DETERMINATION OF SAMPLE SIZES FOR SETTING TOLERANCE LIMITS

By S. S. Wilks

Princeton University, Princeton, N. J.

1. Introduction. In the mass production of a given product or apparatus piece-part, Shewhart[1] has discussed a practical procedure for detecting the existence of assignable causes of variation in a given quality characteristic of the product as measured by a variable x. For example, x may be the thickness in inches of a washer or the tensile strength in pounds of a small aluminum casting made according to a given set of specifications; x varies in value from washer to washer or from casting to casting. Now suppose assignable causes of variability in x have been detected by Shewhart's procedure and have been sufficiently well eliminated by making appropriate refinements in the manufacturing process so that for all practical purposes the remaining variability may be considered "random," thus allowing us to assume that we have a statistical universe U in which x is a random variable with some distribution law $f(x)$. $f(x)$ is, in general, unknown and cannot be determined until long after the refined manufacturing operation has been under way. Two types of situations arise in practice, one in which x is a discrete variable taking on only certain isolated values as for example 1, 2, 3, $\cdots$, etc. with corresponding probabilities $p(1), p(2), \cdots$, the other being that in which x is essentially a continuous variable over some range with a corresponding probability density function $f(x)$. In this paper we shall consider the latter type of variable.

The problem now arises as to how we should calculate a tolerance range (L_1, L_2) for x from a sample, and how large the sample should be in order for the tolerance range to have a given degree of stability. More specifically, *for a given method of calculating tolerance limits, how large should our sample be in order that the proportion P of the universe included between L_1 and L_2 have an average value a, and will be such that the probability is at least p that P will lie between two given numbers, say b and c?* For example, if a tolerance range is obtained by using a *truncated sample range*, that is by letting L_1 be the greatest of the r smallest values in a sample and L_2 the smallest of the r largest values, r being chosen so that $E(P) = .99$, how large should the sample size, say n, be in order for the probability to be .9 that P would lie between .985 and .995? A similar question can be asked when the setting of only one tolerance limit is under consideration.

[1] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*, D. Van Nostrand Company, New York, 1931.