# הרצאה 10

תזכורת מהרצאה 9: (לא היה השנה)

פונקציה μ היא $\sigma - \text{additive}$, (סיגמא-אדטיבית), אם היא מקיימת: $\mu(\cup_{n=1}^{\infty} A_n) = \Sigma_{n=1}^{\infty}\mu(A_n)$
(כלומר, ערך הפונקציה על איחוד בן-מניה שווה לסכום הערכים על כל איברי האיחוד).

סיגמא-אלגברה A על קבוצה X היא משפחה של תתי קבוצות של X שמקיימת:

1) $\emptyset \subseteq A$
2) אם $E \in A$ אז גם $E^C = X\backslash E \in A$ (המשפחה סגורה ללקיחת משלים).

מרחב הסתברות הוא שלישיה: $(\Omega, \mathcal{F}, \mathbb{P})$ כאשר:

1) Ω היא קבוצה לא-ריקה. נקראת **מרחב המדגם.**
2) $\mathcal{F} \subseteq 2^{\Omega}$ היא $\sigma - \text{algebra}$, (סיגמא-אלגברה).
3) $\mathbb{P} : \mathcal{F} \to [0,1]$ היא פונקציית הסתברות סיגמא-אדטיבית שמקיימת $\mathbb{P}(\emptyset) = 0$ ו- $\mathbb{P}(\Omega) = 1$

כאשר Ω סופית או בת מניה, אפשר לקחת $\mathcal{F} = 2^{\Omega}$ ואז אפשר לרשום $(\Omega, \mathbb{P})$.

## 1 משתנים מקריים רציפים:

יהי $X: \mathbb{R} \to \mathbb{R}$ משתנה מקרי במרחב הסתברות $(\mathbb{R}, \mathcal{F}, \mathbb{P})$,
($\mathbb{R}$ היא ה- Ω, ו- $2^{\Omega}$ זה המאורעות (תת הקבוצות של מרחב המדגם), $\mathcal{F} \subseteq 2^{\Omega}$. $\mathcal{F}$ זה המאורעות שעבורם אנחנו יודעים להגדיר את $\mathbb{P}$.)
כאשר $\mathcal{F}$ היא סיגמא-אלגברה שכוללת את כל המקטעים, כלומר $(a,b), (a,b], [a,b] \in \mathcal{F}$ וכו'. אנחנו רוצים ש$\mathcal{F}$ תהיה סיגמא-אלגברה כדי שאם נדע את ההסתברות של מאורע, נוכל לדעת את ההסתברות של המשלים שלו, ואם נדע את ההסתברות של קבוצת מאורעות נוכל לדעת את ההסתברות של האיחוד והחיתוך.
נאמר ש X **רציף** (או רציף בהחלט) אם יש דרך מיוחדת לחשב את ההסתברות של המאורעות שתלויים ב X.
פורמלית: אם קיימת פונקציה $f_X : \mathbb{R} \to \mathbb{R}$ כך ש (1):

$$\mathbb{P}(X \in B) = \int_B f_X(x)dx$$ לכל $B \in \mathcal{F}$.

בפרט, צריך שיתקיים: $\mathbb{P}(\mathbb{R}) = \int_{-\infty}^{\infty} f_X(x)dx = 1$.

הפונקציה $f_X$ נקראת **פונקציית הצפיפות** (PDF – Probability Density Function) של X.

מ- (1) נובע ש: $\mathbb{P}(a \le X \le b) = \int_a^b f_X(x)dx$ לכל שני $a \le b$ ממשיים.

בפרט, $\mathbb{P}(X = c) = \int_c^c f_X(x)dx = 0$ לכל $c \in \mathbb{R}$. ולכן: לכל $a \le b$ ממשיים:

$$\mathbb{P}(a \le X \le b) = \mathbb{P}(a < X \le b) = \mathbb{P}(a \le X < b) = \mathbb{P}(a < X < b) = \int_a^b f_X(x)dx$$

לכל $a \in \mathbb{R}$ נגדיר $F_X(a) = \mathbb{P}(X \le a) = \int_{-\infty}^{a} f_X(x)dx$. בהמשך נראה למה זה ככה.

הפונקציה $F_X$ נקראת **פונקציית ההתפלגות המצטברת** (CDF – Cumulative Distribution Function) של X.

זה נותן לנו גם ש: $\mathbb{P}(a \le X \le b) = \mathbb{P}(X \le b) - \mathbb{P}(X < a) = \mathbb{P}(X \le b) - \mathbb{P}(X \le a) = F_X(b) - F_X(a)$.

**דוגמה 1:** יהי X משתנה מקרי רציף, ותהי פונקציית הצפיפות שלו: (ה-PDF)

$$f_X(x) = \begin{cases} c(1-x^2), & if -1 < x < 1 \\ 0, & otherwise \end{cases}$$ עבור קבוע c.

כדי למצוא את הערך של c נשתמש במשוואה $\int_{-\infty}^{\infty} f_X(x)dx = 1$. מתקיים:

$$1=\int_{-\infty}^{\infty}f_X(x)dx=\int_{-\infty}^{-1}0dx+\int_{-1}^{1}c(1-x^2)dx+\int_{1}^{\infty}0dx=\int_{-1}^{1}c(1-x^2)dx=c\left[x-\frac{x^3}{3}\right]|_{x=-1}^{x=1}=$$
$$=c\left[\left(1-\frac{1}{3}\right)-\left(-1+\frac{1}{3}\right)\right]=\frac{4c}{3}$$

כלומר $c = 3/4$. ועכשיו נוכל לחשב הסתברויות עם X, לדוגמה:

$$\mathbb{P}(X\le 0)=\int_{-\infty}^{0}f_X(x)dx=\int_{-1}^{0}\frac{3(1-x^2)}{4}dx=\left[\frac{3x}{4}-\frac{x^3}{4}\right]|_{x=-1}^{x=0}=0-\left(-\frac{3}{4}+\frac{1}{4}\right)=\frac{1}{2}$$

**דוגמה 2:** יהי X משתנה מקרי רציף עם פונקציית צפיפות $f_X$ ופונקציית התפלגות מצטברת $F_X$.
נרצה למצוא את ה- PDF ו- CDF של המשתנה המקרי $Y = 2X$. לכל $a \in \mathbb{R}$ מתקיים:

$$F_Y(a)=\mathbb{P}(Y\le a)=\mathbb{P}(2X\le a)=\mathbb{P}\left(X\le\frac{a}{2}\right)=F_X\left(\frac{a}{2}\right)$$

נגזור את $F_Y$ לפי $a$ ונקבל: $f_Y(a)=\frac{1}{2}f_X\left(\frac{a}{2}\right)$

## 1.1 תוחלת ושונות של משתנה מקרי רציף:

ניזכר שאם X הוא בדיד, התוחלת היא סכום ההסתברויות על התומך: $\mathbb{E}(X)=\sum_x x\mathbb{P}(X=x)$.

באופן דומה, במ"מ רציף, התוחלת היא: $\mathbb{E}(X)=\int_{-\infty}^{\infty}xf_X(x)dx$.

**דוגמה 3:** יהי X מ"מ רציף עם PDF המוגדרת: $f_X(x)=\begin{cases}2x, & if\ 0\le x\le 1\\ 0, & otherwise\end{cases}$

נשים לב שזה באמת PDF, כי: $\int_{-\infty}^{\infty}f_X(x)dx=\int_0^1 2xdx=x^2|_{x=0}^{x=1}=1$, אז:

$$\mathbb{E}(X)=\int_{-\infty}^{\infty}xf_X(x)dx=\int_0^1 2x^2dx=2\frac{x^3}{3}|_{x=0}^{x=1}=\frac{2}{3}$$

באופן דומה למקרה של מ"מ בדיד, אפשר להשתמש בPDF כי לחשב את התוחלת של פונקציה כלשהי של X.

**טענה 1.1:** יהי X מ"מ רציף עם PDF $f_X$. אזי, לכל פונקציה $g:\mathbb{R}\to\mathbb{R}$ מתקיים:

$$\mathbb{E}\big(g(X)\big)=\int_{-\infty}^{\infty}g(x)f_X(x)dx$$

כלומר נוכל להשתמש בפונקציה של $x$, וההבדל היחיד הוא שכופלים ב- $g(x)$.

**למה 1.2:** הוכחה עבור פונקציות אי-שליליות:

יהי Y מ"מ רציף אי"ש עם PDF $f_Y$. אזי:

$$\mathbb{E}(Y)=\int_0^{\infty}\mathbb{P}(Y>y)dy$$

הוכחה: מכיוון שמתקיים $\mathbb{P}(Y>y)=\int_y^{\infty}f_Y(x)dx$ לכל $y\in\mathbb{R}$, נובע ש:

$$\int_0^{\infty}\mathbb{P}(Y>y)dy=^{\text{א}}\int_0^{\infty}\left(\int_y^{\infty}f_Y(x)dx\right)dy=^{\text{ב}}\int_0^{\infty}\left(\int_0^{x}dy\right)f_Y(x)dx=^{\text{ג}}\int_0^{\infty}xf_Y(x)dx=^{\text{ד}}\mathbb{E}(Y)$$

א – נציב את הנוסחה.

ב – החלפת סדר אינטגרציה: בשלב הראשון, $x \in [y,\infty); y \in [0,\infty)$. כלומר תמיד $0 \le y \le x$. האינטגרל הפנימי הוא לפי $x$ והחיצוני הוא לפי $y$. אנחנו רוצים להחליף את הסדר, כדי שהאינטגרל החיצוני יהיה לפי $x$ והפנימי לפי $y$. במקום לסכום לפי $x$ בתחום $[y,\infty)$, (כאשר $y$ בתחום $[0,\infty)$) נסכום לפי $y$ בתחום $[0,x)$. (כאשר $x$ בתחום $[0,\infty)$).
ואם קבענו בחוץ את $x$, אז $f_Y(x)$ הוא קבוע בזמן שעושים את האינטגרל לפי $y$. אז נוציא אותו החוצה.

ג – האינטגרל שהוא רק $dy$ זה בעצם האינטגרל של הפונקציה 1, שזה פשוט $y$. ואז בחישוב מציבים $x - 0$.

ד – לפי הגדרה, ובגלל ש-$Y$ אי שלילי אז אפשר להתעלם מהתחום השלילי.

הוכחת טענה 1.1:

א – בהנחה ש $g$ אי"ש, מלמה 1.2 נובע:

$$\mathbb{E}\big(g(X)\big) =^{\text{א}} \int_0^\infty \mathbb{P}(g(X) > y)\, dy =^{\text{ב}} \int_0^\infty \int_{x:g(x)>y} f_X(x)\, dx\, dy =^{\text{ג}} \int_{x:g(x)>0} \left( \int_0^{g(x)} dy \right) f_X(x) dx =^{\text{ד}}$$
$$= \int_{x:g(x)>0} g(x) f_X(x)\, dx =^{\text{ה}} \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

ב – לפי (1). האינטגרל על הקבוצה המתאימה.

ג – החלפת סדר אינטגרציה. החיצוני הוא $y$, בתחום $[0,\infty)$. הפנימי $x$ הוא כך ש $y < g(x)$.

ד – האינטגרל הפנימי שווה $g(x)$, ו- $f_X(x)$ הוא קבוע מבחינתו.

ה – כי $g$ אי"ש ואם היא שווה 0, הכל יוצא אפס. אז בכל מקרה נסכום רק על הקבוצה שבה $g(x) > 0$.

**טענה 1.3:** לינאריות התוחלת

יהיו $a, b \in \mathbb{R}$ ויהי X מ"מ רציף. אז מתקיים: $\mathbb{E}(aX + b) = a \cdot \mathbb{E}(X) + b$

הוכחה: תהי $f_X$ ה-PDF של $X$. אז מטענה 1.1 נובע:

$$\mathbb{E}(aX + b) = \int_{-\infty}^{\infty} (ax + b) f_X(x)\, dx = a \underbrace{\int_{-\infty}^{\infty} x f_X(x)\, dx}_{\mathbb{E}(X)} + b \underbrace{\int_{-\infty}^{\infty} f_X(x)\, dx}_{=1} = a\mathbb{E}(X) + b$$

כמו במקרה של מ"מ בדיד, **השונות** של X מוגדרת: $Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$,
וחישוב ישיר נותן: $Var(X) = \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2$

**דוגמה 4:** יהי X מ"מ רציף עם PDF המוגדרת: $f_X(x) = \begin{cases} 2x, & if\ 0 \le x \le 1 \\ 0, & otherwise \end{cases}$

ראינו כבר שזו אכן PDF, ושמתקיים: $\mathbb{E}(X) = 2/3$. לפי טענה 1.1 נקבל:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^1 2x^3\, dx = \frac{x^4}{2}\Big|_{x=0}^{x=1} = \frac{1}{2}$$

ולכן: $Var(X) = \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$

**טענה 1.4:** יהיו $a, b \in \mathbb{R}$ ויהי X מ"מ רציף, אז: $Var(aX + b) = a^2 Var(X)$.

הזזה בסקלר לא משפיעה, כפל בסקלר יוצא בריבוע.

ההוכחה זהה להוכחה מהסתברות 1 על מ"מ בדיד.