ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVII, 4 (1983)

G. S. LINGAPPAIAH (Montreal)

ON THE LINEAR COMBINATIONS OF SPACINGS AND THE RESTRICTED RANGE IN THE EXPONENTIAL POPULATIONS

We consider k independent exponential populations. From each population a sample is taken which is censored both above and below. The restricted range is defined as the difference between the largest and the smallest available observations in the censored sample. The spacing is taken as the usual difference between two consecutive order statistics. The distribution of the sum of k linear combinations of spacings (LCOS) in k populations and, in particular, the distribution of the sum of k restricted ranges, is obtained. The variance of this sum of LCOS is minimized. Finally, the distribution of the ratios of k linear combinations to a specific linear combination is also obtained. For $k = 2$, the probability integral is evaluated.

1. INTRODUCTION

In this paper, two concepts, restricted range and spacings, are discussed in relation to the exponential distribution. The reason for the discussion as related to the exponential model is that the order statistics in the samples from this population are extensively utilized in life tests. One such case is in the problem of prediction. Lawless [5] uses the order statistics in the case of exponential distribution to predict future order statistics. Lingappaiah [6], [7] utilizes order statistics for the same purpose of prediction in exponential and gamma populations. Kaminsky [3] gives rigorous bounds for the results of Lawless [5] and Lingappaiah [6]. A Bayesian approach to prediction of order statistics is used in [2] and [8]. Now, suppose a sample of size n is drawn from a population, where this sample is censored both above and below by s and r, respectively. Then the *restricted range* R is defined as the difference between the largest and the smallest available observations in this censored sample. That is, $R = x_{(n-s)} - x_{(r+1)} = u_{n-s} - u_{r+1}$, where $x_{(i)}$ denotes the i-th order statistics in this sample of size n. Obviously, if there is no censoring, this reduces to the usual range in the sample, i.e., $u_n - u_1$. Next consider the *spacings* $v_i = u_i - u_{i-1} = x_{(i)} - x_{(i-1)}$. In this paper, these ideas of R and v_i are dealt with in relation to k independent exponential populations. First, a linear combination of spacings (LCOS) is considered for a population and then the distribution of the sum of k LCOS from k populations is obtained. Then the variance of this sum is minimized subject to certain conditions. For $k = 2$, the probability integral is obtained for the case of $\theta_1 = \theta_2$. Then the probability is evaluated for the case of $\theta_1 \neq \theta_2$. Finally, ratios of these LCOS are taken up and the distribution of k linear combinations to a specific linear combination is obtained.

2. MAIN RESULTS

2.1. Distribution of the linear combinations. Let the sample of size n be drawn from the population

$$f(x) = \theta \exp(-\theta x), \quad x > 0,\ \theta > 0. \tag{1}$$

Let this sample be censored above by s and below by r. Define the restricted range as

$$R = u_{n-s} - u_{r+1} = x_{(n-s)} - x_{(r+1)}.$$

Then from the joint density of u_{n-s} and u_{r+1} we obtain the distribution of the restricted range

$$f(R) = \sum_{j=0}^{a} (-1)^j [\theta \exp(-Rb\theta)] \frac{m!}{s!j!(a-j)!}, \tag{2}$$

where $a = n-r-s-2$, $b = s+j+1$, $m = n-r-1$. Expression (2) can also be written as

$$f(R) = \frac{m!}{s!a!} \theta e^{-R(s+1)\theta} (1-e^{-\theta R})^a. \tag{3}$$

Now by (3) we have the characteristic function of R in the form

$$\varphi_R(t) = \frac{m!}{s!} \frac{\Gamma(s+1-it/\theta)}{\Gamma(a+s+2-it/\theta)}. \tag{4}$$

Now, consider the linear combination of the spacings v_i, where $v_i = u_i - u_{i-1}$ with $i = 1, 2, \ldots, n$ and $u_0 = 0$. Put

$$z = \sum_{i=r+2}^{n-s} a_i v_i, \qquad a_i > 0. \tag{5}$$

It is easy to see that z is R if all a_i's are equal to 1. It is well known from [1] that these spacings in the exponential case are independent and v_r has also the exponential distribution

$$f(v_r) = \theta(n-r+1)\exp\left(-(n-r+1)\theta v_r\right). \tag{6}$$

By (5) and (6) we have the characteristic function of z in the form

$$\varphi_z(t) = \left[\prod_{j=r+2}^{n-s}\left(1 - \frac{ita_j}{\theta(n-r+1)}\right)\right]^{-1}. \tag{7}$$

If all a_{ij}'s are equal, then (7) reduces to (4) as it should be, since in this case z reduces to R as mentioned above.

Now, inverting (7), we have the distribution of z as

$$f(z) = \frac{m!}{s!}\sum_{k=0}^{a}\frac{(-1)^a\left[\theta\exp\left(-z\theta(m-k)/C(k)\right)\right][1/C(k)]}{\prod\limits_{\substack{j=0\\ j\neq k}}^{a}[(m-k)C(j)/C(k)-(m-j)]}, \tag{8}$$

where $C(j) = a_{r+j+2}$, $C(k) = a_{r+k+2}$, and $C(j)$ and $C(k)$ are such that no terms in the denominator are zero. In (8), the product term for each k represents the residue at each of $a+1$ poles.

2.2. Distribution of the sum of LCOS. Now consider k independent exponential populations

$$f(x) = \theta_i\exp(-\theta_i x), \qquad x > 0,\ \theta_i > 0,\ i = 1, 2, \ldots, k.$$

Let the sample of size n_i be drawn from the i-th population and let this sample be censored above by s_i and below by r_i. Then put

$$w = z_1 + z_2 + \ldots + z_k.$$

Then from (6) again we have the characteristic of w in the form

$$\varphi_w(t) = \left[\prod_{i=1}^{k}\prod_{j=r_i+2}^{n_i-s_i}\left(1 - \frac{\lambda t a_{ij}}{\theta(n_i-j+1)}\right)\right]^{-1}, \tag{9}$$

where λ is as in $\varphi(t) = \mathrm{E}(\lambda t x)$.

Inverting (9), we have the distribution of w as

$$f(w) = \prod_{i=1}^{k}\left[\frac{m_i!}{s_i!}(-1)^{a_i}(-1)^{k-1}\right]\times \tag{10}$$

$$\times \sum_{i=1}^{k}\sum_{k_i=0}^{a_i}\left[\theta_i \exp\left(\frac{-\theta_i w(m_i-k_i)}{C(k_i)}\right)\right]\times$$

$$\times \prod_{\substack{j=1\\ j\neq i}}^{k}\left(\frac{\theta_j}{\theta_i}\right)^{a_j+1}\left[C(k_i)\prod_{\substack{j_i=0\\ j_i\neq k_i}}^{a_i}\left((m_i-k_i)\frac{C(j_i)}{C(k_i)}-(m_i-j_i)\right)\right]^{-1}\times$$

$$\times\left[\prod_{\substack{t=1\\ t\neq i}}^{k}\prod_{j_t=0}^{a_t}\left((m_i-k_i)\frac{C(j_t)}{C(k_i)}-(m_t-j_t)\frac{\theta_t}{\theta_i}\right)\right]^{-1}.$$

In (10) the quantities under the sums represent the residues at each of $a+k$ poles, where $a = a_1+a_2+\ldots+a_k$ with

$$a_i = n_i-r_i-s_i-2,\qquad m_i = n_i-r_i-1$$

and

$$C(j_t) = a_{r_t+j_t+2},\qquad C(j_i) = a_{r_i+j_i+2},\qquad C(k_i) = a_{r_i+k_i+2}.$$

For example, if $n_1 = 5$, $r_1 = s_1 = 1$ ($m_1 = 3$, $a_1 = 1$) and $n_2 = 7$, $r_2 = 1$, $s_2 = 3$ ($m_2 = 5$, $a_2 = 1$), then we deal with ($a_{ij} = 1$ in z) u_4-u_3 in both the samples 1 and 2, and now (10) takes the form

$$f(w) = -(5!)\times \tag{11}$$

$$\times\left[\left(\frac{\theta_2}{\theta_1}\right)^2\begin{pmatrix}[\theta_1\exp(-3\theta_1 w)][(3-5\theta_2/\theta_1)(3-4\theta_2/\theta_1)]^{-1}\\ [-\theta_1\exp(-2\theta_1 w)][(2-5\theta_2/\theta_1)(2-4\theta_2/\theta_1)]^{-1}\end{pmatrix}+\right.$$

$$\left.+\left(\frac{\theta_1}{\theta_2}\right)^2\begin{pmatrix}[\theta_2\exp(-5\theta_2 w)][(5-2\theta_1/\theta_2)(5-3\theta_1/\theta_2)]^{-1}\\ [-\theta_2\exp(-4\theta_2 w)][(4-2\theta_1/\theta_2)(4-3\theta_1/\theta_2)]^{-1}\end{pmatrix}\right].$$

2.3. Restricted range. Obviously, if all $C(j)$ and $C(k)$ in (10) are equal to 1, then we have the distribution of the sum of k restricted ranges $R_i = u_{n_i-s_i}-u_{r_i+1}$ ($i = 1, 2, \ldots, k$) as

$$f(R_0) = \prod_{i=1}^{k}\left[\frac{m_i!}{a_i!s_i!}(-1)^{a_i}(-1)^{k-1}\right]\times$$

$$\times\sum_{i=1}^{k}\sum_{k_i=0}^{a_i}\left[\binom{a_i}{k_i}(-1)^{k_i}\right]\left[\theta_i\exp\left(-\theta_i R_0(m_i-k_i)\right)\right]\times$$

$$\times\left[\prod_{\substack{j=1\\j\neq i}}^{k}\left(\frac{\theta_j}{\theta_i}\right)^{a_j+1}\right]\left[\prod_{\substack{t=1\\t\neq i}}^{k}\prod_{j_t=0}^{a_t}\left((m_i-k_i)-(m_t-j_t)\frac{\theta_t}{\theta_i}\right)\right]^{-1},$$

where $R_0 = R_1+R_2+\ldots+R_k$.

Now, for the same example used for (11) we get the integral

$$\int_w^\infty f(w,\theta_1=\theta_2)dw=\beta$$

as

$$-20e^{-3w}+10e^{-2w}-4e^{-5w}+15e^{-4w}=\beta$$

which for $w_0 = 3$ gives the value $\beta = .0223$.

Table I gives $P(w \geqslant w_0, \theta_1 \neq \theta_2)$ by using (11).

TABLE I

θ_1/θ_2	0.5	1	3	5	7	9
$P(w \geqslant w_0)$	.0118	.0223	.2046	.5378	.6008	.9202

2.4. Illustrative example. The two samples in Table II are simulated from the life tests following the exponential model given by (1) with average life of $1/\theta_1 = 2000$ hours and $1/\theta_2 = 1000$ hours, respectively.

TABLE II

Sample 1 ($n_1 = 5$)	Sample 2 ($n_2 = 7$)
221.41	1510.54
506.38	5430.71
552.29	3297.11
1095.22	2261.35
942.25	1929.30
	341.34
	1009.31

TABLE III

w_0	$P(w \geqslant w_0)$
500	.97341
1000	.83484
1500	.63607
2000	.44939
2500	.30291
3000	.19815
4000	.08025
5000	.03116

Using (11) with the above values of θ_1 and θ_2 we get

$$P(w \geqslant w_0) = -(32/7)\exp(-3w_0/2000)+5\exp(-w_0/1000)- -(3/7)\exp(-w_0/200)+\exp(-w_0/250). \tag{12}$$

Table III is generated by using (12).

By Table II, from sample 1 we have $z_1 = u_4-u_3 = 389.96$ and from sample 2 we obtain $z_2 = 418.76$. Hence $w = z_1+z_2 = 808.72$, and using (12) we get

$$P(w \geqslant 808.72) = .90. \tag{13}$$

In an actual experiment, we can use the enlarged table which can be generated by using (10) and evaluate the probabilities similar to (13).

3. VARIANCE

Now

$$w = \sum_{i=1}^{k} \sum_{j=r_i+2}^{n_i-s_i} a_{ij} v_j. \tag{14}$$

From (6) and (14) we have

$$\mathrm{E}(w) = \sum_i \sum_j a_{ij}/(n_i - j + 1)\,\theta_i,$$

$$\mathrm{Var}(w) = \sum_i \sum_j a_{ij}^2/(n_i - j + 1)^2\,\theta_i^2. \tag{15}$$

Now, we can minimize $\mathrm{Var}(w)$ by setting

$$\sum_j a_{ij}/(n_i - j + 1) = 1.$$

That is, we minimize

$$\sum_j a_{ij}^2/(n_i - j + 1)^2 - \lambda_i \sum_j a_{ij}/(n_i - j + 1) \qquad (i = 1, 2, \ldots, k),$$

where λ_i is the Lagrange multiplier. Then we have

$$\lambda_i = 2/(n_i - r_i - s_i - 1) \qquad (i = 1, 2, \ldots, k),$$

$$a_{ij} = (n_i - j + 1)/(n_i - r_i - s_i - 1). \tag{16}$$

Finally, from (15) and (16) we obtain

$$\min \mathrm{Var}(w) = \sum_{i=1}^{k} 1/(n_i - r_i - s_i - 1)\,\theta_i^2. \tag{17}$$

Now, if all $r_i = s_i$ are equal to 0, then (14) together with (16) reduces to

$$w_0 = \sum_{i=1}^{k} \sum_{t=2}^{n_i} (n_i - t + 1)(x_{(t)i} - x_{(1)i})/(n_i - 1).$$

Also we have

$$w_0 = \sum_{i=1}^{k} [n_i/(n_i-1)]\,\hat{\theta}_i, \tag{18}$$

where

$$\hat{\theta}_i = \sum_{t=2}^{n_i} (x_{(t)i} - x_{(1)i})/n_i$$

is the sufficient estimate of θ_i in $(1/\theta_i)\exp\left(-(x-a_i)/\theta_i\right)$ as given by Sukhathme [9]. It is also known that $2n(\hat{\theta}_i/\theta_i)$ has χ^2-distribution with $2(n_i-1)$ degrees of freedom. Then we have

$$\mathrm{Var}(\hat{\theta}_i) = (n_i-1)\,\theta_i^2/n_i^2$$

and from (18) we get

$$\mathrm{Var}(w_0) = \sum_{i=1}^{k} 1/(n_i-1)\,\theta_i^2,$$

which concurs with (17) for $r_i = s_i = 0$ $(i = 1, 2, \ldots, k)$.

4. DISTRIBUTION OF RATIOS OF LCOS

THEOREM. *Each of the independent random variables* $z_1, \ldots, z_{l+1}$ *has the density function* (8) *if and only if the random variables* $y_i = z_i/z_{l+1}$, $i = 1, 2, \ldots, l$, *have the distribution*

$$f(y_1, \ldots, y_l) = \left[\prod_{i=1}^{l+1} E_i \sum_{k_i=0}^{a_i} \frac{(-1)^{a_i}\Gamma(l+1)(\theta_1 \ldots \theta_l)T}{[\prod_{j_i}(Q_i)]\,C_i\left(1+\sum_{i=1}^{l} v_i\right)^{l+1}}\right], \tag{19}$$

where $E_i = m_i!/s_i!$, $C_i = C(k_i)$, *and*

$$\prod_{j_i}(Q_i) = \prod_{\substack{j_i=0\\ j_i\neq k}}^{a_i} [d_i C(j_i) - (m_i - j_i)],$$

$$v_i = (\theta_i d_i y_i/\theta_{l+1} d_{l+1}), \qquad T = \theta_{l+1}/(\theta_{l+1} d_{l+1})^{l+1}.$$

$$d_i = (m_i - k_i)/C_i.$$

Proof. Necessity. From (8) we have

$$f(z_1, \ldots, z_{l+1}) = \prod_{i=1}^{l+1} E_i \sum_{k_i} \frac{(-1)^{a_i}\exp(-\theta_i d_i z_i)}{C_i \prod_{j_i}(Q_i)}. \tag{20}$$

Make the transformations

(21) $$y_i = z_i/z_{l+1}, \quad y_{l+1} = z_{l+1}.$$

Then we have the Jacobian y_{l+1}^l of the transformation $|J|$. Now using (21) and $|J|$ in (20) and integrating out y_{l+1}, we get (19).

Sufficiency. Let $\varphi(t_1, \ldots, t_l)$ denote the characteristic function of $\log y_1, \ldots, \log y_l$. Also, let $\varphi_i(t_i)$ be the characteristic function of $\log z_i$, $i = 1, 2, \ldots, l+1$. Now, we use the result of Kotlarski [4] to prove this part of the theorem, according to which the characteristic function of the random variables $\log y_1, \ldots, \log y_l$ determines the characteristic function of $\log z_1, \ldots, \log z_{l+1}$ provided the z_i's are positive. From (19) we have

(22)
$$\varphi(t_1, \ldots, t_l) = \prod_{i=1}^{l+1} E_i \sum_{k_i} (-1)^{a_i} \left(\frac{\theta_{l+1} d_{l+1}}{\theta_i d_i}\right)^{\lambda t_i} \frac{\Gamma(1+\lambda t_i)}{[\prod_{j_i} (Q_i)] d_i C_i} \frac{\Gamma(1-\lambda \sum_{i=1} t_i)}{\Gamma(1+\lambda t_{l+1})},$$

where λ is as in $\varphi(t) = \mathrm{E}(e^{\lambda tx})$.

Now, it is clear from the definition of the y_i's that

$$\varphi(t_1, \ldots, t_l) = \varphi_1(t_1) \ldots \varphi_l(t_l) \varphi_{l+1}\left(-\sum_{i=1}^{l} t_i\right).$$

By (8) we get

(23) $$\varphi_i(t_i) = E_i \sum_{k_i} \frac{(-1)^{a_i} \Gamma(1+\lambda t_i)}{C_i [\prod_{j_i} (Q_i)] \theta_i^{\lambda t_i} d_i^{1+\lambda t_i}}$$

and

(24) $$\varphi_{l+1} - \sum_{i=1}^{l} t_i = E_{l+1} \sum_{k_{l+1}=0}^{a_{l+1}} \frac{(-1)^{a_{l+1}} \Gamma(1-\lambda \sum_{i=1}^{l} t_i)}{C_{l+1} \theta_{l+1}^{-\lambda \Sigma t_i} d_{l+1}^{1-\lambda \Sigma t_i} \prod_{j_{l+1}} (Q_{l+1})}.$$

It is easy to observe that the product of (23) and (24) is equal exactly to (22) and the sufficiency follows. It is also trivial to see that from (22) we obtain

(25) $$\varphi(0, \ldots, 0) = \prod_{i=1}^{l+1} \left[E_i \sum_{k_i} (-1)^{a_i} \left(\left[\prod_{j_i} (Q_i)\right] d_i C_i\right)^{-1}\right]$$

which, of course, is equal to 1. For example, if $C_i = C(j_i) = 1$, then (25) is

$$\prod_{i=1}^{l+1} \frac{E_i}{a_i!} \sum_{k_i} (-1)^{a_i}(-1)^{k_i}\binom{a_i}{k_i} \frac{1}{m_i - k_i} = 1.$$

Incidentally, it is easy to obtain the distribution of $y = y_1 + \ldots + y_l$ since $y = \sum_{i=1}^{l} z_i/z_{l+1}$ and we know the distribution of both $z = \sum z_i$ and z_{l+1} from (8) and (10), respectively. Then we can get the distribution of y directly from (19). However, it seems easier to obtain the result using z and z_{l+1}.

Acknowledgement. The author would like to thank the referee and the editor for the suggestion to include the illustrative example.

References

[1] H. A. David, *Order statistics*, J. Wiley, New York 1970.
[2] I. R. Dunsmore, *The Bayesian predictive distribution in life testing models*, Technometrics 16 (1974), p. 455-460.
[3] K. S. Kaminsky, *Comparison of prediction intervals for failure times when life is exponential*, ibidem 19 (1977), p. 83-86.
[4] I. Kotlarski, *On characterizing the chi-square distribution by the Student law*, J. Amer. Statist. Assoc. 61 (1966), p. 976-980.
[5] J. F. Lawless, *A prediction problem concerning samples from the exponential distribution with applications to life testing*, Technometrics 13 (1971), p. 725-730.
[6] G. S. Lingappaiah, *Prediction in exponential life testing*, Canad. J. Statist. 1 (1973), p. 113-117.
[7] — *Prediction in the samples from the gamma distribution as applied to life testing*, Austral. J. Statist. 16 (1974), p. 30-32.
[8] — *Bayesian approach to the prediction problem in the exponential population*, IEEE Trans. Reliab. R-27 (1978), p. 222-225.
[9] P. V. Sukhathme, *Tests of significance for samples of the chi-square population with two degrees of freedom*, Ann. Eugenics 8 (1937), p. 52-56.

DEPT. OF MATHEMATICS
SIR GEORGE WILLIAMS CAMPUS
CONCORDIA UNIVERSITY
MONTREAL, CANADA

Received on 16. 8. 1978;
revised version on 14. 4. 1980

G. S. LINGAPPAIAH (Montreal)

LINIOWE KOMBINACJE ROZSTAWIEŃ I OGRANICZONY ROZSTĘP W POPULACJACH WYKŁADNICZYCH

STRESZCZENIE

Rozważmy k niezależnych populacji wykładniczych. Z każdej pobiera się próbkę uciętą zarówno od góry, jak i od dołu. *Ograniczony rozstęp* (*restricted range*) definiuje się jako różnicę między największą i najmniejszą zaobserwowaną wartością w próbce uciętej. Przez *rozstawienie* (*spacing*) rozumie się, jak zwykle, różnicę między dwiema kolejnymi statystykami porządkowymi. W pracy podaje się rozkład sumy k liniowych kombinacji rozstawień w k populacjach. Jako wynik szczególny otrzymuje się rozkład sumy k ograniczonych rozstępów. Otrzymano także rozkład ilorazów k liniowych kombinacji względem ustalonej liniowej kombinacji. Dla $k = 2$ obliczono odpowiednie prawdopodobieństwa.