The Annals of Statistics
1975, Vol. 3, No. 1, 234–240

THE RATE OF CONVERGENCE OF CONSISTENT POINT ESTIMATORS[1]

BY JAMES C. FU

University of Toronto

The rate at which the probability $P_\theta\{|t_n - \theta| \geqq \varepsilon\}$ of consistent estimator t_n tends to zero is of great importance in large sample theory of point estimation. The main tools available at present for finding the rate are Bernstein–Chernoff–Bahadur's theorem and Sanov's theorem. In this paper, we give two new techniques for finding the rate of convergence of certain consistent estimators. By using these techniques, we have obtained an upper bound for the rate of convergence of consistent estimators based on sample quantities and proved that the sample median is an asymptotically efficient estimator in Bahadur's sense if and only if the underlying distribution is double-exponential. Furthermore, we have proved that the Bahadur asymptotic relative efficiency of sample mean and sample median coincides with the classical Pitman asymptotic relative efficiency.

1. Introduction and summary. Let $\mathscr{X}$ be an abstract sample space with a Borel field $\mathscr{B}$ and σ-finite measure μ on $\mathscr{B}$. Let $\{P_\theta : \theta \in \Theta\}$ be a family of probability measures on $\mathscr{B}$, where the parameter space Θ is an open subset of the real line. We assume, for each $\theta \in \Theta$, that P_θ has a density function $f(\cdot \mid \theta)$. Let $s = (x_1, x_2, \cdots$ ad inf) be a sequence of independent identically distributed (i.i.d.) random variables with common probability measure P_θ, where θ belongs to Θ.

Let t_n be any consistent estimator, and let $\alpha(t_n, \varepsilon, \theta) = P_\theta\{|t_n - \theta| > \varepsilon\}$ for $\varepsilon > 0$. It has been proved ([1], [2] and [4]) that the following inequalities hold:

$$\liminf_{n\to\infty} \frac{1}{n} \log \alpha(t_n, \varepsilon, \theta) \geqq -\inf_{\theta'} \{K(\theta', \theta); |\theta' - \theta| > \varepsilon\}, \tag{1.1}$$

and

$$\liminf_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{\varepsilon^2 n} \log \alpha(t_n, \varepsilon, \theta) \geqq -\tfrac{1}{2} I(\theta), \tag{1.2}$$

where $K(\theta', \theta)$ is the Kullback–Liebler information of P_θ with respect to $P_{\theta'}$ and $I(\theta)$ is Fisher's information of P_θ; and that the maximum likelihood estimator $\hat\theta_n$ is asymptotically optimal in the following sense; i.e.,

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n\varepsilon^2} \log \alpha(\hat\theta_n, \varepsilon, \theta) = -\tfrac{1}{2} I(\theta). \tag{1.3}$$

Received December 1972; revised January 1974.

[1] This work was sponsored in part by the National Research Council of Canada under grant NRC A-3011 and the Air Force of U. S. under AFOSR contact F44620-70-C-066.

AMS 1970 *subject classifications.* Primary 62E20; Secondary 62E10.

Key words and phrases. Consistent estimator, the rate of convergence, asymptotically efficient, location parameter, sample quantiles.

In general, the rate at which the probability $\alpha(t_n, \varepsilon, \theta)$ tends to zero for a consistent estimator is very difficult to obtain. The main tools available at present for finding the rate are the Bernstein–Chernoff–Bahadur theorem [3] and Sanov's theorem [6]. In Section 2, two new techniques for finding the rate at which $\alpha(t_n, \varepsilon, \theta)$ tends to zero are introduced, and a generalization of Bahadur's inequality is obtained. Section 3 studies the rate of convergence of certain consistent estimators for the location parameter, and proves that the sample median is an asymptotically efficient estimator in Bahadur's sense if and only if the underlying distribution is double-exponential. Section 4 shows that the Bahadur asymptotic relative efficiency of the mean $\bar{x}_n$ and the median $x_{(\frac{1}{2})}$ coincides with the classical Pitman asymptotic relative efficiency for the location parameter.

2. Main theorems.

THEOREM 2.1. *Let t_n be an estimator for θ having density function $f_n(t \mid \theta)$ which satisfies the following conditions*:

(A) *for each θ, $f_n(t \mid \theta)$ is continuous and mode(s) of $f_n(t \mid \theta)$ approach θ as $n \to \infty$*;
(B) *for each θ, $r(t, \theta) = \lim_{n\to\infty} n^{-1} \log f_n(t \mid \theta)$ is continuous in t at $\theta \pm \varepsilon$*;
(C) *there exists a sequence $\{a_n\}$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \log a_n = 0 \qquad \text{and} \qquad \limsup_{n\to\infty} \frac{1}{n} \log \frac{P_\theta\{|t_n - \theta| \geqq \varepsilon + a_n\}}{f_n(\theta_n \pm \varepsilon \mid \theta)} = 0 \,.$$

Then

$$\lim_{n\to\infty} \frac{1}{n} \log \alpha(t_n, \varepsilon, \theta) = \max\,(r(\theta + \varepsilon, \theta), r(\theta - \varepsilon, \theta)) \,. \tag{2.1}$$

PROOF. Let δ be an arbitrarily small positive constant. It follows from conditions (A) and (C) that the inequalities

$$\delta f_n(\theta + \varepsilon + \delta \mid \theta) \leqq P_\theta\{t_n > \theta + \varepsilon\} \leqq (a_n + e^{n\delta}) f_n(\theta + \varepsilon \mid \theta) \tag{2.2}$$

and

$$\delta f_n(\theta - \varepsilon - \delta \mid \theta) \leqq P_\theta\{t_n < \theta - \varepsilon\} \leqq (a_n + e^{n\delta}) f_n(\theta - \varepsilon \mid \theta) \tag{2.3}$$

hold for all n sufficiently large. Since δ is arbitrary the result (2.1) follows immediately from condition (B) and equations (2.2) and (2.3). □

If the parameter space Θ is bounded then the result (2.1) holds without condition (C).

In many statistical estimation problems, the log-likelihood ratio statistic

$$\lambda_n(t_n(s), S', \theta) = \frac{1}{n} \log \frac{f_n(t_n(s) \mid \theta')}{f_n(t_n(s) \mid \theta)}$$

of a consistent estimator t_n usually converges (under $P_{\theta'}$) almost surely to some nonzero positive constant, say, $R(\theta', \theta)$. The following theorem shows, for certain types of consistent estimators, that the rate at which $\alpha(t_n, \varepsilon, \theta)$ tends to zero is the same as the rate at which $\exp[-n \inf_{\theta'} \{R(\theta', \theta); |\theta' - \theta| > \varepsilon\} + o(n)]$ tends to zero.

THEOREM 2.2. *Let t_n be a consistent estimator having density function $f_n(t \mid \theta)$ and satisfying the following conditions*:

(D) *for every θ, $\theta' \in \Theta$ there exists a constant $R(\theta', \theta)$ such that* $\lim_{n\to\infty} \lambda_n(t_n(s), \theta', \theta) = R(\theta', \theta)$ a.s. $[P_{\theta'}]$;

(E) *for each n, $\lambda_n(t, \theta', \theta)$ is a non-decreasing function of t if $\theta' > \theta$, and is non-increasing if $\theta' < \theta$* (monotonic likelihood ratio).

Then

$$\lim_{n\to\infty} \frac{1}{n} \log \alpha(t_n, \varepsilon, \theta) = -\inf_{\theta'} \{R(\theta', \theta) : |\theta' - \theta| > \varepsilon\} \geqq -\inf_{\theta'} \{K(\theta', \theta) : |\theta' - \theta| > \varepsilon\} . \tag{2.4}$$

PROOF. The inequality part of (2.4) is an immediate consequence of Theorem 4.1 of Kullback [5], condition (D) and Fatou's lemma.

It follows from condition (D) and the consistency of the estimator t_n that

$$\liminf_{n\to\infty} \frac{1}{n} \log \alpha(t_n, \varepsilon, \theta) \geqq -\inf_{\theta'} \{R(\theta', \theta) : |\theta' - \theta| > \varepsilon\} . \tag{2.5}$$

It follows directly from condition (E) that

$$\limsup_{n\to\infty} \frac{1}{n} \log \alpha(t_n, \varepsilon, \theta) \leqq -\inf_{\theta'} \{R(\theta', \theta) : |\theta' - \theta| > \varepsilon\} . \tag{2.6}$$

The equality part of (2.4) follows from (2.5) and (2.6). □

THEOREM 2.3. *Let t_n be a consistent estimator for θ based on k statistics $u_1(s), u_2(s), \cdots, u_k(s)$. Assume that $u_1(s), \cdots, u_k(s)$ have joint density function $f_n(u_1, \cdots, u_k \mid \theta)$ which satisfies the following condition*

(F) *for every θ, $\theta' \in \Theta$, there exists a constant $R(\theta', \theta)$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{f_n(u_1, \cdots, u_k \mid \theta')}{f_n(u_1, \cdots, u_k \mid \theta)} = R(\theta', \theta) \qquad \text{a.s.} \quad [P_{\theta'}]$$

as $n \to \infty$. Then

$$\liminf_{n\to\infty} \frac{1}{n} \log \alpha(t_n, \varepsilon, \theta) \geqq -\inf_{\theta'} \{R(\theta', \theta) : |\theta' - \theta| > \varepsilon\} \geqq -\inf_{\theta'} \{K(\theta', \theta) : |\theta' - \theta| > \varepsilon\} . \tag{2.7}$$

PROOF. The proof is the same as the first part of the proof of Theorem 2.2.

It is well known (Kullback [5]) that $K(\theta', \theta)$ is a locally convex function of θ': i.e., $K(\theta', \theta) = \frac{1}{2}I(\theta)(\theta' - \theta)^2 + o((\theta' - \theta)^2)$. In many cases, $R(\theta', \theta)$ is also a locally convex function of θ' and there exists a positive constant $\beta(\theta)$ such that

$R(\theta', \theta) = \frac{1}{2}\beta(\theta)(\theta' - \theta)^2 + o((\theta' - \theta)^2)$. If this is the case, we may conclude

$$\text{(2.8)} \qquad \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{\varepsilon^2 n} \log \alpha(t_n, \varepsilon, \theta) \geqq -\tfrac{1}{2}\beta(\theta) \geqq -\tfrac{1}{2}I(\theta) .$$

3. The rate of convergence of certain consistent estimators for the location parameter. Let $\{x_{[i]} : i = 1, \cdots, n\}$ be the ordered observations of a sample of i.i.d. observations $(x_1, \cdots, x_n)$ from a common population having density function $f(x - \theta)$ with location parameter $\theta \in \Theta$. We assume $f(x)$ satisfies the following conditions: (a) $f(x) > 0$ for all x, and $f(\infty) = \lim_{x\to\infty} f(x) = 0$, (b) $f(-x) = f(x)$, (c) $f(x)$ is continuous and piece-wise differentiable, and (d) f has finite Fisher's information $I(f)$ for the location parameter. This family of densities covers Normal, Logistic, Double Exponential, and Cauchy, of interest in both statistical theory and practice.

For every f satisfying conditions (a) through (d) we define, for $\lambda \in [0, \frac{1}{2}]$,

$$\text{(3.1)} \qquad \beta(f, \lambda) = 2\left[\frac{f^2(\overline{1-\lambda})}{\lambda} + \int_0^{\overline{1-\lambda}} \frac{(f'(x))^2}{f(x)}\, dx\right]$$

where $\overline{1-\lambda} = F^{-1}(1-\lambda)$ is the $(1-\lambda)$ 100-percent population quantile.

Let t_n be a consistent estimator based on k sample quantiles, say, $x_{(\lambda_1)}, \cdots, x_{(\lambda_k)}$, where $0 \equiv \lambda_0 \leqq \lambda \equiv \lambda_1 \leqq \lambda_2, \cdots, \lambda_{k-1} \leqq \lambda_k \equiv 1 - \lambda \leqq \lambda_{k+1} \equiv 1$ and $\lambda \in [0, \frac{1}{2}]$.

THEOREM 3.1. *If the condition*

$$\text{(G)} \qquad \sup_{k\geqq i\geqq 2} |\lambda_i - \lambda_{i-1}| \to 0 \quad \textit{as} \quad k \to \infty$$

is satisfied, then

$$\text{(3.2)} \qquad \lim_{k\to\infty} \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{\varepsilon^2 n} \log \alpha(t_n, \varepsilon, \theta) \geqq -\tfrac{1}{2}\beta(f, \lambda) \geqq -\tfrac{1}{2}I(f) .$$

PROOF. Let $f_n(x_{(\lambda_1)}, \cdots, x_{(\lambda_k)} \mid \theta)$ be the joint density function of the k sample quantiles. It follows from the consistency of sample quantiles that

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \log \frac{f_n(x_{(\lambda_1)}, \cdots, x_{(\lambda_k)} \mid \theta')}{f_n(x_{(\lambda_1)}, \cdots, x_{(\lambda_k)} \mid \theta)} \\ &\quad = R(\theta', \theta) \\ \text{(3.3)} \qquad &\quad = \textstyle\sum_{i=1}^{k+1} (\lambda_i - \lambda_{i-1})\{\log (F(\bar{\lambda}_i + \delta \mid \theta') - F(\bar{\lambda}_{i-1} + \delta \mid \theta')) \\ &\qquad - \log (F(\bar{\lambda}_i + \delta \mid \theta) - F(\bar{\lambda}_{i-1} + \delta \mid \theta))\} , \end{aligned}$$

a.s. $[P_{\theta'}]$, where $\delta = \theta' - \theta$ and $\bar{\lambda}_i = F_\theta^{-1}(\lambda_i)$. For δ sufficiently small, Taylor's expansion of $R(\theta', \theta)$ about θ' in the neighborhood of θ is

$$\text{(3.4)} \qquad R(\theta', \theta) = \frac{\delta^2}{2} \beta(f, \lambda, k) + o(\delta^2) ,$$

where

$$\beta(f, \lambda, k) = \frac{f^2(\bar{\lambda} \mid \theta)}{\lambda} + \frac{f^2(\overline{1-\lambda} \mid \theta)}{\lambda} + \sum_{i=2}^{k} \frac{(f(\bar{\lambda}_i \mid \theta) - f(\bar{\lambda}_{i-1} \mid \theta))^2}{\lambda_i - \lambda_{i-1}} .$$

The result (3.2) follows immediately from Theorem 2.3, condition (G), the mean value theorem, and equation (3.4). □

THEOREM 3.2. *Let $x_{(\frac{1}{2})}$ be the estimator for the location parameter θ; then*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n\varepsilon^2}\log\alpha(x_{(\frac{1}{2})},\varepsilon,\theta) = -2f^2(0)\,. \tag{3.5}$$

PROOF. Since θ is a location parameter, the random variable $y = F(x\,|\,\theta)$ is uniformly distributed in [0, 1] for all $\theta \in \Theta$ and $\alpha(x_{(\frac{1}{2})}, \varepsilon, \theta) = P\{|y_{(\frac{1}{2})} - \frac{1}{2}| > \delta\}$, where $\delta = F(\theta+\varepsilon\,|\,\theta) - F(\theta\,|\,\theta) = F(\theta\,|\,\theta) - F(\theta-\varepsilon\,|\,\theta)$. The result (3.5) follows from Theorem 2.1 and the following Taylor's expansion:

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log\alpha(x_{(\frac{1}{2})},\varepsilon,\theta) &= \log 2 + \tfrac{1}{2}\log F(\theta+\varepsilon\,|\,\theta)(1-F(\theta+\varepsilon\,|\,\theta))\\ &= -2f^2(0)\varepsilon^2 + o(\varepsilon^2)\,.\end{aligned}$$

□

THEOREM 3.3. *The median $x_{(\frac{1}{2})}$ is asymptotically efficient in Bahadur's sense; i.e.,*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n\varepsilon^2}\log\alpha(x_{(\frac{1}{2})},\varepsilon,\theta) = -\tfrac{1}{2}I(f)\,, \tag{3.6}$$

if and only if the underlying distribution is a double exponential distribution with density $f(x) = \frac{1}{2}I^{\frac{1}{2}}(f)\exp\{-I^{\frac{1}{2}}(f)|x|\}$.

PROOF. It follows from the definition of $\beta(f, \lambda)$ that $\beta(f, \lambda)$ is a non-increasing function and $I(f) \geqq \beta(f, \lambda) \geqq 4f^2(0)$ for all $\lambda \in [0, \frac{1}{2}]$. If $I(f) = 4f^2(0)$ then $\beta(f, \lambda) \equiv I(f)$ for all $\lambda \in [0, \frac{1}{2}]$; i.e., for all $x \geqq 0$

$$\frac{f^2(x)}{1-F(x)} + \int_0^x \frac{(f'(\xi))^2}{f(\xi)}\,d\xi \equiv \tfrac{1}{2}I(f)\,.$$

The necessity part follows directly by solving the above integral equation. The sufficiency is an immediate consequence of Theorem 3.2. □

REMARK 1. All the results in this section could be obtained under a different set of regularity conditions (e.g., existence of moment generating function) by using Bahadur's technique.

REMARK 2. Theorem 3.3 may also be proved via the Cauchy–Schwarz inequality.

4. A measure of asymptotic relative efficiency for consistent estimators of the location parameter. Let $t_n^{(1)}$ and $t_n^{(2)}$ be two consistent estimators for θ. We denote $\beta_i = \lim_{\varepsilon\to 0}\lim_{n\to\infty}(\varepsilon^2 n)^{-1}\log\alpha(t_n^{(i)}, \varepsilon, \theta)$ for $i = 1, 2$. For the case in which one of the quantities β_1 or β_2 is nonzero (i.e., at least one of the estimators $t_n^{(1)}$ or $t_n^{(2)}$ converges exponentially to θ), the ratio

$$e_{12} = \frac{\beta_1}{\beta_2}\,,\quad \beta_2 > 0 \tag{4.1}$$

serves as a meaningful measure of the asymptotic relative efficiency of $t_n^{(1)}$ relative to $t_n^{(2)}$ (Bahadur's ARE).

THEOREM 4.1. *Assume that the moment generating function* $\phi(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$ *exists for t in an open interval around zero. Then*

$$\beta(\theta) = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n\varepsilon^2} \log \alpha(\bar{x}_n, \varepsilon, \theta) = -\frac{1}{2\sigma^2}, \tag{4.2}$$

where $\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx$.

PROOF. Since θ is a location parameter we can assume that $\theta = 0$. Let $\phi(t, \varepsilon) = e^{-t\varepsilon}\phi(t)$ and $\rho(\varepsilon) = \inf_{t \geq 0} e^{-\lambda\varepsilon}\phi(t)$. The existence of $\phi(t)$ for t in an open interval around zero implies there exists a unique $\tau(\varepsilon) \doteq (\varepsilon/\sigma^2)(1 + o(1))$ such that $\rho(\varepsilon) = \phi(\tau(\varepsilon), \varepsilon)$. From Bahadur's theorem and the expansion of $\log \phi(\tau, \varepsilon)$, we obtain, for ε sufficiently small:

$$\begin{aligned} \lim_{n \to \infty} \frac{1}{n} \log \alpha(\bar{x}_n, \varepsilon, 0) &= \log \rho(\varepsilon) = \log \phi(\tau, \varepsilon) \\ &= k_1(\varepsilon)\tau(\varepsilon) + k_2(\varepsilon)\frac{\tau^2(\varepsilon)}{2} + o(\tau^2(\varepsilon)), \end{aligned} \tag{4.3}$$

where $k_1(\varepsilon) = -\varepsilon$ and $k_2(\varepsilon) = \sigma^2 + \varepsilon^2$ are the first two cumulants of $x - \varepsilon$. This completes the proof.

THEOREM 4.2. *The asymptotic efficiency of* $\bar{x}_n$ *relative to the median* $x_{(\frac{1}{2})}$ *is*

$$e_{\bar{x}_n, x_{(\frac{1}{2})}}(\theta) = (4\sigma^2 f^2(0))^{-1}. \tag{4.4}$$

PROOF. This result is a direct consequence of Theorem 3.2 and Theorem 4.1.

We recognize the result (4.4) as the classical measure of asymptotic relative efficiency of $\bar{x}_n$ with respect to $x_{(\frac{1}{2})}$ based on their asymptotic variances. It shows, in this case, that the classical Pitman measure of asymptotic relative efficiency and the ARE defined by (4.1) coincide.

Acknowledgment. The author wishes to thank Professor Leon J. Gleser and Professor D. A. S. Fraser for their encouragement, comments, and financial support. He is also grateful to the referee for his detailed comments and for pointing out that the results in Section 3 may be obtained by using Bahadur's technique, and that Theorem 3.3 can also be proved via Cauchy-Schwarz inequality.

REFERENCES

[1] BAHADUR, R. R. (1960). Asymptotic efficiency of tests and estimates. *Sankhyā*, **22** 229–252.

[2] BAHADUR, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 304–324.

[3] BAHADUR, R. R. (1971). Some limit theorems in statistics. *Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia.

[4] FU, J. C. (1973). On a theorem of Bahadur on the rate of convergence of point estimators. *Ann. Statist.* **1** 745–749.

[5] KULLBACK, S. (1968). *Information Theory and Statistics*. Dover Publications, New York.
[6] SANOV, I. N. (1957). On the probability of large deviations of random variables. *Select. Transl. Math. Statist. Prob.* **1** 213–244.

DEPARTMENT OF STATISTICS
UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA
CANADA R3T 2N2