The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1695–1699

ON A THEOREM OF BAHADUR ON THE RATE OF CONVERGENCE OF TEST STATISTICS

BY M. RAGHAVACHARI

Carnegie–Mellon University

0. Summary. Let $x_1, x_2, \cdots, x_n$ be n independent and identically distributed random variables whose distribution depends on a parameter θ, $\theta \in \Theta$. Let Θ_0 be a subset of Θ and consider the test of the hypothesis that $\theta \in \Theta_0$. $L_n(x_1, \cdots, x_n)$ is the level attained by a test statistic $T_n(x_1, \cdots, x_n)$ in the sense that it is the maximum probability under the hypothesis of obtaining a value as large or larger than T_n where large values of T_n are significant for the hypothesis. Under some assumptions Bahadur [3] showed that where a non-null θ obtains L_n cannot tend to zero at a rate faster than $[\rho(\theta)]^n$ where ρ is a function defined in terms of Kullback–Liebler information numbers. In this paper this result has been shown to be true without any assumptions whatsoever (Theorem 1). Some aspects of the relationship between the rate of convergence of L_n and rate of convergence of the size of the tests are also studied and an equivalence property is shown (Theorem 2).

1. Introduction and main results. To facilitate reference with the relevant work on this topic, we use the same notation as is employed in [3] and [4]. Let X be a space of points x, $\mathscr{B}$ a σ-field of sets of X and for each point θ in a set Θ, let P_θ be a probability measure on $\mathscr{B}$. Let Θ_0 be a given subset of Θ. If P_θ admits a density with respect to P_{θ_0}, say $dP_\theta = f(x)\, dP_{\theta_0}$, let

$$K(\theta, \theta_0) = E_\theta(\log f(x)); \tag{1}$$

otherwise let $K = \infty$. K is called the Kullback–Liebler information number. It is clear that $0 \leqq K \leqq \infty$ and $K = 0$ if and only if $P_\theta \equiv P_{\theta_0}$. For each θ in Θ, let

$$J(\theta) = \inf\{K(\theta, \theta_0) : \theta_0 \in \Theta_0\}. \tag{2}$$

Then J is well defined over Θ and $0 \leqq J \leqq \infty$. $J = 0$ on Θ_0.

Let $s = (x_1, x_2, \cdots)$ be a sequence of independent and identically distributed observations on x. Let $P_\theta^{(\infty)}$ be the probability distribution of s in its sample space when θ is true. As in [3], for convenience we shall sometimes write P_θ for $P_\theta^{(\infty)}$. For each $n = 1, 2, \cdots$, let $T_n(s)$ be an extended real-valued measurable function of the observations $x_1, \cdots, x_n$. For each θ, let

$$F_n(t, \theta) = P_\theta(T_n(s) < t) \tag{3}$$

and let

$$G_n(t) = \inf\{F_n(t, \theta) : \theta \in \Theta_0\} \qquad -\infty < t < \infty \tag{4}$$

Define

$$L_n(s) = 1 - G_n(T_n(s)). \tag{5}$$

Received February 28, 1969; revised February 16, 1970.

$L_n(s)$ has been called by Bahadur the level attained by T_n and has this interpretation in the framework of tests of hypotheses. Consider the testing problem $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta - \Theta_0$ for which large values of T_n are significant. Then $L_n(s)$ is an index of the performance of this test in the sense it is the maximum probability of obtaining a value of T_n as large or larger than $T_n(s)$ when the hypothesis is true. In many cases $L_n \to 0$ with probability one or in probability as $n \to \infty$ when θ obtains with $\theta \in \Theta - \Theta_0$. Further in some cases there exists a function $c(\theta)$ defined for $\theta \in \Theta - \Theta_0$ such that $0 < c < \infty$ and

$$n^{-1} \log L_n \to -c(\theta)/2 \quad \text{as} \quad n \to \infty \tag{6}$$

with probability one or in probability.

In order to distinguish the two modes of convergence in (6), let us call $c(\theta)$ the strong exact slope if (6) happens with probability one and the weak exact slope if (6) happens in probability only. For details regarding comparison of test statistics through the slopes see [4], [2] and [5].

Bahadur [3] proved that for each θ in $\Theta - \Theta_0$

$$\liminf_{n \to \infty} n^{-1} \log L_n(s) \geqq -J(\theta) \tag{7}$$

with probability one when θ obtains if certain assumptions (Assumption 1 and Assumption 2 of [3]) are satisfied. Bahadur stated that (7) holds under weaker assumptions and posed the question whether (7) holds without any assumptions whatsoever. The following theorem answers this question in the affirmative.

THEOREM 1. *For each* $\theta \in \Theta - \Theta_0$, (7) *holds with probability one when* θ *obtains.*

REMARK 1. Bahadur [3] proved an optimal property of the likelihood ratio statistic under assumptions which include an Assumption 2. In view of Theorem 1 above, Theorem 2 of [3] is valid without the latter assumption.

REMARK 2. Bahadur gives details of the relationship between exact slopes and the size of the test statistics. Assume that the sequence $\{T_n\}$ is such that, for any given p, $0 < p < 1$ and $\theta \in \Theta - \Theta_0$ there exist constants k_n such that $P_\theta\{T_n \geqq k_n\} \to p$ as $n \to \infty$.

For each n, let $\alpha_n = \sup\{P_{\theta_0}(T_n \geqq k_n): \theta_0 \in \Theta_0\}$. α_n is the size of $\{T_n \geqq k_n\}$ in testing Θ_0 against $\Theta - \Theta_0$. It can be readily seen (the method of proof parallels that of Proposition 11 of Bahadur [4]) that Theorem 1 implies $\liminf_{n \to \infty} n^{-1} \log \alpha_n \geqq -J(\theta)$ when θ obtains. This result can also be deduced from the proof of Proposition 9 of Bahadur [4].

Bahadur has shown [4] that if the sequence $\{T_n\}$ has a strong exact slope $c(\theta)$, then $n^{-1} \log \alpha_n \to -c(\theta)/2$ as $n \to \infty$ for every $0 < p < 1$. The following theorem gives necessary and sufficient conditions in terms of the rate of convergence of α_n for a sequence $\{T_n\}$ to have weak exact slope.

THEOREM 2. *Under assumptions and notations given in Remark* 2, $n^{-1} \log L_n(s) \to h(\theta)$ *in probability as* $n \to \infty$ *when* θ *obtains with* $\theta \in \Theta - \Theta_0$ *if and only if* $n^{-1} \log \alpha_n \to h(\theta)$ *as* $n \to \infty$ *for every* p, $0 < p < 1$ *and* $h(\theta)$ *is independent of* p.

The above theorem may be used sometimes to obtain weak exact slopes when they exist. Asymptotic efficiency of test statistics based on weak exact and weak approximate (see, for example, Bahadur [2] or Gleser [5] for definition) slopes have been studied in [2] and [5]. The following simple example illustrates the application of Theorem 2. In fact, for this example the strong exact slope exists and has been computed by Bahadur in [1].

EXAMPLE. $x_1, x_2, \cdots, x_n$ are independent and identically distributed as normal with mean θ and unit variance. Consider testing $H:\theta = 0$ vs. $H_1:\theta > 0$. Let $T_n(x_1, \cdots, x_n) = \bar{x}$. Fix a $\theta > 0$ and a p such that $0 < p < 1$. In this case $k_n = (n^{-\frac{1}{2}}\Phi^{-1}(1-p)+\theta)$, $\alpha_n = 1-\Phi[\Phi^{-1}(1-p)+\theta n^{\frac{1}{2}}]$ where Φ is the cdf of the standard normal distribution. It can be checked easily that $n^{-1}\log\alpha_n \to -\theta^2/2$ as $n \to \infty$ for every p such that $0 < p < 1$. Theorem 2 applies and we have that $n^{-1}\log L_n(s)$ converges to $-\theta^2/2$ in probability.

PROOF OF THEOREM 1. Choose and fix a θ in $\Theta - \Theta_0$. If $J(\theta) = \infty$, (7) holds trivially. Suppose then that $J(\theta) < \infty$. Choose and fix $\varepsilon > 0$. By the definition of J there exists θ_0 in Θ_0 such that

$$K(\theta, \theta_0) < J(\theta) + \varepsilon. \tag{8}$$

Since $K(\theta, \theta_0) < \infty$, P_θ is dominated by P_{θ_0}, say $dP_\theta = f(x)\,dP_{\theta_0}$, $0 \leqq f \leqq \infty$. Let $g(x) = \log f(x)$, $-\infty \leqq g < \infty$. Note that $P_\theta(-\infty < g < \infty) = 1$ and $K(\theta, \theta_0) = E_\theta(g)$.

Choose constants γ, δ and ρ, all in (0, 1), with $\delta < \rho$. Let $h_n = n^{-1}\sum_{i=1}^n g(x_i)$. We have $h_n \to K$ as $n \to \infty$ with $P_\theta{}^{(\infty)}$—probability one. By Egoroff's theorem there exists a set A of sequences s such that $P_\theta{}^{(\infty)}(A) > 1-\gamma$ and such that $h_n(s) \to K$ uniformly on A. Hence there exists a finite constant N such that

$$K-\delta < h_n(s) < K+\delta \quad \text{for} \quad n \geqq N, \qquad s \in A. \tag{9}$$

Now define $L_n{}^*(s) = 1 - F_n(T_n, \theta_0)$. Let

$$B_n = \{s: L_n{}^* < e^{n(-K-\rho)}\} \tag{10}$$

and let

$$C_n = \{s: -\infty < h_n < \infty\}. \tag{11}$$

Then

$$P_\theta(C_n) = 1 \quad \text{and} \quad P_{\theta_0}(B_n) \leqq e^{-n(K+\rho)} \tag{12}$$

for each n; the second part of (12) follows from Bahadur [3] page 20, the first sentence in the proof of Lemma 3. Note also that $dP_\theta{}^{(n)} = e^{nh_n}\,dP_{\theta_0}^{(n)}$ on $X^{(n)}$, and hence

$$dP_{\theta_0}^{(n)} = e^{-nh_n}\,dP_\theta{}^{(n)} \quad \text{on} \quad C_n, \tag{13}$$

by (11). Now for $n \geqq N$

$$
\begin{aligned}
(14)\qquad P_\theta^{(\infty)}(A \cap B_n) &= P_\theta^{(\infty)}(A \cap B_n \cap C_n) && \text{by (12)}\\
&= \int_{A \cap B_n \cap C_n} e^{nh_n} e^{-nh_n}\, dP_\theta^{(\infty)} \\
&\leqq e^{n(K+\delta)} \int_{A \cap B_n \cap C_n} e^{-nh_n}\, dP_\theta^{(\infty)} && \text{by (9)}\\
&\leqq e^{n(K+\delta)} \int_{B_n \cap C_n} e^{-nh_n}\, dP_\theta^{(\infty)} \\
&= e^{n(K+\delta)} \int_{B_n \cap C_n} e^{-nh_n}\, dP_\theta^{(n)} \\
&= e^{n(K+\delta)} P_{\theta_0}^{(n)}(B_n \cap C_n) && \text{by (13)}\\
&\leqq e^{n(K+\delta)} P_{\theta_0}^{(n)}(B_n) \\
&\leqq e^{n(\delta-\rho)} && \text{by (12).}
\end{aligned}
$$

Since $\delta < \rho$, it follows from (14) that $\sum_n P_\theta(A \cap B_n) < \infty$. Hence with $D = \limsup_{n\to\infty} (A \cap B_n)$, we have $P_\theta(D) = 0$. Let $E = \limsup_{n\to\infty} (B_n)$. Then $P_\theta(E) \leqq P_\theta(A^c) + P_\theta(D) = P_\theta(A^c) < \gamma$. Since E does not depend on γ, and γ is arbitrary, $P_\theta(E) = 0$. It now follows from (10), and $L_n(s) \geqq L_n^*(s)$, that with $P_\theta^{(\infty)}$—probability one, $n^{-1} \log L_n \geqq -K - \rho$ for all sufficiently large n. Hence, by (8),

$$(15)\qquad \liminf_{n\to\infty} n^{-1} \log L_n \geqq -J(\theta) - \varepsilon - \rho$$

with probability one when θ obtains. Since ε and ρ are arbitrary, (7) holds with probability one when θ obtains.

Proof of Theorem 2. Suppose $n^{-1} \log L_n(s) \to h(\theta)$ in probability as $n \to \infty$ when θ in $\Theta - \Theta_0$ obtains. Suppose, if possible, $\liminf_{n\to\infty} n^{-1} \log \alpha_n < h(\theta)$. This would imply that there exists a sequence $m_1 < m_2 < \cdots$ of positive integers m_r such that $m_r^{-1} \log \alpha_{m_r} < h(\theta)$. Take ε to be a positive constant. We have

$$
\begin{aligned}
&P_\theta[m_r^{-1} \log L_{m_r}(s) > h(\theta) - \varepsilon] \\
(16)\qquad &\leqq P_\theta[m_r^{-1} \log L_{m_r}(s) > m_r^{-1} \log \alpha_{m_r}] \\
&= P_\theta[L_{m_r}(s) > \alpha_{m_r}] \\
&\leqq P_\theta[T_{m_r}(s) < k_{m_r}]
\end{aligned}
$$

since for each n, $L_n(s) = 1 - G_n(T_n(s))$ and $\alpha_n = 1 - G_n(k_n)$ and $G_n(t)$ is monotone non-decreasing in t. Letting $r \to \infty$ in (16) and observing that $P_\theta[T_{m_r}(s) < k_{m_r}] \to 1-p$, we have $1 \leqq 1-p$ which is not true since $0 < p < 1$. Thus

$$\liminf_{n\to\infty} n^{-1} \log \alpha_n \geqq h(\theta).$$

In a similar fashion it follows that $\limsup_{n\to\infty} n^{-1} \log \alpha_n \leqq h(\theta)$. This proves the "if" part.

Now assume that $n^{-1}\log\alpha_n \to h(\theta)$ as $n\to\infty$ for every p, $0<p<1$, and that $h(\theta)$ is independent of p. Take any $\varepsilon>0$. We have for all sufficiently large n

$$\begin{aligned} &P_\theta[n^{-1}\log L_n(s) < h(\theta)-\varepsilon] \\ (17)\qquad &\leqq P_\theta[n^{-1}\log L_n(s) < n^{-1}\log\alpha_n] \\ &\leqq P_\theta[T_n(s) \geqq k_n]. \end{aligned}$$

Letting $n\to\infty$ in (17) and noting that $P_\theta[T_n(s)\geqq k_n]\to p$, we have

$$(18)\qquad \limsup_{n\to\infty} P_\theta[n^{-1}\log L_n(s) < h(\theta)-\varepsilon] \leqq p$$

for all $0<p<1$. Since $h(\theta)$ is independent of p and $L_n(s)$ clearly does not depend on p, we have from (18) that $\limsup_{n\to\infty} P_\theta[n^{-1}\log L_n(s) < h(\theta)-\varepsilon] = 0$ which implies that $P_\theta[n^{-1}\log L_n(s) < h(\theta)-\varepsilon]\to 0$ as $n\to\infty$. A similar argument shows that for every $\varepsilon>0$, $P_\theta[n^{-1}\log L_n(s) > h(\theta)+\varepsilon]\to 0$ as $n\to\infty$. This establishes the "only if" part and hence the theorem.

Acknowledgment. The author is indebted to Professor R. R. Bahadur for helpful comments on the earlier version of this paper.

REFERENCES

[1] BAHADUR, R. R. (1960a). Simultaneous comparison of the optimum and sign tests of a normal mean in *Contributions to Probability and Statistics. Essays in Honor of Harold Hotelling.* Stanford Univ. 79–88.

[2] BAHADUR, R. R. (1960b). Stochastic comparison of tests. *Ann. Math. Statist.* **31** 276–295.

[3] BAHADUR, R. R. (1967a). An optimal property of the likelihood ratio statistic. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 13–26.

[4] BAHADUR, R. R. (1967b). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–324.

[5] GLESER, L. J. (1964). On a measure of test efficiency proposed by R. R. Bahadur. *Ann. Math. Statist.* **35** 1537–1544.