תיאור חלק א:

חלק זה מהווה מבוא לקורס. מטרתו היא להציג ולהגדיר תהליכים סטוכסטיים כמודלים מתמטים יישומיים המאפשרים לנתח בעיות יישומיות ותיאורטיות שונות. בנוסף מטרתו היא לחזור על מושגים מהסתברות אשר נלמדו בקורס מבוא להסתברות א' ותורת ההתפלגויות בכדי לחזק את ההבנה בנושאים אלו ולהכיר מספר דוגמאות מורכבות יותר אשר ניתוחם דומה לניתוח של בעיות אשר יופיעו בהמשך הקורס. החלק מסתיים עם הנושא של תהליכי ברנולי. לימוד של תהליכים אלו יעזור לפתח הבנה של מושגי יסוד אשר יופיעו בפרקים בהמשך.

חלק א: מבוא

פרק א-1: הגדרת תהליך סטוכסטי, זמן בדיד/רציף, מרחב מצבים, דוגמאות, שימושים וסקירת הקורס.

הגדרת תהליך סטוכטסטי:

בקורס מבוא להסתברות ותורת ההתפלגויות הגדרנו מרחבי הסתברות ע"י השלשה (Ω, Σ, P).
כאשר Ω, מרחב המדגם, הוא אוסף התוצאות האפשריות.
Σ, אוסף המאורעות, הוא אוסף של תתי קבוצות של Ω ($\Sigma \subseteq 2^{\Omega}$).
ו P היא פונקציה: $P : \Sigma \to [0,1]$.

לאחר מכן הגדרנו משתנים מקרים, אלו הם פונקציות הממפות תוצאות ב - Ω למרחב אחר (לדוגמא: המספרים הממשיים, או המספרים השלמים או הוקטורים הממשיים וכו').
ז"א משתנה מקרי X הוא בעצם $X(\omega)$ כאשר $\omega \in \Omega$. ז"א $X : \Omega \to D$ כאשר D הוא $\mathbb{Z}$ או $\mathbb{R}$ או $\mathbb{R}^n$ וכו'.

בקורס זה נדון במרחבי הסתברות ובמשתנים מקריים יותר מעניינים: תהליכים סטוכסטיים.

תהליך סטוכסטי הוא משתנה מקרי כפי שמוגדר לעיל אבל D הוא מרחב של פונקציות. ז"א תהליך סטוכסטי הוא "פונקציה אקראית".

במקום שעבור כל $\omega \in \Omega$, $X(\omega)$ ימופה לערך מספרי, מתקיים כי $X(\omega)$ ממופה לפונקציה. ניתן לרשום פו' זו כ $X(\omega, t)$, כאשר $t \in T$.

דרך אחרת לחשוב על תהליך סטוכסטי היא כעל אוסף של משתנים מקריים סקלרים אשר מאופיינים ע"י הפרמטר $t \in T$. כאשר לרוב קיימת תלות סטטיסטית מסוימת בין המשתנים המקריים הללו.

את T נכנה **מרחב הפרמטר** של הפונקציה האקראית. לרוב T מסמל זמן (ומכאן השם *תהליך* סטוכסטי).
בקורס זה T יהיה לרוב $\mathbb{R}^+ = [0, \infty)$ או $\mathbb{N}$. כאשר T הוא $\mathbb{R}^+$ נאמר כי התהליך הוא **תהליך בזמן רציף**, וכאשר T הוא $\mathbb{N}$, נאמר כי התהליך הוא **תהליך בזמן בדיד**.

הערה: בחלק מהמקומות בקורס נתייחס לטבעיים ($\mathbb{N}$) כמכילים את 0, ובחלק לא, זאת בהתאם לנוחות.

קבוצת הערכים אשר $X(\omega, t)$ מקבלת (הטווח של הפונקציה האקראית) היא **מרחב המצבים** של התהליך.
לפעמים נסמל את מרחב המצבים ב S. מרחב מצבים אפשרי ונוח לשימוש בהרבה מקרים הוא $\mathbb{N}$. מרחב מצבים אחר הוא $\{0,1,\ldots,N\}$ (קבוצה סופית). ניתן גם לבחור את מרחב המצבים להיות רציף ($\mathbb{R}$), אבל כמעט ולא נפגוש תהליכים סטוכסטיים כאלו בקורס.

לסיכום: תהליך סטוכסטי הוא משתנה מקרי $X : \Omega \to D$ כאשר D הוא אוסף הפונקציות המקבלות ערכים ב T (מרחב הפרמטר) וממופות לS (מרחב המצבים). או לחלופין תהליך סטוכסטי הוא אוסף של משתנים מקריים סקלרים אשר מקבלים ערכים במרחב המצבים S, ומאופיינים ע"י פרמטר $t \in T$.

חלק א: מבוא

עבור תוצאה נתונה מתוך מרחב המדגם Ω, $\omega_0 \in \Omega$, נקראה לפונקציה $X(\omega_0,t)$ (של t) המתקבלת **ריאליזציה** של התהליך הסטוכסטי.

תהליכים סטוכסטיים כמודלים מתמטים:

את רוב התהליכים אשר נפגוש בקורס זה ניתן ליישם לבניית מודלים מתמטים עבור תופעות שכיחות. תופעות כגון גודלי אוכלוסיה, הכנסות/הוצאות של בתי עסק, מלאי של בנק דם, שיחות טלפון אשר מגיעות למרכזייה ועוד. בכל הדוגמאות לעיל קיים **המרכיב התהליכי והמרכיב הסטוכסטי** (אקראי).
המרכיב התהליכי: שינוי מצב מערכת על פני זמן.
המרכיב הסטוכסטי: אקראיות בשינוי המערכת.

דוגמא א-1:

נתחיל בתיאור של תהליך דטרמיניסטי (ללא אקראיות) ולאחר מכן "נשדרג" אותו לתהליך סטוכסטי.

הסיפור: חשבון הבנק שלנו.
נבנה תהליך אשר מתאר את כמות הכסף אשר יש לנו בחשבון הבנק. נניח כי אנחנו חיים בזמן בדיד ….n=1,2, (כאשר n=1 הוא יום פתיחת חשבון הבנק שלנו). מרחב המצבים של התהליך יהיה רציף (נמדל אותו כרציף) שקלים (ייתכן גם חיובי וגם שלילי).

המודל הדטרמיניסטי:

- S(n) – ערך חשבון הבנק שלנו בזמן n.
- S(1) – הערך ההתחלתי של חשבון הבנק – נניח 200.
- נניח כי כל שלושים ימים (בדיוק) אנחנו מקבלים משכורת, ערך המשכורת הוא 2000 במשכורת הראשונה ובכל משכורת ישנה עלייה של 1% בשכר ביחס למשכורת של החודש הקודם. נסמן ב P(i) את המשכורת אשר אנו מקבלים ביום ה – i:

$$P(i)=\begin{cases}2000(1.01)^{\frac{i}{30}-1} & i=30k \\ 0 & i\neq 30k\end{cases} \quad k\in\mathbb{N}$$

- נניח בנוסף כי בכל יום יש לנו הוצאה של 70 שקלים באופן קבוע לאורך החיים.
  נסמן הוצאה זו ב E(i). 70=E(i) לכל i.

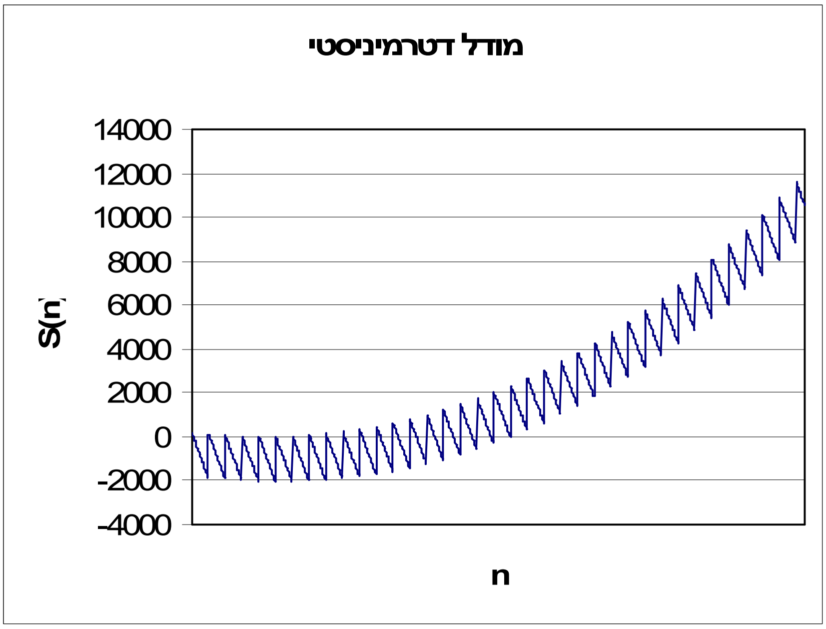
כעת ניתן לרשום את המשוואה עבור תהליך הערך של חשבון הבנק שלנו:

$$S(n)=S(1)+\sum_{i=1}^{n}P(i)-\sum_{i=1}^{n}E(i)$$

את התפתחות התהליך הדטרמיניסטי S(n) ניתן בקלות לחשב (לדוגמא ב Excel). ערך התהליך עבור כל $n\in\{1,...,1095\}$ (שלוש שנים) מוצג באיור בהמשך. ברור כי "שיני המסור" אשר מופיעים באיור הינם תוצאה של ההוצאה היומית האחידה וקפיצות בהכנסה בכל 30 יום עקב קבלת משכורת. רואים כי בחודשים הראשונים, בערך עד חצי התקופה (שנה וחצי בערך) ערך חשבון הבנק לעיתים שלילי ולאחר מכן הוא רק חיובי (המשכורת עולה כל חודש).

חלק א: מבוא

להלן הריאליזציה של המודל הדטרמיניסטי בExcel.



שדרוג למודל הסטוכסטי:

כעת נוסיף מימד אקראי למודל שלנו. את המודל נרשום באופן הבא:

$\widetilde{S}(n)=S(1)+\sum_{i=1}^{n}\widetilde{P}(i)-\sum_{i=1}^{n}\widetilde{E}(i)$ (בדיוק כמו המודל הדטרמיניסטי אבל גדלים אקראיים נסמן עם טילדה(~) מעליהם).

להלן הפרטים:

- S(1) נשאר ללא שינוי (מהמודל הדטרמיניסטי).
- למשכורת נכניס מימד של גידול אקראי. נאמר שאם המשכורת בחודש j היא P אזי המשכורת בחודש i+1 היא $P\cdot(1+U_i)$ כאשר $U_i\sim Uniform(0,0.02)$. מידול זה משקף כי בכל חודש הגידול במשכורת הוא בממוצע כמו במודל הדטרמיניסטי אבל בעל השתנות אקראית. כעת ניתן לרשום את ההכנסה היומית כתוצאה ממשכורת $(\widetilde{P}(i))$ כך:

$$\widetilde{P}(i)=\begin{cases}2000\prod_{j=1}^{\frac{i}{30}-1}(1+U_j) & n=30k\\ 0 & n\neq 30k\end{cases}\quad k\in\mathbb{N}$$

כאשר $\{U_j, j\in\mathbb{N}\}$ היא סדרה של משתנים מקריים אחידים i.i.d. על הקטע $[0,0.02]$.

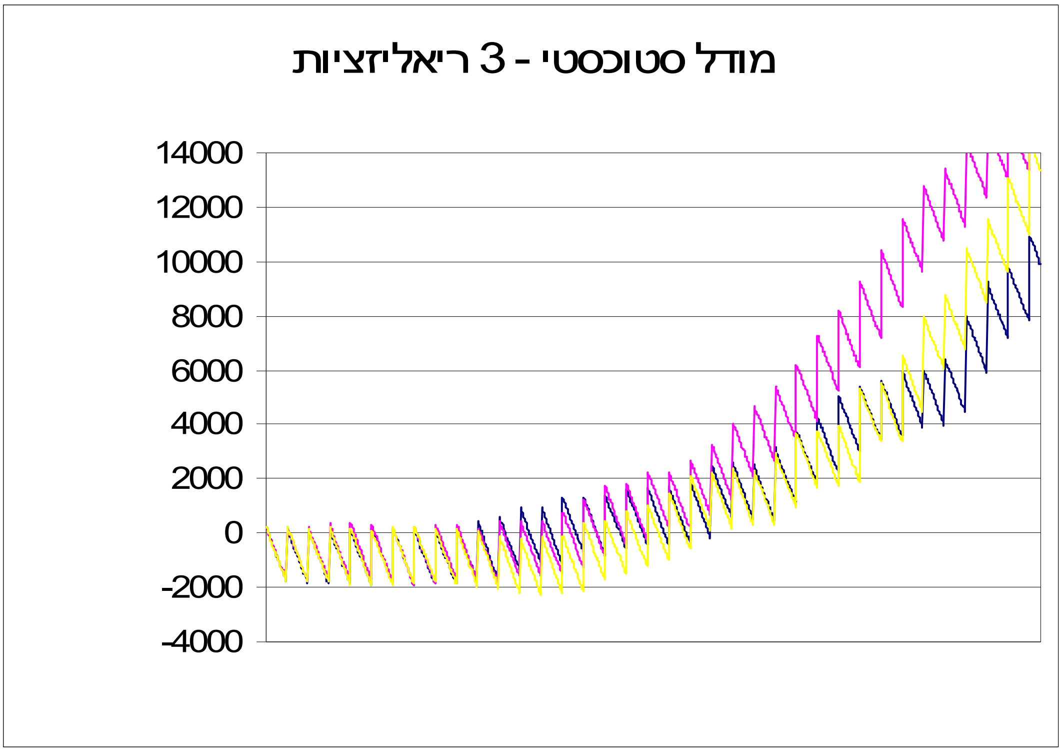
הערה: $\prod_{i=1}^{0}f(i)=1$ (הגדרה).

- נכניס גם מימד של אקראיות להוצאה היומית: $\widetilde{E}(i)\sim Uniform(50,90)$

חלק א: מבוא

נשים לב כי בנינו את המודל הסטוכסטי כך שכל הרכיבים האקראיים שלו הינם בעלי תוחלת הזהה לרכיבים המקבילים במודל הדטרמיניסטי.

להלן 3 ריאליזציות של המודל הסטוכסטי (גם Excel ע"י שימוש בפו' ()rand).



שאלות מעניינות לגבי תכונות של תהליכים סטוכסטיים:

בהינתן מודל של תהליך סטוכסטי מה ניתן לעשות איתו? לרוב שאלה זו תלויה במודל. למרות זאת ישנם מספר מאפיינים של תהליכים סטוכסטיים אשר מעניינים אותנו ביותר:

1. חוק ההסתברות (התפלגות) של ערכי התהליך בזמן מסוים $t \in T$.
2. התוחלת, השונות, המומנטים של ערכי התהליך בזמן מסוים $t \in T$.
3. הקשר הסטוכסטי בין ערך התהליך בשני זמנים מסוימים.
4. ההסתברות כי התהליך יכנס למצב מסוים או אוסף מצבים ולעולם לא יצא ממצב זה.
5. התפלגות הזמן עד הגעה למצב מסוים או אוסף מצבים.
6. התוחלת, שונות, מומנטים של הסעיף הקודם.
7. חוק ההסתברות (התפלגות) מס' החזרות למצב מסוים ביחידת זמן.
8. קיום התפלגות גבולית.
9. ארוגודיות.
10. תכונות נוספות.

דוגמאות/משפחות של תהליכים סטוכסטיים:

חלק א: מבוא

להלן רשימה של משפחות של תהליכים סטוכסטיים:

1. סדרה של משתנים מקריים בלתי תלויים.
2. סדרת משתני ברנולי.
3. מהלך מקרי פשוט.
4. מהלך מקרי כללי.
5. תהליכים גאוסים.
6. תהליכי ספירה.
7. תהליכי חידוש.
8. תהליכי פואסון.
9. שרשראות מרקוב.
10. תהליכי קפיצה מרקובים.
11. תהליכי לידה – מוות.
12. תהליכי הסתעפות.
13. תנועת בראון.

יישומים:

1. מידול אותות בתקשורת רדיו.
2. מידול עומס ברשתות תקשורת חוטיות.
3. מידול מצב תעסוקתי במודלים של פנסיה.
4. מידול תביעות לחברת ביטוח.
5. מידול תהליכי יצור.
6. מידול עומסים על מערכת מחשב מרובת משאבים.
7. מידול של גדלי אוכלוסיות.
8. יישומים רבים נוספים.

סימולציה ע"י מחשב:

בקורס זה, נכיר מספר מודלים בסיסיים של תהליכים סטוכסטיים ונלמד כיצד לענות על שאלות לגבי התהליכים: התפלגות התהליך בזמן מסוים וכו'. הדרך ובה נעבוד היא הדרך האנליטית: נגדיר בכל פעם את התהליך, נגדיר את ההנחות ההסתברותיות ולאחר מכן נשתמש בתכונות מיוחדות של התהליך לצורך חישוב הגדלים אשר מעניינים אותנו.

דרך חלופית להתמודדות עם בעיות מסוג זה היא באמצעות **סימולציה ע"י מחשב**. כאשר חוקרים תכונות של תהליכים סטוכסטיים ע"י סימולציה ממוחשבת פשוט נותנים למחשב להריץ ריאליזציה אחת ארוכה, או אוסף רב של ריאליזציות ובוחנים את ההתנהגות התהליך המתואר ע"י המחשב באמצעות הסקה סטטיסטית.

סקירת חלקי הקורס:

הקורס מחולק לחמישה חלקים (א'-ה') ובכל חלק בין חמישה לעשרה פרקים (סה"כ כ- 35 פרקים), משך ההרצה של כל פרק הוא בערך כשעה. בתחילת הקורס נבצע חזרה מקיפה על הסתברות, זהו חלק (א').

חלק א: מבוא

לאחר מכן נעבור לחלק (ב') ובו נשתעשע בשרשראות מרקוב, מודל שימושי זה נפוץ כמעט בכל תחומי המדע ובפרקטיקה. המושג המרכזי אשר נפגוש בפרק של שרשראות מרקוב הוא מושג הארוגודיות.
בחלק הבא (ג') נרחיב את תחום העניין שלנו לתהליכים בעלי מרחב מצבים רציף מהסוג הפשוט והפופולארי ביותר, אלו הם תהליכים פואסון. תהליך פואסון הינו מושג בסיסי וחשוב בתורת ההסתברות ובאמצעותו ניתן למדל תופעות רבות בטבע. בחלק לאחר מכן (ד') נפגוש בתהליכי קפיצה מרקובים. תהליכים אלו הינם הרחבה של תהליכי פואסון לשרשראות מרקוב בזמן רציף.
בחלק הבא (ה') ננתח תהליכי קפיצה מרקובים בעלי מבנה מסוים, פשוט אך שימושי, אלו הם תהליכי לידה-מוות. ניישם תהליכי לידה ומוות שונים למודלים של תורת התורים. בנוסף נדון בתורת התורים באופן כללי. ניתן להתייחס לפרק זה כמבוא לתורת התורים.
לבסוף בחלק ו' נסכם בקצרה ונציין נושאים אלמנטאריים נוספים בתהליכים סטוכסטיים אשר לא נלמדו בקורס זה.

פרק א-2: חזרה על הסתברות, ותוצאות מתמטיות נוספות שימושיות.

תוצאות מתמטיות שימושיות:

1. הבינום של ניוטון.

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

2. טור גיאומטרי.

$$\sum_{k=0}^{n} a^k = \frac{a^{n+1}-1}{a-1}$$

$$|a|<1, \sum_{k=0}^{\infty} a^k = \frac{1}{1-a}$$

3. טור טלסקופי.

$$\sum_{k=0}^{N} (A_k - A_{k+1}) = A_0 - A_{N+1}$$

4. ייצוג exp(x) ע"י טור טילור.

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

5. נוסחת פסקל:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

6. 4 מקרים קומבינאטורים פשוטים – פיזור k כדורים ל n תאים.
   a. n^k
   b. $n(n-1)...(n-k+1) = \binom{n}{k} k!$
   c. $\binom{n}{k}$
   d. המקרה היותר קשה הוא כדורים זהים ללא הגבלה של מקום בכל תא: $\binom{k+n-1}{k}$

דוגמא א-2:
כמה פתרונות במספרים שלמים אי-שליליים יש למשוואה $a_1 + a_2 + ... + a_n = k$?
שאלת ספירה זו זהה לשאלה בכמה דרכים ניתן לפזר k כדורים לn תאים כאשר אין הגבלה על מספר הכדורים אשר ניתן להכניס לתא והכדורים זהים. (שזו גם בדיוק השאלה, בכמה דרכים ניתן לדגום k פריטים מאוכלוסיה בעלת n פריטים עם החזרות (שקול ליותר מכדור אחד בתא) וסדר הדגימה לא משנה (שקול למקרה בו הכדורים זהים).

7. נוסחת המשולש

המקרה הבדיד:

$$\sum_{k=0}^{N} kf(k) = \sum_{k=1}^{N} kf(k)$$
$$= f(1) +$$
$$f(2) + f(2) +$$
$$f(3) + f(3) + f(3) +$$
$$....$$
$$= \sum_{k=1}^{N}\sum_{j=k}^{N} f(j)$$

המקרה הרציף:

$$\int_0^\infty xf(x)dx = \int_0^\infty \int_0^x 1dt f(x)dx$$
$$= \int_{x=0}^{\infty}\int_{t=0}^{x} f(x)dtdx = \int_{t=0}^{\infty}\int_{x=t}^{\infty} f(x)dxdt$$

8. קירוב Stirling ל – n!.

$$n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$$

הערה: הסימון $f_n \sim g_n$ מסמל כי $\lim_{n\to\infty} \frac{f_n}{g_n} = 1$.

חזרה כללית על הסתברות:

כנאמר בפרק הקודם, משתנה מקרי הוא פונקציה ממרחב המדגם אל מרחב של מספרים כלשהו. לרוב נקטלג את המשתנים המקריים אשר אנחנו מכירים למשתנים מקריים רציפים ולמשתנים מקריים בדידים אבל לפעמים ישנם גם משתנים מקריים מעורבים.

אנו מתעניינים בחוק ההסתברות של משתנים מקריים. ניתן לתאר את חוק ההסתברות במספר דרכים. הדרך המקובלת ביותר היא **פונקציות ההתפלגות המצטברת** וזאת בגלל שתיאור זה מתאים גם למשתנים מקריים בדידים וגם לרציפים (וגם למעורבים).

$$F_X(x) = P(X \le x) = P(\{w : X(w) \le x\})$$

חלק א: מבוא

כזכור זוהי פונקציה לא יורדת ומתקיים כי

$$\lim_{x\to-\infty} F_X(x)=0$$

$$\lim_{x\to\infty} F_X(x)=1$$

לפעמים נוח יותר לעבוד עם המשלים של פונקצית ההתפלגות המצטברת, זוהי **פונקצית השרידות**:

$$\overline{F}_X(x)=P(X>x)=1-F_X(x)$$

עבור משתנים מקריים בדידים קיימת **פונקצית מסת הסתברות**:

$$P_X(x)=P(X=x).$$

ואז מתקיים:

$$F_X(x)=\sum_{k=-\infty}^{x} P_X(k).$$

עבור משתנים מקריים רציפים קיימת **פונקצית צפיפות** $f_X(x)$. לערך פונקציה זו אין משמעות הסתברותית בשל עצמו אבל מתקיים:

$$\Delta x f_X(x)\cong P(x\le X<x+\Delta x)$$

ואז מתקיים:

$$F_X(x)=\int_{-\infty}^{x} f_X(t)dt.$$

התוחלת של משתנה מקרי היא:

$$EX=\int_{-\infty}^{\infty} x f_X(x)dx$$

$$EX=\sum_{x=-\infty}^{\infty} xP_X(x)$$

עבור המקרה הרציף והבדיד בהתאמה.

עבור משתנה מקרי X ניתן להגדיר משתנה מקרי חדש Y=g(X) כאשר g(.) היא פונקציה כלשהי.
את חוק ההסתברות של המתנה המקרי החדש (Y) ניתן לחשב בדרכים אשר נלמדו בקורס תורת ההתפלגויות (לדוגמה אם $Z\sim N(0,1)$ אז $Y=Z^2\sim Gamma(1/2,1/2)\equiv\chi^2(1)$).

את התוחלת של Y ניתן לחשב עלפי הנוסחאות לעיל (ע"י שימוש בצפיפות או פונקצית מסת ההסתברות אשר חושבה עבור Y). דרך אלטרנטיבית (ולרוב יותר קלה) היא להשתמש בתכונה הבא של התוחלת:

$$EY=Eg(X)=\int_{-\infty}^{\infty} g(x)f_X(x)dx$$

ובאופן דומה עבור המקרה הבדיד.

המומנט ה – k (k=1,2,3,…) של משתנה מקרי X הוא EX^k. (ניתן לחשב עפ"י הנוסחה לעיל כאשר $g(x)=x^k$).

השונות של משתנה מקרי היא $Var(X)=E\left\{(X-EX)^2\right\}=E\left\{X^2\right\}-(EX)^2$.

מה יודעים על תוחלות/שונויות של סכומים/מכפלות?

- תוחלת של סכום היא סכום התוחלות (לא משנה אם תלויים או לא).
- תוחלת של מכפלה היא מכפלת התוחלת כאשר המשתנים המקריים בלתי תלויים.
- שונות של סכום של משתנים מקריים **בלתי מתואמים** (יותר חלש מבלתי תלויים) הוא סכום השונויות.

עבור משתנה מקרי חיובי מתקיים השוויון הבא:

$$EX=\int_0^\infty \overline{F}_X(x)dx$$

$$EX=\sum_{k=0}^\infty \overline{F}_X(k)$$

עבור המקרה הרציף והבדיד בהתאמה.

הוכחה עבור המקרה הבדיד:
נשתמש בנוסחת המשולש (הוצגה בסעיף הקודם):

$$EX=\sum_{k=0}^\infty kP_X(k)=\sum_{k=1}^\infty\sum_{j=k}^\infty P_X(j)$$

$$=\sum_{k=1}^\infty P(X\ge k)=\sum_{k'=0}^\infty P(X\ge k'+1)$$

$$=\sum_{k'=0}^\infty P(X>k')=\sum_{k'=0}^\infty \overline{F}_X(k')$$

הוכחה עבור המקרה הרציף:

$$EX=\int_0^\infty xf_X(x)dx$$

$$=\int_{t=0}^\infty\int_{x=t}^\infty f(x)dxdt=\int_{t=0}^\infty \overline{F}_X(t)dt$$

פונקציה יוצרת מומנטים של משתנה מקרי X היא $M_X(t)=Ee^{tX}$.

זוהי פונקציה של t. היא מעניינת בגלל מספר סיבות.

א) פונקצית יוצרת מומנטים מגדירה באופן חד-חד ערכי (חח"ע) את חוק ההסתברות של המשתנה המקרי. זאת אומרת שבמידה וזיהינו את הפונקציה יוצרת מומנטים של משתנה מקרי, אז גילינו את התפלגותו!!!

הערה 1: במקרים חריגים בהם הפו' אינה מוגדרת בסביבה של t=0, אז זה לא מתקיים אבל לא נתעסק במקרים כאלו בקורס זה.

הערה 2: המעבר מצפיפות/מסת הסתברות לפו' יוצרת מומנטים הוא לרוב קל (חישוב

תוחלת), המעבר בחזרה הוא קשה (דורש אינטגרציה במישור מרוכב – "קשה") לכן הדרך הנוחה לעשות זאת היא בעזרת טבלה.

ב) ע"י גזירה k פעמים והצבה t=0 ניתן לקבל את המומנט ה k.

ג) הפונקציה של סכום של שתי מ"מ בלתי תלויים, היא מכפלת הפונקציות. נראה בהמשך.

(וריאציות אחרות הן **התמרת לפלס** $L_X(s)=Ee^{-sX}$ **ופונקציה אופיינית** $\Phi_X(t)=Ee^{itX}$, אבל אנו נשתמש לרוב בפו' יוצרת מומנטים בקורס זה).

עבור משתנים מקריים בדידים אי-שליליים מוגדרת **פונקציה יוצרת הסתברות** (לפעמים נקראת פשוט פונקציה יוצרת): $G_X(t)=Et^X$.

קיים קשר פשוט בין פונקציה יוצרת הסתברות לפונקציה יוצרת מומנטים: $M_X(t)=G_X(e^t)$.

הערה: פו' יוצרת קיימת גם לסדרות באופן כללי (הסדרה לא חייבת להיות פו' מסת הסתברות).

<u>משתנים מקריים ממשפחות מוכרות:</u>

בקורס מבוא להסתברות ותורת ההתפלגויות הכרנו משפחות רבות של משתנים מקריים, בדידים ורציפים. לכל משפחה יש פרמטר אחד או יותר ולפעמים גם ישנו "סיפור" (בעיקר עבור הבדידים). לא נחזור כאן על הפירוט של המשתנים המקריים (ניתן לראות אותם בטבלת התפלגויות) אבל נציין את החשובים והרלוונטיים לקורס זה:

1. ברנולי.
2. בינומי.
3. היפר-גיאומטרי.
4. גיאומטרי – סופר כישלונות.
5. גיאומטרי – סופר ניסיונות.
6. בינומי שלילי – סופר כישלונות.
7. בינומי שלילי – סופר ניסיונות.
8. פואסון.
9. אחיד בדיד.
10. אחיד (רציף).
11. אקספוננציאלי.
12. גאמא.
13. ארלנג (מקרה פרטי של גאמא).
14. נורמאלי.

<u>משתנים מקריים החיים יחדיו במרחבי הסתברות:</u>

עבור שני משתנים מקריים X,Y, כך מוגדרת פונקצית ההתפלגות המשותפת:
$F_{X,Y}(x,y)=P(X\le x,Y\le y)$ (כאשר פסיק בין "מאורעות" מסמל חיתוך מאורעות).

מכאן: $F_X(x)=P(X\le x,Y\le\infty)=\lim_{y\to\infty}F_{X,Y}(x,y)$ (וכנ"ל עבור Y).

עבור שני משתנים מקריים בדידים, פונקצית מסת ההסתברות המשותפת מוגדרת כך:

$P_{X,Y}(x,y)=P(X=x,Y=y)$.

כאשר 2 משתנים מקריים (X_1,X_2) הינם i.i.d. (Independent and identically distributed) אזי $F_{X_1,X_2}(x_1,x_2)=F_{X_1}(x_1)F_{X_2}(x_2)$. או באופן כללי עבור וקטור $\overrightarrow{X}$ של משתנים מקריים i.i.d. באורך n (בעל n משתנים מקריים) מתקיים $F_{\overline{X}}(x_1,...,x_n)=\prod_{i=1}^{n}F_X(x_i)$. אותה תוצאה נכונה עבור פו' מסת הסתברות משותפת של אוסף i.i.d. ועבור הצפיפות המשותפת של אוסף i.i.d.

סכומים של משתנים מקריים:

סכימה של משתנים מקריים הינה תופעה שכיחה ביותר במודלים הסתברותיים ובתהליכים סטוכסטיים.
בעקרון כל תהליך סטוכסטי בזמן בדיד S_n ניתן להגדיר כסכום של משתנים מקריים (לא בהכרח בלתי תלויים). נגדיר:

$$X_0=S_0$$
$$X_n=S_n-S_{n-1}$$

ואז:

$$S_n=\sum_{k=0}^{n}X_k.$$

במידה והמשתנים המקריים X_n הינם בלתי תלויים אזי $\{S_n,n\geq 0\}$ נקרא מהלך מקרי.

ניתן לחשב את חוק ההסתברות המשותף של סכום של משתנים מקריים בלתי תלויים במספר דרכים.
הדרך ה"ישירה" היא באמצעות פעולת הקונבולוציה, דרך אחרת (ולרוב יותר פשוטה) היא ע"י הכפלה של פונקציות יוצרות מומנטים או פונקציות יוצרות הסתברות, לבסוף כאשר הסכום הוא של אוסף רב של משתנים מקריים אז ניתן להשתמש במשפטי גבול.

קונבולוציה:

עבור המקרה הבדיד, קונבולוציה היא פעולה הפועלת על שתי סדרות של מספרים ופולטת סדרה שלישית.
עבור המקרה הרציף, קונבולוציה היא פעולה אשר פועלת על שתי פונקציות ופולטת פונקציה שלישית.

בהינתן סדרה של מספרים $a=\{a_n,n\geq 0\}$ וסדרה נוספת $b=\{b_n,n\geq 0\}$, נגדיר את הקונבולוציה של a ו b כך:

$$c_n=(a\otimes b)_n=\sum_{i=0}^{\infty}a_i b_{n-i}.$$

נוצרה כאן סדרה חדשה של מספרים c_n, כאשר הערך של הסדרה ב $n=n_0$ כלשהו הוא: $\sum_{i=0}^{\infty}a_i b_{n_0-i}$.

לפעולת הקונבולוציה שימושים רבים בניתוח מערכות ליניאריות ובעיבוד אותות, אבל השימושים אשר מעניינים אותנו הם בהסתברות:

חלק א: מבוא

נראה כי עבור שני משתנים מקריים בדידים,חיוביים בלתי תלויים X,Y מתקיים כי פונקצית מסת ההסתברות של Z=X+Y היא הקונבולוציה של פונקציות המסה של X ושל Y.

נסתכל על המאורע: $[X+Y=n]$, מאורע זה ניתן לכתיבה כאיחוד זר של מאורעות אשר מתארים את הערכים הספציפיים אשר X ו Yמקבלים:

$$[X+Y=n]=\bigcup_{i=0}^{\infty}[X=i,Y=n-i].$$

ולכן:

$$\begin{aligned}P_Z(n) &= P_{X+Y}(n)=P(X+Y=n)\\ &= P\left(\bigcup_{i=0}^{\infty}[X=i,Y=n-i]\right)\\ &= \sum_{i=0}^{\infty}P(X=i,Y=n-i)\\ &= \sum_{i=0}^{\infty}P(X=i)P(Y=n-i)\\ &= \sum_{i=0}^{\infty}P_X(i)P_Y(n-i)=(P_X\otimes P_Y)(n)\end{aligned}$$

ניתן להגדיר קונבולוציה באופן דומה עבור פונקציות (פו' צפיפות של משתנים מקריים רציפים) ואין הכרחי כי הפונקציות/סדרות יהיו בעלי תומך חיובי בלבד – לא נעשה זאת כאן.

דוגמא א-3:

$X_1,X_2\sim Bernulli(p)$ בלתי תלויים.

נגדיר $N_2=X_1+X_2$.

אזי $P(N_2=k)=(P_{X_1}\otimes P_{X_2})(k)$.

$$(P_{X_1}\otimes P_{X_2})(k)=\sum_{i=0}^{\infty}P_{X_1}(i)P_{X_2}(k-i)$$

ראשית נבחין כי עבור k>2 ועבור k<0 הסכום לעיל אינו מכיל איברים חיוביים.
נותר עם כך לחשב עבור k=0,1,2.

עבור k=0 הסכום מכיל איבר חיובי יחיד (כאשר i=0): q^2.
עבור k=1 הסכום מכיל 2 איברים חיוביים (כאשר i=0): qp (כאשר i=1): pq. ולכן סה"כ: $2pq$.
עבור k=2 הסכום מכיל איבר חיובי יחיד (כאשר i=1): p^2

קבלנו אם כך כמצופה כי $N_2\sim Bin(p,2)$.

דוגמא א-4:

$X_1\sim Poisson(\lambda_1)$
$X_2\sim Poisson(\lambda_2)$
והם בלתי תלויים.

חלק א: מבוא

כיצד מתפלג $Z = X_1 + X_2$?

$$P(Z=k) = (P_{X_1} \otimes P_{X_2})(k)$$
$$= \sum_{i=0}^{\infty} P_{X_1}(i) P_{X_2}(k-i)$$
$$= \sum_{i=0}^{k} P_{X_1}(i) P_{X_2}(k-i)$$
$$= e^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{k} \frac{\lambda_1^{\,i} \lambda_2^{\,k-i}}{i!(k-i)!}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_{i=0}^{k} k! \frac{\lambda_1^{\,i} \lambda_2^{\,k-i}}{i!(k-i)!}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda_1^{\,i} \lambda_2^{\,k-i}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} (\lambda_1+\lambda_2)^k$$

קבלנו שסכום של שתי מ"מ פואסונים מתפלג פואסונית עם סכום העוצמות (הפרמטרים).

שימוש בפונקציה יוצרת מומנטים:

פעולת הקונבולוציה היא לפעמים מסובכת (הן אנליטית והן נומרית). מדוע אנליטית? כי ככה יוצא. ומדוע נומרית? נניח כי ברצונו לבצע קונבולוציה של שתי פונקציות מסת הסתברות, כל אחת בעלת תומך המכיל כ – 1000 ערכים, אז מספר ההכפלות אשר עלינו לבצע הוא מסדר גודל של כמיליון.

ייצוג חוק ההסתברות ע"י פונקציה יוצרת מומנטים (או גם פונקציה יוצרת הסתברות) יכול להקל על המלאכה (לזכור כי X ו Y בלתי תלויים):

$$M_{X+Y}(t) = Ee^{(X+Y)t} = Ee^{Xt}e^{Yt} = Ee^{Xt}Ee^{Yt} = M_X(t)M_Y(t)$$

החישוב דומה עבור פונקציה יוצרת הסתברות.

זאת אומרת שלאחר המרה של חוקי ההסתברות למרחב t, הפעולה המסובכת של סכום משתנים מקריים (ולכן קונבולוציה פו' מסת הסתברות) מתפשטת לפעולה של מכפלה של פונקציות יוצרות מומנטים/הסתברות.

<u>דוגמא א-5:</u>

$X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$ בלתי תלויים.

כיצד מתפלג $Z = X_1 + X_2$?

ידוע כי $M_{X_i}(t) = e^{\mu_i t + \frac{\sigma_i^2 t^2}{2}}$ (לא קשה לחשב זאת על פי הגדרה).

$$M_Z(t)=M_{X_1}(t)M_{X_2}(t)=e^{\mu_1 t+\frac{\sigma_1^2t^2}{2}}e^{\mu_2 t+\frac{\sigma_2^2t^2}{2}}=e^{(\mu_1+\mu_2)t+\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}$$

וזוהי פו' יוצרת מומנטים של משתנה מקרי $N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

משפטי גבול:

כאשר הסכום המקרי S_n הוא בעל הרבה מחוברים (X_n) (n גדול). והמחוברים הינם .i.i.d. בעלי תוחלת ושונות סופית אזי ניתן ליישם משפטי גבול.

החוק החלש של המספרים הגדולים:

משפט:

$\lim_{n\to\infty}P(|\frac{S_n}{n}-EX|>\varepsilon)=0$ לכל $\varepsilon>0$.

הערה: קל להוכיח משפט זה באמצעות אי שוויון צבישב/מרקוב.

החוק החזק של המספרים הגדולים:

משפט:

$P(\lim_{n\to\infty}|\frac{S_n}{n}-EX|>\varepsilon)=0$ לכל $\varepsilon>0$.

ניתן גם לומר זאת כך:

לכל $\omega\in\Omega$ בעלת הסתברות חיוביות $\lim_{n\to\infty}\frac{S_n}{n}=EX$

הערה: לזכור כי S_n הוא פונקציה של ω.

הערה: החוק החזק והחוק החלש נראים דומים. למרות זאת החוק החזק אומר הרבה יותר (ולכן נקרא חזק). הוא אומר שעבור כל ריאליזציה אפשרית של אוסף משתנים .i.i.d, הממוצע שואף לתוחלת (אין ריאליזציות בעלות הסתברות חיוביות אשר עבורן הממוצע אינו שואף לתוחלת). החוק החלש, בסך הכול טוען כי ככל שמגדילים את n (את המדגם) אז ההסתברות כי הממוצע יהיה שונה מהתוחלת הולכת וקטנה.
במילים אחרות, החוק החזק מסתכל על כל התהליך לעומת החוק החלש אשר מסתכל על הפילוג השולי עבור n נתון.

משפט הגבול המרכזי:

משפט:

$$\lim_{n\to\infty}F_{\frac{S_n-nEX}{\sqrt{nVar(X)}}}\left(x\right)=F_Z(x)$$

כאשר Z משתנה מקרי נורמאלי סטנדרטי.

דוגמא:

את נוסחת סטירלינג ($n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$). ניתן להוכיח במדויק ע"י קירוב של אינטגרל של log.
עכשיו נראה הסבר אלטרנטיבי לנכונותה של הנוסחה.

יהיו $\{X_n, n \geq 1\}$ אוסף משתנים מקריים i.i.d. המתפלגים Poisson(1).
ז"א $EX_i = 1$ וגם $Var(X_i) = 1$.
נסכל על הסכום $S_n = \sum_{k=1}^{n} X_i$.
אזי $ES_n = n$. וגם $Var(S_n) = n$.
בנוסף ידוע כי S_n מתפלג Poisson(n).

עכשיו עבור n גדול $\frac{S_n - n}{\sqrt{n}}$ מתפלג בקירוב נורמאלית סטנדרטית
ולכן ניתן לכתוב:

$$P(S_n = n) = P(n-1 < S_n \leq n) = P(-1 < S_n - n \leq 0)$$

$$= P(-\frac{1}{\sqrt{n}} < \frac{S_n - n}{\sqrt{n}} \leq 0) \approx \int_{-1/\sqrt{n}}^{0} (2\pi)^{-1/2} e^{-x^2/2} dx$$

ערך זה שווה בקירוב ל $\frac{1}{\sqrt{2\pi n}}$ (קירוב של האינטגרל באזור הנקודה 0).

אבל בנוסף מתקיים $P(S_n = n) = \frac{e^{-n} n^n}{n!}$.
ולכן

$$\frac{e^{-n} n^n}{n!} \approx \frac{1}{\sqrt{2\pi n}}.$$

ומכאן מייד נובעת נוסחת סטרלינג.

סטטיסטי סדר:

ניקח אוסף משתנים מקריים i.i.d. $X_1, ..., X_n$ (לפעמים נקרא לאוסף שכזה מדגם מקרי).

נניח כאן לצורך הדיון כי כל ערכי המדגם שונים זה מזה (זה קורה בהסתברות 1 במידה ו X מ"מ רציפים).

סטטיסטי הסדר ה - k של המדגם המקרי הוא הערך של המשתנה המקרי מתוך המדגם כך ש k-1 מהמשתנים במדגם קטנים ממנו ו n-k משתנים במדגם גדולים ממנו. (כך שסטטיסטי הסדר ה – 1 הוא המינימום וסטטיסטי הסדר ה n הוא המקסימום).

ניתן להגדיר זאת במדויק כך:

חלק א: מבוא

$$X_{(1)} = \min_{1\le i\le n} X_i$$
$$\langle 1\rangle = \arg\min_{1\le i\le n} X_i$$

$$X_{(2)} = \min_{\substack{1\le i\le n\\ i\ne\langle 1\rangle}} X_i$$
$$\langle 2\rangle = \arg\min_{\substack{1\le i\le n\\ i\ne\langle 1\rangle}} X_i$$

....

....

$$X_{(k)} = \min_{\substack{1\le i\le n\\ i\ne\langle 1\rangle,\dots,\langle k-1\rangle}} X_i$$
$$\langle k\rangle = \arg\min_{\substack{1\le i\le n\\ i\ne\langle 1\rangle,\dots,\langle k-1\rangle}} X_i$$

...

...

$$X_{(n)} = \min_{\substack{1\le i\le n\\ i\ne\langle 1\rangle,\dots,\langle n-1\rangle}} X_i = \max_{1\le i\le n} X_i$$
$$\langle n\rangle = \arg\min_{\substack{1\le i\le n\\ i\ne\langle 1\rangle,\dots,\langle n-1\rangle}} X_i = \arg\max_{1\le i\le n} X_i$$

אזי $X_{(k)}$ הוא סטטיסטי הסדר ה – k ו $\langle k\rangle$ הוא האינדקס של המשתנה המקרי אשר הוא סטטיסטי הסדר ה-k . (ז"א מתקיים $X_{(k)} = X_{\langle k\rangle}$).

נשאל מספר שאלות לגבי חוק ההסתברות של סטטיסטי הסדר ה k ושל כל הוקטור $(X_{(1)},\dots,X_{(n)})$.

נתחיל לצורך חימום עם מציאת חוק ההסתברות של המינימום:

$$\overline{F}_{X_{(1)}}(x) = P(X_{(1)} > x) = P(X_1 > x,\dots,X_n > x) = P(X_1 > x)\dots P(X_n > x)$$
$$= (P(X > x))^n = (\overline{F}_X(x))^n$$

באופן דומה, חוק ההסתברות של המקסימום:
$$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P(X_1 \le x,\dots,X_n \le x) = (F_X(x))^n .$$

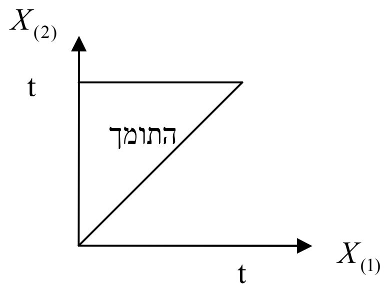
הצפיפות המשותפת של $X_{(1)},\dots,X_{(n)}$ (של הוקטור $(X_{(1)},\dots,X_{(n)})$) היא:
$$f_{(X_{(1)},\dots,X_{(n)})}(x_1,\dots,x_n) = n!\prod_{j=1}^{n} f_X(x_j) I_C^{(x_1,\dots,x_n)}$$

חלק א: מבוא

כאשר הקבוצה C (התומך של הצפיפות המשותפת) מוגדרת כך:
$C = \{(x_1,...,x_n) \in \mathbb{R}^n : x_1 < x_2 < ... < x_n\}$.

דוגמא:

$X_i \sim Uniform(0,t)$ עבור i=1,2 והם בלתי תלויים.
אזי ההתפלגות המשותפת של $(X_{(1)}, X_{(2)})$ היא "אחידה" על התומך:



שטח התומך הוא $\frac{t^2}{2}$ וגובה הצפיפות או $\frac{2}{t^2}$ (על פי הנוסחה לעיל).

הוכחה של נוסחת הצפיפות המשותפת של סטטיסטי הסדר:
תהי A תת קבוצה של המרחב: $A \subseteq \mathbb{R}^n$.
תהי σ פרמוטציה (תמורה) על $\{1,...,n\}$.
תהי Σ קבוצת כל $n!$ הפרמוטציות על $\{1,...,n\}$.
יהי X' הוקטור הסדור $(X_{(1)},...,X_{(n)})$ כאשר $<1>,<2>,...,<n>$ נתונים ע"י σ.
יהי $C = \{(x_1,...,x_n) \in \mathbb{R}^n : x_1 < x_2 < ... < x_n\}$.
אז:

$$P(X' \in A) = P(X' \in A \cap C) = P(X_\sigma \in A \cap C) =$$
$$\sum_{\sigma_0 \in \Sigma} (X_\sigma \in A \cap C, \sigma = \sigma_0)$$

נשים לב שאם $X_{\sigma_0} \in A \cap C$ אזי $\sigma_0 = \sigma$ ואז

$$P(X_{\sigma_0} \in A \cap C) = P(X \in A \cap C)$$

ולכן

$$P(X' \in A) = \sum_{\sigma_0 \in \Sigma} (X \in A \cap C) = n! P(X \in A \cap C) = n! \int_{A \cap C} f_X(x) dx = n! \int_A f_X(x) I_C^{(x)} dx$$

מ.ש.ל.

פרק א-3: הסתברות מותנית, התפלגות מותנית, תוחלת מותנית.

הסתברות מותנית מהווה מרכיב מרכזי בניתוחים של הרבה מהתהליכים הסטוכסטיים אשר נלמד בקורס זה. כנלמד במבוא להסתברות ההגדרה של **ההסתברות המותנית** של מאורע A בהינתן מאורע B (אשר הסתברותו חיובית ממש) היא:

$$P(A\mid B)=\frac{P(AB)}{P(B)}.$$

הגדרה זו משקפת את השינוי של חלוקת ההסתברויות של התוצאות לאור האינפורמציה אשר קיבלנו.

שני מאורעות הינם **בלתי תלויים** אם P(AB)=P(A)P(B). מושג האי-תלות מאפשר לנו לקבץ יחדיו 2 (או יותר) מרחבי הסתברות.
במידה ו A ו B בלתי תלויים אז P(A|B)=P(A).

<u>דוגמא:</u>
האם ייתכנו שתי מאורעות זרים A,B בעלי הסתברות חיוביות ממש בלתי תלויים?
לא, נניח בשלילה כי A ו B בלתי תלויים אז $P(AB)=P(A)P(B)>0$. אבל אם A ו B זרים אז $AB=\varnothing$ ולכן P(AB)=0. סתירה.

בהינתן אוסף מאורעות $\{B_\gamma,\gamma\in\Gamma\}$ זרים בזוגות, מתקיימת **נוסחת ההסתברות השלמה:**
$P(A)=\sum_{\substack{\gamma\in\Gamma\\ P(B_\gamma)>0}}P(A\mid B_\gamma)P(B_\gamma)$ כאשר ניתן להתייחס לסכום כאינטגרל במידה והקבוצה של המאורעות Γ אינה בת-מנייה.

נוסחת בייס היא דרך לחשב את $P(B_\gamma\mid A)$ (כאשר ידוע לנו $P(A\mid B_\gamma)$ ו $P(B_\gamma)$ עבור כל γ).

$$P(B_\gamma\mid A)=\frac{P(B_\gamma A)}{P(A)}=\frac{P(A\mid B_\gamma)P(B_\gamma)}{\sum_{\substack{\gamma\in\Gamma\\ P(B_\gamma)>0}}P(A\mid B_\gamma)P(B_\gamma)}$$

<u>דוגמא א-15:</u>
מערכת תקשורת מכילה משדר המשדר אחד משתי האותות '0' או '1' ומקלט הנמצא במרחק רב מהמשדר ומנסה לפענח את האות אשר המשדר שידר. אותות '0' משודרים בהסתברות 1-p ואותות '1' משודרים בהסברות p. בגלל המרחק הרב בין המשדר למקלט, מתווסף רעש משמעותי לאות הנקלט וישנה הסתברות של α כי המקלט החליט כי קלט '1' למרות שבעצם שודר '0' או החליט כי קלט '0' למרות שבעצם שודר '1'.
נתון כי המקלט קלט '1' , מה ההסתברות שמהשדר באמת שידר '1'?

נסמן T_i - שידור של i.
נסמן R_i - קליטה של i.
אז לפי בייס:

$$P(T_1\mid R_1)=\frac{P(R_1\mid T_1)P(T_1)}{P(R_1\mid T_1)P(T_1)+P(R_1\mid T_0)P(T_0)}=\frac{(1-\alpha)p}{(1-\alpha)p+\alpha(1-p)}$$

דוגמא:

במשחק טלוויזיה משתתף צריך לבחור 1 מ-3 דלתות (מאחורי 2 דלתות אין כלום ומאחורי דלת אחת יש מיליון דולר). המשתתף בחר את דלת מס' 1 ולאחר מכן המארח של המשחק פתח את דלת מס' 2 והראה למשתתף שאין כלום מאחורי דלת זו.
בשלב זה המשתתף נשאל האם ברצונו לדבוק בדלת מס' 1 או לשנות את בחירתו לדלת מס' 3.

א) מהי ההסתברות לזכייה במידה ויישאר עם דלת מס' 1.
ב) מהי ההסתברות לזכייה במידה וישנה את בחירתו לדלת מס' 3.

נסמן ב A_i את המאורע שהמיליון נמצא מאחורי דלת i.
נסמן ב – E את המאורע שהמנחה פתח את דלת מס' 2 לאחר שהמשתתף בחר את דלת מס' 1.
ידוע ש $P(A_i)=\frac{1}{3}$ ולכן גם $P(\overline{A_i})=\frac{2}{3}$.
בנוסף:

$$P(E\mid A_1)=\frac{1}{2}$$

$$P(E\mid A_2)=0$$

$$P(E\mid A_3)=1$$

בנוסף ידוע כי

$$P(E)=P(E\mid A_1)P(A_1)+P(E\mid A_2)P(A_2)+P(E\mid A_3)P(A_3)=\frac{1}{2}*\frac{1}{3}+0*\frac{1}{3}+1*\frac{1}{3}=\frac{1}{2}$$

ולכן:

א) $$P(A_1\mid E)=\frac{P(E\mid A_1)P(A_1)}{P(E)}=\frac{\frac{1}{2}*\frac{1}{3}}{\frac{1}{2}}=\frac{1}{3}$$

ב) $$P(A_3\mid E)=\frac{P(E\mid A_3)P(A_3)}{P(E)}=\frac{1*\frac{1}{3}}{\frac{1}{2}}=\frac{2}{3}$$

העובדה כי חיתוך של מאורעות ניתן לביטוי כמכפלה של הסתברות בהסתברות מותניית ($P(AB)=P(A\mid B)P(B)$) באה לידי ביטוי גם **בנוסחת השרשרת**:

$$\begin{aligned}
&P(B_1...B_n) = \\
&P(B_n \mid B_1...B_{n-1})P(B_1...B_{n-1}) = \\
&P(B_n \mid B_1...B_{n-1})P(B_{n-1} \mid B_1...B_{n-2})P(B_1...B_{n-2}) = \\
&... = \\
&P(B_n \mid B_1...B_{n-1})P(B_{n-1} \mid B_1...B_{n-2})...P(B_2 \mid B_1)P(B_1)
\end{aligned}$$

חוק ההסתברות המותנה של משתנים מקריים:

להלן פונקצית מסת ההסתברות המותניית עבור משתנים מקריים בדידים:

$$P_{X|Y}(x \mid y) = P(X = x \mid Y = y) = \frac{P_{X,Y}(x,y)}{P_Y(y)} .$$

עבור משתנים מקריים בלתי תלויים מתקיים: $P_{X,Y}(x,y) = P_X(x)P_Y(y)$.

ולכן במקרה זה $P_{X|Y}(x \mid y) = P_X(x)$

להלן פונקצית הצפיפות המותניית עבור משתנים מקריים רציפים:

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

כאשר נכפיל את צד שמאל ב Δx ואת צד ימין ב $\frac{\Delta x \Delta y}{\Delta y}$ נקבל.

$f_{X|Y}(x \mid y)\Delta x = \frac{f_{X,Y}(x,y)\Delta x \Delta y}{f_Y(y)\Delta y}$ ולשוויון זה יש משמעות הסתברותית:

$$\cong \frac{P(x < X \le x + \Delta x,\ y < Y \le y + \Delta y)}{P(y < Y \le y + \Delta x)} = P(x < X \le x + \Delta x \mid y < Y \le y + \Delta y)$$

כמו במקרה הבדיד, במקרה הרציף מתקיים כי עבור משתנים מקריים בלתי תלויים:

$$f_{X|Y}(x \mid y) = f_X(x)$$

<u>דוגמא א-8:</u>

$X_1 \sim Poisson(\lambda_1)$, $X_2 \sim Poisson(\lambda_2)$ והם בלתי תלויים.

ידוע כי $X_1 + X_2 = n$, כיצד מתפלג X_1 .

צריך למצוא את ההתפלגות המותניית של X_1 בהינתן $X_1 + X_2 = n$.

ברור שיש לדון בערכים $k = 0,...,n$ (ההסתברות היא אפס עבור ערכים שונים מאלו).

<u>חלק א: מבוא</u>

$$P(X_1=k\mid X_1+X_2=n)$$
$$=\frac{P(X_1=k,X_1+X_2=n)}{P(X_1+X_2=n)}$$
$$=\frac{P(X_1=k,X_2=n-k)}{P(X_1+X_2=n)}$$
$$=\frac{P(X_1=k)P(X_2=n-k)}{P(X_1+X_2=n)}$$
$$=\frac{e^{-\lambda_1}\frac{\lambda_1^{\ k}}{k!}e^{-\lambda_2}\frac{\lambda_2^{\ n-k}}{(n-k)!}}{\frac{e^{-\lambda_1+\lambda_2}(\lambda_1+\lambda_2)^n}{n!}}=\binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-k}$$

קבלנו אם כך כי $X_1\mid X_1+X_2=n\sim Bin(n,\frac{\lambda_1}{\lambda_1+\lambda_2})$

<u>דוגמא:</u>

$X_1\sim Bin(n_1,p)$, $X_2\sim Bin(n_2,p)$ והם בלתי תלויים.

ידוע כי $X_1+X_2=r$, כיצד מתפלג X_1.

$$P(X_1=k\mid X_1+X_2=r)$$
$$=\frac{P(X_1=k,X_1+X_2=r)}{P(X_1+X_2=r)}$$
$$=\frac{P(X_1=k,X_2=r-k)}{P(X_1+X_2=r)}$$
$$=\frac{P(X_1=k)P(X_2=r-k)}{P(X_1+X_2=r)}$$
$$=\frac{\binom{n_1}{k}\binom{n_2}{r-k}p^k(1-p)^{n_1-k}p^{r-k}(1-p)^{n_2-(r-k)}}{\binom{n_1+n_2}{r}p^r(1-p)^{n_1+n_2-r}}$$
$$=\frac{\binom{n_1}{k}\binom{n_2}{r-k}}{\binom{n_1+n_2}{r}}$$

קבלנו אם כך כי $X_1\mid X_1+X_2=r\sim Hyperg(n_1+n_2,n_1,r)$

מקרה פרטי: אם ב – n ניסויים ישנם r הצלחות אז הסיכוי כי ההצלחה בניסוי ה i הוא:

$P(X_i=1|\sum_{j=1}^{n}X_j=r)=\frac{r}{n}$. (זאת ע"י התוצאה הקודמת כאשר $n_1=1, n_2=n-1$ ו k=1).

דוגמא:

יהיו X,Λ שני משתנים מקריים.
נניח כי פונקצית הצפיפות המשותפת שלהם היא:

$$f_{X,\Lambda}(x,\lambda)=\frac{1}{2}\lambda e^{-x\lambda}I_{\{[0,\infty)\}}(x)I_{\{[0,2]\}}(\lambda)$$

זהו למעשה "מודל תערובת" ובו X מתפלג אקספוננציאלית עם פרמטר Λ שהוא בעצמו משתנה מקרי (יוניפורמי [0,2] במקרה זה). בו נראה:

נחשב $f_{X|\Lambda}(x|\lambda)=\frac{f_{X,\Lambda}(x,\lambda)}{f_{\Lambda}(\lambda)}$ עבור $\lambda\in[0,2]$.

ראשית:

$$f_{\Lambda}(\lambda)=I_{\{[0,2]\}}(\lambda)\int_0^{\infty}\frac{1}{2}\lambda e^{-x\lambda}dx=I_{\{[0,2]\}}(\lambda)\frac{1}{2}$$

אז אכן $\Lambda\sim Uniform(0,2)$

ולכן אכן:

$$f_{X|\Lambda}(x|\lambda)=\frac{f_{X,\Lambda}(x,\lambda)}{f_{\Lambda}(\lambda)}=\lambda e^{-x\lambda}I_{\{[0,\infty)\}}(x)$$

תוחלת ושונות מותניית:

נגדיר את התוחלת המותניית להיות התוחלת המחושבת באמצעות חוק ההתפלגות המותנה:

$E[X|Y=y]=\sum_{x=-\infty}^{\infty}xP_{X|Y}(x|y)$ במקרה הבדיד.

או

$E[X|Y=y]=\int_{-\infty}^{\infty}xf_{X|Y}(x|y)dx$ במקרה הרציף.

חשוב לשים לב ש E[X|Y] היא פו' של המשתנה המקרי Y. ולכן E[X|Y] הוא משתנה מקרי בעצמו.

דוגמא:

נתונה סדרה $X_1,...,X_n$ i.i.d. של משתנים ברנולי (בעלי הסתברות הצלחה p). נתון כי מספר ההצלחות בסדרה הוא k ($0\le k\le n$), מה תוחלת מספר ההצלחות ב m הניסיונות הראשונים ($m\le n$)?

ברצוננו לחשב את $E[\sum_{i=1}^{m}X_i|\sum_{j=1}^{n}X_j=k]$.

מתקיים:

<u>חלק א: מבוא</u>

$$E[\sum_{i=1}^{m} X_i \mid \sum_{j=1}^{n} X_j = k] = \sum_{i=1}^{m} E[X_i \mid \sum_{j=1}^{n} X_j = k]$$

ובנוסף מתקיים $E[X_i \mid \sum_{j=1}^{n} X_j = k] = P(X_i = 1 \mid \sum_{j=1}^{i-1} X_j + X_i + \sum_{j=i+1}^{n} X_j = k) = \dfrac{k}{n}$

ולכן $E[\sum_{i=1}^{m} X_i \mid \sum_{j=1}^{n} X_j = k] = m\dfrac{k}{n}$

<u>משפט התוחלת המותנית</u>: $E[E[X \mid Y]] = E[X]$.

כנאמר קודם, E[X|Y] הוא משתנה מקרי ולכן המשמעות של התוחלת החיצונית במשפט היא תוחלת על המשתנה המקרי הנ"ל (זאתי תוחלת המחשבים לפי חוק ההסברות של Y).
הוכחה (למקרה הרציף):

$$E[E[X \mid Y]] = E[\int_{-\infty}^{\infty} x f_{X|Y}(x \mid y) dx] = E[\int_{-\infty}^{\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)} dx] =$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)} dx f_Y(y) dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y) dx dy = \int_{-\infty}^{\infty} x f_X(x) dx = EX$$

ההוכחה למקרה הבדיד דומה.

<u>דוגמא:</u>

$$Var(X) = E[X^2] - (E[X])^2 = EE[[X^2 \mid Y]] - (E[E[X \mid Y]])^2$$

פרק א-4: דוגמאות לשימוש בהתניה.

חישובי תוחלת של משתנים מקריים בעזרת תוחלת מותניית:

כזכור משתנה מקרי גיאומטרי – סופר ניסיונות הוא בעל תוחלת $\frac{1}{p}$ להלן החישוב הישיר:

$$\sum_{k=1}^{\infty} k(1-p)^{k-1} p = p\sum_{k=1}^{\infty}\sum_{j=k}^{\infty}(1-p)^{j-1}$$
$$= p\sum_{k=1}^{\infty}\sum_{j'=0}^{\infty}(1-p)^{j'+k-1}$$
$$= p\sum_{k=1}^{\infty}(1-p)^{k-1}\sum_{j'=0}^{\infty}(1-p)^{j'}$$
$$= \frac{p}{1-p}\sum_{k=1}^{\infty}(1-p)^{k}\sum_{j'=0}^{\infty}(1-p)^{j'}$$
$$= \frac{p}{1-p}\frac{1-p}{1-(1-p)}\frac{1}{1-(1-p)}$$
$$= \frac{1}{p}$$

חישוב זה אינו ארוך מדי, אבל למרות זאת נראה כיצד ניתן להגיע לתוצאה בצורה אלגנטית באמצעות תוחלת מותניית:

יהיה T המשתנה המקרי הגיאומטרי סופר הניסיונות. נגדיר משתנה מקרי חדש Y כך ש Y מקבל 1 אם בניסיון הראשון מצליחים, אחרת Y הוא 0.

על פי משפט התוחלת המותניית:

$$ET = E[E[T \mid Y]] = E[T \mid Y=0]P(Y=0) + E[T \mid Y=1]P(Y=1) = (E[T]+1)q + 1p \, .$$

כעת נפתור עבור ET:

$$ET - qET = q + p$$
$$ET(1-q) = q + p$$
$$ET = \frac{1}{p}$$

סכום אקראי:

יהיו $X_1, X_2, ...$ סדרה של משתנים מקריים .i.i.d.

כבר דנו בפילוג של סכום של n משתנים מקריים כאלו, אבל מה אם מספר המשתנים המקריים אשר אנו מחברים הוא אקראי? (נסמנו N).

זהו מודל מתאים לתביעות אשר מגיעות לחברת ביטוח בפרק זמן נתון (ידוע מה ההתפלגות של גודל כל תביעה, וידועה ההתפלגות של מספר התביעות). מעוניינים בחוק ההסתברות (או לפעמים רק בתוחלת) של סכום התביעות:

$S=\sum_{n=1}^{N}X_n$.

נחשב את E[S]:

$$E[S]=E[\sum_{k=1}^{N}X_k]=E[E[\sum_{k=1}^{N}X_k\mid N]]$$

אבל $E[\sum_{k=1}^{N}X_k\mid N]=N\,E[X]$

ולכן $E[S]=E[N\,E[X]]=E[X]\;E[N]$

ז"א כמצופה, התוחלת של סכום אקראי היא תוחלת מספר האיברים בסכום כפול תוחלת גודל האיברים.

נחשב את Var(S):

כמובן שמתקיים

$$Var(S)=E[S^2]-E[S]^2=E[S^2]-E[N]^2E[X]^2$$

אם כך ראשית נחשב את המומנט השני של S ($E[S^2]$):

$$E[S^2]=E[(\sum_{k=1}^{N}X_k)^2]=E[E[(\sum_{k=1}^{N}X_k)^2\mid N=n]]$$

אבל

$$\begin{aligned}
&E[(\sum_{k=1}^{N}X_k)^2\mid N=n]=E[(\sum_{k=1}^{n}X_k)^2]=E[(X_1+\ldots+X_n)(X_1+\ldots+X_n)]=\\
&E[{X_1}^2+X_1X_2+X_1X_3+\ldots+X_1X_n+\\
&X_2X_1+{X_2}^2+X_2X_3+\ldots+X_2X_n+\\
&+\\
&\ldots\\
&+\\
&X_nX_1+X_nX_2+\ldots+{X_n}^2]
\end{aligned}$$

בגלל האי-תלות של הסדרה $X_1,X_2,\ldots.$ אז גורמים מהסוג $E[X_iX_j]$ כאשר $i\neq j$ הם $E[X_i]E[X_j]$

ולכן סכום זה שווה ל –

$$E[{X_1}^2]+\ldots+E[{X_n}^2]+\sum_{i\neq j}E[X_i]E[X_j]$$

n המחוברים הראשונים בביטוי לעיל כולם שווים ל $E[X^2]$. שאר המחוברים גם הם שווים וכולם שווים ל $E[X]^2$. כמה כאלו יש? $\binom{n}{2}2=n(n-1)$.

אם כך קיבלנו:

$$E[(\sum_{k=1}^{N}X_k)^2\mid N=n]=nE[X^2]+n(n-1)E[X]^2=nE[X^2]+n^2E[X]^2-nE[X]^2$$

ולכן

חלק א: מבוא

$$E[S^2]=E[E[(\sum_{k=1}^{N}X_k)^2\mid N=n]]=E[NE[X^2]+N^2E[X]^2-NE[X]^2]=$$
$$E[N]E[X^2]+E[N^2]E[X]^2-E[N]E[X]^2=E[N]Var(X)+E[N^2]E[X]^2$$

ולכן

$$Var(S)=E[S^2]-E[N]^2E[X]^2=E[N]Var(X)+E[N^2]E[X]^2-E[N]^2E[X]^2=$$
$$E[N]Var(X)+Var(N)E[X]^2$$

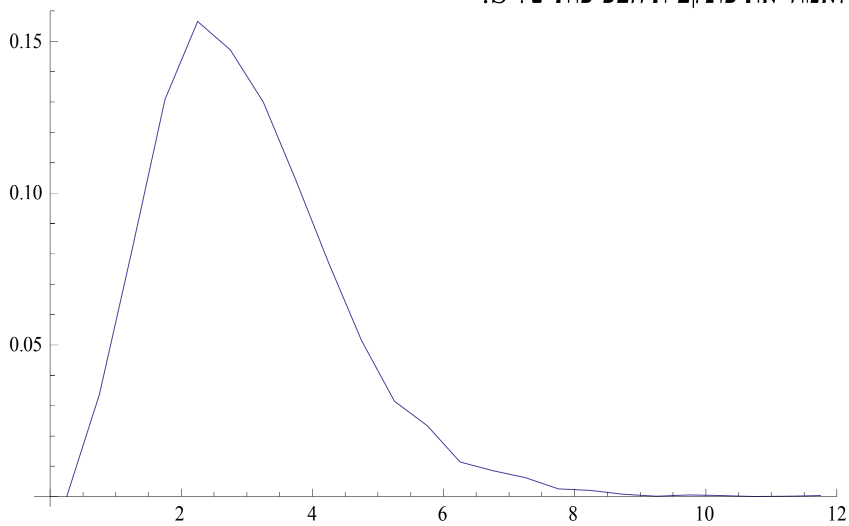
נחשב את חוק ההתפלגות של סכום אקראי:

אז ראינו מהי התוחלת ומהי השונות של סכום אקראי. אבל כיצד מתפלג סכום אקראי?

דוגמא חישובית:

ניקח $X_1,X_2,...\sim \exp(1)$, $N\sim Bin(10,\frac{1}{2})$

ניתן בקלות לסמלץ ערכים של $S=\sum_{i=1}^{N}X_i$ באמצעות מחשב. לאחר סימולציה של 5000 ערכים כאלו, ניתן לאמוד את פונקצית הצפיפות של S:



כאלטרנטיבה, ניתן באופן מפורש לקבל ביטוי עבור הפילוג של S באמצעות פונקציה יוצרת מומנטים. נניח שפונקציות יוצרות המומנטים של X ושל N נתונות: $M_N(.)$, $M_X(.)$.

נחשב את הפונקציה יוצרת מומנטים של S:

$$M_S(t)=Ee^{t\sum_{k=1}^{N}X_k}=E[E[e^{t\sum_{k=1}^{N}X_k}\mid N]]$$

אבל $E[e^{t\sum_{k=1}^{N}X_k}\mid N]=Ee^{t\sum_{k=1}^{N}X_k}=\prod_{k=1}^{N}Ee^{tX_k}=(M_X(t))^N$

ולכן $M_S(t)=E[(M_X(t))^N]=E[e^{N\log(M_X(t))}]=M_N(\log(M_X(t))$.

דוגמא:
נניח כי $N\sim Geometric(p)$ (סופר ניסיונות). ו $X\sim\exp(\lambda)$.
להלן פו' יוצרות המומנטים:
$t<\lambda$ $M_X(t)=\dfrac{\lambda}{\lambda-t}$ (ראה פרק ג-1).

$$M_N(t)=\sum_{k=1}^{\infty}e^{tk}p(1-p)^{k-1}=\frac{p}{1-p}\sum_{k=1}^{\infty}\left(e^t(1-p)\right)^k$$

$$=\left(\frac{p}{1-p}\right)\left(\frac{e^t(1-p)}{1-e^t(1-p)}\right)=\frac{pe^t}{1-e^t(1-p)} \qquad e^t(1-p)<1$$

אם כך קבלנו ש

$$M_S(t)=M_N(\log(M_X(t))=\frac{p\frac{\lambda}{\lambda-t}}{1-\frac{\lambda}{\lambda-t}(1-p)}=\frac{p\lambda}{\lambda-t-\lambda(1-p)}=\frac{p\lambda}{\lambda p-t}.$$

ואם כך $S\sim\exp(p\lambda)$.

ניתוח זמן הריצה הממוצע של אלגוריתם quick sort:

אלגוריתם מיון הוא אלגוריתם אשר מקבל מדגם מקרי $x_1,...,x_n$ ופולט את סטטיסטי הסדר $x_{(1)},...,x_{(n)}$.
במילים אחרות האלגוריתם ממין את הסדרה $x_1,...,x_n$.
ישנם אלגוריתמים רבים אשר יכולים לבצע פעולה זאת (ייתכן ובקורס שפת C או מבוא למדעי המחשב פגשתם מספר אלגוריתמים כאלו). אחד האלגוריתמים הפופולאריים והשימושיים ביותר הוא quick sort.
האלגוריתם מוגדר באופן הבא (נניח כאן כי ערכי המדגם הינם שונים זה מזה):
כאשר n=1 אין מה למיין.
כאשר n=2 האלגוריתם ממיין את שתי הערכים באופן טריוויאלי (ממקם את הקטן לפני הגדול).
כאשר n>2 האלגוריתם פועל כך: באקראי (אחיד בדיד) האלגוריתם בוחר אחד מ – n הערכים (נאמר x_i).
ומשווה את כל אחד מה n-1 ערכים ל x_i. נסמן ב S_i את קבוצת הערכים הקטנים מ - x_i ו ב $\overline{S}_i$ את קבוצת הערכים הגדולים מ x_i. לאחר מכן האלגוריתם חוזר באופן רקורסיבי על הפעולה אשר תוארה לעיל על S_i ועל $\overline{S}_i$ כאילו ואלו היו הקלטים.

הרצת דוגמא:
$X=(10,5,8,2,1,4,7)$
תחילה נבחר אחד מערכים אלו באקראי (לכל ערך יש הסתברות של 1/7 להיבחר).
נניח כי נבחר 4.
עכשיו האלגוריתם משווה את 4 לכל אחד מהערכים האחרים ומקבלים.
$S_i=(2,1)$
$\overline{S}_i=(10,5,8,7)$
כעת יש לחזור על האלגוריתם על S_i ועל $\overline{S}_i$.
על S_i: בגלל ש n=2 המיון הוא פשוט ומקבלים (1,2).

חלק א: מבוא

על $\overline{S}_i$: נבחר באקראי ערך, נאמר כי נבחר 7, אזי מקבלים

$$S'_i = (5)$$

$$\overline{S'}_i = (8,10)$$

כעת יש לחזור על האלגוריתם על S'_i ועל $\overline{S'}_i$.

על S'_i אין מה למיין.

על $\overline{S'}_i$ המיון הוא שוב טריוויאלי (n=2) וגם כאן הערכים כבר ממוינים.
סך הכול קבלנו:

$$(\tilde{S}_i, 4, (\tilde{S}'_i, 7, \tilde{\overline{S}}'_i)) =$$

$$((1,2), 4, ((5), 7, (8,10)))$$

(כאשר הסימון ~ מתאר קבוצה בגודל 2 או 1 אשר מוינה באופן טריוויאלי).
הסוגריים מתארים את הקריאות הרקורסיביות אשר האלגוריתם ביצע. וכאשר מורידים את הסוגריים אז מקבלים את הסדרה הממוינת.

כאשר דנים ביעילות של אלגוריתמי מיון אז נהוג לדון במספר ההשוואות אשר האלגוריתם נדרש לבצע. לרוב נהוג לתאר מספר זה כפונקציה של גודל המדגם (n) (ומחפשים אלגוריתמים אשר בהם מספר ההשוואות לא גדל יותר מדי מהר כאשר n גדל).

כאן ננתח את תוחלת מספר ההשוואות אשר quick sort מבצע. נסמן את ערך זה ב M_n.
נגדיר את $M_{n|j}$ להיות תוחלת מספר ההשוואות על קלט בגודל n כאשר נתון כי הערך אשר נבחר הוא ה j הכי קטן. (j=1,…,n).
ברור כי $M_{n|j} = (n-1+M_{j-1}+M_{n-j})$ (כאשר $M_0 = 0$) וזאת כי יש לבצע n-1 השוואת בין הערך הנבחר לשאר הערכים ולאחר מכן לחזור על הפעולה על הקבוצה S_i ובה j-1 ערכים ועל הקבוצה $\overline{S}_i$ ובה n-j ערכים.
עכשיו, בגלל שהערכים נבחרים על פי התפלגות אחידה בדידה אזי בהנחה שהקלט הוא אקראי לחלוטין, קל לראות כי ההסתברות שהערך הנבחר הוא ה j הכי קטן היא $\frac{1}{n}$. מכאן:

$$M_n = \sum_{j=1}^{n} M_{n|j}\frac{1}{n} = \sum_{j=1}^{n}(n-1+M_{j-1}+M_{n-j})\frac{1}{n} =$$

$$n-1+\frac{1}{n}((M_0+M_{n-1})+(M_1+M_{n-2})+...+(M_{n-2}+M_1)+(M_{n-1}+M_0)) =$$

$$n-1+\frac{1}{n}2\sum_{k=1}^{n-1}M_k$$

מכאן:

$$nM_n = n(n-1)+2\sum_{k=1}^{n-1}M_k$$

או עבור n+1:

$$(n+1)M_{n+1} = (n+1)n+2\sum_{k=1}^{n}M_k$$

<u>חלק א: מבוא</u>

נחסיר את שתי המשוואות לקבל:

$$(n+1)M_{n+1} - nM_n = 2n + 2M_n$$

או:

$$(n+1)M_{n+1} = 2n + (2+n)M_n$$

ולכן:

$$\frac{M_{n+1}}{n+2} = \frac{2n}{(n+2)(n+1)} + \frac{M_n}{n+1}$$

עכשיו קבלנו ביטוי עבור $\frac{M_{n+1}}{n+2}$ אשר מורכב מאותו ערך עבור n יותר קטן ב-1 ותוספת ולכן ניתן לבצע הצבות חוזרות:

$$\frac{M_{n+1}}{n+2} = \frac{2n}{(n+2)(n+1)} + \frac{M_n}{n+1}$$
$$= \frac{2n}{(n+2)(n+1)} + \frac{2(n-1)}{(n-1+2)(n-1+1)} + \frac{M_{n-1}}{n-1+1}$$
$$\dots$$
$$= \sum_{k=0}^{n-1} \frac{2(n-k)}{(n+2-k)(n+1-k)} + \frac{M_1}{1}$$

ולכן (ע"י ארגון מחדש והעובדה ש $M_1 = 0$):

$$M_{n+1} = 2(n+2)\sum_{i=1}^{n} \frac{i}{(i+1)(i+2)}$$

ניתן להראות כי גודל זה שווה בקירוב ל $2(n+2)\log_2(n+2)$.

(ולכן אומרים כי מספר ההשוואות של quick sort על קלט בגודל n הוא בממוצע מסדר גודל של $n\log n$).

פרק א-5: תהליכי ברנולי – I.

נחשוב על ניסוי ובו $\Omega = \{\omega : \omega = (\omega_1, \omega_2, ...), \omega_i = 1,0\}$. ז"א מרחב המדגם הוא אוסף כל הסדרות הבינאריות האינסופיות (ניתן גם לחשוב על 1 כהצלחה ו על 0 ככישלון ואז מרחב המדגם שקול לאוסף כל סדרות הטלות המטבע האינסופיות). נאמר שפונקצית ההסתברות P, היא כזאת הנותנת הסתברות $p \in [0,1]$ לכל מאורע $A_i \in \Sigma$ מהסוג $A_i = \{\omega_i = 1\}$ ו-q=1-p עבור המאורע המשלים.

נגדיר משתנים מקריים $X_1, X_2, ...$ כך ש $X_i(\omega) = I_{A_i}(\omega)$. ז"א $P(X_i = 1) = EX_i = p$. אלו הם כמובן משתנים מקריים ברנולי .i.i.d.

משלב זה והלאה כבר לא נדון ב Ω, Σ אלה רק נדון בתהליך ברנולי .i.i.d ובתהליכים אשר יגזרו ממנו.

<u>הגדרה:</u>
תהליך ברנולי .i.i.d: הוא אוסף משתנים מקריים .i.i.d ברנולי אם פרמטר p. נסמן ב $\{X_n, n = 1,2,...\}$ או ב $\{X_n, n \geq 1\}$.

באמצעות תהליך זה ניתן למדל תופעות רבות לדוגמא סיביות (bits) אשר מגיעות ממקור לא ידוע (דיסק, אינטרנט, תקשורת סלולארית) וזאת במידה וערך כל סיבית אינו תלוי בסיבית הקודמת או ההבאה.

ברור כי:
$$EX_n = p$$

קל לראות גם כי:
$$Var(X_n) = pq$$

בנוסף קל לראות כי השונות של X_n היא מקסימאלית עבור פרמטר $p = \frac{1}{2}$

$$\frac{d}{dp} p(1-p) = \frac{d}{dp}(p - p^2) = 1 - 2p = 0$$
$$p^* = \frac{1}{2}$$
$$\frac{d^2}{dp^2}(p - p^2) = \frac{d}{dp}(1 - 2p) = -2 < 0$$

תהליך זה אינו כל כך מעניין בשל עצמו, הרי הוא בסך הכול סדרת .i.i.d אבל כעת נבנה ממנו תהליכים נוספים בעלי מבנה קצת יותר מורכב.

<u>תהליך ספירה ברנולי (תהליך בינומי):</u>

נבנה כעת את התהליך הבא:

חלק א: מבוא

$$N_n = \begin{cases} 0 & n=0 \\ X_1+...+X_n & n=1,2,.. \end{cases}$$

או:

$$N_0 = 0$$
$$N_n = N_{n-1} + X_n, n \geq 1$$

זהו **תהליך ספירה ברנולי** (נקרא גם תהליך בינומי). הוא סופר את מספר ההצלחות אשר אירעו עד הזמן n.

מהו EN_n?

$.EN_n = E[X_1+...+X_n] = EX_1+...+EX_n = np$

מהו $Var(N_n)$?

$.Var(N_n) = Var(X_1+...+X_n) = Var(X_1)+...+Var(X_n) = npq$

הערה: עבור כל $\omega \in \Omega$ קיימת ריאליזציה של X ולכן קיימת ריאליזציה של N. התוחלת והשונות של תהליכים אלו (ושל כל תהליך סטוכסטי) הם כבר פונקציות דטרמיניסטיות של n. רואים כי שתיהן עולות ליניארית ב-n עבור התהליך N_n וקבועות עבור התהליך X_n.

אז עכשיו כשידוע לנו התוחלת והשונות של N_n, נעבור לחישוב ההתפלגות של N_n עבור כל n:

ממבוא להסתברות אנו בעצם יודעים כי זוהי התפלגות בינומית עם פרמטרים n ו p:

$P_{N_n}(k) = \binom{n}{k} p^k q^{n-k}$ (עבור k=0,1,…,n).

ההסבר הפשוט ביותר לנוסחה זו הוא: עבור המאורע אשר מתאר כי היו k הצלחות מתוך n ניסויים $\{N_n = k\}$, דרוש איחוד של $\binom{n}{k}$ מאורעות זרים, A_i (כל אחד מהם מתאר בחירה שונה של המאורעות המצליחים). וההסתברות של כל אחד מהמאורעות הזרים הללו (A_i) הינה $p^k q^{n-k}$ וזה בגלל אי תלות של המאורעות הבסיסיים יותר (הברנוליים $\{X_j = 1\}$ או $\{X_j = 0\}$) המרכיבים כל מאורע כזה.

הערה: קיבלנו כאן את ההתפלגות של ערך התהליך עבור כל $n \in \mathbb{N}$ נתון. לעיתים התפלגות זאת נקראת **ההתפלגות השולית של** התהליך בזמן n. בהמשך הקורס, זה יהיה הגודל אשר לרוב יעניין אותנו עבור רוב התהליכים הסטוכסטיים אשר ננתח. ז"א, נציג תהליך סטוכסטי כלשהו וננסה להגיע להתפלגות השולית שלו עבור כל זמן (n או t) נתון. לפעמים גם נתעניין בהתפלגות זו כאשר $n \to \infty$, אבל לא כך המצב עם תהליך ספירה ברנולי כי כאן $P(\lim_{n\to\infty} N_n = \infty) = 1$ ולכן אין משמעות להתפלגות של N_n כאשר n שואף ל- ∞.

כעת נגיע לתוצאה זו (התפלגות הבינומית) בדרך קצת שונה מהרגיל. ראשית נבחין כי:

א) $P_{N_{n+1}}(k) = pP_{N_n}(k-1) + qP_{N_n}(k)$ כי הרי:

$$P(N_{n+1}=k\mid X_{n+1}=0)=P(N_n=k)=P_{N_n}(k)$$
$$P(N_{n+1}=k\mid X_{n+1}=1)=P(N_n=k-1)=P_{N_n}(k-1)$$

ולכן:

$$\begin{aligned}P_{N_{n+1}}(k)&=P(N_{n+1}=k)=\\&=P(N_{n+1}=k\mid X_{n+1}=1)P(X_{n+1}=1)+P(N_{n+1}=k\mid X_{n+1}=0)P(X_{n+1}=0)\\&=P_{N_n}(k-1)p+P_{N_n}(k)q\end{aligned}$$

ב) $P_{N_0}(0)=1$ - על פי הגדרה.

א' ו ב' יוצרים נוסחה רקורסיבית לחישוב $P_{N_n}(k)$:

$$P_{N_1}(0)=pP_{N_0}(-1)+qP_{N_0}(0)=q$$
$$P_{N_1}(1)=pP_{N_0}(0)+qP_{N_0}(1)=p$$

ז"א $P_{N_1}(k)=\binom{1}{k}p^kq^{1-k}$

ובהמשך לזאת:

$$P_{N_2}(0)=pP_{N_1}(-1)+qP_{N_1}(0)=q^2$$
$$P_{N_2}(1)=pP_{N_1}(0)+qP_{N_1}(1)=2pq$$
$$P_{N_2}(2)=pP_{N_1}(1)+qP_{N_1}(2)=p^2$$

ז"א $P_{N_2}(k)=\binom{2}{k}p^kq^{2-k}$

ובהמשך לזאת:

$$P_{N_3}(0)=pP_{N_2}(-1)+qP_{N_2}(0)=q^3$$
$$P_{N_3}(1)=pP_{N_2}(0)+qP_{N_2}(1)=3pq^2$$
$$P_{N_3}(2)=pP_{N_2}(1)+qP_{N_2}(2)=3pq^2$$
$$P_{N_3}(3)=pP_{N_2}(2)+qP_{N_2}(3)=p^3$$

ז"א $P_{N_3}(k)=\binom{3}{k}p^3q^{3-k}$

מסתמן כאן חישוב הדומה לחישוב משולש פסקל רק שחישוב זה לא רק מכיל את המקדמים הבינומיים (כמו במשולש פסקל) אלה בנוסף גורר איתו את ההסתברויות p,q.

הסבר:

<u>חלק א: מבוא</u>

נזכר במשולש פסקל:

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$\cdots$
$n=0$	1	0	0	0	0	$\cdots$
$n=1$	1	1	0	0	0	$\cdots$
$n=2$	1	2	1	0	0	$\cdots$
$n=3$	1	3	3	1	0	$\cdots$
$n=4$	1	4	6	4	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

האיברים בטבלה הם $\binom{n}{k}$.

הטבלה מאותחלת עם

$$\binom{0}{0}=1$$

$$\binom{0}{k}=0, k\neq 0$$

לאחר מכן הערכים מחושבים באופן רקורסיבי ע"י $\binom{n-1}{k}+\binom{n-1}{k-1}=\binom{n}{k}$ וכך כל איבר הוא הסכום של האיבר שמעליו והאיבר שמשמאל לאיבר שמעליו (מניחים עוד עמודת אפסים עבור $k=-1$).

באופן דומה הנוסחה הרקורסיבית אשר כתבנו עבור $P_{N_n}(k)$ נותנת את הטבלה הבאה:

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$\cdots$
$n=0$	1	0	0	0	0	$\cdots$
$n=1$	$1q$	$1p$	0	0	0	$\cdots$
$n=2$	$1q^2$	$2pq$	$1p^2$	0	0	$\cdots$
$n=3$	$1q^3$	$3pq^2$	$3p^2q$	$1p^3$	0	$\cdots$
$n=4$	$1q^3$	$4pq^3$	$6p^2q^2$	$4p^3q$	$1p^4$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

כאן, כל איבר הוא q כפול האיבר שמעליו ו-p כפול האיבר שמשמאל לאיבר שמעליו.

קל כך להוכיח באינדוקציה את הנוסחה עבור $P_{N_n}(k)$.

<u>אינקרימנטים סטציונרים ובלתי תלויים:</u>

נסתכל עכשיו על ההפרש $N_{n+m}-N_n$ נכנה משתנה מקרי זה **האינקרימנט** (תוספת) של התהליך. משתנה מקרי זה הוא למעשה הסכום $X_{n+1}+X_{n+2}+...+X_{n+m}$ ברור כי הסכום הנ"ל זהה בהתפלגותו לסכום $X_1+X_2+...+X_m$ שזהו בעצם N_m. ולכן $N_{n+m}-N_n \sim Bin(m,p)$.

חלק א: מבוא

אם כך, אנו רואים כי התפלגות האינקרימנט $N_{n+m}-N_n$ אינה תלויה ב n, (היא זהה לדוגמא להתפלגות $N_{l+m}-N_m$ עבור $l \neq n$). תכונה זו של התפלגות האינקרימנטים של התהליך נקראת **אינקרימנטים סטציונרים**. פילוג האינקרימנט הוא סטציונרי (אינו משתנה) בזמן (ב – n). להלן ההגדרה המדויקת.

הגדרה:
תהליך $\{Z_n, n \geq 0\}$ הוא בעל **אינקרימנטים סטציונרים** אם:
$Z_{k+m}-Z_k \sim Z_{l+m}-Z_l$ לכל $k,l,m \in \mathbb{N}$.

נסתכל עכשיו על תכונה נוספת של תהליך ספירה ברנולי:
עבור כל סדרה של זמנים, $n_0 = 0 < n_1 < n_2 < \ldots < n_j$ המשתנים המקריים (האינקרימנטים):

$$N_{n_1}-N_{n_0},$$
$$N_{n_2}-N_{n_1},$$
$$\ldots$$
$$N_{n_j}-N_{n_{j-1}}$$

הינם בלתי תלויים. הרי כל אינקרימנט הוא על פרק זמן זר ולכן כל אינקרימנט הוא סכום של משתנים מקריים ברנוליים שונים ובלתי תלויים זה בזה. תכונה זו של התהליך נקראת תכונת **אינקרימנטים בלתי תלויים**. להלן ההגדרה המפורטת:

הגדרה:
תהליך $\{Z_n, n \geq 0\}$ הוא בעל **אינקרימנטים בלתי תלויים** אם:
עבור כל סדרת זמנים $n_0 = 0 < n_1 < n_2 < \ldots$ מתקיים $Z_{n_i}-Z_{n_{i-1}}$ בלתי תלוי ב $Z_{n_j}-Z_{n_{j-1}}$ עבור כל $i,j \in \mathbb{N}, i \neq j$.

קבלנו אם כך כי האינקרימנטים של תהליך ספירה ברנולי הינם **אינקרימנטים סטציונרים בלתי תלויים**.
תכונה זו היא חשובה ביותר והיא תופיע במספר תהליכים נוספים אשר נחקור, בעיקר בתהליכי פואסון.

באמצעות תכונת האינקרימנטים הבלתי תלויים ניתן לחשב
חישוב הסתברות משותפת של $N_{n_1}, N_{n_2}, \ldots\ldots$:

דוגמא:
מהי $P(N_5=3, N_9=6, N_{13}=7)$?

$$P(N_5=3, N_9=6, N_{13}=7) = P(N_5=3, N_9-N_5=3, N_{13}-N_9=1)$$

ולפי אינקרימנטים בלתי תלויים:

$$= P(N_5=3)P(N_9-N_5=3)P(N_{13}-N_9=1) =$$

ולפי אינקרימנטים סטציונרים:

$$= P(N_5 = 3)P(N_{4+5} - N_5 = 3)P(N_{9+4} - N_9 = 1)$$
$$= P(N_5 = 3)P(N_4 - N_0 = 3)P(N_4 - N_0 = 1)$$
$$= P(N_5 = 3)P(N_4 = 3)P(N_4 = 1)$$
$$= \binom{5}{3} p^3 q^2 \binom{4}{3} p^3 q^1 \binom{4}{1} p^1 q^3$$

ניתן להשתמש גם בתכונות של אינקרימנטים סטציונרים בלתי תלויים לחישוב תוחלות של מכפלות (כזכור תוחלת של מכפלה של שני משתנים מקריים אינה תמיד שווה מכפלת התוחלות. (תנאי מספיק לכך הוא אי-תלות בין המשתנים המקריים).

דוגמא:

נרצה לחשב את $E[N_3N_5]$, לצערנו N_3 אינו בלתי תלוי ב N_5 ולכן **<u>לא</u>** ניתן להסיק כי:

$$E[N_3N_5] = EN_3EN_5 = 3p5p = 15p^2 \text{ !!!}$$

אבל ע"י ייצוג של N_5 כסכום של N_3 והאינקרימנט, נצליח לחשב את תוחלת המכפלה:

$$E[N_3N_5] = E[N_3(N_5 - N_3 + N_3)] = E[(N_3(N_5 - N_3) + N_3{}^2] = EN_3(N_5 - N_3) + EN_3{}^2$$

עכשיו נשתמש באינקרימנטים בלתי תלויים:

$$= EN_3E(N_5 - N_3) + EN_3{}^2$$

ועכשיו נשתמש באינקרימנטים סטציונרים:

$$= EN_3EN_2 + EN_3{}^2$$

מכאן הכול קל, רק צריך להיזכר כי המומנט השני הוא השונות בתוספת המומנט הראשון בריבוע:

$$= 3p2p + (3pq + (3p)^2) = 15p^2 + 3pq$$

<u>תכונות נוספות של תהליך ספירה ברנולי:</u>

נסתכל על הסתברויות מותנות בין ערכים של התהליך בזמנים שונים. נתחיל מדוגמא. נניח כי ידוע כי בזמן 10 ערך התהליך היה 6, מעניין מהי ההסתברות שערך התהליך בזמן 15 יהיה 8. ננסח בעיה זו באמצעות הסתברות מותנית ונשתמש בתכונת אינקרימנטים סטציונרים בלתי תלויים:

$$P(N_{15} = 8 \mid N_{10} = 6) = \frac{P(N_{15} = 8, N_{10} = 6)}{P(N_{10} = 6)} = \frac{P(N_{10} = 6, N_{15} - N_{10} = 2)}{P(N_{10} = 6)} = \frac{P(N_{10} = 6)P(N_{15} - N_{10} = 2)}{P(N_{10} = 6)} =$$
$$P(N_{15} - N_{10} = 2) = P(N_5 = 2)$$

באופן כללי חזרה על אותו חישוב תיתן:

$$P(N_m = l \mid N_n = k) = \begin{cases} P_{m-n}(l-k) & k \le l \\ 0 & l < k \end{cases}$$ עבור $n \le m$.

כעת נהפוך את היחס בין m ל n. נניח כי ידוע כי שבזמן 15 ערך התהליך הוא 8. ז"א ידוע ששמונה מתוך המשתנים המקריים $X_1,...,X_{15}$ הם 1 (והשאר 0). לאור ידיעה זו נתעניין בהסתברות שבזמן 10 ערך התהליך הוא 6. זו בעצם השאלה: מה ההסתברות שמתוך ה -8 הצלחות, 6 נפלו במשתנים $X_1,...,X_{10}$ ושתיים נפלו במשתנים $X_{11},...,X_{15}$:

$$P(N_{10}=6\,|\,N_{15}=8)=\frac{P(N_{15}=8,N_{10}=6)}{P(N_{15}=8)}=\frac{P(N_{10}=6,N_{15}-N_{10}=2)}{P(N_{15}=8)}=\frac{P(N_{10}=6)P(N_{15}-N_{10}=2)}{P(N_{15}=8)}$$

$$\frac{P(N_{10}=6)P(N_{5}=2)}{P(N_{15}=8)}==\frac{\binom{10}{6}p^6q^{10-6}\binom{5}{2}p^2q^{5-2}}{\binom{15}{8}p^8q^{15-8}}=\frac{\binom{10}{6}\binom{5}{2}}{\binom{15}{8}}$$

נראה מוכר? אכן כן, זהו ערך מתוך ההתפלגות ההיפר-גיאומטרית: דגימה של 8 דגימות ללא החזרה מתוך אוכלוסיה בגודל 15, כאשר 10 מתוך ה – 15 הם מסוג מסוים (ו 5 הם מהסוג השני) ומחושבת ההסתברות שבדיוק 6 (מתוך 8) הדגימות יהיו מהסוג המסוים.

ניתן באופן כללי לרשום את $P(N_m=l\,|\,N_n=k)$ עבור $m<n$ כהסתברות מתוך התפלגות היפר-גיאומטרית אבל לא נעשה זאת כאן.

פרק א-5: תהליכי ברנולי – II.

בפרק הקודם הכרנו את התהליך $\{X_n, n \geq 1\}$, תהליך הברנולי .i.i.d ואת התהליך הנבנה ממנו $\{N_n, n \geq 0\}$, תהליך הספירה הבינומי. בפרק זה נגדיר תהליך נוסף הנבנה מ $\{X_n, n \geq 1\}$ והוא $\{T_k, k \geq 0\}$ והוא תהליך זמני ההצלחה (יקרא גם תהליך בינומי שלילי).

נניח כי נתונה ריאליזציה של תהליך ברנולי .i.i.d. להלן דוגמא לתחילת סדרה כזאת:
$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, ...\} = \{1,0,0,1,1,0,1,0,1,....\}$.
נסמן ב T_k את זמן ההצלחה ה- k. אם כך:

$$T_1 = 1,$$
$$T_2 = 4,$$
$$T_3 = 5,$$
$$T_4 = 7,$$
$$T_5 = 9,...$$

בנוסף, ולמען הנוחות בהמשך נסמן $T_0 = 0$.

הערה: זוהי סדרת משתנים מקריים עולה ממש.

לתהליך זה נקרה **תהליך זמני ההצלחה ברנולי**, נסמנו $\{T_k, k \geq 0\}$.

הערה: עבור רוב התהליכים הסטוכסטיים אשר נפגוש בקורס זה, מרחב הפרמטר הוא בעל משמעות של זמן. לעומת זאת, בתהליך זה מרחב הפרמטר אינו מסמל זמן במובן הישיר אלה מסמן את האינדקס של ההצלחה.

ראינו בדוגמא כיצד $\{T_k, k \geq 0\}$ נבנה מ $\{X_n, n \geq 1\}$. כעת נכתוב זאת במפורש:
$T_k = \min\{n \in \mathbb{N} \mid k \leq X_1 + ... + X_n\}$ עבור $1 \leq k$ ו $T_0 = 0$. ז"א עבור k נתון (מספר הצלחה נתון), ערך התהליך הוא מספר הניסיונות הדרוש לצורך קבלת k הצלחות.

לדוגמא, עבור הריאליזציה אשר צוינה לעיל נסתכל על T_3:
$T_3 = \min\{n \in \mathbb{N} \mid 3 \leq X_1 + ... + X_n\}$. עכשיו, הקבוצה $\{n \in \mathbb{N} \mid 3 \leq X_1 + ... + X_n\}$ היא תת קבוצה של המספרים הטבעיים המקיימת את התנאי: $3 \leq X_1 + ... + X_n$ עבור כל איבר n בקבוצה. מהי אם כך הקבוצה הזאת עבור הריאליזציה הנתונה לעיל?
$\{n \in \mathbb{N} \mid 3 \leq X_1 + ... + X_n\} = \{5,6,7,8,9,10,11,12,13,....\}$ והערך המינימאלי של קבוצה זו הוא 5 ולכן
$T_3 = 5$.

מהו חוק ההסתברות של T_k? בהמשך נפתח חוק זה בדרך אלטרנטיבית אבל ראשית ניזכר בסיפורים אשר הובילו לבניית ההתפלגות הבינומית השלילית (בקורס מבוא להסתברות):

חלק א: מבוא

נתונה סדרה של ניסויי ברנולי .i.i.d, כמה ניסיונות דרושים לצורך קבלת k הצלחות? ערך זה הוא הרי בדיוק T_k. ז"א $T_k \sim NB(k,p)$. ניזכר כעת בפונקציית מסת ההסתברות של משתנים מקריים בינומיים שליליים.
נסתכל על המאורע $\{T_k = n\}$, ברור שעבור $n < k$ ההסתברות של מאורע זה היא 0, הרי לא ניתן לקבל k הצלחות ע"י פחות מ-k ניסיונות. אזי נניח כי $k \le n$. מאחורי מאורע זה, מסתתרת סדרת ברנולי .i.i.d המקיימת $X_n = 1$. והערכים של $X_1,...,X_{n-1}$ הינם או 0 או 1 כך ש n-1 מהערכים הינם אחדים. כי הרי הניסוי האחרון (הניסוי ה-n) חייב להיות הצלחה בשביל שהמאורע $\{T_k = n\}$ יתקיים והניסויים שקדמו לו יכולים להיות או הצלחה או כשלון כל עוד ישנם k-1 הצלחות ביניהם. אם כך ישנן $\binom{n-1}{k-1}$ אפשרויות לבחירת הניסויים המצליחים והנכשלים. עבור כל אפשרות כזאת, הסיכוי שבאמת יהיו הצלחות וכישלונות בהתאמה הוא $p^{k-1}(1-p)^{n-1-(k-1)} = p^{k-1}(1-p)^{n-k}$. ואת כל זה עלינו להכפיל בסיכוי שהניסוי האחרון הוא הצלחה (p) ולכן מקבלים:

$$P(T_k = n) = p\left(\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}\right) = \binom{n-1}{k-1}p^k(1-p)^{n-k}.$$

כזכור, משתנה מקרי גיאומטרי הוא מקרה פרטי של בינומי שלילי $Geom(p) = NB(1,p)$ ואם כך אנו יודעים כי $T_1 \sim Geom(p)$.

הקשר בין תהליך ברנולי .i.i.d, ספירה ברנולי ותהליך זמני ההצלחה ברנולי:

עד כה פגשנו שלושה תהליכים סטוכסטיים אשר מתבססים על אותו מרחב הסתברות: $\{X_n, n \ge 1\}$, $\{N_n, n \ge 0\}$, $\{T_k, k \ge 0\}$. ראינו בפרק הקודם כיצד לבנות את $\{N_n, n \ge 0\}$ בהינתן $\{X_n, n \ge 1\}$ וראינו בפרק זה כיצד לבנות את $\{T_k, k \ge 0\}$ בהינתן $\{X_n, n \ge 1\}$.
קל לראות כי הבניות הנ"ל הינן דו-כיווניות, ז"א ניתן לשחזר את תהליך הברנולי .i.i.d ($\{X_n, n \ge 1\}$) מתוך $\{N_n, n \ge 0\}$ או מתוך $\{T_k, k \ge 0\}$.
שחזור $\{X_n, n \ge 1\}$ מתוך $\{N_n, n \ge 0\}$: $X_n = N_n - N_{n-1}$.
שחזור $\{X_n, n \ge 1\}$ מתוך $\{T_k, k \ge 0\}$: $X_n = I_{\{T_k, k\ge 0\}}(n)$. כי הרי אם $n \in \{T_k, k \ge 0\}$ אזי בזמן n ישנה הצלחה ברנולית ולכן n משתייכת לתהליך זמני ההצלחה ופו' האינדיקאטור תיתן 1. ואם $n \notin \{T_k, k \ge 0\}$ אז בזמן n אין אף הצלחה (לא הראשונה, ולא השנייה ולא השלישית,....) ופו' האינדיקאטור תיתן 0.

אם כך בהינתן כל אחת משלושת הסדרות, ניתן לקבל את השתיים האחרות.

ניתן גם לקשר ישירות בין מאורעות המתוארים על פי התהליך $\{N_n, n \ge 0\}$ למאורעות המתוארים על פי התהליך $\{T_k, k \ge 0\}$, להלן הקשר:
נניח כי מתקיים $k \le N_n$, ז"א עד זמן n היו לפחות k הצלחות ברנולי (מאורע זה כמובן יכול להתקיים אך ורק אם $k \le n$). אם כך יתקיים כי ההצלחה ה-k הייתה לכל היותר בזמן n: $T_k \le n$. קבלנו אם כך את יחס הגרירה הבאה, בין מאורעות: $\{k \le N_n\} \Rightarrow \{T_k \le n\}$.
באופן דומה אם מתקיים $T_k \le n$, ז"א ההצלחה ה-k הייתה לכל היותר בזמן n, אזי יתקיים כי בזמן n היו לפחות k הצלחות: $k \le N_n$. אם כך קבלנו: $\{T_k \le n\} \Rightarrow \{k \le N_n\}$.
יש כאן אם כך גרירה דו-כיוונית ולכן שקילות בין המאורעות: $\{T_k \le n\} \Leftrightarrow \{k \le N_n\}$.

קשר זה בין מאורעות מקנה להתפלגות הבינומית השלילית את שמה:

$$F_{T_k}(n)=P(T_k\le n)=P(k\le N_n)=\sum_{i=k}^{n}P(N_n=i)=\sum_{i=k}^{n}\binom{n}{i}p^i q^{n-i}=\overline{F}_{N_n}(k+1)$$

אם כך פו' ההתפלגות המצטברת של משתנה מקרי בינומי שלילי עם פרמטרים p,k בנקודה n היא פו' השרידות של משתנה מקרי בינומי עם פרמטרים n,p בנקודה k+1 ומכאן ה"שליליות".

קשר נוסף בין מאורעות הוא הקשר הבא: $\{T_k=n\}=\{N_{n-1}=k-1,X_n=1\}$. הרי במידה וזמן ההצלחה ה – k היה בדיוק n אזי עד זמן n-1 היו k-1 הצלחות ובזמן n הייתה הצלחה וכנ"ל בכוון ההפוך.
ברור כי X_n בלתי תלוי ב N_{n-1} ולכן:

$$P(T_k=n)=P(N_{n-1}=k-1,X_n=1)=P(N_{n-1}=k-1)P(X_n=1)=$$
$$=\binom{n-1}{k-1}p^{k-1}q^{n-1-(k-1)}p=\binom{n-1}{k-1}p^k q^{n-k}$$

כפי שקבלנו קודם בדרך הישירה.

<u>אינקרימנטים סטציונרים ובלתי תלויים של תהליך זמני ההצלחה ברנולי:</u>

כעת נרצה לראות שגם התהליך $\{T_k,k\ge 0\}$ הוא בעל אינקרימנטים סטציונרים ובלתי תלויים. במקום להוכיח זאת באופן פרטני נראה תוצאה קצת יותר כללית:

<u>טענה:</u>
תהי $\{Y_n,0\le n\}$ סדרת משתנים מקריים i.i.d. ו- $\{Z_n,0\le n\}$ תהליך סטוכסטי כך ש:

$$Z_n=\begin{cases}0 & n=0\\ Y_1+...+Y_n & n=1,2,..\end{cases}$$

אזי $\{Z_n,0\le n\}$ הוא בעל אינקרימנטים סטציונרים ובלתי תלויים.
הוכחת הטענה פשוטה וזהה להוכחת אינקרימנטים סטציונרים ובלתי תלויים של התהליך $\{N_n,0\le n\}$ כפי שהוכחנו בפרק הקודם.

אם כך, כל תהליך סטוכסטי אשר ניתן לרישום כסכום של משתנים .i.i.d. (תהליך כזה נקרא **הילוך אקראי כללי**) הוא בעל אינקרימנטים סטציונרים ובלתי תלויים. בנוסף דרוש כי $Z_1=Y_1$ ולכן פילוג אינקרימנט על צעד בגודל 1 (Y_i) צריך להיות שווה לערך התהליך בזמן 1 (Z_1).

בכדי להראות כי התהליך $\{T_k,0\le k\}$ הוא בעל אינקרימנטים סטציונרים בלתי תלויים, עלינו להראות כי T_k ניתן לכתיבה כסכום של k משתנים מקריים .i.i.d. קל לראות כי פילוג כל משתנה כזה יהיה Geom(p).

ברור גם כי סכום של k משתנים מקריים בעלי התפלגות Geom(p) מתפלג NB(k,p). בנוסף סכום של שתי משתנים מקריים בעלי התפלגות $NB(k_1,p)$ והתפלגות $NB(k_2,p)$ מתפלגים $NB(k_1+k_2,p)$.

אז כעת אנו יודעים כי $T_{k+1}-T_k\sim Geom(p)$ ועכשיו קל לחשב את התוחלת והשונות של התהליך $\{T_k,0\le k\}$. (שוב הכוונה לתוחלת או שונות של תהליך היא הפו' הדטרמיניסטית המראה את התוחלת והשונות של ההתפלגויות השוליות של התהליך).

חלק א: מבוא

$$ET_k = E[T_0 + (T_1 - T_0) + ... + (T_k - T_{k-1})] = ET_0 + E[T_1 - T_0] + ... + E[T_k - T_{k-1}] =$$
$$= 0 + kE[T_1] = 0 + k\frac{1}{p} = \frac{k}{p}$$

לאחר שניזכר כי השונות של משתנה מקרי גיאומטרי היא $\frac{q}{p^2}$ באופן דומה נקבל כי $Var(T_k) = k\frac{q}{p^2}$

דוגמא א-22:
נניח כי עבור סדרת ניסויי ברנולי אנו משלמים מחיר c עבור "כל הצלחה". (לדוגמא: אנו רוכבים על אופניים כל יום ובכל פעם שיש פנצ'ר החלפת פנימית עולה כ 20 ₪ ואנו קוראים לפנצ'ר הצלחה). אם כך, אנו יודעים כי לאחר n ניסיונות תוחלת העלות היא $EcN_n = cEN_n = cnp$.
מדד מעניין יותר לעלות היא העלות **הערך הנוכחי** של העלויות אשר נגררות מהצלחותינו. כלל ידוע בכלכלה הוא ששקל שיש לנו היום לרוב שווה מחר קצת יותר (צריך להכפיל ב 1 ועוד הריבית שאנו יכולים להרוויח על השקל) – ז"א עדיף לנו שהשקל יהיה בידנו היום מאשר מחר. באופן דומה ניתן לומר כי $0 < \alpha < 1$ שקלים אשר בידינו היום שווים לשקל 1 מחר (כאשר $\alpha = 1/(1+r)$). אם כך α הוא מקדם ההיוון. ולכן מחיר של c אשר נידרש לשלם מחר הוא בעצם אינו כל כך גרוע היום הוא בסך הכול αc היום.

ואם נידרש לשלם c בעוד 150 ימים אזי במונחי היום המחיר הוא $\alpha^{150}c$. כי אם היום יש לנו $\alpha^{150}c$ שקלים ואנו שומרים את אלו בבנק ומכפילים את כספנו בכל יום ב (1+r) אזי כעבור 150 ימים בידינו:
$$(1+r)^{150}\alpha^{150}c = (1+r)^{150}\left(\frac{1}{1+r}\right)^{150}c = c.$$

אם כך, תוחלת הערך הנוכחי מהעלויות הנגררות מ"ההצלחות" היא $E\sum_{k=1}^{\infty} c\alpha^{T_k}$ או $cE\sum_{k=1}^{\infty}\alpha^{T_k}$ או $c\sum_{k=1}^{\infty}E\alpha^{T_k}$.

נבחין כי $E\alpha^{T_k}$ היא הפו' יוצרת ההסתברות של משתנה מקרי $NB(k,p)$, בנקודה α.
אם כך נחשב את $G_{T_k}(\alpha) = E\alpha^{T_k}$.
$$G_{T_k}(\alpha) = E\alpha^{T_k} = E\alpha^{T_0+(T_1-T_0)+(T_2-T_1)+...+(T_k-T_{k-1})} = E\alpha^{T_0}\alpha^{T_1-T_0}\cdot ...\cdot \alpha^{T_k-T_{k-1}} =$$
$$= E\alpha^{T_0}E\alpha^{T_1-T_0}\cdot ...\cdot E\alpha^{T_k-T_{k-1}} = E1\underbrace{E\alpha^{T_1}E\alpha^{T_1}\cdot ...\cdot E\alpha^{T_1}}_{k} = (E\alpha^{T_1})^k$$
עכשיו ידוע לנו כי $T_1 \sim Geom(p)$. ואם כך:
$$G_{T_1}(\alpha) = E\alpha^{T_1} = \sum_{i=1}^{\infty}\alpha^i q^{i-1}p = p\alpha\sum_{i=1}^{\infty}(\alpha q)^{i-1} = p\alpha\sum_{i=0}^{\infty}(\alpha q)^i = p\alpha\frac{1}{1-\alpha q}$$

וזאת עבור $\alpha q < 1$ או $\alpha < \frac{1}{1-p}$.

ולכן $G_{T_k}(\alpha) = \left(\frac{p\alpha}{1-\alpha q}\right)^k$

נותר כעת לחשב את $c\sum_{k=1}^{\infty} E\alpha^{T_k}$:

$$c\sum_{k=1}^{\infty} E\alpha^{T_k} = c\sum_{k=1}^{\infty} G_{T_k}(\alpha) = c\sum_{k=1}^{\infty}\left(\frac{p\alpha}{1-\alpha q}\right)^k = c\frac{\frac{p\alpha}{1-\alpha q}}{1-\frac{p\alpha}{1-\alpha q}} = c\frac{p\alpha}{1-\alpha q}\frac{1}{\frac{1-\alpha q-p\alpha}{1-\alpha q}} = c\frac{p\alpha}{1-\alpha(p+q)} = cp\frac{\alpha}{1-\alpha}$$

נבחין כי כאשר $\alpha = 1/(1+r)$ אז $\frac{1}{r} = \frac{\alpha}{1-\alpha}$

ואז תוחלת העלות המהוונת היא: $cp\frac{1}{r}$.