Sci. J. Fac. Sci. Minufia Univ., 2006, Vol. XX, 133 - 147

LARGE–SAMPLE APPROXIMATION TO THE MEAN, VARIANCE AND COVARIANCE OF ORDER STATISTICS FROM BURR TYPE II–DISTRIBUTION

A.A.JAMJOOM
Department of Mathematics, Girls College of Education,
Jeddah, Saudi Arabia

ABSTRACT

In this paper derivation of exact and approximated moments (mean and variance) of the ith order statistics $X_{i:n}$ from burr type II distribution are considered . An approximate formula for the covariance of two order statistics $X_{i:n}$, $X_{j:n}$ from this distribution is also derived . The method used for approximation is series approximation by David and Johnson (1954) . Some basic characteristics and properties of burr type II distribution are also considered . The results obtained here for burr type II distribution are a generalization of the results for logistic distribution given by Arnold and etal (1992).

INTRODUCTION

Evaluation of the exact moments of the ith order statistic X_i from the probability density function requires numerical integration for many F(x) of interest . However, the ith order statistic from any continuous distribution is a function of the ith order statistics from the uniform

distribution because of the probability – integral transformation . Letting U_i denote the ith order statistic from a uniform distribution over the interval (0,1) . This functional relation can be denoted by

$$X_{i:n} = F^{-1}(U_{i:n}), \qquad (1.1)$$

Since the moments of U_i were easily evaluated , an approximation to the moments of a function in terms of some function of these moments would enable one to approximate the moments of X_i for any specified continuous F_x . Such an approximation to the moments of any function in terms of the moments of the argument is referred to as series approximation . David and Johnson (1954) have given series approximations for the first four cumulant & cross - cumulant of order statistics. These series approximations are given in inverse powers of n + 2 merely for the simplicity and computational ease. Clark and Williams (1958) have developed series approximations similar to those of David and Johnson (1954) by making use of the exact expressions of the central moments of uniform order statistics where the k th central moments is of order $\{(n + 2)\ (n + 3)...\ (n + k)\}^{-1}$ instead of in inverse powers of (n + 2). These developments have been discussed in great detail by David (1981) and Arnold , Balakrishnan and Nagaraga (1992). As pointed out by the last two authors that series approximation can be used in case

of distributions with F^{-1} (u) being available explicitly or not. Application to the Logistic and Normal distributions were considered by Arnold ,Balakrishnan and Nagaraga (1992).

SERIES APPROXIMATION

Letting U_i denote the ith order statistic from a uniform distribution over the interval (0,1) . This functional relation can be denoted by

$$X_{i:n} = F^{-1}(u_{i:n}) \qquad (2.1)$$

and

$$(X_{i:n}, X_{j:n}) \underline{\underline{d}} (F^{-1}(u_{i:n}), F^{-1}(u_{j:n})), \qquad (2.2)$$

Upon expanding $F^{-1}(U_{i:n})$ in a Taylor series around the point $E(U_{i:n}) = i/(n+1) = p_i$,

we get a series expansion for $X_{i:n}$ from (2.1) as

$$X_{i:n} = F^{-1}(p_i) + F^{-1(1)}(p_i)(u_{i:n} - p_i) + \frac{1}{2}F^{-1(2)}(p_i)(u_{i:n} - p_i)^2$$

$$+ \frac{1}{6}F^{-1(3)}(p_i)(u_{i:n} - p_i)^3 + \frac{1}{24}F^{-1(4)}(p_i)(u_{i:n} - p_i)^4 + \ldots\ldots \qquad (2.3)$$

where $F^{-1(1)}(p_i)$, $F^{-1(2)}(p_i)$, $F^{-1(3)}(p_i)$, $F^{-1(4)}(p_i)$, ... are the successive derivatives of $F^{-1}(u)$ evaluated at $u = p_i$.

Now by taking expectation on both sides of (2.3) and using the expressions of the central moments of uniform order statistics [however, written in inverse powers of $n + 2$ by David and Johnson (1954) for computational ease and algebraic simplicity], we derive

$$\mu_{i:n} \approx F^{-1}(p_i) + \frac{p_i q_i}{2(n+2)} F^{-1(2)}(p_i)$$

$$+ \frac{p_i q_i}{(n+2)^2}\left[\frac{1}{3}(q_i - p_i)F^{-1(3)}(p_i) + \frac{1}{8}p_i q_i F^{-1(4)}(p_i)\right]$$

$$+ \frac{p_i q_i}{(n+2)^3}[-\frac{1}{3}(q_i - p_i)F^{-1(3)}(p_i) + \frac{1}{4}\{(q_i - p_i)^2 - p_i q_i\}F^{-1(4)}(p_i)$$

$$+ \frac{1}{6}p_i q_i (q_i - p_i)F^{-1(5)}(p_i) + \frac{1}{48}p_i^2 q_i^2 F^{-1(6)}(p_i)]\,, \qquad (2.4)$$

where $q_i = 1 - p_i = (n - i + 1) / (n + 1)$.

Similarly, by taking expectation on both sides of the series expansion for $X_{i:n}^2$ obtained form (2.3) and then subtracting from it the approximation of the $\mu_{i:n}^2$ obtained from Eqs. (2.4), an approximate formula for the variance of $X_{i:n}$ can be derived as

$$\sigma_{i,i:n} \approx \frac{p_i q_i}{n+2}\{F^{-1(1)}(p_i)\}^2 + \frac{p_i q_i}{(n+2)^2}[2(q_i - p_i)F^{-1(1)}F^{-1(2)}(p_i)$$

$$+ p_i q_i [F^{-1(1)}(p_i)F^{-1(3)}(p_i) + \frac{1}{2}\{F^{-1(2)}(p_i)\}^2]]$$

$$+ \frac{p_i q_i}{(n+2)^3}[-2(q_i - p_i)F^{-1(1)}(p_i)F^{-1(2)}(p_i)$$

$$+ \{(q_i - p_i)^2 - p_i q_i\}[2F^{-1(1)}(p_i)F^{-1(3)}(p_i) + \frac{3}{2}\{F^{-1(2)}(p_i)\}^2]$$

$$+ p_i q_i (q_i - p_i)\{\frac{5}{3}F^{-1(1)}(p_i)F^{-1(4)}(p_i) + 3F^{-1(2)}(p_i)F^{-1(3)}(p_i)\}$$

$$+ \frac{1}{4}p_i^2 q_i^2 [F^{-1(1)}(p_i)F^{-1(5)}(p_i) + 2F^{-1(2)}(p_i)F^{-1(4)}(p_i)$$

$$+ \frac{5}{3}\{F^{-1(3)}(p_i)\}^2]]. \qquad (2.5)$$

Proceeding similarly and taking expectation on both sides of the series expansion for $X_{i:n}X_{j:n}$ obtained from (2.3) and then subtracting

from it the approximation for $\mu_{i:n}\mu_{j:n}$ obtained from (2.4), an approximate formula can be derived for the covariance of $X_{i:n}$ *and* $X_{j:n}$ as

$$
\begin{aligned}
\sigma_{i,j:n} \approx\ & \frac{p_i\, q_j}{n+2} F^{-1(1)}(p_i) F^{-1(1)}(p_j) \\
& + \frac{p_i\, q_i}{(n+2)^2}\Big[(q_i - p_i) F^{-1(2)}(p_i) F^{-1(1)}(p_j) \\
& + (q_j - p_j) F^{-1(1)}(p_i) F^{-1(2)}(p_j) \\
& + \frac{1}{2} p_i q_i F^{-1(3)}(p_i) F^{-1(1)}(p_j) + \frac{1}{2} p_j q_j F^{-1(1)}(p_i) F^{-1(3)}(p_j) \\
& + \frac{1}{2} p_i q_j F^{-1(2)}(p_i) F^{-1(2)}(p_j) \Big] \\
& + \frac{p_i\, q_j}{(n+2)^3}\Big[-(q_i - p_i) F^{-1(2)}(p_i) F^{-1(1)}(p_j) \\
& - (q_j - p_j) F^{-1(1)}(p_i) F^{-1(2)}(P_j) \\
& + \{(q_i - p_i)^2 - p_i q_i\} F^{-1(3)}(p_i) F^{-1(1)}(p_j) \\
& + \{(q_j - p_j)^2 - p_j q_j\} F^{-1(1)}(p_i) F^{-1(3)}(p_j) \\
& + \{\frac{3}{2}(q_i - p_i)(q_j - p_j) + \frac{1}{2} p_j q_i - 2 p_i q_j\} F^{-1(2)}(p_i) F^{-1(2)}(p_j) \\
& + \frac{5}{6} p_i q_i (q_i - p_i) F^{-1(4)}(p_i) F^{-1(1)}(p_j) \\
& + \frac{5}{6} p_j q_j (q_j - p_j) F^{-1(1)}(p_i) F^{-1(4)}(p_j) \\
& + \{p_i q_j (q_i - p_i) + \frac{1}{2} p_i q_i (q_j - p_j)\} F^{-1(3)}(p_i) F^{-1(2)}(p_j) \\
& + \{p_i q_j (q_j - p_j) + \frac{1}{2} p_j q_j (q_i - p_i)\} F^{-1(2)}(p_i) F^{-1(3)}(p_j) \\
& + \frac{1}{8} p_i^2 q_i^2 F^{-1(5)}(p_i) F^{-1(1)}(p_j) + \frac{1}{8} p_j^2 q_j^2 F^{-1(1)}(p_i) F^{-1(5)}(p_j) \\
& + \frac{1}{4} p_i^2 q_i q_j F^{-1(4)}(p_i) F^{-1(2)}(p_j) + \frac{1}{4} p_i p_j q_j{}^2 F^{-1(2)}(p_i) F^{-1(4)}(p_j) \\
& + \frac{1}{12}\{2 p_i^2 q_j^2 + 3 p_i p_j q_i q_j\} F^{-1(3)}(p_i) F^{-1(3)}(p_j) \Big],
\end{aligned}
\tag{2.6}
$$

The evaluation of the derivatives of $F^{-1}(u)$ is rather straightforward in case of distribution with $F^{-1}(u)$ being available explicitly. Fortunately, the evaluation of the derivatives of $F^{-1}(u)$ is not difficult even in case of distributions with $F^{-1}(u)$ not existing in an explicit form. In this case, by noting that

$$F^{-1(1)}(u) = \frac{d}{du}F^{-1}(u) = \frac{dx}{du} = \frac{1}{(du/dx)} = \frac{1}{f(x)} = \frac{1}{f(F^{-1}(u))}, \qquad (2.7)$$

which is simply the reciprocal of the pdf of the population evaluated at $F^{-1}(u)$, one may derive the higher - order derivatives of $F^{-1}(u)$ without great difficult by successively differentiating the expression of $F^{-1(1)}(u)$ in (2.7).

SOME BASIC CHARACTERISTICS AND PROPERTIES OF BURR TYPE II DISTRIBUTION

The probability density function (pdf) of Burr type II variable X , introduced by Burr (1942) , is given by

$$f(x) = \frac{re^{-x}}{(1+e^{-x})^{r+1}}, \qquad -\infty \langle x \langle \infty, \quad -\infty \langle r \langle \infty \qquad (3.1)$$

and the corresponding cumulative distribution function (cdf) is

$$F(x) = \frac{1}{(1+e^{-x})^{r}}, \qquad -\infty \langle x \langle \infty \qquad (3.2)$$

The pdf and cdf of Burr type II have been graphed in figs. 3.1 and 3.2 respectively for several values of the shape parametr r .

The hazard function of Burr type II variable X is given by

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{re^{-x}\left[(1+e^{-x})^r - 1\right]}{(1+e^{-x})^{2r+1}} \tag{3.3}$$

its graph is shown in Fig. (3.3).

The moment generating function of X is

$$M_X(t) = E(e^{tx}) = r\int_{-\infty}^{\infty} \frac{e^{-x(1-t)}}{(1+e^{-x})^{r+1}}\,dx$$

Substitutıng $u = \frac{1}{1+e^{-x}}$, *weget*

$$M_x(t) = r\int_0^1 u^{r+t-1}(1-u)^{1-t-1}\,du$$

$$= rB(r+t, 1-t)$$

$$= \frac{\Gamma(r+t)\Gamma(1-t)}{\Gamma(r)} \tag{3.4}$$

Hence, the cumulant generating function of X is obtained as

$$K_X(t) = \ln M_X(t) = \ln\Gamma(r+t) + \ln\Gamma(1-t) - \ln\Gamma(r) \tag{3.5}$$

The cumulants of X may be derived from (3.5) upon differentiating with respect to t and setting t to zero. For example, we obtain the mean and variance of Burr type II distribution as

in increasing order of magnitude. Then the density function of $X_{i;n}$ $(1 \le i \le n)$ is

$$f_{i:n}(x_i) = \frac{n!}{(i-1)\,(n-i)!}\left[F(x_i)\right]^{i-1}\left[1-F(x_i)\right]^{n-i} f(x_i)\,, \qquad -\infty \langle x_i \langle \infty \qquad (4.1)$$

then the moment generating function of $X_{i:n}$ $(1 \le i \le n)$ is

$$M_{i:n}(t) = E(e^{tx_{i:n}}) = \frac{r}{B(i, n-i+1)} \int_{-\infty}^{\infty} \frac{e^{-x(1-t)}}{(1+e^{-x})^{ri+1}} \left[1-(1+e^{-x})^{-r}\right]^{n-i} dx$$

Expanding $\{1-(1+e^{-x})^{-r}\}^{n-i}$ binomially , the above equation can be written as

$$M_X(t) = \frac{r}{B(i,n-i+1)} \sum_{k=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-k}$$

$$\int_{-\infty}^{\infty} \frac{e^{-x(1-t)}}{(1+e^{-x})^{r(n-k)}}\; dx \qquad (4.2)$$

Substitutig $u = \frac{1}{1+e^{-x}}$ *in the integral in* (4.2), *we obtain*

$$M_X(T) = \frac{r}{B(i,n-i+1)} \sum_{K=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-K} \int_0^1 u^{r(n-K)+t-1}\,(1-u)^{1-t-1}\; du$$

$$M_x(t) = \frac{r}{B(i,n-i+1)} \sum_{K=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-k} B[(r(n-k)+t), (1-t)] \tag{4.3}$$

$$= \frac{r}{B(i,n-i+1)} \sum_{K=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-k} \frac{\Gamma(r(n-k)+t) \,.\, \Gamma(1+t)}{\Gamma(r(n-k)+1)} \tag{4.4}$$

where B (. , .) and Γ (.) are the usual complete beta and gamma functions respectively. If
we denote the single moments E ($X^*_{i:n}$) by $\alpha^*_{i:n}$ for $1 \le i \le n$, $k \ge 1$, then from the
expression of the moment generating function in (4.4) we obtain the following

$$\alpha^*_{i:n} = \frac{1}{B(i,n-i+1)} \sum_{K=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-K} \frac{1}{n-k} [\psi(r(n-k)) - \Gamma'(1)] \tag{4.5}$$

and

$$\alpha^{*(2)}_{i:n} = \frac{1}{B(i,n-i+1)} \sum_{K=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-k} \frac{1}{n-k} [\Psi' [r(n-k)] + \Psi^2 [r(n-k)] - \Gamma''(1)] \tag{4.6}$$

where

$$\Psi(z) = \frac{d}{dz} \ln \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$$

$$\Psi'(z) = \frac{d^2}{dz^2} \ln \Gamma(z) = \frac{\Gamma''(z)}{\Gamma(z)} - \Psi^2(z)$$

are the digamma and trigamma functions respectively . Thus from (4.5), (4.6) one may compute the variance of $X_{i:n}$, the i th order statistics as

$$\beta_{i,i:n} = Var(X_{i:n}) = \frac{1}{B(i,n-i+1)} \sum_{k=0}^{n-i} \binom{n-1}{k} (-1)^{n-i-k} \frac{1}{n-k} \left[\Psi'[r(n-k)] + \Psi^2[r(n-k)] - \Gamma''(1)\right]$$

$$-\left\{\frac{1}{B(i,n-i+1)} \sum_{k=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-k} \frac{1}{n-k} \left[\Psi(r(n-k) - \Gamma'(1)\right]\right\}^2 \qquad (4.7)$$

Thus, from Eqs. (4.5) and (4.7), one may compute the means and variances of order statistics either by using the extensive tables of digamma and trigamma function prepared by Davis (1935) and Abramowitz and stegun (1965) or by using the algorithms given by Bernardo (1976) and Schneider (1978). We note here that the moment generating function in (4.4) may be used to obtain higher - order single moments also by involving polygamma functions. From (4.4) we obtain the cumulant generating function of $X_{i:n}$ $(1 \leq i \leq n)$:

$$K_{i:n}(t) = \ln M_{i:n}(t)$$

$$= \ln[\frac{r}{B(i,n-i+1)} \sum_{k=0}^{n-i} \binom{n-i}{k} (-1)^{n-i-k} \frac{1}{\Gamma[r(n-k)+1]} [\Gamma'[r(n-k)+t]\Gamma(1-t)$$

$$-\Gamma[r(n-k)+t]\Gamma'(1-t) \ , \qquad (4.8)$$

APPLICATION TO BURR TYPE II DISTRIBUTION

Approximate formulae for the mean and variance of the ith order statistic $X_{i:n}$, and the covariance

of two order statistics $X_{i:n}$ and $X_{j:n}$ from burr type II distribution is now derived using series approximation discribed in section 2 Eqs (2.4), (2.5) and (2.6). The evaluation of the derivatives of F^{-1} (u) is rather straight forward in case of burr type II distributions with F^{-1} (u) being available explicitly given by Equ. (3.2) .Then

$$x = F^{-1}(u) = r\ln u - \ln(1-u^r), \tag{5.1}$$

We then obtain the sixth derivatives of F^{-1} (u) as follows

$$F^{-1(1)}(u) = \frac{r}{u} + \frac{r u^{r-1}}{(1-u^r)}, \tag{5.2}$$

$$F^{-1(2)}(u) = \frac{-r}{u^2} + \frac{r u^{r-2}\left[u^r + r - 1\right]}{(1-u^r)^2}, \tag{5.3}$$

$$F^{-1(3)}(u) = \frac{2r}{u^3} + \frac{1}{(1-u^r)^3}\left\{r^3(u^{r-3}+u^{2r-3}) + 3r^2(u^{2r-3}-u^{r-3}) + 2r(u^{r-3}-2u^{2r-3}+u^{3r-3})\right\}, \tag{5.4}$$

$$\begin{aligned} F^{-1(4)}(u) = \frac{-6r}{u^4} + \frac{1}{(1-u^r)^4}\,[\,& r^4(u^{r-4}+4u^{2r-4}+u^{3r-4}+6r^3(u^{3r-4}-u^{r-4}) \\ & +11r^2(u^{r-4}-2u^{2r-4}+u^{3r-4}) \\ & +6r(3u^{2r-4}-u^{r-4}-3u^{3r-4}+u^{4r-4})\,]\,, \end{aligned} \tag{5.5}$$

$$f^{-1(5)}(u)=\frac{24r}{u^5}+\frac{1}{(1-u^r)^5}\,[r^5(u^{r-5}+11u^{2r-5}+11u^{3r-5}+u^{4r-5})$$
$$-10r^4(u^{r-5}+3u^{2r-5}-3u^{3r-5}-u^{4r-5})$$
$$+35r^3(u^{r-5}-u^{2r-5}-u^{3r-5}+u^{4r-5})$$
$$-50r^2(u^{r-5}-3u^{2r-5}+3u^{3r-5}-u^{4r-5})$$
$$+24r\,(u^{r-5}-4u^{2r-5}+6u^{3r-5}-4u^{4r-5}+u^{5r-5})\,], \qquad (5.6)$$

$$f^{-1(6)}(u)=\frac{-120r}{u^6}+\frac{1}{(1-u^r)^6}\,[\,r^6(u^{r-6}+26u^{2r-6}+66u^{3r-6}+26u^{4r-6}+u^{5r-6})$$
$$-15r^5(u^{r-6}+10u^{2r-6}-10u^{4r-6}-u^{5r-6})+85r^4(u^{r-6}+2u^{2r-6}-6u^{3r-6}+2u^{4r-6}+u^{5r-6})$$
$$-225r^3(u^{r-6}-2u^{2r-6}+2u^{4r-6}-u^{5r-6})$$
$$+r^2\left[274(u^{r-6}-4u^{2r-6}+6u^{3r-6}-4u^{4r-6}+u^{5r-6})-120u^{6r-6}\right]$$
$$-120r\,(u^{r-6}-5u^{2r-6}+10u^{3r-6}-10u^{4r-6}+5u^{5r-6}-u^{6r-6})\,, \qquad (5.7)$$

The above six derivatives of $F^{-1}(u)$ of burr type II distribution are the generalization of those of the logistic distribution obtained by Arnold , Balakrishnan and Nagaraja (1992) . Setting (r = 1) in Eqs (2.5) through (2.7) , we obtain their results .

REFERENCES

Abramowitz, M. And Stegun, I. A. (eds.) (1965). Handbook of Mathematical functions With Formulas, Graphs, and Mathematical Tables. Dover, New York.

Arnold, B.C., Balakrishnan, N., and Nagaraja, H.N. (1992). A First Course in Order Statistics. John Wiley & sons, Inc, New York, 128 - 130.

Balakrishnan, N. (1991). Handbook of The Logistic Distribution. Marcel Dekker, Inc, New York.

Bernardo, J. M. (1976). Psi (digamma) function. Algorithm AS103. Appl. Statist. , 25, 315 - 317.

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

clark , C.E. and Williams, G.T. (1958) . Distribution of members of an ordered sample . Ann . Math . Statist. , 29, 862 - 870.

David, F.N. and Johnson, N.L. (1954). Statistical treatment of censored data, I. Fundamental formulae, Biometrika, 41, 228 - 240.

Davis, H. T. (1935). Tables of the Higher Mathematical Functions, Vols. 1,II. Principia Press, Bloomington.

David, H.A. (1981) . "Order Statistics" . 2nd ed . New York : John Wiley and Sons , Inc .

Schneider, B. E. (1978). Trigamma function, Alogorithm AS 121. Appl. Statist. , 27, 97 - 99.