high degree of confidence when they are used as tests of significance for index numbers, since in nearly all time series there exists an appreciable degree of serial correlation, persistence, or lack of independence among successive items of any sample.

**4. Bibliographical note.** Certain aspects of the sampling distribution of the geometric mean have been discussed by Burton H. Camp.[5] Attempts to derive forms for estimating the standard errors of index numbers have been made by Truman L. Kelley[6] and Irving Fisher,[7] and an empirical study of the sampling fluctuations of indexes has been made by E. C. Rhodes.[8] Although various special tests of significance for time series have been proposed,[9] at the present time no generally satisfactory procedure has appeared.

HUNTER COLLEGE,
NEW YORK, N. Y.

[5] Burton H. Camp, "Notes on the distribution of the geometric mean," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 221–226.

[6] Truman L. Kelley, "Certain Properties of Index Numbers," *Quarterly Publications of Am. Stat. Assn.*, Vol. 17, New Series 135, Sept., 1921, pp. 826–841.

[7] Irving Fisher, *The Making of Index Numbers*, Houghton Mifflin Company, New York, 1927, 3d ed., pp. 225–229, 342–345, and Appendix I, pp. 407 and 430 *f*.

[8] E. C. Rhodes, "The precision of index numbers," *Roy. Stat. Soc. Jour.*, Vol. 99 (1936), Part I, pp. 142–146, and Part II, pp. 367–369.

[9] Some of the more recent papers dealing with this matter are: G. Tintner, "On tests of significance in time series," *Annals of Math. Stat.*, Vol. 10 (1939), pp. 139–143; "The analysis of economic time series," *Am. Stat. Assn. Jour.*, Vol. 35 (1940), pp. 93–100; L. R. Hafstad, "On the Bartels technique for time-series analysis, and its relation to the analysis of variance," *Am. Stat. Assn. Jour.*, Vol. 35 (1940), pp. 347–361; and Lila F. Knudsen, "Interdependence in a series," *Am. Stat. Assn. Jour.*, Vol. 35 (1940), pp. 507–514.

# A NOTE ON THE USE OF A PEARSON TYPE III FUNCTION IN RENEWAL THEORY

BY A. W. BROWN

One of the methods suggested by A. J. Lotka[1] for the derivation of the renewal function may be briefly summarized as follows.

The method consists of dissecting the total renewal function into "generations". The original installation constitutes the zero generation, the units introduced to replace disused units of the zero generation constitute the first generation, renewal of these the second, and so on. Let $f(x)$ be the "mortality" function, the same for all generations. $f(x)$ is a function satisfying the usual conditions of a distribution function. Adopting Lotka's notation, let $N$ be the number of units in the original collection, $B_1(t)\,dt$ the number of objects intro-

[1] A. J. Lotka, "A Contribution to the Theory of Self Renewing Aggregates, With Special Reference to Industrial Replacement," *Annals of Math. Stat.*, Vol. 10 (1939), p. 1.

duced between times $t$ and $t + dt$ and belonging to the first generation, $B_2(t)\,dt$ a similar expression for the second generation, etc. $B_1(t)/N$, $B_2(t)/N, \ldots$ may be regarded as renewal density functions for the various generations.

Now, evidently,

$$B_1(t) = Nf(t) \tag{1}$$

$$B_2(t) = \int_0^t B_1(t - x)f(x)\,dx \tag{2}$$

and in general

$$B_{j+1}(t) = \int_0^t B_j(t - x)f(x)\,dx. \tag{3}$$

Summation of the contributions of the successive generations gives for the total renewal at the time $t$

$$B(t) = B_1(t) + \int_0^t B(t - x)f(x)\,dx. \tag{4}$$

In this note we propose to use a Pearson Type III function for $f(x)$ and observe what form our equations then assume. The Pearson Type III function $\frac{c^k}{\Gamma(k)} x^{k-1}e^{-cx}$, $(c > 0,\ k > 0)$, appears to be a reasonable one to use in many practical situations. The two parameters $c$ and $k$ give it a considerable amount of flexibility. The fact that this function has an unlimited range in one direction is relatively unimportant from a practical point of view, as is well known from the experience of fitting curves of this type to skewed data with limited range. Of course the question of whether a Type III curve is appropriate can be answered more objectively by using the usual Pearson curve-fitting criteria, $\beta_1$, $\beta_2$ and $k$. We have, then, substituting in (1)

$$B_1(t) = N \frac{c^k}{\Gamma(k)} t^{k-1} e^{-ct} \tag{5}$$

and from (2)

$$B_2(t) = \int_0^t N \frac{c^k}{\Gamma(k)} (t - x)^{k-1} e^{-c(t-x)} \frac{c^k}{\Gamma(k)} x^{k-1} e^{-cx}\,dx \tag{6}$$

$$= \frac{Nc^{2k}}{\Gamma(k)\Gamma(k)} e^{-ct} \int_0^t (t - x)^{k-1} x^{k-1}\,dx. \tag{7}$$

If, now, we set $x = ty$, the integral in (7) reduces to

$$\int_0^t (t - x)^{k-1} x^{k-1}\,dx = t^{2k-1} \frac{\Gamma(k)\Gamma(k)}{\Gamma(2k)}.$$

Hence,

$$B_2(t) = N \frac{c^{2k}}{\Gamma(2k)} t^{2k-1} e^{-ct} \tag{8}$$

and in general

$$B_j(t) = N \frac{c^{jk}}{\Gamma(jk)} t^{jk-1} e^{-ct}. \tag{9}$$

Summing the contributions of the several generations, we have for the total renewal function

$$B(t) = Nce^{-ct} \left\{ \frac{(ct)^{k-1}}{\Gamma(k)} + \frac{(ct)^{2k-1}}{\Gamma(2k)} + \cdots \right\}. \tag{10}$$

If $k$ is a positive integer $\geq 3$, (10) can be easily summed to a form which shows immediately its damped periodic nature. Even if $k$ is positive but not an integer, it can be shown by continuity considerations that the function $B(t)$ defined by (10) has periodic properties.

Assuming $k$ to be a positive integer, then, and setting $z = ct$, we may write the expression in brackets in (10) as

$$\frac{z^{k-1}}{(k-1)!} + \frac{z^{2k-1}}{(2k-1)!} + \cdots = f(z). \tag{11}$$

Then

$$\frac{d^k f(z)}{dz^k} = f(z)$$

and upon making the trial substitution, $f(z) = Ae^{mz}$, we get

$$Am^k e^{mz} = Ae^{mz}.$$

Hence,

$$m^k = 1.$$

Taking unity in its complex form

$$1 = \cos 2n\pi + i \sin 2n\pi$$

we have that

$$m_n = \sqrt[k]{1} = \cos \frac{2n\pi}{k} + i \sin \frac{2n\pi}{k} \tag{12}$$

where $n = 0, 1, 2, \cdots, k - 1$. Then

$$f(z) = \sum_{n=0}^{k-1} A_n e^{m_n z}$$

and

$$f^j(z) = \sum_{n=0}^{k-1} A_n m_n^j e^{m_n z}.$$

Now setting $z = 0$, we get

$$f(0) = A_0 + A_1 + \cdots + A_{k-1} = 0$$

$$f'(0) = A_0 m_0 + A_1 m_1 + \cdots + A_{k-1} m_{k-1} = 0$$

$$\vdots$$

$$f^{k-1}(0) = A_0 m_0^{k-1} + A_1 m_1^{k-1} + \cdots + A_{k-1} m_{k-1}^{k-1} = 1$$

$k$ equations to determine the $k$ constants. We know that $A_n$ is equal to the ratio of two determinants formed from the coefficients of the above equations. This ratio reduces to

$$A_n = \frac{(-1)^{k+n+1}}{(m_{k-1} - m_n)(m_{k-2} - m_n) \cdots (m_n - m_0)}. \tag{13}$$

We have, then, an expression for the $k$ constants in terms of the $k$ roots of unity. Therefore, for any particular value of $k$ we can obtain the sum of our series from the relation

$$f(z) = \sum_{n=0}^{k-1} A_n e^{m_n z}.$$

Hence, under the assumption that $k$ is a positive integer, we have

$$B(t) = Nce^{-ct} \sum_{n=0}^{k-1} A_n e^{m_n ct}. \tag{14}$$

The forms of $B(t)$ for $k = 1, 2, 3, 4$ are respectively

$$B(t) = Nc$$

$$B(t) = \tfrac{1}{2}Nc(1 - e^{-2ct})$$

$$B(t) = Nce^{-ct}\left[\tfrac{1}{3}e^{ct} - e^{-\frac{1}{2}ct}\left(\tfrac{1}{3}\cos \tfrac{1}{2}\sqrt{3}ct + \frac{1}{\sqrt{3}} \sin \tfrac{1}{2}\sqrt{3}ct\right)\right]$$

$$B(t) = Nce^{-ct}[\tfrac{1}{4}(e^{ct} - e^{-ct}) - \tfrac{1}{2}\sin ct].$$

Although the above procedure is valuable particularly because it brings to light something of the nature of our renewal function, the forms derived above can be used actually to obtain values of $B(t)$ for various values of $t$. However, for extensive numerical work a better method is at hand, which does not even depend on the assumption of an integral value for $k$.

Let us return once again to equation (10) which may be written in the following form

$$B(t) = Nc\left\{\frac{e^{-ct}(ct)^{k-1}}{\Gamma(k)} + \frac{e^{-ct}(ct)^{2k-1}}{\Gamma(2k)} + \cdots\right\}. \tag{15}$$

If $k$ and $c$ are determined by the method of moments, (using two moments), $k$ will not, in general, be a positive integer. However, by using the *Tables of the Incomplete Gamma Function* edited by Karl Pearson, one can compute values of $B(t)$ without much difficulty. In these tables the function $I(u, p)$ is tabulated for various values of $u$ and $p$, where $I(u, p)$ is defined by

$$(16) \qquad I(u, p) = \frac{\int_0^{u\sqrt{p+1}} e^{-v} v^p \, dv}{\Gamma(p + 1)}.$$

If we let $\xi = u_1\sqrt{p + 1} = u_0\sqrt{p}$ then upon integrating by parts we find

$$(17) \qquad \frac{e^{-\xi} \xi^p}{\Gamma(p + 1)} = I(u_0, p - 1) - I(u_1, p).$$

The left hand member of this equation is of the same form as each of the terms of the series in brackets in (15). Hence, the value of the renewal function for a particular time, $t$, is directly obtainable by summation of the right hand member of (17) for successive significant values of the argument $p$.

By way of illustration a numerical example will be considered. The data are taken from E. B. Kurtz' book entitled *Life Expectancy of Physical Property.* In this book the author makes a study of retirement rates of fifty-two different types of physical property, and finds that their replacement curves fall into seven distinct groups. We consider here Group VII which happens to be the largest group, embracing seventeen different types of industrial equipment out of the fifty-two examined. Using Kurtz' replacement data [2] we obtain for the value of the first and second moments

$$\mu_1 = 10.002$$

$$\mu_2 = 121.71$$

and from these by the method of moments, we find

$$k = 4.62$$

$$c = .462.$$

We then proceed to calculate values of $B(t)/N$ by means of Pearson's Tables,[3] obtaining the results shown in the following table.

---

[2] E. B. Kurtz, *Life Expectancy of Physical Property,* Ronald Press, 1930, Table 22, page 86.

[3] With regard to the method of interpolation employed in the calculations, it should be mentioned that it was found advisable to use the Mid-panel Central Difference Formula (*xxiii*) on page *xii* of the introduction to Pearson's Tables; and that it is quite sufficient for our purposes to calculate only first order terms.

| $t$ | $B(t)/N$ | $t$ | $B(t)/N$ |
|---|---|---|---|
| 0 | .0000 | 10 | .1049 |
| 1 | .0016 | 11 | .1043 |
| 2 | .0103 | 12 | .1028 |
| 3 | .0279 | 13 | .1006 |
| 4 | .0486 | 14 | .0990 |
| 5 | .0714 | 15 | .0994 |
| 6 | .0867 | 16 | .1009 |
| 7 | .0980 | 17 | .1013 |
| 8 | .1039 | 18 | .0992 |
| 9 | .1066 | 19 | .0999 |
| | | 20 | .0993 |

In conclusion the author wishes to thank Professor S. S. Wilks for various suggestions he has made in connection with this note.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

# ESTIMATES OF PARAMETERS BY MEANS OF LEAST SQUARES

BY EVAN JOHNSON, JR.

As a criterion for comparing estimates of a parameter of a universe, of known type of distribution, the use of the principle of least squares is suggested. A criterion may be stated in rather general terms. Its application to any given problem presumes a knowledge of the distribution functions of the estimates considered. In the present paper a criterion is set up and application of it is made in the estimation of the mean and of the square of standard deviation of a normal universe.

We shall use the symbol $\theta$ to represent a parameter to be estimated. It is to be remembered that $\theta$ is a constant throughout any problem, that it represents an unknown value, and that observations and functions of observations (called estimates) are the only variables that occur. We shall use the symbols $x_i$, $i = 1, 2, \cdots, n$, to represent observed values of the variable $x$ of the universe, and the symbol $F$ to represent a given function of the observations $x_i$.

If we choose to consider a given function $F$ as an estimate of $\theta$, we are then interested in the error $F - \theta$. This quantity differs from the so-called residual of least square theory, since we are here interested in the difference between computed and true values, rather than in the difference between observed and computed values. To avoid any possible confusion we shall refer to $F - \theta$ as the *error*. Over the set of all samples of $n$ observations, $x_i$, the distribution of the errors $F - \theta$ is expressed by means of the distribution function $f(F)$,