

נוסחאות

הממוצע של X והממוצע של Y (X ו-Y "גדולים"):

$$\bar{X} = \frac{\sum X_i}{n} \quad \bar{Y} = \frac{\sum Y_i}{n}$$

הממוצע לאחר טרנספורמציה ליניארית (הוספת קבוע לכל ערכי המשתנה ו/או הכפלהם בקבוע):

$$\overline{a + bX} = a + b\bar{X} \quad \overline{c + dY} = c + d\bar{Y}$$

סטיות מהממוצע:

$$x_i = X_i - \bar{X} \quad \sum x_i = 0 \quad \bar{x} = 0 \quad \bar{X} \text{ "גדול" מהממוצע}$$

$$y_i = Y_i - \bar{Y} \quad \sum y_i = 0 \quad \bar{y} = 0 \quad \bar{Y} \text{ "גדול" מהממוצע}$$

הנוסחאות הקושרות בין ה"קטנים" ל"גדולים":

$$\sum x^2 = \sum X^2 - n\bar{X}^2 \quad \sum y^2 = \sum Y^2 - n\bar{Y}^2 \quad \sum xy = \sum XY - n \cdot \bar{X} \cdot \bar{Y}$$

השונות של X במדגם היא S_X^2 וסטיית התקן של X במדגם היא S_X (השורש החיובי של השונות):

$$S_X^2 = \frac{\sum (X - \bar{X})^2}{n} \quad S_X^2 = \frac{\sum x^2}{n} \quad S_X^2 = \frac{\sum X^2}{n} - \bar{X}^2$$

ולאחר טרנספורמציה ליניארית:

$$S_{a+bX}^2 = b^2 S_X^2 \quad S_{a+bX} = |b| S_X$$

השונות המשותפת (covariance) של X ושל Y , במדגם:

$$S_{X,Y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{n} \quad S_{X,Y} = \frac{\sum xy}{n} \quad S_{X,Y} = \frac{\sum XY}{n} - \bar{X} \cdot \bar{Y}$$

ולאחר טרנספורמציות ליניאריות:

$$S_{a+bX, c+dY} = b \cdot d \cdot S_{X,Y}$$

מקדם המתאם (של פירסון) בין X לבין Y במדגם:

$$r_{X,Y} = \frac{S_{X,Y}}{S_X \cdot S_Y} \quad r_{X,Y} = \frac{\sum xy}{\sqrt{\sum x^2 \cdot \sum y^2}}$$

לאחר טרנספורמציות ליניאריות ערכו המוחלט לא משתנה,

$$|r_{a+bX, c+dY}| = |r_{X,Y}|$$

ורק אם אחד, ורק אחד, מהכופלים (b או d) הוא שלילי- סימנו מתהפך.

מקדם ההסבר (של פירסון) בין X לבין Y הוא $r_{X,Y}^2$ (מקדם המתאם בריבוע).

ולאחר טרנספורמציות ליניאריות הוא נשאר ללא שינוי:

$$r_{a+bX, c+dY}^2 = r_{X,Y}^2$$

רגרסה פשוטה (מסביר אחד וחותך). קו הרגרסה הוא $\hat{Y} = a + bX$.

עבור כל תצפית (X_i, Y_i) :

$$Y_i = a + bX_i + e_i$$

$$Y_i = \hat{Y}_i + e_i$$

$$e_i = Y_i - \hat{Y}_i$$

הסטייה מקו הרגרסה:

מקדמי הקו נאמדים משטת הריבועים הפחותים הרגילים (OLS), השיטה המביאה למינימום את סכום הסטיות הריבועיות, $\min \sum e_i^2$.

השיפוע:

$$b = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$b = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$b = \frac{\sum Y_i X_i - n \bar{Y} \bar{X}}{\sum X_i^2 - n \bar{X}^2}$$

$$b = r \frac{S_Y}{S_X}$$

$$b = \frac{n \sum Y_i X_i - (\sum Y_i)(\sum X_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

$$b = \frac{S_{XY}}{S_X^2}$$

$$b = \frac{(\sum Y_i X_i)/n - \bar{X} \cdot \bar{Y}}{(\sum X_i^2)/n - \bar{X}^2}$$

$$a = \frac{\bar{Y} \sum X_i^2 - \bar{X} \sum X_i Y_i}{\sum X_i^2 - \bar{X} \sum X_i}$$

$$a = \bar{Y} - b\bar{X}$$

החותך:

תכונת פירוק השונות: $S_Y^2 = S_{\hat{Y}}^2 + S_e^2$ וכן $\frac{\sum y_i^2}{n} = \frac{\sum \hat{y}_i^2}{n} + \frac{\sum e_i^2}{n}$

R^2 מקדם ההסבר (או מקדם ההסבר המרום, או מקדם הקביעה) מוגדר כפרופורציית השונות המוסברת:

$$R^2 = 1 - \frac{\sum e^2}{\sum y^2}$$

סכום הסטיות הריבועיות הוא $\sum e^2 = (1 - R^2) \sum y^2$.

רק ברגרסה פשוטה: $R^2 = r_{X,Y}^2$ (פרופורציית השונות המוסברת שווה למקדם ההסבר של פיסיון)

$$r^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{S_e^2}{S_Y^2} = \frac{S_{\hat{Y}}^2}{S_Y^2}$$

מקדם ההסבר: נסמן את $r_{X,Y}^2$ ב- r^2 .

$$r^2 = b^2 \frac{S_X^2}{S_Y^2}$$

$$r = b \frac{S_X}{S_Y}$$

נציב $S_{\hat{Y}}^2 = S_{a+bX}^2 = b^2 \cdot S_X^2$ ונקבל,

$$S_X = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

$$S_Y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n}}$$

$$b = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

נציב

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \cdot \sum y_i^2}}$$

ונקבל

$$r = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n \sum X_i^2 - (\sum X_i)^2][n \sum Y_i^2 - (\sum Y_i)^2]}}$$

או

$$r = \frac{S_{X,Y}}{S_X \cdot S_Y}$$

ומכאן,

$$S_{X,Y} = \frac{\sum x_i y_i}{n}$$

$$S_Y = \sqrt{\frac{\sum y_i^2}{n}}$$

$$S_X = \sqrt{\frac{\sum x_i^2}{n}}$$

האומדים לשוניות:

$$S^2 = \frac{\sum e_i^2}{n-1-1}$$

האומד ל- S^2 , האומד לשונות ההפרעה המקרית:

$$S_b^2 = \frac{S^2}{\sum x^2}$$

האומד ל- $V(b)$, האומד לשונות של האומד לשיפוע:

$$S_a^2 = S^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum x^2} \right] = S^2 \frac{\sum X^2}{n \sum x^2}$$

האומד ל- $V(a)$, האומד לשונות של האומד לחזית:

האומד ל- $COV(a,b)$, האומד לשונות המשותפת שבין האומד לחזית לבין האומד לשיפוע:

$$S_{a,b} = \frac{-\bar{X} S^2}{\sum x^2}$$

מבחני השערה ורווחי סמך:

$$t\text{-value} = \frac{b}{S_b} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

נוסדה שימושית:

$$\frac{b - b_{H_0}}{S_b} \sim t_{n-1-1}$$

מבחן השערה על b :

$$\frac{a - a_{H_0}}{S_a} \sim t_{n-1-1}$$

מבחן השערה על a :

$$b \pm t_{n-1-1, \alpha/2} \cdot S_b$$

רווח סמך ל- b :

$$a \pm t_{n-1-1, \alpha/2} \cdot S_a$$

רווח סמך ל- a :

ניבוי התוחלת של Y עבור $X = X_0$:

אומד נקודתי ל- $E(Y|_{X=X_0})$ הוא $\hat{Y}|_{X=X_0} = a + bX_0$.

רווח סמך ל- $E(Y|_{X=X_0})$ הוא $\hat{Y}|_{X=X_0} \pm t_{n-1-1, \alpha/2} \cdot S \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x^2}}$

ניבוי Y עבור $X = X_0$ (תחזית):

אומד נקודתי ל- $Y|_{X=X_0}$ הוא $\hat{Y}|_{X=X_0} = a + bX_0$.

רווח סמך ל- $Y|_{X=X_0}$ הוא $\hat{Y}|_{X=X_0} \pm t_{n-1-1, \alpha/2} \cdot S \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x^2}}$

$$X = a' + b'Y + e'$$

רגרסיה הפוכה $\hat{X} = a' + b'Y$:

$$b' = \frac{\sum xy}{\sum y^2} \quad a' = \bar{X} - b'\bar{Y}$$

$$b \cdot b' = r^2$$

$$e_i' = X_i - \hat{X}_i \quad R^2 = 1 - \frac{\sum e'^2}{\sum x^2} \quad S'^2 = \frac{\sum e'^2}{n-2}$$

רגרסה מרובה: מודל עם שני משתנים מסבירים ($k = 2$) וחותך $Y = a + b_{01.2}X_1 + b_{02.1}X_2 + u$

$$\hat{Y} = a + b_{01.2}X_1 + b_{02.1}X_2$$

$$b_{01.2} = \frac{\sum yx_1 \sum x_2^2 - \sum yx_2 \sum x_1x_2}{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2}$$

$$b_{02.1} = \frac{\sum yx_2 \sum x_1^2 - \sum yx_1 \sum x_1x_2}{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2}$$

$$a = \bar{Y} - b_{01.2}\bar{X}_1 - b_{02.1}\bar{X}_2$$

$$b_{02} = b_{02.1} + b_{12}b_{01.2}$$

$$b_{01} = b_{01.2} + b_{21}b_{02.1} \quad \text{הקשר בין מקדמי המרום למקדמי הפשוטה:}$$

$$S^2 = \frac{\sum e^2}{n - k - 1}$$

האומד לשונות ההפרעה המקרית (S^2):

האומדנים למקדמים במודל עם שני משתנים מסבירים וחותך:

$$S_{b_{01.2}}^2 = \frac{S^2}{(1 - r_{12}^2) \sum x_1^2}$$

$$S_{b_{02.1}}^2 = \frac{S^2}{(1 - r_{12}^2) \sum x_2^2}$$

$$S_{b_{01.2}, b_{02.1}} = \frac{-S^2 r_{12}^2}{(1 - r_{12}^2) \sum x_1x_2}$$

$$\frac{b_j - (b_j)_{H_0}}{S_{b_j}} \sim t_{n-k-1}$$

מבחן השערה על מקדם הרגרסה b_j :

$$H_0 : l_1b_1 \pm l_2b_2 = l_3$$

מבחן השערה על קשר ליניארי בין מקדמים:

$$l_1b_1 \pm l_2b_2 - l_3$$

$$\frac{l_1b_1 \pm l_2b_2 - l_3}{\sqrt{l_1^2 S_{b_1}^2 + l_2^2 S_{b_2}^2 \pm 2l_1l_2 \text{cov}(b_1, b_2)}} \sim t_{n-k-1}$$

$$a \pm t_{n-k-1, \alpha/2} \cdot S_a$$

רווח סמך ל- a :

$$b_j \pm t_{n-k-1, \alpha/2} \cdot S_{b_j}$$

רווח סמך ל- b_j :

$$\frac{R^2/k}{(1 - R^2)/(n - k - 1)} \sim F_{k, n-k-1}$$

בחינת מובחנות לרגרסה, $H_0 : r^2 = 0$

$$\frac{(\sum_{RES} e^2 - \sum_{UNRES} e^2) / m}{\sum_{UNRES} e^2 / (n - k - 1)} \sim F_{m, n-k-1}$$

מבחן WALT (לפי רגרסה מוגבלת):

$$\frac{(\sum_{UNRES} R^2 - \sum_{RES} R^2) / m}{\sum_{UNRES} (1 - R^2) / (n - k - 1)} \sim F_{m, n-k-1}$$

ואם ל-RES ול-UNRES יש אותו משתנה מוסבר

כאשר m הוא מספר ההצטרפות ב- H_0 ו- k הוא מספר המשתנים המסבירים ברגרסה הלא מוגבלת.

מבחן LM:

למבחנות הרגרסה: בר"מ a דחים H_0 אם $nR_e^2 > c_{k,a}^2$

למבחנות m מקדמים: מריצים את המודל המצומצם (בלעדיו) ומחשבים את $e = e_1, \dots, e_n$.

מריצים את e לפי כל המשתנים ומחשבים את מקדם ההסבר, R_e^2 . בר"מ a דחים H_0 אם $nR_e^2 > c_{m,a}^2$.

$$\bar{R}^2 = 1 - \frac{\sum e^2 / (n - k - 1)}{\sum y^2 / (n - 1)}$$

מקדם ההסבר המתוקן:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - k - 1}$$