הפקולטה להנדסה
אוניברסיטת תל אביב

אותות אקראיים ורעש

0512.3632

ע״פ סיכום הרצאות מסמסטר א׳ שנת תשס״ח

מרצים:

ד"ר אורי ארז

פרופ' רם זמיר

רישום: יגאל רג'ואן

מהדורה 1.3

עריכה אחרונה 27.10.2011

תוכן עניינים

מבוא

דוגמא 1: נניח שאנו מאזינים לדובר #1, בנוכחות דובר #2 (נסמן את אותות הדיבור שלהם ב- $x_2(t), x_1(t)$ בהתאמה).

שני אותות אלו הם אקראיים מבחינתו. אנו לא יודעים מראש מה הם יגידו בהמשך.

מכיוון שאנו מעוניינים להקשיב רק לדובר #1, אות הדיבור של דובר #2 הוא אות אקראי לא רצוי. נקרא לו "רעש".

האות הנקלט באוזנינו הוא $y(t) = \alpha \cdot x_1(t) + \beta \cdot x_2(t)$

המוח האנושי מסוגל "לסנן" את האות שלא רצוי ו"להיפטר" מהרעש. נרצה לבנות מערכת הנדסית שתחקה פעולה זו.

תחילה, נרצה לאפיין את האותות. האם הם דומים או שונים מבחינה ספקטרלית? (דהיינו, האם התדרים של האותות נמצאים בתחומים שונים, או בערך באותו תחום?)

כיצד נוכל לעשות זאת? נרצה לדעת כיצד לנתח אותות אקראיים בתחום התדר.

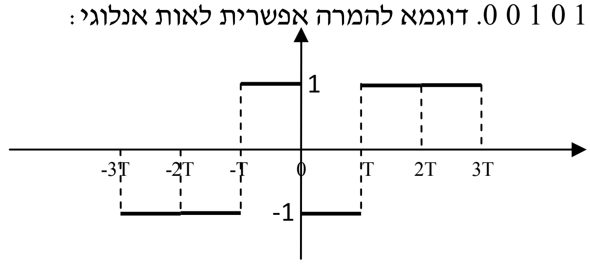
<u>מה ניתן לעשות עם מערכת לינארית, למשל?</u>

אנו יודעים מהי תוצאת מערכת לינארית כאשר מפעילים אותה על אות דטרמיניסטי.
לכן יש לדעת למדל אותות אקראיים, וכן לדעת את הקשר בין הקלט והפלט של המערכת הלינארית עבורם.

דוגמא 2: <u>העברת אינפורמציה</u> ממקום א' למקום ב'.
האינפורמציה היא תהליך אקראי (קריא, סדרה אינסופית של משתנים אקראיים): $b_1 b_2 b_3 \ldots$ (למשל מתח בחוט טלפון).

מכיוון שהעברה של מידע יכולה להיות ע"י אות רציף בלבד (מתח), נרצה להמיר את התהליך מסדרה בדידה לאות רציף.

נניח שהתהליך הוא 1 1 0 1 0 0 0. דוגמא להמרה אפשרית לאות אנלוגי:



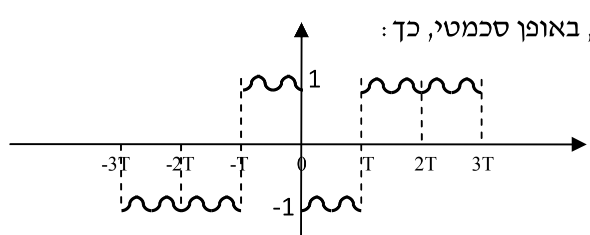
הנוסחא שמתארת את הקשר בין האות הבדיד לאות הרציף בטרנספורמציה היא:

$$x(t) = \sum_n (2b_n - 1) \cdot P(t - nT)$$

כאשר P(t) הוא פולס ברוחב T ובגובה 1, החל מזמן 0.

המודל הבסיסי ביותר לרעש במקלט הוא רעש גאוסי לבן - האות הנקלט הוא $y(t) = \alpha \cdot x(t) +$ $n(t)$, כאשר α קטן ככל שהמרחק אותו האות עובר גדל.

האות הנקלט ייראה, באופן סכמטי, כך:



המקלט שלנו יחליט מהו $\hat{b}_n$ לפי הסימן של ממוצע האות הנקלט על פני איטרוול.

איך נמדוד את טיב המקלט? נרצה ש- $\Pr\{\hat{b}_n \neq b_n\}$ יהיה קטן ככל האפשר. כיצד נמצא הסתברות זו?

חלק א: משתנים אקראיים ווקטורים אקראיים

חזרה על מושגים בסיסיים בתורת ההסתברות

<u>מרחב הסתברות</u>

מרחב הסתברות מוגדר ע״י השלשה $\{\Omega, \mathcal{F}, \mathrm{P}\}$:

- Ω (מרחב המדגם): אוסף כל התוצאות האפשריות בניסוי.
- $\mathcal{F}$ (שדה המאורעות): מאורע הינו תת קבוצה של Ω. שדה המאורעות הוא אוסף של תתי קבוצות של Ω.
  שדה המאורעות סגור תחת חיתוך ואיחוד סופיים או בני מניה ותחת משלים, משמע:
  אם $\forall i \in \mathbb{N}: A_i \in \mathcal{F}$ אזי גם $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ ו- $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.
  בפרט אם $A_1, A_2 \in \mathcal{F}$ אזי גם $A_1 \cup A_2 \in \mathcal{F}$ ו- $A_1 \cap A_2 \in \mathcal{F}$.
- P (פונקצית הסתברות): $P: \mathcal{F} \to [0,1]$ (כלומר, P הינה פונקציה הממפה מאורעות לערכים בין 0 ל-1).

על השלשה $\{\Omega, \mathcal{F}, \mathrm{P}\}$ לקיים את האקסיומות הבאות:

א. $\forall A \in \mathcal{F}, P(A) \geq 0$
ב. $P(\Omega) = 1$ (הסתברות המאורע הוודאי היא 1).
ג. אם $A_1, A_2, \dots$ הינם מאורעות זרים, כלומר $\forall i \neq j: A_i \cap A_j = \emptyset$, אזי $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

מתוך אקסיומות אלו נובעות בקלות התכונות הבאות:

א. $P(\emptyset) = 0$
ב. $\forall A \in F, P(A) \leq 1$
ג. $P(\bar{A}) = 1 - P(A)$

דוגמא 3: הטלת קובייה הוגנת

ישנן 6 תוצאות אפשריות בניסוי, לכן $\Omega = \{1,2,3,4,5,6\}$.

$\mathcal{F}$ יכול אוסף כל תתי הקבוצות של Ω (ישנן 2^6 תתי קבוצות).

מה לגבי P? לכאורה יש להגדיר את P עבור 64 המאורעות, אך בעצם הדבר מיותר.
ניתן להגדיר את P לפי ערכיה על המאורעות הבסיסיים בלבד. את ערכי P עבור המאורעות הנותרים נסיק ע״י האקסיומות של מרחב ההסתברות.

$$\mathrm{P}\{1\} = \mathrm{P}\{2\} = \mathrm{P}\{3\} = \mathrm{P}\{4\} = \mathrm{P}\{5\} = \mathrm{P}\{6\} = 1/6$$

כל מאורע אחר נוכל לקבל כאיחוד מאורעות זרים. למשל: מהו P(B), כאשר B={1,3,5}?

$$\mathrm{P(B)} = \mathrm{P}(\{1,3,5\}) = \mathrm{P}(\{1\}\cup\{3\}\cup\{5\}) = \mathrm{P}(\{1\}) + \mathrm{P}(\{3\}) + \mathrm{P}(\{5\}) = 1/2$$

דוגמא 4: הגרלה אחידה של נקודה בקטע [0,1]

מרחב המדגם במקרה זה הוא כל הנקודות ω הנמצאות בין 0 ל-1 : $\Omega = [0,1]$.

$\mathcal{F}$ יהיה אוסף כל הקטעים המוכלים ב- $[0,1]$ ואיחוד/חיתוך בן מניה של קטעים אלה.

ההסתברות ליפול על נקודה בודדת בקטע היא 0, לכן לא נוכל להגדיר את P ע״י המאורעות הבסיסיים.

P{[a,b]} = b-a עבור קטע, ובאופן דומה אורך כולל (מידה) עבור מאורע כללי.

משתנה אקראי

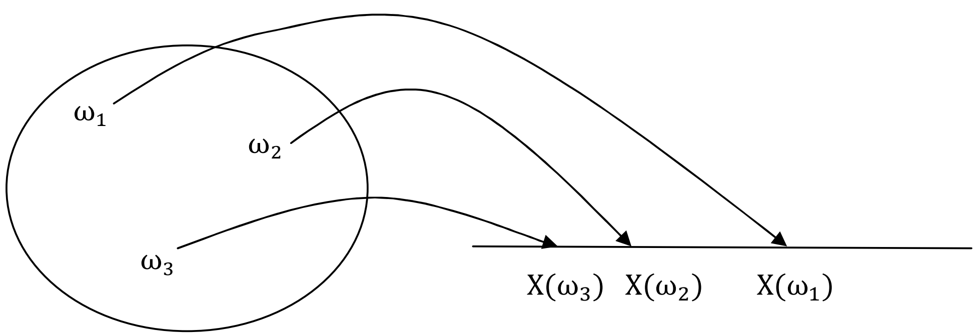
מכיוון שאיברי Ω הינם מופשטים, עבור כל ניסוי הקבוצה Ω בעלת אופי שונה (בהטלת מטבע למשל, $\Omega = \{\text{פלי, עץ}\}$. בהטלת קובייה, $\Omega = \{1,2,3,4,5,6\}$). לכן נרצה למפות את איברי Ω לשפה אחידה : הציר הממשי.

הגדרה: <u>משתנה אקראי</u> X הינו פונקציה המשייכת לכל תוצאת ניסוי ω ∈ Ω מספר ממשי X(ω), כך שהקבוצה $\{\omega : X(\omega) \le x\}$ הינה מאורע, $\forall x \in \mathbb{R}$.

משהגדרנו את X, השם של האיבר (צבע, צורה, מספר סידורי) לא רלוונטי לניתוח ההסתברות.

<u>הערות :</u>

1. מ״א יסומן באות גדולה.
2. הדרישה שהקבוצה $\{\omega : X(\omega) \le x\}$ הינה מאורע, $\forall x \in \mathbb{R}$. הינה לרוב תנאי טכני שיתקיים אם נבחר מיפוי ״טבעי״.



דוגמא 5: הגדרת מ״א בהטלת קובייה

בהטלת קובייה נוכל להגדיר $X(\omega) = \begin{cases} 1 & \omega = 1,3,5 \\ 2 & \omega = 2,4,6 \end{cases}$

כלומר 1 משמע תוצאת הטלה אי זוגית, ו-2 תוצאת הטלה זוגית.

פונקצית התפלגות מצטברת, פונקצית צפיפות ופונקציה אופיינית

עבור משתנה אקראי X ניתן לייצג את חוק ההסתברות בכמה אופנים:

1. פונקצית פילוג מצטבר Cumulative Distribution Function (CDF): $F_X(x) = P(X \le x)$

תכונות של CDF:

(1) $0 \le F_X(x) \le 1$
(2) $\lim_{x\to\infty} F_X(x) = 1$
(3) $\lim_{x\to-\infty} F_X(x) = 0$
(4) $F_X(x_0) = \lim_{\substack{x\to x_0 \\ x>x_0}} F_X(x)$ (רציפות מימין)
(5) $\forall x_2 > x_1:\ F_X(x_2) \ge F_X(x_1)$
(6) $\Pr(a < X \le b) = F_X(b) - F_X(a)$

2. פונקצית צפיפות הסתברות Probability Density Function (PDF)

אם $F_X(x)$ גזירה בנקודה x אזי נגדיר: $f_X(x) = \frac{\partial F_X}{\partial x} = \lim_{\Delta x\to 0} \frac{F_X(x+\Delta x)-F_X(x)}{\Delta x}$

תכונות של PDF:

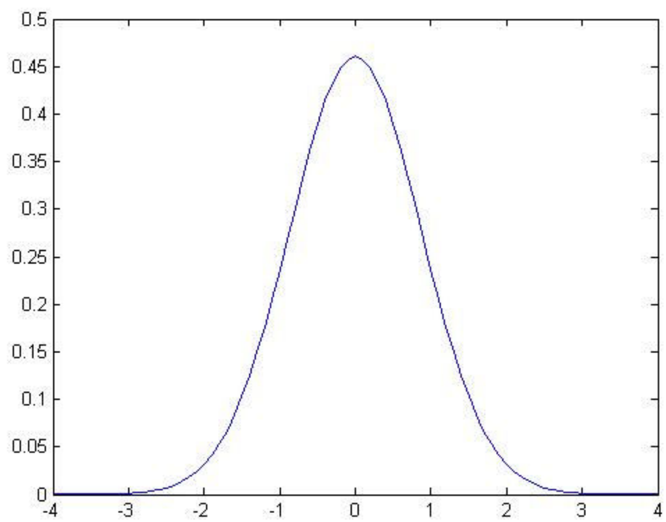
(1) $F_X(x) = \int_{-\infty}^{x} f_X(x)dx$
(2) $\Pr(a \le X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx$
(3) $\forall x: f_X(x) \ge 0$
(4) $\int_{-\infty}^{\infty} f_X(x)dx = 1$

הערה: נשים לב שייתכן $f_X(x) > 1$ (PDF אינה מתארת הסתברות אלא את צפיפות ההסתברות סביב נקודה מסוימת).

מסקנה: כל עוד $f_X(x)$ סופית, ההסתברות לקבל ערך בדיד היא 0.

הגדרה: נאמר ש- X הוא משתנה אקראי רציף אם $F_X(x)$ היא פונקציה רציפה לכל $x \in \mathbb{R}$.

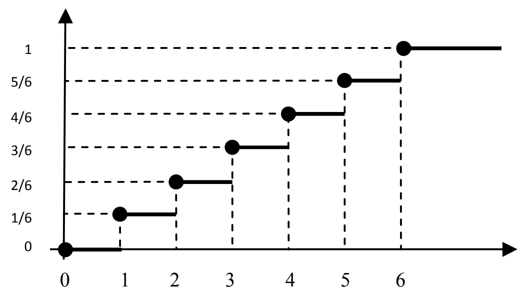
דוגמא 6: פונקצית PDF של משתנה נורמלי:



הגדרה: נאמר ש- X הוא משתנה אקראי בדיד אם $F_X(x)$ היא פונקצית מדרגות.

דוגמא 7: פונקצית הCDF בהטלת קובייה הוגנת:

$$F_X(x) = \frac{1}{6}[U(x-1)+U(x-2)+U(x-3)+U(x-4)+U(x-5)+U(x-6)]$$



את פונקצית ה- PDF של המשתנה האקראי הנ"ל נוכל לכתוב כסכום של הלמים:

$$f_X(x) = \frac{1}{6}[\delta(x-1)+\delta(x-2)+\delta(x-3)+\delta(x-4)+\delta(x-5)+\delta(x-6)]$$

הגדרה: תוחלת של משתנה אקראי: $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$

עבור משתנה אקראי בדיד נוכל לרשום גם: $\eta_x = E[X] = \sum_{k=1}^{\infty} x_k \Pr(X = x_k)$. מכאן שלתוחלת משמעות של "מרכז המסה של הפילוג".

הגדרה: מומנט מסדר n של משתנה אקראי: $m_n = E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)dx$

הגדרה: מומנט מרכזי מסדר n של משתנה אקראי:

$$\mu_n = E[(X-\eta_x)^n] = \int_{-\infty}^{\infty}(x-\eta_x)^n f_X(x)dx$$

הגדרה: שונות של משתנה אקראי: $Var[X] = \sigma_x^2 = \mu_2 = E[(X-\eta_x)^2] = E[X^2] - E^2[X]$

הערה: מומנט ומומנט מרכזי של משתנה אקראי (ובפרט תוחלת ושונות) מוגדרים רק במקרה שבו האינטגרל מוגדר היטב.

הערה: ידיעת מספר סופי של מומנטים (לרוב) נותנת מידע חלקי בלבד על ההתפלגות

הגדרה: פונקציה אופיינית של משתנה אקראי X היא התמרת פורייה ההפוכה של ה- PDF של X:

$$\Phi_X(\omega) = E\left[e^{j\omega X}\right] = \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)dx$$

תכונות של פונקציה אופיינית

(1) $f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-j\omega x}\Phi_X(\omega)d\omega$

(2) $\Phi_X(0) = \int_{-\infty}^{\infty} f_X(x)dx = 1$

(3) $|\Phi_X(\omega)| = \left|\int_{-\infty}^{\infty} e^{j\omega x} f_X(x)dx\right| \le \int_{-\infty}^{\infty} \underbrace{\left|e^{j\omega x}\right|}_{=1} f_X(x)dx = 1$

(4) $\left.\frac{d^n\Phi_X(\omega)}{d\omega^n}\right|_{\omega=0} = j^n \cdot m_n$

$$\left.\frac{d^n\Phi_X(\omega)}{d\omega^n}\right|_{\omega=0} = \left.\int_{-\infty}^{\infty}(jx)^n e^{j\omega x} f_X(x)dx\right|_{\omega=0} = j^n \cdot \int_{-\infty}^{\infty} x^n f_X(x)dx = j^n m_n$$

(5) $Y = aX + b \Longrightarrow \Phi_Y(\omega) = e^{j\omega b}\cdot\Phi_X(a\omega)$

תזכורת- תוחלת של פונקציה של מ״א: יהי Y מ״א שהינו פונקציה של מ״א X:
Y=g(X), אזי:

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

הוכחת תכונה מספר 5:

$$\Phi_Y(\omega) = E\left[e^{j\omega Y}\right] = E\left[e^{j\omega(aX+b)}\right] = e^{j\omega b} \cdot E\left[e^{j\omega aX}\right] = e^{j\omega b} \cdot \Phi_X(a\omega)$$

הגדרה: פונקציה יוצרת מומנטים של משתנה אקראי X

$$M_X(a) = E[e^{aX}] = \int_{-\infty}^{\infty} e^{ax} f_X(x)dx \qquad \Longrightarrow \left.\frac{d^n M_X(a)}{da^n}\right|_{a=0} = E[X^n]$$

הערות:

1. שימוש חשוב של פונקציה יוצרת מומנטים הוא חסם Chernoff (ראו תרגול).
2. ידיעת הפונקציה האופיינית מתארת התפלגות של מ״א באופן מלא, ובכך שקולה לידיעת ה-CDF וה- PDF שלו. האם הכרת $\{m_n\}_{n=0}^{\infty}$ שקולה לכל אלה?

פיתוח Taylor של פונקציה אופיינית

$$\Phi_X(\omega) =_? \Phi_X(\omega = 0) + \left.\frac{d\Phi_X(\omega)}{d\omega}\right|_{\omega=0} \cdot \omega + \cdots + \frac{1}{n!} \cdot \left.\frac{\mathrm{d}^n \Phi_X(\omega)}{d\omega^n}\right|_{\omega=0} \cdot \omega^n + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot \left.\frac{\mathrm{d}^n \Phi_X(\omega)}{d\omega^n}\right|_{\omega=0} \cdot \omega^n$$

הערה: $\Phi_X(\omega)$ אינה תמיד פונקציה אנליטית ולכן פיתוח טיילור לא תמיד מתכנס ל- $\Phi_X(\omega)$.
קיימים בספרות תנאים מספיקים להתכנסות הטור.

וקטורים אקראיים

הגדרה: וקטור אקראי $\underline{X}$ הוא מיפוי מ- Ω ל- $\mathbb{R}^n$: $\underline{X}: \Omega \longrightarrow \mathbb{R}^n$ המקיים ש :
$\{\omega: X_1(\omega) \le x_1, X_2(\omega) \le x_2, \dots\}$
הוא מאורע לכל $\underline{x} \in \mathbb{R}^n$.

כמו במשתנה אקראי יחיד, ניתן לתאר את המידע ההסתברותי על $\underline{X}$ (וקטור אקראי) ע"י :
1. (Joint) CDF
2. (Joint) PDF
3. פונקציה יוצרת מומנטים (משותפת)

הגדרה: <u>CDF של ו"א</u> : $F_{\underline{X}}(\underline{x}) = F_{\underline{X}}(x_1, x_2, \dots, x_n) = \Pr(X_1 \le x_1, \dots, X_n \le x_n)$

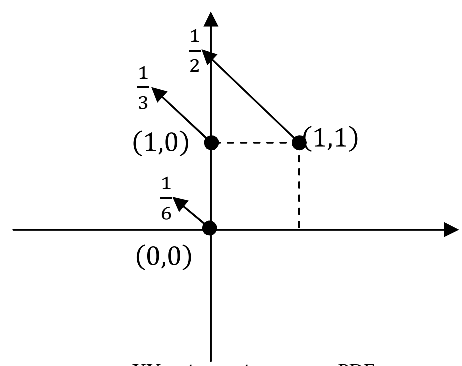
דוגמא 1: הגדרת ו"א בניסוי הטלת קובייה באמצעות CDF

נגדיר משתנה אקראי X שיציין האם תוצאת ההטלה היא מספר זוגי.
נגדיר משתנה אקראי Y שיציין האם תוצאת ההטלה גדולה או שווה ל-2.

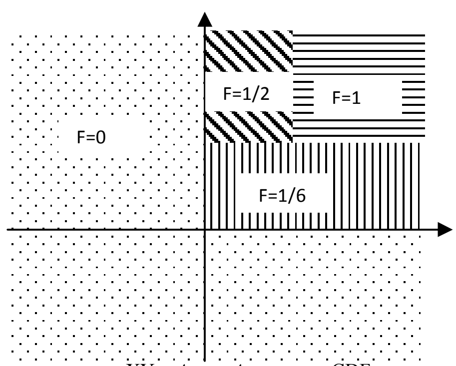
נגדיר את הוקטור האקראי : (X, Y)

$$X = \begin{cases} 1 & \omega = 2,4,6 \\ 0 & \omega = 1,3,5 \end{cases} \qquad Y = \begin{cases} 1 & \omega \ge 2 \\ 0 & \omega = 1 \end{cases}$$

$$Pr(X = 0, Y = 0) = \frac{1}{6} \qquad Pr(X = 1, Y = 1) = \frac{1}{2}$$
$$Pr(X = 1, Y = 0) = 0 \qquad Pr(X = 0, Y = 1) = \frac{1}{3}$$



ה- PDF המשותף של ההתפלגות XY



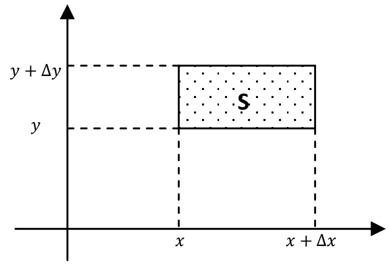
ה- CDF המשותף של ההתפלגות XY

נוכל לרשום את ה- CDF במקרה זה באופן הבא :
$$F_{XY}(x,y) = \frac{1}{6} \cdot U(x-0) \cdot U(y-0) + \frac{1}{3} \cdot U(x-0) \cdot U(y-1) + \frac{1}{2} \cdot U(x-1) \cdot U(y-1)$$

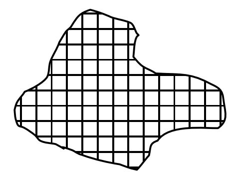
<u>תכונות של CDF של ו"א</u>:

1. $0 \le F_{\underline{X}}(\underline{x}) \le 1$
2. $F_{\underline{X}}(\infty, \infty, \dots, \infty) = 1$
3. $\forall 1 \le k \le n: \ F_{\underline{X}}(x_1, \dots, x_{k-1}, -\infty, x_{k+1}, \dots, x_n) = 0$
   הסבר : זהו בעצם מאורע שדורש $X_k \le -\infty$, לכן הסתברותו היא 0.
4. $\forall 1 \le k \le n: \ F_{\underline{X}}(\infty, \dots, \infty, x_k, \infty, \dots, \infty) = \Pr(X_k \le x_k) = F_{X_k}(x_k)$
5. $F_{\underline{X}}(\underline{x})$ היא פונקציה מונוטונית לא יורדת בכל משתנה.

איך מחשבים הסתברות של מלבן?



$$\Pr\left(\underline{X}\in S\right)=F_{XY}(x+\Delta x,y+\Delta y) - F_{XY}(x,y+\Delta y) - F_{XY}(x+\Delta x,y) + F_{XY}(x,y)$$



ואם התחום S אינו מלבן?

נחלק אותו להמון מלבנים קטנים, שאת ההסתברות של כל-אחד מהם בנפרד אנחנו יודעים לחשב, ואז נסכום.

הגדרה: PDF של ו"א:

$$f_{XY}(x,y)=\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\left(\frac{1}{\Delta x\Delta y}\cdot \Pr\left(x\le X\le x+\Delta x, y\le Y\le y+\Delta y\right)\right)=$$

$$=\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\left(\frac{1}{\Delta x\Delta y}\cdot\left(F_{XY}(x+\Delta x,y+\Delta y)-F_{XY}(x,y+\Delta y)-[F_{XY}(x+\Delta x,y)-F_{XY}(x,y)]\right)\right)$$

$$=\frac{\partial^2 F_{XY}(x,y)}{\partial x\partial y}$$

דוגמא 2: הגדרת ו"א בניסוי הטלת קובייה באמצעות PDF

נשתמש בפונקצית ה- CDF שהגדרנו בדוגמא 1. את הגזירה מבצעים תחילה לפי משתנה אחד, ואז לפי המשתנה השני.

$$f_{XY}(x,y)=\frac{1}{6}\cdot\delta(x)\cdot\delta(y)+\frac{1}{3}\cdot\delta(x)\cdot\delta(y-1)+\frac{1}{2}\cdot\delta(x-1)\cdot\delta(y-1)$$

תכונות של PDF של ו"א:

1. חישוב הסתברות של תחום כלשהו (לאו דווקא מלבני):

   מקרה דו-מימדי:

   $$\Pr\left(\underline{X}\in S\subset\mathbb{R}^2\right)=\iint_{(x,y)\in S} f_{XY}(x,y)dxdy$$

   מקרה n-מימדי:

   $$\Pr\left(\underline{X}\in S\subset\mathbb{R}^n\right)=\iint_{\underline{x}\in S} f_{\underline{X}}\left(\underline{x}\right)d\underline{x}$$

2. פילוג שולי

   מקרה דו-מימדי:

   $$f_X(x)=\int_{-\infty}^{\infty} f_{XY}(x,y)dy$$

במקרה ה n-מימדי $\underline{X}=(x_1,x_2,\dots,x_n)$ נרצה לדעת מהי $f_{X_{i_1},\dots,X_{i_k}}(x_{i_1},\dots,x_{i_k})$ (כאשר $1\le i_1<\dots<i_k\le n$). נסמן את שאר המשתנים ב-$u_1,\dots,u_{n-k}$. אזי:

$$f_{X_{i_1},\dots,X_{i_k}}(x_{i_1},\dots,x_{i_k})=\iiint_{-\infty}^{\infty} f_{X_1,\dots,X_n}(x_1,\dots,x_n)du_1\dots du_{n-k}$$

כלומר, אנו עושים אינטגרציה מ- $(-\infty)$ עד ∞ על כל המשתנים מהם נרצה "להיפטר".

דוגמא 3: ו"א גאוסי דו-מימדי

במקרה הדו-מימדי ("התפלגות דו-נורמלית"), משפחת התפלגות זו מאופיינת ע"י 5 פרמטרים:

η_x, η_y - התוחלות של X ושל Y בהתאמה.

σ_x^2, σ_y^2 - השונויות של X ושל Y בהתאמה.

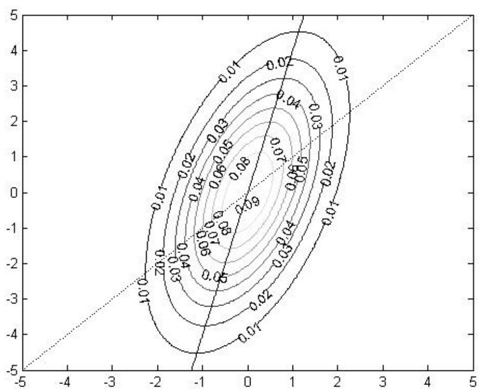
ρ - מקדם המתאם (דרישה: $-1 \le \rho \le 1$).

פונקצית ה- PDF המשותפת (עבור $|\rho| \neq 1$):

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \cdot exp\left\{-\frac{1}{2(1-\rho^2)} \cdot \left[\frac{(x-\eta_x)^2}{\sigma_x^2} - \frac{2\rho(x-\eta_x)(y-\eta_y)}{\sigma_x\sigma_y} + \frac{(y-\eta_y)^2}{\sigma_y^2}\right]\right\}$$

<u>טענה:</u> הפילוגים השוליים בהתפלגות דו-נורמלית הם נורמליים $X\sim N(\eta_x, \sigma_x^2)$, $Y\sim N(\eta_y, \sigma_y^2)$.

(הערה: מידיעת $f_X(x)$, $f_Y(y)$ לא ניתן לקבוע את $f_{XY}(x,y)$).

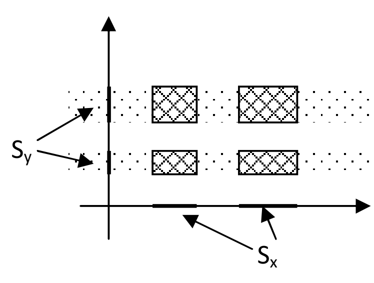


נשים לב שהמקום הגיאומטרי של כל הנקודות המקיימות $f_{XY}(x,y) = C$ הוא אליפסה שמרכזה (η_x, η_y).

בציור משמאל, $\rho = 0.5$, התוחלות הן 0, $\sigma_x = 1, \sigma_y = 2$

תזכורת: <u>פילוג מותנה</u>: אם A,B מאורעות ו- $\Pr(B) \neq 0$, אזי $\Pr(A|B) \triangleq \frac{\Pr(A\cap B)}{\Pr(B)}$

נניח (X,Y) וקטור אקראי. נגדיר שני מאורעות $X \in S_X, Y \in S_Y$.



$$\Pr(X \in S_X | Y \in S_Y) = \frac{\Pr(X\in S_X, Y\in S_Y)}{\Pr(Y\in S_Y)} = \frac{\int_{X\in S_X}\int_{Y\in S_Y} f_{XY}(x,y)dxdy}{\int_{-\infty}^{\infty}\int_{Y\in S_Y} f_{XY}(x,y)dxdy} = \frac{\int_{X\in S_X}\int_{Y\in S_Y} f_{XY}(x,y)dxdy}{\int_{Y\in S_Y} f_Y(y)dy}$$

<u>התניה נקודתית</u>

לעיתים נרצה להתנות מאורע מסוים במאורע נקודתי. נשים לב שעבור מ"א רציף, הסתברות המאורע הנקודתי היא אפס (למשל: $\Pr(X \in S_X | Y = y)$, כאשר Y הוא משתנה אקראי רציף). אזי נגדיר את ההסתברות ע"י גבול:

$$\Pr(X\in S_X|Y=y)$$
$$\triangleq \lim_{\Delta y\to 0}\frac{\Pr(X\in S_X, y\le Y\le y+\Delta y)}{\Pr(y\le Y\le y+\Delta y)} \underbrace{=}_{(1)} \lim_{\Delta y\to 0}\frac{\frac{1}{\Delta y}\cdot\int_{u\in S_X}\int_{y}^{y+\Delta y} f_{XY}(u,v)dvdu}{\frac{1}{\Delta y}\cdot\int_{y}^{y+\Delta y} f_Y(v)dv} \underbrace{=}_{(2)}$$
$$\frac{\lim_{\Delta y\to 0}\left[\frac{1}{\Delta y}\cdot\int_{y}^{y+\Delta y}\left(\int_{u\in S_X} f_{XY}(u,v)du\right)dv\right]}{\lim_{\Delta y\to 0}\left[\frac{1}{\Delta y}\cdot\int_{y}^{y+\Delta y} f_Y(v)dv\right]} \underbrace{=}_{(3)} \frac{\int_{u\in S_X} f_{XY}(u,y)du}{f_Y(y)}$$

הסבר המעברים:

(1) כותבים את ההסתברות ע"י ה- PDF, כופלים ומחלקים ב- $\frac{1}{\Delta y}$.

(2) משנים סדר אינטגרציה.

(3) f_{XY} היא פונקציה רציפה לכן בקטע Δy מאוד קטן נוכל להחליף את האינטגרל בכפל. הכפל ב- Δy מבוטל עם הכפל ב- $\frac{1}{\Delta y}$.

<u>הערה:</u> אם $f_Y(y)=0$ אזי ההתניה הנקודתית אינה מוגדרת.

הגדרה: <u>CDF מותנה</u>: $F_{X|Y}(x|y)\triangleq \Pr(X\le x|Y=y)=\frac{\int_{-\infty}^{x} f_{XY}(u,y)du}{f_Y(y)}$

הגדרה: <u>*PDF* מותנה</u>: $f_{X|Y}(x|Y=y)=\frac{\partial}{\partial x}F_{X|Y}(x|Y=y)=\frac{f_{XY}(x,y)}{f_Y(y)}$

<u>נוסחת *Bayes*</u>

נשים לב שמתקיים: $f_{XY}(x,y)=f_{X|Y}(x|y)\cdot f_Y(y)=f_{Y|X}(y|x)\cdot f_X(x)$. לכן:

$$f_{X|Y}(x|y)=\frac{f_{XY}(x,y)}{f_Y(y)}=\frac{f_X(x)\cdot f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_{XY}(u,y)du}=\boxed{\frac{f_X(x)\cdot f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(x)\cdot f_{Y|X}(y|x)dx}}$$

תזכורת: במקרה הבדיד:

$$\Pr(X=x_i|Y=y_i)=\frac{\overbrace{\Pr(Y=y_i|X=x_i)\cdot \Pr(X=x_i)}^{\Pr(X=x_i,Y=y_i)}}{\underbrace{\sum_m[\Pr(X=x_m)\cdot\Pr(Y=y_i|X=x_m)]}_{\substack{\text{נוסחת ההסתברות השלמה}\\ \Pr(Y=y_i)}}}$$

<u>אי תלות סטטיסטית</u>

נאמר ש- X,Y בלתי תלויים סטטיסטית (בת"ס) אם מתקיים שלכל מאורע $X\in S_X, Y\in S_Y$ מתקיים: $\Pr(X\in S_X, Y\in S_Y)=\Pr(X\in S_X)\cdot\Pr(Y\in S_Y)$.

מכאן נובע שאם X,Y בת"ס אזי:

1. $F_{XY}(x,y)=F_X(x)\cdot F_Y(y)$
2. $f_{XY}(x,y)=f_X(x)\cdot f_Y(y)$
3. $f_{X|Y}(x|y)=\frac{f_{XY}(x,y)}{f_Y(y)}=\frac{f_X(x)\cdot f_Y(y)}{f_Y(y)}=f_X(x)$

<u>טענה</u>: כ"א מהתנאים הנ"ל מהווה הגדרה שקולה לכך ש- X,Y בת"ס.

הכללה לוקטור אקראי n-מימדי: עבור ו"א $\underline{X} = (x_1, x_2, \dots, x_n)$ נאמר שרכיביו בת"ס אס"ם
$F_{\underline{X}}(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_n}(x_n)$

הגדרה: תוחלת מותנית: $\eta_{X|Y=y} \triangleq \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)dx$

הגדרה: שונות מותנית: $\sigma_{X|Y=y}^2 = \int_{-\infty}^{\infty} (x - \eta_{X|Y=y})^2 \cdot f_{X|Y}(x|y)dx$

טענות לגבי התפלגות דו-נורמלית

1. $\eta_{X|Y=y} = \eta_X + \rho \cdot \frac{\sigma_x}{\sigma_y} \cdot (y - \eta_y)$

$\eta_{Y|X=x} = \eta_Y + \rho \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - \eta_x)$

2. $\sigma_{X|Y=y}^2 = \sigma_x^2 \cdot (1 - \rho^2)$

$\sigma_{Y|X=x}^2 = \sigma_y^2 \cdot (1 - \rho^2)$

3. $X_{|Y=y} \sim N\left(\eta_{X|Y=y}, \sigma_{X|Y=y}^2\right)$

$Y_{|X=x} \sim N\left(\eta_{Y|X=x}, \sigma_{Y|X=x}^2\right)$

הערות:
1. נשים לב ש- $\sigma_{X|Y=y}^2$ קבוע לכל y, וכן $\sigma_{Y|X=x}^2$ קבוע לכל x.
2. כאשר $\rho \to 1$ השונויות המותנות מתקרבות ל-0. כאשר גאוסיאן מצטמצם ברוחבו, הוא הופך ל- δ. כלומר: $f_{XY}(x, y) = f_X(x) \cdot \delta(y - \eta_{Y|X=x})$ (למעשה, הפילוג המותנה של Y הופך ל- $\delta\left(y - \eta_{Y|X=x}\right)$).

האם הפילוג המשותף של 2 משתנים אקראיים גאוסיים הוא בהכרח גאוסי? לא.

דוגמא 4: משתנים גאוסיים שהפילוג המשותף שלהם אינו גאוסי:

יהי X משתנה אקראי גאוסי $X \sim N(0,1)$, ויהי B משתנה בת"ס ב- X המקבל את הערכים 1, -1 בהסתברויות שוות.

נגדיר את המשתנה האקראי Y באופן הבא: $Y = B \cdot X$. נראה כי $Y \sim N(0,1)$ וכי הפילוג המשותף של X ושל Y אינו נורמלי.

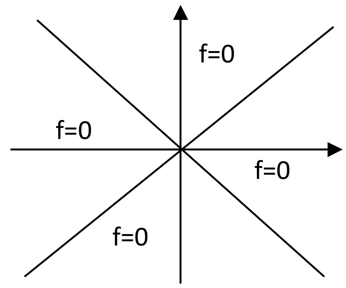
לפי נוסחת ההסתברות השלמה נוכל לתאר את ההתפלגות של Y באופן הבא:
$$F_Y(y) = F_{Y|B}(y|1) \cdot \Pr(B = 1) + F_{Y|B}(y|-1) \cdot \Pr(B = -1)$$

נגזור את המשוואה הנ"ל ונקבל:
$$f_Y(y) = f_{Y|B}(y|1) \cdot \Pr(B = 1) + f_{Y|B}(y|-1) \cdot \Pr(B = -1)$$
$$f_Y(y) = f_{Y|B}(y|1) \cdot \frac{1}{2} + f_{Y|B}(y|-1) \cdot \frac{1}{2}$$
$$f_Y(y) = \frac{1}{2} \cdot \left[f_{Y|B}(y|1) + f_{Y|B}(y|-1)\right] = \frac{1}{2} \cdot [f_X(y) + f_X(-y)] = \frac{1}{\sqrt{2\pi}} \cdot e^{-y^2/2}$$

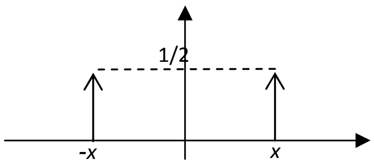
קיבלנו ש- $Y \sim N(0,1)$. כעת נותר להראות שהפילוג המשותף אינו נורמלי:
$$f_{XY}(x, y) = f_X(x) \cdot f_{Y|X}(y|X = x) \underbrace{=}_{(*)} f_X(x) \cdot \frac{1}{2}[\delta(y - x) + \delta(y + x)]$$

(*) בהינתן x, y יכול להיות $-x$ או x בהסתברויות שוות (0.5 כ"א).

על כל אחד מהאלכסונים הפילוג הוא נורמלי בגובה 0.5
בכל שאר התחומים הפילוג הוא 0.



זהו אינו פילוג גאוסי במשותף
כי הפילוג המותנה של y
בהינתן x הוא צמד הלמים!



ואילו בפילוג גאוסי משותף הפילוג המותנה הוא גאוסי. הלם הוא פילוג גאוסי אפשרי רק אם לגאוסיאן שונות אפס, אך בכל מקרה שני הלמים אינם אפשריים לפילוג גאוסי.

<u>פונקציה של וקטור אקראי</u>

נתונה ההתפלגות $f_{XY}(x,y)$. מגדירים את המשתנה האקראי Z באופן הבא:

$$Z:\Omega\to\mathbb{R},\qquad g:\mathbb{R}^2\to\mathbb{R},\qquad Z=g(X,\mathrm{Y})$$

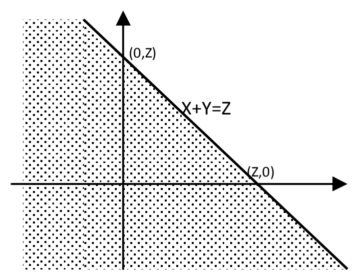
מהי ההתפלגות $F_Z(z)$?

$$F_Z(z)=\Pr(Z\le z)=\Pr(g(X,Y)\le z)=\Pr\left((X,Y)\in g^{-1}(-\infty,z]\right)$$

כאשר בהינתן $S\subset\mathbb{R}$ ו- $g:\mathbb{R}^2\to\mathbb{R}$, מגדירים: $g^{-1}(S)=\{(x,y):g(x,y)\in S\}$.
לכן: $g^{-1}\big((-\infty,z]\big)=\{(x,y):g(x,y)\le z\}$.

$$F_Z(z)=\iint_{(x,y)\in g^{-1}(-\infty,z]}f_{XY}(x,y)dxdy$$

דוגמא 1: $Z=X+Y$:



$$F_Z(z)=\Pr(Z\le z)=\Pr(X+Y\le z)$$
$$F_Z(z)=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{z-x}f_{XY}(x,y)dy\right]dx$$

$$f_Z(z)=\int_{-\infty}^{\infty}f_{XY}(x,z-x)dx$$

עבור X,Y בת״ס, נקבל:

$$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)\cdot f_Y(z-x)dx=f_X(x)*f_Y(y)\Longrightarrow$$
$Convolution$

<u>הערה</u>: אם עושים קונבולוציה בין הרבה מ״א שווי-פילוג ובת״ס, אזי מתכנסים לגאוסיאן.

<u>תכונות של (והערות על) תוחלת של ו״א ותוחלת של פונקציה של ו״א</u>

1. תכונת ההחלקה: נתון ו״א (X,Y), ורוצים לחשב את התוחלת של Y:

$$E[Y]=\int_{-\infty}^{\infty}y\cdot f_Y(y)dy=\int_{-\infty}^{\infty}y\cdot\left[\int_{-\infty}^{\infty}f_{XY}(x,y)dx\right]dy=$$
$$=\int_{-\infty}^{\infty}y\cdot\left[\int_{-\infty}^{\infty}f_X(x)\cdot f_{Y|X}(y|x)dx\right]dy$$
$$=\int_{-\infty}^{\infty}f_X(x)\cdot\left[\int_{-\infty}^{\infty}y\cdot f_{Y|X}(y|x)dy\right]dx=\int_{-\infty}^{\infty}f_X(x)\cdot E[Y|X=x]dx$$
$$=E\big[E[Y|X]\big]$$

2. משפט התוחלת (תוחלת של פונקציה של ו״א):
נניח $Z=g(X,\mathrm{Y})$.

$$E[Z]=\int_{-\infty}^{\infty}z\cdot f_Z(z)dz=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)\cdot f_{XY}(x,y)dxdy$$

ההוכחה זהה למקרה החד מימדי.
הכללה למקרה ה n-מימדי:

$$E\left[g\left(\underline{X}\right)\right] = \iiint_{-\infty}^{\infty} g\left(\underline{X}\right)\cdot f_{\underline{X}}\left(\underline{x}\right)d\underline{x}$$

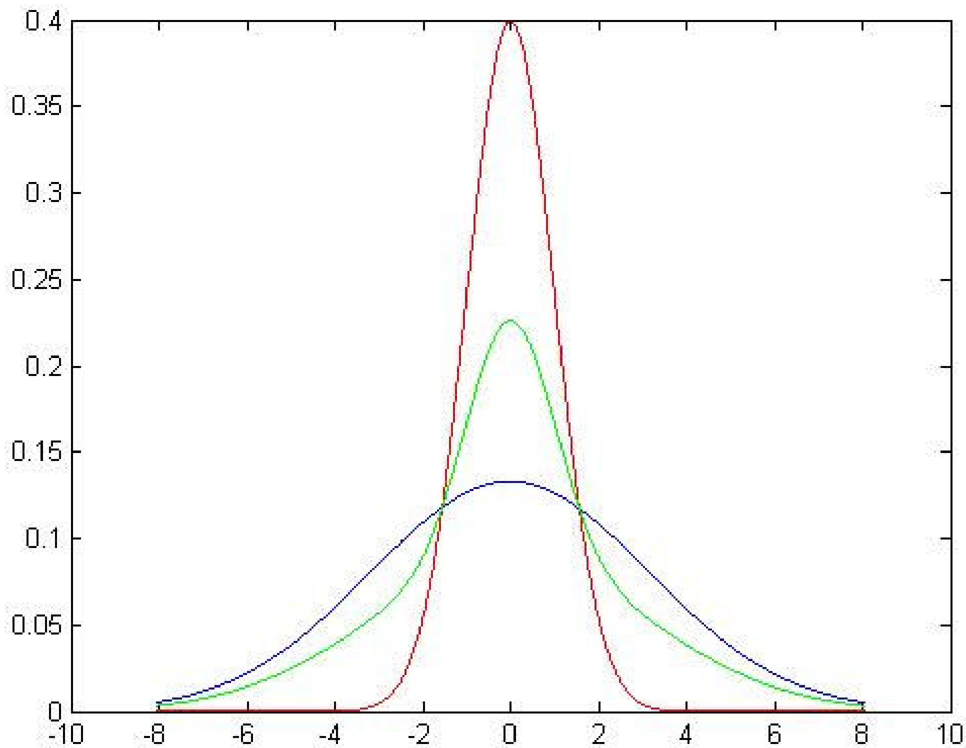
3. משפט ההחלקה עבור פונקציה של ו"א:

$$E[g(X,\mathrm{Y})] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(X,\mathrm{Y})\cdot f_{XY}(x,y)dxdy =$$
$$= \int_{-\infty}^{\infty} f_X(x)\cdot\left[\int_{-\infty}^{\infty} g(x,y)\cdot f_{Y|X}(y|x)dy\right]dx \underbrace{=}_{\text{סימון}} E_X[E_Y[g(X,\mathrm{Y})|X]]$$

דוגמא 2: חישוב מומנטים מסדר כלשהו של Gaussian Mixture:

יהי X מ"א עם הפילוג הבא ($0 \le a \le 1$):

$$f_X(x) = \frac{a}{\sqrt{2\pi s^2}}\cdot e^{-x^2/2s^2} + \frac{1-a}{\sqrt{2\pi\cdot 10s^2}}\cdot e^{-x^2/2\cdot 10s^2}$$



בתרשים הנ"ל, a=0.35, s=1.
הגרף הנמוך מתאים להתפלגות עם שונות 9, והגרף הגבוה להתפלגות עם שונות 1.
הגרף האמצעי מתאר את ה"ממוצע" המשוקלל של שתי ההתפלגויות.

נרצה לחשב את $E[X^n] = \int_{-\infty}^{\infty} x^n\cdot f_X(x)dx$.

ניתן לקבל משתנה $\hat{X}$ בעל פילוג זהה באופן הבא: נגדיר את המ"א הבאים (בת"ס אחד בשני):

$$X_1 \sim N(0,s^2), X_2 \sim N(0,10s^2), I = \begin{cases} 1 & w.p.\ a \\ 0 & w.p.\ 1-a \end{cases}$$

$$\hat{X} = I\cdot X_1 + (1-I)\cdot X_2$$

מתקיים: $f_X(x) = f_{\hat{X}}(x)$. נוכיח:

$$f_{\hat{X}}(x) = \Pr(I=0)\cdot f_{\hat{X}|I=0}(x|I=0) + \Pr(I=1)\cdot f_{\hat{X}|I=1}(x|I=1)$$
$$= (1-a)\cdot f_{X_2}(x) + a\cdot f_{X_1}(x) = f_X(x)$$

יצרנו משתנה אקראי $\hat{X}$ שהוא פונקציה של 3 משתנים אקראיים : $\hat{X} = g(I, X_1, X_2)$.

המסקנה היא ש- $E[X^n] = E[\hat{X}^n] = E[g^n(I, X_1, X_2)]$.

$$E[X^n] = E[\hat{X}^n] = E[g^n(I, X_1, X_2)] = E_I[E[g^n(I, X_1, X_2)]|I] =$$
$$= E_I[E[(I\cdot X_1 + (1-I)\cdot X_2)^n]|I=i] =$$
$$= \Pr(I=0)\cdot E_I[E[(I\cdot X_1 + (1-I)\cdot X_2)^n]|I=0]$$
$$+ \Pr(I=1)\cdot E_I[E[(I\cdot X_1 + (1-I)\cdot X_2)^n]|I=1]$$
$$= (1-\mathrm{a})\cdot E[{X_2}^n] + a\cdot E[{X_1}^n]$$

את התוחלות $E[{X_2}^n]$, $E[{X_1}^n]$ חישבנו בשיעורי הבית.

<u>הערה</u> : את תוצאה זו יכולנו לקבל ישירות גם ע"י חישוב התוחלת לפי הגדרה. יהיו מקרים בהם חישוב התוחלת לפי ההגדרה יהיה מסובך, ובשיטה זו נוכל להתגבר על כך.

<u>מומנטים משותפים</u>

בהינתן זוג משתנים אקראיים X,Y (ע"י פונקצית צפיפות משותפת) נגדיר :

הגדרה: <u>מומנט משותף מסדר n,k</u> :

$$m_{n,k} = E[X^n Y^k] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n y^k \cdot f_{XY}(x,y)dxdy$$

הגדרה: <u>מומנט מרכזי משותף מסדר n,k</u> :

$$\mu_{n,k} = E\left[(X-\eta_x)^n (Y-\eta_y)^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\eta_x)^n (y-\eta_y)^k \cdot f_{XY}(x,y)dxdy$$

דוגמא 3: מומנטים מסדרים שונים

$m_{n,0} = E[X^n]$ $m_{0,k} = E[Y^k]$

$\mu_{2,0} = \sigma_x^2$ $\mu_{0,2} = \sigma_y^2$

$m_{1,1} = E[XY] = R_{XY}$

<u>שונות משותפת (Covariance)</u> :

נגדיר את השונות המשותפת של X ו-Y באופן הבא :

$$Cov(X,Y) = \sigma_{xy} \triangleq \mu_{1,1} = E\left[(X-\eta_x)(Y-\eta_y)\right] = E[XY] - E[X]E[Y]$$

הגדרה: נאמר ש- X,Y <u>חסרי קורלציה</u> (חס"ק) אם $Cov(X,Y) = 0$.

<u>טענה</u> : אם X,Y בת"ס אזי הם חסרי קורלציה.

תזכורת: $Var(X) = E[(X-\eta_x)^2] = E[X^2] - \eta_x^2$

<u>טענה</u> : אם X,Y חס"ק, אזי $Var(X+Y) = Var(X) + Var(Y)$.

<u>הוכחה</u> :

$$Var(X+Y) = E\left[(X+Y-\eta_x-\eta_y)^2\right] = E\left[((X-\eta_x)+(Y-\eta_y))^2\right] =$$
$$= E[(X-\eta_x)^2] + 2\,E\left[(X-\eta_x)(Y-\eta_y)\right] + E\left[(Y-\eta_y)^2\right] =$$

$$= Var(X) + 2Cov(X,Y) + Var(Y) \underbrace{=}_{\text{חס"ק}} Var(X) + Var(Y)$$

מ.ש.ל.

הגדרה: $r = r(X,Y) = \frac{E[XY]}{\sqrt{E[X^2]\cdot E[Y^2]}}$. גודל זה מודד עד כמה ניתן לקרב את התלות של Y ב- X ע"י $Y = aX$ (תלות לינארית בין X ו- Y העוברת דרך ראשית הצירים).

הגדרה: מקדם הקורלציה של X ו-Y : $\rho = \rho(X,Y) = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$. מקדם הקורלציה מודד עד כמה ניתן לקרב את התלות של Y ב- X ע"י $Y = aX + b$ (תלות לינארית כלשהי בין X ו- Y)

הגדרה: משתנים אקראיים X,Y הם <u>אורתוגונליים</u> אם מתקיים $E[XY] = 0$.

<u>טענה</u> : אם $r(X,Y) = 0$ אזי המשתנים X,Y אורתוגונליים.

<u>טענה</u> : אם $\rho(X,Y) = 0$ אזי המשתנים X,Y חס"ק.

<u>טענה</u> : אם X,Y הם אורתוגונליים וידוע ש- $E[X] = 0$ או $E[Y] = 0$, אזי הם חס"ק.

<u>טענה</u> : אם X,Y חס"ק וידוע ש- $E[X] = 0$ או $E[Y] = 0$, אזי הם אורתוגונאליים.

<u>טענה</u> :

1. $|r| \leq 1$
2. $|\rho| \leq 1$

<u>הוכחה</u> : תחילה נוכיח ש- $E^2[XY] \leq E[X^2]E[Y^2]$ (טענה זו נקראת אי שוויון Cauchy-Schwartz).

<u>הוכחת אי שוויון Cauchy-Schwartz</u> :

נתבונן בפונקציה $X - \alpha Y$:

$$0 \underbrace{\leq}_{(*)} E[(X-\alpha Y)^2] = E[X^2] - 2\alpha E[XY] + \alpha^2 E[Y^2] \underbrace{=}_{\text{נגדיר}} g(\alpha)$$

הפונקציה $g(\alpha)$ היא פרבולה מחייכת אי שלילית, לכן בפתרון המשוואה הריבועית המתארת אותה, מתקיים $\Delta \leq 0$ (ישנו פתרון יחיד או שאין פתרונות למשוואה).

$$0 \geq \Delta = b^2 - 4ac = 4E^2[XY] - 4E[X^2]E[Y^2]$$
$$E^2[XY] \leq E[X^2]E[Y^2]$$

(*) אי השוויון מתקיים לכל α. נשים לב כי מתקבל שוויון אס"ם $X = \alpha Y$, כלומר ישנו קשר לינארי בין X ל- Y.

<u>נחזור להוכחת הטענה</u> : לפי אי שוויון Cauchy-Swartz,

1. $E^2[XY] \leq E[X^2]E[Y^2] \Rightarrow \frac{E^2[XY]}{E[X^2]E[Y^2]} \leq 1 \Rightarrow \left|\frac{E[XY]}{\sqrt{E[X^2]\cdot E[Y^2]}}\right| \leq 1$
2. $E^2\left[(X-\eta_x)(Y-\eta_y)\right] \leq E[(X-\eta_x)^2]\cdot E\left[\left(Y-\eta_y\right)^2\right]$

$$Cov^2(X,Y) \leq \sigma_x^2\cdot\sigma_y^2 \Rightarrow \frac{Cov^2(X,Y)}{\sigma_x^2\cdot\sigma_y^2} \leq 1 \Rightarrow \left|\frac{Cov(X,Y)}{\sigma_x\cdot\sigma_y}\right| \leq 1$$

מ.ש.ל.

מסקנות:

1. $E^2[XY]=E[X^2]E[Y^2]$ אס"ם קיימת α כך ש- $X=\alpha Y$ ($|r|=1$).
2. $E^2\left[(X-\eta_x)(Y-\eta_y)\right]=\sigma_x^2\cdot\sigma_y^2$ אס"ם קיימת α כך ש- $X-\eta_x=\alpha(Y-\eta_y)$ ($|\rho|=1$). במקרה זה נאמר שיש תלות לינארית בין X ל-Y.

ההבדל בין מקרה זה למקרה הקודם הוא שבמקרה הקודם הישר שמתאר את התלות הלינארית בין X ל-Y עובר דרך ראשית הצירים, וכאן לא בהכרח.

דוגמא 4: התפלגות דו-נורמלית $(X,Y)\sim N(\eta_x,\eta_y,\sigma_x^2,\sigma_y^2,\rho)$:

את השונות המשותפת של Xושל Y נוכל לחשב באופן הבא: $Cov(X,Y)=\rho\cdot\sigma_x\cdot\sigma_y$.

כזכור, התוחלת המותנית של X בהינתן Y היא $\eta_{X|Y=y}=\eta_X+\rho\cdot\frac{\sigma_x}{\sigma_y}\cdot(y-\eta_y)$. השונות המותנית היא $\sigma_{X|Y=y}^2=\sigma_x^2\cdot(1-\rho^2)$.

כאשר $\rho\to\pm1$, $\sigma_{X|Y=y}^2\to 0$, כלומר הפיזור של $X|Y=y$ הופך לערך יחיד. הערך הזה הוא $X_{|Y=y}=\eta_X\pm\frac{\sigma_x}{\sigma_y}\cdot(y-\eta_y)$

מומנטים מסדר כלשהו של ו"א

יהיה $\underline{X}=[X_1...X_n]$ ו"א. נגדיר את המומנט $\underline{k}_{1\times n}\in N^n$ באופן הבא:

$$m_{\underline{k}}=E\left\{\prod_{m=1}^{n}X_m^{k_m}\right\}$$

ואת המומנט המרכזי $\underline{k}_{1\times n}\in N^n$ באופן הבא:

$$\mu_{\underline{k}}=E\left\{\prod_{m=1}^{n}\left(X_m-E\{X_m\}\right)^{k_m}\right\}$$

סדר המומנט הוא סכום איברי הוקטור $\underline{k}_{1\times n}$.

פונקציה אופיינית משותפת

הגדרה: הפונקציה האופיינית המשותפת של ו"א $\underline{X}=\begin{bmatrix}X_1\\X_2\\\vdots\\X_n\end{bmatrix}$, היא $\Phi_{\underline{X}}(\underline{\omega})=E[e^{j\underline{\omega}^T\underline{X}}]$, כאשר $\underline{\omega}^T=(\omega_1,\omega_2,\dots,\omega_n)$. לפי משפט התוחלת:

$$\Phi_{\underline{X}}(\underline{\omega})=\iiint_{-\infty}^{\infty}e^{j\underline{\omega}^T\underline{x}}\cdot f_{\underline{X}}(\underline{x})d\underline{x}$$

תכונות של פונקציה אופיינית משותפת

1. $\left|\Phi_{\underline{X}}(\underline{\omega})\right|\le 1 \quad \forall\underline{\omega}\in\mathbb{R}^n$. בנוסף: $\left|\Phi_{\underline{X}}(\underline{0})\right|=1$ (כאשר $\underline{0}^T=(\underbrace{0,0,\dots,0}_{n\ times})$).

2. פונקציה אופיינית שולית: $\Phi_{X_k}(\omega_k)=\Phi_{\underline{X}}(0,\dots,0,\underbrace{\omega_k}_{k'th\ place},0,\dots,0)$

3. התמרת פורייה הפוכה:

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^n} \iint_{-\infty}^{\infty} \Phi_{\underline{X}}(\underline{\omega}) \cdot e^{-j\underline{\omega}^{\mathrm{T}}\underline{x}} d\underline{\omega}$$

4. חישוב מומנטים באמצעות הפונקציה האופיינית ($k = k_1 + k_2 + \cdots + k_n$):

$$m_{k_1,k_2,\dots,k_n} = E\left[X_1^{k_1} X_2^{k_2} \dots X_n^{k_n}\right] = \frac{1}{j^k} \cdot \left.\frac{\partial^k \Phi_{\underline{X}}(\underline{\omega})}{\partial \omega_1^{k_1} \omega_2^{k_2} \dots \omega_n^{k_n}}\right|_{\underline{\omega}=\underline{0}}$$

5. השפעת טרנספורמציה לינארית על הפונקציה האופיינית:

$$\underline{Y} = \mathrm{A}\underline{X} + \underline{b}, \quad \mathrm{A} \in \mathbb{R}^{n\times n}, \quad \underline{b} \in \mathbb{R}^n$$
$$\Phi_{\underline{Y}}(\underline{u}) = e^{j\underline{u}^T\underline{b}} \cdot \Phi_{\underline{X}}(A^T\underline{u})$$

6. אם $\underline{X}, \underline{Y}$ ו״א בת״ס אזי $\Phi_{\underline{XY}}(\underline{\omega}, \underline{u}) = \Phi_{\underline{X}}(\underline{\omega}) \cdot \Phi_{\underline{Y}}(\underline{u})$

<u>סטטיסטיקה מסדר שני של וקטורים אקראיים</u>

בהינתן ו״א $\underline{X} = \begin{bmatrix} \mathrm{X}_1 \\ \mathrm{X}_2 \\ \vdots \\ \mathrm{X}_n \end{bmatrix}$, נוכל לחשב את התוחלות, את השונויות ואת השונויות המשותפות של רכיביו באופן הבא:

$$\eta_{X_k} = E[X_k] = m_{0,\dots,0,\underbrace{1}_{\mathrm{k'th\ place}},0,\dots,0}$$
$$E[X_k X_l] = m_{0,\dots,0,\underbrace{1}_{\mathrm{k'th\ place}},0,\dots,0,\underbrace{1}_{\mathrm{l'th\ place}},0,\dots,0}$$
$$\sigma_{X_k X_l} = Cov(X_k, X_l) = E\left[(X_k - \eta_{X_k})(X_l - \eta_{X_l})\right] = \mu_{0,\dots,0,\underbrace{1}_{\mathrm{k'th\ place}},0,\dots,0,\underbrace{1}_{\mathrm{l'th\ place}},0,\dots,0}$$

הגדרה: <u>וקטור התוחלת</u> של $\underline{X}$ הוא $\eta_{\underline{X}} = \begin{bmatrix} \mathrm{E}[\mathrm{X}_1] \\ \mathrm{E}[\mathrm{X}_2] \\ \vdots \\ \mathrm{E}[\mathrm{X}_n] \end{bmatrix}$

הגדרה: <u>מטריצת הקורלציה</u> של $\underline{X}$ היא $R_{\underline{X}} = E\left[\underline{X} \cdot \underline{X}^{\mathrm{T}}\right] = \begin{bmatrix} E[X_1X_1] & \dots & E[X_1X_n] \\ \vdots & \ddots & \vdots \\ E[X_nX_1] & \dots & E[X_nX_n] \end{bmatrix}$

הגדרה: <u>מטריצת ה Covariance</u> של $\underline{X}$ היא

$$C_{\underline{X}} = E\left[(\underline{X} - \eta_{\underline{X}}) \cdot (\underline{X} - \eta_{\underline{X}})^{\mathrm{T}}\right] = \begin{bmatrix} \sigma_1^2 & \dots & \mathrm{Cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & \dots & \sigma_n^2 \end{bmatrix}$$

<u>תכונות של מטריצת הקורלציה ומטריצת ה Covariance</u>:

1. $\mathrm{C}_{\underline{X}} = \mathrm{R}_{\underline{X}} - \eta_{\underline{X}} \cdot \eta_{\underline{X}}^{\mathrm{T}}$ (לפי הקשר: $Cov(X, Y) = E[XY] - E[X]E[Y]$, ראה עמוד 19).

הגדרה: $\mathrm{A} \in \mathbb{R}^{n\times n}$ היא Positive Semi Definite (מוגדרת אי-שלילית) אם היא סימטרית וגם $\forall \underline{\mathrm{b}} \in \mathbb{R}^{\mathrm{n}}: \underline{\mathrm{b}}^{\mathrm{T}}\mathrm{A}\underline{\mathrm{b}} \geq 0$

תזכורת: אם $A \in \mathbb{R}^{n\times n}$ היא Positive Semi Definite אזי
i. הערכים העצמיים של A הם אי-שליליים.
ii. קיימת מטריצה אלכסונית Λ ומטריצה אורתונורמלית P (כלומר $P^T = P^{-1}$) כך ש:

$$\mathrm{P^TAP} = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

2. $\mathrm{R}_{\underline{X}}$ ו- $\mathrm{C}_{\underline{X}}$ הן מטריצות אי-שליליות מוגדרות (Positive Semi Definite), כלומר סימטריות בעלות ערכים עצמיים אי שליליים וניתנות ללכסון אורתונורמלי.

נוכיח ש $\mathrm{R}_{\underline{X}}$ אי שלילית מוגדרת.
אומנם $\mathrm{R}_{\underline{X}}$ סימטרית.
צ״ל: $\forall \underline{\mathrm{b}} \in \mathbb{R}^n: \underline{\mathrm{b}}^T\mathrm{R}_{\underline{X}}\underline{\mathrm{b}} \geq 0$.
יהי $\underline{\mathrm{b}} = \begin{bmatrix} \mathrm{b}_1 \\ \mathrm{b}_2 \\ \vdots \\ \mathrm{b}_n \end{bmatrix}$ כלשהו. נגדיר מ״א $Y = \underline{b}^T\underline{X}$.

$$0 \leq E[Y^2] = E[Y \cdot Y] = E[Y \cdot Y^T] = E[(\underline{b}^T\underline{X}) \cdot (\underline{b}^T\underline{X})^T] = E[\underline{b}^T\underline{X}\underline{X}^T\underline{b}] \underbrace{=}_{(*)}$$

$$\underline{b}^T E\left[\underline{X}\underline{X}^T\right]\underline{b} = \underline{b}^T \mathrm{R}_{\underline{X}}\, \underline{b}$$

(*) עבור מטריצה אקראית $\mathrm{Z}_{n\times m}$ ומטריצות דטרמיניסטיות $\mathrm{A} \in \mathbb{R}^{k\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{m\times l}$
מתקיים:
i. $\mathrm{E[AZ]} = AE[Z]$
ii. $\mathrm{E[ZC]} = E[Z]C$
iii. $\mathrm{E[Z+B]} = \mathrm{E[Z]} + \mathrm{B}$

ניתן להוכיח באותו אופן ש- $\mathrm{C}_{\underline{X}}$ היא Positive Semi Definite ע״י הגדרת $Y = \underline{b}^T(\underline{X} - \eta_{\underline{X}})$.
<u>הערות</u>:
1. אם $\mathrm{R}_{\underline{X}}$ אלכסונית נשים לב שרכיבי $\underline{X}$ אורתוגונליים הדדית.
2. אם $\mathrm{C}_{\underline{X}}$ אלכסונית נשים לב שרכיבי $\underline{X}$ חס״ק באופן הדדי.

הגדרה: <u>מטריצת הקרוס קורלציה</u> של ו״א $\underline{\mathrm{Y}} = \begin{bmatrix} \mathrm{Y}_1 \\ \mathrm{Y}_2 \\ \vdots \\ \mathrm{Y}_n \end{bmatrix}, \underline{X} = \begin{bmatrix} \mathrm{X}_1 \\ \mathrm{X}_2 \\ \vdots \\ \mathrm{X}_n \end{bmatrix}$ היא:

$$\mathrm{R}_{\underline{X},\underline{Y}} = E\left[\underline{X} \cdot \underline{\mathrm{Y}}^{\mathrm{T}}\right] = \begin{bmatrix} E[X_1Y_1] & \dots & E[X_1Y_n] \\ \vdots & \ddots & \vdots \\ E[X_nY_1] & \dots & E[X_nY_n] \end{bmatrix}$$

הגדרה: <u>מטריצת הקרוס Covariance</u> של ו״א $\underline{\mathrm{Y}} = \begin{bmatrix} \mathrm{Y}_1 \\ \mathrm{Y}_2 \\ \vdots \\ \mathrm{Y}_n \end{bmatrix}, \underline{X} = \begin{bmatrix} \mathrm{X}_1 \\ \mathrm{X}_2 \\ \vdots \\ \mathrm{X}_n \end{bmatrix}$ היא:

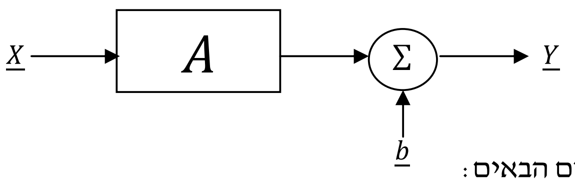
$$\mathrm{C}_{\underline{X},\underline{Y}} = E\left[(\underline{X} - \eta_{\underline{X}}) \cdot (\underline{\mathrm{Y}} - \eta_{\underline{\mathrm{Y}}})^{\mathrm{T}}\right] = \begin{bmatrix} Cov(X_1, Y_1) & \dots & Cov(X_1, Y_n) \\ \vdots & \ddots & \vdots \\ Cov(X_n, Y_1) & \dots & Cov(X_n, Y_n) \end{bmatrix}$$

<u>תכונות של מטריצות קרוס קורלציה וקרוס Covariance</u>:

1. $C_{\underline{X},\underline{X}} = C_{\underline{X}}$,$R_{\underline{X},\underline{X}} = R_{\underline{X}}$
2. $C_{\underline{Y},\underline{X}} = C^{T}_{\underline{X},\underline{Y}}$,$R_{\underline{Y},\underline{X}} = R^{T}_{\underline{X},\underline{Y}}$

מעבר וקטור אקראי דרך מערכת לינארית

יהיו $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$ ו״א, $\underline{b} \in \mathbb{R}^m$ וקטור דטרמיניסטי ו- $A \in \mathbb{R}^{m\times n}$ מטריצה דטרמיניסטית.



נגדיר את הו״א $\underline{Y} = A\underline{X} + \underline{b}$.

בהינתן $C_{\underline{X}}$ (או $R_{\underline{X}}$) ו- $\eta_{\underline{X}}$, נוכל לחשב את הגדלים הבאים:

1. $\eta_{\underline{Y}} = A\eta_{\underline{X}} + \underline{b}$

2. $C_{\underline{Y}} = E\left[(\underline{Y}-\eta_{\underline{Y}})\cdot(\underline{Y}-\eta_{\underline{Y}})^T\right] = E\left[(A\underline{X}+\underline{b}-A\eta_{\underline{X}}-\underline{b})\cdot(A\underline{X}+\underline{b}-A\eta_{\underline{X}}-\underline{b})^T\right]$
$= E\left[A(\underline{X}-\eta_{\underline{X}})\cdot(\underline{X}-\eta_{\underline{X}})^T A^T\right] = AC_{\underline{X}}A^T$

3. $R_{\underline{Y}} = E\left[(A\underline{X}+\underline{b})\cdot(A\underline{X}+\underline{b})^T\right] = AR_{\underline{X}}A^T + A\eta_{\underline{X}}\underline{b}^T + \underline{b}\eta_{\underline{X}}A^T + \underline{b}\cdot\underline{b}^T$

4. $R_{\underline{X},\underline{Y}} = E\left[\underline{X}\cdot(A\underline{X}+\underline{b})^T\right] = R_{\underline{X}}A^T + \eta_{\underline{X}}\underline{b}^T$
$R_{\underline{Y},\underline{X}} = E\left[(A\underline{X}+\underline{b})\cdot\underline{X}^T\right] = AR_{\underline{X}} + \underline{b}\eta^T_{\underline{X}}$

5. $C_{\underline{Y},\underline{X}} = E\left[(A\underline{X}+\underline{b}-A\eta_{\underline{X}}-\underline{b})\cdot(\underline{X}-\eta_{\underline{X}})^T\right] = E\left[A(\underline{X}-\eta_{\underline{X}})\cdot(\underline{X}-\eta_{\underline{X}})^T\right]$
$= AE\left[(\underline{X}-\eta_{\underline{X}})\cdot(\underline{X}-\eta_{\underline{X}})^T\right] = AC_{\underline{X}}$

וקטור גאוסי

בהרצאות הקודמות בחנו התפלגות דו-נורמלית, שהיא בעצם מקרה פרטי של וקטור גאוסי עבור n=2.

הגדרה: יהי $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$ ו״א. $\underline{X}$ ייקרא <u>וקטור גאוסי</u> אם $\forall\underline{a} \in \mathbb{R}^n$ מתקיים שהמשתנה האקראי $< \underline{X}, \underline{a} >= \underline{a}^T\underline{X} = \sum_{i=1}^{n} a_i X_i$ הוא משתנה אקראי גאוסי.

<u>מסקנות מתוך ההגדרה</u>:

1. אם $\underline{X}$ הוא וקטור גאוסי אזי כל אחד מרכיביו הוא משתנה גאוסי.

הוכחה: נובע ישירות מההגדרה ע״י בחירת $\underline{a} = \underline{e_i}$ (וקטורי הבסיס הסטנדרטי: כל רכיביו הם 0 פרט לאינדקס אחד שבו מופיע הערך 1).

2. אם $\underline{X}$ וקטור גאוסי ממימד n, $A \in \mathbb{R}^{m\times n}$ ו- $\underline{b} \in \mathbb{R}^m$ אזי $\underline{Y} = A\underline{X} + \underline{b}$ הוא וקטור גאוסי ממימד m:

$\underline{a}^T\underline{Y}=\underline{a}^TA\underline{X}+\underline{a}^T\mathrm{b}\underbrace{=}_{\substack{\underline{c}^T\triangleq\underline{a}^TA\\ \underline{c}\in\mathbb{R}^n}}\underbrace{\underline{c}^T\underline{X}}_{\substack{\text{משתנה גאוסי}\\ \text{לפי הגדרה}}}+\underbrace{\underline{a}^T\mathrm{b}}_{\text{קבוע}}=$ משתנה גאוסי

<u>טענה</u>: אם $\underline{X}$ וקטור גאוסי בעל תוחלת $\eta_{\underline{X}}$ ומטריצת שונות משותפת $C_{\underline{X}}$, אזי

$$\Phi_{\underline{X}}(\underline{\omega})=e^{j\eta_{\underline{X}}{}^T\underline{\omega}-\frac{1}{2}\underline{\omega}^TC_{\underline{X}}\underline{\omega}}$$

<u>הוכחה</u>: נתבונן במשתנה האקראי $Y=\underline{\omega}^T\underline{X}$. מכיוון ש- $\underline{X}$ הוא וקטור גאוסי, ידוע ש- Y הוא משתנה אקראי גאוסי. הראנו בשיעורי הבית שהפונקציה האופיינית של משתנה גאוסי היא:

$$\Phi_Y(u)=E\left[e^{juY}\right]=e^{ju\eta_Y-\frac{1}{2}u^2\sigma_Y^2}$$

הקשר בין הפונקציות האופייניות של $\underline{X}$ ושל Y הוא:

$$\Phi_{\underline{X}}(\underline{\omega})=E\left[e^{j\underline{\omega}^T\underline{X}}\right]=E\left[e^{jY}\right]=E\left[e^{j\cdot1\cdot Y}\right]=\Phi_Y(1)=e^{j\eta_Y-\frac{1}{2}\sigma_Y^2}$$

לכן נותר לנו להביע η_Y ואת σ_Y^2 במונחים של $\eta_{\underline{X}}$ ו- $C_{\underline{X}}$. את זה עשינו בהרצאה 4:

$$\eta_Y=\underline{\omega}^T\eta_{\underline{X}}$$
$$\sigma_Y^2=\underline{\omega}^TC_{\underline{X}}\underline{\omega}$$

ומכאן ש-

$$\Phi_{\underline{X}}(\underline{\omega})=e^{j\underline{\omega}^T\eta_{\underline{X}}-\frac{1}{2}\underline{\omega}^TC_{\underline{X}}\underline{\omega}}$$

מ.ש.ל

כעת, נניח שידוע ש $\underline{X}\sim N(\eta_{\underline{X}},C_{\underline{X}})$, ו- $C_{\underline{X}}$ הפיכה. אזי ע"י התמרת פורייה הפוכה נקבל:

$$\boxed{f_{\underline{X}}(\underline{x})=\frac{1}{\sqrt{(2\pi)^n|C_{\underline{X}}|}}\cdot e^{-\frac{1}{2}(\underline{X}-\eta_{\underline{X}})^TC_{\underline{X}}^{-1}(\underline{X}-\eta_{\underline{X}})}}$$

<u>הערות</u>:

1. נשים לב שעבור n=2 הנוסחא המתקבלת מתלכדת עם הנוסחא שראינו בעבר להתפלגות דו-נורמלית.
2. אם הוקטור $\underline{X}$ הוא וקטור גאוסי אזי רכיביו בת"ס אס"ם מטריצת השונות אלכסונית: $C_{\underline{X}}=\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \ddots & & \vdots\\ \vdots & & \ddots & 0\\ 0 & \dots & 0 & \lambda_n\end{bmatrix}$:

$$f_{\underline{X}}(\underline{x})=\frac{1}{\sqrt{(2\pi)^n\lambda_1\cdot\ldots\cdot\lambda_n}}\cdot e^{-\frac{1}{2}\sum_{i=1}^n\frac{(X_i-\eta_i)^2}{2\lambda_i}}$$

3. $C_{\underline{X}}^{-1}$ היא אי שלילית מוגדרת מאחר שהערכים העצמיים שלה הינם $1/\lambda_i$:

$$C_{\underline{X}}=P\Lambda P^T\Longrightarrow C_{\underline{X}}^{-1}=P\Lambda^{-1}P^T$$

מכיוון ש $\forall1\le i\le n:\lambda_i>0$, מתקיים גם $\forall1\le i\le n:1/\lambda_i>0$.

4. המשוואה $f_{\underline{X}}(\underline{x})=c$ מתארת אליפסואיד n מימדי (ראה תרשימים בהמשך).

<u>איך מייצרים וקטור גאוסי?</u>

נרצה ליצר וקטור גאוסי $\underline{X}$ ממימד n בעל תוחלת $\eta_{\underline{X}}$ ומטריצת שונות משותפת $C_{\underline{X}}$.

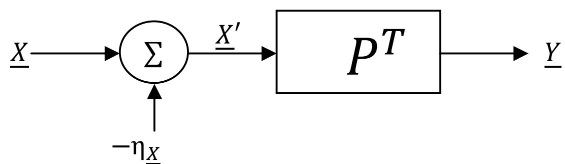
נפתח את הדיון במקרה הפוך: נניח שיש בידינו וקטור גאוסי $\underline{X}$ בעל תוחלת $\eta_{\underline{X}}$ ומטריצת שונות משותפת $C_{\underline{X}}$, וברצוננו ליצר וקטור גאוסי בעל תוחלת $\underline{0}$, שרכיביו בת"ס זה בזה.

תחילה נפחית מ-$\underline{X}$ את וקטור התוחלת שלו. נקבל וקטור גאוסי $\underline{X}'$ שתוחלתו $\underline{0}$, ומטריצת השונות המשותפת שלו היא $C_{\underline{X}}$.

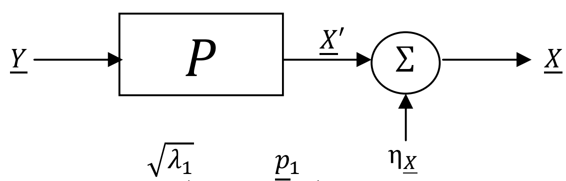
מכיוון ש- $C_{\underline{X}}$ היא אי שלילית מוגדרת, אנו יודעים שקיימות מטריצה אורתונורמלית P ומטריצה אלכסונית Λ, כך ש- $\Lambda = P^T C_{\underline{X}} P$.

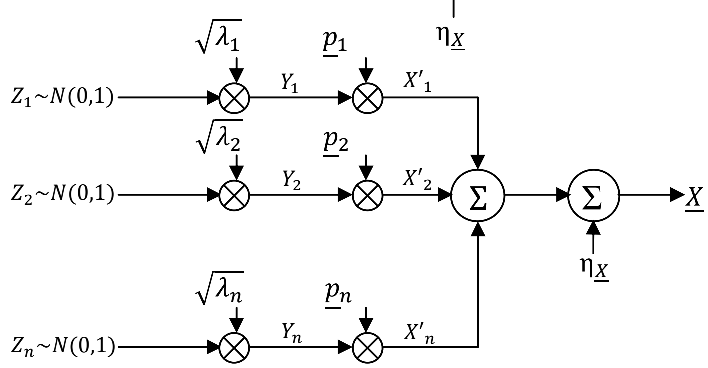
לכן, נפעיל על הקטור החדש שקיבלנו את הטרנספורמציה P^T. נסמן את הוקטור שהתקבל ב- $\underline{Y}$.
אזי מתקיים: $C_{\underline{Y}} = P^T C_{\underline{X}} P = \Lambda$.

מכיוון ש- $C_{\underline{Y}}$ אלכסונית, אנו יודעים שרכיבי הוקטור $\underline{Y}$ בת"ס, כפי שדרשנו מלכתחילה.



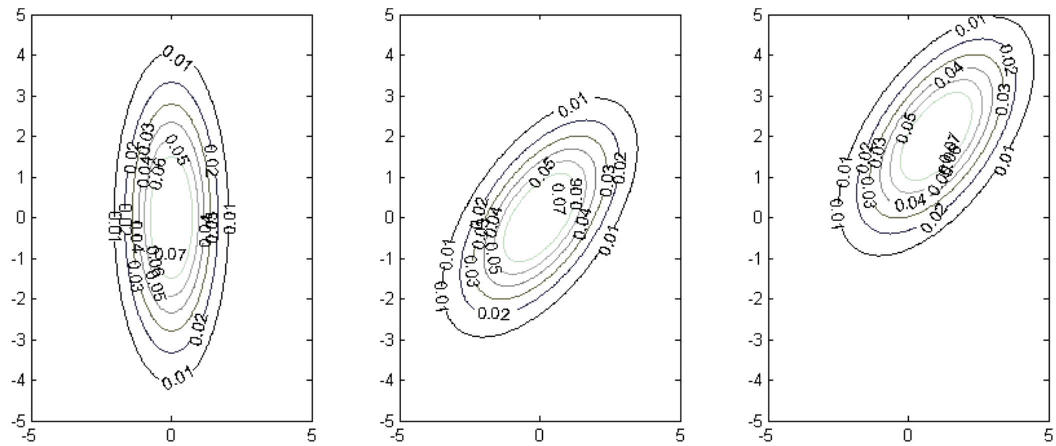
לכן על מנת לייצר וקטור $\underline{X}$ בעל תוחלת $\eta_{\underline{X}}$ ומטריצת שונות משותפת $C_{\underline{X}}$, ניצור וקטור גאוסי $\underline{Y}$ שרכיביו בת"ס זה בזה, ונעביר אותו דרך המערכת ההפוכה למערכת שלעיל:





במקרה הדו מימדי, שלבי הטרנספורמציה ייראו כך:

$\underline{Y}$ $\qquad$ $\underline{X}'$ $\qquad$ $\underline{X}$



<u>הסבר</u>: הוקטור $\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ הוא וקטור גאוסי שרכיביו בת"ס ומתפלגים באופן הבא:

$$Y_1 \sim N(0,1), Y_2 \sim N(0,2)$$

על מנת לקבל את הוקטור $\underline{X}'$, השתמשנו בטרנספורמציה הלינארית הבאה:

$$\underline{X}' = \underbrace{\begin{bmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}}_{P} \cdot \underline{Y}$$

ומכאן שמטריצת השונות של $\underline{X}'$ היא:

$$C_{\underline{X}'} = P \cdot C_{\underline{Y}} \cdot P^T = \begin{bmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} \frac{41}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{34}{25} \end{bmatrix}$$

התוחלת של $\underline{X}'$ היא כשל $\underline{Y}$: $\eta_{\underline{X}'} = \eta_{\underline{Y}} = \underline{0}$

את הוקטור $\underline{X}$ קיבלנו ע"י הוספת וקטור קבועים ל- $\underline{X}'$:

$$\underline{X} = \underline{X}' + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Longrightarrow \eta_{\underline{X}} = \eta_{\underline{X}'} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

מטריצת השונות של $\underline{X}$ שווה למטריצת השונות של $\underline{X}'$.

דוגמא 1: <u>וקטור אקראי</u>

במטוס ישנם 100 נשים ו-200 גברים. נסמן:

- X_i - משקל האישה ה- i
- Y_i - משקל הגבר ה- i

ידוע שמשקלי הנשים מתפלגים $i.i.d$ עם תוחלת η_X ושונות σ_x^2, ומשקלי הגברים מתפלגים $i.i.d$ עם תוחלת η_Y ושונות σ_y^2. משקלי הנשים ומשקלי הגברים בת"ס.

נגדיר:

- $X_T = \sum_{i=1}^{100} X_i$ - משקל הנשים הכולל.
- $Y_T = \sum_{i=1}^{200} Y_i$ - משקל הגברים הכולל.
- $W = X_T + Y_T$ - המשקל הכולל של נוסעי המטוס.

מכיוון שמשקלי הנשים וכן משקלי הגברים מתפלגים $i.i.d$ (ומספרם גדול), נוכל לקרב את ההתפלגות של X_T ואת ההתפלגות של Y_T ע"י התפלגות נורמלית:

$$X_T \sim N(100\eta_X, 100\sigma_x^2), \quad Y_T \sim N\left(200\eta_Y, 200\sigma_y^2\right)$$

כעת, W גם הוא מתפלג נורמלית, כסכום של שני משתנים נורמליים בת"ס.

$$W \sim N\left(100\eta_X + 200\eta_Y, 100\sigma_x^2 + 200\sigma_y^2\right)$$

הוקטור $\begin{bmatrix} X_T \\ Y_T \\ W \end{bmatrix}$ הוא וקטור גאוסי. למה?

כאמור, המשתנים X_T, Y_T הם משתנים גאוסיים בת"ס, לפי משפט הגבול המרכזי. אזי הוקטור $\begin{bmatrix} X_T \\ Y_T \end{bmatrix}$ הוא וקטור גאוסי. נקבל את הוקטור $\begin{bmatrix} X_T \\ Y_T \\ W \end{bmatrix}$ כטרנספורמציה לינארית על הוקטור $\begin{bmatrix} X_T \\ Y_T \end{bmatrix}$, ומכאן שהוא וקטור גאוסי בעצמו:

$$\begin{bmatrix} X_T \\ Y_T \\ W \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}}_{A} \cdot \begin{bmatrix} X_T \\ Y_T \end{bmatrix}$$

נתבונן כעת במקרה פרטי של ההתפלגות: $X_T \sim N(0,1), \quad Y_T \sim N(0,1)$. אזי

$$C_{X_T Y_T} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$C_{X_T Y_T W} = A \cdot C_{X_T Y_T} \cdot A^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

נשים לב ששורותיה (ועמודותיה) של המטריצה $C_{X_T Y_T W}$ תלויות לינארית. יכולנו לצפות עובדה זו מראש, מפני שיצרנו וקטור גאוסי בעל 3 רכיבים מוקטור גאוסי בעל 2 רכיבים. באופן כללי, אם ננסה ליצור וקטור גאוסי בעל n רכיבים מוקטור גאוסי בעל $k<n$ רכיבים, אזי נקבל רכיבים מנוונים בוקטור שיצרנו (כלומר רכיבים שיש ביניהם תלות דטרמיניסטית).

מכיוון שישנה תלות לינארית בין שורותיה של המטריצה $C_{X_T Y_T W}$, היא אינה הפיכה, ולא נוכל להשתמש בנוסחא שתוארה בתחילת השיעור על מנת לחשב את $f_{X_T Y_T W}(x, y, w)$.

לכן נחשב את פונקצית הצפיפות המשותפת ע"י פילוגים מותנים:

$$f_{X_T Y_T W}(x, y, w) = f_{X_T}(x) \cdot f_{Y_T|X_T}(y|x) \cdot f_{W|Y_T, X_T}(w|x, y)$$

נשים לב ש:

1. Y_T ו- X_T בת"ס, לכן $f_{Y_T|X_T}(y|x) = f_{Y_T}(y)$.
2. ההתפלגויות של X_T ושל Y_T הן נורמליות סטנדרטיות.
3. בהינתן X_T ו- Y_T, ערכו של W קבוע: $W = X_T + Y_T$, כלומר פונקצית ההתפלגות המותנית של W היא:

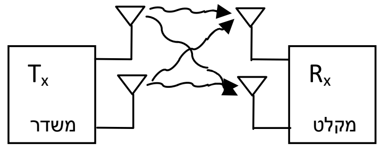
$$f_{W|Y_T, X_T}(w|x, y) = \delta(w - (x + y))$$

מכאן ש:

$$f_{X_T Y_T W}(x, y, w) = \left[\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}\right] \cdot \left[\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}}\right] \cdot \delta(w - (x + y)) =$$

$$= \left[\frac{1}{2\pi} \cdot e^{-\frac{x^2 + y^2}{2}}\right] \cdot \delta(w - (x + y))$$

דוגמא 2: <u>מערכת MIMO (Multiple Input Multiple Output)</u>



תיאור הבעיה: ברשותנו משדר בעל שתי אנטנות ומקלט בעל שתי אנטנות. כל אנטנה של המקלט קולטת סיגנלים המשודרים מכל-אחת מהאנטנות של המשדר.

כלומר, האות הנקלט בכל-אחת מהאנטנות של המקלט הוא צירוף לינארי של האותות המשודרים מהמשדר, בנוסף לרעש.

לכן נוכל לייצג את הוקטור $\underline{Y}$ באופן הבא: $\underline{Y} = H \cdot \underline{X} + \underline{Z}$, כאשר H היא טרנספורמציה לינארית $H \in \mathbb{R}^{2\times 2}$ והוקטור $\underline{Z}$ הוא הרעש.

$$Y_1 = h_{11} \cdot X_1 + h_{12} \cdot X_2 + Z_1$$
$$Y_2 = h_{21} \cdot X_1 + h_{22} \cdot X_2 + Z_2$$

נניח ש- $X_i = \begin{cases} 1 & w.p. \frac{1}{2} \\ -1 & w.p. \frac{1}{2} \end{cases}$, ושׁ- $Z_i \sim N(0, \sigma_z^2)$. כמו-כן X_1, X_2, Z_1, Z_2 בת״ס.

אם H אלכסונית, אזי המשמעות ההנדסית היא שכל אנטנה במקלט קולטת אות המשודר מאנטנה אחרת של המשדר, ונוכל לשערך את האות המשודר ע״פ הערך הנקלט באופן הבא:

$$Y_1 = h_{11} \cdot X_1 + Z_1 \Rightarrow \tilde{Y}_1 = Y_1/h_{11} \Rightarrow \hat{X}_1 = sign(\tilde{Y}_1)$$
$$Y_2 = h_{22} \cdot X_2 + Z_2 \Rightarrow \tilde{Y}_2 = Y_2/h_{22} \Rightarrow \hat{X}_2 = sign(\tilde{Y}_2)$$

כעת נתבונן במקרה כללי יותר בו H היא מטריצה הפיכה, והמטריצה ההפוכה לה היא H^{-1}:

$$\underline{\tilde{Y}} = H^{-1} \cdot \underline{Y} = H^{-1} \cdot \left(H \cdot \underline{X} + \underline{Z}\right) = \underline{X} + \underbrace{H^{-1} \cdot \underline{Z}}_{\underline{\tilde{Z}}}$$

$$\tilde{Y}_1 = X_1 + \tilde{Z}_1$$
$$\tilde{Y}_2 = X_2 + \tilde{Z}_2$$

$$C_{\underline{\tilde{Z}}} = H^{-1} \cdot \begin{bmatrix} \sigma_z^2 & 0 \\ 0 & \sigma_z^2 \end{bmatrix} \cdot (H^{-1})^T$$

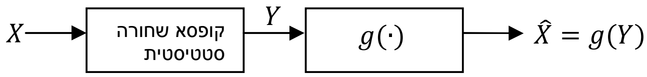
נשים לב ש-$\underline{\tilde{Z}}$ הוא וקטור גאוסי כטרנספורמציה לינארית של וקטור גאוסי, ורכיביו $\tilde{Z}_1, \tilde{Z}_2$ תלויים סטטיסטית.

ניתן לשערך את $\underline{X}$ באותה צורה כמו מקודם (רק שכעת הוקטור $\underline{\tilde{Y}}$ מוגדר אחרת):

$$\hat{X}_1 = sign(\tilde{Y}_1)$$
$$\hat{X}_2 = sign(\tilde{Y}_2)$$

נשים לב שמאורעות השגיאה: $\{\hat{X}_1 \neq X_1\}, \{\hat{X}_2 \neq X_2\}$ תלויים סטטיסטית. מדוע?

שערוך משתנים אקראיים



מטרה: בהינתן Y נרצה לשערך את X.
$\hat{X} = g(Y)$ נקרא המשערך של X.

הגדרה: <u>שגיאת שערוך</u>: $e = X - \hat{X} = X - g(y)$.

<u>הערה</u>: e הוא משתנה אקראי, למרות שאנו מסמנים אותו באות קטנה.

נרצה ש- e יהיה קטן ככל האפשר.

<u>מדדים לטיב השערוך</u>

<u>מינימום הסתברות שגיאה (עבור X מ"א בדיד בלבד)</u>

אם X בדיד, ניתן להגדיר משערך שיביא הסתברות שגיאה למינימום:
$$\Pr(g(Y) \neq X) = \Pr\left(\hat{X} \neq X\right) = \Pr(e \neq 0)$$
כלומר, הקריטריון הוא
$$g(\cdot) = \arg min_{g(\cdot)}\{\Pr(g(Y) \neq X)\}$$

<u>פתרון</u>: מקסימום אפוסטריורי (MAP)
נתבונן ב- $\Pr\,(X = x|Y = y)$ ונבחר את ה- x שמביא את ההסתברות המותנית למקסימום:
$$g(y) = \arg max_x\{\Pr\,(X = x|Y = y)\}$$
<u>הבהרה</u>: לפי הקשר $\Pr\left(\hat{X} \neq X\right) = 1 - \Pr\left(\hat{X} = X\right)$, כדי להביא למינימום את הסתברות השגיאה, עלינו למקסם את ההסתברות שהשערוך יהיה נכון, וזו בדיוק משמעות השוויון שלעיל.

<u>הערות</u>:
(1) נשים לב שבמדד שגיאה זה אין חשיבות לגודל השגיאה.
(2) במקרים רבים נרצה "לקנוס" במידה רבה יותר שגיאות גדולות.
(3) בפרט, אם X רציף, תמיד תהיה שגיאה ולכן למדד זה אין תוקף.

<u>שערוך אופטימלי במובן MMSE (Minimum Mean Square Error)</u>

נגדיר משערך אופטימלי במובן MSE, כלומר נדרוש שתוחלת ריבועי השגיאות תהיה מינימלית:
$$\mathrm{E}[e^2] = E\left[\left(X - \hat{X}\right)^2\right]$$

<u>פתרון</u>: משערך תוחלת מותנית

<u>טענה</u>:
(1) משערך ה- MMSE הוא $g_{opt}(y) = E[X|Y = y]$
(2) עבור $\hat{X} = g_{opt}(Y)$ מתקיים:
$$E\left[e_{opt}^2\right] = E[(X - E[X|Y])^2] \underbrace{=}_{\text{נוכיח}} E_Y[Var_X(X|Y)]$$

<u>הוכחה</u>:
(1) תהי $g(\cdot)$ כלשהי.

$$E\left[\left(X-g(Y)\right)^2\right] \underbrace{=}_{\text{החלקה}} E\left[E\left[\left(X-g(Y)\right)^2|Y\right]\right] =$$
$$= \int_{-\infty}^{\infty} f_Y(y)\, E\left[\left(X-g(Y)\right)^2|Y=y\right] dy =$$
$$= \int_{-\infty}^{\infty} f_Y(y) \underbrace{\left\{\int_{-\infty}^{\infty}\left(x-g(y)\right)^2 f_{X|Y}(x|y)dx\right\}}_{h(y)} dy$$

נרצה להביא את תוצאת האינטגרל למינימום. האיבר היחיד באינטגרל עליו יש לנו שליטה (כלומר, נוכל לקבוע אותו כרצוננו, ואינו נובע מההתפלגות של x ושל y) הוא $\left(x-g(y)\right)^2$. אם לכל y נעשה את הניחוש הכי טוב האפשרי, אזי נקבל את התוצאה הקטנה ביותר האפשרית לאינטגרל.

לכל y ננחש את הערך $\alpha = g(y)$ כך ש- $h(y)$ יהיה מינימלי. נמצא את הנקודה בה מתקבל המינימום ואת ערך המינימום באמצעות נגזרת:

$$q(\alpha) = \int_{-\infty}^{\infty}(x-\alpha)^2 f_{X|Y}(x|y)dx$$
$$q'(\alpha) = -2\int_{-\infty}^{\infty}(x-\alpha) f_{X|Y}(x|y)dx =$$
$$= -2\underbrace{\int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx}_{E[X|Y]} + 2\alpha \underbrace{\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx}_{1} = 0$$

ומכאן ש: $\alpha = \boxed{E[X|Y] = g(Y)}$.

(2) נציב את הערך שקיבלנו ל α בביטוי שקיבלנו ל $E\left[\left(X-g(Y)\right)^2\right]$, כדי למצוא את תוחלת ריבוע שגיאת השערוך האופטימלי:

$$E\left[e_{opt}^2\right] = E\left[\left(X-g(Y)\right)^2\right] =$$
$$\int_{-\infty}^{\infty} f_Y(y) \underbrace{\left\{\int_{-\infty}^{\infty}(x-E[X|Y=y])^2 f_{X|Y}(x|y)dx\right\}}_{\text{נסמן:}Var[X|Y=y]} dy \;= E[Var[X|Y]]$$

<u>**שערוך לינארי אופטימלי במובן MSE**</u>

נגדיר משערך לינארי להיות פונקציה לינארית של Y: $\hat{X} = aY + b$.

משערך לינארי אופטימלי BLE הינו $(a,b) = arg_{a,b}\min E[(X-aY-b)^2]$

<u>פתרון</u>:

$$a_{BLE} = \frac{\sigma_{xy}}{\sigma_y^2}$$
$$b_{BLE} = \eta_X - \frac{\sigma_{xy}}{\sigma_y^2}\cdot\eta_Y$$

<u>טענה</u>:

(1) $\hat{X}_{BLE} = \eta_X + \frac{\sigma_{xy}}{\sigma_y^2}\cdot(Y-\eta_Y)$

(2) $E[e_{BLE}^2] = \sigma_x^2\cdot(1-\rho^2)$

<u>הוכחה</u>: שגיאת השערוך היא:

$$e = X - \hat{X} = X - aY - b$$
$$\eta_e = \eta_X - a\eta_Y - b$$

נשים לב ש:

$$E[e^2] = Var[e] + \eta_e^2$$
$$Var[e] = E\left[\left((X-\eta_X) - a(Y-\eta_Y)\right)^2\right]$$

כלומר $Var[e]$ לא תלוי ב- b. לכן, כדי להביא את $E[e^2]$ למינימום לפי b, עלינו להביא את η_e^2 למינימום לפי b. זה יקרה כאשר $b=\eta_X-a\eta_Y$ ואז נקבל $\eta_e^2=0$.

כלומר לכל a, כדי להגיע לתוחלת שגיאה ריבועית מינימלית אפשרית, נבחר $b=\eta_X-a\eta_Y$.
במקרה זה, תוחלת השגיאה הריבועית תהיה (כפונקציה של a):

$$q(a)=E[e^2]\underbrace{=}_{\eta_e^2=0}Var[e]=E\left[\left((X-\eta_X)-a(Y-\eta_Y)\right)^2\right]=$$
$$=E[(X-\eta_X)^2]-2aE[(X-\eta_X)(Y-\eta_Y)]+a^2E[(Y-\eta_Y)^2]=$$
$$=\sigma_x^2-2a\sigma_{xy}+a^2\sigma_y^2$$

כעת נחשב עבור איזה a הביטוי הנ"ל מקבל ערך מינימלי:

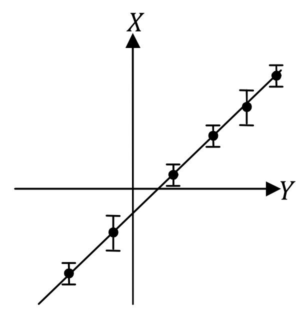
$$q'(a)=-2\sigma_{xy}+2a\cdot\sigma_y^2=0\Longrightarrow a=\frac{\sigma_{xy}}{\sigma_y^2}=\rho\frac{\sigma_x}{\sigma_y}$$

נציב את התוצאה שהתקבלה חזרה ב-$q(a)$ כדי למצוא את תוחלת השגיאה הריבועית המינימלית:

$$E[e_{BLE}^2]=\sigma_x^2-2\left(\rho\frac{\sigma_x}{\sigma_y}\right)\sigma_{xy}+\rho^2\frac{\sigma_x^2}{\sigma_y^2}\sigma_y^2=\sigma_x^2-2\left(\rho\frac{\sigma_x}{\sigma_y}\cdot\frac{\sigma_x}{\sigma_x}\right)\sigma_{xy}+\rho^2\sigma_x^2$$
$$E[e_{BLE}^2]=\sigma_x^2(1-\rho^2)$$

$$\hat{X}_{BLE}=\eta_X+\rho\frac{\sigma_x}{\sigma_y}(Y-\eta_Y)$$

נשים לב שככל ש- ρ מתקרב ל- 1, השגיאה שלנו קטנה ולכן השערוך של X כפונקציה לינארית של Y הוא יותר מדויק (ρ מתקרב ל- 1$\Longleftarrow$ X מתקרב סטטיסטית להיות פונקציה לינארית של Y)



תכונות של משערך MMSE לינארי / אופטימלי

(1) חוסר הטיה: תוחלת המשערך שווה לתוחלת המשוערך:

$$E[e_{BLE}]=0,\ E\left[\hat{X}_{BLE}\right]=\eta_X$$
$$E\left[e_{opt}\right]=0,\ E\left[\hat{X}_{opt}\right]=\eta_X$$

(2) ניצבות:

א. $\hat{X}=aY+b, e=X-aY-B$ הינו BLE אס"ם

$$\forall c,d\in\mathbb{R}:\ E[e\cdot(cY+d)]=0$$

כלומר, $\forall c,d\in\mathbb{R}:\ e\perp cY+d$.
נשים לב שמתקיים: $E[eY]=0$.

הוכחה:

($\Longleftarrow$) ידוע ש- $\hat{X}=aY+b$ הינו BLE. לכן, $\hat{X}_{BLE}=\eta_X+\frac{\sigma_{xy}}{\sigma_y^2}\cdot(Y-\eta_Y)$.

$$E[e_{BLE}\cdot(cY+d)]=cE[e_{BLE}Y]+\underbrace{dE[e_{BLE}]}_{\text{לפי תכונה 1},=0}=cE[e_{BLE}Y]=$$

$$= cE\left[\left(X-\hat{X}_{BLE}\right)\cdot Y\right] = cE[(X-\eta_X-\frac{\sigma_{xy}}{\sigma_y^2}\cdot(Y-\eta_Y))\cdot(Y-\eta_Y+\eta_Y)] =$$
$$= cE[((X-\eta_X)-\frac{\sigma_{xy}}{\sigma_y^2}\cdot(Y-\eta_Y))\cdot(Y-\eta_Y+\eta_Y)] =$$
$$= c\begin{Bmatrix} E[(X-\eta_X)\cdot(Y-\eta_Y)]+\eta_Y E[X-\eta_X] \\ -\frac{\sigma_{xy}}{\sigma_y^2}\cdot E[(Y-\eta_Y)^2]-\frac{\sigma_{xy}}{\sigma_y^2}\cdot E[Y-\eta_Y]\cdot\eta_Y \end{Bmatrix} = 0$$

נסביר את השוויון האחרון:

(i) האיבר הראשון בסכום הוא σ_{xy}, והאיבר השלישי הוא $-\sigma_{xy}$.

(ii) האיבר השני והאיבר הרביעי שווים כ"א ל-0.

($\Longrightarrow$) נתון ש- $E[e\cdot(cY+d)]=0$ $\forall c,d\in\mathbb{R}:$.

יהי $\hat{X}_{BLE}=aY+b$ ו- $\hat{\hat{X}}=a'Y+b'$ משערך לינארי כלשהו. נראה ש:

$$E[(X-\hat{\hat{X}})^2]\geq E[(X-\hat{X})]^2$$

כלומר, ריבוע שגיאת השערוך בכל משערך לינארי אחר גדולה או שווה לריבוע שגיאת השערוך של משערך ששגיאתו מקיימת את הנ"ל. לכן משערך זה הוא האופטימלי:

$$E[(X-\hat{\hat{X}})^2] = E[((X-\hat{X})+(\hat{X}-\hat{\hat{X}}))^2] =$$
$$= E[(X-\hat{X})^2] \quad + 2E[(X-\hat{X})(\hat{X}-\hat{\hat{X}})] + \underbrace{E[(\hat{X}-\hat{\hat{X}})^2]}_{\text{אי שלילי}} \geq$$
$$\geq E[(X-\hat{X})^2] \quad + 2E[(X-\hat{X})(\hat{X}-\hat{\hat{X}})] \underbrace{=}_{(*)} E[(X-\hat{X})^2]$$

$$(*):\ E[(X-\hat{X})(\hat{X}-\hat{\hat{X}})] = E[e\cdot(aY+b-a'Y-b')] =$$
$$= E[e\cdot(\underbrace{(a-a')}_{c}Y+\underbrace{(b-b')}_{d}] = E[e\cdot(cY+d)] \underbrace{=}_{\text{לפי נתון}} 0$$

ב. $\hat{X}=g(Y)$ הינו משערך אופטימלי אס"ם

$$\forall h(\,):\ E\left[\left(X-g(Y)\right)\cdot h(Y)\right]=0$$

(3) <u>פיתגורס</u>:

א. $E[X^2]=E\left[\hat{X}_{BLE}^2\right]+E[e_{BLE}^2]$

ב. $E[X^2]=E\left[\hat{X}_{opt}^2\right]+E\left[e_{opt}^2\right]$

<u>הוכחה</u>:

$$E[X^2]=E[((X-\hat{X})+\hat{X})^2]=E[(X-\hat{X})^2]+2\underbrace{E[(X-\hat{X})\cdot\hat{X}]}_{\text{מתאפס לפי תכונה 2}}+E[\hat{X}^2] =$$
$$= E[e^2]+E[\hat{X}^2]$$

(4) עבור $\begin{bmatrix}X\\Y\end{bmatrix}$ ו"ג, $\hat{X}_{BLE}=\hat{X}_{opt}$.

<u>תמונה גיאומטרית של המשערכים</u>

נבחין שאוסף כל המשתנים האקראיים האפשריים בעלי מומנט שני סופי סגור תחת קומבינציה לינארית, ולכן הוא מרחב וקטורי. בנוסף, ניתן להגדיר מכפלה פנימית באופן הבא:

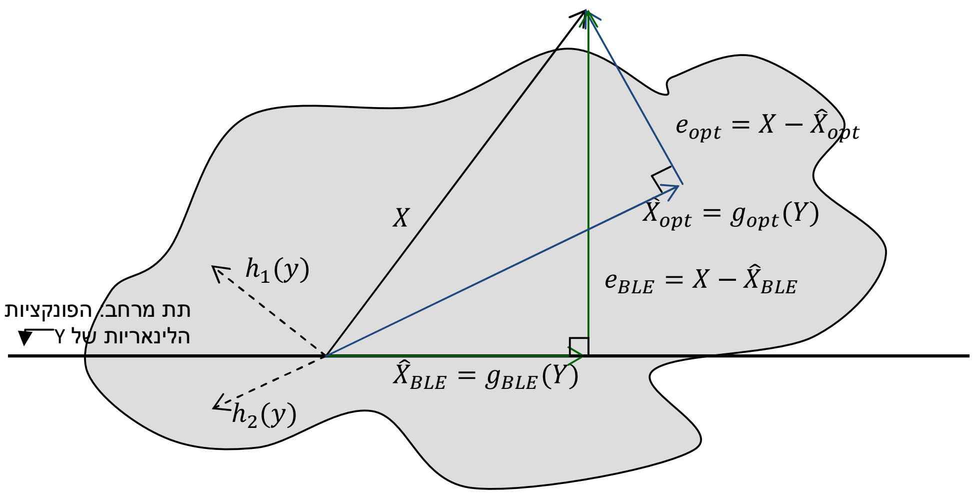
הגדרה: <u>מכפלה פנימית של מרחב הסתברות</u>: $<X,Y>=E[XY]$

קל להראות שהגדרה זו אכן מקיימת את האקסיומות של מכפלה פנימית.

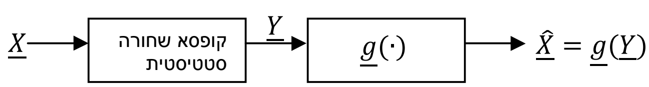
הגדרה: <u>נורמה של משתנה אקראי</u>: $\|X\|^2=<X,X>=E[X^2]$

הגדרה: הזוית בין X ו-Y היא: $\cos(\sphericalangle(X,Y)) = \frac{<X,Y>}{\|X\|\cdot\|Y\|} = \frac{E[XY]}{\sqrt{E[X^2]\cdot E[Y^2]}}$.

כעת נוכל להסביר את תכונת הניצבות באופן גרפי:



שערוך וקטור אקראי



(כעת $\underline{g}$ היא פונקציה וקטורית)

משמעות הסימון $\underline{\hat{X}} = \underline{g}(\underline{Y})$ היא:

$$\underline{\hat{X}} = \begin{bmatrix} \hat{X}_1 \\ \vdots \\ \hat{X}_n \end{bmatrix} = \begin{bmatrix} g_1(Y_1, \dots, Y_m) \\ \vdots \\ g_n(Y_1, \dots, Y_m) \end{bmatrix} = \underline{g}(\underline{Y})$$

משערך ייקרא אופטימלי אם $\forall 1 \le i \le n: E[e_i^2] \to min$.

נשים לב שכדי לדעת כיצד לשערך וקטור אקראי $\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ מתוך וקטור אקראי $\underline{Y}$ מספיק לדעת לשערך כל משתנה אקראי X_i בנפרד. הוקטור המשוערך $\underline{\hat{X}}$ יהיה הוקטור שרכיביו הם כל פונקציות השערוך.

שערוך אופטימלי של משתנה אקראי X מתוך וקטור אקראי $\underline{Y}$

טענה: $\hat{X}_{opt} = E[X|\, \underline{Y}] = \int_{-\infty}^{\infty} x f_{X|\,\underline{Y}}(x|\, \underline{y})dx$

הוכחת הטענה זהה לשערוך של משתנה אקראי מתוך משתנה אקראי יחיד.

שגיאת השערוך במקרה זה היא:

$$E[e_{opt}] = E[Var[X|\underline{Y}]] = \iiint_{-\infty}^{\infty} f_{\underline{Y}}(\underline{y}) \cdot \left(\int_{-\infty}^{\infty} (x-\eta_{X|\underline{Y}})^2 f_{X|\underline{Y}}(x|\underline{y})dx\right) d\underline{y}$$

שערוך לינארי אופטימלי של משתנה אקראי X מתוך וקטור אקראי $\underline{Y}$

הגדרה: משערך לינארי הוא משערך מהצורה $\hat{X} = \underline{a}^T \cdot \underline{Y} + b = \sum_{i=1}^{m} a_i Y_i + b$

שגיאת השערוך היא $e = X - \underline{a}^T \cdot \underline{Y} - b$. נרצה למצוא $\underline{a}^T, b$ כך ש $E[e^2]$ יהיה מינימלי.

$$E[e] = \eta_e = \eta_X - \underline{a}^T \cdot \eta_{\underline{Y}} - b$$
$$E[e^2] = Var[e] + E^2[e]$$

נראה תחילה שכמו במקרה של שערוך מתוך מ"א, גם כאן $Var[e]$ אינו תלוי ב- b. לכן נבחר את ערכו של b להיות כזה שיביא למינימום את $E^2[e]$. דהיינו נבחר b שמאלץ $\eta_e = 0$:

$$0 \underbrace{=}_{\text{דרישה}} \eta_e = \eta_X - \underline{a}^T \cdot \eta_{\underline{Y}} - b \Longrightarrow b = \eta_X - \underline{a}^T \cdot \eta_{\underline{Y}}$$

למעשה, בהינתן $\underline{a}$ כלשהו, ה-b הנ"ל מביא את השגיאה הריבועית למינימום מתוך כל ה- b האפשריים.

כעת נמצא את $\underline{a}^T$ שעבורו $E[e^2]$ מינימלי:

$$E[e^2] \underbrace{=}_{b \text{ לפי בחירת}} Var[e] = E[(e-E[e])^2] = E[(X - \underline{a}^T \cdot \underline{Y} - b - \eta_X + \underline{a}^T \cdot \eta_{\underline{Y}} + b)^2]$$
$$= E[((X-\eta_X) - \underline{a}^T \cdot (\underline{Y} - \eta_{\underline{Y}}))^2] =$$
$$= E[((X-\eta_X) - \underline{a}^T \cdot (\underline{Y} - \eta_{\underline{Y}})) \cdot ((X-\eta_X) - \underline{a}^T \cdot (\underline{Y} - \eta_{\underline{Y}}))^T] =$$
$$= E[(X-\eta_X) \cdot (X-\eta_X)]$$
$$-2\underline{a}^T \cdot E\left[(X-\eta_X) \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right)\right]$$
$$+\underline{a}^T \cdot E\left[\left(\underline{Y} - \eta_{\underline{Y}}\right) \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right)^T\right] \cdot \underline{a}$$
$$= \sigma_x^2 - 2C_{X\underline{Y}} \cdot \underline{a} + \underline{a}^T \cdot C_{\underline{Y}} \cdot \underline{a}$$

נמצא עבור איזה $\underline{a}$ מתקבל מינימום ע"י גזירת הביטוי (לפי $\underline{a}$) והשוואת הנגזרת ל-0:

$$\frac{d(E[e^2])}{d\,\underline{a}} = -2C_{X\underline{Y}} + 2\underline{a}^T \cdot C_{\underline{Y}} = 0 \Longrightarrow \underline{a}_{BLE}^T = C_{X\underline{Y}} \cdot C_{\underline{Y}}^{-1}$$
$$b_{BLE} = \eta_X - C_{X\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot \eta_{\underline{Y}}$$

<u>הערה</u>: את הגזירה לפי הוקטור $\underline{a}$ מבצעים ע"י כתיבת המכפלות באופן מפורש, גזירה, וכינוס חזרה לכתיב וקטורי.

השערוך הלינארי האופטימלי הוא, אם כן:

$$\hat{X}_{BLE} = \left(C_{X\underline{Y}} \cdot C_{\underline{Y}}^{-1}\right) \cdot \underline{Y} + \left(\eta_X - C_{X\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot \eta_{\underline{Y}}\right) = \eta_X + \left(C_{X\underline{Y}} \cdot C_{\underline{Y}}^{-1}\right) \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right)$$

שגיאת השערוך היא:

$$E[e_{BLE}^2] = Var[e_{BLE}] = \cdots = \sigma_x^2 - C_{X\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot C_{\underline{Y}X}$$

<u>תכונות של משערך כללי / לינארי אופטימלי</u> (ללא הוכחה)

(1) <u>חוסר הטיה</u>: תוחלת המשערך שווה לתוחלת המשוערך:

$$E[e_{BLE}] = 0,\ E[\hat{X}_{BLE}] = \eta_X$$
$$E[e_{opt}] = 0,\ E[\hat{X}_{opt}] = \eta_X$$

(2) <u>תכונת הניצבות</u>:

א. $\hat{X} = g(Y_1, \dots, Y_m)$ הוא אופטימלי אס"ם $\forall h(\,):\ E[e \cdot h(Y_1, \dots, Y_m)] = 0$

ב. $\hat{X} = \underline{a}^T \cdot \underline{Y} + b$ הוא BLE אס"ם

$$\forall d \in \mathbb{R}, \underline{c} \in \mathbb{R}^m:\ E\left[e \cdot \left(\underline{c}^T \cdot \underline{Y} + d\right)\right] = 0$$

(3) <u>פיתגורס</u>:

א. $E[X^2] = E\left[\hat{X}_{BLE}^2\right] + E\left[e_{BLE}^2\right]$

ב. $E[X^2] = E\left[\hat{X}_{opt}^2\right] + E\left[e_{opt}^2\right]$

<u>שערוך וקטור אקראי $\underline{X}$ מתוך וקטור אקראי $\underline{Y}$</u>

<u>משערך אופטימלי</u>:

$$\underline{\hat{X}}_{opt} = E\left[\underline{X}\middle|\underline{Y}\right] = \begin{bmatrix} E\left[X_1\middle|\underline{Y}\right] \\ \vdots \\ E\left[X_n\middle|\underline{Y}\right] \end{bmatrix}$$

<u>משערך לינארי אופטימלי</u>:

$$\underline{a}_{1,BLE}^T = C_{X_1\underline{Y}} \cdot C_{\underline{Y}}^{-1}, \qquad b_{1,BLE} = \eta_{X_1} - C_{X_1\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot \eta_{\underline{Y}}$$
$$\vdots$$
$$\underline{a}_{n,BLE}^T = C_{X_n\underline{Y}} \cdot C_{\underline{Y}}^{-1}, \qquad b_{n,BLE} = \eta_{X_n} - C_{X_n\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot \eta_{\underline{Y}}$$
$$\Longrightarrow A_{BLE}^T = C_{\underline{X},\underline{Y}} \cdot C_{\underline{Y}}^{-1}, \quad \underline{b}_{BLE} = \eta_{\underline{X}} - C_{\underline{X},\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot \eta_{\underline{Y}}$$

$$\underline{\hat{X}}_{BLE}^T = \left(C_{\underline{X},\underline{Y}} \cdot C_{\underline{Y}}^{-1}\right) \cdot \underline{Y} + \underline{b}_{BLE} = \eta_{\underline{X}} + C_{\underline{X},\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot (\underline{Y} - \eta_{\underline{Y}})$$
$$C_{\underline{e}} = C_{\underline{X}} - C_{\underline{X},\underline{Y}} \cdot C_{\underline{Y}}^{-1} \cdot C_{\underline{Y},\underline{X}}$$

<u>תכונות של משערך כללי / לינארי אופטימלי</u> (ללא הוכחה)

(1) <u>חוסר הטיה</u>: תוחלת המשערך שווה לתוחלת המשוערך:

$$E[\underline{e}_{BLE}] = 0,\ E[\underline{\hat{X}}_{BLE}] = \eta_{\underline{X}}$$
$$E[\underline{e}_{opt}] = 0,\ E[\underline{\hat{X}}_{opt}] = \eta_{\underline{X}}$$

(2) <u>תכונת הניצבות</u>:

א. $\underline{\hat{X}} = \underline{g}(\underline{Y})$ הוא אופטימלי אס"ם $\forall h(\,), 1 \le i \le n:\ E\left[e_i \cdot h\left(\underline{Y}\right)\right] = 0$.

בנוסף במשערך וקטורי מתקיים $E\left[\underline{e} \cdot \underline{h}^T\left(\underline{Y}\right)\right] = \underline{\underline{0}}$ לכל פונקציה וקטורית של $\underline{Y}$, המסומנת $\underline{h}()$

ב. $\underline{\hat{X}} = A\underline{Y} + \underline{b}$ הוא BLE אס"ם

$$\forall d \in \mathbb{R}, \underline{c} \in \mathbb{R}^m, 1 \le i \le n:\ E\left[e_{i,BLE} \cdot \left(\underline{c}^T\underline{Y} + d\right)\right] = 0$$

ומכאן שמתקיים

$$\forall \underline{d} \in \mathbb{R}^n, C \in \mathbb{R}^{n \times m}:\ E\left[\underline{e}_{BLE} \cdot \left(C\underline{Y} + \underline{d}\right)^T\right] = \underline{\underline{0}}$$

(3) <u>פיתגורס</u>:

א. $C_{\underline{X}} = C_{\underline{e}_{opt}} + C_{\underline{\hat{X}}_{opt}}$

ב. $C_{\underline{X}} = C_{\underline{e}_{BLE}} + C_{\underline{\hat{X}}_{BLE}}$

חלק ב: תהליכים אקראיים

מבוא לתהליכים אקראיים

הגדרה: <u>תהליך אקראי בזמן רציף</u> הוא פונקציה המשייכת לכל זמן t ותוצאת ניסוי $\omega \in \Omega$ מספר ממשי $X: \mathbb{R} \times \Omega \to \mathbb{R}, \mathbb{N}$. נסמן: $X(t, \omega)$ או בקיצור $X(t)$.

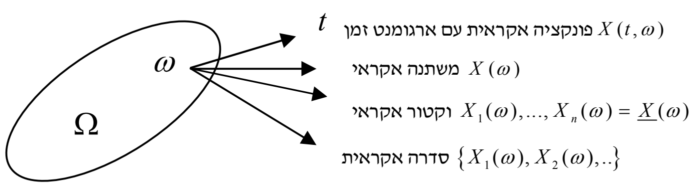
הגדרה: <u>תהליך אקראי בזמן בדיד</u> הוא פונקציה המשייכת לכל זמן n ותוצאת ניסוי $\omega \in \Omega$ מספר ממשי $X: \mathbb{N} \times \Omega \to \mathbb{R}, \mathbb{N}$. נסמן: $X(n, \omega)$. במקרה זה נסמן גם X_n.

הגדרה: <u>פונקצית המדגם</u> (Realization) היא הפונקציה הדטרמיניסטית המתקבלת מהמשתנה האקראי $X(t)$ או X_n ע"י הקפאת משתנה מרחב המדגם $(\omega = \omega_0)$ בעוד שמשתנה הזמן רץ.

<u>הערה</u>: אם מקפיאים את משתנה הזמן $(t = t_0)$, כלומר דוגמים את התהליך האקראי מקבלים משתנה אקראי.

נוכל לחשוב על תהליך אקראי כעל וקטור אינסופי: תהליך אקראי בזמן רציף הוא וקטור ש"אורכו" עוצמת הרצף (א); תהליך אקראי בזמן בדיד הוא וקטור ש"אורכו" אינסופי בן-מניה ($א_0$).

Ω = מרחב המדגם
$\omega \in \Omega$ = תוצאות ניסוי, התרחשות
F = מאורעות = איחוד של ω-ות
P = הסתברות של מאורע בF
מ"א = פונקציה של ω



<u>דוגמאות לתהליכים אקראיים בחיי היומיום ובהנדסה</u>

1. $X(t)$: מיקום של חלקיק של חומר שמתמוסס בחומר אחר (דיפוזיה).
2. $C^0(t)$: טמפרטורת גוף של אדם (או כל אות ביולוגי אחר, למשל: א.ק.ג).
3. רעש תרמי: מתח שנובע מתנועות אלקטרונים עקב חום.
4. $N(t)$: מספר האנשים הממתינים בתחנת אוטובוס, מספר שיחות הטלפון המגיעות למוקד, מספר החבילות המגיעות ליעד מסוים וכן הלאה.
   T_n: זמן ההגעה של: האוטובוס ה-n-י לתחנה, השיחה ה-n שמגיעה למוקד, החבילה ה-n...
5. $\$_n$: שער הדולר היציג ביום ה-n.
6. $A(n,m)$: עוצמת ההארה של הפיקסל ה- (n,m) במסך דו מימדי.
7. B_n: ערך הביט ה- n כשמורידים קובץ JPEG מהאינטרנט.

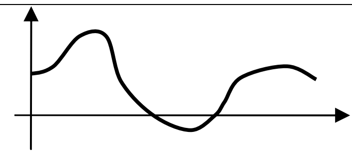
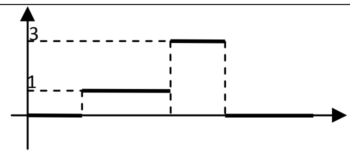
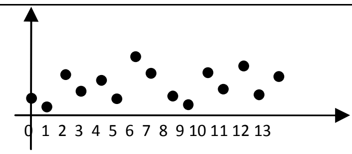
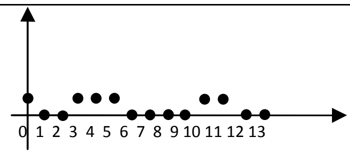
נשים לב שהערך אותו יכול לקבל התהליך האקראי הוא רציף או בדיד, כמו גם ציר הזמן המתאר את התהליך. ייתכנו כל השילובים האפשריים (ערך בדיד בזמן בדיד, ערך בדיד בזמן רציף, ערך רציף בזמן בדיד וערך רציף בזמן רציף).

נמיין את התהליכים שבדוגמא לפי הקריטריונים הנ"ל:

ערך / זמן	רציף	בדיד
רציף	דיפוזיה, רעש תרמי, טמפרטורה (אות ביולוגי)	$N(t)$
בדיד	שער הדולר (*) T_n $A(n,m)$	B_n

(*) נניח שרמת הדיוק אינה מוגבלת ל-3 ספרות אחרי הנקודה העשרונית

נצייר תרשימים המתאימים לתיאור פונקציות מדגם של התהליכים האקראיים, לפי המיון לעיל:

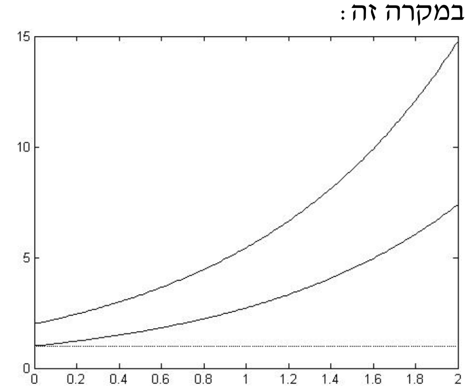
זמן \ ערך	רציף	בדיד
רציף		
בדיד		

דוגמא 1: <u>אקספוננט אקראי</u>: $t \geq 0$, $X(t) = \mathrm{A} \cdot e^{Bt}$; $\mathrm{A, B}$ משתנים אקראיים.

מרחב המדגם: $\left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\}$ $w.p$ $\Omega = \{\omega_1, \omega_2, \omega_3\}$, כאשר המאורעות ω_i הם המאורעות הבאים:

$$\omega_1 = \{A = 1, B = 0\} \Longrightarrow X(t) \equiv 1 \quad w.p\ \frac{1}{2}$$
$$\omega_2 = \{A = 1, B = 1\} \Longrightarrow X(t) = e^t \quad w.p\ \frac{1}{4}$$
$$\omega_3 = \{A = 2, B = 1\} \Longrightarrow X(t) = 2e^t \quad w.p\ \frac{1}{4}$$

נצייר את פונקציות המדגם במקרה זה:

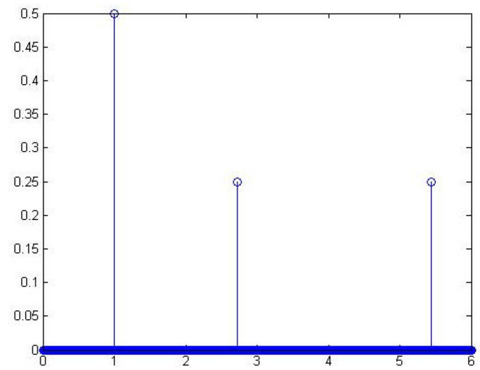


הגרף העליון מתאים לפונקצית המדגם עבור $\omega = \omega_3$.
הגרף האמצעי מתאים לפונקצית המדגם עבור $\omega = \omega_2$.
הגרף התחתון מתאים לפונקצית המדגם עבור $\omega = \omega_1$.

נקפיא את משתנה הזמן ברגע $t_0 = 1$. נקבל משתנה אקראי בעל הפילוג הבא:

$$X(t_0 = 1) = \begin{cases} 1 & w.p \quad \frac{1}{2} \\ e & w.p \quad \frac{1}{4} \\ 2e & w.p \quad \frac{1}{4} \end{cases}$$

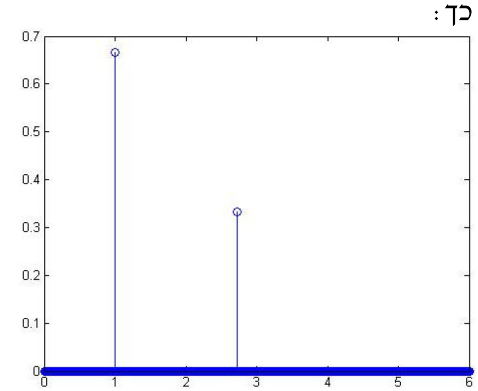
פונקצית הצפיפות של המשתנה האקראי הזה היא:



נחשב כעת את הפילוג המותנה של $X(t_0=1)$, בהינתן שמתקיים $X(t_0=0)=1$:

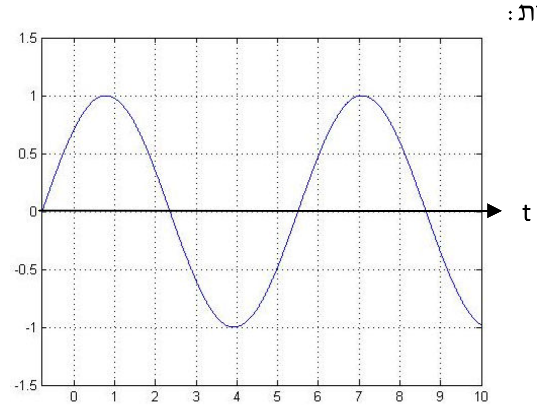
$$\Pr\left(X(1)=1 \,\middle|\, \underbrace{X(0)=1}_{\omega=\omega_2,\omega_1 \text{ מתרחש כאשר}}\right) \underbrace{=}_{Bayes} \frac{\Pr(X(1)=1 \wedge X(0)=1)}{\Pr(X(0)=1)} = \frac{\Pr(\omega=\omega_1)}{\Pr(\omega=\omega_2)+\Pr(\omega=\omega_1)} = \frac{2}{3}$$

ההתפלגות המותנית נראית כך :



דוגמא 2 : <u>סינוס עם פאזה אקראית</u> $t \geq 0$, $X(t)=\sin(\omega_0 t+\Theta)$; $\Theta \sim Unif(-\pi,\pi)$.
ω_0 קבוע, t ציר הזמן.

פונקצית מדגם אופיינית :



פונקצית מדגם אופיינית עבור $\omega_0=1$ ו- $\Theta=\theta_0=\frac{\pi}{4}$

<u>הערה</u> : בשרטוט מתוארת פונקצית המדגם האופיינית מרגע $t=-\theta_0$, אף על פי שהתהליך האקראי מוגדר עבור $t \geq 0$ בלבד.

נוכל לשאול את השאלות הבאות לגבי ההתפלגות של $X(t)$:

1. $\Pr(X(t_0)<3)=?$

ברור שבמקרה הזה ההסתברות היא 1, מפני שפונקציית sin חסומה בין 1- ל 1.

2. $\Pr\left(X(t) \geq 0, \forall 0 \leq t \leq \frac{2\pi}{\omega_0}\right) = ?$

על-פני מחזור שלם בהכרח sin מקבלת גם ערכים שליליים, לכן ההסתברות היא 0.

3. $\Pr\left(X(t) \geq 0, \forall 0 \leq t \leq \frac{\pi}{2\omega_0}\right) = ?$

נשים לב שפונקציית sin מקבלת ערכים חיוביים לאורך חצי מחזור, וערכים שליליים לאורך חצי המחזור הבא. נרצה שהתחום הנתון, שאורכו רבע מחזור, יהיה כולו מוכל בתוך חצי מחזור שבו הערכים של sin חיוביים, ומכאן שההסתברות היא 0.25, שכן המאורע מתרחש עבור $0 \leq \theta \leq \dfrac{\pi}{2}$ - רבע מתחום הערכים של θ.

4. $\Pr\left(\exists 0 \leq t \leq \frac{\pi}{\omega_0} : X(t) = 1\right) = ?$

התשובה במקרה זה היא 0.5 (פעם במחזור מתקבל הערך 1. אנו רוצים שקטע באורך חצי מחזור יכיל נקודה זו)

5. מהי ההתפלגות המותנית: $X(t)|_{X(0)=sin\alpha}$ (כאשר $\alpha \neq \dfrac{\pi}{2}, \dfrac{3\pi}{2}$).

ישנם שני פתרונות (בקטע $(-\pi,\pi)$) למשוואה $siny = sin\alpha$, והם $y_1 = \alpha, y_2 = \pi - \alpha$.
לכן לפי הנתון, נסיק כי $\Theta = \alpha, \Theta = \pi - \alpha$. שתי הזויות הללו מתקבלות בהסתברויות שוות, מכיוון שהפילוג של Θ הוא אחיד.

נסכם:

$$X(t)|_{X(0)=sin\alpha} = \begin{cases} sin(\omega_0 t + \alpha) & w.p \ \ \frac{1}{2} \\ sin(\omega_0 t + \pi - \alpha) & w.p \ \ \frac{1}{2} \end{cases}$$

6. $E[X(t_0)] = E[\sin(\omega_0 t_0 + \Theta)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} sin(\omega_0 t_0 + \theta)\, d\theta = 0$

הסבר: נובע מיידית מהגדרת האינטגרל, ומכך שאינטגרל על פני מחזור של פונקציה מחזורית הינו איווריאנטי להזזה (סופית).

צפיפות הפילוג של דגימת התהליך - $f_{x(t_0)}(x) = \sin(\overbrace{\omega_0 t_0}^{const} + \theta)$.

7. "תוחלת בזמן": $\frac{1}{T}\int_0^T X(t)dt \to_{T\to\infty} 0 \ \ w.p.1$.

הסבר: נשים לב ש- $\int_0^T X(t)dt$ הוא משתנה אקראי שחסום בתחום $\left[-\dfrac{1}{\omega_0}, \dfrac{1}{\omega_0}\right]$ (T הוא פרמטר). לכן לאחר חלוקה ב-T, כאשר $T \to \infty$, מקבלים 0.

נשים לב שהתוחלת, ולמעשה הפילוג השולי של X(t) כולו, קבועים בזמן. גם הפילוג המותנה $P(X(t+\Delta) \mid X(t) = x_0)$ קבוע בזמן. על כן האקראיות של תהליך זה הינה קבועה בזמן. בהמשך נגדיר תכונה זאת כ:סטציואנריות".

<u>דוגמאות נוספות לתהליכים אקראיים</u>:

א. $X(t) = \begin{cases} t & \text{מטבע = "עץ"} \\ t^2 & \text{מטבע = "פאלי"} \end{cases}$

ב. נתונה המד"ר הבאה, עם תנאי התחלה:

$$\ddot{Y}_t + A \cdot \dot{Y}_t + B \cdot Y_t = 0$$
$$Y_0 = C$$
$$\dot{Y}_0 = D$$

כאשר A,B,C,D הם משתנים אקראיים.

ג. $V\sim Unif[0,1]$ מ״א. נפתח את V לייצוג בינארי: $0.X_1X_2X_3\ldots$, אזי $X_1,X_2,X_3,\ldots$ הוא ת״א בזמן בדיד.

נשים לב ש- $\Pr(X_k=1)=\frac{1}{2}$, מפני שכל ביט מציין האם אנו נמצאים מימין או משמאל לתחום מסוים (X_1 מציין האם אנו בין 0 ל- $\frac{1}{2}$, X_2 מציין האם אנחנו בין 0 ל- $\frac{1}{4}$ או בין $\frac{1}{2}$ ל- $\frac{3}{4}$, או שמא בין $\frac{1}{4}$ ל- $\frac{1}{2}$ או בין $\frac{3}{4}$ ל- 1, וכן הלאה). מכיוון שאנו מגרילים את כל הקטע $[0,1]$ בהסתברויות שוות, הנ״ל מתקיים.

בנוסף נוכל לומר על $\{X_k\}_{k=1}^{\infty}$ שהם מתפלגים $i.i.d$ (ראה בהמשך).

כל הדוגמאות עד עכשיו היו של תהליכים אקראיים שכל האקראיות שלהם נובעת מפרמטר אקראי בודד או ממספר סופי של פרמטרים אקראיים. נראה להלן איך בונים תהליך שהאקראיות שלו מתחדשת כל הזמן.

הגדרה של תהליך אקראי ע״י פילוג משותף

תהליך אקראי בדיד $X_1,X_2,X_3,\ldots$

הפילוג מוגדר לחלוטין ע״י הפילוג המשותף של n דגימות לכל n טבעי

$$p(x_1,x_2,x_3,\ldots,x_n),\quad \forall n\in\mathbb{N}$$

(p מייצג פילוג ערך בדיד. נחליף את p ב- f עבור פילוג בעל ערכים רציפים).

יש לוודא כי מתקיימת תכונת העקביות: כל פילוג מסדר גבוה צריך להתלכד עם פילוג נתון מסדר נמוך. כלומר, נבדוק ש:

$$p(x_1,x_2,x_3,\ldots,x_{n-1})=\sum_{x_n}p(x_1,x_2,x_3,\ldots,x_n)$$

תהליך אקראי רציף $\{X(t),\ t\geq 0\}$

הפילוג מוגדר ע״י הפילוג המשותף של כל קבוצה של n דגימות זמן $t_1,t_2,t_3,\ldots,t_n$ לכל n טבעי.

$$p\left(x_{t_1},x_{t_2}x_{t_3},\ldots,x_{t_n}\right),\quad \forall t_1,t_2,t_3,\ldots,t_n,\quad \forall n\in\mathbb{N}$$

גם במקרה זה צריכה להתקיים תכונת העקביות.

הגדרה: פילוג שולי של תהליך אקראי בזמן רציף $X(t)$ הוא הפילוג של המשתנה האקראי $X(t_0)$ (לכל t_0).

הגדרה: פילוג שולי של תהליך אקראי בזמן בדיד X_n הוא הפילוג של המשתנה האקראי X_{n_0} (לכל n_0)

תהליך אקראי גאוסי

הגדרה: $X(t)$ ייקרא תהליך אקראי גאוסי אם כל קבוצת דגימות שלו $X(t_1),X(t_2),\ldots,X(t_n)$ לכל n טבעי היא גאוסית במשותף. כלומר, $\forall a_1,a_2,a_3,\ldots,a_n:\sum_{i=1}^{n}a_i\cdot X(t_i)$ הוא משתנה אקראי גאוסי (ראה הגדרה בעמ׳ 24)

הערה ותזכורת: לא מספיק שהפילוג השולי יהיה גאוסי בכדי לומר שהתהליך האקראי הינו גאוסי!
עבור משתנים אקראיים, פילוג גאוסי מוגדר באופן יחיד ע״י תוחלת ושונות.
עבור וקטורים אקראיים, פילוג גאוסי מוגדר באופן יחיד ע״י וקטור תוחלת ומטריצת קווריאנס.

בשני המקרים מדובר בסטטיסטיקה מסדר שני של המשתנה או של הוקטור.

כיצד מוגדר תהליך גאוסי אקראי? ע"י סטטיסטיקה מסדר שני של תהליכים אקראיים: פונקצית תוחלת ופונקצית אוטו-קווריאנס או פונקצית אוטו-קורלציה.

סטטיסטיקה מסדר שני של תהליכים אקראיים

הגדרה: תוחלת של תהליך אקראי היא פונקציה דטרמיניסטית המקיימת

$$\eta(t) = E[X(t)]$$

הגדרה: פונקצית האוטו-קורלציה של תהליך אקראי היא פונקציה דטרמיניסטית בעלת שני משתנים, המוגדרת באופן הבא:

$$R_X(t_1, t_2) \triangleq E[X(t_1) \cdot X(t_2)]$$

הגדרה: פונקצית האוטו-קווריאנס של תהליך אקראי היא פונקציה דטרמיניסטית בעלת שני משתנים, המוגדרת באופן הבא:

$$C_X(t_1, t_2) \triangleq E\left[\left(X(t_1) - \eta(t_1)\right) \cdot \left(X(t_2) - \eta(t_2)\right)\right]$$

טענה: $C_X(t_1, t_2) = R_X(t_1, t_2) - \eta(t_1) \cdot \eta(t_2)$

טענה: פונקצית האוטו-קורלציה ופונקצית האוטו-קווריאנס הן אי שליליות מוגדרות:
לכל קבוצה של דגימות $t_1, t_2, t_3, \dots, t_n$ ולכל קבוצה של קבועים $a_1, a_2, a_3, \dots, a_n$ מתקיים:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j R_X\left(t_i, t_j\right) \geq 0 ,\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j C_X\left(t_i, t_j\right) \geq 0$$

הגדרה: מומנט שני של תהליך אקראי: $E[X^2(t)] = R_X(t, t)$

הגדרה: שונות של תהליך אקראי: $Var[X(t)] = C_X(t, t)$

מתקיים: $C_X(t, t) = R_X(t, t) - \eta^2(t)$

הגדרה: מקדם הקורלציה של תהליך אקראי:

$$\rho(t_1, t_2) = \frac{C_X(t_1, t_2)}{\sqrt{C_X(t_1, t_1) \cdot C_X(t_2, t_2)}} \underbrace{=}_{\eta(t) \equiv 0} \frac{R_X(t_1, t_2)}{\sqrt{R_X(t_1, t_1) \cdot R_X(t_2, t_2)}}$$

טענה: $-1 \leq \rho(t_1, t_2) \leq 1$. ניתן להוכיח טענה זו באמצעות אי שוויון קושי שוורץ (כמו שהוכחנו עבור זוג משתנים אקראיים בפרק על מומנטים משותפים).

סטאציונאריות של תהליך אקראי

הגדרה: תהליך אקראי נקרא סטאציונארי במובן הצר (*Strict Sense Stationary*, נסמן *S.S.S*) אם הפילוג המשותף שלו לכל וקטור זמנים $\{t_i\}_{i=1}^{n}$ (לכל *n*) הוא קבוע ביחס להזזה בזמן:

$$\forall \Delta \in \mathbb{R}: \quad P\left(X(t_1 + \Delta), \dots, X(t_n + \Delta)\right) = P\left(X(t_1), \dots, X(t_n)\right)$$

בפרט עבור $n = 1$: $P\left(X(t_1 + \Delta)\right) = P\left(X(t_1)\right)$, כלומר הפילוג השולי אינו משתנה בזמן.

נוכל לומר שתהליך סטאציונארי הוא תהליך אקראי בעל אותה אקראיות בכל נקודה בזמן.

הגדרה: תהליך אקראי נקרא סטאציונארי במובן הרחב (*Wide Sense Stationary*, נסמן *W.S.S*) אם הסטטיסטיקה מסדר שני שלו היא קבועה ביחס להזזה בזמן:

$$\forall \Delta \in \mathbb{R}: \quad \eta(t) = \eta(t+\Delta) \Rightarrow \eta(t) = constant$$

$$R_X(t_1+\Delta, t_2+\Delta) =$$
$$R_X(t_1,t_2) \underbrace{=}_{\substack{\text{תלוי רק בהפרש,} \\ \text{נעבור לפונקציה} \\ \text{עם ארגומנט יחיד}}} R(t_1-t_2) \underbrace{=}_{\tau=t_1-t_2} R(\tau) \underbrace{=}_{\text{קבוע ביחס ל-}t} E[X(t+\tau)\cdot X(t)]$$

<u>הערות</u>:

1. $S.S.S$ הוא מקרה פרטי של $W.S.S$, כלומר $SSS \Rightarrow WSS$, אבל לא להפך.
2. תהליך אקראי גאוסי מוגדר לחלוטין ע"י סטטיסטיקה מסדר שני. לכן עבור תהליך אקראי גאוסי שני המובנים של הסטאציונאריות מתלכדים.

הגדרה: תהליך אקראי נקרא <u>סטאציונארי אסימפטוטית</u> (התהליך "שואף להיות" סטאציונארי):

- <u>במובן הצר</u>: אם קיים הגבול:

$$\Pr(X_k = a_1, \dots, X_{k+n} = a_n) \to_{k\to\infty} P_{stat}(a_1, \dots, a_n)$$

עבור איזושהי פונקציה $P_{stat}()$.

- <u>במובן הרחב</u>: אם קיימים הגבולות הבאים:

$$\eta(t) \to_{t\to\infty} \eta_{stat} = Const.$$
$$R(t,t+\tau) \to_{t\to\infty} R_{stat}(\tau)$$

בנייה של תהליכים אקראיים

הגדרה: תהליך אקראי בדיד בזמן $\{X_i\}_{i=1}^{\infty}$ מתפלג *i.i.d* (*Independent Identically Distributed*) אם מתקיימים התנאים הבאים:

i. כולם בעלי אותו פילוג $\forall 1 \leq i: X_i \sim p(x)$ (נחליף את $p(x)$ ב- $f(x)$ במקרה הרציף).

ii. $X_1, X_2, X_3, \ldots$ בלתי תלויים הדדית.

נסמן אי תלות באופן הבא: $X \perp\!\!\!\perp Y$ (X בלתי תלוי ב- Y). אזי:

$$\forall n: (X_1, \ldots, X_{n-1}) \perp\!\!\!\perp X_n$$

<u>הערה</u>: אי תלות בזוגות אינה גוררת אי תלות הדדית!!

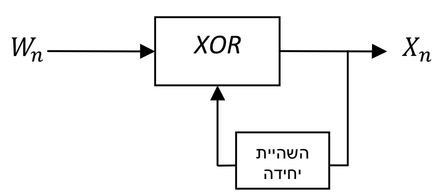
דוגמא 3 (אי תלות בזוגות): נתונים X_1, X_2 בת"ס בעלי הפילוג הבא: $X_i = \begin{cases} 1 & w.p \ \frac{1}{2} \\ 0 & w.p \ \frac{1}{2} \end{cases}$.

נגדיר משתנה אקראי חדש $Y = X_1 \ XOR \ X_2$.
נשים לב שכל זוג של מ"א מגדיר את המשתנה השלישי באופן יחיד, לכן כל משתנה בהחלט תלוי בשניים האחרים, ומצד שני כל זוג של משתנים שנבחר הוא בת"ס (למשל, עבור X_1 קבוע, אם אנו לא יודעים את X_2 אז בפועל אין בידינו שום מידע נוסף על הפילוג של Y).

<u>תהליך אוטו-רגרסיבי (Auto-Regressive Process, AR)</u>

הגדרה: יהי W_n תהליך אקראי *i.i.d* בעל פילוג ידוע, ו X_0 משתנה אקראי (בעל פילוג ידוע) או דטרמיניסטי, וכן X_0 בלתי תלוי ב- $W_1, W_2, W_3, \ldots$, ותהי g פונקציה דטרמיניסטית.
אזי $X_n = g(X_{n-1}, W_n)$ נקרא <u>תהליך אוטו-רגרסיבי</u> (Auto-Regressive Process).

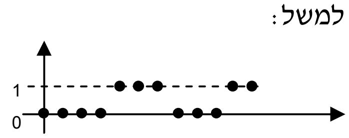
דוגמא 4:



כלומר $X_n = W_n \oplus X_{n-1}$, כאשר $W_n \sim Ber(p)$, ונניח $p \ll 1$.

פונקצית מדגם: מכיוון ש W_n יכול לקבל את הערכים 0 ו- 1, כך שההסתברות לקבל 0 גבוהה הרבה יותר, אנו מקבלים כי $X_n = X_{n-1}$ בהסתברות גבוהה.

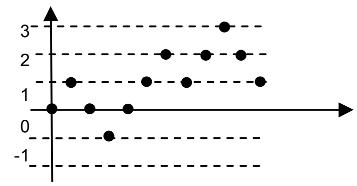
פונקצית מדגם של X_n תראה כך, למשל:



דוגמא 5: "הילוך שיכור" $X_n = X_{n-1} + W_n$, כאשר $W_n = \begin{cases} 1 & w.p \ \frac{1}{2} \\ -1 & w.p \ \frac{1}{2} \end{cases}$.

נשים לב ש: $X_n = W_1 + W_2 + \cdots + W_n$.

פונקצית מדגם של X_n תראה כך, למשל:



הגדרה: <u>תהליך AR ממשי לינארי כללי</u> הוא תהליך AR מהצורה $X_n = \alpha \cdot X_{n-1} + W_n$.

נוכל לחשוב על תהליך זה כאילו W_n נכנס לתוך מערכת לינארית קבועה בזמן (LTI).

$$W_n \longrightarrow \boxed{H(z)} \longrightarrow X_n$$

אזי מתקיים: $h[n] = \alpha^n, \;\; n \geq 0$ ופונקצית התמסורת של המערכת היא:

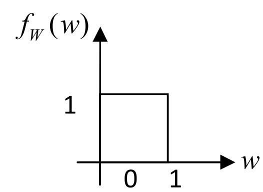
$$H(z) = \frac{1}{1-\alpha \cdot z^{-1}}$$

דוגמא 6: <u>הצגת שרשרת מרקוב במונחים של תהליך AR</u>

<u>תזכורת</u>: תהליך AR הוא תהליך בו מוגדרת פונקצית הרקורסיה g: $g(X_{n-1}, W_n) = X_n$

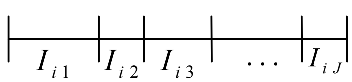
שרשרת מרקוב מקיימת: $P_{ij} = \Pr(X_n = j \mid X_{n-1} = i)$

<u>תהליך הבנייה</u>:



$$X_n \in \{1,2,...,J\}$$
$$W_n \sim Unif\,[0,1)$$

לכל i $(=1,...,J)$ נחלק את הקטע $[0,1)$ ל- J אינטרוולים לא שווים בהכרח הקטעים זרים ועל-כן איחודם נותן את הקטע $[0,1)$.



<u>הגדרה</u>: פונקצית הרקורסיה מוגדרת כך: $g(i, W_n) = j \quad if \;\; W_n \in I_{ij}$

כעת נראה כי אם נקבע את אורך האינטרוול I_{ij} להיות שווה ל- P_{ij}, אזי התהליך ה- AR יתלכד עם שרשרת מרקוב:

$$\Pr(X_n = j \mid X_{n-1} = i) \underset{\substack{\text{מהגדרת פונקצית} \\ \text{הרקורסיה}}}{=} \Pr\{W_n \in I_{ij}\} \underset{\substack{\text{הסיכוי ליפול באינטרוול} \\ \text{הוא אורך האינטרוול}}}{=} |\, I_{ij} \,| = P_{ij}$$

תנאים לסטאציונאריות של תהליך A.R לינארי:

יהי $X_n = \alpha \cdot X_{n-1} + W_n$, $X_0 \sim P_{X_0}(x)$ (W_n תהליך אקראי $i.i.d$) תהליך A.R לינארי כלשהו.

1. אם $|\alpha| \geq 1$ אזי התהליך X_n אינו סטאציונארי.
2. אם $|\alpha| < 1$ אזי התהליך סטאציונארי אסימפטוטית (במובן הצר ובמובן הרחב).
3. אם $|\alpha| < 1$ וגם תנאי ההתחלה של התהליך "מתואמים" אזי התהליך סטאציונארי.
   תנאי ההתחלה מתואמים עבור סטאציונאריות במובן הצר אם $P_{X_0}(x) = P_{stat}(x)$.
   תנאי ההתחלה מתואמים עבור סטאציונאריות במובן ברחב אם: $\eta_{X_0} = \eta_{stat}, \sigma_{X_0}^2 = \sigma_{stat}^2$.
   (נשים לב שאת η_{stat}, σ_{stat}^2 ניתן לחשב מתוך הסטטיסטיקה מסדר II של W_n ומתוך α).

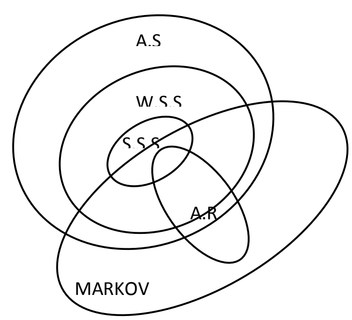
הגדרה: תהליך אקראי ייקרא <u>מרקובי</u> אם הפילוג המותנה של העתיד בהינתן ההווה והעבר שווה לפילוג המותנה של העתיד בהינתן ההווה בלבד:

$$\Pr(X_n = a | X_m = b, X_k = c) = Pr(X_n = a | X_m = b), \quad (n > m > k)$$
$$\Pr(X_n | X_{n-1}, \dots, X_1) = Pr(X_n | X_{n-1})$$

<u>טענה</u>: תהליך A.R הוא מרקובי.

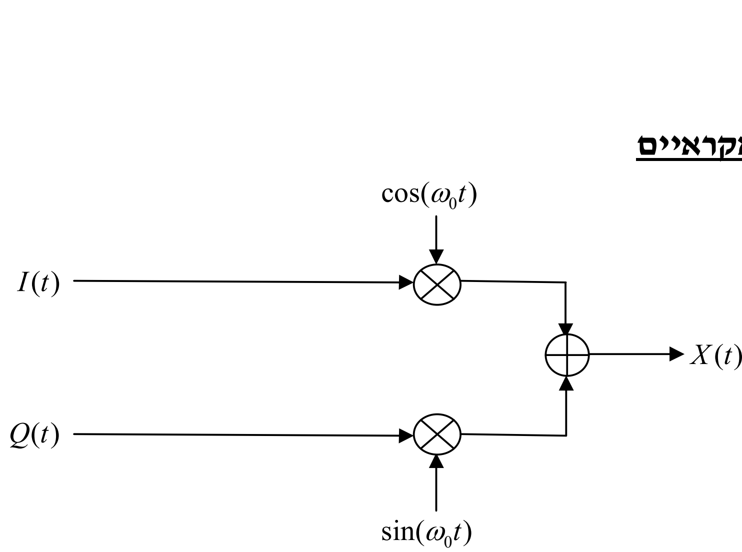
<u>הסבר</u>: נשים לב ש- X_n תלוי ב- X_{n-1} וב- W_n. מכיוון ש- W_n מתפלגים $i.i.d$ אזי ההתפלגות של W_n אינה תלויה ב- W_k ($n > k$) קודמים ולכן אינה תלויה גם ב- X_k ($n > k$) קודמים. מכאן ש- X_{n-1} מסכם בתוכו את כל העבר הרלוונטי עבור X_n. כמו כן עבור כל $n > k$, X_n הוא פונקציה דטרמיניסטית של X_k ו- W_n, ..., W_k, W_{k+1}.

<u>מפת עולם התהליכים האקראיים</u>:



<u>הערה</u>: בשרטוט .A.S הוא תהליך אקראי אסימפטוטי סטציונרי במובן הרחב.

<u>אפנון של אותות אקראיים</u>



מערכת זו מאפננת את האותות (תהליכים) האקראיים $I(t),Q(t)$, שהינם סטציונארים במשותף במובן הרחב (J.W.S.S), על ידי הגלים הנושאים $\cos(\omega_0 t),\sin(\omega_0 t)$ בהתאמה. מוצא המערכת מקיים

$$X(t)=I(t)\cos(\omega_0 t)+Q(t)\sin(\omega_0 t)=r(t)\cos\left(\omega_0 t+\varphi(t)\right)$$

כאשר $r(t)$ הינו אות (תהליך) אקראי שמוגדר $r(t)=\sqrt{I^2(t)+Q^2(t)}$ ו $\varphi(t)$ הינו אות (תהליך) אקראי שמוגדר $\tan\varphi(t)=-\dfrac{Q(t)}{I(t)}$. האות $X(t)$ הינו אות האפנון, כאשר **אפנון האמפליטודה** הינו $r(t)$ **ואפנון הפאזה** הינו $\varphi(t)$.

תחילה, לשם פשטות והבנה מקדימה, נניח כי האותות האקראיים המאפננים הינם קבועים בזמן, משמע משתנים אקראיים - $I(t)=A \quad Q(t)=B$. אילו תנאים צריכים לקיים המ״א A וB על מנת שהתהליך $X(t)=A\cos(\omega_0 t)+B\sin(\omega_0 t)$ יהיה סטציונארי במובן הצר (**S.S.S**), ובמובן הרחב (**W.S.S**)?

<u>תנאים לסטציונאריות במובן הצר של האות המאופנן</u>

א) <u>במובן הצר S.S.S</u>

זמן המחזור של האותות הנושאים $\sin(\omega_0 t),\cos(\omega_0 t)$ מקיים $T=\dfrac{2\pi}{\omega_0}$, כך שנשים לב כי

$$X(t=0)=A \quad X(t=\frac{T}{4})=B \quad X(t=\frac{T}{2})=-A \quad X(t=\frac{3T}{4})=-B$$

נזכיר כי $X(t)$ יהיה S.S.S אם ורק אם לכל $c\in\mathbb{R}$, לתהליכים $X(t)$ ו $Y(t)=X(t+c)$ ישנה אותה התפלגות. בפרט ההתפלגות של המשתנה האקראי $X(t_0)$ עבור t_0 מסוים (התפלגות שולית של התהליך) לא תלויה בזמן, למשל עבור $t_0=0,\frac{T}{4},\frac{T}{2},T$. על כן המשתנים האקראיים A,B בהכרח צריכים להיות בעלי התפלגות זהה, כלומר שהמשתנים האקראיים $A,B,-A,-B$ צריכים להיות בעלי התפלגות זהה שהינה זוגית, זוהי סימטריה של סיבוב ב90 מעלות במישור (a,b), כלומר $f_A(a)=f_B(b)=f_A(-a)=f_B(-b)$.

עתה נדמיין כי אוספים את כל התנאים שA,B בהכרח מקיימים כדי שX(t) יהיה S.S.S בעזרת בדיקת X(t) בזמנים נוספים מעבר ל $0,\frac{T}{4},\frac{T}{2},T$ - כלומר לכל הזמנים שהם בצורה רציפה.

הסימטריה שתתקבל תהיה חזקה יותר – סימטריה לסיבוב במישור a,b, קרי סימטריה מעגלית של ההתפלגות המשותפת. על כן לא נופתע מנכונותה של הטענה הבאה –

<u>**טענה:**</u> $X(t)=A\cos(\omega_0 t)+B\sin(\omega_0 t)$ יהיה תהליך S.S.S **אם ורק אם** הפילוג המשותף של A,B הוא בעל סימטריה מעגלית,כלומר ש - $f_{AB}(a,b)=g(a^2+b^2)$ כאשר $g(x)$ פונקציה דטרמיניסטית כלשהי.

דוגמא להתפלגות משותפת מוכרת שבעלת סימטריה כזו היא ההתפלגות המשותפת של A,B מ״א גאוסיים עם תוחלת אפס לכל אחד, חסרי קורלציה אחד עם השני (ולכן בת״ס וגם גאוסיים במשותף).

א) <u>במובן הרחב W.S.S</u>

אם נסתפק בדרישה כי $X(t)$ יהיה W.S.S, שתאפשר לנו להגדיר בהמשך ספקטרום הספק ל $X(t)$ אז נקבל דרישות מסוימות על המ״א A,B -

$$E\left\{X(t)\right\}=\cos(\omega_0 t)E\left\{A\right\}+\sin(\omega_0 t)E\left\{B\right\}\underset{\forall t}{=}0\Rightarrow E\left\{A\right\}=E\left\{B\right\}=0$$

$$R_X(t_1,t_2)=E\left\{X(t_1)X(t_2)\right\}=E\left\{\left(A\cos(\omega_0 t_1)+B\sin(\omega_0 t_1)\right)\left(A\cos(\omega_0 t_2)+B\sin(\omega_0 t_2)\right)\right\}=$$
$$\cos(\omega_0 t_1)\cos(\omega_0 t_2)E\left\{A^2\right\}+\sin\left(\omega_0(t_1+t_2)\right)E\left\{AB\right\}+\sin(\omega_0 t_1)\sin(\omega_0 t_2)E\left\{B^2\right\}$$

בהכרח צריך להתקיים $E\left\{AB\right\}=0$ (אורתוגונאליות) אחרת ישנה תלות ב t_1+t_2.
$E\left\{A\right\}=E\left\{B\right\}=0$ כך שנקבל דרישה לחס״ק בין A לB. בנוסף הכרחי כי $E\left\{A^2\right\}=E\left\{B^2\right\}$, כלומר $Var\left\{A\right\}=Var\left\{B\right\}=\sigma^2$ במקרה זה, אחרת ישנה תלות ב t_1,t_2. אם אכן תנאים אלו מתקיימים נקבל

$$R_X(t_1,t_2)=\left(\cos(\omega_0 t_1)\cos(\omega_0 t_2)+\sin(\omega_0 t_1)\sin(\omega_0 t_2)\right)\sigma^2=\sigma^2\cos\left(\omega_0(t_1-t_2)\right)$$

משמע $R_X(t_1,t_2)=R_X(\tau)$ ולמעשה הוכחנו את הטענה הבאה –

<u>**טענה:**</u> $X(t)=A\cos(\omega_0 t)+B\sin(\omega_0 t)$ יהיה תהליך W.S.S **אם ורק אם** A ו-B משתנים אקראיים שווי שונות, חסרי קורלציה ובעלי תוחלת אפס כל אחד, כלומר $E\left\{A^2\right\}=E\left\{B^2\right\}$ ו $E\left\{A\right\}=E\left\{B\right\}=E\left\{AB\right\}=0$.

במבט ראשון נדמה כאילו אין קשר בין התנאים לS.S.S וW.S.S , והרי שאנו יודעים כי עבור תהליך מסוים $S.S.S \Rightarrow W.S.S$. אלא שמקרה פרטי של התנאי לS.S.S הינו $f_A(a) = f_B(b) = f_A(-a) = f_B(-b)$ שראינו, ומכאן $E\{B\} \underset{f_a(a)=f_b(b)}{=} E\{A\} \underset{def}{=} \int_{-\infty}^{\infty} \underbrace{af_a(a)}_{odd\ function}\, da = 0$.

בנוסף ראינו כי $\varphi \sim Uniform(-\pi,\pi)$, ומכיוון ש $\tan\varphi = -\frac{B}{A}$ נובע מתנאי זה כי A חסר קורלציה עם B, $\mathrm{cov}(A,B) = E\{AB\} - E\{A\}E\{B\} = 0$. יחד נקבל כי ישנה אורתוגונאליות בין A וB, $E\{AB\} = 0$, כך שאכן התקבלו התנאים לW.S.S בטענה הנ"ל.

<u>תנאים לסטציונאריות של האות המאופנן – אפנון כללי</u>

עתה נחזור לדון במקרה הכללי שבו $X(t) = I(t)\cos(\omega_0 t) + Q(t)\sin(\omega_0 t)$, ונשאל מתי X(t) עתה יהיה תהליך סטציונארי במובן הרחב. נשים לב לדמיון למקרה הקבוע בזמן.

$$E\{X(t)\} = \cos(\omega_0 t)E\{I(t)\} + \sin(\omega_0 t)E\{Q(t)\} \underset{\forall t}{=} 0 \Rightarrow E\{I(t)\} = E\{Q(t)\} \underset{\forall t}{=} 0$$

$$\begin{aligned}
R_X(t_1,t_2) &= E\{X(t_1)X(t_2)\} = E\{(I(t_1)\cos(\omega_0 t_1) + Q(t_1)\sin(\omega_0 t_1))(I(t_2)\cos(\omega_0 t_2) + Q(t_2)\sin(\omega_0 t_2))\} = \\
&(\cos(\omega_0 t_1)\cos(\omega_0 t_2))E\{I(t_1)I(t_2)\} + (\sin(\omega_0 t_1)\cos(\omega_0 t_2))E\{Q(t_1)I(t_2)\} \\
&+ (\sin(\omega_0 t_2)\cos(\omega_0 t_1))E\{I(t_1)Q(t_2)\} + (\sin(\omega_0 t_1)\sin(\omega_0 t_2))E\{Q(t_1)Q(t_2)\} \underset{J.W.S.S}{=} \\
&\cos(\omega_0 t_1)\cos(\omega_0 t_2)R_I(\tau) + \sin(\omega_0 t_1)\cos(\omega_0 t_2)R_{QI}(\tau) + \sin(\omega_0 t_2)\cos(\omega_0 t_1)R_{QI}(\tau) + \sin(\omega_0 t_1)\sin(\omega_0 t_2)R_Q(\tau)
\end{aligned}$$

לכן נרצה לדרוש $R_I(\tau) = R_Q(\tau)$ וגם $R_{IQ}(\tau) = -R_{QI}(\tau) \Leftrightarrow R_{IQ}(\tau) = -R_{IQ}(-\tau)$, אחרת אנו רואים כי התקבלה תלות של פונקצית האוטו-קורלציה בביטויי זמן שאינם $t_1 - t_2$. תחת דרישה זו נקבל -

$$\begin{aligned}
R_X(t_1,t_2) &= (\cos(\omega_0 t_1)\cos(\omega_0 t_2) + \sin(\omega_0 t_1)\sin(\omega_0 t_2))R_I(\tau) + (\sin(\omega_0 t_1)\cos(\omega_0 t_2) - \cos(\omega_0 t_1)\sin(\omega_0 t_2))R_{QI}(\tau) = \\
&\cos(\omega_0(t_1 - t_2))R_I(\tau) + \sin(\omega_0(t_1 - t_2))R_{QI}(\tau)
\end{aligned}$$

משמע $R_X(t_1,t_2) = R_X(\tau)$, $\forall t_1,t_2 \in \mathbb{R}$, והוכחנו את הטענה הבאה –

<u>**טענה:**</u> $X(t) = I(t)\cos(\omega_0 t) + Q(t)\sin(\omega_0 t)$ יהיה תהליך W.S.S **אם ורק אם** $I(t),Q(t)$ תהליכים אקראיים J.W.S.S שפונקציות האוטו-קורלציה שלהם מקיימות $R_I(\tau) = R_Q(\tau)$ ופונקציות הקרוס-קורלציה ביניהם מקיימות $R_{IQ}(-\tau) = -R_{IQ}(\tau)$.

תהליך וינר (*Wiener Process*)

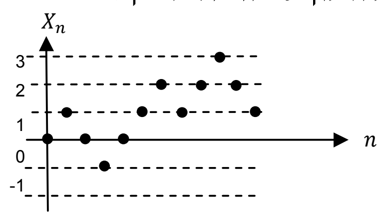
תהליך וינר הוא תהליך גאוסי עם תוחלת 0 ושונות שעולה לינארית בזמן. הוא מייצג תופעה פיזיקלית של דיפוזיה. נפתח את התכונות שלו בשלבים.

<u>הגדרת תהליך אקראי וינר</u>

נגדיר מהו תהליך אקראי וינר בשלבים:

1. "הילוך שיכור": $X_n = X_{n-1} + D_n$, $X_0 = 0 \quad w.p\ \ 1$, $D_n = \begin{cases} 1 & w.p\ 1/2 \\ -1 & w.p\ 1/2 \end{cases}$ $(i.i.d)$ בת"ס ב- X_0.

תזכורת: פונקציית מדגם אופיינית של הילוך שיכור נראית כך:



$E[X_n] = 0$
$Var[X_n] = E[X_n^2] = n$

$$R_X(n,m) = E[X_n \cdot X_m] \underbrace{=}_{\text{הנחה } n>m} E[(X_m + \textstyle\sum_{k=m+1}^{n} D_k) \cdot X_m] = E[X_m \cdot X_m] = m$$

(במעבר הלפני האחרון השתמשנו בכך ש- X_m בת"ס ב- D_k כאשר $k > m$).

אם מאחדים את המקרים ($n > m$, $n < m$, $n = m$) מקבלים:

$R_X(n,m) = \min\{n,m\}$

2. מעבר לזמן רציף (דגימה מהירה עם צעדים קטנים): נגדיר s - גודל הצעד של השיכור; Δ - מרווח הזמן בין צעד לצעד.

$W_\Delta(t) = s \cdot \sum_{n=1}^{\lfloor t/\Delta \rfloor} D_n$

בנוסף נדרוש כי יתקיים הקשר $s^2 = \alpha \cdot \Delta$ שיבטיח עליית שונות לינארית בזמן.

$E[W_\Delta(t)] = 0$
$$Var[W_\Delta(t)] = Var\left(s \cdot \textstyle\sum_{n=1}^{\lfloor t/\Delta \rfloor} D_n\right) = s^2 \cdot Var\left(\textstyle\sum_{n=1}^{\lfloor t/\Delta \rfloor} D_n\right) = s^2 \cdot \textstyle\sum_{n=1}^{\lfloor t/\Delta \rfloor} Var(D_n)$$
$$\underbrace{=}_{t \text{ כפולה שלמה של } \Delta} s^2 \cdot {}^{t}/_{\Delta} = \alpha \cdot t$$

$$R_{W_\Delta}(t_1,t_2) \underbrace{=}_{t \text{ כפולות שלמות של } \Delta} \alpha \cdot \min\{t_1,t_2\}$$

3. מעבר לערכים רציפים (ע"י השאפת קצב הדגימה ל- ∞):

$$\lim_{\Delta\to 0} W_\Delta(t) = \lim_{\Delta\to 0} \underbrace{\sqrt{\alpha \cdot \Delta}}_{s} \cdot \textstyle\sum_{n=1}^{\lfloor t/\Delta \rfloor} D_n \underbrace{=}_{N=t/\Delta} \lim_{\Delta\to 0} \sqrt{\alpha \cdot t} \cdot \frac{1}{\sqrt{N}} \textstyle\sum_{n=1}^{N} D_n$$

לכן, לפי משפט הגבול המרכזי:

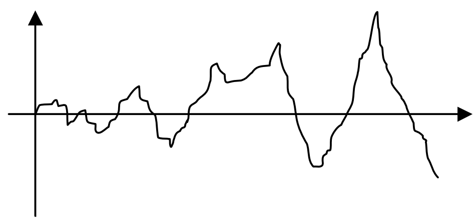
$W(t) = \lim_{\Delta\to 0} W_\Delta(t) \sim N(0, \alpha \cdot t)$

כל דגימה שואפת להיות גאוסית.
האם $W(t)$ תהליך אקראי גאוסי? ניזכר כי כדי שתהליך אקראי יהיה גאוסי נדרוש כי כל קבוצה של דגימות שלו תהיה וקטור גאוסי. היות ש- $W_\Delta(t)$ הוא תהליך עם "תוספות בת"ס" לכל Δ, אזי גם $W(t)$ תהליך עם "תוספות בת"ס". בפרט,

$$W(t_1) \perp\!\!\!\perp W(t_1,t_2) \triangleq W(t_2)-W(t_1)$$

לפיכך $W(t_1,t_2), W(t_1)$ גאוסים במשותף, ומכאן שגם כל זוג קומבינציות לינאריות שלהם גאוסים במשותף. בפרט $W(t_2), W(t_1)$.

פונקצית מדגם אופיינית של תהליך וינר:



<u>ניתוח תהליך הנגזרת של תהליך וינר $W(t)$</u>

נתבונן בתהליך $W_\delta(t)=\frac{W(t+\delta)-W(t)}{\delta}$. נבחין כי $\frac{dW}{dt}=\lim_{\delta\to 0} W_\delta(t)$, וכי זהו תהליך גאוסי גם כן.

במקום לאפיין את התהליך $\frac{dW}{dt}$ (כלומר, את הגבול) ישירות, נאפיין את $W_\delta(t)$ ($\delta>0$), ונשתמש בגבול של המאפיינים (תוחלת ואוטו קורלציה).

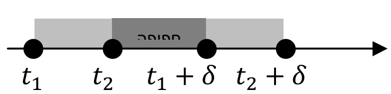
<u>תוחלת</u>: $E[W_\delta(t)]=0 \quad \forall t,\delta$.

<u>פונקצית אוטו קורלציה</u>: $R_{W_\delta}(t_1,t_2)=E[W_\delta(t_1)\cdot W_\delta(t_2)]$

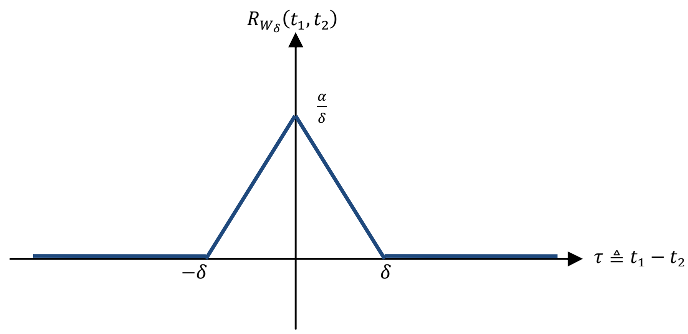
נחלק את חישוב פונקצית האוטו קורלציה ל-2 מקרים:
1. נשים לב ש $W_\delta(t)$ מתאר דגימות של W בזמנים $(t,t+\delta)$. לכן, אם $|t_2-t_1|>\delta$, הדגימות $W_\delta(t_1)$ ו- $W_\delta(t_2)$ מתארות קטעים זרים. לפי תכונת "תוספות בת"ס" של תהליך אקראי וינר, הדגימות $W_\delta(t_1)$ ו- $W_\delta(t_2)$ בת"ס. לכן:

$$R_{W_\delta}(t_1,t_2)=E[W_\delta(t_1)\cdot W_\delta(t_2)]=E[W_\delta(t_1)]\cdot E[W_\delta(t_2)]=0$$

2. נניח כי $|t_2-t_1|\le\delta$. בנוסף נניח בה"כ כי $t_2>t_1$.



$$R_{W_\delta}(t_1,t_2)=E\left[\frac{W(t_1,t_2)+W(t_2,t_1+\delta)}{\delta}\cdot\frac{W(t_2,t_1+\delta)+W(t_1+\delta,t_2+\delta)}{\delta}\right]$$
$$=\frac{1}{\delta^2}\cdot E[W^2(t_2,t_1+\delta)]=\frac{1}{\delta^2}\cdot\alpha\cdot(\delta-|t_1-t_2|)$$



נשים לב שהשטח מתחת לפונקציה לעיל הוא קבוע (שווה ל- α) בלי תלות ב- δ

עבור $\delta \to 0$ פונקצית האוטו קורלציה מתקרבת לפונקצית הלם: $R_{dW/dt}(\tau) = \alpha \cdot \delta(\tau)$.

<u>רעש לבן</u>

הגדרה: תהליך אקראי בעל תוחלת 0 שפונקצית האוטו קורלציה שלו היא פונקצית הלם נקרא <u>תהליך רעש לבן</u>.

משמעות: תהליך אקראי שבו כל שתי נקודות קרובות כרצוננו הן חסרות קורלציה (כלומר התהליך לא "זוכר" מה קרה רגע קודם, עבור רגע קטן כרצוננו).

אם הן בת"ס נאמר שהתהליך הוא <u>רעש לבן במובן החזק</u>.

שרשרות מרקוב

מרקוביות

תזכורת: תהליך $X(t)$ הוא מרקובי אם (עבור $t_1 < t_2$):

$$Pr(X(t_2) = x|\{X(t), t \leq t_1\}) = Pr(X(t_2) = x|X(t_1))$$

כלומר ההסתברות של "העתיד" בהינתן "ההווה" ו"העבר" שווה להסתברות של "העתיד" בהינתן "ההווה" בלבד. דוגמאות לתהליך מרקובי (רציף) הינן תהליך פואסון ותהליך AR.

<u>הערה</u>: ההתניה בביטוי $Pr(X(t_2) = x|\{X(t), t \leq t_1\})$ כוללת את כל הזמנים הקטנים מ- t_1, ולא רק זמן t מסוים.

הגדרה: <u>שרשרת מרקוב</u> היא תהליך מרקובי בזמן בדיד.

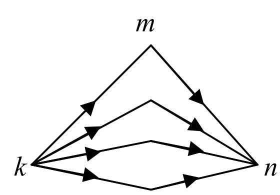
<u>כלל השרשרת</u>:

$$Pr(X_1, X_2, \dots, X_n) = Pr(X_1) \cdot Pr(X_2|X_1) \cdot Pr(X_3|X_2, X_1) \cdot \dots \cdot Pr(X_n|X_1, \dots, X_{n-1}) = \underbrace{=}_{\text{תכונת המרקוביות}} Pr(X_1) \cdot \prod_{i=2}^{n} Pr(X_i|X_{i-1})$$

הגדרה: היות ומספיק להתנות את ההסתברות בערך <u>האחרון</u> של התהליך, נקרא לערך זה ה"<u>מצב</u>" של התהליך.

<u>נוסחת צ'אפמן קולמוגורוב</u>: נניח $k < m < n$:

$$\Pr(X_n = x_n|X_k = x_k) = \sum_{x_m} Pr(X_n = x_n|X_m = x_m) \cdot Pr(X_m = x_m|X_k = x_k)$$



נוכיח:

$$\Pr(X_n = x_n|X_k = x_k) \underbrace{=}_{\text{נוסחת ההסתברות השלמה}} \sum_{x_m} Pr(X_n = x_n, X_m = x_m|X_k = x_k) =$$

$$\underbrace{=}_{\text{חוק בייס}} \sum_{x_m} Pr(X_n = x_n|X_m = x_m, X_k = x_k) \cdot Pr(X_m = x_m|X_k = x_k) =$$

$$\underbrace{=}_{\text{מרקוביות}} \sum_{x_m} Pr(X_n = x_n|X_m = x_m) \cdot Pr(X_m = x_m|X_k = x_k)$$

הגדרה: שרשרת מרקובית תיקרא <u>הומוגנית</u> אם:

$$\forall n: \quad Pr(X_n = a|X_{n-1} = b) = Pr(X_1 = a|X_0 = b)$$

כלומר, פילוג המעבר הוא קבוע בזמן.

הגדרה: שרשרת הומוגנית תיקרא <u>דיסקרטית</u> אם:

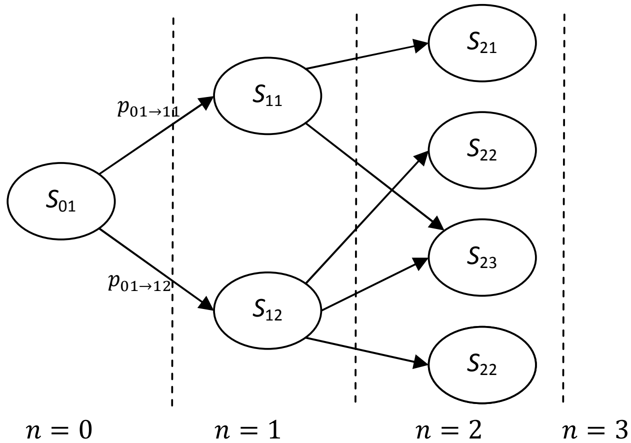
$$X_n \in \{0,1,2, \dots, J_n\}$$

כלומר יש לה מספר סופי (או אינסוף בר-מניה) של מצבים.

ייצוג גרפי לפילוג שרשרת עם מצבים:

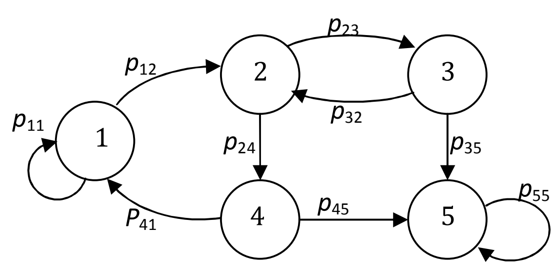
נייצג את הפילוג של שרשרת מרקובית באמצעות גרף שבו הקודקוד (מצב) S_{ni} מייצג את המקרה $X_n = i$. קשתות בגרף יחברו בין מצבים בזמנים עוקבים. בין שני מצבים תופיע קשת אם קיימת הסתברות חיובית למעבר בין שני המצבים. על הקשת נרשום את ערך ההסתברות: $p_{n,i\to n+1,j}$ היא ההסתברות לעבור ממצב i ברגע n למצב j ברגע $n+1$.

גרף אופייני המתאר שרשרת מרקוב יהיה מהצורה הבאה:



נבחין כי סכום הערכים על הקשתות היוצאות מכל מצב הוא 1.

אם השרשרת הומוגנית, אין משמעות לתנועה בציר הזמן. תיאור גרפי של שרשרת מרקוב הומוגנית ייראה כך:



ניתן לתאר את התרשים הנ"ל באמצעות מטריצה בגודל J × J הנקראת <u>מטריצת העבר</u>:

$$\underline{\underline{P}} = \{p_{ij}\}_{i=1,j=1}^{J,J} = \begin{bmatrix} p_{11} & \cdots & p_{1J} \\ \vdots & \ddots & \vdots \\ p_{J1} & \cdots & p_{JJ} \end{bmatrix}$$

כאשר $p_{ij} = Pr(X_{n+1} = j | X_n = i)$; n - זמן, i - מצב נוכחי, j - המצב הבא.

השורה ה i מייצגת את המעברים ממצב i לשאר המצבים, לכן סכום כל שורה הוא 1.

הגדרה: מטריצה שכל איבריה אי-שליליים וסכום כל שורה בה הוא 1 נקראת <u>מטריצה סטוכסטית</u>.

נשים לב כי $\underline{\underline{P}}$ היא מטריצה סטוכסטית.

סיווג מצבים של שרשרת מרקוב הומוגנית

הרבה תופעות מעניינות של שרשרות מרקוב נובעות מכך שבמטריצת המעבר ישנם אפסים, כלומר מעברים אסורים.

1. נגישות: מצב j הוא נגיש (*accessible*) מ- i אם ניתן להגיע מ- i ל- j במספר צעדים (סופי) בגרף.
   נסמן: $i \to j$.

   לדוגמא: בגרף הקודם $1 \to 3$, $5 \nrightarrow 1$.

2. קשירות: מצבים הם קשורים אם הנגישות היא דו צדדית. נסמן: $i \leftrightarrow j$.

   נבחין כי $i \leftrightarrow m, m \leftrightarrow j \Rightarrow i \leftrightarrow j$.

3. מחלקה: קבוצה שלמה של מצבים קשורים.

   לדוגמא: בגרף הקודם המחלקות הן: $\{1,2,3,4\}, \{5\}$.

4. מצב נישנה: מצב הנגיש מכל המצבים הנגישים ממנו.

5. מצב חולף: מצב שאינו נישנה.

6. מחלקה נישנת- מחלקה שכל מצביה נשנים. מחלקה שאינה נשנית היא מחלקה חולפת.

   לדוגמא: המחלקה $\{1,2,3,4\}$ היא מחלקה חולפת, מפני שמצב 5 נגיש מכל אחד מאיברי הקבוצה, אך הם אינם נגישים ממנו. מצב 5 הוא מצב נישנה.

לפני שנעבור לסיווג השישי והאחרון ברשימה ("מחזור של מצב"), נסתכל על האופן שבו מתפתח הפילוג השולי של השרשרת בציר הזמן.

הגדרה: פילוג שולי של שרשרת מרקוב הומוגנית דיסקרטית $X_1, X_2, \ldots, X_n$ ברגע n:

$$\underline{\pi}^{(n)} = [Pr(X_n = 1), Pr(X_n = 2), \ldots, Pr(X_n = J)]$$

אזי לפי משפט ההסתברות השלמה:

$$Pr(X_{n+1} = j) = \sum_{i=1}^{J} Pr(X_n = i) \cdot Pr(X_{n+1} = j | X_n = i) = \sum_{i=1}^{J} Pr(X_n = i) \cdot p_{ij}$$

$$\underline{\pi}^{(n+1)} = \underline{\pi}^{(n)} \cdot \underline{\underline{P}}$$

ומכאן שמתקיים (הוכחה באינדוקציה):

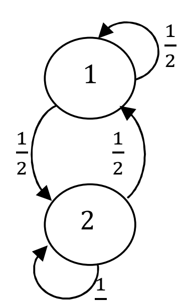
$$\underline{\pi}^{(n+k)} = \underline{\pi}^{(n)} \cdot \underline{\underline{P}}^k$$

כלומר $\underline{\underline{P}}^k$ היא מטריצת המעבר k צעדים קדימה בזמן. בפרט:

$$p_{ij}^{(k)} \triangleq \left[\underline{\underline{P}}^k\right]_{ij} = \Pr\left(X_{n+k} = j | X_n = i\right)$$

דוגמא 1: נתונה שרשרת מרקוב הבאה עם $J = 2$ מצבים (תנאי התחלה: $X_0 = 1$):

$$\underline{\underline{P}} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \underline{\pi}^{(0)} = [1 \quad 0] \qquad \underline{\pi}^{(n)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$



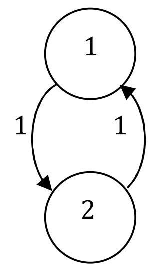
נבחין כי עבור כל תנאי התחלה $\underline{\pi}^{(0)} = [p \quad 1-p]$ עדיין מתקיים

$$\forall n \geq 1: \ \underline{\pi}^{(n)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

דוגמא 2: נתונה שרשרת מרקוב הבאה עם $J = 2$ מצבים (תנאי התחלה: $X_0 = 1$):

$$\underline{\underline{P}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

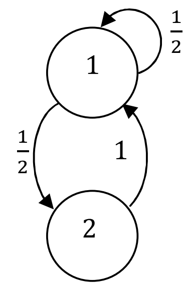
$$\underline{\pi}^{(0)} = [1 \quad 0]$$
$$\underline{\pi}^{(1)} = [0 \quad 1]$$
$$\underline{\pi}^{(2)} = [1 \quad 0]$$



נבחין כי עבור תנאי התחלה $\underline{\pi}^{(0)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ מקבלים

$$\forall n \geq 0: \ \underline{\pi}^{(n)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

דוגמא 3: נתונה שרשרת מרקוב הבאה עם $J = 2$ מצבים (תנאי התחלה: $X_0 = 1$):



$$\underline{\pi}^{(0)} = [1 \quad 0]$$
$$\underline{\pi}^{(1)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$
$$\underline{\pi}^{(2)} = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$
$$\underline{\pi}^{(3)} = \begin{bmatrix} \frac{5}{8} & \frac{3}{8} \end{bmatrix}$$
$$\underline{\pi}^{(4)} = \begin{bmatrix} \frac{11}{16} & \frac{5}{16} \end{bmatrix}$$
$$\vdots$$
$$\underline{\pi}^{(\infty)} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

נבחין כי מתקיים: $Pr(X_{n+1} = 2) = \frac{1}{2} Pr(X_n = 1)$ (האיבר הימני בכל שורה הוא מחצית האיבר השמאלי בשורה שמעליה).

בנוסף מתקיים $Pr(X_{n+1} = 2) = 1 - Pr(X_{n+1} = 1)$ (הסתברות משלימה).

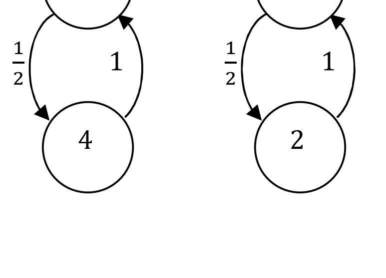
כעת אם נניח שההסתברות $Pr(X_{n+1} = 1)$ מתכנסת לערך q, נקבל כי:
$$lim_{n\to\infty} Pr(X_{n+1} = 1) = lim_{n\to\infty} Pr(X_n = 1) = q$$

לכן מתקיים (נשווה את החישוב הראשון של ההסתברות $Pr(X_{n+1} = 2)$ לחישוב השני, בגבול):
$$\frac{1}{2} q = 1 - q \Rightarrow q = \frac{2}{3}$$

$$lim_{n\to\infty} \underline{\pi}^{(n)} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

דוגמא 4: נתונה שרשרת מרקוב הבאה עם $J = 4$:

$$\underline{\underline{P}} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$



ישנן 2 מחלקות קשירות, ושתיהן נישנות: $\{1,2\},\{3,4\}$

נניח כי נתון תנאי התחלה $X_0 = 1$. אזי:

$$\underline{\pi}^{(0)} = [1 \quad 0 \quad 0 \quad 0]$$
$$\underline{\pi}^{(1)} = \left[\frac{1}{2} \quad \frac{1}{2} \quad 0 \quad 0\right]$$
$$\forall n \geq 1:\ \underline{\pi}^{(n)} = \left[\frac{1}{2} \quad \frac{1}{2} \quad 0 \quad 0\right]$$

נניח כי נתון תנאי התחלה $\underline{\pi}^{(0)} = \left[\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4}\right]$. אזי:

$$\forall n \geq 0:\ \underline{\pi}^{(n)} = \left[\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4}\right]$$

נשים לב שנוכל למיין את הדוגמאות הנ"ל לפי התלות שלהן בתנאי ההתחלה.

הגדרה: נאמר ששרשרת "שוכחת את העבר" אם $lim_{n\to\infty} Pr(X_n = j|X_0 = i) = P_{stat}(j)$ (כלומר, אסימפטוטית אין תלות בעבר).

דוגמא 5: השרשרות בדוגמאות 1,3 שוכחות את העבר. השרשרות בדוגמאות 2,4 אינן שוכחות את העבר.

(ראה בהמשך התייחסות נוספת לנושא זה)

כעת נחזור לסיווג המצבים של שרשרת מרקוב, להגדרה של מחזור.

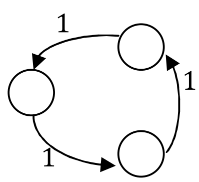
הגדרה: זמני החזרה של מצב i הם הזמנים n שעבורם $Pr(X_n = i|X_0 = i) = p_{ii}^{(n)} > 0$.

הגדרה: מחזור של מצב $d(i)$ הוא המחלק המשותף הגדול ביותר של סדרת זמני החזרה.

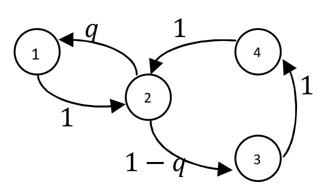
טענה: לכל המצבים באותה המחלקה יש אותו מחזור.

הגדרה: מחלקה היא מחזורית אם $\forall 1 \leq i \leq n:\ d(i) > 1$

דוגמא 6: לכל המצבים בשרשרת הבאה יש מחזור 3:



דוגמא 7: נמצא את המחזורים של המצבים בשרשרת הבאה:



סדרת זמני החזרה של כל המצבים בשרשרת (כולם באותה מחלקה):

$$P_{11}^{(n)} > 0, n = 0,2,4,5,6,7,8,9,10, \dots$$
$$d(i) = 1$$

שרשרת זו אינה מחזורית.

סטאציונאריות של שרשרת מרקוב

ניתן לראות כי אם הוקטור $\underline{\pi}^{(k)} = \underline{\pi}'$ מקיים את המשוואה $\underline{\pi}' = \underline{\pi}' \cdot \underline{\underline{P}}$ אזי $\underline{\pi}^{(n)} = \underline{\pi}'$ (הוכחה באינדוקציה), לכל $n > k$.

כלומר, אם $\underline{\pi}^{(k)} = \underline{\pi}'$ אזי הפילוג השולי של השרשרת קבוע לכל $n \geq k$.
<u>טענה</u>: אם מתקיימת המשוואה $\underline{\pi}^{(0)} = \underline{\pi}^{(0)} \cdot \underline{\underline{P}}$ אזי השרשרת סטאציונארית במובן הצר.

<u>הסבר</u>: מכיוון שהשרשרת הומוגנית (פילוג מעבר קבוע בזמן), על מנת שהפילוג המשותף יהיה קבוע בזמן (והרי זו הדרישה עבור סטאציונאריות במובן הצר) מספיק שהפילוג השולי יהיה קבוע בזמן, מפני ש:

פילוג משותף = פילוג שולי × פילוג מעבר

למעשה, נבחין כי הדרישה בטענה שקולה לכך ש- $\underline{\pi}^{(0)}$ הוא וקטור עצמי שמאלי של המטריצה $\underline{\underline{P}}$, עם ערך עצמי $\lambda = 1$.

"שיכחת העבר"

נחזור לדון כעת בשרשרות ש"שוכחות את העבר": הדרישה לכך ששרשרת מרקוב "שוכחת את העבר" היא ש- $lim_{n\to\infty} Pr(X_n = j | X_0 = i) = P_{stat}(j)$.

לכן אם שרשרת שוכחת את העבר, קיים וקטור $\underline{\pi}^{(\infty)}$ כך ש:
$$lim_{n\to\infty}\underline{\pi}^{(n)} = lim_{n\to\infty}\underline{\pi}^{(0)} \cdot \underline{\underline{P}}^n = \underline{\pi}^{(\infty)}.$$

אם הגבול קיים אזי הוא בהכרח מקיים $\underline{\pi}^{(\infty)} \cdot \underline{\underline{P}} = \underline{\pi}^{(\infty)}$, כלומר הגבול הוא בהכרח וקטור סטאציונארי.

בנוסף, מכיוון שאין תלות בתנאי ההתחלה, ברור שהגבול צריך להתקיים לכל בחירה של וקטור $\underline{\pi}^{(0)}$. כלומר, עבור כל בחירה של $\underline{\pi}^{(0)}$ מתקבל אותו הגבול.

נסמן את וקטורי הבסיס הסטנדרטי מסדר n ב: $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$
$$\underline{e}_i = (0, \dots, 0, \underbrace{1}_{\text{המקום ה} - i}, 0, \dots, 0)$$

וקטורים אלו מייצגים תנאי התחלה דטרמיניסטיים - $x_0 = i, w.p = 1$.
וניזכר שעבור מטריצה $\underline{\underline{A}} \in \mathbb{R}^{n\times n}$ כלשהי, $\underline{e}_i \cdot \underline{\underline{A}}$ היא השורה ה- i במטריצה $\underline{\underline{A}}$. לכן:
$$\text{השורה ה } i - \text{ במטריצה } \underline{\underline{P}}^\infty = lim_{n\to\infty}\underline{e}_i \cdot \underline{\underline{P}}^n = \underline{\pi}^{(\infty)}$$

ומכאן שאם השרשרת שוכחת את העבר, אזי:
$$lim_{n\to\infty}\underline{\underline{P}}^n = \begin{bmatrix} \leftarrow & \underline{\pi}^{(\infty)} & \rightarrow \\ & \vdots & \\ \leftarrow & \underline{\pi}^{(\infty)} & \rightarrow \end{bmatrix}$$

ארגודיות

הגדרה: תהליך אקראי הוא <u>ארגודי</u> (*Ergodic*) אם ניתן ללמוד את הפילוג המשותף המלא שלו (היינו את כל הפרמטרים הסטטיסטיים שלו) מתוך התבוננות על פונקצית מדגם בודדת שלו (בהסתברות 1).

בדרך כלל הכוונה ב"ללמוד" היא למיצוע בזמן. מכאן נובעת ההגדרה הבאה -

הגדרה קונקרטית: תהליך אקראי הוא <u>ארגודי</u> (*Ergodic*) אם ורק אם

$$\frac{1}{N}\sum_{i=1}^{N}\varphi(x_i,x_{i+1},\dots,x_{i+k}) \xrightarrow[N\to\infty]{} E(\varphi(x_1,\dots,x_k))$$

לכל k ולכל פונקציה $\varphi: X^k \to \mathbb{R}$, וכן לכל $i_1, i_2, \dots$.

עבור שרשרת מרקוב הומוגנית דיסקרטית, $X_1, X_2, \dots, X_n$ ($X_i \in \{0,1,2,\dots,J\}$) ארגודיות שקולה ל (2 התנאים יחד):

א. $\hat{\pi}_a = \hat{P}_N(a) = \dfrac{\text{מספר המופעים של } a \text{ ב-} X_1,X_2,\dots,X_N}{N} \to_{N\to\infty} P_{stat}(a)$

ב. $\hat{P}_{ij}^{(N)} \underbrace{=}_{\substack{\text{שערוך הסתברות המעבר} \\ \text{מ-} i \text{ ל-} j \text{ מתוך } N \\ \text{דגימות של השרשרת}}} \dfrac{\text{מספר הפעמים שהופיע } j \text{ אחרי } i \text{ ב-} X_1,X_2,\dots,X_N}{\text{מספר הפעמים שהופיע } i \text{ ב-} X_1,X_2,\dots,X_N} \to_{N\to\infty} P_{ij}$

דוגמא 1: נתונה סדרת הדגימות הבאה של שרשרת מרקוב כלשהי:

$$1,1,1,2,1,1,2,1,1,1,2,2,1 \qquad (N=13)$$

נשערך את $\hat{P}_{N=13}(2)$ ואת $\hat{P}_{i=2,j=2}^{(N=13)}$ לפי הנתון:

$$\hat{P}_{N=13}(2) = \frac{4}{13}$$
$$\hat{P}_{i=2,j=2}^{(N=13)} = \frac{1}{4}$$

<u>טענה</u>: שרשרת מרקוב (הומוגנית) היא ארגודית אם ורק אם היא שוכחת את העבר.

<u>הערה</u>: עבור תהליך אקראי *i.i.d* (מקרה פרטי של שרשרת מרקוב), חוק המספרים הגדולים מבטיח את קיומו של תנאי א לארגודיות. ארגודיות של שרשרת מרקוב כללית - נראה בהמשך.

תזכורת: <u>חוק המספרים הגדולים</u> (*Law of Large Numbers*):

עבור X_i תהליך אקראי *i.i.d* עם שונות סופית מתקיים: $\frac{1}{N}\cdot\sum_{i=1}^{N}X_i \xrightarrow[N\to\infty]{MSE,\ w.p.1} E[X_i]$.

איך חוק המספרים הגדולים מתקשר לתנאי א?

נגדיר $Y_n \triangleq \begin{cases}1 & X_n = a\\ 0 & X_n \neq a\end{cases}$ ("האינדיקטור של a").

הפעלה של חוק המספרים הגדולים נותנת את תנאי א.

דוגמא 2: <u>הטלת קובייה</u>

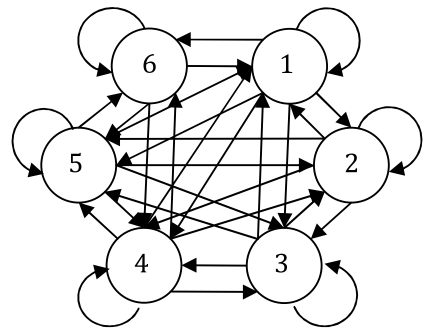
נגדיר שני תהליכים אקראיים (בזמן בדיד המקבלים ערכים בדידים) לפי הטלת קובייה הוגנת באופן הבא:

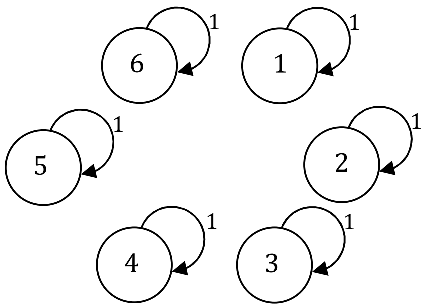
תהליך אקראי 1: בכל רגע מטילים מחדש את הקובייה (באופן בלתי תלוי בהטלה הקודמת).
תהליך אקראי 2: מטילים את הקוביה ברגע 0, והתוצאה שהתקבלה היא הערך שמקבל התהליך האקראי בכל n.

נשים לב כי בשני התהליכים הנ"ל הפילוג השולי הוא $\underline{\pi}^{(n)} = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$ וכי שני התהליכים הנ"ל סטאציונאריים.

נעמוד כעת על ההבדלים בין שני התהליכים:

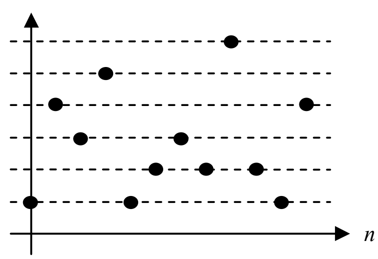
נשרטט דיאגרמות מצבים לכל אחד מהתהליכים:

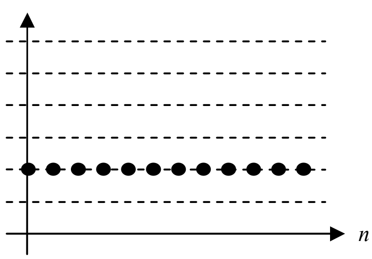




התרשים הימני מתאר את התהליך האקראי הראשון, וההסתברויות על כל החצים שלו הן $\frac{1}{6}$.
התרשים השמאלי מתאר את התהליך האקראי השני.

נצייר פונקציות מדגם אופיינית לכל אחד מהתהליכים (מימין ת"א 1, משמאל ת"א 2):

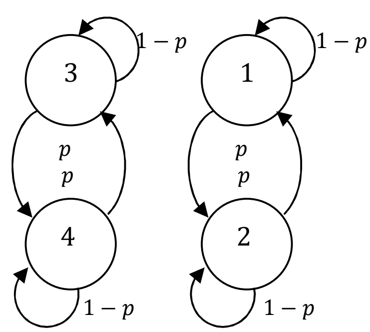




למעשה, התהליך הראשון הוא תהליך ארגודי ואילו התהליך השני אינו ארגודי כי אין שכחה של העבר.

דוגמא 3: בן/בת בוכה/צוחק

בן תמיד נשאר בן, בת תמיד נשארת בת. מצב הרוח של כל אחד מהם יכול להשתנות בהסתברות p או להישאר אותו דבר בהסתברות $1-p$.



נסמן את המצבים: 1 - בן צוחק, 2 - בן בוכה, 3 - בת צוחקת, 4 - בת בוכה.

שרשרת זו אינה ארגודית, מפני שהיא לא שוכחת את העבר (ישנן שתי מחלקות נשנות, ולא ניתן לעבור ביניהן).

התורה של פרון-פרובניניוס (*Perron-Frobenius Theorem*)

הדיון הוא בשרשרות מרקוב הומוגניות סופיות.

זוהי מערכת כללים לקביעת מתי שרשרת מרקוב הומוגנית עם מספר מצבים סופי היא סטאציונארית וארגודית, ומה קורה במקרים שהיא לא. נתייחס לשרשרת עם מטריצת מעבר $\underline{\underline{P}}$.

א. תמיד קיים לפחות פתרון הסתברותי אחד למשוואה $\underline{\pi} = \underline{\pi} \cdot \underline{\underline{P}}$, כלומר ו"ע שמאלי אי שלילי עם ע"ע 1.

ב. אם לשרשרת יש מחלקה נשנית אחת אזי הפתרון הוא יחיד.

ג. אם לשרשרת יש r מחלקות נשנות אזי ישנם r פתרונות בלתי תלויים לינארית. (בהנחה שסידרנו את המצבים בקבוצות לפי המחלקות)

$$\underline{\underline{P}} = \begin{bmatrix} * & * & & & & & & \\ * & * & & 0 & & & 0 & \\ & & * & * & * & & & \\ & 0 & * & * & * & & 0 & \\ & & * & * & * & & & \\ & & & & & * & * & * \\ & 0 & & 0 & & * & * & * \\ & & & & & * & * & * \end{bmatrix} \qquad \begin{aligned} \underline{\pi}_{stat,1} &= (*,*\ ,0,0,0,0,0,0) \\ \underline{\pi}_{stat,2} &= (0,0,*,*,*\ ,0,0,0) \\ \underline{\pi}_{stat,3} &= (0,0,0,0,0,*,*,*) \end{aligned}$$

ד. שרשרת מרקוב היא ארגודית אם כל מצביה שייכים למחלקה נשנית <u>אחת</u> <u>לא מחזורית</u> ($d = 1$).

במקרה זה מתקיים:

$$lim_{n\to\infty}\underline{\underline{P}}^n = \begin{bmatrix} \leftarrow & \underline{\pi}^{(\infty)} & \rightarrow \\ & \vdots & \\ \leftarrow & \underline{\pi}^{(\infty)} & \rightarrow \end{bmatrix}$$

השרשרת סטאציונארית אסימפטוטית לכל $\underline{\pi}^{(0)}$, ומתכנסת ל- $\underline{\pi}^{(\infty)}$.

ה. שרשרת היא <u>ארגודית עם תופעת מעבר</u> (*transient*) אם יש לה מחלקה נשנית אחת לא מחזורית, ומצבים חולפים (כלומר, אם נתחיל במחלקה הנשנית נישאר שם, ואם נתחיל באחד המצבים החולפים אז בהסתברות 1 נעבור למחלקה הנשנית מתישהו).

במקרה זה נוכל ללמוד מפונקצית מדגם בודדת רק על המצבים הנשנים.

<u>מקרים לא ארגודיים:</u>

ו. אם לשרשרת יש 2 מחלקות נשנות או יותר, אזי היא לא ארגודית. $\underline{\underline{P}}^n$ אמנם מתכנסת, אבל שורותיה, בגבול $n \to \infty$, אינן זהות.

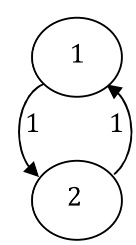
ניתן לראות כי במקרה זה השרשרת אינה שוכחת את העבר: אם מתחילים במחלקה נשנית מסוימת, או מגיעים אליה ממצב חולף - לא ניתן לעבור למחלקה נשנית אחרת. לכן לא ניתן ללמוד את הסטטיסטיקה של כל המצבים מפונקצית מדגם בודדת.

ז. מחלקה נשנית אחת מחזורית $d > 1$:

במקרה זה $\underline{\underline{P}}^n$ אינה מתכנסת (לא קיים גבול), אבל $\underline{\underline{P}}^{n\cdot d}$ כן מתכנסת.

דוגמא 4: נתונה שרשרת מרקוב הומוגנית (נבחין כי $d = 2$):

$$\underline{\underline{P}}^n = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^n$$



הסדרה $\left\{\underline{\underline{P}}^{i}\right\}_{i=1}^{\infty}$ מתנדנדת עם מחזור 2.

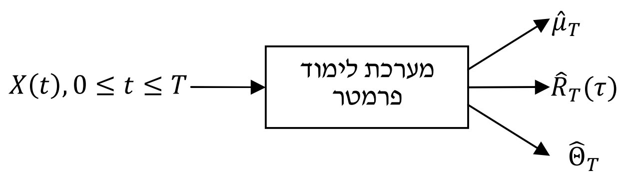
$$\underline{\underline{P}}^{2n}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}^{2n}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}^{n}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\qquad \forall n$$

הסדרה $\left\{\underline{\underline{P}}^{2i}\right\}_{i=1}^{\infty}$ מתכנסת למטריצת היחידה.
השרשרת לא ארגודית.

ארגודיות חלקית (ארגודיות בתוחלת, בקורלציה או בפרמטר אחר)

יהיו $\{X(t),\ 0\le t\le T\}$ תהליך אקראי סטאציונארי ו- Θ פרמטר כלשהו המאפיין את התפלגות התהליך (למשל: α של תהליך אוטו-רגרסיבי, תוחלת של תהליך אקראי, פונקצית אוטו קורלציה וכו').

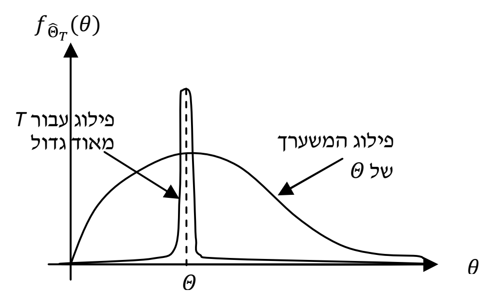
נניח שבידינו מערכת המקבלת בכניסה תהליך אקראי בחלון זמן סופי $\{X(t),\ 0\le t\le T\}$ ומחזירה במוצאה שערוך של הפרמטר Θ (נסמן $\widehat{\Theta}_T$).



<u>הערה</u>:
Θ הוא גודל דטרמיניסטי. $\widehat{\Theta}_T$, המשערך של Θ, הוא משתנה אקראי.

הגדרה: נאמר שמשערך $\widehat{\Theta}_T$ (משערך של הפרמטר Θ) הוא <u>חסר הטיה אסימפטוטית</u> אם
$$E\left[\widehat{\Theta}_T\right]\to_{T\to\infty}\Theta$$

הגדרה: נאמר שמשערך $\widehat{\Theta}_T$ (משערך של הפרמטר Θ) הוא <u>עקבי</u> אם הוא חסר הטיה אסימפטוטית וכן מתקיים:
$$Var\left[\widehat{\Theta}_T\right]\to_{T\to\infty}0$$



המשמעות של הגדרה זו היא שככל ש- T גדל, כך פילוג המשערך הולך ומתקרב להלם בנקודה Θ.

הגדרה: נאמר שהתהליך <u>ארגודי בפרמטר Θ</u> אם קיים משערך $\widehat{\Theta}_T$ שהוא (חסר הטיה אסימפטוטית ו-) עקבי.

ארגודיות בקורלציה:

ישנם 2 משערכים לפונקצית האוטו-קורלציה של $X(t)$:

1. $0 \le t \le T:\ \hat{R}_T(\tau) = \frac{1}{T-\tau} \cdot \int_0^{T-\tau} X(t+\tau) \cdot X(t)dt$ (משערך בלתי מוטה)
2. $0 \le t \le T:\ \hat{R}_T(\tau) = \frac{1}{T} \cdot \int_0^{T-\tau} X(t+\tau) \cdot X(t)dt$ (משערך מוטה)

ארגודיות בתוחלת:

תחת אילו תנאים על $X(t)$ המשערך $\hat{\mu}_T \triangleq \frac{1}{T} \cdot \int_0^T X(t)dt$ הוא עקבי?

חוסר הטיה מובטח לכל T. נוכיח:

$$E[\hat{\mu}_T] = E\left[\frac{1}{T} \cdot \int_0^T X(t)dt\right] \underbrace{=}_{\text{לינאריות התוחלת והאינטגרל}} \frac{1}{T} \cdot \int_0^T E[X(t)]dt = \frac{1}{T} \cdot \int_0^T \mu(t)dt$$

$$\underbrace{=}_{\text{הנחנו סטאציונאריות}} \frac{1}{T} \cdot \int_0^T \mu dt = \mu$$

נותר לבדוק מתי $Var[\hat{\mu}_T] \to_{T\to\infty} 0$.

טענה: $Var[\hat{\mu}_T] \triangleq Var\left[\frac{1}{T} \cdot \int_0^T X(t)dt\right] = \frac{2}{T} \cdot \int_0^T C_X(\tau) \cdot \left(1 - \frac{\tau}{T}\right) d\tau$

מסקנה: תהליך סטאציונארי (במובן הרחב) הוא ארגודי בתוחלת אם

$$(*) \frac{1}{T} \cdot \int_0^T C_X(\tau) \cdot \left(1 - \frac{\tau}{T}\right) d\tau \to_{T\to\infty} 0$$

תנאי סלוצקי: תנאי (*) שקול ל- $\frac{1}{T} \cdot \int_0^T C_X(\tau)d\tau \to_{T\to\infty} 0$.

הגדרה: אם: $\rho(\tau) = \frac{C(\tau)}{C(0)} =$ מקדם הקורלציה בין זוג זמנים במרווח τ $\to_{T\to\infty} 0$ אזי נאמר ש התהליך "שוכח את עצמו" במובן הקורלציה.

נשים לב שתנאי זה מזכיר את התנאי לשכחת העבר של שרשרות מרקוב, אבל הוא יותר חלש מפני שכאן מדובר רק על חוסר קורלציה בעוד שהתנאי לשכחת עבר דורש חוסר תלות ממש בתנאי ההתחלה.

טענה: אם תהליך שוכח את העבר במובן הקורלציה, אזי הוא מקיים את תנאי סלוצקי ולכן ארגודי בתוחלת. ההפך לא בהכרח מתקיים, ראה למשל סינוס עם פאזה אקראית.

הוכחה: לכל $\varepsilon > 0$, נמצא τ_0 כך ש- $C_X(\tau_0) < \varepsilon$ לכל $\tau > \tau_0$ (חייב להימצא כזה $\tau_0 < \infty$ מהתנאי ש- $\rho(\tau) \underset{\tau\to\infty}{\longrightarrow} 0$).

כעת נפרק את האינטגרל בתנאי סלוצקי:

$$\frac{1}{T} \cdot \int_0^T C_X(\tau)d\tau = \frac{1}{T} \cdot \left[\int_0^{\tau_0} C_X(\tau)d\tau + \int_{\tau_0}^T C_X(\tau)d\tau\right] \le$$
$$\le \underbrace{\frac{1}{T} \cdot \int_0^{\tau_0} C_X(\tau)d\tau}_{\underset{T\to\infty}{\rightarrow} 0} + \underbrace{\frac{1}{T} \cdot (T-\tau_0) \cdot \varepsilon}_{\underset{T\to\infty}{\rightarrow} \varepsilon}$$

מכאן שהאינטגרל של סלוצקי קטן מ- ε.

היות ו- ε קטן כרצוננו, תנאי סלוצקי מתקיים. מ.ש.ל.

ספקטרום הספק Power Spectral Density (PSD)

הדיון יתמקד בתהליכים $W.S.S$.

הגדרה: <u>ספקטרום ההספק</u> של תהליך אקראי $W.S.S$ בזמן רציף $X(t)$ בעל פונקצית אוטו קורלציה $R_X(\tau)$ הוא $S_X(\omega)=\mathcal{F}\{R_X(t)\}=\int_{-\infty}^{\infty}R_X(\tau)\cdot e^{-j\omega\tau}\,d\tau$, $-\infty<\omega<\infty$.

נשים לב ש- $R_X(\tau)$ היא פונקצית דטרמיניסטית ולכן כך גם $S_X(\omega)$.

הגדרה: <u>ספקטרום ההספק</u> של תהליך אקראי $W.S.S$ בזמן בדיד X_n בעל פונקצית אוטו קורלציה $R_X(k)$ הוא $S_X(e^{j\Omega})=\mathcal{F}\{R_X(k)\}=\sum_{k=-\infty}^{\infty}R_X(k)\cdot e^{-j\Omega k}$, $-\pi<\Omega<\pi$.

נשים לב ש- $R_X(k)$ היא פונקצית דטרמיניסטית ולכן כך גם $S_X(e^{j\Omega})$.

תזכורת: תכונות של פונקצית אוטו קורלציה $R_X(\tau)$ (עבור תהליך אקראי ממשי):
1. זוגיות: $R_X(\tau)=R_X(-\tau)$
2. $R_X(\tau)$ היא פונקציה אי שלילית מוגדרת. בפרט מתקיים: $\forall\tau:\ R_X(0)\geq|R_X(\tau)|$.

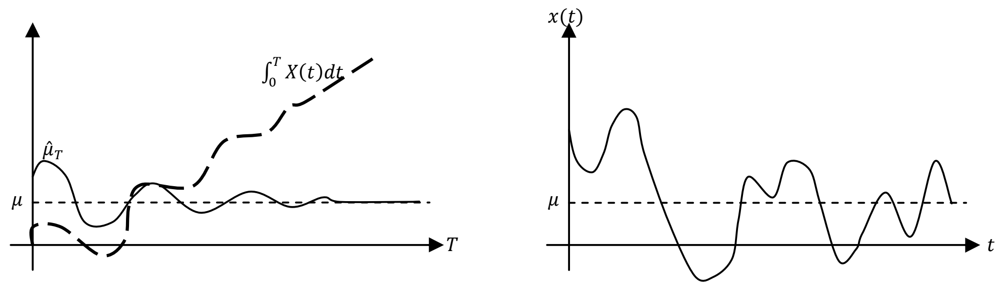
<u>תכונות של ספקטרום הספק $S_X(\omega)$</u> (עבור תהליך אקראי ממשי):
1. $S_X(\omega)$ היא פונקציה ממשית:

$$S_X(\omega)=\int_{-\infty}^{\infty}R_X(\tau)\cdot e^{-j\omega\tau}\,d\tau=\int_{-\infty}^{\infty}R_X(\tau)\cdot(cos\omega\tau+jsin\omega\tau)\,d\tau=$$
$$\underbrace{=}_{\substack{R_X(\tau)\ \text{זוגית}\\ cos\omega\tau\ \text{זוגית}\\ sin\omega\tau\ \text{אי זוגית}}}\int_{-\infty}^{\infty}R_X(\tau)\cdot cos\omega\tau\,d\tau$$

2. $S_X(\omega)$ זוגית ביחס ל- ω (נובע מהזוגיות של $cos\omega\tau$ ומסעיף 1).
3. אי שלילית $\forall\omega:\ \ S_X(\omega)\geq 0$.
<u>הוכחה</u>: ראה דוגמא 2 בסוף השיעור (עמוד 79), או לחלופין - כיוון שהתמרת פורייה של פונקציה אי שלילית מוגדרת (פונקצית האוטו-קורלציה במקרה הזה) היא אי שלילית.

ניזכר ב-2 מושגים מהשיעור הקודם: ארגודיות בתוחלת וארגודיות בקורלציה:

ארגודיות בתוחלת: נניח שהמשערך $\hat{\mu}_T\triangleq\frac{1}{T}\cdot\int_0^T X(t)dt$ ארגודי. אזי הוא חסר הטיה (ללא קשר לארגודיות, כפי שראינו בשיעור הקודם): $E[\hat{\mu}_T]=\mu$; ועקבי: $Var[\hat{\mu}_T]\to_{T\to\infty}0$.



ארגודיות באוטו קורלציה: נתבונן במשערך האוטו קורלציה הבא:

$$M\triangleq E[X^2(t)]=R(0)$$
$$\widehat{M}_T=\frac{1}{T}\int_0^T X^2(t)dt$$

המשערך הנ״ל הוא חסר הטיה: $E\left[\widehat{M}_T\right] = M$. אם התהליך $X(t)$ ארגודי בקורלציה, אזי המשערך הנ״ל הוא עקבי: $Var\left[\widehat{M}_T\right] \to_{T\to\infty} 0$.
נוכל להסיק מהדיון לעיל שלפונקצית מדגם אופיינית של $X(t)$ יש אנרגיה אינסופית, אך מצד שני האינטגרל של $X^2(t)$ בחלון עולה לינארית בגודל החלון (אסימפטוטית), ולכן יש לה הספק חסום.

תזכורת: <u>אותות ומערכות דטרמיניסטיים</u>:

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

$h(t)$ - תגובה להלם שהתרחש ברגע $t = 0$.

<u>התמרת פורייה</u>:

$$x(t) \xleftrightarrow{\text{התמרת פורייה}} \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t) \cdot e^{-j\omega t}\, dt$$
$$h(t) \xleftrightarrow{\text{התמרת פורייה}} \mathcal{F}\{h(t)\} = H(\omega) = \int_{-\infty}^{\infty} h(t) \cdot e^{-j\omega t}\, dt$$

($H(\omega)$ היא פונקציה מרוכבת: כפל בה מורכב מהכפלה בגודל ומהזזת פאזה)

כפל בתדר $\xleftrightarrow{\text{התמרת פורייה}}$ קונבולוציה בזמן

עבור אותות מחזוריים בזמן משתמשים בטורי פורייה.

היינו רוצים ליצור התמרה דומה עבור אותות אקראיים אבל מסתבר שדבר זה בלתי אפשרי. נסביר:

מצד אחד, פונקצית מדגם של תהליך אקראי סטאציונארי איננה דועכת (כלומר איננה בעלת אנרגיה סופית), ולכן אין לה התמרת פורייה. מאידך, יש לה הספק סופי ($\int_0^T X^2(t)dt$ עולה בקירוב באופן לינארי ב- T), אך תנאי זה אינו מספיק עבור קיום טור פורייה: פונקצית מדגם של תהליך אקראי באופן כללי איננה מחזורית (למעט דוגמאות מיוחדות כמו סינוס עם פאזה אקראית), ולכן באופן כללי לא ניתן להציגה באמצעות טור פורייה.

בתור שלב ביניים נתייחס להתמרת פורייה בחלון זמן סופי.

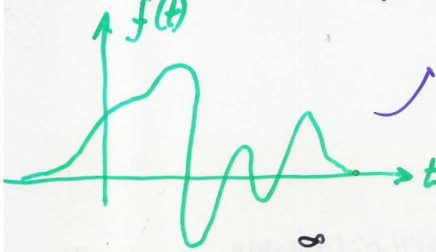
נתבונן בחלון של $X(t)$: $X_T(t) = \begin{cases} X(t) & 0 \le t \le T \\ 0 & \text{אחרת} \end{cases}$

הגדרה: <u>התמרת פורייה בחלון</u> של $X(t)$ היא $X_T(\omega) = \mathcal{F}\{X_T(t)\}$

<u>טענה</u>: יהי $X(t)$ תהליך אקראי סטאציונארי במובן הצר. אזי הפאזה של $X_T(\omega)$ היא משתנה אקראי אחיד: $\measuredangle X_T(\omega) \sim Unif[-\pi, \pi]$.

<u>הסבר</u>: נגריל פונקציות מדגם של $X(t)$. ההסתברות לקבל פונקצית מדגם כלשהי של $X(t)$ שווה להסתברות לקבל פונקצית מדגם זהה עד כדי הזזת זמן. הזזת זמן שקולה לפאזה לינארית במישור התדר. מכאן שעבור תדר נתון (פרט לתדרים מאוד קטנים) הזזת זמן שקולה להזזת פאזה. לכן ההסתברות לקבל פאזה מסוימת של $X_T(\omega)$ שווה להסתברות לקבל כל פאזה אחרת. הערכים האפשריים לפאזה הם $[-\pi, \pi]$, ומכאן שהפאזה של $X_T(\omega)$ מתפלגת אחיד בתחום זה.

למה תהליך סטציונארי דומה - לְאות סופי או מחזורי?



1. פונקציה דטרמיניסטית עם אנרגיה סופית

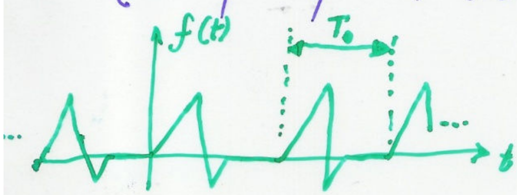
$$\text{אנרגיה} = \int_{-\infty}^{\infty} f^2(t)\,dt < \infty$$

+ התמרת פוריה: $F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t}\,dt$

+ $F^2(\omega)$ מייצג צפיפות האנרגיה ליחידת תדר.

+ משפט פרסבל: $\frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega = \int_{-\infty}^{\infty} f^2(t)\,dt = \text{סה"כ אנרגיה}$

2. פונקציה דטרמיניסטית מחזורית (פונקציה עם "הספק סופי"):



$$\text{אנרגיה} = \int_0^T f^2(t)\,dt \propto T$$

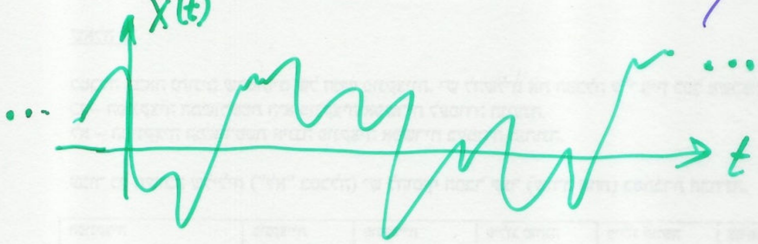
+ טור פוריה: $\omega_0 = \frac{2\pi}{T_0}$; $F_k = \int_0^{T_0} f(t) e^{-jk\omega_0 t}\,dt \,, \quad k = 0,1,2,\ldots$

⇐ התמרת פוריה מורכבת מהלמים בהרמוניות: $F(\omega) = \sum_k F_k \cdot \delta(\omega - k\omega_0)$

+ צפיפות אנרגיה ליחידת תדר:

$$\left|\int_0^T f(t) e^{-j\omega t}\,dt\right|^2 \begin{cases} \simeq \left|F_k \frac{T}{T_0}\right|^2 \propto T^2, & \text{at } \omega = k\omega_0 \\ \longrightarrow 0, & \text{at } \omega \neq k\omega_0 \end{cases}$$

3. תהליך אקראי סטאציונארי (WSS):



+ הוא לעולם לא דועך (השונות שלו קבועה בזמן)

⇐ בדומה לפונקציה מחזורית

⇐ הספק חיובי: $\underset{\text{ממוצעת}}{\text{אנרגיה}} = E\left\{\int_0^T x^2(t)\,dt\right\} \propto T$

+ מצד שני הוא אינו מחזורי

⇐ אין הספק חיובי בתדרים דיסקרטיים, אלא צפיפות הספק ליחידת תדר.

⇐ הנירמול "הנכון" של $|\text{התמרת פוריה}|^2$ גדול מ-1 וקטן מ-T^2 ...

אז למה הוא יותר דומה? ...

הוא באמצע!

הנירמול "הנכון"

$$E\left\{\frac{1}{T}\underbrace{\left|\int_0^T x(t)e^{-j\omega t}\,dt\right|^2}_{\text{צפיפות אנרגיה ליחידת תדר}}\right\} \xrightarrow[T\to\infty]{} \text{קבוע} \triangleq \text{power spectral density}$$

אוטוקורלציה זמנית, פריודוגראמה וספקטרום הספק

* **חלון צפיה סופי**: $X_T(t) = \begin{cases} X(t), & 0 \le t \le T \\ 0, & \text{אחרת} \end{cases}$

* **אוטוקורלציה זמנית** (המשערך המוטה):

$$\hat{R}_T(\tau) \triangleq \frac{1}{T}\int_{-\infty}^{\infty} X_T(t+\tau)\cdot X_T(t)\,dt$$

* **התמרת פורייה בחלון**:

$$X_T(\omega) \triangleq F\{X_T(t)\}$$

* **פריודוגראמה**:

$$\hat{S}_T(\omega) \triangleq \frac{1}{T}\left|X_T(\omega)\right|^2$$

* **זהות**: לכל פונקצית מדגם

$$\hat{S}_T(\omega) = F\{\hat{R}_T(\tau)\}$$

תכונות אסימפטוטיות לתהליך WSS ארגודי:

* אם $X(t)$ ארגודי באוטוקורלציה אזי הגבול הבא מתקיים במובן ריבועי:

$$\hat{R}_T(\tau) \xrightarrow[T\to\infty]{} R(\tau) \qquad \forall \tau$$

כאשר $R(\tau) =$ פ. האוטוקורלציה של $X(t)$.

גבולים
אך דומים..

* אם $X(t)$ מקיים את תנאי (*)

$$E\{\hat{S}_T(\omega)\} \xrightarrow[T\to\infty]{} S(\omega) \qquad \forall \omega$$

כאשר $S(\omega) \triangleq F\{R(\tau)\}$ ספקטרום ההספק של $X(t)$.

תנאי (*): התנאי הוא $\frac{1}{T}\int_0^T \tau R_X(\tau) \xrightarrow{T\to\infty} 0$

הבחינו: תנאי (*) חזק מתנאי סלוצקי (עבור המקרה שהתוחלת שווה ל-0):
$\frac{1}{T}\int_0^T R_X(\tau) \xrightarrow{T\to\infty} 0$.

ספקטרום הספק של תהליך W.S.S.

* תהליך רציף בזמן:

$$S(\omega) \triangleq F\{R(\tau)\} = \int_{-\infty}^{\infty} R(\tau)\cdot e^{-j\omega\tau}\, d\tau$$

לכל $-\infty < \omega < \infty$.

* תהליך WSS בדיד בזמן $X_0, X_1, X_2, \ldots$

$$S(e^{j\Omega}) \triangleq F\{R(k)\} = \sum_{k=-\infty}^{\infty} R(k)\cdot e^{-j\Omega k}$$

לכל $-\pi \le \Omega \le \pi$,

* דוגמאות:

א. רעש לבן:

ב. קבוע אקראי:

$X(t) = A \quad \forall t$

("הדלת קורית בנדנוד")

ג. סינוס עם פאזה:

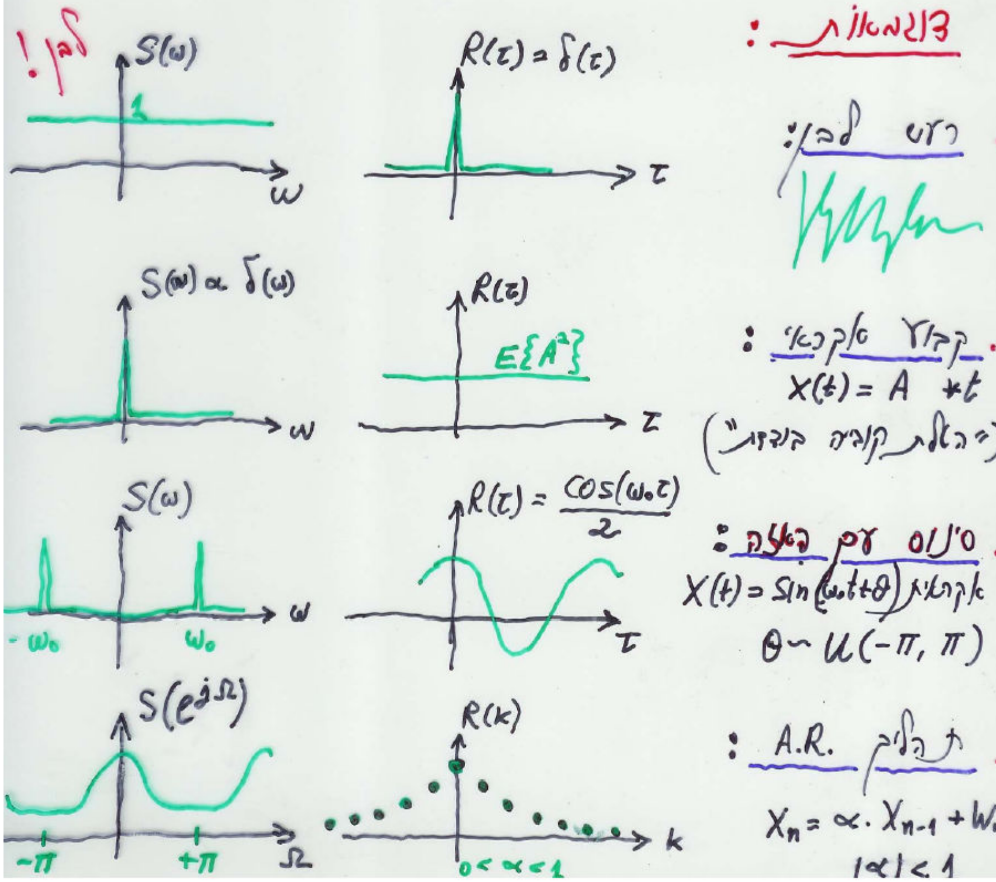
$X(t) = \sin(\omega_0 t + \theta)$ אקראית

$\theta \sim U(-\pi, \pi)$

ד. תהליך A.R.:

$X_n = \alpha \cdot X_{n-1} + W_n$

$|\alpha| < 1$



מעבר תהליך אקראי דרך מסננת לינארית קבועה בזמן

תזכורת: <u>מעבר משתנה אקראי דרך מערכת לינארית $Y = a \cdot X$</u>:

נתונה סטטיסטיקה מסדר שני של X: $\eta_X = E[X],\ R_X = E[X^2]$.

מהי הסטטיסטיקה מסדר שני של המוצא ושל הכניסה-מוצא במשותף?

$$\eta_Y = E[Y] = E[a \cdot X] = a \cdot E[X] = a \cdot \eta_X$$

$$R_Y \triangleq E[Y^2] = E[(a \cdot X)^2] = a^2 \cdot E[X^2] = a^2 \cdot R_X$$

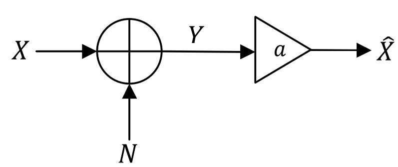
$$R_{XY} \triangleq E[XY] = E[X \cdot (a \cdot X)] = a \cdot E[X^2] = a \cdot R_X$$

<u>שערוך אופטימלי במובן $MMSE$</u>: נתונה סטטיסטיקה מסדר שני של רעש אדטיבי N: $\eta_N = 0,$
$R_N = E[N^2]$.

<u>הערה</u>: המונח "רעש אדטיבי" מתייחס לשני מקרים שונים:
1. הרעש N והאות X חסרי קורלציה ($E[(X-\eta_X)\cdot(N-\eta_N)] = 0$).
2. הרעש N והאות X בת"ס.

במקרה זה נשתמש במונח "רעש אדטיבי" לפי ההגדרה הראשונה.

נרצה לשערך את X לפי Y בצורה אופטימלית במובן MSE, דהיינו למצוא a כך שתוחלת השגיאה הריבועית היא מינימלית.



עבור קשר כללי בין X לY -

$$MSE(a) = E\left[\left(\hat{X} - X\right)^2\right] = E[(a \cdot Y - X)^2] = a^2 R_y - 2aR_{yx} + R_x$$

$$\frac{dMSE}{da} = 2a \cdot R_y - 2R_{yx} = 0 \Longrightarrow a_{opt} = \frac{R_{yx}}{R_y}$$

ובמקרה האדטיבי שלנו -

$$MSE(a) = E[(\hat{X} - X)^2] = E[(a \cdot Y - X)^2] = E[(a \cdot (X + N) - X)^2] =$$
$$= E\left[\left((a-1) \cdot X + a \cdot N\right)^2\right] = (a-1)^2 \cdot E[X^2] + \underbrace{2 \cdot (a-1) E[X \cdot N]}_{=0} + a^2 \cdot E[N^2] =$$

רעש אדטיבי
בעל תוחלת 0

$$= (a-1)^2 \cdot R_X + a^2 \cdot R_N$$

$$\frac{dMSE}{da} = 2 \cdot (a-1) \cdot R_X + 2 \cdot a \cdot R_N = 0 \Longrightarrow a_{opt} = \frac{R_X}{R_X + R_N} \rightarrow \begin{cases} 1 & R_X \gg R_N \\ \frac{R_X}{R_N} \rightarrow 0 & R_X \ll R_N \end{cases}$$

תזכורת: <u>מעבר וקטור אקראי דרך מערכת לינארית $Y = A \cdot X$</u>:

($\underline{X}$ וקטור אקראי בגודל $m \times 1$, $\underline{Y}$ וקטור אקראי בגודל $n \times 1$, $\underline{\underline{A}} \in \mathbb{R}^{n \times m}$ (מטריצה דטרמיניסטית).

נתונה סטטיסטיקה מסדר שני של $\underline{X}$: $\eta_{\underline{X}} = E\left[\underline{X}\right],\ R_{\underline{X}} = E\left[\underline{X}\cdot\underline{X}^T\right]$.

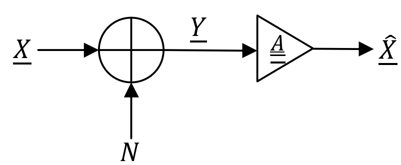
מהי הסטטיסטיקה מסדר שני של המוצא ושל הכניסה-מוצא במשותף?

$$\eta_{\underline{Y}} = E\left[\underline{Y}\right] = E\left[\underline{\underline{A}}\cdot\underline{X}\right] = \underline{\underline{A}}\cdot E\left[\underline{X}\right] = \underline{\underline{A}}\cdot\eta_{\underline{X}}$$

$$R_{\underline{Y}} \triangleq E\left[\underline{Y}\cdot\underline{Y}^T\right] = E\left[\left(\underline{\underline{A}}\cdot\underline{X}\right)\cdot\left(\underline{\underline{A}}\cdot\underline{X}\right)^T\right] = E\left[\underline{\underline{A}}\cdot\underline{X}\cdot\underline{X}^T\cdot\underline{\underline{A}}^T\right] =$$
$$= \underline{\underline{A}}\cdot E\left[\underline{X}\cdot\underline{X}^T\right]\cdot\underline{\underline{A}}^T = \underline{\underline{A}}\cdot R_{\underline{X}}\cdot\underline{\underline{A}}^T$$

$$R_{\underline{Y}\,\underline{X}} \triangleq E\left[\underline{Y}\cdot\underline{X}^T\right] = E\left[\underline{\underline{A}}\cdot\underline{X}\cdot\underline{X}^T\right] = \underline{\underline{A}}\cdot E\left[\underline{X}\cdot\underline{X}^T\right] = \underline{\underline{A}}\cdot R_{\underline{X}}$$

גם במקרה זה ניתן לתאר בעיה של שערוך אופטימלי (כעת $\underline{N}$, הרעש אדטיבי, הוא וקטור בגודל $m\times 1$)



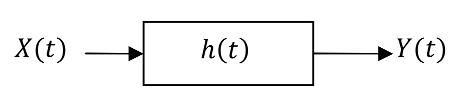
כלומר $\underline{\hat{X}}=\underline{\underline{A}}\underline{Y}$.

$$MSE\left(\underline{\underline{A}}\right) = E\left[\left\|\underline{\hat{X}}-\underline{X}\right\|^2\right] = traceE\left[\left(\underline{\hat{X}}-\underline{X}\right)\left(\underline{\hat{X}}-\underline{X}\right)^t\right] = trE\left[\left(\underline{\underline{A}}\underline{Y}-\underline{X}\right)\left(\underline{\underline{A}}\underline{Y}-\underline{X}\right)^t\right] =$$
$$= tr\left\{\underline{\underline{A}}R_{\underline{y}}\underline{\underline{A}}^t - \underline{\underline{A}}R_{\underline{yx}} - R_{\underline{xy}}\underline{\underline{A}}^t - R_{\underline{X}}\right\} \Rightarrow \underline{\underline{A}}^{opt} = R_{\underline{xy}}R_{\underline{y}}^{\ -1}$$

כאשר המטריצה הנ"ל הינה מסדר mXn.

לאחר שתי התזכורות לעיל, נעבור לנושא העיקרי של השיעור :

נתבונן במעבר תהליך אקראי $X(t)$ דרך מסננת לינארית קבועה בזמן בעלת תגובה להלם $h(t)$.



נרצה למצוא קשר בין כניסת המערכת $X(t)$ לבין מוצאה $Y(t)$. דהיינו נבטא את הסטטיסטיקה מסדר שני של מוצא המערכת ושל הכניסה והמוצא במשותף, ע"י הסטטיסטיקה מסדר שני של כניסת המערכת וע"י תגובת ההלם של המערכת.

נחלק את הדיון לשני מקרים :

1. $X(t)$ הוא תהליך אקראי כלשהו (אינו בהכרח סטאציונארי)
2. $X(t)$ הוא תהליך אקראי סטאציונארי במובן הרחב ($W.S.S$).

<u>הערות</u> :

1. המקרה השני הוא מקרה פרטי של המקרה הראשון.
2. מכיוון שאנו מחשבים סטטיסטיקה מסדר שני של המוצא ושל הכניסה-מוצא במשותף, לפי סטטיסטיקה מסדר שני של הכניסה, מספיק לדרוש במקרה השני כי $X(t)$ הוא $W.S.S$. תהליכים סטאציונאריים במובן הצר הם גם סטאציונאריים במובן הרחב, ולכן הדיון תקף לשני המובנים.

<u>$X(t)$ תהליך אקראי שאינו בהכרח סטאציונארי:</u>

$$Y(t) = h(t) * X(t) = \int_{-\infty}^{\infty} h(\alpha) \cdot X(t-\alpha) \cdot d\alpha$$

תוחלת המוצא:

$$\eta_Y(t) \triangleq E[Y(t)] = E\left[\int_{-\infty}^{\infty} h(\alpha) \cdot X(t-\alpha) \cdot d\alpha\right] \underbrace{=}_{\text{לינאריות התוחלת והאינטגרל}} \int_{-\infty}^{\infty} h(\alpha) \cdot E[X(t-\alpha)] \cdot d\alpha =$$
$$= \int_{-\infty}^{\infty} h(\alpha) \cdot \eta_X(t-\alpha) \cdot d\alpha = h(t) * \eta_X(t)$$

פונקצית קרוס קורלציה כניסה-מוצא:

$$R_{XY}(t_1,t_2) \triangleq E[X(t_1) \cdot Y(t_2)] = E\left[X(t_1) \cdot \int_{-\infty}^{\infty} h(\alpha) \cdot X(t_2-\alpha) \cdot d\alpha\right] =$$
$$= E\left[\int_{-\infty}^{\infty} h(\alpha) \cdot X(t_1) \cdot X(t_2-\alpha) \cdot d\alpha\right] \underbrace{=}_{\text{לינאריות התוחלת והאינטגרל}} \int_{-\infty}^{\infty} h(\alpha) \cdot E[X(t_1) \cdot X(t_2-\alpha)] \cdot d\alpha =$$
$$= \int_{-\infty}^{\infty} h(\alpha) \cdot R_X(t_1, t_2-\alpha) \cdot d\alpha = h(t) * R(t_1,t)\big|_{t=t_2} \triangleq h(t_2) * R_X(t_1,t_2)$$

בדומה ניתן לקבל גם כי:

$$R_{YX}(t_1,t_2) = h(t_1) * R_X(t_1,t_2)$$

נבחין כי מתקיים:

$$R_{YX}(t_1,t_2) = R_{XY}(t_2,t_1)$$

פונקצית האוטו קורלציה של המוצא:

$$R_Y(t_1,t_2) \triangleq E[Y(t_1) \cdot Y(t_2)] = E\left[\int_{-\infty}^{\infty} h(\alpha) \cdot X(t_1-\alpha) \cdot d\alpha \cdot Y(t_2)\right] =$$
$$= E\left[\int_{-\infty}^{\infty} h(\alpha) \cdot X(t_1-\alpha) \cdot Y(t_2) \cdot d\alpha\right] \underbrace{=}_{\text{לינאריות התוחלת והאינטגרל}} \int_{-\infty}^{\infty} h(\alpha) \cdot E[X(t_1-\alpha) \cdot Y(t_2)] \cdot d\alpha =$$
$$= \int_{-\infty}^{\infty} h(\alpha) \cdot R_{XY}(t_1-\alpha, t_2) \cdot d\alpha = h(t_1) * R_{XY}(t_1,t_2)$$

נציב בתוצאה שהתקבלה את תוצאת חישוב פונקצית קרוס קורלציה כניסה-מוצא:

$$R_Y(t_1,t_2) = h(t_1) * R_X(t_1,t_2) * h(t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_X(t_1-\alpha, t_2-\beta) \cdot h(\alpha) \cdot h(\beta) \cdot d\alpha d\beta$$

<u>נניח כעת כי $X(t)$ הוא סטאציונארי במובן הרחב:</u>

הגדרה: נאמר שתהליכים אקראיים $X(t), Y(t)$ הם <u>סטאציונאריים במשותף במובן הצר</u> אם כל אחד מהם סטאציונארי במובן הצר וכן הפילוג המשותף של כל קבוצת דגימות של $X(t)$ עם כל קבוצת דגימות של $Y(t)$ הוא קבוע ביחס להזזה בזמן.

הגדרה: נאמר שתהליכים אקראיים $X(t), Y(t)$ הם <u>סטאציונאריים במשותף במובן הרחב</u> (ונסמן $J.W.S.S$) אם כל אחד מהם סטאציונארי במובן הרחב וכן פונקצית הקרוס קורלציה $R_{XY}(t_1,t_2)$ תלויה רק בהפרש הזמנים $\tau \triangleq t_1 - t_2$.

<u>טענה</u>: אם $X(t)$ הוא $W.S.S$ בעל תוחלת η_X ופונקצית אוטו קורלציה $R_X(\tau)$, ו $Y(t)$ הוא מוצא מערכת LTI בעלת תגובת הלם $h(t)$ שבכניסתה האות $X(t)$ ($Y(t) = h(t) * X(t)$), אזי $X(t), Y(t)$ הם $J.W.S.S$, ומתקיים:

$$R_{XY}(\tau) = R_X(\tau) * h(-\tau)$$
$$R_{YX}(\tau) = R_{XY}(-\tau) = R_X(-\tau) * h(\tau) \underbrace{=}_{R_X \text{ זוגיות של}} R_X(\tau) * h(\tau)$$
$$R_Y(\tau) = R_X(\tau) * h(-\tau) * h(\tau)$$

הוכחה: נראה שהתוחלת של $Y(t)$ קבועה בזמן, וכן שפונקצית הקרוס קורלציה כניסה-מוצא תלויה רק בהפרש הזמנים. את הטענה כי $R_Y(t_1,t_2)$ תלויה רק בהפרש הזמנים $\tau \triangleq t_1 - t_2$ נשאיר ללא הוכחה

$$\eta_Y(t) = \eta_X(t) * h(t) \underbrace{=}_{W.S.S} \eta_X * h(t) = \eta_X \cdot \int_{-\infty}^{\infty} h(\tau)d\tau = \eta_X \cdot H(\omega = 0)$$

$$\begin{aligned} R_{XY}(t_1,t_2) &\triangleq E[X(t_1)\cdot Y(t_2)] = \cdots = \int_{-\infty}^{\infty} h(\alpha)\cdot R_X(t_1,t_2-\alpha)\cdot d\alpha \\ &= \int_{-\infty}^{\infty} h(\alpha)\cdot R_X(t_1-t_2+\alpha)\cdot d\alpha = \int_{-\infty}^{\infty} h(\alpha)\cdot R_X(\tau+\alpha)\cdot d\alpha = \\ &\underbrace{=}_{\beta=-\alpha} \int_{-\infty}^{\infty} h(-\beta)\cdot R_X(\tau-\beta)\cdot d\beta = h(-\tau) * R_X(\tau) \end{aligned}$$

ייצוג מעבר תהליך אקראי *W.S.S* במערכת *LTI* במרחב פוריה ("ציר התדר")

ננסח את הטענה הקודמת מחדש, הפעם תוך שימוש במושג ספקטרום הספק, במקום פונקצית קורלציה:

טענה: אם $X(t)$ הוא *W.S.S* בעל תוחלת η_X וספקטרום הספק $S_X(\omega) = \mathcal{F}\{R_X(t)\}$, ו $Y(t)$ הוא מוצא מערכת *LTI* בעלת תגובת הלם $h(t)$ שבכניסתה האות $X(t)$ ($Y(t) = h(t) * X(t)$), אזי $X(t), Y(t)$ הם *J.W.S.S*, ומתקיים:

$$S_Y(\omega) = S_X(\omega)\cdot|H(\omega)|^2$$
$$S_{XY}(\omega) = S_X(\omega)\cdot H^*(\omega)$$
$$S_{YX}(\omega) = S_X(\omega)\cdot H(\omega)$$

כאשר $S_{XY}(\omega)$ הינו הקרוס-ספקטרום בין תהליכים X(t),Y(t) שהינם J.W.S.S והיא נתונה על ידי התמרת פורייה של $R_{XY}(\tau)$.

נכונות הטענה נובעת מהוכחתה בציר הזמן.

דוגמא 1: ניתוח תהליך *A.R* לינארי כמעבר של תהליך אקראי דרך מסננת

נתבונן בתהליך $X_n = \alpha\cdot X_{n-1} + W_n$ (W_n תהליך אקראי המתפלג *i.i.d* עם תוחלת 0 ושונות σ_W^2, X_0 בלתי תלוי ב- $W_1, W_2, W_3, \ldots$.

$$W_n \longrightarrow \boxed{h_n} \longrightarrow X_n$$

נראה תחילה כי מתקיים: $h_n = \alpha^n,\ \ n \geq 0$:

h_n היא התגובה לכניסת הלם עם תנאי התחלה 0 של המערכת.

$$\delta_n = 1,0,0,0,\ldots \quad ; h_{-1} = 0$$

$$\begin{aligned} h_0 &= \alpha\cdot h_{-1} + \delta_0 = \alpha\cdot 0 + 1 = 1 \\ h_1 &= \alpha\cdot h_0 + \delta_1 = \alpha\cdot 1 + 0 = \alpha \\ h_2 &= \alpha\cdot h_1 + \delta_2 = \alpha\cdot 1 + 0 = \alpha^2 \\ h_3 &= \alpha\cdot h_2 + \delta_3 = \alpha\cdot\alpha + 0 = \alpha^3 \\ &\vdots \\ h_n &= \alpha\cdot h_{n-1} + \delta_n = \alpha\cdot\alpha^{n-1} + 0 = \alpha^n \end{aligned}$$

(הוכחה באינדוקציה)

פונקצית התמסורת של המערכת היא, אם כך:

$$H(e^{j\Omega}) = \frac{1}{1-\alpha\cdot e^{-j\Omega}}$$

פונקצית האוטו קורלציה של W_n היא:

$$R_W(k)=E[W_{n+k}\cdot W_n]=\begin{Bmatrix} k=0: & E[W_n^2]=\sigma_w^2 \\ k\neq 0: & E[W_{n+k}]\cdot E[W_n]=0 \end{Bmatrix}=\sigma_w^2\cdot\delta(k)$$

ספקטרום ההספק של W_n הוא:

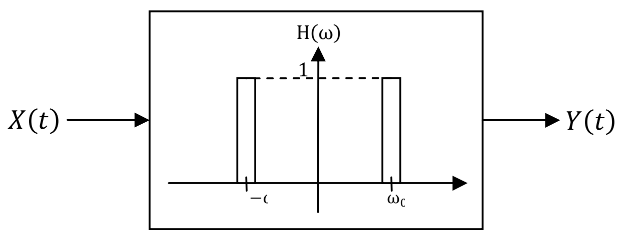
$$S_W\left(e^{j\Omega}\right)=\sigma_w^2$$

אם מתקיים התנאי לסטציונאריות, נוכל כעת לחשב את ספקטרום ההספק של X_n, באמצעות הנוסחא שבטענה:

$$S_X\left(e^{j\Omega}\right)=S_W\left(e^{j\Omega}\right)\cdot|H(\omega)|^2=\sigma_w^2\cdot\left|\frac{1}{1-\alpha\cdot e^{-j\Omega}}\right|^2$$

ומכאן נוכל לחשב את פונקצית האוטו קורלציה של X_n ע"י התמרת פורייה הפוכה.

דוגמא 2: <u>מעבר תהליך אקראי דרך מסננת פס-מעבר (BPF) צרה סביב תדר ω_0</u>



(רוחב המסננת Δ)

$$S_Y(\omega)=\begin{cases} S_X(\omega) & \omega\in\{\omega_0\pm\Delta/2\} \\ 0 & \text{אחרת} \end{cases}$$

נבחין כי $Y(t)$ הוא תהליך אקראי, לכן $R_Y(\tau=0)=E[Y^2(t)]\geq 0$. נחשב את $R_Y(\tau=0)$ ע"י התמרת פורייה הפוכה של ספקטרום ההספק שלו:

$$R_Y(\tau=0)=\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}S_Y(\omega)d\omega=\frac{1}{2\pi}\cdot\left[\int_{-\omega_0-\Delta/2}^{-\omega_0+\Delta/2}S_X(\omega)d\omega+\int_{\omega_0-\Delta/2}^{\omega_0+\Delta/2}S_X(\omega)d\omega\right]=$$

$$\underbrace{=}_{\text{זוגיות של ספקטרום הספק}}\frac{1}{2\pi}\cdot 2\cdot\int_{\omega_0-\Delta/2}^{\omega_0+\Delta/2}S_X(\omega)d\omega$$

נבחר Δ מספיק קטנה, שכך שנוכל בקירוב לומר כי $S_X(\omega)\cong S_X(\omega_0)$ עבור $\omega\in\{\omega_0\pm\Delta/2\}$:

$$R_Y(\tau=0)=\frac{1}{\pi}\cdot\Delta\cdot S_X(\omega_0)$$

ניזכר כי $R_Y(\tau=0)\geq 0$, ונקבל כי $S_X(\omega_0)\geq 0$.
מכיוון שחישוב זה נכון לכל ω, מתקיים: $S_X(\omega)\geq 0$. (ובכך הוכחנו את תכונת האי-שליליות של ספקטרום ההספק).

דוגמא 3: <u>מעבר זוג תהליכים אקראיים דרך מסננת פס-מעבר (BPF) צרה סביב תדר ω_0</u>

יהיו $X_1(t), X_2(t)$ תהליכים אקראיים J.W.S.S. נערוך ניסוי מחשבתי (דומה לזה שבדוגמא הקודמת) ובו נעביר את $X_1(t)$ במסננת $H_1(\omega)$ ואת $X_2(t)$ במסננת $H_2(\omega)$, כאשר המסננות הינן צרות סרט (N.B.F), כלומר מסננות מהצורה הבאה-

$$H_1(\omega) = H_2(\omega) = \begin{cases} 1 & \omega_0 - \frac{\Delta}{2} \le \omega \le \omega_0 + \frac{\Delta}{2} \\ 0 & else \end{cases}$$

כאשר $\Delta \ll \omega_0$ (צרות הסרט של המסננות).

באופן כללי מתקיים עבור פונקציות הקרוס-ספקטרום $S_{Y_1Y_2}(\omega) = H_1(\omega)S_{X_1X_2}(\omega)H_2^*(\omega)$. נוכיח אותו -

$$R_{Y_1Y_2}(\tau) = E\{Y_1(t)Y_2(t-\tau)\} = E\{(X_1(t)*h_1(t))(X_2(t-\tau)*h_2(t))\} = E\left\{\left(\int_{-\infty}^{\infty} h_1(\alpha)X_1(t-\alpha)d\alpha\right)\left(\int_{-\infty}^{\infty} h_2(\beta)X_2(t-\tau-\beta)d\beta\right)\right\} =$$

$$E\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X_1(t-\alpha)X_2(t-\tau-\beta)h_1(\alpha)h_2(\beta)d\alpha d\beta\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{X_1(t-\alpha)X_2(t-\tau-\beta)\}h_1(\alpha)h_2(\beta)d\alpha d\beta =$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{X_1X_2}(\tau-\alpha+\beta)h_1(\alpha)h_2(\beta)d\alpha d\beta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(R_{X_1X_2}(\tau-\alpha+\beta)h_1(\alpha)d\alpha\right)h_2(\beta)d\beta =$$

$$\int_{-\infty}^{\infty}\left(R_{X_1X_2}*h_1\right)_{t=\tau+\beta} h_2(\beta)d\beta \underset{\substack{\gamma=-\beta \\ d\gamma=-d\beta}}{=} \int_{-\infty}^{\infty}\left(R_{X_1X_2}*h_1\right)_{t=\tau-\gamma} h_2(-\gamma)d\gamma = h_1(\tau)*R_{X_1X_2}(\tau)*h_2(-\tau)$$

מהתמרת פורייה על שני האגפים נקבל $S_{Y_1Y_2}(\omega) = H_1(\omega)*S_{X_1X_2}(\omega)*H_2^*(\omega)$,

מקשר זה נובע כי פונקצית קרוס-ספקטרום בין שני המוצאים של מסננות צרות סרט אלו הינו

$$S_{Y_1Y_2}(\omega) = \begin{cases} S_{X_1X_2}(\omega) & \omega_0 - \frac{\Delta}{2} \le \omega \le \omega_0 + \frac{\Delta}{2} \\ 0 & else \end{cases}$$

ועתה נשים לב כי

$$R_{Y_1Y_2}(\tau=0) \underset{1}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{Y_1Y_2}(\omega)e^{j\omega(\tau=0)}d\omega \underset{2}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} H_1(\omega)S_{X_1X_2}(\omega_0)H_2^*(\omega)d\omega \underset{3}{=} \frac{1}{2\pi}\int_{\omega_0-\frac{\Delta}{2}}^{\omega_0+\frac{\Delta}{2}} S_{X_1X_2}(\omega_0)d\omega_0 \underset{4}{\simeq} \frac{\Delta}{2\pi}S_{X_1X_2}(\omega_0)$$

(1) פונקצית הקרוס-קורלציה היא התמרה הפוכה של הקרוס-ספקטרום עבור תהליך מסוים.
(2) הקשר $S_{Y_1Y_2}(\omega) = H_1(\omega)S_{X_1X_2}(\omega)H_2^*(\omega)$.
(3) הצבה של הביטוי ל $S_{Y_1Y_2}(\omega)$ במקרה הזה.
(4) $\Delta \ll \omega_0$ לכן האינטגרנד קבוע בקירוב.

מהניסוי המחשבתי שערכנו, יחד עם הניסוי המחשבתי מהדוגמא הקודמת נובע שלכל ω_0 ("נחזור על הניסוי פעמים רבות" עבור מסננות צרי סרט המרוכזות סביב תדר ω_0 אחר) מתקיים

$R_{Y_1}(\tau=0)\underset{4}{\simeq}\frac{\Delta}{\pi}S_{X_1}(\omega_0), R_{Y_2}(\tau=0)\underset{4}{\simeq}\frac{\Delta}{\pi}S_{X_2}(\omega_0)$ ו $R_{Y_1Y_2}(\tau=0)\underset{4}{\simeq}\frac{\Delta}{2\pi}S_{X_1X_2}(\omega_0)$.

עבור התהליכים האקראיים $Y_1(t),Y_2(t)$ מתקיים, כפי שראינו בעבר, אי שוויון קושי-שוורץ לפונקציות קרוס-קורלציה –

$$\left|R_{Y_1Y_2}(\tau)\right| \le \sqrt{R_{Y_1}(0)R_{Y_2}(0)} \Leftrightarrow \left|E\left\{Y_1(t)Y_2(t+\tau)\right\}\right| \le \sqrt{E\left\{Y_1^2(t)\right\}E\left\{Y_2^{\,2}(t)\right\}}$$

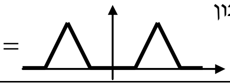
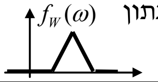
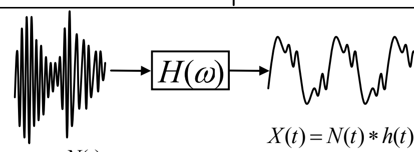
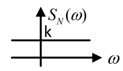
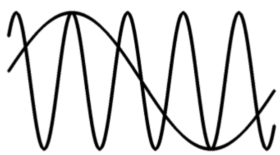
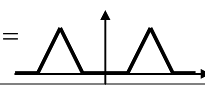
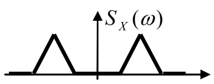
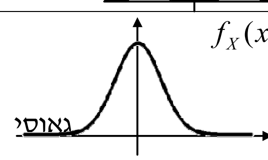
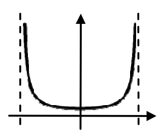
יחד נקבל כי מתקיים שלכל ω_0 ולכל שני תהליכים אקראיים $X_1(t),X_2(t)$ תהליכים אקראיים J.W.S.S - $\left|S_{X_1X_2}(\omega_0)\right| \le \sqrt{S_{Y_1}(\omega_0)S_{Y_2}(\omega_0)}$.

כך למעשה קיבלנו הוכחה לאי שוויון קושי-שוורץ עבור פונקציות הקרוס-ספקטרום –

$$\left|S_{XY}(\omega_0)\right| \le \sqrt{S_X(\omega_0)S_Y(\omega_0)}$$

דוגמא 4: <u>עיצוב ספקטרום בשתי דרכים</u>

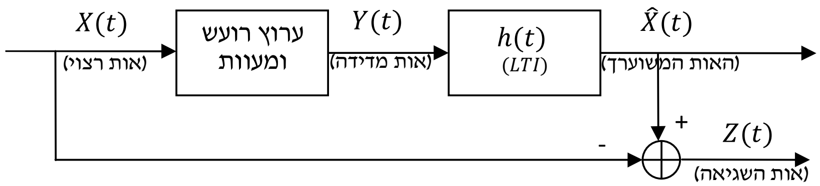
בדומה להשוואה בין תכונות תהליך פואסון ותהליך וינר, נשווה בין שתי דרכים לעיצוב ספקטרום הספק :

	רעש לבן דרך מסננת $H(\omega)$	**אפנון תדר FM (סינוס עם תדר אקראי** W)
	כאשר נתון $\mid H(\omega)\mid^2 =$ 	כאשר נתון $W \sim f_W(\omega)$ ונתון 
תאור ציורי של התהליך	 $N(t)$ רעש גאוסי לבן עם צפיפות ספקטראלית=k 	 $\sin(Wt+\theta)$ where $\theta \sim unif(-\pi,\pi)$
תוחלת	$\mu = 0$	$\mu = 0$
ספקטרום	$S_X(\omega) = k\cdot\mid H(\omega)\mid^2$ נניח $k=1$ אזי נקבל : $S_X(\omega) = \mid H(\omega)\mid^2 =$ 	$S_X(\omega) = \frac{\pi}{2}\left[f_W(\omega) + f_W(-\omega)\right]$ 
פילוג שולי (פילוג דגימה בזמן)	$f_X(x)$  גאוסי	$f_X(x \mid W=\omega_0) \Rightarrow g(\theta) = \sin(\text{קבוע} + \theta)$ $\Rightarrow f_X = \frac{f_\theta}{g'(\theta)} \sim \frac{1}{\cos(\theta)}$ 
ארגודיות	כן (בתנאי שמתקיים תנאי השיכחה המתאים).	לא! (ניתן ללמוד רק את W שהוגרל).
קורלציה זמנית	$\hat{R}_T(\tau) \xrightarrow[T\to\infty]{} F^{-1}\{S_W(\omega)\}$	$\hat{R}_T(\tau) \xrightarrow[T\to\infty]{} \frac{\cos(\omega_0\tau)}{2}$ בהנחה שהתרחש $W=\omega_0$ (כלומר הוגרל תדר ω_0 מסוים).

← כמו בהשוואת Weiner / Poisson, התהליכים דומים בסטטיסטיקה מסדר שני, אבל מאוד שונים!

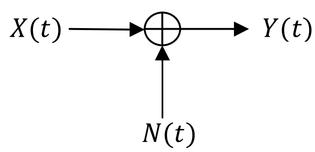
סינון וינר (*Wiener Filter*)

מטרת סינון וינר היא לסנן או להפחית את כמות הרעש והעיוות בתהליך אקראי, כלומר מטרתנו למצוא מסננת "טובה" $h(t)$.



דוגמא 1: <u>דוגמאות לערוץ רועש ומעוות</u>:

א. רעש אדטיבי



ב. מערכת לינארית LTI ואחריה רעש אדטיבי: $Y(t) = X(t) * a(t) + N(t)$

ג. $Y(t) = X^2(t) + N(t)$

נניח שבידינו סטטיסטיקה מסדר שני של התהליכים האקראיים $X(t), Y(t)$:

$$R_X(\tau) \leftrightarrow S_X(\omega), R_Y(\tau) \leftrightarrow S_Y(\omega), R_{XY}(\tau) \leftrightarrow S_{XY}(\omega)$$

נבטא את התלות של $\hat{X}(t)$ באות המדידה $Y(t)$:

$$\hat{X}(t) = h(t) * Y(t) = \int_{-\infty}^{\infty} h(\alpha) \cdot Y(t-\alpha) \cdot d\alpha \underbrace{\cong}_{(1)} \int_{-T}^{T} h(\alpha) \cdot Y(t-\alpha) \cdot d\alpha$$

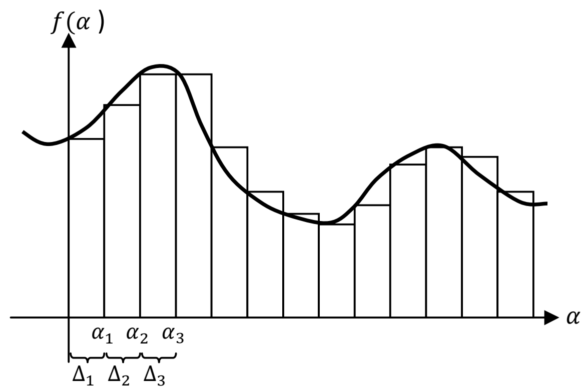
$$\underbrace{\cong}_{(2)} \sum_{i=1}^{N} \underbrace{h(\alpha_i) \cdot Y(t-\alpha_i)}_{f(\alpha_i)} \cdot \Delta_i$$

(1) קירוב קונבולוציה ע"י קונבולוציה בחלון סופי

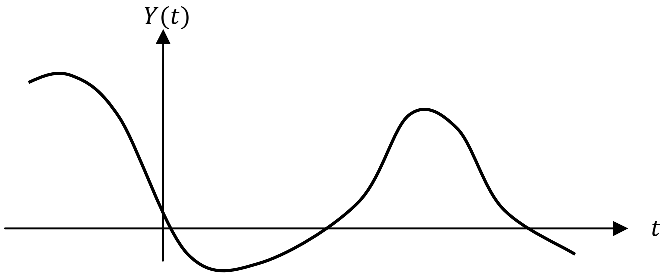
(2) קירוב קונבולוציה ע"י קונבולוציה בזמן בדיד

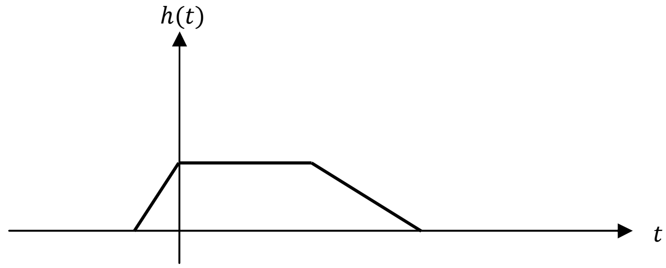
תרשימים המדגימים בצורה גרפית את החישוב שלעיל:

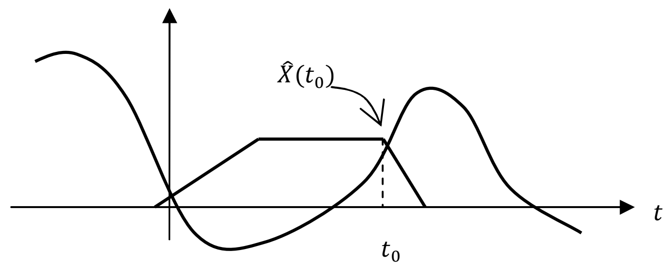
את הקירוב של האינטגרל ע"י סכום ניתן להצדיק ע"י הגדרת אינטגרל רימן, שבו השטח מתחת לפונקציה מקורב ע"י חישוב שטחי מלבנים.



תזכורת: איך קונבולוציה בנקודה t_0 נראית בצורה גרפית:







<u>הערות</u>:

1. הנחת המודל היא ש- $X(t),Y(t)$ הם תהליכים אקראיים $J.W.S.S$.
2. המסננת $h(t)$ אינה בהכרח סיבתית.
3. את השגיאה בשערוך נוכל למדוד לפי מספר קריטריונים. לדוגמא:
   א. שגיאה ריבועית: $MSE = E[Z^2(t)]$
   ב. הסתברות שגיאה: $Pr(Z(t) \neq 0) = Pr(\hat{X}(t) \neq X(t)$
4. אם $X(t),Y(t)$ הם $J.W.S.S$ אזי גם $Z(t) = h(t) * Y(t) - X(t)$ הוא $J.W.S.S$ עם כל אחד מהם. (ראה הוכחה בדיון על מערכת $MIMO$ בהמשך).
5. אנו נתרכז בסינון לינארי <u>טהור</u>: $\hat{X}(t) = h(t) * Y(t)$,

בשונה מסינון לינארי אפיני (סינון לינארי עם תוספת קבוע: $\hat{X}(t) = h(t) * Y(t) + b$), ומסינון לא לינארי: $\hat{X}(t) = g(Y(t), 0 < t < \infty)$.

6. שערוך סקלרי הוא מקרה פרטי (מנוון) של סינון לינארי טהור: $\hat{X}(t) = h \cdot Y(t-\tau)$ (השערוך של X ברגע t הוא הערך ש- Y קיבל τ רגעים קודם לכן, מוכפל בקבוע h, כאשר τ הוא קבוע שנקבע מראש).

מהו h האופטימלי בהינתן τ?

ניזכר כי עבור X,Y משתנים אקראיים, השערוך הסקלרי האופטימלי מקיים:

$$\hat{X} = a \cdot Y$$

$$MSE(a) = min_a E[\underbrace{(a \cdot Y - X)}_{Z}{}^2]$$

לפי עקרון הניצבות מתקיים $E[ZY] = 0$ עבור המשערך הלינארי הטהור האופטימלי. לכן:

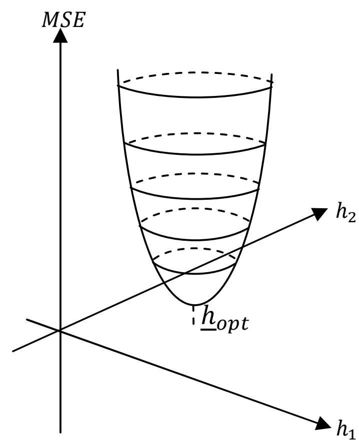
$$0 = E[ZY] = E[(a \cdot Y - X) \cdot Y] = aR_Y - R_{XY}$$

$a_{BLE} = {R_{XY}}/{R_Y}$

מכאן שהמשערך הסקאלרי הלינארי האופטימלי עם השהייה τ עבור תהליכים אקראיים נתון על ידי:

$h_{BLE} = {R_{XY}(\tau)}/{R_Y(0)}$

ניתוח המסננת הלא סיבתית האופטימלית



נחזור כעת לבעיה הכללית. נרצה למצוא את:

$h_{BLE}(t) \triangleq argmin_{h(t)}MSE$

$$MSE = E[Z^2(t)] = E\left[\left(h(t) * Y(t) - X(t)\right)^2\right] \cong$$
$$\cong E\left[\left(\sum_{i=1}^{N} h(\alpha_i) \cdot Y(t-\alpha_i) \cdot \Delta_i - X(t)\right)^2\right]$$

לאחר חישוב התוחלת נקבל תבנית ריבועית של וקטור הדגימות של תגובת ההלם של המסנן:

$\underline{h} = [h_1, h_2, \dots] \triangleq [h(\alpha_1), h(\alpha_2), \dots]$

זהו פרבולואיד שנקודת המינימום שלו היא המסננת הרצויה לנו. היות וזוהי פונקציה קמורה עם מינימום יחיד נמצא את נקודת המינימום ע"י גזירה:

$$0 = \frac{\partial MSE}{\partial h_i} = \frac{\partial}{\partial h_i} \cdot E[Z^2(t)] = E\left[\frac{\partial}{\partial h_i} Z^2(t)\right] = E\left[2 \cdot Z(t) \cdot \left(\frac{\partial}{\partial h_i} Z(t)\right)\right] =$$
$$= 2 \cdot E[Z(t) \cdot Y(t-\alpha_i)] \cdot \Delta_i$$

ומכאן ש:

$\forall i: \ \ R_{ZY}(\alpha_i) = 0$

נוכל להסיק (ע"י השאפת N לאינסוף ו- Δ_i ל-0) שעבור המסננת האופטימלית:

$\forall \tau: \ \ R_{ZY}(\tau) = 0$

וזהו בעצם <u>עקרון הניצבות</u> $\{Z(t)\} \perp \{Y(t)\}$ <u>עבור תהליכים אקראיים</u>.

נראה שזהו גם תנאי מספיק למציאת $h_{BLE}(t)$:

$$0 = R_{ZY}(\tau) = E[Z(t) \cdot Y(t-\tau)] = E\left[\left(h(t) * Y(t) - X(t)\right) \cdot Y(t-\tau)\right] =$$
$$= R_{h*Y,Y}(\tau) - R_{XY}(\tau)$$

ניזכר ש (משיעור שעבר):

$R_{h*Y,Y}(\tau) = R_Y(\tau) * h(\tau)$

לכן:

$0 \equiv R_Y(\tau) * h(\tau) - R_{XY}(\tau)$

על מנת לפתור את בעית הדֶה-קונבולוציה (חילוץ של $h(t)$) כדאי לעבור לציר התדר. התמרה של פונקציה שהיא זהותית 0 בזמן, היא זהותית 0 בתדר.

$0 = \mathcal{F}\{R_Y(\tau) * h(\tau) - R_{XY}(\tau)\} = \mathcal{F}\{R_Y(\tau) * h(\tau)\} - \mathcal{F}\{R_{XY}(\tau)\}$

$$S_Y(\omega)\cdot H(\omega)=S_{XY}(\omega)$$

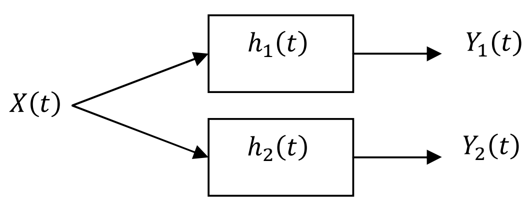
<u>הערה</u>: נשים לב שמאי שוויון קושי שוורץ לתהליכים אקראיים: $S_X(\omega)\cdot S_Y(\omega)\ge |S_{XY}(\omega)|^2$ (נראה בהמשך) נובע שאם $S_Y(\omega)=0$ אזי בהכרח $S_{XY}(\omega)$. לכן אפשר לקבל $\left[\frac{0}{0}\right]$ אבל לא $\left[\frac{*}{0}\right]$.
עבור תדר ω כזה הפתרון האוטומטי הוא ערך כלשהו.

$$H_{BLE}(\omega)=H_{Wiener}(\omega)=\frac{S_{XY}(\omega)}{S_Y(\omega)}\quad\left(\left[\frac{0}{0}\right]\triangleq \text{don't care}\right)$$
$$h_{BLE}(t)=h_{Wiener}(t)=\mathcal{F}^{-1}\left\{\frac{S_{XY}(\omega)}{S_Y(\omega)}\right\}\neq\frac{\mathcal{F}^{-1}\{S_{XY}(\omega)\}}{\mathcal{F}^{-1}\{S_Y(\omega)\}}$$

נבחין כי זהו אותו פתרון שקיבלנו במקרה הסקאלרי, רק לכל תדר בנפרד.
ניתן לפרש זאת כאילו שהאותות $X(t)$ ו- $Y(t)$ ניתנים לפירוק לאותות צרי סרט (לא חופפים) סביב התדרים $\omega_1,\omega_2,\dots,\omega_N$ במרווחים $\Delta_1,\Delta_2,\dots,\Delta_N$ כך שבתדר ω_i האות הרצוי שקול למשתנה אקראי עם מומנט שני $\Delta_i\cdot S_X(\omega_i)$ שנמדד ע"י משתנה אקראי עם מומנט שני $\Delta_i\cdot S_Y(\omega_i)$, ועם קרוס קורלציה ביניהם $\Delta_i\cdot S_{XY}(\omega_i)$.

לפני שנחשב את השגיאה הריבועית הממוצעת שמתקבלת מתוך מסננת וינר שחושבה לעיל, נבצע מספר הרחבות ופרשנויות.

<u>מעבר במערכת SIMO (Single Input Multiple Output)</u>

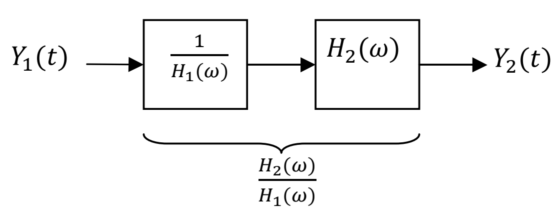


אנו יודעים לחשב סטטיסטיקה מסדר שני של $Y_1(t),Y_2(t)$ בנפרד וביחס ל- $X(t)$.
נרצה לחשב פונקצית אוטו קורלציה בין היציאות:

$$R_{Y_1Y_2}(t_1,t_2)=E[Y_1(t_1)\cdot Y_2(t_2)]=E\left[\left(h_1(t_1)*X(t_1)\right)\cdot\left(h_2(t_2)*X(t_2)\right)\right]=$$
$$=\cdots=h_1(t_1)*R_X(t_1,t_2)*h_2(t_2)=h_1(\tau)*R_X(\tau)*h_2(-\tau)$$

$$S_{Y_1Y_2}(\omega)=\mathcal{F}\left\{R_{Y_1Y_2}(\tau)\right\}=S_X(\omega)\cdot H_1(\omega)\cdot H_2^*(\omega)$$

נבחין כי יכולנו לקבל תוצאה זו גם אם היינו הולכים בכיוון ההפוך דרך המסננת $h_1(t)$, מ- $Y_1(t)$, דרך $X(t)$ ומשם ל- $Y_2(t)$. (נניח $H_1(\omega)\neq 0$)

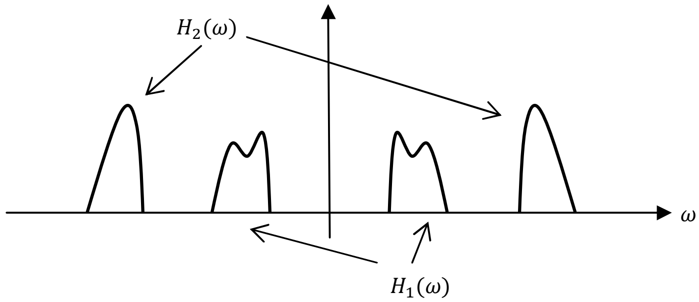


$$S_{Y_1Y_2}(\omega)=S_{\text{מוצא,כניסה}}(\omega)=S_{Y_1}(\omega)\cdot\left[\frac{H_2(\omega)}{H_1(\omega)}\right]^*=S_X(\omega)\cdot|H_1(\omega)|^2\cdot\frac{H_2^*(\omega)}{H_1^*(\omega)}=$$

$$= S_X(\omega)\cdot H_1(\omega)\cdot H_2^*(\omega)$$

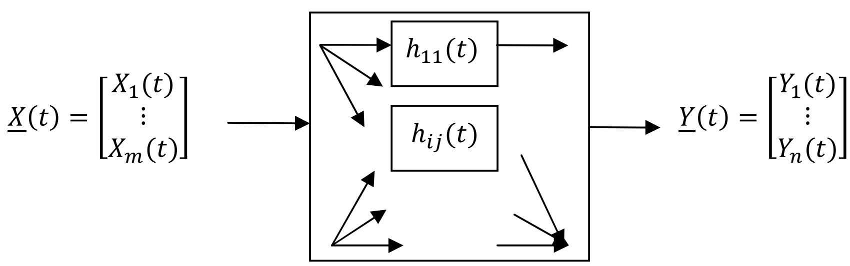
נוסחא זו נכונה גם כאשר ישנם תדרים בהם $H_1(\omega)=0$.

הגדרה: מסננות אורתוגונאליות הן מסננות לא חופפות בתדר, כלומר $H_1(\omega)\cdot H_2(\omega)\equiv 0$ לכל ω (או $h_1(t)*h_2(t)\equiv 0$).



נשים לב שמהנוסחא לעיל נובע שבמערכת *SIMO* עם מסננות אורתוגונאליות הקרוס קורלציה בין המוצאים היא 0.

מעבר במערכת MIMO (Mingle Input Multiple Output)



$$\underset{n\times 1}{\underline{Y}(t)} = \underset{n\times m}{\underline{\underline{H}}(t)} * \underset{m\times 1}{\underline{X}(t)}$$

$$Y_i(t)=\sum_{j=1}^{m} h_{ij}(t)*X_j(t)$$

$$\underline{\underline{R}}_X(\tau)\triangleq\left\{R_{X_iX_j}(\tau)\right\}_{i=1,j=1}^{i=m,j=m}$$

$$\underline{\underline{\mathbb{H}}}(\omega)=\mathcal{F}\left\{\underline{\underline{H}}(t)\right\}$$

<u>טענה</u> : אם וקטור הכניסה $\underline{X}(t)$ הוא *J.W.S.S* ואם המערכת $\underline{\underline{H}}(t)$ היא קבועה בזמן אזי וקטור הכניסה ווקטור היציאה הם *J.W.S.S* ומתקיים :

ציר התדר	ציר הזמן
$\underset{m\times n}{\underline{\underline{S}}_{XY}(\omega)} = \underset{m\times m}{\underline{\underline{S}}_X(\omega)}\cdot\underset{m\times n}{\underline{\underline{\mathbb{H}}}^{*T}(\omega)}$	$\underset{m\times n}{\underline{\underline{R}}_{XY}(\tau)} = \underset{m\times m}{\underline{\underline{R}}_X(\tau)} * \underset{m\times n}{\underline{\underline{H}}^T(-\tau)}$
$\underset{n\times n}{\underline{\underline{S}}_Y(\omega)} = \underset{n\times m}{\underline{\underline{\mathbb{H}}}(\omega)}\cdot\underset{m\times n}{\underline{\underline{S}}_{XY}(\omega)}$	$\underset{n\times n}{\underline{\underline{R}}_Y(\tau)} = \underset{n\times m}{\underline{\underline{H}}(\tau)} * \underset{m\times n}{\underline{\underline{R}}_{XY}(\tau)}$
$= \underset{n\times m}{\underline{\underline{\mathbb{H}}}(\omega)}\cdot\underset{m\times m}{\underline{\underline{S}}_X(\omega)}\cdot\underset{m\times n}{\underline{\underline{\mathbb{H}}}^{*T}(\omega)}$	$= \underset{n\times m}{\underline{\underline{H}}(\tau)} * \underset{m\times m}{\underline{\underline{R}}_X(\tau)} * \underset{m\times n}{\underline{\underline{H}}^T(-\tau)}$

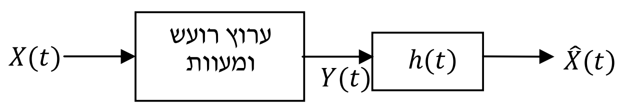
נתייחס ל- *SIMO* כאל מקרה פרטי עבור $m=1, n=2$:

$$\underline{\underline{S}}_Y(\omega) = \begin{bmatrix} S_{Y_1Y_1}(\omega) & S_{Y_1Y_2}(\omega) \\ S_{Y_2Y_1}(\omega) & S_{Y_2Y_2}(\omega) \end{bmatrix} = \begin{bmatrix} H_1(\omega) \\ H_2(\omega) \end{bmatrix} \cdot S_X(\omega) \cdot [H_1^*(\omega),\ H_2^*(\omega)] =$$
$$= S_X(\omega) \cdot \begin{bmatrix} |H_1(\omega)|^2 & H_1(\omega) \cdot H_2^*(\omega) \\ H_2(\omega) \cdot H_1^*(\omega) & |H_2(\omega)|^2 \end{bmatrix}$$

ואכן ניתן לראות שהאיבר המתאים ל- $S_{Y_1Y_2}(\omega)$ מתלכד עם הנוסחא ל- $SIMO$ שפיתחנו קודם.

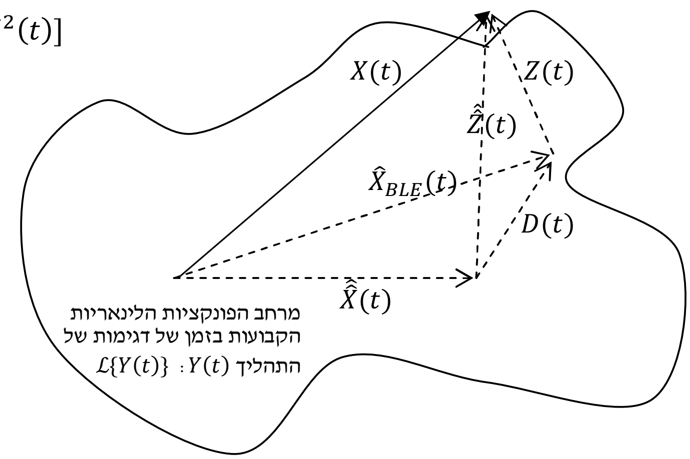
<u>מסננת וינר על-פי משפט פיתגורס</u>

נחזור כעת לתת פרשנות (והוכחה חלופית) לעקרון הניצבות עבור תהליכים רציפים



אות השגיאה: $Z(t) \triangleq \hat{X}(t) - X(t)$

$MSE = E[Z^2(t)]$

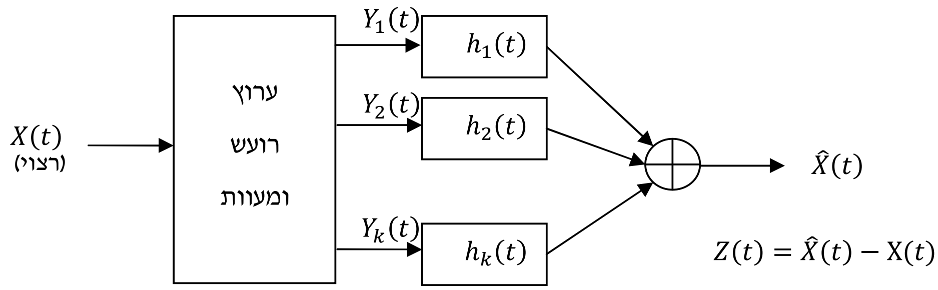


נראה כי המשערך שלנו הוא לינארי אופטימלי אם מתקיימת תכונת האורתוגונאליות, כלומר שגיאת השערוך $Z(t)$ ניצבת לכל תהליך במרחב המשערכים הלינארים הקבועים בזמן של $X(t)$ מתוך $Y(t)$:

$$MSE_{\text{מתחרה}} = E\left[\left(\hat{\hat{X}}(t) - X(t)\right)^2\right] =$$
$$= E\left[\left(\left(\hat{\hat{X}}(t) - \hat{X}_{BLE}(t)\right) + \left(\hat{X}_{BLE}(t) - X(t)\right)\right)^2\right] =$$
$$= E\left[\left(D(t) + Z(t)\right)^2\right] = \underbrace{E[D^2(t)]}_{\geq 0} + \underbrace{2E[D(t)\cdot Z(t)]}_{\text{לפי אורתוגונליות } =0} + \underbrace{E[Z^2(t)]}_{MSE_{BLE}} \geq MSE_{BLE}$$

הסבר: הסיבה להתאפסות האיבר הצולב היא ש $D(t)$ הוא ההפרש בין שני משערכים לינארים קבועים בזמן מתוך אות המדידה $Y(t)$, ולכן הוא בעצמו פונקצית לינארית קבועה בזמן של אות המדידה $Y(t)$, ומכאן שהוא ניצב לאות השגיאה $Z(t)$.

מסננת וינר עבור מספר תהליכי מדידה



($[Y_1(t), Y_2(t), \dots, Y_k(t)]$ - וקטור של תהליכי מדידה)

נרצה למצוא וקטור של מסננות $h_1(t), \dots, h_k(t)$ שיביאו למינימום את $E[Z^2(t)]$.

נניח שהתהליך הרצוי ותהליכי המדידה הם $J.W.S.S$, ושנתונות לנו פונקציות הקרוס-קורלציה בין $X(t)$ ו- $Y_i(t)$ ובין $Y_i(t)$ לבין עצמם.

<u>עקרון הניצבות לוקטור מסננות אופטימלי</u>

$$\forall \tau: \quad Z(t) \perp Y_i(t-\tau) \qquad\qquad i = 1, \dots, k$$

$$\forall i: \quad 0 = E[Z(t) \cdot Y_i(t-\tau)] = E\left[\left(\sum_{j=1}^{k} h_j(t) * Y_j(t) - X(t)\right) \cdot Y_i(t-\tau)\right] =$$

$$\underbrace{=}_{\substack{\text{החלפת סדר} \\ \text{סכום ותוחלת}}} \sum_{j=1}^{k} h_j(\tau) * R_{Y_j Y_i}(\tau) - R_{XY_i}(\tau)$$

בציר התדר:

$$\sum_{j=1}^{k} H_j(\omega) S_{Y_j Y_i}(\omega) - S_{XY_i}(\omega) = 0 \qquad\qquad i = 1, \dots, k$$

ובכתיבה מטריצית אפשר לחלץ

$$\boxed{\underset{k\times 1}{\underline{H}_{BLE}(\omega)} = \underset{k\times k}{\underline{\underline{S}}_Y^{-T}(\omega)} \cdot \underset{k\times 1}{\underline{S}_{XY}(\omega)}}$$

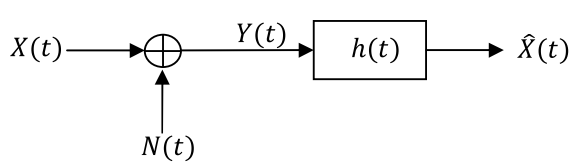
עבור $k = 1$ נקבל:

$$H(\omega) = S_Y^{-1}(\omega) \cdot S_{XY}(\omega) = \frac{S_{XY}(\omega)}{S_Y(\omega)}$$

זהו מסנן וינר רגיל.

דוגמא 2: <u>שערוך אות לאחר מעבר דרך ערוץ עם רעש אדטיבי</u> (דוגמא לסינון וינר מתהליך מדידה <u>יחיד</u>)

$N(t)$ הוא רעש אדטיבי: $\forall t_1, t_2: X(t_1) \perp N(t_2)$



$$h_{BLE}(t) = \mathcal{F}^{-1}\{H_{BLE}(\omega)\}$$

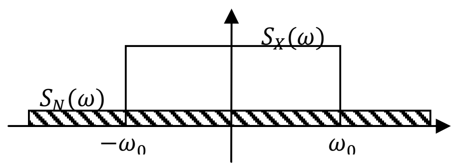
$$H_{BLE}(\omega) = \frac{S_{XY}(\omega)}{S_Y(\omega)} \underbrace{=}_{(1),(2)} \frac{S_X(\omega)}{S_X(\omega) + S_N(\omega)} \rightarrow \begin{cases} 1 & S_N(\omega) \ll S_X(\omega) \\ \frac{S_X(\omega)}{S_N(\omega)} \rightarrow 0 & S_X(\omega) \ll S_N(\omega) \end{cases}$$

$$(1)\quad R_Y(\tau) = R_{X+N}(\tau) \underbrace{=}_{\text{חס"ק}} R_X(\tau) + R_N(\tau) \Rightarrow S_Y(\omega) = S_X(\omega) + S_N(\omega)$$

$$(2)\quad R_{XY}(\tau) = R_{X,X+N}(\tau) \underbrace{=}_{\text{חס"ק}} R_X(\tau) \Rightarrow S_{XY}(\omega) = S_X(\omega)$$

כלומר בכל ω נוסחת המשערך הלינארי הוא כמו שערוך סקאלרי בערוץ רעש אדיטיבי של משתנים אקראיים.

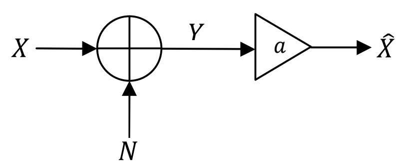
נניח $N(t)$ הוא רעש לבן, כלומר במישור הזמן הוא בעל פונקצית אוטו קורלציה $\delta(t)$ (מוכפלת בקבוע, אולי) ובמישור התדר ספקטרום ההספק הוא גודל קבוע, ויהי $X(t)$ אות $W.S.S$ מוגבל סרט בעל ספקטרום שטוח, דהיינו "חלון" במישור התדר.



המסננת הדרושה לסינון אופטימלי (לינארי) תעביר תדרים בתחום $[-\omega_0, \omega_0]$ בלבד.
מחוץ לתחום זה המסננת תאפס את אות המוצא, ובתוך תחום זה יהיה כפל בקבוע שייקבע על פי היחס $\frac{S_X(\omega)}{S_X(\omega)+S_N(\omega)}$.

נוסחת השגיאה של מסננת וינר

תזכורת: נוסחת השגיאה במקרה הסקלארי:

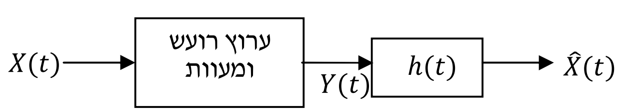


$$a_{BLE} = \frac{R_{XY}}{R_Y} = \frac{R_X}{R_X+R_N}$$

$$E[Z^2] = E[Z \cdot (aY - X)] \underbrace{=}_{\text{ניצבות}} - E[Z_{BLE} \cdot X] = -E[(a_{BLE}Y - X) \cdot X] =$$

$$= E[X^2] - a_{BLE}E[Y \cdot X] = R_X - a_{BLE}R_{YX} = R_X - \frac{R_{XY}}{R_Y}R_{YX} = R_X - \frac{R_{XY}^2}{R_Y}$$

עבור תהליכים אקראיים:



$$MSE = E[Z^2(t)] = E\left[Z(t) \cdot \left(h(t) * Y(t) - X(t)\right)\right] =$$

$$\underbrace{=}_{\forall t_1,t_2: Z(t_1)\perp Y(t_2)} - E[Z(t)X(t)] =$$

$$= -E\left[\left(h_{BLE}(t) * Y(t) - X(t)\right) \cdot X(t)\right] =$$

$$= R_X(\tau = 0) - \underbrace{\int_{-\infty}^{\infty} h_{BLE}(\alpha) \cdot R_{YX}(-\alpha) d\alpha}_{h_{BLE}(\tau) * R_{YX}(\tau)|_{\tau=0}} =$$

$$\underbrace{=}_{(1)} \frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}\left(S_X(\omega)-S_{YX}(\omega)\cdot H_{BLE}(\omega)\right)d\omega =$$
$$\underbrace{=}_{(2)} \frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}\left(S_X(\omega)-S_{YX}(\omega)\cdot\frac{S_{XY}(\omega)}{S_Y(\omega)}\right)d\omega =$$
$$\underbrace{=}_{(3)} \frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}\left(S_X(\omega)-\frac{|S_{XY}(\omega)|^2}{S_Y(\omega)}\right)d\omega$$

(1) $R_X(\tau=0)=\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}S_X(\omega)d\omega$

(2) $H_{BLE}(\omega)=\frac{S_{XY}(\omega)}{S_Y(\omega)}$

(3) $S_{YX}(\omega)=S_{XY}^*(\omega)$

נבחין כי הארגומנט שבתוך האינטגרל בביטוי שהתקבל הוא ה- MSE בכל תדר. נרצה לוודא שהארגומנט הוא אי שלילי.

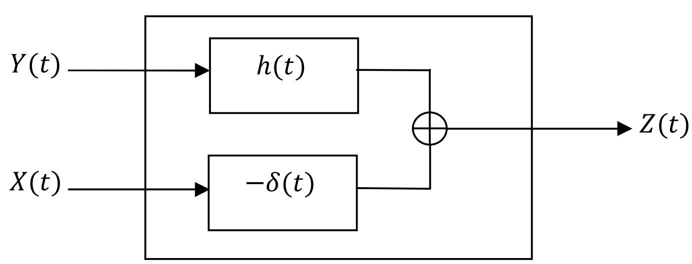
<u>אי שוויון קושי שוורץ לפונקצית אוטו קורלציה</u>: $R_X(0)\cdot R_Y(0)\geq[R_{XY}(\tau)]^2$.

נשים לב שהוכחנו בעמ' XX, על סמך אי שוויון בדיוק, אי שוויון דומה גם עבור ספקטרומי הספק:
$S_X(\omega)\cdot S_Y(\omega)\geq|S_{XY}(\omega)|^2$

ניתן להוכיח כי הארגומנט בתוך האינטגרל מקודם הוא ספקטרום השגיאה:

$$S_Z(\omega)\triangleq\mathcal{F}\{R_Z(t)\}=S_X(\omega)-\frac{|S_{XY}(\omega)|^2}{S_Y(\omega)} \qquad (\text{עבור } h_{BLE}(t))$$

<u>רעיון ההוכחה</u>: ע"י מערכת $MISO$ שמייצגת את הקשר בין האות הרצוי ואות המדידה לאות השגיאה:



נשתמש בנוסחת $MIMO$ עבור $n=1, m=2$, ונקבל $R_Z(t)$ עבור $h(t)$ כללית. כעת ניתן להציב את פתרון וינר $H_{BLE}(\omega)=\frac{S_{XY}(\omega)}{S_Y(\omega)}$ או לבצע מינימיזציה מחדש:

$$MSE|_{h(t)\text{ כללית}}=\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}S_Z(\omega,H(\cdot))d\omega$$

על מנת לקבל את המסננת האופטימלית:

$$H_{BLE}(\omega)=argmin_{H(\omega)}\int_{-\infty}^{\infty}S_Z(\omega,H(\cdot))d\omega\underbrace{=}_{(1)}argmin_{H(\omega)}S_Z(\omega,H(\cdot))$$

(1) ההחלפה מותרת כי $S_Z(\omega,H(\cdot))$ תלוי רק ב- $H(\omega)$ בתדר ω ולא בשאר התדרים ולכן אפשר לבצע מינימיזציה תדר-תדר.

לאחר מינימיזציה מתקבל שאכן עבור מסננת וינר מתקיים שספקטרום אות השגיאה הוא
$S_Z(\omega)=S_X(\omega)-\frac{|S_{XY}(\omega)|^2}{S_Y(\omega)}$

נשים לב שתוצאה זו מתלכדת עם נוסחת השיאה הריבועית בשערוך של משתנים אקראיים $(R_Z = R_X - \frac{{R_{XY}}^2}{R_Y})$ כפי שפותח בעמוד הקודם, רק <u>לכל תדר בנפרד</u>.
תוצאה זו מתאימה להסבר שניתן בשיעור הקודם, שניתן לחשוב על סינון וינר כאילו שיש לנו אוסף של תהליכים צרי סרט בתדרים $\omega_1, \omega_2, \dots, \omega_N$ כך שבתדר ω_i יש לאות הרצוי מומנט שני $\Delta_i \cdot S_X(\omega_i)$, לאות המדידה מומנט שני $\Delta_i \cdot S_Y(\omega_i)$ והקרוס-קורלציה בניהם $\Delta_i \cdot S_{XY}(\omega_i)$.

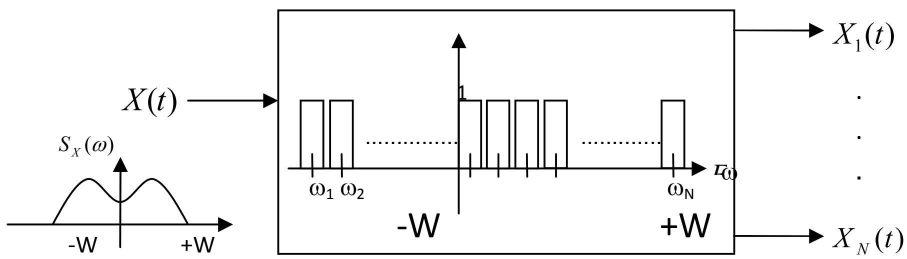
הרחבה של עקרון זה ניתנת במסגרת הדיון על עיבוד מקבילי של פסי תדר.

<u>עיבוד מקבילי של פסי תדר</u>

<u>הגדרה</u>: נניח תהליך $X(t)$, תהליך WSS מוגבל סרט (-W,+W).

$X(t)$ נכנס למסנן המורכב מ- N מסננות, לא חופפות, בגובה 1 ומכסות את כל ציר התדר (-W,+W).

במוצא מתקבלים התהליכים $X_1(t), X_2(t), \dots, X_N(t)$.



Filter Bank

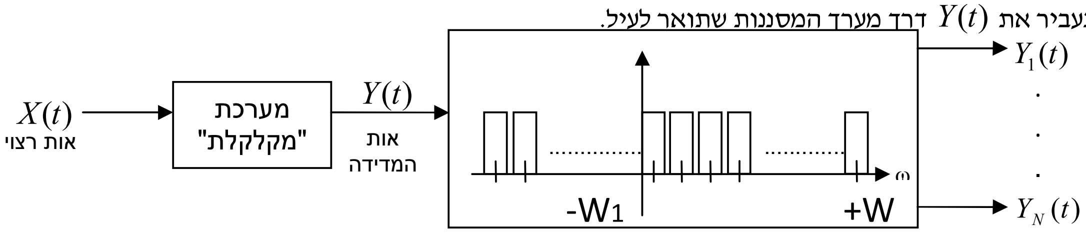
קל לראות כי זוהי בעיית SIMO בה המסננות אורתוגונליות (כלומר הן לא חופפות בתדר ומקיימות $\forall\omega\ ,\ H_1(\omega)\cdot H_2^*(\omega)\equiv 0$). נובע מכך שרכיבי התהליך הם אורתוגונליים: $R_{X_iX_j}(\tau)\equiv 0\ ,\ \forall i\neq j$

<u>מסקנה</u>: תהליך בתדר מסוים חס"ק בתהליך בתדר אחר.

<u>סינון וינר ע"י עיבוד מקבילי של פסי תדר</u>

<u>הגדרה</u>: נניח שהתהליך $X(t)$ מוגבל סרט (-W,+W) ותהליך המדידה $Y(t)$ סטאציונארי איתו במשותף (כלומר $X(t)$ ו-$Y(t)$ הם JWSS).

נעביר את $Y(t)$ דרך מערך המסננות שתואר לעיל.



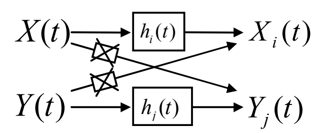
Filter Bank

נדמיין שהיינו מחלקים גם את $X(t)$ ל- N רכיבים ע"י מעבר במסננות, כפי שתואר בתחילת העמוד.

מהתכונה של מערכת MIMO אלכסונית עם מסננות אורתוגונליות נובע: $R_{X_iY_j}(\tau)\equiv 0\ ,\ \forall i\neq j$

<u>הבהרה:</u> מערכת אלכסונית היא מערכת שאין בה מסננות "צולבות" שמחברות בין כניסה i ליציאה j.

במקרה שלנו



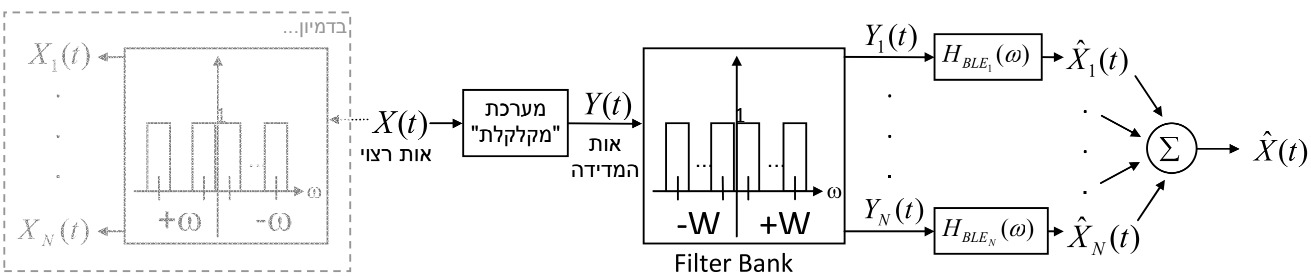
כלומר $\underline{\underline{H}}(t)=\begin{pmatrix} h_i(t) & 0 \\ 0 & h_j(t)\end{pmatrix}$.

מנוסחת ספקטרום המוצא של מערכת MIMO נובע: $S_{X_iY_j}(\omega)=S_{XY}(\omega)\cdot H_i(\omega)\cdot H_j^*(\omega)$

<u>מסקנה:</u> לכל אחד מרכיבי אות המדידה יש מידע (מבחינת קרוס קורלציה) אודות רכיב האות הרצוי שמתאים לו באינדקס.

<u>מסקנה:</u> לצורך סינון לינארי אופטימלי, ב- $Y_j(t)$ אין אינפורמציה מעניינת לגבי $X_i(t)$, $i\neq j$.

המערכת הכוללת תיראה כך:

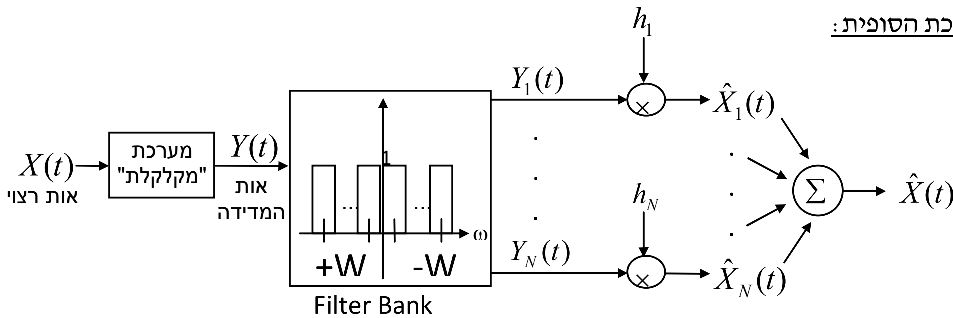


$$H_{BLE_i}(\omega)=\begin{cases}\dfrac{S_{X_iY_i}(\omega)}{S_{Y_i}(\omega)} & ,\ \text{בפס התדר ה}-i\ \ \omega \\ don't\ care & ,\ \text{בתדרים אחרים}\end{cases}\approx\begin{cases}\dfrac{S_{XY}(\omega_i)}{S_Y(\omega_i)}=h_i\ \ (\text{קבוע}) & ,\ \text{בפס התדר ה}-i\ \ \omega \\ don't\ care & ,\ \text{בתדרים אחרים}\end{cases}\approx h_i\ \ \forall\omega$$

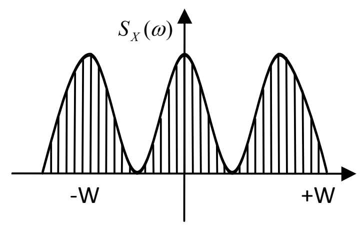
<u>הבהרה ל- don't care:</u> נשים-לב שאם מציבים ω מחוץ לפס התדרים ב-i, בנוסחת וינר מקבלים $\frac{S_{X_iY_i}(\omega)}{S_{Y_i}(\omega)}=\frac{0}{0}$. אבל בפועל, היות ואין לאות $Y_i(t)$ הספק בתדרים מחוץ לפס, אזי המוצא לא יושפע מצורת המסננת מחוץ לפס.

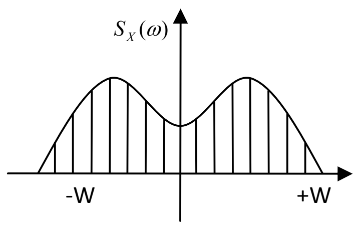
<u>מסקנה:</u> נבחר: $H_{BLE_i}(\omega)=h_i$ לכל ω.

<u>המערכת הסופית</u>:



<u>הערה חשובה</u>: את רוחב פסי התדר נבחר ביחס לשינויים בספקטרום האות. ככל שהתנודתיות בספקטרום האות תגדל, כך רוחב פסי התדר יקטן.





תהליך פואסון (*Poisson Process*)

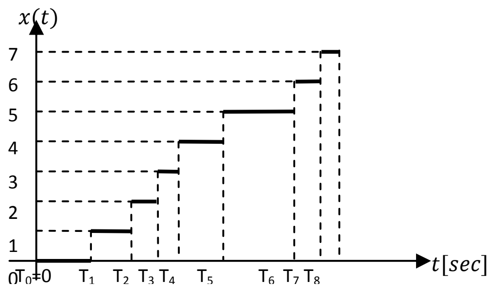
<u>תהליך מנייה (*Counting Process*) בזמן רציף</u>

דוגמא 1: סופרים את כמות השיחות הנכנסות לטלפון סלולרי עד לרגע נתון.

דוגמא 2: מספר ה- Packetים שהתקבלו במחשב מסוים עד לרגע נתון.

שני התהליכים הנ"ל מתאפיינים בכך שהם יכולים לקבל אך ורק ערכים בדידים, ושפונקציות המדגם האופייניות שלהם הן מונוטוניות עולות במובן החלש.

פונקצית מדגם אופיינית של כ"א מהתהליכים הנ"ל יכולה להיראות כך:



הגדרת תהליך פואסון

הגדרה: תהליך מנייה $\{N(t),\ t \geq 0\}$ הוא <u>תהליך אקראי פואסון</u> עם קצב הגעות ממוצע $\lambda \left[\frac{\text{הגעות}}{\text{שנייה}}\right]$ אם:

א. הדגימה $N(t_0)$ ברגע $t = t_0$ היא משתנה אקראי בעל פילוג פואסוני עם תוחלת $\lambda \cdot t_0$ (לכל t_0).

ב. $N(t)$ הוא תהליך עם "תוספות בת"ס":

$$N(t_1, t_2) \triangleq N(t_2) - N(t_1)\ ; \qquad t_1 < t_2$$
$$N(t_1, t_2) \perp\!\!\!\perp \{\text{כל מה שקרה עד רגע } t_1\}$$

בפרט מתקיים:

$$N(t_1, t_2) \perp\!\!\!\perp N(t_1)$$

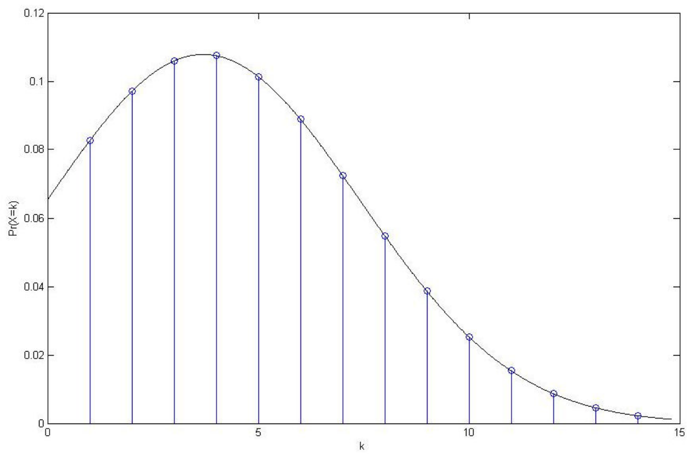
תזכורת: <u>פילוג פואסון</u> $X \sim Poisson(\Lambda)$

$$\Pr(X = k) = e^{-\Lambda} \cdot \frac{\Lambda^k}{k!} \qquad k = 0,1,2,\dots$$

<u>סטטיסטיקה מסדר שני של פילוג פואסון</u>:

$$E[X] = \sum_{k=0}^{\infty} k \cdot \Pr(X = k) = \Lambda$$
$$Var[X] = \sum_{k=0}^{\infty} (k - \Lambda)^2 \cdot \Pr(X = k) = \Lambda$$

כאשר $\Lambda \gg 1$ ההסתברויות $\Pr(X = k)$ נמצאות על גרף שדומה לגרף פונקצית צפיפות של התפלגות נורמלית, בעלת תוחלת Λ ושונות Λ (ללא הזנב משמאל לראשית, כמובן):



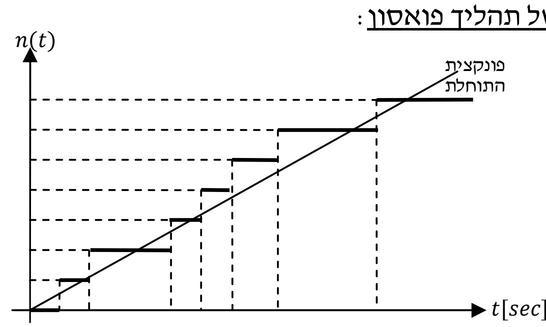
<u>תכונת האדיטיביות של פילוג פואסון</u>: אם X_1, X_2 הם משתנים אקראיים פואסוניים בת״ס בעלי תוחלות Λ_1, Λ_2 בהתאמה אזי $X_1 + X_2 \sim Poisson(\Lambda_1 + \Lambda_2)$.

תכונת האדיטיביות מראה שניתן לקיים את סעיפים א ו ב׳ בהגדרת תהליך פואסון, וכן:

$$N(t_1, t_2) \sim Poisson\left(\lambda \cdot (\mathrm{t}_2 - \mathrm{t}_1)\right)$$

(ראו הוכחה באמצעות פונקציה אופיינית בדפי התרגול).

<u>פונקצית מדגם אופיינית של תהליך פואסון</u>:



פונקצית מדגם אופיינית של תהליך אקראי פואסון היא פונקצית מדגם של תהליך מנייה, שבה המדרגות של הפונקציה נעות סביב פונקצית התוחלת שעולה לינארית בזמן.

<u>**סטטיסטיקה מסדר שני של תהליך אקראי פואסון**</u>

א. <u>תוחלת של תהליך אקראי פואסון</u>:

$$\eta(t) \triangleq E[N(t)] = E[Poisson(\lambda \cdot t)] = \lambda \cdot t$$

ב. <u>פונקצית אוטו קורלציה</u>: (נניח $t_1 < t_2$)

$$R_N(t_1, t_2) = E[N(t_1) \cdot N(t_2)] \underbrace{=}_{t_1<t_2 \text{ :הנחה}} E\big[N(t_1) \cdot \big(N(t_1) + N(t_1, t_2)\big)\big] =$$

$$= E[N(t_1) \cdot N(t_1)] + E\underbrace{[N(t_1) \cdot N(t_1, t_2)]}_{\text{בת"ס, לפי סעיף ב בהגדרה}} =$$

$$\underbrace{=}_{(*)} Var[N(t_1)] + E^2[N(t_1)] + E[N(t_1)] \cdot E[N(t_1, t_2)] =$$

$$= \lambda \cdot t_1 + (\lambda \cdot t_1)^2 + (\lambda \cdot t_1) \cdot \big(\lambda \cdot (t_2 - t_1)\big) = \lambda \cdot t_1 (1 + \lambda \cdot t_2)$$

(*) השתמשנו בכך ש: מומנט שני = שונות + תוחלת בריבוע, ובכך שתוחלת מכפלת תהליכים בת״ס שווה למכפלת התוחלות.

ג. פונקצית אוטו קווריאנס: (נניח $t_1 < t_2$)

$$C_N(t_1,t_2) = R_N(t_1,t_2) - \eta(t_1)\cdot\eta(t_2) = \lambda\cdot t_1 + \lambda^2\cdot t_1\cdot t_2 - \lambda^2\cdot t_1\cdot t_2 = \lambda\cdot t_1$$

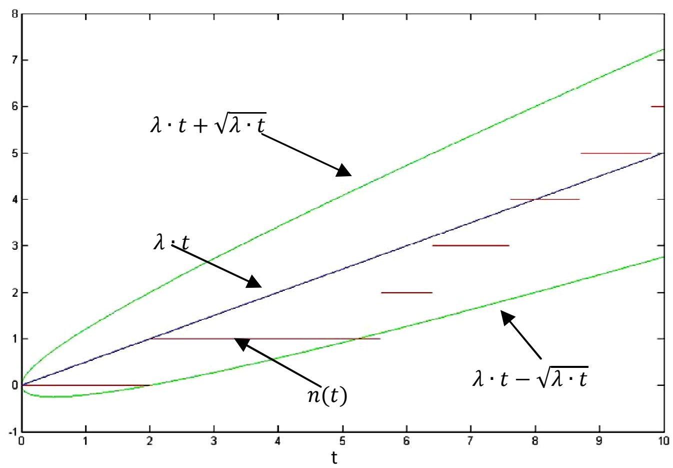
ניזכר שהנחנו כי $t_1 < t_2$. הנוסחא הכללית היא:

$$C_N(t_1,t_2) = \lambda\cdot\min\{t_1,t_2\}$$

ומכאן ש:

$$Var[N(t)] = C_N(t,t) = \lambda\cdot t$$

לכן בפונקצית מדגם אופיינית של תהליך אקראי פואסון, מרבית הפיזור יהיה סביב פונקצית התוחלת $\lambda\cdot t$, במרחק של $\sqrt{\lambda\cdot t}$ לכל כיוון:



תהליך זמני ההגעות

הגדרה: זמן הגעה של תהליך אקראי פואסון $N(t)$ הוא זמן שבו ערכה של פונקצית המדגם גדל בלפחות 1.

הערה: כפי שנראה בהמשך, ההסתברות לעלייה ביותר מ-1 באותה נקודת זמן היא 0.

נסמן T_i - זמן ההגעה ה- i בתהליך אקראי פואסון $N(t)$.
אזי התהליך האקראי $\{T_i\}_{i=1}^n$ מקיים קשר חד-חד ערכי עם התהליך $N(t)$. כלומר, ניתן לחשב הסתברויות של התהליך $\{T_i\}_{i=1}^n$ מתוך הסתברויות של התהליך $N(t)$, ולהיפך.

מתקיים: $\{N(t) < n\} \Leftrightarrow \{T_n > t\}$.

מהו הפילוג של המ״א T_1 (זמן ההגעה הראשון)?
נשים לב שההתפלגות של T_1 רציפה, לכן נחשב במקרה זה את ה- PDF:

$$Pr(t < T_1 \le t+\Delta) = Pr\big(\text{זמן ההגעה הראשון היה באינטרוול}(t,t+\Delta)\big) =$$
$$= Pr(\text{עד רגע } t \text{ לא היו הגעות וגם עד הרגע } t+\Delta \text{ הייתה לפחות הגעה אחת}) =$$
$$= Pr(N(t)=0, N(t+\Delta)\ge 1) = Pr\underbrace{(N(t)=0, N(t,t+\Delta)\ge 1)}_{\text{בת"ס}} =$$
$$= Pr(N(t)=0)\cdot Pr(N(t,t+\Delta)\ge 1) =$$
$$= Pr(Poisson(\lambda\cdot t)=0)\cdot Pr(Poisson(\lambda\cdot\Delta)\ge 1) = e^{-\lambda\cdot t}\cdot(1-e^{-\lambda\cdot\Delta})$$

$$f_{T_1}(t) = \lim_{\Delta\to 0} \frac{Pr(t<T_1\le t+\Delta)}{\Delta} = \lim_{\Delta\to 0} \frac{e^{-\lambda\cdot t}\cdot(1-e^{-\lambda\cdot\Delta})}{\Delta} \underbrace{=}_{\text{טור טיילור},(*)}$$

$$= \lim_{\Delta\to 0} \frac{e^{-\lambda\cdot t}\cdot\left(1-\left(1-\lambda\cdot\Delta+\frac{(\lambda\cdot\Delta)^2}{2!}-\cdots\right)\right)}{\Delta} = \lim_{\Delta\to 0}(\lambda e^{-\lambda\cdot t} + O(\Delta)) = \lambda e^{-\lambda\cdot t}$$

(*) אפשר להוכיח ע״י שימוש במשפט לופיטל

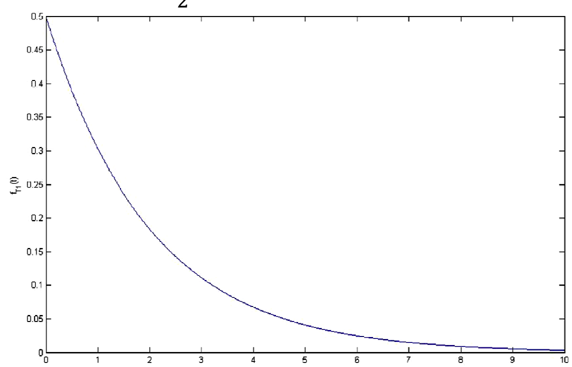
במקרה דנן $g(x) = O\big(f(x)\big)$ אם ורק אם $f(x)$ חוסמת מלעיל את $g(x)$ כאשר $x \to 0$.
פורמלית:
$$\exists M > 0, \exists\delta > 0 : \forall |x| \le \delta,\ \ |g(x)| \le M \cdot |f(x)|$$

נוכל להשתמש גם בהגדרה אחרת עבור הסימון הנ״ל: נאמר ש $g(x) = O(\Delta^n)$ אם $g(x)$ דועכת ל- 0 עבור $x \to 0$ בקצב זהה או מהיר יותר מאשר Δ^n (ובכל אופן מהר יותר מאשר Δ^{n-1}).

התקבל כי $T_1 \sim \exp(\lambda)$.
(כלומר T_1 מתפלג אקספוננציאלית עם תוחלת $1/\lambda$: $f_{T_1}(t) = \lambda \cdot e^{-\lambda t} \cdot u(t)$).

תזכורת: יהי X משתנה אקראי המתפלג אקספוננציאלית עם פרמטר λ. אזי ה- CDF היא:
$$F_X(x) = \Pr(X \le x) = \int_0^x f_X(x)dx = 1 - e^{-\lambda x}$$

פונקצית צפיפות של משתנה אקספוננציאלי בעל תוחלת 2 ($\lambda = \frac{1}{2}$):



נגדיר את התהליך האקראי $X_n = T_n - T_{n-1}$: הפרש בין זמן ההגעה ה- n וזמן ההגעה ה- n-1.
כפי שנראה להלן, תהליך זה הוא תהליך אקראי $i.i.d$ המתפלג $X_n \sim \exp(\lambda)$.

<u>הגדרות חלופיות של תהליך פואסון</u>

כעת נגדיר תהליך אקראי פואסון בשלוש דרכים נוספות:

הגדרה II: <u>הגדרת תהליך אקראי פואסון באמצעות סדרת זמני ההגעות</u>

תהליך מניה שנוצר ע״י סדרת זמני הגעה $T_1, T_2, T_3 \ldots$ עם הפרשים $X_n \triangleq T_n - T_{n-1}$, כך ש- X_n היא סדרת $i.i.d$ עם פילוג אקספוננציאלי $X_n \sim \exp(\lambda)$ מתאר תהליך אקראי פואסון בעל קצב הגעות λ.

<u>הוכחה:</u> בתרגיל הבית.

תכונת חוסר הזכרון של הפילוג האקספוננציאלי

תזכורת: יהי $X\sim\exp(\lambda)$. אזי X הוא "חסר זיכרון": $X-a|_{X>a}\sim exp(\lambda)$, כלומר:
$$\Pr(X-a>b|X>a)=\Pr(X>b)$$

<u>הוכחה</u>:
$$\Pr(X-a>b|X>a)=\frac{\Pr(X>a \text{ and } X-a>b)}{\Pr(X>a)}=\frac{\Pr(X>a+b)}{\Pr(X>a)}=\frac{1-F_X(a+b)}{1-F_X(a)}=$$
$$=\frac{e^{-\lambda(a+b)}}{e^{-\lambda a}}=e^{-\lambda b}=\Pr(X>b)$$

<u>פילוג ההגעות באינטרוול קטן Δ</u>:

$$Pr(N(t,t+\Delta)=0)=Pr(N(\Delta)=0)=e^{-\lambda\cdot\Delta}=1-\lambda\cdot\Delta+O(\Delta^2)$$
$$Pr(N(t,t+\Delta)=1)=Pr(N(\Delta)=1)=\lambda\cdot\Delta\cdot e^{-\lambda\cdot\Delta}=\lambda\cdot\Delta+O(\Delta^2)$$
$$Pr(N(t,t+\Delta)\geq 2)=Pr(N(\Delta)\geq 2)=O(\Delta^2)$$

כלומר, בקירוב מסדר ראשון, באינטרוול קטן בהסתברות קרובה ל-1 לא מגיע כלום, ובהסתברות קטנה יש הגעה אחת (ההסתברות ליותר מהגעה אחת - זניחה).

הגדרה III: <u>בנייה של תהליך אקראי פואסון ע"י קירוב של דגימה מהירה בזמן בדיד</u>

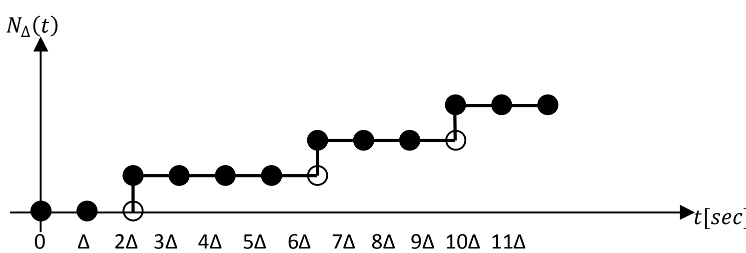
תהא B_n סדרת ברנולי-p: $B_n=\begin{cases}1 & w.p.\ p\\ 0 & w.p\ 1-p\end{cases}$.

נגדיר את התהליך האקראי בזמן רציף: $N_\Delta(t)=\sum_{n=1}^{\lfloor t/\Delta\rfloor}B_n$, כאשר $p=\lambda\cdot\Delta$.

(הערה: $1/\lambda$ = זמן ההגעות הממוצע $\gg\Delta$)

בגבול $\Delta\to 0$, $N_\Delta(t)$ הוא תהליך אקראי פואסון.

נשים לב שכאשר $\Delta\to 0$ קצב הדגימה עולה מחד, והסיכוי להצלחה יורד מאידך.



<u>מאפיינים סטטיסטיים של התהליך $N_\Delta(t)$</u>: (נניח t כפולה שלמה של Δ)
$$E[N_\Delta(t)]=p\cdot{}^{t}/_{\Delta}=\lambda\cdot\Delta\cdot{}^{t}/_{\Delta}=\lambda\cdot t$$
$$Var[N_\Delta(t)]=Var\left(\sum_{n=1}^{\lfloor t/\Delta\rfloor}B_n\right)=\left({}^{t}/_{\Delta}\right)\cdot Var(B_n)=\left({}^{t}/_{\Delta}\right)\cdot(p-p^2)=$$
$$=\lambda\cdot t+O(\Delta)\cong\lambda\cdot t$$

$$Pr(N_\Delta(t)=m)\underbrace{=}_{n=\lfloor t/\Delta\rfloor}Pr\left(\sum_{k=1}^{n}B_k=m\right)=\binom{n}{m}\cdot p^m\cdot(1-p)^{n-m}=\cdots\underbrace{=}_{\Delta\to 0}$$
$$=e^{-\lambda\cdot t}\cdot\frac{(\lambda\cdot t)^m}{m!}\equiv Pr(Poisson(\lambda\cdot t)=m)$$

נוכיח כי הגדרה זו קונסיסטנטית עם הגדרה II. כלומר ההפרשים בין זמני ההגעה לפי הגדרה III הם משתנים אקספוננציאליים. נשים לב שזמני ההגעה לפי הגדרה III הם ערכים בדידים, לכן נצפה לקבל התפלגות שהיא אקספוננציאלית דגומה, ולא אקספוננציאלית רציפה

$$Pr(T_{i+1}-T_i=(k+1)\cdot\Delta)=Pr(\text{היו } k \text{ כשלונות ברצף ואז הצלחה אחת})=(1-p)^k\cdot p$$
$$=(1-\lambda\Delta)^k\cdot\lambda\Delta$$

נשתמש בקירוב הבא: עבור ε קטן מאוד, מתקיים $1-\varepsilon\cong e^{-\varepsilon}$. נקבל:

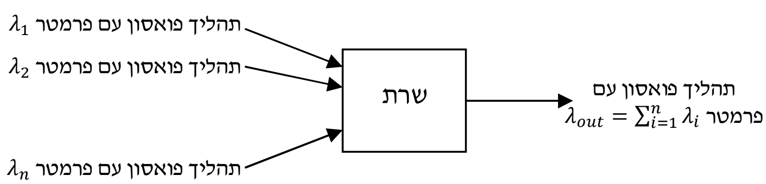
$$Pr(T_{i+1}-T_i=(k+1)\cdot\Delta)\cong\left(e^{-\lambda\Delta}\right)^k\lambda\Delta=\text{אקספוננציאלי דגום}$$

מעבר לזמן רציף:

$$Pr(T_{i+1}-T_i=x)=e^{-\lambda\Delta(k+1)}\cdot e^{-\lambda\Delta}\cdot\lambda\Delta\propto e^{-\lambda\Delta(k+1)}=e^{-\lambda x}$$

תכונות של מיזוג ופיצול תהליך פואסון

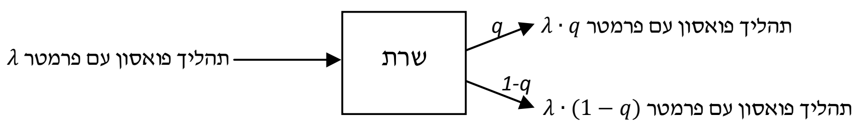
1. סכום של מספר תהליכי פואסון בת"ס הוא תהליך פואסון עם סכום קצבי ההגעות



הוכחה: ע"פ תכונת האדיטיביות + תוספות בת"ס.

הערה: שני תהליכים אקראיים הם בת"ס אם כל קבוצה של דגימות של התהליך הראשון היא בת"ס בכל קבוצת דגימות של התהליך השני.

2. פיצול של תהליך פואסון:



כל אירוע מכוון בהסתברות q למעלה ובהסתברות $1-q$ למטה, כך שתהליך הפיצול בת"ס בתהליך הכניסה וההחלטות בזמנים שונים גם הן בת"ס.

תהליכי הפיצול של תהליך הפואסון שבכניסה הם בעצמם תהליכי פואסון עם קצבים $\lambda\cdot q$ (למעלה) ו- $\lambda\cdot(1-q)$ (למטה), בת"ס זה בזה.

הוכחה: תחת הגדרה III כל אחד מתתי התהליכים הוא סדרת ברנולי עם הסתברות להצלחה קטנה פי q או פי $1-q$ ביחס להסתברות ההצלחה של סדרת הברנולי שבכניסה.

הגדרה IV: הגרלת הגעות אחידה וסידור כרונולוגי

נגדיר $N_n(t)$, $0\le t\le {n}/{\lambda}$ כהגרלה של n נקודות בפילוג אחיד $i.i.d$ באינטרוול הזמן $\left(0,{n}/{\lambda}\right)$.
לאחר ההגרלה נסמן את הנקודות לפי סדר עולה בזמן $T_1,T_2,\dots,T_n$ (סידור "כרונולוגי").

כאשר $n\to\infty$, $N_n(t)$ מתפלג כתהליך פואסון בעל פרמטר λ.

טענת עזר: עבור תהליך פואסון, בהינתן ש- $T_n=t$, הפילוג המותנה של $n-1$ ההגעות הראשונות הוא אחיד באינטרוול $(0,t)$, בת"ס ביניהן, עם סידור כרונולוגי.

הוכחה: ראו תרגול כיתה.

טבלת השוואה וסיכום: תהליכי פואסון ו-וינר

	תהליך פואסון $N(t)$ בעל קצב הגעות λ	תהליך וינר $W(t)$ עם פרמטר α
רציפות	תהליך בזמן רציף המקבל ערכים בדידים	תהליך בזמן רציף המקבל ערכים רציפים
פילוג שולי	$Poisson(\lambda \cdot t)$ $Pr(X(t) = k) = e^{-\lambda \cdot t} \cdot \frac{(\lambda \cdot t)^k}{k!}$	$N(0, \alpha \cdot t)$ $f_X(x) = \frac{1}{\sqrt{2\pi\alpha \cdot t}} \cdot e^{-\frac{x^2}{2\alpha \cdot t}}$
תלות בין זמנים	מרקובי, תוספות בת״ס	מרקובי, תוספות בת״ס
תוחלת	$\eta(t) = \lambda \cdot t$	$\eta(t) = 0$
שונות	$Var[N(t)] = \lambda \cdot t$	$Var[W(t)] = \alpha \cdot t$
פונקצית אוטו קוריאנס	$C_N(t_1, t_2) = \lambda \cdot \min\{t_1, t_2\}$	$C_W(t_1, t_2) = \alpha \cdot \min\{t_1, t_2\}$
סטציונאריות	לא סטציונארי	לא סטציונארי
תהליך הנגזרת	רעש לבן + איבר קבוע הנגזרת היא תהליך אקראי סטאציונארי במובן הצר.	רעש לבן הנגזרת היא תהליך אקראי סטאציונארי במובן הצר.