

## 1 ריגרסיה – מבוא

בסטטיסטיקה, ניתוח ריגרסיה הוא שם כולל למשפחה של מודלים סטטיסטיים להערכת הקשרים בין משתנים. המשותף לכל המודלים הוא קיומם של משתנה מוסבר (המכונה לעיתים בשם המשתנה התלו) ומשתנה מסביר אחד או יותר (המכונים לעיתים בשם המשתנים הבלטי תלויים או המשתנים המנביים). באמצעות מודל ריגרסיה ניתן ללמוד כיצד ערכו של המשתנה המוסבר משתנה כאשר חל שינוי בערכו של אחד המשתנים המסבירים, וערכי שאר המשתנים המסבירים נשאים קבועים. עם זאת, אין בכך די להסביר סיבתיות: השינוי בערכו של המשתנה המסביר לא בהכרח גורם לשינוי בערכו של המשתנה המוסבר.

### 1.1 מודל מתמטי

נניח מודל:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i} + u$$

ונניח כ- $n$  תצפיות. אז:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & x_{2,1} & \dots & x_{k,1} \\ 1 & x_{1,2} & x_{2,2} & \dots & x_{k,2} \\ \vdots & \dots & & & \\ 1 & x_{1,n} & x_{2,n} & \dots & x_{k,n} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

### 1.2 אומד הריבועים הפחותים (OLS)

רוצים להביא למינימום את סכום השאריות. כלומר:

$$\min_{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

### 1.3 טבלה

תא ראשון	תא אמצעי	תא אחרון
תא אחרון	תא אמצעי	תא ראשון

• פריט ראשון

• פריט שני