SOME STATISTICAL CONSIDERATIONS IN THE ASSESSMENT OF COMPLIANCE

S. R. ESTERBY

Lakes Research Branch, National Water Research Institute, Burlington, Ontario, Canada

(Received March 1989)

Abstract. Compliance with criteria limiting the discharge of pollutants or standards of water quality in the receiving water body are assessed by monitoring. In order to set limits or standards, the features of the discharge or the water body to be monitored must be characterized. This generally involves the fitting of a probability distribution to historical data or data from preliminary sampling and the choice of a statistic for the limit or standard. Monitoring data collected to assess compliance are assumed to follow the same distribution as that of the historical or preliminary data. Proper characterization of this latter data will help to ensure that the assumption is met. Statistical methods which either assume a distribution for the quality variable or involve a transformation to a binary variable are compared. The validity of the underlying assumptions in the application of the methods to water or effluent quality data is discussed.

Introduction

Consider the situation where samples are taken to assess whether an acceptable level of quality is being maintained. This involves the setting of a standard or a limit against which the results of the sampling are compared. When the samples do not provide evidence that the source being sampled is at variance with the standard, the source is considered to be in compliance. There are a number of ways in which the standard or limit may be defined and these include external considerations, such as the level which makes the water acceptable for a specific use, and values (determined from either prior sampling of the source to be assessed or another source), considered to represent background levels. It is clear that the definition of the standard or limit involves a measure of location. Due to the inherent variability of effluent or water quality parameters, a measure of the variability must also be incorporated into the method of assessment. A single probability distribution may adequately characterize this variability. However, there will often be structure within a data set or a concomitant variable that will account for some of the variability, and then a model involving both deterministic and random components will be needed.

Methods based on the number of times a quality parameter exceeds a limit and methods which use the distribution of a measure of location of the quality parameter directly have been applied in the assessment of compliance. These will be discussed using examples from the literature. For the methods using the quality parameter distribution directly, one example takes the standard as given, while the other accounts for the variability in both background data and samples taken for the assessment of compliance. The assumptions underlying the methods of analysis will be discussed.

Binomial variables

An application of the theory of hypothesis testing and confidence limits, known as sampling inspection or acceptance sampling (e.g., Brownlee, 1965 or Mandel, 1967), has become standard methodology in the control of the quality of manufactured items. This is based on the binomial distribution. The number of times a limit is exceeded in water or effluent quality assessment has also been treated as a binomial random variable (Warn and Matthews, 1984; Ellis, 1985; Crabtree et al., 1987). The use of the binomial distribution in sampling inspection is briefly reviewed and then the applicability of these methods to the assessment of the compliance of effluent or water quality parameters is considered. This comparison permits clarification of a difficulty expressed by the above mentioned authors. The role of the distribution of the quality variable is shown and estimation of a percentile is considered.

Simple sampling inspection for attributes

These methods have been applied in industry, where items in a manufacturing process can be sampled and classified as defective or nondefective. A random sample of n items is inspected and X, the number of defective items, is compared with a specified standard x₀ (e.g., Brownlee, 1965). If the sample size, n, is small relative to the total number of items, then X follows a binomial distribution with parameters n and p, where p is the fraction of defective items in the batch from which the sample was drawn. The probability that the process is found to be in compliance is

\[ P(X \leq x_0) = \sum_{x=0}^{x_0} \binom{n}{x} p^x (1-p)^{n-x} \]  

By specifying the acceptable fraction defective, p₀, and a fraction defective, p₁, which is considered to be relatively bad, a sampling plan determining n and x₀ in (1), can be found for given type 1 error, α, and type 2 error, β. That is, n and x₀ are chosen so that the probabilities α, the probability of finding the process out of compliance when the fraction defective is acceptable, i.e., p = p₀, and β, the probability of finding the process in compliance when too many items are defective, i.e., p = p₁, are as specified. Alternatively, the operating characteristic curve, a plot of A(p) versus p, where

\[ A(p) = P(X \leq x_0; n, p) \]  

can be used to determine how a specific combination of n and x₀ perform for different values of p.

Binomial methods for quality variables

Let Y be the concentration of the quality parameter in a particular effluent or water body, and L, the upper limit for acceptable concentrations. For example, L might be the 95th percentile determined from a large number of prior samples from the source being assessed. In general, for L defined so that p in the expression