NOTES

A CLASS OF MINIMAX TESTS FOR ONE-SIDED COMPOSITE HYPOTHESES[1]

By S. G. Allen, Jr.

Stanford University

Summary. Fixed sample-size procedures are considered for testing a one-sided composite hypothesis concerning a real, one-dimensional parameter of an exponential distribution (1.1). In particular, conditions are studied such that the minimax tests have a critical region which is a semi-infinite interval on the real line.

1. Statement of the problem. Let X be a real-valued, one-dimensional random variable with the probability density

$$p(x, \theta) = \omega(\theta)\psi(x)e^{\theta x}, \tag{1.1}$$

where

$$\omega(\theta) = \left[\int_{-\infty}^{\infty} e^{\theta x}\psi(x)\, dx\right]^{-1} \tag{1.2}$$

is a positive, bounded, continuous function of the real variable θ and where $\psi(x)$ is a continuous, nonnegative function of the real variable x. Let $X_1, X_2, \cdots, X_n$ denote n independent observations on X, and let $T(X_1, \cdots, X_n)$ denote a fixed sample-size procedure based on the n observations for testing the composite hypothesis $\theta > \theta_0$ against the alternative $\theta < \theta_0$. The loss functions are defined as follows: if the hypothesis is rejected, the loss is $w_1(\theta) \geqq 0$ for $\theta > \theta_0$ and $w_1(\theta) = 0$ otherwise; if the hypothesis is accepted, the loss is $w_2(\theta) \geqq 0$ for $\theta < \theta_0$ and $w_2(\theta) = 0$ otherwise. Furthermore, it is assumed that the function $w_1(\theta)$ is actually positive for at least one value of $\theta > \theta_0$, and $w_2(\theta)$ is positive for at least one value of $\theta < \theta_0$. The problem to be considered is the selection of a minimax test procedure $T(X_1, \cdots, X_n)$ under these conditions.

2. A class of minimax tests. For testing the simple hypothesis $\theta = \theta_2$ in (1.1)

Received 11/5/51, revised 11/2/52.

[1] This paper, representing work done under the sponsorship of the Office of Naval Research, was presented at the Western meetings of the Institute of Mathematical Statistics, June 15–16, 1951, Santa Monica, California. Discussion with members of the Department of Statistics, Stanford University, in particular with Professor M. A. Girshick, was most beneficial in the formulation of the present draft of the paper. The author understands that results similar to those of the present paper were obtained for the sequential case by Milton Sobel in his doctoral thesis, "An essentially complete class of decision functions for certain standard sequential problems."

against the simple alternative $\theta = \theta_1$ with $\theta_1 < \theta_2$, the minimax procedure based on n independent observations on X is well known [1]. The value of the statistic

$$\lambda = \lambda(\theta_1, \theta_2) = \prod_{i=1}^{n} \frac{p(x_i, \theta_2)}{p(x_i, \theta_1)} \tag{2.1}$$

is computed from the observed values of X in the sample. The hypothesis is then accepted if $\lambda > c$ and rejected if $\lambda \leqq c$, where the criterion c satisfies

$$w_1(\theta_2)Pr(\lambda \leqq c \mid \theta_2) = w_2(\theta_1)Pr(\lambda > c \mid \theta_1). \tag{2.2}$$

This value of c is

$$c = \frac{w_2(\theta_1)g}{w_1(\theta_2)(1 - g)}, \tag{2.3}$$

where g is the least favorable a priori probability that $\theta = \theta_1$.

From the form of the density function (1.1), it is clear that an identical procedure to the preceding ratio test specifies acceptance of the hypothesis if and only if $\sum_{i=1}^{n} x_i > k$, where

$$k = \frac{1}{\theta_2 - \theta_1} \log c \left[\frac{\omega(\theta_1)}{\omega(\theta_2)}\right]^n. \tag{2.4}$$

Since the probability density of the statistic $\sum_{i=1}^{n} x_i$ is again of the form (1.1) (see Section 4 of [2]), the discussion of tests like the above is not restricted by an assumption that the sample consists of a single observation on X. Therefore, the number k defined in (2.4) may be determined by a condition equivalent to (2.2) with $n = 1$, namely,

$$w_1(\theta_2)Pr(X \leqq k \mid \theta_2) = w_2(\theta_1)Pr(X > k \mid \theta_1). \tag{2.5}$$

Let $T_k(X)$ denote a test procedure specifying acceptance of the hypothesis $\theta > \theta_0$ if the observed value of X exceeds k, and specifying rejection otherwise. One might ask if such test procedures, which form a class of minimax procedures in the case of the simple dichotomy, retain this property in the more general problem of Section 1. If so, does a condition similar to (2.5) determine the minimax test?

The following theorem supplies an answer.[2]

THEOREM 1. *Let*

$$R_1(k, \theta) = w_1(\theta) \int_{-\infty}^{k} \omega(\theta)\psi(x)e^{\theta x}\, dx, \qquad \theta \geqq \theta_0, \tag{2.6}$$

$$R_2(k, \theta) = w_2(\theta) \int_{k}^{\infty} \omega(\theta)\psi(x)e^{\theta x}\, dx, \qquad \theta \leqq \theta_0. \tag{2.7}$$

[2] The motivating idea for Theorem 1 was a lot acceptance sampling procedure proposed in an unpublished paper by Mr. Norman Rudy of Sacramento State College.

Then $T_k(X)$ is minimax if

$$\max_{\theta \geqq \theta_0} R_1(k, \theta) = \max_{\theta \leqq \theta_0} R_2(k, \theta). \tag{2.8}$$

PROOF. Let $R(T, G)$ denote the expected loss of a test T with respect to the a priori distribution with cdf $G(\theta)$. In particular, for a k_0 satisfying (2.8),

$$\begin{aligned} R(T_{k_0}, G) &= \int_{\theta_0}^{\infty} R_1(k_0, \theta)\, dG(\theta) + \int_{-\infty}^{\theta_0} R_2(k_0, \theta)\, dG(\theta) \\ &\leqq \int_{\theta_0}^{\infty} \max_{\theta \geqq \theta_0} R_1(k_0, \theta)\, dG(\theta) + \int_{-\infty}^{\theta_0} \max_{\theta \leqq \theta_0} R_2(k_0, \theta)\, dG(\theta) \\ &= \max_{\theta \geqq \theta_0} R_1(k_0, \theta) = \max_{\theta \leqq \theta_0} R_2(k_0, \theta). \end{aligned} \tag{2.9}$$

Let θ_1, θ_2 be values of θ such that $\theta_1 \leqq \theta_0 \leqq \theta_2$ and

$$\max_{\theta \geqq \theta_0} R_1(k_0, \theta) = R_1(k_0, \theta_2), \tag{2.10}$$

$$\max_{\theta \leqq \theta_0} R_2(k_0, \theta) = R_2(k_0, \theta_1). \tag{2.11}$$

If G is a distribution concentrating all probability at θ_1 and θ_2, then the equality sign holds throughout (2.9). Therefore

$$\max_{G} R(T_{k_0}, G) = \max_{\theta \geqq \theta_0} R_1(k_0, \theta) = \max_{\theta \leqq \theta_0} R_2(k_0, \theta). \tag{2.12}$$

In particular let G_0 be the distribution given by $g = Pr(\theta = \theta_1)$, $1 - g = Pr(\theta = \theta_2)$, where g satisfies

$$k_0 = \frac{1}{\theta_2 - \theta_1} \log \frac{w_2(\theta_1) g \omega(\theta_1)}{w_1(\theta_2)(1 - g)\omega(\theta_2)}.$$

Clearly T_{k_0} is the Bayes procedure against G_0. (Compare g in (2.3) and (2.4).) Hence

$$\min_{T} R(T, G_0) = R(T_{k_0}, G_0) = \max_{G} R(T_{k_0}, G).$$

Application of the saddle-point theorem of [3] completes the proof.

3. An example based on the normal distribution. Suppose it is desired to test the hypothesis that θ, the mean of a normal distribution with variance one, is positive against the alternative that it is negative, where $w_1(\theta) = \theta$ for $\theta \geqq 0$ and $w_2(\theta) = -\theta$ for $\theta \leqq 0$. The functions defined in (2.6) and (2.7) are

$$R_1(k, \theta) = \int_{-\infty}^{k-\theta} \frac{\theta}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy, \qquad \theta \geqq 0$$

$$R_2(k, \theta) = \int_{k-\theta}^{\infty} -\frac{\theta}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy = \int_{-\infty}^{-k+\theta} -\frac{\theta}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy, \qquad \theta \leqq 0.$$

Since $R_2(k, -|\theta|) \equiv R_1(-k, |\theta|)$, it follows that $\max_{\theta \leqq 0} R_2(0, \theta) = \max_{\theta \geqq 0} R_1(0, \theta)$, provided the latter exist. This is certainly the case, since, by L'Hospital's rule,

$$\lim_{\theta \to \infty} R_1(0, \theta) = \lim_{\theta \to \infty} \frac{\theta^2}{\sqrt{2\pi}} e^{-\frac{1}{2}\theta^2} = 0.$$

4. Remarks on the discrete case. The continuous distributions studied in the preceding sections represent a sub-family of a more general family of distributions of the form $\omega(\theta)e^{\theta x}d\Psi(x)$, where Ψ is a measure on the real numbers and where

$$\omega(\theta) = \left[\int_{-\infty}^{\infty} e^{\theta x}\, d\Psi(x)\right]^{-1}$$

is a positive bounded function of the real variable θ. This family includes many of the most important distributions encountered in statistics, such as the normal, χ^2, binomial, negative binomial, and Poisson distributions.

Suppose the distribution under consideration in this family is a discrete one, and suppose that $\Psi(x)$ assumes jumps at each value of a denumerable, ordered sequence $(x_1, x_2, \cdots)$. For example, if X is the number of successes in n Bernoulli trials, the function $\Psi(x)$ assumes jumps at $x = 0, 1, 2, \cdots, n$. In general, it will not be possible to find a value of k in such a sequence so that condition (2.8) is fulfilled. However, a randomized mixture of two procedures T_k and $T_{k'}$ will be a minimax procedure if there exists a pair (k, k') such that

$$\max_{\theta \geqq \theta_0} R_1(k', \theta) < \max_{\theta \leqq \theta_0} R_2(k', \theta),$$

$$\max_{\theta \geqq \theta_0} R_1(k, \theta) > \max_{\theta \leqq \theta_0} R_2(k, \theta),$$

where k' is the next smaller element than k in the sequence $(x_1, x_2, \cdots)$. In this event, the minimax test procedure consists of the following: reject the hypothesis $\theta > \theta_0$ if $x < k$; accept the hypothesis if $x > k$; if $x = k$, accept the hypothesis with probability f and reject with probability $1 - f$, where f satisfies

$$\max_{\theta \geqq \theta_0} [fR_1(k', \theta) + (1 - f)R_1(k, \theta)] = \max_{\theta \leqq \theta_0} [fR_2(k', \theta) + (1 - f)R_2(k, \theta)].$$

REFERENCES

[1] A. WALD, *Statistical Decision Functions*, John Wiley and Sons, 1950, p. 17.
[2] M. A. GIRSHICK AND L. J. SAVAGE, "Bayes and Minimax Estimates for Quadratic Loss Functions," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951, pp. 53–73.
[3] J. VON NEUMANN AND O. MORGENSTERN, *Theory of Games and Economic Behavior*, 2nd ed., Princeton University Press, 1947, p. 95.