THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE ANALYSIS OF BIOLOGICAL DATA COLLECTED SYSTEMATICALLY AND PSEUDO-SYSTEMATICALLY OVER TIME

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

DONALD EARL PARKER

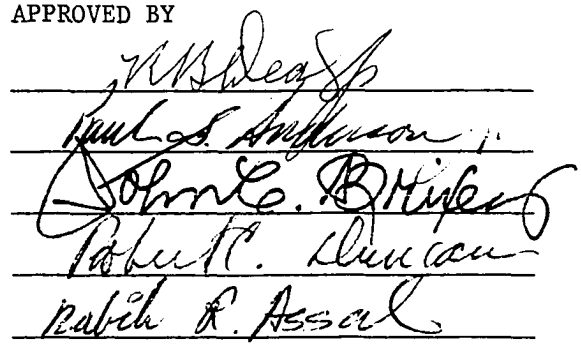
Oklahoma City, Oklahoma

1970

THE ANALYSIS OF BIOLOGICAL DATA COLLECTED SYSTEMATICALLY

AND PSEUDO-SYSTEMATICALLY OVER TIME

APPROVED BY



DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

I am especially indebted to Dr. Robert C. Duncan for the assistance given to me while serving as the Major Advisor for this research and for the knowledgeable insight of many statistical concepts that he provided even before this research was begun.

I am deeply grateful to Dr. R. B. Deal, Jr. for his invaluable assistance in the theoretical portion of this research. Appreciation is also expressed to Drs. Paul S. Anderson, Jr., John C. Brixey, and Nabih R. Assal for serving on the reading committee.

Appreciation is extended to Mrs. Rose Titsworth whose knowledge of form and format has made the preparation of this dissertation possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES—Continued

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS--Continued

THE ANALYSIS OF BIOLOGICAL DATA COLLECTED SYSTEMATICALLY AND PSEUDO-SYSTEMATICALLY OVER TIME

CHAPTER I

INTRODUCTION AND STATEMENT OF PROBLEM

Introduction

Within the past decade the variation over time of many biological phenomena has received a large measure of attention. In 1962 the New York Academy of Science sponsored a conference concerned with the rhythmic phenomena in living systems and, in 1966, another conference with human variation as the theme. The resulting publication from the former with Wolf (1962) as conference editor serves not only to indicate the application in varied biological fields, but also provides a valuable source of bibliographic information. The publication resulting from the latter, with Brozek (1966) as conference chairman, contains several papers concerned with variation in the genetic sense, but also contains papers directly concerned with variation over time. In particular, Vandenberg's (1966) contribution gives a discussion of statistical techniques and computer applications together with a rather extensive list of references.

Although some investigators such as Potthoff and Roy (1964) and Elston and Grizzle (1962) have offered new multivariate techniques

for handling time response curves of the growth curve type, a large part of the theoretical work on time response seems to concentrate on dealing with the analysis of rhythmic data. In discussing these analyses Sollberger (1962) writes,

> The various types of analysis may be obtained by tedious mathematical computation or mechanical, optical, and electronic computers. They include several modifications of straightforward harmonic analysis, auto-correlation (lag correlation, correlating a curve with itself, moved one or several steps out of phase, yielding correlograms ...), and cross correlation. The results may be represented by an amplitude spectrum (periodograms, frequencies on the x axis, corresponding amplitudes on the y axis), a power spectrum (phase plotted against frequencies). Many of these analyses require long stretches of cycle recordings which may be difficult to obtain. ... The application of Fourier analysis to biological rhythm has developed rapidly in the last years, ...

Although advances have been made in techniques since the time of Sollberger's statement, it still serves as a succinct summary of the more sophisticated analyses for rhythmic data. However, many biological investigations are carried out which, although concerned with measurements over time, are not seeking the information yielded by the above analyses nor, in many instances, does the conduct of the experiments yield data which is suitable. The investigations of diurnal variation in blood coagulation times of animals reported by Scheving and Pauly (1967), Nagorra-Stasiak (1963), and Everson (1960) are examples of such studies.

Statistical tools are available to aid in the extraction of information from the time series resulting from these studies, but often the appropriate tools are not those analyses mentioned by Sollberger. It is to some of these less elaborate (though often more appropriate) statistical methods and their application to time series that this dissertation addresses itself. Although the term "time series" has

become so closely associated with analyses involving Fourier series, power spectra, etc., that its usage seems to connote data to be subjected to one of these techniques, this is not the intention here. A time series will mean simply measurements taken serially over time.

For the purpose of discussion the analyses of time series considered here will be partitioned into two classifications - Time Classification Analyses (TCA) and Time Average Analyses (TAA). The appropriateness of a TCA or TAA approach in the extraction of information for a particular study will depend upon the nature of the variable measured and the sampling procedure and the inferences to be made. The definitions, examples, and discussions below will provide the rationale for this partitioning.

Time Classification Analyses

If the analysis regards the time points as levels of a classification or factor, then the analysis will be called a Time Classification Analysis which, for brevity, will be denoted by TCA. Essentially this is a requirement that the subject matter point of view and the conduct of the sampling allow a meaningful way to define an equivalence of time points across a second classification such as individual subjects, groups, treatments, etc.

Many of the investigations concerned with circadian periodicity and seasonal rhythm would meet this criterion. The investigation by Scheving and Pauly (1967) provides a clear example of a set of data for which a TCA would be appropriate.

In many investigations yielding data consistent with a TCA

there are typically several environmental conditions $\{a_1, a_2, \ldots, a_p\}$ measured at times $\{b_1, b_2, \ldots, b_q\}$. The time levels may be actual clock times that are the same for each condition or, the more usual situation, time relative to some specific event. In the article cited above the measurements within each environmental condition were begun at 6 A.M. and taken every two hours until 4 A.M. the following day. Each group was measured on different days, but the time points were regarded as equivalent since they were measured as time elapsed from a specific event, (12 Midnight).

The important point concerning the "levels" of time is that for the purpose of interpretation they are equivalent in some meaningful sense across groups.

The purpose of such experiments is to gain information on the locus of the mean of a population P_{a_i} of subjects over a time period. The population of inference is all subjects (of which a sample has been measured at time points $b_1, b_2, \ldots, b_q$) under environmental condition a_i. Not only is the experimenter uncertain of the exact value of the population mean at a time t_i because he is estimating this from a sample from P_{a_i}, but also there is usually assumed a random variable associated with each measurement. Said another way: if several elements of the sample are measured at time b_i they will not all give the same value for two reasons.

(1) An element x_k has its own individual deviation π_k from the population mean.

(2) There is a random component ε_{k_i}.

Conceptually, the ε is a random variable in the sense that if x_k

were measured again at the same (equivalent) time, the ε would not necessarily be the same. The π_k, however, is a characteristic of the element x_k and would be the same.

Clearly, the problem of extracting information from these experiments is partially a statistical one. However, examination of recent publications concerning studies of this nature reveals that only the simpler statistical techniques are usually applied to the data. As an example consider the following statement from the article by Scheving and Pauly (1967) cited above: "Statistical analyses were made by calculating the standard error of the difference between the means to be compared and the t value." The means compared were, for each environmental condition:

1. High vs. low
2. Overall mean vs. high
3. Overall mean vs. low.

Certainly these are comparisons of interest and are appropriate; however, application of slightly more sophisticated analyses would allow other informative comparisons to be made. Also, since in the original analysis, the pairs to be tested were determined <u>after</u> examining the data, probability levels obtained from Students' t tables could be seriously misleading.

The data were presented as means for each time point together with a standard deviation for each mean. Graphs of the means were presented and inferences made from examining the graphs. The purpose here is not to criticize the discussions and observations presented, but to point out that the application of certain statistical techniques will

allow correct probability levels to be associated with desired comparisons. The authors were kind enough to grant permission for the use of their data to illustrate an analysis by TCA. This expanded analysis is presented in CHAPTER II.

Time Average Analyses

If an analysis summarizes the measurements without specific regard to the time coordinates and inferences are drawn only from the averaged or pooled measurements over time, then the analysis will be called a Time Average Analysis which, for brevity, will be denoted by TAA. The investigation of fibrinogen patterns by Hampton (1966) offers an example of data analysed by TAA. In the analysis presented, not only are estimates of means obtained by averaging over time (the more common TAA found in applied biological literature), but also estimates of within subject variances are obtained by pooling over time.

A TAA is appropriate whenever inferences are to be drawn - (by choice or necessitated by circumstance) only from estimates of the characteristics of the distribution of values attained by a variable over the sampled time interval. The property of when in the interval values are attained enters the problem only in connection with how to sample, not as a factor which changes the population values being estimated. If certain mathematical niceties are ignored, one may say, roughly, that these population values and estimates depend upon the relative frequency with which values are attained, but not upon the pattern exhibited over time in attaining these values. Note, however, that the interaction of the sampling plan and pattern may have important effects on estimates.

Although many investigations conducted in laboratories would use a TCA, many others would use TAA; however, information gathered from large epidemiological studies such as the Framingham, Tecumseh, and Roseto studies would almost exclusively be summarized in a manner that would classify the analysis as TAA. In large sampling situations the difficulty in controlling the activities of subjects, especially human subjects, often prohibits the measurements at precise times which are necessary for a meaningful TCA. Perhaps the best that can be done in survey situations is to control partially for time pattern changes by partitioning the data relative to when the measurements were taken (time of day, season of year, etc.).

In addition to the difficulties in conducting experiments, the variable itself may be such as to almost preclude a TCA approach. The pattern exhibited may be similar in two different subjects, but the time coordinate must be measured relative to each subject if the similarity is to be perceived. The pattern in hormone levels related to the menstrual cycle in women is an example. Of course, when the cycle is known, as in this case, it would be possible to define meaningfully equivalent time points by measuring time relative to onset of menses for each individual, but when such specific knowledge is not available, there may be no way to define meaningfully equivalent time points.

Whether for convenience, nature of the variable under study, availability of data, tradition, or other reasons, many studies are conducted where conclusions are reached by examining only the distribution of values obtained. In the following chapters some of the difficulties and merits of this procedure will be examined.

The greater portion of this dissertation is devoted to TAA topics rather than TCA, not because TAA is felt to be more important, but only because discussions applicable to TCA are found in most standard statistical references. The portion devoted to TCA is intended only to illustrate how standard statistical techniques may be applied to biological time series by regarding time points as levels of a factor. With this view of time in mind, many of the statistical techniques discussed in context other than time series are seen to be appropriate and, therefore, a less detailed treatment of TCA seems sufficient.

CHAPTER II

ANALYSIS BY TCA

When the aim of an investigation is the acquisition of information on the response over time of a variable G under a set of environmental conditions $A = \{a_1, a_2, \ldots, a_p\}$, and measurements are made in each condition at a set of time points $B = \{b_1, b_2, \ldots, b_q\}$, a well known statistical procedure, an analysis of variance (AOV) will be the basic statistical tool appropriate for the analysis in order to interpret the data. Although certain conditions must obtain for an AOV to be theoretically valid, rigorous insistence that the assumptions are exactly met is seldom made in situations where AOV's are used extensively, and a critical attitude toward these assumptions simply because time is regarded as a factor would seem unnecessarily restrictive. In this chapter only analyses which require these assumptions are specifically considered. The same approach may be used, however, with most of the observations and recommendations remaining germane, even if the nonparametric analogs are used rather than the normal theory tactics discussed here. Winer (1962) discusses the nonparametric analogs of AOV's. Siegel (1956) gives many of the nonparametric alternatives together with comparisons of relative efficiency.

For statistical considerations studies in which measurements are taken at the same (equivalent) time points under several conditions

may be considered as experiments conducted with a factorial arrangement of treatments. The time points are regarded as levels of factor B and the conditions as levels of factor A. An appropriate analysis conducted in the factorial context will "answer" the following questions:

1. Are the means (averaged over all time periods) different across conditions? (A main effect).
2. Are the means (averaged over all conditions) different across time periods? (B main effect).
3. Is the pattern of response over time the same in each condition (except possibly for a translation), i.e. are the time response curves within each condition parallel?

In addition to the above three questions (which are the ordinary orthogonal partitions made in a factorial analysis), tests of the simple effects of B within each level of A would "answer" the question (for each condition) "Does the response variable vary over time?". These tests would be concerned with the same information that is being tested by the "t tests" of the investigation of Scheving and Pauly (1967) mentioned in CHAPTER I. The selection of comparisons, however, would not be based upon examination of the data, and a proper probability level would, therefore, be obtained. Note also that an AOV would use information under all times and conditions for the estimation of the error term. Thus a better estimate of error is available than with the t test approach.

Also, subsets of conditions or subsets of times can be compared for parallelism of response curves over time. Of course, care must be taken in selection if independence of tests is to be preserved.

Even if an investigation is confined to a univariate approach

for a time series with time regarded in the factorial context, the possible designs are many. A greater degree of complexity occurs with the introduction of more and more classification factors, replications in the factorial sense, attempts to design so as to estimate residual effects, etc. The set of all such designs, however, may be partitioned into two important subsets - those designs in which each experimental unit is measured under only one treatment combination, and those in which some or all experimental units are measured under more than one treatment combination. The assumed structural model for the analysis and a brief discussion of the two-way cross classification design in each of these subsets is given below for the purpose of pointing up the relevant difference between the two subsets and to offer justification for the conceptual view of time points as levels of a factor.

Both of these analyses and the multivariate approach to the same experimental situation discussed below require the usual assumption of normality, equality of variance, and identical covariance matrices across conditions.

If the structural model assumed is

$$x_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk} \quad ,$$

where x_{ijk} is the observation on the k^{th} subject under condition a_i at time b_j; μ is the overall mean; α_i is the effect of condition a_i; β_j is the effect of time b_j; $\alpha\beta_{ij}$ is the interaction; and e_{ijk} is a random variable distributed $N(0, \sigma_e^2)$ and $E(e_{ijk}, e_{i'j'k'}) = 0$ for ijk not identical to i'j'k', then the analysis will be called a Type I.

This analysis would apply to studies where no repeated measurements are made on the same experimental unit, i.e., a different group of

subjects is measured at each time point and condition. This circumstance would obtain necessarily in many biological studies where the measurement itself requires the sacrifice of an experimental animal or, when the animal is not sacrificed, the measuring process is such that it may in some way influence a measurement at a subsequent time. The latter condition necessitated the use of different rats at each time point in the coagulation study used below as an example of TCA.

Theory and computational procedures for Type I are discussed in most texts on experimental design. Steel and Torrie (1960) and Winer (1962) give extensive and readable coverage of this topic.

If the structural model assumed is

$$x_{ijk} = \mu + \alpha_i + \pi_{k(i)} + \beta_j + \alpha\beta_{ij} + e_{k(ij)} \quad ,$$

where the terms have the same meaning as in Type I, but with the addition of $\pi_{k(i)}$, the effect of subject k, and parentheses in subscripts indicating that the effect is nested under the enclosed indices, then the analysis will be called Type II.

Type II would be appropriate when measurements are made on the same subjects at each time point within a condition, but different subjects for each condition. Indeed, the term "nested" and the use of parentheses in the indexing in the definition above are, respectively, the statistical parlance and often used symbolic representation to state formally this condition.

The repeated measures upon a subject within each condition together with the assumptions of the model lead to a within condition covariance matrix with σ_e^2 along the main diagonal and a constant $r\sigma^2$ elsewhere, i.e., the correlation of measurements at b_i and b_j is a

constant r for all $i \neq j$. This may be regarded as correlation existing only because measurements are taken on the same subjects, but not because of proximity in time. When correlation does exist because of proximity in time, Type II is not strictly proper. Under these circumstances the multivariate view of the time series is appropriate.

Note that the covariance matrix of Type I consists of σ_e^2 along the main diagonal also, but with zeros elsewhere.

Winer (1962) deals at great length with Type II and other more complex designs with repeated measures, as well as with comparisons of Type I and Type II models.

Cole and Grizzle (1966) offer a multivariate alternative to the Type II repeated measure design above. This alternative will be referenced as Type III. In this analysis a slightly different concept of the series of measurements is employed. As in Type II, measurements are taken on the same subject at each time point, and different subjects are used for each condition. However, the series of measurements at time points $b_1, b_2, \ldots, b_q$ for each subject is regarded as an element drawn from a q dimensional multivariate normal population. Whereas the time points are considered levels of a factor in Type II, they act in Type III as an index to designate $1^{st}, 2^{nd}, \ldots, q^{th}$ component of a vector.

As mentioned above, if correlation between pairs of measurements within a subject is a function of their proximity in time, then the assumptions necessary for Type II are not met. This condition does not violate the assumptions of Type III. In the article cited, the authors offer analyses of time series data by both Type II and Type III, contrast the two, and suggest conditions under which one is to be preferred

over the other. All the comparisons yielding main effects, interactions, simple effects, etc., possible in the Type II are possible with Type III. In fact, the method of forming rejection regions (the union intersection method of Roy) allows comparison to be made without the distinction necessary in Type II between pre-planned comparisons and those conceived after examining the data.

Although Type III circumvents certain problems of Type II, the multivariate approach creates difficulties of its own. The complexities of computing procedures is not the least of these. Also, the question of power of these tests does not seem resolved. For a more thorough discussion of the merits and difficulties of Type III, the article cited may be consulted.

Although Types I and II are found in many intermediate level text books in the section covering factorial and factorial with repeated measures on one factor, there is a shift of interest when these methods are applied to time series which is not reflected in the usual text book treatments and examples. Since the interest is in the pattern of response over factor B, and the effect upon this pattern by the levels of A, the AB interaction is the major concern rather than main effects. The problem of extracting information about the response curves across levels of a factor is treated most often in the statistical literature in one of two ways:

(1) Estimation of the curve and a corresponding confidence band.

(2) Partitioning the interactions into comparisons such as linear by linear, quadratic by linear, quadratic by quadratic, etc. Most intermediate level texts discuss the procedures and

interpretation of this trend analysis approach.

While both of these procedures are appropriate for many time series, there are certain problems associated with each for particular applications. In the usual regression fitting of a polynomial to cyclic data such as circadian rhythms, the "closeness of fit" at particular time points may be influenced by the selection of the time to be considered as the first in the series. The fitting of a polynomial model to a set of data with, say 12 noon measurements, considered as the first of the series may yield different values for a given "closeness of fit" criterion than the value obtained by fitting the same polynomial model to the same data with, say 12 midnight measurements, considered as the first of the series. Since the designation of starting time is arbitrary if the purpose is to represent the curve through exactly one cycle, an investigator may be able to find an adequate fit (by whatever standard he wishes to judge "adequate") with the use of a smaller degree polynomial simply by choosing a different time point as the beginning of the cycle. To illustrate the dependency upon starting point, consider a measurement series wherein the response curve is exactly $\sin 2\pi t$. If the first measurement is at $t = .25$ and the last at $t = 1.25$, then a plot of the data would appear parabolic, and a second degree polynomial would roughly approximate the curve. However, if $t = 0$ were the initial point and $t = 1$ the last, a third degree polynomial would be required for even a crude approximation.

Clearly, the ability to describe the data by a lesser degree polynomial would be advantageous where the calculations are to be done on a hand calculator, but it is perhaps not so obvious when a computer

program is to be used; however, the problem of retention of significant digits does present itself in some programs for regression analyses.

Descriptions couched in terms of trend analysis may be difficult for a researcher to interpret meaningfully. Information such as "the general upward (downward) trend is the same under all conditions" - a rough statement of the failure of linear by anything to be significant - may not be helpful. In a specific study there are often other comparisons that would be more useful for extracting information from the data. A close cooperation between a researcher and a statistical consultant may uncover sets of comparisons which may be presented together with proper probability levels that will have a clearer interpretation from a subject matter, point of view and, consequently, be more helpful to the researcher and the readers of his publication.

When the sets of comparisons are not defined a priori but are decided upon only after examining the data, the methods suggested by Scheffe´ (1959) may provide a method of reporting a proper probability level. As mentioned above, Dr. L. E. Scheving and Dr. J. E. Pauly have given their permission to use the data from their coagulation study (1967) to illustrate the applicability of certain statistical, graphical, and data analysis techniques to measurements taken over time. The results of the application of these techniques lead to conclusions that differ in some respects from those of the authors. These differences are in no way intended as criticism, but rather examples of how more thorough analysis may give insight into alternate interpretations and suggest hypotheses for further study.

The following is from the abstract appearing with the article:

> With all other environmental factors rigidly standardized, normal Sprague-Dawley rats were maintained under the following schedules: (1) 12 hours of artificial light 0600 to 1800 alternating with 12 hours of darkness - LD; (2) reversal of the first - DL; (3) constant darkness - DD; and (4) constant illumination - LL.
>
> After the animals had been under a specific lighting regimen for at least three weeks, blood coagulation times were determined on separate groups of 8 to 16 animals at bi-hourly intervals during a 24 hour period.

The following is from the text of the article:

> In the first phase of the study two different colonies were utilized. One colony had been in DD for nearly four months; the second colony had been maintained under a LD regimen for almost three months. Other than the differences in lighting regimens the rats were comparable ... .
>
> The second phase involved sampling the original colony of rats that had been utilized for the DD studies but which, by this time, had been maintained in LL for about one month. However, prior to being placed in LL they were first subjected to LD cycles for seven days ... .
>
> The LL animals subsequently were placed in LD for readjustment and then were subjected to DL.

For a more detailed explanation of the conduct of the experiment the article should be consulted.

For the purpose of analysis the two colonies are regarded as simple random samples from the population of rats to which inferences are to be made (normal Sprague-Dawley rats weighing about 350 grams, etc.), and the selection of rats to be measured at a time point is regarded as a random assignment of time points to the members of the colony.

In the conduct of the study equal numbers of rats were not used at each time period in either the first or second phase. In order to illustrate analyses with equal and unequal cell frequencies the data from the first phase was adjusted by randomly discarding measurements from those time points for which more than 12 rats were measured. Thus,

the cell means presented here for LD and DD rats will differ slightly from those in the article. Table 1 below gives the summary statistics for the data used in this analysis. Table 2 is the basic AOV and simple effects resulting from a Type I analysis with equal cell frequencies. Prior to the computation of the AOV, the homogeneity of variance assumption was tested by the method of Hartley. No evidence was found for rejection of the hypothesis of equality of variance with $\alpha = .01$.

Although this basic AOV could be made more elaborate by further partitioning, the analysis as presented in Table 2 would provide statistical support for several hypotheses concerned with the daily temporal relation of the coagulation time (CT) of laboratory rats. Some possible wordings of these hypotheses are the following:

i. The 24 hr. CT mean is different in LD rats from the 24 hr. mean of DD rats (significance of lighting condition main effect).

ii. Within the LD and within the DD conditions the locus of the CT mean over time is not a constant (significance of within LD and DD simple effects).

iii. The pattern over time of mean CT response of LD rats is not parallel to the pattern of DD rats (Significance of Condition X Time interaction).

iv. The mean CT of LD rats is different from the mean CT of DD rats for each of the time points measured from 0730 through 1730 (significance of the first 6 within time points simple effects).

v. No difference in the mean CT of LD rats and DD rats is

TABLE 1

SUMMARY STATISTICS* FOR COAGULATION TIMES OF RATS UNDER DD AND LD LIGHTING CONDITIONS

Time Period	DD Mean* (Sec.)	DD Standard Error	LD Mean* (Sec.)	LD Standard Error
0730	255.3	12.0	217.8	13.6
0930	136.3	14.7	246.1	13.8
1130	131.8	12.9	276.8	15.1
1330	149.4	19.6	219.9	13.1
1530	163.9	11.0	240.6	21.3
1730	181.2	19.7	265.4	10.9
1930	283.1	12.0	259.3	15.0
2130	288.1	22.3	299.3	18.9
0230	344.8	25.0	302.4	18.8
0130	321.1	17.9	314.3	23.3
0330	306.6	17.2	296.5	23.7
0570	254.5	19.2	257.4	16.9

*Each mean is based on 12 rats for a total of 288 rats.

TABLE 2

ANALYSIS OF VARIANCE FOR COAGULATION TIMES OF RATS UNDER DD AND LD LIGHTING CONDITIONS

Source of Variation	SS	DF	MS	F
Lighting Condition	72168.340	1	72168.34	19.64**
Time Period	672797.733	11	61163.43	16.64**
Condition x Period	258295.524	11	23481.41	6.39**
Within Cell	970248.418	264	3675.18	

Time Periods Within Lighting Condition Simple Effects

Condition	SS	DF	MS	F
LD	137028	11	12457.1	3.38**
DD	794065	11	72187.7	19.64**

Lighting Condition Within Time Periods Simple Effects

Time Period	SS	DF	MS	F
0730	8437.50	1	8437.50	2.29
0930	7238.01	1	7328.01	19.69**
1130	1261.50	1	1261.50	34.32**
1330	2982.15	1	2982.15	8.11**
1530	3526.66	1	3526.66	9.59**
1730	4258.83	1	4258.83	11.58**
1930	3384.37	1	3384.37	.92
2130	7481.70	1	7481.70	.20
2330	1075.26	1	1075.26	2.92
0130	2733.75	1	2733.75	.07
0330	6100.44	1	6100.44	.16
0530	5104.10	1	5104.10	.01

**Indicates significance at .01; * at .05.

detectable at $\alpha = .05$ for each of the time periods measured from 1930 through 0530 (non-significance of last 6 within time points simple effects).

Further partitioning of the within LD simple effect sum of squares could be made so as to lead to more specific comparisons concerning the nature of the pattern of response over time within this lighting condition. In addition to the trend analysis mentioned above, the nature of the subject matter suggests other comparisons such as the mean CT in light versus the mean under darkness, the first x time periods after switching on lights versus the first x time periods after switching off lights, the middle time periods of light versus middle dark periods, etc. Similarly, other meaningful comparisons may be made by further partitions of the DD simple effect or interaction sum of squares.

These comparisons, in the context used here, are preplanned comparisons, i.e., comparisons decided upon without examining the data. Of course all possible comparisons should not be made. The specific questions which a researcher would like the data to answer would govern the selection of the set of comparisons. Ideally, the set selected would consist only of orthogonal comparisons so as to preserve the independence of <u>tests</u>, but departure from complete independence in order to obtain meaningful comparisons should not be condemned. The lack of independence should, of course, be noted in the discussion of results.

Comparisons which are suggested by the data can, of course, be made. The procedure of Scheffe´ (1959) assigns a proper α level to comparisons. He refers to such procedures as "data snooping". The disadvantage of making comparisons after the fact lies in the lessened

ability to detect small differences, i.e., differences in data means must be greater to be declared significant at a given α level if the comparisons are decided upon after examining the data. Comparison of the high mean with the low mean would be an example of this kind of comparison. There are alternatives to Scheffe''s method when only pairs of means (such as the high and low) are to be compared. Duncan's new multiple range test is an excellent alternative which divides the set of all means into groups which may be considered different at a proper α level. Application of this test to the cell means of this study at each time point within each condition shows that the high and low mean in each condition are significantly different for a .05 α level.

Since an AOV is essentially only an arithmetic process for partitioning the sum of squares into identified sources of variation, the AOV presented in Table 2 is certainly not the only possible basic partitioning. The analysis is governed by the basic point of view taken of the experiment. This is necessarily a subject matter decision, not a statistical decision. If the view is taken of the first phase that the experiment is designed to detect the differences in the CT response pattern over time in rats exposed to light during normal daylight hours and the pattern in rats kept in total darkness during (a) the time period of exposure (6 AM to 6 PM) and/or (b) the time when neither is receiving light (6 PM to 6 AM), then the AOV in Table 3 would perhaps be more appropriate. The AOV of Table 3 would directly furnish statistical support for the following conclusions:

i. In the hours of exposure to light the response pattern of LD rats differs from that of DD rats not only in the average

TABLE 3

ALTERNATE AOV FOR LD AND DD CONDITIONS

Source of Variation	DF	SS	MS	F
Among All Cells	23	1003261.516	43620.07	
Between Day and Night Hours	1	543664.720	543664.72	147.9**
Within Day Hours	11	360716.580		
Lighting Conditions	1	201377.460	201377.46	54.79**
Time Periods	5	46073.840	9241.77	2.51*
Lighting x Time	5	113265.276	22653.06	6.16**
Within Night Hours	11	98880.240		
Lighting Conditions	1	4737.960	4737.96	1.29
Time Periods	5	83060.520	16612.21	4.52**
Lighting x Time	5	11081.760	2216.35	< 1
Within All Cells	264	970248.418	3675.18	
Total	287			

**Indicates significance at .01; * at .05.

clotting time for the daylight hours (significant Within Day Hours Lighting Condition main effect), but also in the shape of the pattern (significant interaction Within Day Hours).

ii. In the time interval in which neither colony is receiving light no evidence of either a different average CT (failure of Within Night Hours Lighting Conditions main effect to be significant) or pattern of response (failure of Within Night Hours interaction to be significant) is found between the two colonies; however, the locus of the mean clotting time is not constant over the night hours (significant Within Night Hours Time main effect).

As a computational convenience the data for the day hours and night hours may be regarded as separate sets or experiments in the computations of their respective sums of squares for conditions, times, and interaction, but the information for the error term comes from each cell of both sets. Note that the within time period simple effects and the within LD and DD simple effects (ignoring the day-night classification) would be the same for the analysis of Table 2 and Table 3.

While both analyses convey, in a general way, the same information, slightly different hypotheses are being directly subjected to statistical tests in the two analyses. As an example consider the within night hours lighting condition main effect. This was tested in the second analysis and found to be not significant. One might have suspected that this would be the result since the simple effects within the 1930 to 0530 time periods reported in the first analysis were not significant; however, it is perfectly conceivable that the difference at each time

point would be so small as to be undetectable at any one point, but the information accumulated over all the times (the test of main effect) would be sufficient to detect the small difference.

Although a well constructed statistical analysis may be extremely helpful in extracting information, a large measure of the essential meaning is often more quickly imparted by a plot of the data. For studies in which the main interest is in pattern of response over time this is certainly true. In the plot of the CT response data means for each of the lighting conditions (Figure 1), some of the conclusions of the AOV would seem obvious, but others would not. The significant DD simple effect, for example, would be easily detected from the graph, but the failure of the LD and DD responses to be significantly different during the night hours is not so apparent. In the original article, this similarity is not noted.

Often a smoothing of the data by some process will be an aid in perceiving patterns obscured by variation of the data points. An often used procedure is the least squares fit of a polynomial to the data. Programs for this procedure are available at most computer facilities and if a plotter is available as peripheral equipment to the computer, even the tedium of plotting the large numbers of points necessary to produce a smooth curve is circumvented.

A sixth degree polynomial was fitted to the cell means for both the DD and LD conditions with 0730 considered as first time point (Figure 2). To illustrate the dependency of the fit of a polynomial upon choice of first time point, a sixth degree polynomial was also fitted to the means with 0130 considered as first time point (Figure 3). For

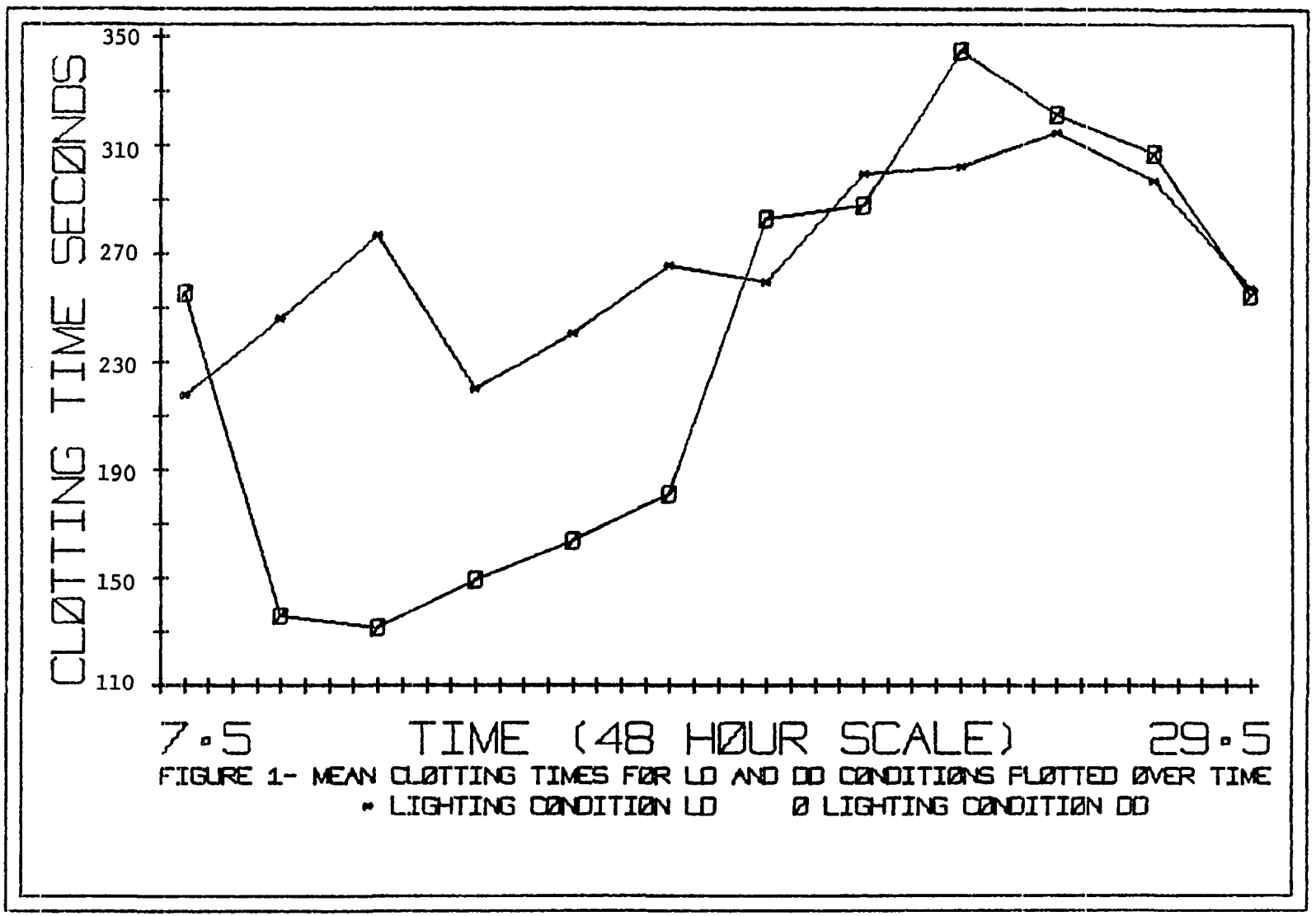


FIGURE 1- MEAN CLØTTING TIMES FØR LD AND DD CØNDITIØNS PLØTTED ØVER TIME

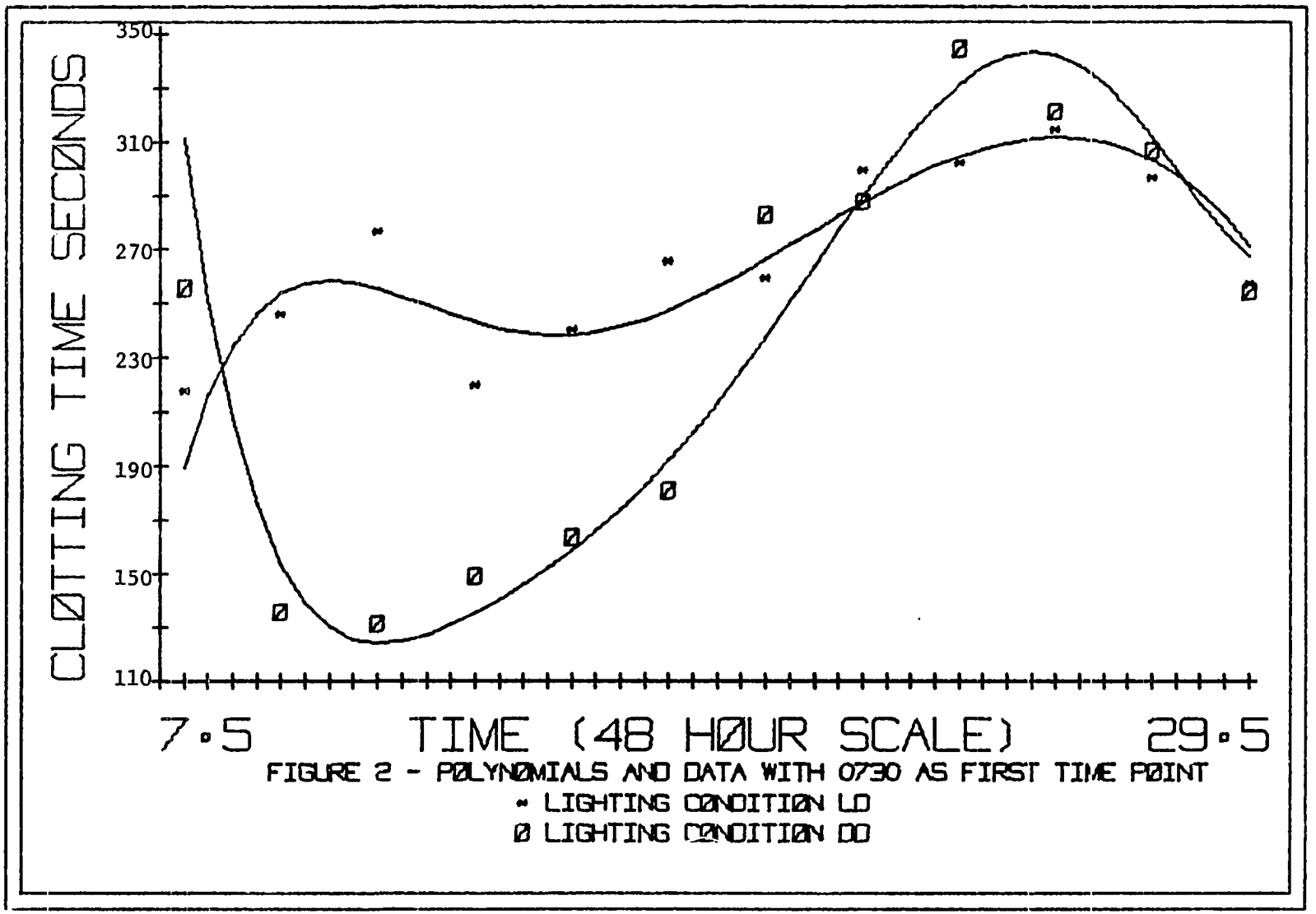


FIGURE 2 - PØLYNØMIALS AND DATA WITH 0730 AS FIRST TIME PØINT
* LIGHTING CØNDITIØN LD
Ø LIGHTING CØNDITIØN DD

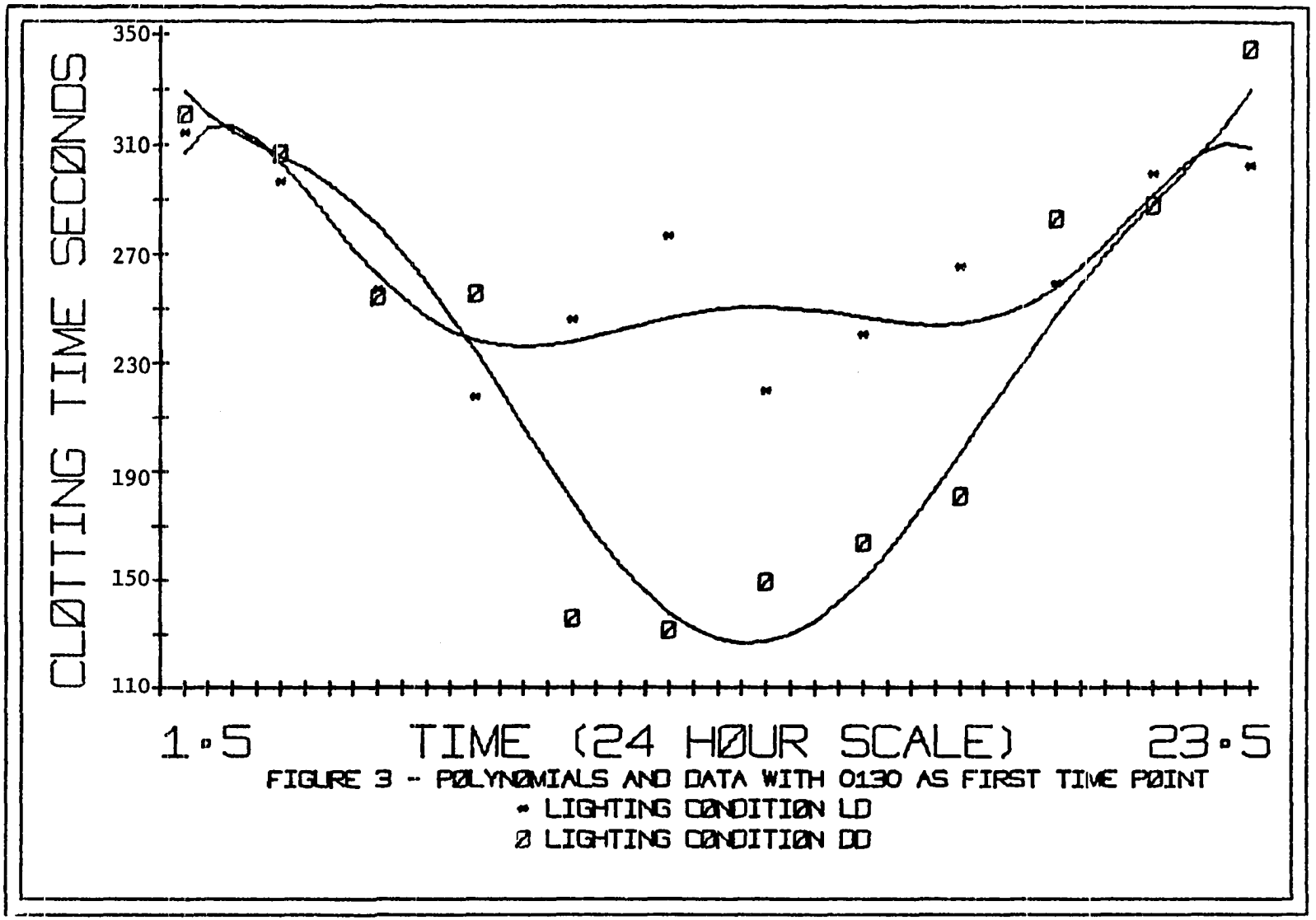


FIGURE 3 - PØLYNØMIALS AND DATA WITH 0130 AS FIRST TIME PØINT
* LIGHTING CØNDITIØN LD
Ø LIGHTING CØNDITIØN DD

the LD condition the 0730 polynomial accounted for 86.56% of the total variation of the cell means while the 0130 fit accounts for only 76.71%. For the DD condition the 0730 start yields a polynomial which accounts for 98.92% of the variation and the other accounts for 94.03%. These differences in the percent of variation accounted for by the regression reflect the difference in the adequacy of fit obtained for different choices of the serial relation of the time points.

The summary statistic for the LL and DL conditions are given in Table 4. Note that the number of rats measured is not the same at each time point. Therefore, the techniques used in the first phase must be modified slightly to account for the unequal cell frequencies. The technique of an unweighted means analysis seems appropriate for these data. This procedure is computationally simpler than the procedure based upon the least squares solution to the unequal cell frequency problem. The computational methods and conditions under which each of these analyses is proper are discussed in Winer (1962).

To illustrate the appearance of an AOV with more than two environmental conditions, the DD condition is included in the analysis for the second phase. The AOV computed by the unweighted means procedure is given in Table 5. Although the sums of squares are computed by a different technique, the interpretation of the effects is analogous to those given with the two previous AOV's.

In the analysis of the first phase data it was found that when the light stimulus was removed, the coagulation times of the rats were not significantly different from those of rats which had received no light. Hence, a comparison which (before observing the data) would seem

TABLE 4

SUMMARY STATISTICS FOR COAGULATION TIMES OF RATS UNDER LL AND DL LIGHTING CONDITIONS

Time Periods	LL Sample Size	LL Mean (Sec.)	LL Standard Error	DL Sample Size	DL Mean (Sec.)	DL Standard Error
0730	12	268.1	12.4	12	217.6	15.0
0930	13	255.0	13.9	12	150.4	26.4
1130	14	251.9	15.1	11	138.5	19.3
1330	14	274.4	12.7	11	149.5	20.8
1530	14	272.9	14.4	12	178.8	26.6
1730	14	284.1	15.6	11	181.3	22.1
1930	14	299.5	15.2	9	221.4	28.6
2130	14	305.8	16.2	12	237.1	24.0
2330	13	284.4	12.2	12	231.5	14.2
0130	12	305.1	16.9	9	216.6	6.2
0330	12	333.4	16.5	11	211.6	12.0
0530	12	292.2	14.4	11	210.9	19.3

TABLE 5

ANALYSIS OF VARIANCE FOR COAGULATION TIMES OF RATS UNDER DD, DL, AND LL LIGHTING CONDITIONS

Source of Variation	SS	DF	MS	F
Lighting Condition	586245.494	2	293122.747	79.03**
Time Period	725098.549	11	65418.050	17.77**
Condition x Period	293504.219	22	13341.101	3.60**
Within Cell	1479891.510	399	3709.00	

Time Periods Within Lighting Condition Simple Effect

Condition	SS	DF	MS	F
DD	791641	11	71967.3	19.40**
DL	155469	11	14133.5	3.81**
LL	71493	11	6499.3	1.75

**Indicates significance at .01; * at .05.

to be of interest is a comparison of measurements during the dark hours for the DL rats with the corresponding times for the DD rats. The results of this comparison are given in Table 6.

Note that the tests in the second phase analysis are not independent of those in the first phase since the DD data are included in both. However, the DD data used in computing the main effects and interaction sums of squares in Table 6 were not used in the computation of sums of squares for the Within Night Hours portion of the analysis in Table 3. The only data common to both comparisons is the contribution to the pooled error term made by the DD data. Thus, joint consideration of these two comparisons is not greatly influenced by the presence of DD data in both.

A possible interpretation of the joint consideration that suggests itself is that upon removal of the light stimulus the rats revert within a short period to a basic time dependent CT pattern manifested in rats which have not been exposed to light for nearly four months. In the discussion portion of the original article, the authors note that the CT pattern in rats does not seem to invert with an inversion of the light-dark cycle, and express concern for this seeming "resistance to inversion in blood coagulation rhythm" since "it previously had been reported that the characteristic rhythms in their (same rats) circulating eosinophils, neutrophils, and leucocytes as well as the rhythm in their plasma corticosterone levels all could be inverted by reversing the environmental LD cycle 180°."

If the interpretation suggested above is accepted, then no inversion would be expected and the dilemma is, at least partially,

TABLE 6

COMPARISON OF DARK HOURS FOR DL RATS
WITH SAME HOURS FOR DD RATS

Source of Variation	SS	DF	MS	F
Lighting Condition	2.942	1	2.94	<1
Time Period	164781.011	5	32956.20	8.89**
Condition x Period	8957.108	5	1791.42	<1
Error Term (From Table 5)			3709.00	

**Indicates significance at .01; * at .05.

resolved. Indeed, one might speculate that some of the other reported inversions may be manifestations of the light stimulus interaction with a basic no-light pattern rather than true inversions. These findings are in no way intended as criticism of the author' work, but are offered as an example of how slightly more sophisticated statistical techniques may be used to extract from the data a larger portion of the information concerning questions which are important to the researcher.

It is not intended that the analysis of the coagulation study presented here be construed as "the most appropriate" analysis possible. Many others would be equally, or perhaps more appropriate. The intention here is to illustrate the applicability of a basically analysis of variance approach to rhythmic time series data.

CHAPTER III

SOME THEORETICAL CONSIDERATIONS FOR TAA

Introduction

Many physiological variables are known to vary within a subject over time as well as among subjects. Often the variable of interest, say G, will in some manner vary within a subject around a central value and the purpose of much research is to obtain estimates for this central tendency. Conceptually, the central value $\overline{G}_A$ for subject A is a number assigned to the subject as a measure of G although if G were measured at a randomly chosen time, the probability of obtaining $\overline{G}_A$ as the actually measured value may be very small (possibly zero). Nevertheless, the concept of a central value is a useful abstraction since it offers a number with some stability over time which characterizes G within A over some time interval. In fact, its use has proven so fruitful that it seems to have dominated the thinking of much biological research. Other characteristics of the relationship of time and G within a subject, such as the variability over time, have often been ignored or used only to establish a measure of reliability for the central tendency number.

Procedures for classification, estimates of susceptibility to disease, etc. often depend upon an estimate of central tendency only. Variability enters the procedure only as a measure of reliability of the estimate and is seldom of itself used as a criterion. The common usage

of "Error Term" for the measure of variability indicates the general attitude toward this characteristic of G. This is not to say that variability is not recognized and used. Indeed, the analysis of variance is a much used tool in biological statistics, but the point of view is still directed toward the relationship of certain mean values rather than the variability over time. A comment heard often in discussions of medical studies is "The patient group seems more variable than the normals." Even though the difference is noted, it is often ignored and tests which require homogeneity of variance are applied. The robustness of these tests is such that little damage is done--perhaps. The point here is not to attack the use of Student's t tests and AOV procedures when variances seem slightly different, but to suggest that perhaps a useful characteristic of the phenomenon under study is being ignored in some investigations.

In practice, the mean is most often used as the measure of central tendency and the variance (or its square root) is used as a measure of variability. Researchers may be better able to use information phrased in these terms because of familiarity. For this reason, the concentration here will be upon exploring possible uses and properties of the usual estimates of means and variances under certain conditions.

Consider the familiar problem "Does group A differ from group B relative to variable G?" A straightforward and often used sampling approach to attempt to answer this question is made by obtaining N_A subjects in group A, N_B subjects in group B, measuring G <u>once</u> in each subject and presenting the usual sample estimates $\overline{G}_A$, $\overline{G}_B$, S_A^2, S_B^2 for each group mean and variance.

Assuming that conditions for the sample and measurements are such that the estimates are unbiased, there is still a problem concerning the variance that should be considered. The expected value of S^2 for each group is composed of two components--the within subject and the among subjects variances. Symbolically,

$$E(S_A^2) = \sigma_{SA}^2 + \sigma_{WA}^2 \quad \text{and}$$

$$E(S_B^2) = \sigma_{SB}^2 + \sigma_{WB}^2$$

where the S subscript denotes the variance among subjects and the W denotes the within subject variance. If only one measurement is made on each subject, the two variance components are completely confounded. Clearly, there are various possible combinations of relative sizes for the population values of these components such that $E(S_A^2) = E(S_B^2)$, but $\sigma_{SA}^2 \neq \sigma_{SB}^2$ and $\sigma_{WA}^2 \neq \sigma_{WB}^2$. Also, corresponding components from each group, say σ_{SA}^2 and σ_{SB}^2 may be equal, or nearly so, and large relative to σ_{WA}^2 and σ_{WB}^2. Under this condition $S_A^2 \simeq S_B^2$ may not be an unusual result from a sampling situation even though σ_{WA}^2 is very different from σ_{WB}^2.

Thus it can be seen that even if the question of difference in the groups is confined to means and variances, the above approach may obscure important information. Note that increased sample size does not mollify the problem.

Another common problem in biological research is the estimation of "normal" values. As in the problem of comparing two groups, this problem is often approached by drawing a sample of N subjects, measuring the quantity G once on each subject and reporting a mean and variance estimate from the sample. A second approach is to take several measurements on each subject, compute a mean for each subject, and use a mean

of the means as a central tendency measure of the population.

This second approach may be viewed as a two stage sampling problem. It may be shown that the second method may, under some circumstances, yield more precise estimates with fewer measurements and/or less cost than the first. The considerations necessary for choosing between these two and the allocation of measurements among subjects and within subjects is considered in Cochran (1965). The point to be made here is that within subject measurements over time may be of value in increasing precision and/or in decreasing cost even when the concern is for a central tendency measure only.

Although many of the rather general remarks above may be relevant for studies in which a Time Classification Analysis is to be used, the primary concern here is for some of the problems in implementing within subject measurements as a method of obtaining estimates of means and variances of values attained over time for Time Average Analysis.

For many biological situations no problem exists. If the variation within a subject around the subject's own central value is entirely random and a measurement at one sampling time may be considered as uncorrelated with any other, then the sample mean $\bar{y}$ and sample variance S^2 are statistics appropriate for the estimates and both have many of the desired properties of estimators. However, if in addition to the random variation, there is a variation which is functionally related with time, then the problem of estimates and meaningful interpretations becomes difficult. In a TCA the pattern over time of the central value around which the random variation occurs is usually observed or taken into account in some manner, but in TAA these two sources would be confounded in a

within subject estimate and complications result. In particular some of the desirable properties of $\bar{y}$ and S^2 as estimators may not be guaranteed.

If mathematical tools are to be employed in perceiving the problems and interpreting results from TAA, the behavior over time of the measured variable and sampling process employed must be couched in precise terms. Definitions of a model for a variable, sampling models, and terminology related to sampling over time are offered below in mathematically rigorous terms together with less rigorous discussions and motivations for the definitions. Following the definitions, some theoretical results are given in the form of an algorithm and theorems concerned with the effect on $\bar{y}$ and S^2 of functional variation over time. Although not all of the results are applicable exclusively to periodic variation, the emphasis is on this form.

In the theorems presented it is shown that under sampling plans often employed in obtaining within subject measurements, S^2 is a biased estimator of the variance when the pattern over time of the variable sampled is periodic. It is further shown that this bias is, in general, dependent upon the pattern, the scheme used to select the times for measurement, and the number of measurements taken within each subject.

The theorems indicate some of the misleading results that may be obtained if the cyclic nature of the measured variable is unsuspected or ignored. Fortunately, however, these theorems together with the computer sampling results given in CHAPTER V provide information for guidelines in constructing sampling schemes to overcome some of the difficulties in obtaining appropriate estimates for within subject means and variances for a TAA even if the pattern over time is cyclic.

A Model for a Physiological Variable Over Time

Before the terms mean, variance, or other characteristics of the distribution of values attained may be applied with specificity to the relationship of a variable G over time, some model must be assumed.

The model assumed and referenced as MI is the following:

$$G(t) = f(t) + \varepsilon_t$$

where ε_t, $G(t)$ are random variables for each real number t, f is a fixed function, and the joint distribution is such that for all marginals at s,t, $K(s,t) = E(\varepsilon_s\ \varepsilon_t)$ exist, $E(\varepsilon_t) = 0$ and $K(t,t) = \sigma_\varepsilon^2$ for all t, with the additional property that there exist a positive real number d* such that $d^* = \inf \{d \mid |\ s-t\ | \geq d \text{ implies } K(s,t) = 0\}$. Additional conditions common to stochastic processes of this kind, that is, "signal plus noise" such as covariance stationarity, where $K(s,t) = k(s-t)$, and $k(s-t)$ is continuous at 0, will often hold in the models of interest but these assumptions are not made here.

This model, then, describes a quantity G which follows the locus of f(t) over time, but which has in addition a random variation around the locus. The assumption of the possibility of covariance of ε's is introduced to account for the belief that in many biological situations measurements taken close together in time tend to be alike, but after a sufficient amount of time, d*, has elapsed, this tendency is no longer detectable. The covariance assumption may be viewed as an attempt to build into the model a belief that in the real physical situation a finite amount of time is required for G to change appreciably. The rapidity with which $E(\varepsilon_s\ \varepsilon_t)$ converges to zero is, roughly a measure of just how fast it is believed G may change relative to f.

Note that as a special case, MI may describe a function G which is an exact mathematical function defined on T by letting $\sigma_\varepsilon^2 = 0$.

Terminology for Sampling of a Function

In the sampling of G, a realization of MI, there must exist some scheme whereby the times for taking the measurements are selected. This scheme may be a rigidly defined schedule such as blood samples drawn every four hours, or a rather loose system such as instructing a subject to "come back in two weeks." No matter how accidentally or rigidly the selections were made, if n measurements are obtained, then some scheme has been used to select n times. This scheme, whatever it may be, will be called the sampling plan.

In the consideration of the mean and variance of G over an interval T, the mean of f over T is regarded as the "true" or population mean and the sum of the variance of f and σ_ε^2 as the population variance. In many realizations the mean of f is easily obtained by integrating f over the interval T and dividing by the length of T. With the mean thus computed, the variance of f is found by integrating f^2 over the interval T, subtracting the mean squared, and dividing this computation by the length of T.

Given below are some of the terms and notations used in considering the sampling of functions.

Let $T_s = \{x_i \mid i = 1, 2, \ldots m\}$ be a set of elements selected from a time interval T by sampling plan M.

Let G be a realization of MI defined on T and $g(t_i)$ the sample value at t_i.

The set $Y = \{g(t_i) \mid t_i \in T_s\}$ is said to be an m-size sample of (from) G over T by M. Forming a set in this manner will be called sampling G by (under) M.

The set of values attained by G over T will be denoted by GY(T).

The distribution function induced on GY(T) by M will be denoted by Q_{MT}.

Sampling G under M_1 over T_1 is said to be equivalent to sampling under M_2 over T_2 if and only if $GY(T_1) = GY(T_2)$ and $Q_{M_1T_1}$ is identical to $Q_{M_2T_2}$.

Definition of RAN

When repeated measurements within subjects are being made, the time points are seldom selected in a truly random manner from the sampled interval T. If a random sequence of times $\{t_i\}$ were selected, not only would the correlation of measurement taken close together in time create difficulty, but also the practical limitation of time required to take a measurement would often prohibit the execution of the randomly selected schedule. The use of automated devices such as the various patient monitoring systems may circumvent the latter objection, but true random sampling within subjects would be rare. There would, however, be sampling situations which are approximated by a random sampling model. If only one measurement is taken per subject where the subjects are selected for sampling from a homogeneous group at random times selected from T, then the function representing the group mean over time might be regarded as a function being randomly sampled.

Ran Sampling. A sampling plan M is said to be random, denoted by RAN (belongs to set RAN), if the set $T_s = \{t_i \mid i = 1, \ldots, m\}$ of time points at which G is to be measured is a random sample from the sampled interval T.

If G is a realization of MI which is a mathematical function ($\varepsilon^2 = 0$) defined on a sampled interval T, then the distribution induced by the RAN sampling scheme, Q_{RT}, may be constructed by the methods given in the algorithm below. In the consideration of a RAN sampling of a function it is often helpful to construct the associated Q_{RT} since an m-size sampling of G under RAN may be regarded as equivalent to drawing a simple random sample of size m from Q_{RT}. Following the algorithm, some of the special techniques in its application are illustrated by specific examples.

Algorithm for the Construction of Distribution and Density Functions of GY Given G

Let G be a bounded, continuous mathematical function defined on an interval T = [L,R] such that T may be partitioned into a finite number m of disjoint subintervals $\{T_k\} = T_1[a_o = L, a_1), T_2 = [a_1,a_2], \ldots, T_m = [a_{m-1}, a_m = R]$ where $i < j$ implies $a_i < a_j$ for $i,j = 0, 1, \ldots, m$ with $\{T_k\}$ having the following properties

(1) The union of all $T_j \in \{T_k\}$ is T.

(2) Within a given subinterval G is:

(a) monotonic strictly increasing,

or (b) monotonic strictly decreasing

or (c) constant.

If G is sampled over T by RAN, then the distribution function F

and the density function f of the set of values attained, GY, may be constructed in the following way:

I. For each $T_i \in \{T_k\}$ form F_i defined by $F_i(\mu) = P[G(t) \leq \mu \mid t \in T_i]$ $-\infty < \mu < +\infty$. F_i, therefore, is the distribution function of GY_i, where GY_i is the set of values attained by G in T_i. Equivalently, F_i may be viewed as the distribution obtained when G is sampled only over T_i.

Let G_i be the function equal to G in T_i with G_i undefined elsewhere. Then G_i^{-1} exists whenever the monotonic strictly increasing (decreasing) property of G in T_i holds.

Let GY_i MAX be the least upper bound and GY_i MIN be the greatest lower bound of GY_i.

Using this notation then, F_i will have one of the following forms.

(a) If $G_i = K_i$, a constant,

$$F_i(\mu) = \begin{cases} 1 & K_i \leq \mu \\ 0 & K_i > \mu \end{cases}$$

and GY_i MIN = GY_i MAX = $G(a_i)$

(b) If G_i is monotonic strictly increasing

$$F_i(\mu) = \begin{cases} 1 & G(a_i) \leq \mu \\ \dfrac{G_i^{-1}(\mu) - a_{i-1}}{a_i - a_{i-1}} & G(a_{i-1}) < \mu < G(a_i) \\ 0 & G(a_{i-1}) \geq \mu, \end{cases}$$

GY_i MIN = $G(a_{i-1})$ and GY_i MAX = $G(a_i)$

(c) If G_i is monotonic strictly decreasing

$$F_i(\mu) = \begin{cases} 1 & G(a_{i-1}) \leq \mu \\ \dfrac{a_i - G_i^{-1}(\mu)}{(a_i - a_{i-1})} & G(a_i) < \mu < G(a_{i-1}) \\ 0 & G(a_i) \geq \mu \end{cases}$$

The validity of the three forms is argued below.

For a given T_i, if $\mu \leq GY_i$ MIN, the probability of obtaining a value less than or equal to μ is certainly zero. If $\mu \geq GY_i$ MAX, the probability of obtaining a value less than or equal to μ is clearly 1. If GY_i MIN $< \mu <$ GY_i MAX, then the probability of obtaining a value less than or equal to μ must be considered separately for (b) and (c). Note that (a) need not be considered since GY_i MIN = GY_i MAX for (a).

If G_i is monotonic strictly increasing, then $G_i(t) < \mu < GY_i$ MAX if and only if $t \in I = [a_{i-1}, G_i^{-1}(\mu)]$, i.e., t must be in the interval with <u>left</u> end point the same as T_i and <u>right</u> end point $G_i^{-1}(\mu)$.

If G_i is monotonic strictly decreasing, then $G_i(t) < \mu < GY_i$ MAX if and only if $t \in I = [G_i^{-1}(\mu), a_i]$, i.e., t must be in the interval with <u>right</u> end point the same as T_i and <u>left</u> end point $G_i^{-1}(\mu)$.

For (b) or (c), the probability of $t \in I$ is the length of I divided by the length of T_i. For (b) then, the required probability is given by

$$\frac{G_i^{-1}(\mu) - a_{i-1}}{(a_i - a_{i-1})}$$

and for (c) by

$$\frac{a_i - G_i^{-1}(\mu)}{(a_i - a_{i-1})}$$

II. Let F be defined by $F(\mu) = P[G(t) < \mu]$ for $-\infty < \mu < +\infty$

by letting

$$F = \sum_{i=1}^{i=m} \frac{(a_i - a_{i-1})}{(R-L)} F_i$$

Since $\{T_k\}$ constitutes a partition of T into mutually exclusive sets, and $P[t \in T_i]$ is clearly the length of T_i divided by length of T, the validity of defining F as the weighted sum of F_i's follows from the relationship given by Parzen (1960):

$$P[B] = P[B|C_1]\ P[C_1] + \ldots + P[B|C_m]\ P[C_m]$$

where B is the event $G(t) < \mu$ and C_i is the event $t \in T_i$.

III. The density function may now be formed by differentiating F.

Sampling by RAN

The difficulty in using a RAN sampling plan in a real situation is briefly discussed above; however, consideration of RAN sampling in an idealized or theoretical situation may, hopefully, lend some insight into the problems of sampling a function.

To remove some of the difficulties, assume G is a realization of MI which is a mathematical function ($\sigma_\varepsilon^2 = 0$) defined on T and further assume that measurements may be taken instantaneously.

Now, if G is sampled over T by RAN, then the set of time points T_s is a simple random sample from T. Since each element of the set of function values Y is a function of an element of this sample of time points, it follows that Y is a simple random sample from the distribution Q_{RT} of the values attained by G over T. Thus, the theory of simple random sampling applies. In particular, the usual estimates for mean and variance using the elements of Y will be unbiased.

Even with these simplifications, the results are often difficult

to interpret. A "seemingly innocent" function which is bounded, continuous, and with derivitives existing at every point of T can have rather bizarre functions for the distribution and density of the values attained. Intuitive ideas of the nature of the distributions are often wrong even when the functioned sampled is known. To illustrate the above remarks and some of the special considerations in the application of the algorithm, four examples are considered.

Examples of Application of Algorithm

I. Let $G(t) = t^3 \qquad -1 \leq t \leq 1.$

Since G(t) is monotonic strictly increasing throughout the interval sampled, only one interval need be considered. Thus

$$F(\mu) = \begin{cases} 1 & 1 \leq \mu \\ \dfrac{\mu^{1/3} + 1}{(1 + 1)} & -1 < \mu < 1 \\ 0 & -1 \geq \mu \end{cases}$$

and $f = \dfrac{dF}{d\mu}$, the density function is found by

$$f = \frac{dF}{d\mu} = \begin{cases} 1/6\ \mu^{-2/3} & -1 < \mu < 1,\ \mu \neq 0 \\ \text{undefined} & \mu = 0 \\ 0 & \text{elsewhere} \end{cases}$$

Note that at the mid-point of its range, zero, the density is undefined and $\lim_{\mu \to 0+} f = \infty$ and $\lim_{\mu \to 0-} f = \infty$. However, f is a density function since

$$\lim_{a \to 0-} \int_{-\infty}^{a} f(\mu)\, d\mu + \lim_{a \to 0+} \int_{a}^{\infty} f(\mu)\, d\mu = 1 \quad \text{and}$$

$f(\mu) \geq 0 \quad -\infty < \mu < +\infty$.

This example was chosen to illustrate the "seemingly innocent"

function which has a discontinuous density. Further, the expected value of μ is zero, a value at which the density is undefined.

The next three functions were chosen because these same functions will be considered in the empirical portion of this investigation discussed below.

II. Let $G(t) = \text{Sin } 2\pi t \qquad 0 \leq t \leq 1.$

G is monotonic strictly increasing in T_1: $0 \leq t \leq 1/4$;

monotonic strictly decreasing in T_2: $1/4 \leq t \leq 3/4$ and

monotonic strictly increasing in T_3: $3/4 \leq t < 1$.

Thus
$$\left.\begin{aligned} G_1^{-1}(\mu) &= 1/2\pi \ \text{Sin}^{-1}\mu \qquad 0 < \mu < 1 \\ G_2^{-1}(\mu) &= 1/2\pi \ *\text{Sin}^{-1}\mu \qquad -1 \leq \mu < +1 \\ G_3^{-1}(\mu) &= 1/2\pi \ *\text{Sin}^{-1}\mu \qquad -1 \leq \mu < 0. \end{aligned}\right\}$$

The * in front of the inverse sine indicates that this is <u>not</u> the principal value range inverse. In the derivation of the algorithm, the range of the inverse function was <u>required</u> to be T_i. These may be written in terms of the principal value functions as

$$G_2^{-1}(\mu) = 1/2\pi(\pi - \text{Sin}^{-1}\mu) \qquad -1 < \mu < +1$$

$$G_3^{-1}(\mu) = 1/2\ (2\pi + \text{Sin}^{-1}\mu) \qquad -1 \leq \mu < 0.$$

In terms of these principal value inverses then

$$F_1(\mu) = \begin{cases} 1 & 1 \leq \mu \\ 2/\pi \ \text{Sin}^{-1}\mu & 0 < \mu < 1 \\ 0 & 0 \geq \mu \end{cases}$$

$$F_2(\mu) = \begin{cases} 1 & 1 \leq \mu \\ 1/2 + 1/\pi \ \text{Sin}^{-1}\mu & -1 < \mu < 1 \\ 0 & -1 > \mu \end{cases}$$

$$F_3(\mu) = \begin{cases} 1 & 0 \leq \mu \\ 1 + 2/\pi \ \mathrm{Sin}^{-1}\mu & -1 < \mu < 0 \\ 0 & -1 \geq \mu \end{cases}$$

The weights associated with the F_i's are 1/4, 1/2, and 1/4, respectively. Thus, for $-1 \leq \mu < 0$, $F(\mu) = 1/4(0) + 1/2(1/2 + 1/\pi \ \mathrm{Sin}^{-1} \mu) +$

$$1/4(1 + 2/\pi \ \mathrm{Sin}^{-1} \mu)$$

$$= 1/2 + 1/\pi \ \mathrm{Sin}^{-1} \mu.$$

For $0 \leq \mu \leq 1$,

$$F(\mu) = 1/4(2/\pi \ \mathrm{Sin}^{-1} \mu) + 1/2(1/2 + 1/\pi \ \mathrm{Sin}^{-1} \mu) +$$

$$1/4(1)$$

$$= 1/2 + 1/\pi \ \mathrm{Sin}^{-1} \mu.$$

Therefore,

$$F(\mu) = \begin{cases} 1 & 1 \leq \mu \\ 1/2 + 1/\pi \ \mathrm{Sin}^{-1} \mu & -1 < \mu < 1 \\ 0 & -1 \geq \mu \end{cases}$$

and

$$f(\mu) = \begin{cases} 1/\pi \ \dfrac{1}{\sqrt{1 - \mu^2}} & -1 < \mu < 1 \\ 0 & \text{elsewhere} \end{cases}$$

This result agrees with that obtained by Parzen (1960) by a slightly different procedure.

Note that F is exactly F_2. This is not a property peculiarly associated with the sine function, but an expression of a more general principle. Whenever a function f_1 on an interval I_1 may be transformed so as to coincide with a function f_2 on I_2, I_1, and I_2 having the same length, by a translation or a reflection in a line parallel to the function

axis, a RAN sampling of f_1 over I_1 and a RAN sampling of f_2 over I_2 are equivalent in the sense that the distribution and density functions will be the same. Roughly speaking, f_1 attains the same values with the same relative frequency over I_1 as does f_2 over I_2.

For the function $G(t) = \text{Sin } 2\pi t$, G in T_1 may be reflected onto G in the first half of T_2 through the line $t = 1/4$; G in T_3 may be reflected onto G in the second half of T_2 through $t = 3/4$. It follows then that sampling Sin $2\pi t$ on [1/4, 3/4] is equivalent to sampling on [0, 1] and the identity of F and F_2 is to be expected. This principle will be referenced as the "reflection principle".

Since a large class of functions will have the necessary symmetry so that the reflection principle may be used, a preliminary examination of the function for reflection lines may greatly reduce the work of constructing the distribution function. Noting the equivalence of sampling [1/4, 3/4] and [0, 1] in the above example would have reduced the work by 2/3.

III. Let $$G(t) = \begin{cases} 4t - 1 & 0 \le t \le 1/2 \\ 3 - 4t & 1/2 \le t \le 1 \end{cases}$$

Therefore, G(t) is monotonic strictly increasing in T_1: $0 \le t \le 1/2$ and monotonic strictly decreasing in T_2: $1/2 \le t \le 1$. Note that G in T_2 may be made to coincide with G in T_1 by reflection through $t = 1/2$. Thus the reflection principle may be used. Since

$$G^{-1}(\mu) = \frac{\mu + 1}{4} \qquad -1 < \mu < +1,$$

$$F(\mu) = \begin{cases} 1 & 1 \le \mu \\ 1/2(\mu + 1) & -1 < \mu < 1 \\ 0 & -1 \ge \mu \end{cases}$$

and

$$f(\mu) = \frac{dF}{d\mu} = \begin{cases} 1/2 & -1 < \mu < +1 \\ 0 & \text{elsewhere} \end{cases}$$

IV. Let $G(t) = \text{Sin}^5 2\pi t \qquad -\infty < t < +\infty.$

Since G is a periodic function, sampling G over the real number line is equivalent to sampling G over one period, say $[-1/2, +1/2]$. Then, by the reflection principle, this is equivalent to sampling G over $[-1/4, 1/4] = T_1$. This 1/2 period is chosen so that the inverses will be the usual principal value inverse functions.

Since $\quad G_1^{-1}(\mu) = 1/2\pi \ \text{Sin}^{-1}(\mu^{1/5}) \quad$ for $\quad -1 < \mu < +1$

$$F(\mu) = \begin{cases} 1 & 1 \leq \mu \\ 1/2 + 1/\pi \ \text{Sin}^{-1}(\mu^{1/5}) & -1 < \mu < 1 \\ 0 & -1 \geq \mu \end{cases}$$

and

$$f = \frac{dF}{d\mu} = \frac{1}{5\pi\mu^{4/5}\sqrt{1 - \mu^{2/5}}} \qquad -1 < \mu < +1 \text{ and } \mu \neq 0.$$

Definition of A and R Sampling

If the problem of selecting the initial time point t_o is ignored for the moment, many sampling plans for the selection of set $T_s = \{t_i | i=1, 2, \ldots m-1\}$ of time points at which the subsequent (m-1) measurements of an n-size sample of G over time are to be made may be described by two sets of models, A and R, defined below.

Let t_o be selected by any means.

$\{H_k\}$ is a sequence of distributions having finite ranges and defined as zero on $(-\infty, 0)$.

$\{d_k\}$ is a sequence of non-negative numbers selected à priori of the

selection of T_s.

A Sampling. Sampling plan M is said to be A (belong to set A) if for every $i > 0$

$$t_i = t_o + x_i + \sum_{j=1}^{j=i} d_j$$

where x_i is a random sample from H_i.

R Sampling. Sampling plan M is said to be R (belong to set R) if for every $i > 0$

$$t_i = t_o + \sum_{j=1}^{j=i} (d_j + x_j) ,$$

where for every j, x_j is a random sample from H_j.

The essential quality which distinguishes A from R is the effect of previous selection of time points on subsequent selection. After t_o has been selected the selection of t_i is independent in the probability sense of all other t's in A sampling, but in R sampling this would not, in general, be true. Note that for measurements taken always at exact times (x's can take on zero value only) A and R are the same.

As an example to illustrate A and R consider a clinical study which schedules a subject to return every 14 days following the initial visit at time t_o. Then for every i, $d_i = 14$. There are many circumstances that may arise such as personal business, illness, whim, etc., such that the first measurement taken after t_o is delayed, say by 2 days. If the remaining visits are still scheduled at the same time regardless of the delay, then the plan would be A. However, if the schedule were revised so that the third time is now 16 + 14 = 30 days after t_o, the 4th visit 14 days after the 3rd, etc., then the plan would be R. It is

assumed here that there exist distributions, though perhaps unknown, such that the occurrence of a delay is a random selection from one of these distributions.

The use of A and R to designate the two models is intended as a mnemonic since A is a sampling analogue of the additive model and R the analogue of the autoregressive model for functions described by Hurwicz (1962). Indeed it may be shown that sampling an additive function model by R is equivalent to sampling an autoregressive model at exact à priori designated intervals.

Definition of PA and PR

Since there is little expectation of knowing the exact distributions $\{H_k\}$ to use in defining a real sampling situation, the approach is to consider certain idealized situations and then, hopefully, be able to make judgements relative to real sampling. To facilitate the discussions, a subset PA of A and a subset PR of R are defined below.

Let MϵA or R and for all i and j

$H_i \equiv H_j \equiv H$ where H is such that x selected from H can take on only non-negative values and H has a finite range h.

PA Sampling. M is said to be a pseudo-systematic sampling type A denoted by PA (belongs to subset PA) if

1. MϵA
2. $d_1 = d_2 = \ldots = d_n = d$
3. $d > h$

PR Sampling. M is said to be a pseudo-systematic sampling type R denoted by PR (belongs to subset PR) if

1. $M \in R$

2. $d_1 = d_2 = \ldots d_n = d$

In the context of A and R, H will be called the "distribution of delays" and a random sample x from H will be called a "delay". $\{d_k\}$ will be called the set of minimum times.

In the clinical study example above, if the assumption is made that the distribution of delays is the same for all visits then the plan M for the study would be PA if (1) no delay greater than 14 could occur and (2) no revision of the schedule of visits is made regardless of the delays, or M would be PR if a reschedule to new 14 day intervals occurred following a non-zero delay.

Although there are many other situations where the sampling may be described by PA, the definition was motivated by studies such as the PA clinical study example above where measurements are planned at exact times, but "random" circumstances cause delays. However, the motivation for the definition of PR was the more loosely designed study in which the subject is measured, after some minimum time has elapsed, "at the next opportunity" following each measurement. In this context H might better be described as a distribution of next opportunities, but the term "delay" will be retained for both PA and PR. There are studies which of necessity have a PR sampling plan. If at each visit an action is taken, a drug administered, a treatment given, etc. and the effect of this is to be measured d time units later, when a delay occurs then the following visit must be rescheduled.

Expected Values of Mean and Variance Estimators

In Theorem 1 below, the expected value of S^2 is exhibited in terms

of sample size, variance of the population sampled, and covariance of the measurements for any sampling scheme such that, at each measurement, the expected value of the measurement is the mean of the population sampled, and the value obtained, when squared, has as expected value the second moment of the population. In Theorem 2 it is shown that in the application of Theorem 1 to the sampling of a realization of MI the possible bias indicated in Theorem 1 may be partitioned into two components - one reflecting the covariance of the function values at the sampling points and one reflecting the covariance of the random components around the locus. Theorem 3 shows that in this sampling the expected value of the sample mean is the population mean, but the variance of this estimator is a function of the bias of S^2.

Let $x_1, x_2, \ldots, x_n$ be a sample from a population P with mean μ and variance σ^2 under a sampling plan M such that for every i=1, 2, ..., n, $E(x_i) = \mu$ and $E(x_i^2) = \mu_2$, the second moment of P. Let

$$\bar{x} = \sum_{i=1}^{n} x_i/n \quad \text{and} \quad S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2/(n-1)$$

denote the estimates of mean and variance respectively. Note that the constraint present in independent sampling that for $i \neq j$ cov $(x_i, x_j) = 0$ is not placed on M.

The following lemma and theorem are listed for further reference. Their proofs follow in a straight forward manner from similar proofs found in elementary texts.

Lemma 1. $E(\sum_{i\neq j} x_i x_j) = \sum_{i\neq j} \operatorname{cov}(x_i, x_j) + n(n-1)\mu^2$.

Theorem 1. $E(S^2) = \sigma^2 - \frac{1}{n(n-1)} \sum_{i\neq j} \operatorname{cov}(x_i, x_j)$.

Clearly then, S^2 is an unbiased estimator of σ^2 if and only if $\sum_{i\neq j} \operatorname{cov}(x_i, x_j) = 0$. The coefficient (including the sign) and the summation will be called the bias.

Let G be a realization of MI. Let $Y = \{g_1, g_2, \ldots, g_n\}$ be the set of values obtained by sampling G under M* over an interval T. Clearly Y is a sample from the population P = GY(T) of all values possible for G to attain on T. It follows from the definition of G that the mean of P is $\bar{f}$ the mean of f over T. Let σ^2 denote the variance of P. Let the constraints on M above obtain for M*.

Theorem 2. If G is sampled over T under M*, then

$$E(S^2) = \sigma^2 - \frac{1}{n(n-1)} \sum_{i\neq j} \operatorname{cov}[f(t_i), f(t_j)] - \frac{1}{n(n-1)} \sum_{i\neq j} \operatorname{cov}(\varepsilon_i, \varepsilon_j).$$

Proof: By Theorem 1

$$E(S^2) = \sigma^2 - \frac{1}{n(n-1)} \sum_{i\neq j} \operatorname{cov}(g_i, g_j).$$

Since for any index k, $g_k = f(t_k) + \varepsilon_k$, the theorem follows from the additive property of covariance of sums.

Theorem 3. If G is sampled over T under M*, then the sample mean, $\bar{g}$, is an unbiased estimator of $\bar{f}$ and $\sigma^2_{\bar{g}}$ the variance of $\bar{g}$ may be expressed as $\sigma^2_{\bar{g}} = \sigma^2/n - (1 - 1/n)B$ where B is the bias given in Theorem 1.

Proof:

$$E(\bar{g}) = E\left(\sum_{i=1}^{n} g_i/n\right) = \sum_{i=1}^{n} E(g_i)/n = \bar{f}.$$

Thus $\bar{g}$ is unbiased. Since

$$\bar{g}^2 = 1/n^2 \left(\sum_{i=1}^{n} g_i^2 + \sum_{i\neq j} g_i g_j\right),$$

$$E(\bar{g}^2) = \mu_2/n + 1/n^2 \sum_{i\neq j} E(g_i g_j)$$

By Lemma 1

$$E(\bar{g}^2) = \mu_2/n + (1 - 1/n)\ \bar{f}^2 + 1/n^2 \sum_{i \neq j} \operatorname{cov}(g_i, g_j)$$

$$= \sigma^2 + \bar{f}^2 - \sigma^2/n + \bar{f}^2 - (1 - 1/n)B.$$

Hence $$\sigma^2_{\bar{g}} = E(\bar{g}^2) - [E(\bar{g})]^2$$

$$= \sigma^2/n + \bar{f}^2 - (1 - 1/n)\ B - \bar{f}^2$$

$$= \sigma^2/n - (1 - 1/n)\ B.$$

Sampling Periodic Functions by PR and PA

Of the infinitude of functions which may be used to define a particular realization of MI, the class of functions which exhibit a rhythmic variation around a central value create special problems for the investigator. Many times it is impossible to distinguish the rhythm of f from the random fluctuations around the locus. When the rhythmic nature of f is unsuspected or ignored and the usual estimate, S^2_G, of variance is made, the results may be biased and misleading interpretations made from the data. Unfortunately for the biological researcher, many of the variables of interest are rhythmic over time.

The purpose here will be to investigate a subset of the rhythmic functions -- those where the rhythm may be regarded as periodic over time. Although physiological variables, whose rhythmic variations have the nice properties of the mathematically precise periodic functions, may not be commonplace, many do exist which exhibit a close approximation to well known periodic functions. The urinary excretion of adrenaline reported by Levi (1968) is an example of a variable over time which closely approximates a sine curve. The various circadian and seasonal variations that

have been reported certainly may be considered periodic functions although, unfortunately, not a function so familiar as Levi's sine curve. Even if the rhythm of interest is very different from the periodic functions to be presented here, knowledge of the behavior of the bias and variance of the estimator for these functions may be used to design sampling plans which may increase accuracy or better allocate time and effort.

Clearly, the results from a PA or PR sample will be highly dependent upon the initial time point. (The special case of $f = c$ a constant is an exception). The investigation here is limited to sampling plans where the selection of t_o may be regarded as a random selection from an interval T whose length is equal to a period of the function sampled. Certainly there are many biological experiments and epidemiological surveys conducted for which the initial measurement time (entry into study, first visit to clinic, etc.) is not related to the position of the function in its cycle. In situations in which the value of the function does not, at least partially, define t_o, the assumption of a randomly selected initial time is not unrealistic.

With the preceding remarks as motivation, the following assumptions are made:

1. G is a realization of MI.
2. f is a continuous periodic function with period λ.
3. In the sampling of G, t_o is a random selection from the uniform distribution $U(0,\lambda)$.
4. The minimum time between measurements is such that the covariance of ε's is zero.

These four assumptions are in force for all theorems and discussions in the remainder of this chapter unless specifically stated otherwise.

Clearly, the assumption of positive time and first sampled point in $(0, \lambda)$ are merely a reflection of the choice of a convenient reference point for measuring time.

Since f is periodic, $E[G(t)] = E[G(t+\lambda)]$. The selection of t as a sample point is equivalent to selecting $t+\lambda$ in the sense that both will yield the same value of f. This property is the motivation for the following definitions.

Equivalence of Time Points. Time points t_1 and t_2 are said to be equivalent if $|t_1 - t_2| = \lambda$.

Equivalence of Intervals. Intervals T_1 and T_2 of the time axis are said to be equivalent if T_1 is T_2 or T_2 is a positive or negative translation of T_1 by an amount which is an integral multiple of λ.

Clique of an Interval. The union of intervals equivalent to an interval T is said to be the clique of T and will be denoted by C(T).

In order to show that under assumptions 1-4, PA and PR sampling of G are such that Theorems 1, 2, and 3 apply, the necessary constraints on the sampling scheme are shown to obtain for both A and R sampling plans and, therefore, for the subsets PA and PR.

In Theorem 4 below it is shown that in A or R sampling, the sample drawn at each time point is equivalent to a random sampling of size one of a periodic function over an interval of length one period. Therefore, the constraints on the sampling scheme in the previous theorems are seen to hold for A and R sampling. This result is established as a

corollary of Theorem 4. Thus, the expression for the possible bias of S^2 as an estimator of the variance of G exhibited in the previous section is seen to apply to A and R sampling of periodic functions. Also, the expected value of the within subject sample mean and the relationship of the variance of this estimator and the bias term are seen to be those given by Theorem 3.

Two preliminary results are established in the two lemmas preceding Theorem 4. The first of these merely develops a convenient expression for the distribution of the i^{th} sample time and the second establishes a sufficient condition for the equivalence of a sampling plan to a random sample over one period.

Lemma 2. If y and v are independent random variables such that y is distributed as the uniform $U(0, \lambda)$ and v has distribution F where F is zero for any negative value and $t = y + v$ has distribution Z, then

$$Z(x) = 1/\lambda \int_{x-\lambda}^{x} F(w)\, dw.$$

Proof: Since y and v are independent, Z may be expressed as the convolution of U and F. Thus

$$Z(x) = U(x) * F(x)$$

$$= \int_0^x \int_0^t \mu(t - w)\, dF(w)\, dt.$$

Since u, the density of y, is zero for all negative values in its domain,

$$Z(x) = \int_0^x [\, 1/\lambda \int_{t-\lambda}^{t} dF(w)]\, dt.$$

$$= 1/\lambda \int_0^x [F(t) - F(t-\lambda)]\, dt.$$

By letting $w = t-\lambda$ and noting that F is zero for any negative value in

its domain

$$\int_0^x F(t-\lambda)dt = \int_0^{x-\lambda} F(w)dw.$$

Thus,

$$Z(x) = 1/\lambda \left[\int_0^x F(t)dt - \int_0^{x-\lambda} F(w)dw\right] = 1/\lambda \int_{x-\lambda}^x F(w)dw.$$

For the next lemma assumption 3 requiring the first sample point to be selected from $U(0,\lambda)$ is not required for M^*.

Lemma 3. Let M^* be any sampling plan which selects exactly one sample point t^* from an interval T^* with length greater than or equal to λ.

Let M be a sampling which selects exactly one sample point t from the interval $T = (0,\lambda)$ by letting t be a random selection from the uniform $U(0,\lambda)$.

If I is a subinterval of T, denote the probability that $t \in I$ by P_I and the probability that t^* in C(I), the clique of I, by P_I^*.

If $P_I = P_I^*$ for every subinterval I of T, then sampling G under M is equivalent to sampling under M^*.

Proof: Since the length of T^* is at least λ, $GY(T^*) = GY(T)$. Let $T_x = \{t \mid G(t) \leq x \text{ and } t \in T\}$. Then, there exist a set $\{I_k\}$ of disjoint subintervals of T such that the union IU of these subintervals is equal to T_x except possibly for a set of measure zero. Hence

$$P[G(t) \leq x] = P[t \in IU] = \sum_k P[t \in I_K].$$

Since G is periodic, $P[G(t^*) \leq x] = \sum_k P[t^* \in C(t_k)]$. But for every k $P[t \in I_k] = P[t^* \in C(I_k)]$ by hypothesis. Therefore, $P[G(t) \leq x] = P[G(t^*) \leq x]$ and the plans are equivalent.

Theorem 4. If an (n+1)-size sample of G is drawn under $M \in A$ or R, then for every positive integer $s \leq n+1$ the sampling at the s^{th} sample

number is equivalent to a random sample of G over $T = (0,\lambda)$.

Proof: Assume $M \varepsilon R$. Choose a sample number $s \leq n+1$. Let $j = s-1$. Then at the s^{th} sample number t_j is the sample point selected and $t_j = t_o + \sum_{i=1}^{j} v_i$ where $v_i = (d_i + x_i)$. Let $v = \sum_{i=1}^{j} v_i$ and v be distributed as F. (Note the distribution of v may be found by the repeated convolution of the distributions for each v_i since v is the sum of independent variables.) By assumption 3, t_o is distributed as the uniform $U(0,\lambda)$. If Z_j denotes the distribution of the time point of the s^{th} sample number, then by Lemma 2

$$Z_j(x) = 1/\lambda \int_{x-\lambda}^{x} F(w)dw.$$

Let I be an arbitrary subinterval of $[0,\lambda]$ with left end point a and right end point b. Then, under a size one random sampling of G, the probability that the sample point is in I is $(b-a)/\lambda$.

Note that for every non-negative interval equivalent to I there exist a non-negative integer k such that the left end point is $a+k\lambda$ and the right end point is $b+k\lambda$.

For every k, therefore, the probability that t_j is in the associated interval is

$$Z_j(b+k\lambda) - Z_j(a+k\lambda) = 1/\lambda \int_{b+(k-1)\lambda}^{b+k\lambda} F(w)dw - 1/\lambda \int_{a+(k-1)\lambda}^{a+k\lambda} F(w)dw$$

$$= 1/\lambda \int_{a+k\lambda}^{b+k\lambda} F(w)dw - 1/\lambda \int_{a+(k-1)\lambda}^{b+(k-1)\lambda} F(w)dw.$$

Since t_j has a finite range, there exist an m such that $w > a+m\lambda$ implies $F(w) = 1$. Hence the probability that $t_j \varepsilon C(I)$ may be written as

$$(1/\lambda) \sum_{k=0}^{k=m} \left\{ \int_{a+k\lambda}^{b+k\lambda} F(w)dw - \int_{a+(k-1)\lambda}^{b+(k-1)\lambda} F(w)dw \right\}$$

$$= (1/\lambda) \int_{a+m\lambda}^{b+m\lambda} F(w)dw = 1/\lambda \ (b+m\lambda - a-m\lambda)$$

$$= (b - a)/\lambda.$$

Therefore, the theorem follows from Lemma 3 if MεR. The same argument with v defined by

$$v = x_j + \sum_{i=1}^{i=j} d_i$$

assures the theorem for MεA.

Corollary 1. Under the hypothesis and notation of Theorem 4, for any function ϕ such that the integrals exist

$$\int_{-\infty}^{\infty} \phi\ [G(x)]\ dZ_j(x) = 1/\lambda \int_0^{\lambda} \phi\ [G(t)]dt.$$

Proof: It is clear that this right hand integral is the expected value of ϕ associated with the random sampling of G over $(0,\lambda)$. The left integral is the expected value of ϕ associated with the sampling of G at the s^{th} sample number. The proposition follows from the equivalence of the two sampling plans as established in Theorem 4.

Corollary 2. If G is sampled under MεA or R, then the constraints upon the sampling scheme in the hypothesis of Theorems 1-3 obtain for M.

Proof: The proposition follows immediately from Theorem 4 and the properties of random sampling of a population .

Since Theorems 1, 2, and 3 apply to A and R sampling, they must apply for PA and PR. Theorem 2 shows that whenever the minimum time is sufficiently large to insure assumption 4 is in force, (the covariance

among the random components is zero) the bias is due entirely to the covariance of the values of f at the sample points. Clearly, investigations into the magnitude of the bias may be conducted by examining the sampling of mathematical functions, i.e., by assuming $E(\varepsilon^2) = 0$. Note, however, that the variance of the sample estimates $\bar{x}$ and S^2 are influenced by the assumption of no random component. Since the value of f at a sample point is independent of ε at that point, the adjustments of the variances for the presence of a random component are merely the usual adjustments for the sum of two independent variables. For the remainder of this chapter, G is assumed to be a mathematical function over time. The notation $T_n = \{t_0, t_1, \ldots, t_{n-1}\}$ will be used to denote the sequence of sampling points for an n-size sample and for $i = 0, 1, \ldots, n-1$, $g_i = G(t_i)$.

In Lemma 4 below it is shown that in PA sampling the covariance of any two sample values (not involving the first) is dependent upon their separation in the sampling sequence and not upon their position in the sequence. The covariance of the values at sample numbers 5 and 8, for example, would be the same as numbers 6 and 9, the same as 7 and 10, etc. This concept may be expressed in terms of the covariance matrix by stating that with the exception of the first row or column all numbers on a diagonal parallel to the main diagonal will be equal.

An analogous result for PR sampling is given in Lemma 5, but for PR sampling covariances involving the initial measurement are not exceptions as in PA sampling. Thus, the above description of the covariance matrix holds for PR sampling, but the numbers in the first row or column also being equal to the other numbers on the associated diagonal.

Theorems5 and 6 use the properties of these two lemmas to establish more compact expressions for the bias terms for PA and PR sampling respectively.

Lemma 4. For an n-size sample of G under McPA there exists, for every positive integer $j \leq n-2$, a constant C_j such that if $i > 0$ and t_i and t_{i+j} are terms of the sequence T_n,

$$\operatorname{cov}(g_i, g_{i+j}) = C_j.$$

Proof: Since covariances are invariant under a translation, it is sufficient to prove the lemma for G with mean zero.

Choose a positive integer $j \leq n-2$. Since G has mean zero, it follows from Corollory 2 and the definition of covariance that if i is as in hypothesis, $\operatorname{cov}(g_i, g_{i+j}) = K_{ij} = E(g_i\, g_{i+j})$. It follows from the definition of PA that $K_{ij} = E[G(t_o + id + x_i)\, G(t_o + id + jd + x_{i+j})]$. Since the distribution of all x's are the same by definition of PA,

$$K_{ij} = E\,[G(t_o + id + x_1)\, G(t_o + id + jd + x_2)].$$

Letting $v_i = t_o + id$ yields

$$K_{ij} = E\,[G(v_i + x_1)\, G(v + jd + x_2)].$$

Note that for every i, the distribution of v_i is the uniform $U(id, id+\lambda)$. Since G is periodic with period λ it follows that for every i the selection of v_i is equivalent to a random sampling from $U(0,\lambda)$. It follows from the definition of expected value that

$$K_{ij} = \int_o^h H(x_1) \int_o^h H(x_2) \{1/\lambda \int_{id}^{id+\lambda} G(v_i + x_1)\, G(v_i + jd + x_2)dv_i\}dx_2\, dx_1.$$

Since x_1 and x_2 are functionally independent of v, it follows from Corollary 1 that the expression in braces may be replaced by

$$1/\lambda \int_o^{\lambda} G(t + x_1)\, G(t + jd + x_2)dt.$$

Thus, the value of K_{ij} does not depend upon the choice of i. Letting $K_{ij} = C_j$ completes the proof.

Theorem 5. For an n-size sample of G under McPA, the bias B has the form

$$B = -2/n(n-1) \left\{ \sum_{j=1}^{n-1} \operatorname{cov}(g_o, g_j) + \sum_{j=1}^{n-1} (n - j-1)\, C_j \right\}$$

where for every j, C_j is the constant defined in Lemma 4.

Proof: By Theorem 1

$$B = -2/n(n-1) \sum_{i< k} \operatorname{cov}(g_i, g_k).$$

Clearly there are n-1 such covariances involving g_o. For each positive integer $j \leq n-2$, the set of integers $I_j = \{1, 2, \ldots, n-j-1\}$ is such that if $i \in I_j$, then $i + j \leq n-1$. Thus, there are exactly n-j-1 covariance terms of the form $\operatorname{cov}(g_i, g_{i+j})$. By Lemma 4, each of these is equal to C_j.

Lemma 5. For an n-size sample of G under McPR there exist, for every positive integer $j \leq n-1$, a constant C_j such that if $i \geq 0$ and t_i and t_{i+j} are terms of the sequence T_n, then

$$\operatorname{cov}(g_i, g_{i+j}) = C_j.$$

Proof: Again, since covariances are invariant under a translation, it is sufficient to prove lemma for G with mean zero.

Choose a positive integer $j \leq n-1$. Since G has mean zero, it follows from Theorem 5 and the definition of covariance that, if i is as in hypothesis,

$$\operatorname{cov}(g_i, g_{i+j}) = K_{ij} = E(g_i\, g_{i+j}).$$

By definition of PR

$$K_{ij} = E[G(t_i)\ G(t_i + jd + \sum_{k=i+1}^{i+j} x_k)].$$

Since by definition of PR all distributions of x's are identical, letting

$$y = \sum_{k=1}^{j} x_k$$

implies

$$K_{ij} = E[G(t_i)\ G(t_i + jd + y)].$$

Let v be the density of t_i and w the density of y. Note w is not functionally related to t_i. It follows from the definitin of expected value that

$$K_{ij} = \int_{-\infty}^{\infty} w(y) \{ \int_{-\infty}^{\infty} G(t_i)\ G(t_i + jd + y)\ v(t_i)dt_i \}dy.$$

Since y is not functionally related to t_i, it follows from Corollary 1 that the expression in braces may be replaced by

$$1/\lambda \int_0^{\lambda} G(t)\ G(t + jd + y)dt.$$

Thus, K_{ij} is not dependent upon the choice of i. Letting $C_j = K_{ij}$ completes the proof.

Theorem 6. For an n-size sample of MεPR the bias B has the form

$$B = -2/n(n-1) \sum_{j=1}^{n-1} (n-j)\ C_j$$

where for every j, C_j is the constant defined by Lemma 5.

Proof: By Theorem 1,

$$B = -2/n(n-1) \sum_{i>j} \operatorname{cov}(g_i, g_j).$$

For each positive integer $j \leq n-1$ the set $I_j = \{0, 1, \ldots, n-j-1\}$ is such that if $i \in I_j$, then $i + j \leq n-1$. Thus, there are n-j covariance

terms of the form $\text{cov}(g_i, g_{i+j})$. By Lemma 5, each of these is equal to C_j.

Note that in PR sampling it is not necessary to consider covariances involving g_o separately as in PA. This difference is illustrated in the calculation of the bias terms for a specific function below.

Theorem 7. Let G = A Sin Bt. Let the delay distribution be the uniform U(0, h) and d be the minimum time between measurements for both MεPR and M*εPA. The bias B_R for an n-size sample of G under M is given by

$$B_R = -2/n(n-1) \sum_{j=1}^{n-1} (n-j)\, A^2/2(2/Bh \text{ Sin } Bh/2)^j \cos Bj(d + h/2),$$

and the bias B_A for an n-size sample of G under M* is given by

$$B_A = -2/n(n-1) \left\{ \sum_{j=1}^{n-1} A^2/Bh \text{ Sin } Bh/2 \cos B(jd + h/2) \right.$$

$$\left. + \sum_{j=1}^{n-2} \frac{2A^2(n-j-1)}{(Bh)^2} \text{ Sin}^2(Bh/2) \cos B\, jd \right\}.$$

Proof: By use of the trigonemetric identity for the product of two sines

$$G(t)\, G(t + jd + y) = A^2/2 \cos B(jd + y) - A^2/2 \cos B(2t + jd + y).$$

Under M sampling for every positive integer $j \leq n-1$, C_j may be calculated using the method and notation of Lemma 5. The above identity implies

$$1/\lambda \int_o^\lambda G(t)\, G(t + jd + y)dt = A^2/\lambda \cos B(jd + y) \int_o^\lambda dt - A^2/2\lambda \int_o^\lambda \cos B(2 + jd + y)\, dt$$

$$= A^2/2 \cos B(jd + y) - 0.$$

The right most integral is clearly zero since it is a cosine function

integrated over twice its period.

Substituting this result in the expression for the calculation of C_j in Lemma 5 yields

$$C_j = A^2/2h^j \int_o^h \cdots \int_o^h \cos B(jd + \sum_{k=1}^{j} x_k)\, dx_1 \cdots dx_j$$

$$= A^2/2h^j \int_o^h \cdots \int_o^h \{1/B(\operatorname{Sin} B(jd + x_j + \sum_{k=1}^{j-1} x_k)|_o^k\} dx_1 \cdots dx_{j-1}$$

$$= A^2/2Bh^j \ (2 \operatorname{Sin} Bh/2) \int_o^h \cdots \int_o^h \cos B(jd + \frac{h}{2} + \sum_{k=1}^{j-1} x_k) dx_1 \cdots dx_{j-1}.$$

Note that after one integration the integrand has exactly the same form as before except that the constant term has been increased by Bh/2. Clearly then, after j such integrations

$$C_j = A^2/2B^j \ (2 \operatorname{Sin} Bh/2)^j \cos Bj \ (d + h/2).$$

The expression for B_R follows from Theorem 6.

Under M* sampling for every $j \leq n - 1$,

$$\operatorname{cov}(g_o, g_j) = E(g_o g_j) = E[G(t_o)\, G(t_o + jd + x_j)].$$

By using the above identity with $t = t_o$ and $y = x_j$ and definition of PA,

$$E(g_o g_j) = 1/\lambda h \int_o^h \int_o^\lambda A^2/2 \cos B(jd + x_j)\, d\lambda dx_j - A^2/2\lambda h \int_o^h \int_o^\lambda \cos(2t_o + jd + x_j)\, d\lambda dx_j$$

$$= A^2/2h \int_o^h \cos B(jd + x_j)\, dx_j - 0$$

$$= A^2/2hB \ [\operatorname{Sin} B(jd + x_j)\ |_o^h$$

$$= A^2/hB \operatorname{Sin} Bh/2 \cos B(jd + h/2)].$$

If $i \neq 0$, then $\operatorname{cov}(g_i, g_{i+j})$ may be calculated using the method and notation of Lemma 4. By use of the trigonemetric identity for the

product of two sines

$$G(t + x_1)\, G(t + jd + x_2) = A^2/2 \cos B(x_1 - x_2) - A^2/2 \cos(2t + jd + x_1 + x_2).$$

Thus,

$$1/\lambda \int_o^\lambda G(t+x_1)\, G(t + jd + x_2) = A^2/2\lambda \cos B(x_1-x_2) \int_o^\lambda d\lambda - A^2/2\lambda$$

$$\int_o^\lambda \cos B(2t + jd + x_1 + x_2)\, d\lambda$$

$$= A^2/2\lambda \cos B(x_1 - x_2) - 0.$$

Substituting this result in the expression for the calculation of C_j in Lemma 4 yields

$$C_j = A^2/2h^2 \int_o^h \int_o^h \cos B(x_1 - x_2)\, dx_2\, dx_1$$

$$= 2A^2/h^2 \quad \mathrm{Sin}^2(Bh/2) \cos Bjd.$$

The expression for B_A follows from Theorem 5.

From the above theorems it is seen that PA and PR sampling may yield a biased estimate of the population variance. In general this bias is a function of n, f, and the sampling plan. The complexity of these relationships are illustrated by the expressions for B_A and B_R in Theorem 7. Note that for n = 2, PA and PR sampling do not differ since there is no possible "rescheduling" and that the two expressions do indeed yield the same value.

Although Theorem 3 shows that the estimator for the mean is unbiased, the variance of this estimator is a function of the bias of the population variance estimator.

If mild regularity conditions were placed on the functions and delay distributions to be considered, then the magnitude of the bias, variance of $\bar{g}$, and other quantities may be found using the techniques

for the sine function in Theorem 7. However, the integration may prove formidable for certain functions and delay distributions.

No further theoretical distributions will be presented, but rather a systematic study of three functions under PA and PR sampling with uniform distribution of delays using computer simulation is described in CHAPTER IV and the results presented in CHAPTER V. The three functions sampled are defined as follows:

FI is the function defined by $FI(t) = \text{Sin } 2\pi t$.

FII is the function with period 1 such that

$$FII(t) = \begin{cases} 4t - 1 & 0 \leq t < 1/2 \\ 3 - 4t & 1/2 \leq t \leq 1. \end{cases}$$

FIII is the function defined by $FIII(t) = \text{Sin}^5 2\pi t$.

As can be seen in Figure 4, FI models a variable whose course over time is such that it is remote from its mean value most of the time, FIII models a variable which is near its mean most of the time, but with short lived excursions to the extremes and FII a variable whose time course is intermediate to the other two. Conclusions based upon the results of the simulated sampling of these functions, then, may act as guidelines in designing sampling schemes for a variable whose time course is not necessarily the same as any of these three functions, but is intermediate to two of them.

In the definitions, discussions, and development of the properties of PA and PR sampling, the terms "scheduled time" and "delay" have been used to describe the departure of these sampling schemes from that of the usual systematic sampling. The failure to have samples taken at these scheduled times has been depicted as some random event the outcome of

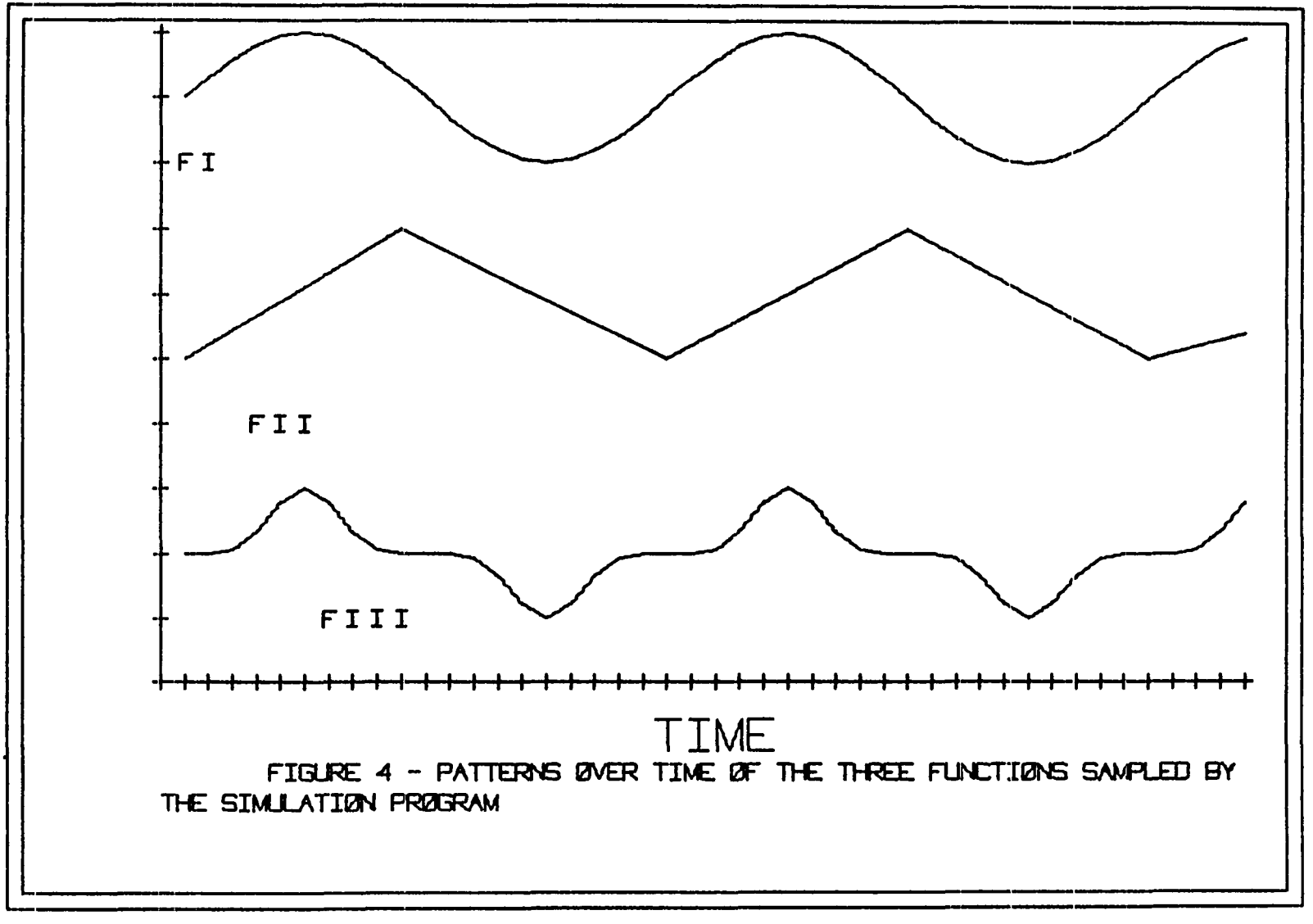


FIGURE 4 - PATTERNS ØVER TIME ØF THE THREE FUNCTIØNS SAMPLED BY THE SIMULATIØN PRØGRAM

which is not known prior to the scheduling. From this perspective the departures are unknown and perhaps even unsuspected during the planning stages and the theoretical and empirical results, when viewed in this context, seem, perhaps, to be meaningful only in ascertaining how much difficulty these "accidental delays" and cyclic nature of the variable have created in estimating the within subject mean and variances of a population. While these difficulties were certainly part of the motivation for this investigation, this seemingly negative purpose was not the only goal. The investigation of sampling with a uniform distribution of delays is intended not only to aid in perceiving the imbroglio of sampling periodic functions, but also to offer guidelines for designing sampling schemes which mollify, to some extent, the difficulties. Although a uniform distribution of the departures from the scheduled times would seldom obtain in a real sampling situation where the departures are dictated by events such as illness of subject, equipment failure, subject's access to clinic, etc., certain general properties may be seen in the sampling with this admittedly artificial distribution that may aid in the design and interpretation of investigations where the distributions are those dictated by the real world. Also, if the random element in the selection of time points is viewed in a mathematically equivalent, but conceptually different manner, the information from the empirical sampling is directly applicable to the problem. As part of the design, the random component may be introduced at the time of scheduling by selecting, for each subject, a sequence $x_1, x_2, \ldots, x_n$ where for every $i \leq n$ x_i is an independent selection from the uniform $U(0,h)$ and these values, together with the minimum time, used to define the times

of measurements. Using these "artificial delays" to select the time points is similar in purpose to the use of randomization in many experimental design situations. This point will be discussed later. Of course, the "accidental delays" from the other point of view given above would be superimposed on the scheduled times derived from the "artificial delays", but if the investigator can keep the range of these departures from schedule small, results from real world sampling should be quite similar to results from the computer simulation.

In the context of "artificial delays", it is, perhaps, more apparent that PA and PR sampling plans have aspects of both random and systematic sampling. Hence, some basis for the choice of the term pseudo-systematic sampling may be seen.

CHAPTER IV

METHODS OF EMPIRICAL INVESTIGATION

The basic program used was written to simulate a survey of 1500 subjects with 25 measurements over time made within each subject on a variable which, within each subject, has the same functional relationship over time. The program has as options, by way of subroutines and parameter cards, the exact function to be sampled and sampling scheme to be employed. The program was designed specifically for this investigation by the author with the assistance of Mr. Gary Haskin at the University of Oklahoma Medical School Computer Facility. The program was written in Fortran suitable for utilization of the Task System for an IBM 1800 computer with typewriter, cards, and two magnetic tape drives as input-output devices.

To avoid certain ambiguities in the discussion, the term "study" will be used to mean the simulation of a given function sampled under a given sampling plan with a given minimum time and a given range for the delay distribution. Hence, a study is completely designated by stating the functions, type of sampling (PA, PR. or RAN), minimum time, and range. The term "sample size" will refer to the number of values obtained within each of the 1500 subjects simulated in each study.

Since the functions sampled have a period of unit length, sampling with a minimum time of d is clearly equivalent to sampling with a

minimum time of d + k where k is any positive integer. In certain of the PA samplings, the requirement in the definition of PA that the minimum time be greater than the range is seemingly violated, but with the equivalence mentioned above and interpreting the minimum time in these studies as the stated value plus one, the difficulty is removed. With this convention used when necessary, each set of the equivalent sampling plans may be identified by using the smallest minimum time to designate the set. Thus, the inferences from the computer simulation of pseudo-systematic sampling is to a much larger set of sampling plans than the 19 distinct plans used. Note that this equivalence holds for the minimum time parameter, but not for the range. Sampling with a range of h is not necessarily equivalent to sampling with range h + k.

Utilizing the program discussed above, 27 studies were conducted using PR sampling, 27 using PA, and 3 using RAN. Each of the three functions FI, FII, and FIII defined in CHAPTER III were sampled under both PA and PR sampling with a separate study conducted for each of the 9 combinations of three minimum times (.25, .5, .75) and three ranges (.25, .5, .75) of a uniform distribution of delays. After some preliminary examination of the output from these studies, one additional sampling of FI under PR with minimum time .5 and range 1.25 was conducted.

For all sampling plans the random component associated with each sampling point was defined by a pseudo-random number generated by the subroutine RANDU furnished in the IBM Scientific Subroutine Package. For each study, five odd integers were selected from a table of random digits as initializing values for RANDU. Since the subroutine generates only 2^{13} numbers before cycling, a new initializing integer was used after each

7,500 sampling points. In this manner the 7,500 x 5 = 37,500 random components associated with each study were simulated. The first 1500 numbers generated were used to define the first sample point for each of the 1500 subjects, the next 1500 define the second within subject time, the next 1500 define the third, etc. Thus, each initializing digit for RANDU defined only 5 within subject values for each subject. Therefore, the within subject samples for a particular subject should not reflect any autocorrelation since the possibility of "recycling" the generator has been avoided.

The basic simulation had two forms of output for each study. As an example of the first, the summary statistics, abbreviated forms of the tables obtained for the RAN sampling of each function are given in Table 7. For each sample size n the entries in the table were computed in the manner described below.

Let x_{in} be the function value obtained for the n^{th} sample on the i^{th} subject. For each subject a running mean $\bar{x}_{in}$ was computed as

$$\bar{x}_{in} = \sum_{j=1}^{n} x_{ij} \; / \; n.$$

An estimate S^2_{in} for the within subject variance was computed by

$$S^2_{in} = \sum_{j=1}^{n} (x_{ij} - \bar{x}_{ij})^2 \; / \; (n - 1).$$

The table for $\bar{x}$ was then computed by

$$\bar{x} = \bar{x}_{.n} = \sum_{i=1}^{1500} \bar{x}_{in} \; / \; 1500.$$

The table entry for Variance of $\bar{x}$ was computed by

$$S^2_{\bar{x}n} = \sum_{i=1}^{1500} (\bar{x}_{in} - \bar{x}_{.n})^2 \; / \; 1499.$$

TABLE 7

SUMMARY STATISTICS OF RAN SAMPLING OF FI, FII, AND FIII

Function	N	$\bar{x}$	Variance of $\bar{x}$	$\bar{S}^2$	Variance of S^2	Bias	Bias/VAR
	1	-.0107	.4890	-	-	-	-
	2	-.0008	.2529	.4980	.3115	-.0020	-.003
	3	.0028	.1664	.4997	.1242	-.0003	-.000
FI	5	.0051	.0988	.5010	.0502	.0010	.002
	10	.0010	.0496	.5002	.0182	.0002	.000
	15	.0011	.0327	.5008	.0106	.0008	.001
	25	.0015	.0197	.5011	.0061	.0011	.0002
	1	.0122	.3285	-	-	-	-
	2	.0056	.1664	.3379	.1562	.0046	.013
	3	-.0064	.1097	.3356	.0664	.0022	.006
FII	5	-.0026	.0652	.3359	.0298	.0026	.007
	10	-.0004	.0326	.3347	.0114	.0014	.004
	15	-.0001	.0222	.3336	.0071	.0003	.000
	25	.0009	.0136	.3328	.0039	-.0005	-.001
	1	-.0068	.2340	-	-	-	-
	2	-.0041	.1233	.2434	.1154	-.0027	-.010
	3	-.0018	.0810	.2458	.0589	-.0003	-.001
FIII	5	.0018	.0489	.2470	.0293	.0009	.003
	10	.0019	.0240	.2456	.0132	-.0004	-.001
	15	.0010	.0163	.2451	.0083	-.0009	-.004
	25	.0005	.0097	.2460	.0049	.0001	.000

The table entry for $\bar{S^2} = S^2_{.n} = \sum_{i=1}^{1500} S^2_{in} / 1500.$

The table entry for Variance of S^2 was computed by

$$VAR(S^2_{in}) = \sum_{i=1}^{1500} (S^2_{in} - S^2_{.n})^2 / 1499.$$

The table entry for Bias was computed by subtracting the known variance of the sampled function from $S^2_{.n}$. The entry for Bias/Var was computed by dividing the empirically computed estimate of bias by the known variance.

The second output of the simulation program consisted of 25 sets of three vectors stored on magnetic tape for use as input to other programs. For each sample size the 1500 sample values were stored as a vector. The corresponding running sample mean and within subject sample variances were also stored. This tape was then used as input to a program which yielded frequency distributions for each sample size of the sample values, sample means, and sample S^2.

The frequency distributions of the sample values were used as a partial check to assure that the simulation was performing as expected. Since for each sample size the frequency distribution should approximate the known frequency distribution of a random sample of size 1500 from the sampled function, a radical departure from this expected distribution by any of the 25 in a study would indicate a possible programming error or a failure of the pseudo-random number generator to be adequate for the purpose of this simulation. One such check was made with PA type sampling and one with PR for each of the three functions.

The frequency distribution for each sample size of the sample

means and S^2 were obtained for each study. Tables and graphs prepared from these frequency distributions are given in CHAPTER V.

For each study four groups of size 20 were selected by sampling without replacement from the 1500 simulated subjects. The magnetic tape output from the simulation program was then used as input to another program which computed, for each sample size, the group mean and pooled within subject estimate of variance of the sampled function. Graphs showing the group means as a function of sample size were plotted using a Calcomp 1627 plotter under control of an IBM 1800 computer system. The four functions were plotted on the same axis set. Similar plots were made for the estimates of variance as a function of sample size. Thus, a set of 58 means and 58 estimates of variance graphs were produced. Selections from these sets are presented in CHAPTER V.

In addition to these graphs, the 25 sample values for a randomly selected subject were also plotted as a function of sample number for several of the studies.

The decision to use 20 subjects per group to investigate the pattern over sample size was based upon two factors. First, 20 subjects is not an unrealistically large number for a group size in clinical studies. Secondly, it was felt that group sizes of much less than 20 would obscure the relationship of bias and sample size because of sampling variation. Prior experience in sampling the sine function was used in making the latter judgement.

Several techniques were used to assure the validity of the simulation program in addition to the frequency distribution checks mentioned above. The expected value of the quantities computed for each of the

functions under a RAN sampling are of course known from statistical theory. The output from each of the RAN samples was carefully examined for abnormal departure from these expected values. Table 7 gives the abbreviated forms of the tabular output from the RAN samplings.

Since certain subroutines are used in the PA and PR sampling that are not used in the RAN sampling, this portion of the program was investigated by using these two sampling types to obtain samples which are theoretically equivalent to RAN sampling. This was accomplished by setting the range of the uniform delay distribution equal to the period of the function. Thus, the delays would have a uniform distribution, U(0,1), and any selection of a delay would define a random selection from an interval with length exactly one period. Note that neither the value of the minimum time, nor the sampling type affects the equivalence to RAN sampling. One study was performed with a PA type sampling for FI with minimum time as .5 and delay range of 1 and another for the same function and delay range with minimum time as .25 and PR sampling type. The results of these two studies were examined for differences one from the other and from the RAN sample. Only minor differences were noted and these, it was felt, could be attributed to sampling variation.

After these checks had provided evidence for the essential validity of the simulation, the remainder of the studies were performed. In the sampling of FI under both PA and PR, further evidence of the adequacy of the simulation was given by the rather close agreement of the empirical estimates of bias as compared with the theoretical values given by the formulas of CHAPTER III.

CHAPTER V

RESULTS OF COMPUTER SAMPLING

Summary Statistics

In evaluating the results of the computer sampling it should be noted that for a given sample size within a given study, the $\overline{S^2}$ and estimate of variance of S^2 computed are based upon a simple random sample of size 1500 from a distribution of within subject S^2's which, in general, is not the same distribution sampled for the computation of $\overline{S^2}$ and estimate of variance of S^2 for any other sample size in the same study nor in any other study. Indeed, examination of the frequency distributions of the values of S^2 obtained in the simulation for samples of size 2 and 3 differ radically from those of larger sample sizes. However, a mean based upon a sample of size 1500 should be a sufficiently accurate estimate of the expected value to allow reliable comparisons if the variance of the sampled distribution is not extremely large. The maximum estimate for the variance of S^2 was .3790 obtained for sample size 2 in the sampling of FI with minimum time .25 and range .5. Thus, even this extreme does not seem to invalidate a high degree of confidence in the accuracy of the computed $\overline{S^2}$ as an estimate of the expected value of S^2 and hence the accuracy of the corresponding estimate of the bias for every sample size of every study. An even greater degree of confidence may be expressed for estimates associated with larger sample sizes since in every

study there was a general downward trend in the estimate of the variance S^2 with increasing sample size. This trend was not quite monotonic in every study, but within each study for a difference of 2 in the sample size the estimate of the variance of S^2 was smaller for the larger sample size. The rate of decrease with sample size seems much greater for the smaller sample sizes. As an example consider the study in which the maximum variance estimate was obtained for sample size 2. At sample size 4 the estimate had decreased to .0653 and for size 5 to .0437. Within each study all sample sizes greater than 5 had an estimate of variance of S^2 which were smaller than that obtained at size 5. The maximum value over all studies at sample size 5 was .0846 for FI under PA with minimum time .5 and range .25, and within each function the maximum was obtained under exactly the same sampling plan. These values were .0530 for FII and .0579 for FIII.

The magnitude of the estimate of bias was also observed to exhibit a somewhat similar pattern over sample size. Within each study, the greater magnitudes were observed for the smaller sample sizes and a general downward trend of magnitude as sample size increased was found. These two trends are illustrated in Table 8 which is a copy of the initial output of the summary statistics obtained from the simulation program for the sampling of FII under PR with minimum time of .25 and range .5. The complete table of summary statistics was selected for presentation as a typical example of the simulation of the pseudo-systematic samplings. Abbreviated forms of the summary statistics for all studies are given in Tables 9, 10 and 11.

In comparing the statistics of these tables it should be noted

TABLE 8

EXAMPLE TABLE OF SUMMARY STATISTICS FOR PSEUDO-SYSTEMATIC SAMPLING SIMULATION

Size	$\bar{x}$	Variance of $\bar{x}$	$\bar{S}^2$	Variance of S^2	Bias	Bias/Var
1	-0.004364	0.343855	0.000000	0.000000	-0.333333	-1.000
2	-0.007446	0.063923	0.532303	0.196210	0.198969	0.596
3	-0.003654	0.049485	0.420263	0.061969	0.086930	0.260
4	-0.002681	0.028692	0.406228	0.040671	0.072894	0.218
5	0.001085	0.022610	0.386772	0.027707	0.053438	0.160
6	-0.002644	0.017649	0.375905	0.021595	0.042571	0.127
7	-0.002282	0.015030	0.369637	0.018068	0.036304	0.108
8	-0.000530	0.012014	0.366137	0.015377	0.032804	0.098
9	-0.002820	0.010692	0.360456	0.013089	0.027122	0.081
10	-0.001225	0.009446	0.357143	0.011771	0.023810	0.071
11	-0.001598	0.008311	0.356035	0.010303	0.022702	0.068
12	-0.000905	0.007452	0.353209	0.009234	0.019876	0.059
13	-0.001462	0.006949	0.351823	0.008175	0.018490	0.055
14	-0.000486	0.006320	0.349426	0.007577	0.016093	0.048
15	0.000348	0.005798	0.348733	0.006981	0.015400	0.046
16	0.000227	0.005437	0.347754	0.006566	0.014421	0.043
17	0.000371	0.004935	0.347115	0.006027	0.013781	0.041
18	-0.000331	0.004736	0.345323	0.005627	0.011989	0.035
19	0.000513	0.004432	0.344660	0.005326	0.011327	0.033
20	0.001117	0.004209	0.344016	0.005056	0.010682	0.032
21	0.001270	0.003956	0.343932	0.004764	0.010599	0.031
22	0.001547	0.003716	0.343839	0.004459	0.010506	0.031
23	0.001370	0.003608	0.343176	0.004228	0.009843	0.029
24	0.001930	0.003350	0.342556	0.004014	0.009223	0.027
25	0.002111	0.003256	0.342409	0.003922	0.009076	0.027

TABLE 9

SUMMARY STATISTICS OF COMPUTER SAMPLING OF FI (SIN 2 PIT)

			PA Sampling				PR Sampling			
Min Time	Range	N	Variance of $\bar{x}$	$\bar{S}^2$	Variance of S^2	Bias	Variance of $\bar{x}$	$\bar{S}^2$	Variance of s^2	Bias
		2	.0882	.8380	.3762	.3380	.0914	.8280	.3664	.3280
		3	.0264	.7105	.0645	.2105	.0249	.7227	.0627	.2227
.25	.25	5	.0349	.5810	.0269	.0810	.0136	.6071	.0242	.1071
		10	.0114	.5444	.0112	.0444	.0058	.5502	.0084	.0502
		15	.0064	.5294	.0061	.0294	.0033	.5331	.0049	.0331
		2	.0896	.8299	.3586	.3299	.0889	.8241	.3790	.3241
		3	.0980	.6023	.1125	.1023	.0662	.6494	.1019	.1494
.25	.5	5	.0688	.5429	.0440	.0429	.0319	.5821	.0437	.0821
		10	.0282	.5302	.0166	.0302	.0132	.5409	.0165	.0409
		15	.0191	.5191	.0102	.0191	.0082	.5267	.0267	.0267
		2	.1922	.6069	.3447	.1069	.1989	.6130	.3553	.1130
		3	.1602	.5003	.1186	.0003	.1129	.5895	.1144	.0895
.25	.75	5	.0884	.5064	.0481	.0064	.0633	.5535	.0428	.0534
		10	.0443	.5039	.0184	.0039	.0309	.5248	.0148	.0248
		15	.0306	.5013	.0112	.0013	.0192	.5177	.0089	.0177
		2	.0931	.8146	.3759	.3146	.0875	.8206	.3672	.3206
		3	.0771	.6297	.1274	.1297	.0264	.7073	.0644	.2073
.5	.25	5	.0350	.5750	.0846	.0750	.0129	.6064	.0239	.1064
		10	.0113	.5381	.0651	.0381	.0059	.5485	.0079	.0485
		15	.0077	.5199	.0581	.0199	.0036	.5321	.0045	.0321

TABLE 9--Continued

Min Time	Range	N	PA Sampling Variance of $\bar{x}$	PA Sampling $\bar{S}^2$	PA Sampling Variance of S^2	PA Sampling Bias	PR Sampling Variance of $\bar{x}$	PR Sampling $\bar{S}^2$	PR Sampling Variance of S^2	PR Sampling Bias
		2	.2535	.5020	.3045	.0020	.2537	.4969	.3097	-.0031
		3	.1227	.5755	.1160	.0755	.1321	.5521	.1197	.0521
.5	.5	5	.0668	.5427	.0442	.0427	.0573	.5522	.0385	.0522
		10	.0343	.5206	.0170	.0206	.0246	.5251	.0158	.0251
		15	.0208	.5171	.0099	.0171	.0150	.5173	.0092	.0173
		2	.3086	.3956	.2565	-.1044	.2982	.3884	.2450	-.1116
		3	.1576	.5174	.1274	.0174	.2092	.4309	.1137	-.0691
.5	.75	5	.0931	.5101	.0532	.0101	.1265	.4654	.0514	-.0346
		10	.0473	.5006	.0209	.0006	.0650	.4825	.0177	-.0174
		15	.0313	.4993	.0149	-.0007	.0464	.4841	.0102	-.0158
		2	.4086	.1813	.0528	-.3187	.4147	.1834	.0521	-.3166
		3	.1709	.4960	.1091	-.0040	.3033	.2901	.0624	-.2099
.75	.25	5	.0355	.5783	.0256	.0783	.1294	.4611	.0459	-.0389
		10	.0244	.5282	.0094	.0282	.0343	.5179	.0080	.0179
		15	.0119	.5236	.0060	.0236	.0192	.5183	.0046	.0183
		2	.4026	.1863	.0526	-.3137	.3956	.1787	.0486	-.3213
		3	.2340	.3971	.1146	-.1029	.3458	.2196	.0477	-.2804
.75	.5	5	.0660	.5444	.0452	.0444	.2804	.2753	.0401	-.2247
		10	.0411	.5137	.0173	.0137	.1772	.3598	.0295	-.1402
		15	.0254	.5112	.0104	.0112	.1295	.3967	.0212	-.1033
		2	.3031	.3787	.2392	-.1213	.3017	.4008	.2475	-.0992
		3	.2133	.4303	.1198	-.0697	.2112	.4314	.1157	-.0686
.75	.75	5	.0937	.5085	.0505	.0085	.1324	.4579	.0484	-.0421
		10	.0484	.5034	.0169	.0034	.0674	.4810	.0180	-.0190
		15	.0335	.5005	.0108	.0005	.0460	.4870	.0108	-.0130

TABLE 10

SUMMARY STATISTICS OF COMPUTER SAMPLING OF FII (TRIANGLE)

			PA Sampling				PR Sampling			
Min Time	Range	N	Variance of $\bar{x}$	$\bar{S}^2$	Variance of S^2	Bias	Variance of $\bar{x}$	$\bar{S}^2$	Variance of S^2	Bias
.25	.25	2	.0625	.5300	.2077	.1967	.0648	.5231	.2004	.1898
		3	.0186	.4678	.0464	.1345	.0181	.4641	.0476	.1307
		5	.0235	.3865	.0184	.0531	.0100	.4047	.0174	.0713
		10	.0076	.3610	.0075	.0277	.0047	.3648	.0061	.0314
		15	.0042	.3513	.0043	.0180	.0027	.3533	.0036	.0199
.25	.5	2	.0650	.5341	.2010	.2008	.0639	.5323	.1962	.1990
		3	.0673	.3983	.0664	.0650	.0495	.4203	.0620	.0869
		5	.0487	.3572	.0274	.0239	.0226	.3868	.0277	.0534
		10	.0186	.3488	.0109	.0154	.0094	.3571	.0117	.0238
		15	.0133	.3442	.0067	.0109	.0058	.3487	.0070	.0154
.25	.75	2	.1352	.3923	.1737	.0590	.1307	.4020	.1837	.0686
		3	.1147	.3254	.0635	-.0079	.0786	.3772	.0660	.0439
		5	.0632	.3363	.0282	.0030	.0454	.3578	.0271	.0244
		10	.0293	.3378	.0113	.0045	.0212	.3445	.0100	.0111
		15	.0211	.3341	.0067	.0008	.0142	.3399	.0062	.0065
.5	.25	2	.0645	.5272	.1991	.1939	.0635	.5361	.2151	.2027
		3	.0538	.4109	.0768	.0776	.0178	.4724	.0487	.1391
		5	.0246	.3786	.0530	.0452	.0099	.4058	.0174	.0724
		10	.0081	.3516	.0411	.0182	.0052	.3653	.0057	.0320
		15	.0056	.3395	.0376	.006	.0030	.3555	.0035	.0222

TABLE 10--Continued

Min Time	Range	N	PA Sampling Variance of $\bar{x}$	PA Sampling $\bar{S}^2$	PA Sampling Variance of S^2	PA Sampling Bias	PR Sampling Variance of $\bar{x}$	PR Sampling $\bar{S}^2$	PR Sampling Variance of S^2	PR Sampling Bias
		2	.1754	.3369	.1605	.0036	.1653	.3333	.1594	.0000
		3	.0840	.3860	.0723	.0527	.0842	.3701	.0687	.0368
.5	.5	5	.0440	.3680	.0293	.0347	.0379	.3696	.0260	.0363
		10	.0225	.3485	.0115	.0152	.0171	.3471	.0107	.0137
		15	.0143	.3441	.0069	.0108	.0112	.3442	.0067	.0108
		2	.2004	.2849	.1322	-.0484	.2007	.2646	.1225	-.0687
		3	.1021	.3608	.0711	.0275	.1408	.2879	.0625	-.0454
.5	.75	5	.0605	.3484	.0322	.0150	.0850	.3081	.0277	-.0252
		10	.0322	.3395	.0145	.0062	.0437	.3226	.0108	-.0108
		15	.0207	.3391	.0099	.0057	.0294	.3263	.0064	-.0070
		2	.2707	.1242	.0173	-.2091	.2647	.1270	.0178	-.2063
		3	.1127	.3328	.0540	-.0005	.2003	.1919	.0243	-.1415
.75	.25	5	.0232	.3872	.0176	.0538	.0891	.2998	.0235	-.0335
		10	.0154	.3521	.0070	.0188	.0220	.3435	.0053	.0102
		15	.0078	.3477	.0044	.0144	.0118	.3412	.0033	.0078
		2	.2671	.1269	.0178	-.2064	.2659	.1274	.0177	-.2060
		3	.1525	.2712	.0562	-.0621	.2277	.1537	.0199	-.1797
.75	.5	5	.0459	.3565	.0273	.0231	.1838	.1874	.0179	-.1459
		10	.0272	.3395	.0107	.0061	.1160	.2423	.0137	-.0911
		15	.0171	.3378	.0068	.0045	.0825	.2691	.0095	-.0642
		2	.2121	.2544	.1203	-.0790	.2069	.2645	.1181	-.0688
		3	.1516	.2812	.0616	-.0521	.1454	.2888	.0600	-.0446
.75	.75	5	.0645	.3371	.0285	.0037	.0840	.3110	.0273	-.0223
		10	.0337	.3318	.0121	-.0015	.0445	.3197	.0117	-.0137
		15	.0222	.3334	.0073	.0001	.0290	.3242	.0067	-.0091

TABLE 11

SUMMARY STATISTICS OF COMPUTER SAMPLING OF FII (SIN^5 2 PIT)

Min Time	Range	N	PA Sampling Variance of $\bar{x}$	PA Sampling $\overline{S^2}$	PA Sampling Variance of S^2	PA Sampling Bias	PR Sampling Variance of $\bar{x}$	PR Sampling $\overline{S^2}$	PR Sampling Variance of S^2	PR Sampling Bias
.25	.25	2	.0651	.3663	.1979	.1202	.0668	.3438	.1726	.0977
		3	.0324	.3273	.0650	.0812	.0319	.3169	.0631	.0708
		5	.0222	.2791	.0251	.0330	.0185	.2795	.2369	.0334
		10	.0091	.2601	.0101	.0140	.0091	.2609	.0090	.0148
		15	.0057	.2549	.0060	.0088	.0059	.2580	.0060	.0119
.25	.5	2	.0637	.3453	.1868	.0991	.0660	.3709	.1958	.1248
		3	.0528	.2792	.0714	.0332	.0046	.3068	.0745	.0607
		5	.0350	.2610	.0317	.0149	.0254	.2816	.0341	.0355
		10	.0151	.2548	.0132	.0087	.0122	.2604	.0147	.0144
		15	.0109	.2489	.0080	.0028	.0079	.2567	.0087	.0106
.25	.75	2	.1085	.2873	.1413	.0413	.1031	.2785	.1309	.0325
		3	.0846	.2487	.0583	.0026	.0632	.2672	.0625	.0212
		5	.0456	.2532	.0296	.0072	.0370	.2545	.0280	.0084
		10	.0214	.2510	.0127	.0049	.0181	.2531	.0129	.0070
		15	.0148	.2466	.0081	.0005	.0122	.2506	.0080	.0046
.5	.25	2	.0637	.3501	.1764	.1040	.0661	.3531	.1807	.1070
		3	.0478	.3023	.0872	.0562	.0304	.3179	.0636	.0718
		5	.0243	.2851	.0579	.0390	.0183	.2794	.0248	.0333
		10	.0090	.2682	.0432	.0221	.0090	.2614	.0089	.0153
		15	.0063	.2615	.0394	.0153	.0058	.2566	.0052	.0105

TABLE 11--Continued

Min Time	Range	N	PA Sampling Variance of $\bar{x}$	PA Sampling $\bar{S}^2$	PA Sampling Variance of S^2	PA Sampling Bias	PR Sampling Variance of $\bar{x}$	PR Sampling $\bar{S}^2$	PR Sampling Variance of S^2	PR Sampling Bias
		2	.1210	.2404	.1145	-.0057	.1248	.2474	.1181	.0013
		3	.0607	.2771	.0696	.0310	.0635	.2746	.0684	.0285
.5	.5	5	.0340	.2624	.0322	.0164	.0326	.2683	.0316	.0222
		10	.0172	.2527	.0134	.0066	.0151	.2560	.0124	.0099
		15	.0108	.2488	.0083	.0028	.0092	.2518	.0079	.0057
		2	.1379	.2056	.0831	-.0405	.1433	.2104	.0920	-.0357
		3	.0774	.2512	.0601	.0051	.1001	.2263	.0530	-.0198
.5	.75	5	.0466	.2512	.0326	.0051	.0604	.2346	.0280	-.0115
		10	.0241	.2440	.0157	-.0021	.0306	.2383	.0126	-.0078
		15	.0153	.2462	.0117	.0001	.0202	.2411	.0079	-.0050
		2	.1730	.1316	.0243	-.1145	.1860	.1289	.0241	-.1172
		3	.0800	.2409	.0478	-.0052	.1321	.1661	.0178	-.0800
.75	.25	5	.0226	.2768	.0250	.0307	.0573	.2364	.0217	-.0097
		10	.0135	.2571	.0097	.0110	.0169	.2535	.0079	.0074
		15	.0076	.2540	.0064	.0079	.0089	.2520	.0051	.0059
		2	.1865	.1345	.0239	-.1116	.1832	.1263	.0231	-.1197
		3	.1070	.2087	.0476	-.0374	.1510	.1433	.0147	-.1028
.75	.5	5	.0355	.2631	.0328	.0170	.1157	.1609	.0121	-.0852
		10	.0194	.2515	.0135	.0053	.0718	.1929	.0094	-.0532
		15	.0128	.2514	.0085	.0053	.0515	.2091	.0074	-.0370
		2	.1407	.1992	.0856	-.0469	.1390	.2021	.0877	-.0440
		3	.0981	.2180	.0504	-.0281	.1008	.2125	.0449	-.0336
.75	.75	5	.0475	.2430	.0287	-.0031	.0504	.2310	.0194	-.0151
		10	.0237	.2446	.0125	-.0015	.0294	.2400	.0116	-.0061
		15	.0157	.2442	.0081	-.0019	.0204	.2423	.0075	-.0037

that although the mean of each of the three sampled functions - FI (Sin 2 PIT), FII (TRIANGLE), and FIII (Sin^5 2 PIT) - is zero, their variances are 1/2, 1/3, and .246093 respectively. These values were computed by evaluating the integral of the function squared over one period. It is these values which are used in the computation of the estimate of bias. Since these differences in variance among the function will be reflected in the distributions of the biases, across function comparisons of bias will also be affected by these differences. If a statistic r defined as the estimate of the bias divided by the variance of the function sampled is used for comparisons rather than the bias itself, different results are obtained. Clearly, if r is used for within function comparisons, the variance acts merely as a scale factor, but in across function comparisons a completely different impression may be given by a comparison of r's than by a comparison of the unweighted biases. Consider the PR sampling with minimum time .25 and range .25 for FI and FII. The size 5 estimate of bias obtained for FI was .1071 while that for FII was .0713. The corresponding r for FI, however, was .2142 while the associate r for FII was .2139. For some purposes the use of r would seem more appropriate. Defined below is a statistic, the critical size, based upon the use of r which proved very helpful in giving some specificity to the rather vague general impression of the relationship of bias and sample size gleaned from studying the summary statistics of each study.

For a given function and a given positive number p, the critical size N_p is the smallest positive integer $n \leq 25$ such that:

(i) If B_n is the absolute value of the bias estimate obtained for sample size n and V the variance of the given function, then

$(B_n / V) \leq p.$

(ii) If $25 \geq m > n$ then $(B_m / V) \leq p.$

If no such integer exist, then N_p is said not to exist.

N_p is, then, an estimate of the smallest sample size such that, for this sample size or any larger size, the absolute value of the bias expressed as a fractional portion of the variance of the function does not exceed a given amount p. Of course, N_p proved useful only because of the general downward trend noted above. Table 12 gives critical size information for the functions sampled.

Although in the discussions of the PA and PR sampling in CHAPTER III the view of these plans as having a fixed scheduled time between sampling times delayed, perhaps, by some positive time was adopted, an equally valid interpretation is given by regarding the scheduled time between samples as the minimum time plus the expected value of the delay distribution with the possibility of a sample being taken early (a delay of less than the expected value) or late (a delay of greater than the expected value). If this latter view is taken, then the sampling plans with a minimum time .25 and range .75 are similar to those with minimum time .5 and range .25 in the sense that the expected time between samples is 5/8. The same relationship exists among those plans with minimum time .5, range .75 and those with minimum time .75, range .25, since the expected time between samples is 7/8. Some insight into the influence of the range may be gained by comparing the results for these similar plans. Table 12 shows that within each function the corresponding N_p is less for the associated similar plan of the same type with the larger range. Thus, it would seem that increasing the range tends to accelerate the downward

TABLE 12

CRITICAL SIZE VALUES FOR P = .1, .05, .03

Min Time	Range	FI (Sin 2 PIT) PA $N_{.10}$	$N_{.05}$	$N_{.03}$	PR $N_{.10}$	$N_{.05}$	$N_{.03}$	FII (TRIANGLE) PA $N_{.10}$	$N_{.05}$	$N_{.03}$	PR $N_{.10}$	$N_{.05}$	$N_{.03}$	FIII (Sin^5 2 PIT) PA $N_{.10}$	$N_{.05}$	$N_{.03}$	PR $N_{.10}$	$N_{.05}$	$N_{.03}$
	.25	9	17	-	10	20	-	9	16	24	10	17	-	7	12	18	6	13	22
.25	.5	6	12	19	8	16	-	4	8	16	8	14	23	4	7	11	7	12	24
	.75	3	3	3	6	10	18	3	3	5	4	8	11	3	3	3	3	5	6
	.25	8	14	18	10	20	-	7	11	13	10	21	-	9	19	-	7	13	21
.5	.5	4	8	16	6	10	19	6	10	16	6	9	16	4	8	10	5	9	13
	.75	3	4	4	4	8	17	3	4	8	4	8	12	3	3	3	3	5	11
	.25	6	17	-	5	9	18	9	16	25	5	6	10	6	13	18	5	9	10
.75	.5	4	8	14	-	-	-	4	6	13	-	-	-	4	6	14	24	-	-
	.75	4	4	4	5	8	12	4	4	4	5	7	14	4	4	5	4	7	9

- Indicates no critical size exist.

trend of the magnitude of the bias for a fixed "scheduled" time between samples. Examination of the tables of summary statistics reinforces this observation.

As a further check upon the effect of increase of range and to observe the results of sampling with a plan having a delay range greater than one, an additional study was conducted for FI under PR with minimum time .5 and range 1.25. The expected time for this plan is 9/8, the same as for a minimum time of .75 and range of .75. An abbreviated table of the summary statistics for this study is given in Table 13. The $N_{.10}$, $N_{.05}$, and $N_{.03}$ for this study were 3, 6, and 7 respectively. From Table 12 the corresponding N_p's for the similar PR sampling of FI are 5, 8, and 12. This additional study, therefore, offers further evidence in support of the hypothesis of larger range values accelerating the downward trend of the bias.

It was shown in CHAPTER III that under PA and PR sampling the expected value of the within subject mean is an unbiased estimator of the function mean. Hence, the mean (of a sample of size 1500) of the within subject running means for each sample size is an unbiased estimator of this expected value. Since the mean of all functions was zero, the expected value of this statistic at each sample size was zero. The results of the computer sampling were in agreement with this theoretical result. The maximum deviation from zero was .02456 obtained for a sample size 1 of FI, however most of the magnitudes of the deviations from zero were much smaller than this. Over all studies, the number of positive deviations was approximately equal to the number of negative deviations as would be expected from the symmetry.

TABLE 13

SUMMARY STATISTICS FOR PR SAMPLING OF FI WITH MINIMUM TIME .5 AND RANGE 1.25

N	Variance of $\bar{g}$	$\bar{S}^2$	Variance of S^2	Bias	Bias/V
2	.2200	.5736	.3458	.0736	.147
3	.1414	.5451	.1210	.0451	.090
5	.0789	.5290	.0464	.0290	.058
10	.0400	.5079	.0166	.0079	.015
15	.0272	.5039	.0100	.0039	.007

The variance $V_{\bar{g}}$ of the distribution of running means was shown in CHAPTER III to be related to the bias by

$$V_{\bar{g}} = \sigma^2/n - (1 - 1/n)B$$

where n is the sample size, σ^2 the variance of the function, and B the bias. Examination of the tables of summary statistics shows that the estimate of the variance of $\bar{g}$ is in agreement with this result even when the estimate of bias is used rather than the exact population figure. Note that the relationship of bias and $V_{\bar{g}}$ is such that if the bias is positive, $V_{\bar{g}}$ is less than the corresponding variance for a random sampling, but if the bias is negative $V_{\bar{g}}$ is greater than that for random sampling. The effect of the bias on $V_{\bar{g}}$ is such that $V_{\bar{g}}$ does not necessarily decrease monotonically with increasing sample size as in random sampling, but a sharp downward trend is found in all studies.

Frequency Distributions

For the frequency distributions of the within subject running means at each sample size, the interval -1 to +1 was partitioned into 13 equal length intervals and a computer program used to tabulate, for each sample size within each study, the number of means which fell into each interval. Since the means for the samples at size one are random samples from the functions over one period, histograms or frequency polygons constructed from the frequency distribution output would be expected to be the density functions derived in CHAPTER III for the sampled functions. This was, indeed, the case for each study. From examination of the frequency distributions obtained it would seem that within each study the densities of within subject running means change progressively with

increasing sample size from the original densities at sample size one to symmetric, unimodal densities with mean zero. The densities associated with the larger sample sizes seem to become very "peaked" about zero. The rapid (in terms of sample size) convergence from the original frequency distributions at sample size one to the unimodal form was one of the more surprising results of the simulation. The pattern described may be seen in Table 14 which gives abbreviated forms of the frequency distributions obtained from three studies. In addition to the tabular values presented, the pattern is illustrated in the graphs of the frequency polygons for sample sizes from 1 to 10 in Figure 5.

The differences in the variances of the three functions presented a problem in selecting intervals for the frequency distribution of S^2. Clearly, intervals suitable for displaying the distribution for FI, say, are not appropriate for FII and FIII. After preliminary trials with several different selections, the decision was made to use 13 intervals for each function, but with different interval widths for each function. Thus, within a given function, frequency counts for intervals are directly comparable, but for across function comparisons the differences in width must be noted. The definitions of the sets of intervals for each function are given below.

For FI the interval from zero to 1.274 was partitioned into 12 equal length intervals and these were used to define the first 12 intervals for the frequency distribution. All values greater than 1.274 were tabulated in interval 13. For FII the interval partitioned was from zero to 1.140 and for FIII from zero to .840. For these functions, as for FI, all values not in any of the first 12 intervals were tabulated in

TABLE 14

FREQUENCY DISTRIBUTIONS OF WITHIN SUBJECT RUNNING MEANS FOR THREE STUDIES

Function	Min Time	Range	Type Sampling	Size	Interval Number 1	2	3	4	5	6	7	8	9	10	11	12	13
				1	296	107	82	77	77	71	66	79	78	75	96	125	271
				2	153	164	121	106	71	81	78	87	84	95	134	168	158
FI	.75	.25	PR	3	58	116	155	120	141	103	91	106	111	146	150	136	67
				4	22	70	110	143	151	158	152	161	149	158	134	71	21
				10	0	1	1	11	155	337	439	384	155	16	0	0	1
				20	0	0	0	1	40	337	727	361	34	0	0	0	0
				1	113	121	93	113	127	117	121	118	112	125	118	108	114
				2	15	29	51	111	187	237	233	201	201	110	71	40	14
FII	.25	.75	PA	3	6	21	48	117	166	226	273	224	194	129	70	22	4
				4	3	11	48	79	187	252	296	275	176	116	41	16	0
				10	0	0	0	17	124	338	493	397	114	16	1	0	0
				20	0	0	0	1	43	373	685	349	49	0	0	0	0
				1	120	44	49	54	54	107	607	96	60	55	63	69	122
				2	11	21	35	145	137	175	415	157	137	188	37	27	15
FIII	.5	.5	PA	3	3	2	15	37	168	276	427	266	200	67	37	2	0
				4	0	1	8	41	141	302	448	318	165	57	17	2	0
				10	0	0	0	2	52	330	690	356	67	3	0	0	0
				20	0	0	0	1	7	247	937	299	9	0	0	0	0

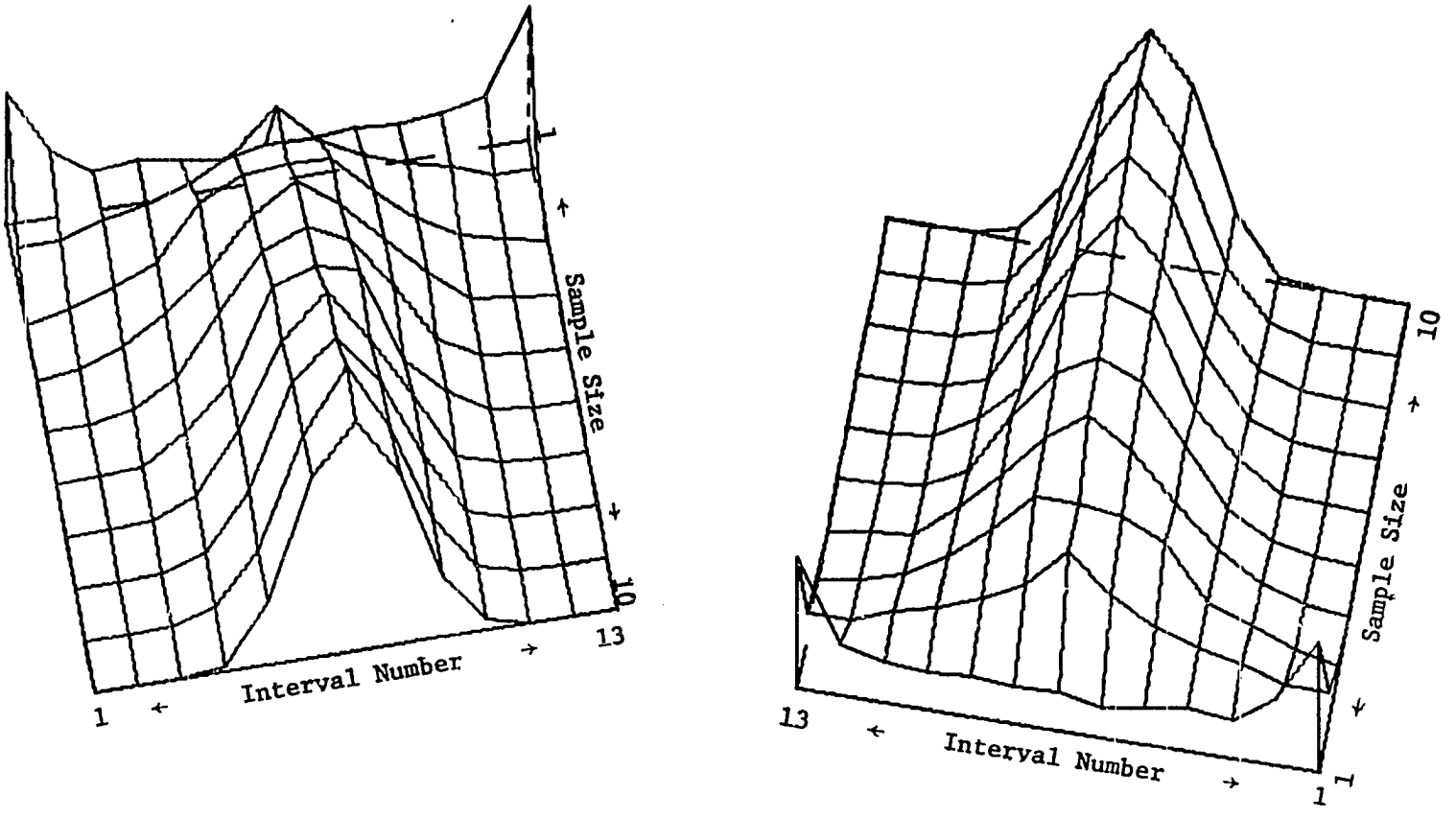


Figure 5--Frequency polygons of within subject running means for FI sampled under PR with minimum time .25 and range .75.

interval 13.

Using these sets of intervals, the variance of FI is a number in interval number 5 while the variances of FII and FIII lie in interval number 4.

With the exception of the studies using a PA sampling with minimum time .5 and range .25, which will be discussed below, the frequency distributions within each study followed a pattern with increasing sample size similar to that for the frequency distributions of the running means in that, as sample sizes become large, the distributions become unimodal with a "peaking" at the interval containing the population value being estimated as the bias becomes small. The symmetry noted for the means, however, was not a characteristic of these distributions. The general pattern of change with increasing sample size observed in each study may be described as progressing from a frequency distribution with excessive counts in the extreme intervals for sample size 2, to a more nearly uniform distribution for size 3 or 4, thence to the unimodal form mentioned above. This pattern may be observed in the three sets of frequency distributions selected as typical and presented in Table 15 and in the graphical presentation of the frequency polygons in Figure 6.

In the PR sampling with minimum time .75 and range .5 the change to a unimodal form of the frequency distribution did not occur until a much larger sample size than for the other studies. Also, the influence of the relatively large negative bias is seen in that even for sample sizes where the distribution may be considered unimodal, the interval having the most counts is not the interval containing the variance of the function sampled until the sample size is quite large. The sets of frequency distributions obtained under this plan given in Table 16 illustrate

TABLE 15

FREQUENCY DISTRIBUTIONS OF S^2 FOR THREE STUDIES

Function	Min Time	Range	Type Sampling	Size	Interval Number 1	2	3	4	5	6	7	8	9	10	11	12	13
				2	248	100	83	73	62	77	68	56	87	67	73	71	435
				3	15	39	58	95	149	211	219	240	223	139	65	38	9
FI	.25	.25	PA	5	1	19	61	157	317	444	272	145	58	17	9	0	0
				10	0	5	27	153	471	576	223	41	4	0	0	0	0
				15	0	0	10	119	611	648	106	6	0	0	0	0	0
				25	0	0	2	85	828	557	27	1	0	0	0	0	0
				2	685	198	156	113	84	51	27	33	29	17	23	16	68
				3	395	288	233	149	122	78	76	64	46	27	12	7	3
FII	.5	.75	PR	5	140	279	309	266	249	147	74	26	8	1	1	0	0
				10	8	138	424	495	311	109	15	0	0	0	0	0	0
				15	0	64	390	679	312	54	1	0	0	0	0	0	0
				25	0	8	342	857	282	11	0	0	0	0	0	0	0
				2	602	132	96	96	82	103	133	74	28	21	15	8	110
				3	382	172	160	202	209	85	51	44	39	29	32	19	76
FIII	.25	.75	PR	5	197	197	318	229	162	126	119	73	44	18	9	4	4
				10	50	208	300	347	300	169	85	33	6	1	1	0	0
				15	22	129	354	451	338	154	40	10	1	1	0	0	0
				25	2	61	360	616	364	82	13	2	0	0	0	0	0

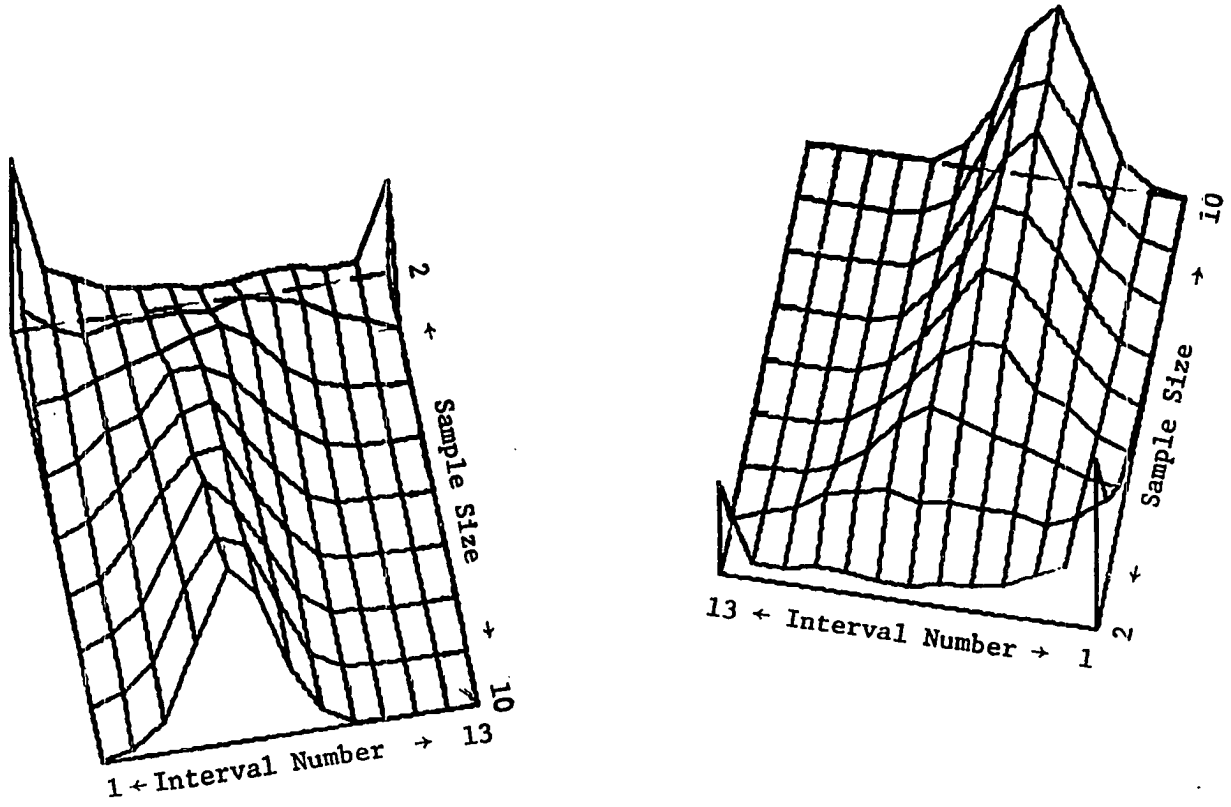


Figure 6--Frequency polygons of within subject variance estimates for FI sampled under PR with minimum time .5 and range 1.25.

TABLE 16

FREQUENCY DISTRIBUTIONS OF S^2 FOR PR SAMPLING
WITH MINIMUM TIME .75 AND RANGE .5

Function	Min Time	Range	Type Sampling	Size	Interval Number 1	2	3	4	5	6	7	8	9	10	11	12	13
				2	830	207	142	102	66	60	42	36	14	1	0	0	0
				3	631	268	200	149	106	50	40	27	27	2	0	0	0
FI	.75	.5	PR	5	376	297	267	213	145	111	63	23	5	0	0	0	0
				10	129	212	267	314	301	215	54	8	0	0	0	0	0
				15	62	124	234	365	449	220	42	4	0	0	0	0	0
				25	10	43	147	409	629	251	11	0	0	0	0	0	0
				2	813	288	107	114	94	21	0	0	0	0	0	0	0
				3	652	390	250	50	30	20	6	5	2	0	0	0	0
FII	.75	.5	PR	5	423	488	285	154	81	48	18	3	0	0	0	0	0
				10	149	426	396	329	159	33	7	1	0	0	0	0	0
				15	59	292	469	489	107	21	0	0	0	0	0	0	0
				25	14	131	536	645	171	3	0	0	0	0	0	0	0
				2	805	203	103	105	79	87	94	24	0	0	0	0	0
				3	528	278	229	276	147	19	7	5	2	5	1	0	3
FIII	.75	.5	PR	5	300	370	479	193	65	43	25	12	7	5	0	1	0
				10	82	391	516	228	168	77	26	10	1	1	0	0	0
				15	31	303	520	344	194	83	21	2	1	1	0	0	0
				25	2	137	527	560	222	49	2	1	0	0	0	0	0

these remarks.

Exceptions to the general pattern were obtained in PA sampling with minimum time .5 and range .25. Under this plan the frequency distributions remained bimodal even for the larger sample sizes with one "mode" lying to the left and one to the right of the variance of the sampled function. This characteristic may be seen in the sets of distributions presented in Table 17.

Four Group Sampling

As described in CHAPTER IV, four groups of 20 subjects each were randomly selected from each study and the group mean and $\bar{s}^2$ plotted against sample size. Each graph for each group, then, represents a plot of the group characteristic as might be obtained in a real investigation where the variable being measured has a pattern over time similar to the sampled function.

Two of the graphs of the group means of the within subject running means are reproduced in Figures 7 and 8. The former is from a study with a positive bias and the latter from a study with a negative bias. At each sample size the four plotted points represent four means based upon a random sampling of size 20 from the distribution associated with that sample size, function, and sampling plan. The spread of the four points, then, reflects the magnitude of the variance. This remark applies to both group means and $\bar{s}^2$ graphs. In Figure 7 the spread among the group means decreases rapidly as sample size increases. This is a reflection of the reduction in variance of the distribution when positive bias is present. The group means are relatively widespread until the

TABLE 17

FREQUENCY DISTRIBUTIONS OF S^2 FOR PA SAMPLING
WITH MINIMUM TIME .5 AND RANGE .25

Function	Min Time	Range	Type Sampling	Size	Interval Number 1	2	3	4	5	6	7	8	9	10	11	12	13
				2	265	99	78	79	82	65	71	60	66	73	74	66	422
				3	122	116	124	134	138	125	107	127	143	173	114	59	18
FI	.5	.25	PA	5	50	152	175	155	156	156	137	170	188	126	35	0	0
				10	12	144	285	154	148	142	161	138	194	22	0	0	0
				15	2	152	287	186	135	169	178	261	128	2	0	0	0
				25	1	174	307	144	152	171	185	300	66	0	0	0	0
				2	336	130	116	94	83	106	123	87	83	64	56	58	164
				3	200	210	208	147	142	140	152	107	88	51	35	15	5
FII	.5	.25	PA	5	140	313	170	166	162	176	188	110	51	21	3	0	0
				10	122	349	187	151	187	219	207	69	9	0	0	0	0
				15	123	377	183	139	192	246	208	32	0	0	0	0	0
				25	137	364	179	156	190	254	212	8	0	0	0	0	0
				2	542	100	92	87	82	101	102	89	44	32	32	28	168
				3	456	111	122	166	140	95	80	54	35	37	49	38	117
FIII	.5	.25	PA	5	412	128	152	130	108	117	115	104	89	60	32	26	27
				10	416	100	110	102	140	172	167	107	92	33	18	3	0
				15	417	143	102	116	136	163	179	139	59	26	2	0	0
				25	420	142	116	115	112	177	219	149	47	8	0	0	0

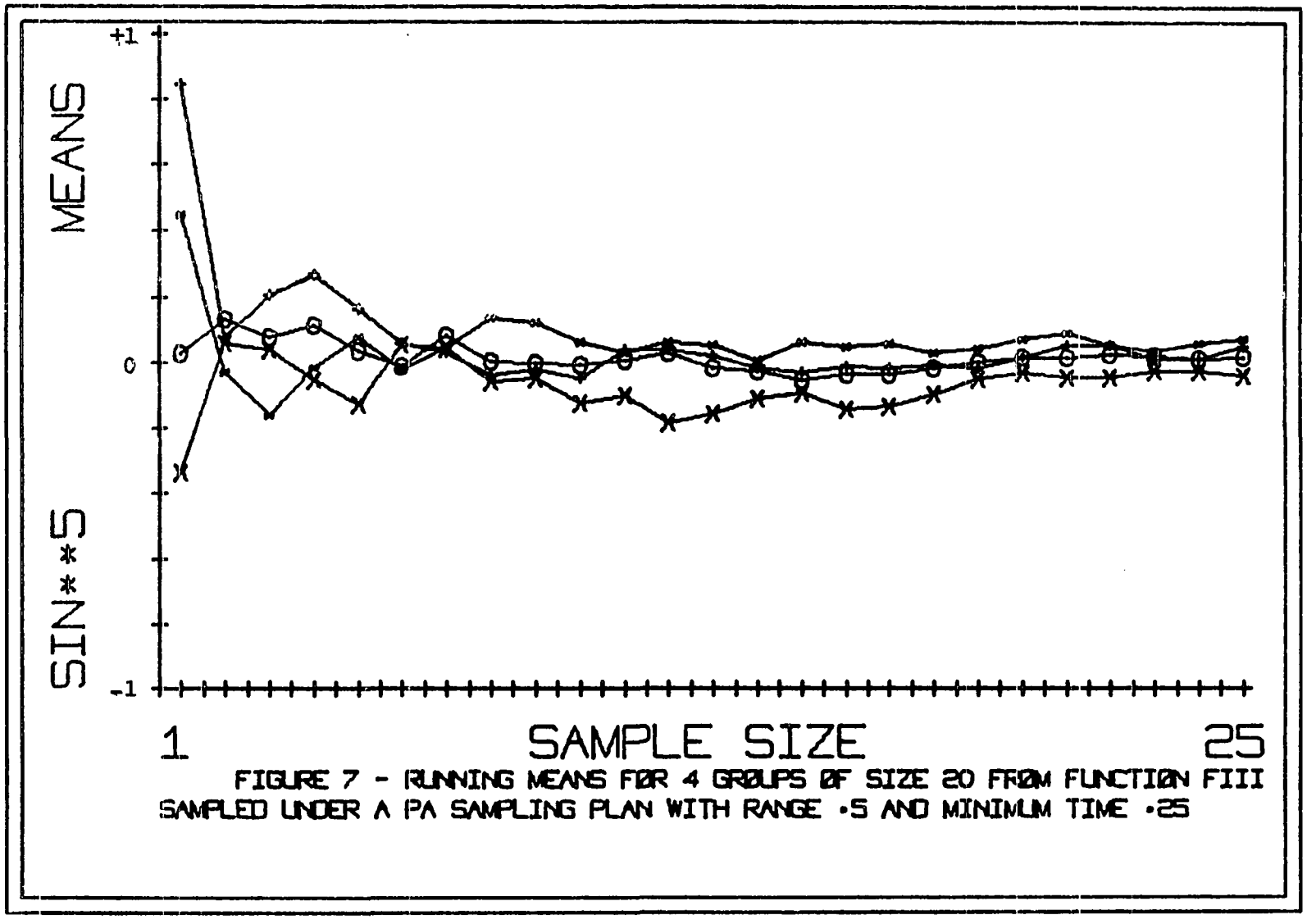


FIGURE 7 - RUNNING MEANS FOR 4 GROUPS OF SIZE 20 FROM FUNCTION FIII SAMPLED UNDER A PA SAMPLING PLAN WITH RANGE .5 AND MINIMUM TIME .25

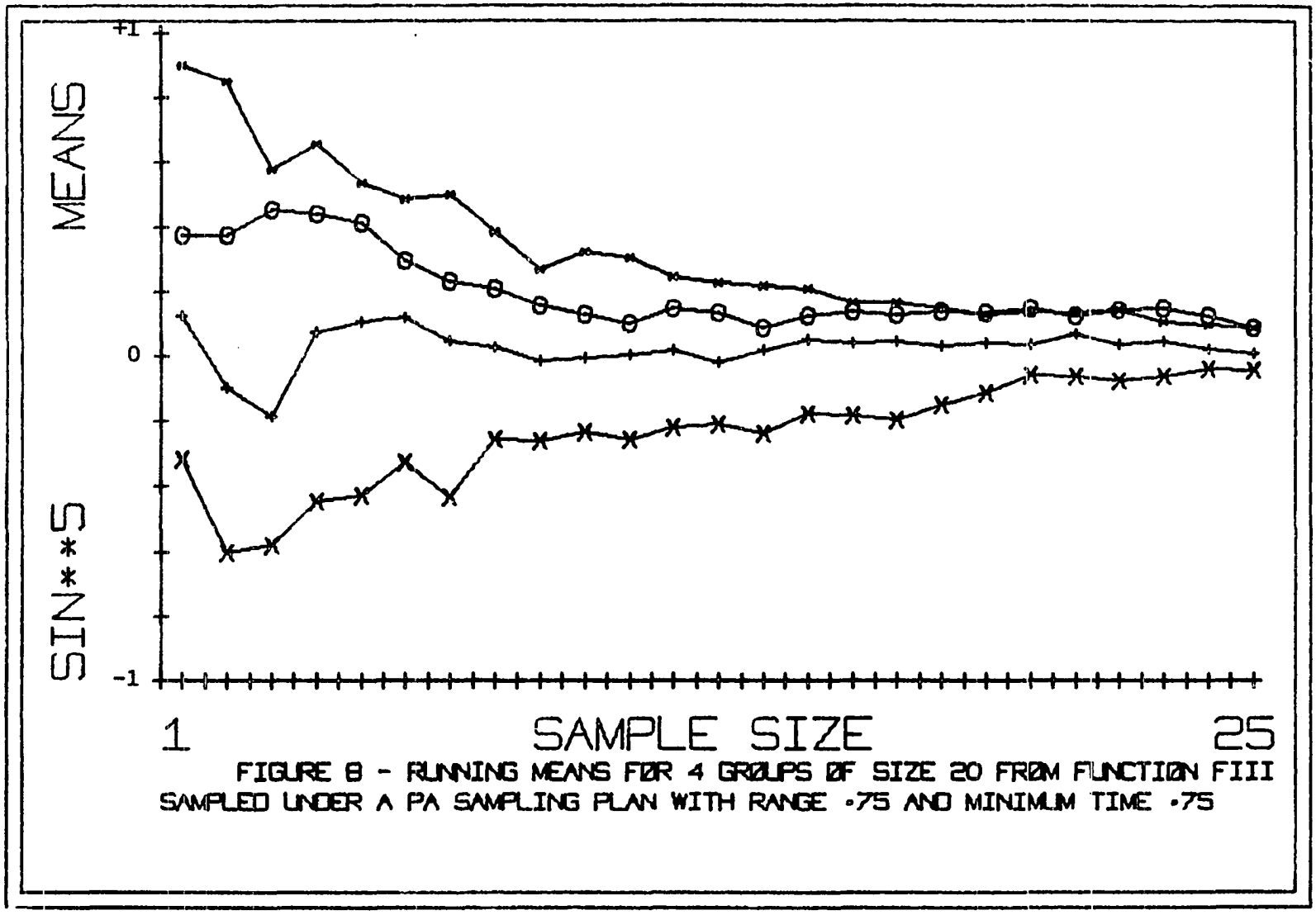


FIGURE 8 - RUNNING MEANS FOR 4 GROUPS OF SIZE 20 FROM FUNCTION FIII SAMPLED UNDER A PA SAMPLING PLAN WITH RANGE .75 AND MINIMUM TIME .75

larger sample sizes in the study with negative bias depicted in Figure 8. Note that in both figures the spread is around zero, the mean of the sampled function, and that the means of each group lie close to this value as sample size becomes large. These results were to be expected from the unbiasedness of the estimator and the increase of precision with sample size noted in the discussion of the summary statistics. The patterns observed in these two graphs are typical of those observed for the graphs of group means in all studies.

At each sample size in the $\overline{S^2}$ graphs, the spread of the four plotted points reflects the variance as in the mean graphs, but the expected value is the same at each sample point for the mean graph, while the expected value is a function of sample size for the $\overline{S^2}$ graphs. Thus, the patterns over sample size reflect the pattern of the expected values of $\overline{S^2}$ over sample size, and the distances from the line parallel to the sample size axis at a distance equal to the variance of the sampled function reflect the relationship of sample size and bias.

In Figures 9 through 16 eight sets of $\overline{S^2}$ graphs are given to illustrate how the pattern over sample size exhibited by $\overline{S^2}$ of the four groups recapitulates the findings concerning S^2 in the above discussion of summary statistics and frequency distributions.

In the discussions above it was noted that in PR sampling of each function with minimum time .75 and range .5 a relatively large negative bias was present even for larger sample sizes, while for the PA sampling with the same parameters the magnitude of the bias becomes small for larger sample size. The effect of the difference in the two sampling types for these parameters is illustrated for FI in Figures 9 and 10.

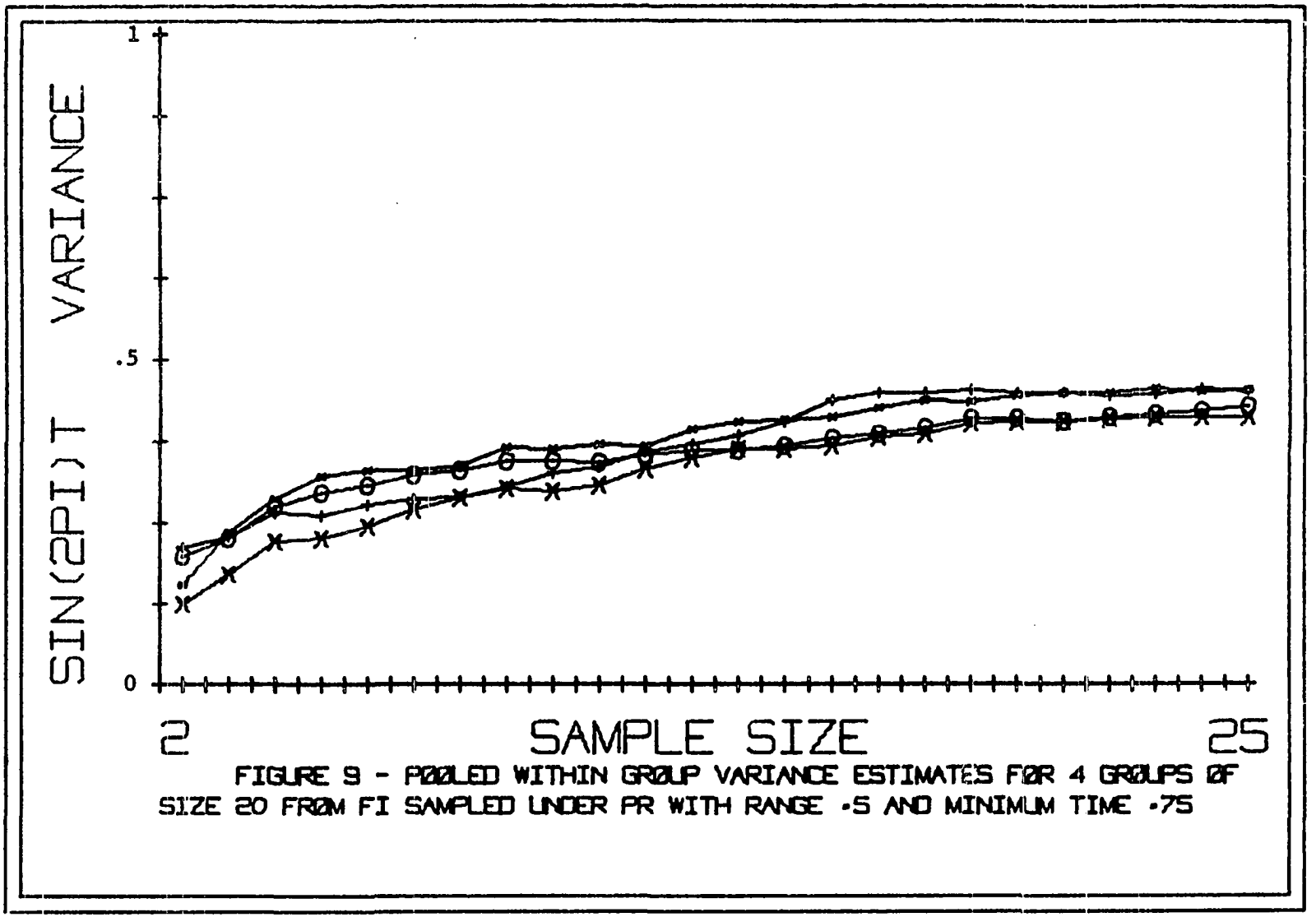


FIGURE 9 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FI SAMPLED UNDER PR WITH RANGE .5 AND MINIMUM TIME .75

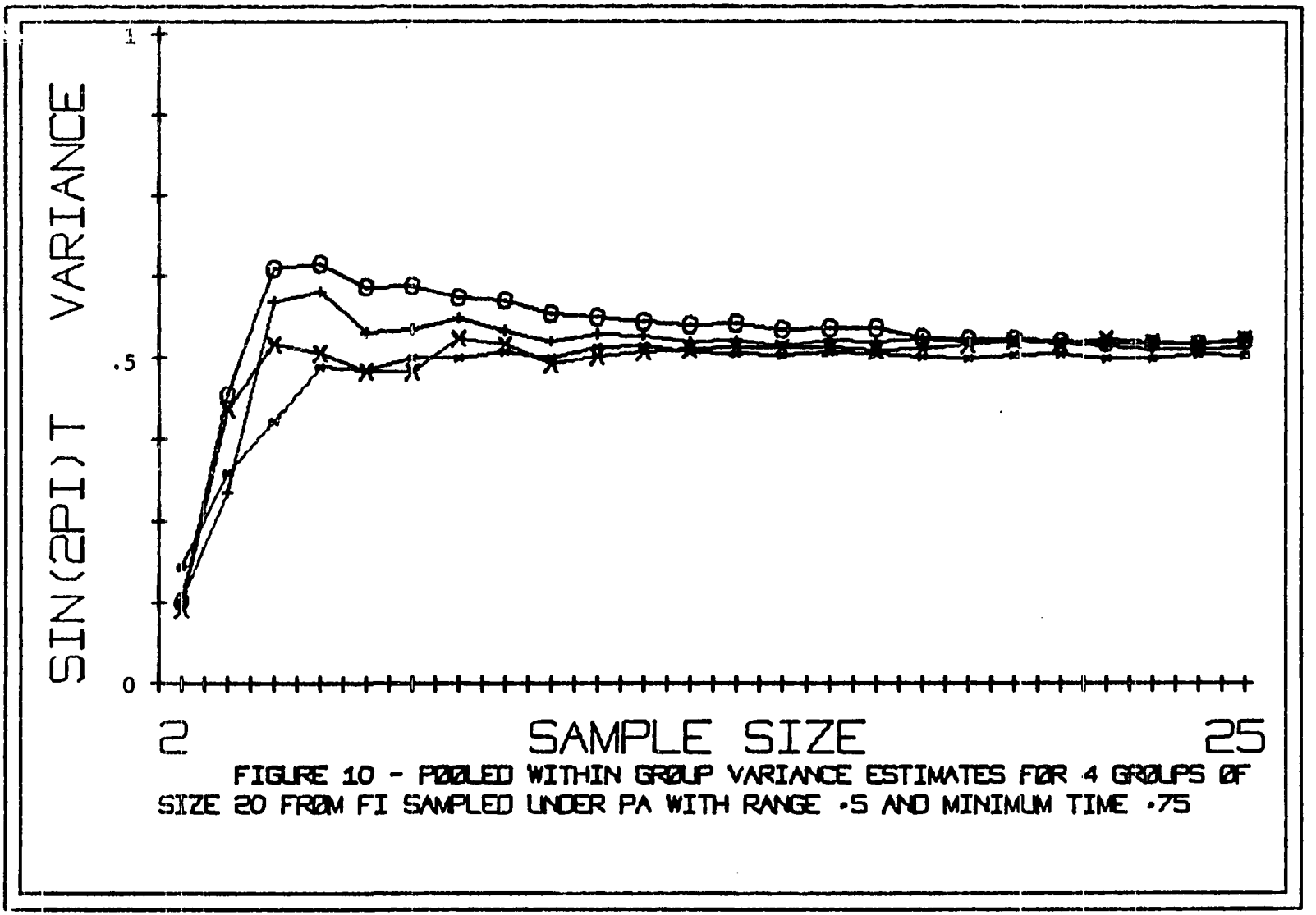


FIGURE 10 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FI SAMPLED UNDER PA WITH RANGE .5 AND MINIMUM TIME .75

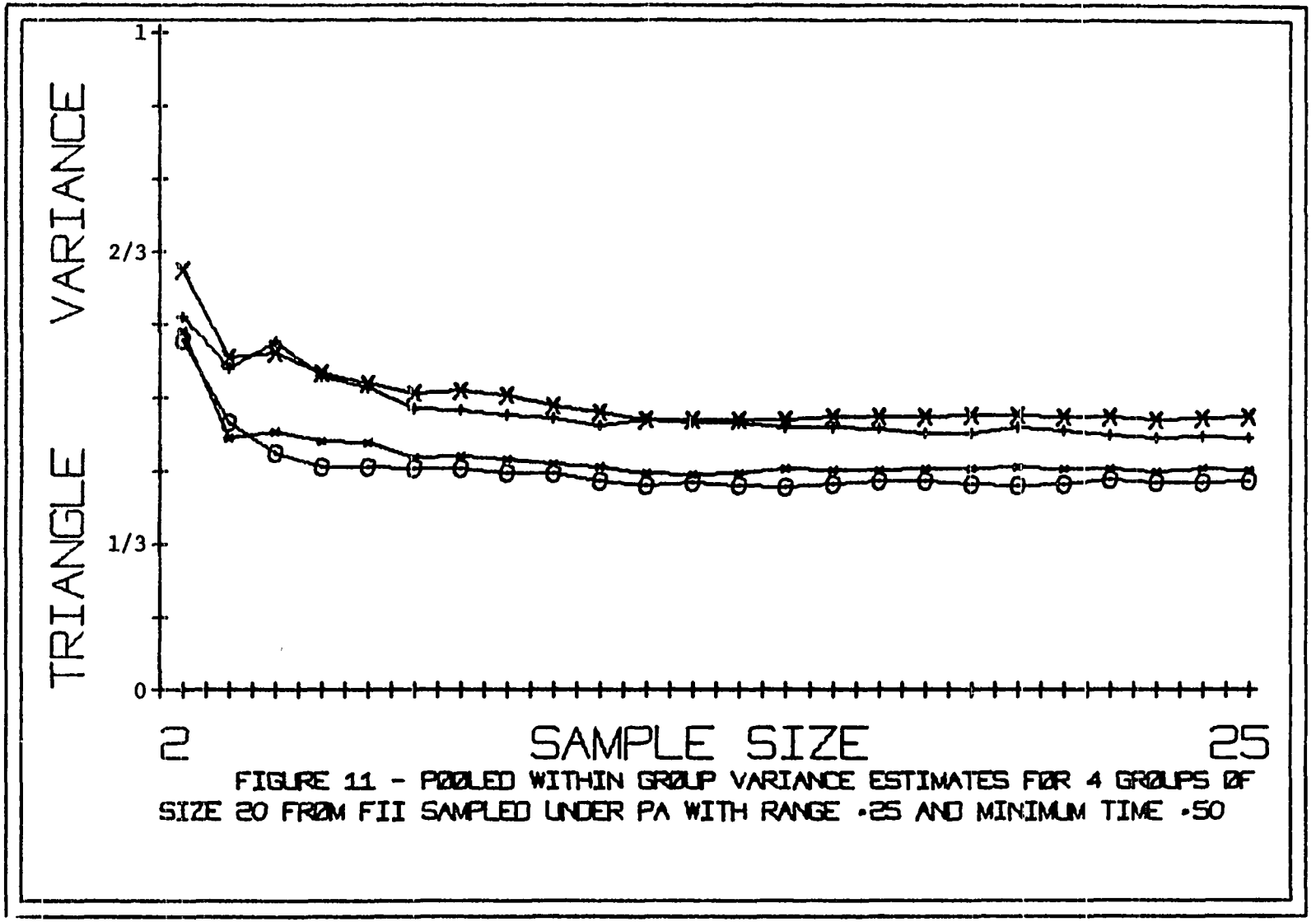


FIGURE 11 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FII SAMPLED UNDER PA WITH RANGE .25 AND MINIMUM TIME .50

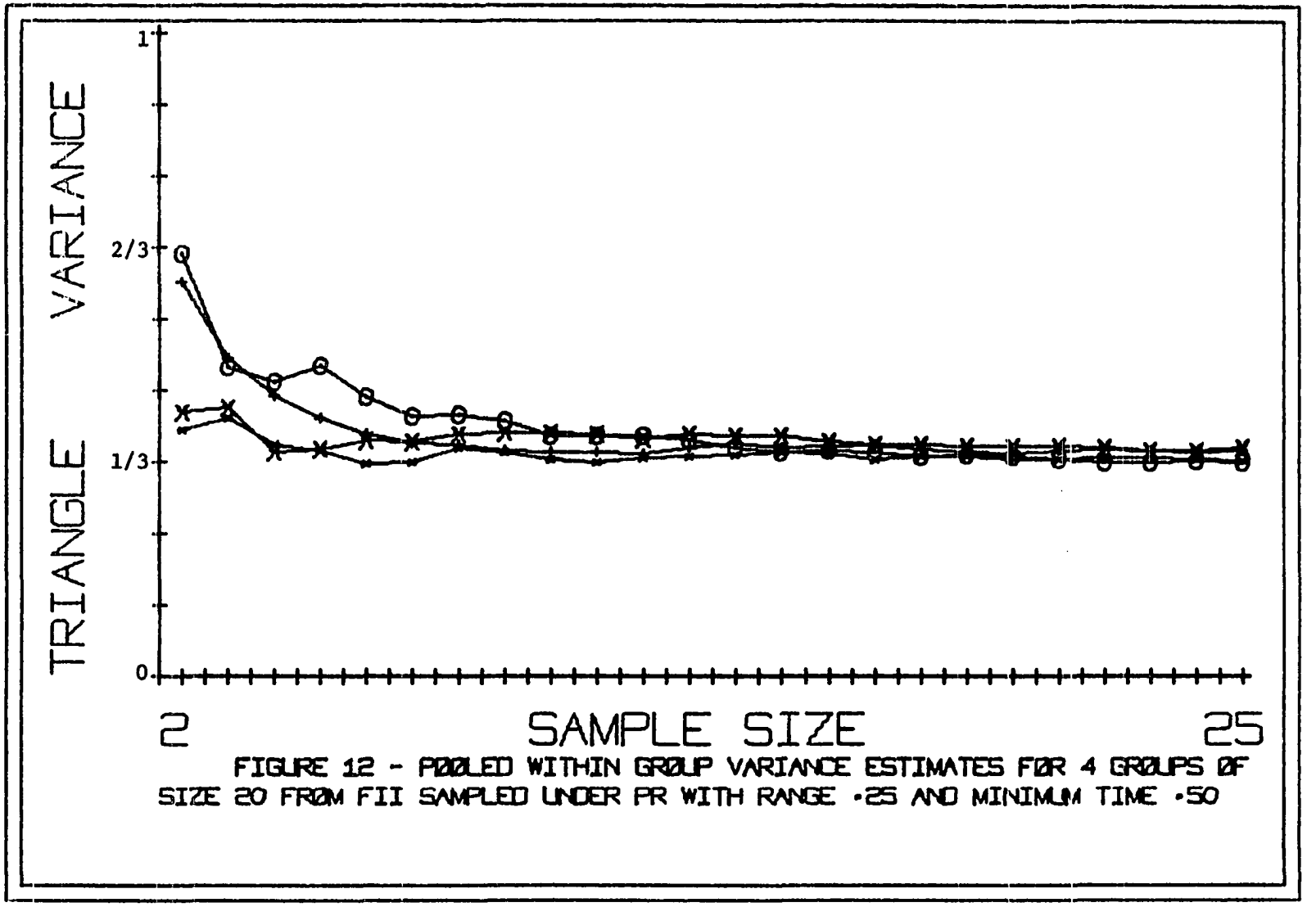


FIGURE 12 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FII SAMPLED UNDER PR WITH RANGE .25 AND MINIMUM TIME .50

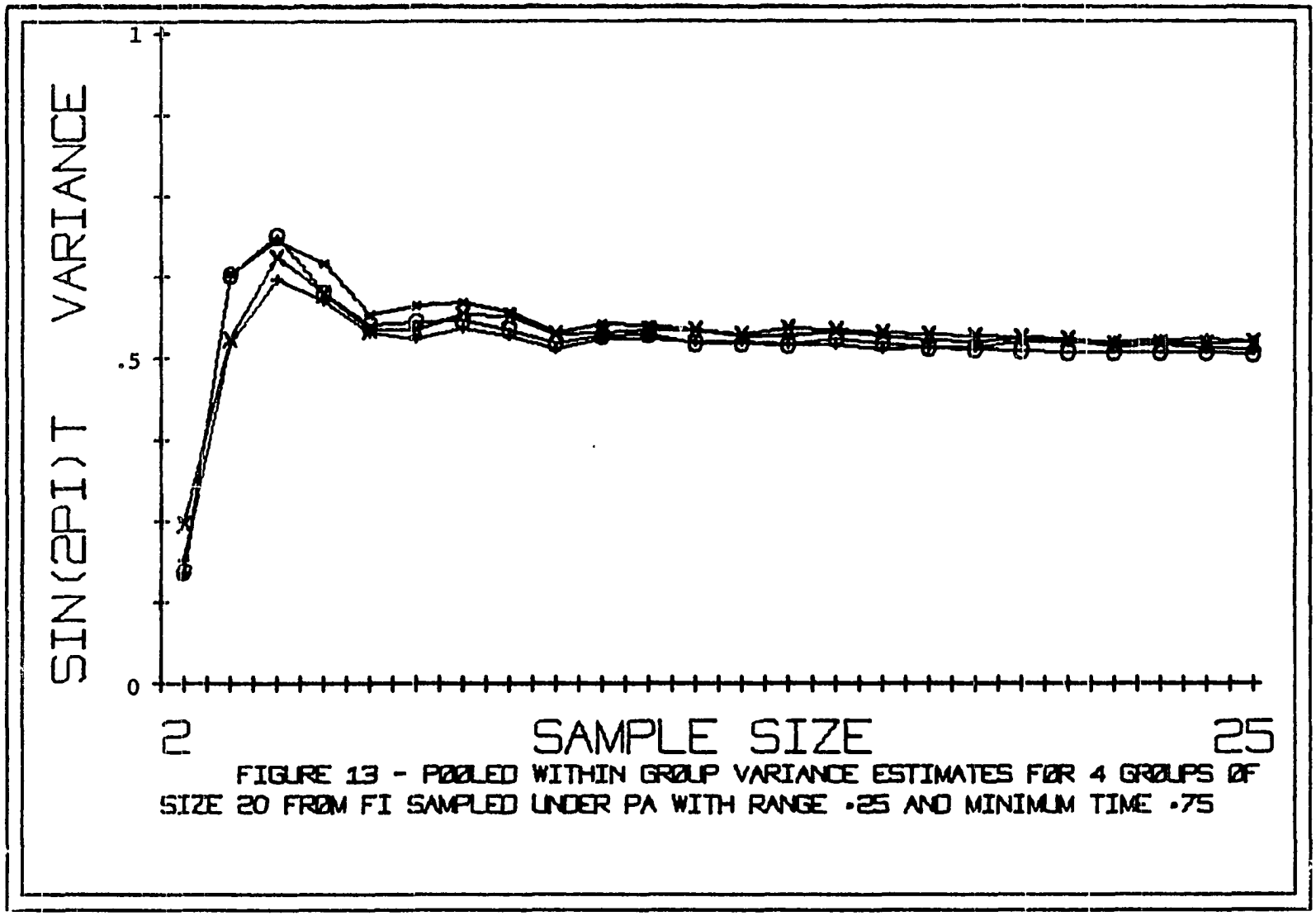


FIGURE 13 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FI SAMPLED UNDER PA WITH RANGE .25 AND MINIMUM TIME .75

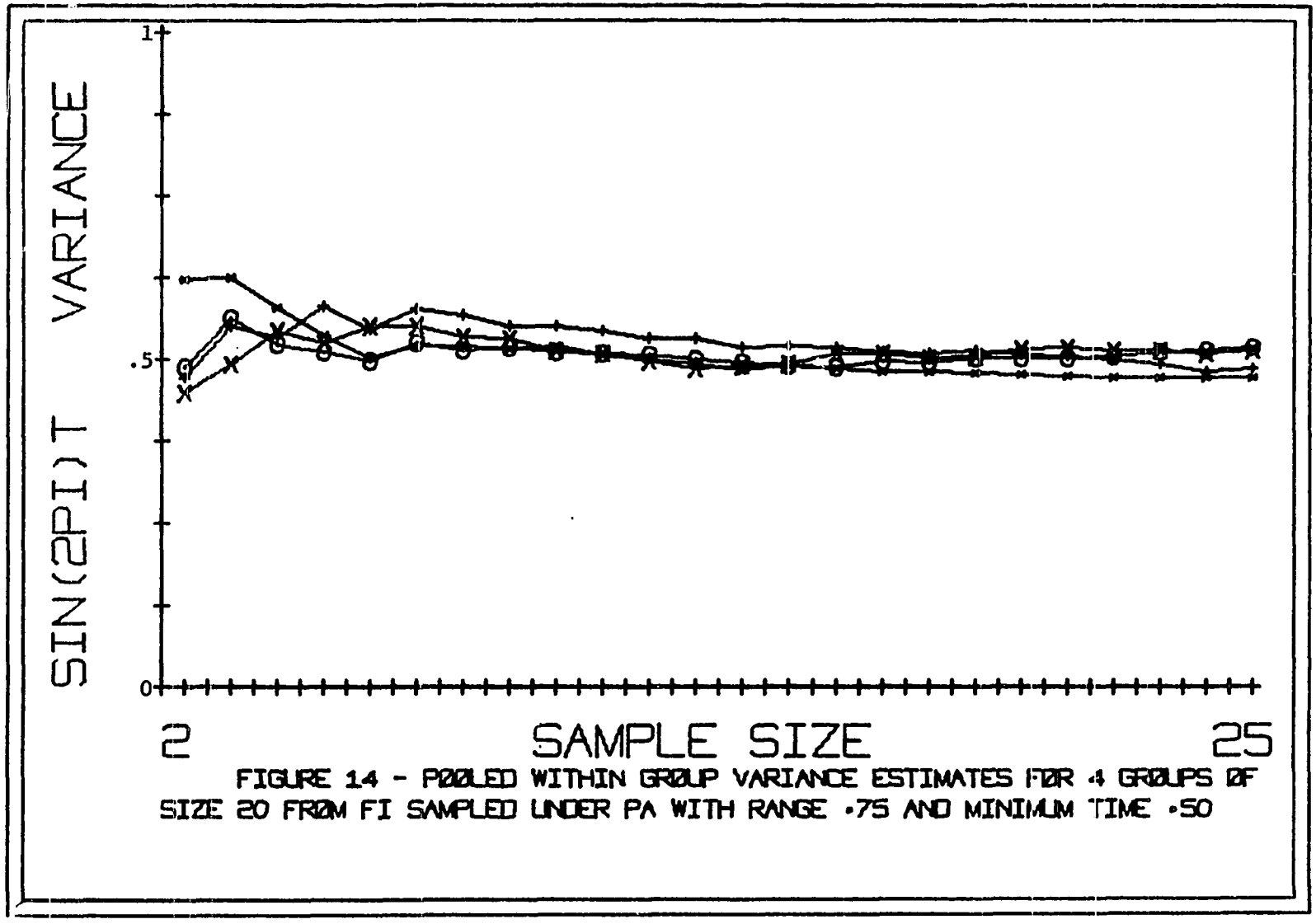


FIGURE 14 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FI SAMPLED UNDER PA WITH RANGE .75 AND MINIMUM TIME .50

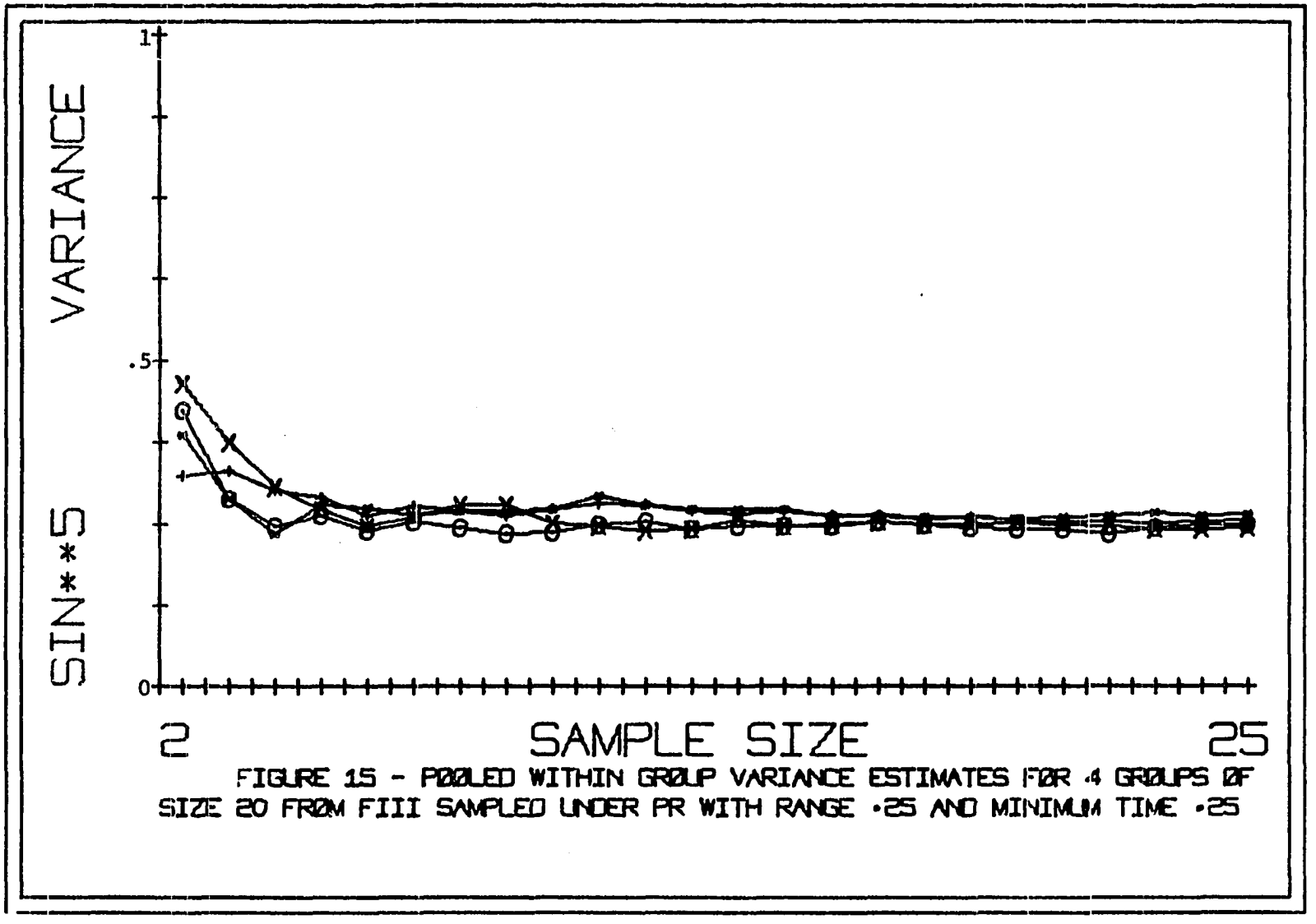


FIGURE 15 - POOLED WITHIN GROUP VARIANCE ESTIMATES FOR 4 GROUPS OF SIZE 20 FROM FIII SAMPLED UNDER PR WITH RANGE .25 AND MINIMUM TIME .25

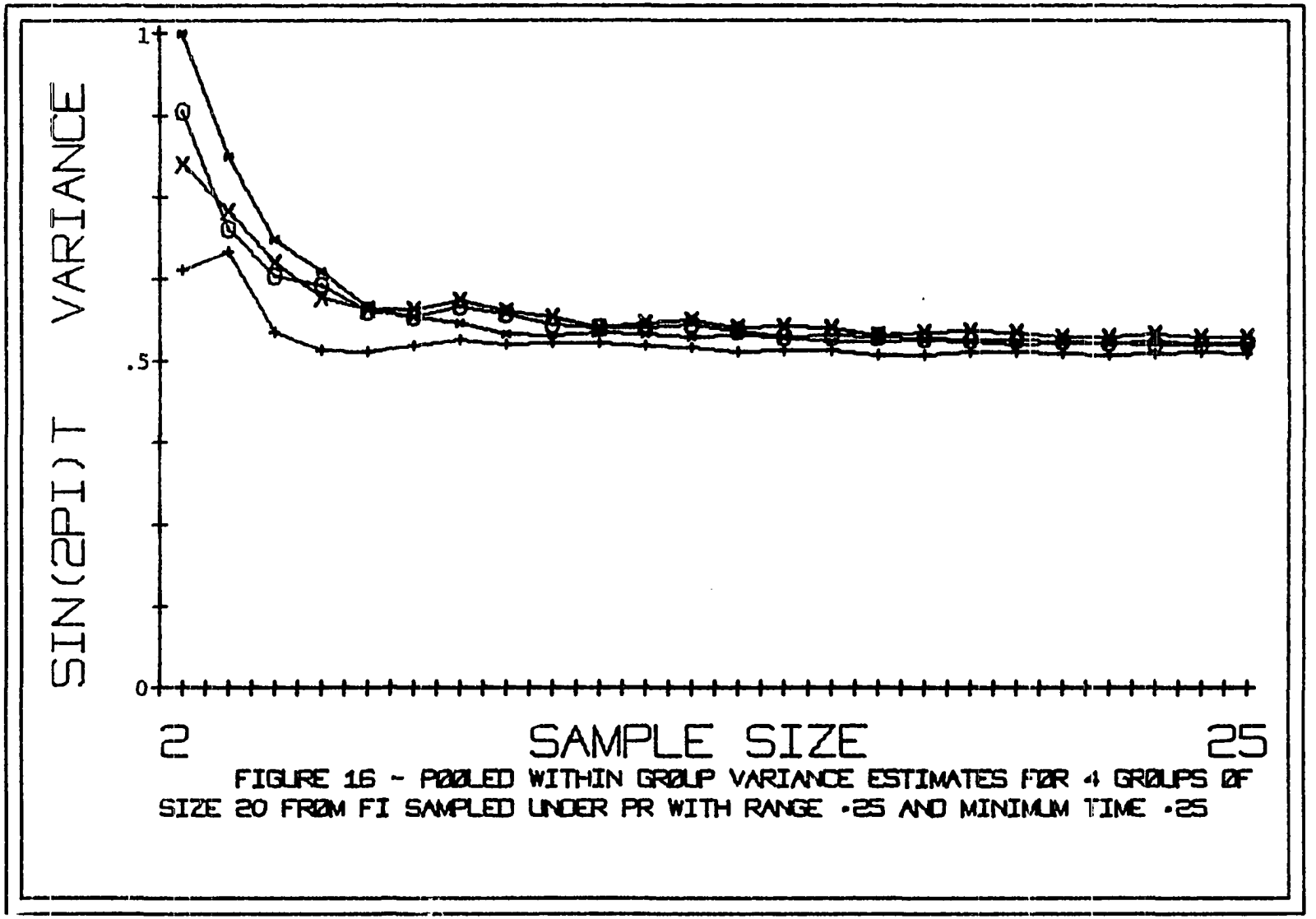


FIGURE 16 ~ P00LED WITHIN GR0UP VARIANCE ESTIMATES F0R 4 GR0UPS 0F SIZE 20 FR0M FI SAMPLED UNDER PR WITH RANGE .25 AND MINIMUM TIME .25

Another sampling plan that was noted to yield results somewhat different from the others was the PA sampling with minimum time .5 and range .25. Figure 11 gives the $\bar{S}^2$ graphs for the four groups from the FII study using this plan. Notice the spread of the points at each sample size remains relatively large even at size 25 whereas the four groups from the PR sampling with the same parameters graphed in Figure 12 are close together for the larger sizes.

Figures 13 and 14 illustrate the greater bias found for the smaller sample sizes for the plan with the smaller range even though the expected time between measurements is the same in both studies.

Figures15 and 16 contrast the pattern for the function with the greatest variance with that of the function having the smallest when both are sampled under the same plan.

Discussion of Results

The findings of the simulation were in agreement with the theoretical results for pseudo-systematic sampling of periodic functions developed in CHAPTER III. The unbiased property of the within subject sample mean as an estimator of the function mean, the biasedness of within subject S^2 as an estimator of function variance, the dependency of bias on sample size, sampling plan, and function sampled, and the relationship of the variance of sample mean and bias of S^2 were demonstrated by the simulation.

In addition to the empirical evidence in support of the theoretical results, the simulation suggests certain characteristics of S^2 as an estimator of the function variance and the within subject sample mean

as an estimator of function mean that were not necessarily investigated directly in CHAPTER III. Given below are several impressions concerning the estimators given by the evidence of the simulation viewed in the framework of the results of CHAPTER III and known properties of estimation and sampling in general. All remarks are in reference to sampling periodic functions under PA and PR with an initial time selected randomly from an interval with length the period of the function and a uniform distribution of delays. For some of the problems and the recommendations made so as to minimize the difficulties, there would, of course, exist combinations of minimum times, ranges, and functions which would circumvent the undesirable aspects, but this, in general, would be possible only if a great deal of specific information concerning the functions were available. Certainly, the impressions are somewhat speculative in nature, but the results of the simulation offer considerable support for each of them.

Estimation of variance based upon sample size less than 5 seem undesirable not only because of the relatively large magnitude of the bias of S^2 often found for the small sample sizes, but also because of the characteristics of the distribution of S^2 for the small number of within subject measurements. Note that characteristics such as a high probability of obtaining extreme values for a given subject is present even in random sampling of many functions. The undesirable characteristics of these distributions (except for the bias) is seemingly inherent in the problem of sampling functions and not, exclusively, a problem of pseudo-systematic sampling. Analogous remarks may be made concerning the estimates of within subject means.

For sample size 5 or greater and range one quarter period or more, S^2 seems to be a relatively stable (although perhaps biased) estimation of the variance of periodic functions if certain conditions are avoided. In PA sampling with minimum time near one-half the period and range less than one-half the period, the variance of S^2 seems excessive and the frequency distribution seems bimodal. Both of these conditions would be undesirable properties for an estimator. With the exception of this one sampling condition, the distribution of S^2 seems to be such that for larger sample sizes S^2 would have the desirable properties of a unimodal density and relatively small variance. Support for the stability of S^2 as an estimator is found in the frequency distributions, the downward trend of the variance of S^2, and the upper bounds of the variance for sample size 5 or more discussed earlier. Also, the graphs of the four groups indicate the generally good precision obtained for groups of size 20.

In the use of S^2 as an estimator for the within subject variance of a group, as in the four group portion of the simulation, two alternatives are available for increasing the precision. The downward trend of the variance of S^2 with increased sample size would indicate that taking more within subject measurements would increase precision. Also, since the group $\overline{S^2}$ is a mean based upon a sample size equal to the number of subjects in the group, increasing the number of subjects would increase precision if the subjects are truly homogeneous relative to the pattern over time of the sampled variable. Once again, note that the expected value of S^2 is, in general, a function of sample size. Therefore, in computing a group $\overline{S^2}$, <u>all</u> subjects <u>must</u> have the same number of within

subject measurements if each within subject S^2 is to have the same expected value.

Although for fixed expected time between measurements, an increase in range may decrease the precision of S^2 for a given sample size, employing a larger range seems to be a method for decreasing bias. As was noted above studies with minimum time .25 and range .75 have the same expected time between measurements as do studies with minimum time .5 and range .25. Examination of the summary statistics and critical sizes for these studies shows the magnitude of the bias much smaller for a fixed sample size in the studies with the larger range. Similar results are noted in contrasting studies with minimum time .5 and range .75 with those studies with minimum time .75 and range .25. Further evidence that larger range aids in decreasing bias may be seen by noting that in every study with PA sampling with range .75 the associated $N_{.03}$ is 5 or less.

At the cost of a small loss in precision, then, the undesirable bias property of S^2 may be alleviated somewhat by increasing the range of the delays. Another alternative is to increase the sample size. Under certain sampling plans, the number of within subject measurements necessary to "drive" the magnitude of the bias down to tolerable limits may be so large as to be impracticable. In the PR sampling with minimum time .75 and range .5, for example, the bias was greater than 10% of the total variance even at sample size 23 for all functions. This sampling plan was, of course, the worst in terms of the decline of magnitude of bias with increased sample size and if this plan is avoided increase in sample size does offer a practical method of reducing bias. The decrease

of magnitude of bias with sample size seems much more pronounced in the PA samplings than in the corresponding PR samplings with the same minimum time and range whenever the range was one-half the period or more.

In the discussion of the empirical portion of this investigation given in CHAPTER III it was noted that the function FI was intended to model a variable whose pattern over time is such that most of the time is spent away from its mean, FII models a variable which spends an equal amount of time in all its states, and FIII models a variable which spends most of the time near its mean, but with pulse-like excursions to its extreme values. It would seem, then, that these functions span, in some sense, a large set of the types of periodic variation and that attributes of PA and PR sampling which hold for all three functions should be attributes of the same type of sampling of variables whose time course is either similar to one of the three or intermediate to two of them. Thus, the results discussed here may be useful as guidelines for the design of studies which are attempting to estimate within subject means and/or variance of physiological quantities whose time course is known or suspected to be periodic. If even a crude approximation of the period is available, then the kind of sampling used in the simulation seems to offer a practical method of obtaining reasonably reliable estimates of mean and variance. Of course, the practical conditions associated with a particular investigation will dictate the feasibility of employing some of the indicated results; however, the recommendations listed below may serve as basic principles in designing sampling plans for such studies using pseudo-systematic sampling with the initial time point regarded as a random selection.

1. At least five measurements per subject should be made.
2. If two or more estimates of variance are to be combined to form estimates for a group, the number of measurements must be the same for each subject.
3. In comparing two or more estimates of variance, all variance estimates must be based upon the same number of within subject measurements for each subject involved.
4. Sampling under PR with expected times between samples near an integral multiple of the period should be avoided.
5. Sampling under PA with expected time between measurements near one-half or an odd multiple of half the period should be avoided, but if circumstances necessitate this condition, then the range of delays should be at least one-half the period.
6. If the range is more than one quarter the period, PA sampling is preferable to PR with the exception of the condition discussed in 5 above. This preference is extended even to the smaller ranges for functions that are intermediate to FI and FII.
7. Table 12 may be used as a guide for determining the number of within subject measurements necessary so as to reduce the bias to tolerable limits.
8. The tables for the summary statistics may be used to indicate the number of measurements necessary to achieve the desired stability of the estimates.
9. The range of the delays should be as large as possible commensurate with the practical conditions of the investigation.

A PA sampling plan with a range approximately three-fourths the

period, minimum time approximately one-fourth the period, and ten measurements per subject would be an example of a plan which incorporates the above recommendations.

CHAPTER VI

SUMMARY

It is shown that an analysis of variance with time viewed as a factor (or classification) and the times of measurements viewed as levels of this factor may be an appropriate and fruitful approach for extracting information from many biological studies concerned with the response pattern of a variable over time whenever a meaningful definition of equivalence of time points across experimental units or subjects may be made. The technique of an analysis of variance is applied to data from a previously published biological investigation to demonstrate the applicability of this approach and to indicate how this analysis not only allows appropriate tests of hypotheses essentially equivalent to those tested by the more elementary statistical methods of the original analyses, but also gives information relevant to questions discussed but left unresolved in the publication.

A second part of the study is the consideration of some of the difficulties in obtaining appropriate estimates of within subject variance and means for variables which exhibit a periodic fluctuation over time. It is shown that under sampling schemes often employed in biological investigations the usual estimator, S^2, of within subject variance is biased with the bias being a function of sample size, sampling scheme, and characteristics of the pattern over time. Under these same

sampling schemes, the usual estimate of within subject mean, $\bar{x}$ is shown to be unbiased, but the variance of the estimator is a function of the bias of S^2.

The theoretical considerations of sampling functions is amplified by computer simulation of the sampling of three periodic functions under various sampling plans. Both the theoretical and empirical findings are used to construct a set of guidelines for the design of sampling schemes which may minimize some of the undesirable aspects of the estimators so that the sampling schemes advocated provide a practical method of obtaining estimates of within subject variance and mean when the variable of interest varies periodically over time. The results of this investigation indicate that these same guidelines would aid in obtaining reliable estimates whenever the variable executes a rhythmic but not necessarily periodic variation. However, this conjecture is not specifically investigated at this time.

LIST OF REFERENCES

Brozek, J. 1966 The biology of human variation, Annals N. Y. Acad. of Sci., 134: 497-1066.

Cochran, W. G. 1953 Sampling Techniques. John Wiley & Sons, Inc., New York.

Cole, J. W. L. and Grizzle, J. E. 1966 Applications of multivariate analysis of variance to repeated measurements experiments, Biometrica, 22: 810-828.

Elston, R. C. and Grizzle, J. E. 1962 Estimation of time response curves and their confidence bands, Biometrics, 18: 148-159.

Everson, R. A. 1960 Daily rhythmic variations of blood coagulation times in four species of rodents, Physiol. Zool. 33: 281-287.

Hampton, J. W. 1966 Plasma fibinogen patterns in patients with coronary atherosclerosis, Circulation, 34: 1098-1101.

Hurwicz, L. 1962 Basic mathematical and statistical considerations in the study of rhythms and near rhythms, Annals N. Y. Acad. of Sci., 98: 851-857.

Levi, L. 1968 Sympatho-adrenomedullary and related biochemical reactions during experimentally induced emotional stress, Endocrinology and Human Behaviour. Oxford University Press, New York: 200-219.

Nagorra-Stasiak, B. 1963 Diurnal variation of the clotting and bleeding time in rabbits, Acta Physiologia Polonica, 14: 500-506.

Parzen, E. 1960 Modern Probability Theory and Its Application. John Wiley & Sons, Inc., New York.

Potthoff, R. F. and Roy, S. N. 1964 Generalized analysis of variance model useful especially for growth curve problems, Biometrics, 51: 313-326.

Scheffe´, H. 1959 The Analysis of Variance. John Wiley & Sons, Inc., New York.

Scheving, L. E. and Pauly, J. E. 1967 Daily rhythmic variation in blood coagulation times in rats, Anat. Rec., 157: 657-666.

Siegal, S. 1956 Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Book Company, New York.

Sollberger, A. 1962 General properties of biological rhythms, Annals N. Y. Acad. of Sci., 98: 757-774.

Steel, R. G. D. and Torrie, J. H. 1960 Principles and Procedures of Statistics. McGraw-Hill Book Company, New York.

Vandenberg, S. G. 1966 Some advances in statistical analysis of human variations, Annals of N. Y. Acad. of Sci., 134: 526-537.

Winer, B. J. 1962 Statistical Principles in Experimental Design. McGraw-Hill Book Company, New York.

Wolf, W. 1962 Rhythmic functions in the living system, Annals N. Y. Acad. of Sci., 98: 753-1326.