הפקולטה להנדסה

אוניברסיטת תל אביב

אותות אקראיים ורעש

0512.3632

על פי הרצאותיו של פרופסור אורי ארז סמסטר א' תשע"ד

סוכם, נכתב ונערך ע"י ערן אבידור

מהדורה 1.1

14/01/14

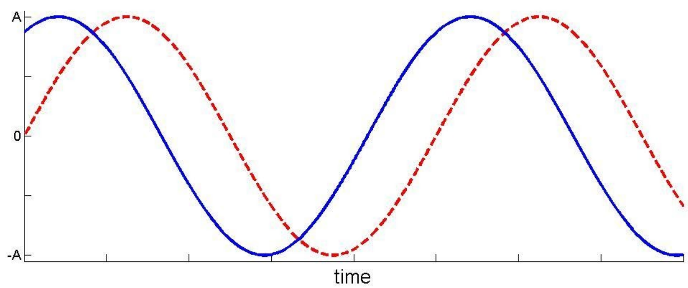
תוכן עניינים

מבוא/מטרה

תהליכים אקראיים נמצאים בכל מקום. אפילו במערכות דטרמיניסטיות לכאורה קיימים דברים שלא ידועים לחלוטין (מידת אקראיות מסויימת). אנו נמדל את אי-הידיעה על דברים אלו כרכיב אקראי של המערכת.

<u>דוגמא 1 - "מעט אקראיות":</u>

נסתכל על מתח בשקע חשמל ביתי ברגע הדגימה. אנו יודעים כי נקבל דגימה מתוך גל סינוס אך לא נדע באיזה מיקום הגל כי הפאזה לא ידועה לנו.



$$V(t) = A \cdot \sin(2\pi ft + \phi)$$
$$\phi \sim Unif\,(0, 2\pi]$$

למעשה $V(t)$ אקראי לכל זמן t מאחר והפאזה ϕ אינה ידועה. מכיוון שאין תיעדוף לפאזה מסויימת מתקיים $\phi \sim Unif\,(0,2\pi]$. נרצה לדעת מה ההסתברות לקבל ערך כלשהו עבור $V(t)$ בזמן דגימה מסויים (עבור פילוג לזמן כלשהו נוכל לדעת את הפילוג לכל שאר הזמנים כי התנהגות סינוס ידועה), נניח $t=0$.

למעשה נשאל איך נראה $f_{V(0)}(v)$. במהלך הקורס נראה כי הסיכוי לקבל $V(0)=A$ גדול מהסיכוי לקבל $V(0)=0$.

<u>דוגמא 2 – "הרבה אקראיות":</u>

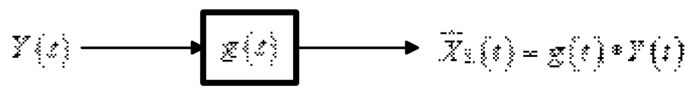
שני דוברים נמצאים בחדר, דובר א' (ילד) ודובר ב' (מבוגר). נסמן את אותות הדיבור שלהם ע"י $X_1(t)$ ו- $X_2(t)$ בהתאמה.

נניח גם שהדוברים נייחים בחדר ולכן נוכל למדל את המערכת (חדר, החזרות וכו') כמערכת ליניארית. נסמן את פונקציות התמסורת הכוללות (סך הגברה והחזרות) של כל אות ע"י $h_1(t), h_2(t)$ (אם הדוברים היו במרחב פתוח אז פונקציות התמסורת היו קבועות בזמן).

נוכל לייצג את אות הדיבור בחדר ע"י האות $Y(t) = h_1(t) * X_1(t) + h_2(t) * X_2(t)$.

נניח שאנו מעוניינים להאזין לדובר א'. שני האותות $X_1(t), X_2(t)$ מכילים מידע אך האחד רצוי (בהתאם לדרישות המערכת, אצלנו $X_1(t)$) והשני רעש.

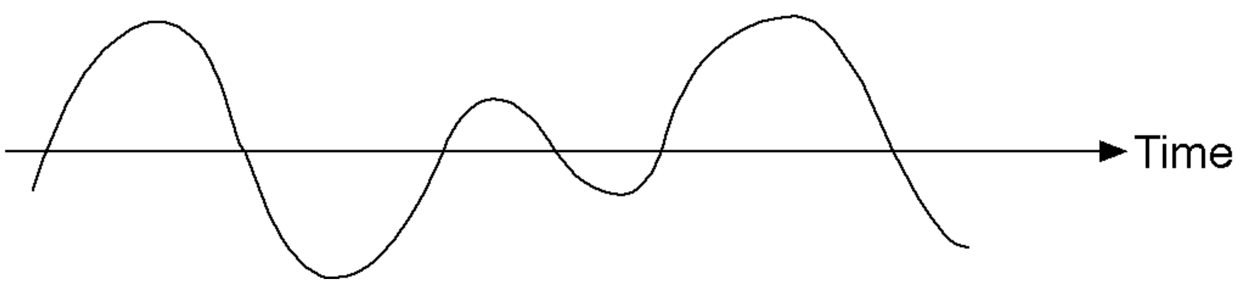
מטרתנו היא למעשה לסנן את הרעש כך שנקבל "שערוך" טוב של אות המידע הרלוונטי $X_1(t)$.
מוח האדם עושה זאת בקלות ובאינטואיטיביות מדהימה. ע"מ לתכנן מערכת הנדסית כזאת נרצה לדעת מידע מקדים על אופי האותות. אנו יודעים שאלו אותות דיבור של ילד ומבוגר ולכן האותות שונים מבחינה ספקטרלית, דהיינו האותות קיימים בתחומי תדר שונים (ילד מדבר בקול גבוה יותר מאשר מבוגר).
ממידע מקדים זה נתכנן מסנן מתאים (אצלנו מדובר במסנן מעביר גבוהים / מנחית נמוכים פשוט) וכך נשערך את אות האינפורמציה.
למעשה המערכת (מסנן) הזאת היא פשוט פונקציה.



נרצה לדעת את הקשר הסטטיסטי בין מוצא המערכת לכניסתה, ז"א איך הפונקציה משפיעה על אות הכניסה (אשר כולל אי-ודאות כלשהו) ועל מאפייניו.
באופן כללי נרצה למדל את האות/התהליך האקראי, לבצע ניתוחים בתחום כלשהו (לדוגמא תחום התדר) ואז לבנות משערך איכותי לפי המדד המבוקש.
במהלך הקורס נלמד להתמודד עם אותות כניסה מהסוגים הבאים:

- משתנה אקראי (מ"א) יחיד
- ווקטור אקראי (ו"א)
- תהליך אקראי (ת"א) בזמן בדיד
- תהליך אקראי (ת"א) בזמן רציף

לצורך קבלת אינטואיציה לגביי המצבים השונים נניח קיום של אות כלשהו $X(t)$ אשר קיימת אי-וודאות לגביו.



ממנו נגדיר את ארבעת המצבים השונים:

- מ"א יחיד יהיה ערך האות בנקודת זמן כלשהי $X(t_0)$.
- ו"א יהיה סט דגימות סופי של האות $\underline{X}(t_0,t_1,...,t_n) = (X(t_0), X(t_1),\dots,X(t_n))$.
- ת"א בזמן בדיד יהיה סט הדגימות האינסופי במרווחי זמן קבועים

$$\{X_i\}_{i=1}^{\infty} = \{X(n) \quad \forall n \in \mathbb{N}\}.$$

- ת"א בזמן רציף הוא למעשה התהליך $X(t)$ בשלמותו (ז"א לכל $t\geq 0$ $t\in\mathbb{R}$).

חלק א': משתנים אקראיים ופעולות עליהם

הסתברות ומשתנים אקראיים

חזרה על מושגי יסוד בהסתברות

מרחב הסתברות (הגדרה)

מרחב הסתברות מוגדר ע"י השלשה $\{\Omega, F, P\}$.

- Ω - מרחב המדגם:
קבוצת כל התוצאות האפשריות בניסוי. מרחב המדגם יכול להיות סופי (לדוגמא הטלת קובייה) או אינסופי (לדוגמא בחירת מספר רציונלי).
- F - שדה המאורעות:
מאורע הינו תת-קבוצה של מרחב המדגם Ω. ז"א מאורע הוא למעשה "תרחיש אפשרי". שדה המאורעות מכיל רק (ואת כל) את השאלות שניתן לשאול לגביי ה"ניסוי".

מלבד כך ש- F מוכל עד כדי שיוויון ב- Ω ($F \subseteq \Omega$) על חלות הדרישות המתמטיות הבאות:

I. שדה המאורעות F סגור תחת איחוד בן-מניה, חיתוך בן-מניה וכן תחת משלים. זאת למעשה על מנת שכל שאלה טבעית של תרחישים בנוגע לניסוי תהיה מוגדרת.

II. האיברים Ω ו- $\varnothing$ נמצאים בשדה המאורעות F (ז"א $\varnothing, \Omega \in F$).

- P - פונקציית מידת ההסתברות:

פונקציית מידת ההסתברות הינה פונקציה מן שדה המאורעות לאינטרוול הסגור $[0,1]$.

פונקציה זו מתארת את הסיכוי לקבלת מאורע. $P: F \to [0,1]$.

על השלשה $\{\Omega, F, P\}$ לקיים את הדרישות הבאות:

1. $\forall A \in F : P(A) \geq 0$
2. $P(\Omega) = 1$, זאת אומרת מתקיים כי הסתברות המאורע הוודאי הינה 1.
3. אם $\{A_i\}$ היא קבוצת מאורעות זרים בזוגות, ז"א $\forall i \neq j \quad A_i \cap A_j = \varnothing$ אזי $P(\cup A_i) = \sum P(A_i)$.

מתוך דרישות אלו נובעות התכונות הבאות:

1. $\forall A \in F : P(A) \leq 1$

נוכיח:

$$1 = P(\Omega) = P\left(A \cup \overline{A}\right) = P(A) + P\left(\overline{A}\right)$$
$$\underset{P(A),P(\overline{A})\geq 0}{\Rightarrow} P(A) \leq 1 \quad \forall A \in F$$

2. $P(\varnothing) = 0$

3. $P\left(\overline{A}\right) = 1 - P(A)$

<u>אי תלות בין מאורעות (הגדרה):</u>

המאורעות $A, B \in F$ הינם בלתי תלויים סטטיסטית (בת"ס) אמ"מ $P(A \cap B) = P(A) \cdot P(B)$.

<u>דוגמא 1:</u>

נסתכל על ניסוי הטלת קובייה הוגנת.

- מרחב המדגם הינו $\Omega = \{1,2,3,4,5,6\}$.
- שדה המאורעות יהיה כל תתי-הקבוצות האפשריות של Ω, ז"א את קבוצת החזקה של Ω.

$$F = \{\varnothing, \{1\}, \{1,2\}, \ldots, \{1,2,3,4,5,6\}\} \quad |F| = 2^{|\Omega|} = 2^6 = 64$$

- פונקציית מידת ההסתברות המלאה P מתאימה מספר עבור כל אחד מ- 64 המאורעות.
למעשה, מספיק להגדיר את P רק על המאורעות האלמנטריים $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$ ומערכים אלו נוכל לחשב מידת ההסתברות עבור כל מאורע שנתעניין בו בעזרת הדרישות והתכונות של מרחב הסתברות.
מתקיים $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$.
ונראה דוגמא:

$$P(even \quad result) = P(\{2,4,6\}) = P(\{2\} \cup \{4\} \cup \{6\}) =$$
$$= P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

<u>הערה:</u>

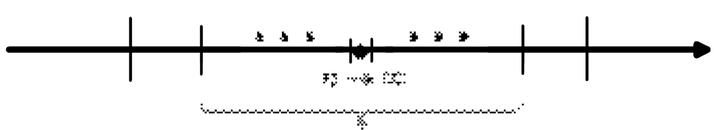
על אותו מרחב מדגם (ואפילו על אותו ניסוי ספציפי) ניתן להגדיר מספר שדות מאורעות שונים.
לדוגמא נגדיר את שדה המאורעות לניסוי זוגי/אי-זוגי $F_2 = \{\varnothing, \{1,3,5\}, \{2,4,6\}, \{1,2,3,4,5,6\}\}$.
אך למעשה אין טעם להגדיר שדה מאורעות כזה מאחר והוא מוכל בשדה המאורעות המלא.
לכן כל עוד מרחב המדגם סופי (או אפילו בן-מניה) נעדיף לעבוד עם קבוצת החזקה של מרחב

המדגם כשדה המאורעות ואת פונקציית מידת ההסתברות נגדיר על המאורעות האלמנטריים בלבד.

<u>דוגמא 2:</u>

הגרלה אחידה של נקודה באינטרוול הסגור $[0,1]$.
למעשה זהו התרחיש הכי "קרוב" לקובייה הוגנת אך עם מרחב מדגם רציף.

- מרחב המדגם הינו $\Omega=[0,1]$.
- שדה המאורעות יהיה כל תתי-הקבוצות האפשריות של Ω , ז"א קבוצת כל האינטרוולים הפנימיים ב- $[0,1]$. חשוב להבחין כי ב- F ישנם גם אינטרוולים לא רציפים. למעשה בשדה המאורעות ישנם כל האינטרוולים מהצורה $[a,b]\subseteq[0,1]$ וכן איחוד/חיתוך/משלים של כל קומבינציה אפשרית מהם.
  למעשה אם נרצה להגיע לנקודה בודדת נבצע סדרת אינסופית של חיתוכים כך שבכל שלב נקטין את האינטרוול אך עדיין נשאר עם אינטרוול ולא נקודה (בגבול נקבל אינטרוול אינפיטיסימלי אשר ייצג נקודה ברמת "דיוק" מסויימת). הסיבה לכך היא שנקודה אינה תת קבוצה של מרחב המדגם ולכן לא ניתן להגדיר את P עבורה (הסיכוי ליפול על נקודה מסויימת הוא 0 ואם היינו מגדירים את שדה המאורעות כאוסף כל הנקודות האפשריות היינו מקבלים שמידת ההסתברות עבור כל מאורע תהיה 0).



- פונקציית מידת ההסתברות P תוגדר עבור מאורעות "בסיסיים" שהם אינטרוולים (לאו דווקא סגורים מאחר ומידת ההסתברות של נקודה יחידה היא אפס) רציפים והיא מקיימת:

$$P([a,b])=\frac{b-a}{1-0},\quad [a,b]\subseteq[0,1]$$

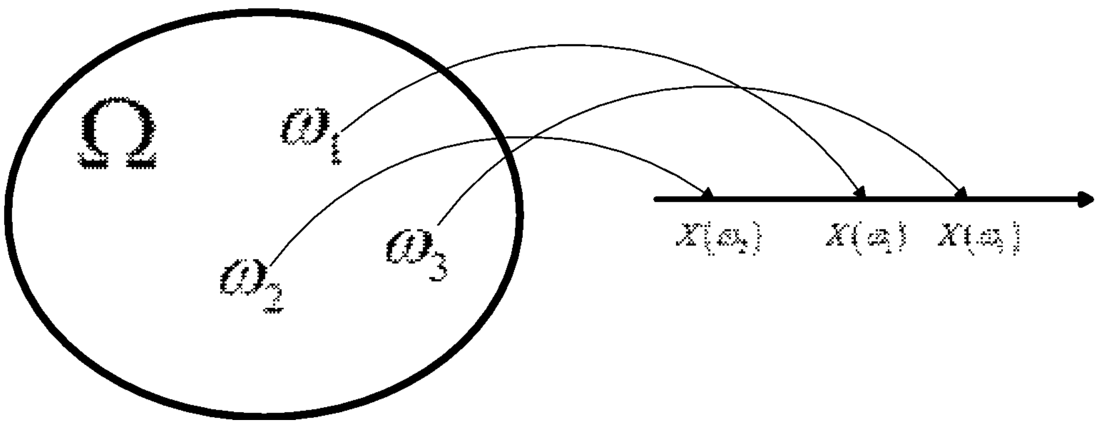
<u>חזרה על מושגי יסוד עבור משתנה אקראי יחיד</u>

נרצה "להיפטר" מהשפה המופשטת של מרחבי הסתברות והשלשות $\{\Omega,F,P\}$ ולדון בשפה יותר הנדסית. "שפה" זו מופשטת מאחר ואיברי Ω הינם מופשטים (עבור כל ניסוי הקבוצה בעלת אופי שונה ולאיברי הקבוצה אין משמעות ערכית/מתמטית לדוגמא בניסוי הטלת מטבע $\Omega=\{Heads,Tail\}$ ובניסוי קובייה $\Omega=\{1,2,3,4,5,6\}$).

לכן, על מנת "להיפטר" משפה מופשטת זו נרצה למפות את איברי Ω לשפה אחידה, שפת המספרים הממשיים.

<u>משתנה אקראי (הגדרה)</u>

משתנה אקראי (מ"א) X הינו מיפוי (פונקצייה) המשייך לכל תוצאת ניסוי אפשרית $\omega \in \Omega$ מספר ממשי $X(\omega) \in \mathbb{R}$, כך שמתקיימת הדרישה $\{\omega : X(\omega) \le x\} \in F$ $\forall x \in \mathbb{R}$. ז"א עבור כל $x \in \mathbb{R}$ הקבוצה $\{\omega : X(\omega) \le x\}$ תהיה מאורע (שייכת לשדה המאורעות).



<u>הערות:</u>

1. מ"א יסומן תמיד באות גדולה.
2. בעיקרון המיפוי יכול להיות שרירותי (כל עוד מתקיימת הדרישה בהגדרה) אך בפועל בהנדסה המיפוי יהיה טבעי ומתבקש (וכך גם הדרישה תתקיים).
3. הדרישה נחוצה ע"מ להגדיר את פונקציית ה-CDF (בהמשך).

<u>דוגמא:</u>

עבור ניסוי הטלת קובייה מתקיים $\Omega = \{1,2,3,4,5,6\}$ ונוכל להגדיר מ"א

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

כלומר מ"א בעל ערכים 1 (עבור תוצאת הטלה אי-זוגית) ו- 2 (עבור תוצאת הטלה זוגית).

פונקציית התפלגות מצטברת ופונקציית צפיפות הסתברות

החלפנו את תוצאות הניסוי המופשטות בערכים מספריים ע"י מיפוי. כעת נרצה לפתח כלים אשר לחישוב מידת ההסתברות עבור ערכים המתאימים למשתנה האקראי. ז"א, נרצה להגדיר פונקציות כך שבעזרתן נוכל לענות על כל שאלה הסתברותית על ניסוי מסויים (לדוגמא $\Pr(X \in S)$).

פונקציית התפלגות מצטברת – Cumulative Distribution Function

פונקציית התפלגות מצטברת מסומנת באות F עם אינדקס שם המשתנה האקראי וארגומנט המייצג את הערך המבוקש. ערכה מייצג את ההסתברות לקבל ערך לא גדול (נמוך וכולל) מן ערך הארגומנט.

עבור מ"א X נגדיר ונסמן את פונקציית ה-CDF באופן הבא:

$$F_X(x) \triangleq \Pr(\{\omega : X(\omega) \le x\}) \overset{*}{=} \Pr(X \le x)$$

(*) חישוב מידת ההסתברות עובד על מאורעות אך נכתוב ע"י ערכי המ"א לשם פשטות ונוחות.

<u>תכונות:</u>

1. $0 \le F_X(x) \le 1$
2. $\lim\limits_{x\to\infty} F_X(x) = 1$
3. $\lim\limits_{x\to-\infty} F_X(x) = 0$
4. $F_X(x_0) = \lim\limits_{x\to x_0^+} F_X(x) = F_X(x_0^+)$ (רציפות מימין)
5. $\forall x_2 > x_1 : F_X(x_2) \ge F_X(x_1)$ (פונקציה לא יורדת)
6. $\Pr(a < X \le b) = F_X(b) - F_X(a)$ מאחר ומתקיים:

$$\Pr(a < X \le b) = \Pr(-\infty < X \le b) - \Pr(-\infty < X \le a) = F_X(b) - F_X(a)$$

<u>הערות:</u>

1. מתקיים $\Pr(X \le a) = F_X(a)$ (נובע מתכונות 6 ו-3 או ממשמעות ההגדרה).
2. עבור גבול הסיכוי הנקודתי $\Pr(X = a) = F_X(a^+) - F_X(a^-) \overset{*}{=} F_X(a) - F_X(a^-)$.

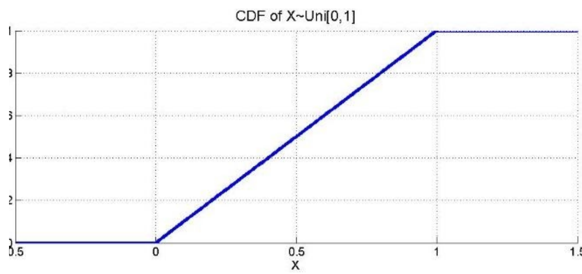
(*) נובע מרציפות מימין (תכונה 4).

3. תכונה 4 הינה תכונה שרירותית אשר נובעת מההגדרה אשר כוללת שיוויון. אילו היינו מסירים את השיוויון אז היינו מקבלים רציפות דווקא משמאל.

<u>דוגמאות:</u>

הגרלה אחידה בקטע $[0,1]$

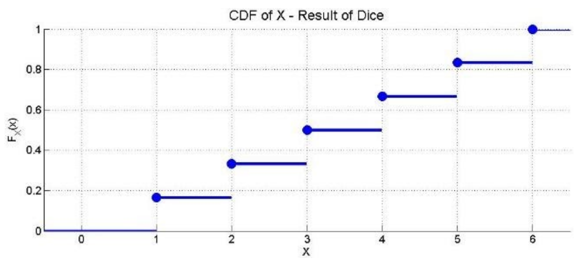
נגדיר מיפוי $X(\omega)=\omega$ (כמובן שמתקיים $\omega\in\Omega=[0,1]$) ונקבל:



הפונקציה לא גזירה בכל נקודה אך רציפה לחלוטין.

הטלת קובייה הוגנת

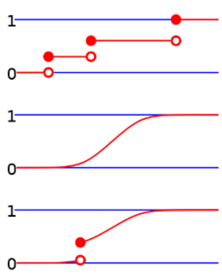
נגדיר מיפוי "טבעי" ונקבל:



$$F_X(x)=\sum_i \Pr(X=x_i)u(x-x_i)=\frac{1}{6}\sum_{i=1}^{6}u(x-i)$$

<u>הבחנה:</u>

- מ"א X ייקרא בדיד אם $F_X(x)$ מורכבת ממדרגות (קפיצות) בלבד.
- מ"א X ייקרא רציף אם $F_X(x)$ רציפה לכל $x\in\mathbb{R}$.
- מ"א X ייקרא מעורב אם $F_X(x)$ מורכבת גם ממדגות (קפיצות) וגם מקטעים רציפים.



נוכל לכתוב את פונקציית ה-CDF עבור כל מ"א ע"י:

$$F_X(x)=\alpha F_X{}^D(x)+(1-\alpha)F_X{}^C(x)\quad 0\le\alpha\le 1$$

כאשר $F_X{}^D(x)$ הינה פונקציית ה-CDF הבדידה (Discrete) ו- $F_X{}^C(x)$ הינה פונקציית ה-CDF הרציפה (Continuous).

<u>הערה:</u>

עבור מ"א כללי, נקודות עם הסתברות לא אפס נקראות אטומים. אלו הנקודות עם הקפיצה ב-CDF ונאמר שיש בהן מסת הסתברות.

<u>פונקציית צפיפות ההסתברות – Probability Density Function</u>

פונקציית צפיפות ההסתברות מסומנת באות f עם אינדקס שם המשתנה האקראי וארגומנט המייצג את הערך המבוקש. ערכה מייצג את צפיפות ההסתברות סביב נקודה מסויימת (ערך הארגומנט).

עבור מ"א X פונקציית ה-PDF שלו תוגדר להיות הפונקציה המקיימת:

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

<u>תכונות:</u>

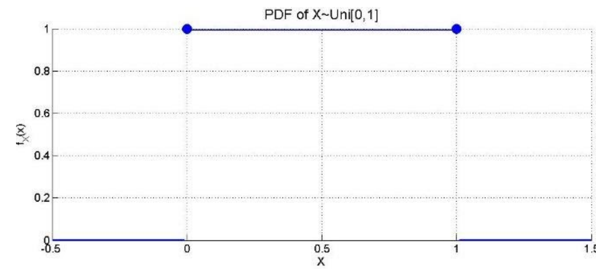
1. $\int_{-\infty}^{\infty} f_X(x)\,dx=1$
2. $\forall x \quad f_X(x)\geq 0$
3. $\Pr(a<X\leq b)=F_X(b)-F_X(a)=\int_{a}^{b} f_X(x)\,dx$ (נובע מהגדרת פונקציית ה-PDF).

<u>הערות:</u>

1. עבור $F_X(x)$ גזירה לכל x (גזירה זו דרישה חזקה יותר מרציפה) אז מתקיים $f_X(x)=\frac{d}{dx}F_X(x)$.
2. $f_X(x)$ אינה יחידה. הדבר נובע מערכים נקודתיים אשר לא ישפיעו על האינטגרל. נקודות אלו קיימות בנקודות אי-גזירות של $F_X(x)$ וניתן לייחס ל- $f_X(x)$ כל ערך סופי בהן.

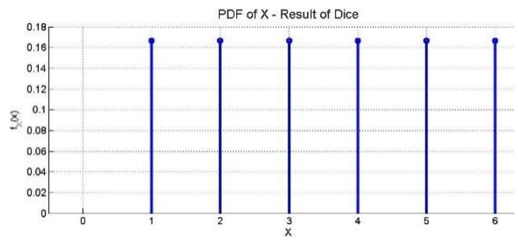
<u>דוגמאות:</u>

הגרלה אחידה בקטע $[0,1]$



בנקודות $x=0^-,1^+$ ניתן לשים כל ערך סופי שנרצה

הטלת קובייה הוגנת



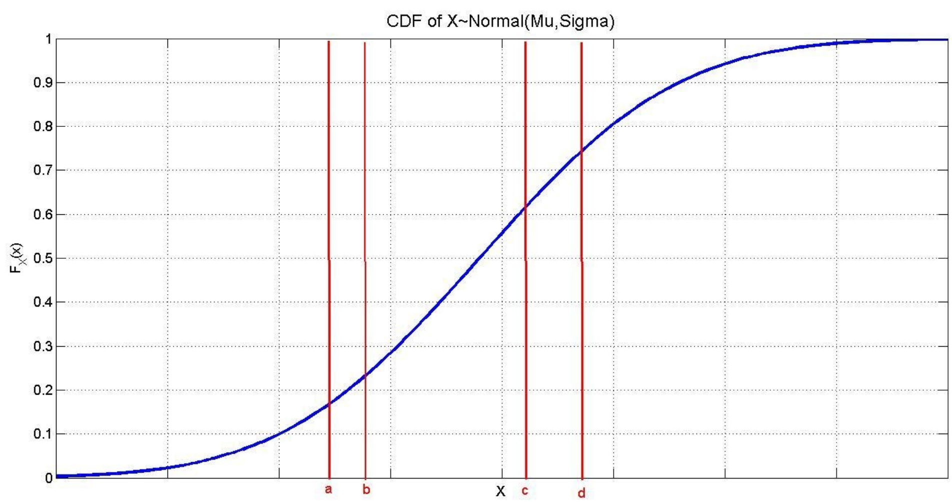
$$f_X(x)=\frac{d}{dx}F_X(x)=\frac{1}{6}\sum_{i=1}^{6}\delta(x-i)$$

<u>סיכום</u>

כעת נוכל לענות כל כל שאלה מהצורה $\Pr(X\in S)$ כאשר $S\in\mathbb{R}$ תוך שימוש בפונקציות $f_X(x),F_X(x)$.

לצורך המחשה נראה דוגמא פשוטה של מ"א רציף (מ"א גאוסי – נלמד בהמשך).

$$S=\{[a,b],[c,d]\}$$

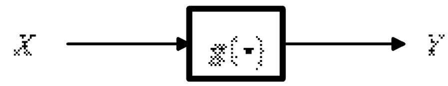


$$\Pr(x\in S)=\Pr(x\in\{[a,b],[c,d]\})=\Pr(x\in[a,b]\cup[c,d])=\Pr(x\in[a,b])+\Pr(x\in[c,d])\overset{*}{=}$$
$$=[F_X(b)-F_X(a)]+[F_X(d)-F_X(c)]$$

כאשר המעבר $(*)$ מוצדק מאחר ומתקיים $\Pr(a<X\le b)=F_X(b)-F_X(a)$ (הוכח בתכונות CDF). ומאחר והמשתנה האקראי רציף אין מסות הסתברות ולכן $\Pr(a\le X\le b)=\Pr(a<X\le b)$.

<u>פונקציות של משתנים אקראיים</u>

משתנה אקראי X עובר טרנספורמציה כלשהי $g(\cdot)$. בידיעת נתוני ההסתברות $F_X(x)$ ו-$f_X(x)$ (ז"א ידועות פונקציות ה-CDF וה-PDF של הכניסה) ואיפיון המערכת $g(\cdot)$ נרצה לדעת את הסטטיסטיקה של מוצא הטרנספורמציה $Y=g(X)$, כלומר את $f_Y(y),F_Y(y)$.



למעשה קיבלנו פונקציה מורכבת:

$$Y=g(X)=g\big(X(\omega)\big)\quad\Rightarrow\quad Y(\omega)=g\big(X(\omega)\big)$$

פונקצייה מורכבת של מ"א הינה עדיין מ"א מאחר והיא עדיין מיפוי מהצורה $\Omega\to\mathbb{R}$.

נרצה לענות על $\Pr(Y\in S)$ לכל $S\in\mathbb{R}$.

$$\Pr(Y\in S)=\Pr\big(g(X)\in S\big)\overset{*}{=}\Pr\big(X\in g^{-1}(S)\big)$$

כאשר $(*)$ נובע מההגדרה $g^{-1}(S)\triangleq\{x:g(x)\in S\}$ (g לאו דווקא הפיכה).

ז"א כל ה- x-ים אשר $g(\cdot)$ ממפה אותם ל- S.

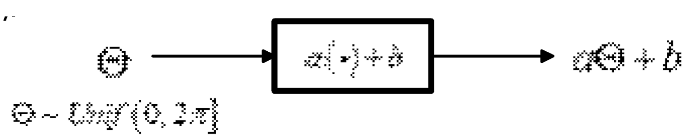
<u>מסקנה:</u>

את פונקציית ה-CDF של Y ניתן לחשב מתוך פונקציית ה-CDF של X:

$$F_Y(y)\triangleq\Pr(Y\le y)=\Pr\big(Y\in(-\infty,y]\big)=\Pr\big(g(X)\in(-\infty,y]\big)=\Pr\Big(X\in g^{-1}\big((-\infty,y]\big)\Big)$$

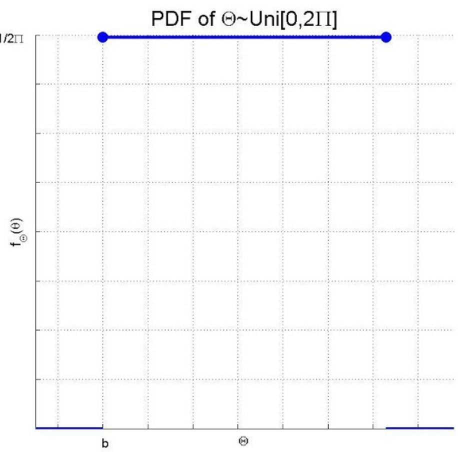
<u>דוגמא:</u>

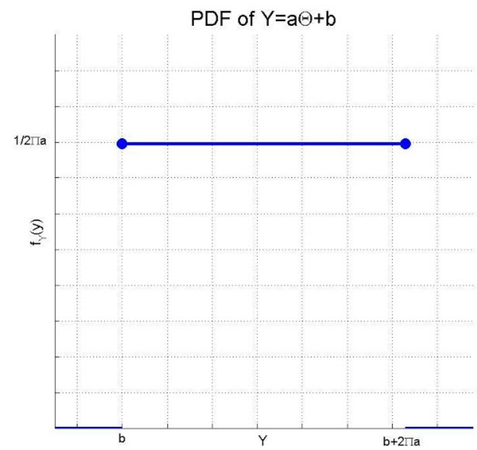
עבור המערכת הבאה נרצה לדעת את הפילוג של המוצא. נבחין בפיתוח ובתוצאות כי יש חשיבות לסימן של a.



$$F_Y(y)=\Pr\left(g(\Theta)\le y\right)=\Pr\left(a\Theta+b\le y\right)=\Pr\left(a\Theta\le y-b\right)=\begin{matrix}a>0\\a<0\end{matrix}\begin{cases}\Pr\left(\Theta\le\dfrac{y-b}{a}\right)\\ \Pr\left(\Theta\ge\dfrac{y-b}{a}\right)\end{cases}=\begin{cases}F_\Theta\left(\dfrac{y-b}{a}\right)\\ 1-F_\Theta\left(\dfrac{y-b}{a}\right)\end{cases}$$

$$\Rightarrow F_Y(y)=\begin{cases}F_\Theta\left(\dfrac{y-b}{a}\right) & a>0\\ 1-F_\Theta\left(\dfrac{y-b}{a}\right) & a<0\end{cases}\quad\underset{\frac{d}{dy}}{\Rightarrow}\quad f_Y(y)=\begin{cases}\dfrac{1}{a}f_\Theta\left(\dfrac{y-b}{a}\right) & a>0\\ -\dfrac{1}{a}f_\Theta\left(\dfrac{y-b}{a}\right) & a<0\end{cases}$$





<u>מומנטים, פונקציה אופיינית ופונקציה יוצרת מומנטים</u>

לא תמיד ידוע CDF/PDF בצורה מדוייקת ואנו יודעים רק חלק מהנתונים הסטטיסטיים. לשם כך נגדיר אפיון סטטיסטי חלקי. מאפיון חלקי נקבל חסם במקום הסתברות מדוייקת.
בנוסף, נוכל לבחון את הרובסטיות (איתנות ויכולת לשמור על מבנה המערכת כנגד שינויים) של מערכת לפי תלות בסטטיסטיקה החלקית (לדוגמא תלות מערכת דרך מומנטים עד סדר שני בלבד או הנחת סדר שני לשם שערוך ליניארי).

<u>מומנט מסדר n של משתנה אקראי</u>

מומנט מסדר n של מ"א מסומן (בקיצור) ע"י m_n ומחושב באופן הבא:

$$m_n(X)=m_n\triangleq E\left[X^n\right]\overset{*}{=}\int_{-\infty}^{\infty}x^n\cdot f_X(x)\cdot dx$$

כאשר $(*)$ מתקיים בתנאי שקיים האינטגרל.

m_n למעשה מייצג את "המרכז המסה של הפילוג" של המ"א X^n.

<u>מומנט מרכזי מסדר n של משתנה אקראי</u>

מומנט מרכזי מסדר n של מ"א מסומן (בקיצור) ע"י μ_n ומחושב באופן הבא:

$$\mu_n(X)=\mu_n=m_n(X-\eta_X)=E\left[(X-\eta_X)^n\right]\overset{*}{=}\int_{-\infty}^{\infty}(X-\eta_X)^n\cdot f_X(x)\cdot dx$$

כאשר $(*)$ מתקיים בתנאי שקיים האינטגרל.

μ_n למעשה מייצג את "המרכז המסה של הפילוג" של המ"א $(X-\eta_X)^n$.

<u>תוחלת</u>

תוחלת של מ"א הינה למעשה מומנט מסדר ראשון של המשתנה האקראי.

תוחלת של מ"א מוגדרת ומסומנת ע"י:

$$\eta_X=m_1(x)=E[X]=\int_{-\infty}^{\infty}x\cdot f_X(x)\cdot dx$$

עבור מ"א בדיד נוכל לרשום:

$$E[X]=\sum_i x_i\cdot\Pr(x_i)$$

לתוחלת משמעות של "מרכז המסה של הפילוג".

<u>הערה/הבחנה:</u> התוחלת הינה פונקצייה ליניארית מתכונות הליניאריות של אינטגרל.

<u>תכונות:</u>

1. משפט התוחלת

אם $Y=g(X)$ אז:

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

2. $E[aX+b]=aE[X]+b$ (ליניאריות התוחלת).

תכונה זו נובעת מההבחנה לגביי ליניאריות התוחלת, משפט התוחלת, ותכונות ה-PDF:

$$E[aX+b]=\int_{-\infty}^{\infty}(ax+b)\cdot f_X(x)\cdot dx=\int_{-\infty}^{\infty}(ax+b)\cdot f_X(x)\cdot dx=\int_{-\infty}^{\infty}ax\cdot f_X(x)\cdot dx+\int_{-\infty}^{\infty}b\cdot f_X(x)\cdot dx=$$

$$=a\int_{-\infty}^{\infty}x\cdot f_X(x)\cdot dx+b\int_{-\infty}^{\infty}f_X(x)\cdot dx=a\cdot E[X]+b\cdot 1=aE[X]+b$$

<u>שונות</u>

שונות של מ"א הינה למעשה מומנט מרכזי מסדר שני של המתשנה האקראי.

שונות של מ"א מוגדרת ומסומנת ע"י:

$$Var(X)=\sigma_x^2=\mu_2(x)=E\left[(X-\eta_X)^2\right]=E\left[X^2-2X\eta_X+\eta_X{}^2\right]=E\left[X^2\right]-2\eta_X E[X]+\eta_X{}^2=$$

$$E\left[X^2\right]-2\eta_X\cdot\eta_X+\eta_X{}^2=E\left[X^2\right]-\eta_X{}^2=E\left[X^2\right]-E^2[X]=m_2-(m_1)^2$$

<u>הערה:</u>

הגדרת כל המומנטים עד סדר 2 של מ"א שקול לידיעת התוחלת והשונות בלבד.

<u>הסבר:</u>

נציג את רשימת כל המומנטים עד סדר שני:

$$m_0 = 1 \quad m_1 = \eta_X \qquad m_2 = E\left[X^2\right]$$

$$\mu_0 = 1 \quad \mu_1 = 0 \quad \mu_2 = \sigma_X^2 = Var\left(X\right) = m_2 - \left(m_1\right)^2$$

מתוך השישה ישנם שלושה טריוויאליים (m_0, μ_0, μ_1 טריוויאליים ונובעים מחישוב פשוט ואינם תלויים בסטטיסטיקה של המשתנה האקראי) ושלושה נוספים אשר מוגדרים כל אחד ע"י השניים האחרים.

<u>סטטיסטיקה מסדר סופי</u>

סטטיסטיקה מסדר k משמעה ידיעת כל המומנטים עד (כולל) סדר k, ז"א ידיעת כל המומנטים $\left\{m_i\right\}_{i=0}^{k}$ (או לחילופין ידיעת $\left\{\left\{\mu_i\right\}_{i=0}^{k}, m_1 = \eta\right\}$, ז"א כל המומנטים המרכזיים עד סדר k ובנוסף את המומנט מסדר ראשון, קרי התוחלת).

אם נתעלם מן הגדלים הטריוויאליים אז סטטיסטיקה מסדר k משמעה ידיעת אחד הסטים:

$$\begin{bmatrix} \left\{m_i\right\}_{i=0}^{k} \\ \left\{\left\{\mu_i\right\}_{i=0}^{k}, m_1 = \eta\right\} \end{bmatrix} = \begin{bmatrix} m_0, m_1, m_2, \ldots, m_k \\ \mu_0, \mu_1, \mu_2, \ldots, \mu_k, m_1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} m_1, m_2, \ldots, m_k \\ \mu_2, \ldots, \mu_k, m_1 \end{bmatrix} = \begin{bmatrix} \left\{m_i\right\}_{i=1}^{k} \\ \left\{\left\{\mu_i\right\}_{i=2}^{k}, m_1 = \eta\right\} \end{bmatrix}$$

<u>הערה:</u>

מבחינה מתמטית ידיעת אחד הסטים שקולה לידיעת הסט השני (מבחינה הנדסית יש הבדל, נניח עבור סדר שני ידיעת X^2 מעידה על הספק בעוד שידיעת $\left(X - \eta_X\right)^2$ מעידה על סטייה ריבועית מן המרכז).

<u>הסבר:</u>

נראה כיוון ראשון, נניח כי ידועים $\left\{\left\{\mu_i\right\}_{i=0}^{k}, m_1 = \eta_X\right\}$. נרצה לבטא את m_k על-ידם.

$$m_k = E\left[X^k\right] = E\left[\left(\left(X - \eta_X\right) + \eta_X\right)^k\right] \overset{*}{=} E\left[\sum_{l=0}^{k} \binom{k}{l} \left(X - \eta_X\right)^l \left(\eta_X\right)^{k-l}\right] = \sum_{l=0}^{k} \binom{k}{l} \mu_l \left(\eta_X\right)^{k-l}$$

כאשר $(*)$ תקף לפי נוסחת הבינום של ניוטון.

נראה כיוון הפוך, נניח כי ידועים $\left\{m_i\right\}_{i=1}^{k}$. נרצה לבטא את μ_k על-ידם.

$$\mu_k = E\left[\left(X - \eta_X\right)^k\right] \overset{*}{=} E\left[\sum_{l=0}^{k} \binom{k}{l} \left(X\right)^l \left(-\eta_X\right)^{k-l}\right] \overset{**}{=} \sum_{l=0}^{k} \binom{k}{l} m_l \left(-m_1\right)^{k-l}$$

כאשר $(*)$ תקף לפי נוסחת הבינום של ניוטון ו- $(**)$ מאחר ומתקיים $\eta_X = m_1$.
כאמור, ידיעת סטטיסטקה חלקית לא תוכל לתת לנו הסתברות מדוייקת אלא רק חסם.

<u>אי-שוויון צ'בישב – Chebyshev's Inequality</u>

אי-שיוויון צ'בישב מתאר את הסיכוי להתרחק מידה מסויימת מן התוחלת.

$$\Pr\left(\left|X-\eta_X\right| \geq a\right) \leq \frac{\sigma_X{}^2}{a^2}, \quad a>0$$

ע"י בחירת $a = k\sigma_X$, ז"א בחירת a כמכפלה שלמה של גודל שורש השונות, אנו שואלים בעצם מה הסיכוי ש- X התרחק מן התוחלת יותר מ- k צעדים של σ_X. במקרה זה נקבל:

$$\Pr\left(\left|X-\eta_X\right| \geq k\sigma_X\right) \leq \frac{1}{k^2}.$$

<u>אי-שוויון מרקוב – Markov's Inequality</u>

אי שוויון מרקוב מתאר את הסיכוי של מ"א אי-שלילי Y לחרוג מרף מסויים $b>0$.

$$\Pr\left(Y \geq b\right) \leq \frac{\eta_Y}{b}$$

<u>הוכחת אי-שוויון מרקוב:</u>

$$\Pr\left(Y \geq b\right) = \int_b^\infty f_Y\left(y\right)dy = \int_b^\infty 1\cdot f_Y\left(y\right)dy \overset{y\geq b}{\leq} \int_b^\infty \frac{y}{b}\cdot f_Y\left(y\right)dy \underset{*}{\overset{\frac{y}{b}>0}{\leq}} \int_0^\infty \frac{y}{b}\cdot f_Y\left(y\right)dy = \frac{1}{b}\int_0^\infty y\cdot f_Y\left(y\right)dy =$$

$$\underset{**}{=} \frac{1}{b}\int_{-\infty}^\infty y\cdot f_Y\left(y\right)dy = \frac{1}{b}E\left[Y\right] = \frac{\eta_Y}{b}$$

כאשר מעבר $(*)$ תקף מאחר והאינטגרנד חיובי ולכן נוכל להרחיב את תחום האינטגרל ומעבר $(**)$ תקף מאחר ומ"א Y הינו אי-שלילי ולכן $\forall y\in\left(-\infty,0\right) \quad f_Y\left(y\right)=0$ (ז"א האינטגרנד מתאפס בכל התחום שהוספנו).

<u>הוכחת אי-שוויון צ'בישב:</u>

עבור מ"א X נגדיר $Y=\left(X-\eta_X\right)^2$. כמובן ש- Y הינו מ"א אי-שלילי.

$$\Pr\left(\left|X-\eta_X\right| \geq a\right) = \Pr\left(\left|X-\eta_X\right|^2 \geq a^2\right) = \Pr\left(Y \geq a^2\right) \overset{\substack{Markov's\\Inequality}}{\leq} \frac{\eta_Y}{a^2} = \frac{\sigma_X{}^2}{a^2}$$

<u>הערה:</u>

ידיעת מומנטים מכל סדר (ז"א ידיעת הסט האינסופי של המומנטים) שקולה לסטטיסטיקה מלאה (ראה הערה בסוף הנושא פונקציה אופיינית ופונקציה יוצרת מומנטים).

<u>פונקציה אופיינית</u>

פונקציה אופיינית (פ"א) של מ"א X היא התמרת פורייה של ה-PDF של X (עד כדי מינוס בחזקה).

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

<u>הערות:</u>

- נבחין למעשה כי זו פונקצייה של מ"א $Y(X)=e^{j\omega X}$, פונקצייה זו יוצרת מ"א מרוכב.
- מאחר והפונקציה האופיינית נוצרת מהתמרת פורייה אז תכונות פורייה תקפות.

<u>תכונות:</u>

1. פ"א הינה תיאור אלטרנטיבי של מידע סטטיסטי מלא ובכך שקולה לידיעת ה-CDF שלו (וגם ה-PDF שלו אך למעשה הוא לא תמיד מוגדר). לכן מבחינת מידע סטטיסטי $\Phi_X(\omega)\Leftrightarrow F_X(x)\Leftrightarrow f_X(x)$.

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

2. $\Phi_X(0)=1$ (הוכחה: $\Phi_X(0)=\int_{-\infty}^{\infty} e^{0} f_X(x)\,dx=\int_{-\infty}^{\infty} f_X(x)\,dx=1$).
3. $\left|\Phi_X(\omega)\right|\le 1 \quad \forall\omega$ (הוכחה: $\left|\Phi_X(\omega)\right|=\left|\int_{-\infty}^{\infty} e^{j\omega x} f_X(x)\,dx\right|\le\int_{-\infty}^{\infty}\left|e^{j\omega x}\right| f_X(x)\,dx \overset{\left|e^{j\omega x}\right|=1}{=} 1$).
4. הקשר בין הפונקציה האופיינית לבין המומנטים נתון ע"י $\left.\frac{d^n\Phi_X(\omega)}{d\omega^n}\right|_{\omega=0}=j^n\cdot m_n$.

הוכחה:

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

5. פונקציה אופיינית הינה פונקציה ליניארית (נובע מליניאריות של אינטגרל ותכונות אקספוננט).

$$Y=a\cdot X+b \quad\Rightarrow\quad \Phi_Y(\omega)=e^{j\omega b}\cdot\Phi_X(a\omega)$$

הוכחה:

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

<u>פונקציה יוצרת מומנטים</u>

פונקציה יוצרת מומנטים (Moment Generating Function) של מ"א X היא התמרת לפלס של ה-PDF של X.

$$M_X(s)=\int_{-\infty}^{\infty}e^{sx}f_X(x)dx=E\left[e^{sX}\right]$$

<u>הערה:</u>
פונקציה אופיינית תמיד רציפה בהחלט (התמרת פורייה) ואילו בפונקציה יוצרת מומנטים יש לעיתים בעיות התכנסות (התמרת לפלס) אך היתרון שלה הוא שאין את המקדם המרוכב j.

<u>תכונות:</u>
1. פונקציה יוצרת מומנטים הינה תיאור אלנרנטיבי של מידע סטטיסטי מלא.
2. הקשר בין הפונקציה היוצרת לבין המומנטים נתון ע"י $\left.\frac{d^n M_X(s)}{ds^n}\right|_{s=0}=m_n$.

<u>חסם צ'רנוף – Chernoff Bound</u>

חסם צ'רנוף מתאר את הסיכוי של מ"א כלשהו X לחרוג מרף מסויים a.

$$\Pr(X\ge a)\le e^{-sa}M_X(s)$$

<u>הערה:</u>
באגף ימין של החסם אנו משתמשים בפונקציה יוצרת מומנטים. ידיעתה שקולה לידיעת סטטיסטיקה מלאה, לכן מדוע שנסתפק רק בחסם אם בידינו כל הסטטיטיקה?
התשובה היא שחישוב החסם דרך הפונקצייה היוצרת הרבה יותר פשוט ומאוד שימושי (מתקבל חסם הדוק) לחישוב הסתברויות עם רף נמוך למ"א מהצורה $X=\sum_i a_i X_i \quad \{X_i\}\ i.i.d$.

<u>i.i.d – תזכורת</u>
סט מ"א הינם Independent Identically Distributed (בת"ס ושווי פילוג) אם לכולם יש אותה פונקציית PDF ויש אי-תלות סטטיסטית ביניהם (לא רק בזוגות אלא בכלל).

<u>הוכחת חסם צ'רנוף:</u>

$$\Pr(X\ge a)\overset{s>0}{=}\Pr(sX\ge sa)=\Pr\left(e^{sX}\ge e^{sa}\right)\overset{\substack{Markov's\\Inequality}}{\le}\frac{M_X(s)}{e^{sa}}=e^{-sa}M_X(s)$$

<u>הערה:</u>
כאמור, ידיעת המומנטים מכל סדר שקולה לסטטיסטיקה המלאה. נניח כי ידועה הקבוצה $\{m_i\}_{i=0}^{\infty}$.

נפתח את הפונקציה האופיינית לטור מק-לורן (טיילור סביב הראשית $\omega=0$):

$$\Phi_X(\omega)=\sum_{n=0}^{\infty}\frac{1}{n!}\frac{d^n\Phi_X(\omega)}{d\omega^n}\bigg|_{\omega=0}\omega^n=\sum_{n=0}^{\infty}\frac{1}{n!}j^n m_n\omega^n=\sum_{n=0}^{\infty}\frac{(j\omega)^n}{n!}m_n=$$

ואכן הבענו את הפונקציה היוצרת (אשר שקולה לידיעת פונקציית ה-PDF) ע"י קבוצת המומנטים. כמובן שפיתוח הטור תקף רק אם הפונקציה האופיינית אנליטית.
תנאים מדוייקים הינו מסובכים ומיותרים לקורס, תנאי מספיק (ואף מחמיר) לאנליטיות הוא שהמ"א X חסום, כלומר קיים A כך שתמיד $|X|\le A$, או $\Pr(-A\le X\le A)=1$.

וקטורים אקראיים

וקטור אקראי הינו הכללה של משתנה אקראי. הוא סט של מספר מ"א (אשר יכולים להיות תלויים אחד בשני, בחלקם, בכולם או כלל לא). למעשה הוא סט של מספר מ"א אשר מוגדרים על אותו מרחב הסתברות $X_1(\omega),\ldots,X_n(\omega)$.

מתמטית, וקטור אקראי $\underline{X}$ מוגדר על אותו מרחב הסתברות $\{\Omega,F,P\}$ כמו מ"א אך עם מיפוי $\underline{X}:\Omega\to\mathbb{R}^n$.

אנו נדרוש שמיפוי יקיים כי הקבוצה $\{\omega:X_1(\omega)\le x_1,\ldots,X_n(\omega)\le x_n\}$ הינה מאורע ב- F לכל $\underline{x}\in\mathbb{R}^n$.

אנו ננהג להציג וקטור אקראי כווקטור עמודה:

$$\underline{X}(\omega)=\begin{bmatrix}X_1(\omega)\\ \vdots\\ X_n(\omega)\end{bmatrix}\qquad \underline{X}:\Omega\to\mathbb{R}^n$$

כמו במ"א נתאר את המידע ההסתברותי על הווקטור האקראי $\underline{X}$ ע"י:

- jCDF
- jPDF
- פונקציה יוצרת מומנטים (משותפת)

הערה: j עבור joint (משותף – לכל המשתנים האקראיים).

פונקציית התפלגות מצטברת משותפת ופונקציית צפיפות הסתברות משותפת

פונקציית התפלגות מצטברת משותפת – Joint Cumulative Distribution Function (jCDF)

פונקציית ה-jCDF מוגדרת באופן הבא:

$$F_{\underline{X}}(\underline{x})\triangleq\Pr\left(\{\omega:X_1(\omega)\le x_1,\ldots,X_n(\omega)\le x_n\}\right)\overset{*}{=}\Pr(X_1\le x_1,\ldots,X_n\le x_n)=\Pr(\underline{X}\le\underline{x})$$

(*) חישוב מידת ההסתברות עובד על מאורעות אך נכתוב ע"י ערכי המ"א לשם פשטות ונוחות.

<u>תכונות:</u>

1. $0\le F_{\underline{X}}(\underline{x})\le 1\quad \forall\underline{x}\in\mathbb{R}^n$
2. $F_{\underline{X}}(\underline{x})$ היא פונקצייה מונוטונית לא-יורדת בכל אחד מן המשתנים $x_i\quad i=1,2,\ldots,n$.
3. $F_{\underline{X}}(\underline{x})$ רציפה מימין בכל אחד מן המשתנים $x_i\quad i=1,2,\ldots,n$.

4. $\lim_{\underline{x}\to"\underline{\infty}"} F_{\underline{X}}\left(\underline{x}\right) = \lim_{x_1\to\infty, x_2\to\infty,\ldots,x_n\to\infty} F_{\underline{X}}\left(x_1,x_2,\ldots,x_n\right)=1$ (הסיבה לכך היא שב- '∞' כל אחד מן המשתנים עבר את המיפוי האפשרי מ- Ω).
5. $\lim_{x_i\to-\infty} F_{\underline{X}}\left(\underline{x}\right)=0$ (הסיבה לכך היא שאין אף מקור מ- Ω שימפה את X_i לערך הקטן מ- '$-\infty$' , הרי $\left\{\omega : X_i\left(\omega\right)\le -\infty\right\}=\varnothing$)
6. פילוג שולי (ניתן לחשב את פונקציית ה-CDF של מ"א יחיד או תת-ווקטור):

$$F_{X_k}\left(x\right)=\lim_{x_i\to\infty \quad i\neq k} F_{\underline{X}}\left(x_1,x_2,\ldots,x_n\right)$$

$$F_{\underline{X}'}\left(\underline{x}'\right)=\lim_{x_i\to\infty \quad X_i\notin\underline{X}'} F_{\underline{X}}\left(x_1,x_2,\ldots,x_n\right)$$

הדבר מבוצע ע"י השאפת הארגומנטים אשר לא שייכים לפילוג השולי הרלוונטי ל- '∞' (לפי אותו עקרון של תכונה 3).

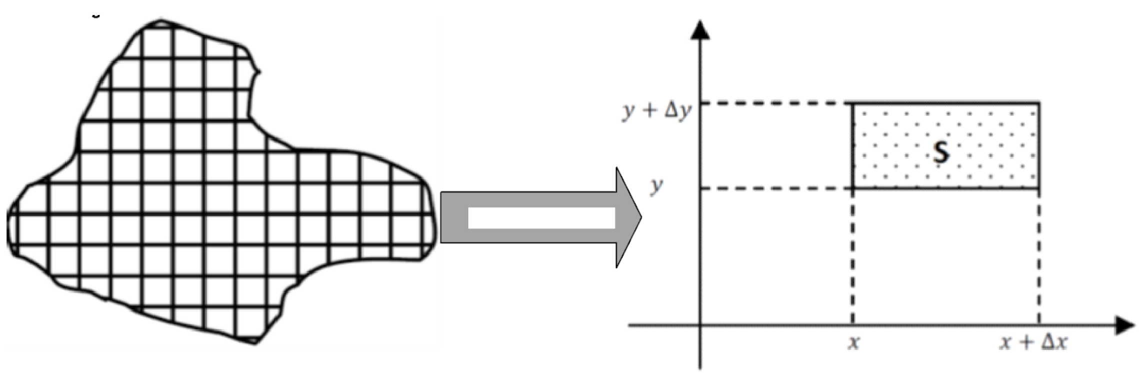
<u>הערה:</u>

הגדרת ה-jCDF מספיקה ע"מ לענות על כל שאלה מהצורה $\Pr\left(\underline{X}\in S\right)$.

<u>הסבר:</u>

נמחיש עבור דו-מימד (נעבוד עם מ"א המסומנים X ו- Y לשם נוחות).

נניח תחילה כי S הוא שטח אשר ניתן לחלק למלבנים. עבור S אשר אינו ניתן לחלוקה פשוטה למלבנים נוכל לחלקו למלבנים אינפיטיסימלים. כעת נרצה לענות מה ההסתברות "ליפול" במלבן בודד (קטן כרצוננו) ואז נוכל פשוט לסכום את ההסתברויות המתאימות לכל המלבנים.



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

הסיבה לכך היא למעשה ש- $F_{XY}(a,b)$ היא למעשה הסיכוי ש- (X,Y) יימצא בכל המלבן האינסופי אשר נמצא משמאל-מתחת לנקודה (a,b) ולכן אנו מחסרים ומחברים חלקי מלבנים אינסופיים אחד מהשני.

פונקציית צפיפות ההסתברות המשותפת – Joint Probability Density Function (jPDF)

עבור ו"א $\underline{X}$ פונקציית ה-jPDF שלו תוגדר להיות הפונקציה המקיימת:

$$F_{\underline{X}}(\underline{x})=\int\int_{-\underline{\infty}}^{\underline{x}}\cdots\int f_{\underline{X}}(\underline{x})\underline{dx}=\int_{x_1'=-\infty}^{x_1}\int_{x_2'=-\infty}^{x_2}\cdots\int_{x_n'=-\infty}^{x_n}f_{\underline{X}}(\underline{x})dx_1'dx_2'\cdot\ldots\cdot dx_n'$$

תכונות:

1. $\int\underset{\underline{X}\in\mathbb{R}^n}{\cdots}\int f_{\underline{X}}(\underline{x})\underline{dx}=1$
2. $\forall\underline{x}\quad f_{\underline{X}}(\underline{x})\geq 0$
3. פילוג שולי (ניתן לחשב את פונקציית ה-PDF של מ"א יחיד או תת-ווקטור):

$$f_{X_k}(x)=\int_{-\infty}^{\infty}\underset{\forall X\neq X_k}{\cdots}\int_{-\infty}^{\infty}f_{\underline{X}}(\underline{x})\underline{dx}$$

$$f_{\underline{X}'}(\underline{x}')=\int_{-\infty}^{\infty}\underset{\forall X_i\notin\underline{X}'}{\cdots}\int_{-\infty}^{\infty}f_{\underline{X}}(\underline{x})\underline{dx}$$

הדבר מבוצע ע"י אינטגרציה בתחום $(-\infty,\infty)$ על כל המ"א אשר לא שייכים לפילוג השולי הרלוונטי.

הוכחה עבור דו-מימד:

$$F_X(x)=F_{XY}(x,\infty)$$

$$\Rightarrow\begin{aligned}&f_X(x)=\frac{d}{dx}F_{XY}(x,\infty)=\frac{d}{dx}\Pr(-\infty\leq X\leq x,-\infty\leq Y\leq\infty)=\\&=\frac{d}{dx'}\int_{y'=-\infty}^{\infty}\int_{x'=-\infty}^{x}f_{XY}(x',y')dx'dy'=\int_{-\infty}^{\infty}f_{XY}(x,y)dy\end{aligned}$$

הערה:

עבור $F_{\underline{X}}(\underline{x})$ גזירה לכל $\underline{x}$ (גזירה זו דרישה חזקה יותר מרציפה) אז מתקיים $f_{\underline{X}}(\underline{x})=\frac{d^n}{dx_1\cdot\ldots dx_n}F_{\underline{X}}(\underline{x})$. כאשר אנו מניחים שהנגזרת מוגדרת היטב ואין חשיבות לסדר הגזירה.

<u>הסבר:</u>

עבור דו-מימד מתקיים (נוכל לראות מפיתוח ההסתברות עבור מלבן ב-jCDF):

$$f_{XY}(x,y)=\lim_{\Delta x\to 0,\Delta y\to 0}\frac{1}{\Delta x\Delta y}\Pr\left(x\le X\le x+\Delta x, y\le Y\le y+\Delta y\right)=\frac{d^2}{dxdy}F_{XY}(x,y)$$

<u>מסקנה:</u>

$$\Pr\left(\underline{X}\in S\right)=\int\underset{\underline{X}\in S}{\cdots}\int f_{\underline{X}}\left(\underline{x}\right)\underline{dx}$$

שני משתנים אקראיים (פילוג מותנה ופילוג שולי)

תחילה נרחיב את הדיון בנושא ו"א ע"י דיון על ו"א בגודל 2, ז"א שני משתנים אקראיים. הדיון על שני משתנים אקראיים תקף גם לו"א בגודל כללי, אנו מתחילים עם שני מ"א בלבד לצורך תהליך הלימוד ופשטות הפיתוח.

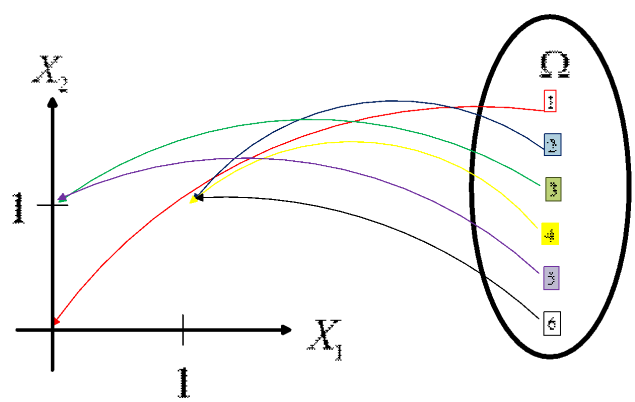
<u>דוגמא:</u>

עבור ניסוי הטלת קוביה הוגנת נגדיר שני משתנים:

1. $X_1(\omega)$ אינדיקטור זוגיות
2. $X_2(\omega)$ אינדיקטור 'האם יצא גדול מ-1'

$$X_1(\omega)=\begin{cases}0 & \omega=1,3,5\\ 1 & \omega=2,4,6\end{cases}\qquad X_2(\omega)=\begin{cases}1 & \omega>1\\ 0 & \omega=1\end{cases}$$

נראה את המיפוי באופן גרפי:



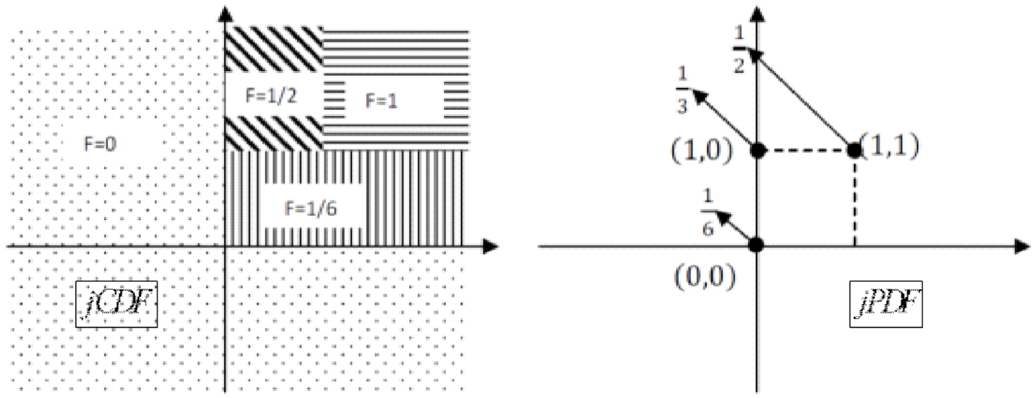
נוכל לחשב את ההסתברויות:

$$\Pr\left(X_1,X_2\right)=\begin{cases}1/6 & \left(X_1,X_2\right)=(0,0)\\ 1/3 & \left(X_1,X_2\right)=(0,1)\\ 0 & \left(X_1,X_2\right)=(1,0)\\ 1/2 & \left(X_1,X_2\right)=(1,1)\end{cases}$$

ניתן להבחין כי קיימת תלות בין המשתנים מאחר ו-

$$\Pr\left(X_1=\alpha,X_2=\beta\right)\neq\Pr\left(X_1=\alpha\right)\cdot\Pr\left(X_2=\beta\right).$$

פונקציות ה-jPDF וה-jCDF המתאימות לווקטור זה הן:



אנו רואים כי פונקציית ה-jPDF מורכבת מהלמים (מומחשים כחצים אך כיוונם מחוץ למישור הגרף) ופונקציית ה-jCDF מורכבת מקפיצות. הדבר נובע מכך שהמשתנים האקראיים אשר מרכיבים את הו"א בדידים.

$$F_{X_1X_2}\left(x_1,x_2\right)=\sum_{X_1=0}^{1}\sum_{X_2=0}^{1}\Pr\left(X_1,X_2\right)U\left(x_1-X_1\right)U\left(x_2-X_2\right)=$$

$$=\frac{1}{6}U\left(x_1-0\right)U\left(x_2-0\right)+\frac{1}{3}U\left(x_1-0\right)U\left(x_2-1\right)+\frac{1}{2}U\left(x_1-1\right)U\left(x_2-1\right)$$

$$\Rightarrow f_{X_1X_2}\left(x_1,x_2\right)=\frac{d^2}{dx_1dx_2}F_{X_1X_2}\left(x_1,x_2\right)=\frac{d}{dx_2}\left[\frac{d}{dx_1}F_{X_1X_2}\left(x_1,x_2\right)\right]=$$

$$=\frac{1}{6}\delta\left(x_1\right)\delta\left(x_2\right)+\frac{1}{3}\delta\left(x_1\right)\delta\left(x_2-1\right)+\frac{1}{2}\delta\left(x_1-1\right)\delta\left(x_2-1\right)$$

<u>פילוג מותנה</u>

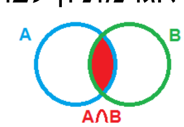
נרצה לדעת לתאר את ההסתברות לגביי מאורעות על אחד המשתנים בהינתן מידע מקדים על המשתנה האקראי הנוסף.

<u>הסתברות מותנית – תזכורת</u>

בהינתן מאורעות A,B כאשר ידוע $\Pr\left(B\right)\neq 0$ אז ההסתברות של מאורע A בהינתן קיום מאורע B היא:

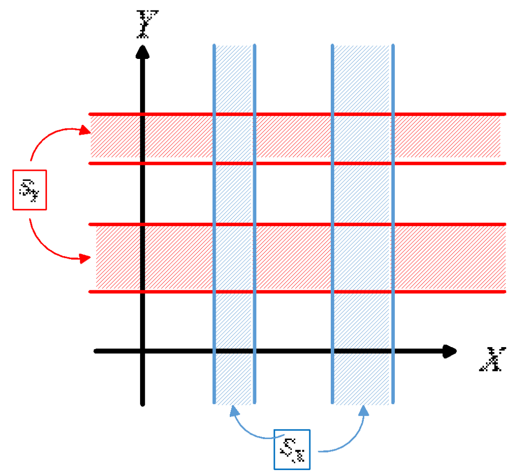
$$\Pr\left(A|B\right)=\frac{\Pr\left(A\cap B\right)}{\Pr\left(B\right)}$$

דיאגרמת ון לצורך המחשה:



הסתברות מותנית עבור שני משתנים אקראיים

לכן עבור ו"א (X,Y) נגדיר שני מאורעות $X \in S_X , Y \in S_Y$. ולצורך ההכללה נניח כי כ"א מהתחומים S_X , S_Y הוא תחום רציף למקוטעין.



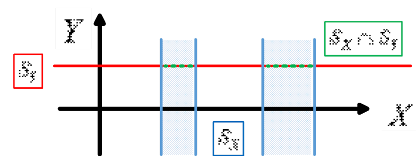
$$\Pr\left(X \in S_X \mid Y \in S_Y\right) = \frac{\Pr\left(X \in S_X \cap Y \in S_Y\right)}{\Pr\left(Y \in S_Y\right)} = \frac{\int\limits_{X \in S_X} \int\limits_{Y \in S_Y} f_{XY}(x,y)\,dxdy}{\int\limits_{Y \in S_Y} f_Y(y)\,dy} = \frac{\int\limits_{X \in S_X} \int\limits_{Y \in S_Y} f_{XY}(x,y)\,dxdy}{\int\limits_{X=-\infty}^{\infty} \int\limits_{Y \in S_Y} f_{XY}(x,y)\,dxdy}$$

התנייה נקודתית

ישנן הסתברויות אשר מתנות במאורע נקודתי מהצורה $Y = y$, ז"א אחד המשתנים ידוע בדיוק.

במקרה הרציף אנו בבעיה מאחר והסיכוי הנקודתי הינו אפס ($\Pr(Y=y)=0$) ולכן לא נוכל להתשמש בנוסחא להסתברות מותנית מאחר שהחלוקה באפס אינה מוגדרת.

$$\Pr\left(X \in S_X \mid Y = y\right) = \frac{\Pr\left(X \in S_X \cap Y = y\right)}{\Pr\left(Y = y\right)} = "\left(\frac{0}{0}\right)"$$



על מנת להתמודד עם הבעיה נגדיר:

$$\Pr\left(X \in S_X \mid Y = y\right) = \lim_{\Delta y \to 0} \Pr\left(X \in S_X \mid Y = \in \left(y, y+\Delta y\right)\right) \overset{S_Y = (y, y+\Delta y)}{=} \lim_{\Delta y \to 0} \Pr\left(X \in S_X \mid Y \in S_Y\right) =$$

$$= \lim_{\Delta y \to 0} \frac{\Pr\left(X \in S_X \cap Y \in S_Y\right)}{\Pr\left(Y \in S_Y\right)} = \lim_{\Delta y \to 0} \frac{\int\limits_{X \in S_X} \int\limits_{Y \in S_Y} f_{XY}(x,y)\,dxdy}{\int\limits_{Y \in S_Y} f_Y(y)\,dy} \cdot \frac{\frac{1}{\Delta y}}{\frac{1}{\Delta y}} = \lim_{\Delta y \to 0} \frac{\frac{1}{\Delta y} \int\limits_{v=y}^{y+\Delta y} \int\limits_{X \in S_X} f_{XY}(u,v)\,dudv}{\frac{1}{\Delta y} \int\limits_{v=y}^{y+\Delta y} f_Y(v)\,dv} =$$

$$\overset{*}{=} \frac{\lim\limits_{\Delta y \to 0} \left[\frac{1}{\Delta y} \int\limits_{v=y}^{y+\Delta y} \left(\int\limits_{X \in S_X} f_{XY}(u,v)\,du \right) dv \right]}{\lim\limits_{\Delta y \to 0} \left[\frac{1}{\Delta y} \int\limits_{v=y}^{y+\Delta y} f_Y(v)\,dv \right]} \underset{\text{L'H\^{o}pital's}}{\overset{**}{=}} \frac{\int\limits_{X \in S_X} f_{XY}(x,y)\,dx}{f_Y(y)} = \int\limits_{X \in S_X} \frac{f_{XY}(x,y)\,dx}{f_Y(y)} = \int\limits_{X \in S_X} f_{X|Y}(x \mid y)\,dx$$

כאשר במעבר $(*)$ הגבול הופרד למונה ולמכנה בהנחה שהוא קיים והוחלף סדר האינטגרציה.

במעבר $(**)$ השתמשנו בהגדרת הנגזרת ובמשמעות כלל לופיטל יחד עם רציפות הפונקציה בקטע האינטגרל (ולכן יכלנו להחליף את האינטגרל על התחום בערך הפונקציה כפול רוחב התחום – אינטגרל רימן/לבג).

<u>פונקציות פילוג (CDF/PDF) מותנה</u>

נגדיר את פונקציות הפילוג המותנה, ז"א נגדיר את הפונקציות בעזרתן נענה על השאלות מסוג $\Pr\left(X \in S_X \mid Y = y\right)$.

<u>CDF מותנה</u>

$$F_{X|Y}\left(x \mid y\right) \overset{*}{=} F_{X|Y}\left(x, y\right) \triangleq \Pr\left(X \in \left(-\infty, x\right) \mid Y = y\right) \overset{S_X = (-\infty, x)}{=} \int_{X \in S_X} \frac{f_{XY}\left(x, y\right) dx}{f_Y\left(y\right)}$$

נבחין כי בשוויון $(*)$ הפונקציה מוצגת בשני אופנים שונים. שני התצוגות זהות מאחר וזוהי פונקציה פשוטה של שני משתנים והרישום יכול להתשנות לצורכי נוחות, אינדקס הפונקציה ($X \mid Y$) הוא החשוב (כנ"ל לגביי PDF מותנה).

<u>PDF מותנה</u>

$$f_{X|Y}\left(x \mid y\right) = f_{X|Y}\left(x, y\right) \triangleq \frac{d}{dx} F_{X|Y}\left(x, y\right) = \frac{f_{XY}\left(x, y\right)}{f_Y\left(y\right)}$$

<u>הערה – חוק בייס:</u>

ע"י החלפת תפקידים בין המשתנים נוכל לקבל את חוק בייס (Bayes' Theorem) עבור PDF מותנה:

$$f_{XY}\left(x, y\right) = f_X\left(x\right) \cdot f_{Y|X}\left(y \mid x\right) = f_Y\left(y\right) \cdot f_{X|Y}\left(x \mid y\right)$$

$$f_{Y|X}\left(y \mid x\right) = \frac{f_Y\left(y\right) \cdot f_{X|Y}\left(x \mid y\right)}{f_X\left(x\right)} = \frac{f_Y\left(y\right) \cdot f_{X|Y}\left(x \mid y\right)}{\int_{-\infty}^{\infty} f_{XY}\left(x, y\right) dy} = \frac{f_Y\left(y\right) \cdot f_{X|Y}\left(x \mid y\right)}{\int_{-\infty}^{\infty} f_Y\left(y\right) f_{X|Y}\left(x \mid y\right) dy}$$

החוק תקף גם במקרים של בדיד ומעורב, עבור משתנה בדיד יהיה $\Pr$ ולא f.

<u>הערה:</u>

פונקציות ה-PDF (כנ"ל לגביי CDF) היא פונקציית PDF לכל דבר ועניין. ניתן לחשוב עליה כעל פונקציית PDF אשר מתארת את המשתנה האקראי $'X \mid Y = y'$.

<u>תוחלת מותנה ושונות מותנה</u>

נוכל להגדיר תוחלת מותנה, שונות מותנה וכנ"ל לגבי כל מומנט שנרצה.

$$\eta_{X|Y=y} = E\left[X \mid Y = y\right] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}\left(x, y\right) dx$$

$$Var\left(X \mid Y = y\right) = \int_{-\infty}^{\infty} \left(x - \eta_{X|Y=y}\right)^2 \cdot f_{X|Y}\left(x, y\right) dx$$

אי-תלות סטטיסטית

נאמר ש- X ו- Y משתנים אקראיים בלתי-תלויים סטטיסטית (או נאמר ש- (X,Y) הינו ו"א בת"ס) אם:

$$\Pr\left(X \in S_X, Y \in S_Y\right) = \Pr\left(X \in S_X\right) \cdot \Pr\left(Y \in S_Y\right) \quad \forall S_X, S_Y \in \mathbb{R}$$

<u>הערה:</u>

הטענות הבאות שקולות:

1. (X,Y) הינו ו"א בת"ס.
2. $F_{XY}\left(x, y\right) = F_X\left(x\right) \cdot F_Y\left(y\right)$
3. $f_{XY}\left(x, y\right) = f_X\left(x\right) \cdot f_Y\left(y\right)$
4. $f_{X|Y}\left(x \mid y\right) = f_X\left(x\right), \quad f_{Y|X}\left(y \mid x\right) = f_Y\left(y\right)$
5. $F_{X|Y}\left(x \mid y\right) = F_X\left(x\right), \quad F_{Y|X}\left(y \mid x\right) = F_Y\left(y\right)$

<u>הוכחה:</u>

$$1 \to 2 \quad F_{XY}\left(x, y\right) \triangleq \Pr\left(X \in \left(-\infty, x\right), Y \in \left(-\infty, y\right)\right) \overset{*}{=} \Pr\left(X \in \left(-\infty, x\right)\right) \cdot \Pr\left(Y \in \left(-\infty, y\right)\right) = F_X\left(x\right) \cdot F_Y\left(y\right)$$

$$2 \to 3 \quad f_{XY}\left(x, y\right) = \frac{d^2}{dxdy} F_{XY}\left(x, y\right) \overset{**}{=} \frac{d^2}{dxdy}\left[F_X\left(x\right) \cdot F_Y\left(y\right)\right] = \frac{d}{dx} F_X\left(x\right) \cdot \frac{d}{dy} F_Y\left(y\right) = f_X\left(x\right) \cdot f_Y\left(y\right)$$

$$3 \to 1 \quad \Pr\left(X \in S_X, Y \in S_Y\right) = \int_{X \in S_X} \int_{Y \in S_Y} f_{XY}\left(x, y\right) dxdy \overset{***}{=} \int_{X \in S_X} \int_{Y \in S_Y} f_X\left(x\right) \cdot f_Y\left(y\right) dxdy =$$

$$= \int_{X \in S_X} f_X\left(x\right) dx \int_{Y \in S_Y} f_Y\left(y\right) dy = \Pr\left(X \in S_X\right) \cdot \Pr\left(Y \in S_Y\right)$$

$$4 \qquad f_{X|Y}\left(x \mid y\right) = \frac{f_{XY}\left(x, y\right)}{f_Y\left(y\right)} \overset{***}{=} \frac{f_X\left(x\right) \cdot f_Y\left(y\right)}{f_Y\left(y\right)} = f_X\left(x\right)$$

$$5 \qquad F_{X|Y}\left(x \mid y\right) = \int_{-\infty}^{x} f_{X|Y}\left(x \mid y\right) dx \overset{****}{=} \int_{-\infty}^{x} f_X\left(x\right) dx = F_X\left(x\right)$$

מעבר $(*)$ עקב חוסר-תלות סטטיסטית. מעבר $(**)$ עקב טענה 2. מעבר $(***)$ עקב טענה 3.

מעבר $(****)$ עקב טענה 4.

<u>הרחבה:</u>

עבור ו"א בת"ס כללי $\underline{X}$ מתקיים:

$$\Pr\left(X_1 \in S_{X_1},\ldots,X_2 \in S_{X_2}\right)=\prod_{i=1}^{n}\Pr\left(X_i \in S_{X_i}\right)$$

מכאן ההכללה מיידית לגביי פונקציות ה-jCDF וה-jPDF (מתפרקות למכפלה).

<u>הערה:</u>

ידיעת פונקציית ה-jPDF של ו"א נותנת גם את ה-PDF-ים של כל המ"א ממנו מורכב הווקטור (או את ה-jPDF של כל תת-ווקטור) ע"י אינטגרציה בתחום $(-\infty,\infty)$ על שאר המשתנים האקראיים.

ההיפך לא נכון אלא אם ידוע כי המשתנים האקראיים בת"ס ואז $f_X(x)$ ו- $f_Y(y)$ נותנים את $f_{XY}(x,y)$ ע"י קשר פשוט של מכפלה, ז"א $f_{XY}(x,y)=f_X(x)\cdot f_Y(y)$.

<u>דוגמא 1:</u>

אי-תלות בזוגות לא גוררת אי-תלות מלאה (לא גוררת בת"ס).

נתון ו"א (X,Y,Z) בעל רכיבים בת"ס בזוגות, ז"א ידוע כי $(X,Y),(X,Z),(Y,Z)$ בת"ס.

ידוע כי:

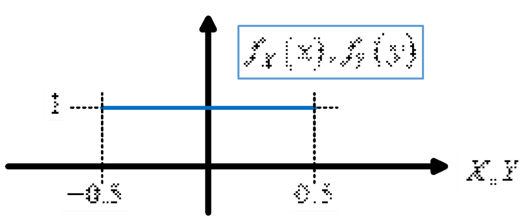
$$X=Y=\begin{cases}1 & w.p.\tfrac{1}{2}\\ 0 & w.p.\tfrac{1}{2}\end{cases}\qquad Z=X\oplus Y=\begin{cases}1 & w.p.\tfrac{1}{2}\{(0,1),(1,0)\}\\ 0 & w.p.\tfrac{1}{2}\{(0,0),(1,1)\}\end{cases}$$

אך הו"א אינו בת"ס מאחר ו-

$$\Pr(X=1,Y=1,Z=1)\neq\Pr(X=1)\cdot\Pr(Y=1)\cdot\Pr(Z=1)$$
$$\Pr(X=1,Y=1,Z=1)=0$$
$$\Pr(X=1)\cdot\Pr(Y=1)\cdot\Pr(Z=1)=\tfrac{1}{8}$$

<u>דוגמא 2:</u>

נתון ו"א (X,Y) כאשר $X,Y\sim Unif(-0.5,0.5)$.

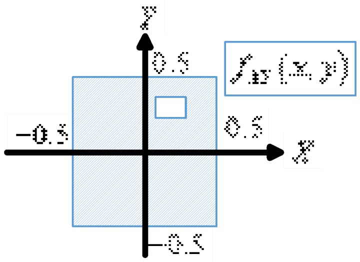


$$f_X(x)=f_Y(y)=\begin{cases}1 & -0.5<y<0.5\\ 0 & else\end{cases}$$

נרצה להראות איך התלות בין שני המשתנים האקראיים משפיעה ע"י שלושה מקרים:

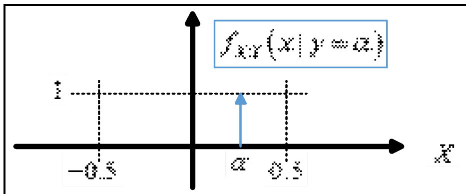
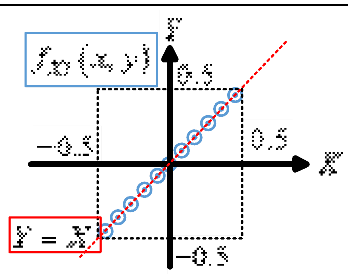
<u>מקרה I –</u>

אם ידוע X,Y בת"ס אז:

$f_{XY}(x,y)=f_X(x)\cdot f_Y(y)=\begin{cases}1 & \begin{cases}-0.5<x<0.5\\-0.5<y<0.5\end{cases}\\0 & else\end{cases}$	

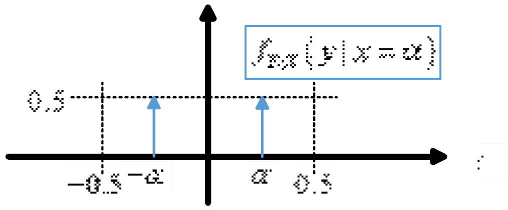
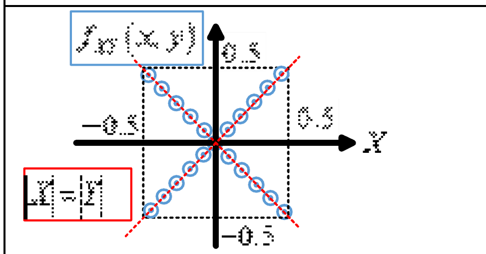
<u>מקרה II –</u>

אם נתון $X=Y$, ז"א הם למעשה אותו מ"א אז:

כמובן שמתקיים (משום שזה אותו מ"א): $f_{X\mid Y}(x\mid y)=\begin{cases}1 & x=y,\lvert y\rvert<0.5\\0 & else\end{cases}=\delta(x-y),\lvert y\rvert<0.5$ הערה - כאשר y ידוע ובעל ערך מתאים לפילוג שלו.	
ולכן נקבל	$f_{XY}(x,y)=f_Y(y)\cdot f_{X\mid Y}(x\mid y)=\begin{cases}1\cdot\delta(x-y) & \begin{cases}-0.5<x<0.5\\-0.5<y<0.5\end{cases}\\0 & else\end{cases}$
שזהו למעשה "קיר-דלתאות" על הישר $X=Y$ בתחום המתאים לפילוג המ"א ($\lvert X\rvert,\lvert Y\rvert<0.5$).	

<u>מקרה III –</u>

נתון כי $Y=BX$ כאשר $B=\pm1$ בהסתברויות שוות ו- B,X בת"ס. ידוע $X\sim Unif(-0.5,0.5)$ ולכן המ"א $'-X'$ מתפלג באופן זהה. מכאן שעבור כל ערך של B ל- Y ול- X פילוג זהה, ז"א $X,Y\sim Unif(-0.5,0.5)$.	$B=\begin{cases}1 & w.p.\tfrac{1}{2}\\-1 & w.p.\tfrac{1}{2}\end{cases}$

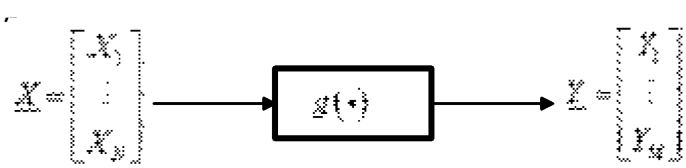
מתקיים $f_{Y\|X}(y\|x)=\begin{cases}\frac{1}{2}\delta(y-x)+\frac{1}{2}\delta(y+x) & \|x\|<0.5\\ 0 & else\end{cases}$	
ולכן נקבל	$f_{XY}(x,y)=f_X(x)\cdot f_{Y\|X}(y\|x)=\begin{cases}1\cdot\left[\frac{1}{2}\delta(y-x)+\frac{1}{2}\delta(y+x)\right] & \begin{cases}-0.5<x<0.5\\-0.5<y<0.5\end{cases}\\ 0 & else\end{cases}$
שאלו למעשה שני "קירות-דלתאות" על הישרים המקיימים את השיווין $\|X\|=\|Y\|$ בתחום המתאים לפילוג המ"א ($\|X\|,\|Y\|<0.5$).	

<u>וקטור של משתנים אקראיים</u>

כעת נרחיב את הדיון לווקטור כללי (בגודל כלשהו) של משתנים אקראיים, אך עדיין נדגים לעיתים עבור שני מ"א.

<u>פונקציה של וקטור אקראי</u>

ווקטור אקראי $\underline{X}$ עובר טרנספורמציה כלשהי $g(\cdot)$, מוצא הטרנספורמציה הינו ו"א $\underline{Y}$ אשר לא בהכרח באותו בגודל/מימד. בידיעת נתוני ההסתברות $F_{\underline{X}}(\underline{x})$ ו- $f_{\underline{X}}(\underline{x})$ (ז"א ידועות פונקציות ה-jCDF וה-jPDF של הכניסה) ואיפיון המערכת $g(\cdot)$ נרצה לדעת את הסטטיסטיקה של מוצא הטרנספורמציה $\underline{Y}=g(\underline{X})$, כלומר את $f_{\underline{Y}}(\underline{y}),F_{\underline{Y}}(\underline{y})$.



למעשה קיבלנו פונקציה מורכבת:

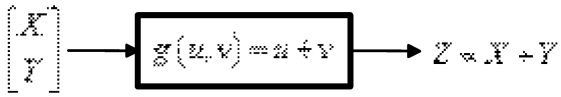
$$\underline{Y}=g(\underline{X})=g(\underline{X}(\omega)) \quad\Rightarrow\quad \underline{Y}(\omega)=g(\underline{X}(\omega))$$

פונקצייה מורכבת של ו"א הינה עדיין ו"א מאחר והיא עדיין מיפוי מהצורה $\Omega\to\mathbb{R}^k$.

נרצה לענות על $\Pr(\underline{Y}\in S)$ לכל $S\in\mathbb{R}^M$.

<u>דוגמא:</u>

עבור המערכת הבאה (כאשר הפילוג המשותף של הכניסה ידוע ולכן גם הפילוג של כ"א מן הרכיבים בנפרד)



נרצה למצוא את פילוג המוצא, כלומר את $F_Z(z)$.

$$F_Z(z)=\Pr(Z\le z)=\Pr(g(X,Y)\le z)\overset{*}{=}\Pr\left((X,Y)\in g^{-1}((-\infty,z])\right)=\iint\limits_{(X,Y)\in g^{-1}(-\infty,z]} f_{XY}(x,y)\,dxdy$$

כאשר בשוויון $(*)$ הגדרנו את הפונקצייה ההופכית להיות:

$$g^{-1}((-\infty,z])=\{(x,y):g(x,y)\in(-\infty,z]\}$$

עד כה הפיתוח היה כללי עבור $g:\mathbb{R}^2\to\mathbb{R}$ כלשהי, אצלנו מתקיים $g(X,Y)=X+Y$ ולכן:

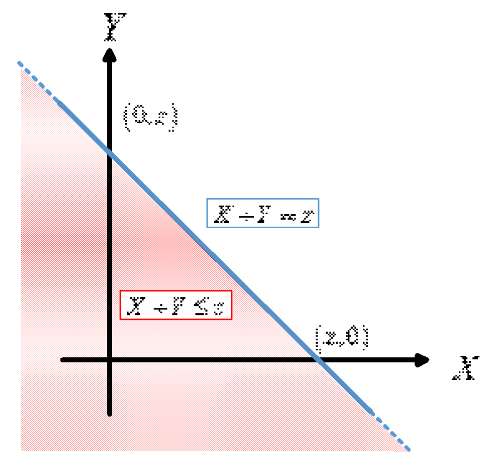
$$F_Z(z)=\Pr(Z\le z)=\Pr(X+Y\le z)=$$

$$=\int\limits_{x=-\infty}^{\infty}\left(\int\limits_{y=-\infty}^{z-x} f_{XY}(x,y)\,dy\right)dx$$

$$\Rightarrow f_Z(z)=\frac{d}{dz}F_Z(z)\overset{*}{=}\int\limits_{x=-\infty}^{\infty} f_{XY}(x,z-x)\,dx$$

$(*)$ לפי המשפט היסודי של האלגברה:

$$\frac{d}{dt}\int_a^b f(t)\,dt=f(b)$$



אם נוסיף כי (X,Y) ו"א בת"ס אז $f_{XY}(x,y)=f_X(x)\cdot f_Y(y)$ ולכן:

$$f_Z(z)=\int\limits_{x=-\infty}^{\infty} f_{XY}(x,z-x)\,dx=\int\limits_{x=-\infty}^{\infty} f_X(x)\cdot f_Y(z-x)\,dx\overset{*}{=}f_X(x)*f_Y(y)$$

כאשר שוויון $(*)$ מתקיים לפי הגדרת הקונבולוציה (והגבול האינסופי לכן נוכל להחליף משתנים בקלות):

$$f(x)*g(x)=\int\limits_{t=-\infty}^{\infty} f(t-x)\cdot g(t)\,dt=\int\limits_{t=-\infty}^{\infty} g(t-x)\cdot f(t)\,dt=g(x)*f(x)$$

<u>מסקנה:</u>

עבור סדרת משתנים אקראיים $\{X_i\}_{i=1}^{N}$ בת"ס, נגדיר מ"א $Z = X_1 + \ldots + X_N = \sum_{i=1}^{N} X_i$ ועבורו מתקיים:

$$f_Z(z) = f_{X_1} * \ldots * f_{X_N}(z)$$

<u>הערה:</u>

אם סדרת משתנים אקראיים $\{X_i\}_{i=1}^{N}$ היא i.i.d נגדיר מ"א Z באותו אופן (סכום שלהם).

פונקציית ה- $f_Z(z)$ היא למעשה קונבולוציה N-עצמית של אותו $f_X(x)$ (כי הם i.i.d).

כאשר $N \to \infty$ נקבל גאוסיאן עבור ללא תלות בפילוג ההתחלתי.

<u>הסבר:</u>

הדבר נובע ממשפט הגבול המרכזי. עבור כל פילוג התחלתי נגיע בסוף לפילוג עם Variance ממורכז ולכן זהו גאוסיאן.

ניתן להמחיש את הדבר אם נתחיל עם פילוג אחיד; עבור מ"א בודד נקבל מלבן, עבור שניים נקבל משולש (קונבולוציה של שני מלבנים) ובגבול נקבל גאוסיאן.

<u>תוחלת של ווקטור אקראי ותוחלת של פונקציה של ווקטור אקראי</u>

נניח קיים ו"א (X,Y) עם $f_{XY}(x,y)$ (או במקרה ה-n מימדי ו"א $\underline{X}$ עם $f_{\underline{X}}(\underline{x})$).

כמו כן נניח טרנספורמציה $g(X,Y)$ (או בהתאמה טרנספורמציה $g(\underline{X})$)

1. תכונת ההחלקה / משפט התוחלת השלמה
נחשב את התוחלת של מ"א כלשהו מתוך הווקטור ע"י:

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

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

ז"א נבצע את התוחלת של Y תחילה ע"י התנייה לפי X (החזקתו בערך קבוע מסויים) ואז נבצע תוחלת לפי כל ערכי X האפשריים (כך בתוחלת הפנימית "נפטרנו" מן המשתנה האקראי Y).

2. משפט ההחלקה / משפט התוחלת השלמה - תוחלת של פונקציה של ווקטור אקראי
נחשב את התוחלת של פונקציה של ו"א ע"י:

$$E\left[g(\underline{X})\right] = \int \underset{\forall \underline{x}}{\cdots} \int g(\underline{X}) f_{\underline{X}}(\underline{x}) \underline{dx}$$

עבור המקרה הדו-מימדי:

$$E\left[g\left(X,Y\right)\right]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g\left(X,Y\right)f_{XY}\left(x,y\right)dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g\left(X,Y\right)f_{X}\left(x\right)f_{Y|X}\left(y\mid x\right)dxdy=$$
$$=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}g\left(X,Y\right)f_{Y|X}\left(y\mid x\right)dy\right)f_{X}\left(x\right)dx=\int_{-\infty}^{\infty}\left[E_{Y}\left[g\left(X,Y\right)\mid X\right]\right]f_{X}\left(x\right)dx=$$
$$=E_{X}\left[E_{Y}\left[g\left(X,Y\right)\mid X\right]\right]\overset{Symmetry}{=}E_{Y}\left[E_{X}\left[g\left(X,Y\right)\mid Y\right]\right]$$

<u>הערות:</u>

1. תכונת ההחלקה נובעת ממשפט ההחלקה כאשר $g\left(X,Y\right)=Y$, ז"א טרנספורמציה אשר מתעלמת מן המשתנה האקראי X.
2. נוכל להכליל את הנוסחאות הנ"ל למקרה n-מימדי באותו האופן, לדוגמא:

$$E\left[Z\right]=E_{X}\left[E_{Y}\left[E_{Z}\left[Z\mid X,Y\right]\right]\right]=E_{X}\left[E_{Y}\left[E_{Z}\left[Z\mid X,Y\right]\right]\right]$$

<u>מומנטים משותפים של ווקטור אקראי</u>

עבור ו"א כללי $\underline{X}$ ממימד N נגדיר את סדר המומנט k ע"י $k=\sum_{i=1}^{N}k_i$.

המומנט מסדר k מוגדר להיות:

$$m_{k_1,\dots,k_N}=E\left[X_1^{k_1}\cdot\ldots\cdot X_N^{k_N}\right]=E\left[\prod_{i=1}^{N}X_i^{k_i}\right]$$

המומנט המרכזי מסדר k מוגדר להיות:

$$\mu_{k_1,\dots,k_N}=E\left[\left(X_1-\eta_{X_1}\right)^{k_1}\cdot\ldots\cdot\left(X_N-\eta_{X_N}\right)^{k_N}\right]=E\left[\prod_{i=1}^{N}\left(X_i-\eta_{X_i}\right)^{k_i}\right]$$

עבור ו"א דו-מימדי $\left(X,Y\right)$ נקבל:

$$m_{nk}=E\left[X^n\cdot Y^k\right],\qquad \mu_{nk}=E\left[\left(X-\eta_X\right)^n\cdot\left(Y-\eta_Y\right)^k\right]$$

נציג את המומנטים (הלא טריוויאליים) עד סדר 2:

$$m_{1,0}=\eta_X\qquad m_{0,1}=\eta_Y\qquad m_{1,1}=E\left[XY\right]=R_{XY}$$
$$\mu_{2,0}=\sigma_X{}^2=Var\left(X\right)\quad \mu_{0,2}=\sigma_Y{}^2=Var\left(X\right)\quad \mu_{1,1}=\sigma_{XY}=Cov\left(X,Y\right)$$

<u>הערה:</u>
סטטיסטיקה משותפת מסדר 2 למעשה נותנת את כל התוחלות והשונויות (ואריאנסים - Var) של המ"א כולל את כל הקרוס-ואריאנסים (Co-Var) בין זוגות מ"א.

<u>שונות משותפת – Covariance</u>

השונות המשותפת של שני מ"א X ו- Y מוגדרת להיות (פיתוח בדומה לאיך שבוצע עבור שונות רגילה):

$$Cov(X,Y)=\sigma_{XY}=\mu_{1,1}=E\left[(X-\eta_X)\cdot(Y-\eta_Y)\right]=E[XY]-E[X]E[Y]$$

<u>חוסר קורולציה</u>

נאמר ששני מ"א X ו- Y הם חסרי-קורולציה (חס"ק) או בלתי-מתואמים אם $Cov(X,Y)=0$.
חוסר קורולציה הינו "החלשה" של אי-תלות סטטיסטית.

<u>טענה:</u>

אם שני מ"א X ו- Y הם בת"ס אז הם גם חס"ק.

<u>הוכחה:</u>

$$Cov(X,Y)=E[XY]-E[X]E[Y]\overset{\text{Independence}}{=}E[X]E[Y]-E[X]E[Y]=0$$

<u>טענה:</u>

אם שני מ"א X ו- Y הם חס"ק אז מתקיים $Var(X+Y)=Var(X)+Var(Y)$.

<u>הוכחה:</u>

ניזכר כי:

$$Var(X)=E\left[(X-\eta_X)^2\right]=E\left[X^2\right]-E^2[X]=E\left[X^2\right]-\eta_X{}^2$$

ולכן:

$$Var(X+Y)=E\left[\left((X-\eta_X)+(Y-\eta_Y)\right)^2\right]\overset{*}{=}E\left[(X-\eta_X)^2\right]+2\cdot E\left[(X-\eta_X)(Y-\eta_Y)\right]+$$
$$+E\left[(Y-\eta_Y)^2\right]=Var(X)+2\cdot Cov(X,Y)+Var(Y)\overset{**}{=}Var(X)+Var(Y)$$

מעבר $(*)$ לפי ליניאריות התוחלת. מעבר $(**)$ מאחר ש- X ו- Y חס"ק ומתקיים $Cov(X,Y)=0$.

<u>אורתוגונליות</u>

נאמר ששני מ"א X ו- Y הם אורתוגונליים אם $R_{XY}=E[XY]=m_{1,1}=0$.

<u>הערה:</u>

באופן כללי אי-תלות סטטיסטית (בת"ס) לא בהכרח גורר אורתוגונליות (אך כפי שראינו כן גורר חס"ק).

<u>טענות (חס"ק ואורתוגנליות):</u>

1. אם זוג משתנים אקראיים X,Y הינם אורתוגונלים ובנוסף ידוע כי $E[X]=\eta_X=0$ או $E[Y]=\eta_Y=0$ (או שניהם), אזי הם חס"ק (הרי נקבל $Cov(X,Y)=E[XY]-E[X]\cdot E[Y]=0-0=0$).
2. אם זוג משתנים אקראיים X,Y הינם חס"ק ובנוסף ידוע כי $E[X]=\eta_X=0$ או $E[Y]=\eta_Y=0$ (או שניהם), אזי הם אורתוגנלים (הרי נקבל $E[XY]=Cov(X,Y)+E[X]\cdot E[Y]=0+0=0$).

<u>מסקנה:</u>

אם $E[X]=\eta_X=0$ או $E[Y]=\eta_Y=0$, אזי אורתוגנליות $\Leftrightarrow$ חס"ק.

<u>מקדמי מתאם ליניארים (מקדם פירסון ומקדם קורולציה)</u>

בהינתן שני משתנים אקראיים נרצה לבטא את הקשר ביניהם.

לא פעם נחפש מתאם בין שני רכיבים בתוך אותו ווקטור-אקראי (נאמר בין X_l לבין X_k).
אנו נתעסק עם מקדמי מתאם ליניארים בלבד, ז"א מדדים אשר מבטאים את יכולת קירוב התלות של מ"א מסויים (נאמר X) במ"א אחר (נאמר Y) ע"י תלות ליניארית ללא מקדם חופשי (עוברת בראשית), ז"א ע"י $Y=\alpha X$.

<u>מקדם המתאם של פירסון</u>

מקדם המתאם של פירסון (מתאם פרסון) הוא מדד למתאם ליניארי בין שתי קבוצות של מספרים ונתון ע"י:

$$r(X,Y)=\frac{E[X\cdot Y]}{\sqrt{E[X^2]\cdot E[Y^2]}}$$

נקבל $|r(X,Y)|=1$ אם"ם קיימת α כך שמתקיים $Y=\alpha X$. ז"א קיים ישר העובר דרך ראשית הצירים אשר מתאר את הקשר בין X ל- Y.

<u>מקדם קורולציה</u>

מקדם קורולציה גם הוא מדד למתאם ליניארי והוא מוגדר מהגדרת מקדם פירסון ע"י:

$$\rho(X,Y)=r(X-\eta_X,Y-\eta_Y)=\frac{E[(X-\eta_X)\cdot(Y-\eta_Y)]}{\sqrt{E[(X-\eta_X)^2]\cdot E[(Y-\eta_Y)^2]}}=\frac{Cov(X,Y)}{\sigma_X\cdot\sigma_Y}$$

נוכל גם לרשום באופן יותר מפורש:

$$\rho\left(X,Y\right)=\frac{E\left[XY\right]-E\left[X\right]\cdot E\left[Y\right]}{\sqrt{E\left[X^2\right]-E^2\left[X\right]}\cdot\sqrt{E\left[Y^2\right]-E^2\left[Y\right]}}$$

נקבל $\left|\rho\left(X,Y\right)\right|=1$ אם"ם קיימת α כך שמתקיים $\left(Y-\eta_Y\right)=\alpha\left(X-\eta_X\right)$. ז"א קיים ישר כלשהו אשר מתאר את הקשר בין X ל- Y.

<u>טענות:</u>

1. זוג משתנים אקראיים X,Y הינם אורותוגונלים אמ"מ $r\left(X,Y\right)=0$ (הרי אז מתקיים $E\left[X\cdot Y\right]=0$).
2. זוג משתנים אקראיים X,Y הינם חס"ק אמ"מ $\rho\left(X,Y\right)=0$ (הרי אז מתקיים $Cov\left(X,Y\right)=0$).

<u>טענה:</u>

עבור כל שני משתנים אקראיים X,Y מתקיים $-1\le r\left(X,Y\right)\le 1$ (ז"א $\left|r\left(X,Y\right)\right|\le 1$).

<u>הוכחה:</u>

נתבונן במטריצה:

$$\begin{bmatrix} E\left[X^2\right] & E\left[X\cdot Y\right] \\ E\left[Y\cdot X\right] & E\left[Y^2\right] \end{bmatrix}$$

זו מטריצה PSD (Positive-Semi-Definite) – נלמד באלגברה ליניארית, תבוצע חזרה בהמשך הקורס.

לכן כל הערכים העצמיים שלה אי-שליליים, ז"א λ_1,λ_2 אי-שליליים. מכאן שגם הדטרמיננטה של המטריצה אי-שלילית ולכן:

$$\det\left(\begin{bmatrix} E\left[X^2\right] & E\left[X\cdot Y\right] \\ E\left[Y\cdot X\right] & E\left[Y^2\right] \end{bmatrix}\right)=E\left[X^2\right]\cdot E\left[Y^2\right]-E^2\left[X\cdot Y\right]\ge 0$$

נעביר אגפים ונקבל את אי-שיוויון קושי-שוורץ (Cauchy-Schwarz Inequality):

$$E^2\left[X\cdot Y\right]\le E\left[X^2\right]\cdot E\left[Y^2\right]$$

מחלוקה באגף ימין ומהגדרת מקדם פירסון נקבל:

$$r^2\left(X,Y\right)=\frac{E^2\left[X\cdot Y\right]}{E\left[X^2\right]\cdot E\left[Y^2\right]}\le 1 \quad\Rightarrow\quad \left|r\left(X,Y\right)\right|\le 1$$

<u>טענה:</u>

עבור כל שני משתנים אקראיים X,Y מתקיים $-1\le\rho(X,Y)\le 1$ (ז"א $|\rho(X,Y)|\le 1$).

<u>הוכחה:</u>

הוכחנו $|r(X,Y)|\le 1$ עבור כל שני משתנים אקראיים X,Y . נגדיר זוג משתנים:

$$\begin{cases} X' = X-\eta_X \\ Y' = Y-\eta_Y \end{cases}$$

וכמובן שמתקיים $|r(X',Y')|\le 1$ ומכאן $|\rho(X,Y)|\le 1$.

<u>פונקציה אופיינית משותפת</u>

פונקציה אופיינית משותפת של ו"א $\underline{X}$ היא התמרת פורייה של ה-PDF של $\underline{X}$ (עד כדי מינוס בחזקה).

$$\Phi_{\underline{X}}(\underline{\omega}) = \int\underset{\underline{x}\in\mathbb{R}^N}{\cdots}\int e^{j\underline{\omega}^T\underline{x}} f_{\underline{X}}(\underline{x})\,d\underline{x} = E\left[e^{j\underline{\omega}^T\underline{X}}\right]$$

<u>הערות:</u>

- $\underline{\omega}$ ו- $\underline{X}$ הינם מאותו מימד, את המכפלה ביניהם ניתן לרשום כסכום
$$\underline{\omega}^T\cdot\underline{X} = \sum_{i=1}^N \omega_i\cdot X_i$$.
- פ"א משותפת הינה תיאור אלטרנטיבי של מידע סטטיסטי מלא של הו"א.
- את הארגומנט של פ"א משותפת נרשום כווקטור שורה או עמודה לפי הנוחות.

<u>תכונות:</u>

1. $\Phi_{\underline{X}}(\underline{0}) = 1$
2. $|\Phi_{\underline{X}}(\underline{\omega})|\le 1$
3. ניתן לחשב פ"א שולית מתוך הפונקציה האופיינית המשותפת ע"י הצבת $\omega_i = 0$ עבור כל אינדקס i אשר אינו רצוי. לדוגמא:

$$\Phi_{X_k}(\omega_k) = \Phi_{\underline{X}}\left(0,\ldots,0,\underbrace{\omega_k}_{k'th\ \ place},0,\ldots,0\right)$$

ז"א הזנת $\underline{\omega I}$ לפ"א משותפת כאשר בווקטור $\underline{I}$ ישנם אחדים רק במקומות הרלוונטים.

4. התמרת פורייה הפוכה:

$$f_{\underline{X}}\left(\underline{x}\right)=\frac{1}{\left(2\pi\right)^{n}}=\int_{\underline{\omega}\in\mathbb{R}^{N}}\cdots\int e^{-j\underline{\omega}^{T}\underline{x}}\Phi_{\underline{X}}\left(\underline{\omega}\right)d\underline{\omega}$$

5. הקשר בין הפונקציה האופיינית המשותפת לבין המומנטים נתון ע"י:

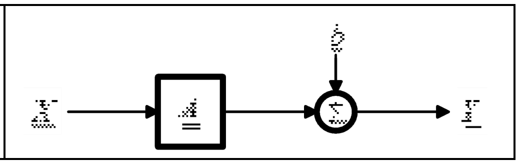
$$m_{k_1,\ldots,k_n}=\frac{1}{j^{k_1+\ldots+k_n}}\left[\frac{d^{k_1+\ldots+k_n}}{d\omega_1^{k_1}\cdot\ldots\cdot d\omega_n^{k_n}}\Phi_{\underline{X}}\left(\underline{\omega}\right)\right]\Bigg|_{\omega=0}$$

6. פונקציה אופיינית משותפת הינה פונקציה ליניארית (נובע מליניאריות של אינטגרל ותכונות אקספוננט).
נראה את השפעת טרנספורמציה ליניארית על הפ"א המשותפת.

נתון ו"א $\underline{X}$ עם פ"א משותפת $\Phi_{\underline{X}}\left(\underline{\omega}\right)$, הו"א עובר טרנספורמציה ליניארית ידועה.
נקבל:

$$\underline{Y}_{M\times1}=\underline{\underline{A}}_{M\times N}\cdot\underline{X}_{N\times1}+\underline{b}_{M\times1}$$

$$\Phi_{\underline{Y}}\left(\underline{\omega}\right)=e^{j\underline{\omega}^{T}\underline{b}}\cdot\Phi_{\underline{X}}\left(\underline{\underline{A}}^{T}\underline{\omega}\right)$$



הסבר:

$$\Phi_{\underline{Y}}\left(\underline{\omega}\right)=E\left[e^{j\underline{\omega}^{T}\underline{Y}}\right]=E\left[e^{j\underline{\omega}^{T}\left(\underline{\underline{A}}\cdot\underline{X}+\underline{b}\right)}\right]\overset{*}{=}e^{j\underline{\omega}^{T}\underline{b}}\cdot E\left[e^{j\underline{\omega}^{T}\cdot\underline{\underline{A}}\cdot\underline{X}}\right]\overset{**}{=}e^{j\underline{\omega}^{T}\underline{b}}\cdot E\left[e^{j\left(\underline{\omega}^{T}\cdot\underline{\underline{A}}\right)^{T^{T}}\cdot\underline{X}}\right]=$$

$$\overset{***}{=}e^{j\underline{\omega}^{T}\underline{b}}\cdot E\left[e^{j\left(\underline{\underline{A}}^{T}\cdot\underline{\omega}\right)^{T}\cdot\underline{X}}\right]=e^{j\underline{\omega}^{T}\underline{b}}\cdot\Phi_{\underline{X}}\left(\underline{\underline{A}}^{T}\underline{\omega}\right)$$

מעבר $(*)$ ע"י הוצאת קבוע מהתוחלת. מעבר $(**)$ ע"י הוספת טרנספוז פעמיים.

מעבר $(***)$ ע"י הכנסת טרנספוז פעם אחת ולכן גם הפיכת סדר המכפלה.

7. ו"א $\left(X,Y\right)$ בעל רכיבים בת"ס אם"ם $\Phi_{XY}\left(\omega_1,\omega_2\right)=\Phi_X\left(\omega_1\right)\cdot\Phi_Y\left(\omega_2\right)$.
הוכחת כיוון 1 – נתון בת"ס:

$$\Phi_{XY}\left(\omega_1,\omega_2\right)=\iint e^{j\left(\omega_1,\omega_2\right)^{T}\left(x,y\right)}f_{XY}\left(x,y\right)dxdy\overset{\text{Independence}}{=}\iint e^{j\left(\omega_1x+\omega_2y\right)}f_X\left(x\right)f_Y\left(y\right)dxdy=$$

$$=\iint e^{j\omega_1x}e^{j\omega_2y}f_X\left(x\right)f_Y\left(y\right)dxdy=\int e^{j\omega_1x}f_X\left(x\right)dx\int e^{j\omega_2y}f_Y\left(y\right)dy=\Phi_X\left(\omega_1\right)\cdot\Phi_Y\left(\omega_2\right)$$

הוכחת כיוון 2 – יש להוכיח בת"ס: אותו דבר במעברים.

<u>דוגמאות לשימוש בתכונות של פונקצייה אופיינית משותפת:</u>

- דוגמא לתכונה מספר 7

ידוע (X,Y) ו"א בת"ס. נגדיר שני מ"א: $V=h(Y)$ $U=g(X)$. נראה כי גם U,V בת"ס.

$$\Phi_{UV}(\omega_1,\omega_2)=E\left[e^{j(\omega_1,\omega_2)^T(U,V)}\right]=\iint e^{j(\omega_1,\omega_2)^T(U,V)}f_{UV}(u,v)\,dudv=$$
$$=\iint e^{j(\omega_1,\omega_2)^T(g(x),h(y))}f_{XY}(x,y)\,dxdy\overset{\text{Independence}}{=}\iint e^{j(\omega_1 g(x),\omega_2 h(y))}f_X(x)f_Y(y)\,dxdy=$$
$$=\int e^{j\omega_1 g(x)}f_X(x)\,dx\int e^{j\omega_2 h(y)}f_Y(y)\,dy=\Phi_U(\omega_1)\cdot\Phi_V(\omega_2)$$

משמעות הדבר היא שפונקציות של מ"א בת"ס מניבות מ"א בת"ס.

- דוגמא לתכונה מספר 6

יהיו מ"א X,Y בת"ס, נגדיר $Z=X+Y$. כבר ראינו כי $f_Z=f_X*f_Y$, נוכיח שוב.

נכתוב את Z כמערכת ליניארית:

$A=(1,1)$ $b=\underline{0}$ $Z=AX+b$	$\begin{pmatrix}X\\Y\end{pmatrix}\rightarrow\boxed{A=(1,1)}\rightarrow Z=X+Y$

$$\Phi_Z(\omega)=\Phi_{XY}\left(A^T\cdot\omega\right)=\Phi_{XY}\left((1,1)^T\cdot\omega\right)=\Phi_{XY}\left(\begin{pmatrix}\omega\\\omega\end{pmatrix}\right)\overset{\text{Independence}}{=}\Phi_X(\omega)\cdot\Phi_Y(\omega)$$

<u>**סטטיסטיקה מסדר II של ווקטורים אקראיים**</u>

על מנת לנתח מערכות ליניאריות נתעניין בסטטיסטיקה מסדר שני בלבד מאחר וסטטיסטיקה מסדר גבוה יותר לא תורמת לניתוח. הדבר יבוא לידי ביטוי בשערוך לינארי (מאחר והמערכת ליניארית) אשר ילמד בהמשך ובאפיון ו"א גאוסי (מאחר ושאר המומנטים מתאפסים).

עבור ו"א $\underline{X}$ נוכל לחשב את התוחלות של המשתנים האקראיים ממנו הוא מורכב, את השונויות שלהם ואת השונויות המשותפות של כל זוג מרכיבים ע"י מומנטים משותפים כפי שראינו לגביי סטטיסטיקה מסדר שני.

$$m_{0,\dots,0,\underbrace{1}_{k'th\ place},0,\dots,0}=E\left[X_k\right]=\eta_{X_k}$$
$$m_{0,\dots,0,\underbrace{1}_{k'th\ place},0,\dots,0,\underbrace{1}_{l'th\ place},0,\dots,0}=E\left[X_k\cdot X_l\right]$$
$$\mu_{0,\dots,0,\underbrace{1}_{k'th\ place},0,\dots,0,\underbrace{1}_{l'th\ place},0,\dots,0}=E\left[(X_k-\eta_k)\cdot(X_l-\eta_l)\right]=\sigma_{X_kX_l}=Cov(X_k,X_l)$$
$$\mu_{0,\dots,0,\underbrace{2}_{k'th\ place},0,\dots,0}=E\left[(X_k-\eta_k)^2\right]=\sigma_{X_k}^2=Var(X_k)$$

הסימונים של המומנטים באופן הנ"ל מייגעים. נעבור לסימון מטריציוני וכך ע"י מטריצות נוכל לייצג את כל המומנטים הנ"ל, ז"א ע"י מטריצות נייצג את הסטטיסטיקה מסדר שני של ו"א.

<u>ווקטור התוחלות</u>

ווקטור התוחלות של ווקטור אקראי $\underline{X}$ מציג את כל המומנטים מסדר ראשון של הווקטור:

$$\eta_{\underline{X}} = E\left[\underline{X}\right] = \begin{bmatrix} E\left[X_1\right] \\ \vdots \\ E\left[X_N\right] \end{bmatrix} = \begin{bmatrix} \eta_{X_1} \\ \vdots \\ \eta_{X_N} \end{bmatrix}$$

<u>מטריצת קורולציה</u>

מטריצת הקורולציה של ווקטור אקראי $\underline{X}$ מציגה את כל המומנטים מסדר שני של הווקטור:

$$R_{\underline{X},\underline{X}} = E\left[\underline{X}\cdot\underline{X}^T\right] = \begin{bmatrix} E\left[X_1^2\right] & \cdots & E\left[X_1\cdot X_N\right] \\ \vdots & \ddots & \vdots \\ E\left[X_N\cdot X_1\right] & \cdots & E\left[X_N^{\ 2}\right] \end{bmatrix}$$

נבחין כי המכפלה $\underline{X}\cdot\underline{X}^T$ מניבה מטריצה במימד $N\times N$ ומחשבים תוחלת על כל איבר במטריצה בנפרד.
הערות:

- נזהה שזו מטריצה סימטרית מאחר שמתקיים $E\left[X_i\cdot X_j\right] = E\left[X_j\cdot X_i\right] \quad \forall i,j$.
- נסמן "גישת אינדקסים" לאיברים במטריצה ע"י $R_{\underline{X}\underline{X}}(i,j) = E\left[X_i\cdot X_j\right]$.
- $R_{\underline{X}\underline{X}}$ אלכסונית אם"ם רכיבי $\underline{X}$ אורתוגונליים הדדית (הרי $E\left[X_i\cdot X_j\right]=0 \quad \forall i\neq j$).

מטריצת הקו-ואריאנס (Covariance)

מטריצת הקו-ואריאנס של ווקטור אקראי $\underline{X}$ מציגה את כל המומנטים המרכזיים מסדר שני של הווקטור:

$$C_{\underline{X},\underline{X}} = R_{\underline{X}-\eta_{\underline{X}},\underline{X}-\eta_{\underline{X}}} = E\left[\left(\underline{X}-\eta_{\underline{X}}\right)\cdot\left(\underline{X}-\eta_{\underline{X}}\right)^T\right] =$$

$$= \begin{bmatrix} E\left[\left(X_1-\eta_{X_1}\right)^2\right] & \cdots & E\left[\left(X_1-\eta_{X_1}\right)\cdot\left(X_N-\eta_{X_N}\right)\right] \\ \vdots & \ddots & \vdots \\ E\left[\left(X_N-\eta_{X_N}\right)\cdot\left(X_1-\eta_{X_1}\right)\right] & \cdots & E\left[\left(X_N-\eta_{X_N}\right)^2\right] \end{bmatrix} =$$

$$= \begin{bmatrix} \sigma_{X_1}^2 & \cdots & Cov\left(X_1, X_N\right) \\ \vdots & \ddots & \vdots \\ Cov\left(X_N, X_1\right) & \cdots & \sigma_{X_N}^2 \end{bmatrix}$$

הערות:

- כמובן שגם זו מטריצה סימטרית וגם עבורה נסמן גישת אינדקסים.
- מתקיים $C_{\underline{X}\underline{X}} = R_{\underline{X}\underline{X}} - \eta_{\underline{X}} \cdot \eta_{\underline{X}}^T$.
- $C_{\underline{X}\underline{X}}$ אלכסונית אם"ם רכיבי $\underline{X}$ חס"ק באופן הדדי (הרי $Cov\left(X_i, X_j\right) = 0 \quad \forall i \neq j$).

תכונות של מטריצת קורולציה ומטריצת קו-ואריאנס

1. $R_{\underline{X}\underline{X}}, C_{\underline{X}\underline{X}}$ הינן מטריצות סימטריות (מיידי) ולכן מתלכסנות אורתונורמלית.
2. $R_{\underline{X}\underline{X}}, C_{\underline{X}\underline{X}}$ הינן מטריצות אי-שליליות מוגדרות (Positive-Semi-Definite) – הוכחה בהמשך.

מטריצה סימטרית – תזכורת:

מטריצה $\underline{\underline{A}}$ תיקרא מטריצה סימטרית אם היא ריבועית ומתקיים $\underline{\underline{A}}^T = \underline{\underline{A}}$.

לכסון מטריצה סימטרית – תזכורת:

מטריצה סימטרית $\underline{\underline{A}}$ מתלכסנת אורתונורמלית, ז"א קיימת מטריצה $\underline{\underline{P}}$ כך ש:

1. $\underline{\underline{P}} \cdot \underline{\underline{P}}^T = \underline{\underline{I}}$
2. $\underline{\underline{P}}^T \cdot \underline{\underline{A}} \cdot \underline{\underline{P}} = \underline{\underline{\Omega}}$ כאשר $\underline{\underline{\Omega}}$ היא מטריצה אלכסונית.

מטריצה $\underline{\underline{P}}$ הזו מורכבת מן הווקטורים העצמיים (פתרונות הפולינום האופייני) של המטריצה $\underline{\underline{A}}$ בעמודות, מטריצה $\underline{\underline{\Omega}}$ היא מטריצה אלכסונית עם הערכים העצמיים של $\underline{\underline{A}}$ על האלכסון הראשי (מסודרים בסדר של הו"ע במטריצה $\underline{\underline{P}}$).

$$\underline{v}_i \cdot \underline{v}_j = 0 \quad \forall i \neq j, \qquad \|\underline{v}_i\| = 1 \quad \forall i, \qquad \begin{gathered} \underline{\underline{P}} = \left[\underline{v}_1 | \cdots | \underline{v}_N\right] \\ \{\lambda_i\} : \det\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) = 0, \\ \left(\underline{\underline{A}} - \lambda_i \underline{\underline{I}}\right) \cdot \underline{v}_i = 0 \end{gathered}$$

מטריצה אי-שלילית מוגדרת (PSD Positive-Semi-Definite) – תזכורת:

מטריצה סימטרית $\underline{\underline{A}}$ תיקרא PSD אם מתקיים $\forall \underline{v} \in \mathbb{R}^N : \quad \underline{v}^T \cdot \underline{\underline{A}} \cdot \underline{v} \geq 0$.

הערות:

- מטריצה $\underline{\underline{A}}$ PSD הינה סימטרית ולכן ניתנת ללכסון.
- למטריצה $\underline{\underline{A}}$ PSD ישנם ערכים עצמיים אי-שליליים ($\lambda_1, \ldots, \lambda_N \geq 0$).
- למטריצה סימטרית $\underline{\underline{A}}$ ישנם ע"ע אי שליליים בלבד ($\forall i: \quad \lambda_i \geq 0$) אם"ם מטריצה $\underline{\underline{A}}$ הינה PSD.

תוחלת של מכפלת מטריצות – תזכורת:

בהינתן מטריצה אקראית $\underline{\underline{X}}_{N\times M}$ ומטריצות דטרמיניסטיות (קבועות) $\underline{\underline{A}}_{K\times N},\underline{\underline{B}}_{M\times L},\underline{\underline{C}}_{N\times M},\underline{\underline{D}}_{K\times L}$ מתקיים:

$$\begin{aligned}&E[A\cdot X]=A\cdot E[X]\\&E[X\cdot B]=E[X]\cdot B\qquad\Rightarrow\quad E[A\cdot X\cdot B+D]=A\cdot E[X]\cdot B+D\\&E[X+C]=E[X]+C\end{aligned}$$

נוכיח את אחד הכללים, נראה עבור איבר יחיד במטריצה (שאר ההוכחות באופן דומה):

$$\left(E[A\cdot X]\right)_{j,k}=E\left[\left(A\cdot X\right)_{j,k}\right]=E\left[\sum_{i=1}^{N}a_{j,i}\cdot x_{i,k}\right]=\sum_{i=1}^{N}E\left[a_{j,i}\cdot x_{i,k}\right]=$$
$$=\sum_{i=1}^{N}a_{j,i}\cdot E\left[x_{i,k}\right]=\left(A\cdot E[X]\right)_{j,k}$$

<u>הוכחה עבור תכונת PSD של מטריצת קורולציה ומטריצת קו-ואריאנס</u>

עבור ווקטור אקראי $\underline{X}$ אנו יודעים כי $R_{\underline{X}\underline{X}}$ סימטרית. נרצה להוכיח כי $\forall\underline{b}\in\mathbb{R}^N:\quad \underline{b}^T\cdot R_{\underline{X}\underline{X}}\cdot\underline{b}\ge 0$.

יהי ווקטור $\underline{b}=(b_1,\dots,b_N)\in\mathbb{R}^N$ כלשהו, נגדיר מ"א $Y=\underline{b}^T\cdot\underline{X}$.

$$0\overset{*}{\le}E\left[Y^2\right]=E[Y\cdot Y]\overset{**}{=}E\left[Y\cdot Y^T\right]=E\left[\left(\underline{b}^T\cdot\underline{X}\right)\cdot\left(\underline{b}^T\cdot\underline{X}\right)^T\right]=E\left[\left(\underline{b}^T\cdot\underline{X}\right)\cdot\left(\underline{X}^T\cdot\underline{b}\right)\right]=$$
$$=E\left[\underline{b}^T\cdot\underline{X}\cdot\underline{X}^T\cdot\underline{b}\right]\overset{***}{=}\underline{b}^T\cdot E\left[\underline{X}\cdot\underline{X}^T\right]\cdot\underline{b}=\underline{b}^T\cdot R_{\underline{X}\underline{X}}\cdot\underline{b}\qquad\Rightarrow R_{\underline{X}\underline{X}}\quad is\quad PSD$$

מעבר $(*)$ כי תמיד מתקיים $\forall Y\in\mathbb{R}:Y^2\ge 0$. מעבר $(**)$ כי נוכל להוסיף טרנספוז על מספר הרי $\forall a\in\mathbb{R}:a^T=a$. מעבר $(***)$ לפי תכונות של תוחלת של מכפלת מטריצות.

נוכיח באותו אופן כי $C_{\underline{X}\underline{X}}$ הינה PSD ע"י הגדרת $Y=\underline{b}^T\cdot\left(\underline{X}-\eta_{\underline{X}}\right)$. מעבר לתהליך ההוכחה מחדש ניתן להגיד כי מטריצת $C_{\underline{X}\underline{X}}$ הינה מקרה פרטי של $R_{\underline{X}',\underline{X}'}$ עבור $\underline{X}'=\underline{X}-\eta_{\underline{X}}$ ולכן תכונותיה של מטריצת R נשמרות.

מטריצת קרוס-קורולציה ומטריצת קרוס-קו-ואריאנס

עבור זוג ו"א $\underline{X}_{N\times1}$ ו- $\underline{Y}_{M\times1}$ נגדיר:

- מטריצת קרוס-קורולציה

$$R_{\underline{X},\underline{Y}} = E\left[\underline{X}\cdot\underline{Y}^T\right] = \begin{bmatrix} E[X_1\cdot Y_1] & \cdots & E[X_1\cdot Y_M] \\ \vdots & \ddots & \vdots \\ E[X_N\cdot Y_1] & \cdots & E[X_N\cdot Y_M] \end{bmatrix}$$

- מטריצת קרוס-קו-ואריאנס

$$C_{\underline{X},\underline{Y}} = R_{\underline{X}-\eta_{\underline{X}},\underline{Y}-\eta_{\underline{Y}}} = E\left[\left(\underline{X}-\eta_{\underline{X}}\right)\cdot\left(\underline{Y}-\eta_{\underline{Y}}\right)^T\right] = \begin{bmatrix} Cov(X_1,Y_1) & \cdots & Cov(X_1,Y_M) \\ \vdots & \ddots & \vdots \\ Cov(X_N,Y_1) & \cdots & Cov(X_N,Y_M) \end{bmatrix}$$

תכונות:

1. $R_{\underline{Y}\underline{X}} = \left(R_{\underline{X}\underline{Y}}\right)^T, \quad C_{\underline{Y}\underline{X}} = \left(C_{\underline{X}\underline{Y}}\right)^T$
2. מתקיים $C_{\underline{X}\underline{Y}} = R_{\underline{X}\underline{Y}} - \eta_{\underline{X}}\cdot\eta_{\underline{Y}}^T$.

הערות:

- מטריצות $R_{\underline{X}\underline{Y}}, C_{\underline{X}\underline{Y}}$ אינן בהכרח ריבועיות (ולכן גם לא בהכרח סימטריות), למעשה הן ממימד $\left[|\underline{X}|\times|\underline{Y}|\right]$.
- זו למעשה הכללה של $R_{\underline{X}\underline{X}}, C_{\underline{X}\underline{X}}$ מאחר ונוכל לבחור $\underline{Y}=\underline{X}$.
- עבור שני ו"א $\underline{X}_{N\times1}$ ו- $\underline{Y}_{M\times1}$ נוכל להגדיר ו"א חדש $\underline{Z}_{N+M\times1}$ אשר בנוי משירשור של $\underline{X}$ ו- $\underline{Y}$.

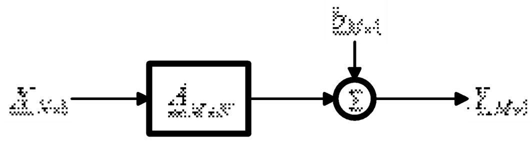
  נוכל להגדיר את מטריצת הקורולציה (והקו-ואריאנס) של $\underline{Z}$ וזו תכלול את תתי-המטריצות הבאות:

$$\underline{Z} = \begin{pmatrix} \underline{X} \\ \underline{Y} \end{pmatrix} \qquad R_{\underline{Z},\underline{Z}} = E\left[\underline{Z}\cdot\underline{Z}^T\right] = E\left[\begin{pmatrix} \underline{X} \\ \underline{Y} \end{pmatrix}\cdot\begin{pmatrix} \underline{X} & \underline{Y} \end{pmatrix}\right] = \begin{bmatrix} R_{\underline{X},\underline{X}} & R_{\underline{X},\underline{Y}} \\ R_{\underline{Y},\underline{X}} & R_{\underline{Y},\underline{Y}} \end{bmatrix}$$

<u>מעבר ווקטור אקראי דרך מערכת ליניארית</u>

<u>סטטיסטיקת המוצא של מערכת ליניארית ווקטורית</u>

יהי $\underline{X}_{N\times1}$ ו"א אשר עובד במערכת ליניארית המתוארת (כאשר $\underline{b}$ ווקטור דטרמיניסטי ו- $\underline{\underline{A}}$ מטריצה דטרמיניסטית):

$\underline{Y}_{M\times1}=\underline{\underline{A}}_{M\times N}\cdot\underline{X}_{N\times1}+\underline{b}_{M\times1}$	

במערכות כאלו נדון בסטטיסטיקה מסדר שני (משום שהמערכת ליניארית).

סטטיסטיקת הכניסה ידועה, כלומר נתונים $C_{\underline{X}\,\underline{X}}$ (או $R_{\underline{X}\,\underline{X}}$) ווקטור התוחלות $\eta_{\underline{X}}$. נרצה לדעת את סטטיסטיקת המוצא ואת קרוס-הסטטיסטיקה בין הכניסה למוצא (ז"א איך נוכל לבטא אותן ע"י סטטיסטיקת הכניסה ונתוני המערכת).

<u>טענות:</u>

1. ווקטור התוחלות של המוצא נתון ע"י:

$$\eta_{\underline{Y}}=E\left[\underline{Y}\right]=E\left[\underline{\underline{A}}\cdot\underline{X}+\underline{b}\right]=\underline{\underline{A}}\cdot\eta_{\underline{X}}+\underline{b}$$

2. מטריצת הקו-ווריאנס של המוצא נתונה ע"י:

$$C_{\underline{Y}\,\underline{Y}}=\underline{\underline{A}}\cdot C_{\underline{X}\,\underline{X}}\cdot\underline{\underline{A}}^T$$

הוכחה:

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

$$=E\left[\left[\underline{\underline{A}}\cdot\left(\underline{X}-\eta_{\underline{X}}\right)\right]\cdot\left[\underline{\underline{A}}\cdot\left(\underline{X}-\eta_{\underline{X}}\right)\right]^T\right]=E\left[\underline{\underline{A}}\cdot\left(\underline{X}-\eta_{\underline{X}}\right)\cdot\left(\underline{X}-\eta_{\underline{X}}\right)^T\cdot\underline{\underline{A}}^T\right]=\underline{\underline{A}}\cdot C_{\underline{X}\,\underline{X}}\cdot\underline{\underline{A}}^T$$

3. מטריצת הקרוס-קו-ואריאנס בין המוצא לכניסה נתונה ע"י:

$$C_{\underline{Y}\,\underline{X}}=\underline{\underline{A}}\cdot C_{\underline{X}\,\underline{X}}$$

הוכחה:

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

$$=E\left[\left[\underline{\underline{A}}\cdot\left(\underline{X}-\eta_{\underline{X}}\right)\right]\cdot\left[\left(\underline{X}-\eta_{\underline{X}}\right)\right]^T\right]=E\left[\underline{\underline{A}}\cdot\left(\underline{X}-\eta_{\underline{X}}\right)\cdot\left(\underline{X}-\eta_{\underline{X}}\right)^T\right]=\underline{\underline{A}}\cdot C_{\underline{X}\,\underline{X}}$$

וכמובן שמתקיים:

$$C_{\underline{X}\,\underline{Y}}=\left(C_{\underline{Y}\,\underline{X}}\right)^T=\left(\underline{\underline{A}}\cdot C_{\underline{X}\,\underline{X}}\right)^T=C_{\underline{X}\,\underline{X}}^{\;T}\cdot\underline{\underline{A}}^T=C_{\underline{X}\,\underline{X}}\cdot\underline{\underline{A}}^T$$

<u>הערה:</u>
באותו האופן נוכל גם לחשב את מטריצות הקורולציה (והקרוס-קורולציה) ע"י חיבור/חיסור של ווקטור התוחלות:

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

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

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

<u>וקטור אקראי גאוסי</u>

ווקטור אקראי $\underline{X}_{N\times 1}$ יקרא ו"א גאוסי (וא"ג) אם לכל $\underline{a}\in\mathbb{R}^N$ מתקיים כי המ"א $Y=\underline{a}^T\cdot\underline{X}$ הינו מ"א גאוסי.

ז"א, אם עבור כל קומבינציה ליניארית של רכיבי הווקטור ($Y=\underline{a}^T\cdot\underline{X}=\sum_{i=1}^{N}a_i\cdot X_i$) נקבל מ"א גאוסי.
למעשה אנו דורשים כי המכפלה הפנימית של הווקטור האקראי עם כל ווקטור מהמימד שלו $\langle\underline{a},\underline{X}\rangle$, ז"א ההיטל של $\underline{X}$ על כל ווקטור $\underline{a}\in\mathbb{R}^N$, תהיה מ"א גאוסי.

עבור $\underline{X}$ וא"ג נסמן $\underline{X}\sim N\left(\eta_{\underline{X}},C_{\underline{X}\underline{X}}\right)$ (מתפלג לפי ווקטור התוחלות ומטריצת הקו-ואריאנס).
<u>מסקנות מההגדרה (הוכחות בהמשך):</u>

1. כל רכיבים של וא"ג הינם מ"א גאוסיים, ז"א עבור $\underline{X}_{N\times 1}$ וא"ג מתקיים X_i מ"א גאוסי לכל $i=1,\dots,N$.
2. וא"ג מקיים סגירות תחת התמרות ליניאריות, ז"א עבור $\underline{X}$ וא"ג גם $\underline{Y}=\underline{\underline{A}}\cdot\underline{X}+\underline{b}$ הינו וא"ג.
3. פונקצייה אופיינית משותפת של וא"ג $\underline{X}\sim N\left(\eta_{\underline{X}},C_{\underline{X}\underline{X}}\right)$ נתונה ע"י:

$$\Phi_{\underline{X}}\left(\underline{\omega}\right) = e^{j\cdot\underline{\omega}^T\cdot\eta_{\underline{X}}-\frac{1}{2}\cdot\underline{\omega}^T\cdot C_{\underline{X}\underline{X}}\cdot\underline{\omega}} = \exp\left(j\cdot\underline{\omega}^T\cdot\eta_{\underline{X}}-\frac{1}{2}\cdot\underline{\omega}^T\cdot C_{\underline{X}\underline{X}}\cdot\underline{\omega}\right)$$

4. jPDF של וא"ג $\underline{X}\sim N\left(\eta_{\underline{X}},C_{\underline{X}\underline{X}}\right)$ (אם מטריצת $C_{\underline{X}\underline{X}}$ הפיכה, ז"א $\det\left(C_{\underline{X}\underline{X}}\right)\neq 0$):

$$f_{\underline{X}}\left(\underline{x}\right) = \frac{1}{\sqrt{\left(2\pi\right)^N\left|C_{\underline{X}\underline{X}}\right|}}\cdot\exp\left(-\frac{1}{2}\left(\underline{x}-\eta_{\underline{X}}\right)^T\cdot\left(C_{\underline{X}\underline{X}}\right)^{-1}\cdot\left(\underline{x}-\eta_{\underline{X}}\right)\right)$$

5. אם מתקיים כי $\underline{Z}=\begin{bmatrix}\underline{X}\\ \underline{Y}\end{bmatrix}$ הינו וא"ג אז הו"א $\underline{Y}|\underline{X}$ הינו וא"ג אשר מקיים:

$$\underline{Y}|\underline{X}\sim N\left(\eta_{\underline{Y}|\underline{X}},C_{\underline{Y}\underline{Y}|\underline{X}}\right)$$
$$\eta_{\underline{Y}|\underline{X}=\underline{x}}=\eta_{\underline{X}\underline{Y}}+C_{\underline{Y}\underline{X}}\cdot\left(C_{\underline{X}\underline{X}}\right)^{-1}\cdot\left(\underline{x}-\eta_{\underline{X}}\right)$$
$$C_{\underline{Y}\underline{Y}|\underline{X}}=C_{\underline{Y}\underline{Y}}-C_{\underline{Y}\underline{X}}\cdot\left(C_{\underline{X}\underline{X}}\right)^{-1}\cdot C_{\underline{X}\underline{Y}}$$

<u>הערה:</u>

אם $\underline{X}\sim N\left(\eta_{\underline{X}},C_{\underline{X}\underline{X}}\right)$ וא"ג וידוע שרכיביו חס"ק (ז"א $C_{\underline{X}\underline{X}}$ הינה מטריצה אלכסונית) אז"י $\underline{X}$ ו"א בת"ס.

<u>הסבר:</u>

ידוע כי מטריצת הקו-ואריאנס אלכסונית, ז"א מהצורה:

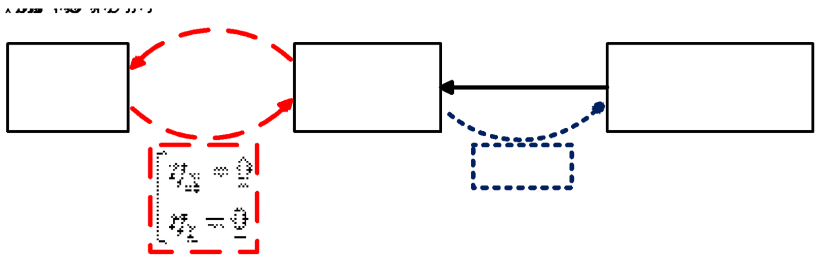
$$C_{\underline{X}\underline{X}}=\begin{bmatrix}\sigma_{X_1}^2 & 0 & \cdots & 0\\ 0 & \ddots & & \vdots\\ \vdots & & \ddots & 0\\ 0 & \cdots & 0 & \sigma_{X_N}^2\end{bmatrix}=\underline{\underline{\Omega}}$$

ולכן ע"י שימוש בנוסחאת הפונקצייה האופיינית המשותפת נקבל:

$$\Phi_{\underline{X}}\left(\underline{\omega}\right)=\exp\left(j\cdot\underline{\omega}^T\cdot\eta_{\underline{X}}-\frac{1}{2}\cdot\underline{\omega}^T\cdot C_{\underline{X}\underline{X}}\cdot\underline{\omega}\right)=\exp\left(j\cdot\sum_{i=1}^{N}\omega_i\cdot\eta_{X_i}-\frac{1}{2}\cdot\underline{\omega}^T\cdot\underline{\underline{\Omega}}\cdot\underline{\omega}\right)=$$
$$=\exp\left(j\cdot\sum_{i=1}^{N}\omega_i\cdot\eta_{X_i}-\frac{1}{2}\cdot\sum_{i=1}^{N}\omega_i^2\cdot\sigma_{X_i}^2\right)=\exp\left(\sum_{i=1}^{N}\left(j\cdot\omega_i\cdot\eta_{X_i}-\frac{1}{2}\cdot\omega_i^2\cdot\sigma_{X_i}^2\right)\right)=$$
$$=\prod_{i=1}^{N}\exp\left(j\cdot\omega_i\cdot\eta_{X_i}-\frac{1}{2}\cdot\omega_i^2\cdot\sigma_{X_i}^2\right)=\prod_{i=1}^{N}\Phi_{X_i}\left(\omega_i\right)$$

<u>סיכום:</u>

אם שני ו"א כלליים $\underline{X}$ ו- $\underline{Y}$ הינם בת"ס הם גם חס"ק. אם בנוסף ווקטור התוחלת של לפחות אחד מהם מתאפס (ז"א $\eta_{\underline{X}},\eta_{\underline{Y}}=\underline{0}$) אז הם אורתונורמלים. אם עבור ווקטורים אקראיים חס"ק ידוע כי הם גאוסיים אז הם גם בת"ס.



הערה - אם ידוע כי X מ"א גאוסי וכי Y מ"א גאוסי ובנוסף כי הם חס"ק אין הדבר אומר כי הם בהכרח בת"ס. לדעת מידע על כל רכיב גאוסי זה לא מספיק – צריך שהרכיבים ירכיבו ווקטור אקראי גאוסי.

ז"א אם ו"א $\underline{X}$ הינו ווקטור שכל רכיביו מא"ג ואפילו חס"ק אין הדבר אומר בהכרח ש- $\underline{X}$ הוא וא"ג. אך אם $\underline{X}$ הוא ו"א עם רכיבים שכל אחד בעצמו מא"ג וכן בת"ס אז $\underline{X}$ הוא וא"ג.

דוגמא - ניקח שני מ"א גאוסיים וחס"ק ונבנה מהם ו"א אשר לא וא"ג.

$$X \sim N(0,1) \qquad B = \begin{cases} 1 & w.p.\, ½ \\ -1 & w.p.\, ½ \end{cases}$$
$$Y = B \cdot X$$

X ו- B הינם בת"ס ומתקבל כי Y הינו מ"א גאוסי בגלל

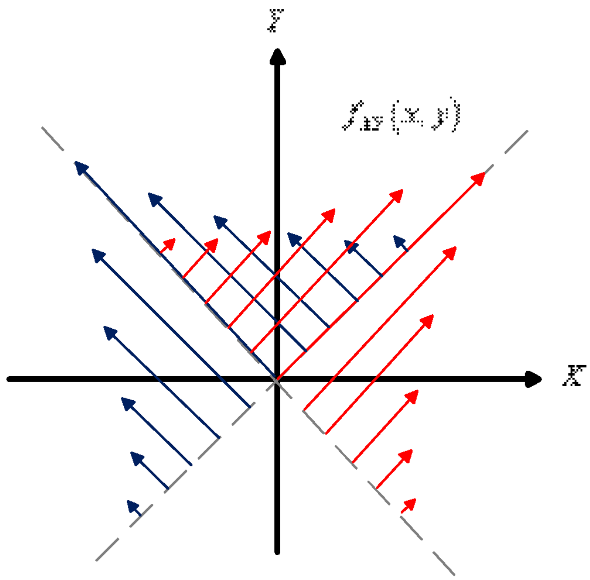
$$(Y \mid B=1) = X \sim N(0,1)$$
$$(Y \mid B=-1) = -X \sim N(0,1)$$

כמו כן מתקיים כי X ו- Y חס"ק:

$$E[X \cdot Y] = E_B\left[E[X \cdot Y \mid B]\right] = \Pr(B=1) \cdot E[X \cdot Y \mid B=1] + \Pr(B=-1) \cdot E[X \cdot Y \mid B=-1] =$$
$$= \tfrac{1}{2} \cdot E\left[X^2\right] + \tfrac{1}{2} \cdot E\left[-X^2\right] = \tfrac{1}{2} \cdot E\left[X^2\right] - \tfrac{1}{2} \cdot E\left[X^2\right] = 0$$
$$Cov(X,Y) = E[X \cdot Y] - E[X] \cdot E[Y] = 0 - 0 \cdot 0 = 0$$

אך הם אינם גאוסיים במשותף. נראה את ה-jPDF שלהם:

$$f_{XY}(x,y) = f_X(x) \cdot f_{Y|X}(y \mid X=x) = f_X(x) \cdot \left[\tfrac{1}{2} \cdot \delta(y-x) + \tfrac{1}{2} \cdot \delta(y+x)\right]$$



הפילוג המותנה אינו פילוג גאוסי משום שהוא לא מורכבת מפונקציית דלתא בודדת (אילו הוא היה אז היינו מתייחסים לדלתא בודדת כמ"א גאוסי אשר בעל תוחלת כמיקום הדלתא ושונות אפס – ז"א מא"ג אשר מתרכז כולו בנקודה אחת - קבוע).
קיבלנו פונקציית jPDF אשר מורכבת משני "הרי-דלתאות" לפי הפילוג הגאוסי של המא"ג X ("הרים" אלו ממוקמים על הישרים $Y=\pm X$ וגובה הדלתאות כמו גאוסיאן) קווי הגובה של פונקציית הצפיפות המשותפת הזו אינם אליפסיים. בנוסף, אילו הפילוג המשותף היה גאוסי היינו יכולים לכתוב ע"י פילוגים גאוסיים.

<u>הלבנה של ווקטור אקראי גאוסי</u>

נתון וא"ג $\underline{X}$ עם ווקטור תוחלות $\eta_{\underline{X}}$ ומטריצת קו-וואריאנס $C_{\underline{X}\,\underline{X}}$ ידועים. נרצה לבצע טרנספורמציה ליניארית כך שהמוצא יהיה בעל אותו מימד עם ווקטור תוחלות $\underline{0}$ ומטריצת קו-וואריאנס אלכסונית, כלומר איברים חס"ק (ומאחר שזהו וא"ג אזי האיברים בת"ס).
התהליך יבוצע באופן הבא:

1. נפחית מ- $\underline{X}$ את ווקטור התוחלות שלו $\eta_{\underline{X}}$, כלומר נגדיר $\underline{X}'=\underline{X}-\eta_{\underline{X}}$.
   כך נקבל וא"ג עם ווקטור תוחלות $\eta_{\underline{X}'}=\underline{0}$ ומטריצת קו-וואריאנס $C_{\underline{X}'\underline{X}'}=C_{\underline{X}\,\underline{X}}$.
2. מכיוון שמטריצת קו-וואריאנס הינה מטריצת PSD אנו יודעים שקיימות מטריצה מלכסנת אורתונורמלית $\underline{\underline{P}}$ ומטריצה אלכסונית $\underline{\underline{\Lambda}}$ כך שמתקיים $\underline{\underline{P}}^T\cdot C_{\underline{X}\,\underline{X}}\cdot\underline{\underline{P}}=\underline{\underline{\Lambda}}$ ($\underline{\underline{P}}$ מלכסנת אורתונורמלית את $C_{\underline{X}'\underline{X}'}=C_{\underline{X}\,\underline{X}}$ ו- $\underline{\underline{\Lambda}}$ מכילה את הערכים העצמיים של $C_{\underline{X}\,\underline{X}}$).

לכן נפעיל על $\underline{X}'$ את הטרנספורמציה $\underline{\underline{P}}^T$ ונסמן את המוצא ב- $\underline{Y}$. יתקיים

$C_{\underline{Y}\underline{Y}} = \underline{\underline{P}}^T \cdot C_{\underline{X}\underline{X}} \cdot \underline{\underline{P}} = \underline{\underline{\Lambda}}$.

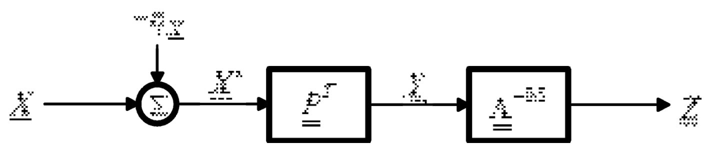
מכיוון שווקטור התוחלות של $\underline{X}'$ מקיים $\eta_{\underline{X}'} = \underline{0}$ אז גם $\eta_{\underline{Y}} = \underline{0}$ (מתכונות מעבר ו"א במערכת ליניארית).

אם בנוסף נרצה שרכיבי ווקטור המוצא יהיו i.i.d אז למעשה נרצה לייצר מטריצת קו-ואריאנס אלכסונית עם איברי אלכסון זהים, לשם פשטות נדרוש כי מטריצת הקו-ואריאנס של המוצא תהיה מטריצת היחידה $\underline{\underline{I}}$.

לשם כך נפעיל על $\underline{Y}$ את הטרנספורמציה $\left(\underline{\underline{\Lambda}}^{0.5}\right)^T = \underline{\underline{\Lambda}}^{-0.5}$ ונסמן את המוצא ב- $\underline{Z}$. נבחין כי נקבל:

$$C_{\underline{Z}\underline{Z}} = \underline{\underline{\Lambda}}^{-0.5} \cdot C_{\underline{Y}\underline{Y}} \cdot \underline{\underline{\Lambda}}^{0.5} = \underline{\underline{\Lambda}}^{-0.5} \cdot \underline{\underline{\Lambda}} \cdot \underline{\underline{\Lambda}}^{0.5} = \underline{\underline{I}}$$

תרשים התהליך:

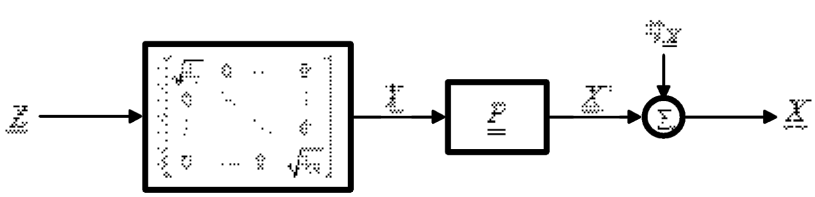


$\underline{X}$	$\eta_{\underline{X}}, \quad C_{\underline{X}\underline{X}}$
$\underline{X}' = \underline{X} - \eta_{\underline{X}}$	$\eta_{\underline{X}'} = \underline{0}, \quad C_{\underline{X}'\underline{X}'} = C_{\underline{X}\underline{X}}$
$\underline{Y} = \underline{\underline{P}}^T \cdot \left(\underline{X} - \eta_{\underline{X}}\right)$	$\eta_{\underline{Y}} = \underline{0}, \quad C_{\underline{Y}\underline{Y}} = \underline{\underline{\Lambda}}$
$\underline{Z} = \underline{\underline{\Lambda}}^{-0.5} \cdot \underline{\underline{P}}^T \cdot \left(\underline{X} - \eta_{\underline{X}}\right)$	$\eta_{\underline{Z}} = \underline{0}, \quad C_{\underline{Z}\underline{Z}} = \underline{\underline{I}}$

<u>ייצור/צביעת ווקטור אקראי גאוסי</u>

נרצה "לייצר" וא"ג כלשהו $\underline{X} \sim N\left(\eta_{\underline{X}}, C_{\underline{X}\underline{X}}\right)$. לשם כך נתחיל עם וא"ג $\underline{Z} \sim N\left(\underline{0}, \underline{\underline{I}}\right)$ מאותו מימד, ז"א וא"ג i.i.d עם תוחלת אפס ושונות אחת. נבצע את התהליך בכיוון ההפוך (כמובן ש- $\{\lambda_1, \ldots, \lambda_N\}$ הם הע"ע של $C_{\underline{X}\underline{X}}$).

תרשים התהליך:



$$\begin{aligned}
&\underline{Z} && \eta_{\underline{Z}} = \underline{0}, \quad C_{\underline{Z}\,\underline{Z}} = \underline{\underline{I}} \\
&\underline{Y} = \underline{\underline{\Lambda}}^{0.5} \cdot \underline{Z} && \eta_{\underline{Y}} = \underline{0}, \quad C_{\underline{Y}\,\underline{Y}} = \underline{\underline{\Lambda}} \\
&\underline{X}' = \underline{\underline{P}} \cdot \underline{\underline{\Lambda}}^{0.5} \cdot \underline{Z} && \eta_{\underline{X}'} = \underline{0}, \quad C_{\underline{X}'\underline{X}'} = C_{\underline{X}\,\underline{X}} \\
&\underline{X} = \underline{\underline{P}} \cdot \underline{\underline{\Lambda}}^{0.5} \cdot \underline{Z} + \eta_{\underline{X}} && \eta_{\underline{X}}, \quad C_{\underline{X}\,\underline{X}}
\end{aligned}$$

<u>תמונה גיאומטרית של תהליך ייצור/הלבנת וא"ג</u>

כזכור פונקציית ה-jPDF של מא"ג נתונה ע"י:

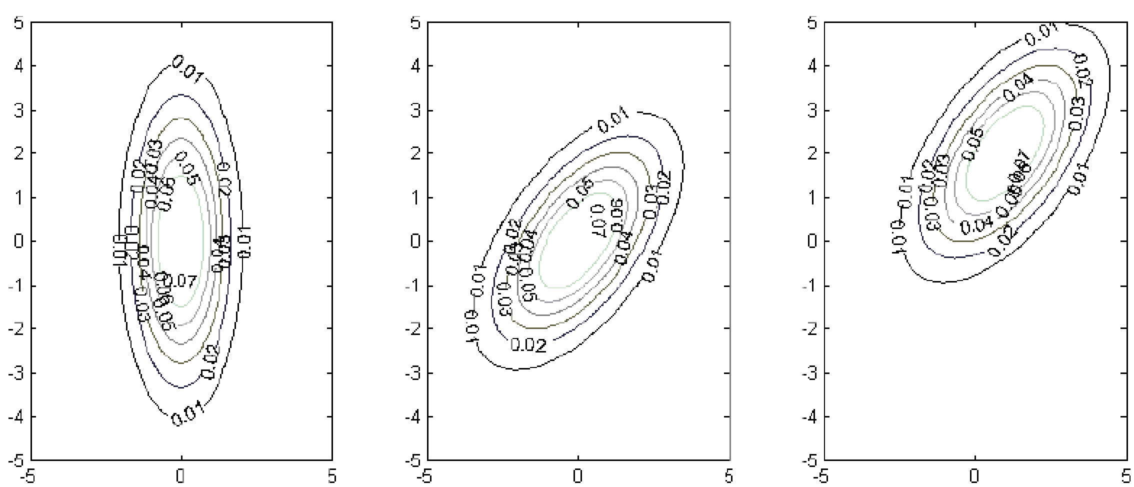
$$f_{\underline{X}}(\underline{x}) = \frac{1}{\sqrt{(2\pi)^N \left|C_{\underline{X}\,\underline{X}}\right|}} \cdot \exp\left(-\frac{1}{2}\left(\underline{x}-\eta_{\underline{X}}\right)^T \cdot \left(C_{\underline{X}\,\underline{X}}\right)^{-1} \cdot \left(\underline{x}-\eta_{\underline{X}}\right)\right)$$

ע"מ להמחיש את שלבי התהליך נרצה לבחון את קווי-הגובה של פונקציות צפיפות הפילוג המשותפות בשלבים השונים (קו-גובה של קבוע מסויים $Const_1$ הינו קבוצת הווקטורים $\left\{\underline{x} \mid f_{\underline{X}}(\underline{x}) = Const_1\right\}$). ממבנה ה-jPDF נוכל לדרוש רק כי ארגומנט האקספוננט יישאר קבוע, ז"א נדרוש כי $\left(\underline{x}-\eta_{\underline{X}}\right)^T \cdot \left(C_{\underline{X}\,\underline{X}}\right)^{-1} \cdot \left(\underline{x}-\eta_{\underline{X}}\right) = Const$.

נזהה כי הדרישה הינה מהתבנית $\underline{X}^T \cdot \underline{\underline{A}} \cdot \underline{X} = Cconst$, תבנית זו הינה תבנית מטריצית ריבועית. עבור תבנית מטריצית ריבועית עם מטריצה $\underline{\underline{A}}$ שהיא PSD פתרון המשוואה הוא אלפסואיד במימד של $\underline{\underline{A}}$.

לשם פשטות נבחן את התהליך עבור דו-מימד, האלפסואיד מתנוון לאליפסה דו-מימדית. נראה את מפת הגבהים עבור השלבים השונים:

הרצאה 13.11.13



<u>הסבר:</u>

- עבור $\underline{Z}=\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ אנו דנים למעשה בשני מא"ג i.i.d בעלי פילוג $Z_1,Z_2\sim N(0,1)$.
דרישת קווי הגובה הנובעת מפונקציית ה-jPDF הינה $z_1^2+z_2^2=Const$, כלומר מעגל סביב הראשית.
- עבור $\underline{Y}=\underline{\underline{\Lambda}}^{0.5}\cdot\underline{Z}=\begin{bmatrix} \sqrt{\lambda_1}\cdot Z_1 \\ \sqrt{\lambda_2}\cdot Z_2 \end{bmatrix}$ נקבל את משוואת קווי הגובה $\frac{y_1^2}{\lambda_1}+\frac{y_2^2}{\lambda_2}=Const$, כלומר אליפסה ניצבת (בעלת צירים $(0,1)$ ו- $(1,0)$) סביב הראשית (בשרטוט הנחנו $\lambda_2>\lambda_1$).
- עבור $\underline{X}'=\underline{\underline{P}}\cdot\underline{Y}=\left[\underline{v}_1,\underline{v}_2\right]\cdot\underline{Y}$ (כאשר $\underline{v}_1,\underline{v}_2$ הם הווקטורים העצמיים של $C_{\underline{X}\underline{X}}$) נקבל למעשה העתקת סיבוב (אנחנו יודעים שהו"ע יוצרים בסיס אוניטרי חדש) כאשר $T\left(\underline{y}=\begin{pmatrix}1\\0\end{pmatrix}\right)=\underline{v}_1$ ו- $T\left(\underline{y}=\begin{pmatrix}0\\1\end{pmatrix}\right)=\underline{v}_2$.
טרנספורמציית סיבוב כזאת נקראת טרנספורמציה אורתונורמלית (יכולה גם ליצור היפוך בסימן).
- עבור $\underline{X}=\underline{X}'+\underline{\eta}_{\underline{X}}=\begin{bmatrix} X_1'+\eta_{X_1} \\ X_2'+\eta_{X_2} \end{bmatrix}$ אנחנו למעשה מקבלים הזזה בקבוע $\begin{bmatrix} \eta_{X_1} \\ \eta_{X_2} \end{bmatrix}$.

<u>הוכחות עבור מסקנות מההגדרה:</u>

1. המסקנה מיידית מהגדרת וא"ג אם נבחר וקטור יחידה באחת הקואורדינטות, לדוגמא עבור $\underline{a}=\underline{e}_i=\left(0,\dots,0,\underset{i'th\ place}{1},0,\dots,0\right)$ נקבל $Y=\underline{a}\cdot\underline{X}=X_i$, ז"א X_i הינו מ"א גאוסי.
2. נבדוק את דרישת ההגדרה עבור ו"א $\underline{Y}$:

$$Z=\left\langle\underline{a},\underline{Y}\right\rangle=\underline{a}^T\cdot\underline{Y}=\underline{a}^T\cdot\left(\underline{\underline{A}}\cdot\underline{X}+\underline{b}\right)=\underline{a}^T\cdot\underline{\underline{A}}\cdot\underline{X}+\underline{a}^T\cdot\underline{b}\overset{\underline{a}'^T=\underline{a}^T\cdot\underline{\underline{A}}}{=}\underline{a}'^T\cdot\underline{X}+\underline{a}^T\cdot\underline{b}$$

וקיבלנו כי $\underline{Y}$ הינו וא"ג כי Z שהתקבל הינו מ"א גאוסי (מאחר ש- $\underline{a}'^T\cdot\underline{X}$ הינו מ"א גאוסי משום ש- $\underline{X}$ וא"ג וקומבינציה ליניארית שלו היא מ"א גאוסי והוספת הקבוע $\underline{a}^T\cdot\underline{b}$ לא משנה).

3. עבור $\underline{\omega} \in \mathbb{R}^N$ כלשהו נתבונן במ"א $Y = \underline{\omega}^T \cdot \underline{X}$. מכיוון ש- $\underline{X}$ הינו וא"ג אז Y הוא מ"א גאוסי.
פונקציה אופיינית של מ"א גאוסי נתונה ע"י:

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

מהקשר בין הפונקציות האופייניות של $\underline{X}$ ושל Y (לפי ההתמרה ביניהם) נסיק:

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

אנו יודעים (מן ההתמרה) כי מתקיים:

$$\eta_Y = \underline{\omega}^T \cdot \eta_{\underline{X}}, \qquad \sigma_Y^2 = \underline{\omega}^T \cdot C_{\underline{X}\underline{X}} \cdot \underline{\omega}$$

ולכן לאחר הצבה נקבל את הביטוי הדרוש.

4. <u>דרך I</u> – נוכל להוכיח ע"י התמרה הפוכה של הפונקציה האופיינית המשותפת $\Phi_{\underline{X}}(\underline{\omega})$.
<u>דרך II</u> –

- שלב 1: נניח $\underline{X}$ וא"ג חס"ק (ולכן בת"ס).
  המקרה המדובר הינו טריוויאלי,

$$f_{\underline{X}}(\underline{x}) = \prod_{i=1}^{N} f_{X_i}(x_i) = \prod_{i=1}^{N}\left(\frac{1}{\sqrt{2\pi \cdot \sigma_{X_i}^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{\left(x_i - \eta_{X_i}\right)^2}{\sigma_{X_i}^2}\right)\right) =$$

$$= \prod_{i=1}^{N}\left(\frac{1}{\sqrt{2\pi \cdot \sigma_{X_i}^2}} \cdot \exp\left(-\frac{1}{2} \cdot \left(x_i - \eta_{X_i}\right)^T \cdot \frac{1}{\sigma_{X_i}^2} \cdot \left(x_i - \eta_{X_i}\right)\right)\right) =$$

$$= \frac{1}{\sqrt{(2\pi)^N \left|C_{\underline{X}\underline{X}}\right|}} \cdot \exp\left(-\frac{1}{2} \cdot \left(\underline{x} - \eta_{\underline{X}}\right)^T \cdot \begin{bmatrix} \frac{1}{\sigma_{X_1}^2} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\sigma_{X_N}^2} \end{bmatrix} \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right) =$$

$$= \frac{1}{\sqrt{(2\pi)^N \left|C_{\underline{X}\underline{X}}\right|}} \cdot \exp\left(-\frac{1}{2}\left(\underline{x} - \eta_{\underline{X}}\right)^T \cdot \left(C_{\underline{X}\underline{X}}\right)^{-1} \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right)$$

- שלב 2: נניח $\underline{X}$ וא"ג כללי.

נרצה להפעיל על $\underline{X}$ טרנספורמציה כזו שיהיה חס"ק/בת"ס, ז"א נרצה לבצע לוא"ג $\underline{X} \sim N\left(\eta_{\underline{X}}, C_{\underline{X}\underline{X}}\right)$ תהליך הלבנה $\underline{Y} = T\left(\underline{X}\right) \sim N\left(0, \underline{\underline{\Lambda}}\right)$ (למעשה הדרישה של תוחלות אפס כלל לא הכרחית כפי שראינו בשלב 1). הטרנספורמציה המתאימה להלבנת וא"ג כללי הינה $\underline{Y} = \underline{\underline{P}}^T \cdot \left(\underline{X} - \eta_{\underline{X}}\right)$. לאחר תהליך ההלבנה $\underline{\underline{\Lambda}}$ הינה מטריצה אלכסונית המכילה ע"ע של $C_{\underline{X}\underline{X}}$.

אנו יודעים כי פונקציית ה-jPDF של $\underline{Y}$ נתונה באופן הבא:

$$f_{\underline{Y}}\left(\underline{y}\right) = \frac{1}{\sqrt{\left(2\pi\right)^N \left|C_{\underline{Y}\underline{Y}}\right|}} \cdot \exp\left(-\frac{1}{2}\left(\underline{y} - \eta_{\underline{Y}}\right)^T \cdot \left(C_{\underline{Y}\underline{Y}}\right)^{-1} \cdot \left(\underline{y} - \eta_{\underline{Y}}\right)\right)$$

נבחן את פונקציית ה-jPDF של $\underline{X}$ (לפי הקשר בין פונקציות צפיפות לאחר טרנספורמציה):

$$f_{\underline{X}}\left(\underline{x}\right) \overset{*}{=} \frac{1}{\left|\underline{\underline{P}}^T\right|} \cdot f_{\underline{Y}}\left(\underline{\underline{P}}^T \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right) =$$

$$= \frac{1}{\sqrt{\left(2\pi\right)^N \left|C_{\underline{Y}\underline{Y}}\right|}} \cdot \exp\left(-\frac{1}{2}\left(\underline{\underline{P}}^T \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right)^T \cdot \left(C_{\underline{Y}\underline{Y}}\right)^{-1} \cdot \left(\underline{\underline{P}}^T \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right)\right) =$$

$$\overset{**}{=} \frac{1}{\sqrt{\left(2\pi\right)^N \left|C_{\underline{X}\underline{X}}\right|}} \cdot \exp\left(-\frac{1}{2}\left(\underline{x} - \eta_{\underline{X}}\right)^T \cdot \left(\underline{\underline{P}} \cdot \left(C_{\underline{Y}\underline{Y}}\right)^{-1} \cdot \underline{\underline{P}}^T\right) \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right) =$$

$$\overset{***}{=} \frac{1}{\sqrt{\left(2\pi\right)^N \left|C_{\underline{X}\underline{X}}\right|}} \cdot \exp\left(-\frac{1}{2}\left(\underline{x} - \eta_{\underline{X}}\right)^T \cdot \left(C_{\underline{X}\underline{X}}\right)^{-1} \cdot \left(\underline{x} - \eta_{\underline{X}}\right)\right)$$

מעבר $(*)$ לפי הקשר $\underline{Y} = \underline{\underline{A}} \cdot \underline{X} + \underline{b} \underset{\left|\underline{\underline{A}}\right| \neq 0}{\Rightarrow} f_{\underline{Y}}\left(\underline{y}\right) = \frac{1}{\left|\underline{\underline{A}}\right|} \cdot f_{\underline{X}}\left(\underline{\underline{A}}^{-1} \cdot \left(\underline{y} - \underline{b}\right)\right)$.

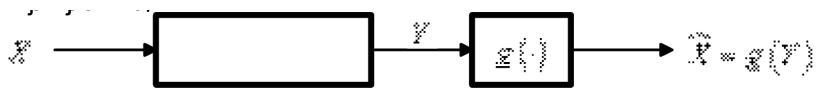
מעברים $(**)$ ו- $(***)$ לפי $\left|C_{\underline{Y}\underline{Y}}\right| = \left|C_{\underline{X}\underline{X}}\right|$ מאחר שמתקיים $C_{\underline{Y}\underline{Y}} = \underline{\underline{P}}^T \cdot C_{\underline{X}\underline{X}} \cdot \underline{\underline{P}}$ ו- $\underline{\underline{P}}$ מטריצה אוניטרית אלכסונית ולכן עם ע"ע אחדים, מכאן ש- $\left|\underline{\underline{P}}\right| = 1$.

5. לא נוכיח כאן, הדבר נובע מתכונות סטטיסטיקה מסדר שני. נוכיח בפרק שערוך.

<u>שערוך</u>

נתונים שני מ"א X ו- Y (או ו"א (X,Y)), אנו יודעים את Y אך מעוניינים ב- X. לכן למעשה מ- Y הגודל המצוי נרצה לדעת את X הגודל הרצוי.
לדוגמא, נשדר אות מידע X בעל פילוג ידוע (לדוגמא X הוא אות בדיד אשר יכול לקבל ערך $X=\pm1$ בסיכוי שווה) בערוץ עם רעש גאוסי $Z\sim N(0,1)$. נקלוט את האות $Y=X+Z$ ונרצה לשערך איזה X שודר.

נשערך את X מתוך Y ע"י פונקצייה כלשהי ונסמן $\hat{X}=g(Y)$.



<u>קריטריוני שגיאה</u>

היינו רוצים לקבל $\hat{X}=X$ אך לא נוכל להבטיח זאת תמיד. הדבר ניתן אך ורק אם הקשר בין X ו- Y הוא חח"ע.
נצטרך להתפשר ונגדיר שגיאת שערוך.

<u>שגיאת שערוך</u>

שגיאת שערוך הינה פונקציה אשר נותנת מדד לטיב השערוך. שגיאת השערוך תסומן באות e (אות קטנה על מנת לא להתבלבל עם סימון התוחלת). ישנן הרבה פונקציות לחישוב שגיאת השערוך.
אנו נעבוד עם שגיאת שערוך הפרשית:

$$e=X-\hat{X}=X-g(Y)$$

X הוא מ"א, גם $\hat{X}$ הוא מ"א ולכן שגיאת השערוך e הינה מ"א.

<u>מדד עיוות</u>

נרצה לתכנן מערכת $g(\cdot)$ כך שנקבל שגיאת שערוך קטנה. e הינו מ"א ולכן הוא אינו קבוע.

לצורך "הסדרת" התנהגותו נגדיר מדד-עיוות. פונקציית $d(e)$ תיקרא מדד עיוות אם $d(e)\geq 0$ לכל e וכן $d(0)=0$. נדגיש כי מאחר ו- e מ"א גם $d(e)$ הינו מ"א.
כמובן שאין דבר כזה מדד עיוות אופטימלי באופן אבסולוטי. הדבר תלוי באופי המערכת.
<u>דוגמאות:</u>

1. מדד "קטיעה" (מתאים עבור מ"א X בדיד בלבד), נקרא גם 'מדד הסתברות שגיאה'	2. מדד שגיאה ריבועית $d(e)=e^2$

$d(e)=\begin{cases}0 & e=0\\ 1 & e\neq 0\end{cases}$	3. מדד שגיאה מוחלטת $d(e)=\lvert e\rvert$

<u>עיוות ממוצע/תוחלת העיוות</u>

כאמור, גם $d(e)$ הוא מ"א. נגדיר עיוות ממוצע (או תוחלת העיוות) ע"י:

$$D=E\left[d(e)\right]=E\left[d\left(X-\hat{X}\right)\right]=E\left[d\left(X-g(Y)\right)\right]$$

כעת נקבל מספר דטרמיניסטי ואיתו נוכל למדוד את "טיב השערוך".

- הערה - מאחר ואנו מתכננים מערכת אשר תעבוד המון פעמים אנו מבצעים תוחלת לפי חוק המספרים הגדולים. התוחלת תהפוך לממוצע זמני בהמשך.
- מטרה - למצוא מערכת $g(\cdot)$ שתביא למינימום את D. אנו מקבלים שהמערכת $g(\cdot)$ תלויה בבחירת $d(e)$.

<u>שערוך אופטימלי</u>

<u>מטרה:</u>

בהינתן מדד עיוות $d(e)$ נרצה למצוא מערכת $g(\cdot)$ שתביא את D למינימום.

$$D=E\left[d\left(X-g(Y)\right)\right]\rightarrow \min$$

<u>פתרון:</u>

נבחר את המערכת אשר עבור כל ערך Y "מנחשת" את α אשר תביא D את למינימום:

$$g_{opt}^{d}(Y=y)=\arg\min_{\alpha}\left(E\left[d\left(X-\alpha\right)\middle|Y=y\right]\right)$$

- הערה - נבחין כי הביטוי $E\left[d\left(X-g(Y)\right)\middle|Y=y\right]$ הוא פונקציה של Y ו- α בלבד.

<u>הוכחה:</u>

$$D=E\left[d\left(X-g(Y)\right)\right]\overset{*}{=}E_Y\left[E\left[d\left(X-g(Y)\right)\middle|Y=y\right]\right]=E_Y\left[h(\alpha,Y)\right]$$

$$D=\int f_Y(y)\left(\int d\left(X-g(Y)\right)f_{X|Y}(x,y)dx\right)dy$$

כאשר מעבר $(*)$ לפי משפט ההחלקה והגדרת $h(\alpha,Y)=E\left[d\left(X-g(Y)\right)\middle|Y=y\right]$.

נרצה להביא את D למינימום כפונקציה של α ולכן נבחר:

$$g_{opt}^{d(e)}\left(Y=y\right)=\alpha^{*}=\arg\min_{\alpha}\left(E\left[d\left(X-\alpha\right)\middle|Y=y\right]\right)$$

ז"א לבדוק את כל האפשרויות ולראות מי נותן פתרון הכי טוב.

<u>דוגמא:</u>

עבור X מ"א דיסקרטי (בדיד) נבחר את מדד העיוות 'מדד הסתברות שגיאה':

$$d\left(e\right)=\begin{cases}0 & e=0\\ 1 & e\neq 0\end{cases}$$

נקבל:

$$D=E\left[d\left(e\right)\right]=\Pr\left(e=0\right)\cdot e+\Pr\left(e=1\right)\cdot e=\Pr\left(e=0\right)\cdot 0+\Pr\left(e=1\right)\cdot 1=\Pr\left(e=1\right)=\Pr\left(\hat{X}\neq X\right)$$

ומכאן שמו, הרי קיבלנו ש- D הוא למעשה הסתברות השגיאה ($D=Pe$ - Probability of error).

<u>משערך אופטימלי תחת מדד הסתברות השגיאה</u>

טענה - המשערך האופטימלי תחת מדד הסתברות השגיאה הוא:

$$\hat{X}_{opt}^{Pe}\left(Y=y\right)=g_{opt}^{Pe}\left(Y=y\right)=\arg\max_{\alpha}\Pr\left(X=\alpha\middle|Y=y\right)$$

זו למעשה בחירה לפי כלל MAP (maximum a posteriori probability) מאחר שמחליטים על הערך המשוערך לאחר "ראיית" Y.

הוכחה -

$$\hat{X}_{opt}^{Pe}\left(Y=y\right)=g_{opt}^{Pe}\left(Y=y\right)=\arg\min_{\alpha}\left(E\left[d\left(X-\alpha\right)\middle|Y=y\right]\right)\overset{*}{=}\arg\min_{\alpha}\left(\Pr\left(X\neq\alpha\middle|Y=y\right)\right)=$$
$$=\arg\max_{\alpha}\Pr\left(X=\alpha\middle|Y=y\right)$$

$$(*)\quad E\left[d\left(X-\alpha\right)\middle|Y=y\right]=\Pr\left(d\left(X-\alpha\right)=0\middle|Y=y\right)\cdot d\left(X-\alpha\right)+\Pr\left(d\left(X-\alpha\right)\neq 0\middle|Y=y\right)\cdot d\left(X-\alpha\right)=$$
$$=\Pr\left(d\left(X-\alpha\right)=0\middle|Y=y\right)\cdot 0+\Pr\left(d\left(X-\alpha\right)\neq 0\middle|Y=y\right)\cdot 1=\Pr\left(d\left(X-\alpha\right)\neq 0\middle|Y=y\right)=$$
$$=\Pr\left(X-\alpha\neq 0\middle|Y=y\right)=\Pr\left(X\neq\alpha\middle|Y=y\right)$$

<u>דוגמא:</u>

אות מידע בדיד X ורעש N נקלט $Y=X+N$ (X ו- N בת"ס).

$$X=\begin{cases}1 & w.p. \quad p\\ -1 & w.p. \quad 1-p\end{cases}$$

$$N\sim N\left(0,\sigma^{2}\right)\qquad\qquad d\left(e\right)=\begin{cases}0 & e=0\\ 1 & e\neq 0\end{cases}$$

$$Y=X+N$$

אם נשתמש במדד הסתברות שגיאה אז אינטואיטיבית:

- אם $p=0.5$ אז נחליט לפי הסף " 0 ". אם $Y\geq 0$ ננחש $\hat{X}=1$, אחרת ננחש $\hat{X}=-1$.
- אם $p>0.5$ אז נתדעף את $\hat{X}=1$ ולכן נחליט לפי ערך סף שלילי כלשהו.

בסה"כ נחליט באופן אופטימלי לפי כלל MAP:

$$\hat{X}_{opt}^{Pe}\left(Y=y\right)=g_{opt}^{Pe}\left(Y=y\right)=\arg\max_{\alpha}\Pr\left(X=\alpha\middle|Y=y\right)$$

ז"א עבור כל $Y=y$ נחשב את ההסתברויות המותנות:

$$\Pr\left(X=1\middle|Y=y\right),\quad \Pr\left(X=-1\middle|Y=y\right)$$

ולבחור את הסביר יותר. לשם החישוב נצטרך לדעת את הפילוג של משתנים אקראיים אלו.

נזהה כי המ"א $Y|X$ מתפלג גאוסית עם מאחר שהרעש N הוא רעש גאוסי ובהינתן X אנו למעשה מבצעים הזזה.

$$\left(Y\middle|X=x\right)\sim N\left(x,\sigma^2\right)\quad\Rightarrow\quad\begin{cases}\left(Y\middle|X=1\right)\sim N\left(1,\sigma^2\right)\\\left(Y\middle|X=-1\right)\sim N\left(-1,\sigma^2\right)\end{cases}$$

ע"מ למצוא את הפילוג המותנה ההפוך נשתמש בחוק בייס:

$$\Pr\left(X=1\middle|Y=y\right)\overset{Bayes}{=}\frac{f_{Y|X=1}\left(y\,|\,1\right)\cdot\Pr\left(x=1\right)}{P_Y\left(y\right)}=\frac{f_N\left(y-1\right)\cdot p}{p\cdot f_N\left(y-1\right)+\left(1-p\right)\cdot f_N\left(y+1\right)}=$$

$$=\frac{p\cdot\exp\left(-\frac{\left(y-1\right)^2}{2\sigma^2}\right)}{p\cdot\exp\left(-\frac{\left(y-1\right)^2}{2\sigma^2}\right)+\left(1-p\right)\cdot\exp\left(-\frac{\left(y+1\right)^2}{2\sigma^2}\right)}$$

$$\Pr\left(X=-1\middle|Y=y\right)\overset{Bayes}{=}\frac{f_{Y|X=1}\left(y\,|\,-1\right)\cdot\Pr\left(x=-1\right)}{P_Y\left(y\right)}=\ldots=\frac{\left(1-p\right)\cdot\exp\left(-\frac{\left(y+1\right)^2}{2\sigma^2}\right)}{p\cdot\exp\left(-\frac{\left(y-1\right)^2}{2\sigma^2}\right)+\left(1-p\right)\cdot\exp\left(-\frac{\left(y+1\right)^2}{2\sigma^2}\right)}$$

נרצה למצוא את הסף המדובר, ז"א לבדוק עבור איזה y -ים מתקיים

$\Pr\left(X=1\middle|Y=y\right)>\Pr\left(X=-1\middle|Y=y\right)$ ונחליט $\hat{X}_{opt}^{Pe}\left(Y\right)=1$, ועבור איזה –ים מתקיים

$\Pr(X=1|Y=y) < \Pr(X=-1|Y=y)$ ונחליט $\hat{X}_{opt}^{Pe}(Y)=-1$ (ההחלטה על הסף עצמו לא קריטית). נמצא את הסף:

$$\Pr(X=1|Y=y) \underset{-1}{\overset{1}{\gtrless}} \Pr(X=-1|Y=y)$$

$$p\cdot\exp\left(-\frac{(y-1)^2}{2\sigma^2}\right) \underset{-1}{\overset{1}{\gtrless}} (1-p)\cdot\exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)$$

$$\exp\left(-\frac{(y-1)^2}{2\sigma^2}+\frac{(y+1)^2}{2\sigma^2}\right) \underset{-1}{\overset{1}{\gtrless}} \frac{(1-p)}{p}$$

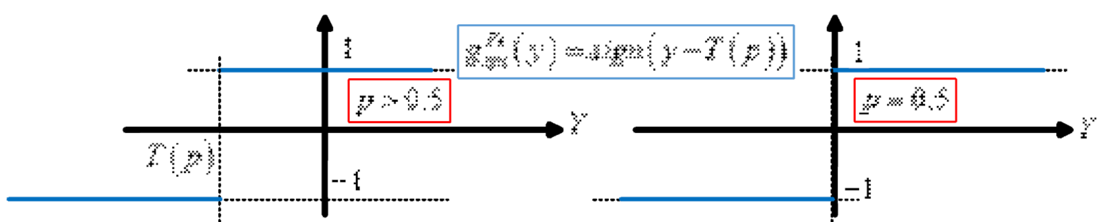
$$\exp\left(\frac{4y}{2\sigma^2}\right) \underset{-1}{\overset{1}{\gtrless}} \frac{(1-p)}{p}$$

מכיוון שאקספוננט הינה פונקציה מונוטונית עולה נוכל להוציא $\ln$ משני האגפים ונקבל:

$$\frac{4y}{2\sigma^2} \underset{-1}{\overset{1}{\gtrless}} \ln\left(\frac{(1-p)}{p}\right)$$

$$y \underset{-1}{\overset{1}{\gtrless}} \frac{1}{2}\cdot\sigma^2\cdot\ln\left(\frac{(1-p)}{p}\right) = T(p)$$

נראה את כלל ההחלטה בצורה גרפית:



שערוך במובן של שגיאה ריבועית ממוצעת מינימלית (MMSE)

ראינו שערוך אופטימלי במובן הסתברות השגיאה, ז"א שערוך אשר מביא למינימום את הסיכוי לשגיאה.

לשערוך זה מספר חסרונות מהותיים:

1. מדד העיוות בינארי ולכן מתאים רק למ"א בדידים.
2. אין שליטה על גודל השגיאה כשהיא קיימת ואנו רק ממזערים את הסיכוי לשגיאה.
3. המערכת מבוצעת לפי כלל MAP, ז"א מותאם כל פעם ערך אופטימלי בהתאם לערך הנקלט.

אם נשתמש במדד עיוות של שגיאה ריבועית $d(e)=e^2$ אז תוחלת העיוות תוגדר להיות:

$$D=E\left[d(e)\right]=E\left[\left(X-\hat{X}\right)^2\right]$$

זהו מדד MSE (Mean Squared Error). נרצה למצוא מערכת שערוך אופטימלית במובן MSE, כלומר נרצה למצוא מערכת $g(\cdot)$ אשר תביא את תוחלת העיוות (העיוות הממוצע) למינימום (MMSE – Minimum MSE).

$$D = E\left[d(e)\right] = E\left[\left(X - \hat{X}\right)^2\right] \to \min$$

נבחין כי תוחלת העיוות למעשה מייצגת את הספק השגיאה.

<u>משערך אופטימלי תחת מדד MSE</u>

טענה - המשערך האופטימלי תחת מדד MSE הוא:

$$\hat{X}_{opt}^{MSE}\left(Y = y\right) = g_{opt}^{MSE}\left(Y = y\right) = E\left[X \middle| Y = y\right]$$

ז"א הערך המשוערך האופטימלי של X עבור Y מסויים הוא למעשה התוחלת המותנת של X בהינתן אותו Y.

הוכחה -

$$\hat{X}_{opt}^{MSE}\left(Y = y\right) = g_{opt}^{MSE}\left(Y = y\right) = \underset{\alpha}{\arg\min}\left(E\left[d\left(X - \alpha\right) \middle| Y = y\right]\right) = \underset{\alpha}{\arg\min}\left(E\left[\left(X - \alpha\right)^2 \middle| Y = y\right]\right)$$

שוב נסמן את הארגומנט של $\underset{\alpha}{\arg\min}$:

$$h(\alpha, Y) = E\left[d\left(X - g(Y)\right) \middle| Y = y\right] = E\left[\left(X - \alpha\right)^2 \middle| Y = y\right] = \int_{-\infty}^{\infty} \left(x - \alpha\right)^2 \cdot f_{X|Y}\left(x \mid y\right) dx$$

נגזור לפי α ונשווה לאפס למציאת קיצון (מאחר וזו משוואה ריבועית של X ניתן להבחין כי זו פרבולה מחייכת ומכאן שמדובר בנקודת מינימום):

$$\frac{\partial}{\partial \alpha} h(\alpha, Y) = -2 \cdot \int_{-\infty}^{\infty} \left(x - \alpha\right) \cdot f_{X|Y}\left(x \mid y\right) dx$$

$$\Rightarrow \quad -2 \cdot \int_{-\infty}^{\infty} \left(x - \alpha^*\right) \cdot f_{X|Y}\left(x \mid y\right) dx = 0$$

$$\int_{-\infty}^{\infty} x \cdot f_{X|Y}\left(x \mid y\right) dx = \int_{-\infty}^{\infty} \alpha^* \cdot f_{X|Y}\left(x \mid y\right) dx$$

$$\int_{-\infty}^{\infty} x \cdot f_{X|Y}\left(x \mid y\right) dx \overset{*}{=} \alpha^* \cdot \int_{-\infty}^{\infty} f_{X|Y}\left(x \mid y\right) dx$$

$$E\left[X \middle| Y = y\right] \overset{**}{=} \alpha^* \cdot 1 = \alpha^*$$

מעבר $(*)$ מאחר וניתן להוציא את הקבוע α^* מהאינטגרל על x. מעבר $(**)$ מאחר שאינגרל בתחום $(-\infty,\infty)$ על כל פונקציית PDF הוא 1.

<u>דוגמא:</u>

אות מידע בדיד X ורעש N נקלט $Y=X+N$ (X ו- N בת"ס).

$$X=\begin{cases}1 & w.p. \quad p\\ -1 & w.p. \quad 1-p\end{cases} \qquad Y=X+N$$

$$N\sim N\left(0,\sigma^2\right) \qquad d(e)=e^2$$

אם נשתמש במדד שגיאה ריבועית (כרגיל עבור e הפרשי):

$$\hat{X}_{opt}^{MSE}\left(Y=y\right)=g_{opt}^{MSE}\left(Y=y\right)=E\left[X\middle|Y=y\right]=1\cdot\Pr\left(\left[X=1\middle|Y=y\right]\right)+(-1)\cdot\Pr\left(\left[X=-1\middle|Y=y\right]\right)=$$

$$=\Pr\left(\left[X=1\middle|Y=y\right]\right)-\Pr\left(\left[X=-1\middle|Y=y\right]\right)=$$

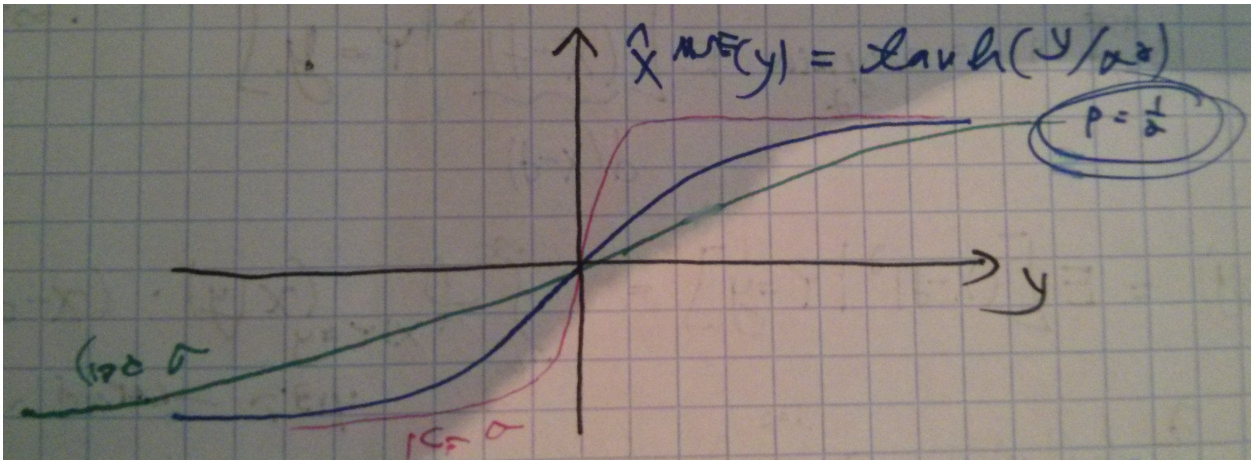
$$=\frac{p\cdot\exp\left(-\frac{(y-1)^2}{2\sigma^2}\right)}{p\cdot\exp\left(-\frac{(y-1)^2}{2\sigma^2}\right)+(1-p)\cdot\exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)}-\frac{(1-p)\cdot\exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)}{p\cdot\exp\left(-\frac{(y-1)^2}{2\sigma^2}\right)+(1-p)\cdot\exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)}=$$

$$=\frac{p\cdot\exp\left(-\frac{(y-1)^2}{2\sigma^2}\right)-(1-p)\cdot\exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)}{p\cdot\exp\left(-\frac{(y-1)^2}{2\sigma^2}\right)+(1-p)\cdot\exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)}=\ldots=\frac{p\cdot\exp\left(\frac{y}{\sigma^2}\right)-(1-p)\cdot\exp\left(-\frac{y}{\sigma^2}\right)}{p\cdot\exp\left(\frac{y}{\sigma^2}\right)+(1-p)\cdot\exp\left(-\frac{y}{\sigma^2}\right)}$$

אם $p=0.5$ אז נקבל מערכת/משערך:

$$\hat{X}_{opt}^{MSE}\left(Y=y\right)=g_{opt}^{MSE}\left(Y=y\right)\overset{p=0.5}{=}\frac{\exp\left(\frac{y}{\sigma^2}\right)-\exp\left(-\frac{y}{\sigma^2}\right)}{\exp\left(\frac{y}{\sigma^2}\right)+\exp\left(-\frac{y}{\sigma^2}\right)}=\tanh\left(\frac{y}{\sigma^2}\right)$$

נראה את כלל ההחלטה בצורה גרפית כאשר $p=0.5$ עבור ערכי σ שונים:



נבחין כי בעוד שכלל ההחלטה אשר נבע מן מדד הסתברות השגיאה היה מאוד נוקשה וחותך, כלל ההחלטה הנובע מן מדד השגיאה הריבועית הינו עדין וחלק יותר.

עבור $\sigma \to 0$ (או אפילו $\sigma \ll 1$) אנחנו מתכנסים לכלל ההחלטה החותך. הדבר הגיוני משום שאם לרעש אין שונות משמעות הדבר שהוא מתרכז סביב התוחלת שלו, שהיא אפס, ואם אין רעש אז ההחלטה נורא קלה סביב אפס. דרך נוספת להסתכל על הדבר הוא שאם נציב σ מאוד קטן אז הארגומנט מאוד גדול ואנו "נזרקים לקצוות הגרף".

עבור $\sigma \gg 1$ הגרף חלק יותר מאחר והרעש "מפוזר" יותר, גם מקרה זה מוצדק הנדסית ומתמטית מאותם שיקולים.

<u>סיכום ביניים – משערכים אופטימליים במובן הסתברות שגיאה ובמובן MSE</u>

תחילה נדגיש כי אנו מתעסקים עם שגיאה הפרשית בלבד, ז"א $e = X - \hat{X} = X - g(Y)$.

בהינתן מדד עיוות מסויים $d(e)$, המשערך האופטימלי מקיים:

$$g_{opt}^{d(e)}(Y = y) = \alpha^* = \arg\min_{\alpha}\left(E\left[d(X-\alpha)\middle|Y = y\right]\right)$$

עבור מדדי עיוות $d(e)$ שונים נקבל מערכות אופטימליות $g_{opt}^{d(e)}(Y)$ שונות:

I. מדד הסתברות שגיאה ומערכת מבוססת כלל MAP (תוחלת העיוות מציגה את הסתברות השגיאה)

$$d(e) = \begin{cases} 0 & e = 0 \\ 1 & e \neq 0 \end{cases} \quad \Rightarrow \quad \begin{aligned} &\hat{X}_{opt}^{Pe}(Y = y) = g_{opt}^{Pe}(Y = y) = \arg\max_{\alpha} \Pr\left(X = \alpha\middle|Y = y\right) \\ &D = \Pr\left(\hat{X} \neq X\right) = Pe \end{aligned}$$

II. מדד שגיאה ריבועית ומערכת תוחלת מותנית (תוחלת העיוות מציגה את הספק השגיאה)

$$d(e)=e^2 \quad \Rightarrow \quad \begin{array}{l} \hat{X}_{opt}^{MSE}(Y=y)=g_{opt}^{MSE}(Y=y)=E\left[X\middle|Y=y\right] \\ D=E\left[\left(X-\hat{X}\right)^2\right]=E\left[e^2\right] \end{array}$$

<u>שערוך ליניארי אופטימלי במובן MSE (LMMSE)</u>

אנו יודעים לתכנן מערכת שערוך אופטימלי במובן MSE לפי:

$$\hat{X}_{opt}^{MSE}(Y=y)=g_{opt}^{MSE}(Y=y)=E\left[X\middle|Y=y\right]$$

לעיתים נרצה לתכנן מערכת שערוך $\hat{X}(Y)=g(Y)$ אשר תהיה ליניארית, ז"א מהצורה $g(Y)=a\cdot Y+b$.

למה שנרצה להתפשר ולהגביל את המשערך להיות ליניארי אם אנחנו יודעים למצוא את המשערך האופטימלי?

I. המימוש פשוט יותר (קל מאוד לבנות מערכת ליניארית).
II. המערכת תלויה (ומושפעת) רק מסטטיסטיקה מסדר שני של הווקטור (X,Y).
III. התבססות על מומנטים מסדרים נמוכים יותר גוררת רובסטיות.

נבחין מנקודה II כי עבור וא"ג המשערך האופטימלי במובן MSE הינו ליניארי הרי לוא"ג קיימים מומנטים (לא אפסים) עד סדר שני בלבד. לכן, ככל שהמידע יותר קרוב לגאוסי אז ה"הפסד" הנוצר משיערון ליניארי יותר קטן.

<u>סימונים:</u>

מעתה נדון בעיקר במשערכים במובן MSE, אופטימלים וליניארים אופטימליים. מטעמי נוחות נגדיר את הסימונים הבאים:

- עבור משערך אופטימלי במובן MSE כללי:

$$\hat{X}_{opt}^{MSE}(Y)=\hat{X}_{opt}(Y)=\hat{X}_{MMSE}(Y)$$

כאשר MMSE = Minimum Mean Squared Error

- עבור משערך ליניארי אופטימלי במובן MSE:

$$\hat{X}_{opt}^{LinearMSE}(Y)=\hat{X}_{LMMSE}(Y)=\hat{X}_{BLE}(Y)$$

כאשר LMMSE = Linear MMSE ו-BLE = Best Linear Estimation.

<u>משערך ליניארי אופטימלי במובן MSE:</u>

נרצה מערכת שערוך ליניארית אופטימלית, ז"א נרצה פונקציה ליניארית של Y אשר תשערך ליניארית את X באופן אופטימלי. נסמן את מערכת השערוך הליניארית האופטימלית הזאת ע"י:

$$\hat{X}_{LMMSE}(Y)=a_{LMMSE}\cdot Y+b_{LMMSE}$$

כאשר מתקיים:

$$\left(a_{LMMSE},b_{LMMSE}\right)=\underset{\alpha,b}{\arg\min}\left(E\left[\left(X-\alpha\cdot Y-b\right)\middle|Y=y\right]\right)$$

נקבל כי:

$$\left(a_{LMMSE},b_{LMMSE}\right)=\left(\frac{\sigma_{XY}}{\sigma_Y^2},\eta_X-\frac{\sigma_{XY}}{\sigma_Y^2}\cdot\eta_Y\right)$$

מפרמטרים אלו נקבל שמערכת שערוך ליניארי אופטימלי במובן MSE נתונה ע"י:

$$\hat{X}_{LMMSE}(Y)=\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right)$$

ומקיימת תוחלת של שגיאה ריבועית ממוצעת (תזכורת $\sigma_{XY}=Cov(X,Y)$):

$$E\left[e_{LMMSE}^2\right]=\sigma_X^2\left(1-\rho_{XY}^2\right)=\sigma_X^2-\frac{\sigma_X^2\cdot\sigma_{XY}^2}{\sigma_X^2\cdot\sigma_Y^2}=\sigma_X^2-\frac{\sigma_{XY}^2}{\sigma_Y^2}$$

נבחין כי מתקיים כי תוחלת העיוות היא הספק השגיאה הריבועית שהוא תוחלת ריבוע העיוות:

$$D=E\left[\left(X-\hat{X}\right)^2\right]=E\left[e^2\right]=MSE$$

<u>הוכחה (מתבססת על תכונות של משערכי LMMSE):</u>

הוכחת נוסחת המשערך נמצאת בהוכחה תכונת הניצבות עבור משערך LMMSE בכיוון

אופטימליות $\leftarrow$ ניצבות.

נוכיח את הביטוי לתוחלת ריבוע השגיאה:

$$\begin{aligned}
E\left[e_{LMMSE}^2\right]&=E\left[\left(X-\hat{X}_{LMMSE}\right)^2\right]=E\left[X^2\right]-2\cdot E\left[X\cdot\hat{X}_{LMMSE}\right]+E\left[\hat{X}_{LMMSE}^2\right]=\\
&=E\left[X^2\right]-2\cdot E\left[\left(X-\hat{X}_{LMMSE}+\hat{X}_{LMMSE}\right)\cdot\hat{X}_{LMMSE}\right]+E\left[\hat{X}_{LMMSE}^2\right]=\\
&=E\left[X^2\right]-2\cdot E\left[\left(e_{LMMSE}+\hat{X}_{LMMSE}\right)\cdot\hat{X}_{LMMSE}\right]+E\left[\hat{X}_{LMMSE}^2\right]=\\
&=E\left[X^2\right]-2\cdot E\left[e_{LMMSE}\cdot\hat{X}_{LMMSE}\right]-2\cdot E\left[\hat{X}_{LMMSE}^2\right]+E\left[\hat{X}_{LMMSE}^2\right]=\\
&\overset{*}{\underset{e\perp\hat{X}(Y)}{=}}E\left[X^2\right]-2\cdot 0-E\left[\hat{X}_{LMMSE}^2\right]=E\left[X^2\right]-E\left[\hat{X}_{LMMSE}^2\right]
\end{aligned}$$

כאשר שיוויון $(*)$ מתקיים לפי תכונת הניצבות (בהמשך). נמשיך ונפתח:

$$E\left[e_{LMMSE}^2\right]=E\left[X^2\right]-E\left[\hat{X}_{LMMSE}^2\right]=E\left[X^2\right]-E\left[\hat{X}_{LMMSE}^2-\eta_X^2+\eta_X^2\right]=$$
$$=E\left[X^2\right]-\eta_X^2-E\left[\hat{X}_{LMMSE}^2-\eta_X^2\right]\underset{E[X]=E[\hat{X}_{LMMSE}]}{\overset{*}{=}}E\left[X^2\right]-\eta_{LMMSE}^2-\left(E\left[\hat{X}_{LMMSE}^2\right]-\eta_{LMMSE}^2\right)=$$
$$=\sigma_X^2-Var\left(\hat{X}_{LMMSE}\right)$$

כאשר שיוויון $(*)$ מתקיים לפי תכונת חוסר ההטייה (בהמשך). נבחין כי:

$$Var\left(\hat{X}_{LMMSE}\right)=E\left[\left(\hat{X}_{LMMSE}-\eta_{LMMSE}\right)^2\right]\underset{E[X]=E[\hat{X}_{LMMSE}]}{=}E\left[\left(\hat{X}_{LMMSE}-\eta_X\right)^2\right]=$$
$$=E\left[\left(\left(\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right)\right)-\eta_X\right)^2\right]=E\left[\left(\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right)\right)^2\right]=\left(\frac{\sigma_{XY}}{\sigma_Y^2}\right)^2\cdot E\left[\left(Y-\eta_Y\right)^2\right]=$$
$$=\left(\frac{\sigma_{XY}}{\sigma_Y^2}\right)^2\cdot\sigma_Y^2=\frac{\sigma_{XY}^2}{\sigma_Y^2}$$

ולכן סה"כ קיבלנו:

$$E\left[e_{LMMSE}^2\right]=\sigma_X^2-\frac{\sigma_{XY}^2}{\sigma_Y^2}=\sigma_X^2-\frac{\sigma_X^2\cdot\sigma_{XY}^2}{\sigma_X^2\cdot\sigma_Y^2}=\sigma_X^2\left(1-\rho_{XY}^2\right)$$

<u>מסקנות משערכי L\MMSE (מתבססות על תכונות של משערכי L\MMSE):</u>

1. אנו יודעים כי $e=X-\hat{X}$, נבחין כי נוכל גם לרשום $\hat{X}=X-e$.

   מתכונות משערכי L\MMSE אנו יודעים כי הזוג $e,\hat{X}$ אורתוגונליים (כי השגיאה e ניצבת לכל פונקציה של המדידות Y). בנוסף אנו יודעים מן התכונות כי מתקיים $E[e]=0$.

   לכן למעשה נוכל להסיק כי הזוג $e,\hat{X}$ הינם חס"ק (עבור שערוך MMSE ו-LMMSE) הרי:

$$Cov\left(e,\hat{X}\right)\triangleq E\left[\left(e-E[e]\right)\cdot\left(\hat{X}-E\left[\hat{X}\right]\right)\right]=E\left[e\cdot\hat{X}\right]-E[e]\cdot E\left[\hat{X}\right]=0-0\cdot E\left[\hat{X}\right]=0$$

2. ראינו כי מתקיים עבור תוחלת ריבוע השגיאה (עבור LMMSE):

$$E\left[e_{LMMSE}^2\right]=\sigma_X^2\left(1-\rho_{XY}^2\right)$$

נזהה כי עבור מדד מתאם $\left|\rho_{XY}\right| \to 1$ תוחלת ריבוע השגיאה של השערוך תתקרב לאפס. בנוסף, אנו יודעים כי תמיד תוחלת שגיאת שערוך LMMSE הינה אפס. לכן למעשה קיבלנו כי השגיאה זהותית אפס:

$$\left|\rho_{XY}\right| = 1 \quad \Rightarrow \quad \begin{array}{c} E\left[e_{LMMSE}^2\right] = 0 \\ E\left[e_{LMMSE}\right] = 0 \end{array} \quad \Rightarrow \quad e \equiv 0$$

הדבר הגיוני כי אם $\rho_{XY} = \pm 1$ אז יש קשר ליניארי של מתיחה בין שני המ"א X ו- Y (תלות דטרמיניסטית ליניארית). מאחר שהמשערך $\hat{X}$ הוא פונקצייה ליניארית של Y אז מתקיים גם קשר ליניארי של מתיחה בין שני המ"א X ו- $\hat{X}$. למעשה בגלל שבמצב זה $e \equiv 0$ אז מתקיים ממש $\hat{X} \equiv X$.

<u>תכונות משערכי MMSE\LMMSE</u>

לשם נוחות, כאשר נרצה להציג תכונה אשר תקפה גם למשערך MMSE וגם למשערך LMMSE נסמן את האינדקס ע"י L\MMSE. משמעות הסימון שהתכונה תקפה לשניהם אך 'רק לאחד מהם כל פעם' (ז"א אין לערבב ביניהם באותו תכונה או משוואה).

1. תכונת ניצבות/אורתוגונליות
   a. משערך MMSE

      משערך $\hat{X}(Y) = g(Y)$ הינו MMSE אם"ם שגיאתו $e = X - g(Y)$ מקיימת $e \perp h(Y)$ לכל פונקציה $h(Y)$. למעשה השגיאה ניצבת לכל פונקציה של המדידות. ז"א $\forall h: \mathbb{R} \to \mathbb{R} \quad E\left[e \cdot h(Y)\right] = 0$.

   b. משערך LMMSE

      משערך ליניארי $\hat{X}(Y) = g(Y) = a \cdot Y + b$ הינו LMMSE אם"ם שגיאתו $e = X - g(Y) = X - (a \cdot Y + b)$ מקיימת $e \perp h(Y)$ לכל פונקציה ליניארית $h(Y)$ (או למעשה נדרוש כי $\forall c, d \in \mathbb{R} \quad e \perp (c \cdot Y + d)$). למעשה השגיאה ניצבת לכל פונקציה ליניארית של המדידות. ז"א

      $$\forall c, d \in \mathbb{R} \quad E\left[e \cdot (c \cdot Y + d)\right] = E\left[\left(X - (a \cdot Y + b)\right) \cdot (c \cdot Y + d)\right] = 0.$$

2. חוסר הטייה (Un-Biasedness)

   משערך L\MMSE מקיים כי תוחלת השגיאה מתאפסת, למעשה הוא מקיים כי תוחלתו של המשערך שווה לתוחלתו של המשתנה המשוערך X.

$$E\left[e_{L\backslash MMSE}\right]=0$$
$$E\left[X-\hat{X}_{L\backslash MMSE}\right]=0$$
$$E\left[\hat{X}_{L\backslash MMSE}\right]=E\left[X\right]=\eta_X$$

3. פיתגורס (Pythagoras)
הספק המ"א אשר אנו רוצים לשערך שווה להספק המשערך של המ"א פלוס הספק השגיאה (כאמור עבור מדד MSE הספק השגיאה מקיים $D=MSE$).

$$E\left[X^2\right]=E\left[\hat{X}^2_{L\backslash MMSE}\right]+\left[e^2_{L\backslash MMSE}\right]$$

נבחין כי אם נתעסק עם משערך לינארי (נגביל את עצמנו) אז הספק השגיאה יגדל על חשבון הספק המשערך.

4. עבור וא"ג $\begin{bmatrix} X \\ Y \end{bmatrix}$ מתקיים $\hat{X}_{MMSE}=\hat{X}_{LMMSE}$, ז"א משערך LMMSE הינו אופטימלי
אם $\begin{bmatrix} X \\ Y \end{bmatrix}$ הוא וא"ג אז המשערך האופטימלי במובן MMSE הינו ליניארי ולמעשה מתקיים:

$$\hat{X}_{MMSE}=E\left[X\middle|Y=y\right]\overset{*}{=}\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(y-\eta_Y\right)=\hat{X}_{LMMSE}$$

כאשר מעבר $(*)$ מתקיים כאשר $\begin{bmatrix} X \\ Y \end{bmatrix}$ הוא וא"ג.

<u>הוכחות:</u>

1. תכונת ניצבות/אורתוגונליות
כיוון I – ניצבות $\leftarrow$ אופטימליות MSE
   a. משערך MMSE
   נתון שמשערך $\hat{X}(Y)=g(Y)$ מקיים $e\perp h(Y)$ לכל פונקציה $h(Y)$. נניח כי $\hat{X}(Y)=g(Y)$ אינו אופטימלי וקיים משערך מתחרה (טוב יותר) $\hat{X}(Y)=\tilde{g}(Y)$.
   נרצה להראות כי הספק השגיאה של המשערך המתחרה לא קטן מזה שמקיים ניצבות:

$$E\left[e^2\right]\le E\left[\tilde{e}^2\right] \quad \Leftrightarrow \quad E\left[\left(X-\hat{X}\right)^2\right]\le E\left[\left(X-\tilde{X}\right)^2\right]$$

נחשב את הספק השגיאה של המשערך המתחרה:

$$E\left[\tilde{e}^2\right]=E\left[\left(X-\hat{X}\right)^2\right]=E\left[\left(X-\tilde{X}\right)^2\right]E\left[\left(X-\tilde{X}\right)^2\right]=E\left[\left[\left(X-\hat{X}\right)+\left(\hat{X}-\tilde{X}\right)\right]^2\right]=$$

$$E\left[\left(X-\hat{X}\right)^2\right]+2\cdot E\left[\left(X-\hat{X}\right)\cdot\left(\hat{X}-\tilde{X}\right)\right]+E\left[\left(\hat{X}-\tilde{X}\right)^2\right]=E\left[e^2\right]+2\cdot E\left[e\cdot\left(\hat{X}-\tilde{X}\right)\right]+$$

$$+E\left[\left(\hat{X}-\tilde{X}\right)^2\right]\overset{*}{=}E\left[e^2\right]+E\left[\left(\hat{X}-\tilde{X}\right)^2\right]\overset{**}{\ge}E\left[e^2\right]$$

כאשר מעבר $(*)$ מאחר ש- $e\perp h(Y)$ וכמובן ש- $\left(\hat{X}-\tilde{X}\right)$ הינה פונקציה של y משום ששני המשערכים הם פונקציות של Y. מעבר $(**)$ מאחר ש- $E\left[\left(\hat{X}-\tilde{X}\right)^2\right]$ הינו גודל אי-שלילי.

b. משערך LMMSE

ההוכחה זהה מכיוון שמתקיים ש- $\left(\hat{X}-\tilde{X}\right)$ הינו פונקצייה ליניארית (כתוצאה של חיסור של פונקציות ליניאריות), ואז $E\left[e\cdot\left(\hat{X}-\tilde{X}\right)\right]=0$ מתכונת הניצבות של e.

כיוון II – אופטימיליות MSE ← ניצבות

a. משערך MMSE

נתון משערך $\hat{X}(Y)=g(Y)$ אופטימלי במובן MSE. יש להראות כי הוא מקיים את תכונת הניצבות, כלומר מקיים $e\perp h(Y)$ לכל פונקציה $h(Y)$. נדרוש כי התכונה תתקיים ונמצא באופן מפורש את המשערך $\hat{X}(Y)=g(Y)$ הזה.

$$\forall h(Y): \quad e \perp h(Y) \quad \Leftrightarrow \quad E\left[e \cdot h(Y)\right] = 0 \quad \Leftrightarrow \quad E\left[\left(X - g(Y)\right) \cdot h(Y)\right] = 0$$

$$E_Y\left[E_{X|Y}\left[\left(X - g(y)\right) \cdot h(y) \middle| Y = y\right]\right] = 0 \qquad \forall h(Y)$$

$$E_Y\left[h(Y) \cdot E_{X|Y}\left[\left(X - g(y)\right) \middle| Y = y\right]\right] = 0 \qquad \forall h(Y)$$

$$E_Y\left[h(Y) \cdot \left(E_{X|Y}\left[X \middle| Y = y\right] - g(Y)\right)\right] = 0 \qquad \forall h(Y)$$

נרצה כי השיוויון יתקיים עבור כל פונקצייה $h(Y)$ ולכן נדרוש (נוכל גם לפתח לפי הגדרת התוחלת לפי אינטגרלים):

$$E_{X|Y}\left[X \middle| Y = y\right] - g(Y) = 0 \quad \Rightarrow \quad g(Y = y) = E_{X|Y}\left[X \middle| Y = y\right] = E\left[X \middle| Y\right]$$

ואכן קיבלנו כי $g(Y)$ נתון לפי נוסחת השונות המותנת, ז"א לפי הגדרת משערך MMSE.

b. משערך LMMSE

נבצע אותו דבר עבור משערך מהצורה $\hat{X}(Y) = g(Y) = a \cdot Y + b$:

$$\forall c, d \in \mathbb{R}: \quad e \perp \left(c \cdot Y + d\right) \quad \Leftrightarrow \quad E\left[e \cdot \left(c \cdot Y + d\right)\right] = 0$$

נבחין כי מספיק לדרוש (כי שניהם מספיקים ע"מ לפרוש את מרחב הפונקציות הליניאריות של Y):

i. $e \perp 1$ (או כל קבוע אחר מ- $\mathbb{R}$)

ii. $e \perp \left(Y - \eta_Y\right)$

$$E\left[e \cdot \left(c \cdot Y + d\right)\right] = E\left[e \cdot \left(c \cdot Y + c \cdot \eta_Y - c \cdot \eta_Y + d\right)\right] = E\left[e \cdot c \cdot \left(Y - \eta_Y\right) + e \cdot \left(c \cdot \eta_Y + d\right)\right] =$$
$$= c \cdot E\left[e \cdot \left(Y - \eta_Y\right)\right] + \left(c \cdot \eta_Y + d\right) \cdot E\left[e \cdot 1\right] = c \cdot 0 + \left(c \cdot \eta_Y + d\right) \cdot 0 = 0$$

כעת נראה מה משמעות זוג דרישות אלו:

$$i \quad e \perp 1 \Leftrightarrow E\left[e \cdot 1\right] = 0$$
$$E\left[X - a \cdot Y - b\right] = 0$$
$$\eta_X - a \cdot \eta_Y - b = 0$$
$$b(a) = \eta_X - a \cdot \eta_Y$$

$$ii \quad e \perp \left(Y - \eta_Y\right) \Leftrightarrow E\left[e \cdot \left(Y - \eta_Y\right)\right] = 0$$
$$E\left[\left(X - a \cdot Y - b\right) \cdot \left(Y - \eta_Y\right)\right] = 0$$

נציב את $b(a)$ ונמשיך:

$$
\begin{aligned}
ii\quad & E\left[\left(X-a\cdot Y-\eta_X+a\cdot\eta_Y\right)\cdot\left(Y-\eta_Y\right)\right]=0\\
& E\left[\left(\left(X-\eta_X\right)-a\cdot\left(Y-\eta_Y\right)\right)\cdot\left(Y-\eta_Y\right)\right]=0\\
& \sigma_{XY}-a\cdot\sigma_Y^2=0 \quad\Rightarrow\quad a=\frac{\sigma_{XY}}{\sigma_Y^2}
\end{aligned}
$$

וכעת "נרכיב" את המשערך שמצאנו (ונקבל שזה משערך $\hat{X}_{LMMSE}$ כפי שהגדרנו):

$$
\hat{X}(Y)=a\cdot Y+b=\begin{bmatrix} a=\dfrac{\sigma_{XY}}{\sigma_Y^2} \\ b(a)=\eta_X-a\cdot\eta_Y \end{bmatrix}=\frac{\sigma_{XY}}{\sigma_Y^2}Y+\eta_X-\eta_Y\frac{\sigma_{XY}}{\sigma_Y^2}=\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right)
$$

2. חוסר הטייה (Un-Biasedness)
   a. משערך MMSE

נראה כי מתקיים $E\left[\hat{X}_{MMSE}\right]=E[X]=\eta_X$.

$$
E\left[\hat{X}_{MMSE}\right]=E_Y\left[\hat{X}_{MMSE}\right]=E_Y\left[E_{X|Y}\left[X|Y=y\right]\right]\overset{*}{=}E[X]
$$

כאשר שיוויון $(*)$ מתקיים לפי נוסחת ההחלקה (בכיוון ההפוך).

   b. משערך LMMSE

במקרה זה ההוכחה מיידית מאחר שדרשנו שיתקיים $e\perp 1$ ולכן $E[e_{LMMSE}]=0$.

3. פיתגורס (Pythagoras)

$$
E\left[X^2\right]=E\left[\left(\left(X-\hat{X}_{L\backslash MMSE}\right)+\hat{X}_{L\backslash MMSE}\right)^2\right]=E\left[\left(e+\hat{X}_{L\backslash MMSE}\right)^2\right]=E\left[e^2\right]+
$$

$$
+2\cdot E\left[e\cdot\hat{X}_{L\backslash MMSE}\right]+E\left[\hat{X}^2_{L\backslash MMSE}\right]\overset{e\perp\hat{X}_{L\backslash MMSE}(Y)}{=}E\left[\hat{X}^2_{L\backslash MMSE}\right]+\left[e^2_{L\backslash MMSE}\right]
$$

כמובן שהאיבר $E\left[e\cdot\hat{X}_{L\backslash MMSE}\right]$ מתאפס עבור מקרה אופטימלי כללי וליניארי (מניצבות השגיאה).

4. עבור וא"ג $\begin{bmatrix}X\\Y\end{bmatrix}$ מתקיים $\hat{X}_{MMSE}=\hat{X}_{LMMSE}$, ז"א משערך LMMSE הינו אופטימלי

ראינו (מסקנות משערכי L\MMSE) כי מתקיים שהזוג $e,\hat{X}$ הינם חס"ק ($Cov\left(e,\hat{X}\right)=0$).

מאחר ש- $\begin{bmatrix} X \\ Y \end{bmatrix}$ וא"ג אז גם $\begin{bmatrix} e_{LMMSE} \\ \hat{X}_{LMMSE} \end{bmatrix}$ וא"ג ולכן הזוג $\hat{X}_{LMMSE}, e_{LMMSE}$ אינו רק חס"ק אלא גם בת"ס. למעשה נבחין כי $\begin{bmatrix} X & Y & \hat{X}_{LMMSE} & e_{LMMSE} \end{bmatrix}$ הוא וא"ג (מדוע?). לכן נסיק:

- הזוג $\hat{X}_{LMMSE}, e_{LMMSE}$ בת"ס
- הזוג Y, e_{LMMSE} בת"ס
- אנו יודעים אפיון מלא של השגיאה $e_{LMMSE} \sim N\left(0,\sigma_X^2\cdot\left(1-\rho_{XY}^2\right)\right)$
- ידוע הפילוג המותנה של $\left(X|Y=y\right)$:

$$\left(X|Y=y\right)=\hat{X}_{LMMSE}\left(y\right)+e_{LMMSE}\left(y\right)$$

$$\hat{X}_{LMMSE}\left(y\right)=\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(y-\eta_Y\right)\sim N\left(\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right),0\right)$$

$$e_{LMMSE}\left(y\right)\sim N\left(0,\sigma_X^2\cdot\left(1-\rho_{XY}^2\right)\right)$$

$$\Rightarrow\quad \left(X|Y=y\right)\sim N\left(\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right),\sigma_X^2\cdot\left(1-\rho_{XY}^2\right)\right)$$

במקביל אנו יודעים איך לחשב את $\hat{X}_{MMSE}\left(Y\right)$ ולכן (יחד עם ידיעת הפילוג המותנה):

$$\hat{X}_{MMSE}\left(Y=y\right)=E\left[X|Y=y\right]=\eta_X+\frac{\sigma_{XY}}{\sigma_Y^2}\left(Y-\eta_Y\right)=\hat{X}_{LMMSE}\left(Y=y\right)$$

<u>תמונה גיאומטרית של משערכי L\MMSE</u>

<u>אקסיומות של מכפלה פנימית:</u>

בהינתן מרחב ווקטורי מכפלה פנימית מקיימת:

1. $\left\langle \underline{X},\underline{Y}\right\rangle=\left\langle \underline{Y},\underline{X}\right\rangle^*$ (מעל אין צורך בצמוד)
2. $\left\langle a\cdot\underline{X},\underline{Y}\right\rangle=a\cdot\left\langle \underline{X},\underline{Y}\right\rangle=\left\langle \underline{X},a\cdot\underline{Y}\right\rangle$
3. $\left\langle \underline{X}+\underline{Y},\underline{Z}\right\rangle=\left\langle \underline{X},\underline{Z}\right\rangle+\left\langle \underline{Y},\underline{Z}\right\rangle$

4. $\langle \underline{X}, \underline{X} \rangle \geq 0, \quad \langle \underline{X}, \underline{X} \rangle = 0 \Leftrightarrow \underline{X} = 0$

<u>מסקנות:</u>

- נבחין שאוסף המ"א (בעלי שונות סופית, ז"א בעלי מומנט שני סופי) הינו מרחב ווקטורי מאחר והוא סגור תחת קומבינציה ליניארית(כל קומבינציה ליניארית שנעשה בהם תיצור מ"א).
- עבור מ"א X, אוסף כל הפונקציות $g(X)$ הינו תת-מרחב ווקטורי של המרחב הווקטורי של המשתנים האקראיים. בפרט עבור אוסף כל הפונקציות הליניאריות של המ"א X.

<u>הגדרת מכפלה פנימית על מרחב המשתנים האקראיים:</u>

נגדיר מכפלה פנימית (מ"פ) על מרחב המשתנים האקראיים ע"י:

$$\langle X, Y \rangle = E[X \cdot Y]$$

נבחן את קיום האקסיומות:

1. טריוויאלי
2. טריוויאלי
3. טריוויאלי
4. כמובן שמתקיים $\langle X, X \rangle \geq 0$ מאחר ש- $\langle X, X \rangle = E[X \cdot X] = E\left[X^2\right]$ ו- X^2 הינו גודל אי-שלילי.

מתקיים $\langle 0,0 \rangle = E[0] = 0$ (כיוון I) וגם אם $\langle X, X \rangle = E\left[X^2\right] = 0$ אז בהכרח $X = 0$ מהגדרת האינטגרל (כיוון II).

<u>הגדרת נורמה של משתנה אקראי:</u>

נגדיר נורמה של מ"א ע"י המ"פ שלו עם עצמו:

$$\|X\| = \langle X, X \rangle = E\left[X^2\right]$$

בעזרת הגדרת הנורמה נוכל להגדיר:

$$\cos\left(\sphericalangle(X,Y)\right) = \frac{\langle X, Y \rangle}{\|X\| \cdot \|Y\|} = \frac{E[X \cdot Y]}{\sqrt{E\left[X^2\right] \cdot E\left[Y^2\right]}} = r(X,Y)$$

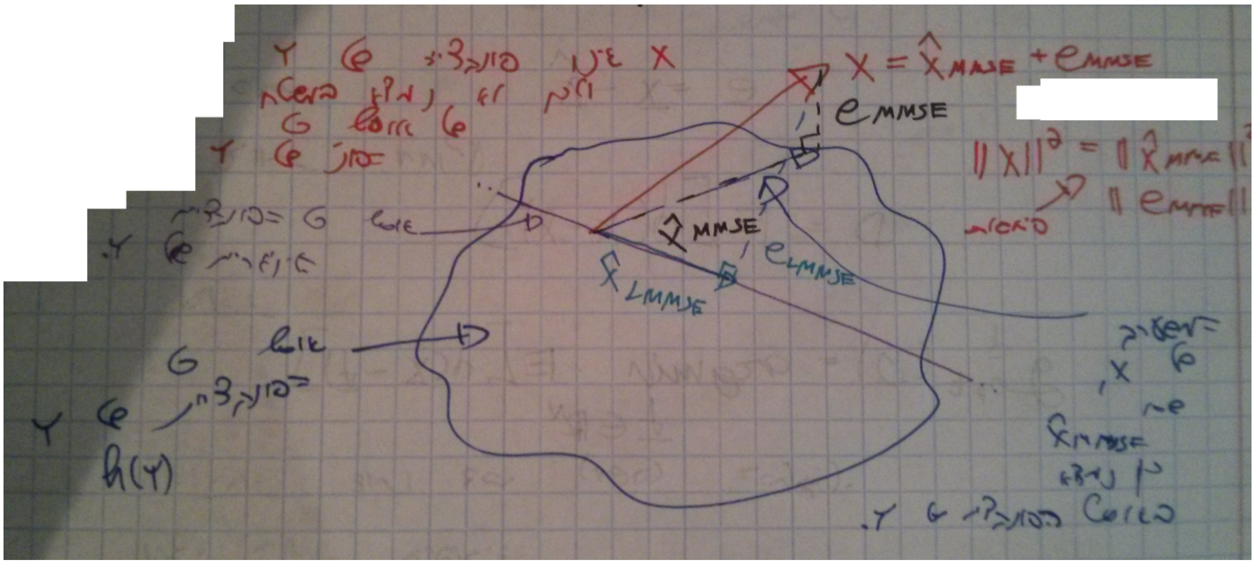
$$\cos\left(\sphericalangle(X-\eta_X, Y-\eta_X)\right) = \ldots = \frac{E\left[(X-\eta_X)\cdot(Y-\eta_Y)\right]}{\sqrt{E\left[(X-\eta_X)^2\right] \cdot E\left[(Y-\eta_Y)^2\right]}} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \rho(X,Y)$$

כזכור מתקיים $-1 \leq r(X,Y) \leq 1$ (וכנ"ל עבור $\rho(X,Y)$). הדבר אכן מסתדר עם הידע על פונקציית $\cos(\cdot)$.

<u>ניצבות של משתנים אקראיים:</u>

נאמר ששני מ"א X ו- Y ניצבים (ונסמן $X \perp Y$) אם"ם מתקיים $\langle X,Y \rangle = 0$ (כלומר $E[X \cdot Y] = 0$).

<u>תמונה גיאומטרית של משערכי L\MMSE:</u>



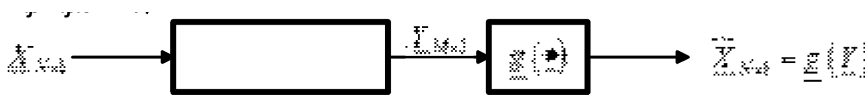
המישור מצייג את אוסף (מרחב) כל הפונקציות האפשריות של Y , ז"א את אוסף כל הפונקציות $h(Y)$.

בתוך מרחב זה קיים תת-מרחב הפונקציות הליניאריות של Y , מסומן ע"י הקו הסגול.

המ"א X אינו פונקצייה של Y ולכן לא על המישור, המשערך $\hat{X}_{MMSE}$ הינו פונקציה של Y ולכן נמצא במישור, המשערך $\hat{X}_{LMMSE}$ הינו פונקציה ליניארית של Y ולכן נמצא על תת-מרחב הפונקציות הליניאריות של Y . בגלל שהמשערכים אופטימליים השגיאה חייבת להיות קטנה ככל האפשר ולכן היא האנך. המשערכים הם למעשה ההיטל של X על מרחב הפונקציות (הכלליות או הליניאריות בהתאם לסוג המשערך). עוד ניתן להבחין כי מאחר שהשגיאות $e_{L\backslash MMSE}$ נתונות לפי ההגדרה על ידי $e_{L\backslash MMSE} = X - \hat{X}_{L\backslash MMSE}$ ומכיוון שאנו מתייחסים למ"א כמרחב ווקטורי השגיאות אנו מבצעים פשוט חיסור ווקטורי ואכן מקבלים ניצבות למרחב הפונקציות המתאים.
באותו אופן היינו יכולים להגביל את המשערך להיות פונקצייה פולינומיאלית עד סדר שני לדוגמא ואז לבצע את ההיטל והאנך לתת-מרחב זה.

שערוך ווקטור אקראי מתוך ווקטור אקראי

למדנו לשערך מ"א מתוך מ"א וכעת נרצה להרחיב את הדיון לשערוך ו"א מתוך ו"א, כלומר ידוע לנו ו"א $\underline{Y}$ ונרצה לשערך ו"א אחר $\underline{X}$. נדגיש כי ווקטורים אלו לא חייבים להיות מאותו מימד (אך $\underline{X}$ ו- $\underline{\hat{X}}$ כן).



המערכת $\underline{g}(\bullet)$ הינה פונקציה ווקטורית אשר במוצאה ווקטור ממימד N . למעשה $\underline{g}(\bullet)$ בנויה מ- N פונקציות שונות, כ"א עבור רכיב אחר בווקטור $\underline{\hat{X}}$, אשר כל אחת מן הפונקציות משתמשת בכל רכיבי הווקטור $\underline{Y}$.

לכן למעשה נוכל לרשום (ולזהות כי המערכת מורכבת מ- N משערכים):

$$\underline{\hat{X}}_{N\times 1}\left(\underline{Y}\right)=\underline{g}\left(\underline{Y}\right)=\begin{bmatrix} g_1\left(\underline{Y}\right) \\ \vdots \\ g_N\left(\underline{Y}\right)\end{bmatrix}=\begin{bmatrix} g_1\left(Y_1,\dots,Y_M\right) \\ \vdots \\ g_N\left(Y_1,\dots,Y_M\right)\end{bmatrix}=\begin{bmatrix} \hat{X}_1\left(Y_1,\dots,Y_M\right) \\ \vdots \\ \hat{X}_N\left(Y_1,\dots,Y_M\right)\end{bmatrix}=\begin{bmatrix} \hat{X}_1\left(\underline{Y}\right) \\ \vdots \\ \hat{X}_N\left(\underline{Y}\right)\end{bmatrix}$$

ווקטור שגיאת שערוך

כעת נגדיר את ווקטור שגיאת השיערוך באופן דומה לזה שהגדרנו עבור מ"א בודד (כרגיל נעבוד עם שגיאת שערוך הפרשית):

$$\underline{e}=\underline{X}-\underline{\hat{X}}=\underline{X}-\underline{g}\left(\underline{Y}\right)$$

גם כעת ווקטור שגיאת השערוך $\underline{e}$ הינו ו"א מאחר ו- $\underline{X}$ וגם $\underline{\hat{X}}$ הוא ו"א (ו- $\underline{e}$ הוא פונקציה שלהם).

מדד עיוות ווקטורי

עבור שערוך מ"א יחיד הגדרנו פונקציית מדד עיוות $d:\mathbb{R}\to\mathbb{R}$ והחלנו עליה מספר דרישות.

עבור שערוך ו"א הפונקציה תתאים עבור ווקטור שגיאת השערוך $\underline{e}$ מספר יחיד, ז"א $d:\mathbb{R}^N\to\mathbb{R}$. פונקצייה $d\left(\underline{e}\right)$ תיקרא מדד עיוות אם $d\left(\underline{e}\right)\geq 0$ לכל $\underline{e}$ וכן $d\left(\underline{0}\right)=0$. נדגיש כי מאחר ו- $\underline{e}$ ו"א אז $d\left(\underline{e}\right)$ הינו מ"א.

<u>מדד עיוות אדיטיבי (חיבורי):</u>

בהינתן מדד עיוות סקלרי $d:\mathbb{R}\to\mathbb{R}$ נאמר שהוא משרה מדד עיוות ווקטורי אדיטיבי ע"י:

$$d\left(\underline{e}\right)=\sum_i d\left(e_i\right)$$

מדדי עיוות אדיטיביים הינם פריקים.
דוגמאות:

1. מדד עיוות סקלרי $d\left(e\right)=e^2$ (מדד MSE) משרה מדד עיוות ווקטורי:

$$d\left(\underline{e}\right)=\sum_i d\left(e_i\right)=\sum_i e_i^2=\left\|\underline{e}\right\|^2$$

2. מדד עיוות סקלרי 'מדד הסתברות השגיאה' (נקרה גם מדד SER) משרה מדד עיוות ווקטורי:

$$d\left(e\right)=\begin{cases}0 & e=0\\ 1 & e\neq 0\end{cases}\quad\Rightarrow\quad d\left(\underline{e}\right)=\sum_i d\left(e_i\right)=\#\left(errors\right)$$

דוגמא למדד עיוות שאינו אדיטיבי:

- נגדיר מדד עיוות בינארי המציג האם קיימת שגיאה כלשהי בשיערוך הווקטור:

$$d_{FER}\left(\underline{e}\right)=\begin{cases}0 & \underline{e}=0\\ 1 & \underline{e}\neq 0\end{cases}$$

מדד כזה מתאים למקרים בהם חשוב לגלות אפילו טעות קטנה בתוך סט מידע (נגיד ביט פגום בקובץ).
נהוג לקרוא לממד זה מדד FER = Frame Error Rate. עבור כל אחת מהדגימות מוגדרת בהתאמה (אנחנו כבר מכירים מדד עיוות זה) מדד SER = Symbol Error Rate:

$$d_{SER}\left(e\right)=\begin{cases}0 & e=0\\ 1 & e\neq 0\end{cases}$$

נציג את $d_{FER}\left(\underline{e}\right)$ ע"י פונקציות ה- $d_{SER}\left(e\right)$ על רכיבי הווקטור השונים:

$$d_{FER}\left(\underline{e}\right)=1-\left(1-d_{SER}\left(e_1\right)\right)\cdot\ldots\cdot\left(1-d_{SER}\left(e_N\right)\right)=1-\prod_{i=1}^{N}\left(1-d_{SER}\left(e_i\right)\right)$$

נבחין כי התצוגה היא של מכפלות וככלל לא אדיטיבית.

<u>עיוות ממוצע/תוחלת העיוות</u>

העיוות הממוצע (או תוחלת העיוות) יחושבו באותו האופן, ז"א ע"י:

$$D=E\left[d\left(\underline{e}\right)\right]=E\left[d\left(\underline{X}-\hat{\underline{X}}\right)\right]=E\left[d\left(\underline{X}-\underline{g}\left(\underline{Y}\right)\right)\right]$$

<u>משערך אופטימלי</u>

<u>טענה:</u>

נבחר את המערכת אשר עבור כל ערך $\underline{Y}$ "מנחשת" את $\underline{\alpha}$ אשר תביא את D למינימום:

$$g_{opt}^{d(\underline{e})}\left(\underline{Y}=\underline{y}\right)=\underline{\alpha}^{*}=\underset{\underline{\alpha}\in\mathbb{R}^N}{\arg}\min\left(E\left[d\left(\underline{X}-\underline{\alpha}\right)\middle|\underline{Y}=\underline{y}\right]\right)$$

<u>הוכחה:</u> באופן זהה כמו עבור שערוך מ"א בודד (בעזרת משפט ההחלקה).

<u>שערוך אופטימלי עבור מדדי עיוות אדיטיביים:</u>

אם נצטמצם למדדי-עיוות אדיטיביים (ולכן פריקים) בלבד נוכל להציג את תוחלת העיוות לפי:

$$D=E\left[d\left(\underline{e}\right)\right]=E\left[\sum_i d\left(e_i\right)\right]\overset{*}{=}\sum_i E\left[d\left(e_i\right)\right]\overset{D_i=E\left[d\left(e_i\right)\right]}{=}\sum_i D_i$$

כאשר שיוויון $(*)$ מתקיים מאחר ותוחלת (כמו אינטגרל) היא פעולה ליניארית לחיבור (קומוטטיביות).

הערה - גם עבור $d\left(\underline{e}\right)$ כיפלי נוכל לבצע אנליזה דומה ע"י הפרדה וזאת בעזרת הפעלת $\log$.

<u>מסקנה:</u>

על מנת להביא את למינימום את D נוכל להביא למינימום כל D_i בנפרד (כי על מנת להביא סכום של איברים לא תלויים למינימום נביא כל אחד מהם למינימים בנפרד). לכן נוכל לכתוב את פונקציית השערוך של כל קואורדינטה באופן הבא:

$$g_{opt,i}^{d(\underline{e})}\left(\underline{Y}=\underline{y}\right)=\alpha^{*}=\underset{\alpha}{\arg}\min\left(E\left[d\left(X_i-\alpha\right)\middle|\underline{Y}=\underline{y}\right]\right)$$

ולכן המשערך כולו ייראה כך:

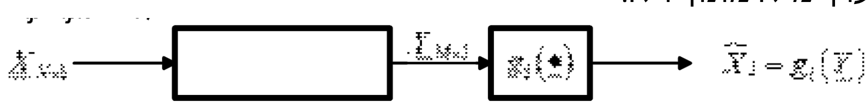
$$\hat{\underline{X}}_{opt}^{d(\underline{e})}\left(\underline{Y}=\underline{y}\right)=\underline{g}_{opt}^{d(\underline{e})}\left(\underline{Y}=\underline{y}\right)=\begin{bmatrix} g_{opt,1}^{d(\underline{e})}\left(\underline{Y}=\underline{y}\right)\\ \vdots \\ g_{opt,N}^{d(\underline{e})}\left(\underline{Y}=\underline{y}\right)\end{bmatrix}$$

<u>מסקנה:</u>

ניתן עתה להספק בתרחיש של שערוך מ"א מתוך ו"א כי שערוך ו"א מתוך ו"א מתפרק ל- N בעיות שערוך מ"א מתוך ו"א כאשר מדד העיוות הוא פריק.

<u>שערוך משתנה אקראי מתוך ווקטור אקראי</u>

כעת נרצה לשערך מ"א מתוך ו"א:



נדון במקרה זה רק עבור מדד עיוות פריקים וכפי שראינו (כרגיל שגיאת שערוך הפרשית):

$$D_i=E\left[d\left(e_i\right)\right]=E\left[d\left(X_i-\hat{X}_i\right)\right]=E\left[d\left(X_i-g_i\left(\underline{Y}\right)\right)\right]$$

כאשר ראינו ששערוך אופטימלי (שוב – עבור מדד עיוות פריק) מבוצע ע"י המערכת:

$$\hat{X}^{d(\underline{e})}_{opt,i}\left(\underline{Y}=\underline{y}\right)=g^{d(\underline{e})}_{opt,i}\left(\underline{Y}=\underline{y}\right)=\alpha^*=\arg\min_{\alpha}\left(E\left[d\left(X_i-\alpha\right)\middle|\underline{Y}=\underline{y}\right]\right)$$

לכן למעשה ל- $\hat{X}_i$ יש תלות ב- X_i וב- $\underline{Y}$ בלבד ולא בשאר רכיבי הווקטורים $\underline{X}$ ו- $\hat{\underline{X}}$.

<u>שערוך מ"א מתוך ו"א באופן אופטימלי תחת מדד הסתברות השגיאה</u>

עבור מדד עיוות של הסתברות שגיאה (מסומן Pe או SER) נרצה לתכנן את המערכת אשר תשערך מ"א מתוך ו"א באופן אופטימלי, ז"א תביא את תוחלת העיוות למינימום.

$$d_{Pe}\left(e\right)=\begin{cases}0 & e=0\\ 1 & e\neq 0\end{cases},\qquad D=E\left[d_{Pe}\left(e\right)\right]\overset{*}{=}\Pr\left(\hat{X}\neq X\right)$$

כאשר את מעבר שיוויון $(*)$ ראינו בשערוך אופטימלי של מ"א מתוך מ"א תחת מדד הסתברות שגיאה.

טענה - המשערך האופטימלי של מ"א מתוך ו"א תחת מדד הסתברות השגיאה הוא:

$$\hat{X}^{Pe}_{opt}\left(\underline{Y}=\underline{y}\right)=g^{Pe}_{opt}\left(\underline{Y}=\underline{y}\right)=\arg\max_{\alpha}\Pr\left(X=\alpha\middle|\underline{Y}=\underline{y}\right)$$

זו למעשה בחירה לפי כלל MAP (maximum a posteriori probability) מאחר שמחליטים על הערך המשוערך לאחר "ראיית" $\underline{Y}$.

הוכחה - זהה למקרה של שערוך מ"א מתוך מ"א (ע"י פיתוח התוחלת המותנית $E\left[d\left(X-\alpha\right)\middle|\underline{Y}=\underline{y}\right]$).

<u>שערוך מ"א מתוך ו"א באופן אופטימלי תחת מדד MSE</u>

ראינו כי מדד עיוות שגיאה ריבועית $d\left(e\right)=e^2$ (מדד MSE) משרה מדד עיוות ווקטורי $d\left(\underline{e}\right)=\left\|\underline{e}\right\|^2$.

נרצה לתכנן את המערכת אשר תשערך מ"א מתוך ו"א באופן אופטימלי, ז"א תביא את תוחלת העיוות למינימום.

<u>שערוך אופטימלי של מ"א מתוך ו"א במובן MMSE:</u>

טענה - המשערך האופטימלי של מ"א מתוך ו"א תחת מדד MSE הוא:

$$\hat{X}_{MMSE}\left(\underline{Y}=\underline{y}\right)=\hat{X}_{opt}^{MSE}\left(\underline{Y}=\underline{y}\right)=g_{opt}^{MSE}\left(\underline{Y}=\underline{y}\right)=E\left[X\middle|\underline{Y}=\underline{y}\right]$$

זו למעשה בחירה לפי כלל MAP (maximum a posteriori probability) מאחר שמחליטים על הערך המשוערך לאחר "ראיית" $\underline{Y}$.

הוכחה - זהה למקרה של שערוך מ"א מתוך מ"א (ע"י פיתוח התוחלת המותנית $E\left[d\left(X-\alpha\right)\middle|\underline{Y}=\underline{y}\right]$).

<u>שערוך אופטימלי של מ"א מתוך ו"א במובן LMMSE:</u>

כעת נרצה להגביל את המשערך להיות פונקציה ליניארית של $\underline{Y}$, ז"א המשערך יהיה מהצורה:

$$\hat{X}\left(\underline{Y}\right)=\underline{a}^T\cdot\underline{Y}+b=\sum_{i=1}^{M}a_i\cdot Y_i+b$$

שגיאת השערוך תהיה:

$$e=X-\hat{X}\left(\underline{Y}\right)=X-\underline{a}^T\cdot\underline{Y}-b$$

נרצה למצוא $\left(\underline{a},b\right)$ שיביאו את תוחלת העיוות למינימום, ז"א יקיימו $D=E\left[d\left(e\right)\right]=E\left[e^2\right]\to\min$.

טענה - המשערך הליניארי האופטימלי של מ"א מתוך ו"א תחת מדד MSE הוא:

$$\hat{X}_{LMMSE}\left(\underline{Y}\right)=\hat{X}_{opt}^{LinearMSE}\left(\underline{Y}\right)=g_{opt}^{LinearMSE}\left(\underline{Y}\right)=\eta_X+C_{X\underline{Y}}C_{\underline{Y}\underline{Y}}^{-1}\left(\underline{Y}-\eta_{\underline{Y}}\right)$$

ועבור השגיאה מתקיים:

$$E\left[e_{LMMSE}^2\right]=\sigma_X^2-C_{X\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\cdot C_{\underline{Y}X}$$

הוכחה - הוכחת נוסחת המשערך נמצאת בהוכחה תכונת הניצבות עבור משערך LMMSE של מ"א מתוך ו"א. הוכחת הביטוי לתוחלת ריבוע השגיאה מבוצע באופן דומה להוכחה עבור משערך LMMSE של מ"א מתוך מ"א.

תכונות משערכי MMSE\LMMSE עבור שערוך מ"א מתוך ו"א:

1. תכונת ניצבות/אורתוגונליות
   a. משערך MMSE

משערך $\hat{X}(\underline{Y}) = g(\underline{Y})$ הינו MMSE אם"ם שגיאתו $e = X - g(\underline{Y})$ מקיימת $e \perp h(\underline{Y})$ לכל פונקציה $h(\underline{Y})$. למעשה השגיאה ניצבת לכל פונקציה של המדידות. ז"א $E\left[e \cdot h(\underline{Y})\right] = 0 \quad \forall h : \mathbb{R}^M \to \mathbb{R}$.

   b. משערך LMMSE

משערך ליניארי $\hat{X}(\underline{Y}) = \underline{a}^T \cdot \underline{Y} + b$ הינו LMMSE אם"ם שגיאתו $e = X - g(\underline{Y}) = X - \left(\underline{a}^T \cdot \underline{Y} + b\right)$ מקיימת $e \perp h(\underline{Y})$ לכל פונקציה ליניארית $h(\underline{Y})$ (או למעשה נדרוש כי $e \perp \left(\underline{c}^T \cdot \underline{Y} + d\right) \quad \forall \underline{c} \in \mathbb{R}^M, d \in \mathbb{R}$). למעשה השגיאה ניצבת לכל פונקציה ליניארית של המדידות. ז"א

$$E\left[e \cdot \left(\underline{c}^T \cdot \underline{Y} + d\right)\right] = E\left[\left(X - \left(\underline{a}^T \cdot \underline{Y} + b\right)\right) \cdot \left(\underline{c}^T \cdot \underline{Y} + d\right)\right] = 0 \quad \forall \underline{c} \in \mathbb{R}^M, d \in \mathbb{R}.$$

2. חוסר הטייה (Un-Biasedness)
משערך L\MMSE מקיים כי תוחלת השגיאה מתאפסת, למעשה הוא מקיים כי תוחלתו של המשערך שווה לתוחלתו של המשתנה המשוערך X.

$$E\left[e_{L\backslash MMSE}\right] = 0$$
$$E\left[X - \hat{X}_{L\backslash MMSE}\right] = 0$$
$$E\left[\hat{X}_{L\backslash MMSE}\right] = E[X] = \eta_X$$

3. פיתגורס (Pythagoras)
הספק המ"א אשר אנו רוצים לשערך שווה להספק המשערך של המ"א פלוס הספק השגיאה (כאמור עבור מדד MSE הספק השגיאה מקיים $D = MSE$).

$$E\left[X^2\right] = E\left[\hat{X}^2_{L\backslash MMSE}\right] + \left[e^2_{L\backslash MMSE}\right]$$

נבחין כי אם נתעסק עם משערך לינארי (נגביל את עצמנו) אז הספק השגיאה יגדל על חשבון הספק המשערך.

4. עבור וא"ג $\begin{bmatrix} X \\ \underline{Y} \end{bmatrix}$ מתקיים $\hat{X}_{MMSE} = \hat{X}_{LMMSE}$, ז"א משערך LMMSE הינו אופטימלי

אם $\begin{bmatrix} X \\ \underline{Y} \end{bmatrix}$ הוא וא"ג אז המשערך האופטימלי במובן MMSE הינו ליניארי ולמעשה מתקיים:

$$\hat{X}_{MMSE} = E\left[X \middle| \underline{Y} = \underline{y}\right] \overset{*}{=} \eta_X + C_{X\underline{Y}} \cdot C_{\underline{Y}\underline{Y}}^{-1} \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right) = \hat{X}_{LMMSE}$$

כאשר מעבר $(*)$ מתקיים כאשר $\begin{bmatrix} X \\ \underline{Y} \end{bmatrix}$ הוא וא"ג.

<u>הוכחות:</u>

נוכיח רק עבור קיום של משערך ליניארי אופטימלי המקיים את תכונת האורתוגנליות (ניצבות) משום שכל שאר ההוכחות מבוצעות באופן דומה להוכחות של תכונות משערכי L\MMSE אופטימליים של מ"א מתוך מ"א.

נרצה למצוא את $(\underline{a}, b)$ כך שתתקיים תכונת הניצבות:

$$\left(X - \left(\underline{a}^T \cdot \underline{Y} + b\right)\right) \perp \left(\underline{c}^T \cdot \underline{Y} + d\right) \quad \forall \underline{c} \in \mathbb{R}^M, d \in \mathbb{R}.$$

טענה - להראות ניצבות כזאת שקולה לשתי הדרישות הבאות:

i. $e \perp 1$ (או כל קבוע אחר מ- $\mathbb{R}$)

ii. $$E\left[e \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right)^T\right] = \underline{0} \quad \Leftrightarrow \quad \begin{matrix} e \perp \left(Y_i - \eta_{Y_i}\right) \\ \forall i = 1, \ldots, M \end{matrix} \quad \Leftrightarrow \quad \begin{cases} e \perp \left(Y_1 - \eta_{Y_1}\right) \\ \vdots \\ e \perp \left(Y_M - \eta_{Y_M}\right) \end{cases}$$

כעת נראה מה משמעות זוג דרישות אלו:

$$\begin{aligned} i \quad & e \perp 1 \Leftrightarrow E\left[e \cdot 1\right] = 0 \\ & E\left[X - \underline{a}^T \cdot \underline{Y} - b\right] = 0 \\ & \eta_X - \underline{a}^T \cdot \eta_{\underline{Y}} - b = 0 \\ & b\left(\underline{a}\right) = \eta_X - \underline{a}^T \cdot \eta_{\underline{Y}} \end{aligned}$$

$$\begin{aligned} ii \quad & E\left[e \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right)^T\right] = \underline{0} \\ & E\left[\left(X - \underline{a}^T \cdot \underline{Y} - b\right) \cdot \left(\underline{Y} - \eta_{\underline{Y}}\right)^T\right] = \underline{0} \end{aligned}$$

נציב את $b(\underline{a})$ ונמשיך:

$$ii \quad E\left[\left(X-\underline{a}^T\cdot\underline{Y}-\eta_X+\underline{a}^T\cdot\eta_{\underline{Y}}\right)\cdot\left(\underline{Y}-\eta_{\underline{Y}}\right)^T\right]=0$$
$$E\left[\left(\left(X-\eta_X\right)-\underline{a}^T\cdot\left(\underline{Y}-\eta_{\underline{Y}}\right)\right)\cdot\left(\underline{Y}-\eta_{\underline{Y}}\right)^T\right]=0$$
$$C_{X\underline{Y}}-\underline{a}^T\cdot C_{\underline{Y}\underline{Y}}=0 \quad\Rightarrow\quad \underline{a}^T=C_{X\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}$$

וכעת "נרכיב" את המשערך שמצאנו (ונקבל שזה משערך $\hat{X}_{LMMSE}$ כפי שהגדרנו):

$$\hat{X}\left(\underline{Y}\right)=\underline{a}^T\cdot\underline{Y}+b=\begin{bmatrix}\underline{a}^T=C_{X\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\\ b\left(\underline{a}\right)=\eta_X-\underline{a}^T\cdot\eta_{\underline{Y}}\end{bmatrix}=C_{X\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\cdot Y+\eta_X-C_{X\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\cdot\eta_{\underline{Y}}=$$
$$\eta_{\underline{X}}+C_{\underline{X}\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\cdot\left(\underline{Y}-\eta_{\underline{Y}}\right)$$

<u>שערוך אופטימלי במובן MSE של ו"א מתוך ו"א</u>

נחזיר את הדיון לשערוך ו"א מתוך ו"א כאשר נצטמצם לשגיאה הפרשית עם מדד עיוות MSE.

<u>משערך MMSE:</u>

$$\underline{\hat{X}}\left(\underline{Y}=\underline{y}\right)=E\left[\underline{X}\middle|\underline{Y}=\underline{y}\right]=\begin{bmatrix}E\left[X_1\middle|\underline{Y}=\underline{y}\right]\\ \vdots\\ E\left[X_N\middle|\underline{Y}=\underline{y}\right]\end{bmatrix}$$

<u>משערך LMMSE:</u>

$$\underline{\hat{X}}_{LMMSE}\left(\underline{Y}\right)=\underline{\hat{X}}_{opt}^{LinearMSE}\left(\underline{Y}\right)=\underline{g}_{opt}^{LinearMSE}\left(\underline{Y}\right)=\eta_{\underline{X}}+C_{\underline{X}\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\cdot\left(\underline{Y}-\eta_{\underline{Y}}\right)$$
$$C_{\underline{e}\underline{e}}=C_{\underline{X}\underline{X}}-C_{\underline{X}\underline{Y}}\cdot C_{\underline{Y}\underline{Y}}^{-1}\cdot C_{\underline{Y}\underline{X}}$$

תכונות של מטריצת הקו-ואריאנס של השגיאה $C_{\underline{e}\underline{e}}$:

1. יהי משערך מתחרה $\underline{\hat{X}'}=\underline{h}\left(\underline{Y}\right)$ עם שגיאה $\underline{e}'=\underline{X}-\underline{\hat{X}'}$ אזי $C_{\underline{e}'\underline{e}'}-C_{\underline{e}_{LMMSE}\underline{e}_{LMMSE}}\geq 0$.
2. יהי משערך ליניארי מתחרה $\underline{\hat{X}''}=\underline{c}^T\cdot\underline{Y}+\underline{d}$ עם שגיאה $\underline{e}''=\underline{X}-\underline{\hat{X}''}$ אזי המטריצה $C_{\underline{e}''\underline{e}''}-C_{\underline{e}_{LMMSE}\underline{e}_{LMMSE}}$ היא PSD (מטריצה אי-שלילית מוגדרת).

הערה - מטריצת הקו-ואריאנס של ווקטור השגיאה מוגדרת ע"י $C_{\underline{e}_{L\backslash MMSE}\underline{e}_{L\backslash MMSE}}=C_{\underline{X}\underline{X}}-C_{\underline{\hat{X}}_{L\backslash MMSE}\underline{\hat{X}}_{L\backslash MMSE}}$.

חלק ב': תהליכים אקראיים ופעולות עליהם

מבוא, הגדרות ותכונות

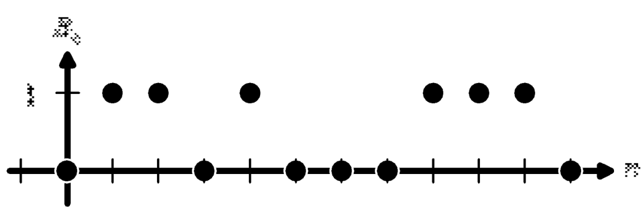
תהליכים אקראיים הם הכללה של ו"א, תהליכים מתרחשים בזמן/במרחב ולכן אינדקס המ"א מקבל משמעות מתאימה ובנוסף יש חשיבות לסדר המשתנים האקראיים (האינדקסים מייצגים את התפתחות התהליך).
אנו נתעסק עם תהליכים אקראיים אשר מתפתחים בזמן. תהליך אקראי (ת"א) יכול להיות בזמן רציף או בדיד ולקבל ערכים רציפים או בדידים.
אם נחשוב על ת"א כעל ו"א אינסופי אז:

- ת"א בזמן רציף יהיה ב"גודל" אינסופי של עוצמת-הרצף (א).
- ת"א בזמן בדיד יהיה ב"גודל" אינסופי בן-מניה ($\aleph_0$).

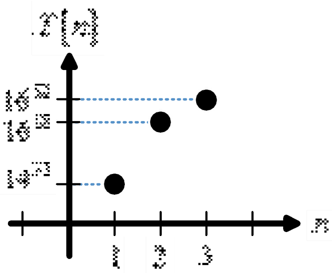
מאחר ותהליכים אקראיים הינם אינסופיים לעיתים במקום לסמנם $\{X_n\}_{n=1}^{\infty}$ (או תחום אינסופי אחר) עבור זמן בדיד נסמן בפשטות X_n (אך אין הכוונה לדגימה בודדת עבור n מסויים) ובמקום לסמנם $\{X(t)\}, t\in[0,\infty)$ (או תחום אינסופי אחר) עבור זמן רציף נסמן בפשטות $X(t)$ (אך אין הכוונה לדגימה בודדת עבור t מסויים).

<u>דוגמאות:</u>

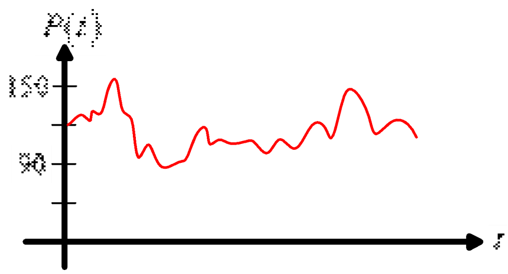
1. סדרת הביטים המתקבלים בהורדת קובץ - $B[n]$
זהו ת"א בדיד בזמן בדיד – הוא מקבל ערכים בדידים (0 או 1) כל הפרש זמן ידוע.



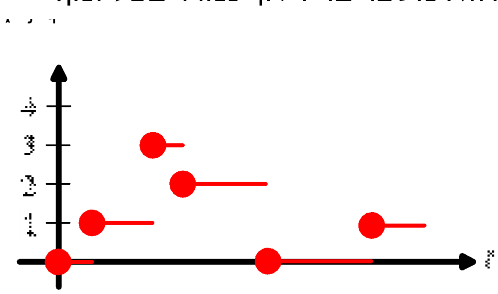
2. זמן הגעת האוטובוס ה-n-י לתחנה - $T[n]$
זהו ת"א רציף ב"זמן" בדיד – זמן ההגעה הינו פרמטר רציף אך ציר המקור מייצג את מספר האוטובוס.



3. מדידת לחץ דם לזמן - $P(t)$
זהו ת"א רציף בזמן רציף – לחץ הדם יכול לקבל כל ערך (בטווח מסויים) והוא נמדד כל הזמן.



4. מספר האנשים בתחנה כפונקציה של הזמן - $N(t)$
זהו ת"א בדיד בזמן רציף – מספר האנשים בתחנה הוא מספר בדיד אך נמדד בכל זמן.



__תהליך אקראי בזמן רציף__

ת"א בזמן רציף $X(\omega,t)$ הוא פונקציה המשייכת לכל זמן רציף t ותוצאת ניסוי $\omega\in\Omega$ מספר ממשי:

$$X:\Omega\times\mathbb{R}\to\mathbb{R}$$

הערה - ייצוג אחר הוא $X:\Omega\to\mathbb{R}^{\mathbb{R}}$, ז"א עבור כל תוצאת ניסוי $\omega_0\in\Omega$ נקבל אינסוף זוגות (t,x) כך ש- $t,x\in\mathbb{R}$. עבור כל ערך t_0 יש את הזוג המתאים עם הערך x_0 אשר מתאים ל- $X(\omega_0,t_0)$.

כאמור, נסמן את התהליך בפשטות ע"י $X(t)$.

<u>תהליך אקראי בזמן בדיד</u>

ת"א בזמן בדיד $X[\omega,n]$ הוא פונקציה המשייכת לכל זמן בדיד n ותוצאת ניסוי $\omega\in\Omega$ מספר ממשי:

$$X:\Omega\times\mathbb{Z}\to\mathbb{R}\quad\leftrightarrow\quad X:\Omega\to\mathbb{R}^{\mathbb{Z}}$$

כאמור, נסמן את התהליך בפשטות ע"י X_n .

<u>פונקציית מדגם (ריאליזציה) של תהליך אקראי</u>

פונקציית מדגם (ריאליזציה – Realization) של ת"א היא תוצאות דגימת התהליך האקראי עבור $\omega=\omega_0$ תוצאת ניסוי מסויימת. כמובן שעבור תוצאת ניסוי ω_0 זו התהליך מקבל ערך בכל נקודת זמן. ז"א עבור $\omega=\omega_0$ (ז"א הקפה של ציר ω) נקבל פונקצייה $X(\omega_0,t)$ שהיא מיפוי של תוצאת ניסוי מסויימת בהתאם לזמן.

אם בנוסף נבחר $t=t_0$ אז נקבל מ"א $X(\omega_0,t_0)$ (זהו מ"א מאחר ש- X הוא ת"א ולא דטרמיניסטי).

מוטיבציה לתהליכים אקראיים

לצורך הבנת יצירה של תהליכים אקראיים נראה מספר דוגמאות פשוטות. במהלך הקורס נבנה ת"א לא טריוויאלים ונצטרך לזהות ולפרקם לתהליכים אקראיים פשוטים יותר.

<u>תהליך אקראי i.i.d</u>

תהליך אקראי בדיד (לצורך ההמחשה, אפשר גם רציף) בזמן X_n יקרא ת"א i.i.d אם מתקיימים התנאים הבאים:

1. כל דגימות התהליך בעלות אותו פילוג, ז"א
$\Pr\left(X_i = x\right) = \Pr\left(X_j = x\right) \quad \forall i \in \mathbb{Z}, j \in \mathbb{Z}, x \in \mathbb{R}$ (או עם פונקציית ה-PDF עבור המקרה של זמן רציף).
2. דגימות התהליך $X_1, X_2, \dots$ בת"ס באופן הדדי (ולא רק בזוגות).
נסמן אי-תלות סטטיסטית של זוג משתנים ע"י $X \perp\!\!\!\perp Y$.
ולכן למעשה אנחנו דורשים $\forall n \in \mathbb{Z} \quad X_1, X_2, \dots, X_{n-1} \perp\!\!\!\perp X_n$.
הערה – אי תלות בזוגות אינה גורת אי תלות הדדית.

<u>דוגמאות קונסטרוקטיביות פשוטות של תהליכים אקראיים</u>

1. <u>ת"א i.i.d:</u>
ת"א i.i.d הוא תהליך שכל דגימותיו הינן i.i.d. לצורך הדגמה נהיה יותר ספציפית ונראה ת"א i.i.d עם דגימות ברנולי עם פרמטר הצלחה p .

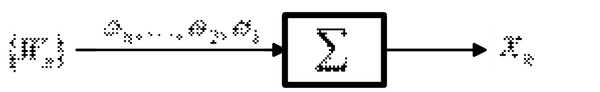
$$W_n = \begin{cases} 1 & p \\ 0 & q = 1 - p \end{cases} \sim Ber\left(p\right), \qquad \left\{W_n\right\} \quad i.i.d$$

נרצה לבחון את הסטטיסטיקה של התהליך עבור k נקודות זמן: $i_1, \dots, i_k$.
הדגימות i.i.d ולכן בת"ס, לכן נוכל לדעת סטטיסטיקה מלאה של הת"א לכל סט זמנים כזה לפי:

$$\Pr\left(W_{i_1} = \omega_1, \dots, W_{i_k} = \omega_k\right) \overset{\text{Independence}}{=} \prod_{j=1}^{k} \Pr\left(W_{i_j} = \omega_j\right)$$

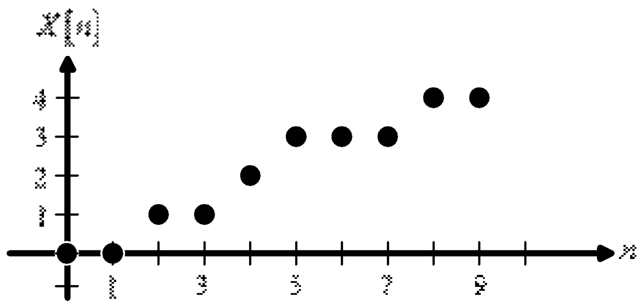
2. <u>ת"א של מנייה:</u>

ניקח את ת"א ברנולי $\left\{W_n\right\}$ ה-i.i.d מקודם ונזין אותה לסוכם:



$$X_n = \sum_{i=1}^{n} W_i$$

פונקציית מדגם אפשרית של התהליך $\left\{X_n\right\}$ תראה כך:



נבחין כי נוכל להציג את התהליך באופן רקורסיבי:

$$X_n = X_{n-1} + W_n \qquad W_n = \begin{cases} 1 & p \\ 0 & q = 1-p \end{cases}, \{W_n\} \quad i.i.d$$
$$X_0 = 0$$

למעשה זוהי הגדרה של ת"א אוטו-רגרסיבי (Auto Regressive Process).

ת"א $\{X_n\}$ אינו i.i.d על אף שהוא נבנה מת"א i.i.d. יתרה מזאת, גם בלי התייחסות לקשר בין הדגימות, דגימותיו כלל לא מתפלגות באופן זהה. אין ספק שיש קשר בין הדגימות הרי זוג דגימות סמוכות בהכרח נבדלות בלכל היותר אחד והסדרה עולה (אך לא עולה ממש).

למעשה, זהו תהליך מנייה של דגימות ברנולי ולכן זה מ"א בינומי $X_n \sim Bin(n,p)$.

נראה:

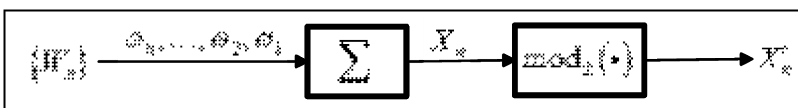
$$E[X_n] = E\left[\sum_{i=1}^{n} W_i\right] = \sum_{i=1}^{n} E[W_i] = \sum_{i=1}^{n} p = n \cdot p$$
$$Var(X_n) = E\left[X_n^2\right] - E^2[X_n] = \ldots = n \cdot p \cdot (1-p)$$

אנו מבחינים כי זהו ת"א עם הרבה זיכרון – כל דגימה מתפלגת שונה ויש הרבה תלות בין הדגימות.

נבחין שאם $p = 0,1$ אז השונות מתאפסת – הדבר מובן כי אז התהליך דטרמיניסטי.

3. <u>ת"א XOR:</u>

על אותו ת"א ברנולי $\{W_n\}$ ה-i.i.d נבנה את הת"א הבא:

	$X'_n = \mathrm{mod}_2\left(\sum_{i=1}^{n} W_i\right)$

פונקציית מדגם אפשרית של התהליך $\{X'_n\}$ תראה כך:

שרטוט אפסים או אחדים

נבחין כי נוכל להציג את התהליך באופן רקורסיבי:

$$X'_n = X'_{n-1} \oplus W_n \qquad W_n = \begin{cases} 1 & p \\ 0 & q = 1-p \end{cases}, \{W_n\} \quad i.i.d \qquad A \oplus B = \begin{cases} 1 & A \neq B \\ 0 & A = B \end{cases}$$
$$X'_0 = 0$$

ז"א אם $W_n = 1$ אז ערכו של X'_n יהיה הפוך מקודמו, אחרת זהה (בכל אופן $X'_n \in \{0,1\}$).

עבור $p = 0.5$ נקבל כי דגימותיו של ת"א זה הן i.i.d (כלומר X'_n יקבל ערך 0 או 1 בהסתברות שווה ללא תלות בדגימות אחרות שלו).

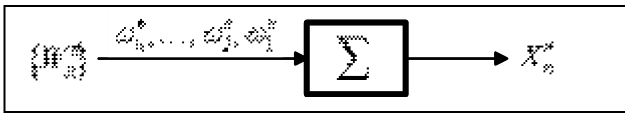
4. <u>ת"א הילוך שיכור:</u>

נרצה ליצור ת"א אשר מדמה תנועה לאחד משני כיוונים אפשריים בכל נקודת זמן. הדבר שונה מן התפלגות ברנולי כי שם או שמתקיים צעד ($W_n = 1$) או שלא ($W_n = 0$).

לשם כך נגדיר מת"א ברנולי $\{W_n\}$ ה-i.i.d את הת"א הבא:

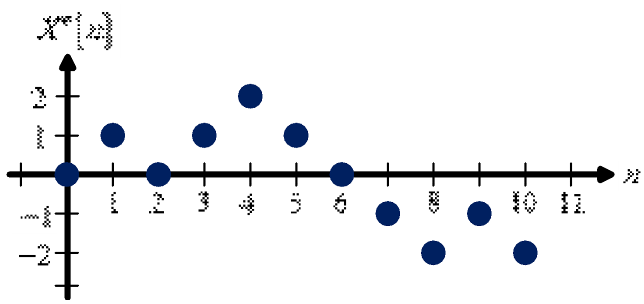
$$W''_n = 2 \cdot W_n - 1 = \begin{cases} 1 & p \\ -1 & q = 1-p \end{cases} \qquad W_n \sim Ber(p), \{W_n\} \quad i.i.d$$

כעת נבצע לת"א $\{W''_n\}$ תהליך סכימה:



$$X''_n = \sum_{i=1}^{n} W''_i$$

פונקציית מדגם אפשרית של התהליך $\{X''_n\}$ תראה כך:



נבחין כי נוכל להציג את התהליך באופן רקורסיבי:

$$X''_n = X''_{n-1} + X''_n \qquad W''_n = \begin{cases} 1 & p \\ -1 & 1-p \end{cases}, \{W''_n\} \quad i.i.d$$
$$X''_0 = 0$$

ננתח את הסטטיסטיקה של הת"א $\{X''_n\}$:

$$E[X''_n] = E\left[\sum_{i=1}^{n} W''_i\right] = \sum_{i=1}^{n} E[W''_i] = \sum_{i=1}^{n} E[2 \cdot W_n - 1] = \sum_{i=1}^{n} (2 \cdot p - 1) = n \cdot (2 \cdot p - 1)$$

$$Var(X''_n) \overset{*}{=} 4 \cdot Var(X_n) = \ldots = 4 \cdot n \cdot p \cdot (1-p)$$

כאשר שיוויון $(*)$ נובע מתכונות השונות ($Var\left(a\cdot X+b\right)=a^2\cdot Var\left(X\right)$).

עבור $p=0.5$ נקבל $E\left[X_n''\right]=0$, הדבר הגיוני כי ה"שיכור" נע לכל אחד מן הכיוונים בסיכוי זהה.

הערה - ת"א זה דומה לתנועת בראון (Brownian Motion), נדון בהמשך הקורס בת"א מתקדמים.

<u>מידע סטטיסטי מלא על תהליך אקראי</u>

עבור ווקטור אקראי $\underline{X}$ ידיעת הסטטיסטיקה של הווקטור האופן מלא הייתה שקולה לידיעת פונקציית ה-jCDF שלו (או ידיעת פונקציית ה-jPDF שלו), ז"א ידיעת פונקציית פילוג משותפת שלו:

$$F_{\underline{X}}\left(\underline{x}\right)=\Pr\left(X_1\le x_1,\ldots,X_n\le x_n\right)=\Pr\left(\underline{X}\le\underline{x}\right)$$

עבור תהליך אקראי גודל הווקטור הינו אינסופי ולכן ניתקל בבעיית רישום ואף חישוב או אחסון של הפונקצייה.

לכן נתאר ת"א מבחינה סטטיסטית ע"י פונקציית ה-jCDF (או פונקציית ה-jPDF) לכל אוסף נקודות זמן.

ז"א נבחר אוסף סופי של נקודות זמן מתוך הציר האינסופי ואז נצטמצם לו"א ונרצה לדעת להציג את ה-jCDF.

עבור ת"א בזמן רציף נרצה לדעת לתת את פונקציית ה-jCDF עבור כל סט זמן $\left\{t_1,\ldots,t_N\right\}$:

$$F\left(t_1,\ldots,t_N,x_1,\ldots,x_N\right)=F_{X(t_1),\ldots,X(t_N)}\left(x_1,\ldots,x_N\right)=F_{t_1,\ldots,t_N}\left(\underline{x}\right)=\Pr\left(X\left(t_1\right)\le x_1,\ldots,X\left(t_N\right)\le x_N\right)$$

עבור ת"א בזמן בדיד נרצה לדעת לתת את פונקציית ה-jCDF עבור כל סט זמן $\left\{k_1,\ldots,k_N\right\}$:

$$F\left(k_1,\ldots,k_N,x_1,\ldots,x_N\right)=F_{X(k_1),\ldots,X(k_N)}\left(x_1,\ldots,x_N\right)=F_{k_1,\ldots,k_N}\left(\underline{x}\right)=\Pr\left(X\left(k_1\right)\le x_1,\ldots,X\left(k_N\right)\le x_N\right)$$

נהוג גם לסמן $X\left(k_i\right)=X_{k_i}$, $X\left(t_i\right)=X_{t_i}$ ואז הסימונים קלים יותר.

נחזור ונדגיש כי מידע סטטיסטי מלא על ת"א משמעותו ידיעת הפונקציה הזאת עבור כל סט זמנים אפשרי כלשהו.

<u>הערות:</u>

- בגלל שאנו רוצים לדעת את ה-jCDF (או את ה-jPDF) עבור כל סט זמנים עדיין נתקשה לרשום את הביטוי לפונקצייה אלא אם יש לנו מידע על אופי התהליך (לדוגמא גאוסי, i.i.d וכדומה).
- פונקציות ה-jCDF לסטים שונים של זמן חייבות לקיימת את דרישת הקונסיסטנטיות (עקביות).

נבחין כי אם ניקח את שני הסטים $\{t_1,t_2\},\{t_2,t_3\}$ אז נקבל את שתי פונקציות ה-jCDF הבאות (לאחר שנסמן לשם פשטות $X_{t_i}=X_i$):

$$F_{X_1,X_2}\left(x_1,x_2\right),F_{X_2,X_3}\left(x_2,x_3\right)$$

לפי תכונות של jCDF מתקיים:

$$F_{X_2}\left(x\right)=F_{X_1,X_2}\left(\infty,x\right),F_{X_2}\left(x\right)=F_{X_2,X_3}\left(x,\infty\right)$$

משמעות הדבר שיש קשר בין פונקציות ה-jCDF של ו"א הנוצרים מסטים של זמן עם חפיפה. לכן עבור כל שני סטים של זמן עם חפיפה כלשהי יש דרישה שיש לקיים דבר היוצר קושי בהגדרה.

תהליך אקראי גאוסי בזמן רציף

ת"א $X\left(t\right)$ ייקרא ת"א גאוסי (תא"ג) בזמן רציף אם לכל קבוצת דגימות שלו (לכל סט זמן שייבחר) מתקיים כי קבוצה זו היא גאוסית במשותף (ז"א הדגימות מרכיבות וא"ג).

ז"א עבור כל סט זמנים $\{t_1,\ldots,t_N\}$ דגימות התהליך $\{X\left(t_1\right),\ldots,X\left(t_N\right)\}$ מקיימות כי כל קומבינצייה ליניארית שלהן היא מא"ג, כלומר $Y=\sum_{i=1}^{N}\alpha_i\cdot X\left(t_i\right)$ הוא מא"ג עבור כל $\{\alpha_i\}_{i=1}^{N},\quad \alpha_i\in\mathbb{R}$.

כאמור, על מנת לתאר ת"א באופן סטטיסטי מלא אנחנו צריכים לדעת עליו מידע מקדים.

עבור תא"ג $X\left(t\right)$ נוכל לרשום את פונקציית ה-jCDF עבור כל סט זמנים $\{t_1,\ldots,t_N\}$ ע"י:

$$F_{X(t_1),\ldots,X(t_N)}\left(x_1,\ldots,x_N\right)=F_{t_1,\ldots,t_N}\left(\underline{x}\right)=\frac{1}{\sqrt{\left(2\pi\right)^N\cdot C_{\underline{t}\underline{t}}}}\cdot\exp\left[-\frac{1}{2}\cdot\left(\underline{x}-\eta_{\underline{t}}\right)^T\cdot C_{\underline{t}\underline{t}}^{-1}\cdot\left(\underline{x}-\eta_{\underline{t}}\right)\right]$$

כאשר בסימונים $\left(\eta_{\underline{t}},C_{\underline{t}\underline{t}}\right)$ הכוונה היא ל- $\left(\eta_{\underline{X}_t},C_{\underline{X}_t\underline{X}_t}\right)$ והם מסומנים כך לצורך פשטות.

על מטריצת $C_{\underline{t}\underline{t}}=C_{\underline{X}_t\underline{X}_t}$ לקיים את דרישת הקונסיסטנטיות. מטריצה זו היא פשוט מטריצת הקו-וריאנס של דגימות התהליך ומחושבת ע"י:

$$C_{\underline{X}_t\underline{X}_t}=\begin{bmatrix}\sigma^2_{X(t_1)} & \cdots & Cov\left(X\left(t_1\right),X\left(t_N\right)\right)\\ \vdots & \ddots & \vdots\\ Cov\left(X\left(t_N\right),X\left(t_1\right)\right) & \cdots & \sigma^2_{X(t_N)}\end{bmatrix}=\begin{bmatrix}\sigma^2_{X_{t_1}} & \cdots & Cov\left(X_{t_1},X_{t_N}\right)\\ \vdots & \ddots & \vdots\\ Cov\left(X_{t_N},X_{t_1}\right) & \cdots & \sigma^2_{X_{t_N}}\end{bmatrix}$$

<u>הערה/תזכורת:</u> לא מספיק שהפילוג השולי של תהליך יהיה גאוסי על מנת לומר שהת"א הינו תא"ג.

עבור מ"א פילוג גאוסי מוגדר ע"י תוחלת ושונות. עבור ו"א פילוג גאוסי מוגדר ע"י וקטור התוחלות ומטריצת הקו-ואריאנס. בשני המקרים מדובר בסטטיסטיקה מסדר שני.

תא"ג מוגדר ע"י סטטיסטיקה מסדר שני של ת"א – כלומר פונקציית תוחלת ופונקציית אוטו-קורולציה או פונקציית אוטו-קו-ואריאנס.

סטטיסטיקה מסדר שני של תהליכים אקראיים

סטטיסטיקה מסדר שני של ת"א $X(t)\backslash X_n$ היא אוסף כל המומנטים עד סדר שני של התהליך.

$$\underline{\text{Continuous}}$$
$$\forall t,t_1,t_2\in\mathbb{R}$$
$$\eta_X(t)=E[X_t]=E[X(t)]$$
$$R_{XX}(t_1,t_2)=E[X(t_1)\cdot X(t_2)]$$
$$C_{XX}(t_1,t_2)=E\left[\left(X(t_1)-\eta_X(t_1)\right)\cdot\left(X(t_2)-\eta_X(t_2)\right)\right]=Cov\left(X(t_1),X(t_2)\right)$$
$$C_{XX}(t_1,t_2)=R_{XX}(t_1,t_2)-\eta_X(t_1)\cdot\eta_X(t_2)$$
$$\underline{\text{Discrete}}$$
$$\forall n,m\in\mathbb{Z}$$
$$\eta_X(n)=E[X_n]=E[X(n)]$$
$$R_{XX}(n,m)=E[X_n\cdot X_m]$$
$$C_{XX}(n,m)=E\left[\left(X_n-\eta_X(n)\right)\cdot\left(X_m-\eta_X(m)\right)\right]=Cov(X_n,X_m)$$
$$C_{XX}(n,m)=R_{XX}(n,m)-\eta_X(n)\cdot\eta_X(m)$$

לכן ידיעת סטטיסטיקה מסדר שני היא ידיעת הנ"ל (כזכור קיימת שקילות בין מומנטים רגילים למרכזיים).

הערות:

- עבור ת"א מטריצות הקורולציה והקו-ואריאנס הינן אינסופיות. הסימונים הנ"ל הם פנייה לתא מסויים. כמובן שהדרישה של ידיעת כל הסטטיסטיקה היא ידיעת כל תא.

$$C_{\underline{X}_t\underline{X}_t}=\begin{bmatrix}C_{\underline{X}_t\underline{X}_t}(t_1,t_1) & \cdots & C_{\underline{X}_t\underline{X}_t}(t_1,t_N)\\ \vdots & \ddots & \vdots\\ C_{\underline{X}_t\underline{X}_t}(t_N,t_1) & \cdots & C_{\underline{X}_t\underline{X}_t}(t_Nt_N)\end{bmatrix},R_{\underline{X}_t\underline{X}_t}=\begin{bmatrix}R_{\underline{X}_t\underline{X}_t}(t_1,t_1) & \cdots & R_{\underline{X}_t\underline{X}_t}(t_1,t_N)\\ \vdots & \ddots & \vdots\\ R_{\underline{X}_t\underline{X}_t}(t_N,t_1) & \cdots & R_{\underline{X}_t\underline{X}_t}(t_Nt_N)\end{bmatrix}$$

- מטריצות $R_{\underline{X}\underline{X}},C_{\underline{X}\underline{X}}$ הינן סימטריות, כלומר מתקיים $A(i,j)=A(j,i)$.
- נבחין כי איברי האלכסון של מטריצות המומנטים השניים מקיימים:

$$R_{XX}(t,t)=E[X(t)\cdot X(t)]=E[X^2(t)]$$
$$C_{XX}(t,t)=E\left[\left(X(t)-\eta_X(t)\right)\cdot\left(X(t)-\eta_X(t)\right)\right]=Var\left(X(t)\right)$$

- מטריצות $R_{\underline{X}\underline{X}},C_{\underline{X}\underline{X}}$ הינן מטריצות אי-שליליות מוגדרות (PSD) ולכן:

$$\forall \underline{a},\underline{t} \in \mathbb{R}^N \qquad \underline{a}^T \cdot R_{\underline{X}\,\underline{X}} \cdot \underline{a} \geq 0 \quad \Rightarrow \quad \sum_{i=1}^{N}\sum_{j=1}^{N} a_i \cdot a_j \cdot R_{\underline{X}\,\underline{X}}\left(t_i,t_j\right) \geq 0$$

או בגבול האינטגרלי:

$$\int_{i=-\infty}^{\infty}\int_{j=-\infty}^{\infty} a(i)\cdot a(j)\cdot R_{\underline{X}\,\underline{X}}(i,j)\cdot dj\cdot di \geq 0$$

<u>מקדמי מתאם ליניאריים בין דגימות של תהליך אקראי</u>

נגדיר את מקדמי המתאם בין זוג דגימות של ת"א $X(t)$ באופן הבא:

$$r(t_1,t_2) = \frac{R_{XX}(t_1,t_2)}{\sqrt{R_{XX}(t_1,t_1)\cdot R_{XX}(t_2,t_2)}} = \frac{E\left[X(t_1)\cdot X(t_2)\right]}{\sqrt{E\left[X^2(t_1)\right]\cdot E\left[X^2(t_2)\right]}}$$

$$\rho(t_1,t_2) = \frac{C_{XX}(t_1,t_2)}{\sqrt{C_{XX}(t_1,t_1)\cdot C_{XX}(t_2,t_2)}} = \frac{Cov\left(X(t_1),X(t_2)\right)}{\sqrt{Var\left(X(t_1)\right)\cdot Var\left(X(t_2)\right)}}$$

וכמובן שמתקיים $-1 \leq r(t_1,t_2),\rho(t_1,t_2) \leq 1$.

נוכל לחלץ ולקבל את אי השיוויון הבא:

$$C_{XX}^2(t_1,t_2) \leq C_{XX}(t_1,t_1)\cdot C_{XX}(t_2,t_2) \quad \Leftrightarrow \quad Cov^2\left(X(t_1),X(t_2)\right) \leq Var\left(X(t_1)\right)\cdot Var\left(X(t_2)\right)$$

ז"א שונות משותפת בריבוע של שתי דגימות של ת"א לא גדולה ממכפלת השונויות של כל דגימה בנפרד.

<u>סטציונריות במובן הצר (SSS) ובמובן הרחב (WSS)</u>

תהליכים אקראיים הינם בעלי תכונת הסטוכסטיות (תהליכים סטוכסטיים), ז"א התפתחות התהליך תלוייה בגורמים מקריים (שאינם דטרמיניסטיים). כלומר ממצב התחלתי נתון של המערכת קיימים מספר מצבים שונים אליהם יכולה המערכת להגיע (כמובן שייתכן כי למצבים מסויימים הסתברות גבוהה יותר).

קיימת תת קבוצה של קבוצת התהליכים האקראיים אשר מקיימת את תכונת הסטציונאריות.

<u>סטציונריות במובן הצר (SSS)</u>

תהליך אקראי יקרא סטציונארי במובן הצר (Strict Sense Stationarity) אם הפילוג השולי שלו (עבור כל סט דגימות אפשרי) לא תלוי בזמן.

כלומר, ת"א $X(t)$ יקרא SSS אם לכל סט זמנים $\{t_1,\dots,t_N\}$ ולכל הזזה בזמן Δ מתקיים:

$$f_{X(t_1),\dots,X(t_N)}\left(x_1,\dots,x_N\right) = f_{X(t_1+\Delta),\dots,X(t_N+\Delta)}\left(x_1,\dots,x_N\right)$$

בפרט אם הפילוג השולי כלל לא תלוי בזמן, ז"א מתקיים $f_{X(t)}(x) = f_{X(t+\Delta)}(x)$ עבור כל t,Δ .

<u>סטציונריות במובן הרחב (WSS)</u>

תהליך אקראי $X(t)$ יקרא סטציונארי במובן הרחב (Wide Sense Stationarity) אם:

1. תוחלת דגימות התהליך לא תלויות בזמן
2. פונקציית הקורולצייה בין כל זוג דגימות לא תלויה בהזזה בזמן (תלוייה בהפרש הזמנים בלבד)

כלומר, ת"א $X(t)$ יקרא WSS אם:

1. $\forall t \in \mathbb{R} \quad \eta_X(t) = \eta_X$
2. $\forall t_1, t_2, \Delta \in \mathbb{R} \quad R_{XX}(t_1,t_2) = R_{XX}(t_1+\Delta, t_2+\Delta)$

<u>הערה –</u>

אם נציג את תנאי 2 עם $\Delta = -t_2$ נקבל כי מתקיים

$R_{XX}(t_1,t_2) = R_{XX}(t_1-t_2, t_2-t_2) = R_{XX}(t_1-t_2, 0)$, כלומר פונקציית הקורולצייה תלויה בהפרש הזמנים $t_1 - t_2$ בלבד. נסמן את הפרש הזמנים $\tau = t_1 - t_2$ ולכן נוכל לרשום:

$$R_{XX}(t_1,t_2) = R_{XX}(t_1-t_2, 0) \overset{\tau=t_1-t_2}{=} R_{XX}(\tau, 0) \overset{*}{=} R_{XX}(\tau)$$

כאשר בשיוויון $(*)$ אנו מסמנים את הפונקציה $R_{XX}(\cdot)$ כפונקציה של משתנה אחד מטעמי נוחות, זו אותו פונקציה של שני משתנים אך נשתמש בארגומנט אחד אשר מייצג את הפרש הארגומנטים כאשר נדון בפונקציית הקורולצייה של ת"א WSS.

<u>תכונות פונקציית הקורולציה עבור תהליך אקראי WSS</u>

אם $X(t)$ הוא תהליך אקראי סטציונרי במובן הרחב (WSS) אז:

1. $R_{XX}(t_1,t_2) = R_{XX}(t_1-t_2, 0) \quad \Rightarrow \quad E\left[X(t_1)\cdot X(t_2)\right] = E\left[X(t_1-t_2)\cdot X(0)\right]$
2. $R_{XX}(0)$ הוא הספק התהליך האקראי

   הפרש הזמנים מתאפס עבור זוג דגימות זהות, ז"א שני הזמנים היו זהים. במצב זה נקבל:

$$R_{XX}(t,t) = E\left[X(t)\cdot X(t)\right] = E\left[X^2(t)\right] = Power\left(X(t)\right)$$

ניתן גם להראות בכיוון ההפוך:

$$R_{XX}(0) = R_{XX}(0,0) \overset{*}{=} R_{XX}(t,t) = E\left[X(t)\cdot X(t)\right] = E\left[X^2(t)\right] = Power\left(X(t)\right)$$

כאשר שיוויון $(*)$ מתקיים מתכונה 2 של ת"א WSS (פונקציית הקורולציה לא רגישה להזזה בזמן).

3. פונקציית הקורולצייה $R_{XX}(\tau)$ היא פונקציה סימטרית
מהגדרת התוחלת ניתן להסיק כי פונקציית הקורולצייה של ת"א WSS היא פונקציה סימטרית

$$R_{XX}(\tau)=R_{XX}(\tau,0)\overset{*}{=}R_{XX}(0,-\tau)=E\left[X(0)\cdot X(-\tau)\right]=E\left[X(-\tau)\cdot X(0)\right]=$$
$$=R_{XX}(-\tau,0)=R_{XX}(-\tau)$$

כאשר שיוויון $(*)$ מתקיים מתכונה 2 של ת"א WSS (פונקציית הקורולציה לא רגישה להזזה בזמן).

4. מתקיים $\left|R_{XX}(\tau)\right|\le R_{XX}(0)\quad\forall\tau\in\mathbb{R}$
ז"א הפונקצייה מקבל ערך (גודל) מקסימלי עבור הפרש זמנים אפס.
כזכור, מהגדרת מקדם מתאם פירסון עבור זוג דגימות של ת"א ע"י:

$$r(t_1,t_2)=\frac{R_{XX}(t_1,t_2)}{\sqrt{R_{XX}(t_1,t_1)\cdot R_{XX}(t_2,t_2)}}\overset{\tau=t_1-t_2}{=}\frac{R_{XX}(\tau,0)}{\sqrt{R_{XX}(0,0)\cdot R_{XX}(0,0)}}=\frac{R_{XX}(\tau)}{\sqrt{R_{XX}^2(0)}}=\frac{R_{XX}(\tau)}{\left|R_{XX}(0)\right|}$$

נסדר את השיוויון ונקבל:

$$R_{XX}(\tau)=r(t_1,t_2)\cdot\left|R_{XX}(0)\right|\quad\Rightarrow\quad\left|R_{XX}(\tau)\right|=\left|r(t_1,t_2)\right|\cdot\left|R_{XX}(0)\right|$$

אנו יודעים כי מתקיים $\left|r(t_1,t_2)\right|\le 1$ ובנוסף כי $\left|R_{XX}(0)\right|$ אי-שלילי (תכונה 2 - הספק) ולכן:

$$\left|R_{XX}(\tau)\right|\le R_{XX}(0)$$

5. $$\int_{i=-\infty}^{\infty}\int_{j=-\infty}^{\infty}a(i)\cdot a(j)\cdot R_{\underline{X}\,\underline{X}}(i-j)\cdot dj\cdot di\ge 0$$

ראינו כי מאחר ומטריצת הקורולציה $R_{\underline{X}\,\underline{X}}$ הינה מטריצת PSD מתקיים:

$$\int_{i=-\infty}^{\infty}\int_{j=-\infty}^{\infty}a(i)\cdot a(j)\cdot R_{\underline{X}\,\underline{X}}(i,j)\cdot dj\cdot di\ge 0$$

ומאחר ש- $X(t)$ הינו ת"א WSS מתקיים $R_{\underline{X}\,\underline{X}}(i,j)=R_{\underline{X}\,\underline{X}}(i-j,0)=R_{\underline{X}\,\underline{X}}(i-j)$ ולכן:

$$\int_{i=-\infty}^{\infty}\int_{j=-\infty}^{\infty} a(i)\cdot a(j)\cdot R_{\underline{X}\,\underline{X}}(i-j)\cdot dj\cdot di\geq 0$$

<u>סטציונריות אסימפטוטית</u>

לעיתים ת"א לא יקיים את דרישות הסטציונריות אך הוא יקיים אותן באופן אסימפטוטי (בגבול).

<u>סטציונאריות אסימפטוטית במובן הצר:</u>

תהליך אקראי $X(t)$ יקרא סטציונארי אסימפטוטי במובן הצר (Asymptotic SSS) אם לפילוג השולי שלו (עבור כל סט דגימות אפשרי $\{t_1,\dots,t_N\}$) קיים הגבול:

$$\lim_{\Delta\to\infty}\left(f_{X(t_1+\Delta),\dots,X(t_N+\Delta)}(x_1,\dots,x_N)\right)$$

נסמן את הגבול כפילוג הסטציונרי:

$$f_{t_1,\dots,t_N}^{Stationary}(x_1,\dots,x_N)$$

<u>סטציונאריות אסימפטוטית במובן הרחב:</u>

תהליך אקראי $X(t)$ יקרא סטציונארי אסימפטוטי במובן הרחב (Asymptotic WSS) אם לפילוג השולי שלו (עבור כל סט דגימות אפשרי $\{t_1,\dots,t_N\}$) קיים הגבול:

$$\lim_{\Delta\to\infty}\left(f_{X(t_1+\Delta),\dots,X(t_N+\Delta)}(x_1,\dots,x_N)\right)$$

נסמן את הגבול כפילוג הסטציונרי:

$$f_{t_1,\dots,t_N}^{Stationary}(x_1,\dots,x_N)$$

<u>סטציונריות משותפת במובן הצר (jSSS) ובמובן הרחב (jWSS)</u>

הסברנו כי תהליכים אקראיים הינם בעלי תכונת הסטוכסטיות. עבור זוג ת"א כל אחד מהם סטוכסטי ונרצה לדעת האם נוכל להגיד כי יחדיו הם מקיימים תכונת סטציונריות כלשהי.

<u>סטציונריות משותפת במובן הצר (jSSS)</u>

זוג תהליכים אקראיים $X(t)$ ו- $Y(t)$ יקראו סטציונאריים במשותף במובן הצר (Jointly Strict Sence Stationarity) אם הפילוג השולי שלו (עבור כל סט דגימות אפשרי) לא תלוי בזמן.

כלומר, ת"א $X(t)$ ו- $Y(t)$ יקראו jSSS אם לכל סט זמנים $\{t_1,\dots,t_N,t_{N+1},\dots,t_{N+M}\}$ ולכל הזזה בזמן Δ מתקיים (כנ"ל לגביי פונקציית ה-jCDF):

$$f_{X(t_1),\dots,X(t_N),Y(t_{N+1}),\dots,Y(t_{N+M})}(x_1,\dots,x_N,y_1,\dots,y_M)=$$
$$=f_{X(t_1+\Delta),\dots,X(t_N+\Delta),Y(t_{N+1}+\Delta),\dots,Y(t_{N+M}+\Delta)}(x_1,\dots,x_N,y_1,\dots,y_M)$$

כמובן שאם זוג תהליכים אקראיים $X(t)$ ו- $Y(t)$ הם jSSS אז זה אומר שכ"א SSS בעצמו (ניתן להבחין כי התנאי מתקיים גם עבור פונקציית ה-jCDF ונוכל לייצר ממנה פונקציית jCDF שולית ע"י הצבת ערכים $t_i=\infty$ עבור כל הזמנים של אחד התהליכים ע"מ לנוון אותו).

<u>סטציונריות משותפת במובן הרחב (jWSS)</u>

זוג תהליכים אקראיים $X(t)$ ו- $Y(t)$ יקראו סטציונאריים במשותף במובן הרחב (jointly Wide Sense Stationarity) אם:

1. כ"א מן התהליכים הוא WSS בעצמו ולכן:
   a. $\forall t\in\mathbb{R}\quad \eta_X(t)=\eta_X,\eta_Y(t)=\eta_Y$
   b. $\forall t_1,t_2,\Delta\in\mathbb{R}\quad R_{XX}(t_1,t_2)=R_{XX}(t_1+\Delta,t_2+\Delta),R_{YY}(t_1,t_2)=R_{YY}(t_1+\Delta,t_2+\Delta)$
2. פונקציית הקרוס-קורולצייה בין כל זוג דגימות לא תלויה בהזזה בזמן (תלוייה בהפרש הזמנים בלבד), כלומר:

$$\forall t_1,t_2,\Delta\in\mathbb{R}\quad R_{XY}(t_1,t_2)=R_{XY}(t_1+\Delta,t_2+\Delta)$$

<u>הערה –</u>
גם עבור פונקציית הקרוס-קורולצייה של שני ת"א jWSS נוכל להגדיר:

$$R_{XY}(t_1,t_2)=R_{XY}(t_1-t_2,0)\overset{\tau=t_1-t_2}{=}R_{XY}(\tau,0)=R_{XY}(\tau)$$

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

יהי $\{W_n\}$ ת"א i.i.d בעל פילוג ידוע $f_W(w)$ ותנאי התחלה משתנה אקראי X_0 בעל פילוג ידוע $f_{X_0}(x)$ (בפרט אם X_0 דטרמיניסטי) כאשר ת"ה X_0 בת"ס ב- $\{W_n\}$ (ז"א בכל $\{W_1,W_2,\dots\}$ יחד ולא בכ"א בנפרד בלבד).

פונקציה דטרמיניסטית $g(\cdot)$ תגדיר את $\{X_n\}$ ת"א A.R. ע"י:

$$X_n=g(X_{n-1},W_n)$$

פונקציה יכולה להיות כל פונקציה דטרמיניסטית, לדוגמא:

$$g_1(u,v)=u+v,\quad g_1(u,v)=u\oplus v$$

<u>מרקוביות של תהליך אקראי אוטו-רגרסיבי</u>

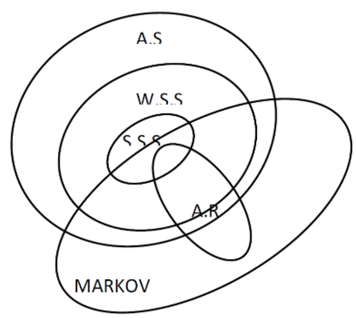
נדון במרקוביות בהרחבה בפרק של שרשראות מרקוב אך נסביר כעת בקצרה.

מערכת מוגדרת כמערכת מרקובית אם סך המידע בהווה אשר יכול לשמש לידיעת העתיד מתמצה במידע הנוכחי, ז"א אין צורך לזכור את העבר ע"מ להיות "חכמים יותר" מאחר שהנתונים הנוכחיים מספקים אותו מידע.

<u>טענה</u> - ת"א A.R. הינו מרקובי.

<u>הסבר</u> -בכל איטרציה X_n תלוי ב- X_{n-1} וב- W_n. אנו יודעים כי $\{W_i\}$ הוא ת"א i.i.d ולכן התפלגות W_n לא תלויה בערכי W_k הקודמים ($W_1,W_2,\dots,W_{n-1}$) וגם לא תלויה בערכי X_k הקודמים ($X_1,X_2,\dots,X_{n-1}$). לכן למעשה X_n מתבסס על X_{n-1} והמשמעות היא ש- X_{n-1} מכיל בתוכו את כל המידע הרלוונטי מן העבר.

<u>הערה</u> - נציג דיאגרמת ון עבור קבוצת התהליכים האקראיים (לא מלאה):



A.S היא קבוצת התהליכים האקראיים אשר WSS אסימפטוטית (ילמד בהמשך).

<u>תהליך אקראי אוטו-רגרסיבי ליניארי</u>

ת"א אוטו-רגרסיבי $\{X_n\}$ יקרא ת"א A.R. ליניארי אם $g(\cdot,\cdot)$ הינה פונקציה דטרמיניסית ליניארית מהצורה:

$$g(u,v)=\alpha\cdot u+v$$

ולכן הגדרת ת"א אוטו-רגרסיבי ליניארי $\{X_n\}$ תהיה מהצורה:

$$X_n=\alpha\cdot X_{n-1}+W_n$$

בקורס לא נתעסק עם תהליכים אקראיים אוטו-רגרסיביים אשר אינם ליניאריים.

<u>הערה</u> - אנו מבחינים כי מן הדוגמאות הקונסטרוקטיביות כל התהליכים הינם תהליכים A.R. אך לא כולם ליניאריים.

1. $X_n=0\cdot X_{n-1}+W_n,\quad X_0=0,\quad \{W_n\}\quad i.i.d$
2. $X_n=1\cdot X_{n-1}+W_n,\quad X_0=0,\quad \{W_n\}\quad i.i.d$
3. $X'_n=X'_{n-1}\oplus W_n,\quad X'_0=0\quad \{W_n\}\quad i.i.d$

4. $X''_n = 1 \cdot X''_{n-1} + X'', \quad X''_0 = 0, \quad \{W''_n\} \quad i.i.d$

<u>סטציונריות של תהליך אקראי אוטו-רגרסיבי ליניארי</u>

<u>טענה:</u>

עבור ת"א אוטו-רגרסיבי ליניארי $\{X_n\}$ אשר מוגדר ע"י:

$$\begin{aligned} &X_n = \alpha \cdot X_{n-1} + W_n && W_i \sim f_W(w) \quad \forall i \\ &X_0 \sim f_{X_0}(x) && \{W_n\} \, i.i.d \end{aligned}$$

מתקיים:

1. אם $|\alpha| \geq 1$ אז הת"א לא סטציונרי.

   ניתן להבחין כי עבור $|\alpha| \geq 1$ מתקיים $Var(X_n) \underset{n\to\infty}{\to} \infty$ וזה מפר דרישת WSS (ולכן גם SSS).

2. אם $|\alpha| < 1$ אז הת"א SSS אסימפטוטית (ולכן גם WSS אסימפטוטית).

   ניתן להראות כי הגבול קיים משום שבזמן "אינסוף" מתקיים $\alpha^n \to 0$.

3. אם $|\alpha| < 1$ ובנוסף תנאי ההתחלה מתואם אז הת"א סטציונרי בהתאם לתנאי ההתחלה.

   a. ת"ה מתואם עבור SSS הוא $f_{X_0}(x) = f_X^{Stat}(x)$ (ז"א הפילוג השולי של רגע אפס יהיה זהה לפילוג הסטציונרי).

   b. ת"ה מתואם עבור WSS הוא $\eta_{X_0} = \eta_{Stat}, \quad \sigma^2_{X_0} = \sigma^2_{Stat}$, ניתן לחשב את $(\eta_{Stat}, \sigma^2_{Stat})$ מתוך הסטטיסטיקה מסדר שני של התהליך (נדרוש כי השונות והתוחלת של רגע אפס יהיו זהים לסטציונרים).

<u>דוגמא:</u>

נבחן את הת"א האוטו-רגרסיבי ליניארי הבא:

$$\begin{aligned} &X_n = \tfrac{1}{2} \cdot X_{n-1} + W_n \\ &X_0 = 0 \end{aligned} \qquad W_n = \begin{cases} 1 & w.p. \tfrac{1}{2} \\ 0 & w.p. \tfrac{1}{2} \end{cases}, \{W_n\} \quad i.i.d$$

אנו רואים כי $\alpha = \frac{1}{2}$ ולכן במקרה הגרוע התהליך SSS אסימפטוטית.

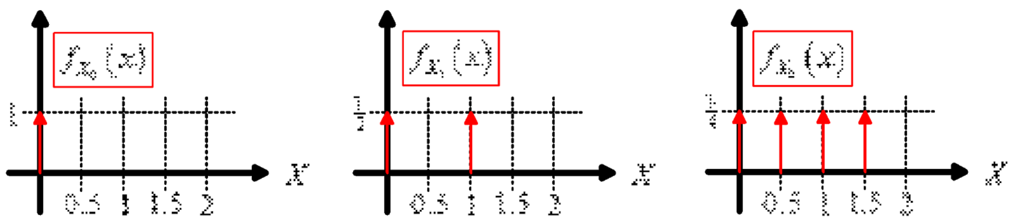
תנאי ההתחלה הינו $X_0 = 0$, ז"א ידוע באופן דטרמיניסטי (ערכו 0 בהסתברות 1) ולכן $f_{X_0}(x) = \delta(x)$.

נחשב עבור הדגימות הבאות:

$$X_1=\tfrac{1}{2}\cdot X_0+W_1=\tfrac{1}{2}\cdot 0+W=W=\begin{cases}1 & w.p.\tfrac{1}{2}\\ 0 & w.p.\tfrac{1}{2}\end{cases}\;\Rightarrow\; f_{X_1}(x)=\tfrac{1}{2}\cdot\delta(x)+\tfrac{1}{2}\cdot\delta(x-1)$$

$$X_2=\tfrac{1}{2}\cdot X_1+W_2=\tfrac{1}{2}\cdot X_1+W=\begin{cases}1.5 & w.p.\tfrac{1}{4}\\ 1 & w.p.\tfrac{1}{4}\\ 0.5 & w.p.\tfrac{1}{4}\\ 0 & w.p.\tfrac{1}{4}\end{cases}\;\Rightarrow\; \begin{aligned}f_{X_2}(x)=\tfrac{1}{4}\cdot\big[\delta(x)+\delta(x-0.5)+\\ +\delta(x-1)+\delta(x-1.5)\big]\end{aligned}$$

נראה גרפית את פילוג הדגימות הללו:



אנו מזהים כי בגבול $n\to\infty$ פילוג הדגימה X_n ישאף לפילוג הרציף $Unif(0,2)$.

נוכל לוודא זאת לפי הרצת האיטרציה ה- $(n+1)$ עבור $n\to\infty$:

$$\boldsymbol{X_{n+1}=\tfrac{1}{2}\cdot X_n+W_{n+1}}$$

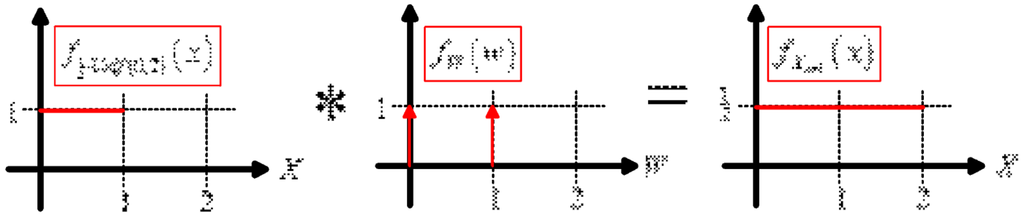
$$f_{X_{n+1}}(x)=\left(f_{\frac{1}{2}\cdot X_n}(x)\right)*\left(f_W(x)\right)=\left(2\cdot f_{X_n}(2x)\right)*\left(f_W(x)\right)=$$

$$=\left(2\cdot\begin{cases}0.5 & 0\le 2x\le 2\\ 0 & else\end{cases}\right)*\left(\tfrac{1}{2}\cdot\delta(x)+\tfrac{1}{2}\cdot\delta(x-1)\right)=\frac{1}{2}\cdot\left(\begin{cases}1 & 0\le x\le 1\\ 0 & else\end{cases}\right)*\left(\delta(x)+\delta(x-1)\right)=$$

$$=\frac{1}{2}\cdot\left[\left(\begin{cases}1 & 0\le x\le 1\\ 0 & else\end{cases}\right)*\delta(x)+\left(\begin{cases}1 & 0\le x\le 1\\ 0 & else\end{cases}\right)*\delta(x-1)\right]=$$

$$=\frac{1}{2}\cdot\left[\left(\begin{cases}1 & 0\le x\le 1\\ 0 & else\end{cases}\right)+\left(\begin{cases}1 & 1\le x\le 2\\ 0 & else\end{cases}\right)\right]\sim\frac{1}{2}\cdot\left[Unif(0,1)+Unif(1,2)\right]=Unif(0,2)$$

נראה גם גרפית:



ואכן קיבלנו כי פילוג שתי דגימות סמוכות באינסוף לא משתנה, לכן זהו הפילוג הסטציונרי.
נוכל גם למצוא משוואה רקורסיבית עבור התהליך:

$$X_n = \tfrac{1}{2}\cdot X_{n-1} + W_n = \tfrac{1}{2}\cdot\left(\tfrac{1}{2}\cdot X_{n-2} + W_{n-1}\right) + W_n = \left(\tfrac{1}{2}\right)^2\cdot X_{n-2} + \tfrac{1}{2}\cdot W_{n-1} + W_n =$$
$$= \ldots = \left(\tfrac{1}{2}\right)^n \cdot X_0 + \sum_{i=1}^{n}\left(\tfrac{1}{2}\right)^{n-i}\cdot W_i$$

אנו מזהים כי הראש דועך אקספוננציאלית ב- n ולכן רק הסכום רלוונטי.

אנחנו מקבלים SSS עבור ת"ה X_0 מתואם ל- $X_n, n\to\infty$, ז"א מתואם לפילוג $Unif\left(0,2\right)$.
אצלנו לא.

הערה - בחישוב הפילוג הסטציונארי ביצענו (נובע מהתמרות של מ"א):

$$X_n = \tfrac{1}{2}\cdot X_{n-1} + W_n$$
$$\Downarrow$$
$$f_X\left(x\right) = f_{\frac{1}{2}\cdot X}\left(x\right) * f_W\left(x\right) = \left(2\cdot f_X\left(2x\right)\right) * f_W\left(x\right)$$

שרשראות מרקוב

שרשראות מרקוב הן כלי לתיאור תהליך מרקובי בזמן בדיד. אנו נגדיר תהליך מרקובי אך נתעסק בתהליכים מרקוביים בזמן בדיד בלבד (ולרוב נתעסק בתהליכים מרקוביים הומוגניים בזמן בדיד בעלי מספר מצבים סופי).
שרשראות מרקוב מציגות התפתחות תהליכים כסדרת מצבים וניתן לחשב מהן הסתברות לסדרת אירועים מסויימת, לחלק את המערכת לסוגי מצבים ועוד.

תהליך מרקובי

תהליכים אקראיים מרקוביים הינם תת-מחלקה מאוד פשוטה ואינטואיטיבית של תהליכים סטוכסטיים (אקראיים).

ת"א $X\left(t\right)$ יקרא מרקובי אם פילוג העתיד תלוי רק במצב ההווה וללא תלות בעבר ("כל החטאים נמחלנו").

למעשה, ת"א $X\left(t\right)$ יקרא מרקובי אם הפילוג המותנה של העתיד בהינתן ההווה והעבר שווה לפילוג המותנה של העתיד בהינתן ההווה בלבד (היות ומספיק להתנות את ההסתברות בערך הנוכחי של התהליך נקרא לערך זה ה"מצב" של התהליך).

עבור ת"א בזמן רציף (עבור כל סדרת זמנים $t_1 < t_2 < \ldots < t_N \in \mathbb{R}$):

$$\Pr\left(X(t_N) \in S \middle| X(t_{N-1}) = x_{N-1}, \ldots, X(t_1) = x_1\right) = \Pr\left(X(t_N) \in S \middle| X(t_{N-1}) = x_{N-1}\right)$$

או בכתיב אחר:

$$\Pr\left(X(t_N) \in S \middle| \left\{X(t), t < t_N\right\}\right) = \Pr\left(X(t_N) \in S \middle| \left\{X(t), t \le t_{N-1}\right\}\right) = \Pr\left(X(t_N) \in S \middle| X(t_{N-1}) = x_{N-1}\right)$$

ועבור ת"א בזמן בדיד:

$$\Pr\left(X_{k_N} \in S \middle| X_{k_{N-1}} = x_{N-1}, \ldots, X_{k_2} = x_2\right) = \Pr\left(X_{k_N} \in S \middle| X_{k_{N-1}} = x_{N-1}\right) \qquad \forall k_1 < k_2 < \ldots < k_N \in \mathbb{Z}$$

<u>כלל השרשרת (עבור שרשראות מרקוב)</u>

עבור ת"א מרקוביים בזמן בדיד מתקיים (עבור סט נקודות זמן עולה ממש $k_1 < k_2 < \ldots < k_N \in \mathbb{Z}$ כלשהו):

$$\begin{aligned}
&\Pr\left(X_{k_1} = x_1, X_{k_2} = x_2, \ldots, X_{k_N} = x_N\right) \overset{*}{=} \Pr\left(X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_2} = x_2, \ldots, X_{k_N} = x_N \middle| X_{k_1} = x_1\right) = \\
&\overset{*}{=} \Pr\left(X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_2} = x_2 \middle| X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_3} = x_3, \ldots, X_{k_N} = x_N \middle| X_{k_2} = x_2, X_{k_1} = x_1\right) = \\
&\overset{**}{=} \Pr\left(X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_2} = x_2 \middle| X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_3} = x_3, \ldots, X_{k_N} = x_N \middle| X_{k_2} = x_2\right) = \\
&= \ldots = \Pr\left(X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_2} = x_2 \middle| X_{k_1} = x_1\right) \cdot \Pr\left(X_{k_3} = x_3 \middle| X_{k_2} = x_2\right) \cdot \\
&\cdot \Pr\left(X_{k_4} = x_4, \ldots, X_{k_N} = x_N \middle| X_{k_3} = x_3\right) = \ldots = \Pr\left(X_{k_1} = x_1\right) \cdot \prod_{i=2}^{N} \Pr\left(X_{k_i} = x_i \middle| X_{k_{i-1}} = x_{i-1}\right)
\end{aligned}$$

שיוויוני $(*)$ לפי הסתברות מותנית. שיוויון $(**)$ לפי מרקוביות (תלות במצב נוכחי ולא בעבר).

<u>נוסחת צ'אפמן קולמוגורוב</u>

עבור ת"א מרקובי בזמן בדיד מתקיים (עבור $k < m < n \in \mathbb{Z}$):

$$\Pr\left(X_n = x_n \middle| X_k = x_k\right) = \sum_m \left[\Pr\left(X_n = x_n \middle| X_m = x_m\right) \cdot \Pr\left(X_m = x_m \middle| X_k = x_k\right)\right]$$

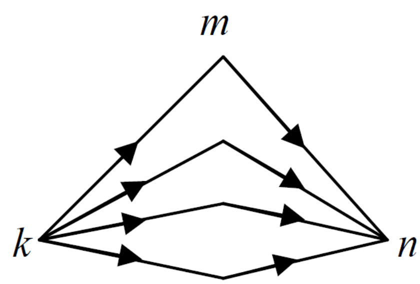
ז"א ההסתברות להגיע למצב x_n מסויים ברגע n אם ידוע כי ברגע k היינו במצב x_k כלשהו

הוא סך ההסתברויות להגיע מהמצב x_k למצב x_n דרך מצב ביניים x_m .

נוכל להציג את הנוסחה גם ע"י סכימת מצבי הביניים האפשריים במקום נקודות זמן-הביניים האפשריות:

$$\Pr\left(X_n = x_n \middle| X_k = x_k\right) = \sum_{x_m} \left[\Pr\left(X_n = x_n \middle| X_m = x_m\right) \cdot \Pr\left(X_m = x_m \middle| X_k = x_k\right)\right]$$

תרשים גרפי של משמעות הנוסחה:



הוכחה:

הדבר הגיוני מנוסחת ההסתברות השלמה. נראה:

$$\Pr\left(X_n = x_n \middle| X_k = x_k\right) \overset{*}{=} \sum_{x_m} \Pr\left(X_n = x_n, X_m = x_m \middle| X_k = x_k\right) =$$
$$\overset{Bayes}{=} \sum_{x_m} \Pr\left(X_n = x_n \middle| X_m = x_m, X_k = x_k\right) \cdot \Pr\left(X_m = x_m \middle| X_k = x_k\right) =$$
$$\overset{**}{=} \sum_{x_m} \Pr\left(X_n = x_n \middle| X_m = x_m\right) \cdot \Pr\left(X_m = x_m \middle| X_k = x_k\right)$$

שיוויון לפי נוסחת ההסתברות השלמה. שיוויון $(**)$ לפי מרקוביות (תלות במצב נוכחי ולא בעבר).

שרשרת מרקוב סופית

שרשרת מרקוב (ת"א מרקובי בזמן בדיד) תקרא סופית אם יש לה מספר סופי (או בר-מניה) של מצבים, זאת אומרת מתקיים כי מצב יכול לקבל ערך מתוך קבוצה סופית (או ברת-מניה).

$$X_n \in \{x_1, x_2, \ldots, x_J\} \overset{*}{=} \{1, 2, \ldots, J\} \qquad \forall n \in \mathbb{Z}$$

שיוויון $(*)$ מבוצע לשם נוחות, נסמן את המצבים עם מספרים לצורך הקלת הדיון.

הערה - מעתה נדון בשרשראות מרקוב (תהליכים מרקוביים בזמן בדיד) סופיות בלבד

פילוג שולי של שרשרת מרקוב

עבור שרשרת מרקוב נרצה לדעת איך מתפלג מצב התהליך ברגע מסויים n. אם השרשרת סופית נוכל להציג את הפילוג השולי של X_n ע"י:

$$\underline{\pi}(n) = \left(\pi_1(n), \ldots, \pi_J(n)\right) = \left(\Pr\left(X_n = 1\right), \ldots, \Pr\left(X_n = J\right)\right)$$

מטריצת מעבר של שרשרת מרקוב

עבור שרשרת מרקוב נרצה לדעת בכל נקודת זמן את הסתברויות המעבר בין המצבים. אם השרשרת סופית נוכל להציג הסתברויות אלו במטריצה:

$$\underline{\underline{P}}(n)=\begin{bmatrix} P_{11}(n) & \cdots & P_{1J}(n) \\ \vdots & \ddots & \vdots \\ P_{J1}(n) & \cdots & P_{JJ}(n) \end{bmatrix} \qquad P_{ij}(n)=\Pr\left(X_n=j \middle| X_{n-1}=i\right)$$

המטריצה מורכבת מתאים $P_{ij}(n)$. הערך $P_{ij}(n)$ מייצג את ההסתברות להגיע בזמן n למצב j אם ידוע שביחידת זמן קודם לכן, משמע בזמן $n-1$, היינו במצב i.

נזהה כי שורה i מייצגת את הסיכויים להגיע ממצב i (בזמן $n-1$) לכל שאר התאים לפי מיקום העמודה בזמן n.

נזהה כי עמודה j מייצגת את הסיכויים להגיע למצב j (בזמן n) מכל שאר המצבים לפי מיקום השורה בזמן $n-1$.

כל איברי המטריצה אי-שליליים (הסתברויות) וסך כל שורה הינו 1 (סך הסתברויות המעבר ממצב מסויים) ולכן מטריצה $\underline{\underline{P}}(n)$ הינה מטריצה סטוכסטית.

<u>מסקנה:</u>

תיאור סטטיסטי מלא של שרשרת מרקוב סופית עד זמן N ניתן ע"י הפילוג השולי של מצב ההתחלה ומטריצות המעברים לכל זמן $n=1,\ldots,N$, ז"א על ידי:

1. $\underline{\pi}(0)$
2. $\forall n=1,\ldots,N \qquad \underline{\underline{P}}(n)$

או באופן כללי ע"י $\underline{\pi}(0)$ ו- $\underline{\underline{P}}(n)$ עבור $n=1,2,\ldots$ ($\forall n\in\mathbb{Z}$).

<u>הסבר:</u>

לפי נוסחת ההסתברות השלמה נקבל את הפילוג השולי של כל מצב עפ"י נוסחת ההסבתרות השלמה:

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

ובפרט מתקיים $\underline{\pi}(1)=\underline{\pi}(0)\cdot\underline{\underline{P}}(1)$ (אינטואיטיבי לגבי פילוג המצב לאחר המעבר הראשון).

נראה כי למעשה מתקיים:

$$\underline{\pi}(n)=\underline{\pi}(n-1)\cdot\underline{\underline{P}}(n)=\underline{\pi}(n-2)\cdot\underline{\underline{P}}(n-1)\cdot\underline{\underline{P}}(n)=\ldots=\underline{\pi}(0)\cdot\prod_{i=0}^{n-1}\underline{\underline{P}}(n-i)=\underline{\pi}(0)\cdot\prod_{i=1}^{n}\underline{\underline{P}}(i)$$

דיאגרמת טרליס (Trellis Diagram)

דיאגרמת טרליס הינה כלי גרפי להצגת פילוג מצבי שרשרת מרקוב בזמן.
בכל עמודת זמן נקצה קודקוד עבור כל מצב אפשרי של השרשרת. נחבר בקשת שני מצבים $i \to j$ בזמנים עוקבים $n-1 \to n$ אם הסתברות המעבר חיובית (כלומר $P_{ij}(n) > 0$) ונרשום (לא חובה) על הקשת את הסתברות המעבר.

<u>דוגמא:</u>

נציג דיאגרמת טרליס עבור שרשרת מרקוב הנתונה לפי:

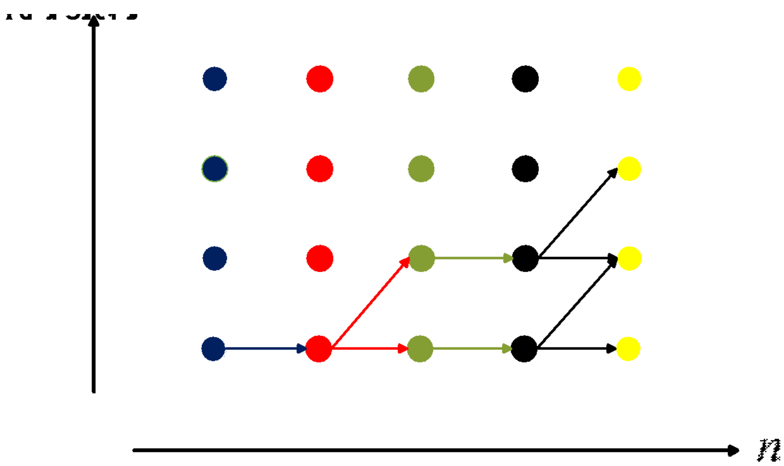
$$W_n = \begin{cases} 0 & n \quad is \quad odd \\ \begin{cases} 1 & w.p.\frac{1}{2} \\ 0 & w.p.\frac{1}{2} \end{cases} & n \quad is \quad even \end{cases} \qquad \begin{aligned} X_n &= \mathrm{mod}_4\left(X_{n-1} + W_n\right) \\ X_0 &= 0 \end{aligned}$$

נבחין בקלות כי $X_n \in \{0,1,2,3\}$ ולכן השרשרת סופית.

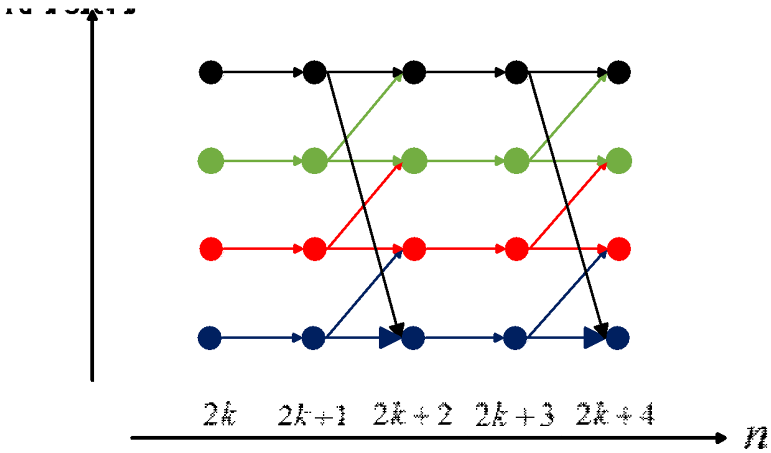
עבור זמנים אי-זוגיים לא מבוצע מעבר כלל כי בוודאות $W_n = 0$, ומנוסחת הנסיגה ומסופיות השרשרת $X_n = X_{n-1}$.

עבור זמנים זוגיים יש סיכוי חצי ל- $W_n = 0$ (ואז כמו קודם אין מעבר) וסיכוי חצי ל- $W_n = 1$ (ואז מבוצע קידום באחד ומודולו 4 על מנת "להחזיר" את ערך השלב לקבוצה הסופית שלו).

נתון שהמצב ההתחלתי של השרשרת הוא $X_0 = 0$ ולכן השרשרת נראית באופן הבא:



נבחין כי השרשרת "מכירה" את המצבים ככל שהיא מתקדמת. דיאגרמת טרליס עבור זמן מתקדם (מתחיל מזמן זוגי) בשרשרת:



נזהה כי השרשרת נהיית מחזורית.

הערה - ערכי ההסתברויות לא רשומים על הקשתות (כל המעברים לזמנים אי-זוגיים הם 1 – משמע וודאיים, כל המעברים לזמנים זוגיים הינם חצי).

שרשרת מרקוב הומוגנית ודיאגרמת מצבים

שרשרת מרקוב (ת"א מרקובי בזמן בדיד) תקרא הומוגנית (בזמן) אם מתקיים:

$$\Pr\left(X_n = x_D \middle| X_{n-1} = x_S\right) = \Pr\left(X_m = x_D \middle| X_{m-1} = x_S\right) \qquad \forall m,n \in \mathbb{Z}$$

כלומר, הפילוג המותנה של מעבר (ביחידת זמן אחת) הוא קבוע בזמן. המשמעות היא שמעבר ממצב מקור מסויים למצב מטרה אחר לא תלוי בזמן.

בפרט נראה כי סיכוי המעבר שווה לסיכוי אילו היה זה המעבר הראשון של השרשרת:

$$\Pr\left(X_n = x_D \middle| X_{n-1} = x_S\right) = \Pr\left(X_1 = x_D \middle| X_0 = x_S\right) \qquad \forall n \in \mathbb{Z}$$

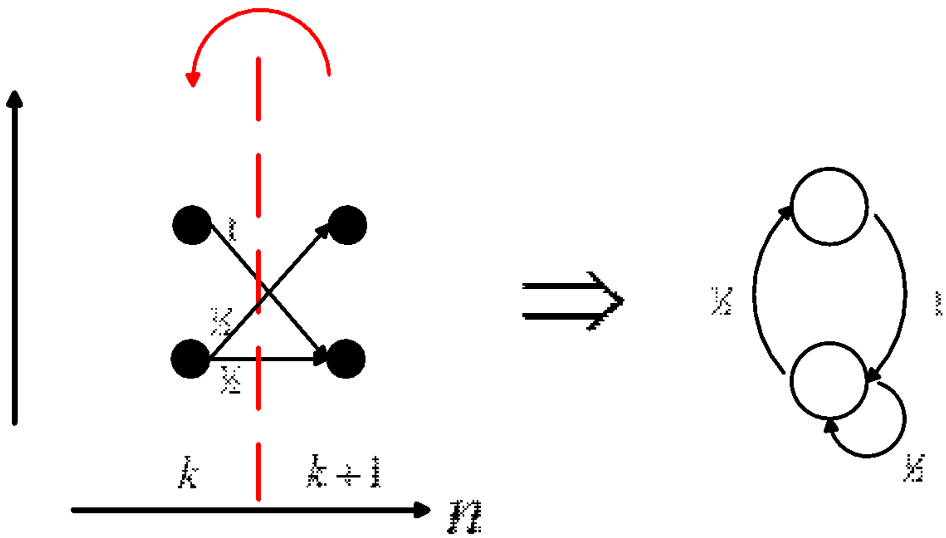
נבחין אם שרשרת מרקוב היא הומוגנית אז מטריצת המעבר שלה לא תלויה בזמן $\underline{\underline{P}}(n) = \underline{\underline{P}}$.

כמו כן מתקיים $\underline{\pi}(n) = \underline{\pi}(0)\cdot\left(\underline{\underline{P}}\right)^n$.

הערה - מעתה נדון בשרשראות מרקוב (תהליכים מרקוביים בזמן בדיד) הומוגניות וסופיות בלבד

נבחין כי עבור שרשרת מרקוב הומוגנית סופית לא נצטרך לצייר בדיאגרמת טרליס יותר מחתך בודד (מהמצב הסופי עבור זמן מתקדם) מאחר שאין שינוי בהסתברויות המעברים בזמן. לכן שרשראות מרקוב הומוגניות וסופיות מוצגות/נתונות ע"י דיאגרמת מצבים ולא דיאגרמת טרליס. למעשה נוכל להמיר חתך יחיד מדיאגרמת טרליס לדיאגרמת מצבים ע"י "קיפולו".

<u>דוגמא:</u>



נמשיך עם דוגמא זו ונתייחס אלייה כאל מקרה של 'חיפושית קופצת'.
כשהחיפושית במצב 1 יש לה סיכוי זהה להישאר או להחליף מצב, לחיפושית לא נוח במצב 2 ולכן תחזור למצב 1.

נניח כי אנו יודעים שבזמן אפס מיקמנו את החיפושית במצב 1, ז"א כי מתקיים $\underline{\pi}(0)=(1,0)$.

נבנה מן דיאגרמת המצבים את מטריצת המעברים ונרכיב את כלל הפילוג השולי של מעבר:

$$\underline{\underline{P}}(n)=\begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix} \quad \underline{\pi}(n)=\underline{\pi}(n-1)\cdot\begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}=\underline{\pi}(0)\cdot\left(\begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}\right)^n=(1,0)\cdot\left(\begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}\right)^n$$

נראה את הפילוגים השוליים של מצבי התהליך בזמן:

$$\underline{\pi}(1)=\left(1/2,1/2\right),\quad \underline{\pi}(2)=\left(3/4,1/4\right),\quad \underline{\pi}(3)=\left(5/8,3/8\right),\quad \underline{\pi}(4)=\left(11/16,5/16\right),\ldots$$

נרצה לדעת את הפילוג השולי של מצב החיפושית עבור $n\to\infty$ (אם יש כזה).

נוכל לחשבו עבור העלאת המטריצה $\underline{\underline{P}}(n)$ בחזקה אינסופית ולקוות להתכנסות, נבצע ונקבל:

$$\left(\underline{\underline{P}}(n)\right)^n=\left(\begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}\right)^n \xrightarrow[n\to\infty]{} \begin{bmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{bmatrix}$$

נבחין כי מתקיים:

$$\lim_{n\to\infty}\underline{\pi}(n)=\lim_{n\to\infty}\left(\underline{\pi}(0)\cdot\left(\begin{bmatrix}1/2 & 1/2\\ 1 & 0\end{bmatrix}\right)^n\right)=(1,0)\cdot\begin{bmatrix}2/3 & 1/3\\ 2/3 & 1/3\end{bmatrix}=\left(2/3,1/3\right)$$

נבחין למעשה כי כלל לא היה חשיבות למצב ההתחלתי. הרי אם איברי כל עמודה במטריצה $\left(\underline{\underline{P}}(n)\right)^n$ זהים וסך איברי ווקטור ההתחלה תמיד יהיה 1 (גם אם המצב ההתחלתי בעל פילוג לא דטרמיניסטי) אז פילוג המצב הסופי יהיה ווקטור שורה מן המטריצה $\left(\underline{\underline{P}}(n)\right)^n$ - כלל לא תלוי בפילוג השולי $\underline{\pi}(0)$.

בדרך נוספת, עם מעט אינטואיציה נוכל לזהות את כיוון ההתכנסות ולנחש כי הסיכוי ב"אינסוף" להיות במצב 1 הוא $2/3$, כלומר $\lim_{n\to\infty}\pi_1(n)=2/3$. נחלץ מכלל המעברים:

$$\underline{\pi}(n)=\underline{\pi}(n-1)\cdot\begin{bmatrix}1/2 & 1/2\\ 1 & 0\end{bmatrix}\quad\Rightarrow\quad\pi_1(n)=\frac{1}{2}\cdot\pi_1(n-1)+1\cdot\pi_2(n-1)$$

אם אכן ההנחה מתקיימת אז עבור n מאוד גדול מתקיים:

$$\pi_1(n)=\pi_1(n-1)=2/3\qquad\pi_2(n-1)=1-\pi_1(n-1)=1/3$$

נציב ונבדוק (ונראה שצדקנו):

$$2/3\overset{?}{=}\frac{1}{2}\cdot 2/3+1\cdot 1/3\quad\Rightarrow\quad True$$

<u>פילוג סטציונרי</u>

כפי שראינו במקרה החיפושית, קיימות שרשראות בהן הפילוג השולי של השרשרת מתכנס (כלומר קבוע לכל $n\ge k$ החל מערך k מסויים).

נאמר שלשרשרת מרקוב הומוגנית סופית יש פילוג סטציונארי אם קיים פילוג כזה, ז"א:

$$\exists k:\underline{\pi}(n)=\underline{\pi}(k)\quad\forall n\ge k$$

<u>ווקטור סטציונרי</u>

ווקטור $\underline{x}$ ייקרא ווקטור סטציונרי עבור מטריצה $\underline{\underline{A}}$ אם $\underline{x}=\underline{x}\cdot\underline{\underline{A}}$.

הערה - נבחין כי מטריצה $\underline{\underline{A}}$ חייבת להיות ריבועית ומאותו מימד של $\underline{x}$.

<u>ווקטור הפילוג הסטציונרי</u>

נסמן את ווקטור הפילוג הסטציונרי (ונקרא לו הפילוג הסטציונרי) של שרשרת מרקוב ע"י $\underline{\pi}'$

(או $\underline{\pi}^{\infty}$).

בשרשרת מרקוב הומוגנית סופית הפילוג הסטציונרי מקיים $\underline{\pi}' = \underline{\pi}' \cdot \underline{\underline{P}}$.

<u>סטציונריות בשרשראות מרקוב</u>

כזכור, הדרישה עבור סטציונריות במובן הצר (SSS) היא שהפילוג המשותף יהיה קבוע בזמן.
בשרשראות מרקוב אם למצב יש פילוג שולי אשר זהה לפילוג הסטציונרי אז גם לכל המצבים אחריו יהיה פילוג שולי זהה.
נגדיר עבור שרשראות מרקוב הומוגניות סופיות:

1. אם $\underline{\pi}(k) = \underline{\pi}'$ אזי מתקיים $\underline{\pi}(n) = \underline{\pi}'$ לכל $n \geq k$ והשרשרת נקראת סצטיונרית אסימפטוטית.

   המשמעות היא שאם בזמן מסויים k הפילוג השולי של המצב X_k זהה לפילוג הסטציונרי אז השרשרת הגיעה למצב סטציונרי והפילוג השולי של כל המצבים הבאים ($\forall n \geq k \quad X_n$) לא ישתנה.

2. אם $\underline{\pi}(0) = \underline{\pi}'$ אזי מתקיים $\forall n \in \mathbb{Z} \quad \underline{\pi}(n) = \underline{\pi}'$ והשרשרת נקראת סטציונרית.

   למעשה זו סטציונריות מלאה במובן הצר מכיוון שאנו מקבלים שהפילוג המשותף קבוע בזמן כי בכל נקודות הזמן הפילוג השולי של המצב קבוע – הוא מתחיל מהפילוג הסטציונרי.

   דרישה שקולה היא שפילוג ההתחלה $\underline{\pi}(0)$ יהיה ו"ע שמאלי של המטריצה $\underline{\underline{P}}$ עם ע"ע $\lambda = 1$.

<u>טענה:</u>

אם עבור שרשרת מרקוב עם פילוג סטציונרי $\underline{\pi}'$ קיים הגבול $\lim_{n\to\infty} \underline{\pi}(n) = \underline{\pi}''$ אזי $\underline{\pi}''$ חייב להתכנס לווקטור הסטציונרי של השרשרת, ז"א מתקיים $\underline{\pi}' = \underline{\pi}''$.

<u>הוכחה:</u>

מקיום הגבול $\lim_{n\to\infty} \underline{\pi}(n) = \underline{\pi}''$ אז גם מתקיים $\lim_{n\to\infty} \underline{\pi}(n+1) = \underline{\pi}''$. ולכן:

$$\lim_{n\to\infty}\underline{\pi}(n)=\lim_{n\to\infty}\underline{\pi}(n+1)$$
$$\lim_{n\to\infty}\underline{\pi}(n)=\lim_{n\to\infty}\left[\underline{\pi}(n)\cdot\underline{\underline{P}}\right]$$
$$\lim_{n\to\infty}\underline{\pi}(n)=\lim_{n\to\infty}\underline{\pi}(n)\cdot\underline{\underline{P}}$$
$$\underline{\pi''}=\underline{\pi''}\cdot\underline{\underline{P}}$$

קיבלנו ש- $\underline{\pi''}$ הוא ווקטור סטציונרי של מטריצה $\underline{\underline{P}}$ ומיחידות הגבול $\underline{\pi'}=\underline{\pi''}$.

<u>אפיון שרשראות מרקוב</u>

נרצה לענות על השאלה הטבעית מתי שרשרת מרקובית תתנהג באופן כזה בו אין תלות במצב ההתחלה (כמו החיפושית), ז"א נרצה לדעת מתי שרשרת תתכנס לפילוג שולי סטציונרי. כמובן שנדון רק בשרשראות מרקוב הומוגניות וסופיות. לצורך אפיון השרשרת נסווג את מצבי המערכת לפי דיאגרמת המצבים (או מטריצת המעברים).

<u>הגדרות סיווג:</u>

1. <u>נגישות:</u> נאמר שמצב j נגיש ממצב i ונסמן $i\to j$ אם יש מסלול כלשהו (במספר שלבים סופי) ממצב i למצב j .
2. <u>קשירות:</u> נאמר שמצב i קשור למצב j ונסמן $i\leftrightarrow j$ אם $i\to j$ וגם $j\to i$.
   הערה – קשירות הינה תכונה טרנזטיבית, משמע אם $i\leftrightarrow j$ וגם $j\leftrightarrow k$ אז $i\leftrightarrow k$.
3. <u>מחלקה:</u> מחקה היא אוסף מקסימלי של מצבים קשורים.
   במחלקה כל המצבים קשורים אחד לשני ואין אף מצב אשר קשור לאיבר אשר לא נמצא במחלקה.
4. <u>מצב נשנה:</u> מצב הנגיש מכל מצב הנגיש ממנו (מצב הנגיש רק למצבים שהוא קשור איתם), ז"א מצב שניתן לחזור אליו אם עזבנו אותו.
5. <u>מצב חולף:</u> מצב שאינו נשנה. ז"א מצב שניתן להמשיך ממנו למצב שלא יאפשר לחזור.
   מצב חולף יופיע מספר סופי של פעמים.
6. <u>מחלקה נשנית:</u> מחלקה שכל איבריה נשנים. על מנת לבדוק האם מחלקה היא מחלקה נשנית מספיק לבדוק עבור מצב בודד בה האם הוא מצב נשנה.
7. <u>מחלקה חולפת:</u> מחלקה שאינה נשנית. מחלקה שכל איבריה חולפים. על מנת לבדוק האם מחלקה היא

מחלקה חולפת מספיק לבדור עבור מצב בודד בה האם הוא מצב חולף.

8. <u>זמני חזרה:</u> זמני חזרה של מצב i הם אוסף אורכי המעגלים היוצאים ממצב i וחוזרים אליו.

זמני החזרה (אם קיימים) הם סדרה אינסופית ומסומנת ע"י $n_1^i, n_2^i, \cdots$.

9. <u>מחזור:</u> מחזור של מצב i הוא המחלק המשותף הגדול ביותר של סדרת זמני החזרה של המצב, ז"א:

$$d(i) = GCD\left(\left\{n_k^i\right\}_{k=1}^{\infty}\right)$$

הערה - ניתן לסווג מחלקות (חולפת/נשנית) לפי מטריצת המעברים עפ"י הזיהוי הבא:

i. מזהים מחלקות נשנות ע"י איתור תת-מטריצות ריבועיות כך שבאותן השורות יש רק אפסים מלבד איברי תת-המטריצה (ממצבי שורות תת-המטריצות לא ניתן לעבור למצבים מחוץ לה).

ii. מזהים מחלקות חולפות ע"י איתור תת-מטריצות ריבועיות כך שבאותם הטורים יש רק אפסים מלבד איברי תת-המטריצה (אל מצבי שורות תת-המטריצות לא להגיע מאף מצב מחוץ לה).

<u>מחלקה מחזורית</u>

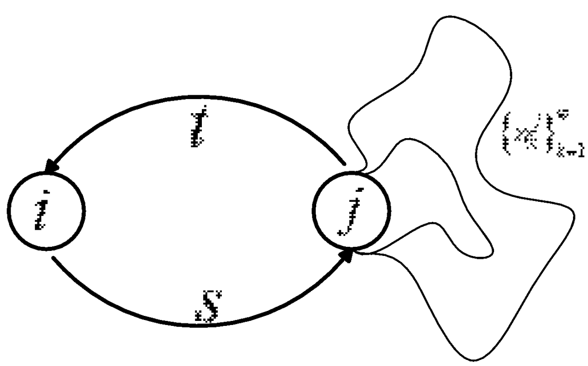
מחלקה תיקרא מחלקה מחזורית אם למצב i אשר חבר בה מתקיים $d(i) > 1$.

<u>טענה</u> - לכל המצבים באותה המחלקה יש אותו מחזור.

<u>הוכחה</u> - ניקח שני מצבים i,j מאותה מחלקה ונסמן $d(i), d(j)$.

נראה שמתקיים שהמחזור $d(i)$ הוא מחלק של המחזור $d(j)$ ולהיפך (ולכן $d(i) = d(j)$).

נאמר שקיים מסלול ממצב i למצב j באורך s וגם כי קיים מסלול הפוך באורך t .



מהגדרת מחזור של מצב אנו יודעים כי $d(i)$ הוא מחלק של כל זמני החזרה של המצב i ולכן:

$$\begin{matrix} d(i) \mid s+t \\ d(i) \mid s+t+n_k^j \quad \forall k \end{matrix} \quad \Rightarrow \quad d(i) \mid n_k^j \quad \forall k \quad \Rightarrow \quad d(i) \mid d(j)$$

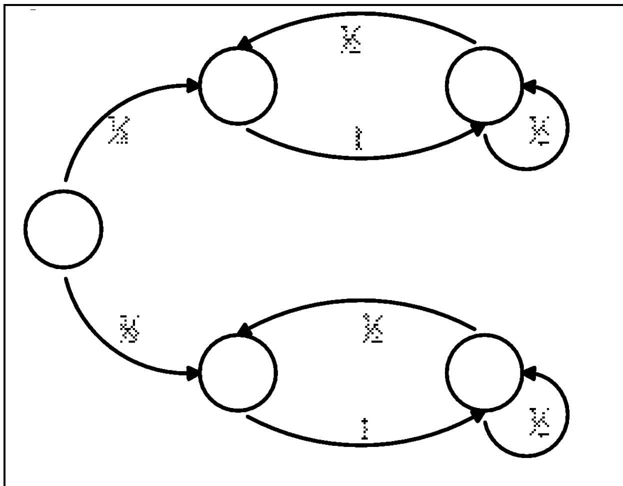
נהפוך תפקידים בין המצבים ונחשב באותו אופן ונקבל $d(j) \mid d(i)$.

מאחר שקיבלנו $d(i) \mid d(j)$ וגם $d(j) \mid d(i)$ (ז"א הם מחלקים אחד של השני) אז $d(i) = d(j)$.

<u>דוגמא 1:</u>

נתונה דיאגרמת המצבים (או המטריצה – ניתנות להמרה פשוטה):

$$\underline{\underline{P}} = \begin{bmatrix} 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$



נבצע סיווג לפי מחלקות (נזכור כי במחלקה חולפת כל המצבים חולפים ובמחלקה נשנית כל המצבים נשנים):

מחלקות (נזהה לפי תתי-מטריצות ריבועיות עם איברי שורות/עמודות אפסים):

- $T_1 = \{1\}$ - מחלקה חולפת
- $T_2 = \{2,3\}$ - מחלקה נשנית
- $T_3 = \{4,5\}$ - מחלקה נשנית

שתי תתי-המטריצות של המחלקות הנשנות זהות. בנוסף לתת-מטריצות אלו קיים פילוג סטציונרי (לא לשרשרת כולה).

$$\lim_{n\to\infty}\left(\left(\begin{bmatrix}0 & 1\\ 1/2 & 1/2\end{bmatrix}\right)^n\right)=\begin{bmatrix}1/3 & 2/3\\ 1/3 & 2/3\end{bmatrix}$$

ועבור המטריצה כולה:

$$\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^n\right)=\begin{bmatrix}0 & 1/9 & 2/9 & 2/9 & 4/9\\ 0 & 1/3 & 2/3 & 0 & 0\\ 0 & 1/3 & 2/3 & 0 & 0\\ 0 & 0 & 0 & 1/3 & 2/3\\ 0 & 0 & 0 & 1/3 & 2/3\end{bmatrix}$$

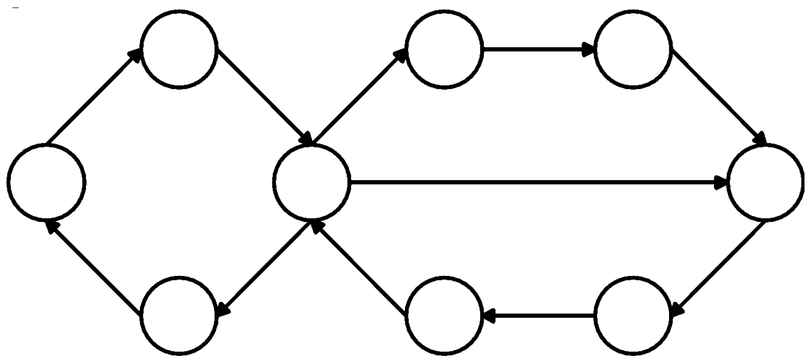
ערך ווקטור השורה הראשונה שווה ל (בהתאם לסיכויי המעבר לכל אחת מהמחלקות והפילוג הסטציונרי בהן):

$$\frac{1}{3}\cdot\underline{\pi}^{\infty}\left(T_2\right)+\frac{2}{3}\cdot\underline{\pi}^{\infty}\left(T_3\right)$$

שתי המחלקות הנשנות אינן מחזוריות מאחר ויש מסלול באורך אחד (ממצב 3 לעצמו וכנ"ל לגביי איבר 5).

<u>דוגמא 2 (חישוב מחזורים בלבד):</u>

נתונה דיאגרמת המצבים:



נזהה כי כל המצבים קשירים וכולם מהווים מחלקה נשנית אחת.

נחשב מסלולים עבור אחד המצבים, נגיד עבור מצב 1:

$$n_1^1 = 4, \quad n_2^1 = 8, \quad n_3^1 = 10, \quad n_4^1 = 12, \quad n_5^1 = 14, \ldots$$

נזהה כי לעולם זמני החזרה ימשיכו להיות זוגיים, ולכן:

$$d(1) = GCD(4,8,10,12,14,\ldots) = 2$$

ז"א המחזור של המצב הוא 2, ולכן מחזור המחלקה כולה הוא 2, מחלקה זו מהווה את כל המצבים בשרשרת ולכן זהו גם מחזור השרשרת.

"שיכחת העבר"

ראינו בדוגמת החיפושית כי אין חשיבות למצב ההתחלתי שלה. כלומר, אין תלות בתנאי ההתחלה.

נגדיר התנהגות זאת כ"שיכחת העבר" (חוסר זיכרון כלפי מצבי העבר).

נאמר ששרשרת מרקובית שוכחת את העבר (ארגודית) אם:

$$\lim_{n\to\infty} \Pr\left(X_n = j \middle| X_0 = i\right) = \pi_j^\infty$$

הדבר צריך להתקיים עבור כל מצב אפשרי $j \in \{1,\ldots,J\}$ בשרשרת ומצב התחלתי i.

לכן אם שרשרת הינה ארגודית אז בכתיב אחר:

$$\lim_{n\to\infty}\left(\underline{\pi}(n)\right) = \lim_{n\to\infty}\left(\underline{\pi}(0)\cdot\left(\underline{\underline{P}}\right)^n\right) = \underline{\pi}^\infty = \left(\underline{\pi}_1^\infty,\ldots,\underline{\pi}_J^\infty\right)$$

כאשר $\underline{\pi}(0)$ הוא וקטור המייצג את תנאי ההתחלה. לרוב נדע את מצבה ההתחלתי של השרשרת באופן דטרמיניסטי ואז יתקיים $\underline{\pi}(0) = \underline{e}_i$, כלומר הפילוג השולי של המצב ההתחלתי הוא ווקטור יחידה ממימד J (ווקטור אפסים עם 1 במקום ה- i אשר מייצג הסתברות 1 להיות במקום ה- i בזמן 0). ייתכן גם פילוג לא דטרמיניסטי.

אם הגבול הנ"ל קיים אז הוא בהכרח מקיים את תכונת הווקטור הסטציונרי של השרשרת, כלומר $\underline{\pi}^{\infty}=\underline{\pi}^{\infty}\cdot\underline{\underline{P}}$.

אם השרשרת ארגודית אז הגבול צריך להתקיים עבור כל פילוג התחלה $\underline{\pi}(0)$, ובפרט עבור פילוגי התחלה המייצג ידיעה דטרמיניסטית של המצב ההתחלתי – משמע ווקטור יחידה מהבסיס הסטנדרטי $\underline{e}_i$. לכן מתקיים:

$$\lim_{n\to\infty}\left(\underline{e}_i\cdot\left(\underline{\underline{P}}\right)^n\right)=\underline{\pi}^{\infty}$$

המשמעות של מכפלת ווקטור יחידה $\underline{e}_i$ במטריצה היא קבלת השורה ה- i מן המטריצה.

ומכאן שכל שורותיה של $\left(\underline{\underline{P}}\right)^n$ צריכות להיות שוות לווקטור הסטציונרי $\underline{\pi}^{\infty}$.
נאמר ששרשרת מרקוב הומוגנית וסופית שוכחת את העבר (ארגודית) אם"ם מטריצת המעברים שלה $\underline{\underline{P}}$ מקיימת:

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

נבחין שהדבר לא מונע מהפילוג השולי של המצב ההתחלתי להיות ווקטור שונה מווקטור יחידה מאחר שסך איבריו הוא אחד ושורות המטריצה $\left(\underline{\underline{P}}\right)^n$ זהות כולן אם השרשרת ארגודית.
כמובן שנוכל להבחין כי הדבר התקיים עבור דוגמת החיפושית.

ארגודיות (שיכחת העבר)

תהליך אקראי יקרא ארגודי (Ergodic - בעל תכונת הארגודיות) אם עבור מדגם ארוך מספיק כל פונקציה סטטיסטית זמנית מתכנסת לסטטיסטיקה של התהליך. במילים אחרות, ת"א הינו ארגודי אם ניתן ללמוד את הפילוג המשותף המלא שלו (משמע את כל הפרמטרים הסטטיסטיים שלו) מתוך התבוננות בפונקציית מדגם (ארוכה מספיק) בודדת שלו.
באופן זה, עבור ת"א ארגודי, נוכל להסיק תכונות סטטיסטיות לפי ממוצע מדגם וללא ידע סטטיסטי מקדים על התהליך. נדגיש כי עבור ת"א ארגודי נקבל התכנסות ממדגם יחיד שלו.
באופן כללי, עיבודים של תהליכים אקראיים (או תכנון מערכות) מתבססים לא פעם על ידיעת הסטטיסטיקה שלהם. עבור ת"א שהסטטיסטיקה שלו לא ידועה (או ידועה באופן לא מספיק) נוכל "למדוד" את הסטטיסטיקה שלו בעזרת מדגם ארוך מספיק אם הוא ארגודי.

ארגודיות בשרשראות מרקוב

ראינו כי ישנן שרשראות מרקוב אשר בעלות תכונת הארגודיות וגם כאלה שלא. נרחיב את הדיון בנושא ארגודיות בשרשראות מרקוב (הומוגניות וסופיות בלבד).

משפט פרון-פרוביניוס (Perron-Frobenius Theorem)

עבור שרשרת מרקוב הומוגנית וסופית בעלת מטריצת מעברים $\underline{\underline{P}}$ מתקיים:

1. תמיד קיים (לפחות אחד) פתרון הסתברותי למשוואה $\underline{\pi}=\underline{\pi}\cdot\underline{\underline{P}}$. כלומר תמיד קיים למטריצת המעברים $\underline{\underline{P}}$ ו"ע שמאלי אי-שלילי עם ע"ע $\lambda=1$ (נובע מתכונות של מטריצה סטוכסטית).
2. אם לשרשרת יש מחלקה נשנית יחידה אז הפתרון למשוואה $\underline{\pi}=\underline{\pi}\cdot\underline{\underline{P}}$ הוא יחיד.
3. אם לשרשרת יש r מחלקות נשנות אז למשוואה $\underline{\pi}=\underline{\pi}\cdot\underline{\underline{P}}$ יש r פתרונות בלתי תלויים ליניאריים.
כל קומבינציה הסתברותית (קומבינציה ליניארית עם מקדמים אי-שליליים שסכומם 1) של פתרונות אלו (הווקטורים הסטציונרים) הינה ו"ע של המטריצה, המשמעות היא שהווקטורים הסטציונרים האלו פורסים את כל הווקטורים העצמיים של המטריצה $\underline{\underline{P}}$.
כאמור, הווקטורים הסטציונרים האלו בת"ל ולכן ייראו מהצורה (סידרנו את המטריצה $\underline{\underline{P}}$ מטעמי נוחות):

$$\underline{\underline{P}}=\begin{bmatrix} * & * & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{bmatrix} \qquad \begin{matrix} \underline{\pi}_1^{Stat}=\left(*,*,0,0,0\right) \\ \underline{\pi}_2^{Stat}=\left(0,0,*,*,*\right) \end{matrix}$$

4. השרשרת היא ארגודית אם יש לה רק מחלקה אחת והיא נשנית ולא מחזורית ($d=1$).
במקרה זה השרשרת סטציונרית אסימפטוטית והיא מתכנסת לפילוג $\underline{\pi}^{\infty}$ עבור כל $\underline{\pi}(0)$.
5. השרשרת היא ארגודית עם תופעת מעבר (transient) אם יש לה מחלקה נשנית יחידה (שאינה מחזורית) ובנוסף יש בשרשרת מצבים חולפים (או מחלקות חולפות כמובן).
במצב זו מובטח כי אם נתחיל במחלקה הנשנית נישאר בה ונתנוון למקרה 4 ואם נתחיל באחד המצבים החולפים אז בנקודת זמן כלשהי בוודאות נגיע למחלקה הנשנית.
במקרה זה נוכל ללמוד מפונקציית מדגם יחידה רק על המחלקה הנשנית (ועל חלק מהחולפים אך לא בוודאות מספיק בשביל סטטיסטיקה).

- עבור שרשרת ארגודית (עם או בלי transient) מתקיים:

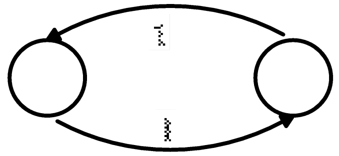
$$\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^n\right)=\begin{bmatrix}\leftarrow & \underline{\pi}^{\infty} & \rightarrow \\ & \vdots & \\ \leftarrow & \underline{\pi}^{\infty} & \rightarrow\end{bmatrix}$$

נרחיב את המשפט עבור זיהוי שרשראות שאינן ארגודיות:

6. אם לשרשרת יש שתי מחלקות נשנות או יותר אז היא לא ארגודית.
במקרה זה אם נתחיל במחלקה נשנית כלשהי אז נישאר בה ואם נתחיל במצב חולף אז נגיע לאחת מהמחלקות הנשנית אשר קשירות ממנו. לכן לא נוכל ללמוד את הסטטיסטיקה של כל המצבים ממדגם יחיד.
אם במקרה זה כל המחלקות הנשנות אינן מחזוריות אז $\left(\underline{\underline{P}}\right)^n$ מתכנסת אבל שורותיה אינן זהות.
7. אם לשרשרת מחלקה נשנית אחת מחזורית d אז היא לא ארגודית. למקרה זה נוכל גם להוסיף תופעת מעבר.
בגלל שהמחלקה הנשנית מחזורית אז $\left(\underline{\underline{P}}\right)^n$ לא מתכנסת (לא קיים הגבול $\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^n\right)$).
תופעת המחזוריות מונעת את הסטציונריות אך נקבל כי $\forall i=0,1,\ldots,d-1$ $\left(\underline{\underline{P}}\right)^{d\cdot n+i}$ כן מתכנסת.

<u>דוגמאות לשרשראות מרקוב לא ארגודיות:</u>

1. נתונה השרשרת המחזורית הבאה:

לשרשרת שני מצבים עם מעברים וודאיים. ברור לנו שהיא לא מתכנסת הרי היא קופצת ויש חשיבות לזוגיות של הזמן. אם $\underline{\pi}(0)=(1,0)$ אז בזמנים א"ז נהיה במצב 2 ובזמנים זוגיים במצב 1.	$\underline{\underline{P}}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$	

לשרשרת מחלקה נשנית יחידה עם מחזור 2, לכן לפי משפט PF השרשרת לא ארגודית.

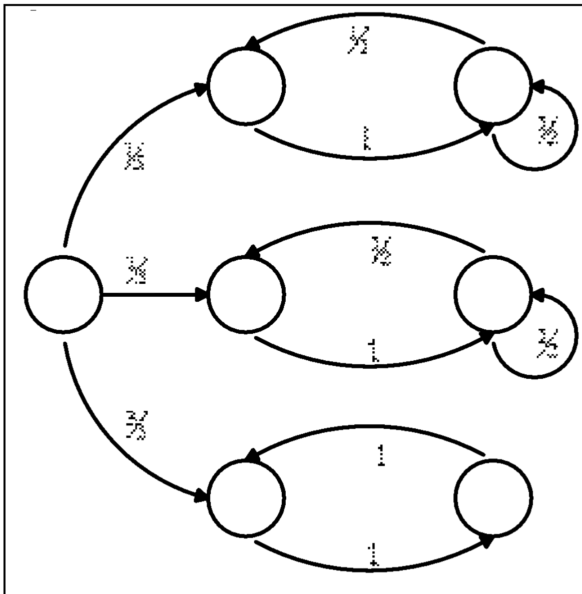
ממטריצת המעברים נראה כי לא קיים הגבול $\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^n\right)$, אך קיימים הגבולות עבור $\left(\underline{\underline{P}}\right)^{2\cdot n},\left(\underline{\underline{P}}\right)^{2\cdot n+1}$.

$$\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^{2\cdot n}\right)=\lim_{n\to\infty}\left(\left(\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)^{2\cdot n}\right)=\lim_{n\to\infty}\left(\left(\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)^{n}\right)=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

$$\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^{2\cdot n+1}\right)=\lim_{n\to\infty}\left(\left(\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)^{2\cdot n+1}\right)=\lim_{n\to\infty}\left(\left(\begin{bmatrix}0&1\\1&0\end{bmatrix}\right)^{2\cdot n}\right)\cdot\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

והגבולות שונים. בכל מקרה קיים למשוואה $\underline{\pi}=\underline{\pi}\cdot\underline{\underline{P}}$ פתרון והוא $\underline{\pi}^{stat}=\left(\tfrac{1}{2},\tfrac{1}{2}\right)$.

2. נתונה השרשרת הבאה:

$\underline{\underline{P}}=\begin{bmatrix}0&\tfrac{1}{5}&0&\tfrac{1}{5}&0&\tfrac{3}{5}&0\\0&0&1&0&0&0&0\\0&\tfrac{1}{2}&\tfrac{1}{2}&0&0&0&0\\0&0&0&0&1&0&0\\0&0&0&\tfrac{1}{2}&\tfrac{1}{2}&0&0\\0&0&0&0&0&1&0\\0&0&0&0&0&0&1\end{bmatrix}$	

נזהה מחלקות לפי מטריצת המעברים (מחלקה חולפת, מחלקות נשנות):

$T_1=\{1\},T_2=\{2,3\},T_3=\{4,5\},T_4=\{6,7\}$

נחשב מחזורים ונקבל $d\left(T_2\right)=1,\quad d\left(T_3\right)=1,\quad d\left(T_4\right)=2$.

זו שרשרת עם תופעת מעבר ומספר מחלקות נשנות. בנוסף, לאחת המחלקות הנשנות שלה יש מחזור.

בגלל המחזור למטריצה $\underline{\underline{P}}$ לא קיים הגבול $\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^n\right)$.

יש לנו שלוש מחלקות נשנות ולכן יש למשוואה $\underline{\pi}=\underline{\pi}\cdot\underline{\underline{P}}$ שלושה פתרונות, אחד עבור כל מחלקה ונסמנם בהתאם לשמות המחלקות $\underline{\pi}_{T_2}^{stat},\underline{\pi}_{T_3}^{stat},\underline{\pi}_{T_4}^{stat}$ (אך $\underline{\pi}_{T_4}^{stat}$ מתנדנד ותלוי בזמן ולכן גם ערכו שונה) ווקטור הפילוג הסטציונרי מורכב מקומבינציה הסתברותית שלהם, כלומר כל ווקטור אשר עונה על:

$$\underline{\pi}^{stat}=p_1\cdot\underline{\pi}_{T_2}^{stat}+p_2\cdot\underline{\pi}_{T_3}^{stat}+p_3\cdot\underline{\pi}_{T_4}^{stat}\quad p_1,p_2,p_3\geq 0\quad p_1+p_2+p_3=1$$

כמובן שלא נשאף לפילוג הסטציונרי הזה עבור כל סט $\left(p_1,p_2,p_3\right)$, הדבר תלוי במקדמים ובתנאי ההתחלה (המשמעות היא רק שכל סט כזה יניב ווקטור אשר פותר את המשוואה $\underline{\pi}=\underline{\pi}\cdot\underline{\underline{P}}$).

נראה את הגבולות עבור $\left(\underline{\underline{P}}\right)^n$:

$$\lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^{2\cdot n}\right)=\begin{bmatrix} \multicolumn{7}{c}{\tfrac{1}{5}\cdot\underline{\pi}_{T_2}^{stat}+\tfrac{1}{5}\cdot\underline{\pi}_{T_3}^{stat}+\tfrac{3}{5}\cdot\underline{\pi}_{T_4}^{stat\ even}} \\ 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}\qquad \lim_{n\to\infty}\left(\left(\underline{\underline{P}}\right)^{2\cdot n+1}\right)=\begin{bmatrix} \multicolumn{7}{c}{\tfrac{1}{5}\cdot\underline{\pi}_{T_2}^{stat}+\tfrac{1}{5}\cdot\underline{\pi}_{T_3}^{stat}+\tfrac{3}{5}\cdot\underline{\pi}_{T_4}^{stat\ odd}} \\ 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & \tfrac{1}{3} & \tfrac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

נבחין בהתנדנדות עבור המחלקה T_4, ולכן גם בפילוג הסטציונרי שלה.

<u>הגדרה פורמלית עבור ארגודיות</u>

ראינו כי ת"א הינו ארגודי אם ניתן ללמוד את הפילוג המשותף המלא שלו (משמע את כל הפרמטרים הסטטיסטיים שלו) מתוך התבוננות בפונקציית מדגם (ארוכה מספיק) בודדת שלו. למעשה תהליך אקראי הינו ארגודי אם"ם המיצוע הזמני של כל פונקציה של ערכי התהליך מתכנס לתוחלת הפונקציה, זאת אומרת אם"ם ממוצע זמני שווה לממוצע סטטיסטי עבור פונקציה $g:\mathbb{R}^M\to\mathbb{R}$ כלשהי ולכל סט זמנים $k_1,\ldots,k_M\in\mathbb{Z}$, כלומר:

$$\lim_{N\to\infty}\left(\frac{1}{N}\cdot\sum_{i=0}^{N-1}g\left(X_{k_1+i},\dots,X_{k_M+i}\right)\right)\overset{*}{=}\lim_{N\to\infty}\left(E\left[g\left(X_{k_1+N},\dots,X_{k_M+N}\right)\right]\right)\overset{**}{=}E\left[g\left(X_{k_1},\dots,X_{k_M}\right)\right]$$

ועבור ת"א בזמן רציף הדבר צריך להתקיים לכל סט זמנים $t_1,\dots,t_M\in\mathbb{R}$:

$$\lim_{T\to\infty}\left(\frac{1}{T}\cdot\int_{\tau=0}^{T}g\left(X\left(t_1+\tau\right),\dots,X\left(t_M+\tau\right)\right)d\tau\right)\overset{*}{=}\lim_{N\to\infty}\left(E\left[g\left(X\left(t_1+\tau\right),\dots,X\left(t_M+\tau\right)\right)\right]\right)=$$
$$\overset{**}{=}E\left[g\left(X\left(t_1\right),\dots,X\left(t_M\right)\right)\right]$$

כאשר שיוויונים $(*)$ נובעים מסטציונריות אסימפטוטית ושיוויונים $(**)$ נובעים מסטציונריות.
למעשה ארגודיות היא הכלל של חוק המספרים הגדולים (Law of Large Numbers) כי מתקיימת שם התכנסות של ממוצע זמני לממוצע סטטיסטי. לדוגמא עבור מערכת עם הסתברות שגיאה 0.1 לביט כאשר סדרת הביטים ארגודית נוכל להבטיח 10% שגיאה כללית בזמן (למשך זמן ארוך מספיק).

<u>חוק המספרים הגדולים (LLN)</u>

עבור סדרה $\{X_n\}$ ת"א i.i.d מתקיים:

$$\lim_{N\to\infty}\left(\frac{1}{N}\cdot\sum_{i=0}^{N-1}X_n\right)=E\left[X_n\right]$$

נבחין כי LLN הוא מקרה פרטי של ארגודיות בו $g(x)=x$ (ואז תוכן הגבול הוא ממוצע זמני אשר שואף לתוחלת).

על מנת לבדוק האם חייבים שהסדרה $\{X_n\}$ תהיה i.i.d בשביל קיום LLN נבחן את הפונקציה $g(x,y)=x\cdot y$ עבור ת"א $\{X_n\}$ סטציונרי וארגודי. נבחר סט זמנים $k_1=1,k_2=3$ ונקבל:

$$\lim_{N\to\infty}\left(\frac{1}{N}\cdot\sum_{i=0}^{N-1}\left(X_{1+n}\cdot X_{3+n}\right)\right)=\lim_{N\to\infty}\left(\frac{1}{N}\cdot\sum_{i=0}^{N-1}\left(X_{n+1}\cdot X_{n+3}\right)\right)\overset{*}{=}E\left[\left(X_1\cdot X_3\right)\right]=$$
$$=R_{XX}\left(n+1,n+3\right)\overset{**}{=}R_{XX}\left(-2\right)\overset{***}{=}R_{XX}\left(2\right)$$

שיוויון $(*)$ מארגודיות, שיוויון $(**)$ מסטציונריות ושיוויון $(***)$ מתכונות של פונקציית קורולוציה (זוגיות).

באופן כללי עבור סט זמנים k_1,k_2 נקבל:

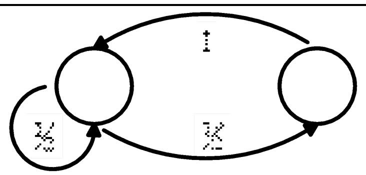
$$\lim_{N\to\infty}\left(\frac{1}{N}\cdot\sum_{i=0}^{N-1}\left(X_{k_1+n}\cdot X_{k_2+n}\right)\right)=R_{XX}\left(\left|k_2-k_1\right|\right)$$

קיבלנו תכונת זיכרון, לכן נוכל להחליש את הדרישה וקיום ארגודיות יאפשר לנו התכנסות ממוצע זמני לממוצע סטטיסטי גם עבור תהליכים אקראיים אשר אינם i.i.d.

מציאת סטטיסטיקה של שרשרת מרקוב ארגודית

מהגדרת הארגודיות, עבור ת"א ארגודי ניתן למצוא את כל הפרמטרים הסטטיסטיים שלו מידיעת פונקציית מדגם יחידה ארוכה ככל שנרצה. נרצה ללמוד פילוג מסדר כלשהו עבור שרשרת מרקוב ארגודית מתוך התבוננות בפונקציית מדגם.

נראה איך לשערך את הפילוג הסטציונרי ומטריצת המעברים בהינתן פונקציית מדגם של שרשרת מרקוב ארגודית. נבצע זאת על שרשרת החיפושית (אנו יודעים עליה שהיא ארגודית וסטציונרית), תזכורת:

	$\underline{\underline{P}}(n)=\begin{bmatrix}1/2 & 1/2\\ 1 & 0\end{bmatrix}$ $\underline{\pi}^{stat}=\left(2/3,1/3\right)$	

ונניח פונקציית מדגם כלשהי שלה (האיברים מסומנים באות קטנה כי זו ריאליזציה של התהליך):

$$x_{n=0,1,\dots}=1,1,2,1,2,1,2,1,1,2,\dots$$

שערוך הפילוג הסטציונרי של שרשרת מרקוב ארגודית

נרצה לשערך את הפילוג הסטציונרי $\underline{\pi}^{stat}$, נחשבו את איברי הווקטור תא-תא. הסתברות כל תא תהיה שכיחות הופעת המצב המתאים בפונקציית המדגם (אנחנו משתמשים בפונקציית אינדיקטור $I_a(x)$):

$$\hat{\pi}_a^{stat}(N)=\frac{1}{N}\cdot\sum_{i=0}^{N-1}I_a(x_i)\qquad\qquad I_a(x)=\begin{cases}1 & x=a\\ 0 & else\end{cases}$$

באופן זה עבור שערוך $\hat{\pi}_a^{stat}(N)$ אנחנו סופרים את מספר הפעמים שהופיע המצב a בתוך N הדגימות ומחלקים ב- N, ז"א ממש מיצוע לכמות ההופעות שלו.

שערוך מטריצת המעברים של שרשרת מרקוב ארגודית (והומוגנית)

נרצה לשערך את מטריצת המעברים $\underline{\underline{P}}$, לשם כך נחשב את איברי המטריצה תא-תא.

נשערך את הסתברות המעבר $\underline{\underline{P}}(a,b)$ (הסתברות המעבר ממצב a למצב b) באופן הבא:

i. נשערך את הפילוג המשותף $\Pr\left(X_n=a,X_{n+1}=b\right)$ ע"י סכימת ההופעות של המצב b אחרי המצב a , נבחין כי עבור סדרה בגודל N יש $N-1$ מעברים (אנחנו משתמשים בפונקציית אינדיקטור כפול $I_{ab}\left(x\right)$):

$$\hat{p}\left(a,b\right)\left(N\right)=\frac{1}{N-1}\cdot\sum_{i=0}^{N-2}I_{ab}\left(x_i,x_{i+1}\right) \qquad I_{ab}\left(x,y\right)=\begin{cases}1 & x=a,y=b\\ 0 & else\end{cases}$$

ii. נחשב מן הסיכוי המשוערך הנ"ל את ההסתברות המותנית $\Pr\left(X_{n+1}=b\middle|X_n=a\right)$ (אשר מתבססת על השערוך כמובן) עפ"י הנוסחה להסתברות מותנית (ושימוש בשערוך ההסתברות למצב a).

$$\Pr\left(X_{n+1}=b\middle|X_n=a\right)=\frac{\Pr\left(X_{n+1}=b\cap X_n=a\right)}{\Pr\left(X_n=a\right)}=\frac{\Pr\left(X_n=a,X_{n+1}=b\right)}{\Pr\left(X_n=a\right)}$$

$$\Rightarrow\quad \hat{P}_{ab}\left(N\right)=\underline{\underline{\hat{P}}}\left(a,b\right)\left(N\right)=\frac{\hat{p}\left(a,b\right)\left(N\right)}{\hat{\pi}_a^{stat}\left(N\right)}$$

<u>דוגמא:</u>

עבור דוגמת החיפושית עם פונקציית המדגם (הקטנה - $N=10$) הנ"ל נשערך את הפילוג הסטציונרי:

$$\underline{\hat{\pi}}^{stat}\left(10\right)=\left(\hat{\pi}_1^{stat}\left(10\right),\hat{\pi}_2^{stat}\left(10\right)\right)=\left(\frac{6}{10},\frac{4}{10}\right)=\left(\frac{3}{5},\frac{2}{5}\right)=\left(0.6,0.4\right)$$

נחשב את סיכוי המעבר עבור תא בודד (המייצג את הסתברות המעבר ממצב 2 לעצמו):

$$\hat{P}_{2,2}\left(10\right)=\frac{\hat{p}\left(2,2\right)\left(10\right)}{\hat{\pi}_2^{stat}\left(10\right)}=\frac{0/9}{0.4}=0$$

עבור הפילוג הסטציונרי אכן מתקיים $\left(0.6,0.4\right)\approx\left(2/3,1/3\right)$ ולגביי התסברות המעבר בהחלט אין סיכוי כזה.

<u>הערות:</u>

- אם יש מחלקה נשנית אחת בלבד אז ניתן ללמוד את הפילוג המשותף מכל סדר (לפחות באופן אסימפטוטי) מתוך פונקציית מדגם בודדת ויתקיים:

$$\lim_{N\to\infty}\left(\hat{\pi}_a^{stat}(N)\right)=\pi_a^{stat}\quad\forall a\in\{1,\ldots,J\}\quad\Rightarrow\quad\lim_{N\to\infty}\left(\hat{\underline{\pi}}^{stat}(N)\right)=\underline{\pi}^{stat}$$

$$\lim_{N\to\infty}\left(\hat{P}_{ab}(N)\right)=\underline{\underline{P}}(a,b)\quad\forall a,b\in\{1,\ldots,J\}$$

אם אין סטציונריות אסימפטוטית (לפחות) אז לא נקבל התכנסות כזאת. נניח עבור מחלקה נשנית בת שני מצבים עם קפיצה עם מחזור 2 (ממצב ראשון למצב שני וממצב שני למצב ראשון) לא קיימת התכנסות ל- N כללי אלא רק בהפרדה לזוגי ואי-זוגי.

- אם יש בשרשרת יותר ממחלקה נשנית אחת אז לא נוכל ללמוד אותה מהסתכלות בפונקציית מדגם בודדת כי בהכרח נבקר בלכל היותר במחלקה נשנית אחת (לכל היותר ולא בהכרח מכיוון שיש את הסיכוי הקלוש להישאר בין מצבים חולפים) אשר אותה נלמד היטב.
- בשרשרת לא ארגודית לעולם לא נוכל ללמוד את הסטטיסטיקה שלה מהתבוננות בריאליזציה בודדת.

<u>הקשר בין ההגדרה האינטואיטיבית והפורמלית של ארגודיות</u>

ראינו בתחילת הנושא כי ארגודיות באופן אינטואיטיבי היא היכולת ללמוד סטטיסטיקה מלאה של ת"א על פי פונקציית מדגם יחידה שלו. בהמשך הנושא הצגנו הגדרה פורמלית בו ארגודיות נתפסה כהתכנסות הממוצע הזמני של פונקציה זמנית כלשהי על ת"א לממוצע הסטטיסטי (תוחלת) של אותה פונקציה על דגימותיו. נראה שההגדרות שקולות.

<u>כיוון ראשון:</u>

ידוע כי הממוצע הזמני מתכנס לממוצע סטטסיטי, נרצה להראות שנוכל ללמוד סטטיסטיקה של התהליך. נדגים עבור פילוג שולי מסדר 2 (שתי נקודות זמן של התהליך, נוכל עבור כל סדר באופן זהה), תוך שימוש בפונקציית אינדיקטור כפול:

$$\frac{1}{N}\cdot\sum_{i=0}^{N-1}I_{ab}\left(X_{k_1+i},X_{k_2+i}\right)\to E\left[I_{ab}\left(X_{k_1},X_{k_2}\right)\right]=\Pr\left(X_{k_1}=a,X_{k_2}=b\right)$$

הראינו שבידיעת העובדה כי ממוצע זמני מתכנס לממוצע סטטיסטי אנו יכולים להשתמש בתכונת הארגודיות על פונקצייה $g(x,y)=I_{ab}(x,y)$ ולקבל את הפילוג השולי $\Pr(x=a,y=b)$.

<u>כיוון שני:</u>

ידוע כי נוכל ללמוד את הסטטיסטיקה של התהליך, נרצה להוכיח שממוצע זמני מתכנס לממוצע סטטיסטי.

נרכיב פונקצייה מסויימת (שוב, נדגים עבור פונקציה של שתי דגימות של הת"א) בעזרת פונקציית אינדיקטור ע"י (אנו מדגימים עבור ת"א סופי מטעמי נוחות, נוכל גם לעבוד עם אינטגרל בתיאוריה):

$$g(x,y)=\sum_{a=1}^{J}\sum_{b=1}^{J}g(a,b)I_{ab}(x,y)$$

הדבר יתאים עבור כל פונקציה כי אנו סוכמים את כל ערכיה האפשריים כפול אינדיקציה רק לערכים הרלוונטים, ז"א רק כאשר יתקיים $(x,y)=(a,b)$ אז איבר הסכימה לא יהיה אפס וערכו יהיה $g(a,b)\cdot 1=g(x,y)$.

ומידיעת העובדה שהסטטיסטיקה מתכנסת נרצה להראות שמיצוע הפונקציה בזמן יתכנס לתוחלת הפונקציה:

$$\begin{aligned}\frac{1}{N}\cdot\sum_{i=0}^{N-1}g\left(X_{k_1+i},X_{k_2+i}\right)&=\frac{1}{N}\cdot\sum_{i=0}^{N-1}\left[\sum_{a=1}^{J}\sum_{b=1}^{J}g(a,b)I_{ab}\left(X_{k_1+i},X_{k_2+i}\right)\right]=\\&=\sum_{i=0}^{N-1}\left[\sum_{a=1}^{J}\sum_{b=1}^{J}g(a,b)\cdot\frac{I_{ab}\left(X_{k_1+i},X_{k_2+i}\right)}{N}\right]\to\sum_{i=0}^{N-1}\left[\sum_{a=1}^{J}\sum_{b=1}^{J}g(a,b)\Pr\left(X_{k_1}=a,X_{k_2}=b\right)\right]=\\&=E\left[g\left(X_{k_1},X_{k_2}\right)\right]\end{aligned}$$

וקיבלנו את ההתכנסות הרצויה.

<u>ארגודיות חלשה</u>

כזכור, ההגדרה הפורמלית עבור ארגודיות היא ההתכנסות של הממוצע הזמני של כל פונקציה לממוצע הסטטיסטי (תוחלת) שלה. ישנם תהליכים אשר לא מקיימים תכונה זאת עבור כל פונקציה אלא רק עבור חלק (ארגודיות חלשה).

נאמר שת"א $X(t)$ (או עבור ת"א בזמן רציף X_n) ארגודי ביחס לפונקציה מסויימת $g:\mathbb{R}^M\to\mathbb{R}$ אם מתקיימת ארגודיות ביחד לפונקציה זו, כלומר מתקיים:

$$\lim_{T\to\infty}\left(\frac{1}{T}\cdot\int_{\tau=0}^{T}g\left(X(t_1+\tau),\dots,X(t_M+\tau)\right)d\tau\right)=E\left[g\left(X(t_1),\dots,X(t_M)\right)\right]\quad\forall t_1,,t_M\in\mathbb{R}$$

עבור ההתכנסות הזאת נדרוש שהתהליך יהיה לפחות סטציונרי אסימפטוטית (ס"א).
מאחר שלא כל תהליך הוא ארגודי חזק למעשה נוכל לשערך גדלים מסויימים באופן לא מוצלח.

<u>שערוך פרמטרים של תהליך אקראי</u>

נניח כי אנו מנסים לשערך פרמטר מסויים של ת"א (מקדם α של ת"א A.R ליניארי, תוחלת של ת"א, פאזה וכדומה). נסמן את הפרמטר המשוערך ב- Θ ואת המשערך שלו ב- $\hat{\Theta}(T)$. נזהה כי Θ הוא גודל דטרמיניסטי (רק שאנחנו לא יודעים אותו) ואילו $\hat{\Theta}(T)$ הוא תהליך אקראי (מתבסס על דגימותיו של הת"א בזמן $[0,T]$).

<u>קיום גבול במובן MSE עבור סדרה אקראית:</u>

נרצה להעמיק את ההבנה לגביי קיום גבול עבור סדרה אקראית ולהבין מה המשמעות של סדרה אקראית אשר מתכנסת לקבוע. למעשה מה המשמעות של:

$$\lim_{T\to\infty}\left(\hat{\Theta}(T)\right)=\Theta$$

נאמר שת"א $X[n]$ (או $X(t)$ עבור ת"א בזמן רציף) שואף לקבוע α כאשר $N\to\infty$ אם מתקיים:

$$\lim_{N\to\infty}E\left[\left(X[N]-\alpha\right)^2\right]=0$$

ונסמן $X[n]\xrightarrow{MSE}\alpha$.

הערה - שאיפה במובן MSE גורר שאיפה בהסתברות.

נניח כי מתקיים $X[n]\xrightarrow{MSE}\alpha$ ולכן:

$$\Pr\left(\left|X[N]-\alpha\right|\geq\varepsilon\right)=\Pr\left(\left(X[N]-\alpha\right)^2\geq\varepsilon^2\right)\overset{\text{Markov's Inequality}}{\leq}\frac{E\left[\left(X[N]-\alpha\right)^2\right]}{\varepsilon^2}\xrightarrow[N\to\infty]{}0$$

קיבלנו שאיפה לאפס לכל ε, אז זו למעשה שאיפה בהסתברות.

כעת, כאשר התכנסנו לערך מסויים נרצה לדעת לבחון את איכות המשערך.

<u>משערך חסר הטיה אסימפטוטית:</u>

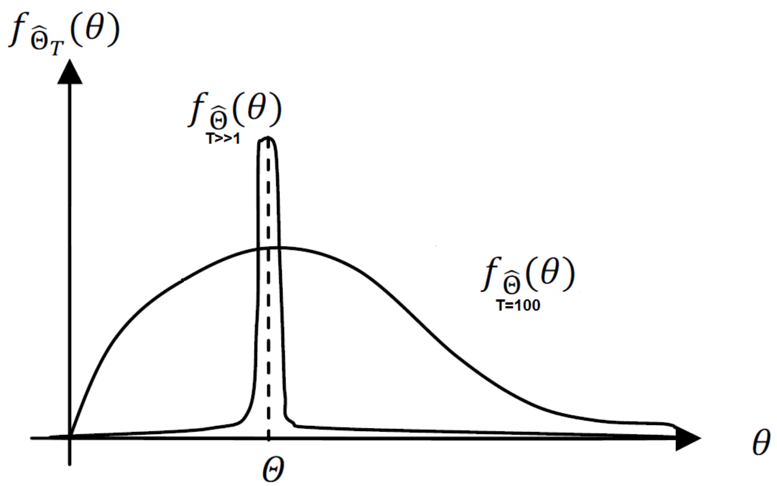
נאמר שהמשערך $\hat{\Theta}(T)$ (של הפרמטר Θ) הוא משערך חסר הטיה אסימפטוטית אם מתקיים $E\left[\hat{\Theta}(T)\right]\xrightarrow[T\to\infty]{}\Theta$.

<u>משערך עקבי:</u>

נאמר שהמשערך $\hat{\Theta}(T)$ (של הפרמטר Θ) הוא משערך עקבי אם:

1. $\hat{\Theta}(T)$ הוא משערך חסר הטיה אסימפטוטית
2. מתקיים $Var\left(\hat{\Theta}(T)\right)\underset{T\to\infty}{\to}0$

המשמעות של ההגדרה היא שמתקיים $\lim_{T\to\infty}\left(f_{\hat{\Theta}(T)}\left(x\right)\right)=\delta\left(x-\Theta\right)$. כלומר ככל ש- T גדל כך פילוג המשערך מתקרב להלם סביב הערך הדטרמיניסטי של הפרמטר.



<u>תהליך ארגודי בפרמטר:</u>

נאמר שת"א הוא ארגודי בפרמטר Θ אם קיים משערך $\hat{\Theta}(T)$ שהוא משערך עקבי של Θ.

<u>**ארגודיות בתוחלת ובקורולציה**</u>

נרחיב את הדיון בארגודיות חלשה ביחס לשתי פונקציות ספציפיות. אנו דורשים שהתהליך האקראי יהיה סטציונרי לפחות אסימפטוטית לצורך קיום הגבול (הפילוג השולי באינסוף לא תלוי בזמן).

<u>ארגודיות בתוחלת:</u>

ת"א $X\left(t\right)$ סטציונרי אסימפטוטי (ס"א) יקרא ארגודי ביחס לתוחלת (או בקצרה – ארגודי בתוחלת) אם הוא ארגודי ביחד לפונקציה הספציפית $g\left(x\right)=x$, כלומר מתקיים:

$$\lim_{T\to\infty}\left(\hat{\eta}_X\left(t_1\right)\left(T\right)\right)=\lim_{T\to\infty}\left(\frac{1}{T}\cdot\int_{\tau=0}^{T}X\left(t_1+\tau\right)d\tau\right)=E\left[X\left(t_1\right)\right]\underset{T\to\infty}{=}\eta_X^{stat}$$

מאחר והתהליך סטציונרי (לפחות אסימפטוטית) נוכל לרשום:

$$\lim_{T\to\infty}\left(\hat{\eta}_X\left(T\right)\right)=\lim_{T\to\infty}\left(\frac{1}{T}\cdot\int_{0}^{T}X\left(t\right)dt\right)=E\left[X\left(t\right)\right]=\eta_X^{stat}$$

<u>ארגודיות בקורולצייה:</u>

ת"א סטציונרי אסימפטוטי (ס"א) יקרא ארגודי ביחס לקורולצייה (או בקצרה – ארגודי בקורולצייה) אם הוא ארגודי ביחד לפונקציה הספציפית $g(x,y)=x\cdot y$, כלומר מתקיים:

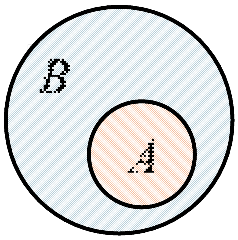
$$\lim_{T\to\infty}\left(\hat{R}_{XX}(t_1,t_2)(T)\right)=\lim_{T\to\infty}\left(\frac{1}{T}\cdot\int_{\tau=0}^{T}X(t_1+\tau)\cdot X(t_2+\tau)\,d\tau\right)=E\left[X(t_1)\cdot X(t_2)\right]=$$
$$=R_{XX}^{stat}(t_1,t_2)=R_{XX}^{stat}\left(\left|t_2-t_1\right|\right)$$

מאחר והתהליך סטציונרי (לפחות אסימפטוטית) נוכל לרשום:

$$\lim_{T\to\infty}\left(\hat{R}_{XX}(\tau)(T)\right)=\lim_{T\to\infty}\left(\frac{1}{T}\cdot\int_{\tau=0}^{T}X(t)\cdot X(t+\tau)\,dt\right)=R_{XX}^{stat}(\tau)$$

<u>תנאים לארגודיות בתוחלת</u>

נציג תנאים הכרחיים ותנאים מספיקים עבור ארגודיות בתוחלת של ת"א.

<table>
<tr>
<td>
<u>תנאי מספיק</u> –
תנאי מספיק הוא כאשר קיום טענה A גוררת בהכרח את קיום טענה B. נסמן ע"י $A \Rightarrow B$, ז"א מספיק לדעת כי $x \in A$ בשביל לקבוע כי $x \in B$.
טענה שקולה עבור $A \Rightarrow B$ היא $\neg B \Rightarrow \neg A$.
<u>תנאי הכרחי</u> –
תנאי הכרחי הוא כאשר על מנת לקבוע את קיום טענה A בהכרח חייבת להתקיים טענה B. לדוגמא עבור קיום $x \in A$ בהכרח חייב להתקיים $x \in B$. נבחין כי תנאי הכרחי הוא לא מספיק – לדוגמא אם מתקיים $x \in B$ אין הדבר אומר בהכרח כי מתקיים $x \in A$.
אם טענה B היא תנאי הכרחי לטענה A אז ניתן להסיק את התנאי המספיק $A \Rightarrow B$.
</td>
<td>

$A \subset B$
</td>
</tr>
<tr>
<td colspan="2">
<u>תנאי הכרחי ומספיק</u> –
קיום של שני התנאים יחד (A הוא תנאי הכרחי ומספיק ל- B) הוא למעשה שקילות לוגית בין טענה A וטענה B. הדבר מתקיים מאחר שמ- A ניתן להסיק את B (מספיק) ומ- B ניתן להסיק את A (הכרחי), כלומר $A \Leftrightarrow B$.
</td>
</tr>
</table>

נסמן את התהליך האקראי $X(t)$ (או $X[n]$ עבור ת"א בזמן בדיד). נניח לשם ה"ניקיון" כי התהליך סטציונרי.

ראינו כי על מנת לדרוש התכנסות אנו למעשה צריכים לדרוש שהמשערך יהיה עקבי (התוחלת תתכנס לערך המשוערך הדטרמיניסטי והשונות תתכנס לאפס). אנו דנים בפרמטר התוחלת.

משערך התוחלת בנוי על הפונקציה $g(x)=x$ ולכן הוא למעשה:

$$\eta_X(T)=\frac{1}{T}\cdot\int_0^T X(t)\,dt$$

נראה כי חוסר ההטיה מובטח עבור כל T (ללא השאפה $T\to\infty$) אם התהליך סטציונרי (תוחלת קבועה בזמן):

$$E\left[\eta_X(T)\right]=E\left[\frac{1}{T}\cdot\int_0^T X(t)\,dt\right]=\frac{1}{T}\cdot E\left[\int_0^T X(t)\,dt\right]=\frac{1}{T}\cdot\int_0^T E\left[X(t)\right]dt=$$

$$=\frac{1}{T}\cdot\int_0^T \eta_X(t)\,dt\overset{Stationarity}{=}\frac{1}{T}\cdot\int_0^T \eta_X dt=\frac{1}{T}\cdot\eta_X\cdot\int_0^T 1\cdot dt=\eta_X\cdot\frac{1}{T}\cdot T=\eta_X$$

נותר לבדוק את התנאי עבור התכנסות השונות לאפס. נרצה לדעת מתי מתקיים

$\lim_{T\to\infty}\left(Var\left(\hat{\eta}_X(T)\right)\right)=0$.

למעשה מתקיים $Var\left(\hat{\eta}_X(T)\right)=Var\left(\frac{1}{T}\cdot\int_0^T X(t)\,dt\right)=\ldots=\frac{1}{T}\cdot\int_{-T}^{T}C_{XX}(t)\cdot\left(1-\frac{|t|}{T}\right)dt$.

נראה כי הדבר אכן מתקיים (עבור ת"א בזמן בדיד מטעמי נוחות):

$$\hat{\eta}_X(T)=\frac{1}{N}\cdot\sum_{n=0}^{N-1}X[n]$$

$$Var\left(\hat{\eta}_X(T)\right)=E\left[\left(\hat{\eta}_X(T)-E\left[\hat{\eta}_X(T)\right]\right)^2\right]\overset{*}{=}E\left[\left(\hat{\eta}_X(T)-\eta_X\right)^2\right]=E\left[\left(\frac{1}{N}\cdot\sum_{n=0}^{N-1}X[n]-\eta_X\right)^2\right]=$$

$$\overset{**}{=}E\left[\left(\frac{1}{N}\cdot\sum_{n=0}^{N-1}\left(X[n]-\eta_X\right)\right)^2\right]=E\left[\frac{1}{N^2}\cdot\left(\sum_{n=0}^{N-1}\left(X[n]-\eta_X\right)\right)^2\right]=$$

$$=\frac{1}{N^2}\cdot E\left[\left(\sum_{n_1=0}^{N-1}\left(X[n_1]-\eta_X\right)\right)\left(\sum_{n_2=0}^{N-1}\left(X[n_2]-\eta_X\right)\right)\right]=\frac{1}{N^2}\cdot\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}E\left[\left(X[n_1]-\eta_X\right)\left(X[n_2]-\eta_X\right)\right]=$$

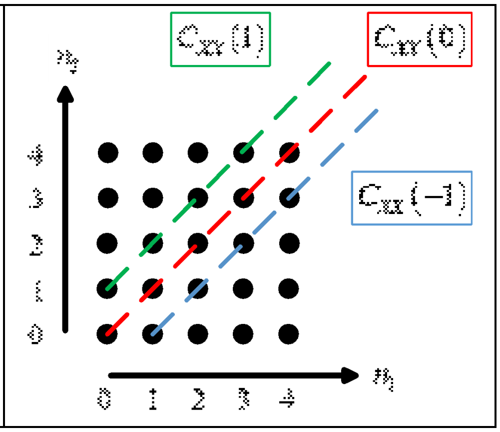
$$=\frac{1}{N^2}\cdot\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}C_{XX}(n_1,n_2)\overset{Stationarity}{=}\frac{1}{N^2}\cdot\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}C_{XX}(n_1-n_2)$$

שיוויון $(*)$ מתכונת חוסר ההטייה של התוחלת עבור ת"א סטציונרי. שיוויון $(**)$ משום שנכניס את קבוע התוחלת לסכום ולכן היא תחולק באורך הסכום ואז נוכל להוציא את השבר בכורם משותף.

אנו צריכים לסכום למעשה את הגודל $C_{XX}(n_1-n_2)$ עבור כל קומבינציה (n_1,n_2) אפשרית כאשר $n_1,n_2\in\mathbb{Z}$.

נבחר לסכום בנפרד את כל הקומבינציות האלו אשר יש להן אותו הפרש. ניתן לראות את נקודות אלו באופן גרפי ע"י אלכסונים (תזוזה מן האלכסון הראשי תקטין את מספר האיברים באלכסון באחד). באלכסון הראשי יש N איברים.

$$\frac{1}{N^2}\cdot\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}C_{XX}\left(n_1-n_2\right)=$$
$$\underset{n=n_1-n_2}{=}\frac{1}{N^2}\cdot\sum_{n=-(N-1)}^{N-1}C_{XX}\left(n\right)\cdot\left(N-\left|n\right|\right)=$$
$$=\frac{1}{N}\cdot\sum_{n=-(N-1)}^{N-1}C_{XX}\left(n\right)\cdot\left(1-\frac{\left|n\right|}{N}\right)$$



<u>טענה א' – סט תנאים:</u>

1. ארגודיות ביחס לתוחלת אם"ם (הכרחי ומספיק) מתקיים:

$$\frac{1}{N}\cdot\sum_{n=-(N-1)}^{N-1}C_{XX}\left(n\right)\cdot\left(1-\frac{\left|n\right|}{N}\right)\underset{N\to\infty}{\to}0,\quad \frac{1}{T}\cdot\int_{-T}^{T}C_{XX}\left(t\right)\cdot\left(1-\frac{\left|t\right|}{T}\right)dt\underset{T\to\infty}{\to}0$$

מאחר שפונקצייה האוטו-קו-ואריאנס הינה פונקציה זוגית תנאי שקול הוא:

$$\frac{2}{N}\cdot\sum_{n=0}^{N-1}C_{XX}\left(n\right)\cdot\left(1-\frac{\left|n\right|}{N}\right)\underset{N\to\infty}{\to}0,\quad \frac{2}{T}\cdot\int_{0}^{T}C_{XX}\left(t\right)\cdot\left(1-\frac{\left|t\right|}{T}\right)dt\underset{T\to\infty}{\to}0$$

וכמובן ש- 2 במכנה לא קריטי.

2. ארגודיות ביחס לתוחלת אם (תנאי מספיק) מתקיים:

$$\frac{1}{N}\cdot\sum_{n=-(N-1)}^{N-1}\left|C_{XX}\left(n\right)\right|\underset{N\to\infty}{\to}0,\quad \frac{1}{T}\cdot\int_{-T}^{T}\left|C_{XX}\left(\tau\right)\right|d\tau\underset{T\to\infty}{\to}0$$

מאחר שמקדם הקורולציה $\rho_{XX}\left(\tau\right)$ דומה לפונקציית הקורולציה עד כדי נירמול בהספק האות נוכל לרשום את התנאי המספיק באופן הבא:

$$\rho_{XX}\left(\tau\right)=\frac{C_{XX}\left(\tau\right)}{C_{XX}\left(0\right)}\quad\Rightarrow\quad\frac{1}{N}\cdot\sum_{n=-(N-1)}^{N-1}\left|\rho_{XX}\left(n\right)\right|\underset{N\to\infty}{\to}0,\quad \frac{1}{T}\cdot\int_{-T}^{T}\left|\rho_{XX}\left(\tau\right)\right|d\tau\underset{T\to\infty}{\to}0$$

3. ארגודיות בתוחלת מתקיימת אם (תנאי מספיק) מקדם הקורולצייה $\rho_{XX}\left(\tau\right)$ מקיים לפחות את אחד התנאים:

a. $\rho_{XX}\left(\tau\right)=\frac{C_{XX}\left(\tau\right)}{C_{XX}\left(0\right)}\underset{\tau\to\infty}{\to}0$

b. אי-התבדרות האינגרל $\int_{0}^{\infty}\left|\rho_{XX}\left(\tau\right)\right|d\tau$, כלומר $\int_{0}^{\infty}\left|\rho_{XX}\left(\tau\right)\right|d\tau<\infty$.

הערות - תנאי 1 הוא דרישה בדיוק לפי הגדרת העקביות של משערך (מביטוי השונות של המשערך).

תנאי 2 נובע מתנאי 1 לפי החלשה שלו:

$$\frac{1}{T}\cdot\int_{-T}^{T}C_{XX}(t)\cdot\left(1-\frac{|t|}{T}\right)dt\le\frac{1}{T}\cdot\int_{-T}^{T}C_{XX}(t)\cdot 1\cdot dt\le\frac{1}{T}\cdot\int_{-T}^{T}\left|C_{XX}(t)\right|dt$$

תנאי 3א' נובע מהגדרת מקדם הקורולציה ומהגדרת הספק אות (סופי).

תנאי 3ב' נובע לפי אי-התבדרות האינטגרל (נסיק ממנו את תנאי 2 אשר כבר הסברנו):

$$\exists A:\int_{0}^{\infty}\left|\rho_{XX}(\tau)\right|d\tau=A\quad\Rightarrow\quad\frac{1}{N}\cdot\int_{0}^{\infty}\left|\rho_{XX}(\tau)\right|d\tau=\frac{A}{N}\underset{N\to\infty}{\to}0$$

<u>טענה ב' – תנאי סלוצקי (Slutsky's Theorem):</u>

עבור ת"א סטציונרי תתקיים ארגודיות בתוחלת לפי תנאי סלוצקי אם"ם (תנאי הכרחי ומספיק) מתקיים:

$$\frac{1}{N}\cdot\sum_{n=-(N-1)}^{N-1}C_{XX}(n)\underset{N\to\infty}{\to}0,\quad\frac{1}{T}\cdot\int_{-T}^{T}C_{XX}(t)\cdot dt\underset{T\to\infty}{\to}0$$

בטענה א' תנאי 1 אנו מבחינים כי אנו לוקחים את פונקציית האוטו-קו-ואריאנס ומכפילים אותה בחלון משולשי ברוחב $2T$ עם גובה מקסימלי 1 , ז"א מנחיתים אותה. לכן תנאי מספיק הוא אם הביטוי מתכנס לאפס ללא ההנחתה.

<u>דיון בארגודיות בתוחלת של שרשרת מרקוב:</u>

נרצה להראות עבור שרשרת מרקובית סטציונרית כי שכחת העבר שקולה להתכנסות ממוצע זמני לממוצע סטטיסטי.

ז"א נרצה להראות כי אם מתקיים:

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

אז מתקיים:

$$\lim_{N\to\infty}\left(\hat{\underline{\pi}}^{stat}(N)\right)=\underline{\pi}^{stat},\qquad\lim_{N\to\infty}\left(\hat{P}_{ab}(N)\right)=\underline{\underline{P}}(a,b)\quad\forall a,b\in\{1,\ldots,J\}$$

למעשה מספיק להראות את הקיום של התכנסות ערך הסתברות המעבר (כי נוסחתו כוללת את הפילוג השולי של מצב המקור – נובע מהסתברות מותנית ופותח בנושא מציאת סטטסיטיקה של שרשרת מרקוב ארגודית).

מן הביטוי עבור משערך התפלגות המעבר נראה שלמעשה נצטרך להראות כי הפילוג השולי המשוערך מתכנס לפילוג השולי הסטטיסטי:

$$\hat{P}(a,b)(N)=\frac{1}{N-1}\cdot\sum_{i=0}^{N-2}I_{ab}\left(x_i,x_{i+1}\right)\underset{N\to\infty}{\to}\Pr\left(x_i=a,x_{i+1}=b\right)$$

לכן למעשה נצטרך להראות שפונקציית האינדיקטור $I_{ab}\left(x,y\right)$ ארגודית ביחס לתוחלת. נראה את הדבר ע"י שימוש בתנאי 3א', כלומר נראה כי מקדם הקורולציה בין זוג דגימות שואף לאפס. אנו דנים בשרשרת מרקוב ולכן עבורנו פונקציית האינדיקטור תמיד פועלת על זוג זמנים סמוכים ולכן למעשה $I_{ab}\left(k\right)=I_{ab}\left(X_k,X_{k+1}\right)$.

$$E\left[I_{ab}\left(X_i,X_{i+1}\right)\right]=\Pr\left(X_i=a,X_{i+1}=b\right)=\pi_a^{stat}\cdot\underline{\underline{P}}\left(a,b\right)$$
$$R_{I_{ab}I_{ab}}\left(k\right)=R_{I_{ab}I_{ab}}\left(i+k,i\right)=E\left[I_{ab}\left(X_{i+k},X_{i+k+1}\right)\cdot I_{ab}\left(X_i,X_{i+1}\right)\right]$$

מתקיים $I_{ab}\left(X_{i+k},X_{i+k+1}\right)\cdot I_{ab}\left(X_i,X_{i+1}\right)=1$ אם"ם $X_i=X_{i+k}=a,X_{i+1}=X_{i+k+1}=b$.

$$R_{I_{ab}I_{ab}}\left(k\right)=1\cdot\pi_a^{stat}\cdot\underline{\underline{P}}\left(a,b\right)\cdot\Pr\left(X_{i+1}=b,X_{i+k}=a\right)\cdot\underline{\underline{P}}\left(a,b\right)=$$
$$=\pi_a^{stat}\cdot\left(\underline{\underline{P}}\left(a,b\right)\right)^2\cdot\Pr\left(X_{i+1}=b,X_{i+k}=a\right)$$

ולכן פונקציית האוטו-קו-ואריאנס המתאימה היא:

$$C_{I_{ab}I_{ab}}\left(k\right)=R_{I_{ab}I_{ab}}\left(k\right)-E\left[I_{ab}\left(X_{i+k},X_{i+k+1}\right)\right]\cdot E\left[I_{ab}\left(X_{i+k},X_{i+k+1}\right)\right]=$$
$$=\pi_a^{stat}\cdot\left(\underline{\underline{P}}\left(a,b\right)\right)^2\cdot\Pr\left(X_{i+1}=b,X_{i+k}=a\right)-\pi_a^{stat}\cdot\underline{\underline{P}}\left(a,b\right)\cdot\pi_a^{stat}\cdot\underline{\underline{P}}\left(a,b\right)=$$
$$=\pi_a^{stat}\cdot\left(\underline{\underline{P}}\left(a,b\right)\right)^2\cdot\Pr\left(X_{i+1}=b,X_{i+k}=a\right)-\left(\pi_a^{stat}\right)^2\cdot\left(\underline{\underline{P}}\left(a,b\right)\right)^2=$$
$$=\pi_a^{stat}\cdot\left(\underline{\underline{P}}\left(a,b\right)\right)^2\cdot\left[\Pr\left(X_{i+1}=b,X_{i+k}=a\right)-\pi_a^{stat}\right]$$

אך משיכחת העבר נקבל $\lim_{k\to\infty}\left(\Pr\left(X_{i+1}=b,X_{i+k}=a\right)\right)=\pi_a^{stat}$. לכן $C_{I_{ab}I_{ab}}\left(k\right)\underset{k\to\infty}{\to}0$ משמע ארגודיות.

ספקטרום הספק

ספקטרום הספק (PSD = Power Spectral Density) היא פונקציה אי-שלילית אשר מציגה את ההספק ליחידת תדר (ומכאן שמה 'צפיפות הספק-ספקטרלי') עבור אותות בזמן (דטרמיניסטיים או סטוכסטיים/אקראיים). אנו נחשב את ספקטרום ההספק עבור ת"א WSS בלבד.

חישוב ספקטרום הספק

ספקטרום הספק הוא למעשה התמרת פורייה (Fourier Transform) של פונקציית האוטו-קורולציה של התהליך.
כידוע פונקציית האוטו-קורולציה של ת"א היא פונקציה דטרמניסטית ולכן גם ספקטרום ההספק תהיה כזו.

ספקטרום הספק של ת"א WSS בזמן רציף

ספקטרום ההספק של תהליך אקראי WSS בזמן רציף $X(t)$ בעל פונקציית אוטו-קורולציה $R_{XX}(\tau)$ הוא:

$$S_{XX}(\omega)=F\left\{R_{XX}(\tau)\right\}=\int_{-\infty}^{\infty}R_{XX}(\tau)\cdot e^{-j\omega\tau}\cdot d\tau$$

<u>ספקטרום הספק של ת"א WSS בזמן בדיד</u>

ספקטרום ההספק של תהליך אקראי WSS בזמן בדיד $X[n]$ בעל פונקציית אוטו-קורולציה $R_{XX}(k)$ הוא:

$$S_{XX}\left(e^{j\omega}\right)=F\left\{R_{XX}(k)\right\}=\sum_{k=-\infty}^{\infty}R_{XX}(k)\cdot e^{-j\omega k}$$

<u>קרוס-ספקטרום הספק של זוג ת"א jWSS</u>

ספקטרום ההספק של זוג ת"א jWSS בזמן רציף $X(t),Y(t)$ בעלי פונקציית קורולציה משותפת $R_{XY}(\tau)$ הוא:

$$S_{XY}(\omega)=F\left\{R_{XY}(\tau)\right\}=\int_{-\infty}^{\infty}R_{XY}(\tau)\cdot e^{-j\omega\tau}\cdot d\tau$$

נראה כי זו הגדרה כללית וניתן להספיק ממנה את הגדרת ה"אוטו-ספקטרום".
בהתאם ניתן להגדיר גם עבור זוג ת"א בזמן בדיד.

<u>תכונות של ספקטרום הספק עבור ת"א ממשי</u>

בקורס הזה אנחנו מתעסקים עם תהליכים ממשיים בלבד (אין ערכים מרוכבים) ולכן הדיון כשאנו עוברים למרחב התדר נהיה פשוט.

<u>תכונות של פונקציית האוטו-קורולציה – תזכורת:</u>

- $R_{XX}(\tau)$ היא פונקציה זוגית, כלומר מתקיים $\forall\tau \quad R_{XX}(\tau)=R_{XX}(-\tau)$.
- עבור תהליך WSS מתקיים $\forall\tau \quad R_{XX}(0)\geq\left|R_{XX}(\tau)\right|$.

<u>תכונות של ספקטרום הספק:</u>

1. $S_{XX}(\omega)$ היא פונקציה ממשית. נראה:

$$S_{XX}(\omega)=\int_{-\infty}^{\infty}R_{XX}(\tau)\cdot e^{-j\omega\tau}\cdot d\tau=\int_{-\infty}^{\infty}R_{XX}(\tau)\cdot(\cos\omega\tau+j\cdot\sin\omega\tau)\cdot d\tau=$$

$$=\int_{-\infty}^{\infty}R_{XX}(\tau)\cdot\cos\omega\tau\cdot d\tau+j\cdot\int_{-\infty}^{\infty}R_{XX}(\tau)\cdot\sin\omega\tau\cdot d\tau\overset{*}{=}\int_{-\infty}^{\infty}R_{XX}(\tau)\cdot\cos\omega t\cdot d\tau$$

שיוויון $(*)$ נובע מאי-זוגיות האינטגרנד (מכפלת פונקציות $\sin\omega\tau$ ו- $R_{XX}(\tau)$) בקטע סימטרי.

2. $S_{XX}(\omega)$ היא פונקציה זוגית (נובע מזוגיות האינטגרנד כמכפלת פונקציות זוגיות $\cos\omega\tau$ ו- $R_{XX}(\tau)$).

3. $S_{XX}(\omega)$ היא פונקציה אי-שלילית, כלומר מתקיים $\forall\omega \quad S_{XX}(\omega)$. מכיוון שהתמרת פורייה של פונקציה אי-שלילית מוגדרת היא פונקציה אי-שלילית.

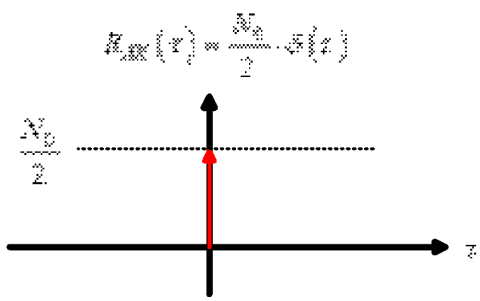
<u>רעש לבן בזמן רציף</u>

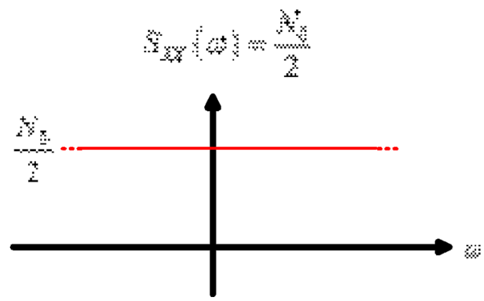
רעש לבן הוא תהליך אקראי בעל צפיפות הספק-ספקטרלי אחידה וקבועה בכל התדרים. כמובן שרעש לבן בעל רוחב סרט אינסופי הוא תיאורטי בלבד אחרת נקבל ת"א עם אנרגיה אינסופית.
ברעש לבן כל זוג דגימות שונות הן בת"ס, על אף שתהליך רעש לבן הוא ת"א בזמן רציף ולא משנה כמה קרובות תהיינה הדגימות (שקול לסדרה i.i.d). למעשה המשמעות של זה היא שת"א של רעש לבן לא רציף בשום נקודה.
רעש לבן הוא מידול של רעש תרמי ולכן נלמד לעבוד איתו. בפועל דגימותיו של רעש תרמי כן תלויות אחת בשניה (והוא גם תהליך רציף) אך בקצב כה מהיר שהמערכות שלנו לא מספיקות לקלוט דגימות תלויות. לא משנה כמה מהר נדגום רעש תרמי זה יהיה מספיק רחוק כך שהדגימות תהיינה i.i.d ולכן נוכל למדל את הרעש התרמי לרעש לבן.

נאמר שתהליך אקראי בזמן רציף $X(t)$ הינו רעש לבן (בזמן רציף) אם הוא WSS ופונקציית האוטו-קורולציה שלו היא מהצורה $R_{XX}(\tau)=A\cdot\delta(\tau)$ (ז"א הלם בגובה A ב- $\tau=0$).
המשמעות היא שכל זוג דגימות בזמנים שונים הינן חס"ק. נחשב את ספיפות ההספק של אות כזה ונקבל ערך קבוע:

$$S_{XX}(\omega)=F\{A\cdot\delta(\tau)\}=A\cdot F\{\delta(\tau)\}=A$$

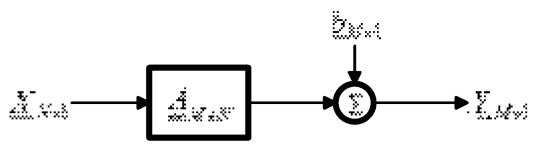
נהוג לסמן $A=\frac{N_0}{2}$ בתור צפיפות הספק חצי ספקטרלי (מאחר שעבור כל תדר יש גם את התדר ה"שלילי").





מעבר תהליכים WSS במערכות ליניאריות

למדנו איך לנתח מעבר של ווקטור אקראי דרך מערכת ליניארית. כמו כן למדנו לחשב סטטיסטיקה של המוצא וסטטיסטיקה של איברי הקרוס בהינתן סטטיסטיקה של הכניסה ואיפיון המערכת (עד סדר שני מאחר שזאת מערכת ליניארית).



$$\underline{Y}_{M\times1}=\underline{\underline{A}}_{M\times N}\cdot\underline{X}_{N\times1}+\underline{b}_{M\times1}$$

נרצה להרחיב את הדיון עבור מעבר של תהליכים אקראיים במערכת ליניארית.

מעבר תהליך אקראי דרך מערכת ליניארית

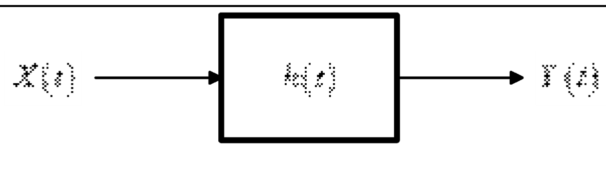
מערכת עבור תהליכים לא יכולה להיות מאופיינת ע"י מטריצה או לעבור עבור מימדים סופיים מאחר והקלט הוא תהליך אקראי. תהליך אקראי הוא סדרה אינסופית. בפועל, סדרה תיכנס למערכת וסדרה תצא. למעשה מערכת כזאת מאופיינת ע"י פונקציית תמסורת ונהוג לקרוא לפונקציה זו מסנן (על אף שהיא לא חייבת לבצע פעולת סינון).

אנו נתעסק רק בתהליכים סטציונריים במובן הרחב (WSS) שעוברים במערכות ליניאריות אשר אינן משתנות בזמן (Linear Time-Invariant), ז"א השפעת הזמן על המערכת הוא רק דרך השפעת הזמן על תהליך הכניסה.

מבנה כללי של מערכת כזאת יראה באופן הבא (או עם $X[n],Y[n],h[n]$ לת"א בזמן בדיד):

עבור מערכות LTI המוצא נתון ע"י קונבולוציה בין הכניסה לפונקציית המסנן של המערכת.

$$Y(t)=X(t)*h(t)$$

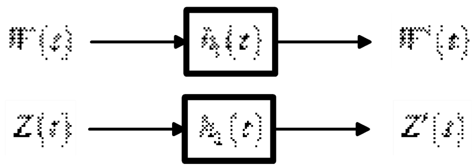


לא הוספנו קבועים (בכניסה או במוצא) וזאת לשם פשטות הדיון. בכל מקרה, גם אם נוסיף אותם, הקבועים יכולים לשנות רק את התוחלת מאחר שהם בהכרח קבועים בזמן כי אנו דנים במערכות LTI.

נרצה לדעת את הסטטיסטיקה מסדר שני של המוצא בהינתן הסטטיסטיקה מסדר שני של הכניסה. לצורך כך נחשב סטטיטיקה עבור מודל כללי יותר של מעבר ת"א במערכת LTI ואז נסיק על המערכת ה'פשוטה' שלנו.

מודל עזר לניתוח הקשר בין הכניסה למוצא במערכת LTI

נתונות שתי מערכות $h_1(t),h_2(t)$, הן מוזנות בת"א $W(t),Z(t)$ בהתאמה. תהליכים אלו הם ת"א jWSS. הסטטיסטיקה מסדר שני $R_{WW}(\tau),R_{ZZ}(\tau),R_{WZ}(\tau)$ ידועה עבורם.



נרצה לדעת את המומנט המשותף מסדר שני של תהליכי המוצא, כלומר את $R_{W'Z'}(\tau)$ (אם הם בכלל jWSS).

$$R_{W'Z'}(t+\tau,t)=E\left[W'(t+\tau)\cdot Z'(t)\right]=E\left[\left(W(t+\tau)*h_1(t)\right)\cdot\left(Z(t)*h_2(t)\right)\right]=$$

$$=E\left[\left(\int_{-\infty}^{\infty}h_1(\alpha)\cdot W(t+\tau-\alpha)\cdot d\alpha\right)\cdot\left(\int_{-\infty}^{\infty}h_2(\beta)\cdot Z(t-\beta)\cdot d\beta\right)\right]=$$

$$\overset{*}{=}\int_{\beta=-\infty}^{\infty}\int_{\alpha=-\infty}^{\infty}h_1(\alpha)\cdot h_2(\beta)\cdot E\left[W(t+\tau-\alpha)\cdot Z(t-\beta)\right]\cdot d\alpha\cdot d\beta=$$

$$=\int_{\beta=-\infty}^{\infty}\int_{\alpha=-\infty}^{\infty}h_1(\alpha)\cdot h_2(\beta)\cdot R_{WZ}(t+\tau-\alpha,t-\beta)\cdot d\alpha\cdot d\beta=$$

$$\overset{**}{=}\int_{\beta=-\infty}^{\infty}\int_{\alpha=-\infty}^{\infty}h_1(\alpha)\cdot h_2(\beta)\cdot R_{WZ}(\tau-\alpha+\beta)\cdot d\alpha\cdot d\beta=$$

$$=\int_{\beta=-\infty}^{\infty}h_2(\beta)\cdot\left(\int_{\alpha=-\infty}^{\infty}h_1(\alpha)R_{WZ}(\tau+\beta-\alpha)\cdot d\alpha\right)\cdot d\beta=$$

$$=\int_{\beta=-\infty}^{\infty}h_2(\beta)\cdot\left[h_1(\tau+\beta)*R_{WZ}(\tau+\beta)\right]\cdot d\beta\underset{\substack{\gamma=-\beta\\ d\gamma=-d\beta}}{=}\int_{\gamma=-\infty}^{\infty}h_2(-\gamma)\cdot\left[h_1(\tau-\gamma)*R_{WZ}(\tau-\gamma)\right]\cdot d\gamma=$$

$$=\int_{\gamma=-\infty}^{\infty}h_2(-\gamma)\cdot\left[h_1*R_{WZ}\right](\tau-\gamma)\cdot d\gamma=h_1(\tau)*R_{WZ}(\tau)*h_2(-\tau)=R_{W'Z'}(\tau)$$

שיוויון $(*)$ מליניאריות של תוחלת ואינטגרל. שיוויון $(**)$ משום ש- $W(t)$ ו- $Z(t)$ הם ת"א jWSS.

קיבלנו כי המוצאים $W'(t),Z'(t)$ הם אכן ת"א jWSS, ז"א עבור שני ת"א jWSS אשר עוברים כ"א (אפילו בנפרד) במערכות LTI מתקיים כי המוצאים הם jWSS.

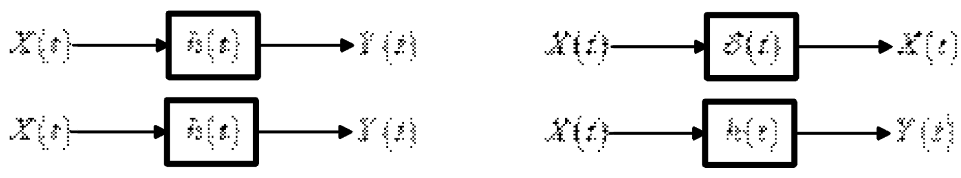
<u>סטטיסטיקה של מוצא מערכת LTI המוזנת בתהליך אקראי WSS</u>

עבור $X(t)$ תהליך אקראי WSS עם סטטיסטיקה מסדר שני $R_{XX}(\tau)$ ידועה ומערכת LTI עם פונקציית תמסורת ידועה $h(t)$ נסמן את מוצא המערכת ב- $Y(t)$ ואנו יודעים כי מתקיים $Y(t)=X(t)*h(t)$.

אנו רוצים לדעת את הסטטיסטיקה מסדר שני של המוצא ואת המשותפת, כלומר את $R_{XY}(\tau),R_{YY}(\tau)$.

לצורך כך נשתמש במודל העזר שפיתחנו וניזכר כי קיבלנו $R_{W'Z'}(\tau)=h_1(\tau)*R_{WZ}(\tau)*h_2(-\tau)$.

נזהה כי נוכל לבחור למודל הנ"ל קונפיגורציות שונות אשר יעזרו לנו לקבל את הסטטיסטיקה הרצויה.



(L) (R)

נראה את הפיתוחים עבור זוג הקונפיגורציות (ימין ושמאל):

$$(L)\quad \begin{array}{lll} W(t)=X(t),h_1(t)=h(t) & \Rightarrow & W'(t)=W(t)*h_1(t)=X(t)*h(t)=Y(t) \\ Z(t)=X(t),h_2(t)=h(t) & \Rightarrow & Z'(t)=Z(t)*h_2(t)=X(t)*h(t)=Y(t) \\ R_{W'Z'}(\tau)=h_1(\tau)*R_{WZ}(\tau)*h_2(-\tau) & \Rightarrow & R_{YY}(\tau)=h(\tau)*R_{XX}(\tau)*h(-\tau) \end{array}$$

$$(R)\quad \begin{array}{lll} W(t)=X(t),h_1(t)=\delta(t) & \Rightarrow & W'(t)=W(t)*h_1(t)=X(t)*\delta(t)=X(t) \\ Z(t)=X(t),h_2(t)=h(t) & \Rightarrow & Z'(t)=Z(t)*h_2(t)=X(t)*h(t)=Y(t) \\ R_{W'Z'}(\tau)=h_1(\tau)*R_{WZ}(\tau)*h_2(-\tau) & \Rightarrow & R_{XY}(\tau)=\delta(\tau)*R_{XX}(\tau)*h(-\tau) \end{array}$$

סה"כ קיבלנו (נוכל גם להציג תוצאה עבור $R_{YX}(\tau)$ אם נהפוך את קונפוגרצית R):

$$R_{XY}(\tau)=R_{XX}(\tau)*h(-\tau)$$
$$R_{YX}(\tau)=R_{XX}(\tau)*h(\tau)$$
$$R_{YY}(\tau)=h(\tau)*R_{XX}(\tau)*h(-\tau)$$

נחשב את ספקטרום ההספק עבור אותות אלו (ניזכר כי קונבולוציה בזמן $\leftrightarrow$ מכפלה בתדר) לפי ההגדרה ע"י טרנספורם פורייה של פונקציית הקורולציה ונקבל:

$$S_{XY}(\omega)=S_{XX}(\omega)\cdot H^*(\omega)$$
$$S_{YX}(\omega)=S_{XX}(\omega)\cdot H(\omega)$$
$$S_{YY}(\omega)=H(\omega)\cdot S_{XX}(\omega)\cdot H^*(\omega)=S_{XX}(\omega)\cdot\left|H(\omega)\right|^2$$

באותו אופן עבור ת"א WSS בזמן בדיד נקבל:

$$\begin{aligned}R_{XY}[n]&=R_{XX}[n]*h[-n]\\R_{YX}[n]&=R_{XX}[n]*h[n]\\R_{YY}[n]&=h[n]*R_{XX}[n]*h[-n]\end{aligned}\quad\underset{FourierTransform}{\overset{Discrete-Time}{\longleftrightarrow}}\quad\begin{aligned}S_{XY}\left(e^{j\omega}\right)&=S_{XX}\left(e^{j\omega}\right)\cdot H^*\left(e^{j\omega}\right)\\S_{YX}\left(e^{j\omega}\right)&=S_{XX}\left(e^{j\omega}\right)\cdot H\left(e^{j\omega}\right)\\S_{YY}\left(e^{j\omega}\right)&=S_{XX}\left(e^{j\omega}\right)\cdot\left|H\left(e^{j\omega}\right)\right|^2\end{aligned}$$

נרצה להחיל את מסקנות הפיתוח גם על פונקציית הקו-ואריאנס, כלומר להראות שהקשרים הנ"ל מתקיימים גם עבור $C_{XX}(\tau),C_{XY}(\tau),C_{YY}(\tau)$.

עבור התהליכים $X(t),Y(t)$ הנ"ל (כניסתה ומוצאה של המערכת $h(t)$) נגדיר את שני התהליכים הבאים:

$$X'(t)=X(t)-\eta_X,\qquad Y'(t)=Y(t)-\eta_Y$$

נזין את התהליך האקראי $X'(t)$ למערכת $h(t)$ ונקבל במוצאה:

$$X'(t)*h(t)=\left(X(t)-\eta_X\right)*h(t)=X(t)*h(t)-\eta_X*h(t)=Y(t)-\eta_Y=Y'(t)$$

קיבלנו את התהליך $Y'(t)$ במוצא.

מכאן שהנוסחאות לפונקציות הקורולציה תקפות עבור התהליכים $X'(t),Y'(t)$.

מכיוון ששני התהליכים שהגדרנו $X'(t),Y'(t)$ הם מופחתי תוחלת אז מתקיים:

$$R_{X'Y'}(\tau)=C_{X'Y'}(\tau),\qquad R_{Y'Y'}(\tau)=C_{Y'Y'}(\tau),\qquad R_{X'X'}(\tau)=C_{XX}(\tau)$$

עבור פוריה טרנספורם של פונקציה הקו-ואריאנס נוסיף סימון על מנת שנוכל להבדיל ולסיכום:

$$\begin{aligned}C_{XY}(\tau)&=C_{XX}(\tau)*h(-\tau)\\C_{YX}(\tau)&=C_{XX}(\tau)*h(\tau)\\C_{YY}(\tau)&=h(\tau)*C_{XX}(\tau)*h(-\tau)\end{aligned}\quad\underset{Transform}{\overset{Fourier}{\longleftrightarrow}}\quad\begin{aligned}S^C_{XY}(\omega)&=S^C_{XX}(\omega)\cdot H^*(\omega)\\S^C_{YX}(\omega)&=S^C_{XX}(\omega)\cdot H(\omega)\\S^C_{YY}(\omega)&=S^C_{XX}(\omega)\cdot\left|H(\omega)\right|^2\end{aligned}$$

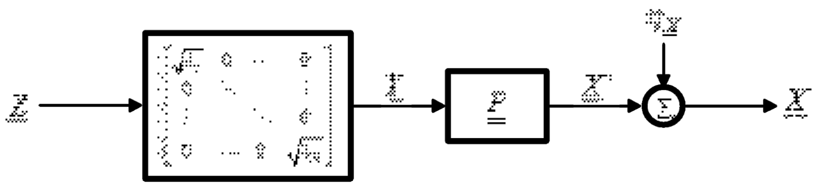
ועבור תהליכים אקראיים WSS בזמן בדיד:

$$C_{XY}[n]=C_{XX}[n]*h[-n] \qquad S_{XY}^{C}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)\cdot H^{*}\left(e^{j\omega}\right)$$
$$C_{YX}[n]=C_{XX}[n]*h[n] \quad \overset{Discrete-Time}{\underset{FourierTransform}{\leftrightarrow}} \quad S_{YX}^{C}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)\cdot H\left(e^{j\omega}\right)$$
$$C_{YY}[n]=h[n]*C_{XX}[n]*h[-n] \qquad S_{YY}^{C}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)\cdot\left|H\left(e^{j\omega}\right)\right|^{2}$$

<u>יצור תהליך אקראי גאוסי סטציונרי</u>

למדנו כי נוכל לייצר וא"ג כרצוננו ע"י העברת וא"ג i.i.d במערכת ליניארית. קראנו לפעולה זו "צביעת" וא"ג (וגם הכרנו את הפעולה ההפוכה של "הלבנת" וא"ג). התחלנו עם וא"ג עם רכיבים i.i.d נסמנו $\underline{Z}\sim N\left(\underline{0},\underline{\underline{I}}\right)$ וקיבלנו במוצא וא"ג באותו מימד עם פילוג $\underline{X}\sim N\left(\eta_{\underline{X}},C_{\underline{X}\underline{X}}\right)$.

מבנה המערכת (כאשר $\underline{\underline{P}}$ הינה מטריצה מלכסנת של $C_{\underline{X}\underline{X}}$):



עבור תהליכים אקראיים נוכל לייצר תא"ג סטציונרי. נבחין ממודל העזר כי פונקציית האוטו-קורולציה של מוצא מערכת LTI בעלת מסנן $h(t)$ וכניסה $X(t)$ WSS היא:

$$R_{YY}(\tau)=h(\tau)*R_{XX}(\tau)*h(-\tau)$$

מכאן (וכפי שראינו במודל) שהמוצא כבר סטציונרי. נרצה לשלוט ו"לעצב" את המוצא.
אם ניקח תהליך מקור עם פונקציית קורולציה קבועה בזמן (רעש לבן) אז נקבל:

$$R_{YY}(\tau)=h(\tau)*\left(\frac{N_0}{2}\cdot\delta(\tau)\right)*h(-\tau)=\frac{N_0}{2}\cdot h(\tau)*h(-\tau)$$

ולכן נצטרך לתכנן רק את המסנן ונקבל אות עם פונקציית קורולציה כרצוננו.
אם נרצה שהמוצא יהיה תא"ג אז נצטרך שהכניסה תהיה תא"ג (מאחר שהוא עובר במערכת LTI). לכן אם הכניסה תהיה תהליך רעש לבן בעל התפלגות נורמלית בזמן ותוחלת אפס (רעש גאוסי לבן) אז נקבל במוצא תא"ג סטציונרי.

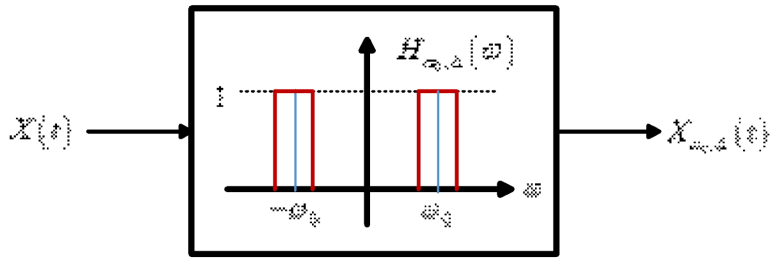
כנ"ל לגביי תא"ג סטציונרי בזמן בדיד עבורו $R_{YY}[n]=\frac{N_0}{2}\cdot h[n]*h[-n]$.

<u>מסנן מעביר סרט</u>

נבחן מעבר של ת"א ממשי WSS דרך מסנן צר סרט (Narrow Band Pass Filter).

נזין ת"א בזמן רציף $X(t)$ אשר ממשי ו-WSS דרך מסנן BFP אשר יסומן $H_{\omega_0,\Delta}(\omega)$.

נציג את המערכת:



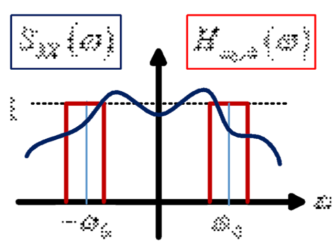
נבחין כי סימנו את המוצא ע"י $X_{\omega_0,\Delta}(t)$ מאחר שהמוצא הוא חיתוך של $X(t)$ בתחום התדר לפי המסנן, ז"א סביב ω_0. חשוב לציין כי תדר המרכז מסומן בתדר זויתי ω ואילו רוחב המסנן בתדר f ולכן (לפי המעבר $\omega = 2\pi f$):

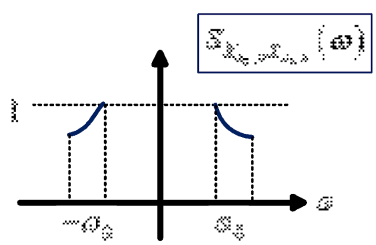
$$H_{\omega_0,\Delta}(f) = \begin{cases} 1 & f \in \left(f_0 - \frac{\Delta}{2}, f_0 + \frac{\Delta}{2}\right) \cup \left(-f_0 - \frac{\Delta}{2}, -f_0 + \frac{\Delta}{2}\right) \\ 0 & else \end{cases}, \quad f_0 = \frac{\omega_0}{2\pi}$$

נניח כי ידוע הספקטרום של התהליך $X(t)$ אז נוכל לחשב את ספקטרום ההספק במוצא ע"י:

$$S_{X_{\omega_0,\Delta}X_{\omega_0,\Delta}}(\omega) = S_{XX}(\omega) \cdot \left|H_{\omega_0,\Delta}(\omega)\right|^2$$

האותות בתדר יראו כך:





כמובן שאם המסנן לא היה בגובה אחד אז היינו מקבלים במוצא צפיפויות הספק בגבהים שונים אך עם אותה התנהגות אם המסנן היה קבוע (בכל מקרה אילו היינו רוצים להחזיר את הערכים היינו מעבירים במסנן אחר ש"יתקן").
נרצה לחשב את הספק אות התהליך במוצא:

$$Power\left(X_{\omega_0,\Delta}(t)\right) = E\left[\left(X_{\omega_0,\Delta}(t)\right)^2\right] = R_{X_{\omega_0,\Delta}X_{\omega_0,\Delta}}(0) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} S_{X_{\omega_0,\Delta}X_{\omega_0,\Delta}}(\omega) e^{j\omega 0} d\omega =$$

$$= \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} S_{X_{\omega_0,\Delta}X_{\omega_0,\Delta}}(\omega) d\omega = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} S_{XX}(\omega) \cdot \left|H_{\omega_0,\Delta}(\omega)\right|^2 d\omega = 2 \cdot \frac{1}{2\pi} \cdot \int_{2\pi\left(f_0 - \frac{\Delta}{2}\right)}^{2\pi\left(f_0 + \frac{\Delta}{2}\right)} S_{XX}(\omega) d\omega$$

ז"א אנחנו סוכמים את צפיפות ההספק הספקטרלי על התדרים בהם קיים המסנן.
אם אנחנו יודעים כי מתקיים אז נוכל להפעיל אינטגרל רימן (רוחב האינטגרל כפול ערך האינטגרנד במרכז הקטע):

$$Power\left(X_{\omega_0,\Delta}(t)\right)=\frac{1}{\pi}\cdot\int_{2\pi\left(f_0-\frac{\Delta}{2}\right)}^{2\pi\left(f_0+\frac{\Delta}{2}\right)} S_{XX}(\omega)d\omega \underset{\Delta\ll}{\overset{\text{Riemann}}{\rightarrow}} \frac{1}{\pi}\cdot 2\pi\Delta\cdot S_{XX}(2\pi f_0)=2\Delta\cdot S_{XX}(\omega_0)$$

נבחין כי אם נעביר אות רעש לבן בזמן רציף עם צפיפות הספק חצי ספקטרלית נקבל במוצא אות בעל הספק:

$$N(t)=\frac{N_0}{2}\delta(t) \quad\Rightarrow\quad Power\left(X_{\omega_0,\Delta}(t)\right)=2\Delta\cdot S_{NN}(\omega_0)=2\Delta\cdot\frac{N_0}{2}=\Delta\cdot N_0$$

הערה - פונקציית האוטו-קורולציה מקיימת תמיד את התנאי שטרנספורם פורייה שלה יהיה אי-שלילי. זה נותן תחושה נוספת לגביי כך שהיא פונקציה אי-שלילית מוגדרת. גם ההפך מתקיים – עבור כל פונקציה אי-שלילית זוגית במרחב התדר טרנספורם פורייה הפוך יניב פונקציית קורולציה.

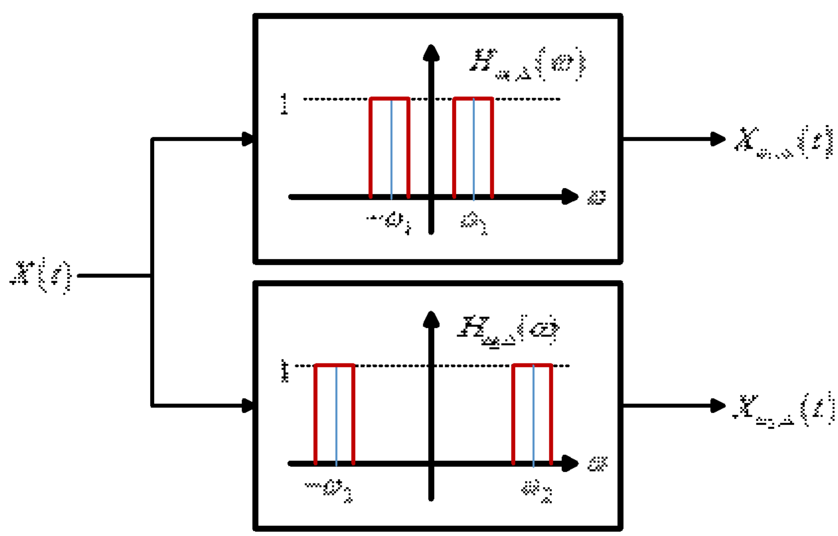
<u>חוסר קורולציה בין פסי-תדר</u>

נדון בתהליכים אקראיים בזמן רציף אך הפיתוח תקף גם לתהליכים אקראיים בזמן בדיד.

<u>חוסר קורולציה בין פסי תדר-זרים של תהליך אקראי WSS</u>

נרצה להראות כי עבור ת"א בזמן רציף $X(t)$ שהוא סטציונרי WSS מתקיים ששני אותות שנוצרו מחיתוך שלו בפסי-תדר זרים הם תמיד חסרי קורולציה.

נציג את המערכת הבאה:



<u>טענה:</u>

עבור חיתוך בפסי-תדר זרים, כלומר עבור $\Delta<|f_1-f_2|$, מתקיים כי אותות המוצא $X_{\omega_1,\Delta}(t),X_{\omega_2,\Delta}(t)$ (שהם jWSS כמובן כפי שראינו בפיתוח במודל העזר) הם חסרי קורולציה.

זאת אומרת עבור חיתוך בפסי-תדר זרים מתקיים:

$$R_{X_{\omega_1,\Delta}X_{\omega_2,\Delta}}(\tau)=E\left[X_{\omega_1,\Delta}(t+\tau)\cdot X_{\omega_2,\Delta}(t)\right]=0 \qquad \forall\tau$$

<u>הוכחה:</u>

נזהה כי נוכל להשתמש במודל העזר ולבנות קונפיגורציה עבור מערכת זו:

$$W(t)=X(t),h_1(t)=h_{\omega_1,\Delta}(t) \quad\Rightarrow\quad W'(t)=W(t)*h_1(t)=X(t)*h_{\omega_1,\Delta}(t)=X_{\omega_1,\Delta}(t)$$

$$Z(t)=X(t),h_2(t)=h_{\omega_2,\Delta}(t) \quad\Rightarrow\quad Z'(t)=Z(t)*h_2(t)=X(t)*h_{\omega_2,\Delta}(t)=X_{\omega_2,\Delta}(t)$$

$$R_{W'Z'}(\tau)=h_1(\tau)*R_{WZ}(\tau)*h_2(-\tau) \quad\Rightarrow\quad R_{X_{\omega_1,\Delta}X_{\omega_2,\Delta}}(\tau)=h_{\omega_1,\Delta}(\tau)*R_{XX}(\tau)*h_{\omega_2,\Delta}(-\tau)$$

$$\Rightarrow\quad S_{X_{\omega_1,\Delta}X_{\omega_2,\Delta}}(\omega)=S_{XX}(\omega)\cdot H_{\omega_1,\Delta}(\omega)\cdot H^*_{\omega_2,\Delta}(\omega)\overset{*}{=}0$$

כאשר שיוויון $(*)$ נובע מכך שבתחום התדר אין למסננים חפיפה.

לכן אם הספקטרום ההספק מתאפס לכל תדר אז התמרת פורייה ההפוכה שלו גם ולכן הוכחנו:

$$R_{X_{\omega_1,\Delta}X_{\omega_2,\Delta}}(\tau)=F^{-1}\left\{S_{X_{\omega_1,\Delta}X_{\omega_2,\Delta}}(\omega)\right\}=F^{-1}\{0\}\equiv 0$$

<u>הערות:</u>

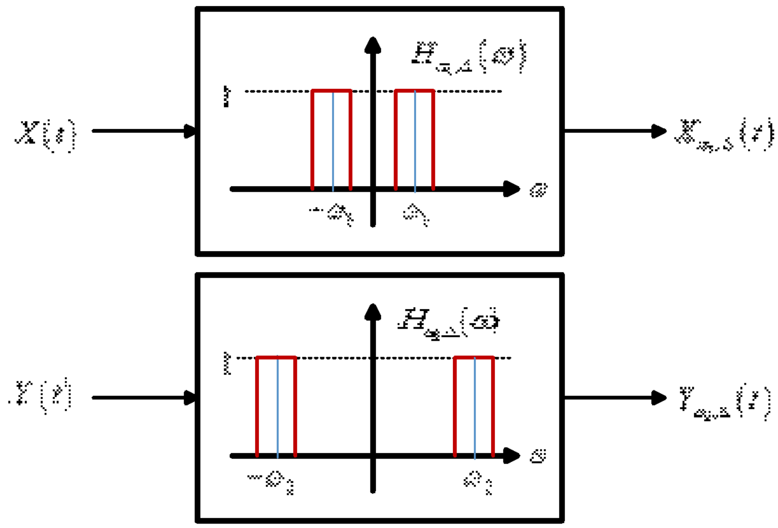
- אילו היה מדובר באות גאוסי אז חס"ק גורר בת"ס.

  לכן למעשה כל תדר בפני עצמו ניתן לאפנון כ- $\sin$ כפול מ"א גאוסי.
- אילו היה מדובר באות כללי אז נוכל לקטום לו קצוות לא רלוונטיים מתחום התדר (ע"י LPF ו-HPF) ואז לא להפריע לאותות אחרים או לתכנן בהתאם ולהשתמש בתדרים לפיצול תקשורת.

<u>חוסר קורולציה בין פסי-תדר זרים של שני תהליכים אקראיים jWSS</u>

נרצה להראות כי עבור שני ת"א בזמן רציף $X(t),Y(t)$ אשר סטציונריים במשותף jWSS מתקיים ששני האותות אשר נוצרו מחיתוך של כל אחד מהם בפס-תדר זר הם תמיד חסרי קורולציה.

נציג את המערכת הבאה:



<u>טענה:</u>

עבור חיתוך בפסי-תדר זרים, כלומר עבור $\Delta < |f_1 - f_2|$, מתקיים כי אותות המוצא $X_{\omega_1,\Delta}(t), Y_{\omega_2,\Delta}(t)$ (שהם jWSS כמובן כפי שראינו בפיתוח במודל העזר) הם חסרי קורולציה.
זאת אומרת עבור חיתוך בפסי-תדר זרים מתקיים:

$$R_{X_{\omega_1,\Delta}Y_{\omega_2,\Delta}}(\tau) = E\left[X_{\omega_1,\Delta}(t+\tau)\cdot Y_{\omega_2,\Delta}(t)\right] = 0 \qquad \forall\tau$$

<u>הוכחה:</u>
נזהה כי המערכת מהצורה של מודל העזר (נוכל להשתמש בו מאחר שהכניסות הן jWSS) ולכן:

$$W(t) = X(t), h_1(t) = h_{\omega_1,\Delta}(t) \quad \Rightarrow \quad W'(t) = W(t) * h_1(t) = X(t) * h_{\omega_1,\Delta}(t) = X_{\omega_1,\Delta}(t)$$
$$Z(t) = Y(t), h_2(t) = h_{\omega_2,\Delta}(t) \quad \Rightarrow \quad Z'(t) = Z(t) * h_2(t) = Y(t) * h_{\omega_2,\Delta}(t) = Y_{\omega_2,\Delta}(t)$$
$$R_{W'Z'}(\tau) = h_1(\tau) * R_{WZ}(\tau) * h_2(-\tau) \quad \Rightarrow \quad R_{X_{\omega_1,\Delta}Y_{\omega_2,\Delta}}(\tau) = h_{\omega_1,\Delta}(\tau) * R_{XY}(\tau) * h_{\omega_2,\Delta}(-\tau)$$
$$\Rightarrow \quad S_{X_{\omega_1,\Delta}Y_{\omega_2,\Delta}}(\omega) = S_{XY}(\omega)\cdot H_{\omega_1,\Delta}(\omega)\cdot H^*_{\omega_2,\Delta}(\omega) \overset{*}{=} 0$$

כאשר שיוויון $(*)$ נובע מכך שבתחום התדר אין למסננים חפיפה.
לכן אם הספקטרום ההספק מתאפס לכל תדר אז התמרת פורייה ההפוכה שלו גם ולכן הוכחנו:

$$R_{X_{\omega_1,\Delta}Y_{\omega_1,\Delta}}(\tau) = F^{-1}\left\{S_{X_{\omega_1,\Delta}Y_{\omega_2,\Delta}}(\omega)\right\} = F^{-1}\{0\} \equiv 0$$

<u>הערה</u> - על מנת שתהיה קורולציה בין האותות נצטרך שתהיה חפיפה בין פסי-התדר (תנאי הכרחי אך לא מספיק כי ייתכן שמאפייניו של אחד האותות מאפסים את הקורולציה ללא קשר לחפיפה בתדר אך לצורך הדיון נתעלם ממקרה זה). במילים

אחרות, אם נעבוד עם חפיפה בפסי-התדר אז התהליכים לא יהיו חס"ק ולכן נוכל לבצע שיערוך ליניארי (כי יש לנו סטטיסטיקה מסדר שני).
אין הדבר נכון בתחום הזמן מאחר שלכל דגימת זמן יש קורולציה כלשהי לדגימת זמן אחרת (לפי ההפרש).

<u>מסקנה</u> - שערוך ליניארי בין שני ת"א תקף אם דוגמים אחד בתדר מסויים ומשערכים את השני באותו התדר. ז"א נאסוף מידע על ת"א $X(t)$ בתדר ω_0 וממנו נסיק מידע על התהליך האקראי $Y(t)$ באותו תדר ω_0.

<u>עיבוד מקבילי של פסי תדר</u>

קיבלנו כי אין קורולציה בין פסי תדר שונים. נדגיש כי הדבר מתקיים בין כל זוג פסי-תדר שונים.
לדוגמא אם ניקח אות ונעביר אותו דרך M מסננות ללא חפיפת תדר אז כל זוג מ- M האותות קטומי-התדר יהיו חס"ק.

<u>שערוך ליניארי אופטימלי של תהליך אקראי במובן MSE (מסנן Wiener)</u>

בהינתן זוג תהליכים אקראיים $X[n],Y[n]$ נרצה לשערך ת"א $X[n]$ מתוך דגימותיו של $Y[n]$ (במובן MSE).

כאשר שערכנו ו"א $\underline{X}$ מתוך ו"א $\underline{Y}$ למעשה שערכנו את רכיביו של $\underline{X}$ אחד-אחד, כאשר כל אחד מהרכיבים $X_i = \underline{X}(i)$ שוערכך בעזרת כל הווקטור $\underline{Y}$.

כעת נרצה לשערך ת"א מתוך ת"א. גם כעת נשערך כל איבר X_i מהסדרה $\{X_n\}$ מתוך כל הידע שיש לנו על איברי סדרת $\{Y_n\}$.

לצורך שערוך כל איבר, נסמנו X_i, נסתכל על אוסף דגימות מחלון זמן כלשהו:

$$\{Y_n\}_{n=i-k}^{i+k} = \{Y_{i-k}, Y_{i-k+1}, \ldots, Y_{i-1}, Y_i, Y_{i+1}, \ldots, Y_{i+k-1}, Y_{i+k}\}$$

בהינתן דגימות אלו נוכל לשערך את X_i באופן אופטימלי במובן MSE כפי ששערכנו מ"א מתוך ו"א (כי כמות האיברים סופית) בעזרת משערך התוחלת המותנית.

המשערך האופטימלי במובן MSE של X_i בהינתן חלון הזמן הוא:

$$\hat{X}_i^{opt,MSE} = \hat{X}_i^{MMSE} = E\left[X_i \middle| Y_{i-k}, \ldots, Y_{i+k}\right]$$

כך נוכל לעשות עבור כל חלון זמן שנרצה (גם לא סימטרי סביב $n=i$) אך ככל שהחלון יגדל המטריצות תגדלנה והחישובים יארכו זמן רב יותר אך נקבל שערוך מדויק יותר.

אם נדע כי התהליכים האקראיים $X[n],Y[n]$ הם jWSS אז נוכל לבצע שערוך פשוט יותר באופן ליניארי (שערוך זה לא בהכרח יהיה MMSE, אנחנו נלמד משערך LMMSE בלבד).

<u>שערוך ליניארי MMSE של ת"א מת"א jWSS</u>

עבור זוג ת"א $X[n],Y[n]$ אשר jWSS משערך LMMSE הינו אופטימלי ונתון ע"י מסנן ה-LTI הבא (נקרה מסנן וינר – Wiener Filter):

$$\hat{X}[n]=\hat{X}^{LMMSE}[n]=Y[n]*h_{LMMSE}[n]+b_{LMMSE}$$

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XY}^{C}\left(e^{j\omega}\right)}{S_{YY}^{C}\left(e^{j\omega}\right)} \qquad S_{YY}^{C}\left(e^{j\omega}\right)\neq 0$$

$$b_{LMMSE}=\eta_X-\eta_Y\cdot\sum_{n=-\infty}^{\infty}h[n]$$

מהגדרת התמרת פורייה נוכל לזהות כי הסכום מקיים למעשה $F\{h[n]\}\big|_{\omega=0}$ ולכן:

$$b_{LMMSE}=\eta_X-\eta_Y\cdot H_{LMMSE}(\omega=0)$$

לסיכום (גם עבור תהליכים jWSS בזמן רציף):

$$\hat{X}^{LMMSE}[n]=Y[n]*h_{LMMSE}[n]+b_{LMMSE} \qquad \hat{X}^{LMMSE}(t)=Y(t)*h_{LMMSE}(t)+b_{LMMSE}$$

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XY}^{C}\left(e^{j\omega}\right)}{S_{YY}^{C}\left(e^{j\omega}\right)} \qquad H_{LMMSE}(\omega)=\frac{S_{XY}^{C}(\omega)}{S_{YY}^{C}(\omega)}$$

$$b_{LMMSE}=\eta_X-\eta_Y\cdot H_{LMMSE}(\omega=0) \qquad b_{LMMSE}=\eta_X-\eta_Y\cdot H_{LMMSE}(\omega=0)$$

<u>הוכחה:</u>

נוכיח ע"י טענת העזר כי משערך ליניארי הינו אופטימלי אם"ם השגיאה ניצבת לכל פונקצייה ליניארית של המדידות.

כלומר אם"ם מתקיים $e_n\perp f^{Linear}\left(\{Y_n\},1\right)$ (ההוכחה זהה למקרים של ו"א או מ"א) כאשר:

$$f^{Linear}\left(\{Y_n\},1\right)=\sum_i a_i\cdot Y_i+c \quad \forall n,a_i,c$$

ניצבות כזאת שקולה לשתי הדרישות הבאות:

i. $e_n\perp 1$ (או כל קבוע אחר מ- $\mathbb{R}$)

ii. $e_n\perp\left(Y_m-\eta_Y\right) \quad \forall n,m$

לכן נרצה לבחור את המסנן הליניארי עם הפרמטרים $h[n],b$ אשר יקיימו דרישות אלו.
כעת נראה מה משמעות זוג דרישות אלו (כמובן שאנחנו משתמשים בשגיאה הפרשית כרגיל):

$$i \quad e_n \perp 1 \Leftrightarrow E\left[e_n \cdot 1\right] = 0$$
$$E\left[X[n] - Y[n] * h[n] - b\right] = 0$$
$$\eta_X - \eta_Y \cdot \sum_{n=-\infty}^{\infty} h[n] - b = 0$$
$$b(h) = \eta_X - \eta_Y \cdot \sum_{n=-\infty}^{\infty} h[n]$$

$$ii \quad E\left[e_n \cdot (Y_m - \eta_Y)\right] = 0$$
$$E\left[\left(X[n] - Y[n] * h[n] - b\right) \cdot (Y_m - \eta_Y)\right] = 0$$

נציב את $b(h)$ ונמשיך:

$$ii \quad E\left[\left(X[n] - Y[n] * h[n] - \eta_X + \eta_Y \cdot \sum_{n=-\infty}^{\infty} h[n]\right) \cdot (Y_m - \eta_Y)\right] = 0$$
$$E\left[\left(X[n] - \sum_{k=-\infty}^{\infty} h[k] \cdot Y[n-k] - \eta_X + \eta_Y \cdot \sum_{k=-\infty}^{\infty} h[k]\right) \cdot (Y_m - \eta_Y)\right] = 0$$
$$E\left[\left(\left(X[n] - \eta_X\right) - \sum_{k=-\infty}^{\infty} \left(h[k] \cdot \left(Y[n-k] - \eta_Y\right)\right)\right) \cdot (Y_m - \eta_Y)\right] = 0$$
$$C_{XY}(n-m) - \sum_{k=-\infty}^{\infty} h[k] \cdot C_{YY}(n-k-m) = 0$$

אך ניזכר כי הדבר צריך להתקיים $\forall n,m$ ולכן:

$$C_{XY}(n) - \sum_{k=-\infty}^{\infty} h[k] \cdot C_{YY}(n-k) = 0$$
$$C_{XY}(n) - h[n] * C_{YY}(n) = 0$$
$$C_{XY}(n) = h[n] * C_{YY}(n)$$
$$S_{XY}^C\left(e^{j\omega}\right) = H\left(e^{j\omega}\right) \cdot S_{YY}^C\left(e^{j\omega}\right) \quad \Rightarrow \quad H\left(e^{j\omega}\right) = \frac{S_{XY}^C\left(e^{j\omega}\right)}{S_{YY}^C\left(e^{j\omega}\right)}$$

וקיבלנו בדיוק את אותם המקדמים כפי שטענו.

<u>מערכת לשערוך LMMSE של שני ת"א jWSS</u>

ראינו כי מתקיים:

$$\hat{X}^{LMMSE}[n]=Y[n]*h_{LMMSE}[n]+b_{LMMSE}$$

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XY}^{C}\left(e^{j\omega}\right)}{S_{YY}^{C}\left(e^{j\omega}\right)},\qquad \begin{aligned} b_{LMMSE}&=\eta_X-\eta_Y\cdot H_{LMMSE}(\omega=0)\\ b_{LMMSE}&=\eta_X-\eta_Y\cdot\sum_{n=-\infty}^{\infty}h[n]\end{aligned}$$

נציב את האיבר החופשי b_{LMMSE} ונקבל:

$$\hat{X}^{LMMSE}[n]=Y[n]*h_{LMMSE}[n]+\eta_X-\eta_Y\cdot\sum_{n=-\infty}^{\infty}h[n]$$

נהפוך את המשוואה ונפתח:

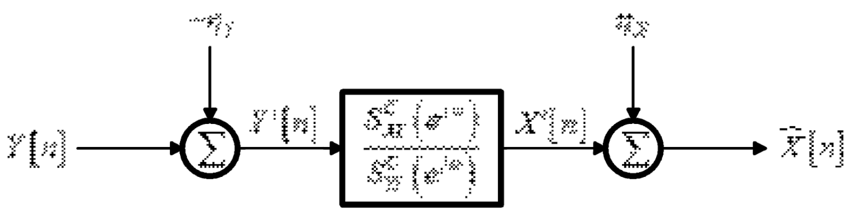
$$Y[n]*h_{LMMSE}[n]+\eta_X-\eta_Y\cdot\sum_{n=-\infty}^{\infty}h[n]=\hat{X}^{LMMSE}[n]$$

$$Y[n]*h_{LMMSE}[n]+\eta_X-\eta_Y*h_{LMMSE}[n]=\hat{X}^{LMMSE}[n]$$

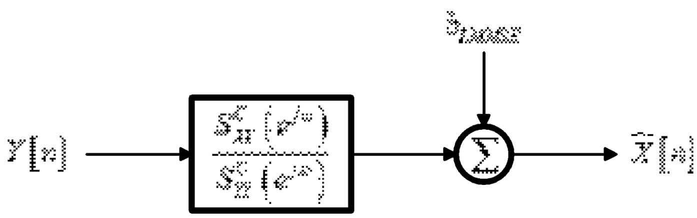
$$\left(Y[n]-\eta_Y\right)*h_{LMMSE}[n]=\hat{X}^{LMMSE}[n]-\eta_X$$

משמע אם נזין את $Y[n]-\eta_Y$ לתוך מערכת $h_{LMMSE}[n]$ כזאת נקבל במוצאה שערוך LMMSE של $X[n]-\eta_X$.

ולכן מערכת השערוך LMMSE של שני ת"א jWSS היא מהצורה:

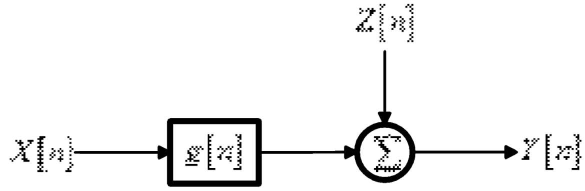


כמובן שנוכל לספק גם את המערכת ה"בסיסית":



<u>דוגמא:</u>

נתונה מערכת LTI המסומנת g אשר הבאה:

$Y[n]=X[n]*g[n]+Z[n]$	

המשוואה הזאת מתארת מערכת עם רעש אדיטיבי $Z[n]$ (חיבורי). נרצה לשחזר (לשערך) את התהליך $X[n]$.

נתון כי ת"א $X[n]$ הוא WSS עם $\eta_X=0$ וספקטרום הספק $S_{XX}\left(e^{j\omega}\right)$. כמו כן נתון כי ת"א $Z[n]$ הוא WSS עם $\eta_Z=0$ וספקטרום הספק $S_{ZZ}\left(e^{j\omega}\right)$. בנוסף ידוע כי התהליכים $X[n],Z[n]$ הם בת"ס.

מאחר והתוחלות של התהליכים $X[n],Z[n]$ הן אפס אז גם תוחלתו של תהליך המוצא $Y[n]$ תהיה אפס מאחר שאין במערכת קבוע חיבורי או תהליך אדיטיבי בעל תוחלת לא אפס.
לכן עבור כל התהליכים מתקיים כי פונקציית (העצמית והקרוס) הקורולציה והקוואריאנס זהות (למעשה הדבר היה מתקיים גם רק אם אחת התוחלות η_X,η_Z היתה אפס אבל אז לא היינו יכולים לקבוע כי $\eta_Y=0$). נראה באופן כללי עבור שני תהליכים כלשהם עם תוחלות אפס:

$$C_{AB}(t_1,t_2)\triangleq E\left[\left(A(t_1)-\eta_A\right)\cdot\left(B(t_2)-\eta_B\right)\right]=E\left[A(t_1)\cdot B(t_2)\right]\triangleq R_{AB}(t_1,t_2)$$

ידוע כי שני התהליכים $X[n],Z[n]$ הם WSS בנפרד ובת"ס מכאן שהם jWSS.
נחשב את פונקציית הקרוס-קורולצייה שלהם:

$$R_{XZ}(t_1,t_2)=E\left[X(t_1)\cdot Z(t_2)\right]\overset{\text{Independence}}{=}E\left[X(t_1)\right]\cdot E\left[Z(t_2)\right]=\eta_X\cdot\eta_Z=0\cdot 0\equiv 0$$

קיבלנו כי היא שווה זהותית לאפס ולכן ניתן להגיד שהיא לא תלויה בזמנים עצמם. מאחר ששני התהליכים WSS בנפרד ופונקציית הקרוס-קורולציה שלהם לא תלויה בזמנים הם jWSS.

מאחר שמתקיים לתהליכים אלו $R(\tau)=C(\tau)$ (עצמי וקרוס) אז מתקיים $S\left(e^{j\omega}\right)=S^C\left(e^{j\omega}\right)$.

נרצה להראות כי התהליכים $X[n],Y[n]$ הם jWSS על מנת להשתמש במסנן וינר (זהו מסנן LMMSE).

תהליך $Y[n]$ הינו WSS מאחר שהוא מעבר של תהליך WSS במערכת LTI עם אדיטיביות של תהליך WSS נוסף.

שני התהליכים $X[n],Y[n]$ הינם jWSS כפי שראינו במודל העזר.
לכן נוכל להשתמש במשערך וינר (משערך ליניארי אופטימלי במובן MSE):

$$\hat{X}^{LMMSE}[n]=Y[n]*h_{LMMSE}[n]+b_{LMMSE}$$

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XY}^{C}\left(e^{j\omega}\right)}{S_{YY}^{C}\left(e^{j\omega}\right)},\qquad \begin{aligned}b_{LMMSE}&=\eta_X-\eta_Y\cdot H_{LMMSE}(\omega=0)\\ b_{LMMSE}&=\eta_X-\eta_Y\cdot\sum_{n=-\infty}^{\infty}h[n]\end{aligned}$$

נפתח את פונקציית התמסורת של המסנן:

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XY}^{C}\left(e^{j\omega}\right)}{S_{YY}^{C}\left(e^{j\omega}\right)}\underset{C_{YY}=R_{YY}}{\overset{C_{XY}=R_{XY}}{=}}\frac{S_{XY}\left(e^{j\omega}\right)}{S_{YY}\left(e^{j\omega}\right)}$$

נמצא את פונקציות האוטוקורלציה על מנת שנוכל למצוא את ספקטרומי ההספק הדרושים לביטוי המסנן תוך שימוש במודל העזר. נסמן $Y'[n]=X[n]*g[n]$ ונקבל:

$$R_{XY}(n)=R_{XY'+XZ}(n)=R_{XY'}(n)+R_{XZ}(n)=R_{XY'}(n)=\delta(n)*R_{XX}(n)*g(-n)=R_{XX}(n)*g(-n)$$
$$R_{YY}(n)=R_{(Y'+Z)(Y'+Z)}(n)=R_{Y'Y'}(n)+R_{Y'Z}(n)+R_{ZY'}(n)+R_{ZZ}(n)$$

אך מתקיים כי $Y'[n],Z[n]$ הינם בת"ס מאחר שהזוג $X[n],Z[n]$ הינם בת"ס ו- $X[n]$ מתקבל ממעבר של במערכת LTI. בנוסף תוחלתו של $Z[n]$ הינה אפס (גם של $Y'[n]$) ולכן מתקיים $R_{Y'Z}(n)=0$. נוכל גם להראות:

$$R_{Y'Z}(n)\triangleq E\left[\left(X[n+\tau]*g[n+\tau]\right)\cdot Z[n]\right]\overset{\text{Independence}}{=}E\left[X[n+\tau]*g[n+\tau]\right]\cdot E\left[Z[n]\right]=0$$

ולכן פונקציית האוטו-קורולציה של $Y[n]$ תהיה:

$$R_{YY}(n)=R_{Y'Y'}(n)+R_{ZZ}(n)=g(n)*R_{XX}(n)*g(-n)+R_{ZZ}(n)$$

ולכן ספקטרומי ההספק יהיו:

$$S_{XY}\left(e^{j\omega}\right)=F\left\{R_{XY}(n)\right\}=S_{XX}\left(e^{j\omega}\right)\cdot G^{*}\left(e^{j\omega}\right)$$
$$S_{YY}\left(e^{j\omega}\right)=F\left\{R_{YY}(n)\right\}=S_{XX}\left(e^{j\omega}\right)\cdot\left|G\left(e^{j\omega}\right)\right|^{2}+S_{ZZ}\left(e^{j\omega}\right)$$

וביטוי המסנן יהיה:

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XX}\left(e^{j\omega}\right)\cdot G^{*}\left(e^{j\omega}\right)}{S_{XX}\left(e^{j\omega}\right)\cdot\left|G\left(e^{j\omega}\right)\right|^{2}+S_{ZZ}\left(e^{j\omega}\right)}$$

נבחן את ביצועי המערכת לפי נתון היחס אות לרעש (SNR = Sound\Signal to Noise Ratio):

- אם $SNR \gg 1$ אז למעשה הספק הרעש $S_{ZZ}\left(e^{j\omega}\right)$ זניח מול הספק האות ומתקיים:

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XX}\left(e^{j\omega}\right)\cdot G^{*}\left(e^{j\omega}\right)}{S_{XX}\left(e^{j\omega}\right)\cdot\left|G\left(e^{j\omega}\right)\right|^{2}+S_{ZZ}\left(e^{j\omega}\right)}\approx\frac{1}{G\left(e^{j\omega}\right)}$$

הדבר הגיוני כי אם הרעש זניח ביחס לאות אז מערכת השערוך רק צריכה "לבטל" את פעולת מערכת $g(\cdot)$.

- אם $SNR \ll 1$ אז הספק האות זניח מול הספק הרעש $S_{ZZ}\left(e^{j\omega}\right)$ ומתקיים:

$$H_{LMMSE}\left(e^{j\omega}\right)=\frac{S_{XX}\left(e^{j\omega}\right)\cdot G^{*}\left(e^{j\omega}\right)}{S_{XX}\left(e^{j\omega}\right)\cdot\left|G\left(e^{j\omega}\right)\right|^{2}+S_{ZZ}\left(e^{j\omega}\right)}\approx 0$$

הדבר הגיוני כי אם הרעש עצום יחסית לאות אז התהליך $Y[n]$ אשר ממנו מנסים לשערך את התהליך $X[n]$ מזוהם מידי ולכן "כדאי" למחוק אותו. במקרה זה השערוך יתבצע רק ע"י הקבוע החיבורי אשר אחראי על התוחלת וכמובן שזה השערוך הנבון ביותר במקרה הזה.

<u>שגיאת שערוך וינר</u>

בשערוך מ"א ראינו כי השגיאה היא מ"א בעצמה. כנ"ל לגביי ווקטורים אקראיים וכמובן לגביי תהליכים אקראיים.

אנו משתמשים בשגיאה הפרשית ולכן תהליך השגיאה נתון ע"י:

$$e[n]=X[n]-\hat{X}[n]$$

נרצה לבחון את תהליך השגיאה.

<u>טענה:</u>

תהליך השגיאה $e[n]$ הינו WSS ומקיים:

1. $E\left[e[n]\right]=0 \quad \forall n$, ז"א תוחלת השגיאה עבור כל דגימה הינה אפס.
2. $R_{ee}(k)=R_{XX}(k)-R_{\hat{X}\hat{X}}(k)=C_{XX}(k)-C_{\hat{X}\hat{X}}(k)$

<u>הוכחה:</u>

1. הדבר נובע מכך שמתקיים $e_n \perp 1$.
2. נראה כי הוא WSS וניתן ביטוי עבור פונקציית האוטו-קורולציה של התהליך. לצורך כך נתחיל עם פונקציית האוטו-קורולציה של תהליך המקור $X[n]$:

$$R_{XX}(k) \triangleq E\left[X(n+k)\cdot X(n)\right] = E\left[\left(\hat{X}[n+k]+e[n+k]\right)\cdot\left(\hat{X}[n]+e[n]\right)\right] =$$

$$= E\left[\hat{X}[n+k]\cdot\hat{X}[n]\right] + E\left[\hat{X}[n+k]\cdot e[n]\right] + E\left[e[n+k]\cdot\hat{X}[n]\right] + E\left[e[n+k]\cdot e[n]\right] =$$

$$\underset{\hat{X}=f^{Linear}(\{Y_n\},1)}{=} E\left[\hat{X}[n+k]\cdot\hat{X}[n]\right] + E\left[e[n+k]\cdot e[n]\right] \overset{*}{=} R_{\hat{X}\hat{X}}(k) + R_{ee}(n+k,n)$$

שיוויון $(*)$ מאחר שהמשערך $\hat{X}$ הוא מוצא של מערכת LTI אשר בכניסתה הת"א $Y[n]$ אשר WSS ולכן גם המוצא הוא ת"א WSS. נעביר אגפים ונקבל את פונקציית האוטו-קורולציה של השגיאה:

$$R_{ee}(n+k,n) = R_{XX}(k) - R_{\hat{X}\hat{X}}(k)$$

קיבלנו שפונקציית האוטוקורולציה של השגיאה תלויה בהפרש הזמנים בלבד ולכן היא WSS.

ניזכר כי לפי טענה 1 מתקיים $E\left[e[n]\right]=0$ ולכן:

$$E\left[e[n]\right]=0 \;\Rightarrow\; E\left[X[n]-\hat{X}[n]\right]=0 \;\Rightarrow\; E\left[X[n]\right]=E\left[\hat{X}[n]\right] \;\Rightarrow\; \eta_X=\eta_{\hat{X}}$$

נשתמש בקשר בין פונקציית הקורולציה לפונקציית הקו-ואריאנס ונקבל:

$$C_{AB}(k) = R_{AB}(k) + \eta_A\cdot\eta_B$$

לכן נוכל להסיק:

$$R_{ee}(k) = R_{ee}(n+k,n) = R_{XX}(k) - R_{\hat{X}\hat{X}}(k) = R_{XX}(k) - R_{\hat{X}\hat{X}}(k) + \eta_X^2 - \eta_X^2 =$$

$$\underset{\eta_X=\eta_{\hat{X}}}{=} R_{XX}(k) - R_{\hat{X}\hat{X}}(k) + \eta_X^2 - \eta_{\hat{X}}^2 = R_{XX}(k) + \eta_X^2 - \left(R_{\hat{X}\hat{X}}(k) + \eta_{\hat{X}}^2\right) = C_{XX}(k) - C_{\hat{X}\hat{X}}(k)$$

<u>מסקנה:</u>
ספקטרום ההספק של תהליך השגיאה נתון ע"י (התמרת פורייה של פונקציית האוטו-קורולציה של תהליך WSS):

$$S_{ee}\left(e^{j\omega}\right) = F\left\{R_{ee}(k)\right\} = S_{XX}\left(e^{j\omega}\right) - S_{\hat{X}\hat{X}}\left(e^{j\omega}\right) = S_{XX}^{C}\left(e^{j\omega}\right) - S_{\hat{X}\hat{X}}^{C}\left(e^{j\omega}\right)$$

אך המשערך נתון ע"י מעבר התהליך $Y[n]$ במסנן $h_{LMMSE}[n]$ (מערכת LTI ולכן קונבולוציה בזמן ומשמעה מכפלה בתדר) ולכן:

$$S_{ee}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)-S_{\hat{X}\hat{X}}^{C}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)-\left|H_{LMMSE}\left(e^{j\omega}\right)\right|^{2}\cdot S_{YY}^{C}\left(e^{j\omega}\right)$$

נציב את ביטוי המסנן ונקבל:

$$S_{ee}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)-\left|\frac{S_{XY}^{C}\left(e^{j\omega}\right)}{S_{YY}^{C}\left(e^{j\omega}\right)}\right|^{2}\cdot S_{YY}^{C}\left(e^{j\omega}\right)=S_{XX}^{C}\left(e^{j\omega}\right)-\frac{\left|S_{XY}^{C}\left(e^{j\omega}\right)\right|^{2}}{S_{YY}^{C}\left(e^{j\omega}\right)}$$

ולכן תוחלת השגיאה הריבועית היא:

$$MSE\triangleq E\left[e^{2}\left[n\right]\right]=R_{ee}\left(0\right)=F^{-1}\left\{S_{ee}\left(e^{j\omega}\right)\right\}\Big|_{n=0}=\frac{1}{2\pi}\cdot\int_{-\pi}^{\pi}S_{ee}\left(e^{j\omega}\right)\cdot e^{j\omega\cdot 0}d\omega=\frac{1}{2\pi}\cdot\int_{-\pi}^{\pi}S_{ee}\left(e^{j\omega}\right)d\omega$$

<u>הערה:</u>

נבחין כי הביטוי מזכיר את מבנה המשערך הרגיל של מ"א מתוך מ"א (שונות $\leftrightarrow$ ספקטרום).

המשמעות של ביטוי ספקטרום הספק השגיאה היא למעשה שכל תדר משוערך בנפרד. בפועל עבור כל תדר מופעל משערך סקלרי.

תהליכים אקראיים מתקדמים

בחלק זה נכיר שני תהליכים אקראיים מוכרים ופשוטים, תהליך וינר ותהליך פואסון. שני תהליכים אלו הם תהליכים בעלי תוספות בת"ס.

<u>תהליך אקראי בעל תוספות בת"ס</u>

ת"א $X[n]$ יקרא בעל תוספות בת"ס (או בקיצור – תהליך תוספות בת"ס) אם מתקיים שתוספות זרות שלו הן בת"ס.

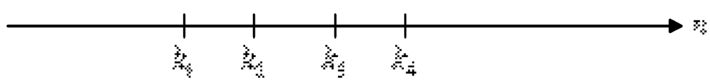
נסמן תוספת של תהליך $X[n]$ מזמן k_1 לזמן k_2 ע"י:

$$X(k_1,k_2)=X[k_2]-X[k_1]$$

תהליך $X[n]$ בעל תוספות בת"ס יקיים כי התוספות $X(k_1,k_2)$ ו- $X(k_3,k_4)$ הן בת"ס אם הקטעים $(k_1,k_2),(k_3,k_4)$ זרים זה לזה (אם הקטעים לא זרים לא נדע להגיד על התוספות כלום).

<u>דוגמא:</u>

עבור ת"א $X[n]$ בעל תוספת בת"ס וציר הזמן המצורף מתקיים:



- התוספות $X(k_1,k_2)$ ו- $X(k_3,k_4)$ הן בת"ס.
- התוספות $X(k_1,k_3)$ ו- $X(k_2,k_4)$ הן לא בהכרח בת"ס.

<u>תהליך Wiener / תנועת Brown</u>

תהליך וינר (Wiener) הוא מודל מתמטי לתיאור תנועת בראון (Brownian Motion) אשר נפגשנו בה כאשר רצינו לקבל מוטיביה על תהליכים אקראים – דוגמת "הילוך שיכור". תנועת בראון היא תנועה מרחבית אשר בכל רגע (זהו תהליך בזמן רציף) יכולה להיות תנועה לאחד משני הכיוונים בכל אחד מן הצירים. על מנת לפתח את המודל נתחיל עם תהליך הילוך שיכור חד-מימדי בזמן בדיד.

<u>תהליך שיכור חד-מימדי בזמן בדיד</u>

תהליך שיכור חד-מימדי בזמן בדיד נתון ע"י:

$$X[n]=X[n-1]+d\cdot W[n],\qquad X[0]=0$$

$$W[n]=W_n=\begin{cases}1 & w.p.\frac{1}{2}\\ -1 & w.p.\frac{1}{2}\end{cases}\qquad \{W_n\}\quad i.i.d$$

בכל רגע מתקבל "החלטה" עבור כיוון הצעד. הצעדים קבועים בגודל d . למעשה מתקיים:

$$X[n]=d\cdot\sum_{k=1}^{n}W_k$$

אנו יכולים לזהות לפי נוסחת הנסיגה כי זהו תהליך A.R. ליניארי עם מקדם $\alpha=1$ ולכן הוא לא סטציונרי.

נבחין כי תהליך השיכור הוא תהליך בעל תוספות בת"ס (נציג זוג תוספת $X(l,m),X(i,j)$ כאשר נניח ללא הגבלת הכלליות $l<m,i<j$):

$$X(l,m)=X[l]-X[m]=d\cdot\sum_{k=1}^{l}W_k-d\cdot\sum_{k=1}^{m}W_k=d\cdot\sum_{k=l+1}^{m}W_k$$

$$X(i,j)=X[i]-X[j]=d\cdot\sum_{k=1}^{i}W_k-d\cdot\sum_{k=1}^{j}W_k=d\cdot\sum_{k=i+1}^{j}W_k$$

מכיוון ש- $\{W_n\}$ סדרה i.i.d אז איבריה בת"ס הדדית. אם הקטעים $(l,m),(i,j)$ הם קטעים זרים אז בתוספות המתאימות לקטעים אלו, אשר בנויות מן הסכומים הנ"ל, אין איברים זהים מן הסדרה $\{W_n\}$.

התוספת $X(l,m)$ היא פונקציה (סכום) של איברים $W_{l+1},\dots,W_m$ אשר בת"ס הדדית. התוספת $X(i,j)$ היא פונקציה (סכום) של איברים שונים $W_{i+1},\dots,W_j$ אשר בת"ס הדדית. מכיוון שאם משתנים הם בת"ס אז גם פונקציות שלהם הן בת"ס (ומאחר שאם הקטעים זרים אין בתוספות איבר W_k משותף) אז נקבל $X(l,m)\perp\!\!\!\perp X(i,j)$.

<u>סטטיסטיקה עד סדר שני של תהליך שיכור חד-מימדי בזמן בדיד:</u>

התהליך נתון ע"י $X[n]=d\cdot\sum_{k=1}^{n}W_k$.

תוחלת התהליך תתאפס לכל n :

$$E\big[X[n]\big]=E\left[d\cdot\sum_{k=1}^{n}W_k\right]=d\cdot E\left[\sum_{k=1}^{n}W_k\right]=d\cdot\sum_{k=1}^{n}E[W_k]=d\cdot\sum_{k=1}^{n}\left(\frac{1}{2}\cdot1+\frac{1}{2}\cdot(-1)\right)=0$$

פונקציית הקורולצייה נתונה ע"י (נניח לשם הפיתוח $n>m$):

$$R_{XX}(n,m)\triangleq E\left[X[n]\cdot X[m]\right]\overset{*}{=}E\left[\left(X[m]+X(m,n)\right)\cdot X[m]\right]=$$
$$=E\left[\left(X[m]\right)^2\right]+E\left[X(m,n)\cdot X[m]\right]\overset{**}{=}E\left[\left(X[m]\right)^2\right]+E\left[X(m,n)\cdot X(0,m)\right]=$$
$$\overset{\text{Independence}}{=}E\left[\left(X[m]\right)^2\right]+E\left[X(m,n)\right]\cdot E\left[X(0,m)\right]=E\left[\left(X[m]\right)^2\right]+0\cdot 0=$$
$$=E\left[\left(X[m]\right)^2\right]=E\left[\left(d\cdot\sum_{k=1}^{m}W_k\right)^2\right]=d^2\cdot E\left[\left(\sum_{k=1}^{m}W_k\right)^2\right]$$

שיוויון $(*)$ מאחר שמתקיים $X[n]=X[n]+X[m]-X[m]=X[m]+X(m,n)$.

שיוויון $(**)$ מאחר שמתקיים $X[m]=X[m]+X[0]-X[0]=X(0,m)+X[0]\underset{X[0]=0}{=}X(0,m)$.

עבור $E\left[\left(\sum_{k=1}^{m}W_k\right)^2\right]$ מתקיים למעשה ריבוע של סכום איברי סדרה i.i.d. כל המכפלות המעורבות תתאפסנה מאחר שמתקיים $E\left[W_i\cdot W_j\right]\overset{\text{Independence}}{=}E\left[W_i\right]\cdot E\left[W_j\right]=0\cdot 0=0$. לכן נישאר רק עם המכפלות הלא-מעורבות ונקבל:

$$R_{XX}(n,m)\triangleq d^2\cdot E\left[\left(\sum_{k=1}^{m}W_k\right)^2\right]\underset{\{W_n\}\ i.i.d}{=}d^2\cdot E\left[\sum_{k=1}^{m}\left(W_k\right)^2\right]=d^2\cdot m$$

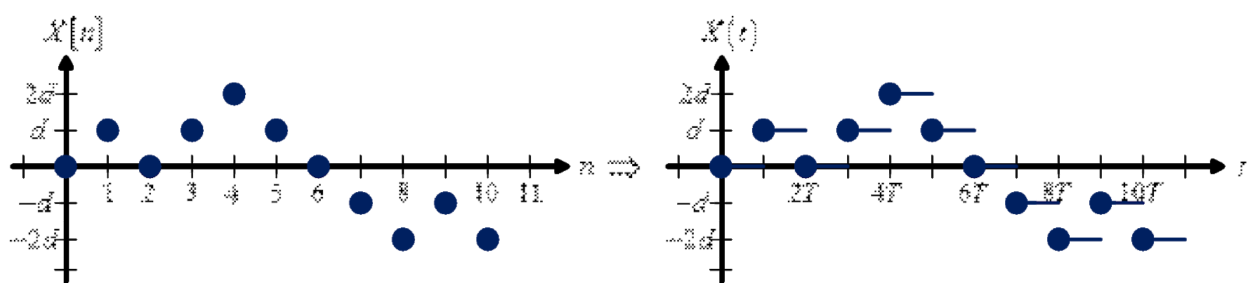
פונקציית הקורולציה תלויה לא רק בהפרש הזמנים ולכן התהליך לא סטציונרי.

באופן כללי עבור n,m כלשהם נקבל:

$$R_{XX}(n,m)\triangleq d^2\cdot\min(n,m)$$

<u>תהליך שיכור חד-מימדי בזמן רציף</u>

נרצה "להרציף" את תהליך השיכור בזמן בדיד כך שהשיכור יבצע את הצעד (עדיין בסיכוי חצי) בזמנים שהם מכפלה של T, כלומר בזמנים nT (ובשאר הזמן לא יזוז). באופן הזה ריאליזציה של התהליך תיראה באופן הבא:



נסמן את התהליך ע"י $X_{T,d}(t)$. נזהה כי עד זמן t למעשה מתבצעים $\left\lfloor \frac{t}{T} \right\rfloor$ צעדים נסמן גודל זה ע"י $N(t) = \left\lfloor \frac{t}{T} \right\rfloor$.

ביטוי מתמטי לתהליך השיכור יהיה:

$$X_{T,d}(t) = X\left[N(t)\right] = X\left[\left\lfloor \frac{t}{T} \right\rfloor\right] = d \cdot \sum_{k=1}^{N(t)} W_k = d \cdot \sum_{k=1}^{\left\lfloor \frac{t}{T} \right\rfloor} W_k = d \cdot \sum_{k=1}^{\infty} W_k \cdot U(t-kT)$$

תפקיד המדרגה הוא "להתעלם" מההחלטות המאוחרות מזמן t (להתחשב רק ב- $N(t)$ ההחלטות הראשונות) הרי המדרגה מתאפסת עבור החלטות k אשר גדולות מהחלטה מספר $N(t)$.

<u>סטטיסטיקה עד סדר שני של תהליך שיכור חד-מימדי בזמן רציף:</u>

התהליך נתון ע"י $X_{T,d}(t) = d \cdot \sum_{k=1}^{\infty} W_k \cdot U(t-kT)$.

כמובן שגם כעת תוחלת התהליך תתאפס לכל n מאותן סיבות:

$$E\left[X_{T,d}(t)\right] = 0$$

פונקציית הקורולצייה נתונה ע"י (עבור n,m כלשהם):

$$R_{XX}(t_1,t_2) = d^2 \cdot \min\left(N(t_1), N(t_2)\right) = d^2 \cdot \min\left(\left\lfloor \frac{t_1}{T} \right\rfloor, \left\lfloor \frac{t_2}{T} \right\rfloor\right)$$

כמובן שהתהליך לא סטציונרי (גם התהליך בזמן בדיד לא וגם ניתן לראות שוב כי הקורולציה תלויה בזמנים עצמם).

אם מתקיים $t_1, t_2 \gg T$ אז למעשה (מאחר שמתקיים $\left\lfloor \frac{x}{A} \right\rfloor \underset{x \gg A}{\approx} \frac{x}{A}$):

$$R_{XX}(t_1,t_2) \underset{t_1,t_2 \gg 1}{\approx} d^2 \cdot \min\left(\frac{t_1}{T}, \frac{t_2}{T}\right)$$

<u>תהליך וינר (פיתוח)</u>

נרצה לדמות את תהליך השיכור לתהליך אקראי בזמן רציף ולא מקוונטט. נרצה להפוך את השיכור לחלקיק אשר "בוחר" תנועות אינפיטיסימליות לאורך הציר כל פרק זמן אשר שואף לאפס.

אנו למעשה רוצים את גבול התהליך כאשר $T \to 0, d \to 0$ (ואז נקבל גרף תנועה חלק) באופן כזה שלסטטיסטיקה מסדר שני של התהליך יהיה גבול.
התנאי לכך הוא כמובן שיתקיים הגבול:

$$\lim_{\substack{T\to 0\\ d\to 0}}\left(d^2\cdot\frac{1}{T}\right)$$

דרישה לכך היא שביטוי הגבול יהיה קבוע חיובי סופי, כלומר $\frac{d^2}{T}=\alpha, \quad \alpha>0$.

למעשה מתקיים $d^2=\alpha\cdot T$. באופן זה אם אחד הפרמטרים T,d שואף לאפס אז גם השני.
נגדיר כעת את תהליך וינר ונסמנו .

נסמן את תהליך וינר ע"י $X_\alpha(t)$. ביטוי מתמטי לתהליך יהיה:

$$X_\alpha(t)=\lim_{\substack{T\to 0\\ d^2=\alpha\cdot T}}\left(X_{T,d}(t)\right)$$

<u>סטטיסטיקה עד סדר שני של תהליך וינר:</u>

התהליך נתון ע"י $X_\alpha(t)=\lim_{\substack{T\to 0\\ d^2=\alpha\cdot T}}\left(X_{T,d}(t)\right)$.

סטטיסטיקה עד סדר שני נתונה ע"י (כמובן שהתהליך לא סטציונרי):

$$E\left[X_\alpha(t)\right]=0 \quad \forall t$$
$$R_{XX}(t_1,t_2)=\alpha\cdot\min(t_1,t_2)$$

<u>הפילוג של תהליך וינר:</u>
נרצה לבחון את הפילוג של תהליך וינר. לצורך כך נבחן את התהליך לפני השאפת הגבול, ז"א את $X_{T,d}(t)$.

נקפיד על היחס $d^2=\alpha\cdot T$, ולכן נוכל לסמן פשוט $X_T(t)$.

$$X_T(t)=X_{T,d}(t)\Big|_{d^2=\alpha\cdot T}=\sqrt{\alpha\cdot T}\cdot\sum_{k=1}^{N(t)}W_k=\sqrt{\alpha\cdot T}\cdot\sqrt{\frac{t}{t}}\cdot\sum_{k=1}^{N(t)}W_k=\sqrt{\alpha\cdot t}\cdot\sqrt{T/t}\cdot\sum_{k=1}^{N(t)}W_k=$$
$$=\sqrt{\alpha\cdot t}\cdot\frac{1}{\sqrt{t/T}}\cdot\frac{\sqrt{N(t)}}{\sqrt{N(t)}}\cdot\sum_{k=1}^{N(t)}W_k=\sqrt{\alpha\cdot t}\cdot\frac{\sqrt{N(t)}}{\sqrt{t/T}}\cdot\frac{1}{\sqrt{N(t)}}\cdot\sum_{k=1}^{N(t)}W_k=$$
$$=\sqrt{\alpha\cdot t}\cdot\frac{\sqrt{\lfloor t/T\rfloor}}{\sqrt{t/T}}\cdot\frac{1}{\sqrt{N(t)}}\cdot\sum_{k=1}^{N(t)}W_k$$

כעת נחיל את הגבול $T \to 0$ מכיוון שאנו רוצים לדון בתהליך וינר.
אם קיים הגבול כולו נוכל להחיל אותו בחלקים לפי מכפלה:

1. $\dfrac{\sqrt{\left\lfloor t/T \right\rfloor}}{\sqrt{t/T}} \underset{T\to 0}{\to} 1$ (מכיוון שמתקיים $\left\lfloor t/T \right\rfloor \underset{T\to 0}{\overset{t\gg T}{\approx}} t/T$).

2. $\lim_{T\to 0}\left(\dfrac{1}{\sqrt{N(t)}}\cdot\sum_{k=1}^{N(t)} W_k\right) \sim N(0,1)$ (לפי משפט הגבול המרכזי).

ולכן קבילנו כי הפילוג של תהליך וינר מקיים:

$$X_\alpha(t) = \lim_{\substack{T\to 0 \\ d^2=\alpha\cdot T}}\left(X_{T,d}(t)\right) \sim \sqrt{\alpha\cdot t}\cdot N(0,1) = N(0,\alpha\cdot t)$$

סה"כ קיבלנו כי תהליך וינר מתפלג נורמלית עם תוחלת $\eta_X = 0$ ושונות $\sigma_X^2 = \alpha\cdot t$.
כמובן שתהליך וינר הוא בעל תוספות בת"ס כמו שתהליך הילוך השיכור כזה ולכן:

$$X_\alpha(t_1,t_2) = X_\alpha(t_2) - X_\alpha(t_1) \overset{*}{\sim} X_\alpha(t_2-t_1) \sim N\left(0,\alpha\cdot(t_2-t_1)\right)$$

הסבר מעבר $(*)$: הפרש הדגימות $X_\alpha(t_2)-X_\alpha(t_1)$ מתפלג כמו דגימת התהליך $X_\alpha(t_2-t_1)$ מאחר שזהו תהליך תוספות בת"ס לכן למעשה ההפרש מייצג התנהגות זהה לזו שהתהליך עשה מזמן t_1 עד זמן t_2.
המשמעות היא שגם התוספות מתפלגות גאוסית.
כעת נראה כי תהליך וינר הינו תהליך אקראי גאוסי (תזכורת: תא"ג הוא ת"א שכל קבוצת דגימות שלו היא קבוצה גאוסית במשותף – כל קומבינציה ליניארית שלה הוא מ"א גאוסי).

נבחן זוג דגימות של התהליך בזמנים t_1,t_2 (נניח $t_1<t_2$) ונבחין כי מתקיים:

$$\begin{aligned} X_\alpha(t_1) &= X_\alpha(0,t_1) \\ X_\alpha(t_2) &= X_\alpha(t_1)+X_\alpha(t_1,t_2) = X_\alpha(0,t_1)+X_\alpha(t_1,t_2) \end{aligned} \quad\Rightarrow\quad \begin{bmatrix} X_\alpha(t_1) \\ X_\alpha(t_2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\cdot\begin{bmatrix} X_\alpha(0,t_1) \\ X_\alpha(t_1,t_2) \end{bmatrix}$$

אנו יודעים כי זוג התוספות $X_\alpha(0,t_1), X_\alpha(t_1,t_2)$ הינן בת"ס כי הן על קטעי זמן זרים, בנוסף אנו יודעים כי כל אחת מן התוספות מתפלגת גאוסית כפי שכבר ראינו.

מאחר שהווקטור $\begin{bmatrix} X_\alpha(t_1) & X_\alpha(t_2) \end{bmatrix}^T$ נוצר מטרנספורמציה ליניארית של וא"ג גם הוא וא"ג ולכן כל איבריו גאוסיים. נוכל לבצע זאת עבור כל סט זמנים ולכן כל דגימותיו של התהליך יכולות להרכיב וא"ג – משמע התהליך הוא תהליך אקראי גאוסי (תא"ג).

<u>תהליך וינר (הגדרה פורמלית)</u>

תהליך וינר $X(t)$ עם מקדם α הוא ת"א בעל תוספות בת"ס גאוסיות המקיימות (ולמעשה תא"ג):

$$X(t_1,t_2)\sim N\left(0,\alpha\cdot(t_2-t_1)\right)$$

ואשר מקיים את תנאי ההתחלה $X(0)=0$.

הסטטיסטיקה של התהליך (מדובר בתא"ג ולכן סטטיסטיקה מסדר שני מאפיינת את התהליך):

$$E\left[X_\alpha(t)\right]=0 \quad \forall t$$
$$R_{XX}(t_1,t_2)=\alpha\cdot\min(t_1,t_2)$$

ניתן לראות כי התהליך לא סטציונרי (פונקציית הקורולציה שלו תלויה בזמנים עצמם).

למעשה כעת ניתן להסביר אינטואיטיבית את הסטטיסטיקה מסדר שני (נניח $t_1<t_2$):

$$\begin{aligned}R_{XX}(t_1,t_2)&\triangleq E\left[X(t_1)\cdot X(t_2)\right]=E\left[X(t_1)\cdot\left(X(t_1)+X(t_1,t_2)\right)\right]=E\left[X^2(t_1)\right]+\\&+E\left[X(t_1)\cdot X(t_1,t_2)\right]=E\left[X^2(t_1)\right]+E\left[X(0,t_1)\cdot X(t_1,t_2)\right]\underset{\text{Independence}}{\overset{*}{=}}E\left[X^2(t_1)\right]+\\&+E\left[X(0,t_1)\right]\cdot E\left[X(t_1,t_2)\right]=E\left[X^2(t_1)\right]+0\cdot 0=E\left[X^2(t_1)\right]\overset{**}{=}Var\left(X(t_1)\right)\overset{***}{=}\sigma_X^2=\alpha\cdot t_1\end{aligned}$$

שיוויון $(*)$ מאחר שזהו תהליך בעל תוספות בת"ס והקטעים $(0,t_1),(t_1,t_2)$ הינם זרים (תוחלת כל תוספת אפס).

שיוויון $(**)$ מאחר שידוע $\eta_X=0$ ולכן $\sigma_X^2=Var\left(X(t_1)\right)=E\left[X^2(t_1)\right]-\eta_X^2=E\left[X^2(t_1)\right]$.

שיוויון $(***)$ מאחר שדגימת התהליך ברגע מסויים t_1 שקולה לתוספת מרגע אפס ועד אותו רגע t_1 .

הדבר אינטואיטיבי מאחר שעבור שתי דגימות של התהליך קיימת קורולציה רק בתחום החפיפה ולכן למעשה מדובר בקורולציה עצמית של תחום זה, תחום החפיפה הינו תוספת כלשהי ולכן מתפלג גאוסית. קורולציה עצמית של מא"ג היא למעשה השונות ולכן קיבלנו $R_{XX}(t_1,t_2)=\alpha\cdot\min(t_1,t_2)$.

<u>תהליך הנגזרת של תהליך וינר</u>

נרצה לבחון את תהליך הנגזרת של תהליך וינר. לצורך כך נתבונן תחילה בתהליך השיפוע של תהליך וינר ונגדיר:

$$X_\varepsilon(t)=\frac{X(t+\varepsilon)-X(t)}{\varepsilon}=\frac{X(t,t+\varepsilon)}{\varepsilon}$$

תהליך השיפוע הינו ת"א (מאחר שהוא פונקציה של ת"א) גאוסי (בנוי כקומבינציה ליניארית של זוג דגימות מתא"ג – כל דגימותיו של תא"ג הינן גאוסיות במשותף).
ולכן תהליך הנגזרת הוא:

$$X'(t)=\frac{dX(t)}{dt}=\lim_{\varepsilon\to 0}X_\varepsilon(t)$$

תהליך הנגזרת גם הוא תא"ג מכיוון שהוא גבול של תא"ג (זו לא הוכחה פורמלית אך טענה נכונה).

מכיוון שתהליך הנגזרת $X'(t)$ הינו תא"ג אז מספיק למצוא סטטיסטיקה עד סדר שני כדי לאפיין אותו.
נמצא את הסטטיסטיקה עבור תהליך השיפוע ואז נפעיל את הגבול.
תוחלת התהליך:

פונקציית הקורולציה של התהליך (נניח לשם פשטות כי $\tau>0$):

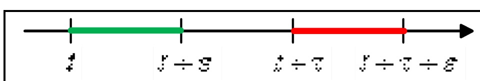
$$E\left[X_\varepsilon(t)\right]=E\left[\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right]=\frac{1}{\varepsilon}\Big(E\left[X(t+\varepsilon)\right]-E\left[X(t)\right]\Big)\underset{\eta_X=0}{\overset{Wiener}{=}}\frac{1}{\varepsilon}(0-0)=0$$

$$R_{X_\varepsilon X_\varepsilon}(t+\tau,t)\triangleq E\left[X_\varepsilon(t+\tau)\cdot X_\varepsilon(t)\right]=E\left[\frac{X(t+\tau+\varepsilon)-X(t+\tau)}{\varepsilon}\cdot\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right]=$$

$$=\frac{1}{\varepsilon^2}E\left[X(t+\tau,t+\tau+\varepsilon)\cdot X(t,t+\varepsilon)\right]$$

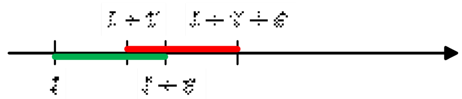
נבחין כי הצגנו את פונקציית הקורולציה ע"י זוג תוספות של תהליך וינר (זהו תהליך בעל תוספות בת"ס).
נפריד את הפיתוח לשני מקרים:

1. $0<\varepsilon<\tau$ (אין חפיפה בין הזמנים של שני תהליכי הנגזרת):

במקרה זה הקטעים זרים. לכן התוספות בת"ס, כ"א בעל תוחלת אפס.	

לכן במקרה זה פונקציית הקורולציה מתאפסת, כלומר $R_{X_\varepsilon X_\varepsilon}(t+\tau,t)=0$.

2. $0<\tau<\varepsilon$ (קיימת חפיפה בין הזמנים של שני תהליכי הנגזרת):

במקרה זה הקטעים לא זרים וקיים חיתוך ביניהם בתחום $(t+\tau,t+\varepsilon)$.	

נפריד את התוספות לחלקים נפרדים:

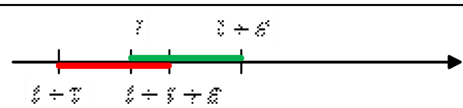
$$X(t,t+\varepsilon)=X(t,t+\tau)+X(t+\tau,t+\varepsilon)$$
$$X(t+\tau,t+\tau+\varepsilon)=X(t+\tau,t+\varepsilon)+X(t+\varepsilon,t+\tau+\varepsilon)$$

לאחר המכפלה כל התוספות של קטעים זרים נופלות כי מדובר בתוספות בת"ס בעל תוחלת אפס ונקבל:

$$R_{X_\varepsilon X_\varepsilon}(t+\tau,t)=\frac{1}{\varepsilon^2}\cdot E\left[X(t+\tau,t+\tau+\varepsilon)\cdot X(t,t+\varepsilon)\right]=\frac{1}{\varepsilon^2}\cdot E\left[X^2(t+\tau,t+\varepsilon)\right]=$$
$$\overset{*}{=}\frac{1}{\varepsilon^2}\cdot E\left[X^2\left((t+\varepsilon)-(t+\tau)\right)\right]=\frac{1}{\varepsilon^2}\cdot E\left[X^2(\varepsilon-\tau)\right]\overset{**}{=}\frac{1}{\varepsilon^2}\cdot a\cdot(\varepsilon-\tau)$$

שיוויונים $(*),(**)$ לפי תכונות תהליך וינר (בעל תוספות בת"ס + פילוג התוספות) כפי שניתן לראות שביצענו בפיתוח הסטטיסטיקה של תהליך וינר תחת ההגדרה הפורמלית.

פיתחנו עבור $\tau>0$, אם מתקיים $\tau<0$ אז הקטעים יראו:

$0<\varepsilon<\lvert\tau\rvert$:	
$0<\lvert\tau\rvert<\varepsilon$:	

גם כעת, במקרה הראשון, אין כלל חפיפה ולכן פונקצייה הקורולציה תתאפס ובמקרה השני יש חפיפה בקטע $(t,t+\tau+\varepsilon)$, אורך קטע זה הוא $\varepsilon+\tau$ בעוד שעבור $\tau>0$ קיבלנו קטע באורך $\varepsilon-\tau$. כזכור במקרה שלנו τ שלילי ולכן בשני המקרים נוכל לרשום באופן שקול $\varepsilon-\lvert\tau\rvert$.

בשני המקרים קיבלנו כי פונקציית הקורולציה תלויה בהפרש הזמנים בלבד ולכן התהליך WSS.
בסה"כ נוכל לרשום:

$$R_{X_\varepsilon X_\varepsilon}(t+\tau,t)=R_{X_\varepsilon X_\varepsilon}(\tau)=\begin{cases}0 & \lvert\tau\rvert>\varepsilon\\ \frac{1}{\varepsilon^2}\cdot a\cdot(\varepsilon-\lvert\tau\rvert) & \lvert\tau\rvert<\varepsilon\end{cases}$$

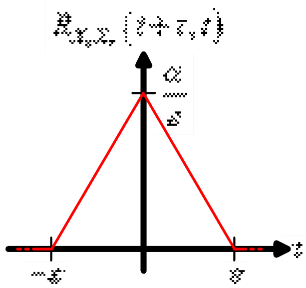
נציג את פונקציית הקורולציה באופן גרפי:

נזהה כי שטח המשולש הוא α וכלל לא תלוי ב- ε אלא ביחס בין הפרמטרים d,T.

בגבול הנגזרת, כאשר $\varepsilon \to 0$, נקבל דלתא בגובה α ז"א:

$$R_{X'X'}(t+\tau,t)=a\cdot\delta(\tau)$$

חשוב להדגיש כי זו לא פונקציית הלם רגילה כי השטח סופי.



אנו מקבלים למעשה:

$$R_{X'X'}(t+\tau,t)=R_{X'X'}(\tau)=\begin{cases}0 & |\tau|>\varepsilon\\ a\cdot\delta(\tau) & |\tau|<\varepsilon\end{cases}$$

כזכור, פונקציית קורולציה כזאת מתאימה לרעש לבן (בין כל זוג דגימות, קרובות ככל שנרצה לא תהיה קורולציה) עם ספקטרום הספק (צפיפות הספק ספקטרלית) α. אנו נוהגים לסמן $\alpha=\frac{N_0}{2}$ ע"מ לייצג צפיפות הספק חצי ספקטרלי.

רעש לבן הוא בעל הספק קבוע (בעוצמת גובה הדלתא) על כל התדרים. הספק הגדול הזה נובע מתהליך השיכור המקורי אשר מכיל בתוכו הרבה אנרגיה.

<u>תהליך אקראי Poisson</u>

תהליך אקראי פואסון הוא תהליך המנייה הפשוט ביותר, התהליך מונה את כמות האירועים בזמן עבור אירועים המתרחשים באופן אקראי לא תלות הדדית. זהו תהליך אקראי בזמן רציף עם ערכים בדידים.

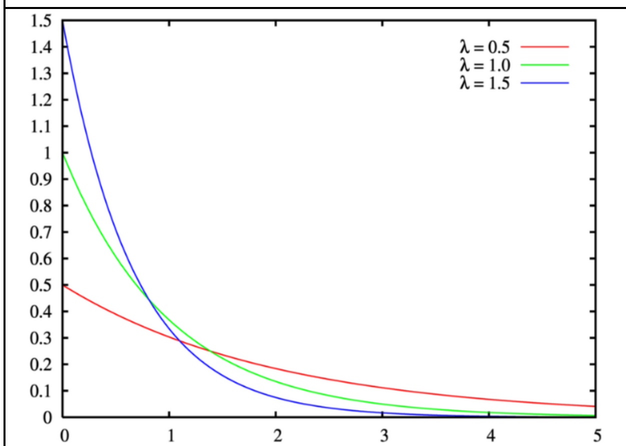
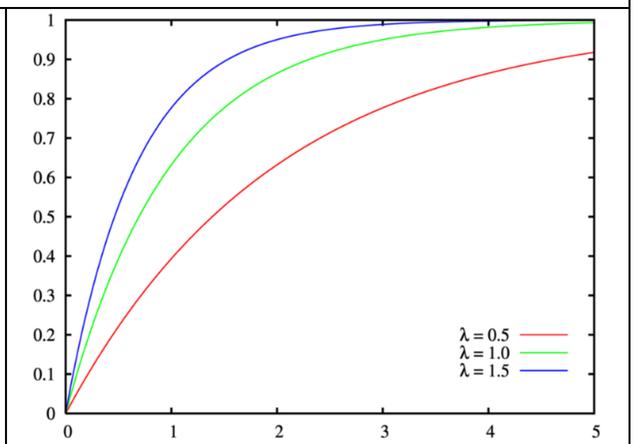
דוגמאות לתהליכים פואסוניים טבעיים:

- מספר טיפות גשם הנופלות בתוך כוס (או חתך עגול מוגדר)
- התפרקויות רדיואקטיביות
- תורים (בקירוב סדר ראשון)

בת"א פואסון התפלגות הזמן בין שני מאורעות עוקבים היא התפלגות מעריכית והתפלגות מספר האירועים המתרחשים בפרק זמן כלשהי היא התפלגון פואסון.

<u>התפלגות מעריכית</u>

התפלגות מעריכית היא התפלגות רציפה על המספרים האי-שליליים. התפלגות זו מאופיינת בתכונת חוסר-זיכרון ולכן מתארת תופעות אקאיות שהסיכוי להתרחשותן קבוע בזמן. להתפלגות פרמטר בודד אשר מייצג את קצב הדעיכה.

$X \sim \exp(\lambda)$	
	
$f_X(x)=\begin{cases}\lambda\cdot e^{-\lambda\cdot x} & x\geq 0\\ 0 & x<0\end{cases}$	$F_X(x)=\begin{cases}1-e^{-\lambda\cdot x} & x\geq 0\\ 0 & x<0\end{cases}$

<u>תכונת חוסר הזיכרון של משתנה אקראי אקספוננציאלי:</u>

כאמור, מ"א אקספוננציאלי מתאפיין בתכונת חוסר הזיכרון, כלומר, עבור $X\sim\exp(\lambda)$ מתקיים:

$$\Pr\left(X>x_1+x_2 \middle| X>x_1\right)=\Pr\left(X>x_2\right)$$

הוכחה -

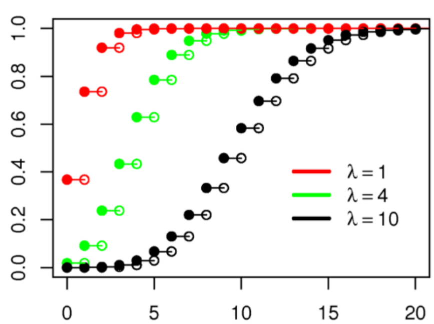
$$\Pr\left(X > x_1 + x_2 \middle| X > x_1\right) = \frac{\Pr\left(X > x_1 + x_2 \cap X > x_1\right)}{\Pr\left(X > x_1\right)} = \frac{\Pr\left(X > x_1 + x_2\right)}{\Pr\left(X > x_1\right)} \underset{X\sim\exp(\lambda)}{=} \frac{1 - \left[1 - e^{-(x_1+x_2)\cdot\lambda}\right]}{1 - \left[1 - e^{-x_1\cdot\lambda}\right]} =$$

$$= \frac{e^{-(x_1+x_2)\cdot\lambda}}{e^{-x_1\cdot\lambda}} = e^{-x_2\cdot\lambda} = \Pr\left(X > x_2\right)$$

<u>התפלגות פואסון</u>

התפלגות פואסון היא התפלגות של משתנה מקרי בדיד. ההתפלגות מתארת את הסיכוי לקיום מספר אירועים בפרק זמן מסויים כאשר האירועים מתרחשים בקצב מוגדר ובאופן בלתי תלוי. להתפלגות פרמטר בודד אשר מייצג את קצב האירועים.

מ"א פואסון עם פרמטר λ מסומן ע"י $X \sim Poisson(\lambda)$.

פונקציית ההסתברות של מ"א פואסון (הסיכוי לקבל k אירועים):	פונקציית ה-CDF של מ"א פואסון:
$\Pr\left(X = k\right) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$	

<u>בניית תהליך פואסון</u>

נבנה את ת"א פואסון ע"י גבול של ת"א בדיד. נתחיל עם ת"א ברנולי i.i.d.

<u>תהליך אקראי "חיי בלון" בזמן בדיד:</u>

חיי הבלון מתרחשים עד אירוע פיצוץ הבלון (פיצוף הבלון משול לאירוע המעניין – טיפה שנפלה בכוס לדוגמא).

נגדיר תהליך אקראי W_n של "ילד מעצבן" אשר בכל פרק זמן כלשהו הוא מנסה לפוצץ את הבלון. כל ניסיון הוא בעל סיכוי הצלחה p והמאורע יסומן $W_n = 1$. אי-הצלחה לא "מחלישה" את הבלון ולכן לא פוגעת בסיכויי הבלון בהמשך מכאן שסדרת הניסיונות הינה i.i.d.

$$W_n = \begin{cases} 1 & w.p. \quad p \\ 0 & w.p. \quad 1-p \end{cases}, \quad \begin{matrix} \{W_n\} \quad is \quad i.i.d \\ W_n \sim Ber(p) \end{matrix}$$

לצורך הגדרת הניסוי נניח שהילד "דפוק" והוא ממשיך עם הניסיונות גם לאחר שהבלון כבר התפוצץ.

נסמן את זמן חיי הבלון ע"י S^D , זמן חיי הבלון הוא מספר הניסיונות עד ההצלחה הראשונה.

מאחר שכל הניסיונות מתפלגים זהה $Ber(p)$ והם i.i.d אז זמן "הפיצוץ" הראשון (הצלחה ראשונה מנקודת המבט של התפלגות ברנולי) מתפלגת גאומטרית עם מקדם p . לכן מתקיים:

$$S^D = \min_n (W_n = 1) \sim Geo(p)$$

$$\Rightarrow \begin{array}{l} E\left[S^D\right] = 1/p \\ Var\left(S^D\right) = \dfrac{1-p}{p^2} \end{array}$$

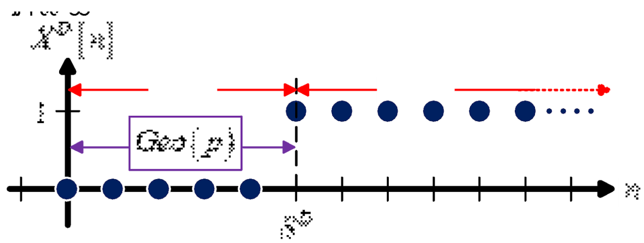
אם נרצה הצלחה ראשונה בזמן n הדבר שקול ל- $(n-1)$ כישלונות קודמים ואז הצלחה:

$$\Pr\left(S^D = n\right) = (p-1)^{n-1} \cdot p$$

נגדיר את תהליך "חיי בלון" בזמן בדיד ע"י:

$$X^D[n] = U\left(n - S^D\right)$$

פונקציית מדגם של התהליך נראית כך:



הערה - נבחין כי בדומה לת"א וינר יצרנו את התהליך מת"א ברנולי, אך בשונה לקחנו זמן מינימלי מבין השינויים בעוד שבת"א וינר סכמנו את השינויים (בנוסף בפיתוח ת"א וינר ל"כישלון" היה סיכוי שווה וערך -1).

<u>תהליך אקראי "חיי בלון" בזמן רציף:</u>

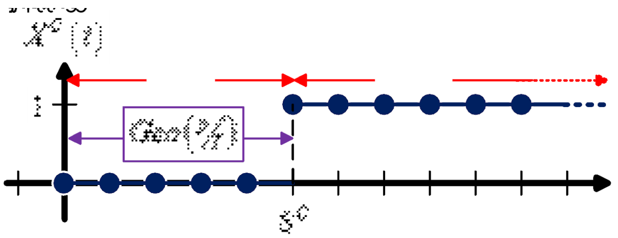
נרצה להרציף את תהליך "חיי בלון" בזמן בדיד. נניח כי ניסויי הפיצוץ מתרחשים בכל פרק זמן T.

לכן זמן הפיצוץ יהיה $S^C = T \cdot S^D$. כעת "צפיפות הצלחת הפיצוץ" היא סיכוי הצלחה p לזמן T ולכן זמן הפיצוץ מתפלג $S^C \sim Geo\left(p/T\right)$.

נגדיר את תהליך אקראי "חיי בלון" בזמן רציף:

$$X_{p,T}^C(t) = X^C(t) = U\left(n - T \cdot S^D\right) = U\left(n - S^C\right)$$

פונקציית מדגם של התהליך נראית כך:



<u>תהליך אקראי "חיי בלון" בגבול:</u>

נרצה להקטין את כך שהילד ינסה לפוצץ את הבלון "כל הזמן" אך באופן כזה שסיכוי ההצלחה שלו ליחידת זמן יישאר קבוע. לצורך כך נצטרך להקטין את T ואת p באותו האופן (אם נקטין רק אחד מהם או שהבלון יתפוצץ ישר או שהוא יחייה לנצח).

בכל יחידת זמן יש $1/T$ ניסיונות פיצוץ בעלי סיכוי p, לכן תוחלת ההצלחות ביחידת זמן זו היא p/T (סכום של $\{W_i\}\, i.i.d$). לכן נסמן את תוחלת מספר הפיצוצים המוצלחים ביחידת זמן (צפיפות) ע"י $\lambda = p/T$.

על מנת לשמור על הדרישה נשמור על הקשר $p = \lambda \cdot T$. באופן זה נוכל להשאיף את אחד מהפרמטרים p,T לאפס וגם השני ישאף לאפס וישמר היחס (מיחס זה נוכל לקבוע את פרקי הזמן T אם ידוע p ומקדם λ הרצוי).

$$X(t) = \lim_{\substack{T\to 0 \\ p=\lambda\cdot T}} \left(X^C_{p,T}(t) \right)$$

אנו יודעים כי התפלגות גיאומטרית (בדידה) שואפת להתפלגות אקספוננציאלית (רציפה) בגבול זה. לצורך כך נחשב את הסיכוי שהפיצוץ הראשון לא יתרחש עד רגע מסויים, כלומר יתרחש אחריו:

$$\Pr\left(S^C > t\right) = \Pr\left(T\cdot S^D > t\right) = \Pr\left(S^D > \frac{t}{T}\right) = \Pr\left(S^D > \left\lfloor \frac{t}{T} \right\rfloor\right) \underset{S^D\sim Geo(p)}{\overset{*}{=}} \left(1-p\right)^{\lfloor t/T \rfloor} \underset{p=\lambda\cdot T}{=} \left(1-\lambda\cdot T\right)^{\lfloor t/T \rfloor} =$$

$$= \left(1-\lambda\cdot T\right)^{\lfloor t/T \rfloor \cdot \frac{t/T}{t/T}} = \left(1-\lambda\cdot T\right)^{\frac{\lfloor t/T \rfloor}{t/T}\cdot (t/T)} \underset{T\to 0}{\to} \left(1-\lambda\cdot T\right)^{1\cdot (t/T)} = \left(\left[\left(1-\lambda\cdot T\right)^{\frac{1}{T}} \right]_{T\to 0} \right)^t \overset{**}{=} \left(e^{-\lambda}\right)^t = e^{-\lambda t}$$

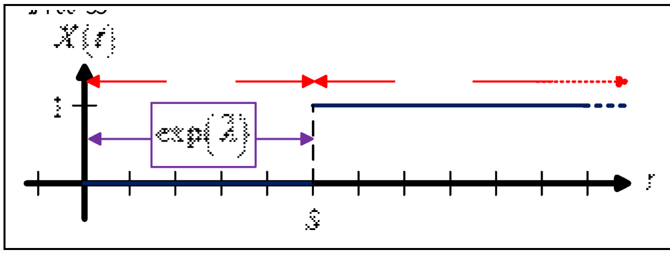
שיוויון $(*)$ מתקיים מאחר ש- $S^D \sim Geo(p)$ ואנו מחפשים את ההיפך מ-CDF ולכן התשובה היא:

$$1-F_{S^D}\left(\left\lfloor t/T \right\rfloor\right)=1-\left[1-(1-p)^{\lfloor t/T\rfloor}\right]=(1-p)^{\lfloor t/T\rfloor}$$

שיוויון $(**)$ מתקיים לפי קיום הגבול (עבור $x=-\lambda,\quad n=1/T$):

$$\lim_{n\to\infty}\left[\left(1+\frac{x}{n}\right)^n\right]=e^x$$

ביטוי ופונקציית מדגם של התהליך:



$$X(t)=\lim_{\substack{T\to 0\\ p=\lambda\cdot T}}\left(X_{p,T}^{C}(t)\right)=U(t-S)$$

$$S\sim\exp(\lambda)\qquad \lambda=\frac{p}{T}$$

<u>תהליך אקראי "חיי N בלונים":</u>

פעם ניקח אוסף של N בלונים בת"ס הדדית (זמני הפיצוץ בלתי תלויים הדדית) ונסמן את זמני הפיצוץ של כל בלון:

$$S_1,\dots,S_N\sim\exp(\lambda),\quad \{S_i\}_{i=1}^{N}\ i.i.d$$

עכשיו יש גם N "ילדים מעצבנים". אם ילד מסויים פוצץ את הבלון שלו הוא לא מקבל בלון חדש ולכן אפקטיבית הוא יצא מן המשחק, שאר הילדים ממשיכים ב"משחק" ללא תלות עד שבסוף כל אחד מצליח לפוצץ את הבלון שלו.
לכן התהליך נתון ע"י (יראה כמו גרם מדרגות):

$$X(t)=\sum_{i=1}^{N}U(t-S_i)$$

נתעניין לדעת סיכויים לגבי זמן התרחשות הפיצוץ הראשון, השני וכן הלאה. באופן כלל נרצה לדעת סיכויים לגבי זמן הפיצוץ ה- i (אנחנו לא מחפשים לדעת מתי הבלון ה- i התפוצץ כי זו פשוט התפלגות אקספוננציאלית מאחר שהם i.i.d אלא מתי הפיצוץ ה- i מבין כל N הבלונים התרחש). אנו מתעניינים בזמנים אלו מאחר שלא מעניין אותנו מתי תיפול לכוס טיפה מסויימת אשר נמצא באוויר אלא מתי תיפול לכוס הטיפה הבאה.

נסמן את סדרת זמני הפיצוץ מבין כל N הבלונים ע"י (זו למעשה דוגמא של Order Statistics):

$$S^1,\dots,S^N$$

נבחין כי מתקיים:

$$S^1 = \min\left(S_1,\ldots,S_N\right), S^2 = \min\left(S_1,\ldots,S_N \setminus S^1\right),\ldots$$

נמצא את הפילוג של זמן הפיצוץ הראשון, כלומר של S^1 :

$$\Pr\left(S^1 > t\right) \overset{*}{=} \Pr\left(S_1 > t \cap S_2 > t \cap \ldots \cap S_N > t\right) \underset{\text{Independence}}{\overset{i.i.d}{=}} \prod_{i=1}^{N} \Pr\left(S_i > t\right) \overset{i.i.d}{=} \left[\Pr\left(S_i > t\right)\right]^N \underset{\text{S}\sim\exp(\tilde{\lambda})}{=} e^{-\tilde{\lambda}\cdot N}$$

שיוויון $(*)$ מאחר שע"מ שהפיצוץ הראשון יהיה אחרי זמן t אז כל הבלונים צריכים להתפוצץ לאחר זמן t .

נגדיר מקדם אפקטיבי $\lambda = \lambda_{eff} = \tilde{\lambda}\cdot N$, אנו מבחינים כי ככל שיהיו יותר ילדים ככה המקדם האפקטיבי יגדל, הדבר הגיוני לתיאור הניסוי כי אם יש יותר ילדים אז סביר שהפיצוץ הראשון יתרחש מוקדם יותר.
סה"כ קיבלנו כי הפילוג של הפיצוץ הראשון הוא:

$$S^1 \sim \exp\left(\lambda\right), \qquad \lambda = \lambda_{eff} = \tilde{\lambda}\cdot N, \quad \left(\tilde{\lambda} = \frac{p}{T}\right)$$

משמעות תכונת חוסר הזיכרון של מ"א אקספוננציאלי בניסוי שלנו היא שהסיכוי שהפיצוץ לא התרחש עד זמן $t_1 + t_2$ בהינתן הידיעה שהוא לא התפוצץ עד זמן t_1 שווה להסתברות שהוא לא יתפוצץ זמן t_2 . כלומר, נשאר לו לשרוד t_2 זמן והעובדה שהוא כבר שרד t_1 זמן לא משפיעה.
כלומר:

$$\Pr\left(S^1 > t_1 + t_2 \middle| S^1 > t_1\right) = \Pr\left(S^1 > t_2\right)$$

בלון התפוצץ ראשון. לא יודעים איזה בלון אבל יודעים שהפיצוץ התרחש בזמן S^1 . נרצה לבחון את המצב של בלון S_i , נזהה כי הטענה $S_i \neq S^1$ שקולה לכך שלמעשה מתקיים $S_i > S^1$ מכיוון שאם זמן הפיצוץ שלו הוא לא זמן הפיצוץ הראשון אז הוא לא התפוצץ עד זמן זה.

בהינתן המצב הזה נרצה לדעת מה הסיכוי שבלון זה ישרוד עוד t זמן, מתכונת חוסר הזיכרון נקבל שהתשובה היא ספירת הזמן מהרגע האחרון שאנו יודעים עליו – כלומר הסיכוי שהוא ישרוד תקופה של t זמן.

$$\Pr\left(S_i > S^1 + t \middle| S_i \neq S^1\right) = \Pr\left(S_i > S^1 + t \middle| S_i \neq S^1\right) = \Pr\left(S_i > S^1 + t \middle| S_i > S^1\right) \underset{S_i\sim\exp(\tilde{\lambda})}{=} \Pr\left(S_i > t\right) = e^{-\tilde{\lambda}\cdot t}$$

לכן הפילוג של זמן הפיצוץ השני, כלומר של S^2 , הוא:

$$\Pr\left(S^2 > S^1 + t\right) \overset{*}{=} \Pr\left(\left(S_1 > S^1 + t \middle| S_1 > S^1\right) \cap \ldots \cap \left(S_{N-1} > S^1 + t \middle| S_{N-1} > S^1\right)\right) =$$

$$\underset{\text{Independence}}{\overset{i.i.d}{=}} \prod_{i=1}^{N-1} \Pr\left(S_i > S^1 + t \middle| S_i > S^1\right) \overset{i.i.d}{=} \left[\Pr\left(S_i > S^1 + t \middle| S_i > S^1\right)\right]^{N-1} \underset{\text{S}\sim\exp(\lambda)}{=} \left[\Pr\left(S_i > t\right)\right]^{N-1} =$$

$$\underset{S_i\sim\exp(\lambda)}{=} \left(e^{-\lambda \cdot t}\right)^{N-1} = e^{-\lambda \cdot (N-1) \cdot t}$$

שיוויון $(*)$ מאחר שע"מ שהפיצוץ השני יהיה אחרי זמן $S^1 + t$ אנו בעצם דורשים שכל הבלונים שנותרו (מספרם לא באמת משנה אלא רק הכמות שנותרה) ישרדו כ"א את הזמן הנדרש. קיבלנו כי הסיכוי שהפיצוץ השני יקרה מאוחר מ- t זמן אחרי הפיצוץ הראשון (ז"א לפחות $S^1 + t$ זמן) הוא למעשה כמו הישרדות הפיצוץ הראשון במשך זמן t במערכת של בלונים $N-1$.

מסקנה:

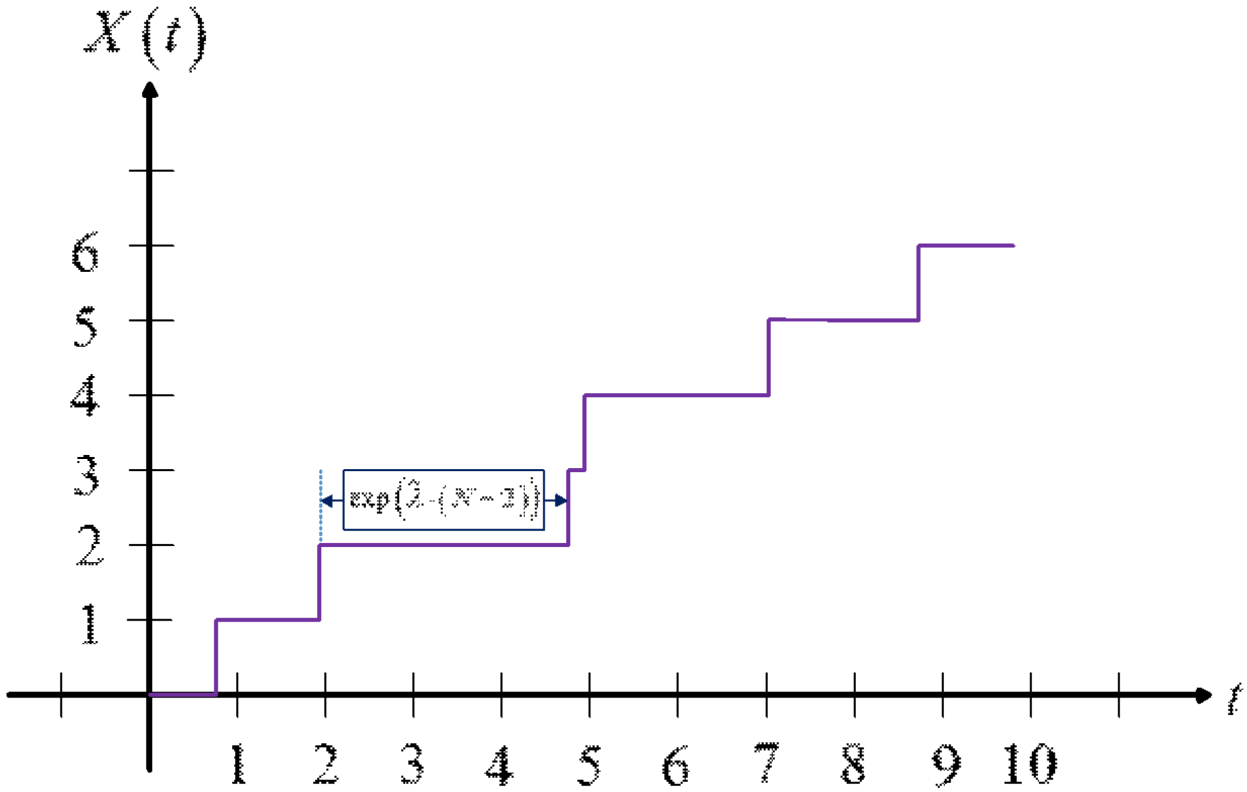
$$\begin{aligned} S^1 &\sim \exp\left(\lambda \cdot N\right) \\ S^2 - S^1 &\sim \exp\left(\lambda \cdot (N-1)\right) \\ &\vdots \\ S^N - S^{N-1} &\sim \exp\left(\lambda\right) \end{aligned} \quad \Rightarrow \quad \begin{aligned} & S^k - S^{k-1} \sim \exp\left(\lambda \cdot (N+1-k)\right) \quad \forall 1 \le k \le N \\ & S^0 = 0 \end{aligned}$$

נראה שעבור הפיצוצים המאוחרים יש קבוע דעיכה קטן יותר משל המוקדמים, לכל ההפרשים בין הפיצוצים הולכים וגדלים. הדבר הגיוני כי יש פחות ילדים שמשחקים ולכן פחות סיכוי (הם בת"ס הדדית). הדבר מתיישב גם מתמטית מאחר שתוחלת של $X \sim \exp(\lambda)$ מקיימת $E[X] = 1/\lambda$, כלומר ככל שקבוע הדעיכה קטן יותר כך ייקח יותר זמן בממוצע.

עכשיו כאשר נתונים לנו פילוגי זמני הפיצוץ נוכל לרשום את התהליך על פיהם ולא על פי זמני הפיצוץ של כל בלון (השוויון תקף מאחר שלמעשה הקבוצות $\{S_i\}_{i=1}^N, \{S^i\}_{i=1}^N$ מחזיקות באותם איברים אך בסדר שונה):

$$X(t) = \sum_{i=1}^{N} U\left(t - S_i\right) = \sum_{i=1}^{N} U\left(t - S^i\right)$$

מכיוון שאנו יודעים את ההתפלגויות נוכל להציג פונקציית מדגם של התהליך באופן איכותי:



<u>תהליך אקראי "חיי N בלונים" עם החזרה:</u>

נניח כעת שכאשר ילד מצליח לפוצץ בלון הוא לא יוצא מהמשחק אלא מקבל מיידית בלון חדש בשביל להמשיך לשחק. כעת זמני הפיצוץ יתפלגו כולם באופן זהה. למשחק זה אין סוף ולכן התהליך ימשיך לנצח (דוגמא טבעית שקולה היא אנודה-קטודה בהן אנו מחזיקים מתחים קבועים והאירועים הם קפיצות האלקטרון). נקבל:

$$S^k - S^{k-1} \sim \exp\left(\lambda \cdot N\right)$$

הדבר שקול למקרה הטבעי בו $N \gg 1$ ואז נקבל שמתקיים $N-1 \approx N$, לכן קבוע הדעיכה של שני זמני הפיצוץ הראשונים שווה $\lambda \cdot (N-1) \approx \lambda \cdot N$. למעשה מתקיים $N-k \approx N$ עד ערך k מסויים. המשמעות היא שכל הפיצוצים ה" k -ראשונים" מתפלגים באותו האופן:

$$S^k - S^{k-1} \underset{approx.}{\sim} \exp\left(\lambda \cdot N\right) \qquad \frac{N-k}{N} \approx 1$$

<u>תהליך אקראי פואסון:</u>

ניקח $N \to \infty, \lambda \to 0$ כך שמתקיים שמכפלתם מניבה קבוע דעיכה קבוע וסופי $\lambda = \lambda_{eff} = N \cdot \lambda$.

נקבל שכל מרווחי המדרגות יתפלגו באופן זהה $\exp\left(\lambda\right)$. התהליך יסומן $X\left(t\right) \sim Poisson\left(\lambda \cdot t\right)$ וינתן ע"י הביטוי:

$$X\left(t\right) = \sum_{i=1}^{\infty} U\left(t - S_i\right) = \sum_{i=1}^{\infty} U\left(t - S^i\right), \qquad S^i - S^{i-1} \sim \exp\left(\lambda\right), S^0 = 0$$

הערה:

נוכל להתייחס לתהליך ככזה אשר נבנה מן התוספות $S^i - S^{i-1}$. תוספות אלו מקיימות:

$$\begin{array}{ll} S^i = S^{i-1} + Z_i & Z_i \sim \exp\left(\lambda\right) \quad \forall i \\ S^0 = 0 & \left\{Z_i\right\} i.i.d \end{array}$$

כאשר $\left\{Z_i\right\}$ הם למעשה מרווחי הפיצוצים. זהו תהליך A.R. ליניארי עם מקדם $\alpha = 1$ ולכן לא סטציונרי.

<u>תהליך תוספות של תהליך אקראי פואסון</u>

עבור תהליך פואסון עם פרמטר λ , $X(t) \sim Poisson(\lambda \cdot t)$, נגדיר תהליך תוספת $X(t_1,t_2) = X(t_2) - X(t_1)$. תוספת $X(t_1,t_2)$ למעשה מייצגת את מספר הפיצוצים באינטרוול זמן $[t_1,t_2]$. תהליך התוספות הזה מתפלג באופן זהה לתהליך כולו (מחוסר הזיכרון ומן הגבול, הרי בכל נקודת זמן יש אותה כמות ילדים עם אותה כמות בלונים וכשלונות/התלחות העבר לא משפיעות) ובנוסף הוא תהליך בעל תוספות בת"ס. לכן מתקיים:

$$X(t_1,t_2) \sim Poisson\left(\lambda\left(t_2 - t_1\right)\right), \qquad X(t_1,t_1+t) \sim Poisson(\lambda t)$$

<u>פילוג של תהליך תוספות של ת"א פואסון:</u>

נפתח את הפילוג של תהליך התוספות של ת"א פואסון על מודל ה"חיי N בלונים" עם החזרה כאשר $N \to \infty$.

$$\Pr\left(X(t_1,t_1+t) = k\right) \overset{*}{=} \Pr\left(X(t) = k\right)$$

שיוויון $(*)$ מתקיים מתכונת חוסר הזיכרון של תהליך פואסון, ז"א הסיכוי שבפרק זמן באורך t התפוצצו k בלונים לא תלוי במיקום הזמן כלומר אין חשיבות אם באינטרוול הזמן $(t_1,t_1+t]$ או באינטרוול הזמן $(0,t]$.

על מנת לקבל בדיוק k פיצוצים אנחנו דורשים ש- $(N-k)$ בלונים ישרדו (לא משנה איזה יתפוצצו/ישרדו).

$$\Pr\left(X(t_1,t_1+t) = k\right) = \Pr\left(X(t) = k\right) = \binom{N}{k} \cdot \left(1 - e^{-\tilde{\lambda} t}\right)^k \cdot \left(e^{-\tilde{\lambda} t}\right)^{N-k} \underset{\lambda = \tilde{\lambda} N}{=} \binom{N}{k} \cdot \left(1 - e^{-\frac{\lambda}{N}t}\right)^k \cdot \left(e^{-\frac{\lambda}{N}t}\right)^{N-k}$$

נרצה לסדר את הביטוי לפני השאפת $N \to \infty$. נסדר ביטויים אשר ישאפו לאחד.

$$\binom{N}{k} \cdot \left(1 - e^{-\frac{\lambda}{N}t}\right)^k \cdot \left(e^{-\frac{\lambda}{N}t}\right)^{N-k} = \frac{N!}{k! \cdot (N-k)!} \cdot \left(1 - e^{-\frac{\lambda t}{N}}\right)^k \cdot \frac{\left(\lambda t / N\right)^k}{\left(\lambda t / N\right)^k} \cdot \left(e^{-\frac{\lambda t}{N}}\right)^N \cdot \left(e^{-\frac{\lambda t}{N}}\right)^{-k} =$$

$$= \frac{1}{k!} \cdot \frac{N \cdot \ldots \cdot (N-k+1)}{N^k} \cdot \frac{(N-k)!}{(N-k)!} \cdot \frac{\left(1 - e^{-\frac{\lambda t}{N}}\right)^k}{\left(\lambda t / N\right)^k} \cdot (\lambda t)^k \cdot e^{-\lambda t} \cdot \left(e^{-\frac{\lambda t}{N}}\right)^{-k}$$

ניזכר כי מתקיים $e^x=1+x+\dots$ ולכן $1-e^x\approx -x$. ולכן נקבל:

$$\Pr\left(X\left(t_1,t_1+t\right)=k\right)\approx\frac{1}{k!}\cdot\frac{N\cdot\ldots\cdot\left(N-k+1\right)}{N^k}\cdot\frac{\left(\frac{\lambda t}{N}\right)^k}{\left(\lambda t/N\right)^k}\cdot\left(\lambda t\right)^k\cdot e^{-\lambda t}\cdot\left(e^{-\frac{\lambda t}{N}}\right)^{-k}=$$

$$=\frac{1}{k!}\cdot\frac{N\cdot\ldots\cdot\left(N-k+1\right)}{N^k}\cdot\left(\lambda t\right)^k\cdot e^{-\lambda t}\cdot\left(e^{-\frac{\lambda t}{N}}\right)^{-k}\xrightarrow[N\to\infty]{}\frac{1}{k!}\cdot 1\cdot\left(\lambda t\right)^k\cdot e^{-\lambda t}\cdot 1=\frac{1}{k!}\cdot\left(\lambda t\right)^k\cdot e^{-\lambda t}$$

וקיבלנו כי לתהליך התוספות פילוג זהה כצפוי.

<u>תהליך פואסון (הגדרה פורמלית)</u>

נוכל לשכוח מתהליך הבניה המפרך שביצענו ו"לקבל" הגדרות (שקולות) לתהליך פואסון.

תהליך פואסון $X(t)\sim Poisson(\lambda\cdot t)$ (מוגדר לכל $t\geq 0$) הוא למעשה:

1. נבנה מתהליך A.R. של זמני פיצוץ (A.R. ליניארי עם מקדם $\alpha=1$ ולכן לא סטציונרי):

$$S^i=S^{i-1}+Z_i \qquad Z_i\sim\exp(\lambda)\quad\forall i$$
$$S^0=0 \qquad \{Z_i\}\ i.i.d$$

כך שזמני הפיצוץ בונים את התהליך עצמו ע"י:

$$X(t)=\sum_{n=1}^{\infty}U\left(t-S^n\right),\quad X(0)=0$$

הערה - מזמני הפיצוץ נזהה כי:

$$S^n=\sum_{i=1}^{n}Z_i$$

הסדרה הינה i.i.d ולכן זמני הפיצוץ הוא סכום של n מ"א i.i.d. מכאן שמתקיים:

$$f_{S^n}=f_{Z_1}*\ldots*f_{Z_n}$$

זהו סכום n של מ"א אספוננציאליים i.i.d. סכום כזה מתפלג לפי פילוג Erlang:

$$f_{S^n}(t)=\frac{\lambda^n\cdot t^{n-1}\cdot e^{-\lambda\cdot t}}{(n-1)!}$$

נוכל לזהות שפילוג ארלנג מתכלס לפילוג אקספוננציאלי עבור $n=1$.

2. תהליך אשר:

   a. מקיים תנאי התחלה $X(0)=0$.

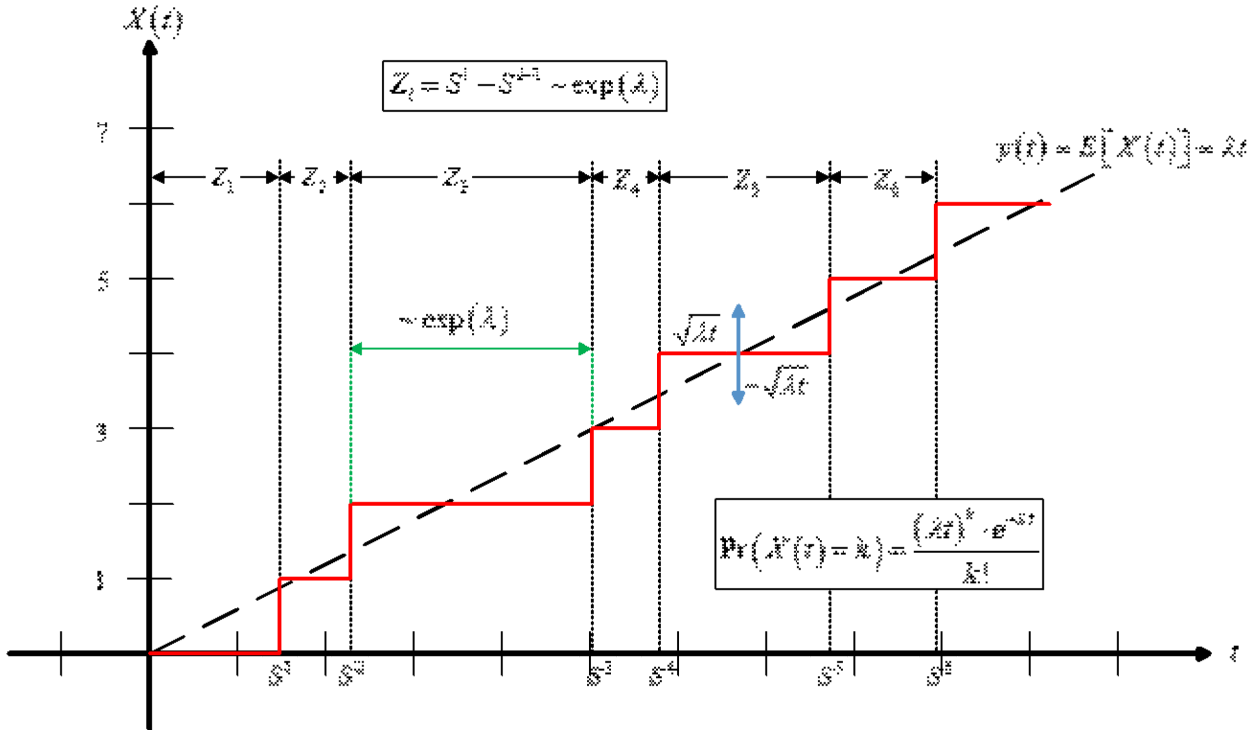
b. Non-negative-integer-valued monotonically non-decreasing step-function
ז"א מקבל ערכים אי-שליליים שלמים בלבד, מונוטוני לא יורד ועשוי ממדרגות.

c. בעל תוספות בת"ס אשר מתפלגות פואסונית $X(t_1,t_2) \sim Poisson(\lambda \cdot (t_2 - t_1))$.

הערה - תוספת $X(t_1,t_2) = X(t_2) - X(t_1)$ למעשה מייצגת את מספר הפיצוצים באינטרוול זמן $[t_1,t_2]$, תהליך התוספות מתפלג בעצמו פואסונית (הרי בכל נקודת זמן יש אותה כמות ילדים עם אותה כמות בלונים וכשלונות/התלחות העבר לא משפיעות).

<u>תכונות (ללא הוכחה):</u>

- $E[X(t)] = \lambda t$, תוחלת התהליך ליניארית בזמן. למעשה נצפה שפונקציית המדרגות תתיישר עם הישר $y(t) = \lambda t$. הדבר הגיוני זה תהליך מונוטוני לא-יורד לכן ברור שהתוחלת צריכה לעלות, בנוסף התהליך מתנהג אותו דבר כל הזמן ולכן גם נצפה לקו ישר (אשר מתחיל בראשית כמו ערך ההתחלה של התהליך).
- $Var(X(t)) = \lambda t$, השונות שווה לתוחלת. למעשה $\sqrt{\lambda t}$ זה הגודל ה"סביר" להתרחק מן ישר התוחלת.
  ניתן להבחין כי התהליך לא סטציונרי (תהליך הנגזרת כן – נראה בהמשך) הרי הואריאנס תלוי בזמן.
- תהליך פואסון מקיים את תכונת האדיטיביות – נראה בהמשך.



<u>בניה אלטרנטיבית של תהליך פואסון מסכום ברנולי</u>

נזהה כי נוכל לבנות את תהליך פואסון ע"י סכימת מ"א ברנולי i.i.d כך שנבצע החלטות ממש מהר עם סיכויים נמוכים מאוד אך גודל החלטה מוצלחת יהיה אחד. לצורך הבנייה נראה כי בתהליך פואסון יכול להתרחש לכל היותר אירוע יחיד בזמן אפסי כלשהו.

ניקח ת"א $X(t) \sim Poisson(\lambda \cdot t)$ וננתח את מספר הפיצוצים בזמן קצר. ניזכר:

$$\Pr(X(t)=k) = \frac{(\lambda t)^k \cdot e^{-\lambda t}}{k!}$$

נראה את הסיכויים לאפס/אחד/יותר פיצוצים כאשר $t = \Delta \ll 1$ ונפתח את כאשר הביטוי $O(\Delta^2)$ מסמן זניח של Δ ולמעשה מייצג את המשך הטור (כולל בתוכו את כל הרכיבים אשר מסדר Δ^2 ומעלה):

$$\Pr(X(\Delta)=0) = \frac{(\lambda\Delta)^0 \cdot e^{-\lambda\Delta}}{0!} = e^{-\lambda\Delta} \underset{Series}{\overset{\text{Taylor}}{=}} 1 - \frac{\lambda\Delta}{1} + O(\Delta^2) = 1 - \lambda\Delta + O(\Delta^2)$$

$$\Pr(X(\Delta)=1) = \frac{(\lambda\Delta)^1 \cdot e^{-\lambda\Delta}}{1!} = \lambda\Delta \cdot e^{-\lambda\Delta} = \lambda\Delta \cdot \left(1 - \lambda\Delta + O(\Delta^2)\right) = \lambda\Delta + O(\Delta^2)$$

$$\Pr(X(\Delta)>1) = 1 - \left(1 - \lambda\Delta + O(\Delta^2)\right) - \left(\lambda\Delta + O(\Delta^2)\right) = O(\Delta^2)$$

למעשה קיבלנו כי בזמן סופי יכול להתרחש לכל היותר אירוע יחיד.

נבחין כי מתקיים עבור זמן $t=\Delta \ll 1$:

$$\Pr\left(X\left(\Delta\right)=0\right)\approx 1-\lambda\Delta,\quad \Pr\left(X\left(\Delta\right)=1\right)\approx \lambda\Delta$$

זה פילוג ברנולי עם סיכוי הצלחה $p=\lambda\Delta$.
לכן נוכל לבנות את התהליך ע"י:

$$X^T\left(t\right)=\sum_{n=1}^{\lfloor t/T\rfloor}B_n=\sum_{n=1}^{\infty}B_n\cdot U\left(t-nT\right)\qquad \begin{matrix}B_n\sim Ber\left(\lambda T\right)\\ \left\{B_n\right\} i.i.d\end{matrix}$$

נבצע הגרלה כל זמן T (שקול לכך שכל אינטרוול T נחליט אם התפוצץ בלון או לא בסיכוי $p=\lambda\Delta$).

ניקח את הגבול $T\to 0$ תוך שמירה על היחס $p=\lambda T$ (ולכן גם $p\to 0$).

נראה כי פילוג התהליך $\lim\limits_{T\to 0,p=\lambda T}\left(X^T\left(t\right)\right)$ שואף לפילוג ת"א פואסון:

$$\Pr\left(X^T\left(t\right)=k\right)=\Pr\left(\sum_{n=1}^{\lfloor t/T\rfloor}B_n=k\right)\underset{Bin\left(\lfloor t/T\rfloor,\lambda T\right)}{=}\binom{\lfloor t/T\rfloor}{k}\cdot\left(1-\lambda T\right)^{\lfloor t/T\rfloor-k}\cdot\left(\lambda T\right)^k$$

נסדר את הביטוי לפני ההשאפה כך שחלקים רבים ישאפו לאחד:

$$\Pr\left(X^T\left(t\right)=k\right)=\frac{\lfloor t/T\rfloor\cdot\left(\lfloor t/T\rfloor-1\right)\cdots\left(\lfloor t/T\rfloor-k+1\right)}{k!}\cdot\left(1-\lambda T\right)^{\lfloor t/T\rfloor}\cdot\frac{1}{\left(1-\lambda T\right)^k}\cdot\left(\lambda T\right)^k=$$

$$=\frac{1}{k!}\cdot\left(\lfloor t/T\rfloor\right)^k\cdot\frac{\lfloor t/T\rfloor\cdot\left(\lfloor t/T\rfloor-1\right)\cdots\left(\lfloor t/T\rfloor-k+1\right)}{\left(\lfloor t/T\rfloor\right)^k}\cdot\left(1-\frac{\lambda}{1/T}\right)^{\lfloor t/T\rfloor}\cdot\left(\frac{\lambda T}{1-\lambda T}\right)^k=$$

$$\xrightarrow[T\to 0]{}\frac{1}{k!}\cdot\left(\frac{t}{T}\right)^k\cdot 1\cdot e^{-\lambda t}\cdot\left(\lambda T\right)^k=\frac{1}{k!}\cdot e^{-\lambda t}\cdot\left(\lambda t\right)^k$$

כלומר קיבלנו: $X\left(t\right)=\lim\limits_{T\to 0,p=\lambda T}\left(X^T\left(t\right)\right),\qquad X\left(t\right)\sim Poisson\left(\lambda t\right)$.

הערה - בניית התהליך כסכום משתני מרנולי מזכיר את בניית תהליך וינר. נבחין כי בפיתוח תהליך וינר עם סדרת ברנולי בכל החלטה היה צעד קטן בכיוון כלשהו, בפיתוח תהליך פואסון עם סדרת ברנולי לא בכל החלטה מתרחשת תנועה הצעד הוא גדול קבוע ושווה לאחד אך בסיכוי נמוך. הסכומים שונים מאוד.

<u>אפשרויות סימלוץ של תהליך אקראי פואסון</u>

מאחר שתהליך פואסון הינו תהליך טבעי נרצה לא פעם לבצע סימולציות ולצורך כך נצטרך להגריל תוצאות של תהליך פואסון. יש מספר דרכים לסמלץ תהליך פואסון ואנו נלמד כמה:

1. נגריל סדרת משתנים אקראיים $Z_i \sim \exp(\lambda), \{Z_i\}\, i.i.d$ ומתוכה נבנה את $\{S^i\}$ לפי נוסחת הנסיגה.
   יתרון: לפי הגדרה. חסרון: הגרלת מ"א אקספוננציאלי היא לא משימה פשוטה.
2. נגריל סדרת מ"א ברנולי B_n לפי הבניה האלטרנטיבית ונסכום אותם $X(t) = \sum_n B_n$.
   יתרון: פשוט למימוש. חסרון: לפי קירוב.
3. לסימלוץ ת"א פואסון עם פרמטר λ עד זמן T_{tot} נגריל מ"א פואסוני בעל פילוג $Poisson(\lambda \cdot T_{tot})$ ונסמן את התשובה $N(T_{tot})$.

   אם $N(T_{tot}) = 0$ אז סיים.

   אחרת, הגרל $n = N(T_{tot})$ משתנים $Uniform(0, T_{tot}]$ אשר i.i.d וסמנם $u_1, \ldots, u_n$. סדר אותם בסדר עולה $(o_1 < \ldots < o_n)$ $o_1, \ldots, o_n,$. סדרה $\{o_i\}_{i=1}^{N(T_{tot})}$ היא סדרת אירועי הפואסון באינטרוול $(0, T_{tot}]$.

   יתרון: מדויק ודורש הגרלה אחת "קשה". חסרון: דורש מיון.
   הוכחת נכונות:

- $N(T_{tot}) = 0$, אין אירועים ולכן המצב טריוויאלית.
- $N(T_{tot}) = 1$, קיים אירוע בודד, נראה שמיקומו מתפלג יוניפורמית בקטע $(0, T_{tot}]$.

$$\Pr\left(o_1 \le s \middle| n=1\right) \overset{Bayes}{=} \frac{\Pr(o_1 \le s \cap n=1)}{\Pr(n=1)} \overset{*}{=} \frac{\Pr\left(X(s)=1 \cap X(T_{tot}) - X(s) = 0\right)}{\Pr\left(X(T_{tot})=1\right)} =$$

$$\overset{**}{=} \frac{\Pr\left(X(s)=1\right) \cdot \Pr\left(X(T_{tot}) - X(s) = 0\right)}{\Pr\left(X(T_{tot})=1\right)} = \frac{e^{-\lambda s}(\lambda s)^1}{1!} \cdot \frac{e^{-\lambda(T_{tot}-s)}\left(\lambda(T_{tot}-s)\right)^0}{0!} \cdot$$

$$\cdot \frac{1!}{e^{-\lambda T_{tot}}(\lambda T_{tot})^1} = \frac{s}{T_{tot}} = CDF\left(Unif\left(0, T_{tot}\right]\right)(s)$$

שיוויון $(*)$ מאחר שהמשמעות היא שהיה אירוע אחד באינטרוול $(0,s)$ ואפס אירועים באינטרוול (s,T_{tot}). שיוויון $(**)$ מאחר שבמונה יש איחוד של קטעים זרים והם בת"ס ולכן ניתנים להפרדה.

- $N(T_{tot})\geq 2$, קיימים לפחות שני אירועים, נראה שמיקומם מתפלג $OrderedUniform(0,T_{tot}]$.

$$\Pr\left(t_1=o_1,\dots,t_n=o_n \middle| N(T_{tot})=n\right)\overset{Bayes}{=}\frac{\Pr\left(t_1=o_1,\dots,t_n=o_n\cap N(T_{tot})=n\right)}{\Pr\left(N(T_{tot})=n\right)}=\Upsilon$$

נחליף את המאורע על הסדרה $t_1,t_2,\dots,t_n$ למאורע המסודר על סדרת ההפרשים $x_1=o_1-0, x_2=o_2-o_1,\dots,x_n=o_n-o_{n-1}$.

אנו מחפשים חיתוך עם כך שיש בדיוק n אירועים עד זמן T_{tot}, המשמעות היא שהאירוע "הבא" התרחש אחרי זמן T_{tot}. ולכן מתקיים:

$$\Upsilon=\frac{\Pr\left(x_1=o_1-0,\dots,x_n=o_n-o_{n-1},x_{n+1}>T_{tot}-o_n\right)}{\Pr\left(N(T_{tot})=n\right)}=$$

$$=\frac{\Pr(x_1=o_1)\cdot\Pr(x_2=o_2-o_1)\cdot\ldots\cdot\Pr(o_n-o_{n-1})\cdot\Pr(x_{n+1}>T_{tot}-o_n)}{\Pr\left(X(T_{tot})=n\right)}=$$

$$\underset{x_i\sim\exp(\lambda)}{=}\lambda e^{-\lambda\cdot o_1}\cdot\lambda e^{-\lambda\cdot(o_2-o_1)}\cdot\ldots\cdot\lambda e^{-\lambda\cdot(o_n-o_{n-1})}\cdot\left[1-F_{\exp(\lambda)}(T_{tot}-o_n)\right]\cdot\frac{n!}{e^{-\lambda T_{tot}}\cdot(\lambda T_{tot})^n}=$$

$$=\lambda^n e^{-\lambda\cdot o_n}\cdot\left[1-\left(1-e^{-\lambda(T_{tot}-o_n)}\right)\right]\cdot\frac{n!}{e^{-\lambda T_{tot}}\cdot(\lambda T_{tot})^n}=\frac{n!}{T_{Tot}^n}\overset{*}{=}OrderedUniform(0,T_{tot}]$$

כאשר שוויון $(*)$ מאחר שידוע $u_i\sim Unif(0,T_{tot}],\{u_i\}\ i.i.d$ ולכן:

$$f_u(x)=\begin{cases}1/T_{tot} & x\in(0,T_{tot}]\\ 0 & else\end{cases}$$

ומכאן שמתקיים עבור הפילוג המשותף:

$$f_{u_1,\dots,u_n}(x_1,\dots,x_n)\overset{i.i.d}{=}f_{u_1}(x_1)\cdot\ldots\cdot f_{u_n}(x_n)=\prod_{i=1}^n f_u(x)=\left(\frac{1}{T_{tot}}\right)^n=\frac{1}{T_{tot}^n}$$

אך עלינו לזכור כי עבור קבלת הסדר הממויינת $o_1,\ldots,o_n$ יש למעשה $n!$
אפשרויות בחירה לפני מיון ולכן:

$$f_{u_1,\ldots,u_n}^{Ordered}\left(x_1,\ldots,x_n\right)=\frac{1}{T_{tot}^n}\cdot n!$$

<u>סטטיסטיקה מסדר שני של תהליך פואסון</u>

נציג את הסטטיסטיקה עד סדר שני של תהליך פואסון עם פרמטר λ , ז"א של $X(t)\sim Poisson(\lambda\cdot t)$.

את התוחלת והשונות כבר ראינו:

$$\left.\begin{aligned}E\left[X(t)\right]=\lambda t\\ Var\left(X(t)\right)=\lambda t\end{aligned}\right\}\Rightarrow E\left[X^2(t)\right]=Var\left(X(t)\right)+E^2\left[X(t)\right]=\lambda t\cdot(1+\lambda t)$$

נפתח את פונקציית הקורולציה (נניח $t_1\le t_2$):

$$\begin{aligned}R_{XX}(t_1,t_2)&\triangleq E\left[X(t_1)\cdot X(t_2)\right]=E\left[X(0,t_1)\cdot\left(X(0,t_1)+X(t_1,t_2)\right)\right]=\\ &=E\left[X^2(0,t_1)+X(0,t_1)\cdot X(t_1,t_2)\right]=E\left[X^2(0,t_1)\right]+E\left[X(0,t_1)\cdot X(t_1,t_2)\right]=\\ &\overset{*}{=}E\left[X^2(0,t_1)\right]+E\left[X(0,t_1)\right]\cdot E\left[X(t_1,t_2)\right]=\lambda t_1\cdot(1+\lambda t_1)+\lambda t_1\cdot\lambda(t_2-t_1)=\\ &=\lambda t_1+\lambda^2t_1^2+\lambda^2\cdot\left(t_1t_2-t_1^2\right)=\lambda t_1+\lambda^2t_1^2+\lambda^2\cdot t_1t_2-\lambda^2t_1^2=\lambda t_1+\lambda^2\cdot t_1t_2\end{aligned}$$

נפתח את פונקציית הקו-ואריאנס (נניח $t_1\le t_2$):

$$C_{XX}(t_1,t_2)=R_{XX}(t_1,t_2)-E\left[X(t_1)\right]\cdot E\left[X(t_2)\right]=\lambda t_1+\lambda^2\cdot t_1t_2-\lambda t_1\cdot\lambda t_2=\lambda t_1$$

ובאופן כללי:

$$R_{XX}(t_1,t_2)=\lambda\min(t_1,t_2)+\lambda^2t_1t_2,\qquad C_{XX}(t_1,t_2)=\lambda\min(t_1,t_2)$$

נזהה כי קיבלנו שלתהליך פואסון ולתהליך וינר יש אותה פונקציית קו-ואריאנס (בתהליך וינר התוחלת בכל נקודת זמן שווה לאפס ולכן פונקציית הקורולציה והקו-ואריאנס של תהליך וינר זהות) אך אלו תהליכים שונים לחלוטין הרי תהליך וינר זז כל הזמן אך נשאר סביב אפס ותהליך פואסון זז "מידי פעם" (בזמני האירועים בלבד) אך תמיד עולה.
המשמעות של הדבר היא שאנו לא יכולים להסיק הרבה על תהליך מהסטטיסטיקה של סדר שני (או מומנט יחיד או קבוצה סופית) בלבד.

<u>מיזוג ופיצול של תהליך פואסון</u>

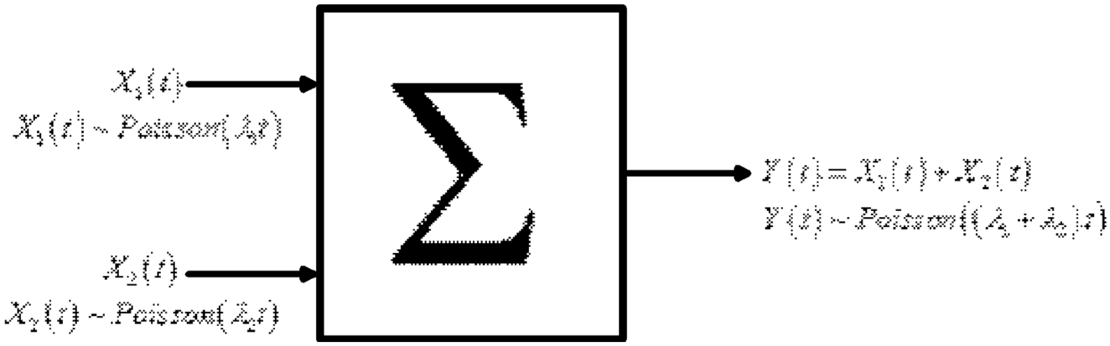
ניזכר בבניה האלטרנטיבית של תהליך פואסון ע"י סדרת ברנולי:

$$X^T(t)=\sum_{n=1}^{\lfloor 1/T \rfloor} B_n=\sum_{n=1}^{\infty} B_n\cdot U(t-nT) \qquad \begin{matrix} B_n\sim Ber(\lambda T) \\ \{B_n\}\ i.i.d \end{matrix}$$

$$X(t)=\lim_{T\to 0,\, p=\lambda T}\left(X^T(t)\right), \qquad X(t)\sim Poisson(\lambda t)$$

<u>מיזוג תהליכי פואסון:</u>

נסתכל על סכימת (מיזוג) שני תהליכי פואסון עם קבועי שונים λ :



נראה כי המוצא הוא גם תהליך פואסון אשר מתפלג עם קבוע $\lambda_1+\lambda_2$.

$$Y^T(t)=X_1^T(t)+X_2^T(t)=\sum_{n=1}^{\infty} B_n^1\cdot U(t-nT)+\sum_{n=1}^{\infty} B_n^2\cdot U(t-nT)=$$

$$=\sum_{n=1}^{\infty}\left(B_n^1+B_n^2\right)\cdot U(t-nT)=\sum_{n=1}^{\infty} B_n\cdot U(t-nT)$$

$$B_n=B_n^1+B_n^2=\begin{cases} 0 & (1-\lambda_1 T)\cdot(1-\lambda_2 T) \\ 1 & \lambda_1 T\cdot(1-\lambda_2 T)+\lambda_2 T\cdot(1-\lambda_1 T) \\ 2 & \lambda_1 T\cdot\lambda_2 T \end{cases}$$

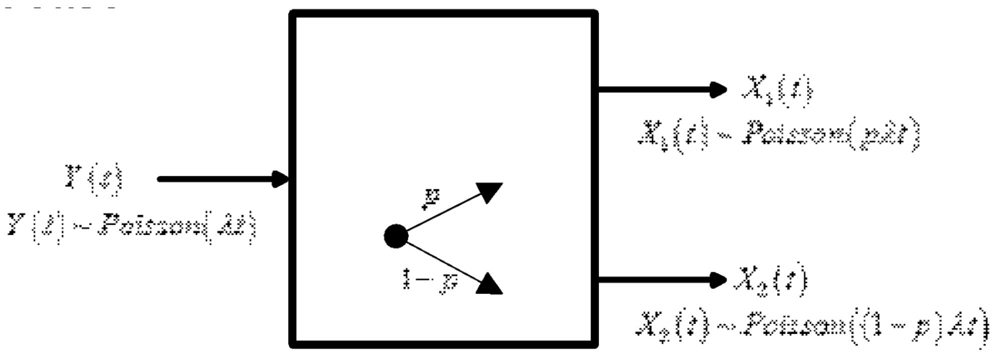
כאשר מרווח הזמן T הוא מאוד קטן ($T\ll 1$) אז נוכל לבצע קירוב סדר ראשון:

$$B_n=\begin{cases} 0 & (1-\lambda_1 T)\cdot(1-\lambda_2 T) \\ 1 & \lambda_1 T\cdot(1-\lambda_2 T)+\lambda_2 T\cdot(1-\lambda_1 T) \\ 2 & \lambda_1 T\cdot\lambda_2 T \end{cases}=\begin{cases} 0 & 1-(\lambda_1+\lambda_2)T+O(T^2) \\ 1 & (\lambda_1+\lambda_2)T+O(T^2) \\ 2 & 0+O(T^2) \end{cases}\approx\begin{cases} 0 & 1-(\lambda_1+\lambda_2)T \\ 1 & (\lambda_1+\lambda_2)T \end{cases}$$

וקיבלנו שהמקדם B_n הוא למעשה מ"א ברנולי עם סיכוי הצלחה $p=(\lambda_1+\lambda_2)T$.

<u>פיצול תהליך פואסון:</u>

נסתכל על המערכת הבאה:



המערכת מוזנת ע"י תהליך פואסון עם מקדם λ ומחליטה לנווט אותו לאחד משני מוצאים לפי Decider.

נגדיר מ"א לייצור ההחלטות:

$$I_n = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

ולכן נוכל להציג את המוצאים באופן הבא (רק אחד מהם מקבל את הכניסה בהסתברות המתאימה):

$$X_1^T(t) = \sum_{n=1}^{\infty} I_n \cdot B_n \cdot U(t-nT), \quad X_2^T(t) = \sum_{n=1}^{\infty} I_n \cdot B_n \cdot U(t-nT)$$

ולכן נוכל לזהות:

$$B_n^1 = I_n \cdot B_n = \begin{cases} 1 & p \cdot \lambda T \\ 0 & (1-p) \cdot \lambda T \end{cases} \qquad B_n^2 = I_n \cdot B_n = \begin{cases} 1 & (1-p) \cdot \lambda T \\ 0 & p \cdot \lambda T \end{cases}$$

ואכן קיבלנו כי שניהם משתנים ברנולים ולכן

$$X_1(t) \sim Poisson(p\lambda t), \quad X_2(t) \sim Poisson((1-p)\lambda t).$$

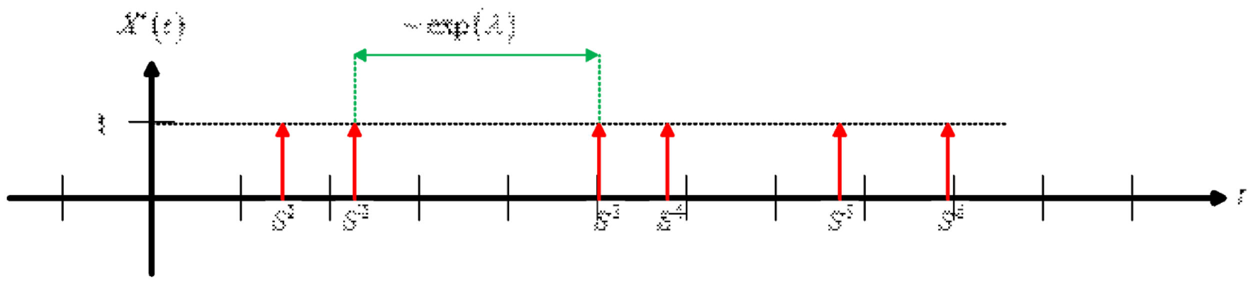
תהליך הנגזרת של תהליך פואסון

כמו שביצענו עבור תהליך וינר נגדיר תחילה את תהליך השיפוע:

$$X_\varepsilon(t)=\frac{X(t+\varepsilon)-X(t)}{\varepsilon}=\frac{X(t,t+\varepsilon)}{\varepsilon}$$

ולכן תהליך הנגזרת הוא $X'(t)=\lim_{\varepsilon\to 0}X_\varepsilon(t)$.

פונקציית מדגם של תהליך הנגזרת ייראה באופן הבא (הלמים בגובה אחד כגובה המדרגות בתהליך פואסון בזמני הקפיצות):



נרצה לבחון את הסטטיסטיקה מסדר שני של תהליך הנגזרת.

לתהליך פואסון יש פונקציית קו-ואריאנס זהה (נחליף α ב- λ) לזו של תהליך וינר ולכן מתקיים כי גם לתהליכי הנגזרת שלהם הדבר תקף:

$$C_{X'X'}(\tau)=\lambda\cdot\delta(\tau)$$

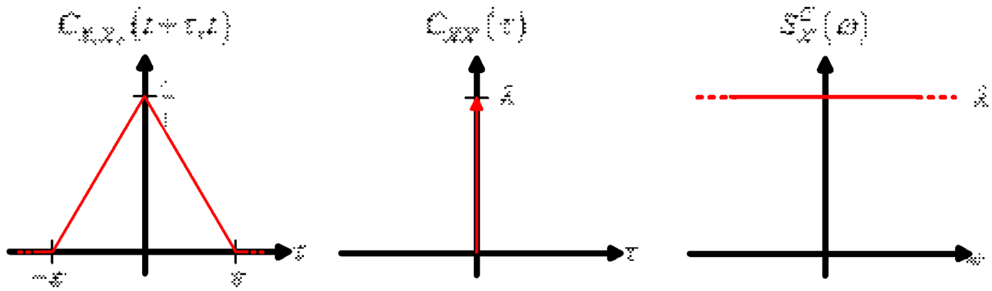
נרצה גם לפתח את פונקציית הקורולציה של תהליך הנגזרת של פואסון (ולשם כך נצטרך את תוחלת תהליך הנגזרת). נראה כי מתקיים שתוחלת תהליך השיפוע לא תלוי ב- ε ולכן היא שווה לתוחלת תהליך הנגזרת.

$$E\left[X_\varepsilon(t)\right]=E\left[\frac{X(t,t+\varepsilon)}{\varepsilon}\right]=\frac{1}{\varepsilon}\cdot E\left[X(t,t+\varepsilon)\right]=\frac{1}{\varepsilon}\cdot\lambda(t+\varepsilon-t)=\lambda$$
$$\Rightarrow\quad E\left[X'(t)\right]=\lambda$$

ולכן:

$$R_{X'X'}(\tau)=R_{X'X'}(\tau,0)=C_{X'X'}(\tau)+E\left[X'(\tau)\right]\cdot E\left[X'(0)\right]=\lambda\cdot\delta(\tau)+\lambda^2$$

נבחין כי מתקיים:



ז"א לתהליך הנגזרת של תהליך פואסון יש ספקטרום הפסק שטוח על כל התדרים, משמע שתהליך הנגזרת של תהליך פואסון הוא רעש לבן.
זהו רעש לבן ממש שונה מרעש גאוסי לבן מאחר שהוא לא קיים כל הזמן אלא רק בזמני אירועי הפואסון (בניגוד לרעש גאוסי לבן אשר תמיד קיים). זהו גם לא רעש לבן טהור מאחר שהתוחלת שלו שונה מאפס.
רעש לבן כזה מתאים למקרים של רמת חלקיק בודד כמו פגיעת לייזר בעובי פוטון בודד או זרימת אלקטרון בחתך בעל שטח פנים בסדר גודל של שטח הפנים של האלקטרון עצמו.

<u>תהליך פואסון כ-Renewal Process</u>

קיימת מחלקה של תהליכים מסוג Renewal Processes. ת"א פואסון שייך למחלקה זו והוא מקרה פרטי שלה.
הסתכלנו על ת"א פואסון כאוסף אינסופי של בלונים עם פיצוצים (התרחשות אירוע) במקביל. נרצה להסתכל על ת"א מנייתי אך בטור, כלומר יש בלון אחד וכאשר הוא מתפוצץ מחליפים אותו באחר.
לצורך הדיון נדון בתהליך של החלפת נורה (כאשר נורה נשרפת היא מוחלפת בחדשה מן האריזה). כמובן שלזמן חיי המנורה יש פילוג כלשהו, נניח $f_Z(z)$. לכן נגדיר את התהליך ע"י:

$$\{Z_i\}\ i.i.d \qquad S^i = S^{i-1} + Z_i \qquad X(t) = \sum_{i=1}^{\infty} U\left(t - S^i\right)$$

כאשר התהליך $X(t)$ יציג בכל זמן את מספר הנורות אשר הוחלפו עד אותו זמן.
תהליכי Renewal מובילים לתורת התורים (לדוגמא נציג בודד אשר מטפל במספר לקוחות באופן טורי כאשר לזמן הטיפול יש פילוג כלשהו).

אופי התהליך תלוי בפילוג הסדרה $\{Z_i\}$, באופן כללי $f_Z(z)$ הוא לא בהכרח פילוג אקספוננציאלי. במקרה בו מתקיים $Z_i \sim \exp(\lambda)$ אז מתקבל ת"א פואסון $X(t) \sim Poisson(\lambda t)$.
במקרה הזה גם נקבל כי התהליך יהיה בעל תוספות בת"ס (הפילוג האקספוננציאלי בעל תכונת חוסר זיכרון והדבר לא נכון עבור כל פילוג, בנוסף הוכחנו כי עבור פילוג תוספות

אקספוננציאלי מתקיימת נוסחת A.R. ליניארית אך אין הדבר אומר שאם מתקיים נוסחת A.R. ליניארית התוספות בהכרח היו מפולגות אקספוננציאלית).

<u>פרדוקס הנורה:</u>

פרדוקס הנורה (מוכר גם כפרדוקס הטרמפיסט/האוטובוס) מראה סוגייה מעניינת לגביי תוחלת של מ"א אקספוננציאלי.

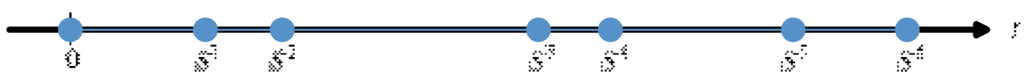
אנו מחילים Renewal Process של החלפת נורה בטור. נניח שזמן חיים של כל נורה מתפלג אקספוננציאלית עם $\frac{1}{\lambda}=1month$. לכן מרגע שהיא חוברה תוחלת החיים שלה היא חודש אחד.

אנו נכנסים לחדר ומצביעים על נורה שעובדת, נרצה לדעת מה תוחלת החיים שלה (מרגע שהיא הוחלפה).

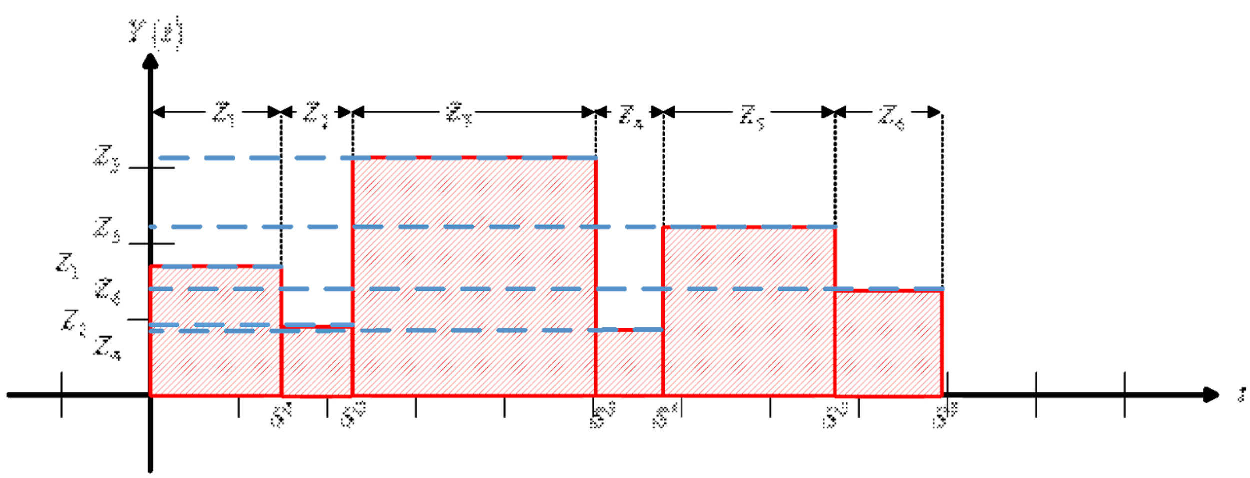
תוחלת החיים של מנורה כזו הוא שני חודשים.

הסברים:

- הפילוג האקספוננציאלי חסר זיכרון ולכן כך גם התהליך. לכן גם הזמן אחורה וגם הזמן קדימה מתפלג $\exp(\lambda)$. מכאן שיש לנו תוחלת של חודש לכל אחד מן כיווני הזמן – סה"כ חודשיים.
- מדובר בתהליך פואסון. אנחנו יודעים כי נקודות האירועים מתפלגות באופן אחיד על הציר (ראינו באפשרות סימלוץ 3). לכן כאשר ניכנס לחדר הדבר שקול לנקודה אקראית על הציר ו"סביר" יותר שניפול על קטע ארוך מאשר קצר (כי יש להם הסתברות מותנית גבוהה יותר כי אורכם ארוך יותר):



- נדגיר ת"א אשר מראה את משך החים של כל מנורה (בכל רגע t יציג את זמן החיים של מנורה נוכחית):



תהליך זה הינו תהליך ריבועים. נניח והתהליך ארוך עבור מספיק זמן והמון נורות.

נבוא ברגע אקראי T מסויים. נרצה לדעת מה ממוצע זמן החיים של המנורות עד זמן זה:

$$\overline{Y}(T)=\frac{1}{T}\cdot\int_0^T T(t)\cdot dt$$

כמובן שהממוצע הזה חסום משני הצדדים ע"י הממוצע בתחילת אותו ריבוע ובסוף אותו ריבוע.

$$\frac{1}{T}\cdot\sum_{i=1}^{X(T)} Z_i^2 \le \overline{Y}(T) \le \frac{1}{T+1}\cdot\sum_{i=1}^{X(T)+1} Z_i^2 \underset{T+1>T}{\le} \frac{1}{T}\cdot\sum_{i=1}^{X(T)+1} Z_i^2$$

$$\frac{1}{T}\cdot\frac{X(T)}{X(T)}\cdot\sum_{i=1}^{X(T)} Z_i^2 \le \overline{Y}(T) \le \frac{1}{T}\cdot\frac{X(T)}{X(T)}\cdot\sum_{i=1}^{X(T)+1} Z_i^2$$

נסתכל על חלק מהביטוי:

$$\frac{1}{T}\cdot\frac{X(T)}{X(T)}\cdot\sum_{i=1}^{X(T)} Z_i^2 = \left(\frac{1}{X(T)}\cdot\sum_{i=1}^{X(T)} Z_i^2\right)\cdot\left(\frac{X(T)}{T}\right) \overset{LLN}{=} E\left[Z_i^2\right]\cdot\left(\frac{X(T)}{T}\right) = E\left[Z_i^2\right]\cdot\frac{1}{E\left[Z_i\right]}$$

עד כה הפיתוח היה כללי עבור כל Renewal Process, נציב פנימה כי $Z_i \sim \exp(\lambda)$ ונקבל:

$$\frac{1}{T}\cdot\frac{X(T)}{X(T)}\cdot\sum_{i=1}^{X(T)} Z_i^2 = E\left[Z_i^2\right]\cdot\frac{1}{E\left[Z_i\right]} = \frac{1/\lambda^2 + 1/\lambda^2}{1/\lambda} = \frac{2}{\lambda} = 2\cdot E\left[Z_i\right]$$

הוכחנו כי התוחלת כפולה.