The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 313–315

A SECOND-ORDER APPROXIMATION TO OPTIMAL SAMPLING REGIONS[1]

BY GIDEON SCHWARZ

University of California, Berkeley, and Hebrew University

1. Introduction. In an earlier paper [3] an asymptotic description of the optimal sequential testing regions for separated hypotheses was given. It involved an asymptotic formula for curves of constant posterior risk. The need for a second-order correction term for this formula was demonstrated by Fushimi [1], who found such a term for normal and binomial sampling distributions, with their conjugate *a priori* distributions, and a truncated linear loss function. In this paper we find the general correction term. While it does not depend on the sampling distributions as long as the latter form an exponential family, it does vary with the loss function, and depends also on some properties of the *a priori* distribution which carry over into the *a posteriori* distributions: the locations of its atoms and the zeros of its density.

2. Preliminaries and statement. We assume a sampling distribution with density $f(x, \theta) = e^{\theta x - b(\theta)}$ with respect to some measure. The parameter θ ranges over the interior Θ of the natural parameter space. There $b(\theta)$ has all derivatives, and since $b'(\theta)$ and $b''(\theta)$ are the expectation and variance of the sampling variable, $b(\theta)$ is a convex function. For expository convenience we assume "one sided" hypotheses H_0: $\theta \leqq M'$, H_1: $\theta \geqq M > M'$ and an *a priori* distribution W that dominates Lebesgue measure on Θ. The local behaviour of W at θ is described by a number $-1 \leqq \tau(\theta) < \infty$, as follows:

(a) if $W(\{\theta_0\}) > 0$, $\tau(\theta_0) = -1$;

(b) if, in a neighbourhood of θ_0, W has a density of the form $|\theta - \theta_0|^\alpha g(\theta)$ with $g(\theta)$ bounded away from 0 and ∞, $\tau(\theta_0) = \alpha$.

We assume the τ is defined at every $\theta \varepsilon \Theta$ either by (a) or by (b). This restricts the generality of the *a priori* distributions somewhat, but it seems general enough for any conceivable application.

The loss for deciding "H_0" when the true parameter value is θ, is given on H_1 by $l(\theta) = (\theta - M)^\eta d(\theta)$ with d bounded away from 0 and ∞. If $\tau(M) > -1$, we assume $\eta + \tau(M) > -1$, to avoid infinite Bayes risks; if $\tau(M) = -1$ any bounded loss function would do, but we assume in this case that it is positive at M, and define $\eta = 0$ independently of the behaviour of $l(\theta)$ near M.

The posterior risk of deciding "H_0" after having made n observations with sum $S_n = kn$ is

$$(*) \qquad R_0(n, S_n) = \int_{\theta \geqq M} l(\theta)\, e^{(k\theta - b(\theta))n}\, dW(\theta) \Big/ \int_\Theta e^{(k\theta - b(\theta))n}\, dW(\theta).$$

Received 18 July 1968.

[1] This paper was prepared with the partial support of National Science Foundation Grant GP-7454.