BDD

Bekhol Derakhekha Daehu

Journal of Torah and Scholarship

36
March 2022

0 1100010752 1
דאנאקוד 110-10752
מחיר מומלץ 65 ₪

Bar-Ilan University Press

BDD

Bekhol Derakhekha Daehu

Journal of Torah and Scholarship

No. 36 – March 2022

Editor

Ely Merzbach

BAR-ILAN UNIVERSITY PRESS, RAMAT GAN

ISSN 0793-3894

Printed in Israel, 2022

CONTENTS

Doron Witztum

Measuring Significance in Torah Codes Research: Text Randomizations vs. Permutation Test

In their paper *Equidistant Letter Sequences in the Book of Genesis* (hereinafter: [1]), Witztum, Rips, and Rosenberg noted that when the Book of Genesis is written as two-dimensional arrays, equidistant letter sequences (ELSs) spelling words with related meaning often appear in *compact configurations*. Quantitative tools for measuring this phenomenon were developed and a sample of pairs of related words on which to run the test was chosen. The sample of word pairs was generated from a list of personalities' names and appellations paired with their Hebrew dates of birth and death. A permutation randomization of the sample of word pairs was used to measure the significance level.

The same permutation test was used in other Torah code experiments as well, including the "communities" experiments of Gans, Inbal, and Bombach (hereinafter: [2]), Simon (hereinafter: [3]), and of the committee to investigate the Gans-Inbal results (hereinafter: [4]). In these experiments the samples of pairs of related words were generated from a list of personalities' names and appellations paired with the names of places or communities, associated with these personalities.

Here we use a more direct approach to evaluate the significance level, by using the same data and measurement with many text randomizations. This test measures the probability that so many compact configurations have *accumulated* for the given sample just by chance.

On the other hand, the permutation test used in the work of Witztum,

Rips, and Rosenberg measured a specific feature of Torah codes. It measured to what extend the convergences of ELSs spelling names and appellations of the personalities with ELSs spelling their dates of birth and death are *more compact* than their convergences with ELSs spelling the dates of birth and death of other personalities.

It follows that text randomization is the appropriate test for experiments like [1]-[4] exploring the statistical evidence for the existence of Torah codes, rather than attempting to measure some specific characteristics of them.

We calculate the significance levels for the experiments [2], [3], and the committee's F list [4] via text randomizations. Some of the results show substantially stronger significance as compared to those derived by a permutation test for the same experiments. These results show that the correlation between the ELSs of names of the personalities and ELSs of names of the communities they are associated with, is manifested significantly no matter whether the data of Gans, Inbal, and Bombach are used, or Simon's or the committee's F list. This implies that even when the committee's F list or Simon's data are used, the encoding of the data as ELSs in the Book of Genesis is manifest at a statistically significant level.

Introduction

A. Background

In their paper *Equidistant Letter Sequences in the Book of Genesis* [1], Witztum, Rips, and Rosenberg (hereinafter: WRR) reported that they noted that when the Book of Genesis (G) is written as two-dimensional arrays, equidistant letter sequences (ELSs) spelling words with related meaning often appear in close proximity. Quantitative tools for measuring this phenomenon were developed. Some necessary details about WRR's methodology are given here.

1. WRR defined the notion of “distance” between any two words, so as to lend meaning to the idea of words in “close proximity.” It is done by function $c(w, w')$, "the corrected distance", for each word pair (w, w'). The corrected distance $c(w, w')$ is so normalized that the maximum distance is 1. A large $c(w, w')$, means that ELS’s representing w are far away from those

representing *w'*, on a scale determined by how far the *perturbed* ELS's for *w* are from those for *w'* (see [1], pgs. 434-435. See there that *c*(*w, w'*) was defined for expressions comprising 5-8 letters).

2. WRR defined two different statistics, *overall proximity measures* P_1 and P_2, which are defined and motivated in the Appendix Section A. Intuitively, each measures overall proximity in a different way. In each case, a small value of P_i indicates that the words in the sample pairs are, on the whole, close to each other.

3. WRR's sample of word pairs was built from a list of personalities *(p)* and the dates (Hebrew day and month) *(p')* of their death or birth. The personalities were taken from the *Encyclopedia of Great Men in Israel* (hereinafter: [5]). At first, the criterion for inclusion of a personality in the sample was simply that his entry contains at least three columns of text and that a date of birth or death be specified. This yielded 34 personalities (the *first list*). In order to avoid any conceivable appearance of having fitted the tests to the data, it was later decided to use a fresh sample, without changing anything else. This was done by considering all personalities whose entries contain between 1.5 and 3 columns of text in the *Encyclopedia;* it yielded 32 personalities (the *second list*). The significance test was carried out on the second sample (SAMPLE2) only.

Note that personality-date pairs *(p, p')* are not word pairs. The personalities each have several appellations, there are variations in spelling and there are different ways of designating dates. Thus each personality-date pair *(p, p')* corresponds to several word pairs *(w, w').* The precise method used to generate a sample of word pairs from a list of personalities is explained in [1] (in the Appendix, Section A.3).

B. Measuring the Significance of the Statistics P_i

The *overall proximity measures* P_i defined in [1] were the statistics used to measure the overall effects. The simplest way to get the significance of P_i was to repeat the same measurement with many "similar" texts, calculate their P'_is, and look how rarely one gets such small values as in *G*. WRR didn't

have the facilities to do it 30 years ago. Instead, another test, a permutation test (*PT*), was agreed upon by Persi Diaconis and Robert Aumann, and was performed in [1].

1. The Permutation Test

The permutation test served to evaluate the significance of the statistics P_i found for *G* measuring SAMPLE2.

SAMPLE2 consists of 32 personalities. For each of the 32! permutations π of these personalities, we define the statistic P_i^{π} obtained by permuting the personalities with accordance to π, so that Personality *k* is matched with the dates of Personality $\pi(k)$. The 32! numbers P_i^{π} are ordered, with possible ties, according to the usual order of the real numbers. If the phenomenon under study were due to chance, it would be just as likely that P_i occupies any one of the 32! places in this order as any other. This is the null hypothesis. To calculate the p-value, we chose 999,999 random permutations π of the 32 personalities. Each of these permutations π determines a statistic P_i^{π}; together with P_i we have thus 1,000,000 numbers. Define the rank order of P_i among these 1,000,000 numbers as the number of P_i^{π} not exceeding P_i; if P_i is tied with other P_i^{π}, half of these others are considered to "exceed" P_i. Let r_i be the rank order of P_i, divided by 1,000,000; under the null hypothesis, r_i is the probability that P_i would rank as low as it does.

2. Text Randomizations

About five and a half years after performing the permutation test, the rapid development of faster computers made Eliyahu Rips think that it may be possible to return to the simpler idea of measuring the significance by comparison to "similar" texts. Eliyahu Rips suggested in a mail [6] to David Kazhdan the following principles for measuring the significance of data list in the *Book of Genesis* (G):

- To use 1000 (or more) texts T_j, each text T_j is constructed by randomly permuting the words within each verse of *G* (the same way text *U* was constructed from *G* in [1], pg. 437).

- To estimate the significance of P_i in G, by comparison to the P'_is in the texts T_j.

Following this suggestion we formulate in detail the proposed test.

Let SAMPLE be a sample of word pairs (w, w'), and TEXT be a text.

(a) We choose the statistics to be used: P_1 or P_2 or both.

(b) We calculate $c(w, w')$ for each word pair (w, w') in SAMPLE for TEXT, and then calculate P_i (as was done in [1]).

(c) We do THE SAME for N "similar texts" T_j. Each text T_j is constructed by randomly permuting the words within each verse of TEXT (the same way text U was constructed from G in [1], pg. 437). For each text T_j we calculate P_i^j.

(d) We find n_i, the number of texts T_j for which $P_i^j \leq P_i$.

(e) The probability to get such small P_i by chance is $p_i = n_i/N$.

(f) In the case that both P_1 and P_2 are chosen, the overall significance for SAMPLE in TEXT is derived from p_1 and p_2 via Bonferroni inequality as in [1], or via less conservative (and more accurate) methods.

The significance test defined in stages (a)-(e) will be denoted *texts' test* (*TXT*).

C. Comparing Permutation Test (PT) and Text Randomizations (TXT)

Each test measures a different feature as is obvious from the current example:

- PT measures to what extend the convergences of ELSs spelling names and appellations of the personalities in SAMPLE2 with ELSs spelling their dates of birth and death are *more "successful"* than their convergences with ELSs spelling the dates of birth and death of other personalities in the same list.
- TXT measures the probability to get just by chance such small P_i; that is, it measures the probability that *so many "successful" convergences have accumulated*.

Therefore, we a priori expect to get different results for these two tests as they measure two different features.

An example:
The data list of word pairs measured by McKay et al. [7] in the book *War and Peace* [8] was prepared by declared manipulation. The manipulation was done specifically to get strong significance in the permutation test. Actually, they achieved a very strong significance for the chosen statistic:

$$p = 7E\text{-}7$$

But when TXT is applied to the same statistic, it turns out that for 12 out of 10,000 "similar texts" (created by randomly mixing their original text) – the statistic was *smaller*, hence:

$$p = 1.2E\text{-}3$$

The difference between these two results is quite huge.

D. Which Test Fits the Current Stage of the Torah Codes Research?

The current scientific research of Torah Codes, as far as it deals with measurement of large lists of word pairs (as in [1]-[4]), is exploring a formal statistical evidence for the *existence* of Torah codes, rather than attempting to measure some specific characteristic of them. This purpose was stated openly by WRR in the Introduction in [1] (on this issue see more in [9]). Therefore, according to the distinction made in Section C, it is clear that only TXT fits works [1]-[4] and not PT.

The Measurements

All our measurements via texts randomizations used (at most) 50,000 "similar" texts. This constraint is posed by our computational capacities. TXT needs several months of computing time for each big sample, as those in [1]-[4], using 50,000 "similar" texts. Therefore, measurements were done only for those samples and P_is that their significance via PT fits our resolution's capability, that is, p>0.00002. Since SAMPLE2 needs greater resolution, we concentrate on samples of [2]-[4].

Here are given the results of TXT for the communities experiment of Gans, Inbal, and Bombach [2], the communities experiment of Simon [3] (henceforth referred to as Sim), and for the "*fresh*" communities experiment

of the committee to investigate the Gans-Inbal results [4] (henceforth referred to as F).

The various communities experiments deal with measuring the proximity of ELS's spelling names and appellations of personalities and ELS's spelling names of places or communities, associated with these personalities.

The three works share in common the list of personalities and the list of their names and appellations. This list is the union of the *first list* and the *second list* of personalities, altogether 66 personalities, their names and appellations, as given in the original paper of WRR ([1], pgs. 432-433). We denote the combined list of names and appellations of the 66 personalities: List12.

Note: The following measurements were done with statistic P_2 (see the reason in the Appendix Section B.).

A. The Communities Experiment of Gans, Inbal, and Bombach

1. A Concise Description (a full description is given in [2])

(a) This experiment has two components:

(1) **The thesis**. The thesis includes an exact definition of the investigation's subject. The experiment was precisely the same as the WRR's experiment [1] except for one thing. A list of communities of birth and death for each personality was substituted for the list of dates of birth and death used in [1]. The personalities, appellations, formulae, and p-level calculations were identical to that of [1]. For the communities, the experiment used the Jewish names (wherever possible) and used the main variants of these names. The spelling of the community names followed fixed rules, those of [1].

(2) **Data collection.** The data used was established by the Inbal Algorithm using the following database: *Encyclopedia of Great Men in Israel* [5], *Encyclopedia Hebraica* [10], and the *New Concordance of the Bible* [11].

(Details and explanations concerning (1) and (2) are given in [2] and [12].)

(b) For each name collected in (a)(2), two new names were produced by using the prefix "קהל" and "קהלת" (Jewish community [of]).

Note: These prefixes denote the Jewish community within a city, as opposed to the city itself. This community was a distinct entity, with its own name, and its own spectrum of religious, social and political institutions. Thus, the *Encyclopedia Hebraica*, volume 29 page 183 under the entry קהלה (= קהל) says "The community was a city within a city".

(c) We denote two lists:

ListGns0 = the list of the community names.

ListGns01= ListGns0 + the names in ListGns0 with prefix "קהל" and "קהלת".

(d) We denote two samples of word pairs (*w, w'*):

Gns0 = the sample of word pairs constructed by taking each name of each personality from List12 and pairing it with each name of that personality's community from ListGns0.

Gns01 = the sample of word pairs constructed similarly from List12 and ListGns01.

The measuring scheme of WRR is designed for expressions comprising 5-8 letters, as mentioned in the Introduction. Abusing our notation somewhat, we continue to denote these samples Gns0 and Gns01, while intending only to word pairs conforming to this restriction.

2. The Measurement

(a) All $c(w, w')$ in both samples were measured in the Book of Genesis (G) and each P_2 was calculated.

(b) Four significance tests were done:

- TXT was done for both samples, running 50,000 competing "similar" texts. In each test we counted the number of competitors (randomized texts) for which the numerical value of the statistic P'_2 didn't exceed that of the original P_2.
- PT was done for both samples, running 50,000 competing random permutations. In each test we counted the number of competitors (random permutations of the personalities) for which the numerical value of the statistic P'_2 didn't exceed that of the original P_2.

3. The Results

The results are given in Table 1.

Table 1
The Number of Competitors n for which $P'_2 \le P_2$ Out of 50,000 Competitors

Sample	Gns0	Gns01
P_2	1.18E-6	4.39E-10
TXT	10	0
PT	249	5

There is *substantial* difference between the results of TXT and PT for sample Gns0: the significance measured via TXT is 25 times stronger than the significance measured via PT. For sample Gns01 we also see stronger significance measured via TXT, but an accurate comparison in this case is beyond our current resolution: it needs much more randomizations.

B. The "Fresh" (F) Experiment of the Committee

1. A Concise Description (a full description is given in [4]):

(a) The data for this experiment was originally intended to be collected by an expert appointed by the committee. Instructions prepared in advance by Dror Bar-Natan et al. were handed to an expert, Rami Reiner, in a written appointment.

(b) The list of places was prepared by the expert and included places where each personality in List12 was born or died, as well as "additional places where the personality stayed during his life." In some cases, the burial place was indicated.

(c) The committee representative marked the names to be used in the "fresh" experiment.

(d) Reiner erred in a pure geographical issue, when he mistakenly indicated that **Tudela,** the name of the small Spanish town in the province (and at the

time – kingdom) of **Navarra**, is just an alternative name for **Toledo** which was the capital of the kingdom of **Castile** and the capital of Spain for several hundred years. Therefore, Reiner added the name **Tudela** to personalities who were linked to **Toledo** and vice versa. We deleted this trivial error.

(e) All the other names marked by the committee representative are taken here *as they are*.

(f) The expert was instructed to choose appropriate prefixes for the names of the places (as did Gans, Inbal, and Bombach, where the prefixes "קהל" and "קהלת" were chosen). But this explicit instruction in the written appointment was ignored both by the expert who prepared the data, and by the committee representative who accepted the expert's list [13].

(g) In the present experiment we complete the missing prefixes.

- We adopt prefix "קהל" and "קהלת" from Section A1(b) above.
- We add two prefixes mentioned by committee members[1] as appropriate for names of Jewish communities: "עדת" and "ק"ק".

(h) We denote four lists:

ListFr0 = the list of places' names described in (b)-(e) above.

ListFr01= ListFr0 + the names in ListFr0 with prefix "קהל" and "קהלת".

ListFr02= ListFr0 + the names in ListFr0 with prefix "קהל" and "קהלת" and "עדת".

ListFr03= ListFr0 + the names in ListFr0 with prefix "קהל" and "קהלת" and "ק"ק".

ListFr04= ListFr0 + the names in ListFr0 with prefix "קהל" and "קהלת" and "עדת" and "ק"ק".

(i) We denote five samples of word pairs (*w, w'*):

Fr0 = the sample of word pairs constructed from List12 and ListFr0.

Fr01 = the sample of word pairs constructed from List12 and ListFr01.

Fr02 = the sample of word pairs constructed from List12 and ListFr02.

1 Hillel Furstenberg mentioned "עדת" while Robert J. Aumann mentioned "ק"ק". The committee's report [4], page 2, part C (#2) says: "Transcripts of the Committee meetings will be made available to the public at the Center for Rationality of the Hebrew University."

Fr03 = the sample of word pairs constructed from List12 and ListFr03.
Fr04 = the sample of word pairs constructed from List12 and ListFr04.

The measuring scheme is designed for expressions comprising 5-8 letters. Abusing our notation somewhat, we continue to denote these samples Fr0, Fr01, Fr02, Fr03 and Fr04, while intending only to word pairs conforming to this restriction.

2. The Measurement

(a) All $c(w, w')$ in these samples were measured in G and five P_2s were calculated.
(b) Ten significance tests were done: TXT was done for all five samples, running 50,000 competing "similar" texts, and PT was done for all five samples, running 50,000 competing random permutations.

3. The Results

The results are given in Table 2.

Table 2
The Number of Competitors n for which $P'_2 \leq P_2$ Out of 50,000 Competitors

Sample	Fr0	Fr01	Fr02	Fr03	Fr04
P_2	6.91E-4	1.699E-5	9.78E-7	1.71E-5	1.15E-6
TXT	205	27	3	43	15
PT	16,084	7,647	12,371	11,168	15,602

Here the significance measured via TXT is dramatically stronger then the significance measured via PT, for all five samples. In computing the overall significance of the five TXT measurements, the strong dependence between the five samples should be taken into account. As explained in the Appendix Section C, the overall significance is $p = 0.00026$.

C. The Communities Experiment of Barry Simon

1. A Concise Description (a full description is given in [3]):

(a) The protocol for this experiment was written by Barry Simon. According to this protocol the required data included the names of the places where the personalities of List12 were born, learned, lived, visited, or died. Simon writes: "I designed the experiment by myself and made only minor changes after consulting my wife, Alec Gindis and Brendan McKay."

(b) The data collection was done using *Encyclopedia of Great Men in Israel* [5], under the supervision of Alec Gindis.

(c) Simon didn't include in his protocol choosing prefixes for the names of the places, hence his list of names includes only 5-8 letters names, to conform to WRR's restriction. But for the present experiment we wanted to complete the missing prefixes as we did in Section B1(g). Therefore,

- We copied from the same Encyclopedia all the missing names of places with 2-4 letters *as they are*, as was done by Simon for the names of places with 5-8 letters. This list is documented in [14].
- We adopted prefix "קהל" and "קהלת" from Section A1(b) above.
- We added two prefixes mentioned by committee members as appropriate for names of Jewish communities: "עדת" and "ק"ק".

(d) We denote the following lists:

ListSim5_8 = Simon's list of names of places *as was published by him*.

ListSim2_4 = List of names of places from the *Encyclopedia* comprising 2-4 letters [14] mentioned in the previous subsection.

ListSim0 = ListSim5_8 + ListSim2_4

ListSim01 = ListSim0 + the names in ListSim0 with prefix "קהל" and "קהלת".

ListSim02 = ListSim0 + the names in ListSim0 with prefix "קהל" and "קהלת" and "עדת".

ListSim03 = ListSim0 + the names in ListSim0 with prefix "קהל" and "קהלת" and "ק"ק".

ListSim04 = ListSim0 + the names in ListSim0 with prefix "קהל" and "קהלת" and "עדת" and "ק"ק".

(e) We denote five samples of word pairs (w, w'):
Sim0 = the sample of word pairs constructed from List12 and ListSim0.
Sim01 = the sample of word pairs constructed from List12 and ListSim01.
Sim02 = the sample of word pairs constructed from List12 and ListSim02.
Sim03 = the sample of word pairs constructed from List12 and ListSim03.
Sim04 = the sample of word pairs constructed from List12 and ListSim04.

The measuring scheme is designed for expressions comprising 5-8 letters. Abusing our notation somewhat, we continue to denote these samples Sim0, Sim01, Sim02, Sim03 and Sim04, while intending only to word pairs conforming to this restriction. (Note that under this restriction, Sim0 = the sample of word pairs constructed from List12 and ListSim5_8 = Simons' original sample.)

2. The Measurement

(a) All $c(w, w')$ in these samples were measured in G and five P_2s were calculated.
(b) Ten significance tests were done: TXT was done for all five samples, running 50,000 competing "similar" texts, and PT was done for all five samples, running 50,000 competing random permutations.

3. The Results

The results are given in Table 3.

Table 3

The Number of Competitors n for which $P'_2 \le P_2$ Out of 50,000 Competitors

Sample	Sim0	Sim01	Sim02	Sim03	Sim04
P_2	2.58E-5	3.216E-7	2.1E-8	1.57E-6	1.26E-7
TXT	33	8	3	20	6
PT	15,211	10,099	10,280	10,825	10,965

Here again, the significance measured via TXT is dramatically stronger then the significance measured via PT, for all five samples. In computing the overall significance of the five TXT measurements, the strong dependence between the five samples should be taken into account. As explained in the Appendix Section C, the overall significance is $p = 0.00020$.

D. A Control Experiment

An additional control experiment was done by changing the order of letters in each of the community names. The details are in the Appendix Section D. The results were not significant.

Conclusions

- There is *substantial* and even dramatic difference between the results of TXT and PT.
- TXT shows that the correlation between the ELSs of names of the personalities and ELSs of names of the communities they were associated with, is manifested significantly no matter whether Gans, Inbal, and Bombach's data are used, or Simon's or the committee's F list.
- In other words: Even when Simon's data or the committee's F list are used – the encoding of the data as ELSs in the Book of Genesis is manifest.

Appendix

A. The Overall Proximity Measures P_1 and P_2

WRR defined two different statistics, *overall proximity measures* P_1 and P_2. Intuitively, each measures overall proximity in a different way. In each case, a small value of P_i indicates that the words in the sample pairs are, on the whole, close to each other. Here are their definitions as given in [1] (pg. 436).

1. Let N be the number of word pairs (w, w') in the sample for which the corrected distance $c(w, w')$ is defined. Let k be the number of such word pairs (w, w') for which $c(w, w') \le 1/5$. Define

$$P_1 \equiv \sum_{j=k}^{N} \binom{N}{j} \left(\frac{1}{5}\right)^j \left(\frac{4}{5}\right)^{N-j}.$$

To understand this definition, note that *if* the $c(w, w')$ were independent random variables that are uniformly distributed over [0, 1], *then* P_1 would be the probability that at least k out of N of them are less than or equal to 0.2. However, we do *not* make or use any such assumptions about uniformity and independence. Thus P_1, though calibrated in probability terms, is simply an ordinal index that measures the number of word pairs in a given sample whose words are "pretty close" to each other [i.e., $c(w, w') \le 1/5$], taking into account the size of the whole sample. It enables us to compare the overall proximity of the word pairs in different samples, so it serves as a statistic for the randomization tests.

2. The statistic P_1 ignores all distances $c(w, w')$ greater than 0.2 and gives equal weight to all distances less then 0.2. For a measure that is sensitive to the actual size of the distances, we calculate the product $\prod c(w, w')$ over all word pairs (w, w') in the sample. We define

$$P_2 \equiv F^N\left(\prod c(w, w')\right),$$

with N as above, and

$$F^N(X) \equiv X\left(1 - \ln X + \frac{(-\ln X)^2}{2!} + \cdots + \frac{(-\ln X)^{N-1}}{(N-1)!}\right).$$

To understand this definition, note first that if $x_1, x_2, \ldots, x_N$ are independent random variables that are uniformly distributed over [0, 1], then the distribution of their product $X \equiv x_1 x_2 \cdots x_N$ is given by

$$\Pr(X \le X_0) = F^N(X_0);$$

this follows from (3.5) in [15], since the $-\ln x_i$ are distributed exponentially, and $-\ln X = \sum_i (-\ln x_i)$. The intuition for P_2 is then analogous to that for P_1: *if* the $c(w,w')$ were independent random variables that are uniformly distributed over [0, 1], *then* P_2 would be the probability that the product $\prod c(w,w')$ is as small as it is, or smaller. But as before, we do not use any such uniformity or independence assumptions. Like P_1, the statistic P_2 is calibrated in probability terms; but rather then thinking of it as a probability, one should think of it simply as an ordinal index that enables us to compare the proximity of the words in word pairs in different samples.

3. In the original work of WRR, **four** *overall proximity measures* were defined: P_1, P_2, P_3 and P_4. According to ([1], pgs. 431, 436):

- Measures P_1 and P_2 were defined as different statistics for SAMPLE2.
- When applied to SAM2, a subsample of SAMPLE2, the statistics P_1 and P_2 are denoted by P_3 and P_4.

In other words: We have two operators $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$.
Applying them to SAMPLE2 gives P_1 and P_2:
$P_1 \equiv \boldsymbol{P}_1$(SAMPLE2), $P_2 \equiv \boldsymbol{P}_2$(SAMPLE2).
Applying them to SAM2 gives P_3 and P_4:
$P_3 \equiv \boldsymbol{P}_1$(SAM2), $P_4 \equiv \boldsymbol{P}_2$(SAM2).
The introduction of P_3 and P_4 was needed only for the permutation test, as is already written and explained in [1] (pgs. 431, 436). Here we don't use the permutation test, and need only P_1 and P_2.

B. The Statistic Used to Measure the "Communities" Experiments

All of the experimenters, Gans, Simon and the Committee chose to use the operator $\boldsymbol{P}_2$ to measure the various "communities" experiments, so we did here.

C. The Overall Significance for Tables 2 and 3

1. An upper bound of the overall significance of the five TXT results in Table 2 may be calculated by Bonferroni procedure, by multiplying the minimum p-value by 5. But this procedure is too conservative especially in our case due to strong dependence between the five samples (they have much of the data in common).

We proceed as follows. For each sample, the 50,000 random texts are ordered according to their P'_2 values in the usual order of the real numbers.

For sample Fr02 there were 3 random texts with $P'_2 \leq P_2$. Therefore the rank order of G is 4. The probability of a text to randomly have a rank order of 1-4 in at least one of the 5 measurements is simply the total number of *all different texts* which have rank orders 1-4 in at least one of the 5 measurements, divided by 50,000+1.

If the samples were not dependent, we would get about 4×5 such different texts. Actually, due to the strong dependence, there were only 13 such different texts (including G). Hence, the overall significance is

$$p = 13/50{,}001 = 0.00026.$$

2. The overall significance of the five TXT results in Table 3 is similarly derived.

The rank order of G for sample Sim02 is 4. There are only 10 different texts (including G) which have rank orders 1-4 in at least one of the 5 measurements. Therefore, the overall significance is

$$p = 10/50{,}001 = 0.00020.$$

D. A Control Experiment

1. We were asked "to form a list of all the community names and randomly permute the letters of each name. Then replace each name in the list with the permuted name. (Thus, if a community name appears several times in the list, the same permuted name will replace each). The prefixes should be used as before without their letters being permuted." This procedure was done with one change: instead of random permutation of letters, a cyclic shift of the letters of each name was used. Thus, the control experiment

can easily be replicated. (The cyclic shift was used by WRR when they were asked by Persi Diaconis to do a single permutation of their data[2] for a control experiment.) For example: "וילנאו" becomes "ילנאו"; "קהל וילנא" becomes "קהל ילנאו".

2. The procedure was applied to all the samples described above. This time TXT was done using 1000 random texts. The probabilities are given below in parallel to the 12 samples in the tables above.

For Table 1: 0.178, 0.092.

For Table 2: 0.289, 0.088, 0.037, 0.080, 0.043.

For Table 3: 0.339, 0.147, 0.049, 0.134, 0.059.

Acknowledgements

The software used for TXT and PT was prepared by Yaakov Rosenberg. A considerable help in computational capacities was given by a friend who wants to remain anonymous.

References

1. 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. Gans, H., Inbal, Z., and Bombach, N. (2006). 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.
3. This reference includes three documents:
   (a) Simon, B. (1998). *An open letter to Harold Gans*. http://www.torahcode.co.il/pdf_files/oppose/openletter.pdf.
   (b) Simon, B. (1998). *Raw Data: The Yeshiva Students' List*.

2 WRR shifted the list of dates with respect to the list of names. See: Witztum, D., Rips, E. and Rosenberg, Y. (1988). *Equidistant letter sequences in the Book of Genesis* ("the blue preprint"), pp. 5, 24.

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

4. 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.
5. Margalioth, M. ed. (1961). *Encyclopedia of Great Men in Israel; a Bibliographical Dictionary of Jewish Sages and Scholars from the 9th to the End of the 18th Century* **1-4**. Joshua Chachick, Tel Aviv.
6. An email dated 15 of May 1997, 16:40:13.
7. 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.
8. Tolstoy, L. N. (1953). *War and Peace*. Hebrew translation by L. Goldberg, Sifriat Poalim, Merhavia.
9. Witztum, D. (2003). *Tzofen Bereshith* (Hebrew), Chapter 3. Hamesorah, Jerusalem.
10. *Encyclopedia Hebraica* (1981). Encyclopedia Publishing Company, Jerusalem.
11. Even-Shoshan, A. ed. (1981). *A New Concordance of the Bible*."Kiryat Sefer", Jerusalem.
12. Gans, H., Inbal, Z., and Bombach, N. (2006). *The Linguistic Protocol and Data used for the Communities Experiment*.
    Available at:
    http://www.torahcode.co.il/pdf_files/pub/communities_data.pdf
13. Aumann, R. J., Furstenberg, H., Lapides, I. and Witztum, D, (2004). *Analyses of the "Gans" Committee Report.* Discussion Paper #365, Center for the Study of Rationality, The Hebrew University of Jerusalem, pp. 10, 19-20.
    Available at: http://www.torahcode.co.il/pdf_files/oppose/dp365.pdf.

14. The list of names of places from the *Encyclopedia of Great Men in Israel* comprising 2-4 letters is available at: http://www.torahcode.co.il/pdf_files/oppose/makom2_4.pdf.
15. Feller, W. (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley and Sons, New York.

מדידת מובהקות במחקר הצפנים בתורה: רנדומיזציה של הטקסט לעומת מבחן פרמוטציות

דורון ויצטום

המאמר "Equidistant Letter Sequences in the Book of Genesis" מאת דורון ויצטום, אליהו ריפס ויואב רוזנברג, 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 אלה מכונים צפני "ELS" חוקרים הבחינו כי כאשר ספר בראשית נכתב בצורה של טבלאות דו־ממדיות, צפני "ELS" של ביטויים בעלי קשר מושגי ביניהם נוטים להופיע על פני הטבלאות במפגשים מכונסים. פותחו כלי מדידה לתופעה זו, ונבחר מדגם של זוגות ביטויים קשורים מושגית. בכל זוג, ביטוי אחד הוא שם אדם או כינויו והביטוי השני הוא התאריך העברי של לידתו או פטירתו. הנתונים שייכים לקבוצה של אישים (גדולי תורה). רמת המובהקות של מדגם זה נמדדה באמצעות מבחן פרמוטציות. כל פרמוטציה מצמידה באקראי את תאריך הלידה או הפטירה של פלוני לשמו או לכינויו של פלמוני (לכל פלוני ופלמוני הנמצאים ברשימת האישים), ונוצר מדגם רנדומלי אשר נמדד בדיוק באותו האופן כמו המדגם המקורי. קביעת רמת המובהקות של תוצאות מדידת המדגם המקורי נעשית על ידי השוואה לקבוצה גדולה של תוצאות מדידות של מדגמים רנדומליים כאלה.

מבחן פרמוטציות דומה שימש למדידת רמת המובהקות בניסויים רבים הקשורים למחקר הצפנים בתורה, ביניהם "ניסויי הקהילות" בספר בראשית של גאנז־ענבל־בומבך, של סיימון ושל הוועדה האוניברסיטאית בראשות ישראל אומן. ניסויי הקהילות עוסקים במדידת מפגשים בין צפני "ELS" המייצגים שמות אישים וכינוייהם לבין אלה המייצגים את שמות המקומות/קהילות בהן חיו.

כאן נעשה שימוש בשיטה ישירה יותר לחישוב רמת המובהקות – רנדומיזציה של הטקסט. אותה מדידה של מפגשי זוגות המדגם הנתון מתבצעת לא רק בספר בראשית, אלא גם במספר רב של טקסטים "דומים", שנוצרו על ידי רנדומיזציה של ספר בראשית. קביעת רמת המובהקות של תוצאות המדידה בספר בראשית נעשית על ידי השוואה לתוצאות המדידות בטקסטים "הדומים". מבחן זה מודד את ההסתברות לכך, שכל כך הרבה מפגשים "מוצלחים" הצטברו במקרה עבור המדגם הנתון בספר בראשית.

לעומת זאת, מבחן הפרמוטציות מתמקד במדידת תכונה ספציפית של "צפני ELS" למשל, באיזו מידה "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).

כל המחקרים שנזכרו לעיל עסקו בהוכחת קיומו של צופן מסוים. במוצהר, הם לא עסקו באפיון מדויק של תכונות ספציפיות של הצפנים. לכן המבחן המתאים לקביעת רמת המובהקות של תוצאותיהם הוא באמצעות רנדומיזציה של הטקסט.

רמות המובהקות של ניסויי הקהילות של גאנז־ענבל־בומבך, של סיימון ושל הניסוי ה"טרי" (fresh) של הוועדה האוניברסיטאית בראשות ישראל אומן, נקבעו כאן באמצעות רנדומיזציה של הטקסט. נמצא כי יש הבדל מהותי ודרמתי בין רמות המובהקות שנמדדו לפי מבחן הפרמוטציות לאלה שנמדדו באמצעות מבחן הטקסטים. מתברר ממבחן הטקסטים, כי הקורלציה בספר בראשית בין "צפני ELS" המייצגים שמות אישים וכינוייהם לבין "צפני ELS" של שמות הקהילות אליהן היו קשורים, מובהקת בהחלט: בין לגבי נתוני גאנז־ענבל־בומבך, בין לגבי נתוני רשימת סיימון ובין לגבי נתוני הניסוי ה"טרי".

במילים אחרות: גם אם היינו משתמשים בנתוני רשימת סיימון או נתוני הניסוי ה"טרי", היינו מזהים הצפנה של הנתונים כ"צפני ELS" בספר בראשית ברמת מובהקות חזקה.