

A STATISTICAL CRITIQUE OF THE WITZTUM *et al* PAPER

A. M. Hasofer

February 18, 1998

Abstract. This paper examines whether the significance test described in the paper *Equidistant Letter Sequences in the Book of Genesis* by Witztum, Rips and Rosenberg was carried out in accordance with accepted procedure, and whether the distance metric used accomplishes its stated purpose. In the light of that examination, it appears that the conclusion of the paper is unfounded.

Key words and phrases: Significance tests, alternative hypothesis, perturbation, Genesis, Equidistant letter sequences.

1 Introduction.

The paper *Equidistant Letter Sequences in the Book of Genesis* by Witztum, Rips and Rosenberg [7] is probably the Mathematical Statistics paper which has enjoyed the widest dissemination ever, having been reproduced *in extenso* in Drosnin's book [2] of which hundreds of thousands of copies have been sold. Wild claims have been made about it. For example, Drosnin writes: "The original experiment *that proved the existence of the Bible code* (emphasis added) was published in a U. S. scholarly journal, Statistical Science, a review journal of the Institute of Mathematical Statistics...In the nearly three years since the Rips-Witztum-Rosenberg paper was published, no one has submitted a rebuttal to the math journal." (p.428). He also writes: "The three referees at the math journal...started out skeptics and ended up believers". In the remainder of this paper we shall refer to it as WRR. Equidistant letter sequences will be referred to as ELS's.

The present paper examines whether the conclusion of the WRR paper is justified.

2 Significance test procedure.

Tests of significance are also known as tests of hypotheses because they are based on setting up a "null hypothesis" and testing its significance [1, p. 64].

The following steps must be followed when setting up a test of hypothesis. Full details and references are given in Appendix A.

1. Set up a sample space.
2. Set up a null hypothesis.
3. Set up a critical region.
4. Set up the alternative hypothesis (or hypotheses).
5. Investigate the power of the test.

We now examine to what extent WRR have carried out the accepted procedure.

3 The sample space and the null hypothesis in WRR

In previous work published on the ELS's (e.g. Michelson [6]) the statistical calculations were based on the following hypothesis H_0 : each letter in the text is an independent discrete random variable taking one of 22 values, each corresponding to a letter of the Hebrew alphabet. The distribution of each discrete random variable is multinomial, with the probability of each letter being equal to the relative frequency of the letter in the whole text. Since the book of Genesis is 78,064 letters long, the sample space contains $22^{78,064}$ texts, almost all of them gibberish.

In the final version of WRR, there are still some vestiges of the earlier null hypothesis, e.g. in the calculation of the expected number of ELS's for a word and the "expected distance" between two words (p.435). However, WRR declare their null hypothesis H'_0 to be as follows: They define four overall measures of proximity P_1, P_2, P_3 and P_4 between the 32 personality- date pairs considered. For each of the $32! (\approx 2.63 \times 10^{35})$ permutations π of the personalities they define the corresponding statistic P_1^π . Thus the sample space is defined as the set of $32!$ permutations of the personalities. The P_1^π are put in ascending order. The null hypothesis is that P_1 is just as likely to occupy any one of the $32!$ places in this order as any other, and similarly for P_2, P_3 and P_4 (p.431). They provide the following motivation for the hypothesis: "If the phenomenon under study were due to chance, it would be just as likely that (the correct order of the names) occupies any one of the $32!$ places... as any other."

It must be emphasized at this point that this equiprobability assumption has no frequency interpretation because we are faced here with a unique object, namely the text of Genesis. The word "chance" used by the authors is meaningless in this context. Because of this, the conclusion that "the proximity of ELS's with related meanings in the book of Genesis is not due to chance" is equally meaningless.

It must also be pointed out that equiprobability in a permutation test usually follows by a symmetry argument from the assumption that the underlying variables are independent and identically distributed [1, p. 182] as the original null hypothesis assumed. But unfortunately here the symmetry argument fails for the following reasons:

1. Each personality may have several appellations, varying between one and eight.
2. Each date takes different forms, varying between one and six.
3. The lengths of the appellations and the dates vary between five and eight.

4. The sample of word pairs is constructed by taking each name of each personality and pairing it with each designation of that personality's date. Thus, when the personalities are permuted, the total number of word pairs in the sample varies. We are not told in the paper of the range of variability. For the identity permutation the sample consists of 298 words ([7, p. 436]), but it appears that that number varies by more than 100 between different permutations.([4]).
5. The corrected distance is defined only when there are more than nine perturbation triples for which a distance can be calculated. If not, the corrected distance is not defined. We are not told how often this happens, or what happens to the considered pair of words. Presumably, they are just dropped from the sample. This would introduce an additional uncontrolled bias in the calculation.

Thus the proposed null hypothesis H'_0 does not have any *a priori* justification even on the basis of the original hypothesis H_0 .

The problem of lack of frequency interpretation has been forcefully highlighted by Matheron [5, p. 23] in the following words: “When we deal with a unique phenomenon and a probabilistic model, that is a space (Ω, α, P) which is put in correspondence with this unique reality,... (an anthropomorphic) illusion incites us to say that everything happens, after all, as if the realized event had been ‘drawn at random’ according to law P in the sample space Ω . But this is a misleadingly clear statement, and the underlying considerations supporting it are particularly inadequate. What is the mechanism of this ‘random choice’ that we evoke, which celestial croupier, which cosmic dice player is shaking the iron dice box of necessity? This ‘random draw’ myth, for it is one, (in the pejorative sense), is both useless and gratuitous. It is gratuitous, for even if we assume that a unique random draw had been performed, once and for all, and an element ω_0 had been selected, we would in any case have no hope of ever reconstructing either the space Ω or the probability P . For since we are dealing with a unique event, the only source of information we possess is the unique element ω_0 , which was chosen at first: all the rest, the whole space Ω and the (almost) infinite ocean of possibilities it contained, will have disappeared, erased forever by this unique draw. It is useless, that is, it has no explanatory value, basically for the same reason: for the properties that we can observe in our universe are contained in this unique element ω_0 , and no longer depend on anything else. Thus we can confidently ignore all the riches which slumbered in the other elements ω , those which have not been chosen. In any case, this element ω_0 , ours, had doubtlessly (an almost) zero probability of being drawn, and thus our universe is ‘almost impossible’: nevertheless, it is the only universe given to us, and the only one we can study.”

Thus, the null hypothesis advanced by WRR must be considered as “entirely hypothetical” ([1, p. 5])

4 The alternative hypothesis in WRR.

In contrast with the situation with the null hypothesis, WRR carefully avoid to make in the paper (or for that matter, apparently anywhere else) any statement at all about the alternative hypothesis or hypotheses. This in itself is a major flaw of the whole investigation, as is abundantly clear from Appendix A, because in the absence of an alternative hypothesis it is impossible to calculate the probability of a Type II error and there can therefore be no grounds for rejecting the null hypothesis, no matter how unlikely it might be in the light of the observations.

The WRR paper has however been widely publicised and those who have made use of it have been far less reticent than the WRR authors. For example, Drosnin [2] concludes: “We do have the scientific proof that some intelligence outside our own does exist, or at least did exist at the time the Bible was written.” (p.50). In “public statements” attributed to two of the WRR authors (E. Rips and D. Witztum) and circulated on the Internet, commenting on Drosnin’s book, the above conclusion was not refuted and the expression “Bible codes” was freely used.

It is therefore appropriate to consider as a possible alternative that the book of Genesis was written by an intelligent being who could predict the future and encoded this information in the text. It must of course be noted that this hypothesis does not entail either that this being is benevolent or that it is still in existence.

Some motivation for the research is given in the Introduction of WRR. The approach is illustrated by the example of determining whether a text written in a foreign language is meaningful or not. Pairs of conceptually related words are considered and WRR write: “A strong tendency of such pairs to appear in close proximity indicates that the text might be meaningful.” They then declare that the purpose of the research is “to test whether the ELS’s in a given text may contain ‘hidden information’.”

It is difficult to avoid the impression that the above reasoning is just a naive anthropomorphism. On what grounds can we make any reasonable assumptions about the thought processes of a being that can predict the future? Why should a text produced by such a being have properties similar to those of an ordinary text in a foreign language written by a human?

But let us suppose that we agree to test the alternative hypothesis that the encoder has actually put the appropriate dates nearest to each of the names according to some distance measure P^* . The alternative hypothesis would then attribute a probability of one to the ordering where the correct match had the smallest distance measure and zero to all others. The optimal critical region would be just that one ordering, the size of the test would be $1/32!$ and its power would be unity. Unfortunately, the data are not consistent with this hypothesis, since none of the proposed measures falls in the proposed critical region.

WRR estimate the rankings of their proposed measures P_i , $i = 1, \dots, 4$ by a Monte-Carlo method, using one million permutations chosen at random. This is perfectly acceptable, but what they do not explicitly state is that even if we take the “best” of their measures of distance, namely P_4 , there are still an estimated $32! \times 3/10^6 \approx 7.89 \times 10^{29}$ permutations where at least two of the correct dates are not the nearest ones to the appropriate names, but whose ranks are smaller than the rank of the correct match! Some motivation for the encoder to choose such a bizarre encoding must be provided by the authors, for otherwise we must conclude that either their measures of distance are wildly off the mark, or else the alternative hypothesis that there is encoded information about the 32 personalities in Genesis obeying their “close proximity” criterion is false. The distance measure used by WRR will be examined more closely in the next Section.

Another problem is the fact that the experiment has been conducted solely on the book of Genesis, while in all the public statements attributed to Rips and Witztum reference is made to “codes in the Torah”. But the word “Torah” refers to all five books of the Pentateuch. In fact it is not known when the Torah was divided into five books and by whom. There are also various opinions as to where each book begins and ends. For example the Gaon of Vilna held that Deuteronomy actually starts from the fifth verse of our present version ([6, p. 10]). The authors must explain why they chose to conduct the experiment solely on Genesis. Moreover the argument provided must be an *a priori* one. This is in view of the unfortunate fact that when conducted on the other four books of the Pentateuch, their experiment (carried out on the second list of personalities) failed

Book	P_1	P_2
Genesis	718	2
Exodus	135,735	193,315
Leviticus	816,860	947,387
Numbers	901,660	920,919
Deuteronomy	790,542	759,428

Table 1: Rank order of P_1 and P_2 among one million P_i^π .

to show any “significant” effect, at least as far as the P_1 and P_2 statistics are concerned, as shown in Table 1 ([4]). The results given for Genesis in the table are not exactly the same as those appearing in the WRR paper because the experiment was carried out with an updated version of the software, provided by WRR.

5 A critique of the WRR “corrected distance”.

In WRR a “corrected distance” between ELS’s is defined on p.435. No mathematical justification for the procedure is given. The “corrected distance” between two words w and w' is supposed to be small when w is “unusually close” to w' , and 1 or almost 1 when w is “unusually far” from w' . There are some problems in the exposition of the procedure:

1. It will not work for skips of 1, 2 or 3 because some of the perturbed ELS’s will not exist.
2. The *distance* between two perturbed ELS’s “is defined as the distance between the ordinary (unperturbed) ELS’s”. This definition is ambiguous because the definition of distance between unperturbed ELS’s involves f and f' , the distances between consecutive letters in the two ELS’s. But in the perturbed ELS’s this distance is not fixed. In what follows we assume that the f and f' of the unperturbed ELS’s, but the actual minimal distance between the letters of the perturbed ELS’s, are to be used.
3. If the number of perturbed ELS’s (out of 125) that actually appear in the text is less than 10, the corrected distance is not defined. The paper does not state what happens to the concerned pair. In the programs supplied by WRR the pair is simply ignored (Source: McKay [4]).

But the really serious problem is that

1. the word “unusually” used in the justification is meaningless, since we are dealing with a unique text,
2. one can easily construct examples where the “corrected distance” yields a result that is totally at variance with common sense. Such an example is discussed further along in this Section. Full details are given in Appendix B.

According to WRR, the “uncorrected proximity”, Ω , “very roughly measures the maximum closeness of the more noteworthy appearances of w and w' as ELS’s in Genesis - the closer they are, the larger is Ω ” (p.435). It can be said in its favour that when applied to

pairs of words which occur only once and are all of the same skip and length it reduces to a monotonically decreasing function of the minimal distance along the text between the letters of the word pairs. This is quite reasonable.

On the other hand, the “corrected distance” has nothing to do with the actual position of the words whose distance is supposed to be measured. It is based on the rank of the “uncorrected distance” Ω between the two original words when compared to the distances of pairs of “perturbed ELS’s”. But the sets of “perturbed ELS’s” are different for different pairs of words and so do not provide any unified standard of comparison of distance.

For our counterexample we focus on just two pairs of four-letter words, each having just one ELS, so as to keep the calculations simple and transparent. They are denoted by w_1 , w'_1 , w_2 and w'_2 . The minimum distance along the text between the letters for the first pair is 90,000 letters and for the second pair 30,000. The “corrected distance” for the first pair turns out to be zero and for the second pair one.

Thus, according to the “corrected distance”, w_1 and w'_1 are “very close”, while w_2 and w'_2 are “very distant”. This is totally at variance with any commonsense definition of distance between ELS’s and contradicts all the examples given in the Introduction of WRR to motivate the work. In addition, it is clear from the details of the counterexample given in Appendix B that by translating the set of perturbed ELS’s without changing the position of the original ELS’s we can vary the “corrected distance” from 0 to 1. The only limitations are due to boundary effects: if the two words are very close there may not be enough space to insert perturbed ELS’s. Conversely, if the two words are near the beginning and the end of the text there may not be enough space to translate the perturbed ELS’s away from the two words.

That the phenomenon described in the counterexample is not artificial can be illustrated by the data of the experiments themselves as follows: all (matching) pairs of words from the two personality lists that did yield an Ω and a c were collected (Source: McKay [4]). There were 320 pairs. As an index of “corrected proximity” we used $1 - c$. The range of Ω was approximately (77 – 60, 365). Since we are interested in the ranking of proximities, a natural measure of the concordance of the two indices is the rank correlation coefficient ([3, p. 494]). The overall value turns out to be 0.603. However, an examination of the scattergram of the ranks indicates that the correlation is mainly due to small and large values of Ω . Indeed, if we select 100 values of Ω from the middle of the range (specifically ranks 101 to 200, corresponding to $1450 < \Omega < 3623$) the rank correlation between them and the corresponding $1 - c$ ’s falls to 0.088. If the pairs were a random sample the hypothesis of zero rank correlation would be accepted at the 20% significance level. The analysis thus supports the conjecture that any association between Ω and $1 - c$ is not inherent but entirely due to boundary effects.

6 Conclusion.

The following flaws in the WRR paper have been documented:

1. The probability distribution embodied in the null hypothesis is purely hypothetical and cannot be justified on grounds of symmetry, as is usually done for permutation tests.
2. No alternative hypothesis is stated, so that the power of the proposed test cannot be evaluated. A powerful reasonable alternative hypothesis, based on the WRR heuristics, is proposed by the writer, but the experiment does not reject the null hypothesis under that proposed alternative.

3. No explanation is given for the choice of the particular book of the Pentateuch on which the experiment was carried out.
4. Some explanation must be given for the failure of the test to show any significant effect in the other four out of five books of the Pentateuch.
5. The definition of “corrected distance”, which is the basic building block of the whole experiment, is shown not to achieve its purpose, for it is easy to construct counterexamples where the “corrected distance” algorithm leads to a result which is contrary to common sense. Moreover, analysis of the correctly matched pairs of words in the two personality lists supports the conjecture that any association between the proximity measure Ω and the “corrected distance” c is not inherent but entirely due to boundary effects.

Until these flaws are remedied, the claims made in the paper must be considered as statistically unfounded.

7 Appendix A: Details of the Hypothesis Testing Procedure.

7.1 Setting up the null hypothesis.

The following quotation from Kendall and Stuart’s *Advanced Theory of Statistics* [3, p. 169] describes what type of null hypothesis is appropriate for a significance test: “The kind of hypothesis which we test in statistics is more restricted than the general scientific hypothesis. It is a scientific hypothesis that every particle of matter in the universe attracts every other particle, or that life exists on Mars; but these are not hypotheses such as arise for testing from the statistical viewpoint. Statistical hypotheses concern the behaviour of observable random variables. More precisely, suppose that we have a set of random variables x_1, \dots, x_n . As before, we may represent them as the co-ordinates of a point (\mathbf{x} say) in the n -dimensional sample space, one of whose axes corresponds to each variable. Since \mathbf{x} is a random variable, it has a probability distribution, and if we select any region, say w , in the sample space W , we may, (at least in principle) calculate the probability that a sample point \mathbf{x} falls in w say $P(\mathbf{x} \in w)$. We shall say that any hypothesis concerning $P(\mathbf{x} \in w)$ is a statistical hypothesis. In other words, any hypothesis concerning the behaviour of observable random variables is a statistical hypothesis.”

7.2 Setting up the critical region.

The next step in the testing procedure is the setting up of the critical region. We quote again from Kendall and Stuart [3, p. 171]: “To test any hypothesis on the basis of a (random) sample of observations, we must divide the sample space (i.e. all possible sets of observations) into two regions. If the observed sample point \mathbf{x} falls into one of these regions, say w , we shall reject the hypothesis; if \mathbf{x} falls into the complementary region, $W - w$, we shall accept the hypothesis. w is known as the *critical region* of the test, and $W - w$ is called the acceptance region.

It is necessary to make it clear at the outset that the rather peremptory terms ‘reject’ and ‘accept’ which we have used of a hypothesis under test are now conventional usage,

to which we shall adhere, and are not intended to imply that any hypothesis is ever finally accepted or rejected in science. If the reader cannot overcome his philosophical dislike of these admittedly inapposite expressions, he will perhaps agree to regard them as code words, ‘reject’ standing for ‘decide that the observations are unfavourable to’ and ‘accept’ for the opposite. We are concerned to investigate procedures which make such decisions with calculable probabilities of error, in a sense to be explained.

Now if we know the probability distribution of the observations under the hypothesis being tested, which we shall call H_0 , we can determine w so that, given H_0 , the probability of rejecting H_0 (i.e. the probability that \mathbf{x} falls in w) is equal to a preassigned value α , i.e.

$$\text{Prob}\{\mathbf{x} \in w|H_0\} = \alpha. \tag{1}$$

...The value α is called the *size* of the test...”

7.3 Setting up the alternative hypothesis.

We continue the quotation: “Evidently, we can in general find many, and often even an infinity, of sub-regions w of the sample space, all obeying (1). Which of them should we prefer to the other? This is the problem of the theory of testing hypotheses. To put it in everyday terms, which sets of observations are we to regard as favouring, and which as disfavouring, a given hypothesis?”

Once the question is put in this way, we are directed to the heart of the problem. For it is of no use whatever to know merely what properties a critical region will have when H_0 holds. What happens when some other hypothesis holds? In other words, we cannot say whether a given body of observations favours a given hypothesis unless we know to what alternative(s) this hypothesis is compared. It is perfectly possible for a sample of observations to be a rather ‘unlikely’ one if the original hypothesis were true; but it may be much more ‘unlikely’ on another hypothesis. If the situation is such that we are forced to choose one hypothesis or the other, we shall obviously choose the first, notwithstanding the ‘unlikeliness’ of the observations. The problem of testing a hypothesis is essentially one of choice between it and some other or others. It follows immediately that whether or not we accept the original hypothesis depends crucially upon the alternatives against which it is being tested.”

7.4 The power of a test.

We continue the quotation: “The (above) discussion ... leads us to the recognition that a critical region (or, synonymously, a test) must be judged by its properties both when the hypothesis tested is true and when it is false. Thus we may say that the errors made in testing a statistical hypothesis are of two types:

1. We may wrongly reject it, when it is true;
2. We may wrongly accept it, when it is false.

These are known as Type I and Type II errors respectively. The probability of a Type I error is equal to the size of the critical region used, α . The probability of a Type II error is, of course, a function of the alternative hypothesis (say, H_1) considered, and is usually denoted by β . Thus

$$\text{Prob}\{\mathbf{x} \in W - w|H_1\} = \beta \tag{2}$$

Pert.No.	x	y	z
1	-1	0	1
2	1	0	-1
3	-1	1	0
4	0	1	-1
5	1	-1	0
6	0	-1	1
7	-2	0	2
8	2	0	-2
9	-2	2	0

Table 2: Perturbation triples.

or

$$\text{Prob}\{\mathbf{x} \in W|H_1\} = 1 - \beta. \quad (3)$$

This complementary probability, $1 - \beta$, is called the *power* of the test of the hypothesis H_0 against the alternative hypothesis H_1 . The specification of H_1 in the last sentence is essential, since power is a function of H_1

We seek a critical region w such that its power, defined at (3), is as large as possible. Then, in addition to having controlled the probability of Type I errors at α , we shall have minimized the probability of a Type II error, β . This is the fundamental idea, first expressed explicitly by J. Neyman and E. S. Pearson, which underlies the theory...

A critical region, whose power is no smaller than that of any other region of the same size for testing a hypothesis H_0 against the alternative H_1 , is called a best critical region (abbreviated BCR) and a test based on a BCR is called a most powerful... test."

8 Appendix B: Construction of the “corrected distance” counterexample.

We construct a string of length 110,000 as follows. We use the first 22 letters of the latin alphabet (written in capitals) to represent the letters of the Hebrew alphabet. We will focus on four “words” w_1 : ABCD, w'_1 : EFGH, w_2 : IJKL, and w'_2 : MNOP. We use the same notation as WRR for an ELS, namely (n, d, k) for the start, the skip and the length.

We set:

w_1 at (11992, 3, 4), w'_1 at (102001, 3, 4)

w_2 at (41992, 3, 4), w'_2 at (72001, 3, 4).

These will be the only ELS’s for these two words. We denote the distance along the text between the last letter of w_1 and the first letter of w'_1 by U_1 . Here $U_1 = 90000$. Similarly we denote the distance along the text between the last letter of w_2 and the first letter of w'_2 by U_2 . Here $U_2 = 30000$.

We now introduce 9 “perturbed” ELS’s for each of the four words, using the triples (x, y, z) given in Table 2:

As we are using a skip of 3, some of the perturbation triples used by WRR will not work, namely those for which $x + y = -3$ and/or $x + y + z = -3$. We avoid them in our list of triples.

Pert.No.	w_1	w'_1	U_1	w_2	w'_2	U_2
1	11962	102031	90060	42022	71971	29940
2	11932	102061	90120	42052	71941	29880
3	11902	102091	90180	42082	71911	29820
4	11872	102121	90240	42112	71881	29760
5	11842	102151	90300	42142	71851	29700
6	11812	102181	90360	42172	71821	29640
7	11782	102211	90420	42202	71791	29580
8	11752	102241	90480	42232	71761	29520
9	11722	102271	90540	42262	71731	29460

Table 3: Starting point of perturbed ELS's.

i	$L_1(h_i)$	$L_2(h_i)$	$f (= f')$
1	30000	10000	1
2	45000	15000	$\sqrt{2}$
3	90000	30000	3
4	90000	30000	3

Table 4: Distances between w_1 and w'_1 for various h_i 's

We note that, for all the triples we use, $x + y + z = 0$. Thus, the perturbation only affects the position of the second and the third letter of the words, but not the position of the first and last.

Table 3 gives the starting point of the perturbed ELS's for the four words and the distances along the text..

The rest of the string is completed arbitrarily with the 6 remaining letters.

We now calculate the minimal distance L_1 between a letter of w_1 and a letter of w'_1 . We first note that since the skip is 3, the only non-zero h_i 's, the nearest integers to $|d|/i$ ($1/2$ rounded up), are $h_1 = 3$, $h_2 = 2$, $h_3 = 1$, $h_4 = 1$. The other h 's contribute zero to $\sigma(w_1, w'_1)$. Here $h'_i = h_i$.

We note that the last letter of w_1 , (D), has rank 12001, so that for all non-zero h 's it is the first letter of the row. Similarly, the first letter of w'_1 , (E), has rank 102001, so that it will also appear as the first letter of the row for all non-zero h 's. The same considerations apply to w_2 and w'_2 , since H will have rank 42001 and I will have rank 72001. The minimal cylindrical distances between the letters of w_1 and w'_1 , denoted by $L_1(h_i)$ and those between the letters of w_2 and w'_2 , denoted by $L_2(h_i)$, are given by Table 4.

It can be seen that when the skip is 3, if the distance U along the text between the last letter of the first word w and the first letter of the second word w' is a multiple of 6, (as is the case in our construction) there are four finite cylindrical distances, namely $U/3$, $U/2$, U and U .

Using the WRR definitions, we find for the proximity measures σ : $\sigma(w_1, w'_1) = 1.5556 \times 10^{-4}$ and $\sigma(w_2, w'_2) = 4.6667 \times 10^{-4}$. This is as it should be, since w_2, w'_2 are much closer than w_1 and w'_1 by any reasonable criterion.

Moreover, since in our case there will be only one ELS (and one perturbed ELS for each perturbation triple) for every word considered, we will have

$$\Omega^{(x,y,z)}(w, w') = \sigma^{(x,y,z)}(w, w') \quad (4)$$

Pert. No.	$10^4 \times \Omega^{(x,y,z)}(w_1, w'_1)$	$10^4 \times \Omega^{(x,y,z)}(w_2, w'_2)$
1	1.5545	4.6760
2	1.5535	4.6854
3	1.5525	4.6948
4	1.5514	4.7043
5	1.5504	4.7138
6	1.5494	4.7233
7	1.5483	4.7329
8	1.5473	4.7425
9	1.5463	4.7522

Table 5: Proximity measures between perturbed ELS's.

for every perturbation triple.

We now calculate the proximity measures of the perturbed ELS's. They are given in 5.

From Table 5 we conclude:

$$c(w_1, w'_1) = 0 \tag{5}$$

$$c(w_2, w'_2) = 1 \tag{6}$$

As pointed out in Section 5 this is totally counterintuitive.

References

- [1] Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. Chapman and Hall.
- [2] Drosnin, M. (1997) *The Bible Code*. Weidenfeld and Nicholson.
- [3] Kendall, M. G. and Stuart, A. (1973) *The Advanced Theory of Statistics*. Vol. II. Charles Griffin & Co. Third Edition.
- [4] McKay, B.(1997) Private communication.
- [5] Matheron, G. (1989) *Estimating and choosing*. Springer-Verlag
- [6] Michelson, D. (1987) Codes in the Torah. *B'Or Ha'Torah*, No.6E, pp.7-39.
- [7] Witztum, D. Rips, E. and Rosenberg, Y. (1994) Equidistant Letter Sequences in the Book of Genesis. *Statistical Science*, **9**, 3, 429-438.