

נספח א 15

מדגם "הלידה"

כאן יוצג המאמר על "מדגם הלידה", שטייטש ראשונה שלו פורסמה באינטרנט בחודש טבת התשנ"ט (ינואר 1999 למניינם).

מפגשים בין שמות אישי ספר בראשית

לתאריכי הלידה שלהם

מאת: דורון ויצטום

תקציר

בספרות חז"ל ישנו מידע לגבי תאריכי הלידה של אישים מרכזיים בספר בראשית. לפי מידע זה הוכנו שני מדגמים, לצורך רפליקציה של הניסוי המתואר במאמרנו [1]:

Equidistant Letter Sequences in the Book of Genesis.

מדידת המפגשים בין השמות לתאריכים נעשתה כמו במאמרנו [2]:

Equidistant Letter Sequences in the Book of Genesis: II. The Relation to the Text.

המובהקות נמדדה ונמצאה חזקה עבור שני המדגמים.

מבוא

מאמרנו *Equidistant Letter Sequences in the Book of Genesis* [1] שהתפרסם ב-*Statistical Science* עסק במתאם שבין שמותיהם של גדולי ישראל לבין תאריכי הלידה והפטירה שלהם, כל זאת בדילוגים שווים של אותיות בספר בראשית. גדולי ישראל אלה חיו אלפי שנים לאחר כתיבת ספר בראשית.

לאחרונה חיפשתי בספרות חז"ל מידע לגבי תאריכי הלידה והפטירה של אישי ספר בראשית עצמם, על מנת למדוד את המפגשים בין שמות האישים לבין התאריכים. מתברר כי אכן קיים מידע כזה, אשר שימש בסיס למדידות שיתוארו להלן: מדידות שהצביעו על מובהקות חזקה של המפגשים הנ"ל.

א. איסוף המידע

המידע לגבי תאריכי הלידה והפטירה של אישים מספר בראשית עשוי להימצא בכמה מקורות בספרות חז"ל; יתר-על-כן, אף ייתכן שישנם מקורות החלוקים זה על זה.

בעבודה קודמת [2] במקרה דומה (לגבי תרגום יונתן) נהגנו כך:

א. בהינתן מקורות אחדים (לעתים סותרים) בספרות חז"ל לגבי הנתונים הנדרשים, בחרנו במקור המספק את מירב הנתונים.

ב. השתמשנו במהדורה המתוקנת של המקור (על פי כתבי יד וכד') [3].

גם במקרה זה פעלתי כך:

א. כדי לאתר את המקור המקיף ביותר בספרות חז"ל לגבי המידע המבוקש, ערכתי סקר ממוחשב באמצעות תקליטור פרויקט השו"ת של "בר אילן" [4]. סרקתי את כל מקורות חז"ל המצויים בו: משנה, תוספתא, מסכתות קטנות, תלמוד בבלי, תלמוד ירושלמי, מדרשי הלכה, מדרשי אגדה וספר הזוהר.

חיפשתי אחר כל המאמרים המצוינים במפורש את תאריך לידתו (או בריאתו) או פטירתו של אחד מאישי ספר בראשית. בדרך זו נמצאו תאריכים של כמה אישים מספר בראשית - כולם תאריכי לידה (לא נמצאו מאמרים המציינים במפורש תאריכי פטירה). נמצא כי:

1. בכמה מקורות נמצא תאריך הלידה של אישיות אחת בלבד: במדרש ויקרא רבה,

בפסיקתא דרב כהנא ובזוהר נמצא תאריך בריאת האדם.

בתלמוד ירושלמי ובמדרש תנחומא. נמצא תאריך לידת יצחק.

2. ב"אוצר מדרשים" נמצאו תאריכי הלידה של עשרה אישים (עשרה מבני יעקב).

3. בילקוט שמעוני נמצאו תאריכי הלידה של שלושה עשר אישים (אדם, יצחק

ואחד עשר מבני יעקב).

נמצא שבמקרה שלפנינו ילקוט שמעוני הוא המקור המקיף מספר רב יותר של נתונים.

ב. בדקתי במהדורה המדעית של ילקוט שמעוני [5], ושם נמצאו ארבעה עשר תאריכים: של אדם הראשון, של יצחק ושל כל שנים עשר בני יעקב:

1. תאריך בריאת האדם מופיע בפרשת פינחס רמז תשפב.

2. תאריך לידת יצחק מופיע בפרשת שמות רמז קעב, ובפרשת בא רמז רי.

3. רשימת תאריכי הלידה של בני יעקב מופיעה בפרשת שמות רמז קסב. (רשימה זהה

נמצאת בביאור רבנו בחיי על התורה. עיינו בהערה [6].)

התאריכים ניתנים בטבלה 1 לקמן.

ב. מה נמדוד

במאמרינו [1] ו[2] וגם במאמרנו: "צפן חברי בדילוג שווה בספר בראשית: מובהקות סטטיסטית של התופעה" [7], הצבענו על שתי תופעות המתרחשות בספר בראשית:

אם עוקבים אחרי הופעות של צפני ELS בספר בראשית, אזי

(א) קיימת נטייה למפגש בין ההופעות הללו, כאשר יש קשר מושגי מובהק בין המלים או הביטויים שהן מייצגות.

(ב) קיימת נטייה למפגש בין ההופעות הללו לבין מלים וביטויים בטקסט (הנמצאים בו כרצף אותיות, דהיינו כאילו ב"דילוג" $d = \pm 1$), כאשר יש קשר מושגי כנ"ל.

במאמרנו [1] טיפלנו במפגשים בין שמות אישים לבין תאריכיהם לפי (א). ברשימה הנוכחית, בה האישים הם מן הדמויות המרכזיות בספר בראשית, הבחירה הטבעית היא לבדוק מפגשים בין שמותיהם לבין תאריכיהם לפי (ב), דהיינו: למדוד מפגשים בין השמות המופיעים בטקסט כרצף של אותיות, לבין התאריכים כצפני ELS.

ב[2] עמדנו בפני מצב דומה לגבי שמות צאצאי נח, גם במקרים אלו הוחלט לבדוק את המפגשים עם השמות לפי (ב).

ג. המדגמים

1. מדגם א:

המדגם מורכב מזוגות ביטויים: בכל זוג – הביטוי האחד הוא שם אחד משמות האישים, והביטוי השני הוא תאריך לידתו.

- כל שם נלקח לפי כל הצורות שהוא נכתב בספר בראשית.
 - כל פרטי כתיבת התאריכים הם כמו ב[1], דהיינו, כל תאריך נכתב בשלוש צורות קבועות. למשל: א' תשרי, בא' תשרי א' בתשרי.
- בטבלה 1 נתונים השמות לצורותיהם (לפי סדר הופעתם בבראשית) והתאריכים לצורותיהם.

טבלה 1

השם	התאריך
אדם	א' תשרי, בא' תשרי, א' בתשרי.
יצחק	ט"ו ניסן, בט"ו ניסן, ט"ו בניסן, י"ה ניסן, בי"ה ניסן, י"ה בניסן.
ראובן	י"ד כסלו, בי"ד כסלו, י"ד בכסלו.
שמעון	כ"א טבת, כ"א טבת, כ"א בטבת.
לוי	ט"ז ניסן, בט"ז ניסן, ט"ז בניסן, י"ו ניסן, בי"ו ניסן, י"ו בניסן.
יהודה	ט"ו סיון, בט"ו סיון, ט"ו בסיון, י"ה סיון, בי"ה סיון, י"ה בסיון.
דן	ט' אלול, בטי' אלול, ט' באלול.
נפתלי	ה' תשרי, בח' תשרי, ה' בתשרי.
גד	י' חשוון, בי' חשוון, י' בחשוון.
אשר	כ" שבט*, בכ" שבט, כ" בשבט.
יששכר	י' אב*, בי' אב*, י' באב*.
זבולון, זבולן	ז' תשרי, בז' תשרי, ז' בתשרי.
יוסף, צפנת פענח	א' תמוז, בא' תמוז, א' בתמוז.
בן אוני, בנימין, בנימן	י"א חשוון, בי"א חשוון, י"א בחשוון.

זו טבלה מקוצרת למדגם א: זוגות המדגם הם זוגות ביטויים, בהם הביטוי הראשון הוא אחד השמות, והביטוי השני הוא אחת מצורות התאריך הרשומה בשורה שבה מופיע השם. כאמור לעיל, המדגם נועד למדידת המפגשים בין השמות המופיעים בטקסט כרצף של אותיות לבין התאריכים בדילוג השווה. לכן, כפי שמבואר ב[2], אין הגבלה לאורך השמות, בעוד המדידה מקיפה רק צורות תאריכים שבהן 5-8 אותיות. צורות התאריכים שיש בהן פחות אותיות סומנו בכוכבית ואינן נכללות במדגם.

2. מדגם הלידה:

כיון שמדובר בתאריכי לידה בלבד, יתכן כי נכון יותר לבדוק את מפגשיהם עם שמות האישים, כפי שהם כתובים בסיפור לידתם בספר בראשית. כל שאר הפרטים זהים למדגם א. כך מתקבל "מדגם הלידה" המתואר בצורה מקוצרת בטבלה 2 (והשונה רק בשלושה שמות מטבלה 1).

טבלה 2

השם	התאריך
אדם	א' תשרי, בא' תשרי, א' בתשרי.
יצחק	ט"ו ניסן, בט"ו ניסן, ט"ו בניסן, י'ה ניסן, בי'ה ניסן, י'ה בניסן.
ראובן	י"ד כסלו, בי"ד כסלו, י"ד בכסלו.
שמעון	כ"א טבת, בכ"א טבת, כ"א בטבת.
לוי	ט"ז ניסן, בט"ז ניסן, ט"ז בניסן, י'ו ניסן, בי'ו ניסן, י'ו בניסן.
יהודה	ט"ו סיון, בט"ו סיון, ט"ו בסיון, י'ה סיון, בי'ה סיון, י'ה בסיון.
דן	ט' אלול, בט' אלול, ט' באלול.
נפתלי	ה' תשרי, בה' תשרי, ה' בתשרי.
גד	י' חשוון, בי' חשוון, י' בחשוון.
אשר	כ' שבט*, בכ' שבט, כ' בשבט.
יששכר	י' אב*, בי' אב*, י' באב*.
זבלון	ז' תשרי, בז' תשרי, ז' בתשרי.
יוסף	א' תמוז, בא' תמוז, א' בתמוז.
בן אוני, בנימין	י"א חשוון, בי"א חשוון, י"א בחשוון.

גם כאן, צורות התאריכים שיש בהן פחות מ-5 אותיות סומנו בכוכבית ואינן נכללות במדגם.

ד. מדידת ערכי $c(w, w')$

1. מדידת המפגשים במדגם א:

מדידת כל המפגשים במדגם א נעשתה כפי שמתואר ב[2]: כאשר הביטוי "הראשון" (w), שהוא השם במקרה שלנו) נלקח כרצף של אותיות, והביטוי "השני" (w'), שהוא התאריך במקרה שלנו) נלקח בדילוגים השווים. עורכים תחרות בין הופעות הביטוי "השני" בדילוגים השווים לבין הופעותיו בדילוגים המשובשים: תחרות ליצירת המפגשים המוצלחים יותר עם הביטוי "הראשון".

כך התקבלו 48 ערכי $c(w, w')$ עבור 48 הזוגות בהם משתתפים תאריכים המופיעים בדילוג

השווה. ערכי P_1 ו- P_2 חושבו והם:

$$P_1 = 1.61 \times 10^{-4} \quad P_2 = 7.07 \times 10^{-3}$$

2. מדידת המפגשים במדגם הלידה:

יש 39 ערכי $c(w, w')$. ערכי P_1 ו- P_2 הם:

$$P_1 = 1.336 \times 10^{-5} \quad P_2 = 2.175 \times 10^{-3}$$

ה. הרנדומיזציה והתוצאות

מבחן הרנדומיזציה שבו השתמשנו ב[1] וב[2], היה מבחן רנדומיזציה של פרמוטציות אקראיות. כאן ניישם אותו בדיוק באותו אופן ועם אותו זרע סטטיסטי. אני ממליץ לקוראים לעיין גם בהערה [8].

1. תוצאות מבחן הפרמוטציות עבור מדגם א:

הדירוג של P_1 מתוך 1,000,000 ערכי P_1 היה 255. כלומר $r_1 = 2.55 \times 10^{-4}$.
הדירוג של P_2 מתוך 1,000,000 ערכי P_2 היה 5,026. כלומר $r_2 = 5.03 \times 10^{-3}$.
המובהקות הסטטיסטית הכוללת היא: $r = 2 \times r_1 = 5.10 \times 10^{-4}$.

2. תוצאות מבחן הפרמוטציות עבור מדגם הלידה:

הדירוג של P_1 מתוך 1,000,000 ערכי P_1 היה 23 כלומר $r_1 = 2.3 \times 10^{-5}$.
הדירוג של P_2 מתוך 1,000,000 ערכי P_2 היה 1,113 כלומר $r_2 = 1.11 \times 10^{-3}$.
המובהקות הסטטיסטית הכוללת היא: $r = 2 \times r_1 = 4.6 \times 10^{-5}$.

הכרת תודה

אני מודה לד"ר שלום סרברניק על הרעיון לערוך ניסוי זה, וליעקב רוזנברג על התוכנה ששימשה לצורך המדידות.

מקורות והערות

1. המאמר

D. Witztum, E. Rips, and Y. Rosenberg. *Equidistant Letter Sequences in the Book of Genesis*. *Statistical Science*, 9 (3): 429-438, August 1994.

2. המאמר

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

3. במאמרנו [2] כתבנו כך: "בספרות התלמודית, בה ישנו זיהוי של חלק מן העמים מ'לוח העמים' ישנן כמה דעות במקורות שונים. כמובן, אין ביכולתנו להכריע בין דעות שונות, ולכן העדפנו להשתמש במקור יחיד. התרגום הארמי לבראשית המזהה את שמות העמים והארצות בהיקף הגדול ביותר, הוא 'תרגום יונתן'. השתמשנו בתרגום יונתן כפי שהוא מודפס ב'תורה שלמה' (עם תיקוני 'תורה שלמה' על פי כתב-יד גינזבורגר ואחרים)". ראה להלן הערה [8].

4. פרוייקט השו"ת, תקליטור מס' 5, אוניברסיטת בר-אילן, רמת-גן, התשנ"ז.

5. רבינו שמעון הדרשן, ילקוט שמעוני על התורה, ספר שמות, על פי כתב יד אוכספורד עם שנויי נוסחאות מכתב יד דפוסים ראשונים, מהדורת הימן – לרר – שילוני, הוצאת מוסד הרב קוק, ירושלים התשל"ג.

6. רבינו בחיי, ביאור על התורה. מוגה על פי דפוסים ראשונים וכתבי יד על ידי הרב חיים דב שעוועל, הוצאת מוסד הרב קוק, ירושלים התשמ"ב.

כאן המקום להעיר, כי רשימת תאריכי הלידה של בני יעקב הנמצאת במקור השני בתקליטור – ב"אוצר מדרשים" (אייזנשטיין) עמ' תעד – היא חסרה ומשובשת. המהדיר עצמו כותב שניסה לתקן את החסר לפי רבנו בחיי (שרשימתו זהה לרשימת ילקוט שמעוני [5]). חשוב לציין, כי גם אחרי תיקוני המהדיר נותרה גרסת "אוצר מדרשים" משובשת, שכן במדרש עצמו מבואר כי לכל חודש בשנה צריך להתאים תאריך של אחד מן השבטים (כלשון המדרש: "באדר ובאייר לא נולדו, נשתיירו לאפרים ולמנשה"). אולם גם לאחר תיקוני המהדיר ההתאמה חסרה.

אין ספק, גם אילולא רצינו לפעול בעקביות לפי [2], היינו חייבים לבחור בילקוט שמעוני [5] כמקור לרשימה, כי הוא המקור המקיף ביותר לעניין זה בספרות חז"ל, והוא היחיד המציג רשימה שלמה וחסרת סתירה של תאריכי בני יעקב (המופיעה במקור חשוב נוסף - [6]). ראו עוד להלן בהערה [8].

7. ד. ויצטום, א. ריפס, י. רוזנברג, צפן חבוי בדילוג שווה בספר בראשית: מובהקות סטטיסטית של התופעה, פרה-פרינט, אביב התשנ"ו.

8. מבקרנו שללו את השימוש בהגיון לצורך בניית מדגמים. למשל, הם טענו, כי אפשר להכליל במדגם נתונים שגויים (תאריכים לא נכונים), להשתמש בכתיב שגוי, או בצורות תאריך נדירות. גישתם נובעת מן ההנחה שאין צופן בספר בראשית. כמובן, טעות לוגית היא לבסס את האנליזה שלהם על הנחה זו.

בבניית המדגם המתואר במאמר, נעשה שימוש מסוים בהגיון. למשל, כאשר איתרתי את ילקוט שמעוני כמקור המקיף ביותר לעניין המידע המבוקש, דאגתי להשתמש במהדורה המדעית והמתוקנת [5], כדי להימנע עד כמה שאפשר ממידע שגוי (גם ב[2] "חטאתי" באותו שימוש בהגיון, ראה הערה 3 לעיל). יתר על כן, הקפדתי לכתוב את השמות כפי שהם כתובים בספר בראשית, ולא לפי הכתיב שבחר בו המעתיק (או המדפיס).

לגבי המקרה שלפנינו מתברר, שגם אילו הייתי מתעלם מן השיקולים הפשוטים שהזכרתי, התוצאה לא היתה שונה מהותית. אילו הייתי משתמש במידע מתוך ילקוט שמעוני שבתקליטור פרוייקט השו"ת, והייתי מעתיק את השמות בדיוק כפי שהם מופיעים בו, הרי ערכו של P_1 היה נעשה $P_1 = 8.4 \times 10^{-5}$. והדירוג של P_1 מתוך 1,000,000 ערכי P_1 היה 26, כלומר $r_1 = 2.6 \times 10^{-5}$.