A Revisit to Le Cam's First Lemma

G. Jogesh Babu & Bing Li

Sankhya A
The Indian Journal of Statistics - Official Journal of Indian Statistical Institute

ISSN 0976-836X

Sankhya A
DOI 10.1007/s13171-020-00223-2

ONLINE FIRST

Springer

Sankhyā A : The Indian Journal of Statistics
https://doi.org/10.1007/s13171-020-00223-2

A Revisit to Le Cam's First Lemma

G. Jogesh Babu and Bing Li
The Pennsylvania State University, University Park USA

Abstract

Le Cam's first lemma is of fundamental importance to modern theory of statistical inference: it is a key result in the foundation of the Convolution Theorem, which implies a very general form of the optimality of the maximum likelihood estimate and any statistic that is asymptotically equivalent to it. This lemma is also important for developing asymptotically efficient tests. In this note we give a relatively simple but detailed proof of Le Cam's first lemma. Our proof allows us to grasp the central idea by making analogies between contiguity and absolute continuity, and is particularly attractive when teaching this lemma in a classroom setting.

AMS (2000) *subject classification.* Primary: 62F12; Secondary: 62E20, 62F10.

1 Introduction

Le Cam's first lemma (Le Cam, 1960; van der Vaart, 1998) plays a critical role in the asymptotic theory of statistical inference. Along with Le Cam's third lemma, it laid the foundation of the Convolution Theorem (Le Cam 1953, 1960; Hájek 1970), which states that every regular estimate can be decomposed into two asymptotically independent pieces, one of which is asymptotically equivalent to the maximum likelihood estimate. A direct consequence of this theorem is that the maximum likelihood estimate, or any estimate asymptotically equivalent to it, has an asymptotic variance smaller than or equal to that of any regular estimate in terms of Louwner's ordering. The publication of this rigorous statement about the optimality of the maximum likelihood estimate resolved some long standing questions about the maximum likelihood estimate ever since its introduction (Fisher 1922, 1925), that is, first, to what extent is the maximum likelihood estimate optimal? and second, beyond the extent to which the maximum likelihood estimate is optimal, how should we interpret the meaning of those estimates whose asymptotic variances are smaller than that of the maximum likelihood estimate? Inevitably, any rigorous course on statistical inference should give adequate answers to these fundamental questions. And, to answer them, it is helpful to have a relatively easy and intuitive proof of Le Cam's first lemma.

Le Cam's first lemma is a set of equivalent conditions for contiguity, which is an asymptotic analogue of absolute continuity. Commonly known proofs of Le Cam's first lemma utilize intricate construction of intermediate functions, and it is difficulty to see the central issue through this construction. van der Vaart (1998) provides a very concise proof of Le Cam's first lemma. What is presented in this note is not largely different from that given in van der Vaart (1998), but a relatively simple and transparent version that we hope can make the proof more easily accessible to graduate students. Especially, this simple proof allows us to grasp the core issue by making direct analogies between sequences of probability measures and individual probability measures. A somewhat different proof, under a stronger assumption, was given in a recent text book on statistical inference by Li and Babu (2019).

Professor C. R. Rao used to start his lectures/presentations with simple basic ideas and then bring in analogies from related problems. This paper is inline with that premise of his.

2 Contiguity

Contiguity is a relation between two sequences of probability measures that resemble the relation of absolute continuity between two probability measures. Recall that a probability measure P is absolutely continuous with respect to a probability measure Q if and only if, for any measurable set A, $Q(A) = 0$ implies $P(A) = 0$. This is symbolically written as $P \ll Q$.

Similarly, if $(\Omega_n, \mathcal{F}_n)$, $n = 1, 2, \ldots$, is a sequence of measurable spaces and $\{P_n\}$ and $\{Q_n\}$ are sequences of probability measures with P_n, Q_n defined on $\mathcal{F}_n$ for each n, then we say that the sequence $\{P_n\}$ is contiguous with respect to the sequence $\{Q_n\}$ if and only if, for each sequence of subsets $\{A_n\}$ with $A_n \in \mathcal{F}_n$, $Q_n(A_n) \to 0$ implies $P_n(A_n) \to 0$. If $\{P_n\}$ is contiguous with respect to $\{Q_n\}$, then we write $P_n \lhd Q_n$.

Despite their conceptual similarity, contiguity and absolute continuity are technically very different: the former characterizes the collective behavior of a pair of sequences of probability measures; the latter that of a pair of probability measures. In particular, even if $P_n \ll Q_n$ holds for every n, this does not imply $P_n \lhd Q_n$. For example, if Z is a standard normal random variable and, for each n, P_n, Q_n are probability distributions of

Z and $1 + (Z/n)$ respectively, then $P_n \ll Q_n$ for each n. However, since the limiting distribution of Q_n is a point mass at 1, if we let A_n be the set $(-\infty, 0)$, then $Q_n(A_n) \to 0$ but $P_n(A_n) = \frac{1}{2}$ for all n. Thus P_n is not contiguous with respect to Q_n.

3 The Main Result

Le Cam's first lemma focuses on necessary and sufficient conditions for contiguity between two sequences of probability measures. Again, it is helpful to make an analogy with the situation where two probability measures are involved. For two probability measures ν and τ, if $\nu \ll \tau$, and if $\frac{d\nu}{d\tau}$ denotes the Radon-Nikodym derivative (Billingsley 1995, Theorem 32.2) then

$$E_\tau\left(\frac{d\nu}{d\tau}\right) = \int \frac{d\nu}{d\tau}\, d\tau = \int d\nu = 1, \tag{3.1}$$

where E_τ denotes the expectation with respect to the probability measure τ. Le Cam's first Lemma is similar to this result when ν and τ are replaced by sequences of probability measures $\{P_n\}$ and $\{Q_n\}$ and absolute continuity $\nu \ll \tau$ is replaced by contiguity $P_n \lhd Q_n$.

We first introduce some notation. Let $X, X_n,\ n = 1, 2, \ldots$ be random vectors in $\mathbb{R}^k$, ν_n the distribution of X_n, and ν the distribution of the random vector X. We use the notation

$$X_n \xrightarrow[\nu_n]{\mathcal{D}} X,$$

to denote weak convergence of X_n to X under the sequence $\{\nu_n\}$; that is, for every bounded and continuous function f on $\mathbb{R}^k$, $E_{\nu_n}(f(X_n)) \to E_\nu(f(X))$. Similarly, if X_n converges in ν_n-probability to a constant a; that is,

$$\nu_n(\|X_n - a\| > \epsilon) \to 0,$$

for every $\epsilon > 0$, then we write $X_n \xrightarrow{\nu_n} a$.

For any two probability measures ν, τ, we have $\nu = \nu^{ac} + \nu^s$ by Lebesgue decomposition (see Billingsley 1995, equation (32.8)) where $\nu^{ac} \ll \tau$, and ν^s and τ are mutually singular (*i.e.*, $\tau(S^c) = 1 = \nu^s(S)$ for some measurable set S). In addition, by Radon-Nikodym Theorem (see Billingsley 1995, Theorem 32.2) $\nu^{ac}(A) = \int_A \frac{d\nu^{ac}}{d\tau}\, d\tau$ for all measurable sets A. For ease of notation we write $\frac{d\nu}{d\tau}$ for Radon-Nikodym derivative $\frac{d\nu^{ac}}{d\tau}$. Analogous to Eq. 3.1, we have for any measurable set B,

$$E_\tau\left(\frac{d\nu}{d\tau} I_B\right) = \int_B \frac{d\nu}{d\tau}\, d\tau = \nu^{ac}(B) \le \nu(B) \le 1. \tag{3.2}$$

The statements (a) and (c) of the theorem given below are comparable to absolute continuity and to Eq. 3.1. For completeness some results on probability measures and convergence of random variables that are needed in the proof of the Theorem are collected following the proof.

Theorem (Le Cam's first lemma)**.** *Let $\{P_n\}$ and $\{Q_n\}$ be sequences of probability measures on measurable spaces $(\Omega_n, \mathcal{F}_n)$. Then the following statements are equivalent:*

(a) $P_n \triangleleft Q_n$;

(b) If $\dfrac{dQ_n}{dP_n} \xrightarrow[P_n]{\mathcal{D}} U$ *along a sub-sequence, then* $P(U > 0) = 1$;

(c) If $\dfrac{dP_n}{dQ_n} \xrightarrow[Q_n]{\mathcal{D}} V$ *along a sub-sequence, then* $E_Q(V) = 1$,

where U, V are random variables defined on probability spaces $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', Q)$ respectively.

Proof of (a) ⇒ (b). We need to show that if $P_n \triangleleft Q_n$ and $\dfrac{dQ_{n'}}{dP_{n'}} \xrightarrow[P_{n'}]{\mathcal{D}} U$ along some subsequence $\{n'\}$, then $P(U > 0) = 1$. Clearly, $P(U \ge 0) = 1$, as $P_{n'}\left(\dfrac{dQ_{n'}}{dP_{n'}} \ge 0\right) = 1$. Thus it suffices to show that $P(U = 0) = 0$. By item (ii) of the Proposition, for any $\epsilon > 0$,

$$\liminf_{n'\to\infty} P_{n'}\left(\frac{dQ_{n'}}{dP_{n'}} < \epsilon\right) \ge P(U < \epsilon) \ge P(U = 0). \tag{3.3}$$

By Lemma 1, there exists a sequence $\epsilon_{n'} \downarrow 0$ such that

$$\liminf_{n'\to\infty} P_{n'}\left(\frac{dQ_{n'}}{dP_{n'}} < \epsilon_{n'}\right) \ge P(U = 0). \tag{3.4}$$

It remains to show that the left-hand side of Eq. 3.4 is 0. Let μ_n be the probability measure $\frac{1}{2}(P_n + Q_n)$. Then $P_n \ll \mu_n$ and $Q_n \ll \mu_n$. Let p_n and

q_n denote the densities of P_n and Q_n with respect to μ_n. By part *(ii)* of Lemma 2,

$$Q_{n'}\left(\left(\frac{dQ_{n'}}{dP_{n'}}<\epsilon_{n'}\right)\cap(p_{n'}>0)\right)=Q_{n'}^{ac}\left(\frac{dQ_{n'}}{dP_{n'}}<\epsilon_{n'}\right)=\int_{\left(\frac{dQ_{n'}}{dP_{n'}}<\epsilon_{n'}\right)}\frac{dQ_{n'}}{dP_{n'}}\,dP_{n'}$$
$$\leq \epsilon_{n'}\int dP_{n'}=\epsilon_{n'}\to 0. \quad (3.5)$$

Since $P_{n'}\lhd Q_{n'}$, Eq. 3.5 and part *(i)* of Lemma 2 imply

$$P_{n'}\left(\frac{dQ_{n'}}{dP_{n'}}<\epsilon_{n'}\right)=P_{n'}\left(\left(\frac{dQ_{n'}}{dP_{n'}}<\epsilon_{n'}\right)\cap(p_{n'}>0)\right)\to 0.$$

Therefore the left-hand side of Eq. 3.4 is 0.

Proof of (b) ⇒ (c). First note that clearly,

$$p_n+q_n=2, \quad \text{and} \quad 0\leq p_n\leq 2,\ 0\leq q_n\leq 2. \qquad (3.6)$$

By Lemma 2,

$$P_n\{p_n=0\}=Q_n\{q_n=0\}=0, \qquad (3.7)$$
$$P_n\left(q_n\neq p_n\frac{dQ_n}{dP_n}\right)=Q_n\left(p_n\neq q_n\frac{dP_n}{dQ_n}\right)=0. \qquad (3.8)$$

Now suppose $\frac{dP_n}{dQ_n}\xrightarrow[Q_n]{\mathcal{D}}V$ along a sub-sequence $\{n'\}$, then we need to establish $E(V)=1$. By Eqs. 3.6 and 3.7, we have for any $K>0$,

$$P_n\left(\frac{q_n}{p_n}>K\right)\leq\frac{1}{K}E_{P_n}\left(\frac{q_n}{p_n}\right)=\frac{1}{K}E_{P_n}\left(\frac{q_n}{p_n}\,I_{\{p_n>0\}}\right)\leq\frac{1}{K}E_{\mu_n}(q_n)\leq\frac{2}{K}.$$

Hence the sequence $\{q_n/p_n\}$ is tight under $\{P_n\}$. By Prohorov's theorem (see Billingsley 1995, Theorem 25.10) there is a further sub-sequence $\{n''\}$ of $\{n'\}$

such that
$$\frac{q_{n''}}{p_{n''}}\xrightarrow[P_{n''}]{\mathcal{D}}U \quad \text{for some random variable} \quad U, \qquad (3.9)$$

which, by Eq. 3.8, implies $dQ_{n''}/dP_{n''}\xrightarrow[P_{n''}]{\mathcal{D}}U$. Moreover, because dP_n/dQ_n converges in distribution to V under Q_n along n', it also converges in distribution to V under Q_n along n''. To summarize, we have

$$\frac{dQ_{n''}}{dP_{n''}}\xrightarrow[P_{n''}]{\mathcal{D}}U, \quad \text{and} \quad \frac{dP_{n''}}{dQ_{n''}}\xrightarrow[P_{n''}]{\mathcal{D}}V. \qquad (3.10)$$

Thus by Eq. 3.2, the second convergence in Eq. 3.10, and item (i) of the Proposition,

$$E_Q(V) \leq \liminf_{n''\to\infty} E_{Q_{n''}}\left(\frac{dP_{n''}}{dQ_{n''}}\right) \leq 1.$$

Since $P_n(p_n = 0) = 0$, we have for any $c > 0$,

$$\begin{aligned} E_{Q_n}\left(\frac{p_n}{q_n} I_{\{p_n/q_n \leq c\}}\right) &\geq E_{\mu_n}\left(p_n I_{\{p_n/q_n\leq c,\ p_n>0,\ q_n>0\}}\right) \\ &= P_n((p_n/q_n) \leq c,\ p_n > 0,\ q_n > 0) \\ &= P_n((q_n/p_n) \geq 1/c,\ p_n > 0,\ q_n > 0) \\ &= P_n((q_n/p_n) \geq 1/c,\ p_n > 0) \\ &= P_n((q_n/p_n) \geq 1/c) \\ &\geq P_n((q_n/p_n) > 1/c). \end{aligned}$$

Hence

$$\begin{aligned} 1 \geq E_Q(V) \geq E_Q\left(V I_{\{V\leq c\}}\right) &\geq \limsup_{n''\to\infty} E_{Q_{n''}}\left(\frac{p_{n''}}{q_{n''}} I_{\{p_{n''}/q_{n''}\leq c\}}\right) \\ &\geq \liminf_{n''\to\infty} P_{n''}((q_{n''}/p_{n''}) > 1/c) \\ &\geq P(U > 1/c), \end{aligned}$$

where the third inequality follows by applying item *(iii)* of the Proposition to the upper semi-continuous function $f(x) = xI_{\{x\leq 1/c\}}$ that is bounded from above, and the last inequality follows by applying item *(ii)* of the Proposition to the open set $(x > 1/c)$. By the continuity of probability, $\lim_{c\to\infty} P(U > 1/c) = P(U > 0)$. Hence $E_Q(V) \geq P(U > 0)$. By the first convergence in Eq. 3.10 and statement (b), $P(U > 0) = 1$. Hence $E_Q(V) = 1$, thus establishing statement (c).

Proof of (c) ⇒ (a). Suppose that statement (c) holds. If $Q_n(A_n) \to 0$ for a sequence of measurable sets $\{A_n\}$, then we need to show that $P_n(A_n) \to 0$. First note that, by Eq. 3.2,

$$Q_n\left(\frac{dP_n}{dQ_n} > K\right) \leq \frac{1}{K} E_{Q_n}\left(\frac{dP_n}{dQ_n}\right) \leq \frac{1}{K},$$

and hence the sequence of Radon-Nikodym derivatives $\left\{\frac{dP_n}{dQ_n}\right\}$ is tight under the sequence of measures $\{Q_n\}$. By Prohorov's Theorem (see Billingsley 1995, Theorem 25.10) every subsequence $\{n'\}$ has a further subsequence $\{n''\}$, and a random variable V, such that $\frac{dP_{n''}}{dQ_{n''}} \xrightarrow[Q_{n''}]{\mathcal{D}} V$.

Let B_n denote the complement of the measurable set A_n. Since $Q_n(A_n) \to 0 \Rightarrow I_{B_n} \xrightarrow[Q_n]{\mathcal{D}} 1$, by a version of Slutsky's theorem (see for example Billingsley 1995, Theorem 25.4)

$$0 \leq \frac{dP_{n''}}{dQ_{n''}} I_{B_{n''}} \xrightarrow[Q_{n''}]{\mathcal{D}} V.$$

Since

$$P_n(B_n) \geq P_n^{ac}(B_n) = \int I_{B_n} \frac{dP_n}{dQ_n} dQ_n,$$

we have by item (i) of the Proposition, and Eq. 3.2,

$$\begin{aligned} E_Q(V) &\leq \liminf_{n'' \to \infty} E_{Q_{n''}} \left(\frac{dP_{n''}}{dQ_{n''}} I_{B_{n''}} \right) \\ &\leq \liminf_{n'' \to \infty} P_{n''}(B_{n''}) = 1 - \limsup_{n'' \to \infty} P_{n''}(A_{n''}). \end{aligned}$$

By statement (c), $E_Q(V) = 1$, so $P_{n''}(A_{n''}) \to 0$. Thus we have shown that every subsequence of $\{P_n(A_n)\}$ contains a further subsequence $\{P_{n''}(A_{n''})\}$ that converges to 0. Hence $P_n(A_n) \to 0$. This completes the proof.

The above detailed proof is simpler than the “standard proof”, say, given in van der Vaart (1998). The proofs of (a) $\Rightarrow$ (b) and (c) $\Rightarrow$ (a) are relatively straightforward and intuitive, but the proof of (b) $\Rightarrow$ (c) is more intricate, and the complexity lies in the passage from U to V. In the proof of (van der Vaart 1998, Lemma 6.4) V is related to U via a third random variable W. In comparison, in our proof V is directly linked to U. This more intuitive proof allows us to see exactly how the sequences of measures P_n and Q_n are interchanged, just like in the two-measure case.

We now present some results on probability measures and convergence of random vectors in a form useful to the proof of the Theorem.

First, a technical lemma is established, which is used only in the proof of $(a) \Rightarrow (b)$ of the theorem.

Lemma 1. *Let a be a real number. Suppose that $g_n : \mathbb{R} \to \mathbb{R}$ is a sequence of functions such that, for any $\epsilon > 0$, $\liminf_{n\to\infty} g_n(\epsilon) \geq a$. Then there is a sequence $\epsilon_n \downarrow 0$ such that $\liminf_{n\to\infty} g_n(\epsilon_n) \geq a$.*

Proof. Since for each integer $k \geq 1$, $\liminf_{n\to\infty} g_n(1/k) \geq a$, there is a positive integer n_k such that, for all $n \geq n_k$, $g_n(1/k) > a - 1/k$. Without loss of generality, we can assume that $n_{k+1} > n_k$ for all $k = 1, 2, \ldots$. Let $\epsilon_n = 1/k$ for $n_k \leq n < n_{k+1}$. Then $g_n(\epsilon_n) > a - \epsilon_n$ for all $n \geq n_1$. Clearly $\epsilon_n \downarrow 0$ and $\liminf_{n\to\infty} g_n(\epsilon_n) \geq a$.

The next lemma connects three probability measures.

Lemma 2. *Let μ, P_1, P_2 be probability measures on a measurable space $(\Omega, \mathcal{F})$, $P_i \ll \mu$, $p_i = \frac{dP_i}{d\mu}$, for $i = 1, 2$. Then the following statements hold.*

(i) $P_i\{p_i = 0\} = 0,\ i = 1, 2.$

(ii) If

$$P_2^{ac}(A) = P_2(A \cap (p_1 > 0)),\ \textit{and}\ \ P_2^s(A) = P_2(A \cap (p_1 = 0)),$$

then $P_2 = P_2^{ac} + P_2^s$, $P_2^{ac} \ll P_1$, and P_2^s and P_1 are mutually singular.

(iii)

$$P_1\left(p_2 \neq \frac{dP_2}{dP_1} p_1\right) = 0.$$

Proof.

(i) This follows from Radon-Nikodym Theorem (see Billingsley 1995, Theorem 32.2).

(ii) The equality $P_2 = P_2^{ac} + P_2^s$ follows by construction. Since, for every $A \in \mathcal{F}$,

$$P_1(A) = \int_A p_1\, d\mu = \int_{A \cap (p_1 > 0)} p_1\, d\mu,$$

by Theorem 15.2(ii) of Billingsley (1995), if $P_1(A) = 0$ then $\mu(A \cap (p_1 > 0)) = 0$. This implies

$$P_2^{ac}(A) = P_2(A \cap (p_1 > 0)) = 0 \ \text{ as } \ P_2 \ll \mu.$$

Therefore $P_2^{ac} \ll P_1$. Since $P_1(p_1 = 0) = 0$ and

$$P_2^s(p_1 > 0) = P_2((p_1 > 0) \cap (p_1 = 0)) = P_2(\emptyset) = 0,$$

P_2^s and P_1 are mutually singular.

(iii) By statement (ii), and Radon-Nikodym Theorem, we have for all $A \in \mathcal{F}$,

$$\int_A p_2\, I_{\{p_1 > 0\}}\, d\mu = P_2(A \cap (p_1 > 0))$$
$$= P_2^{ac}(A) = \int_A \frac{dP_2}{dP_1}\, dP_1 = \int_A \frac{dP_2}{dP_1}\, p_1\, d\mu$$

$$= \int_A \frac{dP_2}{dP_1}\, p_1 \; I_{\{p_1>0\}}\, d\mu,$$

and hence

$$\mu\left(\left(p_2 \neq \frac{dP_2}{dP_1}\, p_1\right) \cap (p_1 > 0)\right) = 0.$$

As $P_1 \ll \mu$, and $P_1(p_1 > 0) = 1$, this leads to

$$P_1\left(p_2 \neq \frac{dP_2}{dP_1}\, p_1\right) = 0.$$

This completes the proof.

The proposition below collects versions of Fatou's lemma and portmanteau theorem.

Proposition. *Let X, X_n be random variables defined on the probability spaces $(\Omega, \mathcal{F}, P)$, $(\Omega_n, \mathcal{F}_n, P_n)$. Suppose $X_n \xrightarrow[P_n]{\mathcal{D}} X$. Then the following statements hold.*

1. $E(|X|) \leq \liminf_{n\to\infty} E_{P_n}(|X_n|)$.
2. *For any open set G,* $\liminf_{n\to\infty} P_n(X_n \in G) \geq P(X \in G)$.
3. *If g is an upper semi-continuous function that is bounded from above, then*
$$\limsup_{n\to\infty} E_{P_n}(g(X_n)) \leq E(g(X)).$$

Proof. By Skorohod's Theorem (see Billingsley 1995, Theorem 25.6) there exist random variables Y, Y_n defined on a common probability space $(\Omega_0, \mathcal{F}_0, P_0)$ such that Y_n has distribution $P_n \circ X_n^{-1}$, Y has distribution $P \circ X^{-1}$ such that $Y_n(\omega) \to Y(\omega)$ for all $\omega \in \Omega_0$. Now (i) follows from Fatou's lemma (see Billingsley 1995, Theorem 16.3) as

$$E(|X|) = E_{P_0}(|Y|) \leq \liminf_{n\to\infty} E_{P_0}(|Y_n|) = \liminf_{n\to\infty} E_{P_n}(|X_n|).$$

Statements (ii) and (iii) are part of Portmanteau Theorem (see for example Billingsley 1995, Theorem 29.1 and Problem 29.1).

Acknowledgments. G. Jogesh Babu thanks the Statistical and Applied Mathematical Sciences Institute (SAMSI), for supporting his research during his visit to SAMSI in Fall 2019. This material was based upon work partially supported by the National Science Foundation under Grant DMS-1638521 to the Statistical and Applied Mathematical Sciences Institute. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Bing Li's work is partially supported by the National Science Foundation Grant DMS-1713078.

References

BILLINGSLEY, P. (1995). *Probablity and Measure*. Third Edition. John Wiley & Sons.

FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A* **222**, 594–604.

FISHER, R. A. (1925). Theory of statistical estimation. *Proc. Cambridge Phil. Soc.* **22**, 700–725.

HÁJEK, J. (1970). A characterization of limiting distributions of regular estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **14**, 323–330.

LE CAM, L. (1953). On some asymptotic Properties of maximum likelihood estimates and related Bayes estimates. *Univ. California Publ. Statistic* **1**, 277–330.

LE CAM, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ Statistic.* **3**, 370–98.

LI, B. and BABU, G. J. (2019). *A Graduate Course on Statistical Inference*. Springer.

VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Publisher's Note. Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

G. Jogesh Babu
Department of Statistics, and
Department of Astronomy &
Astrophysics,
The Pennsylvania State University,
University Park, PA, USA
E-mail: babu@psu.edu

G. Jogesh Babu
Bing Li
Department of Statistics, The
Pennsylvania State University,
University Park, PA, USA
E-mail: bxl9@psu.edu

Paper received: 26 February 2020.