E. W. Stacy

Quasimaximum Likelihood Estimators for Two-parameter Gamma Distributions

Abstract: The Gamma probability distribution is defined by the density function $[\beta^{\alpha}\Gamma(\alpha)]^{-1}$ exp $(-x/\beta)$. This paper presents new estimators for the parameters α^{-1} and β. Required calculations are simple, primarily involving the inner product of certain elementary statistics and their logarithms. Both of the new estimators are shown to be unbiased. The variance formulas for each and their covariance formula are derived—counterparts of these for the method of moments (MM) and the method of maximum likelihood (ML) not being known except in the asymptotic form. Curve-fitting results from samples of size 50 are provided. They support the proposition that the new quasimaximum likelihood estimators (QML) are at least as effective as MM and ML estimators. Illustrative estimation based on samples of size 5 also is presented for comparison; the results highlight relative smaller variances for the ML estimators. The same Monte Carlo results signal a possibility that significant negative bias occurs with both the MM and ML techniques if small samples are used.

Introduction

Because many industrial applications have been found for the Gamma probability distribution, parameter estimation for that distribution has become a matter of practical importance as well as a topic for extensive theoretical research. The distribution (Pearson Type III with zero-valued location parameter) is defined by the density function

$$[\beta^{\alpha}\Gamma(\alpha)]^{-1} \exp\,(x/\beta), \qquad x \geq 0, \alpha > 0, \beta > 0\,, \qquad (1)$$

in which α and β are, respectively, the shape and scale parameters.

The Gamma distribution has found practical application as a model in studies relating to life, fatigue, and reliability characteristics of industrial products. Gupta and Groll [1], for example, treated the distribution in the context of acceptance sampling based upon life tests. Some earlier applications were concerned with the investigation of accident causation; Chambers and Yule [2] were among those who reported on it in this context.

Much research activity has been devoted to the development of theory and computational aids for estimating α and β. The research has dealt with order statistics as well as with complete, censored, and truncated samples. The following partial listing of contributions to the literature spans almost half a century.

Fisher [3] discussed the method of moments (MM) and emphasized certain objections that arise from statistical inefficiency which plagues that method in its application to the Gamma distribution. Gupta and Groll [1] cited likelihood equations that are the basis for the method of maximum likelihood (ML). Some aids in finding approximations to solutions of those equations were developed by Greenwood and Durand [4] and further discussed by Wilk, et al. [5]. Asymptotic ML efficiency (i.e., efficiency expected from large-sample ML estimation) was treated by Harter and Moore [6] in reporting their work on censored-sample theory. Estimation techniques for order statistics were developed by Wilk, et al. [5]. Cohen [7], Des Raj [8], and Chapman [9] treated estimation methods for truncated samples. Blischke, et al. [10] and Blischke [11] reported theory that emphasizes cases in which the shape parameter has values not in excess of two. Probability plots for the Gamma distribution, and informal estimation techniques based upon such plots are described by Wilk, et al. [12].

In the present paper developments in theory are restricted to complete random samples. Both MM and ML concepts are applicable in that context. Estimation of α^{-1} and β is considered, α^{-1} being denoted by δ whenever typographical simplicity is desired. MM estimators have a closed form and are simple to calculate. Monte Carlo results, which are presented in Section 2, lend credence to the notions that those estimators are biased and have variances that are larger than variances for the corresponding ML estimators.

The latter estimators result from the solution of transcendental equations and, hence, are not calculated from closed-form expressions. Computational aids, which were referred to above, make application of the ML techniques fairly simple, provided that the estimated values lie within ranges that are suitable to the relevant aid. Also, solution of the likelihood equations is quite amenable to approximations by electronic computer techniques. Monte Carlo results, previously alluded to, create strong suspicion that ML estimators are biased—indeed, nontrivially biased, if the sample size is small enough.

In this paper a new estimating technique is proposed for use when complete random sample data are available from a Gamma population defined by the distribution in Eq. (1). Two statistics, denoted by $\tilde{\delta}$ and $\tilde{\beta}$, are defined as estimates of α^{-1} and β, respectively. Random sample data are specified by $x_1, x_2, \cdots, x_n$ with the stipulations that 1) $n > 2$; 2) $x_i \neq x_j$ for some value of i and j, and, 3) $x_i > 0$ for $i = 1, \cdots, n$. Then the proposed statistics for estimating δ and β are, respectively,

$$\tilde{\delta} = \frac{n}{n-1} \sum_{i=1}^{n} (z_i - n^{-1}) \ln z_i \,,$$

and

$$\tilde{\beta} = \bar{x}\,\tilde{\delta} \,, \qquad (2)$$

where $\bar{x}$ is the sample mean, $z_i = x_i/n\bar{x}$ for all values of i, and natural logarithms are indicated by the symbol ln.

Definitions (2) do not yield ML estimators. Instead, $\tilde{\delta}$ and $\tilde{\beta}$ stem from ML equations that are associated with a generalization of the distribution (1). Specifically, they arise from ML equations for the generalized gamma (gg) distribution discussed by Stacy and Mihram [13]. [Definition (1) is obtained from the gg distribution by setting one of its parameters, p, equal to unity and reparameterizing the result.] The statistics $\tilde{\delta}$ and $\tilde{\beta}$, consequently, are called quasimaximum likelihood (QML) estimators.

Arithmetic computations implied by (2) involve the use of logarithms, but otherwise are suitable for desk-calculator techniques. It will be shown that each of the estimators $\tilde{\delta}$ and $\tilde{\beta}$ is unbiased (Sections 6 and 8). Variances and the covariance of the proposed estimators are derived as functions of sample size and the two parameters α and β (Sections 7 and 8). It should be recalled in passing that comparable small-sample results are not available for either the MM or ML estimators associated with (1). Some asymptotic distributional properties of the proposed estimators are also discussed (Section 9).

Use of the Greenwood and Durand [4] ML aid is easy after the values for some ancillary statistics are determined. More specifically, natural logarithms of the sample geometric and arithmetic means are substituted into a simple function that yields a necessary argument to use with the tabular aid. Interpolation within the table virtually completes ML estimation whenever the computed argument lies within the range of argument values adopted for the table.

Table 1 Three samples from each of three Gamma populations.

	Gamma population no.								
	1			2			3		
	$\delta=0.5,\ \beta=2.0$			$\delta=1.0,\ \beta=2.0$			$\delta=2.0,\ \beta=2.0$		
	Sample no.								
Item	1	2	3	4	5	6	7	8	9
1	3.78	4.32	2.44	4.15	2.55	2.63	0.321	0.337	2.306
2	1.48	1.22	2.67	0.71	11.60	1.89	1.276	0.606	5.299
3	2.94	5.84	9.98	1.73	0.12	0.85	0.773	0.047	7.101
4	4.50	1.78	3.50	4.11	3.83	3.23	0.731	4.414	1.357
5	2.75	0.49	11.10	4.56	3.55	2.46	0.001	0.332	3.085
6	2.81	5.88	5.10	1.39	2.09	1.23	0.073	0.916	1.129
7	4.48	7.73	2.80	1.14	2.21	1.03	0.007	2.192	0.992
8	4.31	8.71	8.61	2.06	5.41	6.39	0.045	0.676	0.703
9	2.02	4.26	5.75	3.91	1.31	0.09	1.989	0.301	0.214
10	4.45	6.13	3.85	4.34	0.15	1.95	0.373	0.671	0.899
11	2.59	9.60	2.84	0.61	1.33	4.13	2.029	0.533	0.424
12	3.82	8.72	7.50	3.30	3.25	4.18	0.434	0.006	1.934
13	3.54	7.98	7.05	4.65	3.32	3.16	0.632	0.212	1.779
14	4.76	1.01	6.30	4.43	5.20	1.77	1.461	0.239	3.902
15	4.58	2.13	6.84	0.23	4.74	0.05	0.050	0.127	1.601
16	0.85	5.34	1.05	0.40	0.50	3.25	0.002	4.985	0.332
17	1.00	6.70	6.36	0.11	0.22	0.53	0.426	0.805	1.178
18	8.75	10.20	4.50	0.66	0.45	1.73	0.013	0.861	0.559
19	1.54	2.60	2.42	0.64	7.64	3.10	0.621	3.714	0.010
20	6.33	2.80	2.65	1.95	1.78	1.63	1.877	4.368	0.004
21	2.64	6.67	5.53	0.92	1.30	3.74	0.680	0.040	1.140
22	2.40	4.42	2.69	6.75	0.13	0.08	0.922	1.518	1.404
23	0.57	3.39	4.30	3.37	2.05	6.84	0.109	0.046	0.110
24	1.19	4.16	5.20	1.77	0.23	2.74	0.008	0.212	1.350
25	3.53	6.34	5.02	2.47	0.74	7.11	0.711	0.577	0.013
26	2.80	7.64	2.73	0.29	4.77	0.68	0.193	1.620	1.421
27	2.78	0.47	0.59	0.86	0.39	0.49	4.653	0.611	1.384
28	4.49	5.82	5.38	5.23	3.14	2.64	0.908	0.642	3.650
29	0.91	1.77	6.48	0.82	0.21	2.63	0.438	5.452	0.113
30	8.15	6.55	1.82	3.61	0.85	1.63	0.826	1.523	0.091
31	9.43	0.85	2.77	0.96	4.46	2.21	0.956	0.274	0.324
32	7.95	20.14	4.53	0.45	0.11	0.05	0.082	0.018	0.268
33	5.22	4.56	1.78	4.42	2.94	1.58	0.009	0.010	0.848
34	2.14	8.06	5.55	2.44	0.15	1.11	3.102	0.156	0.419
35	2.45	5.62	4.44	0.18	1.74	1.33	3.637	0.043	0.223
36	2.39	3.00	4.58	0.24	0.82	0.80	1.094	3.346	0.020
37	0.95	11.56	3.06	0.25	0.40	1.49	0.006	0.394	1.980
38	1.92	2.76	5.90	4.14	3.39	0.76	0.523	2.863	0.369
39	3.24	5.12	0.76	1.40	0.08	0.18	0.184	0.058	0.005
40	3.08	1.02	7.96	1.77	3.19	0.19	0.148	0.105	0.078
41	1.08	2.51	9.39	4.36	0.63	2.62	0.849	1.190	0.975
42	1.16	6.87	1.84	1.07	3.27	1.65	2.178	0.063	0.010
43	0.74	2.92	5.41	0.20	1.73	0.59	0.239	0.485	0.033
44	4.00	4.17	3.05	3.65	3.63	0.66	0.248	5.769	0.132
45	0.11	2.54	3.80	0.85	0.64	5.24	6.277	0.043	1.059
46	3.81	3.02	2.16	1.35	1.73	0.75	3.021	1.178	0.659
47	2.81	6.93	3.00	0.37	2.35	4.72	0.126	0.077	0.179
48	5.08	2.06	3.32	2.15	1.55	6.07	3.291	0.001	0.491
49	4.85	2.16	1.95	1.34	7.40	0.46	0.676	0.686	0.234
50	7.41	2.05	8.50	4.36	0.85	1.47	0.519	3.446	0.754

Table 2 Summary of population parameter values and estimation results.

	Gamma population no.								
	1			2			3		
	$\delta=0.5,\ \beta=2.0$			$\delta=1.0,\ \beta=2.0$			$\delta=2.0,\ \beta=2.0$		
	Sample no.								
Statistic	1	2	3	4	5	6	7	8	9
μ	4.000			2.000			1.000		
$\dot{\mu}$	3.411	4.972	4.536	2.142	2.342	2.156	0.995	1.176	1.091
$\hat{\mu}$	3.411	4.972	4.536	2.142	2.342	2.156	0.995	1.176	1.091
$\tilde{\mu}$	3.411	4.972	4.536	2.142	2.342	2.156	0.995	1.176	1.091
σ^2	8.000			4.000			2.000		
$\dot{\sigma}^2$	4.808	12.366	6.067	3.008	5.358	3.391	1.727	2.522	1.979
$\hat{\sigma}^2$	5.437	12.103	6.557	3.662	5.775	4.119	2.257	2.577	2.245
$\tilde{\sigma}^2$	5.344	12.294	6.551	3.655	5.810	4.025	1.846	2.688	1.965
δ	0.500			1.000			2.000		
$\dot{\delta}$	0.413	0.500	0.295	0.655	0.976	0.730	1.744	1.825	1.663
$\hat{\delta}$	0.467	0.490	0.319	0.798	1.054	0.886	1.905	1.864	1.648
$\tilde{\delta}$	0.459	0.497	0.318	0.796	1.059	0.866	1.864	1.945	1.651
β	2.000			2.000			2.000		
$\dot{\beta}$	1.410	2.487	1.338	1.404	2.287	1.573	1.735	2.145	1.814
$\hat{\beta}$	1.594	2.434	1.445	1.709	2.468	1.911	2.074	2.191	1.923
$\tilde{\beta}$	1.567	2.473	1.444	1.706	2.480	1.867	1.855	2.287	1.801
$\lvert\Delta\rvert$	0.145	0.195	0.182	0.148	0.139	0.120	0.086	0.093	0.103
$\lvert\dot{\Delta}\rvert$	0.082	0.080	0.089	0.139	0.090	0.078	0.079	0.100	0.061
$\lvert\hat{\Delta}\rvert$	0.099	0.080	0.077	0.132	0.089	0.077	0.074	0.096	0.061
$\lvert\tilde{\Delta}\rvert$	0.097	0.080	0.077	0.132	0.089	0.074	0.073	0.089	0.061

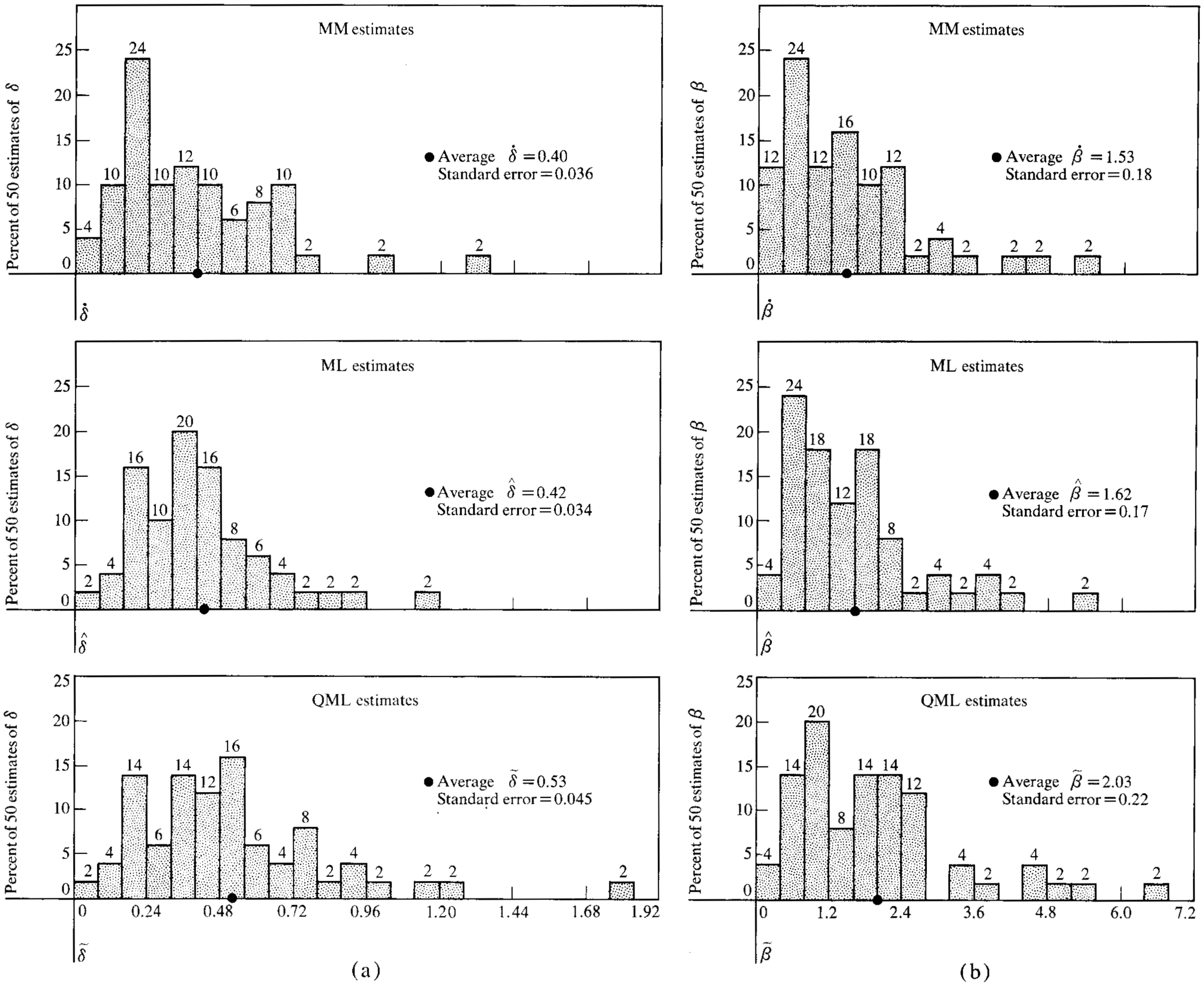


Figure 1 Estimates from samples of size five for a case in which $\delta = 0.5$ and $\beta = 2.0$. Histograms for (a) estimates of δ; (b) estimates of β.

The MM and ML statistics for estimating δ and β are denoted by the respective pairs: $\dot{\delta}$, $\dot{\beta}$ and $\hat{\delta}$, $\hat{\beta}$, with an obvious correspondence to the QML estimators. Formulas for the MM statistics are $\dot{\delta} = s^2/\bar{x}^2$ and $\dot{\beta} = \bar{x}\,\dot{\delta}$, in which s^2 is the sample variance defined according to

$$s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2/n$$

since the method of moments is to be strictly interpreted herein.

2. Summary of empirical findings

Some numerical examples are cited later in the discussion, data from nine random samples being itemized in Table 1. The samples were chosen from three chi-squared populations [special cases of (1)]. The samples were generated by an IBM System/360, Model 40, computer that was programmed in the light of the following well known results:

a) Pairs of independent, random observations from a rectangular (0,1) population are transformed to independent, random standard normal deviates by means of formulas due to Box and Muller [14].
b) The sums of the squares of r such deviates may be regarded as a single random observation from a chi-squared population having r degrees of freedom (a result previously used by Wilk, et al. [12] for similar purposes).

In Table 2, popular parameter values are shown along with estimation results from each of three methods: the MM, ML, and QML methods. The reader will find δ,

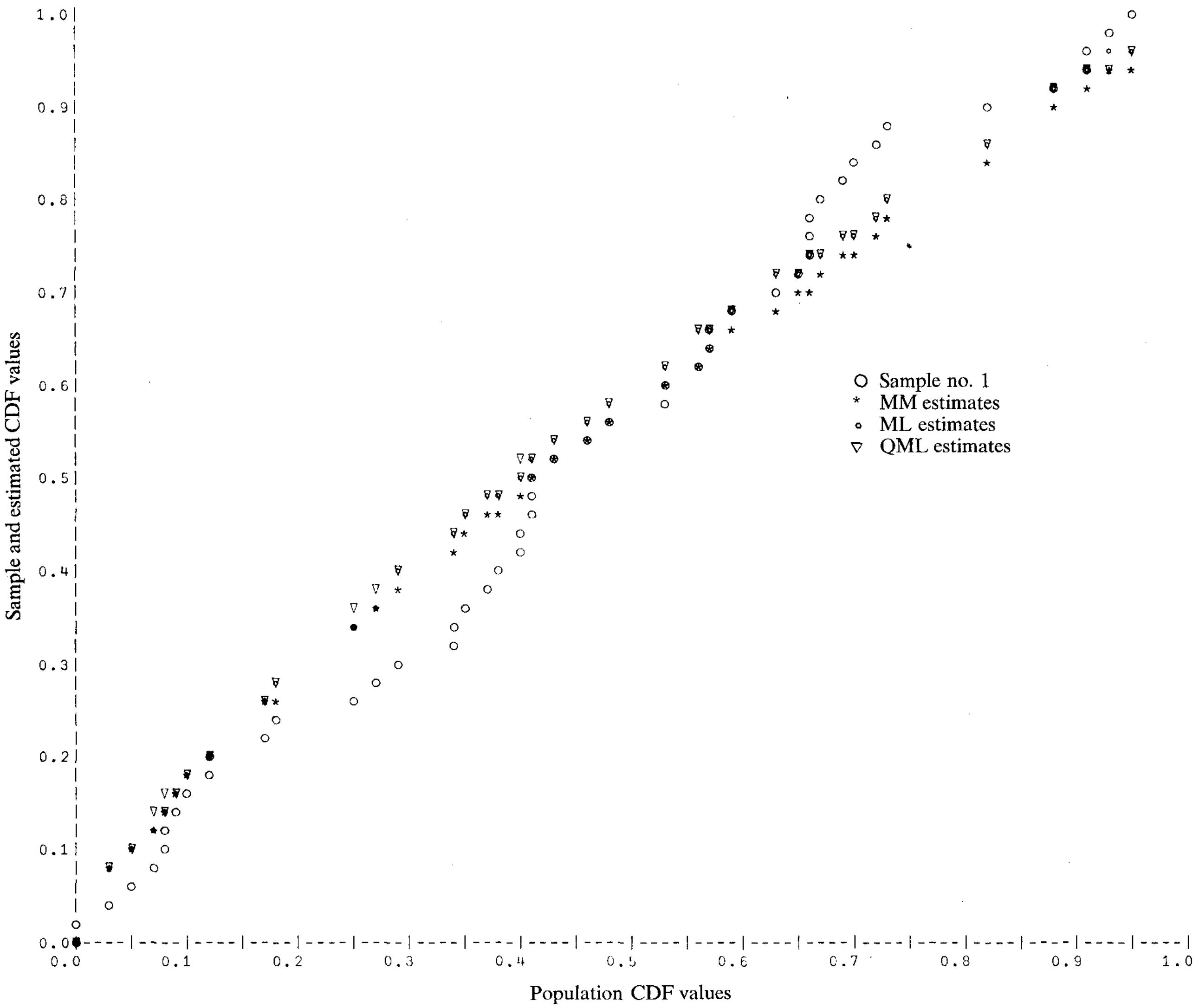


Figure 2 Estimated cumulated distribution functions (CDF) versus population CDF. CDF values were computed from Sample 1 data in Table 1 by the formula $(i-\frac{1}{2})/n$ at population CDF value for x_i; these values, represented by the large circles, are compared with estimated values obtained by the MM, ML, and QML methods.

β, $\dot{\delta}$, $\dot{\beta}$, $\hat{\delta}$, $\hat{\beta}$, $\tilde{\delta}$, and $\tilde{\beta}$ as linear titles therein, for example. Similar notation is adopted to indicate the population means and variances displayed in lines which are labeled μ, $\dot{\mu}$, $\hat{\mu}$, $\tilde{\mu}$ and σ^2, $\dot{\sigma}^2$, $\hat{\sigma}^2$, $\tilde{\sigma}^2$, respectively. Calculations to obtain those particular entries were in accordance with textbook formulas. (See, for example, Rao [15].)

Final lines in Table 2 reflect "goodness of fit." The Kolmogorov-Smirnov (K-S) goodness of fit statistic, showing the maximum deviation (absolute value) between sample and population cumulative distribution functions (CDF's), is indicated by $|\Delta|$. Ordinates for the sample CDF at x_i, $i = 1, \cdots, n$, were taken to be i/n in the usual way and the population CDF was determined by known values of the parameters. Using the previous notational device, the symbols $|\dot{\Delta}|$, $|\hat{\Delta}|$, and $|\tilde{\Delta}|$ have the obvious meaning. Corresponding numerical information provides a means for informal comparison of different estimation techniques with respect to "goodness of fit."

Tables of selected percentiles in the distribution for $|\Delta|$ are generally accessible (see Beyer [21]). Such tables provide an objective criterion for evaluating the goodness-of-fit, which is reflected by $|\dot{\Delta}|$, $|\hat{\Delta}|$ and $|\tilde{\Delta}|$.

The nine numerical examples lead to the informal conclusion that the QML estimators yield a goodness-of-fit that is quite comparable to that produced by ML estimators. Used in conjunction with the criterion implied in the preceding paragraph, Table 2 gives evidence of a slight superiority for the QML curve-fitting results. Perhaps, with samples of this size (50), that finding is not spurious since both $\tilde{\delta}$ and $\tilde{\beta}$ are demonstrably unbiased.

In Section 9, formulas for computing efficiency are given along with illustrative graphs. They show that QML estimators are not efficient, a fact that offsets to some extent the advantages that accrue because those estimators are unbiased. Some further insight is supplied by Monte Carlo results displayed in Figs. 1(a) and (b). Those figures concern a case, $\delta = 0.5$ and $\beta = 2.0$, chosen specifically because it is unfavorable from the standpoint of QML estimator efficiency. The figures summarize estimation results from 50 independent random samples of size five. They support the conjecture that small-sample ML estimation of δ and β leads to biased statistics. The MM estimators also appear to be biased.

In general, the results highlight possible bias and small variance for the ML statistics in contrast to larger variance and lack of bias for QML statistics. The author has calculated, for the data leading to Figs. 1(a) and (b), approximate 95 percent confidence intervals (assuming usual conditions are met for applying the Central Limit Theorem of statistics). Resulting intervals from both the MM and ML estimations do not contain the respective parameter values $\delta = 0.5$ and $\beta = 2.0$. Hence, the observed deviations of the averages from the respective known values appear to be significant for the MM and ML techniques.

3. Probability plots

Figures 2 and 3 are based upon probability plotting techniques that Wilk, et al. [12] have discussed along with relevant computation formulas and error evaluation. Suppose, for the sake of illustration, that a function $F_i(\delta,\beta)$ is defined by

$$F_i(\delta,\beta) = \int_0^{x_i} f(x)\,dx\,,$$

where $f(x)$ is given by (1). Several estimates of $F_i(\delta,\beta)$ are considered for probability plotting, viz., $(i-\frac{1}{2})/n$, $F_i(\dot{\delta},\dot{\beta})$, $F_i(\hat{\delta},\hat{\beta})$ and $F_i(\tilde{\delta},\tilde{\beta})$. Figure 2 refers to the first (left-most) sample in Table 1 and shows each of the above estimates plotted versus $F_i(\delta,\beta)$ for $i = 1,\cdots,n$. Similar graphs for the remaining eight samples could have been generated, of course.

Figure 3 shows corollary information derived from the sample. Specifically, values of a function $x_i(\delta,\beta)$, defined by the equation

$$\frac{i-\frac{1}{2}}{n} = \int_0^{x_i(\delta,\beta)} f(x)\,dx\,,$$

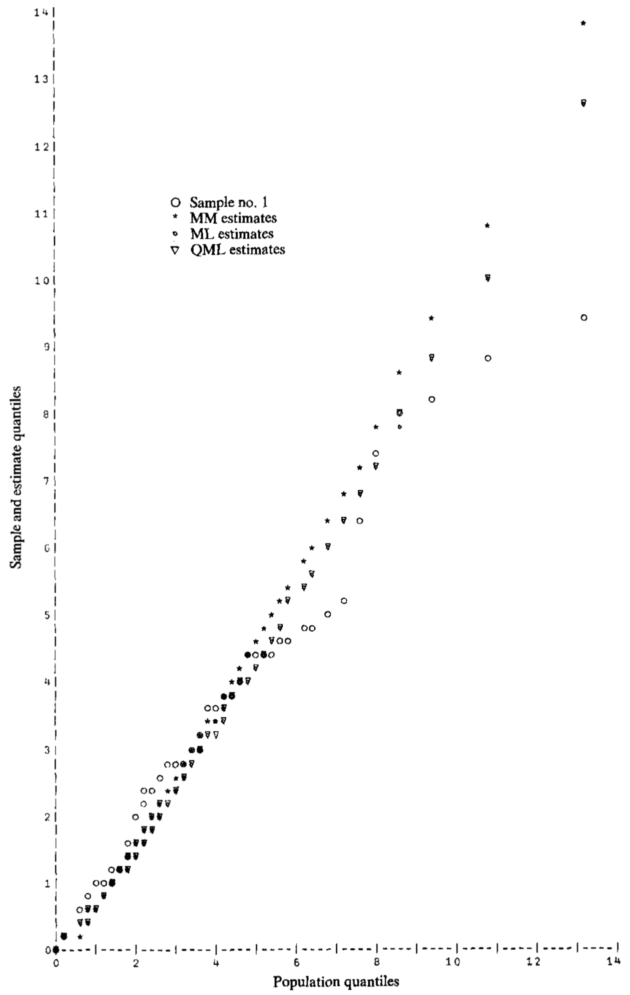


Figure 3 Estimated quantiles versus population quantiles, with values taken from Sample 1 in Table 1 for comparison.

are estimated by x_i, $x_i(\dot{\delta},\dot{\beta})$, $x_i(\hat{\delta},\hat{\beta})$, and $x_i(\tilde{\delta},\tilde{\beta})$. Figure 3 refers to the first sample in Table 1 and shows each of the estimates plotted versus $x_i(\delta,\beta)$.

4. Application of generalized Gamma distribution

The function (1) may be denoted unambiguously by $f(x;\beta,\alpha,1)$ if

$$f(x;\beta,\alpha,p) = \frac{|p|x^{p\alpha-1}}{\beta^{p\alpha}\Gamma(\alpha)} \exp\,[-(x/\beta)^p], \qquad x \geq 0, \tag{3}$$

and the definitions of x,α, and β are unchanged. Temporarily, it is convenient to consider $p \neq 0$ and to work with (3) rather than (1), with a promise to set $p = 1$ at a later time. The logarithm of the likelihood function corresponding to (3) is denoted by L_p so that

$$L_p = \ln\left[\frac{|p|^n\left(\prod_{i=1}^n x_i^{p\alpha-1}\right)}{\beta^{np\alpha}\,\Gamma^n(\alpha)} \exp\left\{-\sum_{i=1}^n (x_i/\beta)^p\right\}\right] \tag{4}$$

Suppose that the three first partial derivatives of L_p with respect to the parameters are equated to zero. This leads to

$$\begin{aligned} &-np\alpha + p\sum_{i=1}^n (x/\beta)^p = 0, \\ &-n\psi(\alpha) + p\sum_{i=1}^n \ln(x_i/\beta) = 0, \\ &\frac{n}{p} - n\alpha \ln\beta + \alpha\sum_{i=1}^n \ln x_i - \sum_{i=1}^n (x_i/\beta)^p \ln(x_i/\beta) = 0. \end{aligned} \tag{5}$$

These may be manipulated to yield the equivalent

$$\beta = (\bar{x}_p\,\alpha^{-1})^{1/p}, \tag{6a}$$

$$\psi(\alpha) - \ln\alpha = n^{-1}\sum_{i=1}^n \ln y_{ip}, \tag{6b}$$

$$\alpha = \left[\sum_{i=1}^n \left(z_{ip} - \frac{1}{n}\right) \ln z_{ip}\right]^{-1}, \tag{6c}$$

where

$$\begin{aligned} \bar{x}_p &= \sum_{i=1}^n x_i^p/n, \\ y_{ip} &= x_i^p/\bar{x}_p, \\ z_{ip} &= y_{ip}/n = x_i^p/\sum_{j=1}^n x_j^p \end{aligned} \tag{7}$$

and $\psi(\alpha)$ is the digamma function $(\partial/\partial\alpha)\ln\Gamma(\alpha)$.

Now consider any special case in which the value of p is known. The equivalent of Eq. (6c) would never arise. Yet, Eq. (6c) suggests an estimate of α whenever the value of p is known; and Eq. (6a) suggests a corresponding estimate of β. Suppose for the moment that $t_i = x_i^p$ for all values of i and that a symbol λ is introduced so that $\lambda = \beta^p$. The t_i values may be regarded [13] as a random sample from a gamma population defined by (1) with x replaced by t and β replaced by λ. Thus, the case in which p has a known value other than unity converts readily to the case $p = 1$ provided α and λ are to be estimated instead of α and β.

5. Independence of $\bar{x}_1$ and the z_{i1}

Equation (3) has served its intended purpose and, with the exception of Section 10, the remaining discussion will be restricted to the case $p = 1$. This justifies simpler notation than was adopted in (7), and we define

$$\bar{x} = \bar{x}_1, y_i = y_{i1}, \text{ and } z_i = z_{i1}.$$

Convenience is served also by defining

$$\tilde{\theta} = \frac{n-1}{n}\tilde{\delta} = \sum_{i=1}^n \left(z_i - \frac{1}{n}\right)\ln z_i \tag{8}$$

and by permitting the symbol x_i to play multiple roles, the specific role being determined according to the context of the discussion. The symbol x_i will represent a sample value, a random variable, or an argument in a density function as the case demands. Similarly, multiple roles will be assigned to the symbol z_i when no ambiguity results.

After appealing to the above expedients, the joint distribution of $x_1, \cdots, x_n$ is found to be

$$[\beta^\alpha\Gamma(\alpha)]^{-n}\left[\prod_{i=1}^n x_i^{\alpha-1}\right]\exp\left[-\sum_{i=1}^n x_i/\beta\right], \tag{9}$$

from which it can be shown that the set of z_i values is independent of $\bar{x}$. One merely transforms (9) by

$$x_i = n\bar{x}z_i, \qquad \text{for } i = 1, \cdots, n-1, \text{ and}$$

$$x_n = n\bar{x}\left(1 - \sum_{i=1}^{n-1} z_i\right),$$

noting that the Jacobian is $(n^n\bar{x}^{n-1})$. The indication is that, with the restriction $\sum_{i=1}^{n-1} z_i \leq 1$,

$$\begin{aligned} &\left[\frac{n^{n\alpha}}{\beta^{n\alpha}\Gamma(n\alpha)}\bar{x}^{n\alpha-1}\exp\left(-n\bar{x}/\beta\right)\right] \\ &\quad\times\frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)}\left(\prod_{i=1}^{n-1} z_i^{\alpha-1}\right)\left(1 - \sum_{i=1}^{n-1} z_i\right)^{\alpha-1} \end{aligned} \tag{10}$$

gives the joint distribution of $\bar{x}$ and $z_1, z_2, \cdots, z_{n-1}$. Hence $\bar{x}$ and the set of z_i variables are statistically independent.

6. Expectations of $\tilde{\theta}$ and $\tilde{\delta}$

The first moment for $\tilde{\theta}$, defined by (8), is obtained after noting that (for $i = 1, \cdots, n-1$)

$$E(z_i^s) = E\left(1 - \sum_{j=1}^{n-1} z_j\right)^s = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)}\frac{\Gamma(\alpha+s)}{\Gamma(n\alpha+s)}, \tag{11}$$

$$E(z_i^s \ln z_1) = E\left(\frac{\partial}{\partial s} z_i^s\right) = \frac{\partial}{\partial s}E(z_i^s), \tag{12}$$

$$E\left[\left(1 - \sum_{j=1}^{n-1} z_j\right)^s \ln\left(1 - \sum_{j=1}^{n-1} z_j\right)\right] = \frac{\partial}{\partial s}E\left(1 - \sum_{i=1}^{n-1} z_i\right)^s, \tag{13}$$

where the operator E indicates the mathematical expectation of the quantity that follows it, and the remaining notation has obvious meaning.

This, along with (8), indicates

$$E(\tilde{\theta}) = E\left(\sum_{i=1}^n z_i \ln z_i - n^{-1}\sum_{i=1}^n \ln z_i\right) \tag{14}$$

$$= n\left[\frac{\partial}{\partial s}E(z_j^s)\right]_{s=1} - \left[\frac{\partial}{\partial s}E(z_j^s)\right]_{s=0}, \tag{15}$$

where j has any of the values 1, 2, $\cdots$, n and $z_n = 1 - \sum_{i=1}^{n-1} z_i$. Referring to Eq. (11), we have:

$$\frac{\partial}{\partial s} E(z_i^s) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+s)}{\Gamma(n\alpha+s)} [\psi(\alpha+s) - \psi(n\alpha+s)]. \tag{16}$$

The results, (15) and (16), together with the recursive relation [16]

$$\psi(\alpha+1) = \psi(\alpha) + \frac{1}{\alpha}, \qquad \text{for } \alpha > 0, \tag{17}$$

leads to

$$E(\tilde{\theta}) = \frac{n-1}{n}\,\delta. \tag{18}$$

It becomes clear that an unbiased estimate of δ is obtained by defining $\tilde{\delta}$ so that

$$\tilde{\delta} = \frac{n}{n-1}\,\tilde{\theta}. \tag{19}$$

7. Variances of $\tilde{\theta}$ and $\tilde{\delta}$.

Notions in the preceding section may be extended to obtain the expected value of $\tilde{\theta}^2$. Toward that end, let

$$\begin{aligned} g(s) &= \frac{\partial^2}{\partial s^2} E(z_i^s) \\ &= \frac{\Gamma(n\alpha)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+s)}{\Gamma(n\alpha+s)} [\psi'(\alpha+s) - \psi'(n\alpha+s) \\ &\quad + \{\psi(\alpha+s) - \psi(n\alpha+s)\}^2] \end{aligned} \tag{20}$$

and, for $i \neq j$

$$\begin{aligned} h(s,t) &= \frac{\partial^2}{\partial s \partial t} E(z_i^s z_j^t) \\ &= \frac{\Gamma(n\alpha)}{\Gamma^2(\alpha)} \frac{\Gamma(\alpha+s)\Gamma(\alpha+t)}{\Gamma(n\alpha+s+t)} [-\psi'(n\alpha+s+t) \\ &\quad + \{\psi(\alpha+s) - \psi(n\alpha+s+t)\} \\ &\quad \times \{\psi(\alpha+t) - \psi(n\alpha+s+t)\}], \end{aligned} \tag{21}$$

where $\psi'(\gamma)$ is the trigamma function [16]

$$\psi'(\gamma) = \frac{\partial^2}{\partial \gamma^2} \ln \Gamma(\gamma) = \sum_{k=0}^{\infty} (\gamma+k)^{-2},$$

the right-hand series being particularly useful in Section 8. Thus,

$$\begin{aligned} E(\tilde{\theta}^2) &= nE\left\{\left(z_i - \frac{1}{n}\right)^2 (\ln z_i)^2\right\} \\ &\quad + n(n-1)\, E\left\{\left(z_j - \frac{1}{n}\right)\left(z_k - \frac{1}{n}\right)(\ln z_j)(\ln z_k)\right\} \end{aligned} \tag{22}$$

with $j \neq k$ in the final expectation.

After it has been recognized that $h(0, 1) = h(1, 0)$, this equation can be simplified to give

$$\begin{aligned} E(\tilde{\theta}^2) &= ng(2) - 2g(1) + \frac{g(0)}{n} \\ &\quad + n(n-1)\left\{h(1,1) - 2\,\frac{h(0,1)}{n} + \frac{h(0,0)}{n^2}\right\}. \end{aligned} \tag{23}$$

Details of further simplification are more easily verified if the terms are reordered and written as Eq. (24):

$$\begin{aligned} E(\tilde{\theta}^2) = {} & \frac{\alpha+1}{n\alpha+1}\left[\psi'(\alpha+2) - \psi'(n\alpha+2) + \left\{\frac{1}{\alpha} + \frac{1}{\alpha+1} - \frac{1}{n\alpha} - \frac{1}{n\alpha+1} + \psi(\alpha) - \psi(n\alpha)\right\}^2\right] \\ & + \frac{(n-1)\alpha}{n\alpha+1}\left[- \psi'(n\alpha+1) + \left\{\frac{1}{\alpha} - \frac{1}{n\alpha} - \frac{1}{n\alpha+1} + \psi(\alpha) - \psi(n\alpha)\right\}^2\right] \\ & - \frac{2}{n}\left[\psi'(\alpha+1) - \psi'(n\alpha+1) + \left\{\frac{1}{\alpha} - \frac{1}{n\alpha} + \psi(\alpha) - \psi(n\alpha)\right\}^2\right] \\ & -2\,\frac{n-1}{n}\left[- \psi'(n\alpha+1) + \left\{\frac{1}{\alpha} - \frac{1}{n\alpha} + \psi(\alpha) - \psi(n\alpha)\right\}\right. \\ & \qquad \left. \times \left\{ - \frac{1}{n\alpha} + \psi(\alpha) - \psi(n\alpha)\right\}\right] \\ & + \frac{1}{n}\,[\psi'(\alpha) - \psi'(n\alpha) + \{\psi'(\alpha) - \psi(n\alpha)\}^2] \\ & + \frac{n-1}{n}\,[- \psi'(n\alpha) + \{\psi(\alpha) - \psi(n\alpha)\}^2]. \end{aligned} \tag{24}$$

The recursive formula [16]

$$\psi'(\gamma+1) = \psi'(\gamma) - \frac{1}{\gamma^2}$$

indicates that each line of (24) can be written to display $\psi'(n\alpha)$ explicitly. It is convenient to note that the aggregate of coefficients outside the squared brackets is zero. Thus, the right-hand side of (24) clearly may be written so that both $\psi'(n\alpha)$ and $\{\psi(\alpha) - \psi(n\alpha)\}^2$ are absent. With similar motivation, it is noted that linear terms that involve $\{\psi(\alpha) - \psi(n\alpha)\}$ also vanish, i.e.,

$$\frac{\alpha+1}{n\alpha+1}\left[\frac{1}{\alpha} + \frac{1}{\alpha+1} - \frac{1}{n\alpha} - \frac{1}{n\alpha+1}\right] + \cdots - 2\,\frac{n-1}{n}\left[\frac{1}{\alpha} - \frac{2}{n\alpha}\right] = 0.$$

The above development, together with somewhat more tedious algebra reduces (24) to

$$E(\tilde{\theta}^2) = \frac{n-1}{n(n\alpha+1)}\left[\psi'(\alpha) + \frac{n}{\alpha}\right]. \tag{25}$$

This is combined with (18) to give the variance of $\tilde{\theta}$ as

$$\mathrm{var}(\tilde{\theta}) = \frac{n-1}{n(n\alpha+1)}\left[\psi'(\alpha+1) + \frac{n\alpha+1}{n\alpha^2}\right]. \tag{26}$$

The indication is that the variance of $\tilde{\delta}$ is

$$\mathrm{var}(\tilde{\delta}) = \frac{1}{(n-1)\alpha^2}\left[1 + \frac{n\alpha^2\,\psi'(\alpha+1)}{n\alpha+1}\right]. \tag{27}$$

8. Mean and variance for $\tilde{\beta}$

It will be recalled from definitions (2) that

$$\tilde{\beta} = \bar{x}\,\tilde{\delta},$$

and from (10), that $\bar{x}$ is independent of the z_i's. Thus $\tilde{\delta}$ and $\bar{x}$ are independent. Accordingly, it is a simple matter to obtain the first and second moment of $\tilde{\beta}$. Specifically,

$$E(\tilde{\beta}) = E(\bar{x})\,E(\tilde{\delta}) = \beta$$

and

$$E(\tilde{\beta}^2) = E(\bar{x}^2)\,E(\tilde{\delta}^2), \tag{28}$$

the final factor on the right-hand side of the last equation being obtainable directly from Eqs. (19)and (25). Thus the variance of $\tilde{\beta}$ is

$$\mathrm{var}(\tilde{\beta}) = \frac{\beta^2/\alpha}{n-1}\,[1 + \alpha + \alpha^2\,\psi'(\alpha+1)]. \tag{29}$$

It is an easy matter to establish the covariance of $\tilde{\delta}$ and $\tilde{\beta}$. One merely observes that the product of the estimators is proportional to $\bar{x}\tilde{\theta}^2$ so that

$$\begin{aligned} E(\tilde{\delta}\,\tilde{\beta}) - \delta\,\beta &= \left(\frac{n}{n-1}\right)^2 E(\bar{x})\,E(\tilde{\theta}^2) \\ &= \frac{n\beta[\alpha\psi'(\alpha) + n]}{(n-1)(n\alpha+1)} - \frac{\beta}{\alpha} \\ &= \beta\left[1 + \frac{n\alpha\psi'(\alpha+1)}{(n-1)(n\alpha+1)}\right]. \end{aligned} \tag{30}$$

9. Some properties of the estimators

The variance of any unbiased estimator of δ must be at least as large as

$$[n\alpha^4\,\psi'(\alpha)]^{-1}, \tag{31}$$

which results from application of the Cramer-Rao theorem [15]. Likewise, the variance of any unbiased estimator for β is at least as large as

$$\beta^2/n\alpha. \tag{32}$$

Direct comparision of these minimum variances with (27) and (29) yields the efficiency of each estimator $\tilde{\delta}$ and $\tilde{\beta}$. Respectively, these efficiences are

$$(n-1)/\left\{n[1 + \alpha^2\,\psi'(\alpha+1)]\left[1 + \frac{n\alpha^2\,\psi'(\alpha+1)}{n\alpha+1}\right]\right\} \tag{33}$$

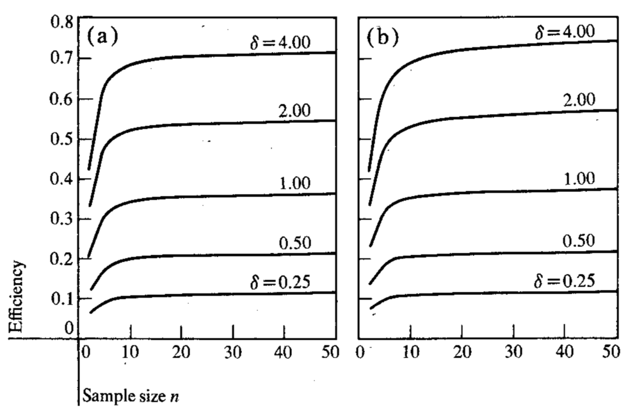
and

$$(n-1)/\{n[1 + \alpha + \alpha^2\,\psi'(\alpha+1)]\}, \tag{34}$$

whose forms result from applying the usual series expression for $\psi'(\alpha)$ and effecting some minor simplifications. Figure 4(a) shows the expression (33) plotted as a function of n for several values of α. Similar information relative to (34) is shown in Fig. 4(b).

It is noted that the joint distribution of the z_i's is Dirichlet [17] with parameters that are obvious from Eq. (10). The z_i's, therefore, are dependent. This makes it convenient to appeal to the Central Limit Theorem for dependent variables [18] to indicate that the asymptotic

Figure 4 Efficiency of QML estimators versus sample size and δ. (a) Efficiency of $\tilde{\delta}$; (b) efficiency of $\tilde{\beta}$.



distribution of $\tilde{\delta}$ is normal with mean δ and variance as shown by the right-hand side of (27). Likewise, it becomes obvious that $\tilde{\beta}$ is asymptotically normal with mean β and variance as given by the right-hand side of (29).

The variance of $\tilde{\delta}$ is a function of α; and the variance of $\tilde{\beta}$ is a function of both α and β. This raises the question of transformations that yield statistics whose variances are nearly independent of nuisance parameters. The Fisher z transformation [3] of the product moment correlation coefficient and transformations found in Chapter 5 of Rao [15], for example, indicate the notion. At this writing, no such transformation has been found for either $\tilde{\delta}$ or for $\tilde{\beta}$; and the problems remain for future research.

10. Miscellany concerning Eqs. (6)

It is instructive to consider other consequences of Eq. (6), which stem from the three-parameter generalized gamma density function (3). The equations admit at least two distinct solutions provided $x_i \neq \beta$ for $i = 1, \cdots, n$. Then, the likelihood function based upon (3) has a smooth local maximum in each of the two parameter space octants $\alpha, \beta, p > 0$ and $\alpha, \beta, -p > 0$.

Rigorous proof of the preceding statements would be lengthy. Hence, only a heuristic argument is presented. Assume $0 < x \neq \beta$, and consider the behavior of the function (3) within the indicated octants of the parameter space. It is a straight-forward matter to show that the function tends to zero when any parameter tends to infinity or when the point (α,β,p) approaches a boundary of either octant from within. Furthermore, the function is a positive, continuous function of its parameters. This makes it clear that the function (3), and consequently the likelihood function, has a smooth maximum in each of the octants defined by $(\alpha,\beta,p > 0)$ and $(\alpha,\beta,-p > 0)$.

Numerical solution of Eqs. (6) is simple in concept. One rearranges (6b) so that its right-hand member becomes zero, and eliminates α from the result by substitution from (6c). That yields an equation with a single unknown, the parameter p. Let this be written $h(p) = 0$. Its roots can be found by iterative computer technique. Each root leads to a unique, simultaneous solution of all equations in (6). At the present writing, however, there is no way to know when all solutions have been found.

This rather unusual situation is illustrated by applying the method to a random sample size 20, taken by Menon [19] from a population in which $p = 0.5$, $\alpha = 1.00$, and $\beta = 2.72$. Figures 5 and 6 show the graph of $h(p)$ for the Menon data and indicate the existence of at least three solutions to the corresponding likelihood equations. Reading from left to right, these are called root 1, root 2, and root 3. Table 3, which accompanies the figures, indicates solutions of Eqs. (6) for each root that has been found, and the corresponding Kolmogorov-Smirnov type maximum deviations, (max $|\hat{\Delta}|$) as defined for Table 2.

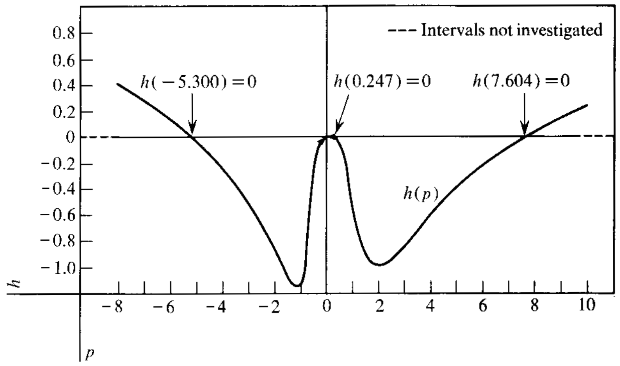


Figure 5 The function $h(p)$ from ML equations (6).

Figure 6 Detail of $h(p)$ near origin.

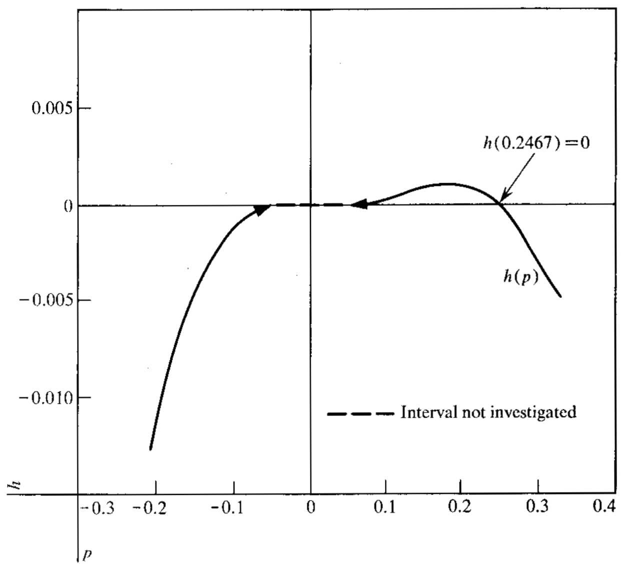


Table 3 Summary of maximum likelihood estimation (Menon's [19] data for generalized gamma distribution).

Statistic	True value	Basis of estimates [Root of $h(p) = 0$] 1	2	3	Sample value
p	0.500	−5.300	0.247	7.604	
α	1.000	0.029	3.986	0.029	
β	2.718	0.001	0.004	61.542	
$\lvert\hat{\Delta}\rvert$	0.270	0.910	0.159	0.385	
μ[13]	5.456	0.000	3.375	11.672	4.422
σ^2[13]	147.781	*	112.935	283.210	154.398

*2/(−5.3) < −0.029 (see [13]).

Solution of Eqs. (6) has been programmed for an IBM System/360. The Newton-Raphson root-finding algorithm (Grove, [20]) was considered, but abandoned in favor of an interval halving convergence technique described in Chapter 1 of [20]. Reference 16 provided ready access to formulas for computing $\psi(\alpha)$, as required in $h(p)$ and for computing the CDF corresponding to (3). Calculations for the CDF values are similar to those discussed by Wilk, et al. [12]. Bias correction to the estimate of β is attempted by substituting the estimated α value in the bias correction factor from formula (8) in Stacy and Mihram [13].

Printout from the program provides a visual comparison of the sample histogram and the continuous curve (3) based upon specific parameter values, e.g. any solution of (6). Similar comparison of the corresponding sample and continuous cumulative distributions is also accommodated. The maximum deviation between these two cumulative curves is included as a part of the program printout. Sample and estimated population percentiles are printed upon demand to the computer.

Acknowledgments
Thanks are extended to Miss Alice Jones and Dr. Alan Jones for vital programming assistance and to Mr. Gerald Ratchford whose data analysis requests motivated early completion of the computer programs.

References
1. S. S. Gupta and P. A. Groll, "Gamma distribution in acceptance sampling based on life tests," *J. Am. Stat. Assoc.* **56**, 942 (1961).
2. E. G. Chambers and G. U. Yule, "Theory and observation in the investigation of accident causation," *Suppl. J. Roy. Stat. Soc.* **7**, 89 (1941).
3. R. A. Fisher, "On the 'probable error' of a coefficient of correlation deducted from a small sample," *Metron* **1**, 1 (1921).
4. J. A. Greenwood and D. Durand, "Aids for fitting the gamma distribution by maximum likelihood," *Technometrics* **2**, 55 (1960).
5. M. B. Wilk, R. Gnanadesikan and M. J. Huyett, "Estimation of parameters of the gamma distribution using order statistics," *Biometrika* **49**, 525 (1962).
6. H. L. Harter and A. H. Moore, "Asymptotic variances and covariances of maximum likelihood estimators, from censored samples, of the parameters of the Weibull and gamma populations," *Ann. Math. Stat.* **38**, 557 (1967).
7. A. C. Cohen, "Estimating parameters of Type III populations from truncated samples," *J. Am. Stat. Assoc.* **45**, 411 (1950).
8. Des Raj, "Estimation of the parameters of Type III populations from truncated samples," *J. Am. Stat. Assoc.* **48**, 336 (1953).
9. D. C. Chapman, "Estimation of the parameters of a truncated gamma distribution," *Ann. Math. Stat.* **27**, 498 (1956).
10. W. R. Blischke, E. J. Brady and P. B. Mundle, "Further results on the estimation of the parameters of the Pearson Type III distribution in the regular and nonregular cases," Aerospace Research Labs., Wright-Patterson Air Force Base, Report ARL 70-0017, AD705207 (1970).
11. W. R. Blischke, "Further results on the estimation of the parameters of the Pearson Type III distribution," Aerospace Research Labs., Wright-Patterson Air Force Base, Report ARL 71-0063, AD728677 (1971).
12. M. B. Wilk, R. Gnanadesikan and M. J. Huyett, "Probability plots for the gamma distribution," *Technometrics* **4**, 1 (1962).
13. E. W. Stacy and G. A. Mihram, "Parameter estimation for a generalized gamma distribution," *Technometrics* **7**, 349 (1965).
14. G. E. P. Box and M. E. Muller, "A note on the generation of random normal variates," *Ann. Math. Stat.* **29**, 610 (1958).
15. C. R. Rao, *Advanced Statistical Methods in Biometric Research*, John Wiley & Sons, Inc., New York, 1952, pp. 176–220.
16. National Bureau of Standards, *Handbook of Mathematical Functions, Applied Mathematics, Series 55*, U.S. Dept. of Commerce, 1964, pp. 255–266.
17. S. S. Wilks, *Mathematical Statistics*, John Wiley & Sons, Inc., New York, 1962, pp. 177–82.
18. W. Hoeffding and H. Robbins, "The central limit theorem for dependent variables, *Duke Math. J.* **15**, 773 (1948).
19. M. V. Menon, "Estimation of the shape and scale parameters of the Weibull distribution," *Technometrics* **5**, 175 (1963).
20. W. E. Grove, *An Introduction to Elementary Numerical Methods*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1966, pp. 18–9.
21. W. H. Beyer, *Handbook of Tables for Probability and Statistics*, The Chemical Rubber Publishing Co., Cleveland, 1966, pp. 321–2.

Received August 31, 1972

The author is located at the IBM System Products Division Laboratory in Endicott, New York 13760.