ביטול ההנחות הקלאסיות: ← נתמקד בגרסת מודל פשוט

1. $E(\varepsilon_i)=0 \quad \forall i$
2. $Cov(\varepsilon_i,\varepsilon_j)=0 \quad \forall i\neq j$
3. $Var(\varepsilon_i)=\sigma^2$ ← Homosced.
4. $\varepsilon_i \sim N$
5. X לא סטוכסטי

שילוב כל ההנחות הללו גורר כי אומד ה OLS הוא BLUE ומתפלג נורמלית.

נתחיל עם ביטול הנחה (4) $\varepsilon_i \sim N$ ← $\boxed{\text{מס' תצפיות גדול פותר את הבעיה.}}$

* לא צריכים להנחה כי בשביל משפט גה"מ (BLUE)
* ההנחה משמעותית רק למבחני t , F ובניית רה"ס.

אמרנו ש $\hat{\beta}$ מתפלג נורמלית כיוון ש:

$$\hat{\beta}=\beta+\frac{\sum(x_i-\bar{x})\varepsilon_i}{\sum(x_i-\bar{x})^2}$$

סכום של משתנים שמתפלגים נורמלית ← מתפלג נורמלית. אבל אף פעם לא משתמשים רק ב ε-ים הבודדים.

לעיתים ניתן להשתמש במשפט הגבול המרכזי:

סכום של משתנים מקריים שמגיעים מאותה התפלגות ובלתי תלויים ביניהם ← הסכום שלהם מתפלג נורמלית.

מה שחשוב הוא שה ε-ים מגיעים מאותה התפלגות ← תוחלת (1) שונות (3) זהות + ב"ת (2)

משפט הגבול המרכזי אומר לנו ש

$$\frac{\hat{\beta}_j-\beta_j^{H_0}}{sd(\hat{\beta}_j)} \overset{a}{\sim} N(0,1)$$

ההתפלגות שואפת לנורמלית כאשר $n\uparrow$, ורמת

אין מספר מדויק לגבול (כמו 30 וכו')

ביטול הנחה קלאסית ① $E(\varepsilon_i)\neq 0$

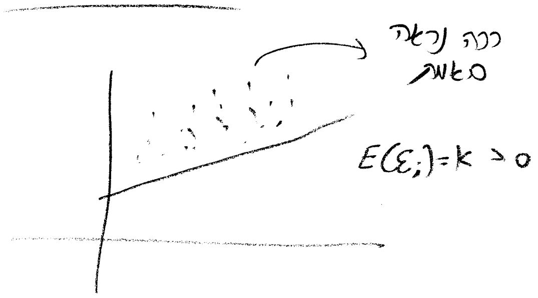
נבחן 2 מקרים: (א) $E(\varepsilon_i)=k$ ← קבוע לכל i.

(ב) $E(\varepsilon_i)=\varepsilon_i$

2 המקרים פוגעים בחוסר הטייה. (א) ← רק בחוסר ההטיה של החותך α.

(ב) ← כל האומדים מוטים וגם השונות לא מינימלית.

מקרה (א) ← לא בעיה:



1) חותך מוטה כלפי מעלה

2) השיפוע נשאר לא מוטה

אפשר לפתור ע"י אומדה המודל הבא

$$y=\underbrace{(\alpha+k)}_{\text{ואפסילון מרכזת את זה.}}+\beta x_i+(\varepsilon_i-k)$$

$\boxed{\text{3) לא פוגע בשונות מינימלית}}$

מקרה (ב): בעיה חמורה

דוגמא מהמקרה של השמטת משתנים:

המודל האמיתי: $Y_i=\beta_0+\beta_1X_{1i}+\beta_2X_{2i}+\varepsilon_i \rightarrow E(\varepsilon_i)=0$

המודל שאמדנו: $Y_i=\beta_0'+\beta_1'X_{1i}+\varepsilon_i' \rightarrow E(\varepsilon_i')=\beta_2X_{2i}+\underbrace{E(\varepsilon_i)}_{0}$

- יתכן שההשפעה היא יותר כל מה שמשפיע על y ולא התחשבנו בו
- ההפרעה משתנה מתצפית לתצפית כי יש x שונה לכל תצפית.

$E(\hat\beta_1')=\beta_1+\beta_2\cdot\frac{Cov(X_1,X_2)}{Var(X_1)}$ ← האומדים מוטים.

לא עקיבים ואין כל שימוש למבחני השערות.

ביטול הנחה 5 ← חוסר סטוכסטיות (x לא מקרי)

זאת הנחה לא ריאליסטית. למזלנו לא צריך אותה. ← מספיק להניח כי $E(u|x)=0$ (ניתן גם חוסר הט"ה. ישנוק')

מה קורה כאשר x משתנה מקרי:

מבדילים בין 2 מקרים:

א. $cov(x_i, \varepsilon_i)=0$ ← אומד עקיב

ב. $cov(x_i, \varepsilon_i)\neq 0$ ← אומד מוטה ולא עקיב

"המציאות": — בעולם האמיתי אין כמעט אומדים חסרי הט"ה.

מבדילים בין אומדים עקיבים ולא עקיבים.

$$E(\hat{\beta})=E(\beta)+E\left(\frac{\sum(x_i-\bar{x})\varepsilon_i}{\sum(x_i-\bar{x})^2}\right)$$

כאן אי אפשר להוציא תוחלת אם לא תלויים.

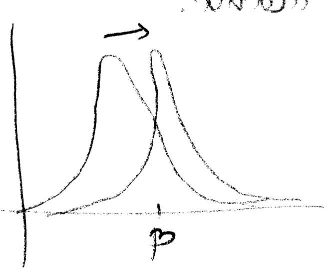
אבל נראה מה קורה כאשר $n\to\infty$

$$E(\hat{\beta})=\beta+E\left(\frac{\overbrace{\sum(x_i-\bar{x})(\varepsilon_i-\bar{\varepsilon})}^{\hat{cov}(x,\varepsilon_i)}}{\sum(x_i-\bar{x})^2}\right)$$

הביטוי הזה שואף ל $cov(x,\varepsilon)$ כאשר $n\to\infty$

הגדרה של עקיבות של אומד (מקביל לחוסר הט"ה אבל לא זהה).

1. השונות של האומד ← 0 כאשר $n\to\infty$
2. ההתפלגות של האומד מתכנסת סביב הפרמטר. כאשר $n\to\infty$



עבור n גדול הט"ה נהיית שולית.

יעילות אסימפטוטית: $Var(\hat{\beta})\to 0$ $n\to\infty$

סיכום:

ב. אם $Cov(x_i,\varepsilon_i)=0$ ← β האומד עקבי ויעיל אסימפטוטית.

$$E(\hat\beta) \to \beta$$
$$n \to \infty$$

אם $Cov(x_i,\varepsilon_i)\neq 0$ ← β האומד מוטה ולא עקבי.

ג. $$E(\hat\beta) \to \beta + \frac{Cov(x_i,\varepsilon_i)}{Var(x_i)}$$

מה הבעיה? אי אפשר לבדוק אם $Cov(x_i,\varepsilon_i)\overset{?}{=}0$

שאלה מכשילה: הרצת רגרסיה $Y_i = \alpha + \beta x_i + \varepsilon_i$

קיבלו אומדים $\hat\alpha, \hat\beta, \hat\varepsilon_i$

אח"כ רצו לבדוק האם $Cov(x,\varepsilon_i)=0$, לכן אמדו:

$$\hat\varepsilon_i = \alpha + \beta x_i + u_i$$

קיבלו ש $\hat\beta=0$ ← האם מזה ניתן לומר כי $Cov(\varepsilon_i,x_i)=0$? לא.

ההתמקדות בשאלה האם $Cov(x,\varepsilon)=0$ היא בד"כ בגלל טעות ספסיפקציה. שכחים לצרף משהו והשמטנו משתנה:

$$\underset{\text{שכר}}{Y} = \beta_0 + \underset{\text{השכלה}}{\beta_1 x_{1i}} + \underset{\text{נסיון}}{\beta_2 x_{2i}} + \ldots + \varepsilon_i$$

אינטליגנציה ← בלתי נצפה

אבל מניחים כי פרטים אינטליגנטים ← יותר השכלה.

$\Rightarrow Cov(x_{1i},\varepsilon_i)\neq 0 \to \hat\beta_1$ מוטה ולא עקבי

מה עושים?? נדבר על זה במקרה של IV.

סיכום:

$Cov(x,\varepsilon)\neq 0$ ← אומד מוטה, רגרסיה לא שווה.

$Cov(x,\varepsilon)=0$ ← אומד עקבי. אפשר להמשיך.

שלב 3 הנחות שלנו ← אומדי OLS הם BLUE ומתפלגים נורמלית

עקביות:
① השונות של $\hat{\beta} \to 0$
② ההתפלגות מתכווצת סביב הפרמטר.

① אפילו בלי הנחה ש $u \sim N$

If you can't get it right as $n \to \infty$, you shouldn't be in this business

עקביות ←

משתמשים במשפט הגבול המרכזי כדי לומר שגם כש ε לא נורמלי t_{stat} שמשתמש מתפלג אסימפטוטית נורמלי.

אפשר לעשות בדיקת השערות

$$\boxed{\hat{\beta} = \sum a_i y_i}$$

ממוצע משוקלל.

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} = \beta_1 + \frac{\frac{\sum (\quad)(\quad)}{n}}{\frac{\sum (x_i - \bar{x})^2}{n}}$$

$\bar{Z}$

שימוש בחוק המספרים הגדולים ←

$$\operatorname{plim} \hat{\beta} = \beta_1 + \frac{Cov(x_1, u)}{Var(x_1)}$$

חוק המספרים הגדולים: הממוצע המדגמי של z של תצפיות $E(z)$

הממוצע הינו אומד עקבי של $E(z)$

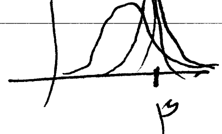
$$\operatorname{plim}(\bar{z}) = E(z) = Cov(x, u)$$

בשביל עקביות מספיק $Cov(x_1, u) = 0$?

אם $E(u|x \ldots) \neq 0$ ← OLS מוטה ולא עקבי ← u אצלו לא כולל הכל

← אפשר לקבל אומדים מוטים אבל עקביים

$\leftarrow E(\hat{\beta}) \neq \beta \quad ; \quad \operatorname{plim} \hat{\beta} = \beta$



Heteroscedasticity הטרוסקדסטיות:

6

בניגוד להנחה 3

Homosce. $\rightarrow Var(\varepsilon_i) \neq \sigma^2 \quad \forall i$

כלומר: $Var(\varepsilon_i)=\sigma_i^2 \leftarrow$ דוגמא: $Var(\varepsilon_i)=\sigma^2\cdot \Pi_i^2$

דוגמאות: (1) אם מסתכלים 2 חברות גדולות וחברות קטנות

ההוצאות ε_i של חברה גדולה יהיה גדול משל חברה קטנה

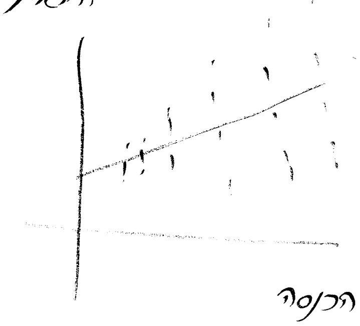
(לדוגמא: מכירות במליוני דולרים רספונסיבית של תהליך פרסום)

(2) דוגמא נוספת: הוצאות של משפחות לפי הכנסה גבוהה

לעומת הכנסה נמוכה - סביר שמשפחות לפי הכנסה נמוכה יוציאו באופן יציב יותר.

expenditure $= f(\quad) + \varepsilon_i$

$\Pi_i = f(x_i)$



(3) דוגמא שכיחה נוספת: Group Data ← נניח שיש נתונים של

ממוצעים של תצפיות ולא תצפיות בודדות: הוצאות של עיר פרטים מסוימת:

$\bar{Y}_i = \beta_0 + \beta_1 \bar{X}_i + \bar{\varepsilon}_i \qquad \bar{Y} = \frac{\sum Y_i}{n_i} \qquad i=1,\dots,n$ יישובים

הכנסה ממוצעת ביישוב

$\Pi_i = f(n_i)$

$$\leftarrow Var(\bar{\varepsilon}_i) = Var\left(\frac{\sum^{n_i} \varepsilon_{ij}}{n_i}\right) = \frac{1}{n_i^2}\cdot E\left(\sum \varepsilon_{ij}^2\right) = \frac{n_i\cdot\sigma_\varepsilon^2}{n_i^2}$$

$$= \frac{\sigma_\varepsilon^2}{n_i}$$

$\rightarrow$ כלומר - ברמת הפרטים אין בעיה של שונות שונה

אבל ברגרסיה ברמת היישוב יש בעיה

השונות קטנה כי שיש יותר תצפיות בעיר מסוימת.

(4) דוגמא של אוכלוסיות שונות: נניח שונות שונה לאוכלוסיות שונות

(גברים מול נשים ← תשובות במבחן לא״פ שאר).

לחשוב איך זה נראה

* $Var(\hat\beta)=E(\hat\beta-\beta)^2=E\left(\frac{\Sigma(x_i-\bar x)\varepsilon_i}{\Sigma x_i^2}\right)^2=\left(\frac{1}{\Sigma x_i^2}\right)^2\cdot E\left(\Sigma x_i\varepsilon_i\right)^2=\frac{\Sigma x_i^2\sigma_i^2}{(\Sigma x_i^2)^2}\neq\dots$ (7)

נקודות חשובות:

(1) שונות שונה אינה פוגעת בתכונת חוסר ההטיה @ אומדי ה OLS. ← לא השתמשנו בכך בשביל להוכיח חוסר הטיה

(2) שונות שונה כן פוגעת ביעילות האומדים - האומדים אינם BLUE.

(3) כמו כן במקרה כזה לא ניתן להסתמך & רב"ס ובדיקת השערות. ← השתמשנו בהנחת השונות הזהה בשביל לחשב את השונות

הערה: שיטת ה OLS נותנת משקל שווה לכל תצפית.

תצפיות עם שונויות גבוהות יותר ← דומיננטיות יותר. אבל הן פחות מהימנות בפועל.

$\Sigma\hat\varepsilon_i^2$ → סכימה עם אותו משקל לכל תצפית

$Var(\hat\beta)=\frac{\Sigma x_i^2\sigma_i^2}{(\Sigma x_i^2)^2}$ → השונות האמיתית

בעצם מה שמעניין אותנו הוא הקשר @ x & y.

אינטואיציה: * לכן נחפש שיטה שתיתן לתצפיות עם שונות גבוהה יותר פחות משקל באמידה.

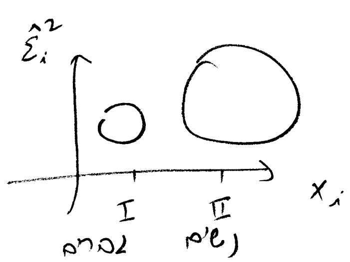
מה עושים?

(1) בוחנים האם יש בעיה @ Heteroscd. - מתחילים עם מבחני ט"ן.

(א) מריצים את המודל $Y_i=\alpha+\beta x_i+\varepsilon_i$ ← במודל פשוט יותר

(ב) מוציאים את סדרת ה resid - $\hat\varepsilon_i$ ומ"רבעים $\hat\varepsilon_i^2$

(ג) משרטטים בגרפית & scatter @ $\hat\varepsilon_i^2$ & x_i שבו חושבים שיש בעיה.



(ד) מבחן סטטיסטי

(ה) תיקון השונות השונה.

ב) מבחנים סטטיסטיים לזיהוי שונות שונה:

מבחן Goldfeld Quandt:

מעין מבחן F שבודק שונות שונה.

מתאים למצבים בהם יודעים את הסיבה לשונות שונה ← לדוגמא:

עבור משתנה ספציפי x_j: $Var(\varepsilon_i)=\sigma^2 x_{j,i}^2$

שלבים:

1. מסדרים את התצפיות בסדר עולה לפי x_j. (כמו הקודם).
2. מחלקים את התצפיות לשלושה חלקים - x_j קטן, בינוני וגבוה.
3. זורקים את הקבוצה האמצעית (בין $\frac{1}{5}$ לשליש מהתצפיות) בגודל sample. ומוודאים שהגדלים של 2 הקבוצות הנותרות זהים.
4. מריצים 2 רגרסיות נפרדות על כל תת מדגם. מחשבים ESS_1, ESS_2 → Error Sum of Sq.
5. מבחן F:

H_0: Homscedast.

H_1: Heter.

הגבוה יותר

$$F_{stat}=\frac{ESS_2 \Big/ \frac{n-m-2k-2}{2}}{ESS_1 \Big/ \frac{n-m-2k-2}{2}}$$

$$\sim F^{(1-\alpha)}_{\left(\frac{n-m-2k-2}{2}, \ \dots\right)}$$

אם $F_{stat} > F_{crit}$ נדחה את השערת האפס ← כלומר יש בעיה של שונות שונה.

מבחן כללי יותר ← מבחן White :

כאשר לא יודעים מהו מקור הבעיה של השונות השונה (σ_i^2)

נניח מודל:
$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

שלב 1: אומדים רגרסיה רגילה.

שלב 2: מחשבים את השאריות $\hat{\varepsilon}_i^2 = (Y_i - \hat{Y}_i)^2$

שלב 3: מריצים את הרגרסיה:

$$\hat{\varepsilon}_i^2 = \gamma_0 + \underbrace{\gamma_1 X_{1i} + \gamma_2 X_{2i} + \gamma_3 X_{1i}^2 + \gamma_4 X_{2i}^2 + \gamma_5 X_{1i} X_{2i}}_{p} + v_i$$

שלב 4: לוקחים את R^2 של הרגרסיה ואת N - מס' התצפיות.
מחשבים $N \cdot R^2$.

שלב 5: מנסחים את H_0 שכל המקדמים (חוץ מהחותך) משלב 3 = 0.

$$H_0: \gamma_1 = \gamma_2 = \dots = \gamma_5 = 0$$

(א) $NR^2 \sim \chi^2_{(p)}$ אם $NR^2 > \chi_{crit(5)}$ אז דוחים את השערת האפס.

מבחן White בתוכנה:

* בחלון של הפלט לוחצים View
* בוחרים residual tests ← White res. test
* מסמנים ☑ with combination
* אם ס.כ.כ נמוך ← אז דוחים שאין שונות שונה

איך מתקנים שונות שונה:

GLS ← שיטת כללית תיקון ← ריבועים פחותים משוקללים.
נותן לכל תצפית משקל שונה בפרופורציה ל π_i^2

← מקרה פרטי של GLS ← WLS:
← מחלקים את המודל המקורי ב π_i

$$Var(\varepsilon_i) = \sigma^2 \cdot \pi_i^2$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_K X_{Ki} + \varepsilon_i$$

$$/ : \pi_i \Rightarrow \frac{Y_i}{\pi_i} = \dots\dots + \beta_K \frac{X_{Ki}}{\pi_i} + \frac{\varepsilon_i}{\pi_i}$$

$$E\left(\frac{\varepsilon_i}{\pi_i}\right) = 0 \ ; \quad Var\left(\frac{\varepsilon_i}{\pi_i}\right) = \frac{1}{\cancel{\pi_i^2}} \cdot \sigma^2 \cancel{\pi_i^2} = \sigma^2$$

דוגמאות:

רגרסיה עם שני שיפועים וללא חותך

(1) $\pi_i^2 = \sigma_i^2$! σ_i ידוע

$$\frac{Y_i}{\sigma_i} = \frac{\beta_0}{\sigma_i} + \beta_1 \frac{X_i}{\sigma_i} + \frac{\varepsilon_i}{\sigma_i}$$

$$Var\left(\frac{\varepsilon_i}{\sigma_i}\right) : E\left(\frac{\varepsilon_i}{\sigma_i}\right) = \frac{1}{\cancel{\sigma_i^2}} \cdot \cancel{\sigma_i^2} = \underline{\underline{1}} \to$$ שונות אחידה

שיפוע

חותך

(2) שונות פרופורציונית לאחד המשתנים:

$$\frac{Y_i}{X_i} = \beta_0 \frac{1}{X_i} + \beta_1 \cdot 1 + \frac{\varepsilon_i}{X_i}$$

$$\pi_i = X_i \to \sigma_i^2 = \sigma^2 X_i^2$$

$$Var\left(\frac{\varepsilon_i}{X_i}\right) = \sigma^2$$

(יא)

③ דוגמא נוספת: $\pi_i = \sqrt{x_i}$

$$\frac{y_i}{\sqrt{x_i}} = \beta_0 \frac{1}{\sqrt{x_i}} + \beta_1 \sqrt{x_i} + \frac{\varepsilon_i}{\sqrt{x_i}}$$

גם כאן רגרסיה ללא חותך

* אם מריצים עם חותך ← האומדים לא יעילים
* לשים לב ← נוצרת גם בעיית מולטיקולינאריות בגלל ש x גם מקיימים את:

תיקון white:

① מריצים את הרגרסיה ממבחן white.

② מחשבים את התחזיות

$$\hat{\hat{\varepsilon}}_i^2 = \hat{\gamma}_0 + \hat{\gamma}_1 x_{1i} + \hat{\gamma}_2 x_{2i} + \hat{\gamma}_3 x_{1i}^2 + \hat{\gamma}_4 x_{2i}^2 + \hat{\gamma}_5 x_{1i} x_{2i}$$

את ערכי ה $\hat{\hat{\varepsilon}}_i^2$ חיובים לפי תחזית מחשבים:

$$T_i = \frac{1}{\sqrt{\hat{V}_i}} \quad ; \quad \hat{V}_i = \hat{\varepsilon}_i^2 - \hat{\hat{\varepsilon}}_i^2$$

זה לא
מספר בשיוני
זה לא מספר גדול ← משהו
קטן של התחזית

③ את T_i מכפילים בכל ערך a הרגרסיה המקורית.

④ מריצים רגרסיה.