תרגיל בית 2 – אמידה בשיטת המומנטים

שאלה 1:

א. הוכיחו שאם $\hat{\theta}$ הינו אומד בשיטת המומנטים ל- θ, אזי $\tau\left(\hat{\theta}\right)$ הינו אומד בשיטת המומנטים ל- $\tau(\theta)$.

ב. הוכיחו שאם T_1 ו-T_2 הם אומדים בשיטת המומנטים ל-θ, אז $T_3 = \frac{T_1+T_2}{2}$ הוא גם אומד בשיטת המומנטים ל-θ.

שאלה 2:

יהי $X_1,...,X_n$ מדגם מקרי מהתפלגות $Gamma(2,\lambda)$, כלומר מהצפיפות:

$$f(x)=\begin{cases}\lambda^2 xe^{-\lambda x} & x\ge 0\\ 0 & otherwise\end{cases}$$

כאשר $\lambda>0$.

1) מצאו אומד ל- θ כאשר $\theta=\frac{1}{\lambda}$ בשיטת המומנטים. אמד זה יסומן ע"י $\hat{\theta}$. חשבו את תוחלתו ואת שונותו. איך מתנהגת שונות האומד כאשר גודל המדגם גדל? האם התנהגות זו הינה אינטואיטיבית? הסבירו.

2) מצאו אומד ל- λ בשיטת המומנטים. אומד זה יסומן ע"י $\hat{\lambda}$.

3) נתונים שלושה אומדים ל- $\frac{1}{\lambda^2}$: $T_1=\left(\frac{\sum_{i=1}^n X_i}{n}\right)^2$ ו- $T_2=\frac{\sum_{i=1}^n X_i^2}{n}$ ו- $T_3=\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}{n}$.

א. לכל אחד משלושת האומדים T_i, מצאו מהו הקבוע c_i כך שהאומד c_iT_i הוא אומד בשיטת המומנטים ל- $\frac{1}{\lambda^2}$, עבור $i=1,2,3$.

ב. חשבו את התוחלת של שלושת האומדים c_iT_i שמצאתם. ציינו מיהם האומדים שהתוחלת שלהם שווה לפרמטר לכל n,λ.

<u>שאלה 3</u>

יהי X משתנה מקרי בעל התפלגות $Geo(p)$. יהיה $X_1,..,X_n$ מדגם מקרי מההתפלגות של X.

יהיו T_1,T_2,T_3 אומדים ל-$Var\left(X\right)$:

$$T_1=\bar{X}(\bar{X}-1),\ \ T_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\ ,\ \ T_3=\frac{1}{2n}\sum_{i=1}^{n}X_i^2-\frac{1}{2}\bar{X}.$$

א. הוכיחו כי שלושת האומדים T_1,T_2,T_3 הם אומדים בשיטת המומנטים ל-$Var(X)$.

ב. חשבו את תוחלות האומדים T_1,T_2,T_3. עבור כל אחד מהאומדים, ענו: האם התוחלת של האומד גדולה/קטנה/שווה ל-$Var(X)$ לכל n,p? הציעו תיקון לאומדים שתוחלתם לא שווה ל-$Var(X)$, כך שתוחלת כל אחד מהאומדים המתוקנים תהיה שווה ל-$Var(X)$ לכל n,p. נסמן את האומד המתוקן של האומד T_i על ידי T_i^*.

ג. הגרילו $N\geq 2000$ מדגמים בגודל $2\leq n\leq 10$ מהתפלגות גיאומטרית עם פרמטר $0.2\leq p\leq 0.8$ (ניתן לבחור את הערכים של N,n,p לפי הטווחים הנתונים). חשבו על סמך כל מדגם את האומדן ל-$Var(X)$ המתקבל לפי T_1,T_2,T_3. העריכו את התוחלת של T_1,T_2,T_3 על ידי ממוצע של N האומדנים המתאימים.

חשבו את תוחלת כל אחד מהאומדים על סמך הסעיף הקודם, עבור הערכים של n,p שבחרתם. השוו את הערכים של התוחלות להערכותיהן המתקבלות בסימולציה. בנוסף, עבור כל אחד מהאומדים הציגו היסטוגרמה של הערכים שהתקבלו. בחרו ערך כלשהו של d והעריכו (בקירוב) על סמך ההיסטוגרמה, מהי ההסתברות שהמרחק המקסימלי בין האומד לערך האמיתי של הפרמטר יהיה לכל היותר d, עבור כל אחד משלושת האומדים.

<u>שימו לב:</u> הפונקציה rgeom(n,p) ב-R מחזירה וקטור של n משתנים מקריים מהתפלגות $Geo_0(p)$, כלומר התפלגות גיאומטרית-0, זוהי ההתפלגות של מספר ה**כשלונות** עד ההצלחה הראשונה (לא כולל) כשמבצעים ניסויי ברנולי בלתי תלויים. ההתפלגות הגיאומטרית $Geo(p)$ זוהי ההתפלגות של מספר ה**נסיונות** עד ההצלחה הראשונה כולל. על מנת לקבל וקטור של n משתנים מקריים מהתפלגות $Geo(p)$, יש להריץ rgeom(n,p)+1.

יש לבצע set.seed(num) לפני הרצת הסימולציה, כאשר num הוא סכום המספרים המתקבלים משתי הספרות האחרונות בת״ז של כל אחד מהמגישים. כלומר, עבור מגיש עם ת״ז 123456789 ומגיש עם ת״ז 987654321, יש לבצע set.seed עם num=89+21=110.

ד. בסעיף זה נתייחס לנתוני השלג שראיתם בהרצאה (תזכורת: הנתונים נמצאים בקובץ ״הרצאה 1 – קוד R״ שפורסם באתר, ראיתם בהרצאה שסביר להניח שהנתונים באים מהתפלגות גאומטרית).

1. מצאו אומדנים לשונות ההצטברות המקסימלית של שלג בירושלים בהתבסס על האומדים T_1, T_2, T_3 ועל סמך האומדים המתוקנים שהצעתם T_i^*.
2. מצאו אומדנים לסטיית התקן של ההצטברות המקסימלית של שלג בירושלים בהתבסס על האומדים T_1, T_2, T_3 ועל סמך האומדים המתוקנים שהצעתם T_i^*. האם יש הבדלים גדולים בין האומדנים לפי האומדים שהוצגו לבין הגרסאות המתוקנות? הסבירו.