ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B 00120 HELSINKI 12 tel. 601322

Keskusteluaiheita
Discussion papers

Pentti Vartia - Yrjö Vartia

DESCRIPTION OF THE INCOME DISTRIBUTION BY THE SCALED F DISTRIBUTION MODEL

No 18 18.2.1978

DESCRIPTION OF THE INCOME DISTRIBUTION BY THE SCALED F DISTRIBUTION MODEL

by Pentti L.I. Vartia and Yrjö O. Vartia

A new income distribution model, the scaled F distribution, is suggested. This distribution is a close approximation to the log-normal distribution for low and medium incomes and to the Pareto distribution for high incomes. Estimation of the F distribution is based on moment- and ML-methods. Its statistical properties and connections with some other well-known distributions (e.g. the Beta distribution) make the model easy to apply.

1. Introduction

The theoretical distributions most commonly in describing the distribution of personal income have been the log-normal and the Pareto distribution. Pareto's law applies to the higher incomes only, while the log-normal distribution often gives a good fit for the lower and medium parts of an observed distribution.

There is plenty of both theoretical and empirical evidence in favour of the Pareto and log-normal distributions as the right type of partial approximations to well-behaved income distributions, see e.g. Cramer (1971, p. 38-75), who gives several references. When trying to describe the entire range of incomes by a single distribution we have regarded appropriate to require that the following desiderata are fulfilled:

1. Correct qualitative behaviour: the distribution should approximate Pareto distribution for high incomes and give a better overall fit than the log-normal model

2. Estimation by standard methods: It should be possible to estimate the model using some familiar method proposed by estimation theory and having desirable properties when data is sampled from the distribution.
3. Easy to apply and manipulate: the statistical characteristics of the estimated model (predicted frequencies, measures of central tendency and inequality, etc.) should be easily determined by standard methods without having to use laborous numerical methods.

No simple set of generally approved rules of contest exist in the art of fitting income distributions. This makes it difficult to evaluate various suggestions. Main problem is how much weight should be given to good fit, how much to simplicity or beauty. Davis (1941), Champernowne (1953), Fisk (1961) and Singh and Maddala (1976) seem to stress the goodness of fit, while the log-normal distribution and e.g. the gamma distribution proposed by Salem and Mount (1976) beat these distributions in simplicity and are easy to estimate. By desiderata 1.-3. we intend to give a tentative clarification of the wishes involved.

We have here followed the traditional fitting approach and not required that an income formation process should be invented to rationalize a particular income distribution, cf. Davis (1941, p. 412) and Cramer (1971). If a distribution approximates well other distributions generated by some theoretical processes (e.g. log-normal and Pareto) a lack of a process of its own is a lesser disadvantage. Some authors, e.g. Salem and Mount (1974), also stress the economic interpretation of the original parameters of the distribution. This desideratum does not concern so much the model but a particular representation of it, because the distribution can often be reparametrized using economically meaningful parameters (e.g. its mean, coefficient of variation etc.). However, the parameters of our model have a natural interpretation.

Denoting money income, a random variable, by $\underline{x}$ we propose the following (scaled and shifted) F distribution model

$$(1) \qquad (\underline{x} - \tau) = A\underline{F} , \qquad (\underline{x} - \tau) \geq 0 ,$$

where $\underline{F}$ is distributed according to Fisher's F distribution with parameters m and n, $\underline{F} \sim F(m,n)$. The shift parameter τ is the minimum income for which the model is used. When $\tau = 0$ we have the (unshifted) scaled F distribution model. The shifted model may be preferred if only incomes exceeding some τ are recorded, as in the case, e.g., with taxed income in Finland. Although some truncated distribution might be more appropriate, the shifting of the origo is often a convenient approximation.

The density function of $\underline{F}$ is

$$(2) \qquad f_{\underline{F}}(F) = C_{m,n} \, F^{\frac{m}{2}-1} \, (1+\frac{m}{n}F)^{-\frac{1}{2}(m+n)} , \quad F \geq 0 ,$$

where $C_{m,n} = \Gamma(\frac{m+n}{2}) \, (\frac{m}{n})^{\frac{m}{2}} / \Gamma(\frac{m}{2})\Gamma(\frac{n}{2})$ and $\Gamma(z)$ is Euler's gamma function. We denote the distribution of the scaled variable $A\underline{F}$ by $F(A,m,n)$, the scaled F distribution. Thus the density function of the scaled F distribution is $\frac{d}{dx}P(A\underline{F} \leq x) = f_{A\underline{F}}(x) = f_{\underline{F}}(\frac{x}{A})/A$.

That the money income exceeding τ distributes according to the scaled F distribution can now be expressed by

$$(3) \qquad (\underline{x}-\tau) \sim F(A,m,n) , \quad (\underline{x} - \tau) \geq 0 .$$

The interpretations given here to the parameters of the scaled F distribution $F(A,m,n)$ are similar to those of the formerly used income distribution measures. The shape parameters m and n measure the inequality of income in low and high income groups in a sense similar to the Pareto parameter.[1)] The scale parameter A is given the dimension of the unit of accounting (e.g. mark) and it is closely connected with the geometric mean of the shifted income $(\underline{x}-\tau)$, e.g. $A = \mathrm{Geom}\ (\underline{x}-\tau)$ if $m=n$. With proportionate growth the relative changes in A are always equal to relative changes in (shifted) income.

2. Connections with other distributions

Let us consider first incomes exceeding x_0 or the conditional variable $(\underline{x}|\underline{x}>x_0)$. If $(\underline{x}|\underline{x}>x_0)$ obeys Pareto's law, $(\underline{x}|\underline{x}>x_0) \sim \text{Pareto } (x_0,\alpha)$, then[2)]

$$P\{(\underline{x}|\underline{x}>x_0)>x\} = (x_0/x)^{\alpha} = P\ (\underline{x}>x|\underline{x}>x_0), \text{ when } x>x_0. \tag{4}$$

We say that a distribution 'has a Pareto tail' if its conditional variable $(\underline{x}|\underline{x}>x_0)$ satisfies (4) when $x_0\to\infty$ or, geometrically, its decreasing distribution function approaches a straight line on double logarithmic paper, see figure 4.

---

1) This corresponds to the idea of 'Two-Tailed Pareto Distribution' presented by Champernowne (1953) to approximate not only the upper tail but also the lower tail of the distribution by a straight line on a double-logarithmic scale.

2) The Pareto model may also be generalized for the shifted income variable $(\underline{x}-\tau)$: $(\underline{x}-\tau|\underline{x}-\tau>x_0) \sim \text{Pareto } (x_0,\alpha)$ and $P(\underline{x}>x|\underline{x}>x_0+\tau) = [x_0/(x-\tau)]^{\alpha}$ for $x>x_0+\tau$. Thus we see that the shift parameter τ and the truncation parameter x_0 are connected with each other in no simple way.

The mean of Pareto (x_0,α), or $E(\underline{x}|\underline{x}>x_0)$, exists only when $\alpha>1$ and equals $x_0(\alpha/\alpha-1)$. The conditional variance $D^2(\underline{x}|\underline{x}>x_0)$ exists only when $\alpha>2$, and similarly for higher moments.

It is well-known that the upper tail of the log-normal distribution does not agree with the Pareto distribution but gives a systematic "undershooting". However, the density function $f_{A\underline{F}}(x)$ approaches $C(1/x)^{\frac{n}{2}+1}$, a Pareto density, where $\alpha=n/2$. F(A,m,n) has thus a Pareto tail as required by desideratum 1. The expectation $E(\underline{F})=n/(n-2)$ exists only for $n>2$, in accordance with the Pareto distribution. If $n\to 2$ the mean of the F distribution approaches infinity. Analogously the k:s moment of $\underline{F}$ exists only if $n>2k$. These facts reflect the shewness of empirical income distributions and should not be regarded as a disadvantage as Fisk (1961, p. 172) does.

The scaled F distribution F(A,m,n) is also a close approximation to the log-normal distribution because

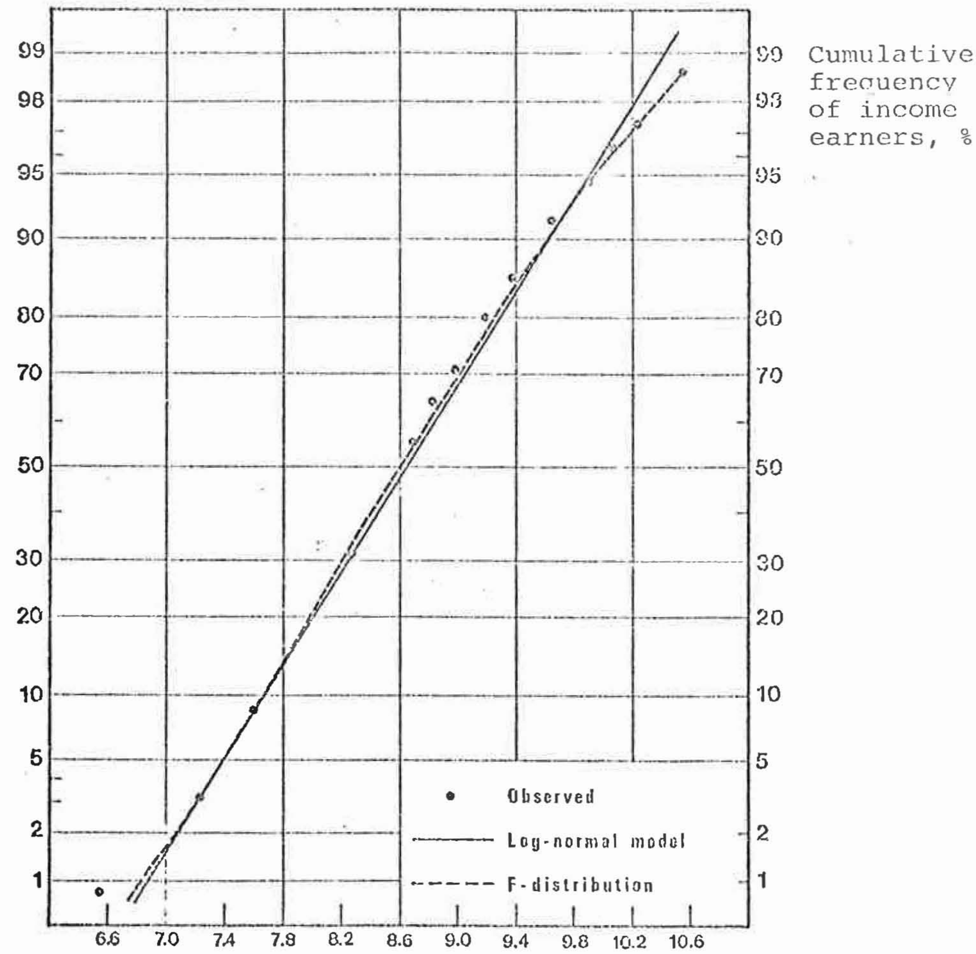
(5) $$\log \underline{x} = \log A + \log \underline{F} = \log A + 2\underline{z} \ ,$$

where $\underline{z}$ is Fisher's z-variable, i.e.,

(6) $$\underline{z} = \frac{1}{2}\log \underline{F} = \frac{1}{2}\log\left[(\underline{x}_m^2/m)/(\underline{x}_n^2/n)\right]$$

the distribution of which is closely normal for moderata m and n. Cumulative distribution functions of $\underline{F}$, plotted on logarithmic probability paper (Figure 1), as presented, e.g., by Hald (1960, p. 377) show the close correspondence between the log-normal and F distributions. From these figures we see that the distribution function of log$A\underline{F}$ on probability paper is S-shaped or its distribution is leptocurtic as required by Rutherford (1955)

Figure 1. Cumulative frequencies of the observed distribution and of the log-normal and F distribution models plotted on logarithmic probability paper, i.e., their "log-normal representations".



the logarithm of shifted income: log(x-2000).

and Lydall (1968, p. 66). Therefore the scaled F distribution seems to fulfil desideratum 1. Our moment method estimation uses the correspondence of $\underline{z}$- and $\underline{F}$-variables.

The flexibility of our F distribution model[1] is reflected by the following results given, e.g., by Hald (1960, p. 384-387):

1. for $n = \infty$, $\underline{F} = \underline{x}_m^2/m$ and $f_{\underline{F}}(F) = C\,F^{\frac{m}{2}-1}e^{-\frac{m}{2}F}$.
   (scaled x^2-variable)
2. for $m = \infty$, $\underline{F} = n/\underline{x}_n^2$ and $f_{\underline{F}}(F) = C\,F^{-(\frac{n}{2}+1)}e^{-\frac{n}{2F}}$.
   (inverse of a scaled x^2-variable)
3. for $m \to \infty$ and $n \to \infty$ the F distribution approaches via normal distributions a singular distribution concentrated on point 1.

Using 1. we get $A x_m^2/m \sim \text{Gamma}(m/2A, m/2) = F(A,m,\infty)$, i.e. the gamma distribution proposed by Salem and Mount (1974) is a special case of the scaled F distribution, when $n=\infty$. Therefore, using any data and any goodness of fit measure some $F(A,m,n)$ fits still better than the best fitting gamma distribution.

---

1) In Vartia and Vartia (1972) we proposed as an income distribution the generalized (scaled) F distribution $F(c,A,m,n)$. This is the distribution of $\underline{x}=A\underline{F}^c$, where A and c are positive parameters and $\underline{F} \sim F(m,n)$. The density function of $F(c,A,m,n)$ is

$$f_{\underline{x}}(x) = \frac{\Gamma(\frac{m+n}{2})\,(\frac{m}{n})^{\frac{m}{2}}}{cA\,\Gamma(\frac{m}{2})\,\Gamma(\frac{n}{2})}\,(\tfrac{x}{A})^{\frac{m}{2c}-1}\,[1+\tfrac{m}{n}(\tfrac{x}{A})^{\frac{1}{c}}]^{-\frac{m+n}{2}}, \quad x>0.$$

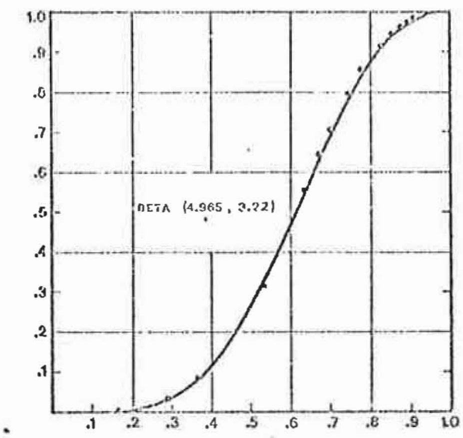
Thus $F(1/\alpha, t_0, 2, 2)$ is Fisk's (1961) sech^2-distribution and $F(1/a_2, (1/a_1a_3)^{1/a_2}, 2, 2a_3)$ is the distribution proposed by Singh and Maddala (1976), both being special cases of the generalized F distribution.

$F(c,A,m,m)$ also approaches log-normal (μ,σ^2), when $m \to \infty$, $A=e^{\mu}$ and $c = \frac{1}{2}\sqrt{m\sigma^2}$. $F(c,A,m,n)$ is a very flexible distribution which satisfies at least our two first desiderata. Estimation of $F(c,A,m,n)$ may be done according to chapter 3.1.

Figure 2. Cumulative frequencies of the observed distribution and the Beta-distributions corresponding to the estimated F distribution models; the unit square representation.

Estimation by moment method

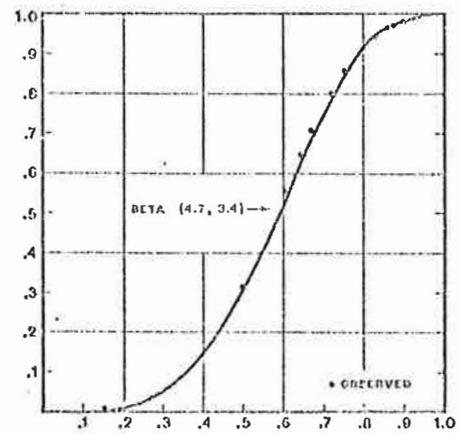
Cumulative frequency of income earners



transformed income λ:
$\lambda = (x-2000)/(1431+x)$

Estimation by ML-method

Cumulative frequency of income earners



transformed income λ:
$\lambda = (x-2000)/(1681+x)$

In the following the ML-estimation of the scaled F distribution model $\underline{x} = A\underline{F}$ is based on the transformation

$$(7) \qquad \underline{\lambda} = \frac{\underline{x}}{\bar{A} + \underline{x}}, \quad \bar{A} = \frac{n}{m} A,$$

which produces a Beta(m/2, n/2) distributed variable $\underline{\lambda}$ in the interval [0,1], see Cramér (1946, p. 242). The distribution function of $\underline{\lambda}$ is the incomplete betafunction $I_x(m/2, n/2)$ tabulated by Pearson (1968), also computer subprograms are often available. The incomplete betafunction is one of the most investigated higher transcendental functions, see e.g. Abramowitz and Stegun (1970). This is a benefit when properties of F(A,m,n) are derived. Note that Thurow (1970) has fitted beta distribution directly to scaled income data.

It is interesting to investigate the distribution functions of the estimated Beta(m/2, n/2) distribution and the corresponding transformed observed distribution in the unit square, as is done in Table 3 and Figure 2. The scaled F-variable $\underline{x}$ on the original income scale in terms of $\underline{\lambda}$ is given by the inverse transformation of (7): $\underline{x} = \bar{A}\underline{\lambda}/(1-\underline{\lambda})$.

3. Estimation of the scaled F distribution

3.1. Moment Method for the log-income

The first three moments of the z-variable, see Fisher (1950),

$$(8) \qquad \mu_1 = E\underline{z} = -\frac{1}{2}(r_1 - r_2) - \frac{1}{6}(r_1^2 - r_2^2) + O(r_1^4 - r_2^4)$$

$$(9) \qquad \mu_2 = D^2(\underline{z}) = \frac{1}{2}(r_1 + r_2) + \frac{1}{2}(r_1^2 + r_2^2) + \frac{1}{3}(r_1^3 + r_2^3) + O(r_1^5 + r_2^5)$$

$$(10) \qquad \mu_3 = E(\underline{z} - E\underline{z})^3 = -\frac{1}{2}(r_1^2 - r_2^2) - (r_1^2 - r_2^3) + O(r_1^4 - r_2^4)$$

(where $r_1 = 1/m$, $r_2 = 1/n$) can be used in estimating the parameters of the scaled F distribution. These equations follow from the moments of $\log(\underline{x}^2/m)$, see Abramowitz and Stegun (1970, p. 943). Since $(1/2)\log\underline{x} - (1/2)\log A$ is approximately normally distributed, the moment methods is here well established. The corresponding moments of the logarithm of income $\log\underline{x} = \log A + 2\underline{z}$ are

(11) $$E(\log\underline{x}) = \log A + 2\mu_1$$

(12) $$D^2(\log\underline{x}) = 4\mu_2$$

(13) $$E(\log\underline{x} - E(\log\underline{x}))^3 = 8\mu_3 \; .$$

Defining the empirical moments of logarithms by

(14) $$\log G = \frac{1}{N}\,\Sigma\log x_i$$

(15) $$s^2 = \frac{1}{N}\,\Sigma(\log x_i - \log G)^2$$

(16) $$m_3 = \frac{1}{N}\,\Sigma(\log x_i - \log G)^3$$

we get the estimating equations

(17) $$\log A = \log G + (r_1 - r_2) + \frac{1}{3}(r_1^2 - r_2^2)$$

(18) $$s^2 = 2(r_1 + r_2) + 2(r_1^2 + r_2^2) + \frac{4}{3}(r_1^3 + r_2^3)$$

(19) $$m_3 = -4(r_1^2 - r_2^2) - 8(r_1^3 - r_2^3) \; .$$

The last two equations determine $r_1 = 1/m$ and $r_2 = 1/n$ as functions of s^2 and m_3 and the first equation can then be used to give the scale parameter A. We have solved the equations (18)-(19) by iteration as follows

(20) $$(r_1+r_2)^{(n+1)} = 2H^{(n)}/(1+\sqrt{1+4H^{(n)}})$$

(21) $$(r_1-r_2)^{(n+1)} = -m_3/[4(r_1+r_2)^{(n+1)}+8(r_1^{(n)}r_1^{(n)}+r_1^{(n)}r_2^{(n)}+r_2^{(n)}r_2^{(n)})]$$

(22) $$r_1^{(n+1)} = \frac{1}{2}[(r_1+r_2)^{(n+1)} + (r_1-r_2)^{(n+1)}]$$

(23) $$r_2^{(n+1)} = \frac{1}{2}[(r_1+r_2)^{(n+1)} - (r_1-r_2)^{(n+1)}]$$

where $H^{(n)} = \frac{1}{2}s^2 + 2r_1^{(n)}r_2^{(n)} - \frac{2}{3}[(r_1^{(n)})^3 + (r_2^{(n)})^3]$ and $r_1^{(0)} = r_2^{(0)} = 0$. Convergence is quite rapid.

3.2. Maximum likelihood method

The close connection of $F(A,m,n)$ with the Beta distribution provides an alternative estimation method. The ML-estimators of the parameters of Beta$(\frac{m}{2},\frac{n}{2})$ are approximated by simple functions in Y. Vartia (1973). These lead to a relatively simple procedure for also estimating the parameters of $F(A,m,n)$.

If $x_1, x_2, \ldots, x_N$ are independent observations from $F(A,m,n)$, the maximum likelihood estimates of m,n and $\bar{A} = \frac{n}{m}A$ are the solutions of the ML-equations

(24) $$\psi(\tfrac{m+n}{2}) - \psi(\tfrac{m}{2}) = \frac{1}{N}\Sigma\log(\frac{1}{\lambda_i}) \hat{=} u$$

(25) $$\psi(\tfrac{m+n}{2}) - \psi(\tfrac{n}{2}) = \frac{1}{N}\Sigma\log(\frac{1}{1-\lambda_i}) \hat{=} v$$

(26) $$\frac{m}{m+n} = \frac{1}{N}\Sigma(\lambda_i) ,$$

where $\lambda_i = x_i/(\bar{A}+x_i) \in [0,1]$ and $\psi(x) = \frac{d}{dx}\log\Gamma(x)$ is the digamma function.

Figure 3. Observed frequency distribution of taxed incomes in 1967 and F distribution models estimated by ML- and moment methods.

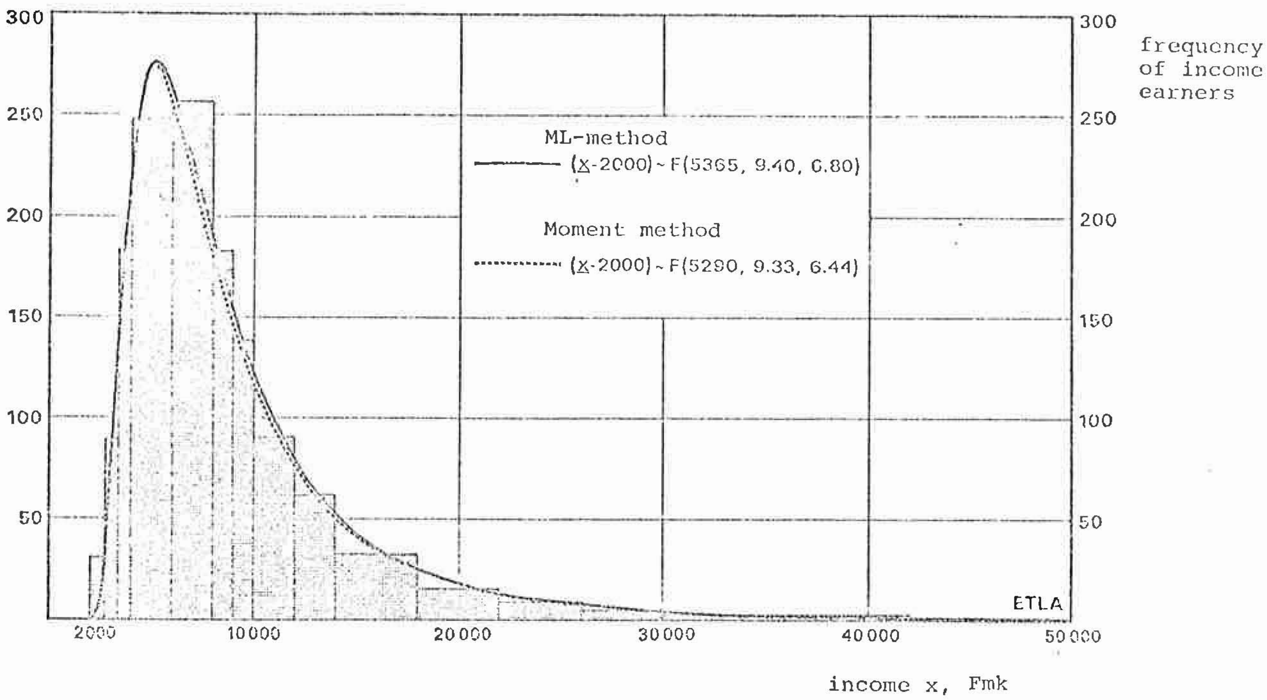
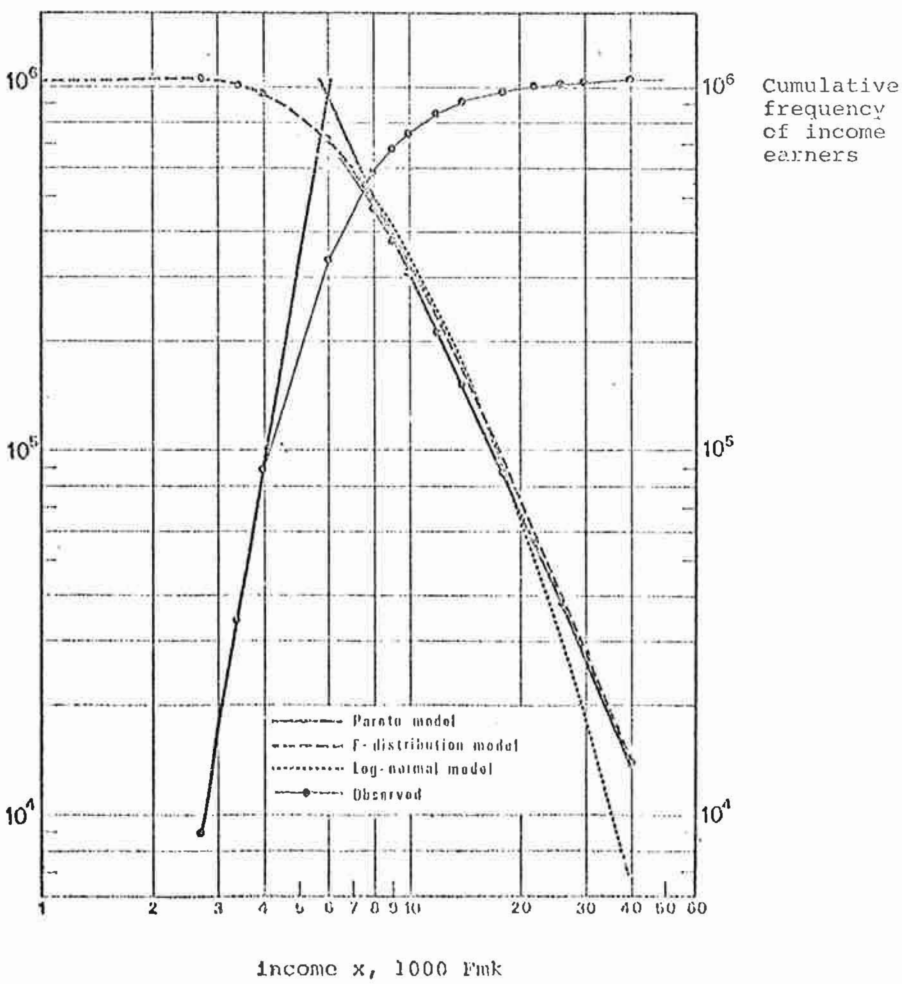


Figure 4. Cumulative frequencies of the observed distribution and of the Pareto, log-normal and F distribution models plotted on double logarithmic paper, i.e., their "Pareto representations".



Here equations (25)-(26) correspond to the ML-equations for the Beta distribution and $\underline{\lambda}$ is a Beta $(\frac{m}{2},\frac{n}{2})$ variable, cf. (7). These cannot be solved explicitly, but are approximated accurately by simple functions given in Vartia and Vartia (1972), which also includes an iterative procedure for solving (25)-(26).

3.3. An empirical illustration

As an illustration we have used the distribution of taxed income in Finland in 1967, see Table 1. For institutional reasons we have chosen in all our examples τ=2000 mk as the value of the shift parameter[1]. We ignore here complications due to problems of estimation with grouped data (cf. Salem and Mount (1974)) and the fact that the figures in Table 1 are based on a stratified sample containing about 200 000 observations. The density functions of the estimated F distribution models together with the empirical frequency distribution are represented in Figure 3. The corresponding cumulative frequencies are given in Table 2.

The moment method estimation gave $\hat{\Lambda}$=5290 marks, $\hat{m}$=9.93 and $\hat{n}$=6.44. This estimated F distribution model is also represented in Figures 1 and 4, in order to compare the fit with those obtained by Pareto and log-normal models[2]. In both representations

---

1) No state tax was levied on incomes less than 2300 marks. Corresponding income earners are not registered in the Finnish taxed income statistics.

2) The log-normal models $\log(\underline{x}-\tau) \sim N(\mu,\sigma^2)$ for τ=0 and τ=2000 were estimated using the ML-method. The Pareto-model in its unshifted form was estimated for the 213780 observations exceeding 12000 mk by ML-method. This Pareto$_2$(12000,2.215) gave an excellent fit, see figure 4, and χ^2=151.2 only (the 0.1 % critical χ^2-value is 20.5 for df=5). The Pareto distribution has fitted excellently to the same data at least from the beginning of 1950's, see Vartia and Vartia (1973).

the F distribution model seems to fit systematically better than the shifted log-normal model $\log(\underline{x}-2000) \sim N(8.632, 0.586)$, as was to be expected. The ML-estimation gave slightly different parameter estimates, i.e. $\hat{A}=5365$ marks, $\hat{m}=9.4$ and $\hat{n}=6.8$, but the density functions approximate each other accurately as shown in Figure 3. The unit square representations for the estimated F distribution models are given in figure 2.

Here, as often in very large samples, deviations between the predicted and observed frequencies clearly cannot result from sampling fluctuations only, though the fit in descriptive sense is very good[1]. E.g. the familiar x^2-measure $x^2=\Sigma(o_i-e_i)^2/e_i= n\Sigma(\hat{p}_i-p_i)^2/\hat{p}_i$ (where n is the total number of observations, $\hat{p}_i=o_i/n$ and $p_i=e_i/n$) is large because it increases with n, although the squared deviations $(\hat{p}_i-p_i)^2$ are small. Here we have an additional problem of the "right" number of observations. However, x^2 remains too large[2] for all reasonable choices of n. For curiosity we report the x^2-values (calculated using n=1 063 065) and the sum of squared deviations between the predicted and observed probabilities used e.g. by Singh and Maddala (1976).

---

1) Ijiri and Simon (1977) give an excellent introduction to these problems. We cite from p. 4: "...our theories are always only approximate theories that do not capture all the fine structure of the phenomena. Hence, with sufficiently large samples of sufficiently good data, the deviations of data from theory almost always reveal themselves. However, we cannot conclude from this that the theory should be rejected; the only valid conclusion to be drawn is that the theory is only a first approximation - hardly surprising - and that the next step in the investigation is to look for an additional mechanism that could be incorporated in the theory so as to lead to a better second approximation. It would be foolish for us to give up the gas laws for ideal gases simply because most gases are not, in fact, ideal; or to give up the law of acceleration in a vacuum because most of the bodies we observe are falling through air. Hence, we shall not be much concerned, in what follows, with significance tests, which are completely inappropriate for testing the fit of data to extreme models. Instead, we will be concerned with how much of the variance in the raw data is explained by the models, and with how sensitive the fit is to changes in assumptions."

2) The 0.1 % critical x^2-value is less than 33 for relevant degrees of freedom df=11 or 12.

	Model	$x^2 = n\Sigma \frac{(\hat{p}_i - p_i)^2}{\hat{p}_i}$	$SD = \Sigma(\hat{p}_i - p_i)^2$
1	$\log \underline{x} \sim N(8.991, 0.3127)$	160 585	0.0246
2	$\log(\underline{x}-2000) \sim N(8.632, 0.586)$	34 975	0.00169
3	$(\underline{x}-2000) \sim F(5290, 9.93, 6.44)$	17 574	0.00133
4	$(\underline{x}-2000) \sim F(5365, 9.4, 6.8)$	22 238	0.00132

Although the SD:s of models 3 and 4 are almost equal, their x^2-values differ because of slightly different tail probabilities (cf. table 2). Different goodness of fit measures[1)] may thus rank the models in different order. Comparing the goodness of fit of good approximate but strictly speaking misspecified models is a difficult philosophical problem, which we must leave aside in this connection, see Ijiri and Simon (1977, p. 109-134).

4. Conclusions

We have tried to demonstrate analytically that the old, well-known F distribution is a relevant alternative in describing empirical income (or similar) distributions. The scaled F distribution approximates the log-normal and Pareto distributions well and has e.g. the gamma distribution as a special case. Furthermore, it has a natural generalization which includes e.g. the distribution proposed by Singh and Maddala and thus Fishk's sech^2-distribution as its special cases. The F distribution models are at least as flexible as the competing models they include.

1) E.g. Rao (1965, p. 288) gives 4 alternatives for the familiar x^2-measure.

As large empirical income data seldom can be interpreted as a random sample from any given theoretical distribution having only a few parameters, we have paid attention not only to good fit and correct qualitative behaviour but also to the general applicability of the model: that it may be estimated by standard methods and is easy to understand and manipulate.

Table 1. Numbers of persons in various taxed income groups in Finland in 1967

Income group (Fmk)	Classmark[1)] (Fmk)	Frequency of income earners
- 2 699	2 520	8 890
2 700 - 3 399	3 033	25 387
3 400 - 3 999	3 691	55 100
4 000 - 5 999	5 089	247 493
6 000 - 7 999	6 940	256 335
8 000 - 8 999	8 469	91 427
9 000 - 9 999	9 474	69 571
10 000 - 11 999	10 925	95 082
12 000 - 13 999	12 921	61 844
14 000 - 17 999	15 702	64 341
18 000 - 21 999	19 805	30 820
22 000 - 25 999	23 809	18 111
26 000 - 29 999	27 852	10 825
30 000 - 39 999	34 209	13 897
40 000 -	63 506	13 942

1) The arithmetic mean income in the income group. Our estimation procedures used only the classmarks and the corresponding frequencies, whereas no continuity correction was applied here.

Table 2. Comparision of the observed and estimated cumulative frequencies of the F distribution model

Upper limit of the income group (Fmk)	Cumulative frequencies		
	Observed	ML-method	Moment method
2 700	8 890	3 614	3 721
3 400	34 277	35 294	34 018
4 000	89 377	90 254	88 979
6 000	336 870	352 406	354 213
8 000	593 205	581 071	580 008
9 000	684 632	674 727	663 671
10 000	754 203	743 082	732 664
12 000	849 285	840 140	830 785
14 000	911 129	906 157	892 975
18 000	975 470	989 607	964 519
22 000	1006 290	1018 204	1004 490
26 000	1024 401	1034 150	1022 031
30 000	1035 226	1042 335	1035 319
40 000	1049 123	1053 816	1048 714
∞	1063 065	1063 065	1063 065

Table 3. Cumulative relative frequencies of the observed distribution and the Beta distribution corresponding to the F distribution model $(\underset{\sim}{x}-2000) \sim F(5365, 9.4, 6.8)$ as estimated by the ML-method; the unit square representation.

Upper class limit of shifted income	Upper class limit of transformed income	Cumulative relative frequency of Beta (4.7,3.4)	Observed cumulative relative frequency
$z_i = x_i - 2000$	$\lambda_i = \frac{z_i}{3881+z_i}$	$P(\underset{\sim}{\lambda} \le \lambda_i)$	$\sum_{x_k \le x_i} f_k / \sum_{k=n}^{K} f_k$
0	.0000	.0000	.0000
700	.1528	.0034	.0084
1 400	.2651	.0332	.0322
2 000	.3401	.0849	.0841
4 000	.5075	.3315	.3169
6 000	.6072	.5466	.5580
7 000	.6433	.6347	.6440
8 000	.6733	.6990	.7095
10 000	.7204	.7903	.7989
12 000	.7556	.8524	.8571
16 000	.8048	.9309	.9176
20 000	.8375	.9578	.9466
24 000	.8608	.9728	.9636
28 000	.8783	.9805	.9738
38 000	.9073	.9913	.9869
∞	1.0000	1.0000	1.0000

REFERENCES

Abramowitz, M. and Stegun, I.A. (1970): Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Government Printing Office, USA.

Champernowne, D.G. (1953): A Model of Income Distribution, Economic Journal, 63, 318-351.

Cramér, H. (1946): Mathematical Methods of Statistics, Princeton University Press, Princeton.

Cramer, J.S. (1971): Empirical Econometrics, North Holland Publishing Company, Amsterdam. London.

Davis, H.T. (1941): The Analysis of Economic Time Series, The Principia Press Inc., Bloomington, Cowles Commission Monograph No. 6.

Fisher, R.A. (1950): Contributions to Mathematical Statistics (Moments and Cumulants in the Spesification of Distributions, 30,.3), Wiley, New York.

Fisk, P.R. (1961): The Graduation of Income Distributions, Econometrica, 29, 171-185.

Ijiri, Y. and Simon, H.A. (1977): Skew Distributions and the sizes of Business Firms, North-Holland Publishing Company, Amsterdam, New York, London.

Hald, A. (1960): Statistical Theory with Engineering Applications, 4th printing, Wiley, New York.

Lydall, H. (1968): The Structure of Earnings, The Clarendon Press, Oxford.

Pearson, K. (1968): Tables of the Incomplete Beta Function, Cambridge University Press, Cambridge.

Rao, C.R. (1965): Linear Statistical Interence and its Applications, John Wiley & Sons, Inc. New York.

Rutherford, R.S.G. (1955): Income distributions; a new model, Econometrica, 23, 277-94.

Salem, A.Z.B. and Mount, T.D. (1974): A Convenient Descriptive Model of Income Distribution: The Gamma Density, Econometrica, 42, 1115-1127.

Singh, S.K. and Maddala, G.S. (1976): A Function for Size Distribution of Incomes, Econometrica, 46, 963-970.

Thurow, L.C. (1970): Analyzing the American Income Distribution, American Economic Review, 60, 261-269.

Vartia, P.L.I. and Vartia, Y.O. (1972): F-distribution as a Model for Income Distribution, Research Reports of the Institute of Economics, University of Helsinki, No. 21.

Vartia, P.L.I. and Vartia, Y.O. (1973): Comments on the discussion held at the conference on income distribution, Yearbook of Finnish Society for Economic Research 1972, Helsinki.

Vartia, Y.O. (1973): Estimation of the Parameters of Beta Distribution by Maximum Likelihood Method (in Finnish), Research Institute of the Finnish Economy A:1.