

השפעה לא ליניארית של משתנים

$$\text{Income} = a + b_1 * \text{Age} + b_2 * \text{Education} + e$$

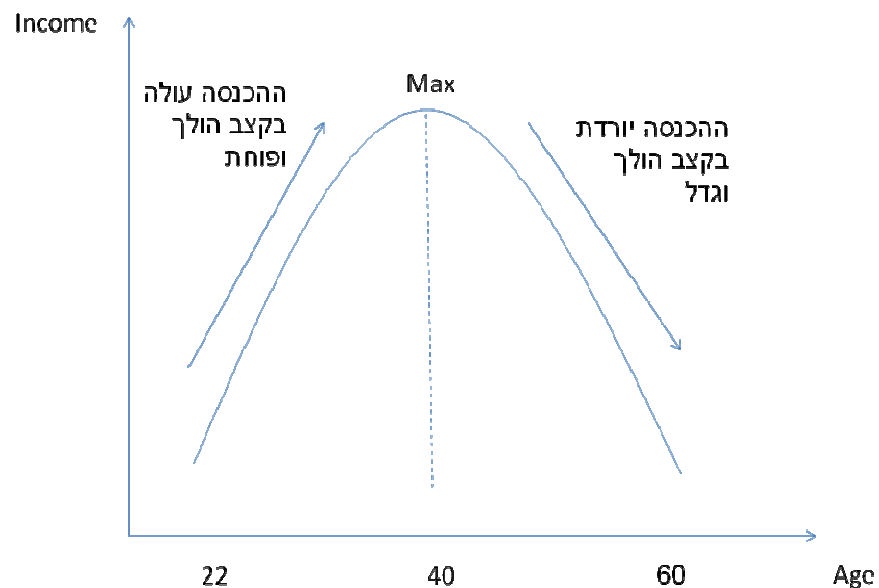
החוקר הריץ רגרסיה (ליניארית בשיטת הריבועים הפחותים) וקיבל:

$$\text{Income} = 1000 + 100 * \text{Age} + 500 * \text{Education} + e$$

- כל שנת גיל מעלה את ההכנסה ב-100.
- כל שנת השכלה מעלה את ההכנסה ב-500.

הבעייתיות במודל זה

לא הגיוני שמעבר מגיל 20 ל-21 הוא כמו השפעה מגיל 60 ל-61. ההשפעה היא לא תמיד השפעה ליניארית.



נאמוד רגרסיה שתיתן לי את האופציה לאמוד U הפוך (השפעה לא ליניארית). איזו פונקציה יכולה לתאר את המצב הזה? ← פונקציה ריבועית $y = ax^2 + bx + c$

כדי למצוא נק' קיצון נגזור את הפונקציה ונשווה ל-0.

$$x = (-b) / 2a \leftarrow y' = 2ax + b = 0$$

על מנת לדעת אם זאת נקודת מינימום או מקסימום, נגזור נגזרת שנייה:

$$y'' = 2a$$

- אם $a > 0$ זוהי נקודת מינימום (פרבולה צוחקת).
- אם $a < 0$ זוהי נקודת מקסימום (פרבולה בוכה).

גיל, השכלה, הכנסה, וותק- אלו משתנים שסימן שלילי הוא לא רלוונטי עבורם. אין משמעות להשכלה שלילית. כדי ש-X (משתנה מסביר) יהיה חיובי הסימן של b צריך להיות ההיפך מהסימן של a.

דוגמא:

חוקר משער שהגיל מעלה את ההכנסה ולאחר מכן מוריד אותה. הוא אמד את הרגרסיה הבאה:

$$\text{Income} = a + b_1 * \text{Age} + b_2 * \text{Age}^2 + b_3 * \text{Education} + e$$

והוא קיבל את התוצאות הבאות:

$$\text{Income} = 200 + 240 * \text{Age} - 3 * \text{Age}^2 + 200 * \text{Education} + e$$

באיזה גיל ההכנסה היא מקסימאלית?

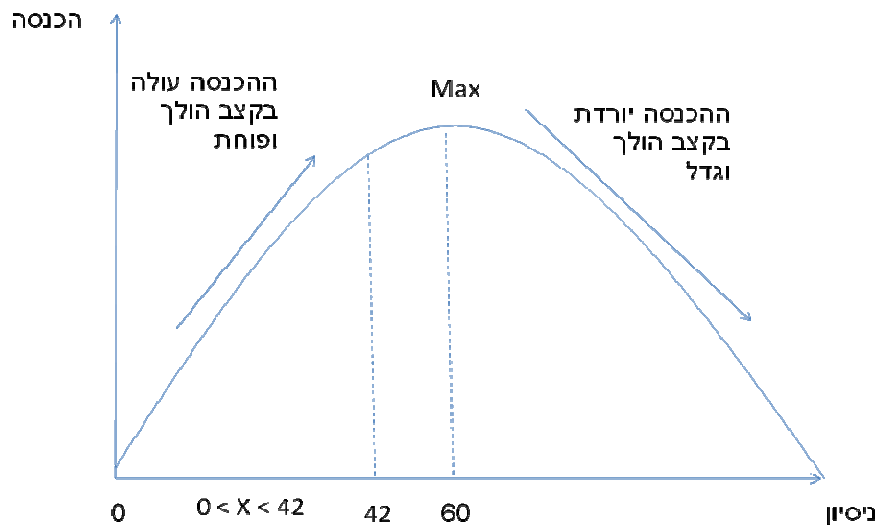
המקדם של Age^2 הוא (-3). אם הוא שלילי זאת פונקציה מקסימום.

- $a = (-3)$
- $b = 240$

$$\text{Age [Max]} = (-240) / (2 * -3) = 40$$

ההכנסה מקסימאלית בגיל 40.

ניסיון



ככל שהניסיון עולה ההכנסה עולה בקצב הולך ופוחת. אני רוצה לאמוד רק את ההשפעה החיובית של ניסיון על הכנסה (בקצב קטן) בלי נקודת קיצון. הדרך הפשוטה ביותר היא להריץ רגרסיה כמו שעשינו קודם, אבל להראות שנקודת הקיצון היא מחוץ לטווח.

דוגמא

חוקר אסף תצפיות על הכנסת הפרט ושנות הניסיון שלו, כאשר שנות הניסיון הן בין 0 ל-42. נאמדה הרגרסיה הבאה:

$$\text{Income} = 3000 + 240X - 2X^2$$

באיזו שנת ניסיון ההכנסה היא מקסימאלית?

$$X = (-240) / (-4) = 60$$

בתחום שבין 0 ל-42 מתקיים שהניסיון משפיע חיובית על ההכנסה אך בקצב הולך ופוחת. כלומר, כל הטווח 0-42 הוא חיובי ולכן ההכנסה המקסימאלית היא בשנת ניסיון 42.

כל עוד אני רוצה אני יכולה להשתמש בפונקציה ריבועית רגילה, אך אני צריכה להוציא את נק' הקיצון מחוץ לתחום ההגדרה.

נתונה הטבלה הבאה:

ניסיון	הכנסה
0	3000
1	3238
2	3472
3	3702
5	4150
10	5200
20	7000
30	8400
40	9400
50	10000

במעבר מהשנה ה-1 ל-2 הפרט הרוויח עוד 238 ₪. לעומת זאת המעבר מ-40 ל-50 נוסף לפרט בממוצע לשנה 60 ₪. מכאן, שככל שהניסיון עולה ההכנסה עולה אבל הקצב הולך ופוחת ← נגזרת של X חיובית ← נגזרת של X^2 שלילית.

כאשר המקדם של X^2 הוא חיובי יש פונקציה מינימום. אם אני רוצה להתייחס רק לחלק היורד של הפונקציה אני צריכה להוציא את נקודת המינימום אל מחוץ לתחום ההגדרה (כמו שעשינו בדוגמא של הניסיון- בפונקציה מקסימום).