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.

Furthermore it may be shown that

$$(2) \qquad F(x) = \int_{-\infty}^{x} F'(S)\, dS, \qquad F'(x) = f(x),$$

where $f(x)$ is the ordinary probability function. Also

$$(3) \qquad P(a \le x \le b) = \int_a^b f(x)\, dx.$$

Similarly for the discrete case,

$$(4) \qquad F(x) = \sum_{i<x} {}_h f(i), \qquad \underset{h}{\Delta} F(x) = \frac{1}{h} f(i),$$

$$(5) \qquad P(a \le x \le b) = F(b + h) - F(a) = \sum_{i=a}^{b} {}_h f(i),$$

where a, b are among the values $nh + d$, and $\underset{h}{\Delta}$ is the usual h-difference. In both cases the percentiles are given by the solutions of the equations

$$(6) \qquad F(x) = n/100.$$

Equations (1), (3) and (5) formulate the advantage to the direct use of $F(x)$. which was mentioned in section 1. Related to this is the fact that the process of finding $f(x)$ from $F(x)$ is at least theoretically much simpler than conversely, as (2) and (4) show. The directness of equation (6) is often an advantage also.

The main problems confronting one in trying to utilize these advantages are (a) to find suitable cumulative functions and (b) to find methods of fitting $F(x)$ directly. These are next discussed.

3. Some special functions $F(x)$. An obvious method of attack is to use (2) or (4) on some $f(x)$. The integration involved is precisely the difficulty the writer wishes to avoid. The cumulative function might be sought directly in probability theory. A differential equation incorporating some of the properties of $F(x)$ given in section 2 is

$$(7) \qquad \frac{dy}{dx} = y(1 - y)g(x, y), \qquad y = F(x),$$

where $g(x, y)$ is to be positive for $0 \le y \le 1$ and x in the range over which the solution is to be used. It is to be noted that (7) is very similar to the differential equation

$$\frac{dy}{dx} = y(m - x)g(x, y), \qquad y = f(x),$$

which generates the Pearson system if $g(x, y) = (a + bx + cx^2)^{-1}$

Equation (7) implies the non-decreasing property for $F(x)$, while for many choices of $g(x, y)$, dy/dx will be zero at $y = 0$ and $y = 1$. When $g(x, y) = g(x)$, (7) becomes

$$F(x) = [e^{-\int g(x)\,dx} + 1]^{-1}. \tag{8}$$

Some functions $g(x)$ whose integrals are such that $F(x)$ increases from 0 to 1 on the interval $-\infty < x < \infty$ are c, cx^{-1}, $[(c - x)x]^{-1}$, $c \sec^2 x$ and $c \cosh x$, where $c > 0$. Generalizations of their corresponding $F(x)$ are given below in (10)–(14) respectively.

Another method of attack is to simply consider functions which have the properties given in section 2. The assumption of high contact provides for the existence of certain integrals to be discussed in section 5. Many functions having the required properties are to be found in tables of definite integrals, particularly Bierens de Haan [1].

A list of particular $F(x)$ is given below. In all cases the number of parameters would be increased by two by letting $x = \gamma x' + \delta$, where γ and δ fix the origin and scale. These parameters are determined by $\bar{x}$ and σ. The range of x over which the given expression is to be used is written to the right when it is not $(-\infty, \infty)$. Constants k, r and c are positive real numbers.

$$F(x) = x, \qquad (0, 1), \tag{9}$$

$$F(x) = (e^{-x} + 1)^{-r}, \tag{10}$$

$$F(x) = (x^{-k} + 1)^{-r}, \qquad (0, \infty), \tag{11}$$

$$F(x) = \left[\left(\frac{c - x}{x}\right)^{1/c} + 1\right]^{-r}, \qquad (0, c), \tag{12}$$

$$F(x) = (ke^{-\tan x} + 1)^{-r}, \qquad \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \tag{13}$$

$$F(x) = (ke^{-c \sinh x} + 1)^{-r}, \tag{14}$$

$$F(x) = 2^{-r}(1 + \tanh x)^r, \tag{15}$$

$$F(x) = \left(\frac{2}{\pi} \operatorname{arc\,tan} e^x\right)^r, \tag{16}$$

$$F(x) = 1 - \frac{2}{k[(1 + e^x)^r - 1] + 2}, \tag{17}$$

$$F(x) = (1 - e^{-x^2})^r, \qquad (0, \infty), \tag{18}$$

$$F(x) = \left(x - \frac{1}{2\pi} \sin 2\pi x\right)^r, \qquad (0, 1), \tag{19}$$

$$F(x) = 1 - (1 + x^c)^{-k}, \qquad (0, \infty), \tag{20}$$

Most of these functions have unimodal probability functions $f(x)$, and all of the functions may be readily handled from the calculational standpoint. To check upon their suitability for practical work, the values of α_3 and α_4 for some special cases were obtained approximately by evaluating $F(x)$ at a convenient regular interval, differencing, and using the results as frequencies of a discrete

TABLE I

Calculated α_3 and α_4 for special functions $F(x)$

Function	Parameters	α_3	α_4
(15)	$r = 1$	0	4.01
(16)	$r = 1$	0	3.24
(17)	$k = 1, r = 2$	−.62	4.50
(17)	$k = 2, r = 1$	0	4.11
(17)	$k = 2, r = 2$	−.54	4.22
(18)[2]	$r = 1$	.63	3.25
(19)[2]	$r = 1$	0	2.41

variable. No correction for grouping was made. The values of α_3 and α_4 for several of the above functions are given in Table I, where

$$\mu_j' = \int_{-\infty}^{\infty} x^j f(x)\,dx, \qquad \sum_{i=-\infty}^{\infty}{}_h\, i^j f(i)$$

(21)
$$\mu_j = \int_{-\infty}^{\infty} (x - \mu_1')^j f(x)\,dx, \qquad \sum_{i=-\infty}^{\infty}{}_h\, (i - \mu_1')^j f(i)$$

$$\alpha_j = \frac{\mu_j}{\sigma^j}, \qquad \sigma^2 = \mu_2 .$$

It will be seen that a variety of values of α_4 appear. The values of α_3 vary considerably in most cases as r varies. These functions show promise of being useful after further investigation. The values of α_3 and α_4 for (20) are convenient and adaptable. This function will be discussed in detail in section 6.

4. Methods of fitting $F(x)$. The problem of graduation of data by a cumulative function involves three steps: (a) the selection of the type of function (b) the determination of the parameters of the function, and (c) the graduation. The first two are often determined by such moment characteristics as α_3 and α_4, as in the Pearson system of frequency functions. The third step involves integration or summation if $f(x)$ is used, whereas, once $F(x)$ is fitted, all that remains to be done is evaluation of the function and differencing.

To fit $F(x)$ by moments, it must be possible to determine the parameters of $F(x)$ from $\bar{x}$, σ, α_3 and α_4. The cumulative moments described in the next section, when they can be evaluated, will lead to the values of the $\bar{x}$, σ, α_3 and α_4 for various values of the parameters. If the relations between the parameters and the moments are difficult or impossible to obtain, then tables may be constructed and interpolation used. The usual process would be to use the α_3

[2] The method of moments of section 5 was used for these values.

and α_4 tables to determine the primary parameters such as c, k and r in (9)–(20). Then for the given values of c, k, r, one computes the corresponding values of $\bar{x}$ and σ from their tables, and these are used to obtain the parameters γ and δ for $x = \gamma x' + \delta$. This procedure is illustrated in section 6.

Even when the cumulative moments cannot be evaluated, this method is still possible. Graduation by a small interval is used to construct tables of $\bar{x}$, σ, α_3 and α_4 for varying values of the parameters. Then the table can be used as described above. Thus it is seen that in practice any $F(x)$ can be fitted by this technique.

The usefulness of a cumulative or a probability function depends upon how wide a range of sets of values of the α_i the function covers, and whether such values occur in practice. In most of the functions (9)–(20), α_3 and α_4 are continuous functions of the parameters. If there is only one parameter then only α_3 (or α_4) can be fitted in the range of values of α_3 which the function possesses, but in the case of two parameters both α_3 and α_4 can be fitted. Three or more parameters permit α_5 etc. to be fitted.

5. Cumulative moment theory for $F(x)$. A moment definition for $F(x)$ is now presented. Since for $n \geq 0$, $\lim_{b \to \infty} \int_a^b x^n F(x)\,dx = \infty$, $\int_{-\infty}^{\infty} x^n F(x)\,dx$ cannot be used. However, it was assumed in section 2 that for some $k > j + 1$, $[1 - F(x)]x^k$ is ultimately bounded. Hence, $\lim_{x \to \infty} [1 - F(x)]x^j = 0$. Thus $1 - F(x)$ can be used as a factor when integrating over any interval (a, ∞), a being finite. But the factor $F(x)$ must be used for an interval of the type $(-\infty, b)$. Two integrals are needed, and we define the *cumulative moment*, $M_j(a)$, by

$$M_j(a) = \int_a^{\infty} (x - a)^j [1 - F(x)]\,dx - \int_{-\infty}^{a} (x - a)^j F(x)\,dx, \tag{22}$$

which exists under the assumptions of section 2. The difference of the integrals is used because, as will be shown, this leads to simpler results than could be obtained by addition. If $a = 0$ in (22) then calling $M_j(0) = M_j$,

$$M_j = \int_0^{\infty} x^j [1 - F(x)]\,dx - \int_{-\infty}^{0} x^j F(x)\,dx. \tag{23}$$

Definitions for the discrete case are similar:

$$M_j(a) = h \sum_{i=a+h}^{\infty}{}_h\, (i - a)^{(j)h} [1 - F(i)] - h \sum_{i=-\infty}^{a}{}_h\, (i - a)^{(j)h} F(i), \tag{24}$$

$$M_j = h \sum_{i=h}^{\infty}{}_h\, i^{(j)h} [1 - F(i)] - h \sum_{i=-\infty}^{0}{}_h\, i^{(j)h} F(i), \tag{25}$$

where $i^{(j)h} = i(i - h) \cdots (i - \overline{j - 1}h)$. This function is used because it has simpler properties in the finite calculus than has i^j.

Various relations between the cumulative moments $M_j(a)$ and M_j, and between these and μ_j', μ_j and α_j of (21) are now developed. To express $M_j(a)$ in terms of M_i's, use $(x - a)^j = \sum_{i=0}^{j} {}_jC_i x^{j-i}(-a)^i$. Thus,

$$M_j(a) = \int_a^\infty (x - a)^j[1 - F(x)]\,dx - \int_{-\infty}^a (x - a)^j F(x)\,dx$$

$$= \int_0^\infty (x - a)^j[1 - F(x)]\,dx - \int_{-\infty}^0 (x - a)^j F(x)\,dx - \int_0^a (x - a)^j\,dx$$

$$M_j(a) = \sum_{i=0}^{j} {}_jC_i(-a)^i M_{j-i} + \frac{(-a)^{j+1}}{j+1}. \tag{26}$$

One reason for the minus sign of (22) may be noted here, because in the contrary case the last term would be $\int_0^a (x - a)^j[2F(x) - 1]\,dx$. By translating the origin in (26) to $x = a$, renaming the moments, and replacing $-a$ by a, one obtains

$$M_j = \sum_{i=0}^{j} {}_jC_i a^i M_{j-i}(a) + \frac{a^{j+1}}{j+1}. \tag{27}$$

To bring in ordinary moments, integration-by-parts and (2) are used.

$$\begin{aligned} M_j(a) &= \left[\frac{(x - a)^{j+1}}{j+1}\{1 - F(x)\}\right]_a^\infty + \int_a^\infty \frac{(x - a)^{j+1}}{j+1} f(x)\,dx \\ &\quad - \left[\frac{(x - a)^{j+1}}{j+1} F(x)\right]_{-\infty}^a + \int_{-\infty}^a \frac{(x - a)^{j+1}}{j+1} f(x)\,dx \\ &= \frac{1}{j+1}\int_{-\infty}^\infty (x - a)^{j+1} f(x)\,dx, \end{aligned} \tag{28}$$

the first and third quantities vanishing because of the contact assumption. A second justification of the minus sign of (22) appears here, since if a positive sign were used, the fourth term would have been subtracted and the integrals would not combine into (28). Expansion of $(x - a)^{j+1}$ in powers of x and $x - \mu_1'$ yields respectively

$$M_j(a) = \frac{1}{j+1}\sum_{i=0}^{j+1} {}_{j+1}C_i(-a)^i \mu_{j+1-i}', \tag{29}$$

$$M_j(a) = \frac{1}{j+1}\sum_{i=0}^{j+1} {}_{j+1}C_i(\mu_1' - a)^i \mu_{j+1-i}. \tag{30}$$

Also setting $a = 0$,

$$M_j = \frac{1}{j+1}\mu_{j+1}' \tag{31}$$

$$M_j = \frac{1}{j+1}\sum_{i=0}^{j+1} {}_{j+1}C_i \mu_1'^i \mu_{j+1-i}. \tag{32}$$

It may be shown that the existence of $M_j(a)$ implies that of the μ'_i $i = 1, \cdots, j + 1$, and conversely if μ'_j exists then so do the $M_i(a)$ $i = 0, \cdots, j - 1$. The following formulas are obtained by the opposite integration by parts, taking two different forms for $\int f(x)\,dx$: $F(x)$ and $-[1 - F(x)]$, to avoid indeterminate situations.

$$\mu'_j = \int_a^\infty x^j f(x)\,dx + \int_{-\infty}^a x^j f(x)\,dx$$

$$= -[x^j\{1 - F(x)\}]_a^\infty$$

$$+ j \int_a^\infty x^{j-1}[1 - F(x)]\,dx + [x^j F(x)]_{-\infty}^a - j \int_{-\infty}^a x^{j-1} F(x)\,dx.$$

The first and third terms vanish by the contact assumption. Then using $(x - a + a)^{j-1}$ for x^{j-1},

$$\mu'_j = j \sum_{i=0}^{j-1} {}_{j-1}C_i\, a^i M_{j-1-i}(a) + a^j, \qquad j > 0. \tag{33}$$

Also in the same manner

$$\mu_j = j \sum_{i=0}^{j-1} {}_{j-1}C_i (a - \mu'_1)^i M_{j-1-i}(a) + (a - \mu'_1)^j, \qquad j > 1,$$

or

$$\mu_j = j \sum_{i=0}^{j-1} {}_{j-1}C_i [-M_0(a)]^i M_{j-1-i}(a) + [-M_0(a)]^j, \qquad j > 1, \tag{34}$$

using (29) $M_0(a) = \mu'_1 - a$. Letting $a = 0$,

$$\mu'_j = jM_{j-1}, \qquad j > 0 \tag{35}$$

$$\mu_j = j \sum_{i=0}^{j-1} {}_{j-1}C_i (-M_0)^i M_{j-1-i} + (-M_0)^j, \qquad j > 1. \tag{36}$$

An interesting graphical property of $F(x)$ may be seen from (35) $j = 1$ by taking $\mu'_1 = 0$. Then $M_0 = 0$ and hence $\int_0^\infty [1 - F(x)]\,dx = \int_{-\infty}^0 F(x)\,dx$. Thus the mean is that ordinate which equates the two areas bounded by (i) $y = F(x)$, $y = 0$ and $x = \mu'_1$ and (ii) $y = F(x)$, $y = 1$ and $x = \mu'_1$.

It is worth noting that the expressions (34) and (36) have the same coefficients, independent of a. This is to be expected because of the invariance of μ_j under translation.

If $a = \mu'_1$ then (30) simplifies to $M_j(\mu'_1) = \frac{1}{j + 1}\mu_{j+1}$. Lastly, expressions for α_j's in terms of the $M_i(a)$'s are given.

$$\alpha_3 = \frac{3M_2(a) - 6M_1(a)M_0(a) + 2M_0^3(a)}{[2M_1(a) - M_0^2(a)]^{3/2}}$$

$$\alpha_4 = \frac{4M_3(a) - 12M_2(a)M_0(a) + 12M_1(a)M_0^2(a) - 3M_0^4(a)}{[2M_1(a) - M_0^2(a)]^2} \tag{37}$$

$$\alpha_j = \frac{j\sum_{i=0}^{j-1} {}_{j-1}C_i[-M_0(a)]^i M_{j-1-i}(a) + [-M_0(a)]^j}{[2M_1(a) - M_0^2(a)]^{j/2}}.$$

The discrete case has been carried through in an exactly similar manner, by the use of finite rather than infinitesimal calculus. Only the results will be stated here. The notation used is that of Steffensen [2].

$$M_j(a) = \sum_{r=0}^{j} {}_jC_{j-r}[a + (r-1)h]^{(r)h}(-1)^r M_{j-r} + \frac{(-1)^{j+1}}{j+1}[a + (j-1)h]^{(j+1)h}, \qquad j > 0 \tag{38}$$

$$M_0(a) = M_0 + a \tag{39}$$

$$M_j = \sum_{r=0}^{j} {}_jC_r a^{(r)h} M_{j-r}(a) + \frac{(a+h)^{(j+1)h}}{j+1} \tag{40}$$

$$M_j(a) = \frac{1}{j+1}\sum_{r=0}^{j+1} \mu_r' \sum_{k=r}^{j+1} {}_kC_r \frac{D^k O^{(-j-1)}}{k!}(-h)^{j+1-k}(h-a)^{k-r}, \qquad j > 0 \tag{41}$$

$$M_0(a) = \mu_1' - a \tag{42}$$

$$M_j(a) = \frac{1}{j+1}\sum_{r=0}^{j+1} \mu_r \sum_{k=r}^{j+1} {}_kC_r \frac{D^k O^{(-j-1)}}{k!}(-h)^{j+1-k}(\mu_1' + h - a)^{k-r}, \qquad j > 0 \tag{43}$$

$$M_j = \frac{1}{j+1}\sum_{r=0}^{j+1} \mu_r' h^{j+1-r} \sum_{k=r}^{j+1} {}_kC_r \frac{D^k O^{(-j-1)}}{k!}(-1)^{k+j+1}, \qquad j > 0 \tag{44}$$

$$M_0 = \mu_1' \tag{45}$$

$$M_j = \frac{1}{j+1}\sum_{r=0}^{j+1} \mu_r \sum_{k=r}^{j+1} {}_kC_r \frac{D^k O^{(-j-1)}}{k!}(-h)^{j+1-k}(\mu_1' + h)^{k-r}, \qquad j > 0 \tag{46}$$

$$\mu_j' = a^j + \sum_{r=0}^{j-1} M_r(a) \sum_{k=r+1}^{j} h^{j-k}\frac{\Delta^k O^j}{k!}\, {}_kC_r(k-r)(a-h)^{(k-r-1)h} \tag{47}$$

$$\mu_j = [-M_0(a)]^j + \sum_{r=0}^{j-1} M_r(a) \sum_{k=r+1}^{j} h^{j-k}\frac{\Delta^k O^j}{k!}\, {}_kC_r(k-r)[-M_0(a) - h]^{(k-r-1)h} \tag{48}$$

$$\mu_j' = \sum_{r=0}^{j-1} M_r h^{j-r-1} \sum_{k=r+1}^{j} \frac{\Delta^k O^j}{k!}\frac{k!}{r!}(-1)^{k-r-1} \tag{49}$$

$$\mu_j = (-M_0)^j + \sum_{r=0}^{j-1} M_r \sum_{k=r+1}^{j} h^{j-k}\frac{\Delta^k O^j}{k!}\, {}_kC_r(k-r)[-M_0 - h]^{(k-r-1)h}. \tag{50}$$

The writer has verified that under certain fairly general conditions the discrete case (38)–(50) approaches the continuous case (26)–(36) as $h \to 0$.

The following three propositions are merely stated without proof since they follow so immediately from (23), (25), (31), (45), (21), (2) and (4).

PROPOSITION 1: *Given a set of functions* $F_i(x)$ *and positive constants* $k_i\, i = 1, \cdots, n$ *for which* $\sum_{i=1}^{r} k_i = 1$, *then for* $F(x) = \sum_{i=1}^{r} k_i F_i(x)$, $M_j = \sum_{i=1}^{r} k_i\, {}_iM_j$ *if all the latter exist.*

PROPOSITION 2. *In the above notation, if all the* ${}_i\mu_1'$ *are equal, then* $\mu_j = \sum_{i=1}^{r} k_i\, {}_i\mu_j$, *when the latter exist.*

PROPOSITION 3. *If in addition to the above hypotheses, all the* ${}_i\mu_2$ *are equal, then*

$$\alpha_j = \sum_{i=1}^{r} k_i\, {}_i\alpha_j. \tag{51}$$

These propositions are sometimes convenient in forming a linear combination of functions $F(x)$, to obtain a function with desired properties. It may be noted that Proposition 1 is still algebraically true even with negative k_i's, but these might give negative derivatives $f(x)$ for $F(x)$.

6. An algebraic function, $F(x) = 1 - \dfrac{1}{(1 + x^c)^k}$. This simple algebraic cumulative function will be discussed in detail. The α_i can be calculated directly by the application of (23), (36) and (21). The resulting α_3 and α_4 values cover a broad range, within which those of many empirical and theoretical distributions lie. A method of finding such cumulative functions with desired α_3 and α_4 will be given. Several graduations are presented for illustration.

This function appears in Bierens de Haan [1] and has the desired properties. The writer has not yet found a probability justification for the function. However, since the α_i are so close to those of functions which can be so supported, it seems that it may eventually prove to be at least some definite approximation to a probability situation.

The complete definition is

$$\begin{aligned} F(x) &= 1 - \frac{1}{(1 + x^c)^k} \qquad & x \geq 0 \\ &= 0 & x < 0, \end{aligned} \tag{52}$$

where $c,\, k \geq 1$ are real numbers. The probability function

$$F'(x) = f(x) = \frac{kcx^{c-1}}{(1 + x^c)^{k+1}}, \tag{53}$$

is unimodal at $x = \left(\dfrac{c - 1}{ck + 1}\right)^{1/c}$ if $c > 1$, and L-shaped if $c = 1$.

Use of (23) on (52) gives

$$M_j = \int_0^\infty \frac{x^j\,dx}{(1 + x^c)^k}, \qquad j < ck - 1. \tag{54}$$

But from Bierens de Haan [1]

$$\int_0^\infty \frac{x^{g-1}\,dx}{(1 + x^c)^k} = \frac{(c - g)^{[k-1]_c}\,\pi}{c^k(k - 1)!\,\sin\,(g\pi/c)} \qquad g < c, \tag{55}$$

where $a^{[r]_c} = a(a + c) \cdots (a + \overline{r - 1}c)$.
Hence

$$M_j = \frac{(c - j - 1)^{[k-1]_c}\,\pi}{c^k(k - 1)!\,\sin \dfrac{j + 1}{c}\,\pi}, \qquad j < c - 1. \tag{56}$$

However, if $j \geq c - 1$ then (55) can still be used through reducing the exponent of x by $x^{j-c}(1 + x^c) - x^{j-c} = x^j$. (56) is only good for integral values of k. A more general formula is obtainable by letting $(1 + x^c) = 1/s$. Then

$$M_j = \frac{1}{c}\int_0^1 (1 - s)^{(j+1)/c-1} s^{k-(j+1)/c-1}\,ds$$

$$= \frac{1}{c}B\left(\frac{j + 1}{c}, k - \frac{j + 1}{c}\right)$$

$$M_j = \frac{\Gamma\left(\dfrac{j + 1}{c}\right)\Gamma\left(k - \dfrac{j + 1}{c}\right)}{c\Gamma(k)}, \tag{57}$$

for $j = 0, 1, \cdots$ up through $j < ck - 1$, and c, k any real numbers ≥ 1. To determine the μ_j values the easiest way is to compute the values of the M_j by (56) or (57), and then to use (36):

$$\mu_2 = 2M_1 - M_0^2, \qquad \mu_3 = 3M_2 - 6M_1M_0 + 2M_0^3,$$

$$\mu_4 = 4M_3 - 12M_2M_0 + 12M_1M_0^2 - 3M_0^4, \text{ etc.}$$

Having these, definitions (21) are used for the α_j.

The results for some integral values of k and c are given in Tables II and III. These computations were made from (56). Formula (57) shows that for a fixed c, M_j for $k + 1$ is obtained by multiplying M_j for k by $\dfrac{kc - j - 1}{kc}$. This recursion relation is very helpful in the computation, because it enables all of the values of the M_j's for a given c to be found from those for the lowest value of k for which M_j exists. The values which need to be copied down in the computation for μ_1', σ, α_3, α_4, by a calculating machine are M_0, M_1, M_2, M_3, M_0^2, M_0^3, M_0^4, $6M_0$, $12M_0$, $12M_0^2$, μ_2, σ, σ^3, σ^4, μ_3, α_3, μ_4, α_4. Because of heavy cancellation, especially in μ_3 and μ_4, it seemed advisable to use eight significant figures throughout. Eight-place sines were obtained from Gifford [3]. The values of the M_j for $k = 11$ were also checked by eight-place logarithms [4]. These verify the values of the M_j for $k < 11$ because of the recurrence calculation.

TABLE II

Mean μ_1', *and Standard Deviation* σ *for* $F(x) = 1 - \dfrac{1}{(1 + x^c)^k}$

(In each cell the upper number is μ_1' *and the lower number is* σ*)*

k \ c	1	2	3	4	5	6	7	8	9	10
1	— —	1.57080 —	1.20920 .97787	1.11072 .58060	1.06896 .42265-	1.04720 .33552	1.03438 .27953	1.02617 .24019	1.02060 .21089	1.01664 .18815
2	1.00000 —	.78540 .61899	.80613 .39533	.83304 .30239	.85517 .24794	.87266 .21116	.88661 .18433	.89790 .16375	.90720 .14742	.91498 .13411
3	.500000 .86603	.58905- .39118	.67178 .29349	.72891 .24029	.76965- .20461	.79994 .17852	.82328 .15847	.84178 .14253	.85680 .12953	.86923 .11872
4	.33333 .47140	.49087 .30393	.59714 .24784	.66817 .21077	.71834 .18344	.75550 .16234	.78408 .14555-	.80671 .13187	.82507 .12053	.84025 .11097
5	.25000 .32275	.42951 .25596	.54737 .22070	.62641 .19269	.68242 .17028	.72402 .15220	.75607 .13743	.78150- .12517	.80215- .11487	.81925- .10610
6	.20000 .24495-	.38656 .22488	.51088 .20220	.59509 .18010	.65513 .16103	.69989 .14505-	.73447 .13168	.76196 .12043	.78432 .11087	.80286 .10266
7	.16667 .19720	.35435- .20274	.48250 .18851	.57029 .17064	.63329 .15403	.68045- .13962	.71698 .12731	.74609 .11682	.76980 .10782	.78948 .10004
8	.14286 .16496	.32904 .18599	.45952 .17783	.54992 .16316	.61520 .14846	.66425- .13528	.70235 .12383	.73276 .11394	.75758 .10539	.77820 .09796
9	.12500 .14174	.30847 .17275	.44038 .16918	.53274 .15704	.59982 .14389	.65041 .13171	.68981 .12095-	.72131 .11156	.74706 .10338	.76848 .09623
10	.11111 .12423	.29134 .16197	.42407 .16197	.51794 .15190	.58649 .14002	.63836 .12869	.67886 .11851	.71130 .10954	.73783 .10168	.75994 .09477
11	.10000 .11055	.27677 .15297	.40993 .15584	.50499 .14749	.57476 .13669	.62772 .12608	.66916 .11640	.70240 .10780	.72964 .10021	.75234 .09351

It will be seen from Table II that in most cases the values of α_3 and α_4 lie within useful ranges. The graph shows the general relationship between α_3, k and c. The curves are the traces of planes $k = 1, 2, \cdots$ upon the surface $\alpha_3 = G(c, k)$. Other traces would contain all pairs (c, k) giving a fixed α_3.

TABLE III

Skewness α_3 *and Kurtosis* α_4 *for* $F(x) = 1 - \dfrac{1}{(1 + x^c)^k}$

(In each cell the upper number is α_3 *and the lower number is* α_4*)*

k \ c	1	2	3	4	5	6	7	8	9	10
1	— —	— —	— —	4.285^- —	2.485 29.56	1.820 14.77	1.458 10.36	1.225^- 8.342	1.060 7.215	.937 6.510
2	— —	4.086 —	1.589 10.81	.956 5.937	.635 4.630	.434 4.106	.294 3.859	.190 3.736	.109 3.673	.044 3.646
3	— —	1.909 12.46	.919 5.132	.513 3.871	.277 3.485	.119 3.358	$.005^-$ 3.329	−.083 3.343	−.152 3.376	−.208 3.418
4	7.071 —	1.432 7.356	.682 4.036	$.335^-$ 3.363	$.125^-$ 3.189	−.019 3.169	−.125 3.205	−.207 3.263	−.271 3.327	−.325 3.393
5	4.648 73.80	1.218 5.832	.559 3.604	.238 3.154	.040 3.070	−.097 3.098	−.199 3.165^-	−.277 3.243	−.340 3.324	−.391 3.401
6	3.810 38.67	1.094 5.118	.484 3.380	.178 3.045	−.013 3.010	−.147 3.065	−.246 3.150^-	−.323 3.241	−.384 3.330	−.435 3.416
7	3.381 27.86	1.014 4.707	.433 3.245^-	.136 2.979	−.051 2.975	−.181 3.048	−.279 3.144	$-.355^-$ 3.244	−.415 3.339	−.465 3.430
8	3.118 22.73	.958 4.443	.396 3.154	.106 2.936	−.078 2.953	−.207 3.039	−.303 3.143	−.378 3.248	−.438 3.349	−.488 3.442
9	2.940 19.76	.916 4.258	.368 3.091	.083 2.906	−.098 2.938	−.226 3.033	−.322 3.143	−.396 3.252	−.456 3.357	$-.505^-$ 3.453
10	2.811 17.83	.884 4.122	.347 3.043	.065 2.883	$-.115^-$ 2.928	−.242 3.030	−.336 3.144	−.410 3.257	−.470 3.364	−.519 3.462
11	2.714 16.48	.858 4.018	.329 3.006	.050 2.866	−.128 2.920	−.254 3.027	−.348 3.146	−.422 3.261	−.481 3.371	−.530 3.470

The surfaces for μ_1', σ and α_4 are more irregular. The problem of determining a cumulative function with $\alpha_3 = a$ and $\alpha_4 = b$ is equivalent to the problem of determining a point of intersection of the curves

$$\alpha_3 = G(k, c), \qquad \alpha_3 = a$$
$$\alpha_4 = H(k, c), \qquad \alpha_4 = b. \tag{58}$$

Direct algebraic solution of this system appears very difficult, and other techniques must be resorted to.

One method is to use only integral values of k, and then for each k interpolate for the value of c giving the desired α_3. For such pairs of c and k, find α_4 by interpolation. Then choosing the pairs having α_4 just above and just below the desired one, the proper linear combination (51) is taken. This gives a combination function which has both α_3 and α_4 at the desired values. This combination

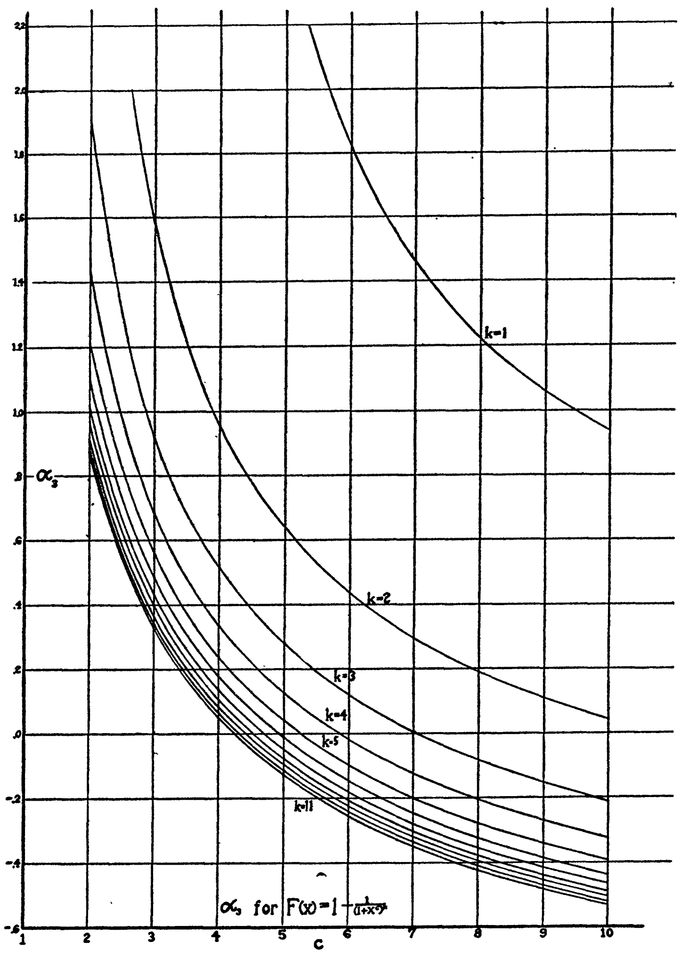

FIGURE I

will be an approximation to the single function with non-integral k, having the given α_3 *and* α_4. This method of linear combinations might be extended to fit α_5 by using three integral values of k.

The interpolations may be done graphically by use of Figure 1 and others like it. Or one may use Stirling's formula [5]. The interpolation for c from α_3 is backwards, while that for α_4 from c is direct. Sometimes it is more accurate to use Newton's formula [5, p. 36] when the values in one direction increase rapidly.

Use of a single function $F(x)$ for a graduation is easily accomplished. First, obtain the c and k to be used so that α_3 is correct and α_4 is as close to the given value as possible. Then determine μ_1' and σ from Table II by interpolation. Change the scale and origin of the original values of the variable X to those x's corresponding to $F(x) = 1 - 1/(1 + x^c)^k$, through

$$\frac{x - \mu_1}{\sigma} = t = \frac{X - M}{S}, \tag{59}$$

where M and S are the mean and standard deviation of the given distribution. Now compute the values of $1/(1 + x^c)^k$ for the various values of x. The differences of these results are equal to the differences of $F(x)$, which by (1) are the probabilities for the given ranges of X. Multiplication by the total frequency will yield the theoretical frequencies, if desired.

If the graduation is to be done by a combination of two functions, the work is carried out for each as described above, and then the frequencies are combined by the same linear combination as that by which the component α_4's must be combined to give the desired α_4. This may readily be seen by considering the separate cumulative functions in terms of the standard variable t, whence the means and σ's are 0 and 1 and (51) is applicable. Then the differences of $G(t) = k_1G_1(t) + k_2G_2(t)$ are sought. But these can be found by taking the same linear combination of the separate differences of the functions $G_1(t)$ and $G_2(t)$. However, these values are merely computed from their respective sets of x values.

For illustration, three graduations are given. The first is a highly normal distribution of heights from Rietz [5, p. 98ff.]. For this distribution, $M = .02085$, $S = 2.5723$, $\alpha_3 = -.0124$, $\alpha_4 = 3.149$. The graduation was done by taking the function $F(x) = 1 - \dfrac{1}{(1 + x^6)^4}$ which has the nearly normal characteristics $\alpha_3 = -.019$, $\alpha_4 = 3.169$. The object was to take a simple cumulative function with integral k and c to show how a satisfactory job can be done on a normal distribution. For this function $\mu_1' = .75550$ and $\sigma = .16234$. Then

$$x = .063110X + .75418,$$

into which are substituted the X class-limits -11.5, -10.5, etc. From these, corresponding values of $\dfrac{8585}{(1 + x^6)^4}$ are calculated and differenced to give the theoretical frequencies for the 8585 cases. The results are given in Table IV.

The fit obtained by use of $F(x)$ is good. One comparison test is that of χ^2. The eight classes -11, -10, -9, 9, 10, 11, 12, 13 were grouped together. The results were

$$\chi_F^2 = 21.210, \qquad \chi_N^2 = 23.479,$$

as compared to

$$P(\chi^2 > 22.31) = .10,$$
$$P(\chi^2 > 19.31) = .20,$$

for 15 degrees of freedom (18 classes minus 3 for linear restrictions). One reason for the somewhat lower χ^2 for $F(x)$ may be that its α_3 and α_4 are closer to

TABLE IV

X	Observed frequency [5]	Graduated frequency by $F(x)$	Graduated frequency of normal [5]
−11.0		.00	.16[3]
−10.0	2	.43	.67
−9.0	4	3.23	2.84
−8.0	14	13.17	10.30
−7.0	41	39.81	32.11
−6.0	83	97.87	86.03
−5.0	169	206.72	198.17
−4.0	394	385.34	392.43
−3.0	669	639.55	668.11
−2.0	990	941.98	977.92
−1.0	1223	1216.47	1230.63
.0	1329	1353.98	1331.41
1.0	1230	1278.39	1238.41
2.0	1063	1013.80	990.33
3.0	646	676.12	680.86
4.0	392	384.41	402.44
5.0	202	190.83	204.51
6.0	79	85.19	89.35
7.0	32	35.24	33.56
8.0	16	13.87	10.84
9.0	5	5.33	3.01
10.0	2	2.01	.72
11.0		.77	.15
12.0		.30	.03
13.0		.19[3]	
Total	8585	8585.00	8584.99

those of the observed distribution than are those of the normal function. This gives a better fit in the tails of the distribution. Nevertheless, this example does illustrate how one of the simplest of the cumulative functions with "normal" characteristics can be used without specifically fitting α_3 and α_4. It may also be mentioned that $F(x)$ for $c = 5$, $k = 6$ has α_3 and α_4 even closer to the normal

[3] Total of stump frequency.

TABLE V

X	Observed frequency	$F(x)$, $k = 4$ $c = 3.228; f_4$	$F(x)$, $k = 5$ $c = 2.944; f_5$	$F(x) = .3063f_4 + .6937f_5$	Type III [5]
−8.0	3	.00	.00	.00	
−7.0	9	.86	.00	.27	2
−6.0	46	39.58	25.07	29.52	27
−5.0	167	180.78	175.27	176.96	142
−4.0	372	433.86	445.79	442.13	410
−3.0	718	768.83	791.72	784.71	799
−2.0	1186	1116.06	1134.52	1128.86	1186
−1.0	1462	1383.06	1384.99	1384.40	1441
.0	1498	1492.04	1477.86	1482.20	1502
+1.0	1460	1419.70	1399.70	1405.83	1385
2.0	1142	1205.81	1190.80	1195.40	1158
3.0	913	926.59	921.47	923.01	891
4.0	642	654.00	656.82	655.96	641
5.0	435	430.66	436.78	434.90	434
6.0	235	268.70	274.63	272.81	280
7.0	167	161.10	165.27	163.99	173
8.0	133	93.99	96.23	95.55	102
9.0	47	53.88	54.77	54.50	59
10.0	29	30.62	30.70	30.68	33
11.0	13	17.37	17.07	17.16	18
12.0	9	9.86	9.46	9.58	9
13.0	5	5.64	5.26	5.38	5
14.0	8	3.26	2.93	3.03	2
15.0	2	1.89	1.66	1.73	1
16.0		1.12	.93	.99	1
17.0		.66	.53	.57	
18.0		.41	.31	.34	
19.0		.24	.18	.20	
20.0		.16	.11	.13	
21.0		.27[4]	.17[4]	.20[4]	
Total........	10701	10701.00	10701.00	10700.99	10701

TABLE VI

Observed [6]	Type III [6]	Type A [6]	Edgeworth [6]	$F(x)$
3	4	5	4	4
20	17	22	17	19
38	42	47	42	42
63	59	60	59	56
51	53	50	53	52
29	33	27	32	34
21	15	13	15	16
4	5	4	6	5
0	1	1	2	1
1	0	0	1	0
230	229	229	231	229
χ^2	4.54	7.55	5.86	4.03

[4] Stump frequency.

values, but it does not give quite as good a fit because it tends to decrease too rapidly on the left.

The second example is also from Rietz [5, p. 108ff.]. For this distribution, $M = .68835$, $S = 2.9480$, $\alpha_3 = .583$ and $\alpha_4 = 3.698$. Two functions were used with $k = 4$ and $k = 5$. By interpolation

		μ_1'	σ	α_3	α_4
$k = 5$	$c = 2.944$	.54200	.22247	.583	3.655
$k = 4$	$c = 3.228$	.61577	.23823	.583	3.795

Because of the rather rapid increases for smaller values of c, Newton's formula [5, p. 36] yields better approximations than Stirling's [5, p. 38 (12)]. The graduation for each function is carried out as above, and since

$$.3063 \cdot 3.795 + .6937 \cdot 3.655 = 3.698,$$

the linear form

$$.3063 f_4^i + .6937 f_5^i = f^i,$$

is used.

Table V gives the component and combined frequencies, and also the frequencies from a Type III. χ^2 for both are very high even though the fit appears reasonably good on a graph. This result is due to classes 6 and 8 which tend to cause a high χ^2 for any distribution function of a small number of parameters. The example, however, does show that $F(x)$ can be used to graduate a skewed distribution.

It is to be further noted that the component functions were used only to obtain an approximation to a single function with $4 < k < 5$, for which α_3 and α_4 are simultaneously correct. When tables more complete than Tables II and III are available, such a single function can be found.

The third example of graduations is from Elderton [6]. The measures were treated as a discrete variable in computing α_3 and α_4. A single function $c = 3.102$, $k = 11$ was used. This function had α_3 at the observed value of .2936, while α_4 was 2.973 as compared to the observed 2.986. The results along with those by classical methods are shown in Table VI. The above χ^2 were obtained by grouping the first and the last three class frequencies. The values are approximate because of rounding. However, they do show that $F(x)$ does a comparable graduation.

Besides aiding in the problem of graduation, this cumulative function should prove of value in the approximation of known or population distributions, as for example, $(p + q)^n$. However much more work needs to be done before this can be more than a conjecture.

7. Conclusion. This paper has stressed the advantages obtained by the direct use of the cumulative function. A number of useful functions have been considered. A general method for fitting any cumulative function by the construction of a table has been suggested. A particular method depending upon the use of certain new cumulative moments has been given. Making use of this theory a certain simple algebraic function has been discussed in detail, and its use in graduations explained.

The writer wishes to convey his sincere thanks to Professor Harry C. Carver, whose counsel was most helpful.

BIBLIOGRAPHY

[1] D. Bierens de Haan, *Nouvelles Tables d'Intégrales Définies*, 1867 (Corrected edition 1939), G. E. Stechert, New York.
[2] J. F. Steffensen, *Interpolation*, 1927, Williams & Wilkins Co., Baltimore, Md., pp. 57 ff.
[3] E. Gifford, *Natural Sines*, 1914, Abel Heywood, Manchester.
[4] J. Bauschinger, J. Peters, *Logarithmic-Trigonometric Tables with Eight Places*, I, II, 1911, Wilhelm Engelmann, Leipzig.
[5] H. L. Rietz, ed., *Handbook of Mathematical Statistics*, 1924, Houghton-Mifflin, Cambridge, Mass.
[6] W. P. Elderton, *Frequency Curves and Correlation*, (3rd edition), 1938, Cambridge University Press, London, p. 134 ff.