

AFFDL-TR-65-57  
AD620987

## GAUSSIAN APPROXIMATIONS TO THE DISTRIBUTION OF SAMPLE COHERENCE

L. D. ENOCHSON  
N. R. GOODMAN

MEASUREMENT ANALYSIS CORPORATION

TECHNICAL REPORT AFFDL-TR-65-57

CLEARINGHOUSE FOR FEDERAL-AID RESEARCH AND TECHNICAL INFORMATION	
Hardcopy	
\$2.00	0.50 35 as
ARO	

JUNE 1965

AIR FORCE FLIGHT DYNAMICS LABORATORY  
RESEARCH AND TECHNOLOGY DIVISION  
AIR FORCE SYSTEMS COMMAND  
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

## FOREWORD

This report was prepared by Measurement Analysis Corporation, Los Angeles, California, for the Aerospace Dynamics Branch, Vehicle Dynamics Division, AF Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio 45433, under Contract No. AF33(615)-1418. The research performed is part of a continuing effort to provide advanced techniques in the application of random process theory and statistics to vibration problems which is part of the Research and Technology Division, Air Force Systems Commands exploratory development program. The contract was initiated under Project No. 1370, "Dynamic Problems in Flight Vehicles," Task No. 137005, "Prediction and Control of Structural Vibration." Mr. R. G. Merkle of the Vehicle Dynamics Division, FDDS, was the project engineer.

This report covers work conducted from March 1964 to January 1965. The contractor's report number is MAC 403-07. Manuscript released by authors February 1965 for publication as an Air Force Flight Dynamics Laboratory Technical Report. One of the coauthors of this report, Dr. N. R. Goodman, is a mathematical-statistical consultant to the Measurement Analysis Corporation.

This technical report has been reviewed and is approved.

*Walter J. Mykytow*  
WALTER J. MYKYTOW  
Asst. for Research & Technology  
Vehicle Dynamics Division  
AF Flight Dynamics Laboratory

## ABSTRACT

This report describes the results of an empirical study to develop a normalizing transformation for sample multiple coherence. The "Fisher z-transformation" is employed. The expected value (including a bias term) and variance of the transformation have been experimentally determined. Numerical values of the transformation which is developed (including the bias term and variance) may be obtained with a reasonable amount of computation. Tables of the Gaussian distribution can then be used to obtain confidence limits and perform statistical tests.

The computational methods and the digital computer program used for the study are described in detail. Flow charts of the program are given. Numerical results from the program results are presented. Examples of the use of the transformation are given for developing confidence limits for multiple coherence. A recommendation for a further theoretical study is presented.

**GAUSSIAN APPROXIMATIONS TO THE  
DISTRIBUTION OF SAMPLE COHERENCE**

**CONTENTS**

	<b>Page</b>
1. Introduction .....	1
2. The Normalizing Transformation .....	2
3. Computational Procedures .....	5
4. Numerical Results .....	11
5. Confidence Bound Examples .....	21
6. Conclusions and Recommendations .....	26
References .....	27

## 1. INTRODUCTION

In measuring coherence when moderately long or long records are available, one often has the "effective degrees-of-freedom" parameter  $n > 20$ . A large set of tables consisting of nine volumes has been developed by the Rocketdyne Division of North American Aviation. These tables provide the sampling distribution of coherence when the effective degrees-of-freedom is roughly less than 20, and when the number of variables involved in the coherence relation is less than eleven. This set of tables is listed as Reference 1 in this report.

In regard to many practical applications, it is desirable to have a tractable and reasonably accurate approximation to the distribution of sample coherence for  $n > 20$ . Such an approximation is especially needed for the range  $20 \leq n \leq 150$ . The degrees-of-freedom parameter in practical spectral analysis often falls in this range, and hence the degrees-of-freedom for coherence functions obtained from the spectral analysis also falls in this range.

The object of the study described in this report was that of obtaining a Gaussian distributed approximation to sample coherence. A reasonable approximation which has a roughly Gaussian distribution has been empirically obtained by a digital computer numerical analysis. This approximation appears to be useful in the range where true coherence varies from roughly 0.4 to 0.95, and the degrees-of-freedom parameter is larger than 20. A slightly more complicated version of the transformation involved in the approximation gives corrections which extend the usefulness of the approximation down to approximately a true coherence value of 0.1, and to the degrees-of-freedom parameter as low as 10. The approximation developed is very useful in that the need to have a large set of tables available for many practical applications is eliminated. One instead can substitute a table of the Gaussian distribution function which is widely and conveniently available. A certain amount of computation is necessary in order to transform the value of sample coherence to the approximately Gaussian distribution variable; however, the amount is not unreasonable. The transformation and the empirical procedures used in obtaining the final version will now be described.

## 2. THE NORMALIZING TRANSFORMATION

The probability density function for sample coherence is given by the following general formula:

$$C(x | n, p, \gamma^2) = \frac{\Gamma(n)}{\Gamma(p-1) \Gamma(n-p+1)} (1 - \gamma^2)^n x^{p-2} (1-x)^{n-p} F(n, n-p; p-1; \gamma^2 x) \quad (1)$$
$$(0 \leq x \leq 1)$$

where

$n$  = effective number of degrees-of-freedom (d. f.)

$p$  = effective number of records

$\gamma^2$  = true coherence

$x = \hat{\gamma}^2$  = sample coherence

$F(n, n-p; p-1; \gamma^2 x)$  = hypergeometric function

A closed form expression for the cumulative distribution function corresponding to Eq. (1) is known. The closed form expression is, however, intractable for  $n$  in the range desired. That is, the computational requirements to obtain a numerical value for a specific  $n$  requires more than a reasonable amount of computation. Recursion relations on  $n$  have also been developed. The tables described in Reference 1 were computed using these relations. These recursion formulas are suitable for obtaining results for  $n+1$  in terms of the results for  $n$ . However, they are not suitable for directly computing a numerical value for any arbitrary value of  $n$ .

The asymptotic ( $n$  approaching infinity) distribution corresponding to Eq. (1) has been investigated. However, comparisons with the existing tables of Reference 1 indicated that approximations based on such asymptotic theory would be far from reasonably accurate for  $20 \leq n \leq 150$ . Other preliminary studies have indicated that simplicity and reasonable accuracy may be achieved by a "normalizing" (Gaussian producing) transformation of the form described below.

Let

$$z = \frac{1}{2} \ln \left( \frac{1 + \sqrt{x}}{1 - \sqrt{x}} \right) = \tanh^{-1} \sqrt{x} \quad (2)$$

$$\zeta = \frac{1}{2} \ln \left( \frac{1 + \gamma}{1 - \gamma} \right) = \tanh^{-1} (\gamma) \quad (3)$$

In Eqs. (2) and (3),  $x = \hat{\gamma} = \sqrt{\text{sample coherence}}$  and  $\gamma = \sqrt{\text{true coherence}}$ .

The random variable  $z$  is, to a close approximation, normally (Gaussian) distributed  $N(\mu, \sigma^2)$ . The mean  $\mu$  is given by an expression of the form

$$\mu = \zeta + \frac{f(\gamma)}{n - p + k} = \zeta + b(\gamma) \quad (4)$$

and the variance by

$$\sigma^2 = \frac{1}{2(n - p + k)} \quad (5)$$

In Eqs. (4) and (5), the constant  $k$  and the function  $f(\gamma)$  are not yet specified.

The choice of the form of the transformation given by Eq. (2) is motivated by a comparison of multiple coherence with the multiple correlation coefficient of classical statistics. When this transformation (termed the 'Fisher z-transformation') is applied to the ordinary correlation coefficient, one obtains a variable which is normally distributed to a very high degree of accuracy. Biases do exist in the transformation of the ordinary correlation coefficient, but they are usually negligible for practical purposes. However, for the case of multiple coherence, the transformation by itself is not quite so suitable. The bias term denoted by  $b(\gamma)$  can be significant. One of the major objects of the study described in this report was to determine the numerical behavior of this function  $b(\gamma)$  and to determine if it could be represented by a reasonably simple form.

The asymptotic ( $n$  approaching infinity) distribution of the random variable  $z$  is normal with a mean of  $\zeta$  and a variance of  $1/2n$ , provided  $\gamma \neq 0$ . Initial comparisons which were made with tabulated results indicated, however, that for  $n$  near 20, the asymptotic distribution is still far from acceptable approximation. One may view Eqs. (4) and (5) as giving finite corrections to the asymptotic distribution of the random variable  $z$ . For various hand calculated examples, corrections of the form indicated by Eqs. (4) and (5) were introduced and acceptable approximations to the distribution of sample coherence were obtained. These hand calculations motivated the performance of a more comprehensive examination by means of a digital computer program.

The object of this program was to determine if  $z$  is a useful transformation for coherence when suitable bias and variance terms were introduced. Thus, determining the form of the bias  $b(\gamma)$  and the value of  $k$  which would fit in the expressions for the mean and variance were to be determined by the digital computer program. The next section of the report describes the computer program which was written, and the section after that presents the results which were obtained.

### 3. COMPUTATIONAL PROCEDURES

The basic method of determining the accuracy of the normalizing transformation was to compute results using the transformation and compare them with the existing tables of sample coherence. These tables are available for  $n$  up to roughly 20. It is assumed that if an acceptable transformation is determined for this small  $n$ , then the transformation will also be acceptable for larger  $n$ . It was decided to determine the bias of the transformation by comparing the deviation corresponding to the 50% point (median) of the sample coherence table from that corresponding to the true value of coherence. It was also decided to establish the variance for the transformation by comparing the 95% point of the distribution. That is, for a Gaussian distribution the 95th percentile should be 1.645 standard deviations away from the mean. Therefore, the difference between the 50th and 95th percentiles was set equal to 1.645 standard deviations. This allowed the determination of the variance.

After the bias and the variance were determined, comparisons for several probability points were then made of the hypothesized normally distributed transformed values with the exact tabulated distribution values. To be more specific, the computational steps proceed as follows.

- (1) Select a value  $\gamma^2$  of true coherence.
- (2) Search the appropriate table of sample coherence for the 50th percentile of the distribution of  $\hat{\gamma}^2$  given  $\gamma^2$ . This procedure actually requires interpolation in the existing table. This interpolation was performed using 3 point quadratic interpolation defined by the equation

$$\hat{\gamma}_{.50}^2 = \frac{(p-p_2)(p-p_3)}{(p_1-p_2)(p_1-p_3)} \hat{\gamma}_1^2 + \frac{(p-p_1)(p-p_3)}{(p_2-p_1)(p_2-p_3)} \hat{\gamma}_2^2 + \frac{(p-p_1)(p-p_2)}{(p_3-p_1)(p_3-p_2)} \hat{\gamma}_3^2 \quad (6)$$

In Eq. (6),

$$p = .50$$

$p_1, p_2, p_3$  are the three distribution function values  
which bracket  $p = .50$

$\hat{\gamma}_1^2, \hat{\gamma}_2^2, \hat{\gamma}_3^2$  are the arguments corresponding to the  
distribution function values  $p_1, p_2$ , and  $p_3$

- (3) The bias at the 50% point is then determined by transforming  
the 50% point of sample coherence and comparing that with the  
transformed value of the true coherence. That is,  $b$  is  
defined by the equation:

$$z_{0.50} = \tanh^{-1} \hat{\gamma}_{0.50}^2, \quad \zeta = \tanh^{-1} \gamma \quad (7)$$

$$\text{bias } \equiv b = z_{0.50} - \zeta$$

- (4) After the bias is determined, then the variance is determined  
by a similar procedure. First, the 95th percentile,  $\hat{\gamma}_{.95}^2$ , is  
obtained by interpolation with a formula the same as Eq. (6),  
but with the following quantities replacing those in Eq. (6).

$$p = .95$$

$p_4, p_5, p_6$  are the three distribution function values which  
bracket  $p = .95$  (these replace  $p_i$ ,  $i = 1, 2, 3$  in  
Eq. (6)).

$\hat{\gamma}_4^2, \hat{\gamma}_5^2, \hat{\gamma}_6^2$  are the arguments corresponding to  $p_4, p_5, p_6$   
and replace  $\hat{\gamma}_i^2$ ,  $i = 1, 2, 3$ .

- (5) The variance for the transformed variable is then determined from the following relation

$$\sigma_{\hat{\gamma}}^2 = \left( \frac{z_{.95} - z_{.50}}{1.645} \right)^2 \quad (8)$$

Points on the normal distribution function with a mean

$$\mu_{\hat{\gamma}} = \xi + b = z_{0.50} \quad (9)$$

and a variance determined by Eq. (8) are next computed for values of  $\hat{\gamma}^2$  in the range 0.05 to 0.95 in steps of 0.05.

- (6) Since the asymptotic form of the variance of the sampling distribution of the transformed coherence is of the form  $\sigma^2 = 1/2n$ , it was decided to obtain an "effective" number of degrees-of-freedom  $n_e$  from the relation

$$n_e = \frac{1}{2\sigma^2} \quad (10)$$

- (7) Finally, the probability values which have been computed, the bias, and the effective degrees-of-freedom are printed out.

For a given degrees-of-freedom parameter  $n$  and a given number of variables  $p$ , the above procedure was repeated for values of true coherence ranging from 0.00 to 0.90 in steps of 0.10, and from 0.91 to 0.99 in steps of 0.01. It was also suspected that the results would vary roughly as a function of the square root of  $n_e$ . Hence, 24 cases were run with  $\sqrt{n_e}$  running from roughly 2.0 to 4.5 in steps of about 0.5. The cases which were actually run are given in Table 1 below. In Table 1,  $n_e = n - p + 1$ .

TABLE 1  
Values of n and p for Which Computer Computations Were Performed

$\sqrt{n_e}$	1.7	2.4	3.0	3.5	4.0	4.4, 4.4, 4.3
$n_e$	3	6	9	12	16	19, 19, 18
n, p	4, 2	7, 2	10, 2	13, 2	17, 2	
	5, 3	8, 3	11, 3	14, 3	18, 3	21, 3
	6, 4	9, 4	12, 4	15, 4	19, 4	22, 4
	7, 5	10, 5	13, 5	16, 5	20, 5	22, 5

This selection of values of n and p provide a coverage of the existing tables for  $2 \leq p \leq 5$  and allowed for the investigation of the behavior of the approximation under a reasonably wide range of values of n.

A flow chart which outlines the computational procedure used is given in Figures 1 and 2. The polynomial approximations given in these figures for the Gaussian distribution function were obtained from Reference 2. These approximations were checked with existing table values, and agreement was found out to at least five decimal places in all cases.

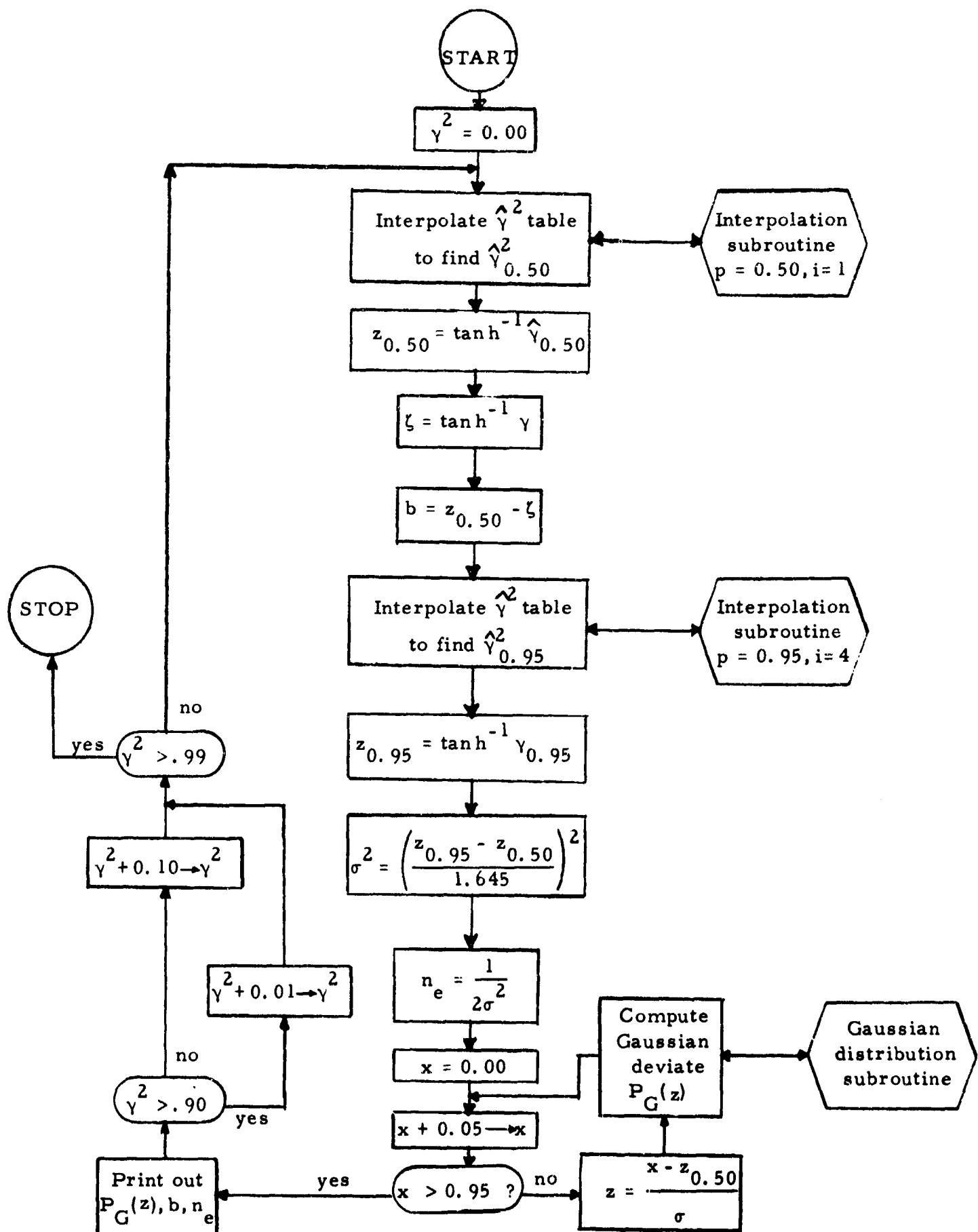


Figure 1. Over-all Computation Procedure

$a_0 = .797884561$   
 $a_1 = .531923007$   
 $a_2 = .319152933$   
 $a_3 = .151968751$   
 $a_4 = .0590540356$   
 $a_5 = .0191982820$   
 $a_6 = .00519877502$   
 $a_7 = .00107520405$   
 $a_8 = .000124818987$   
  
 $b_0 = .999936658$   
 $b_1 = .000535310849$   
 $b_2 = .00214125874$   
 $b_3 = .00535357911$   
 $b_4 = .00977945334$   
 $b_5 = .0116304473$   
 $b_6 = .0105576250$   
 $b_7 = .00654979121$   
 $b_8 = .00203425487$   
 $b_9 = .000794620820$   
 $b_{10} = .00139060284$   
 $b_{11} = .000676904986$   
 $b_{12} = .000019538132$   
 $b_{13} = .00015259290$   
 $b_{14} = .000045255659$

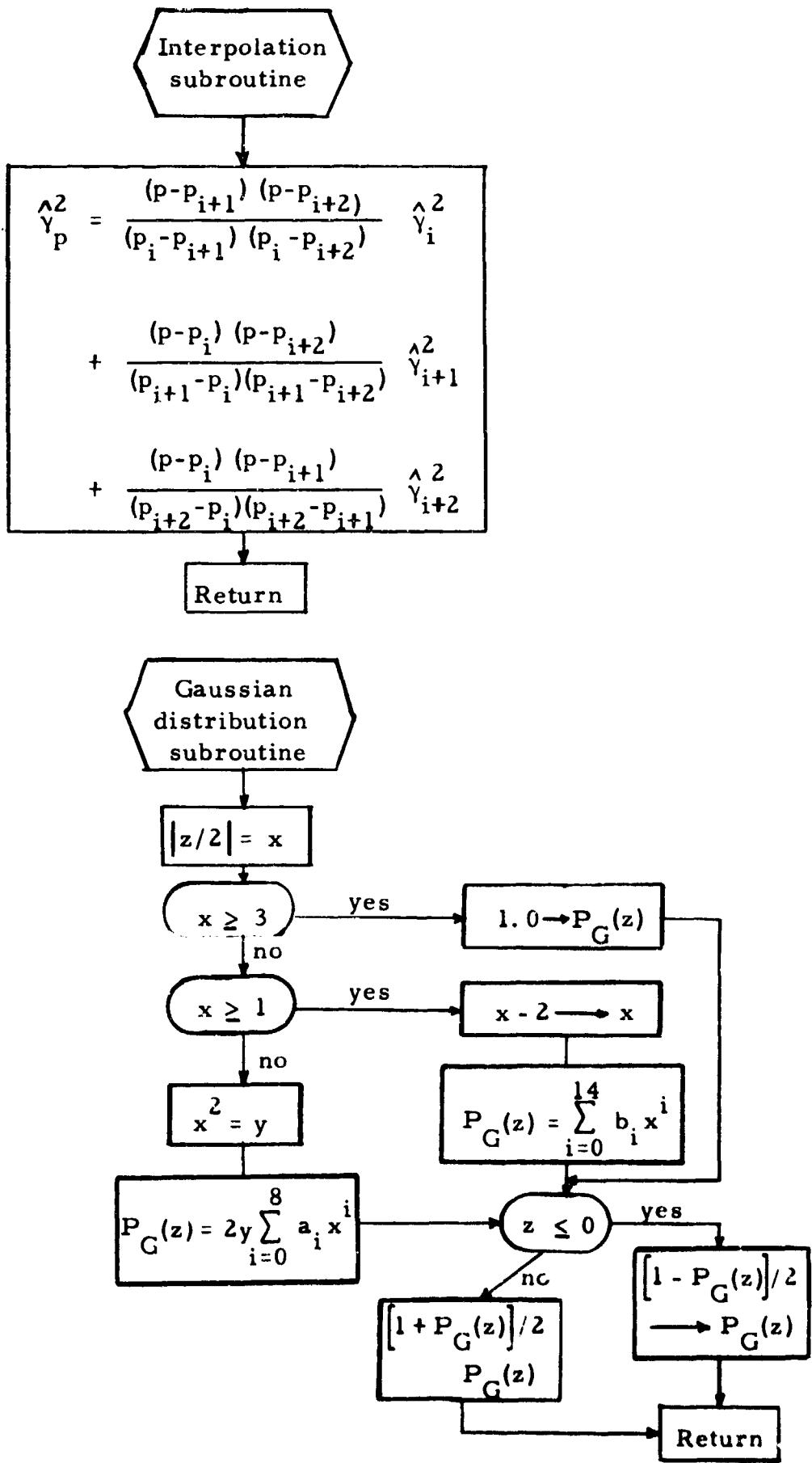


Figure 2. Interpolation and Gaussian Distribution Subroutines

#### 4. NUMERICAL RESULTS

As a general statement, the comparison of the probabilities obtained with the Gaussian approximation (transformation) with the bias and variance adjustments incorporated, agreed with the existing multiple coherence tables to two or three places for fairly large degrees-of-freedom. For smaller degrees-of-freedom, the probabilities were still often quite close except at times toward the tails of the distribution. Even here the errors were seldom worse in magnitude than 0.01 or 0.02. The actual computer output of all the probability values computed is very extensive, and the complete results are not included in this report in the interest of economy of space. However, the bias values obtained for all the cases are plotted in Figure 3 through Figure 8 (located at the end of this section) along with two approximations to the bias correction described below. Also, an example of the agreement in probabilities is given later in this section.

An analysis of the numerical values obtained for the effective degrees-of-freedom  $n_e$  obtained via the variance fitting computations indicates that a reasonable value to use for  $n_e$  is given by

$$n_e = \begin{cases} n - 1 & , p = 2 \\ n - p + 2 & , p > 2 \end{cases} \quad (11)$$

and the variance is

$$\sigma^2 = 1/2n_e \quad (12)$$

This value becomes less accurate as coherence and degrees-of-freedom become smaller and as  $p$  becomes larger. However, it is a quite satisfactory approximation in the range of roughly true coherence between 0.4 and 0.95.

A very simple straightforward constant bias correction, that is, one which is not a function of the true coherence is given by the formula

$$b = \frac{p - 1}{2(n - p + 1)} \quad (13)$$

This value is plotted in Figures 3 through 8 as the constant line. From this it can be seen that for fairly large degrees-of-freedom this value of  $b$  is a quite satisfactory approximation for the bias for true coherence values of roughly 0.4 through 0.95.

One comment should be made about the plots of the bias correction at this time. From the general theory of the approximations, one would expect the singularity which appears to occur near true coherence equal to zero. However, there is no immediate explanation for the apparent singularity which occurs at true coherence equal to one. This leads one to believe that the erratic results obtained as true coherence approaches closely to one may actually be due to computational noise. That is, certain computational instabilities could possibly be occurring in this area which would result in the unexpected results obtained here. This problem has not yet been resolved. Further theoretical studies will tend to shed more light on this question. A modified bias correction gives somewhat better results for lower values of true coherence and lower degrees-of-freedom. The correction is given by the formula

$$\begin{aligned} b_1(\gamma) &= \frac{p - 1}{2(n - p + 1)} \left[ \frac{1}{2} \left( 1 + \frac{1}{(\gamma + .1)} \right) \right] \\ &= \frac{p - 1}{4(n - p + 1)} \left( 1 + \frac{1}{\gamma + .1} \right) \end{aligned} \quad (14)$$

This equation is plotted as the dotted lines on Figures 3 through 8. Note that this bias correction has the disadvantage of having true coherence

appear in the formula. The true coherence will, in general, not be known and hence one would have to use either an estimated value or a value of the sample coherence. The use of this formula would still be more likely to provide a slightly better bias correction term than that given by Eq. (13). Again, further theoretical studies concerning normalizing transformations for coherence would tend to resolve this question.

The plots given in Figures 3 through 8 depict empirically the behavior of the bias term  $b(\gamma)$  defined by Eq. (4). Two approximating forms for  $b(\gamma)$  have been given by Eqs. (13) and (14). It is this empirical behavior whose analytical form should be determined. The computer output for  $n_e = (n-p+k)$  as defined by the variance formula, Eq. (5) is not presented. Equation (11) summarizes the values of  $k$  which were determined as suitable approximations from the computer study.

A typical example is illustrated in Table 2 of the agreement between the sampling distribution probabilities obtained in three ways: (a) the approximation using the bias correction of Eq. (13) and the variance of Eq. (12), (b) the digital computer fitted approximation using the median bias (see Figure 6) and the variance determined from the 95th percentile, and (c) the exact probabilities obtained from the tables of Reference 1. The case chosen is for  $\gamma^2 = 0.70$ ,  $n = 15$ , and  $p = 4$ .

TABLE 2  
Comparison of Approximate and Exact Sample Coherence Probabilities

$\hat{\gamma}^2$	$\hat{\gamma}$	arc tanh $\hat{\gamma}$	Z	(a) P(z)	(b) P(z)	(c) P(z)
.50	.707	.881	-2.321	.010	.008	.008
.55	.742	.955	-1.944	.026	.022	.022
.60	.775	1.043	-1.495	.068	.054	.054
.65	.806	1.117	-1.117	.132	.123	.123
.70	.837	1.211	-.628	.261	.251	.252
.75	.866	1.317	-.097	.461	.456	.456
.80	.895	1.447	.566	.714	.710	.707
.85	.923	1.603	1.362	.913	.916	.910
.90	.950	1.832	2.531	.994	.995	.992

$$\hat{\gamma}^2 = 0.70, n = 15, p = 4$$

- (a) Gaussian Probabilities Obtained Using Eqs. (12) and (13)  
for Variance and Bias. Bias = b = 3/24 = .125,  
Standard Deviation =  $\sigma = 1/\sqrt{26} = .196$
- (b) Gaussian Probabilities Obtained Using Digital Computer  
Fitted Median Bias and 95th Percentile Variance
- (c) Exact Probabilities Obtained Using Tables of Reference 1

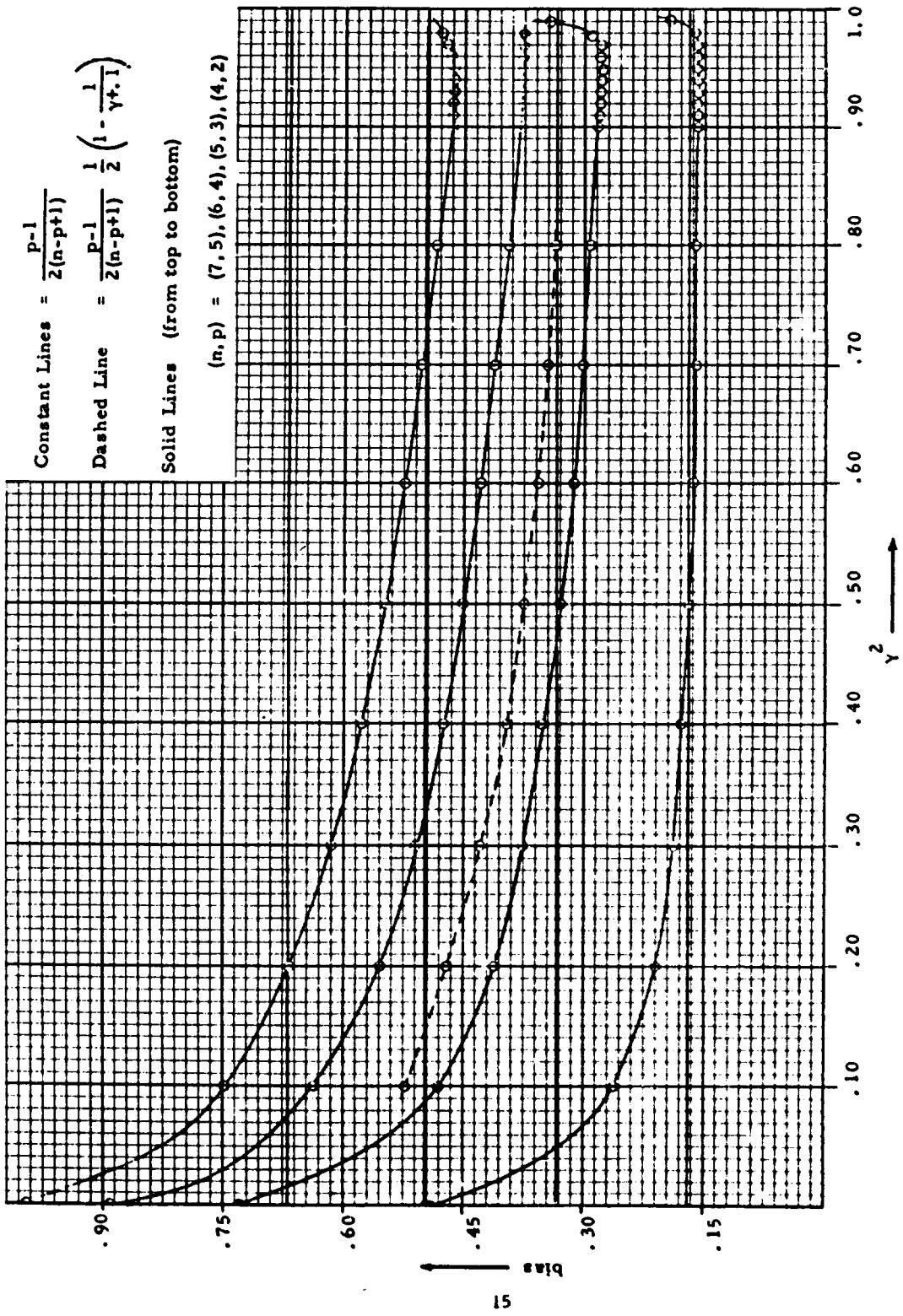


Figure 3. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 3$

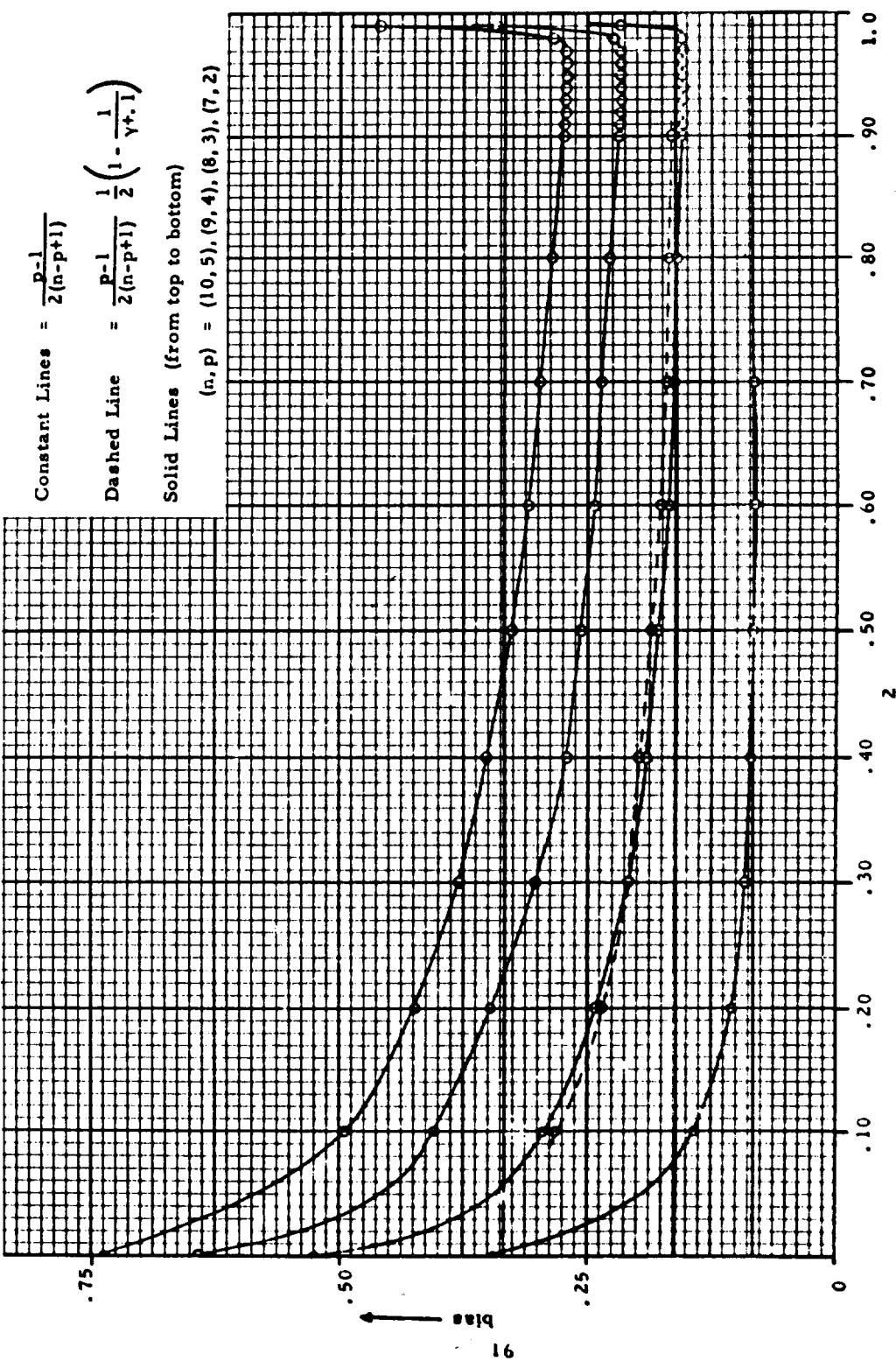


Figure 4. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 6$

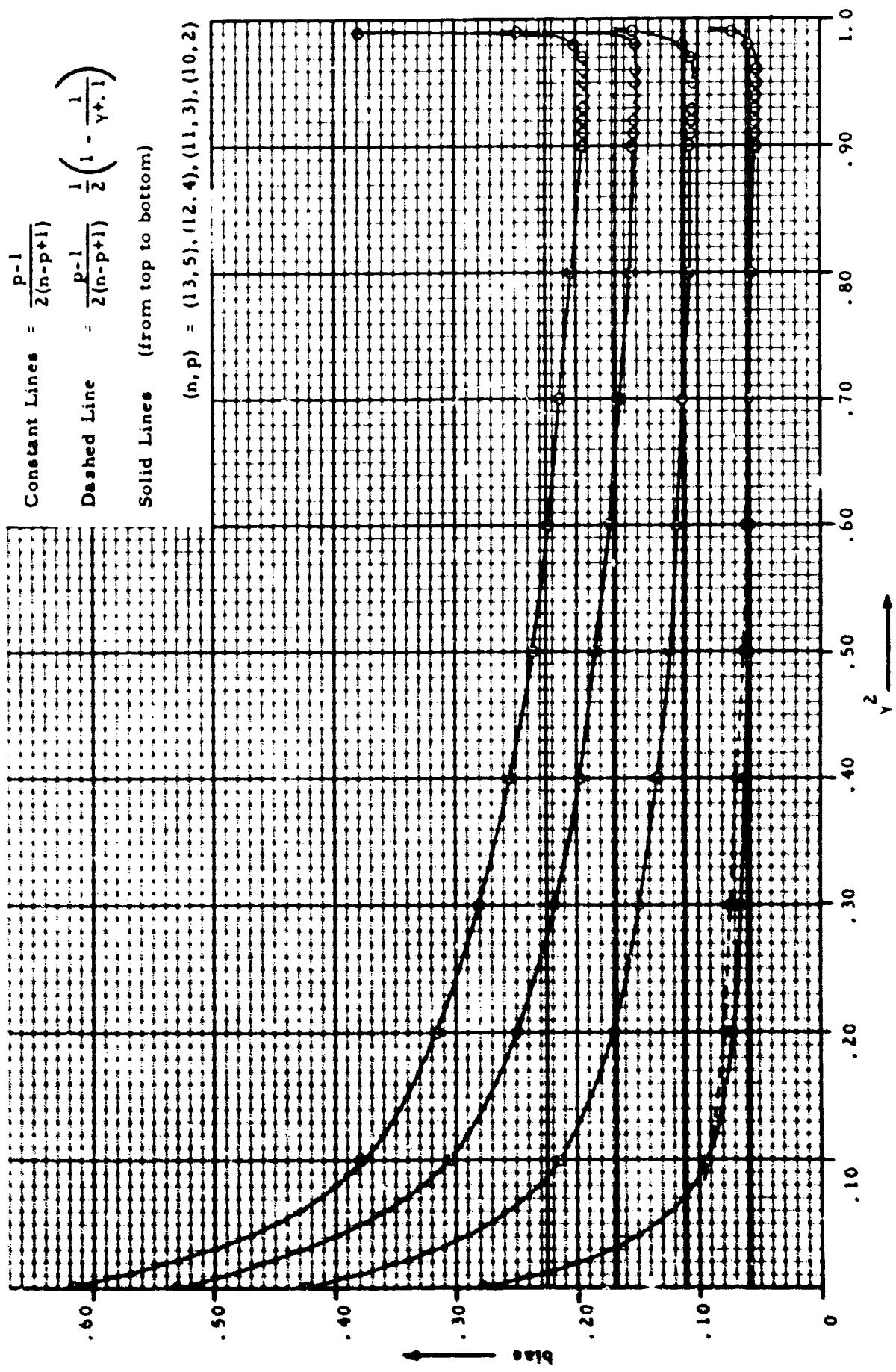


Figure 5. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 9$

Figure 7. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 16$

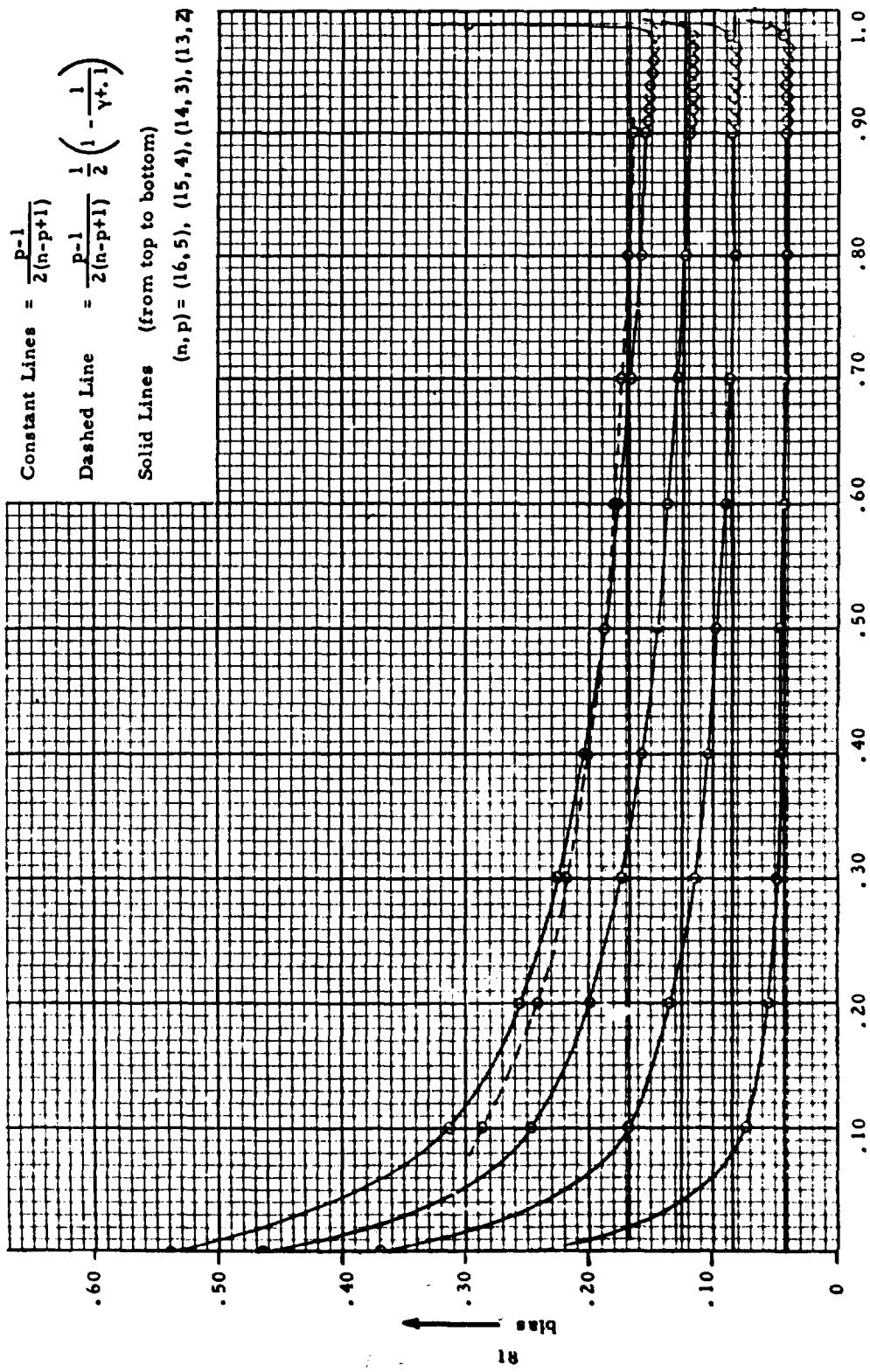


Figure 6. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 12$

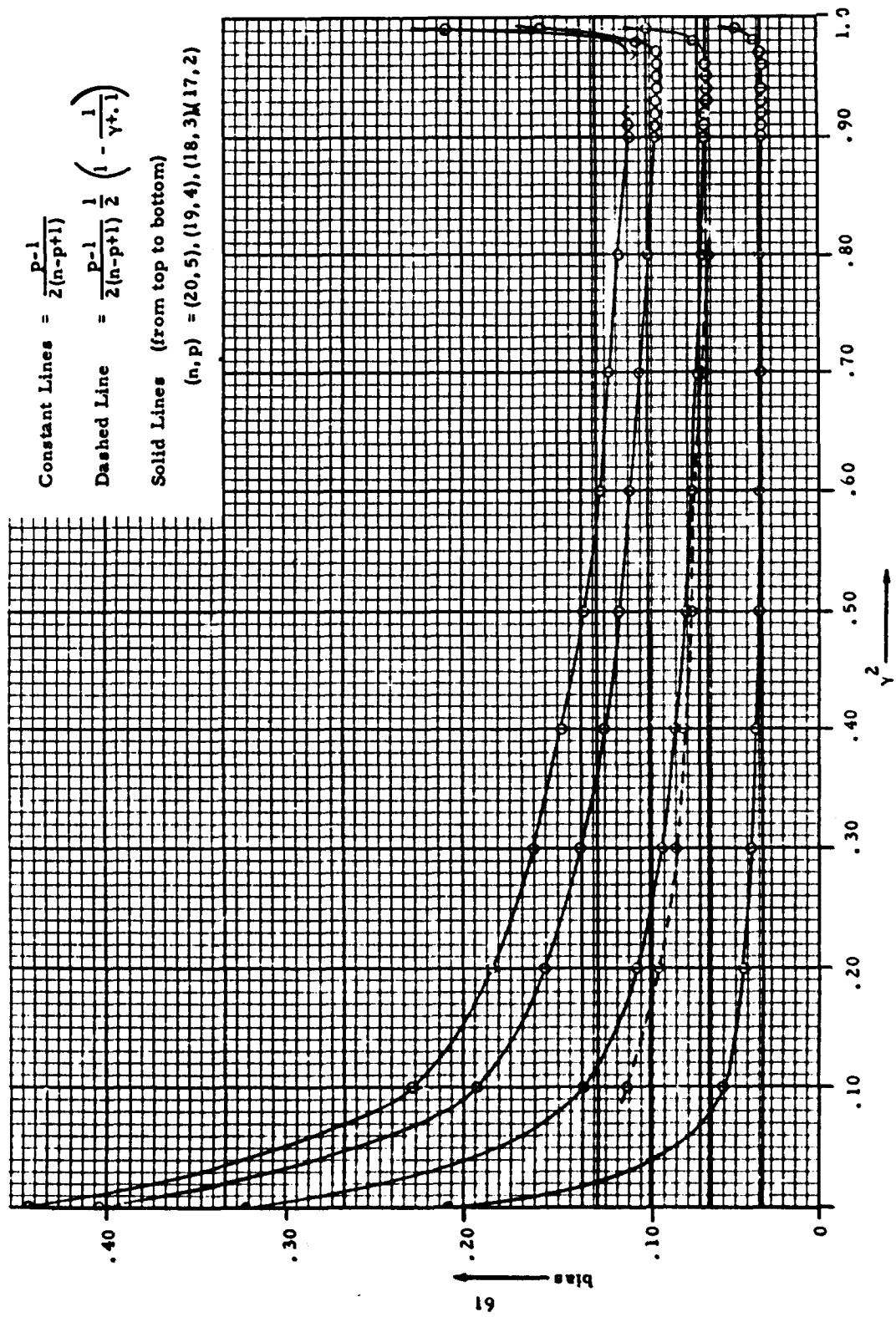


Figure 7. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 16$

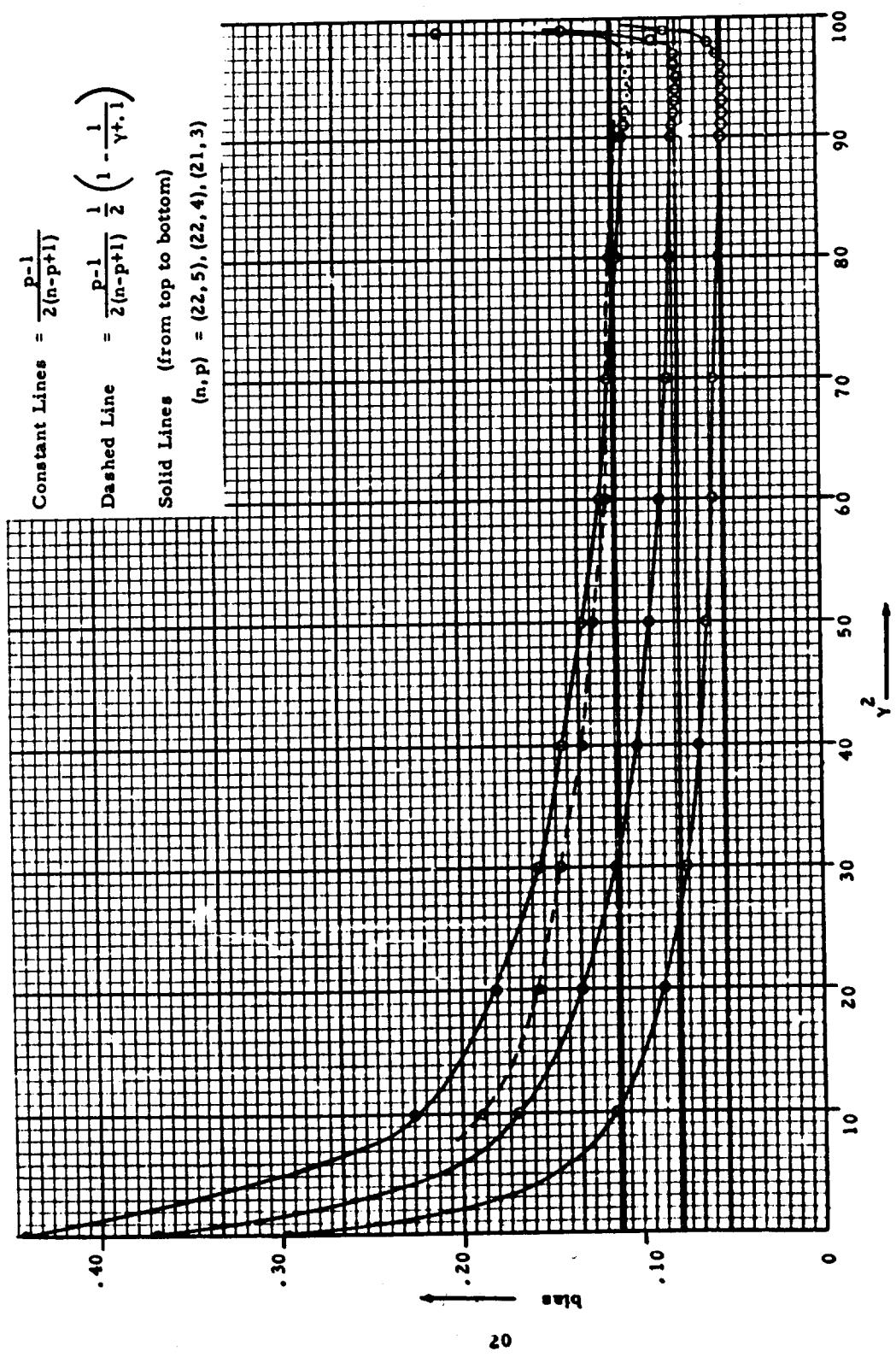


Figure 8. Graphs of Bias Error for Normalizing Transformation for  $n_e = n - p + 1 = 18, 19, 19$

## 5. CONFIDENCE BOUND EXAMPLES

The transformation which is a reasonably accurate normal approximation for moderate values of true coherence ( $0.4 < \gamma^2 < 0.95$ ) and also useful for more extreme values is given by Eq. (2).

$$z = \tanh^{-1} \hat{\gamma} = \frac{1}{2} \ln \frac{1 + \hat{\gamma}}{1 - \hat{\gamma}}$$

where  $\hat{\gamma}$  is the positive square root of sample coherence. In Section 4, it has been empirically determined that  $z$  is approximately normal with an approximate mean value

$$\mu_z = \frac{1}{2} \ln \frac{1 + \gamma}{1 - \gamma} + \frac{p - 1}{2(n - p + 1)} \quad (15)$$

where  $\gamma$  is the square root of true coherence. The variance of  $z$  is approximately

$$\sigma_z^2 = \frac{1}{2(n - p + k)} \quad , \quad \begin{array}{l} k = 1 \text{ for } p = 2 \\ k = 2 \text{ for } p > 2 \end{array} \quad (16)$$

From the above equations, if one selects a confidence level  $(1 - \alpha)$  and is given a value of  $\gamma^2$ , then

$$P\left[\mu_z - \sigma_z Z_{1-\alpha/2} < z \leq \mu_z + \sigma_z Z_{1-\alpha/2}\right] = 1 - \alpha \quad (17)$$

where  $Z_{1-\alpha/2}$  is the  $(1 - \alpha/2)$  percentile of the normal distribution.

Let  $\mu_z$  be broken into two parts

$$\xi = \frac{1}{2} \ln \frac{1 + \gamma}{1 - \gamma}$$

$$b = \frac{p - 1}{2(n - p + 1)}$$

Then, Eq. (17) can be rearranged to give  $(1 - \alpha)$  confidence bounds on  $\gamma$  which are defined by the inequality

$$\tanh(z - b - Z_{1-\alpha/2} \sigma_z) \leq \gamma \leq \tanh(z - b + Z_{1-\alpha/2} \sigma_z) \quad (18)$$

To employ the foregoing equations in developing confidence bounds for the various types of coherence, certain adjustments must be made in the parameters  $n$  and  $p$ . The adjustments are discussed in Reference 3, and will be reviewed here. First, define  $p' =$  the total number of variables being considered,  $n' = BT$  as the number of degrees-of-freedom in the associated spectral analysis. (In many documents, the number of degrees-of-freedom is defined as  $n = 2BT$  exactly double the value used here. One must use care in comparing degree-of-freedom parameters from such documents with this report.)

- (1) For general multiple coherence between one output variable and  $p' - 1$  input variables, set  $n = n'$  and  $p = p'$ .
- (2) For marginal multiple coherence between one output variable and  $p' - 1 - q$  input variables ( $q$  variables being neglected in the relation). set  $n = n'$  and  $p = p' - q$ .
- (3) For conditional (partial) multiple coherence between one "conditioned" output variable and  $p' - 1 - l$  "conditioned" input variables (the linear effect of  $l$  variables has been subtracted out), set  $n = n' - l$  and  $p = p' - l$ .
- (4) For marginal conditional multiple coherence between one conditioned output variable and  $p' - 1 - l - q$  conditioned input variables (the linear effect of  $l$  variables has been subtracted out and  $q$  variables are being neglected), set  $n = n' - l$  and  $p = p' - l - q$ .

For example, for ordinary coherence between one input and one output, set  $p = 2$  and  $n = n'$ . For partial coherence between one input and one output where the effects of a second input are subtracted out, set  $p = 2$  and  $n = n' - 1$ .

Note that the effect of the adjustment in the degrees-of-freedom will be small for small  $p'$  and large for  $n'$ , say  $n' > 30$ . The bias correction is also minor for many practical purposes for small  $p'$  and large  $n'$ . For example, for  $\gamma = 0.9$  ( $\hat{\gamma}^2 \approx 0.95$ ),  $p = 2$ , and  $n = 25$ , the relative bias is

$$\frac{b}{\xi} = \frac{0.02}{1.47} = 0.014 = 1.4\%$$

The bias increases at least linearly with  $p$ . Hence, for larger  $p$ , the percentagewise effect becomes significant.

The use of Eq. (18) is illustrated in Figure 9 and Table 3. The 99% confidence limits for  $n = BT = 50$  and  $p = 2$  are plotted and tabulated there. The dashed lines for the confidence limits where  $\hat{\gamma}^2$  is less than 0.3 is to indicate that the approximation is becoming unreliable in this area. In order that the general shape of the approximate and true confidence limit curves may be compared, true values are presented in Figure 10 for  $n = 21$  and  $p = 2$ . The values for these curves are obtained from the tables of Reference 1.

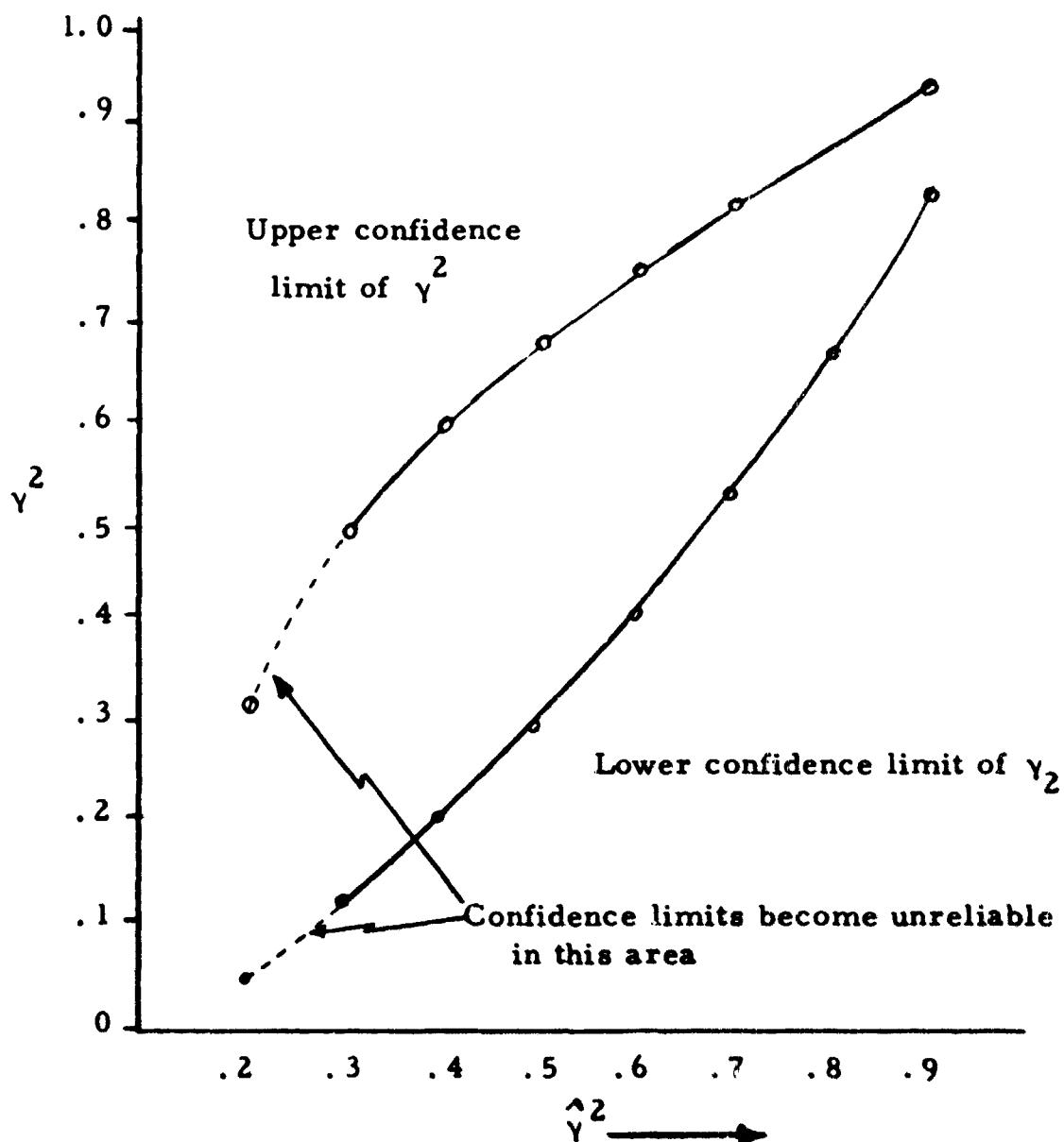
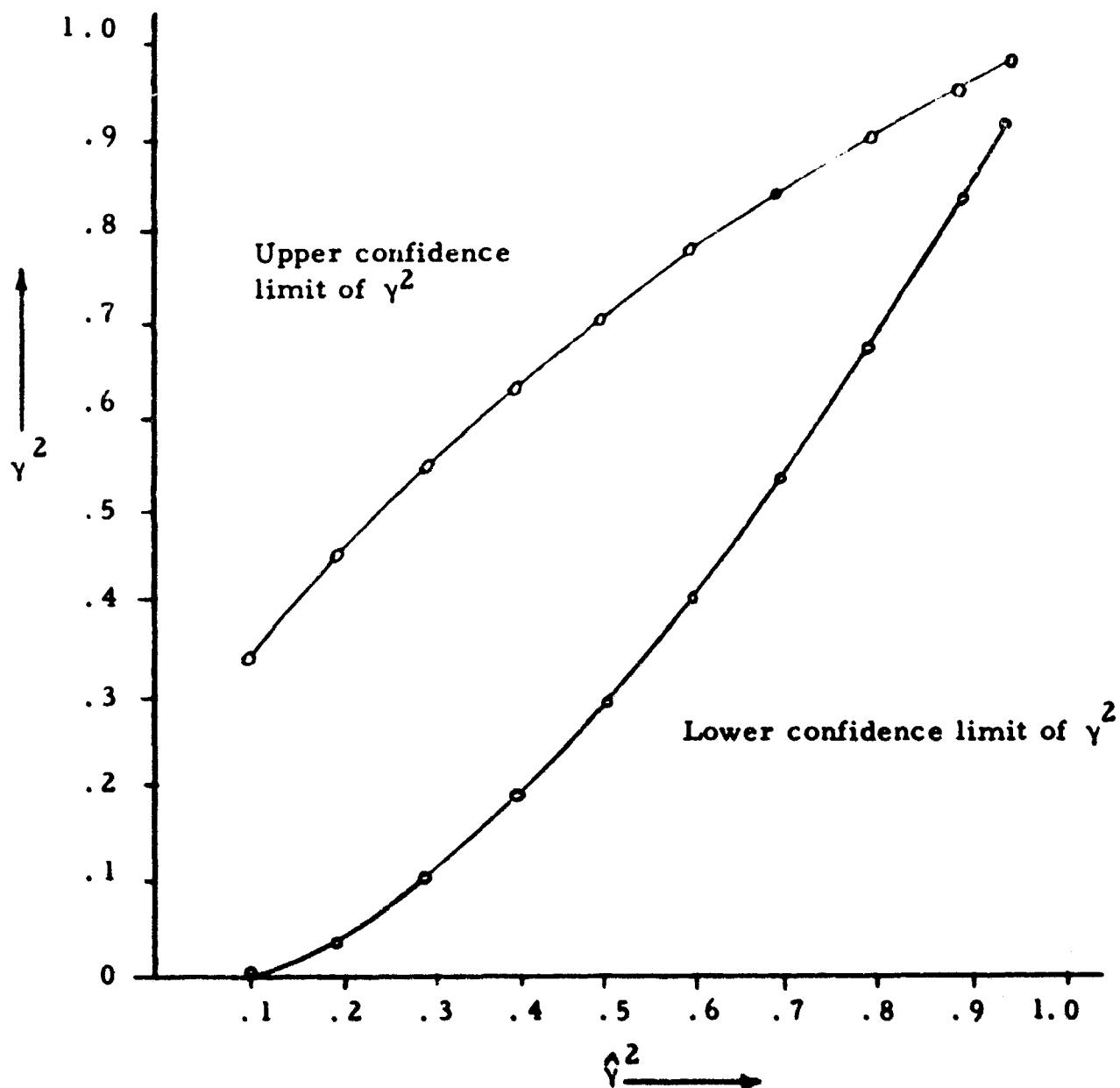


Figure 9. 99% Confidence Limits for Ordinary and Partial Coherence;  $n = 50, p = 2$ . Obtained by Using Equation (18)

TABLE 3  
Confidence Limits for Ordinary and Partial Coherence

$\hat{Y}^2$	.2	.3	.4	.5	.6	.7	.8	.9
Upper 99% Limit	.31	.49	.58	.66	.74	.81	.87	.94
Lower 99% Limit	.04	.11	.19	.29	.41	.54	.67	.83



**Figure 10.** Exact 99% Confidence Limits for Ordinary and Partial Coherence;  
 $n = 21$ ,  $p = 2$  Obtained from Tabulated Distribution Function of  
 Reference 1

## 6. CONCLUSIONS AND RECOMMENDATIONS

A useful normalizing transformation for sample multiple coherence has been empirically developed. The computational effort required to perform the transformation is reasonable and allows the use of Gaussian probability distribution tables.

Several aspects of the transformation are not completely resolved by this empirical approach. The transformed variable is biased. The numerical behavior of this bias has to a certain extent been determined as a function of degrees-of-freedom  $n$  and number of variables  $p$ . This behavior can be approximated in a manner which is satisfactory for multiple coherence roughly between 0.40 and 0.95. However, unexplained numerical behavior of this bias occurs when multiple coherence is close to 1.0. Also, as multiple coherence approaches zero, the bias seems to exhibit singular behavior which is more or less expected from theoretical considerations.

The variance of the transformed variable was empirically developed as a result of this numerical study. This experimentally determined variance is useful for about the same ranges of the parameter values as for the bias.

A theoretical study to attempt to determine precise analytical behavior of various transformations of sample multiple coherence would be a fruitful undertaking. A study of this sort would allow more confidence to be placed in the reliability of transformed values. The functional forms of the bias and variance could be determined to extend the usefulness of transformations to a wider range of parameter values.

The multiple coherence function is closely related to the multiple correlation coefficient of classical statistics. Hence, it is expected that a detailed theoretical study would parallel that performed by Hotelling for the correlation coefficient (Reference 4). There seem to be many complications which enter into an examination of coherence that do not occur for the correlation coefficient. However, the results of such a study for coherence would provide useful information.

## REFERENCES

1. Alexander, M.J. and C.A. Vok, "Tables of the Cumulative Distribution of Sample Multiple Coherence," Rocketdyne Division, North American Aviation, Inc., Research Report 63-67. November 1963.
2. Ibbetson, D., "Algorithm 209, GAUSS" Communications of the ACM, vol. 6, No. 10, October 1963, p. 616.
3. Goodman, N.R., "Measurement of Matrix Frequency Response Functions and Multiple Coherence Functions," AFFDL-TR-65-56, May 1965.
4. Hotelling, Harold, "New Light on the Correlation Coefficient and its Transforms," Journal of Royal Statistical Society, Series B, vol. XV, pp. 193-232. (1953).

UNC CLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Measurement Analysis Corporation 10962 Santa Monica Boulevard Los Angeles, California 90025	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
	2b. GROUP N/A

3. REPORT TITLE

Gaussian Approximations to the Distribution of Sample Coherence

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Report (one of several to be issued under this contract)

5. AUTHOR(S) (Last name, first name, initial)

Enochson, Loren D.  
Goodman, N. R.

6. REPORT DATE June 1965	7a. TOTAL NO. OF PAGES 27	7b. NO. OF REFS 4
8a. CONTRACT OR GRANT NO AF33(615)-1418	9a. ORIGINATOR'S REPORT NUMBER(S) AFFDL-TR-65-57	
8b. PROJECT NO 1370	8b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) MAC 403-07	

10. AVAILABILITY/LIMITATION NOTICES

No Limitation

11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY AirForce Flight Dynamics Laboratory (FDD), Wright-Patterson AFB, Ohio 45433
---------------------------------	--

13. ABSTRACT

This report describes the results of an empirical study to develop a normalizing transformation for sample multiple coherence. The "Fisher z-transformation" is employed. The expected value (including a bias term) and variance of the transformation have been experimentally determined. Numerical values of the transformation which is developed (including the bias term and variance) may be obtained with a reasonable amount of computation. Tables of the Gaussian distribution can then be used to obtain confidence limits and perform statistical tests.

The computational methods and the digital computer program used for the study are described in detail. Flow charts of the program are given. Numerical results from the program results are presented. Examples of the use of the transformation are given for developing confidence limits for multiple coherence. A recommendation for a further theoretical study is presented.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

## UNCLASSIFIED

## Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Random process theory Statistics Correlation analysis Coherence Spectral analysis Linear systems Transformation of random variables						
<b>INSTRUCTIONS</b>						
1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.	imposed by security classification, using standard statements such as:					
2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.	(1) "Qualified requesters may obtain copies of this report from DDC."					
2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.	(2) "Foreign announcement and dissemination of this report by DDC is not authorized."					
3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.	(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."					
4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.	(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."					
5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.	(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."					
6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.	If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.					
7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.	11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.					
7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.	12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.					
8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.	13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.					
8b. & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.	It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).					
9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.	There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.					
9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).	14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, miles, and weights is optional.					
10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those						

UNCLASSIFIED

Security Classification