אקונומטריקה א

תוכן העניינים

אקונומטריקה א

פרק 1 - מבוא לקורס

תוכן העניינים

מבוא לקורס:

רקע:

הגדרות וסימונים:

משתנה אמפירי – תוצאותיו ידועות מראש (למשל: רמת הכנסה, גיל, מסי שנות לימוד במדגם מסוים).
משתנה מקרי – תוצאותיו לא ידועות מראש (כגון תוצאה בהטלת קובייה או בהטלת מטבע). באקונומטריקה נעסוק בעיקר במשתנים מקריים.

שני סוגי המשתנים יסומנו באות לועזית עם אינדקס (למשל: Y_t או X_t).
קבוע – מקבל ערך אחד בלבד (מסומן באות לועזית ללא אינדקס – למשל a או b).
לכל משתנה מקרי X_t יש תוחלת המייצגת את מרכז ההתפלגות (μ_X או $E(X)$).
השונות – מייצגת את מידת הפיזור של ההתפלגות (σ_X^2 או $V(X)$).
סטית התקן – היא השורש של השונות (σ_X).
שונות משותפת (covariance) – מדד להתפלגות המשותפת של שני משתנים מקריים ומייצגת את הכיוון של הקשר ביניהם ($\mathrm{Cov}(X,Y)$):

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

$\mathrm{Cov}(X,Y)>0 \Leftrightarrow$ מתאם חיובי בין המשתנים.

$\mathrm{Cov}(X,Y)<0 \Leftrightarrow$ מתאם שלילי בין המשתנים.

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

מקדם המתאם של פירסון – מדד לכיוון ולעוצמת הקשר הליניארי בין שני משתנים: $\eta_{xy}=\dfrac{\mathrm{cov}(x,y)}{\sigma_x\cdot\sigma_y}$.

$-1\le\eta\le 1$

$\eta=1$ מתאם ליניארי חיובי מלא בין שני המשתנים.

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

$\eta=0$ לא קיים מתאם ליניארי בין שני המשתנים.

אמידה:

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

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

נוסחאות וחוקים בסטטיסטיקה:

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

<u>חוקי הסיגמה</u>:

1. $\sum_{t=1}^{T}X_t=X_1+X_2+X_3+\ldots+X_T$.
2. סכום של קבוע: $\sum_{t=1}^{T}a=Ta$.
3. סכום של קבוע כפול משתנה = לקבוע כפול הסכום: $\sum_{t=1}^{T}aX_i=a\sum_{t=1}^{T}X_t$.
4. סכום של סכום/הפרש = לסכום/הפרש הסכומים: $\sum_{t=1}^{T}(X_t\pm Y)_t=\sum_{t=1}^{T}X_t\pm\sum_{t=1}^{T}Y_t$.
5. יש לשים לב כי: $\sum_{t=1}^{T}X_t^2\neq\left(\sum_{t=1}^{T}X_t\right)^2$

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

הגדרות ופיתוחים:

1. סכום הסטיות מהממוצע = 0: $\sum_{t=1}^{T}\left(X_t-\bar{X}\right)=0$.
2. סכום הסטיות הריבועיות מהממוצע (מונה השונות):
$$S_{XX}=\sum_{t=1}^{T}\left(X_t-\bar{X}\right)^2=\sum_{t=1}^{T}X_t^2-T\bar{X}^2=\sum_{t=1}^{T}\left(X_t-\bar{X}\right)X_t.$$
3. מונה של השונות המשותפת:
$$S_{XY}=\sum_{t=1}^{T}\left(X_t-\bar{X}\right)\left(Y_t-\bar{Y}\right)=\sum_{t=1}^{T}X_tY_t-T\bar{X}\bar{Y}=\sum_{t=1}^{T}\left(X_t-\bar{X}\right)Y_t=\sum_{t=1}^{T}\left(Y_t-\bar{Y}\right)X_t.$$

חוקי התוחלת:

1. תוחלת של קבוע = קבוע: $E(a)=a$.
2. תוחלת של סכום/הפרש = לסכום/הפרש התוחלות:
$$E(X\pm Y)=E(X)\pm E(Y)$$
$$E\left(\sum(X_i)\right)=\sum E(X_i).$$
3. תוחלת של כפל/חילוק ≠ לכפל/חילוק התוחלות:
$$E(X\cdot Y)\neq E(X)\cdot E(Y)$$
$$E\left(\frac{X}{Y}\right)\neq\frac{E(X)}{E(Y)}.$$
$$E\left(X^2\right)\neq\left[E(X)\right]^2$$
4. השפעת טרנספורמציה ליניארית על התוחלת:
$$E\left(a/\frac{1}{a}X\pm b\right)=a/\frac{1}{a}\cdot E(X)\pm b.$$

חוקי השונות:

1. עבור X ו- Y בלתי תלויים/בלתי מתואמים מתקיים:
   שונות של סכום/הפרש = סכום השונויות:
$$V(X\pm Y)=V(X)+V(Y)$$
$$V\sum(X_i)=\sum V(X_i).$$
2. עבור X ו- Y תלויים/מתואמים מתקיים:
   שונות של סכום/הפרש ≠ סכום השונויות:
$$V(X\pm Y)=V(X)+V(Y)\pm 2\cdot Cov(X,Y).$$

3. שונות של קבוע = 0: $V(a)=0$, $V(a\pm x)=V(X)$.

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

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

<u>חוקי השונות המשותפת</u>:

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

המודל האקונומטרי:

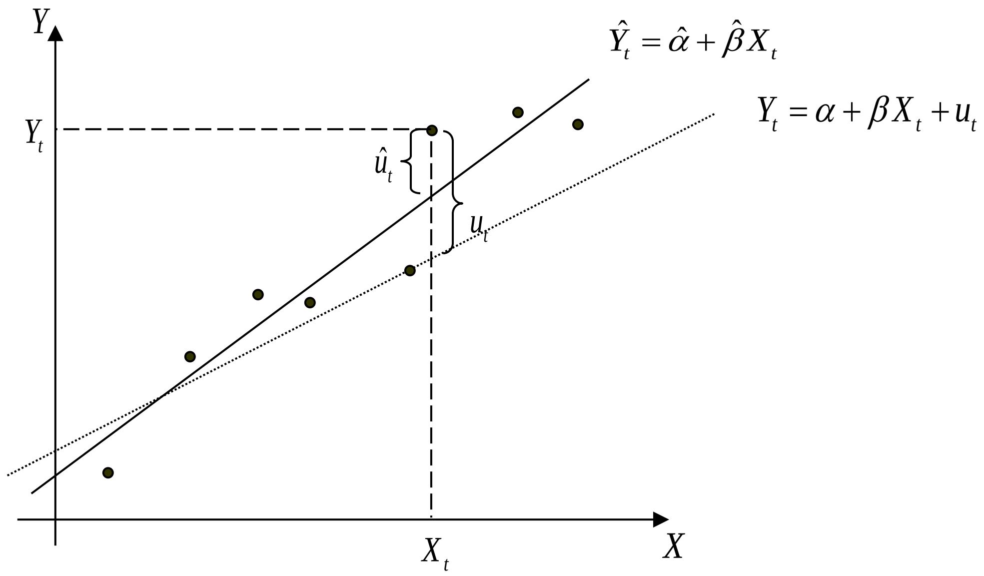
$$Y=\alpha+\beta X+u$$

משתנה מוסבר	פרמטרים: α - חותך, β - שיפוע	משתנה מסביר	הפרעה אקראית

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

4. בעוד α ו-β הם קבועים, $\hat{\alpha}$ ו-$\hat{\beta}$ הם משתנים מקריים כיוון 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 באוכ' לא ידוע.
6. אפשר לדעת את $\hat{u}_t$, שהיא הסטיה מקו הרגרסיה במדגם:

- עבור X_t, הערך הצפוי של המשתנה המוסבר $\left(\hat{Y}_t\right)$ המתקבל לפי הרגרסיה הוא: $\hat{Y}_t = \hat{\alpha} + \hat{\beta}X_t$.
- הסטיה של התצפית (Y_t) מהערך הצפוי לפי הרגרסיה $\left(\hat{Y}_t\right)$ היא: $\hat{u}_t = Y_t - \hat{Y}_t$.



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

שאלות:

1) הבא נוכיח את הזהויות הבאות:

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

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

ג. $\sum\left(X_t-\bar{X}\right)=0$.

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

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

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

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

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

א. $Var(ax)$.

ב. $Var(x+y)$.

ג. $Var(ax+b)$.

ד. $Cov(x,ay)$.

ה. $Cov(x+a,y+b)$.

ו. מקדם המתאם בין x ל- y.

תשובות סופיות:

1) הוכחה.

2) א. $a^2\operatorname{var}(x)$. ב. $\operatorname{var}(x)+\operatorname{var}(y)+2\operatorname{cov}(x,y)$. ג. $a^2\operatorname{var}(x)$. ד. $a\operatorname{cov}(x,y)$.

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

אקונומטריקה א

פרק 2 - אומדי הריבועים הפחותים

תוכן העניינים

אומדי הריבועים הפחותים:

רקע:

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

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

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

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

"המשוואות הנורמליות" מתקבלות בתהליך הגזירה של פונקציית הריבועים הפחותים וחייבות להתקיים על מנת שהפונקציה תתקיים $\left(\sum \hat{u}_t^2 = \min\right)$:

עבור המודל הקלאסי (עם חותך):

בגזירה של α: $\sum \hat{u}_t = 0$.

בגזירה של β: $\sum \hat{u}_t \cdot x_t = 0$.

עבור מודל ללא חותך:

בגזירת β בלבד: $\sum \hat{u}_t \cdot x_t = 0$.

מן המשוואות הנורמליות נובעות:

1. התכונות הגיאומטריות:

   א. $\sum \hat{u}_i = 0$.

   ב. $\sum x_i \hat{u}_i = 0$.

- ברגרסיה <u>ללא שיפוע</u> מתקיימת רק התכונה הגיאומטרית הראשונה.
  ברגרסיה <u>ללא חותך</u> מתקיימת רק התכונה הגיאומטרית השנייה.

2. התכונות האלגבריות:

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

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

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

- התכונות האלגבריות תקפות עבור קו הרגרסיה הקלאסי (עם חותך ושיפוע) במדגם בלבד.

ההנחות הקלאסיות של מודל הרגרסיה:

1. קיים קשר ליניארי בין המשתנה המוסבר למשתנה המסביר.
2. X איננו קבוע: $S_{XX} = \sum_{t=1}^{T} \left(X_t - \bar{X}\right)^2 \neq 0$.
3. תוחלת ההפרעה האקראית היא אפס לכל תצפית: $E(u_t) = 0$ לכל t.
4. X_t אינם משתנים מקריים ⇐ ניתן להוציא אותם מחוץ לתוחלת ולשונות ⇐ $\operatorname{cov}(X_t, u_t) = 0$.

5. הומוסקדסטיות: שונות ההפרעה האקראית קבועה לכל תצפית: $V(u_t)=\sigma_u^2$ לכל t.
6. u_t ב"ת: $\operatorname{cov}(u_t,u_s)=0$ לכל $t\neq s$.
7. ההפרעות האקראיות מתפלגות נורמלית: $u_t \approx N$.

תכונות האומדים:

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

1. לינאריות:

אר"פ ניתנים להצגה כטרנספורמציה לינארית של Y_t.

כדי ש-$\hat{\beta}$ למשל, יהיה אומד ליניארי צריך להתקיים: $\hat{\beta}=\sum W_t\cdot Y_t$.

כאשר W_t היא קומבינציה של ערכי X בדרך כלל. למשל: $\hat{\beta}=\dfrac{\sum X_t\cdot Y_t}{\sum X_t^2}$.

כדי להביא את האומד לצורה: $\tilde{\beta}=\sum w_t\cdot y_i$ נעזר בשוויון: $\dfrac{\sum()}{\sum()}=\sum\dfrac{()}{\sum()}$.

אומד זה ניתן להצגה 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 ליניארי.

- שימו לב כי:

W_t אסור שיכלול את Y_t.

Y_t אסור שיהיה במכנה או בשורש/חזקה (אלא אם כן במודל הנתון הוא מצוי בשורש/חזקה).

2. חוסר הטיה:

אומד $\hat{\vartheta}$ מסוים יהווה אח"ה לפרמטר ϑ אותו הוא אומד באוכלוסייה אם מתקיים: $E(\hat{\vartheta})=\vartheta$.

כיצד יודעים אם אומד הוא חסר הטיה?

1. בשלב הראשון יש לבצע עבודת הכנה – מבטאים את האומד באמצעות הפרמטר האמיתי – מציבים במקום ה- Y_t את המודל ומפתחים אלגברית.

> • יש לזכור כי:
> u_t מהווים משתנים מקריים ⇐ נשארים בתוך התוחלת, השונות
> y_t
> וה- $\sum$.
> x_t איננו משתנה מקרי (על פי הנחה מס' 4) ⇐ יוצא מחוץ לתוחלת ולשונות אך נשאר בתוך ה- $\sum$ ו- α β קבועים ⇐ יוצאים מחוץ לתוחלת, לשונות ול- $\sum$.

2. בשלב השני מפעילים תוחלת על האומד המפותח ואם התוחלת שווה לפרמטר האמיתי אז האומד חסר הטייה.

• חוסר הטייה מחייב את התקיימותן של הנחות (3) $E(u_t)=0$ לכל t ו- (4) $\operatorname{cov}(X_t,u_t)=0$.

3. יעילות:

יעילות פירושה השונות הקטנה ביותר. ככל שהשונות של האומד קטנה יותר, כך יש הסתברות גבוהה יותר שהוא יהיה קרוב לפרמטר האמיתי באוכלוסייה אותו הוא אומד.

$\hat{\vartheta}_1$ יקרא אומד יעיל יותר מ- $\hat{\vartheta}_2$ אם מתקיים שהשונות שלו קטנה יותר: $V(\hat{\vartheta}_1)<V(\hat{\vartheta}_2)$.

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

כיצד מחשבים שונות של אומד?

חייבות להתקיים הנחות (4) $\mathrm{cov}\left(X_t,u_t\right)=0$, (5) $V\left(u_t\right)=\sigma_u^2$ לכל t ו-(6) $\mathrm{cov}\left(u_t,u_s\right)=0$ לכל $t\neq s$. אם הן מתקיימות, מחשבים את השונות של האיברים המכילים את u_t מהפיתוח הקודם (לפי כללי הסיגמא והשונות).

4. עקיבות:

ככל שהמדגם יגדל כן יתקרב האומד לערך האמיתי של הפרמטר.
אם נגדיל את המדגם לאינסוף תצפיות ונחשב את האומד, הוא יהיה שווה לפרמטר האמיתי באוכלוסייה: $\begin{pmatrix}\hat{\theta}\to\vartheta\end{pmatrix}$ $\left(T\to\infty\right)$.

תנאי הכרחי לעקיבות:
האומד חייב להיות פונקציה של גודל המדגם. במילים אחרות, האומד צריך להיות מושפע מגודל המדגם. ברגע שהאומד עונה על תנאי זה הוא יהיה עקיב. אומד המחושב במדגם סופי בהגדרה לא יוכל להיות עקיב לפרמטר באוכלוסייה.

סיכום: השלבים להוכחת התכונות:

1. הוכחת ליניאריות.
2. הכנת האומד $\Leftarrow$ להציב במקום Y_t את המודל האמיתי.
   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, עקיבות.

- ליניאריות מהווה תנאי הכרחי לחוסר הטיה.
- ליניאריות וחוסר הטיה מהוות תנאי הכרחי לבחינת היעילות של האומד לפי משפט גאוס-מרקוב.
- עקיבות איננה תלויה בתכונות האחרות, אלא רק בהיותו של האומד פונקציה של גודל המדגם (לא מחושב על מדגם סופי). כך שאומד לא חייב להיות ליניארי או חסר הטיה כדי להיות עקיב.
- העקיבות משפיעה על היעילות של האומד. עבור אומדים התלויים בגודל המדגם: ככל שגודל המדגם גדול יותר כך שונות האומד קטנה והאומד יהיה יעיל יותר לפרמטר באוכ'.

שאלות:

תרגול ממבחנים:

1) נתון המודל: $Y_t = \alpha + \beta X_t + u_t$, $T=100$, כאשר מתקיימות כל ההנחות הקלאסיות.

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

א. האומד $\tilde{\beta}$ הינו אומד חסר הטיה ל- β. נכון / לא נכון

ב. האומד $\tilde{\beta}$ הינו אומד עקיב ל- β. נכון / לא נכון

ג. האומד $\tilde{\beta}$ הינו אומד ליניארי ל- β. נכון / לא נכון

ד. האומד $\tilde{\beta}$ הינו אומד יעיל ל- β. נכון / לא נכון

ה. השונות האמיתית של $\tilde{\beta}$ היא?

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

נתון האומד: $\tilde{\beta} = \dfrac{\sum Y_t}{\sum X_t}$.

א. האומד $\tilde{\beta}$ הינו אומד מוטה ל- β: נכון / לא נכון / אי אפשר לדעת

ב. על סמך משפט גאוס-מרקוב ניתן להסיק כי $\tilde{\beta}$ איננו אומד יעיל יותר מאומד הריבועים הפחותים: נכון / לא נכון / לא ניתן לדעת

ג. מהי השונות האמיתית של $\tilde{\beta}$?

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

נתון האומד: $\tilde{\beta} = \dfrac{\sum X_t Y_t}{\sum \left(X_t - \bar{X}\right)^2}$.

א. מהי התוחלת של $\tilde{\beta}$?

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

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

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

4) בכל השאלות ההנחות הקלאסיות מתקיימות.
האומדים הם אר"פ, והמודל הוא: $Y_t = \alpha + \beta X_t + u_t$.

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

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

ג. אמידת המודל בשיטת הריבועים הפחותים תיתן את התוצאה: $\sum_{t=1}^{T} u_t = 0$. נכון / לא נכון / אי אפשר לדעת

ד. אם נתון ש- $r_{XY} = 0.57$, אזי $\hat{\beta}$:
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:
i. $\sum_{t=1}^{T}(X_t - \bar{Y})\hat{u}_t = 0$.
ii. $S_{XX} = \sum_{t=1}^{T} X_t^2 - (T\bar{X})^2$.
iii. $\sum_{t=1}^{T} X_t u_t = 0$.
iv. אף אחת מהטענות הנ"ל אינה נכונה בהכרח.

ו. אומדי הריבועים הפחותים אינם חסרי הטיה, אם נתון שהשונות של u_t אינה קבועה. נכון / לא נכון / אי אפשר לדעת

ז. אומד חסר הטיה הוא אינו בהכרח גם אומד עקיב. נכון / לא נכון / אי אפשר לדעת

תשובות סופיות:

1) א. נכון. ב. לא נכון. ג. נכון. ד. לא נכון.

ה. $V\left(\tilde{\beta}\right)=\dfrac{100\sigma_u^2}{\left(\sum\limits_{t=51}^{100}X_t-\sum\limits_{t=1}^{50}X_t\right)^2}$.

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

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

ד. $\dfrac{\sigma_u^2}{\sum X_t^2}$.

4) א. נכון. ב. לא נכון. ג. לא נכון. ד. ii.

ה. i. ו. לא נכון. ז. נכון.

אקונומטריקה א

פרק 3 - מודלים לא ליניאריים

תוכן העניינים

מודלים לא ליניאריים:

רקע:

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

- המשתנה שיש בו LN השינוי בו יהיה באחוזים.

תזכורת של חוקי לוגים:

$$LN\left(e^{x}\right)=X$$
$$LN\left(X^{Y}\right)=Y\cdot LN\left(X\right)$$
$$LN\left(X\cdot Y\right)=LN\left(X\right)+LN\left(Y\right)$$
$$LN\left(\frac{X}{Y}\right)=LN\left(X\right)-LN\left(Y\right)$$

שאלות:

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

1. $MWAGE_t = 139.547 + 118.628 \cdot SCL_t$.
2. $MWAGE_t = -1445.08 + 1239.60 \cdot LN(SCL)_t$.
3. $LN(MWAGE)_t = 5.244 + 0.778 \cdot LN(SCL)_t$.
4. $LN(MWAGE)_t = 6.292 + 0.070 \cdot SCL_t$.

א. הסבירו את המשמעות של β בכל אחד מהמודלים.

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

2) נתונים תוצאות האמידה של המודלים הבאים:

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

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:

1. $Q_i = AK_i^{\beta_1} e^{u_i}$.
2. $Q_i = Ae^{\beta_1 L_i + u_i}$.
3. $Q_i = A + K_i^{\beta_1} + e^{u_i}$.
4. $Q_i = A + \frac{\beta_1}{L_i} + u_i$.
5. $Q_i = A + \beta_1 \sqrt{K_i} + u_i$.
6. $Q_i = e^{A+\beta_1 Ki + u_i}$.
7. $Q_i = A\left(\frac{K_i}{2} + 7\right)^{\beta_1} e^{u_i}$.
8. $Q_i = A + \beta_1 L_i + u_i$.
9. $Q_i = A + \beta_1 \left(\frac{K_i}{L_i}\right) + u_i$.

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. עבור כל אחד מהמודלים קבעו מיהו המשתנה המוסבר ומיהו המסביר במשוואת הרגרסיה הליניארית.
ד. עקומת אנגי'ל מתארת את גמישות הצריכה של הפרט מוצר מסוים ביחס להכנסתו. איזה מהמודלים מתאים כדי לתאר את עקומת אנגי'ל?

4) נתון המודל הבא: $Q_i = \dfrac{A}{K_i^{\beta_1}} e^{u_i}$.

א. האם ניתן לאמוד את המודל בשיטת OLS?
ב. מה המשוואה שצריך לאמוד על מנת לקבל את הפרמטרים למודל זה (כלומר כיצד הופכים את המודל לליניארי 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 עבור: β_1 , A ?

5) נתון כי הקשר באוכלוסייה בין X ל- Y נתון על ידי המודל הבא: $\ln Y = \alpha + \beta \ln X + u$. נתון גם כי עבור המודל הנ"ל כל ההנחות הקלאסיות מתקיימות.

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

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.1. 0.912. 2. 0.77. 3. 0.778. 4. 0.861.

2) א.1. $\hat{\ln}(Y)=4.5+0.05\cdot\ln(X)$. 2. $\hat{\ln}(Y)=4.5+0.05X$.
3. אין צורך. 4. $\ln\left(\frac{1-\hat{Y}}{\hat{Y}}\right)=4.5+0.05X$.
ב.1. 98.45. 2. 121.51. 3. 4.50833. 4. 0.00816.

3) א. מודלים: 4 , 5 , 8 ו-9.
ב. מודלים: 1 , 2 , 6 ו-7.
ג.1. מסביר: $\ln(K_i)$, מוסבר: $\ln(Q_i)$. 2. מסביר: L_i , מוסבר: $\ln(Q_i)$.
3. אינו ליניארי. 4. מסביר: $\frac{1}{L_i}$, מוסבר: Q_i.
5. מסביר: $\sqrt{K_i}$, מוסבר: Q_i. 6. מסביר: K_i , מוסבר: $\ln(Q_i)$.
7. מסביר: $K_i=\frac{K_i}{2}+7$, מוסבר: $\ln(Q_i)$. 8. מסביר: L_i , מוסבר: Q_i.
9. מסביר: $\frac{K_i}{L_i}$, מוסבר: Q_i.
ד. מודלים: 1 ו-7.

4) א. לא. ב. $\ln(Q_i)=\ln(A)-\beta_1\ln(K_i)+u_i$.
ג. $A=20$, $\beta_1=-0.8$.

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

אקונומטריקה א

פרק 4 - מבחני המובהקות וקריאת פלטים - תוכנת SAS

תוכן העניינים

מבחני המובהקות וקריאת פלטים – תוכנת SAS:

רקע:

פלט ניתוח שונות (Analysis of Variance):

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	k	RSS	$RSS/k = MSR$	$F = \frac{MSR}{MSE}$	PF
	+	+			
Error	$T-k-1$	ESS	$ESS/T-k-1 = MSE$		
C Total	$T-1$	TSS			

Root MSE	$\sqrt{MSE} = s_u$	R-square	$R^2 = \frac{RSS}{TSS}$
Dep Mean	$\overline{Y}$	Adj R-sq	$\overline{R}^2 = 1 - \frac{ESS}{TSS} \cdot \frac{T-1}{T-k-1}$
C.V.	$\frac{s_u}{\overline{Y}} \cdot 100$		

פלט מקדמי הרגרסיה (Parameter Estimates):

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
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	$s_{\hat{\alpha}}^2$	$\operatorname{cov}(\hat{\alpha},\hat{\beta})$
X	$\operatorname{cov}(\hat{\alpha},\hat{\beta})$	$s_{\hat{\beta}}^2$

עריכת תחזית וקריאת פלטים (תוכנת SPSS):

אמידה נקודתית:

אמידה נקודתית עבור X_0 מסוים (תחזית).

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

אמידה מרווח ל- $E(Y)$:

אמידה התחזית באוכלוסייה עבור X_0 מסוים. נחשב רווח בר סמך לערך ממוצע של Y באוכ' עבור X_0 מסוים ($E(Y)$) ברמת סמך $1-\alpha$.

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

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

רישום הרב"ס: $p(\_\_\_ \le E(Y) \le \_\_\_) = 1-\alpha$.

אמידה מרווח ל- Y :

אמידה ערך בודד של Y באוכלוסייה עבור X_0 מסוים. נחשב רווח בר סמך לערך בודד של Y באוכ' עבור X_0 מסוים (Y_0) ברמת סמך $1-\alpha$.

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

רישום הרב"ס: $p(\_\_\_ \le Y \le \_\_\_) = 1-\alpha$.

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

שאלות:

פלט ניתוח שונות:

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

Model: MODEL1

Dependent Variable: TAX

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
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 הרגרסיה:

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

Parameter Estimates

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

א. אמדו 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 עדיין תידחה השערת האפס מסעיף ב'?

ד. בדקו את ההשערה כי ככל שההכנסה עולה כך עולה גם המס (שיפוע β חיובי) ברמת מובהקות של 0.01.

ה. בנו רווח-סמך ברמת סמך של 95% עבור β.

ו. בדקו את ההשערה שתוספת של מיליארד $ להכנסה תגדיל את המס ב-0.2 מיליארד $, ברמת מובהקות של 0.05.

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

$$F_{(1,T-2;1-\alpha)} = t^2_{\left(T-2;1-\frac{\alpha}{2}\right)}$$

$$F = t^2_{\hat{\beta}}$$

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

פלט שונויות משותפות:

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

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: 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 מסכמת:

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

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
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 Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
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 נוסף: $EXP=22$.

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 תחזיות:

5) במדגם של 30 דירות מושכרות לסטודנטים ברדיוס של עד 2 ק"מ מסביב למכללה נחקר הקשר בין שכר דירה למספר הסטודנטים הגרים 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

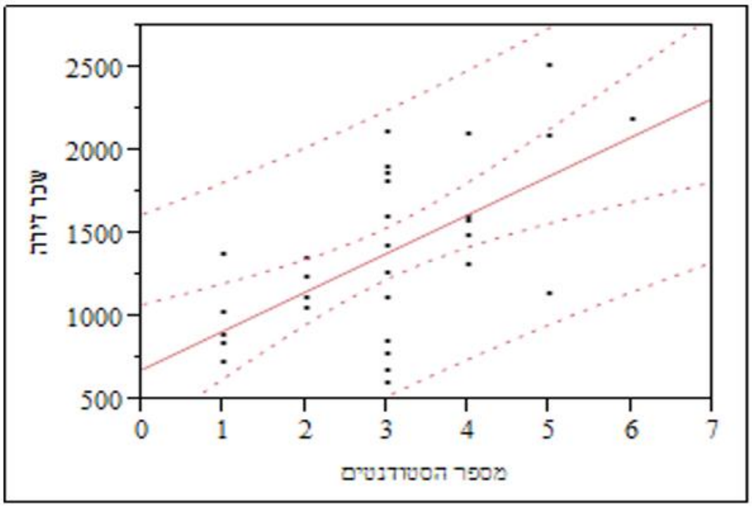
ANOVA[b]

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. אמוד את שכר הדירה הממוצע שישלמו סטודנטים החולקים את הדירה עם שותף אחד בלבד, ברמת בטחון של 95%.

ג. אמוד את שכר הדירה שישלם סטודנט יחיד החולק את הדירה עם שותף אחד בלבד, ברמת ביטחון של 95%.

תשובות סופיות:

1) יש עדות לכך.

2) א. ראה סרטון. ב. יש עדות לכך. ג. $Pt_{\hat{\beta}}=0.0001$.

ד. יש עדות לכך. ה. $P\left(0.1488\le\beta\le0.1554\right)=0.95$.

ו. יש עדות לכך.

3) אין עדות לכך.

4) א. לא נכון. ב. -0.87%. ג. 7.24735.

5) א. 1153.247. ב. $p\left(957.4\le\mu_{Y_{X=2}}\le1349.08\right)=0.95$.

ג. $p\left(282.94\le Y_{X=2}\le2023.55\right)=0.95$.

אקונומטריקה א

פרק 5 - שינוי יחידות מדידה

תוכן העניינים

שינוי יחידות מדידה:

רקע:

טרנספורמציה ליניארית: הוספה/החסרה של קבוע ו/או הכפלה/חילוק של קבוע של אחד או שני המשתנים (התלוי והב"ת).

- טרנספורמציה ליניארית של המשתנים לא תשפיע על: R^2, F, $t_{\hat{\beta}}$ 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.
- $t_{(\hat{\alpha}'=0)} = t_{(\hat{\alpha}=0)}$ רק בהכפלות.

שאלות:

1) חוקר ביקש לאמוד את הקשר בין שכר ב-₪ ($MWAGE$) לבין שנות לימוד (SCL) באמצעות 2 מודלים שונים.
להלן תוצאות האמידה:

א. $MWAGE_t = 139.54 + 118.62 \cdot SCL_t$.

ב. $MWAGE_t = -1445.08 + 1239.60 \cdot LN\left(SCL\right)_t$.

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

1. התברר כי נעשה 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 המשתנה המקורי של מספר שנות הלימוד בריבוע.

2) בהמשך לנתוני השאלה לדוגמא מהפרק החמישי:
החוקר 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%).
המודל הוא: $\ln\left(NET_t\right) = \alpha' + \beta' \cdot EXP_t + v_t$.
מה יהיו ערכי האומדים, סטיות התקן שלהם וטיב ההתאמה באמידת מודל זה?

תשובות סופיות:

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

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

2. לא ניתן לדעת.

2) $S_{\hat{\alpha}'} = S_{\hat{\alpha}} = 0.0688935$, $S_{\hat{\beta}'} = S_{\hat{\beta}} = 0.0026235$, $\hat{\alpha}' = 7.11161$, $\hat{\beta}' = \hat{\beta} = -0.00874$,
$R^2 = 0.0269$.

אקונומטריקה א

פרק 6 - הרגרסיה המרובה

תוכן העניינים

רגרסיה מרובה:

רקע:

מבחן T ו- F:

כאשר יש יותר ממשתנה מסביר אחד, מדובר ברגרסיה מרובה.

המודל הקלאסי: $Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \ldots + \beta_k X_{kt} + u_t$.

- קבוע α יש אחד.
- מספר ה- β טות כמספר המשתנים הב״ת במודל.

מבחן F למובהקות המודל:

השערות: $\begin{matrix} H_0 = \beta_1 = \beta_2 = \ldots = \beta_K = 0 \\ H_1 : OTHERWISE \end{matrix}$.

סטטיסטי המבחן F וכלל ההכרעה:

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

מבחן t למובהקות ה- β טות:

מבחן לבדיקת מובהקות β ספציפית:

השערות: $\begin{matrix} H_0 = \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{matrix}$.

סטטיסטי המבחן t וכלל ההכרעה:

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

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

בכדי להחליט האם כדאי לנו להוסיף למודל משתנה ב״ת מסוים:
נשווה את פרופורציית השונות המוסברת המתוקנת $\bar{R}^2$ בין המודל ללא המשתנה המסביר לבין המודל עם המשתנה המסביר שהוספנו.

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

לפי חוק חיטובסקי – בהוספת משתנה מסביר אחד בלבד למודל ה- $\bar{R}^2$ יעלה 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 (מובהק).

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

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

שאלות:

מבחן T ו-F:

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 Squares	Mean 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 Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
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. השלם 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 המשרה. במבחן למובהקות המשתנה הנוסף התקבל: $t_{\hat{\beta}} = 0.456$. האם ערך $\bar{R}^2$ יעלה/ירד/לא ישתנה בהוספת המשתנה הנוסף למודל?

מבחן 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 Squares	Mean 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 Estimate	Standard Error	T for H0: 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

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

א. מהי השערת האפס?

ב. מהו המודל המוגבל שאותו צריך לאמוד?

להלן אמידת המודל המוגבל:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
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 Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
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?

4) על מנת לאמוד את פונקציית התצרוכת נאספו נתונים על 42 משקי בית בשנת 2007 ונאמדה המשוואה הבאה: $C_t = \alpha + \beta_1 \cdot W_t + \beta_2 \cdot P_t + u_t$.
להלן תוצאות האמידה של המשוואה הנ"ל:

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 שהנטייה השולית לצרוך מתוך ההכנסה זהה לנטייה השולית לצרוך מתוך ההון, נאמדה גם המשוואה הבאה:
$C_t = \alpha + \beta_1 \cdot Y_t + u_t$, כאשר: Y_t = סה"כ ההכנסה של משק בית t.
התקבל: $ESS = 0.4566$.
בדקו את ההשערה בשתי דרכים.

תרגיל מסכם:

5) חוקר אמד את התצרוכת של 500 משקי בית כפונקציה של הכנסה שלהן לפי המשוואה: $EXPENSE_t = \alpha + \beta \cdot INCOME_t + u_t$.

$EXPENSE_t$ - התצרוכת של משק הבית ה-t-י באלפי שקלים.

$INCOME_t$ - ההכנסה של משק הבית ה-t-י באלפי שקלים.

ההפרעות האקראיות מקיימות את כל ההנחות הקלאסיות התקבל הפלט הבא:

Analysis of Variance

Source	DF	Sum of Squares	Mean 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 Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	0.041995	0.054951	0.764236	0.4451
INCOME	1	0.713503	0.008853	80.59618	0.0000

א. מהו Pvalue לבדיקת מובהקות המודל ע"י מבחן F?

ב. מהו אחוז השונות בתצרוכת המוסבר ע"י ההכנסה?

ג. מהו אומדן לתצרוכת ההתחלתית של משק בית?

ד. האם אומדן זה מובהק?

ה. על עוזר מחקר הטיל החוקר לבדוק את ההשערה כי על כל 1000 ₪ נוספים בהכנסה צורך הפרט 700 ₪, כנגד ההשערה כי הוא צורך יותר מ-700 ₪. נסח את השערת האפס ואת ההשערה האלטרנטיבית.

ו. מהו הסטטיסטי t לבדיקת ההשערה?

ז. מהו הסטטיסטי WALD לבדיקת ההשערה?

ח. התברר כי הייתה טעות בנתונים, וכי יש להוסיף 1000 ₪ לתצרוכת של כל משק בית:

i. ההוספה תגדיל את האומד ל- α: נכון / לא נכון / לא ניתן לדעת.

ii. בעקבות ההוספה האומד ל- α יהיה מובהק: נכון / לא נכון / לא ניתן לדעת.

iii. ההוספה תשנה את האומד ל- β: נכון / לא נכון / לא ניתן לדעת.

iv. ההוספה תשנה את R^2: נכון / לא נכון / לא ניתן לדעת.

החוקר טען כי יש להוסיף לפונקציית התצרוכת גם את השפעת העושר. העושר של משק בית מורכב מתוכניות החסכון שלו ($SAVINGS$) ומניירות 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 הטעויות 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"צ מתוך ההכנסה שווה ל-0.6 וכי השפעת ניירות הערך על התצרוכת היא פי 2 מהשפעת תוכניות החסכון.

יא. מהי השערת האפס לבדיקה זו?

יב. המודל המוגבל לבדיקת ההשערה יהיה מהצורה: $Z_t = \gamma_0 + \gamma_1 W_t + v_t$.
בטא את Z_t, ו- W_t באמצעות המשתנים המקוריים.

תשובות סופיות:

1) א.

Analysis of Variance

Source	DF	Sum of Squares	Mean 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 Estimate	Standard Error	T for H0: 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

ב. יש עדות לכך. ג. יש עדות לכך.

2) ירד.

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

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

ה. מקבלים.

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

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

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

ח. i. נכון. ii. נכון. iii. לא נכון. iv. לא נכון.

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

יא. $H_0:\beta_3=2\cdot\beta_2\ ,\ \beta_1=0.6$.

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

אקונומטריקה א

פרק 7 - רגרסיה לוגיסטית

תוכן העניינים

רגרסיה לוגיסטית:

רקע:

מתי נבצע רגרסיה לוגיסטית?

כאשר המשתנה המנובא הוא דיכוטומי (Binary Logistic):
יכול לקבל ערכים של 0 או 1.
הפונקציה הלוגיסטית מתארת את הסיכויים לקבל " 1" במשתנה התלוי כתלות במשתנים הב"ת.

הלוגיקה בניתוח רגרסיה לוגיסטית:

השוואת ניבוי Y ללא המשתנים המנבאים במודל לניבוי Y במודל הכולל את המשתנים המנבאים (סטטיסטי $\chi 2$).

טיב מודל הרגרסיה ("Goodness of fit"):

1. מובהקות המודל:

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	12.225	4	.016
	Block	12.225	4	.016
	Model	12.225	4	.016

מבחן $\chi 2$ - תחת שורת ה-model נמצא את חי בריבוע ואת מובהקות המודל.

2. אחוז שונות מוסברת:

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	96.524	.139	.189

Nagelkerke R^2 – מקביל ל- R^2 כללי ברגרסיה. אחוז שונות Y המוסברת ע"י כל המנבאים יחד (בטווח מוכר של 0-1).

3. דיוק בניבוי:

Classification Table[a]

Observed			Predicted: whether mom believes course will help: no	Predicted: whether mom believes course will help: yes	Percentage Correct
Step 1	whether mom believes course will help	no	46	5	90.2
		yes	17	14	45.2
	Overall Percentage				73.2

a. The cut value is .500

סגוליות (true negative) – ביחס ל- Y =0 במדגם, כמה המודל דייק בניבוי (90.2%).

רגישות (true positive) – ביחס ל- Y=1 במדגם, כמה המודל דייק בניבוי (45.2%).

אחוז הניבוי הכללי – בכמה בסה״כ המודל מדייק בניבוי (73.2%).

מושגים חשובים להבנת טבלת המקדמים:

ODDS:

״הסיכוי להתרחשות אירוע מסוים״- ההסתברות שהאירוע יקרה לעומת ההסתברות שאותו אירוע לא יקרה: $ODDS = \frac{p}{1-p}$

ODDS=1 – הסיכוי שהאירוע יתרחש <u>שווה</u> לסיכוי שהוא לא יתרחש ($\frac{0.5}{0.5}$).

ODDS>1 – הסיכוי שהאירוע יתרחש <u>גבוה</u> מהסיכוי שלא יתרחש (למשל- $\frac{0.75}{0.25}$).

ODDS<1 – הסיכוי שהאירוע יתרחש <u>נמוך</u> מהסיכוי שלא יתרחש (למשל- $\frac{0.25}{0.75}$).

ODDS RATIO (OR):

יחס בין סיכויים - $OR = \frac{ODDS_{(A)}}{ODDS_{(B)}}$

כיצד משתנה ההסתברות במעבר מקבוצה A לקבוצה B.

OR=1 – הסיכוי להתרחשות האירוע <u>שווה</u> בין שתי הקבוצות- אין קשר בין המב״ת למ״ת.

OR>1 – הסיכוי להתרחשות האירוע בקבוצה A <u>גבוה</u> מאשר בקבוצה B – קשר חיובי.

OR<1 – הסיכוי להתרחשות האירוע בקבוצה A <u>נמוך</u> מאשר בקבוצה B – קשר שלילי.

טבלת המקדמים – תרומות ייחודיות של כל מנבא:

(מקביל לטבלת Coefficients ברגרסיות ליניאריות)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	EDU_YRS	-.107	.138	.603	1	.438	.898
	AGE	-.029	.020	2.078	1	.149	.971
	SATISFAC	.118	.175	.457	1	.499	1.126
	BIRTH#	.882	.321	7.530	1	.006	2.415
	Constant	.001	1.796	.000	1	.999	1.001

a. Variable(s) entered on step 1: EDU_YRS, AGE, SATISFAC, BIRTH#.

1. מבחן WALD למובהקות המשתנים:
   מבטא את מובהקות המשתנה מבחינת תרומתו הייחודית לניבוי Y.
2. B – מקדמי המשתנים ב-log odds:
   בטא חיובית – עלייה ב-log odds של Y כפונקציה של עליה ביחידה אחת של X.
   בטא שלילית – ירידה ב-log odds של Y כפונקציה של עליה ביחידה אחת של X.
3. משוואת הרגרסיה:

$$\log\left(\frac{p}{1-p}\right) = \widehat{\alpha} + \widehat{\beta_1}\, x_{1i} + \widehat{\beta_2} x_{2i} + \widehat{\beta_3} x_{3i} + \widehat{\beta_4} x_{4i}$$

חישוב הניבוי במונחי הסתברות (p): $p = \frac{1}{1+e^{-\log odds}}$

חישוב הניבוי במונחי סיכויים (ODDS): $ODDS = e^{\log odds}$

4. Exp (B) - יחס הסיכויים (Odds Ratio):
   מבטא את <u>העלייה</u> (אם גדול מ-1) או את <u>הירידה</u> (אם קטן מ-1) בסיכויים להיות בעלי ערך '1' ב-Y כאשר הערך במשתנה המנבא גדל ביחידה אחת.
5. <u>היחס בין B ל-Exp (B)</u>: $e^{B} = Exp(B)$; $\log Exp(B) = B$

שאלות:

1) חוקרת בחוג למגדר ביקשה לבדוק האם מגדר משפיע על תעסוקה.
היא התבססה על סקר של הלמ"ס שדגם 826 מבוגרים בגילאי העבודה המרכזיים (25-55).
היא הגדירה את המשתנים באופן הבא:
WOMEN - "1" = אישה ; "0" = גבר.
WORKING - "1"= כן ; "0" = לא.
מהצלבה של שני המשתנים התקבלה הטבלה הבאה:

working * women Crosstabulation
Count

		women		Total
		.00	1.00	
working	.00	13	130	143
	1.00	338	345	683
Total		351	475	826

על סמך הטבלה חשבו:
א. מה ההסתברות של אישה לעבוד?
ב. מה הסיכוי של אישה לעבוד?
ג. מה ההסתברות של גבר לעבוד?
ד. מה הסיכוי של גבר לעבוד?
ה. מה יחס הסיכויים (OR) של נשים לעבוד לעומת גברים?
ו. מה הלוגריתם של יחס הסיכויים?
ז. מה יהיה ערך מקדם השיפוע B ברגרסיה הלוגיסטית לניבוי תעסוקה על פי מגדר ומה משמעותו?
ח. מה יהיה ערך Exp(B) ברגרסיה הלוגיסטית ומה משמעותו?

2) במחקר ביקשו לבדוק כיצד מצב משפחתי וגובה המשכורת משפיעים על בעלות על דירה.

משתני המחקר:

apartm - בעלות על דירה: "1"- כן; "0"- לא.

status - מצב משפחתי: status (0) - רווק ; status (1) - בזוגיות ;

status (2) - בזוגיות עם ילדים ; status(3) - פרוד או גרוש.

incom - הכנסה (בעשרות אלפי שקלים).

התקבלו הממצאים הבאים:

Classification Table[a]

Observed			Predicted apartm .00	Predicted apartm 1.00	Percentage Correct
Step 1	apartm	.00	22	11	66.7
		1.00	10	22	68.8
	Overall Percentage				67.7

a. The cut value is .500

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	10.218	4	.037
	Block	10.218	4	.037
	Model	10.218	4	.037

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	79.876[a]	.145	.194

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Status			.682	3	.877	
	Status(1)	-.498	.713	.487	1	.485	.608
	Status(2)	-.520	.784	.441	1	.507	.594
	Status(3)	-.180	.748	.058	1	.810	.835
	income	.000	.000	8.580	1	.003	2.536
	Constant	-2.734	1.079	6.417	1	.011	.065

a. Variable(s) entered on step 1: Status, income.

א. האם ניתן לדחות את השערת האפס הטוענת כי אין קשר בין בעלות על דירה להכנסה ולסטטוס משפחתי?
ב. כמה אחוזים מצליחים המשתנים הב"ת להסביר מהשונות של המשתנה "בעלות על דירה"?
ג. באיזה אחוז מצליח המודל לנבא באופן מדויק בעלות על דירה מתוך כלל המקרים?
ד. באיזה מידה מצליח המודל לנבא בהצלחה בעלות על דירה מתוך בעלי הדירה במדגם? כיצד נקרא המדד המתאים?
ה. באיזה מידה מצליח המודל לנבא בהצלחה אי-בעלות על דירה מתוך אלו שאינם בעלי דירה במדגם? כיצד נקרא המדד המתאים?
ו. מהי המשוואה לניבוי בעלות על דירה על סמך המשתנים הב"ת?
ז. לאיזה מהמשתנים הב"ת יש תרומה ייחודית מובהקת לניבוי בעלות על דירה? מהי משמעות מקדם B ו-Exp(B) של משתנה זה?
ח. על כל עליה ב-10,000 ₪ בהכנסה, בכמה אחוזים יעלה הסיכוי לבעלות על דירה?

i. 53.6%.

ii. 253.6%.

iii. 153.6%.

iv. 93%.

ט. על כל עליה של 20,000 ₪ בהכנסה, בכמה אחוזים יעלה הסיכוי לבעלות על דירה?

i. 307%.

ii. 423%.

iii. 542%.

iv. 642%.

י. מה ההסתברות של רווק המשתכר 20,000 ₪ להיות בעלים של דירה?
יא. האם ההסתברות של אותו רווק להיות בעל דירה גבוהה / שווה / קטנה מההסתברות שלו לא להיות בעל דירה?
יב. מהם הסיכויים (ODDS) שלו להיות בעל דירה?
יג. עבור איזה משכורת הסיכוי (הסתברות) של רווק להיות בעל דירה עולה על הסיכוי שלו לא להיות בעל דירה?
יד. במידה ומשתנה ההכנסה היה נמדד באלפי שקלים (ולא בעשרות אלפי שקלים), כיצד הדבר היה משפיע על ההשפעה השולית של מקדם ההכנסה, אם בכלל?

3) חוקרים בחנו את המאפיינים שעשויים לנבא את הביצוע של חניכים במבחן הסיום של קורס פקחי טיסה.
הביצוע במבחן נמדד על סולם של הצלחה/כשלון והמשתנים הבלתי תלויים כללו מין (1-זכר 0-נקבה), השכלה קודמת (0- ריאלית, 1- לא ריאלית) וביצוע במהלך הקורס (1-7).
להלן תוצאות ניתוח הרגרסיה:

Omnibus Tests of Model Coefficients

		Chi-square	Df	Sig.
Step 1	Step	20.982	3	.000
	Block	20.982	3	.000
	Model	20.982	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	17.209[a]	.503	.699

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	מין	4.445	2.611	2.897	1	.089	85.161
	השכלה קודמת	-.146	2.054	.005	1	.943	.864
	ביצוע במהלך הקורס	2.283	.944	5.846	1	.016	9.810
	Constant	-19.284	8.056	5.731	1	.017	.000

a. Variable(s) entered on step 1: מין, השכלה, ביצוע במהלך הקורס

א. האם למודל הכולל את שלושת המנבאים יכולת הסבר משמעותית?
ב. כמה אחוזים מתוך השונות של Y מצליח המודל להסביר?
ג. מהי משוואת הניבוי?
ד. לאיזה מן המשתנים הב״ת תרומה מובהקת לניבוי?
ה. הסבירו את משמעות המקדמים (b) שהתקבלו עבור המשתנים הב״ת: מגדר, השכלה קודמת והביצוע במהלך הקורס.
ו. בטאו את המקדמים במונחי הסיכויים להצלחה בקורס (odds) והסבירו אותם.
ז. הועלתה הטענה כי ההסתברות ההצלחה של נשים בקורס היא נמוכה ביותר, גם אם הן בעלות השכלה ריאלית ושביצוען במהלך הקורס מקסימאלי. אנא בדקו את הטענה.
ח. עבור זכר, בעל השכלה ריאלית, מהי ההשפעה השולית של עליה ביחידה אחת בדירוג הביצוע במהלך הקורס על הסיכוי להצליח בקורס?

4) לפי מדגם של 20 זוגות נשואים, נאספו נתונים על המשתנה Y השווה ל-1 אם הזוג נוהג לצאת למסעדה לפחות פעם בשבוע ו-0 אחרת.

נאמד המודל : $p=\frac{1}{1+e^{-z}}$ כאשר $p=P(Y=1)$.

התקבלו התוצאות הבאות : $z=-9.456+0.368INCOM-1.207BABY$.

INCOM - ההכנסה של שני בני הזוג (באלפים). ההכנסה במדגם נעה בין 17 אלף ל- 44 אלף.

BABY - משתנה דמי המקבל את הערך '1' אם הזוג צריך להיעזר בשמרטפית ו-'0' אחרת.

ענה נכון/לא נכון :

א. זוג הנעזר בשמרטפית ומשתכר 30.5 אלף, יוצא למסעדה לפחות פעם בשבוע בהסתברות גבוהה מ-0.5.

ב. עבור זוג שאינו נעזר בשמרטפית, עליה של אלף שח בהכנסה, מעלה את ההסתברות לצאת למסעדה ב-0.368.

ג. כל אחד מערכי P הנאמדים כאן איננו גבוה יותר מ-0.99.

ד. הסיכוי של זוג, שהכנסתו עלתה ב-3000 ₪, לצאת למסעדה יעלה ב-200% בערך.

ה. המשכורת שצריך להרוויח זוג, אשר אינו נעזר בשמרטפית, כדי שהסיכוי שלו לצאת למסעדה יהיה שווה לסיכוי שלא לצאת למסעדה הוא 27,000.

ו. זוג, שלא נעזר בשמרטפית, צריך להרוויח יותר מ-28,000 ₪ כדי שהסיכוי שלו לצאת למסעדה יהיה גבוה פי 3 מהסיכוי שלו לא לצאת למסעדה.

ז. עבור odds ratio של משתנה "שמרטפות" התקבל רווח בר סמך הבא : $[0.123\ ;\ 1.01]$ ברמת ביטחון של 95%.

לפיכך ניתן לומר כי למשתנה "שמרטפות" תרומה מובהקת לניבוי הסיכוי לצאת למסעדה.

5) בשנה מסוימת הוגשו 750 בקשות לקבלת משכנתא ורק חלק מהן אושר.

המשתנה התלוי Y=1 אם הבקשה למשכנתא אושרה ול-0 אם נדחתה.

המנבאים :

S משתנה דמי השווה ל-1 אם מבקש המשכנתא הוא רווק ול-0 אחרת.

AGE = גיל בשנים.

המודל הנאמד הינו : $p=\frac{1}{1+e^{-z}}$ כאשר $p=P(Y=1)$.

$z=\alpha+\beta_1 age+\beta_2 age^2+\beta_3 S$.

תוצאות אמידת המודל : $z=-9.3+0.52age-0.006age^2-0.314S$

א. הסבירו את השפעת הגיל והמצב המשפחתי על ההסתברות לאישור המשכנתא.

ב. מה ההסתברות שתאושר משכנתא לרווק בן 30?

ג. עבור איזה גיל ההסתברות של אדם נשוי לקבל משכנתא היא מקסימאלית?

6) משרד הקבלה של האוניברסיטה רצה לבדוק באיזה מידה ניתן לחזות את ההצלחה של הסטודנט בקורס בסטטיסטיקה על סמך נתונים של מבחן פסיכומטרי, ציון ממוצע של תעודת בגרות וסוג תעודת הבגרות: ריאלית או לא ריאלית.
במדגם של 50 סטודנטים נאספו נתונים על המשתנה Y השווה ל-1 אם הסטודנט הצליח במבחן בסטטיסטיקה ו-0 אם נכשל.
כמו כן נרשמו עבור כל סטודנט ציון הפסיכומטרי, ממוצע הבגרות וסוג הבגרות (1 - בגרות ריאלית, 0 - לא ריאלית).
להלן התוצאות שהתקבלו:

	B	S.E.	Wald	df	Sig.
פסיכומטרי	.090	.046	3.723	1	.054
ציון בגרות		2.070	1.089	1	.297
בגרות_ריאלי	4.535	2.519	3.241	1	.072
Constant	-84.892	42.858	3.923	1	.048

א. באיזה שיטת ניתוח הייתם ממליצים להשתמש ומדוע?
ב. נתון כי ההסתברות להצליח בקורס בסטטיסטיקה עבור סטודנט שעשה בגרות הומנית, קיבל 690 בפסיכומטרי וציון 9 בבגרות הינה: 0.034.
ההסתברות של סטודנט שקיבל אותו ציון בפסיכומטרי, עם בגרות הומנית אבל ציונו בבגרות הוא 10 הינה: 0.233
על סמך הנתונים הללו השלם את הערך החסר בפלט המקדמים.
ג. לאיזה משתנים השפעה מובהקת על הסיכוי להצליח במבחן לסטטיסטיקה? (אלפא 10%)
ד. מה ההסתברות של סטודנט להצליח במבחן אם קיבל 680 בפסיכומטרי, ציון 10 בבגרות ולמד במגמה ריאלית?
ה. מהו השינוי בסיכויים (odds) להצליח במבחן בסטטיסטיקה כפונקציה של שינוי ביחידה אחת בפסיכומטרי?
ו. מהי ההשפעה השולית של נקודה נוספת בציון הבגרות על הסיכוי להצליח במבחן בסטטיסטיקה עבור סטודנט שקיבל 640 בפסיכומטרי ולמד במגמה ריאלית?
ז. רותי שיפרה את הפסיכומטרי שלה ב-20 נקודות.
בכמה יעלה הסיכוי שלה להצליח בקורס בסטטיסטיקה?
ח. אם החוקר היה מחליט לקודד בגרות שאינה ריאלית כ-1 ובגרות ריאלית כ-0, האם הדבר היה משפיע על ערכו של Exp(b) של סוג בגרות ועל המשמעות שלו?

תשובות סופיות:

1) א. 0.73. ב. 2.7. ג. 0.96. ד. 24. ה. 0.11.
ו. -2.207. ז. $B=-2.207$. ח. $\text{Exp}(B)=0.11$.

2) א. כן. ב. 19.4%. ג. 67.7%. ד. רגישות = 68.8%.
ה. סגוליות = 66.7%.
ו. $In(odds)=-2.734-0.498status(1)-0.52status(2)-0.18status(3)+0.93\cdot incom$.
ז. משתנה "הכנסה". ח. 3. ט. 3. י. 0.3.
יא. קטנה. יב. 0.42. יג. 29,400. יד. 0.093.

3) א. כן. ב. 69.9%. ג. $\log(odds)=-19.284+4.445x_{1i}-0.146x_{2i}+2.283x_{3i}$.
ד. המשתנה – "ביצוע במהלך הקורס". ה. ראו סרטון.
ו. מגדר - $\text{Exp}(b)=85.19$, השכלה קודמת - $\text{Exp}(b)=0.864$,
ביצוע במהלך הקורס - $\text{Exp}(b)=9.81$. ז. הטענה נכונה $(p=0.035)$.
ח. $\text{Exp}(b)=9.81$.

4) א. נכון. ב. לא נכון. ג. לא נכון. ד. נכון. ה. לא נכון.
ו. נכון. ז. לא נכון.

5) א. ראו סרטון. ב. 0.533. ג. 40.

6) א. רגרסיה לוגיסטית. ב. $B=2.16$. ג. "פסיכומטרי" ו-"בגרות ריאלית".
ד. 0.914. ה. 1.09. ו. 8.67. ז. 504%. ח. ראו סרטון.

אקונומטריקה א

פרק 8 - מבחן 1

תוכן העניינים

מבחן 1:

שאלות:

1) חוקר רצה לבדוק את השפעת התל״ג על ההשקעה במשק לפי המודל הבא: $\ln I_t = \alpha + \beta \ln Y_t + u_t$, כאשר: I_t היא ההשקעה באלפי שקלים, Y_t הוא התוצר באלפי שקלים, וההרעה האקראית, u_t, מקיימת את כל ההנחות הקלאסיות. באמידה התקבל הפלט הבא:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	0.38523	0.38523	72.14	<.0001
Error	199	1.06266	0.00534		
C Total	200	1.44789			

Root MSE	0.073075	R-square	0.733936
Dep Mean	10.01722	Adj R-sq	0.732104
C.V.	0.729494		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|	95% conf. lim.
INTERCEPT	1	3.472013	0.85463	4.06259	0.0002	1.79 – 5.15
lnY	1	0.570042	0.06452	8.493526	0.0000	---- – ----

א. מהו Pvalue לבדיקת מובהקות המודל ע״י מבחן F?
ב. אם נגדיל את התוצר ב-1% בכמה תגדל ההשקעה?
ג. מהו רווח הסמך ל- α? מהו רווח הסמך ל- β?
ד. הועלתה הטענה כי הגמישות שווה ל-0.4. מהן ההשערות לבדיקת הטענה?
ה. מהי הרגרסיה המוגבלת למבחן WALD תחת H0?
ו. מהו הסטטיסטי של WALD למבחן זה (אם ניתן לחישוב)?
ז. אם ההשקעה נמדדת בשקלים במקום באלפי שקלים:
i. המקדם של $\ln Y$ לא ישתנה. נכון / לא נכון / לא ניתן לדעת
ii. החותך לא ישתנה. נכון / לא נכון / לא ניתן לדעת

iii. הסטטיסטי t לבדיקת המובהקות של β לא ישתנה. נכון / לא נכון / לא ניתן לדעת

iv. הסטטיסטי F לבדיקת מובהקות המודל לא ישתנה. נכון / לא נכון / לא ניתן לדעת

v. R^2 לא ישתנה. נכון / לא נכון / לא ניתן לדעת

החוקר טען כי גם גודל האוכלוסייה, P, משפיע על ההשקעה לפי המודל הבא: $\ln I_t = \alpha + \beta_1 \ln Y_t + \beta_2 \ln P_t + u_t$.

ח. מהי השערת האפס לבדיקת הטענה?

התקבל הפלט הבא:

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T
INTERCEPT	1	1.131853	1.43547	0.788489	0.4435
lnY	1	1.035467	0.25756	4.020294	0.0004
lnP	1	-1.77456	0.94657	-1.874727	0.0736

ט. באיזו רמת מובהקות נקבל את טענת החוקר?

י. R^2 של המשוואה החדשה קטן מזה של המשוואה המקורית. נכון / לא נכון / לא ניתן לדעת

במשוואה החדשה הועלתה הטענה כי סכום הגמישויות שווה ל-0.

יא. מהי השערת האפס לבדיקת הטענה?

יב. מהו הסטטיסטי t לבדיקת ההשערה? (נתון כי: $\text{cov}\left(\hat{\beta}_1, \hat{\beta}_2\right) = -0.25$).

יג. האם ניתן לדחות את השערת האפס?

2) ענה על הסעיפים הבאים:

א. ברגרסיה מרובה, כמו ברגרסיה חד משתנית, מבחן F למובהקות המודל שווה לריבוע של מבחן t למובהקות של β. נכון / לא נכון / לא ניתן לדעת

ב. אם הערך 0 נמצא בתוך רווח הסמך ל- β, אזי β מובהקת. נכון / לא נכון / לא ניתן לדעת

ג. בהוספת משתנה לא רלוונטי למודל האומד המתוקן לפרופורציית השונות המוסברת ירד בהכרח. נכון / לא נכון / לא ניתן לדעת

ד. אומדי הריבועים הפחותים אינם חסרי הטיה אם ידוע שהשונות של u_t אינה קבועה (הפרה של הנחה קלאסית). נכון / לא נכון / לא ניתן לדעת

ה. אם דוחים H0 ברמת מובהקות מסוימת, אזי דוחים H0 בכל רמות המובהקות הקטנות יותר. נכון / לא נכון / לא ניתן לדעת

ו. אומד חסר הטיה הוא אינו בהכרח אומד עקיב. נכון / לא נכון / לא ניתן לדעת

3) נתון מודל ללא חותך: $Y_t = \beta X_t + u_t$, ונתון האומד: $\tilde{\beta} = \dfrac{\sum_{t=1}^{T} X_t Y_t}{S_{XX}}$.

א. האומד $\tilde{\beta}$ הוא אר"פ. נכון / לא נכון / לא ניתן לדעת

ב. האומד $\tilde{\beta}$ הוא אומד חסר הטיה. נכון / לא נכון / לא ניתן לדעת

ג. האומד $\tilde{\beta}$ הוא אומד לינארי. נכון / לא נכון / לא ניתן לדעת

ד. אר"פ יעיל יותר מ-$\tilde{\beta}$. נכון / לא נכון / לא ניתן לדעת

ה. מהי השונות של $\tilde{\beta}$?

4) נתון מודל ללא חותך: $Y_t = \beta X_t + u_t$, ונתון האומד: $\hat{\beta} = \dfrac{\sum_{t=1}^{T} X_t Y_t}{\sum_{t=1}^{T} X_t^2}$.

א. האומד $\hat{\beta}$ הוא אר"פ. נכון / לא נכון / לא ניתן לדעת

ב. האומד $\hat{\beta}$ הוא אומד חסר הטיה. נכון / לא נכון / לא ניתן לדעת

ג. האומד $\hat{\beta}$ הוא אומד לינארי. נכון / לא נכון / לא ניתן לדעת

ד. מהי השונות של $\hat{\beta}$?

ה. האומד $\hat{\beta}$ הוא אומד עקיב. נכון / לא נכון / לא ניתן לדעת

תשובות סופיות:

1) א. $PF = 0.0001$. ב. 0.57% ג. $p(1.79 \le \alpha \le 5.15) = 0.95$,
$p(0.026 \le \beta \le 1.11) = 0.95$. ד. $\begin{matrix} H_0 : \beta = 0.4 \\ H_1 : \beta \ne 0.4 \end{matrix}$.

ה. $\ln I_t - 0.4 \ln Y_t = \alpha + u_t$. ו. $WALD_{stat} = 7.054$.

ז. i. לא נכון. ii. לא נכון. iii. נכון. iv. נכון. v. נכון.

ח. $H_0 : \beta_2 = 0$. ט. $Pt_{\hat{\beta}} = 0.0736$. י. לא נכון.

יא. $H_0 : \beta_1 + \beta_2 = 0$. יב. $t = -1.089$. יג. אין סיבה מספקת.

2) א. לא נכון. ב. לא נכון. ג. לא נכון. ד. לא נכון. ה. לא נכון.
ו. נכון.

3) א. לא נכון. ב. לא נכון. ג. נכון. ד. לא ניתן לדעת.
ה. $V(\tilde{\beta}) = \dfrac{\sum X_t^2 \sigma^2}{S^2_{XX}}$.

4) א. נכון. ב. נכון. ג. נכון. ד. $V(\tilde{\beta}) = \dfrac{\sigma_u^2}{\sum X_t^2}$.

ה. נכון.

אקונומטריקה א

פרק 9 - מבחן 2

תוכן העניינים

מבחן 2:

שאלות:

1) חוקר בדק את השפעת שעות העבודה בשבוע (HOURS) על השכר החודשי ברוטו בשקלים (SALARY) לפי המודל: $SALARY_t = \alpha + \beta \cdot HOURS_t + u_t$.
הסטייה המקרית מקיימת את כל ההנחות הקלאסיות.
השלם את הפלט הבא, אם ידוע כי: $S_{XX} = 35079$, $\bar{X} = 46.040873$:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	---	---	---	---
Error	401	402271435	---		
C Total	---	449757359			

Root MSE	---	R-square	---
Dep Mean	1580	Adj R-sq	---
C.V.	---		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEPT	1	---	---	---	0.7476
HOURS	1	36.06745	---	---	0.0001

א. מהו Pvalue לבדיקת מובהקות המודל ע"י מבחן F?
ב. מהו האומדן לשכר התחלתי?

החוקר רצה לבדוק את הטענה כי אם יעבוד שעה אחת נוספת בשבוע, שכרו יגדל ב-40 ₪.
ג. מהן ההשערות לבדיקת הטענה?
ד. מהו הסטטיסטי t למבחן?
ה. מהו הסטטיסטי WALD למבחן?
ו. מהי התחזית לשכר של עובד העובד 55 שעות בשבוע?

ז. החוקר טען כי יש לבדוק את הקשר בין השכר לשעות העבודה ע"י שימוש בנתונים שנתיים, כלומר, שכר שנתי (בהנחה שהשכר החודשי קבוע כל השנה) ושעות עבודה שנתיות (בהנחה ששעות העבודה קבועות בכל 52 השבועות בשנה).
שימוש בנתונים שנתיים:
i. ישנה את הסטטיסטי t לבדיקת המובהקות של α. — נכון / לא נכון / לא ניתן לדעת
ii. יכפיל את האומד של β ב-0.23. — נכון / לא נכון / לא ניתן לדעת
iii. יכפיל את סטית התקן של $\hat{\beta}$ ב-0.23. — נכון / לא נכון / לא ניתן לדעת
iv. ישנה את Pvalue לבדיקת מובהקות המודל. — נכון / לא נכון / לא ניתן לדעת

החוקר טען כי יש להוסיף למשוואה גם את השפעת הגיל (AGE) ומספר שנות הלימוד (SCL). לשם כך הוא אמד את המשוואה הבאה:

$$SALARY_t = \alpha + \beta_1 \cdot HOURS_t + \beta_2 \cdot AGE_t + \beta_3 \cdot SCL_t + u_t.$$

ח. מהי השערת האפס לבדיקת הטענה?
ט. מהו הנתון הנדרש כדי לחשב את הסטטיסטי של WALD לבדיקת טענת החוקר?
י. בפלט האמידה של המשוואה החדשה לא היה ברור אם ערכו של נתון זה הוא 315968434 או 515968434 (בשל בעיה במדפסת). מהו הסטטיסטי של WALD לבדיקת טענת החוקר?
יא. מהם הנתונים הנדרשים לחישוב הסטטיסטי t?

החוקר רוצה לבדוק את הטענה כי השפעת ההשכלה על השכר גדולה פי 8 מהשפעת הגיל על השכר.

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEPT	1	-1995.0275	331.7857	-6.013	0.0001
HOURS	1	36.408461	4.710021	7.730	0.0001
AGE	1	13.674254	3.816426	3.583	0.0004
SCL	1	109.93799	10.63745	10.335	0.0001

יב. הנתונים בפלט אינם מספיקים לבדיקת ההשערה לפי מבחן t.
מהו הנתון החסר? באיזה פלט של SAS ניתן למצוא אותו?
יג. בהנחה שנתון זה הוא 8.3969, חשב את הסטטיסטי t לבדיקת הטענה.
מהי מסקנתך לגבי נכונות הטענה?

יד. אם תרצה לבדוק את הטענה לפי מבחן WALD, יהיה המודל המוגבל:

$Z_0 = \gamma_0 + \gamma_1 \cdot Z_1 + \gamma_2 \cdot Z_2 + v$, כאשר:

$Z_0 =$ __________

$Z_1 =$ __________

$Z_2 =$ __________

טו. אם יש מספיק נתונים, חשב את הסטטיסטי של WALD לבדיקת הטענה?

2) נתון המודל: $Y_t = \alpha + \beta \cdot X_t + u_t$. ידוע כי כל ההנחות הקלאסיות מתקיימות.

נתון האומד: $\tilde{\beta} = \dfrac{S_{XY}}{\sum_{t=1}^{T} X_t^2}$.

א. אומד זה הוא הפתרון של המשוואות הנורמליות: $\sum_{t=1}^{T} \hat{u}_t = 0$, $\sum_{t=1}^{T} \hat{u}_t X_t = 0$.   נכון / לא נכון / לא ניתן לדעת

ב. התוחלת של $\tilde{\beta}$ היא:

i. β.

ii. $\dfrac{\beta \cdot \sum_{t=1}^{T} X_t}{S_{XX}}$.

iii. $\dfrac{\beta \cdot \sum_{t=1}^{T} X_t^2}{S_{XX}}$.

iv. $\dfrac{\beta \cdot S_{XX}}{\sum_{t=1}^{T} X_t^2}$.

v. כל התשובות אינן נכונות.

ג. הטענה כי: $E\left(\tilde{\beta}\right) < \beta$:

i. תמיד נכונה.

ii. אינה נכונה.

iii. נכונה אם ורק אם: $\bar{X} > 0$.

iv. נכונה אם ורק אם: $\bar{X} \neq 0$.

v. כל התשובות אינן נכונות.

ד. אם $\bar{X}=0$ אז השונות של $\tilde{\beta}$ היא:

i. $\dfrac{\sigma^2}{\left(\sum_{t=1}^{T} X_t\right)^2}$.

ii. $\dfrac{\sigma^2}{S_{XX}^2}$.

iii. $\dfrac{\sigma^2 \sum_{t=1}^{T} X_t^2}{S_{XX}}$.

iv. $\dfrac{\sigma^2}{S_{XX}}$.

v. כל התשובות אינן נכונות.

ה. אם $\bar{X}=O$, אז $\tilde{\beta}$ הינו האומד הלינארי חסר ההטיה בעל השונות הקטנה ביותר. נכון / לא נכון / לא ניתן לדעת

(3 נתון המודל: $Y_t=\alpha+\beta\cdot X_t+u_t$. ידוע כי כל ההנחות הקלאסיות מתקיימות. נתון כי $\tilde{\beta}$ הוא אומד לינארי וחסר הטיה ל- β, אך איננו אומד עקיב ל- β. מאחר ש- $\tilde{\beta}$ אינו אומד עקיב, לא נוכל להשתמש במשפט גאוס מרקוב ולקבוע כי: $\hat{\beta}=\dfrac{S_{XY}}{S_{XX}}$ (אר"פ) הינו אומד יעיל יותר.

נכון / לא נכון / לא ניתן לדעת

תשובות סופיות:

1) א. $PF = 0.00$. ב. $\hat{\alpha} = -80.5246$. ג. $H_0: \beta = 40$, $H_1: \beta \neq 40$.

ד. $t_{\hat{\beta}} = -0.75$. ה. $WALD_{stat} = 0.5625$. ו. $SALARY_t = 1903.16$.

ז. i. לא נכון. ii. נכון. iii. נכון. iv. לא נכון.

ח. $H_0: \beta_2 = \beta_3 = 0$. ט. SSE. י. $WALD_{stat} = 54.49$.

יא. לא ניתן לחשב. יב. $S^2_{\hat{\beta}_2,\hat{\beta}_3}$, Covariance of Estimates.

יג. נכונה. יד. $Z_0 = SALARY_t$, $Z_1 = HOURS_t$, $Z_2 = AGE_t + 8 \cdot SCL_t$. טו. $WALD_{stat} = 0.000324$.

2) א. לא נכון. ב. ה. ג. ה. ד. ה. ה. לא נכון.

3) לא נכון.

אקונומטריקה א

פרק 10 - מבחן 3

תוכן העניינים

מבחן 3:

שאלות:

(1 על מנת לאמוד את פונקציית הייצור נאספו נתונים על 150 פירמות בשנת 2007 ונאמדה המשוואה הבאה:

.1 $\ln(Y_t) = \alpha + \beta_1 \cdot \ln(L)_t + U_t$.

כאשר:

$\ln(Y)_t$ - תפוקה שנתית באלפי ₪ בלוגים.

$\ln(L)_t$ - מספר העובדים בלוגים.

U_t - הטעות המקרית המקיימת את כל ההנחות הקלאסיות.

משוואה מס' 1 נאמדה בפלט מס' 1.

Dependent Variable: lnY

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	1	8.54211			0.0001
Error	35969	40.42584			
C Total	35970	48.96795			

Root MSE	0.52264	R-square	0.1744
Dep Mean	5.54003	Adj R-sq	0.1689
C. V.	9.43380		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>ITI
INTERCEP	1	4.389949	0.21003743	20.901	0.0001
lnL	1	0.257487	0.04767276		0.0001

א. סטטיסטי F לבדיקת מובהקות המודל:

i. לא ניתן לחשב את סטטיסטי F בעזרת הנתונים הקיימים.

ii. ניתן לחשב וערכו הוא: ________________

ב. סטטיסטי t לבדיקת מובהקות המודל:

i. לא ניתן להשתמש בסטטיסטי t בהשערה מסוג זה

ii. לא ניתן לחשב בעזרת הנתונים הקיימים.

iii. ניתן לחשב וערכו הוא: ________________.

הועלתה הטענה כי עליה ב-1% במסי העובדים תגדיל את התפוקה בפחות מ-1%.

ג. ההשערות לבדיקת הטענה הן : H_0 : ________ H_1 : ________ .

ד. הסטטיסטי לבדיקת הטענה הינו :

i. לא ניתן לחשבו בנתונים הקיימים.

ii. 5.5.

iii. -5.5.

iv. -15.5.

v. 15.5.

ה. הסטטיסטי של WALD לבדיקת הטענה :

i. לא ניתן לחשבו בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא: ________________.

ו. לאור התשובות לסעיפים הקודמים, אחוז התפוקה קטן ככל שאחוז מסי העובדים גדל: נכון/לא נכון/אי אפשר לדעת

החוקרת טענה כי יש משתנים נוספים המסבירים את תפוקת הפירמה ואמדה את המשוואה הבאה :

2. $\ln(Y_t)=\alpha+\beta_1\cdot\ln(L)_t+\beta_2\cdot\ln(K)_t+\beta_3\cdot\ln(PY)_t+U_t$.

כאשר :

$\ln(K)_t$ - מלאי ההון של הפירמה באלפי ₪ בלוגים.

$\ln(PY)_t$ - הוצאות למחקר ופיתוח באלפי ₪ בלוגים.

משוואה מס' (2) נאמדה בפלט מס' 2 .

Dependent Variable: lnY

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	3	15.63370	5.21123	22.825	0.0001
Error	146	33.33425	0.22832		
C Total	149	48.96795			

Root MSE	0.47783	R-square	0.3193
Dep Mean	5.54003	Adj R-sq	0.3053
C. V.	8.62496		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>lTl
INTERCEP	1	0.542062	1.66317350	0.326	0.7450
lnL	1	0.267771	0.08146608	3.287	0.0013
lnK	1	0.405694	0.09700769	4.182	0.0001
lnPY	1	0.406149	0.30781185	1.319	0.1891

ז. ההשערות לבדיקת הטענה הינן: $H_0 : \_\_\_\_\_\_\_\_\_$. $H_1 : \_\_\_\_\_\_\_\_\_$

ח. הסטטיסטי של WALD לבדיקת הטענה הינו:

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים

ii. ניתן לחישוב וערכו: ________________________________.

ט. הסטטיסטי של t לבדיקת הטענה הינו:

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים

ii. לא ניתן לחשב סטטיסטי t לטענה מסוג זה

iii. ניתן לחישוב וערכו: ________________________________.

החוקרת טוענת כי השפעת הוצאות למחקר ופיתוח אינה מובהקת ולכן יש לאמוד את המשוואה הבאה:

3. $\ln(Y_t) = \alpha + \beta_1 \cdot \ln(L)_t + \beta_2 \cdot \ln(K)_t + U_t$.

כאשר:

משוואה מס' (3) נאמדה בפלט מס' 3.

Dependent Variable: lnY

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	2	15.23620	7.61810	33.199	0.0001
Error	147	33.73175	0.22947		
C Total	149	48.96795			

Root MSE	0.47903	R-square	0.3111
Dep Mean	5.54003	Adj R-sq	0.3018
C. V.	8.64667		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	2.681787	0.37024512	7.243	0.0001
lnL	1	0.177813	0.04470595	3.977	0.0001
lnK	1	0.465154	0.08612163	5.401	0.0001

Covariance of Estimates

COVB	INTERCEP	lnL	lnK
INTERCEP	0.1370814505	-0.003289697	-0.02723683
lnL	-0.003289697	0.0019986217	-0.001270417
lnK	-0.02723683	-0.001270417	0.0074169359

י. ההשערות לבדיקת הטענה הינן : $\begin{matrix} H_0: \_\_\_\_\_\_\_\_ \\ H_1: \_\_\_\_\_\_\_\_ \end{matrix}$.

יא.הסטטיסטי של WALD לבדיקת הטענה הינו :

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים

ii. ניתן לחישוב וערכו: __________________________________.

הועלתה הטענה כי גמישות התפוקה ביחס להון גדולה פי 2 מגמישות התפוקה ביחס לעבודה.
בדקו את הטענה במשוואה (3).

יב. השערת האפס לבדיקת הטענה היא : $H_0: \_\_\_\_\_\_\_\_$.

יג. הסטטיסטי t לבדיקת הטענה הינו :

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים.

ii. ניתן לחישוב וערכו: ____________________________.

יד. הרגרסיה המוגבלת כאשר H_0 נכונה ("תחת H_0") למבחן WALD
הינה: $Z_0 = \gamma_0 + \gamma_1 \cdot Z_1 + V$.
כאשר:
$\begin{matrix} Z_1 = \_\_\_\_\_\_\_\_ \\ Z_0 = \_\_\_\_\_\_\_\_ \end{matrix}$.

טו.הסטטיסטי של WALD לבדיקת הטענה (חשבי ישירות) :

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים.

ii. ניתן לחישוב וערכו: ____________________________.

טז.נטען כי אם נמדוד את המשתנים הב"ת
במודל בדולרים במקום בשקלים, האומדים
ל- β ול- α "יישארו ללא שינוי
(הנח כי שער הדולר הוא 3.5 ₪): נכון/לא נכון/אי אפשר לדעת

יז. נטען שאם נוריד את משתנה PY מהמודל
ה- $\bar{R}^2$ יעלה: נכון/לא נכון/אי אפשר לדעת

2) ענו על כל השאלות הבאות. כל שאלה בפני עצמה. בכל השאלות מונח המודל: $Y = \alpha + \beta X + U$ (ומתקיימות כל ההנחות הקלאסיות).

א. במודל לוגריתמי כפול β מייצגת את
שיעור השינוי השולי: נכון/לא נכון/ אי אפשר לדעת

ב. במודל ללא חותך מתקיימת המשוואה
הנורמאלית: $\sum \hat{u}_t x_t = 0$ בלבד: נכון/לא נכון/אי אפשר לדעת

ג. כאשר מוסיפים משתנה ב"ת למודל, עליה ב- $\bar{R}^2$ מעידה על כך שהמשתנה שהוסף מובהק באוכלוסייה: נכון/לא נכון/אי אפשר לדעת

ד. אם הנחה מס' 3 ($E(\hat{u})=0$ לכל t) איננה מתקיימת, האומדים של המודל לא יהיו יעילים: נכון/לא נכון/אי אפשר לדעת

ה. ככל ש- S_{XX} גדול יותר, קל יותר לדחות את H_0 למובהקות ה- β: נכון/לא נכון/אי אפשר לדעת

ו. $R^2 > \bar{R}^2$ מתקיים תמיד: נכון/לא נכון/אי אפשר לדעת

ז. מבחן F למובהקות המודל מהווה מקרה פרטי של מבחן t למובהקות ה- β: נכון/לא נכון/אי אפשר לדעת

ח. ככל שגודל המדגם גדל כך האומד יהיה יעיל יותר לפרמטר באוכלוסייה: נכון/לא נכון/אי אפשר לדעת

ט. ה-PVALUE גדל ביחס הפוך לרמת המובהקות של המבחן (ה- α): נכון/לא נכון/אי אפשר לדעת

י. אם דחינו את H_0 במבחן t למובהקות ה- β כאשר האומד חיובי, נדחה אותה בהכרח גם ביחס להשערה כי מקדם השיפוע חיובי באוכלוסייה: נכון/לא נכון/אי אפשר לדעת

יא. אם ידוע כי הקשר בין X ל-Y מובהק באוכלוסייה, הדבר מעיד בהכרח על מובהקות המודל: נכון/לא נכון/אי אפשר לדעת

3) נתון המודל: $Y_t = \beta X_t + U_t$.

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

א. $\tilde{\beta}$ הינו אומד חסר הטייה ל- β: נכון/לא נכון/אי אפשר לדעת

ב. שונותו של האומד: ______________________________.

ג. על סמך משפט גאוס מרקוב ניתן להסיק כי אר"פ הינו אומד יעיל יותר מ- $\tilde{\beta}$: נכון/לא נכון/אי אפשר לדעת

ד. המשוואות הנורמאליות: $\sum \hat{u}_t = 0$ ו- $\sum \hat{u}_t x_t = 0$ הינן המשוואות לאמידת הפרמטרים של המודל בשיטת הריבועים הפחותים: נכון/לא נכון/אי אפשר לדעת

ה. אם נתון ש: $\bar{X} = 0$ אזי $\tilde{\beta}$ הינו אומד הריבועים הפחותים: נכון/לא נכון/אי אפשר לדעת

תשובות סופיות:

1) א. ii , $F = 31.273$.
ב. iii , $t = 5.5$.
ג. $H_0: \beta = 1$, $H_1: \beta < 1$.
ד. iv.
ה. i.
ו. לא נכון.
ז. ראו סרטון.
ח. ראו סרטון.
ט. ראו סרטון.
י. $H_0: \beta_3 = 0$, $H_1: \beta_3 \neq 0$.
יא. ii , $WALD_{stat} = 1.74$.
יב. $H_0: \beta_2 = 2 \cdot \beta_1$.
יג. ii , $t = 0.1417$.
יד. $Z_0 = \ln(Y)_t$, $Z_1 = \ln(L)_t + 2\ln(K)_t$.
טו. ii , $WALD_{stat} = 0.585$.
טז. לא נכונה.
יז. לא נכון.

2) א. לא נכון.
ב. נכון.
ג. לא נכון.
ד. לא נכון.
ה. נכון.
ו. נכון.
ז. נכון.
ח. נכון.
ט. לא נכון.
י. נכון.
יא. נכון.

3) א. נכון.
ב. $V(\tilde{\beta}) = \frac{\sigma^2}{S_{XX}}$.
ג. נכון.
ד. לא נכון.
ה. נכון.

אקונומטריקה א

פרק 11 - מבחן 4

תוכן העניינים

מבחן 4:

שאלות:

1) בנק מעוניין לאמוד את סך הפעילות בכרטיסי אשראי של לקוחותיו.
לשם כך אסף נתונים על 35,971 מלקוחותיו ואמד את המשוואה הבאה:

$$CREDIT_t = \alpha + \beta \cdot SAVINGS_t + U_t \quad .1$$

כאשר:

$CREDIT_t$ - סך הפעילות בכרטיסי אשראי ב-₪.

$SAVINGS_t$ - סך הפעילות בחשבונות חיסכון ב-₪.

U_t - סטיה מקרית המקיימת את כל ההנחות הקלאסיות.

משוואה (1) נתונה בפלט מס' 1.

Dependent Variable: credit

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	---	-----	------	------	<0.0001
Error	---	-----	------		
C Total	---	-----			

Root MSE	43859	R-square	0.0106
Dep Mean	7433.60809	Adj R-sq	0.0106
C. V.	589.99662		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0:	Prob>lTl	95% Confidence	
INTERCEP	1	11151.91516	394.35144	2.92	0.0035	378.97	1924.8
savings	1	0.56719	0.02884	19.67		0.51	0.623

א. סטטיסטי F לבדיקת מובהקות המודל הינו:

i. לא ניתן לחשב את סטטיסטי F בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא: ______________________.

ב. PVALUE של סטטיסטי t לבדיקת מובהקות ה- β:

i. לא ניתן לחשבו בעזרת הנתונים הקיימים.

ii. לא ניתן להשתמש בסטטיסטי t בהשערה מסוג זה.

iii. ניתן לחשבו וערכו: ______________________.

הבנק טען שאם יגדילו לקוחותיו את הפעילות בחשבונות חיסכון שלהם אפילו בשקל אחד, הפעילות בכרטיסי אשראי תגדל ביותר מ 40 אגורות.

ג. ההשערות לבדיקת הטענה הינן: $H_0:$ _______ $H_1:$ _______ .

ד. הסטטיסטי לבדיקת טענת הבנק הינו:

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים.

ii. הסטטיסטי לבדיקת הטענה צריך להיות שלילי.

iii. 19.67.

iv. 5.797.

ה. הסטטיסטי של WALD לבדיקת טענת הבנק:

i. לא ניתן לחשב בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו: _______________________.

ו. ברמת ביטחון של 95% מהו טווח הגידול בפעילות בכרטיסי אשראי, על כל שקל נוסף בפעילות בחשבונות חיסכון?

ז. ברמת ביטחון 95% מהו האומד לתוחלת פעילות בכרטיסי אשראי עבור סך פעילות בחשבונות חיסכון של 50,000 ₪?

ח. אם פעילות כרטיסי האשראי של כל לקוח תגדל ב- 1000 ₪:

i. האומד של α ישתנה: נכון/לא נכון/ אי אפשר לדעת

ii. האומד של β ירד: נכון/לא נכון/ אי אפשר לדעת

iii. סטטיסטי F לבדיקת מובהקות המודל לא ישתנה: נכון/לא נכון/ אי אפשר לדעת

נטען שסה"כ פעילות הלקוח בחשבונות חיסכון איננו המשתנה המשפיע על הפעילות בכרטיסי האשראי, אלא הרכב החסכונות.
לשם כך נאמדה המשוואה הבאה:

2. $CREDIT_t = \alpha + \beta_1 \cdot PIKADON1_t + \beta_2 \cdot PIKADON2_t + U_t$.

כאשר:

$PIKADON1_t$ - סה"כ הפקדה לפקדונות יומיים ב-₪.

$PIKADON2_t$ - סה"כ הפקדה לפקדונות חודשיים ב-₪.

משוואה (2) נאמדה בפלט מס' 2.

Dependent Variable: lnY

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	2	1.00791E12	5.003955E11	261.10	0.0001
Error	35968	6.893195E13	1916479937		
C Total	35970	6.993274E13			

Root MSE	43778	R-square	0.0143
Dep Mean	7433.68809	Adj R-sq	0.0143
C. V.	588.90847		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	1259.36230	379.00751	3.32	0.0009
Pikadon1	1	0.07552	0.05539	1.36	0.1728
Pikadon2	1	0.72350	0.03199	22.62	0.0001

Covariance of Estimates

COVB	INTERCEP	Pikadon1	Pikadon2
INTERCEP	143646.69097	-8.178835194	-9.154578973
Pikadon1	-8.176835154	0.0030678685	0.0003564263
Pikadon2	-9.15457897	0.0003564263	0.0010231462

ט. השערת האפס לבדיקת הטענה הינה: _______ : H_0.

י. הסטטיסטי של WALD לבדיקת הטענה:

i. לא ניתן לחשב בעזרת הנתונים הקיימים.

ii. ניתן לחשב וערכו: ________________________.

יא. הסטטיסטי של t לבדיקת הטענה:

i. לא ניתן לחשב בעזרת הנתונים הקיימים.

ii. לא ניתן להשתמש בסטטיסטי t בהשערה מסוג זה.

iii. ניתן לחישוב וערכו: ____________________________.

נטען שהגדלת הפעילות בחשבונות חיסכון של הלקוח על ידי העברה לפקדונות חודשיים משפיעה על הפעילות בכרטיסי אשראי פי 10 מאשר הגדלת הפעילות בחשבונות חיסכון על ידי העברה לפקדונות יומיים.

יב. השערת האפס לבדיקת הטענה הינה: _______ : H_0.

יג. הסטטיסטי t לבדיקת הטענה הינו:

i. לא ניתן לחשב את הסטטיסטי בעזרת הנתונים הקיימים.

ii. ניתן לחשב וערכו הוא: ____________________.

יד. PVALUE של סטטיסטי t מהסעיף הקודם:

i. לא ניתן לחשב את הסטטיסטי t בעזרת הנתונים הקיימים.

ii. ניתן לחשב וערכו הוא: ___________________________.

טו.הרגרסיה המוגבלת כאשר H_0 נכונה למבחן WALD

הינה: $D_0=\gamma_0+\gamma_1\cdot D_1+\gamma_2\cdot D_2+v$.

D_0 : ______

כאשר: D_1 : ______ .

D_2 : ______

טז.על פי משוואה מס' 2, כל שקל שיועבר
לפיקדון הראשון יוסיף כ-0.07552 ₪
לסה"כ הפעילות בכרטיסי אשראי: נכון/לא נכון/אי אפשר לדעת

2) ענו על השאלות הבאות (כל שאלה בפני עצמה, בכל שאלה מונח
המודל: $Y=\alpha+\beta\cdot X+U$ ומתקיימות כל ההנחות הקלאסיות).

א. אם המודל מובהק אזי שיפוע
הרגרסיה מובהק בהכרח: נכון/לא נכון/אי אפשר לדעת

ב. הגמישות במודל חצי לוגרתמי
היא קבועה: נכון/לא נכון/אי אפשר לדעת

ג. אם X_2 מהווה קומבינציה ליניארית
של X_1 לא ניתן לאמוד את הרגרסיה
המרובה: $Y=\alpha+\beta_1X_1+\beta_2X_2+U$: נכון/לא נכון/אי אפשר לדעת

ד. $\bar{R}^2>R^2$ רק בתנאי שהמודל מובהק: נכון/לא נכון/אי אפשר לדעת

ה. ליניאריות וחוסר הטיה של האומדים
מהווים תנאי הכרחי לעקיבותם: נכון/לא נכון/ אי אפשר לדעת

ו. נתון כי רווח הסמך לאמידת β ברמת
סמך של 95% הוא: [-2,-5].
מכך ניתן להסיק כי שיפוע הרגרסיה מובהק
ברמת מובהקות של 5%: נכון/לא נכון/אי אפשר לדעת

ז. ככל שפיזור U_t גדול יותר כך קשה יותר
לדחות את H_0 למובהקות המודל: נכון/לא נכון/ אי אפשר לדעת

ח. מודלים לא ליניאריים מתארים קשרים שאינם
ליניאריים בין המשתנה המסביר למוסבר: נכון/לא נכון/אי אפשר לדעת.

ט. אם הנחה 5 (שונות קבועה) לא מתקיימת, אומדי
הריבועים הפחותים אינם חסרי הטיה: נכון/לא נכון/אי אפשר לדעת

י. אם דחינו את H_0 לבדיקת הטענה כי שיפוע
הרגרסיה הוא שלילי בוודאי שמודל הרגרסיה
הוא מובהק: נכון/לא נכון/ אי אפשר לדעת

3) נתון המודל: $Y_t = \beta \cdot X_t + U_t$, כאשר כל ההנחות הקלאסיות מתקיימות.

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

א. $E\left(\tilde{\beta}\right) = $ _______.

ב. על סמך משפט גאוס מרקוב אומד זה יעיל פחות מאומד הריבועים הפחותים: נכון/ לא נכון/אי אפשר לדעת

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

ד. חשבו את השונות של $\tilde{\beta}$ עבור מודל שבו $\alpha \neq 0$.

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

תשובות סופיות:

1) א. ii , $F = 386.9089$. ב. iii , $PF < 0.0001 = Pt$. ג. $\begin{matrix} H_0 : \beta = 0.4 \\ H_1 : \beta > 0.4 \end{matrix}$.

ד. iv. ה. i. ו. $p\left(0.51 \le \beta \le 0.623\right) = 0.95$.

ז. $p\left(-32,387,174.83 \le E\left(Y\right) \le 32,458,197.67\right) = 0.95$.

ח. i. נכון. ii. לא נכון. iii. נכון. ט. $H_0 : \beta_1 = \beta_2 = 0$.

י. i. יא. ii. יב. $H_0 : \beta_2 = 10 \cdot \beta_1$. יג. ii , $t = -0.0574$.

יד. ii , $PVALUE > 0.1$. טו. $\begin{matrix} D_0 : CREDIT_t \\ D_1 : SAVINGS_t \\ D_2 : PIKADON1_t + 10 \cdot PIKADON2_t \end{matrix}$.

טז. נכון.

2) א. נכון. ב. לא נכון. ג. נכון. ד. לא נכון.
ה. לא נכון. ו. נכון. ז. נכון. ח. לא נכון.
ט. לא נכון. י. לא נכון.

3) א. $E\left(\tilde{\beta}\right) = \dfrac{\beta \sum X_t}{S_{XX}}$. ב. לא ניתן לדעת. ג. נכון.

ד. $V\left(\tilde{\beta}\right) = \dfrac{T\sigma^2}{S_{XX}^2}$. ה. לא נכון.

אקונומטריקה א

פרק 12 - מבחן 5

תוכן העניינים

מבחן 5:

שאלות:

1) על מנת לאמוד את הקשר בין רמת המחירים במשק (P) לכמות הכסף (M), נאספו נתונים חודשיים בשנים 86-94 (סה"כ 105 תצפיות) ונאמדה המשוואה הבאה:

1. $M_t = e^{\alpha} + p^{\beta} + e^{u}$.

כאשר:

m - כמות הכסף במשק לחודש (מזומנים + עו"ש).

p - מדד המחירים לצרכן במשק.

U_t - סטיה מקרית המקיימת את כל ההנחות הקלאסיות.

משוואה מס' (1) נאמדה בפלט מס' 1.

Dependent Variable: lnm

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	1				<0.0001
Error	103				
C Total	104	44.91976			

Root MSE	0.09251	R-square	0.9804
Dep Mean	8.53854	Adj R-sq	0.9802
C. V.	1.08344		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Paramete r=0	Prob>lTl
INTERCE P	1	1.49372	0.09862	15.15	<.0001
lnp	1	1.69267	0.02360		<.0001

א. כתבו את המשוואה בצורה ליניארית בעזרת הטרנספורמציה המתאימה.

ב. האומדן למשוואה (1) הינו: ______________________.

ג. המשמעות הכלכלית של β היא: ____________

ד. גבולות רווח-סמך ברמת סמך של 95% עבור β הינם:

גבול תחתון: ____________________.

גבול עליון: ____________________.

ה. ערך t לחישוב מובהקות ה-β הינו:

i. לא ניתן לחשב ערך זה בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא: ____________________.

ו. אם נגדיל את מדד המחירים לצרכן ביחידה אחת, כמות הכסף במשק תגדל ב:
   i. 71.7233.
   ii. 1.69267.
   iii. 169.267.
   iv. 1.69267%.
   v. אף תשובה איננה נכונה.

הועלתה הטענה שתוספת של אחוז אחד במדד המחירים לצרכן תגדיל את כמות הכסף במשק ביותר מאחוז אחד.

ז. ההשערות לבדיקת הטענה: ____________________.

ח. סטטיסטי t לבדיקת הטענה הינו:
   i. לא ניתן לחשבו באמצעות הנתונים הקיימים.
   ii. ניתן לחשבו וערכו הוא: ________________________________.

ט. על פי התשובות לסעיפים הקודמים ניתן להסיק כי ערכו של סטטיסטי F לבדיקת מובהקות המודל הינו:
   i. לא ניתן לחשב את ערכו של סטטיסטי F על סמך סטטיסטי t.
   ii. 861.4225.
   iii. 5144.23.
   iv. 71.7233 4.

י. אם נוציא שורש ריבועי למדד המחירים לצרכן במשק:
   i. האומד של α ישתנה: נכון/לא נכון/ אי אפשר לדעת
   ii. האומד של β יעלה: נכון/לא נכון/ אי אפשר לדעת
   iii. סטטיסטי F לבדיקת מובהקות המודל לא ישתנה: נכון/לא נכון/ אי אפשר לדעת

הועלתה הטענה כי יש צורך להוסיף למשוואה גם את הפעילות הכלכלית במשק (Y) כמשתנה מסביר, ולכן יש לאמוד את המשוואה הבאה:

2. $LN(M)_t = \alpha + \beta_1 \cdot LN(P)_t + \beta_2 \cdot LN(Y)_t + U_t$.

משוואה (2) נתונה בפלט מס' 2.

Dependent Variable: lnm

Analysis of Variance

Source	DF	Sum of Squares	Mean Squares	F Value	Prob>F
Model	2	44.05069	22.02535	2585.05	<0.0001
Error	102	0.86907	0.00852		
C Total	104	44.91976			

Root MSE	0.09231	R-square	0.9807
Dep Mean	8.53854	Adj R-sq	0.9803
C. V.	1.08104		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob>|T|
INTERCEP	1	0.78242	0.59739	1.31	0.1932
lnp	1	1.63491	0.05332	30.66	<.0001
lny	1	0.20001	0.16568	-----	0.2302

Covariance of Estimates

COVB	INTERCEP	lnp	lny
INTERCEP	0.35687	0.025884	-0.09762
lnp	0.02588	0.002843	-0.00792
lny	-0.09762	-0.00792	0.02745

יא.סטטיסטי t לבדיקת הטענה הינו :

i. לא ניתן לחשב בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא :______________________.

יב. על פי התשובה לסעיף הקודם, ניתן להסיק את ערכו של סטטיסטי F למובהקות המודל. נכון/ לא נכון/ אי אפשר לדעת

יג. על פי התשובה לסעיף יא' ניתן להסיק את ערכו של סטטיסטי WALD לבדיקת הטענה. נכון/לא נכון/ אי אפשר לדעת

הועלתה הטענה כי הגמישות ביחס למחיר גבוהה פי 10 מהגמישות ביחס לפעילות הכלכלית במשק.

יד. סטטיסטי WALD לבדיקת הטענה הינו :

i. לא ניתן לחשב בעזרת הנתונים הקיימים.

ii. ניתן לחשבו וערכו הוא :__________________________.

טו.הרגרסיה המוגבלת כאשר H_0 נכונה למבחן WALD הינה :__________

כאשר : D_0 : _______ ; D_1 : _______ .

טז.

i. איזה מבין המודלים המוצעים במשוואות 1 ו-2 עדיף?
משוואה 1/משוואה2/אין הבדל בין המודלים

ii. אם משתנה רמת המחירים במשק היה מובהק במשוואה מס׳ 1, הוא יהיה מובהק בהכרח גם במשוואה מס׳ 2 : נכון/לא נכון/אי אפשר לדעת

2) ענו על השאלות הבאות (כל שאלה בפני עצמה, בכל שאלה מונח המודל: $Y=\alpha+\beta\cdot X+U$ ומתקיימות כל ההנחות הקלאסיות).

א. $\bar{R}^2 < R^2$ מתקיים תמיד: נכון/לא נכון/אי אפשר לדעת.

ב. אם דוחים H_0 במבחן חד צדדי ברמת מובהקות α, אזי בהכרח גם נדחה H_0 במבחן הדו צדדי באותה רמת מובהקות: נכון/לא נכון/אי אפשר לדעת.

ג. אם ערך האומד ל-β גבוה, השערת האפס למובהקות השיפוע תידחה בוודאות: נכון/לא נכון/אי אפשר לדעת.

ד. הוספת משתנה מסביר למשוואת הרגרסיה עשויה להקטין את R^2: נכון/לא נכון/אי אפשר לדעת.

ה. אם דוחים H_0 במבחן דו צדדי ברמת מובהקות α, אזי בהכרח גם נדחה H_0 במבחן החד צדדי באותה רמת מובהקות: נכון/לא נכון/אי אפשר לדעת.

ו. אם רווח בר סמך לשיפוע כולל את הערך אפס, ניתן לומר כי השערת האפס למובהקות השיפוע מתקבלת בהכרח: נכון/לא נכון/אי אפשר לדעת.

ז. האומדים היעילים ביותר לפרמטרים באוכלוסייה יהיו בהכרח אומדי הריבועים הפחותים: נכון/לא נכון/אי אפשר לדעת.

ח. בהוספת משתנה מסביר מובהק למודל, ערך $\bar{R}^2$ יעלה בהכרח. נכון/לא נכון/אי אפשר לדעת.

ט. מבחן WALD הוא מקרה פרטי של מבחן F למובהקות המודל: נכון/לא נכון/אי אפשר לדעת.

י. שיטת הריבועים הפחותים מביאה למקסימום את $\bar{R}^2$: נכון/לא נכון/אי אפשר לדעת.

3) נתון המודל: $Y_t = \alpha + \beta X_t + U_t$.

נתון כי אר"פ למודל זה הינו: $\hat{\beta} = \frac{S_{XY}}{S_{XX}}$.

א. הוכיחו כי $\hat{\beta}$ אומד ליניארי וחסר הטיה של β.

ב. חשבו את $VAR\left(\hat{\beta}\right)$.

ג. נתון האומד: $\tilde{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2}$.

הוכיחו כי $\tilde{\beta}$ אומד ליניארי אך איננו חסר הטיה ל- β.

ד. מהם התנאים בהם מתקיים: $E\left(\tilde{\beta}\right) = \beta$?

תשובות סופיות:

1) א. $LN(M)_t = \alpha + \beta \cdot LN(P)_t + U_t$. ב. $LN(M)_t = 1.49372 + 1.69267 \cdot LN(P)_t$.

ג. גמישות. ד. גבול תחתון: 1.64527 , גבול עליון: 1.73987.

ה. ii, $t_{\beta=0} = 71.7233$. ו. v. ז. $H_0 : \beta = 1$, $H_1 : \beta > 1$.

ח. ii, $t = 29.35$. ט. i. י. i. לא ניתן לדעת.

ii. אי אפשר לדעת. iii. אי אפשר לדעת. יא. ii, $t = 1.2$.

יב. לא נכון. יג. נכון. יד. ii, $WALD = 0.048$.

טו. $D_0 = LN(M)_t$, $D_1 = 10 \cdot LN(P)_t + LN(Y)_t$. טז. i. משוואה 1.

ii. לא נכון.

2) א. נכון. ב. לא נכון. ג. לא נכון. ד. לא נכון.

ה. לא נכון. ו. לא נכון. ז. לא נכון. ח. נכון.

ט. לא נכון. י. נכון.

3) א. הוכחה. ב. $V(\hat{\beta}) = \frac{\sigma^2}{S_{XX}}$. ג. הוכחה.

ד. ראו סרטון.