Probability in the Engineering and Informational Sciences, **23**, 2009, 449–455. Printed in the U.S.A.
doi:10.1017/S0269964809000278

ORDER STATISTICS FROM DEPENDENT, COMPLETELY MONOTONE DISTRIBUTIONS

MARK BROWN
Department of Mathematics
The City College
CUNY
New York, NY 10031
E-mail: cybergarf@aol.com

We consider order statistics corresponding to $X_1, \ldots, X_n$, where $X_i \sim \lambda_i^{-1}\mathcal{E}$, $i = 1, \ldots, n, \mathcal{E}_1, \ldots, \mathcal{E}_n$ are independent and identically distributed exponentials with mean 1, and $\lambda_1, \ldots, \lambda_n$ are possibly dependent, possibly nonidentically distributed, positive random variables, with $E\lambda_i^{-1} < \infty$. Thus, $\lambda_1, \ldots, \lambda_n$, can be interpreted as random failure rates, and their dependency might be due to common environmental factors.

1. INTRODUCTION

Consider random variables, $X_1, \ldots, X_n$, where $X_i \sim \lambda_i^{-1}\mathcal{E}_i, i = 1, \ldots, n$, the $\mathcal{E}_i$'s being independent and identically distributed (i.i.d.) exponentially with mean 1, and $\lambda_1, \ldots, \lambda_n$ are positive random variables, not necessarily independent nor identically distributed. Thus, each X_i is a mixture of exponentials; the X_i are possibly dependent, but given, $\underline{\lambda} = \underline{\lambda}_0, X_1, \ldots, X_n$ are conditionally independent and expontentially distributed. The dependencies between the X's, embodied in the distribution of $\underline{\lambda}$, can be interpreted as the effects of their common environment.

Define Y_r to be the rth-order statistics from $X_1, \ldots\ldots, X_n$. Y_r represents the time to system failure for an $n - r + 1$ out of n system, with dependent components whose distributions are completely monotone. For $r = 1$, Y_1 is the time to failure for a series system and also the time to death or failure in a competing risk model (see Marshall and Olkin [3, pp. 541–559] for competing risks and [3, pp. 137–182] for time to failure for coherent systems). For $r = 1$, we derive the following comparison

© 2009 Cambridge University Press 0269-9648/09 $25.00
https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press

result (Section 4). Define, $Z_i = (E\lambda_i^{-1})\mathcal{E}_i, i = 1, \ldots, n$. The Z_i's are independent exponentially distributed random variables, with $EZ_i = EX_i, i = 1, \ldots, n$. Then,

$$E(\min X_i) \leq E(\min Z_i) = \left(\sum_1^n (E\lambda_i^{-1})^{-1} \right)^{-1}. \tag{1}$$

Thus, one effect of randomizing the failure rates while keeping each mean constant is to decrease the expected value of the minimum. It is not the case that $\min X_i$ is stochastically smaller than $\min Z_i$. Result (1) is known in the case that $\lambda_1, \ldots, \lambda_n$ are i.i.d. [1, p. 266].

For the case $n = 2$, it immediately follows from (1) that

$$E(\max X_i) \geq E(\max Z_i). \tag{2}$$

Thus, for a parallel system of two components, randomizing the failure rates while keeping each mean constant results in a larger mean survival time.

For general n, I do not know whether (2) always holds. It is known to hold when $\lambda_1, \ldots, \lambda_n$ are i.i.d. [1, p. 266]. In Section 5 (Proposition 1), (2) is proved under the condition, $E\lambda_i^{-1} \equiv c$. This requires that $X_1, \ldots, X_n$ have a common mean, but they still can be dependent, with different distributions. I pose the problem of determing whether (2) is true in the general case.

Ross [5, pp. 313–316] has pointed out that in the famous coupon collecting problem, the mean and variance of the number of coupons required until one collects at least one coupon from each of r different types (where there are n coupon types in all) are simply related to the mean and variance of Y_r, for the case of pure exponentials with $\lambda_i = p_i, i = 1, \ldots, n$ (p_i is the probability of coupon type i). We obtain the survival function of Y_r in Section 2, and the coupon mean and variance in Section 2.

2. DISTRIBUTION OF Y_r

We will show that

$$P_r(Y_r > t) = \sum_{w=n-r+1}^{n} (-1)^{w-(n-r+1)} \binom{w-1}{n-r} B_w^{(t)}, \tag{3}$$

where $B_w^{(t)} = \sum_{\alpha \epsilon C_w} Ł_\alpha(t)$, where C_w is the collection of subsets of size w from $(1, \ldots, n)$ and $Ł_\alpha$ is the Laplace transform of the distribution of $S_\alpha = \sum_{i \epsilon \alpha} \lambda_i$.

For example, if $n = 12$ and $r = 7$, then

$$\begin{aligned} P_r(Y_1 > t) &= \sum_{w=6}^{12} (-1)^{w-6} \binom{w-1}{5} B_w(t) \\ &= B_6(t) - 6B_7(t) + 21B_8(t) - 56B_9(t) + 126B_{10}(t) - 252B_{11}^{(t)} + 462B_{12}^{(t)}. \end{aligned}$$

https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press

To derive (3) first note that

$$\Pr(Y_r > t|\underline{\lambda}) = \sum_{l=0}^{r-1} \sum_{\gamma \epsilon C_l} \left[\prod_{\gamma} (1 - e^{-\lambda_i t}) \right] \left(\prod_{j \epsilon \gamma^c} e^{-\lambda_j t} \right)$$

$$= \sum_{l=0}^{r-1} \sum_{\gamma \epsilon C_l} \sum_{\beta c \gamma} (-1)^{|\beta|} e^{-(S_{\beta \cup \gamma^c})t}, \tag{4}$$

where $S_{\beta \cup \gamma^c} = \sum_{i \epsilon \beta} \lambda_i + \sum_{i \epsilon \gamma^c} \lambda_i$ and $|\beta|$ is the number of elements in β.

Gathering together the various (γ^c, β) with common $\alpha = \beta \cup \gamma^c$,

$$\Pr(Y_r > t|\underline{\lambda}) = \sum_{w=n-r+1}^{n} a_w \tilde{B}_w, \tag{5}$$

where

$$\tilde{B}_w = \sum_{\alpha \epsilon C_w} e^{-S_\alpha t}$$

and

$$a_w = \sum_{l=n-w}^{r-1} (-1)^{w-(n-l)} \binom{w}{n-l}$$

$$= \sum_{l=n-w}^{r-1} (-1)^{w-(n-l)} \left[\binom{w-1}{n-l} + \binom{w-1}{n-l-1} \right]$$

$$= (-1)^{w-(n-r+1)} \binom{w-1}{n-r}.$$

Finally, since $E\tilde{B}_w = B_w$,

$$P_r(Y_r > t) = EP_r(Y_r > t|\underline{\lambda}) = E \sum_{n-r+1}^{n} a_w \tilde{B}_w = \sum_{n-r+1}^{n} a_w B_w,$$

and (3) is proved.

3. APPLICATION TO COUPON COLLECTING

Suppose that there are n coupon types with respective probabilities $p_1, \ldots, p_n$. Let N_r be the number of coupons one needs to collect to obtain at least one coupon from any r different types. Following Ross [5, pp. 313–316], construct n independent Poisson

https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press

processes with rates $p_1, \ldots, p_n$. Define Y_r to be the waiting time until there is at least one event from any r different processes. Then

$$Y_r \stackrel{\text{st}}{=} \sum_{1}^{N_r} V_i \tag{6}$$

where, $\{V_i\}$, are i.i.d. exponentials with mean 1 and $\{V_i\}$ and N_r are independent. It follows from (3) and (6) that

$$EN_r = EY_r = \sum_{n-r+1}^{n} (-1)^{w-(n-r+1)} \binom{w-1}{n-r} K_w, \tag{7}$$

where $K_w = \sum_{\alpha \epsilon C_w} (1/p(\alpha))$, and $p(\alpha) = \sum_{i \epsilon \alpha} p_i$; also,

$$\text{Var } N_r = \text{Var } Y_r - EY_r. \tag{8}$$

From (3),

$$EY_r^2 = \sum_{w=n-r+1}^{n} (-1)^{w-(n-r+1)} \binom{w-1}{n-1} L_w, \tag{9}$$

with $L_w = 2 \sum_{\alpha \epsilon C_w} (1/p^2(\alpha))$.

From, (7)–(9), the variance of N_r can be computed.

Example: Let $n = 6,\ r = 3,\ p_i = i/21,\ i = 1, \ldots, 6$. Then

$$EY_3 = K_4 - 4K_5 + 10 \quad (\text{since } K_6 = 1)$$

and

$$EY_3^2 = L_4 - 4L_5 + 20 \quad (\text{since } L_6 = 2).$$

From these we use (7) and (8) to compute $EN_3 = 3.988$, and $SD(N_3) = 1.327$, both rounded to three decimal places.

4. COMPARISON OF MINIMUMS

Our objective in this section is to derive (1). We first remark that

$$EY_1 = E[E(Y_1|\underline{\lambda})] = E\left(\sum_{1}^{n} \lambda_i\right)^{-1}.$$

Next, $\min Z_i$ is exponential with failure rate $\sum_1^n (E\lambda_i^{-1})^{-1}$; thus,

$$E \min Z_i = \left(\sum_{1}^{n} (E\lambda_i^{-1})^{-1}\right)^{-1}.$$

We assume that $E\lambda_i^{-1} < \infty, i = 1, \ldots, n$.

https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press

Brown [2, p. 222] derived the inequality

$$E\left(\sum_1^n \lambda_i\right)^{-1} \leq \left(\sum_1^n (E\lambda_i^{-1})^{-1}\right)^{-1}, \tag{10}$$

true for all positive random variables, with no independence assumption (but with $E\lambda_i^{-1} < \infty, i = 1, \ldots, n$.) Thus, (1) follows from (10). An interesting discussion and extension of (1) is found in Olkin and Shepp [4].

Y_1 is stochastically larger than an exponential distribution with failure rate $\sum_1^n E\lambda_i$. Thus, an obvious extension of (1) to a two-sided bound gives

$$\left(\sum_1^n E\lambda_i\right)^{-1} \leq EY_1 \leq \left(\sum_1^n (E\lambda_i^{-1})^{-1}\right)^{-1}.$$

Y_1 is not, in general, stochastically larger than $\min Z_i$. Each X_i being decreasing failure rate (DFR) has a smaller survival function than Z_i for t small and a larger survival function than Z_i for t large. Thus, for example, if $X_1, \ldots, X_n$ are independent, then $Y_1 = \min X_i$ has a smaller survival function than $\min Z_i$ for t small and a larger survival function for t large.

5. COMPARISON OF MAXIMUMS

For $n = 2$, it follows from (1) that

$$\begin{aligned} EY_2 &= E\max(X_1, X_2) \\ &= EX_1 + EX_2 - E\min(X_1, X_2) \\ &= EZ_1 + EZ_2 - E\min(X_1, X_2) \\ &\geq EZ_1 + EZ_2 - E\min(Z_1, Z_2), \\ &= E\max(Z_1, Z_2). \end{aligned}$$

Thus, result (2) easily follows from (1) for $n = 2$.

For general n, we get a restricted version of (2), as discussed on Section 1.

Define, $D_{n,i} = (n - i + 1)(Y_i - Y_{i-1}),\ i = 1, \ldots, n$, known as the time on test statistics. We now derive the following result.

LEMMA 1: *$ED_{n,i}$ is nondecreasing in i, for fixed n.*

PROOF: Let F_{i-1} denote the indexes, among $\{1, \ldots, n\}$, for which the first i order statistics, $Y_1, \ldots, Y_i$ (from $X_1, \ldots, X_n$), are achieved. Define, $s = \sum_{j \epsilon F_i^c} \lambda_i$. Then

$$E[D_{n,i} | \underline{\lambda}, F_{i-1}] = \frac{n - i + 1}{s}$$

https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press

and

$$E[D_{n,i+1}|\underline{\lambda}, F_{i-1}] = (n-i)\sum_{j\epsilon F_{i-1}^c}\left(\frac{\lambda_j}{s}\frac{1}{s-\lambda_j}\right) \tag{11}$$

Given, $(\underline{\lambda}, F_{i-1})$ define a random variable U by $P_r(U=\lambda_j) = 1/(n-i+1)$, for $\lambda_j \epsilon F_{i-1}^c$. Then, from (11),

$$\begin{aligned}
E[D_{n,i+1}|\underline{\lambda}, F_{i-1}] &= (n-i+1)\frac{n-i}{s}E\left(\frac{U}{s-U}\right) \\
&\geq (n-i+1)\left(\frac{n-i}{s}\right)EUE(S-U)^{-1} \\
&\text{(since both } U \text{ and } (S-U)^{-1} \text{ are increasing functions of } U) \\
&\geq (n-i+1)\left(\frac{n-i}{s}\right)EU(s-EU)^{-1} \\
&\text{(since } (s-U)^{-1} \text{ is a convex function of } U) \\
&= (n-i+1)\left(\frac{n-i}{s}\right)\left(\frac{s}{n-i+1}\right)\left(\frac{n-i+1}{(n-i)s}\right) \\
&= \frac{n-i+1}{s} = ED_{n,i}(\underline{\lambda}, F_{i-1}).
\end{aligned}$$

Thus,

$$\begin{aligned}
ED_{n,i+1} &= E[E(D_{n,i+1}|\underline{\lambda}, F_{i-1})] \\
&\geq E[E(D_{n,i}|\underline{\lambda}, F_{i-1})] \\
&= ED_{n,i}.
\end{aligned}$$

When $\lambda_1, \ldots, \lambda_n$ are i.i.d., $D_{n,i}$ is known to be stochastically increasing in i for fixed n [1, p. 266]. ■

Examination for the proof of Lemma 1 shows that the inequality in the conclusion is strict as long as $\Pr(\lambda_i \neq \lambda_j)$ for some $i, j > 0$.

We now prove our restricted version of (2).

PROPOSITION 1: *Suppose that* $X_i = \lambda_i^{-1}\mathcal{E}_i$, $i = 1, \ldots, n$, *with* $\mathcal{E}_1, \ldots, \mathcal{E}_n$ *i.i.d. exponential with mean 1, and* $E\lambda_i^{-1} = c < \infty, i = 1, \ldots, n$. *Define* $Z_i = c\mathcal{E}_i, i = 1, \ldots, n$, *so that* $Z_1, \ldots, Z_n$ *are i.i.d. exponential with mean c. Then*

$$E(\max X_i) \geq E(\max Z_i) = c\sum_1^n \frac{1}{i}.$$

PROOF: Define, $b_i = ED_{n,i} = (n-i+1)E(Y_i - Y_{i-1})$, $i = 1, \ldots, n$. Then $\sum_1^n b_i = \sum_1^n EX_i = nc = \sum_1^n EZ_i$.

https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press

By Lemma 1, b_i is nondecreasing. It follows that a weighted average of b_i, with weights increasing in i, is at least as large as the unweighted average, $\frac{1}{n}\sum_1^n b_i = c$. Thus,

$$\left(\sum_1^n \frac{b_i}{n-i+1}\right)\left(\sum_1^n \frac{1}{n-i+1}\right)^{-1} \geq c$$

and

$$\frac{E(\max X_i)}{E(\max Z_i)} = \frac{1}{c}\left(\sum \frac{b_i}{n-i+1}\right)\left(\sum \frac{1}{n-i+1}\right)^{-1} \geq \frac{1}{c}(c) = 1,$$

which proves the result. ■

Acknowledgment

This work was sponsored by the Office of Industry Relations of the National Security Agency.

References

1. Barlow, R.E., Bartholomew, D.J., Bremner, J.M. & Brunk, H.D. (1972). *Statistical inference under order restrictions*. New York: Wiley.
2. Brown, M. (2006). Exploiting the waiting time paradox: Applications of the size-biasing transformation. *Probability in the Engineering and Informationa Sciences* 20: 195–230.
3. Marshall, A.W. & Olkin, I. (2007). *Life distributions, structure of nonparametric, semiparametric, and parametric families*. New York: Springer.
4. Olkin, I. & Shepp, L. (2006). Several colorful inequalities. *The American Mathematical Monthly* 113: 817–822.
5. Ross, S.M. (2007). *Introduction to probability models*, 9th ed. Amsterdam: Elsevier.

https://doi.org/10.1017/S0269964809000278 Published online by Cambridge University Press