The problems of sampling and of the statistical evaluation of microbial numbers in mineral waters

M. D. MORISETTI (*)

Thames Polytechnic, London.

INTRODUCTION

The objective of the microbiological examination of a food or water for the purposes of quality inspection must be to obtain reliable information upon which, if necessary, corrective action may be taken. The application of statistics will be of value in indicating what reliance may be placed on the information gained and when action should be taken. For the microbiological quality control, or better inspection of a process it is necessary to take samples of the population and subject them to test. The test may be to determine the value of some variable *e. g.* the number of microorganisms present in unit volume or weight of the material; or of some attribute *e. g.* the presence or absence of a particular group of microorganisms.

SAMPLING

A sample should resemble as closely as possible the population from which it is drawn *i. e.* it should be representative of the population. It should be taken in such a way that every part of the population should have an equal chance of being selected. This will require the proper mixing of a material such as a liquid, or where the material is presented as a series of separate units, *e. g.* cans or bottles, by allocating numbers to the units and selecting from them using random number tables. If a representative sample cannot

---

(*) Head of School of Biological Sciences, Thames Polytechnic, Wellington Street, London SE 18 6 PF.

Ann. Ist. Super. Sanità (1976) **12**, 191–202

be obtained then any test result must be suspect because the value obtained is unlikely to approximate to the true number in the population from which the sample has been drawn.

As the number of samples taken increases so the precision of the estimate of the population parameter should increase, reaching a maximum when the whole population is sampled. Clearly the latter is not likely to be possible. Much testing is destructive in nature and for microbiological estimates the test procedure is usually time–consuming and expensive so that large scale sampling becomes prohibitively expensive and greater than the profit margin that the process can bear.

A sampling scheme must be then a compromise between the degree of risk which the users of the scheme are prepared to accept and the economics of production and testing of the material concerned. Generally a sample forms only a small fraction of the population (or batch) from which it is drawn and the information gained is related to the size of the sample rather than to the batch size.

When the batch of material to be tested is in the form of a large number of units *e. g.* cans or bottles, one type of control procedure, which is applicable whether the material is homogenous or not, is to grade the items in a sample into acceptable or defective. No estimate of a mean value of a variable is required only whether an item in the sample exceeds or does not exceed some value or possesses or does not possess some characteristic. This is called sampling by attribute. Samples of n items are taken from the batch (of size greater than 100 n) and the batch rejected if the number of defective items, x is greater than some specified value.

Fig. 1 shows the operating characteristics of three such sampling schemes. The operating characteristic shows graphically the risks associated with the given sampling scheme. Greater protection is provided when the size of the sample is increased. When n is five, batches containing 12 % defective items will be accepted on average of no less than 54 % of the occasions of test, when n is ten, on average 31 % of the occasions and when n is twenty on average 10 % of the time. However the degree of protection afforded to the purchaser or consumer is extremely low in all three cases and illustrates that little value may be placed on isolated samples taken from batches of unknown origin. To ensure that a batch containing 1 % defective items is rejected with a probability of 0.99, a single sample of 500 items would have to be taken. Clearly the cost of such a scheme would be prohibitive both in terms of time and money if applied to tests for the presence or absence of coliforms or of *Escherichia coli* in bottled mineral waters. However the operating characteristic ($n = 10, X = 0$) shows also that although it may be small there is always a probability that a batch will be rejected if it contains any defective items. Thus if batches from a producer are sampled

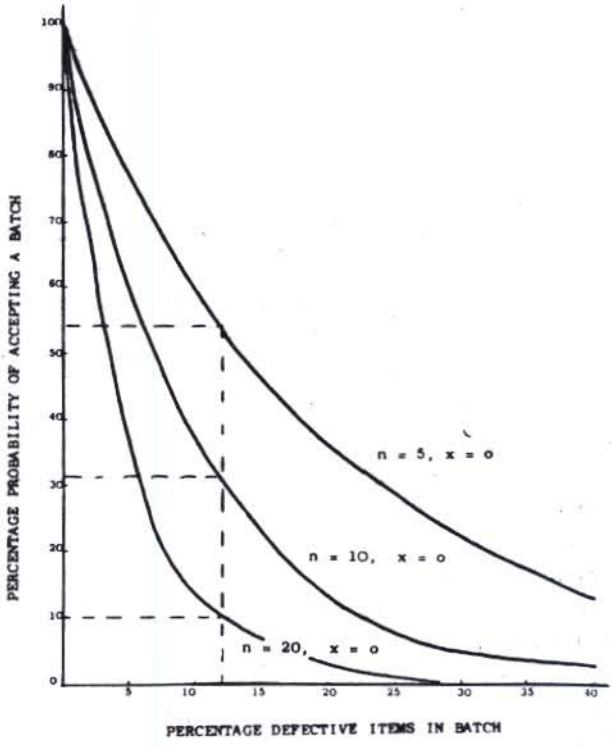

Fig. 1. — Operating characteristics for the sampling schemes indicated

n = number of items in the sample;
x = number of defective items permitted.

continuously using a sampling scheme $n = 10$, $x = 0$ and the batches contain 0.5 % defective items, the batches would be rejected in one in twenty, on average, of the occasions tested. No manufacturer could afford this level of rejection and would therefore need to ensure that proper hygienic conditions of production were maintained by suitable means including the use of physical checks and appropriate sanitation regimes. The microbiological examination of a sample of n bottles from a days' production can give no absolute guarantee that the product is free of a particular organism such as a coliform or an associated pathogen. Indeed an absolute assurance cannot be given with any practicable scheme and one is led to the conclusion that microbiological testing alone is insufficient to control the hygienic condition of a product such as bottled mineral water. The microbiological examination has value, however, as a check on the hygienic handling and bottling of the water.

ANALYTICAL PROCEDURES

Microbiologists may spend a significant proportion of their time carrying out tests in order to determine the numbers of microorganisms present in a material. Unfortunately only too often much of their effort is wasted because no checks are made to determine whether the analytical procedure is in a state of statistical control. Such checks are particularly relevant to microbiological quality inspection and are comparatively simple to carry out. They are essential if any reliance is to be placed on the test results.

Two sources of error enter into an analytical procedure, sampling error and technical error. The latter may be reduced by careful attention to technique and close adherence to the protocol for the test. The former is unavoidable and provides a means of checking on the state of statistical control of the procedure. Under ideal conditions because of sampling error in the number of organisms selected in replicate aliquots the colony count on replicate plates will be distributed as in a Poisson series [1]. The conditions which must obtain are that the particles to be counted are distributed randomly throughout the material; that accurate aliquots are taken and that each particle gives rise to a single colony. Further the development of one colony should not affect that of another.

Adherence to the Poisson series occurs more generally when pure cultures or mixed cultures on selective media are counted rather than with complex populations such as those found in soil when plated on non–selective media [2]. The use of data which departs from a Poisson series is highly questionable and the reasons for abnormal variation should be determined and, if possible, eliminated.

The χ^2 test may be used to determine whether the observed colony counts differ significantly or not from the expected values assuming that colony counts of replicate aliquots conform to a Poisson series. The χ^2 value may be calculated from the relationship

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

Eisenhart and Wilson [2] suggested a control chart procedure to test if counts were in statistical control. They pointed out that the critical value of χ^2 depends only on the number of replicate counts. If the control procedure is to be used then at least duplicate counts must be carried out at each dilution level. Cowell and Morisetti [3] applied the test to data of Huhtanen [4] for counts on raw milk at 2 °C and 37 °C. Figs. 2 and 3 show the control charts using that data. If the procedure is in statistical control then the χ^2 values obtained should lie above (or below) the dotted line ($P = 0.50$)

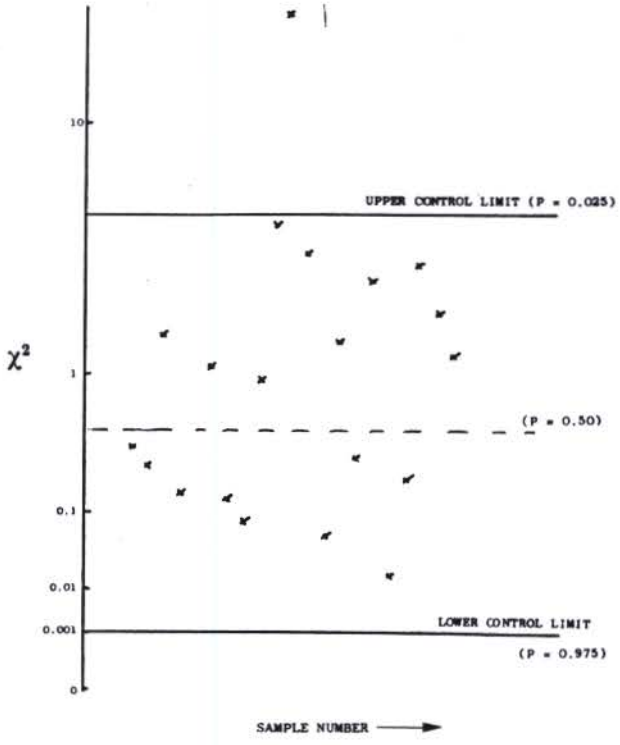

Fig. 2. – Control chart for plate counts of bacteria in milk
Incubation temperature = 2° C
Number of plates counted = 2 (*from*: [3]).

on half the occasions. The upper and lower control limits are set at P = 0.025 and P = 0.975 and therefore results should not fall outside these limits more often than one occasion in twenty on average. It should be noted that agreement between results can be too good as well as too bad. The 2 °C count appears to be in control, approximately equal numbers of the χ^2 values lying on either side of the P = 0.50 line and only one value out of 21 exceeding the limits. On the other hand the 37 °C count is clearly out of control, no less than nine of the values lying outside the limits, and conclusions based on these plate counts would be of dubious value.

One further point is worthy of consideration. Because replicate counts should conform to a Poisson series it permits the limiting precision of the method to be calculated since the data can never be less variable than that due to random sampling alone. The limiting precision may be calculated by making use of the fact that for a Poisson distribution the standard deviation

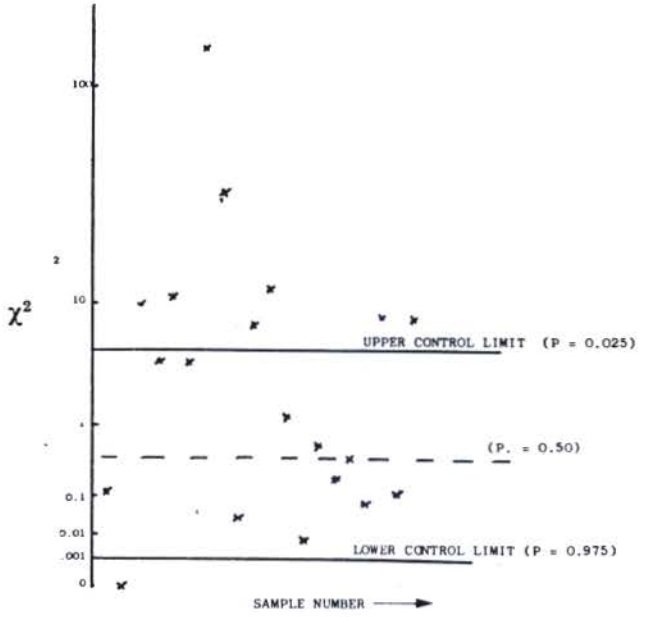

Fig. 3. – Control chart for plate counts of bacteria in milk
Incubation temperature = 37 °C
Number of plates counted = 2 (From: [3]).

is equal to the square root of the mean. Table 1 shows the limiting precision for different values of the colony count. When small numbers of colonies are counted the limiting precision of the method is low.

TABLE 1

The limiting precision (95% confidence levels) of colony counts of various values

True value of colony count	Limiting precision to nearest %	Limits
500	9 %	455–545
400	10 %	360–440
300	11 %	285–355
200	14 %	172–228
100	20 %	80–120
30	37 %	19– 41
20	47 %	11– 29
10	60 %	4– 16
5	95 %	1– 10

Plate counts carried out on natural waters are likely to be low and therefore the estimate of bacterial density obtained is almost inevitably imprecise. The error considered here is due to sampling alone and cannot be avoided. In addition technical error may introduce a further variation into the results which may be as great or greater than the sampling error. Little reliance therefore can be placed on single results and care should be taken in their interpretation particularly if a microbiological standard is involved. It is preferable to consider trends in the values obtained and the next section considers ways in which this may be done.

PROCESS CONTROL CHARTS

Most processes have an inherent variability. If this were not so then there would be no need for more or less sophisticated control procedures because any variation from the expected *i.e.*, an indication of bias, would show up immediately and the process would be in an out–of–control state. The function of the control chart is to distinguish between the inherent variability in a process and the situation when a significant and systematic bias enters into the process, the in–control and out–of–control states.

Fig. 4 shows the form that the Shewart chart may take [5]. There is a target (or population mean) line and upper and lower warning and action limits. Values, either of the sample mean or range, are plotted sequentially over a period of time. Action is taken if any point lies outside the action limit or if two successive points lie between the warning and action limits. When the distribution of the variable is known to be normal the limits may be set so that the probability that they will be exceeded is known. If the distance between the mean and the action limit is set proportional to three

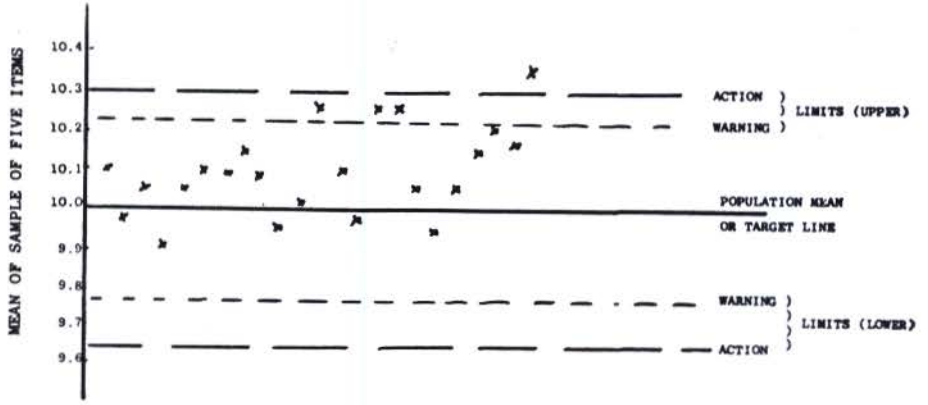

Fig. 4. — Hypothetical shewart chart for control of a variable

standard deviations then 99.7 % of the values obtained may be expected to lie between the action limits and only three observations in a thousand on average to lie outside them. Again if the distance between the warning limit and the mean or target line is made proportional to two standard deviations only one in twenty observations on average would be expected to lie outside these limits. Two successive observations lying between one pair of warning and action limits would be likely to occur, on average only once in eight hundred occasions and is therefore sufficiently improbable to justify action being taken.

The Shewart chart is widely used in industry but is not necessarily suitable for microbiological work. Estimates of viable numbers of microorganisms are time–consuming and expensive to carry–out and therefore it is very unlikely to be possible to carry out the number of tests required when the sample contains a number of items as is usual for the Shewart test. This difficulty will be greatest for the small producer for whom microbiological testing will be a relatively greater expense than for the large producer. Further the distribution of numbers of microorganisms in a particular foodstuff of which samples are taken over a period of time, is not likely to be known. Consequently it would not be possible to set limits to a Shewart type chart for variables other than by arbitrary decision.

An alternative control chart which may be used for variables is based on the cumulative sum test and is called the « cumulative sum» or « Cusum» chart. It is of particular value when observations arise singly and relatively infrequently and when it is not known if the process distribution is normal. These conditions are likely to apply to counts made on bottled waters; on empty bottles immediately after the washing process and other tests which may be carried out to check the hygienic standard of plant used in bottling the water.

The test uses the combined information of any number of observations by using the cumulative sum of deviations from the target rather than the individual observations. The chart will indicate that the process is in control as long as the sum of the deviations from the target value remains with small fluctuations, zero.

Table 2 is of a set of results (slightly modified from Bettes [6]) of plate counts of samples of water taken from the cooling section of a continuous steriliser. One sample was taken on each day of production and a plate count carried out.

Fig. 5 shows the actual counts plotted sequentially over a period of time and Fig. 6 shows the data plotted in the form of a Cusum chart. Bias is indicated more clearly and earlier in the Cusum chart than it would be in a Shewart chart. Each change of slope indicates a change in the average value of the count. Some of these changes will be due to random variation in the

TABLE 2

A cumulative sum analysis of plate counts of water taken from the cooling section of a continuous steriliser

Day and sample number	Count/ml	Count − 30	Cusum of (count − 30)
1	25	− 5	− 5
2	8	− 22	− 27
3	79	49	22
4	60	30	52
5	< 1	− 30	22
6	32	2	24
7	5	− 25	− 1
8	36	6	5
9	18	− 12	− 7
10	12	− 18	− 25
11	9	− 21	− 46
12	2	− 28	− 74
13	41	11	− 63
14	92	62	− 1
15	96	66	66
16	74	44	109
17	40	10	119
18	132	102	221
19	23	− 7	214
20	61	31	245
21	15	− 15	230
22	35	5	235
23	50	20	255
24	105	75	330
25	82	52	382
26	< 1	− 30	352
27	< 1	− 30	322
28	1	− 29	293
29	< 1	− 30	263

count but others will be significant and will give warning of an out–of–control state. Procedures are available for testing for significant increases in slope [7]. A graphical method using a V–shaped mask is simple to use [8]. The mask is laid on the Cusum chart such that the marked position on the mask overlies the last point plotted on the graph (Fig. 7). If the line of the graph cuts across either arm of the mask an out–of–control signal is given.

Van Dobben de Bruyn [9] summarised some of the advantages of the Cusum chart. Cusum charts react more promptly than Shewart charts to a small bias. If an out–of–control situation is indicated it can be seen from

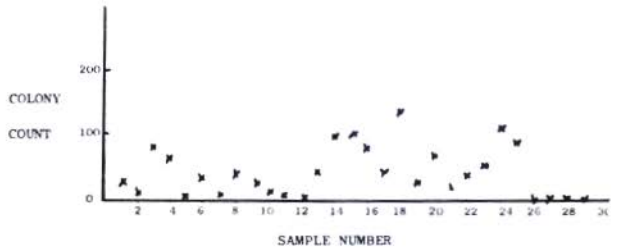

Fig. 5. — Colony counts v. sample number (Data from Table 2) (From: [6]).

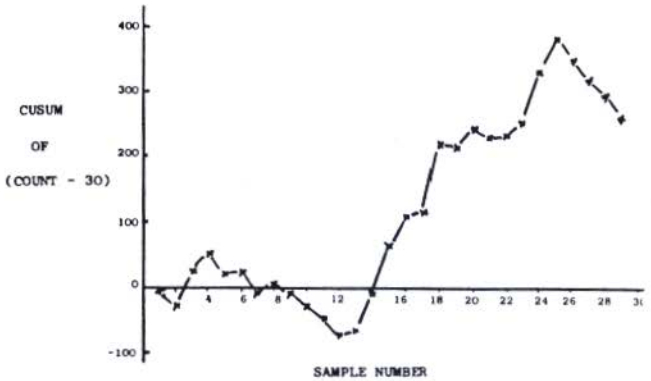

Fig. 6. — Cusum chart of data from Table 2.

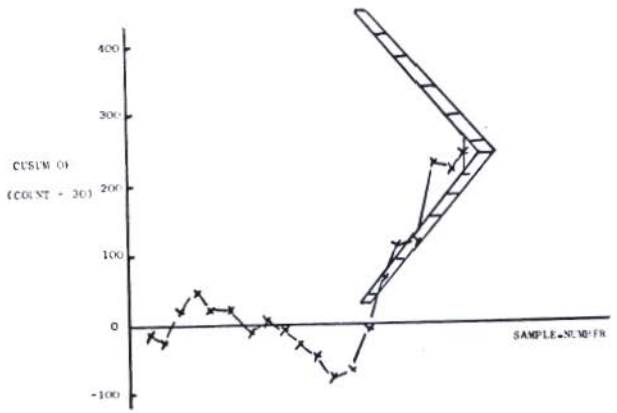

Fig. 7. — Part of Cusum chart shown in figure 6 with V-mask in position.

the graph when the change first occurred and an estimate of the bias can be made by measuring the slope. These plus the fact that the test can be applied to results which arise singly and which do not necessarily conform to a normal distribution make the test particularly suitable for use in monitoring colony counts derived from waters. A change in the colony count may give the earliest warning of the pollution of a water [10].

To summarise; the uncertainty associated with sampling schemes has led to the conclusion that microbiological testing alone is insufficient to guarantee safety and that such testing should be associated with a code of hygienic practice and used as a check on the efficiency of that practice. The low precision of the plate count when it itself is low has been discussed but nevertheless such counts can provide information which, when used cumulatively, may be used as evidence of undesirable changes in a process such as the production of bottled mineral water.

Summary. — Problems in sampling are considered and the implications of various sampling schemes discussed in terms of the risks involved to the users of the schemes. It is concluded that microbiological testing alone cannot guarantee the « safety » of a product. The reliability of analytical procedures of use in the microbiological examination of water is examined and a method of determining the state of statistical control of the procedure described. Finally control charts available for monitoring a manufacturing process are discussed with particular emphasis on the cumulative sum test and its application to colony counts from water.

Résumé (*Les problèmes relatifs à l'echantillonage et à l'estimation statistique de microbes rencontrés dans les eaux minerales*). — Les problèmes posés par la prise d'echantillons sont ici examinés avec les résultats provenants d'une varieté de procédés d'echantillonage relatifs qui se rapportent aux hasards rencontrés au cours de l'exercice. On suppose que des épreuves de microbiologie toutes seules ne garantissent pas que le produit à l'étude soit salubre. On examine l'exactitude de l'analyse microbiologique de l'eau; on propose un procédé qui constate si le règlement represente une vraie evaluation statistique. En dernier lieu on offre des graphiques pour contrôler un procédé de fabrication en soulignant l'importance de la somme cumulative et son application à l'énumeration des microbes rencontrés dans l'eau minerale.

REFERENCES

1. FISHER, R. A., H. G. THORNTON, & W. A. MACKENZIE. 1922. The accuracy of the plating method of estimating the density of bacterial populations. *Ann. Appl. Biol.* **9**: 325–359.

2. EISENHART, C. & P. W. WILSON. 1943. Statistical method and control in bacteriology. *Bacteriol. Rev.*, **7**: 57–137.

3. COWELL, N. & M. D. MORISETTI. 1969. Microbiological techniques – some statistical aspects. *J. Sci. Food Agr.*, **20**: 573–578.

4. HUHTANEN, C. N. 1968. Incubation temperatures and raw milk bacterial counts. *J. Milk Food Technol.* **31**: 154–157.

5. SHEWART, W. A. 1931. *Economic control of quality of manufactured product.* Van Nostrand, New York.

6. BETTES, D. C. 1966. Statistical quality control. *Food Manufacture.* **41** (6): 49–54.

7. PAGE, E. S. 1954. Continuous inspection schemes. *Biometrika.* **41**: 100–115.

8. BARNARD, G. A. 1959. Control charts and Stochastic processes. *J.R.S.S.* B **21**: 239–270.

9. VAN DOBBEN de BRUYN, C. S. 1968. *Cumulative sum tests – theory and practice.* Griffin, London.

10. ANON. 1969. The bacteriological examination of water supplies. Report No. **71**. HMSO, London.