בס"ד

מבוא להסתברות וסטטיסטיקה 88-165

תרגילי בית - אומדנים, רווחי סמך

הערה: במטלב מייצרים מדגם של $n$ משתנים מקריים נורמליים בלתי תלויים מההתפלגות $N(\mu,\sigma)$ על ידי הפקודה $\mu+\sigma*\mathrm{randn}(n,1)$.

1. למשתנה מקרי $X$ יש פונקציית הצטברות

$$F(x)=x^a\ ,\qquad 0<x<1$$

כאשר $a>-1$ הוא קבוע, אבל לא ידוע.

(א) הוכח ש- $\mathbf{E}[X]=\frac{a}{1+a}$.

(ב) מצא את אומדן הנראות המירבית לפרמטר $a$. האם אומדן זה הוא חסר הטייה ?

(ג) לאור התוצאה של סעיף (א) מציעים $\frac{\overline{X}}{1-\overline{X}}$ כאומדן ל- $a$. האם אומדן זה הוא חסר הטייה ?

(ד) במקרים $a=2,4$, העזר במחשב לייצור מדגמים מגודל 10 מההתפלגות הרלוונטית, ובדוק (על ידי הסטוגרמה) את ההתפלגות של שני האומדנים ל-$a$. איזה מהם עדיף ?
הערה: ניתן לייצור מדגמים של ההתפלגות על ידי הפקודה rand(10,1).^ (1/a)

2. $X_1,\dots,X_n$ הוא מדגם מהתפלגות גאומטרית עם פרמטר $p$. מצא את אומדן הנראות המר-בית ל- $p$. האם הוא חסר הטייה ?

3. בוחרים, ללא חזרות, $k$ מספרים מהשלמים $1,2,\dots,n$, כאשר $n$ אינו ידוע. הוכח ששני הסטטיסטיים

$$\frac{k+1}{k}\max_i(X_i)-1\ ,\quad 2\overline{X}-1$$

הם אומדנים חסרי הטייה ל- $n$. על ידי ניסוי במחשב (או אחרת) מצא לאיזה משני אומדנים אלה יש שונות נמוכה יותר.

4. במטלב בחר 500 מדגמים מגודל 5 מתוך ההתפלגות $N(75,10)$ ולכל מדגם חשב את הערכים של $\overline{X}$ ושל $s^2$. תאר את ההתפלגויות של $\overline{X}$ ושל $s^2$ (מומלץ הסטוגרמות). חזור על השאלה למדגמים מגודל 10. מה היא ההשפעה של גודל המדגם על ההתלפגות ?

5. למדגם $X_1,X_2,X_3$ מגודל 3 מההתפלגות $N(\mu,\sigma)$ עם $\sigma$ ידוע, מציעים 3 אומדנים ל- $\mu$:

$$\begin{aligned}W_1&=\tfrac{1}{3}(X_1+X_2+X_3)\\ W_2&=\mathrm{median}(X_1,X_2,X_3)\\ W_3&=\tfrac{1}{2}\left(\max_i(X_i)+\min_i(X_i)\right)\end{aligned}$$

הוכח שכולם חסרי הטייה. על ידי ניסוי במחשב או אחרת, מצא את האומדן עם שונות נמוכה ביותר

חזור על השאלה כאשר עכשיו המדגם הוא מההתפלגות $U\left(a-\frac{1}{2},a+\frac{1}{2}\right)$ עם $a$ לא ידוע, והאומדנים הם ל- $a$.

6. העזר במטלב לייצור 500 מדגמים מגודל 10 מההתפלגות הנורמלית עם $\mu = 15$, $\sigma = 5$. מכל מדגם מצא את הממוצע $\overline{X}$ ואת סטיית תקן של המדגם $s$.

על ידי בדיקת 500 המדגמים וודא ש-

$$\mathbf{P}\left(\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < 1.645\right) \approx 0.90$$

ו-

$$\mathbf{P}\left(\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < 1.96\right) \approx 0.95$$

למה

$$\mathbf{P}\left(\frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}} < 1.645\right)$$

ו-

$$\mathbf{P}\left(\frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}} < 1.96\right)$$

הם יותר נמוכים ? הסבר איך ניתן לשנות את המספרים 1.645 ו- 1.96 כך שהסתברויות אלה יעלו שוב ל- 0.90 ו- 0.95.

7. לוקחים מדגם של 8 רכיבים המיוצרים במכונה האמורה לייצר עם אורך 1.500 ס"מ. מוצאים שהאורכים של הרכיבים הם

$$1.502, 1.501, 1.504, 1.498, 1.503, 1.499, 1.505, 1.504$$

(א) מצא אומדנים לתוחלת ולשונות של האורך של רכיב.

(ב) מצא רווח סמך סימטרי לתוחלת של האורך עם רמת מובהקות 90%

(ג) מצא חסם עליון לסטיית התקן של האורך עם רמת מובהקות 95%.

8. במדגם של גברים נשואים מוצאים שהגיל בזמן החתונה מתפלג כדלהלן:

| גיל (אמצע קטע) | מספר גברים |
|---|---|
| 17.5 | 28 |
| 22.5 | 68 |
| 27.5 | 43 |
| 32.5 | 18 |
| 37.5 | 9 |
| 42.5 | 4 |
| 47.5 | 2 |
| 52.5 | 1 |
| 57.5 | 0 |
| 62.5 | 2 |

מצא אומדנים לתוחלת ולסטיית התקן של הגיל של גברים בזמן החתונה, ומצא רווח סמך עם רמת מובהקות 90% לתוחלת.

בהצלחה!