EFFECT ON THE MINIMAL COMPLETE CLASS OF TESTS OF CHANGES IN THE TESTING PROBLEM[1]

By D. L. Burkholder

University of Illinois

1. Summary and introduction. A question of interest in connection with many statistical problems is the following: Does a slight change in the problem result in a different answer? Here the effect of changes in the testing problem on the minimal complete class of tests is investigated. The effects of such changes are found to be different for the two families of distributions considered: The discrete multivariate exponential family and the continuous multivariate exponential family. In Section 2, it is shown that with respect to the discrete exponential family, the minimal complete class of tests for a standard testing problem is minimal complete for a wide variety of related problems. In Section 3, an example is given showing that with respect to the continuous exponential family, on the other hand, the minimal complete class of tests for a standard problem is not necessarily minimal complete for a slight variation of this problem. Tests that are admissible for the standard problem are not necessarily admissible for the variation.

Partly in a general decision theoretic framework and partly with respect to specific examples, Hoeffding [2] has discussed the effect of changes in the family of probability distributions on the minimax solution and other optimal solutions. He has also given key references to the extensive literature on the performance of standard procedures for families of probability distributions not satisfying all the assumptions under which the standard procedures were derived. Workers in this area have primarily concentrated on the effect of changes in the probability model on a single solution rather than on a class of solutions, for example, the class of admissible solutions, as we do here.

We recall some basic ideas. Consider the probability structure $(\mathfrak{X}, \mathfrak{A}, P, \Omega)$ where $\mathfrak{X}$ and Ω are sets, $\mathfrak{A}$ is a σ-field of subsets of $\mathfrak{X}$, and for each θ in Ω, P_θ is a probability measure on $\mathfrak{A}$. Relative to the above structure, a *testing problem* is an ordered pair (ω_0, ω_1) of disjoint subsets of Ω. A *test* φ is a function from $\mathfrak{X}$ into $[0, 1]$ measurable with respect to $\mathfrak{A}$. The test φ is used in the following way: A random element X with values in $\mathfrak{X}$ having P_θ as its probability distribution is observed. If x is the outcome then the hypothesis H: $\theta \varepsilon \omega_0$ is rejected with probability $\varphi(x)$ in favor of the alternative A: $\theta \varepsilon \omega_1$. If the test φ is used and θ is the parameter, then the probability that H is rejected is $E_\theta\varphi = \int_{\mathfrak{X}} \varphi(x)\, dP_\theta(x)$. If φ and φ^* are tests, then φ is *at least as good* as φ^* if

$$E_\theta\varphi - E_\theta\varphi^* \leqq 0$$

Received September 17, 1959.

[1] This work was supported in part by the National Science Foundation.

for θ in ω_0, $\geqq 0$ for θ in ω_1. The test φ is *better* than φ^* if φ is at least as good as φ^* but φ^* is not at least as good as φ. A test φ is *admissible* if there is no test better than φ. A class of tests is *complete* if to each test not in the class there corresponds a test in the class which is better. A class is *minimal complete* if it is complete and no proper subclass is complete. The notions, "essentially complete" and "minimal essentially complete," are defined similarly with "at least as good" substituted for "better."

2. The discrete exponential case. Let $\mathfrak{X}$ be a countable set, h a function from $\mathfrak{X}$ into the positive numbers, and s and t functions from $\mathfrak{X}$ into the real numbers. Let Ω be the interior of the set of points $\theta = (\theta_1, \theta_2)$ where θ_1 and θ_2 are real and satisfy

$$\sum_{x \varepsilon \mathfrak{X}} h(x) e^{\theta_1 s(x) + \theta_2 t(x)} < \infty. \tag{1}$$

For $\theta \,\varepsilon\, \Omega$, let $k(\theta)$ be the reciprocal of the left hand side of (1),

$$p_\theta(x) = k(\theta) h(x) e^{\theta_1 s(x) + \theta_2 t(x)},$$

$x \,\varepsilon\, \mathfrak{X}$, and $P_\theta(A) = \sum_{x \varepsilon A} p_\theta(x)$ where $A \,\varepsilon\, \mathfrak{A}$, the collection of all subsets of $\mathfrak{X}$. For the testing problems considered here, the parameter θ_1 is the more important one, the parameter θ_2 usually being a nuisance parameter. To keep the discussion more compact we limit our attention here to the case of one nuisance parameter although results similar to those that follow are obtainable in much the same way for the case of more than one nuisance parameter.

Assume throughout this section that any nonempty subset of

$$\mathfrak{Y} = \{t(x) \mid x \in \mathfrak{X}\}$$

contains a least element. For $y \,\varepsilon\, \mathfrak{Y}$ let $A_y = \{x \mid t(x) = y\}$.

Let C be the class of tests such that φ is in C if and only if there is a function c such that $\varphi(x) = 1$ if $s(x) > c(t(x))$, $= 0$ if $s(x) < c(t(x))$, $x \,\varepsilon\, \mathfrak{X}$. Let D be the class of tests such that φ is in D if and only if there is a function c and a function d such that $\varphi(x) = 1$ if $s(x) < c(t(x))$, $= 0$ if $c(t(x)) < s(x) < d(t(x))$, $= 1$ if $s(x) > d(t(x))$, $x \,\varepsilon\, \mathfrak{X}$.

THEOREM 1. *Let* (ω_0, ω_1) *be a testing problem relative to the above described probability structure* $(\mathfrak{X}, \mathfrak{A}, P, \Omega)$ *such that for* $i = 0, 1$, *the set* $\{(e^{\theta_1}, e^{\theta_2}) \mid \theta \,\varepsilon\, \omega_i\}$ *has a limit point* $(M_i, 0)$ *where*

$$\inf\{\theta_1 \mid \theta \,\varepsilon\, \Omega\} < \log M_0 < \log M_1 < \sup\{\theta_1 \mid \theta \,\varepsilon\, \Omega\}.$$

Then no test in C is better than some other test in C. If, in addition,

$$\sup\{\theta_1 \mid \theta \,\varepsilon\, \omega_0\} \leqq \inf\{\theta_1 \mid \theta \,\varepsilon\, \omega_1\},$$

then C is minimal complete.

PROOF. Let C_0 be the class of all tests in C that are functions of x only through s and t. Thus, if φ is in C_0 then φ has the form $\varphi(x) = 1$ if $s(x) > c(t(x))$, $= a(t(x))$ if $s(x) = c(t(x))$, $= 0$ if $s(x) < c(t(x))$. If φ is in C then by the theory of sufficient statistics or by an easy calculation there is a φ_0 in C_0 such that $E_\theta\varphi = E_\theta\varphi_0$, $\theta \varepsilon \Omega$.

Suppose that φ and φ^* are tests in C_0 and that φ is at least as good as φ^*. If we can show that $\varphi = \varphi^*$, implying that φ is not better than φ^*, then the first assertion of the theorem will follow.

Let $\psi(x) = [\varphi(x) - \varphi^*(x)]h(x)$. Suppose that it is not true that $\psi(x) = 0$, $x \varepsilon \mathfrak{X}$. Then let y be the least element z of $\mathfrak{Y}$ not satisfying $\psi(x) = 0$, $x \varepsilon A_z$. Then since φ is at least as good as φ^*,

$$\sum_{x\varepsilon\mathfrak{X}} \psi(x)e^{\theta_1 s(x)+\theta_2(t(x)-y)} \qquad \begin{array}{l} \leqq 0, \ \theta \varepsilon \omega_0, \\ \geqq 0, \ \theta \varepsilon \omega_1, \end{array}$$

so that taking limits of both sides as $(e^{\theta_1}, e^{\theta_2}) \to (M_i, 0)$ with $\theta \varepsilon \omega_i$ gives

$$\sum_{x\varepsilon A_y} \psi(x)M_0^{s(x)} \leqq 0 \leqq \sum_{x\varepsilon A_y} \psi(x)M_1^{s(x)}, \tag{2}$$

provided it is permissible to interchange this limiting operation with the summation here. That this is permissible follows from the dominated convergence theorem, since there are points $\theta^i = (\theta_{i1}, \theta_{i2})$, $i = 0, 1$, in Ω satisfying $\theta_{01} < \log M_i < \theta_{11}$, $i = 0, 1$, and $G = \sum_i he^{\theta_{i1}s+\theta_{i2}(t-y)}$ satisfies $\sum_{x\varepsilon\mathfrak{X}} G(x) < \infty$ and $| \psi e^{\theta_1 s+\theta_2(t-y)} | \leqq G$ for θ satisfying

$$\theta_{01} < \theta_1 < \theta_{11}, \qquad \theta_2 < \min \theta_{i2}.$$

Since φ and φ^* are in C_0 either $\psi(x) \geqq 0$, $x \varepsilon A_y$, or $\psi(x) \leqq 0$, $x \varepsilon A_y$. Thus, by (2) we have that $\psi(x) = 0$ for $x \varepsilon A_y$. But this contradicts the definition of y. Hence $\psi(x) = 0$, $x \varepsilon \mathfrak{X}$, implying $\varphi = \varphi^*$, and the first assertion is proved.

Now suppose, in addition, that $\sup \{\theta_1 \mid \theta \varepsilon \omega_0\} \leqq \inf \{\theta_1 \mid \theta \varepsilon \omega_1\}$. Let $r = \sup \{\theta_1 \mid \theta \varepsilon \omega_0\}$. A general result obtained by Truax [7] implies here that C is essentially complete for the testing problem $(\omega_0(r), \omega_1(r))$ where

$$\omega_0(r) = \{\theta \mid \theta \varepsilon \Omega, \theta_1 \leqq r\}$$

and

$$\omega_1(r) = \{\theta \mid \theta \varepsilon \Omega, \theta_1 > r\}.$$

This could also be shown by slightly modifying arguments contained in a paper by Lehmann and Scheffé [4]. This readily implies that C is essentially complete for (ω_0, ω_1) as follows. For any test φ, $E_\theta\varphi$ is continuous in θ for θ in Ω. Suppose φ^* is not in C. Then there is a test φ in C such that $E_\theta\varphi - E_\theta\varphi^* \leqq 0$ if $\theta \varepsilon \omega_0(r)$, $\geqq 0$ if $\theta \varepsilon \omega_1(r)$ and hence by continuity if $\theta \varepsilon \Omega$, $\theta_1 \geqq r$. Thus, for the problem (ω_0, ω_1), φ is at least as good as φ^* and it follows that C is essentially complete.

We now show that C is complete. Since C is essentially complete, this will follow from showing that, if φ is in C and φ^* is any test satisfying

$$E_\theta\varphi = E_\theta\varphi^*, \qquad \theta \varepsilon \omega_0 \cup \omega_1,$$

then φ^* is in C. Suppose φ and φ^* are such tests where $\varphi(x) = 1$ if $s(x) > c(t(x))$, $= 0$ if $s(x) < c(t(x))$, $x \varepsilon \mathfrak{X}$. Then $\varphi^*(x) = 1$ if $s(x) > c(t(x))$, $= 0$ if

$$s(x) < c(t(x)), \qquad x \varepsilon \mathfrak{X}.$$

For suppose this is not true. Let y be the least element z of $\mathcal{Y}$ not satisfying

$$\begin{aligned} \varphi^*(x) &= 1 \text{ if } s(x) > c(z), \\ &= 0 \text{ if } s(x) < c(z), \end{aligned} \qquad x \varepsilon A_z, \tag{3}$$
$$\sum [\varphi(x) - \varphi^*(x)]h(x) = 0. \qquad s(x) = c(z)$$

Then, by a limiting process similar to one used above,

$$\sum_{x \varepsilon A_y} [\varphi(x) - \varphi^*(x)]h(x)M_i^{s(x)} = 0, \qquad i = 0, 1.$$

But this, with the assumption that $M_0 < M_1$, implies by a standard Neyman-Pearson type calculation that (3) must be true for $z = y$, a contradiction. Thus φ^* has the desired form and is therefore in C.

The minimal completeness of C now follows from the completeness of C and the first assertion of the theorem.

REMARKS.

(i) Theorem 1 indicates that C, minimal complete for the standard problem of testing $\theta_1 \leqq r$ against $\theta_1 > r$, remains minimal complete if the testing problem is changed provided that certain conditions are satisfied. A simple kind of permissible change is the introduction of an indifference zone. For example, C is minimal complete for testing $\theta_1 \leqq r_1$ against $\theta_1 \geqq r_2$ where $r_1 < r_2$.

(ii) It is clear from the theorem and proof that C_0 is minimal essentially complete.

(iii) The inequality $M_0 < M_1$ was not needed in the proof of the first assertion.

(iv) A result obtained by Lehmann [3] is related to this theorem. Lehmann considers only the testing problem $\theta_1 \leqq r$ against $\theta_1 > r$, but in the setting of the general exponential family. He shows that C_0 is minimal essentially complete for testing $\theta_1 \leqq r$ against $\theta_1 > r$.

THEOREM 2. *Let* (ω_0, ω_1) *be a testing problem relative to the above described probability structure* $(\mathfrak{X}, \mathfrak{A}, P, \Omega)$ *such that* $\omega_1 = \omega_2 \cup \omega_3$ *and for* $i = 0, 2, 3$, *the set* $\{(e^{\theta_1}, e^{\theta_2}) \mid \theta \varepsilon \omega_i\}$ *has a limit point* $(M_i, 0)$, *where*

$$\inf \{\theta_1 \mid \theta \varepsilon \Omega\} < \log M_2 < \log M_0 < \log M_3 < \sup \{\theta_1 \mid \theta \varepsilon \Omega\}.$$

Then no test in D is better than some other test in D. If, in addition,

$$\sup \{\theta_1 \mid \theta \varepsilon \omega_2\} \leqq \inf \{\theta_1 \mid \theta \varepsilon \omega_0\}$$

and $\sup \{\theta_1 \mid \theta \varepsilon \omega_0\} \leqq \inf \{\theta_1 \mid \theta \varepsilon \omega_3\}$ *then D is minimal complete.*

The proof exactly parallels the proof of Theorem 1 and is therefore omitted. Theorem 2 implies that D is minimal complete for the standard problem of testing $\theta_1 = r$ against $\theta_1 \neq r$ and is also minimal complete for such problems as testing $r_2 \leqq \theta_1 \leqq r_3$ against $\theta_1 \leqq r_1$ or $\geqq r_4$ where

$$r_1 < r_2 \leqq r_3 < r_4 .$$

EXAMPLES. Theorems 1 and 2 have straightforward applications to testing problems with respect to certain multinomial distributions, distributions arising in contingency table analysis, and also to two-sample binomial, Poisson, negative binomial distributions, and so forth. We examine a little more closely a typical example, the two-sample binomial case. Let $X = (X_1, X_2)$ where X_1 and X_2 are independent random variables and X_i is binomial (n_i, p_i), $i = 1, 2$. Here we may let

$$s(x) = x_1, \qquad t(x) = x_1 + x_2, \qquad x = (x_1, x_2),$$

and

$$\theta_1 = \log\left(\frac{p_1}{1 - p_1}\Big/\frac{p_2}{1 - p_2}\right).$$

Theorem 1 implies that C is minimal complete for any of the following testing problems: $p_1/(1 - p_1) \leqq r_1p_2/(1 - p_2)$ against $p_1/(1 - p_1) \geqq r_2p_2/(1 - p_2)$; $p_1 \leqq r_3p_2$ against $p_1 \geqq r_4p_2$; $p_1 = r_3p_2$ against $p_1 = r_4p_2$; and so forth. Here $r_1 < r_2$, $r_3 < r_4$, $r_3 \leqq 1 \leqq r_4$. It is not hard to see that C is not necessarily minimal complete for such problems as testing $p_1 - p_2 \leqq r_1$ against $p_1 - p_2 \geqq r_2$ where $r_1 < 0 < r_2$. Here the indifference zone is too large in the sense that the first assumption of Theorem 1 is not satisfied. For this problem C contains too many tests, some tests in C being inadmissible. Theorem 2 implies that D is minimal complete for a variety of two-sided testing problems involving the ratio $(p_1/(1 - p_1))/(p_2/(1 - p_2))$ or the ratio p_1/p_2 .

Thus, Theorems 1 and 2 imply that the tests described and investigated by Fisher [1], Tocher [6], Sverdrup [5], and Lehmann [3], among others, in connection with the two sample binomial and related problems are admissible (since they are in C or in D) not only for standard testing problems but also for useful modifications of these problems.

3. Counterexample for the continuous exponential case. In the previous section it was seen that, with respect to the discrete exponential family of distributions, the minimal complete class of tests for a standard problem often proved to be minimal complete for a wide variety of related problems. We now show by an example that the situation is much different for the continuous exponential family of distributions. This example, though dealing only with a special subfamily of the continuous exponential family, does not seem to be atypical and does seem to reveal the essential features of the general case.

Let $X = (X_1, X_2)$ where X_1 and X_2 are independent random variables and X_i has the waiting time density with parameter λ_i , $i = 1, 2$. That is, for $x = (x_1, x_2)$ in the first quadrant, the value of the joint density is

$$\lambda_1\lambda_2 e^{x_1(\lambda_2-\lambda_1)+(x_1+x_2)(-\lambda_2)}.$$

Thus, for the standard problem of testing $\lambda_2 - \lambda_1 \leqq r$ against $\lambda_2 - \lambda_1 > r$, the minimal complete class of tests is C where φ^* is in C if and only if there is a test φ and a function c such that $\varphi(x) = 1$ if $x_1 > c(x_1 + x_2)$, $= 0$ if $x_1 < c(x_1 + x_2)$, and $\varphi(x) = \varphi^*(x)$ a.e. (Lebesgue) for x in the first quadrant. (See Lehmann [3].)

The class C is too large to be minimal complete for the related problem of testing $\lambda_2 - \lambda_1 = -1$ against $\lambda_2 - \lambda_1 = 1$ as we now show. Let φ^* be any test such that $\varphi^*(x) = 1$ if $x_1 > c^*(x_1 + x_2)$, $= 0$ if $x_1 < c^*(x_1 + x_2)$, where for each nonnegative integer n,

$$\begin{aligned} c^*(y) &= 0 && \text{if } 1/2^{2n+1} < y - 2 < 1/2^{2n}, \\ &= \log 4 && \text{if } 1/2^{2n+2} < y - 2 < 1/2^{2n+1}. \end{aligned}$$

Then φ^* is in C and φ^* is inadmissible. For let φ be the test satisfying $\varphi(x) = 1$ if $x_1 > c(x_1 + x_2)$, $= 0$ if $x_1 < c(x_1 + x_2)$, where $c(y) = \log 2$ if $2 < y < 3$, $= c^*(y)$ otherwise. Then φ is better than φ^* as can be seen as follows. Straightforward calculation gives

$$E_\lambda \varphi - E_\lambda \varphi^* = (\lambda_1 \lambda_2)/(\lambda_2 - \lambda_1) \int_2^3 e^{-\lambda_2 y}[e^{(\lambda_2 - \lambda_1)c^*(y)} - e^{(\lambda_2 - \lambda_1)c(y)}]\, dy,$$

where $\lambda = (\lambda_1, \lambda_2)$. Thus, if for $\eta > 0$ and $r = -1, 1$,

$$\int_2^3 e^{-\eta y}[e^{rc^*(y)} - e^{rc(y)}]\, dy > 0, \tag{4}$$

then $E_\lambda\varphi - E_\lambda\varphi^* < 0$ if $\lambda_2 - \lambda_1 = -1$, >0 if $\lambda_2 - \lambda_1 = 1$, implying that φ is better than φ^*. Consider the inequality (4) for $r = 1$, the other case being proved similarly. For $\eta > 0$,

$$\begin{aligned} &\int_2^3 e^{-\eta y}[e^{c^*(y)} - e^{c(y)}]\, dy \\ &\qquad = e^{-2\eta} \sum_{n=0}^{\infty} \left[\int_{2^{-(2n+1)}}^{2^{-2n}} e^{-\eta y}(1 - 2)\, dy + \int_{2^{-(2n+2)}}^{2^{-(2n+1)}} e^{-\eta y}(4 - 2)\, dy \right] > 0, \end{aligned}$$

since the nth term in the series is greater than

$$\begin{aligned} &-2^{-(2n+1)} \exp\ (-\eta/2^{2n+1}) \\ &+2^{-(2n+1)} \exp\ (-\eta/2^{2n+1}) = 0. \end{aligned}$$

REFERENCES

[1] R. A. FISHER, "The logic of inductive inference," *J. Roy. Stat. Soc.*, Vol. 98 (1935), pp. 39–54.

[2] WASSILY HOEFFDING, "The role of assumptions in statistical decisions," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I, University of California Press, 1956, pp. 105–114.

[3] E. L. LEHMANN, "Significance level and power," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 1167–1176.

[4] E. L. LEHMANN AND HENRY SCHEFFÉ, "Completeness, similar regions, and unbiased estimation–Part II," *Sankhyā*, Vol. 15 (1955), pp. 219–236.
[5] ERLING SVERDRUP, "Similarity, unbiasedness, minimaxibility and admissibility of statistical test procedures," *Skand. Aktuarietids.*, Vol. 36 (1953), pp. 64–86.
[6] K. D. TOCHER, "Extension of the Neyman-Pearson theory of tests to discontinuous variates," *Biometrika*, Vol. 37 (1950), pp. 130–144.
[7] DONALD R. TRUAX, "Multi-decision problems for the multivariate exponential family," Stanford Technical Report No. 32 (1955).