תוצאות שוות סיכוי וחוקי הסתברות בסיסיים

הגדרות:

Ω - **מרחב המדגם**: אוסף כל התוצאות האפשריות של הניסוי.

מרחב מדגם בדיד: יש בו מספר סופי או בן מניה של תוצאות.

מרחב מדגם שווה הסתברות: מרחב הסתברות (Ω,P) שבו לכל מאורע פשוט יש אותה הסתברות ו- Ω הוא סופי. דוגמא: הטלת מטבע הוגן. וגם מתקיים: $\forall\, \omega\in\Omega$, הסתברות שווה להתרחש והיא ניתנת ע"י: $P(\{\omega\})=\frac{1}{|\Omega|}$ ועבור מאורע A - $\boxed{P(A)=P\left(\bigcup_{\omega\in A}\{\omega\}\right)=\frac{|A|}{|\Omega|}}$

A-**מאורע**: תת-קבוצה של מרחב המדגם (מסומן באות גדולה).

מאורע פשוט: מאורע המכיל רק תוצאה אחת, לרוב נסמנו- ω. דוגמא: $C=\{5\}$ בהטלת קובייה.

P-**מידת הסתברות**: פונקציה מאוסף המאורעות של Ω ל- $[0,1]$

P צריכה לקיים 3 תנאים:

1. P אי-שלילית, כלומר $P(A)\ge 0$ לכל מאורע A.
2. אם $A_1,\cdots,A_n$ הם חלוקה של A (כלומר A_i זרים ואיחודים הוא A), אזי:
$$P(A)=P(A_1)+P(A_2)+\cdots+P(A_n)$$
3. $P(\Omega)=1$.

תכונות נוספות של מידת ההסתברות:

- $P(A)\le 1$
- $P(A^c)=1-P(A)$
- $P(\varnothing)=0$
- אם $A\subset B$, אז: .1 $P(A)\le P(B)$ .2 $P(B\setminus A)=P(B\cap A^c)=P(B)-P(A)$.

כלל הכלה והפרדה: $\boxed{P(A\cup B)=P(A)+P(B)-P(A\cap B)}$

נוסחה כללית: $\boxed{P(A_1\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{i<j}P(A_i\cap A_j)+\cdots(-1)^{n+1}\cdot P(A_1\cap\cdots\cap A_n)}$

חוק Boole: $\boxed{P\left(\bigcup_{i=1}^{n}A_i\right)\le\sum_{i=1}^{n}P(A_i)}$ $\boxed{P(A\cup B)\le P(A)+P(B)}$

הסתברות מותנית ואי תלות

הסתברות מתנית:

יהיו A, B מאורעות, כך ש- $P(B) > 0$. **ההסתברות המותנית** של A בהינתן B מסומנת ע"י:

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}}$$

כלל הכפל: $\boxed{P(A \cap B) = P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)}$

כלל ההסתברות השלמה: תהי $B_1, B_2, \cdots, B_n$ חלוקה זרה וממצה של Ω. כלומר:

1. $\bigcup_{i=1}^{n} B_i = \Omega$ 2. $\forall i \neq j: \ B_i \cap B_j = \varnothing$ 3. $\forall i: \ B_i \neq \varnothing$

יהיה A מאורע. אז: $\boxed{P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)}$

כלל Beyes: יהיו $B_1, B_2, \cdots, B_n$ חלוקה של Ω ויהא B_j מאורע כלשהו מתוך חלוקה זו. אם A מאורע עבורו מתקיים: $P(A) > 0$ אזי מתקיים:

$$\boxed{P(B_j \mid A) = \frac{P(A \mid B_j).P(B_j)}{P(A)} = \frac{P(A \mid B_j) \cdot P(B_j)}{\sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)}}$$ (פירוק המכנה לפי הסתברות שלמה).

אי תלות:

נאמר שמאורעות A, B בלתי תלויים אם הסיכוי להתרחשות A אינו משופע מהידיעה אם מאורע B התרחש או לא התרחש ולהפך. A, B בלתי תלויים אם מתקיים: $\boxed{P(A \cap B) = P(A) \cdot P(B)}$

הסתברות מותנית עבור מאורעות בלתי תלויים:

$$\boxed{\begin{aligned} P(A) &= P(A \mid B) = p(A \mid B^c) \\ P(B) &= P(B \mid A) = P(B \mid A^c) \end{aligned}}$$

הערות:

- $\varnothing$ ו- Ω תמיד בלתי תלויים עם מאורע אחר.
- מאורעות זרים שאינם ריקים הם תלויים- $P(A \cap B) = P(\varnothing) = 0$ אבל: $P(A) \cdot P(B) > 0$ ולכן לא מתקיימת האי-תלות.
- מאורעות מוכלים שאינם ריקים ואינם שווים ל- Ω - תלויים.
- מאורעות בלתי תלויים שאינם ריקים- נחתכים

חוקי דה-מורגן:

- $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$

- $\overline{A \cup B} = \overline{A} \cap \overline{B}$

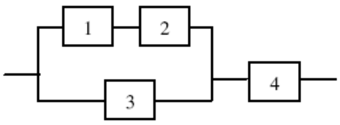
אמינות מערכת:

מערכת מורכבת ממספר רכיבים. תהי p_i ההסתברות שהרכיב i תקין. נניח כי כל הרכיבים תקינים או תקולים באופן בלתי תלוי באחרים. נסמן- a, האמינות של המערכת, להיות ההסתברות שהמערכת תקינה.

1. **מערכת טורית:** n רכיבים מחוברים בטור. דרוש **שכל הרכיבים** יהיו תקינים כדי שהמערכת תהיה תקינה, $\boxed{a = p_1 \cdot p_2 \cdot \cdots \cdot p_n}$ (ההסתברות לחיתוך- מכפלת ההסתברויות שכן הרכיבים ב"ת).

2. **מערכת מקבילית:** n רכיבים מחוברים במקביל. מספיק **שאחד מן הרכיבים** יהיה תקין כדי שהמערכת תהיה תקינה, במילים אחרות- דרוש שכל הרכיבים יהיה תקולים כדי שהמערכת תהיה תקולה. $\boxed{a = 1 - (1 - p_1)(1 - p_2)\cdots(1 - p_n)}$

3. **מערכת מעורבת:** עבור המערכת הנתונה-
$a = \left(1 - (1 - p_3)(1 - p_1 p_2)\right) \cdot p_4$



ניסויים בלתי תלויים- התפלגויות

ניסוי ברנולי:

ניסוי בעל שתי תוצאות אפשריות: הצלחה- S , $\boxed{P(S) = p}$ וכישלון- F , $\boxed{P(F) = 1 - p = q}$

התפלגות בינומית: Bin(n,P)

נתון מרחב המדגם $\Omega = \{0,1,\cdots,n\}$ ועליו מוגדת מידת ההסתברות: $\boxed{P(k) = \binom{n}{k} p^k q^{n-k}}$

מודל בינומי של ניסויים: מבצעים סדרה של n ניסויי ברנולי ב"ת, בעלי הסתברות להצלחה- p .

ההסתברות לקבל בדיוק k הצלחות ב- n הניסויים היא $\binom{n}{k} p^k q^{n-k}$.

התפלגות גיאומטרית: Geo(P)

נתון מרחב המדגם בן-מנייה $\Omega = \{1,2,\cdots\}$ ועליו מוגדת מידת ההסתברות: $\boxed{P(k) = q^{k-1} p}$

מודל גיאומטרי של ניסויים: מבצעים סדרה של ניסויי ברנולי ב"ת על להצלחה ראשונה. ההסתברות שההצלחה הראשונה תתרחש בניסיון ה- k , היא $q^{k-1} p$.

תכונת חוסר הזיכרון למודל גיאומטרי: אם ידוע כי לא הייתה הצלחה עד לניסיון ה- m , ההסתברות שההצלחה הראשונה תתרחש בניסיון ה $m + k$ היא עדיין" $q^{k-1} p$.

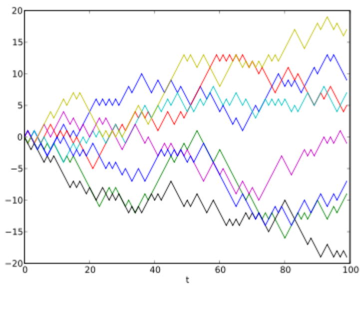
התפלגות בינומית שלילית: NB(m,P)

נתון מרחב המדגם בן-מנייה $\Omega = \{m, m+1,\cdots\}$ ועליו מוגדרת מידת ההסתברות (עבור m מסוים):

$$\boxed{P(k) = \binom{k-1}{m-1} p^m q^{k-m} \quad \forall k \in \Omega}$$

מודל בינומי שלילי: מבצעים סדרה של ניסויי ברנולי ב"ת עד להצלחה ה- m -ית. ההסתברות לכך שהצלחה זו תתקבל בדיוק בניסיון ה- k היא $\binom{k-1}{m-1} p^m q^{k-m}$.

הילוך אקראי: מסלול המכיל שלבים עוקבים וקשורים שבו כל בחירת כל שלב היא אקראית ואינה מושפעת מהשלבים הקודמים.



דגימות מקריות והתפלגות פואסונית

דגימות מקריות:

1. פיזור של r עצמים שונים ל- n תאים: $\boxed{|\Omega| = n^r}$

שקול לדגימת r עצמים מתוך n עם חזרות, עם חשיבות לסדר.

2. פיזור של r עצמים זהים ל- n תאים: $\boxed{|\Omega| = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}}$

שקול לדגימת r עצמים זהים מתוך- n עם חזרות ובלי חשיבות לסדר.

3. פיזור של r עצמים <u>זהים</u> ל- n תאים כך שבל תא יש לכל היותר עצם אחד

$\boxed{r \le n \quad |\Omega| = \binom{n}{r}}$ שקול לדגימת r עצמים מתוך n בלי חזרות, בלי חשיבות לסדר.

התפלגות פואסון:

נתון מרחב מדגם $\Omega = \{0,1,2,\cdots\}$ ועליו מוגדרת מידת ההסתברות $\boxed{P(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad \lambda > 0}$

<u>מודל פואסוני:</u> מספר האירועים ביחידת זמן נתונה, אם ידוע כי האירועים מתרחשים בקצב (מספר אירועים ליחידת זמן) ממוצע קבוע $\lambda > 0$ ובאופן בלתי תלוי בפרק הזמן מאז האירוע האחרון. במקום יחידת זמן נתונה, ניתן לחשוב גם על יחידת מרחק, שטח או נפח נתונה.

התפלגות פואסונית מתקבלת מהתפלגות בינומית $\frac{e^{-\lambda}\lambda^k}{k!} = \lim_{\substack{n \to \infty \\ np=\lambda}} \binom{n}{k} p^n (1-p)^{n-k}$

(כלומר, מבצעים סדרה ארוכה מאוד של ניסויי ברנולי ב"ת כאשר ההצלחה היא בהסתברות מאוד קטנה כך ש- $np = \lambda$).

<u>קירוב פואסוני להתפלגות בינומית:</u> הקשר בין ההתפלגויות מאפשר להשתמש בהתפלגות הפואסונית כקירוב להתפלגות הבינומית (שבמקרים רבים החישוב שלה מסורבל).

תהי התפלגות בינומית עם פרמטרים n ו- p. אם n גדול, p קטן ו- $0 < np < 10$ אז ניתן לקרב את מידת ההתפלגות הבינומית לזו של התפלגות פואסונית עם פרמטר: $\lambda = np$.

התפלגות היפר גיאומטרית:

נתון מרחב מדגם $\Omega = \{0,1,2,\cdots,\min\{D,n\}\}$ ועליו מוגדרת מידת ההסתברות:

$$P(k) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}} \qquad n < N,\ D < N,\quad n,D,N \in \mathbb{N}$$

מודל היפר-גיאומטרי: בתוך כד יש N כדורים שמתוכם D שחורים והשאר לבנים. מוציאים n כדורים באקראי (ללא החזרה). מספר הכדורים השחורים מבין הכדורים שהוצאו מפולג היפר-גיאומטרית עם פרמטרים N, D, n.

סכום סדרה הנדסית אינסופית: $S_\infty = a_1 \frac{q^n - 1}{q-1} = \frac{a_1}{1-q}$

משתנה מקרי בדיד והתפלגותו

הגדרות:

משתנה מקרי: מ"מ X הוא פונקציה ממרחב המדגם Ω אל הישר הממשי R. $X:\Omega\to R$.

סימון מאורע: את המאורע $\{\omega\in\Omega : X(\omega)=x\}$ נסמן- $\{X=x\}$, כלומר: אם יוצאת התוצאה ω, אז X מקבל את הערך $x=X(\omega)$.

מ"מ בדיד: מ"מ שהטווח שלו סופי או בן מניה- כאשר R_X (אוסף כל הערכים שהמ"מ יכול לקבל) מהווה קבוצה סופית או בת מנייה.

הערה: אם לכל קבוצה חלקית B בטווח של משתנה מקרי- X מגדירים מאורע: $P(x\in B)$ אז הסתברותו: $\boxed{P(x\in B)=\sum_{x\in B}P(X=x)}$ (מאורעות זרים וניתן לסכום הסתברויות).

פונקצית הסתברות של מ"מ: פונקצית הסתברות של מ"מ X היא פונקציה $P(X=x)$ או $p_X(x)$. כאשר $P(X=x)$ זאת ההסתברות של מאורע $\{X=x\}$. ותמיד מתקיים:

1. $\forall x,\ 0\le P(X=x)\le 1$
2. $\sum_{x\in R}P(X=x)=1$
3. $\forall A\subset R,\ P(X\in A)=\sum_{x\in A}P(X=x)$

הערה: פונקצית הסתברות אינה אקראית אלא מוגדרת באופן חד-משמעי עבור כל ערך x שמ"מ X יכול לקבל. על-מנת לקבוע את ההתפלגות של מ"מ, מספיק למצוא את ההסתברויות עבור כל ערכי המשתנה.

פונקציה של משתנה מקרי: אם X הוא מ"מ ו $g:R\to R$ פונקציה, אז ניתן להגדיר מ"מ חדש: $Y=g(X)$ המתקבלת ע"י הפעלת הפונקציה g על ערך של X.

$$\boxed{P(Y=y)=\sum_{x\in R_X : g(x)=y}P(X=x)}$$

דוגמאות לפונקציות התפלגות ידועות:

ברנולי: $p_X(x) = \begin{cases} p & x=1 \\ q & x=0 \\ 0 & else \end{cases}$ **סימון:** $\boxed{X \sim Ber(p)}$

בינומית: $p_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & x = 0,1,\cdots,n \\ 0 & else \end{cases}$ **סימון:** $\boxed{X \sim Bin(n,p)}$

היפר-גיאומטרית: $p_X(x) = \begin{cases} \dfrac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} & x = 0,1,\cdots,\min(D,n) \\ 0 & else \end{cases}$ **סימון:** $\boxed{X \sim HG(N,D,n)}$

גיאומטרית: $p_X(x) = \begin{cases} q^{x-1}p & x = 1,2,3,\cdots \\ 0 & else \end{cases}$ **סימון:** $\boxed{X \sim Geo(p)}$

אחידה: $P(X=x) = \begin{cases} \dfrac{k}{N} & x \in \{1,2,\cdots,N\} \\ 0 & otherwise \end{cases}$ **סימון:** $\boxed{X \sim Uni(k,N)}$

פואסונית: $p_X(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & x = 0,1,2,\cdots \\ 0 & else \end{cases}$ **סימון:** $\boxed{X \sim Pois(\lambda)}$

קירוב פואסוני לבינומי: $P(X=x) = \binom{n}{k} p^k (1-p)^{n-k} \cong \dfrac{e^{-\lambda}\lambda^k}{k!}$, $n \to \infty,\ np = \lambda$

מספר משתנים מקריים- שולית, מותנית ואי-תלות

התפלגות משותפת של מ"מ (בדידים):

הגדרה: בהינתן שני משתנים מקריים X ו- Y המוגדרים על אותו מרחב מדגם, ההסתברות המשותפת של הזוג הסדור (X,Y) תסומן ע"י: $P(X=x,Y=y)=P_{XY}(x,y)$. שזהו חיתוך המאורעות.

- $0 \le P_{XY}(x,y) \le 1$
- $\sum_{(x,y)} P_{XY}(x,y)=1$

התפלגות שולית: (זהו למעשה איחוד של מאורעות זרים- $\{X=x\}=\bigcup_{y\in R_Y}\{X=x,Y=y\}$).

פונקצית ההסתברות של X: $\boxed{P_X(x)=\sum_y P_{XY}(x,y)}$

פונקצית ההסתברות של Y: $\boxed{P_Y(y)=\sum_x P_{XY}(x,y)}$

הערה: מהתפלגות משותפת ניתן תמיד למצוא את ההתפלגות השולית ע"י סכימה או אינטגרציה, אך מההתפלגות השולית בד"כ לא ניתן לדעת את ההתפלגות המשותפת.

שווי התפלגות: מ"מ X ו- Y שווי התפלגות אם לכל a ממשי: $P_X(a)=P_Y(a)$.

שוויון: מ"מ X ו- Y שווים אם $P(X=Y)=1 \Rightarrow \sum_a P_{XY}(a,a)=1$.

פונקצית הסתברות מותנית:

$$\boxed{P_{Y|X}(y\,|\,x)=P(Y=y\,|\,X=x)=\frac{P(Y=y,X=x)}{P(X=x)},\quad P(X=x)>0}$$

כלל הכפל: $\boxed{P_{XY}(x,y)=P(Y=y\,|\,X=x)\cdot P(X=x)}$

נוסחת ההסתברות השלמה: $\boxed{P_X(x)=\sum_y P_{X|Y}(x\,|\,y)\cdot P_Y(y)}$

נוסחת בייס: $\boxed{P_{Y|X}(y\,|\,x)=\frac{P_{X|Y}(x\,|\,y)\cdot P_Y(y)}{P_X(x)}}$.

נוסחת החלפת משתנים: יהי $Z=f(X,Y)$ אז Z מ"מ ומתקיים:

$$\boxed{P_Z(z)=P(Z=z)=\sum_{(x,y):f(x,y)=z} P(X=x,Y=y)}$$

תלות: מ"מ בדידים X ו- Y הם בלתי תלויים אם: $\boxed{\forall x, y : P_{XY}(x, y) = P_X(x) \cdot P_Y(y)}$

- ההתפלגות המותנת של Y בהינתן $X = x$ אינה תלויה ב- x.
- ההתפלגות המותנת של X בהינתן $Y = y$ אינה תלויה ב- y.

וקטורים מקריים (מספר רב של משתנים מקריים):

הגדרה: יהיו $(X_1, \cdots, X_n)$ משתנים מקריים בדידים. פונקצית ההסתברות המשותפת:

$$P_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n) = P(X_1 - x_1, X_2 = x_2, \cdots, X_n = x_n)$$

שולית 1-מימדית: $P_{x_j}(x_j) = \sum_{x_1} \cdots \sum_{x_{j-1}} \sum_{x_{j+1}} \cdots \sum_{x_n} P_{x_1 \cdots x_n}(x_1, \cdots, x_n)$

שולית k-מימדית: $P_{x_1 x_2 \cdots x_k}(x_1 \cdots x_k) = \sum_{x_{k+1}} \cdots \sum_{x_n} P_{x_1 \cdots x_n}(x_1, \cdots, x_n)$

הערה: k המשתנים לא חייבים להיות בסדר עוקב כמו בנוסחה.

אי תלות: n מ"מ הם ב"ת אם"ם: $P(X_1 = x_1, \cdots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n)$

הערה: עבור n מ"מ בלתי תלויים, אם נחלקם לקבוצות זרות ונפעיל על כל קבוצה בנפרד פונקציה אחרת שתיתן ערך מסוים:

$$\begin{aligned} y_1 &= f_1(X_1, \cdots, X_{n_1}) \\ y_2 &= f_2(X_{n_1+1}, \cdots, X_{n_2}) \\ &\vdots \\ y_m &= f_m(X_{n_m+1}, \cdots, X_n) \end{aligned}$$

אז $(y_1, \cdots, y_m)$ הם ב"ת.

התפלגות מולטינומית: מבצעים סדרה של n ניסויים ב"ת, בכל ניסוי k תוצאות אפשריות עם הסתברויות: $p_1, \cdots, p_k$, $\sum_{i=1}^{k} p_i = 1$.

נסמן ב- X_1 - מספר הפעמים שהתקבלה תוצאה מסוג-1

$\vdots$

X_k - מספר הפעמים שהתקבלה תוצאה מסוג- k

פונקצית ההתפלגות המשותפת: $X \sim Multi(n, p_1, p_2, \cdots, p_k)$

$$P_{X_1 X_2 \cdots X_k}(x_1, x_2, \cdots, x_k) = \begin{cases} \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}, & \sum_{i=1}^{n} x_i = n \\ 0 & otherwise \end{cases}$$

סטטיסטי הסדר:

הגדרה: נתונים - $X_1, \cdots, X_n$. סטטיסטי הסדר שלהם: $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$.

(כלומר: משתנים המקבלים את הראשון בגודל, השני בגודל וכן הלאה...)

הערות:

- אם הנתונים הם אקראיים אז $X_{(1)} \cdots X_{(n)}$ הם n מ"מ.
- אפילו אם $X_1, \cdots, X_n$ הם ב"ת, סטטיסטי הסדר $X_{(1)} \cdots X_{(n)}$ הם תלויים.

תוחלת ואינדיקטורים

תוחלת:

הגדרה: תוחלת של מ"מ בדיד X: $\boxed{E(X)=\sum_{x\in R_X} xP(X=x)}$.

הערות:

- אם הטווח של X סופי אז הסכום של הטור תמיד מוגדר. אחרת הוא יהיה מוגדר רק אם הטור מתכנס בהחלט: $\sum_x |x|P(X=x)<\infty$.
- ניתן להכליל ולדבר על תוחלת $\pm\infty$ אם הטור מתבדר ו- X מקבל ערכים בסימן קבוע.

תכונות:

1. **כלל הסכום:** לכל X,Y מ"מ $E(X+Y)=EX+EY$
2. אם מ"מ X מקבל ערכים בין a ל- b אז $a\le EX\le b$
3. **ליניאריות:** תוחלת היא אופרטור ליניארי: $E(aX+b)=aE(X)+b$

חוק המספרים הגדולים: עבור X_i ב"ת ו- $M=EX_i$ אז $\frac{1}{n}\sum_{i=1}^{n}X_i\to M$

נוסחת הזנב: עבור מ"מ המקבל ערכים **שלמים אי שליליים** $\boxed{E(X)=\sum_{x=1}^{\infty}P(X\ge x)}$

*מומלץ להשתמש בה כאשר הביטוי $P(X\ge x)$ קל יותר לחישוב מאשר $P(X=x)$.

כלל הכפל: עבור X,Y מ"מ בלתי תלויים $\boxed{E(XY)=EX\cdot EY}$ - ההפך אינו נכון.

זוג מ"מ שמקיימים $EXY=EX\cdot EY$ נקראים בלתי מתואמים (אך לא ב"ת!).

עבור כל X,Y מ"מ: $\boxed{E(XY)=\sum_x\sum_y xyP(X=x,Y=y)}$

משפט:

1. בלתי תלויים $\Leftarrow$ בלתי מתואמים.
2. בלתי מתואמים $\not\Leftarrow$ בלתי תלויים.
3. ב"ת $\leftrightarrow$ $Eh(X)\cdot g(Y)=E(h(X))\cdot E(g(Y))$ לכל h,g רציפות וחסומות.

אינדיקטורים:

הגדרה: I_A אינדיקטור של מאורע A הוא מ"מ המוגדר ע"י: $I_A=\begin{cases}1 & A\ happened\\ 0 & otherwise\end{cases}$

תוחלת של אינדיקטור: $\boxed{E(I_A)=P(A)}$

אם יש אוסף מאורעות A_i, $1\le i\le n$, כאשר X סופר כמה מתוכם קרו $X=\sum I_{A_i}$ אז:

$$\boxed{E(X)=E\left(\sum I_{A_i}\right)=\sum E\left(I_{A_i}\right)=\sum P(A_i)}$$

תוחלת ידועה של התפלגויות:

בינומי: $X\sim Bin(n,p)$ $\boxed{E(X)=np}$

בינומי שלילי: $X\sim NB(m,p)$ $\boxed{E(X)=\frac{m}{p}}$

גיאומטרי: $X\sim Geo(p)$ $\boxed{E(X)=\frac{1}{p}}$

פואסוני: $X\sim Pois(\lambda)$ $\boxed{E(X)=\lambda}$

אי-שוויונים ידועים:

חציון: מספר m המקיים: $P(X\le m)\ge\frac{1}{2},\ P(X\ge m)\ge\frac{1}{2}$.

שכיח: המספר שההסתברות לקבל אותו היא הגדולה ביותר.

הערה: תוחלת, חציון ושכיח יכולים להיות שונים.

אי-שוויון מרקוב: יהיה X מ"מ שמקבל ערכים לא שליליים, אזי לכל $a>0$ $\boxed{P(X\ge a)\le\frac{EX}{a}}$

אי-שוויון בול (מקרה פרטי של מרקוב): $\boxed{P(X\ge 1)\le EX}$

מומנט: התוחלת נקראת גם המומנט הראשון של X. המומנט הה-י הוא: $\boxed{m_n=E\left(X^n\right)}$

דוגמא: $X \sim Uni[1,\cdots,n]$ מ"מ בדיד. אז: $P(X=x)=\frac{1}{n}, \quad x=1,2,\cdots,n$

$$m_1 = EX = \sum \frac{1}{n}x = \frac{1}{n}\cdot\frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$m_2 = EX^2 = \sum \frac{1}{n}x^2 = \frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

אי-שוויון ינסן: אם f פונקציה קמורה אזי: $\boxed{E(f(x)) \geq f(E(x))}$

תוחלת של פונקציה של מ"מ:

במידה ו- $g(X)$ היא פונקציה ניתן לחשב את התוחלת שלה ללא שימוש בהגדרה באופן הבא:

$$\boxed{E(g(X)) = \sum_{x\in R_X} g(x)P(X=x)}$$

ולמספר מ"מ: $\boxed{Eg(X,Y) = \sum_x\sum_y g(x,y)P(X=x,Y=y)}$

חיזוי:

מחפשים ערך b כך ש- $E(X-b)^2$ יהיה מינימאלי. אז $\boxed{b=EX}$

מחפשים ערך b כך ש- $E|X-b|$ יהיה מינימאלי. אז <u>החציון</u> הוא החזאי במקרה זה.

שונות וסטיית תקן

הגדרה: $Var(X)=\min E(X-a)^2$ כלומר אם $EX^2<\infty$ אז: $\boxed{Var(X)=E(X-EX)^2}$

הנוסחה השימושית לחישוב: $\boxed{Var(X)=EX^2-(EX)^2}$

סטיית תקן: $\boxed{\sigma=SD(X)=\sqrt{Var(x)}}$

אלה המדדים המתארים עד כמה ההתפלגות מפוזרת סביב התוחלת. ככל שההתפלגות מרוכזת יותר סביב התוחלת כך קטנה השונות/סטיית התקן. השונות נמדדת ביחידות בריבוע של מ"מ וסטיית התקן באותן היחידות כמו מ"מ.

תכונות של שונות:

- $\forall X,\ Var(X)\geq 0$
- $Var(X)=0$ אמ"מ X הוא מ"מ קבוע.
- יהי X מ"מ המקיים: $Var(X)$ סופי, אזי לכל a ו- b $Var(aX+b)=a^2Var(X)$

נשים-לב כי תכונה זו גוררת כי מתקיים גם $SD(aX+b)=|a|SD(X)$. הכוונה בתוכנה היא כי הזזה של משתנה מקרי (b) אינה משנה שונות אך הכפלה בקבוע (a) כן משנה שונות.

- $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$
- עבור $X_1,\cdots,X_n$ בלתי תלויים: $Var\left(\sum X_i\right)=\sum Var(X_i)$.

שונות של התפלגויות ידועות:

- $X\sim Ber(p)$

$Var(X)=p-p^2=\boxed{p\cdot q}$

- $X\sim Uni[1,N]$ $Var(X)=\boxed{\frac{N^2-1}{12}}$

- $X\sim Bin(n,p)$ $Var(X)=\boxed{n\cdot pq}$

- $X\sim Pois(\lambda)$ $Var(X)=\boxed{\lambda}$

- $X\sim Geo(p)$ $Var(X)=\boxed{\frac{q}{p^2}}$

- $X\sim HG(N,D,n)$

$Var(X)=\boxed{\frac{N-n}{N-1}npq}$

- $X\sim NB(m,p)$ $Var(X)=\boxed{\frac{m(1-p)}{p^2}}$

<u>אי-שוויון צ'בצ'ב:</u> לכל מ"מ X ולכל $k>0$ מתקיים: $\boxed{P\left(|X-E(X)|\ge SD(X)\cdot k\right)\le \frac{1}{k^2}}$

<u>משמעות:</u> ההסתברות ש- X חורג מהתוחלת שלו ביותר מ- k פעמים סטיית התקן, קטנה או שווה ל- $\frac{1}{k^2}$.

צורה אחרת לרישום: $\boxed{P\left(|X-E(X)\ge c\right)\le \frac{Var(X)}{c^2}}$

<u>החוק החלש של המספרים הגדולים:</u> יהיו $X_1,X_2,\cdots$ מ"מ ב"ת (בלתי מתואמים גם תנאי מספיק), שווי-התפלגות עם תוחלת μ ושונות- σ^2. יהי $\bar{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ ממוצע המשתנים האלו. אז לכל $\varepsilon>0$ מתקיים: $\boxed{P\left(|X_n-\mu|\ge\varepsilon\right)\xrightarrow{n\to\infty}0}$

<u>משמעות:</u> ככל שהמדגם גדל, הממוצע מתקרב לתוחלת כרצוננו, בהסתברות שואפת ל-1 ($\lim_{n\to\infty}\bar{X}_n=\mu$).

<u>החוק החזק של המספרים הגדולים:</u> עבור תנאי המשפט הקודם, מתקיים:

$$\boxed{P\left(\lim_{x\to\infty}|X_n-\mu|=0\right)=1}$$

<u>משמעות:</u> כאשר הממוצע מתכנס לתוחלת, אז ההסתברות שמאורע יקרה היא:1.

קווריאנס (שונות משותפת) ומקדם מתאם

הגדרות שונות משותפת:

קווריאנס: $\boxed{Cov(X,Y)=E\left((X-EX)(Y-EY)\right)=EXY-EXEY}$

נוסחה נוספת: $\boxed{Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)}$

הערות:

- שונות משותפת מצביעה על תיאום בין המשתנים המקריים: $Cov(X,Y)=0 \Leftarrow X,Y$ בלתי מתואמים (חסרי קורלציה)
- בלתי תלויים $\Leftarrow$ בלתי מתואמים (ההפך לא נכון).

תכונות של שונות משותפת:

- $Cov(X,X)=Var(X)$.
- $Cov(X,Y+Z)=Cov(X,Y)+Cov(X,Z)$
- $Cov(X,Y)=Cov(Y,X)$
- לכל a קבוע, $Cov(X,a)=0$.
- $Cov(aX+b,cY+d)=acCov(X,Y)$

מקדם המתאם:

מקדם המתאם בין X ל- Y:

$$\boxed{Corr(X,Y)=\rho(X,Y)=\frac{Cov(X,Y)}{SD(X)\cdot SD(Y)}=\frac{Cov(X,Y)}{\sqrt{Var(X)}\cdot\sqrt{Var(Y)}}}$$

תקנון: הורדת תוחלת וחלוקתה בסטיית תקן- $X^*=\frac{X-\mu}{\sigma_X}$ כאשר: $E(X^*)=E(Y^*)=1$ ולכן:

$$\rho(X,Y)=\frac{Cov(X,Y)}{SD(X)\cdot SD(Y)}=\frac{E(X-\mu_X)(Y-\mu_Y)}{SD(X)\cdot SD(Y)}=\boxed{E(X^*Y^*)}$$

תכונות של מקדם המתאם:

$$-1\le Corr(X,Y)\le 1 \qquad Corr(X,X)=1$$

$$Corr(X,Y)=0 \Leftrightarrow Cov(X,Y)=0 \qquad Corr(X,Y)=1 \Leftrightarrow Y=aX+b,\ a>0$$

$$Corr(aX+b,cY+d)=sign(ac)\cdot Corr(X,Y) \qquad Corr(X,Y)=-1 \Leftrightarrow Y=aX+b,\ a<0$$

הערה: אם מקדם המתאם הוא-1 בערכו המוחלט, קיים **קשר ליניארי** בין המשתנים.

<u>אופי התלות בין מאורעות:</u>

<u>**תלות חיובית:**</u> מאורעות A ו- B תלויים חיובית אם $\begin{cases} P(A\mid B) > P(A) \\ P(B\mid A) > P(B) \end{cases}$ (הידיעה ש- A קרה, מגדילה את הסיכוי ש- B יקרה ולהפך).

<u>**תלות שלילית:**</u> מאורעות A ו- B תלויים חיובית אם $\begin{cases} P(A\mid B) < P(A) \\ P(B\mid A) < P(B) \end{cases}$ (הידיעה ש- A קרה, מקטינה את הסיכוי ש- B יקרה ולהפך).

<u>**שימוש באינדיקטורים:**</u>

$Cov(I_A, I_B) > 0 \Leftrightarrow P(A\cap B) > P(A)P(B) \Leftrightarrow$ A ו- B <u>תלויים חיובית</u>.

$Cov(I_A, I_B) = 0 \Leftrightarrow P(A\cap B) > P(A)P(B) \Leftrightarrow$ A ו- B <u>בלתי תלויים</u>.

$Cov(I_A, I_B) < 0 \Leftrightarrow P(A\cap B) > P(A)P(B) \Leftrightarrow$ A ו- B <u>תלויים שלילית</u>.

תוחלת מותנית

הגדרה: התוחלת המותנית של מ"מ Y בהינתן שקרה מאורע A מוגדרת ע"י:

$$\boxed{E(X|Y=y)=\sum_x x\cdot P(X=x|Y=y)=\sum_x xP_{X|Y}(x|y)}$$

הערה: התוחלת המותנית, תלויה ב- y ולכן הינה פונקציה של y: $E(X|Y=y)=g(y)$.

תכונות (כמו במקרה הלא מותנה):

- $E(X+Y|Z=z)=E(X|Z=z)+E(Y|Z=z)$
- $E(aX+b|Z=z)=aE(X|Z=z)+b$

תוחלת מותנית כמשתנה מקרי:

הגדרה: נרכיב את פונקצית התוחלת המותנית על Y ונקבל משתנה מ"מ- $\boxed{E(X|Y)=g(Y)}$

משפט ההחלקה: $\boxed{E(X)=E(E(X|Y))}$ וזה מתקיים לכל $h(X)$, פונקציה של מ"מ X.

במקרה הבדיד: $\boxed{E(X)=E(E(X|Y))=\sum_y E(X|Y=y)\cdot P(Y=y)}$

נוסחאות נוספות:

$\boxed{E(X+Y|Y)=E(X|Y)+Y}$ $\boxed{E(X+Y|Y=y)=E(X|Y=y)+y}$

$\boxed{E(XY|Y=y)=y\cdot E(X|Y=y)}$ $\boxed{E(XY|Y)=Y\cdot E(X|Y)}$

משפט WALD: $X_1,X_2,X_3,\cdots$ ב"ת, כולם בעלי תוחלת $\mu=E(X_j)$. N מ"מ אחר, ב"ת בכל ה- X_j-ים. $S_N=X_1+X_2+\cdots+X_N$ אז מתקיים: $\boxed{E(S_N)=E(X)\cdot E(N)=\mu E(N)}$

הערה: ה- X_j-ים עצמם יכולים להיות תלויים או ב"ת בינם לבין עצמם.

דוגמא: לאדם יהיו N ילדים. לילד מספר k יהיו X_k ילדים משלו, כאשר X_k-ים הם ב"ת ב- N. ו- $E(X_k)=\mu$ לכל k. אז מספר הנכדים של אותו אדם $=S_N$, התוחלת שלו: $E(S_N)=\mu E(N)$.

שונות מותנית:

הגדרה: X,Y מ"מ אז: $\boxed{\sigma_Y^2=E(\sigma_{Y|X}^2)+Var(E(Y|X))}$

פונקציות יוצרות

מומנטים ופונקציות יוצרות מומנטים:

מומנט: יהי X מ"מ. המומנט ה-n-י של X הוא: $E\left(X^n\right)$.

עבור X בדיד: $\boxed{E\left(X^n\right)=\sum_{x\in R} x^n\cdot P\left(X=x\right)}$

פונקציה יוצרת מומנטים: (פי"מ) מסומנת ב- $\boxed{M_X\left(t\right)=E\left(e^{tX}\right)}$

עבור X בדיד: $\boxed{M_X\left(t\right)=\sum_{x\in R} e^{tX}\cdot P\left(X=x\right)}$

הערה: ניתן לקבל מהפונקציה את כל המומנטים של X ע"י חישוב נגזרותיה ומציאת הערך של הנגזרת בנקודה $t=0$, כלומר: $\boxed{\frac{\partial^n M}{\partial t^n}(t=0)=E\left(X^n\right)}$

עבור כמה מ"מ: פונקציה של סכום משתנים מ"מ ב"ת שווה למכפלת הפונקציות יוצרות המומנטים שלהם: $\boxed{M_{X+Y}\left(t\right)=M_X\left(t\right)\cdot M_Y\left(t\right)}$ (וזה נכון גם למספר רב יותר של מ"מ).

תכונה: יהי X מ"מ, a,b קבועים, אזי: $\boxed{M_{aX+b}\left(t\right)=e^{bt}M\left(at\right)}$

משפט: פונקציה יוצרת מומנטים (אם קיימת) מגדירה את התפלגותו של מ"מ באופן חד-משמעי. (אבל כל המומנטים ללא הפונקציה לא כיולים לקבוע באופן חח"ע את ההתפלגות).

דוגמאות:

- $X\sim Ber\left(p\right)$: $M_X\left(t\right)=E\left(e^{tX}\right)=e^t p+e^0\left(1-p\right)=\boxed{pe^t+q}$

- $X\sim Bin\left(n,p\right)$: $M_X\left(t\right)=\sum_{k=1}^{n} e^{tk}\binom{n}{k}\left(pe^t\right)^k q^{n-k}=\boxed{\left(pe^t+q\right)^n}$

- $X\sim Pois\left(\lambda\right)$: $M_X\left(t\right)=E\left(e^{tX}\right)=\sum_{n=0}^{\infty} e^{nt}\cdot\frac{\lambda^n e^{-\lambda}}{n!}=e^{-\lambda}\cdot\sum_{n=0}^{\infty}\frac{\left(e^t\lambda\right)^n}{n!}$
$=e^{-\lambda}\cdot e^{\lambda e^t}=\boxed{e^{\lambda\left(e^t-1\right)}}$

- סכום מ"מ פואסונים: $M_{X+Y}\left(t\right)=e^{\lambda\left(e^t-1\right)}\cdot e^{\mu\left(e^t-1\right)}=e^{\left(\lambda+\mu\right)\left(e^t-1\right)}$

פונקציה יוצרת הסתברות:

הגדרה: יהי X מ"מ המקבל ערכים שלמים ואי-שליליים. $\boxed{g_X(s)=E\left(s^x\right)=\sum_{x\in R_X}s^xP(X=x)}$

תכונות:

- $\frac{\partial^k g_X}{\partial s^k}(s=0)=k!P(X=k)$ (בפרט אם מציגים את g_X כטור חזקות סביב $s=0$ אז המקדם של s^k הוא ההסתברות ש- X יקבל את הערך- k).
- $$g_X(1)=1$$
  $$\frac{\partial g_X}{\partial s}(s=1)=E(X)$$
  $$\frac{\partial g_X}{\partial s^2}(s=1)=E(X(X-1))=E\left(X^2\right)-E(X)$$
- $g^{(n)}(0)=n!\cdot P(n)$.
- יהיו $X_1,\cdots,X_n$ מ"מ ב"ת, אזי: $g_{\sum_{i=1}^n X_i}(t)=\prod_{i=1}^n g_{X_i}(t)$

משפט: פונקציה יוצרת הסתברות של מ"מ (אם קיימת) מגדירה באופן חח"ע את פונקצית ההסתברות של המשתנה המקרי.

פונקציה אופיינית:

הגדרה: $\boxed{\phi_X(\theta)=E\left(e^{i\theta X}\right)=\sum e^{i\theta x}p(x)}$

תהליכי הסתעפות

במצב ההתחלתי יש צומת יחיד בעץ, $i=0$. בשלב ה- i לכל צומת v ברמה ה $i-1$ מתווספים Cv בנים, כאשר Cv הוא משתנה מקרי.

<u>הנחה:</u> כל ה- Cv הם בלתי תלויים ושווי-התפלגות.

<u>סימונים:</u>

g - פונקצית יוצרת הסתברות של המשתנה המקרי Cv.

μ - התוחלת של המשתנה המקרי Cv

Zi - משתנה מקרי שערכו שווה למספר הצמתים הכולל ברמה ה- i בעץ, $Z0=1$.

g_i - פונקציה יוצרת הסתברות של המשתנה המקרי Zi

<u>משפטים:</u>

1. $g_0(s)=s,\ g_i(s)=g_{i-1}(g(s))=g(g(g(\cdots(s)\cdots)))$ עבור $i=1,2,\cdots$
2. $E(Zi)=\mu^i$ עבור $i=0,1,2,\cdots$.

התפלגויות רציפות וצפיפות

מ"מ רציף: מ"מ המגודר על מרחב מדגם Ω ייקרא רציף אם קיימת פונקציה $f_X(x)$ כך ש:

$$\boxed{P(X\in[a,b]) = P(a\le X\le b) = \int_a^b f_X(x)\,dx}$$

$f_x(x)$ נקראת **פונקצית הצפיפות** של מ"מ X.

תכונות:

- כאשר dx קטן מאוד אז מתקיים: $P(X\in[x,x+dx]) = f_X(x)\cdot dx$
- $P(X\le a) = \int_{-\infty}^{a} f_x(x)\,dx$
- $P(a\le X\le b) = P(X\le b) - P(X\le a)$

אילוצים:

- $\forall x,\ f_x \ge 0$
- $P(-\infty\le X\le\infty) = \int_{-\infty}^{\infty} f_x(x)\,dx = 1$

הערות:

- f_x לא מציינת הסתברות ולכן ייתכן כי $f_X(x) > 1$.
- "שטח נקודה אפס": $P(X=x) = \int_x^x f_x(t)dt = 0 \Rightarrow P(a\le X\le b) = P(a<X<b)$.

תוחלת של מ"מ רציף: $\boxed{E(X) = \int_{-\infty}^{\infty} x\cdot f_X(x)\,dx}$ $\boxed{E(g(X)) = \int_{-\infty}^{\infty} g(x)\cdot f_x(x)\,dx}$

הערה: התוחלת קיימת רק כאשר $\int_{-\infty}^{\infty} |x|\cdot f_X(x)\,dx < \infty$ (התכנסות במידה שווה).

שונות של מ"מ רציף: $\boxed{Var(X) = \int_{-\infty}^{\infty} x^2\cdot f_x(x)\,dx - \left(\int_{-\infty}^{\infty} x\cdot f(x)\,dx\right)^2}$

התפלגות אחידה (יוניפורמית):

$X \sim Uni(a,b)$ כאשר $a<b$.
$$f_X(x)=\begin{cases}\frac{1}{b-a} & x\in[a,b]\\ 0 & otherwise\end{cases}$$

$$E(X)=\frac{b+a}{2}\qquad Var(X)=\frac{(b-a)^2}{12}\qquad P(c\le X\le d)=\frac{d-c}{b-a}$$

התפלגות נורמלית:

נורמאלי סטנדרטי: $Z\sim N(0,1)$ $\quad -\infty<z<\infty\quad$ $f_Z(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$

ומתקיים: $P(Z\le a)=\int_{-\infty}^{a}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz$, כאשר לאינטגרל הזה אין פיתרון אנליטי ומסמנים:
$P(Z\le a)=\Phi(a)$, את הערכים מוצאים בטבלה.

נורמאלי כללי: $\mu\in R,\ \sigma>0$ $\quad X\sim N(\mu,\sigma^2)\quad -\infty<x<\infty\quad$ $f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

ניתן לנרמל את X כדי להיות נורמאלי סטנדרטי:
$$P(X\le b)=P\left(\frac{X-\mu}{\sigma}\le\frac{b-\mu}{\sigma}\right)=P\left(Z\le\frac{b-\mu}{\sigma}\right)=\Phi\left(\frac{b-\mu}{\sigma}\right)$$
$$P(a\le X\le b)=P(X\le b)-P(X\le a)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

סכום של משתנים נורמאלים: $X_1\sim N(\mu_1,\sigma_1^2)$, $X_2\sim N(\mu_2,\sigma_2^2)$ אז:
$X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

משפט הגבול המרכזי

יהיו $X_1, X_2, \cdots, X_n$ מ"מ ב"ת ש"ה עם תוחלת μ ושונות σ^2. נגדיר: $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ אזי:

$\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow{n\to\infty} N(0,1)$ כלומר $\overline{X}_n$ בקירוב מתפלג כמו משתנה מקרי מהתפלגות נורמאלית-

$N\left(\mu, \frac{\sigma^2}{n}\right)$, כלומר: $P\left(\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \le a\right) \xrightarrow{n\to\infty} \Phi(a) = P\left(N(0,1) \le a\right)$

באופן שקול אם נגדיר: $S_n = \sum_{i=1}^{n} X_i$ אזי: $\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{n\to\infty} N(0,1)$.

S_n מתפלג בקירוב כמו מ"מ: $N\left(n\mu, n\sigma^2\right)$.

התפלגות אקספוננציאלית והתפלגות גאמא:

התפלגות מעריכית:

סימון: $X \sim \exp(\lambda)$. **פונקצית צפיפות:** $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases}$.

תוחלת: $\boxed{E(X) = \frac{1}{\lambda}}$ **שונות:** $\boxed{Var(X) = \frac{1}{\lambda^2}}$

חישוב הסתברות: $P(X > a) = \int_a^\infty f_X(x)dx = \int_a^\infty \lambda e^{-\lambda x} dx = e^{-\lambda a}$

תכונת חוסר הזיכרון: אם נתון כי עד רגע t (כולל) לא היה מופע, אז "מתחילים מהתחלה", ולכן הזמן עד המופע הבא הוא: $\exp(\lambda)$. $P(X > t+s \mid X > t) = P(X > s) = e^{-\lambda s}$.

משפט: ההתפלגות הרציפה היחידה בעלת תכונת חוסר זיכרון היא ההתפלגות האקספוננציאלית.

התפלגות Gamma:

סימון: $X \sim Gamma(n, \lambda)$.

פונקצית צפיפות: $f_X(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x} (\lambda x)^{n-1}}{\Gamma(n)} = \dfrac{\lambda e^{-\lambda x} (\lambda x)^{n-1}}{(n-1)!} & x \geq 0 \\ 0 & otherwise \end{cases}$

כאשר n, λ וכאשר $\Gamma(n) = \int_0^\infty e^{-y} \cdot y^{n-1} dy$, שזוהי פונקצית גמא.

תכונות:

- עבור n שלם מתקיים: $\Gamma(n) = (n-1) \cdot \Gamma(n-1)$
- $\Gamma(n) = (n-1)!$.
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

תוחלת: $\boxed{E(X) = \frac{n}{\lambda}}$ **שונות:** $\boxed{Var(X) = \frac{n}{\lambda^2}}$

הערה: התפלגות אקספוננציאלית הינה מקרה פרטי של התפלגות $Gamma$:

$$\boxed{Gamma(1,\lambda)=\exp(\lambda)}$$

גאמא- סכום מ"מ אקספוננציאלים: יהיו $X_1,\cdots,X_n$ מ"מ ב"ת, $X_i \sim \exp(\lambda)$. נגדיר: $Y=X_1+\cdots+X_n$. אזי: $Y \sim Gamma(n,\lambda)$.

הקשר בין מ"מ אקספוננציאלי גאמא ופואסון:

N_t - מספר המופעים בקטע- $[0,t]$.

T_i - זמן המופע ה- i

$(T_i - T_{i-1})$ - זמן בין מופעים.

- אם $N_t \sim Pois(\lambda t)$ אז:
  - $(T_i - T_{i-1}) \sim \exp(\lambda)$
  - $T_i \sim Gamma(i,\lambda)$
- אם $(T_i - T_{i-1}) \sim \exp(\lambda)$ וב"ת אזי:
  - $N_t \sim Pois(\lambda t)$

נוסחה לחישוב הסתברויות עבור $Gamma(i,\lambda)$:

$$\boxed{P(T_i > t) = P(N_t \le i-1) = \sum_{j=0}^{i-1} \frac{e^{-\lambda t}(\lambda t)^j}{j!}}$$

משמעות: ההסתברות שהמאורע ה- i יקרה אחרי זמן- t שווה להסתברות לכך שעד לזמן- t קרו לכל היותר $i-1$ מאורעות.

פונקצית התפלגות מצטברת

הגדרה: יהי X מ"מ. פונקצית התפלגות מצטברת של X היא פונקציה $F_X : R \to [0,1]$ המוגדרת ע"י: $F_X(x) = P(X \le x)$.

עבור X בדיד: $F_X(x) = \sum_{x_R \le x} P_X(x)$

עבור X רציף: $\boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \quad \Rightarrow \quad f_X(x) = \frac{\partial F_X(x)}{\partial x}}$

כלומר: ע"י גזירה של פונקצית התפלגות מצטברת, ניתן לקבל את הצפיפות של מ"מ רציף.

תכונות:

- $F(\infty) = 1$
- $F(-\infty) = 0$
- $x < y \Rightarrow F(x) \le F(y)$, F מונוטונית.

הערות:

- פונקצית ההתפלגות המצטברת מגדירה את ההתפלגות של מ"מ. כלומר, אם למשתנים מקריים אותה פונקצית התפלגות מצטברת, אז להם אותה צפיפות ואתה התפלגות על- R.
- עבור X רציף: $\boxed{P(a \le X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)}$

נוסחת הזנב: X מ"מ בדיד, שמקבל ערכים שלמים לא שליליים. $\boxed{E(X) = \sum_{k=1}^{\infty} P(X \ge k)}$

במקרה רציף זה יכול להיות אי-שיוויון ממש ואז מקבלים: $\boxed{E(X) = \int_{0}^{\infty} (1 - F_X(x))\,dx}$

טרנספורמציות של מ"מ רציפים

שינוי משתנה:

שינוי ליניארי: X מ"מ רציף בעל פונקצית צפיפות $f_X(x)$, פונקצית הצפיפות של $Y=a+bX$ ($b\neq 0$) ניתנת ע"י: $\boxed{f_Y(y)=\frac{1}{|b|}f_X\left(\frac{y-a}{b}\right)}$.

הרכבה: X מ"מ רציף בעל פונקצית צפיפות $f_X(x)$. g מונוטונית עולה ממש או יורדת ממש (ואז היא הפיכה בתחום מסוים). פונקצית הצפיפות $Y=g(x)$: $\boxed{f_Y(y)=\left|\left(g^{-1}(y)\right)'\right|\cdot f_X\left(g^{-1}(y)\right)}$

סימולציה:

נתון משתנה אחיד: $U\sim uni[0,1]$.

המטרה: ע"י שימוש ב- U ועל-סמך טרנספורמציה, לקבל <u>תצפית</u> מהתפלגות של מ"מ X, כאשר F פונקצית ההתפלגות המצטברת של- X, נתונה.

נרצה למצוא פונקציה g כך ש- $P\left(g(U)\le x\right)=F(x)$. בדרך זו $g(U)$ ו- X יהיו שווי התפלגות.

בנוסף מתקיים ש- $F(X)\sim Uni(0,1)$. לכן נגדיר $U=F(X)$ ונחלץ את X ($g(U)=F^{-1}(U)$).

על-מנת להגריל מ"מ X, נגריל עתה מ"מ $Uni(0,1)$ ונציב בפונקציה ההופכית שמצאנו.

התפלגות חי בריבוע:

יהיו $X_1,X_2,\cdots,X_n$ מ"מ, ב"ת, כאשר $X_i\sim N(0,1)$. נגדיר: $Y=\sum_{i=1}^{n}X_i^2$.

אז: $Y\sim\chi^2(n)\equiv Gamma\left(\frac{n}{2},\frac{1}{2}\right)$ כלומר: Y מפולג חי בריבוע עם n דרגות חופש.

$$f_X(x,n)=\begin{cases}\frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma(n/2)}\cdot x^{(n/2)-1}e^{-x/2} & x>0\\ 0 & otherwise\end{cases}$$

<u>הערה:</u> כאשר $X_1,X_2\sim N(0,1)$, $\left(X_1^2+X_2^2\right)\sim\chi^2(2)\equiv Gamma\left(2/2,1/2\right)\equiv\exp\left(1/2\right)$

<u>דוגמא:</u> $Y=X^2$, $X\sim N(0,1)=\frac{1}{2\pi}e^{-1/2x^2}$. אז: $f_Y(y)=\frac{y^{-1/2}e^{-y/2}}{\sqrt{2\pi}}$ ו- $Y\sim\chi^2=\Gamma\left(\frac{1}{2},\frac{1}{2}\right)$

התפלגויות וצפיפויות משותפות

התפלגות משותפת רציפה של זוג מ"מ:

ל-מ"מ X,Y המוגדרים על אותו מרחב המדגם יש התפלגות משותפת רציפה אם לכל $B\subseteq R^2$ מתקיים: $P\left(\left(X,Y\right)\in B\right)=\int\limits_B f_{X,Y}\left(t,s\right)dtds$,

צפיפות משותפת: $f_{X,Y}\left(t,s\right)$ היא פונקצית הצפיפות המשותפת של X ו- Y המקיימת:

- $f_{X,Y}\left(t,s\right)\geq 0$ לכל t,s.
- $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f_{X,Y}\left(t,s\right)dtds=1$
- $P\left(x\leq X\leq x+dx, y\leq Y\leq y+dy\right)=f_{X,Y}\left(x,y\right)dxdy$

(מבחינה גיאומטרית, ההסתברות במקרה הדו-מימדי מתוארת ע"י השטח במלבן $dx\times dy$.

שיטות למציאת צפיפות משותפת:

1. ע"י: $\boxed{f_{X,Y}\left(x,y\right)dxdy=P\left(x\leq X\leq x+dx, y\leq Y\leq y+dy\right)}$
2. מציאת $F_{X,Y}\left(x,y\right)$ וגזירתה.

התפלגות מצטברת משותפת: עבור $\left(X,Y\right)$ - זוג מ"מ המוגדרים על אותו מרחב המדגם,

פונקצית ההתפלגות המשותפת: $\boxed{F_{X,Y}\left(x,y\right)=P\left(X\leq x, Y\leq y\right)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f_{X,Y}\left(u,v\right)dudv}$

ומתקיים: $\boxed{f_{X,Y}\left(x,y\right)=\frac{\partial^2}{\partial x\partial y}\left(F_{X,Y}\left(x,y\right)\right)}$

פונקציות צפיפות שוליות:

פונקצית הצפיפות השולית של- X : $\boxed{f_X\left(x\right)=\int\limits_{-\infty}^{\infty} f_{X,Y}\left(x,y\right)dy}$

פונקצית הצפיפות השולית של- Y : $\boxed{f_Y\left(y\right)=\int\limits_{-\infty}^{\infty} f_{X,Y}\left(x,y\right)dx}$

הערה: בכל-פעם עושים אינטגרציה על המשתנה שרוצים להשמיט, נכון גם ליותר מ-2 משתנים.

אי-תלות בין מ"מ רציפים:

X,Y יקראו מ"מ <u>בלתי תלויים</u> אם מתקיים: $f_{X,Y}(x,y)=f_X(x)\cdot f_Y(y)$ לכל $x,y\in R$.

X,Y יקראו מ"מ <u>בלתי תלויים</u> אם מתקיים: $F_{X,Y}(x,y)=F_X(x)\cdot F_Y(y)$ לכל $x,y\in R$.

<u>הערה:</u> אם ניתן להפריד את $f_{X,Y}(x,y)$ למכפלת שתי פונקציות באופן הבא: $f_{X,Y}(x,y)=g(x)\cdot h(y)$ ותחומיהם השוליים של X ו- Y אינם תלויים אחד בשני, כלומר- התחום המשותף של (X,Y) הינו מלבן המקביל לצירים, אזי המתנים המקריים הם ב"ת ו- $g(x)$ ו- $h(y)$ הן פונקציות הצפיפות של X,Y <u>עד כדי קבוע</u>.

<u>כתיבה דרך הסתברות:</u> $P(X\in dx,Y\in dy)=P(X\in dx)\cdot P(Y\in dy)$.

התפלגות אחידה דו-מימדית:

עבור $B\subseteq R^2$, נאמר ש- (X,Y) מפולגים אחיד על B, $(X,Y)\sim Uni(B)$ אם:

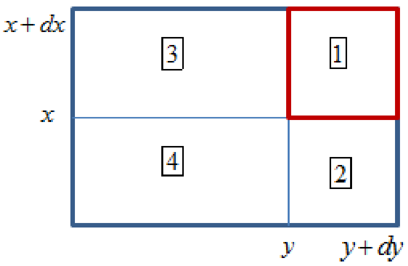
$f_{X,Y}(t,s)=\begin{cases}\frac{1}{area\ B} & (t,s)\in B\\ 0 & otherwise\end{cases}$ כאשר (X,Y) נבחרה באקראי על B.

<u>הערות:</u>

- עבור $B\subseteq A$ מתקיים: $P((X,Y)\in A)=\frac{area\ A}{area\ B}$.

- אם $X\sim Uni[a,b]$ ו- $Y\sim Uni[c,d]$, ו- X,Y ב"ת, אז הם מתפלגים במשותף אחיד דו-מימדי על המלבן: $[a,b]\times[c,d]$.

<u>תכונה:</u>



$$P\Big(\underbrace{(X,Y)\in[x,x+dx]\times[y,y+dy]}_{\boxed{1}}\Big)=\underbrace{F(x+dx,y+dy)}_{\boxed{1}+\boxed{2}+\boxed{3}+\boxed{4}}-\underbrace{F(x+dx,y)}_{\boxed{2}+\boxed{4}}-\underbrace{F(x,y+dy)}_{\boxed{3}+\boxed{4}}+\underbrace{F(x,y)}_{\boxed{4}}$$

סטטיסטי הסדר במקרה הרציף

הגדרה: $X_1, X_2, \dots, X_n$ ב"ת, ש"ה, f_X ~ X_i. $\underset{\min\limits_{1\le i\le n} X_i}{X_{(1)}} < X_{(2)} < \dots < \underset{\max\limits_{1\le i\le n} X_i}{X_{(n)}}$

מינימום, מקסימום והמקרה הכללי:

$$\boxed{\begin{aligned} &F_{(1)} = P(\min X_i < x) = 1 - P(\min X_i > x) = 1 - (1 - F(x))^n \\ &f_{(1)}(x) = \frac{\partial}{\partial x} F_{(1)} = nf(x) \cdot (1 - F(x))^{n-1} \end{aligned}}$$

$$\boxed{\begin{aligned} &F_{(n)} = P(\max X_i < x) = (F(x))^n \\ &f_{(n)}(x) = \frac{\partial}{\partial x} F_{(n)} = nf(x)(F(X))^{n-1} \end{aligned}}$$

$$\boxed{f_{(k)}(x) = n\binom{n-1}{k-1} f(x) F(x)^{k-1} (1 - F(x))^{n-k}}$$

דוגמא: X_i ~ $Uni[0,1]$, $f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$. אז: $f_{(k)}(x) = n\binom{n-1}{k-1} x^{k-1} (1-x)^{n-k}$

פונקצית צפיפות משותפת של $X_{(1)}, \dots, X_{(n)}$: $f_{X_{(1)},\dots,X_{(n)}}(x_1, \dots, x_n) = \begin{cases} n! & 0 \le x_1 \le \dots \le x_n \le 1 \\ 0 & otherwise \end{cases}$

התפלגות בטא:

$f_X(x) = \begin{cases} \dfrac{x^{r-1}(1-x)^{s-1}}{B(r,s)} & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$ כאשר: $B(r,s) = \int_0^1 x^{r-1}(1-x)^{s-1} dx$

דוגמא: סטטיסטי הסדר מתפלג: $\beta(k+1, n-k+1)$

צפיפות מותנית ותוחלת מותנית

צפיפות מותנית:

הגדרה: $f_Y(y) > 0$ כך ש- y המוגדרת עבור כל $Y = y$ בהינתן X הצפיפות המותנית של:

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

כללי הסתברות במקרה הרציף:

- כלל הכפל: $\boxed{f_{X,Y}(x,y) = f_X(x) \cdot f_{Y|X}(y|x)}$

- פונקצית ההתפלגות המצטברת המותנית: $\boxed{F_{X|Y}(x|y) = P(X \le x | Y = y) = \int_{-\infty}^{x} f_{X|Y}(t|y)\,dt}$

- חישוב הסתברות מותנית: $\boxed{P(a \le X \le b | Y = y) = \int_{a}^{b} f_{X|Y}(x|y)\,dx}$

- צפיפות שולית דרך צפיפות מותנית: $\boxed{f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y) \cdot f_Y(y)\,dy}$

-נוסחת בייס לצפיפות: $\boxed{f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)\,f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)\,f_X(x)}{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx}}$

- נוסחת בייס למאורע: $\boxed{f_X(x|A) = \frac{P(A|X = x)\,f_X(x)}{\int_{-\infty}^{\infty} P(A|x)\,f_X(x)\,dx}}$

תוחלת מותנית:

הגדרה: $\boxed{E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx}$

הערה: אם נסמן $h(y) = E(X|Y = y)$ אז: $h(Y) = E(X|Y)$ הוא מ"מ.

משפט ההחלקה במקרה הרציף: $\boxed{E(X) = E(E(X|Y)) = E(h(Y))}$

נוסחה לחישוב שונות: $\boxed{Var(X) = E(Var(X|Y)) + Var(E(X|Y))}$

טרנספורמציות

טרנספורמציות חח"ע של מ"מ רציף:

הגדרה: יהיה X מ"מ רציף המקבל ערכים בקטע $[a,b]$. תהיה $g:R\to R$ פונקציה גזירה ועולה (יורדת) ב-$[a,b]$. אזי למ"מ Y המוגדר להיות $Y=g(X)$ יש את פ' הצפיפות הבאה:

$$f_Y(y)=\begin{cases} f_X\left(g^{-1}(y)\right)\cdot\left|\dfrac{\partial g^{-1}(y)}{\partial y}\right| & g(a)\overset{(\ge)}{\le} y\overset{(\ge)}{\le} g(b) \\ 0 & otherwise\end{cases}$$

טרנספורמציה דו-ממדית:

הגדרה: יהי (X,Y) מ"מ דו-ממדי רציף בעל צפיפות משותפת: $f_{X,Y}(x,y)$. תהיינה $u=g_1(x,y),\ v=g_2(x,y)$ פונקציות ממשיות על המישור, שהן חח"ע וגזירות, כך שקיימות הפונקציות ההפוכות: $x=h_1(u,v),\ y=h_2(u,v)$ והן גזירות. אז נגדיר את היעקוביאן:
$J(u,v)=\det\begin{pmatrix}\frac{\partial h_1}{\partial u} & \frac{\partial h_1}{\partial v}\\ \frac{\partial h_2}{\partial u} & \frac{\partial h_2}{\partial v}\end{pmatrix}$. אז פונקצית הצפיפות המשותפת של $U=g_1(X,Y),\ V=g_2(X,Y)$

נתונה ע"י: $\boxed{f_{U,V}(u,v)=f_{X,Y}\left(h_1(u,v),h_2(u,v)\right)\cdot\left|J(u,v)\right|}$

הערה: עבור מ"מ $U=g_1(X,Y)$ ניתן לכתוב מ"מ נוסף $V=X$ או $V=Y$, לבצע טרנספורמציה דו-ממדית עבור U,V ולמצוא צפיפות שולית של U ע"י אינטגראציה.

מקרה פרטי: מעבר מקואורדינאטות קרטזיות לפולאריות: נניח שלמ"מ (X,Y) יש צפיפות $f_{X,Y}(x,y)$, אזי לקואורדינאטות הפולאריות יש צפיפות משותפת:

$$\boxed{f_{R,\theta}(r,t)=rf_{X,Y}(r\cdot\cos t,r\cdot\sin t)}$$

טרנספורמציות רב-ממדיות:

הגדרה:

$\hat{X}=(X_1,X_2,\cdots,X_n)\sim f_{X_1,X_2,\dots,X_n}(x_1,x_2,\cdots,x_n)$ וגם:

$$\begin{aligned}Y_1&=g_1(X_1,X_2,\cdots,X_n)\\ Y_2&=g_2(X_1,X_2,\cdots,X_n)\\ &\vdots\\ Y_n&=g_n(X_1,X_2,\cdots,X_n)\end{aligned}$$

כל פונקציות g הן חח"ע ולכן הפיכות, ואז קיימות:

$$\begin{aligned} X_1 &= h_1\left(Y_1, Y_2, \cdots, Y_n\right) \\ X_2 &= h_2\left(Y_1, Y_2, \cdots, Y_n\right) \\ &\vdots \\ X_n &= h_n\left(Y_1, Y_2, \cdots, Y_n\right) \end{aligned}$$

אז נגדיר את היעקוביאן: $J\left(y_1, y_2 \cdots, y_n\right) = \begin{pmatrix} \frac{\partial h_1(y)}{\partial y_1} & \cdots & \frac{\partial h_1(y)}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial h_n(y)}{\partial y_1} & \cdots & \frac{\partial h_n(y)}{\partial y_n} \end{pmatrix}$.

ונקבל כי פונקצית הצפיפות המשותפת ל- $\hat{Y} = \left(Y_1, Y_2, \cdots, Y_n\right)$ היא:

$$\boxed{f_{Y_1, Y_2, \ldots, Y_n}\left(y_1, y_2, \cdots, y_n\right) = f_{X_1, X_2, \ldots, X_n}\left(h_1(\hat{y}), h_2(\hat{y}), \cdots, h_n(\hat{y})\right) \cdot \left\| J\left(y_1, y_2, \cdots, y_n\right) \right\|}$$

<u>התפלגות שולית מהמשותפת:</u> $f_{X_1, X_2}\left(x_1, x_2\right) = \underbrace{\int_{x_3} \cdots \int_{x_n}}_{n-2} f_{X_1, X_2 \cdots, X_n}(x)\, dx_3 \cdots dx_n$.

<u>טרנספורמציות ליניאריות:</u>

$\hat{Y} = A\hat{X}$, $\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$. נניח ש- A^{-1} קיימת. אז: $X = A^{-1}Y$.

$$\boxed{f_{\hat{Y}}(\hat{y}) = \frac{1}{|\det A|} \cdot f_{\hat{X}}\left(A^{-1}\hat{y}\right)}$$

סכום של משתנים מקריים

הגדרה: יהיו X,Y מ"מ ויהי $Z = X + Y$.

<u>במקרה הבדיד:</u> פונקצית ההתפלגות של- Z - $\boxed{P_Z\left(z\right) = \sum_x P_{X,Y}\left(x, z-x\right)}$

אם X,Y ב"ת: $\boxed{P_Z\left(z\right) = \sum_x P_X\left(x\right) P_y\left(z-x\right)}$

<u>במקרה הרציף:</u> פונקצית הצפיפות של Z - $\boxed{f_Z\left(z\right) = \int_{-\infty}^{\infty} f_{X,Y}\left(x, z-x\right)}$

אם X,Y ב"ת: $\boxed{f_Z\left(z\right) = \int_{-\infty}^{\infty} f_X\left(x\right) f_Y\left(z-x\right) dx}$

<u>הערה:</u> כאשר X,Y ב"ת זה נקרא קונבולוציה.

התפלגות רב נורמאלית ודו נורמאלית

התפלגות רב-נורמאלית:

הגדרה: יהי $\hat{X}$ וקטור אקראי ממימד k המפולג נורמאלית עם תוחלת $E\hat{X}=\hat{\mu}$ ומטריצת קווריאנס Σ, כלומר: $\hat{X}\sim N(\hat{\mu},\Sigma)$. ניתן לכתוב את פונקצית הצפיפות של $\hat{X}$ בצורה:

$$\boxed{f_{\hat{X}}(\hat{x})=\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}}\cdot\exp\left(-\frac{1}{2}\cdot(\hat{x}-\hat{\mu})^t\cdot\Sigma^{-1}(\hat{x}-\hat{\mu})\right)}$$

מטריצת קווריאנס: יהי $\hat{X}$ וקטור אקראי. Σ נקראת מטריצת הקווריאנס של $\hat{X}$ ומתקיים:

$$\Sigma=E\left((X-EX)(X-EX)^t\right)=$$

$$\begin{pmatrix}Cov(X_1,X_1) & \cdots & Cov(X_1,X_n)\\ \vdots & \ddots & \vdots\\ Cov(X_n,X_1) & \cdots & Cov(X_n,X_n)\end{pmatrix}=\begin{pmatrix}Var(X_1) & \cdots & Cov(X_1,X_n)\\ \vdots & \ddots & \vdots\\ Cov(X_n,X_1) & \cdots & Var(X_n)\end{pmatrix}$$

על האלכסון הראשי מצויות השונויות של איברי $\hat{X}$ בהתאמה, מחוץ לאלכסון נמצאות השונות המשותפת (קווריאנס) של איברי $\hat{X}$, במיקומים המתאימים.

טרנספורמציה ליניארית של וקטור אקראי נורמאלי:

יהי וקטור אקראי $\hat{X}\sim N(\hat{\mu},\Sigma)$. תהא מטריצה A (לאו דווקא ריבועית) ויהי וקטור b. אם $\hat{Y}=A\hat{X}+b$, אזי $\boxed{\hat{Y}\sim N\left(A\hat{\mu}+b,A\Sigma A^t\right)}$

מסקנות:

1. כל התפלגות שולית של וקטור אקראי נורמאלי היא נורמאלית. לדוגמא אם $(X_1,X_2,X_3)^t$ מפולג נורמאלית אז גם $(X_1,X_3)^t$ מפולג נורמאלית. במקרה זה $A=\begin{pmatrix}1&0&0\\0&0&1\end{pmatrix}$ ו- $b=\begin{pmatrix}0\\0\end{pmatrix}$.

2. כל ק"ל של משתנים מקריים, שפילוגם המשותף נורמאלי, היא נורמאלית. לדוגמא: $3X_1-X_2+2X_3$ מפולג נורמאלית. במקרה זה $A=\begin{pmatrix}3&-1&2\end{pmatrix}$ ו- $b=0$.

משפט: משתנים שפילוגם המשותף נורמאלי והם בלתי מתואמים, הינם גם בלתי תלויים.

הערות:

- הכיוון השני אינו נכון, ייתכנו מ"מ בעלי התפלגות שולית נורמאלית שהם בלתי מתואמים אך אינם מפולגים יחד נורמאלית.

- רק פילוג נורמאלי מקיים את המשפט בתנאי החשוב שההתפלגות המשותפת היא נורמאלית.

התפלגות דו-נורמאלית:

הגדרה: יהי $X=\begin{pmatrix}X_1\\X_2\end{pmatrix}$ וקטור המתפלג נורמאלית עם תוחלת- $E\hat{X}=\hat{\mu}$ ומטריצת קווריאנס Σ,

כלומר: $\hat{X}\sim N(\hat{\mu},\Sigma)$. אז ניתן לכתוב את פונקצית הצפיפות של $\hat{X}$ בצורה:

$$\boxed{f_{\hat{X}}(\hat{x})=\frac{1}{2\pi|\Sigma|^{1/2}}\cdot\exp\left(-\frac{1}{2}(\hat{x}-\hat{\mu})^t\Sigma^{-1}(\hat{x}-\hat{\mu})\right)}$$

אם נסמן- $E(X_i)=\mu_i,\ \ Var(X_i)=\sigma_i,\ \ Corr(X_1,X_2)=\rho$, אזי:

$\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$ $\hat{\mu}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$ ופונקצית הצפיפות המשותפת היא:

$$\boxed{\begin{aligned}&f_{X_1X_2}(x_1,x_2)=\\&\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\left(-\frac{1}{2(1-\rho^2)}\cdot\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\cdot\left(\frac{x_1-\mu_1}{\sigma_1}\right)\cdot\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right)\end{aligned}}$$

תכונות:

א. X_1,X_2 מפולגים נורמאלית באופן שולי $X_1\sim N(\mu_1,\sigma_1^2),\ \ X_2\sim N(\mu_2,\sigma_2^2)$.

ב. הכיוון השני לא בהכרח נכון, כלומר לא כל שני משתנים נורמאלים בעלי מקדם מתאם כלשהו ρ, יתפלגו במשותף דו-נורמאלית.

דוגמא: $f_{X,Y}(x,y)=\begin{cases}\frac{2}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2+y^2)} & x\cdot y>0\\ 0 & x\cdot y<0\end{cases}$ זוהי אינה התפלגות נורמאלית, אבל

ההתפלגויות השוליות הן כן נורמאליות: $f_X(x)=\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$

ג. ההתפלגויות המותנות הן נורמאליות:

$$\boxed{(X_2\mid X_1=x_1)\sim N\left(\mu_2+\rho\cdot\frac{\sigma_2}{\sigma_1}(x_1-\mu_1),\sigma_2^2(1-\rho^2)\right)}$$

$$\boxed{(X_1\mid X_2=x_2)\sim N\left(\mu_1+\rho\cdot\frac{\sigma_1}{\sigma_2}(x_2-\mu_2),\sigma_1^2(1-\rho^2)\right)}$$

ד. שני מ"מ ב"ת שפילוגם השולי נורמאלי, מפולגים יחד דו-נורמאלית עם מקדם מתאם אפס.

משפט: אם $(X,Y)\sim N(\mu,\Sigma)$ אזי $\left(\frac{X-\mu_x}{\sigma_x},\frac{Y-\mu_y}{\sigma_y}\right)\sim N\left(0,\begin{pmatrix}1&\rho\\ \rho&1\end{pmatrix}\right)$

דוגמא: $X,Z\sim N(0,1)$ ב"ת, $Y=\rho X+\sqrt{1-\rho^2}\cdot Z$ אז: $Y\sim N(0,\rho^2+1-\rho^2)=N(0,1)$

$$Cov(X,Y)=E(XY)-E(X)E(Y)=E\left(X\left(\rho X+\sqrt{1-\rho^2}Z\right)\right)-0=\rho+\sqrt{1-\rho^2}\underbrace{E(XZ)}_{E(X)\cdot E(Z)=0}=\rho$$

$$E(Y\mid X)==\rho E(X\mid X)+\sqrt{1-\rho^2}E(Z\mid X)=\rho X+0$$

סימולציה של $N(\mu,\sigma^2)$ בהינתן $Uni[0,1]$:

אם $X\sim N(0,1)$ אז $Y=\mu+\sigma X\sim N(\mu,\sigma^2)$

פיתרון-1: $\Phi(x)=P(N(0,1)\le x)$.

פיתרון-2: $(X,Y)\sim N(0,I)$, $Y=R\sin\theta$ $X=R\cos\theta$ $\theta=\arctan\left(\frac{Y}{X}\right)$ $R=\sqrt{X^2+Y^2}$.

$r=$ יעקוביאן. ונקבל: $f_{R\theta}(r,\theta)=\begin{cases}\frac{1}{2\pi}re^{-\frac{1}{2}r^2} & 0<r,\ \ 0\le\theta\le 2\pi\\ 0 & otherwise\end{cases}$

$$\frac{1}{2\pi}\cdot re^{-\frac{1}{2}r^2}=f_\theta(\theta)\cdot f_R(r)\ \Rightarrow\ \theta\sim Uni[0,2\pi],\quad R\sim re^{-\frac{1}{2}r^2}$$

$$F_R(r)=\int_0^r re^{-\frac{1}{2}r^2}=1-e^{-\frac{1}{2}r^2},\quad F^{-1}(v)=\sqrt{-2\ln(1-v)},\quad V\sim Uni[0,1]$$

יצירת 2 מ"מ נורמאלים בהינתן מ"מ אוניפורמיים:

יהיו $U,V\sim Uni[0,1]$. ב"ת אז:

$$U\to 2\pi U=\theta\sim Uni[0,2\pi]$$
$$V\to\sqrt{-2\ln 1-V}=R\sim f_R(r)$$

המ"מ החדשים שיצרנו (θ,R) מקיימים: $(R\cos\theta,R\sin\theta)=(X,Y)\sim N(0,1)$ וב"ת.