוק אותה.
ג. חשב את סטית התקן של המקדם x_{1i} ברגרסיה (1).

7) חוקר מעוניין לאמוד את המודל: $y_i = \alpha + u_i$.

א. חשב את נוסחת אומד הריבועים הפחותים ל- α על ידי פיתרון בעיית המינימיזציה של סכום ריבועי הסטיות.

ב. חשב את נוסחת שונותו של האומד.

8) על סמך מדגם של 45 תצפיות נאמדו המודלים הבאים:

1. $y_i = 5.4 + 1.2x_{2i} + 4.4x_{3i} + u_i \quad R^2 = 0.75$.
2. $y_i = 6.3 + 5.8x_{3i} + u_i \quad R^2 = 0.65$.
3. $y_i = 5.7 + 1.2x_{2i} + u_i \quad R^2 = 0.70$.
4. $y_i = 3.9 + 3.4\ln(x_{2i}) + u_i \quad R^2 = 0.56$.
5. $\ln(y_i) = 2.4 + 1.8x_{2i} + 2.7x_{3i}^2 + 4.2x_{4i}^2 + u_i$.
6. $y_i = 1.3 + 3.1x_{2i} + 0.5x_{3i} + 4.8x_{4i}^2 + 1.5x_{5i}^2 + u_i$.

א. דרג את הרגרסיות על פי מדד ההסבר (מהנמוך לגבוה)

ב. בדוק את ההשערות של משתנים X_2 ו- X_3 ביחד אין השפעה על Y במודל (1).

ג. בדוק בהסתמך על מודל (2) האם המשתנה X_2 מובהק ברגרסיה (1).

ד. ברגרסיה (1) נתונים כעת אומדי הטעויות הסטנדרטיות (סטיות התקן) של מקדמי X_2 ו- X_3 0.5 ו-2.5 בהתאמה. בדוק עבור כל אחד מהמקדמים הנ"ל האם מובהק ומה אפשר ללמוד מרגרסיה (1).

ה. איזו השערה ניתן לבדוק תוך שימוש במשוואות (6) ו-(3)?

9) על מנת לאמוד השפעת מגדר ומצב משפחתי על השכר, נאמדה המשוואה הבאה:

$$WAGE = \alpha_0 + \alpha_1 \cdot GENDER + \alpha_2 \cdot FS + \alpha_3 \cdot (GENDER \cdot FS) + \beta_1 \cdot EDUC + \beta_2 \cdot AGE + U$$

כאשר:

$GENDER$ = מגדר: 1=גבר, 0=אישה.

FS = מצב משפחתי: 1=נשואים, 0=לא נשואים.

$EDUC$ = מס' שנות לימוד של העובד.

AGE = גיל העובד.

$WAGE$ = שכר העובד.

משוואה (1) נאמדה בפלט מס' 1.

בנוסף נאמד גם פלט מס' 2.

א. החוקרת הניחה כי פערי השכר, באים לידי ביטוי בשכר ההתחלתי בלבד: נכון/ לא נכון/ לא ניתן לדעת.

ב. החוקרת הניחה כי הפערים בין נשים לגברים בשכר אינם תלויים בגיל: נכון/ לא נכון/ לא ניתן לדעת.

ג. השערת האפס לבדיקת הטענה היא: _____.

ד. המשתנה המוסבר ברגרסיה מס' 2 הינו: _____.
(כתבו את המודל שבו מחושב המשתנה המוסבר).

ה. הסטטיסטי של LM לבדיקת הטענה:

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא: _____.

ו. המקדם של GENDER בפלט מס' 2 הינו: _____.

הועלתה הטענה כי הפערים בין גברים לנשים בקרב העובדים הנשואים גבוהים ביותר מ-1500 ₪ מאשר הפערים בקרב העובדים שאינם נשואים.

ז. ההשערות לבדיקת הטענה הינן: _____.

ח. הסטטיסטי לבדיקת הטענה:

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא: _____.

ט. ההשערות לבדיקת הטענה הן: _____.

י. המודל המוגבל לבדיקת הטענה הוא: _____.

יא. הסטטיסטי לבדיקת הטענה:

i. לא ניתן לחשבו בעזרת הנתונים הקיימים.

ii. ניתן לחישוב וערכו הוא: _____.

פלט מס' 1 - משוואה 1:

Dependent variable: WAGE
Number of observations used: 17495
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	1.504815E11	30096294654	646.42	<.0001
Error	17489	8.142567E11	46556220		
C Total	17494	9.647382E11			

Root MSE	6823.35843	R-square	0.1560
Dep Mean	7286.58004	Adj R-sq	0.1557
C.V.	93.64281		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	-3642.10108	260.72351	-13.97	<.0001
GENDER	1	2006.13583	187.64224	10.69	<.0001
FS	1	899.68055	159.19316	5.65	<.0001
GENDER*FS	1	1964.31810	227.43348	8.64	<.0001
EDUC	1	428.20041	12.45434	34.38	<.0001
AGE	1	64.72379	4.43948	14.58	<.0001

פלט מס' 2 - מבחן LM:

Dependent variable :
Number of observations used: 17495
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	66653745252	13330749050	286.32	<.0001
Error	17489	8.142567E11	46558220		
C Total	17494	8.809105E11			

Root MSE	6823.35843	R-square	0.0757
Dep Mean	2.29222E-12	Adj R-sq	0.0754
C.V.	2.97675E17		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	-1244.40187	260.72351	-4.77	<.0001
GENDER	1				
FS	1	899.68055	159.19316	5.65	<.0001
GENDER*FS	1	1964.31810	227.43348	8.64	<.0001
EDUC	1	23.18457	12.45434	1.86	0.0627
AGE	1	-38.13257	4.43948	-8.59	<.0001

10) ענה על הסעיפים הבאים:

א. במודל דינאמי עם מתאם סדרתי האומדים מוטים אך עקיבים: נכון/לא נכון/לא ניתן לדעת

ב. בהשמטת משתנה רלוונטי, אם נתון שהמתאם בין המשתנים המסבירים חיובי, והשפעת המשתנה שהושמט שלילית, אזי האומד ל- β במודל הנאמד בעל הטיה חיובית: נכון/לא נכון/לא ניתן לדעת

ג. המודל הבא נאמד במדגם של 1000 תצפיות והתקבלו התוצאות הבאות:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + U_t$$
$$(0.2) \quad (0.8) \quad (0.6)$$

$R^2 = 0.4$.

(המספרים בסוגריים הם ערכי t של המקדמים).

לאור הנתונים שהתקבלו נראה כי במשוואה:

i. אין כל בעיה סטטיסטית ולכן האומדים המתקבלים הם חסרי הטיה.

ii. יש בעיה של: _____, אך האומדים המתקבלים עדיין חסרי הטיה.

iii. יש בעיה של: _____ והאומדים המתקבלים מוטים.

ד. המודל הבא נאמד במדגם של 1000 תצפיות והתקבלו התוצאות הבאות:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + U_t$$

(-6.3) (0.2) (7.8)

לאור הנתונים שהתקבלו נראה כי במשוואה:

i. אין כל בעיה סטטיסטית ולכן האומדים המתקבלים הם חסרי הטיה.

ii. יש בעיה של: _____, אך האומדים המתקבלים עדיין חסרי הטיה.

iii. יש בעיה של: _____ והאומדים המתקבלים מוטים.

11) נתון המודל: $Y = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t$.

א. אם קיים מתאם סדרתי מסדר ראשון, אומדי הריבועים הפחותים חסרי הטיה ועקיבים אך לא יעילים: נכון/לא נכון/לא ניתן לדעת

ב. הוספת המשתנה Y_{t-1} למשוואה, יכולה לפתור את בעיית המתאם הסדרתי מסדר ראשון: נכון/לא נכון/לא ניתן לדעת

ג. רגרסיית עזר למתאם סדרתי מסדר שלישי הינה: _____.

ד. דרגות החופש של הערך הקריטי לבדיקת ההשערה בסעיף ג' הינן: _____.

12) נתון המודל:

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

א. האומדים המתקבלים בשיטת הריבועים הפחותים הם מוטים, אך עקיבים: נכון/לא נכון/אי אפשר לדעת

פתרון הבעיה הקיימת במשוואה (1) ייתכן על ידי אמידת המשוואה הבאה: $(Y_t \cdot W_t) = \delta_0 \cdot W_t + \delta_1 \cdot (X_{1t} \cdot W_t) + \delta_2 \cdot (X_{2t} \cdot W_t) + V_t$.

ב. כאשר W_t הינו: _____.

ג. משוואה (2) בצורתה המפורשת הינה:

_____ + _____ · _____ + _____ · _____ = _____.

מלאו את החסר במונחי המשתנים של משוואה (1).

ד. האומד היעיל ל- σ^2 הינו: _____.

תשובות סופיות:

1) א. כן. ב. מוטה. ג. אי אפשר לדעת. ד. לא.
ה. נכון. ו. $V\left(\tilde{\beta}\right)=\frac{n\sigma^2}{S_{XX}^2}$. ז. לא נכון.

2) א. iii. ב. iii.

3) ד'.

4) ב'.

5) א. מובהק. ב. אינם מובהקים. ג. ראו סרטון.

6) א. מובהק. ב. $H_0:\beta_2=2.2\beta_1$. ג. $S.E=0.00743$.

7) א. $\hat{\alpha}=\bar{y}$. ב. $V\left(\hat{\alpha}\right)=\frac{\sigma_e^2}{n}$.

8) א. $4>2>3>1>6$. ב. יש עדות לכך. ג. יש עדות לכך.
ד. X_2 מובהק, X_3 אינו מובהק. ה. $HO:\beta_2=\beta_4=\beta_5=0$.

9) א. נכון. ב. נכון. ג. $H_0:\alpha_1=\alpha_2=\alpha_3=0$.
ד. ראו סרטון. ה. ראו סרטון. ו. ראו סרטון.
ז. $\begin{matrix}H_0:\alpha_3=1500\\H_1:\alpha_3>1500\end{matrix}$. ח. ii, $t_{stat}=2.04$. ט. $\begin{matrix}H_0:\alpha_1=\alpha_2\\H_1:\alpha_1\neq\alpha_2\end{matrix}$.
י. $WAGE=\alpha_0+\alpha_2\cdot\left(GENDER+FS\right)+\alpha_3\cdot\left(GENDER\cdot FS\right)+\beta_1\cdot EDUC+\beta_2\cdot AGE+U$.
יא. i.

10) א. לא נכון. ב. לא נכון. ג. ii, בעיה של מולטיקולינאריות חלקית.
ד. ii, בעיה של הוספת משתנה לא רלוונטית.

11) א. נכון. ב. נכון. ג. $\hat{u}_t\ /\ X_{1t},X_{2t};\hat{u}_{t-1},\hat{u}_{t-2},\hat{u}_{t-3}$.
ד. 3.

12) א. לא נכון. ב. $W_t=\frac{X_{2t}}{X_{1t}}$.
ג. $\frac{Y_t\cdot X_{2t}}{X_{1t}}=\delta_0\cdot\frac{X_{2t}}{X_{1t}}+\delta_1\cdot X_{2t}+\delta_2\cdot\frac{X_{2t}^2}{X_{1t}}+\frac{u_tX_{2t}}{X_{1t}}$. ד. $S^2=\frac{ESS}{T-k}$.