

שבוע 1 · Conjectures

מטרתו של מדע אמפירי

מנגנון חבוי, נתונים גלויים, והדרך אחורה

פרופ' ניצן שחר

מה נעשה כאן

- נעקוב אחר מטרתה של הסטטיסטיקה במדע אמפירי.
- נגדיר את **המודל המייצר** — איך הטבע מפיק את המספרים שאנחנו מודדים.
- נדבר על יצירת נתונים ועל הנדסה לאחור.
- נבדיל בין **inference** ל-**prediction**.

התמונה הגדולה

מהי מטרתו של מדע אמפירי?

המטרה המרכזית שלנו היא **הערכת חוקיות חבויה בטבע**.

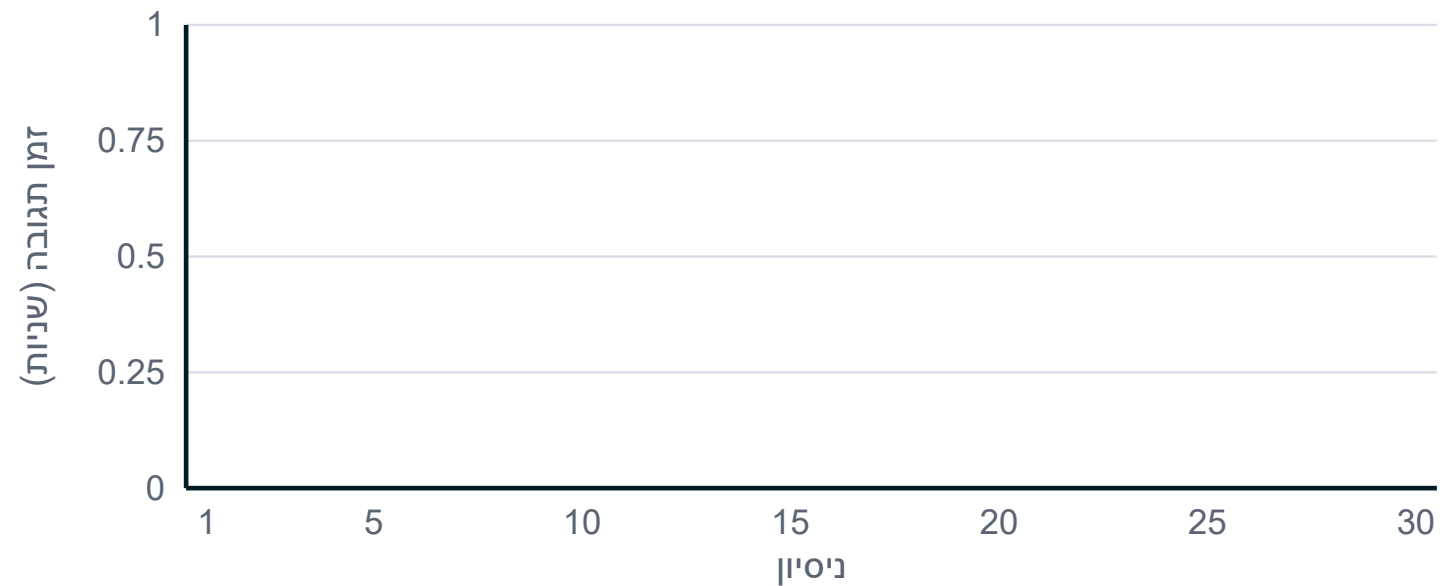
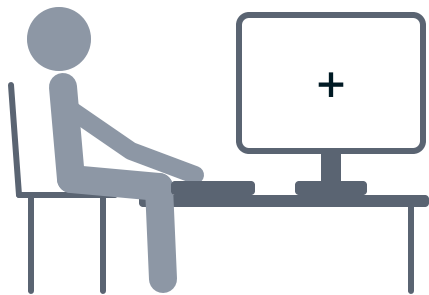
הטבע פועל לפי חוקים מסוימים, אך החוקים האלו נסתרים מעינינו. אנו לא יכולים לגשת אליהם או למדוד אותם באופן ישיר. המטרה שלנו היא ללכת אחורה מהנתונים אל המנגנון האבסטרקטי.

התודעה כמכשיר מדידה רועש

- הכלי העיקרי שדרכו אנו מכירים את העולם הוא התודעה האנושית.
- **ריאליזם נאיבי:** האמונה שהעולם הוא פחות או יותר כפי שהוא נתפס בחושינו.
- **ריאליזם מדעי:** ההבנה שהתודעה והחושים שלנו הם רק מכשירי מדידה, והם מכשירים **רועשים**.
- יש פער מובנה בין המציאות האמיתית (הפרמטר החבוי) לבין מה שאנחנו תופסים (התצפית).

דוגמה 1: זמני תגובה

- אנו מודדים זמני תגובה של נבדק בסדרת ניסיונות: $y_1, y_2, \dots, y_n$.
- מה שרואים (y) : תצפיות המשתנות מניסיון לניסיון.
- מה שקיים בטבע (θ) : מהירות עיבוד אופיינית אחת, לטנטית וקבועה.
- אף תצפית בודדת אינה "המהירות". התפלגות התצפיות מלמדת אותנו על הרעש ועל המרכז.

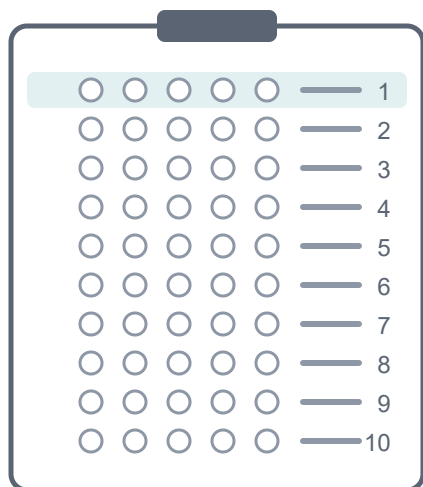


איפוס

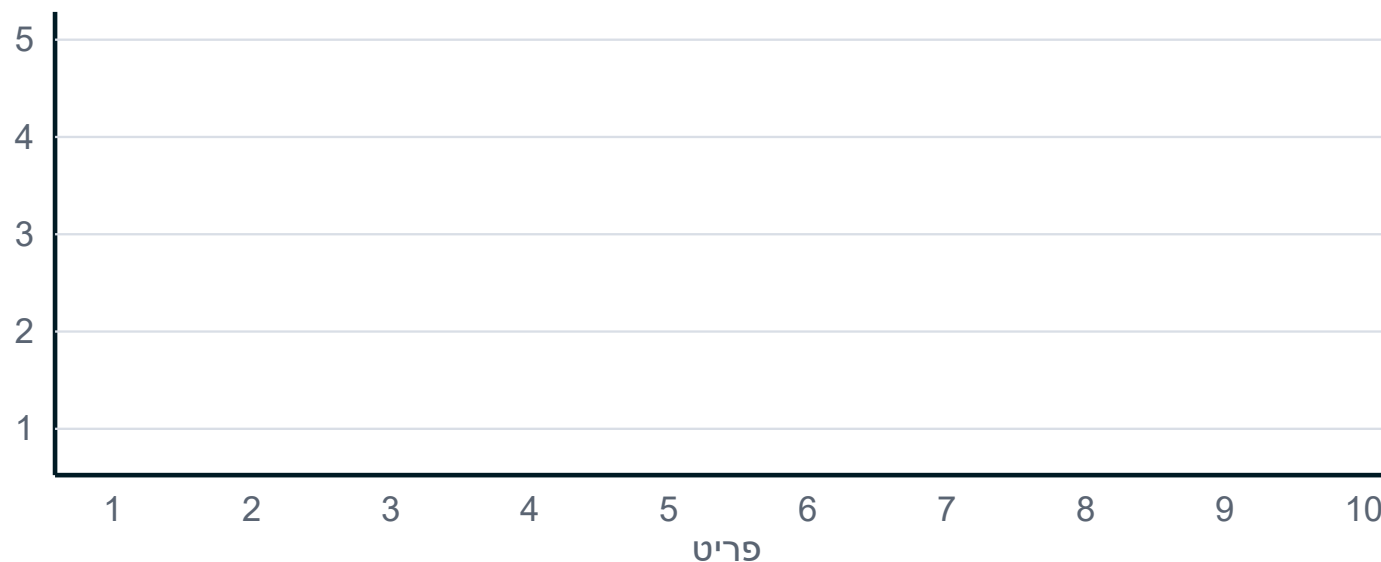
ניסיון חדש

דוגמה 2: רמת חרדה

- מטופל ממלא שאלון עם מספר פריטים ומספק דירוגים: $y_1, y_2, \dots, y_n$.
- מה שרואים (y) : התשובות הגלויות, שמושפעות מגורמים רבים.
- מה שקיים בטבע (θ) : רמת חרדה תכונתית אמיתית אחת.
- שאלה אליכם: גם אם תכונת החרדה (θ) לא תשתנה כלל, האם ציון השאלון (y) יכול להשתנות? ומדוע?



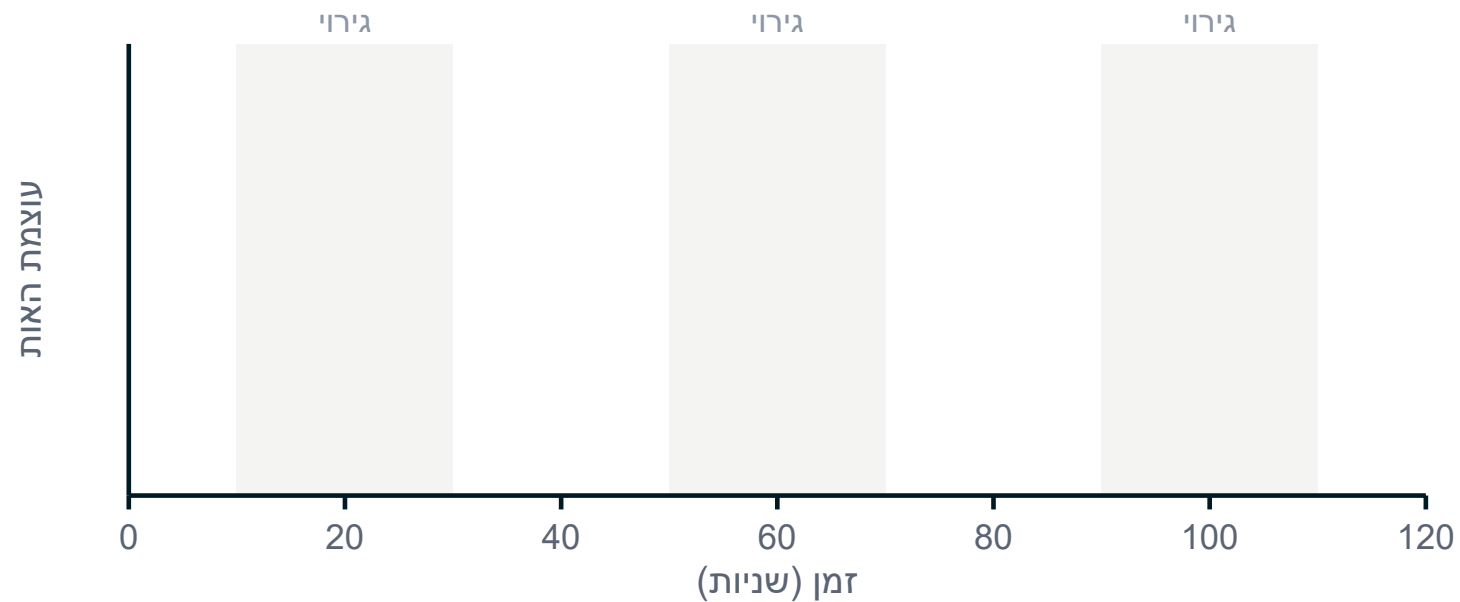
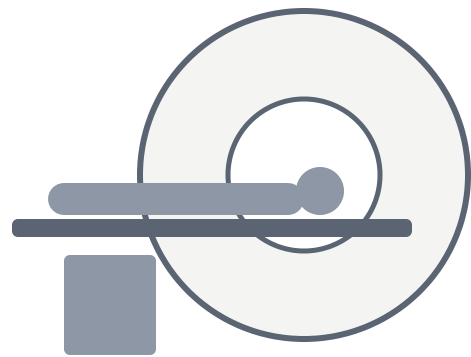
☐	1
☐	2
☐	3
☐	4
☐	5
☐	6
☐	7
☐	8
☐	9
☐	10



פריט הבא

דוגמה 3: קידוד עצבי (fMRI)

- אנו מודדים את עוצמת האות (BOLD signal) באזור מוחי מסוים לאורך זמן: $y_1, y_2, \dots, y_n$.
- מה שרואים (y) : תנודות מתמידות באות הנמדד בסורק.
- מה שקיים בטבע (θ) : רמת האקטיבציה העצבית האמיתית כתגובה לגירוי.
- שוב, התצפית היא תוצר של המנגנון העצבי יחד עם רעש פיזיולוגי ומכשירי.



סריקה של 10 שניות

אותו מבנה, שלוש פעמים

מה שחבוי (θ)	מה שנצפה (y)	
מהירות עיבוד קבועה	מדידות זמן משתנות	זמני תגובה
רמת תכונה אמיתית	תשובות משתנות לשאלון	חרדה
אקטיבציה עצבית טהורה	אותות BOLD רועשים	קידוד עצבי

בכל שלושת המקרים: משהו קבוע וחבוי מייצר משהו משתנה וגלוי דרך התפלגות. זה כל המדע האמפירי.

המודל המייצר (Generative Model)

- פרמטר (θ): קבוע חבוי בטבע. "מספר בסוד". טהור, אינו משתנה בין מדידה למדידה.
- המודל המייצר (**generative model**): החוקיות, או ההתפלגות, שמתארת את המנגנון. הנוסחה שקושרת בין הפרמטר הטהור לבין העולם הגלוי.
- משתנה מקרי (y): התוצאה הגלויה שמופקת בכל מדידה. משתנה כל פעם בגלל הרעש.

$$y \sim \text{Distribution}(\theta)$$

יצירת נתונים והנדסה לאחור

יצירת נתונים (Data Generation)

- כיוון הפעולה של הטבע: מהפרמטר (θ) החבוי אל התצפיות (y) הגלויות.
- ה- θ קבוע, מכוסה ובלתי נראה.
- כל דגימה מוציאה עוד תצפית, ומספר הדגימות שלנו מוגבל.

הנדסה לאחור (Inference)

- כיוון הפעולה של המדען: יש לנו רק את התצפיות (y).
- מהן אנחנו מסיקים אחורה מהו הפרמטר (θ) שיצר אותן.
- ה- θ האמיתי נשאר חבוי. אפשר רק להעריך אותו.

הסקה (inference) מול ניבוי (prediction)

חשוב להבחין בין שתי מטרות שונות:

- **הסקה (Inference):** הרצון להבין את המנגנון הסיבתי ואת החוקיות שיצרה את הנתונים (להבין את θ).
- **ניבוי (Prediction):** היכולת לחזות מה תהיה התצפית הבאה (y_{new}), גם אם אנחנו לא מבינים או לא מתעניינים במנגנון האמיתי שעומד מאחוריה.
- בקורס הזה נתמקד קודם בהסקה, אך נחזור לדבר לעומק על ניבוי בהמשך.

הגולם והינשוף

הגולם

מנוע החישוב. חזק, מדויק, ומציית בלי לשאול. תזין לו הנחה שגויה — הוא יבצע אותה בשלמות ויחזיר מספר משכנע (חישוב המודל).

הינשוף

החשיבה המדעית. שואל מה המודל בכלל מניח, ומה מותר להסיק מהתוצאה. הוא זה שמחליט מה לתת לגולם לעשות.

הקורס הזה מלמד את שניהם — אבל הציון נקבע לפי הינשוף.

התפתחות החשיבה הסטטיסטית

על כתפי ענקים

המאה ה-18	יאקוב ברנולי	הסתברות ותהליכים בינאריים
	תומאס בייס	הסקה הפוכה ועדכון ודאות
המאה ה-19	לפלס וגאוס	התפלגויות ושיטת הריבועים הפחותים
המאה ה-20	פירסון ופישר	סטטיסטיקה שכיחותנית, p-values
המאה ה-21	אנדרו גלמן, ריצ'ארד מקאלרת'	מודלים בסיאניים מודרניים וכוח חישובי

סיכום וחזרה

- המטרה של מדע אמפירי: להעריך חוקיות חבויה בטבע, מתוך הנתונים.
- ריאליזם נאיבי מול ריאליזם מדעי: החושים הם מכשיר מדידה רועש, לא חלון ישיר אל העולם.
- פרמטר (θ): קבוע חבוי, שלא משתנה בין מדידה למדידה.
- משתנה מקרי (y): התצפית הגלויה, שמשתנה בכל מדידה.
- מודל מייצר (**generative model**): ההתפלגות שקושרת בין θ לבין y .
- יצירת נתונים: הכיוון של הטבע, מ- θ אל y .
- הנדסה לאחור (**inference**): הכיוון של המדען, מ- y חזרה אל θ .
- ניבוי (**prediction**): לחזות את y הבא, גם בלי להבין את המנגנון.
- הגולם והינשוף: החישוב מבצע, והחשיבה המדעית מחליטה מה לחשב ומה מותר להסיק.

