

ANALYTICAL PROPERTIES
OF CERTAIN PROBABILITY DISTRIBUTIONS

Thesis presented to the University of Edinburgh
for the degree of Doctor of Philosophy

by W.F. Kibble

1938



CONTENTS

	Preface	
I	Orthogonal polynomials and moment generating functions	1
II	Distribution of the variances in a normally correlated two-variate distribution	22
III	Expression of the series of Laguerre polynomials as a Bessel function	27
IV	Expressions for Bessel functions involving incomplete Gamma functions	33
V	Two-variate distributions of different types in the two variates	62
VI	Distribution functions in any number of variates expressed in orthogonal polynomials	68
VII	Differential equations and difference equations	87
VIII	Selection	94
IX	Curve fitting	102
X	Tables of the psi-log-difference function	117
	References to literature	131
	Index	132

PREFACE

This research has been carried out on Dr Aitken's suggestion. The first chapter is largely a recapitulation of known results which I have learnt from Dr Aitken, here arranged for convenience of reference later in the thesis.

The second chapter is the application of these methods to the deduction of a two-variate Gamma type distribution. Dr Aitken pointed out that the variances in a normally correlated two-variate distribution would give the required distribution, and chapter II is just the carrying out of that suggestion. He also directed me to the papers by Hardy and Wishart and Bartlett which give rise to chapter III. Those two chapters form the centre or core of the thesis from which the other research radiates in three main directions, of varying interest from the points of view of pure mathematics and of statistical applications.

The first is chapter IV which is purely of mathematical interest. It contains the most substantial single piece of research in the thesis. It was perhaps fortunate that on my first searching Watson's Theory of Bessel Functions for theorems involving incomplete Gamma functions, I did not find the paragraph on Hadamard's paper. The result given by Watson would have been sufficient for the purposes of chapter III, but not being satisfied with my own deduction of it, I sought to find it as a special case of a more general theorem, and in doing so have been led to discover a more general result, and incidental results which may be of great interest in themselves.

Secondly, the generalisations in chapters V and VI are of interest as giving forms for statistical distributions, but they will also be of interest for pure mathematics, giving, for example, a generalisation to any number of variables of Mehler's theorem (1866) which has been discussed by Hardy and Watson and others.

The third branch of the thesis is concerned with the actual method of fitting a distribution function such as that discussed in chapters II and III. This is discussed in Chapter IX, and tables to make easy the fitting by the method suggested there of a type III curve are given in chapter X

Other parts of the thesis are chapters VII and VIII, of which

of which VII is concerned with an attempt, not so far successful, to extend the theory to all Pearson's types, instead of type III only. (Similar attempts must have been made before: see Romanovsky, Biometrika, Vol XVI, parts I and II, p. 106) And chapter VIII applies a method which I learnt from Dr Aitken in application to a normal distribution to the distribution of chapter II, and deduces simple formulæ which may be of great practical significance in dealing with problems of selection in distributions in which the coefficient of variation is important and is not small.

Publication

I propose to publish the substance of chapter IV in the Journal of the Indian Mathematical Society. Since ~~xxxx~~ chapter II is a sequel to papers by J.T. Campbell and by Aitken and Gonin, it should probably be published after the latter. Parts of chapter III, V and VI would then follow, and VIII also. Chapter IX might be published, unless it should prove to have been done already by someone else, as soon as the result quoted in § 22.2 in that chapter has been published.

Further lines of research

There are two or three points in the present line of research which remain to be cleared up. On page 61 at the end of chapter IV a constant K is left undetermined. Although its value does not affect the main theorem, it will be of great interest to find it - I think it should be possible in a few weeks. At the end of chapter IX the important question of the best statistic to use to estimate the correlation is left undetermined. A straightforward application of the method of maximum likelihood to the Bessel function form in chapter III should be possible. The methods in the chapter on selection need to be applied to the distributions of V and VI.

Other lines of research include a further attempt to find by the method of chapter VII a continuous distribution to correspond to the double hypergeometric distribution. There must be some partial difference equation satisfied by the frequencies in that distribution, and although so far I have not found it, the work I have done on it will be a help to finding it. We might then arrive at a differential equation of which the Bessel differential equation would be a special case. Another problem is that given in § 7.2, chapter III, p. 31, the integration of a certain ~~xxxxxxx~~ function of the elements of a symmetric matrix with respect to all the elements off the diagonal, throughout the region in which the matrix is positive definite. I suggested this problem to Dr Ledermann when he gave me a copy of his paper in which the function concerned was derived. I carried out the integration when there are only two ~~xxxxxxx~~ variates, in chapter III, obtaining a series of incomplete Gamma functions. For three variates, any such direct integration appears to give

too complicated a result, and one may expect any solution of this problem to be based on general matrix theory, and I do not myself see at present any line of approach to a solution.

The methods of curve fitting suggested in chapter IX it would be well to test by application to a variety of problems. That can best be done by those actually engaged in the practical business of studying naturally occurring distributions.

I wish to thank both Professor Whittaker and Dr Aitken, not only for the suggestions mentioned in this preface, but also for advice and ~~encouragement~~ encouragement during the years in which I have been engaged on this work. It is a great stimulus to associate with them.

Madras, September 1938

Chapter I

ORTHOGONAL POLYNOMIALS , AND MOMENT GENERATING FUNCTIONS

§1.11 Definitions The orthogonal polynomial of degree n , for a distribution function $\varphi(x)$, is ~~xxxxxxx~~ is a polynomial $P_n(x)$, such that

$$\left. \begin{aligned} \int P_r(x) P_s(x) \varphi(x) dx &= 0, \quad r \neq s, \\ &\neq 0, \quad r = s, \end{aligned} \right\} (1)$$

or, if the function $\varphi(x)$ is that of a discontinuous distribution,

$$\left. \begin{aligned} \sum P_r(x) P_s(x) \varphi(x) &= 0, \quad r \neq s, \\ &\neq 0, \quad r = s. \end{aligned} \right\} (2)$$

Alternatively, the functions $P_n(x)$ ~~by~~ can be defined by

$$\int P_n(x) x^r \varphi(x) dx = 0, \quad \text{or} \quad \sum P_n(x) x^r \varphi(x) = 0, \\ r = 0, 1, 2, 3, \dots, n-1.$$

§1.111 These conditions determine the ratios of the coefficients in $P_n(x)$. If the values of the coefficients are so chosen that

$$\int \{P_n(x)\}^2 \varphi(x) dx = 1, \quad (3)$$

the polynomial is called a normalised orthogonal polynomial.

§1.12 The moment generating function for a given distribution function $\varphi(x)$ is defined as

$$\zeta(\alpha) = \int \varphi(x) e^{\alpha x} dx. \quad (4)$$

If $\zeta(\alpha)$ is expanded, the coefficient of $\frac{\alpha^n}{n!}$ is $\int x^n \varphi(x) dx$,

the n^{th} moment of the distribution function $\varphi(x)$.

Thus if the moments are bounded, the integral exists and can be expressed as a series in powers of α in which the coefficient of $\frac{\alpha^n}{n!}$ is the n^{th} moment.

§ 1.13 When $\varphi(x)$ is a discontinuous frequency distribution function, and $\varphi(x)$ exists only for integral values of x , it is convenient to use the same notation $\zeta(\alpha)$ for the factorial moment generating function defined by

$$\zeta(\alpha) = \sum \varphi(x) (1+\alpha)^x \quad (5)$$

in which the coefficient of $\frac{\alpha^n}{n!}$ is the factorial moment

$$\sum \varphi(x) \cdot x^{(n)}, \quad \text{where} \quad x^{(n)} = x(x-1)(x-2)\dots(x-n+1).$$

§ 1.2 It has been pointed out by Dr Aitken that the two-variate distribution functions which correspond to certain well known one-variate distribution functions, can all be expressed in forms which have the one-variate functions as factors, and in which the co-factor is a series bilinear in the respective orthogonal polynomials.

1.21 Examples are (i) the Normal Frequency Function

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (6)$$

for which the orthogonal polynomials are Hermite polynomials :

$$H_n(x) = \frac{\left(-\frac{d}{dx}\right)^n e^{-\frac{1}{2}x^2}}{e^{-\frac{1}{2}x^2}} = x^n - \frac{n(n-1)}{2} x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2 \cdot 4} x^{n-4} - \dots \quad (7)$$

while the corresponding distribution function for two variates is the Normal correlation function

$$\Phi(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{-(x^2 - 2\rho xy + y^2)}{2(1-\rho^2)}\right\}, \quad (8)$$

which can be expressed in the form¹⁾

$$\begin{aligned} \Phi(x, y) &= \frac{1}{2\pi} e^{-\frac{1}{2}(x^2 + y^2)} \left\{ 1 + \rho H_1(x) H_1(y) + \frac{\rho^2}{2!} H_2(x) H_2(y) + \dots \right\} \\ &= \Phi(x) \Phi(y) \left\{ 1 + \rho H_1(x) H_1(y) + \frac{\rho^2}{2!} H_2(x) H_2(y) + \dots \right\} \quad (9) \end{aligned}$$

(22) (ii) the binomial distribution function,

$$\Phi(x) = \binom{n}{x} p^x q^{n-x}, \quad \text{where } \binom{n}{x} = \frac{n(n-1)\dots(n-x+1)}{x!} \quad (10)$$

for which Aitken and Gonin²⁾ have shown the orthogonal polynomials are $G_r(x, b)$ where

$$G_r(x, b) \Phi(x) = (-q)^r \Delta^r x^{(r)} \Phi(x), \quad (11)$$

and for which the two-variate distribution function is

$$\begin{aligned} \Phi(x, y; s) &= \Phi(x, b) \Phi(y, b') \left\{ 1 + \frac{s}{n b q b' q'} G_1(x, b) G_1(y, b') + \right. \\ &\quad \left. + \frac{s^2}{n(n-1)(b q b' q')^2} G_2(x, b) G_2(y, b') + \dots \right\} \\ \text{where } \Phi(x, b) &= \binom{n}{x} p^x q^{n-x} \quad \text{and} \quad \Phi(y, b') = \binom{n}{y} p'^y q'^{n-y} \quad (12) \end{aligned}$$

¹⁾ A.C. Aitken and H.T. Gonin, On fourfold sampling with and without replacement

Dr Aitken kindly gave me a manuscript copy of the paper.

²⁾ G.Mehler, J. für Math., lxvi (1866) p.174.

A limiting form of this is

1.23 (iii) the Poisson distribution function

$$\varphi(x, m) = \frac{e^{-m} m^x}{x!} \quad (13)$$

for which the Charlier polynomials are orthogonal defined by

$$K_r(x) \varphi(x) = (-\nabla)^r \varphi(x), \text{ where } \nabla f(x) = f(x) - f(x-1) \quad (14)$$

and for which the two-variate function is

$$\begin{aligned} \varphi(x, y) &= e^{-\bar{m}-m-m'} \frac{(m-\bar{m})^x (m'-\bar{m})^y}{x! y!} {}_0F_2 \left(-x, -y; \frac{\bar{m}}{(m-\bar{m})(m'-\bar{m})} \right) \\ &= \varphi(x, m) \varphi(y, m') \left\{ 1 + \bar{m} K_1(x) K_1(y) + \frac{\bar{m}^2}{2!} K_2(x) K_2(y) + \dots \right\} \end{aligned} \quad (15)$$

1.24

(iv)

And a generalisation is the hypergeometric distribution

$$\varphi(x) = \binom{n}{x} (Np)^x (Nq)^{n-x} / N^n \quad (16)$$

for which the orthogonal polynomials are $U_r(x)$, where

$$U_r(x) \varphi(x) = (-\Delta)^r \left\{ x^{(r)} (Nq - n + x)^{(r)} \varphi(x) \right\} / (N+r+1)^r \quad (17)$$

$$\begin{aligned} \varphi(x, y; d) &= \varphi(x, y; 0) \left\{ 1 + a_1 \frac{N^{(2)} N^2}{n(N-n) Np Np' Nq Nq'} U_1(x, p) U_1(y, p') + \right. \\ &\quad \left. + a_2 \frac{N^{(4)} \{N-1\}^{(2)2}}{n^{(2)} (N-n)^{(2)} (Np)^{(2)} (Np')^{(2)} (Nq)^{(2)} (Nq')^{(2)}} U_2(x, p) U_2(y, p') + \dots \right\} \end{aligned} \quad (18)$$

and the correlation function is

where a_1, a_2, \dots , are functions of d , given in Aitken and Gouin's paper; $\frac{a_r}{N^r} \rightarrow d^r / r!$ as $N \rightarrow \infty$.

1.31 Other distribution functions which have well-known polynomials as orthogonal polynomials are

(v) the Type III curve of Pearson, or Gamma-type curve

$$\varphi(x) = x^{p-1} e^{-x} / \Gamma(p) \quad (19)$$

generalised

for which the Laguerre polynomials are orthogonal :

$$\varphi(x) L_r(x; p) = \left(-\frac{d}{dx}\right)^r x^r \varphi(x); \quad (20)$$

They are sometimes called Sonine polynomials.

1.32

and (vi) Type I

$$\varphi(x) = \frac{(1-x)^\alpha (1+x)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)} \quad (21)$$

for which Jacobi polynomials are orthogonal :

$$(1-x)^\alpha (1+x)^\beta P_r(x; \alpha, \beta) = \frac{(-1)^r}{2^r r!} \left(\frac{d}{dx}\right)^r \left\{ (1-x)^{\alpha+r} (1+x)^{\beta+r} \right\} \quad (22)$$

1.4

Comparison of the form of the orthogonal polynomials and the moment generating functions

A polynomial of degree n can always be expressed as $f(D)x^n$ or $f(\Delta)x^{(n)}$ where D denotes $\frac{d}{dx}$ and Δ denotes the difference operator : $\Delta f_x = f_{x+1} - f_x$. Dr Aitken points out that when the orthogonal polynomials of various distribution functions are expressed in this way, there is such similarity in form between the function $f(D)$ or $f(\Delta)$ and the moment (or factorial moment) generating function as to suggest that a general method exists for deriving orthogonal polynomials from the generating functions. A theorem showing that this can in fact be done is given below, § 2.2.

§ 1.41

(i) For the Normal Frequency function, $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$,

the moment generating function is

$$\varphi(\alpha) = e^{\frac{1}{2}\alpha^2} \quad (23)$$

and the Hermite polynomials ~~are~~ can be written

$$H_n(x) = e^{-\frac{1}{2}D^2} x^n \quad (24)$$

If $H_n(x)$ is written as $a_{n0}x^n - a_{n1}x^{n-2} + a_{n2}x^{n-4} - a_{n3}x^{n-6} + \dots$,
the coefficients a_{rs} satisfy the recurrence relations

$$a_{r+1,s} = r a_{r-1,s-1} + a_{r,s} \quad (25)$$

and

$H_0(x) =$		1
$H_1(x) =$		x
$H_2(x) =$		$x^2 - 1$
$H_3(x) =$		$x^3 - 3x$
$H_4(x) =$		$x^4 - 6x^2 + 3$
$H_5(x) =$		$x^5 - 10x^3 + 15x$
$H_6(x) =$		$x^6 - 15x^4 + 45x^2 - 15$
$H_7(x) =$		$x^7 - 21x^5 + 105x^3 - 105x$
$H_8(x) =$		$x^8 - 28x^6 + 210x^4 - 420x^2 + 105$
$H_9(x) =$		$x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x$
$H_{10}(x) =$		$x^{10} - 45x^8 + 630x^6 - 3150x^4$ $+ 4725x^2 - 945$
$H_{11}(x) =$		$x^{11} - 55x^9 + 990x^7 - 6930x^5$ $+ 17325x^3 - 10395x$
$H_{12}(x) =$		$x^{12} - 66x^{10} + 1485x^8 - 51975x^6$ $+ 13860x^4 - 62370x^2 + 10395$
$H_{13}(x) =$		$x^{13} - 78x^{11} - 25740x^9 - 270270x^7$ $+ 2145x^5 + 135135x^3 + 135135x$
$H_{14}(x) =$		$x^{14} - 91x^{12} - 45045x^{10} - 945945x^8$ $+ 3003x^6 + 315315x^4 + 945945x^2$ <div style="text-align: right; font-size: small;">$\frac{135135}{-945945}$</div>

The coefficients in these polynomials were calculated from the recurrence relation ⁽²⁵⁾ written above them, those in the last half dozen being calculated with the Archimedes machine. As that is a twelve figure machine it was possible to calculate two or three coefficients with the same set of operations, so long as the coefficients had three or four figures, putting one in the thousands place and one in the millions place, or something of that sort. (The multiplier r in the recurrence relation used, which multiplies coefficients in the last but one of the polynomials already determined, is the same for all the coefficients in the polynomial required.) The whole calculation was checked by calculating the coefficients in the 14th polynomial from

$$a_{2s} = {}^*C_{2s} \cdot 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2s-1). \quad (26)$$

1.42

(ii) For the binomial distribution function,

$$\varphi(x) = \binom{n}{x} p^x q^{n-x} \text{ the factorial moment generating function is}$$

$$G(\alpha) = \sum_{x=0}^n (1+\alpha)^x \varphi(x) = (1+p\alpha)^n; \quad (27)$$

and the orthogonal polynomials can be written as

$$(1+p\Delta)^{-n+r-1} x^{(r)} \quad (28)$$

1.43

(iii) For the Poisson function $\varphi(x) = e^{-m} \frac{m^x}{x!}$

the factorial moment generating function is

$$G(\alpha) = e^{m\alpha} ; \quad (29)$$

and the Charlier polynomials can be written as

$$K_r(x; m) = e^{-m\Delta} x^{(r)} \quad (30)$$

1.44

(iv) For the hypergeometric distribution

$$\varphi(x) = \binom{n}{x} (Np)^{(x)} (Nq)^{(n-x)} / N^{(n)} ; \quad (31)$$

the factorial moment generating function is

$$F(-n, -Np; -N; -\alpha) ; \quad (32)$$

while the polynomials defined by (17) can be expressed with a hypergeometric operator :

$$F(n-r+1, Np-r+1; N-2r+2; -\Delta) x^{(r)} \quad (33)$$

1.45

(v) For Type III or Gamma-type,

$$\varphi(x) = x^{p-1} e^{-x} / \Gamma(p) ; \quad G(\alpha) = (1-\alpha)^{-p} ; \quad (34)$$

and the generalised Laguerre polynomials can be written as

$$L_r(x; p) = (1-D)^{p+r-1} x^r \quad (35)$$

(The simple Laguerre polynomials correspond to the special case in which $\frac{p}{m} = 0$, and $\varphi(x) = e^{-x}$ between 0 and ∞ .)

The general Laguerre polynomial is

$$L_r(x, p) = x^r - (r+p-1)x^{r-1} + \binom{r+p-1}{2}r(r-1)x^{r-2} - \dots \\ + (-1)^r (r+p-1)(r+p-2)\dots p \\ = x^r - r(r+p-1)x^{r-1} + \binom{r}{2}(r+p-1)^{(2)}x^{r-2} - \dots + (-1)^r (r+p-1)^{(r)}r!$$

1.451

$$\int_0^\infty \{L_r(x, p)\}^2 \varphi(x) dx = \int_0^\infty x^r L_r(x, p) \varphi(x) dx = \int_0^\infty \frac{x^{r+p-1} e^{-x}}{\Gamma(p)} dx \\ = r! \frac{\Gamma(r+p)}{\Gamma(p)} = (r+p-1)^{(r)} r! \quad (37)$$

1.452

$$L_r(x, \frac{1}{2}) = x^r - \frac{2r(2r-1)}{2} \cdot \frac{x^{r-1}}{2} + \frac{2r(2r-1)(2r-2)(2r-3)}{2^2 \cdot 2!} \cdot \frac{x^{r-2}}{2^2} - \dots \\ = \frac{1}{2^r} H_{2r} \left\{ \left(\frac{1}{2} x \right)^2 \right\}$$

$$\text{Or, } 2^r L_r \left(\frac{1}{2} x^2, \frac{1}{2} \right) = H_{2r}(x). \quad (38)$$

Similarly,

$$2^{r \times} L_r \left(\frac{1}{2} x^2, \frac{3}{2} \right) = H_{2r+1}(x). \quad (39)$$

The definition of the generalised Laguerre polynomial in the form above in which the coefficient of x^r is 1, (whereas Hille and others use a form in which it is $\frac{1}{r!}$), is a generalisation of Berger's definition of the simple Laguerre polynomial, given by Wigert, Arkiv för Math., etc., 15, No 25 (1921).

1.5 Orthogonal polynomials for the others of Pearson's curves have been given By Romanovsky ¹⁾. For a frequency curve for which the higher moments are infinite, as in Pearson's types IV, V and VI, what is meant when we say that polynomials $p_n(x)$ are orthogonal for a frequency function $\Phi(x)$ is that $\int \Phi(x)p_m(x) x^n dx = 0$ when $m \neq n$, provided the integral converges. It is interesting that for Type IV Romanovsky's polynomials

$$\varphi(x) R_n(x) = \frac{d^n}{dx^n} \left\{ (1+x^2)^n \varphi(x) \right\} ; \varphi(x) = c(1+x^2)^{-\nu} e^{\nu \tan^{-1} x} \quad (1)$$

give, when $2m$ is a whole number, a finite set of polynomials, such that all the integrals $\int \Phi(x)R_m(x)R_n(x)dx$ converge,

To prove this, let $\Phi_r = (1+x^2)^r \exp(\nu \tan^{-1} x)$,

$$\left. \begin{aligned} D \Phi_r &= (2rx + \nu) \Phi_{r-1} \\ D^2 \Phi_r &= 2r \Phi_{r-1} + (2rx + \nu) \{ 2(r-1)x + \nu \} \Phi_{r-2} \end{aligned} \right\} (2)$$

These are valid for all values of r .

Now suppose that $D^{2r-1} \Phi_{r-1} = k_{r-1} \Phi_{-r}$ (3)

This is true when $r = 1$; $k_0 = \nu$. We shall show that if it is true for any value r , then it is also true for $r + 1$.

For $D^{2r+1} \Phi_r = D^{2r+1} \{ (1+x^2) \Phi_{r-1} \}$

$$\begin{aligned} &= (1+x^2) D^{2r+1} \Phi_{r-1} + (2r+1) 2x D^{2r} \Phi_{r-1} + (2r+1) 2r D^{2r-1} \Phi_{r-1} \\ &= \{ (1+x^2) D^2 + (2r+1) 2x D + (2r+1) 2r \} \Phi_r k_{r-1} \end{aligned}$$

by equation (3)

$$\begin{aligned} &= k_{r-1} \left[(1+x^2) \{ 2r(2r+1) \Phi_{r-1} - 2(2r+1) \nu x \Phi_{r-2} + (\nu^2 - 4r(r+1)) \Phi_{r-2} \} \right. \\ &\quad \left. + (2r+1) 2x (-2rx + \nu) \Phi_{r-1} + (2r+1) 2r \Phi_r \right] \end{aligned}$$

¹⁾ Comptes Rendus, vol 188, (1929).

$$\begin{aligned}
&= k_{r-1} \{ \nu^2 - 4\nu(r+1) + 4r(2r+1) \} \varphi_{r-1} \\
&= k_{r-1} (\nu^2 + 4r^2) \varphi_{r-1}
\end{aligned}
\tag{4}$$

So that equation (3) holds for all positive integral r , and

$$\begin{aligned}
k_r &= k_{r-1} (\nu^2 + 4r^2) \\
&= \nu (\nu^2 + 4)(\nu^2 + 16)(\nu^2 + 36) \dots (\nu^2 + 4r^2)
\end{aligned}$$

Now differentiate the equation

$$D^{2r+1} \varphi_r = k_r \varphi_{r-1}
\tag{5}$$

n times :

$$D^{n+2r+1} \left\{ (1+x^2)^r e^{\nu \tan^{-1} x} \right\} = k_r D^n \left\{ (1+x^2)^{r-1} e^{\nu \tan^{-1} x} \right\}
\tag{6}$$

Put $r + 1 = m - n$, $2r + n + 1 = 2m - n - 1$.

Equation (6) gives

$$D^{2m-n-1} \left\{ (1+x^2)^{m-n-1} \varphi(x) \right\} = k_r D^n \left\{ (1+x^2)^n \varphi(x) \right\}$$

where $\varphi(x) = y_0 (1+x^2)^{-m} e^{\nu \tan^{-1} x}$

Thus the polynomials $R_n(x)$ satisfy the relations

$$R_{2m-n-1}(x) \equiv \nu (\nu^2 + 4)(\nu^2 + 16) \dots (\nu^2 + 4(m-n-1)^2) R_n(x).$$

The polynomials from m to $2m - 1$, if $2m$ is even, repeat in the reverse order, with constant multipliers, those from 0 to $m - 1$. If $2m$ is odd, those from $m + \frac{1}{2}$ to $2m - 1$ similarly repeat those from 0 to $m - \frac{3}{2}$.

For the special case in which $\nu = 0$, equation (5) is still true, but with $k_r = 0$, and hence we deduce that

$$R_s(x) \equiv 0, \quad \begin{aligned}
m + \frac{1}{2} \leq s \leq \frac{2m-1}{2}, & \quad 2m \text{ odd} \\
m \leq s \leq 2m-1, & \quad 2m \text{ even.}
\end{aligned}$$

In this case also, then, though polynomials of higher orders occur, special significance attaches to a set of m or $m + \frac{1}{2}$

~~gix~~ polynomials, giving convergent integrals.

This suggests that if a series bilinear in these polynomials is sought, to represent a frequency distribution in two variates, it may be possible when $2m$ is a whole number to find a finite series of that sort.

§2

Theorems connecting moment generating functions with the respective orthogonal polynomials of a frequency function

2.1 If $G(\alpha)$ is the moment generating function of any frequency function $\varphi(x)$, then the m.g.f. of $\frac{d}{dx} \varphi(x)$ is $\int \frac{d}{dx} \varphi(x) e^{\alpha x} dx = [\varphi(x) e^{\alpha x}] - \int \varphi(x) \alpha e^{\alpha x} dx,$

and if $\varphi(x) \cdot e^{\alpha x} = 0$ at the limits of the distribution, this is $-\alpha G(\alpha)$

Further, the moment generating function of $(\frac{d}{dx})^n \varphi(x)$ is $(-\alpha)^n G(\alpha)$ if $\varphi(x)$ and each of its first $n - 1$ derivatives multiplied by $e^{\alpha x}$ vanishes at the limits. A condition such as this is often described as the condition that the curve shall have 'high contact' at the ends. The condition is satisfied by the Normal Frequency Curve, and direct modifications of it, *such as Charlier's* by Pearson's Type III (for a certain range of values of α ; we need only consider $G(\alpha)$ within such a range), but for Type I it is in general only satisfied for values of n less than *either index*.

For a curve (such as Type IV) for which the higher moments are infinite, the function $G(\alpha)$ cannot exist and be finite.

2.2 Since $\frac{d}{d\alpha} \int \varphi(x) e^{\alpha x} dx = \int x \varphi(x) e^{\alpha x} dx,$

$(\frac{d}{d\alpha})^n G(\alpha)$ is the moment generating function of $x^n \varphi(x)$, if the integrals converge uniformly;

So that $P_n(\frac{d}{d\alpha})G(\alpha)$ is the m.g.f. of $P_n(x)\varphi(x)$, where P_n is any polynomial. The condition that $P_n(x)$ should be an orthogonal polynomial with respect to $\varphi(x)$ is

$$\int x^s P_n(x) \varphi(x) dx = 0, \quad s = 0, 1, 2, \dots, n-1,$$

which is the condition that $\left(\frac{d}{d\alpha}\right)^s P_n \left(\frac{d}{d\alpha}\right) G(\alpha) = 0$, when $\alpha = 0$.

Hence the following THEOREM is proved :

I For a frequency function $\varphi(x)$ for which the moment generating function is $G(\alpha)$, the condition that a polynomial of degree n $P_n(x)$ should be one of the orthogonal polynomials for $\varphi(x)$ is that $P_n \left(\frac{d}{d\alpha}\right) G(\alpha)$

should contain no terms in $\alpha^{n-1}, \alpha^{n-2}, \dots, \alpha^1, \alpha^0$.

The condition that $P_n(x)$ should be a normalised orthogonal polynomial is that (coefficient of x^n in $P_n(x)$) (coefficient of x^n in $P_n(x)$) = $\frac{1}{n!}$

This theorem gives a fairly simple method of finding the orthogonal polynomial of any given degree, when the moment generating function is known,

2.3 In order to consider symbolic expressions of the type $f(D)x^n$ for the orthogonal polynomials, we introduce a factor e^{-x^2/σ^2} (which has the effect of making all the moments finite, and yet σ can be chosen so large that $\varphi(x)e^{-x^2/\sigma^2}$ does not differ significantly from $\varphi(x)$ for any values of x for which $\varphi(x)$ is significantly different from zero).

The m.g.f. of e^{-x^2/σ^2} is $\frac{\sigma}{\sqrt{\pi}} e^{i\alpha^2 \sigma^2}$ and therefore that of $x^n e^{-x^2/\sigma^2}$ is $\frac{\sigma}{\sqrt{\pi}} \left(\frac{\partial}{\partial \alpha}\right)^n (e^{i\alpha^2 \sigma^2}) = \frac{\sigma^{n+1}}{\sqrt{\pi}} \hat{H}_n(\alpha\sigma) e^{i\alpha^2 \sigma^2}$

where $\hat{H}_\nu(\cdot)$ denotes the modified Hermite polynomial, which can be described as the Hermite polynomial with all the signs made positive. Thus the m.g.f. of

$$f\left(\frac{\partial}{\partial x}\right)\{x^\nu e^{-i^2 x^2/\sigma^2}\} \text{ is } f(-\alpha) \frac{\sigma^{\nu+1}}{\sqrt{\pi}} \hat{H}_\nu(\alpha\sigma) e^{i^2 \alpha^2 \sigma^2}$$

provided that $f\left(\frac{\partial}{\partial x}\right)\{x^\nu e^{-i^2 x^2/\sigma^2}\}$ is convergent, as it certainly will be if f denotes a polynomial. Now, $f\left(\frac{\partial}{\partial x}\right)\{x^\nu e^{-i^2 x^2/\sigma^2}\}$ can be written symbolically, expanding by Leibniz's theorem :

$$e^{-i^2 x^2/\sigma^2} \left[f(D) x^\nu - \frac{1}{\sigma} H_1\left(\frac{x}{\sigma}\right) f'(D) x^\nu + \frac{1}{2! \sigma^2} H_2\left(\frac{x}{\sigma}\right) f''(D) x^\nu - \dots \right]$$

which, when σ is large, differs by a small quantity from $f(D) x^\nu$. To consider whether $f(D) x^\nu$ can be an orthogonal polynomial for $\varphi(x)$, consider the m.g.f. of $\varphi(x) f(D) x^\nu$ which we may regard as a limiting form of

$$\varphi(x) f(D) \{x^\nu e^{-i^2 x^2/\sigma^2}\} \text{ of which the m.g.f. is } \varphi\left(\frac{\partial}{\partial \alpha}\right) \left\{ f(-\alpha) \frac{\sigma^{\nu+1}}{\sqrt{\pi}} \hat{H}_\nu(\alpha\sigma) e^{i^2 \alpha^2 \sigma^2} \right\}$$

Now the condition that $f(D) x^\nu$ should be an orthogonal polynomial is that the m.g.f. of $\varphi(x) f(D) x^\nu$ should have zero coefficients for all terms up to but not including that in α^ν . Thus we require to find the condition that

$$\varphi\left(\frac{\partial}{\partial \alpha}\right) \left\{ f(-\alpha) \frac{\sigma^{\nu+1}}{\sqrt{\pi}} \hat{H}_\nu(\alpha\sigma) e^{i^2 \alpha^2 \sigma^2} \right\} \text{ should, when } \sigma \text{ is large, have a}$$

form which can be expressed in a series of positive powers of α , beginning with α^ν (or with small coefficients for lower powers than the α^ν). This m.g.f. can be written

$$\text{symbolically } \frac{\sigma^{\nu+1}}{\sqrt{\pi}} e^{i^2 \alpha^2 \sigma^2} \left\{ \hat{H}_\nu(\alpha\sigma) \varphi\left(\frac{\partial}{\partial \alpha}\right) f(-\alpha) + \sigma \hat{H}_{\nu+1}(\alpha\sigma) \varphi\left(\frac{\partial}{\partial \alpha}\right) f(-\alpha) + \dots \right\}$$

which if each \hat{H} is represented by its leading term can

be written $e^{i\alpha^2\sigma^2} \alpha^r \left[\exp\left\{ \alpha\sigma^2 \frac{\partial}{\partial\alpha} \right\} \varphi\left(\frac{\partial}{\partial\alpha}\right) \right] f(-\alpha) \cdot \frac{\sigma^{2r+1}}{\sqrt{\pi}}$

Analogous arguments show that the m.g.f. of $\varphi(x)$ can be represented by $\frac{\sigma}{\sqrt{\pi}} \varphi\left(\frac{\partial}{\partial\alpha}\right) \left\{ e^{i\alpha^2\sigma^2} \right\}$,

when σ is large. This expression is always finite, even if the higher moments of $\varphi(x)$ are infinite, but as $\sigma \rightarrow \infty$ the coefficient of $\frac{\alpha^r}{r!}$ tends to the r^{th} moment of $\varphi(x)$ if that moment is finite.

3.1

Another THEOREM which is included here because

it applies to any form of distribution function is :

II If $x = x_1 + x_2 + x_3 + \dots + x_n$, and if the ^{$x_1, x_2, x_3, \dots, x_n$ being independent variates} distribution function for x_i has a moment generating function $G_i(\alpha)$, then the moment generating function for the distribution function of x is

$$G(\alpha) = G_1(\alpha) \cdot G_2(\alpha) \cdot G_3(\alpha) \cdot \dots \cdot G_n(\alpha).$$

Proof :

$$G(\alpha) = \int \Phi(x) e^{\alpha x} dx,$$

where $\Phi(x) = \int \dots \int_{(n-1)} \phi_1(x_1) \phi_2(x_2) \dots \phi_n(x_n) dx_1 dx_2 \dots dx_{n-1}$,

and the integration is over $x_1 + x_2 + \dots + x_n = x$.

$$\begin{aligned} \text{Therefore } G(\alpha) &= \int \dots \int_{(n)} \phi_1(x_1) \phi_2(x_2) \dots \phi_n(x_n) e^{\alpha(x_1 + x_2 + \dots + x_n)} dx_1 dx_2 \dots dx_n \\ &= \int \phi_1(x_1) e^{\alpha x_1} dx_1 \int \phi_2(x_2) e^{\alpha x_2} dx_2 \dots \int \phi_n(x_n) e^{\alpha x_n} dx_n \end{aligned}$$

Corollary : If the m.g.f. of a certain distribution is

$G(\alpha)$, then the distribution of the sum of a sample of n is $\{G(\alpha)\}^n$.

III The THEOREM above can readily be extended to two-variate distributions, and similarly to any number of variates. The proof for two variates is given :

$$G(\alpha, \beta) = \iint \Phi(x, y) e^{\alpha x + \beta y} dx dy,$$

where $\Phi(x, y) = \int \dots \int_{(2n-1)} \phi_1(x_1, y_1) \phi_2(x_2, y_2) \dots \phi_n(x_n, y_n) dx_1 dx_2 \dots dx_{n-1} dy_1 dy_2 \dots dy_{n-1}$,

and $x = x_1 + x_2 + \dots + x_n$, $y = y_1 + y_2 + \dots + y_n$, so that

$$G(\alpha, \beta) = \iint \phi_1(x_1, y_1) e^{\alpha x_1 + \beta y_1} dx_1 dy_1 \dots \iint \phi_n(x_n, y_n) e^{\alpha x_n + \beta y_n} dx_n dy_n$$

3.2

Series bilinear in orthogonal polynomials occur in expressions for two-variate distributions, and certain simple theorems for a general expression of that sort can be given.

3.21 THEOREM IV If the distribution function for two variates x and y is

$$\varphi(x) \psi(y) \{1 + a_1 P_1(x) P_1(y) + a_2 P_2(x) P_2(y) + \dots\}$$

where $P_1(x)$, $P_2(x)$, \dots , and $P_1(y)$, $P_2(y)$, \dots , are the orthogonal polynomials respectively appropriate to $\varphi(x)$ and $\psi(y)$, then the regression of either variate on the other is linear.

It is convenient to use a form for the orthogonal polynomials in which the coefficient of the highest power is 1.

$$\text{Let } \int x^r P_r(x) \varphi(x) dx = \int \{P_r(x)\}^2 \varphi(x) dx = b_r,$$

$$\text{and } \int y^r P_r(y) \psi(y) dy = \int \{P_r(y)\}^2 \psi(y) dy = b'_r;$$

$$P_1(x) = x + c, \quad P_2(x) = x^2 + c_2 x + \text{const.}; \quad P_1(y) = y + d, \quad P_2(y) = y^2 + d_2 y + \text{const.}$$

Then the regression line of y on x is

$$y = \bar{y} + a_1 b'_1 P_1(x) = -d_1 + a_1 b'_1 P_1(x)$$

That this line passes, as a line of regression must, through the point whose coordinates are the means of x and y , is verified from the fact that $P_1(\bar{x}) = 0$.

3.22 THEOREM ∇ The variance of y for any given x is

$$\overline{y^2} - d_2 a_1 b_1' P_1(x) + a_2 b_2' P_2(x) - \{\bar{y} + a_1 b_1' P_1(x)\}^2$$

It is a constant if and only if $a_2 b_2' = (a_1 b_1')^2$,

and $a_2 b_2' c_1 - 2 a_1^2 b_1'^2 c_1 - a_1 b_1' (d_2 - 2 d_1) = 0$

These conditions are satisfied for Mehler's expression for the normal correlation distribution, but ^{not} for the series in Laguerre polynomials discussed in chapter II.

The mean, for all x , of the variance of y for a given x , is

$$\overline{y^2} - \bar{y}^2 - a_1^2 b_1 b_1'^2 = b_1' (1 - a_1^2 b_1 b_1')$$

and the ratio of this to the variance of y in the distribution $\psi(y)$ is $1 - \rho^2$, where ρ is the geometric mean of the coefficients of regression, $a_1 b_1$ and $a_1 b_1'$.

It does not however follow that for each x the variance of y is less than b_1' .

3.23

We cannot assume that every series of the form in Theorem IV represents a possible frequency function. A necessary condition is that the series should have a positive sum for every pair of values of x and y in the range concerned. One implication of this is that the variance of y for a given x must be positive for each x in the range. If the range for x extends either to $+\infty$ or to $-\infty$ this implies

$$a_1 b_1' \geq (a, b_1')^2 ;$$

and if it extends to $+\infty$ then

$$a_1 b_1' c_1 - 2 a_1^2 b_1'^2 c_1 - a_1 b_1' (d_1 - 2d_2) \geq 0 ,$$

and if to $-\infty$,

$$a_1 b_1' c_1 - 2 a_1^2 b_1'^2 c_1 - a_1 b_1' (d_1 - 2d_2) \leq 0 .$$

3.3 It is evident from theorem IV that an expansion bilinear in orthogonal polynomials of the form given in that theorem can only represent a distribution in certain special cases, namely those in which the regression is linear. But a very large proportion of observed frequency distributions have approximately linear regression, and linearity of regression is very often assumed.

We do however find, in chapter V an expansion of a very similar form for a frequency distribution in which the regression curve is parabolic. It is of the type :

$$\varphi(x) \psi(y) \{ 1 + a_1 P_1(x) P_1(y) + a_2 P_2(x) P_2(y) + \dots \}.$$

That such a series should arise naturally suggests that a much wider class of frequency distributions might be investigated by the use of expansions by means of orthogonal polynomials. It is evident that if the series contained terms $P_r(x) P_s(y)$, and $P_s(x) P_r(y)$,

the regression curve of y on x would be of degree r, and that of x on y of degree s.

Chapter II

DISTRIBUTION OF THE VARIANCES IN A NORMAL CORRELATED TWO-VARIATE DISTRIBUTION

4.1 It is well known that if a variate X has a normal distribution, so that, with suitable choice of the unit, its frequency is given by

$$dP = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \quad (1)$$

then the distribution of the variance of X can be given by the equation

$$dP = \frac{1}{\Gamma(\frac{1}{2})} x^{-\frac{1}{2}} e^{-x} dx$$

and that of the mean variance in a sample of n is

$$dP = \frac{1}{\Gamma(n/2)} x^{\frac{n}{2}-1} e^{-x} dx$$

with suitable choice of the unit.

$$\text{For, putting } x = \frac{1}{2}X^2, \quad dX = \frac{1}{\sqrt{2}} x^{-\frac{1}{2}} dx$$

equation (1) gives

$$dP = \frac{1}{2\sqrt{\pi}} x^{-\frac{1}{2}} e^{-x} dx$$

but two values of X , one positive and one negative, correspond to one value of x , so that the distribution for x is

$$dP = \frac{1}{\sqrt{\pi}} x^{-\frac{1}{2}} e^{-x} dx = \frac{1}{\Gamma(\frac{1}{2})} x^{-\frac{1}{2}} e^{-x} dx, \quad (2)$$

and the moment generating function is

$$\frac{1}{\Gamma(\frac{1}{2})} \int_0^{\infty} x^{-\frac{1}{2}} e^{-x} e^{\alpha x} dx = \frac{1}{(1-\alpha)^{\frac{1}{2}}} \quad (3)$$

4.2 If $x_1, x_2, x_3, \dots, x_n$ each have a distribution of this type, then, by the theorem of §3.1 the moment generating function of $x_1 + x_2 + \dots + x_n$ is $(1-\alpha)^{-n/2}$ and the resultant distribution function is

$$dp = \frac{1}{\Gamma(n/2)} x^{\frac{n}{2}-1} e^{-x} dx \quad (4)$$

Except for the fact that the index of x is restricted to be half an integer, this is a form into which the most general Type III equation

$$dp = \frac{k^{m+1}}{\Gamma(m+1)} e^{kx} (x-c)^m e^{-kx} dx \quad (5)$$

(in which m is a real number greater than -1), can be put, by a suitable choice of origin and unit.

We therefore expect that a similar discussion of the distribution of the variances in a two-variate normal correlated distribution will lead to a two-variate distribution which is the natural extension of the Type III distribution. Consider then two variates X and Y , with frequency function

$$dp = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{X^2 - 2\rho XY + Y^2}{2(1-\rho^2)}\right\} dXdY \quad (6)$$

and put $x = \frac{1}{2}X^2$, $y = \frac{1}{2}Y^2$, so that (6) becomes

$$dp = \exp\left\{-\frac{x - 2\rho\sqrt{xy} + y}{1-\rho^2}\right\} dx dy / (4\pi\sqrt{xy}\sqrt{1-\rho^2}) \quad (7)$$

but there are four pairs of values of X, Y corresponding to one pair of values of x, y ; two with positive and two with negative XY , so that the true expression for dp

—This analysis can be carried out more shortly by ^{putting} adding a factor $e^{i(\alpha X^2 + \beta Y^2)}$ in (6) and integrating, as in ~~chapter~~ ^{VI}

in terms of x and y is not (7) but

$$d\phi = \left\{ \exp\left(-\frac{x-2\rho\sqrt{xy}+y}{1-\rho^2}\right) + \exp\left(-\frac{x+2\rho\sqrt{xy}+y}{1-\rho^2}\right) \right\} dx dy / (2\pi\sqrt{xy}\sqrt{1-\rho^2}) \tag{8}$$

in which the positive sign must be taken with each root.

We shall now find the moment generating function for this distribution. It is defined as

$$G(\alpha, \beta) = \iint \phi(x, y) e^{\alpha x + \beta y} dx dy$$

and putting $A = A/(1-\rho^2)$; $B = B/(1-\rho^2)$,

we have

$$G(\alpha, \beta) = \int_0^\infty \int_0^\infty \exp\left\{-\frac{x(1-A)-2\rho\sqrt{xy}+y(1-B)}{1-\rho^2}\right\} \frac{dx dy}{2\pi\sqrt{xy}\sqrt{1-\rho^2}} \tag{9}$$

+ a similar integral with the opposite sign for ρ .

This last phrase will be written + s.i. (ρ)

Now put $\xi = x(1-A)$, $\gamma = y(1-B)$, $r = \frac{\rho}{\sqrt{1-A}\sqrt{1-B}}$

$$G(\alpha, \beta) = \frac{1}{\sqrt{1-A}\sqrt{1-B}} \int_0^\infty \int_0^\infty \exp\left\{-\frac{\xi-2r\sqrt{\xi\gamma}+\gamma}{1-\rho^2}\right\} \frac{d\xi d\gamma}{2\pi\sqrt{1-r^2}\sqrt{(1-A)(1-B)}\sqrt{\xi\gamma}} + s.i.(r)$$

and, putting $k = \frac{1-r^2}{1-\rho^2}$, $u = k\xi$, $v = k\gamma$, this is

$$\frac{1}{\sqrt{1-A}\sqrt{1-B}} \int_0^\infty \int_0^\infty \exp\left\{-\frac{u-2r\sqrt{uv}+v}{1-r^2}\right\} \frac{du dv}{2\pi\sqrt{kuv}\sqrt{1-r^2}\sqrt{(1-A)(1-B)}} + s.i.(r)$$

Using the fact that the integral of (8) is 1, this shows

that $G(\alpha, \beta) = \frac{1}{\sqrt{k(1-A)(1-B)}} \cdot \frac{\sqrt{1-\rho^2}}{\sqrt{1-r^2(1-A)(1-B)}} = \frac{1}{\sqrt{k(1-A)(1-B)}}$

$$= \sqrt{\frac{1-\rho^2}{(1-A)(1-B)-\rho^2}} = \left\{ (1-\alpha)(1-\beta) - \alpha\beta\rho^2 \right\}^{-\frac{1}{2}} \tag{10}$$

5.12

The theorem of §3.1 shows at once that the m.g.f.

for the distribution of the mean variances in a sample of n

is $\{(1-\alpha)(1-\beta) - \alpha\beta\rho^2\}^{-n/2}$

and we shall therefore take $\{(1-\alpha)(1-\beta) - \alpha\beta\rho^2\}^{-p}$ (11)

as the m.g.f. for a two-variate distribution in which each

of the variates has a distribution $\varphi(x) = x^{p-1} e^{-x} / \Gamma(p)$ and

m.g.f. $(1-\alpha)^{-p}$.

5.13

We can expand $G(\alpha, \beta) = \{(1-\alpha)(1-\beta) - \alpha\beta\rho^2\}^{-p}$ as

$(1-\alpha)^{-p} (1-\beta)^{-p} \{1 + p \frac{\alpha}{1-\alpha} \frac{\beta}{1-\beta} \rho^2 + \frac{p(p+1)}{2!} (\frac{\alpha}{1-\alpha})^2 (\frac{\beta}{1-\beta})^2 \rho^4 + \dots\}$ (12)

and since $(1-\alpha)^{-p} (\frac{\alpha}{1-\alpha})^r = \frac{(-1)^r}{p(p+1)(p+2)\dots(p+r-1)} \alpha^r (\frac{\partial}{\partial \alpha})^r (1-\alpha)^{-p}$

is the m.g.f. of $L_r(x, p) \varphi(x) \frac{\Gamma(p)}{\Gamma(p+r)}$ we see that $\{(1-\alpha)(1-\beta) - \alpha\beta\rho^2\}^{-p}$

is the m.g.f. of a frequency function $\varphi(x, y)$ in which the

co-factor of $\varphi(x) \varphi(y)$ is a series bilinear in the Laguerre

polynomials which are orthogonal with respect to $\varphi(x) = \frac{x^{p-1} e^{-x}}{\Gamma(p)}$:

$\varphi(x) \varphi(y) \{1 + \frac{\rho^2}{p} L_1(x, p) L_1(y, p) + \frac{\rho^4}{2! p(p+1)} L_2(x, p) L_2(y, p) + \dots\}$ (13)

5.2

To prove the convergence of this series we use

a result by E. Hille (1926) who has shown that

$L_r^{(\alpha)}(x) = \frac{1}{\sqrt{\pi}} e^{\frac{1}{2}x} x^{-\frac{1}{4}-\frac{\alpha}{2}} v^{\frac{\alpha}{2}-\frac{1}{4}} \cos[2\sqrt{rx} - \pi(\frac{1}{4} + \frac{\alpha}{2})] + O(r^{\frac{\alpha}{2}-\frac{1}{2}})$

where $L_r^{(\alpha)}(x) = \frac{1}{r!} e^x x^{-\alpha} (\frac{d}{dx})^r [x^{r+\alpha} e^{-x}] = \frac{(\frac{d}{du})^r \{x^r \varphi(x, \alpha+1)\}}{r! \varphi(x, \alpha+1)}$

In our notation, $p = \alpha + 1$, $L_r(x; p) = r! L_r^{(\omega)}(x)$, and

$$L_r(x, p) = r! \left\{ \frac{1}{\sqrt{\pi}} e^{-x/2} x^{\frac{1}{4} - \frac{p}{2}} r^{\frac{p}{2} - \frac{1}{4}} \cos \left[2\sqrt{rx} - \pi \left(\frac{p}{2} - \frac{1}{4} \right) \right] + O\left(r^{\frac{p}{2}-1}\right) \right\}$$

so that

$$\frac{L_r(x, p)L_r(y, p)}{r!(p+1)\dots(p+r-1)} \cdot \frac{(r-1)!p(p+1)\dots(p+r-2)}{L_{r-1}(x, p)L_{r-1}(y, p)} \rightarrow 1$$

as r tends to infinity, and the ratio of any term to the preceding term in (13) tends to ρ^2 . The series is therefore absolutely convergent, if $|\rho| < 1$, for all values of x and p .)

§3

Some deductions from the moment generating function (12) or from the series (13) may be noted :

(i) the product moment for the distribution is $p\rho^2$, and the mean variance for either variate is p . It was evident from the way in which this distribution was deduced from the Normal distribution, that the product moment, when expressed in terms of units in which the mean variances of both x and y are 1, ~~must~~ must be an even function of ρ , equal to 0 when ρ is 0, and to 1 when ρ is 1.

(ii) the line of regression of y on x is

$$y = p + \rho^2 L(x, p) = p(1-\rho^2) + \rho^2 x,$$

~~This is a limiting form of (5), chapter VIII, when $v \rightarrow 0$~~
and the variance of y for any given x is

$$p + 2\rho^2(1-\rho^2)(x-p) - p\rho^4$$

This is a limiting form of (5) in chapter VIII, when $v = 0$, for here $p = h_0 = v_1$

) For the purposes of statistical theory we are interested only in the case in which $|\rho| < 1$. The series clearly diverges when $|\rho| > 1$. When $\rho = \pm 1$ it is likely that it will converge within some region $-\eta < \rho < \eta$, $-\eta < \rho < \eta$, (Hille, Amer. Math. Monthly, April 1938) $-\eta < \rho < \eta$.

Chapter III

EXPRESSION OF THE SERIES OF LAGUERRE POLYNOMIALS

AS A BESSEL FUNCTION

§6 The series in Laguerre polynomials can be reduced to a simple form involving a Bessel function, by two different processes ; first, directly, using a result of Hardy's, and secondly, by a process of integration from probability distributions discussed by Wishart and Bartlett and others. It is of interest to give both, and connect the theory of the distribution function (chapter II, p. 25) with various analytical results.

Hardy has shown that¹⁾

$$\sum_0^{\infty} (-t)^n \chi_n(x) \chi_n(y) = \frac{t^{-\alpha}}{1+t} \exp\left\{-\frac{1}{2}(x+y) \frac{1-t}{1+t}\right\} J_{2\alpha}\left\{\frac{2\sqrt{xyt}}{1+t}\right\} \quad (1)$$

where $\chi_n(x) = \{n! \Gamma(n+1+2\alpha)\}^{-\frac{1}{2}} \frac{e^{-\frac{1}{2}x}}{x^{\alpha}} \left(\frac{d}{dx}\right)^n (e^{-x} x^{n+2\alpha})$.

In our notation, this is $\chi_n(x) = \{n! \Gamma(n+b)\}^{-\frac{1}{2}} e^{-\frac{1}{2}x} x^{\frac{1}{2}(b-1)} L_n(x, b)$;

$$\sum_0^{\infty} (-t)^n \frac{L_n(x, b) L_n(y, b)}{n! \Gamma(n+b)} e^{-\frac{1}{2}(x+y)} (xy)^{\frac{1}{2}(b-1)} = \frac{t^{-\frac{1}{2}(b-1)}}{1+t} \exp\left\{-\frac{1}{2}(x+y) \frac{1-t}{1+t}\right\} J_{b-1}\left\{\frac{2\sqrt{xyt}}{1+t}\right\}.$$

¹⁾ G. H. Hardy, Jour. Lond. Math. Soc., 7 (1932) pp. 138-139

G. N. Watson, Jour. Lond. Math. Soc., 1933. p. 189

A different proof has been given by N. A. Shastri, Jour. Ind. Math. Soc. Vol. 1, No. 7 (1935) p. 238.

Putting $t = -\rho^2$, this shows at once that the expression (13)

$$\begin{aligned}
 &= \frac{1}{\Gamma(p)} e^{-\frac{1}{2}(x+y)} (xy)^{\frac{1}{2}(p-1)} \frac{\rho^{-(p-1)}}{1-\rho^2} \exp\left\{-\frac{1}{2}(x+y) \frac{1+\rho^2}{1-\rho^2}\right\} I_{p-1}\left\{\frac{2\rho\sqrt{xy}}{1-\rho^2}\right\} \\
 &= \frac{\rho^{-(p-1)}}{\Gamma(p)(1-\rho^2)} (xy)^{\frac{1}{2}(p-1)} \exp\left\{-(x+y)/(1-\rho^2)\right\} I_{p-1}\left\{\frac{2\rho\sqrt{xy}}{1-\rho^2}\right\}, \quad (2)
 \end{aligned}$$

$$\text{for } I_{p-1}(z) = e^{-\frac{1}{2}\pi(p-1)i} J_{p-1}(ze^{\frac{1}{2}\pi i}).$$

§ 7.1

The second method of deducing this result starts from results of Wishart and Bartlett¹⁾ who have discussed the distribution of the variances and bivariate correlations in a sample of n sets from a population with p variates; that is, the distribution-function of the $\frac{1}{2}n(n+1)$ variables

$$\begin{aligned}
 v_{11} &= (x_{11}^2 + x_{12}^2 + \dots + x_{1n}^2)/n, & v_{22} &= (x_{21}^2 + x_{22}^2 + \dots + x_{2n}^2)/n, \dots, \\
 v_{12} &= (x_{11}x_{21} + x_{12}x_{22} + \dots + x_{1n}x_{2n})/n, & v_{13} &= (x_{11}x_{31} + x_{12}x_{32} + \dots + x_{1n}x_{3n})/n, \dots,
 \end{aligned}$$

$x_{11}, x_{12}, \dots, x_{1n}; x_{21}, x_{22}, \dots, x_{2n}; \dots; x_{n1}, x_{n2}, \dots, x_{nn}$,
 where $\begin{matrix} x & y & x & y & \dots & x & y \\ \hline & & & & & & \end{matrix}$ are sample-pairs^{sets} from a normally correlated population. They first find the moment generating function of the joint distribution of $x_1, x_2, x_3, \dots, x_p$ and x_1x_2, x_1x_3, \dots . It is

$$\left\{ |R|/|R-A| \right\}^i \exp\{B(\alpha, \alpha)\}, \quad (3)$$

where

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{p1} & \alpha_{p2} & \dots & \alpha_{pp} \end{bmatrix}, \quad R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & r_{pp} \end{bmatrix}$$

¹⁾Wishart and Bartlett, Proc. Camb. Phil. Soc., 29 1932-33, p. 260

the frequency distribution with which the discussion starts being

$$f(x_1, x_2, \dots, x_p) = \pi^{-\frac{1}{2}p} |R|^{\frac{1}{2}} \exp\{-R(x, x)\} \tag{4}$$

where $R(x, x)$ denotes the quadratic form

$$\sum_{\mu, \nu} r_{\mu\nu} x_\mu x_\nu;$$

and $B = (R - A)^{-1}$, $B(\alpha, \alpha)$ being

a quadratic form in $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_p$, which correspond to the first order variates $x_1, x_2, x_3, \dots, x_p$.

It then follows (see chapter I, §3.1) that the m.g.f. for the variances, bivariances, and means is

$$\left\{ |R|/|R-A| \right\}^{\frac{1}{2}} \exp\{n B(\alpha, \alpha)\}. \tag{5}$$

The distribution function for these $p + \frac{1}{2}p(p + 1)$ variables is then to be found by a Fourier integral inversion, and the integration with respect to the p first order variables can be carried out and leaves

$$f(v_{\mu\nu}) = (2\pi)^{-p} \pi^{-\frac{1}{2}p(p-1)} |R|^{\frac{1}{2}(n-1)} \int_{-i\infty}^{+i\infty} \dots \int_{-i\infty}^{+i\infty} |R-A|^{\frac{1}{2}(n-1)} \exp\{-\sum (\alpha_{\mu\nu} v_{\mu\nu})\} d\alpha_{11} d\alpha_{12} \dots d\alpha_{pp} \tag{6}$$

as the distribution function for $v_{11}, v_{12}, v_{13}, \dots, v_{1p}, v_{22}, \dots, v_{pp}$.

This differs only in having $n - 1$ for n from the integral which would be obtained by putting $\alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_p = 0$ in (3) so that the factor $\exp\{n B(\alpha, \alpha)\}$ disappears, (thereby giving the m.g.f. for the distribution of the variances and

bivariances,) AND APPLYING THE Fourier theorem to that m.g.f. ; thus showing that the distribution of the variances and bi-variances from the means, in a sample of n is the same as the distribution of the true variances and bivariances

in a sample of $n - 1$.

The integral (6) has been evaluated by Ingham¹⁾ and by Ledermann²⁾. The result is

$$f(v_{\mu\nu}) = \pi^{-\frac{1}{2}t(t-1)} |R|^{\frac{1}{2}(n-1)} |v_{\mu\nu}|^{\frac{1}{2}(n-t-2)} \frac{\exp\left\{-\sum_{\mu\nu=1}^n r_{\mu\nu} v_{\mu\nu}\right\}}{\prod_{\nu=1}^t \left\{\Gamma\left[\frac{1}{2}(n-\nu)\right]\right\}} \quad (7)$$

where $r_{\mu\nu} = \frac{1}{2} \Delta_{\mu\nu} / (\sigma_\mu \sigma_\nu \Delta)$,
 Δ being $|P_{\mu\nu}|$.

¹⁾ Ingham, A.E., Proc. Camb. Phil. Soc. 29 (1926) p.261.

²⁾ Dr Ledermann kindly gave me a manuscript of his shorter proof.

7.2 The distribution function for the p variances and $\frac{1}{2} p (p - 1)$ bivariances has thus been found by ~~Wishart~~ Wishart and Bartlett. In chapter II (is found) the distribution function for the p variances is found as a multiple series in Laguerre polynomials. It is evident that an expression equivalent to this would be found by integrating (7) with respect to the $\frac{1}{2} p (p - 1)$ variables not on the diagonal in

$$\begin{bmatrix} v_{11} & v_{12} & \dots & v_{1p} \\ v_{12} & v_{22} & \dots & v_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1p} & v_{2p} & \dots & v_{pp} \end{bmatrix} \quad \text{throughout all the space in which this matrix is positive definite.}$$

7.3 To carry out the integration when $p = 2$: then, if $\sigma_1 = \sigma_2 = 1$, $\Delta = 1 - \rho^2$, $r_{12} = r_{21} = -\frac{1}{2} \rho / (1 - \rho^2)$, $r_{11} = r_{22} = \frac{1}{2} \frac{1}{1 - \rho^2}$, $R = \frac{1}{2} \frac{1}{1 - \rho^2}$,

$$f(v_{11}, v_{12}, v_{22}) = \frac{\pi^{-\frac{1}{2}} \{4(1 - \rho^2)\}^{-\frac{1}{2}(n-1)}}{\Gamma\{\frac{1}{2}(n-1)\} \Gamma\{\frac{1}{2}(n-2)\}} \begin{vmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{vmatrix}^{\frac{1}{2}(n-2)} \exp\left\{-\frac{v_{11} - 2\rho v_{12} + v_{22}}{2(1 - \rho^2)}\right\} \quad (8)$$

This is to be integrated with respect to v_{12} throughout $v_{12}^2 < v_{11} v_{22}$. Putting $V^2 = v_{11} v_{22}$, the required frequency function is

$$f(v_{11}, v_{22}) = \frac{\pi^{-\frac{1}{2}} \{4(1 - \rho^2)\}^{-\frac{1}{2}(n-1)}}{\Gamma\{\frac{1}{2}(n-1)\} \Gamma\{\frac{1}{2}(n-2)\}} \exp\left\{-\frac{v_{11} + v_{22}}{2(1 - \rho^2)}\right\} \int_{-V}^{+V} (V^2 - v_{12}^2)^{\frac{1}{2}(n-2)} \exp\left\{\frac{2\rho v_{12}}{2(1 - \rho^2)}\right\} dv_{12} \quad (9)$$

The integral is, putting $t = V - v_{12}$,

$$\int_0^{2V} t^{\frac{1}{2}(n-2)} (2V - t)^{\frac{1}{2}(n-2)} e^{\rho V / (1 - \rho^2)} e^{-\rho t / (1 - \rho^2)} dt$$

which can be expanded in a series of incomplete Gamma functions

$$\begin{aligned}
 & e^{pV/(1-p^2)} \sum_{t=0}^{\infty} \left[(2V)^{\frac{1}{2}(n-4)-t} \binom{\frac{1}{2}(n-4)}{t} \int_0^{2V} t^{\frac{1}{2}(n-4)+t} e^{-pt/(1-p^2)} dt \right] \\
 &= e^{pV/(1-p^2)} \sum \left[(2V)^{\frac{1}{2}(n-4)-t} \binom{\frac{1}{2}(n-4)}{t} \left(\frac{1-p^2}{p}\right)^{\frac{1}{2}(n-4)+t+1} \gamma\left\{\frac{1}{2}(n-4)+t+1, \frac{2VP}{1-p^2}\right\} \right] \\
 &= \left(\frac{2\pi VP}{1-p^2}\right)^{\frac{1}{2}} (2V)^{\frac{1}{2}(n-4)} \left(\frac{1-p^2}{p}\right)^{\frac{1}{2}(n-4)+1} \Gamma_{\frac{1}{2}(n-3)}\left\{\frac{VP}{1-p^2}\right\} \Gamma\left\{\frac{1}{2}(n-2)\right\} \quad (10)
 \end{aligned}$$

by the result quoted in chapter IV

Substituting in (9) we have $f(v_{11}, v_{22}) =$

$$\frac{(V/2p)^{\frac{1}{2}(n-3)}}{4 \Gamma\left\{\frac{1}{2}(n-1)\right\}} \cdot \frac{1}{1-p^2} \exp\left\{-\frac{v_{11}+v_{22}}{2(1-p^2)}\right\} \Gamma_{\frac{1}{2}(n-3)}\left\{\frac{VP}{1-p^2}\right\} \quad (11)$$

Putting $x = \frac{1}{2}v_{11}$, $y = \frac{1}{2}v_{22}$, we obtain an expression identical with (2), when $p = \frac{1}{2}(n-1)$.

Chapter IV

EXPRESSIONS FOR BESSEL FUNCTIONS INVOLVING INCOMPLETE GAMMA FUNCTIONS

§8. The asymptotic expansions of the Bessel function of imaginary argument¹⁾, can be written

$$I_\nu(z) \sim \frac{e^z}{(2\pi z)^{\frac{1}{2}}} \sum_{m=0}^{\infty} \binom{\nu - \frac{1}{2}}{m} \frac{\Gamma(\nu + m + \frac{1}{2})}{\Gamma(\nu + \frac{1}{2})} \cdot \frac{1}{(-2z)^m} \quad (1)$$

It has been pointed out by Hadamard²⁾ that it is possible to modify such an asymptotic expansion so that it becomes a convergent series with a negligible remainder term. One modification which he gives³⁾ for functions of zero order, and which is given by Watson⁴⁾ for functions of any order greater than $-\frac{1}{2}$, is formed by replacing the Gamma function $\Gamma(\nu + m + \frac{1}{2})$ in (1) by the incomplete Gamma function, $\Gamma_{2z}(\nu + m + \frac{1}{2})$ or $\gamma(\nu + m + \frac{1}{2}, 2z)$.

The object of this chapter is to find what other similar expressions involving incomplete Gamma functions can be found as solutions of Bessel's equation.

It is shown that

$$I_\nu(z) = \frac{e^{+z}}{(2\pi z)^{\frac{1}{2}}} \sum_{m=0}^{\infty} \binom{\nu - \frac{1}{2}}{m} \frac{\gamma(\nu + m + \frac{1}{2}, 2z)}{\Gamma(\nu + \frac{1}{2})} \cdot \frac{1}{(-2z)^m} \quad (2)$$

¹⁾ Hankel, Math. Ann. I (1869), pp. 491-495 ; Watson, Theory of Bessel Functions, 1922, chap. VII

²⁾ Trans. Amer. Math. Soc. III Oct. 1902, pp. 421-422

³⁾ Bulletin de la Soc. Math. de France XXXVI, 1908, pp. 77-85;

⁴⁾ Watson pp. 204-205

57

which is Watson's result, when the real part of $\nu > -\frac{1}{2}$
 We also find a series with positive powers of z as a
 solution :

$$z^{-\nu} e^z \sum_{n=0}^{\infty} \frac{(-)^n \Gamma(\nu + \frac{1}{2})}{\Gamma(\nu - n + \frac{1}{2}) \cdot n!} (2z)^n \gamma(\mu - n, 2z)$$

This can similarly be regarded as obtained from the
 convergent series for $z^{\nu} e^{-z} \Gamma_{\nu}(z)$ in the same way as (2) is
 obtained from the asymptotic expansion (1).

It is shown that each of these is a special case of a series
 in positive and negative powers, which is convergent,
 and which can be obtained by changing a Gamma function to
 an incomplete Gamma function, in a series which is in
 general divergent :

$$z^a e^z \sum_{n=-\infty}^{\infty} \frac{(-)^n \Gamma(\nu + \frac{1}{2})}{\Gamma(\nu - n + a + 1) \Gamma(n - a + \frac{1}{2})} (2z)^n \gamma(\mu + n, 2z)$$

It is further shown that for no values of a and p can
 there be more than one ^{convergent} solution of the type

$$z^a e^{bz} \sum_{n=-\infty}^{\infty} c_n (bz)^n \gamma(\mu + n, bz)$$

of the Bessel equation.

so that

$$\left| \gamma[s, \omega] \right| \leq \frac{K}{\omega^2}$$

where K is independent of w and s .

Now
$$\sum_0^{\infty} c_n \omega^n \gamma(\mu - n, \omega)$$

can be written as the difference of two series

$$\sum_0^{\infty} c_n \omega^n \Gamma(\mu - n) - \sum_0^{\infty} c_n \omega^n \gamma(\mu - n, \omega),$$

and the first of these converges if c_n is bounded, and the second is convergent if $\sum_0^{\infty} c_n$ is absolutely convergent. (9)

A simpler discussion of the conditions for the convergence of

$$\sum_{n=-\infty}^{\infty} c_n \omega^n \gamma(m + \mu, \omega)$$

is : substitute from (6), and rearrange the double series so formed :

$$\omega^\mu e^{-\omega} \sum_{n=0}^{\infty} \left\{ \sum_{m=-\infty}^{\infty} \frac{\omega^m \cdot c_m}{(\mu+m)(\mu+m+1) \dots (\mu+m+n)} \right\}$$

Each of these series is convergent if the first is convergent, and thus the necessary and sufficient condition that the series considered should converge is that $\sum_{-\infty}^{\infty} \frac{c_n}{\mu+n}$ should converge.

A finite series

It is of interest to consider first the case in which the series consists of a finite number of terms only ; the analysis is simpler than in the general case.

We shall therefore consider for what values of the constants $a, p, \mu, k, \kappa, c_0, c_1, c_2, \dots, c_n$, the function $F(z)$ defined as

$$z^a e^{kz} \sum_{m=0}^{n+\mu} c_m (kz)^m \gamma(m+\mu, kz) \quad (10)$$

is a solution of the differential equation

$$\frac{d^2 F}{dz^2} + \frac{1}{z} \frac{dF}{dz} - \left(\kappa^2 + \frac{\nu^2}{z^2} \right) F = 0 \quad (11)$$

where n is a positive integer, k is not zero, μ is not zero and not a negative integer, and otherwise the constants may be any complex numbers.

10.2 Substituting in (10) for $\gamma(m+\mu, kz)$ from ~~(6)~~ the series (6),

$$\text{we have } F(z) = z^a e^{(p-k)z} \sum_{m=0}^n \sum_{t=0}^{\infty} c_m \Gamma(m+\mu) (kz)^{m+t} / \Gamma(m+\mu+t+1)$$

the re-arrangement in the order of the powers of z being permissible since the series (6) is absolutely convergent.

Denote the coefficient of $z^{a+\mu+t} e^{(p-k)z}$ in $F(z)$ by a_t .

$$a_t = \sum_{m=0}^n c_m k^{m+t} \Gamma(m+\mu) / \Gamma(m+\mu+t+1) \quad (12)$$

Then

$$F(z) = \sum_{t=0}^{\infty} e^{(p-k)z} z^{a+\mu+t} a_t$$

$$\frac{\partial F}{\partial z} = \sum_{t=0}^{\infty} \left\{ (p-k) a_t + (a+\mu+t+1) a_{t+1} \right\} e^{(p-k)z} z^{a+\mu+t}$$

$$\frac{\partial^2 F}{\partial z^2} = \sum_{t=0}^{\infty} \left\{ (p-k)^2 a_t + 2(p-k)(a+\mu+t+1) a_{t+1} + (a+\mu+t+1)(a+\mu+t+2) a_{t+2} \right\} e^{(p-k)z} z^{a+\mu+t} \quad (13)$$

Substituting from equations (13) in equation (11), omitting the common factor $e^{(p-k)z}$, and equating coefficients of powers of z , we obtain the set of conditions required :

$$\begin{aligned}
 & (p-k)^2 a_t + 2(p-k)(a+\mu+t+1)a_{t+1} + (a+\mu+t+1)(a+\mu+t+2)a_{t+2} \\
 & + (p-k)a_{t+1} + (a+\mu+t+2)a_{t+2} \\
 & - k^2 a_t - v^2 a_{t+2},
 \end{aligned}$$

$$t = -2, -1, 0, 1, 2, 3, \dots, \tag{14}$$

Substituting for a_t, a_{t+1}, a_{t+2} , omitting the common factor k^{u+t} , this becomes, for $t = 0, 1, 2, 3, \dots$, $\left[\frac{(a+\mu+t+2)^2 - v^2}{k^2} \right] / \Gamma(n+\mu+t+3)$

$\sum_{m=0}^n [c_m \Gamma(m+\mu) \{ \frac{(p-k)^2 - k^2}{\Gamma(n+\mu+t+1)} + 2(p-k) \frac{k}{\Gamma(n+\mu+t+2)} + \frac{(a+\mu+t+2)^2 - v^2}{\Gamma(n+\mu+t+3)} \} k^2]$
 Multiplying throughout by $\Gamma(n+\mu+t+3)$, we obtain an equation of degree $n+2$ in t :

$$\begin{aligned}
 & \{ (p-k)^2 - k^2 \} \sum_{m=0}^n \{ c_m \Gamma(m+\mu) (m+\mu+t+1) (m+\mu+t+2) \dots (n+\mu+t+2) \} \\
 & + 2(p-k)(a+\mu+t+\frac{3}{2}) \sum_{m=0}^n \{ c_m \Gamma(m+\mu) (m+\mu+t+2) \dots (n+\mu+t+2) \} k \\
 & + \{ (a+\mu+t+2)^2 - v^2 \} \sum_{m=0}^n \{ c_m \Gamma(m+\mu) (m+\mu+t+3) \dots (n+\mu+t+2) \} k^2 = 0 \tag{15}
 \end{aligned}$$

Equating to zero the coefficients in (15) we obtain $n+3$ equations, which, together with the two equations (14) for which $t = -1$ and $t = -2$, form the necessary and sufficient set of conditions that the function (10) should satisfy the equation (11).

From the coefficient of t^{n+2} we obtain, assuming $c_0 \neq 0$,

$$\begin{aligned}
 & \{ (p-k)^2 - k^2 \} + 2(p-k)k + k^2 = 0 ; \\
 & k^2 = \{ (p-k) + k \}^2 = p^2. \tag{16}
 \end{aligned}$$

10.3 And the coefficient of t^{n+1} ,

$$\begin{aligned}
 & \{ (p-k)^2 - k^2 \} \{ c_0 \Gamma(\mu) \frac{n+2}{2} (n+2\mu+3) + c_1 \Gamma(\mu+1) \} \\
 & + 2(p-k)k \{ c_0 \Gamma(\mu) [a+\mu+\frac{3}{2}] + \frac{n+1}{2} (n+2\mu+4) \} + c_1 \Gamma(\mu+1) \\
 & + k^2 \{ c_0 \Gamma(\mu) [2(a+\mu+2) + \frac{n}{2} (n+2\mu+5)] + c_1 \Gamma(\mu+1) \} = 0 ;
 \end{aligned}$$

which reduces, by (16) to

$$(k^2 - 2kp) \frac{1}{2} (n+6) + 2(p-k)k \{ a+\mu+\frac{3}{2} + \frac{1}{2} (2\mu+4) \} + k^2 2(a+\mu+2) = 0,$$

i.e., $2pk(2a+1) = 0$;

either $p = 0$, or $2a+1 = 0$. (17)

10.3) If $p = 0$, $\kappa = 0$, and the equations ^{corresponding to} (19) have no solution in which $b_0, b_1, b_2, \dots, b_n$ are not all zero unless $\nu^2 = a^2$, and then $b_1 = b_2 = b_3 = \dots = b_n = 0$, so that if $p=0$ there is no solution in the form of a series such as (10). We put $\kappa = 1$, so that (10) becomes Bessel's equation, in the modified form for a purely imaginary argument. And (17) gives $2a + 1 = 0$.

10.4 To reduce the $n + 1$ equations obtained by equating to zero the coefficients of t^s, t^{s-1}, \dots in (15), put

$$b_n = c_n \Gamma(n + \mu) \quad (18)$$

and let $S(r, s)$ denote the sum of all products of r of the numbers $\mu + s + 1, \mu + s + 2, \dots, \mu + n + 2$, so that

$$\begin{aligned} \frac{S(r, s+1)}{S(r, s)} &= S(r, s) - (\mu + s + 1) \cdot S(r-1, s+1), \\ S(0, s) &= 1, \quad s = 0, 1, 2, \dots, n+1. \end{aligned} \quad (19)$$

Then the coefficient in (15) of t^{s-s} is

$$\begin{aligned} & (k^2 - 2kp) \{ b_0 S(s+2, 0) + b_1 S(s+1, 1) + b_2 S(s, 2) + \dots + b_{s+2} S(0, s+2) \} \\ & + 2(p-k)(\mu+1)k \{ b_0 S(s+1, 1) + b_1 S(s, 2) + b_2 S(s-1, 3) + \dots + b_{s+1} S(0, s+1) \} \\ & + 2(p-k)k \{ b_0 S(s+2, 1) + b_1 S(s+1, 2) + b_2 S(s, 3) + \dots + b_{s+2} S(0, s+3) \} \\ & + k^2 \left\{ (\mu + \frac{3}{2})^2 - \nu^2 \right\} \{ b_0 S(s, 2) + b_1 S(s-1, 3) + b_2 S(s-2, 4) + \dots + b_s S(0, s+2) \} \\ & + k^2 \cdot 2(\mu + \frac{3}{2}) \{ b_0 S(s+1, 2) + b_1 S(s, 3) + b_2 S(s-1, 4) + \dots + b_{s+1} S(0, s+3) \} \\ & + k^2 \{ b_0 S(s+2, 2) + b_1 S(s+1, 3) + b_2 S(s, 4) + \dots + b_{s+2} S(0, s+4) \} \end{aligned}$$

Simplifying this by equations (19), we have, for $s = 0, 1, 2, \dots, n$,

$$\begin{aligned} & 2(p-k)(\mu+1) \{ b_0 S(s+1, 1) + b_1 S(s, 2) + \dots \} \\ & + (k-2p) \{ (\mu+1)b_0 S(s+1, 1) + (\mu+2)b_1 S(s, 2) + \dots \} \\ & + 2k(\mu + \frac{3}{2}) \{ b_0 S(s+1, 2) + b_1 S(s, 3) + \dots \} \\ & + k \left\{ (\mu + \frac{3}{2})^2 - \nu^2 \right\} \{ b_0 S(s, 2) + b_1 S(s-1, 3) + \dots \} \\ & - k \{ (\mu+2)b_0 S(s+1, 2) + (\mu+3)b_1 S(s, 3) + \dots \} = 0 \end{aligned} \quad (20)$$

The coefficient of l_m is

$$-2\beta_m S'(s-m+1, m+1) + k[(\mu-m+1)\{S(s-m+1, m+2) - S(s-m+1, m+1)\} + (\mu^2 + 3\mu + \frac{9}{4} - \nu^2)S(s-m, m+2)]$$

$$= -2\beta_m S(s-m+1, m+1) + k(m^2 + m + \frac{1}{4} - \nu^2)S(s-m, m+2),$$

so that the determinant obtained by eliminating l_0, l_1, \dots, l_m from the $n+1$ equations (20) is

$$\begin{vmatrix} k(\frac{1}{4} - \nu^2)S(0, 2) & -2\beta S(0, 2) & 0 & \dots & 0 \\ k(\frac{1}{4} - \nu^2)S(1, 2) & k(\frac{9}{4} - \nu^2)S(0, 3) - 2\beta S(1, 2) & -2\beta \cdot 2S(0, 3) & \dots & 0 \\ k(\frac{1}{4} - \nu^2)S(2, 2) & k(\frac{9}{4} - \nu^2)S(1, 3) - 2\beta S(2, 2) & k(\frac{25}{4} - \nu^2)S(0, 4) - 2\beta \cdot 2S(1, 3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k(\frac{1}{4} - \nu^2)S(n, 2) & k(\frac{9}{4} - \nu^2)S(n-1, 3) - 2\beta S(n, 2) & \dots & k\{(n+\frac{1}{2})^2 - \nu^2\}S(0, n+2) & -2\beta nS(1, n+1) \end{vmatrix}$$

which is of degree $n+1$ in ν^2 , so that the condition that the equations (20) should have a non-zero solution is that ν^2 should have one of the values $(\frac{1}{2})^2, (\frac{3}{2})^2, (\frac{5}{2})^2, \dots, (n+\frac{1}{2})^2$. If, however, $\nu^2 = (m + \frac{1}{2})^2$, where $m < n$, the solution of (20) is one in which $l_j = 0$, for all $j \geq m+1$. So that the only solution which does not in fact correspond to a lower value of n , is $\nu^2 = (n + \frac{1}{2})^2$. (22)

10.5 Solution of (20) now gives

$$l_n = \frac{k}{2\beta} \frac{(n-\frac{1}{2})^2 - \nu^2}{m} l_{m-1} = (-)^m \left(\frac{k}{2\beta}\right)^m l_0 \frac{(n+m)(n+m-1)\dots(n-m+1)}{m!} \quad (23)$$

10.6 We have now found the values which $a, \beta, n, l_0, l_1, l_2, \dots, l_n$

must have in order that those of the equations (14) for which $t = 0, 1, 2, 3, \dots$, may be satisfied. The equations (14) for $t = -1$ and $t = -2$ will give k and μ

$$\text{That for } t = -2 \text{ is } (a + \mu - 1)(a + \mu) + (a + \mu) - \nu^2 = 0$$

$$\text{Hence } \left(n + \frac{1}{2}\right)^2 = \nu^2 = (\mu + a)^2 = \left(\mu - \frac{1}{2}\right)^2 \quad (24)$$

A negative integral or zero value for μ is not permissible, and therefore $n = \mu - 1$.

Then, from (23) and (18) we have

$$c_n = b_n / \Gamma(n + \mu) = (-)^n \left(\frac{k}{2p}\right)^m b_0 \frac{1}{n!(n-m)!} = \left(-\frac{k}{2p}\right)^m c_0 \frac{n!}{n!(n-m)!}$$

It is convenient to put $c_0 = 1$, so that

$$c_n = \binom{n}{m} \left(-\frac{k}{2p}\right)^m \quad (25)$$

10.7 Now the equation (14) for $t = -1$ is

$$(p - k) a_0 + a_1 = 0, \quad (26)$$

$$\text{and } a_0 = \sum_{n=0}^{\infty} \binom{n}{m} \left(-\frac{k}{2p}\right)^m \frac{k^m}{n + \mu}$$

$$a_1 = \sum_{n=0}^{\infty} \binom{n}{m} \left(-\frac{k}{2p}\right)^m \frac{k^{\mu+1}}{(n + \mu + 1)(n + \mu)}$$

$$= \sum_{n=0}^{\infty} \binom{n}{m} \left(-\frac{k}{2p}\right)^m k^{\mu+1} \left(\frac{1}{n + \mu} - \frac{1}{n + \mu + 1}\right)$$

Denoting $\sum \binom{n}{m} \left(-\frac{k}{2p}\right)^m / (n + \mu)$ by $B(n + 1, \mu; \frac{k}{2p}) / \left(\frac{k}{2p}\right)^m$,

$$\left(\text{since } \sum_0^n (-x)^{r+s} \binom{n}{r} \frac{(-1)^s}{r+s} = (-1)^s \sum_0^n \binom{n}{r} \int_0^x (-x)^{r+s-1} dx\right)$$

$$= \int_0^x (1-x)^n x^{s-1} dx$$

the incomplete Beta function),

equation (26) can be written

$$B(n + 1, \mu; \frac{k}{2p}) = B(n + 1, \mu + 1, \frac{k}{2p}) \quad (27)$$

Now

$$\frac{\partial}{\partial x} B(n + 1, \mu; x) = (1-x)^n x^{\mu-1}$$

$$\frac{\partial}{\partial \left(\frac{k}{2p}\right)} \left(B(n+1, \mu; \frac{k}{2p}) - 2 B(n+1, \mu+1; \frac{k}{2p}) \right) = \left(1 - \frac{k}{2p}\right)^n \left(\frac{k}{2p}\right)^{\mu-1} \left(1 - 2 \frac{k}{2p}\right)$$

$$= x \left(\frac{x^2-1}{4}\right)^n \quad \text{if } x = 1 - \frac{k}{p} \quad \text{since } n = \mu - 1.$$

Integrating with respect to x we see that

$$B(n+1, n+1; \frac{k}{2p}) - 2 B(n+1, n+2; \frac{k}{2p})$$

is proportional to $(x^2-1)^{n+1}$, for it is zero when $k=0$, i.e.,

when $x=1$. Thus the only non-zero value of k for which

$$(27) \text{ is satisfied is given by } x = -1, \quad k = 2p. \quad (28)$$

10.8 We see, thus that

$$z^{-\frac{1}{2}} e^z \sum_{m=0}^n \binom{n}{m} (-2z)^{-m} \gamma(m+n+1, 2z)$$

and

$$z^{-\frac{1}{2}} e^{-z} \sum_{m=0}^n \binom{n}{m} (2z)^{-m} \gamma(m+n+1, -2z) \quad (29)$$

the only in this form are solutions of the modified Bessel's equation of order $n + \frac{1}{2}$,

and that if the order is not half an odd integer, no solution in this finite form is possible.)

These series can be written in the reverse order, and

we have then a series in ascending powers of z :

$$z^{-n-\frac{1}{2}} e^{\pm z} \sum_{m=0}^n \binom{n}{m} (\pm 2z)^m \gamma(2n-n+1, \pm 2z) \quad (30)$$

We shall show in the next section that both (29) and (30)

give convergent series when n is not a positive or negative

integer, and, in (30), is not half a positive integer, and

that these ~~solutions~~ series are solutions of the equation of

order $n + \frac{1}{2}$.

The general form of the series

11.1 We shall now consider the conditions that

$$F(z) = z^a e^{pz} \sum_{m=-\infty}^{\infty} c_m (kz)^{-m} \gamma(m+\mu, kz) \quad (31)$$

should be a ^{formal} solution of the modified Bessel's equation

$$\frac{d^2 F}{dz^2} + \frac{1}{z} \frac{dF}{dz} - \left(\kappa^2 + \frac{\nu^2}{z^2} \right) F = 0, \quad (32)$$

postponing consideration of its convergence to

Differentiating (31),

$$\begin{aligned} \frac{dF}{dz} &= z^a e^{pz} \sum c_m (a-m) k^{-m} z^{-m-1} \gamma(m+\mu, kz) \\ &+ p z^a e^{pz} \sum c_m k^{-m} z^{-m} \gamma(m+\mu, kz) \\ &+ z^a e^{pz} \sum c_m k^{\mu} z^{\mu-1} e^{-kz} \end{aligned}$$

$$\begin{aligned} \frac{d^2 F}{dz^2} &= z^a e^{pz} \sum c_m (kz)^{-m} \left\{ \frac{(a-m)(a-m-1)}{z^2} + \frac{2(a-m)p}{z} + p^2 \right\} \gamma(m+\mu, kz) \\ &+ z^{a+\mu-1} e^{(p-k)z} k^{\mu} \sum c_m \left(\frac{a-m}{z} + p \right) \\ &+ \left(\frac{a+\mu-1}{z} + p-k \right) z^{a+\mu-1} e^{(p-k)z} k^{\mu} \sum c_m \end{aligned}$$

and substituting in (32) we get :

$$\begin{aligned} z^a e^{pz} \sum c_m (kz)^{-m} \left[\frac{(a-m)^2 - \nu^2}{z^2} + \frac{2(a-m)p}{z} + p^2 - \kappa^2 \right] \gamma(m+\mu, kz) \\ + z^{a+\mu-1} e^{(p-k)z} k^{\mu} \sum c_m \left(\frac{2a-m+\mu}{z} + 2p-k \right) = 0. \end{aligned} \quad (33)$$

Substituting for $\gamma(m+\mu, kz)$ from the series (5), we may re-arrange according to powers of z , since that series is absolutely

convergent. Equation (33) becomes

$$\begin{aligned} \sum c_m \left\{ \frac{(a-m)^2 - \nu^2}{z^2} + \frac{(2a+1-2m)p}{z} + p^2 - \kappa^2 \right\} \left\{ \frac{1}{m+\mu} + \frac{kz}{(m+\mu)(m+\mu+1)} + \frac{(kz)^2}{(m+\mu)(m+\mu+1)(m+\mu+2)} + \dots \right\} \\ + \sum c_m \left\{ \frac{2a+\mu-m}{z} + \frac{2p-k}{z} \right\} = 0, \end{aligned} \quad (34)$$

omitting the factor $z^{a+\mu} e^{(p-k)z} k^\mu$. Equating to zero the coefficients in (34) of $z^{-2}, z^{-1}, z^0, z, z^2, \dots$, we have

$$\sum c_n \left\{ \frac{(a-n)^2 - v^2}{n+\mu} + 2a + \mu - n \right\} = 0, \quad (35)$$

$$\sum c_n \left\{ \frac{(a-n)^2 - v^2}{(n+\mu)(n+\mu+1)} + \frac{(2a+1-2n)p}{n+\mu} + 2p - k \right\} = 0 \quad (36)$$

$$\text{and } \sum c_n \Gamma(n+\mu) \left[\frac{(a-n)^2 - v^2}{\Gamma(n+\mu+s+3)} k^{s+2} + \frac{(2a+1-2n)p}{\Gamma(n+\mu+s+2)} k^{s+1} + \frac{(p^2-k^2)k^s}{\Gamma(n+\mu+s+1)} \right] = 0 \quad (37)$$

$$s = 0, 1, 2, 3, \dots$$

Equation (35) reduces to

$$\{v^2 - (\mu+a)^2\} \sum \frac{c_n}{n+\mu} = 0, \quad (38)$$

and (36) to

$$\{k[(a+\mu)^2 - v^2] + p(2a+1+2\mu)\} \sum \frac{c_n}{n+\mu} - k\{(a+\mu+1)^2 - v^2\} \sum \frac{c_n}{n+\mu+1} = 0, \quad (39)$$

and (37) can be written, using partial fractions, as

$$\begin{aligned} & \sum c_n k^2 \{(a-n)^2 - v^2\} \left\{ \frac{1}{n+\mu} - \frac{s+2}{n+\mu+1} + \frac{\binom{s+2}{2}}{n+\mu+2} - \dots \pm \frac{1}{n+\mu+s+2} \right\} \frac{1}{(s+2)!} \\ & + \sum c_n k p (2a+1-2n) \left\{ \frac{1}{n+\mu} - \frac{s+1}{n+\mu+1} + \frac{\binom{s+1}{2}}{n+\mu+2} - \dots \mp \frac{1}{n+\mu+s+1} \right\} \frac{1}{(s+1)!} \\ & + \sum c_n (p^2 - k^2) \left\{ \frac{1}{n+\mu} - \frac{s}{n+\mu+1} + \frac{\binom{s}{2}}{n+\mu+2} - \frac{\binom{s}{3}}{n+\mu+3} + \dots \pm \frac{1}{n+\mu+s} \right\} \frac{1}{s!} = 0. \end{aligned}$$

I.e., if we assume, from (38), that $v^2 = (\mu+a)^2$,

$$\begin{aligned} & \sum c_n k^2 \left\{ + \frac{2\mu+2a+1}{n+\mu} \frac{1}{(s+1)!} - \frac{2\mu+2a+3}{n+\mu+1} \frac{1}{s!} + \frac{2\mu+2a+5}{n+\mu+2} \frac{1}{2!(s-1)!} - \dots \right\} \\ & + \sum c_n k^2 \left\{ - \frac{2\mu+2a+1}{n+\mu+1} \frac{1}{(s+1)!} + \frac{2(2\mu+2a+2)}{n+\mu+2} \frac{1}{2!s!} - \frac{3(2\mu+2a+3)}{n+\mu+3} \frac{1}{3!(s-1)!} + \dots \right\} \\ & + \sum c_n (p^2 - k^2) \left\{ \frac{1}{n+\mu} \frac{1}{s!} - \frac{1}{n+\mu+1} \frac{1}{(s-1)!} + \frac{1}{n+\mu+2} \frac{1}{2!(s-2)!} - \dots \right\} = 0, \quad (40) \end{aligned}$$

$$s = 0, 1, 2, 3, \dots$$

successively

Using equation (39), and then subtracting each equation of (40)

from the succeeding equation, we see that the equations (40)

are equivalent to

45

$$\sum c_n k^2 \frac{2\mu+2a+j+2}{n+\mu+2+j} - \sum c_n k^2 \frac{2\mu+2a+2j+3}{n+\mu+j+1} + \sum c_n (k^2 - k^2) \frac{j+1}{n+\mu+j} = 0,$$

$$j = 0, 1, 2, 3, \dots \quad (41)$$

These equations, with (39) form a set ^{of} recurrence formulae from which we can determine uniquely the value of C_j / C_0

where

$$C_j = \sum \frac{c_n}{n+\mu+j} \quad (42)$$

for $j = 1, 2, 3, \dots$

11.2 If in writing (40) we do not make the assumption that $\nu^2 = (\mu+a)^2$ we still obtain a set of recurrence relations for C_j , with $\frac{(\mu+a+j+2)^2 - \nu^2}{j+2}$ instead of $2\mu+2a+j+2$ in the numerator of the first term. (The possibility that $\nu^2 = (\mu+a+j)^2$

for some $j > 0$ need not be further considered when we see

that $k^j z^{a+j} e^{t^2} \sum c_n (kz)^{-n} \gamma(n+\mu-j, kz)$

is another way of writing $z^a e^{t^2} \sum c_n (kz)^{-n} \gamma(n+\mu, kz)$ (43)

and the value of a has yet to be found.)

Thus, unless each C_j is zero, (38) gives

$$\nu^2 = (\mu+a)^2 \quad (44)$$

[1.3] Equation (39) can be included in the set (41), with $j = -1$.

We have to solve the equations

$$k^2 C_{j+2} (2N+j+2) - k^2 C_{j+1} (2N+2j+3) + (k^2 - k^2) C_j (j+1) = 0, \quad (45)$$

where $j = -1, 0, 1, 2, \dots$, and N is written for $\mu+a$.

11.4 It may be verified by substitution that the solution is

$$C_j = C_0 \left(\frac{p}{k}\right)^j \left\{ 1 + \frac{j(j-1)}{2} \frac{\kappa^2}{(2N+2)p^2} + \frac{j(j-1)(j-2)(j-3)}{2 \cdot 4} \frac{\kappa^4}{(2N+2)(2N+4)p^4} + \dots \right\} \quad (46)$$

from which we see that if $\kappa = 0$, no solution is possible in which $C_j/C_{j-1} \rightarrow 1$, and $C_j \rightarrow 0$. We shall therefore without loss of generality put $\kappa = 1$. ((46) gives the solution of (45) even when $p = 0$.)

The coefficients are the coefficients in the Hermite polynomial, but with positive sign throughout.

11.5 We shall now assume that C_j tends to zero as j tends to infinity, a condition which is satisfied if c_m is bounded, and tends to zero as m tends to $+\infty$ or to $-\infty$. And assume further that C_j/C_{j-1} tends to 1 as j tends to infinity, a condition which is satisfied if $\sum_{-\infty}^0 c_m$ and $\sum_0^{\infty} c_m$ are convergent, and $c_m = o\left(-\frac{1}{m}\right)$ as $m \rightarrow -\infty$.

We see from (45) that a necessary condition for C_j/C_{j-1} to tend to 1 is

$$k^2 - 2kp + p^2 - 1 = 0;$$

that is,

$$k = p \pm 1. \quad (47)$$

11.6 To find the sufficient condition, we consider the ratio of the $(s+1)^{\text{th}}$ term in (46) to the s^{th} . This ratio is

$$\frac{(j-2s+2)(j-2s+1)}{2s(2N+2s)} \frac{1}{p^2}$$

and the corresponding ratio in the expansion of

$$\left(1 + \frac{1}{p}\right)^j + \left(1 - \frac{1}{p}\right)^j \quad (48)$$

is

$$\frac{(j-2s+2)(j-2s+1)}{(2s-1)2s} \frac{1}{p^2}$$

Now, given any positive number ϵ we can find r and s ,

1) If $p = 0$, $C_j = 0$, all odd j , so that C_j/C_{j-1} cannot tend to 1. We see thus that no solution is possible with $p = 0$.

functions of j , such that $r < s$, $r \rightarrow \infty$, $(s - r)/r \rightarrow 0$ as $j \rightarrow \infty$, and such that the sum of the terms between the r^{th} and the s^{th} in (48) differs from the total sum by less than ϵ times that sum, and it is evident from the argument ^{above} of §7.6 that the same functions r and s will serve for the series in (46) as for (48). (Since the standard deviation in a binomial distribution varies as \sqrt{n} , $s - r = O(\sqrt{n})$, while $r = O(n)$.) Now the ratio of a typical term in (46), that is, in the series for C_j/C_0 , to the corresponding term in the series for C_{j-1}/C_0 is $(p/k)[j/(j - 2s)]$, and the corresponding ratio for (48) is $j/(j - 2s)$, so that

$$\frac{C_j}{C_{j-1}} \sim \frac{p}{k} \cdot \frac{j}{j-2s} ;$$

and

$$\frac{\left(1 + \frac{1}{k}\right)^j + \left(1 - \frac{1}{k}\right)^j}{\left(1 + \frac{1}{k}\right)^{j-1} + \left(1 - \frac{1}{k}\right)^{j-1}} \sim \frac{j}{j-2s}$$

Thus the condition that C_j/C_{j-1} should tend to 1 is

$$\frac{p}{k} \cdot \frac{\left(1 + \frac{1}{k}\right)^j + \left(1 - \frac{1}{k}\right)^j}{\left(1 + \frac{1}{k}\right)^{j-1} + \left(1 - \frac{1}{k}\right)^{j-1}} \rightarrow 1 \quad \text{as} \quad j \rightarrow \infty.$$

If $k = p + 1$, this condition is (real part of p) > 0 , and if $k = p - 1$, (real part of p) < 0 .)

11.7 If any function of z is a solution of Bessel's equation, the same ^{function} ~~solution~~ of $-z$ is also a solution, and we need therefore only discuss values of p with a positive real part; the form of the theorem for a p with a negative real part can then be deduced by changing the sign of z . For p with positive real part, then, the
 1) If the real part of $p = 0$, the condition cannot be satisfied.

158

necessary and sufficient condition that C_j/C_{j-1} should tend to 1 as j tends to infinity is $k = p + 1$.

11.8 Comparing (46) and (48) we see that the ratio of the $(s + 1)^{\text{th}}$ term in C_j/C_0

to the $(s + 1)^{\text{th}}$ in $\left(\frac{p}{r}\right)^j \left\{ \left(1 + \frac{1}{p}\right)^j + \left(1 - \frac{1}{p}\right)^j \right\}$

is

$$\frac{1 \cdot 3 \cdot 5 \cdot 7 \cdots (2s-1)}{(2N+2)(2N+4) \cdots (2N+2s)}$$

and hence

$$\begin{aligned} \frac{C_j}{C_0} &\sim \left(\frac{p}{r}\right)^j \left\{ \left(1 + \frac{1}{p}\right)^j + \left(1 - \frac{1}{p}\right)^j \right\} \frac{1 \cdot 3 \cdot 5 \cdot 7 \cdots (2s-1)}{(2N+2)(2N+4) \cdots (2N+2s)} \\ &\sim \frac{1 \cdot 3 \cdot 5 \cdot 7 \cdots (2s-1)}{(2N+2)(2N+4) \cdots (2N+2s)} \end{aligned} \quad (49)$$

where s can be taken to be the s of §11.6 or the r of that §11.6 or any number between them. (We may take s such that the s^{th} term in (46) or (48) is the greatest, so that $s \sim j/(p + 1)$.) Hence the necessary and sufficient condition that C_j should tend to zero as j tends to infinity is

$$\text{real part of } \mu + a = \text{real part of } N > -\frac{1}{2}.$$

12.1 We shall now show that if the numbers c_m are such as to satisfy the recurrence relations

$$(m + \mu - 1)(p^2 - 1)c_m + (2a - 2m + 3)kpc_{m-1} - (\mu + 2a - m + 2)k^2c_{m-2} = 0, \quad (50)$$

$$m = \dots, -2, -1, 0, +1, +2, \dots,$$

then the equation (45) is satisfied, $j = -1, 0, 1, 2, \dots$.

The proof depends on the fact that

$$C_j = \sum_{-\infty}^{\infty} \frac{c_m}{\mu+m+j} = \sum_{-\infty}^{\infty} \frac{c_{m-1}}{\mu+m-1+j} = \dots = \sum_{-\infty}^{\infty} \frac{c_{m-j}}{\mu+m} = \dots,$$

$$C_{j+1} = \sum_{-\infty}^{\infty} \frac{c_m}{\mu+m+j+1} = \dots = \sum_{-\infty}^{\infty} \frac{c_{m-j-1}}{\mu+m} = \dots, \dots,$$

$$\sum_{-\infty}^{\infty} c_m = \sum_{-\infty}^{\infty} c_{m-1} = \dots = \sum_{-\infty}^{\infty} c_{m-j} = \dots,$$

provided all these series converge, for equations (50) can be written

$$(m+\mu-j-1)(p^2-1)c_{m-j} + (2\alpha-2m+2j+3)k^2 c_{m-j-1} - (\mu+2\alpha-m+j+2)k^2 c_{m-j-2} = 0;$$

and dividing this by $\mu+m$ and adding for all values of m

$$-(j+1)(p^2-1)C_j + (2\mu+2\alpha+2j+3)k^2 C_{j+1} - (2\mu+2\alpha+j+2)k^2 C_{j+2} + (p^2-1)\sum_{-\infty}^{\infty} c_{m-j} - 2k^2 \sum_{-\infty}^{\infty} c_{m-j-1} + k^2 \sum_{-\infty}^{\infty} c_{m-j-2} = 0 \quad (52)$$

which is equivalent to (45). This analysis holds even if $\sum c_m$ diverges, provided that $c_m \rightarrow 0$, and that all C_j converge

12.2

It has thus been proved that for any values of a and p , if the real part of p is not zero,

$$z^\alpha e^{kz} \sum_{m=-\infty}^{\infty} c_m (kz)^{-m} y(m+\mu, kz) \quad (53)$$

is a formal solution of Bessel's differential equation in the form

$$\frac{d^2 F}{dz^2} + \frac{1}{z} \frac{dF}{dz} - \left(1 + \frac{\nu^2}{z^2}\right) F = 0$$

provided that $k = p + 1$ if the real part of p is positive,

and $k = p - 1$, if the real part of p is negative, that

$(\mu + a)^2 = \nu^2$ that real part of $\mu + a > -\frac{1}{2}$, and that the coefficients c_m satisfy the

recurrence relation (50):

$$(m+\mu-1)(p^2-1)c_m + (2\alpha-2m+3)k^2 c_{m-1} - (\mu+2\alpha-m+2)k^2 c_{m-2} = 0.$$

Convergence of the series

13.1 To discuss the convergence of this series, write (50) in the form

$$(p-1)\left\{(m+\mu-1)c_m - (m-a-\frac{3}{2})c_{m-1}\right\} = (p+1)\left\{(m-a-\frac{3}{2})c_{m-1} - (m-\mu-2a-2)c_{m-2}\right\},$$

with the real part of p positive.

If $p \neq 1$ this can be written

$$\frac{(m+\mu-1)c_m - (m-a-\frac{3}{2})c_{m-1}}{(m-a-\frac{3}{2})c_{m-1} - (m-\mu-2a-2)c_{m-2}} = \frac{p+1}{p-1};$$

and if c_m is bounded as $m \rightarrow \pm \infty$

$$\frac{(m+\mu-1)c_m - (m-a-\frac{3}{2})c_{m-1}}{m c_m - (m-\mu-a-\frac{1}{2})c_{m-1}} \rightarrow 1,$$

and

$$\frac{(m-a-\frac{3}{2})c_{m-1} - (m-\mu-2a-2)c_{m-2}}{(m-1)c_{m-1} - (m-1-\mu-a-\frac{1}{2})c_{m-2}} \rightarrow 1,$$

so that

$$\frac{m c_m - (m-\mu-a-\frac{1}{2})c_{m-1}}{(m-1)c_{m-1} - (m-1-\mu-a-\frac{1}{2})c_{m-2}} \rightarrow \left\{ \frac{p+1}{p-1} \right\}.$$

But $|(p+1)/(p-1)| > 1$, and hence, as $m \rightarrow +\infty$,

$$c_m - \left(1 - \frac{\mu+a+\frac{1}{2}}{m}\right)c_{m-1} \rightarrow \infty.$$

The series therefore does not converge if $p \neq 1$, except in the case in which (50) can be satisfied by a sequence of values for c_m in which $c_m = 0$, for all $m > m_0$.

Without loss of generality we can say $c_m = 0$, all $m > 0$,

and the condition can be satisfied if and only if

$$\mu + 2a = 0, \quad a = -(\mu + a) \text{ or } -\nu.$$

13.2 But if $p = 1$, equation (50) can be written

$$(2a - 2m + 1) c_m = 2(\mu + 2a - m + 1) c_{m-1} \quad (54)$$

from which it is evident that $\sum c_m$ converges as $m \rightarrow \pm\infty$ (56)

if real part of $\mu + a > +\frac{1}{2}$

while $\sum \frac{c_m}{m+\mu}$ converges if

$$\text{real part of } \mu + a > -\frac{1}{2}$$

so that, by (54), the series determined by (54) converges

if real part of $\mu + a > -\frac{1}{2}$

The series determined by (54) is

$$\sum_{n=-\infty}^{\infty} (-)^n z^a e^{\pm 2z} \frac{(\pm 2z)^m \Gamma(N + \frac{1}{2})}{\Gamma(N + \frac{1}{2} - m) \Gamma(m - a + \frac{1}{2})} \gamma(N - a + m, \pm 2z) \quad (55)$$

where $N = \mu + a = \pm \nu$.



Uniqueness of the solution

14.1

If there are more than one set of values of c_n satisfying the conditions for a solution then, by subtracting one such set from another we obtain a set of values of a_n such that

$$\sum_{n=-\infty}^{\infty} \frac{a_n}{\mu+n+i} = 0, \quad \text{where } i \text{ is } \begin{matrix} \text{zero or} \\ \text{any positive integer} \end{matrix} \quad (56)$$

We shall show, with certain assumptions as to the convergence of the series, that no solution of these equations ~~is~~ possible except that ~~is~~ in which $a_n = 0$, all n .

14.2

Consider first the simple case in which the sequence a_n is from 0 to ∞ only and not from $-\infty$ to $+\infty$.

~~and suppose that $\sum a_n$ converges.~~ Let $f(x)$ represent $\sum_0^{\infty} a_n x^n$, which is a continuous function of x , $0 \leq x < 1$.

Equations (56) are then equivalent to

$$\int_0^1 x^{\mu+i-1} f(x) dx = 0, \quad i = \begin{matrix} \text{zero or} \\ \text{any positive integer} \end{matrix}$$

from which it follows that $f(x)$ is identically zero,

$0 \leq x < 1$, and therefore that $a_n = 0$, all n .

14.3

Before proving the proposition in its general form we shall prove a lemma from which we can find the value of the determinant formed by the coefficients of any n of the a_n in any n of the equations (56). It is :

$$\begin{aligned}
 & \left| \begin{array}{cccc} \frac{1}{k} & \frac{1}{k+l_1} & \frac{1}{k+l_1+l_2} & \dots & \frac{1}{k+\sum_1^n l_i} \\ \frac{1}{k+a_1} & \frac{1}{k+a_1+l_1} & \frac{1}{k+a_1+l_1+l_2} & \dots & \frac{1}{k+a_1+\sum_1^n l_i} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{k+\sum_1^n a_i} & \frac{1}{k+\sum_1^n a_i+l_1} & \frac{1}{k+\sum_1^n a_i+l_1+l_2} & \dots & \frac{1}{k+\sum_1^n a_i+\sum_1^n l_i} \end{array} \right| \\
 &= \frac{\prod_{r=1}^n \left\{ \prod_{s=r}^n (a_r + a_{r+1} + a_{r+2} + \dots + a_s) \right\} \prod_{r=1}^n \left\{ \prod_{s=r}^n (l_r + l_{r+1} + l_{r+2} + \dots + l_s) \right\}}{\prod_{r,s=0}^n \left(k + \sum_{i=1}^r a_i + \sum_{i=1}^s l_i \right)} \quad (57)
 \end{aligned}$$

The proof is by induction. The proposition is clearly true when $n = 1$ or 2 . Assuming it is true for $n - 1$, if Δ denotes the determinant of order n ,

$$\begin{aligned}
 \Delta \cdot \prod_{r,s=0}^n \left(k + \sum_{i=1}^r a_i + \sum_{i=1}^s l_i \right) &= k(k+a_1)(k+a_1+a_2) \dots (k+\sum_{i=1}^n a_i) \times \\
 & \left[\frac{(k+l_1)(k+l_1+l_2) \dots (k+\sum_1^n l_i)}{k} a_2(a_2+a_3) \dots (a_2+a_3+\dots+a_n) a_3(a_3+a_4) \dots (a_3+\dots+a_n) \right. \\
 & \quad \cdot a_4(a_4+a_5) \dots (a_4+a_5+\dots+a_n) \dots a_n \\
 & - \frac{(k+a_1+l_1)(k+a_1+l_1+l_2) \dots (k+a_1+\sum_1^n l_i)}{k+a_1} (a_1+a_2)(a_1+a_2+a_3) \dots (a_1+a_2+\dots+a_n) \\
 & \quad \cdot a_3(a_3+a_4) \dots (a_3+\dots+a_n) a_4(a_4+a_5) \dots (a_4+\dots+a_n) \dots a_n \\
 & \left. + \dots + (-)^r \frac{(k+\sum_1^r a_i+l_1)(k+\sum_1^r a_i+l_1+l_2) \dots (k+\sum_1^r a_i+\sum_1^n l_i)}{k+\sum_1^r a_i} \cdot a_1(a_1+a_2) \dots \sum_1^{r-1} a_i \sum_1^{r-1} a_i \dots \sum_1^n a_i \right] \times \\
 & \quad \cdot a_2(a_2+a_3) \dots \sum_2^{r-1} a_i \sum_2^{r-1} a_i \dots \sum_1^n a_i \\
 & \times l_1(l_1+l_2) \dots (l_1+\dots+l_n) l_2(l_2+l_3) \dots (l_2+\dots+l_n) \dots \dots l_n
 \end{aligned}$$

This is a polynomial of degree $2n$ in k , and when

$$k = 0, -a_1, -(a_1 + a_2), \dots, -(a_1 + a_2 + \dots + a_n)$$

this has the value

$$\prod_{r=1}^n \prod_{s=r}^n (a_r + a_{r+1} + \dots + a_s) \cdot \prod_{r=1}^n \prod_{s=r}^n (b_r + b_{r+1} + \dots + b_s) ;$$

but $\Delta \cdot \prod_{r,s=0}^n (k + \sum_i a_i + \sum_i b_i)$ is symmetrical

in a and b, and it must therefore have the same value when $k = -b_1, -(b_1 + b_2), \dots, -(b_1 + b_2 + \dots + b_n)$, and thus, provided these $2n + 1$ values are all different, the proposition is proved for all values of k . The expression () is ~~is~~ a polynomial of degree not greater than n in any one of the b 's, which has the value () for all except a finite number of values for the b , and it is therefore true for all values of that b ~~is~~, and similarly of each of the other b 's.

The lemma is thus proved.

14.32 A corollary obtained by putting all the a 's and b 's equal to 1 is :

$$\begin{vmatrix} \frac{1}{k} & \frac{1}{k+1} & \frac{1}{k+2} & \dots & \frac{1}{k+n} \\ \frac{1}{k+1} & \frac{1}{k+2} & \frac{1}{k+3} & \dots & \frac{1}{k+n+1} \\ \frac{1}{k+2} & \frac{1}{k+3} & \frac{1}{k+4} & \dots & \frac{1}{k+n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{k+n} & \frac{1}{k+n+1} & \frac{1}{k+n+2} & \dots & \frac{1}{k+2n} \end{vmatrix} = \{1!2!3!\dots n!\}^2 \Pi \quad (58)$$

(Cauchy's double alternant)
A.C.A.

where \prod denotes the product of all the ~~element~~ elements in the determinant. The other factor is :-

$$\{1! 2! 3! \dots n!\}^2 = \{G(n+2)\}^2$$

the square of a Barnes's G-function.

The minor of the element in the $(r + 1)$ th row, and the $(s + 1)$ th column in the last determinant, can be found from the lemma, from a determinant of order $n + 1$, putting

$$a_1 = a_2 = \dots = a_{r-1} = 1, a_r = 2, a_{r+1} = \dots = a_n = 1,$$

$$b_1 = b_2 = \dots = b_{s-1} = 1, b_s = 2, b_{s+1} = \dots = b_n = 1.$$

Let Δ_{rs} denote the required minor (without regard to sign), and let Π_{rs} denote the product of all its elements. Then

$$\frac{\Delta_{rs}}{\Pi_{rs}} = \binom{n}{r} \binom{n}{s} \{(n-1)! (n-2)! \dots 3! 2! 1!\}^2$$

and
$$\frac{\Delta_{rs}}{\Delta} = \frac{1}{r!(n-r)! s!(n-s)!} \cdot \frac{\Pi_{rs}}{\Pi} = \frac{\prod_{i=0}^n (k+r+i) \prod_{i=0}^n (k+s+i)}{r!(n-r)! s!(n-s)! (k+r+s)}$$

14.4 Now the equations (56) can be written

$$\left. \begin{aligned} \frac{a_{-l}}{\mu-l} + \frac{a_{-l+1}}{\mu-l+1} + \dots + \frac{a_{-1}}{\mu-1} + \frac{a_0}{\mu} + \dots + \frac{a_l}{\mu+l} &= \epsilon_0, \\ \frac{a_{-l}}{\mu-l+1} + \frac{a_{-l+1}}{\mu-l+2} + \dots + \frac{a_{-1}}{\mu} + \frac{a_0}{\mu+1} + \dots + \frac{a_l}{\mu+l+1} &= \epsilon_1, \\ &\dots \\ &\dots \end{aligned} \right\} (59)$$

where ϵ_r denotes
$$\sum_{i=-\infty}^{-l-1} \frac{a_i}{\mu+i+r} + \sum_{i=l+1}^{\infty} \frac{a_i}{\mu+i+r}$$

Taking $n + 1$ of these equations, with $n = l + h$, eliminate all the a 's except a_0 :

The argument of §14.2 shows that for any fixed l , if all the ϵ_r 's tend to zero as h tends to infinity, then a_0 tends to ~~not~~ zero. That is to say, if, given l and given ϵ we can find h_0 such that $\epsilon_r < \epsilon$ all r , all $n > h_0$, then there exists a function of ϵ , $f(\epsilon)$ such that $f(\epsilon)$ tends to zero as ϵ tends to zero, and such that $|a_0| < f(\epsilon)$. ^{each term in} Now the expression for a_0 given by (60) is a function of which the modulus decreases as l increases if the real part of $\mu + r > 0$, as it is for all except a finite number of terms. ~~Thus if for all l and h greater than some K each of the ϵ_r 's is less than ϵ , the sum of that finite number of terms will be less than ϵ times some constant.~~

And instead of taking the first $n + 1$ equations of (59) we can take those from the j^{th} to the $j + n^{\text{th}}$, where j is sufficiently large for the real part of $\mu + j$ to be positive. If μ is written for $\mu + j$, (60) will still represent the equations to be considered.

Now we assume that $\sum \frac{a_n}{\mu + n + j}$ converges uniformly ^{all j} as it will if $\sum \frac{a_n}{\mu + n}$ converges, and hence given any ϵ we can find l_0 and h_0 such that for all $l > l_0$ and $h > h_0$, all the ϵ_r 's are less than ϵ . The argument above then shows that for $l = l_0$ and $h > h_0$ the value given by (60) for a_0 will be less than $f(\epsilon)$, and hence also for $l > l_0$ and $h > h_0$, $a_0 < f(\epsilon)$. But $f(\epsilon)$ can be made as small as we like by making ϵ sufficiently small. And therefore $a_0 = 0$. The argument can of course be applied to a_r for any finite r , replacing μ by $\mu + r$

We see then that if all the series $\sum \frac{c_m}{\mu+m+j}$ converge as they must if a series $\sum c_m w^{-m} y(\mu+m, w)$ is to converge, ~~then~~ and if $\sum \frac{c_m}{\mu+m+j} = 0$, all j , then all the c 's are 0. It follows then that for given a and p there is not more than one sequence of values of c_m which will make

$$z^a e^{bz} = \sum c_m (kz)^{-m} y(\mu+m, kz)$$

a convergent series, and a solution of the Bessel equation

The corresponding expressions with (complete) Gamma functions

It can readily be shown that the expression (55) remains a formal solution of the modified Bessel equation when the incomplete Gamma functions are replaced by the corresponding Gamma functions. For if in equation (31) $\Gamma(m + \mu)$ is put for $\gamma(m + \mu, kz)$, then equation (33) becomes

$$\sum c_n (kz)^{-n} \left[\frac{(a-n)(a-n-1) + (a-n) - N^2}{z^2} + \frac{2(a-n)p + k + p^2 - 1}{z} \right] \Gamma(m+n) = 0 \quad (61)$$

and putting $c'_n = c_n \cdot \Gamma(m + \mu)$, this is

$$c'_n (p^2 - 1) + p k c'_{n-1} (2a - 2n + 1) + k^2 \{ (a - n + 2)^2 - \nu^2 \} = 0, \quad (62)$$

which agrees exactly with (50), showing that if

$$z^a e^{pz} \sum_{-\infty}^{\infty} c'_n z^n \quad (63)$$

is a formal solution of the equation then

$$z^a e^{pz} \sum_{-\infty}^{\infty} c'_n z^n \frac{\gamma(m + \nu - a, (p \pm 1)z)}{\Gamma(m + \nu - a)} \quad (64)$$

will also be a formal solution. The series (63) will diverge as $m \rightarrow \infty$, and will only give a convergent series when $a = \nu / k$ ^{+ an integer} and the series does not extend to $+\infty$.

Though this result is not lacking in interest, it is the case in which $p^2 = 1$ that is most important, and we see thus that if

$$z^a e^z \sum_{-\infty}^{\infty} c'_n z^n \quad (65)$$

is a formal solution of the equation, then

$$z^a e^{-z} \sum_{-\infty}^{\infty} c_n z^n = \frac{\gamma(m+\nu-a, 2z)}{\Gamma(m+\nu-a)}$$

(66)

is also a solution and is a convergent series, which (65) is not unless a differs from ν by a whole number.

§16

Expression of the series found in terms of $I_\nu(z)$

Substituting in the series (55) for $y(m + N - a, \pm 2z)$ from the series (6), we can re-arrange in powers of z since the series are absolutely convergent. The principal part in the neighbourhood of $z = 0$ is

$$(\pm 2)^{N-a} z^N e^{\mp 2z} \sum_{-\infty}^{\infty} \frac{\Gamma(N+\frac{1}{2})}{\Gamma(N+\frac{1}{2}-m)\Gamma(m-a+\frac{1}{2})} \cdot \frac{(-1)^m}{N-a+m}, \quad (67)$$

while the principal part of $I_\nu(z)$ is also of the order of z^N

Comparison shows that $I_\nu(z) = \text{const.} \cdot z^N e^{\pm 2z} \frac{\Gamma(N-a) (-2z)^{-m}}{\Gamma(N+\frac{1}{2}-m)\Gamma(m-a+\frac{1}{2})} y(\frac{z-a+m}{2z})$

and in order to find the constant we must find the value of

$$\sum_{-\infty}^{\infty} \frac{\Gamma(N-a)}{\Gamma(N+\frac{1}{2}-m)\Gamma(m-a+\frac{1}{2})} \frac{(-1)^m}{N-a+m} \quad (68)$$

§17

Equation (46) which gives $C_j = \sum \frac{c_m}{\mu+m+j}$ in terms of

$C_0 = \sum \frac{c_m}{\mu+m}$ reduces when $p = 1$ to

$$\frac{C_j}{C_{j-1}} = \frac{N+j-\frac{1}{2}}{2N+j}$$

and if C_0 is chosen to be equal to $B(N + \frac{1}{2}, N + \frac{1}{2})$, this gives ~~XXXXXXXXXXXX~~ $C_j = B(N+j+\frac{1}{2}, N+\frac{1}{2})$

and thus for any a and for all j ,

$$\sum_{m=-\infty}^{\infty} (-1)^m \frac{\Gamma(N+\frac{1}{2})}{\Gamma(N+\frac{1}{2}-m)\Gamma(m-a+\frac{1}{2})} \cdot \frac{1}{N-a+m+j} = K \cdot B(N+j+\frac{1}{2}, N+\frac{1}{2})$$

where K may be a function of N or of a , or both, but not of j .

If $a = -\frac{1}{2}$, the expression (68) becomes

$$\sum_{m=0}^{\infty} \binom{N+\frac{1}{2}}{m} \frac{(-1)^m}{N+\frac{1}{2}+m}$$

and the argument of §10.7 shows that this is

$$B(N+\frac{1}{2}, N+\frac{1}{2})$$

and thus in this instance $K = 1$.

The determination of K in general, remains to be completed.

Conclusion

We have seen that the expression derived from the well-known asymptotic expansion for $\Gamma_\nu(z)$ is a special case of an expansion in positive and negative powers of z with incomplete Gamma functions replacing the Gamma functions, giving a convergent in place of a divergent series if the real part of $\nu > -\frac{1}{2}$. Various other results have been found; one interesting theorem discovered incidently giving the evaluation of a determinant, which in the special case when each row and each column form a harmonic progression gives a form in Barnes's G-functions.

Chapter V

~~TWO~~ TWO-VARIATE DISTRIBUTIONS OF DIFFERENT TYPES IN
THE TWO VARIATES

I) A TWO-VARIATE DISTRIBUTION, OF NORMAL TYPE
IN ONE VARIATE, AND GAMMA TYPE IN THE OTHER

§18.1

In order to consider a distribution Normal in one variate, and Gamma type in the other, we will take variates X and y with a normal correlated distribution, and then find the moment generating function for the distribution of y and the variance of X .

$$\text{Put } x = \frac{1}{2} X^2, \quad dX dy = \frac{1}{\sqrt{2}} x^{-\frac{1}{2}} dx dy,$$

$$\varphi(X, y) dX dy = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{X^2 - 2\rho Xy + y^2}{2(1-\rho^2)}\right\} dX dy; \quad (1)$$

$$\psi(x, y) dx dy = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{2x - 2\rho y\sqrt{2x} + y^2}{2(1-\rho^2)}\right\} \frac{dx dy}{\sqrt{2x}}$$

(2)

+ a similar expression with the opposite sign for ρ , since $-X, y$ and $+X, y$ give the same pair of values for x and y . Thus the required moment generating function is

$$G(\alpha, \beta) = \frac{1}{2\sqrt{2}\pi\sqrt{1-\rho^2}} \int_0^\infty \int_{-\infty}^\infty \exp\left\{-\frac{x(1-\alpha(1-\rho^2)) - \rho\sqrt{2x} \cdot y + \frac{1}{2}y^2 - \beta y(1-\rho^2)}{1-\rho^2}\right\} \cdot \frac{dx dy}{\sqrt{x}} + \text{s. i. } (\rho).$$

63

And
$$\int_{-\infty}^{\infty} \exp \left\{ - \frac{[y - \rho\sqrt{2x} - \beta(1-\rho^2)]^2 + [\rho\sqrt{2x} + \beta(1-\rho^2)]^2}{2(1-\rho^2)} \right\} dy$$

$$= \frac{2\sqrt{1-\rho^2}}{\sqrt{\pi}} \exp \frac{\rho\sqrt{2x} + \beta(1-\rho^2)}{2(1-\rho^2)},$$

so that

$$G(\alpha, \beta) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \exp \left\{ - \frac{x[1-\alpha(1-\rho^2)] - \rho^2 x - \rho\beta(1-\rho^2)\sqrt{2x} - \frac{\beta^2}{2}(1-\rho^2)^2}{1-\rho^2} \right\} \frac{dx}{\sqrt{x}}$$

+ s. i. (P)

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp \left\{ - \frac{x^2[1-\alpha(1-\rho^2)] - \rho^2 x^2 - \rho\beta(1-\rho^2)\sqrt{2x} - \beta^2(1-\rho^2)^2}{2(1-\rho^2)} \right\} dx$$

$$= \frac{1}{\sqrt{1-\alpha}} \exp \left\{ \frac{1}{2} \beta^2 \left(1 + \frac{\alpha\rho^2}{1-\alpha} \right) \right\} \quad (3)$$

which of course reduces to $(1-\alpha)^{-\frac{1}{2}} \exp \frac{1}{2} \beta^2$ when $\rho = 0$.

18.2

By the general theorem of §3.1 it follows that the moment generating function for the distribution of x and y , if $2x = X_1^2 + X_2^2 + X_3^2 + \dots + X_n^2$, $y = y_1 + y_2 + y_3 + \dots + y_n$, (X_i, y_i) , (X_2, y_2) , \dots , (X_n, y_n) being sample pairs from a normal population, is

$$(1-\alpha)^{-\frac{n}{2}} \exp \left\{ \frac{1}{2} n \beta^2 \left(1 + \frac{\alpha\rho^2}{1-\alpha} \right) \right\} \quad (4)$$

or, if with a change of scale for y , and $\beta = n/2$,

$$(1-\alpha)^{-\frac{1}{2}} \exp \left\{ \frac{1}{2} \beta^2 \left(1 + \frac{\alpha\rho^2}{1-\alpha} \right) \right\} \quad (5)$$

which can be expanded as a series :

$$(1-\alpha)^{-p} e^{\frac{1}{2}\beta^2} \left\{ 1 + \frac{1}{2} p^2 \frac{\alpha\beta^2}{1-\alpha} + \frac{p^4}{2^2 \cdot 2!} \left(\frac{\alpha\beta^2}{1-\alpha} \right)^2 + \frac{p^6}{2^3 \cdot 3!} \left(\frac{\alpha\beta^2}{1-\alpha} \right)^3 + \dots \right\} \quad (6)$$

from which it is clear that the frequency function for x and y can be expressed with a series linear in Hermite polynomials, and ⁱⁿ the appropriate Laguerre polynomials :

$$\frac{x^{p-1} e^{-x}}{\Gamma(p)} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \left\{ 1 + \frac{1}{2} p^2 H_2(y) L_1(x, p) \frac{\Gamma(p)}{\Gamma(p+1)} + \frac{\left(\frac{1}{2}p^2\right)^2}{2!} H_4(y) L_2(x, p) \frac{\Gamma(p)}{\Gamma(p+2)} + \dots \right\} \quad (7)$$

Hermite polynomials of even degree only occur : it was evident from the way the distribution arose that for any given x , y must be an even function. The curve of regression of x on y is the parabola

$$\begin{aligned} x &= 1 + \frac{1}{2} p^2 H_2(y) \\ &= 1 - \frac{1}{2} p^2 + \frac{1}{2} p^2 y^2. \end{aligned}$$

The convergence of the series (7) follows as in §5.2, using §1.452 §1.751

§ 18.21 We can now deduce an expression for a frequency function for two variates, each of which has a Gamma-type distribution, but, unlike that discussed in *Ch II*, with different indexes for the two variates.

We have seen that the moment generating function for two variates x and y of which x is the mean variance of $X_1, X_2, X_3, \dots, X_n$ and y is the mean of $y_1, y_2, y_3, \dots, y_n$, and $X_1, y_1; X_2, y_2; \dots$, are sample pairs from a normally correlated population, is

$$(1-\alpha)^{-\frac{n}{2}} \exp\left\{\frac{1}{2}\beta^2\left(1+\frac{\alpha\rho^2}{1-\alpha}\right)\right\} = (1-\alpha)^{-\frac{n-1}{2}} \left[\exp\left\{\frac{1}{2}\beta^2\left(1+\frac{\alpha\rho^2}{1-\alpha}\right)\right\} \cdot (1-\alpha)^{-\frac{1}{2}} \right], \quad (8)$$

and the second factor in this expression is the m.g.f. for $\frac{1}{2} X^2$ and y , and we know that if we change the variates to $\frac{1}{2} X^2$ and $\frac{1}{2} y^2$, this moment generating function will become

$$\{(1-\alpha)(1-\beta) - \alpha\beta\rho^2\}^{-\frac{1}{2}} \quad (9)$$

Hence, by theorem it follows that with the same transformation (8) will become

$$(1-\alpha)^{-\frac{n-1}{2}} \{(1-\alpha)(1-\beta) - \alpha\beta\rho^2\}^{-\frac{1}{2}} \quad (10)$$

In other words, for two variates of which one is the mean variance (in a sample of n) of one of two normally-correlated variates, and the other is the variance of the mean of the second, the moment generating function is

$$(1-\alpha)^{-\frac{n}{2}} (1-\beta)^{-\frac{1}{2}} \left\{ 1 - \rho^2 \frac{\alpha\beta}{(1-\alpha)(1-\beta)} \right\}^{-\frac{1}{2}} \quad (11)$$

We can now take a further step, and consider the

mean of m such samples, so that the m.g.f. becomes

$$(1-\alpha)^{-\frac{m\alpha}{2}} (1-\beta)^{-\frac{m}{2}} \left\{ 1 - \rho^2 \frac{\alpha\beta}{(1-\alpha)(1-\beta)} \right\}^{-\frac{m}{2}} \quad (12)$$

We see that ^{if} the general form

$$(1-\alpha)^{-M} (1-\beta)^{-N} \left\{ 1 - \rho^2 \frac{\alpha\beta}{(1-\alpha)(1-\beta)} \right\}^{-P} \quad (13)$$

is ^{to be} the moment generating function for a pair of variates x and y , ^{then} ~~for which~~ the frequency functions ^{will be} ~~are~~

$$\varphi(x) = x^{M-1} e^{-x} / \Gamma(M) \quad \text{and} \quad \psi(y) = y^{N-1} e^{-y} / \Gamma(N)$$

respectively. Since the expression (13) can be expanded :

$$(1-\alpha)^{-M} (1-\beta)^{-N} \left\{ 1 + P\rho^2 \frac{\alpha}{1-\alpha} \frac{\beta}{1-\beta} + \frac{P(P+1)}{2!} \rho^4 \left(\frac{\alpha}{1-\alpha} \frac{\beta}{1-\beta} \right)^2 + \dots \right\}, \quad (14)$$

it follows that the frequency function can be expanded

$$\varphi(x)\psi(y) \left\{ 1 + \frac{P\rho^2 L_1(x, M) L_1(y, N)}{MN} + \frac{P(P+1)\rho^4 L_2(x, M) L_2(y, N)}{2! M(M+1) N(N+1)} + \dots \right\}. \quad (15)$$

The convergence of this series if $|\rho| < 1$ follows exactly as in §18.1. But the line of regression given by (15) is $y = N - P\rho^2 + P\rho^2 x/M$, which gives negative values for y for some values of x if $P\rho^2 > N$. Since the distribution is only over positive values, both of y and of x , this indicates that if $P\rho^2 > N$, or similarly if $P\rho^2 > M$, the function represented by (15) has negative values for some positive values of x and y , and hence cannot then be a frequency function.

68

Chapter VI

DISTRIBUTION FUNCTIONS IN ANY NUMBER OF VARIATES EXPRESSED IN TERMS OF ORTHOGONAL POLYNOMIALS

19.0 Distribution functions in two variates have been expressed by means of series bilinear in the appropriate orthogonal polynomials, in the case of the Normal Distribution, by Mehler, 1866, of various discontinuous distributions by Aitken and Gonin, and Campbell, and of a distribution in which each of the two variates has Gamma-type distribution in chapter II of this thesis.

The use of moment generating functions gives a convenient method of deriving the analogous expressions, when there are n variates instead of two.

The normal distribution

19.1 The moment generating function of a normal distribution in $x_1, x_2, x_3, \dots, x_n$, the unit for each variate being chosen to be the corresponding standard deviation, is

$$\exp \frac{1}{2} (\alpha_1^2 + \alpha_2^2 + \dots + \alpha_n^2 + 2\rho_{12} \alpha_1 \alpha_2 + \dots) \quad (1)$$

Exp $\frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \dots + \alpha_n^2)$ is the moment generating function of $\varphi(x_1) \cdot \varphi(x_2) \dots \varphi(x_n)$ where $\varphi(x)$ denotes $\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$; and $\alpha_1^{k_1} \alpha_2^{k_2} \dots \alpha_n^{k_n} \exp\frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \dots + \alpha_n^2)$ is the moment generating function of

$$(-D_1)^{k_1} (-D_2)^{k_2} \dots (-D_n)^{k_n} [\varphi(x_1) \varphi(x_2) \dots \varphi(x_n)] ,$$

where D_i denotes $\frac{\partial}{\partial x_i}$; that is, of

$$H_{k_1}(x_1) H_{k_2}(x_2) \dots H_{k_n}(x_n) \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) .$$

The moment generating function (1) can be written

$$\exp\left\{\frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \dots + \alpha_n^2)\right\} \left[1 + \sum P_{ij} \alpha_i \alpha_j + \frac{1}{2!} (\sum P_{ij} \alpha_i \alpha_j)^2 + \dots\right]$$

where \sum ~~denotes~~ denotes summation over all values of the suffixes i, j from 1 to n , excluding $i = j$, and this is clearly the moment generating function of the distribution function represented by the series

$$\begin{aligned} &\varphi(x_1) \varphi(x_2) \dots \varphi(x_n) \left[1 + \sum P_{ij} H_1(x_i) H_1(x_j) + \right. \\ &+ \frac{1}{2!} \left\{ \sum P_{ij}^2 H_2(x_i) H_2(x_j) + 2 \sum P_{ij} P_{il} H_2(x_i) H_1(x_j) H_1(x_l) \right. \\ &\quad \left. + 2 \sum P_{ij} P_{lm} H_1(x_i) H_1(x_j) H_1(x_l) H_1(x_m) \right\} \\ &\left. + \frac{1}{3!} \left\{ \sum P_{ij}^3 H_3(x_i) + H_3(x_j) + \dots \right\} + \dots \right] \end{aligned} \tag{2}$$

if this series is convergent, which follows from §5.2 and §1.452

The terms in the series which do not involve the Hermite polynomials in x_i (or, we may say, the terms in $H_0(x_i)$), are

$$\varphi(x_1) \left\{ 1 + \sum' p_{ij} D_i D_j + \frac{1}{2!} (\sum' p_{ij} D_i D_j)^2 + \dots \right\} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n) \quad (3)$$

where \sum' denotes a summation from which $i = 1$ and $j = 1$ are excluded.

The terms in $H_1(x_1)$ are

$$D_1 \varphi(x_1) \sum_{i=2}^n p_{i1} D_i \left\{ 1 + \sum' p_{ij} D_i D_j + \frac{1}{2!} (\sum' p_{ij} D_i D_j)^2 + \dots \right\} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n) \quad (4)$$

Similarly the terms in $H_2(x_1)$ are

$$\frac{1}{2!} D_1^2 \varphi(x_1) \left(\sum_{i=2}^n p_{i1} D_i \right)^2 \left\{ 1 + \sum' p_{ij} D_i D_j + \frac{1}{2!} (\sum' p_{ij} D_i D_j)^2 + \dots \right\} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n)$$

and those in $H_m(x_1)$:

$$\frac{1}{m!} D_1^m \varphi(x_1) \left(\sum_{i=2}^n p_{i1} D_i \right)^m \left\{ 1 + \sum' p_{ij} D_i D_j + \frac{1}{2!} (\sum' p_{ij} D_i D_j)^2 + \dots \right\} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n)$$

From (3) and (4) we see that the regression equation of x_1 on x_2, x_3, \dots, x_n , is

$$x_1 = \frac{\sum_{i=2}^n p_{i1} D_i \left\{ e^{\sum' p_{ij} D_i D_j} \right\} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n)}{e^{\sum' p_{ij} D_i D_j} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n)}$$

We can show by induction that this is equivalent to the usual equation

$$x_1 = x_2 p_{12.34\dots n} + x_3 p_{13.24\dots n} + \dots + x_n p_{1n.23\dots(n-1)}$$

and show that for any given values of x_2, x_3, \dots, x_n , the

distribution of x is normal, with standard deviation

$$(1 - \rho_{12}^2)(1 - \rho_{13.2}^2)(1 - \rho_{14.23}^2) \dots (1 - \rho_{1n.23\dots(n-1)}^2).$$

The 'Type III' distribution

19.21 We shall first consider the distribution of the variances of n variates which have a normal distribution.

Let y1, y2, ..., yn have the distribution function

phi(y1, y2, ..., yn) = 1/(2n)^n * 1/sqrt(R) * exp{- (R11*y1^2 + R22*y2^2 + ... + Rnn*yn^2 + 2R12*y1*y2 + ...) / 2R}

and put x1 = 1/2*y1^2, x2 = 1/2*y2^2, ..., xn = 1/2*yn^2.

Let

psi(x1, x2, ..., xn) dx1 dx2 ... dxn = sum phi(y1, y2, ..., yn) dy1 dy2 ... dyn

where the sign sum indicates a summation over all the 2^n sets of values of +/- y1, +/- y2, ..., +/- yn.

The moment generating function for the x distribution is

is integral integral ... integral psi(x1, x2, ..., xn) e^{alpha1*x1 + alpha2*x2 + ... + alphan*xn} dx1 dx2 ... dxn = integral integral ... integral phi(y1, y2, ..., yn) e^{1/2(alpha1*y1^2 + alpha2*y2^2 + ... + alphan*yn^2)} dy1 dy2 ... dyn = 1/(2n)^n * 1/sqrt(R) * integral integral ... integral exp{- (R11 - alpha1R)y1^2 + (R22 - alpha2R)y2^2 + ... + 2R12*y1*y2 + ...} dy1 ... dyn

where R denotes [matrix of cofactors] and Rij the co-factor of Pij in R.

and this is R^{1/2n} (R, R)^{-1/2}

where Ri denotes

$$\begin{pmatrix}
 R_{11} - \alpha_1 R & R_{12} & R_{13} & \dots & R_{1n} \\
 R_{12} & R_{22} - \alpha_2 R & R_{23} & \dots & R_{2n} \\
 R_{13} & R_{23} & R_{33} - \alpha_3 R & \dots & R_{3n} \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 R_{1n} & R_{2n} & R_{3n} & \dots & R_{nn} - \alpha_n R
 \end{pmatrix} \quad (3)$$

In R_1 the coefficient of $\alpha_1, \alpha_2, \dots, \alpha_n$ is

$$\begin{aligned}
 & (-R)^k \begin{vmatrix} R_{k+1,k+1} & R_{k+1,k+2} & \dots & R_{k+1,n} \\ R_{k+2,k+2} & R_{k+2,k+2} & \dots & R_{k+2,n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{k+1,n} & R_{k+2,n} & \dots & R_{n,n} \end{vmatrix} \\
 & = (-R)^k R^{n-k-1} \begin{vmatrix} 1 & p_{12} & p_{13} & \dots & p_{1k} \\ p_{12} & 1 & p_{23} & \dots & p_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{1k} & p_{2k} & p_{3k} & \dots & 1 \end{vmatrix} \quad (4)
 \end{aligned}$$

so that the required moment generating function $(R/R^{n-1})^{-\frac{1}{2}}$ can be expressed as

$$\left[1 - \sum \alpha_i + \sum \alpha_i \alpha_j \begin{vmatrix} 1 & p_{ij} \\ p_{ij} & 1 \end{vmatrix} - \sum \alpha_i \alpha_j \alpha_k \begin{vmatrix} 1 & p_{ij} & p_{ik} \\ p_{ij} & 1 & p_{jk} \\ p_{ij} & p_{jk} & 1 \end{vmatrix} + \dots \right]^{-\frac{1}{2}} \quad (5)$$

where the Σ sign denotes a summation over all sets of 1, 2, 3, ..., suffixes chosen from 1, 2, 3, ..., n.

The coefficient of $\alpha_1 \alpha_2 \dots \alpha_k$ is in R_1 is

$$\begin{aligned}
 & (-)^k R^k \left\{ 1 - \sum p_{ii}^2 + 2 \sum p_{12} p_{13} p_{23} - \sum (2 p_{12} p_{24} p_{43} p_{31} + 2 p_{13} p_{32} p_{24} p_{41} + \right. \\
 & \left. + 2 p_{14} p_{43} p_{32} p_{21} - p_{12}^2 p_{34}^2 - p_{13}^2 p_{24}^2 - p_{14}^2 p_{23}^2) + \dots \right\}
 \end{aligned}$$

We see thus that the terms independent of the ρ_i in [] in (5) are

$$1 - \sum \alpha_i + \sum \alpha_i \alpha_j - \sum \alpha_i \alpha_j \alpha_k + \dots + (-1)^n \alpha_1 \alpha_2 \dots \alpha_n$$

$$= (1 - \alpha_1)(1 - \alpha_2) \dots (1 - \alpha_n) \tag{6}$$

the coefficient of $\alpha_1 \alpha_2 \rho_{12}^2$ is

$$-1 + \sum \alpha_j - \sum \alpha_j \alpha_k + \dots + (-1)^{n-1} \alpha_3 \alpha_4 \dots \alpha_n = -(1 - \alpha_3)(1 - \alpha_4) \dots (1 - \alpha_n); \tag{7}$$

the coefficient of $\alpha_1 \alpha_2 \alpha_3 \cdot 2 \rho_{12} \rho_{23} \rho_{13}$ is

$$(1 - \alpha_4)(1 - \alpha_5) \dots (1 - \alpha_n)$$

and similarly for the terms of higher degrees in the ρ_i

So that the expression (5) can be written

$$\left[(1 - \alpha_1)(1 - \alpha_2) \dots (1 - \alpha_n) \left\{ 1 - \sum \frac{\alpha_i \alpha_j}{(1 - \alpha_i)(1 - \alpha_j)} \rho_{ij}^2 + \sum \frac{\alpha_i \alpha_j \alpha_k}{(1 - \alpha_i)(1 - \alpha_j)(1 - \alpha_k)} 2 \rho_{ij} \rho_{jk} \rho_{ik} + \dots \right\} \right]^{-\frac{1}{2}} \tag{8}$$

The terms in the bracket can be arranged :

$$\left[\begin{array}{cccc} 1 & \rho_{12} \sqrt{\frac{\alpha_1 \alpha_2}{(1 - \alpha_1)(1 - \alpha_2)}} & \rho_{13} \sqrt{\frac{\alpha_1 \alpha_3}{(1 - \alpha_1)(1 - \alpha_3)}} & \dots & \rho_{1n} \sqrt{\frac{\alpha_1 \alpha_n}{(1 - \alpha_1)(1 - \alpha_n)}} \\ \rho_{12} \sqrt{\frac{\alpha_1 \alpha_2}{(1 - \alpha_1)(1 - \alpha_2)}} & 1 & \rho_{23} \sqrt{\frac{\alpha_2 \alpha_3}{(1 - \alpha_2)(1 - \alpha_3)}} & \dots & \rho_{2n} \sqrt{\frac{\alpha_2 \alpha_n}{(1 - \alpha_2)(1 - \alpha_n)}} \\ \vdots & & & & \\ \rho_{1n} \sqrt{\frac{\alpha_1 \alpha_n}{(1 - \alpha_1)(1 - \alpha_n)}} & & & & 1 \end{array} \right]$$

so that (8) can be more compactly, though possibly less conveniently (be) written

$$\left[\begin{array}{c|cccc} \alpha_1 \alpha_2 \dots \alpha_n & \frac{1 - \alpha_1}{\alpha_1} & \rho_{12} & \rho_{13} & \dots & \rho_{1n} \\ & \rho_{12} & \frac{1 - \alpha_2}{\alpha_2} & \rho_{23} & \dots & \rho_{2n} \\ & \vdots & & & & \\ & \rho_{1n} & \rho_{2n} & \rho_{3n} & \dots & \frac{1 - \alpha_n}{\alpha_n} \end{array} \right]^{-\frac{1}{2}} \tag{9}$$

$$\begin{aligned} & \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) \left[1 + p \left\{ \sum \frac{L_1(x_1)L_1(x_2)p_{12}^2}{p^2} - \sum \frac{L_1(x_1)L_1(x_2)L_1(x_3)}{p^3} 2p_{12}p_{13}p_{23} + \dots \right\} \right. \\ & + \frac{p(p+1)}{2!} \left\{ \sum \frac{L_2(x_1)L_2(x_2)p_{12}^4}{p(p+1)p(p+1)} + 2 \sum \frac{L_1(x_1)L_1(x_2)L_1(x_3)L_1(x_4)}{p^4} (p_{12}^2 p_{34}^2 + p_{13}^2 p_{24}^2 + p_{14}^2 p_{23}^2) \right. \\ & \left. \left. + 2 \sum \frac{L_2(x_1)L_1(x_2)L_1(x_3)}{p(p+1)p} p_{12}^2 p_{13}^2 - \dots \right\} + \dots \right], \quad (11) \end{aligned}$$

where $\varphi(x)$ denotes $x^{b-1} e^{-x} / \Gamma(b)$ and $L_R(x) = L_R(x, p)$.

We can write this more simply by using a symbolic operator

D_x , defined by $p(p+1) \dots (p+k-1) D_x^k \varphi(x) = L_R(x, p) \varphi(x)$,

$$D_x^k \varphi(x) = \left(-\frac{\partial}{\partial x} \right)^k x^k \varphi(x),$$

or $D(D^k) = -\frac{\partial}{\partial x} D^k x$. D_x is to operate only on functions of x , and not on functions of any others of the x_1, x_2, \dots .

Using such operators we see, from (8), that (11) is equivalent

$$\text{to } \left\{ 1 - \sum p_{12}^2 D_1 D_2 + \sum 2p_{12}p_{13}p_{23} D_1 D_2 D_3 - \dots \right\}^{-p} \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) \quad (12)$$

or, from (9), to

$$(D_1 D_2 \dots D_n)^{-p} \begin{vmatrix} D_1^{-1} & p_{12} & p_{13} & \dots & p_{1n} \\ p_{12} & D_2^{-1} & p_{23} & \dots & p_{2n} \\ \vdots & & & & \vdots \\ p_{1n} & p_{2n} & p_{3n} & \dots & D_n^{-1} \end{vmatrix}^{-p} \varphi(x_1) \varphi(x_2) \dots \varphi(x_n), \quad (13)$$

in which, in expanding the power of the determinant, by the binomial theorem, the diagonal term is to be taken as the dominant term, so that terms in negative powers of any of the D will not occur in the series.

The convergence of this series is discussed below, in

From the last expression for the frequency function we can readily find the terms in $L_0(x_1), L_1(x_1), \dots$. Denote the determinant in that expression by Δ and put it $D_1^{-1} \Delta + R$, where R is independent of D_1^{-1} . Then the expression (13) becomes $(D_1 D_2 \dots D_n)^{-p} \{D_1^{-1} \Delta + R\}^{-p} \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) = (D_1 D_2 \dots D_n)^{-p} \{D_1^p \Delta^{-p} - p D_1^{p-1} \Delta^{-p-1} R + \dots\} \varphi(x_1) \varphi(x_2) \dots \varphi(x_n)$, from which it is evident that the term independent of D_1 is

$$\varphi(x_1) (D_1 D_2 \dots D_n)^{-p} \Delta^{-p} \varphi(x_2) \varphi(x_3) \dots \varphi(x_n),$$

and that in D_1 is

$$-p L_1(x_1, p) \varphi(x_1) (D_2 D_3 \dots D_n)^{-p} \Delta^{-p-1} R \varphi(x_2) \varphi(x_3) \dots \varphi(x_n),$$

and that in D_1^m is

$$\binom{-p}{m} L_m(x_1, p) \varphi(x_1) (D_2 D_3 \dots D_n)^{-p} \Delta^{-p-m} R^m \varphi(x_2) \varphi(x_3) \dots \varphi(x_n).$$

This can be put in a form closely resembling that in by noting that

$$R = - \sum_{i=2}^n \rho_{1i} \begin{vmatrix} D_2^{-1} & \rho_{23} & \rho_{24} & \dots & \rho_{2,i-1} & \rho_{12} & \rho_{2,i+1} & \dots & \rho_{2n} \\ \rho_{32} & D_3^{-1} & \rho_{34} & \dots & \rho_{3,i-1} & \rho_{13} & \rho_{3,i+1} & \dots & \rho_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{2i} & \rho_{3i} & \rho_{4i} & \dots & \rho_{i-1,i} & \rho_{1i} & \rho_{i,i+1} & \dots & \rho_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{2n} & \rho_{3n} & \rho_{4n} & \dots & \rho_{i-1,n} & \rho_{1n} & \rho_{i+1,n} & \dots & D_n^{-1} \end{vmatrix}$$

in which the principal term is

$$- \sum_{i=2}^n \rho_{1i}^2 D_2^{-1} D_3^{-1} \dots D_{i-1}^{-1} D_{i+1}^{-1} \dots D_n^{-1}$$

The Multiple Binomial Distribution

19.3 For the double binomial distribution discussed in Aitken and Gonin's paper, the factorial moment generating function is $(1 + p\alpha + p'\beta + p''\alpha\beta)^n$ (1)

where p is the probability of A
 p' of B
 p'' of A and B together.

For a multiple binomial distribution, obtained by selection from a 2^n -fold table, put

p_1 the probability of A_1
 p_2 of A_2
etc.,
 p_{12} of $A_1 A_2$
etc.,
 $p_{12\dots m}$ of $A_1 A_2 \dots A_m$

The factorial moment generating function is then

$$(1 + p_1\alpha_1 + p_2\alpha_2 + \dots + \sum p_{12}\alpha_1\alpha_2 + \sum p_{123}\alpha_1\alpha_2\alpha_3 + \dots + p_{12\dots m}\alpha_1\alpha_2\dots\alpha_m)^n$$
 (2)

Put $d_{12} = p_{12} - p_1 p_2$, etc.,
 $d_{123} = p_{123} - p_1 p_2 p_3$, etc.,
etc.,

so that each d indicates the excess of the corresponding p over what its value would be in the absence of any correlation. The expression (2) is then

$$\left\{ (1 + p_1\alpha_1)(1 + p_2\alpha_2) \dots (1 + p_m\alpha_m) + \sum d_{12}\alpha_1\alpha_2 + \sum d_{123}\alpha_1\alpha_2\alpha_3 + \dots \right\}^n$$
 (3)

Now put $d'_{123} = d_{123} - d_{12}p_3 - d_{13}p_2 - d_{23}p_1$, etc.,
 $d'_{1234} = d_{1234} - d_{12}p_3p_4 - d_{13}p_2p_4 - d_{14}p_2p_3 - d_{23}p_1p_4 - d_{24}p_1p_3 - d_{34}p_1p_2$,
etc.,

so that each d' indicates the excess of the corresponding p over what its value would be in the absence of any special correlation for that particular set of variates,

over and above the correlations between the various ~~subsets~~ pairs of variates involved in that set. Then (3) is

$$\{(1+p_1\alpha_1)(1+p_2\alpha_2)\dots(1+p_m\alpha_m) + \sum d_{12}\alpha_1\alpha_2(1+p_3\alpha_3)(1+p_4\alpha_4)\dots(1+p_m\alpha_m) + \sum d'_{123}\alpha_1\alpha_2\alpha_3 + \sum d'_{1234}\alpha_1\alpha_2\alpha_3\alpha_4 + \dots\}^n$$

Now put $d''_{1234} = d'_{1234} - d'_{123}p_4 - d'_{124}p_3 - d'_{134}p_2 - d'_{234}p_1$, etc.,
 $d''_{12345} = d'_{12345} - d''_{1234}p_5 - \dots$, etc.,

so that $d_{12\dots r}^{(r-2)}$ will indicate the excess of $p_{12\dots r}$ over the value which might be expected by considering the correlations, not only between all pairs of variates in the set of n variates, but between all possible sub-sets of variates.

Then (4) becomes

$$\{(1+p_1\alpha_1)(1+p_2\alpha_2)\dots(1+p_m\alpha_m)\}^n \times \left\{ 1 + \sum d_{12} \frac{\alpha_1\alpha_2}{(1+p_1\alpha_1)(1+p_2\alpha_2)} + \sum d'_{123} \frac{\alpha_1\alpha_2\alpha_3}{(1+p_1\alpha_1)(1+p_2\alpha_2)(1+p_3\alpha_3)} + \dots \right\} \quad (4)$$

The terminating series here given has $(1+p_r\alpha_r)^n \left(\frac{\alpha_r}{1+p_r\alpha_r}\right)^r$ as a factor in each term, with some $r \leq n$, and it is, therefore, the factorial moment generating function of a function, which, divided by $\varphi(x_i) = \binom{n}{x_i} p_i^{x_i} (1-p_i)^{n-x_i}$ gives a function linear in ~~the~~ $\zeta_1(x_i), \zeta_2(x_i), \dots, \zeta_m(x_i)$; and, of course, linear also in $\varphi(x_i)\zeta_1(x_i), \varphi(x_i)\zeta_2(x_i), \dots, \varphi(x_i)\zeta_m(x_i)$, $i = 2, 3, \dots, m$.

This is, in important ways, more general than the functional form found in the previous section, For here, not only can we take a different value of p for each variate, but, further, the nature of the correlation is not

Convergence of the multivariate Laguerre series

19.4 The series to be discussed is

$$\left\{ 1 - \sum p_{11}^2 D_1 D_2 + \sum 2 p_{11} p_{13} p_{15} D_1 D_2 D_3 - \dots (-1)^{r-1} |R-I| D_1 D_2 \dots D_r \right\}^{-p} \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) \tag{1}$$

where $\varphi(x_s) = x_s^{p-1} e^{-x_s} / \Gamma(p)$,

and D_s is an operator on functions of x_s such that

$D_s^r \varphi(x_s) = L_r(x_s, p) \varphi(x_s)$ where L_r is the Laguerre polynomial, and $p^{(r)} = p(p+1)\dots(p+r-1)$

The coefficient of $(-1)^{r-1} D_1 D_2 \dots D_r$ is

$$\begin{vmatrix} 0 & p_{12} & p_{13} & \dots & p_{1r} \\ p_{12} & 0 & p_{23} & \dots & p_{2r} \\ \dots & \dots & \dots & \dots & \dots \\ p_{1r} & p_{2r} & p_{3r} & \dots & 0 \end{vmatrix} = |R_r - I|$$

We shall consider a more general series in which the parameters p in the functions φ are not all the same :

$$\varphi(x_s) = x_s^{p_s-1} e^{-x_s} / \Gamma(p_s), \quad p_s^{(r)} D_s^r \varphi(x_s) = L_r(x_s, p_s) \varphi(x_s) \tag{2}$$

The two-variable series has been seen to be convergent (§5.2) using a theorem of Hille's which will be used here also. (The argument of §5.2 will clearly apply even if we take different p 's in the two functions $\varphi(x)$ and $\varphi(y)$).

The series (1) can be written

$\sum_{k=0}^{\infty} T_k$ if T_k denotes

$$\frac{p(p+1)\dots(p+k-1)}{k!} \left\{ \sum p_{11}^2 D_1 D_2 - \sum 2 p_{11} p_{13} p_{15} + \dots (-1)^{r-1} |R_r - I| \right\}^k \varphi(x_1) \varphi(x_2) \dots \varphi(x_n)$$

Now put

$$L_r(x, p) = \frac{r! \Gamma(p)}{\Gamma(p+r)} \left\{ \frac{1}{\sqrt{\pi}} e^{ix} x^{\frac{1}{2}-\frac{r}{2}} r^{\frac{r}{2}-\frac{3}{2}} \right\}, \tag{3}$$

$$D_i^r \varphi(x_i) = L_r(x_i, p_i) \varphi(x_i),$$

$$J_k = \frac{p(p+1)\dots(p+k-1)}{k!} \left\{ \sum p_{12}^2 D_1 D_2 - \sum 2p_{12} p_{13} p_{23} D_1 D_2 D_3 + \dots (-)^r |R-I| \right\}^e \cdot \varphi(x_1) \varphi(x_2) \dots \varphi(x_n)$$

Then

$$J_k = \frac{p+k-1}{k} \left\{ \sum p_{12}^2 D_1 D_2 - \sum 2p_{12} p_{13} p_{23} D_1 D_2 D_3 + \dots \right\} J_{k-1}$$

We shall show that as k tends to infinity,

$$\frac{J_k}{J_{k-1}} \rightarrow \sum p_{12}^2 - \sum 2p_{12} p_{13} p_{23} + \dots (-)^r |R-I| = 1 - |R|$$

and hence that $\sum J_k$ is absolutely convergent ~~if~~ if R is a positive definite matrix, as it always will be when the series gives a frequency function for variates between which p_{12}, p_{13}, \dots indicate the correlations.

It is evident from the definition (3) that

$$\frac{L_{r-1}}{L_r} \rightarrow 1 \quad \text{and hence} \quad \left| \frac{L_{r-1}}{L_r} - 1 \right| < \epsilon \quad \text{all } r \geq r_0$$

If there are different parameters p, there may be n different values of r_i . We take the largest of them.

$$\text{Let } J_k = J_{k_0} + J_{k_1} + J_{k_2} + \dots + J_{k_n}$$

where J_{k_0} consists of all those terms in the expansion of J_k in which each of the D_i is raised to a power $\geq r_i$,
 J_{k_1} of those terms in which D_1 is raised to a power $< r_1$,
 J_{k_2} of those terms in which D_2 but not D_1 is raised to a power $< r_2$,

and so on.

Then J_{k_0} consists of the first r_0 terms of

$$\frac{p(p+1)\dots(p+k-1)}{k!}$$

$$\left[\left\{ \sum p_{23}^2 D_2 D_3 - \sum 2p_{23} p_{24} p_{34} D_2 D_3 D_4 + \dots \right\} + D_1 \left\{ \sum p_{12}^2 D_2 - \sum 2p_{12} p_{13} p_{23} D_2 D_3 + \dots \right\} \right]^k \cdot \varphi(x_1) \varphi(x_2) \dots \varphi(x_n)$$

where \sum now denotes a summation over all permutations of the suffixes 2, 3, 4, ..., n.

The coefficient of $\frac{p(p+1)\dots(p+k-1)}{k!} D_1^s \varphi(x_1)$

in this expansion is

$$C_s = \binom{k}{s} \left\{ \sum p_{12}^2 D_2 - \sum 2p_{12} p_{13} p_{23} D_2 D_3 + \dots \right\}^s \left\{ \sum p_{23}^2 D_2 D_3 - \sum 2p_{23} p_{34} p_{24} D_2 D_3 D_4 + \dots \right\}^{k-s} \cdot \varphi(x_1) \varphi(x_2) \dots \varphi(x_n)$$

If we assume for $n - 1$ the proposition we have to prove for n , we have

$$\frac{s}{k-s+1} \cdot \frac{C_s}{C_{s-1}} \rightarrow \frac{1-|R|}{1-|R'|} \quad \text{as } k \rightarrow \infty$$

where $|R'|$ denotes

$$\begin{vmatrix} 0 & p_{23} & p_{24} & \dots & p_{2n} \\ p_{13} & 0 & p_{34} & \dots & p_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{1n} & p_{2n} & p_{3n} & \dots & 0 \end{vmatrix}$$

Now $J_k = \sum_0^{k-1} C_s$, $J_k = \sum_0^k C_s$ and since $\frac{C_s}{C_{s-1}}$ has just been shown to tend to a finite limit as $k \rightarrow \infty$, it follows

that $\frac{J_{k+1}}{J_k} \rightarrow 0$. Similarly $\frac{J_{k+2}}{J_k} \rightarrow 0$, and so on.

Therefore $\frac{J_{k+1}}{J_k} \rightarrow 1$ But $\frac{J_{k+1}}{J_{k-1}} \rightarrow 1-|R|$

$\therefore \frac{J_k}{J_{k-1}} \rightarrow 0$, which was required to be proved. (4)

And, more generally,

$$\frac{\left\{ \sum a_1 D_1 + \sum a_{12} D_1 D_2 + \dots + a_{12\dots n} D_1 D_2 \dots D_n \right\} J_k}{J_k}$$

$$\rightarrow \sum a_1 + \sum a_{12} + \dots + a_{12\dots n}, \quad \text{as } k \rightarrow \infty$$

From Hille's result, given any interval $0 \leq x \leq$

$$r^{\frac{p-3}{4}} \cos \left[2\sqrt{rx} - \pi \left(\frac{p}{2} - \frac{1}{4} \right) \right] / L_r(x, p) \rightarrow 1, \quad \text{as } r \rightarrow \infty$$

except for those values of r for which

$$\left| \cos \left[2\sqrt{rx} - \pi \left(\frac{p}{2} - \frac{1}{4} \right) \right] \right| < \frac{K}{r^{\frac{3}{4}}}$$

i.e., those for which

$$\left| 2\sqrt{rx} - \pi \left(\frac{p}{2} - \frac{1}{4} \right) - (2m+1) \frac{\pi}{2} \right| < \frac{K}{r^{\frac{3}{4}}} \quad (6)$$

where m is any integer. K is an arbitrary positive constant. Given m , let r_1 and r_2 be the extreme values for which the inequality is satisfied. That is,

$$2r_1^{\frac{3}{4}}\sqrt{x} - \pi r_1^{\frac{1}{4}} \left(m + \frac{p}{2} + \frac{1}{4} \right) = -K,$$

$$2r_2^{\frac{3}{4}}\sqrt{x} - \pi r_2^{\frac{1}{4}} \left(m + \frac{p}{2} + \frac{1}{4} \right) = +K,$$

$$2\sqrt{x}(r_2 - r_1) = K \left(\frac{1}{r_1^{\frac{1}{4}}} + \frac{1}{r_2^{\frac{1}{4}}} \right) (r_2^{\frac{1}{2}} + r_1^{\frac{1}{2}}) \sim 4Kr_1^{\frac{1}{4}}$$

Thus, corresponding to each m the number of terms for which the inequality (6) holds is $O(r^{\frac{1}{2}}) = O(m^{\frac{1}{2}})$.

That is, given ϵ we can find r_0 , such that

$$\left| \frac{L_r(x, b)}{r^{\frac{p}{2}-\frac{1}{4}} \cos\left[2\sqrt{rx} - \pi\left(\frac{p}{2}-\frac{1}{4}\right)\right]} - 1 \right| < \epsilon, \quad 0 < b \leq x \leq a$$

for all $r \geq r_0$, except for certain values of $r = O(m^2)$, corresponding to each positive integer m , the number of them for any one value of m being $O(m^{\frac{1}{2}})$. For those critical values of r , $L_r(x, b) = O(r^{\frac{p}{2}-1})$

Now,

$$\begin{aligned} \frac{L_{r-1}(x, b)}{L_r(x, b)} &= \frac{p+r-1}{r} \left(\frac{r-1}{r}\right)^{\frac{p}{2}-\frac{3}{4}} \frac{\cos\left[2\sqrt{(r-1)x} - \pi\left(\frac{p}{2}-\frac{1}{4}\right)\right]}{\cos\left[2\sqrt{rx} - \pi\left(\frac{p}{2}-\frac{1}{4}\right)\right]} + O(r^{-\frac{1}{4}}) \\ &= \frac{L_{r-1}}{L_r} \left(1 + \sqrt{\frac{x}{r}} \tan\left[2\sqrt{rx} - \pi\left(\frac{p}{2}-\frac{1}{4}\right)\right]\right) + O(r^{-\frac{1}{4}}) \end{aligned}$$

except for those values of r for which

$$\sqrt{\frac{x}{r}} \tan\left[2\sqrt{rx} - \pi\left(\frac{p}{2}-\frac{1}{4}\right)\right] > K_1 r^{-\frac{1}{4}}$$

or

$$\left| 2\sqrt{rx} - \pi\left(\frac{p}{2}-\frac{1}{4}\right) - (2m+1)\frac{\pi}{2} \right| < \frac{\sqrt{x}}{K_1 r^{\frac{1}{4}}}$$

Clearly the values of r which satisfy this inequality form a set of the same order as those which satisfy the inequality (6), so that

$$\left| \frac{DL_r \varphi}{L_r \varphi} - 1 \right| < \epsilon \quad \text{for all } r \geq r_0 \quad \text{except for a number of}$$

values of r of the order of $m^{\frac{1}{2}}$ in the neighbourhood

of $r = \frac{\pi^2 \left(n + \frac{p}{2} + \frac{1}{4}\right)^2}{4x}$ for every positive integral value of m

We shall now prove that

$$\frac{\left\{ \sum a_n D_n + \sum a_{n_2} D_1 D_2 + \dots + a_{n_2 \dots n} D_1 D_2 \dots D_n \right\} T_k}{T_k}$$

$$\rightarrow \sum a_n + \sum a_{n_2} + \dots + a_{n_2 \dots n}, \text{ as } k \rightarrow \infty,$$

of which (5) was an indication.

It has been shown that given any positive ϵ we can find r_0 such that $D_r(L_r(x)\phi(x))$ lies between $(1 \pm \epsilon)L_r(x)\phi(x)$ for all $r > r_0$, except for the critical values of r found above. The terms in the expansion of T_k for which $r < r_0$ can be dealt with as in § and the terms for which

$$r \equiv \frac{\pi^2 \left(n + \frac{k}{2} + \frac{1}{4} \right)^2}{4x}$$

are of lower order than

neighbouring terms, and the ratio of their number to the total number of terms in T_k tends to zero as k tends to infinity, so that if T_k' denotes T_k with all such terms omitted, $T_k'/T_k \rightarrow 1$ as $k \rightarrow \infty$ and

$$\frac{(\sum a_n D_n + \sum a_{n_2} D_1 D_2 + \dots + a_{n_2 \dots n} D_1 D_2 \dots D_n) T_k'}{(\sum a_n D_n + \sum a_{n_2} D_1 D_2 + \dots + a_{n_2 \dots n} D_1 D_2 \dots D_n) T_k} \rightarrow 1.$$

Hence the proposition is proved, and it follows that

$$\frac{T_k}{T_{k-1}} \rightarrow 1 - |R| \text{ as } k \rightarrow \infty$$

and that $\sum |T_k|$ is convergent, and $\sum T_k$ is convergent, which was required to be proved. It does not follow that the series obtained by taking separately each term in T_k is absolutely convergent.

Chapter VII

DIFFERENTIAL EQUATIONS AND DIFFERENCE EQUATIONS

§20.1 In seeking for the form of a two-variate distribution analogous to any of Pearson's types in one variate, it is natural to enquire what will correspond to the differential equation which distributions of those types all satisfy

$$kx(x+b) \frac{\partial \phi}{\partial x} + (x-a)\phi = 0 \tag{1}$$

which reduces to the form

$$kx \frac{\partial \phi}{\partial x} + (x-a)\phi = 0 \tag{2}$$

in the special case of type III, and to

$$k \frac{\partial \phi}{\partial x} + (x-a)\phi = 0 \tag{3}$$

in the most special case of the ~~linear~~ normal distribution,

The use of curves of Pearson's types has generally been justified on the ground that equation (1) is more or less the simplest differential equation to express the facts that the curve in general touches the axis at the extremities of the range, and that it has not more than one maximum.

The use of them can however be justified on grounds which are perhaps more general, which at least give good reason for having a quadratic and not a cubic function of x in the equation. For the differential equation can be taken as the analogue of the difference equation satisfied by the hypergeometric distribution, and the special importance of type III lies, from this point of view in the fact that its differential equation is analogous to the difference equation of the binomial distribution, of which the frequency generating function is $(p + qt)^n$.

The difference equation is

$$\frac{q}{p}(x+1) \Delta \varphi(x) = \left\{ n - x - \frac{q}{p}(x+1) \right\} \varphi(x) \tag{4}$$

for $\frac{q}{p}(x+1) \varphi(x+1) = (n-x) \varphi(x)$

and that for the hypergeometric distribution is

~~$$\{x^2 + (Nq - n)x\} \Delta \varphi(x) = N(np - x) \varphi(x); \tag{5}$$~~

$$\{(x+1)^2 + (Nq - n)(x+1)\} \Delta \varphi(x) = \{- (N+2)x + N(pn - q) + n - 1\} \varphi(x);$$

for, in it,

~~$$\frac{(x+1)(x+1+Nq-n)}{x(x+Nq-n)} \varphi(x+1) - \{x^2 - (Np+n)x + Npn\} \varphi(x) = 0;$$~~

and if in this we replace the difference by a derivative, it clearly becomes equivalent to (1).

This is not a limiting process. Either of the equations (4) and (5) tends to the form (3) as n tends to infinity. That is to say, if the number in a sample from a binomial or hypergeometric distribution tends to infinity, the continuous distribution to which the distribution tends is always a normal distribution.

01

The procedure of replacing a difference by a derivative is not that of finding the continuous distribution to which a discontinuous one tends, but of finding a continuous function to represent a distribution actually discontinuous.

This seems to be a reasonable procedure for finding a curve to represent many naturally-occurring distributions. It is well known that if observed deviations are due to a very large number of very small nearly independent elementary deviations, the resulting distribution will be nearly normal. In any actual instance, however, though there may be an infinite number of possible sources of deviations of infinitesimal amount, there are generally a measurable number of sources of deviation of appreciable amount. If the deviations from such sources could take only two possible values, the same for each of them, and had the same probabilities for each, then the observed distribution of deviations would be a binomial one. Actually, of course, there is very rarely any such regularity, and the observed distribution is a continuous one. But it is reasonable to suppose in these circumstances that it will resemble the binomial distribution, or, if the deviations are in some way extracted from a not-inexhaustible store, the hypergeometric. This resemblance is expressed by the analogy between the difference equation (5) and the differential equation (1).

20.2 We approach the two-variate problem, then, by considering the ~~two~~ two-variate analogue of the binomial and hypergeometric distributions, which has been discussed by

Aitken and Gonin. For the double-binomial distribution discussed in their paper there exists the recurrence relation :

$$(x+1) \frac{p_{01}}{p_{11}} \varphi(x+1, y) + \left\{ x - y + (x - n + y) \frac{p_{10} p_{01}}{p_{00} p_{11}} \right\} \varphi(x, y) - (n - x + 1) \frac{p_{10}}{p_{00}} \varphi(x-1, y) = 0$$

which can be written

$$\frac{1}{2} \left\{ \left(\frac{p_{01}}{p_{11}} + \frac{p_{10}}{p_{00}} \right) (x+1) - (n+2) \frac{p_{10}}{p_{00}} \right\} \delta^2 \varphi + \left\{ \left(\frac{p_{01}}{p_{11}} - \frac{p_{10}}{p_{00}} \right) (x+1) + (n+2) \frac{p_{10}}{p_{00}} \right\} \delta \varphi + \left\{ \left(1 + \frac{p_{10}}{p_{00}} \right) \left(1 + \frac{p_{01}}{p_{11}} \right) (x+1) - \frac{p_{10}}{p_{00}} \left(1 + \frac{p_{01}}{p_{11}} \right) (n+2) - \left(1 - \frac{p_{10} p_{01}}{p_{00} p_{11}} \right) (y+1) \right\} \varphi = 0 \quad (6)$$

where δ and δ^2 denote central differences :

$$\begin{aligned} \delta \varphi &= \frac{1}{2} \{ \varphi(x+1) - \varphi(x-1) \}, & \delta^2 \varphi &= \frac{1}{2} \{ \varphi(x+1) - 2\varphi(x) + \varphi(x-1) \} \\ &= \frac{1}{2} \{ \Delta \varphi + \nabla \varphi \}, & &= \Delta \varphi - \nabla \varphi \end{aligned}$$

The two-variate distribution found in p. 28 (in which the distribution is either variate alone is a Type III distribution), is

$$\varphi(x, y) = e^{-(x+y)/(1-p^2)} \frac{(xy)^{\frac{1}{2}(b-1)} p^{-b+1}}{(1-p^2) \Gamma(b)} I_{p-1} \left\{ \frac{2p\sqrt{xy}}{1-p^2} \right\}$$

and this satisfies the (partial) differential equation

$$x \frac{\partial^2 \varphi}{\partial x^2} + \left(\frac{2x}{1-p^2} - b \right) \frac{\partial \varphi}{\partial x} + \left(\frac{x-p^2 y}{1-p^2} - b \right) \frac{\varphi}{1-p^2} = 0 \quad (7)$$

When $p = 0$, this must of course reduce to ~~the~~ some form of the differential equation for Type III. We see that it reduces to

$$x \frac{\partial^2 \varphi}{\partial x^2} + (2x - b) \frac{\partial \varphi}{\partial x} + (x - b) \varphi = 0 \quad (8)$$

which can be derived from (2) by differentiating and adding.

In a similar way, if there is no correlation in the double-binomial distribution, the difference equation (6) must reduce to a difference equation for a simple binomial distribution. We see that when $p_{o_1} p_{i_0} = p_{o_0} p_{i_1}$, then (6) reduces, putting $p_{o_1}/p_{i_1} = q/p$, to

$$\frac{1}{2} \{ (p^2 + q^2)(x+1) - p^2(n+2) \} \delta^2 \phi + \{ (q^2 - p^2)(x+1) + p^2(n+2) \} \delta \phi + \{ x+1 - p(n+2) \} \phi = 0 \quad (9)$$

which can be derived from (4) by writing down the backward-difference form of (4), expressing each in terms of central differences by $\Delta \phi = \delta \phi + \frac{1}{2} \delta^2 \phi$, $\nabla \phi = \delta \phi - \frac{1}{2} \delta^2 \phi$ multiplying the two equations by q and p respectively, and adding. The second-order differential equation (8) and the second-order difference equation (9) have been derived from first-order equations which have a definite analogy. It will be seen that in equations (8) and (9) themselves, while the terms correspond, there is not an exact analogy between the coefficients. There is a similar extent of analogy between (6) and (7).

The analogy is not exact, but it is close enough for us to say that a distribution satisfying equation (7) and the same equation with x and y interchanged, will be a suitable type of distribution to represent a continuous distribution similar to the double-binomial distribution, but it does not show that some rather different differential equation would not have an equal suitability, a differential equation, for example, with a different numerical coefficient for the first term, which might depend on P .

Chapter VIII

SELECTION

§ 21.1 This chapter is concerned with a question of great practical importance. When it is desired to compare the distribution of one variate (or set of variates) in two different populations in which another correlated variate (or set of variates) has different distributions, it is important to find what would be the distribution of the first variate in ^a ~~xxxxxxx~~ populations formed by selecting out of one of the given populations a population in which the distribution of the second variate is the same as that in the other of the given populations. The simplest case is that in which the population selected has zero variance in the selected variate - that is, in which we consider the distribution of one variate in those individuals in the population for which the other variate has a particular value. The mean is then given by the curve of regression. It is easy to see that if the regression is linear, then whatever the variance selected, the mean of the other variate in the selected population will be a function of the selected mean given by the equation of regression. Interest in this chapter will therefore centre on the expressions found for the

variance and bi-variance or product-moment.

The problems to which the theory of selection is likely to be most applied are biometrical problems, in which the variates are essentially positive. For such variates, especially if the coefficient of variation is not very small, the Gamma-type curve will be more suited than the Gaussian, and it may therefore be of considerable importance to find some properties of a distribution selected from the distributions discussed in chapters

2). 2 Consider first the distribution which we found could be represented by a Bessel function. Its moment generating function is $(1 - \alpha - \beta + \alpha \beta \sqrt{1 - e^x})^{-p}$ (1)

which is equal to

$$\frac{1}{\{1 - \beta(1 - e^x)\}^p} \int_0^\infty \frac{x^{p-1} e^{-x}}{\Gamma(p)} \exp\left\{x \left(\frac{\beta e^x}{1 + \beta(1 - e^x)} + \alpha\right)\right\} dx \quad (2)$$

We are assuming that the distribution function for ~~each~~ each variate is

$$\frac{x^{p-1} e^{-x}}{\Gamma(p)} dx \quad \text{and not} \quad \int_0^x \frac{x^{p-1} e^{-kx}}{\Gamma(p)} dx$$

since by suitable choice of the unit the second form can be changed into the first. It means a choice of unit which makes the arithmetic mean equal to the variance. We can readily obtain the general forms of the results obtained, at the end of the discussion.

Now replacing $\frac{x^{p-1} e^{-x}}{\Gamma(p)} dx$ in (2) by $\frac{x^{k-1} e^{-x/m}}{m^k \Gamma(k)} dx$

it becomes

$$\begin{aligned}
& \frac{1}{m^k \{1 - \beta(1 - \rho^2)\}^k} \int_0^{\infty} \frac{x^{k-1}}{\Gamma(k)} \exp\left\{x\left(\frac{1}{m} - \frac{\beta\rho^2}{1 - \beta(1 - \rho^2)} - \alpha\right)\right\} dx \\
&= m^{-k} \{1 - \beta(1 - \rho^2)\}^{-k} \left\{\frac{1}{m} - \frac{\beta\rho^2}{1 - \beta(1 - \rho^2)} - \alpha\right\}^{-k} \\
&= \{1 - \beta(1 - \rho^2)\}^{-k} \left\{1 - m\alpha - (1 - \rho^2 + m\rho^2)\beta + m\alpha\beta(1 - \rho^2)\right\}^{-k} \\
&= 1 + km\alpha + \{p(1 - \rho^2) + k m \rho^2\}\beta + \frac{k(k+1)}{2!} m^2 \alpha^2 + \\
&\quad + \{km(p - k - 1)(1 - \rho^2) + k(k+1)m(1 - \rho^2 + m\rho^2)\} \alpha\beta \\
&\quad + \left\{\frac{k(k+1)}{2}(1 - \rho^2)^2 + k \cancel{m} \rho^2(1 - \rho^2) + \frac{k(k+1)}{2} m^2 \rho^4\right\} \beta^2 \\
&\quad + \text{terms in higher powers of } \alpha \text{ and } \beta. \quad (3)
\end{aligned}$$

Thus when from a population in which the mean x is p , and the mean y is p , and the variance of x is also p , and the variance of y is p , a population is selected in which the mean x is km and the variance of x is km^2 , then in the selected population the mean y is $p(1 - \rho^2) + km\rho^2$, the variance of y is $p(1 - \rho^2)^2 + 2km\rho^2(1 - \rho^2) + km^2\rho^4$, and the bivariate $km^2\rho^2$. (The bivariate in the parent population was $p\rho^2$.) Now put $h = km$, the mean x in the selected population, and $v = km^2$, the variance of x in the selected population, and put $h_0 =$ the mean x in the parent population, and $v_0 =$ the variance of x in that population, so that p is replaced by v_0 and the ~~expressions~~ expressions made homogeneous by putting in the factor v_0/h_0 where necessary. The expression for the mean y in the selected population becomes $h_0 + (h - h_0)\rho^2$, and that for the variance

$$v_0(1-r^2) + 2 \frac{r}{h_0} v_0 r(1-r^2) + v r$$

Sofar we have assumed that the mean and variance of y in the parent population are respectively equal to those of x . We can put them equal to h , and v , respectively, where $h^2/v = h_0^2/v_0$, since the type of distribution in y is supposed the same as that in x . With these changes the expression for the new mean y becomes

$$h_1 + (h - h_0) r^2 \sqrt{v_1/v_0} \quad (4)$$

which is the value given by the regression line, since $r^2 \sqrt{v_1/v_0}$ is the regression coefficient. We have already seen that it must be given by the regression line.

The expression for the new variance of y becomes

$$v_1(1-r^2)^2 + 2 \frac{h}{h_0} v_1 r^2(1-r^2) + \frac{v_1 v}{v_0} r^4 \quad (5)$$

and for the bivariate $v r^2 \sqrt{\frac{v_1}{v_0}}$ (6)

so that the coefficient of correlation which was r^2 becomes

$$\sqrt{\frac{v_1}{v_0}} r^2 \left\{ (1-r^2)^2 + 2 \frac{h}{h_0} r^2(1-r^2) + \frac{v}{v_0} r^4 \right\}^{-\frac{1}{2}} \quad (7)$$

§ 21.3

We can get a more general result by using

$$\frac{(x-c)^{k-1} e^{-(x-c)/m}}{m^k \Gamma(k)} \quad \text{instead of} \quad \frac{x^{k-1} e^{-x/m}}{m^k \Gamma(k)}$$

as the distribution of x in the selected population.

The limits of integration are from c to $+\infty$. Strictly c should be positive if a selection of this sort is to be possible, but we can allow a negative c if k is large enough and $-c$ small enough for the frequency function to be small for negative values of x . The effect is to

multiply the moment generating function (3) by

$$\begin{aligned} & \exp \left\{ c \left(\frac{\beta p^2}{1 - \beta(1-p^2)} + \alpha \right) \right\} \\ &= 1 + c \beta p^2 (1 - \beta(1-p^2))^{-1} + c \alpha + \frac{1}{2} c^2 \beta^2 p^4 + \frac{1}{2} c^2 \alpha^2 + c^2 \alpha \beta p^2 + \dots \\ &= 1 + c \alpha + c \beta^2 p^2 + \frac{1}{2} c^2 \alpha^2 + c^2 \alpha \beta p^2 + \left\{ c \beta^2 (1-p^2) + \frac{1}{2} c^2 p^4 \right\} \beta^2 + \dots, \end{aligned}$$

so that the moment generating function becomes

$$\begin{aligned} & 1 + (c + km) \alpha + \left\{ p(1-p^2) + (c + km) p^2 \right\} \beta + \frac{1}{2} \left\{ (c + km)^2 + km^2 \right\} \alpha^2 \\ & + \left\{ (c + km)^2 p^2 + (c + km) p(1-p^2) + km^2 p^2 \right\} \alpha \beta \\ & + \left\{ \frac{1}{2} p(p+1)(1-p^2)^2 + (c + km)(p+1)p^2(1-p^2) + \frac{1}{2} [(km+c)^2 + km^2] p^4 \right\} \beta^2 + \dots \end{aligned} \quad (8)$$

which is the same as (3) with $c + km$ for k_m ,

km^2 being unaltered. But $c + km$ is now the mean x in the selected population, so that the expressions for the mean y and the variance of y , and the bivariate take exactly the same forms as (5) and (6).

§21.4

The problem of selecting a Gaussian population from a Type-III population, or vice versa can be treated as a limiting case of the problem treated above,

for $\frac{(x-c)^{k-1}}{n^k \Gamma(k)} e^{-x/n}$ tends to the Gaussian form as

c tends to infinity, while $c + km$ and km^2 remain finite.

But a separate treatment can be given: if in (2),

$$\frac{x^{p-1}}{\Gamma(p)} e^{-x} \quad \text{is replaced by} \quad \frac{1}{\sqrt{v\pi}} e^{-(x-k)^2/v},$$

and the integration carried out from $-\infty$ to $+\infty$, it becomes

$$\frac{1}{\sqrt{v\pi} \{1 - \beta(1-p^2)\}^p} \int_{-\infty}^{\infty} \exp \left\{ -\frac{(x-k)^2}{2v} + x \left(\frac{\beta p^2}{1 - \beta(1-p^2)} + \alpha \right) \right\} dx$$

$$\begin{aligned}
&= \frac{1}{\{1 - \beta(1 - \rho^2)\}^k} \exp\left\{\left(\frac{\beta \rho^2}{1 - \beta(1 - \rho^2)} + \alpha\right)^2 v + \left(\frac{\beta \rho^2}{1 - \beta(1 - \rho^2)} + \alpha\right) h\right\} \\
&= 1 + \alpha h + \beta\{h\rho^2 + p(1 - \rho^2)\} + \frac{1}{2}(h^2 + v)\alpha^2 \\
&\quad + \{(h^2 + v)\rho^2 + hp(1 - \rho^2)\}\alpha\beta \\
&\quad + \left\{\frac{1}{2}p(p+1)(1 - \rho^2)^2 + h(p+1)\rho^2(1 - \rho^2) + \frac{1}{2}(h^2 + v)\rho^4\right\} \\
&\quad + \dots,
\end{aligned}$$

which agrees with (3) as far as the terms of the second order.

The converse problem is that of the selection of a Type III population from a Gaussian population.

In
$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha^2(1 - \rho^2)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} e^{(\beta + \rho\alpha)x} dx$$

replace $e^{-\frac{1}{2}x^2} dx / \sqrt{2\pi}$ by $\frac{1}{m^k \Gamma(k)} (x - c)^{k-1} e^{-(x-c)/m}$, taking the integral from c to $+\infty$. The required moment

generating function is then

$$\begin{aligned}
&\frac{1}{m^k \Gamma(k)} e^{\frac{1}{2}\alpha^2(1 - \rho^2)} \int_0^{\infty} x^{k-1} e^{-x/m} e^{(\beta + \rho\alpha)(x+c)} dx \\
&= \exp\left\{\frac{1}{2}\alpha^2(1 - \rho^2) + c(\beta + \rho\alpha)\right\} / m^k \left(\frac{1}{m} - \beta - \rho\alpha\right)^k \\
&= 1 + (c + mk)\rho\alpha + (c + mk)\beta\rho + \frac{1}{2}\alpha^2\{(c + mk)^2 + km^2\} + \alpha\beta\rho\{(c + mk)^2 + km^2\} + \dots
\end{aligned}$$

so that the mean y in the selected population is $(c + mk)\rho$ and the variance of y is $1 - \rho^2 + km^2\rho^2$ and the bivariate is $km^2\rho$.

These are the values which ⁽⁴⁾ (5), (6) take when $h = h_0$, ~~$n = n_0$~~

(and $v_1 = v_0 = 1$, since the units have been so chosen that in the Gaussian distribution the variances are both 1) with the requisite shifting of the origin, and change of ρ^2 to ρ .

It is thus verified that in selection of a Gaussian

population from a Type III one, or in selection of a Type III population from a Gaussian one, the formulas (4), (5), (6), (7) still hold, the origin for h , h_0 , and h , being the lower bound of the population from which selection is made, which for the Gaussian distribution is at $-\infty$ so that $h/h_0 = 1$

It is to be expected that the results of this chapter can be further extended by considering selection from a population in which the distributions of the two variates are not the same, such as those discussed in chapter V, and further generalised by considering more than two variates.

214 But the results here obtained may be of practical importance in all problems of selection from distributions in which the coefficient of variation is not small. The new results of importance are that in selection from a population in which the mean x is h_0 , mean y h , and the variance of x is v_0 , and of y v , and if it is assumed that the same type of Type III curve will fit both variates, so that $h_0^2/v_0 = h^2/v$, and if the selected population is determined by the conditions that x shall have a mean h and a variance v , then in the selected population the variance of y is

$$v_1(1-\rho)^2 + 2 \frac{h}{h_0} v_1 \rho(1-\rho) + \frac{v_1 v}{v_0} \rho^2$$

and the bivariate is $v \rho \sqrt{v_1/v_0}$.

The mean y is given by putting $x = h$ in the equation of regression. These formulae can be applied when the mean

variance are known, without carrying through any process of fitting a curve.

Chapter IX

CURVE FITTING

§22.1 The discussion of the function to represent a distribution in which two correlated variates each have a type III distribution, in chapters II and V has been wholly concerned with what R.A. Fisher calls Problems of Specification - the choice of the mathematical form of the population - and not at all with Problems of Estimation. The mathematical form of the population can be chosen on grounds quite independent of whether or not simple and logical means exist for fitting to an observed sample a function to represent the population. The arguments of chapters VI show that in any many-variate population in which each variate has a type=III distribution, it is reasonable to suppose that the combined distribution has the form given in its most general form in §19.4, a series in Laguerre polynomials, which, in the particular instance when there are only two variates, and they have the same distribution, can be represented by a Bessel function. It is important, however, when on purely theoretical grounds we have decided on a functional form, to ~~not~~ enquire what means exist to fit a function of that form, and in this

instances to enquire also in what kinds of problem the parameters involved in the distributions of two correlated variates may be assumed equal, so that the Bessel-function form can be assumed. The object of this chapter is to give some notes on these points, and also, further, on the question of what populations should have a type-III or Gamma-type curve fitted.

22.2 First, I shall quote a result which has been privately communicated to me by Mr D.S. Rajabhushanam, and which seems to me to give the Gamma-type curve a theoretical importance greatly above that of the other Pearson curves, and comparable with that of the Gaussian curve. His argument is as follows : It is well known¹⁾ that the deduction that if the arithmetical mean of a sample always gives the most probable value of a parameter in the population, then the distribution of the population must be Gaussian, given in the standard text^t-books, involves as to the nature of the population, certain assumptions, which though natural to assume in many instances, are not inevitable.

It is assumed that the probability for an observed value x is a function of $x - x_0$ where x_0 is one of the parameters. (It is usually stated to be the parameter which specifies the centre²⁾ or the true mean of the population, but that need not be included in the hypotheses.) As an alternative to this we may assume that it is a function of x/x_0 . This is at least as natural a hypothesis in all instances in which the variate is essentially positive, as it is large numbers of biological and economic data; heights, weights, ages, all kinds of biological measurements, real property, are all

¹⁾Whittaker and Robinson, Combinations of Observations,

things the measures of which must necessarily be positive.

And in any distribution in which the coefficient of variation has significance, and in any J-shaped distribution, it is more natural to assume that the frequency of a particular observed value x is a function of x/x_0 than of $x - x_0$. (The hypothesis does not exclude ~~the~~ J-shaped distributions as does the one which leads to a Gaussian distribution.) We make, then the hypothesis that the probability of an observed value ~~x is to be between x and $x + dx$ is $\varphi(x)dx$~~ between x and $x + dx$ is $\varphi(x)dx$, where $\varphi(x)$ is a function of x/x_0 , x_0 being one of the parameters which specifies the population, and $\varphi(x)$ is not otherwise a function of x . And the hypothesis that the arithmetic mean of an observed set of values x_1, x_2, \dots, x_n

gives the most probable value for x_0 . The condition that x is to be positive gives $\int_0^{\infty} \varphi(x)dx = 1$, from which it

follows that $\frac{x_0 \varphi(x)}{\varphi(x)/x_0}$ is a function of x/x_0 which is $f(x/x_0)/x_0$ not otherwise a function of x_0 . Put $\varphi(x) = \frac{f(x/x_0)}{x_0}$.

Then the condition which the method of maximum likelihood uses to estimate x_0 is that

$$x_0^{-n} f(x_1/x_0) f(x_2/x_0) \dots f(x_n/x_0)$$

shall be a maximum. That is that

$$\sum_{i=1}^n \{ \log f(x_i/x_0) - \log x_0 \}$$

shall be a maximum. The hypothesis of the arithmetic mean asserts that this condition is always equivalent to the condition that

$$\sum (x - x_0) = 0$$

The same argument as is used in the usual deduction of the Gaussian law shows that this is equivalent to ~~the~~ asserting that

$$\frac{\partial}{\partial x_0} \left\{ \log f\left(\frac{x}{x_0}\right) + \log x_0 \right\} = k(x - x_0)$$

for all x , k being some constant.

This leads to

$$\varphi(x) = a x^{-1+kx_0} e^{-kx_0 x}, \quad (1)$$

where

$$a = \frac{k^{kx_0}}{\Gamma(kx_0)}$$

x_0 being positive, k must be positive.

It has thus been shown that the Gamma-type distribution is one in which the arithmetic mean of a sample gives the most probable value for one of the parameters in the distribution function.

22.3 Poincaré has discussed the general form of the distribution function when no hypothesis of the kind here discussed is made, and finds the form

$$\theta(x) e^{Ax + B}, \quad (2)$$

where θ is an arbitrary function of x , and A of x_0 . B is a function of x_0 , which can be determined in terms of A .

This gives a Gaussian function when $\theta(x) = e^{-\frac{1}{2}h^2 x^2}$ and

$A = h^2 x_0$. It gives a Type III function when $\theta(x) = x^m$.

That the arithmetic mean is ~~a sufficient~~ ^{the best} statistic ^{for one parameter} in fitting

a Gamma-type curve is not a wholly new result. The

special interest of Rajabhushanam's result lies in the simplicity and naturalness of the hypothesis made - some special hypothesis would always have to be made before Poincaré's general result could be applied to actual curve fitting - and in the fact that the hypothesis made is suitable for a variate which is essentially positive. The curve fitting is thus not of a general Type-III distribution

$$y = y_0 (x - c)^{-m} e^{-px}$$

in which the three parameters c , m , and p are to be estimated from the sample, but of a function in which the value of c is determined from the natural origin of the distribution. It might be perhaps be convenient to reserve the name Gamma-type for a Type-III curve for which there is such a natural origin.

There are thus in this Gamma-type function two parameters to be estimated from a sample, and since Rajabhushanam's theorem has shown that the arithmetic mean is one statistic satisfying the conditions of the method of maximum likelihood, it is natural to enquire what other statistic the method of maximum likelihood will lead to. The Gamma-type distribution function being

$$\varphi(x) = \frac{k^{n+1}}{\Gamma(n+1)} x^n e^{-kx}, \quad (3)$$

the equations given by the method of maximum likelihood to determine the required statistics are

$$\frac{\partial}{\partial n} \{ \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) \} = 0, \quad \text{and} \quad \frac{\partial}{\partial k} \{ \varphi(x_1) \varphi(x_2) \dots \varphi(x_n) \} = 0;$$

i.e.,

$$\frac{\partial}{\partial m} \left[\frac{k^{n(m+1)}}{\{\Gamma(m+1)\}^n} (x_1 x_2 \dots x_n)^n \right] = 0,$$

and

$$\frac{\partial}{\partial k} \left[k^{n(m+1)} e^{-k(x_1 + x_2 + \dots + x_n)} \right] = 0.$$

These equations are equivalent to

$$n \log k + \sum_{r=1}^n \log x_r - n \frac{\Gamma'(m+1)}{\Gamma(m+1)} = 0,$$

$$\frac{n(m+1)}{k} - \sum_{r=1}^n x_r = 0.$$

That is,

$$\frac{\sum x}{n} = \frac{m+1}{k}, \quad (4)$$

$$\frac{\sum \log x}{n} = \frac{\Gamma'(m+1)}{\Gamma(m+1)} - \log k; \quad (5)$$

so that the most probable values for m and k are to be found by equating certain functions of m and k to the arithmetic and geometric means of the sample. To solve for m and k , write (4) as

$$\log \frac{\sum x}{n} = \log(m+1) - \log k$$

and eliminate k ; putting $m+1 = p$.

$$\log \frac{\text{arith. mean}}{\text{geom. mean}} = \log \frac{\sum x}{n} - \frac{\sum \log x}{n} = \log p - \frac{\Gamma'(p)}{\Gamma(p)}.$$

The process of fitting will thus require a table giving the values of $\log p - \Gamma'(p)/\Gamma(p)$. Having found the arithmetic and geometric means of the sample, find from the table the value

of p for which this function is equal to $\log \frac{\text{arith. mean}}{\text{geom. mean}}$.
The value of p , or of m being thus found, the value of k is readily found from (4)

The Psi-log-difference function

22.4

The function $\log x - \Gamma'(x)/\Gamma(x)$, the difference between the natural logarithm and the logarithmic derivative of the Gamma function of x , I propose to call the psi-log-difference function, and to denote it, if it is necessary to have a short notation, by $\text{pld } x$. Since in the use of the table values of the psi-log-difference function have to be equated to logarithms, and since common logarithms to base 10 will be used, it is convenient to have a table of $\log_{10} x - [\Gamma'(x)/\Gamma(x)] \log_{10} e$; we may call this function the common psi-log-difference function, the other being the natural psi-log-difference function. They can be denoted by $\text{pld}_{10} x$ and $\text{pld}_e x$. Tables are given at the end of this thesis

22.5 Fitting a curve

The Gamma-type curve is J-shaped when $-1 < m < 0$, it touches the vertical axis at the origin and when $0 < m < +1/2$ and to test the practicability of fitting by means of the arithmetic and geometric means, one would like to fit one example of each of those types. I have carried through the calculation for fitting a curve to the data for deaths from diphtheria at different ages in Table XII, p.98 in G.U. Yule's Theory of Statistics, 1929, but the result is a rather poor fit, the curve obtained giving values about ten per cent too ^{low} ~~high~~ in the neighbourhood

of the mode, and much too high values between 5 and 30 years. This may mean that this particular set of data, although having a general resemblance to a Gamma type distribution with $0 < m < 1$, is not in fact one which is suitably fitted by such a curve. Or it may mean that the geometric mean is not accurately calculated from a grouped frequency table of that sort. I have assumed ^{in estimating} ~~for~~ the geometric as well as ~~for~~ the arithmetic mean that the whole frequency in a certain class is concentrated at the mid-point of the class. It would be better to take it at the geometric mean of the lower and upper bounds of the class. But in the table discussed the lowest class is given as 0-1, and it contains about five per cent of the whole sample. If we are to get satisfactory results from an estimate of the geometric mean, we need in preparing a frequency table to aim not at having equal class intervals, but at having a constant ratio of the lower to the upper bound of the class interval- the effects would of course only be significantly different in examples in which the Gamma-type curve fitted has a fairly small value for m , that is those in which the coefficient of variation is large. But the question whether in examples in which the coefficient of variation is small, and a type-III curve can be fairly well fitted by means of the arithmetic mean and standard deviation, it is worth while to use the geometric mean, is one which perhaps can only be answered when a number of trials have been made. It would be of interest if

the data were available to fit a J-shaped Gamma-curve to a distribution of incomes. It is recognised that in dealing with incomes or amounts of wealth a geometric mean may have significance, and of course the arithmetic mean has significance, since it depends on the aggregate income. Hitherto it has been a frequent practice to represent the distribution of incomes over a certain range by Pareto's law¹⁾, in which the frequency is proportional to x^m . This can in no case represent the distribution of incomes over the whole range, since $\int x^m dx$ does not converge for any value of m . It would then be of interest to see if the method of fitting here given would give a satisfactory fit, but published data of incomes generally have too rough a grouping at the lower end for an accurate determination of the geometric mean. To sum up, to test adequately the proposed method of fitting a curve by use of the arithmetic and geometric means, we need data tabulated in a form which is throughout concerned with the ratios of the various values of the variate more than with the differences between them - possibly in some instances two methods of tabulation would be desirable for the calculations of the two means.

Any badness of fit in applying this method to a distribution to which a Gamma-type curve ought to be fitted must be due to defects either of the tabulation or of the method of deriving the arithmetic and geometric means from the frequency table, for the

¹⁾ Bowley, Elements of Statistics

method is on firm theoretical ground, being derived from the method of maximum likelihood.

22.6

And the method of maximum likelihood I regard as derived from the principles of inverse probability. If p_1, p_2, \dots, p_n are the a priori probabilities of various alternative hypotheses, and q_1, q_2, \dots, q_n the probabilities ~~that~~ for the occurrence of an observed event, on these various hypotheses, then the probability for the r^{th} hypothesis is $p_r q_r / \sum p q$. The method of maximum likelihood amounts practically to saying that this is a maximum ~~for~~ when q is a maximum. ~~A~~ similar but more far-reaching hypothesis is really implied in nearly all deduction from statistical data; for whenever an estimate of a parameter of some distribution, or an estimate of some unknown such an element of a planetary orbit, is given with a probable error stated, what is stated is that the probability that the parameter or unknown lies in a certain range is one-half (and of course the distribution of the estimates of the parameter being known or assumed, other statements about the probability that the parameter should lie within various other ranges are also implied); and since the estimate of the probable error, etc., is made from the sample alone, it follows that the assumption is implied that p is a constant. And this assumption is very often fully justifiable by the fact that the a priori probability is fairly constant in the range which the

sample makes possible. That is to say, p is effectively constant within the range outside which q is effectively zero. In statistical investigations the 'hypothesis' is always that some parameter has a certain value, and the observed event is recorded in a set of values of a sample. If on the hypothesis that the parameter has a value m the probability that a value between x and $x + dx$ will be observed is $\varphi(x, m)dx$, then when values x_1, x_2, \dots, x_n have been observed, the probability denoted above by q will be $\varphi(x_1, m) \varphi(x_2, m) \dots \varphi(x_n, m) dx_1 dx_2 \dots dx_n$.

Now the method of maximum likelihood asserts that that value for m is most likely which makes

$\varphi(x_1, m) \varphi(x_2, m) \dots \varphi(x_n, m)$ a maximum, $dx_1 dx_2 \dots dx_n$ being independent of m . (dx_1, dx_2, \dots should not be regarded as arbitrary differentials, but as the unit in ~~the~~ which the measurements are made, and which may differ in different parts of the table - being for ~~example~~ example, in the table from Yule referred to, a year at the lower end of the table and ten years at the upper - but which is independent of the value m of the parameter.)

We see thus that the method of ~~maximum likelihood~~ maximum likelihood is deducible from the principle of inverse probability. Actually the methods of inverse probability give a probability distribution function for the parameter m , $\Phi(m)dm$, and if as may often be the case we wish to consider instead of m some function of it,

m , say, and $\Phi(m)dm = \Phi(m)dm$, then the maximum value of $\Phi(m)$ does not necessarily correspond to the maximum value of $\Phi(m)$. For example, it follows that the most probable value of the square of some parameter is not generally the square of the most probable value. But in practice we generally assume it to be so, and clearly the earlier assumption that the a priori probability is constant within the range of possible values of m is only justifiable if dm/dm is practically constant ~~within that range~~ when m is any reasonable function of m .

Fitting a Type III curve in general

22-7 It may be of interest that for the fitting of a type III curve in which the origin is not supposed determined by natural considerations :

$$\Phi(x) dx = \frac{k^{n+1}}{\Gamma(n+1)} (x-c)^n e^{-k(x-c)} dx,$$

the method of maximum likelihood gives equation which equate the arithmetic, geometric and harmonic means of the sample to those of the population, but the equations are not in a convenient form for solution.

$$\frac{k^{n+1}}{[\Gamma(n+1)]^n} e^{-nkx} \left\{ (x_1-c)(x_2-c) \dots (x_n-c) \right\}^n e^{-k(x_1+x_2+\dots+x_n)}$$

is a maximum when

$$\frac{n}{n} \left(\frac{n}{x_1-c} + \frac{n}{x_2-c} + \dots + \frac{n}{x_n-c} \right) = \frac{k}{m}$$

$$\frac{x_1 + x_2 + \dots + x_n}{n} = \frac{n+1}{k} + c$$

$$\frac{\log\{(x_1 - c)(x_2 - c) \dots (x_n - c)\}}{n} = \frac{\Gamma'(m+1)}{\Gamma(m+1)} - \log k$$

The last two are equivalent to (4) and (5), and the first equates the harmonic mean of the sample to the harmonic mean of the population.

FITTING A TWO-VARIATE DISTRIBUTION

22.8 The Bessel-function distribution found in chapter III (equations 2 and 11) is one in which the distribution is of the same type in each variate ; in which the same values for the parameters occur for each. We cannot therefore assume that Bessel-function form for the distribution is suitable unless we have ground for supposing not only that a Gamma-type distribution will fit each variate, but also that the two distributions will have the same value of m . It is not necessary that they should have the same value of k , for the units in which the variates are measured can be changed, and kx taken instead of x . That is to say, the two variates must have the same ratio of arithmetic to geometric mean, even if they do not have the same arithmetic mean. There are some investigations in which this can be assumed of the population even if it is not exactly true of the sample. For example, in discussing the correlation between heights of father and son, it may be presumed that the distributions are of the same type, but it cannot be presumed, for example in discussing the correlation of height and weight. We have then to consider

how, in a distribution in which we may presume that the Bessel function type of distribution, that is in which we may presume that each of the two variates has the same type of distribution, the same ratio of geometric to arithmetic mean, or the same coefficient of variation, how we are to estimate the four parameters involved.

The four parameters are m , k_1 , k_2 , and ρ where the frequency functions for x and y respectively are

$$x^m e^{-k_1 x} \frac{k_1^{m+1}}{\Gamma(m+1)} \quad \text{and} \quad y^m e^{-k_2 y} \frac{k_2^{m+1}}{\Gamma(m+1)}.$$

The hypothesis of sameness of type requires that m shall be the same in each frequency function, but does not require that k shall be the same in each, since that depends on the unit chosen in which to measure the variate. Now if the units are chosen so as to make the values of k the same in each function, then a sample of N pairs of values of x and y gives us two samples drawn from the same population. Taken together they do not form a random sample of $2N$ values, for the pairs of values are correlated. But the geometric mean of the whole $2N$ values we may, so far as I can see, use in the same way as the geometric mean of a random sample of $2N$ values, and thus equate

$$\log(m+1) - \frac{\Gamma'(m+1)}{\Gamma(m+1)} \quad \text{to} \quad \log(\text{arith. mean}) - \frac{1}{2} \log(\text{geom. mean of } x + \text{g.m. } y)$$

But the ratio of geometric to arithmetic mean is not altered by the change of unit, and we have thus as the

rule by which m is to be estimated

$$\log(m+1) - \frac{\Gamma'(m+1)}{\Gamma(m+1)} = \frac{1}{2} \left\{ \log \frac{\text{a.m. of } x}{\text{s.m. of } x} + \log \frac{\text{a.m. of } y}{\text{s.m. of } y} \right\}$$

Having found m , the values of k_1 and k_2 are found from

$$\text{arith. mean of } x = \frac{m+1}{k_1}, \quad \text{arith. mean of } y = \frac{m+1}{k_2}$$

It remains for some other occasion to discuss the maximum likelihood estimate of ρ . It is clear that the product moment gives a consistent statistic for the estimate of ρ , but it does not follow that it is the best.

If it cannot be presumed that x and y ~~EXPRESSIBLE~~ as a series of Laguerre polynomials have the same frequency function, then the required values for m_1, m_2, k_1, k_2 will be estimated from the respective samples and ρ will be the only parameter in ~~which~~ estimating which both variates have to be used.

Chapter X

TABLES OF THE PSI-LOG-DIFFERENCE FUNCTION

In fitting a gamma-type curve by means of the arithmetic and geometric means, it is necessary to find a value for p such that

$$\log (\text{arith. mean}) - \log (\text{geome. Mean}) = \log p - \frac{\Gamma'(p)}{\Gamma(p)}$$

and for this reason it is desirable to have a table giving values of $\log p - \frac{\Gamma'(p)}{\Gamma(p)}$. I have prepared a table, giving values of this function from $p = 0.02$ to 2.00 at intervals of 0.02 . But since in general logarithms to base ten will be used for the arithmetic and geometric means, I have also prepared tables giving $\log_{10} p - \frac{\Gamma'(p) \log_{10} e}{\Gamma(p)}$ as follows :

II

- A) From 0.0001 to 0.0010 at intervals of 0.0001
- B) From 0.001 to 0.005 at intervals of 0.0005
- C) From 0.01 to 1.00 at intervals of 0.01
- D) From 2.0 to 4.0 at intervals of 0.1
- E) From 4.0 to 10.0 at intervals of 0.5
- F) From 10 to 20 at intervals of 2
- G) From 20 to 80 at intervals of 10

In preparing my first tables (of the natural psi-log-difference function from 0 to 2) I used E. Paarmans tables for the psi function, Tracts for Computers No. I, taking the interval as 0.02 as in her tables ; and Kohler's table of the natural logarithm to eight decimal places (1862), which

the director of the Kodaikanal observatory, Dr A.L. Narayanan, kindly lent me. Since the log tables give logs of whole numbers, one has to subtract $\log 100$. Thus for the part of the table from 1 to 2, the number entered in the table is the sum of three quantities, a logarithm, correct to eight places, $\frac{\Gamma'(x)}{\Gamma(x)} \log e$ correct to eight places, and $\log 100$ correct to a larger number of places, and it will therefore in some instances be wrong by one unit in the eighth place, but it should in no instance be wrong by more ~~than~~ than one unit.

For the part from 0 to 1 it has been necessary to calculate $\frac{\Gamma'(x)}{\Gamma(x)}$ from the equation $\frac{\Gamma'(x)}{\Gamma(x)} = \frac{\Gamma'(x+1)}{\Gamma(x+1)} - \frac{1}{x}$ and for that part of the table we have therefore the sum of four quantities, of which two are given to eight decimal places. The other two I took together, using a value for $\frac{1}{x} - \log 100$ correct to more than eight places, so that in that part of the table also the error should not exceed one unit in the eighth place. All this calculation was done without a machine, and difference calculated as a means of tracing mistakes. The part from 0 to 0.6 I have calculated twice, so as by two independent calculations to find mistakes, and the part from 0 to 0.1 further checked by comparison with the table of the common psi-log-difference function, prepared in Edinburgh subsequently. The unevenness in the fourth differences appears not to be greater than might arise from errors of not more than one unit in the entered values.

For the second set of tables (of the common psi-log-difference function, $= \log_{10} p - \frac{\Gamma'(p)}{\Gamma(p)} \log_{10} e$) I used H.T. Davis table of the psi function) multiplying by $\log e = .4342944819$
 Indiana, 1933

with a multiplying machine, either the electric Muldivo in the Jane Findlay Thomson Commercial Laboratory, or an Archimedes. The electric machine had the disadvantage that it will not take a multiplier of more than eight figures. I decided not to aim at such accuracy as before, but use seven figure logarithms (Shortrede's), giving the tables finally to six figures. One advantage of the electric machine is that when the value of $\log_e e$ is set up as multiplicand, and remains set up through a sequence of operations, and the values of $\frac{\Gamma'(p)}{\Gamma(p)}$ used as multipliers, these multipliers can be allowed to cumulate on the upper register, and at the end of ~~five~~ five or ten multiplications, the sum compared with the sum of the five or ten values of $\frac{\Gamma'(p)}{\Gamma(p)}$ found either on another machine (I used the small Muldivo for that purpose sometimes) or on the same machine subsequently. In this way it is possible to guard against mistakes in entering the multipliers. The differencing has mostly been done on one of the machines, retaining usually eight or nine places, though only seven can be correct - in that way lessening the error in the calculated differences. I have finally given the tables to six places, as more than that is hardly likely to be accurate.

If repeating the work for publication I should ask permission to transcribe the necessary logarithms from Sang's manuscript tables in the library of the Royal Society of Edinburgh (unless the publication of Logarithmica Britannica has been completed by that time) to about eleven places, and use a machine which will take an eleven figure multiplier, giving the final tables to eight places.

p	$\log_e p - \frac{\Gamma'(p)}{\Gamma(p)}$	s^2	s^4	s^6
0	∞			
.02	46.63276630			
.04	22.29439105			
.06	14.33588189			
.08	10.42707069			
.10	8.12116985			
.12	6.60852513	44872949	17347415	19496589
.14	5.54460990	27767701	07668658	05824030
.16	4.75837168	18331111	03813931	02110115
.18	4.15544457	12708452	02069319	877016
.20	3.67960198	09155112	01201723	398885
.22	3.29531051	06803495	733012	206719
.24	2.97904399	05184890	471020	102671
.26	2.71462637	04037305	311699	60792
.28	2.49058180	03201419	213170	35043
.30	2.29855142	02578703	149684	21343
.32	2.13230807	02105671	107541	13426
.34	1.98712143	01740180	78824	8710
.36	1.85933659	01453513	58817	5770
.38	1.74608688	01256663	44580	3923
.40	1.64509380	01 ⁰⁴² 428 393	34266	2732
.42	1.55452465	893389	26684	1913
.44	1.47288939	771069	21015	1384
.46	1.39896482	669764	16830	894
.48	1.33173789	585189	13439	744
.50	1.27036285	514053	10892	559

b	$\log b - \frac{\Gamma'(b)}{\Gamma(b)}$	s^2	s^4	s^6
.50	1.27036285	514053	10892	559
.52	1.21412835	453819	8904	424
.54	1.16243203	402489	7340	302
.56	1.11476060	358499	6078	266
.58	1.07067416	320587	5082	172
.60	1.02979359	287757	4258	172
.62	.99179059	259185	3606	
.64	.95637944	234219	3051	
.66	.92331048	212304	2608	
.68	.89236456	192997	2237	
.70	.86334861	175927	1922	
.72	.83609193	160779	1666	
.74	.81044304	147297	1437	
.76	.78626712	135252	1252	
.78	.76344372	124469	1100	
.80	.74186501	114786	960	
.82	.72143416	106063	852	
.84	.70206394	98192	745	
.86	.68367564	91066	667	
.88	.66619800	84607	585	
.90	.64956643	78733	524	
.92	.63372219	73383	465	
.94	.61861178	68498	424	
.96	.60418635	64037	367	
.98	.59040129	59943	341	
1.00	.57721566	56190	302	

p	$\log p - \frac{\Gamma'(p)}{\Gamma(p)}$	δ^2	δ^4
1.00	.57721566	56190	302
1.02	.56459193	52739	271
1.04	.55249559	49559	250
1.06	.54089484	46629	221
1.08	.52976038	43920	204
1.10	.51906512	41415	187
1.12	.50878401	39097	163
1.14	.49889387	36942	157
1.16	.48937315	34944	140
1.18	.48021187	33086	124
1.20	.47136145	31352	123
1.22	.46283455	29741	102
1.24	.45460506 .44665789	28232	102
1.26	.44665789	26825	92
1.28	.43897897	25510	81
1.30	.43155515	24276	79
1.32	.42437409	23121	71
1.34	.41742424	22037	65
1.36	.41069476	21018	62
1.38	.40417546	20061	59
1.40	.39785677	19163	48
1.42	.39172971	18313	51
1.44	.38578578	17514	46
1.46	.38001699	16761	42
1.48	.37441581	16050	38
1.50	.36897513	15377	37

p	$\log p - \frac{r'(p)}{r(p)}$	s^2	s^4
1.50	.36897513	15377	37
1.52	.36368822	14741	36
1.54	.35854872	14141	29
1.56	.35355063	13520	38
1.58	.34868824	13033	23
1.60	.34395618	12519	30
1.62	.33934931	12035	23
1.64	.33486279	11574	23
1.66	.33049201	11136	23
1.68	.32623259	10721	17
1.70	.32208038	10323	23
1.72	.31803140	9948	17
1.74	.31408190	9590	14
1.76	.31022830	9246	19
1.78	.30646716	8921	14
1.80	.30279523	8610	15
1.82	.29920940	8314	11
1.84	.29570671	8029	16
1.86	.29228431	7760	10
1.88	.28893951	7501	9
1.90	.28566972	7253	11
1.92	.28247246	7016	10
1.94	.27934536	6789	11
1.96	.27628615	6573	8
1.98	.27329267	6365	
2.00	.27036284		

u π
p

$$\log_{10} p - \frac{\Gamma'(b)}{\Gamma(b)} \cdot \log_{10} e$$

.0001	4339.152 001
.0002	2168.023 979
.0003	1944.375 862
.0004	1082.588 661
.0005	865.538 259
.0006	720.852 541
.0007	617.515 969
.0008	540.021 303
.0009	479.753 706
.0010	431.544 450

B.

.0010	431.544 450
.0015	296.955 357
.0020	214.697 526
.0025	171.364 632
.0030	142.447 062
.0035	131.876 393
.0040	106.423 513
.0045	94.411 009
.0050	84.804 989

II C

p	$h_{d,p}$	S^2	S^4
.01	41.673040		
.02	20.252353		
.03	13.183311		
.04	9.682334		
.05	7.601071		
.06	6.225995		
.07	5.252387		
.08	4.528419		
.09	3.970037		
.10	3.526979		
.11	3.167311		
.12	2.870046	047604	
.13	2.620386	37268	2736
.14	2.407993	29669	1919
.15	2.225270	23988	1352
.16	2.066535	19659	974
.17	1.927459	16304	715
.18	1.804687	13664	535
.19	1.695579	11560	406
.20	1.598031	9861	316
.21	1.510344	8479	240
.22	1.431135	7337	196
.23	1.359263	6391	153
.24	1.293782	5598	124
.25	1.233900	4930	101
.26	1.178947	4363	82
.27	1.128358	3878	69
.28	1.081646	3462	57
.29	1.038396	3102	47
.30	0.998248	2790	40

.31	.960890	2518	33
.32	.926050	2279	29
.33	.893488	2069	24
.34	.862996	1884	21
.35	.834388	1720	18
.36	.807500	1574	16
.37	.782186	1444	14
.38	.758316	1328	12
.39	.735774	1223	10
.40	.714455	1129	9
.41	.694266	1045	8
.42	.675121	968	7
.43	.656945	899	7
.44	.639668	836	6
.45	.623226	778	5
.46	.607563	726	4
.47	.592625	678	4
.48	.578366	635	3
.49	.564742	594	3
.50	.551712	557	3
.51	.51 .539239	523	3
.52	.52 .527289	492	2
.53	.53 .515832	463	2
.54	.54 .504838	436	2
.55	.55 .494280	412	2
.56	.56 .484134	389	1
.57	.57 .474377	367	1
.58	.58 .464988	348	1
.59	.59 .455946	329	1
.60	.60 .447234	312	1

p	pld. p	S ²	S ²
61	438 833	296	1
62	430 729	281	1
63	422 906	267	1
64	415 350	254	1
65	408 049	242	1
66	400 989	230	1
67	394 159	220	1
68	387 549	209	1
69	381 148	200	1
70	374 947	191	1
71	368 938	182	
72	363 110	174	
73	357 457	167	
74	351 971	160	
75	346 645	153	
76	341 471	147	
77	336 445	141	
78	331 560	134	
79	326 809	130	
80	322 188	125	
81	317 692	120	
82	313 315	115	
83	309 053	111	
84	304 903	107	
85	300 858	103	
86	296 917	99	
87	293 074	95	
88	289 326	92	
89	285 670	89	
90	282 103	85	

P	Per P	S
.91	278 621	82
.92	275 222	80
.93	271 902	77
.94	268 660	74
.95	265 491	72
.96	262 395	70
.97	259 368	67
.98	256 408	65
.99	253 513	63
1.00	250 682	

λ	D	$\log p$	S^2	S^c
	2.0	.117417		
	2.1	.111441	575	
	2.2	.106039	496	15
	2.3	.101134	431	11
	2.4	.096660	377	8
	2.5	.092563	331	7
	2.6	.088798	293	5
	2.7	.085325	260	5
	2.8	.082113	232	4
	2.9	.079133	208	3
	3.0	.076361	187	3
	3.1	.073776	169	2
	3.2	.071360	153	2
	3.3	.069096	139	1
	3.4	.066971	126	1
	3.5	.064973	116	1
	3.6	.063090	106	1
	3.7	.061313	97	1
	3.8	.059634	89	1
	3.9	.058043	82	
	4.0	.056535		

E

t	pld. p	s'	s''
4.0	.056535	1937	516
4.5	.050033	1339	222
5.0	.044871	964	128
5.5	.040674	718	77
6.0	.037194	548	50
6.5	.034262	428	33
7.0	.031758	341	22
7.5	.029595	276	16
8.0	.027708	226	11
8.5	.026047	188	8
9.0	.024574	158	6
9.5	.023258	134	
10.0	.022077		

F

10.	.022077	1901	
12.	.018347	1078	414
14.	.015695	670	183
16.	.013713	444	91
18.	.012175	310	49
20.	.010948	225	

G

30 30.	.007279		
40 30.	.005451		
50 40.	.004358		
60 50.	.003629		
70 60.	.003109		
80 70.	.002720		
80.			

INDEX

Aitken, Dr A.C., 2, 5, etc.	2, 3, 5, 68, 89
Barnes, E.W.,	55
Bartlett (and Wishart)	28
Berger's definition of Laguerre polynomials	9
Bilinear series	2, 3
Binomial distribution	3, 8, 89
Bessel functions	27
Beta functions	41, 60
- - , incomplete	41
Calculating machines	7, 119
Campbell, J.T.	4, 68
Charlier's polynomials	4, 8
Davis, H.T.	118
Determinants	40, 53
Fisher, R.A. (see also maximum likelihood)	102
Gamma function	33
- - , incomplete	33-60
Gamma type curve	5, 8
Gonin, H.T.	3, 68, 89
Hadamard	33
Hankel	33
Hardy, G.H.	27
Hermite polynomials	2, 3, 6, 7, 64
- - , relation to Laguerre polynomials	9
Hille, Einar	9, 26
Hypergeometric distribution	4, 8, 93
Ingham, A.E.	30
Jacobi's polynomials	5
Kodaikanal observatory	118
Kohler's tables	117
Laguerre polynomials, simple	8
- - , generalised	5, 8, 9, 27, 64, 81
Ledermann	30
Maximum likelihood	106, 111
Mehler, G.	3, 68
Moment generating functions, defined	1, 2, 5
Non-linear regression	21

Orthogonal polynomials, defined	1
- - , normalised	1
Pearson, Karl, differential equation	87, 93
- , Type I	5
- , Type III	5, 87
- , Type IV	10
Pairman, E.	117
Pareto's law	110
Poincare,	106
Poisson distribution	4, 8
Rajabhushanam, D.S.	103, 106
Ramanovsky, V.	10-12
Shastri, N.A.	27
Shortrede's tables	119
Sonine polynomials	5
Tracts for computers	35, 117
Watson, G.N.	27, 33
Whittaker and Watson	35
- and Robinson	103
Wigert	9
Wishart and Bartlett	28
Yule, G.U.	108