

מדידת מובהקות במחקר הצפנים בתורה: תוצאות מדידות במבחן הטקסטים

מאת דורון ויצטום

תקציר

מבחן פרמוטציות שימש למדידת המובהקות בניסוי $WRR1$ [1]. לאחר מכן, נעשה שימוש במבחן זה בניסויים רבים הקשורים למחקר הצפנים בתורה, ביניהם: ניסוי $WRR2$ [2], הניסוי על תאריכי הלידה של אישי ספר בראשית [3], ניסוי הקהילות של גאנז-ענבל-בומבך [4], ניסוי הקהילות של סיימון [5], וניסויי הקהילות של הוועדה האוניברסיטאית בראשות פרופ' אומן [6].

כאן מוצגת המובהקות המתקבלת עבור ניסויי הקהילות [4], [5] ו[6], כאשר משתמשים במבחן טקסטים. במבחן טקסטים מדידת המובהקות נעשית באמצעות השוואה לטקסטים רבים "דומים". בחלק מן המדידות נתקבלה מובהקות חזקה באופן מהותי מן המובהקות שנמדדה באמצעות מבחן הפרמוטציות.

מבוא

א. בעבודת $WRR1$ [1] בוצעו בספר בראשית מדידות של רשימות הנתונים המופיעים בו כמד"שים (מד"ש=מלה בדילוג שווה של אותיות). האפקט הכולל נמדד באמצעות סטטיסטים P_i , המקבלים ערך הקטן ככל שהמפגשים בין המד"שים יותר "מוצלחים". כדי למדוד את מובהקות התוצאות היה צריך לחזור על אותן מדידות בטקסטים רבים "דומים". דבר זה לא היה בהישג ידם של WRR לפני 25 שנים (מבחנית כמות החישובים). הוצע להם, לערוך מבחן פרמוטציות, אשר פרטיו סוכמו בין פרסי דיאקוניס וישראל אומן, ואושרו לפני עריכת הניסוי על ידי ארבעה סטטיסטיקאים נודעים נוספים. מבחן זה בוצע ב-[1]. מבחן הפרמוטציות מעריך את מובהקות הסטטיסטים P_i שנתקבלו בספר בראשית עבור רשימת הנתונים המקורית (מפגשים של שמות אישים עם תאריכי הלידה ו/או הפטירה שלהם), באמצעות השוואה לערכי הסטטיסטים P'_i המתקבלים בספר בראשית עבור הרבה רשימות נתונים "מעורבבים" (מפגשים של אותם שמות אישים עם תאריכי הלידה ו/או הפטירה של אישים אחרים מן הרשימה המקורית).

ב. כחמש וחצי שנים לאחר ביצוע מבחן הפרמוטציות המתואר ב[1], סבר אליהו ריפס שהתקדמות המהירה של אמצעי החישוב מאפשרת לשוב לרעיון הפשוט ביותר של מדידת המובהקות: מדידה באמצעות השוואה לטקסטים רבים "דומים". במכתב [7] לדוד קשדן הציע אליהו ריפס את העקרונות הבאים לגבי מדידת מובהקות של רשימה בספר בראשית:

- להשתמש ב- 1000 (או יותר) טקסטים T_j , כל T_j נוצר ע"י ערבוב אקראי של מלים בתוך פסוקים (הדבר נעשה בדיוק כפי שנעשה ב[1] לגבי טקסט U).
 - להעריך את מובהקות הסטטיסטים P_i בספר בראשית, באמצעות השוואתם לערכי P'_i המתקבלים בטקסטים T_j .
- מבחן מסוג זה יכול להיות "מבחן טקסטים".

ג. השוואה בין מבחן הפרמוטציות למבחן הטקסטים:

כל אחד משני המבחנים מודד דבר שונה:

- מבחן הפרמוטציות מודד עד כמה "מוצלחים" המפגשים של שמות האישים עם תאריכי הלידה ו/או הפטירה שלהם יותר מן המפגשים של אותם שמות אישים עם תאריכים אחרים (שהם תאריכי לידה/פטירה של אישים אחרים מאותה רשימת נתונים).

- מבחן הטקסטים מודד מה הסיכוי לקבל במקרה P_i כה קטנים. דהיינו: מה הסיכוי שהצטברו במקרה כל כך הרבה מפגשים "מוצלחים".

במלים אחרות: מבחן הטקסטים מודד את מובהקות הכמות המצטברת (אקומולציה) של מפגשים "מוצלחים", ואילו מבחן הפרמוטציות מודד את המובהקות של "ספציפיות הצופן".

לכן אנו מצפים אפריורי להבדל בתוצאות שני המבחנים עבור אותה רשימת נתונים, שהרי הם מודדים דברים שונים.

דוגמא:

רשימה הנתונים שנמדדה בעבודת *MBBK* [8] בספר "מלחמה ושלוה" [9] הוכנה על ידי מניפולציה מוצהרת. המניפולציה נעשתה כדי להשיג מובהקות חזקה במבחן הפרמוטציות. ואכן, הושגה מובהקות חזקה מאד עבור הסטטיסטי הנמדד:

$$p = 7E-7.$$

לעומת זאת, כאשר מודדים את המובהקות של אותו סטטיסטי במבחן טקסטים, מתברר כי ב-12 מתוך 10,000 טקסטים "דומים" (שנוצרו מעירבוב אקראי של הטקסט המקורי) – הושגה תוצאה חזקה יותר. לכן, המובהקות חלשה בהרבה:

$$p = 1.2E-3.$$

אכן, קיים הבדל גדול מאד בין התוצאות.

ד. המבחן המתאים למצבו העכשווי של מחקר הצפנים בתורה:

המחקר המדעי של הצפנים בתורה, עד כמה שבא לידי ביטוי במדידת רשימות גדולות של נתונים (כגון אלה שבעבודות [1]-[4]), עסק עד כה בהוכחת קיומו של צופן מסויים. הוא לא עסק באיפיון מדויק או בפיענוח של הצפנים. הדבר הוצהר במפורש, בהקדמת *WRR1* [1] (בעמוד הראשון של המאמר. בעניין זה בהרחבה ראה [10]).

לפי הניתוח בסעיף הקודם ברור, כי לעבודות המצויינות כאן, מתאים דווקא מבחן הטקסטים, ולא מבחן פרמוטציות המודד את המובהקות של "ספציפיות הצופן". (*WRR* השתמשו במבחן פרמוטציות רק באין ברירה, כמבואר בסעיף א לעיל).

ה. מלבד השיקולים הכלליים המורים על הבדל הצפוי אפריורית בין המובהקויות המתקבלות בשתי השיטות, ייתכנו גם סיבות ספציפיות להבדל, עבור סוגים שונים של רשימות נתונים. ניתוח הסיבות הספציפיות, ייעשה במאמר נפרד.

הבדיקות

במחקר המדעי של הצפנים בתורה נעשה מעקב אחר שתי תכונות:

תכונה א: נטייה למפגש בין מד"שים המייצגים "ביטוי א", לבין מד"שים המייצגים "ביטוי ב", כאשר קיים קשר מושגי מובהק בין "ביטוי א" ל"ביטוי ב".

תכונה ב: נטייה למפגש בין מד"שים המייצגים "ביטוי א", לבין "ביטוי ב" הנמצא כרצף אותיות בטקסט עצמו, כאשר קיים קשר מושגי מובהק בין "ביטוי א" ל"ביטוי ב".

במאמר זה, ידווחו תוצאות בדיקות של עבודות המודדות את תכונה א. תוצאות הבדיקות של עבודות המודדות את תכונה ב – יפורסמו במאמר נפרד.

בכל הבדיקות בהן השתמשתי במבחן טקסטים, נעשה שימוש מירבי ב- 50,000 טקסטים (לפי יכולות החישוב העומדות לרשותי, אני זקוק למספר חודשי עבודת מחשב, על מנת לערוך השוואה ב-50,000 טקסטים, עבור רשימת נתונים מסדר גודל של אלו המצויות בעבודות המדווחות כאן). לכן ערכתי בדיקות כדלקמן רק עבור מדגמים, שהמובהקות שלהם במבחן הפרמוטציות תואמת את יכולת הרזולוציה שלנו. כלומר, נבדקו מדגמים ו- P_i כאלה, שבמבחן הפרמוטציות נתקבל עבורם $p > 0.00002$.

בדיקת עבודות המודדות מפגשים בין מד"שים בספר בראשית.

מבוא

כאן יוצגו תוצאות מבחן הטקסטים עבור ניסוי הקהילות של גאנז-ענבל-בומבך [4], ניסוי הקהילות של סיימון [5], וניסוי הקהילות ה"טרי" (fresh) של הוועדה האוניברסיטאית בראשות פרופ' אומן [6]. ניסויי הקהילות עוסקים במדידת מפגשים בין מד"שים המייצגים שמות אישים וכנוייהם לבין מד"שים המייצגים את שמות המקומות/קהילות בהן חיו.

המשותף לעבודות הנ"ל הוא קבוצת האישים, ורשימת שמותיהם וכינוייהם. רשימה זו נוצרה מצירוף הרשימה הראשונה והרשימה השניה של שמות וכינויים, שנוצרו על ידי ש. ז. הבלין ופורסמו על ידי *WRI* [1].

המבדיל בין העבודות הללו הוא:

1. הגדרת היחס בין האיש לבין הקהילה.
 - בניסוי של גאנז-ענבל-בומבך – צוינו הקהילות בהן נולד ו/או נפטר האיש.
 - בניסוי של סיימון – צוינו הקהילות בהן נולד, למד, חי, ביקר ו/או נפטר האיש.
 - בניסוי ה"טרי" של הוועדה – צוינו מקומות הלידה ו/או הפטירה וכן "מקומות נוספים בהם שהתה הדמות במהלך חייה." יש לציין שבכמה מקרים צויין גם מקום קבורה.
2. סוג המקום.
 - בניסוי של גאנז-ענבל-בומבך – צוינו בדרך כלל ערים ולעתים כפרים.
 - בניסוי של סיימון – צוינו ערים, כפרים, חבלי ארץ ולפעמים ארצות.
 - בניסוי ה"טרי" של הוועדה – צוינו בדרך כלל ערים ולעתים של כפרים.
3. שם המקום.
 - בניסוי של גאנז-ענבל-בומבך – צוינו שמות הקהילות כפי שנקראו בפי היהודים.
 - בניסוי של סיימון – צוינו שמות מקומות של ערים, כפרים, חבלי ארץ ולפעמים שמות ארצות.
 - בניסוי ה"טרי" של הוועדה – צוינו בדרך כלל שמות של ערים ולעתים של כפרים.
4. תעתיק שם המקום.
 - בניסוי של גאנז-ענבל-בומבך – תעתיק השמות לעברית נעשה בהתאם לכללים מקובלים שצוינו בעבודת *WRI*.
 - בניסוי של סיימון – תעתיק השמות נעשה ללא כללים והם נכתבו כפי שהועתקו ממאמרים של מחברים שונים באנציקלופדיה (לעתים עם טעויות).
 - בניסוי ה"טרי" של הוועדה – התעתיק ללא כללים.
5. איסוף הנתונים.
 - בניסוי של גאנז-ענבל-בומבך – איסוף הנתונים נעשה באמצעות אלגוריתם מסודר המאפשר שקיפות ומעקב מלא אחרי כל ההחלטות.
 - בניסוי של סיימון – איסוף הנתונים נעשה תוך חריגות רבות מאד מפרוטוקול הניסוי [11].
 - בניסוי ה"טרי" של הוועדה – איסוף הנתונים נעשה תוך סטיות מפרוטוקול הניסוי. אין אפשרות לעקוב אחר ההחלטות [12].

הערה: כל המדידות דלקמן נעשו לפי הסטטיסטי P_2 (הסיבה לכך: ראה בנספח).

א. ניסוי הקהילות של גאנז-ענבל-בומבך:

1. תיאור קצר:

בניסוי זה שני מרכיבים:

- (1) התיזה. התיזה כוללת את ההגדרה המדויקת של הנושא (שנעשתה במקביל ל-*WRI*): כגון, לבדוק את שמות האישים עם שמות הקהילות בהן נולדו או נפטרו, שיש להשתמש בשמות יהודיים (אם קיימים) ולא בשמות המקומות הלועזיים, להשתמש בווריאנטים העיקריים, וכי האיות ייעשה לפי כללים קבועים.

(2) איסוף הנתונים. איסוף הנתונים נקבע על ידי אלגוריתם, דהיינו מנגנון מיכני המבוסס על מספר סביר של כללים ענייניים והגיוניים, תוך שימוש במאגרי המידע הבאים: אנציקלופדיה מרגליות [13], האנציקלופדיה העברית [14] וקונקורדנציה לתנ"ך [15].
(כל הפרטים וההסברים לגבי סעיפים (1) ו-(2) נמצאים ב[4] וב[16]).

נסמן:

$Gns0 =$ רשימת שמות הקהילות.

$Gns0 + Gns0 = Gns01$ כל השמות ב- $Gns0$ עם קידומת "קהל" או "קהלת".

סכמת מדידת המפגשים של WRR מתאימה למד"שים בני 5-8 אותיות [1]. להלן נמשיך להשתמש בסימוני ארבע הרשימות $Gns0$ ו- $Gns01$, גם כאשר הכוונה היא לרשימות הביטויים בני 5-8 אותיות מתוכן.

2. המדידה:

(א) בשלב ראשון נמדדו המפגשים בין שמות האישים וכינוייהם המופיעים בספר בראשית כמד"שים, לבין שמות המקומות המופיעים אף הם כמד"שים, ונקבעו ערכי P_2 עבור שתי הרשימות הנ"ל.
(ב) בשלב שני נערך מבחן טקסטים בו נמדדו ערכי P'_2 ב-50,000 טקסטים, עבור כל אחת משתי הרשימות. במקביל נערך מבחן פרמוטציות בו נמדדו ערכי P'_2 ב-50,000 פרמוטציות המצמידות את שמות האישים וכינוייהם לשמות המקומות באופן אקראי. גם מבחן זה נערך עבור שתי הרשימות. התוצאות מובאות בטבלה 1.

טבלה 1

מספר המתחרים n שלהם ערך $P'_2 \leq P_2$ בתחרות ובה 50,000 מתחרים.

Gns0	Gns01	רשימה
1.18E-6	4.39E-10	ערכי P_2
10	0	מבחן טקסטים
249	5	מבחן פרמוטציות

ב. הניסוי "הטרי" (" F ") של הוועדה:

1. תיאור קצר:

הפרוטוקול לניסוי F נכתב על ידי מתנגדי מחקר הצופן בתורה. רשימת המקומות הוכנה על ידי מומחה, רמי ריינר, וכללה מקומות לידה ו/או פטירה וכן "מקומות נוספים בהם שהתה הדמות במהלך חייה". יש לציין שבכמה מקרים צויין גם מקום קבורה. נציג הוועדה סימן את שמות המקומות הכלולים בניסוי. על הליקויים בעבודה זו ראה ב[12]. [כאן רק נעיר, כי רמי ריינר טעה טעות מצערת בעניין גיאוגרפי גרידא, כאשר סבר בטעות כי טודלה (שם עיר קטנה במחוז נווארה, ספרד) הוא שם אלטרנטיבי לטולידו בירת קסטיליה, שהיתה בירת ממלכת ספרד כחמש מאות שנה. לכן, הוא הצמיד לאישים הקשורים לטולידו גם את השם טודלה, ולהיפך. הדבר דומה למי שיטען, כי דיזנגוף היה ראש העיר הראשון של תל עדשים (במקום: תל אביב). כמובן, מחקנו נתון שגוי זה, שאינו שנוי במחלוקת. כל שאר השמות שסומנו על ידי נציג הוועדה נלקחו כפי שהם.]

לפי הפרוטוקול, המומחה היה צריך לבחור קידומות המתאימות לרשימת המקומות, כפי שנעשה בניסוי הקהילות של גאנז-ענבל-בומבך (שם נבחרו הקידומות "קהל" ו"קהלת"). הדבר לא נעשה. יו"ר הוועדה החליט אד-הוק, תוך חריגה מן הפרוטוקול (ראה [12]), לערוך את הניסוי ללא קידומות. בניסוי שלפנינו, נעשה שימוש בקידומות "קהל" ו"קהלת" הנפוצים מאד בתיאור שמות קהילות יהודיות (ראה [17]). מלבד זאת נעשה שימוש בשתי קידומות, "עדת" ו"ק"ק", שהוצעו בדיוני הוועדה על ידי חבריה (והתנגדו להן, מסיבות ענייניות, מצדדי מחקר הצופן בתורה אשר בוועדה).

נסמן:

$Fr0 =$ רשימת שמות המקומות F (ראה הערה בפיסקה הקודמת).

$Fr0 + Fr0 = Fr01$ כל השמות ב- $Fr0$ עם קידומת "קהל" או "קהלת".

$Fr0 + Fr0 = Fr02$ כל השמות ב- $Fr0$ עם קידומת "קהל" או "קהלת" או "עדת".

$Fr0 = Fr03 +$ כל השמות ב- $Fr0$ עם קידומת "קהל" או "קהלת" או "עדת" או "ק"ק".
 סכמת מדידת המפגשים של WRR מתאימה למד"שים בני 5-8 אותיות. להלן נמשיך להשתמש
 בסימוני ארבע הרשימות $Fr0$, $Fr01$, $Fr02$ או $Fr03$, גם כאשר הכוונה היא לרשימות
 הביטויים בני 5-8 אותיות מתוכן.

2. המדידה:

(א) בשלב ראשון נמדדו המפגשים בין שמות האישים וכינוייהם המופיעים בספר בראשית כמד"שים,
 לבין שמות המקומות המופיעים אף הם כמד"שים, ונקבעו ערכי P_2 עבור ארבע הרשימות הנ"ל.
 (ב) בשלב שני נערך מבחן טקסטים בו נמדדו ערכי P'_2 ב-50,000 טקסטים, עבור שלוש מהרשימות
 הנ"ל. במקביל נערך מבחן פרמוטציות בו נמדדו ערכי P'_2 ב-50,000 פרמוטציות המצמידות את שמות
 האישים וכינוייהם לשמות המקומות באופן אקראי. מבחן זה נערך עבור ארבע הרשימות.
 התוצאות מובאות בטבלה 2.

טבלה 2

מספר המתחרים n שלהם ערך $P'_2 \leq P_2$ בתחרות ובה 50,000 מתחרים.

Fr0	Fr01	Fr02	Fr03	רשימה
6.91E-4	1.699E-5	9.78E-7	1.15E-6	ערכי P_2
*	27	3	15	מבחן טקסטים
16,084	7,647	12,371	15,602	מבחן פרמוטציות

(* מדידה זו לא בוצעה מתוך הערכה כי במקרה זה $n > 100$).

ג. ניסוי הקהילות של סיימון:

1. תיאור קצר:

ניסוי זה נעשה על ידי מתנגדי מחקר הצופן בתורה. הפרוטוקול לניסוי נכתב על ידי סיימון שהוא מתנגד
 נמרץ למחקר זה, ואיסוף הנתונים נעשה תחת פיקוחם של גינדיס ומקי – מתנגדים מושבעים של מחקר
 זה. בניסוי של סיימון צוינו המקומות בהן נולד, למד, חי, ביקר ו/או נפטר האיש. הנתונים הועתקו כפי
 שהם מאנציקלופדיה מרגליות (להלן, א"מ [13]). על הליקויים בהכנת הפרוטוקול ובאיסוף הנתונים
 בניסוי של סיימון ראה ב[11].

נסמן:

$Sim5_8$ = רשימת שמות המקומות של סיימון, כפי שהועתקו לדבריו על ידי עוזריו מא"מ.
 (נלקחו על ידו רק שמות המקומות שהם בני 5-8 אותיות, כדי להתאים לסכימת המדידה של WRR .)

$Sim2_4$ = רשימת המקומות בא"מ ששמותיהם בני 2-4 אותיות [18].

$Sim5_8 + Sim2_4 = Sim0$

$Sim0 = Sim01 +$ כל השמות ב- $Sim0$ עם קידומת "קהל" או "קהלת".

$Sim0 = Sim02 +$ כל השמות ב- $Sim0$ עם קידומת "קהל" או "קהלת" או "עדת".

$Sim0 = Sim03 +$ כל השמות ב- $Sim0$ עם קידומת "קהל" או "קהלת" או "עדת" או "ק"ק".

כאמור, סכמת מדידת המפגשים של WRR מתאימה למד"שים בני 5-8 אותיות. להלן נמשיך להשתמש
 בסימוני ארבע הרשימות $Sim0$, $Sim01$, $Sim02$ או $Sim03$, גם כאשר הכוונה היא לרשימות
 הביטויים בני 5-8 אותיות מתוכן.

2. המדידה:

(א) בשלב ראשון נמדדו המפגשים בין שמות האישים וכינוייהם המופיעים בספר בראשית כמד"שים,
 לבין שמות המקומות המופיעים אף הם כמד"שים, ונקבעו ערכי P_2 עבור ארבע הרשימות הנ"ל.
 (ב) בשלב שני נערך מבחן טקסטים בו נמדדו ערכי P'_2 ב-50,000 טקסטים, עבור כל אחת מארבע
 הרשימות. במקביל נערך מבחן פרמוטציות בו נמדדו ערכי P'_2 ב-50,000 פרמוטציות המצמידות את
 שמות האישים וכינוייהם לשמות המקומות באופן אקראי. גם מבחן זה נערך עבור ארבע הרשימות.
 התוצאות מובאות בטבלה 3.

טבלה 3

מספר המתחרים n שלהם ערך $P_2 \leq P_2'$ בתחרות ובה 50,000 מתחרים.

Sim0	Sim01	Sim02	Sim03	רשימה
2.58E-5	3.216E-7	2.1E-8	1.26E-7	ערכי P_2
33	8	3	6	מבחן טקסטים
15,211	10,099	10,280	10,965	מבחן פרמוטציות

מסקנות

א. יש הבדל מהותי ודרמטי בין התוצאות של מבחן הפרמוטציות לאלה של מבחן הטקסטים.
 ב. מתברר ממבחן הטקסטים, כי הקורלציה בין מד"שים של שמות האישים לבין מד"שים של שמות הקהילות אליהן היו קשורים, שרירה וקיימת בספר בראשית, בין לגבי נתוני גאנז-ענבל-בומבך, ובין לגבי נתוני רשימת סיימון או נתוני רשימה F .
 במלים אחרות: גם אם היינו משתמשים בנתוני רשימת סיימון או נתוני רשימה F , היינו מזהים הצפנה של הנתונים כמד"שים בספר בראשית.

נספח

הסטטיסטי ששימש למדידת מדגמי הקהילות:

בניסוי המקורי של $WRR1$ ([1], עמ' 436) הוגדרו ארבע "מידות כוללות לקירבה" P_1, P_2, P_3 ו- P_4 .
 שם ([1], עמ' 431) מבואר כי:

- מידות P_1 ו- P_2 הוגדרו כשני סטטיסטים שונים לגבי הרשימה השלמה של השמות והכינויים (להלן $LIST 2$).
- כאשר מידות P_1 ו- P_2 מיושמות לגבי תת-רשימה של $LIST 2$ (להלן $L2$), הן מסומנות כ- P_3 ו- P_4 .

במלים אחרות: ישנם שני אופראטורים, P_1 ו- P_2 .
 היישום שלהם על רשימת הנתונים השלמה נותן את P_1 ו- P_2 :

$$P_1 \equiv P_1(LIST 2), P_2 \equiv P_2(LIST 2)$$

והיישום שלהם על תת-הרשימה $L2$ נותן את P_3 ו- P_4 :

$$P_3 \equiv P_1(L2) \quad P_4 \equiv P_2(L2)$$

בניסויי הקהילות בחרו כולם: גאנז, סיימון, וחברי הוועדה, להשתמש באופראטור P_2 .
 לכן נעשה כאן שימוש באופראטור זה בלבד.

הכרת תודה

החשובים נעשו באמצעות תוכנה של יעקב רוזנברג. תודה למי שסייע בתיגבור אמצעי החישוב שנדרשו לעבודות אלה – והפץ בעילום שמו.

מקורות והערות

1. עבודת WRR :

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

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

2. המאמר:

Witztum, D., Rips, E. and Rosenberg, Y. (1995). *Equidistant Letter Sequences in the Book of Genesis: II. The Relation to the Text*, a preprint.

3. הניסוי על תאריכי הלידה של אישי ספר בראשית:
Witztum, D. (2006). *The Hidden Birth Dates of Personalities of Genesis*, Proceedings of the 18th International Conference on Pattern Recognition, August 2006.
Available at: http://www.torahcode.co.il/pdf_files/pub/leida1.pdf.

4. המאמר:
H. Gans, Z. Inbal and N. Bombach, *Patterns of Equidistant Letter Sequence Pairs in Genesis*, Proceedings of the 18th International Conference on Pattern Recognition, August 2006. Available at: http://www.torahcode.co.il/pdf_files/pub/gans.pdf.

5. החומר כולל מכתב:
Simon, B (1998). *An open letter to Harold Gans*.
http://www.torahcode.co.il/pdf_files/oppose/openletter.pdf.

רשימת שמות המקומות:
Simon, B (1998). *Raw Data: The Yeshiva Students' List*.
http://www.torahcode.co.il/pdf_files/oppose/Margcities.pdf.

מאמר:
Simon, B (1998). *Barry Simon's Cities Experiment – An Overview*.
http://www.torahcode.co.il/pdf_files/oppose/BarryS.pdf.

6. דו"ח הוועדה האוניברסיטאית בראשות פרופ' אומן:
Aumann, R. J., Furstenberg, H, (2004) *Findings of the Committee To Investigate the Gans-Inbal Results on Equidistant Letter Sequences in Genesis*. Discussion Paper #364, Center for the Study of Rationality, The Hebrew University of Jerusalem.
Available at: http://www.torahcode.co.il/pdf_files/oppose/dp364.pdf.

7. בדואר אלקטרוני מיום 15 במאי שעה 16:40:13, שנת 1997 (למנינם).

8. עבודת MBBK:
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.

9. ל. נ. טולסטוי. *מלחמה ושלוש*. תרגום לעברית: לאה גולדברג (1953), ספריית פועלים, מרחביה.

10. ד. ויצטום (התשס"ד). *צופן בראשית*, פרק ג. ראה גם: תיאור המחקר המדעי של הצופן בתורה, נמצא באתר *צופן בראשית* בכתובת: http://www.torahcode.co.il/description_heb.htm.

11. ד. ויצטום (התשס"א). על ניסוי הערים של בארי סיימון. נמצא באתר *צופן בראשית* בכתובת: http://www.torahcode.co.il/pdf_files/oppose/sim.pdf.

12. ד. ויצטום (התשס"ג). *הערות על הליקויים בניסוי F*. נמצא בעמוד 150 של Discussion Paper #365, ובכתובת: http://www.torahcode.co.il/pdf_files/oppose/dp365.pdf.

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

14. *האנציקלופדיה העברית*. חברה להוצאת אנציקלופדיות בע"מ, ירושלים, התשמ"א.

15. א. אבן שושן (עורך), *קונקורדנציה חדשה לתורה נביאים וכתובים*, קרית ספר, ירושלים, התשמ"א.

16. תיאור התיזה ואיסוף הנתונים בניסוי של גאנז נמצא בכתובת:
http://www.torahcode.co.il/pdf_files/pub/communities_data.pdf

בעברית: בספר *צופן בראשית* (מס' 10 לעיל), נספח 13, וכן באתר *צופן בראשית* בכתובת:
http://www.torahcode.co.il/pdf_files/research/kehila2.pdf.

17. ראה במבואות למקורות המובאים במס' 16.

18. רשימת המקומות בא"מ ששמותיהם בני 4-2 אותיות, נמצאת בכתובת:
http://www.torahcode.co.il/pdf_files/oppose/makom2_4.pdf.