STATISTICS IN THE WATER INDUSTRY AND IMPLICATIONS FOR TEACHING

Irene M. Calus
Centre for Advancement of Mathematical Education in Technology
University of Technology
Loughborough, Leicestershire, England

Introduction

In his introductory talk to ICOTS I Professor Barnett[1] discussed the question of who should teach statistics to non-statisticians. His view was that, beyond school level, the teaching of such "service courses" is best done by statisticians, with the important proviso that it is directed towards the particular needs of its recipients. Becoming thoroughly familiar with those needs, through books, journals and contacts with practitioners in the particular subject area, can be quite demanding for the teacher who is primarily a statistician. We are now in the International Water Decade which has the declared objective of bringing safe drinking water and adequate sanitation to as many people as possible by 1990. This is an especially appropriate time, therefore, to look at some applications of statistics which are relevant when the recipients of the teaching are people professionally employed in the water industry.

Some Applications in Hydrology

Water in the form of rain or snow is a natural resource and the management of this resource requires a knowledge of hydrology (the study of precipitation and runoff). We are all aware of the uncertainty that is inherent here and so it is not surprising that books on hydrology usually require some understanding of statistical methods. A basic concept is that of the "return period" of hydrologic events. If, for instance, a year's peak flow is such that, in the long run, it is equalled or exceeded on average in 1 year out of 100, it is said to have an average return period (or recurrence interval) of 100 years. The probability of an exceedance in a particular year is then 1/100. In general, if an event has an average return period of T years and p is the probability of an exceedance in a particular year, then $p=1/T$. Haan[2] also shows an approach using the geometric distribution, with each year regarded as a trial in which an exceedance either occurs or does not occur. The probability that the first exceedance occurs at the xth trial is $(1-p)^{x-1}p$ and the mean of this distribution is $1/p$, ie T. This accords with the concept implied by the terms "return period" and "recurrence interval".

Unfortunately hydrologists habitually omit the word "average" and just speak of a flood with a return period of 50 years or simply a 50-year flood. Such terminology gives the impression that every 50 years a flood of this magnitude occurs, thus undermining the efforts of the statistics teacher who has emphasised that this is not what is meant.

The average return period is used in the design of flood protection works, dams, etc. The structure will fail if, say, the flow for which it is designed is exceeded. If the average return period of this flow is T years and the projected life of the structure is n years, then

$$P(\text{failure}) = 1 - (1 - 1/T)^n$$

Economic and other considerations will have to be taken into account in achieving a satisfactory balance between the risk of failure and the value of T used in the design. The designer relates the value of T to the hydrologic variable in question, eg flow, on the basis of past records. Here the use of probability graph paper is common – in fact, hydrologists were pioneers in this. The form most familiar to other users gives a straight line plot when the variable is normally distributed. The version which has a logarithmic scale for the variable is of more use in hydrology, where the lognormal model finds greater applicability. Of especial interest, also, are extreme value distributions as it is the "highs" and the "lows" which cause problems – floods and droughts. The Gumbel (EV1) and log Pearson Type III distributions are widely used, as shown by Schulz[3] in detail, along with the appropriate graph paper.

Some Applications in Construction

Hydrology is used in defining the requirements of a water resource system. In designing and building a structure to meet them, some knowledge of geology and structural engineering is called for. Here, again, statistical methods are of use, particularly in concrete technology, as Bland[4] shows.

To check that a concrete mix will provide adequate strength, specimens of a standard size and type (cubes in the UK, cylinders in the USA) are made from the mix and tested for strength. Minimum strength requirements are expressed in terms of the "characteristic strength". This is that value of strength below which only a stated proportion of the population of all possible strength measurements of the specified concrete are expected to fall. In BS 5328[5] this proportion is 5% (called the 5% defective level). As strength test results are taken to be normally distributed this gives the characteristic strength as $\mu - 1.64\sigma$, where μ and σ are the population mean and standard deviation respectively. The value of σ depends on the conditions under which the concrete is produced – Lydon[6] suggests that where there is excellent on-site control the value would be between 3.5 and 4.5 Nmm^{-2}, for example. The specified characteristic strength then determines the target value for μ. (It is possible to vary the value of μ by altering the mix proportions.) Control charts are used to check that target values are being maintained.

When a defective level is specified, some familiarity with acceptance sampling will be relevant. BS 5328[5] lays down compliance criteria for deciding whether the 5% defective level for a specified characteristic strength is being met. For instance, for grades of concrete for which the characteristic compressive strength is 20 Nmm^{-2} or more, these criteria are that:

(i) the mean strength of any 4 consecutive test results should exceed the specified characteristic strength by 3 Nmm^{-2},

(ii) the strength from any test result should not be less than the specified compressive strength minus 3 Nmm^{-2}.

A British Standard has been used for illustration here, but similar criteria occur in other standards and recommended codes of practice.

Some Applications in Chemistry and Microbiology

It is not only important that water should be available in sufficient quantity but also that it should be of sufficient quality. Monitoring the chemical quality will usually entail measurement of the pH and analysis for constituents such as chloride and nitrate. Everything that was said in a paper[7] at ICOTS I about statistics for analytical chemists will apply here. For instance, the ASTM Standard for Water[8] discusses the ideas of precision and bias (recommending a t-test to decide whether there is evidence of bias in a method), describes the use of control charts in the laboratory and gives a procedure for dealing with suspected outliers, recommending the test based on the statistic $(x - \bar{x})/s$. All these were dealt with in the previous paper so will not be proceeded with here.

In the case of drinking-water it is not only the chemical quality that must be examined but also the microbiological quality. Only samples of the water can be examined, so once again there is a need to appreciate the statistical implications. Where the "colony" method is used for estimating the number of bacteria present, replicate colony counts from the same sample of water are taken to follow a Poisson distribution. In another method that is used, appropriate volumes of water are inoculated into a number of tubes of nutrient. From the number of tubes which subsequently show a positive reaction the most probable number (MPN) of organisms present in 100 ml of water is found from tables. These tables (which also give 95% confidence limits) can be found in Guidelines for Drinking-Water Quality[9] and other publications dealing with microbiology.

Teaching Practising Engineers and Scientists

Many practising engineers and scientists find that they lack the knowledge of statistics that their work requires. Courses of study can be provided for them in various ways. Mention will now be made of the author's experience in teaching those employed in the water industry.

Loughborough's WEDC (Water and Engineering in Developing Countries) group offers an MSc course requiring one or two years full-time study. At the preparatory stage about 30 hours of classroom time are devoted to inculcating a basic knowledge of statistics. Participants come from a variety of countries and differing academic backgrounds. The CAMET coursebook[10,11] gives them whatever individual help they need outside the classroom. With its interactive approach it is like a teacher, presenting material in easy stages and asking questions to ensure that it has been understood. An accompanying study guide meets the needs of this particular course by indicating which parts of the text are relevant, which can be ignored and which examples should be attempted. This ensures maximum benefit from the time spent.

The WEDC students come from overseas (mainly Africa). For others who do not have so far to travel, intensive courses of a few days duration in individual subjects, such as statistics, are possible. Alternatively, if near enough to the university, attending classes one day a week is a possibility. In either case the CAMET coursebook can help people to continue learning when they are away from the classroom.

Whether study is full-time or part-time, courses in statistics are being followed by these people because a knowledge of the subject is necessary to their work. As in all such situations[12], the teaching must therefore focus on the relevant applications in that field. The intention of this paper is to give some indication of how this can be done. It does not claim to be an exhaustive account of the applications of statistics in the water industry.

References

1. Barnett, V. (1983). Why teach statistics? Proceedings of the First International Conference on Teaching Statistics (Vol. I, pp. 3-15). Sheffield: Teaching Statistics Trust.

2. Haan, C.T. (1977). Statistical methods in hydrology, Ames: The Iowa State University Press.

3. Schulz, E.F. (1974). Problems in applied hydrology. Fort Collins: Water Resources Publications.

4. Bland, J.A. (1985). Statistics for construction students. London and New York: Construction Press.

5. BS 5328. (1981). Methods for specifying concrete including ready-mixed concrete. London: British Standards Institution.

6. Lydon, F.D. (1982). Concrete mix design (2nd ed.). London and New York: Applied Science Publishers.

7. Calus, I.M. (1985). Teaching statistics to analytical chemists. Int. J. Math. Educ. Sci. Technol., 16, 493-499.

8. Annual book of ASTM standards. (1986). Section 11: Water and environmental technology. (Volume 11.01 Water (1)). Philadelphia: American Society for Testing and Materials.

9. Guidelines for drinking-water quality. (1984). (Volume 1 Recommendations). Geneva: World Health Organisation.

10. Bajpai, A.C., Calus, I.M., & Fairley, J.A. (1978). Statistical methods for engineers and scientists. Chichester: John Wiley.

11. Bajpai, A.C. (1981). Metodos estadisticos para estudiantes de ingenieria y ciencias. Mexico: Editorial Limusa.

12. Calus, I.M. (1985). The teaching of statistics to engineers and scientists. Int. J. Math. Educ. Sci. Technol., 16, 261-265.