AN ASYMPTOTICALLY EFFICIENT ESTIMATE IN
TIME SERIES ANALYSIS*

BY

RICHARD A. VITALE

Brown University

1. A problem in linear regression. It is well known that often a theoretically best statistical estimate may be of limited value in actual practice. The computation required may be excessive, for instance, or more critically, *a priori* information may be needed which is unavailable to the investigator. As a result, there has been considerable interest traditionally in the formulation of other estimates which are close to optimal and at the same time computationally feasible and less dependent on intimate knowledge of the particular phenomenon being studied. The purpose of this work is to make such an addition to an already rich branch of statistics—time series analysis.

Let us consider the simplest example of linear regression in time series. We have a discrete time process

$$x_k = m + y_k , \qquad k = 1, 2, \cdots \tag{1.1}$$

where m is an unknown real scalar and the y process is

    (i) real-valued,
    (ii) wide-sense stationary (zero mean),
    (iii) with continuous spectral density $f(\lambda)$.

The problem is to estimate the parameter m using observations of the x process $\{x_1 , \cdots , x_n\}$. It is natural to consider linear estimates $m_{EST} = \sum c_k y_k$ where the c_k's are normed to eliminate bias, $\sum c_k = 1$. The most important such estimate is the best linear unbiased (BLU) estimate, which minimizes variance. It is possible to show that if the covariance matrix R of the y process is invertible, then this estimate has the form

$$m_{BLU} = (e^T R^{-1} e)^{-1} e^T R^{-1} x \tag{1.2}$$

where

$$e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Under our conditions for the y process, the BLU estimate exists, is unique, and has variance which tends to zero as $n \to \infty$ (see Grenander and Rosenblatt [2]). However, it illustrates the dilemma mentioned above. The covariance matrix R may not be known in practice and even if it is, its inversion for high dimension will pose computational problems. A common remedy has been to ignore the structure of the y process and to use the straight arithmetic mean which has justification as the least squares estimate:

$$m_{LS} = \frac{1}{n} \sum_1^n x_k .\tag{1.3}$$

It has been known for some time (see Grenander [1]) that in certain cases the least squares estimate is good in the sense that its efficiency tends to one (or its asymptotic efficiency is one):

$$\frac{\text{Var}\ (m_{BLU})}{\text{Var}\ (m_{LS})} \to 1 \quad \text{as } n \to \infty .\tag{1.4}$$

This result does *not* hold for all spectral densities $f(\lambda)$. We may weaken our conditions so that $f(\lambda)$ is only piecewise continuous but the restriction must be added that $f(\lambda)$ be bounded away from zero in a neighborhood of $\lambda = 0$. This restriction is fundamental since the form of the regression distinguishes the frequency $\lambda = 0$ (see Grenander and Szegö [3]). It is at this point that we depart from previous results to examine the case where $f(0) = 0$.

2. An asymptotically efficient estimate. We shall restrict our discussion to the class Y of spectral densities where $f(\lambda)$ is an even, continuous, positive function except for a second-order zero at the origin. We shall often use the factorization $f(\lambda) = (2\pi)^{-1} |1 - e^{i\lambda}|^2 g(\lambda)$ in order to isolate the zero.

The first question to be answered is whether the least squares estimate is asymptotically efficient for the class Y. At this point it is natural to introduce the role which the computer and automatic plotter played in this investigation. A certain amount of preliminary experimentation is reflected in Figs. 2.1–2.8. Fig. 2.1 shows spectral density $f(\lambda) = (2\pi)^{-1} (1 - .1 \cos \lambda)$ for which the least squares estimate is known to be asymptotically efficient (see Figs. 2.2–2.4). Fig. 2.5 shows spectral density $f(\lambda) = (2\pi)^{-1}(1 - .9 \cos \lambda)$ for which the least squares estimate is again asymptotically efficient. But note the slower rate of convergence. This suggests that the condition $f(0) = 0$ might be a limiting case where the least squares estimate breaks down in some way. This hypothesis receives substantiation in the next example. Fig. 2.9 shows spectral density $f(\lambda) = (2\pi)^{-1}(1 - \cos \lambda)$. Note that while the least squares estimate appears to be tending to zero (in variance), its efficiency also is decreasing (Figs. 2.10–2.12). On the other hand, the coefficients in m_{BLU} appear to have a rather nice form (they are plotted in Fig. 2.13 versus their subscripts). In fact, it is not difficult to make these observations more precise.

THEOREM 2.1. *Let* $f(\lambda) = (2\pi)^{-1} (1 - \cos \lambda) = \pi^{-1} |1 - e^{i\lambda}|^2$, *then* $n^3 \text{Var}\ (m_{BLU})$ $\to 12$.

Proof. The requirement $Rc = \mu e$ from (1.2) gives a simple set of equations:

$$2c_1 - c_2 = \mu, \qquad 2c_n - c_{n-1} = \mu,$$

$$2c_k - (c_{k-1} + c_{k+1}) = \mu, \qquad k = 2, \cdots, n - 1.$$

These can be solved with the side condition $\sum c_k = 1$ to yield

$$c_k = \frac{6}{n+2}\left[\frac{k}{n}\left(1 - \frac{k}{n+1}\right)\right].\tag{2.1}$$

A straightforward calculation shows that

$$\text{Var}\,(m_{\text{BLU}}) = \frac{12}{n(n+1)(n+2)}\,.\tag{2.2}$$

Theorem 2.2 shows that the least squares estimate has a slower decrease in variance which produces zero asymptotic efficiency.

THEOREM 2.2. *Let* $f(\lambda) = (2\pi)^{-1}(1 - \cos\lambda) = \pi^{-1}\,|1 - e^{i\lambda}|^2$. *Then* $n^2\,\text{Var}\,(m_{\text{LS}}) \equiv 2$.

Proof.

$$\text{Var}\,(m_{\text{LS}}) = \frac{1}{n^2}\,e^{\text{T}}Re$$

$$= \frac{1}{n^2}\int_{-\pi}^{+\pi}|\sum_1^n e^{ik\lambda}|^2\,\frac{1}{2\pi}\,|1 - e^{i\lambda}|^2\,d\lambda$$

$$= \frac{1}{\pi n^2}\int_{-\pi}^{+\pi}(1 - \cos n\lambda)\,d\lambda.$$

Hence, $n^2\,\text{Var}\,(m_{\text{LS}}) \equiv 2$.

At this point in the investigation it seemed reasonable to abandon the least squares estimate as asymptotically efficient over the class Y. A different approach was suggested by the particularly simple form of the m_{BLU} coefficients in Fig. 2.13. What about using this type of estimate for all spectral density functions in the class Y? An analogy may be made in the following way: the least squares estimate, being the BLU estimate for $f(\lambda) = (2\pi)^{-1}$, is invariant with regard to asymptotic efficiency under multiplication of $f(\lambda)$ by an arbitrary positive function (i.e. for all densities $(2\pi)^{-1}g(\lambda)$). That this type of invariance might hold for the class Y received support in the next example investigated. Figs. 2.14–2.19 show results for $f(\lambda) = (2\pi)^{-1}(1 - \cos 2\lambda)$. The *parabolic estimate* m_p is defined for computational convenience as a slight modification of the BLU estimate for $f(\lambda) = (2\pi)^{-1}(1 - \cos\lambda)$:

$$c_k = \frac{6n}{n^2 - 1}\left[\frac{k}{n}\left(1 - \frac{k}{n}\right)\right].\tag{2.3}$$

The next result shows that this modification is not too drastic.

THEOREM 2.3. *With the assumptions of theorem 2.1,* m_p *is asymptotically efficient.*

Proof. Straightforward calculation shows that

$$\text{Var}\,(m_p) = \frac{12}{n(n^2 - 1)}$$

and hence

$$\text{Var}\,(m_{\text{BLU}})/\text{Var}\,(m_p) \to 1.$$

We are now prepared to show that the parabolic estimate is indeed asymptotically efficient over the class Y. Theorem 2.4 gives a rate of convergence for the parabolic estimate.
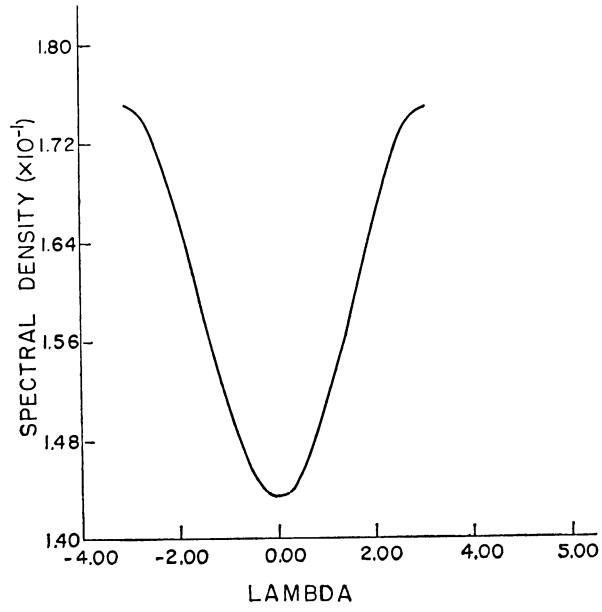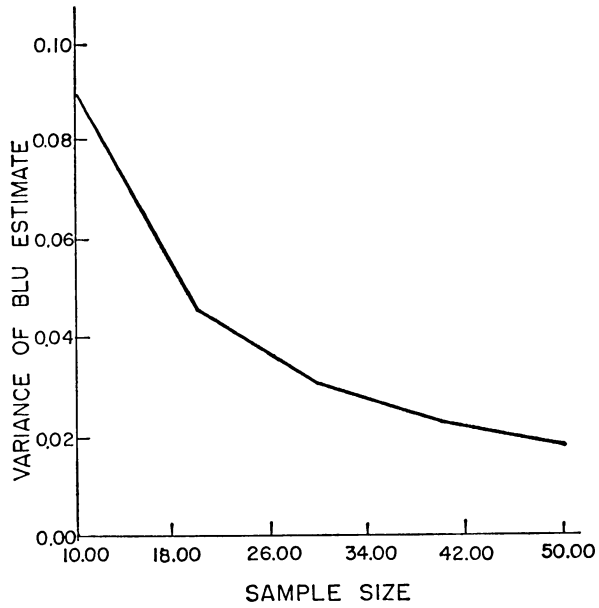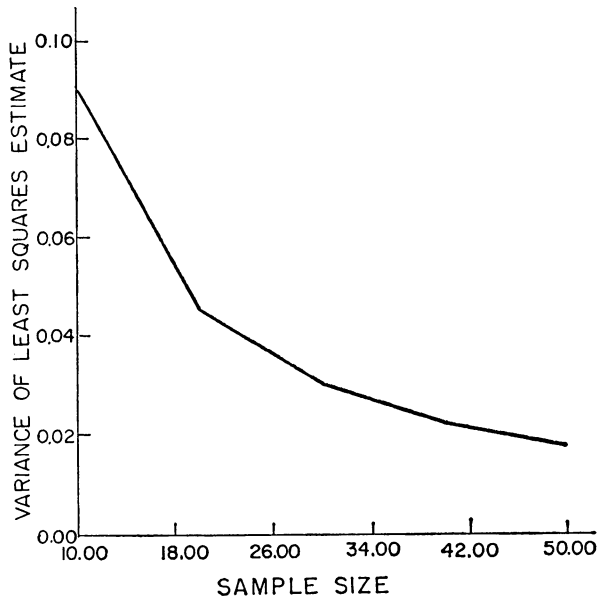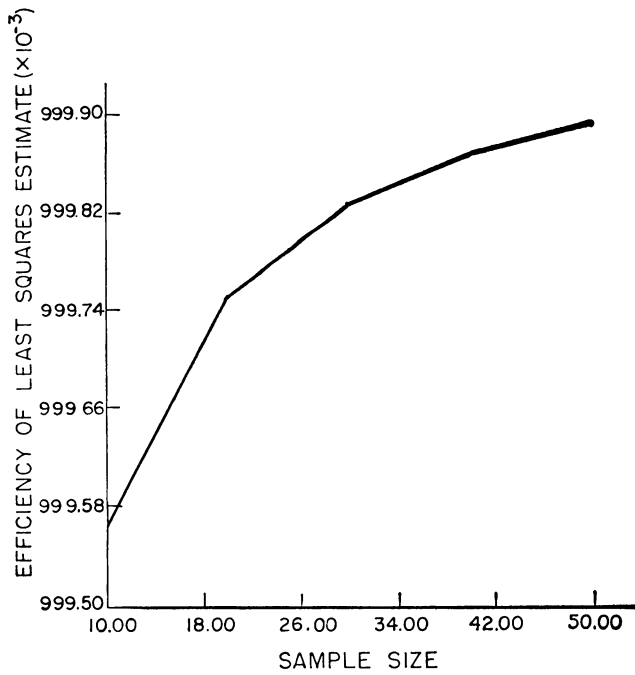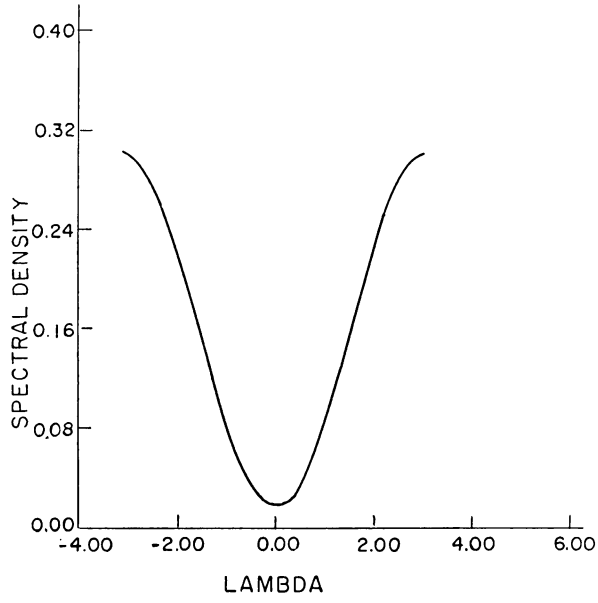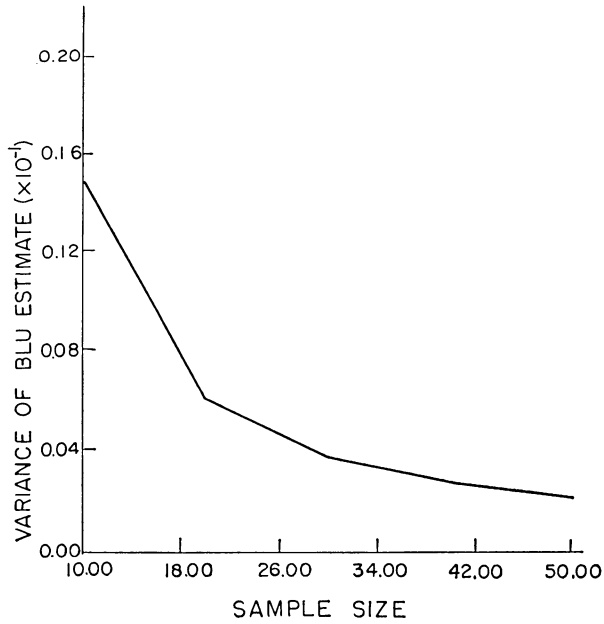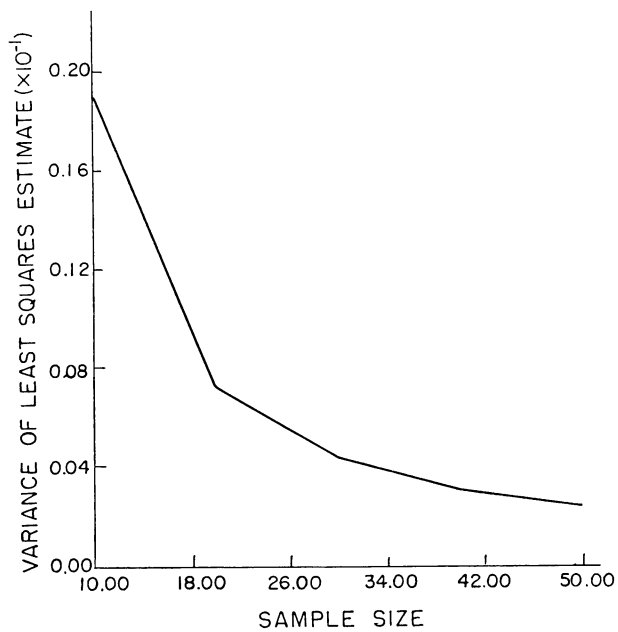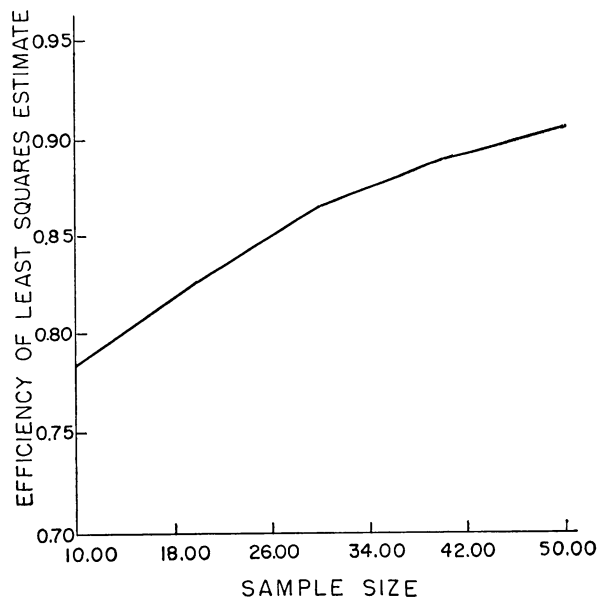
FIG. 2.1

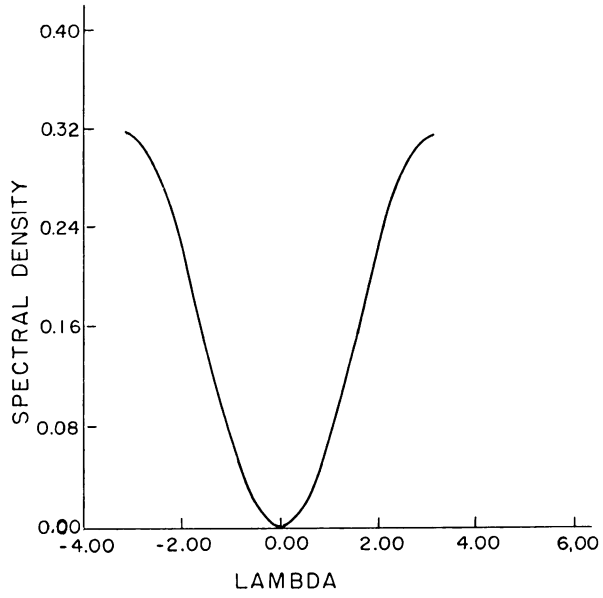FIG. 2.2

FIG. 2.3

FIG. 2.4

FIG. 2.5

FIG. 2.6
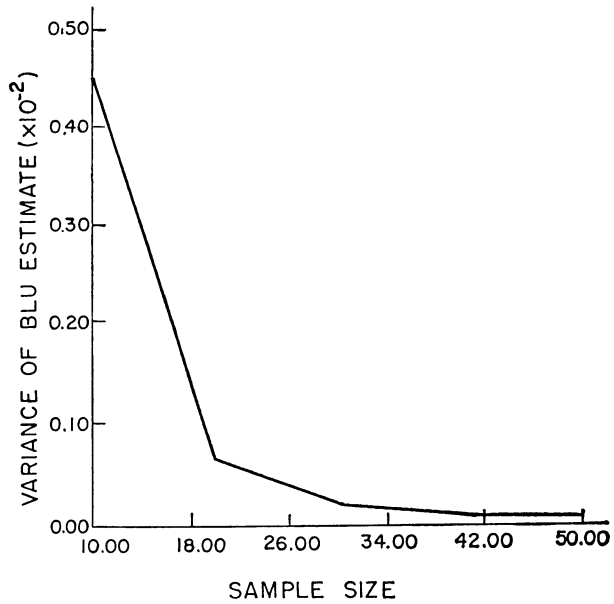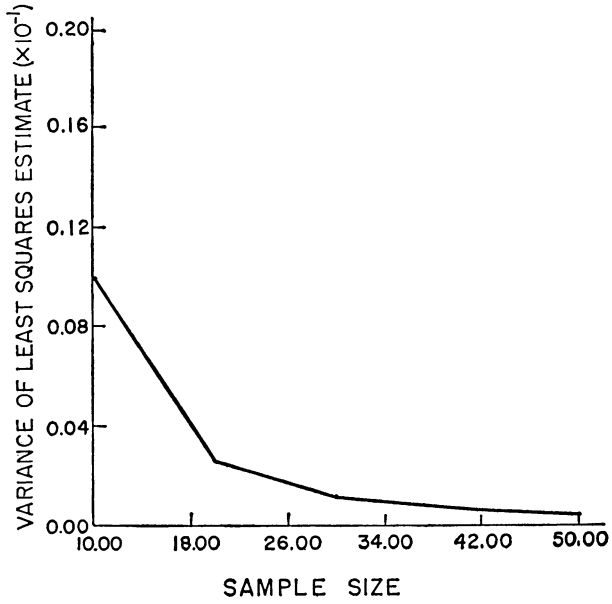
FIG. 2.7

FIG. 2.8

FIG. 2.9
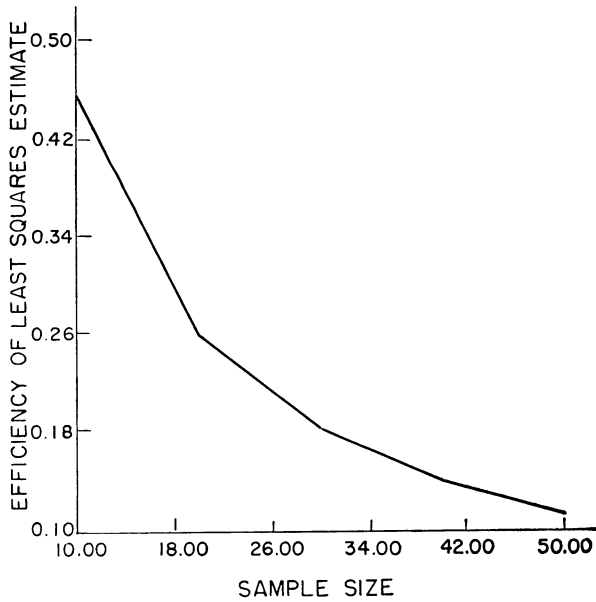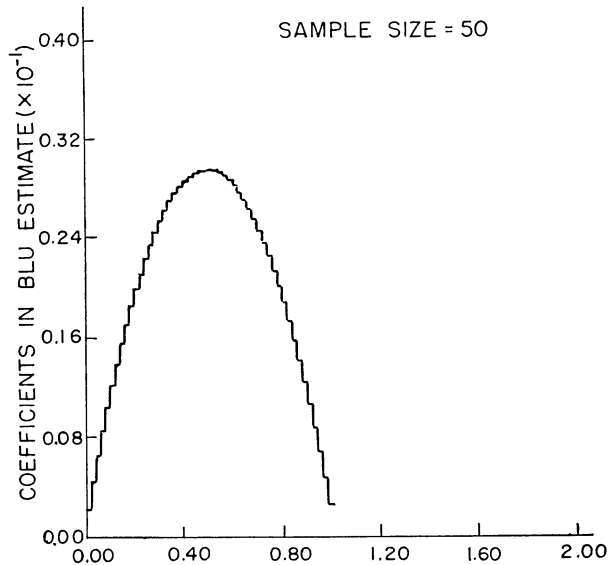

FIG. 2.10

FIG. 2.11

FIG. 2.12

FIG. 2.13

THEOREM 2.4. *Let $f(\lambda) \in Y$. Then*

$$n^3 \operatorname{Var} (m_p) \to 12g(0).$$

Proof. Letting $\{c_k\}$ represent the coefficients of the parabolic estimate, we set

$$p_n(z) = \sum_1^n c_k z^k$$

$$= \frac{6n}{n^2(n-1)} \sum_1^n \frac{k}{n}\left(1 - \frac{k}{n}\right) z^k$$

so that

$$\frac{n^3 \operatorname{Var} (m_p)}{12} = \frac{n^3}{12} \frac{1}{2\pi} \int_{-\pi}^{+\pi} |p_n(e^{i\lambda})|^2 \, |1 - e^{i\lambda}|^2 \, g(\lambda) \, d\lambda.$$

Let

$$h_n(\lambda) = \frac{n^3}{12} \frac{1}{2\pi} |1 - e^{i\lambda}|^2 \, |p_n(e^{i\lambda})|^2$$

and we note that the statement of Theorem 2.3 is equivalent to

$$\int_{-\pi}^{+\pi} h_n(\lambda) \, d\lambda \to 1.$$

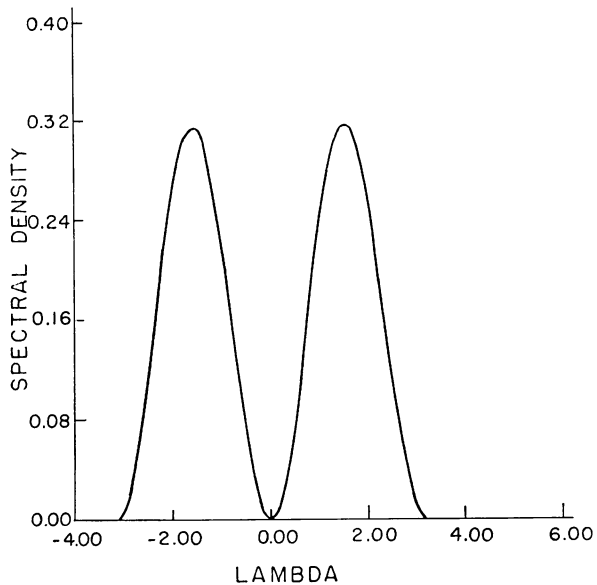If we can show that $h_n(\lambda) \to 0$ uniformly in λ on $I_\epsilon = [-\pi, -\epsilon] \cup [\epsilon, \pi]$ for arbitrary
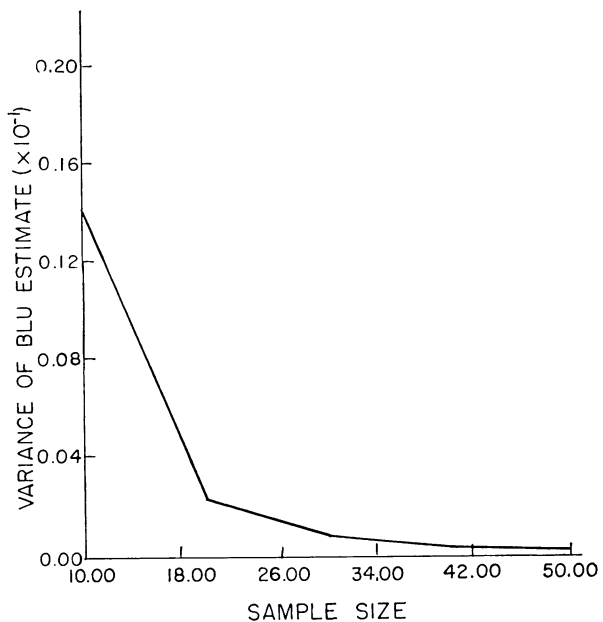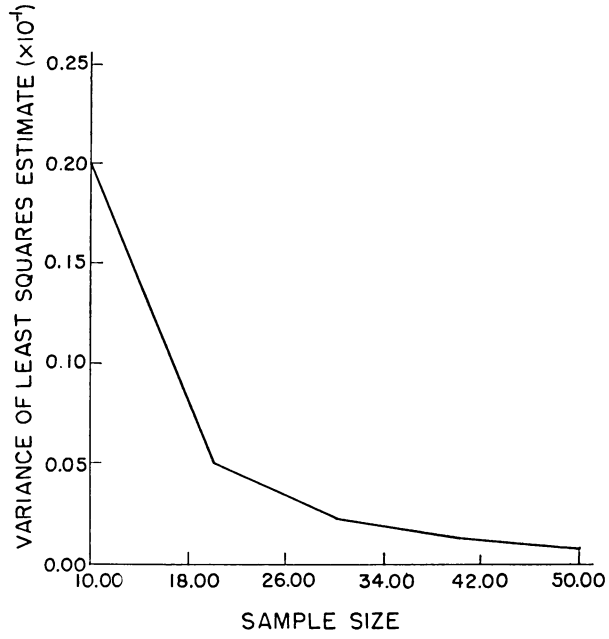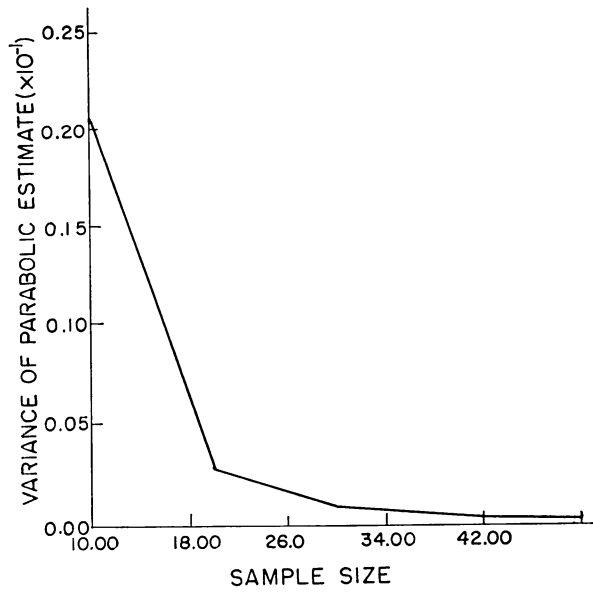
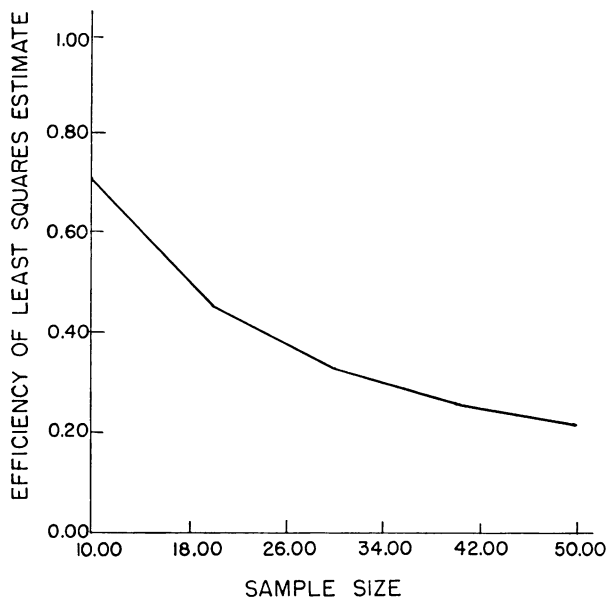FIG. 2.14

FIG 2.15

FIG. 2.16
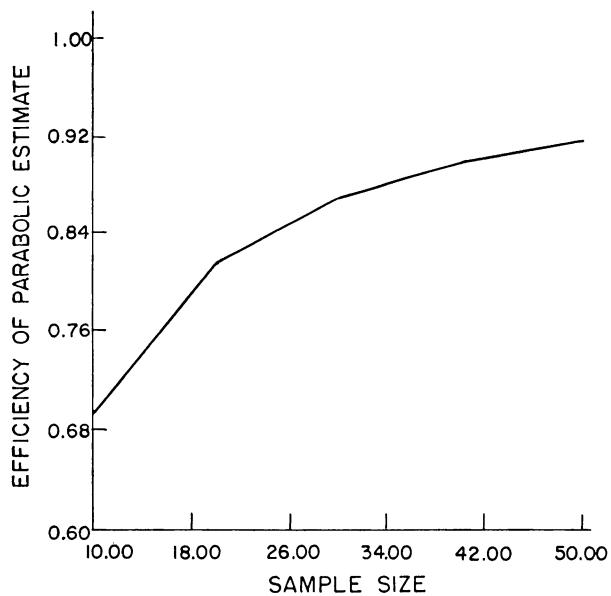

FIG. 2.17

FIG. 2.18

FIG. 2.19

but fixed ϵ in $(0, \pi)$, then the result of the theorem will follow. A bit of computation will show this:

$$
\begin{aligned}
\left(\frac{n^3}{24\pi}\right)^{1/2} p_n(e^{i\lambda}) &= \frac{6n^{5/2}}{(24\pi)^{1/2}(n^2 - 1)} \sum_1^n \frac{k}{n}\left(1 - \frac{k}{n}\right)e^{ik\lambda} \\
&= \frac{6n^{3/2}}{(24\pi)^{1/2}(n^2 - 1)}\left[\frac{1}{i}\frac{d}{d\lambda}\left(\sum_1^n e^{ik\lambda}\right) + \frac{1}{n}\frac{d^2}{d\lambda^2}\left(\sum_1^n e^{ik\lambda}\right)\right] \\
&= \frac{6n^{3/2}}{(24\pi)^{1/2}(n^2 - 1)}\cdot\frac{1}{n(1 - e^{i\lambda})^3}\cdot[(n - 1)e^{i\lambda} + (n + 1)e^{i(n+1)\lambda} \\
&\quad - (n + 1)e^{i2\lambda} - (n - 1)e^{i(n+2)\lambda}]
\end{aligned}
$$

so

$$
|(h_n(\lambda))^{1/2}| \le \frac{6n^{3/2}}{(24\pi)^{1/2}(n^2 - 1)}\frac{4n}{n\,|1 - e^{i\lambda}|^2}\cdot
$$

We fix ϵ, $0 < \epsilon < \pi$ and note that $|1 - e^{i\lambda}|^2$ is bounded away from zero on I_ϵ, say $|1 - e^{i\lambda}|^2 > \delta$. Then

$$
|(h_n(\lambda))^{1/2}| \le \frac{6n^{3/2}}{(24\pi)^{1/2}(n^2 - 1)}\frac{4}{\delta} \quad \text{on} \quad I_\epsilon.
$$

Hence, $h_n(\lambda)$ tends to zero uniformly on I_ϵ and the theorem follows.

Our main result is contained in the next theorem. The method of proof is similar to that used in Grenander and Rosenblatt [2] for the case of positive spectral densities.

THEOREM 2.5.   *Let $f(\lambda) \in Y$. Then m_p is asymptotically efficient.*
Proof.   In view of Theorem 2.4, we only have to show that

$$
n^3 \,\mathrm{Var}\,(m_{\mathrm{BLU}})/12 \to g(0).
$$

The proof will be in two steps: first we verify the result for a dense subclass of densities in Y and then we perform an extension through a continuity argument.

Let us consider a y process of the form

$$
b_0 y_t + b_1 y_{t+1} + \cdots + b_\alpha y_{t+\alpha} = \epsilon_{t+\alpha} - \epsilon_{t+\alpha-1} \tag{2.4}
$$

where $E\epsilon_t\epsilon_s = \delta_{ts}$ and with the roots of the characteristic polynomial $\sum_{k=0}^\alpha b_k z^k$ inside the unit circle (this implies that $Ey_j\epsilon_{j+k} = 0$ for all j and $k > 0$). We consider the space spanned by $y_1, \cdots, y_n$. Now $y_1, \cdots, y_{\alpha-1}$ are orthogonal to $\epsilon_{\alpha+1} - \epsilon_\alpha$, $\epsilon_{\alpha+2} - \epsilon_{\alpha-1}$, $\cdots$, $\epsilon_n - \epsilon_{n-1}$ so that we can perform a change of co-ordinates (using Gram–Schmidt)

$$
\Delta y = \eta
$$

$$
\begin{aligned}
d_{11}y_1 \quad d_{12}y_2 + \cdots + d_{1,\alpha-1}y_{\alpha-1} &= \eta_1 \\
d_{21}y_1 + d_{22}y_2 + \cdots + d_{2,\alpha-2}y_{\alpha-2} &= \eta_2 \\
d_{31}y_1 + d_{32}y_2 + \cdots + d_{3,\alpha-3}y_{\alpha-3} &= \eta_3 \\
\vdots \quad\quad\quad\quad\quad\quad\quad\quad & \\
-b_\alpha y_\alpha &= \eta_\alpha
\end{aligned}
$$

$$b_0 y_1 + b_1 y_2 \qquad\qquad + b_\alpha y_{\alpha+1} \qquad = \eta_{\alpha+1}$$

$$= \epsilon_{\alpha+1} - \epsilon_\alpha$$

$$0 \cdot y_1 + b_1 y_2 + b_2 y_3 \qquad + b_{\alpha+1} y_{\alpha+2} \qquad = \eta_{\alpha+2}$$

$$\cdots \qquad = \epsilon_{\alpha+2} - \epsilon_{\alpha+1}$$

$$\cdots$$
$$\cdots$$

$$0 \cdot y_1 + 0 \cdot y_2 \cdots + b_0 y_{n-a} + \cdots b_a y_n \qquad = \eta_n$$

$$= \epsilon_n - \epsilon_{n-1}$$

such that for $i, j \leq \alpha$, $E\eta_i \eta_j = 0$ and for $i < \alpha$, $E\eta_i^2 = 1$. We note further that for $i, j > \alpha$, $E\eta_i \eta_n = -1$ if $i \neq j$ and $E\eta_i^2 = 2$. Also for $i > \alpha + 1$, $E\eta_\alpha \eta_i = 0$. For $i = \alpha + 1$, we have $E\eta_\alpha \eta_{\alpha+1} = -1$ since

$$E\eta_\alpha \eta_{\alpha+1} = E[-b_\alpha y_\alpha (\epsilon_{\alpha+1} - \epsilon_\alpha)] = -b E y_\alpha \epsilon_{\alpha+1}.$$

But $b_0 y_0 + \cdots + b_\alpha y_\alpha = \epsilon_\alpha - \epsilon_{\alpha-1}$ which implies that $E y_\alpha \epsilon_\alpha = 1/b_\alpha$, hence

$$E\eta_\alpha \eta_{\alpha+1} = -b_\alpha E y_\alpha \epsilon_{\alpha+1} = -1.$$

Therefore the covariance matrix of the vector η is given by

$$Q = E\eta\eta^T = \begin{bmatrix} I_{\alpha-1} & 0 \\ & \\ 0 & J_{n-\alpha+1} \end{bmatrix}$$

where

$$I_{\alpha-1} = \begin{bmatrix} 1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & 1 \end{bmatrix} (\alpha - 1) \times (\alpha - 1)$$

and

$$J_{n-\alpha+1}$$

$$= \begin{bmatrix} \beta & -1 & \cdot & \cdot & & \cdot \\ -1 & 2 & -1 & 0 & & \cdot \\ \cdot & -1 & \cdot & \cdot & & \cdot \\ \cdot & 0 & \cdot & \cdot & & -1 \\ \cdot & \cdot & & \cdot & -1 & 2 \end{bmatrix} (n - \alpha + 1) \times (n - \alpha + 1)(\beta = E\eta_\alpha^2 = b_\alpha^2 E y_\alpha^2).$$

Now the variance of the BLU estimate is given by $(e^T R^{-1} e)^{-1}$ where R is the covariance matrix of the y process. Since $\Delta y = \eta$, we have $\Delta R \Delta^T = Q$ or $R^{-1} = \Delta^T Q^{-1} \Delta$ so that $\text{Var}(m_{\text{BLU}}) = ((\Delta e)^T Q^{-1} \Delta e)^{-1}$. For convenience we let $\gamma = \Delta e$,

$$\gamma_i = \sum_{l=1}^{\alpha-1} d_{il} \qquad i < \alpha,$$

$$= -b_\alpha \qquad\qquad i = \alpha,$$

$$= \sum_{k=1}^{\alpha} b_k = B \qquad i > \alpha.$$

Now

$$\frac{1}{\text{Var}\,(m_{\text{BLU}})} = \gamma^T Q^{-1} \gamma = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_i^T q_{ij}^{(-1)} \gamma_j \ .$$

We split this sum into four parts:

(i)

$$S_1 = \sum_{i=\alpha}^{n} \sum_{j=\alpha}^{n} \gamma_i^T q_{ij}^{(-1)} \gamma_j$$

$$= |B|^2 \sum_{j=\alpha}^{n} \sum_{j=\alpha}^{n} q_{ij}^{(-1)} + \epsilon$$

$$= |B|^2 \sum_{i=\alpha}^{n} \sum_{j=\alpha}^{n} e_i^T q_{ij}^{(-1)} e_j + \epsilon$$

$$= |B|^2 e^T J_{n-\alpha+1}^{-1} e + \epsilon.$$

It is easy to show that $J_{n-\alpha-1} a = e$ where

$$a = \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{h-\alpha+1} \end{pmatrix}$$

is given by

$$a_i = \frac{j^2}{2} + j\left[\frac{1}{2} + \frac{(\beta-1)(l^2+l)}{2[l(\beta-1)+1]}\right] + \frac{(2-\beta)(l^2+l)}{2[l(\beta-1)+1]}$$

where $l = n - \alpha + 1$. We have

$$e^T J_{n-\alpha+1}^{-1} e = e^T J_l^{-1} e = e^T a = (\sum_{1}^{l} a_i = -\frac{1}{2}\left[\frac{l(l+1)(2l+1)}{6}\right]$$

$$+ \frac{l(l+1)}{2}\left[\frac{1}{2} + \frac{(\beta-1)(l^2+l)}{2[l(\beta-1)+1]}\right] + \frac{(2-\beta)(l^3+l^2)}{2[l(\beta-1)+1]},$$

and hence $\lim_{l\to\infty} e^T J_l^{-1} e/l^3 = 1/12$ which implies that $\lim_{n\to\infty} e^T J_{n-\alpha+1}^{-1} e/n^3 = 1/12$. An auxiliary argument shows $\epsilon = o(n^3)$ implying

$$\lim_{n\to\infty} 12 S_1/n^3 = |B|^2. \tag{2.5}$$

(ii)

$$S_2 = \sum_{i=1}^{\alpha-1} \sum_{j=1}^{\alpha-1} \gamma_i^T q_{ij}^{(-1)} \gamma_j$$

$$= \sum_{i=1}^{\alpha-1} \sum_{j=1}^{\alpha-1} \gamma_i^T \gamma_j \ .$$

Thus

$$\lim_{n \to \infty} S_2/n^3 = 0. \tag{2.6}$$

(iii)
$$S_3 = \sum_{i=1}^{\alpha-1} \sum_{j=\alpha}^{n} \gamma_i^T q_{ij}^{(-1)} \gamma_j .$$

For $i \le \alpha - 1$, $j \ge \alpha$, $q_{ij}^{(-1)} = 0$ so that

$$S_3 = 0 \tag{2.7}$$

(iv)
$$S_4 = \sum_{i=\alpha}^{n} \sum_{j=1}^{\alpha-1} \gamma_i^T q_{ij}^{(-1)} \gamma_j .$$

As for S_3,

$$S_4 = 0. \tag{2.8}$$

Combining (2.5)–(2.8), we conclude that

$$\lim_{n \to \infty} \frac{n^3 \operatorname{Var}(m_{\mathrm{BLU}})}{12} = \frac{1}{|B|^2} .$$

Noting that the spectral density of the process defined in (2.4) is

$$f(\lambda) = \frac{1}{2\pi} \frac{|1 - e^{i\lambda}|^2}{\left| \sum_{0}^{\alpha} b_k e^{ik\lambda} \right|^2} = \frac{1}{2\pi} |1 - e^{i\lambda}|^2 \, g(\lambda),$$

we see that $1/|B|^2 = g(0)$ so that by Theorem 2.4, m_p is asymptotically efficient.

Now we suppose that $f(\lambda) \in Y$ is of the general form $(2\pi)^{-1} |1 - e^{i\lambda}|^2 g(\lambda)$. Given any $\delta > 0$ it is possible to choose two spectral densities of the type (2.4), $f_1(\lambda) = (2\pi)^{-1} |1 - e^{i\lambda}|^2 g_1(\lambda)$ and $f_2(\lambda) = (2\pi)^{-1} |1 - e^{i\lambda}|^2 g_2(\lambda)$, such that

$$0 \le g(\lambda) - g_1(\lambda) < \delta, \tag{2.9}$$

$$0 \le g_2(\lambda) - g(\lambda) < \delta. \tag{2.10}$$

Expanding notation a bit, we let

$$\operatorname{Var}(m, f) = c^T R c = \int_{-\pi}^{+\pi} |\sum c_k e^{ik\lambda}|^2 \, f(\lambda) \, d\lambda, \tag{2.11}$$

and m_{BLU}, $m_{\mathrm{BLU}}^{(1)}$, $m_{\mathrm{BLU}}^{(2)}$ be the BLU estimates under the hypotheses $f(\lambda)$, $f_1(\lambda)$, and $f_2(\lambda)$ respectively. We have then evidently

$$\operatorname{Var}(m_{\mathrm{BLU}}^{(1)}, f_1) \le \operatorname{Var}(m_{\mathrm{BLU}}, f_1) \le \operatorname{Var}(m_{\mathrm{BLU}}, f) \le \operatorname{Var}(m_{\mathrm{BLU}}^{(2)}, f) \tag{2.12}$$

$$\le \operatorname{Var}(m_{\mathrm{BLU}}^{(2)}, f_2),$$

where the first and third inequalities follow from the minimal nature of the BLU estimate and the second and fourth inequalities follow from (2.11) and the fact that (2.9) and (2.10) imply $f_1(\lambda) \le f(\lambda) \le f_2(\lambda)$. The first part of the proof of this theorem showed that $n^3 \operatorname{Var}(m_{\mathrm{BLU}}^{(i)}, f_i)/12 \to g_i(0)$, $i = 1, 2$. Hence

$$g_1(0) \le \underline{\lim} \frac{n^3 \operatorname{Var}(m_{\mathrm{BLU}}, f)}{12} \le \overline{\lim} \frac{n^3 \operatorname{Var}(m_{\mathrm{BLU}}, f)}{12} \le g_2(0).$$

Since $\delta > 0$ is arbitrary in (2.9) and (2.10), we have

$$n^3 \operatorname{Var}(m_{\mathrm{BLU}}, f)/12 \to g(0)$$

and the theorem is proved.

An interesting consequence is the following asymptotic property of **the least squares** estimate.

COROLLARY.   *Let* $f(\lambda) \in Y$. *Then*

$$n \cdot \frac{\operatorname{Var}(m_{\mathrm{BLU}})}{\operatorname{Var}(m_{\mathrm{LS}})} \to \frac{12\pi g(0)}{\int_{-\pi}^{+\pi} g(\lambda)\, d\lambda} \ .$$

Proof.   Theorem 2.5 showed that $n^3 \operatorname{Var}(m_{\mathrm{BLU}}) \to 12 g(0)$. A generalization of Theorem 2.2 (requiring only a routine application of the Riemann–Lebesgue lemma) shows that

$$n^2 \operatorname{Var}(m_{\mathrm{LS}}) \to \frac{1}{\pi} \int_{-\pi}^{+\pi} g(\lambda)\, d\lambda$$

and the result follows.

3. Discussion.   We have shown that for the class Y of spectral densities the parabolic estimate is superior to the least squares estimate. It requires somewhat more computation but not a prohibitive amount. Our opinion is that in actual practice the parabolic estimate should be used when $f(0)$ is known to be small and possibly equal to zero. If, in fact, $f(0) > 0$, the following theorem shows that the results cannot be too bad.

THEOREM 3.1.   *Let* $f(\lambda)$ *satisfy all the conditions of* Y *except that* $f(0) > 0$. *Then the asymptotic efficiency of the parabolic estimate is* 5/6.

Proof.   First, we recall the fact that

$$\frac{n}{2\pi} \operatorname{Var}(m_{\mathrm{BLU}}) \to f(0) \tag{3.1}$$

(see, for instance, Grenander and Rosenblatt [2]). We have

$$\operatorname{Var}(m_p) = \int_{-\pi}^{+\pi} |p_n(e^{i\lambda})|^2\, f(\lambda)\, d\lambda, \tag{3.2}$$

where again

$$p_n(z) = \frac{6n}{n^2 - 1} \sum_{1}^{n} \frac{k}{n}\left(1 - \frac{k}{n}\right) z^k.$$

Let $l_n(\lambda) = (5/12\pi)n\, |p_n(e^{i\lambda})|^2$ so that

$$\frac{5n \operatorname{Var}(m_p)}{12\pi} = \int_{-\pi}^{+\pi} l_n(\lambda) f(\lambda)\, d\lambda. \tag{3.3}$$

We now use a singular integral argument similar to that of Theorem 2.4. From there we recall that

$$|p_n(e^{i\lambda})| \le 24n/[(n^2 - 1)|\,1 - e^{i\lambda}|^3],$$

so that

$$|(l_n(\lambda))^{1/2}| \le \left(\frac{5}{12\pi}\right)^{1/2} \frac{24 n^{3/2}}{(n^2 - 1)\,|1 - e^{i\lambda}|^3} ,$$

and hence $l_n(\lambda) \to 0$ uniformly in λ outside any neighborhood $(-\epsilon, \epsilon)$. The second step is to show that

$$\int_{-\pi}^{+\pi} l_n(\lambda) \, d\lambda \to 1.\tag{3.4}$$

It is easily seen that

$$\int_{-\pi}^{+\pi} l_n(\lambda) \, d\lambda = \frac{30n^3}{(n^2 - 1)^2} \sum_{k=1}^{n} \left(\frac{k}{n}\right)^2 \left(1 - \frac{k}{n}\right)^2 \sim 30\left[\frac{1}{n} \sum_{k=1}^{n} \left(\frac{k}{n}\right)^2 \left(1 - \frac{k}{n}\right)^2\right].$$

The normalized sum is a Riemann sum which converges to

$$\int_0^1 x^2(1 - x)^2 \, dx = \frac{1}{30}.$$

Hence, we have shown (3.4) and conclude that

$$\int_{-\pi}^{+\pi} l_n(\lambda) f(\lambda) \, d\lambda \to f(0);$$

or using (3.3),

$$5n \, \text{Var} \, (m_p)/12\pi \to f(0).\tag{3.5}$$

Combining (3.1) and (3.5), we have the desired result.

Figs. 3.1 and 3.2 illustrate this result for two spectral densities introduced before,

$$f(\lambda) = (2\pi)^{-1}(1 - .1 \cos \lambda) \quad \text{and} \quad f(\lambda) = (2\pi)^{-1}(1 - .9 \cos \lambda)$$
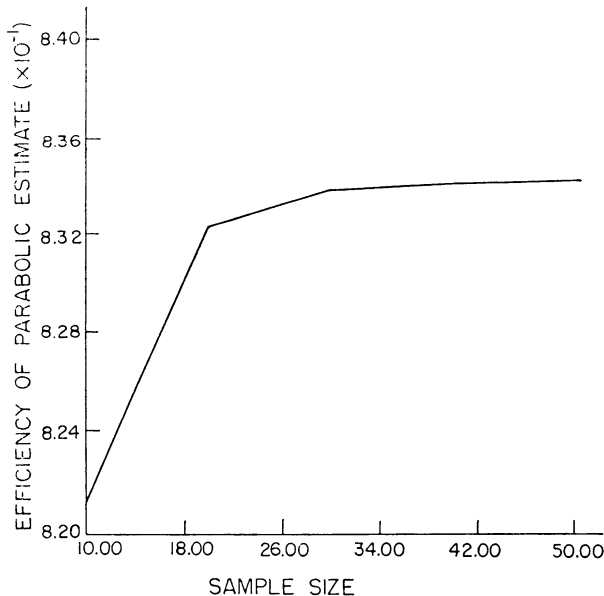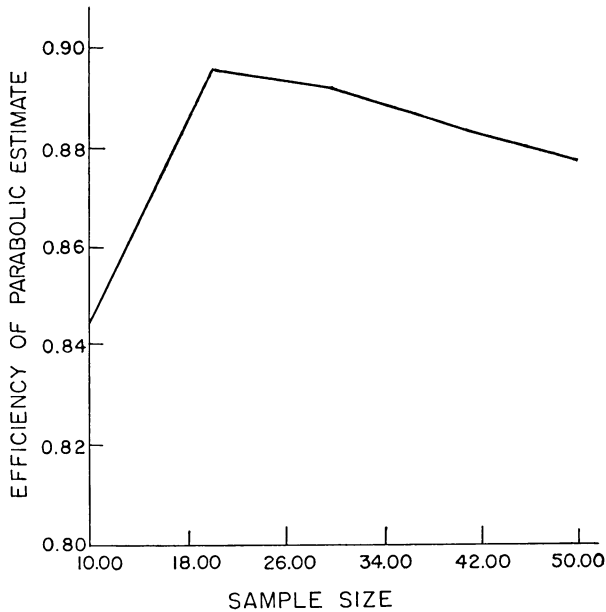
respectively.

FIG. 3.1

FIG. 3.2

We conclude with a conjecture in which there is certain experimental justification and for which a proof by induction seems reasonable.

Conjecture. Let Y_r be the class of spectral densities which are even, continuous, and positive except for a $2r$th order zero at the origin. Then the BLU estimate associated with the spectral density $(2\pi)^{-1} |1 - e^{i\lambda}|^{2r}$ is asymptotically efficient over Y_r.

REFERENCES

[1] U. Grenander, *On the estimation of regression coefficients in an autocorrelated disturbance*, Ann. Math. Statist. 25 (1954)
[2] U. Grenander and M. Rosenblatt, *Statistical analysis of stationary time series*, Wiley, New York, 1957
[3] U. Grenander and G. Szegö, *Toeplitz forms and their applications*, Univ. of California Press, Berkeley, Calif., 1958
[4] E. J. Hannan. *Time series analysis*, Methuen, London, 1967