תרגול 6 – אדם צ'פמן

התפלגות פואסון

משתנה פואסוני הוא משתנה $X \sim P(\lambda)$ המקיים $P(X=i)=\frac{e^{-\lambda}(\lambda^i)}{i!}$ עבור λ נתון.

יש להתפלגות הזאת מספר משמעויות.

משמעות אחת היא מספר מאורעות שקורים בפרק זמן נתון T, כאשר תוחלת מספר המאורעות היא λ וההתפלגות היא חסרת זיכרון. מה זאת אומרת חסרת זיכרון? זה אומר שאם ישנם שני פרקי זמן לא חופפים, אז ידיעת מספר המאורעות שקרו בפרק זמן הראשון לא משפיעה על מספר המאורעות בפרק הזמן השני.

לדוגמא:

אם מספר רעידות האדמה בשבר הסורי-אפריקאי הוא משתנה פואסוני עם מאורע אחד אחת ל10 שנים, אז מספר רעידות האדמה בעשר שנים הוא $X \sim P(1)$. הסיכוי, למשל, שתהיינה שתי רעידות אדמה בעשר שנים נתונות הוא

$$P(X=2)=e^{-1}\cdot\frac{1^2}{2!}=\frac{1}{2e}$$

אם ידוע שבין 1990 לבין 2000 היו 0 רעידות אדמה, הסיכוי שיהיו שתי רעידות אדמה בין 2020 לבין 2030 הוא עדיין $\frac{1}{2e}$ בגלל שההתפלגות היא חסרת זיכרון ואין חפיפה בין שני העשורים.

משמעות נוספת היא קירוב להתפלגות בינומית: אם משתנה בינומי $X \sim B(n,p)$ הוא עם n מאוד גדול וp מאוד קטן, אז הוא בקירוב $X \sim P(\lambda)$ כאשר $\lambda = np$. כלומר, פרק הזמן T עבור ההתפלגות הפואסונית הוא פרק הזמן הנחוץ לביצוע n ניסויי ברנולי עם סיכוי הצלחה של p לכל ניסוי, וX הוא מספר ההצלחות של הניסויים.

לדוגמא:

אם בעמוד שרירותי בספר נתון יש 100 מילים, והסיכוי לשגיאת הדפסה במילה נתונה הוא $\frac{1}{200}$ אז מספר השגיאות בעמוד שרירותי מתפלג (בקירוב) פואסונית $X \sim P(\lambda)$ עם $\lambda = 100\cdot\frac{1}{200}=\frac{1}{2}$.

הערה:

אם מספר המאורעות בזמן T מתפלג פואסונית $X \sim P(\lambda)$, אז מספר המאורעות בכפולה של אותו הזמן aT יתפלג פואסונית $X \sim P(a\lambda)$ עם יחס תוחלות שהוא בדיוק יחס הזמנים (זה נובע מחוסר הזיכרון של ההתפלגות).

לדוגמא:

אם מספר האוטובוסים שעוברים בשעה בתחנה מסוימת מתפלג פואסונית עם $X \sim P(10)$, אז מספר האוטובוסים שעוברים בחצי שעה באותה התחנה יתפלג פואסונית עם $X \sim P(5)$.

שאלה:

מספר המכוניות שעוצרות בתחנת דלק "סולר לעם" מתפלג פואסונית עם 6 מכוניות בממוצע ל15 דקות.

א. מה הסיכוי שבין 17:00 ל17:15 יעצרו לפחות 4 מכוניות?
ב. ידוע כי בין 17:00 ל17:25 הגיעו 10 מכוניות. מה הסיכוי שהגיעו 6 בין 17:00 ל17:15 ועוד 4 בין 17:15 לבין 17:25.

תשובה:

מספר המכוניות שעוצרות ב15 דקות הוא משתנה $X \sim P(6)$. לכן הסיכוי שיעצרו לפחות 4 מכוניות הוא

$$P(X \geq 4) = 1 - P(X=0) - P(X=1) - P(X=2) - P(X=3) = 1 - e^{-6}\left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!}\right)$$
$$\approx 0.848$$

בסעיף ב', נסמן בA את המאורע שבו עברו 6 מכוניות בין 17:00 לבין 17:15 ועוד 4 בין 17:15 לבין 17:25, ובB את המאורע שבו עברו 10 מכוניות בין 17:00 לבין 17:25. מבקשים מאתנו לחשב את $P(A|B) = \frac{P(A \cap B)}{P(B)}$. מכיוון שA הוא תת-מאורע של B (כלומר, אם A קורה אז בטוח שB קורה), מקבלים $A \cap B = A$ ולכן $P(A|B) = \frac{P(A)}{P(B)}$.

כעת, מספר המכוניות שעוברות ב25 דקות מתפלג פואסונית עם תוחלת $\frac{25}{15} \cdot 6 = 10$, כלומר מדובר במשתנה $Y \sim P(10)$, ולכן $P(B) = P(Y=10) = e^{-10} \cdot \frac{10^{10}}{10!} \approx 0.125$.

מספר המכוניות שעוברות ב15 דקות הוא כאמור $X \sim P(6)$ ולכן מספר המכוניות שעוברות ב10 דקות הוא $Z \sim P(4)$. נשים לב שאין חפיפה בין חלון הזמן 17:00-17:15 לבין 17:15-17:25 (קצוות הזמנים זהים, אך הקצה הוא זניח). לכן המאורע $X = 6$ בלתי תלוי במאורע $Z = 4$. מכיוון שמאורע A הוא חיתוך של שני המאורעות האלה

$$P(A) = P(X=6) \cdot P(Z=4) = e^{-6} \cdot \frac{6^6}{6!} \cdot e^{-4} \cdot \frac{4^4}{4!} \approx 0.031$$

ולכן התשובה היא

$$P(A|B) = \frac{P(A)}{P(B)} \approx 0.251$$

שאלה:

מערכת תקשורת מורכבת ממספר רכיבים. הסיכוי של כל רכיב לעבוד הוא p. כדי שהמערכת תעבוד, לפחות חצי מהרכיבים צריכים לעבוד. לאילו ערכי p מערכת עם חמישה רכיבים תעבוד טוב יותר ממערכת עם שלושה רכיבים?

תשובה:

הסיכוי שמערכת עם 3 רכיבים תעבוד הוא

$$p^3 + 3p^2(1-p)$$

הסיכוי שמערכת עם 5 רכיבים תעבוד הוא

$$p^5 + 5p^4(1-p) + 10p^3(1-p)^2$$

נביט באי-השוויון

$$p^5 + 5p^4(1-p) + 10p^3(1-p)^2 > p^3 + 3p^2(1-p)$$

נחלק בp^2 ונקבל

$$p^3 + 5p^2(1-p) + 10p(1-p)^2 > p + 3(1-p)$$

$$p^3 - p + 5p^2(1-p) + 10p(1-p)^2 - 3(1-p) > 0$$

$$p(p-1)(p+1) + 5p^2(1-p) + 10p(1-p)^2 - 3(1-p) > 0$$

המספר $p - 1$ חיובי, ולכן ניתן לחלק בו בלי לשנות את אי-השוויון

$$-p(p+1) + 5p^2 + 10p(1-p) - 3 > 0$$

$$-6p^2 + 9p - 3 > 0$$

$$(2p-1)(1-p) > 0$$

מכיוון ש$p - 1$ חיובי, מקבלים ש

$$p > \frac{1}{2}$$

כלומר, לרכיב צריך סיכוי של לפחות חצי לעבוד כדי שמערכת עם חמישה רכיבים תהיה עדיפה על שלושה.

שאלה:

שתי קבוצות כדורגל משחקות סדרת משחקים עד שאחת מהן מנצחת בשני משחקים. קבוצה א מנצחת כל משחק בסבירות p. מצאו את תוחלת ושונות מספר המשחקים. הראו שתוחלת מספר המשחקים מקסימלית אם $p=\frac{1}{2}$.

תשובה:

מספר המשחקים X האפשרי הוא 2 או 3. הסיכוי שהוא 2 הוא $p^2+(1-p)^2$, והסיכוי שהוא 3 הוא המשלים, דהיינו $1-(p^2+(1-p)^2)=2p-2p^2$. אי-לכך, תוחלת מספר המשחקים היא

$$E(X)=2\cdot(2p^2-2p+1)+3\cdot(2p-2p^2)=-2p^2+2p+2$$

את השונות נחשב גם כן

$$E(X^2)=4(2p^2-2p+1)+9(2p-2p^2)=-10p^2+10p+4$$

$$V(X)=-10p^2+10p+4-(-2p^2+2p+2)^2=-4p^4+8p^3-6p^2+2p$$

נשים לב שהתוחלת כפונקציה של p היא פרבולה בוכה. נקודת הקיצון שלה (שהיא נקודת המקסימום) מתקבלת כאשר

$$-4p+2=0$$

$$p=\frac{1}{2}$$