The Annals of Statistics
1974, Vol. 2, No. 6, 1095–1107

ON LARGE-SAMPLE ESTIMATION FOR THE MEAN OF A STATIONARY RANDOM SEQUENCE

BY ROLF K. ADENSTEDT

Lehigh University

For a wide class of stationary random sequences possessing a spectral density function, the variance of the best linear unbiased estimator for the mean is seen to depend asymptotically only on the behavior of the spectral density near the origin. Asymptotically efficient estimators based only on this behavior may be chosen. For spectral densities behaving like λ^ν at the origin, $\nu > -1$ a constant, the minimum variance decreases like $n^{-\nu-1}$, where n is the sample size. Asymptotically efficient estimators depending on ν are given. Finally, the consequences of over- or under-estimating the value of ν in choosing an estimator are considered.

1. Introduction. The following simple stochastic model is widely applied. A (possibly complex-valued) random sequence $X_t = m + Y_t$, $t = 0, \pm 1, \cdots$ is the sum of a constant mean value m and a disturbance. The disturbance $\{Y_t\}$ is a zero-mean, wide-sense stationary random sequence with a spectral density function $f(\lambda)$, $-\pi \leqq \lambda \leqq \pi$, so that its covariance function is

$$R(\tau) = EY_{t+\tau}\bar{Y}_t = \int_{-\pi}^{\pi} e^{i\tau\lambda} f(\lambda)\, d\lambda\,, \qquad \tau = 0, \pm 1, \cdots. \tag{1.1}$$

We consider in this paper aspects of the estimation of the mean m for this model by unbiased linear estimators

$$\hat{m} = c_0 X_0 + c_1 X_1 + \cdots + c_n X_n\,, \qquad c_0 + c_1 + \cdots + c_n = 1\,, \tag{1.2}$$

formed from a large number of observed values $X_0, X_1, \cdots, X_n$. Of particular interest is the best linear unbiased estimator (BLUE) $\hat{m}_{\mathrm{BLU}}$, i.e., the estimator (1.2) having minimum variance

$$\operatorname{Var} \hat{m} = E|\hat{m} - m|^2 = \sum_{j,k=0}^{n} c_j \bar{c}_k R(j-k)\,.$$

Since, typically, calculation of $\hat{m}_{\mathrm{BLU}}$ or its variance is difficult, adequate approximations are needed in terms of more easily calculated estimators.

There is a substantial literature comparing the BLUE with other estimators, especially the least squares estimator (LSE) $\hat{m}_{\mathrm{LS}} = (n+1)^{-1}(X_0 + \cdots + X_n)$. A good survey of results appears in Anderson (1971). Watson (1967, 1972) and Zyskind (1967) consider coincidence of the BLUE and LSE for coefficients in general linear regression models with a fixed sample size. Kruskal (1968) treats this problem from a coordinate-free Hilbert space viewpoint.

A general Hilbert space approach to linear estimation problems is given by

Received November 1972; revised November 1973.

AMS 1970 *subject classifications.* Primary 62; Secondary 55.

Key words and phrases. Linear estimation, stationary time series, mean, spectral density, efficiency.

Parzen (1961) and in the infinite sample size case by Rozanov (1971). Cleveland (1971) compares the BLUE to an arbitrary linear estimator using Mahalanobis distance.

Studies have also been made with special cases of our model. Fishman (1972) gives closed form expressions for $\hat{m}_{\mathrm{BLU}}$ and its variance when $\{Y_t\}$ is autoregressive of finite order. For order two he compares the BLUE and LSE with regard to small and large sample behavior. The first-order case is treated in detail by Chipman, *et al.* (1968).

We are concerned here with asymptotic or large-sample results and will compare $\hat{m}_{\mathrm{BLU}}$ to competitors other than $\hat{m}_{\mathrm{LS}}$. Grenander (1950) considered asymptotic efficiency of the LSE relative to the BLUE in estimating means and (1954) extended the results to more general regressions. In terms of our model he showed that, as $n \to \infty$, both $\hat{m}_{\mathrm{BLU}}$ and $\hat{m}_{\mathrm{LS}}$ have asymptotic variance $2\pi f(0)/n$ as long as $f(\lambda)$ is positive and continuous. More recently Vitale (1973) considered the case when $f(\lambda)$ is continuous and positive except at the origin, where $\lambda^{-2}f(\lambda) \to L > 0$ as $\lambda \to 0$. Then $\operatorname{Var} \hat{m}_{\mathrm{BLU}} \sim 24\pi L/n^3$ and the estimator (1.2) with coefficients $c_k = 6k(n-k)/[n(n^2-1)]$ is asymptotically efficient, while $\hat{m}_{\mathrm{LS}}$ is not.

Related results have appeared in the applied mathematics literature on smoothing. Greville (1966) and Trench (1971) consider the minimum variance linear estimation of a polynomial $m(t)$ of fixed degree in the model $X_t = m(t) + Y_t$. Trench gives an algorithm for computing the BLUE when $\{Y_t\}$ has a continuous spectral density $f(\lambda)$; while Greville gives a closed form for the BLUE when $f(\lambda) = (\sin^2 \lambda/2)^\alpha$ with α an integer. Both authors are concerned, however, with stability of minimum variance smoothing formulas, rather than asymptotic properties thereof. Trench (1973) also considers estimation of a linear functional of $m(t)$.

The main results of this paper appear in Sections 4 and 5. Sections 2 and 3 contain preliminaries. In Section 4 (Theorem 4.1) we find that, for a large class of spectral densities, the asymptotic form of $\operatorname{Var} \hat{m}_{\mathrm{BLU}}$ is determined solely by the behavior of $f(\lambda)$ near $\lambda = 0$, and that asymptotically efficient estimators based only on this behavior are available. In Section 5 we apply this result to the case of a spectral density $f(\lambda)$ which behaves like λ^ν at the origin, where $\nu > -1$ is any constant. We find (Theorem 5.2) that the $\operatorname{Var} \hat{m}_{\mathrm{BLU}} = O(n^{-\nu-1})$ as $n \to \infty$, and prescribe an asymptotically efficient estimator based only on the value of ν. Thereby we generalize the results of Grenander (1954) and Vitale (1973) described above.

We show in Section 6 that the generating functions associated with the estimators of Section 5 are expressible in terms of Gegenbauer polynomials, then use known properties of these polynomials to strengthen slightly the conclusions of Section 5.

In Sections 7 and 8 we consider cases of over- and under-estimating ν when $f(\lambda) = O(\lambda^\nu)$ as $\lambda \to 0$ and an estimator as prescribed in Section 5 is used. We find (Theorem 7.2) that the optimum rate of decrease of Var $\hat{m}_{\text{BLU}}$ is still attained if too large a value of ν is assumed, but that (Theorem 8.1) Var $\hat{m}_{\text{BLU}}$ decreases at a slower rate if ν is assumed to be smaller than the true value.

2. Uniqueness of the BLUE. From (1.1) it is not difficult to see that the covariance matrix $R = \{R(j-k); j, k = 0, 1, \cdots, n\}$ of $Y_0, \cdots, Y_n$ is positive definite as long as $R(0) = \int_{-\pi}^{\pi} f(\lambda)\, d\lambda > 0$, which we always assume. Then, as is well known, $\hat{m}_{\text{BLU}}$ has a unique representation (1.2). In fact, the row vector c of coefficients in $\hat{m}_{\text{BLU}}$ is given by

$$(2.1) \qquad c = (c_0 \cdots c_n) = (u^T R^{-1} u)^{-1} u^T R^{-1}, \qquad u = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

with

$$(2.2) \qquad \operatorname{Var} \hat{m}_{\text{BLU}} = (u^T R^{-1} u)^{-1} > 0 .$$

From (2.1) and (2.2) we note that $cR = u^T \operatorname{Var} \hat{m}_{\text{BLU}}$. Since R is positive definite, this relation together with $cu = 1$ uniquely determines the coefficient vector c of $\hat{m}_{\text{BLU}}$, as well as Var $\hat{m}_{\text{BLU}}$.

3. Preliminaries. Notation and definitions. It is convenient to reformulate the problem of finding the BLUE in a concise notation that easily lends itself to analysis. For a fixed spectral density $f(\lambda)$, we define an inner product by

$$(3.1) \qquad (\psi_1, \psi_2)_f = \int_{-\pi}^{\pi} \psi_1(e^{i\lambda})\overline{\psi_2(e^{i\lambda})} f(\lambda)\, d\lambda$$

on the space of those (complex-valued) functions $\psi(z)$ for which the associated norm $||\psi||_f$ is finite. Using (1.1), we may then write the variance of an estimator (1.2) as

$$(3.2) \qquad \operatorname{Var}(\hat{m}, f) = \int_{-\pi}^{\pi} |p_n(e^{i\lambda})|^2 f(\lambda)\, d\lambda = ||p_n||_f^2 ,$$

where

$$(3.3) \qquad p_n(z) = c_0 + c_1 z + \cdots + c_n z^n .$$

We employ the notation Var $(\hat{m}, f)$ to indicate dependence on $f(\lambda)$. The problem of finding $\hat{m}_{\text{BLU}}$ may now be stated as

$$(3.4) \qquad ||p_n||_f^2 = \min , \qquad p_n(1) = 1 .$$

DEFINITION 3.1. The polynomial (3.3) that solves (3.4) will be called the *optimal polynomial* (of degree n) for $f(\lambda)$. The minimum itself will be denoted by $\sigma_n^2(f)$.
Thus $||p_n||_f^2 = \sigma_n^2(f) = \operatorname{Var} \hat{m}_{\text{BLU}}$ for the optimal polynomial. Since the optimal $p_n(z)$ is determined by $\hat{m}_{\text{BLU}}$ (and conversely), it is unique.

The square norm $||\psi||_f^2$ is non-decreasing in $f(\lambda)$ for a fixed function $\psi(z)$. Consequently, we have the important property that the minimum variance $\sigma_n^2(f)$ is a non-decreasing functional of $f(\lambda)$, i.e., $\sigma_n^2(f) \geqq \sigma_n^2(g)$ when $f(\lambda) \geqq g(\lambda)$.

Clearly, $\sigma_n^2(f)$ is non-increasing in n. In the sequel we restrict attention to spectral densities for which this minimum variance does not decrease too rapidly.

DEFINITION 3.2. We call $\sigma_n^2(f)$ *slowly decreasing* if

$$\lim_{n\to\infty} \sigma^2_{n+\nu}(f)/\sigma_n^2(f) = 1\,, \qquad \nu = 1, 2, \cdots. \tag{3.5}$$

We also define several classes of functions to be considered in the following sections. All such functions are assumed to be (Lebesgue) measurable on $[-\pi, \pi]$.

DEFINITION 3.3. Let $g(\lambda)$ be defined on $[-\pi, \pi]$. We say that this function belongs to the class

L_0: if $g(\lambda)$ is nonnegative, integrable over $[-\pi, \pi]$, and continuous at $\lambda = 0$;
L_0^+: if $g(\lambda) \in L_0$ and has a positive lower bound;
B_0: if $g(\lambda)$ is nonnegative, bounded, and continuous at $\lambda = 0$;
B_0^+: if $g(\lambda) \in B_0$ and has a positive lower bound.

Furthermore, suppose that $g(\lambda)$ has the form

$$g(\lambda) = h(\lambda)|\lambda - \lambda_1|^{\alpha(1)}|\lambda - \lambda_2|^{\alpha(2)} \cdots |\lambda - \lambda_r|^{\alpha(r)}\,, \tag{3.6}$$

where r is a natural number, $\lambda_1, \lambda_2, \cdots, \lambda_r$ are constants in $[-\pi, \pi]$, and $\alpha(1), \alpha(2), \cdots, \alpha(r)$ are nonnegative constants. We say that $g(\lambda)$ is in ZL_0^+ or ZB_0^+ according as $h(\lambda)$ is in L_0^+ or B_0^+, respectively.

4. Asymptotic behavior of the minimum variance. To relate large-sample behavior of $\operatorname{Var} \hat{m}_{\mathrm{BLU}}$ to behavior of the spectral density near the origin, we study the ratio $\sigma_n^2(fg)/\sigma_n^2(f)$, where $\sigma_n^2(f)$ satisfies (3.5) and $g(\lambda)$ is suitably "nice." In essence we are now taking the spectral density of $\{Y_t\}$ to be $f(\lambda)g(\lambda)$ rather than $f(\lambda)$. The purpose of this section is to prove

THEOREM 4.1. *Let $f(\lambda)$ be a spectral density with slowly decreasing $\sigma_n^2(f)$, and let $g(\lambda)$ be in the class ZB_0^+. Then*

$$\lim_{n\to\infty} \sigma_n^2(fg)/\sigma_n^2(f) = g(0). \tag{4.1}$$

Moreover, if $g(0) > 0$, then the BLUE *$\hat{m}_f$ calculated under the hypothesis that the spectral density is $f(\lambda)$ is asymptotically efficient with respect to $f(\lambda)g(\lambda)$ in the sense that*

$$\lim_{n\to\infty} \sigma_n^2(fg)/\operatorname{Var}(\hat{m}_f, fg) = 1\,. \tag{4.2}$$

Before proceeding with the proof, we obtain some preliminary results. In the remainder of this section, we adhere to the following notation of the theorem: $f(\lambda)$ will denote a spectral density with slowly decreasing $\sigma_n^2(f)$ and $\hat{m}_f$ will denote the BLUE calculated with respect to $f(\lambda)$. Moreover, the optimal polynomials for $f(\lambda)$ will always be denoted by $p_n(z)$.

LEMMA 4.1. *If $t(\lambda)$ is a nonnegative trigonometric polynomial, then*

$$\liminf_{n\to\infty} \sigma_n^2(ft)/\sigma_n^2(f) \geqq t(0)\,.$$

PROOF. The result is trivial if $t(0) = 0$; therefore we take $t(0) > 0$. By the Féjer–Riesz representation theorem, there is a polynomial $p(z)$ of the same degree, say ν, as $t(\lambda)$ for which $t(\lambda) = |p(e^{i\lambda})|^2$. Letting $q_0(z), q_1(z), \cdots$ be the optimal polynomials for $f(\lambda)t(\lambda)$, we define the polynomial $r(z) = p(z)q_n(z)/p(1)$ of degree $n + \nu$, and note that $r(1) = 1$. Therefore

$$t(0)\sigma_{n+\nu}^2(f) \leqq t(0)||r||_f^2 = ||q_n||_{ft}^2 = \sigma_n^2(ft) .$$

The proof is completed upon dividing the extreme inequality by $\sigma_n^2(f)$ and taking lim inf, with use of (3.5).

LEMMA 4.2. *$K_n(\lambda) = |p_n(e^{i\lambda})|^2 f(\lambda)/\sigma_n^2(f)$ satisfies, for any $0 < \delta \leqq \pi$,*

$$\lim_{n\to\infty} \int_{\delta \leqq |\lambda| \leqq \pi} K_n(\lambda)\, d\lambda = 0 . \tag{4.3}$$

PROOF. Let $g(\lambda) = 2$ for $0 \leqq |\lambda| < \delta$ and $g(\lambda) = 1$ for $\delta \leqq |\lambda| \leqq \pi$. Then the integral in (4.3) is just $2 - ||p_n||_{fg}^2/\sigma_n^2(f)$, and it suffices to show that

$$\lim_{n\to\infty} ||p_n||_{fg}^2/\sigma_n^2(f) = 2 . \tag{4.4}$$

Now clearly we can choose a nonnegative trigonometric polynomial $t(\lambda)$ with the properties: $t(\lambda) \leqq g(\lambda)$ and $t(0) = 2$. Then $||p_n||_{fg}^2 \geqq \sigma_n^2(fg) \geqq \sigma_n^2(ft)$. Dividing by $\sigma_n^2(f)$ and taking lim inf, with use of Lemma 4.1 we obtain $\liminf_{n\to\infty} ||p_n||_{fg}^2/\sigma_n^2(f) \geqq t(0) = 2$. But clearly also $||p_n||_{fg}^2/\sigma_n^2(f) \leqq 2$. Thus (4.4) is established and the proof is complete.

LEMMA 4.3. *If $g(\lambda)$ is in B_0, then*

$$\lim_{n\to\infty} \operatorname{Var}(\hat{m}_f, fg)/\sigma_n^2(f) = g(0) .$$

PROOF. With $K_n(\lambda)$ defined as in Lemma 4.2, the assertion is that $\int_{-\pi}^{\pi} K_n(\lambda)g(\lambda)\, d\lambda \to g(0)$ as $n \to \infty$. This follows from a standard integral kernel argument since $K_n(\lambda)$ is nonnegative and satisfies $\int_{-\pi}^{\pi} K_n(\lambda)\, d\lambda = 1$ as well as (4.3).

LEMMA 4.4. *If $g(\lambda)$ is in B_0, then*

$$\limsup_{n\to\infty} \sigma_n^2(fg)/\sigma_n^2(f) \leqq g(0) . \tag{4.5}$$

PROOF. The result is immediate from Lemma 4.3, since $\sigma_n^2(fg) \leqq \operatorname{Var}(\hat{m}_f, fg)$.

LEMMA 4.5. *Let $g(\lambda)$ be in ZL_0^+. Then for any $\varepsilon > 0$ there is a trigonometric polynomial $t(\lambda)$ with the properties: $0 \leqq t(\lambda) \leqq g(\lambda)$ and $t(0) \geqq g(0) - \varepsilon$.*

PROOF. If suffices to prove the result for each of the factors on the right side of (3.6), for the product of the "approximants" will yield an approximant for $g(\lambda)$ in the required sense. We consider the individual factors.

(a) $g(\lambda) \in L_0^+$. For $\varepsilon < g(0)$, we choose $0 < \delta \leqq \pi$ such that $g(\lambda) \geqq g(0) - \varepsilon$ for $|\lambda| \leqq \delta$. Then we may use $t(\lambda) = [g(0) - \varepsilon][\frac{1}{2}(1 + \cos\lambda)]^N$, where the integer N is chosen so large that $[\frac{1}{2}(1 + \cos\delta)]^N \leqq \inf g(\lambda)/[g(0) - \varepsilon]$. For $\varepsilon \geqq g(0)$ we use $t(\lambda) \equiv 0$.

(b) $g(\lambda) = |\lambda - \lambda_0|^{2\nu}$, ν a positive integer. The result follows from (a) by writing $g(\lambda)$ as the product of a positive and continuous function and the trigonometric polynomial $[1 - \cos(\lambda - \lambda_0)]^\nu$.

(c) $g(\lambda) = |\lambda - \lambda_0|^\alpha$, $\alpha > 0$ not an even integer. If $\lambda_0 = 0$ we may use $t(\lambda) \equiv 0$, so we take $\lambda_0 \neq 0$. For any integer $\nu > \alpha/2$ and constant $0 < \delta < |\lambda_0|$, $h(\lambda) = [\max\{\delta, |\lambda - \lambda_0|\}]^{-2\nu+\alpha}$ is positive and continuous. Thus $g_1(\lambda) = h(\lambda)|\lambda - \lambda_0|^{2\nu}$ may be approximated according to (a) and (b). But $g_1(\lambda) \leqq g(\lambda)$ and $g_1(0) = g(0)$, so the approximant for $g_1(\lambda)$ also serves for $g(\lambda)$.

With these preliminary results, the proof of the theorem stated at the beginning of this section is quite short.

PROOF OF THEOREM 4.1. Any function $g(\lambda)$ in ZB_0^+ is in ZL_0^+ and in B_0, hence satisfies (4.5) To prove (4.1) it suffices thus to show that

$$\liminf_{n\to\infty} \sigma_n^2(fg)/\sigma_n^2(f) \geqq g(0) . \tag{4.6}$$

For any $\varepsilon > 0$, choosing a trigonometric polynomial $t(\lambda)$ with the properties in Lemma 4.5, we have $\sigma_n^2(fg) \geqq \sigma_n^2(ft)$, and so with use of Lemma 4.1, the left side of (4.6) is seen to be bounded below by $\liminf_{n\to\infty} \sigma_n^2(ft)/\sigma_n^2(f) \geqq t(0) \geqq g(0) - \varepsilon$. Letting $\varepsilon \to 0$, we obtain (4.6). The second assertion (4.2) follows immediately from (4.1) and Lemma 4.3.

5. Application to certain spectral densities. Theorem 4.1 enables us to obtain the rate of decrease of Var $\hat{m}_{\text{BLU}}$ and asymptotically efficient estimators for large classes of spectral densities, in particular for many spectral densities characterized by a zero (or infinity) of fixed finite order at the origin. As representatives of such spectral densities we take

$$f_\alpha(\lambda) = (2\pi)^{-1}|1 - e^{i\lambda}|^{2\alpha} = 2^{2\alpha-1}\pi^{-1}(\sin^2 \lambda/2)^\alpha . \tag{5.1}$$

Here α (not necessarily an integer) is a constant, with $\alpha > -\frac{1}{2}$ for integrability. The main problem is in obtaining the BLUE for $f_\alpha(\lambda)$.

LEMMA 5.1. *The covariance function corresponding to the spectral density* (5.1) *is*

$$R_\alpha(\tau) = (-1)^\tau \frac{\Gamma(2\alpha + 1)}{\Gamma(\alpha + \tau + 1)\Gamma(\alpha - \tau + 1)}, \qquad \tau = 0, \pm 1, \cdots . \tag{5.2}$$

We take $1/\Gamma(z) = 0$ *for* z *a nonpositive integer.*

PROOF. Since $f_\alpha(\lambda)$ is even, by making the variable change $\lambda = \pi - 2\theta$ in $\int_{-\pi}^{\pi} e^{i\tau\lambda} f_\alpha(\lambda)\, d\lambda$ we derive

$$R_\alpha(\tau) = (-1)^\tau 2^{2\alpha+1}\pi^{-1} \int_0^{\pi/2} \cos 2\tau\theta \cos^{2\alpha}\theta\, d\theta .$$

Since the above integral has value $\pi\Gamma(2\alpha + 1)/[2^{2\alpha+1}\Gamma(\alpha + \tau + 1)\Gamma(\alpha - \tau + 1)]$ as given by Erdélyi (1953), Volume 1, page 12, we obtain (5.2).

The next two combinatorial lemmas are presented in a more general form, needed in Section 7, than presently required. $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p + q)$ represents the beta function.

LEMMA 5.2. *For any constants* $a, b > -1$,

$$\sum_{k=0}^{n} \binom{n}{k} B(a+k+1, b+n-k+1) = B(a+1, b+1) . \tag{5.3}$$

PROOF. With both sides expressed in terms of beta integrals, the identity is obvious.

LEMMA 5.3. *For* $\alpha > -\frac{1}{2}$ *and* β *a nonnegative integer,*

$$\begin{aligned} S_j(n, \alpha, \beta) & \\ &\equiv (-1)^j \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{\Gamma(\alpha+\beta+k+1)\Gamma(\alpha+\beta+n-k+1)}{\Gamma(\alpha+k-j+1)\Gamma(\alpha-k+j+1)} \\ &= (-1)^\beta n! \sum_{\nu=0}^{2\beta} (-1)^\nu \\ &\quad \times \binom{j+\beta}{\nu}\binom{n-j+\beta}{2\beta-\nu} \frac{\Gamma^2(\alpha+\beta+1)}{\Gamma(\alpha+\beta-\nu+1)\Gamma(\alpha-\beta+\nu+1)} \end{aligned} \tag{5.4}$$

for $j = 0, 1, \cdots, n$, *with* $\binom{N}{\nu}$ *taken as zero for* $\nu > N$ *and* $\nu < 0$.

PROOF. With use of $\Gamma(z+1) = z\Gamma(z)$ repeatedly and of Leibnitz's rule, the result follows upon taking the derivative $\partial^{n+2\beta}/\partial x^{j+\beta}\, \partial y^{n-j+\beta}$ of the identity

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} x^{\alpha+\beta+k} y^{\alpha+\beta+n-k} = x^{\alpha+\beta} y^{\alpha+\beta} (y-x)^n$$

and then setting $x = y = 1$.

THEOREM 5.1. *The* BLUE $\hat{m}_\alpha$ *calculated with respect to the spectral density* (5.1) *has coefficients*

$$\begin{aligned} c_k &= c_k(n, \alpha) \\ &= \binom{n}{k} \frac{B(\alpha+k+1, \alpha+n-k+1)}{B(\alpha+1, \alpha+1)}, \qquad k = 0, 1, \cdots, n, \end{aligned} \tag{5.5}$$

and variance

$$\sigma_n^2(f_\alpha) = \mathrm{Var}\,(\hat{m}_\alpha, f_\alpha) = B(n+1, 2\alpha+1)/B(\alpha+1, \alpha+1) . \tag{5.6}$$

PROOF. As noted at the end of Section 2, it suffices to show that

$$\sum_{k=0}^{n} c_k(n, \alpha) = 1 \tag{5.7}$$

and

$$\sum_{k=0}^{n} c_k(n, \alpha) R_\alpha(k-j) = \sigma_n^2(f_\alpha) , \qquad j = 0, 1, \cdots, n , \tag{5.8}$$

with $\sigma_n^2(f_\alpha)$ given by (5.6). (5.7) follows directly from Lemma 5.2 with $a = b = \alpha$. As for (5.8), using (5.2), (5.5) and (5.6) we may reduce the assertion to (5.4) with $\beta = 0$.

A straightforward application of Stirling's formula for the gamma function yields, from (5.6), a

COROLLARY. *As* $n \to \infty$,

$$\sigma_n^2(f_\alpha) \sim n^{-2\alpha-1}\Gamma(2\alpha+1)/B(\alpha+1, \alpha+1) . \tag{5.9}$$

From (5.9) we conclude that $\sigma_n^2(f_\alpha)$ is slowly decreasing, and may immediately apply Theorem 4.1. We state the result in terms of the original model in Section 1.

THEOREM 5.2. *Let the disturbance* $\{Y_t\}$ *have spectral density function* $f(\lambda) = f_\alpha(\lambda)g(\lambda)$, *where* $\alpha > -\frac{1}{2}$ *and* $g(\lambda) \in ZB_0^+$ (Definition 3.3) *with* $g(0) > 0$. *Then*

$$\operatorname{Var} \hat{m}_{\mathrm{BLU}} \sim n^{-2\alpha-1}\Gamma(2\alpha+1)g(0)/B(\alpha+1, \alpha+1)\,, \qquad n \to \infty\,.$$

Moreover, the estimator $\hat{m}_\alpha = \sum_{k=0}^n c_k(n, \alpha)X_k$ *with coefficients given in* (5.5) *is asymptotically efficient in the sense that*

$$\operatorname{Var} \hat{m}_{\mathrm{BLU}}/\operatorname{Var} \hat{m}_\alpha \to 1\,, \qquad n \to \infty\,.$$

For $\alpha = 0$ and $\alpha = 1$ the theorem gives slightly strengthened versions of results of Grenander (1954) and Vitale (1973), respectively. The theorem establishes a conjecture of Vitale for integral α.

6. The optimal polynomials for $f_\alpha(\lambda)$. We denote by $C_n^{(a)}(x)$ the Gegenbauer (or ultraspherical) polynomials, i.e., the polynomials orthogonal on $[-1, 1]$ with respect to the weight function $(1-x^2)^{a-1/2}$ and with the standardization $C_n^{(a)}(1) = \Gamma(n+2a)/[n!\,\Gamma(2a)]$ for $a \neq 0$. We shall relate the optimal polynomials for $f_\alpha(\lambda)$ to the $C_n^{(a)}(x)$ and use the relation to establish a lemma required in Sections 7 and 8. This lemma will also enable us to strengthen Theorem 5.2 slightly. Without further reference we make use of well-known properties of the $C_n^{(a)}(x)$, as given by Erdélyi (1953), Volume 2.

LEMMA 6.1. *The optimal polynomials*

$$p_{n,\alpha}(z) = \textstyle\sum_{k=0}^n c_k(n, \alpha)z^k$$

for $f_\alpha(\lambda)$ *satisfy*

$$p_{n,\alpha}(e^{2i\theta}) = e^{in\theta}C_n^{(\alpha+1)}(\cos\theta)/C_n^{(\alpha+1)}(1)\,. \tag{6.1}$$

PROOF. In terms of the coefficients (5.5), it is known that

$$C_n^{(\alpha+1)}(\cos\theta)/C_n^{(\alpha+1)}(1) = \textstyle\sum_{k=0}^n c_k(n, \alpha)\cos(n-2k)\theta\,.$$

The right side, however, is easily seen to be $e^{-in\theta}p_{n,\alpha}(e^{2i\theta})$, since the $c_k(n, \alpha) = c_{n-k}(n, \alpha)$ and are real.

With use of the relation between the Gegenbauer polynomials and the hypergeometric function $F(a, b; c; x)$, and a quadratic transformation on the hypergeometric function, we can rewrite (6.1) as

$$p_{n,\alpha}(e^{2i\theta}) = e^{in\theta}F(-n/2, n/2+\alpha+1; \alpha+3/2; \sin^2\theta)\,.$$

This result is given essentially by Greville (1966), in a result attributed to Sheppard (1913), for α an integer and n even.

LEMMA 6.2. *For* $\delta > 0$, $n^{\alpha+1}\max_{\delta \le |\lambda| \le \pi}|p_{n,\alpha}(e^{i\lambda})|$ *is bounded uniformly in n.*

PROOF. Since $C_n^{(\alpha+1)}(1) \sim n^{2\alpha+1}/\Gamma(2\alpha+2)$ as $n \to \infty$, and because of (6.1), it suffices to show that $n^{-\alpha}\max_{\varepsilon \le |\theta| \le \pi/2}|C_n^{(\alpha+1)}(\cos\theta)|$ is bounded uniformly in n

when $\varepsilon > 0$. This follows for $-1 < \alpha < 0$ from the inequality

$$(\sin\theta)^{\alpha+1}|C_n^{(\alpha+1)}(\cos\theta)| < (n/2)^\alpha/\Gamma(\alpha+1)\,, \qquad 0 \leqq \theta \leqq \pi\,,$$

and for $\alpha = 0$ from $C_n^{(1)}(\cos\theta) = \sin(n+1)\theta/\sin\theta$. The desideratum may then be obtained inductively for all $\alpha > 0$ from the recurrence formula

$$2\alpha(1-x^2)C_n^{(\alpha+1)}(x) = (2\alpha+n)C_n^{(\alpha)}(x) - (n+1)xC_{n+1}^{(\alpha)}(x)\,.$$

THEOREM 6.1. *The conclusions of Theorem 5.2 remain valid if the condition therein that $g(\lambda) \in ZB_0^+$ is replaced by the weaker condition that $g(\lambda) \in ZL_0^+$.*

PROOF. Theorem 5.2 followed directly from Theorem 4.1. In the proof of Theorem 4.1 boundedness of $g(\lambda)$ entered only through Lemmas 4.3 and 4.4, and these lemmas would be true for unbounded $g(\lambda)$ (in L_0) if (4.3) could be replaced by the stronger condition that

$$\lim_{n\to\infty} \max_{\delta \leq |\lambda| \leq \pi} K_n(\lambda) = 0\,, \qquad \delta > 0\,. \tag{6.2}$$

Since our current $K_n(\lambda) = |p_{n,\alpha}(e^{i\lambda})|^2 f_\alpha(\lambda)/\sigma_n^2(f_\alpha)$ satisfies (6.2), as is seen from Lemma 6.2 and (5.9), the result follows.

7. Overestimating the zero order. In applying the estimator $\hat{m}_\alpha$ of Theorem 5.2 for a spectral density that vanishes at the origin, we may guess incorrectly the true order α of the zero. In this section we consider certain cases when our estimate is too high; the opposite case is studied in Section 8.

We consider initially the following situation: the true spectral density is $f_\alpha(\lambda)$ but we use the estimator $\hat{m}_{\alpha+\beta}$ having coefficients $c_k(n, \alpha+\beta)$. β is restricted to the nonnegative integers; the case of non-integral β remains open. We define the *asymptotic efficiency* (if it exists) of $\hat{m}_{\alpha+\beta}$ relative to $\hat{m}_\alpha$ for the spectral density $f_\alpha(\lambda)$ by

$$e(\alpha, \beta) = \lim_{n\to\infty} \sigma_n^2(f_\alpha)/\mathrm{Var}\,(\hat{m}_{\alpha+\beta}, f_\alpha)\,.$$

With use of (5.2) and (5.5) we may write

$$\begin{aligned} \mathrm{Var}\,(\hat{m}_{\alpha+\beta}, f_\alpha) &= \textstyle\sum_{j,k=0}^n c_j(n, \alpha+\beta)c_k(n, \alpha+\beta)R_\alpha(j-k) \\ &= \frac{\Gamma(2\alpha+1)}{B(\alpha+\beta+1, \alpha+\beta+1)\Gamma(2\alpha+2\beta+n+2)} \\ &\quad \times \textstyle\sum_{j=0}^n c_j(n, \alpha+\beta)S_j(n, \alpha, \beta)\,, \end{aligned} \tag{7.1}$$

where the $S_j(n, \alpha, \beta)$ are given in (5.4). We introduce the notation $\langle x\rangle_r = x(x-1)\cdots(x-r+1)$ for r a nonnegative integer, with $\langle x\rangle_0 = 1$. Writing

$$\binom{j+\beta}{\nu}\binom{n-j+\beta}{2\beta-\nu} = \textstyle\sum_{r=0}^{\nu}\sum_{s=0}^{2\beta-\nu} a_{r,s}\langle j\rangle_r\langle n-j\rangle_s$$

and employing Lemma 5.2, we find that for $\nu = 0, 1, \cdots, 2\beta$ and $n \geqq 2\beta$,

$$\begin{aligned} &\textstyle\sum_{j=0}^n \binom{n}{j}\binom{j+\beta}{\nu}\binom{n-j+\beta}{2\beta-\nu} B(\alpha+\beta+j+1, \alpha+\beta+n-j+1) \\ &\quad = \textstyle\sum_{r,s} a_{r,s}\langle n\rangle_{r+s} \sum_{j=0}^{n-r-s}\binom{n-r-s}{j} \\ &\qquad\quad \times B(\alpha+\beta+j+r+1, \alpha+\beta+n-j-r+1) \\ &\quad \sim a_{\nu,2\beta-\nu}\, n^{2\beta} B(\alpha+\beta+\nu+1, \alpha+3\beta-\nu+1) \end{aligned}$$

as $n \to \infty$. Since $a_{\nu,2\beta-\nu} = [\nu!\,(2\beta-\nu)!]^{-1}$, with (5.4) and (5.5) we therefore obtain

$$\lim_{n\to\infty} (n^{2\beta} n!)^{-1} \textstyle\sum_{j=0}^{n} c_j(n, \alpha+\beta) S_j(n, \alpha, \beta) = S_\beta(2\beta, \alpha, \beta)\Gamma(2\alpha+2\beta+2)/[(2\beta)!\,\Gamma(2\alpha+4\beta+2)]\,. \tag{7.2}$$

Referring to the proof of Lemma 5.3, we see however that

$$S_\beta(2\beta, \alpha, \beta) = (-1)^\beta \frac{\partial^{4\beta}}{\partial x^{2\beta}\,\partial y^{2\beta}} x^{\alpha+\beta} y^{\alpha+\beta}(y-x)^{2\beta}\Big|_{x=y=1}$$

is a polynomial of degree 2β in α that vanishes for $\alpha = -1, -2, \cdots, -2\beta$ and has value $[(2\beta)!]^2\binom{2\beta}{\beta}$ at $\alpha = 0$, so that

$$S_\beta(2\beta, \alpha, \beta) = (2\beta)!\,\binom{2\beta}{\beta}\Gamma(\alpha+2\beta+1)/\Gamma(\alpha+1)\,. \tag{7.3}$$

Combining (7.1), (7.2) and (7.3), we arrive at an asymptotic expression for $\operatorname{Var}(\hat{m}_{\alpha+\beta}, f_\alpha)$. Using (5.9) and asymptotic expressions for those gamma functions with argument involving n, we arrive after a bit of algebra at

THEOREM 7.1. *For $\alpha > -\frac{1}{2}$ and $\beta \geqq 0$ an integer,*

$$e(\alpha, \beta) = \frac{\Gamma(2\alpha+2)\Gamma^2(\alpha+\beta+1)\Gamma(2\alpha+4\beta+2)}{\binom{2\beta}{\beta}\Gamma(\alpha+1)\Gamma(\alpha+2\beta+1)\Gamma^2(2\alpha+2\beta+2)}\,. \tag{7.4}$$

The efficiency of $\hat{m}_{\alpha+\beta}$ for a more general spectral density $f_\alpha(\lambda)g(\lambda)$ may be treated by an integral kernel argument as used in Theorem 6.1.

THEOREM 7.2. *Let $\{Y_t\}$ have spectral density $f(\lambda) = f_\alpha(\lambda)g(\lambda)$, where $g(\lambda) \in ZL_0^+$, $g(0) > 0$, and $\alpha > -\frac{1}{2}$. Then for any integer $\beta \geqq 0$,*

$$\operatorname{Var}(\hat{m}_{\alpha+\beta}, f) \sim g(0)\sigma_n^2(f_\alpha)/e(\alpha, \beta)\,, \qquad n \to \infty \tag{7.5}$$

and $\hat{m}_{\alpha+\beta}$ has the asymptotic efficiency $e(\alpha, \beta)$ in (7.4).

PROOF. (7.5) states that $\int_{-\pi}^{\pi} K_n(\lambda)g(\lambda)\,d\lambda \to g(0)$ as $n \to \infty$, where now

$$K_n(\lambda) = e(\alpha, \beta)|p_{n,\alpha+\beta}(e^{i\lambda})|^2 f_\alpha(\lambda)/\sigma_n^2(f_\alpha)\,.$$

Since $\beta \geqq 0$, from (5.9) and Lemma 6.2 we conclude that $K_n(\lambda)$ satisfies (6.2). Also $K_n(\lambda) \geqq 0$, and $\int_{-\pi}^{\pi} K_n(\lambda)\,d\lambda \to 1$ as $n \to \infty$ by definition of $e(\alpha, \beta)$. Thus (7.5) follows. Since $\sigma_n^2(f) \sim g(0)\sigma_n^2(f_\alpha)$, $n \to \infty$, by Theorem 6.1, the second assertion is also clear.

Since $e(\alpha, \beta) > 0$, we see that overestimation still yields an estimator whose variance decreases at the optimal rate. From (7.4) we observe: (a) $e(\alpha, \beta)$ decreases to $1/\binom{2\beta}{\beta}$ as $\alpha \to \infty$, and (b) $e(\alpha, \beta) \to 1$ as $\alpha \to -\frac{1}{2}$. We note that $e(0, 1) = \frac{5}{6}$, as obtained by Vitale (1973).

8. Underestimating the zero order. We consider now the case when the estimator $\hat{m}_\alpha$, $\alpha =$ integer, is used but the true spectral density has a higher order zero at the origin. The results stand in marked contrast to those of the previous section. We shall prove, after some preliminaries, the following theorem.

THEOREM 8.1. *Let α be a nonnegative integer and let the true spectral density be $f(\lambda) = f_{\alpha+1}(\lambda)g(\lambda)$, where $g(\lambda) \in L_0$. Then*

$$\lim_{n\to\infty} n^{2\alpha+2} \operatorname{Var}(\hat{m}_\alpha, f) = [(2\alpha+1)!/\alpha!]^2 \pi^{-1} \textstyle\int_{-\pi}^{\pi} g(\lambda)\, d\lambda\,. \tag{8.1}$$

Thus, if $f(\lambda) = f_{\alpha+\beta+1}(\lambda)g(\lambda)$ with $\beta \geqq 0$, $g(\lambda) \in ZL_0^+$ and $g(0) > 0$, we see (using Theorem 6.1) that the efficiency of $\hat{m}_\alpha$ is $O(n^{-2\beta-1})$; $\hat{m}_\alpha$ is far from efficient. (8.1) is given by Vitale (1973) for $\alpha = 0$.

In the following, α remains a fixed integer and we suppress dependence on it in some cases. We note that then $R_{\alpha+1}(\tau) = (-1)^\tau \binom{2\alpha+2}{\alpha+1+\tau}$, with $R_{\alpha+1}(\tau) = 0$ for $|\tau| \geqq \alpha + 2$. We may also now write

$$c_k(n, \alpha) = (2\alpha+1)\binom{2\alpha}{\alpha} \frac{(k+1)\cdots(k+\alpha)(n-k+1)\cdots(n-k+\alpha)}{(n+1)(n+2)\cdots(n+2\alpha+1)}\,,$$

which defines a polynomial of degree 2α in k, with zeroes at $k = -1, -2, \cdots, -\alpha$ and $n+1, n+2, \cdots, n+\alpha$. Thus the $c_k(n, \alpha)$ satisfy the difference equation

$$\textstyle\sum_{k=-\alpha-1-\nu}^{\alpha+1-\nu} c_k(n, \alpha) R_{\alpha+1}(k+\nu) = 0\,, \qquad \nu = 0, \pm 1, \cdots. \tag{8.2}$$

LEMMA 8.1. *Define, for $\nu = 0, \pm 1, \cdots$*

$$S_\nu(n) = \textstyle\sum_{k=0}^{n} c_k(n, \alpha) R_{\alpha+1}(k+\nu)\,. \tag{8.3}$$

Then $S_\nu(n) = 0$ for $\nu \geqq \alpha + 2$ and $-n+1 \leqq \nu \leqq -1$, while

$$\lim_{n\to\infty} n^{\alpha+1} S_\nu(n) = (-1)^\nu [(2\alpha+1)!/\alpha!]\binom{\alpha+1}{\nu}\,, \qquad 0 \leqq \nu \leqq \alpha+1\,. \tag{8.4}$$

PROOF. For $\nu \geqq \alpha + 2$, $S_\nu(n)$ involves only the $R_{\alpha+1}(\tau)$ for $\tau \geqq \alpha + 2$, which all vanish. For $-n+1 \leqq \nu \leqq -1$ we may write $S_\nu(n)$ as the left side of (8.2) by adding terms with $c_k(n, \alpha) = 0$ and deleting terms with $R_{\alpha+1}(k+\nu) = 0$. Finally, suppose that $0 \leqq \nu \leqq \alpha + 1$. Note that

$$\lim_{n\to\infty} n^{\alpha+1} c_k(n, \alpha) = (2\alpha+1)\binom{2\alpha}{\alpha}(k+\alpha)!/k!\,. \tag{8.5}$$

Since, for $n \geqq \alpha + 1 - \nu$, the upper summation limit in (8.3) can be taken as $\alpha + 1 - \nu$, as $n \to \infty$ we have

$$\begin{aligned}(-1)^\nu (\alpha!)^2 \nu!\, (\alpha+1-\nu)!\, n^{\alpha+1} S_\nu(n)/[(2\alpha+1)!\,(2\alpha+2)!]\\ \to \textstyle\sum_{k=0}^{\alpha+1-\nu} (-1)^k \binom{\alpha+1-\nu}{k} B(k+\alpha+1, \nu+1) = B(\alpha+1, \alpha+2)\end{aligned}$$

with use of (8.5) and the beta integral, and (8.4) follows.

LEMMA 8.2. *Theorem* 8.1 *holds if $g(\lambda)$ is a nonnegative trigonometric polynomial.*

PROOF. Setting $g(\lambda) = \sum_{\nu=-\gamma}^{\gamma} b_\nu e^{i\nu\lambda}$ we easily find that $f(\lambda) = f_{\alpha+1}(\lambda)g(\lambda)$ has covariance function $R(\tau) = \sum_\nu b_\nu R_{\alpha+1}(\tau+\nu)$. Consequently,

$$n^{2\alpha+2} \operatorname{Var}(\hat{m}_\alpha, f) = \textstyle\sum_{\nu=-\gamma}^{\gamma} b_\nu n^{2\alpha+2} W_\nu(n)\,,$$

where

$$W_\nu(n) = \textstyle\sum_{k=0}^{n} c_k(n, \alpha) S_{\nu-k}(n)$$

and the $S_\nu(n)$ are defined in (8.3). Now, since $c_k(n, \alpha) = c_{n-k}(n, \alpha)$ and $R_{\alpha+1}(\tau) = R_{\alpha+1}(-\tau)$, we find that $S_{-\nu}(n) = S_{-n+\nu}(n)$ and $W_{-\nu}(n) = W_\nu(n)$. Since $c_k(n, \alpha) = 0$ for $-\alpha \leqq k \leqq -1$, from Lemma 8.1 and (8.5) we find that, for $1 \leqq \nu \leqq \gamma$ and $n \geqq \nu$,

$$n^{2\alpha+2}W_\nu(n) = n^{2\alpha+2} \sum_{k=0}^{\alpha+1} c_{\nu-k}(n, \alpha)S_k(n)$$
$$\to A \sum_{k=0}^{\alpha+1} (-1)^k \binom{\alpha+1}{k} (\nu - k + 1)(\nu - k + 2) \cdots (\nu - k + \alpha) = 0$$

as $n \to \infty$, where A depends only on α. The last sum vanishes since $(\nu - k + 1)(\nu - k + 2) \cdots (\nu - k + \alpha)$ is a polynomial of degree α in k. Since $S_{-n}(n) = S_0(n)$, from Lemma 8.1 and (8.5) we also obtain, as $n \to \infty$,

$$n^{2\alpha+2}W_0(n) = 2n^{2\alpha+2}c_0(n, \alpha)S_0(n) \to 2[(2\alpha + 1)!/\alpha!]^2 .$$

The result then follows because $2b_0 = \pi^{-1} \int_{-\pi}^{\pi} g(\lambda)\, d\lambda$.

PROOF OF THEOREM 8.1. For convenience, we write $I(g) = \int_{-\pi}^{\pi} g(\lambda)\, d\lambda$ and define

$$\psi_n(\lambda) = [\alpha!/(2\alpha + 1)!]^2 n^{2\alpha+2} |p_{n,\alpha}(e^{i\lambda})|^2 f_{\alpha+1}(\lambda) .$$

We wish to show that $I(\psi_n g) \to \pi^{-1}I(g)$ as $n \to \infty$. We let $t_\nu(\lambda)$, $\nu = 0, 1, \cdots,$ denote the Féjer means of the Fourier series for $g(\lambda)$, and note that $I(t_\nu) = I(g)$. Then

$$\begin{aligned} &|\pi^{-1}I(g) - I(\psi_n g)| \\ &\quad \leqq |\pi^{-1}I(t_\nu) - I(\psi_n t_\nu)| + \int_{-\delta}^{\delta} \psi_n(\lambda)|t_\nu(\lambda) - g(\lambda)|\, d\lambda \\ &\quad\quad + \int_{\delta \leqq |\lambda| \leqq \pi} \psi_n(\lambda)|t_\nu(\lambda) - g(\lambda)|\, d\lambda \end{aligned} \tag{8.6}$$

for any $0 < \delta \leqq \pi$. The proof will be complete if we show that each term on the right side approaches zero as first $n \to \infty$, then $\nu \to \infty$, then $\delta \to 0$.

By Lemma 8.2, the first term tends to zero as $n \to \infty$ for each fixed ν. By Lemma 6.2, $\psi_n(\lambda)$ is bounded by a constant C_δ independent of n for $|\lambda| \geqq \delta > 0$. Thus the last term in (8.6) is bounded uniformly in n by $C_\delta I(|t_\nu - g|)$, which for fixed δ vanishes as $\nu \to \infty$ by a property of the Féjer means. From Lemma 8.2, $I(\psi_n) \to 2$ as $n \to \infty$. Therefore, as $n \to \infty$, the second term on the right side of (8.6) is bounded by $2 \sup_{|\lambda| \leqq \delta} |t_\nu(\lambda) - g(\lambda)|$. That this expression approaches zero as $\nu \to \infty$, then $\delta \to 0$, follows from the fact that the $t_\nu(\lambda)$ are continuous at $\lambda = 0$ uniformly in ν, as is easily proved using properties of Féjer's kernel.

9. Remarks. The results of this paper may easily be extended to the estimation of m in the model $X_t = me^{i\mu t} + Y_t$, where μ is a fixed constant in $[-\pi, \pi]$. With $\tilde{X}_t = X_t e^{-i\mu t}$ and $\tilde{Y}_t = Y_t e^{-i\mu t}$, this new model reduces to our original: $\tilde{X}_t = m + \tilde{Y}_t$. If $\{Y_t\}$ has spectral density $f(\lambda)$ then $\{\tilde{Y}_t\}$ has spectral density $\tilde{f}(\lambda) = f(\lambda + \mu)$, with definition of $f(\lambda)$ extended outside $[-\pi, \pi]$ by periodicity. In the new model it is thus behavior of $f(\lambda)$ near $\lambda = \mu$ that is pertinent.

The types of "multipliers" $g(\lambda)$ allowed in Theorems 4.1 and 5.2, and defined in Section 3, may seem overly complicated. However, we want to stress that it is only continuity of and behavior near $\lambda = 0$ of the spectral density that are important. It does not matter if there are discontinuities or isolated zeroes away from the origin.

In Sections 5 through 8 we have dealt with spectra behaving like $\lambda^{2\alpha}$ near $\lambda = 0$. We should remark that spectra corresponding to $\alpha > 1$ are unlikely to occur in practice, except possibly through difference filtering. However, spectra with $-\frac{1}{2} < \alpha \leqq 1$ would seem to be practically useful.

Acknowledgments. The author is grateful to the editor and the referees for their helpful comments on the paper.

REFERENCES

ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series.* Wiley, New York.

CHIPMAN, J. S., *et al.* (1968). Efficiency of the sample mean when residuals follow a first-order stationary Markoff process. *J. Amer. Statist. Assoc.* **63** 1237–1246.

CLEVELAND, W. S. (1971). Projection with the wrong inner product and its application to regression with correlated errors and linear filtering of time series. *Ann. Math. Statist.* **42** 616–624.

ERDÉLYI, A., *et al.* (1953). *Higher Transcendental Functions*, **1** and **2**. McGraw-Hill, New York.

FISHMAN, G. S. (1972). Estimating the mean of observations from a wide-sense stationary autoregressive sequence. *J. Amer. Statist. Assoc.* **67** 402–406.

GRENANDER, U. (1950). Stochastic processes and statistical inference. *Ark. Mat.* **1** 195–277.

GRENANDER, U. (1954). On the estimation of regression coefficients in the case of an autocorrelated disturbance. *Ann. Math. Statist.* **25** 252–272.

GREVILLE, T. N. E. (1966). On stability of linear smoothing formulas. *SIAM J. Numer. Anal.* **3** 157–170.

KRUSKAL, W. (1968). When are Gauss-Markov and least squares estimators identical? A coordinate-free approach. *Ann. Math. Statist.* **39** 70–75.

PARZEN, E. (1961). An approach to time series analysis. *Ann. Math. Statist.* **32** 951–989.

ROZANOV, J. A. (1971). *Infinite-Dimensional Gaussian Distributions.* American Math. Soc., Providence.

SHEPPARD, W. F. (1913). Reduction of error by means of negligible differences. *Proc. Fifth International Congress of Mathematicians*, **2** 348–384, Cambridge.

TRENCH, W. F. (1971). Discrete minimum variance smoothing of a polynomial plus random noise. *J. Math. Anal. Appl.* **35** 630–645.

TRENCH, W. F. (1973). Strong discrete minimum variance estimation of a polynomial plus random noise. *J. Math. Anal. Appl.* **41** 20–33.

VITALE, R. A. (1973). An asymptotically efficient estimate in time series analysis. *Quart. Appl. Math.* **30** 421–440.

WATSON, G. S. (1967). Linear least squares regression. *Ann. Math. Statist.* **38** 1679–1699.

WATSON, G. S. (1972). Prediction and the efficiency of least squares. *Biometrika* **59** 91–98.

ZYSKIND, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statist.* **38** 1092–1109.

CENTER FOR THE APPLICATION OF MATHEMATICS
LEHIGH UNIVERSITY
BETHLEHEM, PENNSYLVANIA 18015