密度函数の統計的推定について

赤　池　弘　次

(1964 年 10 月受付)

On the Statistical Estimation of Density Functions

Hirotugu Akaike

In this paper methods of estimation of density functions (of probability and of spectra) are reviewed. It is concluded that the sucessive application of estimation procedures with different characteristics will yield most useful information for hitting a balance between the bias and the variance of an estimate.

Institute of Statistical Mathematics

ま え が き

確率変数についてその分布の型を問題にする場合，あらかじめ想定される分布型に対する観測値の適合度の検定として，よく知られているように χ^2 (カイ二乗) 検定と ω^2 (オメガ二乗) 検定とがある．このふたつの検定法をみると，前者は確率密度函数の考えにより近いところにあり，後者は直接分布函数に関係している．勿論分布函数も確率密度函数も (後者が存在する場合) 理論的には同一の内容のものであるから，このふたつを区別して考えても意味が無いようにみえる．しかし標本値によって考える場合には，分布函数の形の方が取扱いやすい．それは横軸上の座標 x との関係を無視して，分布函数の値 $F(x)$ だけに注目して議論を進めることにすれば，標本分布が本質的に distribution free となることにも関連している．このような事情にもかかわらず実測データは多くヒストグラムの形に整理される．これは単なる慣習によるものであろうか？　この疑問に答えるために，分布函数を知って何に利用しようとしているのかを考えてみると，時に応じて種々のウエイト函数 $W(s, x)$ $(s=1, 2, 3, \cdots)$ (はっきり与えられる場合もあるし，おおよその形が与えられる場合もあるが) について，

$$\gamma(s)=\int W(s, x)\,dF(x) \qquad (s=1, 2, 3, \cdots)$$

のような量を評価することに用いられる場合が殆んどであろうと思われる．このとき x は具体的な意味を持つ数値となり，$F(x)$ の代りに確率密度函数 $f(x)$ を用いて $dF(x)=f(x)dx$ と表現すれば，これが $F(x)$ と x との関係をより直接的に与えるものとなる．このような観点からすると $F(x)$ の推定値が与えられたとしても，結局は適当な $\Delta(>0)$ について $F(x+\Delta)-F(x)$ の推定値を求めることが望まれるようになる．ここで Δ をどのようにえらべばよいかを考えることにすると，すでに密度函数の推定に足をふみ入れているわけで，実際 Neyman-Pearson の fundamental lemma における密度の比較による領域の分割にもみられるように，一般に密度を用いて判断の条件が得られる場合が多いのである．

このようなわけで，理論上からは必らずしもすっきりしない密度函数の推定という問題が実用上からは依然として重要な問題となることがわかる．

さて筆者がはじめて統計的データの処理を試みたとき，実際ヒストグラムを作ろうとして感

じたのは，ひとつのクラスの区間の幅をどのようにとればよいのかという疑問であった．統計学の教科書には“全体のデータを10ないし20程度の区間に分割してヒストグラムを描く”というような表現がされているだけで，筆者の知る限りではどうしてそのようにとらねばならないのかについてふれているものはなかった．そこで，ヒストグラムを描くことは，そのクラスの中心の座標値における確率密度の値を推定することを目指しているものとみなして，区間の幅をどのように定めればよいかをみることにした．数値計算に容易というだけで，あまり一般的とはいえない密度函数を例にとって，この問題を直接数値的に検討してみるという極めて初等的な解析の結果は，やはりこれまでの教科書でいわれているような区間の数のとり方でよいらしいということになった〔1〕．その後この問題に関係する興味深い結果が Rosenblatt〔17〕によって報告され，同時にスペクトル密度の推定に関して同様な議論が発展し，筆者自身ふたたび“推定値の分散と偏りとをどのようにバランスさせるか”という形でこの問題にかかりあうこととなってしまった〔3, 4〕．そこで以下この問題についてのこれまでの結果を紹介し，筆者のこの問題について現在考えるところを述べてみたい．

§ 確率密度函数の推定

Rosenblatt の論文〔17〕は，密度函数に特定の型をあらかじめ想定しないで，直接密度函数の各点の値を推定するという意味から，“密度函数のノンパラメトリックな推定に関する注意 (Remark)” という題名になっている．

はじめに，問題の性質を見やすくするために最も簡単な場合についてみることにする．確率変数 X の分布函数 $F(x)$ が密度函数 $f(x)$ によって，

$$F(x)=\int_{-\infty}^{x}f(y)dy$$

と書ける場合，大きさ n のランダムサンプル $(X_1, X_2, \cdots, X_n)$ から求められる標本分布函数を $F_n(x)$ と表わすことにする．

$$F_n(x)=\frac{1}{n}(x \text{ 以下の値をとった } X_i \text{ の箇数}).$$

$f(y)$ の推定値 $f_n(y)$ として適当な $h(>0)$ について，

$$f_n(y)=\frac{F_n(y+h)-F_n(y-h)}{2h}$$

をとると，

$$E\{f_n\}=\frac{F(y+h)-F(y-h)}{2h}$$

$$D^2\{f_n\}=\left(\frac{1}{2h}\right)^2\cdot\frac{1}{n}\cdot(F(y+h)-F(y-h))(1-F(y+h)+F(y-h))$$

となる．そこで $F(x)$ が十分滑かであるとすれば $h\to0$ のとき，

$$F(y+h)-F(y-h)=2hf(y)+\frac{1}{3}f''(y)h^3+O(|h|^4)$$

従って，

$$E\{f_n(y)\}=f(y)+\frac{1}{6}f''(y)h^2+O(|h|^3)$$

$$(E\{f_n(y)\}-f(y))^2=\frac{1}{36}|f''(y)|^2h^4+o(|h|^4)$$

となり，一方，分散は $n\to\infty$ のとき，

$$D^2\{f_n(y)\}=\frac{1}{2hn}f(y)+o\left(\frac{1}{hn}\right)$$

となる．結局，二乗平均誤差は $h\to 0$ $hn\to\infty$ のとき，

$$E|f_n(y)-f(y)|^2=\frac{f(y)}{2hn}+\frac{h^4}{36}|f''(y)|^2+o\left(\frac{1}{hn}+h^4\right)$$

となる．ここで右辺の主要項についてみると，

$$-\frac{f(y)}{2n}\cdot\frac{1}{h^2}+\frac{h^3}{9}|f''(y)|^2=0$$

をみたす h，すなわち，

$$h=\left(\frac{9}{2}\frac{f(y)}{|f''(y)|^2}\right)^{1/5}\cdot n^{-1/5}$$

で最小値をとることがわかる．n に対して h をこの式で与えることにすると，$n\to\infty$ のとき $h\to 0$，$hn\to\infty$ となって前述の近似のための条件をみたしている．結局，この場合，

$$E|f_n(y)-f(y)|^2\sim\frac{5}{4}2^{-4/5}9^{-1/5}(f(y))^{4/5}|f''(y)|^{2/5}n^{-4/5}$$

となる．最適の h の決定には $f(y)$ と $f''(y)$ とが関係することがわかる．

さて，ここで上記の推定量 $f_n(y)$ の代りに，

$$\int w_n(u)du=1$$

なる $w_n(u)$ をとって，

$$\begin{aligned}\tilde{f}_n(y)&=\int w_n(y-u)dF_n(u)\\&=\frac{1}{n}\sum_{j=1}^{n}w_n(y-X_j)\end{aligned}$$

を用いる場合を考えてみよう．$w_n(u)$ として，

$$w_n(u)=\frac{1}{h}w\left(\frac{u}{h}\right)$$

をとってみる．ただし，

$$\int w(u)du=1$$
$$w(-u)=w(u)$$
$$\int|w(u)|^2du<\infty$$
$$\int|w(u)u^3|du<\infty$$

とする．このとき $nh\to\infty$ とすれば，

$$D^2(f_n(y))=\frac{1}{nh}f(y)\int w^2(u)du+o\left(\frac{1}{nh}\right)$$

が成立し，また偏りについては $h\to 0$ のとき，

$$\begin{aligned}Ef_n(y)-f(y)&=\int w(u)(f(y+hu)-f(y))du\\&=hf'(y)\int w(u)udu+\frac{h^2}{2}f''(y)\int w(u)u^2du+o(|h|^2)\end{aligned}$$

が成立する．従って $h\to 0$，$nh\to\infty$ のとき，

$$E|\tilde{f}_n(y)-f(y)|^2=\frac{f(y)}{nh}\int w^2(u)du+\frac{h^4}{4}|f''(y)|^2\left[\int w(u)u^2du\right]^2+o\left(\frac{1}{nh}+h^4\right)$$

となる．このの主要項は，

$$h=\left(\frac{f(y)}{|f''(y)|^2}\frac{\int w^2(u)du}{\left[\int w(u)u^2du\right]^2}\right)^{1/5}n^{-1/5}$$

で最小値をとることになる．こうして Rosenblatt は結局，

$$E|\tilde{f}_n(y)-f(y)|^2\sim Cn^{-4/5}$$

のオーダーになることを結論し，果して n^{-1} のオーダーになるような $\tilde{f}_n(y)$ が存在するであろうかといっている．

さてこれで密度函数の推定に際しては偏りが問題になり，推定を上手に行なうためには偏りと分散とのバランスを適当に評価しなくてはならないことが明らかになった．通常の不偏な推定量の場合とことなり，サンプルサイズの単純な増大だけでは必らずしも良い推定量が得られないわけで，対象に関する情報（偏りの評価）を何とかしてとり入れて行かなくては実用的な推定値が得られないわけである．

§ スペクトル密度函数の推定

全く同様の事情が，スペクトル密度函数の推定に際して発生する．これについては，ふるくから論文が発表されてきているが，その中で多くの研究者がこの問題についてどのように対処しようとしてきたかをみることにしよう．

連続な時間パラメータ t をもつ弱定常確率過程 $x(t)$ について，$Ex(t)=0$ とし，

$$R(\tau)=E\{x(t+\tau)x(t)\}$$

とする．$R(\tau)$ は連続で $\int|R(\tau)|d\tau<\infty$ と仮定すれば，$x(t)$ のパワースペクトル密度函数 $p(f)$ は，

$$p(f)=\int\exp(-2\pi if\tau)R(\tau)d\tau$$

によって与えられる．この $p(f)$ を $x(t)$ の観測値から推定することがこれから考える問題である．$x(t)$ はガウス過程であるとしてこれからの議論をすすめよう．この場合，観測値が〔$-T$, T〕の時間にわたって得られたものとする．このとき，

$$X\left(\frac{\nu}{2T}\right)=\frac{1}{\sqrt{2T}}\int_{-T}^{T}\exp\left(-2\pi i\frac{\nu}{2T}t\right)x(t)dt$$

とすると，$T\to\infty$ のとき $\frac{\nu}{2T}\to f$ とすれば，

$$E\left|X\left(\frac{\nu}{2T}\right)\right|^2\to p(f)$$

となり，$X\left(\frac{\nu}{2T}\right)$ の成分は互いに無相関に近づくことが示めされる．

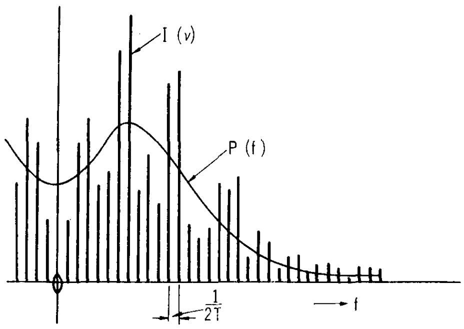
$$I(\nu)\equiv\left|X\left(\frac{\nu}{2T}\right)\right|^2$$

とすると，$2I(\nu)/E(I(\nu))$ $(\nu>0)$ は自由度2のカイ二乗分布に従い，$I(0)/E(0))$ は自由度1のカイ二乗分布に従う．これらの結果と $I(\nu)=I(-\nu)$ なることを考慮に入れると T が十分大きいときには $I(\nu)$ は近似的に

$$E\{I(\nu)\}=p\left(\frac{\nu}{2T}\right)$$

$$Cov\{I(\nu)I(\nu')\}=(\delta_{\nu,\nu'}+\delta_{\nu,-\nu'})p^2\left(\frac{\nu}{2T}\right)$$

をみたすものとして取扱うことができる．ただし $\delta_{\nu,\mu}=1(\nu=\mu)$，$0(\nu\neq\mu)$ である．$I(\nu)$ の分布の様子を模型的に画くと下図のようになる．



ひとつひとつの $I(\nu)$ の変異係数は T の大きさに無関係に一定値（$\nu\neq0$ に対して 1，$\nu=0$ では$\sqrt{2}$）をとるが，T が大となるに従って周波数軸（f 軸）上の $I(\nu)$ の間隔 $1/(2T)$ が小となってくるので，一定の周波数区間内に入る $I(\nu)$ の数は $2T$ に比例して増加する．そこでこれを狭い周波数区間にわたって平均化することにすれば，平均化の範囲がどんなに狭くても T が大となるに従い，十分多数の $I(\nu)$ が平均化に用いられることになり，得られた結果の分散はいくらでも小にできる．平均化の範囲を狭くすれば偏りも小さくすることができる．そこで $p\left(\frac{\mu}{2T}\right)$ の推定値 $\tilde{p}\left(\frac{\mu}{2T}\right)$ として，適当なウエイト $\{w_\nu^*\}$ を用いて得られる

$$\tilde{p}\left(\frac{\mu}{2T}\right)=\sum_\nu w_\nu^* I(\mu-\nu)$$

をとることにすれば，これは前述の確率密度函数の推定の際に用いられた $\tilde{f}_n(y)$ の定義式において，標本分布函数 $F_n(y)$ のかわりに，

$$F_T(f)=\sum_{2Tf\geqq\nu} I(\nu)\frac{1}{2T}$$

によって定義される $F_T(f)$ を用い，$w_n(y)$ の代りに $w_\nu^*=\frac{1}{2T}w_T\left(\frac{\nu}{2T}\right)$ となるような $w_T(f)$ を適用した形になっている．すなわち，

$$\tilde{p}\left(\frac{\mu}{2T}\right)=\int w_T\left(\frac{u}{2T}-f\right)dF_T(f)$$

である．ここで T が十分大であれば，

$$\tilde{p}\left(\frac{\mu}{2T}\right)=\sum_\nu w_T\left(\frac{\mu}{2T}-\frac{\nu}{2T}\right)I(\nu)\frac{1}{2T}$$

として前述の平均および共分散に関する近似を用いると，

$$E\left\{\tilde{p}\left(\frac{\mu}{2T}\right)\right\}=\sum_\nu w_T\left(\frac{\mu}{2T}-\frac{\nu}{2T}\right)p\left(\frac{\nu}{2T}\right)\frac{1}{2T}$$

$$D^2\left\{\tilde{p}\left(\frac{\mu}{2T}\right)\right\}=\frac{1}{2T}\sum_{\nu\geqq0}\left[w_T\left(\frac{\mu}{2T}-\frac{\nu}{2T}\right)+w_T\left(\frac{\mu}{2T}+\frac{\nu}{2T}\right)\right]^2(1+\delta_{\nu,0})^{-1}p^2\left(\frac{\nu}{2T}\right)\frac{1}{2T}$$

が得られる．これはまた近似的に，

$$E\left\{\tilde{p}\left(\frac{\mu}{2T}\right)\right\}=\int p\left(\frac{\mu}{2T}-f\right)w_T(f)df$$

$$D^2\left\{\tilde{p}\left(\frac{\mu}{2T}\right)\right\}=\frac{1}{2T}\int p^2\left(\frac{\mu}{2T}-f\right)w_T{}^2(f)df+\frac{1}{2T}\int p^2\left(\frac{\mu}{2T}-f\right)w_T(f)\,w_T\left(\frac{2\mu}{2T}-f\right)df$$

と評価できる．

ここで $\tilde{p}\left(\frac{\mu}{2T}\right)$ が $p\left(\frac{\mu}{2T}\right)$ の良い推定量となるためには通常予想されるような $p(f)$ に対して，$E\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)$ が $p\left(\frac{u}{2T}\right)$ に近いものとなることが望まれよう．このためには，$w_T(f)$ が f の狭い範囲の外では0に近いものであることが必要であり，従ってまた

$$\int w_T(f)df=1$$

となることが望まれる．さて $p(f)$ の形による $E\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)$ の $p\left(\frac{u}{2T}\right)$ からの変差を小さくとどめるためには $w_T(f)$ のひろがりはできる限り狭いことがのぞまれる．ところが，$w_T(f)$ のひろがりをせばめるにつれて $D^2\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)$ が大となってしまう．これを簡単な場合について見ることにしよう．$w_T(f)$ のひろがりがすでに $p(f)$ の変動に対して十分狭いものとなっているとすれば，$\frac{\mu}{2T}$ が0に近い場合を除けば，上記 $D^2\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)$ の評価式を与える積分の第二項は殆んど消え去って結局近似的に，

$$D^2\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)=\frac{1}{2T}\,p^2\left(\frac{\mu}{2T}\right)\int w_T{}^2(f)df$$

と評価できよう．今 $w_T(f)$ として，

$$w_T(f)\geqq 0$$
$$w_T(f)=0 \qquad |f|>\delta$$

となるものを考える．このとき $\int w_T(f)df=1$ なる条件の下で $\int w_T{}^2(f)df$ を最小にする（従って推定値の分散を最小にする）ような $w_T(f)$ は，$w_T(f)=\frac{1}{2\delta}$ $(|f|\leqq\delta)$ となる．従って一般に $D^2\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)$ は，この $w_T(f)$ に対する値，

$$\frac{1}{2T}\,p^2\left(\frac{\mu}{2T}\right)\frac{1}{2\delta}$$

より大きいものと考えなくてはならない．そこで δ を小さくして行くと，見掛け上この値は無限に増大していってしまう．実際は上記の近似は $\delta\gg\frac{1}{2T}$ の範囲でしか成立しないわけで，この値は $\frac{1}{2T}\,p^2\left(\frac{\mu}{2T}\right)$ に近いものとなっていく（$w_T(f)$ による平均化の範囲内に $I(\mu)$ ただひとつしか入らなくなってしまう）．かくして確率密度函数の推定の場合と同様に T が大となるにつれて，偏りも消失し，分散も小となるような $w_T(f)$ のえらび方が問題となる．

U. Grenander は上記の偏りと分散との相反関係をとりあげて，これを「不確定性原理」と呼んだ〔10, p. 524〕．すなわち，上記 2δ を分解能（resolvability）をあらわすひとつの不確定性 Δ_1（uncertainty）とし，推定値の分散を Δ_2 としてもうひとつの不確定性をあらわすものとすれば，

$$\Delta_1\Delta_2 \geqq \frac{1}{2T}\, p^2\left(\frac{\mu}{2T}\right)$$

となる. Δ_1 のかわりに一般の $w_T(f)$ の場合,

$$\Delta_1' = \left(\int f^2 w_T(f)df\right)^{1/2}$$

のような量をとっても同様な議論が成立する. いずれにしても“分解能”と“分散”とは相反する動きをするものであるというのが Grenander の主張である. この主張をめぐって後日興味深い論争が展開されることになった. この論争は, 日常の言葉であらわされている事柄の内容が, 実は意外に不明確なものであることを示すものであること, 特にそれらしい名前をつけてしまうことによって逆に一定の固定観念を生み出してしまい, その名称を使っている限り真に問題となる点を明らかにみることができなくなってしまう危険が大きいことを示す典型的な例とみることができる.

§ Z. A. Lomnicki および S. K. Zaremba と U. Grenander の論争とスペクトル・ウィンドウの構成法

ことのおこりは1956年11月に Royal Statistical Society で行なわれた「時系列へのスペクトル的接近」に関するシンポジウムにおける Z. A. Lomnicki と S. K. Zaremba の「かくしてわれわれは Grenander の不確定性原理の与える印象とは正反対の決論に到達する. 分解能 (resolvability) と統計的信頼性 (statistical reliability) とは決して相反的ではない. 逆に, 最も信頼できる推定量が, 最高の分解能を持っている」という発言である〔13〕. この発言は次のような考察にもとづいている. $\{x_n\}$ を離散的な時間パラメータを持つ弱定常過程とする.

$$R_k = E\{x_{n+k}x_n\}$$

とすると, この場合のパワースペクトル密度函数 $p(f)$ $\left(-\frac{1}{2}<f<\frac{1}{2}\right)$ は

$$p(f) = \sum_{k=-\infty}^{+\infty} R_k \cos 2\pi fk$$

となる (以下, 級数あるいは積分等の収歛のための条件はみたされているものとする). このとき,

$$C_{k,N} = \frac{1}{N-|k|}\sum_{n=1}^{N-|k|} x_n x_{n+|k|}$$

とすると, 前述の平均化に対応する演算は, R_k の推定値 $C_{k,N}$ を直接上式の R_k に代入するかわりに, これに適当な定数 $\lambda_{k,N}$ を乗じて,

$$\tilde{p}(f) = \sum_{k=1-N}^{N-1} \lambda_{k,N} C_{k,N} \cos 2\pi kf$$

とすることによって実現される. そこで

$$M(\tilde{p}) = \int_{-1/2}^{1/2} E[\tilde{p}(f)-p(f)]^2 df$$

を誤差の総合的な尺度とすると,

$$\begin{aligned} M(\tilde{p}) &= \sum_{k=1-N}^{N-1} E[\lambda_{k,N}C_{k,N}-R_k]^2 + 2\sum_{k=N}^{\infty} R_k^2 \\ &= \sum_{k=1-N}^{N-1} \{\lambda^2_{k,N} D^2(C_{k,N}) + (1-\lambda_{k,N})^2 R_k^2\} + 2\sum_{k=N}^{\infty} R_k^2 \end{aligned}$$

となる. これを最小にする推定量, すなわち, 最も信頼すべき推定量は,

$$\lambda_{k,N} = \frac{R_k^2}{D^2(C_{k,N})+R_k^2} \quad (k=1-N,\ 2-N,\ \cdots,\ N-1)$$

とおくことによって得られる．ところで，分解能とは今の場合 $p(f)$ の変化を検出する能力といえる．そこで $p(f)$ の変化，すなわち $p(f)$ の微係数 $p'(f)$ と考え，これをどのように $\tilde{p}(f)$ の微係数 $\tilde{p}'(f)$ が追跡するかをみると，

$$\int E[\tilde{p}'(f)-p'(f)]^2 df=\sum_{k=1-N}^{N-1}(2\pi k)^{2n}\{\lambda^2_{k,N}D^2(C_{k,N})+(1-\lambda_{k,N})^2R_k^2\}+2\sum_{k=N}^{\infty}(2\pi k)^{2n}R_k^2$$

となることがわかる．明らかに，これを最小にする $\lambda_{k,N}$ の値は前の場合と同一である．ということは $p(f)$ の最も信頼すべき推定量は，また最も分解能の高い推定量となっていることを示している．

これに対して Grenander および M. Rosenblatt はシンポジウムの討論会において反論を加え，更に J. R. S. S. 誌上にその詳細を論文として発表した〔12〕．その中で，「その原理の量的な表現は不確定さの尺度のとりかたに依存するであろう．質的には，そのことの論理的構造はかわらないであろう：すなわち，分解能（バンド幅*，偏り）と統計的信頼性とは相反するものである〔Resolvability (bandwidth, bias) and statistical reliability are antagonistic〕」といっている〔12, p. 155〕．この論文に特に新しい点はないのであるが，以前の自分の論文で用いていた推定量の statistical variability という概念を用いるかわりに，Lomnicki および Zaremba の用いた statistical reliability なる表現をとってそれが直ちに statistical variability の小さいことを示すものとみなしたところは失敗だっといえよう．

これに対する Lomnicki および Zaremba による反論が再び J. R. S. S. 誌上に掲載された〔14〕．そのいうところは，「分解能とは，一般に認められている意味では隣接する周波数でのスペクトル密度の値を弁別する能力をいうのであって，バンド幅と同一視することはできない．分解能とバンド幅の狭さと，偏りの小ささとの三者の区別の不足が誤解を生ずるもとになっている．その上 Grenander が分散が小さいことと統計的信頼性とを区別しそこなっていることから更に混乱が生じている．推定量の統計的な信頼性は分散だけに依存するものではなく，その偏りにも依存するのだ．」というのである．

こうしてみるとこの論争は，結局お互いの言葉のとりちがいが主な原因ということで実につまらないもののようにみえる．しかしこの論争の中で，resolvability, bandwidth, bias の間の区別が明確にされなくてはならないということが強調されたことだけは注目に値する．さて，resolvability を上記の意味で隣接するふたつの周波数でのスペクトル密度の値を弁別する能力と考えることにしよう．このふたつの周波数での推定値のそれぞれの平均化の及ぶ範囲（バンド幅）が互いに重なり合わない（重なり合うところのウエイトが十分小さい）ようになると，それぞれの推定値は互いにほとんど無相関と考えてよいことになる．そうすると，それぞれの推定値の二乗平均誤差が十分抑えられていれば，ふたつの周波数での推定値の差と密度の差との間の二乗平均誤差は，抑えられることになるから，分解能は結局バンド幅と推定値の偏り，そして分散によって総合的に規定できるものと考えられる．そこで (Lomnicki および Zaremba のあげた最適の $\lambda_{k,N}$ のように当面の推定の目標値 $p(f)$ と同等な R_k の値を用いるようなものでない）一般の実用に供しうるような推定値の構成法は，このバンド幅，偏り，および分散の三者を同時に考慮に加えることによってはじめて実現されるであろう，という見透しが得られる．

まず，バンド幅と偏りとが同一内容でないことは，平均化のための函数 $w_T(f)$（スペクトル・ウインドウと呼んでいるが）が必らずしも正の値ばかりをとらなくてもよいものと考えると直ちに明らかになる．

---

* フィルターが一定以上のゲインを示めす周波数範囲の幅で，ここでは実質的に平均化がおよぶ範囲の幅を示すものと理解すればよい．

$$E\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)=\int w_N(f)p\left(\frac{\mu}{2T}-f\right)df$$

において，

$$p\left(\frac{\nu}{2T}-f\right)=p\left(\frac{\mu}{2T}\right)-fp'\left(\frac{\mu}{2T}\right)+\frac{f^2}{2!}p''\left(\frac{\mu}{2T}\right)+r(f)$$

とすると，

$$E\left(\tilde{p}\left(\frac{\mu}{2T}\right)\right)\sim p\left(\frac{\mu}{2T}\right)\int w_T(f)df-p'\left(\frac{\mu}{2T}\right)\int fw_T(f)df$$
$$+\frac{1}{2!}p''\left(\frac{\mu}{2T}\right)\int f^2w_T(f)df+\int r(f)w_T(f)df$$

従って $w_T(f)$ が大きい値をとる範囲で $r(f)$ が f^2 の項に比べて十分小であれば，

$$\int w_T(f)df=1$$
$$\int fw_T(f)df=0$$
$$\int f^2w_T(f)df=0$$

となるような $w_T(f)$ を用いることによって，偏りは十分小にとれることがわかる．同様にして更に高次の変化に対する補正も当然可能である．バンド幅の概念はここでは上記の局所的な近似を有意味ならしめるために用いられている．バンド幅の外では $w_T(f)$ は十分減衰し，上記最終項の影響を低いオーダーにとどめるようなものでなくてはならない．この点の判断を実用上可能にするためには，バンド幅として $w_T(f)$ によるウエイトの分布の特性と直接結びつく指標をとることが望まれるわけである．さて1963年に Bartlett〔6〕は上記偏りの補正が可能であることを指摘し，前述した Rosenblatt の論文の場合，推定値の二乗平均誤差をいくらでも n^{-1} のオーダーに近づけうることを述べ，それをスペクトル密度の推定の場合に適用し，2次の変化まで考慮した場合の $w_T(f)$ として，

$$w_T(f)=\frac{9}{8h}\left(1-\frac{5f^2}{3h^2}\right)\qquad |f|\leqq h$$
$$=0\qquad\qquad |f|>h$$

をとることを提案している．しかし，これは周波数空間で有界な範囲の外で完全に消失する函数であるために，これのフーリエ変換（係数）に相当する $\lambda_{k,N}$ のようなウエイトは無限に多くの k に対して0でない値をとる．結局系列相関係数 $C_{k,N}$ を $k=0,1,2,\cdots,N$ のすべてについて計算する必要が生ずる．これは直接原データ $\{x_n;n=1,2,\cdots,N\}$ について，

$$C_r=\frac{1}{\sqrt{N}}\sum_{n=1}^{N}x_n\cos\left(2\pi\frac{r}{N}n\right)$$
$$S_r=\frac{1}{\sqrt{N}}\sum_{n=1}^{N}x_n\sin\left(2\pi\frac{r}{N}n\right)\qquad r=0,1,\cdots,\left[\frac{N}{2}\right]$$

等を計算することと同等であるが，Bartlett は現在の電子計算機を用いれば，このような計算法も考えてよいのではないかと述べている．この後者の方法の時系列解析上の有用性は筆者も認めるところであるが（欄外文献にその利用の一例がある．その他後の議論に関係する H. E. Daniels の論文〔9〕でもこの方法にふれている），

H. Matsumoto, Y. Saigusa and H. Akaike, "Analytical studies on fluctuations found in time series of daily milk yield", Nat. Inst. Anim. Hlth. Quart. Vol. 2 (1962) pp. 161-171, Vol. 3 (1963) pp. 83-92.

たとえ現在の電子計算機を用いるとしても，このような方法は研究的な場合を除いては日常の工学的問題へのスペクトル解析の実際的な適用を困難なものとしてしまうほどの計算時間を必要とする（現在われわれが通常利用している実際の計算方式で15分で終る計算に8時間を要するということになり，現在の最高速の計算機によっても2分以上の計算ということになる．これでは少なくとも今日においてはあまりにも高価なものとなる）．更にこのような定式化では，$w_T(f)$ が $|f|>h$ の外では0であることを要請するために，$|f|<h$ 内での偏りに対する $w_T(f)$ の寄与が $w_T(f)$ の $f=0$ への集中の必要を無視して評価されている．当然のことながら，上記の偏りの評価は $w_T(f)$ の $f=0$ の附近での $p(f)$ の局所的な近似に依存しているわけで，このためには矢張り $w_T(f)$ を $f=0$ の近傍に集中させる（バンド幅一定としても）ような制約が更に $w_T(f)$ の形に課せられる必要があるのではなかろうか．

さて上記のような偏りの局所的な評価を利用し，推定量の二乗平均誤差を最小にするような設計法は2次の偏りに対する補正を考えない場合，全く前述の Rosenblatt の確率密度函数の推定法と同一であることは明らかであり，また実際その方法が Grenander および Rosenblatt の書物〔11, pp. 153–155〕の中でスペクトル密度の推定法の評価に適用されている．一方 1955 年に Biometrika 誌上に発表された Bartlett および Medhi による論文〔5〕では，同様に“不確定性原理”に基礎をおく評価方式によって最適の $w_T(f)$ が与えられている．ところがこれらの原理ならびに結果に相当するものは，既に1954年に I. R. E. の Proceeding 誌上に発表された通信工学研究者 S. S. L. Chung によるパワースペクトル解析機のフィルターの設計に関する論文〔8〕の中で全くそのままの形で与えられているのである．求められた $w_T(f)$ は

$$
\begin{aligned}
w_T(f)&=A\left\{1-\left(\frac{f}{h}\right)^2\right\} && |f|<h\\
&=0 && |f|>h
\end{aligned}
$$

の形になっている．A は $\int w_T(f)df=1$ にする定常．

勿論この場合は物理的なフィルターであるから $w_T(f)\geqq 0$ となっている．Bartlett も 1963 年の論文ではじめて $w_T(f)$ が負の値をとりうる場合を考えたわけである．$w_T(f)$ が正であるか，またはそれに極めて近いものでなくてはならないとする考え方は，主としてこのような物理的なフィルターの考えに結びついている．結局これらの研究の中では，この Chung による論文が最も早い時期にあらわれ，また極めて実際的な結果を与えるものとなっている．Chung は J. W. Tukey の研究結果〔18〕を利用し，W. J. Pierson（海洋のスペクトル解析に関して著名な研究者）等との接触の下にこの論文を作製している．

これらと少し異なった尺度により，同じく“不確定性原理”にもとずいた $w_T(f)$ の評価基準が E. Parzen〔16〕によって与えられている．しかしこれらの“不確定性原理”にもとずく評価は，多くの場合すでに与えられている $w_T(f)$ の評価にだけ利用され，これまでに引用した限りの文献中でそれが積極的に適切な $w_T(f)$ の構成原理として利用されている例は，前記 Chung の論文〔8〕および Bartlett と Medhi の論文〔5〕以外にはみられない．そしてこの場合の $w_T(f)$ も前述の Bartlett の例と同様に，そのままでは実用に不適なものとなってしまっている．結局ここでスペクトルウインドウの設計に際してもうひとつの条件を考慮に入れる必要が明らかになった．すなわち，計算所要時間（あるいは，ある意味での物理的実現性）である．これをあらわす尺度は，コレログラムを利用して推定値を算出するものとして，推定値の算出に実際に必要とされる系列相関係数 $C_{k,N}$ の総数 H（$|k|>H$ なる k について $\lambda_{k,N}=0$ となるような値）である．これは結局は推定量のいわゆる分解能に依存する量で，推定法の得失を最も明瞭に規定するものとなる．

以上のようにして，バンド幅，偏り，分散および計算所要時間というスペクトルウインドウの評価基準を一応規定したところで，これから1962年の筆者の論文〔3〕におけるウインドウの構成原理について述べてみよう．その前に，更にもうひとつの条件をとり上げることにしたい．これが実用上最も重要な点のひとつとなると思われるが，それは種々の特性のウインドウを逐次適用してその結果を比較してみるという場合のことを考慮に入れるということである．ここで時間軸上での係数 $\lambda_{k,N}$（これをデータ・ウインドウあるいはラグ・ウインドウ(lag window)とよんでいる）と周波数軸上でのウエイト $w_T(f)$ との対応を明確にしておくことにしよう．まず連続パラメータの場合についてみよう．観測値 $\{x(t);\ -T<t<T\}$ が与えられたとする．$x(t+2T)=x(t)$ なる関係によって $x(t)$ を周期 $2T$ を持つ函数とみなし，

$$C_T(\tau)=\frac{1}{2T}\int_{-T}^{T}x(t+\tau)x(t)dt$$

とする．このとき $C_T(\tau)$ のフーリエ係数は，

$$\begin{aligned}\frac{1}{2T}\int_{-T}^{T}&\exp(-2\pi i\frac{\nu}{2T}\tau)C_T(\tau)d\tau\\&=X\left(\frac{\nu}{2T}\right)\overline{X\left(\frac{\nu}{2T}\right)}\frac{1}{2T}\\&=I(\nu)\frac{1}{2T}\end{aligned}$$

である．次に $D(\tau)$ を $-T\leqq\tau\leqq T$ で定義された $\tau=0$ に関して対称な函数とする．これに対して

$$w_T(f)=\int_{-T}^{T}\exp(-2\pi if\tau)D(\tau)d\tau$$

とすると，$D(\tau)$ のフーリエ係数は，

$$w_\nu{}^*=\frac{1}{2T}w_T\left(\frac{\nu}{2T}\right)$$

となる．従って，

$$\begin{aligned}\int_{-T}^{T}&\exp\left(-2\pi i\frac{\mu}{2T}\tau\right)D(\tau)C_T(\tau)d\tau\\&=\sum_\nu w_\nu{}^*I(\mu-\nu)\end{aligned}$$

が成立する．これが前述の $\tilde{p}\left(\frac{\mu}{2T}\right)$ である．実際は $C_T(\tau)$ は τ が適当な Δt (>0) の整数倍にあたる所でしか求められないのが普通である．この場合，上記の積分は，

$$\sum_m\exp\left(-2\pi i\frac{\mu}{2T}m\Delta t\right)D(m\Delta t)\,C_T(m\Delta t)\Delta t$$

でおきかえられる．はじめから観測データが Δt おきにしか与えられていない場合には，

$$X\left(\frac{\nu}{2T}\right)=\frac{\Delta t}{\sqrt{N\Delta t}}\sum_{n=1}^{N}\exp\left(-2\pi i\frac{\nu}{N}n\right)x(n\Delta t)$$

とすれば $\nu=0,1,\cdots\left[\frac{N}{2}\right]$ について前と同様の関係が成立し，

$$C_T(m\Delta t)=\frac{1}{N}\sum_{n=1}^{N}x((n+m)\Delta t)x(n\Delta t)$$

で与えられることになる．従ってこれから $C_T(m\Delta t)$ をふたつの場合に対して区別しないで用いることにする．そこで

$$w_\nu{}^{**}=\sum_k w^*{}_{kN+\nu}$$

とおくと，

$$\sum_m \exp\left(-2\pi i\frac{\mu}{N}m\right)D(m\Delta t)C_T(m\Delta t)\Delta t=\sum_\nu w_\nu{}^{**}I(\mu-\nu)$$

となる．実用される $D(\tau)$ の場合には，

$$\left|w_T\left(\frac{k}{\Delta t}+f\right)\right|\ll|w_T(f)|;\quad |f|\leqq\frac{1}{2\Delta t}$$

となるようなものを用いるから，

$$w_\nu{}^{**}=w_\nu{}^{*}$$

とほぼみなすことができ，従って前出の積分を用いた平均分散の評価式が近似的に成立する．結局前出の $\lambda_{k,N}$ は今の $D(k\Delta t)$ に相当するわけである．更に実用上は $C_T(m\Delta t)$ の定義として，

$$C_T(m\Delta t)=\frac{1}{N}\sum_{n=1}^{N-|m|}x((n+|m|)\Delta t)x(n\Delta t)$$

を採用する．従って $D(\tau)$ が $|\tau|>T_M=H\Delta t$ で 0 であるようなものをとれば，対応する $\lambda_{k,N}=0(k>H)$ となるわけである．今ここで，

$$D(\tau)=\int\exp(2\pi i f\tau)w_T(f)df$$

が成立しているものとする．

さて既に見たように，$p(f)$ の局所的な高次の変化分にもとづく偏りを打ち消そうとすると，$w_T(f)$ の対応する次数のモーメントを0にする必要があった．これは $D(t)$ の $t=0$ における対応する次数の微係数が0にひとしいことを要求する．従ってもし $D(t)$ を適当なクラスに限定して考えると次第に平坦な形を $D(t)$ に要求することになるであろう．$D(t)$ の平坦な範囲が増大すれば，それにつれて推定量の分散が増大することが予想される．そこでやはりここでも偏りと分散との間での取り引きが必要になる．それぞれのウインドウは，あるものは偏りに対する抵抗力が強く，あるものは分散に対して強いというわけであって，対称な $w_T(f)$ に限ればそれぞれの優劣は $p(f)$ の偶数次の変動の大きいときには前者がすぐれ，小さいときには後者がよい結果を与えるということになる．実際の適用に際してこの間の釣合いがどのようになっているかは，あらかじめ完全に判断することはできない．従って最も実際的な方法としては，上記のようにそれぞれの特徴をもつ $w_T(f)$ を何種類か逐次適用してみることにすれば，その結果の間の差異にもとづいて更にどのような対策を講ずればよいかの決定の手掛りがえられるであろう．この場合それぞれの推定値を求めるたびごとに数値的なフーリエ変換を行なうのは計算所要時間の点で不利である．

ところがここで $D(\tau)$ の形を特殊なクラスに限定すると，$w_T(f)$ のちがいは，簡単なウエイト附き算術平均のウエイトの変更だけで実現できることとなってしまう．$D(\tau)$ として，

$$\begin{aligned}D(\tau)&=a_0+2\sum_{n=1}^{k}a_n\cos\left(2\pi\frac{n}{2T_M}\tau\right) && |\tau|<T_M\\ &=0 && |\tau|>T_M\end{aligned}$$

で与えられるものをとると，

$$\begin{aligned}\tilde{p}\left(\frac{\nu}{2T_M}\right)&=\int\exp\left(-2\pi i\frac{\nu}{2T_M}\tau\right)D(\tau)C_T(\tau)d\tau\\ &=\sum_n a_n\bar{p}\left(\frac{\nu-n}{2T_M}\right)\end{aligned}$$

となる．ただし，

$$\bar{p}\left(\frac{\nu}{2T_M}\right)=\int_{-T_M}^{T_M}\exp\left(-2\pi i\frac{\nu}{2T_M}\tau\right)C_T(\tau)d\tau$$

である．従って一旦 $\bar{p}\left(\frac{\nu}{2T_M}\right)$ $(\nu=0, 1, 2, \cdots)$を計算しておけば，$f=\frac{\nu}{2T_M}$ 等における $p(f)$ の推定値はウエイト $\{a_n\}$ を用いて $\bar{p}\left(\frac{\nu}{2T_M}\right)$ を平均することによって求められる〔7〕．従って $D(t)$ の変更は，単にこの ウエイト の変更に帰着してしまう（勿論判定の結果によっては更に T_M の変更が要求される場合もありうる）．

そこで筆者はこのようなクラスに $D(\tau)$ を限定し，その中でのウインドウの設計原理を展開することにした．このような逐次的な適用の効果は，筆者の経験する限りでは極めて明らかで，特に同様な方法を周波数応答函数の推定に利用する場合に著るしい〔4〕．十分スペクトル解析に慣れた後でも，見逃しやすい推定の失敗に対して有効な警報を与えられたことが多い．従って統計的解析の専門家でない一般の利用者に対しては，このような方法をルーチン化して提供することが極めて有効であり，また必要ではないかと思われる．筆者はそのひとつの試みを周波数応答函数の推定に関して行なったが〔4〕，その方法はスペクトル密度のみの推定の場合にも殆んどそのまま適用できる．ところで，このような解析方法の定式化を H. E. Daniels が 1962 年の J. R. S. S. 誌上の論文〔9〕で試みている．この J. R. S. S. 誌の到着が筆者の前記論文〔3〕の原稿提出後であったために，前記論文中ではこの注目すべき論文について言及することができなかったが，筆者と殆んど全く同一の見方に立って偏りの処理が論じられている．Daniels はその方法に対して unsmoothing（次第に高次の変化まで追求する結果，一旦平滑化された曲線を次第に原曲線に近づける）という呼称を与え，逐次的な処理に対する原理的な考察を行なっている．その着想は極めてすぐれたものであるが，実用化との間にはまだ稍々距離があるように思われる．ところで前記筆者の設計法による一連のウインドウの逐次的な適用は，極めて容易にこの着想を実用化するひとつの方法を与えているものとみることができる．

ここで筆者のとった設計原理を極く簡単に箇条書きすると，

1. $D(\tau)$ を上記三角函数の和の形のものに限った．
2. バンド幅 B を，

$$B=2\left(\frac{\int f^2|w_T(f)|^2df}{\int|w_T(f)|^2df}\right)$$

   によって定義した．
3. 0 次，2 次，4 次等の偏りに対する対策を上記のようにしてとった．
4. バンド幅一定の条件の下で推定量の分散を最小にするように $a_0, a_1, \cdots, a_k$ の決定を行なった．
5. 以上の方針に沿って得られたウインドウを，その他の分散最小，バンド幅最小等の考え方から導かれるものと比較し，更に典型的な $R(\tau)$ の形について偏りの影響を評価した．

となる．これで求められた結果によると，以前から Tukey がその使用を提案してきたもの〔7〕と近いウインドウが一定の条件の場合に得られることが認められ，従来主として使用経験に依存してきたそれらの得失の判定を統一的な観点から行なうことができる．

む　す　び

こうしてみるとそれぞれの研究者は意外に同じような時期に同じような発想を行なっていることがわかる．読者はこんなに初等的な問題をめぐって多くの研究者が10年以上の時間をかけてきていることが馬鹿らしいことのように思われるかもしれない．その馬鹿馬鹿しいといった感じには全く正当な面もあるが，それによってひとつの重要な事実を見逃しやすいという危険

な面も含んでいる．その事実とは前述の Lomnicki および Zaremba と Grenander の論争の際に明らかにされていることであるが，この問題をめぐってさまざまな不明確な概念が登場し，それが問題の決定的な解決をおくらせる作用をしてきたということである．この不明確な概念の登場を単なる明確な定義式の欠如によるとだけ考えるのは早計で，その背後にある“ある実体”をなるべく適切に描き出そうとし，それに合った形で問題の解決を求めようとする努力のあらわれとみる必要がある．それではそのある実体とはなにかというと，当然それは“現実のスペクトル解析”ということになる（なお，これに類する意見については Parzen〔15〕を参照されたい）．これまでの多くの研究者の努力が実際的にみてもある程度有効な結果を与えてきたということはこのためであったと理解されると同時に，それはまた前述のような初等的な形に問題がとどまっていた理由をも与えるものであろう．通常実用性をはなれると同時に問題が複雑化する例が多い．この問題に対する適切な解答は現実のスペクトル解析の種々の様相に対する十分な理解を要求していたと考えることが必要ではないかというのが筆者の見方である．そしてこれは事実その通りのように思われる，というのは，スペクトル解析の実際的応用に関して必要な理論的考察として，多くの経験にもとづいて1949年に Woods Hole における自己相関解析の物理的問題への応用に関するシンポジウムに発表された Tukey の先駆的な論文〔18〕をみると，実用上の問題となる点はその後の他のどの論文よりも明快にその中で処理されていることが認められるのである．この内容は更に後日詳細が発表されている〔7〕が，更に最近の論文〔19〕でみるとますます Tukey は本文でみてきたようなウインドウの設計を問題としていないように見える．最も初期に採用した $D(\tau)=\frac{1}{2}+\frac{1}{2}\cos\left(2\pi\frac{\tau}{2T_M}\right)$ 型のもので十分だというのである．筆者自身この意見に殆んど賛成である〔2〕．その意味ではふたたび馬鹿らしい感じにもどることになってしまうが，依然としてこの意見に賛成できない点がある．それはやはり実際的な場面での偏りの処理で，前節で述べた各種ウインドウの逐次的使用の効果およびその簡易性は無視できないということ，および Tukey のいわゆるプリホワイトニング（あらかじめデータをそのパワースペクトル密度に急激なあるいは大きな変化をもたないもの，すなわちホワイトノイズに近いものに変形しておく方法）〔7〕では実際上処理しきれない場面があるということである．本文で述べたような逐次的方法は従来も経験的には多く用いられ，現に Tukey 自身もその有効性をみとめている〔19〕．筆者はここに定式化可能な実用上も重要な意味をもつひとつの具体的な推定あるいは決定の問題があると思う．それは前記 Daniels の論文〔9〕ですでにある程度取扱われているが，スペクトル密度推定法の当面する最も面白い問題がこれではないかと思われる．この問題を一般的に雑音に埋もれた信号の平滑化としてみると，同様な考え方の適用されうる分野は広いものであろうことが予想される．

付記　研究所創立15周年記念号（彙報第7巻第1号）において，筆者は確率過程に関する統計理論の発展の方向を論じ，特に各種研究分野の研究者との協力を理論の発展のための基礎的条件であると述べた．この点はわが国においてはその後の五年間に着実に具体化され，現在は既に他分野の研究者の側から確率過程に関する統計理論そのものの更に深い進歩を強く要求される段階となっている．前回の報告を書く頃の適用分野との関係についての不安感は現在は全く拭い去られて，ただ統計理論の研究への努力の必要を痛感する状態である．従って今回の報告では技術的な面に重点をしぼってスペクトル解析に関する統計的方法の中心的な問題をとりあげて論ずることとした．

統計数理研究所

参 考 文 献

〔1〕 H. Akaike, "An approximation to the density function", Ann. Inst. Stat. Math. Vol. 6(1954) pp. 127-132.

〔2〕 H. Akaike, "Undamped oscillation of the sample autocovariance function and the effect of prewhitening operation", Ann. Inst. Stat. Math. Vol. 13 (1961) pp. 127-143.

〔3〕 H. Akaike, "On the design of lag window for the estimation of spectra", Ann. Inst. Stat. Math., Vol. 14 (1962), pp. 1-21.

〔4〕 H. Akaike, "Statistical measurement of frequency response function", Studies on The Statistical Estimation of Frequency Response Function, Ann. Inst. Stat. Math. Supplement III (1964). pp. 5-17.

〔5〕 M. S. Bartlett and J. Medhi, "On the efficiency of procedures for smoothing periodgrams for time series with continuous spectra", Biometrika, Vol. 42 (1955). pp. 143-150.

〔6〕 M. S. Bartlett, "Statistical estimation of density functions", Sankhyā, A, Vol. 25 (1963). pp. 245-254.

〔7〕 R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering", Bell System Technical Journal, Vol. 37 (1958). pp. 185-282, pp. 458-569 (Dover (1958)).

〔8〕 S. S. L. Chang, "On the filter problem of the power spectrum analyzer", Proc. IRE, Vol. 42 (1954). 1278-1282.

〔9〕 H. E. Daniells, "The estimation of spectral densities", J. R. Statist. Soc., B, Vol. 24 (1962). pp. 185-198.

〔10〕 U. Grenander, "On empirical spectral analysis of stochastic processes", Ark. för Mat., Vol. 1, (1951). pp. 503-531.

〔11〕 U. Grenander and M. Rosenblatt, Statistical Analysis of Stationary Time Series. New York: Wiley (1957)

〔12〕 U. Grenander, "Bandwidth and variance in the estimation of the spectrum", J. R. Statist. Soc., B, Vol. 20 (1958). pp. 152-157.

〔13〕 Z. A. Lomnicki and S. K. Zaremba, "On estimating the spectral density function of a stochastic process", J. R. Statist. Soc., B, Vol. 19 (1957), pp. 13-37. (Discussion pp. 47-63)

〔14〕 Z. A. Lomnicki and S. K. Zaremba, "Bandwidth and resolvability in statistical spectral analysis", J. R. Statist. Soc., B, Vol. 21 (1959). pp. 169-171.

〔15〕 E. Parzen, "On asymptotically efficient consistent estimates of the spectral density function of a stationary time series", J. R. Statist. Soc., Vol. 20 (1958). pp. 303-322.

〔16〕 E. Parzen, "Mathematical considerations in the estimation of spectra", Technometrics, Vol. 3 (1961). pp. 167-190.

〔17〕 M. Rosenblatt, "Remarks on some nonparametric estimates of a density function", Ann. Math. Statistics, Vol. 27 (1956). pp. 832-837.

〔18〕 J. W. Tukey, "The sampling theory of power spectrum estimates", Symposium on Applications of Autocorrelation Analysis to Physical Problems, Woods Hole, Mass., Office of Naval Reseach Washington, D. C. (1949). pp. 47-67.

〔19〕 J. W. Tukey, "Discussion, emphasizing the connection between analysis of variance and spectrum analysis", Technometrics, Vol. 3 (1961). pp. 191-219.