6. Examples. 1) $A = 3.8$, $m = 4$, $\epsilon = .05$. Since A is greater than the value 3.425 in Table I, we compute $a_1 = 2.162$. From Table II we would obtain $A/a_2 = 2.240$ and thus $a_2 = 1.696 < a_1$. 2) $A = 3, m = 4, \epsilon = .02$. Since $A < 4.131$, we read $A/a_2 = 2.600$ from Table II and obtain $a_2 = 1.153$ which will be greater than a_1. 3) $A = 5$, $m = 30$, $\epsilon = .05$. Using the method of section 4 we obtain $\alpha_1 = 1.62$.

REFERENCES

[1] Z. W. Birnbaum, "On random variables with comparable peakedness," *Annals of Math. Stat.*, Vol. 19 (1948), pp. 76–81.
[2] V. J. Francis, "On the distribution of the sum of n sample values drawn from a truncated normal population," *Roy. Stat. Soc. Jour. Suppl.*, Vol. 8 (1946), pp. 223–232

A CERTAIN CUMULATIVE PROBABILITY FUNCTION

By Sister Mary Agnes Hatke, O.S.F.
St. Francis College, Ft. Wayne, Indiana

Graduations of empirically observed distributions show that the cumulative probability function $F(x) = 1 - (1 + x^{1/c})^{-1/k}$ is a practical tool for fitting a smooth curve to observed data. The graduations are comparable with those obtained by the Pearson system, Charlier, and others and are accomplished with simple calculations. Given distributions are graduated by the method of moments. Theoretical frequencies are obtained by evaluation of consecutive values of $F(x)$ by use of calculating machines and logarithms, and by differencing $NF(x)$. No integration nor heavy interpolation is involved, such as may be required in graduation by a classical frequency function. Burr [1] constructed tables of ν_1, σ, α_3, and α_4 values for the function $F(x)$ for certain combinations of integral values of $1/c$ and $1/k$. In these tables curvilinear interpolation must be used in finding an $F(x)$ with desired moments. The writer constructed more extensive tables for the same cumulative function with c and k a variety of real positive numbers less than or equal to one, such that linear interpolation can be used to determine the parameters c and k for an $F(x)$ that has α_3 and α_4 approximately the same as those of the distribution to be graduated. These tables have been deposited with Brown University. Microfilm or photostat copies may be obtained upon request to the Brown University Library.

The writer used the definitions of cumulative moments and the formulas for the ordinary moments ν_1, σ, α_3, and α_4 in terms of cumulative moments as developed by Burr. These latter moments were tabulated for the function $F(x)$ having various combinations of parameters c and k, c ranging from 0.050 to 0.675 and k from 0.050 to 1.000, each at intervals of 0.025. Within these ranges only those combinations of c and k were used which yielded α_3 of approximately 1 or less and α_4 values of 6 or less, since such moments are most common in practice.

It can be verified that over most of the area of the table α_3 values obtained by linear and by curvilinear interpolation on k (or on c) differ by less than 0.001 and values of α_4 by approximately 0.01 or less. If α_3 = constant and α_4 = constant curves are plotted on c, k axes, it will be seen that there exists only one solution (c, k) of the equations $\alpha_3 = B(c, k)$ and $\alpha_4 = C(c, k)$. Furthermore, some α_4 curves intersect two α_3 curves representing the same $|\alpha_3|$. Thus the chance of finding an appropriate function $F(x)$ for graduation is increased since by reversal of scale an $F(x)$ with a positive α_3 may be used to graduate a distribution with a negative α_3, and conversely.

Graduation of an observed frequency distribution is easily accomplished. Linear interpolation on k for a fixed c seems to be the best method for determining

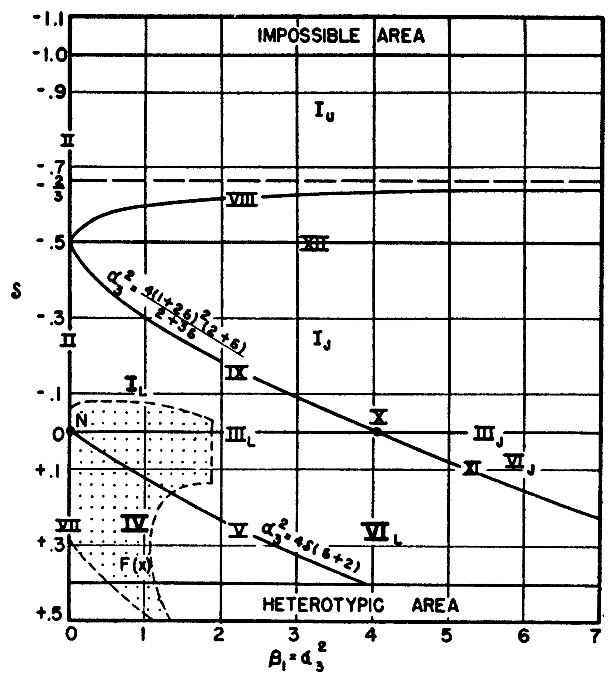

FIG. 1. The α_3^2, δ chart for the Pearson system of frequency curves and the area covered by $f(x) = 1 - (1 + x^{1/c})^{-1/k}$ (subscript L = bell-shaped)

the parameters of an $F(x)$ that has α_3 exactly the same and α_4 nearly the same as the observed α_3 and α_4. If the observed α_3 and α_4 are fairly close to an entry in the table, no interpolation is required. Direct linear interpolation is used to determine ν_1 and σ for the c and k just found. Letting M and S be the mean and standard deviation of the given distribution, the formula,

$$\frac{x - \nu_1}{\sigma} = t = \frac{X - M}{S}$$

is used to translate the class limits X of the given distribution to the corresponding x's of $F(x)$. For any x that is negative the quantity $1 + x^{1/c}$ is taken as one to make $F(-x) = 0$ in accordance with the definition of $F(x)$ [1]. The values of $(1 + x^{1/c})^{-1/k}$ for the various x's are computed by logarithms and differenced to obtain the probabilities for the given class intervals, according to equation

$$P(a \leq x \leq b) = \int_a^b f(x)\,dx = F(b) - F(a).$$

The respective theoretical frequencies are these probabilities multiplied by N, the number of cases.

The headings that proved satisfactory for the columns of the graduation work-sheet are: class intervals (in observed physical units), X (u if unit class-interval is used), f_{obs}, x, $1 + x^{1/c}$, $N/(1 + x^{1/c})^{1/k}$, and f_{th}.

The relation of $F(x)$ to the Pearson system of frequency curves is presented in Figure 1, which is a reproduction of a major part of Craig's chart for α_3^2 and δ [2]. In this chart the parameters of the twelve Pearson curves are expressed in terms of α_3^2 and δ, where $\delta = (2\,\alpha_4 - 3\alpha_3^2 - 6)/(\alpha_4 + 3)$. Values of α_3^2 and δ were computed for $F(x) = 1 - (1 + x^{1/c})^{-1/k}$ in which c and k were assigned the values listed in the α_3, α_4 table. The dotted area superimposed on the Craig chart is that covered by these α_3^2, δ values for $F(x)$. Although it is small in size compared to the total area, it contains a part of the areas representing the three main Pearson curves, I, IV, and VI, as well as the point for the normal curve and part of the line on which lie the points corresponding to the bell-shaped curves of the Type III functions. It also includes transitional Types V and VII. Thus the function $F(x)$ covers part of an important area on the α_3^2, δ chart for the Pearson curves.

The function $F(x)$ was used to graduate satisfactorily several observed distributions classified as Pearson types, including the three main Types, I, IV, and VI, and transitional Types III and VII.

One advantage in the use of this cumulative function $F(x)$ is that it takes but one symbolic form with the area covered, whereas the Pearson-system curves require several different expressions of various complexity requiring identification of type. Furthermore, graduation by a Pearson function generally involves approximate integration or heavy interpolation in the incomplete beta function tables for the evaluation of the integrals of the Pearson functions, whereas graduation by a function $F(x)$ is easily and quickly performed since $F(x)$ only involves two number-parameters readily determined by means of the α_3, α_4 table and straight arithmetic.

The writer is deeply indebted to Professor Irving W. Burr of Purdue University for valuable suggestions in this study.

REFERENCES

[1] I. W. Burr, "Cumulative frequency functions," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 215–232.

[2] C. C. Craig, "A new exposition and chart for the Pearson system of frequency curves," *Annals of Math. Stat.*, Vol. 7 (1936), pp. 16–28.