The Annals of Mathematical Statistics
1969, Vol. 40, No. 5, 1778–1783

MINIMAX RESULTS FOR IFRA SCALE ALTERNATIVES[1]

BY KJELL DOKSUM

University of California, Berkeley

1. Introduction and summary. Recent results by Birnbaum, Esary and Marshall (1966), Barlow and Proschan (1967) and others suggest that the exponential models used be Epstein and Sobel (1953) and others for life testing problems should be extended to models in which the lifetimes have increasing failure rate average (IFRA) distributions. In this paper, IFRA scale models are dealt with. Consider two independent random samples $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ from populations with distributions $F(x)$ and $G(y) = F(\Delta y)$ respectively, where F is a continuous, unknown, IFRA distribution. The null hypothesis $H_0: \Delta \leqq 1$ is to be tested against the alternative $H_1: \Delta > 1$. Since F is unknown, it is not possible to maximize the power $E[\phi \mid F(\cdot), F(\Delta\cdot)]$, $\Delta > 1$. However, it is shown (Theorems 2.1, 2.2, 3.1, 3.2) that the tests that maximize the power for the exponential alternative $F(\Delta y) = 1 - \exp(-\Delta y)$ actually maximize the minimum power $\inf_{F(\cdot)} E[\phi \mid F(\cdot), F(\Delta\cdot)]$. Thus these tests are minimax. They have been computed by Lehmann (1953), Savage (1956), and Rao, Savage and Sobel (1960). The results indicate that in the case of uncensored samples, one should use one of the statistics

$$L = \prod_{i=1}^{n} (N + i - s_{n+1-i}), \quad \text{or}$$

$$S = \sum_{i=1}^{m} J_0(r_i), \quad \text{with} \quad J_0(k) = \sum_{j=N+1-k}^{N} 1/j,$$

where $N = m + n$ and $r_1, \cdots, r_m$ $(s_1, \cdots, s_n)$ are the ordered ranks of the X's (Y's) in the combined sample of X's and Y's. L is minimax for Δ in an interval about two, and S is minimax for Δ in an interval $(1, \delta)$ to the right of one.

The minimax statistics in the case of censored samples are more complicated (see (3.1) and (3.2)) and one might use one of the approximations suggested by Gastwirth (1965) or Basu (1967) (see (3.3)).

Only finite sample size properties are dealt with. Asymptotic results are given in [6].

2. Minimax tests in the two-sample case. The failure rate of a distribution F with density f is defined to be $q(x) = f(x)/[1 - F(x)]$, and the failure rate average is $A(x) = x^{-1} \int_0^x q(x)\, dx = -\log[1 - F(x)]/x$, $x > 0$, $F(0) = 0$. Thus F is said to be an IFRA distribution if $F(0) = 0$ and

$$(2.1) \qquad -\log[1 - F(x)]/x \quad \text{is non-decreasing in} \quad x > 0.$$

Received 7 August 1968; revised 20 February 1969.

[1] This research has been partially supported by the Office of Naval Research under Contracts Nonr-222(83) and Nonr-3656(18) with the University of California, and NSF-GP-7454. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Similarly, F is a DFRA (decreasing failure rate average) distribution if $F(0) = 0$ and $-\log [1 - F(x)]/x$ is non-increasing in $x > 0$.

Let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ be two independent random samples from populations with distributions F and G. The problem of interest is that of testing the null hypothesis that the X's are stochastically smaller than or equal to the Y's against the alternative that the Y's are stochastically smaller than the X's. However, this model without further restrictions is too general to be used in the derivation of optimal tests; moreover, one would like to have tests that are most likely to reject for alternatives that indicate a definite distinction between the distributions of X and Y. When X and Y are time measurements, then one such model is the scale model in which Y has the same distribution as X/Δ for some $\Delta > 0$. Then $[E(X) - E(Y)]/E(Y) = \Delta - 1$, and $(\Delta - 1)$ measures the relative difference of the mean times. One thus tests $H_0{:}\,\Delta \leqq 1$ vs. $H_1{:}\,\Delta > 1$ and considers the power function $\beta(\phi; F; \Delta) = E_{F,G}(\phi)$ of each test ϕ for the scale alternative with $G(y) = F(\Delta y)$. A test ϕ is said to be monotone if $\phi(x_1, \cdots, x_m, y_1', \cdots, y_n') \leqq \phi(x_1, \cdots, x_m, y_1, \cdots, y_n)$ whenever $y_j \leqq y_j'$ for $j = 1, \cdots, n$.

Note that if V_k is defined to be the number of Y's among the k largest observations in the combined sample, then $(V_1, \cdots, V_N)$ is equivalent to the ordered ranks $(r_1, \cdots, r_m)$ of the X's. Next it will be shown that in the class of level α rank tests, the Lehmann level α test ϕ_Δ defined by

$$\begin{aligned} \phi_\Delta &= 1 \quad \text{if } \textstyle\prod_{k=1}^{N} [k + (\Delta - 1)V_k]^{-1} \geqq c_\alpha \\ &= 0 \quad \text{otherwise;} \end{aligned} \tag{2.2}$$

maximizes the minimum power $\inf_F \beta(\phi; F; \Delta)$ for the IFRA scale model whenever $\Delta > 1$. Here, a rank test ϕ is said to be of *level* α if $\beta(\phi; F; 1) = E(\phi \mid (F, F)) = \alpha$ for all continuous distribution functions F.

THEOREM 2.1. *The Lehmann test ϕ_Δ is minimax for the* IFRA *scale alternative in the sense that for each scale parameter $\Delta > 1$ and for F ranging over the class of continuous* IFRA *distributions,*

$$\inf_F \beta(\phi; F; \Delta) \leqq \inf_F \beta(\phi_\Delta; F; \Delta) \tag{2.3}$$

for each level α rank test ϕ. Moreover ϕ_Δ is the unique minimax rank test in the sense that any other level α rank test satisfying (2.3) *coincides with ϕ_Δ* a.e.

PROOF. Using the Neyman-Pearson Lemma, Lehmann (1953), Savage (1956), and Rao, Savage and Sobel (1960, Corollary 3.4) have essentially shown that ϕ_Δ maximizes the power for the exponential scale model, i.e. if $K(x) = 1 - \exp(-x)$, then for each $\Delta > 1$,

$$\beta(\phi; K; \Delta) \leqq \beta(\phi_\Delta; K; \Delta) \tag{2.4}$$

for all level α rank tests ϕ. On the other hand, the definition (2.2) shows that ϕ_Δ is a monotone test, thus since $K^{-1}(x) = -\log(1 - x)$, then the definition of the IFRA property and Theorem 3.1 of [7] implies that ϕ_Δ attains its minimum power for the exponential scale model, i.e.,

$$\inf_F \beta(\phi_\Delta; F; \Delta) = \beta(\phi_\Delta; K; \Delta). \tag{2.5}$$

Since K is an IFRA distribution, then

$$\inf_F \beta(\phi; F; \Delta) \leqq \beta(\phi; K; \Delta) \tag{2.6}$$

and (2.3) follows. Uniqueness holds since if ψ is a level α rank test satisfying (2.3), then the above shows that it must satisfy

$$\beta(\phi; K; \Delta) \leqq \beta(\psi; K; \Delta) \tag{2.7}$$

for all level α rank tests ϕ. Since ϕ_Δ also satisfies this (see (2.3)), then the uniqueness part of the Neyman-Pearson Lemma implies that $\psi = \phi_\Delta$ a.e.

Note that (ϕ_Δ, K) is a saddle point, i.e.,

$$\sup_\phi \beta(\phi; K; \Delta) = \beta(\phi_\Delta; K; \Delta) = \inf_F \beta(\phi_\Delta; F; \Delta), \qquad \Delta > 1. \tag{2.8}$$

In order to be able to use the minimax test ϕ_Δ, one must choose a value of Δ. Savage (1956) suggests using the level α test ϕ_s that maximizes the power for the exponential scale model for Δ in a neighborhood $(1, \delta)$ to the right of one, i.e., when the relative difference of the means of X and Y is close to zero (and positive). This test is defined by

$$\begin{aligned} \phi_s &= 1 \quad \text{if} \quad \textstyle\sum_{i=1}^m J_0(r_i) \geqq c_\alpha', \\ &= 0 \quad \text{otherwise}; \end{aligned} \tag{2.9}$$

where $J_0(k) = \sum_{j=N+1-k}^N 1/j$.

Note that ϕ_s equivalently can be defined to reject for $\sum_1^N V_k/k \leqq c_\alpha''$. It can now be shown that the Savage test ϕ_s is minimax for Δ in an interval of the form $(1, \delta)$.

THEOREM 2.2 *The Savage test ϕ_s is locally minimax for the* IFRA *scale alternative in the sense that there exists $\delta > 1$ such that for F ranging over the class of continuous* IFRA *distributions,*

$$\inf_F \beta(\phi; F; \Delta) \leqq \inf_F \beta(\phi_s; F; \Delta) \tag{2.10}$$

for each level α rank test ϕ and for all Δ in the interval $(1, \delta)$. Moreover, ϕ_s is the unique locally minimax rank test in the sense that any other level α rank test satisfying (2.10) *coincides with ϕ_s* a.e.

PROOF. Savage (1956) has essentially shown that there exists $\delta > 1$ such that

$$\beta(\phi; K; \Delta) \leqq \beta(\phi_s; K; \Delta) \tag{2.11}$$

for all level α tests ϕ and all Δ in $(1, \delta)$. Since ϕ_s is monotone, the remainder of the proof is as the proof of Theorem 2.1.

REMARKS. (i) The Savage test maximizes the minimum power when the relative difference of the means of X and Y, i.e., $(\Delta - 1)$, is close to zero. In most situations, it would be better to use the test that is optimal when $(\Delta - 1)$ is in a neighborhood of some fixed positive number λ (say). Thus one would use the Lehmann test ϕ_Δ defined by (2.2) with $\Delta = \lambda + 1$. For instance, if the relative difference of the means of X and Y is unity, then Lehmann (1953) has shown that $\phi_\Delta = \phi_2$ is equivalent to the test that rejects for large values of

$$\prod_{i=1}^{n} (N + i - s_{n+1-i}). \tag{2.12}$$

Note that in the representation (2.12), the ranks of the stochastically smaller variable, under H_1, must be used.

(ii) Tables of the null distribution of the Savage statistic have been given by Davies (1969) and Hájek (1969). Davies has computed the power of the Savage test and the Lehmann test ϕ_Δ (for various choices of Δ) for the exponential scale alternative. His results indicate that there is no substantial difference in the power of these tests for $\alpha \geqq .01$, $m = n \leqq 10$.

Thus they are all approximately minimax for all values of the scale parameter $\Delta > 1$ ($\alpha \geqq .01$, $m = n \leqq 10$). Since the tests ϕ_2 and ϕ_s defined by (2.12) and (2.9) are the simplest ones, these tests are recommended, ϕ_2 for small and moderate sample sizes, ϕ_s for larger ones. The table of the null distribution of the statistic (2.12) has been partially computed by Davies (1968).

(iii) The uniqueness parts of Theorem 3.1 and 3.2 can be extended as follows. A test ϕ is said to be *distribution-free* (DF) if $\beta(\phi; F, 1) = E(\phi \mid (F, F))$ is independent of F for F continuous. Thus rank tests are DF. For Δ close to one, tests that are not DF have minimum power less than ϕ_Δ and ϕ_s, and they cannot be minimax. To see this, note that if ϕ is of level α, a.e. continuous, and not DF, then there exists a continuous distribution F_0 such that $E[\phi \mid F_0, F_0)] < \alpha = \sup_F E[\phi \mid (F, F)]$. Now for Δ close to one $\beta(\phi; F_0; \Delta) < \alpha$ and ϕ is worse than ϕ_Δ and ϕ_s.

(iv) It is known (Lehmann (1959, page 187), and Bell, Moser and Thompson (1966, page 134)) that if ϕ is a monotone test, then $\beta(\phi; F; \Delta)$ is an increasing function of the scale parameter $\Delta > 0$. This implies that for fixed $\Delta_1 > 1$, ϕ_{Δ_1} is doubly minimax for the scale alternative in the sense that for testing $H_0: \Delta \leqq 1$ against $H_1: \Delta \geqq \Delta_1$, it maximizes $\inf_F [\inf_{\Delta \geqq \Delta_1} \beta(\phi; F; \Delta)] = \inf_{\Delta \geqq \Delta_1} [\inf_F \beta(\phi; F; \Delta)]$. Here, F ranges over the class of continuous IFRA distributions and only level α rank tests are considered. Note that $(1, \Delta_1)$ is an indifference region.

3. Minimax tests based on censored samples. The censored samples considered here arise typically as follows: m objects of one type and n objects of a second type are put on trial at the same time and one waits until a total of $N^* < m + n = N$ of the objects have failed, where N^* is a fixed number. Moreover, the experiment is conducted so that if one waited until all $N = m + n$ objects failed, then the times to failure $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ would be two independent random samples.

Thus the situation is as in Section 2 except that only the first N^* smallest order statistics in the combined sample are observed. Let m^* and n^* denote the total number of X's and Y's observed respectively. Since the unobserved X's and Y's are all larger than the observed ones, it is possible to compute the ordered ranks $r_1 < \cdots < r_{m^*}$ of the X-sample order statistics $X_{(1)} < \cdots < X_{(m^*)}$ in the *un*censored combined sample $X_1, \cdots, X_m; Y_1, \cdots, Y_n$. Rao, Savage and Sobel (1960, Corollary 3.4) have computed the most powerful test ϕ_Δ^* based on $r_1, \cdots, r_{m^*}$ for exponential alternatives. This level α test rejects H_0 if and only if

$$\Delta^{n^*} \prod_{k=1}^{N^*} [A(\Delta) + k + (\Delta - 1)V_k]^{-1} \geqq c_\alpha^* \tag{3.1}$$

where

$$A(\Delta) = (m - m^*) + \Delta(n - n^*)$$

and V_k is as in Section 2. For tests depending only on $r_1, \cdots, r_{m^*}$ one obtains using the arguments of Section 2.

THEOREM 3.1. *The test ϕ_Δ^* is uniquely minimax for the* IFRA *scale alternative in the sense of Theorem* 2.1.

The locally most powerful level α test ϕ_s^* (Rao, Savage and Sobel (1960, Corollary 3.4)) for the exponential scale alternative $G(y) = F(\Delta y) = K(\Delta y)$ rejects H_0 if and only if

$$\sum_{i=1}^{m^*} J_0^*(r_i) + (m - m^*)J_0^*(N^*) - m^* \geqq \hat{c}_\alpha \tag{3.2}$$

where $J_0^*(k) = \sum_{j=N^*-k+1}^{N^*} 1/[j + N - N^*]$.

THEOREM 3.2. *The test ϕ_s^* is uniquely locally minimax for the* IFRA *scale alternative in the sense of Theorem* 2.2.

Again, this is an application of the results of Section 2. Note that these results also can be applied to other (Gastwirth (1965) and Rao, Savage and Sobel (1960)) censoring plans than the one considered here.

REMARK. The tests ϕ_Δ^* and ϕ_s^* are complicated and one may use the approximations suggested by Gastwirth (1965) and Basu (1967). The latter's paper contains tables of rejection points and power for the test based on the statistic

$$\sum_{i=1}^{m^*} J_0(r_i) + (m - m^*)(N - N^*)^{-1} \sum_{k=N^*+1}^{N} J_0(k) - \tfrac{1}{2}N \tag{3.3}$$
$$= \sum_{i=1}^{m^*} J_0(r_i) + (m - m^*)J_0(N^*) - m^* + \tfrac{1}{2}(n - m)$$

where, as in Section 2, $J_0(k) = \sum_{j=N-k+1}^{N} 1/j$.

REFERENCES

[1] BARLOW, R. E. and PROSCHAN, F. (1967). Exponential life test procedures when the distribution has monotone failure rate. *J. Amer. Statist. Assoc.*, **62** 548–560.

[2] BASU, A. P. (1967). On a generalized Savage statistic with applications to life testing. *Ann. Math. Statist.* **39** 1591–1604.

[3] BELL, C. B., MOSER, J. M. and THOMPSON, R. (1966). Goodness criteria for two-sample distribution-free tests. *Ann. Math. Statist.* **37** 133–142.

[4] BIRNBAUM, Z. W., ESARY, J. D. and MARSHALL, A. W. (1966). A stochastic characterization of wear-out for components and systems. *Ann. Math. Statist.* **37** 816–826.

[5] DAVIES, R. B. (1969). The concept of β-optimal tests. Thesis. Univ. of California, Berkeley.

[6] DOKSUM, K. A. (1967). Asymptotically optimal statistics in some models with increasing failure rate averages. *Ann. Math. Statist.* **38** 1731–1739.

[7] DOKSUM, K. A. (1969). Starshaped transformations and the power of rank tests. *Ann. Math. Statist.* **40** 1167–1176.

[8] EPSTEIN, B. and SOBEL, M. (1953). Life testing *J. Amer. Statist. Assoc.* **48** 486–502.

[9] GASTWIRTH, J. L. (1965). Asymptotically most powerful rank tests for the two-sample problem with censored data. *Ann. Math. Statist.* **36** 1243–1247.

[10] HÁJEK, J. (1969). *A Course in Nonparametric Statistics*. Holden-Day, San Francisco.

[11] LEHMANN, E. L. (1953). The power of rank tests. *Ann. Math. Statist.* **24** 23–42.

[12] LEHMANN, E. L. (1959). *Testing Statistical Hypothesis*. Wiley, New York.

[13] RAO, U. V. R., SAVAGE, I. R. and SOBEL, M. (1960). Contributions to the theory of rank order statistics: the two sample censored case. *Ann. Math. Statist.* **31** 415–426.

[14] SAVAGE, I. R. (1956). Contributions to the theory of rank order statistics: two sample case. *Ann. Math. Statist.* **27** 590–616.