Ann. Inst. Statist. Math.
39 (1987), Part A, 289-297

INTERVAL ESTIMATION OF THE CRITICAL VALUE IN A GENERAL LINEAR MODEL*

Y. L. TONG

(Received July 29, 1985; revised Dec. 10, 1985)

Summary

This paper concerns interval estimation of the critical value θ which satisfies $\mu(\theta)=\sup_{x\in\mathcal{X}}\mu(x)$ under the general linear model $Y_i=\mu(x_i)+\varepsilon_i$ $(i=1, 2,\cdots)$, where $\mu(x)=\sum_{j=1}^{p}\beta_j f_j(x)$ for $x\in\mathcal{X}$ and the functional forms of f_j's are known. From an asymptotic expansion it is shown that, under reasonable conditions, the limiting distribution of $\sqrt{n}(\hat{\theta}_n-\theta)$ is normal. Thus in the large-sample case a confidence interval for θ can be obtained. Such a result is useful when one is interested in carrying out a retrospective analysis rather than designing the experiment (as in the Kiefer-Wolfowitz procedure). In Section 3 a sequential procedure is considered for confidence intervals with fixed width $2d$. It is shown that, for a given stopping variable N, $\sqrt{N}(\hat{\theta}_N-\theta)$ is also asymptotically normal as $d\to 0$. Thus the coverage probability converges to $1-\alpha$ (preassigned) as $d\to 0$. An example of application in estimating the phase parameter in circadian rhythms is given for the purpose of illustration.

1. Introduction and assumptions

Consider a general linear model

$$Y_i=\mu(x_i)+\varepsilon_i\;, \qquad i=1, 2,\cdots \tag{1.1}$$

where $\mu(x)$ is a real-valued function of x, $\{x_i\}$ a sequence of real numbers, and $\{\varepsilon_i\}$ a sequence of i.i.d. random variables with means zero and a common unknown variance $\sigma^2<\infty$. Let θ denote the value of x at which $\mu(x)$ reaches its supremum. This paper concerns interval estimation of θ when the parametric form of $\mu(x)$ is known.

---

* Research partially supported by the NSF Grant DMS-8502346.

Key words and phrases: Critical value, confidence interval, sequential estimation.

One of the most well-known methods for estimating θ is, of course, the Kiefer-Wolfowitz procedure (Kiefer and Wolfowitz [10]). Under that procedure the functional form of $\mu(x)$ is assumed to be unknown, and the procedure involves an algorithm for the stochastic approximation of θ. Thus it may be viewed as a "nonparametric" solution to the problem. In certain statistical applications, however, the mean function $\mu(x)$ can be assumed to possess a given parametric form with unknown parameters, such as the polynomial model or the trigonometric model (see e.g., Graybill [7], Sections 8.7 and 8.8). In those cases θ depends on $\mu(x)$ only through the regression parameters β_j's. Thus it can be estimated by estimating the β_j's. Such an approach is of great importance for problems in which one is carrying out a retrospective analysis when the data are already collected.

For the general case we assume that $\mu(x)$ in (1.1) is of the form

$$\mu(x)=\sum_{j=1}^{p} \beta_j f_j(x) , \tag{1.2}$$

where $p\geqq 1$ is fixed, the functional forms of $f_j(x)$ ($j=1, 2, \cdots, p$) are predetermined, and the β_j's are unknown parameters. (In applications the number of terms in $\mu(x)$ (the value of p) may be determined by the usual selection of variables method in regression analysis.) Let $\mathscr{X}$ denote the interval of interest where $\mathscr{X}=[a, b]$ for some finite a, b or $\mathscr{X}=(-\infty, \infty)$, and let us define

$$\mu(\theta)=\sup_{x\in\mathscr{X}} \mu(x) , \tag{1.3}$$

which is a function (although not necessarily explicitly) of

$$\boldsymbol{\beta}=(\beta_1, \cdots, \beta_p)^T . \tag{1.4}$$

For each fixed n and a sequence of predetermined real numbers $\{x_n\}$, $x_n\in\mathscr{X}$, let us rewrite (1.1) in the form

$$\boldsymbol{Y}_n=(Y_1, \cdots, Y_n)^T=\boldsymbol{A}_n\boldsymbol{\beta}+\boldsymbol{\varepsilon}_n$$

where, for $a_{ij}=f_j(x_i)$, $\boldsymbol{A}_n=(a_{ij})$ and $\boldsymbol{\varepsilon}_n=(\varepsilon_1, \cdots, \varepsilon_n)^T$. Let $\boldsymbol{S}_n$ be the $(p\times p)$ matrix $\boldsymbol{A}_n^T\boldsymbol{A}_n$. It is assumed that

CONDITION B1. $\{\varepsilon_i\}$ is a sequence of i.i.d. random variables with means zero and unknown finite variance σ^2.

CONDITION B2. (i) $f_j(x)$ is bounded for $x\in\mathscr{X}$ for all j, (ii) $\boldsymbol{S}_n$ is of full rank for all n and (iii) there exists a positive definite matrix $\boldsymbol{\Sigma}$ such that $\boldsymbol{S}_n/n\to\boldsymbol{\Sigma}$ as $n\to\infty$.

CONDITION B3. (i) θ is an interior point of $\mathscr{X}$, (ii) $\mu^{(3)}(x)$ (the third derivative of $\mu(x)$) exists in a small neighborhood $(\theta-\varepsilon_0, \theta+\varepsilon_0)$

and (iii) for every given $\varepsilon<\varepsilon_0$ there exists an $\delta>0$ such that $|x-\theta|\geqq\varepsilon$ implies $\mu(x)<\mu(\theta)-\delta$.

Note that Conditions B1 and B2 are similar to those in Gleser [6], Perng and Tong [12], [13] and Srivastava [14]. Also note that if $\mu(x)$ is a continuous function of x for $x\in\mathcal{X}$, then Condition B3 (iii) is equivalent to saying that θ is unique.

2. Some large sample properties

For each n let $\hat{\boldsymbol{\beta}}_n=\boldsymbol{S}_n^{-1}\boldsymbol{A}_n^T\boldsymbol{Y}_n$ be the least squares estimator of $\boldsymbol{\beta}$, and

$$\hat{\sigma}_n^2=\frac{1}{n-p}\sum_{i=1}^n\left[Y_i-\sum_{j=1}^p\hat{\beta}_j(n)f_j(x_i)\right]^2 \tag{2.1}$$

the usual estimator of σ^2. Then $\hat{\theta}_n$, the estimator of θ, depends on $\hat{\boldsymbol{\beta}}_n$ through the sample regression function

$$\hat{\mu}_n(x)=\sum_{j=1}^p\hat{\beta}_j(n)f_j(x)\ , \tag{2.2}$$

and it satisfies

$$\hat{\mu}_n(\hat{\theta}_n)=\sup_{x\in\mathcal{X}}\hat{\mu}_n(x)\ . \tag{2.3}$$

Before proving a theorem concerning the asymptotic behavior of $\hat{\theta}_n$ we first observe that

LEMMA 1. *If Conditions* B1 *and* B2 *are satisfied then, as* $n\to\infty$, (a) $\sqrt{n}(\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta})$ *converges to a multivariate normal distribution with mean* 0 *and covariance matrix* $\sigma^2\boldsymbol{\Sigma}^{-1}$, (b) $\hat{\boldsymbol{\beta}}_n\to\boldsymbol{\beta}$ *a.s. and* (c) $\hat{\sigma}_n^2\to\sigma^2$ *a.s.*

PROOF. Condition B2 (i) implies that $(1/\sqrt{n})\max_{i,j}f_j(x_i)\to 0$ as $n\to\infty$, thus the lemma follows from Gleser [6] or Srivastava [14].

THEOREM 1. *If Conditions* B1, B2 *and* B3 *are satisfied then, as* $n\to\infty$, (a) $\hat{\theta}_n\to\theta$ *a.s.*, (b) $\sqrt{n}(\hat{\theta}_n-\theta)\xrightarrow{\mathcal{L}}\eta(0,\tau^2)$, *the normal distribution with mean* 0 *and variance*

$$\tau^2=\tau^2(\sigma^2,\theta)=\sigma^2(\boldsymbol{v}(\theta))^T\boldsymbol{\Sigma}^{-1}\boldsymbol{v}(\theta)/(\mu''(\theta))^2\ , \tag{2.4}$$

where $\boldsymbol{v}(x)=(f_1'(x),\cdots,f_p'(x))^T$.

PROOF. (a) Denote $c=\sup_{x\in\mathcal{X}}\sum_{j=1}^p|f_j(x)|$ which is finite and, for arbitrary but fixed ε in $(0,\varepsilon_0)$, let δ be such that

$$|x-\theta|\geqq\varepsilon \quad \text{implies} \quad \mu(x)<\mu(\theta)-\delta\,.$$

By Lemma 1 for every ω in the sample space (except on a set of probability measure zero) there exists an $N_\delta(\omega)$ such that $n>N_\delta(\omega)$ implies $\max_{1\leqq j\leqq p}|\hat{\beta}_j(n)-\beta_j|\leqq\delta/2c$, which implies

$$|\hat{\mu}_n(x)-\mu(x)|\leqq\delta/2 \qquad \text{uniformly in} \quad x\in\mathcal{X}\,.$$

Now from $\hat{\mu}_n(\theta)\leqq\hat{\mu}_n(\hat{\theta}_n)$ one has

$$\begin{aligned}0\leqq\mu(\theta)-\mu(\hat{\theta}_n)&=[\mu(\theta)-\hat{\mu}_n(\theta)]+[\hat{\mu}_n(\theta)-\hat{\mu}_n(\hat{\theta}_n)]+[\hat{\mu}_n(\hat{\theta}_n)-\mu(\hat{\theta}_n)]\\&\leqq[\mu(\theta)-\hat{\mu}_n(\theta)]+[\hat{\mu}_n(\hat{\theta}_n)-\mu(\hat{\theta}_n)]\,.\end{aligned}$$

Therefore for $n>N_\delta(\omega)$ one has $0\leqq\mu(\theta)-\mu(\hat{\theta}_n)\leqq\delta$, which implies (by Condition B3) $|\hat{\theta}_n-\theta|<\varepsilon$.

(b) For arbitrary but fixed $\varepsilon'>0$ (by (a)) there exists an $N_1(\varepsilon')$ such that $n>N_1(\varepsilon')$ implies $\mathrm{P}\,[|\hat{\theta}_n-\theta|\leqq\varepsilon_0]\geqq1-\varepsilon'$. Now consider the Taylor series expansion of $\hat{\mu}_n'(x)$ about θ for $x\in(\theta-\varepsilon_0,\,\theta+\varepsilon_0)$:

$$\hat{\mu}_n'(x)=\hat{\mu}_n'(\theta)+\hat{\mu}_n''(\theta)(x-\theta)+\frac{1}{2}\hat{\mu}_n^{(3)}(x')(x-\theta)^2\,, \tag{2.5}$$

where $\hat{\mu}_n'$, $\hat{\mu}_n''$ and $\hat{\mu}_n^{(3)}$ are the first three derivatives of $\hat{\mu}_n$ and $|x'-\theta|<|x-\theta|$. From (2.5) and the fact that

$$\hat{\mu}_n'(\hat{\theta}_n)=0\,, \qquad \hat{\mu}_n''(\hat{\theta}_n)<0\,,$$

one can write, when $\hat{\theta}_n$ is in this small neighborhood,

$$\sqrt{n}\,(\hat{\theta}_n-\theta)=-\sqrt{n}\,\hat{\mu}_n'(\theta)/\hat{\mu}_n''(\theta)+\sqrt{n}\,R_n\,, \tag{2.6}$$

where

$$R_n=\hat{\mu}_n'(\theta)\cdot Z_n/\{\hat{\mu}_n''(\theta)\cdot(\hat{\mu}_n''(\theta)+Z_n)\} \tag{2.7}$$

and

$$Z_n=\frac{1}{2}\hat{\mu}_n^{(3)}(x')(\hat{\theta}_n-\theta)\,.$$

It is easy to see that $Z_n\xrightarrow{p}0$ and $\sqrt{n}\,R_n\xrightarrow{p}0$ as $n\to\infty$. On the other hand, applying Lemma 1 and a well-known convergence theorem (see e.g., Anderson [2], p. 76) it can be shown that the limiting distribution of

$$-\sqrt{n}\,\hat{\mu}_n'(\theta)/\hat{\mu}_n''(\theta)=-\sqrt{n}\,\sum_{j=1}^{p}\hat{\beta}_j(n)f_j'(\theta)\Big/\Big[\sum_{j=1}^{p}\hat{\beta}_j(n)f_j''(\theta)\Big]$$

is $\eta(0,\,\tau^2)$. Thus there exists an $N_0(\varepsilon')\geqq N_1(\varepsilon')$ such that $n>N_0(\varepsilon')$ implies

$$
\begin{aligned}
&|P[\sqrt{n}(\hat{\theta}_n-\theta)\leqq\lambda]-P[\eta(0, \tau^2)\leqq\lambda]| \\
&\quad\leqq|P[\sqrt{n}(\hat{\theta}_n-\theta)\leqq\lambda, |\hat{\theta}_n-\theta|<\varepsilon_0]-P[\eta(0, \tau^2)\leqq\lambda]|+\varepsilon' \\
&\quad\leqq|P[-\sqrt{n}\hat{\mu}_n'(\theta)/\hat{\mu}_n''(\theta)\leqq\lambda, |\hat{\theta}_n-\theta|<\varepsilon_0]-P[\eta(0, \tau^2)\leqq\lambda]|+2\varepsilon' \\
&\quad\leqq 3\varepsilon' .
\end{aligned}
$$

Remark. The result given in Theorem 1 may be used to obtain a confidence interval for θ; namely, let $z_{\alpha/2}$ denote the $(1-\alpha/2)$-th percentile of the $\eta(0, 1)$ distribution, then for large n the probability content of the interval

$$
I_n=(\hat{\theta}_n-z_{\alpha/2}\hat{\tau}(\hat{\sigma}_n, \hat{\theta}_n)/\sqrt{n}, \hat{\theta}_n+z_{\alpha/2}\hat{\tau}(\hat{\sigma}_n, \hat{\theta}_n)/\sqrt{n}) \tag{2.8}
$$

is approximately $(1-\alpha)$. Such an approach is particularly important for those problems in which one is carrying out a retrospective analysis—rather than designing an experiment. It is in this sense that our approach is different from the Kiefer-Wolfowitz method.

3. Fixed-width sequential confidence intervals

In certain applications a fixed-width confidence interval for the critical value θ, with confidence probability approximately $(1-\alpha)$ (preassigned), may be desired. Since σ^2 is unknown, clearly there does not exist a solution under single-stage procedures. However, a sequential procedure may be developed under the framework of sequential estimation theory. For preassigned $d>0$ and $\alpha\in(0, 1)$ let $z=z_{\alpha/2}$. Let $\{x_i\}$ be a given sequence of real numbers such that Condition B2 is satisfied. To find a sequential confidence interval for θ one may proceed according to the following

PROCEDURE. (a) Observe the sequence of random variables $Y_1, Y_2, \cdots$, one at a time.
(b) After observing $Y_1, Y_2, \cdots, Y_n$, $n\geqq p+1$, compute $\hat{\beta}_n$, $\hat{\sigma}_n^2$, $\hat{\theta}_n$ and

$$
\hat{\tau}_n^2=\begin{cases} \hat{\sigma}_n^2(v(\hat{\theta}_n))^T\Sigma^{-1}v(\hat{\theta}_n)/(\hat{\mu}_n''(\hat{\theta}_n))^2 & \text{if } \hat{\mu}_n''(\hat{\theta}_n) \text{ exists}, \\ \infty & \text{otherwise}. \end{cases} \tag{3.1}
$$

Stop with $N=n$ where

$$
n=\text{the smallest integer such that } n\geqq z^2\hat{\tau}_n^2/d^2 . \tag{3.2}
$$

(c) When sampling stops, construct the confidence interval

$$
I_N=(\hat{\theta}_N-d, \hat{\theta}_N+d) . \tag{3.3}
$$

Note that in some applications (such as the collection and analysis of certain time series data) the values of x_i are predetermined, thus in the above-stated procedure $\{x_i\}$ is considered fixed. In some other cases if one has the freedom of choosing the x_i's sequentially, then it seems reasonable to do so in such a way that the asymptotic variance τ^2 is minimized. This is a different problem, and will not be studied here.

In the following theorem we show that, under this sequential procedure, the asymptotic efficiency is one and the probability of coverage converges to $1-\alpha$ as $d\to 0$.

THEOREM 2. *If Conditions* B1, B2 *and* B3 *hold, then*

$$\lim_{d\to 0} N/(z\tau/d)^2=1 \quad a.s., \qquad \lim_{d\to 0} \mathrm{E}N/(z\tau/d)^2=1, \tag{3.4}$$

and

$$\lim_{d\to 0} \mathrm{P}[\sqrt{N}(\hat{\theta}_N-\theta)\leqq\lambda]\to\mathrm{P}[\eta(0, \tau^2)\leqq\lambda] \qquad \textit{for all} \quad \lambda. \tag{3.5}$$

Consequently,

$$\lim_{d\to 0} \mathrm{P}[\theta\in I_N]=1-\alpha. \tag{3.6}$$

PROOF. (3.4) follows immediately from Theorem 1, Lemma 1 of Chow and Robbins [5] and a minor modification of Theorem 4.1 of Gleser [6]. To prove (3.5), first note that $\sqrt{N}R_N$ (where R_N is as in (2.7) when n is replaced by N) converges to 0 in probability as $d\to 0$. Hence by (2.6) and Slutsky Theorem the limiting distribution of $\sqrt{N}\cdot(\hat{\theta}_N-\theta)$ is identical to that of $-\sqrt{N}\hat{\mu}'_N(\theta)/\hat{\mu}''_N(\theta)$. Let us denote

$$g(\hat{\boldsymbol{\beta}}_N)=-\hat{\mu}'_N(\theta)/\hat{\mu}''_N(\theta)=-\sum_{j=1}^{p}\hat{\beta}_j(N)f'_j(\theta)\Big/\sum_{j=1}^{p}\hat{\beta}_j(N)f''_j(\theta), \tag{3.7}$$

and show that the limiting distribution of $\sqrt{N}(g(\hat{\boldsymbol{\beta}}_N))$ is $\eta(0, \tau^2)$ as $d\to 0$. For arbitrary but fixed small $\varepsilon>0$ and $\delta>0$ let us define

$$A_{1,n}=\bigcap_{j=1}^{p}[|\hat{\beta}_j(n)-\beta_j|\leqq\varepsilon],$$

$$A_{2,n}=[\mu''(\theta)-\varepsilon\leqq\hat{\mu}''_n(\theta)\leqq\mu''(\theta)+\varepsilon<0],$$

then there exists an ν_1 such that $n>\nu_1$ implies $\mathrm{P}(A_{k,n})\geqq 1-\delta/3$ for $k=1, 2$. On the other hand, by the proof of Theorem 2.2 of Albert [1] there exists a large ν_2 and a small c such that the probability content of the event

$$A_{3,n}=\left[\bigcap_{j=1}^{p}\bigcap_{m=(1-c)n}^{(1+c)n}\{|\hat{\beta}_j(m)-\hat{\beta}_j(n)|\leqq\varepsilon\tau/\sqrt{n}\}\right]$$

is at least $1-\delta/3$ whenever $(1-c)n>\nu_2$ holds. Now if a point ω is in the set $\bigcap_{k=1}^{3} A_{k,n}$, then for all m satisfying $\max(\nu_1, \nu_2) < (1-c)n \leqq m \leqq (1+c)n$ one must have

$$\begin{aligned}|g(\hat{\beta}_m)-g(\hat{\beta}_n)| &\leqq [\mu_n''(\theta)+\varepsilon]^{-2}|\mu_n'(\theta)\hat{\mu}_m''(\theta)-\hat{\mu}_m'(\theta)\hat{\mu}_n''(\theta)| \\ &\leqq [\mu_n''(\theta)+\varepsilon]^{-2}\Bigg[\left|\sum_{j=1}^{p}\hat{\beta}_j(m)f_j''(\theta)\right|\cdot\left\{\sum_{j=1}^{p}|\hat{\beta}_j(m)-\hat{\beta}_j(n)|f_j'(\theta)\right\} \\ &\quad+\left|\sum_{j=1}^{p}\hat{\beta}_j(n)f_j'(\theta)\right|\cdot\left\{\sum_{j=1}^{p}|\hat{\beta}_j(m)-\hat{\beta}_j(n)|f_j''(\theta)\right\} \\ &\quad+\left\{\sum_{j=1}^{p}|\hat{\beta}_j(m)-\hat{\beta}_j(n)|f_j'(\theta)\right\}\left\{\sum_{j=1}^{p}|\hat{\beta}_j(m)-\hat{\beta}_j(n)|f_j''(\theta)\right\}\Bigg] \\ &\leqq [\mu_n''(\theta)+\varepsilon]^{-2}[(\varepsilon\tau/\sqrt{n})(2b_0b_1b_2+b_1b_2\varepsilon\tau/\sqrt{n})] \equiv \varepsilon'\tau/\sqrt{n}\ ,\end{aligned}$$

where

$$b_0=\max_{1\leqq i\leqq p}|\beta_j|+\varepsilon\ , \qquad b_1=\sum_{j=1}^{p}|f_j'(\theta)|\ , \qquad b_2=\sum_{j=1}^{p}|f_j''(\theta)|$$

are finite real numbers. Thus for every fixed $\varepsilon'>0$ and $\delta>0$, there exists a large ν and a small c such that

$$\mathrm{P}\left[\bigcap_{m=(1-c)n}^{(1+c)n}\{|g(\hat{\beta}(m))-g(\hat{\beta}(n))|\leqq\varepsilon'\tau/\sqrt{n}\}\right]\geqq 1-\delta$$

whenever $\nu<(1-c)n$ holds; that is, the C2 condition in Anscombe [4] is satisfied. This together with Theorem 1, (3.4) and the main theorem of Anscombe [4] implies (3.5). (3.6) is then an immediate consequence of (3.4) and (3.5).

4. An example of application

In this section we, for the purpose of illustration, provide an example of application in studying circadian rhythms.

It is well-known that in many biomedical experiments the measurements of a physiological variable taken at different time points show a certain periodic (or rhythmic) pattern, such a pattern is usually called a biorhythm. The study of biorhythms has a fundamental importance in biomedical research, some convenient references in this area are Halberg [8], Halberg, Tong and Johnson [9] and the monograph by the National Institute of Mental Health [11]. A special case of biorhythms is the circadian rhythms in which one concerns the periodic behavior of a physiological variable (such as the oral temperature) as a function of 24-hour local time. Let x denote the local time with 2π being equivalent to 24 hours, and let $Y(x)$ denote the measurement of such a variable at a given time x. Then it is known (e.g., Halberg, Tong and Johnson [9] and Tong [15]) that in most cases one can assume the model

$$(4.1) \qquad Y(x)=\beta_1+A\cos(x-\theta)+\varepsilon\ , \qquad x\in[0, 2\pi]\ ,$$

where β_1, $A>0$ (the amplitude), θ (the phase) are unknown parameters and ε is a random variable with mean zero and finite unknown variance σ^2. Defining

$$(4.2) \qquad \beta_2=A\cos\theta\ , \qquad \beta_3=A\sin\theta$$

and writing

$$(4.3) \qquad \mu(x)=\beta_1+A\cos(x-\theta)=\beta_1+\beta_2\cos x+\beta_3\sin x\ , \qquad x\in[0, 2\pi]\ ,$$

one then has the model

$$(4.4) \qquad Y_i=\mu(x_i)+\varepsilon_i\ , \qquad i=1, 2, \cdots$$

as given in (1.1) when observations on Y are taken at time points $x_1, x_2, \cdots, x_n, \cdots$.

It is known (e.g., Halberg [8]) that in certain cases the effectiveness of treating a patient in a clinic can be significantly improved when the timing of the treatment is made to best fit the patient's rhythm. For this purpose the estimation of the patient's phase parameter θ $(=\tan^{-1}(\beta_3/\beta_2))$, which is the critical value in the linear model (4.4), is of special interest. The estimation of θ can be made based on (x_1, Y_1), $(x_2, Y_2), \cdots, (x_n, Y_n), \cdots$ where $\{x_i\}$ is a sequence of predetermined real numbers. For example, if the oral temperature of a patient is to be observed at 4 : 00 a.m., 8 : 00 a.m., noon, 4 : 00 p.m., 8 : 00 p.m. and midnight for several 24-hour periods, then one has

$$x_1=x_7=\cdots=\frac{2\pi}{6}, x_2=x_8=\cdots=\frac{4\pi}{6}, \cdots, x_6=x_{12}=\cdots=2\pi\ .$$

In this application, it is clear that (i) $f_1(x)=1$, $f_2(x)=\cos x$ and $f_3(x)=\sin x$ are bounded for all $x\in[0, 2\pi]$, (ii) θ is unique and (iii) $\mu^{(3)}(x)$ exists in a small neighborhood of θ. Thus Conditions B1 and B3 are satisfied. Furthermore, there exist many designs for which Condition B2 is also satisfied. For example, if one chooses an equally-spaced design with $x_1=x_{k+1}=\cdots=2\pi/k$, $x_2=x_{k+2}=\cdots=4\pi/k$, $\cdots$, $x_k=x_{2k}=\cdots=2\pi$ for any $k\geqq 3$, then it follows from Anderson ([3], p. 95) that

$$\sum_{j=1}^k \sin\frac{2j\pi}{k}=\sum_{j=1}^k \cos\frac{2j\pi}{k}=\sum_{j=1}^k \sin\frac{2j\pi}{k}\cos\frac{2j\pi}{k}=0\ ,$$

$$\sum_{j=1}^k \sin^2\frac{2j\pi}{k}=\sum_{j=1}^k \cos^2\frac{2j\pi}{k}=\frac{k}{2}\ .$$

Thus $\boldsymbol{S}_n$ in Condition B2 is of full rank for all n and $\boldsymbol{S}_n/n\to\boldsymbol{\Sigma}=(\delta_{ij})$ as $n\to\infty$, where

$$\delta_{11}=1\,, \quad \delta_{22}=\delta_{33}=\frac{1}{2} \quad \text{and} \quad \delta_{ij}=0 \qquad \text{for} \quad i\neq j\,.$$

Now for each given $n>3$, one can apply the least-squares method to obtain $\hat{\beta}_n=(\hat{\beta}_1(n), \hat{\beta}_2(n), \hat{\beta}_3(n))$, the estimator of $\boldsymbol{\beta}=(\beta_1, \beta_2, \beta_3)$, then observe

$$\hat{\mu}_n(x)=\hat{\beta}_1(n)+\hat{\beta}_2(n)\cos x+\hat{\beta}_3(n)\sin x$$

as defined in (2.2). The estimator of θ which satisfies (2.3) is thus given by

$$\hat{\theta}_n=\tan^{-1}(\hat{\beta}_3(n)/\hat{\beta}_2(n))\,.$$

Lemma 1 implies that $\hat{\boldsymbol{\beta}}_n\rightarrow\boldsymbol{\beta}$ a.s., thus $\hat{\theta}_n\rightarrow\theta$ a.s., as $n\rightarrow\infty$, and Theorem 1 yields the fact that $\sqrt{n}(\hat{\theta}_n-\theta)$ is asymptotically normal. Consequently, a confidence interval for the phase parameter θ in a circadian rhythm can be obtained using the results described in Sections 2 and 3.

GEORGIA INSTITUTE OF TECHNOLOGY

REFERENCES

[1] Albert, A. (1966). Fixed size confidence ellipsoids for linear regression parameters, *Ann. Math. Statist.*, **37**, 1602-1630.

[2] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.

[3] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*, Wiley, New York.

[4] Anscombe, F. J. (1952). Large sample theory of sequential estimation, *Proc. Camb. Phil. Soc.*, **48**, 600-607.

[5] Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequence confidence intervals for the mean, *Ann. Math. Statist.*, **36**, 457-462.

[6] Gleser, L. J. (1965). On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters, *Ann. Math. Statist.*, **36**, 463-467. Corrigenda (1966). *Ann. Math. Statist.*, **37**, 1053-1055.

[7] Graybill, F. A. (1976). *Theory and Application of the Linear Model*, Duxbury Press, North Scituate, Massachusetts.

[8] Halberg, F. (1974). Protection by timing treatment according to bodily rhythms—an analogy to protection by scrubbing before surgery, *Chronoblogia*, **1** (Suppl. 1), 27-68.

[9] Halberg, F., Tong, Y. L. and Johnson, E. A. (1967). Circadian system phase—an aspect of temporal morphology, in *The Cellular Aspects of Biorhythms* (ed. H. von Mayersbach), Berlin-Springer, 20-48.

[10] Kiefer, J. and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function, *Ann. Math. Statist.*, **23**, 462-466.

[11] National Institute of Health, H.E.W. (1970). *Biological Rhythms in Psychiatry and Medicine*, Public Health Service Pub. No. 2088, Washington, D.C.

[12] Perng, S. K. and Tong, Y. L. (1974). A sequential solution to the inverse linear regression problem, *Ann. Statist.*, **2**, 535-539.

[13] Perng, S. K. and Tong, Y. L. (1977). Optimal allocation of observations in inverse linear regression, *Ann. Statist.*, **5**, 191-196.

[14] Srivastava, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector, *J. R. Statist. Soc.*, B**29**, 132-140.

[15] Tong, Y. L. (1976). Parameter estimation in studying circadian rhythms, *Biometrics*, **32**, 85-94.