International Journal of Pure and Applied Mathematics
Volume 30 No. 1 2006, 119-131

ON PRODUCT MOMENTS OF ORDER STATISTICS FROM THE WEIBULL DISTRIBUTION

A.M. John[1], A.J. Watkins[2 §]
[1,2]School of Business and Economics
University of Wales Swansea
Singleton Park, Swansea, SA2 8PP, U.K.
[1]e-mail: a.john1@neath-porttalbot.gov.uk
[2]e-mail: a.watkins@swansea.ac.uk

Abstract: We give a result for general product moments of order statistics in the Weibull distribution, which includes those already available as special cases. We illustrate our result for specific values of sample size and shape parameter, provide displays of these theoretical expectations which illustrate the considerable structure to these moments, and note the scope for numerical approximations to exact results.

AMS Subject Classification: 33C05, 33F05, 62E15
Key Words: hypergeometric functions, order statistics, product moments, Weibull distribution

1. Introduction and Motivation

The Weibull distribution is widely used in the statistical analysis of reliability data, and properties of this distribution are of considerable practical importance. This paper outlines a result for general product moments of order statistics from this distribution; the result includes, as special cases, available results for Negative Exponential, Rayleigh and Weibull distributions. The Weibull cumulative distribution function is given by

$$F(x;\theta,\beta) = 1 - \exp\left\{-\left(\frac{x}{\theta}\right)^{\beta}\right\} \tag{1}$$

for $x \geq 0$, where the positive parameters θ and β are, respectively, scale and

Received: June 9, 2006

§Correspondence author

shape parameters. Equation (1) reduces to the Negative Exponential with $\beta = 1$, and results for this special case are well-known; the most elegant discussions in this case are generally those which exploit the lack-of-memory property of the Negative Exponential. These results also cover, via the link between Negative Exponential and Weibull distributions, the theory needed to obtain the Expected Fisher information for the Weibull distribution; see, for example, Watkins and John [11]. However, the general result, in addition to its intrinsic interest, is required when considering the variance-covariance structure of the maximum likelihood estimators of parameters in incorrectly specified reliability models. As outlined in John [7], this analysis has considerable practical relevance given the widespread use of the Weibull distribution as a standard model for reliability data, even when significant improvements in goodness-of-fit are obtainable using similarly parsimonious and tractable models.

2. Notation

We denote variates in a random sample from (1) by $X_1, X_2, \ldots, X_n$, with corresponding order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$; moments for single order statistics are widely documented, and a full description is given in Balakrishnan and Sultan [4]. We next define special functions appearing in our discussion; Abramowitz and Stegun [1] give further information on and properties of these functions. The Gamma function is defined as

$$\Gamma(a+1) = \int_0^\infty u^a \exp(-u)\, du,$$

while the incomplete Gamma function is

$$\gamma(a+1, x) = \int_0^x u^a \exp(-u)\, du = x^{a+1} \sum_{m=0}^{\infty} \frac{(-x)^m}{(a+1+m)\, m!},$$

on exploiting a connection with confluent hypergeometric functions. More generally, hypergeometric functions are defined as

$$F_{p,q}(a_1, a_2, .., a_p; b_1, b_2, ..., b_q; z) = \sum_{m=0}^{\infty} \frac{(a_1)_m (a_2)_m \cdots (a_p)_m}{(b_1)_m (b_2)_m \cdots (b_q)_m} \frac{z^m}{m!},$$

where

$$(a)_m = \frac{\Gamma(a+m)}{\Gamma(a)}$$

is Pochhammer's symbol. We will need linear transformation formula for $F_{2,1}$ and the relation between $F_{2,1}$ and the incomplete Beta function $B_\rho(a,b)$, defined as

$$B_\rho(a,b) = \int_0^\rho t^{a-1}(1-t)^{b-1}\,dt;$$

see, respectively, paragraphs 15.3.4 and 6.6.8 in Abramowitz and Stegun [1]. Finally, it is also convenient to recall here the function

$$\begin{aligned} H(a,b) &\equiv \int_{u=0}^{\infty}\int_{t=0}^{u} t^2u^2\exp\left(-\frac{at^2}{2}\right)\exp\left(-\frac{bu^2}{2}\right)dtdu \\ &= (ab)^{-\frac{3}{2}}\left[\frac{\pi}{2} - \arctan\left\{\left(\frac{b}{a}\right)^{\frac{1}{2}}\right\} + \frac{(ab)^{\frac{1}{2}}(a-b)}{(a+b)^2}\right] \end{aligned}$$

introduced by Dyer and Whisenand [6] for positive a, b, in connection with the Rayleigh distribution; the present discussion generalises this result.

2.1. Previous Results

Wherever appropriate, we take $\theta = 1$ in (1); this reduces the distribution to standard form, and results for the general case can be found by appropriate re-scaling.

— Negative Exponential $\beta = 1$. Obtaining product moments is greatly simplified for the Negative Exponential distribution (in which we take $\beta = 1$ in (1)), since we can exploit the lack-of-memory property of that distribution. Thus, for the standard Negative Exponential distribution, we have (for $1 \le i < j \le n$)

$$\begin{aligned} E\left[X_{(i)}X_{(j)}\right] &= \mathrm{Var}\,,\left(X_{(i)}\right) + E\left[X_{(i)}\right]\times E\left[X_{(j)}\right] \\ &= \sum_{k=1}^{i}\frac{1}{(n+1-k)^2} + \left[\sum_{k=1}^{i}\frac{1}{n+1-k}\right]\left[\sum_{k=1}^{j}\frac{1}{n+1-k}\right]; \qquad (2) \end{aligned}$$

see, for example, Basu and Singh [4].

— Rayleigh $\beta = 2$. Dyer and Whisenand [6] consider the simplifications possible for the Rayleigh distribution, where $\beta = 2$ in (1); these use the expression for $H(a,b)$ above, which, in turn, is based on an integral formula in Ng and Geller [10]. That integral formula appears in a discussion concerned with the standard Normal distribution, which explains their choice of $\theta = \sqrt{2}$ in the integrand.

— General β. Finally, we note that Lieblein [9] considers the important case $E\left[X_{(i)}X_{(j)}\right]$, proving that this expectation is

$$c_{ij:n}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}(-1)^{k+l}\binom{i-1}{k}\binom{j-i-1}{l} \times \psi\left(j-i+k-l, n-j+l+1\right), \quad (3)$$

where

$$c_{i,j:n}=\frac{\Gamma\left(n+1\right)}{\Gamma\left(i\right)\times\Gamma\left(j-i\right)\times\Gamma\left(n-j+1\right)}$$

and

$$\psi\left(t,u\right)=\left(tu\right)^{-r}\times\Gamma\left(2r\right)\times B_{\rho}\left(r,r\right),$$

in which

$$r=1+\frac{1}{\beta},\quad \rho=\frac{t}{t+u}.$$

We present here an explicit form for $E\left[X_{(i)}^{p}X_{(j)}^{q}\right]$ for the Weibull distribution that contains the above results as special cases, and which can be computed using most statistical software.

3. Preliminary Lemma

We start by deriving a useful result with direct applications to the problem under consideration. The lemma can be adapted to obtain similar functions for other reliability distributions, as discussed elsewhere. We define, for arbitrary positive β, a, b, p and q,

$$I_{p,q}\left(a,b,\beta\right)=\int_{u=0}^{\infty}\int_{t=0}^{u}t^{p+\beta-1}u^{q+\beta-1}\exp\left(-at^{\beta}\right)\exp\left(-bu^{\beta}\right)dtdu,$$

and, with

$$p'=\frac{p}{\beta}+1,\quad q'=\frac{q}{\beta}+1,$$

now prove the following.

Lemma 1. *We have*

$$I_{p,q}\left(a,b,\beta\right)=\frac{\Gamma\left(p'+q'\right)}{\beta^{2}p'b^{p'+q'}}\times F_{2,1}\left(p',p'+q';p'+1;-\frac{a}{b}\right).$$

Proof. We write

$$I_{p,q}\left(a,b,\beta\right)=\int_{u=0}^{\infty}u^{q+\beta-1}\exp\left(-bu^{\beta}\right)\left\{\int_{t=0}^{u}t^{p+\beta-1}\exp\left(-at^{\beta}\right)dt\right\}du,$$

in which

$$\int_{t=0}^{u}t^{p+\beta-1}\exp\left(-at^{\beta}\right)dt=\frac{1}{\beta a^{p'}}\gamma\left(p',au^{\beta}\right).$$

Thus, we have

$$I_{p,q}\left(a,b,\beta\right)=\frac{1}{\beta a^{p'}}\int_{u=0}^{\infty}u^{q+\beta-1}\exp\left(-bu^{\beta}\right)\gamma\left(p',au^{\beta}\right)du$$

in which we now write

$$\gamma\left(p',au^{\beta}\right)=a^{p'}u^{p+\beta}\sum_{m=0}^{\infty}\frac{\left(-a\right)^{m}u^{m\beta}}{m!\left(p'+m\right)},$$

so that

$$\begin{aligned}I_{p,q}\left(a,b,\beta\right)&=\frac{1}{\beta}\int_{u=0}^{\infty}u^{p+q+2\beta-1}\exp\left(-bu^{\beta}\right)\left\{\sum_{m=0}^{\infty}\frac{\left(-a\right)^{m}u^{m\beta}}{m!\left(p'+m\right)}\right\}du\\&=\frac{1}{\beta}\sum_{m=0}^{\infty}\left[\frac{\left(-a\right)^{m}}{m!\left(p'+m\right)}\int_{u=0}^{\infty}u^{p+q+(m+2)\beta-1}\exp\left(-bu^{\beta}\right)du\right]\end{aligned}$$

on reversing the order of integration and summation. We thus have

$$\begin{aligned}I_{p,q}\left(a,b,\beta\right)&=\frac{1}{\beta}\sum_{m=0}^{\infty}\left[\frac{\left(-a\right)^{m}}{m!\left(p'+m\right)}\times\frac{\Gamma\left(p'+q'+m\right)}{\beta b^{p'+q'+m}}\right]\\&=\frac{1}{\beta^{2}b^{p'+q'}}\sum_{m=0}^{\infty}\left[\frac{\Gamma\left(p'+q'+m\right)}{\left(p'+m\right)}\times\frac{\left(-\frac{a}{b}\right)^{m}}{m!}\right].\end{aligned}$$

We finally introduce a hypergeometric function by noting that the last summation is

$$\begin{aligned}&\sum_{m=0}^{\infty}\left[\frac{\Gamma\left(p'+m\right)\Gamma\left(p'+q'+m\right)}{\Gamma\left(p'+m+1\right)}\times\frac{\left(-\frac{a}{b}\right)^{m}}{m!}\right]\\&=\frac{\Gamma\left(p'+q'\right)}{p'}\times F_{2,1}\left(p',p'+q';p'+1;-\frac{a}{b}\right),\end{aligned}$$

and hence we obtain

$$I_{p,q}\left(a,b,\beta\right)=\frac{\Gamma\left(p'+q'\right)}{\beta^{2}p'b^{p'+q'}}\times F_{2,1}\left(p',p'+q';p'+1;-\frac{a}{b}\right),$$

as required. □

We remark that this result includes those for $\beta = 1, 2$ as special cases, and note some simplification in the important case $p = q = 1$; we have

$$I_{1,1}(a, b, \beta) = \frac{\Gamma\left(\frac{2}{\beta}+2\right)}{\beta^2\left(\frac{1}{\beta}+1\right)b^{\frac{2}{\beta}+2}} F_{2,1}\left(\frac{1}{\beta}+1, \frac{2}{\beta}+2; \frac{1}{\beta}+2; -\frac{a}{b}\right)$$
$$= \frac{\Gamma\left(\frac{2}{\beta}+1\right)}{\beta^2 a b^{\frac{2}{\beta}+1}}\left[F_{2,1}\left(\frac{1}{\beta}, \frac{2}{\beta}+1; \frac{1}{\beta}+1; -\frac{a}{b}\right) - \left(1+\frac{a}{b}\right)^{-\left(\frac{2}{\beta}+1\right)}\right],$$

on exploiting relations between hypergeometric functions. For $\beta = 1$, we now use

$$F_{2,1}(1, 3; 2; -z) = \frac{1}{2(1+z)} + \frac{1}{2(1+z)^2}$$

to obtain, after some re-arrangement,

$$I_{1,1}(a, b, 1) = \frac{1}{b^2(a+b)^2} + \frac{2}{b(a+b)^3}; \tag{4}$$

for $\beta = 2$, we use

$$F_{2,1}\left(\frac{1}{2}, 2; \frac{3}{2}; -z^2\right) = \frac{1}{2z}\frac{d}{dz}\left[z^2 F_{2,1}\left(\frac{1}{2}, 1; \frac{3}{2}; -z^2\right)\right]$$
$$= \frac{1}{2z}\frac{d}{dz}\left[z^2 \times z^{-1}\arctan z\right] = \frac{1}{2z}\left[\arctan z + \frac{z}{1+z^2}\right],$$

to obtain, again after some simplification,

$$I_{1,1}(a, b, 2) = \frac{1}{8(ab)^{\frac{3}{2}}}\left[\arctan\left\{\left(\frac{a}{b}\right)^{\frac{1}{2}}\right\} + \frac{(ab)^{\frac{1}{2}}(a-b)}{(a+b)^2}\right] = H(2a, 2b), \tag{5}$$

thus taking into account the factor in the denominators in the arguments of the exponentials in H, as we require. The lemma is used in the next section, when we consider the product moments for any pair of order statistics from the standard Weibull distribution H in these summations.

4. The General Result

We now present a general result on the expectations of products of arbitrary powers of $X_{(i)}$ and of $X_{(j)}$, and then use this result to derive numerical values in specific cases. We first make direct use of the definition of joint expectations; with $\theta = 1$ in (1), the joint probability density function of $X_{(i:n)}$ and $X_{(j:n)}$ is

$$c_{i,j:n}\beta^2 t^{\beta-1} u^{\beta-1} \left\{1-\exp\left(-t^{\beta}\right)\right\}^{i-1} \left\{\exp\left(-t^{\beta}\right)-\exp\left(-u^{\beta}\right)\right\}^{j-i-1}$$
$$\times \exp\left\{-\left(t^{\beta}+(n-j+1)\,u^{\beta}\right)\right\}$$

for $0 \le t < u < \infty$; see, for example, Balakrishnan and Sultan [3]. The expectation $E\left[X^{p}_{(i:n)}X^{q}_{(j:n)}\right]$ thus takes the form

$$c_{i,j:n}\beta^2 \int_{u=0}^{\infty}\int_{t=0}^{u} t^{p+\beta-1}u^{q+\beta-1}\left\{1-\exp\left(-t^{\beta}\right)\right\}^{i-1}$$
$$\times \left\{\exp\left(-t^{\beta}\right)-\exp\left(-u^{\beta}\right)\right\}^{j-i-1}\exp\left\{-\left(t^{\beta}+(n-j+1)\,u^{\beta}\right)\right\}dtdu\,.$$

We now expand both brackets inside the integral, and write the expectation as

$$c_{i,j:n}\beta^2 \sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}\binom{i-1}{k}\binom{j-i-1}{l}(-1)^{j-k-l}$$
$$\times \int_{u=0}^{\infty}\int_{t=0}^{u} t^{p+\beta-1}u^{q+\beta-1}\exp\left\{-(i+l-k)\,t^{\beta}\right\}\exp\left\{-(n-i-l)\,u^{\beta}\right\}dtdu$$
$$= c_{i,j:n}\beta^2 \sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}\binom{i-1}{k}\binom{j-i-1}{l}(-1)^{j-k-l}$$
$$\times I_{p,q}\left(i+l-k, n-i-l, \beta\right). \quad (6)$$

Using the lemma, the expectation can also be written as

$$\frac{c_{i,j:n}\Gamma\left(p'+q'\right)}{p'}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}\left[\frac{(-1)^{j-l-k}\binom{i-1}{k}\binom{j-i-1}{l}}{(n-i-l)^{p'+q'}}\right.$$
$$\left. F_{2,1}\left(p', p'+q'; p'+1; -\left\{\frac{i+l-k}{n-i-l}\right\}\right)\right]. \quad (7)$$

We immediately note scope for further simplification, through symmetry in the binomial coefficients, the alternating signs of terms in the summation, and the structure in the arguments of the hypergeometric function, particularly in the important case $p = q = 1$. Unless $i = 1$ or $j = i+1$, when there is reduction to a single summation, (7) involves a double summation. In the next section, we discuss the numerical details of computing such expressions; first, however, we summarise connections with three previous results for $p = q = 1$, noted above.

4.1. Connection with Previous Results for $p = q = 1$

— Negative Exponential $\beta = 1$. Using (6), the expectation becomes

$$c_{i,j:n}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}(-1)^{j-k-l}\binom{i-1}{k}\binom{j-i-1}{l}I_{1,1}(i+l-k,n-i-l,1).$$

$$=c_{i,j:n}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}\left[(-1)^{j-l-k}\binom{i-1}{k}\binom{j-i-1}{l}\times\left\{\begin{array}{c}(n-i-l)^{-2}(n-k)^{-2}\\+2(n-i-l)^{-1}(n-k)^{-3}\end{array}\right\}\right],$$

on exploiting the form of $I_{1,1}(a,b,1)$ in (4). Then, using relations between finite sums of powers of reciprocals with alternating and direct terms - see, for example, John, Johnson and Watkins [8] - this summation reduces to (2) given above.

— Rayleigh $\beta = 2$. Again using (6), the expectation becomes

$$4c_{i,j:n}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}(-1)^{j-k-l}\binom{i-1}{k}\binom{j-i-1}{l}I_{1,1}(i+l-k,n-i-l,2)$$

$$=4c_{i,j:n}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}\left[(-1)^{j-l-k}\binom{i-1}{k}\binom{j-i-1}{l}\times H\left(2\{i+l-k\},2\{n-i-l\}\right)\right]$$

$$=\frac{c_{i,j:n}}{2}\sum_{k=0}^{i-1}\sum_{l=0}^{j-i-1}\left[(-1)^{j-l-k}\binom{i-1}{k}\binom{j-i-1}{l}H(i+l-k,n-i-l)\right]$$

from (5), and using $H(a,b) = 8H(2a,2b)$. This is equivalent to the form given in Dyer and Whisenand [6]; since they consider $\theta = \sqrt{2}$, the divisor of 2 present here is absent from their expression.

— General β. Here, the expression at (7) can be shown to be equivalent to that in Lieblein [9], since we have

$$r = 1 + \frac{1}{\beta} = p'$$

and

$$\frac{\psi(j-i+k-l,n-j+l+1)}{\{(j-i+k-l)(n-j+l+1)\}^{-r}} = \Gamma(2r)\times B_{\frac{j-i+k-l}{n-i+k+1}}(r,r).$$

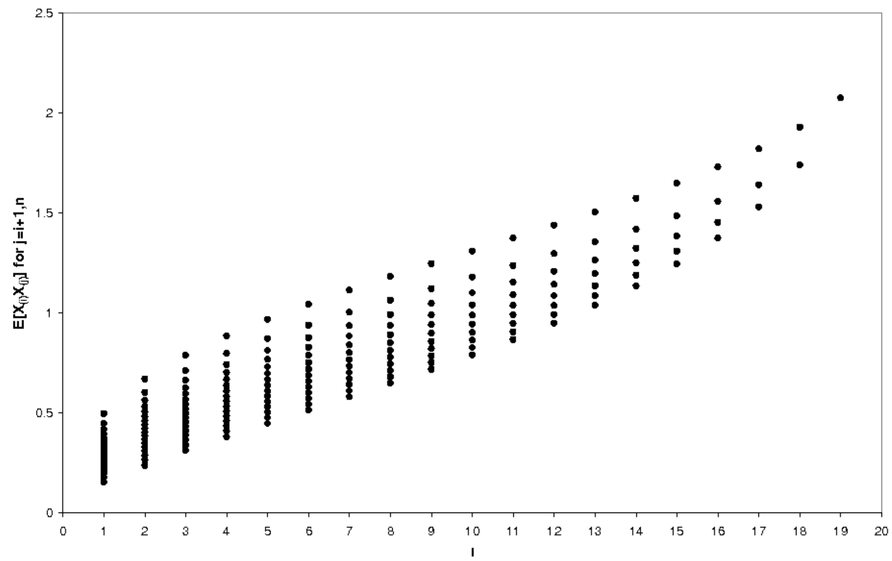

Figure 1: Theoretical values for $E\left[X_{(i)}X_{(j)}\right]$ for $n = 20, \beta = 3$

Then, using paragraphs 6.6.8 and 15.3.4 in Abramowitz and Stegun [1], we write

$$\begin{aligned} B_{\frac{j-i+k-l}{n-i+k+1}}(r,r) &= \frac{1}{r}\left[\frac{j-i+k-l}{n-i+k+1}\right]^r F_{2,1}\left(r, 1-r; r+1; \frac{j-i+k-l}{n-i+k+1}\right) \\ &= \frac{1}{r}\left[\frac{j-i+k-l}{n-j+l+1}\right]^r F_{2,1}\left(r, 2r; r+1; -\left\{\frac{j-i+k-l}{n-j+l+1}\right\}\right), \end{aligned}$$

and, on substituting and simplifying, it can be seen that the two summations (3) and (7) contain the same terms. We next consider briefly some numerical details behind calculating product moments of Weibull order statistics.

5. Some Numerical Details and Discussion

We have given a general expression for expectations of arbitrary products of order statistics for the widely-used Weibull distribution, and shown that this includes, as special cases, some previous results. For illustration, we consider $n = 20$ and $\beta = 3$, and use a straightforward *Mathematica* notebook (available from the authors) to calculate the theoretical expectations with relative ease; these are displayed in Figure 1. This set of theoretical values satisfies the

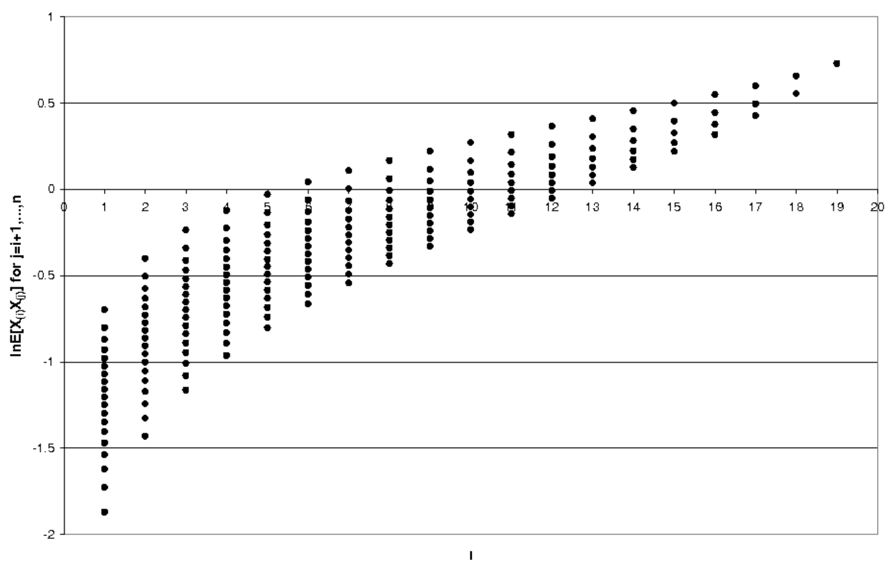

Figure 2: $\ln\left(E\left[X_{(i)}X_{(j)}\right]\right)$ for $n = 20, \beta = 3$

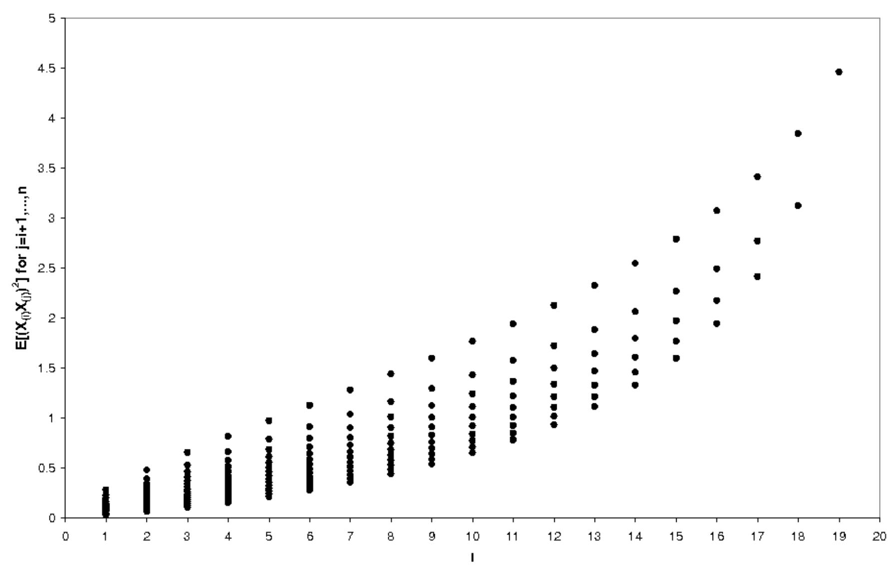

Figure 3: Theoretical values for $E\left[\left\{X_{(i)}X_{(j)}\right\}^2\right]$ for $n = 20, \beta = 3$

various omnibus tests outlined in David and Nagaraja [5], while the accuracy

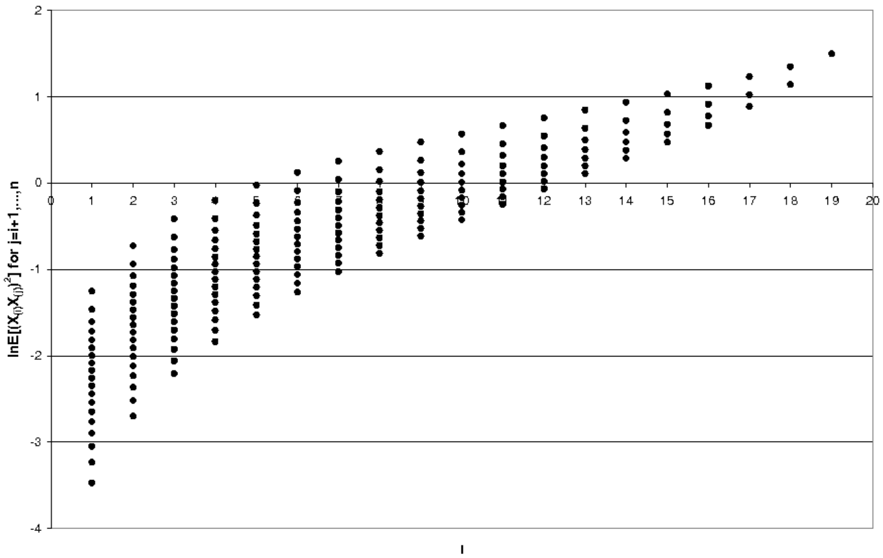

Figure 4: $\ln\left(E\left[\left\{X_{(i)}X_{(j)}\right\}^2\right]\right)$ for $n = 20, \beta = 3$

of individual values can, if necessary, be confirmed by simulation experiments. Figure 2 shows the natural logarithms of the theoretical product moments for all i and j for $n = 20, \beta = 3$; this transformation separates out the original values shown in Figure 1. Figure 2 also indicates that it may be possible to obtain an accurate summary of the structure in the logarithms of the expectations in a straightforward parametric form, which would then form the basis of a suitable numerical approximation to these expectations. Such approximations are clearly useful, as the computational burden associated with (7) increases with n; however, we have successfully calculated expectations from (7) for a wide range of β for $n \leq 500$.

Figures 3 and 4 are the corresponding plots for $n = 20, \beta = 3$ for $p = q = 2$, and show behaviour similar to that observed with $p = q = 1$. Further plots for different n and β confirm this general behaviour. Finally, we remark that an alternative approach to the direct evaluation considered above uses a recurrence relationship in Balakrishnan and Rao [2] to express expectations for $X_{(i)}$ and $X_{(j)}$ in a sample size n in terms of expectations in $X_{(1)}$ and $X_{(j)}$ for sample sizes $\leq n$. This result, which is independent of the underlying distribution, is clearly useful here, since (7) reduces to a single summation when $i = 1$. A detailed comparison between the two approaches will be given elsewhere.agreement for various combinations of sample size n, values of the shape parameter β, and the censoring number m.

References

[1] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, John Wiley and Sons, New York (1972).

[2] N. Balakrishnan, C.R. Rao, Order statistics: An introduction, In: *Handbook of Statistics Volume 16: Order Statistics: Theory and Methods* (Ed-s: N. Balakrishnan, C.R. Rao), Elsevier Science, Amsterdam (1998), 3-24.

[3] N. Balakrishnan, K.S. Sultan, Recurrence relations and identities for moments of order statistics, In: *Handbook of Statistics Volume 16: Order Statistics: Theory and Methods* (Ed-s: N. Balakrishnan, C.R. Rao), Elsevier Science, Amsterdam (1998), 3-24.

[4] A.P. Basu, B. Singh, Order statistics in exponential distribution, In: *Handbook of Statistics Volume 17: Order Statistics: Applications* (Ed-s: N. Balakrishnan, C.R. Rao), Elsevier Science, Amsterdam (1998), 3-24.

[5] H.A. David, H.N. Nagaraja, *Order Statistics*, Third Edition, John Wiley and Sons, New York (2003).

[6] D.D. Dyer, C.W. Whisenand, Best linear unbiased estimator of the parameter of the rayleigh distribution - part I: Small sample theory for censored order statistics, *IEEE Transactions on Reliability*, **22** (1973), 27-34.

[7] A.M. John, *Maximum Likelihood Estimation In Mis-specified Reliability Distributions,* Ph.D. Thesis, University of Wales Swansea (2004).

[8] A.M. John, R. Johnson, A.J. Watkins, A.J. On finite sums of powers of reciprocals, *International Journal of Pure and Applied Mathematics*, **7** (2003), 7-17.

[9] J. Lieblein, On moments of order statistics from the Weibull distribution, *Annals of Mathematical Statistics*, **26** (1955), 330-333.

[10] E.W. Ng, M. Geller, A table of integrals of the error functions, *Journal of Research of the National Bureau of Standards - B. Mathematical Sciences*, **73B** (1969), 1-20.

[11] A.J. Watkins, A.M. John, On the expected fisher information for the Weibull distribution with type II censored data, *International Journal of Pure and Applied Mathematics,* **26** (2006), 93-106.