THE SPACES IN BETWEEN

A multiple test of authorship for Greek writers

This paper describes a set of seven tests of authorship for Greek writers based upon the occurrence of seven words, ἀλλα, γαρ, δε, δια, μεν, μη and ὀυ, all of which behave in the same way. The conjunction *alla* may be taken as typical of the words in the group, most of which tend to occur relatively infrequently and to show some bunching of occurrences in texts. To overcome the difficulties created by the infrequency and the bunching in the pattern of occurrence, it has been necessary to use a new unit of measurement, namely the interval which separates successive occurrences of the conjunction *kai*. To keep the length of the paper within reason the occurrence of *alla* is recorded in a range of authors but the other words in the group are only illustrated as they occur in the 'De Corona' of Demosthenes. All the tests of authorship are then applied to a classic case of disputed authorship, the Pauline Corpus in the New Testament.

It is unfortunately true that nothing is more likely to arouse the suspicions of a scholar without statistical training than the removal of an apparent obstacle to the acceptance of a hypothesis by a change of scale or alteration in units. Such a scholar finds it difficult to believe that the change has any more objective foundation than the ingenuity of the statistician and also suspects that, if he knew as many tricks of the trade, he could affect another transformation which would give good reason for rejecting the hypothesis. Such misgivings are generally without foundation.

It is true that statisticians have a number of transformations which they use but these reflect the nature of the observations and not the prejudices of the observer.

In this case we face two difficulties, that *alla,* and most of the other words in the group, do not occur very often in Greek writing and the occurrences are not evenly spread through texts but tend to come in bunches. The really frequent words in Greek writing are particles and connectives and if the limit of high-frequency vocabulary is set by the appearance of the most-frequent noun, then *alla,* and most of the other words in the group, occur less often than this noun and so would be classed as mid-frequency vocabulary.

This low rate of occurrence means that block samples of the usual size, a few thousand words, are of limited value. Such samples will have in them anything from 5 to 25 occurrences of *alla* and so meet the first requirement of a sample, that it should have, or can be expected to have, not less than five occurrences in it. The reason for this restriction is that no safe prediction can be made about the rate of occurrence of a repeated event until at least five occurrences have been recorded, a rule too often ignored in literary argument. To see how the occurrences vary from one sample to another, a number of samples is required, as many as fifty or sixty for some purposes, and so large amounts of text are needed. The comparison of one author with another by using block samples means that the

comparison is restricted to the number of occurrences of *alla* in the samples and so neglects any idiosyncrasies, such as using pairs of occurrences in certain types of construction, which enable a more detailed comparison to be a more sensitive discriminator between authors.

The first difficulty can be partially overcome by counting the occurrences in a smaller unit than block samples of a thousand words or more. If the occurrences are recorded in sentences, or in blocks of twenty successive words marked off right through a text, then useful judgements can be made if we have even a hundred sentences or small blocks providing that at least five of them contain occurrences of *alla*, and it is likely that some idiosyncrasies can be noted.

However the bunching of the occurrences complicates this proceedure. If words occur at random through a text and we record the occurrences in equal intervals, for example the twenty word blocks just described, then we expect to find a pattern of occurrences known as a *Poisson distribution. ref. 1.* If we record the occurrences in intervals which are not equal, for example sentences which differ in length, then we can expect to find either one of two patterns. If the probability of the occurrence is strictly proportional to the size of the unit, for example if the number of occurrences recorded in sentences with twenty words was exactly twice the number of occurrences in sentences of ten words, the pattern would still be Poisson. The reason for this is that the sum of any two Poisson

distributions is another Poisson distribution and in this situation we have a series of Poisson distributions one for each length of sentence and the sum of all these would be a Poisson distribution. But if the occurrences had a tendency to be more frequent in longer sentences, or to be absent from very short sentences, then this variable probability gives rise to a ***negative binomial distribution. ref. 2.***

In fact neither of these distributions cover what is found in Greek texts when occurrences of ***alla,*** and of many of these other words, are recorded in such units. When occurrences are counted in sentences, or in small blocks, there are always too many which have more than two or three occurrences. The number of sentences or blocks with no occurrence or one occurrence is near to the theoretical prediction but the number with more than two or three is often too large to be reconciled with theory. The excess is not large; it is usually a case of recording two or three sentences or samples with multiple occurrences when theory suggests that there should be less than one, but the differences are too great for the theory to be applicable.

What is needed is a variable unit in which the occurrences can be recorded. This unit would contract when the text is complex and occurrences of this group of words were frequent and expand when the structure of the text is simple and occurrences are rare. Such a unit is the space between successive occurrences of the conjunction *kai. ref. 3.* The length of these

spaces is measured by counting the number of words which separate successive occurrences. Not only do the spaces between occurrences of *kai* vary in the required manner, they also have two other advantages, they occur often and they are independent of all punctuation. *Kai* is generally the most frequent word-form found in Greek texts and, on the average, a thousand words of text will have about fifty occurrences of *kai* and so one space less than this, i.e. forty-nine.

To establish a test of authorship it is necessary to discover some habitual occurrence common to all writers of the class being studied - in this case writers of Greek - which each individual uses in his own way or at his own level. It must be shown that (1) - within any single work of an author and between works of the same author, the variations in the occurrence are such that chance alone will readily explain, and (2) - that between the works of different authors the variations often exceed anything that chance can explain.

An examination cannot cover all Greek writers and must be restricted to a selection of works and authors chosen so as to include the various contingencies which might affect the occurrence. This selection must include works which show differences of subject matter; the passage of periods of time between compositions of the same author; contrasts in literary forms; the differences which appear between careful and conscious stylists and simple natural writers, and so on. Then by a cross comparison

of the samples the effects of these influences can be isolated and measured. When the survey has been completed and it has been shown that the occurrence is consistent within authors, i.e. that the variations within the works of an author are such that chance will explain, while the variations in the occurrence between authors are often too great to be attributed to chance, then the occurrence can be used to determine the authorship of disputed texts. It is important to understand that all judgements about authorship made by these methods are exclusive. It can be shown by such comparisons that two works, dissimilar in respect of some occurrence are not by the same author, but similarity in two works no more implies identity of authorship than the recording of the same height twice in a series of measurements implies that they refer to the same person.

The Principle of Argument

To illustrate the principle of argument used in this paper and to show how comparisons between works are made, Table 1A shows the results of a count of the occurrences of *alla* in a sample taken from the start of each book of the history of Herodotus in the Oxford Classical Text.

In all there are 900 intervals between kais in the samples. In 823 of them there was no occurrence of *alla,* in 73 a single occurrence and in 4 there

were two occurrences. Someone looking at this evidence might say that Herodotus was certainly not consistent in his habit for book 6 has only 4 occurrences of *alla* in the samples while books 7 and 9 have 15, more than three times as many. But this distinction is not as clear cut as it might seem. Suppose that we were to take 900 cards and on 823 of them write - 'no-alla', on 73 write 'one-alla', and on 4 write 'two-allas'. If we shuffled all the cards and drew out cards and recorded the information on them until we had several sets of one hundred observations, then commonsense tells us that we would expect to find, on the average, 91.4 % of the cards were no-alla cards, 8.1 % were one-alla cards and 0.4 % were two-alla cards. We know that we cannot ever get these results exactly, at least not for samples of 100 cards, but we would expect to get figures near to these averages quite often and figures remote from these averages more rarely.

We can bring precision of measurement into the situation and get rid of terms like, 'more often' and 'rarely' by repeating the experiment and tabulating the results so that we could say of any set of differences that they had occurred once in ten trials of the experiment or not even once in one million trials of the experiment. In other words we can measure the size of any set of differences between what is found and the expectation by using as the unit of measurement how often chance alone would generate the differences.

In practise there is no need to perform the experiment or make the tabulation. Statistical mathematics has long been able to supply the results we need. One way of doing this is to use the ***chi squared distribution, ref. 4.*** The chi squared distribution is simple in conception and most useful; it enables one to test the fit of a set of observations to some hypothesis which is supposed to account for them. In such applications, including the present one, each observation is taken and the difference between the observation and the corresponding expectation is arrived at by subtraction, this difference is squared and then divided by the expectation to give chi squared for the observation. The process is repeated for every observation and the sum of all the results is chi squared for the table of observations. Having calculated chi squared we can then consult a table, to be found in all sets of statistical tables, which tells at once how often chance would account for the set of differences displayed in the table.

If we assume that Herodotus habit was consistent, that his habit can be represented by the averages for all the samples and calculate chi squared for the observations we can expect one of three results. It may be that the differences are such that chance will readily explain their appearance. For example if we were to discover that differences as large as these would be found once in every two trials of the shuffling and dealing experiment we could hardly argue that they were of much importance. But if the calculation showed that such differences would be accounted for only once

in a million trials by the action of chance we would correctly infer that some alternative hypothesis to that of a regular habit would have to be adopted. It is possible that the results lay in some middle region when the conclusion might have to be that judgement should be suspended or further testing be made.

The problem is now that of drawing a line at a suitable point to separate chance effects from effects so unlikely to be due to chance that some alternative explanation must be offered. Such a line must be drawn where it represents a balance between the labour involved in making a decision and the consequences of making a wrong decision. There are two *levels of significance,* as these points of decision are called, in general use. The *5 % level of significance* is the point at which chance will account for the differences shown by the observations only once in twenty trials of a random experiment such as our card experiment. The *1 % level of significance* is where the chance will account for the differences shown by the observations only once in one hundred trials of the random experiment. In fact the choice of level is not as important as it may seem because independent tests will rapidly make clear which is the correct conclusion. If we use the 5 % level of significance and use one test, then in one case in twenty it will mislead us in the sense that we would reject as not being due to chance a difference which chance had created. But if we use two tests at the 5 % level, this will happen only once in twenty times twenty trials, i.e. once in four hundred trials and if we use six tests it will only

happen once in 64 million trials.

In other words differences which can be explained by chance more than once in twenty trials, or once in one hundred trials, we will ignore and following statistical practise we will call *random sampling variations;* differences which chance will account less often than this we will call *statistically significant* differences, and for these seek some alternative explanation to that offered by the hypothesis.

To return to the samples from Herodotus, we can calculate chi squared for the table. In doing so we must first ensure that each cell has at least five, or can be expected to have at least five, occurrences of *alla* in it. The figures for *one-alla* and *two-alla* spaces must be added together and treated as *one-or-more-alla* spaces. The figure 4 in book six is in order for the average suggests that we should find 8.6 occurrences of such spaces in a set of one hundred. A correction must then be made to compensate for the fact that chi squared can take any value, such as 1.23, but words only occur in whole units. This correction need not be applied to numbers over about ten but when there are, as in this case, an appreciable proportion of figures less than ten, the differences to be squared in the calculation of chi squared should be each reduced by one half, i.e. 0.5, in this instance from 8.6 to 8.1.

Chi squared is 12.8 for 8 degrees of freedom and chance would create

differences as large as these in one trial in six. The hypothesis that Herodotus wrote the samples with a consistent rate of using *alla* can be accepted, the differences observed between books can be explained by chance.

Table 1B shows a comparable count from each book of Thucydides in the Oxford Classical Text. In the table, occurrences are so few that to keep within the limit of five occurrences we need to add together four samples and so end up with only two samples each of 400 spaces, one having 10 and the other having 8 occurrences. The difference between samples would be created by chance in two out of three trials. Clearly Thucydides has a consistent but low rate of occurrence of *alla.*

Table 1C shows the results of a count in all the orations of Isocrates in the Loeb Text, which are long enough to have at least sixty occurrences of *kai* in them. In all the samples there are 1600 spaces and this number enables the distribution to be determined with some precision; the distribution is Poisson. In the set of samples all that calls for comment is that work 17 has an average low enough to be significant at the 5 % level. This is not surprising, for in a set of 16 samples one difference to be expected once in twenty times is hardly out of place, but it might suggest that the average for the forensic orations is lower than the average for the others. This is not so, the average for the four forensic orations is 0.195 *allas per space;* for the others it is 0.248, a difference which is not statistically significant.

Table 1D records the results of a count in the first ten orations of Demosthenes in the Oxford Classical Text. Included in the set are two, works 7 and 10, thought to be spurious. The average rate of occurrence for the eight genuine orations is 0.172 allas per space, neither work 7 nor work 10 shows statistically significant differences.

Table 1E. In this is set out the results of a count of all the orations of Lysias in the Oxford Classical Text which have more than 60 spaces between *kais*. For the ten works, chi squared is 23.88 for nine degrees of freedom, a difference significant at the 1 % level. Most of the difference arises from work 13; if this is removed from the group no significant differences remain. The differences in work 13 are not due to some feature of the particular sample, the count carried right through the oration shows exactly the same result. That the differences in the occurrence of alla do indicate a difference in authorship can be inferred from other tests in this set and also from independent evidence, for which see K.J. Dover, Lysias and the Corpus Lysiacum, University of California Press, Los Angeles 1968.

Table 1F shows a count in four samples of Plutarch, in the Loeb Text, samples which resemble each other so closely that no calculation is needed to show their homogeneity.

Table 1G sets out the results of a count in six works of Xenophon and in

the spurious Cynegeticus all in the Loeb Text. The differences between the genuine works, if Hiero is excluded, are very small and would arise by chance in more than nine trials in ten, but if Hiero is added statistically significant differences at once appear. The reason is that Hiero is cast in the form of a conversation and almost every occurrence of ***alla*** spoken by one participant is repeated by the other in the course of his reply. From the genuine continuous prose works, Cynegeticus cannot be distinguished by this test.

The samples in Table 1H are from Josephus, again in the Loeb Text. For the five samples chi squared is 7.9 for four degrees of freedom, the differences are not statistically significant.

For the six samples of Aristotle from the Oxford Classical Text in Table 1I chi squared is 5.2 for five degrees of freedom and so the differences are not statistically significant.

The four samples from Diodorus Siculus in the Loeb Text in Table 1J have few occurrences in them and when pairs of samples are added to give the occurrence in 200 spaces there are six occurrences in both samples.

The four samples of Philo Judaeus in Table 1K, taken from the Loeb Text, also show no statistically significant differences.

As was earlier stated, it is impracticable to submit to an editor a paper which contains seven tables each as extensive as table one. The authors therefore decided to print one survey in full and then illustrate the others from a count in Demosthenes 18th Oration, De Corona. This work has 1131 occurrences of *kai* and so 1130 spaces between them. The spaces are grouped in eleven hundreds with a remainder of thirty. The data for all eight words in this oration are set out in *table two.*

The data can be taken in two groups. For six of the seven words there are not enough occurrences to avoid the necessity of grouping together all spaces with one or more occurrences and so the data becomes an 11 x 2 table; eleven samples with two cells, one for no occurrence, the other for one or more occurrences. Only the particle *de* occurs often enough to make an 11 x 3 table with cells of no occurrence, one occurrence and more than one occurrence. For the occurrence of *alla* and *gar* chi squared is 5.2 and 4.55 respectively. There are no statistically significant differences in either table. Nothing is to be gained by lumping samples together.

For *dia* the occurrences are so few that the minimum sample needed to ensure an average of five occurrences is one of two hundred spaces. For the first four such samples chi squared is 1.46, not significant.

For *men, me* and *ou* the data is treated as 11 x 2 table and chi squared is 11.06; 9.41 and 2.82, in no case statistically significant. The occurrences of

de are frequent enough to enable the distribution to be determined and this is shown, in Table 1C, to be negative binomial indicating a slight tendency for the occurrence to bunch together. The tendency is slight for the observed figure for four or more occurrences is below expectation. For the 11 x 3 table of the occurrences of de, chi squared is 20.4 for 18 degrees of freedom, the differences do not reach the 5 % level of significance, they would appear by chance in 30 % of the trials of a random experiment.

The survey just ended, and the others just illustrated, show that a selection of writers of Greek have consistent habits represented by the occurrence of the seven words in this group. The habits are not affected by periods of time, more than sixty years separates some of the works of Isocrates which were examined, nor by the range of subject matter and literary form used by Lysias or Demosthenes. The sole example of a statistically significant difference between two works by the same author came with the comparison of some continuous prose of Xenophon with a conversational exchange between named characters, a comparison which affected the occurrence of one of the words in the group. In most cases of disputed authorship the uncertainty concerns works which might be early or late additions to a central corpus and the conditions which surround them are much less taxing than those covered in the survey.

Objections are offered to methods like that employed here. The most common objection is based upon the fact that occurrence alone is recorded, that the meaning of the words, their context and their usage are completely ignored. It is true that occurrence alone is recorded but occurrence alone is relevant to this method. Occurrence is not misleading. It was earlier mentioned that the sum of two Poisson distributions is another Poisson distribution and the sum of a number of Poisson distributions with variable expectations was a negative binomial distribution. A consequence of these facts is that if, for example, the occurrences of *dia* included ten separate classes which could be distinguished on linguistic or philological or other grounds, then each class would form a Poisson distribution and all of them together would form either a Poisson distribution or a negative binomial distribution. There is no question of one classification masking another and giving rise to a misleading result.

It should also be noted that, for example, K.J. Dover in Greek Word Order, Cambridge, 1960 begins by quoting a sentence of two words and then lists ten reasons which might be considered as having determined the order of the two words. This illustrates the fact that when we record the occurrence of an event we are doing something which is certain; when we begin to classify the occurrences in terms of reasons or motives we are on much more debateable ground.

Another type of objection will argue that if there is an anomaly in the

distribution of the mark-word, say kai, then all seven tests based upon the count-words will be misleading. This objection is based on a misunderstanding of the nature of the negative binomial distribution. All negative binomial distributions have similar shapes, like a long rope hanging loosely from a flag-pole, and they are remarkably insensitive to change. There is no statistically significant difference in the occurrences of *alla* between two works of Demosthenes one of which has just over 3 % of the text made up by occurrences of *kai* and another which has just over 6 % of occurrences of *kai*. For this objection to be valid it would need alterations in the rate of occurrence of the mark-word greater than anything yet observed in any writer. Even so it is sometimes quite impossible to alter the text to give the desired result. The occurrences of *de* in the spaces between *kai* in Hebrews give, for a comparison with the expectation derived from 1st Corinthians and Galatians, for no occurrence observed, 188, expected, 146.6; for one occurrence observed, 40, expected, 47.0; for more than one occurrence observed, 12, expected, 46.3. If this result is to be attributed to some anomaly in the occurrence of *kai*, then the deletions required to get one end of the distribution right will exaggerate the difference recorded at the other end.

In the unlikely event of a real anomaly in the marker word being undetected, the anomaly will only affect the single set of tests for they can all be inverted and converted. They can be inverted, instead of counting the occurrences of *gar* between *kai* we can count the occurrences of *kai* between

gar. They can be converted, instead of counting the occurrences of *gar* and *de* between occurrences of *kai*, we can count occurrences of *gar* between *de*, or of *de* between *gar*.

This set of tests covers a network of words which occur 16,910 times in the New Testament and are 12.3 % of the whole text; they can be arranged in any convenient pattern. The most efficient use is to make the most frequent word the mark-word but it is permissible to use any word as the mark-word and any other as the count-word. The main value of conversions is that they can supply quite dramatic illustrations; the number of occurrences of *kai* between *de* is around 3 for the Pauline epistles, around 6 in Ephesians, 10 in Philippians and 20 in Colossians.

The minimum sample size is determined by the necessity of having an expectation of five occurrences of the count-word in spaces between mark-words. Experimental evidence suggests that samples of as few as twenty spaces which meet this requirement do not show any periodic effects which could be misleading.

A case of disputed authorship; The Pauline Epistles of The New Testament

The Pauline Corpus is too well-known to need any detailed introduction. In this paper, as in all the earlier work of the authors, the Apostle Paul

is defined to be the writer of the epistle to the Galatians and all references to him conform to this axiomatic assumption.

Galatians

In all previous tests Galatians has been found to be entirely homogeneous. With 72 occurrences of kai in the text, the epistle has 71 spaces between occurrences. The data for the seven words are shown in Table 3A. It is possible to divide the epistle into sub-samples, keeping within the rule of not less than five occurrences per sample, into three sub-samples for occurrences of ***alla, gar, dia, me*** and ***ou*** and into six sub-samples for occurrences of ***de***. Samples as short as these will not distinguish one author from another but they will reveal any periodic effects which might exist and so establish the minimum sample size. In none of these divisions do statistically significant differences appear and so it can be assumed that, for this author, the minimum sample size is fixed by the necessity of having an expectation of five occurrences in the sample and is typically about twenty spaces.

1st Corinthians

The previous test showed this to be a Pauline text with two anomalies at

7.25-39, the advice to widows; and at 8.1-13, remarks on food offered to idols. These passages differ from the free composition of the rest of the epistle. Whether the differences are due to the nature of the subject matter, in both cases requiring logical lists, or are due to the use of some underlying source such as a general ecclesiastical pronouncement, only further comparisons with similar material could decide.

1st Corinthians has 277 occurrences of kai and so 276 spaces between them. The epistle is divided into three successive samples of 71 spaces, each equal to Galatians, numbered S1, S2, and S3, and the remainder, R, with 63 spaces. After the examination of the epistle, all the samples are added, with an exception noted later, to Galatians to give a single sample of 347 spaces. The other epistles of the Corpus are first compared with expectations derived from this joint sample of Galatians plus 1st Corinthians.

For all the words but *me* and *ou*, no statistically significant differences appear and, as can be seen from table 4, the data fit either a Poisson or negative binomial distribution. In the occurrence of *me* there is a real anomaly. The data for the combined Galatians - 1st Corinthians sample has a mean of 0.349 and a variance of 0.528. The negative binomial prediction for this mean and variance is : no occurrence 261.8, one occurrence 60.4, more than one occurrence 24.8; not a good fit for the data. The discrepancy is due to a statistically significant difference in sample S2, 6.11 to 10.13 of the text, which has too many single occurrences. There is also a statistically

significant difference, due to a lack of occurrences, in the last samples, R.

For such significant differences there are two alternative explanations. They can be due to chance. We are looking at a set of seven words in five samples and so have 35 figures; it is reasonable to expect that two of them will show differences significant at the 5 % level. If this is the correct explanation of the differences, then the single sample in which the difference appears is not likely to produce a statistically significant difference in the set of samples which show the occurrence of the word in the epistle as a whole and it will also be found that the difference does not appear in the occurrence of any other words. Sample R meets both these conditions but S2 does not and so the alternative explanation of a real anomaly to be explained by a difference of authorship, including references or quotations, must be examined.

In the occurrence of *ou* there is no statistically significant difference which affects a whole sample but in S2 there is a single cell, that of single occurrences, for which chi squared is 3.91 for one degree of freedom. This suggests that the anomaly in S2 is a real one, i.e. not simply due to chance, especially as the occurrences of both *me* and *ou* are high and it is not possible to argue that the use of one has been an alternative to the other. Of this anomaly the passage 7.21-34, part of the advice to widows, is a permissible and practicable explanation.

To return to the last sample in the occurrence of *me*, it might be argued that this is due to some end-effect in the epistle. In the simplest form this argument is untenable for the final request and greetings affect only the last 8 spaces in the sample. But neither Galatians nor 2nd Corinthians show any sign that the end of an epistle differs from the rest of the text.

The final conclusion is that, as in previous examinations, 1st Corinthians is entirely Pauline with a minor anomaly in S2, an anomaly which can be explained by the passage giving advice to widows.

Romans

Though Romans is substantially Pauline, the structure of the epistle is not simple. Neither chapter one, nor chapter fifteen nor sixteen can be considered Pauline and there are other anomalies some of which are doubtless due to the number of quotations in the text. The epistle is divided into six samples : chapter one; two samples of 71 successive spaces S1 and S2; a remainder R of fifty spaces; chapter 15 and chapter 16. Table 3B sets out the data for the epistle.

Chapter one shows a statistically significant difference in the occurrence of *me;* chapter 15 has statistically significant differences in the occurrence of *alla, me* and *ou;* chapter sixteen in the occurrence of *gar, me* and *ou.* S1

has statistically significant differences in the occurrence of *alla, de* and *me,* S2 in the occurrence of *alla, gar, dia* and *ou.* Sample R has a solitary significant difference in the occurrence of *gar.*

It is clear that Romans is not a free composition of the Apostle Paul, there are two different kinds of anomaly in the text, one represented by the first and last two chapters and the other by passages in the body of the epistle.

2nd Corinthians

2nd Corinthians has long been known to be an epistle fabricated from Pauline material. The analysis of such a document must be done with reserve for it is quite possible that some re-ordering of the text has taken place. Wake, *ref. 5* showed that the first chapter of the epistle is distinct from the remainder and offered a suggestion that the first chapter was very like the first chapters of 1st and 2nd Timothy. Other slight anomalies have been noted between chapters 2 and 7 but chapters 10 - 13 have been homogeneous.

The epistle is divided into three samples : chapter one; the text from the beginning of chapter two to the end of chapter nine, omitting 6.14 - 7.1; and chapters ten to thirteen. The data are set out in Table 3C. Chapter one shows statistically significant differences in the occurrence of *de* and *me;*

S1 in the occurrence of *alla* and *dia;* chapters 10 - 13 in the occurrence of *alla, de* and *ou.* The differences in chapter one support the previous conclusion that this passage is no longer in the form in which it left the Apostle's hand; the differences in S1 might be due to redaction or to the re-arrangement of the text; the differences in chapters 10 - 13 may well have the same origin, the revision of a text by omission or insertion.

Ephesians, Philippians and Colossians

Table 3D sets out the data for these three epistles which are treated together for no other reason than that they are isolated documents of about the same length and so present, in statistical terms, much the same problem. Ephesians has been shown to be composite, a view supported by the fact that in both the occurrence of *alla* and of *de* the first half of the spaces, up to 4.17 of the text, show statistically significant differences from the remainder of the epistle.

All three epistles are quite dramatically different from the Pauline expectations. They are much further from the Pauline Expectations than samples from Hebrews for example. Only in the occurrence of *dia* is there no significant differences between these epistles and the Pauline Expectations. When Philippians is compared with Colossians there are statistically significant differences in the occurrence of *alla,* chi squared is 3.34 for one degree of

freedom and in the occurrence of *de,* chi squared 9.87 for two degrees of freedom. The conclusions are that none of the epistles is Pauline and that Ephesians is composite.

1st and 2nd Thessalonians

In previous tests these epistles have not been separable from each other though not Pauline. Their non-Pauline nature is clearly shown by the statistically significant differences which appear in the occurrence of all seven words when they are compared with the Pauline expectations. Even in the rare occurrences of *men,* the rate in 2nd Thessalonians differs significantly from the Pauline expectation.

A new piece of evidence comes to light when comparison is made of the two epistles for they do differ in the occurrence of *gar,* chi squared is 10.31 for 1 degree of freedom with Yates' correction applied. The statistics will not disclose whether the difference is due to the enhanced use of *gar* in 1st Thessalonians or to the avoidance of its use in 2nd Thessalonians, but the absence of the word from the first twenty spaces of 2nd Thessalonians up to verse 4 of chapter 2, has no parallel in the rest of the Epistle nor in all of 1st Thessalonians.

1st and 2nd Timothy and Titus

The Pastoral Epistles have been shown by previous tests to be indistinguishable from each other and to be non-Pauline.

Only in the occurrence of ***alla*** for all three epistles, and of ***dia*** for 2nd Timothy and Titus are there no statistically significant differences from the Pauline expectations. It is not possible to distinguish Timothy from 1st or 2nd Timothy.

The comparison of the 1st chapters of 2nd Corinthians and of Romans with the first chapters of 1st and 2nd Timothy shows that they are similar, no statistically significant differences appear and if the samples are rather short, 31, 24, 17 and 18 spaces respectively, both the first chapters of 2nd Corinthians and of Romans do differ from the rest of the epistles to which they are attached.

Philemon

With only 18 occurrences of ***kai*** in it this epistle offers a sample too short to be distinguished from any other epistle.

Hebrews

This epistle has now been generally recognised as non-Pauline though in many of the habits tested by the methods used in this paper it is much nearer to the Pauline epistles than others, such as Ephesians, Philippians or Colossians which are still accepted.

Wake showed that the last chapter of the epistle differed from the rest of the text. There are 255 occurrences of ***kai*** in the epistle and so 254 spaces. The first occurrence of ***kai*** in chapter 13 is the 243rd and the epistle was therefore divided into three samples, two of 120 successive spaces and chapter 13 with 12 spaces. As no statistically significant difference appeared in any of the occurrences in the two samples these were merged to give a single sample of 240 spaces. In all occurrences but that of ***gar*** and ***men*** there are significant differences from the Pauline expectations.

In the occurrence of ***gar*** there is a statistically significant difference between chapter 13 and the rest of the epistle. The difference is large, chi squared is 8.11 for one degree of freedom, and is due to the large number of occurrences in chapter 13, a fact which would rule out any attempt to explain the difference as due to the extreme brevity of the sample. In the occurrence of ***gar*** chapter 13 does not differ from the Pauline expectation.

The conclusion is that Hebrews is not Pauline, though there is no obstacle

to the hypothesis that chapter 13 might be Pauline, and the hypothesis that chapter 13 is distinct from the rest of the epistle gains support.

Simplification

When an investigation has been made of the distribution of occurrences and the form of the distributions has been established, it is often possible to suggest simplifications which will not be misleading. Seven words have been looked at in this paper and only in two cases, of *de* and *ou*, was it possible to treat the occurrences in three groups, spaces with no occurrences, spaces with one occurrence and spaces with more than one occurrence. If these were treated, as all the other occurrences had to be treated, as simple binomial choices, spaces with no occurrence and spaces with one or more occurrences, little would be lost. In the Pauline Corpus some of the statistically significant differences would diminish when the data is reduced in this way but none would disappear.

Illustration

The wide-spread use of instruments with dials and pointers is evidence of how much easier most people find it to interpret a picture of data than the raw data itself. A simple and useful illustration of the test just described

can be made as follows. In the case of the Pauline Corpus Titus and Philemon which are too short for useful illustration are omitted as are the occurrences of *men* which are too rare to be conveniently shown. This leaves six characteristics, the average rates of occurrence for the remaining six words, to be shown in twelve epistles. Table five shows the average rates of occurrence for all six words. Twelve points are picked out on a sheet of paper one for each epistle. These can be thought of as the centres of twelve clock faces. Each character will be evenly spaced round the dial and with six characters an obvious choice is to position them at 12, 2, 4, 6, 8 and 10 o'clock.

The occurrence of *alla* will be measured along the twelve o'clock radius. From Table 5 the average for Romans is 25 units, to clear the centre of the circle it is wise to add a constant to every measurement and for a sheet of A4 size a feasible scale is to add five to the units then divide by two and count off the distance in millimetres. Thus for *alla* in Romans we have 25 plus 5 equals 30; divide by 2, this gives 15, so measure 15 millimetres out from the centre along the twelve o'clock radius and mark this point. Repeat for all six characters e.g. gar will be 53 plus 5 equals 58, divided by two equals 29, mark at 29 millimetres from the centre along the 2 o'clock radius, and so on. This gives a set of six points for each epistle which are then joined by a firm line to make a polygon. As all measurements have been made on the same scale, identical epistles would be the same size and shape. Between similar epistles there should

be minor variations in size and shape but quite different sizes and shapes show real differences between epistles.

From *Figure 1* drawn as described, the totally different characters of the various groups of epistles can be seen.

Dept. of Computer Science
University of Edinburgh
8 Buccleuch Place
EDINBURGH 8 (Scotland)

Professor S. MICHAELSON
and
Rev. A.Q. MORTON

REFERENCES

1. *The Poisson distribution* is described in almost all textbooks of statistics, for example in J.F. KENNEY and E.S. KEEPING, *The Mathematics of Statistics,* Princeton, 1954, Vol. 1, sec. 10.6.

2. *The negative binomial distribution* is not yet to be found in the majority of statistical textbooks. For the first literary application see : - F. MOSTELLER and D.L. WALLACE, *Inference and Disputed Authorship : The Federalist, Reading,* Massachusetts, 1964. For a simple explanation of the distribution with a fully worked example see : - R. GOODMAN, *Teach Yourself Statistics,* London, 1966.

3. For the spaces between kais in the Pauline epistles see : - A.Q. MORTON and James McLEMAN, *Paul the Man and the Myth,* London and New York, 1965, pp. 68-70. For an example of the negative exponential distribution see : - K.A. BROWNLEE, *Industrial Experimentation,* H.M.S.O., London, 1957.

4. *The chi squared* test is described in all textbooks of statistics, for example in KENNEY and KEEPING, Vol. 2, sec. 3.

5. M. LEVISON, A.Q. MORTON and W.C. WAKE, *Some Statistical Features of the Pauline Epistles, Journal of the Royal Philosophical Society*, July, 1966.

Table One.

The occurrence of alla in a selection of Greek writers.

(A.) Herodotus (Ox. Cl. Text.)

No. of Occ. of alla in space.	Number of spaces in sample which have X occurrences. N.									
X	Books 1	2	3	4	5	6	7	8	9	Total
0	91	94	92	95	94	96	85	91	85	823
1	9	5	7	4	5	4	15	8	15	73
2	-	1	1	1	-	-	-	-	-	4
No. of spaces in sample	100	100	100	100	100	100	100	100	100	900

(B.) Thucydides (Ox. Cl. Text.)

X	Books 1	2	3	4	5	6	7	8	Total
0	97	97	98	98	97	100	98	97	782
1	3	3	2	2	3	-	2	3	18
No. of spaces in sample.	100	100	100	100	100	100	100	100	800

Table One (Contd.)

(C.) 1. The Orations of Isocrates - Loeb Text.

No. of Occ. of alla in space.	Number of spaces in sample which have X Occurrences. N.																	
X Books	2	3	4	5	6	7	8	9	10	11	12	13	14	16	17	18	19	Total
0	83	81	82	79	80	78	78	82	86	78	75	44	72	88	90	76	81	1289
1	13	15	15	18	17	17	20	17	11	15	23	13	24	10	7	19	16	257
2	1	3	2	3	2	3	1	1	3	7	2	1	3	2	3	5	3	44
3	2	1	1	-	1	2	1	-	-	-	-	-	1	-	-	-	-	9
4	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
	100	100	100	100	100	100	100	100	100	100	100	58	100	100	100	100	100	1600

Oration 15 omitted.

(C.) 2.

The observations for the Orations of Isocrates fitted to a Negative Binomial Distribution.

No. of Occurrences of alla.	Number of spaces with X occurrences.	
X	Observed	Expected In a N. B. Distribution with the same mean and variance.
0	1289	1288.0
1	257	259.9
2	44	44.9
3	9	9.1
4	1	
	1600	1599.9

Table One (Contd.)

(D.) Demosthenes. (Ox. Cl. Text.)

No. of occ. of alla.	Number of spaces in sample which have X occurrences.								
X	Orations								
	1	2	3	4	6	7	8	9	10
0	88	90	82	86	90	77	87	89	83
1	7	10	15	14	9	19	12	10	12
2	-	-	3	-	1	3	1	1	5
3	1	-	-	-	-	-	-	-	-
4	-	-	-	-	-	1	-	-	-
No. of spaces in sample	100	100	100	100	100	100	100	100	100

(E.) Lysias. (Ox. Cl. Text.)

No. of occ. of alla.	Number of spaces in sample which have X occurrences.									
X	Orations									
	1	2	3	5	7	12	13	14	18	19
0	88	85	90	80	77	75	94	82	53	78
1	10	14	9	17	8	21	5	17	14	19
2	2	1	1	2	-	3	1	1	-	2
3	-	-	-	1	1	1	-	-	-	1
No. of spaces in sample	100	100	100	100	100	100	100	100	67	100

(F) Plutarch. Loeb Text.

No. of occ. of alla.	Number of spaces in sample which have X occurrences.			
X	On Exile.	Old Men.	Fortune of The Romans.	Divine Vengeance.
0	86	86	183	88
1	14	14	16	11
2			1	1
No. of spaces in sample.	100	100	200	100

Table One (Contd.)

(G.) Xenophon. Loeb Text.

No. of occ. of alla.	Number of spaces in sample which have X occurrences.						
X	Hiero	Agesilaus	Const. of The Lacedaemonians	Ways & Means	Cav. Com.	Horse-manship	Cyneget=icus.
0	84	91	96	93	96	93	94
1	15	8	4	6	4	6	6
2	1	1	-	1	-	1	-
No. of Spaces in Sample	100	100	100	100	100	100	100

(H.) Josephus. Loeb Text.

No. of occ. of alla.	Number of spaces in sample which have X occurrences.				
X	Jewish War Bk.1	Bk.5	Life Bk.1	C. Apion Bk.1	Bk.2
0	94	95	94	86	70
1	6	5	5	14	11
2	-	-	1	-	-
No. of spaces in sample.	100	100	100	100	81

Table One (Contd.)

(I.) Aristotle. Loeb Text.

No. of occ. of alla.	Number of spaces in sample which have X occurrences. N.					
X	De Caelo				Parts of Animals	Progression of Animals
	Bk.1	2	3	4		
0	94	87	88	87	93	181
1	5	12	9	13	7	17
2	1	1	3	-	-	2
No. of Spaces in Sample	100	100	100	100	100	200

(J.) Diodorus Siculus. Loeb Text.

N.

X	Bk.1	2	21	22
0	95	99	95	99
1	5	1	5	1
No. of Spaces in Sample.	100	100	100	100

(K.) Philo Judaeus.

X	On The Creation	Allegorical Inter.	Cherubim	Cain & Abel
0	92	92	93	94
1	6	8	7	6
2	2	-	-	-
No. of spaces in sample.	100	100	100	100

Table Two.

The occurrence of a group of words in Demosthenes 18, De Corona.

1. alla.

No. of occ. in space. | Number of spaces in sample with X occurrences.

X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	87	89	90	86	82	88	85	84	90	85	87	25	978
1	13	9	10	11	14	11	12	12	9	13	12	5	131
2	-	1	-	2	1	1	2	3	1	2	-	-	13
3	-	1	-	1	2	-	1	1	-	-	1	-	7
4					1								1
													1130

2. gar.

X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	87	85	87	85	81	84	85	87	86	88	90	26	971
1	12	12	13	15	16	14	12	10	12	8	10	4	138
2	-	3	-	-	3	2	3	3	2	4	-	-	20
3	1												1
													1130

3. de.

X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	64	71	81	77	76	72	72	68	80	75	79	20	835
1	22	23	15	16	18	16	19	22	17	19	16	6	209
2	8	3	4	6	3	9	4	7	2	4	4	3	57
3	6	2	-	-	1	2	4	3	-	2	1	1	22
4	-	1	-	1	-	1	1	-	-	-	-	-	4
5	-				1								1
6	-				1								1
9	-								1				1
													1130

Table Two (Contd.)

4. dia.

No. of occ. in space. Number of spaces in sample with X occurrences.

X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	97	95	99	95	92	98	99	92	91	99	94	30	1081
1	2	5	1	5	8	2	1	8	9	1	5	-	47
2	1	-	-	-	-	-	-	-	-	-	1	-	2
													1130

5. men.

X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	78	83	85	83	79	75	79	77	82	88	86	24	919
1	16	14	15	12	15	20	16	20	16	10	14	6	174
2	6	3	-	5	5	4	4	3	2	2	-	-	34
3	-	-	-	-	-	1	1	-	-	-	-	-	2
4	-	-	-	-	1	-	-	-	-	-	-	-	1
													1130

6. me.

X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	93	94	95	92	91	90	88	93	95	92	97	29	1049
1	7	6	4	7	8	10	11	6	4	7	3	1	74
2	-	-	1	1	1	-	1	1	1	1	-	-	7
													1130

Table Two (Contd.)

7. ou.

No. of occ. in Space.	Number of spaces in sample with X occurrences.												
X	1	2	3	4	5	6	7	8	9	10	11	R	Total
0	82	81	92	81	77	86	79	79	84	78	85	26	930
1	14	17	8	13	13	11	16	16	14	18	11	3	154
2	3	1	-	4	7	2	4	4	-	3	3	-	31
3	-	1	-	1	2	-	1	1	1	1	-	-	8
4	1	-	-	-	1	1	-	-	-	-	-	1	4
5	-	-	-	1	-	-	-	-	-	-	-	-	1
7	-	-	-	-	-	-	-	-	1	-	1	-	2
													1130

8. The occurrence of de in Demosthenes 18.

The observed data fitted to a negative binomial distribution with the same mean and variance.

No. of occ. in Space.	Number of spaces with X occurrences.	
X	Observed	Expected
0	835	841.9
1	209	197.8
2	57	60.5
3	22	18.4
4 or more	7	11.4
Totals.	1130	1130.0

For this table chi squared is 3.16 for two degrees of freedom.

Table Three.

The occurrence of seven words in the spaces between occurrences of Kai in The Pauline Corpus.

(A) Galatians & 1st. Corinthians.

1. alla.

No. of occs. in space.		Number of spaces in Sample with X occurrences. N.					
		1st. Corinthians.					
X	Gal.	S_1	S_2	S_3	R.	Total	G. & 1C.
0	52	57	55	57	56	225	277
1	16	8	11	10	4	33	49
2	3	5	4	3	3	15	18
3	-	1	1	1	-	3	3
Total	71	71	71	71	63	276	347

2. gar.

X	Gal.	1st. Corinthians. S_1	S_2	S_3	R.	Total	G. & 1C.
0	53	48	57	45	45	195	248
1	7	18	10	24	14	66	73
2	7	1	3	2	3	9	16
3	3	3	-	-	1	4	7
4	1	1	-	-	-	1	2
5	-	-	-	-	-	-	-
6	-	-	1	-	-	1	1
Total	71	71	71	71	63	276	347

Table Three (Contd.) (A) Galatians & 1st Corinthians (Contd.)

3. de.

No. of occs. in space.	Number of spaces in Sample with X occurrences. N.					
		1st. Corinthians.				
X	Gal.	S_1	S_2	S_3	R.	Total	G. & 1C.
0	42	47	46	42	36	171	213
1	15	11	8	17	17	53	68
2	6	8	8	2	4	22	28
3	5	3	7	3	4	17	22
4	3	2	1	3	2	8	11
5	1	-	-	2	-	2	3
6	-	-	-	1	1	-	2	2
9	-	-	-	1	-	1	1
Total	71	71	71	71	63	276	347

4. dia.

		1st. Corinthians.					
X	Gal.	S_1	S_2	S_3	R.	Total	G. & 1C.
0	56	62	61	63	58	244	300
1	14	5	9	6	5	25	39
2	-	4	1	1	-	6	6
3	1	-	-	-	-	-	1
4	-	-	-	1	-	1	1
Total	71	71	71	71	63	276	347

5. men.

		1st. Corinthians.					
X	Gal.	S_1	S_2	S_3	R.	Total	G. & 1C.
0	69	64	69	64	61	258	327
1	1	7	1	6	2	16	17
2	1	-	1	1	-	2	3
Total	71	71	71	71	63	276	347

Table Three (Contd.) (A) Galatians & 1st Corinthians (Contd.)

6. με.

No. of occs. in space.	Number of spaces in Sample with X occurrences. N.						
		1st. Corinthians.					
X	Gal.	S_1	S_2	S_3	R.	Total	G. & 1C.
0							
1	53	52	39	54	56	201	254
2	15	16	25	14	6	61	76
3	2	3	5	1	1	10	12
4	-	-	2	1	-	3	3
5	1	-	-	-	-	-	1
7	-	-	-	1	-	1	1
Total	71	71	71	71	63	276	347

7. ου.

X	Gal.	1st. Corinthians. S_1	S_2	S_3	R.	Total	G. & 1C.
0	49	50	42	50	48	190	239
1	11	10	20	12	10	52	63
2	10	6	4	4	3	17	27
3	-	4	4	2	1	11	11
4	-	-	-	-	1	1	1
5	1	1	-	1	-	2	3
6	-	-	1	-	-	1	1
7	-	-	-	1	-	1	1
8	-	-	-	1	-	1	1
Total	71	71	71	71	63	276	347

Table Three (Contd.)

(B) Romans.

1. Alla.

No. of occs. in space.	Number of spaces in Sample with X occurrences. N.					
X	Chapt. 1.	S_1	S_2	R.	Chapt. 15	Chapt. 16
0	22	55	47	41	35	26
1	2	15	20	7	2	2
2	-	1	3	1	-	-
3	-	-	-	-	-	-
4	-	-	-	-	-	-
5	-	-	1	1	-	-
Total	24	71	71	50	37	28
Chi Squared	2.04	0.25	8.24	0.15	5.06	3.01
Degs. of Freedom.	1	1	1	1	1	1

2. gar.

X	Chapt. 1.	S_1	S_2	R.	Chapt. 15	Chapt. 16
0	19	51	41	29	30	25
1	1	11	15	12	6	3
2	4	9	8	5	1	-
3	-	-	5	3	-	-
4	-	-	-	1	-	-
5	-	-	1	-	-	-
6	-	-	-	-	-	-
7	-	-	1	-	-	-
Total	24	71	71	50	37	28
X^2	0.67	.01	6.64	4.52	1.63	4.38
D/F.	1	1	1	1	1	1

Table Three (Contd.) (B) Romans (Contd.)

3. de.

No. of occs. in space.	Number of spaces in Sample with X occurrences. N.					
X	Chapt. 1.	S_1	S_2	R.	Chapt. 15	Chapt. 16
0	21	49	40	35	28	25
1	3	17	17	9	7	2
2	-	3	5	3	2	-
3	-	2	4	2	-	1
4	-	-	2	-	-	-
5	-	-	-	1	-	-
6	-	-	3	-	-	-
7	-	-	-	-	-	-
Total	24	71	71	50	37	28
X^2	0.69	6.93	0.97	2.11	0.63	1.32
D/F.	1	2	2	2	1	1

4. dia.

X	Chapt. 1.	S_1	S_2	R.	Chapt. 15	Chapt. 16
0	22	48	55	42	31	27
1	1	11	11	5	4	1
2	1	9	4	3	2	-
3	-	2	-	-	-	-
4	-	1	1	-	-	-
Total	24	71	71	50	37	28
X^2	-	21.62	4.94	0.25	0.23	-
D/F.	-	1	1	1	1	-

5. men.

0	23	66	64	46	37	28
1	1	5	7	3	-	-
2	-	-	-	1	-	-
Total	24	71	71	50	37	28

Table Three (Contd.) (B) Romans (Contd.)

6. me.

No. of occs. in space. X	Chapt. 1.	Number of spaces in Sample with X occurrences. N. S_1	S_2	R.	Chapt. 15	Chapt. 16
0	23	59	53	33	35	28
1	1	8	13	11	2	-
2	-	4	4	3	-	-
3	-	-	1	1	-	-
4	-	-	-	-	-	-
5	-	-	-	-	-	-
6	-	-	-	1	-	-
7	-	-	-	1	-	-
Total	24	71	71	50	37	28
X^2	7.06	3.52	0.07	1.32	8.60	10.24
D/F.	1	1	1	1	1	1

7. ou.

X	Chapt. 1.	S_1	S_2	R.	Chapt. 15	Chapt. 16
0	20	46	41	41	-	-
1	3	21	17	5	-	-
2	1	3	7	3	-	-
3	-	-	1	-	-	-
4	-	-	3	1	-	-
5	-	1	1	-	-	-
6	-	-	-	-	-	-
7	-	-	-	-	-	-
8	-	-	1	-	-	-
Total	24	71	71	50	-	-
X^2	1.22	12.21	6.62	-	-	-
D/F.	1	2	2	-	-	-

Table Three (Contd.)

(C) 2nd Corinthians.

1. alla.

X	Ch.1	S_1	S_2
0	28	68	47
1	1	19	7
2	1	3	3
3	1	2	1
4	-	2	-
7	-	1	-
Total	31	95	58
X^2	2.17	3.97	.06
D/F.	1	1	1

2. gar.

X	Ch.1	S_1	S_2
0	26	70	39
1	5	17	10
2	-	6	7
3	-	2	2
4	-	-	-
7	-	-	-
Total	31	95	58
X^2	2.29	0.23	0.53
D/F.	1	1	1

3. de.

0	26	67	44
1	5	18	9
2	-	8	4
3	-	-	-
4	-	1	1
5	-	1	-
Total	31	95	58
X^2	6.84	5.14	6.03
D/F.	1	1	1

4. dia.

0	26	71	52
1	2	21	5
2	3	3	1
3	-	-	-
Total	31	95	58
X^2	-	11.33	0.49
D/F.	-	1	1

5. men.

0	31	92	53
1	-	3	5

Table Three (Contd.)

6. me.

X	Ch.1	S_1	S_2
0	30	72	45
1	1	20	10
2	-	1	2
3	-	2	1
Total	31	95	58
X^2	8.77	0.34	0.60
D/F.	1	1	1

(C) 2nd Corinthians (Contd.)

7. ou.

X	Ch.1	S_1	S_2
0	25	68	32
1	4	20	14
2	1	4	6
3	1	2	4
4	-	1	2
Total	31	95	58
X^2	1.32	4.19	7.47
D/F.	1	2	2

(D) Ephesians, Philippians & Colossians.

1. alla.

X	E.	P.	C.
0	123	95	97
1	13	8	3
2	-	3	-
3	-	-	-
Total	136	106	100
X^2	9.59	4.77	18.25
D/F.	1	1	1

2. gar.

X	E.	P.	C.
0	127	94	94
1	8	11	6
2	-	1	-
3	1	-	-
Total	136	106	100
X^2	32.03	15.34	24.84

Table Three (Contd.)

(D) Ephesians, Phil. & Col. (Contd.)

3. de.

X	E.	P.	C.
0	120	84	95
1	14	19	5
2	1	2	-
3	1	1	-
Total	136	106	100
X^2	44.78	20.79	49.02
D/F.	1	1	1

4. dia.

X	E.	P.	C.
0	119	94	89
1	14	10	10
2	3	2	-
3	-	-	1
Total	136	106	100
X^2	0.13	0.42	0.53
D/F.	1	1	1

5. men.

X	E.	P.	C.
0	135	101	99
1	1	5	1
Total	136	106	100

6. me.

X	E.	P.	C.
0	123	100	93
1	11	6	6
2	1	-	-
3	1	-	-
4	-	-	1
Total	136	106	100
X^2	20.63	24.14	19.99
D/F.	1	1	1

7. ou.

X	E.	P.	C.
0	127	93	92
1	7	13	8
2	2	-	-
Total	136	106	100
X^2	32.64	18.49	20.11
D/F.	2	2	1

Table Three (Contd.)

(E) 1st & 2nd Thess., 1st & 2nd Timothy & Titus.

1. alla.

X	1 Thess.	2 Thess.	1 Tim.	2 Tim.	Titus.
0	91	84	81	55	32
1	7	4	10	12	4
2	1	-	1	-	-
3	-	-	-	-	-
4	1	-	-	-	-
Total	100	88	92	67	36
X^2	7.78	13.27	3.90	0.21	1.87
D/F.	1	1	1	1	1

2. gar.

0	79	84	79	54	30
1	19	3	13	12	6
2	2	1	-	1	-
Total	100	88	92	67	36
X^2	2.76	24.82	9.30	2.73	2.52
D/F.	1	1	1	1	1

3. de.

0	87	79	69	50	29
1	11	8	16	12	7
2	2	1	7	4	-
3	-	-	-	-	-
4	-	-	-	1	-
Total	100	88	92	67	36
X^2	30.26	35.83	8.58	6.95	5.73
D/F.	1	1	1	1	1

Table Three (Contd.) (E) 1st & 2nd Thess., 1st & 2nd Tim. & Titus.

4. dia.

X	1 Thess.	2 Thess.	1 Tim.	2 Tim.	Titus.
0	91	83	86	58	33
1	8	3	6	7	3
2	1	1	-	2	-
3	-	-	-	-	-
4	-	1	-	-	-
Total	100	88	92	67	36
X^2	1.73	4.63	3.81	.02	-
D/F.	1	1	1	1	-

5. men.

0	97	88	92	64	36
1	3	-	-	3	-
Total	100	88	92	67	36

6. me.

0	87	81	74	63	34
1	11	5	15	4	-
2	2	1	1	-	1
3	-	1	1	-	-
4	-	-	1	-	-
6	-	-	-	-	1
Total	100	88	92	67	36
X^2	9.71	15.96	2.42	18.02	8.36
D/F.	1	1	1	1	1

7. ou.

0	86	81	83	56	35
1	11	6	9	10	1
2	2	1	-	1	-
3	1	-	-	-	-
Total	100	88	92	67	36
X^2	12.04	19.09	18.59	8.24	11.77
D/F.	2	2	2	2	1

Table Three (Contd.)

(F) Hebrews.

1. alla.

X	S_1	Ch. 13.
0	226	11
1	13	1
2	1	-
Total	240	12
X^2	30.76	-
D/F.	1	-

2. gar.

X	S_1	Ch. 13.
0	179	4
1	44	6
2	15	2
3	2	-
Total	240	12
X^2	2.66	-
D/F.	1	-

3. de.

0	188	10
1	40	1
2	9	-
3	3	1
Total	240	12
X^2	27.62	-
D/F.	1	-

4. dia.

0	194	8
1	43	3
2	2	-
3	1	1
Total	240	12
X^2	6.60	-
D/F.	1	-

5. men.

0	222	12
1	17	-
2	1	-
Total	240	12
X^2	1.28	-
D/F.	1	-

6. me.

0	210	8
1	28	3
2	1	1
3	-	-
4	1	-
Total	240	12
X^2	25.00	-
D/F.	1	-

7. ou.

0	192	8
1	38	2
2	8	1
3	2	1
Total	240	12
X^2	16.90	-
D/F.	1	-

Table Four.

The fitting of distributions to the data of Table Three for the combined samples of Galatians and 1st Corinthians.

No. of occs. in space.	Expected numbers of spaces with X occurrences.						
X	alla	gar	de	dia	men	me	ou
0	274.0	250.7	206.0	292.7	323.4	244.5	240.8
1	56.9	66.6	75.4	49.7	1+ 23.6	85.6	61.6
2 or more	26.1	2+ 20.2	33.7	2+ 4.6		2+ 16.9	24.2
		3+ 9.5	31.9				3+ 20.4
Form of Dist.	N.B	N.B	N.B	P.D.	P.D.	P.D.	N.B.
Chi Sq.	2.12	1.55	3.05	4.71	0.59	1.53	0.64
D.F.	1	1	2	2	1	1	2

P.D. denotes Poisson distribution: N.B. negative binomial distribution. In all cases the distributions are calculated from the mean for the Poisson distributions; or the mean and variance - for the negative binomial distributions - taken from table three.

Table Five.

The Average rate of occurrences of six words in the Epistles of the Pauline Corpus.

Epistle	Alla	gar	de	dia	me	ou
Romans	25	53	54	53	28	44
1st Cor.	26	38	76	15	35	56
2nd Cor.	35	38	37	23	25	48
Gal.	31	51	83	27	34	51
Eph.	10	8	14	15	12	8
Phil.	13	12	26	13	6	12
Col.	3	6	5	14	10	8
1st Thess.	13	23	15	10	15	18
2nd Thess.	10	10	23	21	23	17
1st Tim.	13	14	33	7	26	10
2nd Tim.	18	21	35	18	6	18
Hebrews	6	36	26	22	16	26
All Epistles	17	26	40	18	20	26

The figures in Table five are averages multiplied by one hundred to get rid of decimals, i.e. the mean rate of occurrence of alla in Romans is 0.25 occurrences per space, multiplied by 100 this gives 25, as in column 1.

FIGURE ONE. THE RATE OF OCCURRENCE OF SIX WORDS IN THE EPISTLES OF THE PAULINE CORPUS.