ORTHOGONAL PARAMETERS FOR
TWO PARAMETER DISTRIBUTIONS

by
John W. Philpot

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of

MASTER OF SCIENCE
in
Statistics

July 1964

Blacksburg, Virginia
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I.</strong> INTRODUCTION</td>
<td>4</td>
</tr>
<tr>
<td>1.1 The General Method</td>
<td>8</td>
</tr>
<tr>
<td>1.2 Parameters of Location and Scale</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Distributions Admitting Jointly</td>
<td>11</td>
</tr>
<tr>
<td><strong>1.4 Limitations of the Method</strong></td>
<td>13</td>
</tr>
<tr>
<td><strong>II.</strong> AN ALTERNATE METHOD</td>
<td>15</td>
</tr>
<tr>
<td>2.1 Derivation</td>
<td>15</td>
</tr>
<tr>
<td>2.2 Application to the Neyman type A</td>
<td>18</td>
</tr>
<tr>
<td>2.3 Limitations of the Method</td>
<td>22</td>
</tr>
<tr>
<td>2.3.1 Expected value of the second derivative</td>
<td>22</td>
</tr>
<tr>
<td>2.3.2 Functional form of $\beta_1$ and $\beta_2$</td>
<td>25</td>
</tr>
<tr>
<td><strong>III.</strong> RESULTS</td>
<td>29</td>
</tr>
<tr>
<td>3.1 Common Characteristics</td>
<td>29</td>
</tr>
<tr>
<td>3.1.1 The class of contagious distributions</td>
<td>33</td>
</tr>
<tr>
<td><strong>IV.</strong> USES OF THE COMMON CHARACTERISTICS</td>
<td>36</td>
</tr>
<tr>
<td>4.1 A Simplified Method for finding Orthogonal Parameters</td>
<td>36</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Applications of the Simplified Method</td>
<td>37</td>
</tr>
<tr>
<td>4.2.1 Neyman type A distribution</td>
<td>37</td>
</tr>
<tr>
<td>4.2.2 Gram-Charlier Negative Exponential Distribution</td>
<td>38</td>
</tr>
<tr>
<td>4.3 Further Applications of the General Method</td>
<td>39</td>
</tr>
<tr>
<td>4.3.1 Neyman type A distribution</td>
<td>39</td>
</tr>
<tr>
<td>4.3.2 Gram-Charlier Negative Exponential Distribution</td>
<td>41</td>
</tr>
<tr>
<td>V. DISTRIBUTIONS WITH MORE THAN TWO PARAMETERS</td>
<td>43</td>
</tr>
<tr>
<td>5.1 Distributions with Three Parameters</td>
<td>43</td>
</tr>
<tr>
<td>5.2 Breakdown of the Alternate Method</td>
<td>45</td>
</tr>
<tr>
<td>5.3 A Class of Distributions Having Orthogonal Parameters</td>
<td>46</td>
</tr>
<tr>
<td>VI. CONCLUSIONS</td>
<td>49</td>
</tr>
<tr>
<td>VII. ACKNOWLEDGEMENTS</td>
<td>53</td>
</tr>
<tr>
<td>VIII. BIBLIOGRAPHY</td>
<td>54</td>
</tr>
<tr>
<td>IX. VITA</td>
<td>55</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>56</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

In the problem of estimating the parameters of a probability distribution, the maximum likelihood procedure is an integral step in the modus operandi of the research worker. Many aspects of this method have been investigated in the literature, providing a firm theoretical justification for the estimates. But although the maximum likelihood procedure is almost always applied, the equations that result are often complex, and solutions may depend on iterative methods. Such likelihood equations often serve only as a practical justification for using other methods of estimation. It appears, then, that a technique to simplify the likelihood equations would have considerable merit. Jeffreys [8] in 1948 proposed orthogonal parameters as a possible solution to this problem.

To define orthogonal parameters, let \( f(x, \beta_i) \) be a probability distribution having \( n \) parameters, \( \beta_i (i = 1, 2, \ldots, n) \). Then the parameters \( \beta_i \) are said to be orthogonal if

\[
E\left\{ -\frac{\partial^2}{\partial \beta_i \partial \beta_j} \log f \right\} = 0 \quad \text{for all } i, j \ (i \neq j) \quad (1).
\]

In general the set \( \{\beta_i\} \) satisfying (1) represents \( n \) independent functions of the original set of parameters \( \{a_i\} \), where the distribution was originally defined as
f(x, a_i), (i = 1, 2, ..., n). The importance of orthogonal parameters lies in the fact that the maximum likelihood equations to estimate them are nearly orthogonal. Hence one might expect any iteration process associated with the likelihood equations to be considerably simplified.

Huzurbazar [6] investigated the general problem of determining orthogonal parameters for a probability distribution, and was successful in developing a procedure for two-parameter distributions. He also showed that orthogonal parameters may always be found when one of the original parameters is of location or scale, location and scale being defined according to Fisher's usage [5]. Furthermore, he pointed out that, for n parameters, the orthogonality condition gives \( \frac{n}{2}(n-1) \) equations. Thus when \( n = 2 \), there is only one restriction; \( n = 3 \) gives three constraints; and for \( n \geq 4 \) there are more constraints than unknowns. This means that theoretically there are an infinite set of orthogonal parameters when \( n = 2 \); for \( n = 3 \) the set is finite; and for \( n \geq 4 \) no solution is possible in general. It may be remarked here that even when \( n = 2 \) we may be unable to find orthogonal parameters that are other than transformations of a basic pair.

The theory was extended by the same author [7] when he showed that, for two-parameter distributions admitting jointly sufficient statistics, the problem of finding
orthogonal parameters is solvable.

A further aspect of orthogonal parameters is that the associated likelihood contours are nearly elliptical, with axes parallel, in the case of two parameters, to $\beta_1 = 0$ and $\beta_2 = 0$ (where $\beta_1$ and $\beta_2$ are orthogonal). If one plots such contours it is a simple matter to obtain a close estimate of the maximum likelihood solution by graphical methods. The original parameters may have likelihood contours that are far from elliptical in shape. This point was brought out by Anscombe [1].

In addition, when the parameters are orthogonal, the information matrix for finding asymptotic variances is reduced to diagonal terms. That the off-diagonal terms are zero is obvious from the definition of orthogonality. The definition is equivalent to stating that the maximum likelihood estimates of the parameters are asymptotically uncorrelated.

An optimistic interpretation of the results would lead one to believe that use of orthogonal parameters combines the accuracy of the maximum likelihood technique with the simplicity of the method of moments. Still, orthogonal parameters have not played even a minor role in practical applications. They have been used in only one case in the literature thus far, when Anscombe [1] applied them to an exponential type law.
Yet there exists an important class of distributions, the contagious distributions, whose maximum likelihood equations are often of a complex form, and whose solutions require iterative techniques. A number of biological applications have been found for these distributions in recent years. An incomplete but representative listing of two-parameter cases includes the Neyman type A, (Douglas [2], Evans [3]), the Poisson-Binomial (McCuirre, et. al. [10], Shumway and Gurland [11]) and the Logarithmic Binomial (Sprott [12]). Moreover, it was pointed out by Feller, [4], that these distributions have some applicability to such diverse topics as accident statistics, telephone traffic, fire damage, sickness and life insurance, and risk theory. It is evident that a procedure for simplifying the maximum likelihood equations in these cases would have some value. Unfortunately, direct application of Huzurbazar's general method to many of the contagious distributions leads to differential equations for determining the orthogonal parameters which are not readily solvable.

Thus the purposes of this paper are: first, to find a method for determining orthogonal parameters for the class of two-parameter distributions typified by the contagious distributions; second, to investigate the common characteristics of this class of distributions; and third, to assess the usefulness of the orthogonal parameters found.
Two-parameter distributions are considered most fully since they seem to be in the most fertile area. However a few observations are directed specifically toward distributions having more than two parameters.

1.1 The General Method.

The development of the method due to Huzurbazar [6] for obtaining orthogonal parameters for two-parameter distributions may be summarized briefly. Let \( \{a_1, a_2\} \) be the original parameters of a probability distribution \( f(x) \), and \( \{\theta_1, \theta_2\} \) the orthogonal parameters. Then set

\[
\theta_1 = a_1
\]

The orthogonality condition requires

\[
E\left(\frac{\partial^2 \log f}{\partial a_1 \partial a_2}\right) \frac{\partial a_2}{\partial \theta_2} + E\left(\frac{\partial^2 \log f}{\partial a_2^2}\right) \frac{\partial a_2}{\partial \theta_2} \cdot \frac{\partial a_2}{\partial \theta_2} = 0 . \tag{2}
\]

Now \( a_2 \) must necessarily depend on \( \theta_2 \), since \( \theta_1 \) and \( \theta_2 \) are independent. Therefore

\[
\frac{\partial a_2}{\partial \theta_2} \neq 0 ,
\]

and equation (2) becomes

\[
E\left(\frac{\partial^2 \log f}{\partial a_1 \partial a_2}\right) \frac{\partial a_2}{\partial \theta_2} + E\left(\frac{\partial^2 \log f}{\partial a_2^2}\right) \frac{\partial a_2}{\partial \theta_2} = 0 . \tag{3}
\]

The solution of equation (3) involves an arbitrary constant which can be set equal to a function of \( \theta_2 \), the second orthogonal parameter.
Thus the pair of orthogonal parameters is given by

\[ \beta_1 = \alpha_1 \]

and the solution of equation (3). It is clear that in the theory the roles of \( \alpha_1 \) and \( \alpha_2 \) are interchangeable. Equation (3) is not always solvable, but for many well-known distributions the variables can be separated and \( \beta_2 \) determined.

1.2 Parameters of Location and Scale.

Following Huzurbazar [6], let us consider finding a pair of orthogonal parameters in the case where the two original parameters are of location and scale, defined in accordance with Fisher's usage [5]. Let

\[ f(x, \alpha_1, \alpha_2) = \frac{1}{\alpha_2} e^{-\frac{x-\alpha_1}{\alpha_2}}, \]

where the location and scale parameters are \( \alpha_1 \) and \( \alpha_2 \) respectively. Writing \( \xi = \frac{x-\alpha_1}{\alpha_2} \), we obtain

\[ E \frac{\partial^2 \log f}{\partial \alpha_1^2} = E_{11} = \frac{k_{11}}{\alpha_2^2}, \]

\[ E \frac{\partial^2 \log f}{\partial \alpha_1 \partial \alpha_2} = E_{12} = \frac{k_{12}}{\alpha_2^2}, \]

and
\[
E \frac{z^2 \log f}{\partial a_2^2} = E_{22} = \frac{k_{22}}{a_2^2},
\]

where
\[
k_{11} = E\{\theta^n(\xi)\},
\]
\[
k_{12} = E\{\xi \theta^n(\xi)\},
\]
and
\[
k_{22} = E\{1 + 2\xi \theta'(\xi) + \xi^2 \theta''(\xi)\}.
\]

Now the \( k \)'s are independent of \( a_1 \) and \( a_2 \). Choosing \( a_2 = \beta_2 \), the general method leads to
\[
k_{12} + k_{11} \frac{2a_1}{\beta_2} = 0,
\]
which gives
\[
a_1 = \beta_1 + k \beta_2, \quad \text{where} \quad k = -\frac{k_{12}}{k_{11}}.
\]

If, instead, we choose \( a_1 = \beta_1 \), we obtain
\[
a_2 = k' \beta_1 + \beta_2, \quad \text{where} \quad k' = -\frac{k_{12}}{k_{22}}.
\]

We see that the parameters of location and scale can be made orthogonal by fixing either of them, and by a linear transformation of the other.

The general method can be applied also to the case where only one of the parameters is of location or scale and the other is a general one. We have
1.3 Distributions Admitting Jointly Sufficient Statistics.

We now consider the relationship between sufficient statistics and orthogonal parameters. The most general form of two-parameter distributions admitting sufficient statistics (Koopman [9]) is

\[ f(x, a_1, a_2) = \exp\{u_1(a_1, a_2) v_1(x) + u_2(a_1, a_2) v_2(x) + A(x) + B(a_1, a_2)\}. \]

We may take \( u_1 \) and \( u_2 \) as new parameters and put \( B \) in terms of \( u_1 \) and \( u_2 \), so that the distribution becomes

\[ f(x, u_1, u_2) = \exp\{u_1 v_1(x) + u_2 v_2(x) + A(x) + B(u_1, u_2)\}. \]

We find
Applying the general method for determining orthogonal parameters, we set \( u_1 = \beta_1 \). Then equation (3) gives

\[
\frac{\partial^2 B}{\partial \beta_1 \partial u_2} + \frac{\partial^2 B}{\partial u_2^2} \cdot \frac{\partial u_2}{\partial \beta_1} = 0.
\]

Now since \( B \) is a function of \( \beta_1 \) and \( u_2 \), where \( u_2 = u_2(\beta_1, \beta_2) \), the above equation can be written as

\[
\frac{\partial}{\partial \beta_1} \left\{ \frac{\partial B}{\partial u_2} \right\} = 0,
\]

which gives

\[
\frac{\partial B}{\partial u_2} = \psi(\beta_2),
\]

where \( \psi(\beta_2) \) is an arbitrary function of \( \beta_2 \).

If we take, for simplicity, \( \psi(\beta_2) = \beta_2 \), our orthogonal parameters are

\[
\beta_1 = u_1,
\]

and

\[
\beta_2 = \frac{\partial B}{\partial u_2}.
\]
In a similar manner it can be shown that

$$\beta_1 = \frac{\beta}{\mu_1}$$

and

$$\beta_2 = \mu_2$$

are also orthogonal.

1.4 Limitations of the Method.

As noted before, obtaining the second orthogonal parameter, \(\beta_2\), using the general method given above hinges on the solution of equation (3). This equation is equivalent to a first order differential equation with variable coefficients. If it is of a type not amenable to a classical form of solution, the second orthogonal parameter cannot be determined from it.

For example, let us consider the Neyman type A distribution, given by

$$P_x = \frac{e^{-m_1} m_2^x}{x!} \sum_{k=0}^{\infty} \frac{(m_1 e^{-m_2})^k}{k!}, \quad x = 0, 1, \ldots, \quad (4)$$

= 0 otherwise.

If we take \(\beta_1 = m_1\), then we find

$$E\left(\frac{3^2 \log P_x}{3m_1^2 3m_2^2}\right) = -(1+m_1) + \frac{1}{m_1 m_2}$$

and
\[
E\left(\frac{\log P}{\frac{m_2}{\partial m_2}}\right) = -\frac{m_1}{m_2} + m_1(1+m_1) - \frac{1}{m_2},
\]

where

\[
\phi = E\left(\frac{(x+1) P_{x+1}}{P_x}\right) \tag{5}
\]

Equation (3) becomes

\[
\frac{1}{m_2^2} \left\{ \frac{1}{\beta_1} - \frac{\partial m_2}{\partial \beta_1} \right\} \phi - \left\{ (1+\beta_1) + \beta_1 \left[ \frac{1}{m_2} - (1+\beta_1) \frac{\partial m_2}{\partial \beta_1} \right] \right\} = 0. \tag{6}
\]

Since in general \( \phi \) is a complicated function, and there is no common factor in the two terms of equation (6) which can be set equal to zero, this method breaks down. A need for another approach to the problem is indicated.
II. AN ALTERNATE METHOD

2.1 Derivation.

For the two-parameter case, \( f(x,a_1,a_2) \), the likelihood expressions can be written in terms of the orthogonal parameters \( \beta_1 \) and \( \beta_2 \), as

\[
\frac{\partial \log f}{\partial a_1} = \frac{\partial \log f}{\partial \beta_1} \cdot \frac{\beta_1}{a_1} + \frac{\partial \log f}{\partial \beta_2} \cdot \frac{\beta_2}{a_1}, \quad (7)
\]

and

\[
\frac{\partial \log f}{\partial a_2} = \frac{\partial \log f}{\partial \beta_1} \cdot \frac{\beta_1}{a_2} + \frac{\partial \log f}{\partial \beta_2} \cdot \frac{\beta_2}{a_2}. \quad (8)
\]

Squaring equation (7) and taking expectations leads to

\[
E\left( \frac{\partial^2 \log f}{\partial a_1^2} \right) + (\frac{\partial \beta_1}{\partial a_1})^2 E\left( -\frac{\partial^2 \log f}{\partial \beta_1^2} \right) + (\frac{\partial \beta_2}{\partial a_1})^2 E\left( -\frac{\partial^2 \log f}{\partial \beta_2^2} \right) = 0. \quad (9)
\]

The expectation of the product of equations (7) and (8) gives

\[
E\left( \frac{\partial^2 \log f}{\partial a_1 \partial a_2} \right) + \frac{\partial \beta_1}{\partial a_1} \frac{\partial \beta_1}{\partial a_2} E\left( -\frac{\partial^2 \log f}{\partial \beta_1^2} \right) + \frac{\partial \beta_2}{\partial a_1} \frac{\partial \beta_2}{\partial a_2} E\left( -\frac{\partial^2 \log f}{\partial \beta_2^2} \right) = 0. \quad (10)
\]
Finally, squaring equation (8) and taking expectations gives

\[
E\left(\frac{\partial^2 \log f}{\partial \beta_1^2}\right) + \left(\frac{\partial^2 \log f}{\partial \beta_2^2}\right) E(-\frac{\partial^2 \log f}{\partial \beta_1^2})
+ (\frac{\partial^2 \log f}{\partial \beta_2^2})^2 E(-\frac{\partial^2 \log f}{\partial \beta_2^2}) = 0.
\]

(11)

It is apparent from equations (9), (10), and (11) that if we can evaluate the first term on the left-hand side of each equation, and then consider the remaining expectation terms as our unknowns, we will have three equations and two unknowns. For the equations to be consistent the determinant of coefficients must equal zero.

Since our aim is to develop orthogonal parameters for distributions like the Neyman type A, we must assume that

\[
E\left(\frac{\partial^2 \log f}{\partial \beta_1 \partial \beta_2}\right) = C_1 + K_1 \phi,
\]

\[
E\left(\frac{\partial^2 \log f}{\partial \beta_1^2}\right) = C_2 + K_2 \phi,
\]

and

\[
E\left(\frac{\partial^2 \log f}{\partial \beta_2^2}\right) = C_3 + K_3 \phi.
\]

(12)

where \( \phi \) is an expectation term which is often complicated structurally, as in equation (5).
Therefore the determinant of coefficients from equations (9), (10), and (11), becomes

\[
\begin{vmatrix}
C_1 + K_1 & \frac{\alpha^2_1}{\alpha_1} & \frac{\alpha^2_2}{\alpha_1} \\
C_2 + K_2 & \frac{\alpha^2_1}{\alpha_2} + \frac{\alpha^2_2}{\alpha_2} \\
C_3 + K_3 & \frac{\alpha^2_1}{\alpha_2} & \frac{\alpha^2_2}{\alpha_2}
\end{vmatrix}
= 0.
\]

Assuming that \(\frac{\alpha^2_1}{\alpha_1} \neq 0\) and \(\frac{\alpha^2_2}{\alpha_1} \neq 0\), we can divide each term in the second and third columns by their first terms. Then expanding the determinant in equation (13) we obtain

\[
\begin{vmatrix}
C_1 & 1 & 1 \\
C_2 & x & y \\
C_3 & x^2 & y^2
\end{vmatrix}
+ \begin{vmatrix}
K_1 & 1 & 1 \\
K_2 & x & y \\
K_3 & x^2 & y^2
\end{vmatrix}
= 0.
\]

where

\[
x = \frac{\alpha^2_1/\alpha_2}{\alpha^2_1/\alpha_1},
\]

and

\[
y = \frac{\alpha^2_2/\alpha_2}{\alpha^2_2/\alpha_1}.
\]

One solution to equation (14) is to set both determinants equal to zero. This leads to a pair of equations
in $x$ and $y$. The resulting equations are

$$C_1(xy) - C_2(x+y) + C_3 = 0 \quad , \quad (15a)$$

and

$$K_1(xy) - K_2(x+y) + K_3 = 0 \quad . \quad (15b)$$

These can be combined to give the quadratic equation:

$$(C_1K_2 - C_2K_1)z^2 + (C_3K_1 - C_1K_3)z + (C_2K_3 - C_3K_2) = 0 \quad .$$

If the two roots of this equation are $z_1$ and $z_2$, we can let

$$z_1 = x = \frac{\partial y_1/\partial a_2}{\partial y_1/\partial a_1} \quad ,$$

and

$$z_2 = y = \frac{\partial y_2/\partial a_2}{\partial y_2/\partial a_1} \quad ,$$

which can be evaluated to obtain $\beta_1$ and $\beta_2$.

2.2 Application to the Neyman type A.

To illustrate the alternate method, consider the Neyman type A distribution defined by equation (14). The first objective is to evaluate

$$E\left(\frac{z^2 \log P}{\partial m_1}\right) \quad , \quad (16a)$$

$$E\left(\frac{z^2 \log P}{\partial m_1 \partial m_2}\right) \quad , \quad (16b)$$

and
To do so we make use of the frequency generating function

\[ G(t) = \sum_x p_x t^x = \exp(m_1 M_2) \]  

(17)

where \( M_2 = \exp(m_2(t-1)) - 1 \).

Note also that when the range is independent of the parameter we have

\[ -E\left( \frac{\partial^2 \log P}{\partial m_2^2} \right) = E\left( \frac{1}{P} \frac{\partial P}{\partial m_1} \right)^2. \]

Differentiating \( G(t) \) with respect to \( m_1 \), and then with respect to \( t \), we have

\[ \frac{\partial G(t)}{\partial m_1} = \left\{ -1 + e^{m_2(t-1)} \right\} \sum_x p_x t^x = \sum_x \frac{\partial p_x}{\partial m_1} t^x, \]  

(18)

and

\[ \frac{\partial G(t)}{\partial t} = m_1 m_2 e^{m_2(t-1)} \sum_x p_x t^x = \sum x t^{x-1} p_x. \]  

(19)

Multiplying equation (18) by \( m_1 m_2 \), and subtracting the result from equation (19) gives:

\[ m_1 m_2 \sum p_x t^x = -m_1 m_2 \sum \frac{\partial p_x}{\partial m_1} t^x + \sum x t^{x-1} p_x. \]

Equating coefficients of \( t^x \) leads to

\[ \frac{\partial p_x}{\partial m_1} = -p_x + \frac{(x+1)}{m_1 m_2} p_{x+1}. \]
Thus in evaluating (16a), we obtain

$$E\left(1 \frac{\partial P}{\partial m_1} \right)^2 = -1 + \frac{\phi}{m_1 m_2}$$

(20)

where $\phi = E\left(\frac{(x+1) P^x}{x} \right)^2$.

Similarly

$$-E\left(\frac{\partial^2 \log P}{\partial m_1 \partial m_2} \right) = E\left(1 \frac{\partial P}{\partial m_1} \frac{\partial P}{\partial m_2} \right) = 1 + m_1 - \frac{\phi}{m_1 m_2}$$

(21)

and

$$-E\left(\frac{\partial^2 \log P}{\partial m_2^2} \right) = E\left(1 \frac{\partial P}{\partial m_2} \right)^2 = \frac{m_1}{m_2} - m_1 (1+m_1) + \frac{\phi}{m_2}$$

(22)

The expanded determinant of coefficients corresponding to equation (14) becomes

$$\begin{vmatrix}
-1 & 1 & 1 \\
1 + m_1 & x & y \\
\frac{m_1}{m_2} - m_1 (1+m_1) & x^2 & y^2 \\
\end{vmatrix}$$

$$+ \phi \begin{vmatrix}
\frac{1}{m_1 m_2} & 1 & 1 \\
-\frac{1}{m_1 m_2} & x & y \\
\frac{1}{m_2} & x^2 & y^2 \\
\end{vmatrix} = 0$$

where

$$x = \frac{\partial \phi / \partial m_2}{\partial \phi / \partial m_1}$$

and
As in equations (15a) and (15b), we now obtain

\[ -l(xy) - (1+m_1)(x+y) + \left( \frac{m_1}{m_2} - m_1(1+m_1) \right) = 0 \quad , \quad (23) \]

and

\[ \frac{1}{m_2} (xy) + \frac{1}{m_1} (x+y) + 1 = 0 \quad . \quad (24) \]

Taking advantage of the binomial coefficients in equation (24), we factor it to obtain

\[ y = -m_1 \quad . \quad (25) \]

Using this in equation (23) gives

\[ x = \frac{m_1}{m_2} \quad . \quad (26) \]

These equations, (25) and (26), in turn give

\[ \frac{dm_1}{m_1} - dm_2 = 0 \quad , \]

and

\[ \frac{dm_1}{m_1} + \frac{dm_2}{m_2} = 0 \quad , \]

leading to the solution

\[ \beta_1 = m_1m_2 \quad \text{and} \quad \beta_2 = m_1e^{-m_2} \quad . \quad (27) \]

That \( \beta_1 \) and \( \beta_2 \) are in fact orthogonal parameters may be checked by substitution into equation (1).
2.3 Limitations of the Method.

2.3.1 Expected value of the second derivative.

We note first that application of the alternate method is restricted to two-parameter distributions, \( f(x, a_1, a_2) \), having terms \( \frac{\partial^2 \log f}{\partial a_1^2} \), \( \frac{\partial^2 \log f}{\partial a_1 \partial a_2} \), and \( \frac{\partial^2 \log f}{\partial a_2^2} \), that are of the general form \( C + K \), as in (12). Consequently the method applies to a limited class of two-parameter distributions.

As an example of a distribution beyond the scope of the alternate method (and the general method) for determining orthogonal parameters, consider a Gram-Charlier type-B distribution of the form

\[
p_x = \{1 + b \left( \frac{x(x-1)}{m^2} - \frac{2x}{m} + 1 \right) \} \frac{e^{-m} m^x}{x!} , \quad x = 0, 1, \ldots ,
\]

\( = 0 \) otherwise ,

where

\[
b = \frac{m^2}{2m[m+1] - m^2 - [m+1][m]} ,
\]

\( m > 0 \),

and \([a]\) represents the integer part of \( a \).

The probability generating function is

\[
p(t) = \{1 + b(t-1)^2\} e^{m(t-1)} .
\]

In applying the alternate method, our aim is to set up the system of equations corresponding to (9), (10), and (11).
Consequently we must determine the first term in each of these equations. We find, in the present case,

\[-E\left(\frac{2 \log P}{ab^2}\right) = -\frac{1}{b^2} + \frac{1}{b^2} \phi_0\quad \text{(28a)}\]

\[-E\left(\frac{2 \log P}{ab^2}\right) = \frac{2}{mb} - \frac{2(1+b)}{mb} \phi_0 + \frac{2}{m^2} \phi_1\quad \text{(28b)}\]

and

\[-E\left(\frac{2 \log P}{ab^2}\right) = \frac{1}{m} - \frac{2(2+b)}{m^2} + \frac{4(1+b)^2}{m^2} \phi_0
- \frac{8b(1+b)}{m^3} \phi_1 + \frac{4b^2}{m^4} \phi_2\quad \text{(28c)}\]

where

\[\phi_0 = E\left(\frac{x^2}{P^2}\right)\]

\[\phi_1 = E\left(\frac{x^2}{P^2}\right)\]

\[\phi_2 = E\left(\frac{x^2}{P^2}\right)\]

and

\[\psi_x = \frac{e^{-m} m^x}{x!}\]

A relationship can be found between the \(\phi\)'s, i.e.

\[\phi_2 = (1+2m)\phi_1 - \frac{m^2}{b} (1+b)\phi_0 + \frac{m^2}{b}\]
so that we can reduce the number of \( n \) terms in equations (28) from three to two.

Now if we write down the determinant of coefficients as in equation (13), and expand it, we are led to a system of three determinants:

\[
\begin{vmatrix}
C_1(b,m) & 1 & 1 \\
C_2(b,m) & x & y \\
C_3(b,m) & x_2 & y_2 \\
\end{vmatrix} + \phi_0 
\begin{vmatrix}
\gamma_1(b,m) & 1 & 1 \\
\gamma_2(b,m) & x & y \\
\gamma_3(b,m) & x^2 & y^2 \\
\end{vmatrix} + \phi_1 
\begin{vmatrix}
\epsilon_1(b,m) & 1 & 1 \\
\epsilon_2(b,m) & x & y \\
\epsilon_3(b,m) & x^2 & y^2 \\
\end{vmatrix} = 0.
\] (29)

No further reduction is possible, since any attempt to eliminate \( \phi_0 \), say, reintroduces \( \phi_2 \). By setting each determinant in equation (29) equal to zero, we obtain three constraints for only two unknowns. Thus the unknowns are overdetermined, and the system has no solution in general.

It is interesting to contrast the above failure with a successful application of the alternate method to another Gram-Charlier type-B distribution given by

\[
P_x = \{1 + a(\frac{x}{m} - 1)\} \frac{e^{-m} m^x}{x!}, \quad x = 0, 1, \ldots,
\] (30)

\[= 0 \text{ otherwise},
\]

where
In this case the expectation terms corresponding to those in equation (12) are:

\[ E \left( \frac{1}{p} \frac{\partial P}{\partial m} \right)^2 = \frac{m+a-1}{m^2} + \frac{(a-1)^2}{m^2} \phi^* , \]

\[ E \left( \frac{1}{p} \frac{\partial P}{\partial m} \frac{\partial P}{\partial a} \right) = \frac{1}{am} + \frac{(a-1)}{am} \phi^* , \]

and

\[ E \left( \frac{1}{p} \frac{\partial P}{\partial a} \right)^2 = -\frac{1}{a^2} + \frac{1}{a} \phi^* , \]

where

\[ \phi^* = E \left( \frac{\psi^2}{p_x} \right) , \]

and

\[ \psi_x = \frac{e^{-m}}{m^x} \frac{m^x}{x!} . \]

Thus we are able to find the orthogonal parameters

\[ \beta_1 = m + a \]

and

\[ \beta_2 = \frac{m(1-a)}{a} . \]

2.3.2 Functional form of \( \beta_1 \) and \( \beta_2 \).

If we look at the examples of orthogonal parameters given
in the tables, we note that in general $\beta_1$ and $\beta_2$ may be defined in terms of the original parameters $a_1$ and $a_2$, in four ways:

\[ \begin{array}{cc}
I & f_1(a_1, a_2) & f_2(a_1, a_2) \\
II & f_3(a_1) & f_4(a_1, a_2) \\
III & f_5(a_2) & f_6(a_1, a_2) \\
IV & f_7(a_1) & f_8(a_2) \\
\end{array} \]

(31)

A perfectly general method for determining orthogonal parameters will not restrict the functional form of the parameters; instead, any such restrictions should come from the distribution itself. Let us consider the alternate method in light of this.

If the only valid orthogonal parameters for a distribution have the functional form given by case I, we are assured of the existence of all terms in the determinant of coefficients given by equation (13) and so the alternate method would clearly apply.

If, however, we assume there are orthogonal parameters of the form given by II, we find that equation (13) becomes:
The fact that \( \frac{\partial \beta_1}{\partial a_1}^2 \) can be factored out of the equation leaves us with information only on \( \beta_2 \), and the two equations that result when the determinant of equation (32) is expanded will be inconsistent in general. The alternate method is clearly inadequate in this case. Similar remarks hold for case III.

Under IV, equation (13) will be:

\[
\begin{vmatrix}
C_1 + K_1 & \frac{\partial \beta_1}{\partial a_1}^2 & 0 \\
C_2 + K_2 & 0 & \frac{\partial \beta_2}{\partial a_1} \\
C_3 + K_3 & 0 & \frac{\partial \beta_2}{\partial a_2}^2
\end{vmatrix} = 0 .
\] (33)

While the alternate method is inapplicable, we can note that (33) implies

\[ C_2 + K_2 \neq 0 , \]

i.e., one can evaluate \( \phi \), and Huzurbazar's method can be used.

From these comments it is clear that the alternate
method is not generally applicable, and that when it is used to determine orthogonal parameters, the method will select only those that are each functions of both original parameters.
The alternate method for finding orthogonal parameters was applied to seven contagious distributions:

Heyman type A,
Poisson-Binomial,
Binomial-Poisson,
Geometric Poisson,
Logarithmic-Binomial,
Logarithmic-Poisson,
Normal Poisson,

as well as to three Gram-Charlier type distributions. The results of this analysis have been tabulated in the Appendix, along with a number of previously known results.

3.1 Common Characteristics.

On examining the tables in the Appendix, one of the first things to be noticed is that in every two-parameter case, one of the orthogonal parameters is the mean of the distribution. This point holds irrespective of the method used, and irrespective of whether the distribution is discrete or continuous.

Complementing this result is the fact that, for all two-parameter distributions tabulated, the first maximum likelihood equation takes the form
where

$$\beta_1 = \beta_1(a_1, a_2) = \mu_1$$

and

$$u_2 = \text{var}(x)$$.

How equation (34) is satisfied is not obvious in all cases, so we return to the Neyman type A to illustrate the point. The orthogonal parameters are known to be

$$\beta_1 = m_1 m_2$$

and

$$\beta_2 = m_1 e^{-m_2}$$.

We have

$$\frac{\partial \log P_x}{\partial \beta_1} = \frac{1}{P_x} \frac{\partial P_x}{\partial \beta_1} = \frac{1}{P_x} \left\{ \frac{\partial P_x}{\partial m_1} \frac{m_1}{\beta_1} + \frac{\partial P_x}{\partial m_2} \frac{m_2}{\beta_1} \right\}$$

which, in this case, becomes

$$\frac{\partial \log P_x}{\partial \beta_1} = \frac{1}{P_x} \left[ \left\{ -P_x + \frac{(x+1)}{m_1 m_2} P_{x+1} \right\} \frac{1}{(1+m_2)} \right.$$ 

$$+ \left\{ \frac{x}{m_2} P_x - \frac{(x+1)}{m_2} P_{x+1} \right\} \frac{1}{m_1(1+m_2)} \right.$$.

We see that the coefficient of $P_{x+1}$ is zero, and so we have

$$-\frac{1}{1+m_2} + \frac{x}{m_1 m_2} \frac{1}{(1+m_2)} = \frac{x - m_1 m_2}{m_1 m_2(1+m_2)}.$$
but since \( \text{var}(x) = m_1 m_2 (1 + m_2) \), we have achieved the form
\[
\frac{x - \beta_1}{\beta_2}
\]
as was to be shown.

Another important fact is that some of the contagious distributions listed cannot be written in terms of their orthogonal parameters, \( \beta_1 \) and \( \beta_2 \). In particular, where \( \beta_2 \) contains an exponential term, or a term raised to the n-th power, the contagious distribution could not be transformed into terms of \( \beta_1 \) and \( \beta_2 \).

Even where transformations can be found, as for the Gram-Charlier type distributions and some contagious distributions, they are of a type which add to the complexity of the probability law.

For example, the Gram-Charlier type B distribution given by equation (30) has orthogonal parameters
\[
\beta_1 = m + a,
\]
and
\[
\beta_2 = \frac{m(1-a)}{a}.
\]
These lead to the transformations
\[
a = \frac{(1 + \beta_1 + \beta_2) - \sqrt{(1 + \beta_1 + \beta_2)^2 - 4 \beta_1}}{2},
\]
and
\[
m = \frac{\beta_1 - (1 + \beta_2) + \sqrt{(1 + \beta_1 + \beta_2)^2 - 4 \beta_1}}{2}.
\]
Such transformations are not likely to be applied in practice. It can be seen that the usefulness of the orthogonal parameters is thereby restricted.

For the contagious distributions we note another common characteristic; i.e., each has a maximum likelihood equation of the form

$$ N \bar{x} = S_1 $$

(35)

where

$$ S_1 = \sum (x+1) \frac{P_{x+1}}{P_x} $$

This result can be illustrated by again considering the Neyman type A distribution given by equation (4). From

$$ \sum \frac{3 \log P}{\beta_1} = \sum \frac{x - \bar{\beta}_1}{\nu_2} $$

we have

$$ \hat{\beta}_1 = \bar{x} $$

Now the second likelihood equation can be written

$$ \sum \frac{3 \log P}{\beta_2} = \sum \left[ \left( -1 + \frac{(x+1) P_{x+1}}{m_1 m_2 \frac{P_x}{x}} \right) \frac{m_2}{(1+m_2)} \right] $$

$$ + \left( \frac{x}{m_2} - \frac{(x+1) P_{x+1}}{m_2 \frac{P_x}{x}} \right) \left( \frac{m_2 e^{m_2}}{m_1 (1+m_2)} \right) $$

Equating this to zero leads to

$$ m_2 (S_1 - N \bar{x}) - (S_1 - N \bar{x}) = 0 $$

(36)

where
So we obtain
\[ N \bar{x} = S_1, \]
corresponding to equation (35).

It should be remarked that \( S_1 \) can be written
\[
S_1 = \sum_{k=0}^{\infty} \left( \hat{\beta}_2 \right)^k \frac{K^{x+1/k}}{K!}.
\]

Clearly our maximum likelihood equation must be solved using an iterative scheme.

For the sake of comparison, the maximum likelihood equations based on the original parameters yield
\[ \hat{m}_1 \hat{m}_2 = \bar{x}, \]
and
\[ N \hat{m}_1 \hat{m}_2 = S_1, \]
where \( S_1 \) is defined as in equation (37).

3.1.1 The class of contagious distributions.

It is seen that the seven contagious distributions listed in Tables I and II have equation (35) as a common maximum likelihood equation. These contagious distributions are from a class of distributions investigated by Sprott [13].
He took as his starting point the maximum likelihood equations of the Poisson-Binomial distribution. The distribution is given by:

\[ P_x = \sum_{k=0}^{\infty} \binom{n}{k} p^k q^{n-k} x^n, \quad x = 0,1,2,\ldots, \]

\[ = 0 \text{ otherwise}, \]

where

\[ 0 < p < 1, \quad m > 0. \]

The likelihood equations are given by:

\[ \sum a_k \frac{(x+1) P_{x+1}}{p_x} = N \bar{x}, \tag{38} \]

and

\[ \hat{N} \hat{m} \hat{p} = \bar{x}, \tag{39} \]

where \( a_k \) is the observed frequency of \( x \). His aim was to obtain the set of distributions having (38) or both (38) and (39) as equations of maximum likelihood, and to indicate the procedure for solving them and for obtaining the information matrix.

Sprott pointed out that equations (38) and (39) are satisfied by the Neyman type A, the General Binomial-General Binomial, and the Poisson-Pascal, as well as by the Poisson-Binomial. However, for our purposes, equation (38) alone leads to the more interesting result, since it defines a broader class of distributions.

It was shown that the general form of the generating
function of distributions having (38) as a maximum likelihood equation is

\[ G(e_1, e_2, t) = G(e_1, (1-t)h(e_2)) \quad (40) \]

where \( h \) is an arbitrary function of \( e_2 \).

It has been pointed out that any distribution compounded with the binomial, Poisson, or hypergeometric distribution will have a maximum likelihood equation given by equation (38), since the form of \( G \) is arbitrary.

It is seen that the probability generating functions for the seven contagious distributions given in Tables Ia and IIa fit the form given by equation (40).
IV. USES OF THE COMMON CHARACTERISTICS

4.1 A Simplified Method for finding Orthogonal Parameters.

The relationship in (34) provides the basis for a simplified method of finding orthogonal parameters. We first assume that our distribution may have its mean as one of its orthogonal parameters, \( \beta_1 \). Furthermore, we assume that (34) holds; i.e.,

\[
\frac{\partial \log f}{\partial \beta_1} = \frac{x - \beta_1}{\nu_2}.
\]

Then

\[
\frac{\partial \log f}{\partial \beta_1} = \frac{\partial \log f}{\partial \alpha_1} \cdot \frac{\partial \alpha_1}{\partial \beta_1} + \frac{\partial \log f}{\partial \alpha_2} \cdot \frac{\partial \alpha_2}{\partial \beta_1} = \frac{x - \beta_1}{\nu_2},
\]

from which we can often determine \( \frac{\partial m_1}{\partial \beta_1} \) and \( \frac{\partial m_2}{\partial \beta_1} \), or their ratio, by inspection.

We also have \( \beta_2 = f_n(\alpha_1, \alpha_2) \). Thus \( \frac{\partial \beta_2}{\partial \beta_1} = 0 \) means

\[
\frac{\partial \beta_2}{\partial \alpha_1} \cdot \frac{\partial \alpha_1}{\partial \beta_1} + \frac{\partial \beta_2}{\partial \alpha_2} \cdot \frac{\partial \alpha_2}{\partial \beta_1} = 0,
\]

or

\[
\frac{\partial \beta_2/\partial \alpha_2}{\partial \beta_2/\partial \alpha_1} = -\frac{\partial \alpha_2/\partial \beta_1}{\partial \alpha_2/\partial \beta_1}.
\]

(42)

Since the right-hand side of equation (42) is now known,
the equation can be used to determine $\beta_2$, the second orthogonal parameter. Naturally the initial assumption that equation (41) holds would bear checking, but if one has been successful in applying (41), the check should be routine.

4.2 Applications of the Simplified Method.

4.2.1 Neyman type A distribution.

Consider the case of the Neyman type A distribution, for which equation (41) becomes:

$$\frac{1}{P_x} \left\{ -x + \frac{x+1}{m_1 m_2} P_{x+1} \right\} \frac{\partial m_1}{\partial \beta_1}$$

$$+ \frac{1}{P_x} \left\{ \frac{xp}{m_2} - \frac{x+1}{m_2} P_{x+1} \right\} \frac{\partial m_2}{\partial \beta_1} = \frac{x-m_1 m_2}{m_1 m_2 (1+m_2)} \cdot (43)$$

For the equality to hold, it is fairly obvious that we must have:

$$\frac{\partial m_2}{\partial \beta_1} = \frac{1}{m_1 (1+m_2)} \quad \text{and} \quad \frac{\partial m_1}{\partial \beta_1} = \frac{1}{1+m_2},$$

which gives

$$\frac{\partial \beta_2}{\partial m_1} = -m_1.$$  \hspace{1cm} (44)

Thus, as before,

$$\beta_2 = m_1 e^{-m_2}.$$  

Since $\beta_1 = \text{mean} = m_1 m_2$ by hypothesis, our pair of orthogonal parameters is completely determined.
4.2.2 Gram-Charlier Negative Exponential Distribution.

As an example of the simplified method applied to a continuous distribution, consider the Gram-Charlier Negative Exponential distribution defined by

\[
f(x) = \begin{cases} 
1 + a(kx - 1) & x > 0 \\
0 & \text{otherwise}
\end{cases} ,
\]

for \(0 < a < 1\), \(k > 0\).

The mean of this distribution is \(\frac{1+a}{k}\), and the variance is \(\frac{1+2a-a^2}{k^2}\). Thus equation (41) becomes

\[
\left\{ \frac{1}{k} - x + \frac{ax}{1+a(kx-1)} \right\} \frac{\partial k}{\partial \beta_1} + \left\{ \frac{kx-1}{1+a(kx-1)} \right\} \frac{\partial a}{\partial \beta_1} = \frac{k^2x-k(1-a)}{1+2a-a^2} .
\]

(46)

One would expect to pick up the \(k^2x\) term in the numerator of the right hand side of equation (46) from the free \(x\) in the first term of the left hand side. Accordingly we let

\[
\frac{\partial k}{\partial \beta_1} = \frac{-k^2}{1+2a-a^2}.
\]

The first term of the LHS of (46) becomes

\[
- \frac{k}{(1+2a-a^2)} + \frac{k^2x}{(1+2a-a^2)} - \frac{ak^2x}{(1+a(kx-1))(1+2a-a^2)} .
\]

Comparing this to the RHS, we see that the second term of the LHS, to be consistent, must be

\[
- \frac{ka[1+a(kx-1)] + ak^2x}{\{1+a(kx-1)\}(1+2a-a^2)} .
\]

(47)
The numerator of equation (47) factors to
\[ ak(1-a)(kx-1) \]
so that
\[ \frac{2a}{2b_1} = \frac{ak(1-a)}{1+2a-a^2} . \]

Thus we obtain
\[ \frac{\beta_2/\partial a}{\beta_2/\partial k} = -\frac{\partial k/\partial b_1}{\partial a/\partial b_1} = \frac{k}{a(1-a)} \]
so that our resulting parameters are
\[ \beta_1 = \frac{1+a}{k} \quad \text{and} \quad \beta_2 = \frac{ak}{1-a} \]

4.3 Further Applications of the General Method.

4.3.1 Neyman type A distribution.

It is of interest to see that the general method can be applied successfully if we are fortunate in our choice of \(a_1\) and \(a_2\). If the parameters of the Neyman type A are redefined so that
\[ a_1 = m_1m_2 \quad (\text{= mean}) \]
and
\[ a_2 = m_2 \]
and we then choose \(\beta_1 = a_1\), the differential equation (3) becomes, after simplification:
\[
(1 + \frac{\beta_1}{\alpha_2} + \frac{\beta_1}{\alpha_2}) - \frac{1}{\beta_1 \alpha_2} (1 + \frac{1}{\alpha_2}) \\
+ \{-\beta_1(1 + \frac{\beta_1}{\alpha_2} + \frac{\beta_1}{\alpha_2})(1 + \frac{1}{\alpha_2}) \\
+ \frac{1}{\alpha_2} (1 + \frac{1}{\alpha_2})^2 \} \frac{\partial a_2}{\partial \beta_1} = 0 ,
\]

when \( \beta \) is defined as in equation (5).

This can be broken up into a term containing \( \beta \), and a second term free of \( \beta \). Each term has the factor:

\[
1-\beta_1(1 + \frac{1}{\alpha_2}) \frac{\partial a_2}{\partial \beta_1} ,
\]

so that equation (48) holds when this is set equal to zero.

We obtain

\[
\frac{\partial a_2}{\partial \beta_1} = \frac{a_2}{\beta_1 (a_2 + 1)} .
\]

A solution of this is

\[
\beta_2 = \frac{\beta_1}{a_2} e^{-a_2} ;
\]
or in terms of the original parameters

\[
\beta_2 = m_1 e^{-m_2} ,
\]
as before.

Thus we found orthogonal parameters for the Neyman type A by assuming that the mean was an admissible orthogonal parameter and making the appropriate transformation before
we applied the general method.

4.3.2 Gram-Charlier Negative Exponential Distribution.

Unfortunately, defining \( a_1 = \text{mean} \) does not guarantee a solution even in those cases where the mean is an admissible orthogonal parameter, because the form of equation (3) still depends on the choice of \( a_2 \). For example, consider the Gram-Charlier Negative Exponential distribution defined in section 4.2.2, where

\[
\mu_1 = \frac{1+a}{k}
\]

The parameters may be redefined such that:

\[
a_1 = \frac{1+a}{k}
\]

and

\[
a_2 = k.
\]

Now, to apply the general method let

\[
\beta_1 = a_1.
\]

From equation (3) we obtain a relationship in terms of \( a_1 \), \( a_2 \), \( \frac{\beta a_2}{\beta a_1} \) and \( \ast \ast \), where

\[
\ast \ast = E \frac{1}{[1+a(kx-1)]^2}.
\]

Again dividing the equation into terms that contain \( \ast \ast \) and those that do not, and setting the common factor equal to zero, we obtain:
The solution of this equation is not apparent, even though we know from our earlier work in section 4.2.2 that one solution must be

$$\beta_2 = \frac{a_k}{1-a}$$

Thus is one is going to assume that the mean is an admissibile orthogonal parameter, the simplified method given in section 4.1 still has much to recommend it.
V. DISTRIBUTIONS WITH MORE THAN TWO PARAMETERS

5.1 Distributions with Three Parameters.

First consider a distribution \( f(x, \alpha_1) \) depending on \( n \) parameters \( \alpha_1 (i = 1, 2, ..., n) \). If \( \{s_j\} \) is a set of \( n \) independent functions of the \( \alpha_1 \), we find by differentiation:

\[
\frac{\partial^2 \log f}{\partial s_k \partial s_L} = \sum_{j} \frac{\partial s_j}{\partial s_k} \frac{\partial s_j}{\partial s_L} \frac{\partial^2 \log f}{\partial \alpha_1 \partial \alpha_j} + \sum_{j} \frac{\partial s_j}{\partial s_k} \frac{\partial^2 \log f}{\partial \alpha_1 \partial \alpha_j}.
\]

(49)

In the three-parameter case, if the \( s_j \) are orthogonal, taking expectations of equation (49) leads to a set of three constraints:

\[
E\left(\frac{\partial^2 \log f}{\partial s_1 \partial s_2}\right) = 0
\]

(50a)

\[
E\left(\frac{\partial^2 \log f}{\partial s_1 \partial s_3}\right) = 0
\]

(50b)

\[
E\left(\frac{\partial^2 \log f}{\partial s_2 \partial s_3}\right) = 0
\]

(50c)

However, a solution of this system of equations is clearly no simple matter.

Nevertheless Jeffreys [8] and Huzurbazar [6] have given examples of three-parameter distributions having orthogonal parameters. These may be found in Table V in the Appendix. It was consideration of Jeffreys' results which suggested to Huzurbazar [6] a special set of conditions under which
orthogonal parameters may be found for three-parameter distributions.

Define $$E_{ij} = E \frac{\partial^2 \log f}{\partial a_i \partial a_j}$$. Then let

$$E_{12} = E_{13} = 0$$,

and

$$E_{23}/E_{33}$$ be independent of $$a_1$$.

Then set $$a_1 = \beta_1$$, and $$a_2 = \beta_2$$. Under the orthogonality the following equations must be satisfied:

$$\frac{\partial a_3}{\partial \beta_1} = 0$$, \hspace{1cm} (51)

and

$$E_{23}/E_{33} \frac{\partial a_3}{\partial \beta_2} = 0$$ . \hspace{1cm} (52)

From equation (51) we see that $$a_3 = a_3(\beta_2, \beta_3)$$. From (52), we find:

$$\frac{\partial a_3}{\partial \beta_2} = - \frac{E_{23}}{E_{33}} = \phi(a_2, a_3) \hspace{1cm} \text{by hypothesis,}$$

so

$$\frac{\partial a_3}{\partial \beta_2} = \phi(\beta_2, a_3)$$ .

This last equation will often be solvable, and we can let the arbitrary constant be a function of $$\beta_3$$, our third orthogonal parameter.
5.2 Breakdown of the Alternate Method,

If we attempt to extend the alternate approach given in section 2.1 to three parameter distributions, we must first assume that

$$E\left( \frac{s^2 \log f}{\sigma_1 \sigma_2} \right) = C_{ij} + \sum_r K(ij) \cdot r \cdot r \cdot r \cdot \cdot \cdot$$

The matrix of coefficients analogous to (13) becomes

$$
\begin{bmatrix}
C_{11} + \sum_r K_{11} \cdot r \cdot a_1^2 & b_1^2 & c_1^2 \\
C_{21} + \sum_r K_{21} \cdot r \cdot a_1^2 & b_2^2 & c_2^2 \\
C_{31} + \sum_r K_{31} \cdot r \cdot a_1^2 & b_3^2 & c_3^2 \\
C_{12} + \sum_r K_{12} \cdot r \cdot a_1 a_2 & b_1 b_2 & c_1 c_2 \\
C_{13} + \sum_r K_{13} \cdot r \cdot a_1 a_3 & b_1 b_3 & c_1 c_3 \\
C_{23} + \sum_r K_{23} \cdot r \cdot a_2 a_3 & b_2 b_3 & c_2 c_3 \\
\end{bmatrix}
$$

(53)

where

$$a_1 = \frac{\partial \beta_1}{\partial a_1}, \quad a_2 = \frac{\partial \beta_1}{\partial a_2}, \quad a_3 = \frac{\partial \beta_1}{\partial a_3}$$

$$b_1 = \frac{\partial \beta_2}{\partial a_1}, \quad b_2 = \frac{\partial \beta_2}{\partial a_2}, \quad b_3 = \frac{\partial \beta_2}{\partial a_3}$$

$$c_1 = \frac{\partial \beta_3}{\partial a_1}, \quad c_2 = \frac{\partial \beta_3}{\partial a_2}, \quad c_3 = \frac{\partial \beta_3}{\partial a_3}$$
This matrix cannot be made square in general; thus there is no determinant of coefficients which one can set equal to zero to give us equations for the unknowns.

5.3 A Class of Distributions Having Orthogonal Parameters.

There exists a class of discrete distributions having more than two parameters for which orthogonal parameters exist. If the original parameters are the $a_i$'s of the set $\{a_i\}_{i=1}^n$, and if a transformation exists that takes the $\{a_i\}$ into $\{\beta_i\}$ such that:

$$
Pr\{X=r\} = g_1(\beta_1) \cdot g_2(\beta_2) \cdots g_n(\beta_n)
$$

(54)

then the $\beta_i$ are clearly orthogonal. For

$$
\log P_x = \psi_1(\beta_1) + \psi_2(\beta_2) \cdots \psi_n(\beta_n)
$$

$$
\frac{\partial \log P_x}{\partial \beta_i} = n(\beta_i), \text{ a function of } \beta_i \text{ above.}
$$

Clearly any second derivative gives:

$$
\frac{\partial^2 \log P_x}{\partial \beta_i \partial \beta_j} = 0 \text{ for } i \neq j.
$$

(55)

This is a much stronger condition than that of orthogonality, given by equation (1).

As an example from this class, consider a discrete case defined on a finite range by:

<table>
<thead>
<tr>
<th>cell</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>1-a-b-c</td>
</tr>
</tbody>
</table>
The transformation
\[ a = \beta_1, \]
\[ b = \beta_2(1-\beta_1), \]
and
\[ c = \beta_3(1-\beta_2)(1-\beta_1), \]
gives

<table>
<thead>
<tr>
<th>cell</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>( \beta_1 )</td>
<td>( \beta_2(1-\beta_1) )</td>
<td>( \beta_3(1-\beta_2)(1-\beta_1) )</td>
<td>( (1-\beta_1)(1-\beta_2)(1-\beta_3) )</td>
</tr>
</tbody>
</table>

the \( \beta_i \)'s now define the distribution as in equation (54) and are thus orthogonal.

If this procedure is extended a more general form of this class of distributions is:

\[
P\{X=x_1\} = p^n_1, \]
\[
P\{X=x_r\} = p^n_r, \]
\[
P\{X=x_{r+1}\} = \{1-\sum_{i=1}^{r} p^n_i\} t^{m_1}, \]
\[
P\{X=x_{r+s}\} = \{1-\sum_{i=1}^{r} p^n_i\} t^{m_s}, \]

\[
P\{X=x_{r+s+1}\} = \{1-\sum_{i=1}^{r} p^n_i\} \{1-\sum_{i=1}^{s} t^{m_i}\} w_1, \]
\[
\vdots \]
\[
P\{X=x_{r+s+u}\} = \{1-\sum_{i=1}^{r} p^n_i\} \{1-\sum_{i=1}^{s} t^{m_i}\} \ldots h_u, \]
\[
P\{X=x_{r+s+u+1}\} = \{1-\sum_{i=1}^{r} p^n_i\} \{1-\sum_{i=1}^{s} t^{m_i}\} \ldots \{1-\sum_{i=1}^{u} h_i\}, \]
where the parameters \( p, t, \ldots h \) are orthogonal. Note that this distribution is defined on a finite range.

It appears that a discrete distribution having orthogonal parameters which is defined over an infinite range could be generated if, beginning with the last term given above in (55), we formed probabilities of the type

\[
F \{ X = x_r + s + \ldots + u + 1 + i \} = P \{ X = x_r + s + \ldots + u \} \cdot P_l (y = i) \quad i = 0, 1, \ldots
\]

where \( y \) is a discrete variate having its distribution defined over an infinite range. A simple example of this type, using the Poisson distribution, is:

\[
\begin{align*}
P[X=0] &= p \\
P[X=1] &= (1-p)t \\
P[X=2] &= (1-p)(1-t)e^{-m} \\
P[X=3] &= (1-p)(1-t)m e^{-m} \\
&\vdots \\
P[X=r] &= (1-p)(1-t) \frac{m^{r-2}}{(r-2)!} e^{-m} \\
&\vdots
\end{align*}
\]

where \( p, t, \) and \( m \) are orthogonal.
VI. CONCLUSIONS

It is pointed out by Huzurbazar [6] that when the number of parameters, \( n \), is greater than three, the problem of finding orthogonal parameters is impossible in general; and when \( n = 3 \) the problem is exceedingly difficult. Although a case was shown, in section 5.3, where an arbitrary number of parameters may be orthogonal, it is clearly not a general case, and the above conclusions require no amplification.

For finding orthogonal parameters in the two-parameter case, one has the general method given in section 1.1; in addition an alternate method is developed in section 2.1 which extends somewhat our ability to find these parameters. The alternate method was applied successfully to seven contagious distributions and three Gram-Charlier type distributions.

While the number of pairs of parameters that will satisfy the orthogonality condition is theoretically infinite, it is by no means obvious whether there are basic solutions to the problem even when limiting ourselves to the two parameter case. For example, for the Neyman-type A, one pair of orthogonal parameters is

\[ \beta_1 = m_1 m_2 \quad \text{and} \quad \beta_2 = m_1 e^{-m_2} \].

The question arises as to whether one can obtain any other
pair that is not merely of the form
\[ \beta_1^* = \eta(\beta_1) \quad \beta_2^* = \eta(\beta_2) \]
e.g., for the Neyman type A it is not clear whether we could find \( \beta_2^* \) given \( \beta_1^* = m_1 \) (or \( m_2 \)).

Finding one pair of orthogonal parameters may be beyond our present methods, as seen in section 2.3.1. An example of a Gram-Charlier type-B distribution was given there which could not be solved by either the general method or the alternate method. One can conclude that even in the two-parameter case, the general problem of finding orthogonal parameters is quite difficult.

Another disadvantage associated with orthogonal parameters is that the physical meanings that may be attached to the original parameters may be lost by the transformations. Hopefully the advantages gained will outweigh this loss of information.

Furthermore, it was brought out in section 3.1 that direct transformations for writing a distribution \( f(x, \alpha_i) \) as \( f(x, \beta_j) \) (where \( \alpha_i \) and \( \beta_j \) are respectively the original and orthogonal parameters) cannot always be found. This is an undesirable feature, and limits the value of such orthogonal parameters.

One advantage given by orthogonal parameters is that a graphical solution to the maximum likelihood equations should be quite easy to obtain from plots of the likelihood
contours, since the contours associated with these parameters are nearly elliptical. This has been illustrated by Anscombe [1]. He showed a plot of the likelihood contours based on the original parameters that led to boomerang-like shapes, which left one in doubt as to the location of the solution. By contrast, with orthogonal parameters the same data gave elliptical contours from which, by picking the center of these contours, one could obtain a close approximation to the maximum likelihood solution.

Another advantage is that the Information matrix for finding asymptotic variances should be simplified, since, if the parameters are orthogonal, the off-diagonal terms will be zero.

Finally, the point raised by Jeffreys [8] can be considered; i.e., that since the equations to determine the orthogonal parameters are nearly orthogonal, their practical solution will be much simplified. In the case of the contagious distributions, if one compares the likelihood equations based on the orthogonal parameters with the likelihood equations given by the original parameters, no clear advantage is apparent. Whether the more rapid convergence expected of the likelihood equations of the orthogonal parameters is significant enough to warrant their use is a moot point. In a practical case it seems unlikely that they will be used except in those cases where the convergence time for
the likelihood equations based on the original parameters
proves uneconomical on a computer.

In summary, orthogonal parameters have been found for
ten distributions not previously considered. It appears
that, in general, the use of orthogonal parameters is
limited, first by our ability to determine them, and second
by the characteristics of the parameters one is able to
determine. In cases where orthogonal parameters can be
found, their application will depend on the goals of the
investigator.
VII. ACKNOWLEDGEMENTS

I am indebted to Dr. Boyd Harshbarger for his advice and encouragement during the course of this work; to Dr. Raymond Myers for his timely assistance in reading the draft of this thesis; and, most of all, to Dr. L. R. Shenton, who guided this work, and whose counsel and aid were invaluable in its preparation.

I would also like to express my gratitude for the financial assistance received from a National Institute of Health scholarship.
VIII. BIBLIOGRAPHY


The vita has been removed from the scanned document
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>DISTRIBUTION LAW</th>
<th>PROB. GEN. FUNCT.</th>
<th>MEAN</th>
<th>VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEYMAN TYPE-A</td>
<td>( e^{-m_1 m_2 x} \sum_{k=0}^{x} \frac{(m_1 e^{-m_2})^k x^k}{k!} ) ( x = 0, 1, 2, \ldots )</td>
<td>( e_{m_2}(t-1) )</td>
<td>( m_1 m_2 )</td>
<td>( m_1 m_2(1+m_2) )</td>
</tr>
<tr>
<td>POISSON-BINOMIAL</td>
<td>( \sum_{k=0}^{x} \frac{e^{-m k}}{k!} \binom{n}{k} p^k q^{n-k} e^{-m k} ) ( x = 0, 1, 2, \ldots )</td>
<td>( e[(pt+q)^n-1] )</td>
<td>( nmp )</td>
<td>( nmp[1+p(n-1)] )</td>
</tr>
<tr>
<td>BINOMIAL-POISSON</td>
<td>( \sum_{k=0}^{n-k} \binom{n}{k} p^k q^{n-k} e^{-m k} ) ( x = 0, 1, 2, \ldots )</td>
<td>( d [e^{m(t-1)} - q]^n )</td>
<td>( nmp )</td>
<td>( nmp[1+mq] )</td>
</tr>
<tr>
<td>GEOMETRIC-POISSON</td>
<td>( \sum_{k=1}^{p(1-p)^{k-1}} e^{-m k} ) ( x = 0, 1, 2, \ldots ) ( x )</td>
<td>( \frac{p e^{m(t-1)}}{1-q \cdot e^{m(t-1)}} )</td>
<td>( \frac{m}{p} )</td>
<td>( \frac{m(p + mq)}{n^2} )</td>
</tr>
</tbody>
</table>

**TABLE I(a). (DISCRETE DISTRIBUTIONS; CONTAGIOUS)**
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta \frac{\log P_x}{\beta_1}$</th>
<th>RESULT FROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEYMAN-TYPE A</td>
<td>$m_1 m_2$</td>
<td>$m_1 e^{-m_2}$</td>
<td>$\frac{x-m_1 m_2}{m_1 m_2 (1+m_2)}$</td>
<td>$N \bar{x} = S_1^+$</td>
</tr>
<tr>
<td>POISSON-BINOMIAL</td>
<td>$nmp$</td>
<td>$m (1-p)^n$</td>
<td>$\frac{x-nmp}{nmp[1+p(n-1)]}$</td>
<td>$N \bar{x} = S_1^+$</td>
</tr>
<tr>
<td>BINOMIAL-POISSON</td>
<td>$nmp$</td>
<td>$\frac{p}{q} e^{-m}$</td>
<td>$\frac{x-nmp}{nmp(1+mq)}$</td>
<td>$N \bar{x} = S_1^+$</td>
</tr>
<tr>
<td>GEOMETRIC POISSON</td>
<td>$\frac{p}{p}$</td>
<td>$q e^{-m}$</td>
<td>$\frac{p^2 x-m p}{m (p+mq)}$</td>
<td>$N \bar{x} = S_1^+$</td>
</tr>
</tbody>
</table>

$S_1 = \sum (x+1) \frac{P_x}{x+1}$

**TABLE I(b). (DISCRETE DIST.: CONTAGIOUS)**
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>DISTRIBUTION LAW, $P_x$</th>
<th>PROB. GEN. FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOGARITHMIC-BINOMIAL</td>
<td>$\sum_{k=1}^{\infty} \frac{\theta^k (kn)^x (1-m)^{kn-x}}{-k \log(1-\theta)}$</td>
<td>$\frac{\log[1-\theta (1-m+mt)^n]}{\log(1-\theta)}$</td>
</tr>
<tr>
<td></td>
<td>$x = 0, 1, 2, \ldots$</td>
<td></td>
</tr>
<tr>
<td>LOGARITHMIC-POISSON</td>
<td>$\sum_{k=1}^{\infty} \frac{\theta^k e^{-km} (km)^x}{-k \log(1-\theta) x!}$</td>
<td>$\frac{\log[1-\theta e^{m(t-1)}]}{\log(1-\theta)}$</td>
</tr>
<tr>
<td></td>
<td>$x = 0, 1, 2, \ldots$</td>
<td></td>
</tr>
<tr>
<td>NORMAL-POISSON</td>
<td>$\int \frac{e^{-\frac{(n-m)^2}{2\sigma^2}} \cdot \frac{2^{-n\theta}(n\theta)^x}{x!}}{\sqrt{2\pi \sigma}} , dn$</td>
<td>$\frac{m^2}{2\sigma^2} \cdot \frac{[\theta e^{2\sigma^2(1-t)}]^2}{e^{2\sigma^2} \cdot e^{2\sigma^2}}$</td>
</tr>
<tr>
<td></td>
<td>$x = 0, 1, 2, \ldots$</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II(a). (DISCRETE DIST.; CONTAGIOUS)**
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>MEAN</th>
<th>VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOGARITHMIC-BINOMIAL</td>
<td>$\frac{-nm\theta}{(1-\theta)\log(1-\theta)}$</td>
<td>$\frac{-nm\theta[(1-m)(1-\theta)1+mn]}{(1-\theta)^2 \log(1-\theta)} - \left(\frac{nm\theta}{(1-\theta)\log(1-\theta)}\right)^2$</td>
</tr>
<tr>
<td>LOGARITHMIC-POISSON</td>
<td>$\frac{-m\theta}{(1-\theta)\log(1-\theta)}$</td>
<td>$\frac{-m\theta[1-\theta+m]}{(1-\theta)^2 \log(1-\theta)} - \left(\frac{m\theta}{(1-\theta)\log(1-\theta)}\right)^2$</td>
</tr>
<tr>
<td>NORMAL-POISSON</td>
<td>$m\theta$</td>
<td>$(m+\theta\sigma^2)$</td>
</tr>
</tbody>
</table>

**TABLE II(a)(Continued). (DISCRETE DIST.; CONTAGIOUS)**
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\frac{\partial \log P}{\partial \theta_1}$</th>
<th>$\frac{\partial \log P}{\partial \theta_2}$</th>
<th>RESULT FROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOGARITHMIC-BINOMIAL</td>
<td>$\frac{-nm\theta}{(1-\theta)\log(1-\theta)}$</td>
<td>$\theta(1-\theta)^n$</td>
<td>$\frac{x-u_1}{u_2}$</td>
<td>$n \bar{x} = S_1$</td>
<td>$\sum \frac{\partial \log P}{\partial \theta_2} = 0$</td>
</tr>
<tr>
<td>LOGARITHMIC-POISSON</td>
<td>$\frac{-m\theta}{(1-\theta)\log(1-\theta)}$</td>
<td>$\theta e^{-m}$</td>
<td>$\frac{x-u_1}{u_2}$</td>
<td>$n \bar{x} = S_1$</td>
<td>$\sum \frac{\partial \log P}{\partial \theta_2} = 0$</td>
</tr>
<tr>
<td>NORMAL-POISSON</td>
<td>$m\theta$</td>
<td>$m-\theta \sigma^2$</td>
<td>$\frac{x-m\theta}{\theta(m+\theta \sigma^2)}$</td>
<td>$n \bar{x} = S_1$</td>
<td>$\sum \frac{\partial \log P}{\partial \theta_2} = 0$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$S_1 = \sum \frac{(x+1)p_{x+1}}{p_x}$</td>
</tr>
</tbody>
</table>

TABLE II(b). (DISCRETE DIST.; CONTAGIOUS)
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>DISTRIBUTION LAW</th>
<th>PROB. GEN. FUNCTION</th>
<th>MEAN</th>
<th>VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEGATIVE BINOMIAL</td>
<td>( \frac{(k+x-1) \cdot p^x}{x \cdot (1+p)^{k+x}} )</td>
<td>( (1+p-tp)^{-p} )</td>
<td>kp</td>
<td>kp(1+p)</td>
</tr>
<tr>
<td>x = 0,1,2,...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRAH-CHARLIER TYPE-B</td>
<td>( \left{1+\frac{a(x-1)}{m}\right} \cdot \frac{e^{-m} \cdot x}{x!} )</td>
<td>( (1+a(t-1)) \cdot e^{m(t-1)} )</td>
<td>m+a</td>
<td>m+a(1-a)</td>
</tr>
<tr>
<td>x = 0,1,...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE III(a). (DISCRETE DIST.; MISC.)**
TABLE III(d) (DISCRETE DIST.; MISC.)

<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>NEGATIVE BINOMIAL</th>
<th>GEM-CHARLIER TYPE-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>( \nu )</td>
<td>( \frac{m(1-a)}{a} )</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>( \frac{x\cdot k \cdot p}{k \cdot p (1+p)} )</td>
<td>( \frac{m+\alpha}{m+\alpha} )</td>
</tr>
<tr>
<td>( 2 \log \frac{P}{\theta_2} )</td>
<td>( \sum \frac{2}{\theta_2} \log \left( \frac{\theta_2}{x-1,1} \right) - N \log \left( 1 + \frac{x}{\theta_2} \right) )</td>
<td>( N \left( \frac{x}{\theta_2} - 1 \right) - \frac{1}{\theta_2} )</td>
</tr>
<tr>
<td>RESULT FROM</td>
<td>( \sum \frac{2}{\theta_2} \log \left( \frac{\theta_2}{x-1,1} \right) - N \log \left( 1 + \frac{x}{\theta_2} \right) )</td>
<td>( N \left( \frac{x}{\theta_2} - 1 \right) - \frac{1}{\theta_2} )</td>
</tr>
</tbody>
</table>

\( \sum \) ranges from \( 1 \) to \( N \).
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>PROB. DENSITY FUNCTION</th>
<th>MOMENT GEN. FUNCTION</th>
<th>MEAN</th>
<th>VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NORMAL</td>
<td>$\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}$</td>
<td>$e^{(tm+\frac{1}{2}t^2\sigma^2)}$</td>
<td>$\mu$</td>
<td>$\sigma^2$</td>
</tr>
<tr>
<td>GRAM-CHARLIER</td>
<td>$(1+a(kx-1))ke^{-kx}$</td>
<td>$(1-a)^{1-t}+a(1-t)^{-2}$</td>
<td>$\frac{1+a}{k}$</td>
<td>$\frac{1+2a-a^2}{k^2}$</td>
</tr>
<tr>
<td>NEG. EXP.</td>
<td>$0 &lt; x &lt; \infty$</td>
<td>$0 &lt; x &lt; \infty$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRAM-CHARLIER</td>
<td>$\left[1+a\left(\frac{k}{n}x-1\right)\right]\frac{k^n x^{n-1}e^{-kx}}{\Gamma(n)}$</td>
<td>$(1-a)^{1-t}+a(1-t)^{-2}(n+1)$</td>
<td>$\frac{n+a}{k}$</td>
<td>$\frac{n+2a-a^2}{k^2}$</td>
</tr>
<tr>
<td>GAMMA (n unk)</td>
<td>$\frac{k^n x^{n-1}e^{-kx}}{\Gamma(n)}$</td>
<td>$(1 - \frac{t}{k})^{-n}$</td>
<td>$\frac{n}{k}$</td>
<td>$\frac{n}{k^2}$</td>
</tr>
<tr>
<td></td>
<td>$0 &lt; x &lt; \infty$</td>
<td>$0 &lt; x &lt; \infty$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE IV(a). (CONTINUOUS DIST.)**
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\frac{\partial \log P}{\partial \beta_1}$</th>
<th>$\frac{\partial \log P}{\partial \beta_2}$</th>
<th>RESULT FROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NORMAL</td>
<td>$\mu$</td>
<td>$\sigma^2$</td>
<td>$\frac{x-\mu}{\sigma^2}$</td>
<td>$\sigma^2 = \frac{\sum(x_i-\bar{x})^2}{n}$</td>
<td>$\sum \frac{\partial \log P}{\partial \beta_2} = 0$</td>
</tr>
<tr>
<td>GHAN-CHARLIER</td>
<td>$\frac{1+a}{k}$</td>
<td>$\frac{ak}{1-a}$</td>
<td>$\frac{k^2x-k(1+a)}{1+2a-a^2}$</td>
<td></td>
<td>$N(1-a) = \sum (1+\hat{\beta}_2 x)^{-1}$</td>
</tr>
<tr>
<td>NEGF. EXP.</td>
<td>$\frac{n+a}{k}$</td>
<td>$\frac{ak}{1-a}$</td>
<td>$\frac{k^2x-k(n+a)}{n+2a-a^2}$</td>
<td></td>
<td>$N(1-a) = \sum n{n+\hat{\beta}_2 x}^{-1}$</td>
</tr>
<tr>
<td>GHAN-CHARLIER</td>
<td>$\frac{n+k}{k}$</td>
<td>$\frac{n}{k}$</td>
<td>$\frac{k^2x-nk}{n}$</td>
<td></td>
<td>$N\log \hat{\beta}_2 - n\log \bar{x} + \sum \log x - \frac{1}{\hat{\beta}_2} \Gamma(\hat{\beta}_2) = 0$</td>
</tr>
</tbody>
</table>

**TABLE IV(b). (CONTINUOUS DIST.)**
<table>
<thead>
<tr>
<th>DISTRIBUTION</th>
<th>PROB. DENSITY FUNCTION</th>
<th>TRANSFORMATIONS TO ORTHOGONAL PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEARSON TYPE II [8]</td>
<td>[ \frac{(m - \frac{1}{2})!}{\frac{1}{2}} \left{ 1 - \frac{m^2(x-\lambda)^2}{2(m+\frac{1}{2})^3 \sigma^2} \right}^m ]</td>
<td>[ \lambda = \beta_1 ] [ \sigma = \beta_2 ] [ m = - \frac{1}{\beta_3} ]</td>
</tr>
<tr>
<td>PEARSON TYPE III [6]</td>
<td>[ \frac{1}{\alpha_3 \alpha_2} \cdot \frac{1}{\alpha_2} \cdot \frac{(x-a_1)^{a_3}}{\left(\frac{x-a_1}{\alpha_2}\right)^{a_3}} ]</td>
<td>[ a_1 = \beta_1 \cdot (\beta_3-1) \cdot \beta_3 \cdot \beta_2 ] [ a_2 = \beta_3 \cdot \beta_2 ] [ a_3 = \varepsilon_3 \ (\text{for } a_3 \text{ large}) ]</td>
</tr>
<tr>
<td>PEARSON TYPE VII [8]</td>
<td>[ \frac{m!}{\frac{1}{2} \left{ 1 + \frac{m^2(x-\lambda)^2}{2(m+\frac{1}{2})^3 \sigma^2} \right}} ]</td>
<td>[ \lambda = \beta_1 ] [ \sigma = \beta_2 ] [ m = \frac{1}{\beta_3} ]</td>
</tr>
</tbody>
</table>

**TABLE V.** (THREE PARAMETER DISTRIBUTIONS)
Orthogonal parameters for a distribution, $f(x, \theta_i)$, $(i = 1, 2, \ldots, n)$, are defined such that

$$E(- \frac{\partial^2 \log f}{\partial \theta_i \partial \theta_j}) = 0 \quad \text{for all } i, j, (i \neq j).$$

It has been pointed out that the problem of estimating the parameters of a distribution by maximum likelihood procedures, when the likelihood equations require iterative schemes for their solution, may be simplified by the use of orthogonal parameters.

Distributions having maximum likelihood equations where iteration is required include a number of two-parameter contagious distributions, like the Neyman type A and the Poisson-Binomial. The general method for finding orthogonal parameters is examined, and is seen to be inappropriate for the contagious distributions. An alternate method is developed by which orthogonal parameters are obtained for the Neyman type A, Poisson-Binomial, Binomial-Poisson, Geometric-Poisson, Logarithmic-Binomial, Logarithmic-Poisson and Normal-Poisson distributions, as well as for three Gram-Charlier type distributions. Some characteristics of the class of distributions to which the alternate method is applicable are discussed.

The limitations of the general and the alternate
methods are examined, and an example given where neither is of any use. It is also pointed out that, in many cases where orthogonal parameters are determined, simple transformations by which one can write the distribution in terms of the orthogonal parameters may not exist. It is concluded that the methods for determining the parameters are somewhat limited in scope; and that although the characteristics of the orthogonal parameters may be useful, the disadvantages associated with them may restrict their application.