PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 7, Fasc. 1 (1986), p. 1-5

BOUNDED STOPPING TIME OF SOME BAYES SEQUENTIAL TESTS FOR THE t-TEST MODEL *

BY

ARTHUR COHEN (NEW BRUNSWICK)

Abstract. For the t-test model the problem is to sequentially test whether the sign of the mean is negative or positive. Consider normal-gamma priors and the following three loss functions:

(i) linear combination of cost and 0-1;

(ii) linear combination of cost and absolute error;

(iii) linear combination of cost and absolute error divided by the standard deviation.

For losses (i) and (iii) the Bayes test is shown to have bounded stopping time and a bound on the maximum sample size is obtainable. For loss (ii) the Bayes test does not have bounded stopping time. Intuitive explanations for these somewhat surprising results are offered.

1. Introduction and summary. Consider the t-test model. That is, X, X_1, $X_2, \ldots$ are independent, identically distributed normal variables with unknown mean μ and unknown variance σ^2. The problem is to sequentially test H: $\mu \leqslant 0$ vs K: $\mu > 0$. Open ended tests based on t-statistics at stage n have been recommended testing the sign of $\lambda = \mu/\sigma$ (see [6] for example). Schwartz [5] suggests a sequential t-test for this model assuming an indifference region separates the hypotheses. He uses the Asymptotic Shapes Method which is appropriate for the case when the cost of sampling approaches 0.

There appears to be very little work done on the model and hypotheses posed here in either the asymptotic or non-asymptotic cases. Bayes tests or properties of such have not previously been studied. In this note we address the issue of bounded stopping times for a class of Bayes tests. Bounded

* Research supported by NSF-MCS-81-18564-01.

stopping time results have appeared in problems dealing with exponential families. References include Ray [4], Berk, Brown, and Cohen [1] and Cohen and Samuel-Cahn [2]. We consider three typical testing loss functions:

(1.1) linear combination of cost and 0-1 (1 for incorrect terminal decision); cost is assumed constant for each observation.

(1.2) linear combination of cost and $|\mu|$ in case of error.

(1.3) linear combination of cost and $|\mu|/\sigma$ in case of error.

The interesting and somewhat surprising results are as follows: For the normal-gamma family of priors (normal with mean 0) and generalized priors which are limiting cases of such, we find that the stopping time is bounded for losses (1.1) and (1.3). Furthermore, the condition determining whether the stopping time is bounded is *independent* of $u = \sum (x_i - \bar{x})^2$. For loss (1.2) the stopping time is not bounded.

Intuitive explanations for the results are discussed in Section 4. Preliminaries are given in Section 2 and the theorems are proved in Section 3.

2. Preliminaries. Let

$$X_n = \sum_{i=1}^{n} X_i, \quad u_n = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2, \quad s_n^2 = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2/(n-1),$$
$$h = 1/\sigma^2, \ \theta = (\mu, h),$$

$\theta \in \Theta$, where $\Theta = \{(\mu, h)\colon -\infty < \mu < \infty, \ 0 < h < \infty\}$. When the meaning is clear, we merely write $\bar{X}$, u instead of $\bar{X}_n$, u_n. The normal c.d.f. is F_θ and the standard normal c.d.f. is Φ. The actions are denoted by the pairs (n, a), where $n = 0, 1, 2, \ldots$ is the stopping time and $a = 1$ or 2 depending on whether H is accepted or rejected. The loss function (1.1) is written as

$$L_1(\theta, (n, 1)) = \begin{cases} cn+1 & \text{for } \mu > 0, \\ cn & \text{for } \mu \leqslant 0; \end{cases} \quad L_1(\theta, (n, 2)) = \begin{cases} cn+1 & \text{for } \mu \leqslant 0, \\ cn & \text{for } \mu > 0, \end{cases} \tag{2.1}$$

where $c > 0$.

The loss function for (1.2) is

$$L_2(\theta, (n, 1)) = \begin{cases} cn+\mu & \text{for } \mu > 0, \\ cn & \text{for } \mu \leqslant 0; \end{cases} \quad L_2(\theta, (n, 2)) = \begin{cases} cn-\mu & \text{for } \mu \leqslant 0, \\ cn & \text{for } \mu > 0. \end{cases} \tag{2.2}$$

The loss function for (1.3) is the same as (2.2) except that μ is replaced in (2.2) by μ/σ.

A normal-gamma prior distribution on (μ, h) with parameters (m, v, n_0, ν), denoted as $f_n(\mu|m, h, n_0) f_{\gamma 2}(h|v, \nu)$, is proportional to

$$e^{(-hn_0/2)(\mu-m)^2} h^{\delta(n_0)/2} e^{-\nu v h/2} h^{\nu/2-1}, \tag{2.3}$$

where $\delta(n_0) = 0$ if $n_0 = 0$ and $\delta(n_0) = 1$ if $n_0 > 0$ ([3], p. 300). The joint posterior distribution of (μ, h), given $\underset{\sim}{x}_n = (x_1, x_2, \ldots, x_n)$ depends on $(\bar{x}, u)$, is denoted by $dP(\mu, h|\bar{x}, u)$ and is proportional to

$$(2.4) \qquad e^{-[h(n_0+n)/2](\mu-[mn_0+n\bar{x}/(n_0+n)]^2} \times \\ \times e^{-(h/2)([vv+n_0m^2]+(u+n\bar{x}^2)-(n_0m+n\bar{x})^2/(n_0+n))} h^{(v+\delta(n_0)+(n-2))/2}.$$

The proof of the theorems in the next section will be based on some results of Ray [4]. A restatement of the relevant theorems appear in Cohen and Samuel-Cahn [2] and we state such theorems below.

Let $L(\theta, (n, \tau)) = cn + W(\theta, a)$, so that $W(\theta, a)$ is the loss due to a terminal decision. Let $L(\theta) = W(\theta, 1) - W(\theta, 2)$. For any given prior distribution ξ let $\tau_{\underset{\sim}{x}_n}\xi$ denote the posterior distribution of θ given $\underset{\sim}{x}_n$. For this problem $\tau_{\underset{\sim}{x}_n}\xi$ depends on $\underset{\sim}{x}_n$ only through $(\bar{x}_n, u_n)$, so we write $\tau_{\bar{x}_n, u_n}\xi$ for $\tau_{\underset{\sim}{x}_n}\xi$. For any distribution ξ and any integrable function $u(\theta)$ let

$$u(\xi) = \int_{\Theta} u(\theta)\, d\xi(\theta) = E_{\xi}\, u(\theta).$$

Let $L(\tau_{\bar{x},u}\xi) = E_{\tau_{\bar{x},n}\xi} L(\theta)$ be denoted by $L(\bar{x}_n, u_n)$. For the family of priors in (2.3) with $m = 0$, giving rise to the posterior distributions in (2.4) and for loss functions (1.1), (1.2) and (1.3), it is clear that, regardless of u_n, for every n, $L(\tau_{\bar{x},u}\xi) = 0$ if and only if $\bar{x} = 0$. Thus $(\bar{x}_n = 0, u_n)$ is called the *neutral boundary*. Now let

$$\varrho_0(\xi) = \min_{\tau} L(\xi, \tau)$$

and define $\lambda(\xi) = \varrho_0(\xi) - E_{\xi}\varrho_0(\tau_x \xi)$. Denote $\lambda(\tau_{\bar{x},u}\xi)$ by $\lambda(\bar{x}, u)$. The following theorems assume loss functions (1.1), (1.2) or (1.3) and priors (2.3) with $m = 0$. They are derivable from [4] and/or [2].

THEOREM 2.1. *We have*

$$(2.5) \qquad \lambda(0, u) = \left|\int_{\Theta} L(\theta) F_{\theta}(0)\, dP(\mu, h\,|\,\bar{x}, u)\right|.$$

If there exists an integer $\bar{n}$ such that $\lambda(0, u) \leqslant c$ uniformly in u for all $n \geqslant \bar{n}$, then the stopping time of the Bayes rule is bounded.

THEOREM 2.2. *Suppose that, for every integer n, $\lambda(0, u) > c$ for some values of u with positive probability. Then the Bayes test does not have bounded stopping time.*

3. Bounded stopping time results.

THEOREM 3.1. *Assume the loss function is (2.1). For prior (2.3) with $m = 0$, including cases where $n_0 = 0$ and/or $v = 0$, the stopping time of the Bayes test is bounded.*

Proof. Use (2.1), the definition of $L(\theta)$, (2.4) and (2.5) to compute

(3.1)
$$\lambda(0, u) = \Big|\int_0^\infty \int_{-\infty}^0 -\Phi(-\mu\sqrt{h})\, dP(\mu, h|0, u) + \int_0^\infty \int_0^\infty \Phi(-\mu\sqrt{h})\, dP(\mu; h|0, u)$$
$$= \int_0^\infty \int_0^\infty [\Phi(\mu\sqrt{h}) - \Phi(-\mu\sqrt{h})]\, dP(\mu, h).$$

Use (2.4) with $m = 0$, $\bar{x} = 0$, so that (3.1) becomes

(3.2)
$$(1/\sqrt{2\pi})\big[(n_0+n)^{1/2}/\Gamma([\nu+\delta(n_0)+(n-1)]/2)\, 2^{(\nu+\delta(n_0)+(n-1))/2}\big] \times$$
$$\times \int_0^\infty \int_0^\infty [\Phi(u\sqrt{h}) - \Phi(-\mu\sqrt{h}]\, e^{-[h(n_0+n)/2]\mu^2}\, e^{-(h/2)[\nu v+u]}\, h^{(\nu+\delta(n_0)+(n-1))/2}\, d\mu\, dh$$
$$= E_{h,n}\Big\{\int_0^\infty [\Phi(y/\sqrt{n_0}+n) - \Phi(-y/\sqrt{n_0}+n)](e^{-y^2/2}/\sqrt{2\pi})\, dy\Big\},$$

where $E_{h,n}$ denotes expectation over the marginal distribution of h which is a $\Gamma(\nu+\delta(n_0)+(n-1), \nu v+u)$ distribution. From (3.2) however we note that the bracketed term is independent of h and so $\lambda(0, u)$ does not depend on u. Clearly, for all n sufficiently large, $\lambda(0, u) < c$ uniformly in u, and the proof follows from Theorem 2.1.

Remark 2.2. The fact that $\lambda(0, u)$ does not depend on u implies that computable bounds on the stopping time can be found as in [4].

THEOREM 3.3. *Assume the loss function is* (2.2). *For prior* (2.3) *with* $m = 0$ *the stopping time of the Bayes test is not bounded.*

Proof. This time use (2.2) in computing (2.5) and follow the steps of the proof in Theorem 3.1 to find that

(3.2)
$$\lambda(0, u) = \int_0^\infty y[\Phi(y/\sqrt{n+n_0}) - \Phi(-y/\sqrt{n+n_0})] \times$$
$$\times (e^{-y^2/2}/\sqrt{2\pi})\, dy\, E_{h,n} h^{-1/2}(n_0+n)^{-1/2}$$
$$= \Big[\int_0^\infty y[\Phi(y/\sqrt{n+n_0}) - \Phi(-y/\sqrt{n+n_0})](e^{-y^2/2}/\sqrt{2\pi})\, dy\Big] \times$$
$$\times [(\nu v+u)/(n_0+n)]^{1/2}\, \Gamma([\nu+\delta(n_0)+(n-2)]/2)/\Gamma([\nu+\delta(n_0)+(n-1)]/2).$$

From (3.2) we see that for any n there exists a u-set of positive probability for which $\lambda(0, u) > c$. Now apply Theorem 2.2.

COROLLARY 3.4. *Assume the loss function is* (1.3). *For prior* (2.3), *with* $m = 0$, *the stopping time of the Bayes test is bounded.*

Proof. The proof follows as in the previous theorems and it turns out that, as in Theorem 3.1, $\lambda(0, u)$ is independent of u.

4. Discussion. We offer some intuitive explanations for the results obtained in Section 3.

Suppose the loss is (1.1). If $X \sim N(\mu, \sigma^2)$, we feel that if σ^2 is small we do not need n to be large to gain information about μ. However, if $\mu|\sigma^2 \sim N(0, \sigma^2)$, small σ^2 means small $|\mu|$ which in turn means we need large n to distinguish between H and K. Hence there is a balancing effect, so that the determination of whether the stopping time is bounded does not depend on s^2.

If the loss is (1.2), the penalty is too severe for an error in the terminal decision for large $|\mu|$. In this case if σ^2 is large, we need large n, and if σ^2 is small, we do not need n large. This isaaccomplished by n depending on s^2, the estimator of σ^2. We see that (3.2) reflects this behavior. In connection with the loss in (1.3) we again find a balancing effect between the distribution of X and the prior distribution.

Another intuitive explanation is as follows: If the loss in terminal decision is bounded, then stop before n gets too big, otherwise the penalty from cn is too severe. On the other hand, for absolute error loss there is a trade off in severity between loss due to terminal decision and samplig cost. If $|\mu|/\sigma$ is the loss due to terminal decision, then, although large σ means large n, the terminal decision loss is neutralized for large σ so that n should not actually have to be large.

The author would like to thank Professor M. Woodroofe for a helpful discussion.

REFERENCES

[1] R. H. Berk, L. D. Brown and A. Cohen, *Bounded stopping times for a class of sequential tests*, Ann. Statist. 9 (1981), p. 834-845.

[2] A. Cohen and E. Samuel-Cahn, *Necessary and sufficient conditions for bounded stopping time of sequential Bayes tests in one parameter exponential families*, Commun. Statist. Sequential Analysis 1 (1982), p. 89-99.

[3] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory*, M. I. T. Press, Cambridge 1972.

[4] S. N. Ray, *Bounds on the maximum sample size of a Bayes sequential procedure*, Ann. Math. Statis. 36 (1965), p. 859-878.

[5] G. Schwartz, *A sequential Student test*, ibidem 42 (1971), p. 1003-1009.

[6] R. A. Wijsman, *Stopping time of invariant sequential probability ratio tests*, Developments in Statistics 2 (1979).

Department of Statistics
Rutgers University
Hill Center, Busch Campus
New Brunswick, NJ 08903
U.S.A.

Received on 13. 3. 1985