תרגיל בית 4 - אמידה נקודתית ותכונות האומדים

שאלה 1:

נתון מדגם מקרי $X_1,...,X_n$ מהתפלגות $Unif\left[0,\theta\right]$. שאלה זו הינה המשך הדוגמה ממצגת 4, שקפים 21-25.

ניתן להיעזר בחישובים שנעשו בהרצאה.

א) מהו הקבוע c שעבורו לאומד $T=cX_{(n)}$ יש MSE מינימלי לכל $\theta>0$?

ב) מלאו את הטבלה הבאה. איזה מבין האומדים בטבלה עדיף על-פני כל אחד מהאומדים האחרים בטבלה משיקולי MSE?

האומדן עבור המדגם הנתון במצגת 4, שקף 21	MSE	האם עקיב?	האם חסר הטיה?	האומד	אומד ל- θ
					אומד בשיטת המומנטים
					אומד נראות מירבית
					θ^* שהוגדר במצגת 4, שקף 25
					$cX_{(n)}$ עם c מסעיף א׳.

שאלה 2:

יהי $X_1,...,X_n$ מדגם מקרי פשוט מהתפלגות נורמלית $N\left(\mu,\sigma^2\right)$.

א) מהו אומד נראות מרבית לשונות σ^2 כאשר התוחלת μ ידועה? האם הוא חסר הטיה?

ב) מהו אומד נראות מרבית לתוחלת μ כאשר השונות σ^2 ידועה?

ג) כעת נניח כי $\mu=\sigma=\theta$, כלומר $X_i \sim N\left(\theta,\theta^2\right)$.

הוצעו שני אומדים לשונות θ^2:

$$S_1=\bar{X}^2, \qquad S_2=\frac{\sum_{i=1}^{n}X_i^2}{n}$$

ציינו לגבי כל אחד מהאומדים האם הוא חסר הטיה והאם הוא עקיב ביחס לשונות θ^2. נמקו.

אם האומד מוטה, מצאו אומד חסר הטיה שהינו פונקציה של האומד המקורי . האם האומד המתוקן הינו אומד עקיב ביחס לשונות θ^2? נמקו.

שאלה 3:

יהי $X_1,\dots,X_n$ מדגם מקרי מהתפלגות לוג-נורמלית, בעלת פונקציית צפיפות הבאה:

$$f(x;\mu,\sigma^2)=\frac{1}{x\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\},\quad x\ge 0,\quad -\infty<\mu<\infty,\quad \sigma^2>0$$

הערה: אם ההתפלגות של X היא לוג-נורמלית עם פרמטרים μ,σ^2, אזי $\ln X\sim N\left(\mu,\sigma^2\right)$.

בשאלה זו נעבוד עם הקובץ Optime.csv איתו עבדתם בהרצאה. צרפו קוד עבור כל סעיף שדורש עבודה עם R.

א. אמדו את μ ואת σ^2 בשיטת הנראות המירבית.

ב. האם אומד נראות מירבית ל-μ הוא חסר הטיה? מהו ה-MSE שלו?

ג. האם אומדי נראות מירבית ל-μ,σ^2 עקיבים? ענו מבלי להסתמך על חישובי MSE.

ד. בהנחה שהנתונים על אורכי זמן הניתוח בקובץ Optime.csv מגיעים מהתפלגות לוג-נורמלית, מהם אומדני נראות מרבית ל-μ,σ^2?

ה. נסמן בתור ζ_p את השברון ה-p של התפלגות לוג-נורמלית עם פרמטרים μ,σ^2, ונסמן בתור Y_p את השברון ה-p של התפלגות נורמלית עם פרמטרים μ,σ^2. כתבו את ζ_p כפונקצייה של Y_p. הוכיחו.

ו. בנו דיאגרמת QQPLOT לבדיקת התאמה להתפלגות לוג-נורמלית, שבו מוצגים השברונים האמפיריים כנגד השברונים התאורטיים של התפלגות לוג-נורמלית עם פרמטרים השווים לאומדנים $\hat{\mu}$, $\widehat{\sigma^2}$ שמצאתם בסעיף ד׳. השתמשו בפונקציה qnorm ובקשר שמצאתם בסעיף הקודם. השוו את הצורה שמתקבלת לקו ישר מתאים.
ניתן להיעזר בפונקציה qlnorm על מנת ליצור דיאגרמת QQPLOT ל<u>בדיקת</u> הדיאגרמה שקיבלתם.

ז. מצאו דרך מתאימה להשתמש בפונקציה qqnorm כדי לבנות דיאגרמת QQPLOT לבדיקת ההתאמה של הנתונים להתפלגות <u>לוג</u>-נורמלית. לאיזה קו ישר יש להשוות את הצורה שיוצרות הנקודות במקרה הזה? הסבירו והציגו את התרשים ביחד עם הקו המתאים.

ח. על-סמך התרשימים בסעיפים ו׳ ו-ז׳, האם סביר להניח כי הנתונים באים מהתפלגות לוג-נורמלית?

ט. יהי $X_1,\dots,X_n$ מדגם מהתפלגות $Gamma(\alpha,\beta)$. מצאו אומדים בשיטת המומנטים לפרמטרים α,β. תחת ההנחה שזמני הניתוח מהווים מדגם מקרי מהתפלגות גאמה, מהם האומדנים ל-α,β?

י. בנו היסטוגרמה של זמני הניתוח כך שבציר ה-Y תוצג השכיחות היחסית, כפי שהודגם בהרצאה:

```
hist(Optime,breaks=15,freq = FALSE)
```

הציגו על ההיסטוגרמה את הצפיפות הלוג-נורמלית עם פרמטרים השווים לאומדנים שחישבתם בסעיף ד׳, וכן את הצפיפות של משתנה גאמה עם אומדנים שחישבתם בסעיף הקודם. דונו בהתאמה למודל הלוג-נורמלי לעומת ההתאמה למודל גאמה.

בסעיפים הבאים הניחו כי הנתונים באים מהתפלגות לוג-נורמלית.

יא. חוקר מעוניין לאמוד את ההסתברות שזמן של ניתוח אחד יעלה על שעתיים. מהו אומד נראות מירבית להסתברות זו? הביעו את האומד באמצעות Φ, פונקציית ההתפלגות המצטברת של משתנה נורמלי סטנדרטי. מהו אומדן נראות מירבית להסתברות זו? פרטו את כל הדרך.

רמז: מצאו את הקשר בין פונקציית ההתפלגות המצטברת של משתנה לוג-נורמלי עם פרמטרים μ, σ^2, לבין פונקציית התפלגות מצטברת של משתנה נורמלי סטנדרטי.

יב. האם האומד שמצאתם בסעיף הקודם הוא אומד עקיב?

יג. נבחן כעת אומד אחר, המבוסס על פונקציית ההתפלגות האמפירית, להסתברות המתוארת בסעיף יא׳:

$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} I\{X_i > 2\}$.

1. מהו האומדן המתקבל?
2. האם האומד עקיב? ענו ללא חישוב MSE.
3. חשבו את ההטיה וה-MSE של האומד. האם ניתן להתבסס על התוצאה שקיבלתם בשביל לענות על סעיף 2?