S.M. SUNOJ (*) - N. UNNIKRISHNAN NAIR (*)

Bivariate distributions with weighted marginals and reliablity modelling

1. INTRODUCTION

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X : \Omega \to H$ be a random variable, where $H = (a, b)$ is a subset of the real line with $a > 0$ and $b > a$ can be finite or infinite. When the distribution function $F(x)$ of X is absolutely continuous with density function $f(x)$ and $w(x)$, a non-negative function satisfying $Ew(X) = \mu < \infty$, the random variable Y with density

$$g(x) = \frac{w(x) f(x)}{\mu}, \qquad x > a \tag{1.1}$$

is said to have weighted distribution corresponding to X. Introduced by Rao (1965), the concept of weighted distributions has been employed in various practical problems in analysis of family size, study of albinism, human heredity, aerial survey and visibility bias, line transcend sampling, renewal theory, cell cycle analysis and pulse labelling, efficacy of early screening for disease, etiological studies, statistical ecology and reliability modelling. An exhaustive work of research in this area is available in Patil and Rao (1977) and Gupta and Kirmani (1990).

(*) Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, India

The bivariate extension of weighted distribution is discussed in Patil *et al.* (1987). For a non-negative random vector $(X_1,\ X_2)$ with density $f(x_1, x_2)$ and weight function $w(x_1, x_2)$, the weighted distribution is specified by the density

$$g(x_1, x_2) = \frac{w(x_1, x_2)\, f(x_1, x_2)}{E\,w(X_1, X_2)} \tag{1.2}$$

provided the expectation in (1.2) is finite. For properties of (1.2) we refer to the above paper, Mahfoud and Patil (1982) and Arnold and Nagaraja (1991). For most choices of $w(x_1, x_2)$, it is difficult to obtain the length biased forms for the marginal distributions if we use equation (1.2) and therefore modelling with prior knowledge about the individual distributions of X_i becomes impracticable. Accordingly in the present paper, we focus attention on bivariate models that have as their marginals weighted distributions of the form (1.1). Obviously, the discussions concerning bivariate laws with specified marginals (see for example, Johnson and Kotz (1972)) can lead to such cases. However, we prefer to give a different model which is a direct extension of the univariate case that can be useful in reliability studies.

2. The model

Let $(X_1,\ X_2)$ be a non-negative random vector admitting absolutely continuous survival function

$$S(x_1, x_2) = P(X_1 > x_1,\ X_2 > x_2)\,. \tag{2.1}$$

in the support of $\{(x_1, x_2) | x_i > a_i,\ i = 1, 2;\ a_i > 0\}$.

We require a random vector $(Y_1,\ Y_2)$ whose marginal distributions have densities

$$g_i(x_i) = \frac{w_i(x_i)\, f_i(x_i)}{E\,w_i(X_i)}\,, \qquad i = 1, 2\,. \tag{2.2}$$

where $f_i(x_i)$ is the density of X_i.

In this paper we consider a joint survival function for the random variable $(Y_1,\ Y_2)$ defined by

$$R_1(x_1, x_2) = \frac{m_2(x_1, x_2) m_1(x_1, a_2)}{m_2(x_1, a_2) m_1(a_1, a_2)} S(x_1, x_2), \qquad x_i > a_i \geq 0\,. \tag{2.3}$$

where

$$m_i(x_1, x_2) = E[w_i(X_i)|X_j > x_j,\ j = 1, 2], \quad i = 1, 2\,.$$

It is easy to see that the marginals of (2.3) are $g_i(x_i)$. By symmetry, one can also define a joint distribution with marginals (2.2) by using the joint survival function

$$R_2(x_1, x_2) = \frac{m_1(x_1, x_2)m_2(a_1, x_2)}{m_1(a_1, x_2)m_2(a_1, a_2)}S(x_1, x_2)\,. \tag{2.4}$$

In general, this does not produce the same distribution as (2.3). One set of conditions under which (2.3) and (2.4) are identical is stated below.

THEOREM 2.1. A necessary and sufficient condition for the survival functions $R_1(x_1, x_2)$ and $R_2(x_1, x_2)$ to be identical is that

$$\frac{m_1(x_1, x_2)}{m_2(x_1, x_2)} = A(x_1)B(x_2) \tag{2.5}$$

for all $x_i \geq a_i > 0$ and some functions $A(\cdot)$ and $B(\cdot)$.

Proof: On equating the expressions in (2.3) and (2.4),

$$\frac{m_1(x_1, x_2)}{m_2(x_1, x_2)} = \frac{m_1(a_1, x_2)}{m_2(a_1, x_2)}\,\frac{m_2(a_1, a_2)}{m_1(a_1, a_2)}\,\frac{m_1(x_1, a_2)}{m_2(x_1, a_2)} \tag{2.6}$$

which is of the form (2.5). Conversely, if (2.5) is true it is easy to see that (2.6) holds and hence

$$R_1(x_1, x_2) = R_2(x_1, x_2)\,.$$

Example. Bivariate Pareto distribution given in equation (4.6).

3. Length biased models

The length biased models result when $w_i(x_i) = x_i$ in equation (2.2). In this case

$$m_1(x_1, a_2) = E[X_1 | X_1 > x_1]$$

is the vitality function of X_1. For a detailed discussion of vitality functions and their use in clarifying the distinction between increasing failure rate and decreasing mean residual life we refer to Kupka and Loo (1989). By analogy, in the two dimensional case

$$m(x_1, x_2) = E(X_2 | X_1 > x_1,\ X_2 > x_2)$$

is a component of the bivariate vitality function

$$V(x_1, x_2) = (m_1(x_1, x_2), m_2(x_1, x_2))\,.$$

In a two-component system, $m_2(x_1, x_2)$ measures the expected age at failure of the second component at age x_2, as the sum of the present age x_2 and residual life conditioned on $X_1 > x_1$. We now give some examples of bivariate distributions possessing length biased marginals

Examples

1. In the Gumbel's bivariate exponential distribution

$$S(x_1, x_2) = \exp[-\lambda_1 x_1 - \lambda_2 x_2 - \theta x_1 x_2],\ \lambda_1, \lambda_2, x_1, x_2 > 0,\ 0 \le \theta < \lambda_1 \lambda_2$$
$$R_1(x_1, x_2) = (1 + \lambda_2\, x_2 + \theta x_1 x_2)(1 + \lambda_1 x_1) S(x_1, x_2)\,.$$

2. For the bivariate Pareto II with

$$S(x_1, x_2) = (1 + a_1 x_1 + a_2 x_2 + b x_1 x_2)^{-c},$$
$$x_1, x_2, a_1, a_2, c > 0,\ 0 \le b \le (c+1) a_1 a_2,$$
$$R_1(x_1, x_2) = [1 + c x_2 (a_2 + b x_1)(1 + a_1 x_1)^{-1}][1 + c a_1 x_1] S(x_1, x_2)\,.$$

3. If (X_1, X_2) has bivariate beta distribution with

$$S(x_1, x_2) = (1 - p_1 x_1 - p_2 x_2 + q x_1 x_2)^d,\ p_1, p_2, d > 0,\ 1 - d < \frac{q}{p_1 p_2} \le 1,$$
$$0 < x_1 < 1/p_1, \quad 0 < x_2 < \frac{(1 - p_1 x_1)}{(p_2 - q x_1)}$$

then

$$R_1(x_1, x_2) = [1 + d x_2 (p_2 - q x_1)(1 - p_1 x_1)^{-1}][1 + d p_1 x_1] S(x_1, x_2)\,.$$

4. Structural relationships

In this section we prove certain relationships that are useful in the context of reliability modelling and form extensions of some of the results in the univariate case established in Gupta and Kirmani (1990).

4.1. Failure rates

The vector failure rate (see for example Johnson and Kotz (1975)) of the random vector $(X_1,\ X_2)$ is

$$(h_1(x_1, x_2), h_2(x_1, x_2))$$

where

$$h_i(x_1, x_2) = -\frac{\partial \log S(x_1, x_2)}{\partial x_i}, \quad i = 1, 2 .$$

Similarly, defining $(k_1(x_1, x_2), k_2(x_1, x_2))$ as the failure rate of $(Y_1,\ Y_2)$ we have from (2.3),

$$k_2(x_1, x_2) = h_2(x_1, x_2) - \frac{\partial \log m_2(x_1, x_2)}{\partial x_2} . \tag{4.1}$$

Using

$$\frac{\partial m_2}{\partial x_2} = h_2(x_1, x_2)[m_2(x_1, x_2) - x_2] ,$$

the identity

$$k_2(x_1, x_2) = x_2 h_2(x_1, x_2)/m_2(x_1, x_2) \tag{4.2}$$

results. Further

$$k_1(x_1, x_2) = h_1(x_1, x_2) - \frac{\partial}{\partial x_1} \log \left[\frac{m_2(x_1, x_2) m_1(x_1, a_2)}{m_2(x_1, a_2)}\right] . \tag{4.3}$$

Equations (4.2) and (4.3) provide tools for mutual characterizations of the distributions of $(X_1,\ X_2)$ and $(Y_1,\ Y_2)$ as the failure rates uniquely determine the corresponding distributions. When $R_1(x_1, x_2) \equiv R_2(x_1, x_2)$, equation (4.1) and (4.2) reduces to

$$k_i(x_1, x_2) = x_i h_i(x_1, x_2)/m_i(x_1, x_2), \quad i = 1, 2 . \tag{4.4}$$

Since $m_1(x_1, x_2)$ is increasing in x_1 and $m_2(x_1, x_2)$ is increasing in x_2 we see that

$$k_i(x_1, x_2) \leq h_i(x_1, x_2), \quad i = 1, 2\,.$$

Further $R(x_1, x_2) \geq S(x_1, x_2)$ and $R(x_1, x_2)/S(x_1, x_2)$ is non-decreasing in x_1 and x_2.

We now present theorems that enable identification of the failure time distribution, through the relationships between the failure rates of $(X_1,\ X_2)$ and $(Y_1,\ Y_2)$.

THEOREM 4.1. The failure rates of $(X_1,\ X_2)$ and $(Y_1,\ Y_2)$ of the distribution with survival function $R_1(x_1, x_2)$ in the length biased case, satisfy the relationship

$$k_i(x_1, x_2) = C_i h_i(x_1, x_2) \tag{4.5}$$

for all $x_i > a_i$, $i = 1, 2$, where C_i is less than unity and independent of x_i if and only if $(X_1,\ X_2)$ has bivariate Pareto distribution specified by

$$S(x_1, x_2) = \left(\frac{x_1}{a_1}\right)^{-\alpha_1} \left(\frac{x_2}{a_2}\right)^{-\alpha_2} \left(\frac{x_1}{a_1}\right)^{-\theta \log(\frac{x_2}{a_2})}, x_i > a_i, \quad \theta < \alpha_1\alpha_2 \tag{4.6}$$

provided

$$\log\left(\frac{x_i}{a_i}\right) > \max\left(0, \frac{1-\alpha_j}{\theta}\right), \quad i \neq j,\ \alpha_j > 0,\ i, j = 1, 2\,.$$

Proof: When $(X_1,\ X_2)$ is distributed as (4.6)

$$h_i(x_1, x_2) = \left(\alpha_i + \theta \log\left(\frac{x_j}{a_j}\right)\right) x_i^{-1}$$

$$m_i(x_1, x_2) = \left(\alpha_i + \theta \log\left(\frac{x_j}{a_j}\right)\right) x_i \Bigg/ \left(\alpha_i + \theta \log\left(\frac{x_j}{a_j}\right) - 1\right)\ i \neq j,\ i, j = 1, 2.$$

so that (4.5) holds with

$$C_i(x_j) = \left(\alpha_i + \theta \log\left(\frac{x_j}{a_j}\right) - 1\right) \Bigg/ \left(\alpha_i + \theta \log\left(\frac{x_j}{a_j}\right)\right) < 1$$

which is independent of x_i.

Conversely, for $i = 2$, (4.2) and (4.5) imply

$$m_2(x_1, x_2) = x_2 / C_2(x_1) \tag{4.7}$$

which gives

$$S(x_1, x_2) = K_1(x_1) \left(\frac{x_2}{a_2}\right)^{-[1-C_2(x_1)]^{-1}} \tag{4.8}$$

Substituting (4.7) into (4.3)

$$k_1(x_1, x_2) = h_1(x_1, x_2) - \frac{\partial \log m_1(x_1, a_2)}{\partial x_1} \tag{4.9}$$

Using (4.5) with $i = 1$ and taking the limit as x_2 tends to a_2,

$$m_1(x_1, a_2) = x_1 / C_1, \quad C_1 = C_1(a_2) \tag{4.10}$$

Now from (4.9),

$$h_1(x_1, x_2) = \frac{\partial \log S}{\partial x_1} = [\{1 - C_1(x_2)\} x_1]^{-1}$$

and

$$S(x_1, x_2) = K_2(x_2) \left(\frac{x_1}{a_1}\right)^{-[1-C_1(x_2)]^{-1}} \tag{4.11}$$

From $m(x_1, a_2)$ in equation (4.10)

$$S(x_1, a_2) = \left(\frac{a_1}{x_1}\right)^{(1-C_1)^{-1}}, \quad x_1 > a_1 .$$

Similarly

$$S(a_1, x_2) = \left(\frac{a_2}{x_2}\right)^{(1-C_2)^{-1}}, \quad x_2 > a_2$$

where $C_2 = C_2(a_1)$.

Now from (4.8) as x_2 tends to a_2

$$K_1(x_1) = S(x_1, a_2)$$

and from (4.11) as x_1 tend to a_1,

$$K_2(x_2) = S(a_1, x_2) .$$

Inserting the expression for $S(x_1, a_2)$ and $S(a_1, x_2)$ in the place of $K_1(x_1)$ and $K_2(x_2)$ in (4.8) and (4.11) and equating

$$\left(\frac{a_1}{x_1}\right)^{(1-C_1)^{-1}} \left(\frac{a_2}{x_2}\right)^{(1-C_2(x_1))^{-1}} = \left(\frac{a_1}{x_1}\right)^{(1-C_1(x_2))^{-1}} \left(\frac{a_2}{x_2}\right)^{(1-C_2)^{-1}}, x_i > a_i .$$

This means that

$$\left(\log\left(\frac{x_i}{a_i}\right)\right)^{-1} [(1 - C_j(x_i))^{-1} - (1 - C_j)^{-1}] = \theta ,$$

a constant. Hence

$$1 - C_j(x_i) = \left(\alpha_j + \theta \log\left(\frac{x_i}{a_i}\right)\right)^{-1} \tag{4.12}$$

with $a_j = (1 - C_j)^{-1}$. Using (4.12) in (4.8) we get (4.6). This completes the proof.

Remarks

1. When C_i is independent of x_j as well, bivariate Pareto with independent marginals is characterized.
2. The Gumbel's (1960) bivariate exponential distribution is characterized by the property

$$h_i(x_1, x_2) - k_i(x_1, x_2) = C_j$$

when the weight functions are $w_i(x_i) = \exp\{-a_i x_i\}$.

THEOREM 4.2. Let (X_1, X_2) be a random vector admitting absolutely continuous survival function $S(x_1, x_2)$ in the support of $\{(x_1, x_2)|x_i > a_i, a_i > 0\}$. Then the following statements are equivalent

(i) (X_1,X_2) has bivariate Pareto distribution with survival function (4.6)

(ii) $R_1(x_1, x_2) = \frac{x_1 x_2}{a_1 a_2} S(x_1, x_2) = R_2(x_1, x_2)$

(iii) $h_i(x_1, x_2) - k_i(x_1, x_2) = x_i^{-1}, \ i = 1, 2.$

Proof: Assume that (i) is true, then from the expressions for $h_i(x_1, x_2)$ and $m_i(x_1, x_2)$ given earlier (ii) follows. By logarithmic differentiation of (ii) with respect to x_i we get (iii). It remain to establish the reverse implications. Suppose (iii) holds. Taking $i = 1$ and integrating with respect to x_1,

$$R_1(x_1, x_2) = S(x_1, x_2)p_1(x_2)x_1$$

Similarly for $i = 2$,

$$R_2(x_1, x_2) = S(x_1, x_2)p_2(x_1)x_2 \,.$$

From the last two equations $p_j(x_i) = Cx_i$. Usimg the boundary conditions $S(a_1, a_2) = 1 = R(a_1, a_2)$, statement (ii) is obtained. Lastly, when (ii) holds,

$$m_i(x_1, x_2) = x_i/C_i(x_j)$$

and hence (i) follows if we proceed as in the proof of Theorem 4.1.

ACKNOWLEDGMENT

The first author acknowledges the financial assistance received from the C.S.I.R. for carrying out this research work.

REFERENCES

ARNOLD, B. C. and NAGARAJA, H. N. (1991) On some properties of bivariate weighted distributions, *Commun. Statist.- Theor & Meth.*, 20, 1853-1860.

GUMBEL, E. J. (1960) Bivariate exponential distribution, *J. Amer. Statist. Assoc.*, 55, 698-707.

GUPTA, R. C. and KIRMANI, S. N. U. A. (1990) The role of weighted distributions on stochastic modelling, *Commun. Statist.- Theor. & Meth.*, 19, 3147-3162.

JOHNSON, N. L. and KOTZ, S. (1972) *Distributions in Statistics, Continuous multivariate distributions*, John Wiley and Sons, New York.

JOHNSON, N. L. and KOTZ, S. (1975) A vector valued multivariate hazard rate, *J. Multiv. Anal.*, 5, 53-66.

KUPKA, J. and LOO, S. (1989) The hazard and vitality measures of ageing, *J. Appl. Prob.*, 26, 532-542.

MAHFOUD, M. and PATIL, G. P. (1982) *On weighted distributions*, In *Statistics and Probability*: Essays in honour of C. R. Rao, G. Kallianpur *et al.* (eds.), North Holland, 479-492.

PATIL, G. P. and RAO, C. R. (1977) *The Weighted distributions: A survey and their applications, Applications of Statistics*, P.R. Krishnaiah (ed.), North Holland Publishing Company, 383-405.

PATIL, G. P., RAO, C. R. and RATNAPARKHI, M. V. (1987) *Bivariate weighted distributions and related applications*, Technical report, Centre for Statistical Ecology and Environmental Statistics.

RAO, C. R. (1965) *On discrete distributions arising out of methods of ascertainment: Classical and Contagious Discrete Distributions*, G. P. Patil (ed.), Statistical Publishing Society and Pergamon Press, 320-332.

Bivariate distributions with weighted marginals and reliablity modelling

SUMMARY

In this paper, a family of bivariate distributions whose marginals are weighted distributions in the original variables is studied. The relationship between the failure rates of the derived and original models are obtained. These relationships are used to provide some characterizations of specific bivariate models.

Distribuzioni bivariate con marginali pesate e modelli di affidabilità

RIASSUNTO

Nel presente articolo, viene studiata una famiglia di distribuzioni bivariate le cui marginali sono distribuzioni pesate nelle variabili originarie. Si ottengono le relazioni tra i tempi di avaria dei modelli originari e quelli dei modelli derivati. Tali relazioni sono usate per fornire alcune caratterizzazioni di determinati modelli bivariati.

KEY WORDS

Weighted distributions; Bivariate models; Reliability; Failure rate; Caracterization.

[Manuscript received September 1998; final version received June 1999.]