1976

An Investigation Of A Family Of Size Distributions.

Katharine Ann Kirk
University of Alabama at Birmingham

78-624

KIRK, Katharine Ann, 1943-
AN INVESTIGATION OF A FAMILY OF
SIZE DISTRIBUTIONS.

The University of Alabama,
Ph.D., 1976
Biostatistics

University Microfilms International, Ann Arbor, Michigan 48106

AN INVESTIGATION OF A FAMILY OF SIZE DISTRIBUTIONS

by

Katharine Ann Kirk

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in the
Department of Biostatistics in The Graduate School
University of Alabama in Birmingham

BIRMINGHAM, ALABAMA
1976

ACKNOWLEDGEMENTS

I wish to take this opportunity to thank the many persons whose help and encouragement throughout the two year period of dissertation writing have been indispensable.

First, thanks must go to my major professor, Dr. Malcolm E. Turner, Jr., for the inspiration of his growth models and for encouragement to use them in the development of the size distribution presented here. In addition, I wish to thank him for his careful reading of earlier drafts of this dissertation, for his general guidance and for his patience--both in checking the mathematics involved and in dealing with a sometimes discouraged graduate student.

Secondly, I must thank Dr. David C. Hurst, chairman of the Department of Biostatistics and my employer, for the critical reading of earlier drafts and the discussion (sometimes lengthy) of the main arguments in the paper. I also wish to thank him for releasing time from my employment duties for use in pursuit of this dissertation and for generous financial support in the form of computing monies. His continuing encouragement in this effort has been much appreciated.

Thirdly, I wish to give my very special thanks to my typist, Mrs. Jaquelyn Henderson, for her expert and cheerful help in executing a very difficult typing assignment.

Lastly, I wish to thank the members of my graduate committee, Dr. S.B. Barker, Dr. C.R. Katholi, Dr. E.L. Bradley, Jr. and Dr. K.M. Pruitt, for their helpful comments and criticisms on a variety of points, and the University of Alabama in Birmingham Educational Assistance Fund for partial financial support in payment of the tuition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Distributions of sizes of things have been of a great deal of interest to men for many centuries. The earliest written accounts of men's lives and cultures are composed of lists of accomplishments of rulers. These are recorded in terms of territorial measures, population counts, amounts of supplies possessed or needed, battles won and enemies confounded. According to the Inca Garcilaso de la Vega, the Inca had, supplied to him each lunar month by his subordinates, records of births and deaths occurring in the territory under their control as well as any conscripts dead in battle. He had lists not only of all inhabitants in each of his provinces, but province by province lists of all goods produced therein (Gheerbrant, 1961).

In Egypt an inventory made in 1164 B.C. under Ramses III shows that the temple to Amon at Karnak owned sixty-five towns or villages, of which seven were in Asia, 433 gardens and orchards, 700,000 areas considered to be fertile, 81,000 servants, 420,000 head of cattle and some forty-six factories of various flavors. Tithes to Amon were registered at 310,000 sacks of grain, 12,000 pounds of gold, 2.2 million pounds of silver and 5.2 million pounds of copper (Langer, 1968).

From biblical times one is told that the dimensions of the ark were three hundred by fifty by thirty cubits, that it had three decks. In Solomon's temple the Hall of the Forest of Lebanon was one hundred by fifty by thirty cubits. The Hall of Pillars was fifty by thirty cubits.

The two bronze pillars reached a height of eighteen cubits and a girth of twelve, with five cubit capitals (Jones, 1966).

Today one views the past partially through the window of size. One considers the size of the wooden horse that could contain enough men to open the gates of a city. One marvels at the precision of the measured dimensions of the Great Pyramid, at the accuracy with which it registered the astronomical dimensions and relationships known in that age. One compares the size and amount of effort involved in constructing the pyramids of Giza versus those of Yucatan. One carefully records the dimensions of the skeletal remains of the past and uses them to discuss the evolution of life forms (including that of man himself) on this earth.

Mining engineers measure sizes of particles obtained in the process of digging coal or iron or copper or gold. Power plant personnel worry about the size of particle which may be allowed in a steam turbine. Farmers are concerned with the size of an ear of corn, the amount of the harvest. Poultry breeders are involved with maximizing the size of mature chickens or, perhaps, of their eggs.

In medical and dental fields, size may record a list of normal values for a patient or it may indicate the degree of abnormality present and the prognosis for the patient given this particular measurement set. The oral surgeon, faced with cleft palate reconstructive surgery, judges hope for success by the size of the total lesions and the degree of involvement of different tissues, e.g., whether hard palate, soft palate or both. A soaring white cell count signals trouble to the physician. Depending upon its actual size, the diagnosis can vary from minor infection to major infection to infectious mononucleosis to leukemia. Size of a tumor can indicate operability and prognosis for recovery. Size of cells in a smear may indicate a pre-cancerous condition to be checked regularly. Enlargement of the heart may be followed by x-ray as a check on a certain type of coronary disease process. The amount of plaque in an artery warns of trouble ahead. A measure of cardiac output forms an overall indication of the pump's efficiency and ability to maintain life. The results of cardiac catheterization, by indicating the amount of heart tissue injured, point toward the advisability of surgery and the likelihood of survival.

There is a sense in which any thing which is measurable has a size. The range of items discussable in this manner is tremendous. It may be a length, a mass, a volume, a diameter. It may be the frequency of word usage, the number of species in genera, the number of individuals in the different species in one genus. It may be the lengths of polymer chains or the amount of products of some chemical reaction, a precipitant or a condensate. It may be the successively measured counts of particles passing through sieves with successively smaller holes so that the resulting data are gathered in the form of an empirical distribution function, not normalized. It may be the population counts for the world's largest cities, the volumes of its seas, the lengths of its river systems, the incomes of its citizens or the yields of its agricultural efforts. It may be the number of crimes of various sorts in American cities, the number of cells in a bacterial colony or the mass of a cancer.

It is not true that all of these possible examples of size distributions have arisen in the same manner. Fundamentally, things have acquired their current size as the by-product of one of several kinds of processes. They may have been produced by the destruction of something larger. Examples could include asteroids, chunks of coal mined and broken for sale, size of particles in geological strata, size of particles of fly ash. On the opposite side, they may have grown to their current form and size. In this case, growth should be interpreted not merely as biologic growth, but as a process of adding units to form the whole, whatever may be the particular type of that unit. From this point of view a ship or a building may grow as well as a single cell. There is provision in such a pure growth process to discuss the respective contributions of both anabolic and catabolic forces. A more complex formulation of the size problem considers not only pure birth (or pure death) processes but the so-called birth-death processes in which both are allowed. In this latter case, the time pattern may not be monotonically increasing or monotonically decreasing, but may pass through an extremum or even be oscillatory in nature.

In this paper are considered distributions of size which follow from an underlying growth process; however, the section on the history of the development of size distributions would be incomplete without some remarks on distributions that have more usually been derived to describe sizes arising from comminution. One might well ask whether it would not be more informative to study the growth process curves themselves to gain knowledge about that process than to discuss distributions of sizes measured cross-sectionally in time on different units. Of course, such an objection is often quite valid. Any experimental scientist, any historian or any archeologist could make the following response. There is frequently a big difference between what one would prefer to measure and the reality which is allowed by the nature of the observation desired, or by other circumstances. As an example, consider the Viking probe on Mars. It has taken samples of the geologic material around it and returned reports of the chemical composition at this time. Scientists have no way of knowing what it was like in the past. They were not there. Nor are they likely to know very far into the future. Equipment wears out and the expense of the mission will limit the number of future attempts. Besides, the time scale involved for the increments of change is rather too vast. If one prefers a biologic example, one may choose that of studies of organ sizes in embryos. Since the taking of the measurement permanently ends growth (the rat or chicken or kitten is sacrificed), only one measurement in time is possible per animal. The data must be cross-sectional. Other laboratory experiments calling for longitudinal studies may simply prove prohibitively expensive, although not technically impossible. If it is not a question of laboratory-controlled experimental measures, but of diagnostic measures taken on a patient admitted to an emergency unit, the past measures may well not have been taken and, even if they were, they have an excellent chance of not being available to those who need them.

Perhaps, for a rather different example, one is a marine biologist interested in cataloging the number of species within each genus of the order Octopoda, found in a given region of the ocean. A count estimate of the current distribution from repeated samples is possible, but such a past census is unknown. It is clear, then, that cross-sectional data are what one often has to deal with in practice. This does not mean that one has no interest in the growth process. It does mean that the data available often will not support direct investigation of the "appropriate" growth curves.

The past scientific literature is full of such data, cross-sectional data. What was often done was to assume away the individual differences in growth function and to average the information from several animals at the same time point. These mean values at different time points with different animals at each point and from one to the next such point were treated as observations on one and the same growth curve. Then the chosen growth function's path through these points was estimated and compared for different choices of the growth function. Many papers were written in this manner with lengthy discussions of how to get the "best" fit and of the relative merits of the models compared. The other type of data, in which there were no time measurements at all, even on different units, was given as a descriptive account or case study and left largely unanalyzed. The exceptions to this would be contained in such early work as Yule's 1924 paper on genus and species distributions or, perhaps, Fisher's 1943 work on abundance of species. About these two important papers, more is said in the second chapter.

The conclusion to this line of thought is as follows. If one were able to construct a system of size distributions at a point in time, which would still contain information about the form of the underlying growth process without requiring longitudinal data to be collected, one might have a very interesting tool indeed. If this system were somehow nested so that the usual growth equations yielded different members of the system according to some hierarchical form, it would be possible to test statistically whether any given function is adequate or even necessary to describe a given data set. Another problem with past attempts is the lack of variability allowed. Each individual organism is assumed to march in step along the growth curve, beginning and ending in the same places. Of course, such a scheme is unrealistic. At the very least, members of such a sample are of different ages, at different points on their individual time scales. If the growth curve were to be assumed in a deterministic fashion and some probability distribution were to be assigned to the distribution of ages in the population under discussion, one would be able, through a simple transformation process, to derive a new distribution of the sizes for a population heterogeneous with respect to age. Such a procedure is actually mentioned as a very rough approximating technique in the derivation of Yule's distribution (Yule, 1924). The parameters of the new distribution of sizes would be related to the parameters of the original growth curve and to those of the age distribution. They are not all necessarily preserved in the transformed distribution in an estimable form, although some may be.

The approach described above is the one used in this dissertation. It is still an oversimplification. A great may factors cause sizes to vary. Growth itself, as defined biologically, is the change of size with time. For young organisms it is usually a monotonically increasing (or, at least, nondecreasing) function. For organisms of adult size, however, changes are often nonmonotonic. The type of growth function differs according to inherent properties of the class of items described thereby. Secondly, even given the "proper" growth function for a type of organism, one may consider its parameters to vary in a multivariate manner according to probabilistic laws governing genetic changes. Thirdly, organisms may be examined at different stages of growth because their ages are different. Then a probability distribution for the ages of the organisms in a colony, for instance, is a relevant factor. A fourth possibility is the existence of a mixture, mixed in either random or fixed proportions, of different growth functions in a population. This could occur in a population in which different species were intermingled. It is also possible that the growth functions are not independent from one organism to another in the population. Competition for a fixed food supply or a limited living space could cause this type of effect. Then, of course, there is the probability that the measurement obtained of a size is not the same as the actual size. Depending upon the variety of measurement and the general situation, such an observation may be an unbiased estimate of the size or it may be, rather, a function of the measurement process. A last consideration is the existence of a birth-death process in which members of the population are selectively added or deleted from the population due to any of a large number of factors, some of which may be size dependent.

So, one sees that many extensions are possible. There is every reason to believe that the parameters of a growth curve are not identical from one organism to another. Genetics, environment and other kinds of random variability all play a part. One could also place a prior probability distribution upon any or all of the parameters of the growth function. Then there are many alternative choices for an "appropriate" age distribution. One can introduce more complications by assigning functional forms to the parameters or restrictions among them--hopefully, for sound theoretical reasons.

In this paper, discussion is limited to a family of size distributions based upon a family of monotonically nondecreasing growth curves and a single distribution of ages in the population. The flexibility of the system of curves derived by the method adopted in this paper is determined partly by the nature of the distribution of ages chosen and partly by the shapes allowed in the system of growth curves chosen. The ability of the derived family of size distributions to describe real data is increased according to the degree of flexibility it retains. It is reasonable that to begin such an investigation one would do well to examine simple hypotheses for the functions considered as age distributions. These simplified assumptions will undoubtedly limit the shape of the resulting curve.

Because of the interest in obtaining as much information on an underlying growth process as possible, it was decided to permit the use of the full range of usual growth functions as incorporated in the family derived by Turner et al. (1976). It was chosen to place the simpler assumptions instead on the distribution of ages, which was considered to be negative exponential in form. This would be characteristic of a population having many young members and few older ones. Shapes produced easily by the resulting size distributions include J shapes, twisted J shapes, U shapes and, more generally, distributions with a definite left-skewness and a very abrupt descent toward the axis on the right. Only one of the limiting cases, which is a form of the Weibull, will really allow much right-skewness. This limitation in form is a direct result of the exponential choice for a distribution of ages. The next step in increasing flexibility would be to choose a Weibull age distribution. This choice would retain a closed form for the resulting cumulative distribution function. The number of parameters to be estimated would be increased for each comparable member in the new system (the corresponding members derived from the corresponding growth curves). Beyond that, perhaps a generalized Gamma distribution should be considered. In an unpublished manuscript, Professor Turner has shown that such a choice yields some of the old chestnuts used as size distributions, e.g., the lognormal. The price for this expanded capability, however, is the confounding of more of the original parameters and the sacrifice of a closed form for the cumulative distribution functions.

In this dissertation, the problem discussed has been limited to distributions of size derived from a negative exponential distribution of ages (or, from another point of view, to a constant acquisition rate for acquiring the size information on a particular item to one's records, whatever its actual age) and to the family of growth curves published by Turner et al. In the second chapter, there is presented a brief review of the historical work on the development of distributions of sizes. This includes both the empirical choices with their rationales and those derived from theoretical modelling considerations. The third chapter contains a brief review of the history of growth curves in the literature and includes sketches of the derivations used for the most popular ones. The fourth chapter centers upon the generic family of growth curves derived by Turner et al. (1976). This is the system of curves upon which our corresponding system of size distributions depends. The whole chapter is devoted to the discussion of the relative shapes and the relative merits of the individual members of this growth curve group. The fifth chapter begins the derivation and discussion of the size distributions themselves. This involves a detailed discussion of the shapes of the curves, followed by a selection of illustrative graphs. The sixth chapter is devoted to two characterizations of the distributions, their moments and their intensity functions. The moments for each of these distributions are displayed and a few words are said about their usefulness--or lack of it. Their respective intensity functions are presented along with a graph and comparisons of the intensity functions to other well-known curves. The seventh chapter is reserved for a discussion of possible estimation techniques: maximum likelihood, weighted and unweighted least squares using the empirical cumulative distribution function and "quick and dirty" methods best used to get starting values for more sophisticated procedures. The eighth chapter contains a selection of examples--some from literature, some from practice--to which some of the derived size distributions have been fitted. The ninth and final chapter presents a brief review of the work contained herein and closes with a thought for future research.

II. THE HISTORY OF SIZE DISTRIBUTIONS

Most of the distributions used as size distributions were chosen because they were well-known and thoroughly studied distributions which, although derived for some other application, possessed a shape that was reminiscent of that generated by data. Such distributions used empirically would include the Gaussian, the Poisson, the Gamma, the log-normal, the harmonic, the so-called Phi and the Rosin-Rammler curve or, as it is usually called, the Weibull. See G. Herdan (1953).

The Gaussian or normal distribution is thought to occur rarely in natural size distributions. It is usually considered in the guise of its role as the limiting distribution of a sum of Bernouilli variates, representing the accumulated effect of a sequence of small changes each of which contributes the same amount of change toward an end product. Such an example of an effect produced by a series of elementary, small causes combining at random would be the sizes in some particulate substance produced by a chemical reaction, a condensate or a precipitant.

The lognormal would be appropriate if one considers that ratios rather than the net differences of equal amounts from a mean value are relevant. This distribution is thought to be useful in describing sizes produced by comminution. Particularly indicating its suitableness in a given data set would be a large ratio between the extreme observations in the sample, showing the large range of observable values.

If, instead, one measures the amount of a substance present in some medium by counts of the numbers of particles found in small observational fields, so small that the probability of finding a particle in the field is very slight, one derives from the binomial distribution the distribution of rare events, the Poisson. See Herdan (1953).

Other possibilities would include the harmonic distribution used by Keynes in 1910 and 1911 in the form:

$$f(x) = c \cdot \exp[-h^2(\mu-x)^2/x]$$

where μ is the population mean and h measures the precision of the distribution of $1/x$. Kapteyn in 1903 developed a generalization of the normal law which would make it invariant for all powers of the variable. Thus diameters, surface areas, weights or volumes would all have the same distributional form. Krumbein preferred the transformation $\phi = -\log_2 x$ which was useful in the sieving analysis of sediments or powders in which the ratio of successive mesh sizes is intended to be constant. See Herdan (1953).

One rather interesting development is contained in a series of papers from 1927 to 1933 by Rosin and Rammler. Their new distribution, later called the Weibull, was proposed for the distribution of broken coal, but was later used, by them, for cement, ores of many varieties, dye-stuffs and glass. It had two advantages over the other distributions used in mining at that time: it allowed for more skewness than did the lognormal and its cumulative distribution function existed in closed form (which made it easily applicable to data from sieving experiments). This latter property also made simple graphical methods of parameter estimation feasible. See Herdan (1953).

The Weibull distribution was rediscovered about 1939 by Waloddi

Weibull as an appropriate distribution for the size effect upon failures in solids. This work was published with seven examples in 1951. The idea was as follows. If X is a random variable indicating an attribute possessed by members of a population, the probability that X is less than or equal to a given value x is denoted by

$$P(X \leq x) = F(x).$$

This F(x) may always be written as $1 - P(X > x)$. If the probability that X is greater than x is written as $e^{-\Psi(x)}$, then it is true that any distribution function may be written as

$$F(x) = 1 - e^{-\Psi(x)}.$$

Next he considered a chain of n links, in which a failure in any one link causes a failure in the whole chain. So the probability of non-failure is equal to the simultaneous non-failure of all links. If the probability of non-failure of each link is $e^{-\Psi(x)}$, the probability of non-failure of the chain is $e^{-n\Psi(x)}$. Then the probability of failure of the chain is of the form $1 - e^{-n\Psi(x)}$. The requirement for $\Psi(x)$ is that it be a positive, non-decreasing function of X vanishing at some point x_μ, not necessarily equal to zero. A power function would be the simplest such function. The resulting function has the form

$$F(x) = 1 - e^{\frac{-(x-x_\mu)^m}{x_0}},$$

where x_0 is a dimensionless parameter. It is interesting to note that four of Weibull's seven examples are, in fact, size distributions: fly ash, length of Cyrtoideae, statures for adult males born in the British Isles, and breadth of beans of Phaseolus Vulgaris. His choice of form for $\Psi(x)$, however, was motivated solely by mathematical simplicity and bore no theoretical relationship, as far as he was concerned, to any underlying growth or deterioration process.

Gittus (1967) rewrote the Weibull distribution in its "standard form" as

$$F(x) = 1-\exp\{-[(x-x_{\mu})/x_0]^{\alpha}\}$$

and pointed out that it is a particular solution of the following differential equation:

$$d\{P(x)\}/dx = kx^{m}\{1-P(x)\}^{n}$$

If n is equal to 1 and x_{μ} is interpreted as the origin of x,

$$P(x) = 1-\exp(-kx^{m+1}/(m+1))$$

which is a form of the Weibull. If, however, n is greater than 1, a more general class of distributions is obtained:

$$P(x) = 1-\{k(n-1)/(m+1)\ x^{m+1}\ +1\}^{1/(1-n)}$$

The restrictions on his parameters are k, positive; m, greater than -1; and n greater than or equal to 1. His approach was, therefore, also purely empirical, intended only to extend the class of Weibull-type distributions. He proposed applications to data from physical processes involving nucleation and growth or relaxation phenomena.

For a new view in the variety of empirical approaches, one might mention the paper of Keiding, Jensen and Ranek (1972) on the size distribution of liver cell nuclei from the observed distribution in a plane section. They defined g(x) to be the density of radii of sphere sections on the surface, f(r), to be the density of sphere radii and m, to be the expected radius. By t, they meant the thickness of the section. Then they assumed

$$g(x) = (2m+t)^{-1}\{\int_x^\infty [2x/(r^2-x^2)^{\frac{1}{2}}]f(r)dr+tf(x)\}.$$

As the thickness approaches zero, this distribution becomes

$$g(x) = m^{-1}\int_x^\infty [x/(r^2-x^2)^{\frac{1}{2}}]f(r)dr.$$

Then for empirical reasons, they assumed that the distribution of nucleus radii for any given cell type in the liver may be approximated by a χ distribution. Since the liver is composed of three different cell types--diploid, tetraploid and octaploid, they saw the distribution of nucleus radii in the liver as a whole as a mixture of three χ distributions with possibly different parameters.

Lastly, among the empirical models, one may mention the work of Vilfredo Pareto (1897) on the distribution of income over a population:

$$N = Ax^{-a}$$

where A and a are parameters of the system and N is the number of persons whose income is greater than some specified value x. Pareto noticed that the logarithm of the number of persons N was linearly related to the logarithm of the size of income x for various sets of data. His law of income distribution follows immediately. The so-called Willis distribution, a discrete counterpart of the Pareto has been used more in species abundance data. This distribution has the form

$$P(x) = [\zeta(\beta+1)]^{-1}y^{-(\beta+1)};\ y = 1,2,3,\ldots$$

for positive β. The zeta function is well tabulated. It is defined as follows:

$$\zeta(s) = \sum_{i=1}^{\infty} i^{-s}$$

The Pareto distribution was later obtained in much the same way for application to population sizes of cities (Zipf, 1949). The rank-size relation thus produced is often called Zipf's law. It was thought by Zipf to have been caused by some vague balancing between the forces of diversification and those of unification, but no more complicated theoretical basis was proposed.

Of these various distributions usually chosen for empirical reasons, one, the lognormal, was developed for the purpose theoretically by Kolmogorov (1941) and Cramér (1946). Kolmogorov followed what would today be called a stochastic method of proof. He defined N(t) to be the total number of pieces of whatever size at time t. Then N(r,t) is the number of those pieces having size less than or equal to r. Q(k) is the number of pieces expected to arise from one piece that was size r at t in the next unit of time from t to t+1. The he set

$$A = [Q(1)]^{-1} \int_0^1 \log k \; dQ(k)$$

and

$$B = [Q(1)]^{-1} \int_0^1 (\log k - A)^2 \; dQ(k)$$

Let P_n be the probability of getting n pieces from one in one time unit of change. If $k_i = r_i/r$ is the ratio of the size of the ith new piece to that of the old piece, then the joint distribution of the n proportions can be written as

$$F_n(a_1, a_2, \ldots, a_n) = P(k_1 \le a_1, k_2 \le a_2, \ldots, k_n \le a_n).$$

If one assumes these pieces are listed in rank order according to relative size, one has

$$Q(k) = \sum_{n=1}^{\infty} P_n \{F_n(k,1,\ldots,1) + F_n(k,k,1,\ldots,1) + F_n(k,k,\ldots,k)\}.$$

Kolmogorov next made three following simple assumptions: that P_n, F_n and the fate of the other pieces are all independent of their previous histories, that Q(1) is finite and greater than 1, and that the integral $\int_0^1 |\log k|^3 \, dQ(k)$ is finite. He added that at t equal to 0 a fixed number, N(0), of pieces are present, distributed according to N(r,0). From these assumptions he deduced that the expectation of the number of pieces at time t is constant: $\overline{N(t)} = N(0)Q^t(1)$, and that the proportion k for sufficiently large t approaches arbitrarily closely

$$T(x,t) = \overline{N}(e^x,t)/\overline{N}(t) = \overline{N}(e^x,t)/(N(0)Q^t(1)).$$

Now the assumptions assure that

$$\overline{N}(r,t+1) = \int_0^1 \overline{N}(r/k,t)dQ(k).$$

Setting Q(k) = Q(1) S(log k), one obtains by substitution that

$$T(x,t+1) = \int_{-\infty}^{0} T(x-\xi,t)dS(\xi)$$

where x corresponds to log r and ξ, to log k. Then according to Liapounoff's theorem for t approaching infinity, it is true that

$$T(x,t) \to (2\pi t)^{-\frac{1}{2}}B^{-1}\int_{-\infty}^{x} \exp\{-(\xi-At)^2/(2B^2t)\}d\xi$$

where $A = \int_{-\infty}^{0} x dS(x)$ and $B^2 = \int_{-\infty}^{0} (x-A)^2 dS(x)$.

Cramér's derivation was similar in spirit, but a bit easier to follow. He assumed that there are n impulses, call them $\xi_1,\xi_2,\ldots,\xi_n$ acting sequentially. The size produced by the first γ such impulses, he named x_γ. Then the next size depends upon the size already reached plus some function of that size times the next impulse:

$$x_{\gamma+1} = x_\gamma + \xi_{\gamma+1} g(x_\gamma).$$

Then for the sum of the n impulses, one has

$$\sum_{i=1}^{n} \xi_i = \sum_{i=0}^{n-1} [(x_{i+1}-x_i)/g(x_i)] \doteq \int_{x_0}^{x_n} [g(t)]^{-1} \, dt.$$

Next assume that the effect of each impulse is directly proportional to its current size, i.e., $g(t) = t$. Then it follows that

$$\sum_{i=1}^{n} \xi_i = \log x_n - \log x_0.$$

So, by the central limit theorem, log x is normally distributed.

The next part of the history of size distributions must concern those which were derived as theoretical models for an underlying growth process. Most of these models were developed in a slightly different context from that of growth of an organism, that of species abundance theory or that of word frequencies in a published work. The earliest and, judging by the frequency with which others cite it, the most important is Yule's 1924 paper. In it, he presented a mathematical theory of species evolution, the increase in the number of species within genera with time. Yule began with N prime species of different genera. If p is the probability of a new species occurring in a time interval and $q = 1-p$, then after one time interval there would be Nq monotypic genera and Np ditypic genera. In the next time interval either set could produce new species by mutation. At the end of the second time period, there would be Nq^2 monotypic genera, $Npq(1+q)$ ditypic genera, $N(2p^2q)$ genera with three species and Np^3 genera with four species. This pattern would continue in the same manner for each successive time interval. Now if the size of the time interval, Δt, is allowed to decrease toward zero and the number of such intervals, n, is to increase in such a way that $n\Delta t = t$ is finite, it is true that

$$p = s\Delta t \text{ and } pn = st.$$

It follows then that one has

$$\lim_{\substack{n\to\infty \\ n\Delta t\to t \\ \Delta t\to 0}} q^n = \lim_{\substack{n\to\infty \\ n\Delta t\to t \\ \Delta t\to 0}} (1-p)^n = \lim_{\substack{n\to\infty \\ n\Delta t\to t \\ \Delta t\to 0}} (1-st/n)^n = e^{-st}.$$

So, if f_1 is the probability of monotypic genera at time t, one has

$$f_1 = e^{-st}.$$

Likewise, define f_2 as the probability of ditypic genera at time t. Then for f_2, one has

$$f_2 = e^{-st}(1-e^{-st}).$$

Through the recursive relationship one obtains finally that

$$f_n = e^{-st}(1-e^{-st})^{n-1}.$$

Since this is a simple geometric sequence, it is easily shown that the mean for this distribution is ${}_tM_s = e^{st}$. By exactly the same argument, Yule developed the distribution of the number of genera belonging to distinct families at time t:

$$f_1 = e^{-gt}$$

$$f_2 = e^{-gt}(1-e^{-gt})$$

$$\vdots$$

$$f_n = e^{-gt}(1-e^{-gt})^{n-1}$$

The mean number of genera in a family at time t is ${}_tM_g = e^{gt}$. The next question would be to find the distribution of those aged x at time τ. The total number of genera at time τ is $Ne^{g\tau}$. The number of new genera appearing in a time interval is $Nge^{-gt}dt$. So the number of age x at time τ can be written as $Nge^{g(\tau-x)}dx$ or $(Ne^{g\tau})(ge^{-gx}dx)$ which implies the distribution of ages at time τ is $ge^{-gx}dx$. So, if one takes the distribution of species all at a given age and compounds that with the distribution of ages, the resulting Yule distribution will be the distribution of species of all ages. The general term of the distribution so obtained is

$$_{\infty}f_n = [\Gamma(1+\rho^{-1})\Gamma(n)]/[\rho\Gamma(n+1-\rho^{-1})] \text{ where } \rho = s/g.$$

This approach of Yule's was for an infinite time scale. For that reason he had been able to ignore in the limit the number of prime genera as compared with the number of derived genera. He then considered the case of finite time and introduced the additive corrections for this case.

Fisher (1943) tackled the problem of repeated sampling (many traps) and the number of species represented by one, two, three, and so forth individuals. He assumed that for any one particular species, the number of individuals caught in a trap would follow a Poisson distribution (because of the relatively small probability of being caught at all) with the parameter varying from one species to the next. Then he assumed a Gamma distribution for the Poisson parameter, partly because its argument is positive and partly because it is the natural conjugate prior for the Poisson. The probability of observing n individuals found in this manner is

$$[(k+n-1)!\ p^n]/[(k-1)!\ n!\ (1+p)^{k+n}]$$

where k and p are the parameters of the Gamma distribution. Now in practice, n equal to zero cannot be observed since there is no information in the traps on the number of species not present. Fisher then proceeded to construct the zero-truncated form of this negative binomial distribution:

$$f(n) = \{(k+n-1)!p^n\}/\{(k-1)!n!(1+p)^n[(1+p)^k-1]\}$$

for $n = 1,2,\ldots$. Next taking the limit of the zero-truncated negative binomial as k tends toward zero and the expression log (1+p) is replaced by the new constant α, he derived the logarithmic series distribution:

$$\lim_{\substack{k \to 0 \\ \log(1+p)=\alpha}} \{(k+n-1)!k\}/\{k!n![(1+p)^k-1]\}[p/(p+1)]^n = (\alpha/n)[p/(p+1)]^n$$

If $p/(p+1)$ is replaced by x, the expected number of species with n individuals is α/nx^n. Summing the series, he found the total number of species expected as $-\alpha \log(1-x)$. Then multiplying the original term by n and summing that series, he found the total number of individuals expected as $\alpha x/(1-x)$.

D. G. Kendall (1948) undertook an exploration of possible modes of population growth which might lead to Fisher's logarithmic series distribution. He mentioned first a simple sampling approach and then presented a more complicated approach in terms of discontinuous Markov processes. For the initial motivation, he noticed that in the limit as k tends to zero, Fisher's distribution of intrinsic abundance (his Gamma prior) becomes a distribution of the form $Aw^{-1}e^{-w/a}dw$ with w chosen so that the integral of total probability converges. It is this form, a continuous analogue of the logarithmic series distribution, that puzzled him. With his first method, he showed that if the number, z, of species in the population is finite, this prior may be replaced by the logarithmic series distribution and Fisher's results re-obtained. Then v is the true number of individuals in a particular species in the population, it is distributed as $X^v/(vY)$ where $v = 1,2,\ldots$. Y is equal to $-\log(1-X)$. Next he let the probability of catching any individual in time t, p, be equal to $1-e^{-\gamma t}$. Using generating functions, he showed that the probability of a species not being represented in the sample, P_0, is equal to $1-y/Y$ where $y = -\log(1-x)$. Now the probability any species will have n individuals in the sample for n greater than zero may be written as

$$P_n = (1-P_0)x^n/(ny^n).$$

The expected number of species in the catch is $\overline{S} = X(1-P_0) = yZ/Y$ and likewise, the expected number of individuals is $\overline{N} = (e^Y-1)Z/Y$. Taking the limit as time, t, increases without bound, one sees that $\overline{S}$ approaches Z, and $\overline{N}$ approaches $(e^Y-1)Z/Y$. Then for all values of t, $\overline{S}$ is equal to $\alpha\log(1+N/\alpha)$ where α is Z/Y, Fisher's diversity index.

For Kendall's stochastic approach, he considered his population's size to be the result of a birth-death process with immigration allowed. By n, he denoted the population size at time t. His three parameters included β, representing an individual's reproductive power by binary fission, μ, representing a mortality factor and k, representing immigration into the colony. The time τ at which any of these might occur is assumed to follow a negative exponential distribution appropriate for the event:

$$\beta e^{-\beta\tau}d\tau,\ \mu e^{-\mu\tau}d\tau \text{ or } ke^{-k\tau}d\tau.$$

So, if n is the current state of the population, the transition probability from n to n+1 is $(n\beta+k)dt$. A move from n to n-1 occurs with probability $n\mu dt$ and n may remain n with probability $1-(n\mu+n\beta+k)dt$. For n greater than 0, the differential difference equations of the process are

$$d/dt\ P_n(t) = (n+1)_\mu P_{n+1}(t)-\{n(\beta+\mu)+k\}P_n(t)+\{(n-1)\beta+k\}P_{n-1}(t)$$

and for n = 0,

$$d/dt\ P_0(t) = P_1(t)-kP_0(t).$$

Using generating functions, Kendall obtained solutions which yield forms of Fisher's logarithmic series distribution when either the k/β ratio becomes negligibly small or in the absence of immigration ($k = 0$).

Mandelbrot (1970) provided a "thermostatistical" theory for taxonomic systems with Willis structure. By this structure, he referred to the property often employed in constructing taxonomic trees that if the number of lowest level items classified follows a designated distribution, the sum of such lowest level items is also distributed in the same form and the items at the next level up likewise. In particular, he commented that the Willis or discrete Pareto was a good approximation to such a stable distribution under certain conditions. If the sum of a large number, I, of Willis variables is divided by $I^{1/\alpha}$, its distribution as I increases will approach that of a Cauchy-Paul Levy random variable. This stable distribution is not available in closed form, although its characteristic function is. (Unless $\alpha = \frac{1}{2}$, in which case the distribution function has closed form also.) So if one had I independent random variables each following a Willis distribution with parameter a_i, then the total number of such categories would follow a Willis distribution with

$$a = \left(\sum_{i=1}^{I} a_i^{\alpha}\right)^{1/\alpha}$$

with exceptions for small values of the random variable increasing with I. If the a_i are all equal to each other, then $a = a_i I^{1/\alpha}$. The median number of species increases in the same manner as $I^{1/\alpha}$. Mandelbrot actually preferred to use the approximating distribution which he called the Modified Willis. It has the advantage of having both its distribution function and its generating function in closed form. The Modified

Willis distribution is obtained from the zero-truncated negative binomial by adding the zeros back in and taking the limit as p tends to zero, q tends to one and β replaces -k:

$$P(0) = \lim_{\substack{p\to 0\\ k=-\beta}} 1+(p\gamma)^{-k}(1-p^{k}) = 1-\gamma^{\beta}$$

$$P(x) = \lim_{\substack{p\to 0\\ k\to -\beta}} \gamma^{-k}(1-p^{k})^{-1}[\Gamma(x+k)]/[x!\Gamma(k)]p^{k}q^{x}$$

$$= -\gamma^{\beta}[\Gamma(x-\beta)]/[\Gamma(-\beta)x!].$$

The generating function can be written as $G(s) = 1-\gamma^{\beta}(1-s)^{\beta}$. This theory allowed for three situations found often in the field: (1) the case of very few common species and a large number of local species, (2) the reverse case of many common species few of which are represented locally, and (3) the case of a median number of common species and, for two adjacent areas x and y, many species found in x and not in y with few species found in y and not in x.

Good (1953) was concerned about the estimation of a population frequency, q_r, for an arbitrary species which occurs r times in a sample size N. His n_r denoted the number of species drawn r times in the sample. About n_0, the number of species not represented in the sample, little is known. He proposed methods for estimating q_r, the proportion of the species in the population actually occurring in the sample and some of the measures of heterogeneity sometimes employed in discussing a population. If p_μ is the true population frequency of the μth species where μ is less than or equal to s, the total number of species in the population, then observing it r times in a sample of N has a likelihood of $\binom{N}{r} p_\mu^{r}(1-p_\mu)^{N-r}$. If the final probabilities are proportional to the initial ones times the likelihoods, the probability that a species represented r times is the μth one is

$$P(q_r=p_\mu|H) = [p_\mu^r(1-p_\mu)^{N-r}]/[\sum_{\mu=1}^{s} p_\mu^r(1-p_\mu)^{N-r}].$$

Then the expected value for q_r^m for any m, a positive integer, is

$$E(q_r^m|H) = [\sum_{\mu=1}^{s} p_\mu^{r+m}(1-p_\mu)^{N-r}]/[\sum_{\mu=1}^{s} p_\mu^r(1-p_\mu)^{N-r}],$$

which may be rewritten as

$$E(q_r^m|H) = (r+m)^{(m)}/(N+m)^{(m)}[E_{N+m}(n_{r+m}|H)]/[E_N(n_r|H)].$$

His result was that the expected value of q_r is approximately equal to $(r+1)n_{r+1}/(n_rN)$ although he would prefer to substitute for n_{r+1} and n_r, n'_{r+1} and n'_r respectively where the primed numbers represent sample frequencies smoothed by some "locally appropriate" method. Because of this approximation, he was able to show that for the successive moments, it is approximately true that

$$E(q_r^m) \doteq E(q_r)E(q_{r+1})\ldots E(q_{r+m-1}),$$

so that the variance may be written as

$$V(q_r) \doteq E(q_r)[E(q_{r+1})-E(q_r)].$$

Accordingly, he found that $(r+1)n_{r+1}/N$ is the approximate value of the expected total chance of all species being represented r times in the sample. Then the expected total chance of all species represented in the sample is roughly $1-n_1/N$, which implies the next animal sampled will come from a new species with rough probability n_1/N.

Simon (1955) developed a class of skewed distribution functions having three properties commonly found in observed distributions of sizes and word frequencies: (1) they have very long upper tails with respect to their lower tails and may in fact be J-shaped with the tails well approximated by $f(i) = (a/i^k)b^i$ where a, b and k are constants, (2) k is usually greater than one and often close to two and (3) in many cases the approximation of this above mentioned function is good for small values of i as well as in the tail areas. His distribution, which is a modification of the Yule distribution, is derived in terms of a stochastic process. His derivation is formulated as a discussion of a written work of length k words, in which the number of different words, each occurring i times is indicated by f(i,k). He made two initial assumptions: the probability that the next word, the (k+1)-st, has already appeared i times is proportional to if(i,k) and that the probability that the next word has not been used before is equal to a constant, α. This may be given in difference equation form as

$$E[f(i,k+1)]-f(i,k) = K(k)\{(i-1)f(i-1,k)-if(i,k)\}$$

where i is between 2 and k+1 or as

$$E\{f(1,k+1)\}-f(1,k) = \alpha-K(k)f(1,k)$$

for i equal to 1. Several relationships are implied:

1) $\sum_{i=1}^{k} K(k)if(i,k) - K(k)\sum_{i=1}^{k} if(i,k) = 1-\alpha,$

2) $\sum_{i=1}^{k} if(i,k) = k$ and

3) $K(k) = (1-\alpha)/k.$

Then he assumed that all the frequencies increase at a rate proportional to k so that their relative size is maintained:

$$\{f(i,k+1)\}/\{f(i,k)\} = (k+1)/k \text{ for i less than or equal to k}$$

and f(i,k) = 0 for i greater than k.

This implies that the ratio $f(i,k)/f(i-1,k)$ is equal to some function $\beta(i)$ which does not contain k. If $f^*(i)$ is used to refer to the relative frequencies, then $\beta(i)$ can be solved for as $f^*(i)/f^*(i-1)$ for $i = 2,3,\ldots k$. Substituting $\rho = 1/(1-\alpha)$, he obtained that

$$f^*(i) = \beta(i,\rho+1)f^*(1) \text{ for } i = 2,3,\ldots,k.$$

where $\beta(i,\rho+1)$ is the Beta function for i and $\rho+1$. Now the total number of different words, n_k, is αk. Recalling that k is the total number of words in the work, he found that $f^*(1)$ is equal to $k\alpha/(2-\alpha)$ or $n_k/(2-\alpha)$. This final distribution is often re-parameterized as

$$P(i) = A(k)\beta(i,\rho+1)$$

where $1 \le i < k$ and $A(k) = k\alpha/(2-\alpha)$.

In reviewing a number of the more important papers on size distributions, the aim has been to indicate a variety of the approaches taken and interest pursued rather than to catalogue every existing paper in the field. There are many other contributions by these and other authors whose work deserves an attentive reading. For a last look at the possibilities already tackled, one may mention one in which no analytical solution was achieved. The work of Saidel (1968) is interesting because of the scope of the factors he incorporated in his model. Consider a bacterial colony. Let $f(x,t)dx$ be the number of cells in the size interval from x to x+dx at time t and x_m be the smallest cell size. Then let G, B and D be rates of growth, birth and death respectively, each of which is dependent on the size x and the time t. By x, the cell size, he denoted a radius in a spherical cell or an axial length if the cell is rod-like (assuming the radius to be constant in that case). Then he set up a system of assumed differential equations:

(1) $[\partial f(x,t)]/(\partial t) + \{\partial[G(x,t)f(x,t)]\}/(\partial x) = B(x,t)-D(x,t)$

where $x \geq x_m$;

(2) $G(x,t) = dx/dt = \beta(t)c_1(t)x^r$

where $\beta(t)$ is the time-dependent growth-rate coefficient,

$c_1(t)$ is the concentration of nutrient present in the solution

and r is 0 for spheres or 1 for rods;

(3) the net rate of production of cells of size x is

$B(x,t) = 2\int\alpha(y)f(y,t)\delta(x-y/2^s)dy-\alpha(x)f(x,t)$ where $\delta(x)$ is the Dirac delta function, $\alpha(x)$ is the division rate coefficient and x = 1/3 for spheres or 1 for rods;

(4) the cell death rate is $D(x,t) = \gamma(c_1,c_2)f(x,t)$ where $\gamma(c_1,c_2)$ is the death rate coefficient which is a function of the concentration of nutrient, c_1, and of the amount of toxic product produced in the solution, c_2.

Combining these equations, he reconstructed the new dynamic equation for the size distribution as

$$df/dt + \beta c_1\{\partial[x^r f]\}/\partial x = 2^{s+1}\alpha(2^s x)f(2^s x)-\alpha f-\gamma f$$

for $x \geq x_m$ and the initial condition

$$f|_{t=0} = \delta(x-x_m).$$

He assumed next that the probability of cell division increases exponentially with size, providing that the cells are large enough to divide at all, that is, are greater than $2^s x_m$: $\alpha(x) = \alpha_0 e^{Ax/x_m}u(x-2^s x_m)$ where u is the unit step function. Next he assumed that the growth rate $\beta(t)$ is a constant, β_0, and the death rate increases if either the nutrient supply drops or the amount of toxic product rises:

$$\gamma(c_1,c_2) = \gamma_1/(\gamma_{s_1}+c_1) + (\gamma_2 c_2)/(\gamma_{s_2}+c_2)$$

where γ_1, γ_2, γ_{s_1} and γ_{s_2} are constants. He postulated that the diffusion rate of nutrient to the cells is a function of the nutrient present and of the surface area of the cells and that the process is diffusion controlled:

$$dc_1/dt = -k_1 c_1 \int x^p f(x,t)dx; \; c_1|_{t=0} = c_1(0)$$

where p = 2 for spheres or 1 for rods. On the other hand, the toxic product is assumed to rise according to the volume of the cells present:

$$dc_2/dt = k_2 \int x^q f(x,t)dx; \; c_2|_{t=0} = 0$$

where q is 3 for spheres and 1 for rods. After he reparameterized and transformed his variables seeking a set of dimensionless groups of equations to characterize his rather complex system, he proceeded to the computer to solve the equations numerically for a few selected cases. The difficulty of such a system, of course, is that the amount of time and money required to solve any given example numerically is often prohibitive. The merit is that it explicitly accounts for much of the variability in a real system.

III. A BRIEF HISTORY OF DETERMINISTIC GROWTH CURVES

The size distributions derived in this paper are based on growth curves. These growth curves are obtained in a deterministic manner from relatively simple postulates. They belong to the genre of "classical" growth processes and are produced as members of a family of growth curves published by Turner et al. (1976). In order to understand the development of such a system, one may read the review in this rather brief chapter of the history of deterministic growth curves of this general type. Developments as birth-death processes and stochastic developments of growth curves are plentiful, but they are beyond the scope of this paper. They are derived in quite a different spirit. Readers who are interested in this approach are referred to the work of Feller (1966), Bartlett (1969, 1955), D. G. Kendall (1948, 1949), Goodman (1953, 1967), Keyfitz (1967) and J. H. Pollard (1966, 1969).

The earliest mathematical model for growth of a population is generally credited to Thomas Malthus in 1798. According to his theory, two things were necessary to the existence of man, food and sex. Food increases in an arithmetic ratio while population increases geometrically at a rate dependent on the current size and an unknown constant. If x represents the population size and dx/dt is the rate of growth of x with respect to t, then his differential equation is $dx/dt = \beta x$ where β is the positive growth rate constant. The solution yields the geometric growth model $x = c \cdot \exp(\beta t)$ where c is the original size of the population at time zero. This model is an unbounded, rising curve consistent with unlimited population growth (Pollard, 1973).

The model known as the logistic was first proposed by Daniel Bernoulli in 1760 as the growth of the proportion of immune persons in a smallpox epidemic (Todhunter, 1865). Verhulst, in 1838, rederived the logistic equation for population growth (Ahuja and Nash, 1967). The idea was that growth could not proceed indefinitely, but was limited according to the needs of the population for food, space and other necessities. Call this upper bound to the supportable size of the population k. Then it must be that some inhibition factor is present that slows the rate of growth down as it approaches that upper limit k. Restated, it could be said that the growth rate increases proportionally with the current size and decreases proportionally as the ratio of the current size to the maximal size approaches unity. In mathematical terms, this is the equation

$$dy/dt = \beta y(1-y/k)$$

whose solution may be written as

$$y = k[1+\alpha \cdot \exp(-\beta t)]^{-1},$$

where α is the constant of integration. The logistic curve was re-examined by Robertson (1923) from the point of view of analogy with an autocatalytic chemical reaction. He postulated a reversible reaction in which three molecules of different substances combine to produce two molecules of the first substance and one of a new substance. Call the forward rate of reaction k_1 and the backward rate k_2. This reaction could be written as a group of four differential equations:

$$\dot{y}_1 = k_1y_2y_3y_1 - k_2y_4y_1^2$$

$$\dot{y}_2 = k_2y_4y_1^2 - k_1y_2y_3y_1$$

$$\dot{y}_3 = k_2y_4y_1^2 - k_1y_2y_3y_1$$

$$\dot{y}_4 = k_1y_2y_3y_1 - k_2y_4y_1^2$$

If one assumed that y_2, y_3 and y_4 are effectively constant because the substances which they represent are present in such abundance, this would mean that the product $k_1y_2y_3$ is also approximately constant (β) and k_2y_4 is approximately constant (β/K). Then the equations would become

$$\dot{y}_1 = \beta y_1 - (\beta/K)y_1^2 = 0$$

$$\dot{y}_2 = \beta/Ky_1^2 - \beta y_1 = 0$$

$$\dot{y}_3 = \beta/Ky_1^2 - \beta y_1 = 0$$

$$\dot{y}_4 = \beta y_1 - (\beta/K)y_1^2 = 0$$

So the approximating differential equation for y_1 is $\dot{y}_1 = \beta y_1 - (\beta/K)y_1^2$, the Verhulst differential equation whose solution is

$$y = K\{1+[(k-y_{10})/y_{10}]\exp(-\beta t)\}^{-1}$$

where y_{10} is the initial concentration of the first substance. Lotka (1925) developed the same equation from the Taylor's series approximation approach. Probably its best known exponents, however, were Pearl and Reed who utilized this curve in fitting a number of sets of growth data including that of the United States population. See Pearl (1921), Pearl, Edwards and Miner (1934) and Pearl and Reed (1920). Since that time the logistic growth equation has been successfully applied to the growth of many small animals although not typically to members of a group as complex as that of humans. See Krause et al. (1967) for examples from the growth of chickens.

Another well-worn curve in the annals of growth is that due to Gompertz (1825). This curve is characterized by an asymmetry in the direction of a much faster initial growth period with a long period of approaching maximal growth very slowly, from not very far away. It is interesting to note that Gompertz did not intend to derive a growth curve when he did his work in 1825, but was, instead, seeking a function with which to graduate mortality tables. His argument went as follows. Suppose by L_x one means to indicate the number of persons who are alive at age x. Then suppose one knows that the difference in the logarithm (Gompertz used common logs) of the number of persons alive at time a and the logarithm of the number of persons alive at time a+r is a constant, m. Then a similar difference for another time period of length r will be mp and for the next such time period will be mp^2. Thus one has made the following set of difference equations:

$$\log(L_a)-\log(L_{a+r}) = m$$

$$\log(L_{a+r})-\log(L_{a+2r}) = mp$$

$$\log(L_{a+2r})-\log(L_{a+3r}) = mp^2$$

$$\vdots$$

$$\log(L_{a+n})-\log(L_{a+\overline{n-r}}) = mp^{n/r-1}.$$

These are terms of a geometric series which when summed give

$$\log(L_a)-\log(L_{a+n}) = m(1-p^{n/r})/(1-p).$$

Next substituting the results of these transformations,

$$p^{1/r} = q,\ q^r = p \text{ and } \log \varepsilon = m/(1-q^r),$$

one has

$$\log(L_a)-\log(L_{a+n}) = (1-q^n)\log \varepsilon$$

which may be rearranged as

$$\log(L_{a+n}) = \log(L_a/\varepsilon)+q^n\log \varepsilon$$

Again transforming the indices using $x = a+n$ or $n = x-a$ and $d = \log(L_a/\varepsilon)$, one finds

$$\log L_x = d+q^{x-a}\log \varepsilon.$$

Now solving for L_x one obtains the original form of the Gompertz curve as

$$L_x = d\varepsilon^{q^{x-a}} = d(\varepsilon^{q^{-a}})^{q^x}$$

$$\text{or } L_x = dg^{q^x} \text{ where } g = \varepsilon^{q^{-a}}.$$

These days, to simplify the estimation of the parameters by bounding them positive, one tends to write the growth form of the equations as $\dot{x} = \beta x\log(k/x)$ and $x = k\exp\{-c\cdot\exp(\beta t)\}$. This curve was later modified by Makeham (1890) as

$$y = c\cdot\exp(At+\beta b^t),$$

which is unbounded for positive values of A.

It was Medawar (1940) who provided an interpretation for the Gompertz curve as a growth model. His derivation introduced the concept of a growth energy of tissue which was defined as the magnitude of the threshold for total growth inhibition. Call this growth energy quantity g. Then if t is the age of the growing organism and m is its mass, his three growth assumptions may be written as

(1) $g = a \cdot \exp(-kt)$, the growth energy declines as a negative exponential with time,

(2) $\phi(m) = g/b$ where $\phi(m)$ is the specific growth rate and b is an unknown constant,

and (3) $\dot{m} = \phi(m) \cdot m$, the rate of change of mass over time is equal to the specific growth rate times the current mass.

Combining these assumptions and solving for m, he rederived the Gompertz model:

$$\dot{m}/m = g/b = (a/b) \cdot \exp(-kt)$$

$$\log m = \log c - [a/(bk)] \cdot \exp(-kt)$$

$$m = c \exp\{[-a/(bk)] \cdot \exp(-kt)\};$$

however, for this version the unknown parameters have more biologic meaning. Medawar then applied his theory to the embryonic development of chicken hearts. For an interesting example of applied Gompertz models, see Laird, Tyler and Barton (1965).

Pütter (1920) proposed a differential equation for growth based upon a theory of anabolism and catabolism. The rate of growth, he said, should be equal to the difference between the forces of anabolism and catabolism. Now if anabolism were proportional to the amount of absorbent surface area (s) present and catabolism were proportional to the weight (w) of the organism, this differential equation could be stated mathematically as $dw/dt = n's - k'w$, where n' and k' are the proportionality constants. Bertalanffy (1938, 1957) offered a particular solution to this equation by assuming that the surface area in turn was proportional to the two-thirds power of weight. This can be formulated as $s = (\eta/\eta')w^{2/3}$ with the resulting rate equation as $\dot{w} = \eta w^{2/3} - k'w$. Then the solution to this equation was found as

$$w = \{\eta/k' - [\eta/k' - w_0^{-1/3}]e^{k't/3}\}^{-3}$$

In general, Bertalanffy investigated the situation in which anabolism is proportional to the mth power of weight (w) and catabolism is proportional to the nth power of weight where m is less than or equal to n. The rate equation became $\dot{w} = \eta w^m - kw^n$ where $m \leq n$ and $\dot{w}$ is the time derivative of weight. Then he assumed n was equal to 1 and listed three metabolic types of solutions: that for $m = 2/3$, which is a surface area type; that for $m = 1$, which is a weight type; that for $2/3 < m < 1$, which is an intermediate type. No solutions were considered for $m < 2/3$ or $m > 1$ or for n not equal to 1. This particular rate equation was $\dot{w} = \eta w^m - kw$ and its solution was

$$w = \{(\eta/k) - [(\eta/k) - w_0^{1-m}]e^{-(1-m)kt}\}^{1/(1-m)},$$

where w_0 is the weight at time $t = 0$. Richards (1959) extended the results of Bertalanffy to include values of m down to zero and to all positive real numbers. The special case of the limiting form as m approaches one and the maximal size remains finite, he noted, was the Gompertz. The case of $m = 2$ was the logistic or autocatalytic curve. As m approaches zero, the monomolecular or Mitscherlich equation appeared: $\dot{w} = k(A-w)$, $w = A(1-be^{-kt})$, where kA equals the η of the Bertalanffy formula. He, as did Bertalanffy, noticed the special limit as m approaches one and there is no upper bound to growth, the exponential or Malthusian growth curve. So values of m between 0 and 1 were considered to shade between monomolecular growth and Gompertzian growth, while values between 1 and 2 were shading between the Gompertzian and the logistic curves. It is often overlooked that the so-called Bertalanffy-Richards growth curve was actually first derived by Verhulst, not only for the logistic case in 1838, but for a general power $m>1$ in 1845, which corresponds partially to the work of Richards (1959). Verhulst should also be credited with the Mitscherlich equation in 1847. See Ahuja and Nash (1967).

A number of writers have since based their work on the set of curves already mentioned. Nelder (1961a) considered the Bertalanffy-Richards case when $dw/dt = kw[1-(w/A)^{1/\theta}]$. This final equation for w was $w = A\{1+\exp[-(\lambda+kt)/\theta]^{-\theta}\}$ where θ was greater than zero. Then Nelder (1962) reparameterized his earlier model to allow for negative values of θ, which he had found necessary for the fitting of data on the growth of carrots. His new equation read

$$w = A\{1+\phi e^{-(\lambda'+k't)}\}^{1/\phi} \text{ where } \phi = 1/\theta$$

of the previous work, $\lambda' = -\ln\theta+\lambda/\theta$ and $k' = k/\theta$. In this version ϕ could be negative as well as positive. ϕ equal to 1 gave the logistic equation. ϕ equal to -1 gave a form of the Mitscherlich equation and the limit as ϕ approached zero gave the Gompertz.

Ahuja and Nash (1967) suggested a generalization for each of three basic, historical growth curves

(1) the Gompertzian: $F_1(t) = e^{-\rho[\exp(-t/\sigma)]}$ for $-\infty<t<\infty$,

(2) $F_2(t) = (1+\rho e^{-t/\sigma})^{-\theta}$ for $-\infty<t<\infty$, which is the logistic if $\theta = 1$ and the Bertalanffy-Richards in general,

and (3) $F_3(t) = (1-\rho e^{-t/\sigma})^{\theta}$ for $\sigma\log\rho<t<\infty$, the Mitscherlich equation for $\theta = 1$.

The structures of these equations are more easily seen if one makes the transformation $s = \rho[\exp(-t/\sigma)]$ or $t = \sigma(\log \rho - \log s)$ with $dt = -\sigma/s\, ds$:

(1) $g_1(s) = e^{-s}$, for $0<s<\infty$;

(2) $g_2(s) = \theta(1+s)^{-(\theta+1)}$, for $0<s<\infty$;

(3) $g_3(s) = \theta(1-s)^{\theta-1}$, for $0<s<1$.

Ahuja and Nash assumed that these distributions of s were respectively special cases of the following distributions when the new, positive parameter ψ was equal to 1:

(1) the gamma distribution:

$$g_1(s;\psi) = 1/[\Gamma(\psi)]s^{\psi-1}e^{-s} \text{ for } 0<s<\infty;$$

(2) the beta distribution of the second kind:

$$g_2(s;\psi) = 1/[\beta(\psi,\theta)]s^{\psi-1}(1+s)^{-(\psi+\theta)} \text{ for } 0<s<\infty;$$

(3) the beta distribution of the first kind:

$$g_3(s;\psi) = 1/[\beta(\psi,\theta)]s^{\psi-1}(1-s)^{\theta-1} \text{ for } 0<s<1.$$

Then the corresponding densities for t were obtained by reversing the transformation process:

$$f_1(t;\rho,\sigma,\psi) = 1/[\sigma\Gamma(\psi)](\rho e^{-t/\sigma})^{\psi}e^{-\rho e^{-t/\sigma}} \text{ for } -\infty<t<\infty,$$

$$f_2(t;\rho,\sigma,\psi) = 1/[\sigma\beta(\psi,\theta)](\rho e^{-t/\sigma})^{\psi}(1+\rho e^{-t/\sigma})^{-(\psi+\theta)} \text{ for } -\infty<t<\infty,$$

and

$$f_3(t;\rho,\sigma,\psi) = 1/[\sigma\beta(\psi,\theta)](\rho e^{-t/\sigma})^{\psi}(1-\rho e^{-t/\sigma})^{\theta-1} \text{ for } \sigma\log\rho<t<\infty.$$

One disadvantage of their general system, however, is that the cumulative distribution functions will, in most cases, lack a closed form and must be evaluated in terms of incomplete beta and gamma functions.

Will (1936) pursued a model generalizing the logistic. He made five basic assumptions: the rate of growth is finite and continuous; its magnitude is a positive real number; growth is restricted to definite bounds; growth is a function of time; the whole growth process is free of cataclysmic disturbances. He postulated three kinds of generalizations which he called types α, β and γ. If L is the lower limit of growth and H is the upper limit, these types are as follows:

Type α: $p = L+H/\{1+\exp[a+bt+s \sin(m(t+q))]\}$;

Type β: $p = L+H/\{1+\exp[a+bt+s \log(1+m^2(t+q)^2)]\}$;

Type γ: $p = L+H/\{1+\exp[a+bt+s\sqrt{1+m^2(t+q)^2}]\}$

There is a whole host of other proposed models for various growth processes. Before beginning a more lengthy discussion of two more complex proposals involving systems of differential equations, a partial list of these other investigators is presented with an indication of their approaches. Nelder (1921b) noted that simple enzyme systems mimic the growth of bacterial cultures. Three of his four curves are in the guise of mixtures of two curves. Using n as the size at time t, his equations were

(1) $n = pe^t+(1-p)(1+t)$,

(2) $n = pe^t+(1-p)$,

(3) $n = pe^t+(1-p)(2-e^{-t})$,

and (4) $n = e^t-t$.

Takahashi and Inouye (1967) discussed the characteristics of an assumed cube root growth in transplantable tumors: $N(t) = (N_0^{1/3}+st)^3$.

Williams (1967) proposed a particular form for the constants of the logistic model in which M is the mass, M_0 is the mass at time t = 0 and C_0 is the initial concentration of nutrient:

$$M = (C_0+M_0)/[1+(C_0/M_0)\exp(-k_1 t/(C_0+M_0))]$$

where k_1 is a constant. Laird (1969) provided an explanation of the Gompertz "fit" to tumor data in terms of human tumor cell kinetics in the National Cancer Institute's Monograph Number 30. Turner et al. (1969) developed a generalization of the logistic growth law in which k, the maximal size, is a function of time and approaches the ultimate maximum size K as t increases without bound. Welch (1970) simply returned to the use of orthogonal polynomials to represent growth curves.

Borg (1971) considered a quantity f, growing over time or as a function of another quantity, α. If f_1 and f_2 correspond to the quantities measured at α_1 and α_2, $\Delta\alpha$ is equal to $\alpha_2-\alpha_1$ and k represents a term containing parameters for any other factors influencing growth, this relationship is written as $f_2 = f_1+f_1K(\Delta\alpha)$. The function representing the ratio of the current value of f to the final value f_f is

$$f/f_f = \exp[\int_{\alpha}^{\alpha_f} K d\alpha] \text{ where } 0\le\alpha\le\alpha_f.$$

One possible choice for the function $K(\Delta\alpha)$ yields

$$f/f_f = \exp\{\int_{\alpha}^{\alpha_f} \sum_n k_n \alpha_f^{n-1}/\alpha^n \, d\alpha\}.$$

Borg applied this system to a panorama of different problems ranging from the "growth" of specific heat with temperature, to the weight of a human brain over time, to the "growth" of oxygenated hemoglobin as dependent upon the concentration of oxygen. Wingert (1971), as did Pearl much earlier, proposed a generalization of the logistic substituting a cubic polynomial for the usual linear exponent of the e term:

$$S(t) = p_1/[1+e^{P(T)}]$$

where $P(T) = P_2+P_3T+P_4T^2+P_5T^3$. Wenk (1973) derived two generalizations of Gompertzian growth for particular application to the volume growth of forest trees:

$$w(t) = \exp\{-c_1[1-\exp(-c_2t(1-\exp(-c_3t)))]\}$$

and

$$w(t) = \exp\{-c_1t[1-\exp(-c_2t)]\}.$$

Blum (1974) modelled the growth of cell populations in cancer. If N is the number of cells present, then N is equal to N_0, the number of initial cancer cells times 2 to the power $r\Delta t$, where r is a constant and Δt is the time change. This relationship may be rewritten in terms of powers of e:

$$N = N_0 2^{r\Delta t} = e^{f(a)} e^{(\log 2)f(r)\Delta t}.$$

Then the growth in volume, by transformation, is

$$V = \exp\{f(s)+(\log 2)f(r)\Delta t\}$$

where f(s) represents the changes in the number of clones and non-duplicating factors with time. The volume at the end point V_c was found to be

$$V_c = \exp\{1+(\log 2)(t_c-t_1)F(D)P_c\}$$

where t_c is the final time point, P_c is the end-point probability and F(D) is a function describing the interaction of the carcinogenic agent with the tissue. He assumed that the natural logarithm of V_c was equal to a constant, K. Fletcher (1974) produced what he called a quadric law of damped exponential growth using the differential equation $y'+ay^2+by+g(y')^2+hyy' = 0$ where y' is the time derivative of y.

In concluding this section on the history of "classical" growth curves, mention should be made of two versions which involve the solution of systems of differential equations, that of Parks (1970) and that of Weiss and Kavanau (1957). First, for the approach of Parks, a few definitions are needed:

W = biomass at time t,

W_0 = initial biomass,

F = cumulative amount of food consumed,

dF/dt = food intake,

A = mature biomass,

$(\ln 2)/\beta$ = feed required for a growth increment of $(A-W)/2$,

t = age,

C = mature feed intake,

D = initial feed intake,

$(\ln 2)\tau$ = time required for a feed intake increment of $(C-dF/dt)/2$.

Next he constructed two postulates:

(1) $dW/dF = AB-BW$ where $W_0 \le W \le A$

and (2) $d^2F/dt^2 = df/dt = (c/\tau)-(1/\tau)f$ where $D \le f \le C$.

Solving these two equations, he obtained, in the first instance, $W-W_0 = (A-W_0)(1-e^{-BF})$ and, in the second instance, $f-D = (C-D)(1-e^{-t/\tau})$. Integrating this second equation again, he obtained an expression for F as

$$F = Ct-(C-D)\tau(1-e^{-t/\tau}).$$

Combining this equation for F with the equation for W, he wrote his final model as

$$W-W_0 = (A-W_0)\{1-\exp[-BCt+B(C-D)\tau(1-e^{-t/\tau})]\},$$

which is a generalization of the Gompertz growth curve.

Lastly, a treatment of the theory due to Weiss and Kavanau (1957) is in order. Before beginning a summary of their theory, one may note that this is a rather long and sophisticated paper whose contents bear a careful scrutiny. For readers who would wish to pursue theories of growth in their own work, a thorough consideration of this work is highly recommended. The assumptions of the Weiss-Kavanau model can be stated in the following manner. Growth is the expression which designates the gain in organic mass. It is the net difference between the mass created and retained and mass destroyed or lost through some process such as excretion. There are two fundamental kinds of mass, generative mass, used in reproduction, and differentiated mass, which is derived from generative mass but does not have reproductive ability. The mechanism by which cells reproduce involves key compounds called templates which act as catalysts for the reproduction of particular cell types. There are also antitemplates produced by each cell and freely diffusible which can act as blocking agents to the action of the templates, thus inhibiting reproduction. A decline in the growth rate is due to an increased release of the antitemplate compounds, which acts as a negative feedback mechanism blocking more of the templates. The concept of a terminal size for growth should be thought of as the point of a stationary equilibrium between the forces for growth and those for destruction. This stationary equilibrium occurs when there is also an equilibrium between the intracellular and the extracellular concentrations of antitemplate compounds. Both types of mass along with the antitemplates are continually renewed and destroyed according to various metabolic cycles. Briefly stated, they have three main differential equations. Firstly, a change in the generative mass over time is equal to the basic rate of reproduction of generative mass over time times the generative feedback term minus the initial rate of conversion of generative mass into differentiated mass over time times the differentiation feedback term. This corresponds to the difference between generative mass formed and generative mass lost. Secondly, a change in differentiated mass over time is equal to the rate of gain of differentiated mass by conversion from generative mass over time minus the rate of loss of differentiated mass due to catabolism. This corresponds to the difference between differentiated mass produced and differentiated mass lost. Thirdly, for the feedback component, the change in the number of inhibitor molecules over time is equal to the difference between the rate of production of inhibitors and the rate of their catabolic loss.

Now, one may show this trio of equations mathematically:

$$(1)\quad dG/dt = (G \log 2)[1-\{b(G^n-G_0^n)\}/\{G_e^n-G_0^n\}]$$

$$-k_1G[1-\{G^n-G_0^n\}/\{G_e^n-G_0^n\}]-k_2G,$$

$$(2)\quad dD/dt = k_1G[1-\{G^n-G_0^n\}/\{G_e^n-G_0^n\}]+k_2G-k_3D,$$

$$(3)\quad dI/dt = k_4G-k_5I \quad \text{or} \quad dI/dt = k_4D-k_5I$$

$$\text{or } dI/dt = k_4dG_g-k_5I \quad \text{or} \quad dI/dt = k_4dG_d-k_5I.$$

Definitions of terms for these equations are as follows:

D = differentiated mass of organ,

e = subscript referring to terminal size,

G = generative mass of organ,

dG_d = change in generative mass as the result of conversion from G to D,

dG_g = change in generative mass via reproduction,

I = number of inhibitor molecules,

k_1 = rate constant for the accretion component of differentiation,

k_2 = rate constant for the maintenance component of differentiation,

k_3 = rate constant for the catabolic loss of D,

k_4 = rate constant for the formation of inhibitor molecules,

k_5 = rate constant for the catabolic loss of inhibitors,

0 = subscript referring to initial size

n = an unknown power parameter,

t = physiological time, the doubling time for a unit of G in pure growth.

The solutions to these equations are

(1) $G = G_e/(1+ae^{-n\alpha\gamma t})^{1/n}$ where $a = (G_e^n - G_0^n)/G_0^n$, $\gamma = (a+1)/a$ and $\alpha = \log 2 - k_1 - k_2$, which is of the Bertalanffy-Richards type of equation,

(2) $D = D_e/e^{k_3 t}$ ·

$$\{e^{n_1 z}[1+n_1(y-1)s_1-k_1 n_1 \gamma s_2/k_2]$$

$$-(a+1)^n[1+n_1(y-1)s_{1a}-k_1 n_1 \gamma s_{2a}/k_2]\}$$

where $\rho = a/(a+e^{n\alpha\gamma t})$, $n_1 = k_3/(n\alpha\gamma)$, $y = 1/n+n_1$, $e^z = a/\rho = (a+e^{n\alpha\gamma t})$ and s_1, s_{1a}, S_2, and S_{2a} represent infinite series. The feedback component solution, however, requires a more complicated discussion which will not be undertaken here.

IV. A GROWTH CURVE BASIS LEADING TO DISTRIBUTIONS OF SIZES

In selecting a group of growth curves upon which size distributions could be based, there were several considerations. One would, of course, like to include most of the oldest, most heavily used curves and to include generalizations thereof which would provide the extra flexibility of shape needed to describe the more complicated forms of growth. One would also prefer that such a group or family of curves have a simple interpretation for its parameters in terms of usual biological meaning. Such an example is the family of growth curves published by Turner _et al_. (1976) in _Mathematical Biosciences_. In the development of this theory, x is defined as the size of the population at any point in time t. This x could be some dimension of a particular organism such as length, height, weight or surface area. Now x could equally well be a measurement on some specific organ within that organism. On the other hand, x could be interpreted as approximately (since continuity of measure is assumed) counting the number of individuals in a population of cells, species, city residents or words in a manuscript. The maximum size to which the population (or an organism) is able to grow is κ. The limitation of growth may be taken as due to factors inherent to the organism or to inhibiting factors present in the environment. For example, a cell culture may exhaust the nutrient solution supplied unless it is constantly renewed. This maximal growth is assumed to be the limit of x as t increases without bound (although in three main cases, this maximum will actually occur at a finite time τ). The rate of growth or rate of change of size per unit of change in time is $\dot{x}$ or dx/dt. One might easily hypothesize that this rate of growth may not be constant. It seems reasonable that growth may depend upon some function of the distance between the origin (just before growth begins) and the current size at time t. For this reason, Turner has defined $\delta_n(0,x) = (x^n-0^n)^{1/n} = x$, a generalized distance from the origin to the current size. Symmetrically, a growth may depend upon the distance between the current size and the maximum size. This generalized distance function is called $\delta_n(x,\kappa)$ and is equal to $(\kappa^n-x^n)^{1/n}$.

For this development of the theory, Turner chooses three basic postulates of growth. Firstly, the rate of growth is assumed to be jointly proportional to monotonically increasing functions both of the distance from the origin to the current size and of the distance from the current size to the maximum size. Formally this may be represented by

$$\dot{x} \propto \phi_1[\delta_n(0,x)]\phi_2[\delta_n(x,\kappa)].$$

Secondly, these monotonically increasing functions ϕ_1 and ϕ_2 are both assumed to be power functions, so that $\phi_1[\cdot] = [\cdot]^{\theta_1}$ and $\phi_2[\cdot] = [\cdot]^{\theta_2}$ where θ_1 and θ_2 are both greater than zero. Under the third assumption θ_1 and θ_2 obey a pair of constraints: $\theta_1 = 1\text{-}np$ and $\theta_2 = n\text{+}np$. These constraints allow for greater mathematical tractability. They also cause the generation of a master or generic curve for which simple limits and/or assignation of particular parametric values can easily reduce to the historically important cases derived by Verhulst in

1838, Gompertz (1825), Bertalanffy (1957) and Richards (1959). Of course, since θ_1 and θ_2 were assumed to be positive, constraints are implied for both n and p. Originally when n was defined in the context of the distance functions, it became apparent that in order for the distance meaning to remain intact, n must be a positive quantity. Coupled with these additional constraints, one sees that $-1<p<1/n$ and $n>0$ so that $\theta_1+\theta_2 = n+1$. Under these definitions and assumptions, the generic rate relation is obtained as

$$\dot{x} \propto x^{1-np}(\kappa^n - x^n)^{1+p}.$$

If the proportionality constant is called δ/κ^n (which may be done without loss of generality), δ may be considered as an intrinsic growth constant measureable in reciprocal time units. Then the generic rate equation becomes

$$\dot{x} = \delta\kappa^{-n}x^{1-np}(\kappa^n - x^n)^{1+p}.$$

In an as yet unpublished manuscript, Professors Turner and Pruitt have derived the same generic curve for growth of an organism considering growth to be a result of the interacting forces of anabolism and catabolism. In this derivation, which follows the approach of Pütter (1920) and Bertalanffy (1957), the growth rate is seen to be proportional to some power, call it N, of the difference between the results of anabolism and those of catabolism. The anabolic part is assumed to be directly proportional to some power, a, of the current size, while the catabolic contribution is directly proportional to another power, b, of the current size. If the following constraints are imposed, the problem becomes mathematically tractable:

$$a = (1-np)/(1+p)$$

$$c = (1+n)/(1+p)$$

$$N = p+1.$$

Supplying proportionality constants as before, they obtain the following form of the generic rate equation:

$$\dot{x} = \delta\kappa^{-n}\{n/|n|[\kappa^{n}x^{(1-np)/(1+p)}-x^{(1+n)/(1+p)}]\}^{1+p}.$$

Note, however, that in this formulation, it is not necessary to place restrictions upon the sign of either n or p. Nor is it necessary that p be less than 1/n, as the other derivation required. Of course, for the case of a positive-valued n, the generic rate equation reduces to its previous form:

$$\dot{x} = \delta\kappa^{-n}x^{1-np}(\kappa^{n}-x^{n})^{1+p},$$

for which the growth equation itself is solvable as

$$x = \kappa\{1+[1+\delta np(t-\tau)]^{-1/p}\}^{-1/n}.$$

τ, the constant of integration, is a function of all of the parameters in the equation and of the smallest value (greater than zero) which can be measured. Let this smallest size, which occurs at t = 0, be called γ. Rewriting τ in this manner may in some instances provide a convenience.

It would seem that a discussion of the subcases and their characteristic shapes is the next order of business. A flow chart of the published list of subcases would have three main branches: the limiting case as n approaches 0 provides the Gompertz type subclass; the limiting case as p approaches 0 provides the Bertalanffy-Richards type subclass; setting n equal to 1 in the generic curve yields a subclass of logistic style models. Of course, these three main subclasses may be further interconnected by selecting parameter values or taking limits involving subclass members. All three lines of development include the case of exponential or geometric increase growth as the bottom special case when the upper limit to growth is allowed to increase without bound. Taking these subcases one at a time, one may start with the Gompertz variety of curve. A curve, termed hyper-Gompertzian, is derivable from the generic rate equation as the limit when n approaches 0 and δn^{1+p} approaches some new constant, call it δ'. The new rate equation becomes $\dot{x} = \delta' x[\log(\kappa/x)]^{1+p}$ for which the solution is

$$x = \kappa \exp\{-[\delta' p(t-\tau)]^{-1/p}\}.$$

If the limiting process is then applied to this curve as p approaches zero, one obtains the curve due to Gompertz (1825). This rate equation has the form $\dot{x} = \delta' x[\log(\kappa/x)]$ which, upon solution, displays the well-known

$$x = \kappa \exp\{-\exp[-\delta'(t-\tau)]\}.$$

The Gompertzian growth function, in turn, reduces to the old exponential growth function with $\dot{x} = \delta x$ and $x = \exp[\delta(t-\tau)]$ as the limiting case when κ approaches infinity and $\delta' \log \kappa$ approaches some constant δ (the same δ as that of the generic curve).

Characteristic of the Gompertz curve, and even more so of the hyper-Gompertzian curve, is a very early and steep rise in the growth function, a maximum growth rate (occurring at the inflection point of the curve) being virtually maintained over an extended segment of the actual growting period which is followed by a relatively abrupt leveling off as a longer period (indicated by the long right-hand tail to the distribution) of nearly maximized size begins. This type of extremely rapid early growth has often been thought to characterize growth of abnormal tissue, particularly that of transplanted tumors. See Laird et al. (1965) and Thurman et al. (1971). It may also prove useful in studying embryonic growth. Geometric increase growth is more useful to describe a portion of such early growth (before the leveling off can be observed) or perhaps the early growth of primary tumors. See Thurman et al. (1971). It has also been used by many researchers to describe the "steady-state" growth of cell cultures for which the nutrient solution is being constantly replenished. There is, for this subcase, no leveling off or indication of reaching a finite growth limit.

A second subclass of curves, logistic in type, may be seen to arise from the generic rate equation with the parameter n set equal to 1. This produces the so-called hyperlogistic curve with rate equation $\dot{x} = \delta/\kappa\, x^{1-p}(\kappa-x)^{1+p}$. The corresponding growth equation is

$$x = \kappa/\{1+[1+\delta p(t-\tau)]^{-1/p}\}.$$

If the limit is taken as p approaches 0 in the hyperlogistic rate equation, the rate equation of the Verhulst logistic appears: $\dot{x} = \delta/\kappa\, x(\kappa-x)$. The growth equation which is probably the most often employed in practice results: $x = \kappa/[1+e^{-\delta(t-\tau)}]$. As κ is allowed to increase without bound in the logistic rate equation, the geometric increase rate equation reappears.

The logistic curve, in contradistinction to the Gompertz curve, is skew symmetric, having a long, slow beginning followed by a gentle rise to its inflection point with a gentle (and equal) falling away from this maximum growth rate and an equally long, slow leveling off as maximum growth is reached. τ is here not only the constant of integration for the curve but indicates the point having the maximum growth rate, the inflection point, and a kind of half-life, the point at which half of maximum growth is obtained. Such curves have often been used to describe the normal growth of populations and of individual organisms. See Pearl and Reed (1920), Pearl (1921), Pearl, Edwards and Miner (1934) and Krause _et al_. (1967). The hyperlogistic curve is no longer necessarily skew symmetric although the τ does still represent the half-way point to ultimate size. It is not necessarily true that the slope of the growth curve shows a maximum at this point, however. On the other hand, δ is equal to $(4/\kappa)\dot{x}_{t=\tau}$, so δ does actually measure that slope, whatever it is. The effect of increasing the value of p upward from zero is to lengthen the right-hand tail differentially.

So far, very little has been said about the effects of a possibly negative p. In this circumstance, maximum growth occurs in finite time at t equal to $\tau-1/(\delta p)$. This causes a much faster rise to the maximum than would have been true for positive values of p. The value of x at t equal to τ is still $\kappa/2$, but the rise to the maximum is fairly abrupt from there. For the left-hand tail, the intersection with the axis at t equal to zero occurs at a larger value of x, perhaps implying a more severe truncation on that side. The left-hand tail will be longer, in general, than the right-hand tail and will retain the more gradual acceleration of slope as the inflection point is approached.

Returning for a moment to the other side of this system of curves based on the generic curve, one should mention the character of the hyper-Gompertz growth function for negative values of p. In this case also maximum size occurs in finite time, at t equal to τ. The point at which half of the maximum size is attained occurs at t equal to $\tau+1/(\delta' p)[\log 2]^{-p}$ and δ' is a rough measure of the slope, the half-life (or half-growth), for fixed p and κ:

$$\delta' = \{\kappa/2[\log 2]^{1+p}\}^{-1} \dot{x}_{x=\kappa/2}.$$

As p approaches -1, the function becomes more and more like the geometric increase curve. The case of a negative value of n, with its subsequent change in the formal manner of constructing the generic curve and its other family members, is not considered in this development.

The third subclass of curves flows out of the generic growth rate if one proceeds to take the limit as p approaches zero. The resulting rate equation is the familiar version due to Bertalanffy (1938, 1957) and to Richards (1959):

$$\dot{x} = \delta\kappa^{-n}x(\kappa^{n}-x^{n}).$$

Solving this equation, one can display the growth curve itself:

$$x = \kappa\{1+e^{-\delta n(t-\tau)}\}^{-1/n}.$$

For this function, t equal to τ yields the half-life point only if n equals 1, the logistic case. In the case of n more generally defined, the point at which half of maximum size is found will be located at t equal to $\tau-1/(\delta n)\log(2^{n}-1)$. As for the parameters δ and n, the reciprocal of n is a measure of the logarithm of the ratio of maximum size to size at t equal to τ:

$$n = \{\log 2\}\{\log[\kappa/x_{t=\tau}]\}^{-1}.$$

The slope at t equal to τ for fixed values of n is still related to δ:

$$\delta = \kappa^{-1}2^{1+1/n}\dot{x}_{t=\tau}.$$

For fixed values of δ and κ, an increase in the size of n causes the value of x at t equal to τ to approach the maximum size κ. This has the effect of steepening the rise of the curve and of shortening the right-hand tail of the function. If n is allowed to become smaller than 1 and to approach zero, the function resembles more closely the Gompertz, the limiting case from the Bertalanffy-Richards as n approaches 0 and δn approaches the constant δ'. The apparent lengthening of the right-hand tail also follows. Here again taking the limit as κ approaches infinity reduces the function to the geometric increase curve.

One other special case of some interest has been obtained from this generic growth system. That is the case in which the rate of growth is proportional to some power of the current size, not necessarily equal to 1. This relationship is obtainable as a limiting case from the hyper-Gompertz, from the hyperlogistic and from the generic curve itself. From the hyper-Gompertz equation, one requires the limit of $\dot{x}$ as κ increases without bound such that $(p+1)/\log \kappa$ approaches a constant, call it α, and $\delta(\log \kappa)^{p+1}$ approaches another constant, call it δ''. The rate equation thus deduced is $\dot{x} = \delta'' x^{1-\alpha}$. The corresponding growth equation can be written as $x = [\alpha\delta''(t-\tau)]^{1/\alpha}$, where α is greater than zero since the lower limit of p allowed in this discussion is -1. From either the generic or the hyperlogistic rate equations, this power function rate equation is derived as the limiting form when κ increases without bound and $\delta\kappa^{np}$ approaches the constant δ''. In this form, the product np contains the only appearance of the two parameters, which are, therefore, no longer separable. Call this product a new parameter, α. In this last derivation, it is not apparent that α must be positive. If, in fact, α is negative, this implies that an infinite size could be approached at a finite time τ. The shape of the growth function varies according to the size and sign of α. If α is equal to 1, growth is linear. If α is greater than 1, growth follows a slowly rising curve with a curvature concave away from the ordinate axis. If α is between 0 and 1, growth follows a swiftly rising curve, concave upward. As α approaches 0, this swiftly rising curve reduces to the limiting form of exponential growth. As α moves farther negative, away from 0, the curve maintains its upward concavity but becomes flatter and flatter at the early time points as α decreases. As t comes nearer to τ, however, the rise in the flat curve becomes quite suddenly steep and each succeeding unit of time pushes the growth curve a considerable step forward toward its infinite asymptote.

V. THE SIZE DISTRIBUTIONS

Now the exploration of the new family of size distributions begins. The family is intended to represent the distribution of observations taken in a cross-section of time from an underlying "growth" process. The word "growth" is placed in quotation marks as a reminder that, by it, is intended any process of accretion, including but not limited to biological growth. The growth curves upon which these distributions are founded are those of the Turner family of curves discussed in the preceding chapter. The general method employed is the simultaneous solution of two differential equations in time and the elimination of time between them. This approach to the size problem was probably first used by Yule (1924) to obtain a rough estimate for his own species abundance distribution. The second of the two differential equations is always one specifying a particular member of the growth curve family. It is a deterministic equation. Every individual in a given population is assumed to follow exactly the same curve with the same starting and ending points and the same scale and shape parameters. It would be possible to insert probability distributions for the parameters and so to generalize. This, however, is a subject for future research.

The first differential equation has two possible interpretations. It may be thought of as an acquisition rate equation for the collection of a new member of the population to some set of records or it may be considered to be the differential equation for the probability distribution of ages of the members of the population. For the acquisition rate explanation t is defined to be the time point in the growth of an organism at which it is acquired to the record set in question. It is not necessary that this time point be the same for all such organisms. Then the acquisition rate equation is of the following form:

$$\dot{F}(t)/[1-F(t)] = g(t).$$

So the acquisition function itself has the form

$$-\log[1-F(t)] \equiv y$$

where y is $\int g(t)dt$ + a constant, c. Clearly dy/dt is g(t). If this function F(t) is suitably normalized so that

$$\begin{aligned} F(t) &= 0, & t &< 0 \\ 0 \le F(t) &\le 1, & 0 \le t &\le \tau \\ \text{and } F(t) &= 1, & t &> \tau \end{aligned}$$

and if, in addition, F(t) is monotonically nondecreasing, then F(t) may be interpreted as a distribution function for t. One could, in this case, call y the log odds of some t_0 being greater than t.

In this paper, the assumption made is that this acquistion rate, g(t), is a constant, c_0. Integrating and applying the initial condition that F(t) = 0 for t = 0, one sees that the log odds or acquisition function becomes

$$-\log[1-F(t)] = c_0 t$$

$$\text{or } y = c_0 t.$$

More and more organisms are acquired as time increases upward from 0. Or, restated, only a few organisms are acquired for small values of t and more are added as t increases toward τ (or toward infinity, depending upon whether maximal growth is postulated to occur in finite time or not). This increasing chance of being acquired to the records as the individual grows over time naturally results in distributions of sizes of individuals which tend to stack at the large end; that is, they tend to be left-skewed. It is interesting to note that the acquisition rate function is sometimes known in other applications as the intensity function or the hazard rate function, while the reciprocal is known as Mill's ratio.

If the first equation is seen as an acquisition rate equation, the development proceeds in this manner. First the two differential equations are solved in terms of t and the appropriate constants determined by the application of the initial conditions. The first equation yields a cumulative distribution function for the time in an organism's growing cycle at which its size may be recorded. The second yields the size of the organism deterministically as a function of its time since "birth". It is the same time point in both equations. By inverting the solution for the growth function, one can easily find an expression for t in terms of x, the current size of the organism. The substitution of this expression into the cumulative distribution function for acquisition time produces a new cumulative distribution function for the sizes of the organisms recorded. Whether or not all the organisms in the recorded set are really of the same age is not relevant here. Either may be true. The observations differ because they were taken at different time points in the growth of the organisms.

On the other hand, the first differential equation may be thought of as the defining equation for a probability distribution of ages in a population of mixed ages. Again, if the growth equation gives a size for any t, or any age, it can be solved backwards for the age t in terms of size x. Making this transformation in the distribution function for ages yields the corresponding distribution function for sizes in a population heterogeneous with respect to age. It is quite possible to encounter cross-sectional data of either variety in practice, although a mixture of these types is probably more common-place. It is useful to have a formulation that is able to cover both circumstances.

All of the size distributions in the family presented here have the same choice for the first differential equation. It is the second, the growth equation, that has been allowed to vary from one member of the family to the next. For the first equation, the simplest possible model was chosen so that the equations might be more easily manipulated and the properties of such a family studied. Therefore, the choice was a constant acquisition rate or a negative exponential distribution of ages, depending upon the interpretation one prefers. This means that the probability of getting another observation in the next interval from t to t+dt given that there will be another observation obtained after time t is the same, constant, for any choice of the t to t+dt interval. From the hazard function point of view in life tables, this would correspond to getting another death in the next interval. For the age distribution version, this would describe a population in which there were very many small children, for instance, and a few adults of advanced age with a gradually declining proportion of persons with ages in between these extremes. One should make haste to add that many other interesting distributions of ages may be postulated. The Weibull distribution as a generalization of the exponential is rather more flexible and would make a nice step upward in complexity. Both distributions have the advantage of possessing a closed form cumulative distribution function. For this investigation, the simpler of the two was used. The resulting distributional forms are more constrained than would have been so otherwise. The possible shapes attained by family members include J-shaped curves, U-shaped curves, twisted J-shaped curves, reversed J-shaped curves and unimodal curves with long tails and a more gentle rise on the left followed by a rapid descent on the right. Now one may advance to the first and simplest case.

The most primitive member of the Turner growth curve family is the case of geometric or Malthusian growth. For the initial conditions, one may state that at the original time, $t = 0$, the cumulative distribution function $F(t)$ is equal to 0 and the initial size of x, the point at which size may first be measured, is equal to some small positive quantity, γ. This assumption will allow the determination of the nature of the constants of integration. Now the equation for a constant acquisition rate is of the form $\dot{F}(t) = \beta\delta(1-F(t))$ where $\dot{F}(t)$ is the time derivative of $F(t)$, the cumulative distribution function of t. Let $\beta\delta$ represent a single positive quantity, the rate constant. It is written here as the product of two positive parameters β and δ to simplify the naming of the corresponding parameter in the final distribution of sizes. No loss of generality is incurred by doing so. Now one solves this equation for $F(t)$:

$$\dot{F}(t) = \beta\delta(1-F)$$

$$-\dot{F}(t)/[1-F(t)] = -\beta\delta$$

$$\log(1-F(t)) = -\beta\delta t+\log c$$

$$1-F(t) = ce^{-\beta\delta t}$$

$$F(t) = 1-ce^{-\beta\delta t}.$$

Applying the initial conditions, one determines that the value of the constant of integration, c, is 1. Thus, the distribution of the acquisition time, t, or of ages in a population heterogeneous with respect to age, is

$$F(t) = 1-e^{-\beta\delta t}.$$

This same derivation and solution are applicable for all the size distributions developed in this paper and will only be repeated in summary form hereafter. The terminal conditions for this member of the family are that, as t approaches infinity, F(t) approaches unity and x, the size, increases without bound. If $\dot{x}$ is the time derivative of size at time t, then the differential equation for Malthusian growth is of the form $\dot{x} = \delta x$ where δ is the positive growth rate constant. Solving this equation for x, one has

$$\dot{x}/x = \delta$$

$$\log x = \delta t+\log c'$$

$$x = c'e^{\delta t}.$$

Again applying the initial conditions, one sees that the constant of integration c' is equal to the minimal size, γ, and the equation becomes

$$x = \gamma e^{\delta t}.$$

Solving this equation, in turn, for t in terms of x, one finds the deterministic relation that

$$t = 1/\delta \log(x/\gamma).$$

This relationship may then be used as an ordinary transformation in the equation for F(t), which becomes an equation for F(x), the cumulative distribution function of sizes given a negative exponential distribution of acquisition times (or of ages) and geometric increase growth:

$$F(x) = 1-e^{-\beta\delta[1/\delta \log(x/\gamma)]}$$

$$F(x) = 1-e^{-\beta \log(x/\gamma)}$$

$$F(x) = 1-(x/\gamma)^{-\beta}.$$

This distribution will be readily recognized as the Pareto, discussed in chapter two. Differentiating with respect to x, one obtains its probability density function,

$$f(x) = \beta/\gamma(x/\gamma)^{-(\beta+1)};\ \gamma<x<\infty,\ \beta>0.$$

The first derivative of f(x) with respect to x is

$$f'(x) = -\beta(\beta+1)\gamma^{\beta}x^{-(\beta+2)}.$$

Since β and γ are positive, $f'(x)$ is negative for all values of x greater than γ. The second derivative of f(x) is positive for all x greater than γ:

$$f''(x) = \beta(\beta+1)(\beta+2)\gamma^{\beta}x^{-(\beta+3)}>0.$$

Thus one can see that f(x) is a J-shaped function declining from $\beta\gamma^{-1}$ at $x = \gamma$ toward 0 as x increases without bound.

For the next size distribution, the assumption, along with the constant acquisition rate, is a logistic growth pattern. The initial conditions remain as before: at t equal zero, F(t) equals zero and x approaches γ. The final conditions are changed to indicate a finite limit for growth: as t approaches infinity, F(t) approaches unity and x approaches the finite quantity κ. The first differential equation and its solution remain unchanged:

$$\dot{F}(t) = \beta\delta(1-F(t))$$

$$F(t) = 1-e^{-\beta\delta t}.$$

The second differential equation becomes that equation which specifies a logistic growth function,

$$\dot{x} = \delta\kappa^{-1}x(\kappa-x).$$

Solving the equation in the usual manner, one has

$$(1/x+1/[\kappa-x])dx = \delta dt$$

$$\log x-\log(\kappa-x) = \delta t+\log c'$$

$$\log(\kappa/x-1) = -\delta t-\log c'$$

$$\kappa/x-1 = 1/c'e^{-\delta t}$$

$$\kappa/x = 1+1/c'e^{-\delta t}$$

$$x = \kappa(1+1/c'e^{-\delta t})^{-1}.$$

Note that the expression $1/c'$ is the same as Turner's $e^{\delta\tau}$ since $e^{-\delta\tau}$ is his form for the constant of integration. Applying the initial conditions, one sees that, in terms of κ and γ, $1/c'$ is the same as $(\kappa/\gamma-1)$. So the final form for the logistic growth curve solution is

$$x = \kappa[1+(\kappa/\gamma-1)e^{-\delta t}]^{-1}.$$

Now one must solve this deterministic equation backward to find t in terms of x:

$$\kappa/x = 1+(\kappa/\gamma-1)e^{-\delta t}$$

$$(\kappa/x-1) = (\kappa/\gamma-1)e^{-\delta t}$$

$$(\kappa/\gamma-1)^{-1}(\kappa/x-1) = e^{-\delta t}$$

$$-\delta t = \log[(\kappa/\gamma-1)^{-1}(\kappa/x-1)]$$

$$t = -1/\delta \log[(\kappa/\gamma-1)^{-1}(\kappa/x-1)].$$

Substituting this expression in the equation for F(t), one obtains F(x) as

$$F(x) = 1-e^{-\beta\delta\{-1/\delta \log[(\kappa/\gamma-1)^{-1}(\kappa/x-1)]\}}$$

or

$$F(x) = 1-e^{\beta \log[(\kappa/\gamma-1)^{-1}(\kappa/x-1)]}$$

or

$$F(x) = 1-(\kappa/\gamma-1)^{-\beta}(\kappa/x-1)^{\beta}.$$

Differentiating, one finds the form of the probability density function:

$$f(x) = \beta\kappa(\kappa/\gamma-1)^{-\beta}x^{-2}(\kappa/x-1)^{\beta-1}, \quad \gamma<x<\kappa.$$

To investigate the behavior of this function, one may produce the first and second derivatives of f(x) with respect to x:

$$f'(x) = -\beta\kappa(\kappa/\gamma-1)^{-\beta}x^{-4}(\kappa/x-1)^{\beta-2}\{2x(\kappa/x-1)+(\beta-1)\kappa\}$$

and

$$f''(x) = \beta\kappa(\kappa/\gamma-1)^{-\beta}x^{-6}(\kappa/x-1)^{\beta-3}\{6x^2-6(\beta+1)\kappa x+(\beta+1)(\beta+2)\kappa^2\}.$$

Setting $f'(x)$ equal to zero, one sees that there is a critical point at $x = \kappa/2(\beta+1)$ and, checking the boundary conditions for x, one sees that this point is in the range from γ to κ for β such that

$$\sup(0,2\gamma/\kappa-1)<\beta<1.$$

Next evaluating the second derivative at the critical point, one finds that for $0<\beta<1$, it is positive, indicating the presence of a minimum if also $\sup(0,2\gamma/\kappa-1)<\beta<1$. There is no maximum at the critical point within the allowable range $\gamma<x<\kappa$. In other words, for β between $\sup(0,2\gamma/\kappa-1)$ and unity, the distribution will be U-shaped. If β is greater than or equal to unity, the curve will always be J-shaped. If, however, it does happen because of the scaling that $(2\gamma/\kappa-1)$ is greater than zero, implying that $\kappa<2\gamma$, and $0<\beta\leq(2\gamma/\kappa-1)$, the form of the curve will be a reversed J with the supremum at κ instead of at γ. This serves to point out an important fact about this system of curves. They are not shape invariant with respect to different choices of γ and κ.

The third member of the family is the case of Gompertzian growth. The acquisition rate equation and solution remain as before:

$$\dot{F}(t) = \beta'\delta'(1-F(t))$$

and $F(t) = 1-e^{-\beta'\delta' t}$.

The growth differential equation becomes instead

$$\dot{x} = \delta' x \log(\kappa/x).$$

In this equation, it is easily seen that $\delta' x$ represents a Malthusian growth piece while for $x<\kappa e^{-1}$ the function $\log(\kappa/x)$ will be greater than one and will accelerate the growth rate above that for straight Malthusian growth. The point $x = \kappa e^{-1}$ would coincide with the Malthusian curve point at that value. Then for $e^{-1}\kappa<x<\kappa$ the rate of growth $\dot{x}$ declines below that of Malthusian growth. Integrating this growth rate equation, and applying the initial conditions, which are the same as for the logistic, one obtains the Gompertzian growth function

$$x = \kappa e^{-[\log(\kappa/\gamma)]e^{-\delta' t}}.$$

Then t may be written as a function of x as

$$t = 1/\delta'[\log\log(\kappa/\gamma)-\log\log(\kappa/x)].$$

Substituting in the acquisition time or age cumulative distribution function, one has

$$F(x) = 1-e^{-\beta'\delta'\{1/\delta'[\log\log(\kappa/\gamma)-\log\log(\kappa/x)]\}}$$

which reduces to

$$F(x) = 1-e^{-\beta'[\log\log(\kappa/\gamma)-\log\log(\kappa/x)]}$$

or

$$F(x) = 1-[\log(\kappa/\gamma)]^{-\beta'}[\log(\kappa/x)]^{\beta'}.$$

The probability density function is found to be

$$f(x) = \beta'[\log(\kappa/\gamma)]^{-\beta'}x^{-1}[\log(\kappa/x)]^{\beta'-1}.$$

As with the logistic, one proceeds now to investigate the nature of this function through its derivatives:

$$f'(x) = -[\log(\kappa/\gamma)]^{-\beta'}\beta' x^{-2}[\log(\kappa/x)]^{\beta'-2}\{\log(\kappa/x)+(\beta'-1)\}$$

and

$$f''(x) = [\log(\kappa/\gamma)]^{-\beta'}\beta' x^{-3}[\log(\kappa/x)]^{\beta'-3}\cdot$$
$$\{2[\log(\kappa/x)]^2+3(\beta'-1)[\log(\kappa/x)]+(\beta'-1)(\beta'-2)\}.$$

Setting the first derivative equal to zero and solving for x, one finds that f(x) has a critical point at $x = \kappa e^{\beta'-1}$. This point is in the range from γ to κ for all β' such that $\sup(0,1+\log(\gamma/\kappa))<\beta'<1$. Evaluating the second derivative, $f''(x)$, at this critical point, one sees that it is greater than zero for β' less than unity. Again one has the presence of a minimum. There is no maximum in the range $\gamma<x<\kappa$ for any value of β'. If β' is greater than or equal to unity, $f'(x)$ will be negative for all x between γ and κ. In this case, the curve will be J-shaped with a supremum at $x = \gamma$. If, however, $1+\log(\gamma/\kappa)$ is greater than zero, which implies $\kappa<\gamma e$, then for $0<\beta'\leq 1+\log(\gamma/\kappa)$ the curve will have a reversed J shape with a supremum at $x = \kappa$. So this curve, as well as the logistic based curve, may be J-shaped, U-shaped or reversed J-shaped. It is never unimodal. The difference in type of J- or U-shaped curve produced by a Gompertz growth pattern versus a logistic form is that the logistic-based curve bends more slowly, tending to put a long section of linear movement in the curve. The Gompertz-based curve has more curvature and tends to have a steeper initial drop in the density function.

The fourth curve considered and the first with more than three parameters is that based upon a constant acquisition rate and Bertalanffy-Richards growth. The acquisition rate equation and its solution are as before:

$$\dot{F}(t) = \beta\delta[1-F(t)]$$

$$\text{and } F(t) = 1-e^{-\beta\delta t}.$$

The Bertalanffy-Richards growth rate equation is

$$\dot{x} = \delta\kappa^{-n}x(\kappa^n-x^n);\ n,\ \delta>0.$$

Solving this equation for x in terms of t and applying the initial conditions, which are the same as for the Gompertz and the logistic, namely at $t = 0$, $x = \gamma$ and $F(t) = 0$, one finds the Bertalanffy-Richards growth equation to be

$$x = \kappa\{1+[(\kappa/\gamma)^n-1]e^{-\delta nt}\}^{-1/n}.$$

Now reversing the solution, one obtains an expression for t in terms of x:

$$t = 1/\delta n\{\log[(\kappa/\gamma)^n-1]-\log[(\kappa/x)^n-1]\}.$$

Transforming the cumulative distribution function F(t), one derives the Bertalanffy-Richards based size distribution function,

$$F(x) = 1-[(\kappa/\gamma)^n-1]^{-\beta/n}[(\kappa/x)^n-1]^{\beta/n}.$$

The size density for this case is

$$f(x) = [(\kappa/\gamma)^n-1]^{-\beta/n}\beta x^{-1}(\kappa/x)^n[(\kappa/x)^n-1]^{\beta/n-1},\ \gamma<x<\kappa.$$

Now the limit for this distribution as n approaches zero is the Gompertz-based distribution while, for n equal to one, the logistic-based distribution is obtained. Clearly, then, for values of n between zero and one, this distribution graduates change between the Gompertz-based and the logistic-based models. One may study this curve's characteristics through its first two derivatives:

$$f'(x) = [(\kappa/\gamma)^n-1]^{-\beta/n}\beta\kappa^n x^{-2(n+1)}[(\kappa/x)^n-1]^{\beta/n-2}\cdot$$
$$\{(n+1)x^n-(\beta+1)\kappa^n\}$$

and

$$f''(x) = [(\kappa/\gamma)^n-1]^{-\beta/n}\beta\kappa^n x^{-(n+3)}[(\kappa/x)^n-1]^{\beta/n-3}\cdot$$
$$\{(\kappa/x)^{2n}(\beta+1)(\beta+2)+(n+1)(n-3\beta-4)(\kappa/x)^n+(n+1)(n+2)\}.$$

One sets the first derivative $f'(x)$ equal to zero and solves for the critical point which is located at

$$x = \kappa[(\beta+1)/(n+1)]^{1/n}.$$

One finds that this critical point is actually in the range from γ to κ for all β such that

$$\sup(0,(n+1)(\gamma/\kappa)^n-1)<\beta<n.$$

Next one evaluates the second derivative, $f''(x)$, at this point:

$$f''(x)\big|_{x=\kappa[(\beta+1)/(n+1)]^{1/n}} = [(\kappa/\gamma)^n-1]^{-\beta/n}\beta\kappa^{-3}\cdot$$

$$[(n+1)/(\beta+1)]^{(\beta-n+3)/n}n(n-\beta).$$

One sees that this second derivative is negative if n is less than β, zero if n is equal to β and positive if n is greater than β. However, only in the last case of n greater than β is the point in the range. So the conclusion for the Bertalanffy-Richards based curve is similar to that for the Gompertz and logistic based forms. The curve is U-shaped with its minimum at $x = \kappa[(\beta+1)/(n+1)]^{1/n}$ for the case of

$$\sup(0,(n+1)(\gamma/\kappa)^n-1)<\beta<n.$$

If n is less than or equal to β, $f'(x)$ is always negative and the curve is J-shaped with its supremum at γ. If it should happen that $[(n+1)\{(\gamma/\kappa)^n\}-1]$ is positive, which implies that $\gamma>\kappa(n+1)^{-1/n}$, and if β is in the range $0<\beta<[(n+1)\{(\gamma/\kappa)^n\}-1]$, then the curve will be of the reversed-J type with its supremum at κ. As n increases above unity, the Bertalanffy-Richards based curve will be pushed increasingly toward the Pareto case of Malthusian growth and exponential acquisition.

These four size distributions have in common that they are a sub-family of J and U shaped curves which may be either right-skewed or left-skewed. None of them may be either of a twisted J shape or unimodal. In discussing the estimation and the choice of model in subsequent chapters, it sometimes proves useful to be able to refer to this set as a group. For this reason, they are dubbed here "the lower four" or "the lower curves". One may also refer to the other members of the general family as "the upper four" or "the upper curves".

To begin the discussion of the upper curves, one may note that for each growth curve used, there are two size distributions generated--

one for values of the parameter p that are greater than zero and one for values between zero and -1. The terminal conditions are different in the two cases. For positive values of p, the initial and terminal conditions are the same as for the Gompertz, the logistic and the Bertalanffy-Richards curves: at $t = 0$, one has $x = \gamma$ and $F(t) = 0$; as t increases without bound, x approaches κ and $F(t)$ approaches unity. For negative values of p, however, x reaches κ and $F(t)$ reaches unity at a finite time whose exact mathematical representation will vary from one growth function to another. The correct one for each member will be discussed as the corresponding size distribution is developed.

For the first derivation in the upper curve group, one may consider the hyperlogistic growth rate. The acquisition rate equation remains $\dot{F}(t) = \beta\delta(1-F(t))$ and its solution, $F(t) = 1-e^{-\beta\delta t}$. The growth rate equation has two terminal conditions depending upon whether p is positive or negative. For both cases, the initial conditions are that for $t = 0$, $x = \gamma$ and $F(t) = 0$. One may start with p greater than zero. Here one has the usual terminal conditions: as t approaches infinity, x approaches κ and $F(t)$ approaches one. The hyperlogistic growth rate equation postulated by Turner et al. and discussed in the previous chapter is

$$\dot{x} = \delta\kappa^{-1}x^{1-p}(\kappa-x)^{1+p}.$$

Now solving this equation for x in terms of t and applying the initial conditions, one produces the hyperlogistic growth equation:

$$x = \kappa\{1+[\delta pt+(\kappa/\gamma-1)^{-p}]^{-1/p}\}^{-1}.$$

Solving this deterministic equation in reverse for t as a function of x, one finds

$$t = 1/(\delta p)[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}].$$

Next one may substitute this expression for t into the cumulative distribution function, F(t), to yield the distribution of sizes assuming hyperlogistic growth:

$$F(x) = 1-e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}.$$

The corresponding probability density function is found to be

$$f(x) = \beta\kappa x^{-2}(\kappa/x-1)^{-(p+1)}e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}.$$

For negative values of p, between zero and -1, the current size x reaches its maximum κ at a finite time. Professor Turner's original expression of his growth function's constant of integration is in terms of τ. If this formulation is used, the finite time at which growth is maximal may be written as $\tau-1/(\delta p)$ for the hyperlogistic. Then the summary of the terminal conditions is that, as t approaches $\tau-1/(\delta p)$, x approaches κ and F(t), given that t is less than or equal to $\tau-1/(\delta p)$, approaches unity. In other words, a conditional cumulative distribution function given that t is less than or equal to $\tau-1/(\delta p)$ is needed. This function may be written as

$$F(t\,|\,t\leq\tau-1/(\delta p)) = (1-e^{-\beta\delta t})/(1-e^{-\beta\delta\tau+\beta/p}).$$

The growth rate equation remains the same

$$\dot{x} = \delta\kappa^{-1}x^{1-p}(\kappa-x)^{p+1}.$$

Its solution, written using Professor Turner's τ, is

$$x = \kappa\{1+[1+\delta p(t-\tau)]^{-1/p}\}^{-1},$$

or, noting that his τ is equal to $-1/(\delta p)[(\kappa/\gamma)^{-p}-1]$, here, one sees that the form of the growth equation also remains unchanged:

$$x = \kappa\{1+[\delta pt+(\kappa/\gamma-1)^{-p}]^{-1/p}\}^{-1}.$$

As before, the reversed equation for t as a function of x is written as

$$t = 1/(\delta p)[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}].$$

Substituting this function for t into the conditional cumulative distribution function, one sees that, for $-1<p<0$, one has

$$F(x|t\leq\tau-1/(\delta p)) = \{1-e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}\}$$
$$/[1-e^{\beta/p(\kappa/\gamma-1)^{-p}}].$$

One may examine the condition in terms of x also:

$$t\leq\tau-1/(\delta p),$$

$$1/(\delta p)[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]\leq-1/(\delta p)[(\kappa/\gamma-1)^{-p}-1]-1/(\delta p),$$

$$1/(\delta p)(\kappa/x-1)^{-p}-1/(\delta p)(\kappa/\gamma-1)^{-p}\leq-1/(\delta p)(\kappa/\gamma-1)^{-p}+1/(\delta p)-1/(\delta p),$$

$$1/(\delta p)(\kappa/x-1)^{-p}\leq 0.$$

Since δ is positive and $(\kappa/x-1)$ is positive or equal to zero for all x, it follows that the condition only reduces to $p<0$, which is the case under consideration. So, one may drop the statement of the condition in terms of t and τ and write that, for $-1<p<0$, one has

$$F(x) = \{1-e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}\}/[1-e^{\beta/p(\kappa/\gamma-1)^{-p}}].$$

The density function for $-1<p<0$ is

$$f(x) = [1-e^{\beta/p(\kappa/\gamma-1)^{-p}}]^{-1}\beta\kappa x^{-2}(\kappa/x-1)^{-(p+1)}\cdot$$
$$e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}.$$

One is able to cover both positive and negative p with the same expression if one makes the following definition:

$$c = 1, \text{ if } p>0,$$

$$\text{and } c = [1-e^{\beta/p(\kappa/\gamma-1)^{-p}}]^{-1}, \text{ if } -1<p<0.$$

Then one may write in general

$$F(x) = c\{1-e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}\}$$

and

$$f(x) = c\beta\kappa x^{-2}(\kappa/x-1)^{-(p+1)}e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}.$$

It would be desirable to be able to study the various shapes of the curves analytically, in the same manner as is possible for all of the lower curves. So, one may proceed as before to find the first derivative, $f'(x)$:

$$f'(x) = c\beta\kappa e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}x^{-4}(\kappa/x-1)^{-2(p+1)}\cdot$$
$$\{-\beta\kappa-2x(\kappa/x-1)^{p+1}+(p+1)\kappa(\kappa/x-1)^{p}\}.$$

Since all of the first pieces of this derivative are positive, the sign of the slope will depend only upon the last expression which is enclosed in braces. In this expression, the first two parts will always be negative while the third will be positive. It is not possible to solve for the critical point explicitly. The implicit relationship may be stated as follows:

$$x_c = \kappa/2[1-p+\beta(\kappa/x_c-1)^{-p}]$$

where x_c stands for the critical point. Secondly, one may note that there is no reason to believe this solution to be unique. There may be more than one such critical point. Applying the boundaries for x, one sees that x_c is in the range for x from γ to κ for all β such that

$$(\kappa/x_c-1)^p(2\gamma/\kappa+p-1)<\beta<(p+1)(\kappa/x_c-1)^p.$$

It is possible to go further and to obtain the second derivative with respect to x. This has been done. The resulting expression, however, is long and complicated. Given the implicit nature of the result for the critical point(s) and the complexity of the second derivative, it is not immediately apparent that any further analytic information is easily obtained. For this reason, curves based upon hyperlogistic growth and exponential acquisition in the range from $\gamma = 1$ to $\kappa = 101$ were simulated. For these values, the following curve types are possible: J, U, reversed J, twisted J, a left-skewed unimodal curve and another form which has a supremum at γ followed by a minimum which in turn is followed by a maximum and another steeply approached minimum. This last curve has somewhat of the appearance of a contour chair and is reminiscent of pictures of mixtures of densities, perhaps of a negative exponential and a negatively skewed Beta. The breakpoints in parameter values which determine the type of curve cannot, it seems, be as neatly found or discussed as for the lower curves. Too many moving parameters are involved.

The last four-parameter model to be developed is based upon a constant acquisition and hyper-Gompertzian growth. This model also has to be considered in two pieces, for p greater than zero and for p between zero and -1. First, one may tackle the case for positive p. At t equal to zero, one has x equal to γ, a small positive quantity and F(t) equal to zero. As t increases without bound, one has x approaching its maximum κ and F(t) approaching unity. The equation for $\dot{F}(t)$ and F(t) are the same as previously used: $\dot{F}(t) = \beta'\delta'(1-F(t))$ and $F(t) = 1-e^{-\beta'\delta' t}$.

Choosing the hyper-Gompertzian rate equation, one writes

$$\dot{x} = \delta' x[\log(\kappa/x)]^{p+1},$$

a form allowing a much steeper initial growth than even the Gompertz with a very long period of growth at close to maximal size. The solution for this equation after applying the initial conditions to determine the constant of integration is

$$x = k\cdot\exp\{-(\delta' pt+[\log(\kappa/\gamma)]^{-p})^{-1/p}\}.$$

Reversing this solution, one finds t in terms of x as

$$t = 1/(p\delta')\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}.$$

If one transforms the cumulative distribution function, F(t), one derives the size distribution function, F(x) as

$$F(x) = 1-e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}$$

with its density as

$$f(x) = \beta' x^{-1}[\log(\kappa/x)]^{-(p+1)}e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}.$$

If, on the other hand, p is between zero and -1, one is in a situation comparable to that of the hyperlogistic development. Maximal size is attained at a finite time point, which for the hyper-Gompertz is the same as Professor Turner's τ. Again one needs the conditional distribution function for t given that t is less than or equal to τ. The initial conditions are unchanged, but the terminal conditions are that as t approaches τ, x approaches its upper limit, κ, and F(t) approaches unity. As before, one has $\dot{F}(t) = \beta'\delta'(1-F(t))$ and $F(t) = 1-e^{-\beta'\delta' t}$. Then one obtains the conditional distribution as

$$F(t|t\leq\tau') = [1-e^{-\beta'\delta' t}]/[1-e^{-\beta'\delta'\tau'}].$$

The assumption of hyper-Gompertzian growth with a negative value of p leaves the same growth rate equation and solution as for positive p:

$$\dot{x} = \delta' x[\log(\kappa/x)]^{p+1}$$

and

$$x = ke^{-[\delta' p(t-\tau')]^{-1/p}},$$

where the solution is written this time using Professor Turner's τ form for the constant of integration. It is this τ' which is the time point of maximal size. By checking the initial conditions, one sees that τ' is equivalent to $-1/(p\delta')[\log(\kappa/\gamma)]^{-p}$. Reversing the solution, one finds t in terms of x and τ' as

$$t = \tau' + 1/(\delta' p)[\log(\kappa/x)]^{-p}.$$

If one then substitutes $\tau' = -1/(p\delta')[\log(\kappa/\gamma)]^{-p}$ into the equations for t and for $F(t|t\leq\tau')$, one has

$$t = 1/(p\delta)\{[\log(\kappa/x)]^{-p} - [\log(\kappa/\gamma)]^{-p}\}$$

and

$$F(t|t\leq\tau') = \{1-e^{-\beta'\delta' t}\}/\{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}.$$

Transforming t, one obtains, $F(x|t\leq\tau)$:

$$F(x|t\leq\tau') = \{1-e^{-\beta'/p\{[\log(\kappa/x)]^{-p} - [\log(\kappa/\gamma)]^{-p}\}}$$

$$/\{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}.$$

Again, one may examine the condition:

$$t\leq\tau'$$

$$1/(p\delta')\{[\log(\kappa/x)]^{-p} - [\log(\kappa/\gamma)]^{-p}\} \leq -1/(p\delta')[\log(\kappa/\gamma)]^{-p}$$

$$1/(p\delta')[\log(\kappa/x)]^{-p} \leq 0.$$

Now since $\delta' > 0$ and $[\log(\kappa/x)]^{-p} \geq 0$, since $\kappa \geq x$, this implies that $p<0$, which is the case under consideration. So, the condition may be removed here to write

$$F(x) = \{1-e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}\}$$

$$/\{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}$$

and for the density function

$$f(x) = \{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}^{-1}\beta' x^{-1}[\log(\kappa/x)]^{-(p+1)}\cdot$$

$$e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}.$$

Both cases could be covered under the same expression by defining a new parameter, c', as follows:

$$c' = 1, \text{ if } p>0,$$

$$\text{and } c' = \{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}^{-1}, \text{ if } -1<p<0.$$

Then one may write

$$F(x) = c'\{1-e^{-\beta/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}\}$$

and

$$f(x) = c'\beta' x^{-1}[\log(\kappa/x)]^{-(p+1)}\cdot$$

$$e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}.$$

In trying to discuss the analytic properties of these hyper-Gompertz based curves, one has the same difficulties that one encountered in the hyperlogistic based set. It is easy enough to display the first derivative function:

$$f'(x) = c'\beta' x^{-2}[\log(\kappa/x)]^{-2(p+1)} \cdot$$

$$e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}} \cdot$$

$$\{-[\log(\kappa/x)]^{p+1}+(p+1)[\log(\kappa/x)]^{p}-\beta'\}.$$

All of the first factors are positive. The sign of the derivative depends only upon the last factor. Of the terms in that factor, the first and last are always negative and the middle one is always positive. If one sets the derivative equal to zero and solves for x, one obtains, however, only an implicit solution:

$$x_c = \kappa e^{\beta'[\log(\kappa/x_c)]^{-p}-p-1},$$

where x_c stands for the critical point. It is also a bit unclear as to exactly when this critical point is within the range from γ to κ, since the bounds deduced for β' must be dependent upon the unknown critical point. The point, x_c, is in the range from γ to κ for all β' such that

$$[p+1+\log(\gamma/\kappa)][\log(\kappa/x_c)]^{p}<\beta'<(p+1)[\log(\kappa/x_c)]^{p}.$$

The second derivative, $f''(x)$, is easily obtainable, but, as with the hyperlogistic based curve, the expression is too complicated to allow simple conclusions to be drawn, particularly since there is no explicit solution. From simulation of the f(x) function between γ equal to 1 and κ equal to 101, one may see a few hints of the behavior. As p increases, the central portion of a unimodal curve will become narrower and steeper. As β increases for a given value of p, the curve is forced toward the Gompertz-based limiting form. The shape, however, is not independent of the range chosen. Curve forms seen are of the same sorts as those producible by the hyperlogistic based form, although they tend to have sharper rises and falls with closer contact to the x-axis in low regions of the density than do those from the hyperlogistic. Shapes seen include a J, a reversed J, a U, a bulging J (as if a wave form were trying to emerge), a twisted J, a left-skewed unimodal curve and a contour chair type curve.

The next curve to be considered is the most complex of the group and the only five parameter model. It is the size distribution based on the generic growth curve itself. In it, one has the γ and κ of the minimum and maximum size values, the n of the Bertalanffy-Richards, the p of the hyper-Gompertz and hyperlogistic forms. One also retains the β of the acquisition function. As with the two higher curves already discussed, it is necessary to deal with two cases: one for values of p between 0 and -1 and one for values of p greater than 0. One may recall the equations for the exponential distribution of ages (or acquisition times):

$$\dot{F}(t) = \beta\delta(1-F(t))$$

$$\text{and } F(t) = 1-e^{-\beta\delta t}.$$

Assuming this distribution, one may discuss the size distribution using a generic growth model for positive values of the parameter, p. The initial and final conditions for this case are the same as for the other distributions for positive p's already developed: at t equal to zero, x approaches γ and F(t) equals zero; as t increases without bound, x approaches κ and F(t) approaches unity. The growth rate equation for the generic curve is

$$\dot{x} = \delta\kappa^{-n} x^{1-np} (\kappa^n - x^n)^{p+1}$$

and its solution is found to be

$$x = \kappa[1+\{\delta npt+[(\kappa/\gamma)^n-1]^{-p}\}^{-1/p}]^{-1/n}.$$

Now flipping the deterministic growth function, one displays t in terms of x as

$$t = 1/(\delta np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}.$$

Making this transformation in F(t), one gets the cumulative distribution for sizes as

$$F(x) = 1-e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}$$

and notes that its corresponding density function is

$$f(x) = \beta x^{-1}(\kappa/x)^n[(\kappa/x)^n-1]^{-(p+1)} \cdot e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}.$$

If one has a negative p between zero and -1, one is again in the situation of reaching maximal size at a finite time. The initial conditions are unchanged. The final conditions are that, as t approaches a finite time which is equal to $\tau-1/(\delta np)$ where τ is Professor Turner's form for the constant of integration, x approaches κ and F(t) approaches unity. So one needs the conditional cumulative distribution function for t given that t is less than or equal to $\tau-1/(\delta np)$:

$$F(t|t\leq\tau-1/(\delta np)) = [1-F(\tau-1/(\delta np))]^{-1}[1-F(t)]$$

or

$$F(t|t\leq\tau-1/(\delta np)) = [1-e^{-\beta\delta(\tau-1/(\delta np))}]^{-1}[1-e^{-\beta\delta t}].$$

The generic growth rate equation and its solution in terms of τ are:

$$\dot{x} = \delta\kappa^{-n}x^{1-np}(\kappa^n - x^n)^{1+p}$$

and $x = \kappa\{1+[1+\delta np(t-\tau)]^{-1/p}\}^{-1/n}$.

Looking at the initial conditions, one sees that τ is equivalent to $-1/(\delta np)\{[(\kappa/\gamma)^n-1]^{-p}-1\}$. The equation for t in terms of x and τ is

$$t = \tau+1/(\delta np)\{[(\kappa/x)^n-1]^{-p}-1\}$$

or, substituting for τ,

$$t = 1/(\delta np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}.$$

Now, replacing t and τ in the conditional distribution function, one gets:

$$F(x|t\leq\tau-1/(\delta np)) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\cdot \{1-e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}\}.$$

Rewriting the condition $t<\tau-1/(\delta np)$ in terms of x and γ, one sees that it reduces to the condition that p is negative. So for negative p, one may write

$$F(x) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\cdot \{1-e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}\}$$

and, for its density,

$$f(x) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\cdot \{\beta x^{-1}(\kappa/x)^n[(\kappa/x)^n-1]^{-(p+1)}\cdot e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}\}.$$

If one defines a parameter c such that

c = 1, for positive p,

and

$$c = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}, \text{ for } -1<p<0,$$

one may write these two functions for all p as:

$$F(x) = c\{1-e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}\}$$

and

$$f(x) = c\beta x^{-1}(\kappa/x)^n[(\kappa/x)^n-1]^{-(p+1)} \cdot$$
$$e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}.$$

The nature of the size curves based upon generic growth is that they may take on any of the shapes allowed in the size densities already covered. It is still restricted somewhat in shape. The unimodal curves tend to be left-skewed, dropping off more sharply on the right than on the left. This drop, however, is not as steep as that in the hyper-Gompertz and hyperlogistic based curves, but is more rounded. It does not have the freer slope on the right that may be found in the Weibull, as will be discussed soon. An implicit equation for critical values is derivable from the first derivative of the density in the usual fashion. One has

$$f'(x) = c\beta\kappa^n x^{-(n+2)}[(\kappa/x)^n-1]^{-2(p+1)} \cdot$$
$$e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}} \cdot$$
$$\{-(n+1)[(\kappa/x)^n-1]^{p+1}+n(p+1)(\kappa/x)^n[(\kappa/x)^n-1]^p-\beta(\kappa/x)^n\}$$

with the implicit solution as

$$x_c = \kappa(n+1)^{-1/n}\{1-np+\beta[(\kappa/x_c)^n-1]^{-p}\}^{1/n}.$$

By using the range for x, one sees that any given critical point is within the range from γ to κ for all β such that

$$\{n[p+(\gamma/\kappa)^n]-[1-(\gamma/\kappa)^n]\}[(\kappa/x_c)^n-1]^p<\beta<n(p+1)[(\kappa/x_c)^n-1]^p.$$

Such a critical point is often not unique and may be either a maximum or a minimum or a point of inflection. For any particular critical point solution, one may easily determine its type by substitution into the second derivative, $f''(x)$. In general, however, this function is rather complicated and it is not possible to come to any simple analytic conclusions, especially in the light of the lack of explicit solutions for x_c. Perhaps the only clue to the nature of the shape change with variation of the parameters is the one mentioned for the hyper-Gompertz based curve (which is also true for the hyperlogistic based curve): as β increases, the function for x is forced more and more toward one of the lower curves, the Bertalanffy-Richards based curve in this case.

The last curve in the series of size distributions considered in this paper is a limiting curve as κ increases without bound. It is obtainable from all of the three other upper curves. Depending upon which of the other curves one believes it to be a limiting form of, the range of the parameter, α, to be discussed, may be positive only (from the hyper-Gompertz based form) or positive and negative, but not zero (from the generic and the hyperlogistic based forms). The α equal to zero case yields the Pareto size distribution, which has already been developed as a limiting case for all of the lower curves. For this last size distribution, one assumes exponential acquisition and a growth rate proportional to some power of x, $1-\alpha$, which is not necessarily unity. For a more complete discussion of the limits involved and the relationship between the n and p parameters and the new

α parameter, see the review in the fourth chapter. As was the situation with the other upper curves, this derivation must treat separately the curves for positive and for negative α. The equations for constant acquisition are as before:

$$\dot{F}(t) = \beta''\delta''(1-F(t))$$

and

$$F(t) = 1-e^{-\beta''\delta''t}.$$

For positive values of α, the initial conditions are that for t equal to zero, x equals γ and F(t) equals zero. The terminal conditions are that as t increases without bounds, x also increases without bounds and F(t) approaches unity. Now the growth rate equation may be stated as

$$\dot{x} = \delta''x^{1-\alpha} \text{ where } \alpha>0.$$

Solving this equation, one has, after the initial conditions are applied to determine the constant of integration:

$$x = (\alpha\delta''t+\gamma^{\alpha})^{1/\alpha}.$$

Now obtaining t in terms of x in this deterministic growth equation, one writes

$$t = 1/(\alpha\delta'')(x^{\gamma}-\gamma^{\alpha}).$$

When one makes this substitution in the cumulative distribution function for t, one finds

$$F(x) = 1-e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}$$

which one recognizes as a nonstandard representation of the Weibull distribution function. Differentiating, one displays the density function as

$$f(x) = \beta'' x^{\alpha-1} e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}, \ \gamma<x<\infty.$$

If one moves lastly to the development of the curve for a negative α, one is once more in the situation of assuming a finite time, τ'', for the end of growth. Growth, however, is assumed to be unbounded. In other words, the initial conditions are the same as for positive α; the terminal conditions are that as t approaches τ'', x increases without bound and F(t) approaches unity. One may recall the conditional acquisition time distribution for t less than or equal to τ'':

$$F(t|t\leq\tau'') = (1-e^{-\beta''\delta''\tau''})^{-1}(1-e^{-\beta''\delta''t}).$$

Next returning to the growth rate equation, one writes its solution in terms of τ'' as

$$x = [\alpha\delta''(t-\tau'')]^{1/\alpha}.$$

Clearly, it is true from the initial conditions that $\tau'' = -1/(\alpha\delta'')\gamma^{\alpha}$. Making the transformation to t in terms of x and τ'', one writes

$$t = 1/(\alpha\delta'')x^{\alpha}+\tau''.$$

Using the transformations for t and τ'' in the conditional distribution for t, one derives for negative α:

$$F(x|t\leq\tau'') = [1-e^{\beta''/\alpha\gamma^{\alpha}}]^{-1}[1-e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}].$$

One notes that the condition reduces in terms of x and γ simplify to the condition that α is negative which is the premise. Therefore, one may remove the condition and write:

$$F(x) = [1-e^{\beta''/\alpha\gamma^{\alpha}}]^{-1}[1-e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}], \ \alpha<0$$

and

$$f(x) = [1-e^{\beta''/\alpha\gamma^{\alpha}}]^{-1}\beta''x^{\alpha-1}e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}, \ \gamma<x<\infty, \ \alpha<0.$$

If one is willing to define a parameter c'', one may write the two cases for α as:

$$c'' = 1, \text{ for } \alpha>0,$$

$$\text{and } c'' = [1-e^{\beta''/\alpha\gamma^{\alpha}}]^{-1}, \text{ for } \alpha<0.$$

Then for the Weibull form one has derived, one has

$$F(x) = c''[1-e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}]$$

and

$$f(x) = c''\beta''x^{\alpha-1}e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}, \ \gamma<x<\infty.$$

The Weibull size distribution is the only one of the upper curves whose analytic properties are as easily discussed as those of the lower curves. First one may obtain the first derivative of the density function with respect to x:

$$f'(x) = c''\beta''x^{\alpha-2}e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}\{(\alpha-1)-\beta''x^{\alpha}\}.$$

It is obvious that the sign of the last factor determines the sign of the derivative or the slope along the curve. The second term in this factor will always be negative since β'' is positive as is x. If α is less than or equal to 1, the first term in the factor will also be negative. So for α less than or equal to 1, the density will be J-shaped with supremum at x equal to γ. The equation may be solved explicitly for a unique critical point which is

$$x_c = \{(\alpha-1)/\beta''\}^{1/\alpha}.$$

Using the range of x from γ to ∞, one sees that this critical point falls within the range for all β'' such that

$$0<\beta''<(\alpha-1)\gamma^{-\alpha}.$$

Secondly, one may obtain the second derivative of the density with respect to x as

$$f''(x) = c''\beta''x^{\alpha-3}e^{-\beta/\alpha(x^{\alpha}-\gamma^{\alpha})}\{(\alpha-1)(\alpha-2)-3\beta''\alpha x^{\alpha}$$
$$+3\beta''x^{\alpha}+(\beta'')^2x^{2\alpha}\}.$$

Again the first factors are all positive, so the sign of the derivative depends only upon the last factor. One may call this factor H. Now one has:

$$H = (\alpha-1)(\alpha-2)-3\beta''\alpha x^{\alpha}+3\beta''x^{\alpha}+(\beta'')^2x^{2\alpha}$$

Evaluating H at the critical point, x_c, one finally gets:

$$H\Big|_{x_c=\{(\alpha-1)/\beta\}^{1/\alpha}} = -\alpha(\alpha-1)$$

Summarizing, one may state that H, for α greater than 1, is always negative and that for α greater than 1 and β'' between 0 and $(\alpha-1)\gamma^{-\alpha}$ a maximum will exist in the range from γ to infinity. If β'' is greater than or equal to $(\alpha-1)\gamma^{-\alpha}$ or α is less than or equal to unity, no maximum or minimum will exist in the range. The Weibull size density never has a true minimum. It is either J-shaped or unimodal. Its skewness, unlike the other curves in this set, may be either to the left, or to the right.

This completes the assembly of size distributions based upon an exponential distribution of ages (or acquisition times) and the growth curves which are members of the Turner family. In the following chapter are presented a couple of characterizations of these size distributions through development and discussion of their moments and of their intensity functions. Then in the seventh chapter, the problem of estimation is approached. There is a discussion of some standard techniques, one not-so-standard technique and a method of obtaining starting values for use in any of these other procedures. The eighth chapter considers a few interesting examples from "real" data. For a few selected illustrations of the types of size distributions which can be produced by these family members, one may study the following seven graphs before continuing with Chapter VI.

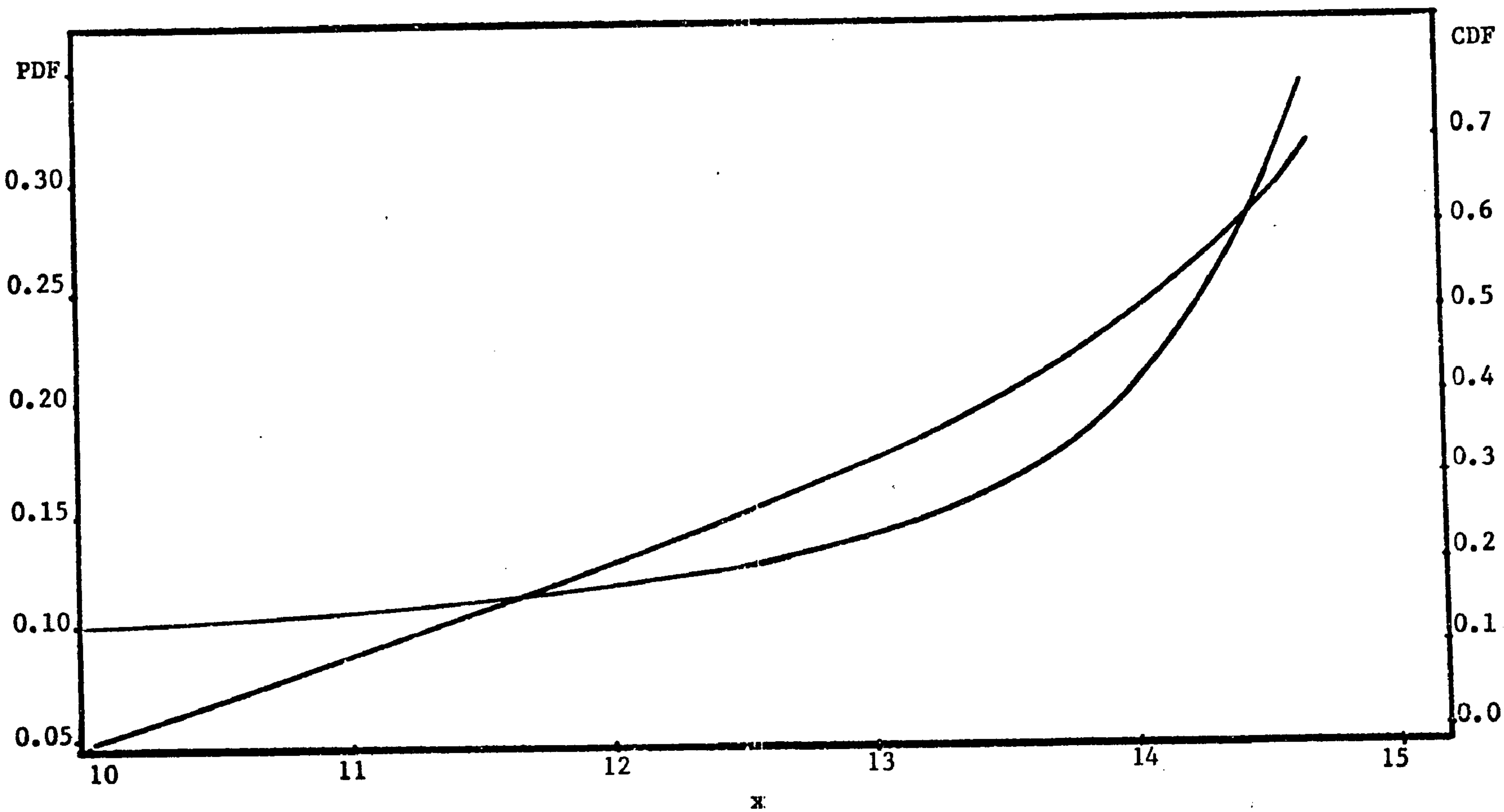


Fig. 1 Exponential-Gompertzian distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 10$, $\kappa = 15$ and $\beta' = 0.4$.

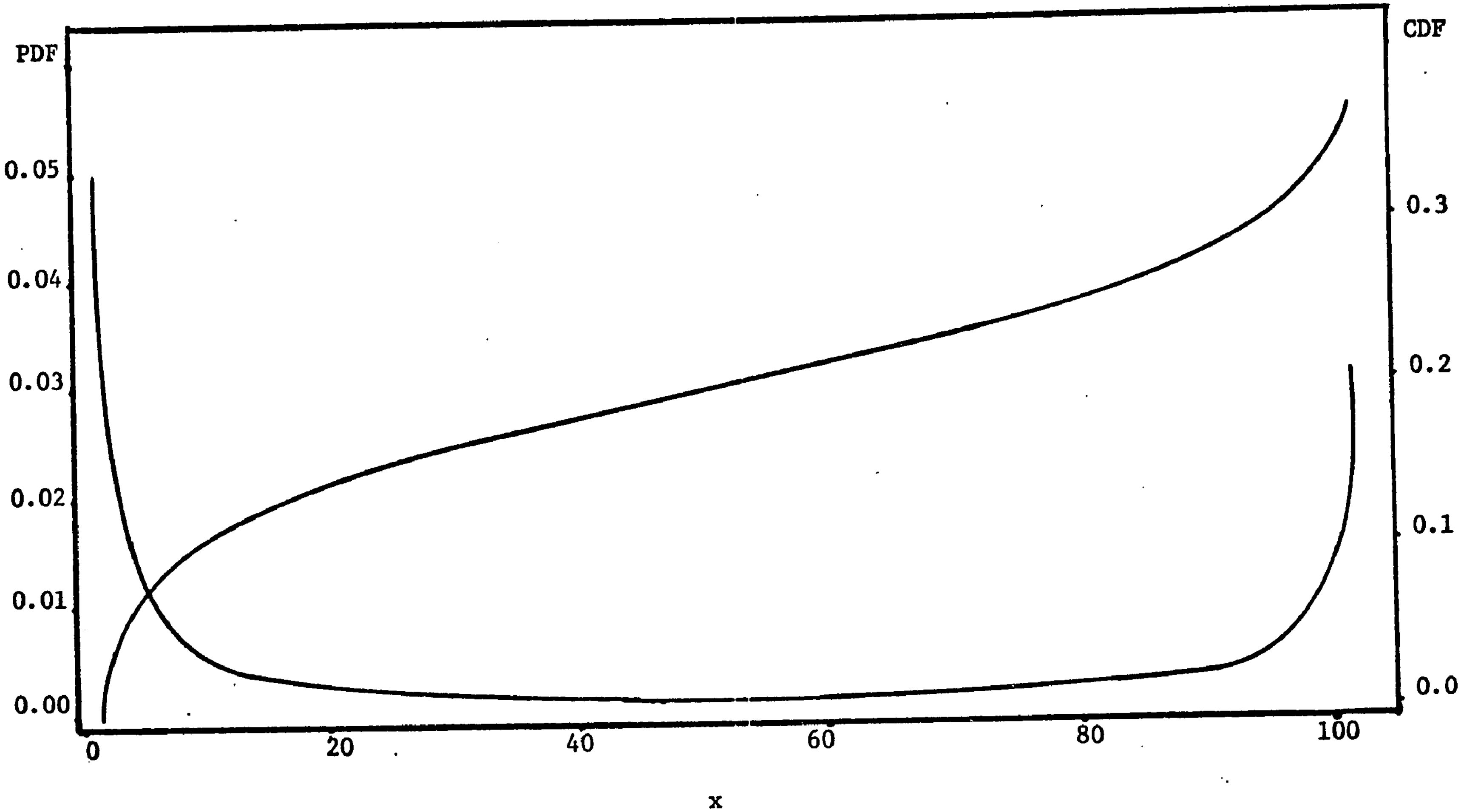


Fig. 2 Exponential-logistic distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 1$, $\kappa = 101$ and $\beta = 0.05$.

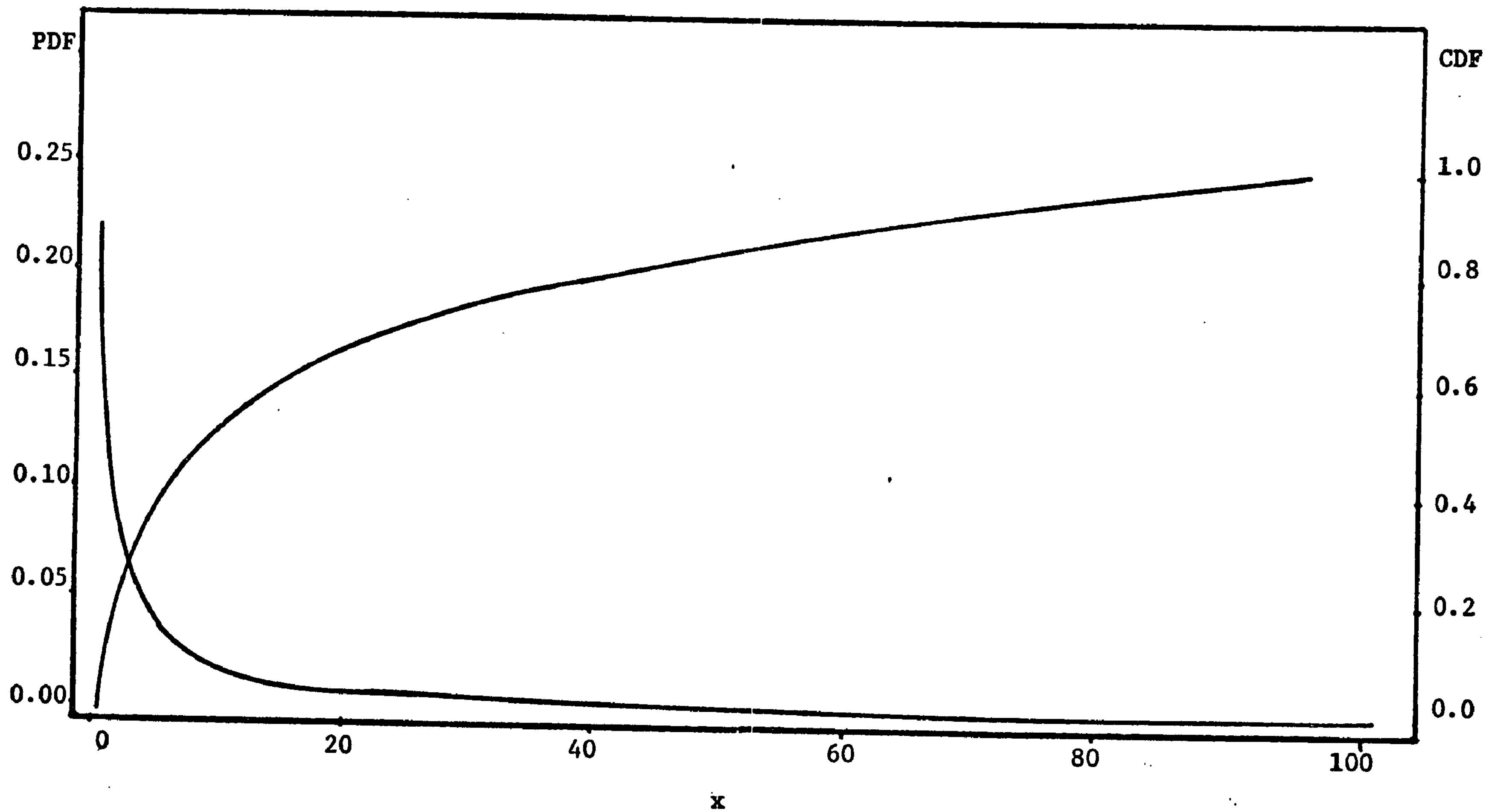


Fig. 3 Exponential-Gompertzian distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 1$, $\kappa = 101$ and $\beta' = 1$.

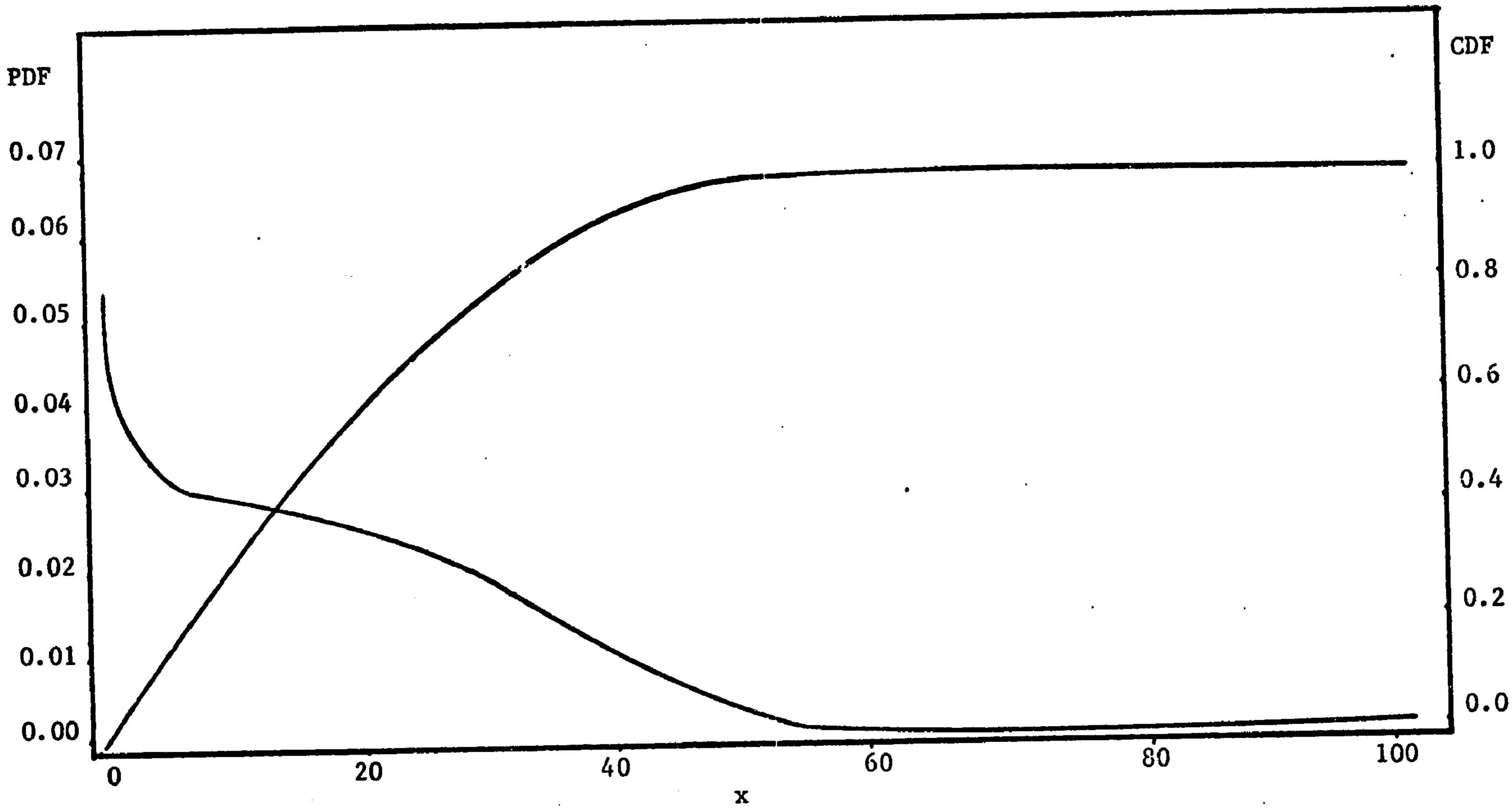


Fig. 4 Exponential-hyper-Gompertzian distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 1$, $\kappa = 101$, $p = 2$ and $\beta' = 5$.

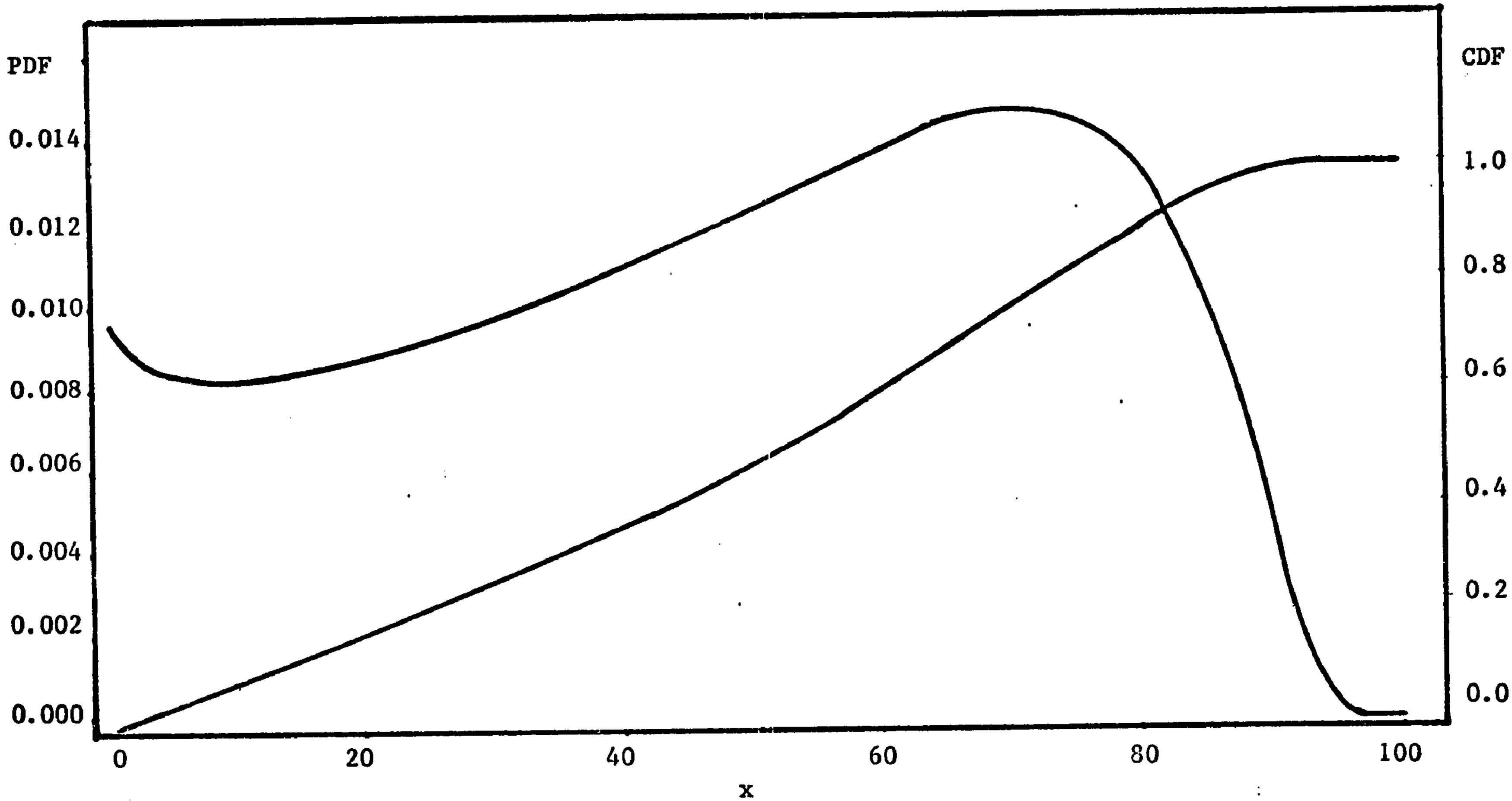


Fig. 5 Exponential-hyperlogistic distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 1$, $\kappa = 101$, $p = 0.9$ and $\beta = 0.6$.

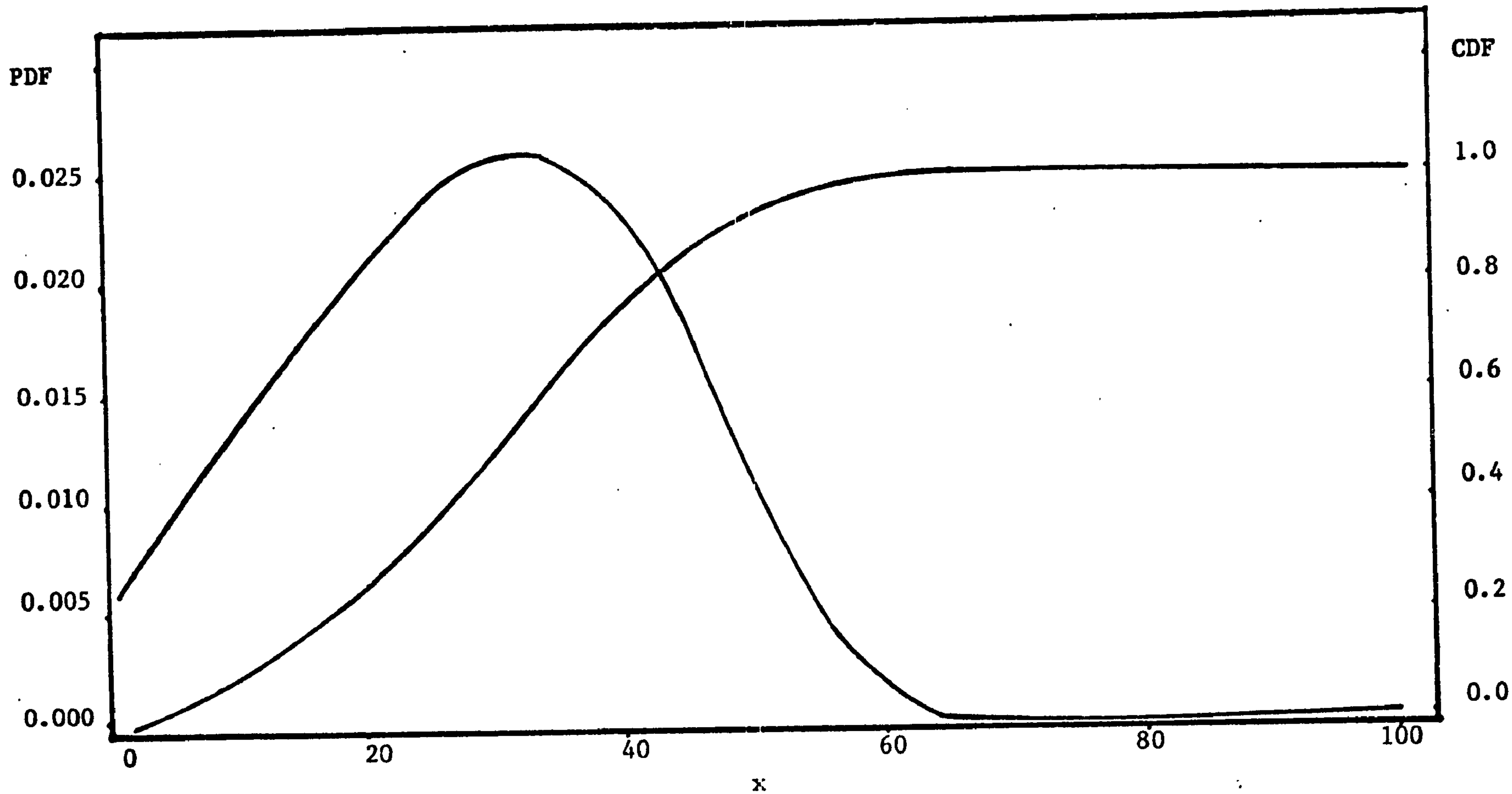


Fig. 6 Exponential-generic distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 1$, $\kappa = 101$, $n = 0.5$, $p = 2$ and $\beta = 0.5$.

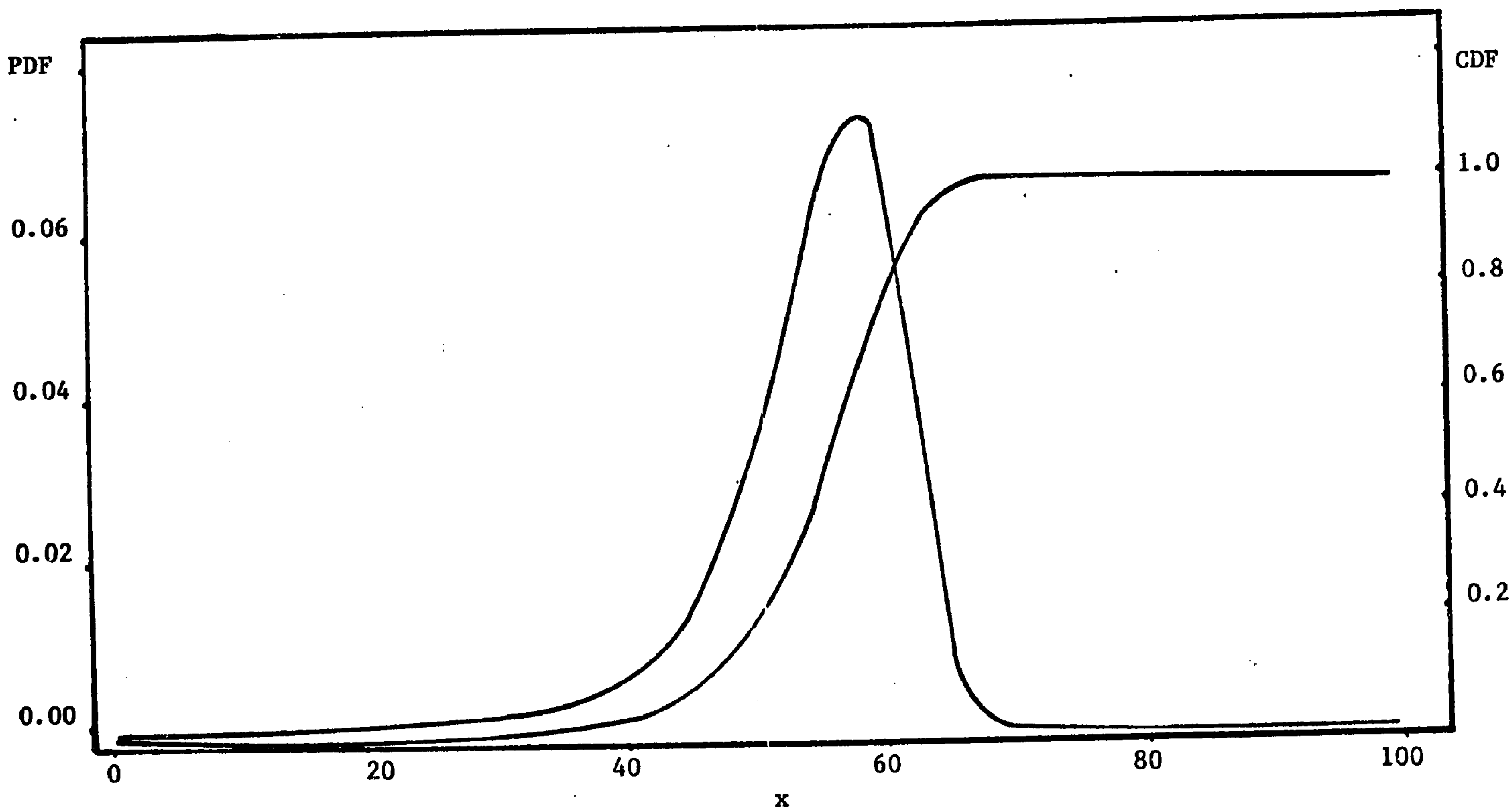


Fig. 7 Exponential-hyperlogistic distribution function (CDF) and probability density function (PDF) with parameter values $\gamma = 1$, $\kappa = 101$, $p = 5$ and $\beta = 1$.

VI. TWO CHARACTERIZATIONS OF THE CURVES

A. The Moments

Since the time of Karl Pearson, it has become customary to describe a distribution or family of distributions in terms of certain mathematical expectations, the moments, and of their ratios. The rth such moment is designated by $E(x^r)$. By constructing ratios of these moments, one would hope to obtain information on the degree of skewness and of kurtosis for different values of the curve's parameters. For the family of size distributions developed in this dissertation, the moments, although obtainable, do not prove to be very useful. In general, for cases of finite upper limit, κ, they do exist, but in infinite series form. In the well-known cases of the Pareto and the Weibull, the moments will sometimes be infinite.

First one may work through the moments for the Pareto distribution:

$$E(x^r) = \beta/\gamma \int_{\gamma}^{\infty} x^r (x/\gamma)^{-(\beta+1)} dx$$

$$E(x^r) = \gamma^r \{\beta/\gamma \int_{\gamma}^{\infty} (x/\gamma)^{r-\beta-1}\} dx.$$

Now there are three situations to be considered:

1. $r-\beta=0$
2. $r-\beta<0$
3. $r-\beta>0$.

If $(r-\beta)=0$, one has

$$E(x^r) = \gamma^r \beta [\lim_{x\to\infty} \log x - \log \gamma] \to \infty.$$

If $(r-\beta)<0$, one has

$$E(x^r) = (\beta\gamma^r)/(r-\beta)\{\lim_{x\to\infty} (x/\gamma)^{r-\beta}-1\} = (-\beta\gamma^r)/(r-\beta).$$

If $(r-\beta)>0$, one has

$$E(x^r) = (\beta\gamma^r)/(r-\beta)\{\lim_{x\to\infty} (x/\gamma)^{r-\beta}-1\} \to \infty.$$

This means that as long as the order of the moment (r) is less than the value of β, the moment will exist and will be equal to $(\beta\gamma^r)/(\beta-r)$. Once the moment order becomes equal to or greater than β, that moment and all of the moments higher than it, will be infinite. So, for instance, if β is less than or equal to unity, no moments will be finite. For β less than or equal to 2 and greater than 1, only the mean, $E(x)$, will be finite.

The exponential-logistic size distribution has a finite range from γ to κ and its moments always exist and are finite. They may be written as

$$E(x^r) = \beta\kappa(\kappa/\gamma-1)^{-\beta} \int_\gamma^\kappa x^{r-2}(\kappa/x-1)^{\beta-1}.$$

Through a proper choice of transformations and expansions, it is possible to find a series solution for the moments as

$$E(x^r) = \beta\kappa^{r-\beta}\gamma^\beta \frac{\Gamma(\beta)}{\Gamma(\beta+1)} \left\{\frac{\Gamma(\beta+1)}{\Gamma(1-r+\beta)\Gamma(\beta)} \cdot \sum_{j=0}^{\infty} \frac{\Gamma(j+1-r-\beta)\Gamma(j+\beta)}{\Gamma(j+\beta+1)} \frac{[1-\gamma/\kappa]^j}{j!}\right\}$$

which may be recognized as an incomplete Beta integral or as a Gaussian hypergeometric series,

$$E(x^r) = \kappa^{r-\beta}\gamma^{\beta}F(\beta,1-r+\beta;\beta+1;(1-\gamma/\kappa)).$$

If β is an integer, the series will reduce to a polynomial in $(1-\gamma/\kappa)$ for all r greater than or equal to $\beta+1$. The degree of that polynomial will be $1-r+\beta$.

The exponential-Gompertzian size distribution also has a finite range from γ to κ and finite moments. They may be displayed as

$$E(x^r) = [\log(\kappa/\gamma)]^{-\beta'}\beta'\int_{\gamma}^{\kappa} x^{r-1}[\log(\kappa/x)]^{\beta'-1}dx.$$

Transforming the integral, one finds that the solution in series form is

$$E(x^r) = \beta'\kappa^r \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}\,\frac{[r\,\log(\kappa/\gamma)]^i}{i+\beta'}\,.$$

This series may be recognized as an expansion of the Pearson incomplete Gamma function; therefore, an equivalent form for the solution is

$$E(x^r) = [\log(\kappa/\gamma)]^{-\beta'}\kappa^r r^{-\beta'}\Gamma(\beta'+1)I(\{r\,\log(\kappa/\gamma)\}/\sqrt{\beta'},\beta'-1).$$

The moments clearly exist for all β' greater than zero. The series reduces to a finite Poisson sum for integral β'. As a confluent hypergeometric series, the notation changes to

$$E(x^r) = \kappa^r M(\beta';\beta'+1;-r\,\log(\kappa/\gamma))$$

or

$$E(x^r) = \kappa^r(\kappa/\gamma)^{-r}M(1;\beta'+1;r\,\log(\kappa/\gamma))$$

$$= \gamma^r M(1;\beta'+1;r\,\log(\kappa/\gamma)).$$

The moments for the exponential-Bertalanffy-Richards size distribution are closely related to those of the exponential-logistic. The moment equation is

$$E(x^r) = \beta[(\kappa/\gamma)^n-1]^{-\beta/n} \int_\gamma^\kappa x^{r-1}(\kappa/x)^n[(\kappa/x)^n-1]^{\beta/n-1}dx.$$

This may be transformed to an incomplete Beta integral form:

$$E(x^r) = \beta/n\kappa^r[(\kappa/\gamma)^n-1]^{-\beta/n} \int_0^{1-(\gamma/\kappa)^n} w^{\beta/n-1}(1-w)^{(r-\beta)/n-1}dw.$$

The solution may be written as a series in the following way:

$$E(x^r) = \beta/n\kappa^r(\gamma/\kappa)^\beta \frac{\Gamma(\beta/n)}{\Gamma(\beta/n+1)} \left\{\frac{\Gamma(\beta/n+1)}{\Gamma(\beta/n)\Gamma(1-(r-\beta)/n)} \cdot \right.$$

$$\left. \sum_{j=0}^{\infty} \frac{\Gamma(j+1-(r-\beta)/n)\Gamma(\beta/n+j)}{\Gamma(\beta/n+j+1)} \frac{[1-(\gamma/\kappa)^n]^j}{j!}\right\}.$$

In a Gaussian hypergeometric series notation, the solution becomes

$$E(x^r) = \gamma^\beta\kappa^{r-\beta}F(\beta/n,1-(r-\beta)/n;\beta/n+1;[1-(\gamma/\kappa)^n]).$$

If $1-(r-\beta)/n$ is zero or a negative integer, the series terminates as a polynomial in $[1-(\gamma/\kappa)^n]$.

For all of the upper curves, in which the distribution will take two forms according to the sign of p or α, the moments also must be found separately for the two situations.

For the exponential-hyperlogistic size distributions, if p is negative, one has

$$E(x^r) = [1-e^{\beta/p(\kappa/\gamma-1)^{-p}}]^{-1}\beta\kappa \int_\gamma^\kappa x^{r-2}(\kappa/x-1)^{-(p+1)} \cdot$$

$$e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}dx$$

and, if p is positive, one has

$$E(x^r) = \beta\kappa \int_\gamma^\kappa x^{r-2}(\kappa/x-1)^{-(p+1)}e^{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]}dx.$$

These two functions may be evaluated by careful selection of transformations and expansions. The moments for negative values of p will be the more closely related to those of the exponential-logistic size distributions, as should be expected from the restriction of shapes for negative p distributions. For p between -1 and 0, the integral may be written as a double series:

$$E(x^r) = [1-e^{\beta/p(\kappa/\gamma-1)^{-p}}]^{-1}\beta\kappa^r \sum_{i=0}^{\infty} \frac{(-\beta/p)^i(1-\gamma/\kappa)^{-p(i+1)}}{i![-p(i+1)]} \cdot$$

$$\sum_{j=0}^{\infty} \frac{\Gamma[-p(i+1)+1]\Gamma[1-r-p(i+1)+j]\Gamma[j-p(i+1)](1-\gamma/\kappa)^j}{\Gamma[-p(i+1)]\Gamma[1-r-p(i+1)]\Gamma[j-p(i+1)+1]j!} .$$

The second series is a Gaussian hypergeometric series:

$$E(x^r) = [1-e^{\beta/p(\kappa/\gamma-1)^{-p}}]^{-1}\kappa^r \sum_{i=0}^{\infty} \frac{(-\beta/p)^{i+1}(1-\gamma/\kappa)^{-p(i+1)}}{(i+1)!} \cdot$$

$$F[-p(i+1),1-r-p(i+1);-p(i+1)+1;(1-\gamma/\kappa)].$$

So it is possible to represent the rth moment of the distribution for negative values of p as an infinite sum of Gaussian hypergeometric series. For p greater than zero, the integral may be found as the sum of the upper ends of incomplete Gamma integrals, which always exist:

$$E(x^r) = e^{\beta/p(\kappa/\gamma-1)^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{\Gamma(1-r)(\beta/p)^{i/p}}{\Gamma(1-r-i)i!} \cdot$$

$$\int_{\beta/p(\kappa/\gamma-1)^{-p}}^{\infty} u^{-i/p}e^{-u}du$$

This last integral may be replaced by $\Gamma[1-i/p,\ \beta/p(\kappa/\gamma-1)^{-p}]$. Then the expectation may be written as

$$E(x^r) = e^{\beta/p(\kappa/\gamma-1)^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{\Gamma(1-r)(\beta/p)^{i/p}}{\Gamma(1-r-i)i!} \cdot$$

$$\Gamma[1-i/p,\beta/p(\kappa/\gamma-1)^{-p}]$$

or

$$E(x^r) = e^{\beta/p(\kappa/\gamma-1)^{-p}} \kappa^r \sum_{i=0}^{\infty} \frac{\Gamma(r+i)}{\Gamma(r)i!} [-(\beta/p)^{1/p}]^i \cdot$$

$$\Gamma[1-i/p, \beta/p(\kappa/\gamma-1)^{-p}].$$

The exponential-hyper-Gompertzian size distribution has for its moments

$$E(x^r) = \{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}^{-1} \beta' \int_{\gamma}^{\kappa} x^{r-1}[\log(\kappa/x)]^{-(p+1)} \cdot$$

$$e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}} dx$$

for negative values of p and

$$E(x^r) = \beta' \int_{\gamma}^{\kappa} x^{r-1}[\log(\kappa/x)]^{-(p+1)} \cdot$$

$$e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}} dx$$

for positive values of p. Again it is the moment form for negative p which has more kinship to the limiting case as p approaches zero, the exponential-Gompertzian distribution. If p is between zero and -1, the expectation, as a double series, is

$$E(x^r) = \{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}^{-1} \beta' \kappa^r \sum_{i=0}^{\infty} \frac{(-\beta'/p)^i}{i!} \cdot$$

$$[\log(\kappa/\gamma)]^{-p(i+1)} \sum_{j=0}^{\infty} \frac{\Gamma[j-p(i+1)]}{\Gamma[j-p(i+1)+1]} \frac{[-r \log(\kappa/\gamma)]^j}{j!}$$

In the notation for Pearson's incomplete Gamma function I(.,.), it is written as

$$E(x^r) = \{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}^{-1} \beta' \kappa^r \sum_{i=0}^{\infty} \frac{(-\beta'/p)^i}{i!} \cdot$$

$$r^{p(i+1)}\Gamma[-p(i+1)]I[\{r \log(\kappa/\gamma)\}/\sqrt{-p(i+1)},-p(i+1)-1]$$

or as a sum of confluent hypergeometric series in either of two forms:

$$E(x^r) = \{1-e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\}^{-1}\beta'\kappa^r \sum_{i=0}^{\infty} \frac{(-\beta'/p)^i}{i!} \cdot$$

$$[-p(i+1)]^{-1}[\log(\kappa/\gamma)]^{-p(i+1)} \cdot$$

$$M[-p(i+1),1-p(i+1),-r \log(\kappa/\gamma)]$$

or

$$E(x^r) = \{1-e^{\beta'/p \log(\kappa/\gamma)^{-p}}\}^{-1}\beta'\kappa^r \sum_{i=0}^{\infty} \frac{(-\beta'/p)^i}{i!} \cdot$$

$$[-p(i+1)]^{-1}[\log(\kappa/\gamma)]^{-p(i+1)} \cdot$$

$$e^{p(i+1)}M[1,1-p(i+1),r \log(\kappa/\gamma)].$$

For positive values of p, one again finds the moments as an infinite sum of the upper ends of incomplete Gamma distributions, which always exist. In terms of the incomplete Gamma integral, the function is

$$E(x^r) = e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{(-r)^i}{i!} (\beta'/p)^{i/p} \cdot$$

$$\int_{\beta'/p[\log(\kappa/\gamma)]^{-p}}^{\infty} u^{-i/p}e^{-u}du \quad .$$

Collapsing the integral notationally to $\Gamma(1-i/p,\beta'/p[\log(\kappa/\gamma)]^{-p})$, one may write the expression for the moments as

$$E(x^r) = e^{\beta'/p[\log(\kappa/\gamma)]^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{[-r(\beta'/p)^{1/p}]^i}{i!} \cdot$$

$$\Gamma(1-i/p,\beta'/p[\log(\kappa/\gamma)]^{-p})$$

For the most complicated of the upper curves, the exponential-generic growth size distribution, the results are much the same as for the exponential-hyperlogistic distribution. For negative p, the moment equation is

$$E(x^r) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\beta \int_\gamma^\kappa x^{r-1}(\kappa/x)^n \cdot$$

$$[(\kappa/x)^n-1]^{-(p+1)} e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}dx.$$

For positive p, the moment equation is

$$E(x^r) = \beta \int_\gamma^\kappa x^{r-1}(\kappa/x)^n[(\kappa/x)^n-1]^{-(p+1)} \cdot$$

$$e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}dx.$$

Through several transformations and two carefully chosen expansions, one sees that the double series representation for the moments in the negative p situation is

$$E(x^r) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\beta/n\kappa^r \sum_{i=0}^{\infty} \frac{(-\beta/(np))^i}{i!} \cdot$$

$$\frac{\Gamma[-p(i+1)][1-(\gamma/\kappa)^n]^{-p(i+1)}}{\Gamma[-p(i+1)+1]} \left\{\frac{\Gamma[-p(i+1)+1]}{\Gamma[1-r/n-p(i+1)]\Gamma[-p(i+1)]} \cdot\right.$$

$$\left.\sum_{j=0}^{\infty} \frac{\Gamma[1-r/n-p(i+1)+j]\Gamma[j-p(i+1)][1-(\gamma/\kappa)^n]^j}{\Gamma[j-p(i+1)+1]j!}\right\}.$$

Replacing the last summation with the notation for the Gaussian hypergeometric series, one has

$$E(x^r) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\kappa^r \sum_{i=0}^{\infty} \frac{\{-\beta/(np)[1-(\gamma/\kappa)^n]^{-p}\}^{i+1}}{(i+1)!} \cdot$$

$$F[-p(i+1),1-r/n-p(i+1);-p(i+1)+1;1-(\gamma/\kappa)^n]$$

Now if p is greater than zero, one obtains, as for the two upper curves already discussed, a sum of the upper ends of incomplete Gamma integrals. One has

$$E(x^r) = e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{\Gamma(1-r/n)(\beta/(np))^{i/p}}{\Gamma(1-r/n-i)i!} \cdot$$

$$\int_{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}^{\infty} u^{-i/p}e^{-u}\,du.$$

This equation can be rewritten either as

$$E(x^r) = e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{\Gamma(1-r/n)(\beta/(np))^{i/p}}{\Gamma(1-r/n-i)i!} \cdot$$

$$\Gamma(1-i/p,\beta/(np)[(\kappa/\gamma)^n-1]^{-p})$$

or

$$E(x^r) = e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\kappa^r \sum_{i=0}^{\infty} \frac{\Gamma(r/n+i)[-(\beta/(np))^{1/p}]^i}{\Gamma(r/n)i!} \cdot$$

$$\Gamma(1-i/p,\beta/(np)[(\kappa/\gamma)^n-1]^{-p}).$$

The last of the upper curves, the Weibull size distribution, is like the Pareto in having an infinite upper bound for growth and, like the Pareto, some of its moments may not be finite. If its parameter, α, is less than zero, $E(x^r)$ will exist for all r less than $-\alpha$. For r greater than or equal to $-\alpha$, $E(x^r)$ will be infinite. The equation for the rth moment is

$$E(x^r) = (1-e^{\beta/\alpha\gamma^\alpha})^{-1}\beta'' \int_{\gamma}^{\infty} x^{r+\alpha-1}e^{-\beta''/\alpha(x^\alpha-\gamma^\alpha)}dx,$$

if α is negative, and is

$$E(x^r) = \beta'' \int_{\gamma}^{\infty} x^{r+\alpha-1}e^{-\beta''/\alpha(x^\alpha-\gamma^\alpha)}dx,$$

if α is positive. For negative α, the series form for the moments for r less than $-\alpha$ is

$$E(x^r) = (-\alpha)^{-1}\beta''\gamma^{\alpha+r}e^{\beta''/\alpha\gamma^\alpha}(1-e^{\beta''/\alpha\gamma^\alpha})^{-1} \cdot$$

$$\sum_{i=0}^{\infty} \frac{\Gamma(i+1+r/\alpha)}{\Gamma(i+2+r/\alpha)} \frac{(-\beta''/\alpha\gamma^{\alpha})^{i}}{i!} .$$

In the notation of the confluent hypergeometric series, this solution may be recorded as

$$E(x^r) = (-\alpha)^{-1}\beta''\gamma^{\alpha+r}e^{\beta''/\alpha\gamma^{\alpha}}(1-e^{\beta''/\alpha\gamma^{\alpha}})^{-1} \cdot$$

$$(r/n+1)^{-1}M(r/\alpha+1;r/\alpha+2;-\beta''/\alpha\gamma^{\alpha}).$$

If the incomplete Gamma form is preferred, it is

$$E(x^r) = -(1-e^{\beta''/\alpha\gamma^{\alpha}})^{-1}(\beta''/\alpha)^{-r/\alpha}e^{\beta''/\alpha\gamma^{\alpha}} \cdot$$

$$\Gamma(r/\alpha+1)I(\{\beta''/\alpha\gamma^{\alpha}\}/\sqrt{1+r/\alpha},r/\alpha).$$

The moments for positive values of α are a multiple of a single upper end of an incomplete Gamma function. As an integral, the expectations are

$$E(x^r) = (\alpha/\beta'')^{r/\alpha}e^{\beta''/\alpha\gamma^{\alpha}} \int_{\beta/\alpha\gamma^{\alpha}}^{\infty} z^{(r+\alpha)/\alpha-1}e^{-z}dz.$$

Replacing the integral notationally by $\Gamma(1+r/\alpha,\beta''/\alpha\gamma^{\alpha})$, one has

$$E(x^r) = (\alpha/\beta'')^{r/\alpha}e^{\beta''/\alpha\gamma^{\alpha}}\Gamma(1+r/\alpha,\beta''/\alpha\gamma^{\alpha}).$$

These moments exist for all r with α greater than zero.

This completes the list of moments for this family of size distributions. In general, one sees that the moments for the upper curves with p between zero and -1 (excluding the Weibull) bear more resemblance to those of the lower curves since they are sums of elements which are of the same form as the moments of the corresponding lower curves. The lower curves other than the Pareto all have moments which are able to be written in terms of a single summation. This is also true of the

Weibull. The upper curves other than the Weibull have their moments in terms of a double summation. For positive values of p, they are a weighted sum of upper ends of incomplete Gamma integrals.

B. The Intensity Functions

Another method of characterizing distributions used especially in the realm of reliability theory and in the theory of extremes (Gumbel, 1958) is to discuss the nature of the intensity or hazard functions derivable from the distributions. This would be more easily done in distributions whose cumulative distribution functions exist in closed form, as do those of this size distribution family. The intensity function is defined to be the relative proportion of items falling between x and x+dx out of those known to be larger than some specified value x:

$$I(x) = \frac{f(x)dx}{1-F(x)} .$$

I(x) is not, in general, a density function and, since the integral of I(x) over the range is often infinite, cannot be normalized as one.

For the Pareto case, one obtains

$$I(x) = \frac{\beta/\gamma(x/\gamma)^{-(\beta+1)}}{1-[1-(x/\gamma)^{-\beta}]}, \quad \gamma<x<\infty,$$

which reduces to $I(x) = \beta x^{-1}$. Its first and second derivatives are

$$\frac{dI(x)}{dx} = -\beta x^{-2}$$

and $$\frac{d^2I(x)}{dx^2} = 2\beta x^{-3}.$$

The slope is negative for all $x>\gamma>0$; therefore I(x) is a J-shaped curve, decreasing from $\beta\gamma^{-1}$ toward 0 as x increases without bound.

The second intensity function is that from the exponential-logistic size distribution. Here one finds that, after simplification,

$$I(x) = \beta\kappa x^{-2}(\kappa/x-1)^{-1}, \quad \gamma<x<\kappa.$$

Its first two derivatives with respect to x are

$$\frac{dI(x)}{dx} = -\beta\kappa x^{-2}(\kappa-x)^{-1}+\beta\kappa x^{-1}(\kappa-x)^{-2}$$

and

$$\frac{d^2I(x)}{dx^2} = 2\beta\kappa x^{-3}(\kappa-x)^{-1}-2\beta\kappa x^{-2}(\kappa-x)^{-2}+2\beta\kappa x^{-1}(\kappa-x)^{-3}.$$

Setting the first derivative equal to zero and solving for x, one finds a unique solution at $x = \kappa/2$. Evaluating the second derivative at that point, one sees that it is equal to $2^5\beta\kappa^{-3}$, which is always positive. So I(x) has a minimum at $\kappa/2$. This will be in the range from γ to κ if γ is less than $\kappa/2$ and the function I(x) will be U-shaped. Otherwise, the function will be a rising curve (reversed J-shape) from γ to κ. Checking the second derivative for solutions, one sees that it is equal to zero only at a pair of imaginary points. Therefore, I(x) has no real points of inflection.

For the exponential-Gompertzian size distribution, the intensity function simplifies to

$$I(x) = \beta' x^{-1}[\log(\kappa/x)]^{-1}.$$

Its first and second derivatives are, respectively,

$$\frac{dI(x)}{dx} = -\beta' x^{-2}[\log(\kappa/x)]^{-1}+\beta' x^{-2}[\log(\kappa/x)]^{-2}$$

and $$\frac{d^2I(x)}{dx^2} = 2\beta' x^{-3}[\log(\kappa/x)]^{-1}-3\beta' x^{-3}[\log(\kappa/x)]^{-2}$$

$$+2\beta' x^{-3}[\log(\kappa/x)]^{-3}$$

Setting the first derivative equal to zero, one finds a critical point at $x = \kappa e^{-1}$. Evaluating the second derivative there, one sees that

$$\left.\frac{d^2 I(x)}{dx^2}\right|_{x=\kappa e^{-1}} = \beta' e^3 \kappa^{-3},$$

which is always positive for β' and κ positive. So $I(x)$ has a minimum at $x = \kappa e^{-1}$. This point is in the range from γ to κ if κ is greater than γe. Otherwise, the point would fall to the left of γ and the curve would be a rising curve, concave upward, from γ to κ, rather than U-shaped. $I(x)$ has no real points of inflection, as may be seen by setting the second derivative equal to zero.

The exponential-Bertalanffy-Richards size distribution has as its intensity function,

$$I(x) = \beta[x-\kappa^{-n}x^{n+1}]^{-1}.$$

Its first derivative with respect to x is

$$\frac{dI(x)}{dx} = -\beta[x-\kappa^{-n}x^{n+1}]^{-2}[1-(n+1)\kappa^{-n}x^{n}]$$

which, when equated to zero and solved, yields a critical point at $x = \kappa(n+1)^{-1/n}$. The second derivative is easily found to be

$$\frac{d^2 I(x)}{dx^2} = \frac{\beta n(n+1)x^{n-1}[x-\kappa^{-n}x^{n+1}]+2\beta\kappa^{n}[1-(n+1)\kappa^{-n}x^{n}]^2}{\kappa^{n}[x-\kappa^{-n}x^{n+1}]^3}.$$

If the second derivative is evaluated at the critical point, the result is seen to be $\beta n^{-1}\kappa^{-3}(n+1)^{3/n+2}$ which is positive for all β, n and κ greater than zero. Thus one has that $I(x)$ has a minimum at $x = \kappa(n+1)^{-1/n}$. This minimum falls within the range from γ to κ if κ is greater than $(n+1)^{1/n}\gamma$. By studying the second derivative, one sees that a rare occurrence would be the finding of a point of inflection in the range from γ to κ, which may happen if n is greater than 7. Such a point would be found at

$$x = \kappa \left\{ \frac{-1 \pm \sqrt{(n-4)^2 - 8(n+2)/(n+1)}}{2(n+2)} \right\}^{1/n} .$$

The intensity functions for the upper curves (excluding the Weibull) are more complicated to describe than are those for the lower curves. The properties for negative values of p are not easily found at all. For the hyperlogistic based curve, the intensity function, if p is positive, is

$$I(x) = \beta\kappa x^{-2}[\kappa/x-1]^{-(p+1)} .$$

The first derivative with respect to x and its solution when set equal to zero are

$$\frac{dI(x)}{dx} = \beta\kappa x^{-4}(\kappa/x-1)^{-(p+2)}[2x+\kappa(p-1)]$$

and

$$x = \frac{\kappa(1-p)}{2} .$$

Clearly this point cannot be in the range for p greater than or equal to 1. For p between 0 and 1, this point will only be in the range from γ to κ if κ is greater than 2γ and is less than $1-2\gamma/\kappa$. Differentiating again with respect to x, one produces the second derivative as

$$\frac{d^2I(x)}{dx^2} = \beta\kappa x^{-6}(\kappa/x-1)^{-(p+3)} \cdot$$

$$\{6(\kappa-x)^2 - 6\kappa(p+1)(\kappa-x) + (p+1)(p+2)\kappa^2\}.$$

If one evaluates this derivative at $x = \kappa(1-p)/2$, one sees that it is positive for p between zero and unity, zero for p equal to unity, and negative for p greater than one. So there is a minimum at this point for I(x) if $0<p<1-2\gamma/\kappa<1$. Otherwise, the point will not be in the range from γ to κ. If the second derivative is also checked for solutions, one sees that an inflection point may occur at

$$x = \frac{-3(p-1)\kappa+\kappa\sqrt{3(p^2-1)}}{6},$$

if p is greater than 1 and less than 2. The intensity function for the hyperlogistic based curve for negative p is

$$I(x) = \beta x^{-1}(\kappa/x)(\kappa/x-1)^{-(p+1)}\{1-e^{\beta/p(\kappa/x-1)^{-p}}\}^{-1}.$$

Because of this function's complexity, no analytic discussion of its properties will be attempted; however, see Figure 8.

For the case of a positive p and the exponential-hyper-Gompertzian size distribution, the intensity function has the form

$$I(x) = \beta' x^{-1}[\log\ (\kappa/x)]^{-(p+1)}.$$

Its first and second derivatives respectively may be written as

$$\frac{dI(x)}{dx} = \beta' x^{-2}[\log(\kappa/x)]^{-(p+2)}[p+1-\log(\kappa/x)]$$

and

$$\frac{d^2I(x)}{dx^2} = \beta' x^{-3}[\log(\kappa/x)]^{-(p+3)}\{2[\log(\kappa/x)]^2-3(p+1)\cdot$$

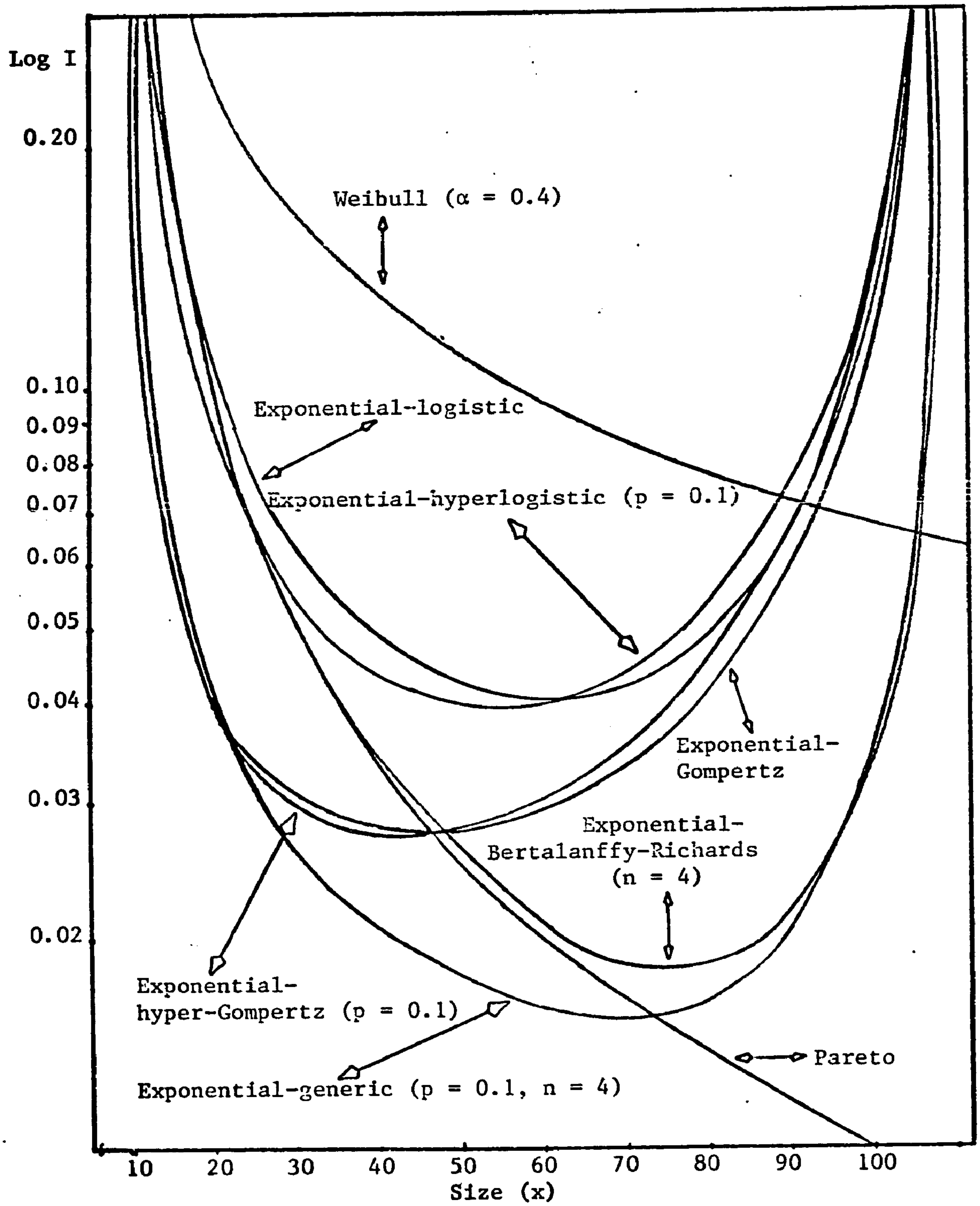
$$[\log(\kappa/x)]+(p+1)(p+2)\}$$

The critical point found by setting the first derivative equal to zero and solving is at

$$x = \kappa e^{-(p+1)}.$$

Evaluating the second derivative there, one sees that it is equal to

Fig. 8 Graphs of selected intensity functions (I) with parameter values $\gamma = 1$, $\kappa = 101$ and $\beta = \beta' = \beta'' = 1$.



$\beta' \kappa^{-3} e^{3(p+1)} (p+1)^{-(p+2)}$, which is positive for all β', κ and p positive. So one finds that the point $x = \kappa e^{-(p+1)}$ is a minimum which is in the range from γ to κ for all p between $\{[\log(\kappa/\gamma)]-1\}$ and zero if $\kappa > \gamma e$. A look at the solution for the second derivative tells one that inflection points may occur at

$$x = \kappa \cdot \exp \left\{ \frac{-3(p+1) \pm \sqrt{(p+1)(p-7)}}{4} \right\}$$

and that these points will be real for p greater than or equal to 7. If p is equal to 7, only one such point will exist. The point is in the range from γ to κ if one has

$$\log(\kappa/\gamma) > \frac{3(p+1) \pm \sqrt{(p+1)(p-7)}}{4} > 0.$$

The intensity function for negative p is

$$I(x) = \beta' x^{-1} [\log(\kappa/x)]^{-(p+1)} \{1 - e^{\beta'/p[\log(\kappa/x)]^{-p}}\}^{-1}.$$

No discussion of properties has been attempted for this case. Again see Figure 8.

The intensity function for the exponential-generic growth size distribution with p greater than zero is

$$I(x) = \beta \kappa^n x^{-(n+1)} [(\kappa/x)^n - 1]^{-(p+1)}.$$

From the first derivative,

$$\frac{dI(x)}{dx} = \beta \kappa^n x^{-(n+2)} [(\kappa/x)^n - 1]^{-(p+2)} \cdot$$

$$\{-(n+1)[(\kappa/x)^n - 1] + (p+1)n(\kappa/x)^n\},$$

one obtains a critical point at $x = \kappa[(1-np)/(n+1)]^{1/n}$. This point is within the range from γ to κ if κ is greater than $(n+1)^{1/n}\gamma$ and p is between zero and $1/n[1-(n+1)(\gamma/\kappa)^n]$. The second derivative is a complicated function which finally simplifies to

$$\frac{d^2I(x)}{dx^2} = \beta\kappa^n x^{-(3+n)}[(\kappa/x)^n-1]^{-(p+3)} \cdot$$

$$\{[n(1+pn)+2(1+pn)](\kappa/x)^{2n}$$

$$-[n^2+5n+4n^2p+4+3np+n^2p^2](\kappa/x)^n$$

$$+[3n^2+5n+n^2p+np+2]\}.$$

By checking the sign of the first derivative on either side of the root, one sees that it changes from negative on the left to positive on the right, implying the presence of a minimum. The second derivative is not particularly useful for this most involved situation. The intensity function for negative p in the generic growth case is

$$I(x) = \beta x^{-1}(\kappa/x)^n[(\kappa/x)^n-1]^{-(p+1)}\{1-e^{\beta/(np)[(\kappa/x)^n-1]^{-p}}\}^{-1}.$$

The last of the upper curves, the Weibull, has a simpler intensity function than any of the other size family member except the Pareto. For positive values of the parameter α, the function is $I(x) = \beta''x^{\alpha-1}$. Its first and second derivatives are

$$\frac{dI(x)}{dx} = \beta''(\alpha-1)x^{\alpha-2}$$

and $$\frac{d^2I(x)}{dx^2} = \beta''(\alpha-1)(\alpha-2)x^{\alpha-3},$$

respectively. I(x) has no critical points or points of inflection between γ and infinity. For α greater than unity, I(x) is an unbounded rising curve moving upward from $\beta''\gamma^{\alpha-1}$. If α is equal to unity, I(x) is a constant, horizontal line at I(x) equals β'' for all x between γ and κ. If α is less than one, I(x) is a J-shaped curve decreasing from $\beta''\gamma^{\alpha-1}$ toward zero. For the situation of a negative α, the intensity function is

$$I(x) = \beta''x^{\alpha-1}[1-e^{\beta''/\alpha x^{\alpha}}]^{-1}.$$

Its first and second derivatives are found to be

$$\frac{dI(x)}{dx} = \beta''x^{\alpha-2}(1-e^{\beta''/\alpha x^{\alpha}})^{-1} \cdot$$

$$[\alpha-1+\beta''x^{\alpha}e^{\beta''/\alpha x^{\alpha}}(1-e^{\beta''/\alpha x^{\alpha}})^{-1}]$$

and

$$\frac{d^2I(x)}{dx^2} = \{\beta''(\alpha-1)(\alpha-2)x^{\alpha-3}(1-e^{\beta''/\alpha x^{\alpha}})^{-1}$$

$$+3(\beta'')^2(\alpha-1)x^{2\alpha-3}e^{\beta''/\alpha x^{\alpha}}(1-e^{\beta''/\alpha x^{\alpha}})^{-2}$$

$$+(\beta'')^3x^{3(\alpha-1)}e^{\beta''/\alpha x^{\alpha}}(1-e^{\beta''/\alpha x^{\alpha}})^{-2}$$

$$+2(\beta'')^3x^{3(\alpha-1)}e^{\beta''/\alpha x^{\alpha}}(1-e^{\beta''/\alpha x^{\alpha}})^{-3}\}.$$

An implicit solution for a critical point or points may be found from the first derivative as

$$x_c = [-\alpha/\beta''\log(1+\{\beta''x^{\alpha}\}/\{1-\alpha\})]^{1/\alpha}.$$

If the boundary conditions are checked, one sees that the point is always within the range for $x>\gamma>0$ if

$$x_c > [(1-\alpha)/\beta''(e^{-\beta''/\alpha\gamma^{\alpha}}-1)]^{1/\alpha}.$$

Further information on this intensity function is not easily obtained by these means.

By way of comparison with other frequently used functions, one may note that the intensity functions presented here tend to be concave upward as U-shaped or reversed J-shaped curves for the cases having a finite upper limit κ. The two cases of unlimited size, the Pareto and the Weibull, tend to have J-shaped curves. The intensity function for the normal or Gaussian distribution, however, is a bounded rising curve concave downward. That for the log-normal rises quickly to a maximum and then declines more gradually. That for the logistic is similar to the intensity curve for the normal, with a more gentle slope. That for the exponential is a horizontal line and the function for the Gamma is close to those of the logistic and the normal. The uniform's intensity function is a concave upward rising curve as are some of those from this family of size distributions. (See Gumbel, 1958).

VII. ESTIMATION

A. Starting Values

One of the convenient properties of the generic family of size distributions is that the distribution functions are available in closed form. A second is that the various models bear specific relationships through nesting or through limits with each other. These two together make it possible to estimate the parameters from an empirical distribution function in a straight-forward manner. Then for the parameters γ and κ which are the lower and upper bounds respectively, one may always construct simple estimates in terms of the minimum and maximum sample observations. For γ, one can choose either the minimum observation itself or perhaps some multiple of it based upon the size of the sample, N. For example, one may choose $(N-1)/N \min(\underline{x})$, where $\underline{x}$ represents the vector of observations $\{x_1, x_2, \ldots, x_N\}$. However, it is not unusual in real data to find that the least squares estimate (weighted or unweighted) for γ subject to the restriction that it is at least less than or equal to the smallest observation is indeed the minimum of $\underline{x}$. For maximum likelihood estimation $\hat{\gamma}$ is the minimum of $\underline{x}$ in all cases, for all family members. One should also notice here that γ is interpretable either as the smallest measureable size or as a lower truncation point. The forms of the functions will remain the same in either situation. The presence of an upper truncation point which is less than the upper end of the range, by contrast, does change the forms of the distributions. Besides, the parameter κ is heavily intertwined with the parameter β in determining the shapes of the various functions. Although one might wish to use a multiple of the maximum sample observation to estimate κ, one may not use that maximum observation by itself. The distribution functions and the density functions are not defined at x equal to κ and may actually be infinite there. This difficulty has more significance in estimation by maximum likelihood techniques. A possible recommendation for a starting value for κ is $(N+1)/N \max(\underline{x})$, which at least avoids the problem for a first guess. Another point to mention here is that this initial guess for κ is generally farther from the mark than that for γ. The reason is that the value of κ does not only indicate an end-point of the range for x as does that of γ. It is also very intimately connected to the shapes and characteristics of the functions which contain it.

First one should examine the estimation for the simplest case, the two-parameter Pareto. Recalling that the distribution function has the form $F(x) = 1-(x/\gamma)^{-\beta}$, one may invert it to find an expression for the value of x corresponding to any particular point, c, of the distribution function: for example, a median, a quartile or a decile. For the Pareto, this expression is $x_c = \gamma(1-c)^{-1/\beta}$. For the median, this becomes $x_{\frac{1}{2}} = \gamma(\frac{1}{2})^{-1/\beta}$. If one were to construct an empirical distribution function with its graph, it would be possible to read off a value for $\tilde{x}_{\frac{1}{2}}$. In the case of the median, however, the value is easily calculable from a frequency distribution of the variable.

If $x_{(r)}$ designates the value of the rth ranked (in ascending order) observation in a set of N observations with r being the average rank for tied values, the empirical distribution function $\hat{F}(x)$ may be estimated by $\hat{F}(x_{(r)}) = r/(N+1)$. If one solves for β in terms of $x_{\frac{1}{2}}$ and γ for the Pareto and, in addition, one uses $(N-1)/N\ x_{(1)}$ as the γ estimate and $\tilde{x}_{\frac{1}{2}}$ for $x_{\frac{1}{2}}$, one may obtain a simple starting value for β by substituting in the following equation:

$$\beta = (\log 2)/[\log x_{\frac{1}{2}} - \log \gamma].$$

So for the Pareto one has that $\tilde{\gamma} = (N-1)N\ x_{(1)}$ and $\tilde{\beta} = (\log 2)/[\log \tilde{x}_{\frac{1}{2}} - \log((N-1)/N\ x_{(1)})]$.

Using the same methods for the exponential-logistic size distributions, one may solve for any percentile point, c, as

$$x_c = \kappa\{1+(1-c)^{1/\beta}(\kappa/\gamma-1)\}^{-1}.$$

In this case, there are only three parameters to estimate γ, β and κ. For $\tilde{\gamma}$ and $\tilde{\kappa}$, one may use $(N-1)/N\ x_{(1)}$ and $(N+1)/N\ x_{(N)}$ respectively. Then if one again uses the relationship for the median as estimated from an empirical distribution function, one gets

$$x_{\frac{1}{2}} = \kappa\{1+(\tfrac{1}{2})^{1/\beta}(\kappa/\gamma-1)\}^{-1}$$

which, upon substituting for γ, κ and $x_{\frac{1}{2}}$, yields an estimate for β as

$$\tilde{\beta} = (\log 2)\{\log(\tilde{\kappa}-\tilde{\gamma}) - \log(\tilde{\gamma}) - \log(\tilde{\kappa}-\tilde{x}_{\frac{1}{2}}) + \log \tilde{x}_{\frac{1}{2}}\}^{-1}.$$

Now checking the exponential-Gompertzian distribution, one sees that for any percentile point, c, the relationship for x_c is

$$x_c = \kappa(\gamma/\kappa)^{(1-c)^{1/\beta'}}.$$

If one solves for β' instead, the corrresponding relationship is

$$\beta' = \log(1-c)/\log\{(\log \kappa - \log x_c)/(\log \kappa - \log \gamma)\}.$$

Again one may use $(N-1)N\, x_{(1)}$ for $\tilde{\gamma}$ and $(N+1)/N\, x_{(N)}$ for $\tilde{\kappa}$. Using the relationship at the median and estimating $x_{\frac{1}{2}}$ from the empirical distribution function, one finds

$$\tilde{\beta}' = (-\log 2)/\log[(\log \tilde{\kappa} - \log \tilde{x}_{\frac{1}{2}})/(\log \tilde{\kappa} - \log \tilde{\gamma})].$$

As is already apparent, these models are presented in hierarchical form partly as a help in understanding their characteristics and partly so that this relationship may be exploited to give clues for estimation. The exponential-Bertalanffy-Richards is the most complicated of the lower curves and, because of its shape limitations already discussed, the simplest of the curves having more than three parameters. It is related to the Gompertz based curve as a form which produces it as the limiting distribution when n approaches 0 and β/n approaches the β' of the exponential-Gompertz. This is the parallel limit to that shown for the growth curves themselves. Nonlinear fitting has well-known difficulties which increase with the number of parameters in the model. The three curves presented earlier are relatively easy to fit. For this reason, it would be better to proceed by fitting the Gompertzian and logistic based functions to a given set of data. The fitted estimates can then be used in conjunction with knowledge of the curve structure to get better starting values for the four-parameter Bertalanffy-Richards based curve. Suppose, for instance, that one were to have a data set for which the Gompertzian and the logistic based models had been fitted. If, by some measure, the Gompertzian based curve were judged to have given the better fit, but systematic deviation from fit remained, it would be reasonable to try a Bertalanffy-Richards based model with a starting value for n between 0 and 1. Placing it at $\tilde{n}=\frac{1}{2}$ is a good choice. The fitting of this model is not particularly sensitive to small misses for the starting value of n. Then for the estimate, there are several choices. One may choose to use the Gompertzian estimate for β' and the limiting relationship to calculate $\tilde{\beta}$ as $\tilde{n}\tilde{\beta}'$ or, in this case, $\frac{1}{2}\tilde{\beta}'$. One might also reach for the empirical distribution function. The equation for β in terms of any percentile point c is

$$\beta = n \log(1-c)\{\log[(\kappa/x_c)^n-1]-\log[(\kappa/\gamma)^n-1]\}.$$

Using $\tilde{n} = \frac{1}{2}$ and the estimates from the Gompertzian based curve for γ and κ along with an estimate of $x_{\frac{1}{2}}$ from the empirical distribution function, one obtains

$$\tilde{\beta} = -\tfrac{1}{2}(\log 2)\{\log[(\tilde{\kappa}/\tilde{x}_{\frac{1}{2}})^{\frac{1}{2}}-1]-\log[(\tilde{\kappa}/\tilde{\gamma})^{\frac{1}{2}}-1]\}.$$

If the logistic-based curve appeared to fit better than the Gompertz-based curve, one could simply start the more complicated fit from the estimates for the exponential-logistic with $\tilde{n} = 1$. If, on the other hand, one prefers not to start by fitting a lower model, one might use $(N+1)/N\ x_{(N)}$ to estimate κ, $(N-1)N\ x_{(1)}$ to estimate γ and an iterative estimate for both β and n (not necessarily a converged estimate, but one pursued several cycles) based upon two points, perhaps at 1/3 and 2/3, of the empirical distribution function. To begin this iterative cycle, it is usually acceptable to use $\tilde{n} = 1$. Then the two equations at 1/3 and 2/3 are

$$\tilde{\beta} = \tilde{n}\ \log(2/3)/\{[\log(\tilde{\kappa}/\tilde{x}_{1/3})^{\tilde{n}}-1]-[\log(\tilde{\kappa}/\tilde{\gamma})^{\tilde{n}}-1]\}$$

and

$$\tilde{n} = \tilde{\beta}\{\log[(\tilde{\kappa}/\tilde{x}_{2/3})^{\tilde{n}}-1]-\log[(\tilde{\kappa}/\tilde{\gamma})^{\tilde{n}}-1]\}/\log(1/3).$$

This last method for obtaining starting values is preferable if there is reason to suspect an n much larger than one.

Now the exponential-hyperlogistic curve for negative p is not hard to estimate. The value of p may only be between zero and -1. Selecting a $\tilde{p}$ of -1/2 is quite adequate. This used with the estimates of γ, κ and β from the logistic-based function will usually converge quite rapidly. Again, if one would rather not start from the logistic, but use the empirical distribution function, one may use the 1/3 and 2/3 points with $\tilde{\beta} = 1$ and $\tilde{p} = -1/2$ and iterate a few cycles for a pair of starting values. With these and (N-1)N $x_{(1)}$ for $\tilde{\gamma}$ and (N+1)/N $x_{(N)}$ for $\tilde{\kappa}$, the two equations are

$$\tilde{\beta} = \tilde{p}(\tilde{\kappa}/\tilde{\gamma}-1)^{\tilde{p}}\{\log(2/3)-\log[e^{-\tilde{\beta}/\tilde{p}(\tilde{\kappa}/\tilde{x}_{1/3}-1)^{-\tilde{p}}}-1/3]\}$$

and

$$\tilde{p} = \tilde{\beta}(\tilde{\kappa}/\tilde{\gamma}-1)^{-\tilde{p}}\{\log(1/3)-\log[e^{-\tilde{\beta}/\tilde{p}(\tilde{\kappa}/\tilde{x}_{2/3}-1)^{-\tilde{p}}}-2/3]\}^{-1}.$$

For positive values of p and the hyperlogistic based model, one might arbitrarily choose to start at $\tilde{p} = 1$ and the estimates from the logistic-based fit of γ, κ and β. In many cases that would be quite adequate; one might also try a $\tilde{p}$ of 1 and a $\tilde{\beta}$ of 1 with $\tilde{\gamma}$ = (N-1)/N $x_{(1)}$ and $\tilde{\kappa}$ = (N+1)/N $x_{(N)}$. Or one might wish to use the empirical distribution function at 1/3 and 2/3 to iterate a few cycles for the starting values for $\tilde{\beta}$ and $\tilde{p}$ beginning with either of the above suggestions.

These equations are

$$\tilde{\beta} = -\tilde{p}[(\tilde{\kappa}/\tilde{x}_{1/3}-1)^{-\tilde{p}}-(\tilde{\kappa}/\tilde{\gamma}-1)^{-\tilde{p}}]^{-1}\log(2/3)$$

and

$$\tilde{p} = -\tilde{\beta}[(\tilde{\kappa}/\tilde{x}_{2/3}-1)^{-\tilde{p}}-(\tilde{\kappa}/\tilde{\gamma}-1)^{-p}]/\log(1/3).$$

The estimation for the exponential-hyper-Gompertzian distribution is similar to that just considered. For negative values of p beginning with $\tilde{p} = -1/2$ and the $\tilde{\gamma}$, $\tilde{\kappa}$ and $\tilde{\beta}'$ from the exponential-Gompertzian form is usually good. One could use also $\tilde{p} = -1/2$, $\tilde{\beta}' = 1$, $\tilde{\gamma} = (N-1)/N\ x_{(1)}$ and $\tilde{\kappa} = (N+1)/N\ x_{(N)}$. Particularly in this latter case one often gains from using the empirical distribution function estimates of $x_{1/3}$ and $x_{2/3}$ to improve the starting values iteratively. For this purpose, the two equations are

$$\tilde{\beta}'= \tilde{p}[\log(\tilde{\kappa}/\tilde{\gamma})]^{\tilde{p}}\{\log(2/3)-\log[e^{-\tilde{\beta}'/\tilde{p}[\log(\tilde{\kappa}/\tilde{x}_{1/3})]^{-\tilde{p}}}-1/3]\}$$

and

$$\tilde{p} =\tilde{\beta}'[\log(\tilde{\kappa}/\tilde{\gamma})]^{-p}\{\log(1/3)-\log[e^{-\tilde{\beta}'/\tilde{p}[\log(\tilde{\kappa}/\tilde{x}_{2/3})]^{-\tilde{p}}}-2/3]\}^{-1}.$$

For positive values of p, one can consider either set of starting values for $\tilde{\gamma}$, $\tilde{\kappa}$ and $\tilde{\beta}'$ which were covered above along with an estimate of p at 1. Using the empirical distribution function estimates of $x_{1/3}$ and $x_{2/3}$ to improve these estimates iteratively yields the following pair of equations for estimates of $\tilde{\beta}'$ and $\tilde{p}$:

$$\tilde{\beta}'= -\tilde{p}\{[\log(\tilde{\kappa}/\tilde{x}_{1/3})]^{-\tilde{p}}-[\log(\tilde{\kappa}/\tilde{\gamma})]^{-\tilde{p}}\}^{-1}\log(2/3)$$

and

$$\tilde{p} =-\tilde{\beta}'\{[\log(\tilde{\kappa}/\tilde{x}_{2/3})]^{-\tilde{p}}-[\log(\tilde{\kappa}/\tilde{\gamma})]^{-\tilde{p}}\}/\log(1/3).$$

Like the exponential-hyperlogistic distribution, this distribution's ability to yield converged estimates in a data set is relatively insensitive to small misses in the estimate of p. In both cases, it is more important (but not crucial) for the relative magnitudes of $\tilde{\beta}'$ and $\tilde{p}$ to be correct.

The exponential-generic growth size distribution is approached cautiously, as a model of last resort. It has five parameters. Most data contain insufficient information to make this model very practical or the estimates, if at all obtainable, very stable. When it is used, it is because one of the four-parameter models fits better than the others, but leaves too much unexplained systematic deviation from the theoretical curve. The method used to produce starting values for the generic-based curve must depend in part upon which of the four-parameter curves gave the best fit. It is not the best approach to proceed from the empirical distribution function in view of the fact that fairly good estimates of four, if not five of the parameters should be obtainable from the best four-parameter model. From a fit of the exponential-hyper-Gompertzian model, one would have perfectly proper estimates for γ, κ, p and β', in either the positive or negative p situation. With a reasonable starting value for n, an estimate for β may be found by noticing that the β' of the hyper-Gompertz based distribution is the limit of βn^{-1-p} as β and n approach zero. So β could be estimated as $\tilde{\beta}'\tilde{n}^{1+p}$, given estimates for β', n and p. The problem is the estimation of n. Of course, if the exponential-Bertalanffy-Richards has also been used, the estimate of n obtained therein could be tried. If this has not been found, but the hyper-Gompertz based model looks better than the hyperlogistic-based model, $\tilde{n} = 1/2$ would seem appropriate. If, instead, one finds that the hyperlogistic-based model is the best of the four-parameter models, estimates of γ, κ, β and p are present from that model and may be used as starting values for the five-parameter version. Again a value for $\tilde{n}$ can be taken from the exponential-Bertalanffy-Richards if it has also been fit. If not, a value of $\tilde{n} = 1$ is probably not so bad. If the exponential-Bertalanffy-Richards curve was the best of the three four-parameter models, it gives good starting values for $\tilde{\gamma}$, $\tilde{\kappa}$, $\tilde{\beta}$ and $\tilde{n}$ for the exponential-generic curve. It is the estimate of p in this case which gives the most difficulty. If both the other models have also been tried, one could take the $\tilde{p}$ from the better of them. If they seem to be roughly equivalent, one could use the average $\tilde{p}$ of the two. If neither model has been used, one could evaluate the situation (and measures of fit) for $\tilde{p} = -1/2$ and $\tilde{p} = 1$ and proceed using the better of the two. In this case more than the others mentioned, it may do some good to iterate a few cycles for some or all of the estimates using points from the empirical distribution function. For negative p and the c_1, c_2 and c_3 points, these equations are:

$$\tilde{\beta} = \tilde{n}\tilde{p}[(\tilde{\kappa}/\tilde{\gamma})^{\tilde{n}}-1]^{\tilde{p}}\{\log(1-c_1)-\log[e^{-\tilde{\beta}/\tilde{n}\tilde{p}[(\tilde{\kappa}/x_{c_1})^{\tilde{n}}-1]^{-\tilde{p}}}-c_1]\},$$

$$\tilde{n} = \tilde{\beta}/\tilde{p}[(\tilde{\kappa}/\tilde{\gamma})^{n}-1]^{-p}\{\log(1-c_2)-\log[e^{-\tilde{\beta}/\tilde{n}\tilde{p}[(\tilde{\kappa}/\tilde{x}_{c_2})^{\tilde{n}}-1]^{-\tilde{p}}}-c_2]\}^{-1}$$

and

$$\tilde{p} = \tilde{\beta}/\tilde{n}[(\tilde{\kappa}/\tilde{\gamma})^{n}-1]^{-\tilde{p}}\{\log(1-c_3)-\log[e^{-\tilde{\beta}/\tilde{n}\tilde{p}[(\tilde{\kappa}/\tilde{x}_{c_3})^{\tilde{n}}-1]^{-\tilde{p}}}-c_3]\}^{-1}.$$

If only one is to be used, one could set that c to 1/2. If two are to be used, one could choose 1/3 and 2/3. If all three are to be used, 1/4, 1/2 and 3/4 are reasonable choices for c_1, c_2 and c_3. For positive values of p, the same situation is present, but the three equations are:

$$\tilde{\beta} = -\tilde{n}\tilde{p}[\log(1-c_1)]\{[(\tilde{\kappa}/\tilde{x}_{c_1})^{\tilde{n}}-1]^{-\tilde{p}}-[(\tilde{\kappa}/\tilde{\gamma})^{\tilde{n}}-1]^{-\tilde{p}}\}^{-1},$$

$$\tilde{n} = -\tilde{\beta}/\tilde{p}[\log(1-c_2)]^{-1}\{[(\tilde{\kappa}/\tilde{x}_{c_2})^{\tilde{n}}-1]^{-\tilde{p}}-[(\tilde{\kappa}/\tilde{\gamma})^{\tilde{n}}-1]^{-\tilde{p}}\}$$

and

$$\tilde{p} = -\tilde{\beta}/\tilde{n}[\log(1-c_3)]^{-1}\{[(\tilde{\kappa}/\tilde{x}_{c_3})^{\tilde{n}}-1]^{-\tilde{p}}-[(\tilde{\kappa}/\tilde{\gamma})^{\tilde{n}}-1]^{-\tilde{p}}\}.$$

In discussing the possible starting values for the Weibull parameters γ, β'' and α, one should recall the relationship with the exponential-hyper-Gompertz, the exponential-hyperlogistic and the exponential-generic curves. The Weibull has been obtained here as a limiting case from each of these. Depending upon which of these others is considered the best, one can use the nature of the appropriate limiting forms to use its parameters to estimate the β'' and α of the Weibull. The γ estimate from the more general curve may be used intact. If the hyper-Gompertz based curve is the "best fitting" one, the limit to the Weibull is achieved as κ increases without bound, $(p+1)/\log \kappa$ approaches α and $\beta'[\log \kappa]^{-(p+1)}$ approaches β'' (or $\delta'[\log \kappa]^{p+1}$ approaches δ'' in the growth curve). So, if there are estimates of κ, p and β' from the hyper-Gompertz based curve, these may be used to estimate α and β'':

$$\tilde{\alpha} = (p+1)/\log(\tilde{\kappa})$$

and $\tilde{\beta}'' = [\log(\tilde{\kappa})]^{-(p+1)}\tilde{\beta}'$.

If the Weibull is found, instead, as a limiting form from the hyper-logistic or generic growth based curves, the appropriate estimates for p, n and β may be used to estimate α and β'' with the limiting relationships for these cases which state that as κ increases without bound, np approaches α and $\beta\kappa^{-np}$ approaches β'' (or $\delta\kappa^{np}$ approaches δ'' in the growth curves). Remember here that n is equal to 1 in the exponential-hyperlogistic. The equations for estimating the Weibull parameters this way are:

$$\tilde{\alpha} = \tilde{n}\tilde{p}$$

and

$$\tilde{\beta}'' = \tilde{\beta}\tilde{\kappa}^{-\tilde{n}\tilde{p}}.$$

The estimate for γ may remain that from the generic or hyperlogistic curves.

Note that in the limit from the hyper-Gompertz based curve α should only be positive while, in the cases of the hyperlogistic or generic based curves, α may be positive or negative (not necessarily greater than -1). If α were to approach 0, the curve would reduce to that of the Pareto, a limiting case of the Weibull also.

The Weibull distribution is notoriously difficult to fit to real data. This is because it is one of the most sensitive distributions to the choice of starting values for non-linear estimation. It is by far the most sensitive, in that way, of the members of this size family. In particular, many attempts to apply the Weibull in different data sets show the importance of correctly assessing the relative magnitudes of α and β''. In this formulation, β'' is often quite small, perhaps of order 10^{-6} in magnitude and yet significantly different from 0. The parameter α, however, is usually of a "more reasonable" size. If α is greater than 1, the distribution function has a point of inflection at $\{(\alpha-1)/\beta''\}^{1/\alpha}$. It is sometimes possible, if this point is distinct enough in an empirical distribution function, to solve for $\tilde{\beta}''$ in terms of $\tilde{\alpha}$ and the estimated inflection point, $\tilde{x}_I$:

$$\tilde{\beta}'' = (\tilde{\alpha}-1)\tilde{x}_I^{-\tilde{\alpha}}.$$

For this point to be within the range, one should also notice that $\tilde{\beta}''$ must be less than $(\tilde{\alpha}-1)\tilde{\gamma}^{-\alpha}$, $\alpha>1$. From these, one may gain some knowledge of the relative magnitude of α and β''. The above approach is more to be recommended if the other upper curves have not been tried. The parameter γ can be estimated as before, as $(N-1)/N\ x_{(1)}$. A rough guess starting point for α, if α is positive, can be used in equations from the empirical distribution function to produce iterative estimates of α and β''. If the empirical distribution function has no apparent point of inflection, α must be less than or equal to 1. So $\alpha = 1/2$ makes a good trial value. If it appears to have a point of inflection, α must be greater than 1. One might try $\alpha = 2$ for a trial value. In either case, these values may be improved using

$$\tilde{\beta}'' = -\tilde{\alpha}(\tilde{x}_{1/3}^{\tilde{\alpha}}-\tilde{\gamma}^{\tilde{\alpha}})^{-1}\log(2/3)$$

and

$$\tilde{\alpha} = -\tilde{\beta}''(\tilde{x}_{2/3}^{\tilde{\alpha}}-\tilde{\gamma}^{\tilde{\alpha}})/\log(1/3).$$

If one thinks that α is more likely to be negative, the distribution function will have no useful point of inflection. One could try guesses for β'' and α at 1 and -1 respectively and try to improve them using the empirical distribution function:

$$\tilde{\beta}'' = \tilde{\tilde{\alpha}}\gamma^{-\tilde{\alpha}}\{\log(2/3)-\log(e^{-\tilde{\beta}''/\tilde{\tilde{\alpha}}x_{1/3}^{\tilde{\alpha}}}-1/3)\}$$

and

$$\tilde{\alpha} = \tilde{\tilde{\beta}}\gamma^{\tilde{\alpha}}\{\log(1/3)-\log(e^{-\tilde{\beta}''/\tilde{\tilde{\alpha}}x_{2/3}^{\tilde{\alpha}}}-2/3)\}^{-1}.$$

B. Maximum Likelihood Estimation

In the case of the Pareto size distribution, the maximum likelihood estimators of the parameters γ and β are the simplest of the estimators considered in this paper. Recalling the density function,

$$f(x) = \beta/\gamma(x/\gamma)^{-(\beta+1)}, \quad \gamma<x<\infty,$$

one may form the likelihood function for a sample of size N:

$$\phi(\gamma,\beta|\underline{x}) = \beta^N\gamma^{-N\beta}(\prod_{i=1}^{N} x_i)^{-(\beta+1)}.$$

The log likelihood function is shown to be

$$L(\gamma,\beta|\underline{x}) = N \log \beta - N\beta \log \gamma - (\beta+1) \sum_{i=1}^{N} \log x_i.$$

Now taking the first partial derivatives of L with respect to the parameters, one has

$$\frac{\partial L}{\partial \beta} = \frac{N}{\beta} - N \log \gamma - \sum_{i=1}^{N} \log x_i$$

and

$$\frac{\partial L}{\partial \gamma} = \frac{N\beta}{\gamma}.$$

Clearly L is maximized for $\hat{\gamma}$ as large as possible subject to the restriction that it is at least less than or equal to the smallest sample value for x, $x_{(1)}$, and for $\hat{\beta}$ equal to $(\overline{\log x} - \log x_{(1)})^{-1}$ where

$$\overline{\log x} = \frac{1}{N} \sum_{i=1}^{N} \log x_i .$$

So the explicit solution for the Pareto maximum likelihood estimators is the pair

$$\hat{\gamma} = x_{(1)},$$

$$\hat{\beta} = (\overline{\log x} - \log x_{(1)})^{-1}.$$

The matrix of second partial derivatives is

$$\begin{bmatrix} \frac{\partial^2 L}{\partial \gamma^2} & \frac{\partial^2 L}{\partial \gamma \partial \beta} \\ \frac{\partial^2 L}{\partial \beta \partial \gamma} & \frac{\partial^2 L}{\partial \beta^2} \end{bmatrix} = \begin{bmatrix} \frac{-N\beta}{\gamma^2} & \frac{N}{\gamma} \\ \frac{N}{\gamma} & \frac{-N}{\beta^2} \end{bmatrix}$$

So, for this distribution, one may easily write the information matrix as

$$-E \begin{bmatrix} \frac{-N\beta}{\gamma^2} & \frac{N}{\gamma} \\ \frac{N}{\gamma} & \frac{-N}{\beta^2} \end{bmatrix} = \begin{bmatrix} \frac{N\beta}{\gamma^2} & \frac{-N}{\gamma} \\ \frac{-N}{\gamma} & \frac{N}{\beta^2} \end{bmatrix}$$

Since the estimate for γ cannot be obtained from the partial derivative of L with respect to γ in the "usual" manner by equating that derivative to zero and solving for $\hat{\gamma}$, one should really only consider the conditional likelihood given that $\hat{\gamma} = x_{(1)}$ and then solve for the $\hat{\beta}$ which maximizes this conditional function. So the only relevant information matrix is the conditional one:

$$I = -E(-N/\beta^2) = N/\beta^2.$$

Since $-N/\beta^2$ is negative, the conditional likelihood is indeed maximized.

It is necessary to note here that with both range parameters, γ and κ, to be estimated, it is not known whether or not the usual asymptotic properties of maximum likelihood estimators hold in these cases. The usual regularity conditions are not met. See Rao, 1973. It is also not known whether these properties hold in the likelihood estimation conditional on $\hat{\gamma}$ equal to the minimum observation, particularly since the first partial derivatives of the log likelihood function are undefined at κ equal to the maximum observation. Thus the properties of maximum likelihood estimators could be a good subject for future research.

Now for the logistic based size distribution, the density is

$$f(x) = \beta\kappa x^{-2}(\kappa/\gamma-1)^{-\beta}(\kappa/x-1)^{\beta-1}.$$

The likelihood function for a sample of size N is

$$\phi(\gamma,\beta,\kappa|\underline{x}) = \beta^N\kappa^N\gamma^{N\beta}(\kappa-\gamma)^{-N\beta}(\prod_{i=1}^{N} x_i)^{\beta+1}[\prod_{i=1}^{N}(\kappa-x_i)]^{\beta-1}$$

and the log likelihood is

$$L(\gamma,\beta,\kappa|\underline{x}) = N\log\beta+N\log\kappa+N\beta\log\gamma-N\beta\log(\kappa-\gamma)-(\beta+1)\cdot\sum_{i=1}^{N}\log x_i+(\beta-1)\sum_{i=1}^{N}\log(\kappa-x_i).$$

The partial derivatives are seen to be:

$$\frac{\partial L}{\partial\gamma} = \frac{N\beta}{\gamma}+\frac{N}{\kappa-\gamma}$$

$$\frac{\partial L}{\partial\beta} = \frac{N}{\beta}+N\log\gamma-N\log(\kappa-\gamma)-\sum_{i=1}^{N}\log x_i+\sum_{i=1}^{N}\log(\kappa-x_i)$$

$$\frac{\partial L}{\partial \kappa} = \frac{N}{\kappa} - \frac{N\beta}{\kappa-\gamma} + (\beta-1)\sum_{i=1}^{N}(\kappa-x_i)^{-1}.$$

Again the estimate for γ is $x_{(1)}$, although this is not found by using the partial derivative. Considering the conditional likelihood given that $\hat{\gamma} = x_{(1)}$, one finds that it is possible to solve for $\hat{\beta}$ explicitly in terms of $\hat{\gamma}$, $\hat{\kappa}$ and the observations:

$$\hat{\beta} = \{\overline{\log x} - \overline{\log(\hat{\kappa}-x)} + \log(\hat{\kappa}-\hat{\gamma}) - \log \hat{\gamma}\}^{-1}$$

where

$$\overline{\log x} = \frac{1}{N}\sum_{i=1}^{N}\log x_i \text{ and } \overline{\log(\hat{\kappa}-x)} = \frac{1}{N}\sum_{i=1}^{N}\log(\hat{\kappa}-x_i).$$

An estimate for κ, however, is only available implicitly:

$$\hat{\kappa} = \{\hat{\beta}[(\hat{\kappa}-\hat{\gamma})^{-1} - \frac{1}{N}\sum_{i=1}^{N}(\kappa-x_i)^{-1}] + \frac{1}{N}\sum_{i=1}^{N}(\kappa-x_i)^{-1}\}^{-1}$$

In order to use the method of scoring to maximize the conditional log likelihood, one needs the matrix of second partial derivatives whose elements are:

$$\frac{\partial^2 L}{\partial \beta^2} = -N\beta^{-2},$$

$$\frac{\partial^2 L}{\partial \beta \partial \kappa} = \frac{\partial^2 L}{\partial \kappa \partial \beta} = -N(\kappa-\hat{\gamma})^{-1} + \sum_{i=1}^{N}(\kappa-x_i)^{-1}$$

and

$$\frac{\partial^2 L}{\partial \kappa^2} = -N\kappa^{-2} + N\beta(\kappa-\hat{\gamma})^{-2} - (\beta-1)\sum_{i=1}^{N}(\kappa-x_i)^{-2}.$$

The expected values of these two last elements are not easily found and, since they exist as series, would not be very useful. For the exponential-logistic distribution, the best starting values for maximum likelihood estimation are $\hat{\kappa}=(N+1)/N\ x_{(N)}$ and the $\hat{\beta}$ of the Pareto.

The exponential-Gompertzian distribution has a density function of the form

$$f(x) = \beta'[\log(\kappa/\gamma)]^{-\beta'} x^{-1}[\log(\kappa/x)]^{\beta'-1}.$$

The likelihood function and the log likelihood functions are

$$\phi(\gamma,\beta',\kappa|\underline{x}) = \beta'^{N}[\log \kappa-\log \gamma]^{-N\beta'}[\prod_{i=1}^{N} x_i^{-1}] \cdot \prod_{i=1}^{N} [\log \kappa-\log x_i]^{\beta'-1}$$

and

$$L(\gamma,\beta',\kappa|\underline{x}) = N \log \beta'-N\beta'\log[\log \kappa-\log \gamma]-\sum_{i=1}^{N} \log x_i + (\beta'-1)\sum_{i=1}^{N} \log[\log \kappa-\log x_i].$$

The vector of first partial derivatives with respect to the parameters is

$$\begin{bmatrix} \dfrac{\partial L}{\partial \gamma} \\ \\ \dfrac{\partial L}{\partial \beta'} \\ \\ \dfrac{\partial L}{\partial \kappa} \end{bmatrix} = \begin{bmatrix} N\beta'\gamma^{-1}[\log \kappa-\log \gamma]^{-1} \\ \\ N(\beta')^{-1}-N\log[\log \kappa-\log \gamma] + \sum_{i=1}^{N} \log[\log \kappa-\log x_i] \\ \\ -N\beta'\kappa^{-1}[\log \kappa-\log \gamma]^{-1}+(\beta'-1)\kappa^{-1} \cdot \sum_{i=1}^{N} [\log \kappa-\log x_i]^{-1} \end{bmatrix}$$

Again the maximum likelihood estimate of γ is $x_{(1)}$, but this may not be found by setting the partial derivative to zero and solving. Considering the conditional likelihood given this γ estimate, one finds an explicit expression for $\hat{\beta}'$ in terms of $\hat{\gamma}$, $\hat{\kappa}$ and the observations:

$$\hat{\beta}' = \{\log[\log \hat{\kappa}-\log \hat{\gamma}] - \overline{\log[\log \hat{\kappa}-\log x]}\}^{-1}$$

where

$$\overline{\log[\log \kappa-\log x]} = \frac{1}{N}\sum_{i=1}^{N} \log[\log \kappa-\log x_i].$$

An implicit relationship for $\hat{\kappa}$ may be written as

$$\hat{\kappa} = \hat{\gamma}[\exp\{N\beta'(\beta'-1)^{-1}[\sum_{i=1}^{N} (\log \hat{\kappa}-\log x_i)^{-1}]^{-1}\}].$$

If one wishes to use the method of scoring for estimation of β' and κ, the matrix of second partial derivatives is needed. Its elements, for the conditional likelihood, are:

$$\frac{\partial^2 L}{\partial \beta'^2} = -N\beta'^{-2},$$

$$\frac{\partial^2 L}{\partial \beta' \partial \kappa} = \frac{\partial^2 L}{\partial \kappa \partial \beta'} = -N\kappa^{-1}(\log \kappa-\log \hat{\gamma})^{-1}+\kappa^{-1}\sum_{i=1}^{N}(\log \kappa-\log x_i)^{-1}$$

and

$$\frac{\partial^2 L}{\partial \kappa^2} = \kappa^{-2}\{N\beta'[(\log \kappa-\log \hat{\gamma})^{-1}+(\log \kappa-\log \hat{\gamma})^{-2}]-(\beta'-1)\cdot$$

$$\sum_{i=1}^{N}[(\log \kappa-\log x_i)^{-1}+(\log \kappa-\log x_i)^{-2}]\}$$

The expected values of these second partials are not of simple form; therefore the information matrix is not displayed here. For starting values, one of the techniques from the earlier discussion may be used, or one might use $\hat{\gamma} = x_{(1)}$, $\hat{\kappa} = (N+1)/N\ x_{(N)}$ and $\hat{\beta}' = [\log \hat{\kappa}]\hat{\beta}$, where $\hat{\beta}$

is the maximum likelihood estimate of β in the Pareto. This approach uses the limiting relationship between the Pareto and the exponential-Gompertz.

The exponential-Bertalanffy-Richards has a density of the form

$$f(x) = \beta\kappa^{n}[(\kappa/\gamma)^{n}-1]^{-\beta/n}x^{-(n+1)}[(\kappa/x)^{n}-1]^{\beta/n-1}.$$

Its likelihood and log likelihood functions are

$$\phi(\gamma,\beta,n,\kappa|\underline{x}) = \beta^{N}\kappa^{nN}[(\kappa/\gamma)^{n}-1]^{-\beta/nN}\left\{\prod_{i=1}^{N} x_i^{-(n+1)}\right\} \cdot \left\{\prod_{i=1}^{N}[(\kappa/x_i)^{n}-1]^{\beta/n-1}\right\}$$

and

$$L(\gamma,\beta,n,\kappa|\underline{x}) = N\log\beta+nN\log\kappa-\beta N/n\log[(\kappa/\gamma)^{n}-1] -(n+1)\sum_{i=1}^{N}\log x_i+(\beta/n-1)\sum_{i=1}^{N}\log[(\kappa/x_i)^{n}-1].$$

The first partial derivatives of L are as follows:

$$\frac{\partial L}{\partial\gamma} = \beta N\gamma^{-1}[1-(\gamma/\kappa)^{n}]^{-1},$$

$$\frac{\partial L}{\partial\beta} = -Nn^{-1}\log[(\kappa/\gamma)^{n}-1]+N\beta^{-1}+n^{-1}\sum_{i=1}^{N}\log[(\kappa/x_i)^{n}-1],$$

$$\frac{\partial L}{\partial n} = \beta Nn^{-2}\log[(\kappa/\gamma)^{n}-1]-\beta Nn^{-1}[(\kappa/\gamma)^{n}-1]^{-1}(\kappa/\gamma)^{n}\log(\kappa/\gamma) +N\log\kappa-\sum_{i=1}^{N}\log x_i-\beta n^{-2}\sum_{i=1}^{N}\log[(\kappa/x_i)^{n}-1] +(\beta-n)n^{-1}\sum_{i=1}^{N}[(\kappa/x_i)^{n}-1]^{-1}(\kappa/x_i)^{n}\log(\kappa/x_i)$$

and

$$\frac{\partial L}{\partial \kappa} = -\beta N\kappa^{-1}[1-(\gamma/\kappa)^n]^{-1}+nN\kappa^{-1}+(\beta-n)\kappa^{-1}\sum_{i=1}^{N}[1-(x_i/\kappa)^n]^{-1}.$$

The same situation as before prevails with respect to $\hat{\gamma}$ and $\hat{\beta}$. For $\hat{\gamma}$ the estimate is $x_{(1)}$. For $\hat{\beta}$ there is an explicit expression in terms of the other parameter estimates and the observations:

$$\hat{\beta} = \hat{n}\{\log[(\hat{\kappa}/\hat{\gamma})^{\hat{n}}-1]-N^{-1}\sum_{i=1}^{N}\log[(\hat{\kappa}/x_i)^{\hat{n}}-1]\}^{-1}.$$

The expressions for $\hat{n}$ and $\hat{\kappa}$ are only available implicitly. The second partial derivatives for use with the method of scoring and the likelihood function conditional on the estimate of γ being $x_{(1)}$ are:

$$\frac{\partial^2 L}{\partial \beta^2} = -N\beta^{-2},$$

$$\frac{\partial^2 L}{\partial\beta\partial n} = \frac{\partial^2 L}{\partial n\partial\beta} = n^{-2}\{N\log[(\kappa/\hat{\gamma})^n-1]-\sum_{i=1}^{N}\log[(\kappa/x_i)^n-1]\}$$

$$-n^{-1}\{N[(\kappa/\hat{\gamma})^n-1]^{-1}(\kappa/\hat{\gamma})^n\log(\kappa/\hat{\gamma})-\sum_{i=1}^{N}$$

$$[(\kappa/x_i)^n-1]^{-1}(\kappa/x_i)^n\log(\kappa/x_i)\},$$

$$\frac{\partial^2 L}{\partial\beta\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial\beta} = -N\kappa^{-1}[1-(\hat{\gamma}/\kappa)^n]^{-1}+\kappa^{-1}\sum_{i=1}^{N}[1-(x_i/\kappa)^n]^{-1},$$

$$\frac{\partial^2 L}{\partial n^2} = -2\beta Nn^{-3}\log[(\kappa/\hat{\gamma})^n-1]+2\beta Nn^{-2}[(\kappa/\hat{\gamma})^n-1]^{-1}(\kappa/\hat{\gamma})^n\log(\kappa/\hat{\gamma})$$

$$+\beta Nn^{-1}\{[(\kappa/\hat{\gamma})^n-1]^{-2}(\kappa/\hat{\gamma})^{2n}[\log(\kappa/\hat{\gamma})]^2-[(\kappa/\hat{\gamma})^n-1]^{-1}\cdot$$

$$(\kappa/\hat{\gamma})^n[\log(\kappa/\hat{\gamma})]^2\}+2\beta n^{-3}\sum_{i=1}^{N}\log[(\kappa/x_i)^n-1]-2\beta n^{-2}\cdot$$

$$\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-1}(\kappa/x_i)^n\log(\kappa/x_i)-(\beta-n)n^{-1}\cdot$$

$$\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-2}(\kappa/x_i)^{2n}[\log(\kappa/x_i)]^2+(\beta-n)n^{-1}\cdot$$

$$\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-1}(\kappa/x_i)^n[\log(\kappa/x_i)]^2,$$

$$\frac{\partial^2 L}{\partial n\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial n} = -\beta N\kappa^{-1}[1-(\hat{\gamma}/\kappa)^n]^{-2}(\hat{\gamma}/\kappa)^n\log(\hat{\gamma}/\kappa)+N\kappa^{-1}-\kappa^{-1}\cdot$$

$$\sum_{i=1}^{N}[1-(x_i/\kappa)^n]^{-1}+(\beta-n)\kappa^{-1}\sum_{i=1}^{N}[1-(x_i/\kappa)^n]^{-2}\cdot$$

$$(x_i/\kappa)^n\log(x_i/\kappa)$$

and

$$\frac{\partial^2 L}{\partial\kappa^2} = \beta\kappa^{-2}\{N[1-(\hat{\gamma}/\kappa)^n]^{-1}-\sum_{i=1}^{N}[1-(x_i/\kappa)^n]^{-1}\}-\beta n\kappa^{-2}\{N(\hat{\gamma}/\kappa)^n\cdot$$

$$[1-(\hat{\gamma}/\kappa)^n]^{-2}-\sum_{i=1}^{N}(x_i/\kappa)^n[1-(x_i/\kappa)^n]^{-2}\}+n\kappa^{-2}\cdot$$

$$\{\sum_{i=1}^{N}[1-(x_i/\kappa)^n]^{-1}-N\}-n^2\kappa^{-2}\{\sum_{i=1}^{N}(x_i/\kappa)^n[1-(x_i/\kappa)^n]^{-2}\}.$$

The density function for the exponential-hyperlogistic for the p greater than zero case is

$$f(x) = \beta\kappa x^{-2}(\kappa/x-1)^{-(p+1)}\exp\{-\beta/p[(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}]\}.$$

The likelihood function and the log likelihood functions are

$$\phi(\gamma,\beta,p,\kappa|\underline{x}) = \beta^N\kappa^N e^{\beta/pN(\kappa/\gamma-1)^{-p}}\{\prod_{i=1}^{N}x_i^{-2}(\kappa/x_i-1)^{-(p+1)}\}$$

$$e^{-\beta/p\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}}$$

and $L(\gamma,\beta,p,\kappa|\underline{x}) = N\log\beta+N\log\kappa+\beta/p\;N(\kappa/\gamma-1)^{-p}-2\sum_{i=1}^{N}\log x_i$

$$-(p+1)\sum_{i=1}^{N}\log(\kappa/x_i-1)-\beta/p\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}.$$

Now the first partial derivatives of $L(\gamma,\beta,p,\kappa|\underline{x})$ with respect to the parameters are seen to be:

$$\frac{\partial L}{\partial \gamma} = \beta\kappa\gamma^{-2}N(\kappa/\gamma-1)^{-(p+1)},$$

$$\frac{\partial L}{\partial \beta} = Np^{-1}(\kappa/\gamma-1)^{-p}+N\beta^{-1}-p^{-1}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p},$$

$$\frac{\partial L}{\partial p} = -\beta Np^{-2}(\kappa/\gamma-1)^{-p}-\beta Np^{-1}(\kappa/\gamma-1)^{-p}\log(\kappa/\gamma-1)$$

$$+\beta p^{-2}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}+\beta p^{-1}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}\log(\kappa/x_i-1)$$

$$-\sum_{i=1}^{N}\log(\kappa/x_i-1),$$

and

$$\frac{\partial L}{\partial \kappa} = \beta\gamma^{-1}N(\kappa/\gamma-1)^{-(p+1)}+N\kappa^{-1}-\beta\sum_{i=1}^{N}x_i^{-1}(\kappa/x_i-1)^{-(p+1)}$$

$$-(p+1)\sum_{i=1}^{N}x_i^{-1}(\kappa/x_i-1)^{-1}.$$

After noticing that $\hat{\gamma}$ must be $x_{(1)}$ as for the lower curves, one may again consider the conditional likelihood. Solving for $\hat{\beta}$ in terms of the other parameter estimates and the observations, one obtains an explicit solution for $\hat{\beta}$ as

$$\hat{\beta} = \hat{p}\{1/N\sum_{i=1}^{N}(\hat{\kappa}/x_i-1)^{-\hat{p}}-(\hat{\kappa}/\hat{\gamma}-1)^{-\hat{p}}\}^{-1}.$$

Then implicit expressions for $\hat{p}$ and $\hat{\kappa}$ are

$$\hat{p} = \hat{\beta}[\sum_{i=1}^{N}\log(\hat{\kappa}/x_i-1)]^{-1}\{\sum_{i=1}^{N}(\hat{\kappa}/x_i-1)^{-\hat{p}}[\hat{p}^{-1}+\log(\hat{\kappa}/x_i-1)]$$

$$-N(\hat{\kappa}/\hat{\gamma}-1)^{-\hat{p}}[\hat{p}^{-1}+\log(\hat{\kappa}/\hat{\gamma}-1)]\}$$

and

$$\hat{\kappa} = \{\hat{\beta}N^{-1}\sum_{i=1}^{N} x_i^{-1}(\hat{\kappa}/x_i-1)^{-(\hat{p}+1)}+(\hat{p}+1)N^{-1}\sum_{i=1}^{N} x_i^{-1}(\hat{\kappa}/x_i-1)^{-1}$$

$$-\hat{\beta}\hat{\gamma}^{-1}(\hat{\kappa}/\hat{\gamma}-1)^{-(\hat{p}+1)}\}^{-1}.$$

The elements of the matrix of second partial derivatives for use in the method of scoring are:

$$\frac{\partial^2 L}{\partial \beta^2} = -N\beta^{-2},$$

$$\frac{\partial^2 L}{\partial \beta \partial p} = \frac{\partial^2 L}{\partial p \partial \beta} = p^{-2}\{\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}-N(\kappa/\hat{\gamma}-1)^{-p}\}+p^{-1}\cdot$$

$$\{\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}\log(\kappa/x_i-1)-N(\kappa/\hat{\gamma}-1)^{-p}\log(\kappa/\hat{\gamma}-1)\},$$

$$\frac{\partial^2 L}{\partial \beta \partial \kappa} = \frac{\partial^2 L}{\partial \kappa \partial \beta} = -N\hat{\gamma}^{-1}(\kappa/\hat{\gamma}-1)^{-(p+1)}+\sum_{i=1}^{N} x_i^{-1}(\kappa/x_i-1)^{-(p+1)},$$

$$\frac{\partial^2 L}{\partial p^2} = 2\beta p^{-3}\{N(\kappa/\hat{\gamma}-1)^{-p}-\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}\}+2\beta p^{-2}\cdot$$

$$\{N(\kappa/\hat{\gamma}-1)^{-p}\log(\kappa/\hat{\gamma}-1)-\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}\log(\kappa/x_i-1)\}$$

$$+\beta p^{-1}\{N(\kappa/\hat{\gamma}-1)^{-p}[\log(\kappa/\hat{\gamma}-1)]^2-\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}$$

$$[\log(\kappa/x_i-1)]^{-2}\},$$

$$\frac{\partial^2 L}{\partial p \partial \kappa} = \frac{\partial^2 L}{\partial \kappa \partial p} = \beta\{\sum_{i=1}^{N} x_i^{-1}(\kappa/x_i-1)^{-(p+1)}\log(\kappa/x_i-1)-N\hat{\gamma}^{-1}\cdot$$

$$(\kappa/\hat{\gamma}-1)^{-(p+1)}\log(\kappa/\hat{\gamma}-1)\}-\sum_{i=1}^{N} x_i^{-1}(\kappa/x_i-1)^{-1},$$

and

$$\frac{\partial^2 L}{\partial \kappa^2} = \beta(p+1)\{\sum_{i=1}^{N} x_i^{-2}(\kappa/x_i-1)^{-(p+2)} - N\hat{\gamma}^{-2}(\kappa/\hat{\gamma}-1)^{-(p+2)}\}$$

$$+(p+1)\sum_{i=1}^{N} x_i^{-2}(\kappa/x_i-1)^{-2}.$$

Now for negative values of p, the equations are more complicated, but somewhat similar to those already found. The density, one recalls, has the form

$$f(x) = \{1-e^{\beta/p(\kappa/\gamma-1)^{-p}}\}^{-1}\beta\kappa x^{-2}(\kappa/x-1)^{-(p+1)}\cdot$$

$$e^{-\beta/p\{(\kappa/x-1)^{-p}-(\kappa/\gamma-1)^{-p}\}}.$$

So the likelihood and log likelihood functions are

$$\phi(\gamma,\beta,p,\kappa|\underline{x}) = \{1-e^{\beta/p(\kappa/\gamma-1)^{-p}}\}^{-N}\beta^N\kappa^N\{\prod_{i=1}^{N} x_i^{-2}(\kappa/x_i-1)^{-(p+1)}\}\cdot$$

$$e^{-\beta/p\{\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}-N(\kappa/\gamma-1)^{-p}\}}$$

and

$$L(\gamma,\beta,p,\kappa|\underline{x}) = -N\log\{1-e^{\beta/p(\kappa/\gamma-1)^{-p}}\}+N\log\beta+N\log\kappa$$

$$-2\sum_{i=1}^{N}\log x_i-(p+1)\sum_{i=1}^{N}\log(\kappa/x_i-1)-\beta/p\cdot$$

$$\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}+N\beta/p(\kappa/\gamma-1)^{-p}.$$

The first partial derivatives of the log likelihood function with respect to the parameters are as follows:

$$\frac{\partial L}{\partial \gamma} = \beta\kappa N\gamma^{-2}(\kappa/\gamma-1)^{-(p+1)}[1-e^{\beta/p(\kappa/\gamma-1)^{-p}}]^{-1},$$

$$\frac{\partial L}{\partial \beta} = N\beta^{-1}+Np^{-1}[(\kappa/\gamma-1)^{-p}-1/N\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}]+Np^{-1}(\kappa/\gamma-1)^{-p}\cdot$$

$$[e^{-\beta/p(\kappa/\gamma-1)^{-p}}-1]^{-1},$$

$$\frac{\partial L}{\partial p} = -\beta Np^{-1}(\kappa/\gamma-1)^{-p}[p^{-1}+\log(\kappa/\gamma-1)]-\sum_{i=1}^{N}\log(\kappa/x_i-1)$$

$$+\beta p^{-1}[p^{-1}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}+\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}\log(\kappa/x_i-1)]$$

$$-N\beta p^{-1}[e^{-\beta/p(\kappa/\gamma-1)^{-p}}-1]^{-1}(\kappa/\gamma-1)^{-p}[p^{-1}+\log(\kappa/\gamma-1)]$$

and

$$\frac{\partial L}{\partial \kappa} = [e^{\beta/p(\kappa/\gamma-1)^{-p}}-1]^{-1}+N\kappa^{-1}-(p+1)\sum_{i=1}^{N}(\kappa-x_i)^{-1}$$

$$+\beta\sum_{i=1}^{N}x_i^{-1}(\kappa/x_i-1)^{-(p+1)}.$$

One finds that the maximum likelihood estimate of γ is still $x_{(1)}$ and one must consider the conditional likelihood function given that $\hat{\gamma}$ is $x_{(1)}$. There is no explicit solution possible for any of the other parameter estimates. One may easily go on to find the elements of the matrix of second partial derivatives. They are

$$\frac{\partial^2 L}{\partial \beta^2} = -N\beta^{-2}+Np^{-2}(\kappa/\hat{\gamma}-1)^{-2p}e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}[e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-2},$$

$$\frac{\partial^2 L}{\partial \beta \partial p} = \frac{\partial^2 L}{\partial p \partial \beta} = -Np^{-1}(\kappa/\hat{\gamma}-1)^{-p}[p^{-1}+\log(\kappa/\hat{\gamma}-1)]$$

$$+p^{-1}\sum_{i=1}^{N}\{(\kappa/x_i-1)^{-p}[p^{-1}+\log(\kappa/x_i-1)]\}$$

$$-Np^{-1}(\kappa/\hat{\gamma}-1)^{-p}[p^{-1}+\log(\kappa/\hat{\gamma}-1)][e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-1}$$

$$-N\beta p^{-2}(\kappa/\hat{\gamma}-1)^{-2p}[p^{-1}+\log(\kappa/\hat{\gamma}-1)]e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}$$

$$[e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-2},$$

$$\frac{\partial^2 L}{\partial\beta\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial\beta} = -p^{-1}(\kappa/\hat{\gamma}-1)^{-p}e^{\beta/p(\kappa/\hat{\gamma}-1)^{-p}}[e^{\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-2}$$

$$+\sum_{i=1}^{N} x_i^{-1}(\kappa/x_i-1)^{-(p+1)},$$

$$\frac{\partial^2 L}{\partial p^2} = 2\beta Np^{-3}(\kappa/\hat{\gamma}-1)^{-p}+2\beta Np^{-2}(\kappa/\hat{\gamma}-1)^{-p}\log(\kappa/\hat{\gamma}-1)$$

$$+\beta Np^{-1}(\kappa/\hat{\gamma}-1)^{-p}[\log(\kappa/\hat{\gamma}-1)]^2-2\beta p^{-3}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}$$

$$-2\beta p^{-2}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}\log(\kappa/x_i-1)$$

$$-\beta p^{-1}\sum_{i=1}^{N}(\kappa/x_i-1)^{-p}[\log(\kappa/x_i-1)]^2$$

$$+N[e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-2}e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}[\beta p^{-2}(\kappa/\hat{\gamma}-1)^{-p}$$

$$+\beta p^{-1}(\kappa/\hat{\gamma}-1)^{-p}\log(\kappa/\hat{\gamma}-1)]^2+N[e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-1}\cdot$$

$$\{2\beta p^{-3}(\kappa/\hat{\gamma}-1)^{-p}+2\beta p^{-2}(\kappa/\hat{\gamma}-1)^{-p}\log(\kappa/\hat{\gamma}-1)$$

$$+\beta p^{-1}(\kappa/\hat{\gamma}-1)^{-p}[\log(\kappa/\hat{\gamma}-1)]^2\},$$

$$\frac{\partial^2 L}{\partial p\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial p} = \beta p^{-1}(\kappa/\hat{\gamma}-1)^{-p}[p^{-1}+\log(\kappa/\hat{\gamma}-1)]e^{\beta/p(\kappa/\hat{\gamma}-1)^{-p}}$$

$$[e^{\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-2}-\sum_{i=1}^{N}(\kappa-x_i)^{-1}-\beta\sum_{i=1}^{N}x_i^{-1}$$

$$(\kappa/x_i-1)^{-(p+1)}\log(\kappa/x_i-1)$$

and

$$\frac{\partial^2 L}{\partial \kappa^2} = -\beta\hat{\gamma}^{-1}(\kappa/\hat{\gamma}-1)^{-(p+1)}e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}[e^{-\beta/p(\kappa/\hat{\gamma}-1)^{-p}}-1]^{-2}$$

$$-N\kappa^{-2}+(p+1)\sum_{i=1}^{N}(\kappa-x_i)^{-2}-\beta(p+1)\sum_{i=1}^{N}x_i^{-2}(\kappa/x_i-1)^{-(p+2)}.$$

Using these derivatives and the method of scoring, one may obtain iterative estimates of the parameters β, p and κ given $\hat{\gamma}$ is $x_{(1)}$.

For the exponential-hyper-Gompertzian size distribution and a positive value of the parameter p, the density function is

$$f(x) = \beta' x^{-1}[\log(\kappa/x)]^{-(p+1)}e^{-\beta'/p\{[\log(\kappa/x)]^{-p}-[\log(\kappa/\gamma)]^{-p}\}}.$$

The likelihood and log likelihood functions respectively are found to be

$$\phi(\gamma,\beta',p,\kappa|\underline{x}) = (\beta')^N\{\prod_{i=1}^{N}x_i^{-1}[\log(\kappa/x_i)]^{-(p+1)}\}\cdot$$

$$e^{-\beta'/p\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}-N[\log(\kappa/\gamma)]^{-p}}$$

and

$$L(\gamma,\beta',p,\kappa|\underline{x}) = N\log\beta'-\sum_{i=1}^{N}\log x_i-(p+1)\sum_{i=1}^{N}\log[\log(\kappa/x_i)]$$

$$-\beta'/p\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}+N\beta'/p[\log(\kappa/\gamma)]^{-p}.$$

To obtain estimates of the parameters, one may first display the first partial derivatives with respect to the parameters. For simplicity, β' is written as β:

$$\frac{\partial L}{\partial \gamma} = N\beta\gamma^{-1}[\log(\kappa/\gamma)]^{-(p+1)},$$

$$\frac{\partial L}{\partial \beta} = N\beta^{-1}-p^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}+Np^{-1}[\log(\kappa/\gamma)]^{-p},$$

$$\frac{\partial L}{\partial p} = \beta p^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p} - N[\log(\kappa/\gamma)]^{-p}\} - \sum_{i=1}^{N}\log[\log(\kappa/x_i)]$$

$$+\beta p^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\log[\log(\kappa/x_i)] - N[\log(\kappa/\gamma)]^{-p}\cdot$$

$$\log[\log(\kappa/\gamma)]\}$$

and

$$\frac{\partial L}{\partial \kappa} = -N\beta\kappa^{-1}[\log(\kappa/\gamma)]^{-(p+1)} - (p+1)\kappa^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-1}$$

$$+\beta\kappa^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)}.$$

The maximum likelihood estimate for γ is, as in the cases already discussed, $x_{(1)}$. Then if the conditional likelihood is considered, one can find an explicit solution for $\hat{\beta}$ in terms of the estimates of the other parameters and the observations as

$$\hat{\beta} = N\hat{p}\{\sum_{i=1}^{N}[\log(\hat{\kappa}/x_i)]^{-p} - N[\log(\hat{\kappa}/\hat{\gamma})]^{-\hat{p}}\}^{-1}.$$

No explicit solution of this kind for $\hat{p}$ or $\hat{\kappa}$ is possible. The elements of the matrix of second partials are

$$\frac{\partial^2}{\partial \beta^2} = -N\beta^{-2}$$

$$\frac{\partial^2 L}{\partial \beta \partial p} = \frac{\partial^2 L}{\partial p \partial \beta} = p^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p} - N[\log(\kappa/\hat{\gamma})]^{-p}\}$$

$$+p^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\log[\log(\kappa/x_i)] - N[\log(\kappa/\hat{\gamma})]^{-p}\cdot$$

$$\log[\log(\kappa/\hat{\gamma})]\},$$

$$\frac{\partial^2 L}{\partial\beta\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial\beta} = \kappa^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)} - N[\log(\kappa/\hat{\gamma})]^{-(p+1)}\},$$

$$\frac{\partial^2 L}{\partial p^2} = 2\beta p^{-3}\{N[\log(\kappa/\hat{\gamma})]^{-p} - \sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\}$$

$$+2\beta p^{-2}\{N[\log(\kappa/\hat{\gamma})]^{-p}\log[\log(\kappa/\hat{\gamma})] - \sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\cdot$$

$$\log[\log(\kappa/x_i)]\} + \beta p^{-1}\{N[\log(\kappa/\hat{\gamma})]^{-p}\{\log[\log(\kappa/\gamma)]\}^2$$

$$-\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\{\log[\log(\kappa/x_i)]\}^2\},$$

$$\frac{\partial^2 L}{\partial p\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial p} = -\beta\kappa^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)}\log[\log(\kappa/x_i)]$$

$$-N[\log(\kappa/\hat{\gamma})]^{-(p+1)}\log[\log(\kappa/\hat{\gamma})]\} - \kappa^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-1}$$

and

$$\frac{\partial^2 L}{\partial\kappa^2} = -\kappa^{-2}\beta\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)} - N[\log(\kappa/\hat{\gamma})]^{-(p+1)}\}$$

$$+\kappa^{-2}(p+1)\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-1} - (p+1)\kappa^{-2}\beta\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+2)}$$

$$-N[\log(\kappa/\hat{\gamma})]^{-(p+2)}\} + (p+1)\kappa^{-2}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-2}$$

If p is between zero and -1, the form of the density is

$$f(x) = \{1-e^{\beta/p[\log(\kappa/\gamma)]^{-p}}\}^{-1}\beta x^{-1}[\log(\kappa/x)]^{-(p+1)}\cdot$$

$$e^{-\beta/p\{[\log(\kappa/x)]^{-p} - [\log(\kappa/\gamma)]^{-p}\}}.$$

In this case, the likelihood and log likelihood functions are written as

$$\phi(\gamma,\beta,p,\kappa|\underline{x}) = \{1-e^{\beta/p[\log(\kappa/\gamma)]^{-p}}\}^{-N}\beta^N\{\prod_{i=1}^{N} x_i^{-1} \cdot$$

$$[\log(\kappa/x_i)]^{-(p+1)}\} \cdot$$

$$e^{-\beta/p\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}-N[\log(\kappa/\gamma)]^{-p}\}}$$

and

$$L(\gamma,\beta,p,\kappa|\underline{x}) = -N\log\{1-e^{\beta/p[\log(\kappa/\gamma)]^{-p}}\}+N\log\beta-\sum_{i=1}^{N}\log x_i$$

$$-(p+1)\sum_{i=1}^{N}\log[\log(\kappa/x_i)]-\beta/p\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}$$

$$+\beta/pN[\log(\kappa/\gamma)]^{-p}.$$

As before, one may easily display the first partial derivatives, which are:

$$\frac{\partial L}{\partial\gamma} = N\beta\gamma^{-1}[\log(\kappa/\gamma)]^{-(p+1)}\{1-e^{\beta/p[\log(\kappa/\gamma)]^{-p}}\}^{-1},$$

$$\frac{\partial L}{\partial\beta} = N\beta^{-1}+Np^{-1}[\log(\kappa/\gamma)]^{-p}-p^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}+Np^{-1}\cdot$$

$$[\log(\kappa/\gamma)]^{-p}\{e^{-\beta/p[\log(\kappa/\gamma)]^{-p}}-1\}^{-1},$$

$$\frac{\partial L}{\partial p} = \beta p^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}-N[\log(\kappa/\gamma)]^{-p}\}-\sum_{i=1}^{N}\log[\log(\kappa/x_i)]$$

$$+\beta p^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\log[\log(\kappa/x_i)]-N[\log(\kappa/\gamma)]^{-p}\cdot$$

$$\log[\log(\kappa/\gamma)]\}-N\{e^{-\beta/p[\log(\kappa/\gamma)]^{-p}}-1\}^{-1}\beta p^{-1}[\log(\kappa/\gamma)]^{-p}\cdot$$

$$\{p^{-1}+\log[\log(\kappa/\gamma)]\}$$

and

$$\frac{\partial L}{\partial \kappa} = -N\beta\kappa^{-1}[\log(\kappa/\gamma)]^{-(p+1)}-N\beta\kappa^{-1}[\log(\kappa/\gamma)]^{-(p+1)}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\gamma)]^{-p}}-1\}^{-1}-(p+1)\kappa^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-1}$$

$$+\beta\kappa^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)}.$$

By examining the likelihood function, one sees that the estimate for γ must be as large as possible subject to the restriction that it is not greater than the smallest observation, $x_{(1)}$. So, $\hat{\gamma}$ is $x_{(1)}$. The other parameter estimates should maximize the conditional likelihood function given $\hat{\gamma}$ is $x_{(1)}$. No explicit solutions for $\hat{\beta}$, $\hat{p}$ and $\hat{\kappa}$ are possible. One may solve for them iteratively, however, using the method of scoring. The second partial derivatives, for this purpose, are as follows:

$$\frac{\partial^2 L}{\partial \beta^2} = -N\beta^{-2}+Np^{-2}[\log(\kappa/\hat{\gamma})]^{-2p}e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-2},$$

$$\frac{\partial^2 L}{\partial \beta \partial p} = \frac{\partial^2 L}{\partial p \partial \beta} = p^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}-N[\log(\kappa/\hat{\gamma})]^{-p}\}$$

$$+p^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\log[\log(\kappa/x_i)]-N[\log(\kappa/\hat{\gamma})]^{-p}\cdot$$

$$\log[\log(\kappa/\hat{\gamma})]\}-Np^{-1}[\log(\kappa/\hat{\gamma})]^{-p}\{p^{-1}+\log[\log(\kappa/\hat{\gamma})]\}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-1}-Np^{-2}\beta[\log(\kappa/\hat{\gamma})]^{-2p}.$$

$$\{p^{-1}+\log[\log(\kappa/\hat{\gamma})]\}e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-2},$$

$$\frac{\partial^2 L}{\partial\beta\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial\beta} = \kappa^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)}-N[\log(\kappa/\hat{\gamma})]^{-(p+1)}\}$$

$$-N\kappa^{-1}[\log(\kappa/\hat{\gamma})]^{-(p+1)}\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-1}$$

$$-N\beta p^{-1}\kappa^{-1}[\log(\kappa/\hat{\gamma})]^{-(2p+1)}e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-2},$$

$$\frac{\partial^2 L}{\partial p^2} = -2\beta p^{-3}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}-N[\log(\kappa/\hat{\gamma})]^{-p}\}$$

$$-2\beta p^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\log[\log(\kappa/x_i)]-N[\log(\kappa/\hat{\gamma})]^{-p}\cdot$$

$$\log[\log(\kappa/\hat{\gamma})]\}-\beta p^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-p}\{\log[\log(\kappa/x_i)]\}^2$$

$$-N[\log(\kappa/\hat{\gamma})]^{-p}\{\log[\log(\kappa/\hat{\gamma})]\}^2\}+N\beta^2p^{-2}[\log(\kappa/\hat{\gamma})]^{-2p}\cdot$$

$$\{p^{-1}+\log[\log(\kappa/\hat{\gamma})]\}^2e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\cdot$$

$$\{e^{-\beta/p\ \log(\kappa/\hat{\gamma})\ ^{-p}}-1\}^{-2}+N\{e^{-\beta/p\ \log(\kappa/\hat{\gamma})\ ^{-p}}-1\}^{-1}\cdot$$

$$\beta p^{-1}[\log(\kappa/\hat{\gamma})]^{-p}\{2p^{-2}+2p^{-1}\log[\log(\kappa/\hat{\gamma})]+\{\log[\log$$

$$(\kappa/\hat{\gamma})]\}^2\},$$

$$\frac{\partial^2 L}{\partial p\partial\kappa} = \frac{\partial^2 L}{\partial\kappa\partial p} = -\beta\kappa^{-1}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)}\log[\log(\kappa/x_i)]$$

$$-N[\log(\kappa/\hat{\gamma})]^{-(p+1)}\log[\log(\kappa/\hat{\gamma})]\}+N\beta^2p^{-1}\kappa^{-1}\cdot$$

$$[\log(\kappa/\hat{\gamma})]^{-(2p+1)}\{p^{-1}+\log[\log(\kappa/\hat{\gamma})]\}e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-2}+N\beta\kappa^{-1}[\log(\kappa/\hat{\gamma})]^{-(p+1)}\cdot$$

$$\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-1}-\kappa^{-1}\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-1}$$

and

$$\frac{\partial^2 L}{\partial\kappa^2} = N\beta\kappa^{-2}\{[\log(\kappa/\hat{\gamma})]^{-(p+1)}+(p+1)[\log(\kappa/\hat{\gamma})]^{-(p+2)}\cdot$$

$$\{1-e^{\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\}^{-1}+N\beta^2\kappa^{-2}[\log(\kappa/\hat{\gamma})]^{-2(p+1)}\cdot$$

$$e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}\{e^{-\beta/p[\log(\kappa/\hat{\gamma})]^{-p}}-1\}^{-2}$$

$$+(p+1)\kappa^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-1}+\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-2}\}$$

$$-\beta\kappa^{-2}\{\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+1)}+(p+1)\sum_{i=1}^{N}[\log(\kappa/x_i)]^{-(p+2)}\}.$$

Now for the exponential-generic curve with its five parameters, the equations for maximum likelihood estimation become exceedingly complicated. For that reason, only the first partial derivatives are given (or, indeed, obtained) in both the positive and the negative p cases. Of course, to estimate the parameters by the method of scoring, one may either obtain them or construct the matrix as the product of the column vector of first partial derivatives and its transpose. For positive values of p, the density is written as

$$f(x) = \beta x^{-1}(\kappa/x)^{n}[(\kappa/x)^{n}-1]^{-(p+1)}\cdot$$

$$e^{-\beta/(np)\{[(\kappa/x)^{n}-1]^{-p}-[(\kappa/\gamma)^{n}-1]^{-p}\}}\cdot$$

Then the likelihood function is

$$\phi(\gamma,\beta,n,p,\kappa|\underline{x}) = \beta^N\{\prod_{i=1}^{N} x_i^{-1}(\kappa/x_i)^n[(\kappa/x_i)^n-1]^{-(p+1)}\}$$

$$e^{-\beta/(np)\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}-N[(\kappa/\hat{\gamma})^n-1]^{-p}\}}$$

and the log likelihood function is

$$L(\gamma,\beta,n,p,\kappa|\underline{x}) = N\log\beta-\sum_{i=1}^{N}\log x_i+nN\log\kappa-n\sum_{i=1}^{N}\log x_i$$

$$-(p+1)\sum_{i=1}^{N}\log[(\kappa/x_i)^n-1]-\beta/(np)\cdot$$

$$\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}+\beta N/(np)[(\kappa/\gamma)^n-1]^{-p}.$$

The first partial derivatives of the log likelihood function with respect to the parameters are:

$$\frac{\partial L}{\partial\gamma} = \beta N\gamma^{-1}(\kappa/\gamma)^n[(\kappa/\gamma)^n-1]^{-(p+1)},$$

$$\frac{\partial L}{\partial\beta} = N\beta^{-1}-1/(np)\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}-N[(\kappa/\gamma)^n-1]^{-p}\},$$

$$\frac{\partial L}{\partial n} = \beta n^{-2}p^{-1}\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}-N[(\kappa/\gamma)^n-1]^{-p}\}$$

$$+\beta n^{-1}\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-(p+1)}(\kappa/x_i)^n\log(\kappa/x_i)$$

$$-N[(\kappa/\gamma)^n-1]^{-(p+1)}(\kappa/\gamma)^n\log(\kappa/\gamma)\}+N\log\kappa$$

$$-\sum_{i=1}^{N}\log x_i-(p+1)\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-1}(\kappa/x_i)^n\log(\kappa/x_i),$$

$$\frac{\partial L}{\partial p} = \beta n^{-1}p^{-2}\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}-N[(\kappa/\gamma)^n-1]^{-p}\}+\beta n^{-1}p^{-1}.$$

$$\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}\log[(\kappa/x_i)^n-1]-N[(\kappa/\gamma)^n-1]^{-p}\cdot$$

$$\log[(\kappa/\gamma)^n-1]\}-\sum_{i=1}^{N}\log[(\kappa/\gamma)^n-1]$$

and

$$\frac{\partial L}{\partial\kappa}=\beta\{\sum_{i=1}^{N}x_i^{-1}(\kappa/x_i)^{n-1}[(\kappa/x_i)^n-1]^{-(p+1)}-N\gamma^{-1}(\kappa/\gamma)^{n-1}\cdot$$

$$[(\kappa/\gamma)^n-1]^{-(p+1)}\}+Nn\kappa^{-1}-(p+1)n\sum_{i=1}^{N}x_i^{-1}(\kappa/x_i)^{n-1}\cdot$$

$$[(\kappa/x_i)^n-1]^{-1}.$$

The maximum likelihood estimate of γ remains $x_{(1)}$ as in all the previous cases. An explicit solution for $\hat{\beta}$ in terms of the observations and the other parameter estimates is obtainable as

$$\hat{\beta}=\hat{n}\hat{p}\{1/N\sum_{i=1}^{N}[(\hat{\kappa}/x_i)^{\hat{n}}-1]^{-\hat{p}}-[(\hat{\kappa}/\hat{\gamma})^{\hat{n}}-1]^{-\hat{p}}\}^{-1}.$$

It is not possible to solve for the other parameter estimates explicitly. Of course, only the conditional likelihood given $\hat{\gamma}$ is $x_{(1)}$ should be considered in solving the equations iteratively.

If p is between zero and -1, the density function, one recalls, is

$$f(x)=\{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\beta x^{-1}(\kappa/x)^n[(\kappa/x)^n-1]^{-(p+1)}\cdot$$

$$e^{-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}}.$$

The likelihood and log likelihood functions for this case are, respectively:

$$\phi(\gamma,\beta,n,p,\kappa|\underline{x}) = \{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-N}\beta^N\{\prod_{i=1}^{N} x_i^{-1}(\kappa/x_i)^n \cdot$$

$$[(\kappa/x_i)^n-1]^{-(p+1)}\}$$

$$e^{-\beta/(np)\{\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}-N[(\kappa/\gamma)^n-1]^{-p}\}}$$

and

$$L(\gamma,\beta,n,p,\kappa|\underline{x}) = -N\log\{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}+N\log\beta$$

$$-\sum_{i=1}^{N}\log x_i+nN\log\kappa-n\sum_{i=1}^{N}\log x_i-(p+1)\cdot$$

$$\sum_{i=1}^{N}\log[(\kappa/x_i)^n-1]-\beta/(np)\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}$$

$$+N\beta/(np)[(\kappa/\gamma)^n-1]^{-p}$$

The first partial derivatives of the log likelihood function with respect to the parameters are:

$$\frac{\partial L}{\partial\gamma} = N\beta\gamma^{-1}(\kappa/\gamma)^n[(\kappa/\gamma)^n-1]^{-(p+1)}\{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1},$$

$$\frac{\partial L}{\partial\beta} = N\beta^{-1}-n^{-1}p^{-1}\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}+Nn^{-1}p^{-1}[(\kappa/\gamma)^n-1]^{-p}\cdot$$

$$\{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1},$$

$$\frac{\partial L}{\partial n} = -N\{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}\beta n^{-1}[(\kappa/\gamma)^n-1]^{-p}\cdot$$

$$\{n^{-1}p^{-1}+[(\kappa/\gamma)^n-1]^{-1}(\kappa/\gamma)^n\log(\kappa/\gamma)\}+N\log\kappa$$

$$-\sum_{i=1}^{N}\log x_i-(p+1)\sum_{i=1}^{N}[(\kappa/x_i)^n-1](\kappa/x_i)^n\log(\kappa/x_i)$$

$$+\beta n^{-1}\{n^{-1}p^{-1}\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}+\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-(p+1)}\cdot$$

$$(\kappa/x_i)^n\log(\kappa/x_i)\},$$

$$\frac{\partial L}{\partial p} = -N\beta n^{-1}p^{-1}\{p^{-1}+\log[(\kappa/\gamma)^n-1]\}[(\kappa/\gamma)^n-1]^{-p}\cdot$$

$$[1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}]^{-1}-\sum_{i=1}^{N}\log[(\kappa/x_i)^n-1]+\beta n^{-1}p^{-1}\cdot$$

$$\{p^{-1}\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}+\sum_{i=1}^{N}[(\kappa/x_i)^n-1]^{-p}\log[(\kappa/x_i)^n-1]\}$$

and

$$\frac{\partial L}{\partial \kappa} = -N\beta\kappa^{-1}\{1-e^{\beta/(np)[(\kappa/\gamma)^n-1]^{-p}}\}^{-1}[(\kappa/\gamma)^n-1]^{-(p+1)}(\kappa/\gamma)^n$$

$$+Nn\kappa^{-1}-(p+1)n\kappa^{-1}\sum_{i=1}^{N}(\kappa/x_i)^n[(\kappa/x_i)^n-1]^{-1}$$

$$+\beta\kappa^{-1}\sum_{i=1}^{N}(\kappa/x_i)^n[(\kappa/x_i)^n-1]^{-(p+1)}.$$

Again the maximum likelihood solution for $\hat{\gamma}$ is $x_{(1)}$. The other parameters must be estimated conditional on $\hat{\gamma}$ being $x_{(1)}$ and in an iterative manner.

The last of the upper curves, the Weibull, has a simpler set of maximum likelihood equations. The density, for a positive value of α, is

$$f(x) = \beta''x^{\alpha-1}e^{-\beta''/\alpha(x^\alpha-\gamma^\alpha)}.$$

Then the likelihood and log likelihood function are

$$\phi(\gamma,\beta'',\alpha|\underline{x}) = (\beta'')^N(\prod_{i=1}^{N} x_i^{\alpha-1})e^{-\beta''/\alpha\sum_{i=1}^{N}(x_i^\alpha-\gamma^\alpha)}$$

and

$$L(\gamma,\beta'',\alpha|\underline{x}) = N \log \beta''+(\alpha-1) \sum_{i=1}^{N} \log x_i$$

$$-\beta''/\alpha \sum_{i=1}^{N} x_i+(\beta''N)/\alpha\gamma^{\alpha}.$$

Differentiating, one finds the first partial derivatives with respect to the parameters:

$$\frac{\partial L}{\partial \gamma} = \beta''N\gamma^{\alpha-1},$$

$$\frac{\partial L}{\partial \beta''} = N(\beta'')^{-1}-\alpha^{-1} \sum_{i=1}^{N} x_i^{\alpha}+\alpha^{-1}N\alpha^{\gamma},$$

and

$$\frac{\partial L}{\partial \gamma} = \sum_{i=1}^{N} \log x_i+\beta\alpha^{-2}[\sum_{i=1}^{N} x^{\alpha}-N\gamma^{\alpha}]-\beta\alpha^{-1}[\sum_{i=1}^{N} x_i^{\alpha} \log x_i-N\gamma^{\alpha} \log \gamma].$$

Now it is apparent that the estimate for γ in the Weibull has two cases according to whether α is known to be greater than unity or not. If α is greater than 1, $\hat{\gamma}$ must be 0. If α is less than or equal to 1, $\hat{\gamma}$ is $x_{(1)}$ as for all the other size curves developed in this dissertation. An explicit expression for β'' in terms of $\hat{\gamma}$ and $\hat{\alpha}$ is obtainable as

$$\hat{\beta}'' = \hat{\alpha}\{1/N \sum_{i=1}^{N} x_i^{\hat{\alpha}}-\hat{\gamma}^{\hat{\alpha}}\}^{-1}.$$

It is not possible to get an explicit expression for $\hat{\alpha}$ in terms of $\hat{\beta}''$, $\hat{\gamma}$ and the observations only. In order to estimate $\hat{\beta}''$ and $\hat{\alpha}$ conditionally given $\hat{\gamma}$ is $x_{(1)}$ by the method of scoring, one can easily find the second partial derivatives as

$$\frac{\partial^2 L}{\partial(\beta'')^2} = -N(\beta'')^{-2}$$

$$\frac{\partial^2}{\partial\beta''\partial\alpha} = \frac{\partial^2 L}{\partial\alpha\partial\beta''} = \alpha^{-2}\{\sum_{i=1}^{N} x_i^\alpha - N\hat{\gamma}^\alpha\} - \alpha^{-1}\cdot$$

$$\{\sum_{i=1}^{N} x_i^\alpha \log x_i^\alpha - N\hat{\gamma}^\alpha \log \hat{\gamma}\}$$

and

$$\frac{\partial^2 L}{\partial\alpha^2} = -2\beta''\alpha^{-3}[\sum_{i=1}^{N} x_i^\alpha - N\hat{\gamma}^\alpha] + 2\beta''\alpha^{-2}[\sum_{i=1}^{N} x_i^\alpha \log x_i - N\hat{\gamma}^\alpha \log \hat{\gamma}]$$

$$-\beta''\alpha^{-1}[\sum_{i=1}^{N} x_i^\alpha(\log x_i)^2 - N\hat{\gamma}^\alpha(\log \hat{\gamma})^2].$$

If, however, α is known to be greater than 1 (a reasonable guess if the empirical distribution function appears to be that of a J-shaped density function), $\hat{x}$ is 0. Then the conditional log likelihood becomes

$$L(\beta'',\alpha|\hat{\gamma}=0,\underline{x}) = N \log \beta'' + (\alpha-1)\sum_{i=1}^{N} \log x_i - \beta''\alpha^{-1}\sum_{i=1}^{N} x_i^\alpha$$

and the first partial derivatives become

$$\frac{\partial L}{\partial\beta''} = N(\beta'')^{-1} - \alpha^{-1}\sum_{i=1}^{N} x_i^\alpha$$

and

$$\frac{\partial L}{\partial\alpha} = \sum_{i=1}^{N} \log x_i + \beta'' a^{-2}\sum_{i=1}^{N} x_i^\alpha - \beta''\alpha^{-1}\sum_{i=1}^{N} x_i^\alpha \log x_i.$$

So solving for β'', one finds

$$\hat{\beta}'' = \hat{\alpha}(1/N \sum_{i=1}^{N} x_i^{\hat{\alpha}})^{-1}.$$

There is still no explicit solution for $\hat{\alpha}$. The second partial derivatives become

$$\frac{\partial^2 L}{\partial(\beta'')^2} = -N(\beta'')^{-2},$$

$$\frac{\partial^2 L}{\partial\beta''\partial\alpha} = \frac{\partial^2 L}{\partial\alpha\partial\beta''} = \alpha^{-2}\sum_{i=1}^{N} x_i^{\alpha} - \alpha^{-1}\sum_{i=1}^{N} x_i^{\alpha}\log x_i$$

and

$$\frac{\partial^2 L}{\partial\alpha^2} = -2\beta''\alpha^{-3}\sum_{i=1}^{N} x_i^{\alpha} + 2\beta\alpha^{-2}\sum_{i=1}^{N} x_i^{\alpha}\log x_i$$

$$-\beta''\alpha^{-1}\sum_{i=1}^{N} x_i^{\alpha}(\log x_i)^2.$$

If α is less than zero, the form of the density is changed to

$$f(x) = \{1-e^{\beta''/\alpha\gamma^{\alpha}}\}^{-1}\beta'' x^{\alpha-1} e^{-\beta''/\alpha(x^{\alpha}-\gamma^{\alpha})}.$$

With the corresponding change, the likelihood and log likelihood functions become, respectively,

$$\phi(\gamma,\beta'',\alpha|\underline{x}) = \{1-e^{\beta''/\alpha\gamma^{\alpha}}\}^{-N}(\beta'')^{N}\left(\prod_{i=1}^{N} x_i^{\alpha-1}\right)\cdot$$

$$e^{-\beta''/\alpha\left(\sum_{i=1}^{N} x_i^{\alpha} - N\gamma^{\alpha}\right)}$$

and

$$L(\gamma,\beta'',\alpha|\underline{x}) = -N\log\{1-e^{\beta''/\alpha\gamma^{\alpha}}\} + N\log\beta'' + (\alpha-1)\cdot$$

$$\sum_{i=1}^{N}\log x_i - \beta''/\alpha\left[\sum_{i=1}^{N} x_i^{\alpha} - N\gamma^{\alpha}\right].$$

Taking the first partial derivatives, one obtains

$$\frac{\partial L}{\partial\gamma} = N\beta''\gamma^{\alpha-1}\{1-e^{\beta''/\alpha\gamma^{\alpha}}\}^{-1},$$

$$\frac{\partial L}{\partial \beta} = N(\beta'')^{-1} + N\alpha^{-1}\gamma^{\alpha}\{1-e^{\beta''/\alpha\gamma^{\alpha}}\}^{-1} - \alpha^{-1}\sum_{i=1}^{N} x_i^{\alpha}$$

and

$$\frac{\partial L}{\partial \alpha} = -N\{1-e^{\beta''/\alpha\gamma^{\alpha}}\}^{-1}\beta''\alpha^{-1}\{\alpha^{-1}\gamma^{\alpha} - \gamma^{\alpha}\log\gamma\} + \sum_{i=1}^{N}\log x_i$$

$$+\beta''\alpha^{-1}\{\alpha^{-1}\sum_{i=1}^{N} x_i^{\alpha} - \sum_{i=1}^{N} x_i^{\alpha}\log x_i\}.$$

It is clear that for negative α, $\hat{\alpha}$ is $x_{(1)}$. No explicit solutions for $\hat{\beta}$ or $\hat{\alpha}$ are able to be found. The second partial derivatives of the conditional likelihood given $\hat{\gamma}$ is $x_{(1)}$ are as follows:

$$\frac{\partial^2 L}{\partial(\beta'')^2} = -N(\beta'')^{-2} + N\hat{\gamma}^{2\alpha}e^{\beta''/\alpha\hat{\gamma}^{\alpha}}\alpha^{-2}[1-e^{\beta''/\alpha\hat{\gamma}^{\alpha}}]^{-2}$$

$$\frac{\partial^2}{\partial\beta''\partial\alpha} = \frac{\partial^2 L}{\partial\alpha\partial\beta''} = \alpha^{-2}\sum_{i=1}^{N} x_i^{\alpha} - \alpha^{-1}\sum_{i=1}^{N} x_i^{\alpha}\log x_i$$

$$-N\alpha^{-1}\hat{\gamma}^{\alpha}(\alpha^{-1} - \log\hat{\gamma})\{1-e^{\beta''/\alpha\hat{\gamma}^{\alpha}}\}^{-1}$$

$$-N\alpha^{-2}\beta''\hat{\gamma}^{2\alpha}e^{\beta''/\alpha\hat{\gamma}^{\alpha}}\{\alpha^{-1} - \log\hat{\gamma}\}\{1-e^{\beta''/\alpha\hat{\gamma}^{\alpha}}\}^{-2}$$

and

$$\frac{\partial^2 L}{\partial\alpha^2} = N\{1-e^{\beta''/\alpha\hat{\gamma}^{\alpha}}\}^{-2}e^{\beta''/\alpha\hat{\gamma}^{\alpha}}(\beta'')^{2}\alpha^{-2}\hat{\gamma}^{2\alpha}\{\alpha^{-1} - \log\hat{\gamma}\}^{2}$$

$$-N\{1-e^{\beta''/\alpha\hat{\gamma}^{\alpha}}\}^{-1}\beta''\alpha^{-1}\hat{\gamma}^{\alpha}\{-2\alpha^{-2} + 2\alpha^{-1}\log\hat{\gamma} - (\log\hat{\gamma})^{2}\}$$

$$-2\beta''\alpha^{-3}\sum_{i=1}^{N} x_i^{\alpha} + 2\beta''\alpha^{-2}\sum_{i=1}^{N} x_i^{\alpha}\log x_i - \beta''\alpha^{-1}\sum_{i=1}^{N} x_i^{\alpha}(\log x_i)^{2}.$$

These derivatives may be used to produce the conditional maximum likelihood estimates of β'' and α by the method of scoring.

C. Least Squares

Since all of these size distributions have a simple, closed form representation of their respective distribution functions, another possibility for a fitting method is to minimize the sum of squared weighted deviations of the estimated parametric distribution function evaluated at the ordered observation points from its expectation, the empirical cumulative distribution function, by means of nonlinear least squares. Some notation must be introduced here to facilitate the discussion. By the column vector $\underline{\theta}$ is meant the vector of parameters for the particular size distribution under consideration. For the Pareto, $\underline{\theta}$ contains only γ and β. For the exponential-generic curve, $\underline{\theta}$ contains γ, β, n, p and κ. For the Weibull, it has γ, β'' and α. If $x_{(r)}$ is the rth observation in the ordered sample arranged from smallest to largest, then the empirical distribution function, $S(x_{(r)})$, has the value r/(N+1) corresponding to that $x_{(r)}$. In case of tied observations, the average rank is assigned to those in the tie. $F(x_{(r)})$ is the size distribution evaluated for $x_{(r)}$. It is dependent upon the parameters included and could well be written as $F(x_{(r)}|\underline{\theta}')$. Let

$$\underline{S} = \begin{bmatrix} S(x_{(1)}) \\ S(x_{(2)}) \\ \vdots \\ S(x_{(N)}) \end{bmatrix} \quad \text{and} \quad \underline{F} = \begin{bmatrix} F(x_{(1)}) \\ F(x_{(2)}) \\ \vdots \\ F(x_{(N)}) \end{bmatrix}$$

Now let $\nabla F(x_{(r)})$ be the column vector of partial derivatives of $F(x_{(r)})$ with respect to the parameters. So one may write $\nabla F(x_{(r)})=\frac{\partial}{\partial \underline{\theta}}F(x_{(r)})$.

An asterisk beside a vector or an element of a vector is used to indicate the evaluation of the expression for a particular choice of $\underline{\theta}$, call it $\underline{\theta}^*$. Accordingly let

$$\underline{F}^* = \underline{F}\big|_{\underline{\theta}=\underline{\theta}^*} \text{ and } \nabla F^*(x_{(r)}) = \nabla F(x_{(r)})\big|_{\underline{\theta}=\underline{\theta}^*}.$$

Now since $F(x)$ is a distribution function, it is well-known that the expectation of $F(x_{(r)})$ is $r/(N+1)$ or $S(x_{(r)})$. The variance-covariance matrix for $F(x_{(1)}), F(x_{(2)}), \ldots, F(x_{(N)})$ is also well-known and has the form

$$V = (N+1)^{-2}(N+2)^{-1} \begin{bmatrix} N & N-1 & N-2 & N-3 & \cdots & 1 \\ N-1 & 2(N-1) & 2(N-2) & 2(N-3) & \cdots & 2 \\ N-2 & 2(N-2) & 3(N-2) & 3(N-3) & \cdots & 3 \\ N-3 & 2(N-3) & 3(N-3) & 4(N-3) & \cdots & 4 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & 4 & \cdots & N \end{bmatrix}.$$

The inverse of the variance-covariance matrix, V^{-1}, has a convenient tridiagonal form which greatly simplifies calculational problems:

$$V^{-1} = (N+1)(N+2) \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ 0 & 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 \end{bmatrix}$$

In terms of the usual nonlinear least squares, the model may be written as

$$\underline{S} = \underline{F} + \underline{\varepsilon}$$

where the vector $\underline{\varepsilon}$ is a vector of errors such that the expectation of $\underline{\varepsilon}$, $E(\underline{\varepsilon})$, is the zero vector $\underline{0}$ and the variance-covariance matrix of the ε_i is V. In other words,

$$\varepsilon(\underline{\varepsilon}) = \underline{0} \text{ and } E(\underline{\varepsilon}\ \underline{\varepsilon}') = V.$$

Now the least squares solution for linear unbiased estimators of the parameters requires the minimization of the expression

$$(\underline{S}-\underline{F})'V^{-1}(\underline{S}-\underline{F}) \text{ or } \underline{\varepsilon}'V^{-1}\underline{\varepsilon}.$$

Since F(x) is a nonlinear function of the parameters, it may be approximated by a Taylor's series expansion about $\underline{\theta}^*$ which is truncated after the first order derivative terms:

$$F(x_{(r)}) \doteq F^*(x_{(r)})+(\underline{\theta}^*-\underline{\theta})'\nabla F^*(x_{(r)}).$$

Then the fit may be made iteratively using one of the many nonlinear fitting techniques descended from the Newton-Raphson approach. Because of the high correlation often encountered among the parameter estimates, a Marquardt-type algorithm is probably to be preferred. If the model is written in terms of the approximation, it becomes

$$\underline{S} = \underline{F}^*+\nabla F^*(\underline{\theta}^*-\underline{\theta})+\underline{\varepsilon}$$

where

$$\nabla F^* = (\nabla F^*(x_{(1)})\nabla F^*(x_{(2)}) \ \ldots \ \nabla F^*(x_{(N)}))'.$$

Then the expression to be minimized by some iterative method is

$$[\underline{S}-\underline{F}^*-\nabla F^*(\underline{\theta}^*-\underline{\theta})]'V^{-1}[\underline{S}-\underline{F}^*- F^*(\underline{\theta}^*-\underline{\theta})].$$

Since the solution to this problem often requires programming which is different from that available in standard computer libraries, one may sometimes wish to settle for an approximate solution which assumes that

V is the identity matrix so that only $\underline{\varepsilon}'\underline{\varepsilon}$ is minimized. It may, however, yield quite satisfactory estimates of the parameters for many sets of data.

Either of these two least squares methods requires the first partial derivatives of the distribution function. These are given for all the size distributions studies in this paper. The second comment to be made is that, as with the maximum likelihood estimation procedures, one would prefer least squares solutions which are restricted so that the estimate of γ is always less than or equal to $x_{(1)}$ and that of κ is greater than $x_{(N)}$. It is not in general true that the least squares estimate for γ is $x_{(1)}$. If the estimate for γ appears to be stuck on the bound at $x_{(1)}$, it is necessary to consider the conditional least squares solution for the other parameters given that $\tilde{\gamma}$ is $x_{(1)}$, in the same manner as for the maximum likelihood solutions. $\tilde{\gamma}$ must also be greater than zero. If the estimate is stuck on that bound, $\tilde{\gamma}$ may be arbitrarily set to some very small positive number and the conditional solution there obtained. If $\tilde{\kappa}$ appears to be stuck on the bound at $x_{(N)}$, it may be set to $x_{(N)}$ and the conditional solutions found; however, in this case, any calculations at $x_{(N)}$ itself must be handled as special points to be separately calculated since the usual computer routines can not proceed for the limiting forms involved. An alternative choice is to set $\tilde{\kappa}$ instead at $x_{(N)}$ plus some very small positive number and to obtain the least squares conditional solutions for the other parameters at this bound. In the following pages, the function and its partial derivatives are given for all of these size distributions.

For the Pareto, the distribution function is

$$F(x) = 1-(\gamma^{-1}x)^{-\beta}.$$

Taking the first partial derivatives with respect to γ and β, one finds them to be

$$\frac{\partial F(x)}{\partial \gamma} = -\beta\gamma^{-1}(\gamma^{-1}x)^{-\beta}$$

and

$$\frac{\partial F(x)}{\partial \beta} = (\gamma^{-1}x)^{-\beta}\log(\gamma^{-1}x).$$

For the exponential-logistic, the distribution function is

$$F(x) = 1-(\kappa/\gamma-1)^{-\beta}(\kappa/x-1)^{\beta}.$$

and the three first partial derivatives are

$$\frac{\partial F(x)}{\partial \gamma} = -\beta\kappa\gamma^{-2}(\kappa/\gamma-1)^{-(\beta+1)}(\kappa/x-1)^{\beta},$$

$$\frac{\partial F(x)}{\partial \beta} = (\kappa/\gamma-1)^{-\beta}(\kappa/x-1)^{\beta}[\log(\kappa-\gamma)+\log x-\log \gamma-\log(\kappa-x)]$$

and

$$\frac{\partial F(x)}{\partial \kappa} = \beta\gamma^{-1}(\kappa/\gamma-1)^{-(\beta+1)}(\kappa/x-1)^{\beta}-\beta x^{-1}(\kappa/\gamma-1)^{-\beta}(\kappa/x-1)^{\beta-1}.$$

For the exponential-Gompertzian size distribution, the cumulative distribution function has the form

$$F(x) = 1-[\log(\kappa/\gamma)]^{-\beta'}[\log(\kappa/x)]^{\beta'}.$$

Then the first partial derivatives with respect to γ, β' and κ are

$$\frac{\partial F(x)}{\partial \gamma} = -\beta'\gamma^{-1}[\log(\kappa/\gamma)]^{-(\beta'+1)}[\log(\kappa/x)]^{\beta'},$$

$$\frac{\partial F(x)}{\partial \beta'} = [\log(\kappa/\gamma)]^{-\beta'}[\log(\kappa/x)]^{\beta'}\{\log[\log(\kappa/\gamma)]$$
$$-\log[\log(\kappa/x)]\}$$

and

$$\frac{\partial F(x)}{\partial \kappa} = \beta' \kappa^{-1} [\log(\kappa/\gamma)]^{-(\beta'-1)} [\log(\kappa/x)]^{\beta'-1} \{[\log(\gamma/x)]/$$

$$[\log(\kappa/\gamma)]^2\}.$$

For the last of the lower curves, the exponential-Bertalanffy-Richards size distribution, it simplifies the expressions to introduce some collapsing notation. Let $A = (\kappa/\gamma)^n-1$ and $Y = (\kappa/x)^n-1$ so that $(\kappa/\gamma)^n$ is $(A+1)$ and $(\kappa/x)^n$ is $(Y+1)$. Then the cumulative distribution function may be written as

$$F(x) = 1-A^{-\beta/n}Y^{\beta/n}.$$

Likewise, the first partial derivatives with respect to the parameters γ, β, n and κ may be written as

$$\frac{\partial F(x)}{\partial \gamma} = -\beta\gamma^{-1}(A+1)A^{-(\beta/n+1)}\ Y^{\beta/n},$$

$$\frac{\partial F(x)}{\partial \beta} = n^{-1}A^{-\beta/n}Y^{\beta/n}[\log(A)-\log(Y)],$$

$$\frac{\partial F(x)}{\partial n} = \beta n^{-1}A^{-\beta/n}Y^{\beta/n}[n^{-1}\log(Y/A)-\{(Y+1)/Y\}\log(\kappa/x)$$

$$+\{(A+1)/A\}\log(\kappa/\gamma)]$$

and

$$\frac{\partial F(x)}{\partial \kappa} = -\beta Y^{\beta/n-1}A^{-(\beta/n+1)}[x^{-1}(\kappa/x)^{n-1}A-\gamma^{-1}(\kappa/\gamma)^{n-1}Y].$$

The first of the upper curves to be presented is the exponential-hyperlogistic size distribution. Again one may make use of "accordion" mathematics. Let $A = \kappa/\gamma-1$ and $Y = \kappa/x-1$ so that $(A+1)$ is (κ/γ) and $(Y+1)$ is (κ/x). Then one may also define an expression H such that

$$H = 1-\exp[\beta p^{-1}(A^{-p}-Y^{-p})].$$

For positive values of the parameter p, the cumulative distribution function is

$$F(x) = H.$$

The partial derivatives of F(x) with respect to the parameters γ, β, p and κ are

$$\frac{\partial F(x)}{\partial \gamma} = -(1-H)\beta\kappa\gamma^{-2}A^{-(p+1)},$$

$$\frac{\partial F(x)}{\partial \beta} = (1-H)p^{-1}(Y^{-p}-A^{-p}),$$

$$\frac{\partial F(x)}{\partial p} = (1-H)\beta p^{-1}[(A^{-p}-Y^{-p})p^{-1}+A^{-p}\log(A)-Y^{-p}\log(Y)]$$

and

$$\frac{\partial F(x)}{\partial \kappa} = (1-H)\beta[\gamma^{-1}A^{-(p+1)}-x^{-1}Y^{-(p+1)}].$$

For negative values of the parameter p, one may add the shorthand expression D such that $D = 1-\exp(\beta p^{-1}A^{-p})$ and (1-D) is $\exp(\beta p^{-1}A^{-p})$. Then the cumulative distribution function becomes

$$F(x) = D^{-1}H.$$

The first partial derivatives with respect to γ, β, p and κ are

$$\frac{\partial F(x)}{\partial \gamma} = \beta\kappa\gamma^{-2}A^{-(p+1)}D^{-1}(D^{-1}H-1),$$

$$\frac{\partial F(x)}{\partial \beta} = D^{-1}p^{-1}[(D^{-1}H-1)A^{-p}+(1-H)Y^{-p}],$$

$$\frac{\partial F(x)}{\partial p} = -D^{-2}(1-D)H\beta p^{-1}A^{-p}[p^{-1}+\log(A)]-D^{-1}(1-H)\beta p^{-1}\cdot$$
$$[p^{-1}(Y^{-p}-A^{-p})+Y^{-p}\log(Y)-A^{-p}\log(A)]$$

and

$$\frac{\partial F(x)}{\partial \kappa} = -D^{-2}(1-D)H\beta\gamma^{-1}A^{-(p+1)}-D^{-1}(1-H)\beta\cdot$$
$$[x^{-1}Y^{-(p+1)}-\gamma^{-1}A^{-(p+1)}].$$

The second of the upper curves, the exponential-hyper-Gompertzian distribution, requires some similarly defined expressions. Let $A = [\log(\kappa/\gamma)]^{-p}$ and $Y = [\log(\kappa/x)]^{-p}$. Then H is defined such that

$$H = 1-\exp[\beta' p^{-1}(A-Y)].$$

If p is positive, the distribution function, F(x), is H. For the first partial derivatives of F(x) with respect to γ, β', p and κ, one obtains

$$\frac{\partial F(x)}{\partial \gamma} = -(1-H)\beta'\gamma^{-1}[\log(\kappa/\gamma)]^{-(p+1)},$$

$$\frac{\partial F(x)}{\partial \beta'} = -(1-H)p^{-1}(A-Y),$$

$$\frac{\partial F(x)}{\partial p} = (1-H)\beta' p^{-1}\{(A-Y)p^{-1}+A\log[\log(\kappa/\gamma)]-Y\log[\log(\kappa/x)]\}$$

and

$$\frac{\partial F(x)}{\partial \kappa} = (1-H)\beta'\kappa^{-1}\{[\log(\kappa/\gamma)]^{-(p+1)}-[\log(\kappa/x)]^{-(p+1)}\}.$$

Again for negative values of p, one may denote, by D, the expression $[1-\exp(\beta' p^{-1}A)]$. So for this situation the distribution function becomes

$$F(x) = D^{-1}H.$$

The first partial derivatives are

$$\frac{\partial F(x)}{\partial \gamma} = \beta'\gamma^{-1}D^{-1}(D^{-1}H-1)[\log(\kappa/\gamma)]^{-(p+1)},$$

$$\frac{\partial F(x)}{\partial \beta'} = p^{-1}D^{-1}[(D^{-1}H-1)A+(1-H)Y],$$

$$\frac{\partial F(x)}{\partial p} = -D^{-2}(1-D)H\beta' p^{-1}A\{p^{-1}+\log[\log(\kappa/\gamma)]\}-D^{-1}(1-H)\beta' p^{-1}\cdot$$
$$\{p^{-1}[Y-A]+Y\log[\log(\kappa/x)]-A\log[\log(\kappa/\gamma)]\}$$

and

$$\frac{\partial F(x)}{\partial \kappa} = -D^{-2}(1-D)\beta'\kappa^{-1}H[\log(\kappa/\gamma)]^{-(p+1)} - D^{-1}(1-H)\beta'\kappa^{-1}$$

$$\{[\log(\kappa/x)]^{-(p+1)} - [\log(\kappa/\gamma)]^{-(p+1)}\}.$$

For the exponential-generic curve, one may make the parallel definitions of expressions:

$$A = (\kappa/\gamma)^n - 1, \qquad A+1 = (\kappa/\gamma)^n$$

$$Y = (\kappa/x)^n - 1, \qquad Y+1 = (\kappa/x)^n$$

and

$$H = 1-\exp[\beta n^{-1}p^{-1}(A^{-p}-Y^{-p})].$$

Then for positive values of the parameter p, the cumulative distribution function is

$$F(x) = H.$$

Differentiating with respect to the five parameters, γ, β, n, p and κ, one obtains

$$\frac{\partial F(x)}{\partial \gamma} = -(1-H)\beta\gamma^{-1}(A+1)A^{-(p+1)},$$

$$\frac{\partial F(x)}{\partial \beta} = -(1-H)n^{-1}p^{-1}(A^{-p}-Y^{-p}),$$

$$\frac{\partial F(x)}{\partial n} = (1-H)\beta n^{-1}\{(A^{-p}-Y^{-p})n^{-1}p^{-1} + [A^{-(p+1)}(A+1)\log(\kappa/\gamma)$$

$$-Y^{-(p+1)}(Y+1)\log(\kappa/x)]\},$$

$$\frac{\partial F(x)}{\partial p} = (1-H)\beta n^{-1}p^{-1}[(A^{-p}-Y^{-p})p^{-1} + A^{-p}\log(A) - Y^{-p}\log(Y)]$$

and

$$\frac{\partial F(x)}{\partial \kappa} = \beta(1-H)[(\kappa/\gamma)^{n-1}A^{-(p+1)}\gamma^{-1} - (\kappa/x)^{n-1}Y^{-(p+1)}x^{-1}].$$

To take care of the exponential-generic curve for a negative value of p, one may define D such that

$$D = 1-\exp(\beta n^{-1} p^{-1} A^{-p}).$$

Now the cumulative distribution function is

$$F(x) = D^{-1}H.$$

The set of first partial derivatives is as follows:

$$\frac{\partial F(x)}{\partial \gamma} = \beta\gamma^{-1}(A+1)A^{-(p+1)}D^{-1}(D^{-1}H-1),$$

$$\frac{\partial F(x)}{\partial \beta} = n^{-1}p^{-1}D^{-1}[(D^{-1}H-1)A^{-p}+(1-H)Y^{-p}],$$

$$\frac{\partial F(x)}{\partial n} = D^{-2}(1-D)H\beta n^{-1}[n^{-1}p^{-1}A^{-p}+A^{-(p+1)}(A+1)\log(\kappa/\gamma)]$$
$$+D^{-1}(1-H)\beta n^{-1}[n^{-1}p^{-1}(A^{-p}-Y^{-p})+A^{-(p+1)}(A+1)\cdot$$
$$\log(\kappa/\gamma)-Y^{-(p+1)}(Y+1)\log(\kappa/x)],$$

$$\frac{\partial F(x)}{\partial p} = -D^{-2}(1-D)H\beta n^{-1}p^{-1}A^{-p}[p^{-1}+\log(A)]-D^{-1}(1-H)\beta n^{-1}p^{-1}$$
$$[p^{-1}(Y^{-p}-A^{-p})+Y^{-p}\log(Y)-A^{-p}\log(A)]$$

and

$$\frac{\partial F(x)}{\partial \kappa} = -D^{-2}(1-D)H\beta\kappa^{-1}(A+1)A^{-(p+1)}-D^{-1}(1-H)\beta\kappa^{-1}$$
$$[(Y+1)Y^{-(p+1)}-(A+1)A^{-(p+1)}].$$

The last of the size distributions covered here is the Weibull. Here one may use the letter H to represent the expression

$$\{1-\exp[\beta''\alpha^{-1}(\gamma^{\alpha}-x^{\alpha})]\}.$$

So the cumulative distribution function for the case of positive α is

$$F(x) = H.$$

The first partial derivatives with respect to γ, β'' and α are

$$\frac{\partial F(x)}{\partial \gamma} = -(1-H)\beta''\gamma^{\alpha-1},$$

$$\frac{\partial F(x)}{\partial \beta''} = -(1-H)\alpha^{-1}(\gamma^{\alpha}-x^{\alpha})$$

and

$$\frac{\partial F(x)}{\partial \alpha} = -(1-H)\beta''\alpha^{-1}[\alpha^{-1}(x^{\alpha}-\gamma^{\alpha})+\gamma^{\alpha}\log(\gamma)-x^{\alpha}\log(x)].$$

If α is negative, one may define a D such that

$$D = 1-\exp(\beta''\alpha^{-1}\gamma^{\alpha}).$$

Then the cumulative distribution function is written as

$$F(x) = D^{-1}H.$$

The corresponding set of first partial derivatives is

$$\frac{\partial F(x)}{\partial \gamma} = \beta''\gamma^{\alpha-1}D^{-1}(D^{-1}H-1),$$

$$\frac{\partial F(x)}{\partial \beta''} = \alpha^{-1}D^{-1}[(D^{-1}H-1)\gamma^{\alpha}+(1-H)x^{\alpha}]$$

and

$$\frac{\partial F(x)}{\partial \alpha} = -D^{-2}(1-D)H\beta''\alpha^{-1}\gamma^{\alpha}(\alpha^{-1}-\log\gamma)-D^{-1}(1-H)\beta''\alpha^{-1}\cdot$$
$$[\alpha^{-1}(x^{\alpha}-\gamma^{\alpha})+\gamma^{\alpha}\log(\gamma)-x^{\alpha}\log(x)].$$

This completes the list of functions needed in order to use either the weighted or the unweighted least squares methods of parameter estimation discussed at the beginning of this section. In the next chapter, several interesting examples are presented and practical aspects of the nonlinear fitting are discussed.

VIII. EXAMPLES AND PRACTICAL CONSIDERATIONS

Before proceeding with the presentation of seven examples of size distributions fitted to data, a few practical considerations must be stated. The lower curves are relatively simple to handle. Convergence can often be easily obtained even with very poor initial values. The methods for acquiring starting values discussed in the preceding chapter, however, provide good starting values for data following shapes covered by the lower curves. In turn, the upper curves are easily handled from the estimates found for the parameters of the lower curves in this case. These general comments are true whichever the method of fitting is chosen. If the curve form needed is not one of those typical of the lower curves, the process becomes much more difficult for several reasons. These curves tend to have values for p or α which are much farther from zero than is tolerable for the estimates of γ, κ and β or β' to bear much relationship to those taken from the lower curves. It is better in such cases to use a method of getting starting values which assigns $\tilde{\gamma}$ and $\tilde{\kappa}$ and constrains the initial estimates to pass through particular percentile points of the empirical distribution function. With modern hand calculators, an iterative process of this kind may be quickly accomplished. For the upper curves, it is especially important that these starting values be quite close to the actual estimates. The reason for this is to be found in the close relationship to the Weibull. This distribution is notoriously difficult to fit because very small changes in the parameter which designates the power of the random variable create dramatic changes in the values of the function to be fitted. The usual approach necessary for such a situation, even with adequate initial guesses for the parameter estimates, is to step the values assigned to that parameter systematically and to obtain converged estimates of the other parameters at each of these stepped points. It is usually possible by this method to get close enough to the estimates sought to turn the fitting process loose and to obtain converged values for all of the parameter estimates. Sometimes, however, one finds oneself in the situation of having many local minima (or maxima) or of having an extremely flat surface in the parameter space (usually due to overparameterization). Either of these problems can be spotted with mapping the adjacent points for the function to be maximized or minimized.

The other problems which plague attempts to fit the upper curves are numerical in nature. There are primarily two such types of problems. In those of the first sort, there tends to be a greater range in the sizes in the data so that the estimates of the κ to γ ratio is larger than 2. If the estimate required for p is fairly large (perhaps greater than 6) or if the initial guess for p is too large, the sequence of events encountered in the fitting process tends to be of this variety. First the exponential piece of the function which contains the parameter γ becomes effectively zero. There is no contribution being made from the normalizing term. This causes an apparent singularity with respect to the estimate for γ. It also causes the loss of the terms involving this correction factor from the calculation of the partial derivatives. As a direct result, the possibility of the partial derivatives' ever becoming near zero is lost. They are forced to continue to change the parameters' estimates in the same direction in constant or even increasing steps. The estimates for all of the remaining parameters tend to become quite large, although a pseudoconvergence may appear as the limits of the accuracy of the calculations are reached. If this difficulty has been caused by a poor guess for p (or, even more likely, by a poor combination of initial guess for p and β or β'), this state of affairs may often be overcome by returning to those initial guesses and mapping the function to be minimized or maximized around those points for fixed values of $\tilde{\gamma}$ and $\tilde{\kappa}$. If this is successful, better staring values will be obtained. If the range of the data is indeed large, however, one may be forced to abandon the upper curves which contain κ in favor of the Weibull's simpler assumptions. One should note that this solution must be adopted if the data in question appear to run out when the distribution function is still rising. In this case, there is insufficient information in the data to allow the estimation of the parameter κ. If the problem appears to be purely a numerical one so that the estimate of β or β' is large and very highly correlated with those of p and κ, a model retaining κ but placing a restriction upon β or β' may give a preferable solution. What has happened in these cases is most easily shown in the exponential-hyperlogistic function in the least squares case. The factor $(\kappa/\gamma-1)^{-p}$ is gone from the equations and the effect upon the function to be minimized of the smaller observations is also greatly decreased so that the fit of the estimated function to the smaller observations is actually worsened with each successive iteration. The estimate of κ is forced upward until the effective difference between $(\kappa/x-1)$ and (κ/x) is lost. It is as if the expression $\beta\kappa^{-p}$ has become a new parameter; the parameters β and κ are no longer separable. In this situation, a useful restriction for β might have the form $\beta = (\kappa/\gamma-1)^{p}$. Of course, if this choice is made, it is necessary to reconstruct the function to be maximized or minimized with its partial derivatives accordingly. The sixth example considered in this chapter, that of the weight of male chickens, is a good illustration of this phenomenon.

The second numerical trouble tends to appear conversely when the range covered by the data is shorter so that the κ to γ ratio is between 1 and 2 and closer to 1. Estimates of κ that are near the maximum sample value in size cause the appearance of large negative values in the expression to be exponentiated. For this situation the estimate of the distribution is essentially unity for large observations. When this has happened, it is common for the parameter estimate for p to be large, for that of β or β' to be small and for that of κ to be reasonably close to the maximum observation. Although there is a considerable loss of accuracy both in the function estimation and in the estimates of the partial derivatives, the fit obtained may still yield quite an adequate description of the data. There will be some systematic deviation from fit in that the estimate of the function will be slightly too large for large values of the observations (symptomatic of the trouble with rounding error in the machine) and those for smaller values of x will be slightly too small. It may be possible to check the convergence or to adjust the values slightly by mapping around the apparently converged values. One should be aware that if the values which cause this problem are not really necessary to describe the data but are an artifact of a poor choice of starting values, most of the nonlinear fitting techniques will be unable to recover and to proceed, unaided, toward better estimates. They can not pull out of this sort of numerical trouble. It must only get worse until a pseudoconvergence is reached. Again the way out of trouble involves mapping the function for values of the parameters removed from those which led to difficulty. The importance of the initial estimate for p or for α in the Weibull can not be overemphasized if a reasonable fit of one of the upper curves to real data is to be obtained. It is only slightly less important for the ratio of β to p in the exponential-hyperlogistic function or in the exponential-generic function, that of β' to p in the exponential-hyper-Gompertz function and that of β'' to α in the Weibull to be approximately correct, certainly within the correct order of magnitude. Both of these requirements are more relevant for the ability of a particular program to converge than are the initial estimates for γ and κ. The remainder of this chapter is devoted to the illustration of these points through examples.

Three methods for fitting these size distributions to data have been discussed in the seventh chapter. Maximum likelihood and weighted nonlinear least squares with a nondiagonal weight matrix are both more difficult to use than unweighted nonlinear least squares since they cannot be done using standard computer programs without much additional specialized programming for a given model. For this reason, the three techniques are all used on only one of the examples, the first. All of the estimates included in this section were obtained using the Statistical Analysis System (SAS) of Barr et al. (1976). The maximum likelihood estimates were obtained using the method of scoring. Both the maximum likelihood and the weighted nonlinear least squares techniques were implemented with special purpose programs using the matrix manipulation procedure in SAS. For the unweighted nonlinear least squares estimates, the nonlinear procedure already available in SAS was used. The algorithms utilized varied with the examples. They included Gauss-Newton, steepest descent, a Marquardt algorithm and grid search. The Marquardt approach has usually proved the most efficient; however, it may accentuate the numerical problems if they are present. In this event, Gauss-Newton together with careful grid searches is better.

The first example, taken from the Pocket Data Book USA 1973 issued by the U.S. Department of Commerce (Lerner et al., 1973), is the list of sizes measured in thousands of persons of the one hundred largest metropolitan areas in the United States according to the 1970 census. The smallest observation is Binghampton, N.Y. - Pa. at 303,000 persons. The largest is New York City at 11,529,000. One point which is brought out by this example is the nature of the family of size distributions when the data are truncated from the left, on the small side. There is no change in the form of the distributions. Only the interpretation of the parameter γ is affected. Instead of representing the smallest measurable size, γ becomes the truncation point for the distribution. Firstly, the unweighted nonlinear least squares solutions for the estimates of the parameters in all of the size distributions in this family are considered. The empirical distribution appears to be J shaped in form so that the fitting of the lower curves is reasonable. The starting values for the upper curves are taken from the estimates of the appropriate lower curves in each case. Note that the data are recorded in thousands. The total uncorrected sum of squares is 33.168. For a summary of the various models for this example, see Table 1. For the Pareto, the simplest model, the estimate of γ is 296.522 with an asymptotic standard error of 2.071. The estimate of β is 0.949 with an asymptotic standard error of 0.013. The sum of squared residuals is 0.074 with 98 degrees of freedom. For the exponential-logistic model, the parameter estimates are $\tilde{\gamma}$ = 294.442 with asymptotic standard error 2.001, $\tilde{\beta}$ = 0.894 with standard error 0.020 and $\tilde{\kappa}$ = 11933.878 with standard error 3556.091. The sum of squared residuals is 0.057 with 97 degrees of freedom. The increase in the sum of squares explained by the model over the Pareto is 0.015 which is clearly significant. For the exponential-logistic model, there is a clear improvement in fit over the Pareto. Next the exponential-Gompertz model is tried. The unweighted least squares estimates are found to be $\tilde{\gamma}$ = 282.850 with standard error 2.199, $\tilde{\beta}'$ = 2.798 with standard error 0.227 and $\tilde{\kappa}$ = 12012.494 with asymptotic standard error 2808.188. The residual sum of squares is 0.035. This is an improvement of 0.039 in the model sum of squares over the Pareto which is a greater improvement than that produced by the logistic based model. The Bertalanffy-Richards based model produces these estimates: $\tilde{\gamma}$ = 284.370 with standard error 2.684; $\tilde{\beta}$ = 0.176 with standard error

Table 1 Summary of the nonlinear unweighted least squares results for the sizes (in thousands of persons) of the 100 largest metropolitan areas in the United States according to the 1970 census.

Model	$\tilde{\gamma}$	$\tilde{\beta}, \tilde{\beta}', \tilde{\beta}''$	$\tilde{n}$	$\tilde{p}$	$\tilde{\kappa}$	df	SSE
Pareto	296.522 (2.071)	0.949 (0.013)	0.000	0.000	infinite	98	0.074
Exponential-logistic	294.442 (2.001)	0.894 (0.020)	1.000	0.000	11933.878 (3556.091)	97	0.057
Exponential-Gompertz	282.850 (2.199)	2.798 (0.227)	0.000	0.000	12012.495 (2808.188)	97	0.035
Exponential-Bertalanffy-Richards	284.370 (2.684)	0.176 (0.608)	0.070 (0.290)	0.000	11805.614 (10732.159)	96	0.035
Exponential-hyperlogistic	283.788 (2.707)	1.828 (0.933)	1.000	0.244 (0.087)	12000.011 (9896.118)	96	0.035
Exponential-hyper-Gompertz	284.135 (2.577)	2.560 (2.975)	0.000	-0.082 (0.598)	12002.673 (19707.088)	96	0.035
Exponential-generic	284.361 (3.509	0.175 (5.460)	0.070 (2.873)	-0.000 (5.079)	11805.614 (61480.811)	95	0.035
Weibull	282.470 (2.477)	0.109 (0.024)	$\tilde{\alpha}$ = 0.337 (0.034)		infinite	97	0.036

df = degrees of freedom for residual sum of squares SSE = residual (or error) sum of squares

estimate = first number in cell, asymptotic standard error of the estimate = parenthesized number in cell

0.608; $\tilde{n}$ = 0.070 with standard error 0.290 and $\tilde{\kappa}$ = 11805.614 with standard error 10732.159. The residual sum of squares is 0.035 with 96 degrees of freedom. This is an improvement in explaining power over the logistic based model of 0.022 in the sum of squares which is significant. It is not, however, a significant improvement over the exponential-Gompertz model since the sum of squared residuals differs only in the fourth decimal place. The exponential-hyperlogistic model yields the following estimates: $\tilde{\gamma}$ = 283.788 with standard error 2.707, $\tilde{\beta}$ = 1.828 with standard error 0.933, $\tilde{p}$ = 0.244 with standard error 0.087 and $\tilde{\kappa}$ = 12000.011 with standard error 9896.118. Again the residual sum of squares is 0.035 with 96 degrees of freedom. Apparently this model is a significant improvement over the logistic based model but not very different in explaining power from the exponential-Bertalanffy-Richards. The exponential-hyper-Gompertz model gives estimates from the formulation for negative values of p as $\tilde{\gamma}$ = 284.135 with an asymptotic standard error of 2.577, $\tilde{\beta}'$ = 2.560 with standard error 2.975, $\tilde{p}$ = -0.082 with standard error 0.598 and $\tilde{\kappa}$ = 12002.673 with standard error 19707.088. The residual sum of squares with 96 degrees of freedom is 0.035. There is no significant improvement in sum of squares explained by the model for the exponential-hyper-Gompertz over the exponential-Gompertz. The fit is essentially equivalent to that from the exponential-Bertalanffy-Richards model. The attempt to fit the exponential-generic model with all five parameters returned to the Bertalanffy-Richards based model. With the presence of the parameter n, it is not possible to add the parameter p. This estimation try gives these results: $\tilde{\gamma}$ = 284.361 with standard error 3.509, $\tilde{\beta}$ = 0.175 with standard error

5.460, $\tilde{n}$ = 0.070 with standard error 2.873, $\tilde{p}$ = -0.000 with standard error 5.079 and $\tilde{\kappa}$ = 11805.614 with asymptotic standard error 61480.811. The residual sum of squares again is 0.035, now with 95 degrees of freedom. The last of the size models, the Weibull, was also used. The estimates it gives are: $\tilde{\gamma}$ = 282.470 with an asymptotic standard error of 2.477, $\tilde{\beta}''$ = 0.109 with an asymptotic standard error of 0.024 and $\tilde{\alpha}$ = 0.337 with an asymptotic standard error of 0.034. The residual sum of squares is 0.036 with 97 degrees of freedom. This is a significant improvement over the Pareto model with a reduction in the error sum of squares of 0.038. Of the collection of models fitted here by unweighted nonlinear least squares, the best model to describe the data, from the point of view of residual sum of squares and of economy in terms of number of parameters needed, is the exponential-Gompertz distribution. Figure 9 presents the results of this model graphically.

The exponential-Gompertz model was also fitted to these data using the maximum likelihood technique. It is always the case for this family of distributions that the maximum likelihood estimate for γ is the smallest observation. For the cities this means that $\tilde{\gamma}$ = 303. Maximization of the log likelihood function conditional upon this value for γ yields a log likelihood value of -755.319. The estimates for β and κ are: $\hat{\beta}$ = 4.539 with an asymptotic standard error of 0.454 and $\hat{\kappa}$ = 53029.7 with an asymptotic standard error of 20719.0. The asymptotic information matrix conditional on $\hat{\gamma}$ = 303 is

$$\begin{bmatrix} 4.855 & -1.032 \times 10^{-4} \\ -1.032 \times 10^{-4} & 2.329 \times 10^{-9} \end{bmatrix}$$

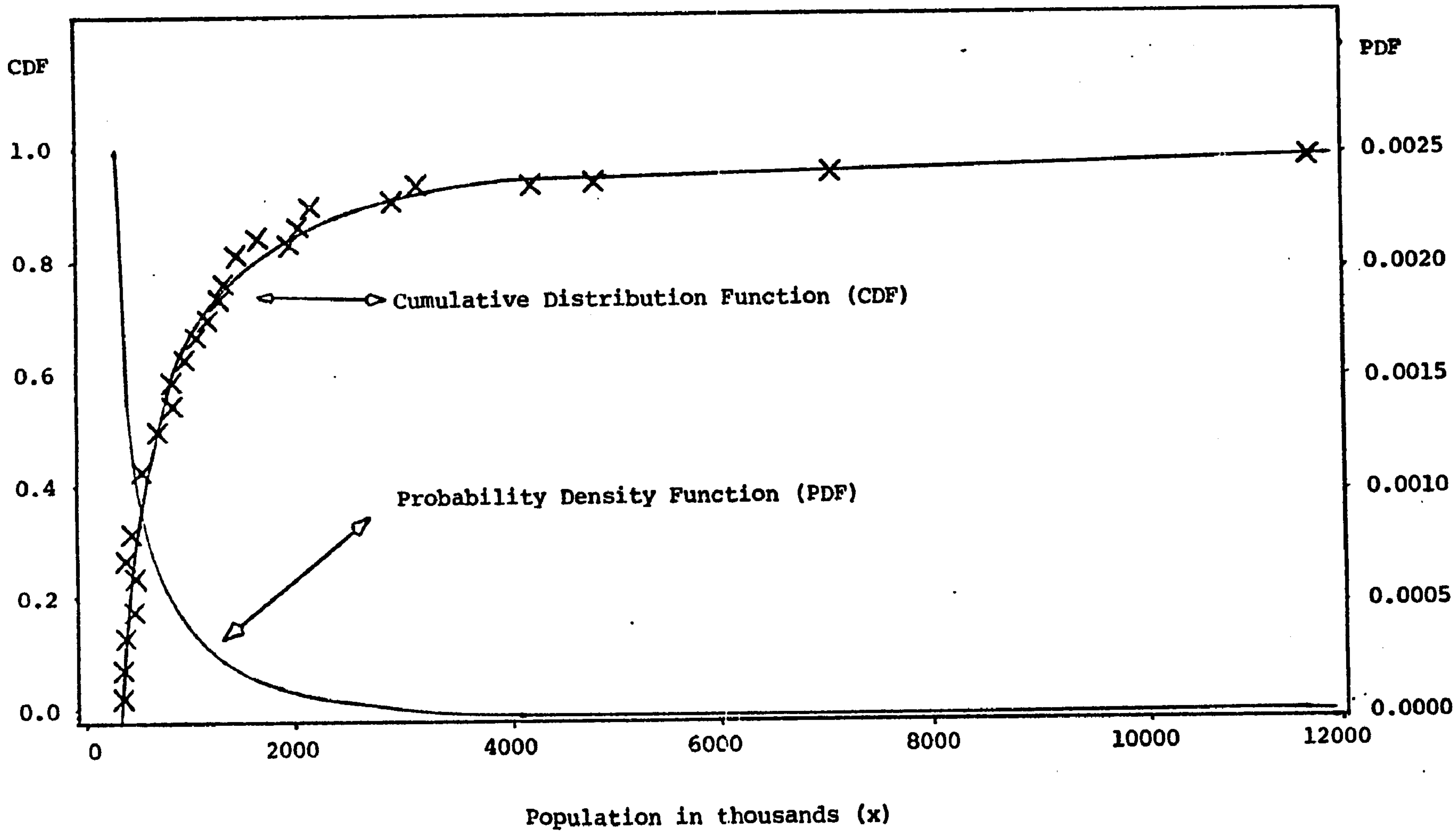


Fig. 9 Size of the 100 largest metropolitan areas in the United States according to the 1970 census -- unweighted nonlinear least squares fit of the exponential-Gompertz size distribution function.

and convergence was obtained after a total of 51 iterations.

The weighted nonlinear least squares approach gives the following estimates: $\tilde{\gamma}$ = 284.594 with standard error 5.522, $\tilde{\beta}$ = 3.150 with asymptotic standard error 0.142 and $\tilde{\kappa}$ = 16970.9 with standard error 8905.2. The variance-covariance matrix of the parameter estimates has a determinant of approximately 2247963. Its trace is 79301848 approximately and its sphericity is roughly 4.956×10^{-6}. Its eigenvalues are 79301840, 7.886 and 3.595×10^{-3}. If the log likelihood function is evaluated for these estimates, it is found to be -760.999. The weighted sum of squares for the model with 3 degrees of freedom is 140257 with mean sum of squares 46752.5. The error sum of squares is 799.418 with 97 degrees of freedom and the mean error sum of squares is 8.241. The total weighted sum of squares (uncorrected) is 154386. These estimates used in the unweighted formulation produce a residual sum of squares of 0.035 so that they differ in the unweighted sum of squares for error from the unweighted estimates only beyond the third decimal place.

The second example is the number of murders and/or instances of manslaughter in 16 selected cities in the United States as recorded in the United States Statistical Abstract for 1970 and cited by Hartigan (1975). The sizes per 100,000 population ranged from 2.5 for Hartford to 18.1 for Dallas. These data follow an empirical distribution function which has a slight inflection point toward the lower end, but is relatively linear on the whole. There is a good deal of scatter in the observations. In this case, nonlinear unweighted least squares fits are made for several of the size models. The lower curves are computed mostly to get starting values for the upper curves.

They should not provide an adequate description of the data since some slightly S shaped appearance is present. The Pareto estimate for γ is stuck upon the bound at 2.5. The estimate for β is 0.637 with an asymptotic standard error of 0.097. The residual sum of squares is 0.215 with 14 degrees of freedom. The exponential-logistic model also has an estimate for γ which is fixed at the bound, 2.5. The estimate for β is 0.387 with a standard error of 0.032 and the estimate of κ is 18.140 with a standard error of 0.147. The residual sum of squares is 0.052 with 13 degrees of freedom. This model is a highly significant improvement over the Pareto. It reduces the error sum of squares by 0.163. The exponential-Gompertz fit yields an estimate for γ at 2.368 with an asymptotic standard error of 0.246. Then $\tilde{\beta}' = 0.635$ with standard error of 0.051 and $\tilde{\kappa} = 18.379$ with standard error 0.492. The error sum of squares is 0.030, which is a reduction of 0.185 over the Pareto. The exponential-Bertalanffy-Richards fit goes to its limit as n approaches 0 and returns the exponential-Gompertz fit. One finds $\tilde{\gamma} = 2.368$ with standard error 0.237, $\tilde{\beta} = 6.000 \times 10^{-8}$ with standard error 0.000, $\tilde{n} = 9.000 \times 10^{-8}$ with standard error 0.000 and $\tilde{\kappa} = 18.379$ with asymptotic standard error 0.475. Again the error sum of squares is 0.030. There is a clear indication that the Gompertz based fit is to be preferred over the logistic based fit. The hyper-Gompertz based model produces a $\tilde{\gamma}$ of 1.441 with standard error 1.022, a $\tilde{\beta}'$ of 2.252 with a standard error of 5.926, a $\tilde{p}$ of 1.667 with a standard error of 2.112 and a $\tilde{\kappa}$ of 34.384 with a standard error of 33.382. The residual sum of squares is 0.021 with 12 degrees of freedom. If the significance of this model over the exponential-Gompertz is studied by means of a pseudo-F statistic with 1 and 12 degrees of freedom, one obtains a value of 4.763 which appears to be significant. The exponential-hyperlogistic does not do quite as well. The estimates are as follows: $\tilde{\gamma}$ = 1.325 with standard error 1.268, $\tilde{\beta}$ = 1.360 with standard error 1.809, $\tilde{p}$ = 0.964 with standard error 0.534 and $\tilde{\kappa}$ = 27.233 with a standard error of 12.358. The residual sum of squares is 0.022 with 12 degrees of freedom. It is not possible to fit the exponential-generic curve in this case. Apparently the exponential-hyper-Gompertz distribution gives the best description of the murder data.

Since much of the development of size distributions historically has been concerned with data on the number of individuals in a group of related species or the number of species in a group of related genera, it seems desirable to include one example of that type. This third example is that of Dr. J. C. Willis on the number of genera of Cerambycinae with 1, 2, 3 or more species, used by Yule (1924). The data range from 469 genera with only 1 species to 1 genus with 125 species. The total number of observations is 1024. The data appear to follow a steep J shaped curve. The Pareto distribution gives an error sum of squares of 0.081 with 1022 degrees of freedom. The parameter estimates are $\tilde{\gamma}$ = 0.704 with a standard error of 0.001 and $\tilde{\beta}$ = 0.738 with a standard error of 0.001. The exponential-logistic model produces a $\tilde{\gamma}$ of 0.700 with an asymptotic standard error of 0.001, a $\tilde{\beta}$ of 0.721 with a standard error of 0.001 and a $\tilde{\kappa}$ of 125.000, which is the lower bound for κ. The error sum of squares is 0.040, half of that for the Pareto. The exponential-Gompertz model has an error sum of squares of 0.017, which is much smaller than that for logistic based model and less than one-fourth the size of that for the Pareto. The parameter estimates are $\tilde{\gamma}$ = 0.673 with a standard error of 0.001, $\tilde{\beta}'$ = 6.220 with a standard error of 0.093 and $\tilde{\kappa}$ = 10219.558 with a standard error of 1288.364. The exponential-Bertalanffy-Richards model, however, yields the best fit to these data. Its error sum of squares is only 0.009 with 1020 degrees of freedom. The parameters are estimated as $\tilde{\gamma}$ = 0.683 with a standard error of 0.001, $\tilde{\beta}$ = 0.613 with a standard error of 0.006, $\tilde{n}$ = 0.452 with a standard error of 0.022 and $\tilde{\kappa}$ = 125.000 with a standard error of 8.927 (not actually stuck on the bound in later decimal places). The attempt to fit the exponential-hyperlogistic is a significant improvement over the logistic based model, but not quite as good as that of the exponential-Bertalanffy-Richards. The error sum of squares is 0.012 with 1020 degrees of freedom. Estimates are $\tilde{\gamma}$ = 0.676 with a standard error of 0.001, $\tilde{\beta}$ = 1.003 with a standard error of 0.020, $\tilde{p}$ = 0.082 with a standard error of 0.003 and $\tilde{\kappa}$ clings to the bound at 125. The attempt to fit the exponential-hyper-Gompertz fails and the attempt to fit the exponential-generic curve essentially returns to the Bertalanffy-Richards based model with an estimated value for p which is smaller than 0.0002.

The fourth example is the diameter at breast height in inches of a stand of 220 Shorea leprosula trees, ranging in size from 6 to 22 inches. The data are truncated on the left. This data set and other related ones were communicated to H. Fairfield Smith in 1954 by G.G.K. Setten of the Forest Research Institute in Selangor, Malaya. The measurements were recorded there in 1952. The model that seems appropriate as a first try is the exponential-hyperlogistic. This actually fits the data quite well. The parameter estimates are the following: $\tilde{\gamma}$ = 1.664 with an asymptotic standard error of 3.697, $\tilde{\beta}$ = 34.560 with an asymptotic standard error of 17.112, $\tilde{p} = 2.885$ with an asymptotic standard error of 0.115 and $\tilde{\kappa} = 53.785$ with an asymptotic standard error of 4.698. The error sum of squares is 0.012 with 216 degrees of freedom.

The fifth example is the size of brain cells located behind the optic nerve in a kitten which has had one eyelid sutured shut from shortly after birth. This study of brain cells in light-deprived kittens was done by Dr. Terry Hickey of the University of Alabama in Birmingham School of Optometry. The data set consists of sizes (areas) of 100 such cells ranging in size from 99.9 to 767.6 square microns. The empirical distribution function is an example of an S shaped curve which is still rising when the data terminate. Although an attempt was made to fit both the exponential-hyperlogistic curve and the exponential-hyper-Gompertz distribution, it became obvious in both cases that there is insufficient information in the data to do so. For this reason the Weibull fit, which describes the curve beautifully, is given here. The estimates are these: $\tilde{\gamma} = 99.9$ (the upper bound for γ), $\tilde{\alpha} = 2.134$ with an asymptotic standard error of 0.071 and $\tilde{\beta}'' = 0.737 \times 10^{-5}$ with a standard error of 0.295×10^{-5}. The error sum of squares is 0.078.

The sixth example is rather different from those already presented in that the data are calculated measures of size produced by an assumed model. They are taken from the doctoral dissertation of Gary F. Krause (1963). The model assumed is the logistic. Longitudinal data on the weights of each of 78 male, Athens-Canadien chickens, in grams, was fitted by the logistic. The least squares parameter estimates for each of the three parameters of the logistic were recorded for each bird.

The data used here consist of the estimates of the upper limit to this juvenile growth cycle for each of these birds. The model chosen to describe these sizes is the exponential-hyperlogistic distribution. One should know that the empirical distribution function follows an S shaped rising curve. There are several points in the flatter area on the right end. There should be enough information in the data to get an idea of the size of κ. It is also true, however, that the p value required is at least greater than 5, because of the slope of the rise. This curve is an example of the numerical problem encountered when p is this large and the κ estimate is such that the $\tilde{\kappa}$ to $\tilde{\gamma}$ ratio is greater than 2. There is an apparent singularity with respect to γ. The estimate of γ is on the bound at 2249 grams. The estimate of β is 24480.654 with a standard error of 445777.076 and the estimate of p is 6.601 with an asymptotic standard error of 3.272. The estimate of κ is 12664.043 with a standard error of 20360.926. This is roughly 28 pounds, more like a large turkey. The error sum of squares is 0.027 with 74 degrees of freedom. This is one of the occasions when applying the β constraint that $\beta = (\kappa/\gamma-1)^p$ provides a stabilizing influence. The estimate of γ remains on its upper bound at 2249 grams. The estimate of p is lowered slightly to 6.276 with a standard error of 1.346. The κ estimate is also lowered slightly to 11595.173 grams with a standard error of 6142.840. The residual sum of squares is increased by only 0.001 which is not at all a significant change. There are 75 error degrees of freedom. Again all of the estimates for this example are the unweighted nonlinear least squares estimates subject to the constraint that the estimate of γ must be less than or equal to the smallest observation.

The last example is a unimodal curve of the sort quite representative of this size family. It has an abrupt descent to the size axis on the right side and a more gradual rise, although still a steep one, on the left. The data are the by-product of a hypertensive survey conducted by optometrists in the southeastern part of the United States under the direction of Dr. John R. Pierce of the University of Alabama in Birmingham School of Optometry. The sample included 5326 white males over 18 years of age. The data ranged from 48 inches to 83 inches. Unweighted nonlinear least squares estimates are obtained using the exponential-hyperlogistic model. Note that the value of p required is high, greater than 9. The value of κ required is reasonably near the maximum observation. The result is the second type of numerical problem discussed. There still appears to be some systematic deviation from fit, but it is not possible to improve the situation due to the noise levels of the techniques involved and the unit of rounding error in the machine (16^{-13} in double precision in an IBM 370/158). Again the estimate for γ is on the bound at 48 inches. The estimates of the other parameters are as follows: $\tilde{\beta} = 0.010$ with an asymptotic standard error of 0.002, $\tilde{p} = 9.694$ with a standard error of 0.246 and $\tilde{\kappa} = 105.900$ with a standard error of 1.377. It is amusing to notice that, according to the Guinness Book of World Records (McWhirter and McWhirter, 1975), the actual height of the world's tallest man was 107.1 inches. The sum of squared deviations from fit for this model is 1.174 with 5322 degrees of freedom. For a picture of the empirical distribution function and the fitted curve, see Figure 10.

This completes the list of chosen examples for the various members

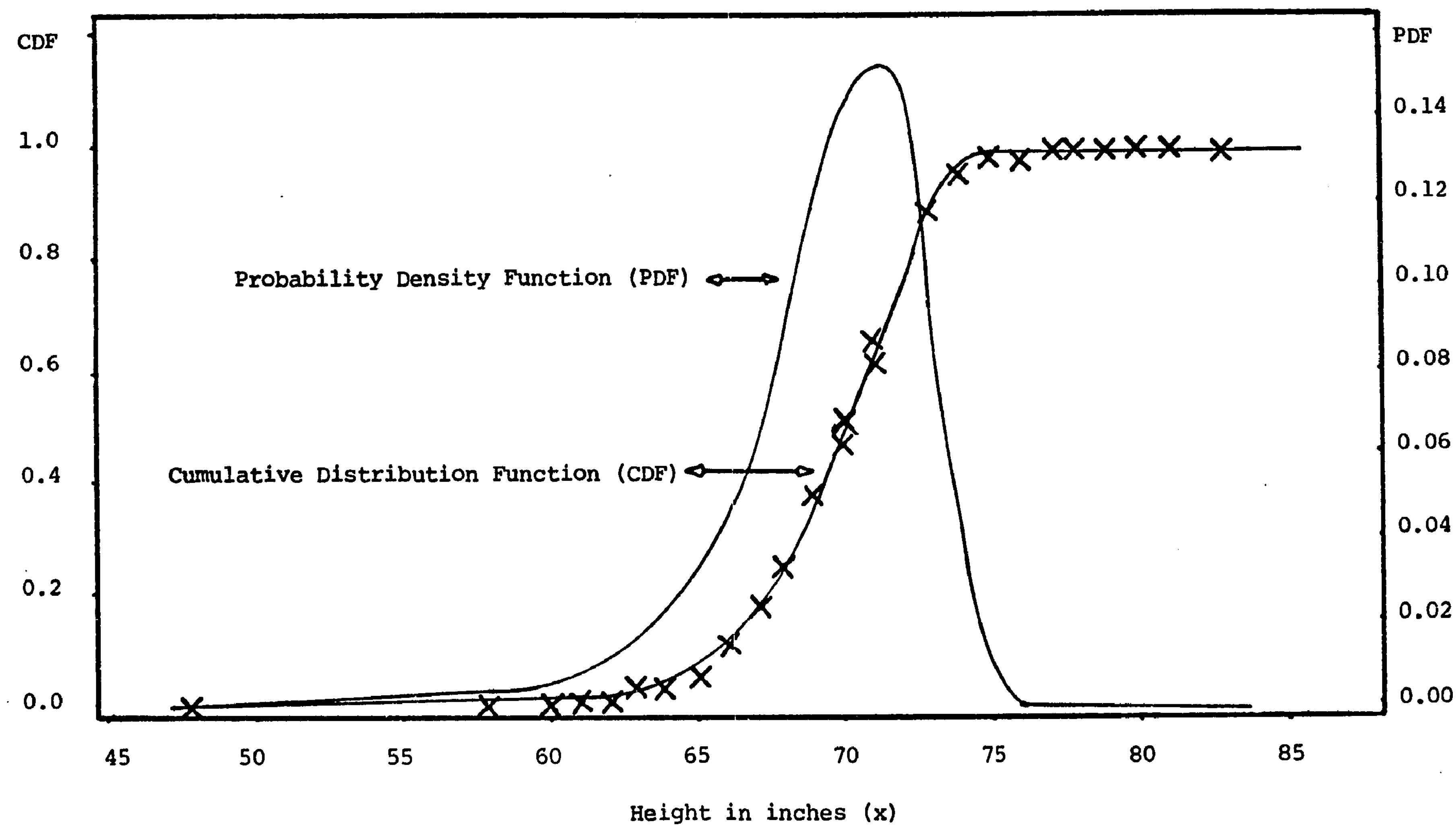


Fig. 10 Size (height in inches) of 5326 white males contacted as the result of a hypertensive screening program sponsored by the University of Alabama School of Optometry -- unweighted nonlinear least squares fit of the exponential-hyperlogistic distribution function.

of the family of size distributions derived and discussed in this dissertation. The following chapter provides a brief summary of this work.

IX. SUMMARY AND FINAL COMMENTS

The purpose of this dissertation is to derive and to describe a new family of distributions of sizes of individual units which have arrived at their current sizes by some process of accretion. There are two basic assumptions, out of the host of such, which are possible, which are actually utilized. The first assumption is that of an exponential distribution of ages for the units in the population. This may be thought of alternatively as the resulting distribution if one assumes a constant rate of acquisition for a single unit to one's records, whatever the age of the unit may be. The second assumption is that growth occurs in a deterministic manner with each member of the group to be studied following the same growth curve and that this growth curve is a member of the family indexed by the generic growth curve of Turner et al. (1976). The general member of this growth family, the generic curve, has the form $x = \kappa\{1+[1+\delta np(t-\tau)]^{-1/p}\}^{-1/n}$. Included as special members of this family are the geometric increase curve, the Gompertz curve, the logistic curve and the Bertalanffy-Richards curve. The family of size distributions arising from these two assumptions has two forms for the generic size curve depending upon the sign of the parameter p. If p is greater than 0, the distribution function has the following form:

$$F(x) = 1-\exp(-\beta/(np)\{[(\kappa/x)^n-1]^{-p}-[(\kappa/\gamma)^n-1]^{-p}\}).$$

If, instead, p is negative, between 0 and -1, the distribution must be renormalized (maximal growth may occur at a finite time). In this case, the distribution function assumes the form:

$$F(x) = \{1-\exp(\beta/(np)[(\kappa/\gamma)^{n}-1]^{-p})\}^{-1}\cdot$$

$$\{1-\exp(-\beta/(np)\{[(\kappa/x)^{n}-1]^{-p}-[(\kappa/\gamma)^{n}-1]^{-p}\})\}.$$

There are basically eight members of this size family, obtained as limiting forms when n = 1, when n approaches 0 and when p approaches 0. Two of these members are well-known curves: the Pareto and the Weibull. The lower curves (the Pareto, the exponential-Gompertz, the exponential-logistic and the exponential-Bertalanffy-Richards) may be J shaped, U shaped or reversed J shaped. The shapes for the other size curves include these. They may also be unimodal or look somewhat like a mixture of a unimodal curve and a negative exponential. The unimodal curves tend to have long tails to the left and an abrupt drop toward the axis on the right for the density functions. The moments exist for all of the curves with finite limits although the form changes for those containing p according to whether p is positive or negative. These moments are, however, typically in terms of one or more infinite series. Some of the upper moments for the Pareto and the Weibull do not always exist. The intensity functions tend to be J shaped or U shaped.

Estimation is discussed here in three guises: maximum likelihood, weighted nonlinear least squares fitting the distribution function against the empirical distribution function using a nondiagonal weight matrix, and unweighted nonlinear least squares fitting the distribution function against the empirical distribution function. This last technique is much the easiest to apply since standard nonlinear programs may be used without the excessive amounts of additional programming required for the other methods. It should be noted that these techniques are all used here as ad hoc methods. The estimation properties are not well-known for either of these least squares approaches or for the maximum likelihood approach in the event that the usual regularity conditions are not met. This dissertation concludes with the presentation of eight examples of which the first, the size of the one hundred largest metropolitan areas in the United States, is the most thoroughly explored.

As with most works of this type, there remain many unresolved questions and many areas which would lend themselves well to future research. For the set of size distributions developed here, the logical next step is the study of the properties of the three estimation methods. One might prefer instead to create another new family of size distributions, hopefully with more flexibility of shape, by choosing a different distribution of ages. The Weibull or perhaps a three-parameter Gamma distribution might be a good choice. Then, too, one may investigate the effect of allowing one or more of the growth curve parameters to follow a random distribution in the population. For each of these possibilities for extending or generalizing the family of size distributions, the problems of description and estimation will be correspondingly magnified. In any of these latter proposals, it would be helpful to have some idea of the estimation properties of the techniques used in this paper. They could easily prove relevant for new distributions to be derived.

A last point which should be mentioned is that the nested structure of this exponential-generic family of size distributions, even as it currently stands without generalizations, is quite appealing from the point of view of hypothesis testing. Knowledge of the properties of the various estimators would certainly be necessary to the proper discussion of tests of hypothesis. Any expansion of the family of size distributions based upon the Turner growth curves would still have a nested structure including these curves as a single branch. A whole system of such similar branches is a possible product of future research.

REFERENCES

Ahuja, J.C. and Nash, S.W. (1967). The generalized Gompertz-Verhulst family of distributions. Sankhya, Series A, 29, 141-156.

Barr, A.J., Goodnight, J.H., Sall, J.P. and Helwig, J.T. (1976). A User's Guide to SAS76. Sparks Press, Raleigh, North Carolina.

Bartlett, M.S. (1955). An Introduction to Stochastic Processes with Special Reference to Methods and Applications. Cambridge University Press, Cambridge, England.

Bartlett, M.S. (1969). Distributions associated with cell populations. Biometrika, 56, 391-400.

Bartlett, M.S., Gower, J.C. and Leslie, P.H. (1960). A comparison of theoretical and empirical results for some stochastic population models. Biometrika, 47, 1-11.

Bertalanffy, L. von. (1938). A quantitative theory of organic growth. Human Biology, 10, 181-213.

Bertalanffy, L. von. (1957). Quantitative laws in metabolism and growth. Quart. Rev. Biology, 32, 565-599.

Blum, H.F. (1974). Uncertainty of growth of cell populations in cancer. Journal of Theoretical Biology, 46, 143-166.

Borg, S.F. (1971). Growth: physical, chemical-biological-medical and social applications. Transactions of the New York Academy of Sciences, Series II, 33, 246-258.

Cramér, H. (1946). Mathematical Methods of Statistics. Princeton University Press, Princeton, New Jersey. pp. 218-220.

Feller, W. (1966). An Introduction to Probability Theory and Its Applications, Vol. II. John Wiley and Sons, New York City, New York.

Fisher, R.A. (1943). A theoretical distribution for the apparent abundance of different species. Journal of Animal Ecology, 12, 54-58.

Fletcher, R.I. (1974). The quadric law of damped exponential growth. Biometrics, 30, pp. 111-124.

Gheerbrant, A., ed. (1961). The Incas: The Royal Commentaries of the Inca Garcilaso de la Vega: 1539-1616. The Orion Press, New York City, New York.

Gittus, J.H. (1967). On a class of distribution functions. Applied Statistics, 16, 45-50.

Gompertz, G. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos. Trans. Roy. Soc., 115, 513-585.

Good, I.J. (1953). The population frequencies of species and the estimation of population parameters. Biometrika, 40, 237-264.

Goodman, L.A. (1953). Population growth of the sexes. Biometrics, 9, 212-225.

Goodman, L.A. (1967). On the reconciliation of mathematical theories of population growth. Journal of the Royal Statistical Society, Series A, 130, 541-553.

Gumbel, E.J. (1958). Statistics of Extremes. Columbia University Press, New York City, New York. pp. 20-26.

Hartigan, J.A. (1975). Clustering Algorithms. John Wiley and Sons. New York City, New York. p. 28.

Herdan, G. (1953). Small Particle Statistics. Elsevier Publishing Company, Amsterdam, The Netherlands. pp. 1-505.

Jones, A., ed. (1966). Genesis, 6: 15-16 and Second Chronicles, 3-4. The Jerusalem Bible. Doubleday and Company, Garden City, New York. pp. 22, 532-533.

Keiding, N., Jensen, S.T. and Ranek, L. (1972). Maximum likelihood estimation of the size distribution on liver cell nuclei from the observed distribution in a plane section. Biometrics, 28, 813-829.

Kendall, D.G. (1948). On some modes of population growth leading to R.A. Fisher's logarithmic series distribution. Biometrika, 35, 6-15.

Kendall, D.G. (1949). Stochastic processes and population growth. Journal of the Royal Statistical Society, Series B, 2, 230-264.

Keyfitz, N. (1967). Reconciliation of population models: matrix integral equation and partial fraction. Journal of the Royal Statistical Society, Series A, 130, 61-83.

Kolmogorov, A.N. (1941). Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS, 31, 99-101.

Krause, G.F. (1963). Functional Observations: a Biostatistical Study of Growth (doctoral dissertation). Virginia Polytechnic Institute, Blacksburg, Virginia.

Krause, G.F., Siegel, P.B. and Hurst, D.C. (1967). A probability structure for growth curves. Biometrics, 23, 217-225.

Laird, A.K. (1969). Dynamics of growth in tumors and in normal organisms. National Cancer Institute Monograph 30, pp. 15-28.

Laird, A.K., Tyler, S.A. and Barton, A.D. (1965). Dynamics of normal growth. Growth, 29, 233-248.

Langer, W.L., ed. (1968). Western Civilization: Paleolithic Man to the Emergence of European Powers. American Heritage Publishing Company, Inc., New York City, New York.

Lerner, W. and Statistical Compendia Staff. (1973). Pocket Data Book U.S.A. 1973, U.S. Bureau of the Census, U.S. Department of Commerce, Washington, D.C., pp. 46-47.

Lotka, A.J. (1925). The empirical elements in population forecasts. Journal of the American Statistical Association, 20, 569.

Makeham, W.M. (1890). On the further development of Gompertz's law. J. Inst. Actuar. 28, pp. 152-159, 185-192, 316-332.

Mandelbrot, B. (1970). On the language of taxonomy: an outline of a 'thermostatistical' theory of systems of categories with Willis (natural) structure. Third Symposium on Information Theory. pp. 135-145.

McWhirter, N. and McWhirter, R. (1975). Guinness Book of World Records. Sterling Publishing Company, New York City, New York.

Medawar, P.B. (1940). The growth, growth energy, and ageing of the chicken's heart. Proceedings of the Royal Society, Series B, 129, 332-355.

Nelder, J.A. (1961a). The fitting of a generalization of the logistic curve, Biometrics, 17, 89-110.

Nelder, J.A. (1961b). A note on some growth patterns in a simple theoretical organism. Biometrics, 17, 220-228.

Nelder, J.A. (1962). An alternative form of a generalized logistic equation. Biometrics, 18, 614-616.

Pareto, V. (1897). Cours d'Economie Politique. Rouge et Cie, Lausanne and Paris, France.

Parks, J.R. (1970). Growth curves and the physiology of growth: I, II and III. American Journal of Physiology, 219, 833-842.

Pearl, R. (1921). The biology of death -- VII -- natural death, public health and the population problem. The Scientific Monthly, 13, 194-213.

Pearl, R., Edwards, T.I. and Miner, J.R. (1934). The growth of *Cucumis melo* seedlings at different temperatures. Journal of General Physiology, 17, 687-700.

Pearl, R. and Reed, L.J. (1920). On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proceedings of the National Academy of Sciences, 6, 275-288.

Pollard, J.H. (1966). On the use of the direct matrix product in analyzing certain stochastic population models. Biometrika, 53, 397-415.

Pollard, J.H. (1969). Continuous time and discrete time models of population growth. Journal of the Royal Statistical Society, Series A, 132, 80-88.

Pollard, J.H. (1973). *Mathematical Models for the Growth of Human Populations*. Cambridge University Press, Cambridge, England, pp. 1-177.

Pütter, A. (1920). Studien über physiologische Ähnlichkeit, VI, Wachstumsähnlichkeiten. Arch. ges. Physiol., 180, 298-340.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*. John Wiley and Sons, New York City, New York.

Richards, F.J. (1959). A flexible growth function for empirical use. Journal of Experimental Botany, 10, 290-300.

Robertson, T.B. (1923). *The Chemical Basis of Growth and Senescence*. Lippincott, Philadelphia, Pennsylvania.

Saidel, G.M. (1968). Bacterial cell populations in a continuously changing environment. Journal of Theoretical Biology, 19, 287-296.

Simon, H.A. (1955). On a class of skew distribution functions. Biometrika, 42, 425-440.

Takahashi, M. and Inouye, K. (1967). Characteristics of cube root growth and transplantable tumours. Journal of Theoretical Biology, 14, 275-283.

Thurman, G.B., Mays, C.W., Taylor, G.N., Christensen, W.R., Rehfeld, C.E. and Dougherty, T.F. (1971). Growth dynamics of beagle osteosarcomas. Growth, 35, 119-125.

Todhunter, I. (1865). *A History of the Mathematical Theory of Probability*. Chelsea Publishing Company, Bronx, New York.

Turner, M.E., Jr., Blumenstein, B.A. and Lebaugh, J.L. (1969). A generalization of the logistic law of growth. Biometrics, 25, 577-580.

Turner, M.E., Jr., Bradley, E.L., Jr., Kirk, K.A. and Pruitt, K.M. (1976). A theory of growth. Mathematical Biosciences, 29, 367-373.

Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of Applied Mechanics, 18, 293-297.

Weiss, P. and Kavanau, J.L. (1957). A model of growth and growth control in mathematical terms. Journal of General Physiology, 41, 1-47.

Welch, Q.B. (1970). Fitting growth and research data. Growth, 34, 293-312.

Will, H.S. (1936). On a general solution for the parameters of any function with application to the theory of organic growth. The Annals of Mathematical Statistics, 7, 165-190.

Williams, F.M. (1967). A model of cell growth dynamics. Journal of Theoretical Biology, 15, 190-207.

Wingert, F. (1971). Eine Verallgemeinerung der logistischen Wachstumsfunktion. Biometrische Zeitschrift, 13, 34-78.

Yule, G.U. (1924). A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, F.R.S. Philosophical Transactions of the Royal Society of London, Series B, 213, 21-87.

Zipf, G.K. (1949). Human Behavior and the Principle of Least Effort. Addison-Wesley Press, Inc., reprinted by Hafner Publishing Company, New York City, New York.

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA IN BIRMINGHAM
DISSERTATION APPROVAL FORM

Name of Candidate Katharine Ann Kirk

Major Subject Biostatistics

Title of Dissertation An Investigation of a Family of Size Distributions

Dissertation Committee:

M. F. Turner, Jr., Chairman

Daniel C. Hurst

Edwin Bradley, Jr.

Kenneth McInnis, Jr.

Charles R. Katholi

Director of Graduate Program Daniel C. Hurst

Dean, UAB Graduate School S B Barker

Date Nov. 15, 1976