Acknowledgement is made to the World Health Organization (European Region) for permission to use material from the chapter entitled 'Design of Sampling Programmes' in their forthcoming publication *Manual of Analysis for Water Pollution Control.*

General Principles of Sampling and Accuracy of Results 1980

Methods for the Examination of Waters and Associated Materials

Contents

London Her Majesty's Stationery Office

Contents (continued)

Note: From time to time, specialized booklets dealing with specific sampling methods and sample types will also be issued in this series, as for instance ref 184 and 186. This is in addition to the extra information contained in specific methods and essay reviews on analytical techniques.

Warning to users

The analytical procedures given in this booklet should only be carried out by competent trained persons, with adequate supervision when necessary. Local Safety Regulations must be observed. Laboratory procedures should be carried out only in a properly equipped laboratory. Field operations should be conducted with due regard to possible local hazards, and portable safety equipment should be carried. Care should be taken against creating hazards for others. Lone working, whether in the laboratory or field, should be discouraged. Reagents of adequate purity must be used, along with properly maintained apparatus and equipment of correct specification. Specifications for reagents, apparatus and equipment are given in manufacturers' catalogues and various published standards. If contamination is suspected, reagent purity should be checked before use.

There are numerous handbooks on first aid and laboratory safety. One such publication is *Code of Practice for Chemical Laboratories* issued by the Royal Society of Chemistry, London. Another such publication, which includes biological hazards, is *Safety in Biological Laboratories* (editors E Hartree and V Booth), Biochemical Society Special Publication No 5, The Biochemical Society, London.

Where the committee have considered that a special unusual hazard exists, attention has been drawn to this in the text so that additional care might be taken beyond that which should be exercised at all times when carrying out analytical procedures. It cannot be too strongly emphasized that prompt first aid, decontamination, or administration of the correct antidote can save life, but that incorrect treatment can make matters worse. It is suggested that both supervisors and operators be familiar with emergency procedures before starting even a slightly hazardous operation, and that doctors consulted after any accident involving chemical contamination, ingestion, or inhalation, be made familiar with the chemical nature of the injury, as some chemical injuries require specialist treatment not normally encountered by most doctors. Similar warning should be given if a biological or radiochemical injury is suspected. Some very unusual parasites, viruses and other micro-organisms are occasionally encountered in samples and when sampling in the field. In the latter case, all equipment including footwear should be disinfected by appropriate methods if contamination is suspected.

The best safeguard is a thorough consideration of hazards and the consequent safety precautions and remedies well in advance. Without intending to give a complete checklist, points that experience has shown are often forgotten include: laboratory tidiness, stray radiation leaks (including ultra violet), use of the correct protective clothing and goggles, removal of toxic fumes and wastes, containment in the event of breakage, access to taps, escape routes, and the accessibility of the correct and properly maintained first aid, fire-fighting, and rescue equipment. If in doubt it is safer to assume that a hazard may exist and take reasonable precautions rather than to assume that no hazard exists until proved otherwise.

First published 1980

ISBN 0 11 751491 8

About this series

This booklet is one of a series intended to provide recommended methods for the determination of water quality. In the past, the Department of the Environment and its predecessors, in collaboration with various learned societies, has issued volumes of methods for the analysis of water and sewage culminating in *Analysis of Raw, Potable and Waste Waters*. These volumes inevitably took some years to prepare, so that they were often partially out of date before they appeared in print. The present series will be published as individual methods, thus allowing for the replacement or addition of methods as quickly as possible without need of waiting for the next edition. The rate of publication will also be related to the urgency of requirement for that particular method, tentative methods being issued when necessary. The aim is to provide as complete and up to date a collection of methods and reviews as is practicable, which will, as far as possible, take into account the analytical facilities available in different parts of the Kingdom, and the quality criteria of interest to those responsible for the various aspects of the water cycle. Because both needs and equipment vary widely, where necessary, a selection of methods may be recommended for a single determinand. It will be the responsibility of the users – the senior analytical chemist, biologist, bacteriologist etc, to decide which of these methods to use for the determination in hand. Whilst attention of the user is drawn to any special known hazards which may occur with the use of any particular method, responsibility for proper supervision and the provision of safe working conditions must remain with the user.

The preparation of this series and its continuous revision is the responsibility of the Standing Committee of Analysts (to review Standard Methods for Quality Control of the Water Cycle). The Standing Committee of Analysts is one of the joint technical committees of the Department of the Environment and the National Water Council. It has nine Working Groups, each responsible for one section or aspect of water cycle quality analysis. They are as follows:

1.0 General principles of sampling and accuracy of results
2.0 Instrumentation and on-line analysis
3.0 Empirical and physical methods
4.0 Metals and metalloids
5.0 General non-metallic substances
6.0 Organic impurities
7.0 Biological methods
8.0 Sludge and other solids analysis
9.0 Radiochemical methods

The actual methods etc are produced by smaller panels of experts in the appropriate field, under the overall supervision of the appropriate working group and the main committee. The names of those associated with this method are listed inside the back cover.

Publication of new or revised methods will be notified to the technical press, whilst a list of Methods in Print is given in the current HMSO Sectional Publication List No 5, and the current status of publication and revision will be given in the biennial reports of the Standing Committee of Analysts.

TA DICK
Chairman

LR PITTWELL
Secretary

20 July 1977

1 Introduction

1.1 Aim of this Publication

Large amounts of time, money and effort are involved in the sampling and examination of waters, effluents, and other types of samples in order to provide information on their qualities. It is clearly desirable that such work be planned so that the required information is obtained with adequate accuracy and maximum efficiency. Time spent on this planning will be well-rewarded, particularly when, as is often the case, the programmes are of a continuing nature. The aim of this publication is, therefore, to provide general advice on the design of measurement programmes. Of course, many short-term information needs arise, and it is realized that it will then usually be impracticable to apply all the detailed considerations described below. Nevertheless, it is always desirable to attempt to follow the recommended approach to the maximum extent possible.

1.2 Topics Considered

The main topics discussed in the following sections are:

i. definition of the information required from the measurement programme;
ii. collection of samples;
iii. examination of samples;

In addition, certain statistical techniques are involved in ii. and iii. (Sections 4 and 5), and for those readers unfamiliar with such techniques, a simplified account of their basic concepts and methods will be found in Section 3. The important aspect of data-handling and interpretation is not included only because it is beyond the scope of this publication. A few introductory words on the above sections will help to indicate their scope and relation to other publications in this series.

If the information required on quality is not carefully defined, it is obvious that the measurement programmes may be inappropriate or inefficient or both. It seems that insufficient attention has often been given in the past to precise definition of the required information and generally-important points are, therefore, discussed in Section 2.

Given a clear statement of the information required, decisions must be made on where and when samples are to be obtained, and on the procedures and equipment used to collect and transport the samples to the location where they will be examined. The importance of ensuring that the samples are adequately representative of the water or other materials of interest needs no emphasis. However, it is stressed that, for many of the applications with which this publication is concerned, a number of factors can lead to grossly unrepresentative samples. To control such errors, several important principles should always be borne in mind and applied as appropriate; those principles are described in Section 4. Detailed recommendations on sampling procedures for individual determinands are given in the other parts of this series of publications dealing with the examination procedures for the various determinands. Note that throughout this publication the term determinand is used to signify 'that which is to be determined'.

The examination procedures applied to samples will also introduce errors in the results. The procedures must, therefore, be chosen so that they are capable of the required accuracy, and tests to ensure that that accuracy is achieved are also necessary. These topics are discussed in Section 5.

2 Definition of the Required Information

2.1 General Principles

As the expenditure of large amounts of time and effort will be governed by the definition of the information required, it is essential that the definition be unambiguous. This will be facilitated by preparing a written description of the required information, and the more detailed this description, the greater the chance of achieving an optimum programme of sampling and examination. In general, it is suggested that information needs should be restricted to those where the user of the results knows beforehand the use to which the results will be put. Otherwise, the situation in which large numbers of needless measurements are made may become perpetuated, probably to the detriment of the essential information needs. Of course, in defining the required information, full use should be made of existing data.

The definition of information needs does not, in principle, involve questions of sampling and examination. However, those latter aspects may, in practice, have very important effects on the information that can be obtained. Detailed discussions between those requiring information on quality and those who will supply it are, therefore, generally recommended when planning measurement programmes.

It should always be borne in mind that the time involved in, and the cost of, sampling and examination often increase markedly as the information needs become more exacting, for example, as greater accuracy, measurement of smaller concentrations or greater numbers of samples are requested. Care is required, therefore, to ensure that the cost of the measurement programme does not exceed the benefits that may accrue from its results. A formal cost-benefit analysis will often be difficult, if possible at all, but the concept is important, and it is a sound approach not to define the information needs to be any more exacting than appears essential.

Information on quality can be required for so many different purposes and applications that no attempt has been made here to consider all conceivable situations. With the above general points in mind, information needs are best defined by those responsible for each particular application. However, a few generally-important points of detail are worth brief mention, and are discussed in Section 2.2 to 2.5.

2.2 Need for Quantitative Definition of Information Needs

Information needs should be defined as quantitatively as possible. As an extreme and rather artificial example of a badly-defined objective consider the statement – 'to obtain information on the quality of rivers'. Such a statement is almost completely useless for the following reasons:

i. the rivers and determinands of interest are not specified;

ii. the required analytical sensitivity and accuracy are neither stated nor implied so that inappropriate methods of sampling and analysis may be used;

iii. no indication is given of the time-scale or sampling frequency so that too few or too many samples may be analysed;

iv. no indication is given on how quality is to be expressed (for example, as an average or median or maximum) and the tolerable uncertainty on any such parameters is not stated so that, again, inappropriate sampling and analytical techniques may be used;

v. no indication of the use of the data is given so that inefficient data-handling techniques may be employed.

2.3 Defining the Determinands of Interest

The quality-parameters must be defined unambiguously so that appropriate sampling and analytical techniques can be ensured. Three aspects are generally important, and their discussion jointly by the planners and by analytical experts will usually be beneficial.

1. Many substances can exist in water in a variety of different chemical and physical forms. For example, phosphorus may be present as ortho-phosphate, condensed inorganic phosphates and organic phosphorus compounds; dissolved and undissolved forms may also be present. Metals, carbon, nitrogen and silicon are other common elements that may exist in many forms. The response of an analytical method often depends on the form of the substance in the sample, for example, methods for ortho-phosphate do not determine many organic phosphorus compounds. Thus, whenever, the possibility of different forms of the same substance arises, those forms of interest must be appropriately defined so that suitable analytical methods can be selected. There is often interest in differentiating between dissolved and undissolved forms, and then the method of differentiation (for example, the type of filter) must also be precisely defined.

2. Some quality-parameters are often expressed in such a way that one parameter is a class of compounds from which a number of individual compounds may be present in samples. For example, phenolic compounds are often quoted as a quality-parameter though many different individual phenolic substances may be present. Other examples are pesticides, organic matter, and polycyclic aromatic hydrocarbons. For such parameters, it is usually desirable to specify the individual compounds of interest so that, again, appropriate analytical methods can be selected. However, circumstances can arise in which the measurement of a whole class of compounds is required, for example, total organic carbon may be of interest.

3. Some quality-parameters are overall properties of a sample rather than a particular substance, for example, biochemical oxygen demand, colour, turbidity, taste and odour. Such parameters must be carefully defined so that the particular property of interest is identified completely.

2.4 Location, Time and Frequency of Sampling

2.4.1 Sampling locations

The general locations for obtaining samples should always be revealed by the statement of objectives. Sometimes this statement alone will also define the exact locations for sampling, for example, when the quality of the influent to, or effluent from, a water-treatment plant is of interest. On other occasions the objectives may be such that the exact sampling locations remain to be chosen, for example, when the effect of an effluent discharge on the quality of a river-water is of concern. In all instances, it must be ensured that the locations at which samples are taken are representative of the water of interest; factors affecting this representativeness are considered in Section 4.

2.4.2 Time and frequency of sampling

Any essential requirements for the time and frequency of sample collection and analysis must be defined. Thus, in process control, the need to control changes in quality may be so important that this alone governs the desired sampling frequency. The time of sampling can also be of particular concern when the quality of the water shows more or less regular variations, for example, diurnal or tidal variations of the dissolved oxygen content of rivers. As an aid in deciding time and frequency of sampling a discussion of the principal factors involved is given in Section 4.

It should be noted that although the ideal approach might often be to obtain a continuous record of the concentrations of determinands of interest, this is, at present, generally impracticable. Suitable analytical instrumentation for all determinands is either not available or else is too costly or insufficiently accurate or reliable. However, when suitable instruments can be obtained their use can be advantageous in providing a large sampling frequency. This, and other advantages of on-line analysis as well as the instruments available are discussed in another part of this series of publications. When such instruments are to be installed, suitable systems must also be established to ensure that either appropriate control actions are taken on the basis of the instruments' readings or that the data produced are treated efficiently to provide the required information. Otherwise, on-line analysis may do little more than provide a mass of undigested and relatively-useless analytical information.

2.5 Requirements for Analytical Results

Certain fundamental points are mentioned below; they are treated in greater detail elsewhere (1).

2.5.1 Range of concentrations to be determined

The concentration* range of interest can markedly affect the choice of analytical method. For example, techniques suitable for the determination of concentrations in the milligrams per litre range are often quite unsuitable for concentrations of micrograms per litre. The most vital point here is to specify the lowest concentration of interest because this will govern the limit of detection required of the analytical method. In choosing this lower limit, it must be borne in mind that demands for extremely low limits will often require greater analytical sophistication and effort. As a general rule, a lower limit of approximately 10% of the smallest concentration of importance would seem to be reasonable. Existing standards for water-quality are, thus, of value in choosing the lower limits for analysis. When such standards do not exist, a lower limit must still be specified, but bearing in mind the point made above that as limits decrease greater analytical complexity and cost will usually result.

2.5.2 Accuracy required of analytical results

Three points are worth noting.

i. It is usually important to specify separate values for systematic and random errors because their effects differ.

ii. The expression of tolerable error is commonly achieved by statements such as 'the error should not exceed a given proportion (for example, 10%) of the result'. Such statements often overlook the fact that analytical errors (expressed as a proportion of the result) increase markedly as the limit of detection of the analytical method is approached. For example, at the limit of the detection, the random error (95% confidence limits) is approximately 50% of the limit. Further, as concentrations decrease towards the limit of detection, it is often the case that the tolerable percentage error increases. One should, therefore, consider whether requirements on errors are better expressed by statements of the form – 'the error should not exceed c mg/l or p% of the concentration whichever is the greater'. The values of c and p for both random and systematic errors are chosen for the particular application.

iii. Random errors can be defined quantitatively only for a given confidence level (see Section 3), and the level(s) must be chosen appropriately. The 95% confidence level is often used but greater confidence levels (for example, 99%) may sometimes be necessary, for example, in controlling some crucial aspect of water-quality.

2.6 Reference

1. Cheeseman RV and Wilson AL, *Water Research Centre Technical Report TR66*. The Centre, Medmenham, 1978

* When the determinand is not expressed in concentration units, for example, conductivity, odour, the range of values of interest should be understood.

3 Introduction to some Concepts of Statistics

3.1 Introduction

3.1.1 The need

Most investigations of water quality, whether they are concerned with chemical concentrations or with the population densities of biological organisms, have to face two difficulties which harden the task of drawing correct conclusions from the experimental results.

i. The quantity of water actually examined is usually only a minute fraction of that for which information is being sought. The sampling scheme must therefore recognize and attempt to allow for possible variability both throughout the body of water and through time.

ii. For many determinands the procedures for collecting and examining samples are not free from error. The accuracy may be limited by technology or expense; whatever the reason, the possible sizes of the errors must be considered in interpreting the results of measurements.

Common sense can go some way towards surmounting these difficulties. Just the recognition of their existence may be enough if all that is needed is a purely qualitative assessment. Often, however, some kind of quantitative statement is required of the inaccuracies or uncertainties associated with the results. This may be in order to reach conclusions as objectively as possible, or to assess the relative risks of alternative decisions, or to ensure that the results of one investigation may be validly compared with those of another.

These needs can be met by establishing what are essentially common sense notions in an unambiguous mathematical framework. It is this mathematically based approach to the study of variability, error and uncertainty which is known as 'statistics'. Statistical methods aid the interpretation of data that are subject to random variability or variability due to uncontrolled factors, and such methods provide the most powerful attack on the difficulties outlined in i. and ii. above. This is true whatever the scale of the problem. Whilst few people would argue against the need for statistical aid in a complicated multi-factor investigation, it should be stressed that the simplest experiment too requires a proper statistical assessment if a purely subjective interpretation of its results is to be bettered.

What statistics provides is a language, a framework of concepts, which enables the problems of random variability to be worked on and communicated. It embraces a set of techniques for helping to answer types of questions which are of wide-spread applicability to the experimental sciences. Statistics is a discipline in its own right, like chemistry or biology. However, in a scientific application it should always play a strictly supporting role, and the pursuit of statistical methods should not overshadow the fundamental aims of the experiment. The aim of applied statistics is to bring its methods to bear on problems in ways which reflect as closely as possible the aims of the primary investigation it is serving. Statistics is an adjunct to, not a substitute for, the scientific method.

3.1.2 Scope of this section

This section does not attempt to survey the whole subject of statistics, nor even just the parts of it which are relevant to water analysis. The aim is simply to introduce and illustrate some basic statistical concepts, as far as possible in non-technical language. No formulae or 'cook-book' methods of analysis are offered; these are readily located in the many statistical texts which are available. Here the objective is to orient the reader to the statistical approach so that he is encouraged to apply it himself, and can recognize when to seek further help elsewhere. The need for specialized advice will always remain, for an important aspect of applied statistics is the right choice of method for the particular situation, and this can only come from knowledge and experience. Nevertheless, there are many simple statistical applications where a grasp of the fundamental underlying principles will greatly illuminate what would otherwise be an indigestible numerical exercise.

3.2 The Statistical Approach

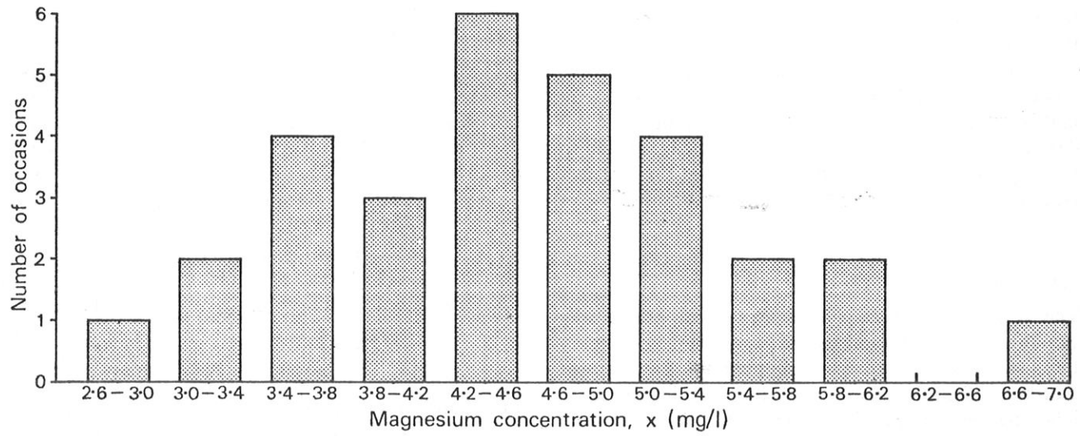
3.2.1 Types of variation

Statistics deals with two kinds of variability: random and systematic. A set of measurements varying in a haphazard and unpredictable manner about some central value would be exhibiting *random* variability. If, however, those measurements taken on day 1 show a persistent tendency to be higher (say) than those taken on day 2, they would be indicating a *systematic* variation between days 1 and 2. One of the general purposes of statistical analysis is to distinguish or separate systematic differences or trends from a background of random variability. This task crops up again and again in trying to answer questions such as:

'Is the average concentration of nitrate in River *X* greater than in River *Y*?'.

'Is the bacterial population density in the water leaving the sewage works below such and such a limit?'.

If the nitrate levels in each of the two rivers vary from place to place or at different times, and the laboratory

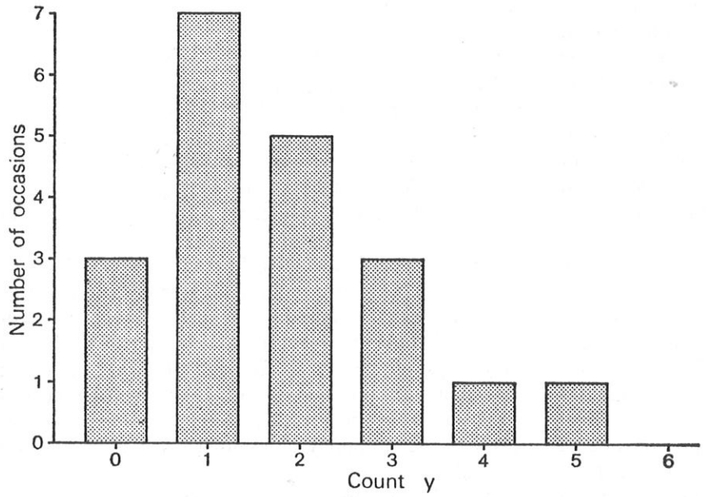


$\bar{x} = 4{\cdot}57$
$s = 0{\cdot}85$

Results of magnesium determinations, mg/l:

3·98	4·38	4·55	4·95	3·62	5·03
5·23	3·79	4·57	3·30	4·50	4·72
3·78	4·65	5·12	5·84	6·62	4·91
3·61	4·34	5·62	2·91	3·02	5·44
4·12	4·87	5·22	5·90	4·42	4·15

Figure 1 Results of magnesium determinations



$\bar{y} = 1{\cdot}75$
$s = 1{\cdot}30$

Counts observed in 10-ml portions of water from River T:

2	1	2	1
2	2	1	0
3	0	3	4
1	1	1	5
2	3	0	1

Figure 2 Micro-organism counts

cannot determine nitrate without experimental error, then the answer must take into account the uncertainties from these sources. Similar remarks may be made about the bacterial counts.

In other situations the question of interest may be posed differently: 'What is the average density of E-coli in this stream?' or 'Within what range may the bacterial density be reasonably assumed to lie?'. The answer to these questions too demand knowledge of the size and type of randomness in the samples and methods of measurement.

3.2.2 Random variability

The way random variability is dealt with statistically can be introduced by some simple examples. To simplify the two following examples, suppose for the moment that the difficulties associated with sampling from an inhomogeneous body of water may be ignored, and that the only statistical problems are ones of repeatability within the laboratory. Two examples of typical data are presented in Figures 1 and 2.

Figure 1 shows the results of a trial of a new method for determination of magnesium. Under normal laboratory conditions 30 analyses were performed on portions of a well mixed synthetic solution. The results are presented in histogram form.

Figure 2 shows the counts of micro-organism *M*, caught in 20 successive dips of a 10-ml trap into a well stirred bucket of water from River *T*. Their histogram is also shown.

Both of these sets of results show a scatter which is typical of many kinds of scientific observations. The result from the magnesium determination is not exactly the same every time, nor is the micro-organism count. Each is subject to what is called random variability, the chief characteristic of which is its unpredictability from observation to observation. Random variability governs *precision* which is a quantitative measure of the variability of observations. If the random variability in a set of results is high, then the precision will be low, and conversely the lower the random variability the higher the precision. Further discussion of precision in the context of analytical quality-control is contained in Sections 5.2.2.4 and 5.3.2.

There is one important difference between Figures 1 and 2. Each chemical result, while there may be no merit in trying to express it to more than three significant digits, can in principle take any value on a continuous scale. In contrast, the counts must be integers; fractional observations are impossible. Histograms provide a convenient way of summarising both kinds of data.

In both figures there is a tendency for the data to cluster about a central region. It is useful to be able to define the 'middle' of the data in an unambiguous way, and a number of measures like the *mode* (the most frequently occurring value) or the *median* (the value exceeded by half the data) are useful in certain circumstances, but by far the most widely used is the *arithmetic mean* – the familiar 'average' of everyday life. It is commonly denoted by a bar: thus $\bar{x}$ is the average of a set of x-values.

The other essential requirement is to know how spread out the data values are. Measures which summarise this are called *measures of dispersion*. Again there is a choice of such measures, but the one of fundamental importance is the *standard deviation* s (or, similarly, the *variance* s^2).

The means and standard deviations for the two illustrations are shown on Figures 1 and 2.

If another set of 30 magnesium analyses were performed, this would almost certainly produce a different set of data, a somewhat different looking histogram and different values of $\bar{x}$ and s. Any one such set of data is called a sample. For different samples the histograms will probably be different in detail but roughly similar in general shape. The haphazardness inherent in the method of chemical analysis is not completely chaotic; though individual results are unpredictable, in total they appear to show a pattern. It should be stressed that this cannot be proved. However, a great deal of experience accumulated in very diverse fields provides ample empirical evidence for it, and this 'order behind the chaos' is taken as a basic statistical assumption. Note the word 'sample' has similar meanings in both the experimental sciences and statistics. In the former it refers to 'a portion of the material of interest'; in the latter it means 'a subset of data values from the whole set of possible values'. The context will make it clear which interpretation of the word is intended.

3.2.3 Populations and probability distributions

The concept of a *population* is central to the whole of statistical reasoning. The assumption is made that the results of an experiment, such as those outlined in Figures 1 and 2, have arisen from a process which produces individual results in random order (that is, completely scrambled) but which has an inbuilt propensity to generate results of different sizes in particular relative proportions. For example, suppose the experiment consists of rolling a pair of dice in the hope of getting a double six. Provided the dice are not loaded, it is reasonable to imagine the underlying process as generating successes and failures in the ratio 1 : 35. This is not to say that in 36 experiments there will be exactly 1 success and 35 failures; that would be rather fortuitous. However, there will be an underlying tendency for the results to be generated in these relative proportions in the long run.

An equivalent but slightly different view is to imagine the results being drawn randomly from a very large body of data containing particular proportions of results of different sizes. In some applications this set may physically exist, as for example when sampling fish from a certain reservoir at a given time (the set then being the entire stock of fish in the reservoir). In other cases the set is purely conceptual, as it is in the dice example.

Whether the process idea or the large set idea is preferred, the totality of possible experimental results, together with their relative frequency weightings, is known as the *population*. The word is happily chosen as its statistical and everyday usages to some extent overlap.

These ideas are expressed in a more tangible mathematical way with the help of what are called *probability distributions*. Probability distributions suitable for the data in Figures 1 and 2 are shown in Figures 3 and 4. It will be remembered that the data are continuous for Figure 1 and discrete for Figure 2. This distinction is carried through to their respective probability distributions. One way to interpret the continuous distribution

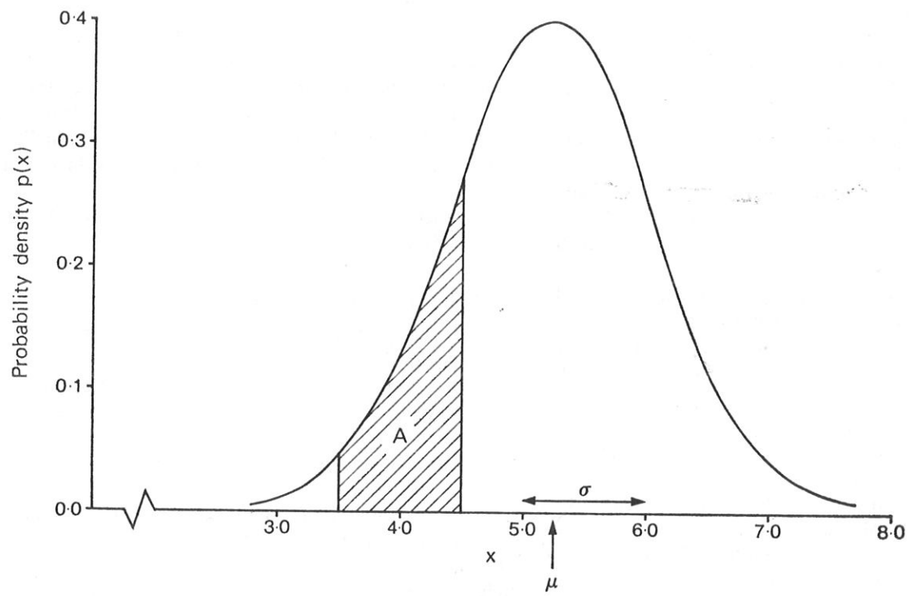


Figure 3 Member of the family of Normal probability distributions

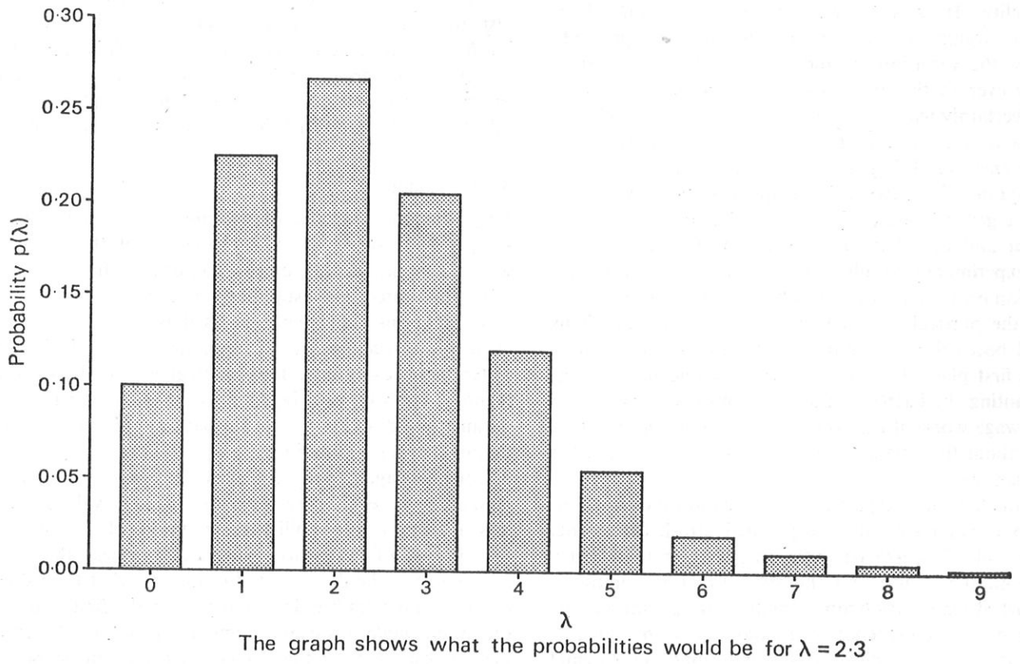


Figure 4 Member of the family of Poisson probability distributions

in Figure 3 is in terms of areas. The total area under the curve is arranged to be unity, and the area between two particular values of x is the probability that a result will fall in that interval. The height of the curve at any point is known as the probability density.

In the case of Figure 4 the distribution is discrete, and the vertical axis measures probability directly as there are now definite non-zero probabilities of observing the discrete values 0, 1, 2, 3, etc.

The probability distributions of Figures 3 and 4 each come from a *family* of distributions, that is distributions which all have the same general mathematical form and are specified by particular values of the *parameters* in the general equation. Figure 3 shows a Normal distribution; this is characterized by two parameters μ and σ. Note that population parameters are usually denoted by Greek letters, and sample statistics by Roman letters. The Poisson distribution of Figure 4 contains one parameter λ. There are a number of other commonly occurring families of distributions; the following remarks apply equally to all of them.

Just as the mean and standard deviation are used to summarize a sample, of data, so it is possible to introduce these concepts to describe a population. In this context the terms *population mean* and *population standard deviation* are used. Knowing the mathematical expression for a particular probability distribution, it is possible to work out the population mean and standard deviation in terms of the distribution's parameters. Thus in the case of the *Normal distribution* (Figure 3) the mean is μ and standard deviation σ. For the distribution in Figure 4 the mean can be shown to be λ and the standard deviation $\sqrt{\lambda}$.

A probability distribution can be regarded as a mathematical model of a system showing random variability. It is a convenient foundation on which further argument may be developed, but precisely because the variation is random, data from the system hardly ever fit the model exactly. This inability of the data (certainly with small samples) to provide conclusive corroboration of the assumed model heightens the importance of being able to justify the underlying assumptions. This can be attempted at two levels.

At a general level, rejection of the notions of population and distribution leads directly to the attitude that experimental results are special to the particular occasion on which they were obtained and all inferences from the particular to the general are impossible. This would be to deny the purpose of obtaining the results in the first place. For example, there would be no point in counting the bacteria in a 2-ml sample of water from the sewage works if the count could give no information at all about the bacterial density in water other than the 2-ml sample.

At the level of the particular application the choice of distributional form must be justified afresh each time. There will of course be occasions when it is obvious, but both the choice and the strength of reasoning to support it can vary from situation to situation. For example, on the evidence of the data alone in Figure 1, there is scanty justification for choosing the Normal distribution, though in cases like this it is often possible to draw upon previous evidence of Normality. A more practical point is that the Normal curve is very well understood mathematically and in the absence of better suggestions it is a good one to opt for. The Normal curve is in fact one of the most useful and widely used distributions because many situations yield data values which are nearly enough Normal for practical purposes. The immense convenience of the Normality assumption should not, however, be interpreted as a licence for its indiscriminate acceptance; it is only an accident that the word 'Normal' has an everday meaning of 'usual'.

For the data in Figure 2 the case is less in doubt. Provided the micro-organisms are small and swim independently the choice of the Poisson distribution can be given a theoretical justification.

3.2.4 Systematic errors

So far the discussion has concentrated on random or haphazard errors. Measurements may, however, also be subject to systematic error – or bias, as it is often known. Systematic error will be described here using the example in Figure 1; it is discussed in relation to laboratory analyses in Sections 5.2.2.7, 5.3.1 and 5.5.2.7.

In the discussion of population and distribution, μ was introduced as the conceptual mean of the population of experimental observations. Suppose now it is revealed that the synthetic solution used for the trial had been carefully made up to a magnesium concentration of 5 mg/l. It would appear from the results that μ is less than 5. This cannot be stated unequivocally because it is suggested only by sample data. If however the true μ were less than 5 the distribution would no longer be centred about the hoped for value and the new method of magnesium determination is said to show a *systematic error*. In this case the systematic error would be synonymous with what chemists call incomplete recovery.

The recovery of the method would not be improved just by doing more analyses. This would, however, improve the precision of estimation of μ. This illustrates the crucial difference between random and systematic errors. In the long run the former tend to balance each other out; the latter maintain a constant influence.

3.2.5 Estimation

Very often the purpose of making measurements on a sample of water is to infer properties of the body of water from which the sample was drawn. In these and other laboratory investigations a limited number of determinations may have to be used as a basis for saying what the concentration of, say, lead really is in the water being examined. In statistical terms, the experimental results have to be used to *estimate* the values of parameters describing the probability distribution of the underlying population.

Consider again the examples in Figures 1 and 2. Assuming in each case that the sample values have arisen from a probability distribution of a specified family but of unknown parameter values, the next step is to use the data to try to estimate what the values of the parameters are. It so happens in the first example that the *sample* mean and standard deviation, $\bar{x}$ and s, provide estimates of the *population* mean and standard deviation, μ and σ, which themselves define the Normally distributed population. It is also the case that the mean in the second example supplies the best estimate of the Poisson parameter λ. Common sense suggests that this

should be so, and much of basic statistical method is indeed essentially common sense in a mathematical framework. It should be mentioned, though that the estimation of population parameters from sample statistics is not always so trivial as it has been here.

Having estimated the parameters, the population (as it is believed to be) is now completely described, and so probabilistic statements can be made about it. For example it can be derived for the magnesium determination experiment that the chances of seeing an individual result lower than 2.5 mg/l are 0.011, or about 1 in 90. It is interesting to note that such a statement could not have been made from these data without the construction of a probability distribution. The sample itself contained no results below 2.91 mg/l. This kind of information which can now be gained has been bought at the price of the assumptions which were made about the distributional form and – the biggest assumption of all – that the population is a good enough model of the system which operates in the real laboratory.

3.2.6 Interval estimates

Estimation as so far described has been concerned with making a single 'best guess' about each unknown parameter. This is called a *point estimate*. If a new sample of data is obtained the estimates will change; to reiterate an earlier remark, if the experiment in Figure 1 were repeated different values of $\bar{x}$ and s would almost certainly be observed and so the estimates of μ and σ would be different. Thus a point estimate of a parameter will rarely be exactly right and so it is more informative to have in addition to this estimate a measure of the range within which the parameter is believed to lie. For example, an assertion that the magnesium concentration is likely to lie in the range 4.27 to 4.87 mg/l clearly provides much more information than does the single point estimate of 4.57 mg/l. A range such as this is called an *interval estimate*.

The most widely used method of interval estimation is that of confidence limits, and in order to explain this concept it is convenient to return to the example given in Figure 2. The population model adopted there enables formulae to be derived, for any chosen probability α (say 99%), by which an interval (a, b) can be calculated from the sample data so that the interval is likely to include* the true value of λ at the probability level α. As the numerical values of the data will vary from sample to sample so too will a and b. The formulae are designed to have the property that the interval they produce will contain λ on a proportion α of occasions of use in the long run. A numerical illustration will clarify this:

α	a	b
0.95	1.22	2.43
0.99	1.08	2.67

The interval (1.22, 2.43) is called the 95% confidence interval for λ, and (1.08, 2.67) the 99% confidence interval. The end points are often referred to as *confidence limits*. The confidence level may be set at any value, but it is obvious that the greater is the desired confidence the wider will be the corresponding interval, as the example illustrates.

Great care must be taken in the interpretation of confidence statements. To quote the 95% confidence limits for λ as 1.22 and 2.43 means that one is 95% confident that λ lies in the interval from 1.22 to 2.43. This is equivalent to saying that the probability of the interval covering the true value of λ is 0.95. Expressed in another way, if a large number of samples were taken from the population and a confidence interval for λ calculated for each, then on average 95% of the intervals would contain λ. The confidence statement does **not** mean that the probability that λ lies in any particular interval is 0.95, although this may seem to be saying the same thing. This last statement is incorrect because λ is a parameter, and so for a given population it has a fixed value which is either inside the interval or outside it. It cannot be inside a given interval for 95% of the time.

All experimental measurements are subject to random error, and any result derived from these measurements, whether a mean or something more complicated like an estimate of variance, will therefore also contain an element of uncertainty. If this unpalatable fact is ignored it is very easy to arrive at incorrect conclusions. Only by using confidence intervals (or the related techniques of hypothesis testing) is it possible to quantify the doubt surrounding any estimate and hence make rational decisions in the face of uncertainty.

3.2.7 Testing hypotheses

The discussion of the data in Figure 1 in connection with systematic error raised an interesting question. Is the difference between the sample mean 4.57 mg/l and the 5.00 mg/l of the synthetic solution attributable to the method of analysis not recovering all the magnesium (that is, having a μ less than 5.00), or is the difference attributable purely to the random variability of the determinations?

Statistics has a well developed set of techniques for dealing with questions such as this. For this example the argument runs as follows. Under the hypothesis that μ really were 5.00 what would be the probability of observing an $\bar{x}$ equal to 4.57 or less? Assuming the Normal probability model discussed in Section 3.2.3, this can be shown to be rather small, between 1 in 100 and 1 in 200. On the other hand if the alternative explanation is entertained that μ really is something less than 5.00, the probability of observing an $\bar{x}$ <4.57 is larger. How much larger this is depends on how small μ really is. For example, if μ were equal to 4.57 the probability of observing $\bar{x}$ <4.57 would be one in two.

The striking difference between the chance of observing $\bar{x}$ <4.57 under the *null hypothesis*, as it is usually known, and the chance of observing it under the alternative leads to rejection of the null hypothesis at the 1% significance level.

If the case were less clear-cut, and there was say a 1 in 10 chance of observing the data under the null hypothesis, it might well be decided that the hypothesis ought to stand. At precisely what level of probability a hypothesis should be accepted or rejected is a matter of judgement and context; traditionally the rejection levels 5%

* Strictly speaking, many intervals (a, b) may be calculated which satisfy this condition. However, it is customary to select the one which makes the probabilities that the interval is either entirely too high or entirely too low both equal to $\frac{1}{2}(100 - \alpha)\%$.

1% and 0.1% are commonly used criteria of increasing significance.

3.2.8 Explanatory variables

The discussion so far has dwelt entirely on the case of a single variable (magnesium analysis, or micro-organism count) which is subject to random variation from sample to sample. In many statistical applications the situation is complicated by the presence of other measurable variables which influence the original variable being considered. These are known as *explanatory variables*. In a study of oxygen demand in a river, for example, suitable explanatory variables might be the time of day and the distance downstream from a fixed reference point. Oxygen demand will of course also be subject to random fluctuations in addition to the sytematic influences of time and distance. Thus the statistical problem is to identify the systematic effects in the presence of random variation. It is not appropriate here to become enmeshed in the details of how this is done. The important point is that the fundamental task still consists of estimating certain parameters and then testing various hypotheses about them; and if the basic principles are clearly understood, their extension to more than one variable raises few new ideas.

3.2.9 The planning of experiments

The supporting nature of the role played by applied statistics in scientific investigations has already been stressed. Statistics simply provides the quantitative means of dealing with problems and ideas which are largely already embedded in scientific method in general. This is especially true in experimental design. A scientist embarks on an experiment seeking information about the effects of certain factors, and if he plans an experiment (or series of experiments) with his objectives clearly in mind, the statistician will probably find his design basically satisfactory. Statistical considerations can help to extract the maximum quantity of information from a given amount of experimental effort; for example 3 replicates from each of 6 batches might give a more precise estimate of within-batch variability than would 2 replicates from each of 9 batches. However, the idea that between-batch error must be assessed by sampling from a number of batches, and within-batch error by doing repeat analyses on the one batch, is not the sole prerogative of the statistician and would arise naturally in any well-planned scientific study.

One aspect which the statistician is perhaps able to perceive more sharply than others is the relationship between the assumptions underlying the method of statistical analysis and the conclusions which may be drawn. For example suppose single nitrate values are obtained at 3 depths at each of 4 sites in a lake. If the rather strong assumption can be made that the depth effect and the site effect (if they exist) operate independently, it is possible to test whether either effect is significant. But if it is felt that a depth/site interaction may exist, it is no longer possible to assess the depth and site effects unless repeat samples are taken at some of the depth/site locations. Once again, it is a matter of common sense that the weaker the assumptions are, the stronger must the experimental design be to answer a particular question.

3.3 Further Reading

There are a great many statistical textbooks, both at the introductory and the more advanced levels, and it would be possible to give a lengthy bibliography of just the best of these. However, this would only serve to confuse the non-specialist reader with limited time at his disposal, and the aim has therefore been to provide only a very selective list of books. These works expose the principles of statistical reasoning in a realistic context, and mathematical aspects though properly handled are not made the main issue.

Among the 'paperback' statistics for the layman, Moroney's classic (1) still stands supreme. A newer work (2) by Cormack is also very illuminating and deserves attention. Both of these would be good for anyone seeking a fuller introduction than in Section 3.2 above.

From the many works written for scientists, (3) and (4) are two of the best. Their respective motivations, the chemical industry and biology, suit the two background disciplines involved in water examination, though it is not being suggested that chemists must exclusively read the one and biologists the other.

Experimental design has a succinct chapter in (2) but is not extensively treated in (3) and (4). The reader who wishes to pursue it further should first turn to (5) which itself contains a short general bibliography on this topic.

3.4 References – Short list for further reading

1. Moroney MJ, *Facts from Figures*, 4th edition. Penguin, Harmondsworth, 1969.
2. Cormack RM, *The Statistical Argument*. Oliver & Boyd, Edinburgh, 1971.
3. Davies OL and Goldsmith PL (eds), *Statistical Methods in Research and Production*, 4th edition. Oliver & Boyd, Edinburgh, 1972.
4. Campbell RC, *Statistics for Biologists*, 2nd edition. CUP, Cambridge, 1974.
5. Cox DR, *Planning of Experiments*, Wiley, New York, 1958

4 Sampling

4.1 Introduction

Section 1 has emphasized the vital importance of sampling. If sampling techniques are not selected with great care, the results of the measurements may be partially or completely invalid for their intended uses. Much time and effort can be devoted to long-term sampling programmes, and it is important to ensure that the routine measurements will provide the required information in the most efficient manner. Therefore, it is worthwhile expending appreciable time and effort on the planning and design stage of programmes; the cost of this will usually be well rewarded. Most programmes generate large amounts of data, and it is important that efficient techniques of handling and interpreting the data are used so that the data can be reviewed efficiently and the sampling programme modified if necessary to achieve the objectives.

The aim of this Section is to discuss the factors of general significance in sampling. The emphasis is placed on principles because it is impossible to give definite practical advice that is suitable for all situations. Much of the discussion will also apply to short term and research programmes.

The basic aim of sampling is to collect samples (usually extremely small fractions) of water* of interest whose quality represents the quality of that water. To achieve that aim two aspects of sampling should be considered. Firstly, a set of samples must provide a true representation of the temporal and spatial variations of the quality of the water body for the duration of the measurement programme. To achieve this the sampling location and the time and frequency of sampling must be considered. Secondly, all the determinands of interest in the sample must have the same values as the water body being sampled at the point and time of collection. Thus, it is also important to consider the method of sample collection and the transportation, storage and preservation of samples. In-situ measurements commonly ensure that the second criterion is met but the choice of sampling positions and the times and frequencies of measurements remains just as important.

The requirements for sampling given above involve the logical, sequential consideration of several factors discussed in Sections 4.4 and 4.5. Figure 5 summarizes the order in which it is recommended those factors be considered when designing sampling programmes. The figure gives initial priority to the need for the fullest possible definition of the objectives of the sampling programme. Thereafter a logical sequence is followed, and this should be continually reviewed in the light of the sampling experience. This may enforce or suggest some change in objectives or procedures. It is important to stress that such a design process is not a once-off happening, but part of a dynamic process of review and control.

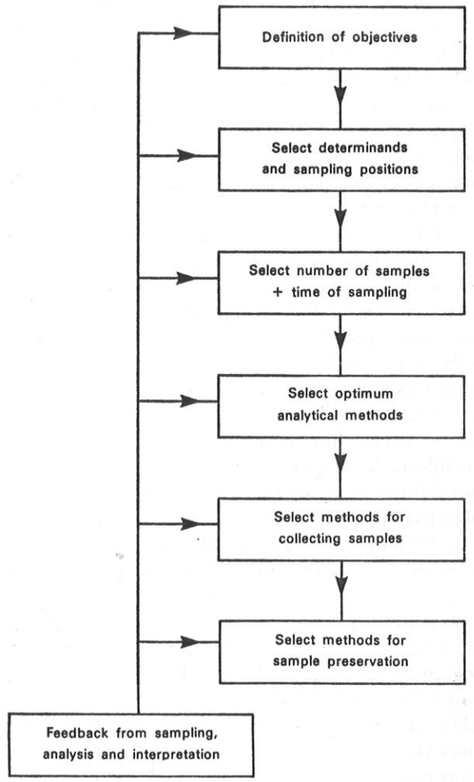


Figure 5 Sequential consideration of factors involved in the design of sampling programmes

Not only may feedback from sampling, analysis, and interpretation affect the objectives of the programme, but feedback from each decision box may affect the selections made in the previous boxes, for example, the analytical method may influence the method for collecting samples.

The prominence given to definition of objectives also implies the early involvement of statistical techniques in the design of the programme. This is necessary not only to formulate confidence limits on mean results, for example, but also to assist in the objective choice between sampling strategies.

* For simplicity, discussion is generally in terms of water samples; the sampling of sediments is subject to the same principles, and special sections on sediments are incorporated in the text.

When setting up a sampling programme, consideration must be given to a particularly important aspect – safety. This is discussed in Section 4.3.

Many other publications deal with the design of sampling programmes and some useful discussions are given in references 1–9. In addition the proceedings of two conferences contain many papers of interest (10, 11).

4.2 Economic Considerations and Optimization Techniques

The growing emphasis on water quality and pollution control is creating increasing demands for information on the quality of all types of water systems. This development is most marked for surface waters, and can lead to vast amounts of data being required for the management of water quality in large areas such as river basins. This can lead to situations where the initial designs of sampling programmes require more resources and effort for sampling and analysis than are available. Sometimes this problem can be overcome by adopting more economic sampling and/or analytical techniques, for example, use of the minimum tolerable sampling frequency (Section 4.5) use of composite samples (Section 4.6), the application of automatic sampling devices (Section 4.8.2), continuous on-line monitoring (Section 4.5.1), remote sensing techniques or the use of automatic methods of analysis. Such approaches are very useful and should always be considered in detail; however, they do not offer a general solution. Often, it will be necessary to re-consider the detailed design of sampling programmes in order to decide how best to reduce the amount of effort involved. Most workers will have experienced this problem, but there appears to have been little quantitative and objective discussion of methods for its general solution. Although it is not possible at present to give any solution, it is useful to discuss some of the aspects involved.

The design of a sampling programme must be chosen such that the information required by the objectives of the programme is obtained. There is little or no point in routine implementation of a programme known to be incapable of providing this information. Thus, the objectives are the prime factor in governing the resources and effort required for sampling and analysis. Therefore, it is essential that those responsible for defining the objectives should do so with great care. In particular, it is suggested that they should initially specify only essential rather than desirable objectives (it is easy to expand the programme when effort allows). In addition, the objectives should be defined as quantitatively as possible (12) so that the possibility of needless sampling and analysis is minimized. As a control on requesting inessential information on quality, information should be called for only when it has already been decided exactly how it will be used, and when systems to ensure its intended use have been established (13). As an aid in ensuring sound objectives, Kittrell (2) has suggested that they should be recorded in writing. This is an extremely useful suggestion. It is also essential that the information obtained from routine programmes is regularly reviewed so that sampling and analysis requirements can be reduced as soon as and whenever possible. The importance of these aspects of objectives cannot be overemphasized, and it is well worthwhile devoting appreciable time and effort to their definition and periodic review. Several papers deal with various aspects of the problem in connection with water-quality management in river basins (14–16).

When dealing with complex water systems such as river basins, it is generally not easy to decide precisely what information is required in order to achieve the objectives. This difficulty may perhaps be reduced by the use of appropriate techniques for the optimization of systems, for example, systems analysis (17). The principles involved in the application of this technique to river basins have been described in detail (3), but there seems to be insufficient experience of its use at present to judge its real value.

Given that an initial set of objectives has been carefully defined and that they require more information on quality than available effort can provide, some restriction of the objectives is then required. Ward and Vanderholm (18) have described a method that may be useful for dealing with this problem. In essence, they estimate the costs of sampling programmes for given degrees of effectiveness. Their approach is based on the idea that, once sampling locations and methods of sampling and analysis have been chosen, the cost of the programme is influenced mainly by sampling frequency. Thus, by first estimating the relationship between sampling frequency and effectiveness, the cost-effectiveness relationship can be deduced. For example, for programmes whose objectives involve quality-characterization (for example, for detecting long-term trends), the effectiveness is equated with the magnitude of the confidence limits of annual means (see Section 4.5.2). This leads to conclusions such as by reducing the sampling frequency by a given amount the confidence limits will be increased by a calculable amount. In this way, some quantitification of the benefits sacrificed by reduction of sampling frequency can be made. Ward and Vanderholm also considered programmes whose objectives are primarily to detect accidental pollutions of rivers. Their approach can be extended to programmes intended for both quality-characterization and quality-control, so that the best allocation of total available effort between the two types of objective can be estimated. Beckers and Chamberlain (19) have recently described a more comprehensive approach to the design of cost-effective surveillance systems.

This cost-effectiveness approach is a valuable development, and is worth detailed consideration by those concerned with water quality management and/or the design of sampling programmes. However, it does not include what is considered to be the key aspect, that is, what is the economic value to be attached to the achievement of the objectives of a sampling programme. For example, there is generally no point in spending a certain amount of money in executing a satisfactory programme when complete achievement of all objectives could result only in saving a much smaller expenditure or in a trivial improvement in water quality. That is to say, a cost-benefit analysis of sampling programmes is desirable. This approach has been applied to hydrological data (20) but no application to water quality sampling programmes has been described so far as is known. This is not surprising in view of the formidable problems in quantifying the value of information on water quality. Such problems involve economic and management considerations that are beyond the scope of this Section. Accordingly, no practical and helpful advice can be given here, but investigations of the approach seem highly desirable.

4.3 Safety Aspects of Sampling

Those involved in sampling waters and effluents can encounter a wide range of conditions and be subject to a variety of safety and health risks. Every effort must be made to minimize these risks.

Sampling from potentially unsafe sites such as insecure banks should be avoided. If unavoidable, the operation should be carried out by more than one person and appropriate safety precautions taken. Similar precautions should be taken when sampling deep wells or chambers in treatment plant.

Reasonable access to the sampling location in all weathers is important and is essential for routine sampling. If instruments or other equipment are installed on a river bank, locations subject to flooding or vandalism should be avoided; appropriate precautions should be taken where this ideal cannot be achieved.

When samples are to be taken by wading into a stream, river, pond or estuary, the sampler should take into account the presence of soft mud, quicksands, deep holes and swift currents. A wading rod or similar probing instrument is essential to safe wading. By probing ahead the sampler can estimate the current and locate holes, benches, soft mud and quicksand. If a sampler has any doubt on his ability to wade a stream or as to the presence or absence of soft mud, etc, he should attach a safety line to a secure object on the bank or shore.

Traffic is the most serious hazard when working from bridges. Sometimes the bridges have footpaths for pedestrian traffic or catwalks suspended at the side of the bridge but more often than not the sampler must work in the traffic lanes. If it is necessary to interfere with traffic, suitable arrangements may need to be made in advance with the police or local authority. The appropriate warning signs and lights should be used. Personnel should wear fluorescent clothing, and when sampling in darkness or poor visibility carry lights or torches. Even when all precautions are taken, the sampler should still beware of approaching vehicles.

When sampling from bridges over navigable streams, care must also be taken not to cause injury to others. People in boats may not see the small line suspending the sampling device or other equipment until a collision occurs, and it is usual to attach warning pennants to the line. Care must be taken not to lower sampling devices on to passing boats.

When sampling takes place from a boat, several special factors must be taken into account. The craft used must have good stability even if working in sheltered waters, and care must be taken at all times regarding the hazard that might be posed by other ships operating in the same location, for example by their wash or by running down. It is also important that due regard be paid to the normal rules of navigation and common courtesy – a small boat in the middle of a narrow dredged shipping lane will not be very popular with the skipper of a large tanker trying to get past. All sampling personnel should realize the dangers of sampling at sea, and life jackets should be worn at all times while working on deck or on sampling platforms overhanging the side of the vessel; such precautions should be regarded as mandatory in small boats and on larger vessels in bad weather.

Weather conditions should be considered in order to ensure the safety of personnel and equipment. Life jackets and life lines should be worn when sampling large masses of water. Before sampling from ice-covered waters, the location and extent of weak ice should be carefully checked. If self-contained underwater breathing apparatus or other diving equipment is used, it should always be checked and maintained to ensure reliability. The Underwater Associations *Code of Practice for Scientific Diving* (published 1972 and distributed by NERC) lays down the precautions necessary for safe diving operations in terms of personnel and recovery requirements.

Many other situations arise during the sampling of water when special precautions have to be taken to avoid accidents, and this should always be borne in mind. For example, some industrial effluents can be corrosive or contain toxic or flammable materials (sometimes gaseous), and the dangers associated with sewage should not be overlooked. Gas protection equipment, breathing apparatus, resuscitation apparatus and other safety equipment must be available when staff have to enter hazardous atmospheres. In addition, the concentration of oxygen and of any toxic vapour or gas likely to be present should be measured before staff enter enclosed spaces. Intrinsically electrically-safe sampling equipment should be used when sampling in hazardous atmospheres.

In the sampling of steam and hot discharges special care is necessary. The handling of radiological samples also requires special care and the special techniques laid down (21) should be used.

The use of electrical sampling equipment in or near water can present special electrocution hazards. Work procedures, site design and equipment maintenance should be planned so as to minimize these hazards. Special care must be taken when electro-fishing (22).

Further discussion of the safety aspects of sampling can be found in references 23 (natural waters), 24 (potable water treatment plant), 25 (sewage and industrial effluent treatment plant) and 26 (sewers and sewage treatment plant).

4.4 Sampling Position

In choosing the exact position from which samples are required, two aspects are generally involved:

i. the location within the system,*

ii. the exact position at the chosen location.

Some general principles relevant to these aspects are described in Sections 4.4.1 and 4.4.2, respectively, and sampling positions for particular systems of common interest are briefly considered in Section 4.4.3.

* The word system is used here to include, among others, river-basins, streams, rivers, lakes, reservoirs, potable water, sewage and industrial effluent treatment plant, distribution systems, estuarine, coastal and sea-waters, sediments.