THEOREM 2.4

If $\mathbf{X}$ *is multivariate lognormal and* $\mathbf{b}$ *is a (column) vector of constants with transpose* $\mathbf{b}'$, *then the product* $c\prod_{j=1}^{n} X_j^{b_j}$ *is* $\Lambda(a+\mathbf{b}'\boldsymbol{\mu}, \mathbf{b}'\mathbf{V}\mathbf{b})$, *where* $c=e^a$ *is a positive constant.*

To Cramér's theorem on the normal distribution [45] there corresponds the following:

THEOREM 2.5

If X_1 *and* X_2 *are two independent positive variates such that their product* X_1X_2 *is a* Λ*-variate, then both* X_1 *and* X_2 *are* Λ*-variates (or, as a special case, one of the variates may be constant and the other a* Λ*-variate).*

This is a converse of the reproductive property of Theorem 2.2 and may be extended to the case of a finite number of independent positive variates, but not to an infinite sequence as is evident from a consideration of § 2.6. Levy's corollary [137] to Cramér's theorem may also be reframed to apply to Λ-distributions.

2.5. Moment Distributions: Gini's Coefficient of Mean Difference

The property now to be discussed has no analogue in normal theory since it involves the concept of *moment distributions* which may be defined meaningfully for positive variates only. This concept will be found important in many practical applications and will be discussed at greater length in Chapters 11 and 12. The jth moment distribution function of $\Lambda(\mu, \sigma^2)$ is defined by

$$\Lambda_j(x \mid \mu, \sigma^2) = \frac{1}{\lambda_j'}\int_0^x u^j \, d\Lambda(u \mid \mu, \sigma^2), \tag{2.15}$$

and the fundamental theorem of the moment distributions is

THEOREM 2.6

The j-th moment distribution of a Λ*-distribution with parameters* μ *and* σ^2 *is also a* Λ*-distribution with parameters* $\mu+j\sigma^2$ *and* σ^2 *respectively.*

Proof

$$\begin{aligned}
\Lambda_j(x \mid \mu, \sigma^2) &= \frac{1}{\lambda_j'}\int_0^x u^j \, d\Lambda(u \mid \mu, \sigma^2) \\
&= e^{-j\mu-\frac{1}{2}j^2\sigma^2}\int_0^x e^{j\log u}\frac{1}{u\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2\sigma^2}(\log u-\mu)^2\right\}du \\
&= \int_0^x \frac{1}{u\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2\sigma^2}(\log u-\mu-j\sigma^2)^2\right\}du \\
&= \Lambda(x \mid \mu+j\sigma^2, \sigma^2),
\end{aligned}$$

using (2.5) and (2.6).

This simple result allows us to obtain an explicit expression for Gini's [89] coefficient of mean difference:

Britain in 1935 and 1948 was found to be approximately 0·04, or small in relation to σ^2 (about 0·5). The conclusion of this section, then, is that there exist a number of models of generation which lend plausibility to the assumption of lognormality as a simple description of income distributions.

11.4. Statistical Analysis of Data

The second criterion concerns ease of handling in statistical analysis. It will be convenient to remark briefly under three headings:

(i) The estimation of parameters: on this we need only say that there is a wide choice of methods for the lognormal distribution from which the statistician may choose rationally according to his need for speed and accuracy.

(ii) The comparison of two or more distributions, and more general applications of the analysis of variance: here the link that the lognormal distribution provides with normal theory is of great value and brings to the statistician the full facilities of existing normal test statistics. These properties of the distribution are discussed more fully in Chapter 8.

(iii) The introduction of the distribution of incomes into econometric models: it is often necessary in this field to investigate the consequences of averaging behaviouristic relationships over the distribution of incomes. Here the lognormal hypothesis seems to have considerable advantages over most other candidates. This point is taken up more fully in Chapter 12.

11.5. Interpretation of the Parameters of the Lognormal Distribution

Thirdly, there is the interpretation of the parameters of a lognormal distribution of incomes. The interpretation of the location parameter μ is straightforward, since (in the two-parameter case) it is the logarithm of the geometric mean income and is also the logarithm of the median income. It is to be noted that since the arithmetic mean $\alpha = e^{\mu+\frac{1}{2}\sigma^2}$ involves both the location and dispersion parameters it is not a pure measure of the level of incomes under the lognormal hypothesis: for this the geometric mean or median is to be preferred. The dispersion parameter σ^2 is of greater interest by virtue of its relation to the concept of *concentration of incomes* as defined by Lorenz[141].

In the Lorenz diagram (Fig. 11.1) the proportion of income receivers having income less than x is measured along the horizontal scale and the proportion of total income accruing to the same income receivers along the vertical scale. The points plotted for the various values of x trace out a curve below the 45° line sloping upwards to the right from the origin. In statistical terms the curve describes the relation between the distribution function $F(x)$ and the first-moment distribution function $F_1(x)$, defined by

$$F_1(x) = \int_0^x t\,dF(t) \Big/ \int_0^\infty t\,dF(t). \tag{11.4}$$

The measure of income concentration which is naturally suggested by the Lorenz diagram is the ratio of the shaded area between the Lorenz curve and the 45° line to the area of the triangle under the 45° line. The measure varies from zero, when all persons have the same income (so that the 45° line may be termed the diagonal of equal distribution), to unity, when all the available income accrues to one person.

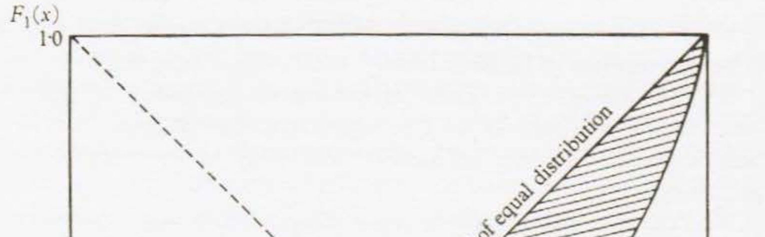

Fig. 11.1. The Lorenz diagram for a two-parameter distribution of incomes. This curve with index $L=0{\cdot}383$ is derived from $\Lambda(\mu, 0{\cdot}5)$.

The formal definition of the measure is

$$L = 1 - 2\int_0^\infty F_1(x)\,dF(x). \tag{11.5}$$

Substituting in equation (11.5) the explicit form for $F_1(x)$ given by Theorem 2.6 we obtain, for the lognormal hypothesis,

$$\begin{aligned} L &= 1 - 2\int_0^\infty \Lambda(x \mid \mu+\sigma^2, \sigma^2)\,d\Lambda(x \mid \mu, \sigma^2) \\ &= 1 - 2\Lambda(1 \mid \sigma^2, 2\sigma^2) \\ &= 1 - 2N\left(\frac{-\sigma}{\sqrt{2}}\,\middle|\, 0, 1\right) \\ &= 2N\left(\frac{\sigma}{\sqrt{2}}\,\middle|\, 0, 1\right) - 1; \end{aligned} \tag{11.6}$$

which shows that the measure of concentration L is monotonically related to the value of σ^2 and is independent of μ.† It will also be noted, from Theorem 2.7, that there is a strong similarity between equation (11.6) and the expression for Gini's coefficient of mean difference. In fact, denoting Gini's coefficient by G, we have in general that

$$G = 2\alpha L, \tag{11.7}$$

where α, as before, is the arithmetic mean income.

It follows that the parameter σ^2 may be interpreted as a measure of the concentration of incomes in a sense which is generally acceptable; and that since the value of σ^2 may be estimated from samples within calculable confidence limits so too can Lorenz's measure of concentration.

Since many empirical data have been described and analysed by means of the Lorenz diagram it is of some interest to discuss the shape of the Lorenz curve resulting from a lognormal distribution.

First, the two-parameter case. The diagonal line drawn at right angles to the diagonal of equal distribution, and defined by the equation

$$F(x) = 1 - F_1(x),$$

cuts the Lorenz curve in this case at the point $\{F(\alpha), 1 - F_1(\alpha)\}$ corresponding to the arithmetic mean income. For

$$\begin{aligned} 1 - \Lambda_1(\alpha) &= 1 - N\left(\mu + \frac{\sigma^2}{2} \middle| \mu + \sigma^2, \sigma^2\right) \\ &= 1 - N\left(-\frac{\sigma}{2} \middle| 0, 1\right) \\ &= N\left(\frac{\sigma}{2} \middle| 0, 1\right) \\ &= \Lambda(\alpha), \quad \text{from equation (5.59);} \end{aligned} \tag{11.8}$$

or, in words, the proportion of persons with less than the mean income is the complement of the proportion of income held by these persons.‡ It also follows from the symmetry properties of the normal distribution that the Lorenz curves in this case (for all values of σ^2) are symmetrical with respect to the diagonal defined above; and that, at the points defined by (11.8), the tangents to the curves are parallel to the diagonal of equal distribution. Also no two curves of the family can intersect. These properties furnish simple tests of the two-parameter hypothesis from the appearance of the Lorenz curves.

Lorenz curves can, however, intersect in two cases of interest. First, if the data plotted on the diagram arise from a two-parameter parent distribution, but these data are available only in truncated form, as is

† Values of L tabulated against σ are given in Appendix Table A1.

‡ This proportion is tabulated against σ in Appendix Table A1.