

תרגול כיתה 10 – מבוא להסתברות וסטטיסטיקה
התפלגויות רציפות מיוחדות, מ"מ דו-מימדיים רציפים, קונבולוציה

נוסחאות:

(1). התפלגות אחידה רציפה $X \sim U(a, b)$, $a, b \in \mathbb{R}$.

פונקציית הצפיפות:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{else} \end{cases}$$

התוחלת והשונות: $E(X) = \frac{a+b}{2}$ $V(X) = \frac{(b-a)^2}{12}$

(2). התפלגות מעריכית $X \sim \text{Exp}(\lambda)$, $\lambda > 0$.

פונקציית הצפיפות:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

התוחלת והשונות: $E(X) = 1/\lambda$ $V(X) = 1/\lambda^2$

הערה: בהתפלגות זו (בדומה להתפלגות הגיאומטרית) קיימת תכונת "חוסר הזכרון":
 $P(X > s+t | X > s) = P(X > t)$, לכל $s, t \geq 0$.

שאלה 1

דני מחכה משעה 16:00 לפגישה עם חברו. ידוע כי החבר יגיע ברגע כלשהו בין 16:00 ל- 16:40 באופן שווה.

א. מהי ההסתברות שדני יצטרך להמתין למעלה מ- 20 דקות?

ב. אם עד 16:20 לא הגיע החבר, מה הסיכוי שיצטרך עדיין להמתין עוד 10 דקות לפחות?

פתרון:

ההתפלגות היא אחידה. נסמן ב- X את הזמן בו יגיע האוטובוס (בדקות) מרגע שהאדם הגיע לתחנה, אזי פונקציית

הצפיפות האחידה-

$$f_X(x) = \begin{cases} \frac{1}{40} & , 0 \leq x \leq 40 \\ 0 & , \text{else} \end{cases}$$

א). הסיכוי שדני יחכה 20 דקות לפחות הוא $\int_{20}^{40} \frac{1}{40} dx = \frac{x}{40} \Big|_{20}^{40} = \frac{1}{2}$

(ב). בהנחה כי דני כבר חיכה 20 דקות, הסיכוי כי יחכה עוד 10 דקות לפחות הוא

$$P(X \geq 30 | X \geq 20) = \frac{P(X \geq 30 \cap X \geq 20)}{P(X \geq 20)} = \frac{P(X \geq 30)}{P(X \geq 20)}$$

נציב את פונקציית הצפיפות:

$$\frac{\int_{30}^{40} \frac{dx}{40}}{\int_{20}^{40} \frac{dx}{40}} = \frac{\left. \frac{x}{40} \right|_{30}^{40}}{\left. \frac{x}{40} \right|_{20}^{40}} = \frac{\frac{1}{4}}{\frac{1}{2}} = \boxed{\frac{1}{2}}$$

שאלה 2

אורך החיים של נורה דלוקה הוא משתנה מקרי המתפלג מעריכית עם ממוצע 100 שעות.

- מה ההסתברות שהנורה תדלק יותר מ-80 שעות אך פחות מ-110 שעות?
- חשב את ההסתברות שאורך חיי נורה יעלה על 110 שעות בהינתן שהיא דולקת כבר 80 שעות.
- מהי שונות זמן חיי נורה דלוקה?

פתרון:

(א). נסמן ב- X את אורך חיי הנורה, אזי $\lambda = 1/100 \Rightarrow E(X) = 100$, לכן הצפיפות של X היא

$$f_X(x) = \frac{1}{100} e^{-\frac{x}{100}} \text{ אזי}$$

ההסתברות שהנורה תדלק יותר מ-80 שעות אך פחות מ-110 שעות

$$\int_{80}^{110} \frac{1}{100} e^{-\frac{t}{100}} dt = -e^{-\frac{t}{100}} \Big|_{80}^{110} = e^{-\frac{80}{100}} - e^{-\frac{110}{100}} \approx 0.1165$$

(ב). נעזר בתכונת חוסר הזכרון של ההתפלגות המעריכית

$$P(x > 110 | x > 80) = P(X > 30)$$

$$P(X > 30) = \int_{30}^{\infty} \frac{1}{100} e^{-\frac{t}{100}} dt = -e^{-\frac{t}{100}} \Big|_{30}^{\infty} = e^{-\frac{30}{100}} \approx 0.741$$

(ג). השונות (בשעות):

$$V(X) = 1/\lambda^2 = \frac{1}{(1/100)^2} = 10,000$$

מ"מ דו-מימדיים רציפיםשאלה 3נתונה פונקציית הצפיפות המשותפת של X ו- Y :

$$f(x, y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, \quad 0 < y < \infty \\ 0 & \text{else} \end{cases}$$

חשב את ההסתברויות הבאות:

א. $P(X > 1, Y < 1)$

ב. $P(X < Y)$

ג. $P(X < a)$ ($a > 0$ – קבוע)

פתרון:

$$\begin{aligned} P(X > 1, Y < 1) &= \int_{y=0}^{y=1} \int_{x=1}^{x=\infty} 2e^{-x}e^{-2y} dx dy = \int_0^1 2e^{-2y} \left(-e^{-x} \Big|_1^\infty \right) dy \quad \text{א.} \\ &= e^{-1} \int_0^1 2e^{-2y} dy = e^{-1} (1 - e^{-2}) \end{aligned}$$

$$\begin{aligned} P(X < Y) &= \iint_{x < y} 2e^{-x}e^{-2y} dx dy = \int_{y=0}^{\infty} \int_{x=0}^y 2e^{-x}e^{-2y} dx dy \quad \text{ב.} \\ &= \int_0^{\infty} 2e^{-2y} (1 - e^{-y}) dy = \int_0^{\infty} 2e^{-2y} dy - \int_0^{\infty} 2e^{-3y} dy = 1 - \frac{2}{3} = \frac{1}{3} \end{aligned}$$

$$\begin{aligned} P(X < a) &= \int_0^a \int_0^{\infty} 2e^{-x}e^{-2y} dy dx = \int_0^a \left[-e^{-x-2y} \right]_{y=0}^{y=\infty} dx \quad \text{ג.} \\ &= \int_0^a e^{-x} dx = \left[-e^{-x} \right]_0^a = 1 - e^{-a} \end{aligned}$$

שאלה 4פונקציית הצפיפות המשותפת של X, Y נתונה ע"י:

$$f(x, y) = \begin{cases} c(1-x-y) & 0 \leq x, 0 \leq y, x+y \leq 1 \\ 0 & \text{else} \end{cases}$$

מצא את:

א. הקבוע c .

ב. $E(X), E(Y), V(X), V(Y)$

ג. $\rho(X, Y)$

פתרון:

$$1 = \iint f(x, y) dy dx \quad \text{א. צריך להתקיים}$$

$$1 = \int_0^1 \int_0^{1-x} c(1-x-y) dy dx \quad \text{כלומר:}$$

$$1 = \int_0^1 c \left[y - xy - \frac{y^2}{2} \right]_0^{1-x} dx = \int_0^1 \frac{c}{2} (1-x)^2 dx = \frac{c}{6} \Rightarrow \boxed{c=6}$$

ב. בכדי למצוא את התוחלת של X נמצא תחילה את הצפיפות השולית שלו:

$$f_X(x) = \int_0^{1-x} 6(1-x-y) dy = 3 \left[2y - 2xy - y^2 \right]_0^{1-x} = 3(x-1)^2$$

$$\text{מטעמי סימטריה } f_Y(y) = 3(y-1)^2$$

$$E(Y) = E(X) = \int_0^1 x \cdot 3(x-1)^2 dx \quad \text{לכן התוחלת:}$$

$$= \int_0^1 (3x^3 - 6x^2 + 3x) dx = \left[\frac{3x^4}{4} - \frac{6x^3}{3} + \frac{3x^2}{2} \right]_0^1 = \boxed{\frac{1}{4}}$$

$$V(Y) = V(X) = E(X^2) - [E(X)]^2 \quad \text{והשונות:}$$

$$= \int_0^1 x^2 \cdot 3(x-1)^2 dx - \left(\frac{1}{4}\right)^2 = \left[\frac{3x^5}{5} - \frac{6x^4}{4} + \frac{3x^3}{3} \right]_0^1 - \left(\frac{1}{4}\right)^2 = \boxed{\frac{3}{80}}$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X) \cdot V(Y)}} = \frac{E(X \cdot Y) - E(X) \cdot E(Y)}{\sqrt{V(X) \cdot V(Y)}} \quad \text{ג.}$$

$$E(X \cdot Y) = \int_0^1 \int_0^{1-x} x \cdot y \cdot 6(1-x-y) dy dx$$

$$= \int_0^1 \left[-xy^2(2y-3+3x) \right]_0^{1-x} dx = \int_0^1 -x(1-x)^3 dx$$

$$= \left[\frac{-x^5}{5} + \frac{3x^4}{4} - x^3 + \frac{x^2}{2} \right]_0^1 = \frac{1}{20}$$

$$\rho(X, Y) = \frac{\frac{1}{20} - \left(\frac{1}{4}\right)^2}{\sqrt{\frac{3}{80} \cdot \frac{3}{80}}} = \boxed{-\frac{1}{3}} \quad \text{נציב את המרכיבים בנוסחת מקדם המתאם:}$$

קונבולוציה

שימוש חשוב של קונבולוציה בהסתברות הוא חישוב סכום של מ"מ בלתי תלויים. כאשר Y, X מ"מ ב"ת וסכומם $Z = X + Y$ אזי –
עבור משתנים בדידים:

$$P(Z = k) = P(X + Y = k) = \sum_i P(X = i, Y = k - i) = \sum_i P(X = i) P(Y = k - i)$$

עבור משתנים רציפים:

$$P(Z \leq k) = P(X + Y \leq k) \Rightarrow f_Z(k) = f_{X+Y}(k) = \int_{-\infty}^{\infty} f_X(k - y) f_Y(y) dy$$

שאלה 5

יהיו X, Y מ"מ ב"ת כך ש- $Y \sim \text{Exp}(\lambda), X \sim \text{Exp}(\lambda)$. מצא בעזרת קונבולוציה את התפלגות $Z = X + Y$.

פתרון:

פונקציות הצפיפות של X ו- Y הן-

$$f_X(x) = f_Y(y) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , otherwise \end{cases}$$

נעזר בנוסחת הקונבולוציה:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy = \int_0^z \lambda e^{-(z-y)} \lambda e^{-\lambda y} dy = \int_0^z \lambda^2 e^{-\lambda z} dy = \lambda^2 z e^{-\lambda z}$$

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & z \geq 0 \\ 0 & else \end{cases} \quad \text{קיבלנו:}$$

הערה: צפיפות זו ניתנת לזיהוי כהתפלגות גמא עם הפרמטרים $Z \sim \text{Gamma}(\alpha = 2, \lambda)$.

שאלה 6

יהיו X, Y מ"מ ב"ת כך ש- $Y \sim \text{Bin}(m, p), X \sim \text{Bin}(n, p)$. מצא ע"י קונבולוציה את התפלגות $Z = X + Y$.

פתרון:

$$\begin{aligned} P(X + Y = k) &= \sum_{i=0}^n P(X = i, Y = k - i) = \sum_{i=0}^n P(X = i) P(Y = k - i) \\ &= \sum_{i=0}^n \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\ &= p^k (1-p)^{n+m-k} \sum_{i=0}^n \binom{n}{i} \binom{m}{k-i} \quad \left\{ (*): \sum_{i=0}^n \binom{n}{i} \binom{m}{k-i} = \binom{n+m}{k} \right\} \\ &= p^k (1-p)^{n+m-k} \binom{n+m}{k} \\ &= \text{Bin}(n+m, p) \end{aligned}$$

כאשר השתמשנו בזהות הבינומית (*) לעיל וכן בכך ש- $\binom{b}{a} = 0$ כאשר $a > b$.