תרגיל בית 5

שאלה 1:

יהי $X_1,....,X_n$ מדגם מקרי מהתפלגות Weibull עם הפרמטר $\theta>0$ ו- $\alpha>0$ ידועה.

$$F_X\left(x\right)=1-e^{-\frac{1}{\theta}x^{\alpha}},x>0$$

א) הראו כי פונקציית הצפיפות היא:

$$f_X(x)=\begin{cases}\dfrac{1}{\theta}\alpha \mathrm{x}^{\alpha\text{-}1}\mathrm{e}^{-\frac{1}{\theta}\mathrm{x}^{\alpha}} & x>0\\ 0 \quad \text{otherwise} & \end{cases}$$

ב) מהו אומד נראות מרבית ל- θ?

i. האם אומד נראות מרבית ל- θ הינו אומד חסר הטיה?

ii. האם הנו עקיב? ענו ללא חישוב ה-MSE של אומד נראות מרבית ל- θ. הסבירו.

iii. חשבו MSE של אומד נראות מרבית ל- θ. ענו על סעיף ב׳-(ii) על סמך ה-MSE שקיבלתם.

הדרכה: הראו כי $Y_i=X_i^{\alpha}\sim\exp\left(\frac{1}{\theta}\right)$.

ג) הוצע לאמוד את θ ע״י $\tilde{\theta}=Y_1=X_1^{\alpha}$.

i. האם אומד זה חסר הטיה?

ii. האם הוא עקיב? בדקו לפי ההגדרה של עקיבות.

iii. חשבו MSE של האומד המוצע.

ד) איזה אומד ל- θ עדיף לפי קריטריון MSE: $\tilde{\theta}$ או אומד נראות מרבית?
האם התוצאה שקיבלתם היא הגיונית? הסבירו.

ה) מהו אומד נראות מרבית ל- $\frac{1}{\theta}$?

שאלה 2:

בהרצאה עבדנו עם נתונים על אורכי זמן של ניתוחים (X). תחת ההנחה שהנתונים באים מהתפלגות $Gamma(\alpha,\beta)$, מצאנו את אומדני הנראות המירבית לפרמטרים α,β ולהסתברויות $p_1=P(Y_2<4),p_2=P(Y_3>8)$ כאשר $Y_k=\sum_{i=1}^{k}X_i$.

<u>חלק I</u>

בתרגיל בית קודם מצאתם אומדים ואומדנים בשיטת המומנטים לפרמטרים α,β. על סמך התוצאות שקיבלתם, מצאו את האומדים והאומדנים להסתברויות p_1 ו – p_2 . פרטו ונמקו.

חלק II

ההשוואה התאורטית לפי המדדים שלמדנו בין האומדים לפי שיטת הנראות המירבית ולפי שיטת המומנטים עשויה להיות מאתגרת במקרה זה (זכרו שאין צורה סגורה לאומד נראות מירבית). בשאלה זו נשתמש בסימולציה בשביל ללמוד על התכונות של האומדים ובשביל להשוות ביניהם.

הגרילו N מדגמים בגודל n מהתפלגות $Gamma(\alpha,\beta)$. בחרו את הערך של N כך שיהיה לפחות 1000, את הערך של α מהקבוצה $\{4.3, 4.4. 4.5, 4.6, 4.7\}$ ואת הערך של β מהקבוצה $\{1.8, 1.9, 2, 2.1, 2.2\}$. בצעו זאת פעם אחת עבור n=10, פעם אחת עבור n=100 ופעם אחת עבור n=1000. עבור כל מדגם חשבו את הערכים של האומדנים לפרמטרים α,β בשיטת הנראות המירבית ובשיטת המומנטים, $\hat{\alpha}_{MLE}, \hat{\alpha}_{moments}, \hat{\beta}_{MLE}, \hat{\beta}_{moments}$. בנוסף, חשבו את האומדנים בשיטת הנראות המירבית ובשיטת המומנטים להסתברויות p_1, p_2, שיסומנו ב- $\widehat{p_1}_{MLE}, \widehat{p_2}_{MLE}, \widehat{p_1}_{moments}, \widehat{p_2}_{moments}$. <u>שימו לב להנחיות והדרכה בסוף התרגיל</u>.

כתבו מהם ערכי α,β שבחרתם, וחשבו עבורם את הערכים של p_1, p_2 (בעזרת הפונקציה pgamma ב-R).

א. הציגו את החלק בקוד המממש את הסימולציה ומחשב את הוקטורים המכילים את האומדנים לפרמטרים.

ב. עבור כל אחד מגדלי המדגם, חשבו את האומדן להטיה ול-MSE של כל אחד מהאומדים כאשר הערכים של α,β הם כפי שבחרתם בסימולציה. רשמו את האומדנים למדדים אלו בטבלאות הבאות (עגלו ל-3 ספרות אחרי הנקודה). הציגו קטע קוד ב-R לחישוב האומדנים להטיה ול-MSE על-סמך וקטור של אומדנים שהתקבל בסימולציה.

הטיה	$\hat{\alpha}_{moments}$	$\hat{\alpha}_{MLE}$	$\hat{\beta}_{moments}$	$\hat{\beta}_{MLE}$	$\widehat{p_1}_{moments}$	$\widehat{p_1}_{MLE}$	$\widehat{p_2}_{moments}$	$\widehat{p_2}_{MLE}$
n=10								
n=100								
n=1000								

MSE	$\hat{\alpha}_{moments}$	$\hat{\alpha}_{MLE}$	$\hat{\beta}_{moments}$	$\hat{\beta}_{MLE}$	$\widehat{p_1}_{moments}$	$\widehat{p_1}_{MLE}$	$\widehat{p_2}_{moments}$	$\widehat{p_2}_{MLE}$
n=10								
n=100								
n=1000								

על סמך התוצאות שבטבלאות, ענו על הסעיפים הבאים:

ג. ענו לגבי כל אחד מהאומדים <u>ונמקו היטב</u>:
1. האם נראה כי האומד חסר הטיה? התייחסו לכל אחד מגדלי המדגם שבטבלה (10, 100, ו-1000).
2. האם נראה כי ההטיה של האומד מושפעת מהגדלת המדגם? אם כן, כיצד?
3. האם נראה כי ה-MSE של האומד מושפע מהגדלת המדגם? אם כן, כיצד?
4. האם הטבלה נותנת עדות כי האומד עקיב?

ד. השווה את האומד בשיטת הנראות המירבית לאומד בשיטת המומנטים עבור כל אחד מהפרמטרים α,β,p_1,p_2, ועבור כל אחד מגדלי המדגם 10, 100, ו-1000, תוך התייחסות לשאלות הבאות:
תחת ההנחה שהערכים האמיתיים של α,β הם כפי שבחרתם: האם נראה שקיים הבדל בין האומד לפי שיטת הנראות המירבית לבין האומד לפי שיטת המומנטים מבחינת ההטיה? מבחינת ה-MSE? האם אומד אחד נראה עדיף על-פני השני משיקולי MSE?
מה קורה להבדלים בין המדדים של שתי השיטות כאשר גודל המדגם גדל?

<u>הנחיות והדרכה:</u>

הגדירו set.seed(a) כך ש-a הינו 4 ספרות אחרונות של ת.ז. המגיש.

לקובץ פתרון התרגיל יש להעתיק את הקוד אותו הרצתם כולל הגדרת set.seed.

הדרכה לכתיבת הסימולציה:

מומלץ להיעזר בקטעי קוד R שהוצגו בהרצאות. <u>שימו לב</u>: בהרצאה הוצגה דרך לפתרון נומרי של משוואה כאשר ידוע שהפתרון נמצא בטווח מסוים. בסימולציה כדאי לחפש את הפתרון בטווח מספיק גדול, למשל בטווח (0.1, 100).

מכיוון שהתוצאות אקראיות וייתכנו ערכים מחוץ לטווח, ניתן להרחיב אותו במידת הצורך. תוכלו לזהות שבחרתם טווח לא מספיק רחב אם תקבלו את הודעת השגיאה הבאה:

```
Error in uniroot(dif.gam, c(0.1, 100)) :
  f() values at end points not of opposite sign
```

שאלה 3:

נתון מדגם מקרי פשוט $Y_1,\dots,Y_5,Y_6$ מהתפלגות נורמלית סטנדרטית $N(0,1)$. נגדיר $W=\sum_{i=1}^{5}Y_i^2\,,\ \bar{Y}_5=\frac{\sum_{i=1}^{5}Y_i}{5}$

1. מצאו כיצד מתפלגים המשתנים המקריים הבאים. נמקו.
   א) W
   ב) $W+Y_6^2$
   ג) $5\left(\bar{Y}_5\right)^2+Y_6^2$
2. הסיקו על סמך ההתפלגויות שמצאתם בסעיף הקודם מהי התוחלת ומהי השונות של כל אחד מהמשתנים המקריים.
3. חשבו את תוחלות המשתנים המקריים ישירות, ללא שימוש בתכונות ההתפלגות שמצאתם, ובדקו שמתקבלת אותה התוצאה.

שאלה 4:

הוכיחו כי:

$$\chi^2_{(1),1-\alpha}=\left(z_{1-\frac{\alpha}{2}}\right)^2$$