TESTING MULTIPARAMETER HYPOTHESES[1]

By E. L. Lehmann

Stanford University and University of California, Berkeley

1. Summary. Let the distribution of some random variables depend on real parameters $\theta_1, \cdots, \theta_s$ and consider the hypothesis $H: \theta_i \leqq \theta_i^*, i = 1, \cdots, s$. It is shown under certain regularity assumptions that unbiased tests of H do not exist. Tests of minimum bias and other types of minimax tests are derived under suitable monotonicity conditions. Certain related multidecision problems are discussed and two-sided hypotheses are considered very briefly.

2. Introduction. The extensive literature on optimum tests has been concerned mainly with hypotheses specifying a set of values for a single real valued parameter. Important exceptions are some cases that can be reduced to the one-parameter situation by the principle of invariance, such as the linear (univariate) hypothesis and Hotelling's T^2-problem. These have been used to illustrate a number of different principles, the successful application of which however seems to rest on the symmetry whose full exploitation makes the problems uniparametric. Another exception is the theory of tests with local optimum properties, initiated by Neyman and Pearson [1] and recently developed further by Isaacson [2].

We shall here concern ourselves mainly with hypotheses which, rather than specifying the values of the parameter in question, state that these parameters do not exceed certain bounds. The following examples illustrate the way in which such problems arise.

Example 2.1. Let p and p' denote the number of major and minor defects in a lot. Then the lot will be considered acceptable provided $p \leqq p_0$ and $p' \leqq p_0'$, where $p_0 < p_0'$.

Example 2.2. It may be desired to compare some new treatments with a standard one. Here the hypothesis would specify that none of the new treatments is better by more than a given amount than the standard.

Example 2.3. Let $x_1, \cdots, x_n$ be a sample from a normal distribution with mean ξ and variance σ^2. The population in question may be considered adequate if $\xi \leqq \xi_0$ and $\sigma \leqq \sigma_0$.

In some of the above examples we are dealing with *bona fide* testing problems while in others we are faced with a choice among more than two decisions. Which of these is the case cannot always be seen from the mathematical formulation alone. Thus in Example 2.1 it clearly depends on the disposition that is made of a rejected lot. If there is complete screening, the reason for rejection is immaterial. If on the other hand a lot rejected for major defects is treated differently from one rejected only for minor defects the decision problem becomes more complicated.

[1] Work done under the sponsorship of the Office of Naval Research.

We shall in the following concern ourselves mainly with the one-sided case of the straightforward testing problem. The two-sided situation and the multi-decision problem will be discussed only rather briefly. For simplicity we shall take the number of parameters to be two. The extension to the higher dimensional cases is immediate.

3. Unbiasedness and the minimax principle. The success of the concept of unbiasedness in the one-parameter case suggests the use of this approach also for the present problems. Unfortunately it turns out that in general unbiased tests of the hypotheses in question do not exist. Let us consider the case of two parameters θ_1, θ_2 and the hypothesis $H:\theta_1 \leqq \theta_1^*, \theta_2 \leqq \theta_2^*$. We shall assume that the power function $\beta(\theta_1, \theta_2)$ of any test is analytic in θ_1 and θ_2 in the sense that it can be expanded in an absolutely convergent double power series. Then we shall show that for any unbiased test we have $\beta(\theta_1, \theta_2) \equiv \alpha$, so that any unbiased test is equivalent to the trivial one that rejects with probability α regardless of the observations. This incidentally proves this trivial and most unsatisfactory test to be admissible for the problem under consideration.

Without loss of generality assume that $\theta_1^* = \theta_2^* = 0$. Then unbiasedness states that $\beta(\theta_1, \theta_2) \leqq \alpha$ in the third quadrant of the θ_1, θ_2-plane, and $\geqq \alpha$ in the other three quadrants. By continuity we have $\beta(\theta_1, 0) = \alpha$ for $\theta_1 \leqq 0$ and hence by analyticity $\beta(\theta_1, 0) = \alpha$ for all θ_1. Analogously $\beta(0, \theta_2) \equiv \alpha$. Consider now $\beta(\theta_1, \theta_2)$ for any fixed $\theta_2 > 0$ as a function of θ_1. It has a minimum at $\theta_1 = 0$ so that $\partial\beta(\theta_1, \theta_2)/\partial\theta_1 \mid_{\theta_1=0} = 0$ for all $\theta_2 \geqq 0$. Since $\partial\beta(\theta_1, \theta_2)/\partial\theta_1 \mid_{\theta_1=0}$ is again analytic in θ_2, it follows that $(\partial\beta(\theta_1, \theta_2/\partial\theta_1) \mid_{\theta_1=0}$ is identically zero. Consider now $\beta(\theta_1, \theta_2)$ for some fixed value $\theta_2 \leqq 0$. Since $\beta(\theta_1, \theta_2) \gtreqless \alpha$ as $\theta_1 \gtreqless 0$ and since at $\theta_1 = 0$ the derivative is zero, $\beta(\theta_1, \theta_2)$ must have a point of inflection at 0 and consequently the second derivative $\partial^2\beta(\theta_1, \theta_2)/\partial\theta_1^2 \mid_{\theta_1=0} = 0$ for all $\theta_2 < 0$ and hence for all θ_2. Since the, order of the first non-vanishing derivative $\partial^k\beta(\theta_1, \theta_2)/\partial\theta_1^k \mid_{\theta_1=0}$ must be even for $\theta_2 > 0$ and odd for $\theta_2 < 0$ we see in this manner that for any fixed θ_2 $\partial\ \beta(\theta_1, \theta_2)^k/\partial\theta_1^k \mid_{\theta_1=0} = 0$ for all $k = 1, 2, \cdots$. By analyticity it follows for each fixed θ_2 that $\beta(\theta_1, \theta_2)$ must be a constant, that is, be independent of θ_1. By symmetry it now follows that $\beta(\theta_1, \theta_2)$ must be identically constant, as was to be proved.

We digress for a moment from our search for a reasonable test of the hypothesis $\theta_1, \theta_2 \leqq 0$ to point out that there do exist non-trivial tests of H satisfying the condition of similarity

$$\beta(\theta_1, 0) = \beta(0, \theta_2) = \alpha$$

for all θ_1, θ_2. Suppose for example that X and Y are independently distributed with joint density $f_{\theta_1}(x)f_{\theta_2}(y)$ and that $\alpha = 1/m$ where m is an integer. Then we can obtain a particularly simple class of similar regions as follows. Let $S_1, \cdots, S_m$ be mutually exclusive and exhaustive sets on the real line such that

$$\int_{S_i} f_0(x)\, dx = \alpha, \qquad i = 1, \cdots, m.$$

In the x, y-plane define the set w_i to be the Cartesian product of S_i with itself, and let $w = w_1 + \cdots + w_m$. Now when $\theta_2 = 0$, X is a sufficient statistic for θ_1 and for every x we have

$$P_{\theta_1,0}((X, Y) \varepsilon W \mid x) = \alpha,$$

and hence $P_{\theta_1,0}((X, Y) \varepsilon W) \equiv \alpha$.

We now return to the original problem, and investigate the test that maximizes the minimum power over a certain class ω' of alternatives. For ω' we take the set of points (θ_1, θ_2) for which either $\theta_1 \geqq \theta_1^{**}$ or $\theta_2 \geqq \theta_2^{**}$. Let us consider first as an example the case that X and Y are independently, normally distributed, with known variances and with means θ_1 and θ_2, respectively. Then it would seem as if any reasonable test should satisfy the following two conditions:

(i) $\beta(\theta_1, \theta_2) \leqq \beta(\theta_1', \theta_2')$ whenever $\theta_1 \leqq \theta_1'$, $\theta_2 \leqq \theta_2'$, if ϕ denotes the critical function,

(ii) $\phi(x, y) \leqq \phi(x', y')$ whenever $x \leqq x'$, $y \leqq y'$. It is easily seen that (ii) implies (i); we shall now show that the test ϕ that maximizes $\inf \omega' \beta(\theta_1, \theta_2)$ does not possess property (i) and hence also not (ii), provided $\theta_1^{**} - \theta_1^*$ and $\theta_2^{**} - \theta_2^*$ are sufficiently large so that ϕ is not identically equal to α. Let β denote the power function of ϕ and suppose that $\inf \omega' \beta(\theta_1, \theta_2) = \gamma > \alpha$. Then condition (i) implies that under the hypothesis $\beta(\theta_1, \theta_2) = \alpha$ only when $\theta_1 = \theta_1^*$ and $\theta_2 = \theta_2^*$. For if $\beta(\theta_1, \theta_2) = \alpha$ also for some other point in H, it would also equal α on the line segment connecting these two points and hence by analyticity on the whole line containing this segment. But this would imply that $\beta(\theta_1, \theta_2) = \alpha$ also for points in ω' where by assumption $\beta(\theta_1, \theta_2) \geqq \gamma > \alpha$. Another consequence of condition (i) is that $\beta(\theta_1, \theta_2) > \gamma$ for all points in ω' so that the minimum point γ is never attained in ω' and is approached only as either θ_1 or θ_2 tend to $-\infty$. For if, for example, $\beta(\theta_1', \theta_2') = \gamma$ for some point with $\theta_1' \geqq \theta_2^{**}$ and finite θ_1' we would have $\beta(\theta_1, \theta_2') = \gamma$ for all $\theta_1 \leqq \theta_1'$ and hence for all θ_1. This would imply $\beta(\theta_1, \theta_2) = \gamma$ for all (θ_1, θ_2) with $\theta_1 \geqq \theta_1^{**}$, $\theta_2 \leqq \theta_2'$ and therefore $\beta(\theta_1, \theta_2) \equiv \gamma$.

From these two remarks and Theorem 3.10 of Wald's book *Statistical Decision Functions* [3], it can be shown that there exists a sequence λ_i of probability distributions over ω' with the following properties: (a) For any real number A the probability under λ_i of the intersection of ω' with the quadrant $\{\theta_1, \theta_2 \mid \theta_1, \theta_2 \geqq A\}$ tends to zero as $i \to \infty$. (b) The power of the most powerful level α test for testing $H': \theta_1 = \theta_1^*$, $\theta_2 = \theta_2^*$ against the simple alternative $\int_{\omega'} p_{\theta_1,\theta_2}(x, y) d\lambda_i(\theta_1, \theta_2)$ tends to γ as $i \to \infty$. But from (a) it follows easily that as $i \to \infty$, $\int p_{\theta_1,\theta_2}(x, y) d\lambda_i(\theta_1, \theta_2)$ can be distinguished arbitrarily well from $p_{\theta_1^*,\theta_2^*}(x, y)$. This leads to the contradiction $\gamma = 1$.

We have given the proof explicitly only for the case of independent normal variables. However it applies equally well to any problem in which, in addition to the analyticity assumption of the present section, also the assumptions of Theorem 4.1 are satisfied.

4. Monotone regions. In the previous section we showed that for the hypotheses under consideration neither unbiasedness nor the minimax principle lead to desirable results. In order to arrive at a reasonable test we now impose the following preliminary conditions suggested by the negative results of the last section. We ask first that the test be nonrandomized, so that we can speak of a region w of rejection. The second restriction is one of monotonicity. Let us assume that we are concerned with two random variables X and Y whose joint distribution is given by $p_{\theta_1,\theta_2}(x, y)$. Then we shall say that the region w of rejection for the hypothesis $H: \theta_1 \leqq \theta_2^*, \theta_2 \leqq \theta_2^*$ is *monotone* (nondecreasing in x and y) if

$$(x, y) \in w, \; x \leqq x', \; y \leqq y' \quad \text{imply} \quad (x', y') \in w, \tag{4.1}$$

that is, if its critical function is nondecreasing in both variables.

The restriction to monotone regions is of course suitable only in certain problems, namely, roughly speaking, when increased values of the parameters lead to higher values for the observations. To make this precise let $\theta_1 \leqq \theta_1'$, $\theta_2 \leqq \theta_2'$ and let F and G be the cumulative distribution functions of X and Y corresponding to (θ_1, θ_2) and (θ_1', θ_2'), respectively. Then we shall consider the condition of monotonicity appropriate provided for every monotone non-decreasing region w

$$\int_w dF \leqq \int_w dG. \tag{4.2}$$

Frequently the simplest way to prove (4.2) is to establish the existence of functions $x' = f(x, y)$, $y' = g(x, y)$ with $x' \geqq x$, $y' \geqq y$ and such that when F is the cumulative distribution function of (X, Y), that of (X', Y') is G. Sometimes it is more convenient instead to prove the existence of random variables, say $Z_1, \cdots, Z_r$, and functions $X = f(Z_1, \cdots, Z_r)$, $Y = g(Z_1, \cdots, Z_r)$, $X' = f'(Z_1, \cdots, Z_r)$, $Y' = g'(Z_1, \cdots, Z_r)$ such that $X \leqq X'$, $Y \leqq Y'$ and the cdf's of (X, Y) and (X', Y') are F and G respectively. Both of these conditions clearly assure the validity of (4.2) since for any w that is nondecreasing in x and y they imply

$$\int_w dF = P((X, Y) \in w) \leqq P((X', Y') \in w) = \int_w dG. \tag{4.3}$$

A remark is required also in connection with the restriction to nonrandomized tests. When dealing with discrete problems, for example binomial distributions, we must permit a certain rather trivial kind of randomization. A formal way of handling the distinction is provided by a representation of randomized tests due to M. Eudey [4]. Let X denote the number of successes in n binomial trials, and let U be uniformly distributed over $[0, 1]$. Then any randomized test in X is equivalent to a non-randomized test in $X + U$, and we shall consider monotone non-randomized tests in the continualized variables $X + U$, $Y + V$. Here monotonicity insures that no very heavy use is made of randomization. In fact, in the original variables X, Y randomization will occur only on the boundary of the critical region.

We shall now derive that test of the hypothesis $H: \theta_1 \leqq \theta_1^*$, $\theta_2 \leqq \theta_2^*$ that among all monotone regions maximizes $\inf_{\omega'} \beta(\theta_1, \theta_2)$ where ω' is the set of points (θ_1, θ_2) with $\theta_1 \geqq \theta_1^{**}$ or $\theta_2 \geqq \theta_2^{**}$. Here $\theta_1^{**} \geqq \theta_1^*$ and $\theta_2^{**} \geqq \theta_2^*$. It may be of interest to point out that if we let $\theta_1^{**} = \theta_1^*$, $\theta_2^{**} = \theta_2^*$, we get the monotone region with minimum bias.

THEOREM 4.1. *Let the joint density of X and Y be $p_{\theta_1,\theta_2}(x, y)$ where the parameter-space is a finite or infinite open rectangle $\underline{\theta}_1 < \theta_1 < \bar{\theta}_1$, $\underline{\theta}_2 < \theta_2 < \bar{\theta}_2$, and the positive sample space also is an open rectangle $\underline{x} < x < \bar{x}$, $\underline{y} < y < \bar{y}$ independent of the θ's. Suppose that (4.2) holds, that the marginal distribution of X depends only on θ_1, and that of Y only on θ_2, and that X tends in probability to $\underline{x}$ as $\theta_1 \to \underline{\theta}_1$, while Y tends to $\underline{y}$ as $\theta_2 \to \underline{\theta}_2$. Then the test that among all monotone nonrandomized tests of H maximizes the minimum power against ω' is given by the region of acceptance S:*

$$x \leqq a, \qquad y \leqq b, \tag{4.4}$$

where a and b are determined by the conditions

$$P(X \leqq a, Y \leqq b \mid \theta_1^*, \theta_2^*) = 1 - \alpha \tag{4.5}$$

and

$$P(X \leqq a \mid \theta_1^{**}) = P(Y \leqq b \mid \theta_2^{**}). \tag{4.6}$$

PROOF. We point out first that for any $x > \underline{x}$, $y > \underline{y}$

$$\lim_{\theta_1 \to \underline{\theta}_1} P(X \leqq x, Y \leqq y \mid \theta_1, \theta_2) = P(Y \leqq y \mid \theta_2) \tag{4.7}$$

$$\lim_{\theta_2 \to \underline{\theta}_2} P(X \leqq x, Y \leqq y \mid \theta_1, \theta_2) = P(X \leqq x \mid \theta_1).$$

For

$$P(X \leqq x, Y \leqq y) = P(X \leqq x) - P(X \leqq x, Y \geqq y)$$

while

$$0 \leqq P(X \leqq x, Y > y) \leqq P(Y > y)$$

and

$$\lim_{\theta_2 \to \underline{\theta}_2} P(Y > y \mid \theta_2) = 0.$$

For any monotone test the limit of the power $\beta(\theta_1, \theta_2)$ as $\theta_1 \to \underline{\theta}_1$ clearly exists and we shall denote it by $\beta(\underline{\theta}_1, \theta_2)$. The minimum power in ω' is then the smaller of the two quantities $\beta(\underline{\theta}_1, \theta_2^{**})$ and $\beta(\theta_1^{**}, \underline{\theta}_2)$. Since for the test given by (4.5) we have $\beta(\underline{\theta}_1, \theta_2^{**}) = \beta(\theta_1^{**}, \underline{\theta}_2)$ we could, if the theorem were false, increase both $\beta(\underline{\theta}_1, \theta_2^{**})$ and $\beta(\theta_1^{**}, \underline{\theta}_2)$.

Clearly any monotone test has a region of acceptance S' of the form $y \leqq g(x)$

or equivalently $x \leqq h(y)$ where g and h are non-increasing. If $S' \neq S$ is to be of size $\leqq \alpha$ we have $P(S' \mid \theta_1^*, \theta_2^*) \geqq P(S \mid \theta_1^*, \theta_2^*)$ and hence either $\lim_{x\to\underline{x}} g(x) > b$ or $\lim_{y\to\underline{y}} h(y) > a$. Suppose that the first of these conditions is satisfied and let us denote the complement of S and S' by $\bar{S}$ and $\bar{S}'$, respectively. Then there exists a constant k such that $x \geqq k$ for all points in $S \wedge \bar{S}'$, and a subset S'' of $\bar{S} \wedge S'$ such that $P(S'' \mid \theta_1, \theta_2) > 0$ for all θ_1, θ_2 and $x \leqq k$ for all points in S''. Hence by (4.7)

$$P(S \wedge \bar{S}' \mid \underline{\theta}_1, \theta_2^{**}) = 0$$

and

$$P(\bar{S} \wedge S' \mid \underline{\theta}_1, \theta_2^{**}) = P(b \leqq Y \leqq \lim_{x\to\underline{x}} g(x) \mid \theta_2^{**}) > 0,$$

which leads to

$$P(S' \mid \underline{\theta}_1, \theta_2^{**}) > P(S \mid \underline{\theta}_1, \theta_2^{**}),$$

and hence to the desired result.

The theorem becomes particularly simple if the joint density of X and Y is symmetric in its two variables when $\theta_1 = \theta_2$. For then if $\theta_1^* = \theta_2^* = \theta^*$, $\theta_1^{**} = \theta_2^{**} = \theta^{**}$, it is seen from (4.6) that $a = b$. Thus the test accepts if $\max(X, Y) \leqq a$ where a is determined by (4.5), and hence is independent of θ^{**}.

The assumptions made in Theorem 4.1 concerning the shape of the parameter and sample spaces are unnecessarily restrictive. The theorem remains valid if we assume that both the parameter space and the positive sample space are convex open sets. The proof is essentially the same, however the notation becomes considerably more complicated.

If the roles of hypothesis and class of alternatives are interchanged, we obtain

THEOREM 4.2. *For testing the hypothesis $H': \theta_1 \geqq \theta_1^{**}$ or $\theta_2 \geqq \theta_2^{**}$ against the class of alternatives $\omega: \theta_1 \leqq \theta_1^*$, $\theta_2 \leqq \theta_2^*$, let S be the region of rejection $x \leqq c$, $y \leqq d$, where c and d are determined by*

$$P(X \leqq c \mid \theta_1^{**}) = P(Y \leqq d \mid \theta_2^{**}) = \alpha. \tag{4.8}$$

Then under the assumptions of Theorem 4.1 S is uniformly most powerful among all regions of rejection that are monotone non-increasing in both variables.

PROOF. Consider any monotone region given by $x \leqq g(y)$ or $y \leqq h(x)$ with g and h non-increasing. Since the probability of rejection must be not greater than α at $(\theta_1^{**}, \underline{\theta}_2)$ and $(\underline{\theta}_1, \theta_2^{**})$ we must have

$$\lim_{x\to\underline{x}} g(x) \leqq c, \qquad \lim_{y\to\underline{y}} h(y) \leqq d. \tag{4.9}$$

But any monotone region satisfying (4.9) is contained in S and hence is uniformly less powerful than S.

5. Examples. In the present section we shall apply Theorem 4.1 to some specific problems. All other assumptions being trivially satisfied we shall in each case only check condition (4.2).

Example 5.1. Let X, Y have a multinomial distribution

$$(4.10) \quad P(X = x, Y = y) = \frac{n!}{x!y!(n - x - y)!} \theta_1^x \theta_2^y (1 - \theta_1 - \theta_2)^{n-x-y}.$$

To see that (4.2) holds let $Z_1, \cdots, Z_n$ be uniformly distributed on (0, 1) and let U, V denote the number of Z's in the intervals $(0, \theta_1)$ and $(1 - \theta_2, 1)$ respectively. If U', V' are defined analogously with respect to θ_1', θ_2' it is seen that (U, V) has the distribution (4.10) while (U', V') has the corresponding distribution for θ_1', θ_2'. Since $U \leqq U'$, $V \leqq V'$ the validity of (4.2) follows.

The same proof works of course in the case that X and Y are independently distributed, each according to a binomial distribution.

Example 5.2. Let $X_1, \cdots, X_n$ be independently and normally distributed with mean ξ and variance σ^2, and consider the hypothesis $H: \xi \leqq \xi_0$, $\sigma \leqq \sigma_0$. Since $\bar{X}$ and $S^2 = \sum(X_i - \bar{X})^2$ are sufficient for ξ and σ, we may restrict attention to these statistics. However, if we try to set $X = \bar{X}$, $Y = S^2$, $\theta_1 = \xi$, $\theta_2 = \sigma^2$ we encounter certain difficulties. First the distribution of X does not depend only on θ_1 as we require in Theorem 4.1. While this is not a very important condition of the theorem, a second consideration shows that it is impossible to apply the monotonicity restriction at all to the present set-up. For the joint cumulative distribution function of $\bar{X}$ and S does not satisfy condition (4.2).

This exhibits an unpleasant feature of the present approach. In a given problem it is not known a priori whether there will exist variables X, Y and a choice of the parameters θ_1, θ_2 so that (4.2) will be satisfied. On the other hand, when such variables and parameters have been found, it is not clear that these are the only possible choices. While it would of course be interesting to investigate existence and uniqueness questions, the monotonicity condition is an extraneous restriction anyway, whose suitability must be judged for each problem in terms of the choices for X, Y, θ_1, and θ_2.

In the present case we may take $X = (\bar{X} - \xi_0)/S$, $Y = S^2$, $\theta_1 = (\xi - \xi_0)/\sigma$ and $\theta_2 = \sigma^2$. To check condition (4.2) assume without loss of generality that $\xi_0 = 0$ and let $\theta_1 = \xi/\sigma < \xi'/\sigma' = \theta_1'$, $\sigma < \sigma'$. If $\sigma' = k\,\xi' = k\xi + c$, let $X_i' = kX_i + c$, so that $S' = kS$, $\bar{X}' = k\bar{X} + c$. Since $k > 1$ and $\xi < \xi'/k$ we see that $c > 0$ and $\bar{X}' > \bar{X}$, $Y' > Y$, so that (4.2) follows.

As a last problem let us consider one in which nuisance parameters are present.

Example 5.3. Let $X_1, \cdots, X_m$; $Y_1, \cdots, Y_n$ be independently normally distributed with common variance σ^2; let $E(X_i) = \xi$, $E(Y_j) = \eta$, and consider the hypothesis $H: \xi \leqq \xi_0$, $\eta \leqq \eta_0$. This time $\bar{X}$, $\bar{Y}$, and $S^2 = \sum(X_i - \bar{X})^2 + \sum(Y_j - \bar{Y})^2$ from a set of sufficient statistics, and again the question arises how to choose X, Y, θ_1, and θ_2. Here the principle of invariance (see [5]) leads to a solution very simply. If one rewrites the hypothesis in the form: $(\xi - \xi_0)/\sigma \leqq 0$, $(\eta - \eta_0)/\sigma \leqq 0$ it is seen that $X = (\bar{X} - \xi_0)/S$, $Y = (\bar{Y} - \eta_0)/S$ constitute a maximal invariant under a suitable group of transformations. The corresponding parameter invariants are of course $\theta_1 = (\xi - \xi_0)/\sigma$, $\theta_2 = (\eta - \eta_0)/\sigma$.

It remains, once more, to check (4.2). For this purpose let ξ, η, ξ', η', σ be numbers such that

$$\frac{\xi}{\sigma} = \theta_1, \qquad \frac{\eta}{\sigma} = \theta_2, \qquad \frac{\xi'}{\sigma} = \theta_1', \qquad \frac{\eta'}{\sigma} = \theta_2'$$

and let $U_1, \cdots, U_m$; $V_1, \cdots, V_n$ be independently normally distributed with common variance σ^2 and means $E(U_i) = \xi$, $E(V_j) = \eta$. If $T^2 = \sum(U_i - \bar{U})^2 + \sum(V_j - \bar{V})^2$, the joint distribution of $U = \bar{U}/T$, $V = \bar{V}/T$ is that corresponding to (θ_1, θ_2). Let $\xi' - \xi = c$, $\eta' - \eta = d$ and let $U_i' = U_i + c$, $V_j' = V_j + d$. Then $U' = \bar{U}'/T$, $V' = \bar{V}'/T$ have the joint distribution corresponding to (θ_1', θ_2'). Also $U' = \bar{U}+c/T > U$ and $V' = \bar{V}+c/T > V$.

Exactly analogously we can treat the hypothesis $H': \xi/\sigma \leqq \gamma$, $\eta/\sigma \leqq \delta$, and the corresponding problems in which the two variances are not assumed to be equal.

6. A multidecision problem. As was pointed out in the introduction, some of the problems considered here really involve the choice between more than two decisions. We shall now indicate, by discussing an example, one method of treating such multidecision problems through successive reduction to problems of the simpler type.

Let us once more consider the hypothesis H : $\theta_1, \theta_2 \leq 0$ and let us assume that in case of rejection we wish to decide whether $\theta_2 > 0 \geq \theta_1$, $\theta_1 > 0 \geq \theta_2$ or whether θ_1 and θ_2 are both > 0[2]. Let us denote these three regions of the parameter space by ω_1, ω_2, and ω_3 and the associated decisions by d_1, d_2, d_3. The set $\theta_1, \theta_2 \leq 0$ will be denoted by ω_0 and the associated decision of accepting H by d_0.

We shall assume that each of the four pairs of random variables $(\pm X, \pm Y)$ is monotone with respect to the corresponding pair of parameters $(\pm\theta_1, \pm\theta_2)$ and that $p_{\theta,\theta}(x, y)$ is symmetric in x and y so that the region of acceptance is given by

$$\max(x, y) \leqq a. \tag{6.1}$$

We must now consider how to divide up the complementary region between d_1, d_2, and d_3. Here we again impose the natural monotonicity restrictions. We ask that the region for d_1 be monotone non-increasing in x and non-decreasing in y, and that the analogous conditions be satisfied by the regions for d_2 and d_3. Suppose the problem concerns a standard treatment and two new ones, where θ_1 and θ_2 measure in some way the differences between the new treatments and the standard. The circumstances are such that the most serious error consists in incorrectly rejecting the standard treatment in favor of one of the others. By proper choice of a this probability is controlled so that it is not greater than α for all $(\theta_1, \theta_2) \, \varepsilon \, \omega_0$.

Next in importance seems to come the possibility of reaching decision d_2 in

[2] A similar multidecision problem has recently been treated by Paulson [6] from a somewhat different point of view.

ω_1 or d_1 in ω_2, but we shall set these aside for the moment. The next important error presumably consists in deciding on d_3 when the parameter point lies either in ω_1 or ω_2. This error can be controlled in the usual manner by making the d_3-region sufficiently small. Thus we may select a number $0 < \beta < 1$ and impose bounds

$$P(d_3 \mid \omega_1) \leqq \beta, \qquad P(d_3 \mid \omega_2) \leqq \beta. \tag{6.2}$$

Subject to these conditions we wish to maximize $P(d_3)$ in ω_3. Let us now restrict attention to d_3-regions of the type $y \geqq g(x)$ or $x \geqq h(y)$ with g and h non-increasing. Then it is seen by the argument used to prove Theorem 4.2 that among all monotone d_3-regions satisfying (6.2) there exists one that uniformly maximizes $P(d_3)$ over ω_3. If $\mathrm{P}(X > a \mid \theta_1 = 0) \leqq \beta$ it consists of the points satisfying either $x \geqq a, y \geqq b$ or $x \geqq b, y \geqq a$ where a, b are determined by

$$P(X \leqq a, Y \leqq a \mid \theta_1 = \theta_2 = 0) = 1 - \alpha,$$
$$P(X > b \mid \theta_1 = 0) = \beta.$$

If on the other hand $P(X > a \mid \theta_1 = 0) \geqq \beta$ the optimum d_3-region is given by $x \geqq b, y \geqq b$.

Let the remainder of the sample space be divided up symmetrically in the obvious manner between d_1 and d_2. It then follows from monotonicity that $P(d_1 \mid \omega_2)$ and $P(d_2 \mid \omega_1)$ both take on their maximum value at $\theta_1 = \theta_2 = 0$. Hence $P(d_1 \mid \omega_2) \leqq \frac{1}{2}\alpha$, $P(d_2 \mid \omega_1) \leqq \frac{1}{2}\alpha$, so that these errors also are controlled in a satisfactory manner.

7. Convex regions.[3] If we try to apply the results of Section 4 to specific examples, we occasionally find an obstacle in the condition that X and Y should tend in probability to $\underline{x}$ and $\underline{y}$ as $\theta_1 \to \underline{\theta}_1$ and $\theta_2 \to \underline{\theta}_2$, respectively. We shall now show that by restricting the acceptance region to be convex as well as monotone we can prove a result analogous to Theorem 4.1 without assuming degeneracy of the distribution at θ_1 and θ_2.

Let us consider again the joint density $p_{\theta_1,\theta_2}(x, y)$ satisfying (4.2), the hypothesis $H: \theta_1 \leqq \theta_1^*$, $\theta_2 \leqq \theta_2^*$ and the set of alternatives $\theta_1 \geqq \theta_1^{**}$ or $\theta_2 \geqq \theta_2^{**}$. Putting

$$r_\pi(x, y) = \frac{p_{\theta_1^*,\theta_2^*}(x, y)}{\pi p_{\theta_1,\theta_2^{**}}(x, y) + (1 - \pi) p_{\theta_1^{**},\theta_2}(x, y)},$$

we shall assume:

(i) For any $0 < \pi < 1$ and any C the region

$$r_\pi(x, y) \geqq C \tag{7.1}$$

is convex and non-decreasing in x and y.

[3] I am indebted to the referee for several valuable suggestions with regard to this section.

(ii) Given any two points (x', y') and (x'', y'') with $x' < x''$, $y' > y''$ there exists $0 < \pi < 1$ and C such that both points lie on the boundary of the set of points (x, y) satisfying (7.1).

(iii) $p_{\theta_1,\theta_2}(x, y) > 0$ for all θ_1, θ_2, x, y under consideration.

Consider now the set of alternatives $\omega' : \theta_1 \geqq \underline{\theta}_1$, $\theta_2 \geqq \theta_2^{**}$ or $\theta_2 \geqq \underline{\theta}_2$, $\theta_1 \geqq \theta_1^{**}$ where θ_1 and θ_2 may now be any numbers less than θ_1^* and θ_2^*, respectively. Let a and b be determined by

$$P(X \leqq a, Y \leqq b \mid \theta_1^*, \theta_2^*) = 1 - \alpha,$$

$$P(X \leqq a, Y \leqq b \mid \underline{\theta}_1, \theta_2^{**}) = P(X \leqq a, Y \leqq b \mid \theta_1^{**}, \underline{\theta}_2).$$

Then the acceptance region S: $x \leqq a$, $y \leqq b$ maximizes $\inf_{\omega'} \beta(\theta_1, \theta_2)$ among all monotone and convex level α tests.

For let S' be any other acceptance region satisfying the conditions that have been imposed. The boundary curves of S and S' have in common either one or

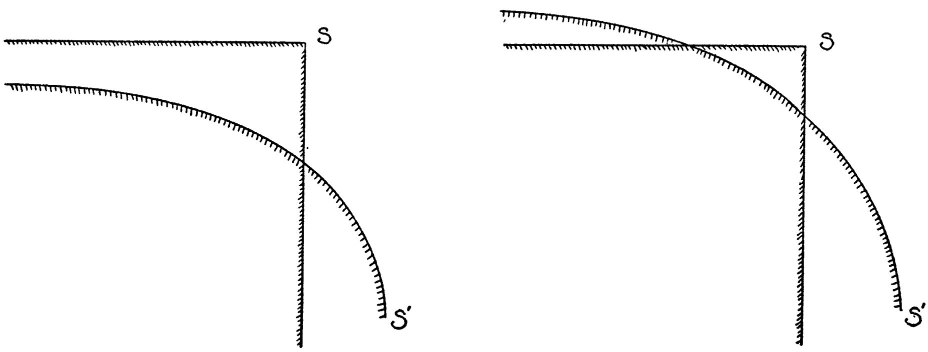

FIG. 1.

two points or an interval. Let us consider the case of two points, say (x', y') and (x'', y''). We may then assume $x' < x''$, $y' > y''$ so that there exist π and C such that the boundary of (7.1) passes through these two points. From (i) it follows that $r(x, y) \geqq C$ for all points in $S \wedge \bar{S}'$ and $\leqq C$ in $\bar{S} \wedge S'$. Since we have $P(S') \leqq P(S) = \alpha$ when the density of X, Y is $p_{\theta_1^*,\theta_2^*}(x, y)$ it follows from the fundamental lemma of Neyman and Pearson that $P(S') > P(S)$ when the density is given by $\pi p_{\theta_1^*,\theta_2^{**}}(x, y) + (1 - \pi)\, p_{\theta_1^{**},\theta_2}(x, y)$. It is therefore impossible that $P(S') \geqq P(S)$ for both $(\underline{\theta}_1, \theta_2^{**})$ and $(\theta_2^{**}, \underline{\theta}_2)$ as was to be proved.

The same argument applies also in the case that the boundaries of S and S' have an interval in common. Consider finally the case of one common point, say (a, y_0). For each n let (π_n, C_n) satisfy (ii) for $(-n, b)$ and (a, y_0). Then

$$0 \leqq C_n \leqq \max \left\{ \frac{p_{\theta_1^*,\theta_2^*}(a, y_0)}{p_{\underline{\theta}_1,\theta_2^{**}}(a, y_0)}, \frac{p_{\theta_1^*,\theta_2^*}(a, y_0)}{p_{\theta_1^{**},\underline{\theta}_2}(a, y_0)} \right\},$$

so that there is a subsequence of $\{(\pi_n, C_n)\}$ which converges, say, $\pi_n \rightarrow \pi^*$, $C_n \rightarrow$

C^*. It is easily seen from the monotonicity and convexity of the regions $r_{\pi_n}(x, y) \geqq C_n$ that the boundary of $r_{\pi^*}(x, y) \geqq C^*$ is the line $y = y_0$. The remainder of the argument is completely analogous to the two point case.

As an example let

$$p_{\theta_1,\theta_2}(x, y) = C(\theta_1, \theta_2)\, e^{\theta_1 x + \theta_2 y}\, h(x, y), \tag{7.2}$$

and assume that (4.2) holds. Suppose that a priori lower bounds $\underline{\theta}_1, \underline{\theta}_2$ are given for θ_1, θ_2 such that $p_{\underline{\theta}_1,\underline{\theta}_2}(x, y)$ is again a density. If we let $\theta_1^{**} = \theta_1^*$, $\theta^{**} = \theta_2^*$, the region (7.1) becomes

$$a\, e^{(\theta_1 - \overset{*}{\theta}_1)x} + b\, e^{(\theta_2 - \overset{*}{\theta}_2)y} \leqq k$$

and conditions (i), (ii), and (iii) are easily checked.

8. Two-sided problems. We shall discuss only one rather trivial two-sided problem, which is enough however to indicate that the type of result one obtains here is entirely different from what we found in the one-sided case.

Let X and Y be independently, normally distributed with unit variance and means ξ and η, respectively, and consider the hypothesis $H: \xi = \eta = 0$. We shall determine the test ϕ that maximizes $\inf_{\omega'} \beta(\xi, \eta)$ where ω' is the set of points for which either $|\xi|$ or $|\eta|$ is $\geqq \gamma$, $(\gamma > 0)$. Any reasonable test for this problem would presumably attain its minimum power over ω' at the four points $(0, \gamma)$, $(0, -\gamma)$, $(\gamma, 0)$, $(-\gamma, 0)$. We therefore expect ϕ to be the most powerful test of H against the simple alternative that assigns probability $\frac{1}{4}$ to each of these 4 points. The region of acceptance for this problem is given by S:

$$e^{\gamma x} + e^{-\gamma x} + e^{\gamma y} + e^{-\gamma y} \leqq k.$$

It is easily checked that this has the following properties:

(i) S is convex.

(ii) If $0 \leqq x' \leqq x$, $0 \leqq y' \leqq y$ and $(x, y) \varepsilon S$, the point (x', y') also lies in S.

(iii) For any fixed η_0 the probability of S decreases with $|\xi|$ and for fixed ξ_0 decreases in $|\eta|$.

From (iii) it follows that ϕ is the test we are looking for, and it seems to be entirely satisfactory. In fact, if we utilize the symmetry of the situation to reduce the variables to $|X|$, $|Y|$ and the parameters to $|\xi|$, $|\eta|$ we are faced essentially with a one-sided situation and it is seen from (i) and (ii) that the acceptance region, when interpreted in this way, is both monotone and convex.

9. A general concept of monotonicity. In Sections 4–6 we made use of the notion of monotonicity, and we shall conclude this paper by indicating how this concept may be extended to the general decision problem.

Suppose that there is defined a partial ordering $\leqq$ in the sample space and a partial ordering $\dot{\leqq}$ in the parameter-space. In analogy to condition (4.2) we shall assume that if $\mathcal{W}$ is any monotone non-decreasing region in the sample space we have for any two parameter-points $\theta \dot{\leqq} \theta' : P_\theta(\mathcal{W}) \leqq P_{\theta'}(\mathcal{W})$.

Suppose that there is also defined a partial ordering in the decision space, to be denoted by $\stackrel{\cdot}{\leqq}$. We shall assume that the loss function W satisfies the conditions:

(i) $d_1 \stackrel{\cdot}{\leqq} d_2 \stackrel{\cdot}{\leqq} d_3$ and $W(\theta, d_1) < W(\theta, d_2)$ implies $W(\theta, d_2) \leqq W(\theta, d_3)$,

(ii) $\theta_1 \stackrel{\cdot}{\leqq} \theta_2 \stackrel{\cdot}{\leqq} \theta_3$ and $W(\theta_1, d) < W(\theta_2, d)$ implies $W(\theta_2, d) \leqq W(\theta_3, d)$. Under these assumptions it seems natural to restrict consideration to monotone decision functions, where we shall call δ monotone if $x \leqq x'$ implies $\delta(x) \stackrel{\cdot}{\leqq} \delta(x')$.

REFERENCES

[1] J. Neyman and E. S. Pearson, "Contributions to the theory of testing statistical hypotheses. Part III." *Statistical Research Memoirs*, Vol. 2 (1938), pp. 36–57.

[2] Stanley L. Isaacson, "On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 217–234.

[3] A. Wald, *Statistical Decision Functions*, John Wiley and Sons, New York, 1950.

[4] Mark W. Eudey, "On the treatment of discontinuous random variables," *Technical Report No. 13, ONR Project NR-042-036*, 1949.

[5] E. L. Lehmann, "Some principles of the theory of testing hypotheses," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 1–26.

[6] Edward Paulson, "On the comparison of several experimental categories with a control," *Annals of Math. Stat.*, Vol. 23 (1952), pp. 239–246.