APPLICATIONES MATHEMATICAE
33,2 (2006), pp. 185–193

Saralees Nadarajah (Manchester)
B. M. Golam Kibria (Miami, FL)

DROUGHT MODELS BASED ON BURR XII VARIABLES

Abstract. Burr distributions are some of the most versatile distributions in statistics. In this paper, a drought application is described by deriving the exact distributions of $U = XY$ and $W = X/(X+Y)$ when X and Y are independent Burr XII random variables. Drought data from the State of Nebraska are used.

1. Introduction. The Burr distribution is one of the most versatile distributions in statistics. As shown by Rodriguez (1977) and Tadikamalla (1980), the Burr distribution contains the shape characteristics of the normal, log-normal, gamma, logistic and exponential distributions as well as a significant portion of the Pearson type I, II, V, VII, IX and XII families. It has received applications in life testing (see Wingo (1983, 1993)) and many other areas.

The aim of this paper is to provide a drought application by deriving the exact distributions of $U = XY$ and $W = X/(X+Y)$ when X and Y are independent Burr XII random variables with the pdfs

$$f_X(x) = \frac{kcx^{c-1}}{(1+x^c)^{k+1}} \tag{1}$$

and

$$f_Y(y) = \frac{ldy^{d-1}}{(1+y^d)^{l+1}}, \tag{2}$$

respectively, for $x > 0$, $y > 0$, $k > 0$, $l > 0$, $c > 0$ and $d > 0$.

Products and ratios of random variables arise naturally in many hydrological problems. They arise in particular with respect to drought modeling.

2000 *Mathematics Subject Classification*: 33C90, 62E99.

Key words and phrases: Burr XII distribution, drought modeling, products of random variables, ratios of random variables.

For example, if X and Y denote the drought intensity and the drought duration then $U = XY$ will represent the magnitude of drought. If X and Y denote the drought duration and the successive non-drought duration then $W = X/(X+Y)$ will represent the proportion of drought events (see Section 4).

This paper is organized as follows. In Sections 2 and 3, explicit expressions for the pdfs of $U = XY$ and $W = X/(X+Y)$ are derived. In Section 4, an application of the results to drought data from Nebraska is provided. The calculations of this paper involve the generalized hypergeometric function defined by

$$ {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;x) = \sum_{k=0}^{\infty} \frac{(a_1)_k(a_2)_k\cdots(a_p)_k}{(b_1)_k(b_2)_k\cdots(b_q)_k}\frac{x^k}{k!}, $$

where $(c)_k = c(c+1)\cdots(c+k-1)$ denotes the ascending factorial. The properties of the generalized hypergeometric function can be found in Prudnikov *et al.* (1986) and Gradshteyn and Ryzhik (2000).

2. Distribution of the product. Theorem 1 expresses the pdf of $U = XY$ as a finite linear combination of generalized hypergeometric functions.

THEOREM 1. *Suppose X and Y are independent Burr XII random variables with pdfs* (1) *and* (2), *respectively. If $c/d = p/q$* (*where $p \geq 1$ and $q \geq 1$ are co-prime integers*) *then the pdf of $U = XY$ can be expressed as*

$$ f_U(u) = \begin{cases} klcu^{d-1}(V_1^+ + V_2^+) & \text{if } u \leq 1, \\ klcu^{d-1}(V_1^- + V_2^-) & \text{if } u \geq 1, \end{cases} \tag{3} $$

where

$$ V_1^+ = \sum_{j=0}^{q-1} \frac{(-1)^j}{j!} E_1^+ u^{ds}\ {}_nF_{n-1}(1, \Delta(p,a_1), \Delta(q,b_1); \Delta(p,c_1), \Delta(q,1+j); z), $$

$$ V_2^+ = \sum_{h=0}^{p-1} \frac{(-1)^h}{h!} E_2^+ u^{dh}\ {}_nF_{n-1}(1, \Delta(p,a_2), \Delta(q,b_2); \Delta(p,1+h), \Delta(q,c_2); z), $$

$$ V_1^- = \sum_{h=0}^{p-1} \frac{(-1)^h}{h!} E_1^- u^{d\alpha}\ {}_nF_{n-1}(1, \Delta(p,a_3), \Delta(q,b_3); \Delta(p,1+h), \Delta(q,c_3); 1/z), $$

$$ V_2^- = \sum_{j=0}^{q-1} \frac{(-1)^j}{j!} E_2^- u^{d\beta}\ {}_nF_{n-1}(1, \Delta(p,a_4), \Delta(q,b_4); \Delta(p,c_4), \Delta(q,1+j); 1/z), $$

$$ E_1^+ = \frac{\Gamma(1+k+j)\Gamma(l+c(1+j)/d)\Gamma(1-c(1+j)/d)}{\Gamma(1+l)\Gamma(1+k)}, $$

$$E_2^+ = \frac{d\Gamma(1-d(1+h)/c)\Gamma(k+d(1+h)/c)\Gamma(1+l+h)}{c\Gamma(1+k)\Gamma(1+l)},$$
$$E_1^- = \frac{d\Gamma(1+d(l+h)/c)\Gamma(k-d(l+h)/c)\Gamma(1+l+h)}{c\Gamma(1+k)\Gamma(1+l)},$$
$$E_2^- = \frac{\Gamma(1+j+k)\Gamma(l-c(j+k)/d)\Gamma(1+c(j+k)/d)}{\Gamma(1+k)\Gamma(1+l)},$$

$$n = p+q+1, \quad s = \frac{c(j+1)}{d} - 1, \quad \alpha = -l-1-h, \quad \beta = -1 - \frac{c(j+k)}{d},$$
$$\begin{aligned} a_1 &= \frac{c(j+1)}{d} + l, & b_1 &= 1+j+k, & c_1 &= \frac{c(j+1)}{d}, \\ a_2 &= 1+l+h, & b_2 &= \frac{d(h+1)}{c} + k, & c_2 &= \frac{d(h+1)}{c}, \\ a_3 &= l+1+h, & b_3 &= 1 + \frac{d(l+h)}{c}, & c_3 &= \frac{d(l+h)}{c} - k + 1, \\ a_4 &= l + \frac{c(j+k)}{d}, & b_4 &= k+1+j, & c_4 &= \frac{c(j+k)}{d} - l + 1 \end{aligned}$$

and

$$z = (-1)^{p+q} u^{pd}.$$

The symbol $\Delta(k,a)$ *denotes the sequence* $a/k, (a+1)/k, \ldots, (a+k-1)/k$.

Proof. Transform $(X,U) = (X, XY)$. Under this transformation, the joint pdf of (X,U) can be expressed as

$$f(x,u) = (1/x) f_X(x) f_Y(u/x) = \frac{klcdu^{d-1} x^{c+dl-1}}{(1+x^c)^{k+1}(x^d+u^d)^{l+1}}.$$

Thus, the pdf of U can be expressed as

$$\begin{aligned} f_U(u) &= klcdu^{d-1} \int_0^\infty x^{c+dl-1}(1+x^c)^{-k-1}(x^d+u^d)^{-l-1} dx \\ &= klcu^{d-1} \int_0^\infty y^{l+c/d-1}(1+y^{c/d})^{-k-1}(y+u^d)^{-l-1} dy, \end{aligned} \tag{4}$$

which follows by setting $y = x^d$. The result of the theorem follows by using equation (2.2.2.6) in Prudnikov *et al.* (1986, Volume 1) to calculate the integral in (4). ∎

3. Distribution of the ratio. Theorem 2 expresses the pdf of $W = X/(X+Y)$ as a finite linear combination of generalized hypergeometric functions.

THEOREM 2. *Suppose X and Y are independent Burr XII random variables with pdfs (1) and (2), respectively. If $c/d = p/q$ (where $p \geq 1$ and $q \geq 1$ are co-prime integers) then the pdf of $W = X/(X+Y)$ can be expressed as*

$$f_W(w) = \begin{cases} kldw^{-ck-1}(1-w)^{ck-1}(V_1^+ + V_2^+) & \text{if } w \geq 1/2, \\ kldw^{-ck-1}(1-w)^{ck-1}(V_1^- + V_2^-) & \text{if } w \leq 1/2, \end{cases} \tag{5}$$

where

$$V_1^+ = \sum_{j=0}^{q-1} \frac{(-1)^j}{j!} E_1^+ \left(\frac{1-w}{w}\right)^{sc} \times {}_nF_{n-1}(1, \Delta(p, a_1), \Delta(q, b_1); \Delta(p, c_1), \Delta(q, 1+j); z),$$

$$V_2^+ = \sum_{h=0}^{p-1} \frac{(-1)^h}{h!} E_2^+ \left(\frac{1-w}{w}\right)^{hc} \times {}_nF_{n-1}(1, \Delta(p, a_2), \Delta(q, b_2); \Delta(p, 1+h), \Delta(q, c_2); z),$$

$$V_1^- = \sum_{h=0}^{p-1} \frac{(-1)^h}{h!} E_1^- \left(\frac{1-w}{w}\right)^{\alpha c} \times {}_nF_{n-1}(1, \Delta(p, a_3), \Delta(q, b_3); \Delta(p, 1+h), \Delta(q, c_3); 1/z),$$

$$V_2^- = \sum_{j=0}^{q-1} \frac{(-1)^j}{j!} E_2^- \left(\frac{1-w}{w}\right)^{\beta c} \times {}_nF_{n-1}(1, \Delta(p, a_4), \Delta(q, b_4); \Delta(p, c_4), \Delta(q, 1+j); 1/z),$$

$$E_1^+ = \frac{\Gamma(1+l+j)\Gamma(1+d(1+j)/c)\Gamma(k-d(1+j)/c)}{\Gamma(1+l)\Gamma(1+k)},$$

$$E_2^+ = \frac{c\Gamma(1-c(k+h)/d)\Gamma(l+c(k+h)/d)\Gamma(1+k+h)}{d\Gamma(1+k)\Gamma(1+l)},$$

$$E_1^- = \frac{c\Gamma(1+c(1+h)/d)\Gamma(l-c(1+h)/d)\Gamma(1+k+h)}{d\Gamma(1+k)\Gamma(1+l)},$$

$$E_2^- = \frac{\Gamma(1+j+l)\Gamma(1-d(j+l)/c)\Gamma(k+d(j+l)/c)}{\Gamma(1+k)\Gamma(1+l)},$$

$$n = p+q+1, \quad s = \frac{d(j+1)}{c} - k, \quad \alpha = -k-1-h, \quad \beta = -k - \frac{d(j+l)}{c},$$

$$a_1 = \frac{d(j+1)}{c} + 1, \quad b_1 = 1+j+l, \quad c_1 = \frac{d(j+1)}{c} - k + 1,$$

$$a_2 = 1+k+h, \quad b_2 = \frac{c(h+k)}{d} + l, \quad c_2 = \frac{c(h+k)}{d},$$

$$a_3 = 1+k+h, \quad b_3 = 1 + \frac{c(1+h)}{d}, \quad c_3 = \frac{c(1+h)}{d} - l + 1,$$

$$a_4 = k + \frac{d(j+l)}{c}, \qquad b_4 = 1 + l + j, \qquad c_4 = \frac{d(j+l)}{c}$$

and

$$z = (-1)^{p+q} \left(\frac{1-w}{w} \right)^{pc}.$$

The symbol $\Delta(k,a)$ denotes the sequence $a/k, (a+1)/k, \ldots, (a+k-1)/k$.

Proof. Transform $(R, W) = (X+Y, X/R)$. Under this transformation, the joint pdf of (R, W) can be expressed as

$$f(r,w) = r f_X(rw) f_Y(r(1-w)) = \frac{klcd r^{c+d-1} w^{c-1} (1-w)^{d-1}}{\{1+(rw)^c\}^{k+1} \{1+(r(1-w))^d\}^{l+1}}.$$

Thus, the pdf of W can be expressed as

$$(6) \qquad f_W(w)$$
$$= klcd w^{c-1} (1-w)^{d-1} \int_0^\infty r^{c+d-1} \{1+(rw)^c\}^{-k-1} \{1+(r(1-w))^d\}^{-l-1}\, dr$$
$$= kld w^{-ck-1} (1-w)^{ck-1} \int_0^\infty y^{d/c} (1+y^{d/c})^{-l-1} \left(y + \left(\frac{1-w}{w} \right)^c \right)^{-k-1} y,$$

which follows by setting $y = (rw)^c$. The result of the theorem follows by using equation (2.2.2.6) in Prudnikov *et al.* (1986, Volume 1) to calculate the integral in (6). ∎

4. Application. Here, we return to the drought problem discussed in Section 1 and provide an application of the model given by (1)–(2). We use the drought data from the State of Nebraska. The data consists of the monthly modified Palmer Drought Severity Index (PDSI) from the period from January 1895 to December 2004. A drought is said to have happened when PDSI is below 0 and is defined by the theory of runs (Yevjevich, 1967). The State of Nebraska is divided into eight climate divisions numbered 1, 2, 3, 5, 6, 7, 8 and 9—there is no climate division 4 for Nebraska. Some statistics of the observed drought for the eight climatic divisions are summarized in Table 1.

Using the PDSI data, data on drought duration, non-drought duration and drought intensity were obtained for each climate division. The interest is in determining the distributions of

1. the magnitude of droughts (U) = drought intensity × drought duration;
2. the proportion of droughts (W) = drought duration/(drought duration + non-drought duration).

Table 1. Basic drought statistics for Nebraska PDSI data

Climate division	No of droughts	Drought frequency (number/year)	Mean drought duration (months)
1	83	0.75	6.0
2	66	0.60	8.6
3	89	0.81	6.3
5	81	0.74	6.3
6	90	0.82	6.3
7	81	0.74	6.1
8	76	0.69	6.5
9	74	0.67	7.5

The distribution of U was determined by fitting the model given by (1) and (2) to the observed values of drought duration (X) and non-drought duration (Y), respectively, and using equation (3) to compute the fitted pdf. The distribution of W was determined by fitting (1) and (2) to the observed values of drought intensity (X) and drought duration (Y), respectively, and using equation (5) to compute the fitted pdf. The fitting of (1) and (2) was performed by the method of maximum likelihood. The quasi-Newton algorithm nlm in the R software package (Dennis and Schnabel, 1983; Schnabel *et al*, 1985; Ihaka and Gentleman, 1996) was used to maximize the likelihood. The parameter estimates of (1)–(2) for data on drought intensity, drought duration and non-drought duration are shown in Tables 2–4.

Table 2. Parameter estimates of (1)–(2) for drought intensity data

Climate division	$\widehat{k}$ ($\widehat{l}$)	$\widehat{c}$ ($\widehat{d}$)
1	0.608	1.135
2	0.579	0.992
3	0.669	1.046
5	0.631	1.066
6	0.587	1.216
7	0.567	1.289
8	0.527	1.337
9	0.504	1.208

The fitted pdfs of U and W for the eight climate divisions are shown in Figures 1 and 2. The pdfs $\tilde{f}_U(u)$ were computed using (3) with the parameter values specified by Tables 2 and 3 while the pdfs $\tilde{f}_W(w)$ were computed using (5) with the parameter values specified by Tables 3 and 4. The fitted pdfs are overlayed with the histograms of the observed data on U and W. It is evident from both the figures that there is little difference between the climate divisions. This is what one would expect given the geography of the State of Nebraska.

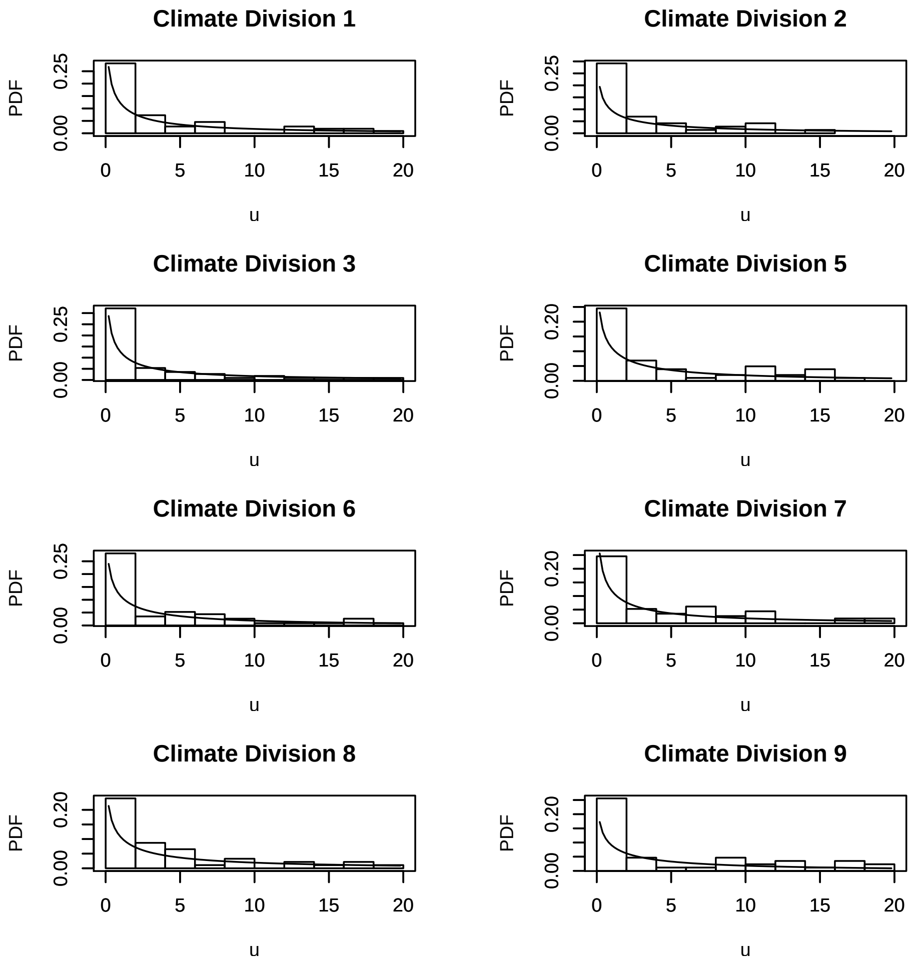

Fig 1. Fitted values of the pdf (3) of $U = XY$ in the eight climate divisions of Nebraska (X = drought intensity and Y = drought duration)

Table 3. Parameter estimates of (1)–(2) for drought duration data divided by 100

Climate division	$\widehat{k}$ ($\widehat{l}$)	$\widehat{c}$ ($\widehat{d}$)
1	15.612	0.932
2	10.944	0.896
3	13.812	0.894
5	15.801	0.949
6	15.726	0.951
7	15.070	0.915
8	14.866	0.926
9	13.015	0.941

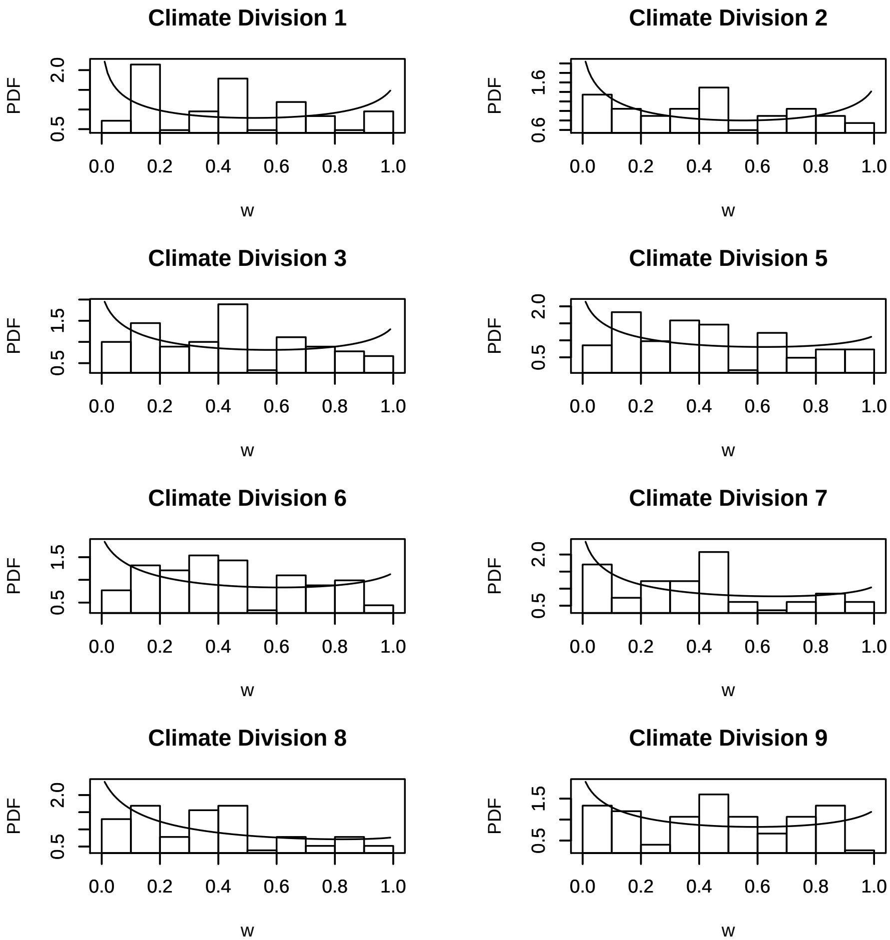

Fig. 2. Fitted values of the pdf (5) of $W = X/(X+Y)$ in the eight climate divisions of Nebraska (X = drought duration and Y = successive non-drought duration)

Table 4. Parameter estimates of (1)–(2) for non-drought duration data divided by 100

Climate division	$\widehat{k}$ ($\widehat{l}$)	$\widehat{c}$ ($\widehat{d}$)
1	8.135	0.768
2	7.847	0.827
3	10.457	0.861
5	8.935	0.846
6	10.885	0.894
7	8.904	0.855
8	9.458	0.943
9	8.901	0.872

Acknowledgments. The authors would like to thank the editor and the referee for carefully reading the paper and for their great help in improving it.

References

J. E. Dennis and R. B. Schnabel (1983), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ.

I. S. Gradshteyn and I. M. Ryzhik (2000), *Table of Integrals, Series, and Products*, 6th ed., Academic Press, San Diego, CA.

R. Ihaka and R. Gentleman (1996), *R: A language for data analysis and graphics*, J. Comput. Graphical Statist. 5, 299–314.

A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev (1986), *Integrals and Series*, Volumes 1, 2 and 3, Gordon and Breach, Amsterdam.

R. N. Rodriguez (1977), *A guide to the Burr type XII distributions*, Biometrics 64, 129–134.

R. B. Schnabel, J. E. Koontz and B. E. Weiss (1985), *A modular system of algorithms for unconstrained minimization*, ACM Trans. Math. Software 11, 419–440.

P. R. Tadikamalla (1980), *A look at the Burr and related distributions*, Internat. Statist. Rev. 48, 337–344.

D. R. Wingo (1983), *Maximum likelihood methods for fitting the Burr type XII distribution to life test data*, Biometrical J. 25, 77–84.

D. R. Wingo (1993), *Maximum likelihood methods for fitting the Burr type XII distribution to multiply* (*progressively*) *censored life test data*, Metrika 40, 203–210.

V. Yevjevich (1967), *An objective approach to definitions and investigations of continental hydrologic droughts*, Hydrologic Paper No. 23, Colorado State Univ., Fort Collins.

School of Mathematics
University of Manchester
Manchester M13 9PL, UK
E-mail: Saralees.Nadarajah@manchester.ac.uk

Department of Statistics
Florida International University
Miami, FL 33199, U.S.A.
E-mail: kibriag@fiu.edu

Received on 22.12.2005;
revised version on 23.6.2006 (1798)