בס"ד טיוטה אחרונה: כ"ג טבת התשע"ט

ניסוי חדש בתחום המורחב: חלקים ב-ג

מאת דורון ויצטום

תקציר
המאמר "*Equidistant letter sequences in the Book of Genesis*" [1] מאת דורון ויצטום, אליהו ריפס ויואב רוזנברג (להלן ור"ר), אשר פורסם בכתב העת Statistical Science בשנת התשנ"ד, הציג תוצאות מחקר מדעי של הרמז בדילוג השווה בספר בראשית. מחקר זה עסק בקבוצה של מופעי ביטויים בדילוגים שווים, שהם מינימליים בקטעים גדולים בטקסט הנחקר. מופעים כאלה נכנה כאן "צפני *ELS*". במוקד המחקר היה ניסוי שבו נמדדה הקורלציה בין צפני *ELS* של השמות והכינויים של קבוצת אישים (גדולי תורה) לבין צפני *ELS* של תאריכי הלידה והפטירה שלהם. המדידה הראתה "סיגנאל" ברור. המובהקות הסטטיסטית של הסיגנאל נמדדה באמצעות מבחן פרמוטציות, דהיינו, רנדומיזציה של הנתונים. התוצאה היתה מובהקת ביותר, ברמה של 0.00002. תאור מפורט בעברית של העבודה פורסם בבד"ד [2].

במאמרנו "סיגנאל של מערכת צפני *ELS*: דעיכה מול הגברה" [3] נמדדה הקורלציה בין צפני *ELS* עבור הנתונים הנ"ל בתחום מדידה מורחב. תחום המדידה המורחב הינו קטע מתחילת ספר התורה, הכפול באורכו מתחום המדידה המקורי, שהיה ספר בראשית בלבד. התברר, כי בתחום המדידה המורחב הסיגנאל מתחזק באופן בולט ומובהק. מסקנת המאמר היתה, כי תחום ההצפנה של הנתונים הנ"ל קרוב יותר לתחום המורחב, מאשר לתחום המקורי. במאמרנו "סיגנאל של מערכת צפני *ELS*: ניסויים חדשים ב"תחום המורחב" [4] דווח על ניסויים נוספים בנושא רשימת הנתונים הנ"ל בתחום המורחב. בכולם נתקבלה מובהקות חזקה. בעבודות [3] ו[4] נמדדה המובהקות באמצעות רנדומיזציה של הטקסט.

במאמרנו "ניסוי חדש בתחום המורחב" [5] מתואר ניסוי בנושא רשימת הנתונים הנ"ל בתחום המורחב, שנעשה בעקבות העבודות [3] ו[4]. הניסוי מבוסס על רשימת נתונים, ששימשה את מבקרי המחקר בניסוי העיקרי מתוך שלושה ניסויים "החשובים ביותר" ("Perhaps the most important class of experiments we have conducted") [6] (עמ' 163-164). רשימת נתונים זו, שנועדה לחקות את רשימת הנתונים ששימשה בניסוי [1] של ור"ר, הוכנה לדבריהם על ידי מומחה בלתי תלוי, שמחה עמנואל. הם מדווחים כי תוצאות הניסוי על רשימת "החיקוי" בספר בראשית אינן מובהקות במבחן הפרמוטציות, והם רואים בכך ראייה מכרעת לטענתם המרכזית כי הצלחת הניסוי של ור"ר נבעה אך ורק ממניפולציה של הנתונים.

אולם, במדידת מובהקות רשימת "החיקוי" באמצעות רנדומיזציה של הטקסט נתברר [5], כי התוצאות בספר בראשית מובהקות בהחלט, וכי התוצאות בתחום המורחב מובהקות ביותר. מתוצאות אלה נובעת מסקנה ההפוכה מטענת מבקרי המחקר. לפנינו הוכחה, שהתוצאה המקורית של ור"ר איננה תולדה של מניפולציה של שמות וכינויים.

במאמרנו [5] נמדדה רשימת "החיקוי" כפי שהיא. במאמר הנוכחי, שהוא המשך ל[5], נבדקה טענתנו שהוצבה החל מתחילת הוויכוח: כי רשימת נתונים שתיווצר על ידי מומחה בלתי תלוי עבור 32 האישים, שהיו כלולים בניסוי ב[1], תצליח בספר בראשית ותיכשל ב"מלחמה ושלום" [7] (זה הספר בו השתמשו מתנגדי המחקר לצורך פרודיה על ניסוי ור"ר).

אכן מתברר, כי כאשר נתוני עמנואל נלקחים עבור אותם 32 אישים, שהיו בניסוי ב[1], התוצאות מתחזקות עוד יותר מאלה ב[5], הן בספר בראשית והן בתחום המורחב. לעומת זאת – מתקבל כשלון עבור "מלחמה ושלום."

עוד מתברר, כי אם עמנואל אכן היה פועל לפי הכללים שהוצגו לפניו, כפי שנדרש ממנו ע"י מעסיקיו, התוצאות היו מוסיפות ומתחזקות עוד יותר, הן בספר בראשית והן בתחום המורחב.

אכן, כפי שכתבו מתנגדי המחקר, לרשימת "החיקוי" שנוצרה על ידי מומחה מטעמם יש חשיבות גדולה מאד: תוצאות מדידותינו ב[5] וגם כאן, מפריכות את טענתם המרכזית הנ"ל. הן מוכיחות, כי יש בתורה תופעה אמיתית של הצפנת נתונים, וכי התוצאה המקורית של ור"ר איננה תולדה של מניפולציה של שמות וכינויים.

הערה

מאמר זה הוא מאמר המשך למאמרנו "ניסוי חדש בתחום המורחב" (להלן: חלק א). כדי להבין את הנאמר כאן יש לקרוא בראשונה את חלק א. בפרט, לא נחזור כאן על הגדרות של מונחים ומושגים הנמצאות בחלק א ובנספח שם.

ניסוי חדש בתחום המורחב - חלק ב:
מדידת נתוני עמנואל עבור 32 האישים המקוריים בעבודת ור"ר

מבוא

למדידות שנעשו בחלק זה חשיבות מיוחדת לאור הפולמוס שנתפתח סביב עבודת ור"ר [1]. להבנת חשיבות המדידות נקדים כמה ידיעות ומושגים נחוצים.

א. הרשימה הראשונה של גדולי חכמי ישראל (מסומנת: *LIST 1*):

1. ייעודה: נועדה לבדוק בספר בראשית את המפגשים של צפני *ELS* המייצגים שמות וכינויים של גדולי חכמי ישראל ובין צפני *ELS* המייצגים את תאריכי הלידה והפטירה שלהם.
2. בחירת האישים: הוכנה רשימה של גדולי חכמי ישראל לפי ה"אנציקלופדיה לתולדות גדולי ישראל" [8], כאשר האישים "הגדולים ביותר" נבחרו לפי הקריטריון הבא: ברשימה נכללו אלה שערכם האנציקלופדי בן שלושה טורים ומעלה, ואשר תאריך הלידה ו/או הפטירה מצויין בו. ברשימה היו 34 אישים.
3. קביעת השמות והכינויים: רשימה של השמות והכינויים של האישים הנ"ל הוכנה מראש, באופן אפריורי, ולפי כללים מקצועיים על ידי פרופסור שלמה זלמן הבלין, ראש המחלקה לביבליוגרפיה וספרנות באוניברסיטת בר אילן באותה העת.
4. כללי האיות וצורת כתיבת התאריכים נקבעו מראש, באופן אפריורי, על ידי הבלשן יעקב אורבך ז"ל.
5. מדידת המפגשים העלתה כי קיימת נטיה חזקה מאד למפגשים בין שמות האישים ובין התאריכים הקשורים אליהם. ור"ר תיארו עבודתם זו במאמר [9]. הנתונים נמצאים בהרחבה גם ב[2].

ב. הרשימה השניה של גדולי חכמי ישראל (מסומנת: *LIST 2*):

1. בתגובה להצלחת *LIST 1*, דרש הסטטיסטיקאי פרסי דיאקוניס להכין רשימה חדשה של אישים מפורסמים, ולבדוק אותה בדיוק באותה התוכנה.
2. לצורך הרשימה החדשה לקחו הפעם ור"ר את האישים שערכם ב"אנציקלופדיה לתולדות גדולי ישראל" הוא בן 1.5 - 3 טורים[1], ואשר תאריך הלידה ו/או הפטירה מצויין בו.
3. התאריכים נכתבו בדיוק באותה המתכונת שכבר נקבעה. גם הפעם הוכנה רשימת השמות והכינויים באופן אפריורי על ידי הבלין, לפי אותם כללי איות וכללים מקצועיים. [מאוחר יותר, הבהיר הבלין במסמך [10] מה היו הכללים המקצועיים לפיהן פעל בקביעת הרשימות.]
4. הכללים הביבליוגרפיים וכללי האיות מהווים אילוצים כבדים על בחירת השמות והכינויים של האישים.

---

[1] למעשה נאלצו ור"ר לקחת קבוצת אישים זו, שהיא ברירת המחדל. כל בחירה אחרת של קבוצת אישים היתה נתקלת בטענה, כי זו רק בחירה אחת מהרבה בחירות דומות, ולכן צריך להכפיל את התוצאה במספר גדול (כרצון המבקר).

5. גם הפעם העלתה מדידת המפגשים, כי קיימת נטיה חזקה מאד למפגשים בין שמות האישים ובין התאריכים הקשורים אליהם. מבחן הפרמוטציות שיזם דיאקוניס למדידת מובהקות הסיגנאלים בוצע לפי דרישתו על *LIST 2*, כפי שפורסם ב[1].
6. בניסוי זה נמצא, כי תת-הרשימה *L2* היא שתרמה להצלחת הניסוי (ראו במבוא לחלק א, סעיף ג).

ג. הטענה המרכזית של מבקרי המחקר:
1. מבקרי המחקר ברנדן מק-קי, דרור בר-נתן, מיה בר-הלל וגיל קלעי (להלן מבב"ק) טוענים [6], כי חופש תמרון במסגרת הכללים הביבליוגרפיים והלשוניים ששימשו ליצירת הרשימות הנ"ל, בשילוב עם חריגות מאותם כללים, איפשרו בחירת השמות והכינויים "המצליחים" לתת-הרשימה *L2* והשמטת "הנכשלים", באופן שתיווצר הצלחה מלאכותית של *L2*.
2. להמחיש זאת, הם יצרו פארודיה על עבודת ור"ר והדגימו "הצלחה" מלאכותית של רשימת נתונים "דומה" בספר "מלחמה ושלום". אומנם, עבודתם זו לקויה מכמה בחינות [11] וחסרה אפשרות לכימות, ולכן חסרה תוקף מדעי [12].
3. לכן, עשו מבב"ק ניסוי "חשוב ביותר" בו רצו לחקות את יצירת *L2*. לדבריהם [6] הם נתנו לעמנואל משימה להרכיב רשימת נתונים חדשה לפי האילוצים של "הרשימה הראשונה"[2]. לטענתם, רשימה זו (המסומנת כאן ב- *LIST E2*) נכשלה במבחן הפרמוטציות בספר בראשית. לדעתם, תוצאה זו מוכיחה, כי רק חופש התימרון במסגרת האילוצים או חריגה מהם, גרמו להצלחת הניסוי של ור"ר.

ד. הרשימה *LIST E2*
1. מבב"ק פירסמו רשימת נתונים של 33 אישים בשמו של שמחה עמנואל (היא מועתקת בנספח לחלק א, סעיף ג). אולם, כבר לפני זמן רב התברר (הפרטים ב[13]), כי לאמיתו של דבר עמנואל הכין לבקשת מבב"ק רשימה גדולה יותר של 35 אישים, מהם 32 האישים שהיו אישי *LIST L2* של ור"ר, ועוד שלושה אישים. מבב"ק השמיטו מרשימת עמנואל המקורית שני אישים הכלולים ב-*LIST L2*, הותירו את שלושת האישים הנוספים, ופירסמו את הרשימה "המשופצת" כרשימת עמנואל (המסומנת כאן *LIST E2*). מלבד זאת התברר, כי מבב"ק התערבו מבחינות נוספות בהכנת הרשימה (הפרטים ב[13]).
2. למרות זאת, בדקנו בחלק א את תת-הרשימה *E2* כפי שהיא הן בספר בראשית והן בתחום המורחב.

ה. הרשימה *LIST EL2*
1. כבר בשלב מוקדם של הוויכוח הצענו מבחן מדעי, פשוט ונכון לבדיקת טענת המבקרים: הצענו שמומחה חיצוני אובייקטיבי יפעל באופן מוסכם על שני הצדדים ויכין עבור 32 אישי *LIST 2* תת-רשימה חדשה על פי האילוצים הנ"ל. תת-הרשימה החדשה תיבדק בספר בראשית, ואם היא תצליח יוכח, שלא חופש תימרון ולא שבירת הכללים הם שיצרו את ההצלחה. (פרטים על כך נמצאים ב[12]: במבוא, בפרק הראשון ובנספח לפרק הראשון.)
2. המבקרים סרבו לעמוד באתגר (פרטים ב[12] שם).
3. בהעדר עבודה אובייקטיבית כנ"ל, נבדוק את נתוני עמנואל עבור 32 האישים המקוריים בעבודת ור"ר, בהתעלמנו מהתערבות מבב"ק בעבודתו (ראו סעיף ד1 לעיל). נסמן רשימה זו ב- *LIST EL2*.

מדידה 1: מדידת הסיגנאל S_2 של *EL2* בספר בראשית.

א. המדגמים
כאמור לעיל, נתוני עמנואל עבור 32 האישים המקוריים בעבודת ור"ר מסומנים כאן ב- *LIST EL2*. בהקבלה גמורה למדידה בחלק א, נתמקד בתת-הרשימה *EL2*, ונגדיר עבורה את המדגמים הבאים:
מדגם [א1]: *EL2(AB)* הוא המדגם הנוצר מתת-הרשימה *EL2* כאשר התאריכים בה ניתנים בצמד התצורות AB.
מדגם [ב1]: *EL2(ABC)* הוא המדגם הנוצר מתת-הרשימה *EL2* כאשר התאריכים בה ניתנים

[2] בלשונם:
Emanuel was shown the spelling rules and table of appellations for WRR's first list as they first appeared in WRR (1986). He then compiled a parallel table of appellations for our list of 33 rabbis, attempting to follow the rules and practices of WRR's first list.

בשלש התצורות ABC.

מדגם [ג1]: *EL2(ABCD)* הוא המדגם הנוצר מתת-הרשימה *EL2* כאשר התאריכים בה ניתנים בארבע התצורות ABCD.

ב. מה נמדוד:

כאן נשוב ונמדוד את מובהקות הסיגנאל S_2 בספר בראשית, בדיוק כפי שנעשה בחלק א במדידה 1.

ג. מדידת המדגמים:

1. כל אחד מן המדגמים [א1], [ב1] ו[ג1] נמדד ב-G באותו אופן בדיוק כפי שנמדד המדגם המקורי ב[1]. כך נקבע ערכו הגולמי של הסיגנאל S_2 עבור כל מדגם.
2. כדי לקבוע את מובהקות הסיגנאלים, נעשתה השוואה לאוסף של 50,000 טקסטים T_j, שכל אחד מהם נוצר על ידי עירבוב אקראי של G, דהיינו, עירבוב של מלים בתוך הפסוקים, ונקבע ערכו של n_2, מספר המתחרים שלהם ערך $S_2^j \geq S_2$.
3. לפי n_2, חושבה מובהקות הסיגנאל S_2: $p_2 = n_2 / 50{,}000$.

התוצאות מוצגות בטבלה 1.

טבלה 1

מספר המתחרים n_2 שלהם ערך $S_2^j \geq S_2$ בתחרות ובה 50,000 מתחרים.

המדגם	*EL2(ABCD)*	*EL2(ABC)*	*EL2(AB)*
ערכי S_2	6.189E+2	7.647E+2	3.29E+3
n_2	149	100	16
p_2	2.98E−3	2.00E−3	3.2E−4

יש לשים לב לכך, שהתוצאות אינן בלתי תלויות. אדרבא, קיימת תלות חזקה בין תוצאות *EL2(AB)* לבין תוצאות *EL2(ABC)* ו-*EL2(ABCD)*.

ד. חישוב המובהקות הכוללת של S_2 עבור *EL2* בספר בראשית:

אם מדרגים את כלל הטקסטים בתחרות עבור *EL2(AB)*, הסיגנאל S_2 בספר בראשית מדורג במקום ה-17. השאלה העומדת לפנינו היא: מהו הסיכוי של טקסט להיות מדורג באקראי במקום 1-17 לפחות באחת משלוש התחרויות שבטבלה 1? – התשובה היא: סך כל הטקסטים השונים המאכלסים את 17 המקומות הראשונים בכל אחת משלושת התחרויות המסוכמות בטבלה 1, מחולק במספר הטקסטים שהשתתפו בתחרות. אם התחרויות היו בלתי תלויות, היינו מקבלים 17×3 טקסטים שונים המאכלסים את 17 המקומות הראשונים (או קרוב לזה). לעומת זאת, בגלל התלות בין המדגמים, כאן יש רק 36 טקסטים שונים כאלה. לכן, המובהקות הכוללת של S_2 עבור *EL2* בספר בראשית היא[3]:

$$\rho_2 = 36/50{,}000 = 7.2\text{E}{-4}$$

ה. לסיכום:

1. הסיגנאל S_2 של כל אחד משלושת המדגמים [א1], [ב1] ו[ג1] נמצא מובהק בספר בראשית יותר מאשר אותו סיגנאל עבור כל אחד משלושת המדגמים [א], [ב] ו[ג] בחלק א, בהתאמה.
2. מובהקות הסיגנאל החזקה ביותר היא עבור מדגם [א1] וערכה $p_2 = 3.2\text{E}{-4}$.
3. המובהקות הכוללת של S_2 עבור *EL2* בספר בראשית היא: $\rho_2 = 7.2\text{E}{-4}$.

מדידה 2: מדידת הסיגנאל S_2 של *EL2* בספר "מלחמה ושלום".

א. הטקסט:

מדידה זו נעשתה בהקבלה גמורה למדידה 2 בחלק א ובאותו טקסט: *WP*.

ב. מדידת המדגמים:

1. כל אחד מן המדגמים [א1], [ב1] ו[ג1] נמדד ב-*WP* באותו אופן בדיוק כפי שנמדד המדגם המקורי ב[1]. כך נקבע ערכו הגולמי של הסיגנאל S_2 עבור כל מדגם.

[3] ראו עוד בנספח לחלק א, סעיף ד.

2. כדי לקבוע את מובהקות הסיגנאלים, נעשתה השוואה לאוסף של 1000 טקסטים T_j, שכל אחד מהם נוצר על ידי עירבוב אקראי של *WP*, ונקבע ערכו של n_2, מספר המתחרים שלהם ערך $S_2^j \geq S_2$.
3. לפי n_2, חושבה מובהקות הסיגנאל S_2: $p_2 = n_2 / 1000$.
התוצאות מוצגות בטבלה 2.

טבלה 2

מספר המתחרים n_2 שלהם ערך $S_2^j \geq S_2$ בתחרות ובה 1,000 מתחרים.

EL2(AB)	*EL2(ABC)*	*EL2(ABCD)*	המדגם
5.208	5.848	3.731	ערכי S_2
157	132	203	n_2
0.157	0.132	0.203	p_2

תוצאות אלה מצביעות על חוסר מובהקות.

ג. מסקנות:
עבור תת-הרשימה *EL2*:
1. יש הבדל מהותי בין ערכו הגולמי של הסיגנאל S_2 בספר בראשית לבין ערכו ב"מלחמה ושלום".
2. יש הבדל מהותי בין מובהקות הסיגנאל S_2 בספר בראשית לבין מובהקותו ב"מלחמה ושלום".

מדידה 3: מדידת תת-הרשימה *EL2* בתחום המורחב.

א. מה נמדוד:
מדידה זו נעשתה בהקבלה גמורה למדידה 3 בחלק א.
כאן נבדקה תת-הרשימה *EL2* בתחום המורחב.

ב. הטקסט:
הטקסט הוא הקטע [1, 156128] מן התורה: $D' = ([1, 156128], TORAH)$.

ג. מדידת המדגמים:
1. כל אחד מן המדגמים [א1], [ב1] ו[ג1] נמדד בקטע D' באותו אופן בדיוק כפי שנמדד המדגם המקורי ב[1]. כך נקבע ערכם הגולמי של הסיגנאלים S_1 ו- S_2 עבור כל מדגם.
2. נחשב את מדד ההגברה Q_i לגבי הסיגנאלים (ראו הסבר במבוא לחלק א סעיף א5), וכן את מדד ההגברה המירבית $Q \equiv \max\{Q_1, Q_2\}$, ביחס לסיגנאלים המקבילים בקטע D (ספר בראשית).
התוצאות מוצגות בטבלה 3.

טבלה 3

EL2(AB)		*EL2(ABC)*		*EL2(ABCD)*		המדגם
1	2	1	2	1	2	*i*
1.03E+5	4.448E+5	5.463E+3	1.339E+4	1.303E+3	3.664E+3	ערכי S_i
1255.5	135.2	20.9	17.5	3.7	5.9	Q_i
1255.5		20.9		5.9		Q

ד. מובהקות המדגמים:
1. כדי לקבוע את מובהקות הסיגנאלים, נעשתה השוואה לאוסף של 50,000 טקסטים T_j שכל אחד מהם נוצר על ידי עירבוב אקראי של D', ונקבע ערכו של n_i, מספר המתחרים שלהם ערך $S_i^j \geq S_i$.
2. לפי n_i, חושבו מובהקויות הסיגנאלים S_i: $p_i = n_i / 50{,}000$.
התוצאות מוצגות בטבלה 4.

טבלה 4

מספר המתחרים n_i שלהם ערך $S_i^j \geq S_i$ בתחרות ובה 50,000 מתחרים.

EL2(AB)		*EL2(ABC)*		*EL2(ABCD)*		המדגם
S_1	S_2	S_1	S_2	S_1	S_2	סיגנאל
1	0	47	15	151	39	n_i
2E−5	< 2E−5	9.4E−4	3E−4	3.02E−3	7.8E−4	p_i

ה. סיכומי תוצאות:

1. המובהקות הכוללת עבור *EL2(ABC)* בתחום המורחב:

כאן חישבנו את המובהקות הכוללת באמצעות ספירת כלל הטקסטים השונים המאכלסים את 16 המקומות הראשונים בכל אחת משתי התחרויות: עבור הסיגנאלים S_1 ו- S_2. אם התחרויות היו בלתי תלויות, היינו מקבלים 16×2 טקסטים שונים כאלה (או קרוב לזה). לעומת זאת, בגלל התלות בין הסיגנאלים, כאן יש רק 25 טקסטים שונים כאלה. לכן, המובהקות הכוללת עבור *EL2(ABC)* בתחום המורחב היא:

$$\rho = 25/50{,}000 = 5\text{E}{-4}$$

2. המובהקות הכוללת עבור *EL2* בתחום המורחב:

המובהקות החזקה ביותר במבחן הטקסטים נתקבלה עבור הסיגנאל S_2 של *EL2(AB)* והיא $p_2 < 2\text{E}{-5}$. גם כאן ישנה תלות בין התוצאות, ולכן רק 5 טקסטים שונים הגיעו לראש הדירוג בשש התחרויות. על כן, המובהקות הכוללת עבור *EL2* בתחום המורחב היא:

$$\rho = 5/50{,}000 = 1\text{E}{-4}$$

ו. לסיכום:

1. הסיגנאלים של כל אחד משלושת המדגמים [א1], [ב1] ו[ג1] התחזקו בתחום המורחב.
2. הסיגנאלים של כל אחד משלושת המדגמים [א1], [ב1] ו[ג1] נמצאו מובהקים בתחום המורחב.
3. המובהקות החזקה ביותר היא בסיגנאל S_2 עבור מדגם [א1] וערכה $p_2 < 2\text{E}{-5}$.
4. המובהקות הכוללת של הסיגנאלים בתחום המורחב היא: $\rho = 1\text{E}{-4}$.

מסקנות חלק ב

א. תת-הרשימה *EL2* מצליחה בספר בראשית ונכשלת בספר "מלחמה ושלום".

ב. הסיגנאל של תת-הרשימה *EL2* מתגבר במעבר לתחום המורחב.

ג. מובהקות הסיגנאל של תת-הרשימה *EL2* בתחום המורחב חזקה מאד.

ד. בסך הכל, במעבר מתת-הרשימה *E2* לתת-הרשימה *EL2*, מתחזקת מובהקות הסיגנאלים הן בספר בראשית והן בתחום המורחב.

ה. לתת-הרשימה *EL2*, שנוצרה על ידי מומחה מטעם מבב"ק, יש חשיבות גדולה מאד: התוצאות א-ג מוכיחות כי יש כאן תופעה אמיתית של הצפנת נתונים, וכי התוצאה המקורית של ור"ר איננה תולדה של מניפולציה של שמות וכינויים.

בחלק ב טיפלנו בתת-הרשימה *EL2* כפי שהיא, דהיינו, הנתונים שלפי מבב"ק הכין שמחה עמנואל עבור אותם 32 אישים שהיו ברשימה *LIST 2* של ור"ר.

בחלק ג נתעניין בשאלה: מה היו התוצאות אם המומחה מטעם מבב"ק אכן היה פועל לפי הכללים שהוצגו לפניו, כפי שנדרש ממנו ע"י מבב"ק?

ניסוי חדש בתחום המורחב - חלק ג: מדידת נתוני עמנואל המתוקנים עבור 32 האישים המקוריים בעבודת ור"ר

מבוא

א. כתיב מנוקד

לדברי מבב"ק [6] הם הטילו על עמנואל את המשימה להרכיב רשימת נתונים חדשה לפי האילוצים של "הרשימה הראשונה"[4]. אולם, עמנואל התעלם מכלל איות יסודי שהוצג לפניו בפירוש (ואף היה יכול להסיק אותו בעצמו מהתבוננות בנתוני הרשימה הראשונה עצמה[5]). הכוונה לכלל האיות הקובע כי על השמות, הכינויים והתאריכים להכתב בכתיב מנוקד. הרציונל מאחורי כלל זה, שנקבע אפריורי לפני הרכבת הרשימה הראשונה, הוא פשוט: סביר ביותר כי ביטויים הרמוזים בטקסט המקראי, כתובים בהתאם לכללי הכתיב המקראי. הכתיב המנוקד משקף את הכתיב הנורמטיבי במקרא. על כן, יש לאיית את הנתונים, שנועדו לבדיקת הופעותיהם בדילוגים שווים בטקסט של ספר בראשית, בהתאם לכתיב המנוקד.

ב. הרשימה *LIST ECL2*

בחלק זה נעשה דבר פשוט: נתקן את נתוני עמנואל ברשימה *LIST EL2* בהתאם לכתיב המנוקד. הדבר נוגע לארבעה נתונים בלבד[6]. נסמן את הרשימה המתוקנת ב- *LIST ECL2*. להלן נסתפק בשתי מדידות: בספר בראשית ובתחום המורחב.

מדידה 1: מדידת הסיגנאל S_2 של *ECL2* בספר בראשית.

א. המדגמים

בהקבלה גמורה למדידות בחלקים א-ב, נתמקד בתת-הרשימה *ECL2*, ונגדיר עבורה את המדגמים הבאים:

מדגם [א2]: *ECL2(AB)* הוא המדגם הנוצר מתת-הרשימה *ECL2* כאשר התאריכים בה ניתנים בצמד התצורות AB.

מדגם [ב2]: *ECL2(ABC)* הוא המדגם הנוצר מתת-הרשימה *ECL2* כאשר התאריכים בה ניתנים בשלש התצורות ABC.

מדגם [ג2]: *ECL2(ABCD)* הוא המדגם הנוצר מתת-הרשימה *ECL2* כאשר התאריכים בה ניתנים בארבע התצורות ABCD.

ב. מה נמדוד:

כאן נשוב ונמדוד את מובהקות הסיגנאל S_2 בספר בראשית, בדיוק כפי שנעשה בחלק א במדידה 1.

ג. מדידת המדגמים:

1. כל אחד מן המדגמים [א2], [ב2] ו[ג2] נמדד ב-G באותו אופן בדיוק כפי שנמדד המדגם המקורי ב[1]. כך נקבע ערכו הגולמי של הסיגנאל S_2 עבור כל מדגם.
2. כדי לקבוע את מובהקות הסיגנאלים, נעשתה השוואה לאוסף של 50,000 טקסטים T_j, שכל אחד מהם נוצר על ידי עירבוב אקראי של G, דהיינו, עירבוב של מלים בתוך הפסוקים, ונקבע ערכו של n_2, מספר המתחרים שלהם ערך $S_2^j \geq S_2$.
3. לפי n_2, חושבה מובהקות הסיגנאל S_2: $p_2 = n_2 / 50{,}000$.

התוצאות מוצגות בטבלה 1.

---

[4] בלשונם:

Emanuel was shown the spelling rules and table of appellations for WRR's first list as they first appeared in WRR (1986). He then compiled a parallel table of appellations for our list of 33 rabbis, attempting to follow the rules and practices of WRR's first list.

[5] למשל, ברשימה הראשונה נמצא "איבשיץ" ולא "אייבשיץ" (הכתוב כך בטקסטים לא מנוקדים). וכן "אידלש" ולא "איידלש" וכן "מימון" ולא "מיימון".

[6] "אופנהים" במקום "אופנהיים"; "עיאש" במקום "עייאש"; "עאיאש" במקום "עאייאש"; "מגני שלמה" במקום "מגיני שלמה".

טבלה 1

מספר המתחרים n_2 שלהם ערך $S_2^j \geq S_2$ בתחרות ובה 50,000 מתחרים.

ECL2(AB)	*ECL2(ABC)*	*ECL2(ABCD)*	המדגם
5.466E+3	1.707E+3	1.247E+3	ערכי S_2
14	55	83	n_2
2.8E−4	1.1E−3	1.66E−3	p_2

יש לשים לב לכך, שגם כאן התוצאות אינן בלתי תלויות. אדרבא, קיימת תלות חזקה בין תוצאות *ELC2(AB)* לבין תוצאות *ELC2(ABC)* ו-*ELC2(ABCD)*.

ד. חישוב המובהקות הכוללת של S_2 עבור *ECL2* בספר בראשית:

כאן חישבנו את המובהקות הכוללת של S_2 בספר בראשית באמצעות ספירת כלל הטקסטים השונים המאכלסים את 15 המקומות הראשונים בכל אחת משלושת התחרויות המסוכמות בטבלה 1. בגלל התלות בין המדגמים, כאן יש רק 30 טקסטים שונים כאלה. לכן, המובהקות הכוללת של S_2 עבור *ECL2* בספר בראשית היא:

$$\rho_2 = 30/50{,}000 = 6\text{E}{-4}$$

ה. לסיכום:

1. הסיגנאל S_2 של כל אחד משלושת המדגמים [א2], [ב2] ו[ג2] נמצא מובהק בספר בראשית יותר מאשר אותו סיגנאל עבור כל אחד משלושת המדגמים [א1], [ב1] ו[ג1], בהתאמה.
2. מובהקות הסיגנאל החזקה ביותר היא עבור מדגם [א2] וערכה $p_2 = 2.8\text{E}{-4}$.
3. המובהקות הכוללת של S_2 עבור *ECL2* בספר בראשית היא: $\rho_2 = 6\text{E}{-4}$.

מדידה 2: מדידת תת-הרשימה *ECL2* בתחום המורחב.

א. מה נמדוד:

מדידה זו נעשתה בהקבלה גמורה למדידה 3 בחלקים א-ב.
כאן נבדקה תת-הרשימה *ECL2* בתחום המורחב.

ב. הטקסט:

הטקסט הוא הקטע [1, 156128] מן התורה: $D' = ([1, 156128], TORAH)$.

ג. מדידת המדגמים:

1. כל אחד מן המדגמים [א2], [ב2] ו[ג2] נמדד בקטע D' באותו אופן בדיוק כפי שנמדד המדגם המקורי ב[1]. כך נקבע ערכם הגולמי של הסיגנאלים S_1 ו- S_2 עבור כל מדגם.
2. נחשב את מדד ההגברה Q_i לגבי הסיגנאלים (ראו הסבר במבוא לחלק א, סעיף א5), וכן את מדד ההגברה המירבית $Q \equiv \max\{Q_1, Q_2\}$, ביחס לסיגנאלים המקבילים בקטע D (ספר בראשית).

התוצאות מוצגות בטבלה 2.

טבלה 2

ECL2(AB)		*ECL2(ABC)*		*ECL2(ABCD)*		המדגם
1	2	1	2	1	2	i
7.963E+5	1.057E+6	8.156E+4	6.211E+4	1.617E+4	1.277E+4	ערכי S_i
4238.5	193.4	140.9	36.4	38	10.5	Q_i
4238.5		140.9		38		Q

ד. מובהקות המדגמים:

1. כדי לקבוע את מובהקות הסיגנאלים, נעשתה השוואה לאוסף של 50,000 טקסטים T_j שכל אחד מהם נוצר על ידי עירבוב אקראי של D', ונקבע ערכו של n_i, מספר המתחרים שלהם ערך $S_i^j \geq S_i$.
2. לפי n_i, חושבו מובהקויות הסיגנאלים S_i: $p_i = n_i / 50{,}000$.

התוצאות מוצגות בטבלה 3.

טבלה 3

מספר המתחרים n_i שלהם ערך $S_i^j \geq S_i$ בתחרות ובה 50,000 מתחרים.

המדגם	*EL2(ABCD)*		*EL2(ABC)*		*EL2(AB)*	
סיגנאל	S_2	S_1	S_2	S_1	S_2	S_1
n_i	16	20	3	6	0	0
p_i	3.2E−4	4E−4	6E−5	1.2E−4	< 2E−5	< 2E−5

ה. סיכומי תוצאות:

1. המובהקות הכוללת עבור *ECL2(ABC)* בתחום המורחב:

כאן חישבנו את המובהקות הכוללת באמצעות ספירת כלל הטקסטים השונים המאכלסים את 4 המקומות הראשונים בכל אחת משתי התחרויות: עבור הסיגנאלים S_1 ו- S_2. אם התחרויות היו בלתי תלויות, היינו מקבלים 2×4 טקסטים שונים כאלה (או קרוב לזה). לעומת זאת, בגלל התלות בין הסיגנאלים, כאן יש רק 6 טקסטים שונים כאלה. לכן, המובהקות הכוללת עבור *ECL2(ABC)* בתחום המורחב היא:

$$\rho = 6/50{,}000 = 1.2\text{E}{-4}$$

2. המובהקות הכוללת עבור *ECL2* בתחום המורחב:

המובהקות החזקה ביותר במבחן הטקסטים נתקבלה עבור הסיגנאלים S_1 ו-S_2 של *ECL2(AB)* והיא $p_2 < 2\text{E}{-5}$. גם כאן ישנה תלות בין התוצאות, ולכן רק 5 טקסטים שונים הגיעו לראש הדירוג בשש התחרויות. על כן, המובהקות הכוללת עבור *EL2* בתחום המורחב היא:

$$\rho = 5/50{,}000 = 1\text{E}{-4}$$

ו. לסיכום:

1. הסיגנאלים של כל אחד משלושת המדגמים [א2], [ב2] ו[ג2] התחזקו בתחום המורחב.
2. הסיגנאלים של כל אחד משלושת המדגמים [א2], [ב2] ו[ג2] נמצאו מובהקים בתחום המורחב.
3. המובהקות החזקה ביותר היא בסיגנאלים S_1 ו-S_2 עבור מדגם [א2] וערכה $p_2 < 2\text{E}{-5}$.
4. המובהקות הכוללת של הסיגנאלים בתחום המורחב היא: $\rho = 1\text{E}{-4}$.

מסקנות חלק ג

א. בסך הכל, במעבר מתת-הרשימה *EL2* לתת-הרשימה *ECL2*, מתחזקת מובהקות הסיגנאלים בספר בראשית ובמיוחד בתחום המורחב.

ב. יש לשים לב לכך, כי כלל הכתיב המנוקד, שנקבע ע"י יעקב אורבך ז"ל עבור ניסוי בספר בראשית, מוכיח את עצמו בתחום המורחב!

הכרת תודה

החישובים נעשו באמצעות תוכנה של יעקב רוזנברג; ישֵׁרון יצחק לוי ומיכל לוי עזרו בהכנת הקטע שנוסף ל"מלחמה ושלום" - תודה להם.

ביבליוגרפיה

1. עבודת ור"ר:

Witztum, D., Rips, E. and Rosenberg, Y. (1994). Equidistant letter sequences in the Book of Genesis. Statist. Sci. 9 No. 3 429-438.

נמצא גם באתר "צופן בראשית" בקישור:

http://www.torahcode.co.il/pdf_files/pub/wrr.pdf

2. ד. ויצטום. "על הרמז בדילוג השווה: מדידה חדשה של מדגם גדולי חכמי התורה." בד"ד, חוברת 7, קיץ התשנ"ח. נמצא גם באתר "צופן בראשית" בקישור:

http://www.torahcode.co.il/pdf_files/pub/bdd.pdf

3. ד. ויצטום, סיגנאל של מערכת צפני *ELS*: דעיכה מול הגברה, התשע"ה. נמצא באתר הנ"ל בקישור:

http://www.torahcode.co.il/pdf_files/research/hagbara1.pdf

4. ד. ויצטום, סיגנאל של מערכת צפני *ELS*: ניסויים חדשים ב"תחום המורחב", התשע"ה. נמצא באתר הנ"ל בקישור:
http://www.torahcode.co.il/pdf_files/research/murchav1.pdf

5. ד. ויצטום, ניסוי חדש בתחום המורחב, התשע"ט. נמצא באתר הנ"ל בקישור:
http://www.torahcode.co.il/pdf_files/research/nisui_chadash1.pdf

6. עבודת מבב"ק:
McKay, B. D., Bar-Natan, D., Bar-Hillel, M. and Kalai, G. (1999). Solving the Bible Code puzzle. Statist. Sci. 14 No. 2 150-173.

7. "מלחמה ושלום" מאת ל. נ. טולסטוי. תרגום לעברית: לאה גולדברג (1953), ספריית פועלים, מרחביה.

8. מ. מרגליות (עורך), אנציקלופדיה לתולדות גדולי ישראל, תל-אביב, 1961.

9. המאמר:
Witztum, D., Rips, E. and Rosenberg, Y. (1986). Equidistant letter sequences in the Book of Genesis. (preprint).

10. ש. ז. הבלין, חוות דעת, התשנ"ז. נמצא באתר הנ"ל, במדור "ההתנגדות למחקר", בקישור:
http://www.torahcode.co.il/pdf_files/about/havlin1.pdf

11. ד. ויצטום, סקירה של המאמצים להפריך את מחקר הצפנים בתורה, התשנ"ט (עדכון אחרון: התשס"ט), פרק ג (והמאמרים והמסמכים המקושרים אליו). נמצא באתר הנ"ל, במדור "ההתנגדות למחקר", בקישור: http://www.torahcode.co.il/rev1_heb.htm. במיוחד ראו שם:

ד. ויצטום, הפרכה מופרכת, או: איך נכשלה רשימת הרבנים ב"מלחמה ושלום", התשנ"ח.
חלק א בקישור: http://www.torahcode.co.il/pdf_files/oppose/maindoch.pdf,
וחלק ב בקישור: http://www.torahcode.co.il/pdf_files/oppose/maindoc2.pdf,
וכן המסמכים הנלווים.

12. ד. ויצטום, על מדע ועל פרודיה: הפרכה גמורה של הטענה המרכזית של מבב"ק, התשס"א. נמצא באתר הנ"ל, במדור "ההתנגדות למחקר", בקישור: http://www.torahcode.co.il/paro_heb.htm

13. ד. ויצטום, הוכחה סטטיסטית חדשה לקיום צופן בספר בראשית, התשס"א. נמצא באתר הנ"ל בקישור: http://www.torahcode.co.il/eman_heb.htm
ראו עוד: http://www.torahcode.co.il/pdf_files/oppose/ema1.pdf