1975

An Evaluation of Bartlett's Chi-Square Approximation for the Determinant of a Matrix of Sample Zero-Order Correlation Coefficients

Stephen M. Hattori
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports
Part of the Applied Statistics Commons, and the Mathematics Commons

Recommended Citation
https://digitalcommons.usu.edu/gradreports/1210

This Report is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact dylan.burns@usu.edu.
AN EVALUATION OF BARTLETT' S CHI-SQUARE APPROXIMATION FOR THE
DETERMINANT OF A MATRIX OF SAMPLE ZERO-ORDER CORRELATION
COEFFICIENTS

by

Stephen M. Hattori

A report submitted in partial fulfillment of the
requirements for the degree

of

MASTER OF SCIENCE

in

Applied Statistics

Plan B

UTAH STATE UNIVERSITY
Logan, Utah
1975
ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to Professor Ronald V. Canfield, my major professor, for the enormous amount of time and guidance extended toward the organization of this paper. His assistance and encouragement were instrumental toward the completion of this paper.

Thanks are also extended to Dr. Rex L. Hurst and Professor Elwin G. Eastman, members of my graduate committee, for their assistance and guidance.

Appreciation is also extended to Dr. Bartell C. Jensen for his suggestions and insight into the problem during the initial research and application of the problem.

I would also like to extend my sincere thanks to my parents for their patience and understanding as I completed my graduate study at Utah State University.

Stephen M. Hattori
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .............................................. i
LIST OF TABLES ................................................ iv
LIST OF FIGURES ............................................... v

Chapter

I. INTRODUCTION ......................................... 1
   1.1 Origin and nature of the problem
   1.2 Objective
   1.3 Approach

II. MOMENTS ................................................. 7
   2.1 Background
   2.2 Application in the problem

III. MATHEMATICAL APPROACH .......................... 10
   3.1 Initial Formulas
   3.2 Mathematical Manipulation
       Wilks' formula
       Bartlett's formula
   3.3 Derived Formulas

IV. COMPUTER APPLICATION ........................... 18
   4.1 Computer program
   4.2 Program usage
   4.3 Program output

V. RESULTS .................................................. 23
   5.1 Computed results
   5.2 Interpretation of results

VI. CONCLUSION ........................................... 27
   6.1 Summary

LITERATURE CITED ......................................... 28
APPENDIX ................................................... 29
VITA ......................................................... 34
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1st moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^-5)</td>
<td>30</td>
</tr>
<tr>
<td>2.</td>
<td>2nd moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^-5)</td>
<td>31</td>
</tr>
<tr>
<td>3.</td>
<td>3rd moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^-5)</td>
<td>32</td>
</tr>
<tr>
<td>4.</td>
<td>4th moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^-5)</td>
<td>33</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Computer program for the computation of the moments.</td>
<td>19</td>
</tr>
<tr>
<td>2.</td>
<td>Sample of the output.</td>
<td>23</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.1 Origin and nature of the problem

The single equation least-squares regression model has been extensively studied by economists and statisticians alike in order to determine the problems which arise when particular assumptions are violated. Much literature is available in terms of the properties and limitations of the model. However, on the multicollinearity problem, there has been little research, and consequently, limited literature is available when the problem is encountered. Farrar & Glauber (1967) present a collection of techniques to use in order to detect or diagnose the occurrence of multicollinearity within a regression analysis. They attempt to define multicollinearity in terms of departures from a hypothesized statistical condition, and then fashion a series of hierarchical measures for its presence, severity, and location in a data set. Since the problem is of a statistical rather than of a mathematical nature, the question of existence or nonexistence is ignored and the focus is on the severity of the problem.

Regression analysis involves the estimation of the parameters of a dependency relationship. A regression model is developed in order to predict a dependent variable, $Y$, from a set of independent variables $X$. The regression model is of the form

$$Y = X \beta + E$$

(1.1.1)

where $\beta$ is a vector of true (structural) coefficients and $E$ is a true (unobserved) error term. The sum of the squared residuals can be found using the following formula:

$$\sum_{i=1}^{n} e_i^2 = E^t E$$

$$= (Y - \beta X)^t (Y - \beta X)$$

$$= Y^t Y - 2 \beta^t X^t Y + \beta^t X^t X \beta \ (1.1.2)$$

since $\beta^t X^t Y$ is a scalar and equal to its transpose $Y^t X \beta$. 

To find $\beta$ which minimizes the squared residual total, we differentiate (1.1.2).

$$\frac{\partial}{\partial \beta} \left( E^t E \right) = -2 X^t Y + 2 X^t X \beta.$$  \hspace{1cm} (1.1.3)

Equating (1.1.3) to zero, we get

$$X^t X \beta = X^t Y$$  \hspace{1cm} (1.1.4)

Let us assume that the number of variables is less than the number of observations on $Y$, then least square regression analysis leads to the estimates

$$\beta' = (X^t X)^{-1} X^t Y$$  \hspace{1cm} (1.1.5)

with variance-covariance matrix

$$V(\beta') = \sigma_E^2 (X^t X)^{-1}$$  \hspace{1cm} (1.1.6)

where $\sigma_E^2$ is the population variance of $E$.

However, multicollinearity in regression arises when an interdependency condition exists within the set of independent variables. This condition threatens the proper specification and effective estimation of the regression model sought.
The difficulties in estimating the parameters are dependent upon the severity of the problem. As interdependency increases within the independent variables, the correlation matrix, \((X^t X)\), approaches singularity and the inverse elements of the correlation matrix explode. As perfect singularity is achieved, the set of parameter estimates are completely indeterminant. This means that the diagonal elements of the inverse matrix, \((X^t X)^{-1}\), corresponding to the linearly dependent members of \(X\), become infinite. Thus, when regression techniques are employed on multicollinear, independent variables, the parameter estimates are markedly sensitive to changes in model specification and to sample coverage.

Model specification is also affected by multicollinearity though it is less dramatic and less easily detected. Correct model specification is more important than a correct estimating procedure. Correct specification entails recognition of all the variables in the problem in order to explain the behavior of a given dependent variable. As interdependency within the independent variable set, \(X\), increases, the stability and sample significance of each independent variable's contribution decreases. This data limitation leads to underspecification of the model. Thus all the variables necessary to define the model entirely are not included. Reliance upon the data to determine if the model is complete is dangerous since the data samples cover only a limited range of experience. Thus the process becomes one of modifying the hypothesis through trial and error until an acceptable equation is generated.

With the development of high speed digital computers, the programmer's approach to singularity in regression analysis has been applied to multicollinearity by the economist. In order to estimate the parameters for a standard regression equation, the inverse, \((X^t X)^{-1}\), is required. If the set of independent variables, \(X\), is multicollinear, \((X^t X)\) is singular, and the determinant is non-existant.

\[
(X^t X)^{-1} = \frac{1}{\text{Adj} (X^t X)} \frac{1}{|X^t X|} \quad (1.1.7)
\]
To test for a singular matrix, the determinant is zero. Thus, if the determinant, $|X^t X|$, approaches zero, $X$ approaches singularity. If $|X^t X|$ approaches one, $X$ is nearly an orthogonal independent variable set. However, the gradient between the limits,

$$0 \leq |X^t X| \leq 1,$$

(1.1.8)

is not well defined.

Now that we are able to detect the occurrence of multicollinearity, is it possible to determine its location within the data set? Near singularity can be due to pairwise correlation or a more complex linkage among several independent variables. One attempt has been to indicate each explanatory variable's dependence on other members of the independent set. Let us define $r_{ii}$ as the diagonal element of the inverse $(X^t X)^{-1}$ corresponding to the $i$th variable. Then we have

$$r_{ii} = \frac{|(X^t X)_{ii}|}{|X^t X|}$$

(1.1.9)

where $(X^t X)_{ii}$ denotes the correlation matrix excluding the $i$th variable, $x_i$. If $x_i$ is orthogonal to the other members of $X$, $r_{ii} = 1$. If $x_i$ is dependent on the other members of $X$, $r_{ii} = \infty$ since the denominator vanishes. Thus the location of the singularity in $X$ is found.
Farrar & Glauber define multicollinearity in terms of departures from orthogonality in an independent variable set. There are two advantages with this definition. One, it clearly distinguishes between the problem’s essential nature; the lack of independence or the interdependency in the independent set and the effect on the dependent relationship produced. The second is that orthogonality easily lends itself to statistical hypothesis formulation and then to development of test statistics against which the severity of the departures can be calibrated against. Such statistics developed sufficiently in detail can provide insight into the location and pattern as well as the severity of the multicollinearity within the data set. Let us assume that X, the independent variable set, is also multivariate normal since existing distributional theory is based almost entirely upon this assumption. Thus if we can attach distributional properties to the determinant \( |X^t X| \) or a transformation of \( |X^t X| \), then the statistic resulting could be quite useful in measuring the presence and severity of multicollinearity within an independent variable set since the heuristic relationship between orthogonality and the determinant \( |X^t X| \) has been previously discussed.

1.2 Objective

During the course of their presentation, Farrar and Glauber introduced two distribution functions involving \( |X^t X| \). The first was derived by S. S. Wilks (Wilks, 1931). He was able to obtain the moments and distribution of determinants for sample correlation matrices \( |X^t X| \). The second was derived by M. S. Bartlett. He was able to obtain a transformation from Wilks' distribution function that is distributed approximately Chi-Square. A more detailed discussion follows in Chapter III. The objective is to determine if Wilks' distribution function for the moments of \( |X^t X| \) and Bartlett's Chi-Square approximation for \( |X^t X| \) are asymptotically equivalent.
1.3 Approach

In order to achieve the objective, a process of three steps is followed. The first step is to mathematically reduce the two distributional functions into more easily handled forms. The second step is to compute the moments for each distribution function for comparison. The third step is to interpret the results and document the findings. In the course of the initial research, it was determined that the mathematics would revolve around a comparison of the moments of the distribution functions under test conditions. The difference would be calculated and compared against some tolerance value. This tolerance value is an arbitrary value which is acceptable in order to claim the moments are approximately equal in value. Also a correlation coefficient would be determined between the two sets of moments. A computer is used to compute the moments since it is less time consuming and more conditions can be evaluated. For limiting conditions, the computation were for 1-10 moments, 2-20 independent variables, and sample sizes of 50, 100, 500, 1000, 2000, 3000, 4000, 5000, and 9000.
2.1 Background

The problem is often encountered in statistical analysis where it is difficult or not feasible to determine completely the cumulative distribution function of a random variable. In many cases, moments and certain functions of moments of the random variable can be used to describe the distribution of the random variable. Or maybe the interest is in certain functions of the population parameters rather than in the parameters themselves. Again, moments of the distribution may often be used for the solution. Given an observed frequency distribution, it may be desirable to measure the nature and amount of departure from normal. Again, we are interested in moments.

Let \( X \) denote a random variable, a function whose domain is a sample space, \( S \), and whose range is a set of real numbers. The 1st moment about the origin is the expected value of \( X \). The \( k \)th moment, \( \mu_k \), about the origin is the expected value of \( X^k \) (Ostle, 1954).

For \( X \) continuous,
\[
\mu_k = \int_{-\infty}^{\infty} X^k f(X) \, dX = \text{E}(X^k). \tag{2.1.1}
\]

For \( X \) discrete,
\[
\mu_k = \sum_{i=1}^{N} X_i^k f(X_i). \tag{2.1.2}
\]

However, for our problem, we are only interested in \( X \) continuous. There are innumerable possible moments depending upon the distribution, but the first four moments are of primary interest. The 1st moment about the origin is the mean of the random variable \( X \). The 2nd moment about the mean is the variance of \( X \). The square root of the 2nd moment is the standard deviation of \( X \). The 3rd moment about the mean indicates skewness. The 4th moment about the mean indicates kurtosis.
Earlier, we mentioned a situation where a distribution deviates from normality. To measure the nature and amount of departure, we emphasize two indicators, skewness and kurtosis. Skewness or asymmetry indicates that one tail of the curve is drawn out more than the other. The mean and the median of the curve do not coincide. A negative valued 3rd moment indicates skewness to the left. A positive value indicates skewness to the right.

Kurtosis or peakedness of a curve is subdivided into two sections, leptokurtic and platykurtic. A leptokurtic curve has more sample points near the mean and at the tails with fewer sample points in between relative to a normal curve. A positive valued 4th moment indicates this situation of leptokurtosis. A platykurtic curve has more of the sample points in between and fewer at the mean and the tails. A negative valued 4th moment indicates this situation of platykurtosis. A bimodal distribution is an extreme platykurtic distribution (Sokal & Rohlf, 1969).

If we become interested in the nomenclature of elementary mechanics, the mean of $X$, 1st moment about the origin, can be interpreted as the center of gravity of $R_1$. The variance of $X$, 2nd moment about the mean, can be interpreted as the moment of inertia about the center of gravity or an indication of the amount by which the probability mass spreads or concentrates about the center of gravity (Wilks, 1962).

2.2 Application in the problem

Moments or the application of them are the keys to attaining the objective of this paper. Moments help to describe the form a distribution function takes. Thus, we can compare the distribution functions through their moments and see what occurs. Equivalence can be tested in this manner since,

"if $X$ is non-negative and integer valued and all its moments exist, then the distribution of $X$ is completely determined by the moments. In other words, if two distributions over $0, 1, 2, \ldots$ have the same moments and all moments are finite, then the distributions are identical. \ldots"  

(Dwass, 1970).
Now, all we have to do is to determine the moments of each distribution function under each set of conditions and note the difference between the corresponding moments. If the difference is less than some tolerance value, we can assume the moments are equivalent and that the distribution functions being analyzed are equivalent.
CHAPTER III
MATHEMATICAL APPROACH

3.1 Initial Formulas

Before we delve mathematically into this multicollinearity problem, let us first define some variable names we will encounter:

N is the sample size.
m is the number of independent variables.
k is the order of the moment.
\(X\) is the matrix of N observations on one dependent and m independent variables, each of which is normalized (by sample size and standard deviation) to unit length.

\(X^t X\) is a zero order correlation matrix, where \(X^t\) is the transpose of X.

Now, let us ponder over the two distribution functions of interest which Wilks and Bartlett have generated.

While S. S. Wilks was involved in some work concerning analysis of variance, he encountered the work of J. Wishart (Wishart, 1928) who had shown that the simultaneous distribution of the variances and covariances of a sample of N items from a m-variate normal population was according to a specific frequency distribution which bears his name, the Wishart distribution (Wilks, 1967). Wilks was able to derive from the Wishart distribution, in an analytic tour de force, the moments and distribution (in open form) of the determinant of sample covariance matrices (Wilks, 1932). Subsequent introduction of the assumption of orthogonality enabled him to further extend his work until he was able to obtain the moments and distribution of the determinants for sample correlation matrices, \(\left| X^t X \right|\), as well. The formula that was derived for the kth moment of \(\left| X^t X \right|\) is as follows.
Theoretically, it should be possible to derive the density function for $|X^t X|$ from (3.1.1), especially since it is in open form. However, for $n>2$, explicit solutions have not been obtained.

However, M. S. Bartlett, through comparison with the lower moments of (3.1.1) and with those of the Chi-Square distribution was able to obtain a transformation of $|X^t X|$,\[ x^2 \left( \frac{\gamma}{|X^t X|} \right) = - \left[ N-1 - \frac{1}{6} \left( 2m + 5 \right) \right] \log(|X^t X|) \] (3.1.2)

that is distributed approximately as a Chi-Square with $\gamma = m (m-1)/2$ degrees of freedom (Bartlett, 1950).

### 3.2 Mathematical Manipulation

The two distribution functions, (3.1.1) and (3.1.2), outwardly appear to be a little too complex for easy generation of data. So let us put into operation the first step in achieving our objective, reduction of the distribution functions to simpler forms.

#### Wilks' formula

The distribution function (3.1.1) contains, in the numerator and denominator, Gamma functions which are quite complex in themselves for solution. But the Gamma function under certain conditions can be replaced by the use of factorials. The value of the Gamma function to be found can be divided into two group types. When $N$ is an odd value, the value is an integer, and the factorial form can be used in place of the Gamma. When $N$ is an even
number, the value is a multiple of 1/2. However, the use of factorials for values that are multiples of 1/2 is feasible until the end or \( \Gamma(1/2) \). However, 
\[ \Gamma(1/2) = \sqrt{\Pi} \]. Thus, the factorial form can again be used in place of the Gamma. Therefore, using the algebraic relationship, \( \Gamma(n) = (n-1)! \), the distribution function (3.1.1) can be rewritten as

\[
M_k (|X^t X|) = \frac{\left[\left(\frac{N-1}{2} - 1\right)\right]^{m-1} \prod_{i=2}^{m} \left(\frac{N-i}{2} + k-1\right)!}{\left[\left(\frac{N-1}{2} + k-1\right)\right]^{m-1} \prod_{i=2}^{m} \left(\frac{N-i}{2} - 1\right)!} \quad (3.2.1)
\]

Continuing our mathematics, we have two algebraic relationships which are applicable.

\[
\frac{S^a}{T^a} = \left(\frac{S}{T}\right)^a \quad \text{and} \quad \frac{\prod_{i=n}^{m} S_i}{\prod_{i=n}^{m} T_i} = \frac{m}{\Pi} \frac{S_i}{T_i} \quad (3.2.2)
\]

Applying these ideas to (3.2.1) and combining the terms within each factorial, we now have the form

\[
M_k (|X^t X|) = \frac{\left[\left(\frac{N-3}{2} \right)!\right]^{m-1} \prod_{i=2}^{m} \left(\frac{N-i-2}{2} + k\right)!}{\left[\left(\frac{N-3}{2} + k\right)!\right]^{m-1} \prod_{i=2}^{m} \left(\frac{N-i-2}{2}\right)!} \quad (3.2.2)
\]

The formulas for the 1st through 4th moments can be readily generated, but it seems to be relevant to generate the formula for the kth moment, the general
case. However, the kth moment of \( \int X^t X \) has the same form as distribution function (3.2.2), so let us use it. Notice that by using the definition of factorials, we can produce the following relation,

\[
( z+k)! = (z-k+1)(z-k+2) \cdots (z+2)(z+1)z!
\]  

(3.2.3)

but notice further as we rearrange the terms, (3.2.3) becomes

\[
\frac{z!}{(z+k)!} = \frac{1}{(z+k)(z+k-1)(z+k-2) \cdots (z+2)(z+1)}
\]

(3.2.4)

However compare the front fraction of equation (3.2.2) and notice that it is of the same form as (3.2.4). Thus, taking it separately, we change the front of (3.2.4) to the form

\[
\left(\frac{N-3}{2}\right)!
\]

\[
\frac{1}{(N-3-k)(N-3-k-1)(N-3-k-2) \cdots (N-3-k-2)(N-3-k-1)(N-3-k)}
\]

(3.2.5)

but there are k quantities in the denominator and we can factor out k values of 2 and place it in the numerator.

\[
\frac{2^k}{(N-3+2k)(N-5+2k)(N-7+2k) \cdots (N+1)(N-1)}
\]

(3.2.5)

Similarly, if we apply the reciprocal form of (3.2.4) to the back fraction of (3.2.2) and using the same method as in (3.2.5), we get

\[
\left(\frac{N-1-2}{2}+k\right)! = \left(\frac{1}{2}\right)^k(N-1-2+2k)(N-1-4+2k) \cdots (N-1+2)(N-1)
\]

(3.2.6)

Now substituting (3.2.5) and (3.2.6) into the appropriate section of (3.2.2), we now have the form
After close examination, it becomes evident that there are \( k ( m - 1 ) \) 2's in the numerator and a like number in the denominator, so we can cancel them out. Note also that there are \( ( m - 1 ) \) terms in the coefficient fraction and a like number in the product fraction. Combining these, (3.2.7) now becomes

\[
M_k \left( |X^t X| \right) = \frac{2^k}{m \prod_{i=2}^{m} \frac{(N-i-2+2k)(N-i-4+2k)\ldots(N-i+2)(N-i)}{(N-3+2k)(N-5+2k)\ldots(N+1)(N-1)}}
\]

Closely investigate the fraction in (3.2.8). It can be broken down into \( k \) smaller fractions which then can be rewritten in a product form.

\[
\prod_{j=1}^{k} \frac{(N-i+2j-2)}{(N+2j-3)}
\]

Substituting (3.2.9) for the fraction in (3.2.8), we have

\[
M_k \left( |X^t X| \right) = \prod_{i=2}^{m} \prod_{j=1}^{k} \frac{(N-i+2j-2)}{(N+2j-3)}
\]

The distribution function (3.2.10) was as far as I was able to reduce it. This will be the formula used for evaluating moments.

**Bartlett's formula**

Now, let us shift our attention to the second of our distribution functions. M. S. Bartlett was able to derive a transformation for \( |X^t X| \) that approximates a Chi-Square distribution with \( \gamma = m ( m - 1 ) / 2 \) degrees of freedom. However, recall that we are going to compare each distribution function through their moments in order to show equivalence. The form (3.2.10) does generate moments for us. But Bartlett’s formula in its present state does not. What is necessary is to substitute into a density
function in order to produce moments. Since (3.2.2) does approximate a Chi-Square distribution, we use the density function from the Chi-Square with \( \gamma \) degrees of freedom (CRC, 1965).

\[
f(x^2) = \frac{\chi^2 \gamma^{-2}}{2 \gamma/2} \frac{e^{-x^2/2}}{\Gamma(\gamma/2)}, \quad 0 < x^2 < \infty \tag{3.2.11}
\]

We use the moment-generating function for a random variable, say \( Y \), (Brunk, 1965) which is

\[
M_Y(t) = E(e^{Yt}).
\]

If \( Y \) is discrete with probability function \( f(y) = \Pr(Y=y) \),

\[
M_Y(t) = E(e^{Yt}) = \sum_y e^{Yt} f(y).
\]

If \( Y \) is continuous with density function \( f(y) \),

\[
M_Y(t) = \int_{-\infty}^{\infty} e^{Yt} f(y) dy. \tag{3.2.12}
\]

Since the random variable we are interested in is continuous, the latter form (3.2.12) is used.

For convenience sake, let \( a = (N-1-(2m+5)/6) \) and substitute into the distribution function (3.2.2) which becomes

\[
x^2 = -a \log(\left|X^t X\right|), \tag{3.2.13}
\]

divide both sides by \(-a\)

\[
\log(\left|X^t X\right|) = -x^2/a \tag{3.2.14}
\]

and take the exponential value of both sides.

\[
\left|X^t X\right| = e^{-x^2/a} \tag{3.2.15}
\]

We now evaluate the moments of \( \left|X^t X\right| \) or equivalently \( e^{-x^2/a} \). For the kth moment of a random variable \( X \), we use \( E(e^{-x^2/a}) \), where \( k = 1, 2, 3, \ldots \) assuming the expected value exists (i.e. \( E(e^{-x^2/a})<\infty \)). Therefore, the kth moment has the form

\[
E(\left|X^t X\right|) = \frac{1}{\gamma/2} \frac{\Gamma(\gamma/2)}{\Gamma(\gamma/2)} \int_0^\infty \frac{X^{k/a}}{e^{-X/2}} \frac{X^{(\gamma-2)/2}}{e^{-X/2}} dX \tag{3.2.16}
\]
But we have an algebraic relationship that states
\[ e^A \cdot e^B = e^{(A + B)} \]
and so applying it to (3.2.16)
\[ E(\begin{bmatrix} X^T X \end{bmatrix}) = \frac{1}{2^{\gamma/2}} \int_0^\infty X^{(\gamma-2)/2} e^{-(k/a + 1/2) X} \, dX. \]  

(3.2.17)

Let us again simplify our equation by allowing
\[ Y/2 = (k/a + 1/2) \]
and then manipulating it further to give
\[ X = \frac{Y}{2 (k/a + 1/2)} = \frac{Y}{2 k/a + 1} \]
which if one takes the derivative of both sides ends up with
\[ dX = (k/a + 1/2)^{-1} \left(1/2\right) \, dY. \]  

(3.2.19)

Substituting (3.2.18) and (3.2.19) into (3.2.17), we now have
\[ E(\begin{bmatrix} X^T X \end{bmatrix}) = \frac{(2 k/a + 1)^{-\gamma/2}}{\Gamma(\gamma/2)} \int_0^\infty e^{-Y/2} Y^{(\gamma-2)/2} \, dY. \]  

(3.2.20)

Turning to our table of definite integrals (CRC, 1965), we find the following form to apply to (3.2.20).
\[ \int_0^\infty X^h e^{-bX} \, dX = \frac{\Gamma(h+1)}{b^{h+1}} \text{ for } h > -1, \ b > 0. \]  

(3.2.21)

However, in our case, we can use either form since this is the same problem we encountered with Wilks' formula when we looked at the value of the Gamma function to be found. For convenience sake, let us use (3.2.21) and substitute (3.2.20) into it.
Our resulting equation is then

\[
E ( \left| X^t X \right| ) = \frac{(2k/a + 1)^{-\gamma/2}}{2^{\gamma/2} \Gamma(\gamma/2) (1/2)^{\gamma/2}} \Gamma(\gamma/2) (\gamma/2) \] (3.2.23)

Finally, let us cancel out the Gamma values and substitute the following values of \(a\) and \(\gamma\) into (3.2.23).

\[
E ( \left| X^t X \right| ) = \left( \frac{2k}{6N - 2m - 11} + 1 \right)^{-\frac{1}{4}m(m-1)} \] (3.2.24)

3.3 Derived Formulas

After much examination and deliberation, the forms (3.2.10) and (3.2.24) are the most manipulable forms we could generate. Outwardly, they do not appear to be equivalent, and there is no easy mathematical method to show equivalence, off hand. Therefore, we undertake the second step in our process by using (3.2.10) and (3.2.24) in a computer program and compute the two sets of moments under various conditions in order to see if there is actually an equivalence between them.
CHAPTER IV
COMPUTER APPLICATION

4.1 Computer program

The program to be described in this chapter was written to eliminate the tedious amount of computations required for the comparison of the two distribution functions for equivalence. With the use of this program, more conditions could be evaluated than if the work was done by hand. The program is written in Fortran IV and was run on an IBM-360 Model 44.

Essentially, the program is broken into two segments. The first segment reads in the necessary parameters and sets up the appropriate headings for output. Input will be further discussed in Section (4.2) and output in Section (4.3). The second segment, using the input parameters, computes the moments for Wilks' and Bartlett's distribution functions and the difference between the two. As the ten moments for a specific number of independent variables are computed, they are printed and then the next group is computed.

Since accuracy was desired to $10^{-8}$, the program values and computations were all in double precision. Initially, the program used single precision. But when certain sections were done by hand to check the results, there were some discrepancies which were cleared up satisfactorily by switching to double precision. A note of caution. The program is written in such a manner as to accept only one set of parameters per run. This situation can be easily remedied by slight modification of the program. A program listing can be found in Figure 1.
IMPLICIT REAL*8 (A-H,O-Z)
DIMENSION EM(10),EV(10),DIFF(10)
DATA EM/10*1.0/
IRD=5
IPR=6
LUA=9
REWIND LUA
100 FORMAT (3I4)
READ (IRD,100) N,M,K
101 FORMAT ('1N=','I6,' FOR 2-','I2,'VARIABLES: '// 'VARIABLES MOMENT', '112X,'EV', '15X,'EM', '25X,'DIFF')
WRITE (IPR,101) N,M
AN=N
DO 20 I=2,M
AM=I
B=1.0
DO 10 J=1,K
AK=J
EV(J)=(((AK*12.0)/((6.0*AN)-AM-AM-11.0))+1.0)**(-AM*(AM-1.0)*0.25)
B=B*(AN-AM+J+J-2.0)/(AN+J+J-3.0)
EM(J)=EM(J)*B
10 DIFF(J)=EV(J)-EM(J)
102 FORMAT (' ',I5,Il0,F26.8,F17.8,F24.8)
WRITE (IPR,102) (I,J,EV(J),EM(J),DIFF(J),J=1,K)
103 FORMAT ('----------------')
WRITE (IPR,103)
104 FORMAT (F26.8,F17.8)
WRITE (LUA,104) (EV(J),EM(J),J=1,K)
20 CONTINUE
STOP
END

Figure 1. Computer program for the computation of the moments.
To show the equivalence or association of the two distribution functions, a mathematical or statistical procedure had to be found that was acceptable and relevant. In statistics, such a test is the calculation of the correlation coefficient. The correlation tests the degree of association between two groups of data. The coefficient varies in value from $-1$ to $+1$. The closer the value is to $-1$ or $+1$, the better is the association between the two groups. If the value is closer to 0, then there is very little association. Available on the Utah State University computer library is a set of statistical programs written by Dr. Rex Hurst, Head of the Department of Applied Statistics - Computer Science. Among these statistical programs is one for computing the correlation coefficient for groups of data. The author's program was run in conjunction with this correlation program. The entire submitted program consists of three sections. The first section sets up the necessary files and storage needed for the correlation test. The second section contains the author's program for the computation of the moments and their difference, output of the results, and storage on tape of the results. The third section reads the tape and runs the correlation test for the two sets of moments. The results will be discussed in Chapter V.

4.2 Program usage

In order to use the program, three pieces of information need to be provided: the sample size, the number of independent variables, and the number of moments to be calculated. Once these parameters have been determined, they must be read into the program in a certain format.

The format is as follows:

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1-4)</td>
<td>Sample Size, $N = (2-9999)$</td>
</tr>
<tr>
<td>(5-8)</td>
<td>Number of Independent Variables, $m = (2-9999)$</td>
</tr>
<tr>
<td>(9-12)</td>
<td>Number of Moments, $k = (1-9999)$ if they exist.</td>
</tr>
</tbody>
</table>
Remember, the program as listed will not accept more than one input card per run. The program must be modified to accept more.

4.3 Program output

The first heading of the output indicates the sample size and the maximum number of independent variables used in the computations. The next set of headings indicates the number of independent variables in use, the order of the moment, Bartlett's value, Wilks' value, and the difference between the two. After ten moments are calculated for a specific number of independent variables, the data set is separated by a broken line from the next set. EV is used to indicate the column containing those moments from Bartlett's distribution function. EM refers to those moments from Wilks' distribution function. An example of the output can be found in Figure 2.

When the ten moments for a specific number of independent variables are computed, they are printed as a set. The next set of moments for the next number of independent variables is computed and printed. This is repeated until all the sets of moments have been printed.
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>MOMENT</th>
<th>EV</th>
<th>EM</th>
<th>DIFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0.97958969</td>
<td>0.97959184</td>
<td>-0.00000215</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.96038018</td>
<td>0.96038415</td>
<td>-0.00000397</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.94225817</td>
<td>0.94226370</td>
<td>-0.00000552</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.92512478</td>
<td>0.92513163</td>
<td>-0.00000685</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.90889326</td>
<td>0.90890125</td>
<td>-0.00000799</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>0.89348717</td>
<td>0.89349615</td>
<td>-0.00000897</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>0.87883885</td>
<td>0.87884867</td>
<td>-0.00000982</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>0.86488814</td>
<td>0.86489869</td>
<td>-0.00001055</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>0.85158138</td>
<td>0.85159255</td>
<td>-0.00001118</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0.83887049</td>
<td>0.83888222</td>
<td>-0.00001173</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.93960783</td>
<td>0.93960850</td>
<td>-0.00000066</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.88505873</td>
<td>0.88505991</td>
<td>-0.00000118</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.83559076</td>
<td>0.83559234</td>
<td>-0.00000158</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.79056514</td>
<td>0.79056703</td>
<td>-0.00000189</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>0.74944277</td>
<td>0.74944489</td>
<td>-0.00000212</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>0.71176581</td>
<td>0.71176812</td>
<td>-0.00000230</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>0.67714326</td>
<td>0.67714569</td>
<td>-0.00000244</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>0.64523934</td>
<td>0.64524188</td>
<td>-0.00000254</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>0.61576432</td>
<td>0.61576692</td>
<td>-0.00000261</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>0.58846696</td>
<td>0.58846962</td>
<td>-0.00000266</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.88209650</td>
<td>0.88208145</td>
<td>0.00001505</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.78202285</td>
<td>0.78199771</td>
<td>0.00002514</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.69653216</td>
<td>0.69650044</td>
<td>0.00003172</td>
</tr>
</tbody>
</table>

Figure 2. Sample of the output.
5.1 Computed results

The objective of this paper was to show that the two distribution functions, one by Wilks and the other by Bartlett, were asymptotically equivalent. The first attempt was to mathematically manipulate the formulas for the distribution functions to show conclusively that they were indeed equivalent. However, the resulting equations indicated nothing of this nature. It became apparent that we would need to compute values from these functions that would indicate that they were asymptotically equivalent. Using the reduced formulas generated by our initial attempt of solution, it was decided to compute the first four moments for each formula. However, since the varying parameters were the sample size and the number of independent variables, it became immediately apparent that much work was needed to generate enough data for conclusive proof. Since this is the age of machines, a computer program was written to generate this data needed. Using the same parameter values, moments were generated from both functions and the difference was computed. If it could be shown that the differences were less than an acceptable tolerance value, then it would indicate that the functions were indeed asymptotically equivalent. The program was written with some limiting factors as has been previously indicated. Also, for each of the nine sample sizes, a correlation coefficient was calculated using the two sets of generated moments as data. The values were roughly $0.99999\ldots$ to $1.0$. Since the interest is the difference in moment values, tables were constructed to best illustrate the aspects of the values needed for comparison. These four tables are located in the Appendix. Table 1 contains the 1st moment differences, Table 2 contains the 2nd moment differences, Table 3 contains the 3rd moment differences, and Table 4 contains the 4th moment differences. Each table is restricted for nine
sample sizes and (2-20) independent variables. The difference is calculated by taking the moment computed from Bartlett's distribution function and subtracting the moment computed from Wilks' distribution function from it.

Analyzing Table 1 in the Appendix, 1st moment differences, it is evident that when 2 and 3 independent variables are used regardless of sample size, Wilks' distribution function produces the larger moments. As we increase the sample size, we see that the difference becomes smaller for each number of independent variables except 2, 3, and 20. If we think of the difference as the absolute difference, then the cases of 2 and 3 independent variables follow the trend. For 20 independent variables, the differences in the samples of 50 are smaller than those in the samples of 100. They peak in the samples of 100 and then decrease. As we increase the number of independent variables, the difference increases within all the sample sizes except 50. For samples of 50, the difference increases in size until we have 14 independent variables. From then on, the difference decreases. The large difference values occur in the smaller sample sizes.

Looking at Table 2 in the Appendix, 2nd moment differences, we see that Wilks' distribution function produces the larger moments for 2 and 3 independent variables. Except for 2, 3, and 17-20, the difference decreases in size as we increase the sample size for a fixed independent variable. If we interpret the difference as being the absolute difference, then we can include the values for 2 and 3 independent variables in the above mentioned trend. At 17-20 independent variables, the difference is smaller in the samples of 50 than in the samples of 100. The peak is in the samples of 100 and decrease from there. If we increase the number of independent variables, the value of the difference increases in size in all cases except for the samples of 50 and 100. In these two cases, there is a peak point. For the samples of 50, the peak is at 10 independent variables. For the samples of 100, the peak is at 15 independent variables. Again, note that the larger differences are in the smaller sample sizes.
Turning to Table 3 in the Appendix, 3rd moment differences, the patterns that were established in the first two tables is repeated. For 2 and 3 independent variables regardless of sample sizes, Wilks' distribution function produces the larger moments. For a fixed number of independent variables, as we increase the sample size, the difference gets smaller except for 13-20 independent variables. For the samples of 50 and 13-20 independent variables, the differences are smaller than those for the samples of 100. Within each sample size, the difference increases in size as we increase the number of independent variables except for the samples of 50 and 100. For the samples of 50, the difference increases and peaks at 9 independent variables. For the samples of 100, the difference peaks at 12 independent variables. Again, note that the larger differences are found in the smaller sample sizes.

Looking at Table 4 in the Appendix, 4th moment differences, the pattern seen in the previous three tables is repeated. Wilks' distribution function produces larger moments for 2 and 3 independent variables. If we interpret the differences as the absolute difference, then we can include these two cases and say that the trend for a fixed number of independent variables and an increase in sample size is a decrease in difference except for 11-20 independent variables. For the samples of 500, the differences for 18-20 independent variables are larger than those in the samples of 50 and 100. The difference peaks in the samples of 500. For samples of 100, the difference peaks for 11-17 independent variables. Note again that the larger differences occur in the smaller sample sizes.

5.2 Interpretation of results

As we analyze the differences in the two sets of moments computed, several trends become obvious. It is evident that for 2 and 3 independent variables regardless of sample size, the moments computed from Wilks' distribution function are larger. The pattern also develops that as we increase sample size, the difference in moments decreases for a fixed number of
independent variables except for the larger numbers of independent variables and the smaller sample sizes. The difference is exceedingly small for the large sample sizes. Noteworthy is the fact that as we increase the number of independent variables, the difference increases within each sample size except in the area of larger number of moments and smaller sample sizes. Again for large samples, the difference gets exceedingly small. These patterns and trends indicated seem to lead to the ultimate evaluation that indeed the two distribution functions derived by Wilks and Bartlett are indeed asymptotically equivalent. All that we have accomplished is to prove that Bartlett did indeed obtain a transformation of $X^\top X$ that is distributed approximately Chi-Square with $\gamma = n(n-1)/2$ degrees of freedom by comparing the lower moments of Wilks' distribution function and those of the Chi-Square distribution.
CHAPTER VI
CONCLUSION

6.1 Summary

The primary objective of this study was to show that Wilks' distribution function for the moments of $|X^t X|$ and Bartlett's Chi-Square approximation for $|X^t X|$ were asymptotically equivalent. Unable to show this case mathematically, we undertook to compute the moments for each distribution function for comparison. We observed that as we increased the sample size, the difference between the moments decreased in size. As we increased the number of independent variables, the difference increased within each sample size. However, for large sample sizes and large numbers of independent variables, the differences were extremely small. The larger differences occur in the smaller sample sizes. These trends were general trends. We then concluded from the trends across the sample sizes and within each sample size that the two distribution functions were asymptotically equivalent.

Since we were able to show asymptotic equivalence, the determinant of the correlation matrix receives new meaning. Prior, $|X^t X|$ was used to indicate either the presence of orthogonality or the presence of multicollinearity within the independent variable set. By transforming $|X^t X|$ into an approximate Chi-Square statistic, we now possess a calibrated scale against which the departures from orthogonality can be measured. We possess a gradient between singularity and orthogonality. A large Chi-Square value indicates substantial multicollinearity. However, further research is needed to indicate exactly how the Chi-Square statistic is calibrated to show the amount of departure from orthogonality. Another topic that requires research is to determine the pattern of interdependence within the independent variables. This study is a discussion of an initial step into multicollinearity in regression analysis and therefore leaves many questions unanswered.
LITERATURE CITED

Bartlett, M. S. Tests of Significance In Factorial Analysis. British Journal of Psychology, Statistical Section. 3 (1950)


Wilks, S. S. Certain Generalizations In the Analysis of Variance. Biometrika. 24 (1932)


Table 2. 2nd moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^-8).

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>m 50</th>
<th>100</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
<th>9000</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-397</td>
<td>-50</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>3</td>
<td>-118</td>
<td>-16</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>4</td>
<td>2514</td>
<td>340</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>8882</td>
<td>1293</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>19221</td>
<td>3080</td>
<td>30</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>32243</td>
<td>5823</td>
<td>62</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>45441</td>
<td>9473</td>
<td>112</td>
<td>15</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>55945</td>
<td>13803</td>
<td>185</td>
<td>25</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>61519</td>
<td>18433</td>
<td>285</td>
<td>39</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>61259</td>
<td>22895</td>
<td>415</td>
<td>57</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>55742</td>
<td>26712</td>
<td>579</td>
<td>82</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>46634</td>
<td>29485</td>
<td>779</td>
<td>112</td>
<td>15</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>36022</td>
<td>30950</td>
<td>1014</td>
<td>150</td>
<td>20</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>25766</td>
<td>31017</td>
<td>1285</td>
<td>195</td>
<td>27</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>17098</td>
<td>29767</td>
<td>1589</td>
<td>249</td>
<td>35</td>
<td>11</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>10540</td>
<td>27422</td>
<td>1923</td>
<td>311</td>
<td>44</td>
<td>14</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>6039</td>
<td>24294</td>
<td>2283</td>
<td>382</td>
<td>55</td>
<td>17</td>
<td>7</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>3217</td>
<td>20730</td>
<td>2662</td>
<td>461</td>
<td>68</td>
<td>21</td>
<td>9</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>1593</td>
<td>17057</td>
<td>3054</td>
<td>550</td>
<td>82</td>
<td>26</td>
<td>11</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3. 3rd moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^{-8}).

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>50</th>
<th>100</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
<th>9000</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-552</td>
<td>-72</td>
<td>-1</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>3</td>
<td>-158</td>
<td>-22</td>
<td>0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>4</td>
<td>3172</td>
<td>446</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>10357</td>
<td>1706</td>
<td>17</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>20292</td>
<td>3868</td>
<td>43</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>30175</td>
<td>6885</td>
<td>88</td>
<td>12</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>36898</td>
<td>10439</td>
<td>158</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>38561</td>
<td>14028</td>
<td>257</td>
<td>36</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>35197</td>
<td>17095</td>
<td>388</td>
<td>55</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>28435</td>
<td>19170</td>
<td>555</td>
<td>81</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>20507</td>
<td>19977</td>
<td>757</td>
<td>114</td>
<td>16</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>13277</td>
<td>19481</td>
<td>993</td>
<td>155</td>
<td>22</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>7745</td>
<td>17868</td>
<td>1260</td>
<td>205</td>
<td>29</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>4080</td>
<td>15474</td>
<td>1552</td>
<td>263</td>
<td>38</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>1944</td>
<td>12689</td>
<td>1862</td>
<td>330</td>
<td>49</td>
<td>15</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>838</td>
<td>9875</td>
<td>2181</td>
<td>406</td>
<td>62</td>
<td>20</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>327</td>
<td>7306</td>
<td>2502</td>
<td>490</td>
<td>76</td>
<td>24</td>
<td>11</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>115</td>
<td>5146</td>
<td>2813</td>
<td>581</td>
<td>93</td>
<td>30</td>
<td>13</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>37</td>
<td>3454</td>
<td>3106</td>
<td>680</td>
<td>112</td>
<td>36</td>
<td>16</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 4. 4th moment differences for 9 sample sizes and (2-20) independent variables. (difference x 10^{-8}).

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>50</th>
<th>100</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
<th>9000</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-685</td>
<td>-92</td>
<td>-1</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>3</td>
<td>-189</td>
<td>-28</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>4</td>
<td>3582</td>
<td>570</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>10841</td>
<td>2007</td>
<td>22</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>19303</td>
<td>4333</td>
<td>56</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>25564</td>
<td>7271</td>
<td>112</td>
<td>15</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>27274</td>
<td>10288</td>
<td>198</td>
<td>28</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>24352</td>
<td>12770</td>
<td>317</td>
<td>46</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>18588</td>
<td>14228</td>
<td>470</td>
<td>70</td>
<td>10</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>12286</td>
<td>14435</td>
<td>659</td>
<td>102</td>
<td>14</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>7089</td>
<td>13466</td>
<td>879</td>
<td>142</td>
<td>20</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>3590</td>
<td>11631</td>
<td>1126</td>
<td>191</td>
<td>23</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>1600</td>
<td>9347</td>
<td>1392</td>
<td>249</td>
<td>37</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>629</td>
<td>7015</td>
<td>1667</td>
<td>315</td>
<td>48</td>
<td>15</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>218</td>
<td>4931</td>
<td>1941</td>
<td>389</td>
<td>62</td>
<td>20</td>
<td>9</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>67</td>
<td>3253</td>
<td>2202</td>
<td>471</td>
<td>77</td>
<td>25</td>
<td>11</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>2018</td>
<td>2440</td>
<td>559</td>
<td>94</td>
<td>31</td>
<td>14</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>1178</td>
<td>2646</td>
<td>651</td>
<td>114</td>
<td>38</td>
<td>17</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>648</td>
<td>2812</td>
<td>747</td>
<td>136</td>
<td>46</td>
<td>20</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>
VITA

Stephen M. Hattori

Candidate for the Degree of

Master of Science

Thesis: An Evaluation of Bartlett's Chi-Square Approximation for the Determinant of a Matrix of Sample Zero-Order Correlation Coefficients

Major Field: Applied Statistics

Biographical Information:

Personal Data: Born at Ogden, Utah, December 10, 1947 son of Mas and Ruth Ota Hattori; single.

Education: Attended grammar school in Lawrence, Massachusetts; graduated from Lawrence High School in 1965; received a Bachelor of Science degree from Northeastern University with a major in secondary education with a minor in mathematics, in June 1970; completed requirements for the Master of Science degree at Utah State University in 1974.

Professional Experience: 1971 to 1972, consultant in applied statistics and computer science, Department of Applied Statistics - Computer Science, Utah State University; 1968 to 1969, co-op student, Wang Laboratories, Inc.