CUMULATIVE FREQUENCY FUNCTIONS[1]

By Irving W. Burr

Purdue University

1. Introduction. The traditional attack upon the problem of determining theoretical probabilities and expected frequencies has been through the use of the ordinary frequency function. Many such functions have been developed for a wide variety of empirical and theoretical situations. The usual procedure is to find the "best" function of an appropriate type, and then to integrate (either infinitesimal or finite calculus) for the probabilities over the given class intervals or other ranges.

The cumulative frequency function would seem to be theoretically much better adapted to this problem. By definition the cumulative frequency function gives the expected number of cases less than a given value. Hence expected frequencies in any given range are found simply by taking the difference between two values of this function. On the other hand, once the ordinary frequency function has been determined, these expected frequencies must still be obtained by an often times difficult integration. The aim of this paper is to make a contribution toward the direct use of the cumulative function so as to utilize this theoretical advantage.

Some properties and theory of the cumulative function will be presented and the problem of fitting the function considered. A new cumulative function possessing considerable practicability will be discussed and examples given.

2. Characteristics of the cumulative function $F(x)$. Let $F(x_0)$ be the probability that $x < x_0$. Since probabilities are non-negative, $F(x)$ is non-decreasing from $F(-\infty) = 0$ to $F(\infty) = 1$. The two ordinary cases will be considered: (1). $F(x)$ continuous in $(-\infty, \infty)$ and with $F'(x)$ continuous except for a denumerable set of points in $(-\infty, \infty)$, (2). $F(x)$ a step-function with all its discontinuities at the points $nh + d$, $h > d \geq 0$, $n = \cdots, -2, -1, 0, 1, 2, \cdots$.

It is assumed that $F(x)$ has high contact with its asymptotes. Specifically for some j (commonly 3 in practice), there is to exist a $k > j + 1$ such that $F(x) \cdot x^k$ and $[1 - F(x)]x^k$ are ultimately bounded as x tends to $-\infty$ and $+\infty$ respectively. These conditions are obviously satisfied when, as is often convenient, a particular expression is used for $F(x)$ over a range, bounded in one or both directions, while $F(x)$ is defined $\equiv 0$ below, or $\equiv 1$ above such finite lower or upper limits.

For the continuous case (1), the definition gives at once

$$P(a \leq x \leq b) = F(b) - F(a). \tag{1}$$

[1] Presented at the joint meeting of the Institute of Mathematical Statistics, the American Mathematical Society and the Econometric Society at Chicago, Ill., September 2, 1941.