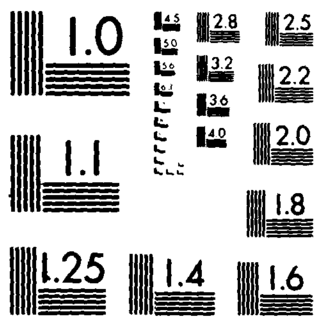
AD-A111 664   CALIFORNIA UNIV SAN DIEGO LA JOLLA DEPT OF MATHEMATICS   F/G 12/1
DECONVOLUTION AND ESTIMATION OF TRANSFER FUNCTION PHASE AND COE--ETC(U)
NOV 81   K S LII, M ROSENBLATT   N00014-81-K-0003

UNCLASSIFIED   NL

END
DATE
FILMED
4-82
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

Deconvolution and estimation of transfer function phase and coefficients for nonGaussian linear processes[1]

by

K. S. Lii and M. Rosenblatt

University of California, Riverside and San Diego

[1]Research supported in part by Office of Naval Research Contract N00014-81-K-0003. Research of the second author was carried out in part at the University of Manchester Institute of Science and Technology.

Key words: nonGaussian, linear process, phase, bispectrum,

asymptotics, deconvolution

Summary

Deconvolution and estimation of transfer function phase and coefficients for nonGaussian linear processes

NonGaussian linear processes are considered. It is shown that the phase of the transfer function can be estimated under broad conditions. This is not true of Gaussian linear processes and in this sense Gaussian linear processes are atypical. The asymptotic behavior of a phase estimate is determined. The phase estimates make use of bispectral estimates. These ideas are applied to a problem of deconvolution which is effective even when the transfer function is not minimum phase. A number of computational illustrations are given.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER --	2. GOVT ACCESSION NO. --	3. RECIPIENT'S CATALOG NUMBER --
4. TITLE (and Subtitle) DECONVOLUTION AND ESTIMATION OF TRANSFER FUNCTION PHASE AND COEFFICIENTS FOR NON-GAUSSIAN LINEAR PROCESSES		5. TYPE OF REPORT & PERIOD COVERED Research 6. PERFORMING ORG. REPORT NUMBER --
7. AUTHOR(s) K. S. Lii and M. Rosenblatt		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0003
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics University of California, San Diego La Jolla, California 92093		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS --
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE November 4, 1981 13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) --		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE --

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

--

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

NonGaussian, linear process, phase, bispectrum asympt· deconvolution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

NonGaussian linear processes are considered. It is shown that the phase of the transfer function can be estimated under broad conditions. This is not true of Gaussian linear processes and in this sense Gaussian linear processes are atypical. The asymptotic behavior of a phase estimate is determined. The phase estimates make use of bispectral estimates. These ideas are applied to a problem of deconvolution which is effective even when the transfer function is not minimum phase. A number of computational illustrations are given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

1. Introduction. Assume that the random variables $v_t, t = \ldots, -1, 0, 1, \ldots$ are independent and identically distributed with mean zero, $Ev_t \equiv 0$, and variance one $Ev_t^2 \equiv 1$. Let $\{\alpha_j\}$ be a sequence of real constants with

$$\sum_{j=-\infty}^{\infty} \alpha_j^2 < \infty \quad .$$

Consider the linear process generated by $\{\alpha_j\}$ and $\{v_t\}$

$$x_t = \sum_{j=-\infty}^{\infty} \alpha_j v_{t-j} \quad . \tag{1}$$

Let $\alpha(z) = \sum_j \alpha_j z^j$ be the z-transform corresponding to the process $\{x_t\}$. Then

$$\alpha(e^{-i\lambda}) = \sum_j \alpha_j e^{-ij\lambda}$$

is called the frequency response function or transfer function. We are concerned with the estimation of $\alpha(e^{-i\lambda})$ on the basis of observations only on the process $\{x_t\}$.

The spectral density of $\{x_t\}$ is

$$f(\lambda) = \frac{1}{2\pi} |\alpha(e^{-i\lambda})|^2 \quad .$$

In the Gaussian case (when $\{x_t\}$ is normally distributed) the full probability structure of $\{x_t\}$ is determined by $f(\lambda)$ or equivalently by the modulus of $\alpha(e^{-i\lambda})$, $|\alpha(e^{-i\lambda})|$. The phase information in $\alpha(e^{-i\lambda})$ is not identifiable in the Gaussian case.

If $\alpha(z)$ is a rational function

$$\alpha(z) = \frac{A(z)}{B(z)}$$

with $A(z)$, $B(z)$ polynomials

$$A(z) = \sum_{k=0}^{q} a_k z^k , \qquad a_0 \neq 0 ,$$

$$B(z) = \sum_{k=0}^{p} b_k z^k , \qquad b_0 = 1 ,$$

the process $\{x_t\}$ is a finite parameter autoregressive moving average process, that is,

$$\sum_{j=0}^{p} b_j x_{t-j} = \sum_{k=0}^{q} a_k v_{t-k} \quad . \tag{2}$$

If $\{x_t\}$ is a Gaussian process satisfying (2), then any root $z_j \neq 0$ of $A(z)$ or $B(z)$ can be replaced by its conjugated inverse $\bar{z}_j^{-1}$ without changing the probability structure of $\{x_t\}$. This follows since $|e^{i\lambda} - z_0| = |z_0|^{-2} |e^{i\lambda} - \bar{z}_0^{-1}|$. If all the roots are distinct there are 2^{p+q} ways of specifying the roots without changing the structure of $\{x_t\}$. To ensure unique determination of the coefficients a_k and b_j of (2) (since there is a different specification of these coefficients corresponding to each of the 2^{p+q} root specifications) in the Gaussian case, it is the custom to assume that all the roots of $A(z)$ and $B(z)$ are outside the unit circle $|z| < 1$ in the complex plane.

However, for a nonGaussian stationary process satisfying (2) (in which case the independent v_t's are nonGaussian) the different 2^{p+q} specifications of roots mentioned above generally correspond to different probability structures and different processes. As a simple example, consider the moving average

$$x_t = 6v_t - 5v_{t-1} + v_{t-2}$$

with the roots o' $A(z)$ 2 and 3, and the moving average

$$y_t = 3v_t - 7v_{t-1} + 2v_{t-2}$$

having a polynomial $A(z)$ with roots 1/2 and 3. Both $\{x_t\}$ and $\{y_t\}$ have the same spectral density but if the independent random sequence $\{v_t\}$ is exponentially distributed, the marginal distributions of the $\{x_t\}$ and $\{y_t\}$ sequences are different. In the problem of deconvolution where one wishes to recover the process $\{v_t\}$ (assumed nonGaussian which is most often the case in applications) in some sense, the proper specification of roots (which are inside and which are outside) becomes crucial (see Rosenblatt (1974)). There is a discussion concerning the distribution of roots as related to prediction problems in Rosenblatt (1980).

There are results on the estimation of the coefficients a_j and b_k of (2) (corresponding to roots assumed outside the unit circle) in Box and Jenkins (1976). In the Gaussian case these are essentially equivalent asymptotically to maximum likelihood procedures. In the nonGaussian case the computations are carried out as if the process were Gaussian. One has a least squares but not a maximum likelihood solution in the nonGaussian case. The coefficients estimated are those corresponding to roots outside the unit circle even though the actual structure of the process may not be one with all the roots outside the unit circle. Thus one will typically not be able to resolve the actual structure using these procedures in the non-Gaussian case. Of course, if one knows the actual nonGaussian distribution of the v_t, one can use the maximum likelihood estimate or an asymptotically equivalent procedure to estimate the coefficients in (2) even if the roots z_j are not all outside the unit circle (see Bhasawa, Feigen and Heyde (1975)). Higher order spectral methods discussed in the next section do not require this knowledge. Our discussion follows that of Rosenblatt (1980).

2. Higher order spectral method

Assume that $\{x_t\}$ is a linear process (see (1)), with the independent random variables $\{v_t\}$ nonGaussian and having all moments finite. Actually we only require that some cumulant γ_k of order $k > 2$ be nonzero. Also let

$$\sum_j |j| |\alpha_j| < \infty$$

and assume

$$\alpha(e^{-i\lambda}) \neq 0$$

for all λ. We will see that $\alpha(e^{-i\lambda})$ is essentially identifiable (as contrasted with the Gaussian case) if one only observes the process $\{x_t\}$.

Since the v_t's are assumed nonGaussian with all moments finite, there must be a cumulant of v_t, $\gamma_k \neq 0$ of smallest subscript $k > 2$. The k^{th} order cumulant spectral density of the process $\{x_t\}$ is given by

$$b_k(\lambda_1, \ldots, \lambda_{k-1}) = \frac{1}{(2\pi)^{k-1}} \sum_{j_1, \ldots, j_{k-1}} \operatorname{cum}(x_t, x_{t+j_1}, \ldots, x_{t+j_{k-1}}) \exp\left(- \sum_{s=1}^{k-1} i j_s \lambda_s\right) \tag{3}$$

$$= \frac{\gamma_k}{(2\pi)^{k-1}} \alpha\left(e^{-i\lambda_1}\right) \cdots \alpha\left(e^{-i\lambda_{k-1}}\right) \alpha\left(e^{i(\lambda_1 + \cdots + \lambda_{k-1})}\right).$$

Let

$$h(\lambda) = \arg\left\{\alpha(e^{-i\lambda}) \frac{\alpha(1)}{|\alpha(1)|}\right\} .$$

Then

$$\left\{\frac{\alpha(1)}{|\alpha(1)|}\right\}^k \gamma_k = (2\pi)^{\frac{k}{2}-1} b_k(0, \ldots, 0) / \{f(0)\}^{\frac{k}{2}} ,$$

and

$$h(\lambda_1) + \cdots + h(\lambda_{k-1}) - h(\lambda_1 + \cdots + \lambda_{k-1})$$

$$= \frac{1}{i} \log\left[(2\pi)^{\frac{k}{2}-1} \left\{\frac{\alpha(1)}{|\alpha(1)|}\right\}^k \gamma_k^{-1} b_k(\lambda_1, \ldots, \lambda_{k-1}) \{f(\lambda_1) \cdots f(\lambda_{k-1}) f(\lambda_1 + \cdots + \lambda_{k-1})\}^{-1/2}\right] \tag{4}$$

since

$$h(-\lambda) = -h(\lambda) \ .$$

Further

$$h'(0) - h'(\lambda) = \lim_{\Delta \to 0} \frac{1}{(k-2)\Delta} \{ h(\lambda) + (k-2)\, h(\Delta) - h(\lambda + (k-2)\Delta) \} \ . \tag{5}$$

Now

$$h(\lambda) = \int_0^{\lambda} \{ h'(u) - h'(0) \} du + c\lambda = h_1(\lambda) + c\lambda$$

where

$$c = h'(0) \ .$$

In particular

$$h(\pi) = h_1(\pi) + c\pi \ .$$

Since the α_j's are real we must have

$$h(\pi) = k\pi$$

for some integer k. Set

$$h_1(\pi)/\pi = \delta \ .$$

Then

$$h(\pi) = k\pi = (\delta + c)\pi$$

so that

$$c = k - \delta \ .$$

The integer k cannot be determined without further assumptions since it corresponds to reindexing or subscripting the v_t's. The sign of $\alpha(1)$ is also intrinsically undecidable since one can multiply α_j's and v_t's by (-1)

without changing the observed process $\{x_t\}$. Thus, under the conditions specified above for a nonGaussian linear process $\{x_t\}$, $\alpha(e^{-i\lambda})$ is identifiable up to the integer k and the sign on the basis of observations on $\{x_t\}$ only and is given by

$$\alpha(e^{-i\lambda}) = |2\pi f(\lambda)|^{1/2} \exp\{ih(\lambda)\}$$

with

$$h(\lambda) = \int_0^\lambda \{h'(u) - h'(0)\}\, du + c\lambda \tag{6}$$

$$= h_1(\lambda) - \frac{h_1(\pi)}{\pi}\lambda \;.$$

Notice that $h_1(\lambda)$ can actually be computed.

3. Phase estimation and convergence of estimates

There are many discussions concerned with the estimation of the second order spectral density $f(\lambda)$ (see Anderson (1971) or Jenkins and Watts (1968)). We will concentrate on the estimation of $h(\lambda)$. For simplicity of discussion we will assume that the third order cumulant γ_3 of v_t is nonzero. The program in the higher order case can be carried out in a similar manner. Equation (5) becomes

$$h'(0) - h'(\lambda) = \lim_{\Delta \to 0} \frac{1}{\Delta}\{h(\lambda) + h(\Delta) - h(\lambda+\Delta)\}$$

when $k = 3$. For (4) we find that up to a sign

$$h(\lambda) + h(\Delta) - h(\lambda+\Delta) = \arg\{b_3(\lambda, \Delta)\} \;.$$

From this point on we will drop the subscript and understand that we are dealing with the bispectral density $b(\lambda,\mu)$. Let ${}_n b(\lambda,\mu)$ be an estimate of $b(\lambda,\mu)$ based on a sample of size n. Then an estimate of

$$\theta(\lambda,\mu) = \arg b(\lambda,\mu)$$

can be given by

$$\theta_n(\lambda,\mu) = \arctan(\mathrm{Im}\ {}_n b(\lambda,\mu)/\mathrm{Re}\ {}_n b(\lambda,\mu)) \quad .$$

We note that for a complex number

$$z = x + iy = re^{i\theta}$$

with $r = |z|$ and $\theta = \arctan(y/x)$ a principal value determination, one has

$$\frac{\partial\theta}{\partial y} = \frac{x}{r^2} \, , \quad \frac{\partial\theta}{\partial x} = -\frac{y}{r^2}$$

$$\frac{\partial^2\theta}{\partial x^2} = \frac{2xy}{r^4} \, , \quad \frac{\partial^2\theta}{\partial y^2} = -\frac{2xy}{r^4}$$

and

$$\frac{\partial^2\theta}{\partial x\,\partial y} = \frac{1}{r^2} - \frac{2x^2}{r^4} = -\frac{1}{r^2} + \frac{2y^2}{r^2} \quad .$$

Therefore

$$\begin{aligned}\theta_n(\lambda,\mu) - \theta(\lambda,\mu) = &-\frac{\mathrm{Im}\ b(\lambda,\mu)}{|b(\lambda,\mu)|^2}\{\mathrm{Re}\ {}_n b(\lambda,\mu) - \mathrm{Re}\ b(\lambda,\mu)\} \\ &+ \frac{\mathrm{Re}\ b(\lambda,\mu)}{|b(\lambda,\mu)|^2}\{\mathrm{Im}\ {}_n b(\lambda,\mu) - \mathrm{Im}\ b(\lambda,\mu)\} \\ &+ o_p({}_n b(\lambda,\mu) - b(\lambda,\mu)) \ . \qquad (7)\end{aligned}$$

Let us consider estimating

$$h(\lambda) = h_1(\lambda) - \frac{h_1(\pi)}{\pi}\lambda \quad .$$

Set $\Delta = \Delta(n)$, $k\Delta = \lambda$, and let $\Delta = \Delta(n) \to 0$ as $n \to \infty$. Assume for convenience that $b(0,0)$ is positive. A simple modification indicated later takes care of the case in which $b(0,0)$ is negative. Now

$$\begin{aligned} h_1(\lambda) &= h(\lambda) - h'(0)\lambda \\ &\cong h(k\Delta) - \frac{h(\Delta)}{\Delta}k\Delta \\ &= -\sum_{j=1}^{k-1}\{h(j\Delta) + h(\Delta) - h((j+1)\Delta)\} \\ &= -\sum_{j=1}^{k-1} \arg b(j\Delta, \Delta) \quad . \end{aligned}$$

This suggests

$$H_n(\lambda) = -\sum_{j=1}^{k-1} \arg {}_n b(j\Delta, \Delta) \tag{8}$$

as an appropriate estimate of $h_1(\lambda)$. Assume that sixth moments are finite and that bispectral estimates ${}_n b(\lambda, \mu)$ of the type obtained by weighted averages of 3^{rd} order periodogram values (see Brillinger and Rosenblatt (1967)) are employed. It has been shown in the paper just cited that if the bispectral density is continuously differentiable up to second order and one has a symmetric bandlimited weight function with bandwidth Δ, that then

$$E\,{}_n b(\lambda,\mu) - b(\lambda,\mu) \sim \iint (uD_\lambda + vD_\mu)^2\, b(\lambda,\mu)\, w(u,v)\, du\, dv\, \Delta^2 + o(\Delta^2) \tag{9}$$

and

$$\sigma^2({}_n b(\lambda,\mu)) \sim \frac{f(\lambda)\, f(\mu)\, f(\lambda+\mu)}{\Delta^2 n} \int w^2(u,v)\, du\, dv \ , \tag{10}$$

if $\Delta^2 n \to \infty$ as $n \to \infty$, $\Delta(n) \to 0$. Here D_λ and D_μ represent partial derivatives with respect to λ and μ respectively. Further estimates ${}_n b(\lambda,\mu)$, ${}_n b(\lambda',\mu')$, $0 \le \mu \le \lambda$, $0 \le \mu' \le \lambda'$, are asymptotically uncorrelated as $n \to \infty$ if (λ,μ) and (λ^1,μ^1) are distinct. Using (7) and (8) we can write

$$H_n(\lambda) - h_1(\lambda) = R_n(\lambda) + o_p(H_n(\lambda) - h_1(\lambda)) \tag{11}$$

and show by employing (9) that

$$\begin{aligned} ER_n(\lambda) \sim \int_0^\lambda -\frac{\operatorname{Im} b(u,0)}{|b(u,0)|^2} [A D_u^2 \operatorname{Re} b(u,0) &+ 2B\, D_u D_v \operatorname{Re} b(u,0) \\ &+ C D_v^2 \operatorname{Re} b(u,0)]\, du\, \Delta \\ + \int_0^\lambda \frac{\operatorname{Re} b(u,0)}{|b(u,0)|^2} [A D_u^2 \operatorname{Im} b(u,0) &+ 2B D_u D_v \operatorname{Im} b(u,0) \\ &+ C D_v^2 \operatorname{Im} b(u,0)]\, du\, \Delta \\ + o(\Delta) \end{aligned}$$

where A, B, and C are the second moments of the weight function w of the bispectral estimates

$$A = \int u^2 w(u,v)\, du\, dv \ , \qquad B = \int uv w(u,v)\, du\, dv$$

$$C = \int v^2 w(u,v)\, du\, dv \ .$$

Further, by using (10) it follows that

$$\operatorname{cov}(R_n(\lambda), R_n(\mu)) \cong \frac{f(0)}{\Delta^3 n} \int_0^{\min(\lambda,\mu)} \{f^2(u)/|b(u,0)|^2\} \, du \quad w^2(u,v)\, du\, dv. \tag{12}$$

Assuming the existence of all moments and (11), one can show that $H_n(\lambda)$ is asymptotically normal with mean $h_1(\lambda)$ and variance given by (12). The mean square error of $R_n(\lambda)$ is of order

$$c_1 \Delta^2 + \frac{c_3}{\Delta^3 n}$$

and the optimal rate of convergence is $n^{-2/5}$ when $\Delta(n) \sim n^{-1/5}$.

Assume that the bispectral density is continuously differentiable up to third order. Further let the weight function of the estimate be band-limited with first and second order moments zero. Such a weight function cannot be nonnegative everywhere. Then

$$E\, {}_n b(\lambda, \mu) - b(\lambda, \mu) = 0(\Delta^3).$$

The mean square error of $R_n(\lambda)$ is of the order

$$c_1 \Delta^4 + \frac{c_2}{\Delta^3 n} \quad .$$

The optimal rate of convergence is then $n^{-4/7}$ with $\Delta(n) \sim n^{-1/7}$.

Generally we will estimate $h_1(\lambda)$ and hence $h(\lambda)$ for a whole range of λ values. The sign of $b(0,0)$ may not be positive. We estimate it by noting the real part of ${}_n b(\Delta, \Delta)$. If it is negative we multiply all ${}_n b(j\Delta, \Delta)$

with a minus sign. The estimate $H_n(\lambda)$ is then given by

$$H_n(\lambda) = - \sum_{j=1}^{k-1} \arg\{ - {}_n b(j\Delta, \Delta)\} \quad .$$

4. Computations using spectral methods

We remark on the computational aspect of phase estimation of $\alpha(e^{-i\lambda})$ and give a few illustrative examples to indicate its effectiveness.

Given a sample $\{x_t\}$ of size $n = kN$, we center and normalize it so that it has mean zero and variance one. Break up the sample into k disjoint subsections of equal length N so that the variance of the bispectral estimate from each section is not too large. Then choose a grid of points $\lambda_j = j\Delta$ in $(0, 2\pi)$, $j = 1, \ldots, M$, $\Delta = 2\pi L/N$. Though the symmetry condition $h(\lambda) = -h(-\lambda)$ implies that one need only deal with λ in $(0, \pi)$, there may be some advantage in considering $\lambda \in (0, 2\pi)$. We will comment on this point later on. Form the bispectral estimate ${}_N b(j\Delta, \Delta)$ of the type discussed above with a weight function of bandwidth Δ from each subsection. Average the estimates from the different subsections so as to arrive at a final estimate ${}_n b(j\Delta, \Delta)$. A detailed discussion of this kind of algorithm can be found in Helland and Lii (1981). Compute $\theta_n(j) = \arg\{{}_n b(j\Delta, \Delta)\} + 2k\pi$ where the integer k is chosen to ensure continuity of $H_n(\ell\Delta) = H_n(\lambda_\ell) = -\sum_{j=1}^{\ell-1} \theta_n(j)$, $\ell = 2, \ldots, M+1$ (neighboring values are as close to each other as possible). Since the upper index is $\ell - 1$ we start with $\ell = 2$. Since $h(0) = 0$ one sets $H_n(0) = 0$ and estimates $H_n(\Delta) = H_n(\lambda_1)$ by an interpolation between 0 and $H_n(\lambda_2)$, $\lambda_2 = 2\Delta$. $H_n(\pi)$ is also computed by an interpolation procedure. This amounts to a complete procedure for estimating $h(\lambda)$.

Since

$$\alpha_k = \frac{1}{2\pi}\int_0^{2\pi} \alpha(e^{-i\lambda})\, e^{ik\lambda}\, d\lambda$$

and estimate $\hat{\alpha}_k$ of α_k is given by

$$\hat{\alpha}_k = \frac{1}{2\pi}\int_0^{2\pi} \hat{\alpha}(e^{-i\lambda})\, e^{ik\lambda}\, d\lambda$$

$$\cong \frac{1}{2\pi(M+2)} \sum_{j=0}^{M+1} \sqrt{2\pi f_n(\lambda_j)}\, \exp\left\{ i\left(H_n(\lambda_j) - \frac{H_n(\pi)}{\pi}\lambda_j + k\lambda_j \right)\right\} \tag{13}$$

and this computation can be carried out by using the fast Fourier transform.

The α_k's are real numbers and so the $\hat{\alpha}_k$'s may or may not be real. If the symmetric property of $f(\lambda)$ and $h(\lambda)$ is used and the integration is carried out from $-\pi$ to π almost real $\hat{\alpha}_k$'s will be obtained. The imaginary part of the α_k's will only be the size of rounding errors. In practice there is no indication of how good or bad the estimates are apart from asymptotic results. In actual practice $\{j\Delta\}_{j=0}^{M+1}$ may not be symmetric about π. If the estimates $H_n(\lambda_j)$ are reasonably good the estimated $\hat{\alpha}_j$'s from (13) should still be almost real. The size of the imaginary part reflects the level of noise. When the estimates $H_n(\lambda_j)$ are not good the imaginary part of the $\hat{\alpha}_k$'s becomes comparable to (or larger than) its real part. This can serve as a direct indication of the quality of the estimation.

If the linear process is one-sided with a finite number of parameters one has a moving average of order q

$$x_t = \sum_{j=0}^{q} \alpha_j v_{t-j}, \qquad \alpha_0 \neq 0 .$$

We could estimate $\alpha(z) = \sum_{j=0}^{q} \alpha_j z^j$ by $\hat{\alpha}(z) = \sum_{j=0}^{q} \hat{\alpha}_j z^j$. In deconvolution we try to recover the process $\{v_t\}$, $v_t = \frac{1}{\alpha(B)} x_t$ (B is the backward shift operator so that $B^j x_t = x_{t-j}$) by computing the approximation $\hat{v}_t = \frac{1}{\hat{\alpha}(B)} x_t$. If all the roots of $\alpha(z)$ (and $\hat{\alpha}(z)$) are outside the unit circle (the frequency function is minimum delay) then $\hat{\alpha}^{-1}(z)$ has a one-sided expansion $\sum_{j=0}^{\infty} \alpha_j' B^j$. In the computation, the series is truncated after a certain number of terms. If some of the roots of $\hat{\alpha}(z)$ have modulus less than one we can still expand $\hat{\alpha}^{-1}(B)$ with a Laurent series expansion. Once the roots of $\hat{\alpha}(z)$ are computed, one can easily get the Laurent series expansion of $\hat{\alpha}^{-1}(B)$ by partial fractions as described in Rosenblatt (1974) or Henrici (1974).

Another way to find the inverse weights in deconvolution is to use a least squares criterion as described in Wiggins (1978). Another general method of deconvolution will be mentioned in the section on computation.

5. Other possible computational methods

We briefly discuss two other possible methods of estimating the coefficients of a nonGaussian moving average process of order q.

$$x_t = \sum_{j=0}^{q} \alpha_j v_{t-j} \quad . \tag{14}$$

As noted earlier, second order moments will not allow us to determine the location of the roots of

$$\alpha(z) = \sum_{j=0}^{q} \alpha_j z^j \quad . \tag{15}$$

Higher order moments will be used in the first method which makes use of a <u>least squares procedure</u>. Assume $Ev_t \equiv 0$, $Ev_t^3 \equiv \gamma \neq 0$. Consider

$$\begin{aligned} c_k &= E\, x_t x_{t+k}^2 \\ &= \gamma \sum_j \alpha_j \alpha_{j+k}^2 , \quad k = -q, -q+1, \ldots, q \quad . \end{aligned} \tag{16}$$

Estimate c_k by

$$\hat{c}_k = \frac{1}{n} \sum_{k=1}^{n} x_t x_{t+k}^2$$

and solve the extremal problem

$$\min_{a_j} \sum_{k=-q}^{q} \left(\hat{c}_k - \gamma \sum_{\ell} a_\ell a_{\ell+k}^2 \right)^2 . \tag{17}$$

There are $q+2$ unknowns $a_0, \ldots, a_q$ and γ in (17). Due to the homogeneity of the α_j's we have to normalize the problem appropriately; all the α_j's can be multiplied by a constant $c \neq 0$ and γ can be divided by c^3 without changing (17). There are a number of ways of carrying out such a normalization. One could set $E v_t^3 = \gamma = 1$. Alternatively $a_0 = 1$ could be the normalization condition. Some comments on the asymptotic distribution of the $\hat{c}_k$'s are given in Appendix 2.

The second method is a searching procedure. One uses a typical second order method to estimate the roots of $\alpha(z)$, r_j, $j = 1, \dots, q$, assuming all the roots have modulus greater than one. An accurate estimate of the distribution of roots is obtained by taking the conjugated inverse of an appropriate number of the r_j's. Suppose all of the r_j's are real and distinct. Then there are 2^q possible sets of roots that give the same second order structure. Each of these sets yields a distinct set of $\hat{\alpha}_j$'s which in turn lead to a distinct set of the c_k's. Choose the set of $\hat{\alpha}_j$'s which determine the set of c_k's minimizing

$$\sum_{k=-q}^{q} (\hat{c}_k - c_k)^2$$

among all the possible sets of $\{\hat{\alpha}_j\}$. If some of the roots r_j are complex, the inverse complex conjugates are taken in pairs. If there are multiple roots, the solution of roots in terms of coefficients is unstable. Some comments on this question are made in Appendix 1. The initial set of coefficient estimates corresponding to roots all outside the unit circle can be obtained by the method described in Box and Jenkins (1976). Alternatively, one could try to obtain the roots directly by solving for the roots of the polynomial

$$p(z) = z^q g(z)$$

where

$$g(z) = \alpha(z)\,\alpha(z^{-1})$$

$$= \sum_{j=-q}^{q} \beta_j z^j$$

with

$$\beta_{|j|} = E\, x_t x_{t+|j|}$$

$$= \sum_{\ell} \alpha_\ell \alpha_{\ell+|j|} \quad .$$

We estimate β_j by

$$\hat{\beta}_j = \frac{1}{n} \sum_{t=1}^{n-|j|} x_t x_{t+|j|} \quad .$$

The roots of $p(z)$ with modulus greater than one are the initial set of roots.

6. Examples

We will consider a few simple examples generated by Monte Carlo simulation to illustrate the computation and to give a qualitative feeling of the effectiveness of the theory. Details and possible "fine tuning" of the computational method will be considered elsewhere.

We generate $x_t = v_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2}$, $t = 1, \ldots, 640$ where $v_t = v_t' - 1$ and v_t''s are independent exponentially distributed random deviates with mean one obtained from the GGZEN subroutine in the international Mathematical and Statistical Library (IMSL). Then

$$E v_t = 0 \ , \quad \text{[illegible]} \quad E v_t^2 = 1 \ ,$$

$$E v_t^3 = 2 \ .$$

We partition $\{x_t\}_{t=1}^{640}$ into five sections, each of which has 128 points. Compute the bispectrum estimate $b_{128}^{(i)}(j\Delta, \Delta)$, $j = 1, 13$; $i = 1, \ldots, 5$, by the algorithm described in Lii and Helland (1981). Here we set $\Delta = \frac{18\pi}{128} = 0.442$. Our final bispectrum estimate is

$$b_{640}(j\Delta, \Delta) \doteq \frac{1}{5} \sum_{i=1}^{5} b_{128}^{(i)}(j\Delta, \Delta) \quad .$$

Compute

$$\hat{\theta}_n(j) = \arg\{\hat{b}_n(j\Delta, \Delta)\}$$

$$= \arctan(\operatorname{Im} \hat{b}_n(j\Delta, \Delta) / \operatorname{Re} \hat{b}_n(j\Delta, \Delta))$$

by taking the principal value as well as

$$-H_n(j\Delta) \equiv -H_n(\lambda_j) = \sum_{i=1}^{j-1} \theta_n(i) \ , \ j = 2, \ldots, 14 \quad .$$

Let $$H_n(0) \equiv H_n(\lambda_0) = 0 ,$$

$$H_n(\Delta) \equiv H_n(\lambda_1) = \frac{1}{2} H_n(\lambda_2) ,$$

$$H_n(\pi) = \frac{1}{2}[H_n(7\Delta) + H_n(8\Delta)]$$

$$= \frac{1}{2}[H_n(3.093) + H_n(3.534)]$$

and $$\hat{c} = -H_n(\pi)/\pi = -\delta .$$

Recall that $\hat{c}$ is an estimate of

$$c = \lim_{\Delta \to 0} \frac{h(\Delta)}{\Delta} = h'(0)$$

up to an integer. We will use $c = h(\Delta)/\Delta$ instead of $h'(0)$ to compare with $\hat{c}$ in the following examples. From formula (9) we compute $\hat{\alpha}_k$'s. A standard smoothed periodogram with uniform weights and bandwidth Δ was used to compute $f_n(\lambda)$ as an estimate of the spectrum $f(\lambda)$ of $\{x_t\}$ ($f_n(0)$ is obtained by a linear extrapolation). These examples are as follows:

Model: $x_t = v_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2}$ with four cases specified given below:

Case	Coefficients			Roots	
	α_0	α_1	α_2	r_1	r_2
1	1.0	-0.833	0.167	2.0	3.0
2	1.0	-2.333	0.667	0.5	3.0
3	1.0	-3.50	1.50	2.0	0.333
4	1.0	-5.0	6.0	0.5	0.333

Case 1. $c = 1.29$, $\hat{c} = 1.605$

λ	Length $\lvert\alpha(e^{-i\lambda})\rvert$	Est. Length	Argument $h(\lambda)$	Argument by Sum $-H_n(\lambda)$	Argument at J $\theta_n(\lambda)$	Adjusted $H_n(\lambda) + \hat{c}\lambda$
0.	0.3333	0.181	0.	0.	0.	0.
0.442	0.4194	0.399	0.5732	0.3125	0.	C.3966
0.884	0.6509	0.617	0.8309	0.6249	0.6249	0.7933
1.325	0.9776	1.050	0.8428	1.3195	0.6945	0.8078
1.767	1.3393	1.452	0.7180	0.9188	0.5994	0.9176
2.209	1.6685	1.663	0.5199	2.7462	0.8273	0.7994
2.651	1.9032	1.715	0.2830	3.5893	0.8431	0.6654
3.093	1.9990	2.241	0.0286	4.4842	0.8949	0.4795
3.534	1.9376	1.740	-0.2274	5.6009	1.1166	0.0720
3.976	1.7307	1.702	-0.4697	6.2542	0.6534	0.1277
4.418	1.4176	1.455	-0.6788	7.1482	0.8939	-0.0571
4.860	110573	1.174	-0.8240	7.9014	0.7532	-0.1012
5.301	0.7172	0.655	-0.8502	8.1520	0.2506	0.3573
5.743	0.4599	0.445	-0.6584	9.3127	1.1607	-0.0943
6.185	0.3377	0.235	-0.1461	9.4044	0.0917	0.5230

$\hat{\alpha}_0 = 0.9593$, $\hat{\alpha}_1 = -0.5816$, $\hat{\alpha}_3 = 0.1158$

Here and from this point on all $\hat{\alpha}_k$'s are adjusted by sign and index shift.

Case 2. $c = -1.388$, $\hat{c} = 0.4094$

λ	Length $\lvert\alpha(e^{-i\lambda})\rvert$	Est. Length	Argument $h(\lambda)$	Argument by Sum $-H_n(\lambda)$	Argument at J $\theta_n(\lambda)$	Adjusted $H_n(\lambda)+\hat{c}\lambda$	$H_n(\lambda)$ $+(\hat{c}-1)\lambda$
0.	0.6667	0.473	0.0000	0.	0.	0.	0.
0.442	0.8389	0.894	-0.6125	-0.2713	0.	0.0904	-0.352
0.884	1.3018	1.315	-1.0828	-0.5425	-0.5425	0.1808	-0.703
1.325	1.9553	1.888	-1.4915	-0.4851	0.0574	-0.0575	-1.382
1.767	2.6785	2.748	-1.8895	-0.6904	-0.2053	-0.0331	-1.80
2.209	3.3369	3.348	-2.2893	-0.6481	0.0423	-0.2563	-2.465
2.651	3.8064	3.520	-2.6920	-1.0097	-0.3615	-0.0756	-2.727
3.093	3.9980	3.809	-3.0966	-1.0884	-0.0787	-0.1778	-3.27
3.534	3.8752	3.951	-3.5014	-1.4842	-0.3958	0.0371	-3.497
3.976	3.4614	3.297	-3.9047	-1.6797	-0.1956	0.0518	-3.924
4.418	2.8351	2.802	-4.3051	-1.5538	0.1259	-0.2550	-4.693
4.860	2.1146	2.168	-4.7031	-1.8428	-0.2890	-0.1469	-5.007
5.301	1.4344	1.429	-5.1070	-1.8785	-0.0357	-0.2921	-5.593
5.743	0.9199	0.936	-5.5558	-2.3478	-0.4693	-0.0037	-5.747
6.185	0.6755	0.443	-6.1366	-2.6185	-0.2706	0.0861	-6.099

$$\hat{\alpha}_0 = 0.7164, \quad \hat{\alpha}_1 = -2.175, \quad \hat{\alpha}_3 = 0.7605$$

As an indication of discretization error, notice that if we use the exact $h(\lambda)$ and $\lvert\alpha(e^{-i\lambda})\rvert$ instead of estimated ones, we get

$$\hat{\alpha}_0 = 0.9136, \quad \hat{\alpha}_1 = -2.247, \quad \hat{\alpha}_3 = 0.5977$$

Case 3. $c = -0.613$, $\hat{c} = 0.6363$

λ	Length $\lvert\alpha(e^{-i\lambda})\rvert$	Est. Length	Argument $h(\lambda)$	Argument by Sum $-H_n(\lambda)$	Argument at J $\theta_n(\lambda)$	Adjusted $H_n(\lambda)+\hat{c}\lambda$	$H_n(\lambda)$ $+(\hat{c}-1)\lambda$
0.	1.0000	0.055	0	0	0.	0	0.
0.442	1.2583	1.280	-0.2711	0.0062	0.	0.2749	-0.167
0.884	1.9527	2.505	-0.6843	0.0125	0.0125	0.5498	-0.334
1.325	2.9330	2.922	-1.1592	0.3266	0.3141	0.5168	-0.808
1.767	4.0178	4.386	-1.6448	0.5819	0.2553	0.5426	-1.224
2.209	5.0054	5.272	-2.1286	1.0098	0.4280	0.3958	-1.813
2.651	5.7097	6.732	-2.6094	1.1805	0.1707	0.5062	-2.145
3.093	5.9971	5.440	-3.0884	1.8244	0.6439	0.1435	-2.95
3.534	5.8129	6.670	-3.5672	2.1738	0.3494	0.0752	-3.459
3.976	5.1922	5.610	-4.0475	2.5575	0.3836	-0.0273	-4.003
4.418	4.2527	4.653	-4.5306	2.8902	0.3328	-0.0790	-4.497
4.860	3.1720	3.040	-5.0162	2.9963	0.1061	0.0961	-4.764
5.301	2.1516	2.630	-5.4959	3.3573	0.3610	0.0161	-5.215
5.743	1.3799	1.456	-5.9306	3.3021	-0.0552	0.3525	-5.391
6.185	1.0132	0.282	-6.2334	3.3388	0.0367	0.5970	-5.588

$\hat{\alpha}_0 = 0.7561$, $\hat{\alpha}_1 = -3.334$, $\hat{\alpha}_2 = 1.778$

Case. 4. $c = -3.29$, $\hat{c} = -1.18$

λ	Length $\lvert\alpha(e^{-i\lambda})\rvert$	Est. Length	Argument $h(\lambda)$	Argument by Sum $-H_n(\lambda)$	Argument by J $\theta_n(\lambda)$	Adjusted $H_n(\lambda)+\hat{c}\lambda$	$H_n(\lambda)$ $+(\hat{c}-2)\lambda$
0.	2.0000	0.969	0	0	0.	0	0.
0.442	2.5167	2.403	-1.4567	0.0542	0.	-0.5757	-1.4597
0.884	3.9053	3.837	-2.5981	0.1085	0.1085	-1.1514	-2.9194
1.325	5.8659	5.832	-3.4935	-0.5056	-0.6140	-1.0589	-3.7089
1.767	8.0356	6.112	-4.2523	-1.0588	-0.5533	-1.0270	-4.561
2.209	10.0109	9.601	-4.9377	-1.7068	-0.6480	-0.9006	-5.3186
2.651	11.4194	10.311	-5.5844	-2.5605	-0.8537	-0.5683	-5.8903
3.093	11.9942	10.165	-6.2136	-3.2706	-0.7101	-0.3797	-6.5657
3.534	11.6258	10.822	-6.8412	-4.1459	-0.8753	-0.0259	-7.0939
3.976	10.3843	9.952	-7.4825	-4.9293	-0.7834	0.2360	-7.716
4.418	8.5055	6.902	-8.1569	-5.6188	-0.6895	0.4040	-8.432
4.860	6.3440	5.925	-8.8953	-6.0445	-0.4257	0.3083	-9.412
5.301	4.3033	4.360	-9.7527	-6.6180	-0.5735	0.3603	-10.242
5.743	2.7597	2.560	-10.8280	-6.7895	-0.1716	0.0104	-11.495
6.185	2.0265	0.760	-12.2239	-6.8955	-0.1060	-0.4051	-12.775

$$\hat{\alpha}_0 = 0.9603, \quad \hat{\alpha}_1 = -3.904, \quad \hat{\alpha}_2 = 4.966$$

When we increase the number of parameters to four, we observe qualitatively the same type of result as in the three parameter case. But generally speaking, the instability is increased. We give one example: the model is

$$x_t = v_t - 4.25\, v_{t-1} + 4.75\, v_{t-2} - 0.938\, v_{t-1}$$

with roots $\frac{2}{3}$, $\frac{1}{2.5}$, 4.0, we get:

$c = -3.6887$, $\hat{c} = -1.825$

λ	Length $\lvert\alpha(e^{-i\lambda})\rvert$	Est. Length	Argument $h(\lambda)$	Argument by Sum $-H_n(\lambda)$	Argument by J $\theta_n(\lambda)$	Adjusted $H_n(\lambda)+\hat{c}\lambda$	$H_n(\lambda)$ $+(\hat{c}-2)\lambda$
0.	0.5625	0.	0.	0.	0.	0.	0.
0.442	0.9470	0.925	-1.6304	-0.3495	0.	-0.4569	-1.341
0.884	2.0233	2.202	-2.6631	-0.6990	-0.6990	-0.9138	-2.682
1.325	3.7839	3.439	-3.4612	-1.1051	-0.4062	-1.3140	-3.964
1.767	6.0495	6.768	-4.1782	-2.1441	-1.0389	-1.0814	-4.615
2.209	8.3719	8.149	-4.8647	-3.2201	-1.0760	-0.8118	-5.23
2.651	10.1662	11.574	-5.5391	-4.0550	-0.8349	-0.7832	-6.085
3.093	10.9296	10.548	-6.2089	-5.0367	-0.9818	-0.6079	-6.794
3.534	10.4383	11.975	-6.8783	-6.4316	-1.3949	-0.0193	-7.087
3.976	8.8369	8.413	-7.5511	-7.2046	-0.7730	-0.0527	-8.005
4.418	6.5805	7.820	-8.2339	-8.1175	-0.9129	0.0538	-8.782
4.860	4.2557	3.946	-8.9415	-8.8852	-0.7677	0.0152	-9.705
5.301	2.3572	2.518	-9.7138	-9.7556	-0.8704	0.0792	-10.[illegible]
5.743	1.1270	1.160	-10.6702	-10.0662	-0.3106	-0.4166	-11.9
6.185	0.5829	0.	-12.1458	-10.7554	-0.6893	-0.5337	-12.0

$\hat{\alpha}_0 = 1.106$, $\hat{\alpha}_1 = -3.572$, $\hat{\alpha}_2 = 4.293$, $\hat{\alpha}_3 = -0.9762$

These simple examples indicate that one can estimate the unknown coefficients reasonably well and one is able to discriminate different models even though they have the same spectral structure. The question of determining how many roots are inside of a unit circle can be answered very reliably using this method. One simply takes the absolute value of an estimate $H_n(\lambda) + c(\lambda)$ of $h(\lambda)$ near $\lambda = 2\pi$ and divides it by 2π, rounding the result to its nearest integer. This integer is the winding number given by $\alpha(e^{-i\lambda})$ which gives the number of zeros of $\alpha(z)$ inside of a unit circle. One can see that this is clearly the case in all these examples.

Graphs 1 through 5 compare the theoretical $h(\lambda)$ with the estimate obtained by using our techniques in the five successive cases considered. Graphs 6 through 10 are concerned with deconvolution. We consider the moving average

$$x_t = v_t - 2.333\, v_{t-1} + 0.667\, v_{t-2}$$

with the v_t's independent exponential variables with mean one. The v_t's are generated as pseudo-random variates. $F(x)$ is the exponential distribution function with mean one. Graph 6 is a plot of the sample distribution function of the generated v_t's and of $F(x)$. Let $\alpha(B) = 1 - 2.333B + 0.667\, B^2$. $\hat{v}_t$ is obtained by truncating the exact deconvolution formula as applied to the generated time series

$$\hat{v}_t = \sum_{j=-9}^{9} a_j\, x_{t-j} \cong \frac{1}{\alpha(B)}\, x_t\,. \tag{18}$$

Graph 7 has a plot of the sample distribution function of the $\hat{v}_t$'s with $F(x)$. The estimated model (using our bispectral techniques) is

$$x_t = 0.7164\, v_t - 2.175\, v_{t-1} + 0.7605\, v_{t-2} .$$

Let $\alpha(B) = 0.7164 - 2.175\,B - 0.7605\,B^2$. Let $\hat{\hat{v}}_t = \frac{1}{\hat{\alpha}(B)}\, x_t$ with the same truncation as that employed in (18). Graph 8 is a plot of the sample distribution function of the $\hat{\hat{v}}_t$ with $F(x)$. In these three graphs a one sample 95% confidence band using the Kolmogorov-Smirnov statistic is indicated. Graph 9 has plots of the sample distribution functions of the $\hat{v}_t$'s and v_t's respectively. Graph 10 plots the sample distribution functions of the $\hat{\hat{v}}_t$'s and the v_t's. Two sample 95% confidence bands are given on these two graphs.

A general way to find the deconvolution weights can be described as follows. We have an estimate.

$$\hat{\alpha}(e^{-i\lambda}) = \sqrt{2\pi \hat{f}_n(\lambda)}\, \exp\{i(H_n(\lambda) + \hat{c}\lambda)\} .$$

To find $\hat{\alpha}^{-1}(e^{-i\lambda})$ we compute $b(e^{-i\lambda}) = \exp\{-\ln(\hat{\alpha}(e^{-i\lambda}))\}$. Then

$$\hat{\omega}_k = \frac{1}{2\pi}\int_0^{2\pi} b(e^{-i\lambda})\, e^{ik\lambda}\, d\lambda$$

give the coefficients of series expansion of $\hat{\alpha}^{-1}(e^{-i\lambda})$ which are the weights desired in deconvolution. This method is general. We do not require knowledge of the order of the moving average process and there is no need to compute the roots from the estimated coefficients and how the roots are distributed is irrelevant.

References

Anderson, T. W. (1971). The Spectral Analysis of Time Series. John Wiley, New York.

Bhasawa, I. V., Feigen, P. D. and Heyde, C. C. (1975). Asymptotic properties of maximum likelihood estimates for stochastic processes. Sankhya, Series A, 38, 259-270.

Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control. Holden-Day,

Brillinger, D. and Rosenblatt, M. (1967). Asymptotic theory of estimates of kth order spectra. In Spectral Analysis of Time Series (ed. B. Harris) 153-188. John Wiley, New York.

Henrici, P. (1974). Applied and Computational Complex Analysis, vol. 1, John Wiley, New York.

Jenkins, G. M. and Watts, D. G. (1968). Spectral Analysis and Its Applications. Holden-Day.

Lii, K. S. and Helland, K. N. (1981). Cross-bispectrum computation and variance estimation. ACM Trans. Math. Software 7, #3.

Rosenblatt, M. (1974). Random Processes. Springer,

Rosenblatt, M. (1980). Linear processes and bispectra. J. Appl. Prob., 17, 265-270.

Wiggins, R. A. (1978). Minimum entropy deconvolution. Geoexploration 16.

Appendix 1

In this appendix we consider the relationship between the coefficients of a polynomial and the roots of the polynomial, at least locally. The polynomial is

$$\sum_{j=0}^{p} a_j z^j = \prod_{j=1}^{p} (z - z_j)$$

where the roots are z_j, $j = 1, \ldots, p$, and the coefficients a_j, $j = 0, 1, \ldots, p$, with $a_p = 1$. It is well known that

$$a_{p-1} = -\sum_j z_j \quad ,$$

$$a_{p-2} = \sum_{j \neq k} z_j z_k$$

$$a_{p-3} = -\sum_{j \neq k \neq \ell} z_j z_k z_\ell$$

$$\cdots$$

$$a_0 = (-1)^p z_1 \ldots z_p \quad .$$

Let us consider the relationship between the differentials of the coefficients a_j and the differentials of the roots z_k. Now

$$\frac{\partial q_{p-1}}{\partial z_\ell} = -1 \; ,$$

$$\frac{\partial a_{p-2}}{\partial z_\ell} = \sum_{j \neq \ell} z_j = -a_{p-1} - z_\ell \quad ,$$

$$\frac{\partial a_{p-3}}{\partial z_\ell} = - \sum_{j \neq k \neq \ell} z_j z_k = -a_{p-2} - 2a_{p-1} z_\ell - 2z_\ell^2 \quad ,$$

$\ell = 1, \ldots, p$. Thus

$$da_{p-1} = - \sum_j dz_j \quad ,$$

$$da_{p-2} = -a_{p-1} \sum_j dz_j - \sum_j z_j dz_j$$

$$da_{p-3} = -a_{p-2} \sum_j dz_j - 2a_{p-1} \sum_j z_j dz_j - 2 \sum_j z_j^2 dz_j$$

$$\ldots$$

and this can be written in matrix notation as

$$\begin{pmatrix} da_{p-1} \\ da_{p-2} \\ da_{p-3} \\ \vdots \end{pmatrix} = UV \begin{pmatrix} dz_1 \\ dz_2 \\ dz_3 \\ \vdots \end{pmatrix}$$

where

$$U = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -a_{p-1} & -1 & 0 & 0 \\ & -2a_{p-1} & -2 & 0 \\ & \ldots & & \end{pmatrix}$$

and

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ z_1 & z_2 & z_3 & \cdots & z_n \\ z_1^2 & z_2^2 & z_3^2 & \cdots & z_n^2 \\ & & \cdots & & \end{pmatrix} .$$

U is a triangular nonsingular matrix and V is the Vander Monde matrix.

V is nonsingular as long as the roots z_i are distinct.

Appendix 2

The object is to remark on some aspects of the asymptotic behavior of estimates of the c_k's in the context of a general linear process. Let x_t be a nonGaussian linear process (1) with

$$E\, v_t \equiv 0\,, \quad E\, v_t^2 \equiv 1\,, \quad E\, v_t^3 \equiv \gamma_3$$

$$\mathrm{cum}(v_t^4) = \gamma_4\,, \quad \mathrm{cum}(v_t^6) = \gamma_6 \quad .$$

Set

$$y_t = x_t^2 \quad .$$

For convenience we introduce

$$g_u = \mathrm{cov}(x_t\,,\ y_{t-u})\ ,$$

$$r_u = \mathrm{cov}(x_t\,,\ x_{t-u})\ ,$$

$$h_u = \mathrm{cov}(y_t\,,\ y_{t-u})\ .$$

Consider the estimates

$$\hat{g}_a = \frac{1}{N} \sum_{t=1}^{N} x_t\, y_{t+a}$$

of $E\, x_t\, y_{t+a}$. Then

$$\mathrm{cov}(\hat{g}_a, \hat{g}_b) = \frac{1}{N^2} \sum_{t,\tau=1}^{N} \mathrm{cov}(x_t\, y_{t+a}\,,\ x_\tau\, y_{\tau+b})$$

where

$$\operatorname{cov}(x_t y_{t+a},\ x_\tau y_{\tau+b})$$

$$= r_{t-\tau} c_{t-\tau+a-b} + g_{t-\tau-b}\, g_{t-\tau+a}$$

$$+ \operatorname{cum}(x_t,\ x_\tau,\ y_{\tau+b}) \operatorname{cum}(y_{t+a})$$

$$+ \operatorname{cum}(x_t,\ x_\tau,\ y_{t+a}) \operatorname{cum}(y_{t+b})$$

$$+ \operatorname{cum}(x_t,\ y_{t+a},\ x_\tau,\ y_{\tau+b}) \ .$$

Now

$$r_u = \sum_j \alpha_j\, \alpha_{j-u} \quad ,$$

$$g_u = \gamma_3 \sum_j \alpha_j\, \alpha^2_{j-u} \quad ,$$

$$h_u = \gamma_4 \sum \alpha_j^2\, \alpha^2_{j-u} + 2\left(\sum_j \alpha_j\, \alpha_{j-u}\right)^2 \quad .$$

Further

$$\operatorname{cum}(y_{t+a}) = \sum_j \alpha_j^2 \quad ,$$

$$\operatorname{cum}(x_t,\ x_\tau,\ y_{\tau+b}) = \gamma_4 \sum_j \alpha_j\, \alpha^2_{j+b}\, \alpha_{j+(t-\tau)} = k(b, t-\tau) \quad ,$$

$$\operatorname{cum}(x_t,\ x_\tau,\ y_{t+a}) = \gamma_4 \sum_j \alpha_j\, \alpha^2_{j+a}\, \alpha_{j+(\tau-t)} = k(a, \tau-t)^j \quad ,$$

$$\operatorname{cum}(x_t,\ y_{t+a},\ x_\tau,\ y_{\tau+b}) = \gamma_6 \sum_u \alpha_u\, \alpha_{u+(\tau-t)}\, \alpha^2_{u+a}\, \alpha^2_{u+b+(\tau-t)}$$

$$+ 4\gamma_4 \left(\sum_u \alpha_u\, \alpha_{u+(\tau-t)}\, \alpha_{u+a}\, \alpha_{u+b+(\tau-t)}\right)\left(\sum_u \alpha_u\, \alpha_{u+(\tau-t)+b-a}\right)$$

$$+ 2\gamma_3^2 \left(\sum_u \alpha_u \alpha_{u+(\tau-t)} \alpha_{u+a} \right) \left(\sum \alpha_u^2 \alpha_{u+(t-\tau)+a-b} \right)$$

$$+ 2\gamma_3^2 \left(\sum_n \alpha_u \alpha_{u+(t-\tau)} \alpha_{u+b} \right) \left(\sum \alpha_u^2 \alpha_{u+(\tau-t)+b-2} \right)$$

$$= k(a, b, t-\tau) .$$

It is then clear that

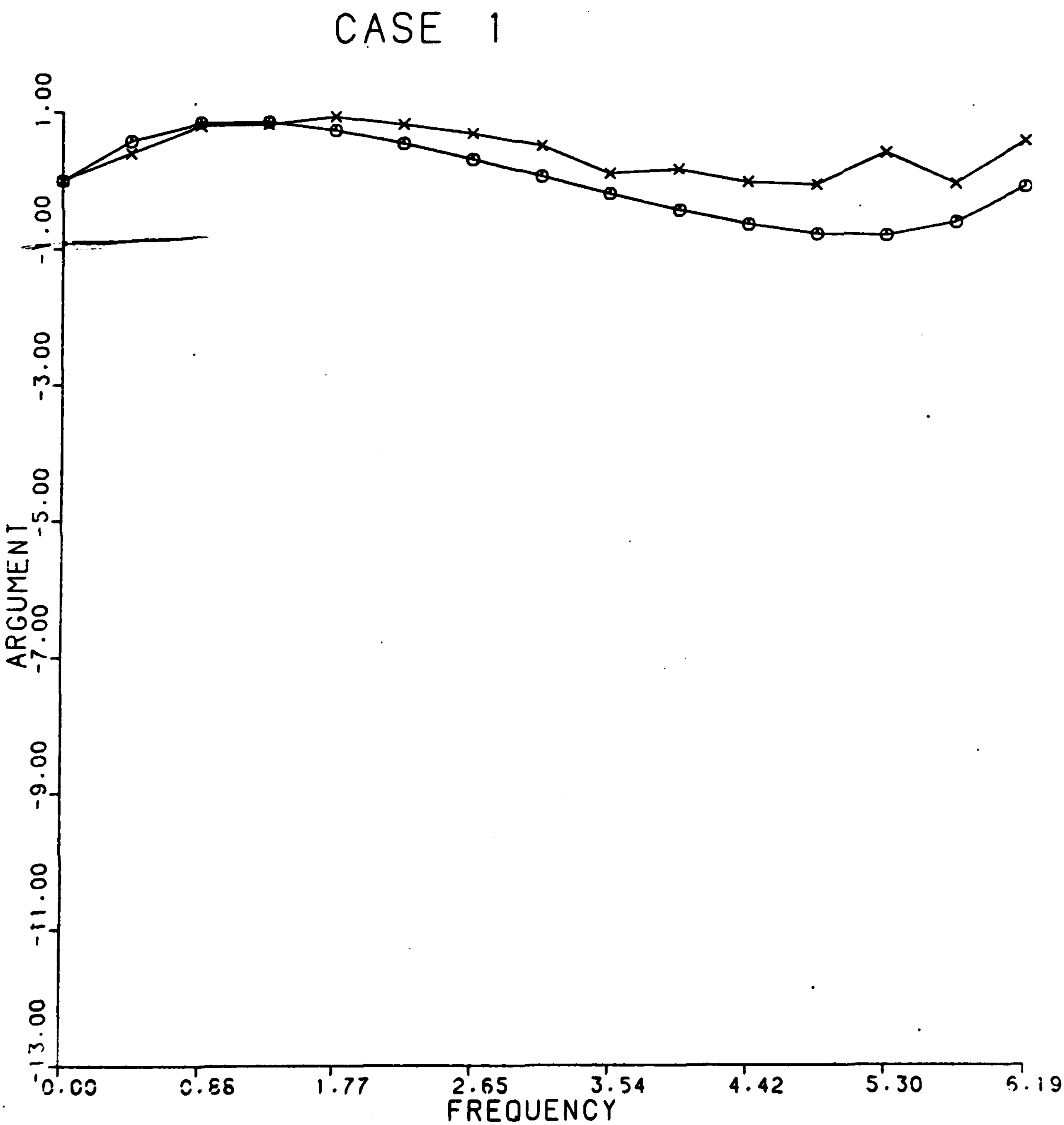
$$\lim_{N \to \infty} N \operatorname{cov}(\hat{g}_a, \hat{g}_b) = \sum_s r_s c_{s+a-b} + \sum_s g_{s-b} g_{s+a}$$

$$+ r_0 \sum_s \{k(b, s) + k(a, s)\} + \sum_s k(a, b, s) .$$

Under the assumption that $\Sigma | \alpha_j | < \infty$, with a truncation argument like that employed in Anderson (1970), one can show that

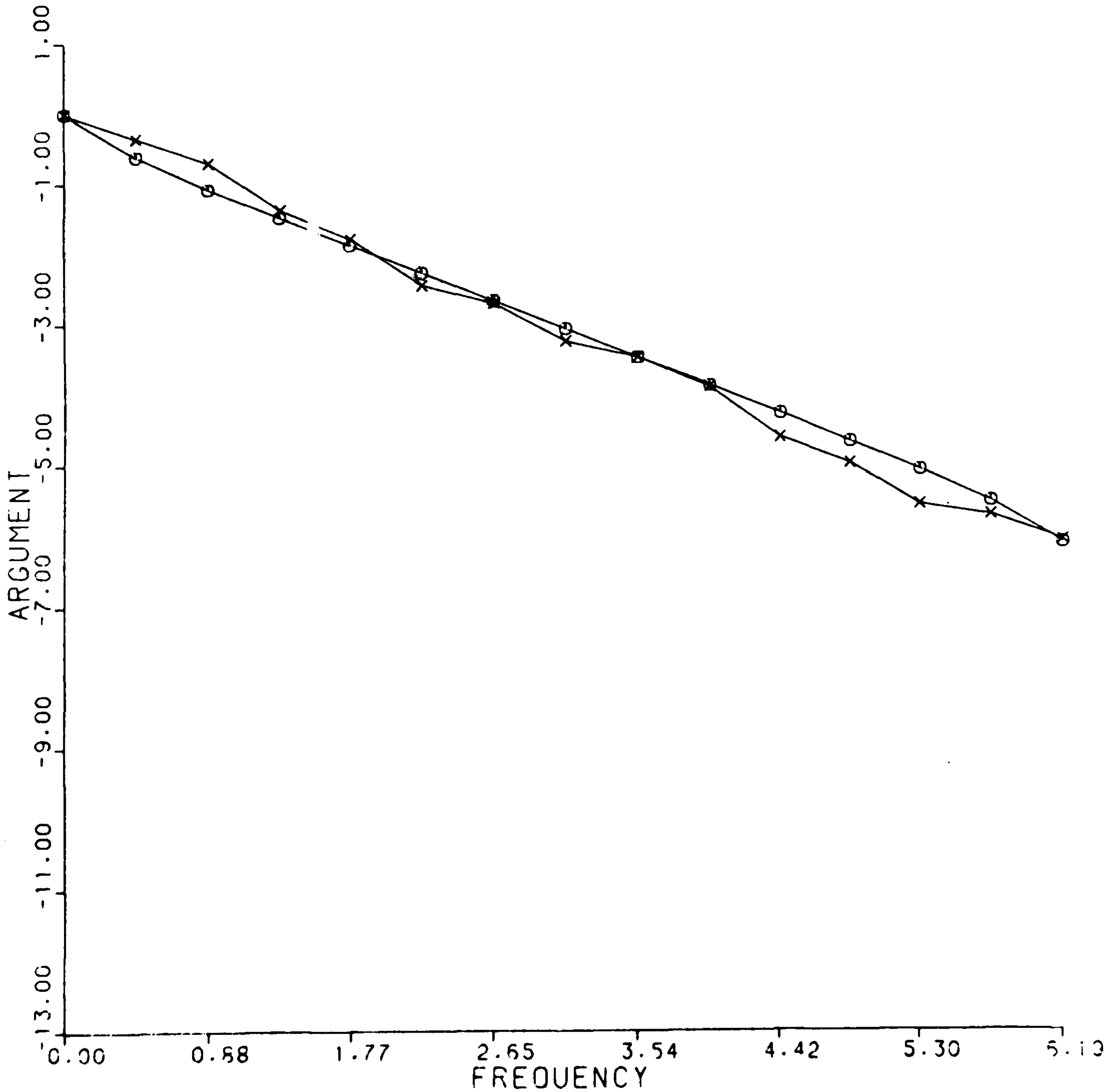
$$\sqrt{N} (\hat{g}_a - g_a)$$

are asymptotically normally distributed with covariance structure given by the preceding formula.

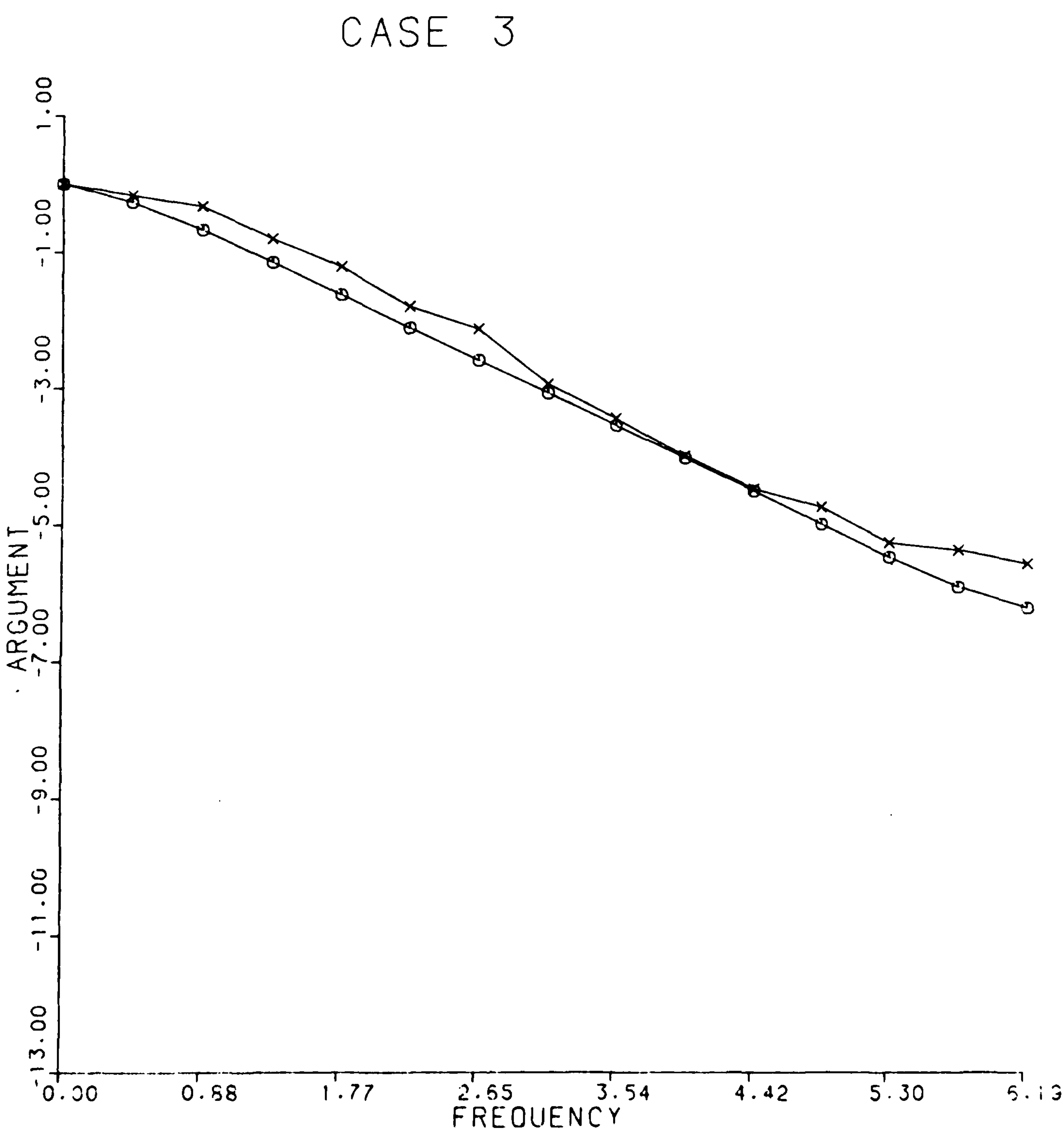

CASE 1
ARGUMENT
1.00
-1.00
-3.00
-5.00
-7.00
-9.00
-11.00
-13.00
0.00
0.86
1.77
2.65
3.54
4.42
5.30
6.19
FREQUENCY

Graph 1

CASE 2

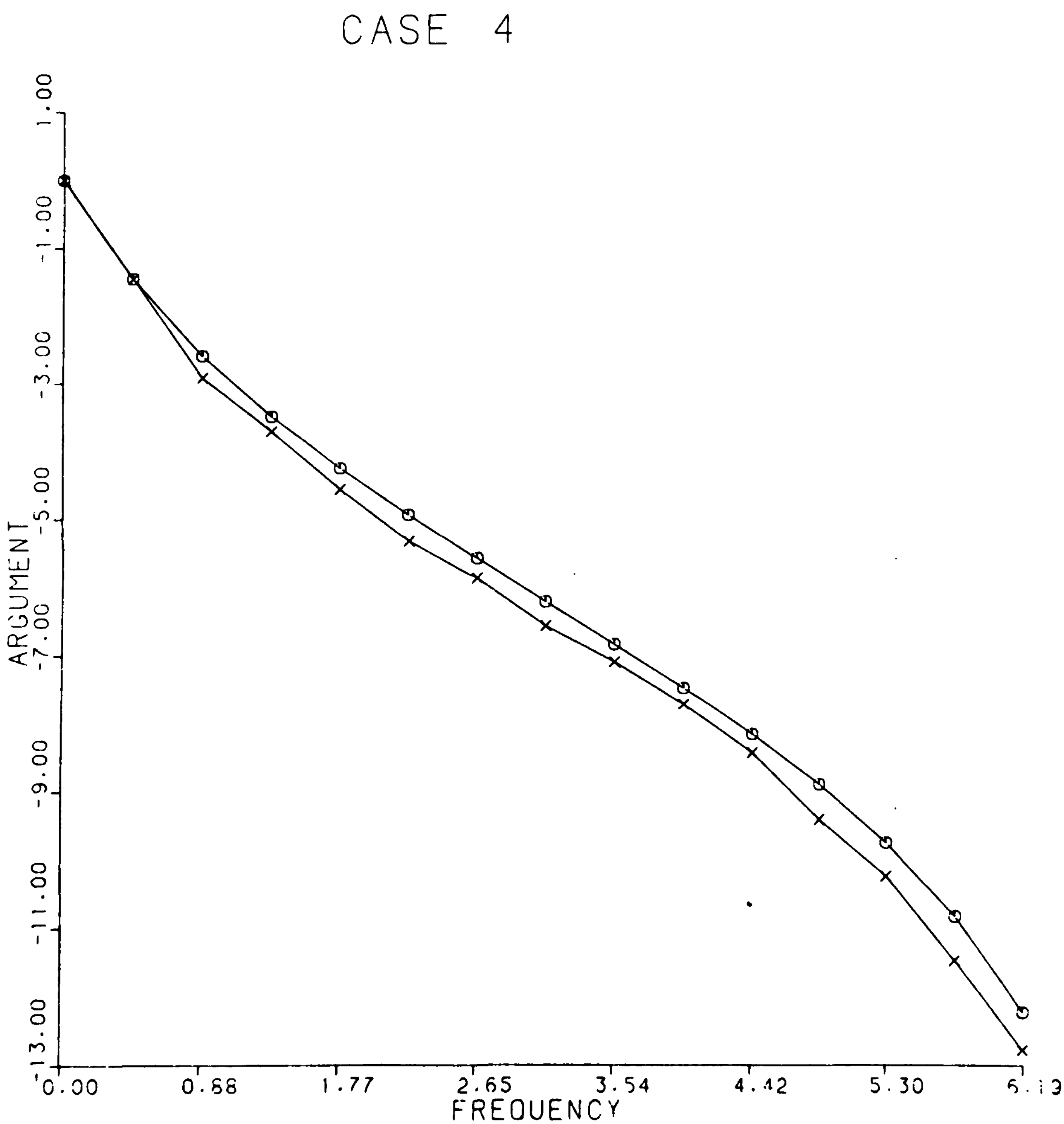


Graph 2

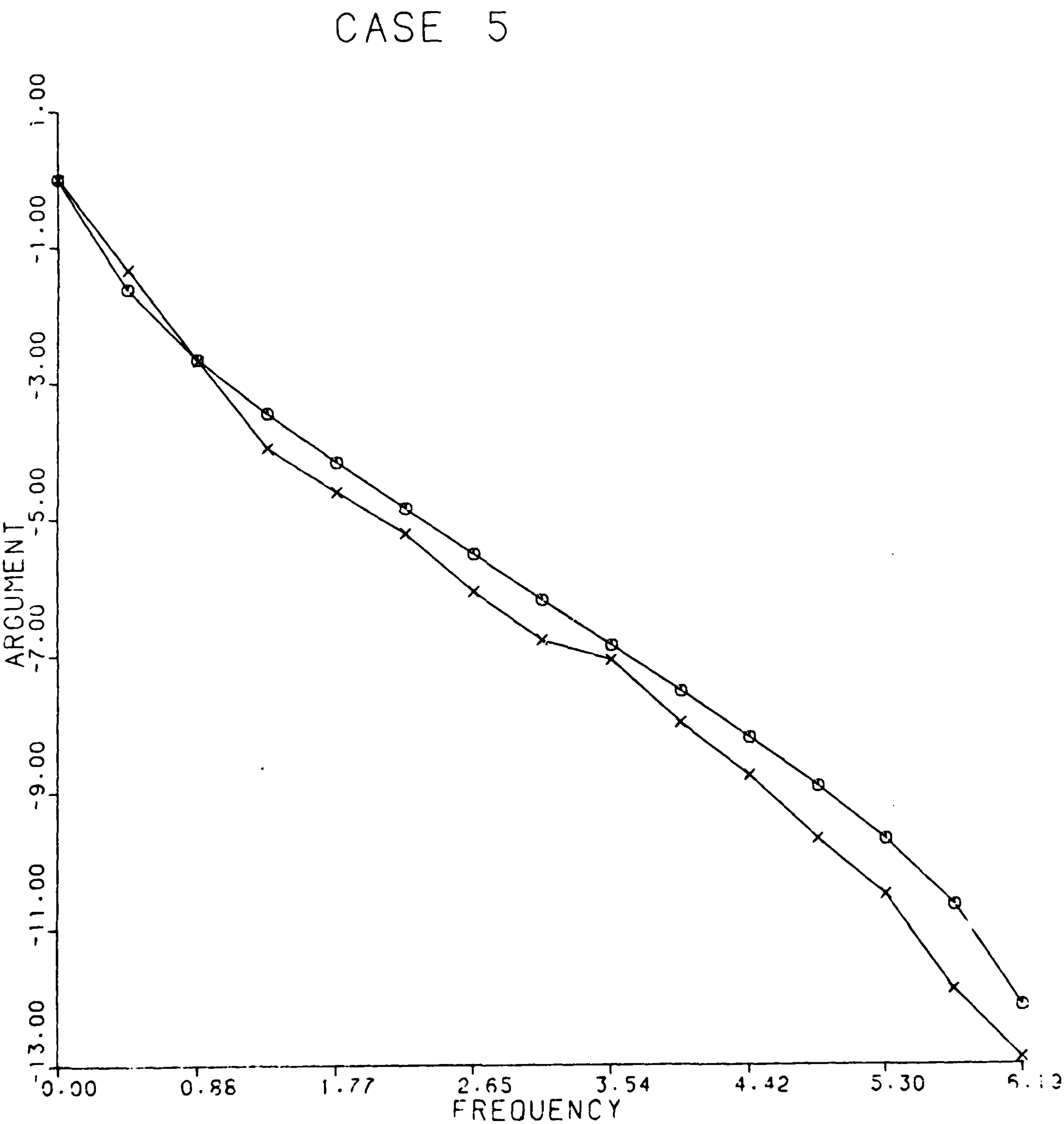


CASE 3
ARGUMENT
1.00
-1.00
-3.00
-5.00
-7.00
-9.00
-11.00
-13.00
FREQUENCY
0.00
0.88
1.77
2.65
3.54
4.42
5.30
6.13

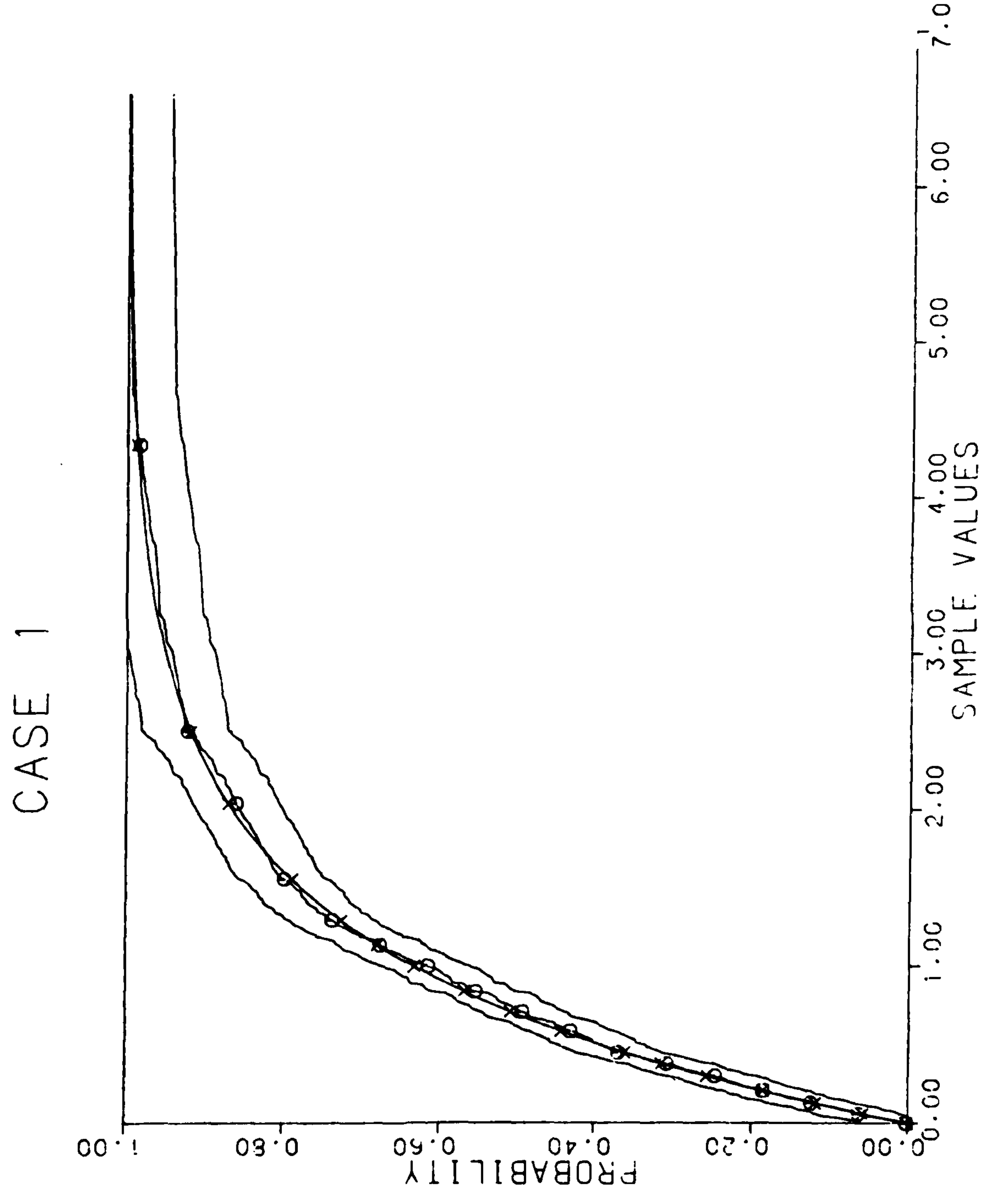
Graph 3


CASE 4
ARGUMENT
1.00
-1.00
-3.00
-5.00
-7.00
-9.00
-11.00
-13.00
0.00
0.88
1.77
2.65
3.54
4.42
5.30
6.19
FREQUENCY

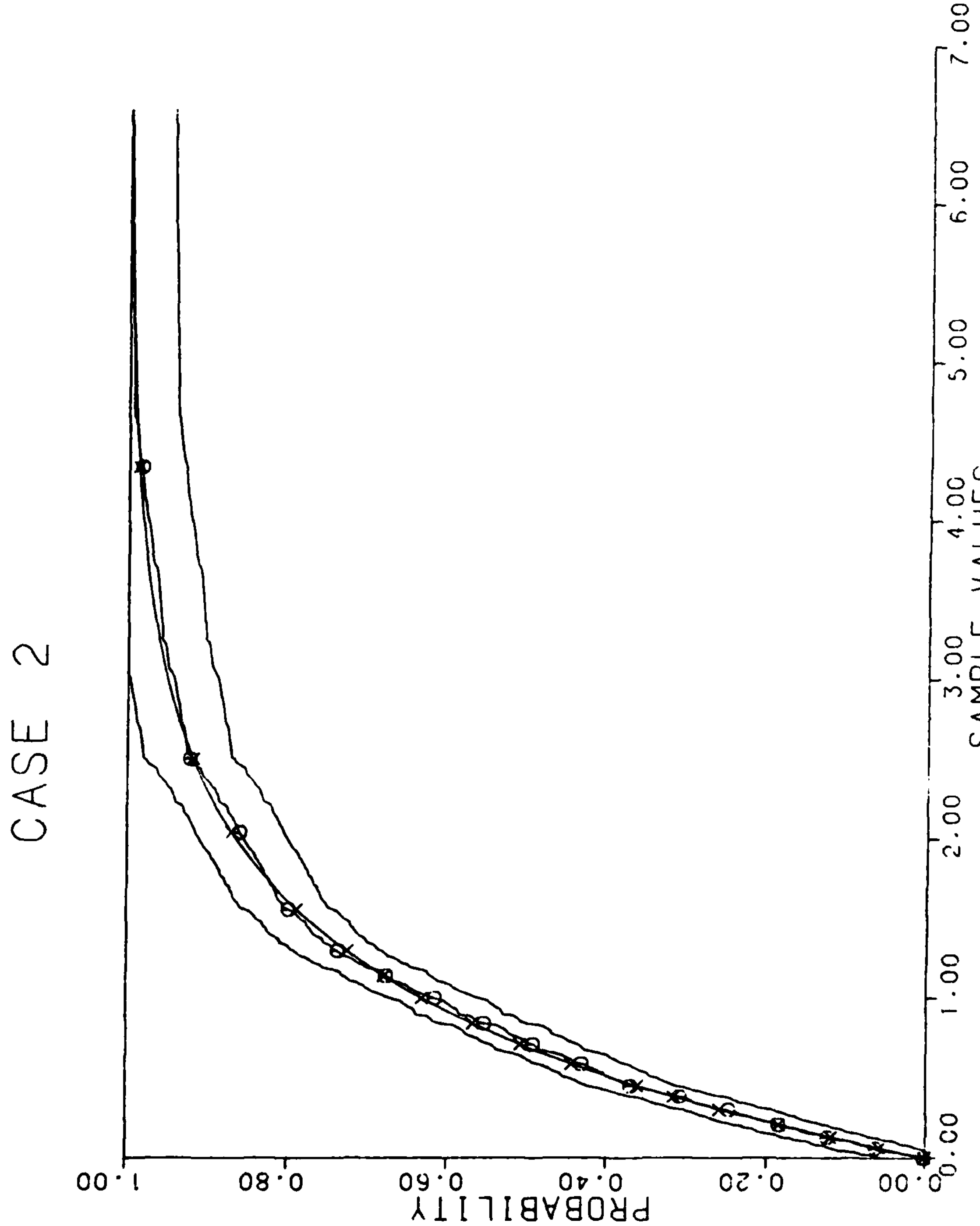
Graph 4


CASE 5
ARGUMENT
1.00
-1.00
-3.00
-5.00
-7.00
-9.00
-11.00
-13.00
0.00
0.88
1.77
2.65
3.54
4.42
5.30
FREQUENCY

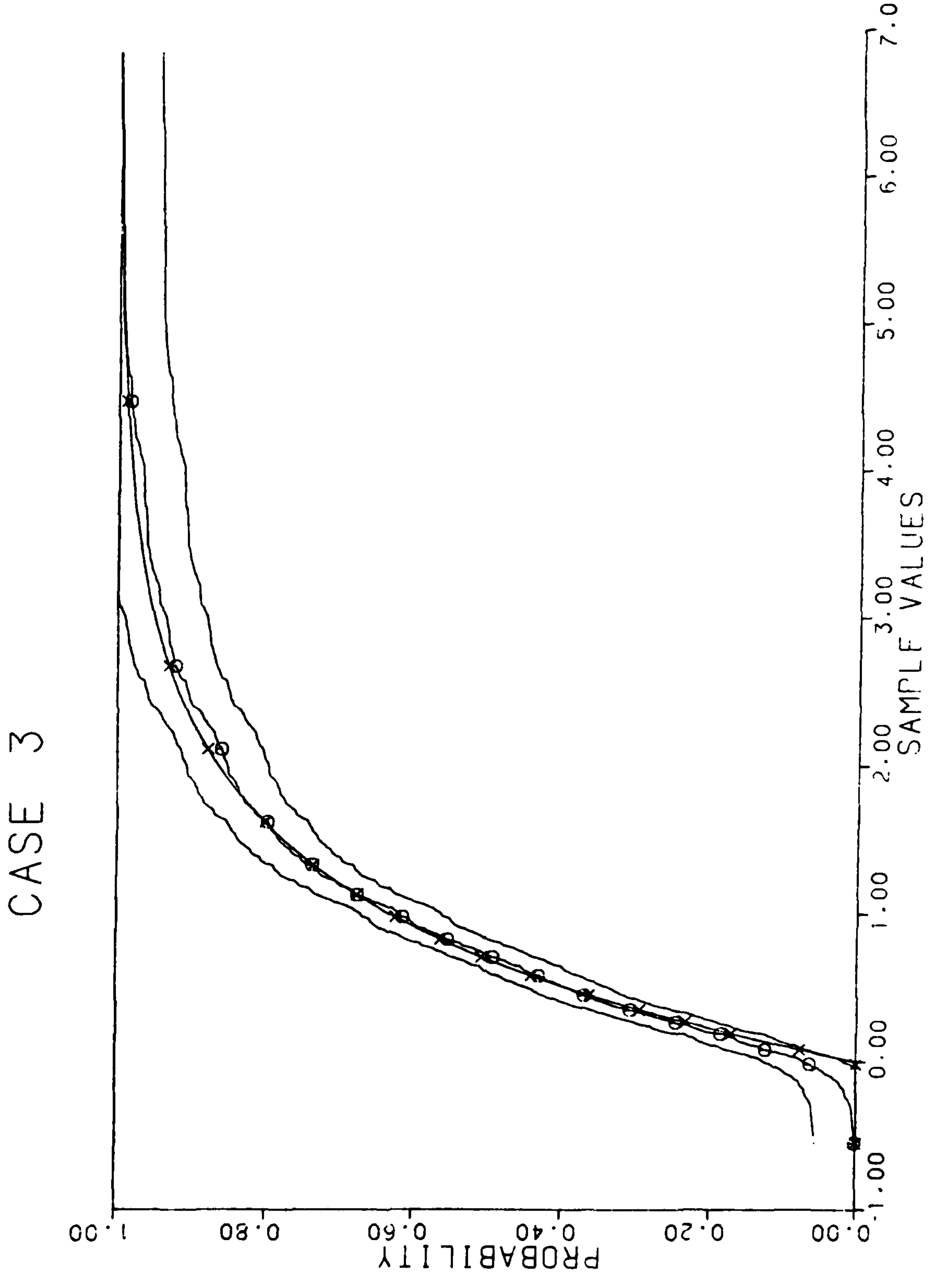
Graph 5


CASE 1
PROBABILITY
0.00
0.20
0.40
0.60
0.80
1.00
SAMPLE VALUES
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00

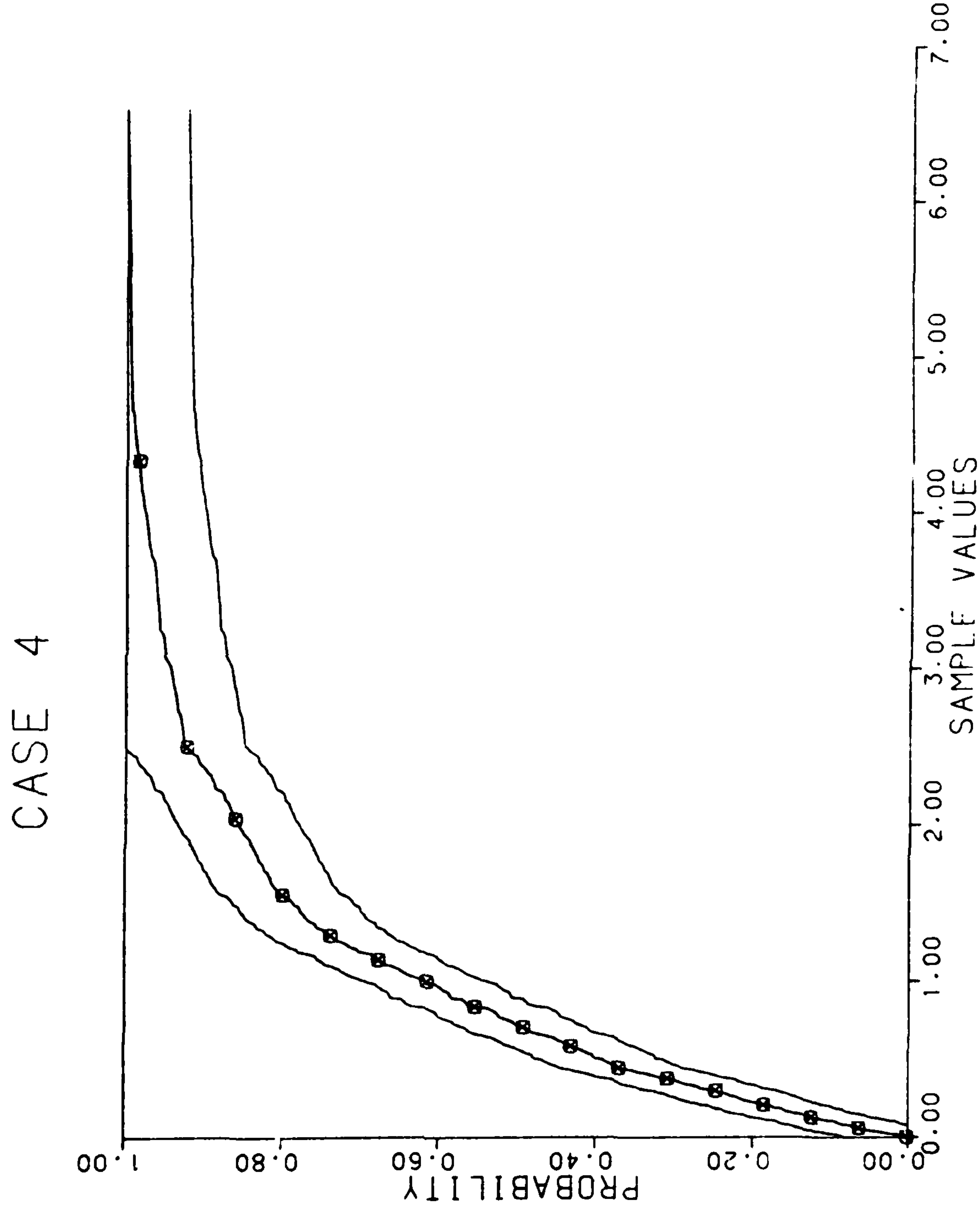
Graph 6


CASE 2
PROBABILITY
SAMPLE VALUES
0.00
0.20
0.40
0.60
0.80
1.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00

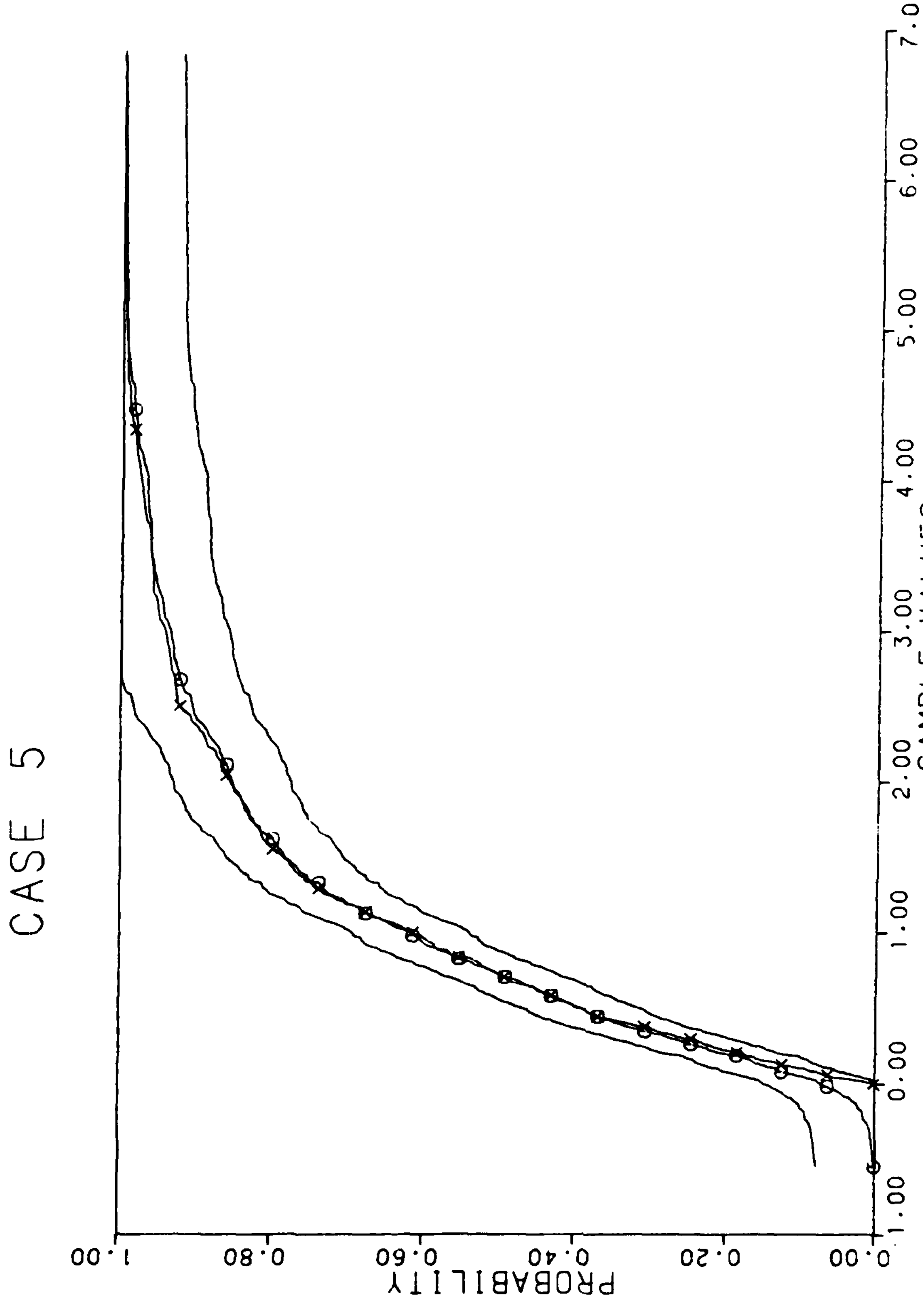
Graph 7


CASE 3
SAMPLE VALUES
PROBABILITY
-1.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
0.00
0.20
0.40
0.60
0.80
1.00

Graph 8


CASE 4
SAMPLE VALUES
PROBABILITY
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
0.20
0.40
0.60
0.80

Graph 9


CASE 5
SAMPLE VALUES
PROBABILITY
-1.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
0.00
0.20
0.40
0.60
0.80
1.00

Graph 10

ILMED

8