תיאור חלק ג:

בחלק זה מוגדר ונחקר תהליך הפואסון ותהליכים דומים הנגזרים מתהליך זה. עבור תהליך פואסון מוצגות 4 הגדרות שקולות ומתבצעים חישובים נלווים לתהליך וגם פיצול ומיזוג של תהליכים. בנוסף מוצגים וריאנטים של התהליך: תהליך פואסון מורכב ותהליך אשר אינו הומוגני בזמן. החלק נפתח ע"י חזרה על התפלגות אקספוננציאלית וארלנג והצגת המושג של קצב Hazard.

פרק ג-1: תכונות של ההתפלגות האקספוננציאלית והתפלגות ארלנג.

להלן התכונות הבסיסיות עבור משתנה מקרי: $X \sim \exp(\lambda)$:

פונקצית צפיפות, התפלגות ושרידות.

צפיפות: $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$ עבור פרמטר $\lambda > 0$.

התפלגות: $P(X \le x) = F_X(x) = \int_{-\infty}^{x} \lambda e^{-\lambda x} I_{[0,\infty)}(x)dx = I_{[0,\infty)}(x)(1-e^{-\lambda x})$

שרידות: $P(X > x) = 1 - F_X(x) = \overline{F}_X(x) = e^{-\lambda x} I_{[0,\infty)}(x) + I_{(-\infty,0)}(x)$

תכונת חוסר זיכרון. עבור *s, t* חיוביים.

$$P(X > s+t \mid X > t) = \frac{P(X > s+t, X > t)}{P(X > t)} = \frac{P(X > s+t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s)$$

פירוש: נניח שמשתנה מקרי X מודד זמן המתנה בשניות (או אורך חיים – זמן המתנה למוות), אז בהינתן שזמן ההמתנה גדול מ t שניות (ידוע כי המתנו לפחות t) אז ההסברות שנמתין עוד לפחות s שניות שווה להסתברות שנמתין לפחות s שניות מתחילת ההמתנה. ז"א ההמתנה לא מתקצרת/מתארכת עקב העובדה שידוע שכבר המתנו.

זהו גם המשנה המקרי הרציף היחיד בעל תכונת חוסר הזיכרון:
יהי Y מ"מ רציף כך ש $P(Y > s+t \mid Y > t) = P(Y > s)$ אז:

או: $\frac{P(Y > s+t)}{P(Y > t)} = P(Y > s)$:

$\overline{F}_Y(s+t) = \overline{F}_Y(s)\overline{F}_Y(t)$,

ניתן להוכיח כי הפתרון היחיד של המשוואה הפונקציונאלית (משוואה אשר הנעלמים בה הם פונקציות) הנ"ל הוא מהסוג: $\overline{F}_Y(x) = e^{Cx}$.

ההוכחה מתבססת על תוצאה יחסית כבדה האומרת כי הפתרון היחיד למשוואה מהסוג $g(s+t) = g(s) + g(t)$ במרחב הפו' הרציפות הוא $g(y) = Cy$ (כאשר C, הוא קבוע כלשהו).

ניקח ln על המשוואה הפונקציונאלית של $\overline{F}_Y(.)$ ונקבל

$\ln \overline{F}_Y(s+t) = \ln \overline{F}_Y(s) + \ln \overline{F}_Y(t)$.

ואז לפי התוצאה "הכבדה" נקבל כי $\ln \overline{F}_Y(y) = Cy$

ולכן $\overline{F}_Y(x) = e^{Cx}$. בשביל ש $\overline{F}_Y(.)$ תהיה פו' שרידות אז דרוש ש C יהיה שלילי ולכן ניקח $C = -\lambda$.

הערה: עבור משתנים מקריים בדידים המשפחה הגיאומטרית הינה המשפחה היחידה בעלת תכונת חוסר זיכרון.

הערה: תכונת חוסר הזיכרון נכונה גם עבור הזזות של משתנים מקריים אקספוננציאלים /גיאומטריים.

הערה: ניתן לייחס את תכונת חוסר הזיכרון לעובדה שאם מסתכלים על העקומה היורדת של הצפיפות האקספוננציאלית או של מסת ההסתברות הגיאומטרית, אז העקומה הינה בעלת אותה צורה ככל שמסתכלים הלאה אל הזנב.

קשר למשתנה מקרי גיאומטרי.

כידוע, גם משתנה מקרי גיאומטרי הוא חסר זיכרון. קשר זה בין האקספוננציאלי לגיאומטרי אינו סתמי:
יהי $X \sim \exp(\lambda)$, נגדיר $G = \lceil X \rceil$ (הערך השלם העליון). אז עבור k=1,2,...

$$P(G=k) = P(\lceil X \rceil = k) = P(X \in (k-1,k]) = \int_{k-1}^{k} \lambda e^{-\lambda x} dx = \lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_{x=k-1}^{x=k} = -(e^{-\lambda k} - e^{-\lambda(k-1)})$$

$$= e^{-\lambda(k-1)}(1-e^{-\lambda}) = (e^{-\lambda})^{k-1}(1-e^{-\lambda})$$

נגדיר $q = 1-p = e^{-\lambda}$
אזי $P(G=k) = (1-p)^{k-1} p$ וזו התפלגות גיאומטרית (סופרת ניסיונות)
עם פרמטר "סיכוי להצלחה" $1-e^{-\lambda}$.

פונקציה יוצרת מומנטים.

$$M_X(t) = Ee^{Xt} = \int_0^\infty e^{xt} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{x(t-\lambda)} dx = \lambda \left. \frac{e^{x(t-\lambda)}}{t-\lambda} \right|_{x=0}^{x=\infty} = \frac{\lambda}{t-\lambda} \left(\lim_{x\to\infty} e^{x(t-\lambda)} - 1 \right) = \frac{\lambda}{\lambda - t}$$

עבור $t-\lambda < 0$ או $t < \lambda$.

תוחלת, מומנטים, שונות.

$X \sim \exp(\lambda)$. נחשב את EX במספר דרכים.

דרך א:

$$EX = \int_0^\infty x\lambda e^{-\lambda x} dx = x(-e^{-\lambda x})\Big|_{x=0}^{x=\infty} - \int_0^\infty 1(-e^{-\lambda x}) dx = (0 - \lim_{x\to\infty} \frac{x}{e^{\lambda x}}) + \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^\infty = 0 + 0 - \frac{1}{-\lambda} = \frac{1}{\lambda}$$

דרך ב:

$$EX = \int_0^\infty \overline{F}_X(x) dx = \int_0^\infty e^{-\lambda x} dx = 0 - \frac{1}{-\lambda} = \frac{1}{\lambda}$$

דרך ג:

$$EX = \left. \frac{dM_X(t)}{dt} \right|_{t=0} = (\lambda(\lambda-t)^{-1})'\Big|_{t=0} = \lambda(-1)(-1)(\lambda-t)^{-2}\Big|_{t=0} = \frac{\lambda}{(\lambda-0)^2} = \frac{1}{\lambda}$$

נחשב את השונות:

ראשית נחשב את המומנט השני ע"י גזירה פעמיים של פו' יוצרת מומנטים והצבה באפס.

$$EX^2 = \left. \frac{d^2 M_X(t)}{dt^2} \right|_{t=0} = (\lambda(\lambda-t)^{-1})''\Big|_{t=0} = (\lambda(-1)(-1)(\lambda-t)^{-2})'\Big|_{t=0} = \lambda(-2)(-1)(\lambda-t)^{-3}\Big|_{t=0} = \frac{2}{\lambda^2}$$

כעת: $Var(X) = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - (\frac{1}{\lambda})^2 = \frac{1}{\lambda^2}$.

הערה: אם כך סטיית התקן של מ"מ אקספוננציאלי שווה לתוחלתו ולכן מקדם ההשתנות (CV) של משתנה מקרי זה הוא 1. ($CV(X)=\frac{\sqrt{Var(X)}}{EX}$).

התפלגות מינימום.

בתחילת הקורס נזכרנו כי פו' השרידות של המינימום של n משתנים מקריים בלתי תלויים הינה מכפלת פו' השרידות (נרענן זאת כאן):

$$\overline{F}_{Min(X_1,...,X_n)}(x)=P(Min(X_1,...,X_n)>x)=P(X_1>x)...P(X_n>x)=\prod_{i=1}^{n}\overline{F}_{X_i}(x)$$

(במידה והמשתנים המקריים הינם שווי התפלגות אז המכפלה הופחת לחזקה בגובה n).

כאשר נתונים n משתנים מקריים אקספוננציאלים, כאשר $X_i \sim \exp(\lambda_i)$ אזי מקבלים:

$$\overline{F}_{Min(X_1,...,X_n)}(x)=\prod_{i=1}^{n}\overline{F}_{X_i}(x)=\prod_{i=1}^{n}e^{-\lambda_i x}=e^{-(\lambda_1+...+\lambda_n)x}.$$

ז"א המינימום הוא גם אקספוננציאלי בעל סכום העוצמות.

הערה: תוצאה זו תהייה מרכזית בהבנתנו של תהליכי קפיצה מרקובים, בפרק הבא.

תחרות בין משתנים מקריים אקספוננציאלים.

נדון כעת במצב בו ישנם שני משתנים מקריים אקספוננציאלים בלתי תלויים:

$$X \sim \exp(\lambda_X)$$

$$Y \sim \exp(\lambda_Y)$$

נחשוב על X ועל Y כמסמלים את הזמן מתחילת החיים עד אשר שתי תאומות מתחתנות. (התאומות הן X ו Y). (ייתכן והנחת האי-תלות במקרה זה היא קצת לא מציאותית).

בסעיף הקודם ראינו כי הזמן עד החתונה הראשונה ((Min(X,Y)) מתפלג $\exp(\lambda_X+\lambda_Y)$.

כעת נדון בהסתברות $P_X=P(X<Y)=P(X=Min(X,Y))=1-P_Y$ (ההסתברות שהתאומה X תתחתן ראשונה).

אינטואיטיבית רואים שאם $\lambda_X=\lambda_Y$ אזי הסתברות זאת צריכה להיות חצי. בנוסף נצפה כי P_X תהיה פו' עולה ב λ_X.

<u>דרך חישוב א':</u>

הצפיפות המשותפת של X ו Y היא: $f_{X,Y}(x,y)=\lambda_X e^{-\lambda_X x}I_{[0,\infty)}(x)\lambda_Y e^{-\lambda_Y y}I_{[0,\infty)}(y)$.

עכשיו, נבצע אינטגרציה על כל הזוגות (x,y) ב $\mathbb{R}^2$ אשר עבורם x<y.

$$P_X=\int_{x=0}^{\infty}\int_{y=x}^{\infty}f_{X,Y}(x,y)dydx=\int_{x=0}^{\infty}\int_{y=x}^{\infty}\lambda_X\lambda_Y e^{-\lambda_X x}e^{-\lambda_Y y}dydx=\int_{x=0}^{\infty}\lambda_X e^{-\lambda_X x}\int_{y=x}^{\infty}\lambda_Y e^{-\lambda_Y y}dydx$$

$$=\int_{x=0}^{\infty}\lambda_X e^{-\lambda_X x}\overline{F}_Y(x)dx=\lambda_X\int_{x=0}^{\infty}e^{-\lambda_X x}e^{-\lambda_Y x}dx=\frac{\lambda_X}{\lambda_X+\lambda_Y}\int_{x=0}^{\infty}(\lambda_X+\lambda_Y)e^{-(\lambda_X+\lambda_Y)x}dx=\frac{\lambda_X}{\lambda_X+\lambda_Y}1=\frac{\lambda_X}{\lambda_X+\lambda_Y}$$

<u>דרך חישוב ב':</u>

על פי נוסחת ההסתברות השלמה:

$$P_X = P(X<Y) = \int_{x=0}^{\infty} P(X<Y \mid X=x) f_X(x)dx$$

ולכן זה שווה ל –

$$\int_{x=0}^{\infty} \overline{F}_Y(x) f_X(x)dx = \int_{x=0}^{\infty} e^{-\lambda_Y x}\lambda_X e^{-\lambda_X x}dx = \frac{\lambda_X}{\lambda_X+\lambda_Y}\int_{x=0}^{\infty}(\lambda_X+\lambda_Y)e^{-(\lambda_X+\lambda_Y)x}dx = \frac{\lambda_X}{\lambda_X+\lambda_Y}$$

כעת נרחיב תוצאה זו לn משתנים מקריים $X_i \sim \exp(\lambda_i)$ בלתי תלויים.

נראה כי $P_i = P(X_i = Min(X_1,...,X_n)) = \frac{\lambda_i}{\lambda_1+...+\lambda_n}$

נגדיר $U = Min(X_1,...,X_{i-1},X_{i+1},...,X_n)$ אזי נסמן $\lambda_U = \lambda_1+...+\lambda_{i-1}+\lambda_{i+1}+...+\lambda_n$

ואז $U \sim \exp(\lambda_U)$.

ועכשיו $P_i = P(X_i < U) = \frac{\lambda_i}{\lambda_i+\lambda_U} = \frac{\lambda_1}{\lambda_1+...+\lambda_n}$.

אי-תלות של "מי ניצח" ו"ערך המנצח"

נסתכל שוב על תחרות בין שני משתנים מקריים $X \sim \exp(\lambda_X)$ $Y \sim \exp(\lambda_Y)$.

ראינו כיצד מתפלג $M = Min(X,Y)$ (הערך המנצח). וראינו מהו הסיכוי ש X או לחלופין Y ינצחו.

תוצאה מעניינת נוספת היא ששתי התוצאות הללו בלתי תלויות: נגדיר

$I = \begin{cases} 1 & X<Y \\ 0 & otherwise \end{cases}$. זהו משתנה מקרי המקבל 1 במידה ו – X ניצח או 0 במידה ו Y ניצח.

נראה כי המשתנים המקריים M ו - I הינם בלתי תלויים.

$$P(I=1, Min(X_1,X_2)>t) = P(X_1>t, X_2>X_1) = \int_{x=t}^{\infty}\int_{y=t}^{\infty}\lambda_X\lambda_Y e^{-\lambda_X x}e^{-\lambda_Y y}dydx$$

$$= \int_{x=t}^{\infty}\lambda_X e^{-\lambda_X x}e^{-t\lambda_Y}dx = \frac{\lambda_X}{\lambda_X+\lambda_Y}\int_{x=t}^{\infty}(\lambda_X+\lambda_Y)e^{-\lambda_X x}e^{-t\lambda_Y}dx = P(I=1)P(Min(X_1,X_2>t)$$

$$P(I=2, Min(X_1,X_2)>t) = P(X_2>t, X_1>X_2) = ... = P(I=2)P(Min(X_1,X_2>t)$$

תוצאה זו נכונה גם עבור תחרות של n משתנים מקריים אקספוננציאלים בלתי תלויים.

סכום של משתנים מקריים בעלי קצב זהה מתפלג ארלנג.

בסעיף זה ניזכר בהתפלגות גאמא ובפרט במקרה פרטי שלה, התפלגות ארלנג. ניזכר בכך שסכום של גאמות בלתי תלויות (כאשר פרמטר הקצב זהה) מתפלגות גאמא ונישם למקרה הפרטי הארלנגי אשר יהיה שימושי ביותר בהמשך הפרק.

X מתפלג גאמא עם פרמטר צורה α ופרמטר קצב λ אם צפיפות X מכילה את הגרעין $x^{\alpha-1}e^{-\lambda x}$.

(הגרעין של פו' צפיפות הוא החלק "עם הבשר", החלק אשר משתנה על פי הערך x, בניגוד לקבוע, במידה והגרעין הוא K(x) אזי, **הגורם המנרמל** הוא $(\int K(x)dx)^{-1}$, וכך מכפלת הגרעין בגורם המנרמל נותן פו' צפיפות תקנית, ז"א האינטגרל מעל התומך הוא 1).

חלק ג: תהליכי פואסון

במקרה של התפלגות גאמא, הגורם המנרמל הוא $\frac{\lambda^\alpha}{\Gamma(\alpha)}$. כאשר $\Gamma(.)$ היא פו' גאמא:

$$\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt .$$

לפו' גאמא מס' תכונות, לא נדון בהן כאן פרט לכך ש $\Gamma(n)=(n-1)!$ (עבור $n\in\mathbb{N}$). קל להראות זאת ע"י אינטגרציה בחלקים.

אם כך:

$$f_X(x)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}I_{[0,\infty)}(x) .$$

כעת נראה כי אם $\begin{matrix} X_1\sim gamma(\alpha_1,\lambda) \\ X_2\sim gamma(\alpha_2,\lambda)\end{matrix}$ בלתי תלויים, אזי $X_1+X_2\sim gamma(\alpha_1+\alpha_2,\lambda)$.

דרך א': קונבולוציה.

לצורך חישוב זה ניזכר בהתפלגות בטא המייגעת:

$X\sim Betta(\alpha,\beta)$ אם $f_X(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}I_{[0,1]}(x)$

נבצע כעת את הקונבולוציה ו"נשלים" להתפלגות בטא במהלך החישוב (ע"י הצבה u=s/x).

$$f_{X_1+X_2}(x)=f_{X_1}(x)\otimes f_{X_2}(x)=\int_{s=0}^{\infty}f_{X_1}(s)f_{X_2}(x-s)ds=\int_{s=0}^{x}f_{X_1}(s)f_{X_2}(x-s)ds=$$

$$=\int_{s=0}^{x}\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}s^{\alpha_1-1}e^{-\lambda s}\frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}(x-s)^{\alpha_2-1}e^{-\lambda(x-s)}ds=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_{s=0}^{x}s^{\alpha_1-1}(x-s)^{\alpha_2-1}e^{-\lambda x}ds$$

$$=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-\lambda x}\int_{s=0}^{x}s^{\alpha_1-1}(x-s)^{\alpha_2-1}ds=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}e^{-\lambda x}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_{u=0}^{1}(ux)^{\alpha_1-1}(x-ux)^{\alpha_2-1}xdu$$

$$=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}x^{\alpha_1-1}x^{\alpha_2-1}xe^{-\lambda x}\int_{u=0}^{1}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}u^{\alpha_1-1}(1-u)^{\alpha_2-1}du$$

$$=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}x^{\alpha_1+\alpha_2-1}e^{-\lambda x}\int_{u=0}^{1}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}u^{\alpha_1-1}(1-u)^{\alpha_2-1}du=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}x^{\alpha_1+\alpha_2-1}e^{-\lambda x}$$

דרך ב': פו' יוצרת מומנטים.

נתון $X\sim gamma(\alpha,\lambda)$

$$M_X(t)=\int_{s=0}^{\infty}e^{st}\frac{\lambda^\alpha}{\Gamma(\alpha)}s^{\alpha-1}e^{-\lambda s}ds=\frac{\lambda^\alpha}{(\lambda-t)^\alpha}\int_{s=0}^{\infty}\frac{(\lambda-t)^\alpha}{\Gamma(\alpha)}s^{\alpha-1}e^{-(\lambda-t)s}ds=\left(\frac{\lambda}{\lambda-t}\right)^\alpha$$

עבור $\lambda-t>0$, ז"א $t<\lambda$.

כעת $M_{X_1+X_2}(t)=M_{X_1}(t)M_{X_2}(t)=\left(\frac{\lambda}{\lambda-t}\right)^{\alpha_1+\alpha_2}$.

וזו אכן הפו' יוצרת מומנטים של $gamma(\lambda,\alpha_1+\alpha_2)$.

הרחבת תוצאה זו ל n משתנים מקריים היא מיידית.

חדי העין הבחינו כי $gamma(1,\lambda)\equiv\exp(\lambda)$ (כן, יוצא ש $\Gamma(1)=1$).
ולכן אם $X_1,...,X_n\sim\exp(\lambda)$ בלתי תלויים אזי $T=X_1+...+X_n\sim gamma(n,\lambda)$.
להתפלגות gamma כאשר פרמטר הצורה הוא שלם קוראים ארלנג (Erlang) – ולכן סכום של n משתנים מקריים אקספוננציאלים .i.i.d מתפלג $erlang(n,\lambda)$ וצפיפותו:

$$f_T(x)=\frac{\lambda^n}{(n-1)!}x^{n-1}e^{-\lambda x}I_{[0,\infty)}(x).$$

אין טעם לחשב באופן ישיר את התוחלת והשונות, יותר קל לעשות זאת ע"י העובדה כי משתנה מקרי זה הוא סכום של n אקספוננציאלים ולכן:

$$ET=(\underbrace{\frac{1}{\lambda}+...+\frac{1}{\lambda}}_{n})=\frac{n}{\lambda}$$

ובאותו אופן $Var(T)=\frac{n}{\lambda^2}$.

פרק ג-2: קצב Hazard (סיכון).

אנו יודעים כי ניתן לייצג את ההתפלגות של משתנים מקריים במספר דרכים:

$$F_X(x), \overline{F}_X(x), f_X(x), P_X(x), M_X(t), G_X(t)$$

וזאת כאשר $f_X(x)$ מוגדרת רק עבור מ"מ רציפים ו $G_X(t)$ ו $P_X(x)$ הן פונקציות רק עבור מ"מ בדידים.
ז"א כל האפשרויות הנ"ל עומדות לרשותנו לייצג את חוק ההסתברות (ההתפלגות) של משתנים מקריים.

בפרק זה, נכיר דרך נוספת אשר באמצעותה ניתן לתאר את חוק ההתפלגות של משתנים מקריים אי שליליים רציפים, זוהי פונקצית קצב ה Hazard או קצב הסיכון.

נגדיר את קצב ה Hazard כך:

$$h_X(x) = \frac{f_X(x)}{\overline{F}_X(x)}.$$

לפני שנתעמק במשמעות של ערך זה, נחשב את קצב ה Hazard של משתנה מקרי אקספוננציאלי כדוגמא.

$$X \sim \exp(\lambda)$$

$$h_X(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$$

לצורך הסבר של המשמעות של קצב ה Hazard, נתייחס למשתנה מקרי X כמשתנה מקרי אי-שלילי אשר מייצג אורך חיים של רכיב כלשהו (יכול להיות גם בן-אדם ביישומים אקטואריים). נשתמש במינוחים של כשלון ושרידות, ז"א אם X>x אז שרדנו יותר מ x יחידות זמן וכו'.

כדי להבין את המשמעות של קצב הסיכון נסתכל על הגודל הבא: $P(x < X \le x+\Delta x \mid X > x)$, זוהי ההסברות ששרדנו עד זמן x ולאחר מכן בפרק זמן בגודל Δx נכשלנו.

ניתן לראות כי:

$$P(x < X \le x+\Delta x \mid X > x) = \frac{P(x < X \le x+\Delta x, X > x)}{P(X > x)} = \frac{P(x < X \le x+\Delta x)}{P(X > x)} = \frac{P(x < X \le x+\Delta x)}{\overline{F}_X(x)}$$

אם כך, עבור Δx קטן מתקיים:

$$P(x < X \le x+\Delta x \mid X > x) \approx \frac{f_X(x)\Delta x}{\overline{F}_X(x)} = h_x(x)\Delta x$$

ז"א, קצב הסיכון בזמן x מוכפל ב Δx הוא ההסתברות להיכשל בזמן זה בהינתן שעדיין לא נכשלנו עד זמן זה. (ביישומיים אקטוארים, ערך זה נקרא עוצמת התמותה).

חשוב לשים לב כי קצב הסיכון בשל עצמו אינו מהווה הסתברות (הוא יכול להיות גדול מ 1) וזאת בדיוק כפי שפונקצית הצפיפות אינה מהווה הסתברות. אבל כאשר מכפילים בערך קטן Δx אז קצב הסיכון (כמו כן גם

פונקצית צפיפות) מקרבים גודל הסתברותי.

נחזור למקרה האקספוננציאלי ובו ראינו כי קצב הסיכון הוא קבוע λ. תוצאה זו מראה שבמידה ואורך החיים הוא בעל התפלגות אקספוננציאלית אז ההסתברות להיכשל בכל רגע נתון (בהנחה שעדיין לא נכשלנו) היא קבועה. זאת אם כך הסתכלות נוספת על תכונת חוסר הזיכרון של משתנים מקריים אקספוננציאלים.

דוגמא:
יהי $X \sim Uniform(0,1)$ מהו קצב הסיכון?
נצפה כאן שקצב הסיכון יעלה ככל שמתרחקים מ -0 ומתקרבים ל -1:

$$h_X(x) = \frac{f_X(x)}{1-F_X(x)} = \frac{1}{1-x}$$

מעבר מקצב סיכון לפונקצית שרידות

ראינו כיצד ניתן למצוא את קצב הסיכון בהינתן פונקצית שרידות (זאת על פי הגדרה):

$$h_X(x) = \frac{-\overline{F}_X'(x)}{\overline{F_X}(x)}$$ (הרי הצפיפות היא הנגזרת השלילית של פונקצית השרידות).

נראה כעת כיצד פונקצית השרידות ניתנת לחישוב על פי קצב הסיכון.
כזכור $\log(g)' = \frac{1}{g} g'$
ולכן

$$-h_X(x) = \log(\overline{F}_X(x))'$$

ניקח אינטגרל על שני צדדי המשוואה ונקבל:

$$-\int_0^x h_X(s)ds = \log(\overline{F_X}(x))$$

ולכן

$$\overline{F_X}(x) = e^{-\int_0^x h_X(s)ds}$$

אם כך יש בידינו נוסחה המאפשרת לקבל את פונקצית השרידות (ובכך כמובן גם את פונקצית ההתפלגות המצטברת) מפונקצית קצב הסיכון.

כמו כן ניתן שוב לראות שעבור המקרה האקספוננציאלי (קצב סיכון קבוע λ), הערך של האינטגרל באקספוננט הוא $-\lambda x$ כצפוי.

הערה: כאשר נדון במערכות תורים (בהמשך הקורס) נתייחס אל משתנים מקריים אי-שלילים רציפים כזמנים בין הגעות של לקוחות במערכת וכזמני שרות לקוחות בדוגמאות אלו, קצב הסיכון ישמש כקצב הגעה או קצב שרות. ז"א לקצב הסיכון יכולה להיות גם משמעות של קצב שרות/הגעה מעבר לקצב התמותה/כשלון אשר הוצג בפרק זה.

חלק ג: תהליכי פואסון

דוגמא, התפלגות Weibull:

פונ' הצפיפות היא: $f_X^{(x)} = \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}}$,

ההתפלגות מוגדרת באופן טבעי ע"י קצב הסיכון $h_X(x) = \alpha\lambda x^{\alpha-1}$. כך הפרמטר α קובע האם הסיכון עולה או יורד ככל ש x גדל. לדוגמא עבור $\alpha = 1$ קצב הסיכון הוא קבוע (ואז ההתפלגות היא אקספוננציאלית),
עבור $\alpha > 1$ קצב הסיכון עולה ועבור $\alpha < 1$ קצב הסיכון יורד.
קל לקבל את הצפיפות, השרידות והתוחלת של ההתפלגות מקצב הסיכון.

התפלגות זו משמשת פעמים רבות למידול אמינות של רכיבים וזאת בגלל שלפעמים יותר נוח להתאים סטטיסטית קצב סיכון מאשר צפיפות או שרידות וקצב הסיכון של התפלגות זו הוא הגיוני למידול (עולה קבוע או יורד ובקצבים שונים).

פרק ג-3: מבוא לתהליך פואסון.

הגדרת תהליך ספירה:

לפני שנתאר ונגדיר במפורט מהו תהליך פואסון, נגדיר את המונח של תהליך ספירה. תהליך סטוכסטי $\{N_t, t \geq 0\}$ הוא **תהליך ספירה בזמן רציף** אם N_t מסמל את סך המאורעות אשר הגיעו/אירעו/נספרו/נולדו עד זמן t. ריאליזציה של תהליך ספירה נראית כמו פונקצית מדרגה לא יורדת, רציפה מימין ובעלת גבול משמאל. באמצעות תהליך ספירה ניתן לתאר:

א) מרכזיית טלפון אשר מקבלת שיחות טלפון.
ב) מונה גייגר אשר סופג חלקיקים.
ג) תביעות לחברת ביטוח.
ד) שאילתות http לאתר אינטרנט.

<u>הגדרה (מפורטת):</u>
התהליך $\{N_t, t \geq 0\}$ הוא **תהליך ספירה בזמן רציף** אם:

1) $N_t \geq 0$

2) המשתנים המקריים N_t מקבלים ערכים ב $\mathbb{N}$ (כולל 0).

3) אם s<t אזי $N_s \leq N_t$

4) עבור s<t הערך $N_t - N_s$ הוא **האינקרימנט** ומסמל את **מספר המאורעות** אשר אירעו באינטרוול (s,t].

5) $N_0 = 0$. (הערה: בספרות לפעמים סעיף זה אינו חלק מההגדרה).

אם כך, ריאליזציה של תהליך ספירה נראית כמו פונקצית מדרגה לא יורדת כאשר כל "קפיצה" בפונקצית המדרגה בזמן t_0 מספרת כי בזמן זה, הגיעו $N_{t_0} - N_{t_0^-}$ ספירות חדשות.

הערה: לפעמים נכנה תהליך מסוג זה, **תהליך לידה טהור** ונתייחס לערך התהליך כמצב (כפי שהתייחסנו למצב בשרשראות מרקוב) ונאמר כי השינויים במצב יכולים להיות "רק כלפי מעלה". המונח "לידה טהור" נובע מהעובדה כי בהמשך נפגוש תהליכים אשר נקראים תהליכי לידה מוות ובהם לא רק נולדים (עולים כלפי מעלה במצב) אלה גם מתים (יורדים כלפי מטה במצב).

הקשר בין תהליך ספירה לתהליך זמני ההגעה:

הערה: נושא זה גם תואר בפרק א-5.

בסעיף הקודם תיארנו תהליך ספירה בצורה הישירה, ערך התהליך בזמן t. המאורע $\{N_t = n\}$ מתאר כי עד זמן t נספרו n הגעות והמאורע $\{N_t \leq n\}$ מתאר כי עד זמן t נספרו לכל היותר n הגעות. ניתן לתאר תהליך ספירה גם בצורה עקיפה, זאת ע"י תיאור של זמני הגעות/זמני הספירה. ע"י ידע של כל זמני ההגעה, ניתן הרי לשחזר את מספר ההגעות בכל זמן נתון.

נסמן בתהליך $\{T_k, k = 1,2,...\}$ את **תהליך זמני ההגעה/המופע**. מרחב המצבים של תהליך זה הוא רציף בעוד שמרחב הפרמטר הוא בדיד. המשמעות של מאורע מהסוג $\{T_k = t_0\}$ הוא "זמן הספירה ה k הוא t_0".

חלק ג: תהליכי פואסון

הערה: נקבע כי $T(0)\equiv 0$.

לפעמים נוח להגדיר את תהליך זמני ההגעה כך:

$T_k=\inf\{t\in\mathbb{R}\,|\,k\le N_t\}$ - זהו הזמן המינימאלי מבין כל הזמנים אשר בו תהליך הספירה הוא לפחות k.

באופן דומה, ניתן להגדיר את תהליך הספירה כך: $N_t=\sup\{k\in\mathbb{N}\,|\,T_k\le t\}$ - זהו ערך הספירה המקסימאלי מבין כל ערכי הספירה אשר ההגעה שלו הייתה לכל היותר בזמן t.

קשר בסיסי בין תהליך ספריה לתהליך זמני ההגעות המתאים הוא השקילות בין שתי המאורעות הבאים: $\{N_t\le k\}\Leftrightarrow\{T_k\ge t\}$. (זאת כמובן עבור $\omega\in\Omega$ נתונה).

הוכחה לכוון $\Leftarrow$: אם זמן ההגעה ה n הוא לאחר או שווה לזמן t אזי בזמן t היו לכל היותר n ספירות.

הוכחה לכוון $\Rightarrow$: אם בזמן t היו לכל היותר n ספירות אזי הספירה ה n חייבת להיות לפחות בזמן t.

תיאור אינטואיטיבי של תהליך פואסון כקרוב של תהליך בינומי:

נשים לב שהגדרתנו לתהליך ספירה לא ייחסה חוק הסתברות כלשהו לתהליך (לאוסף הריאליזציות האפשריות של התהליך) אלה בסך הכול תיארה תכונות כלליות של כל ריאליזציה של התהליך. כעת נתאר את חוק הסתברות עבור המקרה הפרטי המעניין: תהליך פואסון.

בפרק הבא נגדיר במדויק את תהליך פואסון, כאן נקבל תחושה בלבד. נניח כי אנו מבצעים את התרגיל הבא, מתבוננים בעיניים של חברנו ורושמים את הזמנים של המצמוצים בעיניים. "הגעות" המצמוצים יכולים להוות תהליך ספירה (כאשר בזמן t, ערך התהליך מספר כמה מצמוצים היו בפרק הזמן $[0,t]$). נניח כי הופעת מצמוץ ברגע מסוים אינה תלויה בהופעות המצמוצים אשר אירעו ברגעים הקודמים, במילים אחרות, נניח כי ציר הזמן מחולק ל"רגעים" שהם יחסית קטנים (נאמר עשירית שנייה), והופעת מצמוץ ברגע מסוים היא בעצם כמו הצלחה בניסוי ברנולי ברגע זה (אי-הופעת מצמוץ היא כמו כשלון בניסוי הברנולי). בנוסף נניח כי כל ניסויי הברנולי הינם בלתי תלויים. כאשר מניחים כי ההסתברות ליותר ממצמוץ אחד ב"רגע" היא אפס, ומשאיפים את גודל הרגעים ל 0 בצורה מתאימה, אז מקבלים תהליך פואסון.

אופן השאיפה צריך להיות כזה: ברור כי כאשר גודל "רגע" שואף לאפס אז גם ההסתברות להצלחה צריכה לשאוף לאפס. אבל אם משמרים את המנה של הסתברות ההצלחה וגודל הרגע קבוע אז מקבלים תהליך פואסון.

נראה בהמשך שתהליך פואסון קשור ישירות להתפלגות הפואסון – ערך התהליך בזמן t מתפלג פואסונית עם פרמטר λt, כאשר λ הוא תוחלת מספר המופעים ליחידת זמן.

כאן ניזכר ביחס בין התפלגות פואסון וההתפלגות הבינומית:

יהי X_n משתנה מקרי המתפלג בינומית עם פרמטרים n ו p_n. אם נשאיף את n לאינסוף ואת p_n ל 0 כך שנשמר את היחס $np_n=\lambda$ או $p_n=\frac{\lambda}{n}$ אזי נקבל שהתפלגות X_n שואפת להתפלגות פואסונית עם פרמטר λ. קל להוכיח תוצאה זו:

$$P(X_n=k)=\binom{n}{k}{p_n}^k(1-p_n)^{n-k}=\frac{n!}{(n-k)!k!}(\frac{\lambda}{n})^k(1-\frac{\lambda}{n})^{n-k}=\frac{n(n-1)(n-2)...(n-k+1)}{\underbrace{n....n}_{k}}\frac{1}{(1-\frac{\lambda}{n})^k}(1-\frac{\lambda}{n})^n\frac{\lambda^k}{k!}$$

$$\lim_{n\to\infty}P(X_n=k)=1\cdot 1\cdot e^{-\lambda}\frac{\lambda^k}{k!}$$

מה אנו לומדים מקשר זה בין ההתפלגויות? נניח כי ציר הזמן שלנו הוא סופי (נניח 1). ויש לנו תהליך ספירה או תהליך הגעות על ציר זמן זה. נניח בנוסף כי ההסתברות הרגעית להגעה היא קבועה (קצב הסיכון) והיא λ. כעת נניח כי אנו מקרבים את תהליך ההגעות ע"י חלוקה של ציר הזמן הרציף ל n יחידות זמן, ובכל יחידת זמן יתכן אחד משני מקרים. א) אין הגעות. ב) יש הגעה אחת. ההסתברות להגעה היא $p_n=\frac{\lambda}{n}$. אז במקרה זה מס' ההגעות מתפלגת בינומית כמו המשתנה המקרי X_n. אך כאשר נשאיף את n לאינסוף נקבל כפי שצוין לעיל כי ההתפלגות הגבולית הינה פואסונית וזה בצירוף עם התוצאה אשר תיארנו לעיל כי ההתפלגות של תהליך פואסון בזמן t היא פואסונית עם פרמטר λt.

הגדרת אינקרימנטים בלתי תלויים:

הערה: המקרה הבדיד של הגדרה זו הוצג בפרק א-5.

<u>הגדרה:</u>
תהליך ספירה הוא בעל **אינקרימנטים בלתי תלויים** אם מספר המאורעות אשר מופעים בקטעי זמן זרים הם בלתי תלויים.

בפרק הבא נראה כי תהליך פואסון הוא בעל אינקרימנטים בלתי תלויים.

הגדרת אינקרימנטים סטציונרים:

הערה: הגדרה זו גם הוצגה בפרק א-5.

<u>הגדרה:</u>
תהליך ספירה הוא בעל **אינקרימנטים סטציונרים** אם עבור כל אינטרוול $(t_1,t_2]$ מתקיים כי $N_{t_2}-N_{t_1}$ מתפלג כמו $N_{t_2+s}-N_{t_1+s}$ עבור כל s חיובי.

משמעות ההגדרה היא שהתפלגות מספר אינקרימנטים/הגעות/מאורעות באינטרוול תלויה רק באורך האינטרוול.

הערה: בחלק הקודם של הקורס (שרשראות מרקוב) למדנו מהו תהליך סטציונרי. על נא נבלבל בין המושגים. תהליך בעל אינקרימנטים סטציונרים אינו בהכרח תהליך סטציונרי. אבל תהליך סטציונרי הוא בעל אינקרימנטים בלתי תלויים.

חלק ג: תהליכי פואסון

הגדרה טכנית o(h):

הגדרה:

נאמר שפונקציה $f:\mathbb{R}\to\mathbb{R}$ **היא o(h)** אם מתקיים: $\lim_{h\to 0}\frac{f(h)}{h}=0$.

הערה: באופן מדויק, o(h) היא קבוצת כל הפונקציות אשר מקיימות את התנאי שצוין לעיל. למרות זאת, נוח פשוט לומר כי פונקציה "היא o(h) " ובכך הכוונה היא ששאיפת הפונקציה ל 0 היא בקצב גבוהה מליניארי וגבולה ב 0 – הוא 0.

דוגמא:

הפונקציה $f(x)=x^2$ היא o(h) כי $\lim_{h\to 0}\frac{f(h)}{h}=\lim_{h\to 0}\frac{h^2}{h}=\lim_{h\to 0}h=0$.

הפונקציה $f(x)=x$ אינה o(h) כי $\lim_{h\to 0}\frac{f(h)}{h}=\lim_{h\to 0}\frac{h}{h}=\lim_{h\to 0}1=1\neq 0$

באופן כללי $f(x)=x^{\alpha}$ היא o(h) אם $\alpha>1$.

טענה: כל קומבינציה ליניארית סופית של פונקציות שהן o(h), גם היא o(h).

הוכחה:

צריך להוכיח כי אם $f_1,\dots,f_n$ הן o(h) אזי גם $g(x)=c_1f_1(x)+\dots+c_nf_n(x)$ היא גם o(h).

מתקיים כי:

$$\lim_{h\to 0}\frac{c_if_i(h)}{h}=c_i\lim_{h\to 0}\frac{f_i(h)}{h}=c_i\cdot 0=0$$

ולכן c_if_i היא o(h).

בנוסף מתקיים כי $\lim_{h\to 0}\frac{f(h)+g(h)}{h}=\lim_{h\to 0}\frac{f(h)}{h}+\lim_{h\to 0}\frac{g(h)}{h}=0+0=0$

ולכן כל סכום של שתי פו' שהן o(h) הוא o(h) ולכן גם כל סכום סופי הוא o(h).