שיעור מס' 1 - הקדמה:

א) סילבוס.

ב) בכלכלה מנסים להסביר "תופעות רגילות" ← אינפלציה, צמיחה
באמצעות **מודלים**

(סטיבן לוויט) ← למה סוחרי סמים גרים עם האמהות שלהם

ג) מודלים מדגימים של קשרים סיבתיים:

* שער ריבית במשק משפיע על היעילות ההשקעה
* העלאת תקציב הרווחה מעלה את רמת המרירות
* הצלחת נבחרת משפיעה על מחירי מניות
  ← תחושה רעה משפיעה יותר מתחושה טובה
* תמריצים: שווי האופציות הניתן למנהלים משפיע על ביצועי החברה
* השכלה משפיעה על שכר של הפרט

ד) חלק מתאוריה ומודל צריך להיות לוגים ומתמטיים שמאפשרים להסיק מסקנות (לא בהכרח אינטואיטיביות)

ה) יש גם מודלים שהנחות לא מתקיימות המציאות (כמו רציונליות) אבל המודל **עובד**. מסתכלים **כאילו** מתנהגים כך.
BR

ו) בסופו של דבר יש המון תאוריות. רובן לא נכונות.

ז) אקונומטריקה: **ארגז כלים** שמאפשר להעביר תאוריות רלוונטיות לבחינה במציאות.

ח) אתגרים לגבי נתונים. ← הנתונים של הכלכלה:

קושי בניבוי התנהגות → מורכבות התופעות שחוקרים ← יותר מורכבות מתופעות הפיזיקה.

ט) ניסויי מעבדה קשה יותר לחקור

י) יש גם הרבה הבדלים רגשיים.

II

← האקונומטריקה שמה למטרה לשער את הפונ'ות הללו.

לנסח מודלים שמאפשרים

① לבחון את המסקנות של המודלים הכלכליים. → $C = f(r, y)$

consumption תצרוכת

② לתפוס לרמה מספרית גבולים שהתיאוריה הם רק משתנה (ידועה רק כיוון)

f לינארית / קוואדרטית

③ לספק תחזיות ← מה יקרה אם הממשלה תעלה את האינפלציה

←

↓

האם המודל נכון?

④ לאפשר להבין מגמות.

שיעור מס' 1: חזרה על מושגי יסוד בסטטיסטיקה

(1)

חזרה על

אופרטור סכום: $\Sigma \leftarrow$ סמל שאומר לנו שאנחנו סוכמים את האיברים בסדרה: נדרש לחישובים

$$\sum_{i=1}^{n} x_i = x_1 + \dots + x_n$$

X- משתנה [illegible]

[illegible]

תכונות האופרטור:

(1) $\sum_{i=1}^{n} c \,, (c = \text{conts}) = n \cdot c$

(2) $\sum c \cdot x_i = c \cdot x_1 + \dots + c x_n = c \cdot \sum x_i$

(3) $\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{x_1 + x_2 + \dots + x_n}{N}$

(4) $\sum x_i = n \bar{x}$ [illegible]

(5) $\sum (x_i + y_i) = (x_1 + y_1) + \dots + (x_N + y_N) = \sum x_i + \sum y_i$

נזכיר: בדוגמא:

(6) $\sum (x_i - \bar{x}) = 0 \quad = \sum x_i - \sum \bar{x} = N\bar{x} - N\bar{x} = 0$

* $\sum (x_i - \bar{x})^2 \neq 0$
* $\sum (x_i - \bar{x}) y_i \neq 0$

שים לב

לדוגמא:

$y_1 = 1 \,, y_2 = 3$

$x_1 = 10 \quad x_2 = 20$

(7) $\sum (x_i - \bar{x}) \cdot c = 0$

$= \sum x_i \cdot c - \sum \bar{x} \cdot c = cN\bar{x} - cN\bar{x} = 0$

(8) $\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum (x_i - \bar{x}) y_i$

$\rightarrow \sum (x_i - \bar{x}) y_i - \sum (x_i - \bar{x}) \bar{y} = \_\_ - \bar{y} \cancel{\sum} x_i + \bar{y} \cancel{\sum} \bar{x}$

(9) $$\sum(ax_i+by_i) = a\sum x_i + b\sum y_i$$ קבועים a,b

(10) $$\sum_{i=1}^{N}(ax_i+by_i)^2 = a^2x_1^2+b^2y_1^2+2abx_1y_1$$
$$+a^2x_2^2+b^2y_2^2+2abx_2y_2$$
$$\ldots\ldots$$
$$= a^2\sum x_i^2 + b^2\sum y_i^2 + 2ab\sum_{i=1}^{N}x_iy_i$$

$$\boxed{\sum x_i^2 \neq \left(\sum x_i\right)^2}$$

לא קבוע

(11) $$\left(\sum a_ix_i\right)^2 = (a_1x_1+a_2x_2+\ldots+a_nx_n)^2$$
$$= a_1^2x_1^2+a_2^2x_2^2+\ldots+a_n^2x_n^2+2a_1a_2x_1x_2+\ldots 2a_{n-1}a_nx_{n-1}x_n$$
$$= \sum a_i^2x_i^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}a_ia_jx_ix_j$$

$2a_1a_2x_1x_2$ $+2a_1a_3x_1x_3$

הגדרות פורמליות בסטטיסטיקה

אוכלוסיה: אוסף כל הפרטים/יחידות שעליהם מוגדרת התכונה שמעניינת את המשתנה.

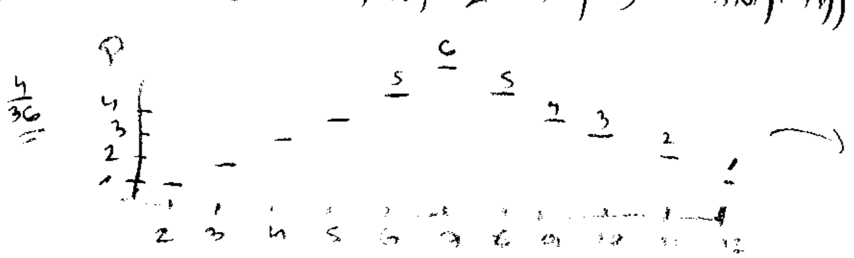
דוגמא: תוצר לאומי: האוכלוסיה היא כל התוצרים של המשק הישראלי מהקמתו ועד היום.

מדגם: חלק מהאוכלוסיה.

ולא ניתן לחזות אותו בודאות מוחלטת

(**משתנה מקרי:** משתנה שיכול לקבל ערכים שונים בהסתברויות כלשהן. → יכול להיות רציף או בדיד

נניח קוביה: זורקים 2 קוביות. הסכום הוא משתנה מקרי.



התפלגות

משתנים בלתי תלויים: y,x ← y לא נותן לנו שום אינפורמציה לגבי ההסתברות של x (זריקת קובייה)

במקרה הבדיד → $P(x\cdot y) = P(x)\cdot P(y)$; $P(x|y) = P(x)$

וגם

דוגמא למשתנים תלויים: שפה ומקום מגורים [illegible]

המקרה הרציף:

פונקציית צפיפות - ההסתברות לקבל ערך מסוים

$$\int_{-\infty}^{\infty} f(x)dx = 1 \qquad 0 \le f(x) \le 1$$ → התפלגות הכוללת

בנוסף:

$$P_r(4 < x \le 8) = \int_4^8 f(x)dx$$

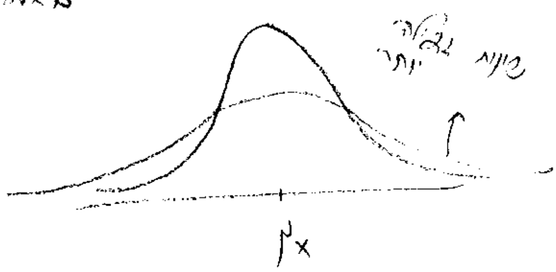
נאמר ש x מתפלג נורמלית אם פונק' הצפיפות שלו היא:

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma_x}\right)$$

מה יפה בה? רציפה, בצורת פעמון, סימטרית וניתנת לתיאור מלא באמצעות התוחלת והשונות שלה.

$$x \sim N(\mu_x, \sigma_x^2)$$

← [illegible]



תכונה שימושית: אם a,b קבועים

(1) $y = a + bx \Rightarrow y \sim N(a + b\mu_x, b^2\sigma_x^2)$

(2) אם $x \sim N(\mu_x, \sigma_x^2)$; $y \sim N(\mu_y, \sigma_y^2)$ אזי $ax + by \sim N(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy})$

התפלגות z ← התפלגות נורמלית סטנדרטית

$$z = \frac{X - \mu_x}{\sigma_x} \sim N(0,1)$$

<u>אופרטור תוחלת:</u> (זה <u>לא</u> הממוצע במדגם) ← ממוצע משוקלל של כל התוצאות האפשריות:

$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$$ ← מקרה רציף

$$E(x) = \sum_i p_i x_i$$ ← מקרה בדיד

$$E\left(\frac{x}{y}\right) \neq \frac{E(x)}{E(y)}$$ אם x ו-y תלויים.

① $E(a) = a$ — נניח a משתנה לא מקרי. (אני יודע מראש מהו בדיוק)

x_i משתנה מקרי

② $E(a \cdot x_i) = aE(x_i) = a\mu_x$

③ $E(x_1 + x_2) = E(x_1) + E(x_2) = 2\mu_x$

$$\Rightarrow \boxed{E\left(\sum a_i x_i\right) = E(a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n) = a_1E(x_1) + \dots + a_nE(x_n) = \sum a_i E(x_i)}$$

④ $E(XY) = E(X)\cdot E(Y) = \mu_x \mu_y \Rightarrow \sigma_{xy} = 0$ כאשר x, y ב"ת.

<u>שונות:</u> מדד הפיזור סביב התוחלת;

$$Var(x) = \sigma_x^2 = E(X - \underbrace{E(x)}_{\mu_x})^2$$

<u>שונות משותפת:</u>

$$\sigma_{xy} = E[(x-\mu_x)(y-\mu_y)]$$

① $Var(c) = 0$

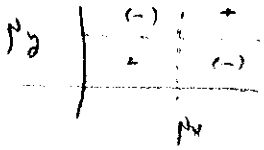
② $Var(a + bX) = b^2 Var(X)$

③ $Var(ax \pm by) = a^2V(x) + b^2V(y) \pm 2ab\,cov(x,y)$

④ $Var(x+y+z) = Var\,x + Var\,y + Var\,z + 2cov(x,y) + 2C(x,z) + 2C(y,z)$

<u>סטיית תקן:</u> השורש של השונות

$$\sqrt{Var(x)} = \sigma_x$$



תכונות של שונות משותפת:

(1) $cov(x_i, c) = 0 \rightarrow E[(x_i - \mu_x)(\underbrace{c - E(c)}_{0})] = 0$

(2) $Cov(ax_i, by_i) = ab\,Cov(x,y) = Cov(c+ax, d+by)$

(הבא) עמוד

הגדרת מקדם המתאם: $\rho_{xy} = \frac{Cov(x,y)}{\sqrt{V(x)V(y)}}$ ← אומדן מוטה

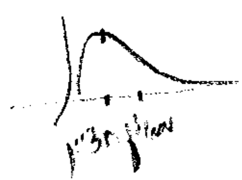
$= \frac{\sigma_{xy}}{\sigma_x \sigma_y} \longrightarrow 0 \le \rho_{xy} \le 1$

לא רק מעניינים איתנו התפלגויות רגילים בין משתנים

* אם ρ שווה ל-1 אזי יש לנו קשר לינארי מושלם בין x ל-y.
* $\rho = -1$ ← מושלם (הפוך)
* בהתפלגות נורמלית: אם שני משתנים בלתי תלויים ← $\rho = 0$.

אבל: אם $\rho = 0$ זה לא אומר שהמשתנים ב"ת.

חציון: הערך באמצע כך שמימינו ומשמאלו נמצאים מחצית האוכלוסיה.
(מתאים להתפלגויות לא סימטריות כמו הכנסה)



שעור מס 2 :

מדגם מקרי: (מדגם מ"ב)

איסוף הנתונים מתוך ההתפלגות שנבחר באופן בלתי תלוי. X_1 לא תלוי ב X_2...

תכונות:

(1) $E(X_1) = E(X_2) = E(X_3) = \dots = \mu_x$

(2) $V(X_1) = \dots = V(X_n) = \sigma_x^2$

(3) $Cov(X_i, X_j) = 0, \ \forall i \neq j$

פרמטר: מספר שמתאר מאפיין של פונקציות אוכלוסיה לדוגמא שונות, תוחלת או קשר בין שני

אומד: פונק' של משתנים מקריים הנותנת ערך לכל תוצאה אפשרית של המדגם. נסמן את האומד לפי ב $\hat{\theta}$.

אומד ללמד אותנו משהו על הפרמטר.

אומדן: הערך המספרי של אומד עבור מדגם מסוים.

בדוגמא: V ממוצע (בחרים) הוא אומד לתוחלת ← אנחנו בוחרים אותו

כי הוא מקיים 2 תכונות: (1) חסר הטיה

(2) בעל שונות מינימלית (="יעיל")

אומדים חשובים:

פרמטר	אומד
$E(x) = \mu_x$	$\bar{x} = \frac{\sum x_i}{n}$ ממוצע
$Var(x) = E(x-\mu_x)^2$	אומד לשונות: $\hat{Var}(x_i) = \frac{\sum(x_i-\bar{x})^2}{n-1} = \hat{\sigma}_x^2 = S_x^2$
$Cov(x,y) = E(x-\mu_x)(y-\mu_y)$	$\hat{Cov}(x,y) = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n-1} = \hat{\sigma}_{xy} = S_{xy}$

עם שורה חסר הטיה

*

$r = \frac{Cov(x,y)}{\sqrt{V(x)}\sqrt{V(y)}}$ (מיוצג)

$\hat{\rho} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y}) / \cancel{n-1}}{\sqrt{\frac{\sum(x_i-\bar{x})^2}{\cancel{n-1}}} \cdot \sqrt{\ }}$

① הגדרת חוסר הטיה של אומד $\hat{\theta}$: $E(\hat{\theta}-\theta)=0$ ← מצגמם

← נשים לב שבהגדרה השע לבינן. $\Rightarrow E(\hat{\theta})=\theta$

② שונות מינימלית: / יעילות

ניח, שני אומדים חסרי הטיה ← $\hat{\theta}_1, \hat{\theta}_2$,
(תסבה)

נאמר כי $\hat{\theta}_1$ יותר (יעיל) מ $\hat{\theta}_2$ אם:

$$Var(\hat{\theta}_1) < V(\hat{\theta}_2)$$

← לפעמים נעדיף אומדים מוטים אבל עם שונות קטנה יותר.

הסבר לכך שממוצע הוא האומד הכי טוב לתוחלת ← רוצים שהוא חסר הטיה ובעל השונות המינימלית.

(א) נראה ש $\bar{x}$ חסר הטיה.

(ב) נתאר את כל האומדים חסרי ההטיה (הליניארים) האפשריים לתוחלת. ← עבור מדגם עם 2 תצפיות

(ג) נראה ש $\bar{x}$ הוא בעל השונות המינימלית מביניהם.

(ב) שונות ⟶ למקרה עם $\underline{n}$

(א) $E(\bar{X}) = E\left(\frac{X_1+\dots+X_n}{n}\right) = \frac{1}{n}\cdot n\cdot \mu_x = \mu_x$

(ב) עבור מדגם $n=2$ נניח $\hat{\theta}=\lambda_1X_1+\lambda_2X_2$

כך מנת שתהיה חסר הטיה נדרוש: $E(\lambda_1X_1+\lambda_2X_2)=\mu_x$

$$=(\lambda_1+\lambda_2)\left[E(x)\right]=\mu_x \Rightarrow \underline{\lambda_1+\lambda_2=1}$$

(ג) נמצא אומד בעל שונות מינימלית:

$$\min Var(\hat{\theta}) = Var(\lambda_1X_1+\lambda_2X_2)=\lambda_1^2V(X_1)+\lambda_2^2V(X_2)+2\lambda_1\lambda_2 Cov(X_1,X_2)$$

$$s.t.\ \lambda_1+\lambda_2=1$$

תצפיות בלתי תלויות $=0$

שקול

$$\Rightarrow (\lambda_1^2+[1-\lambda_1]^2)V(x)$$

$$\frac{d}{d\lambda_1} = V(x)\left[2\lambda_1+2[1-\lambda_1]\cdot(-1)\right]=0$$

$$\Rightarrow 2\lambda_1-2+2\lambda_1=0$$

$$4\lambda_1=2 \Rightarrow \boxed{\lambda_1=\frac{1}{2}}$$

[illegible]

משתנה מקרי

שונות $\bar{X}$ לשונה עם n תצפיות:

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2} Var\left(\sum X_i\right)$$

$$= \frac{1}{n^2}\left[Var(X_1+X_2+\dots+X_n)\right] = \frac{1}{n^2}\left[VarX_1+\dots+VarX_n+\underbrace{2Cov(\dots)}_{0}\right]$$

$$= \frac{1}{n^2}\cdot n\,Var X = \boxed{\frac{Var X}{n}}$$

(מה קורה לפרמטר כשמגדילים מדגם)

הגדרת תכונה נוספת: עקביות $\bar{X}$ יהיה קרוב לתוחלת μ_x בהסתברות גבוהה יותר ככל שמספר התצפיות n גדל.

$$\lim_{N\to\infty} P\left[|\hat{\theta}-\theta|<\varepsilon\right]=1$$ עבור ε קטן כרצוני.

(9) נראה כי $E(\hat{Var}(x_i)) = \sigma_x^2$:

$$(n-1)\cdot \hat{Var}(x_i) = \sum (x_i - \bar{x})^2 = \sum (x_i - \mu_x + \mu_x - \bar{x})^2$$

$$= \sum \left((x_i - \mu_x) - (\bar{x} - \mu_x)\right)^2 = \sum (x_i - \mu_x)^2 + \sum (\bar{x} - \mu_x)^2 - 2\sum (x_i - \mu_x)\underbrace{(\bar{x} - \mu_x)}_{\text{מספר}}$$

$$= \sum (x_i - \mu_x)^2 + \sum (\bar{x} - \mu_x)^2 - 2(\bar{x} - \mu_x)\sum (x_i - \mu_x)$$

$$= \sum (x_i - \mu_x)^2 + \sum (\bar{x} - \mu_x)^2 - 2(\bar{x} - \mu_x)\cdot \left[\overbrace{x_1 + \dots + x_n}^{n\bar{x}} - N\mu_x\right]$$

$$= \sum (x_i - \mu_x)^2 + \sum (\overbrace{\bar{x} - \mu_x}^{\text{מספר}})^2 - 2\cdot N\cdot (\bar{x} - \mu_x)^2 \longrightarrow n(\bar{x} - \mu_x)^2$$

$$\boxed{= \sum (x_i - \mu_x)^2 - N(\bar{x} - \mu_x)^2}$$

$$E(\hat{Var}(x_i)) = \frac{1}{N-1}\cdot E\left(\sum (x_i - \mu_x)^2 - N(\bar{x} - \mu_x)^2\right)$$

$$= \frac{1}{N-1}\cdot E\left(\sum (x_i - \mu_x)^2\right) - \frac{N}{N-1}\cdot \underbrace{E(\bar{x} - \mu_x)^2}_{\frac{\sigma_x^2}{N}}$$

$$= \frac{1}{N-1}\cdot \sum E(x_i - \mu_x)^2 - \frac{\sigma_x^2}{N-1}$$

$$= \frac{1}{N-1}\cdot N\sigma_x^2 - \frac{\sigma_x^2}{N-1} = \frac{\cancel{(N-1)}\sigma_x^2}{\cancel{(N-1)}} = \sigma_x^2$$