EXPANSION OF BIVARIATE DISTRIBUTIONS
BY ORTHOGONAL POLYNOMIALS

By

Hilal Abbud Taki al-Bayyati

Submitted in Partial Fulfillment for the Requirements
of the Degree Master of Science in the
Mathematics Department of the
American University of Beirut
Beirut, Lebanon,

1967

EXPANSION OF BIVARIATE DISTRIBUTIONS BY ORTHOGONAL POLYNOMIALS

Hilal al-Bayyati

AKNOWLEDGMENTS

I am very grateful to Professor M. A. Hamdan for his illuminating suggestions and encouragement.

I like to express my thanks to Professor Khamis and Mr. Khoury for their help in suggesting references and sending photo-copies of some papers.

I thank Miss Mona Jabbour for her careful and neat typing.

ABSTRACT

This thesis deals with the structure of bivariate distributions. The properties of a bivariate distribution can be studied easily by an expansion in a canonical form, i.e. in a series bilinear in the appropriate orthogonal polynomials. Mehler (1866) gave an expansion of the bivariate normal distribution in terms of the Hermite-Chebyshev polynomials. Similar expansions of other bivariate distributions are spread over the period (1900 - 1967).

In Chapter I of this thesis, some special orthogonal polynomials are discussed briefly. These are the polynomials used in the canonical forms of the bivariate distributions.

In Chapter II, the various derivations of the Mehler identity for the bivariate normal are given (1900-1958). Some of these proofs originated in the course of the study of correlation in fourfold tables (Pearson(1900)) and the properties of the generating functions of Hermite - Chebyshev polynomials (Hardy 1933).

The aim of Chapter II is to bring these proofs together for easier reference and comparison. Besides, these proofs are reproduced in a simple and straight forward manner.

Chapter III deals with the expansion of the bivariate gamma distribution in terms of the Laguerre polynomials. This

expansion is due to Kibble (1941).

Chapter IV deals with the canonical forms of the bivariate Poisson (Campbell (1934)), the bivariate binomial and the bivariate Hypergeometric distributions (Aitken and Gonin (1935)). Campbell's derivation of the bivariate Poisson frequency function is indirect. An alternative direct derivation was given by Hamdan (1963). The author gives another direct derivation based on the limiting form of the bivariate binomial distribution.

Finally, Chapter V gives a series form of the bivariate beta distribution (Hamdan (1963)). This form is used by the author to give series forms of the bivariate t and F distributions.

The thesis does not give any statistical applications of these canonical forms. However, we refer to their use in the choice of classes in the Chi-Square test (Hamdan (1963)) and in the estimation of correlation in contingency tables with non-measurable characters (Lancaster and Hamdan (1964)).

TABLE OF CONTENTS

CHAPTER I

SOME ORTHOGONAL POLYNOMIALS AND THEIR GENERATING FUNCTIONS

1. Orthogonal Polynomials. Let $\alpha(x)$ be a fixed non-decreasing function, not constant, in $[a,b]$.

Definition 1.1. The class of functions $f(x)$ which are measurable with respect to $\alpha(x)$ and for which the Stieltjes-Lesbegue integral $\int_a^b |f(x)|^p \, d\alpha(x)$ exists is denoted by $L_\alpha^p(a,b)$.

Definition 1.2. $g_0(x)$, $g_1(x)$, ... $g_k(x)$, k finite or infinite, is said to be an orthonormal set with respect to $\alpha(x)$ if

$$(g_n, g_m) = \int_a^b g_n(x)\, g_m(x)\, d\alpha(x) = \delta_{nm}, \quad n,m = 0,1,\ldots,k$$

where

$$\delta_{nm} = 1 \quad \text{if } n = m$$

and

$$\delta_{nm} = 0 \quad \text{if } n \neq m$$

Such functions are necessarily linearly independent.

Theorem 1.1. Let

(1.1) $$f_0(x),\ f_1(x),\ f_2(x),\ \ldots\ f_k(x)$$

be linearly independent real valued functions belonging to the class $L_\alpha^2(a,b)$. Then an orthonormal set

(1.2) $$g_0(x),\ g_1(x),\ g_2(x),\ \ldots\ g_k(x)$$

exbsts such that, for $n = 0,1,2,\ldots,k$

(1.3) $$g_n(x) = a_{n0}f_0(x) + a_{n1}f_1(x) + \dots + a_{nn}f_n(x), \quad a_{nn} > 0$$

and the set is uniquely determined, (Szegö, (1939)).

Definition 1.3. The procedure of deriving (1.2) from (1.1) is called orthogonalization, commonly referred to as Schmidt's Process of orthogonalization, (Jackson (1941)).

Definition 1.4. Let $\alpha(x)$ be fixed non-decreasing function with infinitely many points of increase in $[a, b]$ and let the moments,

(1.4) $$C_n = \int_a^b x^n \, d\alpha(x) \quad , \quad n = 0,1,\dots \quad \text{exist.}$$

If we orthogonalize the set of non-negative powers of x:

(1.5) $1, x, x^2, \dots, x^n, \dots$, which are linearly independent,

we obtain a set of polynomials,

(1.6) $p_0(x), p_1(x), p_2(x), \dots p_n(x), \dots$, uniquely determined by the following conditions:

(a) $p_n(x)$ is a polynomial of degnee n, where the coefficient of x^n is positive;

(b) the system $\{p_n(x)\}$ is orthonormal, i.e.

$$\int_a^b p_n(x)\, p_m(x)\, d\alpha(x) = \delta_{nm} \quad , \quad n,m = 0,1,2,\dots$$

The existence of (1.4) is equivalent to the functions x^n, $n = 0,1,2,\dots$, belonging to the class $L_\alpha(a,b)$.

A similar definition holds if we have a density function $w(x)$, which is integrable over (a,b) with the following properties: (i) $w(x)$ is actually positive on a set of points such that its definite integral over (a,b) is positive; (ii) the moments must exist; (iii) $w(x)$ is continuous. The set $\{p_n(x)\}$ is called the set of orthogonal polynomials associated with $\alpha(x)$ (or $w(x)$). If the distribution is of type $w(x)$, the system

$$\left\{[w(x)]^{\frac{1}{2}}\, p_n(x)\right\}, \quad n = 0,1,2,\ldots.$$

is orthonormal in the usual sense. (see Jackson (1941)).

If $w(x)$ is a non-negative weight function, integrable over (a,b) and the integral is positive, the product $\{[w(x)]^{\frac{1}{2}}\, x^k\}$, $k = 0,1,2,\ldots$ will be taken as the functions $f_n(x)$, the corresponding functions $g_n(x)$ which are linear combinations of these will be $[w(x)]^{\frac{1}{2}}\, p_n(x)$, which are the normalized orthogonal polynomials, i.e.

$$\int_a^b p_m(x)\, p_n(x)\, w(x)dx = \delta_{nm}, \quad m,n = 0,1,2,\ldots$$

Each polynomial is of degree indicated by its subscript.

2. The Hermite-Chebyshev Polynomials. Let $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$ be the standardized normal frequency function.

Definition 1.4. The n^{th} standardized Hermite-Chebyshev polynomial, $H_n(x)$ is defined by (Szegö 1939)

$$(1.11) \qquad H_n(x)\, \phi(x) = (-d/dx)^n\, \phi(x) \,/\, \sqrt{n!}, \quad \text{hence}$$

$$(1.12) \qquad H_n(x) = \frac{1}{\sqrt{n!}} \left\{x^n - \frac{n(n-1)}{2.1!}\, x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2^2 \cdot 2!}\, x^{n-4} + \ldots\right\}.$$

Alternatively,

$$(1.13) \qquad H_n(x) = \frac{1}{\sqrt{n!}} \sum_{h=0}^{[n/2]} (-1)^h \frac{n!\; x^{n-2h}}{(n-2h)!\; 2^h h!}$$

If $H_n(x)$ is the coefficient of $t^n/\sqrt{n!}$ in the expansion of $k(x,t)$, then

$$k(x,t) = k(t) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \sum_{h=0}^{[n/2]} (-1)^h \frac{n!\, x^{n-2h}}{(n-2h)!\, 2^h\, h!} \frac{t^n}{\sqrt{n!}}$$

$$= \sum_{n=0}^{\infty} \sum_{h=0}^{[n/2]} (-1)^h \frac{(x)^{n-2h}\, t^{n-2h}\, (t^2)^h}{(n-2h)!\, h!\, 2^h}$$

$$= \sum_{n=0}^{\infty} \sum_{h=0}^{[n/2]} \frac{(xt)^{n-2h}}{(n-2h)!} \frac{(-t^2)^h}{2^h\, h!}$$

$$= \sum_{h=0}^{\infty} \sum_{n=2h}^{\infty} \frac{(xt)^{n-2h}}{(n-2h)!} \frac{(-\frac{1}{2} t^2)^h}{h!}$$

$$= \exp(xt) \cdot \exp(-\tfrac{1}{2} t^2)$$

(1.14) $$= \exp(xt - \tfrac{1}{2} t^2).$$

<u>Lemma 1.1.</u> Let $G(t)$ be the generating function of the set $\{p_n(x)\}$ orthogonal with respect to the weight function $w(x)$. Then $\int_{-\infty}^{+\infty} p_n(x)\, p_m(x)\, w(x)dx$ = coefficient of $(t)^n (u)^m$ in the expansion of $\int_{-\infty}^{+\infty} G(t)\, G(u)\, w(x)dx$.

Proof:

$$\int_{-\infty}^{\infty} G(t)\, G(u)\, w(x)dx = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} t^n p_n(x)dx \sum_{m=0}^{\infty} u^m p_m(x) \cdot w(x)dx$$

$$= \int \sum_n \sum_m t^n u^m p_n(x)\, p_m(x)\, w(x)dx$$

$$= \sum_n \sum_m t^n u^m \int p_n(x)\, p_m(x)\, w(x)dx,$$ it follows

$\int_{-\infty}^{\infty} p_n(x)\, p_m(x)\, w(x)dx$ is the coefficient of $t^n u^m$ in the expansion of

$$\int_{-\infty}^{\infty} G(t)\, G(u)\, w(x)dx.$$

<u>Theorem 1.2.</u> The set $\{H_n(x)\}$ is orthonormal on the standardized normal distribution.

Proof: Using Lemma 1.1, it follows that (1.15) $\int_{-\infty}^{+\infty} H_r(x) H_s(x) \phi(x)\,dx$ is the coefficient of $t^r u^s / \sqrt{r!\,s!}$ in the expansion of

$$\int_{-\infty}^{+\infty} k(t)\ k(u)\ \phi(x)\,dx \quad , \quad \text{which is equal to}$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(xt - \tfrac{1}{2}t^2)\ \exp(xu - \tfrac{1}{2}u^2)\ \exp(-\tfrac{1}{2}x^2)\,dx$$

$$= \frac{1}{\sqrt{2\pi}} \int \exp\left\{-\tfrac{1}{2}(t^2 + u^2)\right\} \exp(xt + xu - \tfrac{1}{2}x^2)\,dx$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}(t^2 + u^2)\right\} \int \exp\left\{-\tfrac{1}{2}\left[-2x(t+u) + x^2\right]\right\} dx$$

$$= \exp\left\{-\tfrac{1}{2}(t^2 + u^2)\right\} . \exp\left\{\tfrac{1}{2}(t+u)^2\right\}$$

$$= \exp(tu)$$

i.e. (1.15) is the coefficient of $\frac{t^r u^s}{\sqrt{r!\,s!}}$ in exp(tu)

$$= \delta_{rs} .$$

3. <u>The Laguerre Polynomials.</u> Let

(1.16) $g(x) = e^{-x} x^{p-1} / \Gamma(p)$, $0 \leq x < \infty$ be the gamma frequency function with parameter p.

<u>Definition 1.5.</u> The r^{th} Laguerre polynomial denoted by $L_r^{(p-1)}(x)$ is defined by (Szegö (1939))

(1.17) $L_r^{(p-1)}(x)\ g(x) = (\frac{-d}{dx})^r \left[x^r g(x)\right] / r!$. So that

$$L_0^{(p-1)}(x) = 1,$$

$$L_1^{(p-1)}(x) = x-p,$$

$$L_2^{(p-1)}(x) = \tfrac{1}{2}\left[x^2 - 2x(p+1) + p(p+1)\right] \ldots \text{etc.}$$

Generally,

(1.18) $$L_r^{(p-1)}(x) = \sum_{h=0}^{r} \binom{r+p-1}{r-h} \frac{(-x)^h}{h!} .$$

The generating function of the set $\{L_r^{(p-1)}(x)\}$ is

$$k(t) = \sum_{r=0}^{\infty} L_r^{(p-1)}(x)(-t)^r = \sum_{r=0}^{\infty} \sum_{h=0}^{r} \binom{r+p-1}{r-h} \frac{(-x)^h}{h!} t^r$$

$$= \sum_{r=0}^{\infty} \sum_{h=0}^{r} \binom{r+p-1}{r-h} \left(\frac{-xt}{1-t}\right)^h \frac{t^{r-h}(1-t)^h}{h!}$$

$$= \sum_{h=0}^{\infty} \sum_{r=h}^{\infty} \left\{\left(\frac{-xt}{1-t}\right)^h / h!\right\} \binom{r+p-1}{r-h} t^{r-h} (1-t)^h .$$

But

$$\sum_{r=h}^{\infty} \binom{r+p-1}{r-h} (1-t)^{h+p} t^{r-h} (1-t)^{-p}$$ is equal to

$= (1-t)^{-p}$. It follows.

$$k(t) = \sum_{h=0}^{\infty} \left\{\left(\frac{-xt}{1-t}\right)^h / h!\right\} (1-t)^{-p}$$

(1.19) $k(t) = (1-t)^{-p} . \exp(-xt / 1-t)$, where $L_r^{(p-1)}(x)$ is the coefficient of $(-t)^r$ in $k(t)$.

Theorem 1.3. The set of Laguerre polynomials $\{L_r^{(p-1)}(x)\}$ is orthogonal on the gamma distribution.

Proof: Using Lemma 1.1., it follows that,

(1.20) $$\int_0^{\infty} L_r^{(p-1)}(x)\, L_s^{(p-1)}(x)\, g(x)dx = \text{coefficient of } (-t)^r(-u)^s$$

in the expansion of

$$\int_0^{\infty} k(t)k(u)g(x)dx = \int_0^{\infty} (1-t)^{-p}(1-u)^{-p} \exp\left(\frac{-xt}{1-t} - \frac{xu}{1-u}\right) g(x)dx.$$

$$= (1-t)^{-p}(1-u)^{-p} \int_0^{\infty} \exp\{-x(t/(1-t) + u/(1-u) + 1)\}\, x^{p-1}/\Gamma(p)\, dx.$$

$$= (1-t)^{-p}(1-u)^{-p} \int_0^{\infty} \exp\{-x(1-ut)/(1-t)(1-u)\}\, x^{p-1}/\Gamma(p)\, dx.$$

$$(1.21) \qquad (1-t)^{-p}(1-u)^{-p}\left[\frac{1-ut}{(1-t)(1-u)}\right]^{-p} = (1-ut)^{-p}.$$

It follows (1.20) is equal to the coefficient of $(-t)^r(-u)^s$ in the expansion of (1.21). Hence L. H. S. of (1.20) becomes $\binom{p-1+r}{r}\delta_{rs} = \frac{\Gamma(p+r)}{\Gamma(p)\,r!}\delta_{rs}$.

The moment generating function of $L_r^{(p-1)}(x)\ g(x)$:

The moment generating function of the product $L_r^{(p-1)}(x)\ g(x)$ is

$$G_r(\alpha) = \int_0^\infty e^{\alpha x} L_r^{(p-1)}(x)\ g(x)dx$$

which is equal to the coefficient of $(-t)^r$ in

$$(1.22) \qquad \frac{1}{\Gamma(p)}\int_0^\infty e^{\alpha x}(1-t)^{-p}\exp(-xt/1-t)\ e^{-x}x^{p-1}dx$$

$$= (1-t)^{-p}\int_0^\infty \exp\left\{\alpha x - \frac{xt}{(1-t)} - x\right\}x^{p-1}/\Gamma(p).\ dx$$

$$= (1-t)^{-p}\int_0^\infty \exp\left\{-x\left[1/(1-t)-\alpha\right]\right\}x^{p-1}\ dx/\Gamma(p)$$

$$= \left(\frac{1}{1-t}-\alpha\right)^{-p}\frac{(1-t)^{-p}}{\Gamma(p)}\int_0^\infty \exp\left\{-x\left(\frac{1}{1-t}-\alpha\right)\right\}.$$

$$\left[x\left(\frac{1}{1-t}-\alpha\right)\right]^{p-1}dx\left(\frac{1}{1-t}-\alpha\right)$$

The quantity under the integral is equal to $\Gamma(p)$. Therefore (1.22) is equal to

$$(1-t)^{-p}\left(\frac{1}{1-t}-\alpha\right)^{-p} = \left[1-\alpha(1-t)\right]^{-p}$$

$$= \left[(1-\alpha)+\alpha t\right]^{-p}$$

$$= (1-\alpha)^{-p}\left(1+\frac{\alpha t}{1-\alpha}\right)^{-p}$$

It follows that $G_r(\alpha)$ is the coefficient of $(-t)^r$ in

$$(1-\alpha)^{-p}\left(1+\frac{\alpha t}{1-\alpha}\right)^{-p} = \binom{p+r-1}{r}(1-\alpha)^{-p}\left(\frac{\alpha}{1-\alpha}\right)^r.$$

It follows

(1.23) $$G_r(\alpha) = \frac{\Gamma(p+r)}{\Gamma(p)r!}(1-\alpha)^{-p}\left(\frac{\alpha}{1-\alpha}\right)^r$$

In particular, the moment generating function of $g(x)$ (i.e. where $r = 0$) is $G_0(\alpha) = (1-\alpha)^{-p}$.

The relation of Hermite polynomials to those of Laguerre, established by Szegö (1939), is

(1.24) $$H_{2m}(x) = (-1)^m 2^m m!\, L_m^{(-\frac{1}{2})}(x^2/_2) \;/\; \sqrt{2m!}$$

(1.25) $$H_{2m+1}(x) = (-1)^m m!\, 2^m x\, L_m^{(\frac{1}{2})}(x^2/_2) \;/\; \sqrt{(2m+1)!}$$

4. Jacobi polynomials. Consider the weight function

(1.26) $$w(x) = (1-x)^{\alpha}(1+x)^{\beta},\; x\in[-1,1],\; \alpha>-1,\; \beta>-1$$

so that $w(x)$ is integrable for $x \in [-1,1]$.

Definition 1.6. The n^{th} Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ is (Szegö (1939))

(1.27) $$P_n^{(\alpha,\beta)}(x) = (-1)^n/(2^n n!)\left\{(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}\left[(1-x)^{\alpha+n}(1+x)^{\beta+n}\right]\right\}$$

or

(1.28) $$P_n^{(\alpha,\beta)}(x) = \sum_{h=0}^{n}\binom{n+\alpha}{n-h}\binom{n+\beta}{h}\left(\frac{x-1}{2}\right)^h\left(\frac{x+1}{2}\right)^{n-h}$$

$$= \binom{n+\alpha}{n}\left(\frac{x+1}{2}\right)^n\sum_{h=0}^{n}\frac{n(n-1)\dots(n-h+1)}{(\alpha+1)(\alpha+2)\dots(\alpha+h)}\binom{n+\beta}{h}\left(\frac{x-1}{x+1}\right)^h$$

$$= \binom{n+\alpha}{n}\left(\frac{x+1}{2}\right)^n F\left(-n, -n-\beta; \alpha+1; \frac{x-1}{x+1}\right).$$

Another form of (1.27) (Szegö (1939) is

$$(1.29) \qquad P_n^{(\alpha,\beta)}(x) = \frac{1}{2\pi i}\int \left(1+\frac{x+1}{2}z\right)^{n+\alpha}\left(1+\frac{x-1}{2}z\right)^{n+\beta} z^{-n-1}\,dz,$$

where we assume that $x \not= \mp 1$. The integration is extended in the positive sense along a closed curve around the origin, such that the points $-2(x \pm 1)^{-1}$ lie neither on it nor in its interior. (We define the first and second factors of the integrand to be 1 for $z = 0$).

Using formula (1.29) the generating function of $P_n^{(\alpha,\beta)}(x)$, $\sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)\; t^n$, is

$$\begin{aligned} Q(t) &= \sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)t^n \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi i}\int \left(1+\frac{x+1}{2}z\right)^{\alpha+n}\left(1+\frac{x-1}{2}z\right)^{\beta+n} t^n(z)^{-n-1}\,dz \\ &= \frac{1}{2\pi i}\int \left(1+\frac{x+1}{2}z\right)^{\alpha}\left(1+\frac{x-1}{2}z\right)^{\beta} z\sum_{n=0}^{\infty} z^{-1}\left[\left(1+\frac{x+1}{2}z\right)\left(1+\frac{x-1}{2}\right)z\right]^n dz \end{aligned}$$

But $\frac{1}{z}\sum_{n=0}^{\infty}\left[\left(1+\frac{x+1}{2}z\right)\left(1+\frac{x-1}{2}z\right)\frac{t}{z}\right]^n$ is an infinite geometric series with sum

$$\frac{1}{z}\left\{1\Big/\left[1-\left[\left(1+\frac{x+1}{2}z\right)\left(1+\frac{x-1}{2}z\right)\right]\right]\right\}.$$

Hence the generating function is

$$\begin{aligned} Q(t) &= \frac{1}{2\pi i}\int \frac{\left(1+\frac{x+1}{2}z\right)\left(1+\frac{x-1}{2}z\right)}{z-t\left(1+\frac{x+1}{2}\right)\left(1+\frac{x-1}{2}z\right)}\,dz \\ (1.30) \qquad &= \frac{1}{2\pi i}\int \frac{\left(1+\frac{x+1}{2}z\right)\left(1+\frac{x-1}{2}z\right)}{z-t\left(1+\frac{x+1}{2}\right)\left(1+\frac{x-1}{2}z\right)}\cdot dz \end{aligned}$$

The denomenator of the integrand is

$$\frac{1}{4}(x^2-1)\;tz^2 - z(xt-1)-t = \frac{1}{4}(1-x^2)\,t(z-z_0)(z-Z_0)$$

Where

$$(1.31) \qquad z_0 = z_0(t) = \frac{2}{1-x^2} \cdot \frac{xt-1+R}{t},$$

$$(1.32) \qquad R = R(t) = (1-2xt+t^2)^{1/2},$$

$$(1.33) \qquad Z_0 = Z_0(t) = \frac{2}{1-x^2} \cdot \frac{xt-1-R}{t}.$$

z_0 and R are regular analytic functions of t provided $/t/$* is sufficiently small, taking $R(0) = 1$. At $t = 0$, $z_0(t)$ has a zero and $Z_0(t)$ has a pole. For sufficiently small $/t/$, z_0 lies in the interior, and Z_0 in the exterior, of the integration curve of (1.30). So by Cauchy's theorem for integration,

$$(1.34) \qquad \sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)\, t^n = [\tfrac{1}{4}(1-x^2)t]^{-1}(1+\frac{x+1}{2}z_0)^{\alpha}(1+\frac{x-1}{2}z_0)^{\beta} \cdot (z_0-Z_0)^{-1}, \text{ where}$$

$$1+\frac{x+1}{2}z_0 = 2(1-t+R)^{-1},$$

$$1+\frac{x-1}{2}z_0 = 2(1+t+R)^{-1},$$

$$z_0 - Z_0 = 4t^{-1}(1-x^2)^{-1}R, \quad \text{it follows}$$

$$(1.35) \qquad Q(t) = \sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)\, t^n = 2^{\alpha+\beta} R^{-1}(1-t+R)^{-\alpha}(1+t+R)^{-\beta}$$
$$= 2^{\alpha+\beta}(1-2xt+t^2)^{-1/2}\{1-t+(1-2xt+t^2)^{1/2}\}^{-\alpha} \cdot \{1+t+(1-2xt+t^2)^{1/2}\}^{-\beta}$$

Therefore (Szegö (1939));

$$\int_{-1}^{+1}(1-x)^{\alpha}(1+x)^{\beta} P_n^{(\alpha,\beta)}(x)\, P_m^{(\alpha,\beta)}(x)\,dx = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1}\frac{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}\delta_{nm}$$

---

* / / Stands for the absolute value.

Hence the set $\{ P_n^{(\alpha, \beta)}(x) \}$ is orthogonal with respect to the function $w(x) = (1 - x)^{\alpha} (1 + x)^{\beta}$.

5. The Tchebichef polynomials and Legendre polynomials.

Definition 1.7. The n^{th} Tchebichef polynomial of the first kind is (Szegö (1939)).

(1.36) $T_n(x) = \cos n\theta$, if $x = \cos\theta$, which is a special case of the n^{th} Jacobi polynomial with $\alpha = \beta = -\frac{1}{2}$.

Definition 1.8. The Tchebichef n^{th} polynomial of the second kind is defined by (Szegö (1939))

(1.37) $U_n(x) = \sin(n+1)\theta / \sin\theta$ if $x = \cos\theta$ which is a special case of the n^{th} Jacobi polynomial with $\alpha = \beta = \frac{1}{2}$.

Definition 1.9. Legendre n^{th} polynomial is defined as

(1.38) $$P_n(x) = \frac{1}{2^n\, n!} \frac{d^n}{dx^n} (x^2 - 1)^n ,$$ which is a special case of Jacobi polynomial with $\alpha = \beta = 0$.

The relations of polynomial (1.36) and (1.37) to the Jacobi polynomials are (Szegö (1939))

(1.39) $$P_n^{(\frac{1}{2},\frac{1}{2})}(x) = \frac{1.3.\ldots.(2n-1)}{2.4\ldots 2n} \frac{\sin[(2n+1)\theta/2]}{\sin(\theta/2)}$$

and

(1.40) $$P_n^{(\frac{1}{2},\frac{1}{2})}(x) = \frac{1.3.\ldots.(2n-1)}{2.4\ldots\ldots 2n} \frac{\cos[(2n+1)\theta/2]}{\cos(\theta/2)}$$

(1.41) $$P_n^{(-\frac{1}{2},-\frac{1}{2})}(x) = \frac{1.3.\ldots.(2n-1)}{2.4\ldots.2n} T_n(x)$$

(1.42) $$P_n^{(\frac{1}{2},\frac{1}{2})}(x) = 2\,\frac{1.3.\ldots.(2n+1)}{2.4.\ldots(2n+2)} U_n(x)$$

Definition 1.10. Jacobi polynomials with $\alpha = \beta$ are called ultraspherical polynomials. It follows that Tchebichef and Legendre polynomials are ultraspherical polynomials.

The Tchebichef polynomials $\{T_n\}$ and $\{U_n\}$ are orthogonal with respect to

$$w_1(x) = (1-x)^{-\frac{1}{2}}(1+x)^{-\frac{1}{2}}, \quad w_2(x) = (1-x)^{\frac{1}{2}}(1+x)^{\frac{1}{2}}$$

respectively, i.e. for $n \neq m$ (Szegö (1939))

$$\int_{-1}^{1} T_n(x)\, T_m(x)\, (1-x^2)^{-\frac{1}{2}}\, dx = \int_0^{\pi} \cos n\theta \cos m\theta\, d\theta = 0$$

and

$$\int_{-1}^{1} U_n(x)\, U_m(x)\, (1-x^2)^{\frac{1}{2}} dx = \int_0^{\pi} \sin(n+1)\theta \ \sin(m+1)\theta\, d\theta = 0.$$

So substituting for $\alpha = \beta = 0$ in (1.35), we get

(1.43) $\qquad H'(x,t) = (1 - 2xt + t^2)^{-\frac{1}{2}}$ as the generating function of Legendre polynomial.

6. The Poisson - Charlier Polynomials. Let x be a Poisson variable with parameter m, i.e. with frequency function

(1.44) $\qquad p(x) = p(x,m) = e^{-m}\, m^x / x!\,, \quad x = 0,1,2,\ldots$

Definition 1.11. The r^{th} Charlier polynomial $k_r(x,m) = k_r(x)$ is defined by

(1.45) $\qquad k_r(x)\, p(x) = (-1)^r \ \nabla^r\ p(x)$

where

$$\nabla p(x) = p(x) - p(x-1) = \Delta p(x-1).$$

It follows that

$$k_r(x) = \left[x^{(r)} - rmx^{(r-1)} + \tbinom{r}{2} m^2\, x^{(r-2)} + \ldots + (-1)^r m^r\right] / m^r$$

where

$$x^{(r)} = x(x-1)\ \ldots\ldots\ (x-r+1).$$

An alternative form of $k_r(x)$ (Szegö (1939)) is:

(1.46) $$k_r(x) = \sum_{h=0}^{r} (-1)^{r-h} \binom{r}{h}\binom{x}{h} h!\, m^{-h}$$

<u>Definition 1.12.</u> Define $k'_r(x)$ as

$$k'_r(x) = \frac{m^{r/2}}{(r!)^{\frac{1}{2}}} k_r(x).$$

The generating function of $k'_r(x)$ is $G(x,t) = G(t)$

$$G(t) = \sum_{r=0}^{\infty} \sum_{h=0}^{r} (-1)^{r-h} \frac{m^{r/2}}{(r!)^{\frac{1}{2}}} \binom{r}{h}\binom{x}{h} h!\, \frac{m^{-h}\, t^r}{m^{r/2}(r!)^{\frac{1}{2}}}$$

where $k'_r(x)$ is the coefficient of $t^r / m^{r/2}\sqrt{r!}$ in the expansion of $G(t)$.

$$G(t) = \sum_{r=0}^{\infty} \sum_{h=0}^{r} (-1)^{r-h} \binom{r}{h}\binom{x}{h} \frac{h!\, m^{-h}\, t^r}{(r!)}$$

$$= \sum_{h=0}^{x} \left\{ \sum_{r=h}^{\infty} \frac{(-1)^{r-h}\, t^{r-h}}{(r-h)!} \right\} \binom{x}{h} t^h m^{-h}$$

$$= \sum_{h=0}^{x} \exp(-t).\binom{x}{h}.(t/m)^h$$

(1.48) $$G(t) = \exp(-t)\,(1 + t/m)^x$$

<u>Theorem 1.4.</u> The set $\{k'_r(x)\}$ is orthonormal on the Poisson distribution.

Proof:

(1.49) $\sum_x p(x)\, k'_r(x) k'_s(x)$ is the coefficient of $\dfrac{t^r\, u^s}{m^{r/2} m^{s/2}\sqrt{r!\, s!}}$

in

$$\sum_x p(x)\; g(t)G(u) = \sum_x e^{-m} m^x (x!)^{-1} e^{-t}(1+t/m)^x e^{-u}(1+u/m)^x$$

$$= e^{-m-t-u} \exp\left[m(1+t/m)(1+u/m)\right]$$

(1.50) $$= \exp(tu/m).$$

It follows $\sum_x p(x)\ k'_r(x)\ k'_s(x)$ = to the coefficient of $\dfrac{t^r\ u^s}{m^{r/2} m^{s/2}\ \sqrt{r!s!}}$

in (1.50) and (1.49) becomes equal to

(1.51) $$= \delta_{rs}\ .$$

Therefore the set $\{k'_r(x)\}$ is orthonormal on the Poisson distribution, and when x is referred to the mean as origin and to the standard deviation as a unit, the Charlier polynomials tend to the Hermite - Chebychev polynomials as $m \to \infty$.

Lemma 1.2. If P(t) is the probability generating function (p.g.f.) of an integral valued random variable X, then P(1 + t) is the factorial moment generating function (f.m.g.f.) of X.

Proof: let $p_r(X = j) = f_j$ so that

(1.52) $$P(t) = \sum_{j=0}^{\infty} f_j\ t^j\ .$$

Hence

$$P(1+t) = \sum_{j=0}^{\infty} f_j (1+t)^j$$

$$= \sum_{j=0}^{\infty} f_j \sum_{i=0}^{\infty} \binom{j}{i} t^i$$

$$= \sum_{i=0}^{\infty} \frac{t^i}{i!} \sum_{j=0}^{\infty} f_j\ j(j-1)\ldots(j-i+1)$$

(1.53) $$= \sum_{i=0} t^i\ \mu'_{[i]} / i!$$

where $\mu'_{[i]}$ is the i^{th} factorial moment. Hence P(1 + t) is the f.m.g.f., the factorial moment of order i being the coefficient of $t^i / i!$ in P(1+ t).

Lemma 1.3. The factorial moment generating function of the Poisson distribution with parameter m is $\exp(m\alpha)$

Proof: By lemma 1.2, the f.m.g.f. of p(x) is

$$F(\alpha) = \sum_{x=0}^{\infty} (1+\alpha)^x e^{-m} m^x (x!)^{-1}$$

$$= \sum_{x=0,1,\ldots}^{\infty} [(1+\alpha)m]^x \exp(-m-m\alpha+m\alpha) / x!$$

$$= \exp(m\alpha) \sum_x [(1+\alpha)m]^x \exp[-m(1+\alpha)]/x!$$

(1.54) $$= \exp(m\alpha)$$

A bivatiate f.m.g.f. corresponding to a bivariate frequency function p(x,y) is similarly defined as

$$F(\alpha,\beta) = \sum_x \sum_y (1+\alpha)^x (1+\beta)^y \; p(x,y)$$

$$= P(1+\alpha,\; 1+\beta).$$

Where P(t,u) is the bivariate p.g.f.

Lemma 1.4. The f.m.g.f. of $k_r(x)$ p(x) is $\alpha^r \exp(m\alpha)$. This is called Campbell's lemma because it was proved by Campbell (1932).

Proof: By lemma 1.3,

$$F(\alpha) = \sum_x (1+\alpha)^x p(x) = \exp(m\alpha).$$

Hence

$$\sum_x (1+\alpha)^x k_r(x)\; p(x) = \sum_x (1+\alpha)^x (-1)^r \nabla^r p(x)$$

$$= (-1)^r \sum_x (1+\alpha)^x [p(x) - \tbinom{r}{1} p(x-1) + \tbinom{r}{2} p(x-2) + \ldots + (-1)^r p(x-r)].$$

$$= (-1)^r \exp(m\alpha)[1 - \tbinom{r}{1}(1+\alpha) + \tbinom{r}{2}(1+\alpha)^2 + \ldots + (-1)^r (1+\alpha)^r]$$

$$= (-1)^r \exp(m\alpha)[1-(1+\alpha)]^r \quad \text{(by the binomial theory)}$$

$$= (-1)^r \exp(m\alpha)\; (-\alpha)^r$$

(1.55) $$= \alpha^r \exp(m\alpha).$$

7. <u>Krawtchouk's polynomials and the Factorial Moments of the Binomial Distribution.</u> Let x be a random variable with frequency function

(1.56) $$b(x) = b(x;n,p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0,1,2,\dots n,$$

where

$$0 < p < 1 \quad \text{and} \quad q = 1 - p.$$

<u>Definition 1.13.</u> The r^{th} Krawtchouk polynomial $G_r(x,n,p) = G_r(x)$ is defined by

(1.57) $$G_r(x)\, b(x) = (-q)^r \Delta^r \; x^{(r)} b(x)$$, it follows that

(1.58) $$G_r(x) = x^{(r)} - \binom{r}{1} p(n-r+1)x^{(r-1)} + \binom{r}{2} p^2(n-r+2)^{(2)} x^{(r-2)} + \dots + (-1)^r p^r n^{(r)}.$$

An alternative form for (1.58), is

(1.59) $$G_r(x) = r! \sum_{h=0}^{r} (-1)^{r-h} \binom{x}{h} \binom{n-h}{r-h} p^{r-h} \quad x = 0,1,2,\dots n.$$

The generating function of the set $\{G_r(x)\}$ is

(1.60) $$k(x,t) = \sum_{r=0}^{n} G_r(x) t^r / r!$$

in the sense that $G_r(x)$ is the coefficient of $t^r / r!$ in $k(x,t)$.

We have

$$k(x,t) = \sum_{r=0}^{n} \sum_{h=0}^{r} (-1)^{r-h} \binom{x}{h} \binom{n-h}{r-h} (pt)^{r-h} t^h$$

$$= \sum_{h=0}^{\infty} \left\{ \sum_{r=h}^{\infty} (-1)^{r-h} \binom{n-h}{r-h} (pt)^{r-h} \right\} \binom{x}{h} t^h$$

$$= \sum_{h=0}^{\infty} \binom{x}{h} (1 - pt)^{n-h} t^h$$

$$= (1 - pt)^n \sum_{h=0}^{\infty} \binom{x}{h} \left(\frac{t}{1 - pt}\right)^h$$

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

$$= (1-pt)^{n-x} (1 - pt + t)^x$$

(1.61) $$= (1-pt)^{n-x} (1 + tq)^x$$

<u>Theorem 1.5:</u> The set $\{G_r(x)\}$ is orthonormal on the binomial distribution.

Proof: $G_r(x)G_s(x)b(x)$ is the coefficient of $u^r v^s / r!\, s!$ in the expansion of

(1.62) $$\sum_{x}^{\infty} k(x,u)k(x,v)b(x)$$

$$= \sum_x \binom{n}{x} p^x q^{n-x} (1 + qu)^{x} (1 + qv)^x (1-pv)^{n-x} (1-pu)^{n-x}$$

$$= \sum_x \binom{n}{x} \left[p(1 + qu)(1 + qv)\right]^x \left[q(1-pu)(1-pv)\right]^{n-x}$$

$$= \left[p(1 + qu)(1 + qv) + q(1-pu)(1-pv)\right]^n$$

(1.63) $$= (1 + pquv)^n.$$

It follows that

$$G_r(x)G_s(x)b(x) = \text{coefficient of } u^r v^s / r!\, s! \text{ in } (1.63)$$

$$= \delta_{r,s}, = 0,1,2,\ldots n$$

The factorial m.g.f. corresponding to $b(x)$ is given by

(1.64)
$$F(\alpha) = \sum_{x=0}^{n} (1+\alpha)^x \, b(x)$$
$$= \sum_{x=0}^{n} (1+\alpha)^x \binom{n}{x} p^x q^{n-x}$$
$$= \sum_{x=0}^{n} \binom{n}{x} \left[p(1+\alpha)\right]^x q^{n-x}$$
$$= \sum_{x} \binom{n}{x} \left[p(1+\alpha)\right]^x (1-p)^{n-x}$$
$$= (p + p\alpha + 1 - p)^n = (1 + p\alpha)^n$$

(1.65)
$$F(\alpha) = (1 + p\alpha)^n$$

The f.m.g.f. of the product $G_r(x)\ b(x)$ is

$$\sum_{x=0}^{n}(1+\alpha)^x G_r(x)\ b(x) = \sum_{x=0}^{n} (1+\alpha)^x (-q)^r \Delta^r\left[x^{(r)} b(x)\right]$$
$$= \sum_{x} (1+\alpha)^x q^r \alpha^r (x+r)^{(r)}\ b(x+r)$$
$$= \sum_{x} (1+\alpha)^x q^r \alpha^r (x+r)^{(r)} \binom{n}{x+r} p^{x+r} q^{n-x-r}$$
$$= \alpha^r q^r p^r \sum_{x} (x+r)^{(r)} \frac{n!}{(n-x-r)!\,(x+r)!} p^x q^{n-r-x} (1+\alpha)^x$$
$$= (\alpha q p)^r \sum_{x} \frac{(x+r)^{(r)} x!\, n!}{(x+r)!\,(n-r-x)!\,x!} \left[p(1+\alpha)\right]^x q^{n-r-x}$$
$$= (\alpha q p)^r \sum_{x} \frac{n^{(r)} (n-r)!}{(n-r-x)!\,x!} \left[p(1+\alpha)\right]^x q^{n-r-x}$$
$$= (\alpha q p)^r n^{(r)} \left[q + p(1+\alpha)\right]^{n-r}$$

(1.66)
$$= (\alpha q p)^r n^{(r)} (1 + p\alpha)^{n-r}$$

In particular, the f.m.g.f. of $b(x)$ itself is $(1 + p\alpha)^n$, as may be seen directly by substitution $1+\alpha$ for t in the p.g.f. $(pt+q)^n$.

8. The Aitken-Gonin polynomials and the Factorial moments of the hypergeometric distribution.

Under simple sampling n times without replacement from a population of size N of which Np indivivals possess the character A and Nq possess $\bar{A}$, the probability of x indivivals possessing A is the hypergeometric frequency distribution.

$$(1.67) \qquad h(x) = h(x;n,N,p) = \binom{Np}{x}\binom{Nq}{n-x} \Big/ \binom{N}{n}$$

Definition 1.14. The r^{th} Aitken-Gonin polynomial $U_r(x) = U_r(x;n,N,p)$ is defined (Aitken and Gonin (1935)) as

$$(1.68) \qquad U_r(x)\, h(x) = (-1)^r \Delta^r \left[x^{(r)} (Nq-n+x)^{(r)}\, h(x)\right] / (N-r+1)^{(r)}.$$

It follows

$$(1.69) \qquad U_r(x) = (x)^{(r)} - \frac{r(n-r+1)(Np-r+1)}{(N-2r+2)} x^{(r-1)} + \binom{r}{2}\frac{(n-r+2)^{(2)}}{(N-2r+3)^{(2)}} x^{(r-2)} + \dots\dots + (-1)^r \frac{n^{(r)}(Np)^{(r)}}{(N-r+1)^{(r)}}.$$

$$= \sum_{h=0}^{r} (-1)^h \binom{r}{h} (x)^{(r-h)} \frac{(Nq-n+x+h)^{(h)}(Np-r+h)^{(h)}}{(N-2r+h+1)^{(h)}}.$$

Alternatively (1.69) can be written symbolically (Aitken & Gonnin (1935)) as

$$(1.70) \qquad U_r(x) = F(n-r+1,\ Np-r+1,\ N-2r+2;\ -\Delta)x^{(r)}$$

Theorem 1.6. The set $\{U_r(x)\}$ is orthogonal on the hypergeometric distribution.

Proof: To verify the orthogonality it is enough to consider

$$(1.71) \qquad \sum_{0}^{n} (N-r+1)^{(r)}\, x^{(s)}\, U_r(x)\, h(x).$$

Applying summation by parts (Σ denoting indefinite summation) in the form (Milne Thomson (1933))

$$\Sigma u_x v_x = u_x \Sigma v_x - \Delta u_x \Sigma^2 v_{x+1} + \Delta^2 u_x \Sigma^3 v_{x+2} + \cdots ,$$

We derive, for $s < r$, the expression

$$(1.72) \qquad \Big[x^{(s)} \Delta^{r-1} x^{(r)} (Nq - n + x)^{(r)} h(x) - s x^{(s-1)} \Delta^{r-2} (x+1)^{(r)} (Nq - n + x + 1)^{(r)} h(x+1) + \cdots + (-1)^s s! \Delta^{r-s-1} (x+s)^{(r)} (Nq - n + x + s)^{(r)} h(x+s) \Big]_0^{n+1} .$$

Now $h(x)$ vanishes for integer values of $x > n$, and so the product

$$x^{(r)} (N_q - n + x)^{(r)} h(x)$$

and all its differences vanish also for these values. Hence at the upper limit all terms in (1.72) vanish. At the lower limit all terms except the last vanish through having x as factor. But

$$(x+s)^{(r)} (Nq - n + x + s)^{(r)} h(x+s) = 0, \quad x = 0, 1, \ldots r-s-1,$$

and so when $x = 0$

$$\Delta^{r-s-1} (x+s)^{(r)} (Nq - n + x + s)^{(r)} h(x+s) = 0.$$

Hence all terms vanish at both limits, and (1.71) becomes zero for $s < r$.

Again, when $r = s$, summation by parts yields terms which vanish as in (1.72), except for the last term, which takes the form

$$(-1)^r \; r!\sum_{0}^{n} (x+r)^{(r)} \; (Np-n+x+r)^{(r)} h(x+r),$$

and this reduces without difficulty to

$$\frac{r!\, n^{(r)} (Np)^{(r)} (N-n)^{(r)} (Nq)^{(r)}}{N^{(2r)}} \sum h(x;n-r,N-2r,Np-r)$$

(1.73) $$= r!\, n^{(r)} \; (Np)^{(r)} \; (N-n)^{(r)} (Nq)^{(r)} \; / \; N^{(2r)}$$

The orthogonal properties may be therefore expressed as

(1.74) $$\sum_{0}^{n} U_r(x) U_s(x) h(x) = \delta_{rs} \frac{r!\, n^{(r)} (Np)^{(r)} (N-n)^{(r)} (Nq)^{(r)}}{N^{(2r)} \; (N-r+1)^{(r)}} \; .$$

Factorial Moments in the Hypergeometric case.

The f.m.g.f. of $U_r(x)$ is

$$F(\alpha) = \frac{n^{(r)} (Np)^{(r)} (N-n)^{(r)} (Nq)^{(r)}}{N^{(2r)} \; (N-r+1)^{(r)}} \alpha^r \, F(-n+r,-Np+r, -N+2r;-\alpha).$$

Proof: We use Campell's lemma: if $F(\alpha)$ is the f.m.g.f. of a function $f(x)$, then $\alpha^r F(\alpha)$ is the f.m.g.f. of $(-1)^r \Delta^r f(x-r)$. Applying this to the present case, we have

$$F(\alpha) = \sum_{0}^{n} (1+\alpha)^x \; U_r(x) \; h(x)$$

$$= \sum (1+\alpha)^x \;\; \Delta^r \, x^{(r)} (Nq-n+x)^{(r)} h(x)/(N-r+1)^{(r)}$$

$$= \sum (1+\alpha)^x \; (-1)^r (x+r)^{(r)} (Nq-n+x+r)^{(r)} h(x)/(N-r+1)^{(r)}$$

(1.75) $$= \frac{n^{(r)} (Np)^{(r)} (N-n)^{(r)} (Nq)^{(r)} \alpha^r}{N^{(2r)} \; (N-r+1)^{(r)}}$$

$$\times F(-n+r,-Np+r,-N+2r; \; -\alpha).$$

The case $r = 0$ gives the f.m.g.f. of $h(x)$ which will be $F(-n,-Np,-N;-\alpha)$, which is the coefficient of z^n in the expansion of

$$(1.76) \qquad [1+(1+\alpha)z]^{Np}(1+z)^{Nq}/\binom{N}{n}$$

Proof for (1.76): The f.m.g.f. of $h(x)$ is

$$F(\alpha) = \sum_{x=0}^{n}(1+\alpha)^x \binom{Np}{x}\binom{Nq}{n-x} \Big/ \binom{N}{n}.$$

It follows

$$\binom{N}{n} F(\alpha) = \sum_{x=0}^{n}(1+\alpha)^x \binom{Np}{x}\binom{Nq}{n-x}$$

Hence

$$\sum_{n=0}^{\infty}\binom{N}{n}F(\alpha)z^n = \sum_{n=0}^{\infty}\sum_{x=0}^{n}(1+\alpha)^x\binom{Np}{x}\binom{Nq}{n-x}z^n$$

$$= \sum_{x=0}^{\infty}\sum_{n=x}^{\infty}\binom{Nq}{n-x}z^{n-x}\binom{Np}{x}[(1+\alpha)z]^x$$

$$= \sum_{x=0}^{\infty}(1+z)^{Nq}\binom{Np}{x}[(1+\alpha)z]^x$$

$$= (1+z)^{Nq}[1+(1+\alpha)z]^{Np}.$$

It follows

$F(\alpha)$ is the coefficient of z^n in

$$(1+z)^{Nq}[1+(1+\alpha)z]^{Np} \Big/ \binom{N}{n}.$$

9. Bessel's functions.

Definition 1.15. The Bessel function $I_n(x)$ is defined as

$$(1.77) \qquad I_n(x) = \frac{x^n}{2^n\, n!}\left[1 - \frac{x^2}{2(2n+2)} + \frac{x^4}{2.4(2n+2)(2n+4)} + \ldots\ldots\right]$$

Alternatively

(1.78) $$I_n(x) = \sum_{i=0}^{\infty} \frac{(-1)^i \; (x/2)^{2i+n}}{i! \; \Gamma(i+n+1)}$$

Lemma 1.5. $I_{-n}(x) = (-1)^n I_n(x)$

Proof: $$I_{-n}(x) = \sum_{i=0} \frac{(-1)^i \; (x/2)^{2i-n}}{i! \; \Gamma(i-n+1)}$$

$$= \sum_{i=0}^{n-1} \frac{(-1)^i (x/2)^{2i-n}}{i! \, \Gamma(i-n+1)} + \sum_{i=n} \frac{(-1)^i (x/2)^{2i-n}}{i! \; \Gamma(i-n+1)} \quad .$$

But the first term is equal to zero, it follows that

$$I_{-n}(x) = \sum_{i=n}^{\infty} \frac{(-1)^i \; (x/2)^{2i-n}}{i! \; \Gamma(i-n+1)}$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^{k+n} (x/2)^{2k+n}}{(k+n)! \; \Gamma(k+1)}$$

$$= (-1)^n \sum_{k=0}^{\infty} \frac{(-1)^k \; (x/2)^{2k+n}}{(k+n)! \; \Gamma(k+1)}$$

Therefore $$I_{-n}(x) = (-1)^n \sum_{i=0}^{\infty} \frac{(-1)^i \; (x/2)^{2i+n}}{i! \; \Gamma(i+n-1)}$$

(1.79) $$= (-1)^n \; I_n(x).$$

$$I_{\frac{1}{2}}(x) = \sum_{i=0} \frac{(-1)^i \; (x/2)^{2i+\frac{1}{2}}}{i! \; \Gamma(\frac{2}{2}+i)}$$

$$= \sum_{i=0} \frac{(-1)^i \; x^{2i+1}}{i! \; \sqrt{x} \; 2^{2i+\frac{1}{2}}(i+\frac{1}{2})(i-\frac{1}{2})}$$

$$= \sqrt{\frac{2}{x\pi}} \sum_{i=0}^{\infty} \frac{(-1)^i \; x^{2i+1}}{(2i+1)!}$$

(1.80) $$= \sqrt{\frac{2}{\pi x}} \quad \sin x.$$

Similarly we can show

$$I_{-\frac{1}{2}}(x) = \left(\frac{2}{\pi x}\right)^{\frac{1}{2}} \cos x \tag{1.81}$$

Jackson (1941) showed that the set $\{x\, I_n(\lambda x)\}$ is orthogonal on the interval (0,1). Hence, the set $\{I_n(\lambda x)\}$ is orthogonal with respect to x as a weight function.

$$\int_0^1 x\, I_n(\lambda x)\, I_n(ux)\, dx = o, \lambda \neq u \text{ and}$$

$$\int_0^1 x \left[I_n(\lambda x)\right]^2 dx = \text{ a constant.} \tag{1.82}$$

CHAPTER II

THE BIVARIATE NORMAL DISTRIBUTION

1. Introduction. Let x,y be two random variables jointly normally distributed with zero means, and unit variances, and a coefficient of correlation ρ.

The univariate normal distribution has an exponential function $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ with a quadratic exponent that is never positive. To obtain a bivariate extension, it seems natural to make the exponent quadratic in two variables the result is the bivariate normal distribution $f(x,y,\rho)$

$$(2.1) \qquad = \frac{1}{(2\pi)(1-\rho^2)^{\frac{1}{2}}} \exp\{-\tfrac{1}{2}(x^2+y^2-2\rho xy)/1-\rho^2\}.$$

In general, (see Anderson), the density function for a multivariate normal distribution is

$$(2.2) \qquad \frac{1}{(2\pi)^{p/2}|\Sigma|^{\frac{1}{2}}} \exp\{-\tfrac{1}{2}(\underline{X}-\underline{U})'\Sigma^{-1}(\underline{X}-\underline{U}),$$

where $\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, $\underline{U} = \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$ and Σ is the covariance matrix.

A special case of (2.2) is the bivariate normal distribution when $\underline{u} = \binom{0}{0}$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ and (2.1) will follow by substitution.

Mehler (1866) derived a series expansion of the bivariate normal frequency function, known as the Mehler identity.

$$(2.3) \qquad f(x,y,\rho) = \phi(x)\phi(y) \sum_{i=0}^{\infty} \rho^i H_i(x)H_i(y)$$

where $\{H_n(x)\}$ and $\{H_n(y)\}$ are the sets of standardized Hermite Chebyshev polynomials in x and y and they are given by (1.11). In this chapter five proofs of the Mehler identity spread over the period 1900-1958, will be reproduced with a reference to a sixth proof.

2. First proof of the Mehler Identity. This proof was given by Pearson (1900). It depends on expanding $f(x,y,\rho)$ as a power series in ρ.

$$f(x,y,\rho) = \frac{1}{2\pi}(U_0' + U_1'\rho + \frac{U_2'}{2!}\rho^2 + \ldots) .$$

Let

$$U_n = \exp[\tfrac{1}{2}(x^2+y^2)]\ U_n' \ , \quad \text{it follows}$$

$$f(x,y,\rho) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} (U_0 + \frac{U_1}{1!}\rho + \frac{U_2}{2!}\rho^2 + \ldots) \tag{2.4}$$

where

$$U_n = \exp \tfrac{1}{2}(x^2+y^2) \ . \ (\partial^n f/\partial\rho^n)_{\rho=0}$$

Diffrentiating $f(x,y,\rho)$ logarithmically with respect to ρ, we get

$$\frac{\partial \log f(x,y,\rho)}{\partial\rho} = \frac{1}{f(x,y,\rho)} \frac{\partial f(x,y,\rho)}{\partial\rho}$$

$$= \frac{xy + (1-x^2+y^2)\rho + xy\rho^2 - \rho^3}{(1-\rho^2)^2} .$$

It follows that

$$(1-\rho^2)^2 \frac{\partial f}{\partial\rho} = [xy + \rho(1-x^2+y^2) + \rho^2 xy - \rho^3] f(x,y,\rho). \tag{2.5}$$

Differentiating (2.5) n times with respect to ρ and putting $\rho = 0$, we get

$$(2.6) \qquad U_{n+1} = n(2n-1-x^2-y^2)\; U_{n-1} - n(n-1)(n-2)U_{n-3} + xy\left[U_n + n(n-1)U_{n-2}\right]$$

so that

$$U_0 = 1$$

$$U_1 = xy$$

$$U_2 = (x^2-1)(y^2-1)$$

$$U_3 = x(x^2-3)y(y^2-3), \; \ldots \text{ etc.}$$

Generally, we write

$$U_n = V_n(x)\; V_n(y) \quad , \text{ where}$$

$$(2.7) \qquad V_n(x) = x^n - \frac{n(n-1)}{2.1\;!}\, x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2^2 \;.\; 2!}\, x^{n-4} + \ldots\ldots .$$

Substituting for the values of U_n in (2.4), we get

$$(2.8) \qquad f(x,y,\rho) = \frac{1}{2\pi}\, e^{-\frac{1}{2}(x^2+y^2)}\left[1 + \frac{\rho}{1!}\, V_1(x)V_1(y) + \frac{\rho^2}{2!} V_2(x)V_2(y) + \ldots\right]$$

Obviously,

$$(2.9) \qquad V_n(x) = \sqrt{n!}\;\; H_n(x),$$

it follows

$$(2.10) \qquad H_n(x) = V_n(x) \,/\, \sqrt{n!}$$

so (2.8) becomes

$$f(x,y,\rho) = \phi(x)\phi(y)\left[1 + \rho\, H_1(x)H_1(y) + \rho^2 H_2(x)H_2(y) + \ldots\right]$$

$$= \phi(x)\phi(y) \sum_{i=0}^{\infty} \rho^i\, H_i(x)\; H_i(y),$$

which is the Mehler identity.

3. Second Proof of the Mehler Identity.

Watson (1933) gave two proofs of the Mehler identity. The first proof is due to Hille (1926); it uses the relations (1.24) and (1.25) between the Hermite Chebyshev and the Laguerre polynomials and it used the following result. If

$$(2.11) \qquad k(x,y,t) = \sum_{n=0}^{\infty} \frac{t^n\, n!\, e^{-\frac{1}{2}(x+y)}}{\Gamma(n+\alpha+1)} (xy)^{\frac{1}{2}\alpha} L_n^{(\alpha)}(x) L_n^{(\alpha)}(y)$$

then $k(x,y,t)$ can be written in terms of the Bessel functions of imaginary argument (definition 1-15) in the form

$$(2.12) \qquad k(x,y,t) = \frac{t^{-\frac{1}{2}\alpha}}{1-t} \exp\left\{-\tfrac{1}{2}(x+y)\frac{1+t}{1-t}\right\} I_\alpha\left(\frac{2\sqrt{xyt}}{1-t}\right)$$

Hardy obtained this result but it had been discovered earlier by Wigert (when $\alpha = 0$) (1921) and by Hille (1926) for general values of α. But Watson (1933) gives a direct proof of (2.12) merely by expanding (2.12) as a series of powers of t.

The second proof is due to Watson himself, it involves only (i) rearrangement of absolutely convergent multiple series, (ii) the use of the formula of Saalschutz for generalized hypergeometric functions

$$(2.13) \qquad \frac{1}{\Gamma(d)\Gamma(e)}\ {}_3F_2\binom{a,b,c}{d,e} = \frac{\Gamma(1-d+c)\,\Gamma(1-e+c)}{\Gamma(d-b)\,\Gamma(e-b)\,\Gamma(d-a)\,\Gamma(e-a)},$$

where $d + e = a + b + c + 1$ and one of a,b is a negative integer.

Omitting the trivial factor $\frac{1}{\sqrt{\pi}} \exp\left[-\frac{1}{2}(x^2+y^2)\right]$ from the Mehler identity, we see that we have to prove when $|\rho| < 1$,

$$\frac{1}{\sqrt{1-\rho^2}} \exp\{-\tfrac{1}{2}[-2xy\rho + (x^2+y^2)\rho^2]/(1-\rho^2)\}$$

(2.14)

$$= \sum_{n=0}^{\infty} \rho^n H_n(x)H_n(y) = \sum_{n=0}^{\infty} \sqrt{n!}\,\rho^n \Big[\sum_r \frac{(-1)^r x^{n-2r}}{r!(n-2r)!\,2^r}\Big]\Big[\sqrt{n!}\sum_s \frac{(-1)^s y^{n-2s}}{s!(n-2s)!\,2^s}$$

where the summations with respect to r and s extend over each integral values as do not give rise to negative factorials in the denominators.

Now, when $|\rho| \leq t < 1$ and $|x| \leq A < \infty$, $|y| \leq B < \infty$, the expansion of

$$\frac{1}{\sqrt{1-\rho^2}} \exp\{-\tfrac{1}{2}[-2xy\rho + (x^2+y^2)\rho^2]/(1-\rho^2)\}$$

which is obtained by writing it in the form

$$(2.15) \qquad \sum_{N=0}^{\infty} (-\tfrac{1}{2})^N \frac{(-2xy\rho + x^2\rho^2 + y^2\rho^2)^N}{N!\,(1-\rho^2)^{N+\frac{1}{2}}}$$

and then expanding the numerator by the multinomial theorem and the expressions $(1-\rho^2)^{-N-\frac{1}{2}}$ by the binomial theorem, is dominated by the corresponding convergent expansion of $\frac{1}{1-t^2} \exp\Big[\frac{2ABt + (A^2+B^2)t^2}{-2(1-t^2)}\Big]$, so it is permissible to expand the function

$$(2.16) \qquad \frac{1}{\sqrt{1-\rho^2}} \exp\Big\{\frac{2xy\rho - (x^2+y^2)\rho^2}{-2(1-\rho^2)}\Big\}$$

into the quadruple series just described and then to arrange the terms of the quadruple series in any convenient manner. Writing $\alpha(\alpha+1)$ $\ldots (\alpha+m+1) = (\alpha)^{(m)}$ for brevity, we find that (2.16) will be equal to

$$(2.17)\qquad \sum_{N=0}^{\infty}\sum_{M=0}^{N}\sum_{m=0}^{\infty}(-1)^{N-M}\frac{(2xy\rho)^{N-M}(x^2\rho^2+y^2\rho^2)^M(N+\frac{1}{2})^{(m)}\rho^{2m}}{2^N\,M!\,(N-M)!\,m!}$$

$$=\sum_N\sum_M\sum_m\sum_{R=0}^{M}\frac{(-1)^N(N+\frac{1}{2})^{(m)}\rho^{N+M+2m}x^{N+M-2R}y^{N-M+2R}}{2^M\,(N-M)!\,(M-R)!\,R!\,m!}$$

Write,

$M+N+2m = n$	$M = r+s-2m$
$R+m = r$	$N = n-r-s$
$M-R+m = s$	$R = r-m$

The summations with respect to m,n,r,s then range over all such integral values as do not give rise to negative factorials in the denominators, and we have,

$$\frac{1}{\sqrt{1-\rho^2}}\ \exp\left\{-\tfrac{1}{2}\left[\frac{-2xy\rho+(x^2+y^2)\rho^2}{1-\rho^2}\right]\right\}$$

$$=\sum_{n=0}^{\infty}\sum_r\sum_s\sum_m\frac{(-1)^{r+s}(n-r-s+\frac{1}{2})^{(m)}\,2^{2m-r-s}\rho^n x^{n-2r}y^{n-2s}}{(n+2m-2r-2s)!\,(s-m)!\,(r-m)!\,m!}$$

$$(2.18)\qquad =\sum_{n=0}^{\infty}\sum_r\sum_s\frac{(-1)^{r+s}2^{-r-s}\rho^n x^{n-2r}y^{n-2s}}{(n-2r-2s)!\,r!\,s!}\ {}_3F_2\left({-r,-s,n-r-s+\frac{1}{2}\atop \frac{1}{2}n+\frac{1}{2}-r-s,\frac{1}{2}n+1-r-s}\right).$$

Now, so long as n is not a negative integer, it follows from the formula of Saalschütz that

$$(2.19)\qquad \frac{1}{\Gamma(n-2r-2s+1)}\ {}_3F_2\left({-r,-s,n-r-s+\frac{1}{2}\atop \frac{1}{2}n+\frac{1}{2}-r-s,\frac{1}{2}n+1-r-s}\right)=\frac{\Gamma(n+1)}{\Gamma(n-2r)\ \Gamma(n-2s)}.$$

Then when n tends to an integral value, (2.19) becomes

$$=\frac{n!}{(n-2r)!\,(n-2s)!}$$

So long as $|\rho| < 1$. (2.18) becomes equal to

$$\sum_{n=0}^{\infty} \sum_{r} \sum_{s} \rho^n \frac{n!\,(-1)^{r+s} 2^{-r-s} x^{n-2r} y^{n-2s} n!}{r!\, s!\, (n-2r)!\, (n-2s)!} \quad \text{as required.}$$

4. Third Proof of the Mehler Identity.

Watson (1933) gave a proof due to Hardy(1932) who gave his proof in lectures on Orthogonal polynomials. This proof involves a use of absolutely convergent infinite integrals. We have

$$\int_{-\infty}^{+\infty} \exp(itx - \tfrac{1}{2}t^2)\,dt = \sqrt{2\pi}\exp(-\tfrac{1}{2}x^2),$$

so that

$$\exp(-\tfrac{1}{2}x^2) = \int_{-\infty}^{\infty} \exp(ixt - \tfrac{1}{2}t^2)\,dt \,/\, \sqrt{2\pi} \tag{2.20}$$

Differentiating (2.20) n times with respect to x using definition (1.4) of Hermite polynomials, we get

$$H_n(x) = (-i)^n \exp(+\tfrac{1}{2}x^2) \int_{-\infty}^{\infty} t^n \exp(ixt - \tfrac{1}{2}t^2)\,dt \,/\, \sqrt{2\pi\, n!}$$

and hence

$$\exp\left[-\tfrac{1}{2}(x^2+y^2)\right] \sum_{n=0}^{\infty} \rho^n H_n(x) H_n(y) \tag{2.21}$$
$$= \sum_{n=0}^{\infty} (-\rho)^n \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (tu)^n \frac{\exp(ixt - \tfrac{1}{2}t^2 + iyu - \tfrac{1}{2}u^2)}{2\pi\, n!}\,dt\,du.$$

The convergence of

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\{ \sum_{n=0}^{\infty} \frac{(-\rho t u)^n}{n!} \right\} \exp(At + Bu - \tfrac{1}{2}t^2 - \tfrac{1}{2}u^2)\,dt\,du$$

for $|\rho| < 1$, (A and B are constants), shows that rearrangement of the order of the summation and integration in (2.21) is permissible.

Hence we get

$$\exp\left[-\tfrac{1}{2}(x^2+y^2)\right] \sum_{n=0}^{\infty} \rho^n H_n(x)\, H_n(y)$$

(2.22)
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(ixt - \tfrac{1}{2}t^2 + iyu - \tfrac{1}{2}u^2 - \rho tu)\,dt\,du$$

we integrate first with respect to t after completing the square in the exponent:

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left[\tfrac{1}{2}(ixu - \rho u)^2 - \tfrac{1}{2}u^2 + iyu\right] du$$

$$= \frac{\exp(-\frac{1}{2}x^2)}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left\{-\tfrac{1}{2}\left[u^2(1-\rho^2) - 2u(iy - ix\rho)\right]\right\} du.$$

Next, we complete the square in the exponent, and integrate with respect to u:

$$= \frac{1}{\sqrt{1-\rho^2}} \exp\left[(-\tfrac{1}{2}x^2 - \tfrac{1}{2}y^2 + xy\rho - \tfrac{1}{2}x^2\rho^2)/(1-\rho^2)\right]$$

$$= 2\pi\, f(x,y,\rho).$$

It follows

$$f(x,y,\rho) = \frac{e^{-\frac{1}{2}(x^2+y^2)}}{2\pi} \sum_{n=0}^{\infty} \rho^n H_n(x) H_n(y),$$

as required.

5. Fourth Proof of the Mehler Identity.

This proof was given in A.C. Aitken's lectures for a number of years. Later Kendall (1941) presented the same proof unaware of Aitken's proof. The proof depends on the 1-1 correspondance between distribution functions and characteristic functions.

The characteristic function corresponding to $f(x,y,\rho)$ is

$$(2.23) \qquad \Psi(t,u) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(itx + iuy)\ f(x,y,\rho)\,dx\,dy$$

$$= \exp\left\{-\tfrac{1}{2}(t^2 + 2ut\rho + u^2)\right\}$$

$$(2.24) \qquad = \exp\left(-\tfrac{1}{2}t^2\right) \exp\left(-\tfrac{1}{2}u^2\right) \sum_{n=0}^{\infty} (-\rho)^n\, t^n\, u^n/n!$$

so that $\Psi(t,u)$ is an infinite sum of functions of the form

$$(2.25) \qquad \frac{\exp(-\tfrac{1}{2}t^2)(it)^n}{\sqrt{n!}} \cdot \frac{\exp(-\tfrac{1}{2}u^2)(iu)^n}{\sqrt{n!}} .$$

The idea of the proof is to show that the function $\frac{\exp(-\frac{1}{2}t^2)(it)^n}{\sqrt{n!}}$ is itself a characteristic function. Now, if $X(t)$ is the characteristic function of $\phi(x)\ H_n(x)$ then

$$(2.26) \qquad X(t) = \int_{-\infty}^{+\infty} \exp(itx)\ \phi(x)\ H_n(x)\,dx.$$

Using the generating function (1.14) of the Hermite polynomials, it follows

$$X(t) = \text{coefficient of } \frac{s^n}{\sqrt{n!}} \text{ in } \int_{-\infty}^{\infty} \exp(itx)\phi(x)\exp(xs-\tfrac{1}{2}s^2)\,dx.$$

The integral is equal to

$$\int_{-\infty}^{\infty} \exp\left\{(it+s)x\right\}\phi(x)\ \exp(-\tfrac{1}{2}s^2)\,dx$$

$$= \exp(-\tfrac{1}{2}s^2)\ \exp\left\{\tfrac{1}{2}(it+s)^2\right\}$$

$$(2.27) \qquad = \exp(-\tfrac{1}{2}t^2 + its).$$

It follows

$X(t)$ is the coefficient of $\frac{s^n}{\sqrt{n!}}$ in (2.27). But (2.27) is equal to $\sum_{n=0}^{\infty} \frac{(sti)^n}{n!}\ e^{-\frac{1}{2}t^2}$, it follows $X(t) = \frac{\exp(-\frac{1}{2}t^2)(it)^n}{\sqrt{n!}}$.

Hence the product (2.25) is the characteristic function of $\phi(x)\phi(y)H_n(x)H_n(y)$; and hence $\Psi(t,u)$ is the characteristic function of

$$\phi(x)\ \phi(y)\ \sum_{n=0}^{\infty} \rho^n H_n(x)\ H_n(y),$$

but $\Psi(t,u)$ is the characteristic function of $f(x,y,\rho)$.

Therefore;

$$f(x,y,\rho) = \phi(x)\phi(y) \sum_{n=0}^{\infty} \rho^n H_n(x)H_n(y)$$

as required.

6. Fifth Proof of the Mehler Identity.

This proof was given by Lancaster (1958). Pearson (1904) introduced ϕ^2 as the "mean square contingency" of a bivariate distribution in order to derive a measure of association independent of the sample size, N. He wrote $\phi^2 = \frac{X^2}{N}$. Pearson saw that X^2 (or rather ϕ^2) has a use as a descriptive measure, where it was usually thought of as a criterion of goodness of fit.

For a bivariate distribution with a distribution function $F(x,y)$ and marginal distribution functions, $G(x)$ and $H(y)$, Lancaster (1958) defines:

Definition 2.1.

$$\phi^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [dF(x,y)]^2 \ / \ [dG(x)dH(y)] - 1 \qquad (2.28)$$

$$= \iint \Omega^2(x,y)\ dG(x)dH(y) - 1, \quad \text{where}$$

$$\Omega(x,y) = dF(x,y)\ /\ [dG(x)\, dH(y)]$$

is to be taken as zero if the point (x,y) does not correspond to a point of increase of both $G(x)$ and $H(y)$. ϕ^2 can be regarded as $\sum_{ij} f_{ij}^2 / (f_{i.} \cdot f_{.j}) - 1$ where f_{ij} is the weight of the bivariate distribution corresponding to marginal sets, A_i and B_j, and where $f_{i.}$ and $f_{.j}$ are the weights of the marginal distributions corresponding to the same sets.

In the case of the bivariate joint normal distribution. We may write $g(x)dx$ and $h(y)dy$ in place of $dG(x)$ and $dH(y)$ respectively and $f(x,y)dx\,dy$ in place of $dF(x,y)$ i.e.

$$\phi^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^2(x,y)/[g(x)h(x)]\ dx\ dy - 1.$$

By completion of square and integration with respect to x first and then with respect to y, we get

$$(2.29) \qquad \phi^2 = \frac{\rho^2}{1-\rho^2} \quad \text{where } /\rho/ < 1.$$

If $/\rho/ = 1$, the bivariate normal distribution is singular and ϕ^2 is unbounded. Indeed ϕ^2 is unbounded for any bivariate distribution along a straight line, with infinitely many points of increase.

<u>Definition 2.2.</u> Let $\{x^{(i)}\}$ and $\{y^{(i)}\}$ be complete sets of orthonormal functions defined on $G(x)$ and $H(y)$ respectively, i.e.

$$(2.30) \qquad \int x^{(i)}x^{(j)}\ dG(x) = \int y^{(i)}y^{(j)}\ dH(y) = \delta_{ij}$$

and let ρ_{ij} be the coefficient of correlation of $x^{(i)}$ and $y^{(i)}$, i.e.

$$(2.31) \qquad \rho_{ij} = \text{Corr.}\ (x^{(i)}, y^{(j)}) = \int x^{(i)}y^{(j)}\ dF(x,y)$$

so that $\rho_{00} = 1, \rho_{i0} = \rho_{0i} = 0$ for $i \neq 0$.

Theorem 2.1. If $F(x,y)$ is a ϕ^2 - bounded distribution function and if $S_{mn} = S_{mn}(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{n} \lambda_{ij}\, x^{(i)} y^{(j)}$,

then

$$(2.32) \qquad Q_{mn} = \iint (\Omega - S_{mn})^2 \, dG(x)\, dH(y)$$

is minimized by taking $\lambda_{ij} = \rho_{ij}$ for $i = 0,1,\ldots,m$ and $j = 0,1,\ldots n$.

Taking the limit as $m \to \infty$ and $n \to \infty$, we get

$$\Omega(x,y) = S(x,y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \lambda_{ij}\, x^{(i)} y^{(j)}$$

$$(2.33) \qquad = 1 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho_{ij}\, x^{(i)} y^{(j)}$$

almost everywhere.

Squaring (2.33) and multiplying both sides by $dG(x)\ dH(y)$ and then integrating with respect to x and y, we get

$$(2.34) \qquad \phi^2 = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho_{ij}^2$$

which is called Parseval equality.

Lancaster gave a definition of canonical variables which is an extension of Fisher's (1940) definition.

Definition 2.3. The canonical variables (or functions) are two sets of orthonormal functions defined on the marginal distributions in a recursive manner such that the correlation between corresponding members of the two sets is maximal. i.e. $\{x^{*(i)}\}$ and $\{y^{*(i)}\}$ are called the canonical variables if

$$
\begin{aligned}
&\int x^{*(i)}\, dG(x) = \int y^{*(i)}\, dH(y) = 0 \quad , \; i = 1,2,\ldots \\
(2.35) \qquad &\int (x^{*(i)})^2\, dG(x) = \int (y^{*(i)})^2\, dH(y) = 1 \,, i = 1,2,\ldots \\
&\int x^{*(i)}\, x^{*(j)}\, dG(x) = \int y^{*(i)} y^{*(j)}\, dH(y) = 0
\end{aligned}
$$

for $i \neq j$ and $\rho_i = \text{Corr}\,(x^{*(i)}, y^{*(i)}) = \iint x^{*(i)} y^{*(i)}\, dF(x,y)$ is maximal for each i.

The ρ_i are called the canonical correlation of the canonical variables.

Theorem. 2.2. The canonical variables obey a second set of orthogonal conditions.

$$
(2.36) \qquad \text{corr.}(x^{*(i)}, y^{*(j)}) = \iint x^{*(i)}\, y^{*(j)}\, dF(x,y) = 0 \;, \; i \neq j
$$

Proof. Let $j > i$, by definition 2.3, $\text{corr.}(x^{*(i)}, y^{*(i)}) = \rho_i$ and is maximal. Suppose $\text{Corr.}(x^{*(i)}, y^{*(j)}) = \rho_i \tan\theta \neq 0$. $y^{*(j)}$ has been defined according to 2.35 and so the function $\text{Cos}\theta\; y^{*(i)} + \text{Sin}\theta\; y^{*(j)}$, obeys all the necessary orthogonal and normalizing conditions, and its correlation with $x^{*(i)}$ is equal to

$$
\begin{aligned}
&\iint x^{*(i)} (y^{*(i)} \cos\theta + y^{*(j)} \sin\theta)\, dF(x,y) \\
&= \rho_i \cos\theta + \rho_i \frac{\sin\theta}{\cos\theta} \cdot \sin\theta \\
&= \rho_i \frac{\cos^2\theta + \sin^2\theta}{\cos\theta} \\
&= \frac{\rho_i}{\cos\theta} = \rho_i \sec\theta > \rho_i \quad \text{contradiction,}
\end{aligned}
$$

because ρ_i is maximal.

Now, in terms of the canonical variables, (2.33) becomes

$$(2.37) \qquad dF(x,y) = \left[1 + \sum_{i=1}^{\infty} \rho_i \, x^{*(i)} y^{*(i)}\right] dG(x)\, dH(y)$$

where

$$(2.38) \qquad \phi^2 = \sum_{i=1}^{\infty} \rho_i^2 \,.$$

The result expressed by (2.37) is a generalization of the work by Fisher(1940) and later by Maung (1941) and Williams(1952), where G(x) and H(y) are restricted to have finitely many points of increase.

The converse of the result (2.37) is also true, i.e. if a bivariate distribution can be written in the form (2.37) where $\{x^{*(i)}\}$ and $\{y^{*(i)}\}$ are complete sets of orthonormal functions defined on the marginal distributions and $\sum_i \rho_i^2$ is finite, then the ρ_i are the canonincal correlations, $x^{*(i)}$ and $y^{*(i)}$ are the canonical variables and $\sum_i \rho_i^2 = \phi^2$. (See Lancaster (1958))

In the case of the bivariate normal distribution with coefficient of correlation ρ, we have shown $\phi^2 = \dfrac{\rho^2}{1-\rho^2}$ so that F(x,y) is ϕ^2 - bounded for $/\rho/<1$. The canonical variables in this case are the standardized Hermite-Chebyshev polynomials (Lancaster (1957)), defined by 1.4. The Canonical correlations are ρ^i, since

$$\phi^2 = \sum \rho_i^2 = \frac{\rho^2}{1-\rho^2}$$

it follows

$$\sum \rho_i^2 = \sum (\rho^i)^2, \quad \rho_i = \rho^i$$

By (2.37), it follows that

$$f(x,y,\rho) = \left[1 + \sum_{i=1}^{\infty} \rho^i \, H_i(x)H_i(y)\right]\phi(x)\phi(y)$$

$$= \phi(x)\phi(y) \sum_{i=0}^{\infty} \rho^i \, H_i(x)H_i(y)$$

as required.

CHAPTER III

THE BIVARIATE GAMMA DISTRIBUTION

1. Introduction. Kibble(1941) showed that a two-variate distribution function in which each of the variates, x,y has the frequency function

$$(3.1) \qquad g(x) = \frac{x^{p-1} e^{-x}}{\Gamma(p)} \qquad 0 \leq x < \infty$$

may be represented by

$$(3.2) \qquad h(x,y,\rho) = g(x)g(y)\left[1 + \sum_{r=1}^{\infty} \frac{r!\,\rho^{2r} \cdot}{\Gamma(p)\,\Gamma(p+r)} L_r^{(p-1)}(x)\,L_r^{(p-1)}(y)\right]$$

where $\left\{L_r^{(p-1)}(x)\right\}$ is the set of Laguerre polynomials defined by 1.5, and ρ is the coefficient of correlation. Krishnamoorthy and Parthasarathy (1951) generalized Kibble's work to the multivariate gamma distribution. Other derivations of the bivariate gamma distribution (not in canonical form) are due to Wicksell (1933), who applies Fourier's inversion theorem to derive an integral form of the distribution; and Cherian (1941), who uses the additive property of gamma variables. As we are interested in the canonical forms of bivariate distributions, Kibble's derivation of (3.2) is given in this chapter and the analogy with Mehler's identity is pointed out.

2. The Bivariate Gamma Frequency Function. If X is a standardized normal variate, then $d\phi(X) = \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2}X^2}\, dX$. Now if we let $x = \frac{1}{2}X^2$, then

$$g(x) = x^{-\frac{1}{2}} e^{-x} / \Gamma(\tfrac{1}{2}),$$ hence x is a gamma variable with parameter $p = \frac{1}{2}$.

Generally, if $X_1, X_2, \ldots X_n$ are mutually independent standardized normal variates, then $\frac{1}{2}(X_1^2 + X_2^2 + \ldots + X_n^2)$ is a gamma variable with parameter $\frac{n}{2}$. So the bivariate distribution of the squares of two variables normally correlated would lead to a bivariate gamma distribution.

Let (X,Y) be a bivariate normal variable with probability element

$$(3.3) \qquad f(X,Y,\rho)dX\,dY = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left[-\tfrac{1}{2}(X^2-2\rho XY+Y^2)/(1-\rho^2)\right] dXdY;$$

and make the transformation

$$x = \tfrac{1}{2} X^2, \qquad y = \tfrac{1}{2} Y^2$$

so that

$$dX\,dY = dx\,dy \,/\, 2(xy)^{\frac{1}{2}}.$$

Noting that there are four pairs of values of X and Y corresponding to one pair of values of x and y, two with positive and two with negative (XY), the joint probability element of x and y is

$$(3.4) \qquad g(x,y,\rho)dx\,dy = \frac{(xy)^{-\frac{1}{2}}}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left[-(x-2\rho\sqrt{xy}+y)/(1-\rho^2)\right]dxdy$$

$$+ \frac{(xy)^{-\frac{1}{2}}}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left[-(x+2\rho\sqrt{xy}+y)/(1-\rho^2)\right]dxdy$$

where $0 \leq x < \infty$ and $0 \leq y < \infty$ and the positive sign must be taken with each root.

Now, the joint moment generating function of x and y is

$$(3.5)\qquad G_0(t,u) = \int_0^{\infty}\int_0^{\infty} \exp(tx + uy)\; g(x,y,\rho)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[\tfrac{1}{2}(tX^2 + uY^2)\right] f(X,Y,\rho)\,dX\,dY$$

$$= \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{-\tfrac{1}{2}\left[X^2(1-t(1-\rho^2))\right.\right.$$

$$\left.\left.-2\rho XY\right]/(1-\rho^2)\right\} \cdot \exp\left\{\tfrac{1}{2}\left[Y^2u(1-\rho^2)-Y^2\right]/(1-\rho^2)\right\}\, dX\,dY.$$

By completing the square in the exponent then integrating with respect to X, we get

$$= \frac{1}{(2\pi)^{\frac{1}{2}}\sqrt{1-t(1-\rho^2)}} \int_{-\infty}^{\infty} \exp\left\{-\tfrac{1}{2}\left[\frac{Y^2(1-u-t+tu(1-\rho^2))}{1-t(1-\rho^2)}\right]\right\} dY$$

$$(3.6)\qquad G_0(t,u) = (1-u-t+tu-tu\rho^2)^{-\frac{1}{2}} = \left[(1-t)(1-u)-tu\rho^2\right]^{-\frac{1}{2}}.$$

Generally, if (X_1,Y_1), (X_2,Y_2), ... (X_n,Y_n) is a sample of n mutually independent observations from a bivariate normal population with frequency function $f(X,Y,\rho)$, then the joint moment generating function of $x = \frac{1}{2}\sum_i X_i^2$ and $y = \frac{1}{2}\sum_i Y_i^2$ is

$$(3.7)\qquad G(t,u) = \left[(1-t)(1-u) - tu\rho^2\right]^{-n/2}$$

where each of x and y is a gamma variable with parameter $p = \frac{n}{2}$.

3. The Canonical Form of the Bivariate Gamma Distribution.

Let G(t,u) be written in the following form

$$(3.8)\qquad G(t,u) = (1-t)^{-p}(1-u)^{-p}\left[1 - \frac{tu\rho^2}{(1-t)(1-u)}\right]^{-p}.$$

By the negative binomial expansion we get

$$(3.9) \qquad G(t,u) = (1-t)^{-p}(1-u)^{-p} \sum_{r=0}^{\infty} \binom{p+r-1}{r} \left[\frac{tu\rho^2}{(1-t)(1-u)}\right]^r$$

$$= (1-t)^{-p}(1-u)^{-p} \sum_{r=0}^{\infty} \frac{\Gamma(p+r)}{r!\,\Gamma(p)} \rho^{2r} \frac{t^r u^r}{(1-t)^r(1-u)^r}$$

$$= (1-t)^{-p}(1-u)^{-p}\left[1 + \sum_{r=1}^{\infty} \frac{\Gamma(p+r)}{r!\,\Gamma(p)} \left(\frac{u}{1-u}\right)^r \left(\frac{t}{1-t}\right)^r \rho^{2r}\right].$$

But by (1.23) $G_r(t) = \frac{\Gamma(p+r)}{\Gamma(p)r!}(1-t)^{-p}\left(\frac{t}{1-t}\right)^r$ is the moment generating function of $L_r^{(p-1)}(x)\ g(x)$ and $\frac{\Gamma(p+r)}{\Gamma(p)\ r!}(1-u)^{-p}\left(\frac{u}{1-u}\right)^r$ is the moment generating function of $L_r^{(p-1)}(y)\ g(y)$, and by one-to-one correspondence between moment generating functions and distribution functions, it follows that the bivariate gamma frequency function $h(x,y,\rho)$ has the series expansion

$$(3.10) \qquad h(x,y,\rho) = g(x)g(y)\left[1 + \sum_{r=1}^{\infty} \frac{r!\,\Gamma(p)}{\Gamma(p+r)} \rho^{2r} L_r^{(p-1)}(x) L_r^{(p-1)}(y)\right]$$

where $\quad 0 \leq x, \ y < \infty$.

Now let us derive the regression lines of x on y and y on x. The regression line of x on y can be found by finding $E(x/y)$ (Kendall, V.II p.285).

$$(3.11) \qquad E(x/y) = \int_0^{\infty} xh(x/y)\,dx = \int_0^{\infty} \frac{xh(x,y,\rho)}{g(y)}\,dx$$

$$= \int_0^{\infty} xg(x)\left[1 + \sum_{r=1}^{\infty} \frac{r!\,\Gamma(p)}{\Gamma(p+r)} \rho^{2r} L_r^{(p-1)}(x) L_r^{(p-1)}(y)\right]dx$$

$$= p + \sum_{r=1}^{\infty} \frac{r!\,\Gamma(p)}{\Gamma(p+r)} \rho^{2r} \int_0^{\infty} x\, L_r^{(p-1)}(x)\,dg(x)\, L_r^{(p-1)}(y).$$

The integral is always zero except when $r = 1$ because of the orthogonality of $L_r^{(p-1)}(x)$ (Theorem (1.3)). Hence (3.10) becomes

$$(3.11) \qquad = p + \frac{\Gamma(p)}{\Gamma(p+1)} \rho^2 \int_0^{\infty} x \, L_1^{(p-1)}(x)\, dg(x) \; L_1^{(p-1)}(y).$$

By (1.17) $L_1^{(p-1)}(x) = (p - x)$ and $L_1^{(p-1)}(y) = p - y$, it follows (3.11) becomes

$$(3.12) \qquad p + \frac{\rho^2}{p} (p-y) \int_0^{\infty} (xp - x^2) g(x)\, dx.$$

Becuase the first and second moments of the gamma distribution are p and p(p + 1) respectively, then (3.12) becomes

$$= p + \frac{\rho^2 (p-y)}{p} \left[p^2 - p(p+1) \right]$$

$$E(x/y) = p + \rho^2 (y-p).$$

So the regression of x on y is

$$(3.13) \qquad (x-p) = \rho^2 (y - p).$$

Similarly the regression of y on x, is

$$(3.14) \qquad (y - p) = \rho^2 (x - p)$$

(3.13) and (3.14) are straight lines passing through the double mean (p, p). So, the coefficient of correlation, R, is defined in the usual way as the geometric mean of the regression coefficients is given by (Kendall, V.II, p. 287).

$$(3.15) \qquad R = \rho^2.$$

Hence

$$(3.16) \qquad h(x,y,\rho) = g(x) g(y) \left[1 + \sum_{r=1}^{\infty} \frac{r!\, \Gamma(p)}{\Gamma(p+r)} R^r L_r^{(p-1)}(x) L_r^{(p-1)}(y) \right].$$

To compare (3.16) with the Mehler identity, we let the set $\left\{L_r^{(p-1)}(x)\right\}$ be standardized in the form

$$(3.17) \qquad \overset{*}{L}_r^{(p-1)}(x) = L_r^{(p-1)}(x) \Big/ \left[\frac{\Gamma(p+r)}{\Gamma(p)\ r!}\right]^{\frac{1}{2}}$$

with $\overset{*}{L}_r^{(p-1)}(y)$ defined similarly. It follows that (3.16) becomes

$$(3.18) \qquad h(x,y,\rho) = g(x)g(y)\left[1 + \sum_{i=1}^{\infty} R^i \ \overset{*}{L}_i^{(p-1)}(x) \ \overset{*}{L}_i^{(p-1)}(y)\right]$$

so that the canonical correlations are

$$(3.19) \qquad R^i = \text{corr.} \left[\overset{*}{L}_i^{(p-1)}(x), \ \overset{*}{L}_i^{(p-1)}(y)\right] .$$

4. Representation of the Bivariate Gamma Function in terms of Bessel's Function. Using (2.12) (Watson(1933))

$$(3.20) \qquad \sum_{n=0}^{\infty} \frac{\rho^{2n}\ n!\ e^{-\frac{1}{2}(x+y)}}{\Gamma(n+p)} (xy)^{-\frac{1}{2}(p-1)} L_n^{(p-1)}(x) L_n^{(p-1)}(y)$$

$$= \frac{\rho^{-(p-1)}}{1-\rho^2} \exp\left\{-\tfrac{1}{2}(x+y)\ \frac{1+\rho^2}{1-\rho^2}\right\} I_{(p-1)}\left(2\frac{\sqrt{\rho xy}}{1-\rho^2}\right),$$

it follows that

$$(3.21) \qquad h(x,y,\rho) = \sum_{n=0}^{\infty} \frac{(\rho^2)^n\ n!}{\Gamma(p)\ \Gamma(n+p)} \ L_n^{(p-1)}(x)\ L_n^{(p-1)}(y)$$

$$= \frac{(xy)^{\frac{1}{2}(p-1)}}{\Gamma(p)(1-\rho^2)\ \rho^{(p-1)}} \exp\left\{-\tfrac{1}{2}(x+y)\frac{1+\rho^2}{1-\rho^2}\right\}$$

$$\cdot I_{(p-1)} \left(\frac{2\sqrt{xy\rho}}{1-\rho^2}\right).$$

5. Extensions of the Bivariate Gamma Distribution.

Kibble (1941) (Hamdan (1963)) extends the above analysis to derive the canonical form of the bivariate distribution where one of the marginals is gamma and the other is normal. Moreover, he derives the canonical form of the bivariate gamma distribution with different marginal parameters.

Let (X,Y) be a bivariate normal variables with frequency function $f(X,Y,\rho)$. Making the transformation $x = \frac{1}{2}X^2$ and $y = Y$, we notice here that $(-X,Y)$ and (X,Y) give the same pair of values for x and y, we derive the joint frequency of x and y as

$$(3.22) \qquad k(x,y,\rho) = \frac{(2x)^{-\frac{1}{2}}}{2\pi(1-\rho^2)^{\frac{1}{2}}}\Big[\exp\left\{-\tfrac{1}{2}(2x - 2\rho y\sqrt{2x}+y^2)/(1-\rho^2)\right\}$$
$$+ \exp\left\{-\tfrac{1}{2}(2x+ 2\rho y\sqrt{2x}+y^2)/(1-\rho^2)\right\}\Big]$$

where $0 \leq x < \infty$ and $-\infty < y < \infty$, and the positive sign must be taken with each root. Obviously, the marginal distribution of x is gamma with parameter $\frac{1}{2}$, while the marginal distribution of y is normal $(0,1)$. It follows that the joint moment generating function of x and y is

$$(3.23) \qquad G_0(t,u) = \int_0^{\infty}\int_{-\infty}^{\infty}\exp(tx+uy)k(x,y,\rho)\,dxdy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(\frac{tX^2}{2} + uY)\; f(X,Y,\rho)\,dXdY.$$
$$= \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(\frac{tX^2}{2} - \frac{\frac{1}{2}X^2+\rho XY}{1-\rho^2}).$$
$$\cdot \exp(uY - \frac{Y^2}{2(1-\rho^2)})\,dXdY$$

Integrating with respect to X first, we get

$$\frac{1}{\sqrt{2\pi}\sqrt{1-t(1-\rho^2)}}\int_{-\infty}^{\infty}\exp\left\{-\tfrac{1}{2}\Big[Y^2\Big(\frac{1}{1-\rho^2} - \frac{\rho^2}{(1-\rho^2)[1-t(1-\rho^2)]}\Big)-2uY\Big]\right\}dY.$$

The exponent can be written as

$$\exp\left\{-\tfrac{1}{2}\left\{\frac{(1-t)Y^2}{[1-t(1-\rho^2)]} - 2uY + u^2\,\frac{[1-t(1-\rho^2)]}{(1-t)}\right\}\right\}$$

$$\cdot\ \exp\tfrac{1}{2}\left\{u^2\left[\frac{1-t(1-\rho^2)}{1-t}\right]\right\}.$$

Next integrating with respect to Y we get

$$(3.24)\qquad G_0(t,u) = \exp\left[\tfrac{1}{2}u^2\left(1+\frac{t\rho^2}{1-t}\right)\right]/(1-t)^{\frac{1}{2}}$$

In general, if (X_1,Y_1), $(X_2,Y_2),\ldots,(X_n,Y_n)$ is a sample of n mutually independent observations from a bivariate normal population with frequency function $f(X,Y,\rho)$ and if $x = \frac{1}{2}\sum_i X_i^2$ and $y = \sum_i Y_i/\sqrt{n}$, then the joint moment generating function of x and y is

$$(3.25)\qquad G(t,u) = (1-t)^{-p}\exp\left[\tfrac{1}{2}u^2\left(1+\frac{t\rho^2}{1-t}\right)\right]$$

where $p = \frac{n}{2}$, x is a gamma variate with parameter p and y is a standardized normal variable.

Let us expand $G(t,u)$ in the form

$$G(t,u) = (1-t)^{-p}\exp(\tfrac{1}{2}u^2)\left[1+\sum_{r=1}^{\infty}\left[\tfrac{1}{2}\frac{t\rho^2u^2}{(1-t)}\right]^r/r!\right]$$

$$(3.26)\qquad = (1-t)^{-p}\exp(\tfrac{1}{2}u^2)\left[1+\sum_{r=1}^{\infty}\rho^{2r}\left(\frac{t}{1-t}\right)^r\frac{u^{2r}}{2^r\,r!}\right].$$

Since $(1-t)^{-p}\left(\frac{t}{1-t}\right)^r$ is the moment generating function of $\frac{\Gamma(p)\,r!}{\Gamma(p+r)}L_r^{(p-1)}(x)g(x)$ (See(1.23)) and $u^{2r}\exp(\frac{1}{2}u^2)$ is the moment generating function of $\sqrt{(2r)!}\,H_{2r}(y)\phi(y)$, and by the uniqueness theorem, it follows that $k(x,y,\rho)$ has the canonical expansion

$$(3.27)\qquad k(x,y,\rho) = g(x)\phi(y)\left[1 + \sum_{r=1}^{\infty} \frac{\sqrt{(2r)!}}{2^r}\frac{\Gamma(p)}{\Gamma(p+r)}\rho^{2r} L_r^{(p-1)} H_{2r}(y)\right]$$

for $0 \leq x < \infty$ and $-\infty < y < \infty$.

Now, we derive the line of regression of x on y (Kendall),

$$(3.28)\qquad E(x/y) = \int_0^{\infty} \frac{x\,k(x,y,\rho)}{\phi(y)}\,dx$$

$$= \int_0^{\infty}\left\{x\,g(x)\left[1 + \sum_{r=1}^{\infty} \frac{\sqrt{(2r)!}\,\Gamma(p)}{2^r\,\Gamma(p+r)}\rho^{2r} L_r^{(p-1)}(x) H_{2r}(y)\right]dx\right\}$$

$$= p + \sum_{r=1}^{\infty} \frac{\sqrt{(2r)!}\,\Gamma(p)}{2^r\,\Gamma(p+r)}\rho^{2r} H_{2r}(y)\int_0^{\infty} x\,L_r^{(p-1)}(x)\,g(x)\,dx\,.$$

The integral is equal to zero for all r except when $r=1$, this is so because of the orthogonality of $L_r^{(p-1)}(x)$ and (3.28) becomes

$$(3.29)\qquad = p + \frac{\sqrt{2}}{2p}\rho^2\,H_2(y)\int_0^{\infty} x(x-p)\,g(x)\,dx. \quad \text{But}$$

$$(3.30)\qquad H_2(y) = \frac{1}{\sqrt{2}}(y^2-1)$$ (See (1.12)), it follows, by substituting (3.30) in (3.29) and integrating, that

$$E(x/y) = p + \frac{1}{2}\rho^2(y^2-1),$$

hence $x - p = \frac{1}{2}\rho^2(y^2 - 1)$ is the line of regression of x on y, which is a parabola. Similarly the regression of y on x is derived by

(3.31) $$E(y/x) = \int_{-\infty}^{\infty} yk(x,y,\rho)/g(x)dy$$

(3.32) $$= \sum_{r=0}^{\infty} \frac{(2r)!\,\Gamma(p)}{2r\,\Gamma(p+r)} \rho^{2r} L_r^{(p-1)}(x) \int_{-\infty}^{\infty} yH_{2r}(y)\phi(y)dy.$$

Because of the orthogonality of $H_{2r}(y)$ (Theorem 1.2), the integral is always zero i.e.

(3.33) $$\int_{-\infty}^{\infty} y\, H_{2r}(y)\phi(y)dy = 0$$

$$r = 0, 1, 2, \ldots$$

Hence (3.32) becomes

(3.34) $$E(y/x) = 0 \quad \text{i.e.}$$

It follows that

(3.35) $y = 0$ is the straight line of regression of y on x.

Hence the coefficient of correlation of x and y is zero. (This is an example of a zero correlation not implying independence).

Finally, Kibble (Hamdan (1963)) derives the moment generating function of a pair of gamma variables x and y with different parameters M and N respectively in the form

(3.36) $$(1-t)^{-M}\,(1-u)^{-N}\left[1 - \frac{\rho^2 tu}{(1-t)(1-u)}\right]^{-p}.$$

The derivation of (3.36) is similar to that in the first proof but here the parameters are different.

Hence the canonical form of the corresponding distribution has the form

(3.37) $$g(x,M)g(y,N)\left[1+\sum_{r=1}^{\infty} \frac{r!\,\Gamma(p+r)\;\Gamma(M)\;\Gamma(N)}{\Gamma(p)\;\Gamma(M+r)\;\Gamma(N+r)} \rho^{2r}.L_r^{(M-1)}(x)L_r^{(N-1)}(y)\right].$$

Proof: The expansion of (3.36) is

$$= (1-t)^{-M}\,(1-u)^{-N}\sum_{r=0}^{\infty}\frac{\rho^{2r}}{r!}\left(\frac{t}{1-t}\right)^r\left(\frac{u}{1-u}\right)^r.$$

But

$$\frac{\Gamma(M+r)}{\Gamma(M)\,r!}\,(1-t)^{-M}\,\left(\frac{t}{1-t}\right)^r$$

is the moment generating function of $L_r^{(M-1)}(x)g(x)$ (See (1.23)) and similarly $\frac{\Gamma(N+r)}{\Gamma(N)\,r!}\,(1-t)^{-N}\,\left(\frac{u}{1-u}\right)^r$ is the moment generating function of $L_r^{(N-1)}(y)g(y)$, and by the one-to-one correspondence between moment generating functions and distribution functions, it follows that the bivariate gamma frequency function $h^*(x,y,\rho)$ can be written in the form

$$h^*(x,y,\rho) = g(x,M)g(y,N)\Big[1+\sum_{r=1}\frac{r!\,\Gamma(p+r)\,\Gamma(M)\,\Gamma(N)}{\Gamma(p)\,\Gamma(M+r)\,\Gamma(N+r)}\rho^{2r} \cdot L_r^{(M-1)}(x)\,L_r^{(N-1)}(y)\Big] \tag{3.38}$$

where p is half a positive integer and $M,N > p$, so that (3.38) remains positive for every positive x and y.

The regression line of x on y is derived by finding $E(x/y)$ (Kendall V.II)

$$E(x/y) = \int_0^{\infty} x\,h^*(x,y,\rho)/g(y,N)\,dx$$

$$= M+\sum_{r=1}^{\infty}\frac{r!\,\Gamma(p+r)\,\Gamma(M)\,\Gamma(N)}{\Gamma(p)\,\Gamma(M+r)\,\Gamma(N+r)}\rho^{2r}L_r^{(N-1)}(y)\int_0^{\infty} xL_r^{(M-1)}(x)g(x)\,dx. \tag{3.39}$$

Because of the orthogonality of the set $\{L_r^{(M-1)}(x)\}$ (Theorem 1.3), the integral is zero except when $r = 1$, (3.39) becomes

$$= M+\frac{p}{MN}\rho^2\,L_1^{(N-1)}(y)\int_0^{\infty} x\,L_1^{(m-1)}(x)g(x)\,dx. \tag{3.40}$$

But

(3.41) $$L_1^{(N-1)}(y) = y - N \quad \text{and} \quad L_1^{(M-1)}(x) = x - M$$

(see (1.17)). Substituting (3.41) in (3.40), then integrate, we get

$$E(x/y) = M + \left[p\,\rho^2\,(y - N)\right]/N.$$

Hence (Kendall V.II)

(3.42) $$X - M = p\,\rho^2\,(y - N)/N$$ is the regression line of x on y.

Similarly, the line of regression of y on x is

(3.43) $$y - N = p\,\rho^2\,(x - M)\,/\,M$$

It follows that the coefficient of correlation is

(3.44) $$R = p\,\rho^2\,/\sqrt{MN}$$

In order to compare (3.38) with the Mehler identity, let the set $\left\{L_r^{(M-1)}(x)\right\}$ be standardized in the form

(3.45) $$L_r^{*(M-1)}(x) = \sqrt{r!}\left(\frac{M}{p}\right)^{\frac{r}{2}}\left[\frac{\Gamma(p+r)}{\Gamma(p)}\right]^{\frac{1}{2}}\frac{\Gamma(M)}{\Gamma(M+r)}\,L_r^{(M-1)}(x).$$

and

(3.46) $$L_r^{*(N-1)}(y) = \sqrt{r!}\left(\frac{N}{p}\right)^{\frac{r}{2}}\left[\frac{\Gamma(p+r)}{\Gamma(p)}\right]^{\frac{1}{2}}\frac{\Gamma(N)}{\Gamma(N+r)}\,L_r^{(N-1)}(y)\;.$$

It follows that (3.38) becomes

(3.47) $$h^{*}(x,y,R) = g(x,M)g(y,N)\left[1+\sum_{r=1}^{\infty} R^i\,L_r^{*(M-1)}(x)L_r^{*(N-1)}(y)\right]$$

so that the canonical correlations are

(3.48) $$R^i = \text{Corr.}\left[L_i^{*(M-1)}(x),\;L_i^{*(N-1)}(y)\right].$$

CHAPTER IV

THE BIVARIATE POISSON, BINOMIAL AND HYPERGEOMETRIC DISTRIBUTIONS

1. Introduction. Campbell (1934) derived the bivariate Poisson frequency function by taking the limiting form of the factorial moment generating function (f.m.g.f.) corresponding to fourfold sampling with replacement. Campbell then uses a theorem relating the f.m.g.f. of the Poisson distribution to Charlier's polynomials (Campbell (1932)) to expand the bivariate Poisson frequency function as a series bilinear in Charlier's polynomials. Aitken and Gonin (1935) derived the bivariate binomial frequency function as a series bilinear in Krawtchouk's polynomials (Krawtchouk (1929); and derived the canonical form of the bivariate hypergeometric frequency function.

Campbell's derivation of the bivariate Poisson frequency function, like the earlier derivations of Wicksell (1916) and McKendrick (1926), is rather indirect. Aitken and Gonin's series for bivariate binomial and hypergeometric frequency functions are incorrect (Hamdan (1963)), because of two minor algebraic mistakes.

In this chapter, a direct derivation (Hamdan (1963)) of the bivariate Poisson frequency function is given; and hence Campbell's method to obtain the corresponding canonical form is given. The correct form of the series for the bivariate binomial and hypergeometric distributions are also derived (Hamdan (1963)).

2. The Bivariate Poisson Frequency Function. When each individual of a population of N- numbers can be classified as being either A or $\bar{A}$, and at the same time either B or $\bar{B}$, the relative proportions, or probabilities, of four types AB, $A\bar{B}$, $\bar{A}B$ and $\overline{AB}$ can be set out in the fourfold table

	B	$\bar{B}$	
A	p_{11}	p_{10}	p_1
$\bar{A}$	p_{01}	p_{00}	q_1
	p_2	q_2	1

The probability of type AB is p_{11}, that of $A\bar{B}$ is p_{10}, and so on - summing the rows and entering the sums marginally we denote by p_1 and q_1 the total probabilities of A and $\bar{A}$ when B and $\bar{B}$ are disregarded; thus,

$$p_{11} + p_{10} = p_1, \; p_{01} + p_{00} = q_1, \; p_1 + q_1 = 1. \tag{4.1}$$

In the same way, summing the columns, we enter marginally p_2 and q_2.

Under random sampling n times, the numbers of occurences of A and B are jointly distributed in a bivariate binomial distribution if replacement is permitted and in a bivariate hypergeometric distribution if replacement is not permitted. Campbell (1934) derived the bivariate Poisson frequency function by using the fact that, if p_{11}, p_1 and p_2 are all of order n^{-1}, then in the case of replacement the limiting distribution as $n \to \infty$ is the bivariate Poisson.

Now, since the p.g.f. of the bivariate binomial distribution is

$$(4.2) \qquad \Big[\sum_{i,j} p_{ij}\, s_1^i\, s_2^j\Big]^n = (p_{11}\, s_1\, s_2 + p_{10}s_1 + p_{10}s_2 + p_{00})^n.$$

The substitutions $s_1 = 1 + t$, $s_2 = 1 + u$ give the f.m.g.f. $F_n(t,u)$, which in virtue of the marginal sum relations of the table takes the form

$$(4.3) \qquad F_n(t,u) = (1 + p_1 t + p_2 u + p_{11} tu)^n$$

$$= (1 + p_1 t)^n\, (1 + p_2 u)^n \Big[1 + \frac{(p_{11} - p_1 p_2)\, tu}{(1 + p_1 t)(1 + p_2 u)}\Big]^n \quad .$$

Under the assumption that p_{11}, p_1 and p_2 are all $O(n^{-1})$, we have

$$F_n(t,u) = (1 + p_1 t)^n (1 + p_2 u)^n \Big[1 + (p_{11} - p_1 p_2)\, tu(1\ p_1 t)^{-1}(1\ p_2 u\Big]^{-1}$$

$$= (1 + p_1 t)^n (1 + p_2 u)^n \Big[1 + (p_{11} - p_1 p_2)\, tu(1 + p_1 t + (p_1 t)^2_+ \ldots)$$

$$(1 + p_2 u + (p_2 u)^2_+ \ldots)\Big]^n$$

$$(4.4) \qquad = (1 + p_1 t)^n (1 + p_2 u)^n \big[1 + (p_{11} - p_1 p_2)\, tu + O(n^{-2})\big]$$

and as $n \to \infty$ we get the bivariabe Poisson f.m.g. in the form

$$(4.5) \qquad F(t,u) = \exp(np_1 t + np_2 u + n(p_{11} - p_1 p_2)\, tu)$$

$$= \exp(m_1 t + m_2 u + \tilde{m}\ tu)$$

where

$$(4.6) \qquad m_1 = np_1,\ m_2 = np_2 \text{ and } \tilde{m} = n(p_{11} - p_1 p_2).$$

By definition of the f.m.g.f. the joint Poisson frequency function

$p(x,y;\bar{m})$ satisfies the equation

$$(4.7) \qquad \exp(m_1 t + m_2 u + \bar{m} tu) = \sum_x \sum_y (1+t)^x (1+u)^y p(x,y;\bar{m})$$

writing

$$(4.8) \qquad m_1 t + m_2 u + \bar{m} tu = (m_1-\bar{m})(1+t) + (m_2-\bar{m})(1+u)$$

$$+ \bar{m}(1+t)(1+u) - (m_1+m_2-\bar{m})$$

we get

$$\exp\left\{(m_1-\bar{m})(1+t) + (m_2-\bar{m})(1+u) + \bar{m}(1+t)(1+u) - (m_1+m_2-\bar{m})\right\}$$

$$(4.9) \qquad = \exp -(m_1+m_2-\bar{m}) \sum_i^{\infty} \frac{(m_1-\bar{m})^i (1+t)^i}{i!} \sum_j^{\infty} \frac{(m_2-\bar{m})^j}{j!} (1+u)^j$$

$$\sum_r \frac{(1+t)^r (1+u)^r}{r!} \bar{m}^r.$$

To find the corresponding bivariate Poisson frequency, $p(x,y;\bar{m})$, we have to find the coefficient of $(1+t)^x(1+u)^y$ in (4.9). Hence in (4.9) the sum of i and r must be x and j and r must be y. It follows that the coefficient of $(1+t)^x(1+u)^y$ in (4.9) is

$$(4.10) \qquad p(x,y;\bar{m}) = e^{-(m_1+m_2-\bar{m})} \sum_{r=0} \frac{(m_1-\bar{m})^{x-r}}{(x-r)!} \frac{(m_2-\bar{m})^{y-r}}{(y-r)!} \frac{\bar{m}^r}{r!}$$

where the upper limit of the summation is the lesser of x and y. Putting $\bar{m}$ equal to zero and summing over all values of x (y) we get the marginal distribution of x (y) as a Poisson with parameter m_1(or m_2).

3. A Direct Derivation of the Bivariate Poisson Frequency Function. We shall give now a direct and easy derivation of $p(x,y;\bar{m})$

using the additive property of the Poisson distribution.

Let u_1, u_2, u_3 be three mutually independent Poisson variates with parameters Q_1, Q_2, Q_3 respectively, i.e. with joint frequency function

$$(4.11) \qquad f(u_1, u_2, u_3) = e^{-(Q_1 + Q_2 + Q_3)} Q_1^{u_1} Q_2^{u_2} Q_3^{u_3} / (u_1!\, u_2!\, u_3!).$$

Make the transformation

$$(4.12) \qquad \begin{aligned} x &= u_1 + u_2 \\ y &= u_2 + u_3 \\ u_2 &= u_2 \end{aligned}$$

so that x and y are Poisson variates with parameters $m_1 = Q_1 + Q_2$ and $m_2 = Q_2 + Q_3$; moreover, we have

$$\begin{aligned} m_{11} = E(xy) &= E(u_1 + u_2)(u_2 + u_3) \\ &= E(u_1\, u_2) + E(u_2^2) + E(u_2\, u_3) + E(u_1\, u_3) \\ &= Q_1\, Q_2 + Q_2 + Q_2^2 + Q_2\, Q_3 + Q_1\, Q_3 \end{aligned}$$

$$(4.13) \qquad = (Q_1 + Q_2)(Q_2 + Q_3) + Q_2$$

and

$$\begin{aligned} \overline{m} &= E(xy) - E(x)\, E(y) \\ &= (Q_1 + Q_2)(Q_2 + Q_3) + Q_2 - (Q_1 + Q_2)(Q_2 + Q_3) \end{aligned}$$

$$(4.14) \qquad = Q_2$$

Since the Jacobian of the transformation (4.12) is unity, the joint frequency function of x,y and u_2 is

$$(4.15) \qquad f(x,y,u_2) = e^{-(Q_1+Q_2+Q_3)} \frac{Q_1^{x-u_2} Q_3^{y-u_2} Q_2^{u_2}}{(x-u_2)!(y-u_2)!u_2!}$$

$$(4.16) \qquad = e^{-(m_1+m_2-\bar{m})} \frac{(m_1-\bar{m})^{x-u_2}(m_2-\bar{m})^{y-u_2}\,\bar{m}^{u_2}}{(x-u_2)!\,(y-u_2)!\,u_2!}$$

Summing $f(x,y,u_2)$ for all values of u_2, we get the joint frequency function of x and y in the form

$$(4.17) \qquad p(x,y;\bar{m}) = e^{-(m_1+m_2-\bar{m})} \sum_{u_2=0}^{\infty} \frac{(m_1-\bar{m})^{x-u_2}(m_2-\bar{m})^{y-u_2}\,\bar{m}^{u_2}}{(x-u_2)!\,(y-u_2)!\,u_2!}$$

where the upper limit of the summation is the lesser of x and y.

4. Another Direct Proof of the Bivariate Poisson Distribution:

Let

(4.18)

	B	$\bar{B}$	
A	p_{11}	p_{10}	p_1
$\bar{A}$	p_{01}	p_{00}	q_1
	p_2	q_2	1

be the fourfold table given in section 2 of this chapter.

We may regard this distribution (Kendall V. I) as a multinomial arrayed by

$$(4.19) \qquad (p_{00} + p_{01} + p_{10} + p_{00})^n.$$

The probability of having x A's and y B's, where we get i (AB)'s, $\{i \leq \min(x,y)\}$, is

$$(4.20)\qquad \binom{n}{i,x-i,y-i} p_{11}^{i}\, p_{10}^{x-i}\, p_{01}^{y-i}\, p_{00}^{n+i-x-y}$$

Hence, the probability of x A's and y B's is

$$(4.21)\qquad \sum_{i} \binom{n}{i,x-i,y-i} p_{11}^{i}\, p_{10}^{x-i}\, p_{01}^{y-i}\, p_{00}^{n+i-x-y} \ .$$

For fixed x and y let $n \to \infty$ and $p_{11}, p_{10}, p_{01} \to 0$ i.e. they are of order $\frac{1}{n}$, in such a manner that $np_{11} = \bar{m}$, $np_{10} = m_1 - \bar{m}$ and $np_{01} = m_2 - \bar{m}$ remain fixed, then with

$$(4.22)\qquad p_{11} = \frac{\bar{m}}{n},\ p_{10} = \frac{m_1 - \bar{m}}{n} \quad \text{and } p_{01} = \frac{m_2 - \bar{m}}{n}$$

$$p_{00} = 1 - \frac{(m_2 + m_1 - \bar{m})}{n}$$, it follows that

$$(4.23)\qquad \lim_{n\to\infty} \frac{n}{i!\,(x-i)!\,(y-i)!\,(n+i-x-y)!} \left(\frac{\bar{m}}{n}\right)^{i} \left(\frac{m_1-\bar{m}}{n}\right)^{x-i} \left(\frac{m_2-\bar{m}}{n}\right)^{y-i} \times \left[1 - \frac{(m_2+m_1-\bar{m})}{n}\right]^{n+i-x-y}$$

$$= \lim_{n\to\infty} \frac{n(n-1)\ldots(n-i+1)(n-i)\ldots(n-x+1)(n-x)\ldots(n-x-y+i+1)(n-x-y+i)!}{i!\,(x-i)!\,(y-i)!\,(n+i-x-y)!} \times \left(\frac{\bar{m}}{n}\right)^{i} \left(\frac{m_1-\bar{m}}{n}\right)^{x-i} \left(\frac{m_2-\bar{m}}{n}\right)^{y-i} \left[1 - \frac{(m_1+m_2-\bar{m})}{n}\right]^{n+i-x-y}$$

$$(4.24)\qquad = \lim_{n\to\infty} \frac{1\ldots\left(1-\frac{i-1}{n}\right)\left(1-\frac{i}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)\left(1-\frac{x}{n}\right)\ldots\left(1-\frac{x+y-i-1}{n}\right)}{i!\,(x-i)!\,(y-i)!} \times (\bar{m})^{i} (m_1-\bar{m})^{x-i} (m_2-\bar{m})^{y-i} \left[1 - \frac{(m_1+m_2-\bar{m})}{n}\right]^{n+i-x-y} .$$

But

$$(4.25) \qquad \lim_{n\to\infty} \left[1 - \frac{(m_1 + m_2 - \bar{m})}{n}\right]^n = e^{-(m_1 + m_2 - \bar{m})}$$

Hence (4.24) becomes

$$(4.26) \qquad = e^{-m_1 - m_2 + \bar{m}} \frac{(\bar{m})^i}{i!} \frac{(m_1 - \bar{m})^{x-i}}{(x-i)!} \frac{(m_2 - \bar{m})^{y-i}}{(y-i)!}$$

Therefore; (4.21) becomes

$$(4.27) \qquad e^{-m_1 - m_2 + \bar{m}} \sum_i \frac{(\bar{m})^i}{i!} \frac{(m_1 - \bar{m})^{x-i}}{(x-i)!} \frac{(m_2 - \bar{m})^{y-i}}{(y-i)!}$$

where the upper limit of summation is $\min(x,y)$ (4.27) is exactly the same as (4.17).

5. The Canonical Form of the Bivariate Poisson Distribution.

Let the bivariate Poisson f.m.g.f., $F(t,u)$ be expressed in the following form

$$F(t,u) = \exp(m_1 t + m_2 u + \bar{m} tu)$$

$$= \exp(m_1 t + m_2 u) \sum_{r=0}^{\infty} (tu)^r \bar{m}^r / r!$$

$$(4.28) \qquad = \sum_{r=0}^{\infty} \left[t^r \exp(m_1 t)\right] \left[u^r \exp(m_2 u)\right] . \bar{m}^r / r!$$

By lemma 1.4, the f.m.g.f. of $k_r(x)p(x)$ is $t^r \exp(mt)$. It follow that $F(t,u)$ is the f.m.g.f. of

$$(4.29) \qquad \sum_{r=0}^{\infty} \left[k_r(x)\, p(x)\right] \left[k_r(y)\, p(y)\right] \bar{m}^r / r!$$

But $F(t,u)$ is the f.m.g.f. of $p(x,y;\bar{m})$; so by the uniqueness theorem, it follows that

(4.30) $$p(x,y;,\bar{m}) = p(x;m_1)\ p(y;m_2)\left[1+\sum_{r=1}^{\infty}k_r(x;m_1)k_r(y;m_2)\bar{m}^r\ /r!\right]$$

which is the canonical form of $p(x,y;\bar{m})$.

Theorem 4.1 $\bar{m}$ is the covariance of x and y. Proof:

(4.31) $$E(xy) = \sum_x\sum_y xy\ p(x;m_1)p(y;m_2)\left[1+\sum_{r=1}^{\infty}k_r(x;m_1)k_r(y;m_2)\bar{m}^r/r!\right]$$

Because of the orthogonality of the set $\{k_r(x)\}$ (Theorem 1.4) it follows that the summation in (4.31) is zero for all r except when $r = 1$.

(4.32) $$E(xy) = m_1m_2+\bar{m}\sum_x\sum_y k_1(x;m_1)k_2(y,m_1)p(x;m_1)p(y;m_2)xy.$$

Using (1.46), we can find

(4.33) $$k_1(x;m_1) = \frac{x - m_1}{m_1}$$

and

(4.34) $$k_1(y;m_2) = \frac{y-m_2}{m^2}.$$

Substituting (4.23) and (4.24) in (4.32) and suming over all values of x and y, we get (4.34) $E(xy) = m_1m_2+\bar{m}$.

Hence covariance of $xy = E(xy) - E(x)E(y) = \bar{m}$.

In order to find the regression line of x on y we first find $E(x/y)$

(4.35) $$E(x/y) = \sum_x xp(x/y) = \sum_x xp(x;m_1)\left[1+\sum_{r=1}^{\infty}k_r(x;m_1)k_r(y,m_2)\frac{\bar{m}^r}{r!}\right]$$

$$= m_1+\sum_x xk_r(x)k_r(y)p(x;m_1)\frac{\bar{m}^r}{r!}$$

$$= m_1+\sum_x xk_1(x)k_1(y)p(x;m_1)\bar{m}$$

Using (4.33) & (4.34) we get

$$= m_1 + \bar{m}\left(\frac{y-m_2}{m_2}\right)\sum_x x\,\frac{x-m_1}{m_1}\;p(x;m_1)$$

(4.36) $$= m_1 + \bar{m}\left(\frac{y-m_2}{m_2}\right)\cdot 1$$

Therefore

(4.37) $$= x - m_1 = \frac{\bar{m}}{m_2}(y - m_2)$$ is the regression line of x on y.

Similarly

(4.38) $$y - m_2 = \frac{\bar{m}}{m_1}(x-m_1)$$ is the regression line of y on x.

Hence the coefficient of correlation ρ, if defined in the usual manner as the geometric mean of the regression coefficients, is given by

(4.39) $$\rho = \bar{m} / (m_1 m_2)^{\frac{1}{2}} .$$

To compare (4.19) with the Mehler identity let the set $\{k_r(x;m_1)\}$ be standardized in the form

(4.40) $$k_r^*(x;m_1) = k_r(x;m_1)(m_1^r / r!)^{\frac{1}{2}}$$

with $k_r^*(y;m_2)$ defined similarly. It follows (4.29) becomes

(4.41) $$p(x,y;\rho) = p(x;m_1)p(y;m_2)\left[1+\sum_{r=1}^{\infty}\rho^r k_r^*(x;m_1)k_r^*(y,m_2)\right]$$

so that the canonical correlations are

(4.42) $$\rho^i = \text{Corr.}\left[k_i^*(x;m_1),\ k_i^*(y,m_1)\right].$$

6. <u>The Canonical Form of the Bivariate Binomial Distribution.</u>

We have (equation (4.3)) the f.m.g.f. of the bivariate binomial distribution is given by

$$F_n(t,u) = (1+p_1t)^n(1+p_2u)^n\left[1+\frac{(p_{11}-p_1p_2)\,tu}{(1+p_1t)(1+p_2u)}\right]^n . \tag{4.43}$$

Putting

$$d = p_{11} - p_1p_2 \tag{4.44}$$

we expand $F_n(t,u)$ in the form

$$F_n(t,u) = \sum_{r=0}^{n} \binom{n}{r} (tud)^r (1+p_1t)^{n-r}(1+p_2u)^{n-r} . \tag{4.45}$$

By (1.66)

$$n^{(r)} p_1^r q_1^r t^r(1+p_1t)^{n-r} \tag{4.46}$$

is the f.m.g.f. of the product $G_r(x;n,p_1)b(x;np_1)$, therefore;

$F_n(t,u)$ is the f.m.g.f. of

$$\sum_{r=0}^{n} \binom{n}{r} d^r \left[\frac{G_r(x)b(x)}{n^{(r)}p_1^r q_1^r}\right]\left[\frac{G_r(y)b(y)}{n^{(r)}p_2^r q_2^r}\right] \tag{4.47}$$

and by the uniqueness theorem, the bivariate binomial frequency function, $b(x,y;d)$ say, should be identical with the series (4.47), i.e.

$$b(x,y;d) = b(x;n,p_1)b(y;n,p_2)\left[1+\sum_{r=1}^{n} \frac{d^r}{r!\,n^{(r)}(p_1p_2q_1q_2)^r} \cdot G_r(x)G_r(y)\right] \tag{4.48}$$

which is the canonical form of the bivariate binomial frequency function. Aitken and Gonin (1935) made an algebraic mistake (discovered by Hamdan (1963)) in the derivation of (4.48), which led to a wrong series with no (r!) in the denomenator of the general term.

Equation (4.48) shows that when $d = 0$ or $p_{11} = p_1p_2$, x and y are statistically independent.

In order to find the regression line of x on y, we need to find $E(x/y)$.

$$(4.49) \qquad E(x/y) = \sum_x \frac{x\, p(x,y;d)}{p(y)} .$$

Using equation (4.48), we get

$$(4.50) \qquad E(x/y) = \sum_x xp(x) + \sum_x \sum_{r=1}^{n} \frac{xd^r\, G_r(x)G_r(y)\,p(x)}{r!\, n^{(r)}(p_1p_2q_1q_2)^r} .$$

Because of the orthogonality of the set $\{G_r(x)\}$ (Theorem 1.5), we get

$$(4.51) \qquad E(x/y) = np_1 + G_1(y)\, \frac{d}{n(p_1p_2q_1q_2)} \sum_x x\, G_1(x)\,p(x) .$$

But $G_1(y)$ is given by (1.58) as

$$G_1(y) = (y - np_2) .$$

It follows that (4.51) becomes

$$E(x/y) = np_1 + (y - np_2) \;\frac{d}{n(p_1p_2q_1q_2)}\; (np_1q_1)$$

$$(4.52) \qquad = np_1 + \frac{d}{(p_2\, q_2)} \; (y - np_2)$$

Hence

$$(4.53) \qquad x - np_1 = \frac{d}{p_2\, q_2}\, (y - np_2)$$

is the regression line of x on y.

Similarly

$$(4.54) \qquad y - np_2 = \frac{d}{p_1\, q_1}\; (x - np_1)$$

is the regression line of y on x. It follows that by (4.53) and (4.54) the coefficient of correlation ρ, if defined in the usual manner as the geometric mean of the regression coefficients is given by

$$(4.55) \qquad \rho = d \,/\, (p_1 p_2 q_1 q_2)^{\frac{1}{2}}.$$

To compare (4.48) with the Mehler identity, let the set $\{G_r(x,n,p_1)\}$ be standardized in the form

$$(4.56) \qquad G_r^*(x;n,p_1) = G_r(x;n,p_1) \,/\, (r!\, n^{(r)} p_1^r \, q_1^r)^{\frac{1}{2}}$$

with $G_r^*(y;n,p_2)$ definied similarly. It follows that (4.48) becomes

$$(4.57) \qquad b(x,y;\rho) = b(x;n,p_1)\, b(y;n,p_2) \left[1 + \sum_{r=1}^{n} \rho^r \, G_r^*(x;n,p_1)\, G_r^*(y;n,p_2)\right]$$

so that the canonical correlations are

$$(4.58) \qquad \rho^i = \text{corr.}\left[G_r^*(x;n,p_1),\; G_r^*(y;n,p_2)\right].$$

7. <u>The Canonical Form of the Bivariate Hypergeometric Distribution</u>. Generalizing the result (1.76), the f.m.g.f. of the bivariate hypergeometric distribution (resulting from fourfold sampling without replacement) is the coefficient of z^n in the expansion of

$$(4.59) \qquad F_n(a,b) = \left[1 + (1+a)(1+b)\, z\right]^{Np_{11}} \left[1 + (1+a)z\right]^{Np_{10}} \left[1 + (1+b)z\right]^{Np_{01}} (1+z)^{Np_{00}} \Big/ \binom{N}{n}.$$

Let us write d for $p_{11} - p_1p_2 = p_{00} - q_1q_2 = -(p_{10} - p_1q_2)$

$$= -(p_{01} - p_2q_2).$$

The f.m.g.f. is then the coefficient of z^n in

$$(4.60) \qquad (1+z)^{N-Np_1-Np_2}(1+z+az)^{Np_1}(1+z+bz)^{Np_2} \left\{1+\frac{abz}{(1+z+az)(1+z+bz)}\right\}^{Np_1p_2+Nd} \Big/ \binom{N}{n} .$$

For brevity we write

$$w = abz(1+z+az)^{-1}(1+z+bz)^{-1}$$

$$Q_r = (Np_1 - r)(Np_2 - r)\ (N - 2r)^{-1} - Np_1p_2$$

so that $F_n(a,b)$ becomes the coefficient of z^n in

$$(1+z)^{N-Np_1-Np_2}(1+z+az)^{Np_1}(1+z+bz)^{Np_2}(1+w)^{Np_1p_2+Nd} \Big/ \binom{N}{n}.$$

Now let

$$(4.61) \qquad (1+w)^{Nd} = 1 + c_1w(1+w)^{Q_1} + c_2w^2(1+w)^{Q_2} + \ldots.$$

This gives a set of equations for the c_r, in terms of Nd, for which the coefficients c_r can be determined successively. For example

$c_1 = Nd$, $c_2 = \binom{Nd}{2} - Q_1Nd$, ... Hence (4.60) becomes

$$(4.62) \qquad (1+z)^{N-Np_1-Np_2}(1+z+az)^{Np_1}(1+z+bz)^{Np_2}(1+w)^{Np_1p_2} \left[1+c_1w(1+w)^{Q_1} + \ldots \ldots + c_rw^r(1\ w)^{Q_r} + \ldots\right] \Big/ \binom{N}{n}$$

If $d = 0$, x and y are uncorrelated and hence the coefficient of z^n

in
$$(1+z)^{N-Np_1-Np_2}(1+z+az)^{Np_1}(1+z+bz)^{Np_2}(1+w)^{Np_1Np_2}/\binom{N}{n}$$

is

(4.63) $$F(-n,-Np_1,-N;-a)F(-n,-Np_2,-N;-b),$$

General term in (4.62) is

(4.64) $$= (1+z)^{N-Np_1-Np_2}(1+z+az)^{Np_1}(1+w)^{\frac{(Np_1-r)(Np_2-r)}{(N-2r)}} c_r w^r/\binom{N}{n}.$$

Using the identity

$$\binom{N}{n} = \frac{N^{(r)}}{(N-n)^{(r)}\,n^{(r)}}\binom{N-2r}{n-r}$$

(4.64) becomes

(4.65) $$= c_r a^r b^r z^r (1+z)^{(N-2r)-(Np_1-r)-(Np_2-r)}(1+z+az)^{Np_1-r}(1+z+bz)^{Np_2-r}$$

$$\cdot\,(1+w)^{\frac{(Np_1-r)(Np_2-r)}{N-2r}}\cdot\frac{n^{(r)}(N-n)^{(r)}}{N^{(2r)}}\Big/\binom{N-2r}{n-r}.$$

By (4.63) and (4.65), it follows that the coefficient of z^n in the general term of (4.62) is

(4.66) $$c_r a^r b^r\,\frac{n^{(r)}(N-n)^{(r)}}{N^{(2r)}}\,F(-n+r,-Np_1+r,-N+2r,-a)$$

$$\cdot F(-n+r,-Np_2+r,-N+2r,-b).$$

Hence $F_n(a,b)$ is

(4.67) $$= \sum_{r=0}^{\infty} c_r a^r b^r\,\frac{n^{(r)}(N-n)^{(r)}}{N^{(2r)}}\,F(-n+r,-Np_1+r,-N+2r;-a)$$

$$\times F(-n+r,-Np_2+r,-N+2r;-b)$$

where the upper limit of the summation is the smaller of n and (N-n).

By (1.75), the f.m.g.f. of $U_r(x)\ h(x)$ is

$$F(a) = \frac{n^{(r)}(Np_1)^{(r)}(N-n)^{(r)}(Nq_1)^{(r)}a^r}{N^{(2r)}(N-r\ \ 1)^{(r)}}\ F(-n+r, -Np_1+r, -N+2r;\ -a). \tag{4.68}$$

and then by the uniqueness theorem, it follows that the bivariate hypergeometric frequency function, $h(x,y;d)$, has the canonical expansion

$$h(x,y;d) = h(x)h(y)\left[1+\sum_{r=1}^{\infty} c_r\, \frac{N^{(2r)}\left\{(N-r+1)^{(r)}\right\}^2 U_r(x)U_r(y)}{n^{(r)}(N-n)^{(r)}(Np_1)^{(r)}(Np_2)^{(r)}(Nq_1)^{(r)}(Nq_2)^{(r)}}\right] \tag{4.69}$$

where the coefficients c_r are given by (4.61) and the upper limit of the summation is the smaller of n and (N-n). Aitken and Gonin (Hamdan (1963)) made an algebraic mistake in the derivation of the above series (4.69), thus getting an incorrect series with a factor $(N-r+1)^{(r)}$ missing from the general term. To find the regression line of x on y we need to find

$$E(x/y) = \sum_x x\, \frac{h(x,y;d)}{h(y)}\ . \tag{4.70}$$

Substituting for $h(x,y;d)$ by (4.69) we get

$$E(x/y) = \sum_x x\, h(x) + \sum_x \sum_{r=1}^{\infty} x c_r \cdot \frac{N^{(2r)}\left\{(N-r+1)^{(r)}\right\}^2 U_r(x)U_r(y)}{n^{(r)}(N-n)^{(r)}(Np_1)^{(r)}(Np_2)^{(r)}(Nq_1)^r(Nq_2)^r}\ . \tag{4.71}$$

By theorem (1.6) it follow that (4.71) becomes

$$= np_1 + \frac{c_1\, N^{(2)}\ N^2\ U_1(y)}{n(N-n)(Np_1)(Np_2)(Nq_1)(Nq_2)} \sum_x x\, U_1(x)h(x). \tag{4.72}$$

By (1.69) $U_1(x)$, $U_1(y)$ are given by

$$U_1(x) = x - np_1 \quad \text{and} \quad U_1(y) = y - np_2 \,.$$

Substituting $c_1 = Nd$ we get

$$E(x/y) = np_1 + \frac{d(N-1)(y-np_2)}{n(N-n)(p_1p_2q_1q_2)} \sum_x x(x-np_1)h(x)$$

(4.73) $$= np_1 + \frac{d(y-np_2)}{p_2q_2} \,.$$

It follows that

(4.74) $$x-np_1 = \frac{d}{p_2q_2}(y-np_2)$$

is the regression line of x on y.

Similarly

(4.75) $y - np_2 = \frac{d}{p_1q_1}(x-np_1)$ is the regression line of y on x.

(4.74) and (4.75) are the same regression lines as those of the bivariate binomial distribution given by (4.52) and (4.53).

(4.74) and (4.75) are straight lines, and pass through the double mean (np_1, np_2) of the distribution; while the correlation coefficient ρ, if defined in the usual way as the geometric mean of the regression coefficients, is given by

(4.76) $$\rho = d(p_1p_2q_1q_2)^{-\frac{1}{2}}.$$

which is identical with that of the bivariate binomial distribution given by equation (4.55).

CHAPTER V

THE BIVARIATE BETA, t AND F DISTRIBUTIONS

1. <u>Introduction</u>. Watson (1933) derived an expression for the sum of a series, bilinear in Legendre Polynomials and generalized the result to ultraspherical polynomials (definition 1.10)[1]. Rice (1945) derived the characteristic function of the bivariate distribution of two sine waves of the form

$$(5.1) \qquad x = \cos w\, t$$

$$y = \cos\ w(t + \theta)$$

and hence derived the corresponding frequency function $f(x,y,w)$. Barret and Lampard (1955) expanded $f(x,y,w)$ in the canonical form

$$(5.2) \qquad f(x,y,w) = B(x)\, B(y) \left[1 + \sum_{n=1}^{\infty} 2 \cos(n\, w\, \theta)\, T_n(x)\, T_n(y)\right.$$

where

$$B(x) = (1 - x^2)^{-\frac{1}{2}} / \pi\ , \ -1 \leq x \leq 1$$

and $\{T_n(x)\}$ is the set of Chebyshev polynomials of the 1st kind (definition 1.7). Leipnik (1958) derived a bivariate frequency function in the form of a series bilinear in the transformed Legendre Polynomials $p_n(x)$ defined by 1.9 (orthogonal on (0,1),

---

1. They are sometimes called Gegenbauer's Polynomials.

$$(5.3) \qquad h(x,y,\rho) = 1 + \sum_{n=1}^{\infty} \rho_n \, p_n(x) \, p_n(y)$$

$$= 1 - 2\rho - 2\rho \log G(x,y), \quad 0 \leq x,y \leq 1$$

where $G(x,y)$ is Green's kernel

$$(5.4) \qquad G(x,y) = x(1-y), \quad x \geq y$$

$$= y(1-x), \quad x \leq y$$

and

$$(5.5) \qquad \rho_n = \rho \, 2^{2n+1}(n!)^4 / \left\{ i(i+1) \left[(2i)! \right]^2 \right\}$$

and $0 \leq \rho \leq \frac{1}{2}$ so that $h(x,y,\rho) \geq 0$ for all $0 \leq x,y \leq 1$

Now, the following question arises: "Is there a generalization of the above results in the form of a series bilinear in the Jacobi polynomials (definition 1.6)?" (Hamdan (1963)). Because of the complexity of the generating function of the Jacobi polynomials (equation 1.35), it is difficult to approach the problem by using the g.f. similar to the approach used in the previous chapters.

From a statistical point of view, (5.2) is a special form of the bivariate beta distribution since by change of variables $x = 2X - 1$ and $y = 2Y - 1$ the marginal frequency functions become $X^{-\frac{1}{2}}(1 - X)^{-\frac{1}{2}} / \beta(\frac{1}{2},\frac{1}{2})$ and $Y^{-\frac{1}{2}}(1 - Y)^{-\frac{1}{2}} / \beta(\frac{1}{2},\frac{1}{2})$ with orthogonal polynomials $P_n^{(-\frac{1}{2},-\frac{1}{2})}(2X - 1)$ and $P_n^{(-\frac{1}{2},-\frac{1}{2})}(2Y-1)$, $0 \leq X, Y \leq 1$.

In this chapter, we give a series form of the bivariate beta distribution in general by transformation from the bivariate gamma distribution. This approach is due to Hamdan (1963). However, the result is not in canonical form. This result is used to derive series forms for the bivariate t and F distributions.

2. The Bivariate Beta Distribution. It is well known that if x and y are independent gamma variables with parameters p and q respectively, then the variable $z = x/(x+y)$ has a beta distribution with frequency function

$$B(z;\, p,\, q) = z^{p-1}(1-z)^{q-1}/\beta(p,q), \quad 0 \le z \le 1. \tag{5.6}$$

Now, let x_1 and x_2 be independent gamma variables with parameters p and m; and let y_1 and y_2 be another pair of independent gamma variables with parameters q and m, such that x_1 and y_2 are independent, x_2 and y_1 are independent, but x_2 and y_2 are distributed in a bivariate gamma distribution with coefficient of Correlation ρ. The joint probability element will be

$$\frac{1}{\Gamma(p)\,[\Gamma(m)]^2\,\Gamma(q)}\, e^{-(x_1+x_2+y_1+y_2)}\, x_1^{p-1} x_2^{m-1} y_1^{q-1} y_2^{m-1} \sum_{r=0}^{\infty} \frac{r!\,\Gamma(m)}{\Gamma(m+r)}\, \rho^r\, L_r^{(m-1)}(x_2) L_r^{(m-1)}(y_2)\; dx_1 dx_2 dy_1 dy_2. \tag{5.7}$$

Make the following transformation

$$X = x_1/(x_1+x_2)$$

$$Y = y_1/(y_1+y_2)$$

$$x_2 = x_2 \tag{5.8}$$

$$y_2 = y_2$$

so that

$$(5.9) \qquad x_1 = Xx_2/(1-X),\ y_1 = Yy_2/(1-Y) \text{ and } \frac{\partial x_1}{\partial X}\frac{\partial y_1}{\partial Y} = \frac{x_2 \cdot y_2}{(1-X)^2(1-Y)^2}\ .$$

Hence, the joint frequency function of X, Y, x_2 and y_2 is

$$(5.10) \qquad \frac{1}{\Gamma(p)\,[\Gamma(m)]^2\,\Gamma(q)} \exp\left(-\frac{x_2}{1-X} - \frac{y_2}{1-Y}\right) x_2^{m\ \ p-1}\, y_2^{m+q-1}$$
$$\frac{X^{p-1}\,Y^{q-1}}{(1-X)^{p+1}(1-Y)^{q+1}} \sum_{r=0}^{\infty} \frac{r!\,\Gamma(m)}{\Gamma(m+r)} \rho^r\, L_r^{(m-1)}(x_2)\, L_r^{(m-1)}(y_2)$$

We integrate (5.10) first with respect to x_2 from 0 to ∞, we get (leaving out quantities not containing x_2)

$$(5.11) \qquad \int_0^\infty \exp\left(\frac{-x_2}{1-X}\right) x_2^{m+p-1}\, L_2^{(m-1)}(x_2)\, dx_2\ .$$

By equation (1.19), $L_r^{(m-1)}(x_2)$ is the coefficient of $(-t)^r$ in $(1-t)^{-m} \exp\left[-x_2 t\ /\ (1-t)\right]$. It follows that (5.11) is

$$= \text{coefficient of } (-t)^r \text{ in } (1-t)^{-m} \int_0^\infty \exp \frac{-x_2(1-tX)}{(1-t)(1-X)}\ x_2^{m+p-1}\, dx_2$$

$$= \text{coefficient of } (-t)^r \text{ in } \Gamma(m+p)(1-X)^{m+p}\,(1-t)^p/(1-tX)^{m\ \ p}$$

$$(5.12) \qquad = \Gamma(m+p)(1-X)^{m\ \ p}\, A_r^{(p,m)}\ (X)$$

where

$$A_r^{(p,m)}(X) = \text{coefficient of } (-t)^r \text{ in } (1-t)^p/(1-tX)^{m+p}$$

$$(5.13) \qquad = \sum_{i=0}^{r} (-1)^i \binom{m+p+i-1}{i}\binom{p}{r-i} X^i$$

is a polynomial of degree r in X.

We integrate (5.10) again with respect to y_2, thus getting the bivariate frequency function of X and Y in the form

$$(5.14)\qquad k(X,Y,\rho) = \frac{X^{p-1}(1-X)^{m-1}}{(p,m)} \quad \frac{Y^{q-1}(1-Y)^{m-1}}{(q,m)}$$

$$\left[1+\sum_{r=1}^{\infty}\frac{r!\,\Gamma(m)}{\Gamma(m+r)}\rho^r A_r^{(p,m)}(X)\,A_r^{(q,m)}(Y)\right]$$

$$(5.15)\qquad = B(X,p,m)\,B(Y,q,m)\left[1+\sum_{r=1}^{\infty}\frac{r!\,\Gamma(m)}{\Gamma(m+r)}\rho^r A_r^{(p,m)}(X)\,A_r^{(q,m)}(Y)\right].$$

It can be easily verified that each of the marginals is a beta distribution, since (Hamdan (1963))

$$(5.16)\qquad \int_0^1 A_r^{(p,m)}(X)\,B(X,p,m)dX$$

$$= \sum_{i=0}^{r}(-1)^i\binom{m+p+i-1}{i}\binom{p}{r-i}\beta(p+i,m)/\beta(p,m)$$

$$= \sum_{i=0}^{r}(-1)^i\binom{p}{r-i}\binom{p\quad i-1}{i} = \sum_{i=0}^{r}\binom{-p}{i}\binom{p}{r-i}$$

$$(5.17)\qquad = \text{coefficient of } t^r \text{ in } (1+t)^p\,(1+t)^{-p} = 0 \text{ for } t \neq 0.$$

Now if we make the transormation $x = 2X - 1$ in the Jacobi polynomials we get

$$(5.18)\qquad P_r^{(m-1,p-1)}(2X-1) = \sum_{i=0}^{r}(-1)^i\binom{r+m-1}{r-i}\binom{r+m-1}{i}(1-X)^i X^{r-i}$$

which is orthogonal on the marginal distribution of X.

Obviously, the set $\left\{A_n^{(p,m)}(X)\right\}$ is not identical with the set $\left\{P_n^{(m-1),p-1)}(2X-1)\right\}$. Hence (5.14) is not in canonical form. The problem of deriving a bivariate beta in canonical form remains to be solved.

2. <u>Bivariate t Distribution</u>. Siddiqui (1967) derived a form for the bivariate distribution, he used the joint distribution of $(\bar{x}, \bar{y}, s_1, s_2, r)$ to work out the distribution of (t_1, t_2, r) where t_1 corresponds to the x-observations and t_2 to y-observations from a bivariate normal distribution with a correlation coefficient ρ.

Siddiqui evaluated the exact distributions only for $n = 1$, and 3 (sample size $N = 2$ and 4). He indicated that the exact distribution for arbitrary n can be worked out, following the method for $n = 3$.

The exact distribution for $n = 1$ is:

$$(5.19) \qquad h_1(t_1, t_2; \rho) = \left[(1-\rho^2)\ \text{Cosec}^2\theta \,/\, 4\pi^2(1+t_1^2)(1+t_2^2)\right]$$

$$\cdot \left[1 + (\pi - \theta)\ \text{Cot}\ \theta\right],$$

where

$$\text{Cos}\ \theta = 2\rho(1-\rho^2)(1+t_1\ t_2)(1+t_1^2)^{-\frac{1}{2}}(1+t_2^2)^{-\frac{1}{2}}$$

and

$$\theta = \theta\ (t_1, t_2) \text{ is between } 0 \text{ and } \pi.$$

Making use of the bivariate gamma distribution and following the lines of the proof in section one we can find a simpler form by an easier method without using the distribution of $(\bar{x}, \bar{y}, s_1, s_2, r)$.

Let $x, y_1, \ldots y_n$ be a sample of $n + 1$ from the normal distribution $(0, \sigma^2)$. A t variate with n degrees of freedom is defined as

$$(5.20) \qquad t = \frac{x}{(\sum y_i^2/n)^{\frac{1}{2}}} \quad \text{or } nt^2 = \frac{x^2}{\sum y_i^2}$$

with a density function

$$(5.21) \qquad g(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

It follows that

$$(5.22) \qquad \frac{t^2}{n} = \frac{\chi_1^2}{\chi_n^2} .$$

It can be easily shown that a χ^2 variate with n degrees of freedom is a gamma variable with parameter n / 2, hence (5.22) can be expressed as

$$(5.23) \qquad \frac{t^2}{n} = \frac{\Gamma_{\frac{1}{2}}}{\Gamma_{\frac{n}{2}}} .$$

Now, let x_1, x_2 be independent gamma variables with parameters $\frac{1}{2}$ and $\frac{n}{2}$ respectively, and let y_1, y_2 be another pair of independent gamma variables with parameters $\frac{1}{2}$ and $\frac{n}{2}$, such that x_1 and y_2 are independent, x_2 and y_1 are independent, but x_2 and y_2 are distributed in a bivariate gamma distribution with a coefficient of correlation ρ. The joint probability element will be

$$(5.24) \qquad \frac{1}{[\Gamma(\frac{n}{2})\ \Gamma(\frac{1}{2})]^2}\, e^{-(x_1+x_2+y_1+y_2)}\, x_1^{-\frac{1}{2}}\, x_2^{\frac{n}{2}-1}\, y_1^{-\frac{1}{2}}\, y_2^{\frac{n}{2}-1}$$

$$\cdot \sum_{r=0}^{\infty} \frac{r!\,\Gamma(\frac{n}{2})}{\Gamma(\frac{n}{2}+r)}\, \rho^r\, L_r^{(\frac{n}{2}-1)}(x_2)\, L_r^{(\frac{n}{2}-1)}(y_2)\, dx_1 dx_2 dy_1 dy_2 .$$

Let us make the following transormation

$$\frac{t_1^2}{n} = \frac{x_1}{x_2} ,$$

(5.26) $$\frac{t_2^2}{n} = \frac{y_1}{y_2},$$

$$x_2 = x_2,$$

and

$$y_2 = y_2.$$

It follows that

$$x_1 = \frac{t_1^2}{n} x_2,$$

(5.26) $$y_1 = \frac{t_2^2}{n} y_2 \quad \text{and} \quad \frac{\partial x_1}{\partial t_1} \frac{\partial y_1}{\partial t_2} = \frac{4t_1 t_2}{n^2} x_2 y_2.$$

Noting that there are two values of t_1 and two values of t_2 corresponding to each pair of values of (x_1, y_1) and (x_2, y_2), it follows that the joint probability element of t_1, t_2, x_2 and y_2 is

(5.27) $$= \frac{1}{4n^2 \pi [\Gamma(\frac{n}{2})]^2} e^{-(x_2 \frac{t_1^2}{n} + x_2 + y_2 \frac{t_2^2}{n} + y_2)}$$

$$\cdot (\frac{t_1^2}{n} x_2)^{-\frac{1}{2}} (\frac{t_2^2}{n} y_2)^{-\frac{1}{2}} y_2^{\frac{n}{2}-1} x_2^{\frac{n}{2}-1}$$

$$\cdot \sum_{r=0}^{\infty} \frac{r! \Gamma(\frac{n}{2})}{\Gamma(\frac{n}{2}+r)} \rho^r L_r^{(\frac{n}{2}-1)}(x_2) L_r^{(\frac{n}{2}-1)}(y_2)$$

$$\cdot 4 t_1 t_2 x_2 y_2 dt_1 dt_2 dx_2 dy_2$$

(5.28) $$= \frac{1}{n\pi [\Gamma(\frac{n}{2})]^2} \exp[-x_2(1+\frac{t_1^2}{n}) - y_2(1+\frac{t_2^2}{n})] y_2^{\frac{n}{2}-\frac{1}{2}} x_2^{\frac{n}{2}-\frac{1}{2}}$$

$$\cdot \sum_{r=0}^{\infty} \frac{\Gamma(\frac{n}{2}) r!}{\Gamma(\frac{n}{2}+r)} \rho^r L_r^{(\frac{n}{2}-1)}(x_2) L_r^{(\frac{n}{2}-1)}(y_2) dt_1 dt_2$$

$$\cdot dx_2 dy_2 .$$

Integrating first with respect to x_2 from 0 to ∞ , we get an integral of the form (leaving out quantities not containing x_2)

$$(5.29)\qquad \int_0^\infty \exp\left[-x_2\left(1+\frac{t_1^2}{n}\right)\right] x_2^{\frac{n+1}{2}-1} L_r^{(\frac{n}{2}-1)}(x_2)\, dx_2$$

$$= \text{coefficient of } (-u)^r \text{ in } (1-u)^{-\frac{n}{2}}\, \Gamma\left(\frac{n+1}{2}\right) \left[1+\frac{t_1^2}{n}+\frac{u}{1-u}\right]^{-(\frac{n+1}{2})}$$

$$= \text{coefficient of } (-u)^r \text{ in } \Gamma\left(\frac{n+1}{2}\right)\left(1+\frac{t_1^2}{n}\right)^{-\frac{n+1}{2}} (1-u)^{\frac{1}{2}} \cdot \left[1-\frac{u\frac{t_1^2}{n}}{(1+\frac{t_1^2}{n})}\right]^{-(\frac{n+1}{2})}$$

$$(5.30)\qquad = \text{coefficient of } (-u)^r \text{ in } \Gamma\left(\frac{n+1}{2}\right)\left(1+\frac{t_1^2}{n}\right)^{-\frac{n+1}{2}} A_r^{(\frac{1}{2},\frac{n}{2})}(t_1)$$

where

$$A_r^{(\frac{1}{2},\frac{n}{2})}(t_1) = \text{coefficient of } (-u)^r \text{ in } (1-u)^{\frac{1}{2}} \cdot \left[1-\frac{u\frac{t_1^2}{n}}{(1+\frac{t_1^2}{n})}\right]^{-\frac{n+1}{2}}$$

$$(5.31)\qquad = \sum_{i=0}^{r} (-1)^i \binom{\frac{1}{2}}{r-i}\binom{r+\frac{n+1}{2}-1}{i}\left(\frac{t_1^2}{n}\right)^i\left(1+\frac{t_1^2}{n}\right)^{-i}.$$

Then we integrate (5.28) with respect to y_2 and the result is

$$(5.32)\qquad = \frac{\left[\Gamma(\frac{n+1}{2})\right]^2}{n\pi\left[\Gamma(\frac{n}{2})\right]^2}\left(1+\frac{t_1^2}{n}\right)^{-\frac{n+1}{2}}\left(1+\frac{t_2^2}{n}\right)^{-\frac{n+1}{2}} \cdot \sum_{r=0}^{\infty}\frac{\Gamma(\frac{n}{2})\,r!}{\Gamma(\frac{n}{2}+r)}\rho^r A_r^{(\frac{1}{2},\frac{n}{2})}(t_1) A_r^{(\frac{1}{2},\frac{n}{2})}(t_2).$$

Hence,

$$(5.33) \qquad f(t_1,t_2) = g(t_1)g(t_2) \sum_{r=0}^{\infty} \frac{r!\,\Gamma(\frac{n}{2})}{\Gamma(\frac{n}{2}+r)} \rho^r A_r^{(\frac{1}{2},\frac{n}{2})}(t_1) A_r^{(\frac{1}{2},\frac{n}{2})}(t_2).$$

Unfortunately $A_r^{(\frac{1}{2},\frac{n}{2})}(t_1)$ is not a special case of the Jacobi polynomial

$$(5.34) \qquad P_r^{(\frac{n}{2}-1,\,-\frac{1}{2})}\left(\frac{t_1^2}{n}-1 \Big/ \frac{t_1^2}{n}+1\right) = \sum_{i=0}^{r} (-1)^i \binom{r+\frac{n}{2}-1}{r-i} \binom{r+\frac{1}{2}-1}{i} \left(\frac{t_1^2}{n}\right)^{r-i} \left(\frac{t_1^2}{n}+1\right)^r$$

which is orthogonal on the marginal density function

$$g(t_1^2) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \qquad \frac{(t_1^2/n)^{-\frac{1}{2}}}{(1+\frac{t_1^2}{n})^{\frac{n+1}{2}}} \,.$$

If we let x_2 and y_2 at the biginning of the Proof a pair of gamma variable with different parameters n/2 and m/2 respectivwly, the joint probability element (5.25) becomes equal to (using equation 3.37 for the bivariate gamma distribution)

$$(5.35) \qquad \frac{1}{\pi\,\Gamma(\frac{n}{2})\,\Gamma(\frac{m}{2})}\, e^{-(x_1+x_2+y_1+y_2)}\, x_1^{-\frac{1}{2}}\, x_2^{\frac{n}{2}-1}\, y_1^{-\frac{1}{2}}\, y_2^{\frac{n}{2}-1}$$

$$\left[1+\sum_{r=1}^{\infty} \frac{r!\,\Gamma(p+r)}{\Gamma(p)} \frac{\Gamma(\frac{m}{2})\,\Gamma(\frac{n}{2})}{\Gamma(\frac{m}{2}+r)\,\Gamma(\frac{n}{2}+r)} \rho^{2r}\right.$$

$$\left. L_r^{(\frac{n}{2}-1)}(x_2)\, L_r^{(\frac{m}{2}-1)}(y_2)\right] dx_1\, dx_2\, dy_1\, dy_2.$$

If we follow the same proof we will get

$$f(t_1,t_2) = g(t_1)g(t_2)\left[1+\sum_{r=1}^{\infty}\frac{r!\,\Gamma(p+r)\,\Gamma(\frac{m}{2})\,\Gamma(\frac{n}{2})}{\Gamma(p)\,\Gamma(\frac{n}{2}+r)}\rho^{2r}\right. \tag{5.36}$$

$$\left. A_r^{(\frac{1}{2},\frac{n}{2})}(t_1)A_r^{(\frac{1}{2},\frac{m}{2})}(t_2)\right].$$

where

$A_r^{(\frac{1}{2},\frac{n}{2})}(t)$ is given by (5.31) and $g(t_1)$, $g(t_2)$ are the marginal density functions with D.F. n an m respectively.

3. <u>Bivariate F Distribution</u>. Generalizing the work in the previous sections, we can derive a series form for the bivariate F distribution.

Let $x_1, \ldots, x_n$ and $y_1, \ldots y_m$ be independent samples of n and m from the normal distribution $0, \sigma^2$. We consider the following ratio of mean squares:

$$F = \frac{(x_1^2 + \ldots + x_n^2)/n}{(y_1^2 + \ldots + y_m^2)/m} \quad . \tag{5.37}$$

Letting X_1^2 and X_2^2 designate independent Chi-Square random variables with n and m D.F. respectively, we can represent the distribution of F by

$$F = \frac{\sigma^2 X_1^2 / n}{\sigma^2 X_2^2 / m} = \frac{X_1^2 / n}{X_2^2 / m} \tag{5.38}$$

from which it is apparent that the parameter σ^2 does not effect the distribution.

Because a X^2 variable with n D F is a gamma variable with parameter n/2 it follows that

$$(5.39) \qquad F = \frac{\Gamma_{n/2}/n}{\Gamma_{m/2}/m} \quad \text{or} \quad G = \frac{n}{m} F = \frac{\Gamma_{n/2}}{\Gamma_{m/2}}$$

with probability element.

$$(5.40) \qquad d\,K(G) = \frac{\Gamma(\frac{n+m}{2})}{\Gamma(\frac{n}{2})\,\Gamma(\frac{m}{2})} \;\frac{G^{n/2-1}}{(1+G)^{\frac{m+n}{2}}}\; dG.$$

If we let x_1, x_2, y_1, y_2 be the same variables as in the previous section with parameters $n_1/2$, $m_1/2$, $n_2/2$ and $m_2/2$ respectively. The joint probability element of these variates is

$$(5.41) \qquad \frac{1}{\Gamma(\frac{n_1}{2})\,\Gamma(\frac{m_1}{2})\,\Gamma(\frac{n_2}{2})\,\Gamma(\frac{m_2}{2})}\, e^{-(x_1+x_2+y_1+y_2)}\, x_1^{\frac{n_1}{2}-1}\, x_2^{\frac{m_1}{2}-1}$$

$$\cdot\, y_1^{\frac{n_2}{2}-1}\, y_2^{\frac{m_2}{2}-1}\Big[1+\sum_{r=1}^{\infty}\frac{r!\,\Gamma(p+r)\,\Gamma(\frac{m_1}{2})\,\Gamma(\frac{m_2}{2})}{\Gamma(p)\,\Gamma(\frac{m_1}{2}+r)\,\Gamma(\frac{m_2}{2}+r)}\,\rho^{2r}$$

$$L_r^{(\frac{m_1}{2}-1)}(x_2)\; L_r^{(\frac{m_2}{2}-1)}(y_2)\Big]\, dx_1\, dx_2\, dy_1\, dy_2.$$

Make the following transformation

$$G_1 = \frac{n_1}{m_1} F_1 = \frac{x_1}{x_2}, \quad x_2 = x_2$$

$$(5.45) \qquad G_2 = \frac{n_2}{m_2} F_2 = \frac{y_1}{y_2}, \quad y_2 = y_2$$

and

$$\frac{\partial x_1}{\partial G_1}\;\frac{\partial y_1}{\partial G_2} = x_2\, y_2\,.$$

The joint probability element becomes

(5.43)
$$= \frac{(G_1)^{\frac{n_1}{2}-1}(G_2)^{\frac{n_2}{2}-1}}{\Gamma(\frac{n_1}{2})\,\Gamma(\frac{n_2}{2})}\, e^{-\left[x_2(1+G_1)+y_2(1+G_2)\right]}\; x_2^{\frac{n_1+m_1}{2}-1}\, y_2^{\frac{n_2+m_2}{2}-1}$$

$$\left[1+\sum_{r=1}^{\infty}\frac{r!\,\Gamma(p+r)}{\Gamma(\frac{m_1}{2}+r)\,\Gamma(\frac{m_2}{2}+r)}\,\rho^{r}\, L_r^{(\frac{m_1}{2}-1)}(x_2)\, L_r^{(\frac{m_2}{2}-1)}(y_2)\right]$$

$$\cdot dG_1\, dG_2\, dx_2\, dy_2.$$

We integrate with respect to x_2 (leaving out the terms without x_2) we get

(5.44)
$$\int_0^{\infty} e^{-x_2(1+G_1)}\, x_2^{\frac{n_1+m_1}{2}-1}\, L_r^{(\frac{m_1}{2}-1)}(x_2)\, dx_2$$

$$= \text{coefficient of } (-t)^r \text{ in}$$

$$(1-t)^{-\frac{m_1}{2}}\int_0^{\infty}\exp\left[-x_2(1+G_1-t\,G_1)\,/\,(1-t)\right] x_2^{\frac{n_1+m_1}{2}-1}\, dx_2.$$

$$= \text{coefficient of } (-t)^r \text{ in } \Gamma(\frac{n_1+m_1}{2})(1-t)^{\frac{n_1}{2}}$$

$$(1+G_1)^{-\frac{n_1+m_1}{2}}\left(1-\frac{tG_1}{1+G_1}\right)^{-\frac{n_1+m_1}{2}}$$

(5.45)
$$= \text{coefficient of } (-t)^r \text{ in } \Gamma(\frac{n_1+m_1}{2})\,(1+G_1)^{-\frac{n_1+m_1}{2}}\, B_r^{(\frac{n_1}{2},\frac{m_1}{2})}(G_1)$$

where

(5.46)
$$= B_r^{(\frac{n_1}{2},\frac{m_1}{2})}(G_1) = \text{coefficient of } (-t)^r \text{ in}$$

$$\left(1-\frac{tG_1}{1+G_1}\right)^{-\frac{n_1+m_1}{2}}(1-t)^{n_1/2}$$

$$(5.47) \qquad = \sum_i \binom{r + \frac{n_1 + m_1}{2} - 1}{i} \left(\frac{G_1}{1 + G_1}\right)^i \binom{\frac{n_1}{2}}{r-i}.$$

Similarly, we integrate with respect to y_2. Hence (5.43) becomes

$$(5.48) \qquad = \frac{G_1^{\frac{n_1}{2}-1} G_2^{\frac{n_2}{2}-1} \Gamma(\frac{n_1+m_1}{2}) \Gamma(\frac{n_2+m_2}{2})}{\Gamma(\frac{n_1}{2}) \Gamma(\frac{n_2}{2}) (1+G_1)^{\frac{n_1+m_1}{2}} (1+G_2)^{\frac{n_2+m_2}{2}}}$$

$$\cdot \; 1 + \sum_{r=1}^{\infty} \frac{r! \, \Gamma(p+r)}{\Gamma(\frac{m_1}{2}+r) \Gamma(\frac{n_2}{2}+r) \Gamma(p)} \rho^{2r} \, B_r^{(\frac{n_1}{2}, \frac{m_1}{2})}(G_1) \cdot B_r^{(\frac{n_2}{2}, \frac{m_2}{2})}(G_2)$$

$$(5.49) \qquad = k(G_1)\, k(G_2) \left[1 + \sum_{r=1}^{\infty} \frac{r! \, \Gamma(p+r) \Gamma(\frac{m_1}{2}) \Gamma(\frac{m_2}{2})}{\Gamma(p) \Gamma(\frac{m_1}{2}+r) \Gamma(\frac{m_2}{2}+r)} \rho^{2r} \right.$$

$$\left. B_r^{(\frac{n_1}{2}, \frac{m_1}{2})}(G_1) \, B_r^{(\frac{n_2}{2}, \frac{m_2}{2})}(G_2) \right].$$

REFERENCES

1. Aitken, A.C. and Gonin, H.T. (1935). "Fourfold sampling with and without replacement." Proc. Roy. Soc. Edinburgh, V.55 114 - 125.

2. Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc., New York, London, Sydney.

3. Barret, J.F. and Lampard, D.G. (1955). "An expansion for some second-order Probability diestributions and its application to noise Problems". IRE Trans. of the Professional Group on Information Theory, I - 1, 10 - 15.

4. Campbell, J.T. (1932). "Factorial moments and frequencies of Charlier's polynomials of type B." Proc. Edinburgh Math. Soc. (ser 2), V. 3, 99 - 106.

5. Campbell, J.T. (1934). "The Poisson Correlation Function" Proc. Edinburgh Math. Soc. (ser. 2), V. 4, 18 - 26.

6. Cherian, K.C. (1941). "A bivariate Correlated Gamma-Type Distribution Function." J. Indian Math. Soc. (N.S.), V. 5, 133-144.

7. Fisher, R.A. (1940) "The precision of discriminant functions". Annals of Eugenice, V. 10, 422 - 429.

8. Hamdan, M.A. (1963), Some Statistical Applications of Orthogonal Polynomials and Orthonormal Function. Unpublished Ph.D. Dissertation, University of Sydney.

9. Hamdan, M.A. (1963), "The Number and Width of Classes in the Chi-Square Test". J. of The American Statistical Association, V. 58, 678 - 689.

10. Hardy, G.H. (1932), "Summation of a series of polynomials of Laguerre." J. London Math. Soc., V. 7, 138 - 139.

11. Hille, E. (1926), "On Laguerre's series, Proc. Nat. Acad. Sci. V. 12, 261 - 269, 348 - 352.

12. Jackson, D. (1941). Fourier Series and Orthogonal Polynomials, Published by The Mathematical Association of America.

13. Kendall, M.G. (1941). "Proof of relations connected with the tetrachonic series and its generalization." Biomettika, V. 32, 196 - 198.

14. Kendall, M.G. and Stuart, A. (1958). The Advanced Theory of Statistics, Vol. 1. Hafner Publishing Co., New York.

15. Kendall, M.G. and Stuart, A. (1961). The Advanced Theory of Statistics, Vol. 2. Charles Griffin and Co. Ltd., London.

16. Kibble, F.W. (1941). "A two-variate gamma-type distribution". Sankhya, V. 5, 137 - 150.

17. Krawtchouk, M. (1929). "Sur une generalisation des polynomes d'Hermite." Comptes Rendus de l'Academie des Sciences, V. 189, 620-622.

18. Krishnariah, P.R. and Rao, M.M. (1961). "Remarks on a multivariate gamma distribution." Amer. Math. Monthly, V. 68, 342-346.

19. Krishnamoorthy, A.S. and Parthasarathy, M. (1951). "A multivariate gamma type distribution." Ann. Math. Stat. , V. 22, 549 - 557.

20. Lancaster, H.O. (1957). "Some properties of the bivariate normal distribution considered in the form of a contingency table." Biometrika, V. 44, 289 - 292.

21. Lancaster, H.O. (1958). "The Structure of bivariate distributions." Ann. Math. Stat. , V. 29, 719 - 736.

22. Lancaster, H.O., and Hamdan, M.A. (1964), "Estimation of The Correlation Coefficient In Contingency Tables with Possibly Nonmetrical Characters." Psychometrika, V. 29, No. 4, 383 - 391.

23. Leipnik, R. (1958). "The effect of instantaneous nonlinear devices on cross-correlation." IRE Trans. of the Professional Group on Information Theory. IT - 4, 73 - 76.

24- Maung, K. (1941). "Measurement of association in a contingency table with special references to the pigmentation of hair and eye colours of Scoottish school children." Ann. Eugenics, V. 11, 189 - 223.

25. McKendrick, A.G. (1926). "Application of Mathematics to medial Problems". Proc. Edinburgh Math. Soc, V. 44, 106 - 130.

26. Mehler, F.G. (1866). "Uber die Entwicklung einer Function von beliebig vielen variablen neach Laplaceschen Funktionen hoherer Ordnung." J. Fur Mathematik, V. 66, 161 - 176. Ibid.

27. Milne-Thomson (1933). The Calculus of finite differences, London, Macmillan.

28. Pearson, K. (1900). "Mathematical Contribution to the theory of evelution VII: on the correlation of characters not quantitatively measurable." Philos. Trans. Roy. Soc. London., V. 195 A, 1 - 47.

29. Pearson, K. (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have a risen from random sampling." Phil. Mag. V. 50, 157 - 175.

30. Pearson, K. (1904). "Mathematical Contribution to the theory of evolution XIII: on the theory of contingency and its relation to association and normal correlation." Drapers Company Research Memoirs, Biometric Series No.1.

31. Rice, S.O. (1945). "Mathematical analysis of random noise, III: Statistical Properties of random noise currents." Bell System Tech. J., V. 24, 48 - 156.

32. Siddiqui, M.M. (1967). "Bivariate t Distribution". The Annals of Mathematical Statistics, V. 38, 162- 166.

33. Szegö, G. (1939). Orthogonal Polynomials. Colloquim Publ. No. 23, Amer. Math. Soc.

34. Watson, G.N. (1933). "Notes on generating functions of polynomials: (1) Laguerre polynomials." J. London Math. Soc., V. 8, 189 - 192.

35. Watson, G.N. (1933). "Notes on generating functions of polynomials: (2) Hermite polynomials." J. London Math. Soc. V. 8, 194-199.

36. Watson, G.N. (1933). "Notes on generating functions of polynomials: (3) Polynomials of Legendre and Gegenbauer." J. London Math. Soc., V. 8, 284 - 292.

37. Wicksell, S.D. (1916). "Some theorems in the theory of probability". Suenska Aktuarie foreningers Tidskr., 165 - 213.

38. Wicksell, S.D. (1933). "On correlation functions of type III." Biometrika, V. 25, 121 - 133.

39. Williams, E.J.(1952). "Use of Scores for the analysis of association in contingency tables." Biometrika, V.39, 274-289.