

## EFFICIENT ESTIMATION OF INCOME DISTRIBUTION PARAMETERS\*

Teun KLOEK and Herman K. VAN DIJK

*Econometric Institute, Erasmus University, Rotterdam, The Netherlands*

Received October 1976, final version received October 1977

The parameters of several families of distributions are estimated by means of minimum  $\chi^2$ ; use is made of random samples taken from Dutch income-earning groups in 1973. The numerical search routine used, is the Complex method due to Box. The  $\chi^2$  function is evaluated by standard numerical integration procedures. The lognormal and the Gamma families are rejected because of a poor fit. The log  $t$  and the log Pearson IV families are introduced. This results in a considerable improvement of  $\chi^2$  critical levels. The generalized Gamma and the Champernowne function describe the income distribution reasonably well in some cases.

### 1. Introduction

The most attractive way to characterize an empirical distribution is by means of a small number of parameters. Preferably, these parameters should have a clear interpretation and define a theoretical distribution with a good fit to the empirical data. These statements hold for statistical distributions in general.

In this paper we study size distributions of income. In this context theoretical distributions might serve several purposes, of which we mention the following one. Suppose one is interested in the consequences of a systematic change of an income distribution, which can be described in terms of changing one parameter. If one wants to study the consequences of such a change for aggregate consumption of individual commodities, one does not only need Engel curves but also a carefully specified income distribution.<sup>1</sup>

The best-known families of theoretical distributions, used so far in the field of income distributions, are the Pareto, lognormal, and Gamma families; see

\*A preliminary version of this paper was presented at the European Meeting of the Econometric Society in Helsinki, August 1976. The authors are indebted to a referee for a number of helpful comments.

<sup>1</sup>This was already emphasized by de Wolff (1941). We are indebted to J.S. Cramer for drawing our attention to this reference.

Cramer (1969, ch. 4), Aitchison and Brown (1957), and Salem and Mount (1974). These families, however different in other respects, have in common that their members are characterized by (only) two parameters. The interpretation of these parameters does not present great problems, but as regards fit, their performance is rather poor,<sup>2</sup> which is probably due to oversimplification. More precisely, members of the Pareto family very well fit the higher incomes, but they are inappropriate for the middle and lower incomes. Members of the lognormal family perform well in the central part of the distribution but their tails are not heavy enough to fit empirical data well; compare Aitchison and Brown (1957, pp. 117, 118) and Cramer (1969, fig. 22). The weak points of the Gamma family are similar to that of the lognormal family.

In this paper we explore possibilities to obtain considerable improvements in fit by adding a small number of parameters.<sup>3</sup> Starting from the lognormal family of distributions we tried to reach a greater flexibility by replacing the normal family by the Student-*t* family. The latter allows for heavier tails than the former but is still symmetric in log income. Skewness may be introduced by using the Pearson type IV family. Starting from the Gamma family we investigated the generalized Gamma family, described by Stacey (1962). Finally, we gave some attention to a type of distribution proposed by Champernowne. A more detailed discussion of these families of distributions will be given in section 3.

Our data are all of the class frequency type. The distribution parameters were estimated using the criterion of minimum chi-square. The numerical optimization approach we used is well-known in other contexts but so far it seems not to have been used in the field of income distributions. More details on the estimation method employed will be given in section 2. In section 4 we present our empirical results. Section 5 contains a number of conclusions. Details on the data are given in the appendix.

## 2. An efficient estimation method using minimum chi-square

The data we consider are the results of a random sample of size  $n$ . They are recorded in the form of class frequencies. Let  $n_i$  ( $i=1, \dots, r$ ) denote the observed class frequencies corresponding to  $r$  mutually exclusive and exhaustive income classes and  $p_i(\theta)$  the corresponding (theoretical) probabilities,

<sup>2</sup>Salem and Mount (1974) report positive results with respect to the Gamma family, but our results do not confirm their conclusion; see section 4, below. This might be due to the fact that we studied a different data set; compare also footnote 7.

<sup>3</sup>For completeness we emphasize that our goal is merely descriptive. The much more ambitious goal of explaining income differences by means of analyzing supply and demand in the labor market requires much more and better data than usually are available; see Somermeyer (1977) for a theoretical analysis. For the estimation of a simple reduced form equation based on data from the same research project, we refer to Jansen and Kloek (1976).

where  $\theta$  represents the  $s$  parameters of the theoretical distribution considered. Then we have

$$p_i(\theta) = \int_{a_i}^{a_{i+1}} f(y; \theta) dy, \quad (1)$$

where  $y$  is income,  $(a_i, a_{i+1})$  is the  $i$ th income interval and  $f(y; \theta)$  is the density of the theoretical distribution.

A well-known criterion for fitting theoretical distributions to empirical distributions is Pearson's chi-square ( $\chi^2$ ) criterion,

$$\chi^2(\theta) = \sum_{i=1}^r \frac{\{n_i - np_i(\theta)\}^2}{np_i(\theta)}, \quad (2)$$

which can be interpreted as a generalized least-squares or minimum-distance criterion. For large samples, the criterion (2) follows a  $\chi^2$  distribution with  $r-1$  degrees of freedom, if the theoretical probabilities  $p_i(\theta)$  are known; see Cramér (1946, sec. 30.1). In our case the  $s$ -vector  $\theta$  is to be estimated from the sample. Then the criterion (2) follows a  $\chi^2$  distribution with  $r-s-1$  degrees of freedom; see Cramér (1946, sec. 30.3) and the references cited there. Estimators of  $\theta$  obtained by minimizing (2) are known to be asymptotically normal and efficient.

The derivatives of (2) with respect to  $\theta_j$  can easily be expressed in terms of the derivatives  $\partial p_i / \partial \theta_j$  but the latter require numerical integration for all distributions considered. Due to recent advances in computer technology and numerical optimization methods it is now possible to minimize (2) straightforwardly by direct search. This means that one need not provide expressions for derivatives but just evaluates the criterion function in a certain number of points.<sup>4</sup> Several search methods of this type are available; see, e.g., Beveridge and Schechter (1970) or Himmelblau (1972). The one we used is called the Complex method and is due to Box (1965).

The evaluation of  $p_i(\theta)$  has been carried out by making use of standard numerical integration routines. We experimented with an iterative ninth-degree Newton-Cotes formula; see Abramowitz and Stegun (1964). Experiments with the Gauss-Legendre quadrature yielded the same results. For certain distributions, use was made of the restriction  $\sum p_i = 1$  to find the multiplicative constant, while for other cases, where the integration constant is known, a check on the accuracy of the integration process was possible.

Since the chi-square values are not comparable if the numbers of parameters (and hence the degrees of freedom) are different, we computed critical

<sup>4</sup>An implication of this approach is that no linearization is required such as proposed by Aigner and Goldberger (1970) and used by Kakwani and Podder (1976).

levels, which are comparable. Such a critical level is defined as

$$P(\chi^2 \geq x) = \int_x^\infty \frac{1}{2^{\frac{1}{2}v} \Gamma(\frac{1}{2}v)} e^{-\frac{1}{2}t} t^{\frac{1}{2}v-1} dt, \quad (3)$$

where  $x$  is the empirical minimum  $\chi^2$  value and  $v$  is the number of degrees of freedom (in our case  $v = r - s - 1$ ). These integrals were computed by means of Gauss-Laguerre quadrature.

These critical levels can be used for testing purposes, in the sense that one may choose to reject theoretical distributions which have critical levels below 0.05. For more complicated tests comparing separate families of hypotheses see Cox (1961) and Goldfeld and Quandt (1972, ch. 5, and the references cited there). We have confined ourselves, however, to computing the simple statistics mentioned above.

### 3. Some families of theoretical distributions

As announced in the Introduction, we studied generalizations of both the lognormal and the Gamma families. The first of these is the log-Student family. This is obtained by applying the transformation  $t = \log y$ , where  $y$  is income, to a general Student density. The resulting density formula reads

$$f(y; \mu, \tau, v) = \frac{v^{\frac{1}{2}v}}{B(\frac{1}{2}, \frac{1}{2}v) \tau} \cdot \frac{1}{y} \left[ v + \frac{(\log y - \log \mu)^2}{\tau^2} \right]^{-\frac{1}{2}(v+1)}, \quad (4)$$

where  $y > 0$ ,  $\mu > 0$ ,  $\tau > 0$ ,  $v > 0$ . The location of such a distribution is described by the median  $\mu$ , the dispersion by the standard deviation ( $\sigma$ ) of log income which is obtained from  $\sigma^2 = v\tau^2/(v-2)$  if  $v > 2$ . The degrees of freedom parameter  $v$  characterizes the tails. At  $v=1$  we have a heavy-tailed log-Cauchy distribution; as  $v \rightarrow \infty$  the distribution tends to a light-tailed lognormal distribution. For  $v > 4$ , the kurtosis coefficient equals  $6/(v-4)$ . The symmetry property of the distribution of log income is preserved.

The next step is the introduction of the possibility of skewness of the distribution of log income. This can conveniently be done by considering the family of distributions which is known as Pearson's Type IV; see Elderton and Johnson (1969, pp. 42, 43). The density function of such a distribution is proportional to the product of the right hand side of (4) and

$$\exp\{-\lambda \arctan(x/a)\}, \quad (5)$$

where  $x = \log(y/\mu)$  and  $a = \tau\sqrt{v}$ . Clearly,  $\lambda$  is a skewness parameter. For  $\lambda = 0$  we obtain the symmetric Student family for the distribution of log income. For  $\lambda < 0$  the distribution of log income is skew to the right, for  $\lambda > 0$  it is

skew to the left. As before,  $\mu$  is a location parameter,  $\tau$  a scale parameter and  $\nu$  a kurtosis parameter. Simple interpretations in terms of medians and variances are no longer possible. It is easily verified that the mode of the distribution of log income is given by

$$\text{mode} = \log \mu - \lambda \tau \nu^{1/2} / (\nu + 1). \quad (6)$$

Note that the mode of log income is not equal to the log of the mode of income. The expression for the latter is more complicated than (6). It follows from (6) that, if the mode has a small sampling variance, the correlation between the estimators  $\log \hat{\mu}$  and  $\hat{\lambda}$  will be positive and may be large in absolute value. For that reason it might be advisable to replace  $\mu$  by the mode as a parameter in the computations.<sup>5</sup> We shall denote this family of distributions by the term log Pearson Four.

Another well-known class of theoretical density functions is the Gamma family. In its standardized form the density formula reads

$$f(z; r) = \frac{e^{-z} z^{r-1}}{(r-1)!}, \quad z \geq 0, \quad (7)$$

where  $r > 0$  and  $(r-1)! \equiv \Gamma(r)$ . The mean and variance are equal to  $r$  and the mode equals  $\max\{0, r-1\}$ . So the skewness of the function is positive. Note that this function is applied to the distribution of income, rather than log income. Recently Salem and Mount (1974) used the transformation  $z = y\beta$ , with  $\beta > 0$  and fitted the resulting density function to class frequencies of income data. Our results with this transformation were not very successful; see section 4.

For that reason we considered a more general transformation, mentioned by Stacey (1962), which includes several well-known transformations of the standard Gamma function. Let  $z = (y/a)^p$  with  $a > 0$ ,  $p \neq 0$ , then we have

$$f(y; r, a, p) = \frac{|p| y^{pr-1} e^{-(y/a)^p}}{a^{pr} (r-1)!}, \quad y \geq 0. \quad (8)$$

The interpretation of the parameters is not so easy. Obviously,  $a$  is a scale parameter. The degrees of freedom parameter  $r$  determines the skewness of

<sup>5</sup>The high correlation between the estimators  $\log \hat{\mu}$  and  $\hat{\lambda}$  gives rise to an elongated valley in the surface of the chi-square function [the term valley is used here in the sense as one finds in books on nonlinear optimization; see, e.g., Himmelblau (1972, p. 83)]. In order to avoid search problems along such a valley one may reparameterize the function as indicated above. We hope to discuss this type of problems in more detail in a forthcoming paper.

the function, but so does<sup>6</sup>  $p$ . The mode of the function is given by

$$\text{mode} = \max \{0, a(r-1/p)^{1/p}\}, \quad (9)$$

when  $p=1$ , one obtains the so-called Gamma-1 function. Similarly for  $p=-1$  and  $p=-2$  one obtains the Gamma-1 inverse function and the Gamma-2 inverse function, respectively; see Raiffa and Schlaifer (1961, ch. 7). Our results with (8) were not as good as the results based on the log Pearson Four family.

It is interesting to observe that, if  $p < 0$ , the right-hand tail is asymptotically Pareto in the following sense. Let  $h > 1$  be a fixed number. Then

$$\lim_{y \rightarrow \infty} \log \frac{f(hy; r, a, p)}{f(y; r, a, p)} = \begin{cases} (pr-1) \log h & \text{if } p < 0, \\ -\infty & \text{if } p > 0. \end{cases} \quad (10)$$

A similar exercise for the Pareto function

$$f(y; \alpha, y_0) = \frac{\alpha}{y_0} \left( \frac{y}{y_0} \right)^{-\alpha-1} \quad \text{if } y > y_0, \quad (11)$$

yields

$$\log \frac{f(hy; \alpha, y_0)}{f(y; \alpha, y_0)} = (-\alpha-1) \log h. \quad (12)$$

We conclude that  $pr$  plays the same role in the right tail of (8) as  $-\alpha$  in the Pareto density. So one may reparameterize (8) in terms of  $\alpha$  and  $r$  by using  $p = -\alpha/r$ .

Another family of income distributions was suggested by Champernowne (1937). We studied the following four-parameter density function:

$$f(y; y_0, \alpha, \beta, \gamma) \propto \left[ \left( \frac{y}{y_0} \right)^{1-\gamma} + 2 \cos \pi \beta \left( \frac{y}{y_0} \right)^{1-\frac{\gamma-\alpha}{2}} + \left( \frac{y}{y_0} \right)^{1+\alpha} \right]^{-1}, \quad (13)$$

where  $y > 0$ ,  $y_0 > 0$ ,  $\alpha > 0$ ,  $0 < \beta < 2$ ,  $\gamma > 1$ . The parameter  $y_0$  is a scale parameter, and the parameters  $\alpha$  and  $\gamma$  have interpretations of inequality measures for high and low incomes, respectively. The parameter  $\beta$  apparently

<sup>6</sup>The numerical search routine described in section 2 broke down on collinearity between  $a$  and  $p$ . So we used the mode instead of  $a$  as an independent parameter in the numerical optimization procedure. Compare also footnote 5.

has been introduced to improve the flexibility of this family in the middle income range, but it is hard to give a clear interpretation. The right-hand tail of the density function (13) is asymptotically Pareto, in the sense indicated above.

Finally, we mention some families of distributions which we do not consider in the empirical part of our paper. The first of these is the so-called sech square family which has been suggested in the literature as a possible description of the size distributions of income. This family is related to the Champernowne family; see the discussion by Fisk (1961) of Champernowne's article (1952). Another family of distribution functions, related to the Champernowne function, is obtained by application of the Box-Cox transformation to a general Student- $t$  family; see Box and Cox (1964). Consider

$$t = \frac{(y/\mu)^\lambda - 1}{\lambda}, \quad (14)$$

where  $t$  is distributed according to the general Student family with zero median. The variable  $y$  stands for income and is restricted to be positive. The resulting density of  $y$  resembles (13). When  $\lambda \rightarrow 0$  one obtains the log Student- $t$  function. Now, consider the difference

$$D = \frac{(y/\mu)^\lambda - 1}{\lambda} - \log(y/\mu), \quad (15)$$

with  $y > 0$ ,  $\lambda \neq 0$ . For  $y \neq \mu$  this difference is positive when  $\lambda > 0$  and negative when  $\lambda < 0$ . So the Box-Cox transformation introduces more flexibility with respect to the behavior of the tails of theoretical distributions. However, this transformation introduces also a restriction on the range of the random variable  $t$ . Since  $y$  is restricted to be positive, it follows that (14) maps the interval  $(0, \infty)$  into  $(-1/\lambda, \infty)$  if  $\lambda > 0$ , and into  $(-\infty, -1/\lambda)$  if  $\lambda < 0$ . For instance, when  $\lambda = 1$ , the Student- $t$  variable should be greater than  $-1$  with probability one. The error introduced by these assumptions, of course, depends on the size of  $\lambda$  and  $\tau$ . This means that  $\lambda$  should be close to zero unless  $\tau$  is small.

#### 4. Empirical results

The data used result from random samples taken from several populations of income earning groups in the Netherlands in the spring of 1973. They were collected for a research project entitled 'Economic Consequences of Taxation', a joint project of two institutes of the Erasmus University, Rotterdam. These data are listed in the appendix.

It seems appropriate to comment briefly on the sampling procedure. As a starting point a random sample of municipalities in the Netherlands was taken. From each of these, a sample of residences was taken at random. Finally one income-earning person per residence was selected at random and interviewed. In this way income earners from residences with two or more income earners are underrepresented. One could correct for this by a reweighting procedure and a modification of the  $\chi^2$ -criterion. We preferred, however, to delete all families with two or more income earners not only for simplicity, but mainly in the interest of homogeneity of the subpopulation.<sup>7</sup> It was possible to split this representative sample into four subsamples, i.e.,

Table 1  
Chi-square critical levels ( $\times 10^2$ ) of several families of income distributions.

Distribution	No. of par.	Wage earners (9) <sup>a</sup>	Self-employed (11)	Free professions (9)	Old-aged (6)	Representative sample (12)
Lognormal	2	0.003	1.637	1.247	7.925	0.000
Gamma	2	0.000	0.000	50.347	0.718	0.000
Log <i>t</i>	3	0.028	13.109	48.638	19.668	6.020
Generalized						
Gamma	3	1.956	13.520	63.496	25.125	0.000
Champernowne	4	1.405	22.805	80.917	23.577	3.763
Log Pearson						
Four	4	1.254	20.102	80.536	21.110	6.295

<sup>a</sup>Numbers between brackets denote numbers of income intervals; see appendix.

wage earners, self-employed persons, free-profession members and old-age persons. Since three of these subsamples were very small, they were augmented with the results from some additional samples taken from the relevant groups. So we ended up with five samples, see the appendix.

The first stage of our numerical work consisted of estimating the lognormal and Gamma families, by means of the minimum  $\chi^2$  method discussed in section 2.<sup>8</sup> We used the optimum  $\chi^2$  value found and computed corresponding critical levels. Numerical results are reported in table 1. The results indicate that the two-parameter distributions considered exceed the five percent critical level in only one case each. This poor result was already mentioned in the introduction. We conclude that these families are too simple to be realistic.

<sup>7</sup>Notice that this procedure implies that our incomes refer to personal incomes and family incomes at the same time. This does not hold for other data sets, as a rule. One should be aware of this fact when comparing different empirical studies.

<sup>8</sup>The Beta-2 inverse function was also used in the first stage; see Raiffa and Schlaifer (1961, ch. 7). Since the results were not very good, except for the self-employed, we decided to drop this family.



Table 2  
Parameter estimates of the lognormal, log  $T$  and log Pearson Four families.<sup>a</sup>

	$\mu$	$\tau^2$	$\nu$	$\lambda$	$\chi^2$
<i>Wage earners</i>					
Lognormal	20.65	0.21			30.95
Log $t$	20.14	0.15	5.50		23.46
Log Pearson Four	10.95	0.14	10.20	-5.00	12.75
<i>Self-employed</i>					
Lognormal	25.88	0.48			18.73
Log $t$	24.64	0.31	3.71		13.40
Log Pearson Four	16.61	0.30	4.55	-1.38	8.54
<i>Free professions</i>					
Lognormal	63.37	0.53			16.25
Log $t$	71.05	0.26	1.68		8.49
Log Pearson Four	108.11	0.25	2.69	1.29	1.62
<i>Old-aged</i>					
Lognormal	9.65	0.24			6.78
Log $t$	9.28	0.10	1.40		3.25
Log Pearson Four	8.26	0.11	1.61	-0.39	1.56
<i>Representative group</i>					
Lognormal	17.54	0.44			55.35
Log $t$	17.26	0.24	3.51		14.95
Log Pearson Four	18.74	0.24	3.35	0.28	13.40

<sup>a</sup>For the case of the lognormal,  $\tau^2$  is to be interpreted as the variance ( $\sigma^2$ ) of log income.

Table 3  
Parameter estimates of the generalized Gamma family.

Group	Mode	$r$	$p$	$a$	$\chi^2$
Wage earners	15.62	2.98	-1.39	39.96	13.14
Self-employed	15.17	3.38	-0.85	91.18	11.08
Free professions	51.01	0.92	1.77	91.06	3.42
Old-aged	6.91	1.22	-2.17	8.77	2.76
Representative group	11.04	22.41	-0.33	201,817.16	50.94

In the second stage we considered three- and four-parameter families. The corresponding chi-square critical levels are also reported in table 1. The generalized Gamma and Champernowne families score three times out of five a critical level greater than five percent, while the log  $t$  and the log Pearson Four reach the score of four out of five. It is remarkable that none of our distributions was able to reach a good fit in the case of the wage earners. Another noteworthy point is that the critical levels of the generalized Gamma function are higher than those of the log  $t$  function in all cases except in the case of the representative sample. The same holds for the

Table 4  
Parameter estimates of the Champernowne family.

Group	$y_0$	$\alpha$	$\beta$	$\gamma$	$\chi^2$
Wage earners	15.87	2.90	2.00	5.65	12.49
Self-employed	19.80	1.76	1.52	2.84	8.14
Free professions	93.70	2.91	1.40	1.31	1.60
Old-aged	8.59	1.44	1.21	2.19	1.41
Representative group	18.13	2.23	1.39	1.97	14.88

critical levels of the Champernowne function as compared with the log Pearson Four function. Details on the parameter estimates are presented in tables 2, 3, and 4.

First, consider the estimates for the lognormal and the log  $t$  families; see table 2. It appears that the estimates of  $\mu$  are not very sensitive to the change from lognormal to log  $t$ , except in the case of the free professions. The minimum chi-square values indicate that the introduction of log  $t$  results in substantial improvements in fit, in particular for the case of the representative group. Next, consider the estimates of the log Pearson Four family. The parameter  $\mu$  no longer has the interpretation of a median. In addition, it appears that  $\mu$  and  $\lambda$  are highly correlated. This is seen by comparing the unconditional estimates of  $\mu$  in the log Pearson Four and the corresponding conditional estimates given that  $\lambda=0$ , in other words, the estimates of  $\mu$  for the log  $t$ . Because of these complications we do not consider this family a very attractive one, in spite of the improved goodness of fit.

As we already mentioned in section 3, the parameters of the generalized Gamma family have not a clear interpretation. In addition we found that the chi-square surfaces showed elongated valleys.<sup>9</sup> In the case of the representative group, we were unable to determine an unrestricted minimum of the chi-square function. The reported estimates are conditional upon  $p = -0.33$ . Rounding errors started to dominate for smaller values of  $p$ .

The estimates for the parameters  $\gamma$  and  $\alpha$  of the Champernowne family give some indication about income inequality with respect to low and high incomes, respectively. Notice the very high value of  $\gamma$  for the wage earners. This result may be explained by institutional arrangements.

The density functions of several families of income distributions are plotted for the case of the representative sample; see figs. 1 and 2. We used two scales: income in fig. 1 and log income in fig. 2. The scale of log income has the advantage that the symmetry of the lognormal and log  $t$  with respect to their medians is shown. The skewness of the log Pearson Four was rather

<sup>9</sup>Compare footnotes 5 and 6.

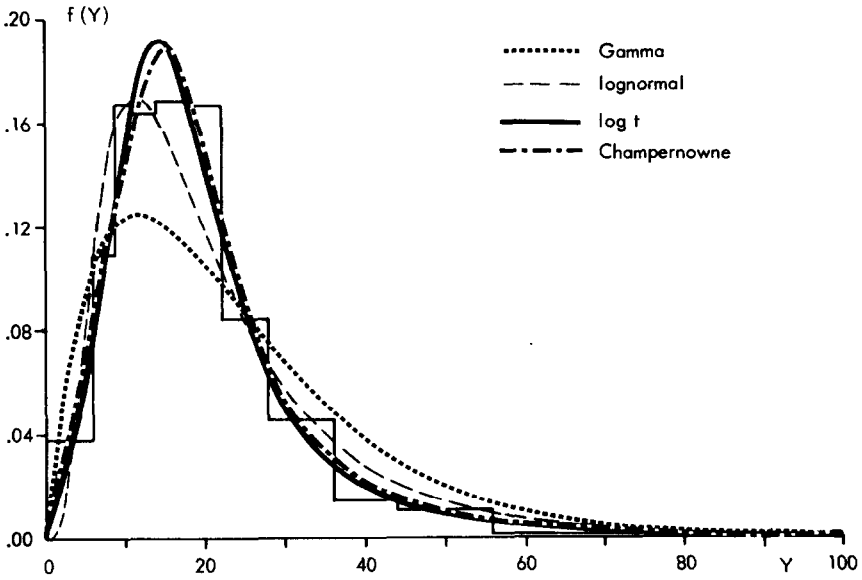


Fig. 1. Some distributions fitted to the representative sample.

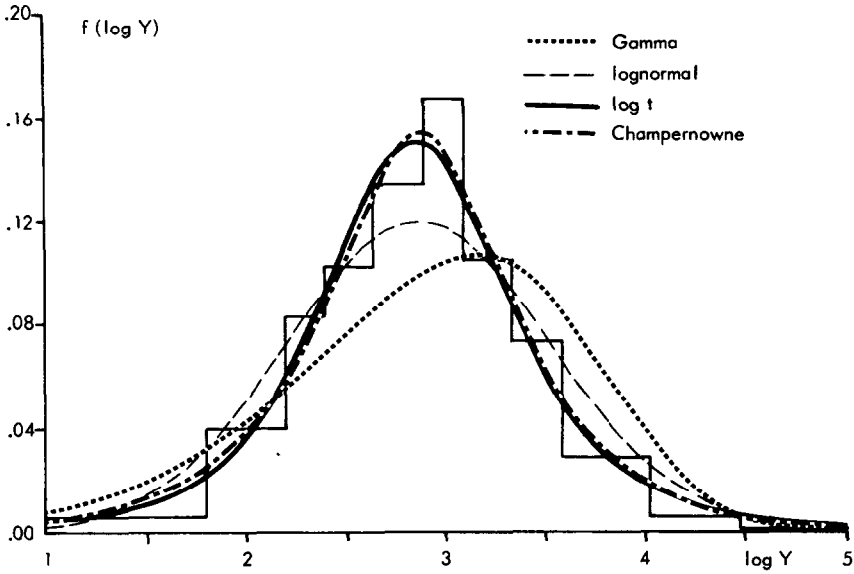


Fig. 2. Some distributions fitted to the representative sample (log income scale).

small in this case. A preliminary plot showed that the log Pearson Four can not clearly be distinguished from the log  $t$ . So we decided to omit it from figs. 1 and 2.

Graphical inspection also indicates that the lognormal and Gamma do not fit very well in the middle range<sup>10</sup> and that a substantial improvement in fit is obtained by using three- and four-parameter families of distributions like the log  $t$  and the Champernowne.

Finally, we turn to the problem of sensitivity of the minimum chi-square values to the class interval groupings. In our case the observed income classes were exogenously determined, but we combined classes with less than ten observations; see the appendix. So the combination decisions were dependent on the sample outcomes. For that reason we tried to obtain information on the sensitivity of the class interval groupings by reestimating the parameters of several families of distributions in a number of cases. The results (not reported here in detail in order to save space) indicate that the chi-square values and their critical levels are sensitive to some extent. However, in all cases but one the ranking of the critical levels did not change. More generally, our main conclusions turned out not to be affected by the way we combined the income classes.

## 5. Conclusions

(i) Optimization methods based on direct search, which use a standard numerical integration routine in each (optimizing) step, are quite feasible on modern computing machinery and as such allow one to apply the minimum chi-square estimation criterion (or the maximum likelihood criterion) to many families of theoretical distribution functions.

(ii) Using the minimum chi-square estimation criterion, we obtained empirical evidence that the lognormal and Gamma families are not flexible enough to describe the size distribution of income accurately. Considerable improvements in fit were obtained by introducing three-parameter families of distributions. The transition from three- to four-parameter families produced some additional improvements but these were much less convincing. The small samples yielded better fits than the large ones. One might conjecture that very large sample sizes will lead to a rejection of (*almost*) any family of distributions if the five percent rule is adopted irrespective of sample size.

(iii) The interpretation of the parameters of the lognormal, the log  $t$  and the Gamma families is straightforward and clear. The interpretation of the parameters of the generalized Gamma, the log Pearson IV, and the Champernowne families gives a number of problems. So there is a certain

<sup>10</sup>This contrasts with the results cited in the introduction. The difference is probably due to the weights implied by the minimum  $\chi^2$  criterion.

trade-off between the two requirements of clear interpretability of the parameters and a good fit to empirical data.

(iv) The generalized Gamma family gave serious estimation problems in a number of cases. In particular, it is hard to obtain reliable estimates of the scale parameter  $a$ . Similar but less serious problems were encountered when estimating the parameters of the log Pearson Four. We give more attention to this problem in Kloek and van Dijk (1977).

### Acknowledgement

The authors are indebted to Professor W.H. Somermeyer of the Econometric Institute of the Erasmus University, Rotterdam for allowing them to use and publish the data, presented in the appendix. We are also indebted to two groups of students, including E.A. Bolhuis, G.J. van 't Eind, G. Gelauff, P. Rietveld, and S. van Wijnbergen, who participated in preparing preliminary discussion papers on this subject. Valuable programming advice was provided by Mr. A.S. Louter of the Econometric Institute. Finally, Mr. M.M. Stefanski of the Erasmus University, Rotterdam assisted in plotting figs. 1 and 2.

### Appendix: Data

Table 5  
Class frequencies of income data (*source*: see section 4).

Interval <sup>a</sup> (guilders × 1000)	Wage earners	Self- employed	Free professions	Old-aged	Representative group
0-3	] 32	] 12	] 20	] 19	] 46
3-6					
6-9	] 68	] 15		41	] 65
9-11				26	
11-14	14	97			
14-18	98	28	12	133	
18-22	116	43	] 17	12	132
22-28	82	38		100	
28-36	57	33	15	73	
36-44	19	22	20	23	
44-56	24	16	27	16	27
56-72	] 13	] 19	25	] 10	
72-88			32		
88-112	] 15	] 17	17	] 14	
112-144			19		
144-2,000					
	509	256	192	128	786

<sup>a</sup>The brackets indicate which income classes were combined in order to have at least ten individuals per interval.

## References

- Abramowitz, Milton and Irene A. Stegun, 1964, *Handbook of mathematical functions* (National Bureau of Standards, Washington, DC).
- Aigner, Dennis J. and Arthur S. Goldberger, 1970, Estimation of Pareto's law from grouped observations, *Journal of the American Statistical Association* 65, 712-723.
- Aitchison, J. and J.A.C. Brown, 1957, *The lognormal distribution* (Cambridge University Press, Cambridge).
- Beveridge, Gordon S.G. and Robert S. Schechter, 1970, *Optimization: Theory and practice* (McGraw-Hill, New York).
- Box, M.J., 1965, A new method of constrained optimization and a comparison with other methods, *Computer Journal* 8, 42-52.
- Box, George E.P. and D.R. Cox, 1964, An analysis of transformations, *Journal of the Royal Statistical Society, Series B*, 26, 211-252.
- Champernowne, D.G., 1937, The theory of income distributions, *Econometrica* 5, 379-381.
- Champernowne, D.G., 1952, The graduation of income distributions, *Econometrica* 20, 591-615.
- Cox, D.R., 1961, Tests of separate families of hypotheses, in: J. Neyman, ed., *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, (University of California Press, Berkeley, CA) 105-123.
- Cramér, Harald, 1946, *Mathematical methods of statistics* (Princeton University Press, Princeton, NJ).
- Cramer, Jan S., 1969, *Empirical econometrics* (North-Holland, Amsterdam).
- Elderton, William P. and Norman L. Johnson, 1969, *Systems of frequency curves* (Cambridge University Press, Cambridge).
- Fisk, Peter R., 1961, The graduation of income distributions, *Econometrica* 29, 171-185.
- Goldfeld, Stephen M. and Richard E. Quandt, 1972, *Nonlinear methods in econometrics* (North-Holland, Amsterdam).
- Himmelblau, David M., 1972, *Applied nonlinear programming* (McGraw-Hill, New York).
- Jansen, Rudolf and Teun Kloek, 1976, *Income differentials: Effects of personal characteristics on individual incomes*, Manuscript, April (Econometric Institute, Erasmus University, Rotterdam).
- Kakwani, N.C. and N. Podder, 1976, Efficient estimation of the Lorenz curve associated inequality measures from grouped observations, *Econometrica* 44, 137-148.
- Kloek, Teun and Herman K. van Dijk, 1977, Further results on efficient estimation of income distribution parameters, *Economie Appliquée, Archives de l'ISMEA, Tome XXX*, 3, 439-459.
- Raiffa, Howard and Robert Schlaifer, 1961, *Applied statistical decision theory* (Graduate School of Business Administration, Harvard University, Cambridge, MA).
- Salem, A.B.Z. and T.D. Mount, 1974, A convenient descriptive model of income distribution: The gamma density, *Econometrica* 42, 1115-1127.
- Somermeyer, William H., 1977, A general market model of labor income distribution: An outline, Report 7708 (Econometric Institute, Erasmus University, Rotterdam).
- Stacey, E.W., 1962, A generalization of the gamma distribution, *Annals of Mathematical Statistics* 33, 1187-1192.
- Wolff, Pieter de, 1941, Income elasticity of demand, A micro-economic and a macro-economic interpretation, *Economic Journal* 51, 140-145.