AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

MONASH

W.P. no. 9/89

MONASH

UNIVERSITY

AUSTRALIA

UNIVARIATE AND MULTIVARIATE BURR DISTRIBUTIONS: A SURVEY

T. R. L. Fry

GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS
LIBRARY
WITHDRAWN
DEC 04 1989

Working Paper No. 9/89

October 1989

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 960 9

UNIVARIATE AND MULTIVARIATE BURR DISTRIBUTIONS: A SURVEY

T. R. L. Fry

Working Paper No. 9/89

October 1989

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS AND POLITICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

UNIVARIATE AND MULTIVARIATE BURR DISTRIBUTIONS: A SURVEY.

Version *2.1*

September *1989*

T.R.L. FRY[1]

Department of Econometrics
Monash University
Clayton, Victoria *3168*
Australia.

Abstract: The genesis of the work on this family of distributions was the paper by Burr 1942, in which he aimed to generate distributions that could take on a wide variety of shapes and yet remain tractable to work with. Subsequent work with the family has been sporadic and has concentrated on the univariate case. It is also spread over a wide range of disciplines. Here we survey the literature concerning the Burr family of distributions and summarise and extend the results concerning two members of the family, the Burr types II and XII. We also develop some new univariate and multivariate Burr type II distributions, which being a generalisation of the Logistic, may prove useful in future applications.

Keywords: Burr distributions, multivariate, univariate.

---

[1] Part of this work was supported by ESRC studentship *A00428424097*. I wish to acknowledge the useful comments made on earlier versions of this paper by Richard Blundell, Nicola Crichton, Len Gill, Grant Hillier, Max King, Richard Smith and participants at seminars at Exeter, Manchester and Monash universities. All remaining errors are my responsibility.

1. Introduction

The genesis of the work on this family of distributions was the paper by Burr, 1942, in which he was interested in generating distributions that could take on a wide variety of shapes and yet remain tractable to work with. Subsequent work with this family has been sporadic and has concentrated on the univariate case. It is also spread over a wide range of disciplines. The purpose of this paper is to pull together many of the strands of work with these distributions and to propose some new forms of Burr distributions.

The Burr family are defined by solutions to a differential equation in the cumulative distribution function. One particular solution, which gives rise to the Burr Type XII (B12) distribution, has been applied in a variety of areas. Indeed the B12 distributions are often simply called Burr distributions in the literature. Burr introduced the family to fit histograms, but since the B12 is a particularly flexible distribution the applications have proved to be much wider. Applications may be found in the areas of quality control, duration or failure time modelling, income distribution modelling, bioassay and hypothesis testing.

This paper surveys the properties and applications of the B12 and Burr Type II (B2) families of distributions in both the univariate and multivariate cases. In contrast to B12 distributions, relatively little attention has been paid in the literature to B2

distributions except in the area of binary choice models. The B2 distribution is of considerable interest in these models because it is a one parameter generalisation of the Logistic distribution. It is a flexible distribution and has wide moment coverage. The B2 distribution may well be useful in many other univariate and multivariate applications.

Useful summaries of the properties of the B12 distribution in the univariate case may be found in Rodriguez 1977, 1982, Tadikamalla 1980b and Vodã 1982. Comparatively few papers have been written about multivariate versions of any Burr distributions. They include work by Takahasi 1965, Durling 1969, 1975 and Johnson 1987 on the B12 distribution and by Rodriguez 1980, and Rodriguez and Taniguchi 1980 on the bivariate B3 distribution. This paper reviews and extends this literature by presenting a number of new results on B2 and multivariate Burr distributions, which may prove useful in further applications.

Section 2 of this paper summarises the properties of univariate B12 and B2 distributions, considers some of the links between the Burr and other well known distributions, and proposes a new generalised B2 distribution, which turns out to coincide with the generalised-F distribution. In section 3 we summarise some of the diverse applications of the B12 distribution and mention the one area where the B2 has found some application, namely binary choice modelling. The properties of multivariate B12 and, in particular, multivariate B2 distributions are found in section 4. Again new,

more flexible, variants on the B2 form are proposed with a view to future applications. The section concludes by investigating the link between multivariate B12, B2 and other distributions. We then consider the few applications of multivariate B12 and B2 distributions in section 5, and finally section 6 includes some concluding remarks.

2. Univariate Distributions

Burr 1942 described a number of forms of cumulative distribution functions which might prove to be useful for fitting data. The purpose of using cumulative distribution functions was to facilitate mathematical analysis whilst attaining a reasonable fit to the data using a method of moments. The system of distributions was introduced by considering distribution functions satisfying the following differential equation:

$$dF = F(1 - F)g(x)dx,$$

where $0 \le F \le 1$ and $g(x)$ is a suitable function, non-negative over the domain of x. The solution to this differential equation, for given $g(x)$, is obtained as:

$$F(x) = \frac{1}{(1 + e^{-G(x)})},$$

where

$$G(x) = \int_{-\infty}^{x} g(u)du .$$

Burr gave 12 solutions to this differential equation (corresponding to choices of $g(x)$) which we list below:

(I) $F(x) = x$ $\qquad 0 < x < 1$

(II) $F(x) = (1 + e^{-x})^{-k}$

(III) $F(x) = (1 + x^c)^{-k}$ $\quad x > 0$

(IV) $F(x) = \left[1 + \left(\frac{c-x}{x}\right)^{1/c}\right]^{-k}$ $\quad 0 < x < c$

(V) $F(x) = (1 + ce^{-\tan(x)})^{-k}$ $\quad -\frac{\pi}{2} < x < \frac{\pi}{2}$

(VI) $F(x) = (1 + e^{-k\sinh(x)})^{-k}$

(VII) $F(x) = 2^{-k}(1 + \tanh(x))^k$

(VIII) $F(x) = \left[\frac{2}{\pi} \tan^{-1}\left(e^x\right)\right]^k$

(IX) $F(x) = 1 - \dfrac{2}{2 + c[(1 + e^x)^k - 1]}$

(X) $F(x) = \left(1 - e^{-x^2}\right)^k$ $\quad x > 0$

(XI) $F(x) = \left[x - \frac{1}{2\pi} \sin(2\pi x)\right]^k$ $\quad 0 < x < 1$

(XII) $F(x) = 1 - (1 + x^c)^{-k}$ $\quad x > 0$

where c, k are positive parameters and $-\infty < x < \infty$ unless otherwise stated. Putting $x = (z - \mu)/\sigma$ would introduce scale and location shift parameters, if required. We have chosen not to include them. The twelve distributions defined by these distribution functions may conveniently be called the Burr Type I - XII distributions.

The Burr Type XII Distribution

If X is distributed as Burr Type XII with parameters c and k we will use the notation $X \sim B12$. For this distribution we have:

$$F(x) = \begin{cases} 1 - (1 + x^c)^{-k} & x > 0 \\ 0 & x \le 0 \end{cases} \qquad c,k > 0,$$

$$f(x) = \frac{kcx^{c-1}}{(1 + x^c)^{k+1}} \qquad x > 0.$$

This density function is unimodal with mode at:

$$x = \left[\frac{c - 1}{(kc + 1)} \right]^{1/c}$$

if $c > 1$ and L shaped if $c \leq 1$. Its moments are given by:

$$E(X^r) = kB(r/c + 1 \ , \ k - r/c)$$

where $B(m, k)$ is the Beta function. Notice that there is an existence condition for the moments of X, that $ck > r$.

It can be shown by means of a standardised third and fourth moment $(\sqrt{\beta_1}, \beta_2)$ coverage diagram that the B12 distribution covers a wide range in the $\sqrt{\beta_1}, \beta_2$ plane (see Rodriguez 1977, Tadikamalla 1980b - both these papers have an incorrect limit value for β_2 in their figures (it should be 5.4, see Sugiura and Gomi 1985)).

As mentioned above the B12 distribution may be fitted by the method of moments in which values of c and k are matched to the sample mean, variance, skewness $\sqrt{\beta_1}$ and kurtosis β_2 measures to produce the best fit for the data (Burr 1973 gives a table of values for this purpose). We can also note the result that the upper and lower tail area integrals for this density are Gauss Hypergeometric functions (see Abramowitz and Stegun 1965).

Dubey 1968 has provided an alternative derivation of the B12 distribution. He assumes that, conditional on θ, the random variable X follows a Weibull process with scale parameter θ and that θ follows a Gamma distribution. The resultant unconditional density is of the B12 form.

The Burr Type II Distribution

If X is distributed as Burr Type II with parameter k we will use the notation X ~ B2. For this distribution we have:

$$F(z) = (1 + e^{-z})^{-k} \qquad -\infty < z < \infty \qquad k > 0,$$

$$f(z) = \frac{ke^{-z}}{(1 + e^{-z})^{k+1}} .$$

Notice that this distribution is a generalisation of the Logistic collapsing to that of the Logistic when k = 1. Indeed it is sometimes termed the Generalised Logistic distribution (see *inter alia* Dubey 1969, Zelterman 1987). The introduction of this extra parameter allows for a variety of shapes for the pdf, as shown in Figures 1 and 2.

The Moment Generating Function (mgf) for this distribution is given by:

$$M(t) = kB(1 - t, k + t),$$

from which we obtain:

$$\mu_1 = \mu = \psi(k) - \psi(1)$$

$$\mu_2 = \sigma^2 = \psi'(k) + \psi'(1)$$

$$\sqrt{\beta_1} = \frac{(\psi''(k) - \psi''(1))}{(\psi'(k) + \psi'(1))^{3/2}}$$

$$\beta_2 = \frac{(\psi'''(k) + \psi'''(1))}{(\psi'(k) + \psi'(1))^2} + 3$$

(details of the polygamma functions ($\psi^{(r)}$) above may be found in Abramowitz and Stegun 1965), which enables us to draw the moment coverage diagram (Figure 3). This shows that the B2 distribution has a wide range of third and fourth moment coverage as traced out

by the curve from (-2, 9), as $k \to 0$, to (1.14, 5.4), as $k \to \infty$, passing through (0, 4.2) corresponding to the Logistic when $k = 1$. Thus the B2 distribution generalises the coverage of the Logistic distribution. As with the B12 distribution the upper and lower tail area integrals for this distribution are found to be Gauss Hypergeometric functions.

The B2 distribution may also be derived through a mixing argument. Assume that conditionally on θ, the random variable Z follows a Log Weibull (or Extreme Value) process and that θ follows a Gamma distribution then the resultant unconditional density is of the B2 form. Such a mixing argument allows the use of the following formula to compute the moments of the distribution:

$$E(Z^r) = E_\theta E_Z(Z^r|\theta).$$

Finally note that Dubey 1969 uses a Generalised Extreme Value distribution mixed with a Gamma distribution to yield a slightly more general B2 distribution than that derived here.

Links With Other Distributions

It is useful to state some of the relationships between particular Burr distributions (i.e., Type XII, Type III and Type II) and some other well known distributions (see Fry 1988 for details). Tadikamalla 1980b investigated the relationships between Burr and related distributions, concentrating his attention on the Type III and Type XII distributions. His work may be viewed as a complement to the list given here.

It is not surprising that since the B2, B3 and B12

distributions derive from particular solutions to the same differential equation that there is an inter-relationship between them. Namely that if $X \sim B12$ then $Y = X^{-1} \sim B3$ and $Z = c\log Y \sim B2$.

The B12 distribution is related through variable transformations to a wide range of well known distributions. If $X \sim B12$ then:

i) $U = (1 + X^c)^{-1}$ has a Beta Type 1 distribution with parameters k and 1.

ii) $V = X^c$ has a Beta Type 2 (Pearson Type VI) distribution with parameters k and 1.

iii) $W = kX^c$ has an F distribution with (2, 2k) degrees of freedom.

Analogously, we may link the B2 distribution to a range of non-Burr system distributions. If $Z \sim B2$ then:

i) $S = (1 + e^{-Z})^{-1}$ has a Beta Type 1 distribution with parameters k and 1.

ii) $T = e^{-Z}$ has a Beta Type 2 distribution with parameters k and 1.

iii) $R = ke^{-Z}$ has an F distribution with (2, 2k) degrees of freedom.

iv) $Q = \log(1 + e^{-Z})$ has an Exponential distribution with mean 1/k.

Another family of distributions to which this version of the B2 distribution is related is the one parameter exponential family. For a distribution to be a one parameter exponential family

distribution we must be able to write it as:

$$f(z) = a(k)b(z)e^{[c(k)d(z)]},$$

where a(k), c(k) are functions of the parameter k alone and b(z), d(z) are functions of z alone. It is straightforward to verify that this may be done for the B2 distribution with the following choice of functions:

$$a(k) = k \; ; \; b(z) = e^{-z} \; ; \; c(k) = -(k + 1) \; ; \; d(z) = \log(1 + e^{-z}).$$

Note however, that it would appear that two or three parameter B2 distributions cannot be written as members of the exponential family of distributions.

A Generalised Burr Type II

We have seen that the B2 distribution can be related to a Beta Type 1 distribution with parameters k and 1. An obvious way to generalise the B2 distribution would therefore be to relate a Beta Type 1 with parameters k and m to the B2 distribution via the appropriate transformation.

Let S ~ Beta Type 1 with parameters k and m. If we consider the (Logistic) transformation:

$$Z = \log\left(\frac{S}{1 - S} \right),$$

we have:

$$f(z) = \frac{1}{B(k, m)} \frac{e^{-mz}}{(1 + e^{-z})^{k+m}} \qquad k,m > 0, \quad -\infty < z < \infty.$$

We may regard this as a Generalised Burr Type II (GB2) distribution with parameters k and m.

It is straightforward to find the mgf of the GB2 distribution as:

$$M(t) = \frac{1}{B(k,\ m)} B(m - t,\ k + t),$$

from which we obtain the following moments:

$$\mu = \psi(k) - \psi(m)$$

$$\sigma^2 = \psi'(k) + \psi'(m)$$

$$\sqrt{\beta_1} = \frac{(\psi''(k) - \psi''(m))}{(\psi'(k) + \psi'(m))^{3/2}}$$

$$\beta_2 = 3 + \frac{(\psi'''(k) + \psi'''(m))}{(\psi'(k) + \psi'(m))^2} .$$

Thus the GB2 distibution extends the moment coverage properties of the B2 distribution and is a more flexible distribution.

To derive the cumulative distribution function of the GB2 distribution we use f(z) defined above and make the obvious change of variable in the integration to yield:

$$F(b) = \frac{1}{B(k,\ m)} \quad B_{1-c}(k,\ m)$$

in which B_{1-c} is the incomplete Beta function and where $1 - c = (1 + e^{-b})^{-1}$. Since the incomplete beta function can also be written as a Gauss Hypergeometric function we have an alternative form for the distribution function:

$$F(b) = \frac{1}{mB(k,\ m)(1 + e^{-b})^m} \quad {}_2F_1(m,\ 1-k;\ m+1;\ (1 + e^{-b})^{-1}).$$

Prentice 1975, 1976 has suggested the use of the pdf given above and points out that the density is symmetric if m = k, negatively skewed for m < k and positively skewed for m > k. Limiting distributions are Normal ($m,k \to \infty$), Double Exponential ($m,k \to 0$),

Exponential ($m \neq 0$, $k \to 0$), and Reflected Exponential ($m \to 0$, $k \neq 0$). The Logistic distribution coincides with $m = k = 1$.

3. Univariate Applications

Psuedo Random Sampling

The B12 distribution covers a wide range of distributions in terms of third and fourth moments. In particular when $c = 4.873717$ and $k = 6.157568$ we have $\sqrt{\beta_1} = 0$ and $\beta_2 = 3$ which coincide with the Normal distribution (these values of c and k yield $\mu = 0.644717$ and $\sigma = 0.16199$). The fact that an appropriate choice of parameter values yields an approximation to the normal distribution led Burr 1967a to suggest the use of the B12 distribution to simulate random sampling from a Normal distribution.

The use of the B12 distribution in simulation is not confined to the Normal distribution. The B12 (and B3) distribution has the advantage of having simple closed forms for both the distribution function and its inverse. This fact allows random samples from the distribution to be easily obtained by the inverse transformation method. Tadikamalla 1977 used a four parameter distribution (including scale and location parameters) as an approximation to the Gamma distribution and also suggested its use to simulate any non-normal distribution (see Tadikamalla 1980a).

Quality Control

An attractive property of the B12 distribution is its ability to fit data which is non-normal i.e $\sqrt{\beta_1} \neq 0$ and/or $\beta_2 \neq 3$. Zimmer and

Burr 1963, and Burr 1967b used this fact to develop sampling plans for variables when the population from which the sample is drawn may not be normal. They were concerned with sampling plans which base a decision about a lot (or process) on the proportion of pieces beyond a single specification limit, and deal with the cases of σ known and σ unknown. For the former $\overline{X} \pm \xi\sigma$ is used in the criterion for acceptance and in the latter $\overline{X} + \xi s$, where ξ is a constant to be determined.

They presented tables which enable the choice of appropriate values of ξ and n (the sample size) for a variety of non-normal distributions if the degree of skewness and/or kurtosis in the population is known or can be estimated/approximated. It was found that, whilst non-normality has some effect on ξ, it is more strongly felt in the sample size n.

Normality Test

Bera 1982 used the B12 distribution to construct tests for normality of observations and regression disturbances. Interest was centered only upon the shape of the distribution. The proposed test of normality is achieved by testing:

$$H_0 : c = 4.873717 \quad \text{and} \quad k = 6.157568$$

after transforming the data into a set of positive observations with $\overline{X} = 0.644714$ and $s = 0.16199$. No details of the proposed transformation are given, although inability to make such a transformation is itself deemed an indication of non-normality. It is also claimed that the proposed test is likely to have high

power to detect non-normality for symmetric distributions.

The paper discusses Lagrange multiplier (LM) and likelihood ratio (LR) tests. When testing for the normality of observations, it is found that no estimation is required to calculate the LM test statistic. When testing regression disturbances, the LM test is again found to be easily applied while the LRT is much more difficult to apply.

Duration Models

Lomax 1954 analysed data on business failure assuming it is reasonable to expect monotonically decreasing conditional probabilities of failure for business. In other words, the longer a business survives, other things being equal, the smaller becomes the probability of failure. This implies a particular form for the failure rate (or hazard) function which corresponds to a B12 distribution with $c = 1$ and scale parameter $1/a$. Dubey 1966b showed that Lomax's data can be better fitted by using another form of the scaled B12 distribution given by $k = 1$ and c free to vary.

Wingo 1983 investigated the use of the B12 distribution for life test data in biometric applications. He considered the maximum likelihood estimation of the complete sample case (i.e. no censoring) of n independent observations, $x_i > 0$, $i = 1,\ldots,n$, with a B12 probability density. The paper gives expressions for the asymptotic variances and covariances of the parameters and for both joint and separate parameter confidence intervals and applies

the methodology to life test data arising in a clinical setting.

The mixing approach to deriving the B12 distribution given above has been used in deriving duration models in the social sciences (see for example Morrison and Schmittlein 1980, Lancaster 1979, 1985). If we assume that each individuals' duration follows a Weibull distribution and let the Weibull scale parameter θ be distributed according to some distribution, $g(\theta)$, across the population of individuals, then we may view the mixing distribution, $g(\theta)$, as capturing the heterogeneity of failure rates in the population. When this mixing distribution is a Gamma distribution, then the resultant density is the B12.

Lancaster 1979, 1985 and Brännäs 1986 also mixed Weibull durations with a Gamma distribution for a random heterogeneity component, thus producing B12 distributions. Their work includes explanatory variables in the model while that of Morrison and Schmittleins does not. Furthermore, the rationale given for the existence of heterogeneity in these papers is as a consequence of omitted variables.

A Model for the Size Distribution of Income

In the literature on this subject, the B12 distribution is often termed the Singh-Maddala distribution after its introduction by Singh and Maddala 1976. They derived the B12 distribution both from an argument concerning failure rates and one concerning decay rates. They also fitted the B12 distribution to U.S data using non-linear least squares and found that it provides a good fit

when compared with the Pareto and Log-normal distributions.

Schmittlein 1983 extended the Singh-Maddala analysis by deriving the large sample properties of the 3 parameter B12 distribution. In particular, he derived the asymptotic variance-covariance matrix of the parameters for both the complete sample case and for the more usual situation, in income data, of grouped, censored observations. His results are therefore generalisations of those of Wingo 1983. He presented the implied values of four common inequality measures (the Gini Index, variance of log-transformed income, Theil's Entropy and the Pietra Ratio) when income follows the B12 distribution and the asymptotic variances for the first three measures are also derived. The paper concludes with a re-analysis of the U.S data used by Singh and Maddala using MLE and a brief look at the impact of the types of grouping commonly used in income data.

Further evidence on the usefulness of the B12 distribution in this area comes from McDonald 1984 who found that it provides a better fit to income data than one of his generalised Beta distributions and was only bettered, in his comparisons, by his generalised Beta of the second kind. Furthermore he concluded that the closed form of the B12 greatly facilitates estimation and analysis of results.

Binary Choice Model

The application of the B2 distribution in a model of binary choice has been discussed in Poirier 1980, Fry 1988 and Smith 1989 and such a model may conveniently be termed a Burrit model.

Poirer was concerned with developing the Burrit model as the alternative to the Logit model in a LM test for skewness in binary choice models. Fry developed the Burrit model within a latent variable framework and discussed both its estimation and a LRT for symmetry. The Burrit model has also been used by Smith 1989 in the context of developing LM tests, computed as nR^2 from an auxiliary least squares regression, for distributional mispecification.

Finally we note that Seibert 1970, Drane, Owen and Seibert 1978 suggested the B12 distribution as a candidate distribution for quantal choice bio-assay problems.

4. Multivariate Distributions

We now investigate some multivariate versions of the B12 and B2 distributions. We begin with the bivariate Burr distributions since these provide an insight into the properties of multivariate distributions. Bivariate distributions are often generated by transformations of given univariate marginal distributions. We consider a variety of possible transformations for both the B12 and the B2 distributions.

Bivariate Burr XII Distributions

We have already seen that the B12 distribution may be derived using a mixing approach. Takahasi 1965 extended this approach to generate a multivariate B12 distribution (TB12). Here we consider a bivariate version of the Takahasi-Burr distribution which omits scale and location shift parameters.

Assume that X_1 and X_2 have independent Weibull distributions with common parameter θ and that this parameter θ follows a Gamma distribution. The joint pdf for X_1 and X_2 is found to be:

$$f(x_1, x_2) = \frac{\Gamma(k+2)c_1c_2x_1^{c_1-1}x_2^{c_2-1}}{\Gamma(k)\left[1 + x_1^{c_1} + x_2^{c_2}\right]^{k+2}} \qquad x_1, x_2, c_1, c_2, k > 0.$$

The marginal pdfs of X_1 and X_2 may be obtained in the usual manner. We find that these are univariate B12 distributions with parameters c_1, c_2 and k. The conditional densities for this distribution are found to be scaled B12 distributions.

The joint moments are given by:

$$E\left(X_1^{r_1}, X_2^{r_2}\right) = \frac{\Gamma(r_2/c_2 + 1)\Gamma(r_1/c_1 + 1)\Gamma(k - r_2/c_2 - r_1/c_1)}{\Gamma(k)}$$

with existence conditions of:

$$r_i/c_i + 1 > 0 \ , \ i = 1,2 \ ; \ k + 1 - r_2/c_2 > 0 \ ;$$
$$k - r_1/c_1 - r_2/c_2 > 0,$$

and the conditional moments are:

$$E\left(X_1^{r_1} \mid X_2 = x_2\right) = \frac{\left[1 + x_2^{c_2}\right]^{r_1/c_1}\Gamma(r_1/c_1 + 1)\Gamma(k + 1 - r_1/c_1)}{\Gamma(k+1)}$$

with an analogous result for the moments of X_2 conditional on values of X_1.

If we were to hold c_1, c_2, and k constant (e.g. to approximate the

bivariate normal) then the correlation coefficient is restricted to a point defined by a function of the c_1, c_2, and k.

The Durling-Burr Distribution.

Durling 1969, 1975 introduced a generalisation of the Takahasi-Burr bivariate distribution where the correlation coefficient ρ would not be restricted to a point if the c_1, c_2. and k parameters were held constant. Comparison of the Takahasi-Burr bivariate distribution function and that of the product of two independent Burr distributions suggested the use of the following joint cdf:

$$F(x_1, x_2) = 1 - (1 + x_1^{c_1})^{-k} - (1 + x_2^{c_2})^{-k}$$

$$+ (1 + x_1^{c_1} + x_2^{c_2} + \alpha x_1^{c_1} x_2^{c_2})^{-k}$$

$$x_1, x_2 \geq 0 \; ; \quad c_1, c_2, k > 0 \; ; \; 0 \leq \alpha \leq k+1$$

$$= 0 \text{ otherwise.}$$

This cdf reduces to that of the Takahasi-Burr distribution if $\alpha = 0$ and to the product of two independent Burr distributions if $\alpha = 1$. The joint density function is obtained by differentiation and the marginal distributions are the same as those for the Takahasi-Burr distribution.

Durling 1969, 1975 gives the conditional density as:

$$f(X_i = x_i | x_j) = \frac{(k+1)(1 + \alpha x_i^{c_i})A_j c_i x_i^{c_i - 1}}{\left[1 + A_j x_i^{c_i}\right]^{k+2}}$$

$$- \frac{\alpha c_i x_i^{c_i - 1}}{\left[1 + A_j x_i^{c_i}\right]^{k+1}} \qquad i \neq j = 1,2$$

where $A_j = (1 + \alpha x_j^{c_j})/(1 + x_j^{c_j})$. From which we find:

$$E(X_1^{r_1} | x_2) = (k + 1)A_2^{-r_1/c_1} B(1 + r_1/c_1, k + 1 - r_1/c_1)$$

$$+ \frac{\alpha r_1}{c_1} A_2^{-(r_1/c_1 + 1)} B(1 + r_1/c_1, k - r_1/c_1),$$

and

$$E(X_1^{r_1}, X_1^{r_2}) = k^2 \alpha^{-r_1/c_1} B(1 + r_1/c_1, k - r_1/c_1)$$

$$B(1 + r_2/c_2, k - r_2/c_2)$$

$${}_2F_1(r_1/c_1, k - r_1/c_1; k; 1 - \alpha^{-1}).$$

Thus the correlation coefficient ρ is a complicated function of c_1, c_2, k and α and its range is difficult to determine. It is, however, more general than that of the Takahasi-Burr distribution.

The Durling-Burr distribution as introduced lacked a formal justification, but Hutchinson 1981 provided a derivation of a version of the Durling-Burr distribution which shows it may be derived within a mixing framework.

Morgenstern Burr XII Distribution.

Another way to generate a bivariate B12 distribution is to use the system attributed to Morgenstern in Johnson and Kotz 1972. The

joint cdf is given by:

$$F(x_1, x_2) = F_1(x_1)F_2(x_2)[1 + \alpha\{1 - F_1(x_1)\}\{1 - F_2(x_2)\}],$$

where $-1 \le \alpha \le 1$ is a 'mixing' parameter and independence of X_1 and X_2 corresponds to $\alpha = 1$.

One drawback with this family of bivariate distributions is that it can be shown that for *any* such bivariate distributions with finite variances, the correlation coefficient is bounded in absolute value by 1/3 (see Schucany, Parr and Boyer 1978).

If we assume that the marginal distributions are both univariate B12 distributions then we find that:

$$F(x_1, x_2) = \{1 - (1 + x_1^{c_1})^{-k_1}\}\{1 - (1 + x_2^{c_2})^{-k_2}\}.$$

$$[1 + \alpha(1 + x_1^{c_1})^{-k_1}(1 + x_1^{c_2})^{-k_2}] \qquad -1 \le \alpha \le 1.$$

We find the joint moments to be:

$$\begin{aligned} E(X_1^{r_1}, X_2^{r_2}) = {} & (\alpha+1)k_1k_2B(1+r_1/c_1, k_1-r_1/c_1)B(1+r_2/c_2, k_2-r_2/c_2) \\ & - 2\alpha k_1k_2[B(1+r_1/c_1, 2k_1-r_1/c_1)B(1+r_2/c_2, k_2-r_2/c_2) \\ & + B(1+r_1/c_1, k_1-r_1/c_1)B(1+r_2/c_2, 2k_2-r_2/c_2)] \\ & + 4\alpha k_1k_2B(1+r_1/c_1, 2k_1-r_1/c_1)B(1+r_2/c_2, 2k_2-r_2/c_2) \end{aligned}$$

Bivariate Burr II Distributions

In the univariate case, the mixing of a Log-Weibull (or Extreme Value) distribution with a Gamma distribution yields a B2 distribution. In this section, we use the Takahasi approach to

obtain a bivariate B2 (TB2) distribution.

Assume that Z_1 and Z_2 have, conditionally on a common scale parameter θ, independent Log-Weibull distributions and that this parameter θ ($\theta \geq 0$) follows a Gamma distribution. The joint pdf is given by:

$$f(z_1, z_2) = \frac{\Gamma(k+2)e^{-z_1}e^{-z_2}}{\Gamma(k)\left[1 + e^{-z_1} + e^{-z_2}\right]^{k+2}} \qquad \begin{array}{l} -\infty < z_i < \infty \\ k > 0 \; ; \; i = 1,2. \end{array}$$

This joint pdf could also have been obtained by making the appropriate change of variables in the TB12 distribution given above. For this distribution it is easy to verify that the marginal densities are univariate B2 and that the conditional densities are 'scaled' B2 distributions.

Given the mathematical form of the joint pdf of the TB2 distribution it is natural to consider the joint moment generating function (mgf):

$$M(t_1, t_2) = \frac{\Gamma(1 - t_1)\Gamma(1 - t_2)\Gamma(k + t_1 + t_2)}{\Gamma(k)}$$

from which joint moments may be obtained. For example,

$$E(Z_1, Z_2) = \left.\frac{\partial^2 M(t_1 \; t_2)}{\partial t_1 \partial t_2}\right|_{t_1=t_2=0}$$

$$= \psi'(k) + [\psi(k) - \psi(1)]^2.$$

In finding the conditional moments, it is again natural to consider the appropriate mgf:

$$M(t_1 | Z_2 = z_2) = \left[1 + e^{-z_2}\right]^{-t_1} \frac{\Gamma(1-t_1)\Gamma(k+1+t_1)}{\Gamma(k+1)}$$

This distribution was considered by Satterthwaite and Hutchinson 1978 who derived the regression of Z_2 on Z_1 (and of Z_1 on Z_2), which is non-linear and homoscedastic.

The correlation between Z_1 and Z_2 is found to be:

$$\rho = \frac{\psi'(k)}{\psi'(k) + \psi'(1)}$$

$$= \frac{\psi'(k)}{\psi'(k) + \pi^2/6},$$

which varies between 0 and 1 and tends to zero as k tends to infinity. Now if k = 1 we have a bivariate Logistic distribution and $\rho = 1/2$. This is also the formula given in Johnson and Kotz 1972 as an approximation to the correlation between X_i and X_j in a multivariate Burr XII distribution. In addition, as Sattertwaite and Hutchinson 1978 point out, since the regression curves are non-linear, the product moment correlation coefficient may not be the most appropriate measure of association. Their paper suggests other measures.

Morgenstern Bivariate Burr II

We have already seen how we may generate such a bivariate distribution for the B12 distribution. Here we assume that the marginal distributions are univariate B2 distributions and hence

produce a bivariate B2 distribution with cdf:

$$F(z_1, z_2) = \frac{1}{\left[1 + e^{-z_1}\right]^{k_1}\left[1 + e^{-z_2}\right]^{k_2}} \left[(\alpha + 1) \right.$$

$$+ \frac{4\alpha}{\left[1 + e^{-z_1}\right]^{k_1}\left[1 + e^{-z_2}\right]^{k_2}} - \frac{2\alpha}{\left[1 + e^{-z_1}\right]^{k_1}}$$

$$\left. - \frac{2\alpha}{\left[1 + e^{-z_2}\right]^{k_2}} \right] \qquad \begin{array}{l} -\infty < z_i < \infty \ ; \ k_i > 0 \ , \\ i = 1,2 \ ; \ -1 \le \alpha \le 1 \ . \end{array}$$

If we consider finding the joint mgf for this distribution. The resultant expression, after some simplification is:

$$M(t_1, t_2) = (\alpha + 1) \frac{\Gamma(1-t_1)\Gamma(k_1+t_1)}{\Gamma(k_1)} \frac{\Gamma(1-t_2)\Gamma(k_2+t_2)}{\Gamma(k_2)}$$

$$+ \alpha \frac{\Gamma(1-t_1)\Gamma(2k_1+t_1)}{\Gamma(2k_1)} \frac{\Gamma(1-t_2)\Gamma(2k_2+t_2)}{\Gamma(2k_2)}$$

$$- \alpha \frac{\Gamma(1-t_1)\Gamma(2k_1+t_1)}{\Gamma(2k_1)} \frac{\Gamma(1-t_2)\Gamma(k_2+t_2)}{\Gamma(k_2)}$$

$$- \alpha \frac{\Gamma(1-t_1)\Gamma(k_1+t_1)}{\Gamma(k_1)} \frac{\Gamma(1-t_1)\Gamma(2k_2+t_2)}{\Gamma(2k_2)} \ .$$

'Triangular' Bivariate Burr II

Another way in which a bivariate B2 distribution can be generated is to take the product of two independent variables with univariate B2 distributions and a 'triangular' linear transformation. If Z_i ~ independent B2, then their joint pdf is given by the product of the marginal pdfs. Now if we consider a 'triangular' linear transformation from the Z_i's to new variables

W_i (i = 1,2), defined by:

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{Az},$$

then the joint density of the W_i's is:

$$f(w_1\ w_2) = \frac{a^{11}a^{22}k_1k_2e^{-(a^{11}+a^{21})w_1}e^{-a^{22}w_2}}{\left[1 + e^{-a^{11}w_1}\right]^{k_1+1}\left[1 + e^{-a^{21}w_1}e^{-a^{22}w_2}\right]^{k_2+1}}$$

$$-\infty < w_1,\ w_2 < \infty\ ;\ k_1,\ k_2 > 0\ .$$

It is possible to find a closed form expression for the marginal density of W_1 by integrating out W_2 (the result is a scaled B2 density). However, the integrals defining both the marginal density of W_2 and the joint cdf do not have closed form solutions. Fry 1988, therefore, used an alternative approach to find the relevant moments:

$$\mu_{11} = a_{11}a_{21}(\psi'(k_1) + \psi'(1)) = \mathrm{Cov}(W_1\ W_2)$$

$$\mu_{10} = a_{11}(\psi(k_1) - \psi(1)) = E(W_1)$$

$$\mu_{20} = a_{11}^2(\psi'(k_1) + \psi'(1)) = \mathrm{Var}(W_1)$$

$$\mu_{01} = a_{21}(\psi(k_1) - \psi(1)) + a_{22}(\psi(k_2) - \psi(1)) = E(W_2)$$

$$\mu_{02} = a_{21}^2(\psi'(k_1) + \psi'(1)) + a_{22}^2(\psi'(k_2) + \psi'(1)) = \mathrm{Var}(W_2).$$

Thus the effect of the linear transformation is to 'free up' the correlation coefficient ρ such that $-1 \le \rho \le 1$.

Multivariate Burr XII Distributions

It is possible to construct other multivariate 'Burr' distributions. For example, the construction of an m-variate Durling-Burr distribution, requiring either an m-variate Gamma

mixing distribution or a generalisation of Durling's approach to the m-variate situation, could be attempted but would not be straightforward. Since attention in the literature has concentrated upon it, we choose only to consider the Takahasi-Burr distribution (see Takahasi 1965, Johnson and Kotz 1972). Johnson 1987 includes many examples of the generation of versions of the Takahasi-Burr distribution and may therefore be viewed as a complement to the distributions given here.

If we assume that $X_1,\ldots,X_m$ have, conditional upon a common scale parameter θ, independent Weibull distributions and that the parameter θ follows a Gamma distribution we find:

$$f(x_1,\ldots,x_m) = \frac{\Gamma(k+m)\prod c_j x_j^{c_j-1}}{\Gamma(k)\left[1 + \sum x_j^{c_j}\right]^{k+m}}$$

$$k > 0 \; ; \; c_j > 0 \; ; \; x_j > 0 \quad j = 1,\ldots,m$$

Takahasi 1965 calls this distribution "multivariate Burr's distribution" (henceforth TB12) and gives the following two theorems with outline proofs, which are elaborated in Fry 1988.

<u>Theorem 1:</u> Any marginal distribution of the TB12 distribution is also (multivariate) TB12.

<u>Theorem 2:</u> Any conditional distribution of the TB12 distribution is also (multivariate) TB12.

Conditional moments are given in Johnson and Kotz 1972 (p290) as:

$$E(X_1^{r_1} \mid x_2,\ldots,x_m) = \left[1 + \sum_{j=2}^{m} x_j^{c_j}\right]^{r_1/c_1} \frac{\Gamma(k+m)}{\Gamma(k+m-1)}$$

$$B(1+r_1/c_1,\ k+m-1-r_1/c_1).$$

The joint moments may be derived as:

$$E(X_1^{r_1}\ldots X_m^{r_m}) = \frac{1}{\Gamma(k)} \prod_{j=1}^{m} \Gamma(1 + r_j/c_j)\Gamma(k - \sum_j r_j/c_j)$$

with the following moment existence conditions

$$1 + r_j/c_j > 0 \ ;\ k > \sum_j r_j/c_j \qquad j = 1,\ldots,m.$$

Since the marginal distributions are univariate B12 the covariance between any two variables X_i and X_j $i \neq j = 1,\ldots,m$ is found to be:

$$\mathrm{Cov}(X_i,X_j) = \frac{\Gamma(1 + 1/c_i)\Gamma(1 + 1/c_j)\Gamma(k - 1/c_i - 1/c_j)}{\Gamma(k)}$$

$$- k^2 B(1 + 1/c_i,\ k - 1/c_i)B(1 + 1/c_j,\ k - 1/c_j)$$

Thus the correlation coefficient is a function of c_i, c_j and k and its range is difficult to determine.

Multivariate Burr II Distributions

The extension of the argument used in generating the bivariate TB2 distribution to generate a multivariate distribution is straightforward. That is, we assume that $Z_1,\ldots,Z_m$ have, conditional upon a common scale parameter θ, independent Log-Weibull (Extreme Value) distributions and that the scale parameter θ has a Gamma distribution. The usual mixing arguments

will then yield:

$$f(z_1,\ldots,z_m) = \frac{\Gamma(k+m)\prod e^{-z_j}}{\Gamma(k)\left[1 + \sum e^{-z_j}\right]^{k+m}}$$

$$k > 0 \; ; \; -\infty < z_j < \infty \; , \; j = 1,\ldots,m.$$

We may call this distribution the multivariate Takahasi-Burr II distribution (TB2). It can also be shown (Fry 1988) that the two theorems relating to the TB12 distribution stated above also apply to the TB2 distribution.

As in the bivariate case, when finding moments, it is natural to consider the joint moment generating function:

$$M(\mathbf{t}) = M(t_1,\ldots,t_m) = \frac{\Gamma(k + \Sigma\, t_j)}{\Gamma(k)} \prod_{j=1}^{m} \Gamma(1 - t_j),$$

with existence conditions:

$$k + \sum_j t_j > 0 \quad ; \quad 1 - t_j > 0 \quad \forall\, j = 1,\ldots,m.$$

The correlation between Z_i and Z_j is given by:

$$\rho_{ij} = \rho = \frac{\psi'(k)}{\psi'(k) + \psi'(1)} \qquad \forall\, i \neq j = 1,\ldots,m.$$

Therefore, the TB2 distribution has an equi-correlated structure. The correlation coefficient ρ is that encountered in the bivariate TB2 distribution. For many applications this correlational structure may be restrictive and we would like to generalise it further to the 'free' structure found in the multivariate Normal distribution. It is useful to write the joint pdf in vector

notation as:

$$f(\mathbf{z}) = \frac{\Gamma(k+m)e^{-\mathbf{i}'\mathbf{z}}}{\Gamma(k)\left[1 + \mathbf{i}'ev(-\mathbf{z})\right]^{k+m}},$$

where $\mathbf{z} = (z_1 \ \ldots \ z_m)'$, $ev(-\mathbf{z}) = \left[e^{-z_1}\ldots e^{-z_m}\right]$, and $\mathbf{i}$ is an $m \times 1$ vector of ones. Therefore:

$$E(\mathbf{z}) = \underline{\mu} = [\psi(k) - \psi(1)]\mathbf{i}$$

$$Var(\mathbf{z}) = \Sigma = \psi'(1)\mathbf{I}_m + \psi'(k)\mathbf{i}\mathbf{i}' .$$

This vector notation enables us to consider the distribution of a linear transformation $\mathbf{w} = \mathbf{A}\mathbf{z}$ (where $\mathbf{A}$ may or may not be triangular). We find the pdf of $\mathbf{w}$ to be:

$$f(\mathbf{w}) = \frac{\Gamma(k+m)|\mathbf{A}^{-1}|e^{-\mathbf{i}'\mathbf{A}^{-1}\mathbf{w}}}{\Gamma(k)\left[1 + \mathbf{i}'ev(-\mathbf{A}^{-1}\mathbf{w})\right]^{k+m}},$$

with

$$E(\mathbf{w}) = \mathbf{A}E(\mathbf{z}) = \mathbf{A}\underline{\mu} = [\psi(k) - \psi(1)]\mathbf{A}\mathbf{i}$$

$$Var(\mathbf{w}) = \mathbf{A}\Sigma\mathbf{A}' = \psi'(1)\mathbf{A}\mathbf{A}' + \psi'(k)\mathbf{A}\mathbf{i}\mathbf{i}'\mathbf{A}'.$$

This generalises the distribution in that its correlation structure will no longer be equi-correlated but will be a fully free structure.

Links With Other Distributions

Cook and Johnson 1981 proposed a general family of distributions for modelling multivariate data, which includes the TB12 and TB2 distributions as special cases. Using the fact that multivariate distributions are usually classified by their associated marginals, they propose a 'standard form' multivariate distribution with uniform (on (0,1]) marginals from which other distributions may be obtained by the appropriate transformation.

If $\mathbf{U} = (U_1 \ldots U_m)'$ follow a Cook-Johnson distribution then:

$$F(u_1,\ldots,u_m;\alpha) = \left[\sum_{j=1}^{m} u_j^{-1/\alpha} - (m-1)\right]^{-\alpha} \qquad 0 < u_j \le 1 \; ; \; \alpha > 0.$$

The parameter α measures the strength of agreement between U_i and U_j and does not affect the marginal distributions. Also the distribution of $\mathbf{U}$ is invariant under permutations.

Multivariate Burr distributions are obtained from the standard Cook-Johnson form by transforming the marginal distributions. In particular:

$$Z_j = -\log(U_j^{-1/\alpha} - 1) \qquad j = 1,\ldots,m \; ,$$

$$\mathbf{Z} \sim \text{TB2},$$

$$X_j = \left[d_j^{-1}(U_j^{-1/\alpha} - 1)\right]^{1/c_j} \qquad j = 1,\ldots,m \; ,$$

$$\mathbf{X} \sim \text{TB12}.$$

Earlier we considered the relationships between univariate Burr distributions and other univariate distributions and hence the Cook-Johnson framework enables us to find the links between multivariate Burr distributions and other multivariate distributions. Cook and Johnson 1986 and Johnson 1987 considered extensions of the family which yield a more general form and which could be used to generate 'Generalised' Burr distributions.

One distribution which the TB2 (hence the TB12) is related to, and which does not fit easily into the Cook-Johnson framework, is the Dirichlet distribution. This distribution is of interest since Woodland 1979 has proposed its use in budget share models.

Consider the variables w_j $(0 < w_j < 1)$ $j = 1,\ldots,m$ defined by:

$$w_j = \frac{e^{-z_j}}{1 + \sum_{i=1}^{m} e^{-z_i}} .$$

We find the density of $\mathbf{w} = (w_1 \ldots w_m)'$ to be:

$$f(\mathbf{w}) = \frac{\Gamma(k+m)}{\Gamma(k)} (1 - \mathbf{i}'\mathbf{w}) = \frac{\Gamma(k+m)}{\Gamma(k)} (1 - w_1 - \ldots - w_m).$$

That is, $\mathbf{w}$ has a Dirichlet distribution (see Wilks 1962). This relationship links a multivariate Burr (TB2) distribution to a multivariate Beta distibution through a 'generalised' logistic transformation. This is the multivariate analogue of the relationship found in the univariate section of this paper.

5. Multivariate Applications

We now consider the application of multivariate Burr distributions. So far only the Takahasi-Burr and Durling-Burr distributions have found any application in the literature. Attention has centred upon multivariate failure time (duration) models and hence upon the B12 distributions. Notice however that analysis of log failure times will involve the B2 distribution.

Time To Failure Under Dependence

Johnson and Kotz 1981 constructed a model for time to failure under dependence. They were interested in determining the distribution of the time to failure, T, of a replacement component taken from a stock which has been stored for some time. For example, when the second component is taken from the same product batch as the first and there is batch to batch variation such that we are unable to assume independence between the times to failure

of the first and second components. When the TB12 distribution, with $c_1 = c_2 = c$, is used in this model they find that the survival function of the time to failure (T) of the second component is a function which depends upon k but not upon c.

A Model For Repeated Failure Time Measurements

The standard model for repeated failure time measurements is based upon the multivariate Normal distribution. Crowder 1985 developed a corresponding model for Weibull distributed failure times. Suppose that a response time is measured on an individual on several occasions, giving a data vector $\mathbf{t} = (t_1,\ldots,t_m)'$. The joint distribution of $\mathbf{t}$ is defined from the assumption that, conditional upon $\theta \sim \text{Gamma}(1,k)$, the t_j's are independent Weibull. The resultant joint distribution is a TB12 distribution with 'scale' parameters ξ_j.

To encourage future applications he discussed the properties of the TB12 distribution in the context of failure time modelling and used the joint and marginal moments of $\text{lv}(\mathbf{t}) = (\log(t_1)\ldots\log(t_m))'$ to suggest a method of moments estimation method for n non-censored iid TB12 distributed vectors $\mathbf{t}_1,\ldots,\mathbf{t}_n$. Finally, he applied this repeated failure time model to data on the response times of rats and found it provided an adequate fit to the data.

Quantal Choice

Durling 1969 suggested the use of bivariate TB12 and Durling-Burr distributions in a model of quantal choice. He is concerned with the biological assay of a mixture of two stimulants. Three

distributions for the stimulants (bivariate Normal, Logistic and Burr (Takahasi and Durling versions)) were fitted to seven data sets using a modified non-linear least squares procedure. It was found that the analytic model utilising the Durling-Burr distribution performed best in terms of minimising the sum of squared errors in the estimation. On this basis the Durling-Burr distribution is recommended for use in applications where the number of parameters to be estimated is not of concern (in such cases the bivariate Normal which fits 3 less parameters than the Durling-Burr is recommended).

A Model For Disease Transmission

The bivariate TB2 distribution has been used in the fitting of a multifactorial model of disease transmission to data on the clustering in families of hysteria and sociopathy (see Hutchinson and Satterthwaite 1977). In the paper they consider fitting the bivariate Normal distribution, a TB2 distribution and a bivariate Pareto distribution. They found that the fits were equally good and agreement on threshold levels close, but that the two non-Normal models gave different estimated values for the correlation coefficient r.

6. Concluding Remarks

In this paper we have reviewed the literature concerning the Burr family of distributions. We have summarised some of the results concerning two members of the family, the Burr types II and XII. We have also developed some new univariate and multivariate variants of the Burr type II distribution, which being a

generalisation of the Logistic, may prove an attractive distributional assumption in future applications.

We have not dwelt upon the estimation of models based upon either the B12 or the B2 distributions. In the univariate case the results concerning estimation of the B12 distribution may be found in, *inter alia*, Schmittlein 1983, Wingo 1983 and Al-Marzoug and Ahmad 1985. Zelterman 1987 and Fry 1988 both consider the estimation of models based upon the univariate B2 distribution. Notice however, that since the one parameter B2 distribution is exponential family estimation should be straightforward. In the multivariate case, both Crowder 1985 and Scallan 1987, have considered the estimation of models for repeated measure data based upon the TB12 distribution. Little is known about the estimation of models based upon multivariate B2 distributions. However, Fry 1988 outlines the estimation of a multivariate binomial choice model based upon the TB2 distribution.

Burr's original motivation to find a flexible distribution, which remains tractable to deal with, remains a desirable attribute of any candidate distribution for, particulary multivariate, applications. It is hoped that by reviewing and extending the literature on the Burr family this paper will stimulate further work with these distributions.

Abramowitz, M and Stegun, I. A. eds. (1965), *Handbook of Mathematical Functions*, New York: Dover.

Al-Marzoug, A.M. and Ahmad, M. (1985), "Estimation of Parameters of Burr Probability Model using Fractional Moments", *Pakistan Journal of Statistics*, 1, 67-77.

Awad, A. M. and Gharraf, M. K. (1986), "Estimation of P(Y < X) in the Burr Case: A Comparative Study", *Communications in Statistics - Simulation and Computation*, 15, 389-403.

Bera, A. K. (1982), "A New Test for Normality", *Economics Letters*, 9, 263-268

Brännäs, K. (1986), "Small Sample Properties in the Heterogeneous Weibull Model", *Economics Letters*, 21, 17-20.

Burr, I. W. (1942), "Cumulative Frequency Functions" *Annals of Mathematical Statistics*, 13, 215-232.

Burr, I. W. (1967a), "A Useful Approximation to the Normal Distribution Function, with Application for Simulation", *Technometrics*, 9, 647-651.

Burr, I. W. (1967b), "The Effects of Non-Normality on Constants for $\bar{X}$ and R Charts", *Industrial Quality Control*, 24, 563-569.

Burr, I. W. (1968), "On a General System of Distributions: III. The Sample Range", *Journal of the American Statistical Association*, 63, 636-643.

Burr, I. W. (1973), "Parameters for a General System of Distributions to Match a Grid of α_3 and α_4", *Communications in Statistics*, 2, 1-21.

Burr, I. W. and Cislak, P. J. (1968), "On a General System of Distributions: I. Its Curve-Shape Characteristics; II. The Sample Median", *Journal of the American Statistical Association*, 63, 627-635.

Cook, R. D. and Johnson, M. E. (1981), "A Family of Distributions for Modelling Non-elliptically Symmetric Multivariate Data", *Journal of the Royal Statistical Society*, Series(B), 43, 201-218.

Cook, R. D. and Johnson, M. E. (1986), "Generalised Burr - Pareto - Logistic Distributions with Applications to a Uranium Exploration Data Set", *Technometrics*, 28, 123-131.

Crowder, M. (1985), "A Distributional Model for Repeated Failure Time Measurements", *Journal of the Royal Statistical Society*, Series (B), 47, 447-452.

Drane, J. W., Owen, D. B. and Seibert, G. B. (1978), "The Burr Distribution and Quantal Responses", *Statistischen Hefte*, 19, 204-210.

Dubey, S. D. (1966a), "Characterization Theorems for Several Distributions and their Applications", *The Journal of the Industrial Mathematics Society*, 16, 1-22.

Dubey, S. D. (1966b), "Transformations for Estimation of Parameters", *Journal of the Indian Statistical Association*, 4, 109-124.

Dubey, S. D. (1968), "A Compound Weibull Distribution", *Naval Research Logistics Quarterly*, 15, 179-188.

Dubey, S. D. (1969), " A New Derivation of the Logistic Distribution", *Naval Research Logistics Quarterly*, 16, 37-40.

Dubey, S. D. (1970), "Compound Gamma, Beta and F Distributions", *Metrika*, 17, 27-31.

Durling, F. C. (1969), "Bivariate Probit, Logit and Burrit Analysis", *THEMIS Signal Analysis Statistics Research Program*, Technical Report 41, Southern Methodist University, Dallas, Texas.

Durling, F. C. (1975), "The Bivariate Burr Distribution", in G. P. Patil et al (eds), Statistical Distributions in Scientific Work, Vol. 1, Dordrecht: Reidel.

Evans I. G. and Ragab, A. S. (1983), "Bayesian Inferences Given a Type-2 Censored Sample from a Burr Distribution", *Communications in Statistics - Theory and Methodology*, 12, 1569-1580.

Fry, T. (1988), "On the Use of some Burr Family Distributions in Econometrics", *Unpublished PhD. Thesis*, University of Manchester.

Harris, C. M. and Singpurwalla, N. D. (1969), "On Estimation in Weibull Distributions with Random Scale Parameters", *Naval Research Logistics Quarterly*, 16, 405-410.

Hatke, M. A. (1949), "A Certain Cumulative Probability Function", *Annals of Mathematical Statistics*, 20, 461-463.

Hutchinson, T. P. (1981), "Compound Gamma Bivariate Distributions", *Metrika*, 28, 263-271.

Hutchinson, T.P. and Satterthwaite, S. P. (1977), "Mathematical Models for Describing the Clustering of Sociopathy and Hysteria in Families: A Comment on the Recent Paper by Cloninger et al", *British Journal of Psychiatry*, 130, 294-297.

Johnson, M.E. (1987), *Multivariate Statistical Simulation*, New York: John Wiley and Sons.

Johnson, N. L. and Kotz, S. (1970), *Distibutions in Statistics: Continuous Univariate Distributions - 2*, New York: John Wiley and Sons.

Johnson, N. L. and Kotz, S. (1972), *Distributions in Statistics: Continuous Multivariate Distributions*, New York: John Wiley and Sons.

Johnson, N. L. and Kotz, S. (1981), "Dependent Relevations: Time - to - Failure Under Dependence", *American Journal of Mathematical and Management Science*, 1, 155-165.

Kabe, D. G. (1971), "Some Distributions of Ordered Values from Burr and Beta Distributions", *Canadian Mathematical Bulletin*, 14, 167-173.

Lancaster, T. (1979), "Econometric Methods for the Duration of Unemployment", *Econometrica*, 47, 939-956.

Lancaster, T. (1985), "Generalised Residuals and Heterogeneous Duration Models: With Applications to the Weibull Model", *Journal of Econometrics*, 28, 155-169.

Lomax, K. S. (1954), "Business Failures: Another Example of the Analysis of Failure Data", *Journal of the American Statistical Association*, 49, 847-852.

McDonald, J. B. (1984), "Some Generalized Functions for the Size Distribution of Income", *Econometrica*, 52, 647-663.

Morrison, D. G. and Schmittlein, D. C. (1980), "Jobs, Strikes and Wars: Probability Models for Duration", *Organizational Behavior and Human Performance*, 25, 224-251.

Papadopoulos, A. S. (1978), "The Burr Distibution as a Failure Model from a Bayesian Approach", *I.E.E.E. Transactions on Reliability*, R-27, 369-371.

Pearson, E. S., Johnson, N. L. and Burr, I. W. (1979), "Comparisons of the Percentage Points of Distributions with the same First Four Moments, Chosen from Eight Different Systems of Frequency Curves", *Communications in Statistics - Simulation and Computation*, 13, 191-229.

Poirier, D. J. (1980), "A Lagrange Multiplier Test for Skewness in Binary Logit Models", *Economics Letters*, 5, 141-143.

Prentice, R. L. (1975), "Discrimination Among Some Parametric Models", *Biometrika*, 62, 607-614.

Prentice, R. L. (1976), "A Generalization of the Probit and Logit Methods for Dose Response Curves", *Biometrics*, 32, 761-768.

Rodriguez, R. N. (1977), "A Guide to the Burr Type XII Distributions", *Biometrika*, 64, 129-134.

Rodriguez, R. N. (1980), "Multivariate Burr III Distributions. Part 1: Theoretical Properties", *General Motors Research Laboratories*, Research Report GMR-3232.

Rodriguez, R. N. (1982), "Burr Distributions", in *Encyclopedia of Statistical Science, Vol 1*, S. Kotz and N. L. Johnson eds, New York: John Wiley and Sons.

Rodriguez, R. N. (1983), "Systems of Frequency Surfaces", in *Encyclopedia of Statistical Science, Vol 3*, S. Kotz and N. L. Johnson eds, New York: John Wiley and Sons.

Rodriguez, R. N. and Taniguchi, B. Y. (1980), "A New Statistical Model for Predicting Customer Octane Satisfaction Using Trained Rater Observations", *Transactions of the Society of Automotive Engineers*, SAE Technical Paper No. 801356, 89, 4213-4240.

Satterthwaite, S. P. and Hutchinson, T. P. (1978), "A Generalisation of Gumbel's Bivariate Logistic Distribution", *Metrika*, 25, 163-170.

Scallan, A.J. (1987), "A GLIM Model for Repeated Measurements", *GLIM Newsletter*, 15, 10-22.

Schmittlein, D. C. (1983), "Some Sampling Properties of a Model for Income Distribution", *Journal of Business and Economic Statistics*, 1, 147-153.

Schucany, W. R., Parr, W. C. and Boyer, J. E. (1978), "Correlation Structure in Farlie-Gumbel-Morgenstern Distributions", *Biometrika*, 65, 650-653.

Seibert, G. B. Jr. (1970), "Estimation and Confidence Intervals for Quantal Response or Sensitivity Data", Technical Report No. 66, *THEMIS Signal Analysis Statistics Research Program*, Department of Statistics, Southern Methodist University, Dallas, Texas.

Shah, S. M. and Kabe, D. G. (1981), "Characterizations of Exponential, Pareto, Power Function, Burr and Logistic Distributions by Order Statistics", *Biometrical Journal*, 23, 141-146.

Singh, S. K. and Maddala, G. S. (1976), "A Function for Size Distribution of Incomes", *Econometrica*, 44, 963-970.

Smith, R. J. (1989), "On the Use of Distributional Misspecification Checks in Limited Dependent Variable Models", *Economic Journal*, 99, 178-192.

Stuart, A. S. and Ord, J. K. (1987), *Kendall's Advanced Theory of Statistics. Volume I, Distribution Theory (Fifth Edition)*, London: Charles Griffin and Company Ltd.

Sugiura, N. and Gomi, A. (1985), "Pearson Diagrams for Truncated Normal and Truncated Weibull Distributions", *Biometrika*, 72, 219-222.

Tadikamalla, P. R. (1977), "An Approximation to the Moments and Percentiles of the Gamma Order Statistics", *Sankya*, Series B, 39, 372-381.

Tadikamalla, P. R. (1980a) "On Simulating Non-Normal Distributions", *Psychometrika*, 45, 273-279.

Tadikamalla, P. R. (1980b) "A Look at the Burr and Related Distributions", *International Statistical Review*, 48, 337-344.

Takahasi, K. (1965), "Note on the Multivariate Burr's Distribution", *Annals of the Institute of Statistical Mathematics*, 17, 257-260.

Vodã, V. G. (1982), "Burr Distribution Revisited", *Revue Roumaine de Mathematiques Pures et Appliquees*, 27, 885-893.

Wilks, S. S. (1962), *Mathemtical Statistics*, New York: John Wiley and Sons.

Wingo, D. R. (1983), "Maximum Likelihood Methods for Fitting the Burr Type XII Distribution to Life Test Data", *Biometrical Journal*, 25, 77-84.

Woodland, A. D. (1979), "Stochastic Specification and the Estimation of Share Equations"' *Journal of Econometrics*, 10, 361-383.

Zelterman, D. (1987), "Parameter Estimation in the Generalized Logistic Distribution", *Computational Statistics and Data Analysis*, 5, 177-184.

Zimmer, W. J. and Burr, I. W. (1963), "Variables Sampling Plans Based on Non-Normal Populations", *Industrial Quality Control*, 20, 18-26.

FIGURE ONE

B2(K) Densities for varying values of parameter K

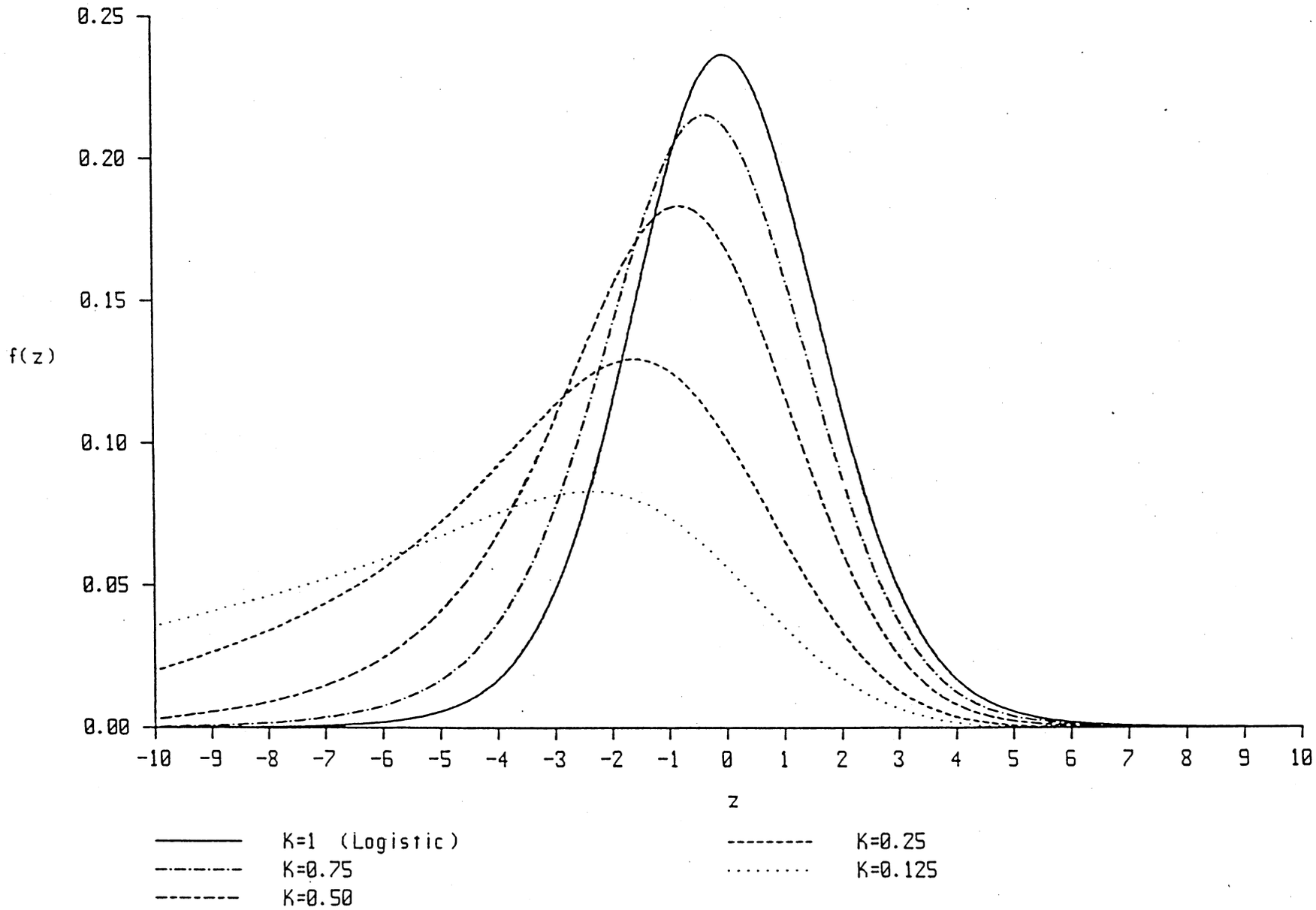


FIGURE TWO

B2(K) Densities for varying values of parameter K

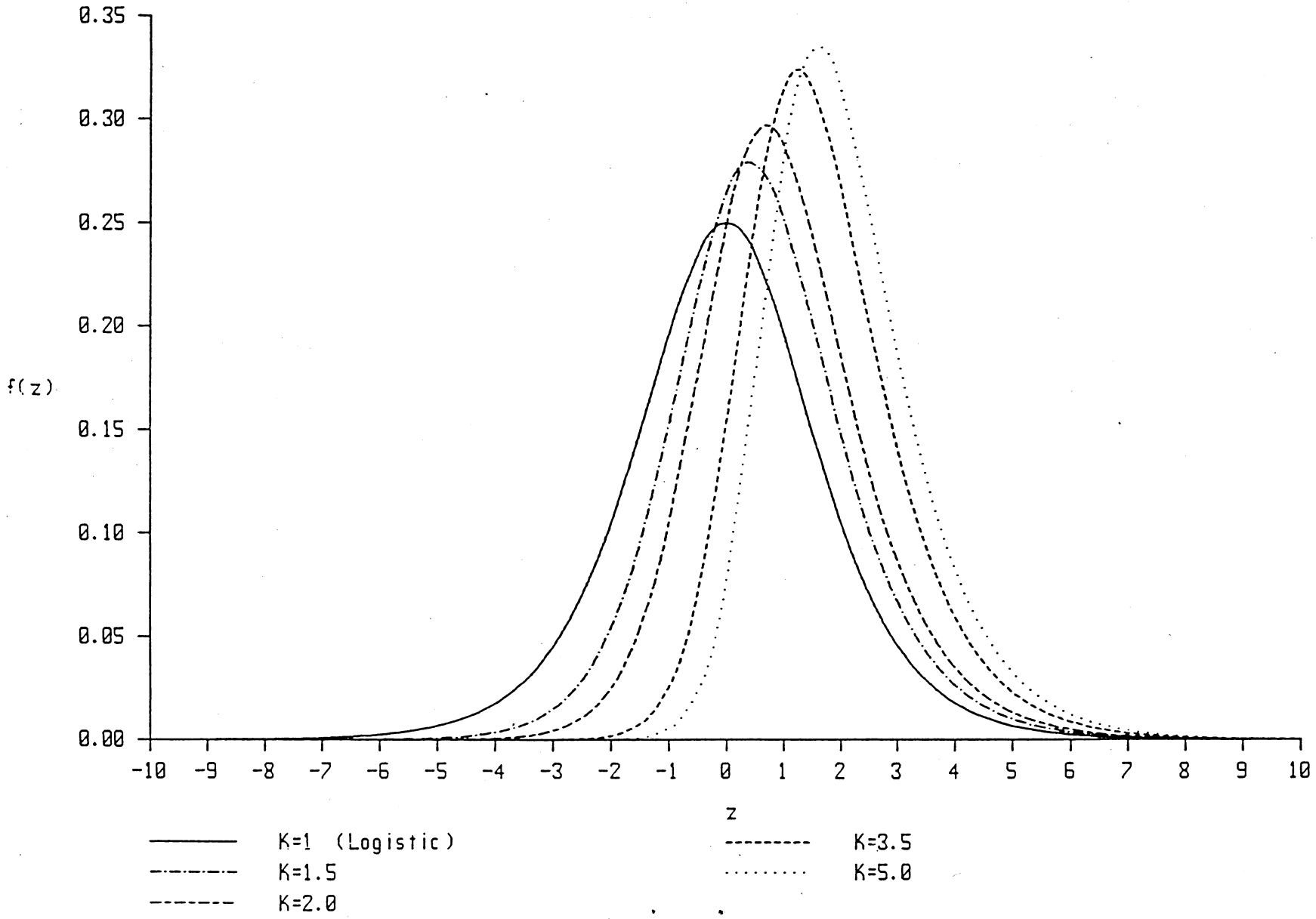
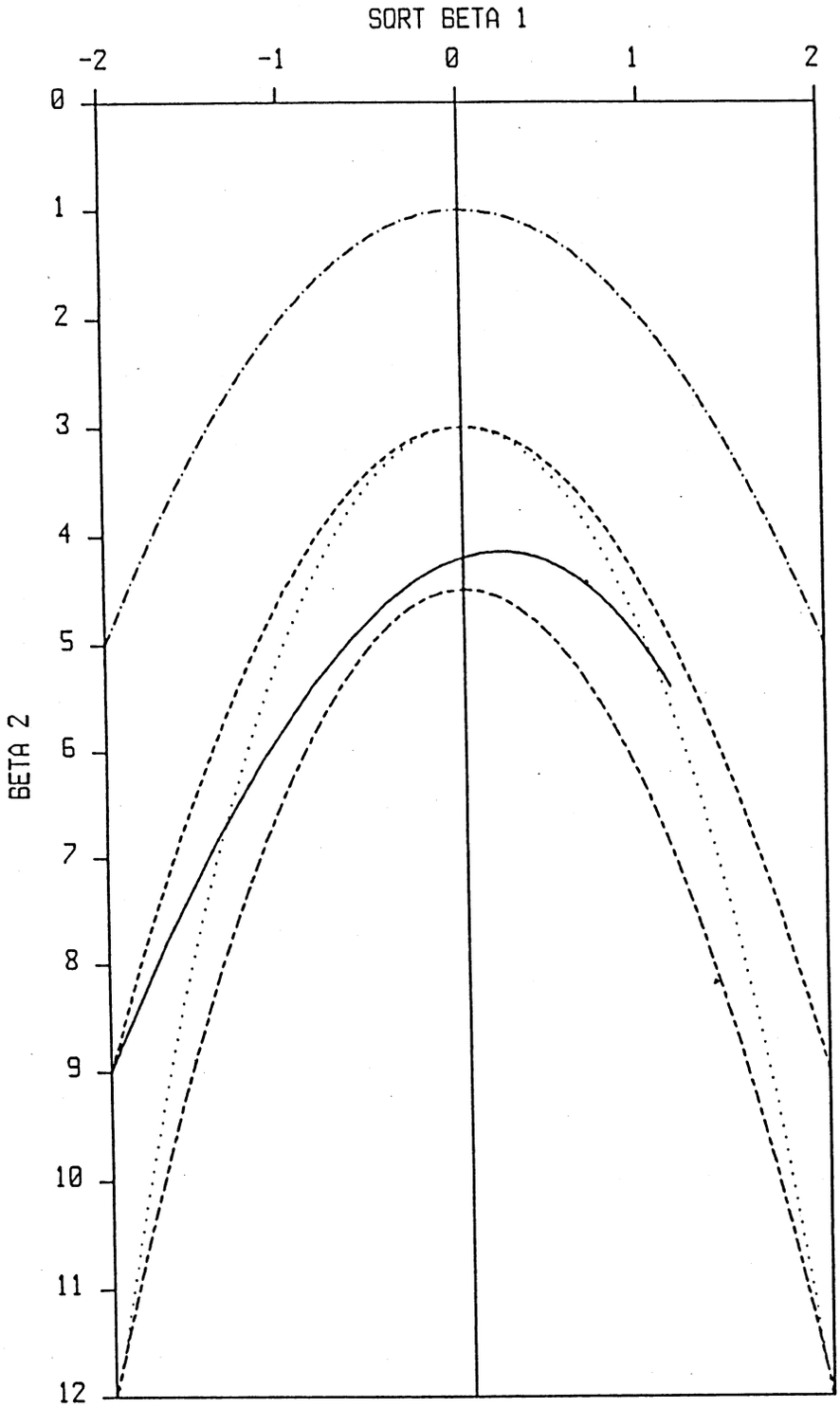


FIGURE THREE

3rd and 4th Moment Coverage Diagram



BURR TYPE II

PEARSON TYPE III

PEARSON TYPE V

LIMIT FOR ALL DISTRIBUTIONS

HETEROTYPIC BOUNDARY

MONASH UNIVERSITY

DEPARTMENT OF ECONOMETRICS

WORKING PAPERS

1988

6/88 Grant H. Hillier, "On the Interpretation of Exact Results for Structural Equation Estimators".

7/88 Kuldeep Kumar, "Some Recent Developments in Non-Linear Time Series Modelling.

8/88 Maxwell L. King, "Testing for Fourth-Order Autocorrelation in Regression Disturbances When First-Order Autocorrelation is Present".

9/88 Ralph D. Snyder, "Kalman Filtering with Partially Diffuse Initial Conditions".

10/88 Asraul Hoque, "Indirect Rational Expectations and Estimation in a Single Equation Macro Model".

11/88 Asraul Hoque, "Efficiency of OLS Relative to C-O for Trended x and Positive Autocorrelation Coefficient".

12/88 Russel J. Cooper & Keith R. McLaren, "Regular Alternatives to the Almost Ideal Demand System".

13/88 Maxwell L. King, "The Power of Student's t Test: Can a Non-Similar Test Do Better?".

14/88 John Hamilton, "On-Line Management of Time Series Databases: Database Retrieval Program DBR".

1989

1/89 R. D. Snyder, C. Chah & C. Lehmer, "Multi-series Heuristics for Exponential Smoothing.

2/89 Maxwell L. King and Phillip M. Edwards, "Transformations for an Exact Goodness-of-Fit Test of Structural Change in the Linear Regression Model".

3/89 Muhammad I. Bhatti and Maxwell L. King, "A Beta-Optimal Test of the Equicorrelation Coefficient".

4/89 Muhammad I. Bhatti, "Null Distribution of the Small Sample Mean Correlation Coefficient: An Application to Medical Research."

5/89 Grant H. Hillier, "On the Normalisation of Structural Equations: Properties of Direction Estimators".

6/89 Brett A. Inder, "A New Test for Autocorrelation in the Disturbances of the Dynamic Linear Regression Model".