

סטטיסטיקה 3- שיעור 6

משוואה מתוקנת

דוג':

$$\hat{y}_i = 60 + 1.2x_{1i} + 0.5x_{2i}$$

נניח כי x_1 - עוגות, נמדד בגרמים

נניח כי x_2 - קולה, נמדד בליטרים

יש לנטרל את סולם המדידה ע"י משוואה מתוקנת.

Handwritten notes on lined paper showing the derivation of the standardized regression equation:

$z \sim (0, 1)$ (with \hat{z} above the 0 and \hat{z} above the 1)

$\hat{y}_i = \hat{z} + \hat{\beta} \cdot x_i$

$\hat{z} = \bar{y} - r \cdot \frac{\hat{s}_y}{\hat{s}_x} \cdot \bar{x} = 0 = \hat{z}$

(Arrows point from \bar{y} to 0 and from \bar{x} to 0. A red bracket underlines the term $r \cdot \frac{\hat{s}_y}{\hat{s}_x} \cdot \bar{x}$.)

$\hat{\beta} = r \cdot \frac{\hat{s}_y}{\hat{s}_x} \leftarrow 1 = \hat{\beta} = r$

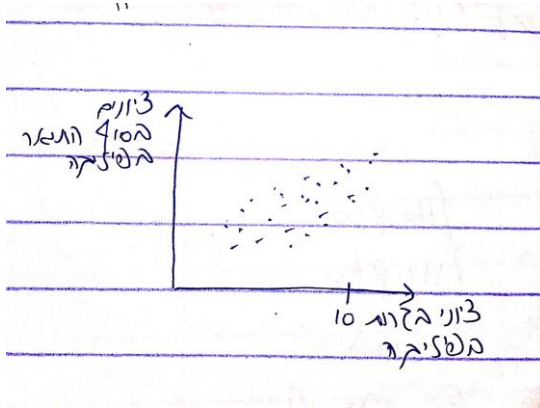
(A red box contains the final standardized equation: $z_{yi} = r \cdot z_{xi}$)

Additional handwritten notes at the bottom left: "אחוז סכום" and "מכפלה והתקן".

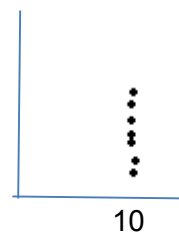
קיצוץ תחום

$$r_{(x,y)} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)\hat{s}_x\hat{s}_y}$$

דוג':



שאלה: מה המתאם בין ציון בגרות 10, לציון בסוף התואר?



ייראה כך:

תשובה: 0, אין קשר.

שאלה: מה המתאם בין ציון בגרות 9-10, לציון בסוף התואר?

תשובה: סביר שנקבל חלש בגלל שאין מספיק לקו הניבוי, מקום מאוד צר, הקשר לא יכול להתפתח.

- הקיצוץ פוגע- אך תלוי מהי מידת קיצוץ הערכים.
- קבלת החלטות על פי קיצוץ תחום- בדרך כלל החלטות שגויות

הסקה לאוכלוסייה:

- ממוצע הטעויות = 0

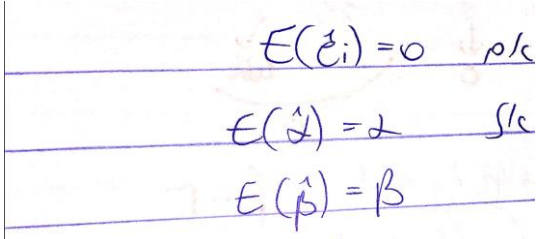
$$\sum_{i=1}^n (y_i - \hat{y}) = 0$$

קו המדגם הוא קו של ממוצעים- אומדן חסר הטיות לקו האוכלוסייה, לכן ההסקה אפשרית.

הנחות במונחי y

1. דגימה מקרית: תצפיות ב"ת
2. שוניות המשתנה התלוי שוות עבור כל ערך של x
3. התלוי מתפלג נורמלית עבור ערך של x

טעויות:

1. טעויות ב"ת
2. 

$\hat{\alpha}, \hat{\beta}$ - אומדנים חסרי הטיות ויעילים ל α, β בהתאמה

3. שוניות הטעויות שוות עבור כל ערך של x
4. טעויות מתפלגות נורמלית עבור כל ערך של x

הנחות(לכל ההשערות ברגרסיה פשוטה)

1. תצפיות ב"ת

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0$$

עבור כל $i \neq j$

2. $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

עבור כל ערך קבוע x_i .

(בהנחה זו מתבטאות הנחה 2,3,4)

מבחנים:

1. מתאם:

השערות:

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

מבחן Z

(לפי הפלט)

(האם יש מתאם באוכי')

2. טיב הניבוי/ מובהקות המשוואה

השערות: $H_0 : \beta = 0$

$$H_1 : \beta \neq 0$$

מבחן F

(האם יש פרופורציה מוסברת באוכי')

3. שיפוע

השערות: $H_0 : \beta = 0$

$$H_1 : \beta \neq 0$$

מבחן t

(האם יש שיפוע באוכי')

- אם דחיתי במבחן 1- אדחה בשני המבחנים האחרים, כך גם אם קיבלתי במבחן 1. מתקיים רק ברגרסיה פשוטה.

פירוט מבחן טיב הניבוי/מובהקות המשוואה

מקור השונות	סכומי ריבועי סטיות SS	דרגות חופש df	ממוצעי סכומי ריבועים ms	תוחלת E	F סטטיסטי
קו הרגרסיה	$SSREG = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ $= r^2 \cdot SSy$	1	$MSREG = \frac{r^2 \cdot SSY}{1}$	$\sigma^2 e \cdot + \beta^2 SSX$	$F = \frac{MSREG}{MSRES} = \frac{r^2 \cdot (n-2)}{1-r^2} \stackrel{H_0}{\sim} F_{1,n-2}$
טעויות/ שאריות	$SSRES = \sum_{i=1}^n (y_i - \hat{y})^2$ $(1-r^2)SSy$	n-2	$MSRES = \frac{(1-r^2) \cdot SSy}{n-2}$	$\sigma^2 e$	
סה"כ	$SSY = \sum_{i=1}^n (y_i - \bar{y})^2$ $SSy = SS\hat{y} + SSRES$	n-1			

איזור דחייה:

$$F_{1,n-2} > F_{1,n-2}(\alpha)$$

מסקנה: אם דוחים אז הפרופורציה המוסברת שונה מאפס באוכל' ואז יש שיפוע, יש מתאם, יש ניבוי, יש משוואה.

- תחת נכונות H0 MSREG הוא אומד חסר הטיה לשונות הטעויות..
- תחת נכונות H1 MSREG הוא אומד מוטה לשונות הטעויות משום שמכיל אפקט.

משוואה מתוקננת, קיצוץ תחום, לרגרסיה פשוטה, אח"ה, אומדן רגרסיה, מודל הרגרסיה, מבחן F, הסקה לאוכלוסיה: מובהקות המשוואה