ON FISHER'S BOUND FOR ASYMPTOTIC VARIANCES[1]

By R. R. Bahadur

University of Chicago

1. Introduction and summary. For each n let t_n be an estimate (based on n independent and identically distributed observations) of a real valued parameter θ. Suppose that, for each θ, $n^{\frac{1}{2}}(t_n - \theta)$ is asymptotically normally distributed with mean zero and variance $v(\theta)$. According to R. A. Fisher we then have

$$v(\theta) \geqq I^{-1}(\theta), \tag{1}$$

where I is the information contained in a single observation. It is known however that, in the absence of regularity conditions on the sequence $\{t_n\}$, (1) does not necessarily hold for each θ. On the other hand, according to LeCam (1952, 1953, 1958) the set of points θ for which (1) fails is always of Lebesgue measure zero. This note gives a simple proof of the stated result of LeCam, along the following lines. First a sufficient condition for the validity of (1) at a given value of θ, say θ^0, is obtained. This is a little weaker than the condition that t_n be asymptotically median-unbiased (i.e. $P(t_n < \theta \mid \theta) \to \frac{1}{2}$ as $n \to \infty$) uniformly for θ in some neighborhood of θ^0. It is then shown that the sufficient condition is automatically satisfied at almost all θ^0.

The main propositions are stated in the following paragraphs of this section, and the proofs are given in Section 2. The proofs depend on the Neyman-Pearson lemma concerning the optimality of the likelihood ratio test of a simple hypothesis against a simple alternative. This lemma is made available in the present context by means of the well known considerations that an estimate of θ can provide a test of the value of θ, and that the quality of the resulting test is heavily dependent on the quality of the estimate. A similar application of the Neyman-Pearson lemma to estimation theory is made in Bahadur (1960). It is shown there that if instead of asymptotic variances one considers quantities called asymptotic effective variances, Fisher's bound becomes valid for all θ.

Now let $X = \{x\}$ be a sample space of points x, $\mathcal{B}$ a σ-field of sets of X, and $\{P(\cdot \mid \theta) : \theta \varepsilon \Theta\}$ a set of probability measures $P(\cdot \mid \theta)$ on $\mathcal{B}$, where θ is a real parameter and Θ is an open interval on the real line. It is assumed that the following conditions (i)–(iv) are satisfied.

(i) There exists a σ-finite measure on $\mathcal{B}$, say μ, such that, for each θ, $P(\cdot \mid \theta)$ admits a probability density with respect to μ, $f(x \mid \theta)$ say, i.e.,

$$P(B \mid \theta) = \int_B f(x \mid \theta)\, d\mu \qquad \text{for all } B \varepsilon \mathcal{B}, \theta \varepsilon \Theta. \tag{2}$$

Received 26 June 1962; revised 9 June 1964.

[1] This research was supported by Research Grant No. NSF GP 609 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation, and by the Statistics Branch, Office of Naval Research.

(ii) For each $x \varepsilon X$,

(3) $$L(\theta \mid x) = \log f(x \mid \theta)$$

is a twice-differentiable function of θ and the second derivative is continuous in θ.

(iii) With dashes on L denoting partial differentiation with respect to θ,

(4) $$0 < E(\{L'(\theta \mid x)\}^2 \mid \theta) \equiv I(\theta) < \infty, \qquad E(L'(\theta \mid x) \mid \theta) = 0,$$

and

(5) $$E(L''(\theta \mid x) \mid \theta) = -I(\theta)$$

for every θ.

(iv) For any given θ^0 in Θ, there exists a $\delta > 0$ and a $\mathcal{B}$-measurable function $M(x)$ such that $|L''(\theta \mid x)| \leqq M(x)$ for all $x \varepsilon X$ and all $\theta \varepsilon (\theta^0 - \delta, \theta^0 + \delta)$, and such that $E(M(x) \mid \theta^0) < \infty$; δ and M are, of course, allowed to depend on the given θ^0.

The above conditions are a simplification, along the lines of LeCam (1953), of the conditions formulated by Cramér (1946) for his analysis of the likelihood equation. It may be added that the present regularity conditions are in a sense weaker than those of LeCam (1953, 1958), since the method of proof of the latter papers requires local conditions such as (ii)–(iv) and also the existence and consistency of maximum likelihood estimates based on independent observations on x.

Let $x_1, x_2, \cdots$ denote a sequence of independent and identically distributed observations on x. For each $n = 1, 2, \cdots$ write $x^{(n)} = (x_1, \cdots, x_n)$. Let $X^{(n)}$ denote the sample space of $x^{(n)}$, and $\mathcal{B}^{(n)}$ the σ-field of sets of $X^{(n)}$ which is determined by the given $\mathcal{B}$ in the usual way. For any measure Q on $\mathcal{B}$, the corresponding product measure on $\mathcal{B}^{(n)}$ will be denoted by $Q^{(n)}$. For simplicity, $Q^{(n)}$ is abbreviated to Q in cases where the domain of $Q^{(n)}$ is plain from the context.

Now let there be given a sequence $\{t_n\}$ such that t_n is a $\mathcal{B}^{(n)}$-measurable function on $X^{(n)}$ into Θ $(n = 1, 2, \cdots)$. It is assumed that for each θ in Θ there exists a positive constant $v(\theta)$ such that, as $n \to \infty$, $n^{\frac{1}{2}}(t_n - \theta)$ is asymptotically normally distributed with mean 0 and variance $v(\theta)$ when θ obtains. (For a treatment of the case when the present assumption $0 < v < \infty$ is weakened to $0 \leqq v < \infty$, cf. the last paragraph of Section 3.) The given sequence $\{t_n\}$ will remain fixed throughout. We note that

(6) $$\lim_{n\to\infty} P(t_n < \theta \mid \theta) = \tfrac{1}{2}$$

for each θ in Θ.

PROPOSITION 1. *If θ^0 is a point in Θ, and if*

(7) $$\liminf_{n\to\infty} P(t_n < \theta^0 + n^{-\frac{1}{2}} \mid \theta^0 + n^{-\frac{1}{2}}) \leqq \tfrac{1}{2},$$

then (1) *holds for* $\theta = \theta^0$.

It follows from Proposition 1 by symmetry that if

(8) $$\liminf_{n\to\infty} P(t_n > \theta^0 - n^{-\frac{1}{2}} \mid \theta^0 - n^{-\frac{1}{2}}) \leqq \tfrac{1}{2},$$

then also (1) holds for $\theta = \theta^0$.

Another consequence of Proposition 1 is that if (6) holds uniformly for θ in some open interval of Θ then (1) holds for each θ in that interval. A somewhat weaker conclusion concerning the sufficiency of uniform convergence for (1) has been obtained independently by Rao (1963).

The sequence $\{t_n\}$ is said to be superefficient if $v(\theta) \leqq I^{-1}(\theta)$ for all θ and the inequality is strict for at least one θ. Examples of superefficient estimates were discovered by J. L. Hodges, Jr. (cf. LeCam (1953)). General studies bearing on superefficiency, using methods different from the present ones, were carried out by LeCam (1953, 1958). An informal discussion along lines similar to those of LeCam was given independently by Wolfowitz (1953). It is shown in LeCam (1953) that if $\{t_n\}$ is superefficient then $v(\theta) = I^{-1}(\theta)$ for almost all θ in Θ; the following more general conclusion is given in LeCam (1958):

PROPOSITION 2. *The set of all θ in Θ for which* (1) *does not hold is of Lebesgue measure zero.*

It was observed by Chernoff (1956) that the asymptotic variance of an estimate is always a lower bound to the asymptotic expected squared error; in view of Proposition 2, this observation yields:

PROPOSITION 3. $\liminf_{n\to\infty} \{nE[(t_n - \theta)^2 \mid \theta]\} \geqq I^{-1}(\theta)$ *for almost all θ in Θ.*

The conclusions stated in this section can be extended to the case when θ is a p dimensional parameter; a brief account of these extensions is given in Section 3. An extension to sampling frameworks more general than the present one of independent and identically distributed observations is described in Section 4.

2. Proofs. For any θ^0 in Θ and any n let

$$\theta_n^0 = \theta^0 + n^{-\frac{1}{2}}. \tag{9}$$

It is convenient to begin with some computations concerning the likelihood ratio test of θ^0 against θ_n^0 based on $x^{(n)}$. For any $\theta \varepsilon \Theta$, n, and $x^{(n)} \varepsilon X^{(n)}$, let

$$L_n(\theta \mid x^{(n)}) = \sum_{i=1}^{n} L(\theta \mid x_i). \tag{10}$$

Now choose and fix a particular θ^0, and let

$$K_n = K_n(x^{(n)}, \theta^0) = [L_n(\theta_n^0 \mid x^{(n)}) - L_n(\theta^0 \mid x^{(n)}) + \tfrac{1}{2}I]/I^{\frac{1}{2}}, \tag{11}$$

where $I = I(\theta^0)$, and θ_n^0 is given by (9). It should be noted that K_n is a strictly increasing function of the likelihood ratio statistic for testing θ^0 against θ_n^0 when the datum is $x^{(n)}$. Let

$$H_n(z) = P(K_n < z \mid \theta^0). \tag{12}$$

Let Φ denote the standard normal distribution function.

LEMMA 1. *As $n \to \infty$, $H_n(z) \to \Phi(z)$ for each z.*

PROOF. It follows from (9) by Taylor's theorem that

$$L_n(\theta_n^0 \mid x^{(n)}) = L_n(\theta^0 \mid x^{(n)}) + n^{-\frac{1}{2}}L_n'(\theta^0 \mid x^{(n)}) + \tfrac{1}{2}n^{-1}L_n''(\theta_n^* \mid x^{(n)}), \tag{13}$$

where $\theta^0 < \theta_n^*(x^{(n)}) < \theta_n^0$. Let us put $n^{-1}\,|L_n''(\theta_n^* \mid x^{(n)}) - L_n''(\theta^0 \mid x^{(n)})| = \xi_n$.

It is shown in the following paragraph that $\xi_n \to 0$ with probability one, and it is plain from (5) and (10) that $n^{-1}L_n''(\theta^0 \mid x^{(n)}) \to -I$ with probability one. Consequently, it follows from (11) and (13) that $K_n = (nI)^{-\frac{1}{2}}L_n'(\theta^0 \mid x^{(n)}) + \eta_n$, where $\eta_n \to 0$ with probability one. The desired conclusion now follows from (4) and (10) by the central limit theorem.

To show that $\xi_n \to 0$, for any $\delta > 0$ let $A(x, \delta) = \sup \{|L''(\theta \mid x) - L''(\theta^0 \mid x)|: \theta \varepsilon \Theta, |\theta - \theta^0| < \delta\}$ and let $m(\delta)$ be the expected value of A. It follows from conditions (ii) and (iv) of Section 1 that $A \downarrow 0$ as $\delta \downarrow 0$ for each x, and that $m \downarrow 0$ as $\delta \downarrow 0$. Given $\epsilon > 0$, choose a $\delta > 0$ such that $m(\delta) \leqq \epsilon$. Then, for each $n > \delta^{-1}$ and all $x^{(n)}$, we have $|\xi_n| \leqq n^{-1} \sum_{i=1}^{n} A(x_i, n^{-1}) \leqq n^{-1} \sum_1^n A(x_i, \delta)$, by (10) and $|\theta_n^* - \theta^0| < n^{-1}$. Hence $\lim \sup \xi_n \leqq m(\delta) \leqq \epsilon$ with probability one. Since ϵ is arbitrary, this completes the proof.

Let k be a constant, and for each n let

$$C_n = \{x^{(n)}: K_n \geqq k\}, \tag{14}$$

where K_n is given by (11).

LEMMA 2. *$P(C_n \mid \theta^0) = 1 - \Phi(k) + o(1)$ and $P(C_n \mid \theta_n^0) = 1 - \Phi(k - I^{\frac{1}{2}}) + o(1)$ as $n \to \infty$.*

PROOF. The first part of Lemma 2 is plain from (14) and Lemma 1. To establish the second part, we note that

$$\begin{aligned} 1 - P(C_n \mid \theta_n^0) &= P(K_n < k \mid \theta_n^0) \\ &= \int_{K_n<k} \exp\,[L_n(\theta_n^0 \mid x^{(n)})]\cdot d\mu^{(n)} \\ &= \int_{K_n<k} \exp\,[L_n(\theta_n^0 \mid x^{(n)}) - L_n(\theta^0 \mid x^{(n)})]\cdot dP^{(n)}(\cdot \mid \theta^0) \\ &= e^{-I/2} \int_{K_n<k} \exp\,(I^{\frac{1}{2}} K_n)\cdot dP^{(n)}(\cdot \mid \theta^0) && \text{by (11)} \\ &= e^{-I/2} \int_{-\infty<z<k} \exp\,(I^{\frac{1}{2}} z)\cdot dH_n && \text{by (12).} \end{aligned} \tag{15}$$

It follows easily from Lemma 1 that

$$\begin{aligned} \int_{-\infty<z<k} \exp\,(I^{\frac{1}{2}} z)\, dH_n &= \int_{-\infty<z<k} \exp\,(I^{\frac{1}{2}} z)\, d\Phi + o(1) \\ &= e^{I/2}\,\Phi(k - I^{\frac{1}{2}}) + o(1). \end{aligned} \tag{16}$$

The conclusion required follows from (15) and (16).

The referee has pointed out that the preceding proof exemplifies a computation carried out previously by LeCam (1960), and that the second part of Lemma 2 is deducible from Theorem 2.1 of the latter paper.

PROOF OF PROPOSITION 1. For each n let

$$D_n = \{x^{(n)}: t_n \geqq \theta_n^0\}. \tag{17}$$

Choose and fix a constant $k > I^{\frac{1}{2}}$, and let C_n be defined by (14). Then $P(C_n \mid \theta_n^0)$ tends to a limit $<\frac{1}{2}$, by Lemma 2, while $\limsup_{n\to\infty} P(D_n \mid \theta_n^0) > \frac{1}{2}$ by (7) and (17). Hence there exists a sequence of positive integers, say $m_1 < m_2 < \cdots$ such that

$$P(D_n \mid \theta_n^0) > P(C_n \mid \theta_n^0) \tag{18}$$

for $n = m_r$ ($r = 1, 2, \cdots$). For each n regard C_n and D_n as alternative critical regions for testing θ^0 against θ_n^0. Since C_n is an optimum critical region, (18) implies

$$P(D_n \mid \theta^0) > P(C_n \mid \theta^0). \tag{19}$$

Hence (19) holds for $n = m_r$ ($r = 1, 2, \cdots$). It follows from (9) and (17) that, with $v = v(\theta^0)$, $P(D_n \mid \theta^0) = 1 - \Phi(v^{-\frac{1}{2}}) + o(1)$ as $n \to \infty$. Consequently, by letting $n \to \infty$ through the sequence $\{m_r\}$ in (19), it follows from Lemma 2 that

$$1 - \Phi(v^{-\frac{1}{2}}) \geqq 1 - \Phi(k). \tag{20}$$

Hence $v^{-\frac{1}{2}} \leqq k$. Since $k > I^{\frac{1}{2}}$ is arbitrary, we conclude that (1) holds at θ^0.

Now regard θ as a real variable and, for any n and any θ, let

$$\begin{aligned} f_n(\theta) &= |P(t_n < \theta \mid \theta) - \tfrac{1}{2}| \quad \text{if} \quad \theta \varepsilon \Theta \\ &= 0 \qquad\qquad\qquad\qquad\quad\; \text{if} \quad \theta \not\varepsilon \Theta. \end{aligned} \tag{21}$$

LEMMA 3. *For each n, f_n is a Borel-measurable function of θ.*

PROOF. For given $x^{(n)}$, the likelihood function $\exp L_n(\theta \mid x^{(n)})$ is continuous in θ. It follows hence from the theorem of Scheffé (1947) that the set of discontinuities of $P(t_n < \theta \mid \theta)$ is a countable subset of Θ. Hence the set of discontinuities of f_n is countable; hence f_n is Borel-measurable. The details of this proof are omitted.

It follows from (6) and (21) that

$$0 \leqq f_n(\theta) \leqq \tfrac{1}{2}, \qquad \lim_{n\to\infty} f_n(\theta) = 0 \tag{22}$$

for each θ. Let

$$g_n(\theta) = f_n(\theta + n^{-\frac{1}{2}}), \qquad 0 \leqq g_n \leqq \tfrac{1}{2}. \tag{23}$$

LEMMA 4. *There exists a set N of Lebesgue measure zero, and a sequence $m_1 < m_2 < \cdots$ of positive integers m_r, such that $\lim_{r\to\infty} g_{m_r}(\theta) = 0$ for each $\theta \not\varepsilon N$.*

PROOF.

$$\begin{aligned} \int_{-\infty}^{+\infty} g_n(\theta)\, d\Phi(\theta) &= \int_{-\infty}^{+\infty} f_n(\theta + n^{-\frac{1}{2}})\, d\Phi(\theta) \\ &= \int_{-\infty}^{+\infty} f_n(\theta) \cdot \exp\left(-\frac{1}{2n} + n^{-\frac{1}{2}}\theta\right) \cdot d\Phi(\theta) = o(1) \end{aligned}$$

as $n \to \infty$, by (22), (23) and Lebesgue's dominated convergence theorem. Hence $g_n \to 0$ in Φ-measure. Hence there exists a sequence $\{m_r\}$ such that, except for a

Φ-null set, $g_{m_r} \to 0$ as $r \to \infty$. This completes the proof since Φ-measure is equivalent to Lebesgue measure.

It would be interesting to know whether, in Lemma 4, one can always take $m_r = r$ for each r.

PROOF OF PROPOSITION 2. Lemma 4 implies that $\lim \inf_{n\to\infty} g_n(\theta) = 0$ almost everywhere; it now follows by referring to (21) and (23) that (7) holds for almost all θ^0, and Proposition 2 follows from Proposition 1.

PROOF OF PROPOSITION 3. For each n and θ,

$$\begin{aligned} nE\{(t_n - \theta)^2 \mid \theta\} &= n \int_0^\infty P\{(t_n - \theta)^2 > y \mid \theta\}\, dy \\ &= \int_0^\infty P\{n(t_n - \theta)^2 > z \mid \theta\}\, dz. \end{aligned} \tag{24}$$

It follows from (24) by Fatou's lemma that

$$\lim \inf_{n\to\infty} nE\{(t_n - \theta)^2 \mid \theta\} \geqq \int_0^\infty 2\Phi(-[z/v(\theta)]^{\frac{1}{2}})\, dz = v(\theta) \tag{25}$$

for each θ; now Proposition 2 applies.

3. The multiparameter case. Consider the previous framework $X = \{x\}$, $\mathcal{B}$, $\{P(\cdot \mid \theta) : \theta \varepsilon \Theta\}$, $dP(\cdot \mid \theta) = f(x \mid \theta)\, d\mu$ concerning a single observation x, but suppose now that θ is a p dimensional parameter, say $\theta = (\phi_1, \cdots, \phi_p)$ and that Θ is an open subset of p dimensional Euclidean space. The following regularity conditions are assumed to hold. For each x in X, all second-order partial derivatives of $L = \log f$ with respect to the coordinates of θ exist and are continuous. With $L_i = \partial L/\partial\phi_i$ and $L_{ij} = \partial^2 L/\partial\phi_i\partial\phi_j$, $L_i(\theta \mid x)$ is square integrable when θ obtains, and $E(L_i) = 0$, $E(L_iL_j) = -E(L_{ij}) = I_{ij}(\theta)$ say $(i, j = 1, \cdots, p)$. $\{I_{ij}(\theta)\} = I(\theta)$ is a positive definite matrix. For any given θ^0 in Θ, there exists a neighborhood of θ^0 and a function $M(x)$ with $E(M \mid \theta^0) < \infty$ such that $|L_{ij}(\theta \mid x)| \leqq M(x)$ for all x in X, θ in the neighborhood, and $i, j = 1, 2, \cdots, p$.

For each $n = 1, 2, \cdots$ let $t_n = (u_{1n}, \cdots, u_{pn})$ be a function on $X^{(n)}$ into Θ. Suppose that, for each θ in Θ, there exists a $p \times p$ symmetric positive semi-definite matrix $v(\theta) = \{v_{ij}(\theta)\}$ such that $n^{\frac{1}{2}}(t_n - \theta)$ tends in distribution to the p-variate normal distribution with mean vector zero and covariance matrix $v(\theta)$ when θ obtains. It can then be shown that there exists a subset of Θ of Lebesgue measure zero, say N, such that

$$v(\theta) - I^{-1}(\theta) \text{ is positive semi-definite} \tag{26}$$

for all θ in $\Theta - N$. It follows that v is necessarily positive definite for almost all θ. It also follows that if the components of v and I are continuous functions of θ then (26) holds for all θ in Θ.

The following is a well known consequence of (26). Suppose it is required to estimate $m(\theta)$ where m is a real-valued function of θ. If we take $s_n = m(t_n)$ as the estimate of m, and if m is a sufficiently smooth function of θ, then

$n^{\frac{1}{2}}(s_n - m(\theta))$ is asymptotically normally distributed with zero mean and variance $\sigma^2(\theta) = \sum_{i,j=1}^{p} m_i(\theta) \cdot m_j(\theta) \cdot v_{ij}(\theta)$ when θ obtains, where $m_i = \partial m / \partial \varphi_i$. Hence $\sigma^2 \geqq \sum m_i m_j I^{ij}$ for every θ for which (26) holds.

That (26) holds for almost all θ can be shown as follows. Suppose first that v is positive definite for each θ. Consider a fixed θ^0 in Θ. Let $a = (a_1, \cdots, a_p)$ and $b = (b_1, \cdots, b_p)$ be arbitrary non-zero vectors. For each n, let

$$\theta_n^0 = \theta^0 + n^{-\frac{1}{2}} a, \tag{27}$$

and let K_n be defined by (11) with I replaced by aIa', where a' denotes the transpose of a. Then Lemma 1 of Section 2 holds. Consequently, with C_n defined by (14), Lemma 2 also holds, with I again replaced by aIa'. Next, for each n let

$$D_n = \{x^{(n)} : t_n b' \geqq \theta_n^0 b'\}, \tag{28}$$

and assume that

$$\limsup_{n \to \infty} P(D_n \mid \theta_n^0) \geqq \tfrac{1}{2}. \tag{29}$$

Then with $k > +(aIa')^{\frac{1}{2}}$ in the present definition of C_n, (18) holds for infinitely many n; hence (19) holds for infinitely many n. Since $P(D_n \mid \theta^0) = P(n^{\frac{1}{2}}(t_n - \theta^0)b' \geqq ab' \mid \theta^0)$ by (27) and (28), it follows as in Section 2 that $k \geqq ab'(bv(\theta^0)b')^{-\frac{1}{2}}$; hence

$$ab'(bv(\theta^0)b')^{-\frac{1}{2}} \leqq (aI(\theta^0)a')^{\frac{1}{2}}. \tag{30}$$

We observe next that $\lim_{n \to \infty} P(t_n b' \geqq \theta b' \mid \theta) = \frac{1}{2}$ for each θ in Θ. It follows hence (cf. Section 2) that (29) is satisfied for all θ^0 in $\Theta - N^0$, where N^0 is a set of Lebesgue measure zero. (The set N^0 depends, possibly, on a and b.) Hence (30) holds for almost all θ^0 in Θ.

Let N be the set of all θ^0 in Θ such that (30) fails for some pair a and b such that all co-ordinates of a and b are rational. Then N is of Lebesgue measure zero, and for each θ in $\Theta - N$, (30) holds for all non-zero a and b. The validity of (30) for all non-zero a and b is equivalent to (26).

Now consider the case when v is not necessarily positive definite for each θ. For each n, let $t_n^* = t_n + \epsilon n^{-\frac{1}{2}} z$, where $\epsilon \neq 0$ is a constant, and z is normally distributed independent of the x_i, with $E(z) = 0$ and with $E(z'z) = w$ (say) positive definite and independent of θ and of ϵ. Then $n^{\frac{1}{2}}(t_n^* - \theta)$ is asymptotically normally distributed with mean zero and covariance $v^*(\theta) = v(\theta) + \epsilon^2 w$ when θ obtains. Since v^* is positive definite for each θ, the preceding argument (or, rather, a trivial extension thereof) shows that $v(\theta) + \epsilon^2 w - I^{-1}(\theta)$ is positive semi-definite for almost all θ. Since $\epsilon \neq 0$ is arbitrary, it follows easily that in fact (26) holds for almost all θ.

4. Another extension. Professor L. LeCam has pointed out to the writer that the arguments and conclusions of this paper continue to hold in sampling frameworks much more general than the one treated in the preceding sections. It will suffice to consider here the case when θ is a real parameter taking values in an open interval Θ. For each n let $X^{(n)}$ be a space of points $x^{(n)}$ and $\mathcal{B}^{(n)}$ a σ-field of sets of $X^{(n)}$. For each n and each θ in Θ let $P^{(n)}(\cdot \mid \theta)$ be a probability measure on $\mathcal{B}^{(n)}$. Suppose that the following conditions are satisfied.

(a) For each n, there exists a σ-finite measure $\mu^{(n)}$ on $\mathcal{B}^{(n)}$ such that, for each θ in Θ, $P^{(n)}(\cdot \mid \theta)$ is absolutely continuous with respect to $\mu^{(n)}$, say $dP^{(n)}(\cdot \mid \theta) = \exp(L_n(\theta \mid x^{(n)}))\, d\mu^{(n)}$.

(b) For each θ^0 in Θ there exists a positive constant $I(\theta^0)$ such that with $\theta_n^0 = \theta^0 + n^{-\frac{1}{2}}$ and K_n defined by (11) (with the present definition of L_n), K_n tends in distribution to a standard normal variable as $n \to \infty$ when θ^0 obtains.

(c) For each n, at least one of the following two conditions is satisfied: $L_n(\theta \mid x^{(n)})$ is continuous in θ for each $x^{(n)}$, or L_n is an $\mathfrak{M} \times \mathcal{B}^{(n)}$-measurable function of $(\theta, x^{(n)})$ where $\mathfrak{M}$ is the class of Borel measurable sets of Θ.

Now, the argument of Section 2 uses the special structure of the framework described in Section 1 only in the proof of Lemma 1. Since Lemma 1 is valid by assumption in the more general framework being treated here, it follows that the arguments of Section 2 hold verbatim for any sequence $\{t_n\}$ of $\mathcal{B}^{(n)}$-measurable functions on $X^{(n)}$ into Θ such that $n^{\frac{1}{2}}(t_n - \theta)$ is asymptotically normally distributed with mean 0 and variance $v(\theta)$. (The first condition of assumption (c) ensures that the proof of Lemma 3 outlined in Section 2 goes through in the present case; the less drastic second condition of assumption (c), if satisfied, would ensure the measurability of g_n by means of Fubini's theorem instead.) It is thus seen that the propositions of Section 1 hold in the present case.

It seems that assumptions such as (b) play a central role in various asymptotic studies, and general sufficient conditions for the validity of such assumptions have been given by LeCam (1960). As may be seen from the paper just cited, it is not necessary to suppose that $x^{(n)} = (x_1, \cdots, x_n)$ where for each θ the x_i have a common sample space and are independently and identically distributed therein, the common distribution being independent of n. It may be added that even if this last is the case under consideration, the regularity conditions (ii)–(iv) of Section 1 can be replaced by others which also ensure that assumption (b) is satisfied.

REFERENCES

Bahadur, R. R. (1960). On the asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–252.

Chernoff, H. (1956). Large sample theory: parametric case. *Ann. Math. Statist.* **27** 1–22.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press.

LeCam, L. (1952). On sets of parameter points where it is possible to achieve superefficiency of estimates (abstract). *Ann. Math. Statist.* **23** 148.

LeCam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. *Univ. California Publ. Statist.* **1** 277–330.

LeCam, L. (1958). Les propriétés asymptotiques des solutions des Bayes. *Publ. Inst. Statist. Univ. Paris* **7** fascicule 3–4, 17–35.

LeCam, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.* **3** 37–98.

Rao, C. R. (1963). Criteria of estimation in large samples. *Sankhyā Ser. A* **25** 189–206.

Scheffé, H. (1947). A useful convergence theorem. *Ann. Math. Statist.* **18** 434–438.

Wolfowitz, J. (1953). The method of maximum likelihood and the Wald theory of decision functions. *Indag. Math.* **15** 114–119.