

הגישה במאמר – רקע כללי

- הגישה המוצגת במאמר מבוססת על התאמת וגרסיה עבור התוחלת ועבור סטיית התקן של המדדים השונים כתלות בגיל העובר.
- הנחה: בכל שבוע הירionario, המשטנה שאנו מודדים מתפלג נורמלית עם **תוחלת ו-סטיית התקן**, כאשר שתיהן תלויות בגיל ההירionario.
- מוטיביציה: לחשב עקומת אחוזון. שכן ידוע:

$$\text{centile} = \text{mean} + K * \text{SD}$$

באשר K הוא score-z של האחוזון הרצוי בהתפלגות נורמלית סטנדרטית.

- הערה: אם רואים שסטיית התקן **גדלה מאוד עם הגיל**, זה יכול להקשות על ההתאמה (הפיזור המשטנה במצב לא ליניארי). במקרה זה, במאמר מציעים לנסות טרנספורמציה לוגריתמית על המשטנה המוסבר שלנו. אחרי שמחשבים את עקומות האחוזונים על סקלת הלוג, פשוט עושים **אנטילוג** כדי לחזור ליחידות המקוריות.
- **אבל**, אם הטרנספורמציה יוצרת **עיוות בתפלגות**, למשל "זנב" ארוך לערבים נמוכים, אז עדיף לוותר על הלוג ולעבוד עם הערכים המקוריים.

איך בונים את מודלי הרגression?

1. מקשרים בין המשתנה המוסף לגיל העובר בעזרת פולינום ממולה שלישית:

$$Y = a + b \cdot \text{age} + c \cdot \text{age}^2 + d \cdot \text{age}^3 + \text{error}$$

כasher מוצאים את המקדמים a, b, c, d בעזרת שיטת הריבועים הפחותים.

2. בודקים האם המקדם של age^3 , במקורה שלנו d , מובהק סטטיסטי, כלומר האם יש לו תרומה אמיתית. הקритריון לכך:

$$\frac{|d|}{SE(d)} < 2$$

3. אם אי השווון הב"ל מתקיים, אז $-d$ אין תרומה אמיתית. אז מסירים אותו. עושים התאמת חדשה, הפעם עם פולינום ריבועי:

$$Y = a + b \cdot \text{age} + c \cdot \text{age}^2$$

4. שוב בודקים עם c מובהק. אם לא אז נסיר גם אותו ונישאר עם קו ישר.

דיברנו רק על התוחלת, מה עם המודל לסטיתת תקן?

1. לוקחים כל תצפית ומחסרים ממנה את הערך הצפוי בהתאם לשבוע הנוכחי

$$r_i = Y_i - \mu(GA_i)$$

כasher זו הוא ממד קלשו נצפה (היקף בטן, היקף ראש וכו') $\mu - (GA_i)$ היא הצבה בנוסחה שהתקבלה בעזרת השלב הקודם.

2. לוקחים ערך מוחלט ומכפילים ב 1.25

$$s_i = 1.25 \times |r_i|$$

3. נשאר לבדוק האם הפיזור משתנה עם הגיל?

נציר גרף של s מול GA

אם הנקודות מפוזרות בלי מגמה ברורה אז סימן שהפיזור קבוע ולכן SD קבוע.

אם הנקודות עלות עם הגיל אז סימן שסטיית התקן גדלה עם הגיל ולכן צריך מודל שמנבא את SD לפי GA

4. אם התשובה לסעיף 3 חיובית אז נעשה רגסיה של s על GA

$$s_i = \alpha + \beta \times GA_i + \varepsilon_i$$

איך בודקים שהמודלים באמת מתאימים לנתונים?

ראשית, נחשב ציוני Z לכל תצפית. לאחר מכן, ישן שתי דרכים לבדוק האם המודלים מתאימים לנתונים

1. Normal Probability Plot: משרטטים גרף שבו על הציר האופקי שמים את הערכים הצפויים מהתפלגות נורמלית, ועל הציר האנכי את ציוני ה-Z. אם הנקודות יושבות פחות או יותר על קו ישר סימן שהתפלגות ה-Z שלנו אכן נורמלית והמודל מתאים טוב

2. צוני Z מול גיל הריון: נصفה לפיזור אקריאי סביב 0