שיעור מס' 5 :

חזרה מהירה:

הנחנו את הנחות הקלאסיות של ה PRF:

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

במקרה שלנו את
פיתחנו אומד ה OLS,
כאשר מניחים שהמודל
האמיתי הוא:

$$Y_i=\alpha+\beta X_i+\varepsilon_i$$

משוואות נורמליות:

$\min \sum \hat{\varepsilon}_i^2$ חות

1. $\sum \hat{\varepsilon}_i=0$
2. $\sum \hat{\varepsilon}_i X_i=0$

$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X} \Rightarrow$ עובר בנק' הממוצעים → וישר

$$\hat{\beta}=\frac{\hat{Cov}(X,Y)}{\hat{Var}(X)}$$

כעת רצינו להראות את הקשר
בין האומדים לפרמטרים: (מציבים את ה Y האמיתי לפי המודל).

$$\boxed{\hat{\beta}=\beta+\frac{\sum(X_i-\bar{X})\varepsilon_i}{\sum(X_i-\bar{X})^2}} \rightarrow E(\hat{\beta})=\beta$$

$$\rightarrow Var(\hat{\beta})=\frac{\sigma^2_\varepsilon}{\sum(X_i-\bar{X})^2}$$

$$\boxed{\hat{\alpha}=\alpha+\sum\left[\left(\frac{1}{N}-\frac{\bar{X}(X_i-\bar{X})}{\sum(X_i-\bar{X})^2}\right)\varepsilon_i\right]} \rightarrow E(\hat{\alpha})=\alpha$$

$$\rightarrow Var(\hat{\alpha})=\frac{\sigma^2_\varepsilon\sum X_i^2}{n\cdot\sum(X_i-\bar{X})^2}$$

ובנוסף נרצה אומד לשונות של ההפרעה
האקראית

$\sigma^2_\varepsilon=?$ ה $\boxed{\hat{\sigma}^2_\varepsilon=\frac{\sum\hat{\varepsilon}_i^2}{n-2}}$

אפשר לחבר את זה לכך שהקו
עובר דרך נק' הממוצעים

תרגיל השהי:

נראה ש $Cov(\hat{\alpha},\hat{\beta})<0$ $\leftarrow$ ~~[illegible]~~ הקשר שלילי בין החותך
לשיפוע.

$$Cov(\hat{\alpha},\hat{\beta}) = E\left(\hat{\alpha}-E(\hat{\alpha})\right)\left(\hat{\beta}-E(\hat{\beta})\right)$$

$$= E\left(\underbrace{\frac{\sum(x_i-\bar{x})\varepsilon_i}{\sum(x_i-\bar{x})^2}}_{w_i}\right)\left(\sum\underbrace{\left(\frac{1}{n}-\frac{\bar{x}(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}\right)}_{h_i}\varepsilon_i\right)$$

$$= E\left(\sum h_i\varepsilon_i\right)\left(\sum w_i\varepsilon_i\right) = E\left(\sum h_iw_i\varepsilon_i^2+\sum\sum h_iw_j\varepsilon_i\varepsilon_j\right)$$

$$= \sum h_iw_iE(\varepsilon_i^2)+\cancel{\sum\sum h_iw_j\underbrace{E(\varepsilon_i\varepsilon_j)}_{0}}$$

$$= \sigma_\varepsilon^2\underbrace{\sum h_iw_i}_{\text{פרשו נוסחי}}$$

$$= \sigma_\varepsilon^2\cdot\sum_{i=1}^{n}\underbrace{\left[\frac{1}{n}-\frac{\bar{x}(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}\right]}_{h_i}\underbrace{\left[\frac{(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}\right]}_{w_i}$$

$$= \sigma_\varepsilon^2\cdot\left[\underbrace{\cancel{\frac{\sum(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}}}_{0}\cdot\frac{1}{n}\right]-\sigma_\varepsilon^2\cdot\frac{\bar{x}\cancel{\sum(x_i-\bar{x})^2}}{\left[\sum(x_i-\bar{x})^2\right]^{\cancel{2}}}$$

$$\Rightarrow Cov(\hat{\alpha},\hat{\beta}) = \boxed{-\frac{\sigma_\varepsilon^2\cdot\bar{x}}{\sum(x_i-\bar{x})^2}}$$

משפט GM :
Gauss Markov Theorem ← אומד ה OLS הוא BLUE ($\hat{\beta}, \hat{\alpha}$ שמקיימים)

(כלומר הם בעלי השונות המינימלית, אומדים יעילים מבין האומדים הליניאריים חסרי ההטייה)

הוכחה:

ליניאריים בפרמטרים

א) נראה ש $\hat{\beta}$ ליניארי ביחס ל Y. ← אומד ליניארי מקיים:

$$\hat{\theta} = \sum a_i Y_i$$

הראינו ש w_i קבוע (משקולת) מכיוון ש x אינו סטוכסטי.
ואינו תלוי ב Y

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

דוגמא למודל שהוא לא ליניארי בפרמטרים:

$$Y_i = \frac{1}{\alpha + \beta X_i}$$

ב) חסר הטייה ← הראינו.

שיטת ההוכחה: ניקח את כל האומדים הליניאריים חסרי ההטייה.
נראה ש OLS הוא בעל השונות המינימלית.
(כמו שעשינו בדוגמא עם ממוצע ← אומד ליניארי).

ניקח נניח את $\breve{\beta}$:

$$\breve{\beta} = \sum q_i Y_i$$ כל אומד ליניארי

$$= \sum (w_i + \Delta_i) Y_i = \sum (w_i + \Delta_i)(\alpha + \beta X_i + \varepsilon_i)$$

נרצה להראות שזה שווה לאפס.

$$= \underbrace{\sum (w_i \alpha)}_{0} + \beta \underbrace{\sum w_i x_i}_{1} + \sum w_i \varepsilon_i + \sum \Delta_i \alpha + \beta \sum \Delta_i x_i + \sum \Delta_i \varepsilon_i$$

$$\Rightarrow \breve{\beta} = \beta + \alpha \sum \Delta_i + \beta \sum \Delta_i x_i + \sum (\Delta_i + w_i) \varepsilon_i$$

$$E(\breve{\beta}) = E(\beta) + E(\alpha \sum \Delta_i) + E(\beta \sum \Delta_i x_i) + \sum (\Delta_i + w_i) \underbrace{E(\varepsilon_i)}_{0}$$

הנחה קלאסית

לכן בשביל חוסר הטיה צריך להתקיים:

$$E(\tilde{\beta}) = \beta + K\underbrace{\sum \Delta_i}_{\text{הפ״ע}} + \beta \sum \Delta_i X_i = \beta$$

נדרוש:

$$\boxed{\begin{aligned} &\sum \Delta_i = 0 \\ &\sum \Delta_i X_i = 0 \end{aligned}}$$

* לציין למה מספיק לא לדרוש

$K\sum \Delta_i = -\beta \sum \Delta_i X_i$

לכן משפחת האומדים הליניאריים חסרי ההטיה ל-β מקיימת:

$$\tilde{\beta} = \sum q_i x_i$$

$$\sum \Delta_i = 0$$

$$\sum \Delta_i x_i = 0$$

נראה שונות מינימלית.

$$Var(\tilde{\beta}) = E(\tilde{\beta} - \beta)^2 = E\left(\sum (\Delta_i + w_i)\varepsilon_i\right)^2$$

$$= E\left(\sum (\Delta_i + w_i)^2 \varepsilon_i^2 + 2\sum\sum (w_i + \Delta_i)(w_j + \Delta_j)\overset{E}{\underbrace{\varepsilon_i \varepsilon_j}_{0\ \text{בהנחות}}}\right)$$

$$= \sigma_\varepsilon^2 \sum (\Delta_i + w_i)^2$$

$$= \sigma_\varepsilon^2 \left(\sum \Delta_i^2 + \sum \overset{\frac{(x_i - \bar{x})^2}{\cdots}}{w_i^2} + 2\sum w_i \Delta_i\right) \qquad 2\sum \overset{0}{x_i \Delta_i} - 2\sum \overset{0}{\Delta_i \bar{x}}$$

$$= \sigma_\varepsilon^2 \cdot \left[\frac{\cancel{\sum (x_i - \bar{x})^2}}{\left(\sum (x_i - \bar{x})^2\right)^{2}} + \sum \Delta_i^2 + 2\frac{\sum (x_i - \bar{x})\Delta_i}{\sum (x_i - \bar{x})^2}\right]$$

$$= \frac{\sigma_\varepsilon^2}{\sum (x_i - \bar{x})^2} + \sigma_\varepsilon^2 \cdot \sum \Delta_i^2 \quad \Rightarrow \quad \tilde{\beta} = \hat{\beta}_{OLS}$$

לכן נדרוש

שכל Δ_i שווה ל-0

ש עבורו נקבל אומד מינימלי.

שיעור מס 6 :

(I) מודל ללא חותך:

אם אני מריץ רגרסיה ללא חותך ← המודל לשונות משתנה (שימי לב ← ללא הישר לשרטוט המודל האמיתי).

$$\min_{\hat\beta} \sum \hat\varepsilon_i^2 = \sum (Y_i - \hat\beta X_i)^2$$

$$\frac{\partial}{\partial \hat\beta} \Rightarrow -2\sum_{i=1}^{n}(Y_i - \hat\beta X_i)X_i = 0$$

$$\Rightarrow \sum Y_i X_i = \hat\beta \sum X_i^2 \Rightarrow \hat\beta = \frac{\sum X_i Y_i}{\sum (X_i^2)}$$

* שימו לב מה היה קורה אם $\bar X = 0$ ומריצים מודל עם חותך.

(II) שינוי יחידות מדידה: ← טרנספורמציה לינארית

מה קורה אם מכפילים את Y ב c
את X ב c'

נכתוב את הרגרסיה החדשה:

$$(cY) = \hat\alpha' + \hat\beta'(c'X) + \hat\varepsilon'$$

$$\hat\beta' = \frac{\hat{Cov}(c'X, cY)}{\hat{Var}(c'X)} = \frac{c' \cdot c\, \hat{Cov}(X,Y)}{c'^2\, \hat{Var}(X)} = \frac{c}{c'}\hat\beta \rightarrow \text{במודל המקורי}$$

אם מעבירים את Y, X לאחוזים ← $c = c'$ ← המודל לשונות לא משתנה.

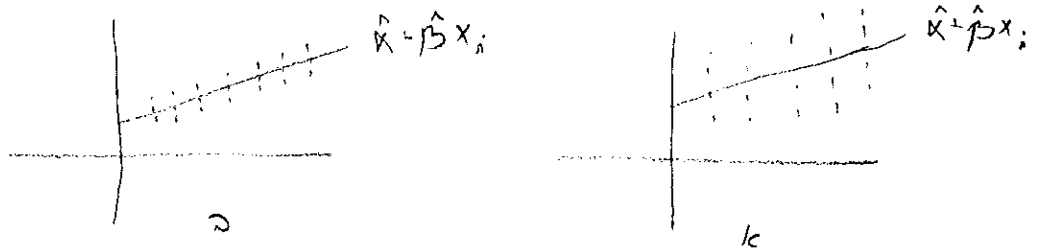
$$\hat\alpha' = (\overline{cY}) - \hat\beta'(\overline{c'X}) \Rightarrow \hat\alpha' = c \cdot \bar Y - \frac{c}{\cancel{c'}}\hat\beta \cdot \cancel{c'}\,\bar X$$

$$\Rightarrow c(\bar Y - \hat\beta \bar X) = c \cdot \hat\alpha$$

* תחשבו בבית מה קורה אם מוסיפים c.
* מה קורה ל $\sum \hat\varepsilon_i^2$

R^2 מדד טיב ההתאמה: (או "איכות הרגרסיה")

המדד מראה כמה מתוך השונות של המשתנה התלוי Y עצמו
להסביר המשתנה — המשתנים הבלתי. (במודל)



הקו במודל ב' מתאים יותר טוב כיוון שהפיזור של התצפיות סביבו קטן יותר.

האומד לשונות של Y

$$\frac{\sum(Y_i-\bar Y)^2}{n-1}=\frac{\sum(\hat Y_i+\hat\varepsilon_i-\bar Y)^2}{\cdot\cdot}=\frac{\sum(\hat Y_i-\bar Y)^2+\sum\hat\varepsilon_i^2+2\sum(\hat Y-\bar Y)\hat\varepsilon_i}{n-1}$$

$$=\frac{\sum(\hat Y_i-\bar Y)^2}{n-1}+\frac{\sum(\hat\varepsilon_i^2)}{n-1}+\frac{2\sum(\hat\alpha+\hat\beta x_i-\hat\alpha-\hat\beta\bar x)\hat\varepsilon_i}{n-1}$$

אני משתמש שעוברים דרך נקודת הממוצע

$$=\quad\downarrow\quad+\quad\downarrow\quad+\quad 2\hat\beta\underbrace{\cancel{\sum(x_i-\bar x)\hat\varepsilon_i}}_{0}$$

$$\Rightarrow\quad\underbrace{\sum(Y_i-\bar Y)^2}_{TSS}=\underbrace{\sum(\hat Y_i-\bar Y)^2}_{\text{Reg Sum of Squared}}+\underbrace{\sum(\hat\varepsilon_i^2)}_{\text{Error Sum of Squares}}\qquad \Big|:TSS$$

$$1=\underbrace{\frac{\sum(\hat Y-\bar Y)^2}{\sum(Y_i-\bar Y)^2}}_{R^2}+\frac{\sum(\hat\varepsilon_i)^2}{\sum(Y_i-\bar Y)^2}$$

שימו לב ש OLS מביא לידי מקסימום R^2.

* להראות ספלט,
* לפתור תרגיל ממבחן מועד א' סמסטר ב' 2009 שאלה פתוחה.

* אם קיבלנו $R^2=0.8$ ← נאמר שהמסבירים 80% מהשונות של Y. (במדגם)

* אין כל קשר בין R^2 לבין מובהקות הפרמטרים ולתנאים של ההנחות הקלאסיות.

* להראות בספר. * R^2 לא יכול להגיד לנו אם המודל הוא נכון

* דרך אחרת להראות את R^2:

$$R^2=\frac{\sum(\hat{Y}_i-\bar{Y})^2}{\sum(Y_i-\bar{Y})^2}=\frac{\sum\left(\hat{\alpha}+\hat{\beta}X_i-\hat{\alpha}-\hat{\beta}\bar{X}\right)^2}{\sum(Y_i-\bar{Y})^2}$$

$$=\hat{\beta}^2\frac{\sum(X_i-\bar{X})^2}{\sum(Y_i-\bar{Y})^2}=\frac{\hat{\beta}^2\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}\longrightarrow$$

עבור מודל עם חותך (כאשר הרגרסיה עוברת דרך נקודת הממוצע)

נציב את $\hat{\beta}^2$

$$R^2=\frac{(\hat{\sigma}_{XY}^2)^2}{(\hat{\sigma}_X^2)^{\cancel{2}}}\cdot\frac{\cancel{\hat{\sigma}_X^2}}{\hat{\sigma}_Y^2}=\frac{\hat{\sigma}_{XY}^2}{\hat{\sigma}_X^2\,\hat{\sigma}_Y^2}=\underline{\hat{\rho}^2}$$

מקדם המתאם פירסון

↓

זה קורה רק כשיש משתנה מסביר אחד

בבית: הראו כי $R^2_{XY}=R^2_{YX}$

$$\hat{Y}=\hat{\alpha}+\hat{\beta}X\neq\hat{X}=\hat{\gamma}+\hat{\lambda}Y$$

אם יש לנו 2 רגרסיות ← על מנת לדעת
מי מהן מסבירה טוב יותר את המשתנה התלוי ← "התופעה"
נשווה R^2. אבל ← ניתן להשוות R^2 רק כאשר
TSS זהה לשני המודלים.

* הוספת משתנים

נניח:

$Y_i = \hat{\alpha} + \hat{\beta} X_i$ מודל ראשון

$(Y_i - X_i) = \hat{\gamma} + \hat{\lambda} X_i$ מודל שני

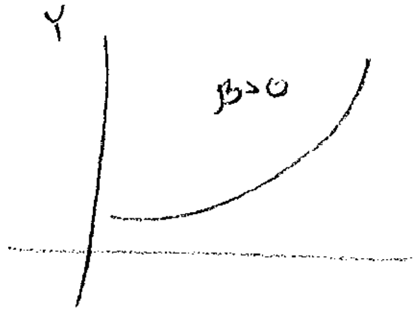
מודלים עם אותה אינפורמציה אבל נקבל R^2 שונה.

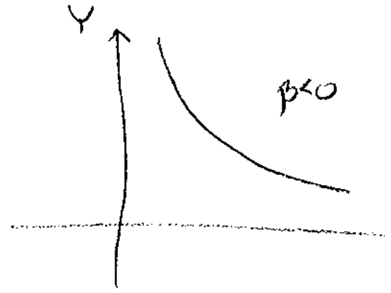
הדוגמא שהכי מתאים להראות זאת
זה טרנספורמציות לוגריתמיות;

טרנספורמציה לוגריתמית

נניח שיש קשר מהצורה הבאה: $Y_i = e^{\alpha+\beta X_i+\varepsilon_i}$

זהו קשר לא ליניארי. אבל ← בעזרת
טרנספורמציה של המשתנים נוכל לרשום בעזרת רגרסיה ליניארית
פרמטרים קשר לא ליניארי.



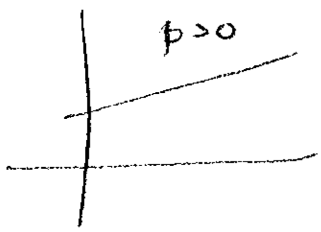


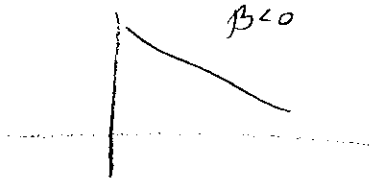
למה זה טוב?

אם נבצע טרנספורמציה ln על הצדדים של המשוואה, נקבל:

$\ln(Y_i) = \alpha+\beta X_i+\varepsilon_i$ ← קשר סמי לוגריתמי

ואז הגרפים יראו:





אבל! ניתן לבצע טרנספורמציה
רק עבור ערכים של $Y > 0$.

פרשנות למקדמים: (טרנספורמציות לוגריתמיות)

* הערה ← המודל לינארי בפרמטרים אבל לא לינארי במשתנים.

אין בעיה

לוגים

$x_1, x_2 > 0$

$$\log(x_1 x_2) = \log(x_1) + \log(x_2)$$

$$\log(x_1/x_2) = \log(x_1) - \log(x_2)$$

$$\log(x_1^c) = c\log(x_1)$$

$$\boxed{\log(x_1) - \log(x_0) \approx \frac{x_1 - x_0}{x_0} = \frac{\Delta X}{x_0}}$$

עבור שינוי קטן מספיק

מכפילים במאה

$$100\,\Delta\log X \approx \overbrace{\%\,\Delta X}^{\text{אחוז שינוי ב-}X}$$

שינוי X הופכי ביחידה אחת

אם ה שינוי ביחידת X, Y עולה ב $\beta \times 100$ אחוזים.

$$\log(y) = \beta_0 + \beta_1 X$$

$$\beta_1 = \frac{\Delta\log y}{\Delta X} \rightarrow \approx \frac{\Delta y}{y_0} \quad 0{,}08$$

שינוי ביחידה אחת $= 1$

אם ה שינוי באחוז X, Y עולה ב β אחוזים.

$$\log(y) = \beta_0 + \beta_1\log(x) \rightarrow \beta_1 = \frac{\Delta\log y}{\Delta\log x} \Rightarrow \approx \frac{\frac{\Delta y}{y_0}}{\frac{\Delta x}{x_0}} \rightarrow \frac{0{,}08}{0{,}01}$$

על כל שינוי
ב 1 אחוז
ב x, y משתנה
ב $\frac{\beta_1}{100}$ יחידות.

$$y = \beta_0 + \beta_1 \log(x)$$

$$\beta_1 = \frac{\Delta y}{\Delta \log x} = \frac{\Delta y}{\frac{\Delta x}{x_0}}$$

$\frac{\Delta x}{x_0} \rightarrow$ שינוי במאית

$1 = \frac{\Delta x}{x_0} \leftarrow$ אחוז