ON APPROXIMATE NORMALIZING TRANSFORMATIONS

by

Mario M.A. D'Avirro

A thesis submitted to the Faculty of Graduate Studies and Research, in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics
McGill University
Montreal, Quebec

August, 1974

ON APPROXIMATE NORMALIZING TRANSFORMATIONS

by

Mario M.A. D'Avirro
Department of Mathematics
M.Sc. Degree

ABSTRACT

In this thesis, we present a system of normalizing transformations defined by the equation, μ being the mean of x:

$$y = \{1 + a_1(x-\mu) + a_2(x-\mu)^2\}^h .$$

We make y approximately normal by expanding the expression above and then solving for the constants a_1, a_2, and h, so that the third, fourth and possibly the fifth cumulants of y are made as close as possible to zero, the cumulants of the normal distribution. We derive several theorems on the validity of an expansion of y, and prove certain properties of the cumulants of y. Numerous comparisons are made with other approximations for a variety of distributions.

These seem to establish that the proposed transformation provides comparable accuracy in the cases of the Negative Binomial, the χ^2, the non-central t, the non-central F, and the correlation coefficient in normal samples, while in the case of the non-central χ^2 distribution, it provides accuracy better than any other considered. Finally, in the appendix, a computer program is presented that will calculate the

constants a_1, a_2, and h; given four or five cumulants of any random variable, and will then yeild either the appropriate percentile level or the quantile.

McGill University
Montreal

August 1974

SUR LES TRANSFORMATIONS APPROXIMATIVEMENT NORMALISANTES

par

Mario M.A. D'Avirro
Département de Mathématiques
M.Sc. Degré

RESUME

Dans cette thèse, nous présentons un système de transformations normalisantes défini par l'équation, μ étant la moyenne de x:

$$y = \{1 + a_1(x-\mu) + a_2(x-\mu)^2\}^h .$$

Nous rendons y approximativement normal en développant cette expression et en solutionnant pour les constantes a_1, a_2 et h, et ainsi le troisième, le quatrième et possiblement le cinquième cumulant de y est rendu aussi près que possible de zéro, le cumulant de la distribution normale.

Nous déduisons plusieurs théorèmes sur la validité du développement de y et nous prouvons certaines propriétés des cumulants de y. Plusieurs comparaisons sont faites avec d'autres approximations pour différentes distributions.

Ces comparaisons semblent établir que la transformation proposée fournit une aussi bonne précision pour les distributions suivantes: Binomiale Négative, Chi-carré, t non-centré, F non-centré et le coefficient de corrélation dans un échantillon normal, tandis que pour la distribution du χ^2 non-centré cette transformation fournit une approximation supérieure à toutes les autres.

Finalement dans l'appendice un program est inclus qui calculera, lui étant fournit quatre ou cinq cumulants, les constantes a_1, a_2 et h et produira ensuite le niveau de percentile approprié ou le quantile.

Université McGill
Montréal

août 1974

ACKNOWLEDGEMENTS

Firstly, I would like to thank Professor H. Ruben for suggesting the topic for this thesis and for his advice throughout the various stages leading to the final manuscript.

Thanks are also due to Professor G.P.H. Styan for his invaluable help throughout my graduate years, and to Professor E. Kounias for assistance in the preparation of certain theorems.

Last but not least, I offer my sincerest thanks to my wife, without whose advice, help, and constant encouragement this would not have been possible.

TABLE OF CONTENTS

		Page
CHAPTER I:	Introduction	1
1.1	Moments and Cumulants	2
1.2	Asymptotic Expansion of Distribution	5
1.3	The Wilson-Hilferty Transformation	10
1.4	The Haldane Transformations	12
1.5	Other Normalizing Transformations	16
CHAPTER II:	Expansion of Transformation	24
2.1	Expansion into Moments	26
2.2	Expansion into Cumulants	29
2.3	Validity of Expansions	34
2.4	Cumulants of a Transformed Variate	42
2.5	Problem of Order	46
CHAPTER III:	General Solution for Constants	52
3.1	Solution of Case A	56
3.2	Solution of Case B	59
3.3	Solution of Case C	62
	3.3.1 Subcase C1	62
	3.3.2 Subcase C2	63
	3.3.3 Subcase C3	63
3.4	Conclusion	64
CHAPTER IV:	Cumulants of Distributions	69
4.1	Distributions of Cumulative Statistics	71
	4.1.1 The Binomial Distribution	71
	4.1.2 The Negative Binomial Distribution	73
	4.1.3 The Poisson Distribution	75
	4.1.4 The Gamma Distribution (Including χ^2)	76
	4.1.5 The Non-Central χ^2 Distribution	77

Page

CHAPTER IV: Continued

4.2 Other Distributions 78

4.2.1 The Hypergeometric Distribution 79

4.2.2 The Beta Distribution 80

4.2.3 "Student's" Distribution 81

4.2.4 The F Distribution 82

4.2.5 The Non-Central t Distribution 84

4.2.6 The Non-Central F Distribution 85

4.2.7 The Distribution of the Sample Correlation Coefficient in Normal Samples 86

CHAPTER V: Results and Conclusions 88

5.1 Symmetric Distributions 89

5.2 Results for the Cumulative Statistics 90

5.2.1 The Binomial Distribution 91

5.2.2 The Negative Binomial Distribuiton 95

5.2.3 The Poisson Distribution 97

5.2.4 The Gamma Distribution (Including χ^2) 102

5.2.5 The Non-Central χ^2 Distribution 104

5.3 Results for Other Distribution 107

5.3.1 The Hypergeometric Distribution 111

5.3.2 The Beta Distribution 113

5.3.3 "Student's" Distribution 115

5.3.4 The F Distribution 119

5.3.5 The Non-Central t Distribution 122

5.3.6 The Non-Central F Distribution 123

5.3.7 The Distribution of the Sample Correlation Coefficient in Normal Samples 128

5.4 Final Conclusions 129

REFERENCES AND BIBLIOGRAPHY 135

APPENDIX A 140

APPENDIX B 146

LIST OF TABLES

TABLE		Page
2.1	Groupings of the cumulants of x	51
3.1	Assumed ordering of the cumulants	53
3.2	Orders of γ_i of y	68
5.1	Binomial Distribution - solution for constants	92
5.2	Binomial Distribution - comparisons of approximations	94
5.3	Negative Binomial Distribution - solution for constants.	96
5.4	Negative Binomial Distribution - comparisons of approximations (1)	98
5.5	Negative Binomial Distribution - comparisons of approximations (2)	99
5.6	Poisson Distribution - solution for constants	100
5.7	Poisson Distribution - comparisons of approximations	101
5.8	Gamma Distribution - solution for constants	103
5.9	Gamma Distribution - comparisons of approximations (1)	105
5.10	Gamma Distribution - comparisons of approximations (2)	106
5.11	Non-central χ^2 Distribution - comparisons of approximations (1)	108
5.12	Non-Central χ^2 Distribution - comparisons of approximations (2)	109
5.13	Non-Central χ^2 Distribution - comparisons of approximations (3)	110
5.14	Hypergeometric Distribution - comparisons of approximations	114
5.15	Beta Distribution - comparisons of approximations	116
5.16	"Student's" Distribution - comparisons of approximations	118
5.17	F Distribution - comparisons of approximations	121
5.18	Non-Central t Distribution - comparisons of approximations	124
5.19	Non-Central F Distribution - comparisons of approximations (1)	126
5.20	Non-Central F Distribution - comparisons of approximations (2)	127
5.21	Distribution of Correlation Coefficient - comparisons of approximations (1)	130
5.22	Distribution of Correlation Coefficient - comparisons of approximations (2)	131

CHAPTER I

INTRODUCTION

In statistics, one is often confronted with the problem of determining the probability that the divergence of a random variable from its expected value should be due to sampling error. We thus require to tabulate or compute values of the distribution function of the random variable in question. In recent years, especially in Monte-Carlo testing and simulation, a quick and easy method for generating a particular random variable has also been needed.

We approach the problem by considering the following transformation:

$$y = (1 + a_1(x - \mu) + a_2(x - \mu)^2)^h, \qquad (1.1)$$

where a_1, a_2, and h are constants and $\mu = E(x)$. We then derive the values of a_1, a_2, and h such that the new transformed random variable y is approximately normal. Thus since the normal distribution is tabulated and normal deviates can easily be generated, a " good " approximation may then be easily obtained. Our method for deriving a_1, a_2, and h is that of equating the higher cumulants (that is, the third, fourth, and so on, if feasible) of y to those of the normal distribution, which are zero.

Out setup is motivated by the papers of Wilson and Hilferty

(1931) and of Haldane (1938) which we shall review in Sections 1.3. and 1.4. Section 1.1 will be devoted to a brief exposition of some useful properties of moments and cumulants. In Section 1.2 we shall review a transformation by Cornish and Fisher (1937) and in Section 1.5 some other transformations in the current literature.

Chapter II shall be devoted to the expansion of (1.1) and determining the cumulants of y. We shall also examine certain assumptions to be made and certain questions of validity. In Chapter III a general solution for the constants in the transformation (1.1) will be presented. In Chapter IV we shall derive the cumulants of distributions, or classes of distributions, which x may represent.

In Chapter V we shall present a solution for each of the cases outlined in Chapter IV, together with the results and the conclusions in each case. In Appendix A, we present several of the extensive formulae that we have developed. Appendix B shall be devoted to a listing of subroutines used which have been developed, and may prove to be of interest.

Throughout the thesis we will use the following notations: K_i will denote the ith cumulant of x while $K_i(y)$, the ith cumulant of y. We will use μ as the mean of x, σ^2 as its variance. In all cases where ambiguity occurs, (y) shall be suffixed to the appropriate property of y. y will always denote the transformed variable, approximately normal, x the given non-normal random variable, and z is a standardized normal random variable.

1.1. Moments and Cumulants

We define a series of coefficients α_r, r = 1, 2, ..., by the

relation

$$\alpha_r = \int_{-\infty}^{\infty} (x-a)^r df .$$

α_r is called the rth *moment of the distribution F(x) about the point a.* In particular when a = 0, α_r is denoted by μ'_r and is called plainly the rth *moment of F(x).* Also when a = μ'_1, or more simply μ, then α_r is denoted by μ_r and called the rth *central moment.* We may also define the rth *factorial moment about an arbitrary origin* by the equation $\mu'_{(r)} = \int_{-\infty}^{\infty} (x-a)_{(r)} dF$ where $x_{(r)} = x(x-1)(x-2) \ldots (x-r+1)$.

The moments are a set of descriptive constants pertaining to a distribution, and useful, under appropriate conditions, in ascertaining its properties or completely specifying it. They are not, however the unique set of constants with these properties. There is another series of coefficients, namely the cumulants, which have certain useful theoretical properties.

Formally then, we define a series of coefficients K_r , r = 1, 2, ..., called *the cumulants* by the following relation in (it)

$$\exp\{K_1(it) + K_2 \frac{(it)^2}{2!} + \ldots + K_r \frac{(it)^r}{r!} + \ldots\}$$

$$= 1 + \mu'_1(it) + \ldots + \mu'_r \frac{(it)^r}{r!} + \ldots \qquad (1.2)$$

$$= \int_{-\infty}^{\infty} e^{itx} dF$$

$$= \phi_x(t).$$

$\phi_x(t)$ is denoted as the *characteristic function* of the random variable x. We then define $\psi_x(t) = \log\phi_x(t)$ as the *cumulant generating function* of x, since it has the property that K_r is the coefficient of $\frac{(it)^r}{r!}$ in its power series expansion. Furthermore, K_r exists whenever μ'_r exists.

Recursive formulae between the moments and cumulants may be obtained by differentiating (1.2) with respect to (it). We then have

$$\sum_{\ell=1}^{\infty} \mu'_\ell \frac{(it)^{\ell-1}}{(\ell-1)!} = \sum_{\ell=1}^{\infty} K_\ell \frac{(it)^{\ell-1}}{(\ell-1)!} \exp\{\sum_{\ell=1}^{\infty} K_\ell \frac{(it)^\ell}{\ell!}\}$$

$$= \sum_{\ell=1}^{\infty} K_\ell \frac{(it)^{\ell-1}}{(\ell-1)!} \sum_{\ell=0}^{\infty} \mu'_\ell \frac{(it)^\ell}{\ell!} .$$

Expanding and setting all the coefficients of $(it)^r$ equal, we have

$$\mu'_{j+1} = j! \sum_{\ell=0}^{j} \frac{K_{\ell+1}\,\mu'_{j-\ell}}{\ell!(j-\ell)!} \quad j = 0, 1, 2, \ldots, \qquad (1.3)$$

and

$$K_1 = \mu'_1$$

$$K_{j+1} = \mu'_{j+1} - j! \sum_{\ell=0}^{j-1} \frac{K_{\ell+1}\mu'_{j-\ell}}{\ell!(j-\ell)!} \quad j = 1, 2, 3, \ldots . \qquad (1.4)$$

Another interesting property is revealed if we consider the linear transformation of the type

$$\xi = \ell x + m.$$

This leaves the cumulants unchanged so far as m is concerned, while it multiples the cumulant K_r by ℓ^r. The sole exception is the first cumulant which besides being multiplied by ℓ, is also increased by m. Furthermore, if m = 0, the moment μ_r is also multiplied by ℓ^r. This property is critical to the developments that follow. In fact, in all the transformations to be considered, it is this property that assures us of the convergence of the expansions under consideration.

Under general conditions, a knowledge of the moments, when they all exist, or similarly the cumulants, is equivalent to a knowledge of the distribution function, that is, that it is theoretically possible to exhibit all the properties of the distribution. In fact, under appropriate conditions, there are the Gram-Charlier and the Edgeworth expansion of any density, which are in terms of the moments and cumulants of the distribution. Thus distributions which have a finite number of lower moments or cumulants in common will, in a sense, be approximations of one another, and, in practice, approximations of this kind often turn out to be remakably good, even when the first three or four moments or cumulants are equated.

1.2. Asymptotic Expansions of Distributions

It has long been known that in many cases several important

theoretical distributions occurring in statistics depend on some variable n in such a way that the distribution tends to normality when n is large. When n is large, it may be sufficient to assume the distribution approximately normal, however for small n, the approximation may be inadquate.

However, using the fact that the deviations from normality can sometimes be conveniently expressed by means of Hermitian polynomials, Cornish and Fisher (1937) developed explicit asymptotic expansions, expressing any desired percentile point of such distributions in terms of known cumulants. They constructed the transformation:

$$y = c_0 + c_1 x + c_2 x^2 + \ldots \qquad (1.5)$$

where the orders of the c_i decrease for larger i. When the c_i are chosen appropriately, the distribution of y is made very close to normality enabling us to find the quantile of x from that of y, which we then assume normal.

Cornish and Fisher noted that the effect of the operator

$$\exp\{ \frac{a_r}{r!} \left(- \frac{d}{dx} \right)^r \}$$

when acting upon the distribution of x is to increase the rth cumulant by a_r, but otherwise leaves the distribution unchanged. This enabled them to consider the Edgeworth expansion of the density of x

$$\exp \{ - \frac{K_1 - m}{1!} D + \frac{K_2 - s^2}{2!} D^2 - \frac{K_3}{3!} D^3 + \frac{K_4}{4!} D^4 - \ldots \} \beta(x) \qquad (1.6)$$

where $\beta(x) = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

and $D = \frac{d}{dx}$.

The constants m, s^2 are the mean and variance of the Edgeworth expansion which are not necessarily equal to μ, σ^2 since the approximation may perhaps be slightly improved by taking them differently.

Letting then

$$s\ell_1 = K_1 - m$$

$$s^2\ell_2 = K_2 - s^2$$

$$s^r\ell_r = K_r \qquad \text{for } r > 2,$$

and since by choice of s, s^2 is of order K_2, we have

$$\ell_1 = 0(n^{-1/2}) \quad , \quad \ell_2 = 0(n^{-1}),$$

$$\ell_r = 0(n^{1-r/2}) \qquad \text{for } r > 2.$$

Then, (1.6), the density of x may be written as

$$\exp\{ -\ell_1 sD + \tfrac{1}{2}\ell_2 s^2 D^2 - \tfrac{1}{6}\ell_3 s^3 D^3 + \tfrac{1}{24}\ell_4 s^4 D^4 + \ldots\} \quad \beta(x).$$

Expanding the operator by means of the exponential expansion, and noting that

$$s^r D^r \beta(x) = (-1)^r H_r\left(\frac{x-m}{s}\right) \quad ,$$

where $H_r(x)$ is the Tchebycheff-Hermite polynomial, we obtain the distribution function by integrating term by term. Then the frequency less than or equal to $m + sx$, may be expressed in terms of the corresponding normal probability integral and a series of adjustments of decreasing order of magnitude. To order at least n^{-2}, we have,

$$\int_{-\infty}^{x} \alpha(x)dx + \alpha(x)\Big\{ - (\ell_1 + \tfrac{1}{6}\ell_3 H_2) - (\tfrac{1}{2}\ell_1{}^2 H_1 + \tfrac{1}{6}\ell_1\ell_3 H_3 + \tfrac{1}{2}\ell_2 H_1$$

$$+ \tfrac{1}{24}\ell_4 H_3 + \tfrac{1}{72}\ell_3^2 H_5) - (\tfrac{1}{6}\ell_1^3 H_2 + \tfrac{1}{2}\ell_1\ell_2 H_2 + \tfrac{1}{12}\ell_1^2\ell_3 H_4 + \tfrac{1}{12}\ell_2\ell_3 H_4 + \tfrac{1}{24}\ell_1\ell_4 H_4$$

$$+ \tfrac{1}{120}\ell_5 H_4 + \tfrac{1}{72}\ell_1\ell_3^2 H_6 + \tfrac{1}{144}\ell_3\ell_4 H_6 + \tfrac{1}{1296}\ell_3^3 H_8) - (\tfrac{1}{24}\ell_1^4 H_3 + \tfrac{1}{8}\ell_2^2 H_3 + \tfrac{1}{4}\ell_1^2\ell_2 H_3$$

$$+ \tfrac{1}{36}\ell_1^3\ell_3 H_5 + \tfrac{1}{12}\ell_1\ell_2\ell_3 H_5 + \tfrac{1}{48}\ell_1^2\ell_4 H_5 + \tfrac{1}{120}\ell_1\ell_5 H_5 + \tfrac{1}{720}\ell_6 H_5 + \tfrac{1}{144}\ell_1^2\ell_3^2 H_7$$

$$+ \tfrac{1}{144}\ell_2\ell_3^2 H_7 + \tfrac{1}{1152}\ell_4^2 H_7 + \tfrac{1}{144}\ell_1\ell_3\ell_4 H_7 + \tfrac{1}{720}\ell_3\ell_5 H_7 + \tfrac{1}{129}\ell_1\ell_3^3 H_9 + \tfrac{1}{1728}\ell_3^2\ell_4 H_9$$

$$+ \tfrac{1}{31104}\ell_3^4 H_{11} + \tfrac{1}{48}\ell_2\ell_4 H_5)\Big\} , \qquad (1.7)$$

where $\alpha(x) = \dfrac{1}{\sqrt{2\Pi}} e^{-x^2/2}$.

Assuming y to be a normal variate with $\Phi(y)$ its distribution function, we have, by Taylor's theorem

$$\Phi(y) = \Phi\{x + (y-x)\} = \Phi(x) + \sum_{r=1}^{\infty} \frac{(y-x)^r}{r!} \frac{d^r}{dx^r}\Phi(x)$$

$$= \Phi(x) + \sum_{r=1}^{\infty} \frac{(y-x)^r}{r!} H_{r-1}(x)\alpha(x) ,$$

and this is equal to (1.7). Now by considering the terms of each order of magnitude in succession, and identifying the coefficients in (1.5) we develop, in terms of successive polynomials in x,

$$x - y = \ell_1 + \frac{1}{6}\ell_3(x^2-1) + \frac{1}{2}\ell_2 x - \frac{1}{3}\ell_1\ell_3 x + \frac{1}{24}\ell_4(x^3-3x)$$

$$- \frac{1}{36}\ell_3^2(4x^3 - 7x) - \frac{1}{2}\ell_1\ell_2 + \frac{1}{6}\ell_1^2\ell_3 - \frac{1}{12}\ell_2\ell_3(5x^2 - 3) - \frac{1}{8}\ell_1\ell_4(x^2 - 1)$$

$$+ \frac{1}{120}\ell_5(x^4 - 6x^2 + 3) + \frac{1}{36}\ell_1\ell_3^2(12x^2 - 7) - \frac{1}{144}\ell_3\ell_4(11x^4 - 42x^2 + 15)$$

$$+ \frac{1}{648}\ell_3^3(69x^4 - 187x^2 + 52) - \frac{3}{8}\ell_2^2 x + \frac{5}{6}\ell_1\ell_2\ell_3 x + \frac{1}{8}\ell_1^2\ell_4 x - \frac{1}{48}\ell_2\ell_4(7x^3 - 15x)$$

$$- \frac{1}{30}\ell_1\ell_5(x^3 - 3x) + \frac{1}{720}\ell_6(x^5 - 10x^3 + 15x) - \frac{1}{3}\ell_1^2\ell_3^2 x + \frac{1}{72}\ell_2\ell_3^2(36x^3 - 49x)$$

$$- \frac{1}{384}\ell_4^2(5x^5 - 32x^3 + 35x) + \frac{1}{36}\ell_1\ell_3\ell_4(11x^3 - 21x) - \frac{1}{360}\ell_3\ell_5(7x^5 - 48x^3 + 51x)$$

$$- \frac{1}{324}\ell_1\ell_3^3(138x^3 - 187x) + \frac{1}{864}\ell_3^2\ell_4(111x^5 - 547x^3 + 456x)$$

$$- \frac{1}{7776}\ell_3^4(948x^5 - 3628x^3 + 2473x). \qquad (1.8)$$

This is the required expression of the variate y, in terms of the variate x. To order n^{-2} at least, y will be normally distributed. Also, (1.8) may be inverted to express x in terms of y, by means of Lagrange's inversion formula.

The importance of the Cornish-Fisher expansion is that given enough terms or adjustments, y as defined in (1.5), as a polynomial in any other distribution x, can be made as close as one would want to the normal distribution: Thus if f(x) is any other transofrmation, and f(x) admits a Taylor series expansion, a term by term comparison of the respective orders of magnitude is possible. In fact, any transformation f(x) may be viewed as an attempt to simplify (1.5) into a concise form, without losing too much accuracy in the process.

1.3. The Wilson-Hilferty Transformation

The main purpose of Wilson and Hilferty's (1931) transformation was to find an approximation to the χ^2 distribution better than Fisher's variance stabalizing transformation $\sqrt{2\chi^2}$. Hence, for purposes of this Section, x will denote a χ^2 random variable.

Consider the distribution of

$$y = \left(\frac{x}{n}\right)^h , \qquad (1.9)$$

where n is the degrees of freedom of the χ^2 random variable x. Noting that the mean of x = n, (1.9) may be written as

$$y = \left(1 + \frac{(x-n)}{n}\right)^h$$

$$= 1 + h\frac{(x-n)}{n} + h\frac{(h-1)}{2!}\frac{(x-n)^2}{n^2} + \ldots , \qquad (1.10)$$

by the Taylor or binomial expansion. We note that (1.10) converges when $|x - n| < n$.

Taking the expectations of (1.10) and noting that $E(x-n)^r = \mu_r$ where $\mu_2 = 2n$

$$\mu_3 = 8n$$

$$\mu_4 = 12n(n+4)$$

$$\mu_5 = 32n(5n+12)$$

$$\mu_6 = 40n(3n^2+52n+96)$$

we find, after some reduction,

$$E(y) = 1 + \frac{h(h-1)}{2!}\frac{\mu_2}{n^2} + \frac{h(h-1)(h-2)}{3!}\frac{\mu_3}{n^3} + \ldots$$

$$= 1 + h(h-1)n^{-1} + \frac{1}{6}h(h-1)(h-2)(3h-1)n^{-2} + \frac{1}{6}h^2(h-1)^2(h-2)(h-3)n^{-3}$$

$$+ 0(n^{-4}). \qquad (1.11)$$

We then find $E(y^r)$ by merely substituting rh for h in (1.11). Then,

$$K_2(y) = 2h^2n^{-1} + 2h^2(h-1)(3h-1)n^{-2} + 0(n^{-3})$$

$$K_3(y) = 4h^3(3h-1)n^{-2} + 0(n^{-3}).$$

We now choose $h = \frac{1}{3}$ so that the leading term of $K_3(y)$, the term of order n^{-2} vanishes. It is interesting to note that the terms of order

n^{-2} in both $K_2(y)$ and $K_1(y)$ also vanish. This suggests that to a considerable degree of accuracy, (1.9) with $h = \frac{1}{3}$ is almost normally distributed with mean $1 - \frac{2}{9n}$ and variance $\frac{2}{9n}$.

1.4. The Haldane Transformations

Haldane (1938) considers a variable so distributed that when n tends to infinity, the orders of magnitude of its cumulants are effectively as follows:

$$K_1 = K_2 = K_3 = 0(n), \quad K_r \leq 0(n^{r-3}) \quad \text{for } r \geq 4.$$

He claims to consider most cumulative and derived cumulative statistics, and some others. He excludes the cases where $K_3 = 0$, and a number of distributions which tend to normality with n. However, he only examines two cases or applications, that of the χ^2 and the binomial distribution. Nowhere does he name or produce results for any other distribution.

Haldane proceeds similar to Wilson and Hilferty, he lets

$$y = \left(\frac{x}{K_1}\right)^h \quad . \tag{1.12}$$

Then, noting that the mean of x, $\mu = K_1$, (1.12) reduces to

$$y = \left(1 + \frac{(x-\mu)}{K_1}\right)^h$$

$$= 1 + h\frac{(x-\mu)}{K_1} + \frac{h(h-1)}{2!}\frac{(x-\mu)^2}{K_1^2} + \ldots \quad ,$$

which converges when $|x-\mu| < K_1$. Again taking expectations, we find,

$$E(y) = 1 + \frac{h(h-1)}{2}\frac{\mu_2}{K_1^2} + \frac{h(h-1)(h-2)}{3!}\frac{\mu_3}{K_1^3} + \dots \quad . \tag{1.13}$$

Expressing the moments in terms of cumulants, and noting that the rth moment of y is merely the expression (1.13) with h replaced by rh, we derive the following cumulants of y after some considerable reduction, expressed in decreasing orders of magnitude:

$$K_1(y) = 1 + h(h-1)\frac{K_2}{2K_1^2} + h(h-1)(h-2)\frac{[4K_1K_3 + 3(h-3)K_2^2]}{24K_1^4}$$

$$+ h(h-1)(h-2)(h-3)\frac{[2K_1^2K_4 + 4(h-4)K_1K_2K_3 + (h-4)(h-5)K_2^3]}{48K_1^6} + 0(n^{-4})$$

$$K_2(y) = h^2\frac{K_2}{K_1^2} + h^2(h-1)\frac{[2K_1K_3 + (3h-5)K_2^2]}{2K_1^4}$$

$$+ h^2(h-1)\frac{[(7h-11)K_1^2K_4 + 4(h-2)(7h-12)K_1K_2K_3 + 2(h-2)(7h^2-30h+32)K_2^3]}{12K_1^6}$$

$$+ 0(n^{-4})$$

$$K_3(y) = h^3\frac{[K_1K_3 + 3(h-1)K_2^2]}{K_1^4}$$

$$+ h^3(h-1)\frac{[3K_1^2K_4 + 3(7h-10)K_1K_2K_3 + (17h^2-55h+44)K_2^3]}{2K_1^6} + 0(n^{-4})$$

$$K_4(y) = h^4 \frac{[K_1^2 K_4 + 12(h-1)K_1K_2K_3 + 4(h-1)(4h-5)K_2^3]}{K_1^6} + 0(n^{-4}) . \quad (1.14)$$

We note that in (1.14) the additional assumption $K_4 = 0(n)$, $K_r \leq 0(n^{r-4})$ was made for reasons to become apparent. This does not affect the result below.

Therefore, if we set $h = 1 - \frac{K_1K_3}{3K_2^2}$ then the leading term of $K_3(y)$ in (1.14) vanishes. Then we can say that when $h = 1 - \frac{K_1K_3}{3K_2^2}$, (1.12) is almost normally distributed with mean and variance $K_1(y)$ and $K_2(y)$, respectively, as defined in (1.14). Haldane refers to this as transformation A. One should note that transformation A is equivalent to the Wilson-Hilferty approximation when x is chi-square.

At this point, Haldane notes that the leading term in $K_4(y)$ may be set to zero by adding a suitable constant to x in (1.12) so that K_1 assumes any desired value g, while the other cumulants remain unchanged. So then under the revised assumptions,

$$K_2 = K_3 = K_4 = 0(n), \quad K_r \leq 0(n^{r-4}) \qquad \text{for } r \geq 4 ,$$

Haldane makes the following transformation,

$$y = \left(1 + \frac{x - K_1}{g}\right)^h . \quad (1.15)$$

The cumulants of y then are those in equation (1.14) with K_1 replaced by g. Then setting the leading term of $K_3(y)$, the term of order n^{-2}, and the leading term of $K_4(y)$, the term of order n^{-3}, equal to zero, we have

$$g = \frac{12K_2^2K_3}{20K_3^2 - 9K_2K_4} \qquad h = \frac{16K_3^2 - 9K_2K_4}{20K_3^2 - 9K_2K_4} \quad . \tag{1.16}$$

Then y will be so distributed that its third and fourth cumulants tend to zero with n^{-3} and n^{-4} respectively. Hence the distribution of (1.15) with g and h as in (1.16) would be very close to the normal distribution, with mean and variance defined ~~by~~ (1.14). This transformation will be referred to as transformation B.

One should note that in transformation B, Haldane neglects to make an assumption of order on g. Upon inspection of (1.13) with K_1 replaced by g, we realize that the series would not converge unless a specific order were to be assumed upon g. In fact the solutions (1.16) are possible with the assumption that g is of the same order as K_1, that is, $g = O(n)$.

Haldane also computes and compares the following measures of skewness and kurtosis

$$\gamma_1 = \frac{K_3}{K_2^{3/2}} \quad , \quad \gamma_2 = \frac{K_4}{K_2^2} \tag{1.17}$$

which indicates a relative "proximity" to the normal distribution. In the case of the Wilson-Hilferty transformation or Haldane's transformation A, we have

$$\gamma_1 = O(n^{-3/2}) \quad , \qquad \gamma_2 = O(n^{-1}).$$

In the case of Haldane's transformation B, we have

$$\gamma_1 = O(n^{-3/2}) \quad , \qquad \gamma_2 = O(n^{-2}) \quad .$$

These measures indicate a very good approximation to normality. In fact, under the assumptions made upon the cumulants, that $K_2 = K_3 = K_4 = O(n)$, we have, that for any distribution x,

$$\gamma_1 = O(n^{-1/2}) \quad , \qquad \gamma_2 = O(n^{-1}) .$$

Haldane notes that other transformations are possible. In particular, h and g may be chosen such that the terms of order n^{-2} and n^{-3} vanish in $K_3(y)$. However, he does not pursue the matter further.

In particular applications of transformations A and B, Haldane only discusses two distributions, the χ^2 and the binomial distributions. In the case of χ^2, he finds that for $n > 10$, transformation B is decidedly superior to transformation A, which is in turn also very good. With regards to the binomial distribution, transformation B does not seem to be well suited, but transformation A shows a considerable degree of accuracy. We shall examine these points in greater detail and display numerical results in Chapter V .

1.5. Other Normalizing Transformations

Transformations of statistical variables were used in statistics mainly for two purposes, variance stabilization, and normalization. Stabilization of variance involves the following. Let $\{T_n\}$, $n = 1, 2, \ldots$,

be a sequence of statistics, such that

$$\sqrt{n}(T_n-\theta) \xrightarrow{L} Y \sim N(0,\sigma^2) \quad ,$$

where σ^2 is a function of θ. Then if g is differentiable at θ with $g'(\theta)\neq 0$,

$$\sqrt{n}[g(T_n) - g(\theta)] \xrightarrow{L} Y \sim N(0,[g'(\theta)\sigma]^2), \qquad (1.18)$$

Now transformations which stabalize the variance make the variance of the transformed variable approximately constant and independent of the parameter, for example θ . Then we must choose g such that

$$g'(\theta)\sigma = c \qquad \text{or} \qquad g = \int \frac{c}{\sigma}\, d\theta$$

This has led to the introduction of the inverse sine square root transformation used on the relative frequency of successes in n Bernoulli trials, that is, effectively in the binomial distribution, the square root transformation for a Poisson, a gamma, or a chi-square variate, the inverse hyperbolic sine transformation for a negative binomial random variable, and the inverse hyperbolic tangent transformation for the correlation coefficient in normal samples.

However, these transformations introduced for variance stabilization have been applied to normalization, that is, finding the tail probabilities, and confidence limits. This is due to the fact that, as shown in (1.18), $g(T_n)$ also tends to normality. In fact, since $g(T_n)$ tends to normality

with a constant variance, the results are more stable, and thus in stabilizing the variance, we usually succeed in bringing the transformed variate closer to normality. Empirical tests have also shown that the transformed random variable is usually the better approximation. (See, for example, Paulson, 1947; Hald, 1952). Many modifications in order to improve the accuracy, have been made to the transformations cited above, notably Anscombe (1948), Freedman and Tukey(1950), and Hotelling (1953). Kendall in his discussion on Hotelling's paper presents a general method for arriving at successive refinements of variance stabalizing transformations.

Blom (1954) derives a transformation of the binomial distribution and then applies the same idea to the Poisson, the χ^2, and the negative binomial distributions, such that it gives the Cornish-Fisher expansion, the property that the term of order n^{-1} is as small as possible. The application of this principle leads to the desired transformation as a solution of a differential equation. In fact, the differential equations lead to most of the variance stabalizing and normalizing transformations, namely the inverse sine square root, the square root, the Wilson-Hilferty or the cube root, and the logarithmic, together with some new, but unworkable, transformations. This principle however, has a major set-back in that the term of order n^{-1} contains the normal random variable as a parameter, and when trying to evaluate the tail probabilities our main task consists in calculating the normal random variable. In effect, if g is the required transformation, the Cornish-Fisher inverse expansion is of the form

$$g(x) = g(\alpha) + a_1 n^{-1/2} + a_2 n^{-1} + \dots \quad .$$

In the case of the binomial and the negative binomial

$$a_2 = \frac{1}{2} y^2 \alpha\beta \, [g''(\alpha) + h(y)(\frac{1}{\alpha} - \frac{1}{\beta}) \, g'(\alpha)],$$

where

$$h(y) = 1 - \frac{y^2 + 2}{3y^2}$$

and α, β are parameters of the distribution of x. Here g has the following solution apart from a multiplicative constant,

$$g = \int y^{-h} \, (1-y)^{-h} dx. \qquad (1.19)$$

In the case of the Poisson and χ^2 distributions, we have

$$a_2 = \frac{1}{2} y^2 [\alpha g''(\alpha) + h(y) g'(\alpha)],$$

which has the solution.

$$g(u) = \int_{u_0}^{u} y^{-h} dy = \begin{cases} \frac{1}{1-h} \, u^{1-h} + c_1 & h < 1 \\ \log u + c_2 & h = 1, \end{cases} \qquad (1.20)$$

where u_0, c_1, c_2 are constants.

The solutions (1.19) and (1.20) were arrived at by assuming a

particular value upon $h(y)$ and treating it as a constant. Thus with some prior knowledge of the area of the tail, (1.19) and (1.20) become approximately soluble. Blom examines three general areas, when $y = 1, 2$, and when y is very large. Then, (1.19) produces no transformation in the first case, the inverse sine transformation in the second case, and an unpractical one in the case when y is very large. Similarly, (1.20) produces no transformation when $y = 1$, the square root transformation when $y = 2$ and the cube root transformation when $y \to \infty$. One should note that, although the motivating principle seems to insure good transformations, the method of solving for g may produce an unsatisfactory result. This may be seen by the fact that in the case of χ^2, the Wilson-Hilferty or the cube root transformation empirically is better than the square root transformation throughout the range of y (as shown in their paper), which in turn is better than no transformation at all.

Moore (1957) examines the properties of the class of transformations

$$y = x^r \quad \text{where } 0 < r < 1, \quad x \geq 0,$$

with the assumption that the class is designed to transform the variate x into some form of normal distribution. Thus instead of considering the properties of x^r and solving for r, he examines the normal variate y, and examines the distribution of $\frac{1}{r}^{th}$ power of it. Finally Moore calculates γ_1 and γ_2 of $y^{1/r}$ as defined in (1.17), for $r = 3/4, 2/3, 1/2, 1/3$, and $1/4$. These are expressions in terms of a coefficient of variation (defined as the standard deviation divided by the mean). The power r is then chosen

such that γ_1 and γ_2 of x are closest to that of $y^{1/r}$ insuring at the same time that the variable has a suitable coefficient of variation by adding a constant to the original variable if necessary. Moore considers only three cases and finds that in the case χ^2, $r = \frac{1}{3}$ is best (the Wilson-Hilferty transformation), in the case of the distribution of the range ω in samples from a normal population, $r = 2/3$ is best, and in the case of a Poisson, $y = (x + 1/4)^{2/3}$ is the best transformation. Empirical tests show that these results work fairly well.

Peizer and Pratt (1968), and Pratt (1968) construct a new normal approximation to the beta distribution and its relatives, in particular, the binomial, Pascal, negative binomial, F, t, Poisson, gamma, and χ^2 dsitributions. They construct two sets of two transformations y_1, and y_2 (a refinement of y_1). For the binomial, Pascal, negative binomial, beta and F distributions, they have

$$y_i = d_i \left\{ \frac{1 + qg\left[\frac{S}{np}\right] + pg\left[\frac{T}{nq}\right]}{\left[n + \frac{1}{6}\right]pq} \right\}^{1/2} \qquad (i=1, 2), \quad (1.21)$$

and for the Poisson, gamma, and χ^2 distribution

$$y_i = d_i \left\{ \frac{1 + g\left[\frac{S}{M}\right]}{M} \right\}^{1/2} \qquad (i=1,2,), \quad (1.22)$$

where

$$g(x) = (1-x)^{-2} \; (1 - x^2 + 2x \log x) \qquad x > 0, \; x \neq 1$$

$$g(0) = 1, \quad g(1) = 0, \quad (1.23)$$

and M, p, q, S, T, d_1, and d_2 are constants or expressions depending on each specific distribution. The transformation for the t distribution is derived from the transformation of F with equal degrees of freedom in the numerator and the denominator.

The motivation of the formulae (1.21) and (1.22) stem from an approximation to the binomial distribution. We start from the approximation

$$y_o = \frac{S-np}{\sigma} ,$$

where $\sigma = (npq)^{1/2}$, S is the number of successes, and n, p, q the appropriate binomial constants. However y_o misplaces the median, and asymptotic investigation suggests the modifications, replacing n by n + 1/6 in σ such that $\sigma^* = [(n+1/6)pq]^{1/2}$ and adding a correction factor to the mean so that

$$y_3 = \frac{d_1}{\sigma^*} = (S-np+\frac{(q-p)}{6})/\sigma^* ,$$

which makes y_3 substantially more accurate to z, the normal deviate, about z = 0. Pratt also notices that

$$z/y_3 = f(p, \frac{y_o}{\sigma}) ,$$

where f is a gently varying function of its arguments. f is then chosen as the limit of $\frac{z}{y_3}$ as $\sigma \to \infty$ for p and $\frac{y_o}{\sigma}$ fixed. He then has

$$y_1 = \frac{d_1}{\sigma^*} G^{1/2} ,$$

where $G = 1 + pg(1-p\frac{y_o}{\sigma}) + qg(1+q\frac{y_o}{\sigma})$,

and again where g is defined in (1.23). The formula is then generalized to the beta distribution from which they derive transformations for the binomial, the Pascal, the negative binomial and the F. The Poisson, the gamma and the chi-square distributions are then taken as limiting cases.

Peizer and Pratt are quick to point out the astonishing accuracy of the transformations, especially in the extreme tails where the relative error in approximation to z approaches 0 as $\sigma \to \infty$. In fact $z - y_1$ (or y_2) is of order $O(\sigma^{-3})$.

Many other normalizing transfromations have been constructed for specific distributions, however few have presented a specific method to normalize any given distribution.

Finally, it is worthwhile to note that Fisher's square root transformation of the χ^2 random variable, is merely a specific case of transformation (1.1), when $a_2 = 0$, $a_1 = \frac{1}{n}$, $h = \frac{1}{2}$. However, Fisher's transformation was derived as a variance stabalizing and not a normalizing transformation. This suggests the very interesting fact that by suitable choice of a_1, a_2, and h, we may employ transformation (1.1) as a variance stabilizing transformation. That is, we may choose a_1, a_2, and h such that the second, the third and the fourth leading terms in the expansion of the second cumulant of y, are zero.

CHAPTER II

EXPANSION OF THE TRANSFORMATION

This chapter will be devoted to the twofold expansion of transformation (1.1). We will first take the Taylor series expansion of (1.1) about μ the mean of x, and then examine where this expansion is valid. By taking the integral of this series, we are then able to find the moments of y, from which we then calculate the first five cumulants.

Our primary aim is, as stated earlier, to set y as close as possible to a normal random variable. One way of accomplishing this is to equate the higher cumulants, from the third on, to zero. This is hardly possible, since we only have three parameters a_1, a_2, and h which may be given appropriate values. We therefore choose three leading terms and solve algebraically for a_1, a_2, and h, by equating them to zero. Also, by noting that the cumulants of y are polynomials in a_1, a_2, and h, we may use an iterative method for the numerical solution of the non-linear equations. However, in all cases, an assumption on the order of the terms in the expansion of the cumulants, must be made; a topic which shall be discussed in the final Section. Sections 3 and 4 are devoted to a theoretical study of the expansion and the cumulants of y.

The series expansions presented in this chapter are valid for all but a few random varibles x. The only restriction is that K_5 or μ_5', or in another case, K_4 or μ_4' exist. Symmetrical distributions, such as "Student's" t distribution, are, in constrast to Haldane's transformation, included.

Our transformation

$$y = (1 + a_1(x-\mu) + a_2(x-\mu)^2)^h.$$

is a generalization of Haldane's transformation, which in turn is a generalization of that of Wilson and Hilferty. We could have considered a more general transformation

$$y^* = b_o + \left[b_1 + b_2 \frac{(x-\mu)}{\sigma} + b_3 \frac{(x-\mu)^2}{\sigma^2} + b_4 \frac{(x-\mu)^3}{\sigma^3}\right]^h .$$

However, the addition of the last term $b_4 \frac{(x-\mu)^3}{\sigma^3}$, would make the amount of algebra involved in the series expansions of y^* and its cumulants, in practice, insurmountable. Furthermore, even if the series expansions were calculated, b_4 could not be expressed in terms of a closed expression, forcing a numerical solution in each case, and even then, b_4 might only have complex solutions. In fact, by including only b_3, the amount of algebra required is enormous. The remaining transformation (with $b_4 = 0$) is not, in fact, a generalization of (1.1), since as we noted in Section 1 of Chapter I,

$$\mu_i'\left(\frac{x-\mu}{\sigma}\right) = \sigma^{-i}\mu_i \quad \text{and} \quad K_i\left(\frac{x-\mu}{\sigma}\right) = \sigma^{-i}K_i \quad \text{for} \quad i \geq 2.$$

We then have

$$y = \frac{y^*-b_o}{b_1^h}, \quad a_1 = \frac{b_2}{b_1\sigma}, \quad a_3 = \frac{b_3}{b_1\sigma^2} .$$

y^* merely has a different mean and variance, but leads to an equivalent quantile or probability level.

One should note that our particular form of transformation is not unique. Obviously, if one views all approximate normalizing transformations as an attempt to approximate Cornish-Fisher expansions (1.5) or (1.8), with as few parameters as possible, then other types of transformations or functions are possible. Some possibilities may include

$$y = e^{a_1(x-\mu) + a_2(x-\mu)^2}$$

$$y = \log \{1 + a_1(x-\mu) + a_2(x-\mu)^2\}$$

$$y = \sin \{a_o + a_1(x-\mu) + a_2(x-\mu)^2\}$$

which may be viable alternatives to (1.1).

However, as mentioned in Chapter I, we were motivated by the amazingly accurate results of Wilson and Hilferty, and Haldane, and since our transformation is of the same form, we expect results at least as accurate. One should also note that the form of our transformation simplifies the analysis since $E(y^r)$ is readily available from $E(y)$, while with other transformations, this may not hold.

2.1. Expansion into Moments

Transformation (1.1) may be expanded as an infinite series in terms of x, either by the binomial or Taylor series expansion. We then have, using the binomial expansion,

$$y = \left\{ 1 + \left[a_1(x-\mu) + a_2(x-\mu)^2\right]\right\}^h$$

$$= 1 + \sum_{i=1}^{\infty} \frac{h^{(i)}}{i!}\left[a_1(x-\mu) + a_2(x-\mu)^2\right]^i$$

$$= 1 + \sum_{i=1}^{\infty} \frac{h^{(i)}}{i!} \sum_{j=0}^{i} \binom{i}{j} a_1^{\,i-j} a_2^{\,j} (x-\mu)^{i+j}$$

$$= 1 + \sum_{i=1}^{\infty} \frac{h^{(i)}}{i!} \sum_{m=i}^{2i} \binom{i}{m-i} a_2^{\,m-i} a_1^{\,2i-m} (x-\mu)^m$$

$$= 1 + \sum_{i=1}^{\infty} \sum_{m=i}^{2i} \binom{i}{m-i} \frac{m(m-1)\ldots(i+1)}{(h-m+1)\ldots(h-1)} a_2^{\,m-i} a_1^{\,2i-m} \frac{h^{(m)}(x-\mu)^m}{m!}$$

$$= 1 + a_1 h(x-\mu) + \sum_{m=2}^{\infty}\left\{a_1^m + \sum_{i=\left[\frac{m+1}{2}\right]}^{m-1} \binom{i}{m-i} \frac{m(m-1)\ldots(i+1)}{(h-m+1)\ldots(h-i)} a_1^{\,2i-m} a_2^{\,m-i}\right\} \frac{h^{(m)}(x-\mu)^m}{m!}, \tag{2.1}$$

where

$$h^{(m)} = h(h-1)(h-2)\ \ldots\ (h-m+1),$$

and $\left[\frac{m+1}{2}\right]$ is defined as the largest integer less than or equal to $\frac{m+1}{2}$. The expansion (2.1) may also be obtained using the Taylor series expansion and noting,

$$f^{(m)}(\mu) = a_1 f^{(m-1)}(\mu)(h-m+1) + 2a_2 f^{(m-2)}(\mu)\left\{(m-1)h - \frac{(m-1)(m-2)}{2}\right\}, \tag{2.2}$$

where

$$y = f(x) \text{ and } f^{(m)}(x) = \frac{d^m f}{dx^m}.$$

However, (2.1) does not converge for all x, and thus is only valid in some region R. This is important since in order to obtain the expected value of y, we must integrate throughout the ranges of x and y, including possibly some set where (2.1) is not valid. This problem shall be dealt with at length in Section 3, and, under certain conditions E(y) as defined in term by term integration of (2.1), will exist and be valid even though (2.1) does not converge for all x.

Integrating (2.1) term by term, we then find

$$E(y) = K_1(y) = 1 + \left(a_1^2 + \frac{2a_2}{h-1}\right)\frac{h^{(2)}}{2!}\mu_2 + \left(a_1^3 + \frac{6a_1a_2}{h-2}\right)\frac{h^{(3)}}{3!}\mu_3$$

$$+ \left(a_1^4 + \frac{12a_1^2a_2}{h-3} + \frac{12a_2^2}{(h-2)(h-3)}\right)\frac{h^{(4)}}{4!}\mu_4 + \dots \quad . \qquad (2.3)$$

We then use the recursive relation (1.3) with $\mu_1' = 0$ in order to solve for μ_i in terms of K_i, and obtain

$$\mu_2 = K_2 , \qquad \mu_3 = K_3 ,$$

$$\mu_4 = K_4 + 3K_2^2, \qquad \mu_5 = K_5 + 10K_3K_2 ,$$

$$\mu_6 = K_6 + 15K_4K_2 + 10K_3^2 + 15K_2^3 \dots \quad .$$

Substituting in (2.3), we find

$$E(y) = K_1(y) = 1 + \left(a_1^2 + \frac{2a_2}{h-1}\right)\frac{h^{(2)}}{2!}K_2 + \left(a_1^3 + \frac{6a_1a_2}{h-2}\right)\frac{h^{(3)}}{3!}K_3$$

$$+ \left(a_1^4 + \frac{12a_1^2 a_2}{h-3} + \frac{12a_2^2}{(h-2)(h-3)}\right) \frac{h^{(4)}}{4!} (K_4 + 3K_2^2)$$

$$+ \left(a_1^5 + \frac{20a_1^3 a_2}{h-4} + \frac{60a_1 a_2^2}{(h-3)(h-4)}\right) \frac{h^{(5)}}{5!} (K_5 + 10K_3K_2)$$

$$+ \left(a_1^6 + \frac{30a_1^4 a_2}{h-5} + \frac{180a_1^2 a_2^2}{(h-5)(h-4)} + \frac{120a_2^3}{(h-5)(h-4)(h-3)}\right) \frac{h^{(6)}}{6!} (K_6 + 15K_4K_2$$

$$+ 10K_3^2 + 15K_2^3) + \ldots \quad . \tag{2.4}$$

The rth moment of y, $E(y^r)$, is then derived by simply replacing h by rh in (2.4). The same comments concerning validity as with E(y), apply here to $E(y^r)$.

2.2. Expansion into Cumulants

We then obtain the cumulants of y. The simplest formulae to use are those supplied by (1.4), which are

$$K_2(y) = \mu_2'(y) - K_1^2(y)$$

$$K_3(y) = \mu_3'(y) - K_1(y)\mu_2'(y) - 2K_1(y)K_2(y)$$

$$K_4(y) = \mu_4'(y) - K_1(y)\mu_3'(y) - 3K_2(y)\mu_2'(y) - 3K_3(y)K_1(y)$$

$$K_5'(y) = \mu_5'(y) - K_1(y)\mu_4'(y) - 4K_2(y)\mu_3'(y) - 6K_3(y)\mu_2'(y) - 4K_4(y)K_1(y).$$

(2.5)

where $\mu_i'(y) = E(y^i)$. The algebra is elementary, since it merely consists in multiplying series together and collecting terms, however, it is very very long and tedious. As in (2.4) for simplicity the series expansions for the cumulants of y will be written out only to the point where the sum of the suffixes in the cumulants or in a_1 and a_2, equals six. Longer expansions will be needed in certain cases, and these are contained in Appendix A.

Then using (2.4) and (2.5), we have

$$
\begin{aligned}
K_2(y) = {} & a_1^2 h^2 K_2 \\
& + K_3 h^2 a_1 (a_1^2 (h-1) + 2a_2) \\
& + K_4 \frac{h^2}{12} [a_1^4 (h-1)(7h-11) + 36 a_1^2 a_2 (h-1) + 12 a_2^2] \\
& + K_2^2 \frac{h^2}{2} [a_1^4 (h-1)(3h-5) + 16 a_1^2 a_2 (h-1) + 4 a_2^2] \\
& + K_5 \frac{h^2 (h-1)}{12} a_1 [a_1^4 (h-2)(3h-5) + 4 a_1^2 a_2 (7h-11) + 36 a_2^2] \\
& + K_3 K_2 \frac{h^2 (h-1)}{3} a_1 [a_1^4 (h-2)(7h-12) + 3 a_1^2 a_2 (22h-35) + 84 a_2^2] \\
& + K_6 \frac{h^2 (h-1)}{360} [a_1^6 (h-2)(31h^2 - 132h + 137) + 150 a_1^4 a_2 (h-2)(3h-5) \\
& \quad + 180 a_1^2 a_2^2 (7h-11) + 360 a_2^3] \\
& + K_4 K_2 \frac{h^2 (h-1)}{12} [a_1^6 (h-2)(15h^2 - 64h + 67) + 2 a_1^4 a_2 (h-2)(109h-183)
\end{aligned}
$$

$$+12a_1^2a_2^2(51h-80) + 168a_2^3]$$

$$+ K_3^2 \frac{h^2(h-1)}{36} [a_1^6(h-2)(30h^2-129h+135) + 6a_1^4a_2(h-2)(73h-123)$$

$$+ 72a_1^2a_2^2(17h-27) + 360a_2^3]$$

$$+ K_2^3 \frac{h^2(h-1)}{6} [a_1^6(h-2)(7h^2-30h+32)+6a_1^4a_2(h-2)(17h-29)$$

$$+ 18a_1^2a_2^2(16h-25)+72a_2^3] \quad + \ldots \qquad (2.6)$$

$$K_3(y) = a_1^3h^3K_3$$

$$+ (K_4+2K_2^2) \frac{3h^3}{2} a_1^2 (a_1^2(h-1) + 2a_2)$$

$$+ K_5 \frac{h^3}{4} a_1 [a_1^4(h-1)(5h-7) + 24a_1^2a_2(h-1) + 12a_2^2]$$

$$+K_3K_2 \frac{3h^3}{2} a_1 [a_1^4(h-1)(7h-10) + 34a_1^2a_2(h-1) + 16a_2^2]$$

$$+ K_6 \frac{h^3}{8} [a_1^6(h-1)(6h^2-19h+15) + 10a_1^4a_2(h-1)(5h-7) + 72 a_1^2a_2^2(h-1)$$

$$+ 8a_2^3]$$

$$+ K_4K_2 \frac{h^3}{4} [a_1^6(h-1)(41h^2-131h+104) + 4a_1^4a_2(h-1)(86h-121)$$

$$+ 492 a_1^2a_2^2(h-1) + 48a_2^3] .$$

$$+ K_3^2 \frac{h^3}{4} [a_1^6(h-1)(28h^2-89h+71) + 6a_1^4a_2(h-1)(39h-55)$$

$$+ 336a_1^2a_2^2(h-1) + 40a_2^3]$$

$$+ K_2^3 \frac{h^3}{2} [a_1^6(h-1)(17h^2-55h+44) + 12a_1^4a_2(h-1)(12h-17)$$

$$+ 204a_1^2a_2^2(h-1)+16a_2^3] + \ldots \qquad (2.7)$$

$$K_4(y) = a_1^4h^4K_4$$

$$+ 2\ (K_5+6K_3K_2)\ h^4a_1^3(a_1^2(h-1)+2a_2)$$

$$+ (K_6 + 12K_4K_2)\ \frac{h^4}{6}\ a_1^2[a_1^4(h-1)(13h-17) + 60a_1^2a_2(h-1)+36a_2^2]$$

$$+ 6K_3^2h^4a_1^2[a_1^4(h-1)(3h-4) + 14a_1^2a_2(h-1) + 8a_2^2]$$

$$+ 4K_2^3h^4a_1^2[a_1^4(h-1)(4h-5) + 18a_1^2a_2(h-1) + 12a_2^2] + \ldots \qquad (2.8)$$

$$K_5(y) = a_1^5h^5K_5$$

$$+ 5(3\,K_6 + 3K_3^2 + 4K_4K_2)\ h^5a_1^4(a_1^2(h-1) + 2a_2) + \ldots \qquad (2.9)$$

We note that equations (1.14) are specific cases of equations (2.4), (2.6), (2.7), (2.8), and (2.9), with $a_2 = 0$ and $a_1 = K_1^{-1}$.

In contrast to the amount of algebra involved, the resulting expressions for $K_i(y)$ simplify notably as i gets larger. In fact, in order to obtain merely $a_1^5 h^5 K_5$, the first term of $K_5(y)$, we must simplify

$$\left(a_1^5 + \frac{20a_1^3a_2}{5h-4} + \frac{60a_1a_2^2}{(5h-4)(5h-3)}\right) \frac{5h(5h-1)(5h-2)(5h-3)(5h-4)}{5!}(K_5+10K_3K_2)$$

$$- \left(a_1^5 + \frac{20a_1^3a_2}{4h-4} + \frac{60a_1a_2^2}{(4h-4)(4h-3)}\right) \frac{4h(4h-1)(4h-2)(4h-3)(4h-4)}{5!}(K_5+10K_3K_2)$$

$$- \left(a_1^5 + \frac{6a_1^3a_2}{4h-2} + \frac{2a_1^3a_2}{h-1} + \frac{12a_1a_2^2}{(h-1)(4h-2)}\right) \frac{4h^2(h-1)(4h-1)(4h-2)}{2 \cdot 3!} K_2K_3$$

$$- \left(a_1^5 + \frac{2a_1^3a_2}{4h-1} + \frac{6a_1^3a_2}{h-2} + \frac{12a_1a_2^2}{(h-2)(4h-1)}\right) \frac{4h^2(h-1)(h-2)(4h-1)}{2 \cdot 3!} K_2K_3$$

$$- \left(a_1^5 + \frac{20a_1^3a_2}{h-4} + \frac{60a_1a_2^2}{(h-4)(h-3)}\right) \frac{h(h-1)(h-2)(h-3)(h-4)}{5!} (K_5 + 10K_3K_2)$$

$$- 4\left(a_1^5 + \frac{6a_1^3a_2}{3h-2}\right) \frac{3h^3(3h-1)(3h-2)}{3!} K_2K_3$$

$$- 4\left(a_1^5(h-1) + \frac{2a_1^3a_2(h-1)}{3h-1} + 2a_1^3a_2 + \frac{4a_1a_2^2}{3h-1}\right) \frac{3h^3(3h-1)}{2} K_3K_2$$

$$- 4K_5 \frac{h^2(h-1)}{12} \left(a_1^5(h-2)(3h-5) + 4a_1^3a_2(7h-11) + 36a_1a_2^2\right)$$

$$- 4K_3K_2 \frac{h^2(h-1)}{3} \left(a_1^5(h-2)(7h-12) + 3a_1^3a_2(22h-35) + 84a_1a_2^2\right)$$

$$- 6\left(a_1^5 + \frac{2a_1^3}{2h-1}\right) \frac{2h^4(2h-1)}{2} K_2K_3$$

$$- 6K_5 \frac{h^3}{4} \left(a_1^5(h-1)(5h-7) + 24a_1^3a_2(h-1) + 12a_1a_2^2\right)$$

$$- 6K_3K_2 \frac{3h^3}{2} (a_1^5(h-1)(7h-10) + 34a_1^3a_2(h-1) + 16a_1a_2^2)$$

$$- 8 (K_5 + 6K_3K_2) h^4(a_1^5(h-1) + 2a_1^3a_2) .$$

A very interesting result in the series expansions of $K_i(y)$ is that certain terms cancel out progressively. For example, the term containing K_2 vanished in $K_3(y)$ and higher cumulants of y, K_3 and K_2^2 vanish in $K_4(y)$ and higher cumulants of y, and so on. This shall form the subject of a theorem in Section 4.

We now come upon the problem of equating the higher cumulants of y to zero. In order to do this, however, we must assume an order of magnitude of the terms in the expansion of the cumulants. This problem shall be discussed in the final Section. The solutions shall be presented in Chapter III.

2.3. Validity of Expansions

The power series expansion of y as written in (2.1) has the radius of convergence

$$R = \{ x : |a_1(x-\mu) + a_2(x-\mu)^2| < 1\} .$$

This region R may be simplified to $R = R_1 \cup R_2$ where

$$R_1 = \left\{ x : a_1^2 \geq 4|a_2| \text{ and } \left[\frac{-a_1}{2a_2} - \frac{\sqrt{a_1^2 + 4|a_2|}}{2|a_2|}\right. < x - \mu < \right.$$

$$\frac{-a_1}{2a_2} - \frac{\sqrt{a_1^2 - 4|a_2|}}{2|a_2|}$$

$$\text{or } \frac{-a_1}{2a_2} + \frac{\sqrt{a_1^2 - 4|a_2|}}{2|a_2|} < x - \mu < \frac{-a_1}{2a_2} + \frac{\sqrt{a_1^2 + 4|a_2|}}{2|a_2|} \Bigg] \Bigg\},$$

$$R_2 = \Bigg\{ x : a_1^2 < 4|a_2| \text{ and } \Bigg[\frac{-a_1}{2a_2} - \frac{\sqrt{a_1^2 + 4|a_2|}}{2|a_2|} < x - \mu <$$

$$\frac{-a_1}{2a_2} + \frac{\sqrt{a_1^2 + 4|a_2|}}{2|a_2|} \Bigg] \Bigg\}.$$

In both cases, the region of convergence always contains the point $x = \mu$.

Neither Wilson and Hilferty, nor Haldane discuss this problem. However, Wallace (1958) in discussing the general δ-method for asymptotic moment expansions, finds that the expansions are known to be valid only in special cases. Such a case is when y, the transformed random variable is a function of sample moments, and is uniformly bounded by a power of the sample size, a theorem proved by Cramér (1963).

We will show that, under appropriate conditions, for any function $y = f(x)$, a valid asymptotic expansion of the rth moment of y, in terms of the moments of x, exists. Furthermore, this expansion is equal to the series obtained by integrating term by term the Taylor series expansion of y, regardless of the fact that the Taylor expansion might not converge for all x. The problem therefore will no longer be whether the radius of convergence R is large enough so that the difference between $\int_{x \in R} y^r dF$ and $\int_{\{x\}} y^r dF$ is small enough, but whether the moment series

expansion of y^r converges. Finally, we will produce a simplified condition for the convergence of the expansion of $E(y^r)$ in the case when y is defined as in (1.1).

Theorem 2.3.1. Let y be defined as in (1.1). Then $E|y^r|$ exists if and only if $E|x|^{2rh}$ exists.

Proof: Define $T = \{x: |x-\mu|<c\}$ for some constant c. Now if $x \in T$, we have

$$|Y|^{1/h} = |1 + a_1(x-\mu) + a_2(x-\mu)^2|$$

$$\leq 1 + |a_1|\ |x-\mu| + |a_2|\ |x-\mu|^2$$

$$1 + c|a_1| + c^2|a_2|$$

If $x \notin T$, and c is chosen such that both $|\frac{a_2}{a_1}|\ c > 1$ and $|a_2|c \geq 1$, then

$$|Y|^{1/h} = |a_2|\ |x-\mu|^2\ \left|1 + \frac{1}{a_2(x-\mu)/a_1} + \frac{1}{a_2(x-\mu)^2}\right|$$

$$\leq 3|a_2|\ |x-\mu|^2.$$

Silimarly, if $x \notin T$, and c is chosen such that $\frac{1}{2}|a_2|c^2 - |1 \pm a_1c| \geq 0$, then

$$|Y|^{1/h} \geq |a_2(x-\mu)^2| - |1+a_1(x-\mu)|$$

$$\geq \frac{1}{2} |a_2| \; |x-\mu|^2 \quad ,$$

and hence $|x-\mu|^2 \leq 2|a_2|^{-1} \; |Y|^{1/h}$

If $E|x|^{2rh}$ exists, we have

$$E|y|^r = \int_{x\varepsilon T} |y^r| dF + \int_{x\notin T} |y^r| dF$$

$$< (1+c|a_1|+c^2|a_2|)^{rh} \int_{x\varepsilon T} dF + 3^{rh} |a_2|^{rh} \int_{x\notin T} |x-\mu|^{rh} dF$$

which exists since each of the components are bounded. Conversely if $E|y|^r$ exists, we have

$$E|x-\mu|^{2rh} = \int_{x\varepsilon T} |x-\mu|^{2rh} dF + \int_{x\notin T} |x-\mu|^{2rh} dF$$

$$\leq \int_{x\varepsilon T} |x-\mu|^{2rh} dF + 2^{rh} |a_2|^{-rh} \int_{x\notin T} |y|^r dF.$$

Then since T is a bounded interval, and $E|y|^r$ is bounded, $E|x-\mu|^{2rh}$ and hence $E|x|^{2rh}$ exists.

<u>Theorem 2.3.2.</u> If $y = f(x)$ and its first $m + 1$ derivatives are continuous in a closed interval containing μ, $E|x|^m$ exists, and

$$R_m = \int_{\{x\}} R_m^* \, dF. < \infty \tag{2.10}$$

where

$$R_m^* = \frac{1}{m!}\int_{\mu}^{x} (x-t)^m g^{(m+1)}(t)dt$$

and where $g(x) = [f(x)]^r$, then

$$E(y^r) = g(\mu) + \sum_{i=2}^{m} g^{(i)}(\mu)\frac{\mu_i}{i!} + R_m \quad . \qquad (2.11)$$

Proof: The fact that $f(x)$ and hence $g(x)$, and its first $m + 1$ derivatives are continuous in a closed interval containing μ, ensure a Taylor series expansion of y^r with remainder. Then

$$E(y^r) = \int_{\{x\}} y^r dF$$

$$= \int_{\{x\}} \left\{ g(\mu) + \sum_{i=1}^{m} g^{(i)}(\mu) + \frac{(x-\mu)^i}{i!} R_m^* \right\} dF \ .$$

But since, by assumption, R_m^* is integrable and $E|x|^m$ exists, the result follows using term by term integration.

Theorem 2.3.3. Let $f(x)$, $g(x)$, and R_m be defined as above. If $E|x|^m$ exists for all m, and $R_m \to 0$ as $m \to \infty$, then

$$E(y^r) = g(\mu) + \sum_{i=2}^{\infty} g^{(i)}(\mu)\frac{\mu_i}{i!} \qquad (2.12)$$

Proof: The proof follows immediately from Theorem 2.3.2.

We note that (2.11) and (2.12) show that a moment expansion of y^r may be constructed, even though R_m^* may not exist for all x. The only condition that needs to be verified is that R_m (or in fact any other form of remainder) tends to zero for large m. This condition however merely ensures the convergence of the resulting moment expansion, and any other condition which implies convergence may be substituted in Theorem 2.3.2 and Theorem 2.3.3. We also note that the condition upon R_m (that it is bounded, or it tends to zero for large m), no longer involves x, as it did with R_m^*, but only upon the moments of x, and a_1, a_2, and h. The expansion (2.11) is particularly important with distributions where the higher moments do not exist as with the t-distribution.

<u>Corollary 2.3.1.</u> If $y = \{1 + a_1(x-\mu) + a_2(x-\mu)\}^h$, and $E|x|^m$ exists for all m, $a_1^2 \geq 4a_2$ and

$$\lim_{m \to \infty} \left| \frac{\mu_{m+1}}{\mu_m} \right| < \frac{2}{|a_1 \pm \sqrt{a_1^2 - 4a_2}|} \tag{2.13}$$

then
$$E(y^r) = 1 + \sum_{m=2}^{\infty} \left\{ a_1^m + \sum_{i=\left[\frac{m+1}{2}\right]}^{m-1} \binom{i}{m-1} \frac{m(m-1)\ldots(i+1)}{(rh-m+1)\ldots(rh-i)} a_1^{2i-m} a_2^{m-i} \right\} \times \frac{(rh)^{(m)}\mu_m}{m!} . \tag{2.14}$$

<u>Proof:</u> Since y^r and all its derivatives are continuous about μ, and (2.14) is merely (2.12) with $f(x) = y$ as defined above, then applying

Theorem 2.3.3., it suffices to show that (2.13) implies that R_m, as defined by (2.10), tends to zero as m tends to zero, or equivalently that (2.14) converges. Rewriting (2.2) and (2.14), we have

$$\frac{g^{(m+1)}(\mu)}{g^{(m)}(\mu)} = a_1(rh-m) + 2a_2 \frac{g^{(m-1)}(\mu)}{g^{(m)}(\mu)} \left(mrh - \frac{m(m-1)}{2}\right),$$

and

$$E(y^r) = 1 + \sum_{m=2}^{\infty} \frac{g^{(m)}(\mu)\mu_m}{m!} = 1 + \sum_{m=2}^{\infty} v_i ,$$

where $g(x) = y^r$. Then by d'Alembert's ratio test, if there exists the limit

$$t = \lim_{m \to \infty} \frac{v_{m+1}}{v_m}$$

where $|t|<1$, then (2.14) is absolutely convergent. If $t = 0$, the result follows immediately, so we assume t non-zero. Examining the ratio, we have

$$\frac{V_{m+1}}{V_m} = \frac{g^{(m+1)}(\mu)\mu_{m+1}}{(m+1)g^{(m)}(\mu)_m} = \left\{ a_1(rh-m)+2a_2\frac{g^{(m-1)}(\mu)}{g^{(m)}(\mu)}\left(mrh- \frac{m(m-1)}{2}\right)\right\}$$

$$\times \frac{\mu_{m+1}}{(m+1)\mu_m} .$$

Hence $\lim_{m \to \infty} \frac{V_{m+1}}{V_m} = t \in (-1,1)$

$$\Longleftrightarrow \lim_{m \to \infty} \frac{V_{m+1}}{V_m} = \lim_{m \to \infty} \frac{V_m}{V_{m-1}}$$

$$= \lim_{m \to \infty} \frac{g^{(m)}(\mu)\mu_m}{mg^{(m-1)}(\mu)\mu_{m-1}} = t \in (-1,1)$$

$$\Longleftrightarrow t = \lim_{m \to \infty} a_1 \frac{(rh-m)\mu_{m+1}}{(m+1)\mu_m} + \frac{2a_2}{t} \lim_{m \to \infty} \left(mrh - \frac{m(m-1)}{2}\right) \frac{\mu_{m+1}}{m(m+1)\mu_m} \in (-1,1) .$$

Solving for t we obtain

$$t = \lim_{m \to \infty} \left(\frac{a_1(rh-m)\mu_{m+1}}{2(m+1)\mu_m} \pm \frac{1}{2}\sqrt{a_1^2 \frac{(rh-m)^2\mu_{m+1}}{(m+1)^2\mu_m} + 8a_2\left(rh - \frac{(m-1)}{2}\right)\frac{\mu_{m+1}}{(m+1)\mu_{m-1}}} \right).$$

Taking limits and noting that $\lim_{m \to \infty} \frac{\mu_m}{\mu_{m-1}} = \lim_{m \to \infty} \frac{\mu_{m+1}}{\mu_m}$, we obtain

(2.13).

By the same argument, (2.14) diverges whenever

$$\lim_{n \to \infty} \left| \frac{\mu_{m+1}}{\mu_m} \right| > \frac{2}{|a_1 \pm \sqrt{a_1^2 - 4a_2}|}$$

Also, some very interesting properties are brought forth by (2.13). It shows that h does not in any way enter into the convergence of (2.14). The form of (2.13) also tells us something about the relative orders of magnitude of a_1 and a_2, and this will be discussed further in Section 5.

2.4. Cumulants of a Transformed Variate

As we mentioned in Section 2, certain terms in the cumulant expansion of $K_i(y)$, always seem to cancel, K_2 in $K_i(y)$ when $i \geq 3$, K_3 and K_2^2 in $K_i(y)$ when $i \geq 4$, and so on. We will now give a general proof of the result. First, however, we make the following definition. We define the *level* of the cumulant $K_i^{\alpha_i} K_j^{\beta_j} \ldots K_m^{\gamma_m}$ as the number $\alpha_i(i-1) + \beta_j(j-1) + \ldots + \gamma_m(m-1)$, ($\alpha_i, \beta_j, \gamma_m$ are integers greater or equal to zero) that is, the sum of the suffixes minus one, or equivalently the sum of the suffixes minus the sum of the exponents. Then for example K_2 is of level one and $K_4 K_3^2$ is of level seven. The definition has no apparent theoretical significance, but is used as a means of identifying and grouping certain cumulants together. In particular as can be seen in equations (2.7), (2.8) and (2.9), level one cumulants vanish in $K_3(y)$, levels one and two in $K_4(y)$, levels one to three in $K_5(y)$, and so on.

Theorem 2.4.1. Let R_m be defined as in (2.10). If $y = f(x)$ and its first $m + 1$ derivatives are continuous in a closed interval containing μ, $E|x|^m$ exists, and R_m is bounded, then $K_r(y)$ where $r \leq m$ does not contain, in its cumulant expansion, cumulants of x with level less than $(r-1)$.

Proof: Since y and its first $m + 1$ derivatives are continuous about μ, then y has a Taylor expansion about μ with remainder R_m^*,

$$y = c_0 + c_1(x-\mu) + c_2(x-\mu)^2 + \ldots + R_m^*.$$

Without loss of generality, we now assume that $\mu = 0$ and $c_o = 0$. We then write $y = \sum_{j=1}^{m} V_i + R_m^*$, where $V_j = c_j x^j$. Denoting $\mu'_{(j_1 \cdots j_r)}(V)$ for the moment $E(V_{j_1} \cdots V_{j_r})$, (some of the j_p may be equal) and $K_{(j_1 \cdots j_r)}(V)$ for the corresponding cumulant, we have, using (1.2)

$$\frac{(it)K_1(y)}{1!} + \frac{(it)^2 K_2(y)}{2!} + \ldots + \varepsilon_1 = \log\left[1 + \frac{(it)\mu_1'(y)}{1!} + \frac{(it)^2\mu_2'(y)}{2!} + \ldots + \varepsilon_2\right]$$

$$= \log\left[1 + \sum_{p=1}^{m} (it_p)\mu_p'(V) + \frac{1}{2!}\sum_{p,q=1}^{m} (it_p)(it_q)\mu'_{(pq)}(V) + \ldots + R_\varepsilon + \varepsilon_2\right]$$

$$= \sum_{p=1}^{m} (it_p)K_p(V) + \frac{1}{2!}\sum_{p,q=1}^{m} (it_p)(it_q)K_{(pq)}(V) + \ldots + R_\varepsilon^* + \varepsilon_2^*,$$

where $t_1 = t_2 = \ldots = t_m = t$ and $\varepsilon_1, \varepsilon_2 = o(t^m)$.

Then identifying coefficients, we have for $r \leq m$,

$$K_r(y) = \sum_{p_1 \cdots p_r = 1}^{r} K_{(p_1 \cdots p_r)}(V). \tag{2.15}$$

It is therefore sufficient to show that $K_{(p_1 \cdots p_r)}(V)$ does not contain, in its cumulant expansion, cumulants of x with level less than (r-1). To show this, suppose that $x_1, \ldots, x_s$ is a sample of n independent values from the distribution of x, and let $V_p = c_p(\sum_{j=1}^{s} x_j)^p$. This coincides with our earlier definition of V_p when $s = 1$ and $x_1 = x$. Then V_p can be regarded as statistics of the sample $x_1, \ldots x_n$, and hence, each V_p is a homogeneous polynomial symmetric function of the sample values of degree p.

However, James (1955a) has shown that Fisher's rules for obtaining the sampling cumulants of the k-statistics, may be adapted to any system of statistics $V_1, V_2 \ldots$ (if V_p is a homogeneous polynomial symmetric function of degree p in the observations), by simply replacing the column factors occurring in the evaluation of the patterm functions by the coefficient of $x_1^{\beta_1} \ldots x_\alpha^{\beta_\alpha}$ in V_p, ($\sum_{j=1}^{\alpha} \beta_j = p$).

We then apply Fisher's third rule. It states that in finding the cumulants of the k (or in our case the V-) statistics in terms of the population cumulants (in our case x), we are to neglect every array in which the numbers in the body of the array fall into two or more blocks each confined to separate rows or columns. Now, any array which does not fall into blocks each confined to separate rows of columns, may be built up column by column in such a way that at each stage, the new column does not form an isolated block by itself. Then if $\alpha_1, \ldots, \alpha_r$ are the numbers of non-zero elements in the $1^{st}, \ldots, r^{th}$ columns, and γ is the total number of rows in the array, we have

$$\gamma \leq \alpha_1 + (\alpha_2 - 1) + \ldots + (\alpha_r - 1) = \sum_{j=1}^{r} \alpha_j - r + 1. \tag{2.16}$$

Now, if this array is one of those contributing to the coefficient of $K_{t_1} \ldots K_{t_\gamma}$ in $K_{(p_1 \ldots p_r)}(V)$, where $\sum_{j=1}^{\gamma} t_j = \sum_{j=1}^{r} p_j$ according to rule one, then the corresponding term is of level $\sum_{j=1}^{\gamma} (t_j - 1) = \sum_{j=1}^{r} p_j + \gamma$, since each K_{t_j} is of level $(t_j - 1)$. But (2.16) shows that

$$\sum_{j=1}^{r} p_j + \gamma \geq \sum_{j=1}^{r} (p_j - \alpha_j) + r - 1 \geq r - 1,$$

since α_j is the number of parts in the partition of r_j and so does not exceed r_j. Therefore, this particular term of $K_{(p_1 \ldots p_r)}(V)$ has cumulants of x of level r - 1 or higher and hence the same is true of $K_{(p_1 \ldots p_r)}(V)$ itself. Using (2.15), the result follows.

The proof given above, is very similar to the proof of a theorem on the order of $K_i(y)$ by James (1955b). We however use the fact that the rules for Fisher's k-statistics do not presume any order on any of the variables, so that we have terms cancelling independant of the order. James' theorem then becomes a special case of the above theorem, and an important corollary.

Corollary 2.4.1. Let R_m be defined as in (2.10). If $y = f(x)$ and all its derivatives are continuous in a closed interval containing μ, K_r exists and is of order n^{1-r}, for some dominant parameter n, and R_m is bounded, then $K_r(y)$ is of order n^{1-r} or lower.

Proof: The proof follows immediately from the fact that level i is synonymous with order n^{-i} in this case. Hence the remaining terms in $K_r(y)$ of level r-1 or higher have order n^{1-r} or less.

2.5. Problem of Order

As mentioned earlier, the method by which we transform the random variable x to an approximately normal random variable y, is by setting, as much as possible, the higher cumulants of y equal to zero. There are several methods of doing this. Since the expansions of $K_3(y), K_4(y)$ and $K_5(y)$ are polynomial equations in the three parameters a_1, a_2, and h, it is theoretically possible to arrive at some solution, say a_1^*, a_2^*, and h^*, by use of Newton's method (or equivalently by a superior numerical technique by Brown (1969), a quadratically convergent Newton-like method based on Gaussian elimination). However, not only are the cumulant expansions of $K_i(y)$ infinite series in the cumulants of x, but also the amount of algebra needed to calculate more terms than we have done (as in Appendix A) is overwhelming. Even if we truncate the series at some manageable level, we must assume a certain order upon the terms, or at least, we must decide which terms are larger than the others, so that they can be included, and the others, whose magnitude is at most equal to the truncated terms, may be discarded.

There is one way, however, to avoid assigning an order of magnitude to terms. This is using Brown's method where the non-linear equations $K_3(y)$, $K_4(y)$, $K_5(y)$ are calculated using direct numerical integration of

$$E(Y^r) = \int_{\{x\}} y^r dF .$$

This method however will rely very heavily upon the simplicity of the

distribution function, and, in almost every case, the amount of computer time needed is enough to prohibit its use. Furthermore, in both cases cited above, unless a very accurate initial guess is known, the routine for solving the non-linear equations will not converge. Also, as with other non-linear methods, each particular case will demand a separate solution, and thus will generally be impractical. Another setback is that the accuracy gained will probably be wasted because of the errors caused by the interplay of $K_6(y)$ and the other higher cumulants. Finally, none of these methods guarantee the possibility of a noncomplex solution.

If we however limit our scope, and assume a certain order of magnitude upon the terms, we may then group together terms of equal orders of magnitude (in $K_3(y)$, $K_4(y)$ and $K_5(y)$) and then set three of them equal to zero. That is, we may write,

$$K_3(y) = b(a_1,a_2,h) + d(a_1,a_2,h) + g(a_1,a_2,h) + \ldots$$

$$K_4(y) = c(a_1,a_2,h) + f(a_1,a_2,h) + \ldots$$

$$K_5(y) \doteq e\,(a_1,a_2,h) + \ldots$$

where b, d, g, c, f, and e, functions of a_1, a_2, and h are the grouped terms of descending orders of magnitude. The choice of which three functions are to be set equal to zero, is based upon the orders of the grouped terms, and the relative importance of $K_3(y)$, $K_4(y)$ and $K_5(y)$, in the distribution of y. Hence two choices seem to promise the best

results, b, c, e, and b, d, c. At this point, one may obtain better results by using the solution of a_1, a_2 and h, as an initial guess and using either of the two methods described above. However, the same setbacks apply, and usually the work involved is not compensated by the small accuracy gained.

The fact that an order of magnitude must be assumed is readily seen from Corollary 2.3.1. and the comments thereafter. For example, if x is χ^2 with n degrees of freedom, $\left(K_i = n\, 2^{i-1}(i-1)!\right.$, and hence $\mu_i = 0\left(\left[\frac{i}{2}\right]_n\right)\Big)$ and we assume that a_1, a_2, and h are constants, then as is seen by Corollary 2.3.1., the expansion of $E(y^r)$ does not converge, and hence $K_i(y)$ do not exist. Even if we were to divide by the standard deviation, $\left(K_i(\frac{x}{\sigma}) = 0(n^{1-i/2})\right.$ and hence $\mu_i(\frac{x}{\sigma}) = 0(1)$ if i is even and $\left.0(n^{-1/2})\right.$ if i is odd$\Big)$ the expansion of $E(y^r)$ may or may not converge. We should note here that, as shown earlier, the result remains unchanged if x is divided by any number. What does change the result, is the different grouping of terms which a different order of a_1, a_2, h and K_r might imply.

What is important to note at this point, is the fact that the problem of the orders of magnitude of the terms in the expansions is essentially one of choice. Different orders of a_1, a_2, and h imply and is implied by a different grouping of terms (as will be seen more explicitly later). Corollary 2.3.1. gives a necessary condition and a good indication of what orders our parameters should be. First, since the convergence of $E(y^r)$ is independant of h, we have that $h = 0(1)$, as otherwise h would affect the convergence. Secondly since both a_1 and $\sqrt{a_2}$ appear simultaneously in the bound (2.13), it indicates that $0(a_2) = 0(a_1^2)$. If this is not so, then either a_1 or a_2 would be the

dominant parameter in the expansion, which would imply results very similar to the Haldane B transformation, since then either $Y = (1 + a_1(x-\mu) + \varepsilon)^h$ or $Y = (1 + a_2[(x-\mu)^2] + \varepsilon)^h$. The interplay between $(x-\mu)$ and $(x-\mu)^2$ is at its fullest, and hence the results are best, only when one term does not over-shadow the other, that is when a_1^2 and a_2 are of the same order.

Thus, if n is some "dominant" parameter of the distribution, we must make an assumption of the form

$$\left.\begin{array}{lll} a_1^p a_2^q K_i = O(n^{(1-ti)\delta}) & & p + 2q = i,\ i = 2, 3, 4, 5, \delta > 0 \\ \text{and} \quad a_1^p a_2^q K_i \leq O(n^{(1-6t)\delta}) & & i = 6, 7, \ldots \end{array}\right\} \quad (2.17)$$

Another equivalent way of stating the problem is, by what power of n must we divide x, and thus K_i, such that the expansions contain terms with orders as in (2.17). From now on, for the sake of simplicity, we assume $\delta = 1$, and we will omit the factors $a_1^p a_2^q$ when referring to the term containing K_i which is of order $O(n^{(1-ti)\delta})$. We have $K_2^2 = O(n^{2-4t})$, $K_3 K_2 = O(n^{2-5t})$, $K_2^3 = O(n^{3-6t})$. The cumulants with their orders are displayed, for different values of t, in Table 2.1. We see immediately from Table 2.1, that in the case $t = 1/2$, not only does it not converge in certain cases, but also, a grouping of terms of equal orders of magnitude is impossible. Also, in all other cases except $t = 1$, each cumulant has a different order and then technically, if all the larger orders would be set to zero, we would obtain the degenerate case $a_1 = a_2 = h = 0$. This problem may be partially solved by grouping the orders themselves, "in some way", and then solving.

Finally, Theorem 2.4.1. and Corollary 2.4.1, also give us an indication of the orders of a_1, a_2, and h. The cancellation of terms is most efficient (that is, the terms with the largest orders are first to cancel completely) only when $t = 1$, $O(a_1^2) = O(a_2)$, and $h = O(1)$. It thus seems as (2.17) with $t = 1$, is the best choice for the order of a_1, a_2, and h. One should note from Table 2.1, that small deviations from $t = 1$, are irrelevant since, with proper grouping, similar equations will be derived.

We should finally realize that the assumption of order is not necessarily needed. In fact, all that is needed is some assumption with regards to the relative magnitude of the terms. It should also be pointed out, that both Haldane and Wilson and Hilferty, implicitly made the assumption (2.17) in their arguments.

$t = 1/2$

Orders	Cumulants
n^{0}	$K_2, K_2^2, K_2^3, \ldots$
$n^{-1/2}$	$K_3, K_3 K_2^i, \ldots$
n^{-1}	$K_4, K_3^2, K_4 K_2^i, \ldots$
$n^{-3/2}$	$K_5, K_4 K_3, K_3^3, K_5 K_2^i, \ldots$

$\forall i$

$t = 3/4$

Orders	Cumulants
$n^{-1/2}$	K_2
n^{-1}	K_2^2
$n^{-5/4}$	K_3
$n^{-3/2}$	K_2^3
$n^{-7/4}$	$K_3 K_2$
n^{-2}	K_2^4
$n^{-5/2}$	K_3^2

$t = 1$

Orders	Cumulants
n^{-1}	K_2
n^{-2}	K_3, K_2^2
n^{-3}	$K_4, K_2 K_3, K_2^3$
n^{-4}	$K_5, K_4 K_2, K_3^2, K_3 K_2^2, K_2^4$

$t = 2$

Orders	Cumulants
n^{-3}	K_2
n^{-5}	K_3
n^{-6}	K_2^2
n^{-7}	K_4
n^{-8}	$K_2 K_3$
n^{-9}	K_5, K_2^3

TABLE 2.1

CHAPTER III

GENERAL SOLUTION FOR CONSTANTS

In this Chapter we will derive a general solution for the constants a_1, a_2, and h. We make the assumption that t = 1 in the ordering condition (2.17), which becomes (δ being set equal to one without loss of generality)

$$\left.\begin{array}{lll} a_1^i K_\ell = O(n^{1-\ell}) & & i + 2j = \ell, i = 2, 3, 4, 5 \\ a_1^i a_2^j K_\ell \leq O(n^{-5}) & & i = 6, 7, \ldots \end{array}\right\} \quad (3.1)$$

This general type of assumption is necessary, as noted in the previous chapter. A different assumption would lead to a different grouping of terms, and hence a different solution. However, as noted earlier, this seems to be the most logical assumption to make. For example, in the case that x is χ^2 with n degrees of freedom, $K_i = O(n)$, and then $a_1 = O(n^{-1})$ and $a_2 = O(n^{-2})$. As in the previous chapter, we will for the sake of simplicity omit the constants $a_1^i a_2^j$ when referring to the order of the term $a_1^i a_2^j K_\ell$. We also assume n to be some dominant parameter in the distribution of x.

Table 3.1. will then give the orders of the terms (which are functions of the cumulants of x) in the expansions of the cumulants of y. This enables us to make a precise grouping of the terms in the expansions

Order	Cumulants
n^{-1}	K_2
n^{-2}	K_3, K_2^2
n^{-3}	K_4, K_2K_3, K_2^3
n^{-4}	$K_5, K_4K_2, K_3^2, K_3K_2^2, K_2^4$
n^{-5}	$K_6, K_5K_2, K_4K_3, K_4K_2^2, K_3K_2^2, K_3^2K_2, K_2^5$

Table 3.1

of the cumulants y, which may be expressed as follows, in descending sequence of powers of n,

$$\left.\begin{aligned} K_3(y) &= b(n^{-2}) + d(n^{-3}) + g(n^{-4}) + 0(n^{-5}) \\ K_4(y) &= \qquad\quad c(n^{-3}) + f(n^{-4}) + 0(n^{-5}) \\ K_5(y) &= \qquad\qquad\qquad\quad e(n^{-4}) + 0(n^{-5}) \end{aligned}\right\} \quad (3.2)$$

where b, c, d, e, f, and g are functions of a_1, a_2, and h and whose orders are those indicated in brackets.

We must then choose which three functions to set equal to zero, thereby solving for a_1, a_2, and h. Heuristically, setting b, c, and e equal to zero (which will henceforth be referred to as case A), and setting b, c, and d equal to zero (henceforth referred to as case B), seem to be the best choice. This is because the lower cumulants of y,

that is those cumulants which have the most effect on the distribution of y, are reduced to lower orders.

A further rationale for case A and case B may be derived by considering the Cornish-Fisher expansion y. Standardizing y, we have

$$y^* = \frac{y - K_1(y)}{\sqrt{K_2(y)}}, \tag{3.3}$$

and then

$$K_1(y^*) = 0, \quad K_i(y^*) = \frac{K_i(y)}{K_2^{i/2}(y)} = 0(n^{1-i/2}),$$

for $i = 2, 3, 4, \ldots$. Setting $\ell_1 = \ell_2 = 0$, and $\ell_i = K_i(y^*)$, for $i = 3, 4, \ldots$, we have from (1.8), when z is a standard normal random variable,

$$\begin{aligned} y^* - z &= \frac{1}{6} K_3(y^*)(y^{*2}-1) \\ &+ \frac{1}{24} K_4(y^*)(y^{*3}-3y^*) - \frac{1}{36} K_3^2(y^*)(4y^{*3}-7y^*) \\ &+ \frac{1}{120} K_5(y^*)(y^{*4}-6y^{*2}+3) - \frac{1}{144} K_3(y^*)K_4(y^*)(11y^{*4}-42y^{*2}+15) \\ &+ K_3^3(y^*)(69y^{*4}-187y^{*2}+52) \qquad (3.4) \\ &+ 0(n^{-2}). \end{aligned}$$

We then note that setting $b = 0$ we have $K_3(y^*) = 0(n^{-3/2})$, setting $c = 0$ we have $K_4(y^*) = 0(n^{-2})$, setting $e = 0$ we have $K_5(y^*) = 0(n^{-5/2})$, and setting both b and d equal to zero we have $K_3(y^*) = 0(n^{-5/2})$. Then with case A we have

$$y^* - z = 0(n^{-3/2}),$$

and with case B, we have

$$y^* - z = 0(n^{-3/2}).$$

The equation (3.4) suggests a further approximation in the sense that it provides us with successive refinements of y^*, as defined by (3.3). For example, in case A,

$$y^* - z - \frac{1}{6} K_3(y^*)(y^{*2}-1) = 0(n^{-2}) , \tag{3.5}$$

and in case B,

$$y^* - z - \frac{1}{120} K_5(y^*)(y^{*4}-6y^{*2}+3) = 0(n^{-2}) . \tag{3.6}$$

Even better refinements may be arrived at, by including in (3.4) terms of higher orders as in (1.8).

Section 1 will be devoted to the solution of case A, and Section 2 of case B. In Section 3, we will consider the special cases

(denoted as case C) when $a_2 = 0$, and which will include the Haldane A and B transformations found in Chapter I. Another case (case D) has also been considered. This consists of solving numerically the simultaneous equations

$$\left.\begin{aligned} b(n^{-2}) + d(n^{-3}) + g(n^{-4}) &= 0 \\ c(n^{-3}) + f(n^{-4}) &= 0 \\ e(n^{-4}) &= 0 \end{aligned}\right\} \qquad (3.7)$$

by a Newton-like method based on Gaussian elimination, by Brown (1969), using as initial guesses the solutions obtained from cases A,B, and C. The numerical results are all contained in Chapter 5.

3.1. Solution of Case A

Expressing the simultaneous equations of case A in full we have

$$b = a_1K_3 + 3K_2^2\,(a_1^2(h-1) + 2a_2) = 0 \qquad (3.8)$$

$$c = a_1^2K_4 + 12K_3K_2a_1(a_1^2(h-1)+2a_2) + 4K_2^3(a_1^4(h-1)(4h-5)+18a_1^2a_2$$

$$(h-1)+12a_2^2) = 0 \qquad (3.9)$$

$$e = a_1^3 K_5 + 5(3K_3^2 + 4K_4K_2)\, a_1^2(a_1^2(h-1) + 2a_2)$$

$$+ 30K_3K_2^2 a_1 (a_1^4(h-1)(5h-6) + 22a_1^2 a_2(h-1) + 16a_2^2)$$

$$+ 5K_2^4\, (a_1^6(h-1)(25h^2-65h+42) + 36a_1^4 a_2(h-1)(5h-6) + 300a_1^2 a_2^2$$

$$(h-1) + 96a_2^3) = 0 . \qquad (3.10)$$

The best way to solve (3.8), (3.9), and (3.10) analyticaly seems to be solving for a_2 in (3.8), for h in (3.9), and last for a_1 in (3.10). We obtain from solving for a_2 in (3.8),

$$a_2 = -\frac{a_1}{2}\left\{ a_1(h-1) + \frac{K_3}{3K_2^2} \right\} . \qquad (3.11)$$

Substituting (3.11) into (3.9), e then simplifies to

$$a_1^2\,(h-1)(2h-1) + a_1(h-1)\frac{K_3}{K_2^2} + \frac{2K_3^2}{3K_2^4} - \frac{K_4}{4K_2^3} = 0. \qquad (3.12)$$

This is a quadratic equation in h which has the solution

$$h = 1 + \frac{-(a_1 + \frac{K_3}{K_2^2}) \pm \frac{1}{K_2^2}\sqrt{(K_2^2 a_1 + K_3)^2 - \frac{2}{3}(8K_3^2 - 3K_4K_2)}}{4a_1} . \qquad (3.13)$$

We now substitute (3.11) into (3.10), and we obtain

$$\frac{K_5}{5K_2^4} - \frac{K_3^3}{K_2^6} - \frac{4K_3K_4}{3K_2^5} - \frac{6K_3}{K_2^2}\left(a_1^2(h-1)(2h-1) + a_1(h-1)\frac{K_3}{K_2^2} - \frac{4K_3^2}{9K_2^4}\right)$$

$$-\left(a_1^3(h-1)(h-3)(2h-1) - 2a_1^2(h-1)\frac{K_3}{K_2^2}(4h-1) - \frac{13a_1(h-1)K_3^2}{3K_2^4}\right.$$

$$\left. + \frac{4K_3^3}{9K_2^6}\right) = 0 .$$

We first use (3.12) to reduce the power of h and then we replace h by the expression (3.13) to obtain

$$a_1^2\left\{\frac{K_3}{8K_2^2}\right\} + a_1\left\{-\frac{4K_3^2}{3K_2^4} + \frac{9K_4}{16K_2^3}\right\} + \left\{\frac{151K_3^3}{72K_2^6} - \frac{79K_3K_4}{48K_2^5}\right.$$

$$\left. + \frac{K_5}{5K_2^4}\right\} \pm \sqrt{(K_2^2a_1+K_3)^2 - \tfrac{2}{3}(8K_3^2-3K_4K_2)}$$

$$\left\{-\frac{K_3}{8K_2^4}a_1 + \frac{K_3^2}{8K_2^6} - \frac{K_4}{16K_2^5}\right\} = 0 . \qquad (3.14)$$

Multiplying (3.14) by its conjugate, we obtain

$$a_1^3\left\{-\frac{K_3^3}{3K_2^6} + \frac{K_4K_3}{8K_2^5}\right\}$$

$$+ a_1^2 \left\{ \frac{29K_3^4}{12K_2^8} - \frac{47K_4K_3^2}{24K_2^7} + \frac{5K_4^2}{16K_2^6} + \frac{K_5K_3}{20K_2^6} \right\}$$

$$+ a_1 \left\{ - \frac{311K_3^5}{54K_2^{10}} + \frac{995K_4K_3^3}{144K_2^9} - \frac{121K_4^2K_3}{64K_2^8} - \frac{8K_5K_3^2}{15K_2^8} + \frac{9K_5K_4}{40K_2^7} \right\}$$

$$+ \left\{ \frac{1447K_3^6}{324K_2^{12}} - \frac{3025K_4K_3^4}{432K_2^{11}} + \frac{397K_4^2K_3^2}{144K_2^{10}} + \frac{151K_5K_3^3}{180K_2^{10}} - \frac{79K_5K_4K_3}{120K_2^9} - \frac{K_4^3}{128K_2^9} + \frac{K_5^2}{25K_2^8} \right\}$$

$$= 0 \quad , \tag{3.15}$$

which is a cubic in a_1 since the coefficient of a_1^4 is zero, and may thus be solved analytically using Cardan's formulae.

We note that (3.11),(3.13), and (3.15) imply six distinct solutions of a_1, a_2, and h. However, three of these are extraneous due to the fact that (3.14) was multiplied by its conjugate. In practice, therefore, one must verify that $e = 0$ for each solution. To each solution of a_1, corresponds only one value of h, and thus the number of solutions will depend upon the number of real roots of (3.15) or equivalently, the sign of the discriminant of (3.15).

3.2. Solution of Case B

The first two simultaneous equations of case B are equivalent to those of case A. The third expression then becomes

$$d = 3K_4a_1^2\{(a_1^2(h-1) - 2a_2\} + 3K_3K_2a_1\{a_1^4(h-1)(7h-10) + 34a_1^2a_2$$

$$(h-1) + 16a_2^2\}$$

$$+ K_2^3\{a_1^6(h-1)(17h^2-55h+44) + 12a_1^4a_2(h-1)(12h-17) + 204a_1^2a_2^2$$

$$(h-1) + 16a_2^3\} = 0 \quad . \qquad (3.16)$$

Our solutions of a_2 and h are therefore the same as in case A, that is (3.11) and (3.13).

In order to solve for a_1 we employ the same technique as in case A. We obtain

$$\frac{K_4K_3}{K_2^2} + 3K_3K_2\left\{3a_1^2(h-1)(2h-1) + \frac{3a_1(h-1)K_3}{K_2^2} - \frac{4K_3^2}{9K_2^4}\right\}$$

$$+ K_2^3\left\{3a_1^3(h-1)(h-3)(2h-1) - \frac{2a_1^2(h-1)K_3(4h+1)}{K_2^2}\right.$$

$$\left. - \frac{5a_1(h-1)K_3^2}{K_2^4} + \frac{2K_3^3}{27K_2^6}\right\} = 0$$

by substituting (3.11) into (3.16). Again, we use (3.12) to reduce the power of h and then we substitute for h to obtain

$$a_1^2\left\{\frac{3K_3}{8K_2^2}\right\} + a_1\left\{-\frac{4K_3^2}{K_2^4}+\frac{27K_4}{16K_2^3}\right\} + \left\{\frac{695K_3^3}{216K_2^6} - \frac{27K_3K_4}{16\ K_2^5}\right\}$$

$$\pm\sqrt{(K_2^2a_1+K_3)^2 - \frac{2}{3}\ (8K_3^2-3K_4K_2)}\left\{-a_1\frac{3K_3}{8K_2^4} + \frac{3K_3^2}{8K_2^6} - \frac{3K_4}{16K_2^5}\right\} = 0, \tag{3.17}$$

which we then multiply by its conjugate to obtain

$$a_1^3\left\{-\frac{3K_3^3}{K_2^6} + \frac{9K_4K_3}{8K_2^5}\right\}$$

$$+\ a_1^2\left\{\frac{175K_3^4}{9K_2^8} - \frac{243K_3^2K_4}{16K_2^7} + \frac{45K_4^2}{16K_2^6}\right\}$$

$$+\ a_1\left\{-\frac{1471K_3^5}{54K_2^{10}} + \frac{413K_3^3K_4}{16K_2^9} - \frac{387K_4^2K_3}{64K_2^8}\right\}$$

$$+\left\{\frac{15983K_3^6}{1458K_2^{12}} - \frac{47K_3^4K_4}{4K_2^{11}} + \frac{105K_3^2K_4^2}{32K_2^{10}} - \frac{9K_4^3}{128K_2^9}\right\} = 0, \tag{3.18}$$

which again is a cubic in a_1, and is solved using Cardan's formulae.

Just as in case A, the equations (3.11), (3.13), and (3.18) give six different solutions of which three are extraneous since (3.17) was multiplied by its conjugate and we are again compelled to verify $d = 0$ for each solution. We finally note that in the case that $K_5 = 0$ (3.15) becomes (3.18), and hence solutions A and B are equivalent.

3.3. Solution of Case C

Case C involves the solution of

$$\left.\begin{aligned} b(n^{-2}) + d(n^{-3}) &= 0 \\ c(n^{-3}) &= 0 \end{aligned}\right\} \quad (3.19)$$

a limited version of (3.3) or (3.2), in the special case when $a_2 = 0$. We consider the following three subcases, C1, C2, and C3. Subcase C1 involves setting b and c equal to zero and solving for a_1 and h. This is, in effect, Haldane's Transformation B. In subcase C2, we solve for a_1 and h by setting b and d equal to zero. Finally, subcase C3 is a special case of C1 and C2 when $a_1 = K_1^{-1}$, and thus involves solving for h by setting b equal to zero.

3.3.1. Subcase C1

We solve the simultaneous equations

$$K_3 + 3K_2^2\, a_1(h-1) = 0$$

$$K_4 + 12K_3K_2a_1(h-1) + 4K_2^3a_1^2\,(h-1)(4h-5) = 0,$$

from which we obtain the solution

$$h = 1 - \frac{K_3}{3K_2^2 a_1} \qquad a_1 = \frac{5K_3}{3K_2^2} - \frac{3K_4}{4K_3K_2}$$

or

$$a_1 = \frac{K_3}{3K_2^2(h-1)} \qquad h = \frac{9K_4K_2 - 16K_3^2}{9K_4K_2 - 20K_3^2} . \tag{3.20}$$

which is identical to Haldane's solution (1.16).

3.3.2. Subcase C2

Setting b and d equal to zero, we have

$$K_3 + 3K_2^2 a_1 (h-1) = 0$$

$$3K_4 + 3K_3K_2a_1(7h-10) + K_2^3a_1^2\,(17h^2 - 55h + 44) = 0.$$

This leads to the solution

$$\left.\begin{aligned} h &= 1 - \frac{K_3}{3K_2^2 a_1} \\ a_1 &= \frac{K_3}{6K_2^2} \pm \frac{1}{2}\sqrt{\frac{95K_3^2}{27K_2^4} - \frac{2K_4}{K_2^3}} . \end{aligned}\right\} \tag{3.21}$$

3.3.3. Subcase C3

This is the special case of C1 and C2 when $a_1 = \frac{1}{K_1}$.

We then set b equal to zero to obtain,

$$K_3 + 3 \frac{K_2^2}{K_1} (h-1) = 0,$$

which yeilds the solution

$$h = 1 - \frac{K_3 K_1}{3K_2^2} . \tag{3.22}$$

This is identical to the solution of Haldane's transformation A. In the case x is χ^2, (3.22) becomes the Wilson-Hilferty transformation, which is in turn used to approximate a host of other distributions which can be related to it. Also, it becomes the Sankaran (1959) approximation when x is non-central χ^2.

3.4. Conclusion

We have thus solved for the constants a_1, a_2 and h in our transformation (1.1) in the general case (for a large class of random variables). We therefore have that y is almost normally distributed with mean $K_1(y)$ and variance $K_2(y)$. However, $K_1(y)$ and $K_2(y)$ are rather complicated expressions as shown in (2.4) and (2.6), and thus any practical application would demand a simplification of these expressions to a manageable level.

We then have from (2.6),

$$
\begin{aligned}
K_2(y) = {} & a_1^2 h^2 K_2 \\
& + K_3 h^2 (a_1^2 (h-1) + 2a_2) + K_2^2 \frac{h^2}{2} (a_1^4 (h-1)(3h-5) + 16 a_1^2 a_2 (h-1) + 4a_2^2) \\
& + K_4 \frac{h^2}{12} (a_1^4 (h-1)(7h-11) + 36 a_1^2 a_2 (h-1) + 12 a_2^2) \\
& + K_3 K_2 \frac{h^2 (h-1)}{3} a_1 (a_1^4 (h-2)(7h-12) + 3a_1^2 a_2 (22h-35) + 84 a_2^2) \\
& + K_2^3 \frac{h^2 (h-1)}{6} (a_1^6 (h-2)(7h^2-30h+32) + 6a_1^4 a_2 (h-2)(17h-29) \\
& \qquad + 18 a_1^2 a_2^2 + (16h-25) + 72 a_2^3) \\
& + 0(n^{-4})
\end{aligned}
$$

which when utilizing (3.11) and (3.12) of cases A and B, simplifies to

$$
\begin{aligned}
K_2(y) = h^2 \Bigg\{ & a_1^2 \left[K_2 - \frac{K_4}{4K_2} + \frac{7K_3^2}{18K_2^2} - \frac{K_4}{12} \left(\frac{K_4}{K_2^3} - \frac{3K_3^2}{K_2^4} \right) \right] \\
& - K_2^3 \frac{(h-1)}{6} \Big[a_1^6 (h-2)(35h^2-84h+40) + 6a_1^4 a_2 (63h^2-160h+95) \\
& + 18 a_1^2 a_2^2 (56h-73) + 936 a_2^3 \Big] \Bigg\} \\
& + 0(n^{-4})
\end{aligned}
$$

or less accurately

$$K_2(y) = a_1^2 h^2 \left[K_2 - \frac{K_4}{4K_2} + \frac{7K_3^2}{18K_2^2} - \frac{K_4^2}{12K_2^3} + \frac{K_3^2 K_4}{4K_2^4} \right]$$

$$+ 0(n^{-3}). \tag{3.23}$$

Similarly from (2.4), $K_1(y)$ reduces, in cases A and B, to

$$K_1(y) = 1 - \frac{a_1 h K_3}{6K_2}$$

$$- \frac{a_1^2 h}{8} \left[\frac{1}{3} K_3 a_1 (h-1)(5h-1) + \frac{K_4}{4K_2} (h-3) + \frac{K_3^2}{3K_2} (h+3) \right]$$

$$+ 0(n^{-3}). \tag{3.24}$$

In case C, $K_2(y)$ and $K_1(y)$ are simplified to

$$K_2(y) = a_1^2 h^2 K_2 [1 - \frac{K_2}{2} a_1^2 (h-1)(3h-1)] + 0(n^{-3})$$

$$K_1(y) = 1 - \frac{1}{2} a_1^2 h(1-h) K_2 [1 - \frac{a_1 K_2}{4} (h-2)(3h-1)] + 0(n^{-3}) \tag{3.25}$$

We finally compare the following measures of deviation from normality of the random variable y,

$$\gamma_i = \frac{K_i(y)}{K_2^{i/2}(y)} = K_i(y^*)$$

These will serve as a good indication of how good the approximations to normality will be.

Table 3.2 will display the orders of the γ_i (i = 1, 2, 3) of the transformed variable y in the cases A, B, C1, C2, and C3, and a random variable x whose cumulants $K_i(y) = 0(n)$, for the range of values i = 2, 3, 4, and 5 (as for example with the χ^2 distribution).

Orders of γ_i of y in the cases C1, C2, C3, A, and B

	γ_1	γ_2	γ_3
A	$n^{-3/2}$	n^{-2}	$n^{-5/2}$
B	$n^{-5/2}$	n^{-2}	$n^{-3/2}$
C1	$n^{-3/2}$	n^{-2}	$n^{-3/2}$
C2	$n^{-5/2}$	n^{-1}	$n^{-3/2}$
C3	$n^{-3/2}$	n^{-1}	$n^{-3/2}$
X	$n^{-1/2}$	n^{-1}	$n^{-3/2}$

Table 3.2.

CHAPTER IV

CUMULANTS OF DISTRIBUTIONS

This Chapter will be devoted exclusively to a brief exposition of a portion of the distributions to which transformation (1.1) may be applied, and to a derivation of their cumulants. As is readily seen by our normalizing principle, a knowledge of the cumulants of the distribution of x, at least numerically, is essential, in order to solve for the constants in our transformation.

Our transformation (1.1) may generally be applied to any distribution that approaches normality as $n \to \infty$, where n is some dominant parameter. As with Haldane, this will include an important class of statistics known as the *cumulative statistics*. These have the property that under convolution, they are reproductive, and more importantly, that the distribution or the random variable is additive. That is, for example, if x_{n_1}, and x_{n_2} are values of the statistic derived from n_1 members of a population (or n_1 experiments) and n_2 different members or experiments then $x_{n_1+n_2} = x_{n_1} + x_{n_2}$. Then if we define $K_r(x_n)$ as the rth cumulant of the distribution of x_n, it is readily seen that $K_r(x_n) = nK_r$ and therefore $K_r(x_n) = O(n)$. The best known examples are the binomial, the negative binomial, the Poisson, the gamma (which includes the χ^2 distribution as a special case), and the non-central χ^2. These distributions will all be considered separately in Section 1.

Another important class of distributions which are considered, as with Haldane, are the *derived cumulative statistics*. They have the property that when multiplied by some power of n, they become cumulative statistics. That is, x_n is a derived cumulative statistic if for some γ, $n^{\gamma}K_r = O(n)$. The mean of a sample clearly has this property.

The generality of our ordering condition (2.17), however, allows us to consider any distribution with the property

$$n^{\gamma}K_r = O(n^{\delta}) \qquad r = 2, 3, 4, 5.$$

$$\text{and} \quad n^{\gamma}K_r \leq O(n^{\delta}) \qquad r = 6, 7, \ldots, \tag{4.1}$$

where γ and $\delta>0$ are constants. Thus whenever the distribution of x has this property, the standardized transformed random variable y^* will in cases A and B, have the following asymptotic property,

$$y^* - z = O(n^{-3/2\delta}). \tag{4.2}$$

However, following the remarks made in the last section of Chapter II, transformation (1.1) may be applied and will generally normalize to a greater degree any distribution x which already tends to normality with some number n, and in practice, approximations of this kind often turn out to be remarkably good, as we shall see in Chapter V. Section 2 will thus be devoted to a study of some of the best known distributions of this kind, namely the hypergeometric, the beta, the F, the t, the

non-central F, the non-central t, and the distribution of the correlation coefficient in samples from the normal population. For the sake of simplicity we shall not distinguish between the random variable and the possible values it may assume.

4.1. Distributions of the Cumulative Statistics

We shall consider a number of well-known distributions of cumulative statistics. These are the binomial, the negative binomial, the Poisson, the gamma (which includes the χ^2 as a special case) and the non-central χ^2. Each distribution shall be examined separately in a different subsection.

4.1.1. The Binomial Distribution

The *binomial distribution with parameters n, p* is defined as the distribution of a random variable x for which $(0 < p < 1)$,

$$P_B(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad (x = 0, 1, 2, \ldots, n). \qquad (4.3)$$

The characteristic function of the distribution is given by

$$\phi_B(t) = (1 - p + p\, e^{it})^n.$$

Taking logarithms and expanding, we have for the cumulant generating function

$$\psi_B(t) = n \log \left(1 + p(it) + p\frac{(it)^2}{2!} + \ldots + p\frac{(it)^m}{m!} + \ldots\right)$$

Hence, the cumulants of x are n times the cumulants of a Bernoulli distribution B^* whose moments are all p. Using (1.4) we have

$$K_1 = np$$

$$K_{m+1} = n\left\{\mu'_{m+1}(B^*) - \sum_{j=0}^{m-1}\binom{m}{j}K_{j+1}(B^*)\,\mu'_{m-j}(B^*)\right\}$$

$$= np\left\{1 - \frac{1}{n}\sum_{j=0}^{m-1}\binom{m}{j}K_{j+1}\right\}.$$

Hence, the first six cumulants are ($q = 1-p$)

$$\left.\begin{aligned}
K &= np \\
K_2 &= npq \\
K_3 &= npq(q-p) \\
K_4 &= npq(1-6pq) \\
K_5 &= npq(q-p)(1-12pq) \\
K_6 &= npq(1-30pq+120p^2q^2),
\end{aligned}\right\} \qquad (4.4)$$

and $\gamma_1 = \dfrac{q-p}{(npq)^{1/2}}$, $\gamma_2 = \dfrac{1-6pq}{npq}$, $\gamma_3 = \dfrac{(q-p)(1-12pq)}{(npq)^{3/2}}$.

The cumulants (4.4) could also have been derived using the recurrence relation

$$K_{r+1} = pq \frac{dK_r}{dp} .$$

4.1.2. The Negative Binomial Distribution

The *negative binomial distribution with parameters n, p* is defined as the distribution of a random variable x for which $(0 < p < 1)$

$$P_N(x) = \binom{x-1}{n-1} p^n (1-p)^{x-n} \qquad (x = n, n+1, \ldots). \qquad (4.5)$$

The characteristic function may then be written as

$$\begin{aligned} \phi_N(t) &= (pe^{it})^n (1 - e^{it}(1-p))^{-n} \\ &= (1 - \frac{1}{p} + \frac{1}{p} e^{it})^{-n} \\ &= (1 - \frac{1}{p} it + \frac{1}{p}(it)^2 + \ldots + \frac{1}{p}(-1)^m \frac{(it)^m}{m!} + \ldots)^{-n} \end{aligned}$$

Hence, the cumulants of x are (-n) times the cumulants of a random variable N^* whose odd moments are $-\frac{1}{p}$ and even moments $\frac{1}{p}$. Again using (1.4) we have

$$K_1 = \frac{n}{p}$$

$$K_{m+1} = \frac{n}{p}\left\{(-1)^m + \frac{1}{n}\sum_{j=0}^{m-1}\binom{m}{j}(-1)^{m-j+1}K_{j+1}\right\},$$

from which we obtain ($q = 1 - p$)

$$\left.\begin{aligned}
K_1 &= \frac{n}{p} \\
K_2 &= \frac{n}{p^2}q \\
K_3 &= \frac{n}{p^3}q(1+q) \\
K_4 &= \frac{n}{p^4}q(1+4q+q^2) \\
K_5 &= \frac{n}{p^5}q(1+11q+11q^2+q^3) \\
K_6 &= \frac{n}{p^6}q(1+26q+66q^2+26q^3+q^4),
\end{aligned}\right\} \quad (4.6)$$

and

$$\gamma_1 = \frac{1+q}{(nq)^{1/2}} \qquad \gamma_2 = \frac{(1+4q+q^2)}{nq}$$

$$\gamma_3 = \frac{(1+11q+11q^2+q^3)}{(nq)^{3/2}}$$

The cumulants (4.6) could also have been obtained by using

$$K_{r+1} = -(1-p)\frac{dK_r}{dp}.$$

We note that unlike the binomial distribution, n need not be an integer, and when it is, (4.5) is sometimes called the *Pascal distribution* Also when n = 1, (4.5) becomes the *geometric distribution.*

4.1.3. The Poisson Distribution

A random variable is said to have a *Poisson distribution with parameter* λ if

$$P_p(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad (x = 0, 1, 2, \ldots; \lambda > 0). \qquad (4.7)$$

The characteristic function is

$$\phi_p(t) = \exp\left\{\lambda(e^{it} - 1)\right\},$$

and hence

$$\psi_p(t) = \lambda(e^{it} - 1)$$

$$= \lambda\sum_{j=1}^{\infty}\frac{it}{j!},$$

from which we have

$$K_r = \lambda \qquad \forall r \tag{4.8}$$

and

$$\gamma_1 = \lambda^{-\frac{1}{2}}, \quad \gamma_2 = \lambda^{-1}, \quad \gamma_3 = \lambda^{-3/2} .$$

4.1.4. The Gamma Distribution (Including χ^2)

A random variable has a *gamma distribution* if its probability density function is of the form

$$f_g(x) = \frac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1} \qquad (\alpha, p, x > 0) . \tag{4.9}$$

The main importance of the gamma distribution in statistical theory is the fact that if $z_1, z_2, \ldots, z_n$ are independant unit normal variables, the distribution of $\sum_{j=1}^{n} z_j^2$ is of the form (4.9) with $\alpha = \frac{1}{2}$ and $p = \frac{n}{2}$ and is called a *chi-square distribution with* n *degrees of freedom.* Also, if p is a positive interger, (4.9) is known as an *Erlang distribution,* and if p = 1 we have an *exponential distribution.*

The characteristic function may be written as

$$\phi_g(t) = \left(\frac{\alpha}{\alpha - it}\right)^p ,$$

from which we obtain the cumulant generating function

$$\psi_g(t) = p \log\left(\frac{\alpha}{\alpha - it}\right)$$
$$= -p \log\left[1 - \frac{it}{\alpha}\right]$$
$$= p\left[\frac{it}{\alpha} + \frac{(it)^2}{2\alpha^2} + \dots + \frac{(it)^m}{m\alpha^m} + \dots\right].$$

Then identifying coefficients we obtain the cumulants

$$K_r = p\,\alpha^{-r}\,(r-1)! \qquad \forall r, \tag{4.10}$$

and

$$\gamma_r = \frac{r!}{p^{r/2}} \qquad \forall r.$$

4.1.5. <u>The Non-Central χ^2 Distribution</u>

If $z_1, z_2, \dots z_\nu$ are independant unit normal variables and $\delta_1, \delta_2, \dots, \delta_\nu$ are constants, then the distribution of

$$\chi^2_\nu(\lambda) = \sum_{j=1}^{\nu} (z_j + \delta_j)^2 \tag{4.11}$$

is called the *non-central* χ^2 *distribution with* ν *degrees of freedom and non-centrality parameter* $\lambda = \sum_{j=1}^{\nu} \delta_j^2$.

The characteristic function is

$$\phi_x(t) = (1-2it)^{-\nu/2} \exp\left\{\frac{it\lambda}{1-2it}\right\},$$

and hence the cumulants are

$$K_r = 2^{r-1}(\nu + r\lambda)(r-1)! \qquad \forall r \tag{4.12}$$

and $$\gamma_r = \frac{2^{r/2-1}(\nu+r\lambda)\,(r-1)!}{(\nu+2\lambda)^{r/2}} \quad . \qquad \forall r$$

4.2. Other Distributions

In this section we will examine certain other distributions to which transformation (1.1) may be applied. We shall limit ourselves only to certain better known distributions, and therefore only the hypergeometric, the beta, the F, the t, the non-central t, the non-central F, and the distribution of the correlation coefficient, will be considered. A subsection shall be devoted to each distribution.

With the F, the non-central F, and the distribution of the correlation coefficient, we use the widely known technique of applying our transformation not directly to the random variable, but to a function of the random variable. For the F and non-central F we have $Z = \frac{1}{2}\log(x)$ and for the correlation coefficient we have $Z = \frac{x}{\sqrt{(1-x^2)}}$. The technique is employed since poor results are obtained when applying the transformation directly. As shall be seen in Chapter V, the approximation is excellent if applied on Z. This is due to the fact that our transformation merely makes a random variable which tends to normality, approach normality much faster. In this case, Z approaches normality faster than x, and hence the transformation applied to Z gives better results.

The problem is also one of convergence. Corrolary 2.3.1 gives sufficient conditions for the convergence of our moment, and hence cumulant expansions. However, in some of the distributions, the condition that all moments exist, may not hold. Most of all, condition (2.13) may not hold. It is this condition which dictates the convergence or non-convergence of our expansions.

We note here that the distributions considered below are far from exhaustive. Among the derived cumulative statistics, we have the mean $\bar{x}$ of a sample of n, since

$$K_i(\bar{x}) = n^{1-i} K_i ,$$

and hence in $nK_i(\bar{x}) = O(n)$. Similarly the cumulants of n times the second moment of a sample of n all tend to infinity with n. In fact, it may be shown, using Fisher's k-statistics, that for any statistic which is expressible as a symmetric function of the observations, if $K_{(\pi_1 \ldots \pi_s)}(a)$ contains r partitions, then every term in it is of order n^{1-r}. Among distributions which tend to normality as $n \to \infty$, we have the distributions of the regression, or the sample multiple correlation coefficients in normal samples.

4.2.1. The Hypergeometric Distribution

The *hypergeometric distribution with parameters* n, N_1, N is defined as the distribution of a random variable for which

$$P_H(x) = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} \qquad \max(0, n-N+N_1) \le x \le \min(n, N_1). \tag{4.13}$$

The characteristic function, in this case, does not provide us with an easy method of finding the cumulants. The simplest method is through the factorial moments which may be determined by noting that

$$\sum_x x^{(r)}\binom{N_1}{x}\binom{N-N_1}{n-x} = N_1^{(r)} \sum_x \binom{N_1-r}{x-r}\binom{N-N_1}{n-x} ,$$

and hence $$\mu'_{(r)} = \frac{N_1^{(r)}\, n^{(r)}}{N^{(r)}} .$$

We then have the moments about zero by

$$\mu'_r = \sum_{i=1}^{r} \mu'_{(r)}\, B^{(-i)}_{(r-i)}\ (0)$$

$$= \sum_{i=1}^{r} \frac{N_1^{(r)} n^{(r)}}{N^{(r)}}\, B^{(-i)}_{(r-i)}\ (0) ,$$

where $B_r^{(n)}(x)$ is the Bernoulli polynomial of order n and degree r in x. The cumulants are then produced using (1.4),

$$\left.\begin{aligned}
K_1 &= \frac{nN_1}{N} \\
K_2 &= \frac{nN_1(N-N_1)(N-n)}{N^2(N-1)} \\
K_3 &= \frac{nN_1(N-N_1)(N-2N_1)(N-n)(N-2n)}{N^3(N-1)(N-2)} \\
K_4 &= \frac{nN_1(N-N_1)(N-n)}{N^4(N-1)^2(N-2)(N-3)} \{ N^3(N+1)(N-1) - 6N^2 n(N-n)(N-1) \\
&\quad + 3N_1(N-N_1)[N^2(N-1)(n-2) - N(N-1)n^2 + 6n(N-1)(N-n) \\
&\quad - n(N-n)(N-2)(N-3)]\} .
\end{aligned}\right\} \quad (4.14)$$

The remaining cumulants and γ_i are calculated best numerically.

4.2.2. The Beta Distribution

A random variable is said to have a *beta distribution* if its probability density function is of the form

$$f_b(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \qquad 0 \le x \le 1 \quad . \qquad (4.15)$$

If b = 1, the distribution is sometimes called a *power function* distribution. The moments about zero, obtained by direct integration are,

$$\mu_r' = \frac{\Gamma(a+b)\ \Gamma(a+r)}{\Gamma(a)\ \Gamma(a+b+r)}$$

Using (1.4), we then derive the cumulants which are

$$\left.\begin{aligned}
K_1 &= \frac{a}{a+b} \\
K_2 &= \frac{ab}{(a+b)^2\ (a+b+1)} \\
K_3 &= \frac{2ab\ (b-a)}{(a+b)^3 (a+b+1)(a+b+2)} \\
K_4 &= \frac{6ab(b^3+a^3+b^2+a^2-2ab^2-2a^2b-4ab)}{(a+b)^4\ (a+b+1)(a+b+2)(a+b+3)} \\
K_5 &= \frac{24ab\ (b-a)(b^3-4ab^2+4a^2b+a^3+b^2-10ab+a^2)}{(a+b)^5\ (a+b+1)^2(a+b+2)(a+b+3)(a+b+4)}
\end{aligned}\right\} \qquad (4.16)$$

4.2.3. <u>"Student's" Distribution</u>

If z and χ_n^2 are independent standard normal and chi-square random variables with n degrees of freedom, then the distribution of

$$t_n = \frac{z}{(\chi^2/n)^{\frac{1}{2}}} \qquad (4.17)$$

is defined as the *t distribution with n degrees of freedom*. The moments

about zero may be found by

$$\mu_r' = E\left(\frac{n^{\frac{1}{2}} z}{\chi_n^2}\right)^r = n^{r/2}\, E(z)^r\, E(\chi_n^2)^{-r}$$

$$= \begin{cases} n^{r/2} \dfrac{\Gamma\left\{\frac{i+1}{2}\right\} \Gamma\left\{\frac{n-r}{2}\right\}}{\Gamma\left\{\frac{1}{2}\right\} \Gamma\left\{\frac{n}{2}\right\}} & r \text{ even} \\ 0 & r \text{ odd,} \end{cases}$$

for $r < n$. The cumulants are then found using (1.4),

$$\left.\begin{aligned} K_i &= 0 && \text{for } i \text{ odd} \\ K_2 &= \frac{n}{n-2} \\ K_4 &= \frac{6n^2}{(n-2)^2(n-4)} \\ K_6 &= \frac{240\, n^3}{(n-2)^3(n-4)(n-6)} \end{aligned}\right\} \quad (4.18)$$

and $\gamma_1 = \gamma_3 = 0 \qquad \gamma_2 = 6(n-4)^{-1}$.

4.2.4. <u>The F Distribution</u>

If $\chi^2_{\nu_1}$ and $\chi^2_{\nu_2}$ are independant chi-square random variables with ν_1 and ν_2 degrees of freedom respectively, then the distribution of

$$F_{\nu_1\nu_2} = \frac{\chi^2_{\nu_1}}{\chi^2_{\nu_2}} \frac{\nu_2}{\nu_1} \qquad (4.19)$$

is defined as the *F distribution with* ν_1, ν_2 *degrees of freedom.* The moments about zero may be obtained using the relation

$$\mu_1^r = E\left\{\frac{\chi^2_{\nu_1}}{\chi^2_{\nu_2}}\frac{\nu_2}{\nu_1}\right\}^r = \left[\frac{\nu_2}{\nu_1}\right]^r E\left[\chi^2_{\nu_1}\right]^r E\left[\chi^2_{\nu_2}\right]^{-r}$$

$$= \left[\frac{\nu_2}{\nu_1}\right]^r \frac{\Gamma\left\{\frac{1}{2}\nu_1 + r\right\}\Gamma\left\{\frac{1}{2}\nu_2 - r\right\}}{\Gamma\left\{\frac{1}{2}\nu_1\right\}\Gamma\left\{\frac{1}{2}\nu_2\right\}} \qquad \nu_2 > 2r.$$

However, since we are applying the transformation to $Z = \frac{1}{2}\log F_{\nu_1,\nu_2}$ we need to find the cumulants of Z. Noting as Wishart (1947) has done that the characteristic function of $\frac{1}{2}\log(\chi^2/n)$ is

$$\phi_Z(it) = \frac{\Gamma\{\frac{1}{2}(n+it)\}}{\Gamma(\frac{1}{2}n)} \exp\{-\tfrac{1}{2}it\log(\tfrac{1}{2}n)\},$$

we obtain the cumulant generating function of Z which is

$$\tfrac{1}{2}it\log\left[\frac{\nu_2}{\nu_1}\right] + \log\Gamma\{\tfrac{1}{2}(\nu_1+it)\} + \log\Gamma\{\tfrac{1}{2}(\nu_2-it)\} - \log\Gamma(\tfrac{1}{2}\nu_1) - \log(\tfrac{1}{2}\nu_2).$$

We then obtain the cumulants in terms of Psi and Zeta functions,

$$\left.\begin{aligned} K_1 &= \tfrac{1}{2}\left\{\log\left[\frac{\nu_2}{\nu_1}\right] + \psi(\tfrac{1}{2}\nu_1) - \psi(\tfrac{1}{2}\nu_2)\right\} \\ K_r &= 2^{-r}(r-1)!\left\{\zeta(r,\tfrac{1}{2}\nu_2) + (-1)^s\zeta(r,\tfrac{1}{2}\nu_1)\right\}, \end{aligned}\right\} \qquad (4.20)$$

where

$$\zeta(s,a) = \sum_{j=0}^{\infty}\frac{1}{(a+j)^s}, \quad \text{and } \psi(x) = \frac{d\{\log\Gamma(x)\}}{dx}$$

4.2.5. The Non-Central t Distribution

The distribution of the ratio

$$t_\nu(\lambda) = \frac{z+\lambda}{(\chi_\nu^2/\nu)^{\frac{1}{2}}} \tag{4.21}$$

where z and χ_ν^2 are independant random variables distributed as standard normal and chi-square with ν degrees of freedom respectively, and λ a constant, is called the *non-central t distribution with ν degrees of freedom and non-centrality parameter λ*. The moments about zero are found by

$$\mu_r' = \nu^{r/2}\, E(\chi_\nu^2)^{-r/2}\, E(z+\lambda)^r$$

$$= \left[\frac{\nu}{2}\right]^{r/2} \frac{\Gamma([\nu-r]/2)}{(\nu/2)} \sum_{0\le j\le r/2} \binom{r}{2j} \frac{(2j)!}{2^j\, j!}\, \lambda^{r-2j} ,$$

from which we derive the cumulants using (1.4) for $r < \nu$,

$$\left.\begin{aligned}
K_1 &= \left[\frac{\nu}{2}\right]^{\frac{1}{2}} \lambda \frac{\Gamma([\nu-1]/2)}{(\nu/2)} \\
K_2 &= \frac{\nu}{\nu-2}\,(1+\lambda^2) - K_1^2 \\
K_3 &= K_1 \left\{ \frac{\nu(2\nu-3+\lambda^2)}{(\nu-2)(\nu-3)} - 2K_2 \right\} \\
K_4 &= \frac{\nu^2(3+6\lambda^2+\lambda^4)}{(\nu-2)(\nu-4)} - K_1^2 \left\{ \frac{\nu[(\nu+1)\lambda^2+3(3\nu-5)]}{(\nu-2)(\nu-3)} - 3K_2 \right\} - 3K_2^2 .
\end{aligned}\right\} \tag{4.22}$$

The remaining cumulants and the γ_i are then calculated numerically.

4.2.6. The Non-Central F Distribution

If $\chi^2_{\nu_1}(\lambda)$ and $\chi^2_{\nu_2}$ are independant non-central and central chi-square random variables with ν_1 and ν_2 degrees of freedom respectively, then the distribution of

$$F_{\nu_1,\nu_2}(\lambda) = \frac{\chi^2_{\nu_1}(\lambda)}{\chi^2_{\nu_2}} \frac{\nu_2}{\nu_1} \tag{4.23}$$

is called the *F distribution with* ν_1, ν_2 *degrees of freedom and non-centrality parameter* λ. Again, the moments about zero are found using the relation

$$\mu'_r = \left[\frac{\nu_2}{\nu_2}\right]^r E(\chi^2_{\nu_1}(\lambda))^r E(\chi^2_{\nu_2})^{-r}$$

$$= \left[\frac{\nu_2}{2\nu_1}\right]^r \frac{\Gamma\left\{\frac{1}{2}\nu_2 - r\right\}}{\Gamma\left\{\frac{1}{2}\nu_2\right\}} \sum_{i=0}^{r} \binom{r}{i} \frac{2^{r-i}\Gamma\left\{\frac{1}{2}\nu_1 + r - i - \frac{1}{2}\right\}}{\Gamma((\nu_1-1)/2)} \sum_{j=0}^{i} \binom{2i}{2j} \frac{(2j)!}{2^j\, J!} \lambda^{i-j}$$

Again, since we are applying our transformation (1.1) to $Z^* = \frac{1}{2}\log F_{\nu_1,\nu_2}(\lambda)$, we must find the cumulants of our new variable Z^*. However Barton, David, and O'Niell (1960) found that the random variable $\log\{\frac{1}{2}\chi^2_{\nu_1}(\lambda)\}$ is the sum of two independant random variables, $\log(\frac{1}{2}\chi^2_{\nu_1})$ and another random variable with crude moments

$$M_r = r! \sum_{i=r}^{\infty} \frac{(\frac{1}{2}\lambda)^i (-1)^{i-r}}{i! \left\{\frac{\nu_1}{2}\right\}^{[i]}} \left| S_i^r \right|$$

where $a^{[r]} = a(a+1)\ldots(a+r-1)$

and $\left| S_i^r \right|$ is Stirling's number of the first kind in modulus.

We then have

$$K_r(Z^*) = K_r(Z) + 2^{-r} R_r \tag{4.24}$$

where R_r is the rth cumulant of the random variable whose moments are M_r (using (1.4)), and $K_r(Z)$ are given by (4.20).

4.2.7. The Distribution of the Sample Correlation Coefficient in Normal Samples

The statistic known as the *sample correlation coefficient* based on n pairs of observed values (X_i, Y_i) $(i = 1,\ldots,n)$ is

$$r = \frac{\sum_{i=1}^{n} (X_i-\overline{X})(Y_i-\overline{Y})}{\left[\sum_{i=1}^{n} (X_i-\overline{X})^2(Y_i-\overline{Y})^2\right]^{\frac{1}{2}}} \tag{4.25}$$

When (X_i, Y_i) and (X_j, Y_j) are mutually independant for $i \neq j$ and the joint distribution of X_i and Y_i is the bivariate normal distribution of all i, the distribution of the sample correlation coefficient has been explicitly derived. The moments of r may be expressed in terms of hypergeometric functions which may then be expanded as power series in $\frac{1}{n}$ (Hotelling (1953)) or in $\frac{1}{n+6}$ (Ghosh (1966)).

Here again we apply our transformation not to r but to

$$\tilde{r} = \frac{r}{\sqrt{1-r^2}}$$

The cumulants are then obtained from the moments of $\tilde{r}$ which have been

derived by Ruben (1966), that is

$$\left.\begin{aligned} E(\tilde{r}^m) &= \frac{1}{\langle n-m-2\rangle_{\frac{1}{2}(m+1)}} \sum_{k=0}^{\frac{1}{2}(m-1)} \langle 1\rangle_k \langle n-2\rangle_{\frac{1}{2}(m+1)-k} \binom{m}{2k} \tilde{\rho}^{m-2k} \\ &\qquad\qquad \text{for } m \text{ odd} \\ &= \frac{1}{\langle n-m-2\rangle_{\frac{1}{2}m}} \sum_{k=0}^{\frac{1}{2}m} \langle 1\rangle_k \langle n-1\rangle_{\frac{1}{2}m-k} \binom{m}{2k} \tilde{\rho}^{m-2k} \\ &\qquad\qquad \text{for } m \text{ even} \end{aligned}\right\}$$

where $\langle a\rangle_0 = 1$, $\langle a\rangle_j = a(a+2)\ldots(a+2j-2)$ for $j = 1,2,\ldots$ and $\tilde{\rho} = \dfrac{\rho}{\sqrt{1-\rho^2}}$.

CHAPTER V

RESULTS AND CONCLUSIONS

In this Chapter, we shall apply the analytic solutions, obtained in Chapter III, to each of the distributions examined in Chapter IV. Numerical solutions of the constants and certain points on the distribution function, are obtained for a number of variations in the parameters for each distribution. These are summarized into tables and given after each distribution is considered. Analytic solutions and results shall also be presented in certain special cases.

Numerical comparisons shall also be made with certain leading approximations. In all cases, the Gram-Charlier, the Cornish-Fisher, and the Haldane A and B transformations shall be calculated and comparisons made with our values. In addition, certain of the more successful approximations made to each particular distribution shall also be considered. However, the scope of our research does not permit an exhaustive study of all the competitive approximations, and hence, our choice of the leading approximations was arbitrary. The reader is referred to the three volumes of Johnson and Kotz (1969 and 1970) for an excellent exposition of the available approximations.

When examining the tables in this chapter, the following must be kept in mind. Firstly, Cases A, B, and C2 all may have more than one solution. The choice of the solution to be inserted in the numerical comparisons was made on the basis of which solution caused the higher order terms in the cumulants of y to approach zero most rapidly, or equivalently, the solution with the smallest maximum absolute deviation.

In certain instances, some of our cases may not have any real solutions, and are hence omitted. All other approximations are those considered here in each subsection. All the deviations are the result of the subtraction of the approximation from the true value. Finally, it should be noted that the maximum and average absolute deviations are calculated on the basis of a maximum of fifty values of x, and in most cases, only a small subset is displayed. GC and CF shall denote the Gram-Charlier and Cornish-Fisher approximations.

It should be remembered when examining the results that our approximation, like the ones of Gram-Charlier, Cornish-Fisher, and Haldane, is distribution-free, and does not take into account the particular form of the distribution function, but only its first few cumulants. However, when the distribution function is unknown, and the first few cumulants may be obtained, even numerically, our transformation provides excellent approximations.

As in Chapter IV, the distributions are divided into two categories, the cumulative statistics and the others, with each distribution considered separately in a subsection. Identical characters and notation shall be used as in the equivalent sections in Chapter IV. Section 1 shall deal with symmetric distributions, while in Section 4, we will make a final analysis, and draw final conclusions. All calculations were performed on the IBM 360/75 computer at McGill University in double precision arithmetic using the Fortran IV WATFIV compiler.

5.1. Symmetric Distributions

When x is symmetric, our solutions for the constants a_1, a_2,

and h simplify immensely. Symmetry implies that

$$K_i = 0 \qquad \text{for } i = 3, 5, 7, \ldots .$$

Hence, as observed in Chapter 3, whenever $K_5 = 0$, we have Case A reducing to Case B. Also we have $K_3 = 0$, and therefore equations (3.8), (3.9), and (3.16) reduce to

$$a_1^2 (h-1) + 2a_2 = 0$$

$$K_4 - 4K_2^3 \{a_1^2 (h-1)(2h-1)\} = 0$$

$$2h^2 - 7h + 3 = 0$$

from which we obtain, as the only possible solutions,

$$a_2 = \frac{K_4}{40K_2^3} , \quad a_1 = \pm\sqrt{-a_2} , \quad h = 3 . \tag{5.1}$$

5.2. Results for the Cumulative Statistics

A particularly interesting result follows when there is some factor, say n, common to the first few cumulants of x (the first five cumulants for Case A, the first four for Case B, and the first three for Case C). Then we may write

$$K_i = nt_i \qquad i = 2, 3, 4, 5 .$$

By substitution into the derived solutions of a_1, a_2, and h in Cases

A, B, and C, that is into equations (3.11), (3.13), (3.15), (3.18), (3.20), and (3.22), we may then write the solutions of a_1, a_2 and h as

$$\left.\begin{aligned} a_1 &= \frac{a_1^*}{n} \\ a_2 &= \frac{a_2^*}{n^2} \\ h &= h^* \end{aligned}\right\} \qquad (5.2)$$

where a_1^*, a_2^*, and h^* are constants dependant solely on t_i ($a_2 = 0$ in Case C). This fact is especially important, since with most cumulative statistics, n is independant of t_i, and therefore a_1^*, a_2^*, and h are independant of n. Thus solutions for different n may be easily obtained from one, and thus one dimension is reduced from any tabulation.

5.2.1. The Binomial Distribution

As a direct result of (5.2), since the parameter n is seperable, and is a common factor to all the cumulants, we are then able to construct Table 5.1 giving the constants a_1^*, a_2^*, and h for different values of p (from .05 to .95 in steps of .05) for each of the Cases A, B, C1, C2, and C3.

There are many approximations to binomial distribution. First there is the classical normalizing transformation (with continuity correction)

$$\frac{x - np + \frac{1}{2}}{npq} \qquad (5.3)$$

TABLE 5.1

SOLUTION OF CONSTANTS FOR THE BINOMIAL DISTRIBUTION (n = 1)

	Case A			Case B			Case C1		Case C2		Case C3	
P	a_1^*	a_2^*	h	a_1^*	a_2^*	h	a_1^*	h	a_1^*	h	a_1^*	h
.05	24.196757	126.957784	.30529639	48.188255	594.360204	.35702071	19.035088	.66820276	15.701448	.59775751	20.000000	.68421053
.10	11.266773	25.667246	.33261855	26.174025	175.263716	.37513816	10.023148	.70438799	7.892920	.62460495	10.000000	.70370370
.15	6.957326	8.638297	.38003650	19.401732	96.176787	.39467714	7.175537	.74495771	5.306619	.65513533	6.666667	.72549020
.20	4.756025	2.495141	.51655962	16.675050	70.884608	.41518158	5.937500	.78947368	4.029195	.68976431	5.000000	.75000000
.25	14.654400	54.466146	.43209492	15.910479	64.326496	.43590998	5.444444	.83673469	3.278685	.72888963	4.000000	.77777778
.30	15.823074	63.321503	.45404972	16.707713	70.650656	.45580999	5.496031	.88447653	2.795210	.77285408	3.333333	.80952380
.35	18.965331	90.618385	.47294209	19.571270	96.521564	.47355703	6.208791	.92920353	2.468279	.82191627	2.857142	.84615384
.40	26.451329	175.597574	.48755658	26.831378	180.636182	.48768746	8.263888	.96638655	2.243716	.87619745	2.500000	.88888889
.45	50.703204	643.373386	.49682278	50.886643	648.038444	.49683135	15.370370	.99123767	2.092773	.93564513	2.222222	.93939394
.50	†	†	†	†	†	†	†	†	†	†	†	†
.55	-50.703204	643.373386	.49682278	-50.886643	648.038444	.49683135	-15.370370	.99123767	-2.092773	.93564513	1.818182	1.07407407
.60	-26.451329	175.597574	.48755658	-26.831378	180.636182	.48768746	-8.263888	.96638655	-2.243716	.87619745	1.666666	1.16666666
.65	-18.965331	90.618385	.47294209	-19.571270	96.521564	.47355703	-6.208791	.92920353	-2.468279	.82191627	1.538461	1.23571428
.70	-15.823074	63.321503	.45404972	-16.707713	70.650656	.45580999	-5.496031	.88447653	-2.795210	.77285408	1.428571	1.44444444
.75	-14.654400	54.466146	.43209492	-15.910479	64.326496	.43590998	-5.444444	.83673469	-3.278685	.72888963	1.333333	1.66666667
.80	-4.756025	2.495141	.51655962	-16.675050	70.884608	.41518158	-5.937500	.78947368	-4.029195	.68976431	1.250000	2.00000000
.85	-6.957326	8.638297	.38003650	-19.401732	96.176787	.39467714	-7.175537	.74495771	-5.306619	.65513533	1.176471	2.55555556
.90	-11.266773	25.667246	.33261855	-26.174025	175.263716	.37513816	-10.023148	.70438799	-7.892920	.62460495	1.111111	3.66666667
.95	-24.196757	126.957784	.30529639	-48.188255	594.360204	.35702071	-19.035088	.66820276	-15.701448	.59775751	1.052632	7.00000000

$$P_B(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots n$$

Then $a_1 = \frac{a_1^*}{n}$ $\quad a_2 = \frac{a_2^*}{n^2}$

Note: Cases A, B, and C may have more than one solution. In such cases, the best was chosen, where best is defined as the one in which the higher orders of terms in the cumulants of y approach zero most rapidly.

† No real solutions

A number of variance stabalizing transformations have also been proposed. Besides the classical inverse sin, Anscombe (1948) and Freeman and Tukey (1950) have proposed various modifications to the inverse sin. Transformations of this nature, although they are generally better than the classical normal, must be used with caution, when employing them as normalizing transformations. Camp (1951), however, produced a very satisfactory transformation developed as a corollary to Paulson's (1942) approximation to the F. The variate

$$\frac{\frac{1}{3}\left\{\left(\frac{9s-1}{s}\right)\left(\frac{s}{x+1}\,\frac{q}{p}\right)^{1/3} - \frac{9x+8}{x+1}\right\}}{\left\{\frac{1}{s}\left(\frac{s}{x+1}\,\frac{q}{p}\right)^{2/3} + \frac{1}{x+1}\right\}^{1/2}} \tag{5.4}$$

where $s = n-x$, is approximately standard normal. Also, as seen in Chapter 1, Peizer and Pratt (1968) proposed the following excellent approximation,

$$d_2\left\{\frac{1 + qg\left(\frac{x+\frac{1}{2}}{np}\right) + pg\left(\frac{n-x-\frac{1}{2}}{nq}\right)}{\left(n+\frac{1}{6}\right)pq}\right\}^{1/2} \tag{5.5}$$

where $d_1 = x + \frac{2}{3} - (n+\frac{1}{3})p$

$$d_2 = d_1 + .02\left[\frac{q}{x+1} - \frac{p}{n-x} + \frac{q-\frac{1}{2}}{n+1}\right]$$

and $g(x) = (1-x)^{-2}(1-x^2+2n\log x)$.

The approximations above have been compared for several values of the parameters n and p. A portion has been condensed into Table 5.2. It may generally be concluded that Piezer and Pratt's approximation is

TABLE 5.2

BINOMIAL DISTRIBUTION: ACCURACY OF APPROXIMATIONS (True value-approximation)

(1) n = 4, p = .2

X	Pr	N	C	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
0	.4096	.0558	-.0019	.0037	-.0006	.0134	-.0111	.0042	-.0185	.0005	-.0085
1	.8192	.0100	-.0005	.0001	.0132	.0134	.0040	.0091	-.0006	.0022	.0010
2	.9728	-.0104	.0009	-.0001	.0050	-.0107	.0060	.0080	.0070	.0002	.0060
3	.9984	-.0012	.0002	-.0001	.0006	-.0016	.0017	.0022	.0021	.0022	.0022
4	1.0000	.0004	.0003	.0000	.0001	.0000	.0002	.0003	.0003	.0003	.0003
Maximum Absolute Deviation		.0558	.0019	.0037	.0132	.0134	.0111	.0091	.0185	.0022	.0085
Average Absolute Deviation		.0156	.0007	.0010	.0039	.0052	.0046	.0047	.0055	.0011	.0036

(2) n = 10, p = .1

X	Pr	N	C	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
0	.3478	.0496	-.0011	.0038	.0052	-.0019	-.0075	-.0028	-.0086	.0017	-.0085
1	.7361	.0352	-.0022	.0002	.0084	.0091	-.0002	.0048	-.0013	.0006	-.0013
2	.9298	-.0133	.0006	.0001	.0035	.0020	.0050	.0069	.0048	.0041	.0048
3	.9672	-.0086	.0008	.0001	.0027	-.0122	.0028	.0031	.0029	.0029	.0029
4	.9984	-.0015	.0003	.0000	.0005	-.0016	.0008	.0008	.0009	.0010	.0009
Maximum Absolute Deviation		.0496	.0021	.0038	.0084	.0122	.0075	.0069	.0086	.0041	.0085
Average Absolute Deviation		.0216	.0010	.0011	.0041	.0054	.0033	.0037	.0037	.0020	.0037

(3) n = 10, p = .9

X	Pr	N	C	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
6	.0128	.0086	-.0008	-.0001	-.0027	.0122	-.0030	-.0031	-.0029	-.0029	-.0051
7	.0702	.0133	-.0006	-.0001	.0035	-.0020	-.0051	-.0069	-.0048	-.0042	.0049
8	.2639	-.0352	.0022	-.0002	.0084	-.0091	.0002	-.0048	.0013	-.0006	.0141
9	.6513	.0496	.0011	-.0038	-.0052	.0019	.0075	.0028	.0087	.0017	-.0293
10	1.0000	.2991	.0000	.0000	.0246	.0204	.0001	.0854	.0189	.0047	.0236
Maximum Absolute Deviation		.2991	.0022	.0038	.0246	.0204	.0075	.0854	.0189	.0047	.0293
Average Absolute Deviation		.0811	.0009	.0011	.0089	.0091	.0032	.0206	.0073	.0028	.0154

Note: N is the classical normal approximation (5.3)
C is Camp's approximation (5.4)
PP is the Piezer and Pratt approximation (5.5)
GC is the Gram Charlier approximation using six terms
CF is the Cornish Fisher approximation using six adjustments

best, followed closely by Camp's approximation, with the proposed approximation following third. The variance stabalizing transformations yield poor results and are not included in the tabulation. With regards to the proposed transformation, it is noted that Cases A and B do not perform as well as Case C2, which seems to give the better results. This is especially true as the distribution becomes less skew.

5.2.2. The Negative Binomial Distribution

Again, since n is a separable parameter in all the cumulants of the Negative Binomial, we may then tabulate (Table 5.3) the required constants a_1^*, a_2^*, and h for the range of values of p from .05 to .95 in steps of .05.

The sum of negative binomial terms can be expressed as a sum of binomial terms, and thus computation of probabilities can thus be reduced to calculation of the corresponding binomial probabilities. Separate approximations, however, have been developed. Beall (1942) proposed a variance stabalizing transformation which was later modified by Anscombe (1948). Paulson's or Camp's approximation may also be applied here. The variate

$$\frac{\frac{1}{3}\left\{\frac{9n+8}{n+1} - (9x-1)\left[\frac{x}{p(n+1)}\right]^{1/3}\right\}}{\left\{\frac{1}{x}\left[\frac{x}{p(x+1)}\right]^{2/3} + \frac{1}{n+1}\right\}^{-1/2}} \tag{5.6}$$

is approximately standard normal. Piezer and Pratt also extend their set of approximations to the negative binomial, yielding,

TABLE 5.3

SOLUTION OF CONSTANTS FOR THE NEGATIVE BINOMIAL DISTRIBUTION (n=1)

	CASE A			CASE B			CASE C1		CASE C2		CASE C3	
P	a_1^*	a_2^*	h	a_1^*	a_2^*	h	a_1^*	h	a_1^*	h	a_1^*	h
.05	.05454616	.00006385	.32989251	.05457169	.00006160	.33173767	.05561740	.38489536	.05408027	.36741211	.05000000	.31578947
.10	.11242639	.00025808	.33323950	.11240571	.00026173	.33253078	.11457115	.38579328	.11139663	.36928997	.10000000	.29629630
.15	.17449629	.00056727	.33909567	.17425342	.00063079	.33393859	.17764308	.38740351	.17269741	.36986009	.15000000	.27450980
.20	.24186998	.00094279	.34759740	.24107271	.00121333	.33602564	.24583333	.38983051	.23893596	.37221673	.20000000	.25000000
.25	.31600889	.00129310	.35878889	.31409800	.00207667	.33884490	.32043651	.39310085	.31130946	.37545840	.25000000	.22222222
.30	.39885993	.00147149	.37262283	.39493938	.00332607	.34243010	.40315126	.39760292	.39150666	.37968580	.30000000	.19047619
.35	.49302385	.00125753	.38896635	.48572089	.00513218	.34678576	.49624126	.40320592	.48155014	.38499894	.35000000	.15384615
.40	.59203402	.00002582	.40761163	.58933681	.00778178	.35187459	.60277778	.41013825	.58430882	.39149377	.40000000	.11111111
.45	.73073742	.00180735	.42827756	.70969927	.00177484	.35760139	.72701612	.41854485	.70367539	.39925612	.45000000	.06060606
.50	.88595246	-.00588077	.45062018	.85235633	.01801704	.36379100	.08750000	.42857143	.84511903	.40836737	†	†
.55	1.07751640	-.01204947	.47423598	1.02533096	.02827497	.37015841	1.05557152	.44035934	1.01655532	.41887989	.55000000	.07407407
.60	1.32013018	-.02520054	.49866981	1.24067897	.04582022	.27625836	1.28214285	.45400390	1.22986741	.43093296	.60000000	.16666667
.65	1.63677031	-.04555963	.52342481	1.51728354	.07810602	.38139399	1.57509206	.46972177	1.50373001	.44423914	.65000000	-.28571429
.70	2.06549784	-.07996389	.54797448	1.83722146	.14214866	.38441012	1.97286324	.48749052	1.86925679	.45908389	.70000000	.44444444
.75	2.67441835	-.14006902	.57177503	2.40804894	.28362470	.38308401	2.53750000	.50738916	2.38242472	.47532445	.75000000	.66666667
.80	3.59905654	-.25152783	.59427542	3.20015836	.66023062	.37108621	3.40000000	.52941176	3.15513623	.49289036	.80000000	-1.00000000
.85	5.15444900	-.48285058	.61[illegible]2115	11.20420735	32.04261933	.29562444	4.86491546	.55349230	4.44840850	.51168553	.85000000	-1.55555556
.90	8.28348287	-1.08172452	.63314661	1.84474314	86.96717578	.31000428	7.84772727	.57949609	7.04513670	.53159177	.90000000	-2.66666667
.95	17.69525213	-3.78180097	.64834843	40.70854424	423.69130695	.32506386	16.93035714	.60721443	14.85945021	.55244735	.95000000	-6.00000000

$$P_N(x) = \binom{x-1}{n-1} p^n (1-p)^{x-n} \qquad x = n, n+1, \ldots$$

Then $a_1 = \frac{a_1^*}{n}$ $\qquad a_2 = \frac{a_2^*}{n}$

Note: Cases A, B, and C2 may have more than one solution. In such cases, the best one was chosen, where best is defined as the one in which the higher order of terms in the cumulants of y approach zero most rapidly.

† No solution

$$d_f \left[\frac{1 + pg\left(\frac{x-n+\frac{1}{2}}{x(1-p)}\right) + qg\left(\frac{x-\frac{1}{2}}{xp}\right)}{\left[x+\frac{1}{6}\right] p(1-p)} \right]^{1/2} \tag{5.7}$$

where $d_1 = x - n + \frac{2}{3} - \left[x+\frac{1}{3}\right](1-p)$

$$d_2 = d_1 + .02\left[\frac{p}{x-n+\frac{1}{2}} - \frac{1-p}{n} + \frac{p-\frac{1}{2}}{x+1}\right].$$

and $g(x)$ is as given earlier.

Comparisons have been made of the approximations for a range of values of the parameters n and p. The numerical results suggest that the proposed approximation is of the same order of accuracy as that of Peizer and Pratt and Paulson-Camp. The Paulson-Camp approximation yields inferior results for smaller values of p, while Peizer and Pratt's approximation is not as good for high values of p. A small portion of the comparisons is presented in Tables 5.4 and 5.5.

5.2.3. The Poisson Distribution

With the Poisson distribution, since all the cumulants are equal to λ, a_1^*, a_2^*, and h are constants. Hence a closed solution, given below in Table 5.6, is possible.

There is a direct relationship between a Poisson and a χ^2 distribution. This implies that approximations of one are also approximations of the other. This fact can be used to extend the Wilson and Hilferty approximation to the Poisson, giving us,

$$3\left[\left\{\frac{\lambda}{x+1}\right\}^{1/3} - 1 + \frac{1}{9(x+1)}\right](x+1)^{1/2}. \tag{5.8}$$

TABLE 5.4

NEGATIVE BINOMIAL: ACCURACY OF APPROXIMATIONS

(1) n = 4, p = .1

X	Pr	N	P	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
4	.00010	.03057	-.00005	.00000	.00880	-.00151	.00004	.00004	.00010	.00007	-.00049	.00004
5	.00046	.03405	-.00016	.00001	.00552	-.00215	.00009	.00008	.00036	.00012	-.00084	.00007
9	.00833	.04564	-.00069	.00010	-.01020	-.00417	.00013	.00012	.00088	.00013	-.00203	.00005
11	.01853	.04800	-.00079	-.00009	-.01784	-.00431	.00009	.00008	.00084	.00008	-.00208	-.00001
15	.05555	.04275	-.00047	-.00010	-.02915	-.00325	-.00002	-.00002	.00047	.00007	-.00100	-.00009
17	.08264	.03520	-.00012	-.00008	-.03219	-.00231	-.00005	-.00006	.00026	.00012	-.00011	-.00007
20	.13295	.01908	-.00046	-.00006	-.03327	-.00124	-.00010	-.00010	-.00003	.00022	.00123	-.00002
22	.17192	.00526	-.00079	-.00005	-.03170	-.00074	-.00011	-.00011	-.00017	.00028	.00196	.00003
26	.25905	.02068	-.00121	-.00004	-.02350	-.00023	-.00011	-.00011	-.00035	.00035	.00280	.00012
28	.30543	.03821	.00127	-.00005	-.01714	-.00012	-.00009	-.00009	-.00039	.00034	.00238	.00016
31	.37617	.04909	.00118	-.00005	-.00547	-.00007	-.00008	-.00008	-.00042	.00026	.00263	.00018
33	.42305	.05710	.00104	-.00006	.00324	-.00005	-.00007	-.00006	-.00042	.00019	.00230	.00018
37	.51355	.06597	.00058	-.00008	.02071	-.00003	-.00004	-.00004	-.00041	.00001	.00139	.00015
39	.55623	.06679	.00039	-.00009	.02835	-.00002	-.00003	-.00003	-.00039	-.00008	.00091	.00013
43	.63520	.06203	-.00004	-.00011	.03870	-.00002	-.00002	-.00002	-.00036	-.00026	-.00001	.00006
44	.65345	.05971	-.00014	-.00011	.03994	-.00002	-.00002	-.00001	-.00035	-.00030	-.00021	.00004
48	.72014	.04721	-.00045	-.00012	.03870	-.00002	-.00001	.00000	-.00031	-.00040	-.00084	-.00002
50	.74971	.03770	-.00056	-.00012	.03444	-.00003	.00000	.00000	-.00028	-.00043	-.00105	-.00004
54	.80154	.02391	-.00069	-.00013	.02029	-.00012	.00001	.00002	-.00023	-.00044	-.00129	-.00008
56	.82399	.01625	-.00071	-.00013	.01152	-.00022	.00001	.00001	-.00020	-.00043	-.00134	-.00010
59	.85369	.00573	-.00070	-.00012	-.00186	-.00050	.00001	.00001	-.00016	-.00040	-.00132	-.00011
61	.87103	-.00039	-.00067	-.00011	-.00999	-.00078	.00001	.00001	-.00013	-.00036	-.00126	-.00012
65	.90044	-.01007	-.00058	-.00010	-.02204	-.00170	.00001	.00001	-.00009	-.00029	-.00108	-.00012
67	.91279	-.01358	-.00052	-.00010	-.02535	-.00236	.00001	.00001	-.00006	-.00025	-.00098	-.00012
70	.92877	-.01726	-.00042	-.00009	-.02675	-.00358	.00001	.00001	-.00003	-.00019	-.00080	-.00011
72	.93789	-.01874	-.00036	-.00008	-.02553	-.00455	.00001	.00001	.00000	-.00014	-.00068	-.00010
76	.95003	-.01978	-.00024	-.00007	-.01932	-.00676	.00001	.00000	.00003	-.00007	-.00046	-.00009
78	.95924	-.01953	-.00019	-.00006	-.01510	-.00792	.00000	.00000	.00004	-.00004	-.00036	-.00008
82	.96946	-.01800	-.00009	-.00005	-.00634	-.01015	.00000	.00000	.00006	.00001	-.00019	-.00007
83	.97260	-.01746	-.00007	-.00005	-.00431	-.01049	.00000	.00000	.00007	.00002	-.00015	-.00007
87	.97887	-.01499	-.00002	-.00004	.00233	-.01166	.00000	.00000	.00008	.00006	-.00002	-.00005
89	.98180	-.01366	.00003	-.00004	.00460	-.01178	.00000	.00000	.00009	.00007	.00003	-.00004
93	.98655	-.01104	.00007	-.00003	.00700	-.01104	.00000	.00000	.00009	.00009	.00010	-.00003
95	.98846	-.00982	.00008	-.00003	.00728	-.01035	-.00001	-.00001	.00008	.00009	.00012	-.00004
ximum Absolute Deviation		.06679	.00127	.00013	.04058	.01178	.00013	.00012	.00088	.00044	.00288	.00018
erage Absolute Deviation		.02917	.00047	.00008	.01842	.00344	.00003	.00003	.00024	.00020	.00103	.00008

te: N is the classical normal approximation
P is the extension of the Paulson-Camp approximation (5.6)
PP is the Peizer-Pratt approximation (5.7)

TABLE 5.5

NEGATIVE BINOMIAL: ACCURACY OF APPROXIMATION

n = 4 p = .8

	Pr	N	P	PP	CC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
	.4096	.0822	-.0019	-.0043	-.0047	.0042	-.0047	.0020	-.0018	.0036	.0438	†
	.7372	.0646	-.0014	-.0008	.0789	.0072	-.0037	.0021	-.0012	-.0001	-.0105	†
	.9011	-.0090	.0002	-.0001	-.0341	-.0026	.0015	.0039	.0025	.0021	-.0130	†
	.9666	-.0206	.0005	.0000	-.0128	-.0321	.0024	.0022	.0023	.0021	-.0033	†
	.9895	-.0095	.0003	.0000	-.0128	-.0105	.0016	.0007	.0012	.0012	.0010	†
	.9969	-.0030	.0001	.0000	.0014	-.0031	.0009	.0002	.0005	.0006	.0018	†
m Absolute Deviation		.0822	.0019	.0043	.0789	.0321	.0047	.0038	.0025	.0036	.0438	†
e Absolute Deviation		.0315	.0007	.0009	.0241	.0099	.0024	.0018	.0015	.0016	.0122	†

n = 30 p = .3

	Pr	N	P	PP	CC	CF	Case A	Case B	Case C1	Case C3	Case C3	Case D
	.001284	-.004576	-.000016	-.000001	.000567	-.000043	-.000003	-.000007	.000002	-.000016	-.000138	-.000005
	.004362	-.007593	-.000028	-.000001	.000459	-.000066	-.000006	-.000014	.000001	-.000028	-.000231	-.000011
	.009551	-.010045	-.000033	-.000001	.000163	-.000077	-.000010	-.000019	-.000001	-.000036	-.000277	-.000015
	.023349	-.012557	-.000028	-.000003	-.000483	-.000082	-.000018	-.000028	-.000009	-.000040	-.000248	-.000028
	.049243	-.012728	-.000006	-.000004	-.001033	-.000077	-.000030	-.000036	-.000023	-.000032	-.000076	-.000039
	.079279	-.010512	.000023	-.000003	-.001294	-.000069	-.000037	-.000040	-.000033	-.000016	.000132	-.000044
	.135846	-.004185	.000065	-.000003	-.001244	-.000058	-.000045	-.000043	-.000046	.000007	.000427	-.000045
	.211546	.004958	.000093	-.000001	-.000740	-.000045	-.000046	-.000040	-.000052	.000026	.000624	-.000037
	.279267	.012271	.000097	.000000	-.000157	-.000032	-.000042	-.000034	-.000050	.000031	.000647	-.000027
	.379760	.020319	.000075	.000000	.000690	-.000014	-.000031	-.000021	-.000040	.000024	.000501	-.000010
	.485326	.024439	.000031	.000000	.001359	.000006	-.000015	-.000005	-.000025	.000005	.000214	.000007
	.563328	.024216	-.000006	-.000001	.001595	.000019	-.000002	.000009	-.000012	-.000010	-.000026	.000019
	.660520	.019932	-.000049	-.000002	.001442	.000031	.000015	.000024	.000005	-.000025	-.000295	.000029
	.745628	.012625	-.000074	-.000003	.000802	.000038	.000027	.000035	.000020	-.000030	-.000444	.000035
	.799927	.006516	-.000079	-.000005	.000147	.000038	.000033	.000039	.000027	-.000027	-.000470	.000035
	.859120	-.000848	-.000070	-.000005	-.000666	.000032	.000036	.000040	.000032	-.000016	-.000410	.000033
	.904095	-.006113	-.000051	-.000005	-.001122	.000017	.000035	.000036	.000034	-.000002	-.000287	.000027
	.929640	-.008388	-.000033	-.000005	-.001144	.000001	.000033	.000032	.000033	.000009	-.000182	.000023
	.954696	-.009397	-.000012	-.000004	-.000820	-.000027	.000027	.000025	.000030	.000020	-.000057	.000017
	.971686	-.008717	.000003	-.000003	-.000313	-.000058	.000022	.000018	.000025	.000025	.000033	.000012
	.980466	-.007578	.000010	-.000003	.000035	-.000080	.000017	.000013	.000021	.000026	.000075	.000009
	.988373	-.005766	.000020	-.000002	.000331	-.000101	.000013	.000008	.000017	.000026	.000103	.000006
	.993256	-.004045	-.000017	-.000002	.000406	-.000115	.000008	.000004	.000012	.000021	.000105	.000003
	.995591	-.002961	.000016	-.000001	.000357	-.000117	.000007	.000003	.000010	.000019	.000097	.000003
a Absolute Deviation		.024681	.000097	.000005	.001594	.000117	.000046	.000042	.000051	.000040	.000652	.000045
e Absolute Deviation		.009589	.000037	.000002	.000703	.000048	.000021	.000022	.000020	.000020	.000243	.000020

† no solution found

TABLE 5.6

	a_1^*	a_2^*	h
Case A	1.30413937	.38739637	.28885204
Case B	2.25856068	1.30552708	.34055201
Case C1	.91666667	0	.63636363
C2	.78280758	0	.57418229
C3	1.0	0	.66666667

Variance stabalizing transformations have also been proposed by Bartlett (1936), Freeman and Tukey (1950), and Anscombe (1948). Peizer and Pratt's approximation is also extended to the Poisson resulting in the following variate

$$d_1 \left\{ \frac{1+g\left\{\frac{x+\frac{1}{2}}{\lambda}\right\}}{\lambda} \right\}^{\frac{1}{2}} \tag{5.9}$$

where $d_1 = x + \frac{2}{3} - \lambda$

$d_2 = d_1 + \frac{.02}{x+1}$.

Numerical comparisons reveal that, although the proposed approximation gives very good results, both the extended Wilson and Hilferty, and the Peizer and Pratt approximations are better. The latter, in particular, is remarkably accurate for all values of λ. The various stabalizing transformations provide only fair results, Some of the numerical comparisons are reproduced in Table 5.7.

We note that Case C3 produces the transformation $\lambda = x^{2/3}$ which has also been proposed, among others, by Moore (1957) and Blom (1954).

TABLE 5.7

POISSON DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) $\lambda = 4$

X	Pr	N	WH	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
0	.0183	-.0217	.0003	.0004	-.0042	-.0031	-.0008	-.0208	-.0004	-.0020	.0012	-.0011
1	.0915	-.0141	.0012	.0003	-.0043	-.0031	-.0027	-.0035	-.0016	-.0015	-.0015	-.0020
3	.4334	.0321	-.0004	.0001	.0010	.0000	-.0019	-.0007	-.0014	-.0003	-.0022	.0011
5	.7851	.0118	-.0009	.0000	.0028	.0023	.0027	.0020	.0015	.0011	.0018	.0017
7	.9488	-.0111	.0001	.0000	-.0001	.0007	.0025	.0015	.0015	.0013	.0016	.0004
9	.9918	-.0052	.0002	.0000	.0008	-.0016	.0008	.0004	.0005	.0006	.0005	.0000
ximum Absolute Deviation		.0321	.0012	.0004	.0043	.0031	.0028	.0208	.0019	.0020	.0029	.0020
erage Absolute Deviation		.0129	.0005	.0000	.0016	.0016	.0017	.0029	.0010	.0008	.0012	.0008

(2) $\lambda = 10$

X	Pr	N	WH	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
2	.0027	-.0061	-.0002	.00000	-.00010	-.0004	-.0001	-.0012	-.0001	-.0004	.0001	-.0002
4	.0292	-.0117	.0000	.00002	-.0012	-.0009	-.0005	-.0010	-.0005	-.0007	-.0003	-.0007
6	.1301	-.0040	.0005	.00002	-.0013	-.0011	-.0011	-.0010	-.0009	-.0006	-.0011	-.0009
8	.3328	.0152	.0003	.00001	-.0003	-.0005	-.0009	-.0006	-.0008	-.0003	-.0011	-.0003
12	.7915	.0061	-.0004	.00000	.0010	.0009	.0009	.0009	.0008	.0005	.0009	.0008
14	.9165	-.0061	-.0001	.00001	.0004	.0007	.0010	.0008	.0008	.0006	.0009	.0006
16	.9729	-.0071	.0001	.00000	.0002	.0003	.0006	.0004	.0005	.0005	.0005	.0003
18	.9928	-.0036	.0001	.00000	.0003	.0000	.0003	.0002	.0002	.0003	.0002	.0001
kimum Absolute Deviation		.0207	.0005	.00002	.0014	.0011	.0011	.0019	.0009	.0007	.0013	.0009
erage Absolute Deviation		.0080	.0002	.00000	.0005	.0005	.0005	.0008	.0004	.0004	.0005	.0004

(3) $\lambda = 30$

X	Pr	N	WH	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
15	.00194	-.00211	-.00004	.000000	.00000	-.00005	-.00003	-.00007	-.00003	-.00007	.00000	-.00004
18	.01293	-.00495	-.00004	.000001	-.00016	-.00014	-.00010	-.00016	-.00010	-.00017	-.00005	-.00013
21	.05444	-.00590	.00005	.000002	-.00037	-.00029	-.00015	-.00027	-.00024	-.00024	-.00024	-.00027
24	.15724	-.00041	.00017	.000002	-.00037	-.00035	-.00036	-.00033	-.00034	-.00021	-.00042	-.00033
27	.33286	.00883	.00010	.000001	-.00015	-.00020	-.00027	-.00022	-.00025	-.00011	-.00035	-.00018
30	.54835	.01199	-.00008	.000000	.00016	.00010	.00002	.00006	.00002	.00003	.00002	-.00010
33	.74444	.00585	-.00017	.000000	.00033	.00029	.00026	.00028	.00025	.00015	.00031	.00028
36	.88037	-.00196	-.00009	.000001	.00026	.00030	.00034	.00032	.00031	.00022	.00036	.00028
39	.95374	-.00484	.00001	.000000	.00012	.00019	.00023	.00021	.00022	.00020	.00023	.00018
42	.98518	-.00358	.00005	.000000	.00010	.00009	.00013	.00011	.00012	.00013	.00011	.00009
imum Absolute Deviation		.01208	.00017	.000002	.00040	.00035	.00036	.00033	.00034	.00024	.00043	.00033
rage Absolute Deviation		.00466	.00007	.000000	.00018	.00017	.00018	.00018	.00017	.00014	.00019	.00016

Note: N is the classical normal approximation
WH is the Wilson-Hilferty approximation (5.8)
PP is the approximation by Piezer-Pratt (5.9)

5.2.4. The Gamma Distribution (Including χ^2)

The large majority of work and transformations of the gamma distribution have been developed in connection with the special case of the χ^2 distribution. Generalization to the gamma is then made through the simple transformation $x = \alpha v$. This relationship produces an interesting result for our transformation. As noted in the introductory remarks in Chapter II, a linear transformation does not generalize our transformation, and a simple relationship between constants may be established. In this case, we have

$$a_1 = \alpha\, a_1'$$

$$a_2 = \alpha^2\, a_2'$$

Also, as with the previous cumulative distributions, relationship (5.2) may also be applied. Then we have, combining with the above result,

$$\left.\begin{aligned} a_1 &= \frac{\alpha\, a_1^*}{p} \\ a_2 &= \frac{\alpha^2\, a_2^*}{p^2}, \end{aligned}\right\} \qquad (5.10)$$

where a_1^*, a_2^* are single unique constants, and h remains unchanged. Table 5.8, below, provides the simple set of constants a_1^*, a_2^*, and h for each of the cases. We note finally that in the special case of a χ^2 random variable, $\alpha = \frac{1}{2}$ and $p = \frac{n}{2}$, and the same constants may be employed by using (5.10).

TABLE 5.8

	a_1^*	a_2^*	h
Case A	1.062287794	.024593178	.328836348
Case B	1.062997246	.023363866	.331489204
Case C1	1.083333333	0.0	.338615385
C2	1.053415633	0.0	.367138055
C3	1.0	0.0	.333333333

Among the best known approximations to the χ^2 distribution are Fisher's (1922) transformation

$$\sqrt{2x} - \sqrt{2n-1} \tag{5.11}$$

and the Wilson and Hilfertz (1931) transformation

$$\left[\left\{\frac{x}{n}\right\}^{\frac{1}{3}} - 1 + \frac{2}{9n}\right]\sqrt{\frac{9n}{2}} \tag{5.12}$$

which are both approximately standard normal. More recently, approximations have been proposed by Gray, Thompson, and McWilliams (1969), Wallace (1958) and Peizer and Pratt. The latter's approximation given by

$$d_2\left[\frac{1+g\left\{\frac{n-1}{x}\right\}}{x/2}\right]^{\frac{1}{2}} \tag{5.13}$$

where $d_2 = x/2 - n/2 + 1/3 - .04/n$

is extremely accurate.

The approximations have been evalued for $\alpha = 1$ and for a range of values of p. Different values of α produce the same results.

The comparisons suggest that the proposed approximation is of the same order of accuracy as that of Peizer and Pratt, and is far superior to the Wilson-Hilferty (special Case C3) and Fisher approximations. A sample of the numerical results reproduced in Tables 5.9 and 5.10, demonstrates the remarkable accuracy of the approximation. For example, for p as low a 1, the maximum error is .3%, and for p = 10, the maximum error is .001%.

5.2.5. The Non-Central χ^2 Distribution

Although the non-central χ^2 is a cumulative distribution, neither of the reproductive parameters ν or λ are separable, and hence (5.2) may not be applied.

Many approximations have been proposed to the non-central χ^2 distribution. Among the simplest was one by Patnaik (1949), which sought to approximate the non-central χ^2 by some multiple of the central χ^2 with different degrees of freedom, that is,

$$c\,\chi_f^2 \tag{5.14}$$

where $c = \dfrac{\nu+2\lambda}{\nu+\lambda}$ $\quad f = \dfrac{(\nu+\lambda)^2}{(\nu+2\lambda)}$.

Pearson (1959) suggested an improvement by introducing an additional constant b. He then approximates the non-central χ^2 with

$$b + c\,\chi_f^2 \tag{5.15}$$

where $b = \dfrac{\lambda^2}{\nu+3\lambda}$, $c = \dfrac{\nu+3\lambda}{\nu+2\lambda}$, $f = \dfrac{(\nu+2\lambda)^3}{(\nu+3\lambda)^2}$.

TABLE 5.9

GAMMA DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) a = 1, p = 1

X	Pr	F	WH	N	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
0.0800	.0769	-.0448	-.0078	-.0919	.0018	-.4445	-.0495	.0032	.0036	.0286	.0066	.0000	.0020
0.1600	.1479	-.0278	-.0018	-.0526	.0010	-.4281	-.0283	.0018	-.0021	-.0070	-.0020	.0080	-.0001
0.3200	.2739	-.0002	.0045	-.0256	-.0009	-.3734	-.0086	-.0029	-.0034	-.0151	-.0039	.0147	.0023
0.4000	.3297	.0095	.0055	.0554	-.0017	-.3314	-.0049	-.0027	-.0032	-.0151	-.0037	.0151	.0033
0.5600	.4288	.0191	.0056	.0088	-.0028	-.2058	-.0023	-.0021	-.0026	-.0133	-.0033	.0132	.0041
0.6400	.4727	.0253	.0052	.1133	-.0031	-.1415	-.0019	-.0017	-.0022	-.0119	-.0030	.0117	.0040
0.8000	.5507	.0320	.0036	.0799	-.0035	.0249	-.0017	-.0013	-.0016	-.0044	-.0024	.0079	.0035
0.8800	.5852	.0280	.0028	.1330	-.0036	.1123	-.0015	-.0010	-.0013	-.0081	-.0025	.0060	.0031
1.0400	.6465	.0258	.0012	.1306	-.0036	.2775	-.0014	-.0007	-.0009	-.0078	-.0023	.0026	.0021
1.1200	.6737	.0240	.0005	.1260	-.0036	.3481	-.0013	-.0006	-.0007	-.0048	-.0022	.0013	.0016
1.2800	.7220	.0198	-.0007	.1111	-.0035	.4494	-.0013	-.0004	-.0005	-.0031	-.0020	-.0015	.0005
1.3600	.7433	.0175	-.0011	.1028	-.0033	.4756	-.0013	-.0002	-.0003	-.0023	-.0019	.0025	.0001
1.5200	.7813	.0128	-.0018	.0828	-.0031	.4736	-.0015	-.0001	-.0002	-.0010	-.0017	-.0041	-.0007
1.6000	.7981	.0107	-.0020	.0624	-.0029	.4462	-.0017	.0000	.0000	-.0004	-.0015	-.0046	-.0010
1.7600	.8280	.0064	-.0024	.0516	-.0027	.3465	-.0034	.0001	.0001	.0005	-.0013	-.0055	-.0016
1.8400	.8412	.0045	-.0025	.0416	-.0026	.2795	-.0050	.0001	.0001	.0009	-.0012	-.0058	-.0018
2.0000	.8647	.0011	-.0025	.0233	-.0023	.1282	-.0103	.0001	.0002	.0015	-.0009	-.0061	-.0022
2.0800	.8751	-.0004	-.0025	.0151	-.0022	-.0507	-.0143	.0002	.0002	.0018	-.0008	-.0062	-.0023
2.2400	.8935	-.0028	-.0024	.0010	-.0019	-.0912	-.0251	.0002	.0003	.0022	-.0006	-.0061	-.0024
2.3200	.9017	-.0039	-.0023	-.0048	-.0018	-.1504	-.0317	.0002	.0003	.0024	-.0005	-.0060	-.0025
2.4800	.9163	-.0056	-.0021	-.0143	-.0016	-.2364	-.0455	.0002	.0003	.0026	-.0003	-.0058	-.0025
2.5600	.9227	-.0063	-.0020	-.0180	-.0016	-.2616	-.0514	.0002	.0003	.0026	-.0002	-.0054	-.0026
2.7200	.9341	-.0072	-.0017	-.0180	-.0013	-.2771	-.0574	.0002	.0004	.0028	.0000	-.0052	-.0025
2.8000	.9391	-.0077	-.0017	-.0249	-.0013	-.2693	-.0571	.0002	.0003	.0027	.0000	-.0051	-.0025
2.9600	.9482	-.0081	-.0014	-.0269	-.0011	-.2296	-.0479	.0002	.0003	.0028	.0002	-.0046	-.0024
3.0400	.9521	-.0082	-.0013	-.0272	-.0010	-.2014	-.0478	.0002	.0003	.0028	.0002	-.0044	-.0023
3.2000	.9592	-.0083	-.0010	-.0268	-.0010	-.1371	-.0407	.0002	.0003	.0027	.0004	-.0039	-.0021
3.2800	.9623	-.0083	-.0010	-.0263	-.0009	-.1043	-.0376	.0001	.0002	.0027	.0004	-.0038	-.0021
3.4400	.9679	-.0081	-.0008	-.0247	-.0007	-.0436	-.0321	.0002	.0003	.0026	.0005	-.0033	-.0019
3.5200	.9704	-.0079	-.0006	-.0237	-.0006	-.0177	-.0296	.0002	.0003	.0026	.0005	-.0029	-.0018
3.6800	.9748	-.0076	-.0005	-.0216	-.0006	.0229	-.0253	.0001	.0002	.0024	.0005	-.0028	-.0017
3.7600	.9767	-.0073	-.0004	-.0204	-.0005	.0374	-.0233	.0001	.0002	.0024	.0006	-.0026	-.0016
3.9200	.9801	-.0069	-.0003	-.0181	-.0005	.0551	-.0199	.0001	.0002	.0022	.0006	-.0023	-.0015
4.0000	.9817	-.0067	-.0003	-.0170	-.0005	.0591	-.0184	.0000	.0001	.0021	.0005	-.0022	-.0015
Maximum Absolute Deviation		.0448	.0078	.1330	.0036	.4837	.0827	.0032	.0036	.0286	.0066	.0151	.0040
Average Absolute Deviation		.0118	.0021	.0507	.0019	.2252	.0225	.0006	.0007	.0048	.0014	.0054	.0020

Note: F is Fisher's approximation (5.11)
WH is an extension to Wilson-Hilferty approximation (5.12)
PP is the Peizer-Pratt approximation (5.13)

TABLE 5.10

GAMMA DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(2) a = 1, p = 5

X	Pr	F	WH	N	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
.2342	.000004	-.000344	-.000013	-.016526	.000000	.011048	-.000277	.000004	.000004	.000004	.000004	-.000012	.000004
.9367	.002777	-.004913	-.000479	-.031817	-.000006	.001855	-.001981	.000082	.000095	.000410	.000081	-.000448	.000055
1.6391	.025904	-.010159	-.000673	-.040513	-.000012	-.013138	-.002498	.000026	.000038	.000321	-.000009	-.000530	-.000025
2.3416	.080693	-.008331	.000139	-.028555	-.000014	-.022040	-.001234	-.000057	-.000003	-.000191	.000034	.000410	-.000048
3.0441	.192206	-.000105	.001094	.001335	-.000018	-.020065	-.000296	-.000074	-.000003	-.000535	.000175	.001362	.000025
3.7466	.321798	.008903	.001246	.034238	-.000028	-.008176	-.000053	-.000048	-.000070	-.000585	.000199	.001512	.000088
4.4491	.458212	.014006	.000793	.055512	-.000041	.009203	-.000030	-.000016	-.000034	-.000458	.000067	.000936	.000092
5.1516	.585694	.014062	.000111	.058674	-.000052	.023901	-.000027	.000007	-.000005	-.000282	-.000123	.000116	.000050
5.8541	.694940	.010500	-.000434	.046134	-.000057	.027140	-.000026	.000020	.000015	-.000118	-.000262	-.000540	.000000
6.5566	.782587	.005537	-.000099	.025761	-.000057	.016380	-.000061	.000025	.000024	.000011	-.000005	-.000873	-.000037
7.2591	.849359	.003949	-.000709	.005529	-.000052	-.001492	-.000281	.000025	.000028	.000103	-.000266	-.000910	-.000053
7.9616	.899146	-.002345	-.000566	.009178	-.000044	-.015325	-.000952	.000021	.000026	.000160	-.000180	-.000764	-.000054
8.6641	.932558	-.004163	-.000368	-.016766	-.000036	-.018366	-.002283	.000014	.000021	.000188	-.000084	-.000545	-.000045
9.3666	.956213	-.004758	-.000182	-.018364	-.000027	-.012138	-.004217	.000008	.000015	.000193	-.000003	-.000329	-.000034
10.0691	.972026	-.004538	-.000041	-.016277	-.000021	-.003088	-.006290	.000002	.000009	.000182	.000053	-.000156	-.000023
10.7715	.982390	-.003686	.000005	-.012685	-.000015	.003250	-.007677	-.000003	.000004	.000161	.000084	-.000037	-.000015
11.4740	.989060	-.003089	.000098	-.009045	-.000010	.005226	-.007599	-.000005	.000001	.000137	.000095	.000034	-.000009
imum Absolute Deviation		.014549	.001265	.059628	.000057	.027687	.007836	.000082	.000095	.000395	.000305	.001561	.000036
rage Absolute Deviation		.005997	.000454	.025030	.000028	.012496	.002101	.000025	.000032	.000236	.000119	.000557	.000038

(3) a = 1, p = 10

X	Pr	F	WH	N	PP	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
1.2721	.000000	.000000	.000000	.000000	.000000	.001874	-.000009	.000000	.000000	.000000	.000000	-.000001	.000000
2.4105	.000208	-.000697	-.000045	.007989	.000000	-.003008	-.000147	.000005	.000006	.000027	.000002	-.000044	.000004
3.5490	.003648	-.002971	-.000204	-.017827	-.000002	.001976	-.000481	.000011	.000014	.000092	-.000015	-.000195	.000005
4.6874	.021853	-.005097	-.000254	-.024625	-.000004	-.002678	-.000560	.000002	.000006	.000084	-.000031	-.000224	-.000007
5.8258	.072453	-.005691	.000050	-.020964	-.000005	-.006611	-.000295	-.000013	.000011	-.000024	.000025	.000110	-.000011
6.9642	.165893	-.000869	.000469	-.002634	-.000004	-.006409	-.000072	-.000017	-.000019	-.000129	.000120	.000545	.000003
8.1026	.296189	.005706	.000599	.021996	-.000003	-.001953	-.000010	-.000009	-.000017	-.000161	.000144	.000664	.000018
9.2411	.444318	.009916	.000357	.032153	-.000005	.004314	-.000005	.000000	-.000009	-.000133	.000061	.000389	.000018
10.3795	.588559	.009100	-.000032	.040801	-.000008	.008503	-.000004	.000006	-.000001	-.000063	-.000063	-.000038	.000007
11.5179	.712959	.006560	-.000318	.027773	-.000010	.007385	-.000005	.000009	.000006	-.000025	-.000148	-.000354	-.000005
12.6563	.810268	.002979	-.000404	.010723	-.000011	.001568	-.000024	.000008	.000009	.000005	-.000161	-.000457	-.000011
13.7947	.880520	-.001125	-.000334	-.000334	-.000011	-.004372	-.000120	.000006	.000010	.000034	-.000119	-.000388	-.000011
14.9332	.927950	-.003104	-.000197	-.012670	-.000010	-.006213	-.000364	.000003	.000008	.000052	-.000057	-.000245	-.000008
16.0716	.958203	-.003672	-.000067	-.016368	-.000007	-.003838	-.000767	.000001	.000005	.000060	-.000002	-.000105	-.000004
17.2100	.976579	-.003944	.000020	-.012117	-.000005	-.000360	-.001237	-.000001	.000002	.000057	.000032	-.000007	-.000002
18.3484	.987279	-.002604	.000063	-.008575	-.000003	.001654	-.001615	-.000002	.000000	.000050	.000046	.000045	.000000
19.4868	.993282	-.001831	.000074	-.005368	-.000002	.001781	-.001755	-.000002	-.000001	.000040	.000046	.000062	.000000
imum Absolute Deviation		.010381	.000617	.042070	.000011	.008792	.001755	.000022	.000019	.000161	.000164	.000677	.000019
rage Absolute Deviation		.003876	.000273	.016113	.000005	.003760	.000405	.000007	.000007	.000062	.000061	.000228	.000006

Abdel-Aty (1954) produced a transformation similar to Wilson and Hilferty in attempting to stabalize the variance. He proposed

$$\left[\left\{\frac{x^2}{\nu+\lambda}\right\}^{\frac{1}{3}} - 1 + \frac{2(\nu+2\lambda)}{9(\nu+\lambda)^2}\right]\left[\frac{9(\nu+\lambda)^2}{2(\nu+2\lambda)}\right]^{\frac{1}{2}} . \tag{5.16}$$

Sankaran (1959 and 1963) proposed a number of approximations, the last and most accurate of which was obtained from the Haldane A transformation (described in Chapter I). As such, his transformation corresponds to Case C3 of the proposed transformation. Johnson (1959) and Bol'shev and Kuznetzov (1963) suggested other approximations. The former's approximation is only moderately accurate, while the latter's proposal although very accurate for very small λ (same order as Patnaik), gives very poor results for larger λ.

Tables 5.11, 5.12, and 5.13 illustrate the accuracy of the above results and compare them with the porposed transformation. The numerical results suggest that the proposed approximation provides better results than any of the above transformations. Moreover, it provides remarkable accuracy and is far superior to all other transformations, when the non-centrality parameter λ is larger. It should finally be noted that both Patnaik and Pearson's approximations are difficult to apply due to the need of the percentiles of a χ^2 distribution with non-integral degrees of freedom.

5.3: Results for Other Distributions

In this section, we shall examine the results of applying transformation (1.1) to the distribution whose cumulants we have found

TABLE 5.11

NON CENTRAL χ^2 DISTRIBUTION: ACCURACY OF APPROXIMATION

$n = 4, \lambda = .5$

	Pr	Pa	Pe	A	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
6993	.03892	.00001	-.00297	-.00204	-.09748	-.01579	.00050	.00072	.00330	.00102	-.00171	.00010
3987	.12793	-.00010	-.00303	.00142	-.10637	-.00682	-.00053	-.00070	-.00226	-.00049	.00269	-.00027
0980	.23780	-.00048	-.00231	.00344	-.09126	-.00196	-.00056	-.00089	-.00387	-.00050	.00535	.00034
4967	.35121	-.00049	-.00076	.00346	-.05522	-.00062	-.00035	-.00069	-.00374	-.00035	.00553	.00079
1960	.45851	-.00077	-.00071	.00233	-.00245	-.00041	-.00016	-.00044	-.00298	-.00036	.00415	.00105
6954	.55491	-.00073	-.00012	.00087	.05603	-.00035	-.00003	-.00023	-.00209	-.00048	.00220	.00070
2941	.63859	-.00061	.00031	-.00040	.10361	-.00031	.00005	-.00008	-.00127	-.00061	.00035	.00041
9934	.76837	-.00027	.00078	-.00180	.11155	-.00037	.00009	.00008	-.00006	-.00070	-.00207	-.00016
6928	.81664	-.00010	.00088	-.00197	.06955	-.00084	.00009	.00012	.00033	-.00064	-.00259	-.00035
914	.91205	.00005	.00086	-.00135	-.06938	-.01091	.00003	.00012	.00090	-.00026	-.00243	-.00054
7 09	.93178	.00031	.00081	-.00102	-.07707	-.01828	.00002	.00011	.00095	-.00012	-.00210	-.00051
4901	.94727	.00033	.00073	-.00072	-.06423	-.02578	.00000	.00009	.00095	-.00001	-.00175	-.00047
8888	.96881	.00034	.00058	-.00022	-.01474	-.02898	-.00003	.00006	.00087	.00016	-.00110	-.00036
5381	.97011	.00032	.00051	-.00005	.00487	-.02370	-.00004	.00004	.00080	.00020	-.00084	-.00031
2875	.98175	.00030	.00044	.00007	.01591	-.01824	-.00004	.00003	.00073	.00024	-.00062	-.00026
9868	.98609	.00027	.00037	.00016	.01911	-.01390	-.00005	.00002	.00065	.00025	-.00044	-.00021
m Absolute Deviation		.00076	.00303	.00345	.12455	.03011	.00055	.00068	.00386	.00101	.00553	.00087
e Absolute Deviation		.00033	.00095	.00126	.05839	.01038	.00013	.00022	.00136	.00038	.00209	.00038

$n = 10, \lambda = .5$

	Pr	Pa	Pe	A	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
2286	.00035	.00000	-.00002	-.00016	.00783	-.00100	.00005	.00006	.00009	.00007	-.00016	.00003
4571	.00696	.00000	-.00015	-.00062	-.00312	-.00244	.00007	.00010	.00048	.00007	-.00061	.00004
6357	.03314	-.00002	-.00033	-.00052	-.01507	-.00235	.00000	.00002	.00025	-.00002	-.00045	-.00004
1428	.17203	-.00011	-.00046	.00100	-.02079	-.00040	-.00007	-.00010	-.00051	.00014	.00127	.00001
3714	.27616	-.00014	-.00038	.00129	-.01278	-.00009	-.00005	-.00008	-.00060	.00021	.00159	.00008
5999	.38946	-.00015	-.00025	.00104	.00029	-.06004	-.00002	-.00006	-.00033	.00030	.00153	.00009
0571	.60464	.00009	.00005	-.00009	.02500	-.00003	.00002	.00001	-.00025	-.00014	.00000	.00004
2856	.69439	-.00004	.00015	-.00053	.02668	-.00005	.00003	.00002	-.00011	-.00025	-.00054	.00000
5142	.76907	.00000	.00021	-.00076	.01862	-.00005	.00003	.00003	.00000	-.00030	-.00085	-.00003
9713	.87560	.00007	.00025	-.00068	-.00939	-.00054	.00002	.00003	.00014	-.00022	-.00087	-.00005
1999	.91094	.00008	.00023	-.00051	-.01736	-.00131	.00002	.00003	.00018	-.00015	-.00071	-.00005
4284	.93714	.00008	.00020	-.00014	-.01755	-.00256	.00000	.00002	.00019	-.00007	-.00053	-.00005
8356	.96984	.00008	.00015	-.00002	-.00442	-.00592	.00000	.00001	.00019	.00005	-.00019	-.00002
1141	.97945	.00007	.00012	.00005	.00163	-.00729	.00000	.00001	.00017	.00008	-.00007	-.00001
3427	.98613	.00006	.00010	.00011	.00469	-.00779	.00000	.00001	.00015	.00010	.00001	-.00001
5712	.99072	.00005	.00008	.00013	.00504	-.00717	-.00001	.00000	.00013	.00010	.00005	.00000
m Absolute Deviation		.00015	.00046	.00129	.02668	.00779	.00007	.00010	.00060	.00030	.00159	.00009
e Absolute Deviation		.00006	.00019	.00045	.01210	.00212	.00002	.00003	.00023	.00011	.00055	.00003

Pa is Patrick's approximation (5.14)
Pe is Pearson's approximation (5.15)
A is Abdel-Aty's approximations (5.16)

TABLE 5.12

NON CENTRAL χ^2 DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(3) $n = 30, \lambda = .5$

X	Pr	Pa	Pe	A	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
9.2406	.000081	.000000	-.000001	-.000014	.001281	-.000036	.000001	.000001	.000006	-.000001	-.000014	.000001
11.6024	.000860	.000000	-.000003	-.000061	-.000016	-.000119	.000003	.000004	.000024	-.000007	-.000060	.000002
13.9646	.004836	-.000001	-.000009	-.000132	.001059	-.000214	.000004	.000006	.000045	-.000020	-.000120	.000001
18.6890	.047976	-.000008	-.000039	-.000042	-.002794	-.000160	.000005	-.000004	.000012	-.000002	-.000026	-.000005
21.0512	.103117	-.000013	-.000051	.000168	-.003694	-.000070	.000007	-.000009	-.000033	.000050	.000192	-.000003
23.4134	.185921	-.000017	-.000054	.000349	-.002890	-.000018	.000006	-.000009	-.000066	.000097	.000380	.000003
28.1378	.412399	-.000015	-.000025	.000259	.002064	-.000002	.000001	-.000002	-.000070	.000060	.000252	.000001
30.5000	.534033	-.000009	-.000002	.000045	.004124	-.000002	.000003	.000001	-.000053	-.000011	.000056	.000004
32.8622	.647364	-.000002	.000017	-.000153	.004451	-.000002	.000005	.000003	-.000033	-.000076	-.000156	-.000001
37.5866	.822015	.000010	.000037	-.000282	-.000073	-.000012	.000004	.000004	.000001	-.000108	-.000304	-.000004
39.9488	.860746	.000013	.000038	-.000227	-.002499	-.000045	.000003	.000004	.000015	-.000082	-.000251	-.000004
42.3110	.922879	.000013	.000034	-.000141	-.003333	-.000117	.000002	.000003	.000024	-.000045	-.000165	-.000002
47.0354	.970701	.000010	.000021	.000004	-.000926	-.000370	.000000	.000001	.000031	.000014	-.000012	.000000
49.3976	.982707	.000008	.000016	.000041	.000052	-.000503	-.000001	.000000	.000029	.000028	.000023	.000000
51.7598	.990056	.000006	.000011	.000056	.001022	-.000598	-.000001	.000000	.000026	.000032	.000047	.000001
54.1220	.994418	.000004	.000007	.000055	.000958	-.000634	-.000002	-.000001	.000020	.000030	.000049	.000000
Maximum Absolute Deviation		.000017	.000054	.000385	.004451	.000634	.000007	.000009	.000077	.000108	.000416	.000007
Average Absolute Deviation		.000007	.000023	.000137	.001930	.000160	.000002	.000003	.000031	.000043	.000146	.000002

(4) $n = 4, \lambda = 5.0$

X	Pr	Pa	Pe	A	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
1.2437	.016882	.005749	-.006158	.005234	.024836	-.008669	.000422	-.000507	.002204	-.000184	-.001495	.000397
2.4875	.067530	.009185	-.003673	.010801	-.001461	-.004961	-.000232	.000049	.000536	.000432	.000426	-.000138
3.7312	.145908	.006698	-.000464	.010645	-.028665	-.001577	-.000309	.000458	-.000783	.000096	.002209	-.000013
6.2186	.346404	-.003872	.002088	.000197	-.006133	.000083	.000077	.000426	-.002063	.000535	.002302	.000370
7.4624	.450193	-.007367	.001792	-.004849	.009860	.000084	.000253	.000228	-.002067	-.000076	.001266	.000370
8.7061	.547594	-.003792	.001160	-.008004	.024894	.000071	.000351	.000006	-.001781	-.000608	.000157	.000272
11.1935	.710009	-.007083	.000149	-.008794	.034219	.000070	.000323	-.000212	-.000817	-.001063	-.001288	-.000012
12.4373	.773978	-.005140	-.000037	-.007416	.024190	.000033	.000237	-.000260	-.000330	-.001003	-.001519	-.000128
13.6810	.825986	-.003122	-.000047	-.005580	.007717	-.000211	.000134	-.000270	.000082	-.000827	-.001498	-.000208
16.1684	.900587	.000127	.000197	-.001983	-.020769	-.002540	-.000056	-.000213	.000609	-.000359	-.001049	-.000255
17.4122	.926072	.001170	.000339	-.000604	-.024229	-.005144	-.000127	-.000168	.000730	-.000149	-.000767	-.000257
18.6559	.945552	.001839	.000453	.000420	-.020249	-.005918	-.000175	-.000121	.000776	.000020	-.000506	-.000236
21.1433	.971224	.002309	.000560	.001513	-.003364	-.014985	-.000213	-.000037	.000718	.000221	-.000118	-.000174
22.3871	.979324	.002254	.000557	.001697	.003063	-.015482	-.000211	-.000006	.000647	.000261	.000002	-.000144
23.6308	.985249	.002089	.000527	.001721	.006254	-.013527	-.000198	.000017	.000565	.000273	.000081	-.000116
24.8745	.989545	.001863	.000480	.001639	.006641	-.010320	-.000180	.000033	.000430	.000265	.000125	-.000092
Maximum Absolute Deviation		.009185	.006158	.010801	.034275	.015482	.000422	.000545	.002204	.001063	.002788	.000397
Average Absolute Deviation		.003905	.001149	.004342	.016997	.005010	.000202	.000191	.000921	.000473	.000951	.000191

TABLE 5.13

NON CENTRAL χ^2 DISTRIBUTION: ACCURACY OF APPROXIMATION

(5) $n = 10, \lambda = 5.0$

X	Pr	Pa	Pe	A	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
1.6987	.00021	.00007	-.00024	.00001	.00603	-.00067	.00005	.00003	.00015	.00004	-.00014	.00005
3.3974	.00463	.00072	-.00098	.00046	-.00024	-.00182	-.00007	.00004	.00043	.00001	-.00057	.00006
5.0960	.02396	.00170	-.00139	.00174	-.00821	-.00199	-.00002	-.00001	.00032	-.00005	-.00057	-.00004
8.4934	.14128	.00103	-.00035	.00304	-.01298	-.00045	-.00008	-.00003	-.00035	.00022	.00108	-.00003
10.1921	.23772	-.00061	.00025	.00178	-.00835	-.00010	-.00004	-.00002	-.00053	.00031	.00156	.00004
11.8908	.34825	-.00218	.00052	-.00017	-.00069	-.00001	.00003	.00000	-.00056	.00024	.00139	.00008
15.2881	.57060	-.00322	.00034	-.00310	.01444	.00001	.00008	.00000	-.00036	-.00015	.00013	.00006
16.9868	.66725	-.00271	.00016	-.00341	.01614	.00000	.00007	-.00002	-.00023	-.00031	-.00048	.00001
18.6855	.74907	-.00181	.00007	-.00304	.01188	-.00001	.00006	-.00002	-.00009	-.00037	-.00083	-.00002
22.0829	.86714	-.00002	.00008	-.00140	-.00537	-.00029	.00001	-.00002	.00011	-.00027	-.00088	-.00005
23.7816	.90632	.00061	.00014	-.00059	-.01067	-.00078	.00000	-.00001	.00018	-.00017	-.00070	-.00004
25.4802	.93515	.00099	.00020	.00004	-.01098	-.00162	-.00002	.00000	.00021	-.00007	-.00049	-.00004
28.8776	.97041	.00117	.00024	.00069	-.00257	-.00400	-.00003	.00000	.00021	.00006	-.00014	-.00002
30.5763	.98045	.00108	.00024	.00079	.00130	-.00504	-.00003	.00001	.00019	.00010	-.00002	-.00001
32.2750	.98725	.00094	.00022	.00077	.00316	-.00555	-.00002	.00001	.00016	.00011	.00004	.00000
33.9737	.99178	.00077	.00019	.00069	.00323	-.00531	-.00003	.00000	.00013	.00011	.00008	-.00001
Maximum Absolute Deviation		.00322	.00139	.00340	.01614	.00555	.00008	.00003	.00055	.00037	.00156	.00008
Average Absolute Deviation		.00127	.00035	.00140	.00737	.00150	.00004	.00001	.00023	.00014	.00054	.00003

(6) $n = 4, \lambda = 50.0$

X	Pr	Pa	Pe	A	GC	CF	Case A	Case B	Case C1	Case C2	Case C3	Case D
15.0600	.000255	.000196	-.000091	.000185	.000642	-.000057	.000004	-.000009	.000007	-.000020	.000000	.000002
19.2867	.001662	.000932	-.000238	.000874	.000687	-.000123	-.000005	-.000015	.000026	-.000046	.000006	.000004
23.7134	.007035	.002550	-.000403	.002422	.000198	-.000167	-.000001	-.000015	.000044	-.000061	.000022	.000004
32.3667	.052725	.005544	-.000264	.005551	-.001349	-.000084	.000011	.000017	.000073	.000068	.000072	.000001
36.6934	.106846	.004261	.000063	.004520	-.001421	-.000027	.000007	.000029	.000027	.000170	.000067	-.000001
41.0200	.186723	.000759	.000351	.001221	-.000834	-.000001	-.000002	.000032	-.000047	.000214	.000025	-.000001
49.6733	.406589	-.006926	.000301	-.-06594	.001039	.000002	-.000012	.000011	-.000166	.000043	-.000109	.000007
54.0000	.527372	-.008161	.000057	-.008096	.001624	.000001	-.000009	-.000002	-.000165	-.000093	-.000144	.000006
58.3267	.641115	-.007230	-.000155	-.007414	.001607	.000001	-.000003	-.000012	-.000124	-.000185	-.000140	.000003
66.9800	.820041	-.002173	-.000214	-.002536	-.000157	-.000001	.000005	-.000021	-.000001	-.000172	-.000047	-.000004
71.3066	.880779	.000148	-.000105	-.000160	-.001061	-.000017	.000005	-.000019	.000044	.000106	-.000003	-.000005
75.6333	.924250	.001676	-.000017	.001464	-.001347	-.000058	.000003	-.000015	.000068	.000037	-.000039	-.000006
84.2866	.972799	.002473	.000162	.002436	-.000321	-.000202	.000001	-.000004	.000068	.000045	.000061	-.000004
88.6133	.984565	.002181	.000171	.002193	.000228	-.000273	.000001	.000001	.000055	.000055	.000055	-.000003
92.9400	.991532	.001730	.000152	.001767	.000455	-.000319	.000000	.000004	.000041	.000052	.000044	-.000002
97.2666	.995501	.001269	.000121	.001314	.000405	-.000328	.000001	.000006	.000029	.000043	.000032	-.000001
Maximum Absolute Deviation		.008161	.000442	.008096	.001624	.000327	.000012	.000031	.000166	.000214	.000144	.000007
Average Absolute Deviation		.003030	.000196	.003036	.000779	.000092	.000004	.000012	.000063	.000087	.000053	.000003

in Chapter IV. These include the hypergeometric, the beta, the F, the t, the non-central t, the non-central F, and the distribution of the correlation coefficient. A subsection shall be devoted to each distribution.

As mentioned in Chapter IV, special transformations were made on the F, the non-central F, and the distribution of the correlation coefficient. This is due to the fact that once the solution for the constants in transformation (1.1) was found, the terms in the expansion of the cumulants of y do not behave as in (3.2). Heuristically, this means that either the expansions of the cumulants of y do not converge rapidly, or the ordering of the terms as defined in (3.1) does not accurately represent the magnitude of the leading terms. In effect, our method of solving for constants so that the leading terms in the cumulants of the transformed variate are set to zero, does not reduce the required cumulants of y to a sufficiently small number. With the special transformation on all three distributions, our variate y provides excellent approximations.

5.3.1. The Hypergeometric Distribution

The hypergeometric distribution may be approximated by a binomial or Poisson distribution. However, when the parameters are large, the terms of the binomial and Poisson are similarly difficult to calculate. The simplest approximation is the ordinary normal approximation

$$\left(x - \frac{N_1 n}{N} + \frac{1}{2}\right)\left(\frac{N_1 n\left\{(N_1-1)(n-1)N+N(N-1)-N_1 n(N-1)\right\}}{N^2(N-1)}\right)^{-\frac{1}{2}} \qquad (5.17)$$

Feller (1957) derived a refined but much more complicated formula which is later improved by Nicholson (1956). He suggests as an approximation to the distribution function,

$$\left(\frac{N+1}{N+2}\right) \exp(R) \left\{ \Phi(V_1) - \Phi(V_2) \right\} \tag{5.18}$$

where $R = \dfrac{5(1-p_1 q_1)(1-st)}{36\sigma^2} + \dfrac{2}{3(N+2)}$

$$V_1 = \frac{(x+1-(n+1)p_1)}{\sigma} + \frac{a}{\sigma}\left\{\frac{x+1-(n+1)p_1}{\sigma}\right\}^2 + \frac{2a}{\sigma}$$

$$V_2 = \frac{(w-(n+1)p_1)}{\sigma} + \frac{a}{\sigma}\left\{\frac{w-(n+1)p_1}{\sigma}\right\}^2 + \frac{2a}{\sigma}$$

$$w = \max(0, n+N_1-N), \quad p_1 = \frac{N_1+1}{N+2}, \quad q_1 = 1-p_1$$

$$s = \frac{n+1}{N+2}, \quad t = 1-s,$$

$$a = \frac{1}{6}(p_1-q_1)(t-s), \quad \sigma = \sqrt{(n+1)tp_1q_1}$$

and $\Phi(x)$ is the normal distribution function.

This rather complicated expression is accurate only for very large N and when the distribution is highly skewed. Wise (1954) proposed another approximation in terms of incomplete betas. He approximates the distribution function by

$$I_h(n-x, x+1) + \frac{n(n-1)}{24\,M^2} I_2 \tag{5.19}$$

where $M = N - \frac{1}{2}n + \frac{1}{2}$, $h = \left\{ N - N_1 - \frac{1}{2}n + \frac{1}{2}x + \frac{1}{2} \right\} / M$

$$I_2 = (n+1) I_h(n-x, x+1) - (n-x+1) I_h(n-x-2, x+1)$$
$$- (x+2) I_h(n-x, x+1) + 2 I_h(n-x-1, x)$$

and $I_p(a,b)$ is the incomplete beta function.

Numerical comparisons are made of the above approximations with the proposed one. The results reproduced in part in Table 5.14 suggest that Wise's approximation provides the best approximation in most cases. The proposed transformation however, provides very good results throughout the range, and for all values of the parameters. It should be noted, however, that Wise's approximation has the disadvantage that it depends heavily on the availability of tables of the incomplete beta function, and thus its applicability is limited.

5.3.2. The Beta Distribution

Since there is a direct relationship between the beta, the F and the binomial distribution if a and b are integers, approximations of one are also approximations of the other. Therefore only a few of the multitude of approximations are considered. The Paulson-Camp approximation may be extended to give us,

$$\frac{1}{3}\left\{ \left(\frac{bx}{a(1-x)}\right)^{1/3} \left(1 - \frac{1}{9b}\right) - 1 + \frac{1}{9a} \right\} \left[\left(\frac{bx}{a(1-x)}\right)^{2/3} \frac{1}{9b} + \frac{1}{9a} \right]^{-1/2} . \quad (5.20)$$

The Peizer Pratt system also applies here, and we obtain

TABLE 5.14

HYPERGEOMETRIC DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) $n = 3$, $N = 10$, $N_1 = 7$

X	Pr	N	F-N	W	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
0	.008	.003	.016	.000	-.004	-.004	-.006	-.010	-.008	-.008	-.011
1	.183	.012	.052	.009	-.024	-.019	.027	-.019	-.005	-.010	-.003
2	.708	.008	.057	.000	.019	.015	.086	-.035	.033	.000	.009
Maximum Absolute Deviation		.012	.057	.009	.024	.019	.086	.040	.035	.010	.015
Average Absolute Deviation		.007	.041	.003	.011	.009	.032	.021	.016	.007	.005

(2) $n = 5$, $N = 15$, $N_1 = 10$

X	Pr	N	F-N	W	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
0	.000	.000	-.001	.000	.000	.000	-.001	-.001	-.001	-.002	-.001
1	.016	-.003	-.013	-.001	-.005	-.006	-.008	-.008	-.007	-.007	-.007
2	.166	-.008	-.033	-.001	-.013	-.012	-.012	-.014	-.007	-.005	-.006
3	.566	-.008	-.037	-.014	-.005	.002	.006	.004	.005	.001	.001
4	.916	.011	-.042	.000	.012	.012	.030	.030	.021	.007	.007
Maximum Absolute Deviation		.011	.042	.014	.013	.012	.030	.030	.021	.007	.007
Average Absolute Deviation		.006	.024	.002	.006	.006	.011	.011	.009	.004	.004

(3) $n = 20$, $N = 50$, $N_1 = 5$

X	Pr	N	F-N	W	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
0	.067	-.009	-.018	.000	-.006	-.008	-.028	-.028	-.023	-.004	-.004
1	.325	.009	-.023	.001	-.008	-.006	-.020	-.019	-.019	-.001	-.002
2	.630	.007	-.023	.001	.007	.007	.005	.004	-.007	.001	.001
3	.924	.001	-.032	.001	.008	.008	.010	.010	.005	.004	.004
4	.992	.001	-.047	.000	.002	.002	.005	.005	.004	.004	.004
Maximum Absolute Deviation		.009	.047	.001	.008	.008	.028	.028	.023	.004	.004
Average Absolute Deviation		.005	.028	.000	.004	.004	.009	.009	.005	.002	.002

(4) $n = 50$, $N = 500$, $N_1 = 100$

X	Pr	N	F-N	W	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
2	.0008	-.0018	-.0005	.0000	-.0001	-.0001	-.0001	-.0038	.0002	-.0002	-.0001
4	.0143	-.0059	-.0031	.0000	-.0008	-.0007	-.0006	-.0098	-.0001	-.0008	-.0006
6	.0915	-.0037	-.0058	.0000	-.0012	-.0014	-.0013	-.0034	-.0011	-.0009	-.0010
8	.2955	-.0073	-.0065	.0000	-.0010	-.0010	-.0010	-.0015	-.0014	-.0004	-.0009
10	.5551	.0113	-.0064	.0000	.0004	.0006	.0006	-.0004	.0001	.0001	.0001
12	8254	.0114	-.0070	.0000	.0014	.0013	.0012	.0011	.0011	.0006	.0009
14	.9499	-.0041	-.0088	.0000	.0009	.0009	.0008	.0013	.0010	.0009	.0009
16	.9896	-.0026	-.0113	.0003	.0002	.0002	.0002	.0007	.0004	.0005	.0004
18	.9985	-.0007	-.0127	-.0004	.0001	.0000	.0000	.0002	.0001	.0002	.0001
Maximum Absolute Deviation		.0116	.0127	.0009	.0014	.0014	.0013	.0126	.0014	.0009	.0010
Average Absolute Deviation		.0043	.0065	.0001	.0006	.0006	.0006	.0023	.0006	.0005	.0005

Note N is the classical normal approximations (5.17)
F-N is the Feller-Nicholson approximation (5.18)
W is the approximation by Wise (5.19)

$$d_2 \left\{ \frac{1 + xg\left(\frac{b - \frac{1}{2}}{(a+b-1)(1-x)}\right) + (1-x)g\left(\frac{a - \frac{1}{2}}{(a+b-1)x}\right)}{\left\{a+b-\frac{5}{6}\right\}x(1-x)} \right\}^{\frac{1}{2}} \tag{5.21}$$

where $d_1 = b - \frac{1}{3} - \left\{a+b-\frac{2}{3}\right\}(1-x)$

$$d_2 = d_1 + .02\left\{\frac{x}{b} - \frac{1-x}{a} \quad \frac{x-\frac{1}{2}}{a+b}\right\}$$

The best results as shown by Table 5.15 are generally obtained by the Paulson-Camp approximation, with Peizer-Pratt following close behind. The proposed distribution, however produces valuable results, especially when the distribution is very skew. In this case, both the Paulson-Camp and the Peizer-Pratt do not perform as well, and transformation (1.1) gives best results. It should also be noted that for large a and b, the Cornish-Fisher expansion gives excellent results.

5.3.3. "Student's" Distribution

Intense study has been conducted into the search of possible approximations to the t distribution, with the result that many approximations with very high accuracy have been produced. The first well-known approximation was given by Fisher (1925), who gave as the cumulative distribution,

$$\int_t^\infty v\,dt + v\left\{\frac{t}{4n}(t^2+1) + \frac{t}{96n^2}(3t^6 - 7t^4 - 5t^2 - 3)\right.$$
$$+ \frac{t}{384n^3}(t^{10} - 11t^8 + 14t^6 + 6t^4 - 3t^2 - 15)$$
$$\left. + \frac{t}{92160n^4}(15t^{14} - 375t^{12} + 2225t^{10} - 2141t^8 - 939t^6 - 213t^4 - 915t^2 + 945)\right\} \tag{5.22}$$

TABLE 5.15

BETA DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) a = 1.5, b = 5.0

X	Pr	N	P-C	P-P	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
.0138	.014	-.065	-.005	-.002	-.033	-.030	-.112	-.113	-.039	-.004	-.018
.0692	.139	-.007	.000	-.008	.005	-.012	-.071	-.071	.022	.006	.004
.1246	.293	.048	.003	-.010	.025	-.001	-.032	-.033	-.001	.013	.020
.3462	.782	.009	-.001	-.006	-.010	.003	-.003	-.002	-.011	-.009	-.009
.4015	.853	-.013	.000	-.004	-.001	.000	-.004	-.004	-.010	-.010	-.010
.4569	.905	-.024	-.001	-.003	.004	-.007	-.004	-.003	-.008	-.008	-.009
.5123	.942	-.024	-.000	-.002	.004	-.015	-.002	-.002	-.005	-.005	-.005
.5677	.967	-.018	.000	-.001	.002	-.020	.000	.000	-.001	-.001	-.002
.6231	.982	-.012	.000	-.001	-.001	-.016	.001	.001	.000	.000	.000
.6785	.992	-.006	.001	.000	.000	-.007	.002	.003	.002	.002	.002
Maximum Absolute Deviation		.073	.005	.010	.033	.030	.112	.113	.044	.013	.020
Average Absolute Deviation		.029	.000	.005	.008	.009	.017	.017	.008	.006	.008

(2) a = 1.5, b = 20.0

X	Pr	N	P-C	P-P	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
.0046	.0204	-.0921	-.0054	-.0026	-.0855	-.0319	-.0924	.0022	.0126	.0161	-.0115
.0231	.1856	-.0067	.0016	-.0095	-.0510	-.0099	-.0570	-.0048	-.0026	-.0019	.0010
.0416	.3672	.0675	.0038	-.0115	-.0050	-.0038	-.0731	-.0024	-.0054	-.0005	.0077
.0785	.6537	.0892	.0004	-.0094	.0637	-.0004	-.0608	.0009	-.0024	-.0004	.0025
.0970	.7522	.0585	-.0009	-.0076	.0559	.0002	-.0327	.0008	-.0019	-.0012	-.0007
.1154	.8257	.0233	-.0014	-.0058	.0168	-.0010	-.0057	.0004	-.0018	-.0018	-.0025
.1339	.8793	-.0045	-.0014	-.0044	-.0258	-.0069	.0167	-.0001	-.0016	-.0019	-.0030
.1524	.9176	-.0204	-.0012	-.0032	-.0439	-.0197	.0337	-.0004	-.0013	-.0017	-.0027
.1709	.9446	-.0254	-.0007	-.0022	-.0331	-.0340	.0462	-.0004	-.0009	-.0012	-.0020
.1893	.9632	-.0237	-.0004	-.0015	-.0102	-.0346	.0550	-.0004	-.0005	-.0008	-.0013
.2078	.9759	-.0190	-.0002	-.0010	.0070	-.0240	.0613	-.0003	-.0002	-.0003	-.0006
.2263	.9844	-.0038	.0000	-.0007	.0122	-.0155	.0660	-.0002	.0000	-.0001	-.0002
Maximum Absolute Deviation		.0987	.0054	.0115	.0855	.0366	.0975	.0048	.0192	.0161	.0115
Average Absolute Deviation		.0400	.0013	.0055	.0338	.0137	.0524	.0010	.0022	.0015	.0030

(3) a = 15.0, b = 22.5

X	Pr	N	P-C	P-P	GC	CF	Case A	Case B	Case C1	Case C2	Case C3
.1821	.00107	-.00182	-.00005	-.00001	-.00007	-.00007	-.00354	-.00353	-.00171	-.00032	-.00054
.2295	.01103	-.00435	-.00009	-.00002	-.00022	-.00011	-.01000	-.01000	-.00321	-.00071	-.00105
.2768	.05574	-.00363	-.00001	-.00008	.00024	-.00008	-.01718	-.01719	.00013	.00032	.00024
.3242	.17218	.00366	.00015	-.00013	.00043	.00001	-.02169	-.02127	-.00002	.00246	.00289
.4663	.79801	-.00153	-.00013	-.00008	.00034	-.00002	-.00011	-.00016	-.00233	-.00271	-.00297
.5137	.92058	-.00448	-.00005	-.00003	-.00028	-.00004	.00046	.00043	-.00056	-.00114	-.00129
.5611	.97656	-.00276	.00002	-.00001	-.00012	-.00006	.00090	.00089	.00061	.00047	.00047
.6084	.99503	-.00082	.00003	.00000	-.00009	-.00005	.00069	.00069	.00065	.00070	.00074
Maximum Absolute Deviation		.00590	.00017	.00014	.00067	.00011	.02257	.02266	.00355	.00276	.00328
Average Absolute Deviation		.00277	.00007	.00006	.00024	.00005	.00540	.00540	.00176	.00116	.00135

Note: N is the classical normal approximation
P-C is the extended Paulson-Camp approximation (5.20)
P-P is the approximation of Peitzer and Pratt (5.21)

where $v = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} t^2\right)$.

Some simpler but very accurate normalizing transformations were produced by Hendricks (1936),

$$\sqrt{2(n+1)}\ ct \left\{t^2 + 2n\right\}^{-\frac{1}{2}} \tag{5.23}$$

where $c = 1 - \frac{3}{4}(n-1) - \frac{7}{32}(n+1)^{-2}$,

by Wallace (1959)

$$\left[1 - \frac{2}{8t+3}\left\{1 - \exp(-S^2)\right\}^{\frac{1}{2}}\right] \tag{5.24}$$

where $S = .183(8n+3)\ n^{-1}\left[\log\left\{1+t^2n^{-1}\right\}\right]^{-\frac{1}{2}}$,

by Peizer Pratt

$$\left\{n - \frac{2}{3} + \frac{1}{10}n^{-1}\right\}\left(\frac{\log\ (1+t^2n^{-1})}{n - 5/6}\right)^{\frac{1}{2}} \tag{5.25}$$

by Elfving (1955), Hotelling and Frankel (1938), and finally by Hill (1969)*,

$$u + \frac{1}{48}(u^3 + 3u)\ a^{-2} - \frac{1}{23040}(4u^{-7} + 33u^5 + 240u^3 + 855u)\ a^{-4} \tag{5.26}$$

where $u = \left[a \log\left\{1 + t^2n^{-1}\right\}\right]^{\frac{1}{2}}$

$a = n - \frac{1}{2}$.

Table 5.16 demonstrates the relative accuracy of the above results, and compares them with the proposed transformation. Generally the following observations may be made. Hill's approximation is by

---

* Hill, G.W. (1969) Progress Results on Asymptotic approximation for Student's t and chi-squared, Personal Communication found in Johnson and Kotz, p. 108.

TABLE 5.16

STUDENT'S DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) n = 7

X	Pr	F	H	W	P-P	Hi	GC	CF	Case B
-3.408	.0056	-.0001	-.0019	.0000	-.0001	-.0000003	-.0061	.0056	-.0284
-2.982	.0102	-.0001	-.0020	.0000	-.0001	-.0000004	-.0103	.0102	-.0241
-1.846	.0537	-.0001	-.0013	.0000	.0000	-.0000003	-.0345	.0315	-.0066
-0.710	.2503	-.0001	-.0001	.0002	.0002	-.0000002	.2360	-.2042	.0103
1.420	.9006	.0000	.0007	-.0001	-.0001	.0000004	.0547	.0606	-.0054
1.938	.9564	.0001	.0016	.0001	.0001	.0000003	-.0950	.0357	.0083
2.414	.9767	.0000	.0019	.0001	.0001	.0000004	.9023	-.0231	.0175
3.124	.9916	-.0004	.0020	.0001	.0001	.0000003	.9845	-.0083	.0373
3.408	.9943	-.0004	.0017	.0000	.0001	.0000003	.0061	-.0056	.0510
Maximum Absolute Deviation		-.0005	.0020	.0002	.0002	.0000004	.2360	.2070	.0554
Average Absolute Deviation		.0001	.0010	.0000	.0000	.0000002	.0959	.0678	.0150

(2) n = 10

X	Pr	F	H	W	P-P	Hi	GC	CF	Case B
-3.354	.00365	-.00001	-.00078	-.00002	-.00003	-.0000003	-.00213	.00294	-.00534
-2.683	.01148	-.00001	-.00092	-.00003	-.00004	-.0000003	.00414	.00597	-.00357
-2.012	.03593	-.00001	-.00079	-.00003	-.00003	-.0000004	.01221	.00589	.00084
-1.342	.10469	-.00001	-.00033	.00005	.00004	-.0000004	-.00388	.00443	.00834
-0.537	.30161	-.00001	-.00001	.00009	.00008	-.0000002	-.02452	-.01213	
0.134	.55203	.00000	.00000	-.00003	-.00003	.0000000	-.00817	-.00390	.00961
0.939	.81510	.00000	.00001	-.00009	-.00009	.0000003	.01066	.01154	-.00346
1.744	.94413	.00000	.00062	.00000	.00001	.0000004	-.00967	-.00281	-.01189
2.415	.98181	.00001	.00092	.00003	.00004	.0000004	-.00848	-.00676	-.00279
3.086	.99423	-.00007	.00085	.00002	.00004	.0000004	.00056	-.00409	.00252
Maximum Absolute Deviation		.00007	.00092	.00009	.00009	.0000004	.02551	.01296	.01203
Average Absolute Deviation		.00001	.00049	.00004	.00004	.0000003	.01030	.00638	.00577

(3) n = 20

X	Pr	F	H	W	P-P	Hi	GC	CF	Case B
-2.909	.00433	-.00001	-.00019	-.00001	-.00001	.0000000	-.00020	.00038	-.00068
-2.277	.01696	-.00001	-.00022	-.00001	-.00001	.0000000	.00063	.00043	-.00038
-1.644	.05786	-.00001	-.00014	.00001	.00000	-.0000001	.00105	.00014	.00085
-1.012	.16182	-.00001	-.00004	.00002	.00001	.0000000	-.00124	-.00046	.00214
-0.379	.35416	-.00001	.00000	.00002	.00001	.0000000	-.00197	-.00046	.00146
0.253	.59856	.00000	.00000	-.00002	-.00002	.0000001	.00145	.00032	.00193
0.885	.80677	.00000	.00002	-.00002	-.00002	.0000001	.00175	.00053	-.00226
1.518	.92765	.00000	.00012	-.00001	.00000	.0000000	-.00080	-.00002	-.00115
2.150	.97803	.00000	.00022	.00001	.00001	.0000001	-.00085	-.00041	.00027
2.783	.99425	.00000	.00020	.00001	.00001	.0000000	.00011	-.00041	.00070
Maximum Absolute Deviation		.00001	.00022	.00002	.00002	.0000001	.00230	.00058	.00276
Average Absolute Deviation		.00000	.00011	.00001	.00001	.0000000	.00090	.00035	.00104

Note: F is Fisher's approximation (5.22)
H is Hendrick's approximation (5.23)
W is the approximation by Wallace (5.24)
PP is the Peitzer-Pratt approximation (5.25)
Hi is Hill's approximation (5.26)

far the most accurate approximation available. The maximum error, for example, when $n = 20$ is .000 000 001. The transformations by Wallace and Peizer and Pratt also provide excellent approximations, followed closely by Fisher and then Hendricks. The proposed approximation provides only fair results. As has been mentioned earlier, transformation (1.1), produces better results as the distribution is more skew, and in this case, when we always have symmetry, poor results are to be expected. We note here that the proposed transformation has a solution only in Case B. Since $K_3 = 0$, Case A reduces to Case B, and Case C has no solution.

5.3.4. The F Distribution

Because of the direct relationship between the F distribution, the beta distribution and the binomial distribution when the degrees of freedom are even, approximations to one may be applied to the other. Many approximations, however are not made directly to the F ratio (4.19), but to $Z = \frac{1}{2} \log F_{\nu_1, \nu_2}$ as we have done...

A very simple approximation may be arrived at by using the standard normal transformation to obtain,

$$\frac{Z + \frac{1}{2}\delta}{\sqrt{\frac{1}{2}\gamma}} \tag{5.27}$$

where $\delta = \frac{1}{\nu_1} - \frac{1}{\nu_2}$ and $\gamma = \frac{1}{\nu_1} + \frac{1}{\nu_2}$.

Carter (1947), Wishart (1957), and Cochran (1940) all derived formulae of the Cornish-Fisher type for the quantile of Z. These all used approximate formulae for the cumulants of Z, and only a few terms in the expansion were used. Aroian (1947) made a number of numerical comparisons, and showed that the one suggested by Paulson (1942) gives generally a better approximation. Paulson's approximation, based on the Wilson-Hilferty approximation to the χ^2, may be written as

$$\frac{\left(1 - \frac{2}{9\nu_2}\right) X^{\frac{1}{3}} - 1 + \frac{2}{9\nu_1}}{\left[\frac{2}{9\nu_2} X^{\frac{2}{3}} + \frac{2}{9\nu_1}\right]^{\frac{1}{2}}} \tag{5.28}$$

The Peizer-Pratt system also applies here, and yields

$$d_2 \left[\frac{1 + (1-v)\, g\left(\frac{\frac{1}{2}(\nu_2 - 1)}{\left[\frac{\nu_1+\nu_2}{2} - 1\right] v}\right) + v\, g\left(\frac{\frac{1}{2}(\nu_1 - 1)}{\left[\frac{\nu_1+\nu_2}{2} - 1\right](1-v)}\right)}{\frac{\nu_1+\nu_2}{2} - \frac{5}{6} \quad x\,(1-x)} \right]^{\frac{1}{2}} \tag{5.29}$$

where $v = \frac{\nu_2}{\nu_1 x + \nu_2}$

$$d_1 = \frac{1}{2}(\nu_2 - 1) + \frac{1}{6} - \left[\frac{\nu_1+\nu_2}{2} - \frac{2}{3}\right] v$$

$$d_2 = d_1 + .02\left[\frac{1-x}{\nu_2/2} - \frac{x}{\nu_1/2} + \frac{\frac{1}{2} - x}{\frac{\nu_1+\nu_2}{2}}\right].$$

The above approximations together with the proposed transformation have all been calculated, and comparisons made. These reveal, as illustrated in Table 5.17, that the Peizer-Pratt and the Cornish-Fisher expansion

TABLE 5.17

F DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) $\nu_1 = 2,\ \nu_2 = 13$

X	Pr	N	P	P-P	CC	CF	Case A	Case B	Case C1	Case C3
0 1089	.1023	.0549	-.0056	.0013	.0305	-.0019	.0012	.0005	.0031	.0239
0 5443	.4070	-.0245	.0045	-.0022	- 0636	-.0002	-.0021	-.0016	.0041	.1329
0 9797	.5985	-.0475	.0021	-.0032	.0180	-.0011	-.0037	-.0024	.0023	.1922
1 4152	.7220	- 0413	-.0001	-.0039	.0590	-.0018	- 0044	-.0029	.0003	.2157
1 8506	.8037	-.0294	-.0011	-.0035	.0702	-.0021	.0043	-.0029	- 0014	.2367
2 2860	.8539	-.0185	-.0014	-.0031	.0567	-.0022	-.0035	-.0025	-.0029	.2474
2 7215	.8970	-.0105	-.0013	-.0026	.0569	-.0021	-.0023	-.0018	-.0040	.2507
3 1569	.9237	.0047	-.0010	-.0021	.0450	-.0020	-.0009	- 0009	- 0047	.2493
3 5923	.9427	- 0007	-.0008	-.0017	.0332	-.0019	.0003	-.0000	-.0051	.2450
4.0278	.9564	-.0019	-.0006	-.0017	.0223	- 0017	.0016	.0007	- 0052	.2390
4 4632	.9665	.0036	-.0004	-.0011	.0126	-.0015	.0028	.0015	-.0051	.2320
4.8986	.9740	.0046	-.0003	-.0009	.0044	-.0014	.0038	.0022	-.0049	.2246
5 3341	.9796	.0051	-.0013	-.0008	-.0025	- 0012	.0047	.0028	-.0046	.2171
Maximum Absolute Deviation		0549	.0056	.0039	.0844	.0022	.0049	.0030	.0052	2507
Average Absolute Deviation		.0169	.0013	.0021	0378	.0011	.0027	.0017	.0037	.2100

(2) $\nu_1 = 17,\ \nu_2 = 30$

X	Pr	N	P	P-P	CC	CF	Case A	Case B	Case C1	Case C3
0 2.52	.03012	.01563	- 00102	-.00009	.00185	- 00000	.00007	.00013	-.00066	.01964
0 4704	.1[illegible]06	.01123	.00080	-.00006	.00007	-.00014	.00052	.00039	.00112	.10138
0.7356	.3[illegible]32	-.01059	.00126	-.00006	-.00703	-.00020	.00548	.00001	.00185	.20818
0 9[illegible]08	.5[illegible]08	-.02362	.00054	- 00012	-.00354	-.00015	.00010	-.00040	.00161	.29421
1 2259	.6[illegible]32	-.01081	.00015	- 00018	.00202	-.00008	-.00028	-.00058	00098	.36467
1.4711	.78481	-.01327	-.00048	-.00021	.00503	- 00000	-.00052	-.00057	-.00030	.36766
1.7163	.85710	-.00602	-.00051	-.00021	.00534	.00004	-.00055	-.00043	-.00027	.36882
1 9615	.90571	- 00077	.00040	- 00019	.00413	.00007	-.00044	-.00026	-.00068	.35583
2.2067	.93768	.00241	-.00026	-.00016	.00241	.00008	-.00032	-.00009	-.00092	.33487
2.4519	.95902	.00403	-.00013	-.00012	.00080	.00009	-.00021	00005	-.00101	.31016
2 6971	.97256	.00460	-.00005	-.00009	-.00044	.00008	-.00011	.00015	- 00101	28442
2 9423	.98194	.00456	.00001	-.00007	-.00129	.00008		.00022	-.00094	.25923
Maximum Absolute Deviation		.02379	.00131	.00022	.00703	.00020	.00059	.00060	.00186	.37046
Average Absolute Deviation		00925	.00043	.00012	.00286	.00006	.00033	.00027	.00089	.25712

(3) $\nu_1 = 15,\ \nu_2 = 30$

X	Pr	N	P	P-P	CC	CF	Case A	Case B	Case C1	Case C3
.0516	.00000	.00000	.00000	-.00000	.00000	.00000	.00000	-.00000	.00000	.00000
.2581	.00385	.00129	-.00021	-.00001	-.00017	.00003	.00013	-.00008	- 00195	.00382
.4645	.05932	.00881	-.00009	-.00003	.00055	.00001	.00009	- 00002	-.00068	.05815
6710	.20587	.00208	.00039	-.00000	-.00007	-.00001	-.00020	.00011	.00090	.20147
.8775	.40645	-.00737	.00032	-.00000	-.00094	-.00002	-.00038	.00011	.00135	.38340
1.7033	.89544	-.00170	-.00006	- 00005	.00059	.00001	-.00015	-.00011	.00001	.71066
1 9098	.93582	- 00008	- 00007	-.00004	.00030	.00001	.00009	-.00001	-.00028	.69668
2.1163	.96071	.00070	.00000	- 00002	.00002	.00002	.00030	.00008	-.00045	.66621
2.3227	.97590	.00096	.00004	-.00002	-.00015	.00002	.00043	.00017	-.00051	.62664
2.5292	.98514	.00096	.00005	-.00001	-.00023	.00002	.00050	.00024	-.00050	.58286
Maximum Absolute Deviation		.01000	.00042	.00005	.00094	.00003	.00053	.00025	.00283	.71066
Average Absolute Deivation		.00371	.00014	.00035	.00001	.00002	.00028	.00009	.00069	.45025

Note: N is the standard normal approximation to the transformed variate, (5.27)
P is Paulson's approximation (5.28)
PP is the Peitzer-Pratt approximation (5.29)

(to six terms) are of equal order of accuracy, and both provide excellent approximation throughout the range. The proposed and then the Paulson-Camp approximations, although slightly less accurate as the ones above, provide simple and very accurate results. Again, transformation (1.1) provides excellent approximations which are better then all other when the distribution is very skew, that is, when $\frac{\nu_1}{\nu_2}$ is very small or very large.

5.3.5. The Non-Central t Distribution

The simplest approximation to use is the standard normal approximation

$$\frac{x - K_1}{\sqrt{K_2}} \tag{5.30}$$

where K_1, K_2 are the cumulants of $t_\nu(\lambda)$ given by (4.22). Jennett and Welch (1939) introduced further refinements to obtain the following normalizing transformation,

$$\left\{1 + t^2 \nu^{-1} a\right\}^{-\frac{1}{2}} \left\{-\lambda + t\, b\right\} \tag{5.31}$$

where $a = \nu(1-b^2)$

$$b = \left[\frac{2}{\nu}\right]^{\frac{1}{2}} \Gamma\left[\frac{1}{2}(\nu+1)\right] / \Gamma\left[\frac{\nu}{2}\right] .$$

Azorin (1953) proposed a variance stabalizing transformation, however it gives poor results as a normalizing transformation, while Merrington and Pearson (1958) fit a Pearson type IV distribution using the first

four moments, and thus obtaining the distribution quantiles.

Many numerical comparisons reproduced in part, in Table 5.18 have been made of the above approximations with the proposed transformation. The comparisons reveal generally that the Jennett and Welch approximation produces very accurate results when the non-centrality parameter is very small, while the proposed transformation is in turn most accurate when the non-centrality parameter is larger. This is to be expected since when λ is small, the distribution is close to the central t where our transformation produced only marginal results, while if λ is large, then the distribution is very skew, which seems to be the condition under which our transformation works best.

5.3.6 The Non-Central F Distribution

Because the non-central F is a ratio of central and non-central χ^2 random variables, Patnaik (1949) used his approximation to the non-central χ^2 to obtain the following approximation to the non-central F ,

$$\left\{1 + \frac{\lambda}{\nu_1}\right\} F_{\nu_1,\nu_2} \tag{5.32}$$

where $\nu = \dfrac{(\nu_1 + \lambda)^2}{(\nu_1 + 2\lambda)}$.

Here again, the complication arises of the need for a percentile of an F distribution with non-integral degrees of freedom. Severo and

TABLE 5.18

NON-CENTRAL T DISTRIBUTION: ACCURACY OF APPROXIMATION

(1) $\nu = 15$, $\lambda = 2.5$

X	Pr	J-W	N	GC	CF	Case A	Case B	Case C1	Case C3
-0.9477	.0003	.0000	-.0010	.0003	-.0001	-.0001	-.0028	.0003	.0000
0.4851	.0219	.0000	-.0140	.0072	-.0011	.0002	-.0012	.0050	.0082
1.9179	.2807	-.0003	.0065	-.0230	-.0029	.0013	.0048	-.0024	-.0159
3.3507	.7510	-.0007	.0252	.0283	.0056	-.0028	-.0032	-.0160	.0073
4.7836	.9538	.0016	-.0120	-.0186	-.0006	.0005	-.0010	.0015	.0030
6.2164	.9932	.0014	-.0054	-.0049	-.0022	.0014	.0001	.0047	-.0012
Maximum Absolute Deviation		.0016	.0252	.0335	.0056	.0028	.0053	.0161	.0159
Average Absolute Deviation		.0006	.0103	.0126	.0024	.0010	.0020	.0067	.0053

(2) $\nu = 30$, $\lambda = 5.0$

X	Pr	J-W	N	GC	CF	Case A	Case B	Case C1	Case C3
1.4001	.0001	.0000	-.0012	.0001	.0000	-.0001	-.0002	.0000	.0001
2.8919	.0229	.0001	-.0130	.0011	.0002	-.0001	-.0002	.0009	.0021
4.3836	.2841	-.0008	.0099	-.0047	-.0002	.0003	.0008	-.0010	-.0054
5.8754	.7443	-.0009	.0186	.0068	.0006	-.0004	-.0007	-.0019	.0030
7.3672	.9542	.0012	-.0098	-.0048	-.0002	.0000	-.0001	.0007	.0010
8.8589	.9946	.0009	-.0040	.0016	-.0001	.0002	.0001	.0009	-.0006
Maximum Absolute Deviation		.0012	.0186	.0092	.0006	.0004	.0011	.0021	.0054
Average Absolute Deviation		.0006	.0094	.0034	.0002	.0001	.0003	.0009	.0018

(3) $\nu = 30$, $\lambda = 10.0$

X	Pr	J-W	N	GC	CF	Case A	Case B	Case C1	Case C3
5.0912	.0000	.0000	.0013	.0004	.0000	.0000	.0001	.0000	.0000
7.1583	.0173	.0010	-.0186	.0022	-.0001	-.0001	-.0004	.0005	.0013
9.2254	.2893	-.0023	.0151	-.0093	-.0002	.0001	.0011	-.0008	-.0022
11.2925	.7495	-.0022	.0237	.0138	.0010	-.0004	-.0008	-.0011	.0004
13.3597	.9514	.0025	-.0126	-.0101	-.0008	.0001	-.0005	.0003	.0001
15.4268	.9972	.0058	-.0014	.0035	.0002	.0008	.0005	.0012	.0005
Maximum Absolute Deviation		.0058	.0237	.0175	.0021	.0020	.0018	.0025	.0022
Average Absolute Deviation		.0023	.0121	.0067	.0005	.0003	.0007	.0007	.0009

Note: J-W is the approximation by Jennett and Welch (5.31)
N is the standard normal approximation (5.30)

Zelen (1960) used Paulson's approximation to the F distribution to obtain

$$\frac{\left[1 - \frac{2}{9\nu_2}\right]\left[\frac{\nu_1 x}{\nu_1+\lambda}\right]^{\frac{1}{3}} - 1 + \frac{2(\nu_1 + 2\lambda)}{9(\nu_1 + \lambda)^2}}{\left\{\frac{2(\nu_1 + 2\lambda)}{9(\nu_1 + \lambda)} + \frac{2}{9\nu_2}\left[\frac{\nu_1 x}{\nu_1 + \lambda}\right]^{\frac{2}{3}}\right\}^{\frac{1}{2}}} \tag{5.33}$$

Tiku (1965) attempted to approximate the non-central F with a linear function of a central F with non-integral degrees of freedom, choosing the constants so as to make the first three moments agree. He then proposes

$$b + c\, F_{\nu, \nu_2} \tag{5.34}$$

where $\nu = \frac{1}{2}(\nu_2 - 2)\left\{\sqrt{\frac{H^2}{H^2 - 4M^3}} - 1\right\}$,

$$c = \frac{\nu H}{\nu_1 M(2\nu + \nu_2 - 2)}$$

$$b = -\nu_2\left\{c - 1 - \frac{\lambda}{\nu_1}\right\}(\nu_2 - 2)^{-1}$$

$$H = 2(\nu_1+\lambda)^3 + 3(\nu_1+\lambda)(\nu_1+2\lambda)(\nu_2-2) + (\nu_1+3\lambda)(\nu_2-2)^2$$

$$M = (\nu_1+\lambda)^2 + (\nu_2-2)(\nu_1+2\lambda)$$

Again difficulty is encountered with an F distribution with non-integral degrees of freedom.

Tables 5.19 and 5.20 illustrate the accuracy of the proposed approximation in comparison to above transformations. It is seen that the Severo-Zelen approximation is more accurate when the non-centrality parameter is small, while the Tiku approximation, in turn, yields the

TABLE 5.19

NON-CENTRAL F DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(1) $\nu_1 = 4$, $\nu_2 = 14$, $\lambda = 2.5$

X	Pr	P	S-Z	T	CG	CF	Case A	Case B	Case C1	Case C3
.2231	.0280	-.0016	.0030	-.0067	.0054	-.0006	-.0001	.0007	.0027	.0279
.6692	.1825	-.0093	.0065	-.0021	.0078	-.0006	.0006	.0015	.0043	.1501
1.1154	.3668	-.0113	.0021	.0000	-.0232	-.0005	.0006	-.0000	.0062	.2108
1.5616	.5278	-.0087	-.0015	.0003	-.0087	-.0005	.0001	-.0013	.0057	.1937
2.0077	.6536	-.0047	-.0030	.0003	-.0072	-.0005	-.0006	-.0022	.0042	.1413
2.4539	.7471	-.0014	.0031	.0004	.0163	-.0004	-.0012	-.0025	.0024	.0391
2.9000	.8150	.0008	-.0026	.0004	.0188	-.0003	-.0015	-.0025	.0007	.0495
3.3462	.8641	.0023	-.0019	.0006	-.0173	-.0001	-.0016	-.0021	-.0006	.0229
3.7924	.8994	.0030	-.0012	.0006	.0138	-.0002	-.0014	-.0017	-.0016	.0066
4.2385	.9249	.0032	-.0007	.0005	.0097	-.0001	-.0011	-.0011	-.0024	-.0026
4.6847	.9435	.0031	-.0003	.0005	.0057	.0002	-.0008	-.0006	-.0028	-.0073
5.1308	.9572	.0030	-.0000	.0006	.0022	.0003	-.0004	-.0001	-.0030	-.0094
5.5770	.9672	.0027	.0000	.0005	-.0006	.0004	-.0001	.0002	-.0031	-.0098
6.0232	.9747	.0024	.0002	.0004	-.0028	.0005	.0001	.0005	-.0030	-.0094
6.4693	.9803	.0021	.0002	.0004	-.0044	.0005	.0003	.0008	-.0028	-.0086
Maximum Absolute Deviation		.0113	.0065	.0067	.0232	.0006	.0016	.0025	.0062	.2108
Average Absolute Deviation		.0038	.0017	.0009	.0095	.0004	.0007	.0012	.0029	.0606

(2) $\nu_1 = 7$, $\nu_2 = 14$, $\lambda = .5$

X	Pr	P	S-Z	T	CG	CF	Case A	Case B	Case C1	Case C3
.1336	.00498	.00195	-.00075	.00161	.00068	.00030	-.00019	-.00010	.00058	.00498
.4009	.09715	.01398	.00012	.01206	.00228	-.00003	.00030	.00057	.00082	.09715
.6682	.27062	.01552	.00099	.01359	.00271	-.00016	.00061	.00045	.00955	.27062
.9355	.44911	.00939	.00061	.00833	.00354	-.00015	.00038	-.00009	.00728	.44911
1.2027	.59733	.00286	.00003	.00265	-.00091	-.00009	-.00001	-.00051	.00459	.59733
1.4700	.70958	-.00151	-.00030	.00193	.00160	-.00003	-.00036	-.00071	.00205	.70958
1.7373	.79110	-.00383	-.00041	.00122	.00282	.00001	-.00056	-.00073	.00008	.79110
2.0045	.84926	-.00472	-.00037	-.00329	.00293	.00005	-.00063	-.00064	-.00130	.84926
2.2718	.89053	-.00489	-.00027	-.00411	.00239	.00008	-.00060	-.00050	-.00220	.89053
2.5391	.91985	-.00444	-.00015	-.00418	.00162	.00011	-.00052	-.00034	.00271	.91985
2.8064	.94079	-.00392	-.00005	-.00656	.00084	.00013	-.00041	-.00019	-.00294	.94079
3.0736	.95584	-.00338	.00002	-.00297	.00018	.00014	-.00030	-.00006	-.00297	.05584
3.3409	.96677	-.00264	.00008	-.00250	-.00032	.00015	-.00020	.00004	-.00288	.96677
3.6082	.97476	-.00237	.00012	-.00209	-.00067	.00016	-.00010	.00012	-.00271	.97476
3.8754	.98065	-.00198	.00015	-.00175	-.00089	.00017	-.00003	.00019	-.00250	.98065
Maximum Absolute Deviation		.01630	.00099	.01418	.00384	.00030	.00063	.00074	.00855	.98301
Average Absolute Deviation		.00501	.00029	.00438	.00151	.00011	.00033	.00034	.00322	.68676

Note: P is Patnaik's approximation (5.32)
S-Z is the approximation by Severo and Zelen (5.33)
T is Tiku's approximation (5.34)

TABLE 5.20

NON CENTRAL F DISTRIBUTION: ACCURACY OF APPROXIMATIONS

(3) $\nu_1 = 25$, $\nu_2 = 14$, $\lambda = 2.5$

X	Pr	P	S-Z	T	GC	CF	Case A	Case B	Case C1	Case C3	Case D
.1125	.00000	.00000	-.00000	.00000	-.00000	-.00000	-.00000	-.00000	.00000	-.00001	-.00000
.3375	.00481	.00026	-.00004	-.00008	.00018	-.00002	-.00004	-.00007	.00062	-.01331	.00001
.5625	.06921	.00129	.00014	-.00030	-.00032	-.00002	.00010	.00007	.00119	-.05616	.00001
.7875	.22529	.00147	.00024	-.00033	-.00057	-.00000	.00016	.00018	-.00028	-.07621	-.00012
1.0125	.41316	.00084	-.00003	-.00019	.00042	.00001	.00004	.00010	-.00208	-.07060	-.00009
1.2376	.57921	.00014	-.00034	-.00004	.00093	.00002	-.00008	-.00002	-.00326	-.05425	.00002
1.4626	.70617	.00029	-.00046	.00006	.00063	.00002	-.00015	-.00012	-.00363	-.03735	.00009
1.6876	.79673	-.00048	-.00041	.00010	.00004	.00001	-.00016	-.00016	-.00332	-.02476	.00009
1.9126	.85930	-.00052	-.00026	.00012	-.00041	.00000	-.00014	-.00016	-.00255	-.01638	.00006
2.1376	.90201	-.00048	-.00009	.00011	-.00060	-.00000	-.00011	-.00014	-.00155	-.01105	.00003
2.3626	.93113	-.00042	.00004	.00010	-.00059	-.00001	-.00007	-.00011	-.00044	-.00771	.00000
2.5876	.95108	-.00036	.00016	.00009	-.00046	-.00002	-.00003	-.00007	.00075	-.00558	-.00001
2.8126	.96485	-.00029	.00023	.00005	.00029	-.00003	-.00001	-.00004	.00244	-.00420	-.00003
3.0376	.97446	-.00023	.00028	.00006	-.00012	-.00003	.00001	-.00001	.00341	-.00327	-.00003
3.2626	.98123	-.00018	.00030	.00007	.00000	-.00004	.00002	.00000	.00291	-.00262	-.00003
Maximum Absolute Deviation		.00153	.00046	.00034	.00093	.00004	.00016	.00018	.00363	.07621	.00012
Average Absolute Deviation		.00046	.00020	.00010	.00036	.00002	.00007	.00008	.00187	.02476	.00004

(4) $\nu_1 = 25$, $\nu_2 = 26$, $\lambda = 15.0$

X	Pr	P	S-Z	T	GC	CF	Case A	Case B	Case C1	Case C3	Case D
.1284	.0000	.0000	-.0000	.0000	.0000	.0000	.0000	-.0000	-.9999	.0000	†
.3851	.0001	.0000	-.0000	-.0001	.0000	.0000	-.0000	-.0000	-.9998	.0000	†
.6419	.0096	.0010	.0000	.0000	-.0003	-.0003	-.0005	-.0001	-.9901	.0006	†
.8987	.0673	.0026	.0001	.0005	-.0002	-.0002	-.0006	.0000	-.9319	.0002	†
1.1554	.1971	.0025	.0002	.0007	.0010	.0103	.0009	.0004	-.8030	-.0022	†
1.4122	.3696	.0011	-.0016	.0006	.0091	.0091	.0021	.0003	-.6316	-.0022	†
1.6689	.5399	-.0004	-.0020	.0001	-.0031	-.0031	.0016	-.0000	-.4612	-.0008	†
1.9257	.6811	-.0011	-.0011	-.0001	-.0113	-.0114	.0001	-.0003	-.3191	.0013	†
2.1825	.7862	-.0013	-.0000	-.0003	-.0111	-.0110	-.0012	-.0004	-.2131	.0026	†
2.4392	.8596	-.0011	.0007	-.0003	-.0061	-.0059	-.0019	-.0003	-.1393	.0026	†
2.6960	.9088	-.0009	.0008	-.0003	-.0004	-.0003	-.0019	-.0002	-.0901	.0020	†
2.9527	.9410	-.0007	.0008	-.0002	.0034	.0035	-.0015	-.0001	-.0581	.0012	†
3.2095	.9618	-.0005	.0006	-.0002	.0051	.0055	-.0010	.0000	-.0376	.0004	†
3.4663	.9752	-.0004	.0004	-.0001	.0052	.0058	-.0004	.0000	-.0244	.0000	†
3.7230	.9838	-.0003	.0002	-.0001	.0044	.0052	-.0000	.0001	-.0161	.0004	†
Maximum Absolute Deviation		.0028	.0008	.0007	.0129	.0129	.0021	.0004	.9999	.0022	†
Average Absolute Deviation		.0009	.0003	.0002	.0050	.0052	.0009	.0001	.4174	.0008	†

† no solution found

better results when the non-centrality parameter is large and both are more accurate then Patnaik's approximation. The numerical comparisons also show that the proposed transformation provides excellent results and is consistently more accurate than any of the above approximations throughout the range and for all values of the parameters. A surprising result is that the Cornish-Fisher expansion to six terms very often provides the best approximation, better than either the proposed transformation, or any of the approximations above.

5.3.7. The Distribution of the Sample Correlation Coefficient in Normal Samples

The distribution of r is so complicated that it is very difficult to use without a simple and accurate approximation.

The simplest approximation is one suggested by Fisher (1921),

$$\frac{1}{2}\left[Z - \log\left\{\frac{1+\rho}{1-\rho}\right\}\right]\sqrt{n-3} \tag{5.35}$$

where $Z = \log\left\{\frac{1+r}{1-r}\right\} = \tanh^{-1}(r)$

This is in fact the variance stabalizing transformation for r. Hotelling (1953) added refinements to Z to improve the variance stabalization, and as he conjectured, to improve the normality. He proposed

$$Z - \frac{3Z + r}{4n} - \frac{23Z + 33r - 5r^3}{96n^2} \tag{5.36}$$

Ruben (1966) found that (5.36) gives very good results as a normalizing transformation,* and further proposed the following transformation

$$\frac{\tilde{r}\left(n-\frac{5}{2}\right)^{\frac{1}{2}}-\tilde{\rho}\left(n-\frac{3}{2}\right)^{\frac{1}{2}}}{\left[1+\frac{\tilde{r}^2}{2}+\frac{\tilde{\rho}^2}{2}\right]^{\frac{1}{2}}} \tag{5.37}$$

where $\tilde{r}=\frac{r}{\sqrt{1-r^2}}$, $\tilde{\rho}=\frac{\rho}{\sqrt{1-\rho^2}}$.

The above approximations, together with the proposed transformation have been computed and a comparison made. As illustrated in Tables 5.21 and 5.22, Ruben's and Hotteling's are far superior to Fisher's approximation, and provide excellent approximations especially when p is small. However, for larger p and n, the proposed approximation seems to provide the best accuracy. When p is very small, the distribution of r approaches a t distribution, and as we have seen earlier, the proposed approximation does not provide its best results for symmetric distributions. Again, the Cornish-Fisher expansion to six terms provides excellent approximation of relatively the same order as the proposed approximation.

5.4. Final Conclusions

In the final analysis, it appears that the proposed transformation provides excellent approximations for a wide range of

*Our calculations of Hotelling's approximation does not display the same accuracy as that shown in Ruben's paper.

TABLE 5.21

DISTRIBUTION OF CORRELATION COEFFICIENT: ACCURACY OF APPROXIMATIONS

(1) p = .3, n = 25

X	Pr	F	R	H	GC	CF	Case A	Case B	Case C1	Case C3	Case D
-.550	.00002	.00002	.00000	.00002	-.00003	-.00001	-.00004	-.00013	.00001	.00001	.00000
-.450	.00014	-.00005	-.00001	.00008	-.00014	-.00005	-.00012	-.00024	.00010	.00007	.00000
-.350	.00089	.00028	-.00002	.00036	.00000	-.00009	-.00023	-.00044	.00047	.00036	.00004
-.250	.00408	.00006	-.00003	.00102	.00073	-.00012	-.00030	-.00059	.00121	.00138	.00016
-.150	.01460	-.00075	-.00007	.00218	.00131	-.00016	-.00017	-.00045	.00190	-.00401	.00024
-.050	.04271	-.00314	-.00010	.00314	.00017	-.00025	.00033	.00020	.00152	.00770	.00011
.050	.10452	-.00727	-.00010	.00350	-.00258	-.00008	.00101	.00112	-.00072	-.00419	-.00032
.150	.21721	-.01175	-.00007	.00270	-.00356	-.00037	.00128	.00149	-.00429	-.01231	-.00067
.250	.39662	-.01321	-.00001	.00095	.00022	.00004	.00060	.00072	-.00756	-.00997	-.00033
.350	.59259	-.01056	.00000	-.00129	.00464	.00058	-.00071	-.00063	-.00901	.00031	.00045
.450	.78776	-.00560	.00015	-.00346	.00193	.00052	-.00136	-.00120	-.00626	.00722	.00062
.550	.92287	-.00341	.00050	-.00406	-.00297	-.00021	-.00063	-.00057	.00012	.00401	.00002
.650	.98308	-.00156	.00000	-.00242	-.00008	-.00067	.00029	.00012	-.00532	-.00099	-.00023
.750	.99872	-.00035	.00024	-.00049	.00027	-.00036	.00026	.00013	-.00113	-.00066	-.00004
Maximum Absolute Deviation		.01321	.00053	.00406	.00464	.00067	.00136	.00149	.00901	.01255	.00067
Average Absolute Deviation		.00337	.00011	.00144	.00114	.00022	.00040	.00045	.00227	.00308	.00018

(2) p = .8, n = 100

X	Pr	F	R	H	GC	CF	Case A	Case B	Case C1	Case C3	Case D
.580	.00001	.00001	.00000	.00001	-.00003	.00000	.00000	.00000	.00000	.00000	.00000
.600	.00003	.00000	.00000	.00001	-.00003	.00000	.00000	.00000	.00000	.00001	.00000
.620	.00011	.00000	-.00001	.00002	.00000	.00000	.00000	.00000	.00000	.00003	.00000
.640	.00037	-.00002	-.00003	.00006	.00007	.00000	.00000	.00000	.00002	.00009	.00000
.660	.00118	-.00011	-.00007	.00012	.00019	.00000	.00000	.00000	.00004	.00019	.00000
.680	.00357	-.00040	-.00015	.00028	.00029	.00000	-.00001	-.00002	.00006	.00032	.00000
.700	.01022	-.00113	-.00024	.00041	.00028	.00000	-.00001	-.00002	.00010	.00044	.00000
.720	.02721	-.00278	-.00033	.00063	.00000	.00000	.00000	-.00001	.00014	.00035	.00000
.740	.06644	-.00585	-.00032	.00079	-.00054	.00000	.00000	.00001	.00013	-.00014	.00000
.760	.14603	-.01027	-.00020	.00075	-.00086	-.00001	.00002	.00005	.00004	-.00093	.00000
.780	.28539	-.01462	-.00005	.00045	-.00027	.00000	.00003	.00007	-.00012	-.00127	-.00001
.800	.48387	-.01613	-.00001	-.00005	.00093	.00002	.00001	.00003	-.00026	-.00043	.00000
.820	.70370	-.01306	.00003	-.00061	.00081	.00002	-.00002	-.00004	-.00025	.00080	.00000
.840	.87905	-.00625	.00028	-.00096	-.00064	.00000	-.00002	-.00006	-.00005	.00081	.00001
.860	.96992	-.00251	.00041	-.00075	-.00026	.00000	.00000	.00000	.00011	.00000	.00000
.880	.99635	-.00048	.00019	-.00026	.00020	.00000	.00000	.00000	.00007	-.00018	.00000
.900	.99985	-.00003	.00002	-.00002	-.00002	.00000	.00000	.00000	.00001	-.00003	.00000
Maximum Absolute Deviation		.01613	.00041	.00096	.00115	.00002	.00003	.00007	.00028	.00127	.00001
Average Absolute Deviation		.00373	.00012	.00030	.00026	.00000	.00000	.00001	.00007	.00030	.00000

Note: F is Fisher's approximation (5.35)
R is the approximation by Ruber (5.37)
H is Hotelling's approximation (5.36)

TABLE 5.22

DISTRIBUTION OF CORRELATION COEFFICIENT: ACCURACY OF APPROXIMATIONS

(3) p = .2, n = 200

X	Pr	F	R	H	GC	CF	Case A	Case B	Case C1	Case C3	Case D
-.080	.000040	.000005	.000001	.000007	.000000	.000001	.000001	.000001	.000007	.000019	.000001
-.040	.000330	.000002	-.000004	.000021	-.000006	-.000004	-.000005	-.000007	.000021	.000120	-.000004
.000	.002210	-.000007	.000000	.000091	.000006	.000000	-.000002	-.000006	.000076	.000450	.000000
.040	.011040	-.000152	-.000008	.000222	.000010	-.000005	-.000008	-.000013	.000128	.000547	-.000001
.080	.042040	-.000655	-.000011	.000380	-.000008	-.000005	-.000002	-.000007	.000091	.000158	.000000
.120	.122820	-.001624	-.000002	.000428	-.000038	.000003	.000016	.000012	-.000093	-.000812	.000000
.160	.278230	-.002320	-.000002	.000269	-.000013	-.000001	.000012	.000011	-.000317	-.001237	-.000010
.200	.497160	-.002340	-.000004	-.000008	.000052	-.000003	-.000004	-.000003	-.000413	-.000231	-.000005
.240	.720280	-.002151	.000003	-.000282	.000024	.000002	-.000012	-.000010	-.000338	.000908	.000009
.280	.882210	-.001222	.000016	-.000439	-.000036	.000003	-.000007	-.000005	.000000	.000735	.000007
.320	.964240	-.000365	.000016	-.000382	-.000018	-.000005	-.000004	-.000002	.000659	-.000006	-.000007
.360	.992540	-.000204	.000014	-.000191	.000011	-.000002	.000002	.000005	.000544	-.000228	-.000003
.400	.998990	-.000044	.000012	-.000048	.000010	.000006	.000009	.000011	-.000002	-.000100	.000007
.440	.999910	-.000012	-.000003	-.000013	-.000005	-.000004	-.000003	-.000003	-.000023	-.000025	-.000004
Maximum Absolute Deviation		.002854	-.000018	.000493	.000052	.000006	.000017	.000015	.001121	.001237	.000010
Average Absolute Deviation		.000756	.000006	.000176	.000015	.000002	.000006	.000006	.000182	.000357	.000004

(4) p = .9, n = 400

X	Pr	F	R	H	GC	CF	Case A	Case B	Case C1	Case C3	Case D
.858	.000080	-.000007	-.000006	.000003	-.000004	-.000001	-.000001	-.000001	-.000001	.000003	†
.863	.000350	-.000024	-.000012	.000017	.000007	.000000	.000000	.000000	.000001	.000015	†
.868	.001370	-.000092	-.000027	.000035	.000027	.000001	.000001	.000001	.000005	.000035	†
.873	.004830	-.000320	-.000063	.000068	.000037	-.000003	-.000003	-.000004	.000004	.000055	†
.878	.015250	-.000908	-.000100	.000124	.000017	-.000004	-.000005	-.000005	.000007	.000062	†
.883	.042510	-.002167	-.000113	.000187	-.000044	.000000	.000000	.000000	.000015	.000017	†
.888	.103340	-.004302	-.000091	.000206	-.000109	-.000001	.000000	.000001	.000008	-.000108	†
.893	.216740	-.006919	-.000042	.000155	-.000069	-.000005	-.000003	-.000001	-.000011	-.000218	†
.898	.388970	-.008003	-.000017	.000043	.000073	-.000012	-.000011	-.000009	-.000036	-.000162	†
.903	.596010	-.008611	-.000012	-.000085	.000131	-.000010	-.000011	-.000012	-.000039	.000055	†
.908	.786450	-.006281	.000030	-.000196	-.000012	-.000005	-.000008	-.000010	-.000020	.000182	†
.913	.915280	-.003206	.000093	-.000241	-.000117	-.000011	-.000012	-.000014	-.000005	.000096	†
.918	.976460	-.001209	.000099	-.000181	-.000026	-.000015	-.000014	-.000015	-.000002	-.000038	†
.923	.995760	-.000293	.000050	-.000080	.000023	-.000010	-.000009	-.000009	-.000003	-.000049	†
.928	.999540	-.000051	.000005	-.000026	-.000008	-.000010	-.000010	-.000010	-.000008	-.000024	†
.933	.999970	-.000007	-.000002	-.000006	-.000007	-.000004	-.000004	-.000004	-.000004	-.000006	†
Maximum Absolute Deviation		.009000	.000113	.000241	.000141	.000015	.000015	.000015	.000039	.000218	†
Average Absolute Deviation		.002278	.000041	.000086	.000039	.000004	.000004	.000005	.000009	.000058	†

† no solution found

distributions. Besides giving very accurate approximations to every distribution considered, it provides accuracy comparable to any other approximation in the cases of the negative binomial, the gamma, the non-central t, the non-central F, and the correlation coefficient in normal samples. In fact, it yields an approximation more accurate than any other considered in the case of the non-central χ^2 distribution. In terms of accuracy therefore, the approximation is excellent and yields remarkable results.

However, in any evaluation of an approximation, the speed and simplicity of an approximation must also be taken into account. Our method involves the lenghty task of solving a cubic equation for each of Cases A and B. However, this need only be done once for a solution of the constants. After this stage, our method is extremely fast and simple. A computer program listed in Appendix B, is presented which will calculate the constants and will either find the probability, or the quantile of a random variable given the first few cumulants. The speed of the routine (without Case D) has been tested, with the following results. The subroutine which will calculate the constants for every case, and will choose the best solution (this is the one where the higher order terms in the cumulants of y approach zero quickest), takes approximately 3.25 sixtieths of a second for each solution. The function which calculates the probability level executes in approximately .04 sixtieths of a second while the one which calculates the quantile takes approximately .07 sixtieths of a second.

These execution times would compare very favorably with most of the approximations examined. A few comparisons have been made with

the Peizer-Pratt approximations, with the result that the proposed transformation executes invariably in less than one-half the time for all the distributions in the Peizer-Pratt system. With regards to the hypergeometric distribution, our approximation is by far much faster than Wise's approximation which involves several time-consuming calculations of the incomplete beta function. In the case of the non-central χ^2, a comparison of the computer time involved between our approximation and that of Pearson, is quite difficult. The whole problem centers around the calculation of a χ^2 with non integral degrees of freedom, which is difficult and takes an enormous amount of time. Even interpolation of two central χ^2 random variables is too time consuming, since the mere calculation of a central χ^2 probability needs more computer time than a complete computation of our transformation, and even then, it depends to some extent on the particular routine used. In fact most of the time in our approximation is consumed by raising the quadratic to the power of h, and calculating the normal integral, both of which are very fast. If another approximation is to be faster, it must be extremely simple, consisting of at most multiplications or divisions, and must preclude at least one of the steps above.

As mentioned before, an important feature of our approximation is that it is distribution-free. Thus in the case when the distribution of a random variable is unknown, but the first few cumulants are known, our transformation provides a very accurate and quick method for finding the probability or quantile. It can therefore be used to fit a distribution to any random variable given just the moments.

Furthermore, our transformation may be used with observational data. It may be used to give very good representations of actual data, besides being employed to give percentiles and quantiles. Basically this would involve determining the numerical values of the first five moments from the data, and hence calculating the cumulants, which are then input into the system.

Finally, since the transformation is invertible, it may be used to generate random variables of any distribution given its cumulants and any uniform or normal random variable. Several tests were made to validate the accuracy of the random variables generated. One thousand of these random variables were generated for a χ^2 with various degrees of freedom. A Kolomogorov-Smirnov test was then made, and in none of the instances did it reject the hypothesis that the random variables were distributed as χ^2. Moreover, the method is extremely fast in relation to the computer time involved.

REFERENCES AND BIBLIOGRAPHY

[1] Abdel-Aty, S.H., (1954). Approximate formulae for the percentage points and the probability integral of the non-central χ^2 distribution. *Biometrika*, 41, 538-540.

[2] Anscombe, F.J. (1948). The transformations of Poisson, binomial, and negative binomial data. *Biometrika*, 35, 246-254.

[3] Aroian, L.A. (1947). Note on the cumulants of Fisher's Z distribution. *Biometrika*, 32, 359-360.

[4] Azorin, P.F. (1953). Sobre la distribution t no central. I, II, *Trabajos de Estadistica*, 4, 137-198 and 307-337.

[5] Bartlett, M.S. (1936). The square root transformation in the analysis of variance. *Journal of the Royal Statistical Society Supplement*, 3, 68.

[6] Barton, D.E., David, F.N., and O'Neill, Anne F. (1960). Some properties of the distribution of the logarithm of non-central F. *Biometrika*, 47, 417-431.

[7] Battacharyya, A. (1945-46). On a measure of divergence between two multinomial populations. *Sankhya*, 7, 401-406.

[8] Beall, G. (1942). The transformation of data from entomological field experiments so that the analysis of variance becomes applicable. *Biometrika*, 32, 243-262.

[9] Blom, G. (1954). Transformations of the binomial, negative binomial, Poisson and the χ^2 distributions. *Biometrika*, 41, 302-316.

[10] Bol'shev, L.N., and Kuznetov, P.I. (1963). On computing the integral p(x,y) = ... *Zhurnal Vychislitenoj Matematiki i Matematicheskoi Fiziki*, 3, 419-430.

[11] Brown, K.M. (1969). A quadratically convergent Newton-like method based on Gaussian elimination. *SIAM Journal of Numerical Analysis*, 6, 560-569.

[12] Camp, B.H. (1951). Approximation to the point binomial. *Annals of Mathematical Statistics*, 22, 130-131.

[13] Carter, A.H. (1947). Approximation to percentage points of the Z distribution. *Biometrika*, 34, 352-358.

[14] Cochran, W.G. (1940). Note on an approximate formula for the significance levels of Z. *Annals of Mathematical Statistics*, 11, 93-95.

[15] Cornish, E.A. and Fisher, R.A. (1937). Moments and cumulants in the specification of distributions. *Revue de L'institut International de Statistique*, 5, 307.

[16] Cramer, H. (1963). *Mathematical Methods of Statistics*. Princeton. Princeton University Press.

[17] Elfving, G. (1955). An expansion principle for distribution functions with an application to the "Student's" statistics. *Annals Academia Scientarum Fennicae, Series A*, 204, 1-8.

[18] Feller, W. (1957). *An Introduction to Probability Theory and Its Applications*. Vol. 1. New York: John Wiley and Sons.

[19] Fisher, R.A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron*, 1, 3-32.

[20] Fisher, R.A. (1922). On the interpretation of χ^2 from contingency tables and calculation of P. *Journal of the Royal Statistical Society, Series A*, 85, 87-94.

[21] Fisher, R.A. (1925). Expansion of "Student's" integral in powers of n^{-1}. *Metron*, 5, 109-112.

[22] Freeman, M.F., and Tukey, J.W. (1950). Transformations related to the angular and square root. *Annals of Mathematical Statistics*, 21, 607-611.

[23] Fisher, R.A., and Cornish, E.A. (1960). The percentile points of distributions having known cumulants. *Technometrics*, 2, 209.

[24] Gray, H.L., Thompson, R.W., and McWilliams, G.V. (1969). A new approximation for the chi-square integral. *Mathematics of Computation*, 23, 85-89.

[25] Haldane, J.B.S. (1938). The approximate normalization of a class of frequency distributions. *Biometrika*, 29, 392-404. (Correction 31, 220).

[26] Hendricks, W.A. (1936). Approximation to "Student's" distribution. *Annals of Mathematical Statistics*, 7, 210-221.

[27] Hotelling, H. (1953). New light on the correlation coefficient and its transforms. *Journal of the Royal Statistical Society, Series B*, 15, 193-225. (Discussion 225-232)

[28] Hotelling, H., and Frankel, L.R. (1938). The transformation of statistics to simplify their distribution. *Annals of Mathematical Statistics*, 9, 87-96.

[29] International Mathematics and Statistical Libraries (1973). *IMSL Library 1, Fortran IV, S/370-360*, Second Edition. Manual LIB1-0002, IMSL, Houston, Texas.

[30] James, G.S. (1955a). On moments and cumulants of systems of statistics. (Unpublished).

[31] James, G.S. (1955b). Cumulants of a transformed variate. *Biometrika*, 42, 529-531.

[32] Jennett, W.J., and Welch, B.L. (1939). The control of proportion defective as judged by a single quality characteristic varying on a continuous scale. *Journal of the Royal Statistical Society, Series B*, 6, 80-88.

[33] Johnson, N.L. (1959). On an extension of the connexion between Poisson and χ^2 distributions. *Biometrika*, 46, 352-363.

[34] Johnson, N.L., and Kotz, S. (1969). *Discrete Distributions*. Honghton Mifflin Co., Boston.

[35] Johnson, N.L., and Kotz, S. (1970). *Continuous Univariate Distributions - 1*. Houghton Mifflin Co., Boston.

[36] Johnson, N.L., and Kotz, S. (1970). *Continuous Univariate Distributions - 2*. Houghton Mifflin Co., Boston.

[37] Kendall, M.G., and Stuart, A. (1963). *The Advanced Theory of Statistics*. Vol. 1, Charles Griffin and Company, London.

[38] Kimball, A.W., and Leach, E. (1959). Approximate linearization of the incomplete B-function. *Biometrika*, 46, 214-218.

[39] Matusita, K. (1953). On the estimation by the minimum-distance method. *Annals of the Institute of Statistical Mathematics*, 5, 59-65.

[40] Merrington, Maxine, and Pearson, E.S. (1958). An approximation to the distribution of non-central t. *Biometrika*, 45, 484-491.

[41] Moore, P.G. (1957). Transformations to normality using fractional powers of the variable. *Journal of the American Statistical Association*, 52, 237-246.

[42] Nicholson, W.L. (1956). On the normal approximation to the hypergeometric distribution. *Annals of Mathematical Statistics*, 27, 471-483.

[43] Patnaik, P.B. (1949). The non-central χ^2 and F distributions and their applications. *Biometrika*, 36, 202-232.

[44] Paulson, E. (1942). An approximate normalization of the analysis of variance distribution. *Annals of Mathematical Statistics*, 13, 233-235.

[45] Paulson, E. (1947). *Selected Techniques of Statistical Analysis*. McGraw-Hill, New York.

[46] Pearson, E.S. (1959). Note on an approximation to the distribution of non-central χ^2. *Biometrika*, 46, 364.

[47] Pearson, K. (1922). *Tables of the Incomplete Γ function*. H.M. Stationary Office, London.

[48] Peizer, D.B., and Pratt, J.W. (1968). A normal approximation for binomial, F, beta, and other common related tail probabilities, I. *Journal of the American Statistical Association*, 63, 1416-1456.

[49] Pratt, J.W. (1968). A normal approximation for binomial, F, beta, and other common related tail probabilities, II. *Journal of the American Statistical Association*, 63, 1457-1458.

[50] Ruben, H. (1963). Probability of a positive sample correlation (abstract). *Annals of Mathematical Statistics*, 34, 694.

[51] Ruben, H. (1966). Some new results on the distribution of the sample correlation coefficient. *Journal of the Royal Statistical Society, Series B*, 28, 513-525.

[52] Sankaran, M. (1959). On the non-central chi-square distribuiton. *Biometrika*, 46, 235-237.

[53] Sankaran, M. (1963). Approximations to the non-central chi-square distribution. *Biometrika*, 50, 199-204.

[54] Severo, N., and Zellen, M. (1960). Normal approximation to the chi-square and non-central F probability functions. *Biometrika*, 47, 411-416.

[55] Tiku, M.L. (1965). Laguerre series forms of non-central χ^2 and F distributions. *Biometrika*, 52, 415-427.

[56] Tiku, M.L. (1966). A note on approximating to the non-central F distributions. *Biometrika*, 53, 606-610.

[57] Wallace, D.L. (1958). Asymptotic approximations to distributions. *Annals of Mathematical Statistics*, 29, 635-654.

[58] Wallace, D.L. (1959). Bounds on normal approximations to "Student's" and the chi-square distributions. *Annals of Mathematical Statistics*, 30, 1121-1130.

[59] Wilson, E.B., and Hilferty, M.M. (1931). The distribution of chi-square. *Proceedings of the National Academy of Science*, 17, 684-688.

[60] Wishart, J. (1947). The cumulant of the Z and of the logarithmic χ^2 and t distributions. *Biometrika*, 34, 170-178.

[61] Wishart, J. (1957). An approximate formula for the cumulative Z distribuiton. *Annals of Mathematical Statistics*, 28, 504-510.

[62] Wise, M.E., (1954). A quickly convergent expansion for cumulative hypergeometric probabilities, direct and inverse. *Biometrika*, 41, 317-329.

APPENDIX A

We present the expansions of the cumulants of y up to the eight cumulant of x. These are the expressions used to solve for the constants a_1, a_2, and h. They should prove to be of interest if a grouping different from that defined by (3.1) is chosen.

Appendix A.

ON APPROXIMATE NORMALIZING TRANSFORMATIONS

$$Y = (1 + a_1(x-\mu) + a_2(x-\mu)^2)^h$$

$$\begin{aligned}
E(y) = 1 &+ \left(a_1^2 + \frac{2a_2}{(h-1)}\right)\frac{h^{(2)}}{2!}K_2 \\
&+ \left(a_1^3 + \frac{6a_1a_2}{(h-2)}\right)\frac{h^{(3)}}{3!}K_3 \\
&+ \left(a_1^4 + \frac{12a_1^2a_2}{(h-3)} + \frac{12a_2^2}{(h-3)(h-2)}\right)\frac{h^{(4)}}{4!}(K_4 + 3K_2^2) \\
&+ \left(a_1^5 + \frac{20a_1^3a_2}{(h-4)} + \frac{60a_1a_2^2}{(h-4)(h-3)}\right)\frac{h^{(5)}}{5!}(K_5 + 10K_3K_2) \\
&+ \left(a_1^6 + \frac{30a_1^4a_2}{(h-5)} + \frac{180a_1^2a_2^2}{(h-5)(h-4)} + \frac{120a_2^3}{(h-5)(h-4)(h-3)}\right)\frac{h^{(6)}}{6!}(K_6 + 15K_4K_2 + 10K_3^2 + 15K_2^3) \\
&+ \left(a_1^7 + \frac{42a_1^5a_2}{(h-6)} + \frac{420a_1^3a_2^2}{(h-6)(h-5)} + \frac{840a_1a_2^3}{(h-6)(h-5)(h-4)}\right)\frac{h^{(7)}}{7!}(K_7 + 21K_5K_2 + 35K_4K_3 + 105K_3K_2^2) \\
&+ \left(a_1^8 + \frac{56a_1^6a_2}{(h-7)} + \frac{840a_1^4a_2^2}{(h-7)(h-6)} + \frac{3360a_1^2a_2^3}{(h-7)(h-6)(h-5)} + \frac{1680a_2^4}{(h-7)\ldots(h-4)}\right)\frac{h^{(8)}}{8!}(K_8 + 28K_6K_2 + 56K_5K_3 + 35K_4^2 + 210K_4K_2^2 + 280K_3^2K_2 + 105K_2^4) \\
&+ \left(a_1^9 + \frac{72a_1^7a_2}{(h-8)} + \frac{1512a_1^5a_2^2}{(h-8)(h-7)} + \frac{10080a_1^3a_2^3}{(h-8)(h-7)(h-6)} + \frac{15120a_1a_2^4}{(h-8)\ldots(h-5)}\right)\frac{h^{(9)}}{9!}(K_9+36K_7K_2+84K_6K_3+126K_5K_4+378K_5K_2^2+1260K_4K_3K_2+280K_3^3+1260K_3K_2^3) \\
&+ \left(a_1^{10} + \frac{90a_1^8a_2}{(h-9)} + \frac{5040a_1^6a_2^2}{(h-9)(h-8)} + \frac{25200a_1^4a_2^3}{(h-9)(h-8)(h-7)} + \frac{75600a_1^2a_2^4}{(h-9)\ldots(h-6)} + \frac{30240a_2^5}{(h-9)\ldots(h-5)}\right)\frac{h^{(10)}}{10!}(K_{10}+45K_8K_2+120K_7K_3+630K_6K_2^2+126K_5^2+2520K_5K_3K_2+1575K_4^2K_2 \\
&\qquad + 210K_6K_4+2100K_4K_3^2+3150K_4K_2^3+6300K_3^2K_2^2+945K_2^5) \\
&+ \ldots
\end{aligned}$$

Note: $E(y^r)$ is the expression above with h replaced by rh.

Appendix A.

$$\sigma^2(y) = \overline{E(y^2)} - \overline{E^2(y)}$$

$$\overline{K_2(y)} = a_1^2 h^2 K_2$$

$$+K_3 h^2 a_1 (a_1^2(h-1) + 2a_2)$$

$$+K_4 \frac{h^2}{12} [a_1^4(h-1)(7h-11) + 36a_1^2 a_2 (h-1) + 12a_2^2] + K_2^2 \frac{h^2}{3} [a_1^4(h-1)(3h-5) + 16a_1^2 a_2(h-1) + 4a_2^2]$$

$$+K_5 \frac{h^2(h-1)}{12} a_1[a_1^4(h-2)(3h-5) + 4a_1^2 a_2 (7h-11) + 36a_2^2] + K_3 K_2 \frac{h^2(h-1)}{3} a_1[a_1^4(h-2)(7h-12) + 3a_1^2 a_2(22h-35) + 84a_2^2]$$

$$+K_6 \frac{h^2(h-1)}{360} [a_1^6(h-2)(21h^2-132h+137) + 150a_1^4 a_2 (h-2)(3h-5) + 180a_1^2 a_2^2 (7h-11) + 360a_2^3]$$

$$+K_4 K_2 \frac{h^2(h-1)}{12} [a_1^6(h-2)(15h^2-64h+67) + 2a_1^4 a_2(h-2)(109h-183) + 12a_1^2 a_2^2(5h-80) + 168a_2^3]$$

$$+K_3^2 \frac{h^2(h-1)}{36} [a_1^6(h-2)(30h^2-129h+135) + 6a_1^4 a_2(h-2)(73h-123) + 72a_1^2 a_2^2(17h-27) + 360a_2^3]$$

$$+K_2^3 \frac{h^2(h-1)}{6} [a_1^6(h-2)(7h^2-30h+32) + 6a_1^4 a_2(h-2)(17h-29) + 18a_1^2 a_2^2 (16h-25) + 72a_2^3]$$

$$+K_7 \frac{h^2(h-1)}{2520} a_1[a_1^6(h-2)(h-3)(63h^2-273h+294) + 42a_1^4 a_2(h-2)(31h^2-132h+137) + 2100a_1^2 a_2(h-2)(3h-5) + 840a_2^3(7h-11)]$$

$$+K_5 K_2 \frac{h^2(h-1)}{60} a_1[a_1^6(h-2)(h-3)(31h^2-134h+145) + 5a_1^4 a_2(h-2)(128h^2-545h+567) + 20a_1^2 a_2^2 (h-2)(155h-258) + 180a_2^3(16h-25)]$$

$$+K_4 K_3 \frac{h^2(h-1)}{36} a_1[a_1^6(h-2)(h-3)(31h^2-135h+146) + 6a_1^4 a_2 (h-2)(107h^2-457h+476) + 84a_1^2 a_2^2 (h-2)(37h-62) + 24a_2^3(121h-191)]$$

$$+K_3 K_2^2 \frac{h^2(h-1)}{6} a_1[a_1^6(h-2)(h-3)(15h^2-65h+71) + 2a_1^4 a_2(h-2)(155h^2-662h+693) + 2a_1^2 a_2^2(h-2)(752h-1257) + 12a_2^3(116h-181)]$$

$$+K_8 \frac{h^2(h-1)}{20160} [a_1^8(h-2)(h-3)(127h^3-1002h^2+2573h-2178) + 56a_1^6 a_2(h-2)(h-3)(63h^2-273h+294) + 840a_1^4 a_2^2(h-2)(31h^2-132h+137) + 8800a_1^2 a_2^3(h-2)(3h-5) + 1680a_2^4(7h-11)]$$

$$+K_6 K_2 \frac{h^2(h-1)}{360} [a_1^8(h-2)(h-3)(63h^3-498h^2+1272h-1379) + 4a_1^6 a_2(h-2)(h-3)(437h^2-1890h+2038) + 30a_1^4 a_2^2(h-2)(430h^2-1829h+1897) + 120a_1^2 a_2^3(h-2)(208h-345) + 120a_2^4(46h-75)]$$

$$+K_5 K_3 \frac{h^2(h-1)}{360} [a_1^8(h-2)(h-3)(126h^3-995h^2+2559h-2170) + (2a_1^6 a_2 (h-2)(h-3)(1751h^2-7599h+8200) + 60a_1^4 a_2^2(h-2)(431h^2-1837h+1917) + 120a_1^2 a_2^3(h-2)(417h-697) + 1680a_2^4(7h-11)$$

$$+K_4^2 \frac{h^2(h-1)}{288} [a_1^8(h-2)(h-3)(63h^3-498h^2+1281h-1086) + 8a_1^6 a_2 (h-2)(h-3)(270h^2-951+1227) + 24a_1^4 a_2^2(h-2)(539h^2-2291+2390) + 96a_1^2 a_2^3(h-2)(263h-436) + 48a_2^4(172h-191)]$$

$$+K_4 K_2^2 \frac{h^2(h-1)}{24} [a_1^8(h-2)(h-3)(31h^2-244+626h-533) + 4a_1^6 a_2(h-2)(h-3)(215h^2-930h+1206) + 6a_1^4 a_2^2(h-2)(1058h^2-4503h+4675) + 24a_1^2 a_2^3(h-2)(512h-847) + 24a_2^4(116h-181)]$$

$$+K_3^2 K_2 \frac{h^2(h-1)}{36} [a_1^8(h-2)(h-3)(62h^3-489h^2+1258h-1071) + 2a_1^6 a_2(h-2)(h-3)(861h^2-3785h+4044) + 30a_1^4 a_2^2(h-2)(424h^2-1807h+1881) + 120a_1^2 a_2^3(h-2)(205h-342) + 120a_2^4(48h-75)]$$

$$+K_2^4 \frac{h^2(h-1)}{24} [a_1^8(h-2)(h-3)(15h^3-118h^2+303h-260) + 8a_1^6 a_2(h-2)(h-3)(52h^2-225h+245) + 12a_1^4 a_2^2(h-2)(256h^2-1091h+1135) + 192a_1^2 a_2^3(h-2)(31h-51) + 48a_2^4(27h-42)]$$

$$+ \ldots$$

Appendix A.

$$K_3(y) = u_3(y) + u_3'(y) - u_1'(y)\,[u_2'(y) + 2K_2(y)]$$

$$= a_1^3 h^3 K_3$$

$$\frac{3}{2}(K_4 + 2K_2^2)\, h^3\, (a_1^4(h-1) + 2a_1^2 a_2)$$

$$+K_5 \frac{h^3}{4}[a_1^5(h-1)(5h-7) + 24a_1^3 a_2(h-1) + 12a_1 a_2^2]$$

$$+K_8 K_2 \frac{3h^3}{2}[a_1^5(h-1)(7h-10) + 34a_1^3 a_2(h-1) + 16a_1 a_2^2]$$

$$+K_6 \frac{h^3}{8}[a_1^6(h-1)(6h^2-19h+15) + 10a_1^4 a_2(h-1)(5h-7) + 72a_1^2 a_2^2(h-1) + 8a_2^3]$$

$$+K_4 K_2 \frac{h^3}{4}[a_1^6(h-1)(41h^2-131h+104) + 4a_1^4 a_2(h-1)(86h-121) + 492a_1^2 a_2^2(h-1) + 48a_2^3]$$

$$+K_3^2 \frac{h^3}{4}[a_1^6(h-1)(28h^2-89h+71) + 6a_1^4 a_2(h-1)(39h-55) + 336a_1^2 a_2^2(h-1) + 40a_2^3]$$

$$+K_2^3 \frac{h^3}{2}[a_1^6(h-1)(17h^2-55h+44) + 12a_1^4 a_2(h-1)(12h-17) + 204a_1^2 a_2^2(h-1) + 16a_2^3]$$

$$+K_7 \frac{h^3(h-1)}{120}[a_1^7(h-2)(43h^2-141h+116) + 90a_1^5 a_2(6h^2-19h+15) + 300a_1^3 a_2^2(5h-7) + 720a_1 a_2^3]$$

$$+K_5 K_2 \frac{h^3(h-1)}{8}[a_1^7(h-2)(57h^2-188h+155) + 2a_1^5 a_2(359h^2-1140h+901) + 4a_1^3 a_2^2(499h-698) + 936a_1 a_2^3]$$

$$+K_4 K_3 \frac{h^3(h-1)}{8}[a_1^7(h-2)(97h^2-319h+264) + 4a_1^5 a_2(305h^2-968h+767) + 4a_1^3 a_2^2(847h-1190) + 1632a_1 a_2^3]$$

$$+K_3 K_2^2 \frac{3h^3(h-1)}{8}[a_1^7(h-2)(87h^2-289h+240) + 4a_1^5 a_2(275h^2-877h+696) + 4a_1^3 a_2^2(765h-1072) + 1408a_1 a_2^3]$$

$$+K_8 \frac{h^3(h-1)}{480}[a_1^8(h-2)(69h^3-395h^2+747h-469) + 28a_1^6 a_2(h-2)(43h^2-141h+116) + 900a_1^4 a_2^2(6h^2-19h+15) + 1200a_1^2 a_2^3(5h-7) + 720a_2^4]$$

$$+K_2^4 \frac{h^3(h-1)}{8}[a_1^8(h-2)(101h^3-582h^2+1111h-702) + 6a_1^6 a_2(h-2)(295h^2-977h+808) + 12a_1^4 a_2^2(665h^2-2109h+1666) + 72a_1^2 a_2^3(121h-168) + 864a_2^4]$$

Appendix A.

$$K_4(y) = \mu_4'(y) - \mu_1'(y)\mu_3'(y) - 3K_2(y)\mu_2'(y) - 3K_3(y)\mu_1'(y)$$

$$= a_1^4 h^4 K_4$$

$$+ 2(K_5 + 6K_3K_2)h^4(a_1^5(h-1) + 2a_1^3a_2)$$

$$+ (K_6 + 12K_4K_2)\frac{h^4}{6}[a_1^6(h-1)(13h-17) + 60a_1^4a_2(h-1) + 36a_1^2a_2^2]$$

$$+ 6K_3^2 h^4[a_1^6(h-1)(3h-4) + 14a_1^4a_2(h-1) + 8a_1^2a_2^2]$$

$$+ 4K_2^3 h^4[a_1^6(h-1)(4h-5) + 18a_1^4a_2(h-1) + 12a_1^2a_2^2]$$

$$+K_7\frac{h^4}{6}[a_1^7(h-1)(10h^2 - 29h + 21) + 6a_1^5a_2(h-1)(13h-17) + 120a_1^3a_2^2(h-1) + 24a_1a_2^3]$$

$$+ 6K_3K_2^2h^4[a_1^7(h-1)(21h^2-61h+44) + 2a_1^5a_2(h-1)(82h-107) + 252a_1^3a_2^2(h-1) + 48a_1a_2^3]$$

$$+K_8\frac{h^4}{240}[a_1^8(h-1)(243h^3-1157h^2+1833h-967) + 280a_1^6a_2(h-1)(10h^2-29h+21) + 600a_1^4a_2(h-1)(13h-17) + 4800a_1^2a_2^3(h-1) + 240a_2^4]$$

$$+K_2^4h^4[a_1^8(h-1)(71h^3-339h^2+535h-279) + 4a_1^6a_2(h-1)(205h^2-593h+426) + 24a_1^4a_2^2(h-1)(95h-124) + 1392a_1^2a_2^3(h-1) + 48a_2^4]$$

Appendix A.

$$K_5(y) = \mu_5'(y) - \mu_1'(y)\mu_4'(y) - 4K_2(y)\mu_3'(y) - 6K_3(y)\mu_2'(y) - 4K_4(y)\mu_1'(y)$$

$$= a_1^5 h^5 K_5$$

$$+ 5(3K_6 + 3K_3^2 + 4K_4K_2)h^5(a_1^6(h-1) + 2a_1^4a_2)$$

$$+ 5K_7 \frac{h^5}{6} [a_1^7(h-1)(4h-5) + 18a_1^5a_2(h-1) + 12a_1^3a_2^2]$$

$$+ 30K_3K_2^2h^5 [a_1^7(h-1)(5h-6) + 22a_1^5a_2(h-1) + 16a_1^3a_2^2]$$

$$+ 5K_8 \frac{h^5}{24} [a_1^8(h-1)(15h^2-41h+28) + 28a_1^6a_2(h-1)(4h-5) + 180a_1^4a_2^2(h-1) + 48a_1^2a_2^3]$$

$$+ 5K_2^4h^5[a_1^8(h-1)(25h^2-65h+42) + 36a_1^6a_2(h-1)(5h-6) + 300a_1^4a_2^2(h-1) + 96a_1^2a_2^3]$$

APPENDIX B

We present a listing of a subroutine and several functions developed for the application of our transformation. The subroutine YTRANS will solve for the constants in our transformation, while the functions PROB, and CENT, or CENTY will calculate the probability or the quantile. A number of comment cards have been added to the program to make their usage and purpose self-explanatory.

After the programs, a sample output is given of the subroutine YTRANS, which illustrates the effect of our transformation on the cumulants of y. The example is chosen from the non-central χ^2 distribution with parameters $n = 10$, $\lambda = 5.0$ (see also table 5.13 for numerical comparisons).

```
C                                                                       YTRA0010
C                                                                       YTRA0020
      SUBROUTINE YTRANS(XK,NCUM,A1,A2,H,EXY,SDY,IER)                    YTRA0030
C                                                                       YTRA0040
C        Y = ( 1 + A1 * ( X - U ) + A2 * ( X - U ) ** 2 ** H            YTRA0050
C                                                                       YTRA0060
C  PURPOSE:                                                             YTRA0070
C        TO CALCULATE THE CONSTANTS A1,A2,H IN THE EQUATION ABOVE       YTRA0080
C        TOGETHER WITH THE MEAN EXY AND STANDARD DEVIATION SDY OF Y     YTRA0090
C        GIVEN THE CUMULANTS OF X SO THAT  THEY CAN BE USED IN THE      YTRA0100
C        FOLLOWING FUNCTIONS:                                           YTRA0110
C             1. CENT(P)                                                YTRA0120
C                  - CALCULATES THE PERCENTAGE POINTS OF X GIVEN        YTRA0130
C                    THE PROBABILITY P                                  YTRA0140
C                CENTY(Y)                                               YTRA0150
C                  - CALCULATES THE PERCENTAGE POINTS OF X GIVEN        YTRA0160
C                    THE CORRESPONDING PERCENTAGE POINT OF A NORMAL     YTRA0170
C                    RANDOM VARIABLE                                    YTRA0180
C             2. PROB(X)                                                YTRA0190
C                  - CALCULATES THE PROBABILITY  OF THE UNKNOWN         YTRA0200
C                    DISTRIBUTION AT POINT X                            YTRA0210
C  PARAMETERS:                                                          YTRA0220
C        XK   - CUMULANTS OF X (DIMENSIONED ARRAY)                      YTRA0230
C        NCUM - NO OF CUMULANTS OF X  (MUST BE GREATER THAN 3)          YTRA0240
C        A1   - SOLUTION CONSTANT                                       YTRA0250
C        A2   - SOLUTION CONSTANT                                       YTRA0260
C        H    - SOLUTION CONSTANT                                       YTRA0270
C        EXY  - SOLUTION CONSTANT - EXPECTED VALUE OF Y                 YTRA0280
C        SDY  - SOLUTION CONSTANT - STANDARD DEVIATION OF Y             YTRA0290
C        IER  - INDICATOR                                               YTRA0300
C             INPUT VALUE:                                              YTRA0310
C                  0. - PRINTS NOTHING                                  YTRA0320
C                  1. - PRINTS A COMPLETE ANALYSIS OF THE DIFFERENT     YTRA0330
C                       SOLUTIONS FOR EACH OF THE FOUR CASES            YTRA0340
C             OUTPUT VALUE:                                             YTRA0350
C                  9. - UNABLE TO FIND SOLUTION                         YTRA0360
C                                                                       YTRA0370
C                                                                       YTRA0380
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA0390
      DIMENSION  XK(1),AA1(3),AH(3),AA2(3),BA1(3),BH(3),BA2(3),C2A1(2), YTRA0400
     & EAA(3),FAA(3),GAA(3),BBB(3),CBB(3),DBB(3),EBB(3),FBB(3),GBB(3),  YTRA0410
     & BC2(2),CC2(2),DC2(2),EC2(2),FC2(2),GC2(2),C2H(2),BAA(3),CAA(3)   YTRA0420
      DIMENSION YMA(3),YMB(3),YMC2(2),YSDA(3),YSDB(3),YSDC2(2),DAA(3),  YTRA0430
     & YM2A(3),YM3A(3),YM4A(3),YM1B(3),YM2B(3),YM3B(3),YM4B(3),YV1A(3), YTRA0440
     & YV2A(3),YV3A(3),YV4A(3),YV1B(3),YV2B(3),YV3B(3),YV4B(3),YM1C2(2),YTRA0450
     & YM2C2(2),YM3C2(2),YM4C2(2),YV1C2(2),YV2C2(2),YV3C2(2),YV4C2(2)   YTRA0460
      DIMENSION YM1A(3),XF(3,5),WA(3,10),IERF(5), XKK(5),               YTRA0470
     & YM1F(5),YM2F(5),YM3F(5),YM4F(5),YV1F(5),YV2F(5),YV3F(5),YV4F(5), YTRA0480
     & YMF(5),YSDF(5),FB(5),FC(5),FD(5),FE(5),FF(5),FG(5)               YTRA0490
      COMMON /COM/CA1,CA2,CH,CEXY,CSDY,SG,U                             YTRA0500
      COMMON XKK                                                        YTRA0510
      EXTERNAL FFF                                                      YTRA0520
C                                                                       YTRA0530
C       INITIALIZE                                                      YTRA0540
C                                                                       YTRA0550
      MCUM=MIN0(NCUM,5)                                                 YTRA0560
      DO 133 I=1,MCUM                                                   YTRA0570
  133 XKK(I)=XK(I)                                                      YTRA0580
```

```
      A1=0.D0                                                           YTRA0590
      A2=0.D0                                                           YTRA0600
      H=0.D0                                                            YTRA0610
      REF=1.D+70                                                        YTRA0620
      NO = 0                                                            YTRA0630
      IF(DABS(XK(3)).GT.1.D-25 ) NO=1                                   YTRA0640
      IF(NCUM.GE.5) GO TO 1                                             YTRA0650
      NA=0                                                              YTRA0660
      IF(NCUM.EQ.4) GO TO 13                                            YTRA0670
      NB=0                                                              YTRA0680
      IF(NCUM.EQ.3) GO TO 18                                            YTRA0690
      IF(IER.NE.0) WRITE(6,150)                                         YTRA0700
  150 FORMAT('0       ERROR ** INSUFFICIENT NUMBER OF CUMULANTS')       YTRA0710
      IER=9                                                             YTRA0720
      RETURN                                                            YTRA0730
C                                                                       YTRA0740
C         CASE  A                                                       YTRA0750
C                                                                       YTRA0760
    1 IF(DABS(XK(5)).LE.1.D-25) GO TO 11                                YTRA0770
      CONST1=-    XK(3)**3/( 3.D0*XK(2)**6)+XK(4)*XK(3)/(8.D0*XK(2)**5)  YTRA0780
      CONST2= 29.D0*XK(3)**4/( 12.D0*XK(2)**8)-  47.D0*XK(4)*XK(3)*XK(3)YTRA0790
     &/(  24.D0*XK(2)**7)+   5.D0*XK(4)**2/(  16.D0*XK(2)**6)+XK(5)*     YTRA0800
     & XK(3)/(20.D0*XK(2)**6)                                           YTRA0810
      CONST3=- 311.D0 XK(3)**5/( 54.D0*XK(2)**10)+ 995.D0*XK(4)*XK(3)**3YTRA0820
     &/(144.D0*XK(2)**9)-121.D0*XK(4)*XK(4)*XK(3)/( 64.D0*XK(2)**8)-8.D0YTRA0830
     &*XK(5)*XK(3)*XK(3)/(15.D0*XK(2)**8)+9.D0*XK(5)*XK(4)/(40.D0*XK(2)*YTRA0840
     &*7)                                                               YTRA0850
      CONST4= 1447.D0 XK(3)**6/( 324.D0*XK(2)**12)-3025.D0*XK(4)*XK(3)**YTRA0860
     &4/( 432.D0*XK(2)**11)+397.D0*XK(4) XK(4)*XK(3)*XK(3)/(144.D0*XK(2)YTRA0870
     &* 10)+151.D0 XK(5)*XK(3)**3/(180.D0*XK(2)**10)-XK(4)**3/(128.D0*XKYTRA0880
     &(2)**9)-79.D0*XK(5)*XK(4)*XK(3)/(120.D0*XK(2)**9)+XK(5)*XK(5)/(25.YTRA0890
     &D0 XK(2)**8)                                                      YTRA0900
      IF(CONST1.NE.0)GO TO 10                                           YTRA0910
      CONST5=CONST3*CONST3-4.D0*CONST2*CONST4                           YTRA0920
      IF(CONST5)11,40,12                                                YTRA0930
   11 NA=0                                                              YTRA0940
      GO TO 13                                                          YTRA0950
   12 AA1(1)=(-CONST3+DSQRT(CONST5))/(2.D0*CONST2)                      YTRA0960
      AA1(2)=(-CONST3-DSQRT(CONST5))/(2.D0*CONST2)                      YTRA0970
      NA=2                                                              YTRA0980
      GO TO 14                                                          YTRA0990
   40 AA1(1)=-CONST3/(2.D0*CONST2)                                      YTRA1000
      NA=1                                                              YTRA1010
      GO TO 14                                                          YTRA1020
   10 CONST2=CONST2/CONST1                                              YTRA1030
      CONST3=CONST3/CONST1                                              YTRA1040
      CONST4=CONST4/CONST1                                              YTRA1050
      CALL  DCUBIC(CONST2,CONST3,CONST4,AA1(1),AA1(2),AA1(3),ADIS)      YTRA1060
      NA=3                                                              YTRA1070
      IF(ADIS.GT.0.D0)NA=1                                              YTRA1080
   14 CONTINUE                                                          YTRA1090
      DO 7 I=1,NA                                                       YTRA1100
      CONST5=(XK(2)*XK(2)*AA1(I)+XK(3))**2-.6666666666666666D0*(8.D0*XK(3YTRA1110
     &)*XK(3)-3.D0*XK(4)*XK(2))                                         YTRA1120
      IF(CONST5.LT.0.D0) GO TO 11                                       YTRA1130
      CONST5=DSQRT(CONST5)/(XK(2)*XK(2)*4.D0*AA1(I))                    YTRA1140
      AH(I)     =1.D0-(AA1(I)+XK(3)/(XK(2)*XK(2)))/(4.D0*AA1(I))+CONST5 YTRA1150
      AA2(I)=-.5D0*AA1(I)*(AA1(I)*(AH(I)-1.D0)+XK(3)/(3.D0*XK(2)*XK(2)))YTRA1160
```

```
      EAA(I)=E(AA1(I),AA2(I),AH(I))                                     YTRA1170
      IF(DABS(EAA(I)).LE.1.D-10) GO TO 2                                YTRA1180
      AH(I)     =1.D0-(AA1(I)+XK(3)/(XK(2)*XK(2)))/(4.D0*AA1(I))-CONST5 YTRA1190
      AA2(I)=-.5D0*AA1(I)*(AA1(I)*(AH(I)-1.D0)+XK(3)/(3.D0*XK(2)*XK(2)))YTRA1200
      IF(IER.EQ.0)GO TO 2                                               YTRA1210
      EAA(I)=E(AA1(I),AA2(I),AH(I))                                     YTRA1220
    2 YM1A(I)=YM1(AA1(I),AA2(I),AH(I))                                  YTRA1230
      YM2A(I)=YM2(AA1(I),AA2(I),AH(I))                                  YTRA1240
      YV1A(I)=YV1(AA1(I),AA2(I),AH(I))                                  YTRA1250
      YV2A(I)=YV2(AA1(I),AA2(I),AH(I))                                  YTRA1260
      YMA(I)=1.D0+YM1A(I)+YM2A(I)                                       YTRA1270
      YSDD=YV1A(I)+YV2A(I)                                              YTRA1280
      IF(YSDD.LE.0.D0) YSDD=AA1(I)*AA1(I)*AH(I)*AH(I)*XK(2)             YTRA1290
      YSDA(I)=DSQRT(YSDD)                                               YTRA1300
      DES=DABS(YM2A(I))+DABS(YV2A(I))                                   YTRA1310
      IF(REF.LE.DES) GO TO 3                                            YTRA1320
      REF=DES                                                           YTRA1330
      A1=AA1(I)                                                         YTRA1340
      A2=AA2(I)                                                         YTRA1350
      H=AH(I)                                                           YTRA1360
      EXY=YMA(I)                                                        YTRA1370
      SDY=YSDA(I)                                                       YTRA1380
    3 IF(IER.EQ.0) GO TO 7                                              YTRA1390
      YM3A(I)=YM3(AA1(I),AA2(I),AH(I))                                  YTRA1400
      YM4A(I)=YM4(AA1(I),AA2(I),AH(I))                                  YTRA1410
      YV3A(I)=YV3(AA1(I),AA2(I),AH(I))                                  YTRA1420
      YV4A(I)=YV4(AA1(I),AA2(I),AH(I))                                  YTRA1430
      BAA(I)=B(AA1(I),AA2(I),AH(I))                                     YTRA1440
      CAA(I)=C(AA1(I),AA2(I),AH(I))                                     YTRA1450
      DAA(I)=D(AA1(I),AA2(I),AH(I))                                     YTRA1460
      FAA(I)=F(AA1(I),AA2(I),AH(I))                                     YTRA1470
      GAA(I)=G(AA1(I),AA2(I),AH(I))                                     YTRA1480
    7 CONTINUE                                                          YTRA1490
C                                                                       YTRA1500
C             CASE  B                                                   YTRA1510
C                                                                       YTRA1520
   13 CONST1=-3.D0*XK(3)**3/XK(2)**6+ 9.D0*XK(4)*XK(3)/( 8.D0*XK(2)**5) YTRA1530
      CONST2=175.D0*XK(3)**4/( 9.*XK(2)**8)-243.D0*XK(3)*XK(3)*XK(4)/   YTRA1540
     &(16.D0 XK(2)**7)+45.D0* XK(4)*XK(4)/(16.D0*XK(2)**6)              YTRA1550
      CONST3=-1471.D0*XK(3)  5/(54.D0*XK(2)**10)+413.D0*XK(3)**3*XK(4)/(YTRA1560
     &16.D0*XK(2) **9)-387.D0*XK(4)*XK(4)*XK(3)/(64.D0*XK(2)**8)        YTRA1570
      CONST4=15983.D0 XK(3)* 6/(1458.D0*XK(2)**12)- 47.D0*XK(3)**4*XK(4)YTRA1580
     &/(4.D0*XK(2) **7)+105.D0*XK(3)*XK(4)*XK(4)/(32.D0*XK(2)**10)YTRA1590
     & -9.D0*XK(4)**3/(128.D0*XK(2)**9)                                 YTRA1600
      IF(CONST1.NE.0.D0) GO TO 15                                       YTRA1610
  100 CONST5=CONST3*CONST3-4.D0*CONST2*CONST4                           YTRA1620
      IF(CONST5)16,41,17                                                YTRA1630
   16 NB=0                                                              YTRA1640
      GO TO 18                                                          YTRA1650
   17 BA1(1)=(-CONST3+DSQRT(CONST5))/(2.D0*CONST2)                      YTRA1660
      BA1(2)=(-CONST3-DSQRT(CONST5))/(2.D0*CONST2)                      YTRA1670
      NB=2                                                              YTRA1680
      GO TO 19                                                          YTRA1690
   41 BA1(1)=-CONST3/(2.D0*CONST2)                                      YTRA1700
      NB=1                                                              YTRA1710
      GO TO 19                                                          YTRA1720
   15 CONST2=CONST2/CONST1                                              YTRA1730
      CONST3=CONST3/CONST1                                              YTRA1740
```

```
      CONST4=CONST4/CONST1                                              YTRA1750
      CALL  DCUBIC(CONST2,CONST3,CONST4,BA1(1),BA1(2),BA1(3),BDIS)      YTRA1760
      NB=3                                                              YTRA1770
      IF(BDIS.GT.0.D0)NB=1                                              YTRA1780
   19 CONTINUE                                                          YTRA1790
      DO 8 I=1,NB                                                       YTRA1800
      CONST5=(XK(2)*XK(2)*BA1(I)+XK(3))**2-.66666666666666D0*(6.D0*XK(3YTRA1810
     &) XK(3)-3.D0*XK(4)*XK(2))                                         YTRA1820
      IF(CONST5.LT.0.D0) GO TO 16                                       YTRA1830
      CONST5=DSQRT(CONST5)/(XK(2)*XK(2)*4.D0*BA1(I))                    YTRA1840
      BH(I)     =1.D0-(BA1(I)+XK(3)/(XK(2)*XK(2)))/(4.D0*BA1(I))+CONST5 YTRA1850
      BA2(I)=-.5D0*BA1(I)+(BA1(I)*(BH(I)-1.D0)+XK(3)/(3.D0*XK(2)*XK(2)))YTRA1860
      DDD(I)=D(BA1(I),BA2(I),BH(I))                                     YTRA1870
      IF(DABS(DDD(I)).LE.1.D-10) GO TO 4                                YTRA1880
      BH(I)     =1.D0-(BA1(I)+XK(3)/(XK(2)*XK(2)))/(4.D0*BA1(I))-CONST5 YTRA1890
      BA2(I)=-.5D0*BA1(I)+(BA1(I)*(BH(I)-1.D0)+XK(3)/(3.D0*XK(2)*XK(2)))YTRA1900
      IF(IER.EQ.0) GO TO 4                                              YTRA1910
      DDD(I)=D(BA1(I),BA2(I),BH(I))                                     YTRA1920
    4 YM1B(I)=YM1(BA1(I),BA2(I),BH(I))                                  YTRA1930
      YM2B(I)=YM2(BA1(I),BA2(I),BH(I))                                  YTRA1940
      YV1B(I)=YV1(BA1(I),BA2(I),BH(I))                                  YTRA1950
      YV2B(I)=YV2(BA1(I),BA2(I),BH(I))                                  YTRA1960
      YMB(I)=1.D0+YM1B(I)+YM2B(I)                                       YTRA1970
      YSDD=YV1B(I)+YV2B(I)                                              YTRA1980
      IF(YSDD.LE.0.D0) YSDD=BA1(I)*BA1(I)*BH(I)*BH(I)*XK(2)             YTRA1990
      YSDB(I)=DSQRT(YSDD)                                               YTRA2000
      DES=DABS(YM2B(I))+DABS(YV2B(I))                                   YTRA2010
      IF(REF.LE.DES) GO TO 5                                            YTRA2020
      REF=DES                                                           YTRA2030
      A1=BA1(I)                                                         YTRA2040
      A2=BA2(I)                                                         YTRA2050
      H=BH(I)                                                           YTRA2060
      EXY=YMB(I)                                                        YTRA2070
      SDY=YSDB(I)                                                       YTRA2080
    5 IF(IER.EQ.0) GO TO 8                                              YTRA2090
      YMDB(I)=YM3(BA1(I),BA2(I),BH(I))                                  YTRA2100
      YM4B(I)=YM4(BA1(I),BA2(I),BH(I))                                  YTRA2110
      YV3B(I)=YV3(BA1(I),BA2(I),BH(I))                                  YTRA2120
      YV4B(I)=YV4(BA1(I),BA2(I),BH(I))                                  YTRA2130
      BBB(I)=B(BA1(I),BA2(I),BH(I))                                     YTRA2140
      CBB(I)=C(BA1(I),BA2(I),BH(I))                                     YTRA2150
      EBB(I)=E(BA1(I),BA2(I),BH(I))                                     YTRA2160
      FBB(I)=F(BA1(I),BA2(I),BH(I))                                     YTRA2170
      GBB(I)=G(BA1(I),BA2(I),BH(I))                                     YTRA2180
    8 CONTINUE                                                          YTRA2190
C                                                                       YTRA2200
C             CASE  C                                                   YTRA2210
C                                                                       YTRA2220
   18 NC2=-2                                                            YTRA2230
      IF(ND.EQ.0) GO TO 69                                              YTRA2240
      C1A1=5.D0*XK(3)/(3.D0*XK(2)*XK(2))-3.D0*XK(4)/(4.D0*XK(3)*XK(2))  YTRA2250
      C1H=1.D0-XK(3)/(3.D0*XK(2)*XK(2)*C1A1)                            YTRA2260
      YM1C1=YM1(C1A1,0.D0,C1H)                                          YTRA2270
      YM2C1=YM2(C1A1,0.D0,C1H)                                          YTRA2280
      YV1C1=YV1(C1A1,0.D0,C1H)                                          YTRA2290
      YV2C1=YV2(C1A1,0.D0,C1H)                                          YTRA2300
      YMC1=1.D0+YM1C1+YM2C1                                             YTRA2310
      YSDD=YV1C1+YV2C1                                                  YTRA2320
```

```
      IF(YSDD.LE.0.D0) YSDD=YV1C1                                       YTRA2330
      YSDC1=DSQRT(YSDD)                                                 YTRA2340
      DES=DABS(YM2C1)+DABS(YV2C1)                                       YTRA2350
      IF(REF.LE.DES) GO TO 32                                           YTRA2360
      REF=DES                                                           YTRA2370
      A1=C1A1                                                           YTRA2380
      A2=0.D0                                                           YTRA2390
      H=C1H                                                             YTRA2400
      EXY=YMC1                                                          YTRA2410
      SDY=YSDC1                                                         YTRA2420
   32 IF(IER.EQ.0) GO TO 67                                             YTRA2430
      YM3C1=YM3(C1A1,0.D0,C1H)                                          YTRA2440
      YM4C1=YM4(C1A1,0.D0,C1H)                                          YTRA2450
      YV3C1=YV3(C1A1,0.D0,C1H)                                          YTRA2460
      YV4C1=YV4(C1A1,0.D0,C1H)                                          YTRA2470
      BC1=B(C1A1,0.D0,C1H)                                              YTRA2480
      CC1=C(C1A1,0.D0,C1H)                                              YTRA2490
      DC1=D(C1A1,0.D0,C1H)                                              YTRA2500
      EC1=E(C1A1,0.D0,C1H)                                              YTRA2510
      FC1=F(C1A1,0.D0,C1H)                                              YTRA2520
      GC1=G(C1A1,0.D0,C1H)                                              YTRA2530
   67 CONSTS=     96.D0*XK(3)*XK(3)/(27.D0*XK(2)**4)-2.D0*XK(4)/XK(2)**3YTRA2540
      IF(CONSTS)37,38,24                                                YTRA2550
   37 NC2=0                                                             YTRA2560
      GO TO 25                                                          YTRA2570
   38 C2A1(1)=XK(3)/(6.D0*XK(2)*XK(2))                                  YTRA2580
      NC2=1                                                             YTRA2590
      GO TO 53                                                          YTRA2600
   24 CONSTS=DSQRT(CONSTS)*.5D0                                         YTRA2610
      NC2=2                                                             YTRA2620
      C2A1(1)= XK(3)/(6.D0*XK(2)*XK(2))+CONSTS                          YTRA2630
      C2A1(2)= XK(3)/(6.D0*XK(2)*XK(2))-CONSTS                          YTRA2640
   53 CONTINUE                                                          YTRA2650
      DO 9 I=1,NC2                                                      YTRA2660
    9 C2H(I)=1.D0-XK(3)/(3.D0*XK(2)*XK(2)*C2A1(I))                      YTRA2670
      DO 9 I=1,NC2                                                      YTRA2680
      YM1C2(I)=YM1(C2A1(I),0.D0,C2H(I))                                 YTRA2690
      YM2C2(I)=YM2(C2A1(I),0.D0,C2H(I))                                 YTRA2700
      YV1C2(I)=YV1(C2A1(I),0.D0,C2H(I))                                 YTRA2710
      YV2C2(I)=YV2(C2A1(I),0.D0,C2H(I))                                 YTRA2720
      YMC2(I)=1.D0+YM1C2(I)+YM2C2(I)                                    YTRA2730
      YSDD=YV1C2(I)+YV2C2(I)                                            YTRA2740
      IF(YSDD.LE.0.D0) YSDD=YV1C2(I)                                    YTRA2750
      YSDC2(I)=DSQRT(YSDD)                                              YTRA2760
      DES=DABS(YM2C2(I))+DABS(YV2C2(I))                                 YTRA2770
      IF(REF.LE.DES) GO TO 54                                           YTRA2780
      REF=DES                                                           YTRA2790
      A1=C2A1(I)                                                        YTRA2800
      A2=0.D0                                                           YTRA2810
      H=C2H(I)                                                          YTRA2820
      EXY=YMC2(I)                                                       YTRA2830
      SDY=YSDC2(I)                                                      YTRA2840
   54 IF(IER.EQ.0) GO TO 9                                              YTRA2850
      YM3C2(I)=YM3(C2A1(I),0.D0,C2H(I))                                 YTRA2860
      YM4C2(I)=YM4(C2A1(I),0.D0,C2H(I))                                 YTRA2870
      YV3C2(I)=YV3(C2A1(I),0.D0,C2H(I))                                 YTRA2880
      YV4C2(I)=YV4(C2A1(I),0.D0,C2H(I))                                 YTRA2890
      BC2(I)=B(C2A1(I),0.D0,C2H(I))                                     YTRA2900
```

```
      CC2(I)=C(C2A1(I),0.D0,C2H(I))                                     YTRA2910
      DC2(I)=D(C2A1(I),0.D0,C2H(I))                                     YTRA2920
      EC2(I)=E(C2A1(I),0.D0,C2H(I))                                     YTRA2930
      FC2(I)=F(C2A1(I),0.D0,C2H(I))                                     YTRA2940
      GC2(I)=G(C2A1(I),0.D0,C2H(I))                                     YTRA2950
    9 CONTINUE                                                          YTRA2960
   25 C3H=1.D0-XK(3)-XK(1)/(3.D0*XK(2)*XK(2))                           YTRA2970
      XK1=1.D0/XK(1)                                                    YTRA2980
      YM1C3=YM1(XK1,0.D0,C3H)                                           YTRA2990
      YM2C3=YM2(XK1,0.D0,C3H)                                           YTRA3000
      YV1C3=YV1(XK1,0.D0,C3H)                                           YTRA3010
      YV2C3=YV2(XK1,0.D0,C3H)                                           YTRA3020
      YMC3=1.D0+YM1C3+YM2C3                                             YTRA3030
      YSDD=YV1C3+YV2C3                                                  YTRA3040
      IF(      YSDD .LE.0.D0) YSDD=YV1C3                                YTRA3050
      YSDC3=DSQRT(YSDD)                                                 YTRA3060
      IF(NA+NB.NE.0) GO TO 20                                           YTRA3070
      DES=2.D0*(DABS(YM3C3)+DABS(YV3C3))                                YTRA3080
      IF(REF.LE.DES) GO TO 20                                           YTRA3090
      REF=DES                                                           YTRA3100
      A1=XK1                                                            YTRA3110
      A2=0.D0                                                           YTRA3120
      H=C3H                                                             YTRA3130
      EXY=YMC3                                                          YTRA3140
      SDY=YSDC3                                                         YTRA3150
   20 IF(IER.EQ.0) GO TO 69                                             YTRA3160
      YM3C3=YM3(XK1,0.D0,C3H)                                           YTRA3170
      YM4C3=YM4(XK1,0.D0,C3H)                                           YTRA3180
      YV3C3=YV3(XK1,0.D0,C3H)                                           YTRA3190
      YV4C3=YV4(XK1,0.D0,C3H)                                           YTRA3200
      BC3=B(XK1,0.D0,C3H)                                               YTRA3210
      CC3=C(XK1,0.D0,C3H)                                               YTRA3220
      DC3=D(XK1,0.D0,C3H)                                               YTRA3230
      EC3=E(XK1,0.D0,C3H)                                               YTRA3240
      FC3=F(XK1,0.D0,C3H)                                               YTRA3250
      GC3=G(XK1,0.D0,C3H)                                               YTRA3260
C                                                                       YTRA3270
C             CASE  D                                                   YTRA3280
C                                                                       YTRA3290
   69 NJ=1                                                              YTRA3300
      NF=0                                                              YTRA3310
      JJ=1                                                              YTRA3320
      NFI=0                                                             YTRA3330
      NFT=NA+NB+NC2+2                                                   YTRA3340
  310 IF(NFI.GE.NA) GO TO 301                                           YTRA3350
      NFI=NFI+1                                                         YTRA3360
      XF(1,NJ)=AA1(NFI)                                                 YTRA3370
      XF(2,NJ)=AA2(NFI)                                                 YTRA3380
      XF(3,NJ)=AH(NFI)                                                  YTRA3390
      GO TO 300                                                         YTRA3400
  301 IF(NFI.GE.NA+NB) GO TO 302                                        YTRA3410
      NFI=NFI+1                                                         YTRA3420
      XF(1,NJ)=BA1(NFI-NA)                                              YTRA3430
      XF(2,NJ)=BA2(NFI-NA)                                              YTRA3440
      XF(3,NJ)=BH(NFI-NA)                                               YTRA3450
      GO TO 300                                                         YTRA3460
  302 IF(NFI.GE.NA+NB+1) GO TO 303                                      YTRA3470
      IF(ND.EQ.0) GO TO 65                                              YTRA3480
```

```
      NFI=NFI+1                                                         YTRA3490
      XF(1,NJ)=C1A1                                                     YTRA3500
      XF(2,NJ)=0.D0                                                     YTRA3510
      XF(3,NJ)=C1H                                                      YTRA3520
      GO TO 300                                                         YTRA3530
  303 IF(NFI.GE.NA+NB+1+NC2) GO TO 304                                  YTRA3540
      NFI=NFI+1                                                         YTRA3550
      XF(1,NJ)=C2A1(NFI-NA-NB-1)                                        YTRA3560
      XF(2,NJ)=0.D0                                                     YTRA3570
      XF(3,NJ)=C2H(NFI-NA-NB-1)                                         YTRA3580
      GO TO 300                                                         YTRA3590
  304 IF(NFI.EQ.NFT)GO TO 65                                            YTRA3600
      NFI=NFI+1                                                         YTRA3610
      XF(1,NJ)=XK1                                                      YTRA3620
      XF(2,NJ)=0.D0                                                     YTRA3630
      XF(3,NJ)=C3H                                                      YTRA3640
  300 ITMAX=30                                                          YTRA3650
      CALL ZSYST4(FFF,10.D-14,13,3,XF(1,NJ),ITMAX,WA,IERF(NJ))          YTRA3660
      IF(XF(1,NJ).GT.1.D-25) GO TO 357                                  YTRA3670
      IF(XF(2,NJ).GT.1.D-25) GO TO 357                                  YTRA3680
      IF(XF(3,NJ).GT.1.D-25) GO TO 357                                  YTRA3690
      GO TO 79                                                          YTRA3700
  357 IF(IERF(NJ).NE.0) GO TO 79                                        YTRA3710
      IF(NF.EQ.0) GO TO 70                                              YTRA3720
      DO 72 K=2,NJ                                                      YTRA3730
      DO 71 J=1,3                                                       YTRA3740
      IF(DABS(XF(J,NJ)-XF(J,K-1)).GT.1.D-09) GO TO 72                   YTRA3750
   71 CONTINUE                                                          YTRA3760
      GO TO 79                                                          YTRA3770
   72 CONTINUE                                                          YTRA3780
   70 NF=NF+1                                                           YTRA3790
      IF(NF.GE.5) GO TO 65                                              YTRA3800
      NJ=NF+1                                                           YTRA3810
   79 JJ=JJ+1                                                           YTRA3820
      GO TO 310                                                         YTRA3830
   65 IF(NF.EQ.0) GO TO 166                                             YTRA3840
      DO 66 I=1,NF                                                      YTRA3850
      YM1F(I)=YM1(XF(1,I),XF(2,I),XF(3,I))                              YTRA3860
      YM2F(I)=YM2(XF(1,I),XF(2,I),XF(3,I))                              YTRA3870
      YV1F(I)=YV1(XF(1,I),XF(2,I),XF(3,I))                              YTRA3880
      YV2F(I)=YV2(XF(1,I),XF(2,I),XF(3,I))                              YTRA3890
      YMF(I)=1.D0+YM1F(I)+YM2F(I)                                       YTRA3900
      YSDD=YV1F(I)+YV2F(I)                                              YTRA3910
      IF(YSDD.LE.1.D-40)YSDD=DMAX1(1.D-40,YV1F(I))                      YTRA3920
      YSDF(I)=DSQRT(YSDD)                                               YTRA3930
      DLS=DABS(YM2F(I))+DABS(YV2F(I))                                   YTRA3940
      IF(REF.LE.DLS)GO TO 28                                            YTRA3950
      REF=DLS                                                           YTRA3960
      A1=XF(1,I)                                                        YTRA3970
      A2=XF(2,I)                                                        YTRA3980
      H =XF(3,I)                                                        YTRA3990
      EXY=YMF(I)                                                        YTRA4000
      SDY=YSDF(I)                                                       YTRA4010
   28 IF(IER.EQ.0) GO TO  66                                            YTRA4020
      YM3F(I)=YM3(XF(1,I),XF(2,I),XF(3,I))                              YTRA4030
      YM4F(I)=YM4(XF(1,I),XF(2,I),XF(3,I))                              YTRA4040
      YV3F(I)=YV3(XF(1,I),XF(2,I),XF(3,I))                              YTRA4050
      YV4F(I)=YV4(XF(1,I),XF(2,I),XF(3,I))                              YTRA4060
```

```
      FB(I)=B(XF(1,I),XF(2,I),XF(3,I))                                  YTRA4070
      FC(I)=C(XF(1,I),XF(2,I),XF(3,I))                                  YTRA4080
      FD(I)=D(XF(1,I),XF(2,I),XF(3,I))                                  YTRA4090
      FE(I)=E(XF(1,I),XF(2,I),XF(3,I))                                  YTRA4100
      FF(I)=F(XF(1,I),XF(2,I),XF(3,I))                                  YTRA4110
      FG(I)=G(XF(1,I),XF(2,I),XF(3,I))                                  YTRA4120
   66 CONTINUE                                                          YTRA4130
  166 CA1=A1                                                            YTRA4140
      CA2=A2                                                            YTRA4150
      CH=H                                                              YTRA4160
      U=XK(1)                                                           YTRA4170
      SG=A1*H                                                           YTRA4180
      IF(REF.NE.1.D+70) GO TO 160                                       YTRA4190
      IF(ICR.NE.0) WRITE(6,170)                                         YTRA4200
  170 FORMAT('0      NO SOLUTION FOUND IN ANY OF THE CASES')            YTRA4210
      IER=9                                                             YTRA4220
      RETURN                                                            YTRA4230
C                                                                       YTRA4240
  160 CEXY=EXY                                                          YTRA4250
      CSDY=SDY                                                          YTRA4260
      IF(IER.EQ.0) RETURN                                               YTRA4270
      WRITE(6,110)(I,XK(I),I=1,NCUM)                                    YTRA4280
  110 FORMAT('1'///20X,'Y = ( 1 + A1 * X + A2 * X**2 ) ** H'////30X,     YTRA4290
     & '    X    IS    ANY    RANDOM    VARIABLE'                       YTRA4300
     & //10X,'CUMULANTS OF X'//(10X,'K(',I2,')=',F25.10/))              YTRA4310
      WRITE(6,101)                                                      YTRA4320
  101 FORMAT(/////10X,'ASSUMING THAT THE TERM CONTAINING K(I) = O(R**(I-YTRA4330
     &1)) FOR SOME DOMINANT PARAMETER R'//13X,'OR EQUIVALENTLY K(I)=O(R*YTRA4340
     &I), A1= O(R**-1), AND A2= O(R**-2)' ////20X,'THEN',T30,'KY(1) = 1YTRA4350
     &.0 + M1 (R**-1) + M2 (R**-2) + M3 (R**-3) + M4 (R**-4)'//         YTRA4360
     & T30,'KY(2) =',7X,'V1 (R**-1) + V2 (R**-2) + V3 (R**-3) + V4 (R**-YTRA4370
     &4)'//T30,'KY(3) =',21X,'B (R**-2) +  D (R**-3) +  G (R**-4)'//T30 YTRA4380
     &,'KY(4) =',34X,'C (R**-3) +  F (R**-4)'//T30,'KY(5) =',47X,'E (R*YTRA4390
     &-4)')                                                             YTRA4400
      WRITE(6,102)                                                      YTRA4410
  102 FORMAT('1'///5X,'CASE :  A            SETTING   B = 0  .  C = 0  . YTRA4420
     & E = 0')                                                          YTRA4430
      IF(NA.NE.0)GO TO 104                                              YTRA4440
      WRITE(6,103)                                                      YTRA4450
  103 FORMAT(//////////////40X,'THE EQUATIONS HAVE NO REAL SOLUTION ')  YTRA4460
      GO TO 105                                                         YTRA4470
  104 WRITE(6,106)NA                                                    YTRA4480
  106 FORMAT(////20X,'THERE ARE',I3,' SOLUTIONS (A1,A2,H)')             YTRA4490
      DO 107 I=1,NA                                                     YTRA4500
      WRITE(6,108)I,AA1(I),AA2(I),AH(I)                                 YTRA4510
  108 FORMAT(////10X,'(',I1,')',5X,'(',3F20.14,')')                     YTRA4520
  107 WRITE(6,109)YM1A(I),YM2A(I),YM3A(I),YM4A(I),YV1A(I),YV2A(I),      YTRA4530
     & YV3A(I),YV4A(I),BAA(I),DAA(I),GAA(I),CAA(I),FAA(I),EAA(I)        YTRA4540
  109 FORMAT(/10X,'THE CONSTANTS IN THE CUMULANTS OF Y ARE'//T15,'KY(1)=YTRA4550
     & 1.0 + M1=',F17.13,' + M2=',F17.13,' + M3=',F17.13,' + M4=',F17.  YTRA4560
     & 13,' + R**-5'//T15,'KY(2)=',7X,'V1=',F17.13,' + V2=',F17.13,' + YTRA4570
     &3=',F17.13,' + V4=',F17.13,' + R**-5'.//T15,'KY(3)=',T51,'B=',    YTRA4580
     & F18.14,    ' + D=',F18.14,    ' + G=',F18.14,    ' + R**-5'//T15 YTRA4590
     & ,'KY(4)=',T74,'C=',F18.14,    ' + F=', F18.14,    ' + R**-5',   YTRA4600
     & //T15,'KY(5)=',T97,'E=',F18.14,    ' + R**-5')                  YTRA4610
  105 WRITE(6,122)                                                      YTRA4620
  122 FORMAT('1'////5X,'CASE :  B            SETTING   E = 0  .  C = 0  .YTRA4630
     &  D = 0')                                                         YTRA4640
```

```
      IF(NC.NE.0)GO TO 124                                              YTRA4650
      WRITE(6,103)                                                      YTRA4660
      GO TO 125                                                         YTRA4670
  124 WRITE(6,106)NB                                                    YTRA4680
      DO 127 I=1,NB                                                     YTRA4690
      WRITE(6,108)I,BA1(I),BA2(I),BH(I)                                 YTRA4700
  127 WRITE(6,109)YM1B(I),YM2B(I),YM3B(I),YM4B(I),YV1B(I),YV2B(I),      YTRA4710
     & YV3B(I),YV4B(I),BBB(I),DBB(I),GBB(I),CBB(I),FBB(I),EBB(I)        YTRA4720
  125 WRITE(6,142)                                                      YTRA4730
  142 FORMAT('1'///5X,'CASE :  C            SPECIAL CASES WHEN A2 = 0'  YTRA4740
     & //7X,'C1 :  (HALDANE TRANSFORMATION B)   SETTING B = 0  .  C = 0YTRA4750
     &')                                                                YTRA4760
      IF(ND.NE.0) GO TO 144                                             YTRA4770
      WRITE(6,121)                                                      YTRA4780
  121 FORMAT(////40X,'THE EQUATIONS HAVE NO SOLUTION'      //////7X,    YTRA4790
     & 'C2 :          SETTING B = 0  .  D = 0'////40X,'THE EQUATIONS HAVE YTRA4800
     &NO SOLUTION'//////7X,'C3 :   (HALDANE TRANSFORMATION A)   SETTING YTRA4810
     & B = 0  .  A1 = 1 / K(1)'////40X,'THE EQUATIONS HAVE NO SOLUTION')YTRA4820
      GO TO 184                                                         YTRA4830
  144 WRITE(6,146)                                                      YTRA4840
  146 FORMAT(/ 20X,' THE SOLUTION IS  (A1,A2,H)' )                      YTRA4850
      WRITE(6,148)C1A1,C1H                                              YTRA4860
  148 FORMAT(/33X,'(',F20.14,10X,'0.0',7X,F20.14,')')                   YTRA4870
      WRITE(6,109)YM1C1,YM2C1,YM3C1,YM4C1,YV1C1,YV2C1,YV3C1,YV4C1,      YTRA4880
     &BC1,DC1,GC1,CC1,FC1,EC1                                           YTRA4890
      WRITE(6,152)                                                      YTRA4900
  152 FORMAT(//////7X,'C2 :          SETTING   B = 0  .  D = 0')        YTRA4910
      IF(NC2.NE.0)GO TO 154                                             YTRA4920
      WRITE(6,153)                                                      YTRA4930
  153 FORMAT(/////        40X,'THE EQUATIONS HAVE NO REAL SOLUTION ')   YTRA4940
      GO TO 155                                                         YTRA4950
  154 WRITE(6,156)NC2                                                   YTRA4960
  156 FORMAT(// 20X,'THERE ARE',I3,' SOLUTIONS (A1,A2,H)')              YTRA4970
      DO 157 I=1,NC2                                                    YTRA4980
      WRITE(6,158)I,C2A1(I),C2H(I)                                      YTRA4990
  158 FORMAT(/25X,'(',I1,')',5X,'(',F20.14,5X,'0.0',F25.14,')')         YTRA5000
  157 WRITE(6,109)YM1C2(I),YM2C2(I),YM3C2(I),YM4C2(I),YV1C2(I),YV2C2(I),YTRA5010
     & YV3C2(I),YV4C2(I),BC2(I),DC2(I),GC2(I),CC2(I),FC2(I),EC2(I)      YTRA5020
  155 WRITE(6,162)                                                      YTRA5030
  162 FORMAT(/////7X,'C3 :   (HALDANE TRANSFORMATION A)   SETTING   B = YTRA5040
     &0   :   A1 = 1 / K(1)')                                           YTRA5050
      WRITE(6,146)                                                      YTRA5060
      WRITE(6,148)XK1,C3H                                               YTRA5070
      WRITE(6,109)YM1C3,YM2C3,YM3C3,YM4C3,YV1C3,YV2C3,YV3C3,YV4C3,BC3,  YTRA5080
     &DC3,GC3,CC3,FC3,EC3                                               YTRA5090
  184 WRITE(6,192)NFI,NF                                                YTRA5100
  192 FORMAT('1'///5X,'CASE : D'//10X,'WE USE BROWNS METHOD OF SOLVING'/YTRA5110
     &///10X,'THE FOLLOWING NON-LINEAR SIMULTANEOUS EQUATIONS'//////15X,YTRA5120
     & '(1) B + D + G = 0'/15X,'(2)    C + F = 0'/15X,'(3)',9X,'E = 0'/YTRA5130
     &///10X,'USING ',I2,' INITIAL GUESSES. WE OBTAIN',I2,' SOLUTIONS'/)YTRA5140
      IF(NF.EQ.0) GO TO 193                                             YTRA5150
      DO 194 I=1,NF                                                     YTRA5160
      WRITE(6,108)I,XF(1,I),XF(2,I),XF(3,I)                             YTRA5170
  194 WRITE(6,109)YM1F(I),YM2F(I),YM3F(I),YM4F(I),YV1F(I),YV2F(I),      YTRA5180
     & YV3F(I),YV4F(I),       FB(I),FD(I),FG(I),FC(I),FF(I),FE(I)      YTRA5190
  193 RETURN                                                            YTRA5200
      END                                                               YTRA5210
```

```
C                                                                       YTRA5220
C                                                                       YTRA5230
      FUNCTION FRCB(X)                                                  YTRA5240
C                                                                       YTRA5250
C  PURPOSE:                                                             YTRA5260
C        TO FIND CUMULATIVE PROBABILITY AT X GIVEN ONLY THE             YTRA5270
C        CUMULANTS. THE METHOD IS TO APPLY THE TRANSFORMATION           YTRA5280
C        Y=(1+A1*(X-U)+A2*(X-U)**2)**H. THE SOLUTIONS OF THE            YTRA5290
C        CONSTANTS A1,A2,H TOGETHER WITH THE MEAN AND STANDARD          YTRA5300
C        DEVIATION OF Y (PASSED THROUGH THE COMMON BLOCK CCM) ARE       YTRA5310
C        OBTAINED FROM THE SUBROUTINE YTRANS.                           YTRA5320
C  PARAMETERS:                                                          YTRA5330
C        X     - THE PERCENTILE POINT AT WHICH THE PROBABILITY IS       YTRA5340
C        TO BE EVALUED                                                  YTRA5350
C                                                                       YTRA5360
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA5370
      COMMON /CCM/A1,A2,H,EXY,SDY,SG,U                                  YTRA5380
      XX=X-U                                                            YTRA5390
      YC=1.D0+A1*XX+A2*XX*XX                                            YTRA5400
      YZ=DSIGN(DABS(YC)**H,YC)                                          YTRA5410
      IF(H*(A1+2.D0*A2*XX)*SG.GT.0.D0) GO TO 1                          YTRA5420
      XA=-A1/.5D0/A2                                                    YTRA5430
      YZZ=1.D0+A1*XA+A2*XA*XA                                           YTRA5440
      YZ=2.D0*DSIGN(DABS(YZZ)**H,YZZ)-YZ                                YTRA5450
    1 YZ=(YZ-EXY)/SDY                                                   YTRA5460
      PROB=.5D0+.5D0*DERF(.7071067811865475*YZ)                         YTRA5470
      IF(SG.LT.0.D0) PROB=1.D0-PROB                                     YTRA5480
      RETURN                                                            YTRA5490
      END                                                               YTRA5500
C                                                                       YTRA5510
C                                                                       YTRA5520
C                                                                       YTRA5530
      FUNCTION CENT(P)                                                  YTRA5540
C                                                                       YTRA5550
C  PURPOSE:                                                             YTRA5560
C        TO FIND THE PERCENTAGE POINT OF A DISTRIBUTION                 YTRA5570
C              GIVEN  1.- THE PROBABILITY                               YTRA5580
C                     2. - A NORMAL RANDOM VARIABLE WITH THE SAME       YTRA5590
C                          PROBABILITY LEVEL                            YTRA5600
C                          ENTRY CENTY(Y) USED IN THIS CASE             YTRA5610
C        AND GIVEN ONLY THE CUMULANTS OF THE DISTRIBUTION.              YTRA5620
C        THE METHOD USED IS TO INVERT THE NORMALIZING TRANSFORMATION    YTRA5630
C        Y=(1+A1*(X-U)+A2*(X-U)**2)**H WHERE THE CONSTANTS ARE          YTRA5640
C        SUPPLIED BY THE SUBROUTINE  YTRANS.                            YTRA5650
C  PARAMETERS:                                                          YTRA5660
C        P     - THE PROBABILITY WHOSE PERCENTILE IS TO BE CALCULATED   YTRA5670
C  IN CASE OF ENTRY CENTY(Y)                                            YTRA5680
C        Y     - THE NORMAL VARIATE CORRESPONDING IN PROBABILITY        YTRA5690
C                TO THE REQUIRED RANDOM VARIABLE.                       YTRA5700
C                                                                       YTRA5710
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA5720
      COMMON /CCM/A1,A2,H,EXY,SDY,SG,U                                  YTRA5730
      Y=PHINV(P)                                                        YTRA5740
      GO TO 3                                                           YTRA5750
      ENTRY CENTY(Y1)                                                   YTRA5760
      Y=Y1                                                              YTRA5770
    3 Y=Y*SDY+EXY                                                       YTRA5780
      IF(A2.LE.0) GO TO 2                                               YTRA5790
```

```
      C=A1*A1-4.D0*A2 (1.D0-DSIGN(DABS(Y)**(1.D0/H).Y))          YTRA5800
      IF(C.LT.0.D0) GO TO 1                                      YTRA5810
      CENT=(-A1+DSIGN(DSQRT(C).A1))/(2.D0*A2)+U                  YTRA5820
      RETURN                                                     YTRA5830
    1 WRITE(6.9)                                                 YTRA5840
    9 FORMAT('0 UNABLE TO FIND PERCENTILE DUE TO NEGATIVE '.     YTRA5850
     & 'DISCRIMINANT')                                           YTRA5860
      CENT=1.D-74                                                YTRA5870
      RETURN                                                     YTRA5880
    2 CENT=(DSIGN(DABS(Y)**(1.D0/H).Y)-1.D0)/A1+U                YTRA5890
      CENTY=CENT                                                 YTRA5900
      RETURN                                                     YTRA5910
      END                                                        YTRA5920
C                                                                YTRA5930
C                                                                YTRA5940
C                                                                YTRA5950
      FUNCTION PHINV(P)                                          YTRA5960
C                                                                YTRA5970
C     INVERSE NORMAL                                             YTRA5980
C                                                                YTRA5990
      IMPLICIT REAL*8 (A-H.O-Z)                                  YTRA6000
      F(C)= .5D0+.5D0*DERF(.707106781186547 5*C)-P               YTRA6010
      X=0.D0                                                     YTRA6020
      DO 2 I=1.1000                                              YTRA6030
      EPS=F(X)/(.39894228040143*DEXP(-.5D0*X*X))                 YTRA6040
      X=X-EPS                                                    YTRA6050
      IF(DABS(EPS).LT.1.D-15) GO TO 1                            YTRA6060
    2 CONTINUE                                                   YTRA6070
      WRITE(6.9)                                                 YTRA6080
    9 FORMAT('0 INVERSE NORMAL DOES NOT CONVERGE AFTER 1000 ITERATIONS')YTRA6090
      PHINV=-1.D50                                               YTRA6100
      RETURN                                                     YTRA6110
    1 PHINV=X                                                    YTRA6120
      RETURN                                                     YTRA6130
      END                                                        YTRA6140
```

```
C
C
C   THE FOLLOWING ARE FUNCTIONS OF THE DIFFERENT ORDERS OF THE
C          CUMULANTS OF Y
C        KY(1) = 1 + YM1 + YM2 + YM3 + YM4
C        KY(2) =     YV1 + YV2 + YV3 + YV4
C        KY(3) =             B   +  D   +  G
C        KY(4) =                     C   +  F
C        KY(5) =                            E
C
C
      FUNCTION B(A1,A2,H)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XK
      DIMENSION XK(5)
      B          =A1*A1*H*H*H*(A1*XK(3)+3.D0*XK(2)*XK(2)*(A1*A1*(H-1.D0)
     & +2.D0 A2))
      RETURN
      END
C
      FUNCTION C(A1,A2,H)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XK
      DIMENSION XK(5)
      C          =A1*A1*H*H*4*(A1*A1*XK(4)+12.D0*XK(3)*XK(2)*A1*(A1*A1*(
     & H-1.D0)+2.D0*A2)+4.D0*XK(2)*XK(2) XK(2)*(A1**4*(H-1.D0)*(4.D0*H-
     & 5.D0)+18.D0*A1*A1*A2*(H-1.D0)+12.D0*A2*A2))
      RETURN
      END
C
      FUNCTION D(A1,A2,H)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XK
      DIMENSION XK(5)
      D          =.5D0 H H*H*(3.D0*XK(4)*A1*A1*(A1*A1*(H-1.D0)+2.D0*A2)
     & +3.D0*XK(3)*XK(2)*A1*A1*4*(H-1.D0)*(7.D0*H-10.D0)+34.D0*A1*A1*
     & A2*(H-1.D0)+16.D0*A2*A2)+XK(2)*XK(2)*XK(2)*(A1**6*(H-1.D0)*(17.D0
     &  H*H-55.D0 H+44.D0)+12.D0*A1 4*A2*(H-1.D0)*(12.D0*H-17.D0)+204.
     & D0 A1*A1*A2*A2*(H-1.D0)+16.D0*A2*A2*A2))
      RETURN
      END
C
      FUNCTION E(A1,A2,H)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XK
      DIMENSION XK(5)
      E          =A1*A1 H*H*5*(A1*A1*A1*XK(5)+5.D0*(3.D0*XK(3)*XK(3)+4.D0
     &  XK(4)*XK(2))*A1*A1*(A1*A1*(H-1.D0)+2.D0*A2)+30.D0*XK(3)*XK(2)
     & *XK(2)*A1 (A1* 4*(H-1.D0)*(5.D0 H-6.D0)+22.D0*A1*A1*A2*(H-1.D0)
     & +16.D0*A2 A2)+5.D0*XK(2)*+4*(A1**6*(H-1.D0)*(25.D0*H*H-65.D0*H+
     & 42.D0)+36.D0*A1*A1*A2*(H-1.D0)*(5.D0*H-6.D0)+300.D0*A1*A1*A2*A2
     & *(H-1.D0)+96.D0*A2*A2*A2))
      RETURN
      END
C
      FUNCTION F(A1,A2,H)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XK
```

```
      DIMENSION XK(5)                                                   YTRA6730
      F=2.D0-XK(5) A1 A1*A1*H**4*(A1*A1*(H-1.D0)+2.D0* A2)+2.D0*XK(4)*  YTRA6740
     & XK(2) A1*A1*H* 4*(A1 -4*(H-1.D0)*(13.D0*H-17.D0)+60.D0*A1*A1*A2  YTRA6750
     & *(H-1.D0)+36.D0*A2*A2)+6.D0*XK(3)*XK(3)*A1*A1*H**4*(A1**4 (H-    YTRA6760
     & 1.D0) (3.D0*H-4.D0)+14.D0*A1*A1*A2*(H-1.D0)+8.D0*A2*A2)+6.D0*XK  YTRA6770
     & (3)*XK(2) XK(2)*A1*H* 4*(A1**5 (H-1.D0)*(21.D0*H*H-61.D0*H+44.D0 YTRA6780
     & )+2.D0 A1**4*A2 (H-1.D0)*(82.D0*H-107.D0)+252.D0*A1*A1*A2*A2*(   YTRA6790
     & H-1.D0)+48.D0*A2*A2*A2)                                          YTRA6800
      F=F+XK(2)  4*H *4*(A1**6*(H-1.D0)*(71.D0*H*H+H-339.D0 H*H+535.D0  YTRA6810
     & *H-279.D0)+4.D0 A1**6*A2*(H-1.D0)*(205.D0*H*H-553.D0*H+426.D0)   YTRA6820
     & +24.D0*A1**4*A2 A2*(H-1.D0)*(95.D0 H-124.D0)+1392.D0*A1*A1*A2    YTRA6830
     &  A2 A2*(H-1.D0)+48.D0 A2**4)                                     YTRA6840
      RETURN                                                            YTRA6850
      END                                                               YTRA6860
C                                                                       YTRA6870
      FUNCTION G(A1,A2,H)                                               YTRA6880
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA6890
      COMMON XK                                                         YTRA6900
      DIMENSION XK(5)                                                   YTRA6910
      G=.25D0*H* 3 (XK(5)* A1*(A1**4*(H+1.D0)*(5.D0*H-7.D0)+24.D0*A1*A1 YTRA6920
     & *A2 (H-1.D0)+12.D0*A2*A2)+XK(4)*XK(2)*(A1**6*(H-1.D0)*(41.D0*H   YTRA6930
     & *(H-131.D0*H+104.D0)+4.D0*A1**4*A2*(H-1.D0)*(86.D0*H-121.D0)+492 YTRA6940
     & .D0 A1*A1 A2*A2*(H-1.D0)+48.D0*A2 A2*A2)+ XK(3)*XK(3)*(A1**6*    YTRA6950
     & (H-1.D0) (23.D0 H H-89.D0*H+71.D0)+6.D0*A1**4*A2*(H-1.D0)*(39.   YTRA6960
     & D0*(H-55.D0)+336.D0*A1*A1*A2*A2*(H-1.D0)+40.D0*A2*A2*A2))        YTRA6970
      G=G+.375D0 XK(3)*XK(2) XK(2)        (H*H*H*                       YTRA6980
     & A1 (H-1.D0)*(A1**6 (H-2.D0)*(87.D0*H*H-289.D0*H+240.D0)+4.D0*A1**YTRA6990
     & *A2 (275.D0 H H-677.D0 H+636.D0)+4.D0 A1*A1*A2*A2*(765.D0*H     YTRA7000
     & -1072.D0)+1408.D0 A2 A2*A2)                                      YTRA7010
      IF(A1.LE.1.D-10.OR.A2.LE.1.D-20) GO TO 1                          YTRA7020
      G=G+.125D0*XK(2)*XK(2)*H*H H*(H-1.D0) (A1**8*(H-2.D0)*(101.D0*H*H*H YTRA7030
     & -582.D0 H H+1111.D0*(H-702.D0)+6.D0*A1  6*A2*(H-2.D0)*(295.D0*H  YTRA7040
     & *(H-977.D0*(H+803.D0)+12.D0*A1**4*A2*A2*(665.D0*H*H-2109.D0*H+1666.YTRA7050
     & D0)+72.D0*A1*A1*A2*A2*A2 (121.D0 H-168.D0)+864.D0*A2**4)         YTRA7060
    1 RETURN                                                            YTRA7070
      END                                                               YTRA7080
C                                                                       YTRA7090
      FUNCTION YM1(A1,A2,H)                                             YTRA7100
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7110
      COMMON XK                                                         YTRA7120
      DIMENSION XK(5)                                                   YTRA7130
      YM1=        (A1*A1+2.D0*A2/(H-1.D0))*H*(H-1.D0)*XK(2)/2.D0        YTRA7140
      RETURN                                                            YTRA7150
      END                                                               YTRA7160
C                                                                       YTRA7170
      FUNCTION YM2(A1,A2,H)                                             YTRA7180
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7190
      COMMON XK                                                         YTRA7200
      DIMENSION XK(5)                                                   YTRA7210
      YM2=(A1*A1+6.D0*A2/(H-2.D0))*A1*H*(H-1.D0)*(H-2.D0)*XK(3)/6.D0    YTRA7220
     & +(A1* 4+12.D0 A1*A1*A2/(H-3.D0)+12.D0*A2*A2/((H-3.D0)*(H-2.D0)))YTRA7230
     & *H*(H-1.D0)*(H-2.D0)*(H-3.D0)* XK(2)*XK(2) / 6.D0                YTRA7240
      RETURN                                                            YTRA7250
      END                                                               YTRA7260
C                                                                       YTRA7270
      FUNCTION YM3(A1,A2,H)                                             YTRA7280
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7290
      COMMON XK                                                         YTRA7300
```

```
      DIMENSION XK(5)                                                   YTRA7310
      YM3=(A1**4+12.D0*A1*A1*A2/(H-3.D0)+12.D0*A2*A2/((H-3.D0)*(H-2.D0))YTRA7320
     & )*H*(H-1.D0)*(H-2.D0)*(H-3.D0)*XK(4)/24.D0                       YTRA7330
     &  +(A1**4+20.D0*A1*A1*A2/(H-4.D0)+60.D0*A2*A2/((H-4.D0)*(H-3.D0   YTRA7340
     & )))*A1*H*(H-1.D0)*(H-2.D0)*(H-3.D0)*(H-4.D0)*XK(3)*XK(2)/12.D0   YTRA7350
     & +  (A1**6+30.D0*A1*A1*A1*A1*A2/(H-5.D0)+180.D0*A1*A1*A2*A2/((H-5.YTRA7360
     & D0)*(H-4.D0))+120.D0*A2*A2*A2/((H-5.D0)*(H-4.D0)*(H-3.D0)))*H*(H YTRA7370
     & -1.D0)*(H-2.D0)*(H-3.D0)*(H-4.D0)*(H-5.D0)*XK(2)**3/48.D0        YTRA7380
      RETURN                                                            YTRA7390
      END                                                               YTRA7400
C                                                                       YTRA7410
      FUNCTION YM4(A1,A2,H)                                             YTRA7420
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7430
      COMMON XK                                                         YTRA7440
      DIMENSION XK(5)                                                   YTRA7450
      YM4=(A1**4+20.D0*A1*A1*A2/(H-4.D0)+60.D0*A2*A2/((H-4.D0)*(H-3.D0  YTRA7460
     & )))*A1*H*(H-1.D0)*(H-2.D0)*(H-3.D0)*(H-4.D0)*XK(5)/120.D0        YTRA7470
     & +  (A1**6+30.D0*A1*A1*A1*A1*A2/(H-5.D0)+180.D0*A1*A1*A2*A2/((H-5 YTRA7480
     & .D0)*(H-4.D0))+120.D0*A2*A2*A2/((H-5.D0)*(H-4.D0)*(H-3.D0)))*H*( YTRA7490
     & H-1.D0)*(H-2.D0)*(H-3.D0)*(H-4.D0)*(H-5.D0)*(3.D0*XK(4)*XK(2)   YTRA7500
     & +2.D0*XK(3)*XK(3))/144.D0                                        YTRA7510
      YM4= YM4          +  (A1**6+42.D0*A1*A1*A1*A1*A2/(H-6.D0)+420.D0 YTRA7520
     & *A1*A1*A2*A2/((H-6.D0)*(H-5.D0))+840.D0*A2*A2*A2/((H-6.D0)*(H-   YTRA7530
     & 5.D0)*(H-4.D0)))*A1*H*(H-1.D0)*(H-2.D0)*(H-3.D0)*(H-4.D0)*(H-   YTRA7540
     & 5.D0)*(H-6.D0)*XK(3)*XK(2)*XK(2)/48.D0  +  (A1**8+56.D0*A2*A1    YTRA7550
     & *A1/(H-7.D0)+840.D0*A1*A1*A1*A1*A1*A2*A2/((H-7.D0)*(H-6.D0))+    YTRA7560
     & 3360.D0*A1*A1*A2*A2*A2/((H-7.D0)*(H-6.D0)*(H-5.D0))+1680.D0*A2   YTRA7570
     & *A2*A2*A2/((H-7.D0)*(H-6.D0)*(H-5.D0)*(H-4.D0)))*H*(H-1.D0)*     YTRA7580
     & (H-2.D0)*(H-3.D0)*(H-4.D0)*(H-5.D0)*(H-6.D0)*(H-7.D0)*XK(2)*     YTRA7590
     & XK(2)*XK(2)*XK(2)/384.D0                                         YTRA7600
      RETURN                                                            YTRA7610
      END                                                               YTRA7620
C                                                                       YTRA7630
      FUNCTION YV1(A1,A2,H)                                             YTRA7640
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7650
      COMMON XK                                                         YTRA7660
      DIMENSION XK(5)                                                   YTRA7670
      YV1=A1*A1*H*H*XK(2)                                               YTRA7680
      RETURN                                                            YTRA7690
      END                                                               YTRA7700
C                                                                       YTRA7710
      FUNCTION YV2(A1,A2,H)                                             YTRA7720
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7730
      COMMON XK                                                         YTRA7740
      DIMENSION XK(5)                                                   YTRA7750
      YV2=XK(3)*H*H*A1*(A1*A1*(H-1.D0)+2.D0*A2)+XK(2)*XK(2)*H*H*(A1**4*(YTRA7760
     & H-1.D0)+(3.D0*H-5.D0)+16.D0*A1*A1*A2*(H-1.D0)+4.D0*A2*A2)/2.D0   YTRA7770
      RETURN                                                            YTRA7780
      END                                                               YTRA7790
C                                                                       YTRA7800
      FUNCTION YV3(A1,A2,H)                                             YTRA7810
      IMPLICIT REAL*8 (A-H,O-Z)                                         YTRA7820
      COMMON XK                                                         YTRA7830
      DIMENSION XK(5)                                                   YTRA7840
      YV3=XK(4)*H*H*(A1**4*(H-1.D0)*(7.D0*H-11.D0)+36.D0*A1*A1*A2*(H-   YTRA7850
     & 1.D0)+12.D0*A2*A2)/12.D0+XK(3)*XK(2)*H*H*(H-1.D0)*A1*(A1**4*(H-2 YTRA7860
     & .D0)*(7.D0*H-12.D0)+3.D0*A1*A1*A2*(22.D0*H-35.D0)+84.D0*A2*A2)/3.YTRA7870
     & D0+XK(2)*XK(2)*XK(2)*H*H*(H-1.D0)*(A1**6*(H-2.D0)*(7.D0*H*H-30.  YTRA7880
```

```
     & D0*H+52.D0)+6.D0*A1**4+A2*(H-2.D0)*(17.DC *H-29.DC)+18.DC*A1*A1   YTRA7890
     & *A2*A2*(16.D0 H-25.D0)+72.DQ*A2*A2*A2)/6.D0                       YTRA7900
      RETURN                                                             YTRA7910
      END                                                                YTRA7920
C                                                                        YTRA7930
      FUNCTION YV4(A1,A2,H)                                              YTRA7940
      IMPLICIT REAL*8 (A-H,O-Z)                                          YTRA7950
      COMMON XK                                                          YTRA7960
      DIMENSION XK(5)                                                    YTRA7970
      YV4    =    XK(5)*H*H*(H-1.D0)*A1*(A1*A1*A1*A1*(H-2.D0)*(3.D0*H-   YTRA7980
     & 5.D0)+4.D0*A1*A1*A2*(7.D0*H-11.D0)+36.D0*A2*A2)/12.D0             YTRA7990
     & +XK(4) XK(2)*H*H*(H-1.D0)*(A1**6*(H-2.D0)*(15.DC*H*H-64.D0*H+ ,   YTRA8000
     & 67.(0)+2.DD A1 A1*A1*A1*A2*(H-2.D0)*(109.C0*H-183.D0)+12.D0*A1*A1YTRA8010
     & *A2*A2*(51.DC H-80.DC)+168.D0*A2 A2*A2)/12.DC  +  XK(3)*XK(3)*    YTRA8020
     & H H (H-1.DD)*(A1**6*(H-2.DD)*(30.DD H*H-129.DC*H+135.DC)+6.D0*    YTRA8030
     & A1*A1*A1*A1*A2*(H-2.CC)*(73.DC*H-123.D0)+72.D0*A1*A1*A2*A2*(17.DCYTRA8040
     & *H-27.DD)+35C.D0*A2*A2 A2)/36.D0                                  YTRA8050
      YV4    =    YV4          +              XK(3)*XK(2)*XK(2) H*H*(H-   YTRA8060
     & 1.DD)*A1*(A1**6*(H-2.DD)*(H-3.CD)*(15.DD*H*H-65.DD*H+71.DD)+2.D0*YTRA8070
     & A1*A1*A1*A1*A2 (H-2.DD)*(155.DC H*H-662.D0*H+693.CD)+2.DC*A1*A1* YTRA8080
     & A2*A2 (H-2.DC) (752.CC*H-1257.DD)+12.DD*A2*A2*A2*(116.D0*H-181    YTRA8090
     & .DD))/6.DD  +  XK(2)*XK(2)*XK(2)*XK(2)*H*H*(H-1.DC)*(A1**6*       YTRA8100
     & (H-2.DD)*(H-3.DD) (15.DD*H H*H-118.DD*H*H+30C.DC*H-260.D0)+8.D0   YTRA8110
     & *A1 **6*A2*(H-2.DC)*(H-3.DD)*(52.DD*H*H-225.DC*H+245.DC)+12.D0*   YTRA8120
     & A1*A1 A1*A1 A2*A2*(H-2.DD)*(256.DD H*H-1091.CC*H+1135.D0)+192.D0 YTRA8130
     & *A1*A1 A2 A2*A2*(H-2.DD)*(31.CD*H-51.D0)+48.DC*A2*A2*A2*A2*(27.DCYTRA8140
     & *H-42.DD))/24.DD                                                  YTRA8150
      RETURN                                                             YTRA8160
      END                                                                YTRA8170
C                                                                        YTRA8180
C                                                                        YTRA8190
      SUBROUTINE DCUBIC(A,B,C,R1,R2,R3,DIS)                              YTRA8200
C                                                                        YTRA8210
C   FINDS THE REAL ROOTS R1,R2,R3, AND THE DISCRIMINANT DIS              YTRA8220
C     OF THE CUBIC        X**3 + A * X**2 + B * X + C = 0                YTRA8230
C                                                                        YTRA8240
      IMPLICIT REAL*8 (A-H,O-Z)                                          YTRA8250
      P=B-A A/3.D0                                                       YTRA8260
      Q=C-B A/3.D0+2.D0*A*A*A/27.D0                                      YTRA8270
      DIS=4.D0*P P*P+27.D0*Q Q                                           YTRA8280
      IF(DIS)3,3,1                                                       YTRA8290
    1 S=DSQRT(DIS/108.D0)                                                YTRA8300
      S1=DSIGN(1.D0,S-Q/2.D0)*DABS(S-Q/2.D0)** .3333333333333333         YTRA8310
      S2=DSIGN(1.D0,-S-Q/2.D0)*DABS(-S-Q/2.D0)**.3333333333333333        YTRA8320
      R1=S1+S2-A/3.D0                                                    YTRA8330
      R2=-.5D0 (S1+S2)-A/3.D0                                            YTRA8340
      R3=(S1-S2)*.3660254037844386                                       YTRA8350
      RETURN                                                             YTRA8360
    3 Y=DSQRT(-DIS/27.D0)/(-Q)                                           YTRA8370
      PHI=(DATAN(Y)+(1.DC-DSIGN(1.D0,Y))*1.5707963267948897)/3.D0        YTRA8380
      T=2.D0 DSQRT(-P/3.CD)                                              YTRA8390
      R1=T DCOS(PHI)-A/3.CD                                              YTRA8400
      R2=-T*DCOS(1.047197551196598-PHI)-A/3.D0                           YTRA8410
      R3=-T*DCOS(1.047197551196598+PHI)-A/3.D0                           YTRA8420
      RETURN                                                             YTRA8430
      END                                                                YTRA8440
C                                                                        YTRA8450
C                                                                        YTRA8460
```

```
      FUNCTION FFF(X,K)                                                 YTRA8470
C                                                                       YTRA8480
C     FUNCTION  REQUIRED BY THE IMSL ROUTINE ZSYSTM                     YTRA8490
C                                                                       YTRA8500
      IMPLICIT REAL*8 (A-H,C-Z)                                         YTRA8510
      DIMENSION X(1)                                                    YTRA8520
      GO TO (1,2,3),K                                                   YTRA8530
    1 FFF=B(X(1),X(2),X(3))+D(X(1),X(2),X(3))+G(X(1),X(2),X(3))         YTRA8540
      RETURN                                                            YTRA8550
    2 FFF=C(X(1),X(2),X(3))+F(X(1),X(2),X(3))                           YTRA8560
      RETURN                                                            YTRA8570
    3 FFF=L(X(1),X(2),X(3))                                             YTRA8580
      RETURN                                                            YTRA8590
      END                                                               YTRA8600
```

```
Y = ( 1 + A1 * X + A2 * X**2 ) ** H
```

X IS A NON-CENTRAL CHI-SQUARE DISTRIBUTION

PARAMETERS N= 10. D= 5.00

CUMULANTS OF X

K(1)=	15.0000000000
K(2)=	40.0000000000
K(3)=	200.0000000000
K(4)=	1440.0000000000
K(5)=	13440.0000000000
K(6)=	153600.0000000000

ASSUMING THAT THE TERM CONTAINING K(I)= O(R**(1-I)) FOR SOME DOMINANT PARAMETER R

OR EQUIVALENTLY K(I)=O(R**I). A1= O(R**-1). AND A2= O(R**-2)

```
THEN    KY(1) = 1.0 + M1 (R**-1) + M2 (R**-2) + M3 (R**-3) + M4 (R**-4)
        KY(2) =       V1 (R**-1) + V2 (R**-2) + V3 (R**-3) + V4 (R**-4)
        KY(3) =                    B (R**-2) +  D (R**-3) +  G (R**-4)
        KY(4) =                                 C (R**-3) +  F (R**-4)
        KY(5) =                                            E (R**-4)
```

```
CASE :  A          SETTING   B = 0  ,  C = 0  ,  E = 0

          THERE ARE  3 SOLUTIONS:(A1,A2,H)

(1)     (     0.15479838420561     0.00605680006204     0.22531029720086)
THE CONSTANTS IN THE CUMULANTS OF Y ARE
    KY(1)= 1.0 + M1= -0.0290647249596 + M2= -0.0006030930429 + M3=  0.0001122569540 + M4=  0.0001402790555 + R**-5
    KY(2)=       V1=  0.0486580744501 + V2=  0.0008785485665 + V3= -0.0000482179808 + V4= -0.0001550988457 + R**-5
    KY(3)=                            B= -0.00000000000000 + D= -0.00001822789910 + G=  0.00009088918272 + R**-5
    KY(4)=                                                 C= -0.00000000000000 + F= -0.00002507221354 + R**-5
    KY(5)=                                                                        E= -0.00000000000000 + R**-5

(2)     (     0.07173434595325     0.00009688184140     0.38149857604948)
THE CONSTANTS IN THE CUMULANTS OF Y ARE
    KY(1)= 1.0 + M1= -0.0228054590292 + M2= -0.0004732132749 + M3= -0.0000140627538 + M4= -0.0000054237792 + R**-5
    KY(2)=       V1=  0.0299571241842 + V2=  0.0005408925200 + V3=  0.0000227778262 + V4=  0.0000089143319 + R**-5
    KY(3)=                            B=  0.00000000000000 + D= -0.00000880550803 + G= -0.00000758317559 + R**-5
    KY(4)=                                                 C= -0.00000000000000 + F=  0.00000306987533 + R**-5
    KY(5)=                                                                        E=  0.00000000000000 + R**-5

(3)     (     0.08248900897157     0.00112919853134     0.16298153145596)
THE CONSTANTS IN THE CUMULANTS OF Y ARE
    KY(1)= 1.0 + M1= -0.0112034875087 + M2= -0.0002324723658 + M3=  0.0000175589067 + M4=  0.0000057585689 + R**-5
    KY(2)=       V1=  0.0072298444238 + V2=  0.0001305388577 + V3= -0.0000041140337 + V4= -0.0000046897521 + R**-5
    KY(3)=                            B=  0.00000000000000 + D= -0.00000104399319 + G=  0.00000196201233 + R**-5
    KY(4)=                                                 C= -0.00000000000000 + F= -0.00000037709700 + R**-5
    KY(5)=                                                                        E= -0.00000000000000 + R**-5
```

```
SE :  B            SETTING   B = 0  .  C = 0  .  D = 0

          THERE ARE  3 SOLUTIONS (A1.A2.H)


(1)     (     0.15661487602136     0.00621306069123     0.22734949021634)
THE CONSTANTS IN THE CUMULANTS OF Y ARE
    KY(1)= 1.0 + M1= -0.0296719268531 + M2= -0.0005838226348 + M3=  0.0001713213630 + M4=  0.0002011485613 + R**-5
    KY(2)=       V1=  0.0507123788070 + V2=  0.0009156401729 + V3= -0.0001012947946 + V4= -0.0002303578928 + R**-5
    KY(3)=                            B= -0.0000000000000 + D= -0.0000000000000 + G=  0.0001439365398 4 + R**-5
    KY(4)=                                                C= -0.0000000000000 + F= -0.0000464067009 2 + R**-5
    KY(5)=                                                                    E=  0.0000032784452 9 + R**-5


(2)     (     0.07138240438957     0.00012903016282     0.36564412135018)
THE CONSTANTS IN THE CUMULANTS OF Y ARE
    KY(1)= 1.0 + M1= -0.0217504637774 + M2= -0.0004279605141 + M3= -0.0000023520057 + M4= -0.0000018557786 + R**-5
    KY(2)=       V1=  0.0272495620531 + V2=  0.0004920059815 + V3=  0.0000070775708 + V4=  0.0000030947268 + R**-5
    KY(3)=                            B=  0.0000000000000 + D=  0.0000000000000 + G= -0.0000020438943 9 + R**-5
    KY(4)=                                                C= -0.0000000000000 + F=  0.0000000092989 1 + R**-5
    KY(5)=                                                                    E=  0.0000006938758 9 + R**-5


(3)     (     0.08055746999164     0.00102212345083     0.16776281968626)
THE CONSTANTS IN THE CUMULANTS OF Y ARE
    KY(1)= 1.0 + M1= -0.0112621235938 + M2= -0.0002215927096 + M3=  0.0000213793105 + M4=  0.0000060412127 + R**-5
    KY(2)=       V1=  0.0073057206437 + V2=  0.0001319088450 + V3= -0.0000074158479 + V4= -0.0000052038362 + R**-5
    KY(3)=                            B=  0.0000000000000 + D= -0.0000000000000 + G=  0.0000023731917 4 + R**-5
    KY(4)=                                                C= -0.0000000000000 + F= -0.0000005436588 7 + R**-5
    KY(5)=                                                                    E=  0.0000000258250 2 + R**-5
```

```
ASE :  C          SPECIAL CASES WHEN A2 = 0

 C1 :  (HALDANE TRANSFORMATION B)   SETTING  B = 0  .  C = 0

            THE SOLUTION IS  (A1.A2.H)

                    (    0.07333333333333          0.0          0.43181818181818)

   THE CONSTANTS IN THE CUMULANTS OF Y ARE

      KY(1)= 1.0 + M1= -0.0263888888889 + M2= -0.0006575231481 + M3= -0.0000695372878 + M4= -0.0000255728359 + R**-5

      KY(2)=       V1=  0.0401111111111 + V2=  0.0007242283951 + V3=  0.0001069443930 + V4=  0.0000440357146 + R**-5

      KY(3)=                             B=  0.00000000000000 + D= -0.00006656711248 + G= -0.00004415416565 + R**-5

      KY(4)=                                                   C=  0.00000000000000 + F=  0.00002631000436 + R**-5

      KY(5)=                                                                          E= -0.00000707618599 + R**-5


 C2 :        SETTING  B = 0  .  D = 0

            THERE ARE  2 SOLUTIONS (A1.A2.H)

                (1)     (    0.07077542943435     0.0          0.41128344964240)

   THE CONSTANTS IN THE CUMULANTS OF Y ARE

      KY(1)= 1.0 + M1= -0.0242573023064 + M2= -0.0004514316133 + M3= -0.0000164047227 + M4= -0.0000070083433 + R**-5

      KY(2)=       V1=  0.0338928027946 + V2=  0.0004674624964 + V3=  0.0000228376539 + V4=  0.0000118291415 + R**-5

      KY(3)=                             B=  0.00000000000000 + D=  0.00000000000000 + G= -0.00001050963495 + R**-5

      KY(4)=                                                   C= -0.00001958880460 + F=  0.00000435439361 + R**-5

      KY(5)=                                                                          E=  0.00000000922973 + R**-5

                (2)     (   -0.02910876276768     0.0          2.43141317957108)

   THE CONSTANTS IN THE CUMULANTS OF Y ARE

      KY(1)= 1.0 + M1=  0.0589795245286 + M2= -0.0013570174608 + M3= -0.0000582229502 + M4= -0.0000084496536 + R**-5

      KY(2)=       V1= -0.2003664564647 + V2= -0.0305922052948 + V3=  0.0033779297438 + V4= -0.0001947391947 + R**-5

      KY(3)=                             B= -0.00000000000000 + D=  0.00000000000000 + G= -0.00086683338848 + R**-5

      KY(4)=                                                   C= -0.02741809218915 + F=  0.01596949232912 + R**-5

      KY(5)=                                                                          E=  0.00394103592583 + R**-5
```

```
C3 :   (HALCANE TRANSFORMATION A)   SETTING   B = 0   ;   A1 = 1 / K(1)

          THE SOLUTION IS  (A1.A2.H)

                    (    0.06666666666667          0.0          0.37500000000000)

  THE CONSTANTS IN THE CUMULANTS OF Y ARE

     KY(1)= 1.0 + M1= -0.0208333333333 + M2= -0.0001880787037 + M3=  0.0000347423161 + M4=  0.0000051812526 + R**-5
     KY(2)=       V1=  0.0250000000000 + V2=  0.0001736111111 + V3= -0.0000503150720 + V4= -0.0000084069439 + R**-5
     KY(3)=                            B=  0.00000000000000 + D=  0.00004918981481 + G=  0.00000939790702 + R**-5
     KY(4)=                                                 C= -0.00002777777778 + F= -0.00000745081019 + R**-5
     KY(5)=                                                                        E=  0.00000321180556 + R**-5
```

CASE : D

WE USE BROWNS METHOD OF SOLVING

THE FOLLOWING NON-LINEAR SIMULTANEOUS EQUATIONS

(1) B + D + G = 0
(2) C + F = 0
(3) E = 0

USING 11 INITIAL GUESSES, WE OBTAIN 2 SOLUTIONS

(1) (0.07157543779526 0.00907715377302 0.38978077716366)

THE CONSTANTS IN THE CUMULANTS OF Y ARE

KY(1)= 1.0 + M1= -0.0231675711843 + M2= -0.0004690478943 + M3= -0.0000148434455 + M4= -0.0000058936533 + R**-5
KY(2)= V1= 0.0311335649104 + V2= 0.0005270732129 + V3= 0.0000230220382 + V4= 0.0000097512813 + R**-5
KY(3)= B= 0.00001520009331 + D= -0.00000681191678 + G= -0.00000838816653 + R**-5
KY(4)= C= -0.00000341563778 + F= 0.00000341563778 + R**-5
KY(5)= E= -0.00000000000000 + R**-5

(2) (0.08074840521623 0.00190274418232 0.17596968559476)

THE CONSTANTS IN THE CUMULANTS OF Y ARE

KY(1)= 1.0 + M1= -0.0118513491643 + M2= -0.0002570077593 + M3= 0.0000182018572 + M4= 0.0000057724625 + R**-5
KY(2)= V1= 0.0020761358376 + V2= 0.0001613099037 + V3= -0.0000044356593 + V4= -0.0000050231429 + R**-5
KY(3)= B= -0.00000049860120 + D= -0.00000177985629 + G= 0.00000227845750 + R**-5
KY(4)= C= 0.00000048471301 + F= -0.00000048471301 + R**-5
KY(5)= E= 0.00000000000000 + R**-5