SOME DISTRIBUTIONS OF ORDERED RANDOM INTERVALS WITH APPLICATIONS

D. G. KABE

(Received May 6, 1968; revised Jan. 27, 1970)

1. Introduction and summary

Let $0<\theta_{(1)}<\theta_{(2)}<\cdots<\theta_{(N-1)}<1$ be $N-1$ random variates from the uniform distribution $f(\theta)=1$, $0<\theta<1$. Then it is known that the joint distribution of the N random (intervals) variates $U_j=\theta_{(j)}-\theta_{(j-1)}$, $j=1, \cdots, N$; $\theta_{(N)}=1$, $\theta_{(0)}=0$, is a Dirichlet's distribution. Further, if $0<U_{(1)}<U_{(2)}<\cdots<U_{(N-1)}<1$ are the ordered U's, then it may be shown that, see e.g., Karlin ([2], p. 263, example 7), the distributions of the ordered U's are identical with those of the distributions of the reduced exponential order statistics. Laurent [3] uses this result implicitly to derive the distributions of the reduced ith order exponential statistic. The implicit argument contained in Laurent's paper is this: the probability that at least k values of the U's exceed a specified constant α, i.e., $\mathrm{P}\{U_{(N-k+1)}>\alpha\}$, is the probability that the $(N-k+1)$th reduced exponential order statistic exceeds α. Laurent did not bring out this argument explicitly and hence his derivation of the distribution of the reduced ith order exponential statistic appears to be involved. Further, it may be shown that any one of the U's and hence any one of the reduced exponential statistic has a beta distribution. This beta distribution is termed by Laurent as an exponential analog of Thompson's [7] known symmetrical beta distribution. Our main purpose in this paper is to explicitly bring out the relation between the distribution theory of reduced exponential order statistics and the order U variates, implicitly contained in Laurent's paper. Note that the distribution theory of U's, i.e. of rectangular ordered variates may be studied by studying the distribution theory of reduced exponential variates. We show that it is easier to study first the distributions of reduced exponential order statistics and then to study the distribution of ordered U's. Thus our derivations of the distributions of ordered U's are simpler than the ones available in the literature, see e.g., Aoyama [1], and the references in his paper. Finally we consider a few applications of our results to survivor probability estimation problems of Failure Theory.

2. Distribution of $U_{(i)}$

It is known that the joint distribution of U's is

$$(1)\qquad f(U_1,\cdots,U_{N-1})=(N-1)!\,,\qquad 0<\sum U_i<1\,.$$

The equation (1) is a particular case of the Dirichlet's distribution, Wilks ([8], p. 177)

$$(2)\qquad f(z_1,\cdots,z_k)=\frac{\Gamma(\alpha_1+\cdots+\alpha_k+\beta)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)\Gamma(\beta)}z_1^{\alpha_1-1}\cdots z_k^{\alpha_k-1}(1-\sum z_k)^{\beta}\,,$$

where $0<z_i<1$, $\sum z_i<1$. It follows from (1) that

$$(3)\qquad f(U_i)=(N-1)(1-U_i)^{N-2}\,,\qquad i=1,\cdots,N-1\,.$$

Thus $\mathrm{P}\{U_i>\alpha\}=(1-\alpha)^{N-1}$. Similarly the probability that a specified j of the U's each exceed α is $S_j=(1-j\alpha)^{N-1}$. Thus the probability that exactly k of U's each exceed α is $P_{[k]}$ as given by Waring's theorem, Wilks ([8], p. 28, example 1.11)

$$(4)\qquad \begin{aligned} P_{[k]}&=\binom{N}{k}\sum_{j=k}^{N}(-1)^{j-k}\binom{N-k}{j-k}S_j\\ &=\binom{N}{k}\sum_{j=k}^{N}(-1)^{j-k}\binom{N-k}{j-k}(1-j\alpha)^{N-1}\,.\end{aligned}$$

It follows that the probability P_k of at least k of U's each exceed α is

$$(5)\qquad \begin{aligned} P_k&=P_{[k]}+\cdots+P_{[N]}=\mathrm{P}\{U_{(N-k+1)}\geqq\alpha\}\\ &=\binom{N}{k}\sum_{j=k}^{N}(-1)^{j-k}\binom{N-k}{j-k}\left(\frac{k}{j}\right)(1-j\alpha)^{N-1}\,.\end{aligned}$$

Now we shall derive the result (5) by using exponential order statistic theory. Let $x_1, x_2,\cdots,x_N$ be a sample of size N from

$$(6)\qquad f(x)=\exp\{-x\}\,,\qquad 0<x<\infty\,,$$

and let $0<x_{(1)}<x_{(2)}<\cdots<x_{(N)}$ be their ordered values, then it is known that, Karlin ([2], p. 263, example 7), the distributions of the reduced variables $x_{(j)}/S$, $j=1,\cdots,N-1$, $S=x_{(1)}+\cdots+x_{(N)}$ are identical with those of the ordered $U_{(j)}$'s, $j=1,\cdots,N-1$. The transformation

$$(7)\qquad x_{(j)}=\sum_{i=1}^{j}y_i\,,\qquad i,j=1,\cdots,N\,,$$

with Jacobian unity transforms the ordered x variates to the unordered y variates, $0<y_i<\infty$, and the y's are independently exponentially distributed. Now we are interested in finding the distribution of the statistic

$$(8) \qquad W=\frac{x_{(i)}}{x_{(1)}+\cdots+x_{(N)}}=\frac{\sum_{j=1}^{i} y_j}{\sum_{j=1}^{N}(N-j+1)y_j}=\frac{u}{v}, \qquad \text{say},$$

where u and v denote, respectively, the numerator and denominator of the second member on right-hand side of (8). Obviously the joint characteristic function (c.f.) $\phi(it_1, it_2)$ of u and v is

$$(9) \qquad \begin{aligned}\phi(it_1, it_2) &= N!\int_0^\infty\cdots\int_0^\infty \exp\left\{it_1\sum_{j=1}^{i} y_j\right. \\ &\qquad \left.-\sum_{j=1}^{N}(N-j+1)(1-it_2)y_j\right\}dy_1\cdots dy_N \\ &= \{N!/(N-i)!(1-it_2)^{N-i}\} \\ &\quad \cdot\left\{\prod_{k=1}^{i}[(N-k+1)(1-it_2)-it_1]\right\}^{-1} \\ &= \sum_{j=1}^{i}\frac{N!(-1)^{j-1}(1-it_2)^{1-N}}{(N-i)!(j-1)!(i-j)![(N-i+j)(1-it_2)-it_1]}.\end{aligned}$$

On inverting the c.f. (9) we find that

$$(10) \qquad f(u, v)=\sum_{j=1}^{i}\frac{N!(-1)^{j-1}\exp\{-v\}\{v-(N-i+j)u\}^{N-2}}{(N-i)!(N-2)!(i-j)!(j-1)!}.$$

It follows that

$$(11) \qquad \begin{aligned}f(W) &= \sum_{j=1}^{i}\frac{N!(N-1)(-1)^{j-1}(1-(N-i+j)W)^{N-2}}{(N-i)!(j-1)!(i-j)!} \\ &= \sum_{t=N-i+1}^{N}\frac{N!(N-1)(-1)^{t-(N-i+1)}(1-tW)^{N-2}}{(N-i)!(t-(N-i+1))!(N-t)!}.\end{aligned}$$

On setting $N-i+1=k$, and integrating $f(W)$ over the range $W=\alpha$ to $W=t^{-1}$ we find $\mathrm{P}\{W>\alpha\}$ and this result agrees with (5). The equation (11) gives us the density of $U_{(N-k+1)}$, which could also be derived by differentiating (5) with respect to W. The result (11) has been derived by a number of authors, see e.g., Aoyama ([1], p. 244, equation 2), Laurent ([3], p. 656, equation 20) and Likes [5]. Incidentally note that equation (4) provides a solution to a problem in Karlin's book ([2], p. 265, problem 18).

Further, note that $U_i=\theta_{(i)}-\theta_{(i-1)}$ has the same distribution as x_i/S, $i=1,\cdots, N-1$. Now any linear function of ordered θ's is a linear function of U's and hence a linear function of reduced exponential statistics. Hence the distributions of the linear functions of ordered rectangular variates may be easily studied by studying the distribution theory of linear functions of reduced exponential statistics.

3. Simultaneous distribution of $U_{(i)}$ and $U_{(j)}$

Now we have to find the joint density of the ith and jth reduced exponential order statistics. Thus setting

$$(12)\quad \begin{aligned} W_1&=x_{(i)}/S=\sum_{k=1}^{i} y_k\Big/\sum_{k=1}^{N}(N-k+1)y_k=u/v=U_{(i)}\\ W_2&=(x_{(j)}-x_{(i)})/S=\sum_{k=i+1}^{j} y_k\Big/\sum_{k=1}^{N}(N-k+1)y_k=z/v=U_{(j)}-U_{(i)}\end{aligned}$$

we find the joint c.f. $\phi(it_1, it_2, it_3)$ of u, z, and v to be

$$(13)\quad \begin{aligned} &\phi(it_1, it_2, it_3)\\ &=N!\int_0^\infty\cdots\int_0^\infty \exp\Big\{it_1\sum_{k=1}^{i} y_k+it_2\sum_{k=i+1}^{j} y_k\\ &\qquad -\sum_{k=1}^{N}(N-k+1)(1-it_3)y_k\Big\}dy_1\cdots dy_k\\ &=\{N!/(N-j)!(1-it_3)^{N-j}\}\Big\{\prod_{k=1}^{i}[(N-k+1)(1-it_3)-it_1]^{-1}\Big\}\\ &\quad\cdot\Big\{\prod_{k=i+1}^{j}[(N-k+1)(1-it_3)-it_2]^{-1}\Big\}\\ &=\frac{N!}{(N-j)!}\Big\{\sum_{r=i+1}^{j}\frac{(-1)^{j-r}(1-it_3)^{2-N}}{(r-1-i)!(j-r)![(N-r+1)(1-it_3)-it_2]}\Big\}\\ &\quad\cdot\Big\{\sum_{k=1}^{i}\frac{(-1)^{i-k}}{(k-1)!(i-k)![(N-k+1)(1-it_3)-it_1]}\Big\}.\end{aligned}$$

On inverting the c.f. (13), we have that

$$(14)\quad f(u,z,v)=\sum_{k=1}^{i}\sum_{r=i+1}^{j}\frac{N!(-1)^{j-r+i-k}\exp\{-v\}}{(N-j)!(r-i-1)!(j-r)!}\cdot\frac{[v-(N-k+1)u-(N-r+1)z]^{N-3}}{(k-1)!\Gamma(N-2)(i-k)!}.$$

The density of W_1 and W_2, and hence that of $U_{(i)}$ and $U_{(j)}$ now is

$$(15)\quad \begin{aligned} f(U_{(i)},U_{(j)})&=\sum_{r=i+1}^{j}\sum_{k=1}^{i}\frac{N!(N-1)(N-2)(-1)^{j-r+i-k}}{(N-j)!(r-i-1)!(j-r)!}\\ &\quad\cdot\frac{[1-(r-k)U_{(i)}-(N-r+1)U_{(j)}]^{N-3}}{(k-1)!(i-k)!},\\ &\qquad 0<U_{(i)}<U_{(j)}<1.\end{aligned}$$

On setting $k=i-t$ and $r=j-q$ the equation (15) reduces to

$$(16)\quad f(U_{(i)},U_{(j)})=\sum_{t=0}^{i-1}\sum_{q=0}^{j-i-1}(N-1)(N-2)\binom{N}{j}\binom{j}{i+1}i(i+1)$$

$$\cdot \binom{j-i-1}{q}\binom{i-1}{t}(-1)^{t+q}[1-(j-q-i+t)U_{(i)}$$
$$-(N-j+q+1)U_{(j)}]^{N-3},$$

a result which agrees with the one given by Aoyama ([1], p. 245, equation 3).

Since any linear function of ordered U's is the same linear function of $x_{(j)}/S$, $j=1,\cdots,N$, the distribution of this linear function may be easily obtained by using the method of this section.

4. Application

We shall consider some applications of the distribution theory of $U_{(i)}$ to the survivor probability estimation problem considered by Laurent. For this purpose we use the following result of Patil and Wani's [6]. Let $x_1,\cdots,x_N$ be a random sample of size N from the distribution function $F(x,\theta)$, and let $t(x_1,\cdots,x_N)$ be complete and sufficient statistic for θ. Then $\mathrm{P}\{x_1\geqq\alpha_1,\ x_2\geqq\alpha_2,\cdots,x_k\geqq\alpha_k\,|\,t\}$ is the unbiased minimum variance estimate (UMVE) of the product $\prod_{i=1}^{k}[1-F(\alpha_i,\theta)]$, $k\leqq N$.

In particular if x has the density

$$f(x,m,\sigma)=(1/\sigma)\exp\{-(x-m)/\sigma\}\,, \qquad x>m,\ \sigma>0\,, \tag{17}$$

then we require an UMVE of $(S(\alpha))^k$, where

$$S(\alpha)=\int_\alpha^\infty f(x,m,\sigma)dx\,. \tag{18}$$

If $x_{(1)},x_2,\cdots,x_N$, with $x_{(1)}$ smallest, are N observation from (17), then we know that the pair $(x_{(1)},Y=(x_2+\cdots+x_N-(N-1)x_{(1)})/N)$ is complete and sufficient for the pair (m,σ). The conditional density of $x_2,\cdots,x_N$, given $x_{(1)}$, is

$$f(x_2,\cdots,x_N\,|\,x_{(1)})=N^{-1}\sigma^{1-N}\exp\left\{-\sum_{i=2}^{N}(x_i-x_{(1)})/\sigma\right\}\,. \tag{19}$$

The distribution of the reduced statistics $\delta_j=(x_j-x_{(1)})/NY$, $j=2,\cdots,N$ is identical with that of $(N-1)$ random intervals on a unit line, and hence

$$f(\delta_2,\cdots,\delta_{N-1})=(N-2)!\,, \qquad \sum_{j=1}^{N-1}\delta_j=1\,. \tag{20}$$

It follows that the density of any one δ_j is

$$f(\delta_j)=(N-2)(1-\delta_j)^{N-3}\,, \qquad j=2,\cdots,N\,, \tag{21}$$

or the density of ξ_j, $(\xi_j-x_{(1)})/NY=\delta_j$, $j=2,\cdots,N$ is

$$(22)\qquad f(\xi_j \mid x_{(1)}, Y)=(N-2)(NY-(\xi_j-x_{(1)}))^{N-3}(NY)^{2-N},$$

where $(\xi_j-x_{(1)})\leqq NY$. The density (22) is termed by Laurent as an exponential analogue of Thompson's [7] symmetric beta density. A certain class of statistics associated with exponential type populations and independently distributed of complete and sufficient statistics have beta type distributions, which are termed by Laurent as generalizations of Thompson's distributions. These generalized Thompson's distributions have wide applications in UMVE problems, outliers rejection problems, nuisance parameter removal tests construction problems, and distribution characterization problems. For the exponential population Laurent [3] uses (22) for survivor probability UMVE estimation and also as outlier rejection criterion. The suspected outlier being x, where x is any one of the values $x_2, \cdots, x_N$. Since $(\xi_j-x_{(1)})/NY$ is nuisance parameter σ free, the distribution of $(\xi_j-x_{(1)})/NY$ may be used to test $m=0$, σ unknown. Recently Laurent and Gupta [4] have utilized the independence of $(\xi_j-x_{(1)})/(NY)$, and the pair of complete and sufficient statistics $(x_{(1)}, NY)$ to characterize the exponential population. Since all the statistics used in above problems by Laurent are reduced exponential (ordered or not) statistics, the problems may be solved by using the distribution theory of random intervals. From (22) note that $\mathrm{P}\{\xi_j>\alpha \mid x_{(1)}, Y\}=(1-(\alpha-x_{(1)})/NY)^{N-2}$ is the UMVE of $S(\alpha)$. Similarly the probability $(1-j(\alpha-x_{(1)})/NY)^{N-2}$, that a specified j of ξ's each exceed α, given $x_{(1)}$ and Y, is the UMVE of $(S(\alpha))^j$. It follows that the probability ψ_k that at least k of the random intervals from (21) will each exceed α where, by using (5),

$$(23)\qquad \psi_k=\binom{N-1}{k}\sum_{j=k}^{N-1}(-1)^{j-k}\binom{N-1-k}{j-k}\left(\frac{k}{j}\right)\left(1-\frac{j(\alpha-x_{(1)})}{NY}\right)^{N-2}$$
$$=\mathrm{P}\{\delta_{(N-k+1)}\geqq\alpha\},$$

is the UMVE of the probability that at least k values out N from (17) will each exceed α. Note that $\delta_{(N-k+1)}$ is the $(N-k+1)$th reduced exponential order statistics considered by Laurent.

Similarly by using expressions of the type

$$(24)\qquad R(\alpha, \beta)=\mathrm{P}\{x_{(i)}\geqq\alpha,\ x_{(j)}\leqq\beta\}$$
$$=\sum_{r=0}^{i-1}\sum_{S=0}^{N-j}\frac{N![F(\alpha, \theta)]^r[F(\beta, \theta)-F(\alpha, \theta)]^{N-r-S}[1-F(\beta, \theta)]^S}{r!\,S!\,(N-r-S)!},$$

where $\alpha<\beta$, and $x_{(1)}<x_{(2)}<\cdots<x_{(N)}$ are ordered variables from $F(x, \theta)$, we may be able to derive UMVE for $R(\alpha, \beta)$, provided θ admits a complete and sufficient statistic. Thus using (16) with $i=N-k+1$, $j=N-k+2$, we may give UMVE for the probability that at least k observa-

tions from (17) will each exceed α, while at least $k-1$ of these will be less than β.

ST. MARY'S UNIVERSITY, HALIFAX, N.S., CANADA

REFERENCES

[1] Aoyama, H. (1962). Note on ordered random intervals and its application, *Ann. Inst. Statist. Math.*, **13**, 243-250.

[2] Karlin, S. (1966). *A First Course in Stochastic Processes*, Academic Press, New York.

[3] Laurent, A. G. (1963). Conditional distribution of order statistics and distribution of the reduced *i*th order statistic of the exponential model, *Ann. Math. Statist.*, **34**, 652-657.

[4] Laurent, A. G. and Gupte, R. C. (1969). A characterization of the exponential distribution, (Abstract), *Ann. Math. Statist.*, **40**, 1865.

[5] Likes, J. (1968). Difference of two ordered observations divided by standard deviation in the case of negatively exponential distribution, *Metrika*, **12**, 161-172.

[6] Patil, G. P. and Wani, J. K. (1966). Minimum variance unbiased estimation of the distribution function admitting a sufficient statistic, *Ann. Inst. Statist. Math.*, **18**, 39-47.

[7] Thompson, W. R. (1935). On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation, *Ann. Math. Statist.*, **6**, 214-219.

[8] Wilks, S. S. (1962). *Mathematical Statistics*, John Wiley, New York.