

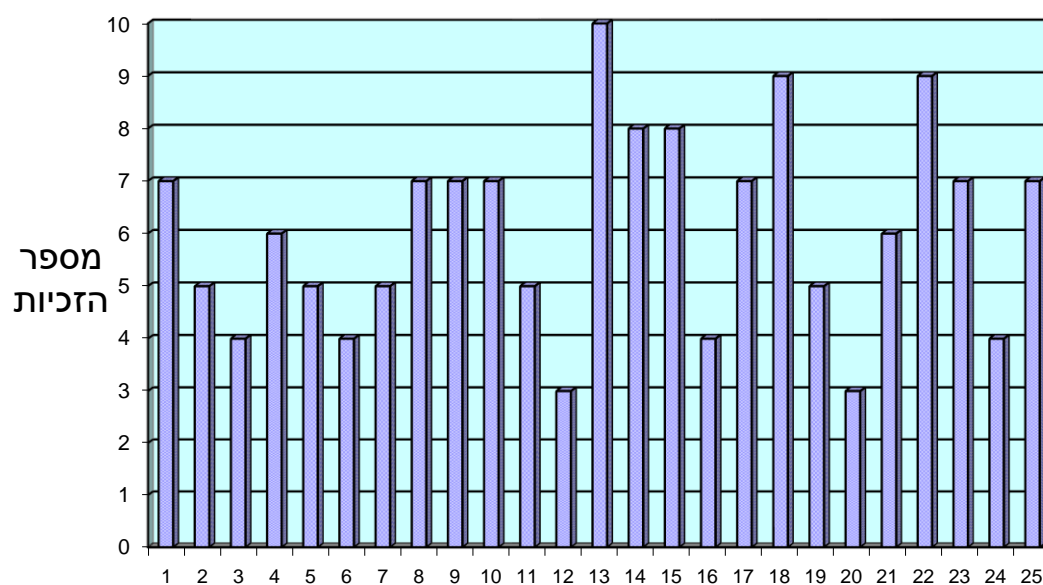
פרק שלושה עשר

הניסוי הגדול הראשון

הנתונים הוקלדו למחשב והוא החל לבצע את הוראות התוכנה. הוא חיפש בטקסט של ספר בראשית את צפני ELS של הזוג הראשון, ופעל לפי המתואר לעיל (בפרק על כימות המפגשים). דהיינו, הוא חישב את המפגשים בין המד"שים על פני סדרת הטבלאות הדו-ממדיות שהם קובעים, ומצא את ערכה של "מידת הקרבה המכילת" עבור הזוג הראשון. כך נעשה עבור כל זוגות המדגם. משחלפו הימים ועברו השבועות והסתיימה ההרצה, התברר, כי הניסוי הצליח. נדון כאן בתוצאות הסופיות של הניסוי ובמשמעותן. התוצאות הסופיות העלו 152 מספרים. מספרים אלה היו ערכי "מידת הקרבה המכילת" של זוגות הביטויים במדגם, אשר הופיעו בדילוג שווה בספר בראשית. כמבואר בנספח א2, ערכו של כל מספר כזה הוא בין 0 ל-1. אם המפגשים של צפני ה-ELS מוצלחים במיוחד – ערכה של מידה זו קרוב ל-0. לעומת זאת, אם המפגשים של צפני ה-ELS חלשים במיוחד – ערכה של מידה זו קרוב ל-1.

אם אין כאן תופעה – דהיינו, ההופעה של צפני ELS בספר בראשית אינה אלא מקרית – יש לצפות שמספר התוצאות הקרובות ל-0 יהיה דומה למספר התוצאות הקרובות ל-1; נכון יותר – למספר התוצאות הקרובות לכל ערך בין 0 ל-1. מצב זה, בו הכל מתקבל באקראי, דומה להגרלה אקראית של מספרים. למשל, נדמיין לעצמנו רולטה בת 25 מספרים שאינה "מזייפת", דהיינו, לכל אחד מן המספרים סיכוי שווה לעלות בהגרלה. כיצד ייראו התוצאות של 152 הגרלות? - אנו מצפים למהו דומה לאיור הבא:

התפלגות הזכיות



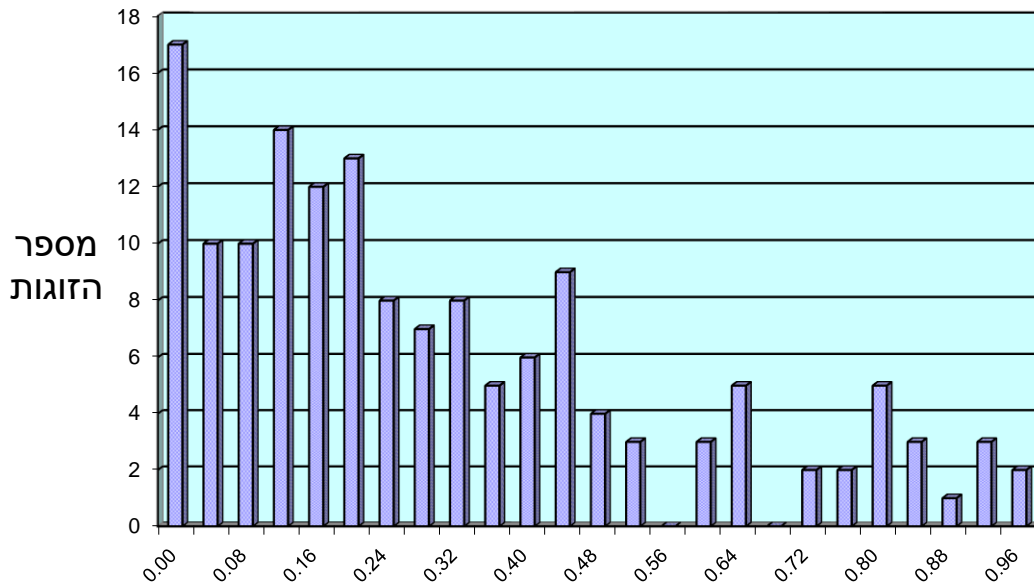
המספר הזוכה

איור יג-1

לפי האיור, אנו רואים כי המספר 1 עלה בגורל 7 פעמים (גובה העמודה הראשונה משמאל), המספר 2 עלה בגורל 5 פעמים, המספר 3 – 4 פעמים, וכן הלאה. במוצע, יש לצפות ל-6 זכיות בערך לכל מספר (תוצאת פעולת חילוק של 152 הגרלות ב-25 מספרים שווי סיכוי). הערך המתקבל למעשה עבור כל מספר מתנדנד בסביבות הערך הממוצע.

במקביל לדוגמא זו, נראה כיצד מתפלגים 152 המספרים שקיבלנו בניסוי. נחלק את המרווח שבין 0 ל-1 ל-25 מרווחים משניים. באמצעות גובה העמודה נסמן בכל מרווח כזה את מספר הזוגות שתוצאותיהם נפלו באותו מרווח. אם הכל הוא מקרי בלבד, אכן צפוי כי יתקבל גרף דומה לאיור יג-1. אבל מה קרה באמת? הרי לפנינו התוצאה:

התפלגות התוצאות בבראשית



ערכי "מידת הקרבה המכילית"

איור יג-2

במרווח הראשון (העמודה הראשונה משמאל) נרשמו התוצאות שבין 0 ל-0.04. במדגם היו 17 זוגות, שערכי "מידת הקרבה המכילית" שלהם התקבלו במרווח זה. במרווח השני, נרשמו התוצאות שבין 0.04 ל-0.08: התקבלו 10 תוצאות כאלו. וכן הלאה. לפנינו אפוא הצטברות חזקה של תוצאות "באזור ההצלחה" (קרוב ל-0), לעומת התרוקנות משמעותית של תוצאות "באזור הכשלון" (קרוב ל-1).⁽¹⁾

את מידת "הנטייה הכוללת לקרבה" של הקבוצה כולה, של כל 152 הזוגות, נחשב באמצעות שתי המידות (הסטטיסטיים) שהוגדרו בשלב ב, וסומנו P_1 ו- P_2 . מידה אחת, שסומנה ב- P_1 , היא פשוטה ביותר. מונים את מספר התוצאות "באזור ההצלחה", אשר הוגדר (שרירותית) כמרווח בין 0 ל-0.2. למשל, מתוך 152 התוצאות התקבלו 63 במרווח זה (באיור יג-2 זהו סך 5 העמודות משמאל). כמה צפויים לעלות במקרה? – $152 \times 0.2 = 30.4$ תוצאות במוצע. מה הסיכוי לקבל סטייה כה גדולה? – קטן מאד. אם התוצאות הן שוות-סיכוי ובלתי תלויות זו בזו, הסיכוי לקבל סטייה כזו במקרה, הוא בסדר גודל של אחד למיליארד!¹ המידה השנייה, שסומנה ב- P_2 , מניבה ערך שהוא בסדר גודל כזה. כלומר, נמדדה כאן נטייה חזקה של בני הזוג למפגשים "בקרב מקום".

¹ זה הערך לפי ההתפלגות הבינומלית. אולם, כפי שכתבנו בנספח א3, באותו זמן עדיין לא חישבנו את P_1 במדויק, והסתפקנו בקירוב הנורמלי להתפלגות הבינומלית. עבור הנתונים הנ"ל מקבלים 6.61 סטיות תקן. זו תוצאה חזקה בהרבה מאחד למיליארד.

בזה נסתיים שלב ג.

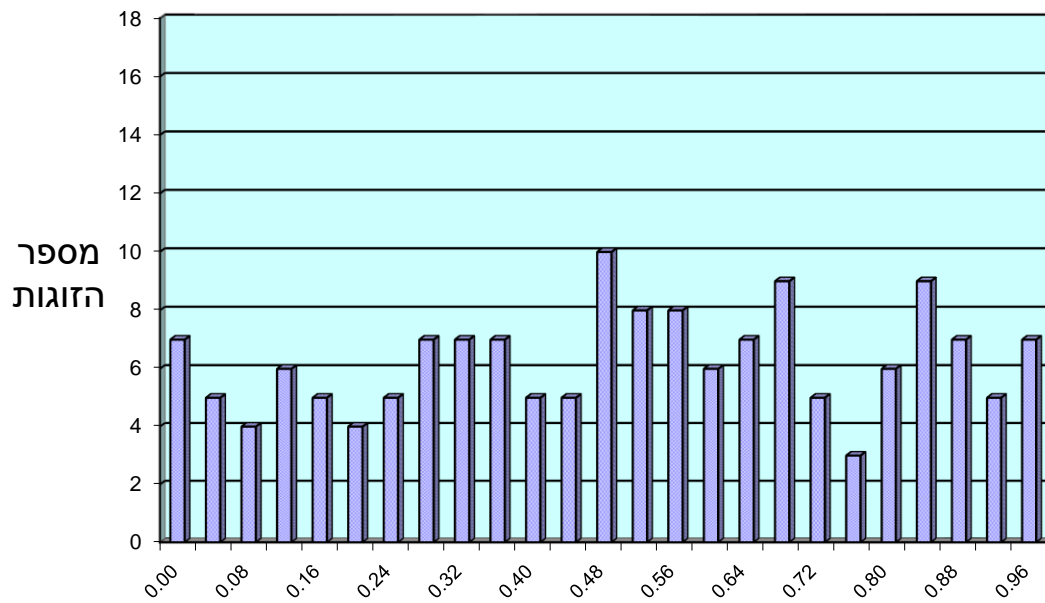
נזכיר, כי שתי המידות P_1 ו- P_2 נועדו למדוד את "ההסתברות הגולמית" (לא מנורמלת) של הצטברות אפשרית כזאת. כדי למצוא את ההסתברות המדוייקת (מנורמלת) היה צורך להפעיל את שלב ד: – לבדוק במבחן מובהקות האם הערך של הסטטיסטים עבור המדגם אכן קטן ממה שצפוי לקרות באקראי, ומה הסיכוי לקבל ערך כל כך קטן.

מבחן המובהקות המתבקש הוא להשוות את ערכי הסטטיסטים P_1 ו- P_2 לאלה המתקבלים באוסף גדול של טקסטים "דומים". אולם שלב ד לא בוצע: לפי אמצעי החישוב שעמדו לרשותנו משך העבודה לכל טקסט היה שלשה שבועות לפחות. (1000 טקסטים = 58 שנים...)

בתנאים שעמדו לרשותנו, הסתפקנו בבדיקת טקסטים אחדים לביקורת. בשלב ראשון, נעשה הדבר לגבי שני טקסטים, ותנאי הניסוי היו זהים: היו אלה טקסטים באורך ספר בראשית, רשימת הזוגות במדגם היתה אותה רשימה שבדקנו בספר בראשית, ואף התוכנה היתה אותה התוכנה שבניסוי המקורי. הטקסט האחד היה ספר בראשית עצמו, שעורבב כב"מיקסר" (באופן פסודו-רנדומלי) על ידי המחשב. הטקסט השני היה אוסף מאמרים של פרופסור שלום רוזנברג (אביו של יואב), שהיה זמין כקובץ מחשב. הרצת ניסויי הביקורת, שניצלה את כל יכולת החישובים שעמדה לרשותנו, היתה אמורה להתמשך לאורך זמן רב. שוב נקפו השבועות, והתוצאות הצטברו באיטיות. אלא, שהפעם היו צפויות התוצאות שהראה המחשב. התפלגות התוצאות הראתה בשני המקרים, כי מפגשי הזוגות בשני הטקסטים הללו התרחשו באקראי. כמות ההצלחות והכשלונות היתה כצפוי בהתרחשות אקראית. ואכן, משחושבו שתי המידות של "הנטיה הכוללת לקרבה" עבור רשימת התוצאות בכל טקסט, התברר שערכיהן אינם קטנים, ושהם משקפים העדר נטייה למפגשים.

אציג כאן תוצאות של ניסויי ביקורת מאוחרים יותר בשני טקסטים המבוססים על ספר בראשית. הטקסט האחד, אותו נכנה "טקסט R", הוא ערבוב אקראי של ספר בראשית (שונה מזה הנ"ל). הפרטים הטכניים בנספח א7. (שימו לב, כי מספר הזוגות המופיעים כמד"שים משתנה מטקסט לטקסט):

התפלגות התוצאות בטקסט R



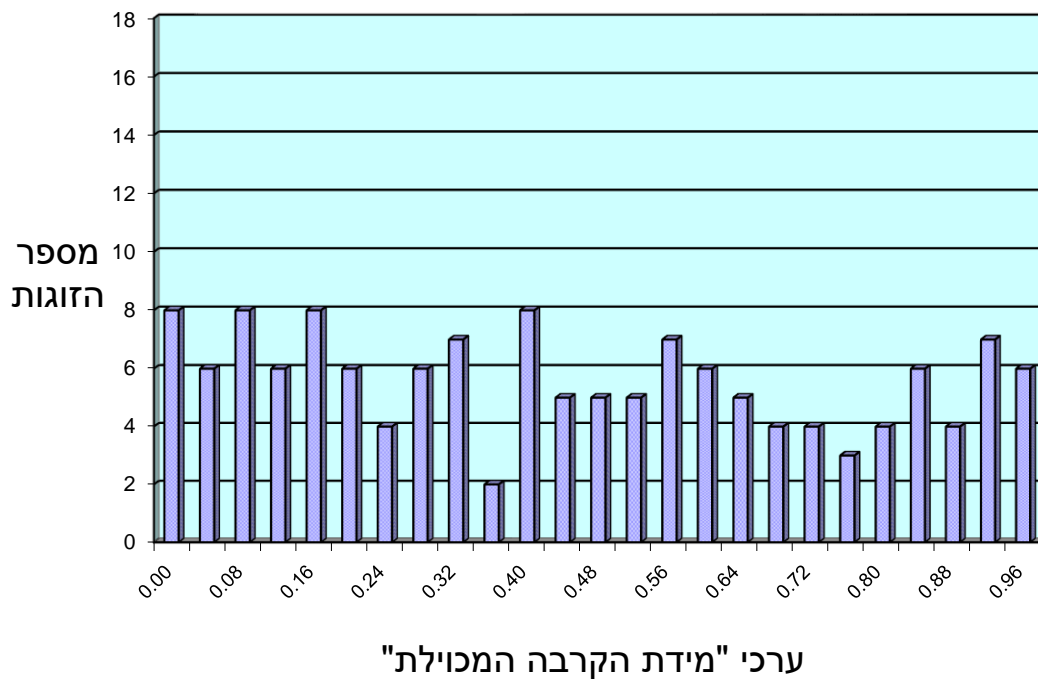
ערכי "מידת הקרבה המכילת"

איור יג-3

הטקסט השני – "טקסט M" – נוצר על ידי מחיקה של ה"א הידיעה הראשונה בכל 1000 אותיות בספר בראשית (בסך הכל נמחקו 78 אותיות ה"א²). מובן, שהטקסט השני דומה ביותר לספר בראשית. באיור הבא נראה את התוצאות עבור טקסט M:

² מתוך 6283 אותיות ה"א בספר בראשית.

התפלגות התוצאות בטקסט M



איור יג-4

כמובן, אין שום דבר מרגש בתוצאות מעין אלו. אך ההשוואה בין התפלגויות התוצאות עבור זוגות המדגם בטקסטים אלה, לבין התפלגות התוצאות בספר בראשית – ממחישה את ההבדל העצום!

הניסוי הסתיים באביב שנת התשמ"ו (1986 למניינם). סיכמנו את שנעשה במאמר. המאמר תיאר בפירוט, לאורך עשרות עמודים, את התופעה ואת שיטת המדידה: את "מידת הקרבה", את "מידת הקרבה המכוילת" ואת "הנטייה הכוללת לקרבה". גם ניתן בו תיאור מקיף של המדגם, כך שכל אחד יכול היה לנסות לשחזר את הניסוי. מלבד כל זאת צירפנו למאמר אוסף של דוגמאות שהבהירו היטב – כך קיוונו – במה המדובר. המאמר יצא לאור כתדפיס מחקר (preprint), בסתיו (ממש בתחילת שנת התשמ"ז), תחת הכותרת "צפני ELS בספר בראשית"³.

³ במקור:

D. Witztum, E. Rips, and Y. Rosenberg, *Equidistant Letter Sequences in the Book of Genesis*. Preprint. 1986.