Back to Part 1 of paper

7. THE EXCEPTIONALLY VARIABLE SETS IN CUTLER'S DATA

The records of the exceptionally variable sets of plates which occurred in Cutler's data, when identified by the method of the preceeding section, were studied individually with a view to gaining light upon the cause of their occurrence. As it is not necessary to reproduce the whole of the statistical tests which were applied, we shall confine ourselves to the main facts which emerged, and which served to justify the previous conclusions, as well as to indicate the nature of the disturbing cause.

The following facts appear to be unquestionable:—

(1) The proportion of exceptionally variable sets is the same for the sets of three, four and five plates in each portion of the total period.

(2) The proportion of exceptionally variable sets varies greatly at different periods, the exceptions occurring in well marked epidemics.

The evidence for these statements may be put in the form of a triple contingency table (see Fig. 2)

Table XI

	Excessively variable				Not excessively variable				Total				
Period	5	4	3	Total	5	4	3	Total	5	4	3	Total	χ^2
1	1	—	1	2	9	9	9	27	10	9	10	29	·967
2	3	2	1	6	7	12	6	25	10	14	7	31	1·728
3	—	—	1	1	12	18	4	34	12	18	5	35	6·176
		(10)		(14)		(11)		(20)					
4	3	9	1	13	5	12	4	21	8	21	5	34	·818
	(6)	(5)		(12)	(6)	(14)		(24)					
5	5	4	1	10	7	15	4	26	12	19	5	36	1·733
6	—	—	—	—	19	18	—	37	19	18	—	37	—
7	—	—	—	—	22	13	1	36	22	13	1	36	—
8	—	—	1	1	20	11	5	36	20	11	6	37	5·310
9	1	—	—	1	17	12	6	35	18	12	6	36	1·029
10	2	1	1	4	23	20	4	47	25	21	5	51	1·299
	(16)	(18)		(41)	(140)	(138)		(321)					
Total	15	16	7	38	141	140	43	324	156	156	50	362	19·060

in which the whole of the 362 observations are divided,

(1) according to the number of plates observed,

(2) in ten periods of time of alternately 36 and 37 days, into which the year was divided,

(3) according as they are judged to be exceptionally variable, or not, solely upon the evidence of the χ^2 index. The subdivision which would be made taking also into account the evidence for epidemics is shown in brackets, but in discussing the evidence for epidemics these modifications are ignored.

To test the first point, each line of Table XI is treated as a 2×3 contingency table, and the value of χ^2 calculated from it. It has been shown (Fisher, 1922(6)) that as in such a table there are two degrees of freedom, χ^2 will be distributed, if there is no association, as in Elderton's Tables when $n' = 3$. To show that at no period is there significant association, the values of χ^2 for the 10 periods are added, and the resulting quantity should be distributed as in Elderton's Tables when $n' = 21$. Since in two consecutive periods no exceptionally variable sets occurred, these periods have been omitted, and n' is taken to be 17. It will be seen from the table that all the values of χ^2 are less than 2, except in two periods in which only a single exceptionally variable set occurred. Such cases are evidently beyond the range of effective application of the χ^2 test, but even including these high values, $P = \cdot 266$, and therefore there is no significant departure from the rule that sets of three, four and five plates show equal proportions of exceptions in all sections of the period of observations.

This fact confirms the justness of the criterion by which the exceptions have been identified, for any error in the method of identification would naturally show itself in the proportion of cases regarded as exceptions; in the second place it indicates that the cause of exceptional variability is not connected with the causes which lead to the rejection of individual plates (contamination, development of fungi or overgrowth by *B. dendroides*), and in the third place it shows that the exceptions are not caused by the exceptional deviation of a single plate, for in this case the proportion of 5-plate sets would necessarily be highest. The third conclusion is borne out by an examination of the numbers counted on individual plates, and both it and the second conclusion are more decisively drawn from the contingency table by ignoring the period of occurrences.

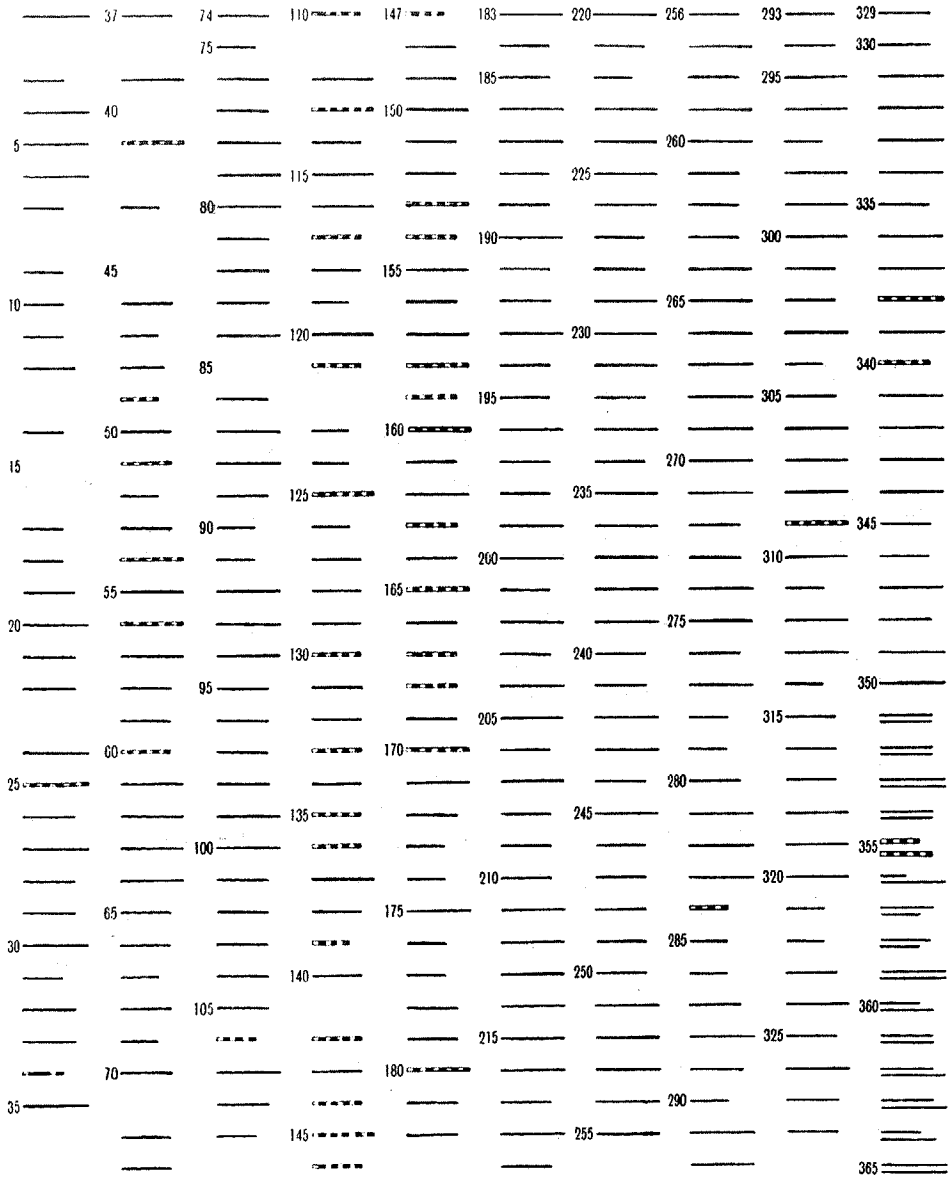
Table XII

No. of plates	5	4	3	Total
Exceptionally variable ...	16	18	7	41
Not exceptionally variable ...	140	138	43	321
Total	156	156	50	362

The numbers in the smaller groups are here sufficient to make a satisfactory test, and the value of P, ·739, shows distinctly that there is

Cutler's Data

Sets of plates for each day of the year



EXCEPTIONALLY VARIABLE

NOT EXCEPTIONALLY VARIABLE

5 Plate

4 "

3 "

Fig. 2.

no significant difference in the proportion of exceptions between the several groups of observations.

Similarly the distribution of the exceptions in time, in which we have shown the different groups to agree, may be best shown by taking the totals, irrespective of the number of plates in each set. If this is done we have a 2×10 contingency table, of which the value of χ^2 proves to be 57·826.

Since $n' = 10$, the chance of such a distribution occurring under conditions of random occurrence in time is about 4×10^{-9}. It is indeed obvious from inspection of Fig. 2 that the exceptional values occur in groups together, although perfectly normal values continue to occur throughout the worst of these epidemics. During the first outbreak seven exceptions occurred with 14 normal values among them; the second epidemic period was more prolonged and included 27 exceptions and 46 normal values. In the second half year of the experiment only six exceptions occurred, of these two occurred on the same day (355) during the last fortnight, when duplicates were taken, and two others, 338 and 340, were but two days apart.

Bearing these points in mind, we have no hesitation in concluding, on purely statistical evidence, that the exceptionally variable sets of platings were due to two causes:—(*a*) a predisposing cause which is at work throughout the epidemic period, and (*b*) some additional circumstance, in the absence of which the counts obtained will still be normal.

8. Special Organisms which affect the Number of Colonies developing

In the daily counts above considered, a uniform technique was followed throughout, and fresh batches of medium were made up at frequent intervals. It is conceivable that occasional differences in plating technique, in the medium, or in counting the plates may by chance have occurred on certain days. It is however most unlikely that any such differences can have extended over the long periods covered by the epidemics of high variance, without the fact being noticed. In seeking a predisposing cause of variance, covering these periods, therefore, one's attention is naturally drawn to possible changes in the soil itself or in its population.

It is known that certain micro-organisms, when growing on the medium, exert an inhibitory action on the development of colonies by other forms. The appearance of such an organism in the soil population, during certain periods, might therefore give rise to periods of higher

variation between parallel plates, for unless present in very large numbers it would not appear on all the plates or even in every batch of five plates.

An example of high variation between parallel plates, that was actually traced to such an organism, is given to illustrate this cause of inaccuracy.

The soil used in this case was from the Leeds Experimental Farm, and had received a treatment of naphthalene. Thirty parallel platings of this soil were made on Thornton's agar. The counts of colonies on these plates are given in Table XIII.

Table XIII

Parallel plates of Leeds soil

Plate No.	Number of colonies	Plate No.	Number of colonies
1	*240*	16	126
2	*209*	17	126
3	*177*	18	126
4	*158*	19	121
5	157	20	120
6	*154*	21	119
7	151	22	118
8	137	23	117
9	136	24	114
10	132	25	113
11	131	26	109
12	131	27	99
13	130	28	*91*
14	128	29	*91*
15	127	30	*87*

χ^2 Index. Whole series = 230·17
Minus the italicised plates = 27·81

It will be seen that the variation between parallel plates in the whole series is excessive. In examining the plates, some were found to contain an organism forming a growth between the agar and the bottom of the dish. This organism occurred on the plates italicised in Table XIII. It is a motile organism and apparently spreads in the water film underlying the agar. On plates 28, 29 and 30, the growth of this organism was sheet-like and from the low counts obtained it would appear that its growth has reduced colony development. On plates 1, 2, 3, 4 and 6, it has produced a number of separate colonies underlying the agar. These colonies were probably produced by individuals which had multiplied and migrated along the bottom of the dish after the agar had set,

but could not be separated from other colonies in counting the plate. The counts on these plates are therefore excessive. The presence of this organism on the bottom of the plates has thus produced an abnormal variation in the whole series. It will be seen that, if plates on which it occurs are ignored, the χ^2 index for the remaining 22 plates falls within the expectation of random sampling.

A pure culture of this organism was obtained and a plating from a sample of Rothamsted soil was made, a small loopful of suspension of the organism being added to the first dilution flask. Table XIV, Series A, shows the colonies developing on six parallel plates of the soil thus treated, compared with a control series of plates of the same soil not inoculated, Series B, which were made at the same time.

Table XIV

Effect of Leeds soil organism on colony development from suspension of Rothamsted soil

Series A. Suspension inoculated			Series B. Control	
Plate No.	Number of colonies	Area of bottom spreading	Plate No.	Number of colonies
1	85	nil	1	95
2	79	nil	2	90
3	78	nil	3	86
4	70	nil	4	85
5	58	nil	5	85
6	60	2·25 sq. cms.	6	82
7	56	7·75 ,,	7	81
8	45	27·0 ,,	8	77
9	41	54·5 ,,	9	73
10	39	56·5 ,,		
χ^2 Index, Plates 1 to 5 = 5·86 Plates 1 to 10 = 40·01			χ^2 Index = 1·89	

In this case the organism formed a spreading growth over the bottom. The area of this spreading growth, where it occurred, was measured and is shown in Table XIV. It will be seen that the reduction in colony development is clearly related to the amount of spreading growth. In this series of plates it is also evident that the variation is greatly increased by the occurrence of the organism on certain of the plates.

From an abnormally variable series of plates of Rothamsted soil a second organism has been isolated, whose frequent habit it is to spread on the under surface of the agar, and which has a similar inhibitory action on the development of other colonies. Table XV shows two sets

of plates of a suspension of Rothamsted soil, one set of which was inoculated with this organism. The reduction of, and increased variation in colony numbers are again well seen.

Table XV

Effect of toxic organism from Rothamsted soil on colony development from a soil suspension

Series A Plates inoculated		Series B Control	
Plate No.	Number of colonies	Plate No.	Number of colonies
1	192	1	179
2	168	2	171
3	147	3	168
4	130	4	150
5	127	5	150
6	113		
Mean 146·1 χ^2 Index = 29·47		Mean 163·6 χ^2 Index = 4·17	

It is of course impossible to decide, with certainty, from a simple record of colony numbers, whether the presence in the soil of some such organism was the cause of the epidemics of variable plate-sets in Cutler's series. However, the above two cases of high variance between parallel plates, which have been traced to the presence of definite organisms, show that this factor, though apparently of infrequent occurrence, is capable of causing a disturbance in the colony numbers of precisely the kind actually observed. It is important to notice that this, probably like all other causes, that produce a sensible departure from the Poisson Series, seriously disturbs the mean value.

9. THE OCCURRENCE OF SUBNORMAL VARIATION

It has been shown that in a small proportion (about 34 cases) of Cutler's data, the variation between parallel plates has been apparently lowered by some disturbing agency. The same phenomenon in a much aggravated form appears in Owen's data (section 10), and has from time to time occurred in Thornton's work. For example the 20 plates shown in Table I display an unduly low variation, and though this fact does not detract from the value of the data in proving the equivalence of parallel dilutions, it does throw suspicion on the value of the mean as an estimate of bacterial density. A similar depression appears in Table XIV, Series B.

Unlike the excessively variable sets, the sets with subnormal variance cannot be identified individually in Cutler's data, and we have therefore less evidence upon which to put forward a biological explanation of the phenomenon; certain facts, however, concerning observations made in the course of 1921, suggest that additional precautions in the preparation of the medium, may be effective in eliminating the disturbing cause.

The additional data were accumulated in the Bacteriological Department[1] in the summer and autumn of 1921 in the course of some work on the relationship of bacterial numbers to nitrate content in the field soil. In each of these experiments a series of some 45 samples of soil were taken from a plot 9 by 15 feet in area and the bacterial numbers in each sample estimated by the plate method using Thornton's agar medium. The first experiment was carried out with the dunged plot in Barnfield. The technique used was similar to that employed in Cutler's work, five parallel platings being made of each sample and the colonies counted after an incubation of seven days at 20° C.

Of the 33 sets available, three show excessive variance, the remainder are distributed as in Table XVI.

Table XVI

$\chi^2 = 3{\cdot}08$ $P = {\cdot}381$

χ^2	5-plate	4-plate	3-plate	Total	Expected	x^2/m
·5	—	—	1	1	9·78	·79
1·5	4	2	—	6		
2·5	2	—	—	2	9·39	·21
3·5	4	2	—	6		
4·5	4	—	1	5	5·56	·37
5·5	—	2	—	2		
6·5	3	1	—	4		
7·5	3	—	—	3	5·06	1·71
8·5	1	—	—	1		
Total	21	7	2	30		3·08

It will be seen that these agree well with the Poisson Series, and show no sign of subnormal variation.

A second experiment was carried out at Kingsthorpe Hall, Northampton. The soil is here of a markedly different type from the heavy Rothamsted soil, being a light ferruginous loam. In this experiment the technique was varied in that the colonies on each plate were counted twice, after seven and twelve days' incubation. It will be sufficient to compare the observed and expected values of the total, $S(\chi^2)$, for different groups of plates.

[1] The authors wish to acknowledge their indebtedness for the assistance rendered by other Departments at Rothamsted in this work.

Table XVII

Number of plates per set	Medium	After 7 days		After 12 days	
		Expected	Observed	Expected	Observed
4	A	18	13·85	24	27·31
5	A	152	109·33	144	133·96
9	A	8	1·96	8	8·73
Total	A	178	125·14	176	170·00
20	B	19	19·45	19	25·34

In all these groups where medium A is used the variance is distinctly subnormal after 7 days, but is apparently normal after 12 days. With medium B, the variance is normal at both counts. Now the sets of 9 and of 20 plates were parallel dilutions of the same sample, and the mean count from medium A was only 75 per cent. of that obtained on medium B. The abnormality of medium A was afterwards traced to the temperature at which it was filtered, a technical detail which has an important bearing on the ability of the medium to support bacterial growth (Thornton, 1922(11)).

In the comparison given by Thornton(11) of the two batches of medium, identical save that one was filtered at 50° C. and the other at 100° C., 10 plates being prepared from each, the former gave a mean count 79 per cent. of the latter; in this case also the defective medium showed subnormal variance giving a value $\chi^2 = 3{\cdot}2$ (after eight days), whereas the normal medium gave a value 10·3. The former would only occur once in 22 trials by chance, and therefore represents clearly a subnormal condition.

Whatever the biological explanation of subnormal variance may be, it is therefore sometimes indicative of a serious error in the value of the mean. In this respect it is a danger signal which cannot be disregarded. When a set of plates shows excessive variability no one will be tempted to lay too much stress upon their mean; it is obvious in such cases that there is a large probable error, and it has been seen (Section 8), that there will usually be also a considerable systematic error in such cases. A set of plates with abnormally low variance on the other hand, may appear to be particularly good data, although, as we have just seen, this type of abnormality is also indicative of large systematic errors. It is therefore of practical importance that such departures from the Poisson distribution should be detected, whenever they occur. Since subnormal

variation cannot be detected with certainty in a small set of plates, we recommend that occasional sets of 10 or 20 plates should be prepared from time to time, and that if necessary every batch of medium prepared should be tested in this way, the colonies being counted after seven days.

10. The χ^2 Index of Variability applied to other Bacterial Count Data

It has been shown by the use of the χ^2 index of variability, that the great bulk of Cutler's data on soil bacteria appears to be true samples from the Poisson Series, and that therefore the accuracy of these results is known with precision; also that, by the same method, a small proportion of exceptions may be detected in which some definite disturbing cause has interfered with the accuracy of the results. It is therefore desirable to apply the same test to other sufficiently extensive bodies of material, in order to ascertain if, by other methods, a similar degree of accuracy can be obtained, and failing that, if further light can be thrown on the problems of the dilution method. Data from four sources have been examined in this way.

(A) Buddin's counts of soil bacteria at Rothamsted, using a gelatine medium.

(B) Counts of soil bacteria published by Engberding (1909 (12)).

(C) Breed and Stocking's tests of the accuracy of counting *B. coli* in milk (1920 (13)).

(D) W. Owen's bacterial counts in sugar refinery products (1914 (14)).

In the aggregate we have tested over 1000 sets of parallel plates; owing to the bulk of the total examined it is possible that a small proportion of arithmetical errors has been included, although the application of the method is much more expeditious than that of the preliminary investigation of Cutler's data. Only the obvious and unquestionable features of each body of data will be dealt with.

(A) *Buddin's data*

A very large number of bacterial counts were made at Rothamsted by W. Buddin, to whom we are indebted for permission to make use of these data. The actual plate counts, though not published, formed the basis of bacterial number estimations used in Buddin's work on the effect of antiseptics on soil (15).

The platings in this work were made on a nutrient gelatine having the following composition:—Witte's peptone 40 grams, Lemco 20 grams, NaCl 20 grams, gelatine 480 grams, distilled water 4000 c.c.

The counts therefore supply an example of the degree of accuracy obtained with a gelatine medium, where a considerable source of variance is produced by the occurrence of liquefying organisms on the plates.

From the mass of data available, 100 sets of triplicate platings were extracted. The expected and observed values of χ^2 in this series are shown in Table XVIII.

Table XVIII

χ^2	Expected	Observed	Difference
·5	39·3	25·5	− 13·8
1·5	23·9	26	+ 2·1
2·5	14·5	12	− 2·5
3·5	8·8	10·5	+ 1·7
4·5	5·3	6	+ ·7
5·5	3·2	4	+ ·8
6·5	2·0	3	+ 1·0
7·5	1·2	4	+ 2·8
over 8	1·8	9	+ 7·2
Mean	2·0	3·04	

There is a marked deficiency below 1, and an increasing excess above 3. No distinct class of exceptionally high values can be detected, only three values exceed 10, and none exceed 15. The causes of additional variability probably affect all observations in some degree, and are therefore systematic rather than sporadic. The mean variance is about 50 per cent. in excess of that due to random sampling. As in Cutler's 3-plate data the departure from expectation is best shown by dividing the distribution at the quintiles as in Table XIX.

Table XIX

$\chi^2 = 17{\cdot}4$ $\quad P = {\cdot}0017$

χ^2	Expected m	Observed $m + x$	x^2
0	20	12	64
·4464	20	15	25
1·0126	20	23	9
1·8326	20	15	25
3·2190	20	35	225
Total	100	100	348

Such a departure from expectation would occur by chance but once in 600 tests; it is therefore clearly significant. The technique used here did not therefore give results of such accuracy that the variance between parallel plates could approximate to the Poisson Series.

(B) *The data of Engberding* (12)

The parallel platings given by this author were made to test various points connected with the plate method of counting soil bacteria. Some of the sets of platings were made on a variety of gelatine and agar media, as a test of these. The majority, however, were poured on an agar medium, containing "Nahrstoff-Heyden," that was considered by the author to be the best of the media tested.

Engberding gives 24 sets of plates; of these, 14 are of six plates each, six of five plates, three of four plates and one of nine plates. Nearly all the sets show excessive variability; only three values out of the 24 are below the expected average for the corresponding number of plates. The total of the 24 values is 5·36 times the expected total. No further test is necessary; random sampling must be regarded as one of the smaller causes of variation in these data.

(C) *The data of Breed and Stocking* (13)

We next come to a very thorough attempt made by Breed and Stocking to test and improve the methods used in the bacterial analysis of milk. The medium used in the platings here considered had the following composition:—"Difco" peptone 1 per cent., lactose 1 per cent., "Lemco" ·3 per cent., air dried agar 1·5 per cent. A single batch of medium was used throughout each experiment, so that ability to reproduce the medium, is not here tested. Parallel samples of normal milk, and of milk inoculated with *B. coli*, were analysed by different analysts and at different stations. Two series of these records have been examined by comparing the different plates of each separate analysis. Each series yielded 132 sets of three numbers, the duplicate counts of the same set of plates being reckoned as two. If the duplicate counts had closely agreed, this would tend to give us a bad fit between observation and expectation, to the extent of doubling χ^2. Though the agreement is not sufficiently great to have this effect, the tendency is to be borne in mind.

The expected and observed distributions are shown in Table XX.

As with Buddin's data, though to a less extent, there is a small systematic excess of the larger values; the mean variance in series B is about 30 per cent. in excess of expectation, while in series C it is only

about 20 per cent. Series B also shows certain other irregularities and possibly the occurrence of sporadic causes of variation. Series C, which represents the final perfection of the technique employed, shows no excessively variable sets of plates.

Table XX

χ^2	Expected	Series "B"	Series "C"
·5	51·9	46	43
1·5	31·5	35·5	30
2·5	19·1	14	24
3·5	11·6	6·5	12
4·5	7·0	10	10
5·5	4·3	3	4
6·5	2·6	5	4
7·5	1·6	2	2
over 8	2·4	8	5
Mean	2·00	2·65	2·45

It is, we believe, possible to indicate the cause of the small systematic excess of variance in this exceptionally fine body of data. As has been observed, the duplicate counts, which are recorded in full, do not agree very closely, and it is possible that what may be called "error of counting" is responsible for the existing discrepancy. If we consider such a typical pair of duplicate counts such as that shown in Table XXI, we may regard

Table XXI

Plate	First count	Second count	Difference	Departure from mean
1	70	68	+ 2	+ 8
2	61	72	− 11	− 5
3	54	63	− 9	− 3
Mean			− 6	

the mean difference, as due to the personal equation of the analyst; and the departures from the mean as made up of the several "errors of counting" of the set. If the standard "error of counting" is σ, then the mean value of the sum of the squares of the three departures will be $4\sigma^2$. In this way the standard "error of counting" was estimated for each of the main groups of observations in Series C, divided according to the mean number of colonies per plate, and the additional variance ascribable to "errors of counting" expressed as a percentage of the expected variance.

Table XXII

Percentage variance due to "errors of counting"

Colonies per plate about ...	36	62	82	161	364	All
Increased variance per cent. ...	16 %	24 %	13 %	17 %	59 %	22 %

The effect is thus seen to be a fairly uniform one, though distinctly more prominent among the more crowded plates, of which eight pairs of triplets were available. The higher value in the second group is perhaps due to the fact that these contain the counts of the mixed bacterial population in normal milk, while the others are counts of a practically pure culture of *B. coli*.

The effect ascribable to "errors of counting" is thus of just the right magnitude to explain the additional variance observed in Series C. Since all the groups are affected similarly and nearly to an equal extent, we may anticipate that if this explanation is correct, the actual values of Series C will fit the theoretical expectation if a uniform allowance of 20 per cent. is made for the additional cause of variation. The distributions are so compared in equal intervals of χ^2 in Table XXIII, and by sextiles in Table XXIV.

Table XXIII

χ^2	Expectation with 20 % allowance	Observed
·6	51·9	47·5
1·8	31·5	35·5
3·0	19·1	21
4·2	11·6	12
5·4	7·0	5
6·6	4·3	5
7·8	2·6	3
9·0	1·6	—
over 9	2·4	3

Table XXIV

$\chi^2 = 7{\cdot}545$, $P = {\cdot}185$ ($P = {\cdot}584$)

χ^2	Expectation m with 20 % allowance	Observed $m + x$	x^2
0	22	14	64
·4378	22	28	36
·9732	22	24	4
1·6634	22	21	1
2·6366	22	28	36
4·3003	22	17	25
Total	132	132	166

The distribution shown in Table XXIII shows a remarkably close agreement with expectation. A more exact test of agreement is afforded by the division at the sextiles (Table XXIV); the actual figures show but a moderately good fit with $\chi^2 = 7{\cdot}545$, and $P = {\cdot}185$; since however

duplicate counts of the same plates have been taken as independent observations, χ^2 has been increased by this cause to some extent short of doubling, so that we may say that in reality χ^2 lies between 3·77 and 7·54, while P lies between ·584 and ·185; neither value could be taken as indicating a significant departure from expectation.

We believe, therefore, that in this material, at all events in Series C, the somewhat severe conditions under which the Poisson Series is produced, were in reality fulfilled, and that the departure of the observations from expectation could have been eliminated had precautions been taken to secure a sufficiently accurate counting of the colonies. It must however be borne in mind that the material employed consisted in nearly all cases of almost pure cultures of *B. coli* in milk. The case cannot therefore be compared closely to the different problem of counting such a mixed bacterial flora as occurs in soil, where many different types of organisms, whose growth may be mutually harmful, occur on the plates.

The interference on the plates between dissimilar organisms cannot here be seen, neither can the capability of the medium to check this interference be studied. In this material, for example, there would be little danger of frequent interference by "spreading" organisms, whose growth, had they occurred, would probably have been stimulated by such a medium as was used, containing peptone and meat extract.

The lessened accuracy in counting a mixed flora on this medium is illustrated in Table XXII, where the second group of platings, which contains counts of uninoculated milk, shows a noticeably higher variance in counting than the adjoining groups made from milk cultures of *B. coli*.

The data show, however, that when such a simplified flora is studied, an agreement between parallel platings comparable with the expectations of random sampling can be obtained.

(D) *The data of W. Owen* (14)

One of the most remarkable bodies of data which we have examined is that provided by W. Owen in his investigation of various culture media for the counting of micro-organisms in cane sugar products. In this work, a variety of different media were employed, varying in composition, reaction and osmotic pressure. These were tested in counting bacteria from a variety of sugar refinery products. From the variety of media employed, and from the fact that most of them were new and of untested value, it was to be expected that a rather high variance between parallel platings would be found over the whole series taken together. Had this been the case, separate tests would have been needed of the

indices of variance on the separate media. In fact, however, no such remarkably high variance was found.

The analyses were performed with sets of six plates, and we have chosen the first 100 of these sets for examination. The expected and observed numbers are shown in Table XXV.

Table XXV

χ^2	Expected	Observed	Expected 43%
·5	3·7	38	1·6
1·5	11·3	15	4·9
2·5	14·9	6	6·5
3·5	15·0	9·5	6·5
4·5	13·4	6	5·8
5·5	11·0	3·5	4·8
6·5	8·5	3	3·7
7·5	6·4	3	2·8
8·5	4·7	1	3·5
9·5	3·4	1	
10·5	2·4	—	
11·5	1·7	1	
12·5	1·1	2	
13·5	·8	—	3·4
14·5		—	
15·5	1·6	1	
over 16		10	

The excess of highly variable sets occasions no surprise; we have met with this feature in about the same proportion in Cutler's data. What is astonishing in this case is the immense excess of sets less variable, and in the majority of cases much less variable, than would be the case under undisturbed conditions of random sampling.

In the fourth column we have shown the expected distribution fitted to the total number in the range from 2 to 14. This seems to agree with the distribution observed within this range. We are unwilling to lay much stress on this explanation since the agreement is based on only 36 observations. If it were accepted it would imply that the conditions which lead to the Poisson Series were really operative in about 44 per cent. of the cases, that in at least 10 and probably 11 per cent. excessive variability has been produced, and in the remaining 45 per cent. the variability has been abnormally depressed.

The extent to which the differences between the counts of parallel plates is diminished seems to put the phenomenon beyond the reach of the ordinary explanations; there are some indications, for example, that the plates have not been in all cases completely counted, but it is

difficult to imagine that this cause could be responsible for any such bias as is observed, in view of the fact that a probable error is calculated separately from each set. Severe competition between colonies on the plate is admittedly a possible cause of diminished variability, but we cannot imagine it acting with such severity as would be necessary to explain these results, especially as in the 38 cases in which χ^2 is less than one, the mean number of colonies per plate is always less than 100, and in 15 cases is less than 10.

In more than one instance all the six plates have an equal number of colonies; in samples from a Poisson Series, this would occur but very rarely. For 13 colonies on each plate for example, as is recorded in one instance, the most favourable assumptions will only allow such a coincidence once in some 25,000 trials. Since in the majority of these counts we clearly are not dealing with undisturbed conditions of random sampling, the point cannot be pressed further. We do not agree, however, with the statement that, when such a coincidence occurs, the probable error is zero.

In reviewing the foregoing data, it seems probable that the action of liquefying bacteria, and the development of rapidly growing organisms, unchecked by the medium employed, were the main causes of excessive variance between parallel platings in the work of Buddin and Engberding respectively.

It appears, however, that the conditions of accuracy, such that the development of colonies on parallel platings will form a Poisson Series, can be fulfilled in dealing with a simplified bacterial flora (Breed and Stocking), and have been approached in dealing with the mixed micro-flora of soil, where the medium used has been so devised as to check the excessive development of spreading organisms, as in the case of Thornton's medium. It is possible that these conditions of accuracy would be fulfilled with greater certainty in the case of a mixed micro-flora, if the medium could be further improved so that it checked the growth of such harmful organisms as that found in the Leeds soil (p. 345).

CONCLUSIONS

(1) Under ideal conditions the bacterial counts on parallel plates will vary in the same manner as samples from a Poisson Series. When these conditions are fulfilled the mean count of a number of plates is a direct measure of the density of the bacterial population considered (though not, of course, of the total bacterial flora); and the accuracy of such an estimate is known with precision.

(2) For any considerable body of records of sets of parallel plates, agreement with this theoretical distribution may be tested by means of the index of dispersion

$$\chi^2 = \frac{1}{\bar{x}} S(x - \bar{x})^2,$$

where $\bar{x}$ is the mean, and x any individual number of colonies counted on a plate (see Section 5).

(3) From an examination of several large bodies of data we conclude that accurate conformity with the theoretical distribution, though rare, is not unattainable. In particular with a carefully improved technique, and a relatively simple bacterial flora, we believe that the conditions have probably been fulfilled by Breed and Stocking; secondly, by the aid of a specially adapted medium Cutler and Thornton have shown that these conditions have been accurately reproduced, in the great majority of cases, even with the mixed bacterial flora of the soil.

(4) *Any significant departure from the theoretical distribution is a sign that the mean may be wholly unreliable.*

(5) Excessive variance may be produced by the occurrence of certain soil organisms, which have been isolated, and which exert a toxic influence on other forms, and in one case disturb the counts by multiple colony formation.

(6) Subnormal variance is in our experience indicative of some defect in the composition of the medium.

REFERENCES

(1) Poisson (1837). *Recherches sur la probabilité des Jugements.* Paris.

(2) "Student" (1907). On the error of counting with haemocytometer. *Biometrika*, v. 351.

(3) Fisher, R. A. (1921). On the mathematical foundations of theoretical statistics. *Phil. Trans.* A. ccxxii. 309.

(4) Elderton, P. (1902). Tables for testing goodness of fit. *Biometrika*, i. 155.

(5) Pearson, K. (1914). *Tables for statisticians and biometricians.* Camb. Univ. Press.

(6) Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of *P*. *J.R.S.S.* lxxxv. 87.

(7) Soper, H. E. (1914). Tables of Poisson's exponential binomial limit. *Biometrika*, x. 25.

(8) Whitaker, L. (1914). On the Poisson law of small numbers. *Biometrika*, x. 36.

(9) Bateman, H. (1910). Note on the probability distribution of α particles. *Phil. Mag.* Series vi. vol. xx. p. 704.

(10) Bortkiewicz, L. von (1898). *Das Gesetz der kleinen Zahlen.* Leipzig.

(11) THORNTON, H. G. (1922). On the development of a standardised agar medium for counting soil bacteria, with especial regard to the repression of spreading colonies. *Annals of Applied Biology*, IX.

(12) ENGBERDING, D. (1909). Vergleichende Untersuchungen über die Bakterienzahl im Ackerboden in ihrer Abhängigkeit von ausseren Einflussen. *Centralblatt für Bakteriologie*, Abt. II. Bd. 23, p. 569.

(13) BREED, R. S. and STOCKING, W. A. (1920). The accuracy of bacterial counts from milk samples. New York Agr. Exp. Station. *Technical Bulletin*, No. 75.

(14) OWEN, W. (1914). Investigation of the comparative values of various culture media for the quantitative determination of microorganisms in cane sugar products. *Centralblatt für Bakteriologie*, Abt. II. Bd. 42, p. 335.

(15) BUDDIN, W. (1914). Partial sterilisation of soils by volatile and non-volatile antiseptics. *Journal of Agricultural Science*, VI. 417.

(16) WYANT, Z. N. (1921). A comparison of the technic recommended by various authors for quantitative bacteriological analysis of soil. *Soil Science*, XI. 295.

(17) CUTLER, D. W., CRUMP, L. M., and SANDON, H. (1922). A quantitative investigation of the bacterial and protozoan population of the soil, with an account of the protozoan fauna. *Phil. Trans. Roy. Soc.* B.

(*Received July 21st*, 1922.)