COMPARISONS OF ORDER STATISTICS
AND OF SPACINGS FROM HETEROGENEOUS
DISTRIBUTIONS\(^1\)

by
Gordon Pledger and Frank Proschan
FSU Statistics Report M212
AFOSR Technical Report No. 2

June, 1971
The Florida State University
Department of Statistics
Tallahassee, Florida

\(^1\)Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR-71-2058 and the National Science Foundation under Grant GU 2612. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.
COMPARISONS OF ORDER STATISTICS AND OF SPACINGS
FROM HETEROGENEOUS DISTRIBUTIONS

by Gordon Pledger and Frank Proschan
The Florida State University

Abstract

Given two sets of independent, possibly unlike, components, conditions involving majorization are given which insure that any k-out-of-n system constructed of components in the first set will have reliability at least as great as that of a corresponding system constructed of components in the second set. Since the ordered failure times of the components represent order statistics from heterogeneous distributions, we obtain stochastic comparisons between the order statistics from one set of underlying distributions \( \{F_1, \ldots, F_n\} \) and those from another set \( \{F_1^*, \ldots, F_n^*\} \) under both parametric and nonparametric assumptions. As a sample result, if one vector of component hazards \((-\log[1-F_1(t)], \ldots, -\log[1-F_n(t)]\)) majorizes a second such vector \((-\log[1-F_1^*(t)], \ldots, -\log[1-F_n^*(t)]\)) for each \( t \geq 0 \), then for \( k = 1, \ldots, n \), the \( k^{th} \) order statistic from the set \( \{F_1, \ldots, F_n\} \) is stochastically larger than the \( k^{th} \) order statistic from the set \( \{F_1^*, \ldots, F_n^*\} \). Results of this type can be used to find bounds for the reliability of a k-out-of-n system of unlike components in terms of a k-out-of-n system of like components.

Under more restrictive hypotheses, among them the assumption that \( F_1^* = \cdots = F_n^* \), both stochastic and expected value comparisons for the
spacings between order statistics are obtained. It is also shown that in the case of proportional concave hazards, the successive normalized spacings from one set of heterogeneous distributions are stochastically increasing.
COMPARISONS OF ORDER STATISTICS AND SPACINGS
FROM HETEROGENEOUS DISTRIBUTIONS

by Gordon Pledger and Frank Proschan
The Florida State University

1. Introduction. A great body of theory and methods exists for order statistics and their spacings from a single underlying distribution. See for example David (1970), Sarhan and Greenberg (1962), and Pyke (1965, 1970), and the references contained therein.

A far smaller set of results is available for the case of order statistics from underlying heterogeneous distributions. See for example Sen (1970). One motivation for considering underlying heterogeneous distributions arises in reliability theory, where so-called "k-out-of-n" systems are studied.

A system of n components is called a k-out-of-n system if it functions if and only if at least k components function. See Barlow and Proschan (1965), Chapter 7, and the references contained therein. Note that the time of failure of a k-out-of-n system of independent components with respective life distributions \( F_1, \ldots, F_n \) corresponds to the \((n-k+1)\) th order statistic from the set of underlying heterogeneous distributions \( \{F_1, \ldots, F_n\} \). With this fact in mind, we may wish to approximate the life distribution of a k-out-of-n system of stochastically unlike components by the life distribution of a k-out-of-n system

---

1Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR-71-2058 and the National Science Foundation under Grant GU 2612. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.
consisting of stochastically like components. Thus we would be comparing order statistics from underlying heterogeneous populations with corresponding order statistics from some appropriate underlying homogeneous population.

Sen (1970) derives such comparisons, taking as the homogeneous population the equally weighted mixture of the original $F_1, \ldots, F_n$. In the present paper, we present additional comparisons involving not only order statistics but also spacings between order statistics. In some of our comparisons, the underlying heterogeneous distributions are compared against a single underlying homogeneous distribution, while in others, they are compared against another set of distributions, less heterogeneous in the sense of majorization.

2. Comparisons for $k$-out-of-$n$ Systems and Corresponding Order Statistic Implications. The simplest comparisons of $k$-out-of-$n$ systems may be made by taking fixed component reliabilities, $p_1, \ldots, p_n$, rather than time-dependent component reliabilities $\bar{F}_1(t), \ldots, \bar{F}_n(t)$. For component reliability $p_i$ we define the corresponding component hazard $R_i$ by:

$$R_i = -\log p_i.$$  \hspace{1cm} (2.1)

Many of our comparisons involve the notion of majorization. Vector $a$ majorizes vector $b$ (written $a \succ_b b$) if, possibly after reordering components, $a_1 \geq \cdots \geq a_n, b_1 \geq \cdots \geq b_n$, $\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i$ for $k = 1, \ldots, n-1$, and $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$. Throughout this paper we shall assume, without loss of
generality, that each vector \( \mathbf{a} \) is already ordered in accordance with this definition (so that \( a_1 \geq \ldots \geq a_n \)).

To obtain our comparisons, we also need the notion of a Schur function. A function \( f(\mathbf{a}) \) with continuous partial derivatives of first order is a Schur function if for all \( i \neq j \),

\[
(a_i - a_j) \left( \frac{\partial f}{\partial a_i} - \frac{\partial f}{\partial a_j} \right) \geq 0.
\]

The notions of majorization and Schur function are linked by the following result of Ostrowski (1952). This result, as well as others relating to majorization, can be found in Mitrinović (1970), pp. 162-170.

**Theorem 2.1.** (Ostrowski) Let \( f \) be a symmetric function with continuous partial derivatives of first order. Then a necessary and sufficient condition for \( f(\mathbf{a}) > f(\mathbf{b}) \) whenever \( \mathbf{a} \succ \mathbf{b} \) is that \( f \) be a Schur function.

Using these notions, we may state a comparison for a \( k \)-out-of-\( n \) system; we denote its system reliability \( h_k(p_1, \ldots, p_n) \), as a function of component reliabilities \( p_1, \ldots, p_n \).

**Theorem 2.2.** Let \( \mathbf{R} = (R_1, \ldots, R_n) \) be a vector of component hazards which majorizes \( \mathbf{R}^\# = (R_1^\#, \ldots, R_n^\#) \), a second vector of component hazards. Then the corresponding reliabilities for a \( k \)-out-of-\( n \) system satisfy

\[
(2.2a) \quad h_k(p) > h_k(p^\#) \quad \text{for } k = 1, \ldots, n-1, \text{ and}
\]
\[
(2.2b) \quad h_n(p) = h_n(p^\#).
\]
Proof. (2.2a) It will suffice to prove for \( k = 1, \ldots, n-1 \), that
\[
\log h_k(e^{-1}, \ldots, e^{-n}) \text{ is a Schur function of } (R_1, \ldots, R_n),
\]
since by Ostrowski's Theorem, this would imply
\[
\log h_k(e^{-1}, \ldots, e^{-n}) > \log h_k(e^{-1}, \ldots, e^{-n}),
\]
which implies (2.2a). (The function \( n_k(R) = -\log h_k(p) \) is called the hazard transform of the \( k \)-out-of-\( n \) system; it represents the system hazard as a function of component hazards \( R_1, \ldots, R_n \). See Esary, Marshall, and Proschan (1970) for a systematic treatment of hazard transforms and their applications in reliability theory.)

Write for \( k = 1, \ldots, n \),
\[
n_k(R) = -\log h_k(p) = \log(e^{-1} [1 + \sum_{i=1}^{n} R_i (e^{-1} - 1) + \cdots + \sum_{i_1 < \cdots < i_{n-k}} \prod_{i_{k+1} \leq i_k} (e^{-1} - 1) \cdots (e^{-1} - 1)]) = \sum_{i=1}^{n} R_i - \log C_k,
\]
where \( C_k = 1 + \sum_{i=1}^{n} R_i (e^{-1} - 1) + \cdots + \sum_{i_1 < \cdots < i_{n-k}} \prod_{i_{k+1} \leq i_k} (e^{-1} - 1) \cdots (e^{-1} - 1) > 0 \)
for \( k = 1, \ldots, n-1 \), and \( C_k = 1 \) for \( k = n \).

For \( k = 1, \ldots, n-1 \), to show that \( -n_k(R) \) is a Schur function, we determine
\[
\frac{\partial n_k}{\partial R_j} = 1 - \frac{1}{C_k} [e^j + e^j \sum_{i \neq j} R_i (e^{-1} - 1) + \cdots + e^j \sum_{i_1 < \cdots < i_{n-k-1}} \prod_{i_{k+1} \leq i_k} (e^{-1} - 1) \cdots (e^{-1} - 1)].
\]
Suppose $j > m$, so that $R_j < R_m$. Define $\Delta_k = \frac{\partial \eta_k}{\partial R_j} - \frac{\partial \eta_k}{\partial R_m}$. Then

$$\Delta_{n-1} \quad \text{sgn} \quad \frac{R_m - e^j}{e^m - e^j} \geq 0,$$

where $x \quad \text{sgn} \quad y$ means $x$ and $y$ have the same sign. A straightforward induction shows that

$$\Delta_k \quad \text{sgn} \quad \frac{R_m - e^j}{e^m - e^j} \sum_{i_{l+1} < \cdots < i_{n-k-l}} (e^{i_{l+1}} - \cdots - e^{i_{n-k-l}}) > 0,$$

for $k = 1, \ldots, n-2$. Hence $-\eta_k(R)$ is a Schur function for $k = 1, \ldots, n-1$, so that (2.2a) holds.

$$-\sum_{i} R_i \quad -\sum_{i} R_i^*$$

(b) $\eta_n(p) = e^l_1 = e^l_1 = h_n(p^*)$. (We note incidentally that $-\eta_n(R)$ is a Schur function.) ||

A case of special interest occurs when $R_1^* = \cdots = R_n^* = \frac{1}{n} \sum_{i} R_i$, or equivalently, when $p_1^* = \cdots = p_n^* = (\prod_{i} p_i)^{1/n}$, the geometric mean of the $p_1, \ldots, p_n$. Since we are assuming $R_1 \geq \cdots \geq R_n$, it follows that $R > R^*$. Thus we may apply Theorem 2.2 to obtain

**Corollary 2.3.** Let $p$ be the vector of component reliabilities for a $k$-out-of-$n$ system, and let $p_G = (\prod_{i} p_i)^{1/n}$. Then

$$h_k(p_1, \ldots, p_n) \geq h_k(p_G, \ldots, p_G) \quad \text{for} \quad k = 1, \ldots, n.$$

This bound is sharp.

Conclusion (2.3) gives a lower bound for the reliability of $k$-out-of-$n$ systems of unlike components in terms of the corresponding reliability
in the case of like components. The bound is sharp in the following sense. If any value \( p > p_G = (\prod_{i=1}^{n} p_i)^{1/n} \) is used for common component reliability, then (2.3) cannot hold for \( k = 1, \ldots, n \). This is a consequence of the fact that equality holds in (2.3) for \( k = n \).

The following lemma provides a means of deriving new Schur functions.

**Lemma 2.4.** Let \( f(y) \) be a Schur function, with \( \frac{\partial f}{\partial y} \leq (>) 0 \) for \( i = 1, \ldots, n \). Let \( y_i = g(x_i) \), where \( g \) is increasing, differentiable, and concave (convex). Then \( h(x) = f(g(x_1), \ldots, g(x_n)) \) is a Schur function.

**Proof.** The conclusion follows from inspection of

\[
\frac{\partial h}{\partial x_j} = \frac{\partial h}{\partial x_i} = \frac{\partial f}{\partial y_j} \frac{\partial y_i}{\partial x_j} - \frac{\partial f}{\partial y_i} \frac{\partial y_i}{\partial x_i}.
\]

Using Lemma 2.4 we may make comparisons of \( k \)-out-of-\( n \) systems when components have comparable odds ratios. We define the **odds ratio** \( r_i \) in favor of failure of component \( i \) by

\[(2.4) \quad r_i = \frac{1-p_i}{p_i}.
\]

It follows from (2.4) that

\[
r_i = e^{R_i} - 1, \text{ or } R_i(r_i) = \log (1+r_i),
\]

a differentiable, increasing, concave function of \( r_i \). This last identity, in conjunction with Lemma 2.4 and Theorem 2.2, yields

**Theorem 2.5.** If \( \bar{r} > \bar{r}^* \), then

\[
h_k(p) > h_k(p^*) \text{ for } k = 1, \ldots, n.
\]
Note that the conclusions of Theorem 2.2 and Theorem 2.5 are the same for \( k = 1, \ldots, n-1 \), while the hypotheses differ. A careful examination shows that neither of the hypotheses \( R > R^* \) and \( r > r^* \) implies the other. Thus there may arise cases where Theorem 2.2 applies while Theorem 2.5 does not, and vice versa.

However in one important case, Theorem 2.2 yields a stronger result than does Theorem 2.5. If in Theorem 2.5 we take \( r^* = \cdots = r^* = \frac{1}{n} \sum_{i=1}^{n} r_i \), then the corresponding common component reliability is

\[
P_H = \left[ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{p_i} \right]^{-1},
\]

the harmonic mean of the \( p_i \). Recall that the harmonic mean \( p_H \leq p_G \), the geometric mean. Thus a \( k \)-out-of-\( n \) system of components each with reliability \( p_H \) is less reliable than a \( k \)-out-of-\( n \) system of components each with reliability \( p_G \).

Our results thus far have involved comparisons for a fixed \( k \)-out-of-\( n \) system. Next we present a set of inequalities among system reliabilities \( h_k(p) \), \( k = 1, \ldots, n \), for a fixed vector \( p \) of component reliabilities.

**Theorem 2.6.** Let \( h_k(p) \) be the reliability of a \( k \)-out-of-\( n \) system of independent components with respective reliabilities \( p_1, \ldots, p_n \). Then

\[
h_k^2(p) \geq h_{k-1}(p) h_{k+1}(p) \quad \text{for} \quad k = 2, \ldots, n-1,
\]

i.e., for fixed \( p \), \( h_k(p) \) is a Pólya frequency sequence of order 2 \( (PF_2) \) in the index \( k \).
Proof. Write

\[ h_k(p) = \sum_{j=-\infty}^{\infty} e_{n-k-j} H_j, \]

where

\[ e_j = \begin{cases} 
0 & \text{for } j \leq -1 \\
1 & \text{for } j = 0 \\
\left(\prod_{i_1, \ldots, i_j} (1-p_{i_1}) \cdots (1-p_{i_j}) \right) & \text{for } j > 1 
\end{cases} \]

and

\[ H_j = \begin{cases} 
0 & \text{for } j < 0 \\
n & \Pi p_i \text{ for } j > 0. \\
1 & \text{for } j = 0 
\end{cases} \]

\{H_j\} is clearly PF_2 and \{e_j\} is PF_2 by Theorem 2, p. 95, of Mitrovic (1970). Therefore, their convolution \( h_k \) is PF_2 (Karlin, 1968, p. 394).

Up to now we have been studying the reliability of \( k \)-out-of-\( n \) systems of components having fixed reliabilities (i.e., non time-dependent). Next we consider time-dependent models and present stochastic comparisons for order statistics. We say \( X \) is stochastically greater than \( Y(X >_S Y \) or \( Y >_S X) \) if \( P[X > x] > P[Y > x] \) for each real \( x \). We say \( X \) is stochastically equal to \( Y(X =_S Y) \) if \( P[X > x] = P[Y > x] \) for each real \( x \).

In some cases the results follow immediately by setting \( p_i = F_i(t) \) (and \( p_i^* = F_i^*(t) \)), where \( F_i(t) = 1 - F_i(t) \) represents the survival probability corresponding to life distribution \( F_i(t), i = 1, \ldots, n \). Such
results are numbered in a way which indicates their connection with earlier theorems.

Our model is as follows. We assume independent observations, one observation from distribution \( F_i(F_i) \), \( i = 1, \ldots, n \). The ordered observations are denoted by \( Y_1 \leq \cdots \leq Y_n \) (\( Y_1^* \leq \cdots \leq Y_n^* \)). From Theorem 2.2 we immediately obtain:

\[ \text{Theorem 2.2'}. \quad (a) \text{Let } (-\log F_1^*(t), \ldots, -\log F_n^*(t)) \]
\[ \overset{\text{m}}{\geq} (-\log F_1^*(t), \ldots, -\log F_n^*(t)) \text{ for each } t \geq 0. \text{ Then } Y_1^{\text{st}} = Y_1^* \]
and \( Y_k^{\text{st}} \geq Y_k^* \) for \( k = 2, \ldots, n \).

(b) Let \( (-\log F_1^*(t), \ldots, -\log F_n^*(t)) \overset{\text{m}}{=} (-\log F_1^*(t), \ldots, -\log F_n^*(t)) \)
for each \( t \geq 0 \). Then \( Y_k^{\text{st}} \leq Y_k^* \) for \( k = 1, \ldots, n-1 \), and \( Y_n^{\text{st}} = Y_n^* \).

Note that (b) is the dual of (a), obtained by interchanging \( F_i \) and \( F_i^* \), and \( P[Y_k > t] \) and \( P[Y_{n-k+1} < t] \).

Just as Corollary 2.3 represents a case of special interest obtained from Theorem 2.2, so Corollary 2.3' below represents a case of special interest obtained from Theorem 2.2'.

\[ \text{Corollary 2.3'}. \quad (a) \text{Let } \bar{F}_1^*(t) = \cdots = \bar{F}_n^*(t) = [\prod_{t} \bar{F}_1(t)]^{1/n} \]
for each \( t \geq 0 \). Then

(2.6a) \( Y_k^{\text{st}} \geq Y_k^* \) for \( k = 1, \ldots, n \); in particular \( Y_1^{\text{st}} = Y_1^* \).

(b) Let \( F_1^*(t) = \cdots = F_n^*(t) = [\prod_{t} F_1(t)]^{1/n} \) for \( t \geq 0 \). Then

(2.6b) \( Y_k^{\text{st}} \leq Y_k^* \) for \( k = 1, \ldots, n \); in particular, \( Y_n^{\text{st}} = Y_n^* \).

These bounds are sharp.
As above, the sharpness of the bounds is a consequence of the fact that if in (a) we replace the common value $[\Pi_i F_i(t)]^1/n$ by a larger value for some $t$, then the conclusion $Y_{1\text{ st}} > Y_1^*$ would no longer hold; similarly for (b).

In a similar fashion we may obtain the analogue of Theorem 2.5 above, with stochastic comparisons of order statistics, rather than reliability comparisons for k-out-of-n systems.

\textbf{Theorem 2.5'.} (a) Let \( \left( \frac{F_1(t)}{F_{1 n}(t)}, \ldots, \frac{F_n(t)}{F_{n n}(t)} \right) \) \( \succ \) \( \left( \frac{F_1^*(t)}{F_{1 n}^*(t)}, \ldots, \frac{F_n^*(t)}{F_{n n}^*(t)} \right) \) for each $t > 0$. Then $Y_{k \text{ st}} \succ Y_k^*$ for $k = 1, \ldots, n$.

(b) Let \( \left( \frac{F_1(t)}{F_{1 n}(t)}, \ldots, \frac{F_n(t)}{F_{n n}(t)} \right) \) \( \succ \) \( \left( \frac{F_1^*(t)}{F_{1 n}^*(t)}, \ldots, \frac{F_n^*(t)}{F_{n n}^*(t)} \right) \) for each $t > 0$.

Then $Y_{k \text{ st}} \prec Y_k^*$ for $k = 1, \ldots, n$.

Next we apply these theorems to several broad parametric families of distributions. We consider the case of proportional hazards, discussed in Esary, Marshall, and Proschan (1970). In keeping with the definition (2.1) of hazard in the non time-dependent case, we may define the hazard function (or more simply, hazard) $R(t)$ corresponding to survival probability $\bar{F}(t)$ in the time-dependent case by

\begin{equation}
(2.7) \quad R(t) = -\log \bar{F}(t) \quad \text{for } t \geq 0.
\end{equation}

Thus survival probability may be expressed in terms of the hazard function by
(2.8) $\overline{F}(t) = e^{-R(t)}$ for $t \geq 0$.

Hazard functions are proportional if hazard $R_i(t)$ may be expressed as

(2.9) $R_i(t) = \lambda_i R(t)$ for $t \geq 0$, $\lambda_i > 0$, $i = 1, \ldots, n$,

where $R(t)$ is a hazard function.

Corollary 2.7. Assume proportional hazards $\lambda_i R(t)$ (alternately $\lambda_i^* R(t)$), $i = 1, \ldots, n$, where $R(t)$ is a hazard function. Let $\lambda > \lambda^*$. Then $Y_{1}^{st} = Y_1^*$ and $Y_{k}^{st} > Y_k^*$ for $k = 2, \ldots, n$.

Proof. Observe that $\lambda > \lambda^*$ implies that

$(-\log \overline{F}_1(t), \ldots, -\log \overline{F}_n(t)) > (-\log \overline{F}_1^*(t), \ldots, -\log \overline{F}_n^*(t))$ for each $t \geq 0$. The result follows by Theorem 2.2'. ||$

Important special cases of proportional hazards are:

(i) Weibull survival probability $\overline{F}_i(t) = e^{-\lambda_i t^\alpha}$, $\alpha > 0$, $i = 1, \ldots, n$.

(ii) Specializing further, with $\alpha = 1$, exponential survival probability

$\overline{F}_i(t) = e^{-\lambda_i t}$, $i = 1, \ldots, n$.

Another interesting comparison can be obtained when survival probability is logarithmically convex in the parameter $\lambda$.

Theorem 2.8. For $t \geq 0$ let $F(t, \lambda)$, $(F(t, \lambda))$ be differentiable, monotone, and log convex in $\lambda > 0$. If $\lambda > \lambda^*$, then $Y_{k}^{st} > (\leq) Y_k^*$ for $k = 1, \ldots, n$. 
Proof. Assume $F(t, \lambda)$ is differentiable, monotone, and log convex. For $t \geq 0$, define $\phi_k(\lambda) = P[Y_k > t]$. Then

$$\phi_1 = \prod_{l=1}^{n} \bar{F}(t, \lambda_i^l)$$

and

$$\frac{\partial \phi_1}{\partial \lambda_j} = \left( \frac{\partial F(t, \lambda_j)}{\bar{F}(t, \lambda_j)} \right) \left( \frac{\partial \lambda_j}{\partial \lambda_j} \right) .$$

Since $\bar{F}(t, \lambda)$ is log convex, it follows that $\phi_1(\lambda)$ is a Schur function.

Also

$$\phi_2 = \prod_{l=1}^{n} \bar{F}(t, \lambda_i^l) \left[ 1 + \sum_{i=1}^{n} \frac{F(t, \lambda_i^l)}{\bar{F}(t, \lambda_i^l)} \right],$$

and

$$\frac{\partial \phi_2}{\partial \lambda_j} = \prod_{i=1}^{n} \bar{F}(t, \lambda_i^l) \cdot \frac{\partial F(t, \lambda_j)}{\bar{F}(t, \lambda_j)} \cdot \left[ 1 + \sum_{i=1}^{n} \frac{F(t, \lambda_i^l)}{\bar{F}(t, \lambda_i^l)} - \frac{1}{\bar{F}(t, \lambda_j)} \right]$$

$$= \prod_{i=1}^{n} \bar{F}(t, \lambda_i^l) \cdot \frac{\partial F(t, \lambda_j)}{\bar{F}(t, \lambda_j)} \cdot \sum_{i \neq j} \frac{F(t, \lambda_i^l)}{\bar{F}(t, \lambda_i^l)} .$$

Since $\bar{F}(t, \lambda)$ is log convex and $\bar{F}(t, \lambda)$ is monotone in $\lambda$, it follows that $\phi_2(\lambda)$ is a Schur function. Finally for $k = 2, \ldots, n-1,$
\[ \frac{\delta \phi_{k+1}}{\delta \lambda_j} = \prod_{i=1}^{n} \frac{\delta F(t, \lambda_i)}{\delta \lambda} \sum_{i_1 < \cdots < i_k} \frac{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_k})}{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_k})} \]

\[ - \frac{F(t, \lambda_j)}{F(t, \lambda_j)} \sum_{i_1 < \cdots < i_{k-1}} \frac{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})}{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})} \]

Suppose \( F(t, \lambda) \) is decreasing in \( \lambda \). Since \( \log F(t, \lambda) \) is convex in \( \lambda \), then for \( \lambda_j \geq \lambda_m \),

\[ \frac{\delta F(t, \lambda_j)}{\delta \lambda} > \frac{\delta F(t, \lambda_m)}{\delta \lambda}, \]

with both ratios negative. Also

\[ 0 \leq \sum_{i_1 < \cdots < i_k} \frac{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})}{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})} \]

\[ - \frac{F(t, \lambda_j)}{F(t, \lambda_j)} \sum_{i_1 < \cdots < i_{k-1}} \frac{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})}{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})} \]

\[ \leq \sum_{i_1 < \cdots < i_k} \frac{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_k})}{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_k})} \]

\[ - \frac{F(t, \lambda_m)}{F(t, \lambda_m)} \sum_{i_1 < \cdots < i_{k-1}} \frac{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})}{F(t, \lambda_{i_1}) \cdots F(t, \lambda_{i_{k-1}})} . \]
It follows that \( \phi_{k+1} \) is a Schur function for \( k = 2, \ldots, n-1 \). A similar argument applies if \( \tilde{F}(t, \lambda) \) is increasing in \( \lambda \). It follows by Theorem 2.1 that for \( k = 1, \ldots, n, \phi_k(\lambda^*) \geq \phi_k(\lambda^*), \) and so, \( Y_{n-k} \geq Y_k \).

A similar proof holds for the case \( F(t, \lambda) \) differentiable, monotone, and log convex in \( \lambda \). For this case, we show that \( P[Y_{n-k} < t] \) is a Schur function of \( \lambda \), since it has the same form as \( \phi_{k+1} \) with each \( \tilde{F}_i \) replaced by \( F_i \).

A familiar example of a class of distributions satisfying the hypothesis of Theorem 2.8 is obtained by considering survival probability \( \tilde{F}(t, \lambda) = \tilde{G}(\lambda t) \), where distribution \( G \) has decreasing failure rate (DFR) and \( \lambda \) occurs as a scale factor. The hypothesis is satisfied since a DFR survival probability is log convex (Barlow and Proschan, 1965, Chapter 2.)

We now give some examples of survival probabilities \( \tilde{F}(t, \lambda) \) which are log convex in \( \lambda \), so that Theorem 2.6 holds.

(i) \( \tilde{F}(t, \lambda) = e^{-\lambda t} \) for \( 0 < \alpha \leq 1 \). This parameterization of the Weibull distribution is different from that following Corollary 2.7. In the latter case, \( \log \tilde{F} \) is linear in \( \lambda \), and so the restriction to \( \alpha \leq 1 \) is not necessary.

(ii) \( F(t, \lambda) = \int_0^t \frac{\alpha}{\Gamma(\alpha)} \frac{\alpha-1}{x} e^{-\lambda x} \, dx \) for \( 0 < \alpha \leq 1 \), a gamma distribution with shape parameter \( \alpha \leq 1 \).

(iii) \( F(t, \lambda) = \sum_{i=1}^n p_i (1 - e^{-a_i \lambda t}), a_i > 0, p_i > 0, \sum_{i=1}^n p_i = 1 \).

A mixture of exponential distributions has decreasing failure rate; see Barlow and Proschan (1965), Chapter 2.
(iv) More generally, let \( \bar{f}_i(t, \lambda) \) be log convex in \( \lambda \), \( p_i > 0 \), and \( \sum_{i=1}^{n} p_i = 1 \). Then \( \bar{F}(t, \lambda) = \sum_{i=1}^{n} p_i \bar{f}_i(t, \lambda) \) is log convex in \( \lambda \). See Barlow and Proschan (1965), Chapter 2.

The last result of this section holds for the smallest and largest order statistics, or equivalently, for series and parallel systems.

**Theorem 2.9.** Let \( F(t, \lambda) \) (\( \bar{F}(t, \lambda) \)) be log concave and differentiable in \( \lambda > 0 \) for each fixed \( t > 0 \). If \( \lambda > \lambda^* \), then \( Y_n \stackrel{st}{\leq} Y^* \), \( Y_{\lambda} \stackrel{st}{\leq} Y^* \).

**Proof.** Assume \( F(t, \lambda) \) is log concave and differentiable in \( \lambda > 0 \). Then

\[
\frac{\partial}{\partial \lambda_j} P[Y_n > t] = \frac{\partial}{\partial \lambda_j} \left[ 1 - \prod_{i=1}^{n} F(t, \lambda_i) \right] = - \prod_{i=1}^{n} F(t, \lambda_i) \frac{\partial \lambda}{F(t, \lambda_j)}.
\]

If \( \lambda_j > \lambda_m \), then

\[
\frac{\partial}{\partial \lambda_j} P[Y_n > t] - \frac{\partial}{\partial \lambda_m} P[Y_n > t] \overset{sgn}{=} \frac{\partial \lambda_t}{F(t, \lambda_m)} - \frac{\partial \lambda_t}{F(t, \lambda_j)} \geq 0
\]

since \( \log F(t, \lambda) \) is concave in \( \lambda > 0 \).

The dual part of the theorem is proved in a similar fashion. \( \| \|

Note that if the density is log concave, or equivalently, \( PF_2 \), in \( t \), then both \( F \) and \( \bar{F} \) are log concave in \( t \). If, in addition, \( F \) is a function of \( \lambda t \), then \( F(\lambda t) \) and \( \bar{F}(\lambda t) \) are log concave in \( \lambda \) and thus both conclusions of Theorem 2.9 hold. Examples of such \( PF_2 \) densities are:
(i) **Weibull.** $f(t, \lambda) = \alpha \lambda^\alpha t^{\alpha-1} e^{-(\lambda t)^\alpha}$ for $\alpha \geq 1$.

(ii) **Gamma.** $f(t, \lambda) = \frac{\lambda t^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda t}$ for $\alpha > 1$.

(iii) **Truncated Normal.** $f(t, \lambda) = \alpha e^{-\frac{1}{2}(t-\mu)^2}$, $t > 0$, $\lambda > 0$,

where $\alpha > 0$ is such that $\int_0^\infty f(t, \lambda)dt = 1$. (Here the density is log concave in $t - \mu$.)

3. **Spacings.** In this section we present comparisons under majorization for spacings between order statistics. In addition, we present stochastic comparisons among the spacings arising from a single set of order statistics. In each case we assume the underlying heterogeneous distributions have proportional hazards $\lambda_1 R(t), \ldots, \lambda_n R(t)$.

Let $D_{1,n} = Y_1, D_{2,n} = Y_2 - Y_1, \ldots, D_{n,n} = Y_n - Y_{n-1}$

denote the spacings between order statistics $Y_1 < \ldots < Y_n$ computed from independent observations, one from each of $n$ heterogeneous distributions. Recall that when the observations come from a single underlying distribution $F$, an exponential, then the normalized spacings $nD_{1,n},$ $(n-1)D_{2,n}, \ldots, D_{n,n}$ are independently distributed according to the same exponential distribution $F$. More generally, we can show that when the underlying distributions have concave proportional hazards, the normalized spacings increase stochastically.

**Theorem 3.1.** Let $f_i(t) = e^{-\lambda_i t}$, $\lambda_i > 0$, for $i = 1, \ldots, n$, where $R(t)$ is a concave, differentiable hazard function. Then $nD_{1,n} \leq (n-1)D_{2,n} \leq \ldots \leq D_{n,n}$.
Proof. First we show \( nD_{1,n} \leq (n-1) D_{2,n} \) for each \( n \); then we use induction on \( n \). Let

\[
\lambda = \sum_{i=1}^{n} \lambda_i, \quad \overline{\lambda} = \lambda / n, \quad r = \frac{dR}{dt}.
\]

For \( x > 0 \),

\[
P[(n-1)D_{2,n} > x] = \sum_{i=1}^{n} \int_{0}^{\infty} \lambda_i r(t) e^{-\lambda_i R(t)} \left( (\lambda - \lambda_i) R(t+ \frac{x}{n-1}) \right) dt
\]

\[
= \int_{0}^{\infty} \sum_{i=1}^{n} \lambda_i \left( R(t+ \frac{x}{n-1}) - R(t) \right) e^{-\lambda R(t+ \frac{x}{n-1})} dt.
\]

Since \( e^{\lambda y} \) is a convex function, by Jensen's inequality we have

\[
P[(n-1) D_{2,n} > x] \geq \int_{0}^{\infty} \lambda \left( R(t+ \frac{x}{n-1}) - R(t) \right) e^{-\lambda R(t+ \frac{x}{n-1})} dt
\]

\[
= \int_{0}^{\infty} -(n-1) \lambda R(t+ \frac{x}{n-1}) - \lambda R(t) dt
\]

\[
\geq \int_{0}^{\infty} -(n-1) \lambda R(t) - (n-1) \overline{\lambda R(t)} - \overline{\lambda R(t)} dt,
\]

since \( R \) is subadditive. Thus

\[
P[(n-1)D_{2,n} > x] \geq e^{-\lambda R(\frac{x}{n-1})} \int_{0}^{\infty} \lambda r(t) e^{-\lambda R(t)} dt
\]

\[
= e^{-\lambda R(\frac{x}{n-1})}
\]

\[
\geq e^{-\lambda R(\frac{x}{n})}, \text{since } R \text{ is concave.}
\]

But the last expression gives \( P[nD_{1,n} > x] \). Thus we have proved

\( nD_{1,n} \leq (n-1) D_{2,n} \) for each \( n \).
Next suppose that for fixed \( i \), \( (n-i+1)D_{i,n} \uparrow (n-i+2)D_{i-1,n} \) for all \( n > i \). Write

\[
P[(n-i)D_{i+1,n} > x]
= \int_0^\infty \sum_{j=1}^n \lambda_j r(t)e^{-\lambda_j R(t)} P[(n-i)D_{i,n-1} > x|\lambda_j \text{ deleted}, R_t],
\]

where \( P[(n-i)D_{i,n-1} > x|\lambda_j \text{ deleted}, R_t] \) indicates that \( F_j \) is no longer present and that hazard \( R(x) \) is to be replaced by \( R_t(x) = F(t+x) - R(t) \). Using the inductive hypothesis, we have

\[
P[(n-i)D_{i+1,n} > x] \geq \int_0^\infty \sum_{j=1}^n \lambda_j r(t)e^{-\lambda_j R(t)} P[(n-i+1)D_{i-1,n-1} > x|\lambda_j \text{ deleted}, R_t] \, dt = P[(n-i+1)D_{i,n} > x].
\]

Thus \( (n-i)D_{i+1,n} \uparrow st (n-i+1)D_{i,n} \), and the proof is complete by induction. \( \Box \)

Next we compare spacings from two sets of underlying exponential distributions. We drop the second subscript in the spacings \( D_{i,n} \) and define \( D^*_i \), \( i = 1, \ldots, n \), to be the spacings arising from \( \{F^*_1, \ldots, F^*_n\} \), the second set of exponential distributions.

**Theorem 3.2.** Let \( \tilde{F}_i(t) = e^{-\lambda_i t} \) and \( \tilde{F}^*_i(t) = e^{-\tilde{\lambda} t} \) for \( i = 1, \ldots, n \), where \( \tilde{\lambda} = \frac{1}{n} \sum_{i=1}^n \lambda_i \). Then \( D_{1} \uparrow st D^*_1 \) and \( D_{i} \uparrow st D^*_i \) for \( i = 2, \ldots, n \).

**Proof.** By Theorem 3.1, we know that

\[
(3.1) \quad nD_{1} \uparrow st (n-1)D_{2} \uparrow st \ldots \uparrow st D_{n}.
\]
It is also well known that the normalized spacings from a single exponential distribution are stochastically alike, i.e.,

\[ n \text{D}_1^* = (n-1) \text{D}_2^* \leq \cdots \leq \text{D}_n^* \]

See Epstein and Sobel (1953, 1954) or Sukhatme (1936, 1937). Finally, we may verify directly that for \( x > 0 \),

\[ P[nD_1 > x] = \prod_{i=1}^{n} e^{\frac{-\lambda_i x}{n}} = e^{-\lambda x} = P[nD_* > x], \]

so that

\[ nD_1^* \leq nD_*^*. \]

From (3.1), (3.2), and (3.3), we obtain the conclusion of the theorem. \( \| \)

Note that the mean \( \bar{\lambda} \) of the \( \lambda_1, \ldots, \lambda_n \) is used for the comparison.

Is the same conclusion possible using a vector \( \lambda_* \), where \( \lambda \geq \lambda_* \)? That the conclusion need not hold when \( \lambda \geq \lambda_* \) may be shown by the counterexample:

Let \( n = 3 \), with \( \lambda_1 = 10, \lambda_2 = 1, \lambda_3 = 1 \), and \( \lambda_* = 7, \lambda_* = 7, \lambda_* = 1 \).

Then, letting \( \lambda = \lambda_1 + \lambda_2 + \lambda_3 \),

\[ P(D_3 > 1) = \left( \frac{\lambda_1 \lambda_2}{\lambda - \lambda_1} + \frac{\lambda_2 \lambda_1}{\lambda - \lambda_2} \right) e^{-\lambda \lambda} + \left( \frac{\lambda_1 \lambda_3}{\lambda - \lambda_1} + \frac{\lambda_3 \lambda_1}{\lambda - \lambda_3} \right) e^{-\lambda \lambda} \]

\[ + \left( \frac{\lambda_2 \lambda_3}{\lambda - \lambda_2} + \frac{\lambda_3 \lambda_2}{\lambda - \lambda_3} \right) e^{-\lambda \lambda} \]

\[ < P(D_* > 1). \]

In Theorem 3.2 we averaged the exponential failure rates to obtain a stochastic comparison. Next we average the exponential means to obtain
a comparison; the comparison is not stochastic, but rather is for expectations. We will first need to show

Lemma 3.3. Let $x > x^*$ and $y > y^*$. Then

$$
\frac{1}{k} \sum_{i=1}^{k} x_i y_i > \frac{1}{k} \sum_{i=1}^{k} x_i^* y_i^* \quad \text{for } k = 1, \ldots, n.
$$

Proof. Define $\phi(x) = \frac{1}{k} \sum_{i=1}^{k} x_i y_i$ for fixed $k$, $1 \leq k \leq n$, and fixed $y_1 \geq \ldots \geq y_k$. Then

$$
\frac{\partial \phi}{\partial x_i} = \begin{cases}
  y_i & \text{for } i \leq k \\
  0 & \text{for } i > k.
\end{cases}
$$

If $x_i > x_j$, then $y_i > y_j$, and thus $\phi$ is a Schur function. By Theorem 2.1,

$$
\frac{1}{k} \sum_{i=1}^{k} x_i y_i > \frac{1}{k} \sum_{i=1}^{k} x_i^* y_i^*.
$$

Using the same argument with fixed $x^*$ and $y > y^*$, we obtain the desired conclusion. 

Theorem 3.4. Let $0 \leq a_1 \leq \ldots \leq a_n$, $F_i(t) = e^{-t/\mu_i}$, $F^*_i(t) = e^{-t/\mu^*_i}$, where $\mu^*_i = \frac{1}{n} \sum_{i=1}^{n} \mu_i$, $i = 1, \ldots, n$. Then

$$
E \sum_{i=1}^{n} a_i (n-i+1) D_i \geq E \sum_{i=1}^{n} a_i (n-i+1) D^*_i.
$$

Proof. $D^*_1, \ldots, D^*_n$ are spacings from a common exponential distribution, and thus by (3.2),

$$
E n D^*_1 = E(n-1) D^*_2 = \cdots = ED^*_n.
$$
By Theorem 3.1, $ED_n > E2D_{n-l} > \cdots > EnD_1$. Since

$$E(nD_1 + \cdots + D_n) = E \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \mu_i = E \sum_{i=1}^{n} Y^*_i = E(nD^*_1 + \cdots + D^*_n),$$

it follows that

$$(ED_n, \ldots, EnD_1)^m > (ED^*_n, \ldots, EnD^*_1).$$

By Lemma 3.3 the desired conclusion follows. ||

**Remark.** Let $\tau(Y_i) = nD_1 + (n-1)D_2 + \cdots + (n-i+1)D_i$, the "total time on test" accumulated by the time of the $i$-th failure; $\tau$ is used commonly in life testing under the exponential assumption. Note that

$$E \tau(Y_i) < E \tau(Y^*_i) \quad \text{for } i = 1, \ldots, n,$$

since

$$(ED_n, E2D_{n-l}, \ldots, EnD_1)^m > (ED^*_n, E2D^*_{n-l}, \ldots, EnD^*_1),$$

as shown in the proof of Theorem 3.4.

**Corollary 3.5.** Let $\bar{F}_1(t) = e^{-t/\mu_1}$, $\bar{F}^*_1(t) = e^{-t/\mu^*_1}$, where

$$\mu^*_i = \frac{1}{n} \sum_{i=1}^{n} \mu_i, \quad i = 1, \ldots, n.$$ 

Then $E(Y_n - Y_i) > E(Y^*_n - Y^*_i)$ for $i = 0, \ldots, n-1$, where $Y_0 \equiv 0 \equiv Y_0^*$.

**Proof.** In Theorem 3.4, choose $a_1 = \cdots = a_{i-1} = 0$,

$$a_i = \frac{1}{n-i+1}, \quad a_{i+1} = \frac{1}{n-i}, \ldots, \quad a_n = 1.$$ 

Then

$$\sum_{j=1}^{n} a_j (n-j+1)D_j = Y_n - Y_i.$$
4. Applications. In this section we describe some applications of the results of Sections 2 and 3 to reliability problems.

Typically, a system (say a relay network) is to be designed using like units. However due to random fluctuations in the units, the individual unit reliabilities vary. From a knowledge of the average (not necessarily arithmetic mean) reliability of components, we wish to predict conservatively the reliability of k-out-of-n systems of the unlike components. The various theorems in Section 2 describe conditions under which we may obtain such conservative predictions.

A statistical application in the same spirit may be made as follows. Suppose for a sample of m systems, the first r successive failure times are observed for the components. Suppose further that the underlying component distributions are exponential with differing failure rates. From this data, we may obtain a conservative estimate of the life distribution of a k-out-of-n system, where k may be greater than r. Theorems and applications illustrating how conservative confidence bounds may be derived are presented in Barlow, Madansky, Proschan, and Scheuer (1968).

Statistical applications of this type as well as other statistical applications of the results of Sections 2 and 3 are being developed; a forthcoming report will give details.
REFERENCES


Comparisons of Order Statistics and of Their Spacings from Heterogeneous Distributions

Technical Report, June, 1971

Gordon Pledger and Frank Proschan

June, 1971

23

14

AFOSR-71-2058

9749

042-004

Originator's Report Number(s)

AFOSR Technical Report No. 2

Other Report No(s).

FSU Statistics Report M212

Releasable without limitations on dissemination

Sponsoring Military Activity
United States Air Force
Air Force Office of Scientific Research
1400 Wilson Blvd.
Arlington, Virginia 22209

Given two sets of independent, possibly unlike, components, conditions involving majorization are given which insure that any k-out-of-n system constructed of components in the first set will have reliability at least as great as that of a corresponding system constructed of components in the second set. Since the ordered failure times of the components represent order statistics from heterogeneous distributions, we obtain stochastic
comparisons between the order statistics from one set of underlying distributions \( \{F_1, \ldots, F_n\} \) and those from another set \( \{F_1^*, \ldots, F_n^*\} \) under both parametric and nonparametric assumptions. As a sample result, if one vector of component hazards \((-\log[1-F_1(t)], \ldots, -\log[1-F_n(t)]\)) majorizes a second such vector \((-\log[1-F_1^*(t)], \ldots, -\log[1-F_n^*(t)]\)) for each \( t \geq 0 \), then for \( k = 1, \ldots, n \), the \( k \)-th order statistic from the set \( \{F_1, \ldots, F_n\} \) is stochastically larger than the \( k \)-th order statistic from the set \( \{F_1^*, \ldots, F_n^*\} \). Results of this type can be used to find bounds for the reliability of a \( k \)-out-of-\( n \) system of unlike components in terms of a \( k \)-out-of-\( n \) system of like components.

Under more restrictive hypotheses, among them the assumptions that \( F_1^* = \ldots = F_n^* \), both stochastic and expected value comparisons for the spacings between order statistics are obtained. It is also shown that in the case of proportional concave hazards, the successive normalized spacings from one set of heterogeneous distributions are stochastically increasing.

<table>
<thead>
<tr>
<th>Key Words</th>
<th>Link A</th>
<th>Link B</th>
<th>Link C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Role</td>
<td>Role</td>
<td>Role</td>
</tr>
<tr>
<td></td>
<td>Wt</td>
<td>Wt</td>
<td>Wt</td>
</tr>
<tr>
<td>Order statistics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spacings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneous distributions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparisons</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( k )-out-of-( n ) systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majorization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>