AN OPTIMAL PROPERTY OF THE LIKELIHOOD RATIO STATISTIC

R. R. BAHADUR
UNIVERSITY OF CHICAGO

1. Introduction

Let $s = (x_1, x_2, \cdots, \text{ad inf})$ be a sequence of independent and identically distributed observations on a variable x with distribution depending on a parameter θ taking values in a set Θ. Let Θ_0 be a subset of Θ and consider the null hypothesis that θ is in Θ_0. For each n, let $T_n = T_n(x_1, \cdots, x_n)$ be a real-valued statistic such that, in testing the hypothesis, large values of T_n are significant. For any given s, let $L_n(s)$ be the level attained by T_n in the given case; that is, $L_n(s)$ is the maximum probability (consistent with θ in Θ_0) of obtaining a value of T_n as large or larger than $T_n(s)$. Then, in typical cases, L_n is asymptotically distributed uniformly over $(0, 1)$ in the null case, and L_n tends to zero in probability, or perhaps even with probability one, in the nonnull case. The rate at which L_n tends to zero when a given nonnull θ obtains is a measure of the asymptotic efficiency of T_n against that θ. It is shown in this paper (under very mild restrictions on the family of possible distributions of x) that L_n cannot tend to zero at a rate faster than $[\rho(\theta)]^n$ when a nonnull θ obtains; here ρ is a parametric function defined in terms of the Kullback-Leibler information numbers such that, in typical cases, $0 < \rho < 1$ (theorem 1). It is also shown (under much more restrictive conditions on the distributions of x) that if $\hat{T}_n$ is (any strictly decreasing function of) the likelihood ratio statistic of Neyman and Pearson [1], and $\hat{L}_n$ is the level attained by $\hat{T}_n$, then $\hat{L}_n$ tends to zero at the rate $[\rho(\theta)]^n$ in the nonnull case (theorem 2). In short, the likelihood ratio statistic is an optimal sequence in terms of exact stochastic comparison as described and exemplified in [2], [3], and [4].

Theorems 1 and 2 are stated more precisely in section 2. Section 3 contains a discussion of these theorems. Proofs are given in sections 4 and 5.

2. Theorems

Let X be a space of points x, $\mathcal{B}$ a σ-field of sets of X, and for each point θ in a set Θ, let P_θ be a probability measure on $\mathcal{B}$. Let Θ_0 be a given subset of Θ.

This research was supported in part by Research Grant No. NSF-GP3707 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation and in part by the Statistics Branch of the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.