Ann. Inst. Statist. Math.
Vol. 53, No. 2, 189–202 (2001)

OPTIMAL THRESHOLD FOR THE k-OUT-OF-n MONITOR WITH DUAL FAILURE MODES

MASAAKI SIBUYA[1] AND KAZUYUKI SUZUKI[2]

[1] *Takachiho University, Suginami-ku, Tokyo 169-8508, Japan*
[2] *The University of Electro-Communications, Chōfu City, Tokyo 182-8585, Japan*

(Received January 4, 1999; revised December 6, 1999)

Abstract. A monitor consists of n identical sensors working independently. Each sensor measures a variate of output or environment of a system, and is activated if a variate is over a threshold specified in advance for each sensor. The monitor alarms if at least k out of n sensors are activated. The performance of the monitor, the probabilities of failure to alarm and false alarming, depends on the number k, the threshold values and the probability distributions of the variate at normal and abnormal states of the system. In this paper, a sufficient condition on the pair of the distributions is given under which the same threshold values for all the sensors are optimal. The condition motivates new orders between probability distributions. Solving an optimization problem an explicit condition is obtained for maximizing or minimizing a symmetric function with the constraint of another symmetric function.

Key words and phrases: Dose response, increasing hazard function ratio, indicator variable, Lagrangian multiplier method, monotone likelihood ratio, Neyman-Pearson lemma, stochastic order.

1. Introduction

1.1 *Design of a monitor*

Many systems fail not only to operate when required to do, but also fail to be idle when required to be so. For example, a monitor, measuring a characteristic variate of a system such as pressure or temperature, or an environment variate such as radioactivity or intensity of leaked gas, may fail to detect its abnormal state or may signal false alarms. Suppose that a monitor consists of n sensors, which are of the same type and work independently of each other. To reduce failure probabilities of the dual modes, failure to alarm and false alarming, the monitor alarms if at least k out of n sensors, $1 \leq k \leq n$, detect the variate being over a threshold value specified for each ("k-out-of-n monitor").

Another simpler example is "open-circuit" failure and "short-circuit" failure in an electronic circuit. To decrease these failures, it is designed to close when k or more relays close, or open when $n - k + 1$ or more relays open. Being open in closed mode corresponds to failure to alarm in the abnormal state, and being closed in open mode corresponds to false alarming in the normal state. Since a monitor watches a variate and it has adjustable thresholds, it is more general than a circuit.

The performance of a k-out-of-n monitor depends on k, threshold values and the probability distributions of the variates at normal and abnormal states. In specifying n

threshold values of each sensor, it seems natural to set all values the same because of the identical properties of the sensors. There is a hope, however, that by properly changing the threshold values one might be able to design a better monitor. In this paper, it is shown that a monitor with the same threshold values of every sensor is optimal under some conditions on the two distributions of the variate.

1.2 *Previous Research*

Previous papers on systems with dual failure modes assume the failure probabilities of each component are fixed and known, and investigate the optimal structure. In particular, many authors investigated a system with identical components. Phillips (1980) showed that the k-out-of-n structure minimizes the sum of the dual failure probabilities. Systems with non-identical components were studied by Kohda *et al.* (1982) and Assaf *et al.* (1986). Lešanovský (1993) surveyed the papers in this field.

Suzuki and Tachikawa (1995) studied a system of components with variable thresholds. For the case $n = 2$, they showed a sufficient condition on the two distributions under which the same threshold values are optimal. The present paper extends their results and improves their condition.

Related problems can be found in Gleser (1975), and Boland and Proschan (1983). They investigated the maximization of the k-out-of-n system reliability under a constraint of components reliability. Boland and Proschan (1994) dealt with, among others, the lifetime of k-out-of-n systems of components with proportional hazard life. Boland (1998) studied the same lifetime when the components have an increasing hazard function.

1.3 *Modeling and notations*

Let Y_j, $1 \le j \le n$, denote a system characteristic or an environmental variate measured by the j-th sensor, which has a threshold value x_j, and is activated if the event $Y_j > x_j$ occurs. This event is expressed by $T_j = \mathrm{I}[Y_j > x_j]$, that is, T_j is equal to 1 if it is activated, and 0 otherwise. $\boldsymbol{T} = (T_1, \ldots, T_n)$ is a state of the monitor taking a value $\boldsymbol{t} = (t_1, \ldots, t_n) \in \{0,1\}^n$. At a critical moment, if the system is safe enough, Y_j is assumed to follow an identical cumulative distribution function (cdf) F, and if dangerous, another identical cdf G. Both F and G are known, and, from the practical view point, we assume that G is stochastically larger than F. Further, to simplify the discussion, we assume F and G are smooth and monotone increasing. Hence,

$$(1.1) \qquad P\{\boldsymbol{T} = \boldsymbol{t} \mid F, \boldsymbol{x}\} = \prod_{j=1}^{n} \bar{F}^{t_j}(x_j) F^{1-t_j}(x_j), \qquad \boldsymbol{x} = (x_1, \ldots, x_n),$$

where $\bar{F}(y) = 1 - F(y)$, is the survival distribution function (sdf) of F. $P\{\boldsymbol{T} = \boldsymbol{t} \mid G, \boldsymbol{x}\}$ is expressed in the same way.

We define a set $R \subset \{0,1\}^n$ such that if the event $T \in R$ happens we make the monitor alarm and do some action on the system, otherwise we let the system keep running without any action. The problem is to maximize the probability of appropriate alarm

$$(1.2) \qquad p(R; \boldsymbol{x}; G) := \sum_{\boldsymbol{t} \in R} P\{\boldsymbol{T} = \boldsymbol{t} \mid G, \boldsymbol{x}\}$$

under the constraint on the probability of false alarm

$$p(R;\boldsymbol{x};F) := \sum_{\boldsymbol{t}\in R} P\{\boldsymbol{T}=\boldsymbol{t} \mid F,\boldsymbol{x}\} = \alpha, \qquad 0<\alpha<1. \tag{1.3}$$

This optimization problem is similar to a statistical test problem,

$$\mathrm{H}_0 : F \quad \text{vs.} \quad \mathrm{H}_1 : G, \qquad \bar{F}(y) \le \bar{G}(y), \qquad -\infty < y < \infty. \tag{1.4}$$

The point is that we observe $T_j = \mathrm{I}[Y_j > x_j]$, not Y_j directly, and the choice of $\boldsymbol{x}$ is a part of the design. Such a design problem can arise in a destructive test, or testing a statistical hypothesis in the dose-response model. The test (1.4) of a simple hypothesis vs. a simple alternative can be extended to the one sided test of a single parameter model: $Y_j \sim K(\cdot;\theta)$, which is stochastically increasing in θ,

$$\mathrm{H}_0 : \theta=\theta_0 \quad \text{vs.} \quad \mathrm{H}_1 : \theta = \theta_1, \qquad \theta_0 < \theta_1. \tag{1.5}$$

1.4 *Formulation of problem*

Suppose, F, G, $\boldsymbol{x}$ and α are given. The optimal rule R can be determined by the Neyman-Pearson fundamental lemma as will be shown in Section 4. In the special case of "homogeneous threshold", that is $\boldsymbol{x}$ with the same component values, the k-out-of-n rule $(1 \le k \le n)$ is optimal for some values of α.

Instead, suppose only F, G and α are given, and find simultaneously optimal $\boldsymbol{x}$ and $R \subset \{0,1\}^n$. The problem is challenging, and a few results will be shown in Section 4.

In this paper, we restrict ourselves to the "k-out-of-n rule",

$$R(k,n) := \left\{ \boldsymbol{t} = (t_1,\dots,t_n) : \sum_{j=1}^{n} t_j \ge k \right\}, \quad 1 \le k \le n. \tag{1.6}$$

The objective is to show that a homogeneous threshold is optimal for this rule, under some conditions on (F,G), or on $K(\cdot;\theta)$. The restriction to k-out-of-n rule is, first of all, to reduce the search for R. Note that the choice of k is still an open problem. Second, it is in practical use, and finally it works in some related situations as we have seen.

Section 2 shows the conditions on (F,G) under which a homogeneous threshold is optimal for the k-out-of-n rule. It also proves, as a basic lemma, a maximal or minimal condition in the Lagrangian multiplier method for optimizing a symmetric function with a symmetric constraint. Section 3 lists typical families of distributions which satisfy the conditions of the optimality. In other cases there are many situations and they are shown by examples. The monotone likelihood ratio is shown to be irrelevant to the optimality of the homogeneous threshold. Section 4 discusses some related problems.

2. Main results

2.1 *Rules n-out-of-n and 1-out-of-n*

First, we study the simple case $R(n,n) = \{\boldsymbol{t} = (1,\dots,1)\}$ and $R(1,n) = \{\boldsymbol{t} \ne (0,\dots,0)\}$. For these action rules, a homogeneous threshold is globally optimal under the following conditions C_h and C_r, respectively.

For a pair of distributions (F, G), having the probability density functions (pdf's) f and g, respectively, we introduce the following conditions.

$$C_h : \frac{g(y)}{\bar{G}(y)} \Big/ \frac{f(y)}{\bar{F}(y)} \text{ is a non-decreasing function of } y. \tag{2.1}$$

$$C_r : \frac{g(y)}{G(y)} \Big/ \frac{f(y)}{F(y)} \text{ is a non-decreasing function of } y. \tag{2.2}$$

The increasing ratio of hazard functions was introduced by Kalashnikov and Rachev (1986) and studied by Sengupta and Deshpande (1994). The ratio of a pdf to its cdf, a sort of dual of the hazard function, was named as "reverse hazard ratio (function)" and explained in Shaked and Shanthikumar (1994).

If the ratios are strictly increasing, the maximum in the following theorems is strictly maximum. The conditions C_h and C_r, orders of probability distributions, will be further discussed in Section 3.

THEOREM 1. *If the condition C_h is satisfied, the homogeneous threshold $\boldsymbol{x} = \boldsymbol{a} = (a, \dots, a)$ such that*

$$\bar{F}^n(a) = \alpha,$$

maximizes globally $p(R(n,n); \boldsymbol{x}; G)$ under the constraint $p(R(n,n); \boldsymbol{x}; F) = \alpha$, for any $n = 2, 3, \dots$ and any $0 < \alpha < 1$.

PROOF. Put

$$z_j := \log(1 - F(x_j)), \qquad \varphi(u) := G(F^{-1}(u)), \quad u := F(x),$$

and

$$\psi(z) := \log(1 - G(x)) = \log(1 - \varphi(1 - e^z)), \qquad x = F^{-1}(1 - e^z).$$

The problem is to maximize $\sum_{j=1}^n \psi(z_j)$ under the constraint $\sum_{j=1}^n z_j = \log \alpha$. Since

$$\frac{d}{dz}\psi(z) = \frac{d}{dz}\log(1 - \varphi(1 - e^z)) = \frac{\varphi'(u)(1-u)}{1 - \varphi(u)},$$

which is equal to $\frac{d}{dx}(\log \bar{G}(x)) \big/ \frac{d}{dx}(\log \bar{F}(x))$, $\frac{d}{dz}\psi(z)$ is non-decreasing in x and non-increasing in z provided that the condition C_h holds. Hence, the function $\psi(z)$ is concave. For a general concave function ψ, the inequality

$$n^{-1}\sum_{j=1}^n \psi(z_j) \le \psi(\bar{z}), \qquad \bar{z} = n^{-1}\sum_{j=1}^n z_j,$$

is valid, and the equality holds if $z_1 = \cdots = z_n$, hence the sum $\sum_j \psi(z_j)$ is maximum at $z_1 = \cdots = z_n = n^{-1}\log\alpha$. □

THEOREM 2. *If the condition C_r is satisfied, the homogeneous threshold $\boldsymbol{x} = \boldsymbol{a} = (a, \dots, a)$ such that*

$$1 - F^n(a) = \alpha,$$

maximizes globally $p(R(1,n); \boldsymbol{x}; G)$ under the constraint $p(R(1,n); \boldsymbol{x}; F) = \alpha$, for any $n = 2, 3, \dots$ and any $0 < \alpha < 1$.

The proof is similar to that of Theorem 1 and omitted.

Remarks. 1. In Theorem 1 (or 2), if the homogeneous threshold is globally optimal for any $n = 2, 3, \ldots$ and any $0 < \alpha < 1$, the condition C_h (or C_r) is necessary because the function $\psi(z)$ in the proof should be concave.

2. Note that Theorems 1 and 2, as well as the following Theorem 3, can be extended to the one sided test (1.5), as shown in the beginning of Section 3.

2.2 *Optimization of a symmetric function*

We state a lemma on the optimization of a symmetric function with a symmetric equality constraint, which is the basis of Theorem 3 on the optimal threshold of the k-out-of-n rule.

Let $\mathcal{X}$ be an open region in $\mathcal{R}^n$, and let $f : \mathcal{X} \to \mathcal{R}$ and $g : \mathcal{X} \to \mathcal{R}$ be C^2 and symmetric, that is, invariant with respect to the permutation of the variables. The problem is to optimize g under the constraint $f(\boldsymbol{x}) = b$, where b is a constant. Let the Lagrangian be denoted by

$$\Phi(\boldsymbol{x}, \lambda) := g(\boldsymbol{x}) + \lambda(b - f(\boldsymbol{x})). \tag{2.3}$$

Because of the symmetry of g and f, the equation

$$\frac{\partial}{\partial x_j}\Phi(\boldsymbol{x}, \lambda) = \frac{\partial}{\partial x_j} g(\boldsymbol{x}) - \lambda \frac{\partial}{\partial x_j} f(\boldsymbol{x}) = 0, \qquad j = 1, \ldots, n, \tag{2.4}$$

has a solution $\boldsymbol{x} = \boldsymbol{a} := (a, \ldots, a)$ and $\lambda = \lambda_0 := \frac{\partial}{\partial x_1} g(\boldsymbol{a}) / \frac{\partial}{\partial x_1} f(\boldsymbol{a})$, provided that $\boldsymbol{a}$, such that $f(\boldsymbol{a}) = b$, is in $\mathcal{X}$. The following lemma is a necessary and sufficient condition for the stationary point $\boldsymbol{a}$ to be locally maximum or minimum.

LEMMA. *The object function g is locally minimum at $\boldsymbol{x} = \boldsymbol{a}$ under the constraint $f(\boldsymbol{x}) = b$, if and only if the following inequality holds at $(\boldsymbol{x}, \lambda) = (\boldsymbol{a}, \lambda_0)$.*

$$\frac{\partial^2 g}{\partial x_1^2} - \frac{\partial^2 g}{\partial x_1 \partial x_2} - \lambda \left(\frac{\partial^2 f}{\partial x_1^2} - \frac{\partial^2 f}{\partial x_1 \partial x_2} \right) > 0. \tag{2.5}$$

If the reverse inequality holds, g is locally maximized.

Proof is given in the Appendix A.

2.3 *Rules k-out-of-n*

In the general case $1 < k < n$, we will show the local optimality of a homogeneous threshold, assuming a general condition on the distributions.

For a pair of distributions (F, G) we introduce the following conditions.

(i) F and G are C^2. (ii)

$$C(k,n) : \frac{g^{n-1}(y)}{\bar{G}^{k-1}(y) G^{n-k}(y)} \Big/ \frac{f^{n-1}(y)}{\bar{F}^{k-1}(y) F^{n-k}(y)}, \quad 1 < k < n, \quad n = 3, 4, \ldots \tag{2.6}$$

is a non-decreasing function of y.

Note that $C(n,n)$ and $C(1,n)$ are equivalent to C_h and C_r, respectively. Note also that C_h and C_r imply $C(k,n)$.

THEOREM 3. *Assume the above conditions* (i), *the homogeneous threshold* $\boldsymbol{x} = \boldsymbol{a} = (a, \dots, a)$ *such that*

$$p(R(k,n); \boldsymbol{a}; F) = \alpha,$$

maximizes locally $p(R(k,n); \boldsymbol{x}; G)$ *under the constraint* $p(R(k,n); \boldsymbol{x}; F) = \alpha$, *if and only if* $C(k,n)$ *holds.*

Proof is given in the Appendix B.

2.4 *Bayes solution*

Returning to the design of a monitor, let L_0 denote the loss due to a false alarm (the error of the first kind) at $\mathrm{H}_0 : F$, and L_1 that due to a lack of alarm (the error of the second kind) at $\mathrm{H}_1 : G$. Furthermore let us assume that the safe state $\mathrm{H}_0 : F$ happens at a crucial time with a prior probability $\omega, 0 < \omega < 1$, and the dangerous state $\mathrm{H}_1 : G$ with a prior probability $1 - \omega$. That is, the Bayes risk, R, F, G being fixed, is

$$\begin{aligned} r &= r(\omega, L_0, L_1, \boldsymbol{x}) \\ &= \omega L_0 p(R; \boldsymbol{x}; F) + (1-\omega) L_1 (1 - p(R; \boldsymbol{x}; G)) \\ &= -(1-\omega) L_1 \left(p(R; \boldsymbol{x}; G) + \frac{\omega L_0}{(1-\omega) L_1} (\alpha - p(R; \boldsymbol{x}; F)) \right) + \omega L_0 \alpha. \end{aligned} \tag{2.7}$$

Hence, minimizing r with respect to $\boldsymbol{x}$ is equivalent to maximizing the Lagrangian

$$p(R; \boldsymbol{x}; G) + \lambda(\alpha - p(R; \boldsymbol{x}; F)), \qquad \lambda = \omega L_0 / (1-\omega) L_1.$$

Due to Theorem 3, the solution is $\boldsymbol{x} = \boldsymbol{a} = (a, \dots, a)$, where a is determined, as is shown in the Appendix B, from

$$\begin{aligned} \lambda = \lambda(a) &= \frac{\partial}{\partial x_1} p(R; \boldsymbol{x}; G) \Big/ \frac{\partial}{\partial x_1} p(R; \boldsymbol{x}; F) \\ &= g(a) \bar{G}^{k-1}(a) G^{n-k}(a) / f(a) \bar{F}^{k-1}(a) F^{n-k}(a). \end{aligned} \tag{2.8}$$

Hence, the following proposition is justified.

PROPOSITION 1. *Under the conditions* (i) *and* (ii) *of Subsection* 2.3, *the Bayes risk* (2.7) *is locally minimized by choosing* $\boldsymbol{x} = \boldsymbol{a}$ *such that the right-hand side of* (2.8) *is equal to* $\omega L_0/(1-\omega)L_1$. *Such* $\boldsymbol{a}$ *is uniquely determined if the right-hand side of* (2.8) *is strictly increasing.*

A sufficient condition for the uniqueness of the solution for any k is, in addition to $C(k,n)$, that both $\bar{G}(y)/\bar{F}(y)$ and $G(y)/F(y)$ are increasing in y, or for example that $g(y)/f(y)$ is increasing (a stronger property).

Table 1. Orders of typical families of distributions

distr. on $(-\infty, \infty)$	parameter	distr. on $(0, \infty)$	parameter	orders
logistic	location	Pareto	scale	C_h and C_r
normal	location	lognormal	scale	C_h and C_r
Gumbel	location	Fréchet	scale	C_h and C_r^*
negative Gumbel	location	Weibull(exponential)	scale	C_h^* and C_r

* denotes constant ratio

3. Examples and counterexamples

3.1 *Examples of increasing C_h and C_r*

In this section we discuss the one-sided test (1.5), and define

$$R_h(y;\theta_0,\theta_1) := (k(y;\theta_1)/\bar{K}(y;\theta_1))\,/(k(y;\theta_0)/\bar{K}(y;\theta_0))$$
$$R_r(y;\theta_0,\theta_1) := (k(y;\theta_1)/K(y;\theta_1))\,/(k(y;\theta_0)/K(y;\theta_0))$$

where $K(y;\theta)$, $\theta \in \Theta \subset \mathcal{R}$, is a cdf and $k(y;\theta)$ is its pdf. For this family of distributions, the condition C_h (or C_r) means $R_h(y;\theta_0,\theta_1)$ (or $R_r(y;\theta_0,\theta_1)$) is non-decreasing in y for any $\theta_0 < \theta_1$. That is, its hazard function (or reverse hazard function) is TP_2 (Total Positivity of order 2, Karlin (1968)).

For example, for the Weibull family of probability distributions with scale parameter (abbreviated as "Weibull (scale)"), $\bar{K}(y;\theta) = \exp(-(y/\theta)^\gamma)$, $y > 0$, $\theta > 0$, we get $R_h(y;\theta_0,\theta_1) = (\theta_0/\theta_1)^\gamma = \text{const}$ and $R_r(y;\theta_0,\theta_1)$ is increasing in y. Hence, the power of n-out-of-n rule is the same for any choice of $(x_1,\ldots,x_n)$ such that $\prod_{j=1}^n \bar{F}(x_j;\theta_0) = \alpha$. A homogeneous threshold maximizes globally the power of 1-out-of-n rule for any α, θ_0 and θ_1. Further, since $C(k,n)$, $1 \le k \le n$, is satisfied, a homogeneous threshold maximizes locally $R(k,n)$ test power for $1 \le k \le n$.

Table 1 is a list of typical parametric families which are both increasing C_h and C_r. Therefore, they are also $C(k,n)$ increasing, $1 < k < n$, $n = 3,4,\ldots$. In this list the other parameters are fixed. The symbol C_h^*(or C_r^*) means $R_h(y;\theta_0,\theta_1)$ (or $R_r(y;\theta_0,\theta_1)$) is constant. Here we note that C_h and C_r properties hold valid under exponential and logarithmic transformations of random variables. The derivation of these properties, as well as those of the following counterexamples, are discussed in detail in a companion paper by Sibuya (1998).

All the families in Table 1 are increasing both C_h and C_r. However an increasing C_h (or C_r) family is not always increasing C_r (or C_h). For example, the Weibull distribution with reciprocal power parameter,

$$\bar{K}(y;\theta) = \exp(-y^\theta), \qquad y > 0; \quad \theta > 0, \tag{3.1}$$

is increasing C_h but not increasing C_r. If a random variable Y follows (3.1), $1/Y$ follows the Fréchet distribution, which is increasing C_r but not increasing C_h.

For the Weibull distribution (3.1) with $\theta_0 < \theta_1$,

$$\bar{K}(y;\theta_1) \gtrless \bar{K}(y;\theta_0) \iff y \lessgtr 1.$$

Hence, the very first assumption on the test (1.4), G is stochastically larger than F, is violated. Further, $k(y;\theta)$ is not TP_2.

3.2 *Decreasing* C_h

The proof of Theorem 1 shows that if a family is "decreasing" C_h the homogeneous threshold has the "minimum" power. An example is the Weibull (scale) with different power parameters,

$$\bar{K}(y;\theta,\gamma) = \exp(-(y/\theta)^{1/\gamma}), \qquad y > 0; \quad \theta > 0, \ \gamma > 0.$$

Since,

$$\log \frac{\bar{K}(y;\theta_1,\gamma_1)}{\bar{K}(y;\theta_0,\gamma_0)} = \frac{y^{1/\gamma_1}}{\theta_0^{1/\gamma_0}} \left(y^{1/\gamma_0 - 1/\gamma_1} - \frac{\theta_0^{1/\gamma_0}}{\theta_1^{1/\gamma_1}} \right),$$

if $\gamma_1 > \gamma_0$,

$$\bar{K}(y;\theta_1,\gamma_1) \geq \bar{K}(y;\theta_0,\gamma_0) \Longleftrightarrow y \geq \left(\frac{\theta_0^{\gamma_1}}{\theta_1^{\gamma_0}} \right)^{1/(\gamma_1-\gamma_0)}.$$

The last expression is a decreasing function of θ_1; a desirable property to test (1.5). Now,

$$R_h(y;\theta_0,\gamma_0,\theta_1,\gamma_1) = \frac{\gamma_0}{\gamma_1} \frac{\theta_0^{1/\gamma_0}}{\theta_1^{1/\gamma_1}} y^{1/\gamma_1 - 1/\gamma_0}$$

is decreasing in y. Hence, in n-out-of-n rule, provided that the homogeneous threshold $x_1 = \cdots = x_n = a$ is greater than 1, and

$$\exp(n/\theta_0^{1/\gamma_0}) > \alpha \Longleftrightarrow \theta_0 > \left(\frac{n}{-\log \alpha} \right)^{\gamma_0},$$

the homogeneous threshold minimizes the power.

3.3 *The role of* TP_2 *pdf*

All the families in Table 1 also have a TP_2 pdf, or monotone likelihood ratio, that is, $k(y;\theta_1)/k(y;\theta_0)$, $\theta_0 < \theta_1$, is non-decreasing in y, which implies that $K(\cdot;\theta)$ is stochastically increasing in θ. However, TP_2 pdf does not imply C_h nor C_r. A counterexample is a family of histogram distributions with the density

$$k(y;\theta) = \begin{cases} 1/(1+\theta+\theta^2), & 0 < y \leq 1, \\ \theta/(1+\theta+\theta^2), & 1 < y \leq 2, \\ \theta^2/(1+\theta+\theta^2), & 2 < y \leq 3, \\ 0, & \text{otherwise}; \quad \theta > 0. \end{cases}$$

It is TP_2, but neither C_h nor C_r. To see what happens about the family, let us restrict ourselves to the 2-out-of-2 rule to test H_0: $\theta = \theta_0$ vs. H_1: $\theta > \theta_0$.

Let $P_0(\theta)$ denote the power for the threshold $x_1 = x_2 = a$, $1 < a < 2$, and $P_1(\theta)$ denote that for $1 < x_1 < a < x_2 = 2$. These conditions restrict α, the level, within the range

$$\left(\frac{\theta_0^2}{1+\theta_0+\theta_0^2} \right)^2 < \alpha < \frac{\theta_0^2(\theta_0+\theta_0^2)}{(1+\theta_0+\theta_0^2)^2}.$$

Now,

$$P_0(\theta) = (\theta^2 + (2-a)\theta)^2/(1+\theta+\theta^2)^2$$
$$P_1(\theta) = (\theta^2 + (2-x_1)\theta)\theta^2/(1+\theta+\theta^2)^2$$

and

$$P_0(\theta_0) = P_1(\theta_0) = \alpha.$$

From the last condition, we get

$$2 - x_1 = (2\theta_0 + 2 - a)(2 - a)/\theta_0,$$

and

$$P_1(\theta) = P_0(\theta)\frac{\theta^2 + (2 - x_1)\theta}{(\theta + (2 - a))^2} > P_0(\theta)$$

if $\theta > \theta_0$. Hence, a homogeneous threshold may not be optimal in the 2-out-of-2 rule unless C_h is satisfied. Incidentally, $R_h(y;\theta_0,\theta_1)$, $\theta_0 < \theta_1$, is constant on $2 < y < 3$. The power for the threshold $x_1 = x_2 = a$, $2 < a < 3$, with

$$\alpha > \frac{\theta_0^2(\theta_0 + \theta_0^2)}{(1 + \theta_0 + \theta_0^2)^2},$$

is exactly the same as the power for any thresholds $2 < x_1 < a < x_2 < 3$. Also, this family can be used to show the case of the 1-out-of-2 rule and that C_r is not satisfied.

Hence, the following proposition is obtained,

PROPOSITION 2. *The monotone likelihood ratio condition is not sufficient to imply that a homogeneous threshold is optimal for all rules k-out-of-n, $1 < k < n$, and for all α and $\theta_0 < \theta_1$.*

3.4 *Restricted level of test*

Even if C_h does not hold, the homogeneous threshold may have the maximum power for the restricted value of α, depending on n. This happens if $R_h(y;\theta_0,\theta_1)$ is increasing for larger values of y, and for smaller values of α.

An example is a bilateral exponential or Laplace distribution (location), $k(y;\theta) = \frac{1}{2}\exp(-|y-\theta|)$. $R_h(y;\theta_0,\theta_1)$ is decreasing if $y < \theta_0$, increasing if $\theta_0 < y < \theta_1$, and equal to 1 if $y > \theta_1$. Consider the n-out-of-n rule with the level, α, to test (1.5). For $\alpha < 2^{-n}$, the homogeneous threshold $x_1 = \cdots = x_n = a$ which satisfies

$$P_0(\theta_0) = \left(\frac{1}{2}e^{-(a-\theta_0)}\right)^n = \alpha,$$

is given by $a = \theta_0 - \log 2\alpha^{1/n} > 0$. Hence the homogeneous threshold is optimal and the maximum power $P(\theta)$ is given by

$$P(\theta) = \begin{cases} \alpha e^{n(\theta-\theta_0)}, & \theta \le a \\ \left(1 - \dfrac{1}{4\alpha^{1/n}}e^{-(\theta-\theta_0)}\right)^n, & \theta > a. \end{cases}$$

For $\alpha > 2^{-n}$, the homogeneous threshold a is smaller than θ_0, and it is not optimal. In the 2-out-of 2 rule with $\alpha > 1/4$, numerical results show that the thresholds $x_1 < a < x_2 = \theta_0$ are better.

Incidentally, the pdf of the bilateral exponential family is TP_2, and this is another example that a TP_2 pdf does not imply C_h nor C_r increasing.

From the above discussion, we get

PROPOSITION 3. *For the restricted values of α, a homogeneous threshold may have the maximum power for the n-out-of-n rule without the condition of increasing C_h. The restriction of α depends on n.*

4. Some related problems

4.1 *Optimization of the rule R given $\boldsymbol{x}$*

In this subsection we discuss briefly the optimization of R given (F, G), α and $\boldsymbol{x}$. The optimal R is, by the Neyman-Pearson fundamental lemma,

$$R = \left\{ \boldsymbol{t} = (t_1, \ldots, t_n) : \prod_{j=1}^{n} \left(\frac{\bar{G}(x_j)}{G(x_j)} \Big/ \frac{\bar{F}(x_j)}{F(x_j)} \right)^{t_j} \geq c \right\} \subset \{0,1\}^n. \tag{4.1}$$

The number c is determined by α. The following discussions will give an insight on the structure of R.

Without loss of generality, we can order $(x_1, \ldots, x_n)$ such as

$$r_1 \geq r_2 \geq \cdots \geq r_n, \qquad r_j := \frac{\bar{G}(x_j)}{G(x_j)} \Big/ \frac{\bar{F}(x_j)}{F(x_j)}.$$

Recall that we have assumed

$$F(y) \geq G(y) \Leftrightarrow \bar{F}(y) \leq \bar{G}(y) \Leftrightarrow \frac{\bar{F}(y)}{F(y)} \leq \frac{\bar{G}(y)}{G(y)} \Rightarrow \frac{\bar{G}(x_j)}{G(x_j)} \Big/ \frac{\bar{F}(x_j)}{F(x_j)} \geq 1.$$

Hence, $t_j = 1$ for the smaller j means the larger product.

Let $\boldsymbol{t}$ and $\boldsymbol{s}$ be elements in $\{0,1\}^n$. We define that the former is "directly larger" than the latter, denoted by $\boldsymbol{t} \succ \boldsymbol{s}$, if and only if

$$t_i = s_i, \forall i (i \neq j, j+1), \quad \text{and} \quad t_j = s_{j+1} = 1, \quad t_{j+1} = s_j = 0, \quad \text{or}$$
$$t_i = s_i, \forall i < n, \quad \text{and} \quad t_n = 1, \quad s_n = 0.$$

This definition introduces a lattice in $\{0,1\}^n$, which depends on (F, G) only through $(r_1, \ldots, r_n)$. The order $\boldsymbol{t} \succ \boldsymbol{s}$ is extended to the lattice. This means

$$\prod_{j=1}^{n} r_j^{t_j} \geq \prod_{j=1}^{n} r_j^{s_j}. \tag{4.2}$$

$R(k,n)$ is an upper set in this lattice, i.e., for any $\boldsymbol{s} \in R(k,n)$, $\boldsymbol{t} \in R(k,n)$ if $\boldsymbol{t} \succ \boldsymbol{s}$. The least elements of $R(k,n)$ is $\boldsymbol{t}_0 = (0, \ldots, 0, 1, \ldots, 1)$ with the last k components 1, and the greatest elements outside of $R(k,n)$ is $\boldsymbol{s}_0 = (1, \ldots, 1, 0, \ldots, 0)$ with the first $k-1$ components 1. $R(k,n)$ is more powerful than any upper sets, which does not include $R(k,n)$, if and only if $\boldsymbol{t}_0 \succ \boldsymbol{s}_0$, or

$$\prod_{j=n-k+1}^{n} r_j \geq \prod_{j=1}^{k-1} r_j. \tag{4.3}$$

If the threshold is homogeneous (4.3) is satisfied for any k. Assaf *et al.* (1986) discussed the optimality of the k-out-of-n rules including the condition (4.3) along a similar line. To compare the rules, however, we have to change $\boldsymbol{x}$ to keep the test level the same. Hence the discussion of this subsection cannot be applied.

4.2 *Comparison among k-out-of-n rules*

In this subsection we discuss the choice of k of the k-out-of-n monitor. First, to compare rules k-out-of-3, $k = 1, 2, 3$, with homogeneous thresholds, the powers for the families of distributions in Table 1 are numerically examined. It turns out that there are two types of families summarized as follows. It looks difficult to develop a general theory.

A. normal (location)-lognormal (scale), logistic (location)-Pareto (scale).

The powers cannot be compared among $k = 1, 2, 3$. The order changes by both α and θ.

B1. negative Gumbel (location)-Weibull, exponential (scale).

The powers are in the order $k = 1, 2, 3$; $k = 3$ is uniformly most powerful.

B2. Gumbel (location)-Fréchet(scale).

The powers are in the order $k = 3, 2, 1$; $k = 1$ is uniformly most powerful.

The pairs are exponential and logarithmic transformations between random variables, and C_h and C_r properties as well as power relations are the same. B1 and B2 are a negation relationship, which will be discussed in the following part of this subsection in relation with symmetric tests.

In a special symmetric condition on (F, G) and on $\boldsymbol{x}$, the performance of the n-out-of-n rule is the same as 1-out-of-n rule. We use the d.f. K, which is introduced in the test problem (1.5), and put $F(y) = K(y - \theta_0)$ and $G(y) = K(y - \theta_1)$, $\theta_0 < \theta_1$. Let the optimal threshold of the n-out-of-n rule be $\boldsymbol{x} = (a_n, \ldots, a_n)$ and that of the 1-out-of-n rule be $\boldsymbol{x} = (a_1, \ldots, a_1)$.

PROPOSITION 4. *Assume that* (i) $K(-y) = 1 - K(y), -\infty < y < \infty$, (ii) *The levels of both rules are the same*: $\bar{F}^n(a_n) = 1 - F^n(a_1) =: \alpha$, (iii) *The parameters are specified so that* $\bar{G}^n(a_n) = 1 - \alpha$, *which is satisfied if* $\theta_0 + \theta_1 = a_1 + a_n$. *Then* $1 - G^n(a_1) = 1 - \alpha$, *that is*, $p(R(n, n); a_n\mathbf{1}; G) = p(R(1, n); a_1\mathbf{1}; G)$.

The proof is simple and omitted. An analogous proposition holds, if a d.f. K is such that $K(0) = 0, K(1/x) = 1 - K(x)$, for $F(y) = K(y/\theta_0)$ and $G(y) = K(y/\theta_1), 0 < \theta_0 < \theta_1$. The proposition is also extended to the equal performance of $R(k, n)$ and $R(n - k + 1, n)$.

4.3 *Optimal reliability in different time points*

The result of the present paper can be applied for maximizing a reliability of a k-out-of-n system at a time subject to the reliability at a different time point as follows. Suppose a k-out-of-n system is composed of components of which lifetime, Y_j, follows cdf $K(y_j/\theta_j)$, $j = 1, \ldots, n$. For two different time points, $\tau_0 < \tau_1$, the reliabilities of j-th component are given by

$$(4.4) \quad p_j(\tau_0) := P\{Y_j > \tau_0\} = \bar{K}(\tau_0/\theta_j) = \bar{K}(\xi_j/\theta_a) \quad \text{where} \quad \xi_j = \tau_0 \cdot \theta_a/\theta_j, \text{ and}$$

$$(4.5) \quad p_j(\tau_1) := P\{Y_j > \tau_1\} = \bar{K}(\tau_1/\theta_j) = \bar{K}(\xi_j/\theta_b) \quad \text{where} \quad \theta_b = \theta_a \cdot \tau_0/\tau_1 < \theta_a.$$

Hence, the problem of maximizing the system reliability at τ_1, subject to the reliability at τ_0, corresponds to an optimization problem of $(\xi_1, \ldots, \xi_n)$ in the statistical testing;

$$\mathrm{H}_0 : \theta_a \quad \text{vs.} \quad \mathrm{H}_1 : \theta_b \quad (\theta_b < \theta_a)$$

which is solved in the present paper. This model was discussed by Boland and Proschan (1994).

Appendix

A. *Proof of Lemma in Subsection* 2.2

Because of the symmetry of f, the gradient of f at $\boldsymbol{a}$ is $\frac{\partial}{\partial \boldsymbol{x}} f(\boldsymbol{a}) = \frac{\partial}{\partial x_1} f(\boldsymbol{a})\mathbf{1}$. Hence the tangent hyperplane of the surface $b - f(\boldsymbol{x}) = 0$ at $\boldsymbol{a}$ is $T = \{\boldsymbol{v} : \boldsymbol{v}^t\mathbf{1} = 0\}$ in the local coordinate system with the origin at $\boldsymbol{a}$.

Let Ψ be the Hessian matrix of the Lagrangean function $\Phi(\boldsymbol{x}, \lambda)$ at $(\boldsymbol{a}, \lambda_0)$. g is locally minimum at $\boldsymbol{a}$, under the constraint $f(\boldsymbol{x}) = b$, if and only if

$$\boldsymbol{v}^t\Psi\boldsymbol{v} \geq 0, \qquad \forall \boldsymbol{v} \in T.$$

The Hessian matrix Ψ has the identical diagonal elements and the identical off-diagonal elements, and hence has the form

$$\Psi = \frac{\partial^2\Phi(\boldsymbol{a}, \lambda_0)}{\partial x_1^2} I + \frac{\partial^2\Phi(\boldsymbol{a}, \lambda_0)}{\partial x_1 \partial x_2}(J - I),$$

where I is the identity matrix and $J = \mathbf{1}\mathbf{1}^t$ is the square matrix whose elements are all one. The quadratic differential form of $\Phi(\boldsymbol{x}, \lambda)$ on T is, therefore,

$$\boldsymbol{v}^t\Psi\boldsymbol{v} = \left(\frac{\partial^2\Phi}{\partial x_1^2} - \frac{\partial^2\Phi}{\partial x_1 \partial x_2}\right) \|\boldsymbol{v}\|^2,$$

since $J\boldsymbol{v} = \mathbf{0}$. The coefficient of $\|\boldsymbol{v}\|^2$ determines maximality or minimality.

See, e.g., Craven (1981), Theorems 3.5 and 3.6, or Fleming (1976), Section 4.8, for the argument in the proof.

B. *Proof of Theorem* 3 *in Subsection* 2.3

Recall that

$$p(\boldsymbol{x}; F) := p(R(k, n); \boldsymbol{x}; F) = \sum_{j=k}^{n} \sum_{\sum t_i = j} \prod_{i=1}^{n} \bar{F}^{t_i}(x_i) F^{1-t_i}(x_i).$$

We apply Lemma in Subsection 2.2 to maximize $p(\boldsymbol{x}, G)$ under the constraint $p(\boldsymbol{x}, F) = \alpha$. Firstly differentiate only the term with $\sum t_i = j$ to obtain

$$\frac{\partial}{\partial x_1}\left(\bar{F}(x_1) \sum_{\sum t_i = j-1} + F(x_1) \sum_{\sum t_i = j}\right) \prod_{i=2}^{n} \bar{F}^{t_i}(x_i) F^{1-t_i}(x_i)$$
$$= f(x_1)\left(-\sum_{\sum t_i = j-1} + \sum_{\sum t_i = j}\right) \prod_{i=2}^{n} \bar{F}^{t_i}(x_i) F^{1-t_i}(x_i).$$

Further differentiation with respect to x_1 results in the same form with $f(x_1)$ replaced by $f'(x_1)$. Further differentiation with respect to x_2 results in

$$\frac{\partial^2}{\partial x_1 \partial x_2}\left(\bar{F}(x_1)\bar{F}(x_2) \sum_{\sum t_i = j-2} + (\bar{F}(x_1)F(x_2) + F(x_1)\bar{F}(x_2))\right.$$

$$\cdot \sum_{\sum t_i=j-1} + F(x_1)F(x_2) \sum_{\sum t_i=j} \Bigg) \prod_{i=3}^{n} \bar{F}^{t_i}(x_i) F^{1-t_i}(x_i)$$

$$= f(x_1)f(x_2) \left(\sum_{\sum t_i=j-2} -2 \sum_{\sum t_i=j-1} + \sum_{\sum t_i=j} \right) \prod_{i=3}^{n} \bar{F}^{t_i}(x_i) F^{1-t_i}(x_i).$$

Now, put $\boldsymbol{x} = \boldsymbol{a}$ and sum over $j = k, \ldots, n$ to obtain (the argument $\boldsymbol{a}$ being omitted)

$$\begin{aligned} D_1 &:= \frac{\partial}{\partial x_1} p(\boldsymbol{a}; F) \\ &= \sum_{j=k}^{n} f \left(- \binom{n-1}{j-1} \bar{F}^{j-1} F^{n-j} + \binom{n-1}{j} \bar{F}^{j} F^{n-j-1} \right) \\ &= -f \binom{n-1}{k-1} \bar{F}^{k-1} F^{n-k}, \\ D_{11} &:= \frac{\partial^2}{\partial x_1^2} p(\boldsymbol{a}; F) = -f' \binom{n-1}{k-1} \bar{F}^{k-1} F^{n-k}, \end{aligned}$$

and

$$\begin{aligned} D_{12} &:= \frac{\partial^2}{\partial x_1 \partial x_2} p(\boldsymbol{a}; F) = \sum_{j=k}^{n} f^2 \Bigg(\binom{n-2}{j-2} \bar{F}^{j-2} F^{n-j} - 2 \binom{n-2}{j-1} \bar{F}^{j-1} F^{n-j-1} \\ &\qquad\qquad + \binom{n-2}{j} \bar{F}^{j} F^{n-j-2} \Bigg) \\ &= f^2 \left(\binom{n-2}{k-2} \bar{F}^{k-2} F^{n-k} - \binom{n-2}{k-1} \bar{F}^{k-1} F^{n-k-1} \right). \end{aligned}$$

The same expressions are obtained for $p(\boldsymbol{a}; G)$. Applying Lemma, we obtain a sufficient condition

$$\begin{aligned} &\binom{n-1}{k-1} \frac{g'}{g} + g \left(\binom{n-2}{k-2} \frac{1}{\bar{G}} - \binom{n-2}{k-1} \frac{1}{G} \right) \\ &\qquad > \binom{n-1}{k-1} \frac{f'}{f} + f \left(\binom{n-2}{k-2} \frac{1}{\bar{F}} - \binom{n-2}{k-1} \frac{1}{F} \right), \end{aligned}$$

or

$$(n-1)\frac{g'}{g} + (k-1)\frac{g}{\bar{G}} - (n-k)\frac{g}{G} > (n-1)\frac{f'}{f} + (k-1)\frac{f}{\bar{F}} - (n-k)\frac{f}{F}.$$

This condition is equivalent to

$$\frac{d}{dy} \log \left(\frac{g^{n-1}(y)}{\bar{G}^{k-1}(y) G^{n-k}(y)} \Big/ \frac{f^{n-1}(y)}{\bar{F}^{k-1}(y) F^{n-k}(y)} \right) > 0,$$

which is the condition $C(k, n)$.

Boland and Proschan (1983) computed D_{12} in a different way. His computation shows why the summation in D_{12} reduces to the term $j = k$.

References

Assaf, D., Ben-Dov, Y. and Krieger, A. M. (1986). Optimal design of systems subject to two types of error, *Operations Research*, **34**(4), 550–553.

Boland, P. J. (1998). A reliability comparison of basic systems using hazard rate functions, *Appl. Stochastic Models and Data Anal.*, **13**, 377–384.

Boland, P. J. and Proschan, F. (1983). The reliability of k-out-of-n systems, *Ann. Probab.*, **11**, 760–764.

Boland, P. J. and Proschan, F. (1994). Stochastic order in system reliability theory, *Stochastic Orders and Their Applications* (eds. M. Shaked and J. G. Shanthikumar), 465-508, Academic Press, San Diego, California.

Craven, B. D. (1981). *Functions of Several Variables*, Chapman and Hall, London.

Fleming, W. (1976). *Functions of Several Variables*, 2nd ed., Springer, New York.

Gleser, L. (1975). On the distribution of the number of successes in independent trials, *Ann. Probab.*, **3**(1), 182–188.

Kalashnikov, V. V. and Rachev, S. T. (1986) Characterization of queuing models and their stability, *Probability Theory and Mathematical Statistics*, Vol. 2 (eds. Prohorov, Yu. V., Statulevicius, V. A., Sazonov, V. V. and Grigelionis, B.), 37–53, VNU Science Press, Utrecht, the Netherland.

Karlin, S. (1968). *Total Positivity*, Stanford University Press, Stanford, California.

Kohda, T., Kumamoto, H., Inoue, K. and Takami, I. (1982). Optimal structure of sensor systems composed of nonidentical sensors, *Microelectronics and Reliability*, **22**(3), 445–456.

Lešanovský, A. (1993). Systems with two dual failure modes—A survey, *Microelectronics and Reliability*, **33**(10), 1597–1626.

Phillips, M. J. (1980). k-out-of-n: G systems are preferable, *IEEE Transactions on Reliability*, **R-29**(2), 166–169.

Sengupta, D. and Deshpande, J. V. (1994). Some results on the relative ageing of two life distributions, *J. Appl. Probab.*, **31**, 991–1003.

Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic Orders and Their Applications,* Academic Press, San Diego, California.

Sibuya, M. (1998). Increasing hazard function ratio and related orders of probability distributions (*submitted*).

Suzuki, K. and Tachikawa, K. (1995). Optimal thresholds of two monitors in condition monitoring maintenance, *Journal of the Japanese Society for Quality Control*, **25**(2), 191–197 (in Japanese).