

## מודלים סטטיסטיים ויישומיהם 52518 תשע"ח – תרגיל 4

להגשה עד 27.11.17 בשעה 23:55

1. נתון מדגם מקרי של נתונים:

$$X = [-9.8, -6.2, -1.7, 9.1, 7.4, 9.4, -3.6, -10.4, -5.4, -3.7]$$

א. חשבו ושרטטו ידנית את פונקציית ההתפלגות האמפירית  $\hat{F}_n$

ב. חשבו את סטטיסטי קולמוגורוב-סמירנוב-ליליפורס, המוגדר כ-

$$D = \max_{t \in [-\infty, \infty]} \left| \hat{F}_n(t) - \Phi\left(\frac{t - \bar{X}}{s}\right) \right|$$

2. בחברת 'סוכר סגול' נערך ניסוי בתרופה להעלאת רמת הסוכר בדם. 10 נסיינים השתתפו בניסוי: 5 קיבלו את התרופה

1- 5 קיבלו פלצבו. הנתונים בדלהלן:

Drug: 2.59, 3.75, 2.08, 2.23, 2.19

Placebo: 2.65, 0.78, 2.46, 2.45, 1.19

א. מצאו את וקטור הדרגות של הנתונים.

ב. חשבו ידנית את סטטיסטי סכום הדרגות של ווילקוקסון ( $W_S$ ).

ג. חשבו ידנית את התוחלת והשונות של  $W_S$ .

ד. נגדיר את הסטטיסטי

$$Z = \frac{W_S - E[W_S]}{\sqrt{Var(W_S)}}$$

השתמשו בקירוב נורמלי עבור  $Z$  לחשב p-value חד-צדדי לבדוק את ההשערה  $H_0$ : לתרופה אין השפעה על רמת הסוכר לעומת ההשערה  $H_1$ : התרופה גורמת לעלייה ברמת הסוכר בדם.

3. השאלה הזאת עוסקת בסטטיסטי סכום הדרגות של ווילקוקסון.

א. כתבו פונקציה ב-R אשר מקבלת גדלים של שתי קבוצות  $n_1, n_2$  ומחשבת את  $P(W = w)$  לכל ערכי  $w$  האפשריים על ידי מנייה ישירה של כל התת-מדגמים האפשריים בגודל  $n_2$  מתוך המספרים  $1, \dots, N$ , כאשר  $N = n_1 + n_2$ . הריצו את הפונקציה עבור המקרה  $n_1 = n_2 = 5$ .

ב. כתבו פונקציה נוספת, המממשת את נוסחת הרקורסיה שהוצגה עבור מספר תתי-המדגמים המקיימים  $W = w$ . הריצו את הפונקציה עבור המקרה  $n_1 = n_2 = 5$ . השוו עם התוצאה של הסעיף הקודם (אמור להיות זהה).

ג. השתמשו בתוצאות של הסעיפים הקודמים לחשב p-value מדויק עבור הנתונים של שאלה 2. השוו מול התוצאה שקיבלתם בשאלה 2 ע"פ קירוב נורמלי.

4. בכיתה הצגנו את הנוסחא הבאה לאומד פונקציית צפיפות:

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right)$$

כאשר  $K$  הינה פונקציה אי-שלילית נתונה המקיימת  $\int_{-\infty}^{\infty} K(u) du = 1$  כגון:

$$K_1(u) = \frac{1}{2} I(|u| \leq 1), \quad K_2(u) = \frac{3}{4} (1 - u^2) I(|u| \leq 1)$$

א. כתבו פונקציה ב-R הקוראת מדגם של נתונים ומחשבת את  $\hat{f}_n(x)$ .

ב. השתמשו בפונקציה rgamma ב-R כדי ליצור מדגם מקרי בגודל 200 מהתפלגות גמא  $\Gamma(3,1)$ . השתמשו בפונקציה שכתבתם כדי לחשב את  $\hat{f}_n(x)$  עבור הנתונים האלה. שרטטו את  $\hat{f}_n(x)$  ואת פונקציית הצפיפות האמתית באותו גרף. בצעו זאת עבור  $K_1$  ו- $K_2$  הנתונים מעלה ועבור הערכים הבאים של  $h$ : 0.2, 0.4, 0.6, 0.8. דונו בתוצאות.