DOCTORAL DISSERTATION SERIES

PUBLICATION: 7715

AUTHOR: Robert Winter Royston, Ph. D., 1954
University of Michigan

TITLE: A FREQUENCY FUNCTION WHICH CAN BE TRANSFORMED INTO A GAMMA TYPE FUNCTION BY A QUADRATIC TRANSFORMATION OF THE VARIABLE

University Microfilms, Ann Arbor, Michigan

A FREQUENCY FUNCTION WHICH CAN BE TRANSFORMED INTO A GAMMA TYPE FUNCTION BY A QUADRATIC TRANSFORMATION OF THE VARIABLE

by

Robert Winter Royston

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1953

Doctoral Committee:

Professor Cecil C. Craig, Chairman
Professor Paul S. Dwyer
Professor Cecil J. Nesbitt
Professor Willard C. Olson
Professor George Y. Rainich

A Frequency Function which can be Transformed into a Gamma Type Function by a Quadratic Transformation of the Variable

by

Robert Winter Royston

The random variable under consideration, z, has as its probability density function

$$f(z)dz = \frac{2a^p}{\sqrt{p}} (z + B)^{2p-1} e^{-a(z+B)^2} dz, \quad -B \leq z < \infty .$$

It is obvious that the variable $x = a(z + B)^2$ has a Gamma type distribution.

The frequency curve f(z) is bell-shaped for $p > 1/2$ and J-shaped for $p \leq 1/2$, with a finite ordinate for $p = 1/2$ and an infinite ordinate for $p < 1/2$.

A criterion for testing the hypothesis that a set of observations is a random sample from a population which is distributed according to the Z-function and the formula for the sampling variance of this criterion are developed.

The functions $f_o(p) = \frac{\sqrt{(p + 1/2)}}{\sqrt{p}}$ and $\phi_o^2 = p - f_o^2(p)$ occur in the equations for the estimation of the parameters by the method of moments. A detailed study of them is necessary before an investigation of the roots of these equations can be made. It is then proven that for any set of observations there will exist a unique set of estimates for the parameters by the method of moments. Expressions for the variances of these estimates are derived by the δ-method.

In the event that estimates for the parameters are obtained by both the Z-function and the Gamma type function, it is shown that the percentage variances for the estimates by the Z-function are many times smaller. A possible exception is in the case of what may be regarded as the "scaling" parameter. In this case the percentage variances are approximately equal.

The functions $R(k) = \sum_{i=1}^{n} \frac{1}{(w_i + k)}$ and $P(k) = \prod_{i=1}^{n} (w_i + k)^{\frac{2}{n}}$, w_i representing the z variate when measured about its mean, occur in the equations for the estimation of the parameters by the method of maximum likelihood. A detailed study of these functions is necessary for the investigation of the roots for these equations and the subsequent theory concerning the maximum likelihood estimators.

It is proved that there will exist a set of estimates for the parameters by the method of maximum likelihood unless the frequency curve is J-shaped. It is further shown that these estimators meet all the necessary conditions for maximizing the likelihood function, and that they cannot minimize it.

The validity of the conventional formulas for the variances of the maximum likelihood estimators is established, even though the range of the variable is a function of the parameters.

The actual determination of the parameters by each of the methods referred to above is illustrated by numerical examples in graduation.

Acknowledgments

I am greatly indebted to Professor C. C. Craig for his advice in making this study. His constant interest in reading the manuscript at its various stages of development and many helpful suggestions have been invaluable. I wish to express my appreciation also to Professor H. C. Carver for reading the manuscript.

TABLE OF CONTENTS

Page

I. Introduction.

1.1 Statement of problem 1

1.2 Conditions imposed on the parameters 3

II. Analysis of the Z-function.

2.1 Variation in the parameters resulting from transformation of the z variates 7

2.2 Effect on the cumulative distribution function of varying the parameters..... 7

2.3 Forms assumed by the frequency curve f(z) 9

III. Estimation by method of moments.

3.1 Equations for the determination of estimates of the parameters and procedure for their solution 13

3.2 Study of the functions $f_o(p)$ and $\phi_o^2(p)$ 15

3.3 The existence of unique moment estimates of the parameters 27

3.4 Formulas for the variance and covariance of the moment estimates of the parameters 33

3.5 Comparison of the variances of the moment estimates of the independent parameters as obtained from the Z-function and as obtained from the Gamma type function 38

IV. Estimation by method of maximum likelihood

4.1 Equations for the determination of maximum likelihood estimates of the parameters and outline of the procedures for their solution 46

4.2 Study of the functions $R = \sum_{i=1}^{n} \frac{1}{w_i + k}$ and $P = [\prod_{i=1}^{n}(w_i + k)]^{2/n}$ 49

4.3 Roots of the maximum likelihood equations and the existence of an admissible solution 66

4.4 Existence of a sufficient set of estimates 77

4.5 Analysis of the roots of the maximum likelihood equations relative to their maximizing the likelihood function 78

4.6 Formulas for the variance and covariance of maximum likelihood estimates 82

V. Test criterion for the Z-function.

5.1 Criterion for testing the hypothesis that a given random sample comes from a populations whose probability density function is $f(z)\,dz$ 92

5.2 The variance of the testing criterion 94

VI. Numerical illustrations.

6.1 By method of moments 97

6.2 By maximum likelihood method 105

Appendix I. Tables of values for $f_o(p)$, $f_o^2(p)$, and $\phi_o^2(p)$..... 108

Appendix II. Technique to be employed in a numerical problem by method of moments for a negative skewness 117

List of references 119

PART I

Introduction

1.1 Statement of problem. There are an infinite number of types of distribution functions and a given random variable may obey any one of them. The amount of information and developed theory concerning both the population and sample characteristics vary widely with the type of distribution function which the variable obeys. Frequently one may known the form or type of the distribution function, $f(x)dx$, for x, the random variable involved, but $f(x)dx$ may be among that vast class of distribution functions about which little or even nothing is known. Now suppose x is transformed into some new random variable, $\varphi(x)$, with distribution function, $f[\varphi(x)]\, d[\varphi(x)] = g(x)$, then any additional information or theory which may be available about $g(x)$ can be utilized. This is especially beneficial if $g(x)$ is a member of that limited class of highly developed distribution functions. For this reason, the problem of transformation of a random variable along with the accompanying transformation of its distribution function is of importance in statistical theory.

The problem of the transformation of a variable obeying a known distribution function has been treated in a variety of ways by several authors. The papers referred to here in no sense form an exhaustive list. Rather they are selected as typifying the work which has been done along this line. Baker (Ref. 1); Edgeworth (Ref. 8, 9, & 10); Kapetyn (Ref. 15 and 16); Rietz (Ref. 23); Wicksell (Ref. 26); and Yuan (Ref. 27) consider the distribution function resulting from the transformation of normally distributed variates or the inverse of this problem. The transformations considered in detail are of two general types: 1) powers of the variates and 2) logarithms of the variates. Some of the writers consider in detail the determination of a transformation such that the

1.1

transformed variate is normally distributed. Olshen (Ref. 21) deals with the distribution function resulting from the transformation $x' = x^n$, in which x is distributed according to the Pearson type III function.

This paper will be concerned with a three parameter transformation of a more general algebraic nature than the one used by Olshen. The emphasis will be placed on the transformed function, the determination of the estimates of the parameters, and the performance of the inverse transformation in a numerical problem in graduation.

Let z, the random variable under consideration, be measured in its natural units about zero and let it take on only real values over a range to be specified later. Further, the probability density function of z, $f(z)dz$, to be referred to as the Z-function, is such that if each z variate is subjected to the transformation given below, the new variate, t, will be distributed according to a Gamma type distribution function. The transformation and its inverse are:

$$t = a'z^2 + b'z + c' \tag{1.1.1}$$

or

$$z = \frac{-b' \pm \sqrt{(b')^2 - 4a'(c' - t)}}{2a'}. \tag{1.1.1a}$$

The probability density function* (Ref. 5) for t is given by

$$f(t)dt = \frac{A^{A^2}(A + t)^{A^2 - 1} e^{-A(A + t)} dt}{\sqrt{A^2!}} \tag{1.1.2}$$

in which t is measured in standard units about its mean and $-A \leq t < \infty$.

The following points relative to the Z-functions will be treated in detail:

---

* This distribution function is also known as the Pearson type III function. However, the name Gamma type function will be used in this paper.

1.1-2

1) Conditions to be imposed on the parameters.

2) Analysis of the several forms which the frequency curve can assume.

3) Determination of the parameters from a given sample by method of moments and by method of maximum likelihood, along with the variance of these estimates.

4) Proof of the existence of solutions of the equations for the determination of the parameters by each of the two methods.

5) Development of a criterion for testing the hypothesis that a given sample comes from a population whose distribution function is the Z-function.

6) Numerical illustrations.

1.2 Conditions to be imposed on the parameters. The substitution of (1.1.1) in (1.1.2) gives

$$(1.2.1) \qquad f(z)dz = \frac{[Aa'z^2 + Ab'z + A(A+c')]^{A^2-1}}{\sqrt{A^2}}\, e^{-[Aa'z^2 + Ab'z + A(A+c')]}\, A(2a'z + b')dz$$

By writing $Aa' = a$; $Ab' = b$; $(A + c')A = c$; and $A^2 = p$; then (1.1.1a) becomes

$$(1.2.2) \qquad z = \frac{-b \pm \sqrt{b^2 - 4a\,c - A(A+t)}}{2a}$$

or

$$(1.2.2a) \qquad A(A + t) = az^2 + bz + c.$$

Then (1.1.2) becomes

$$(1.2.3) \qquad f(z)dz = \frac{1}{\sqrt{p}}(az^2 + bz + c)^{p-1}\, e^{-(az^2 + bz + c)}\, (2az + b)\, dz\,.$$

The independent parameters of the Z-function are now a, b, c and p.

1.2

Since a', b', c' and A are real and finite, a, b, c and p are also real and finite with p greater than zero.

In order to determine the conditions to be imposed on a, b and c to ensure a one-to-one correspondence between values of z and t, write $A(A + t) = x$. The range of values for x will be $0 \leq x < \infty$, and (1.2.2) becomes

$$x = az^2 + bz + c. \tag{1.2.4}$$

A one-to-one correspondence between x and z will give a like correspondence between z and t. Further the relationships connecting a, b, c and p with a', b', c' and A are single valued. Thus a one-to-one correspondence between x and z in (1.2.4) will ensure a like correspondence between t and z in (1.1.1) or (1.1.1a).

From (1.2.4) the corresponding values of x and z will lie along a parabola. See figure 1. In order that the transformation be single valued only those values of x and z which lie along the upper branch of the parabola will be considered. That is

$$z = \frac{-b + \sqrt{b^2 - 4a(c - x)}}{2a}. \tag{1.2.5}$$

To ensure that z remains real over the entire range for x, i.e. $0 \leq x < \infty$, it is necessary and sufficient that:

1) $a > 0$. It is sufficient that $a \geq 0$, but for $a = 0$ the transformation becomes linear.

2) $b^2 - 4ac \geq 0$.

By writing $\theta = \frac{b^2 - 4ac}{4a}$, $f(z)dz$ from (1.2.3) becomes

$$f(z)dz = \frac{1}{\sqrt{p}}\left[a\left(z + \frac{b}{2a}\right)^2 - \theta\right]^{p-1} e^{-\left[a\left(z + \frac{b}{2a}\right)^2 - \theta\right]} (2az + b)dz, \tag{1.2.6}$$

and the transformation (1.2.4) becomes

1.2

(1.2.7) $$x = a(z + \frac{b}{2a})^2 - \theta .$$

Further, putting $x + \theta = y$ gives

(1.2.8) $$f(y)dy = \frac{1}{\Gamma p}(y - \theta)^{p-1} e^{-(y - \theta)} dy.$$

Thus $f(y)dy$ is a Gamma type function with the lower end of the range at θ.

This paper will be concerned only with that class of Z-functions which can be transformed into Gamma type functions the range of whose variables are constant with respect to the parameters of transformation. It is obvious from (1.2.8) that this is equivalent to keeping θ fixed. This constant value of θ can be taken as zero without loss of generality. Other fixed and known values of θ can be achieved by a translation of the Gamma type distribution into which the z distribution is transformed. The conditions being imposed on the parameters of transformation thus become:

1'. $a > 0$.

2'. $b^2 - 4ac = 0$.

Under these two conditions the transformation (1.2.5) will give a one-to-one correspondence between x and z over their respective ranges and consequently a one-to-one correspondence between t and z over their respective ranges. Further, the extremities of the ranges are corresponding values in both cases. The ranges of values are:

$$0 \leq x < \infty; \quad -A \leq t < \infty; \quad -\frac{b}{2a} \leq z < \infty .$$

1.2

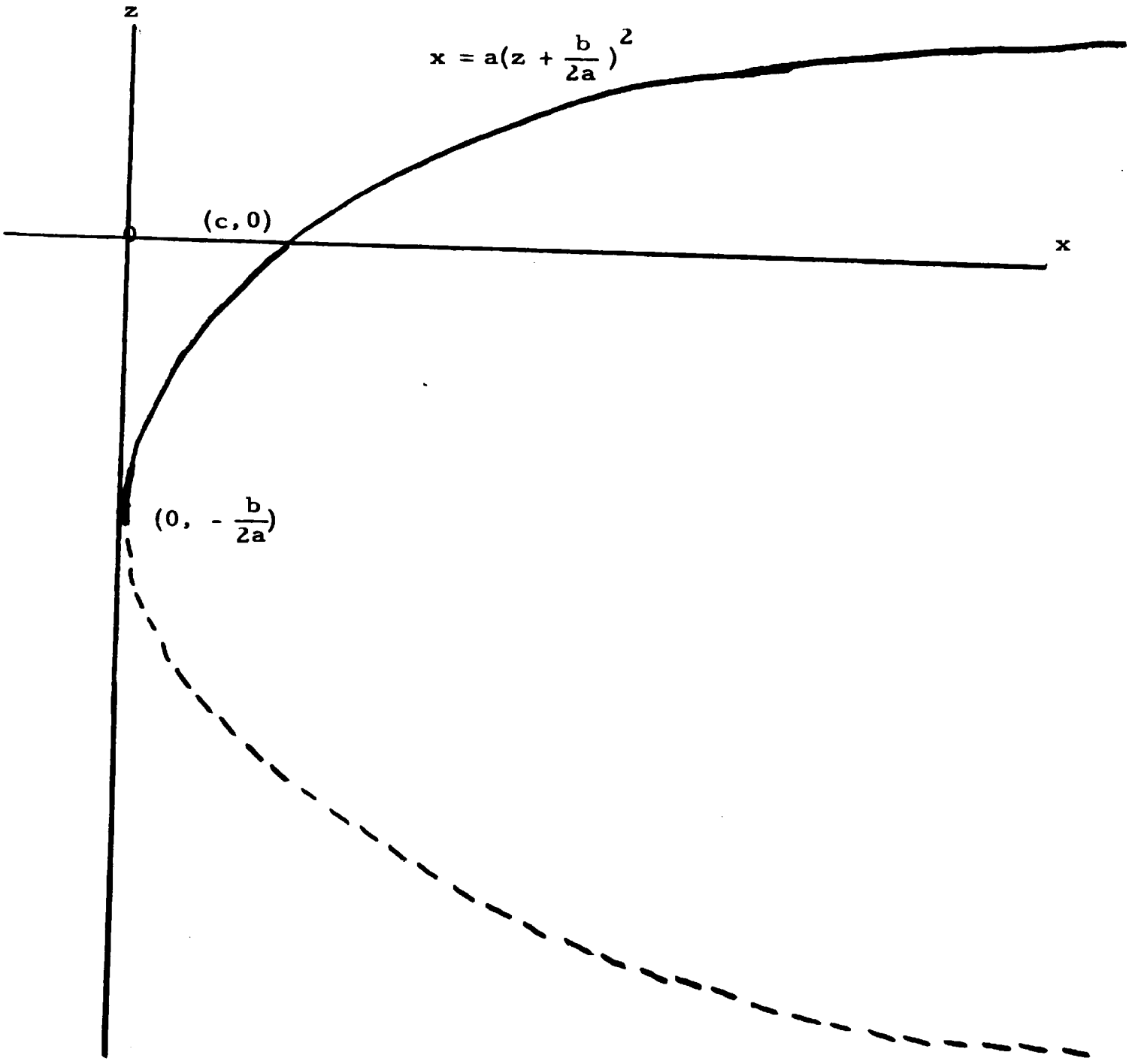


Figure 1

"The parabola of transformation". Values of x and z to be considered lie along the solid branch. The vertex may be above or below the x-axis depending on the sign of b. $c = \frac{b^2}{4a}$.

PART II

Analysis of the Z-function

2.1 Variation in the parameters resulting from a transformation of the z variates.

Theorem 2.1.1. c and θ are invariant under a change in scale for z. b and a are multiplied by the reciprocal of the scaling factor and the square of its reciprocal respectively.

Proof: Let $H(z, a, b, c) = az^2 + bz + c$, then from (1.2.3) $df(z, a, b, c) \equiv df(z/k, k^2a, kb, c)$ and thus z/k obeys the same probability distribution as z with a and b replaced by k^2a and kb respectively. However, if $a_1 = k^2a$; $b_1 = kb$; and $c_1 = c$; then $\theta_1 = \theta$, θ_1 being a function of a_1, b_1, and c_1.

Theorem 2.1.2. a and θ are invariant under a change in origin for z. b and c are replaced by b_2 and c_2 respectively in which $b_2 = b + 2ad$ and $c_2 = ad^2 + bd + c$, d being the amount by which z is displaced.

Proof: If θ_2 is a function of a, b_2, and c_2, then $\theta_2 = \theta$. Also if $H(z, a, b, c)$ is defined as in the previous theorem, then $H(z + d, a, b, c) \equiv H(z, a, b_2, c_2)$, and it follows $df(z + d, a, b, c) \equiv df(z, a, b_2, c_2)$. Thus z-d obeys the same probability distribution function as z with a and θ invariant, but b and c replaced by b_2 and c_2 respectively.

2.2 The effect on the cumulative distribution function of varying the parameters. This section is concerned with the parameters of the transformation and thus p will be regarded as fixed. Further, the variation in a, b, and c will be subject to the restriction that θ is to be kept constant.

2.2

Lemma 2.2.1 The probability of z lying within any finite interval of the lower end of the range is a monotonic increasing function of a, but is independent of b and c.

Proof: Let z_o be the lower end of the range for z and let α be any real number such that $\alpha > 0$. Then the required probability is given by

$$P(z_o \leq z \leq z_o + \alpha) = \frac{1}{\Gamma(p)} \int_{z_o}^{z_o+\alpha} (az^2 + bz + c)^{p-1} e^{-(az^2 + bz + c)} (2az + b)dz$$

in which $z_o = \frac{-b + \sqrt{4a\theta}}{2a}$.

Putting $y = az^2 + bz + c$ gives

$$P(z_o \leq z \leq z_o + \alpha) = \frac{1}{\Gamma(p)} \int_0^{a^{1/2}\alpha [a^{1/2}\alpha + 2\sqrt{\theta}]} y^{p-1} e^{-y} dy .$$

The above is obviously an increasing function of a for any $\alpha > 0$. It is equally obvious, since θ is constant, that the above function is independent of b and c.

No such specific conclusion can be isolated relative to the effect of a variation in b and/or c. The result of varying one or both of these is dependent upon the simultaneous variation taking place in the other parameters.

In general a change in b and/or c will result in a displacement of the lower end of the range. The magnitude and direction of this displacement cannot be classified according to the variation in any one parameter, but is dependent upon the joint variation of two or three parameters. It is a relatively simple operation to analyze this displacement for a given set of variations. It is also obvious from Theorem 2.1.1 that a change in b would be equivalent to a change in scale for the variate, unless accompanied by a counteracting change in a.

2.3

2.3 Forms assumed by the frequency function, f(z).

Subject to the condition 2' of section 1.2, the transformation (1.2.4) becomes

(2.3.1) $x = a(z + B)^2$, in which $B \equiv \frac{b}{2a}$, and the Z-function as given by (1.2.3) becomes

(2.3.2) $$f(z)\,dz = \frac{2a^p}{\sqrt{p'}}(z + B)^{2p-1} e^{-a(z+B)^2} dz. \quad -B \leq z < \infty.$$

The independent parameters are now a, p, and B, with the fourth parameter, c, being determined from the relationship $b^2 - 4ac = 0$.

The equation for the frequency curve is now given by

(2.3.3) $$f(z) = \frac{2a^p}{\sqrt{p'}}(z + B)^{2p-1} e^{-a(z+B)^2}$$ with the range for z being as in (2.3.2), i.e. $-B \leq z < \infty$.

If f'(z) and f''(z) represent $\frac{df(z)}{dz}$ and $\frac{d^2f(z)}{d^2(z)}$ respectively, then

(2.3.4) $$f'(z) = f(z)[(2p - 1)(z + B)^{-1} - 2a(z + B)]$$ and

(2.3.5) $$f''(z) = f(z)[(2p - 1)(z + B)^2(2p - 2) - 2a(4p - 1) + 4a^2(z + B)^2].$$

Table 1 below gives the values of f(z), f'(z), and f''(z) at the lower end of the range for values of p in various critical sub-intervals of the range for p, i.e. $0 < p < \infty$.

	p> 1.5	p=1.5	1<p<1.5	p=1	.5<p<1	p=.5	0<p<.5
f(-B)	0	0	0	0	0	$2a^{1/2}\pi^{-1/2}$	∞
f'(-B)	0	0	0	2a	∞	0	$-\infty$
f''(-B)	0	$8a^{3/2}\pi^{-1/2}$	∞	0	$-\infty$	$-4a^{3/2}\pi^{-1/2}$	$-\infty$

Table 1

2.3

Relative to the behavior of the frequency curve towards its right extremity, which is infinite, it is obvious that $f(z)$, $f'(z)$, and $f''(z)$ all three approach zero as z becomes infinite.

In the investigation of the behavior of the curve within the interval $-B < z < \infty$, let a critical value of z be defined as any value of z for which $f'(z)$ vanishes and let the corresponding point on the frequency curve be defined as a critical point of $f(z)$. Thus a critical point of $f(z)$ may be either a maximum point, a minimum point, or a point of inflection, depending on the sign of $f''(z)$.

Since $f(z) \neq 0$ for $z > -B$, in order to have a critical point for $f(z)$ inside the range from (2.3.4) it is necessary that $(2p - 1)(z + B)^{-1} - 2a(z + B)$ vanish. Then, since $z + B \neq 0$, the roots of $2a(z + B)^2 - (2p - 1) = 0$ will be critical values of z. Thus $z = \frac{-(2a)^{1/2} B \pm \sqrt{2p - 1}}{(2a)^{1/2}}$ will represent either a maximum point, a minimum point, or a point of inflection for $f(z)$.

Since $\frac{-(2a)^{1/2}B - \sqrt{2p - 1}}{(2a)^{1/2}} < -B$ and thus lies outside the range of admissible values for z, there will be at most one critical point for $f(z)$ within the range. That is $z_1 = \frac{-(2a)^{1/2}B + \sqrt{2p - 1}}{(2a)^{1/2}}$, if real, will be the abscissa of a critical value for $f(z)$.

If $p < 1/2$, z_1 is imaginary and if $p = 1/2$, $z_1 = -B$, the lower end of the range for z. Under neither of these conditions will there be a critical point for $f(z)$ within the range.

If $p > 1/2$, z_1 is within the range $-B < z < \infty$. From (2.3.5) $f''(z_1) = -4af(z_1)$. Thus since $f(z_1) > 0$ and from condition I[3] section 1.2, $a > 0$, $f''(z_1) < 0$, and $f(z_1)$ is a relative maximum value for the frequency curve.

2.3

The frequency curve takes on three distinct forms depending on the value of p. These forms are:

I. $p < 1/2$. A J-shaped curve with infinite ordinate at the lower end of the range and monotonically approaching the horizontal axis as an asymptote as z becomes infinite.

II. $p = 1/2$. The curve has a finite ordinate of $2a^{1/2}\pi^{-1/2}$, with slope of zero, at the lower end of the range. Further the curve monotonically approaches the horizontal axis as an asymptote as z becomes infinite. The curve has a point of inflection at a distance of $(2a)^{-1/2}$ from the lower end of the range.

III. $p > 1/2$. $f(-B) = 0$ and $f(z) \longrightarrow 0$ as $z \longrightarrow \infty$; also $f(z)$ has a modal ordinate of

$$\frac{2a^{1/2}}{\sqrt{p'}} (p - \frac{1}{2}) p^{-1/2} e^{-(p - 1/2)}$$

at a distance of $a^{-1/2}(p - 1/2)^{1/2}$ units to the right of the lower end of the range.

A typical curve of each of the five distinct forms which can be assumed is shown in Figure 2.

2.3

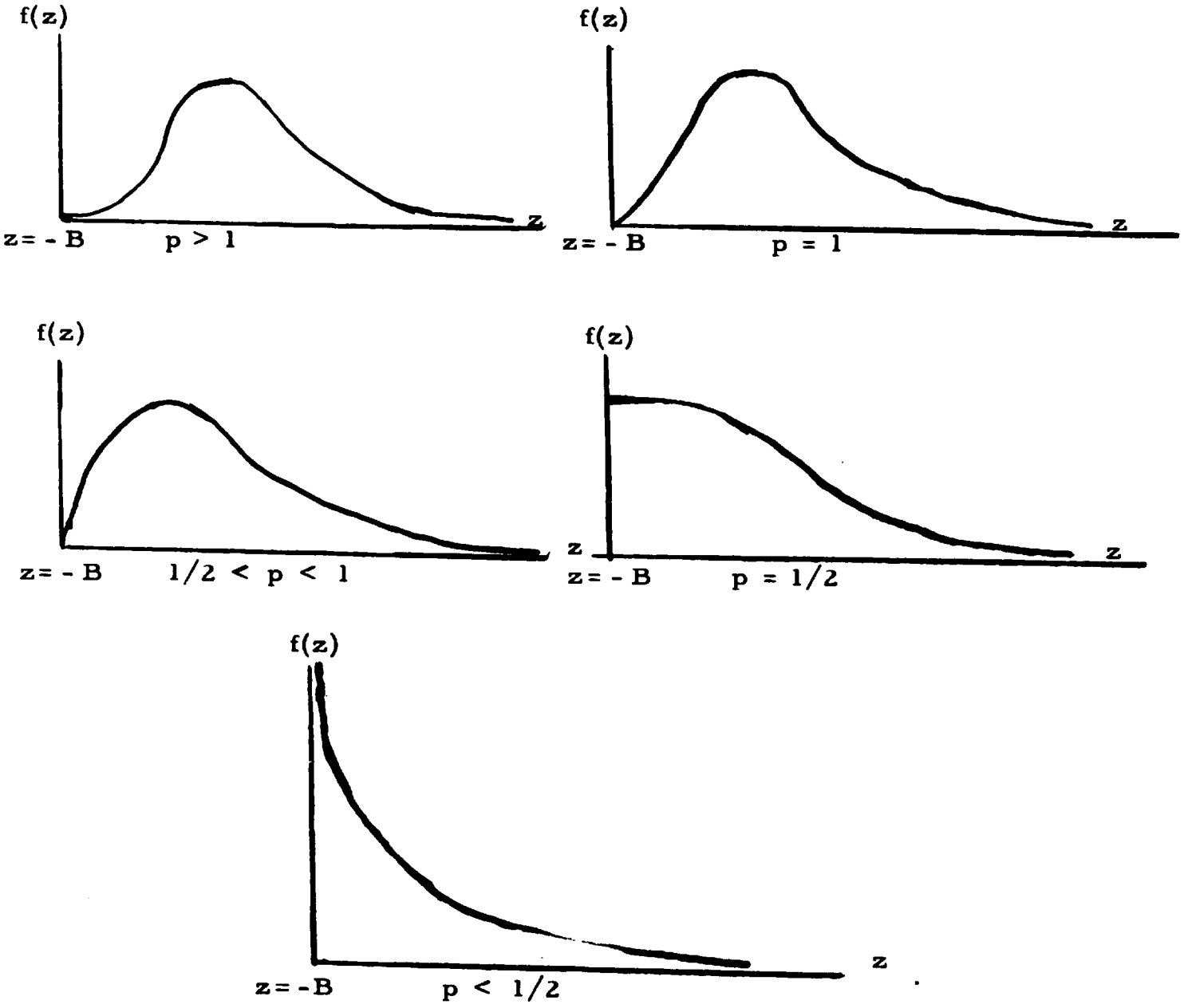


Figure 2.

Typical frequency curves for values of p in critical intervals. Ordinates are shown at the lower end of the range for z, i.e. z = - B, in all cases.

PART III

Estimation by Method of Moments

3.1 Equations* for the determination of estimates of the parameters and procedure for their solution. If ν_s represents the s^{th} population moment for z about $-B$, then

$$\nu_s = \frac{2a^p}{\Gamma(p)} \int_{-B}^{\infty} (z+B)^s (z+B)^{2p-1} e^{-a(z+B)^2} dz, \quad \text{or}$$

$$\nu_s = \frac{a^{-s/2}\,\Gamma(p + s/2)}{\Gamma(p)} \quad \text{which for } s = 1, 2, \text{ and } 3 \text{ gives}$$

$$\nu_1 = \frac{a^{-1/2}\,\Gamma(p+1/2)}{\Gamma(p)}\ ; \quad \nu_2 = a^{-1}p\ ; \text{ and } \quad \nu_3 = \frac{a^{-3/2}\,\Gamma(p+3/2)}{\Gamma(p)}$$

If μ'_s and μ_s represent the s^{th} moment for z about the origin and mean respectively, the relationships between the central moments and the parameters become

(3.1.1) $$f_o = a^{1/2}k\ .$$

(3.1.2) $$p = (\mu_2 + k^2)a$$

(3.1.3) $$(p + 1/2)f_o = (\mu_3 + 3\mu_2 k + k^3)a^{3/2} \quad \text{in which}$$

$$f_o = \frac{\Gamma(p+1/2)}{\Gamma(p)} \quad \text{and} \quad k = (\mu'_1 + B)$$

It will be noticed that $k = \nu_1$ is the first moment about $-B$. However, k will be treated as the independent parameter and B as the dependent one. The independent parameters to be estimated are a, p, and k. The moment estimates for these parameters will be the solution for the system of equations (3.1.1), (3.1.2), and (3.1.3). In the theoretical

* These equations will be referred to as the moment equations, and their solutions will be referred to as the moment estimates for the parameters.

3.1

work with these equations the population moments will continue to be used. However, in an applied problem in graduation or estimation from an observed sample the moments for the sample, $\bar{z}$, m_2, and m_3 would be used.

The moment equations, due primarily to the presence of the function f_o, do not lend themselves to an exact algebraic solution. Approximation to any desired degree of accuracy can be obtained by the procedure outlined below.

The elimination of k between (3.1.1) and (3.1.2) gives

(3.1.4) $\qquad \phi_o^2 = a\,\mu_2, \quad$ in which $\quad \phi_o^2 = p - f_o^2$.

The elimination of f_o between (3.1.1) and (3.1.3) with the aid of (3.1.2) gives, since $a \neq 0$ and can be used as a divisor,

(3.1.5) $\qquad k = 2a(\mu_3 + 2\mu_2 k)$.

Eliminating k from (3.1.5) gives

(3.1.6) $\qquad f_o = 2a(\mu_3 a^{1/2} + 2\mu_2 f_o)$.

Finally, the elimination of a gives as the equation for the determination of p

(3.1.7) $\qquad 2\alpha_3 \phi_o^3 + 4 f_o \phi_o^2 - f_o = 0 \quad$ in which $\quad \alpha_3 = \dfrac{\mu_3}{\mu_2^{3/2}}$.

The Tables of Appendix I give the value for each f_o, f_o^2, and ϕ_o^2 for values of p at small intervals from 0 through 50. By use of the formula $f_o(p+1) = \dfrac{p + 1/2}{p} f_o(p)$ the tables can easily be extended to other intermediate values of p or to values of p beyond the table.

Solving (3.1.7) for α_3 gives

(3.1.8) $\qquad \alpha_3 = \dfrac{f_o(1 - 4\phi_o^2)}{2\phi_o^3}$.

3.1-2

By use of the tables of the Appendix two values of p which yield from (3.1.8) a value of α_3 above and below the sample value can easily be located. A higher degree of accuracy can be obtained either by interpolation or actual computation of values of f_o and ϕ_o^2 for intermediate values of p. Thus by successive approximation the moment estimate of p can be obtained to any desired degree of accuracy. The existence and uniqueness of such an estimate will be discussed in a later section.

Knowing the moment estimate for p the moment estimates for the other parameters are easily obtained from the appropriate relationships. These relationships are all algebraic, in fact they are linear if proper selection and sequence is followed. The slight modification in technique made necessary in event of a negative skewness is discussed in Appendix II.

3.2 Study of the functions $f_o(p)$ and $\phi_o^2(p)$. Before investigating the roots of (3.1.7) it will be necessary to establish some of the properties of the functions of p which are involved in this equation. The study of these properties will be organized as a series of lemmas. Where there is no danger of ambiguity the argument, p, will be omitted.

Lemma 3.2.1. $\phi_o^2 = p - f_o^2$ is a continuous function of p over the interval $0 < p < \infty$.

Proof: Since $f_o \equiv \frac{\Gamma(p + 1/2)}{\Gamma p}$ and $p > 0$, each $\Gamma(p + 1/2)$ and Γp is continuous and Γp has no zero in the interval $0 < p < \infty$. Thus f_o^2 is continuous over the interval and ϕ_o^2 being the difference of two continuous functions is itself continuous.

Lemma 3.2.2. $\frac{d(\phi_o^2)}{dp} \leq 1$ for all $p > 0$ and $\lim_{p \to 0} \frac{d(\phi_o^2)}{dp} = 1$.

Proof: From the definition of ϕ_o^2

(3.2.1) $\frac{d(\phi_o^2)}{dp} = 1 - 2 f_o \frac{df_o}{dp}$ and from the definition of

3.2

f_o', $\log f_o = \log \Gamma(p + 1/2) - \log \Gamma(p)$, which gives

(3.2.2) $\quad \frac{df_o}{dp} = f_o [\psi(p + 1/2) - \psi(p)]$ in which

$$\psi(p + 1/2) = \frac{d[\log \Gamma(p + 1/2)]}{dp}$$

Replacing $\psi(p)$ and $\psi(p + 1/2)$ by their series representation (Ref. 7) gives

$$\frac{df_o}{dp} = f_o \left\{ - C - \frac{1}{p+1/2} + (p+1/2) \sum_{s=1}^{\infty} \frac{1}{s(p+1/2+s)} + C + \frac{1}{p} - p \sum_{s=1}^{\infty} \frac{1}{s(p+s)} \right\}$$

in which C is the Euler constant. Thus

(3.2.3) $\quad \frac{df_o}{dp} = f_o \sum_{s=o}^{\infty} \frac{1}{2(p + s)(p + s + 1/2)}$. Substituting this in

(3.2.1) gives

(3.2.4) $\quad \frac{d(\phi_o^2)}{dp} = 1 - f_o^2 \sum_{s=o}^{\infty} \frac{1}{(p + s)(p + s + 1/2)} \leq 1$ for all $p > 0$.

Thus the first part of the lemma is established.

Now to establish the second part of the lemma.

$$f_o^2 \sum_{s=o}^{\infty} \frac{1}{(p + s)(p + s + 1/2)} =$$

$$= f_o \left\{ \frac{\Gamma(p + 1/2)}{(p + 1/2)\Gamma(p + 1)} + \frac{\Gamma(p + 1/2)}{\Gamma(p)} \sum_{s=1}^{\infty} \frac{1}{(p + s)(p + s + 1/2)} \right\}$$

$\lim_{p \to 0} \frac{\Gamma(p + 1/2)}{(p+1/2)\Gamma(p+1)} = 2\sqrt{\pi}$, and $\lim_{p \to 0} f_o = \lim_{p \to 0} \frac{\Gamma(p + 1/2)}{\Gamma(p)} = 0.$

$$\lim_{p \to 0} \sum_{s=1}^{\infty} \frac{1}{(p + s)(p + s + 1/2)} = \sum_{s=1}^{\infty} \frac{1}{s(s + 1/2)} < \sum_{s=1}^{\infty} \frac{1}{s^2} = \frac{\pi^2}{6}.$$

Combining the above limits give

$$\lim_{p \to 0} f_o^2 \sum_{s=o}^{\infty} \frac{1}{(p + s)(p + s + 1/2)} = 0.$$

Consequently from (3.2.4)

3.2

$$\lim_{p \to 0} \frac{d(\phi_o^2)}{dp} = 1$$ and the lemma is established.

Lemma 3.2.3. ϕ_o^2 is a monotonic increasing function of p.

Proof: From (3.2.1) $\frac{d(\phi_o^2)}{dp} = 1 - 2f_o \frac{df_o}{dp}$ and thus the lemma will be established by showing $2f_o \frac{df_o}{dp} < 1$ for all $p > 0$. It will be necessary to divide the proof into two cases according to whether $p \leq 1$ or $p > 1$.

Case 1. $p \leq 1$. By writing $p + 1/2 = q$, (3.2.3) becomes

$$2f_o \frac{df_o}{dp} = f_o^2 \sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)} .$$

Differentiating the right member with respect to p gives

$$(3.2.5) \qquad \frac{d[2f_o \frac{df_o}{dp}]}{dp} =$$

$$= f_o^2 \left\{ \left[\sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)} \right]^2 - \left[2 \sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)^2} + \frac{1}{2} \sum_{s=o}^{\infty} \frac{1}{(p+s)^2(q+s)^2} \right] \right\}$$

In order to show that the right member of (3.2.5) is positive for all $p \leq 1$ its terms will first be considered individually.

Neglecting the exponent, the first term can be written

$$\sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)} =$$

$$= \frac{1}{pq} + \frac{1}{(p+1)(q+1)} + \frac{1}{(p+2)(q+2)} + \sum_{s=3}^{\infty} \frac{1}{(p+s)(q+s)} .$$

Since $q + s < p + s + 1$, it follows that

$$\sum_{s=3}^{\infty} \frac{1}{(p+s)(q+s)} > \sum_{s=3}^{\infty} \frac{1}{(p+s)(p+s+1)} .$$

Resolving the right member into partial fractions and summing by telescoping gives

3.2

$$\sum_{s=3}^{\infty} \frac{1}{(p+s)(p+s+1)} = \frac{1}{p+3} .$$

Combining the above gives for the first term

$$\sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)} > \frac{1}{pq} + \frac{1}{(p+1)(q+1)} + \frac{1}{(p+2)(q+2)} + \frac{1}{p+3} .$$

For the second term the following inequality is obvious:

$$2\sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)^2} < 2\left[\frac{1}{pq^2} + \frac{1}{(p+1)(q+1)^2} + \frac{1}{(p+2)(q+2)^2} + \sum_{s=3}^{\infty} \frac{1}{(p+3)^3}\right] .$$

By the use of Euler's summation formula (Ref. 18)

$$\sum_{s=3}^{\infty} \frac{1}{(p+3)^3} < \frac{1}{2(p+3)^2} + \frac{1}{2(p+3)^3} + \frac{1}{4(p+3)^4} .$$

Combining these two gives for the second term

$$2\sum_{s=o}^{\infty} \frac{1}{(p+s)(q+s)^2} < 2\left[\frac{1}{pq^2} + \frac{1}{(p+1)(q+1)^2} + \frac{1}{(p+2)(q+2)^2}\right]$$

$$+ \frac{1}{(p+3)^2} + \frac{1}{(p+3)^3} + \frac{1}{2(p+3)^4}$$

In like manner it can be shown for the third term

$$\frac{1}{2}\sum_{s=o}^{\infty} \frac{1}{[(p+s)(q+s)]^2} < \frac{1}{2}\left[\frac{1}{p^2q^2} + \frac{1}{(p+1)^2(q+1)^2} + \frac{1}{(p+2)^2(q+2)^2}\right.$$

$$\left. + \frac{1}{3(p+3)^3} + \frac{1}{2(p+3)^4} + \frac{1}{3(p+3)^5}\right] .$$

Combining the above in (3.2.5) and collecting into one fraction gives

3.2

$$\frac{d\left[2f_o \frac{df_o}{dp}\right]}{dp} > -f_o^2\left[896p^{14} + 20,544p^{13} + 211,872p^{12} + 1,297,776p^{11}\right.$$
$$+ 5,241,400p^{10} + 14,585,196p^9 + 28,041,886p^8 + 34,929,921p^7$$
$$+ 19,140,159p^6 - 22,896,777p^5 - 68,822,245p^4 - 85,558,440p^3$$
$$\left. - 63,487,620p^2 - 27,293,760p - 5,248,800\right]$$
$$\left[768p^2(p+1)^2(p+2)^2(p+3)^5 q^2(q+1)^2(q+2)^2\right]^{-1}.$$

It is obvious that the right member above is positive for any value of p such that $p \leq 1$. Thus $2f_o \frac{df_o}{dp}$ is a monotonic increasing function of p for $p \leq 1$.

Evaluating $f_o^2(1)$ from the table of Appendix I and $\psi(1.5)$ and $\psi(1)$ from any table of digamma functions and putting these values in (3.2.2) gives

$$2f_o \frac{df_o}{dp}\Big|_{p=1} = .9645 < 1.$$

Thus since $2f_o \frac{df_o}{dp}$ is a monotonic increasing function of p for $p \leq 1$ and has a value less than one for $p = 1$ it follows that

$$2f_o \frac{df_o}{dp} < 1 \text{ for all } p \leq 1$$

and thus from (3.2.1)

$$\frac{d\phi_o^2}{dp} = 1 - 2f_o \frac{df_o}{dp} > 0 \text{ for } p \leq 1.$$

Consequently ϕ_o^2 is a monotonic increasing function of p for all $p \leq 1$.

Case 2. $p > 1$. From (3.2.3) $\frac{df_o}{dp} = f_o \sum_{s=o}^{\infty} \frac{1}{2(p+s)(q+s)}$.

Resolving into partial fractions gives

$$\frac{df_o}{dp} = f_o\left\{\sum_{s=o}^{\infty}\left[\frac{1}{(p+s)} - \frac{1}{(q+s)}\right]\right\}.$$

Using the Euler summation formula gives as an asymptotic expansion

3.2

$$(3.2.6)\quad \frac{df_o}{dp} = f_o\left[\log\left(1+\frac{1}{2p}\right) + \frac{1}{4pq} + \sum_{r=1}^{\infty} \frac{B_{2r}\,(p^{2r} - q^{2r})}{2r\,(pq)^{2r}}\right]$$

in which B_{2r} is the (2r)th. Bernoulli number with the ordering being as given in Knopp (Ref. 18). Since the right member of (3.2.6) is an asymptotic expansion,

$$(3.2.7)\quad \frac{df_o}{dp} < f_o\left[\log\left(1+\frac{1}{2p}\right) + \frac{1}{4qp} + \frac{4p+1}{48(pq)^2}\right] \quad \text{and}$$

$$(3.2.8)\quad \frac{df_o}{dp} > f_o\left[\log\left(1+\frac{1}{2p}\right) + \frac{1}{4pq} + \frac{4p+1}{48(pq)^2} - \frac{32p^3+24p^2+8p+1}{120p^4(2p+1)^4}\right] .$$

Since $f_o > 0$ both sides of (3.2.7) can be multiplied by $2f_o$ without changing the sense of the inequality. This gives

$$(3.2.9)\quad 2f_o\frac{df_o}{dp} < f_o^2\left[2\log\left(1+\frac{1}{2p}\right) + \frac{1}{2pq} + \frac{4p+1}{24(pq)^2}\right] .$$

By definition $f_o = \frac{\sqrt{q}}{\sqrt{p}} = \frac{p\sqrt{(q+1)}}{q\sqrt{(p+1)}}$.

Replacing the Gamma functions of f_o^2 by their values as given by Stirling's formula and simplifying gives

$$(3.2.10)\quad f_o^2 = p\left(1+\frac{1}{2p}\right)^{2p} e^{-1}\, e^{\frac{p\theta_1 - q\theta_2}{6pq}} \qquad \begin{matrix} 0 < \theta_1 < 1 \\ 0 < \theta_2 < 1 . \end{matrix}$$

Let the right member of (3.2.9) be designated by N(p). Then replacing f_o^2 in N(p) by its value as given by (3.2.10) and taking the limit gives

$$(3.2.11)\quad \lim_{p\to\infty} N(p) = 1$$

$$\frac{dN(p)}{dp} = 2f_o\frac{df_o}{dp}\left[2\log\left(1+\frac{1}{2p}\right) + \frac{1}{2pq} + \frac{4p+1}{24p^2q^2}\right]$$

$$- f_o^2\left[\frac{1}{pq} + \frac{4p+1}{4p^2q^2} + \frac{12p^2+6p+1}{24p^3q^3}\right]$$

3.2

The coefficient of $\frac{df_o}{dp}$ is positive in the above expression and thus replacing $\frac{df_o}{dp}$ by the right member of (3.2.8) gives

$$\frac{dN(p)}{dp} > f_o^2 \left\{ [A(p)]^2 - B(p) - C(p) \right\} \quad \text{in which}$$

$$A(p) = 2 \log \left(1 + \frac{1}{2p}\right) + \frac{1}{2pq} + \frac{4p+1}{24p^2q^2}$$

$$B(p) = \left[\frac{32p^3 + 24p^2 + 8p + 1}{960\, p^4 q^4}\right] \left[A(p)\right]$$

$$C(p) = \frac{1}{pq} + \frac{4p+1}{4p^2q^2} + \frac{12p^2 + 6p + 1}{24p^3 q^3} .$$

Since $p > 1$, $\frac{1}{2p} < 1$, and thus

$$2 \log \left(1 + \frac{1}{2p}\right) > \frac{1}{p} - \frac{1}{4p^2} + \frac{1}{12p^3} - \frac{1}{32p^4}$$

by the property of alternating series since the first term omitted is positive. Similarly,

$$2 \log \left(1 + \frac{1}{2p}\right) < \frac{1}{p} - \frac{1}{4p^2} + \frac{1}{12p^3}$$

since the first term omitted is negative. Thus

$$A(p) > \frac{1}{p} - \frac{1}{4p^2} + \frac{1}{12p^3} + \frac{1}{2pq} + \frac{4p+1}{24p^2q^2} - \frac{1}{32p^4}$$

$$B(p) < \left[\frac{32p^3 + 24p^2 + 8p + 1}{960\, p^4 q^4}\right]\left[\frac{1}{p} - \frac{1}{4p^2} + \frac{1}{12p^3} + \frac{1}{2pq} + \frac{4p+1}{24p^2q^2}\right] .$$

Clearing the above of fractions and collecting like terms of A(p), B(p), and C(p) gives

$$[A(p)]^2 - B(p) - C(p) \geq (184320p^{10} + 737280p^9 + 731136p^8 + 79872p^7 - 285888p^6 - 210240p^5 - 70224p^4 - 11488p^3 - 324p^2 + 236p + 45)(320 \cdot 96^2 p^8 q^6)^{-1} .$$

3.2

From the above it is obvious that $[A(p)]^2 - B(p) - C(p) > 0$ for $p > 1$ and consequently $\frac{dN(p)}{dp} > 0$ for $p > 1$. Thus $N(p)$ is a monotonic increasing function of p for all $p > 1$. But from (3.2.11) $\lim_{p\to\infty} N(p) = 1$ and consequently $N(p) < 1$ for all p such that $p > 1$.

From the definition of $N(p)$ and (3.2.9) for all $p > 1$

$$2f_o \frac{df_o}{dp} < N(p) < 1 \quad \text{for } p > 1 \text{ and thus}$$

$$\frac{d\phi_o^2}{dp} = 1 - 2f_o \frac{df_o}{dp} > 0 \quad \text{for all } p > 1.$$

Therefore ϕ_o^2 is a monotonic increasing function of p for $p > 1$.

Combining the results of case 1 and case 2 establishes the lemma.

Lemma 3.2.4 $\phi_o^2 > \frac{p}{4p+1}$ and $\lim_{p\to o} \phi_o^2 = 0$.

Proof: From Lemma 3.2.3 $\phi_o^2(p+1) - \phi_o^2(p) > 0$ for all $p > 0$ and from definition $\phi_o^2 = p - f_o^2(p)$ and thus

$$\phi_o^2(p+1) = p + 1 - f_o^2(p+1) = p + 1 - \frac{q^2}{p^2} f_o^2(p) \quad \text{in which } q = p + \frac{1}{2}$$

and thus

$$\phi_o^2(p+1) - \phi_o^2(p) = 1 - \frac{4p+1}{4p^2} f_o^2 > 0.$$

Consequently

$$f_o^2 < \frac{4p^2}{4p+1} \quad \text{or} \quad \phi_o^2(p) > p - \frac{4p^2}{4p+1}$$

which gives

$$\phi_o^2 > \frac{p}{4p+1}$$

and establishes the first part of the lemma.

To establish the second part, it is obvious from the definition of f_o that its limit as $p \to 0$ is zero and thus

3.2

$$\lim_{p\to o}[\phi_o^2] = \lim_{p\to o}(p - f_o^2) = 0.$$

Lemma 3.2.5 $\phi_o^2 = p - f_o^2$ is a bounded function of p for all $p > 0$.

Proof: Lemma 3.2.4 establishes zero as the lower bound. Consequently it is necessary to show only that ϕ_o^2 has an upper bound. Further, as a result of Lemma 3.2.3 it is necessary to consider only large values of p, or specifically the limit of $\phi_o^2(p)$ as p becomes infinite.

From (3.2.10) is obtained

(3.2.12) $$f_o^2(p) = p\left(1 + \frac{1}{2p}\right)^{2p} e^{-1} + \left(1 + \frac{1}{2p}\right)^{2p} e^{-1}\delta \qquad \text{in which}$$

$$\delta = \sum_{i=1}^{\infty} \frac{(p\theta_1 - q\theta_2)^i}{i!(6q)^i p^{i-1}} .$$

This series converges for all $p > 0$ and consequently δ is bounded for all such p.

For $p > \frac{1}{2}$

$$\log\left(1 + \frac{1}{2p}\right) = 1 - \frac{1}{2p}\sum_{i=o}^{\infty}(-1)^i \frac{1}{(2+i)(2p)^i} = 1 - \frac{1}{2p}Q(p)$$

in which Q(p) is finite for all $p > \frac{1}{2}$. Then

(3.2.13) $$\left(1 + \frac{1}{2p}\right)^{2p} e^{-1} = e^{-\frac{Q(p)}{2p}} \qquad \text{which gives}$$

$$e^{-\frac{Q(p)}{2p}} = 1 - \frac{Q(p)}{2p}\sum_{i=1}^{\infty}(-1)^i\left[\frac{Q(p)}{2p}\right]^{i-1}\frac{1}{i!} .$$

Since Q(p) is finite for all $p > \frac{1}{2}$ the above series will converge absolutely over this range. Thus (3.2.13) can be written

(3.2.14) $$\left(1 + \frac{1}{2p}\right)^{2p} e^{-1} = 1 - \frac{\varepsilon}{p}$$

in which ε is a bounded function of p for all $p > \frac{1}{2}$. Then the substi-

3.2

tution of (3.2.14) in (3.2.12) gives,

$$f_o^2(p) = p - \varepsilon + \delta - \frac{\varepsilon\delta}{p}, \quad \text{and from definition of } \phi_o^2,$$

$$\phi_o^2 = \varepsilon - \delta + \frac{\varepsilon\delta}{p}.$$

Since each ε and δ remain finite for all values of $p > \frac{1}{2}$, ϕ_o^2 remains finite for all $p > \frac{1}{2}$ and specifically the limit of ϕ_o^2 is finite as p becomes infinite. Thus ϕ_o^2 is a bounded function of p.

Lemma 3.2.6. $\phi_o^2 = p - f_o^2 < \frac{1}{4}$ for all $p > 0$ and $\lim_{p \to \infty} \phi_o^2 = \frac{1}{4}$.

Proof: Let the function ε_m of p and m be defined as follows: For any positive integer, m,

$$\phi_o^2(p + m) = p + m - f_o^2(p + m) = \frac{1}{4} - \varepsilon_m.$$

Thus ε_o is a function of p alone, but for any other value of m, ε_m will be a function of both variables. Let p take on any value in the interval $0 < p \leq 1$ and for each value of p in this interval let m take on the values 0, 1, 2, - - - - - -. In this way p + m will cover the interval $0 < p + m < \infty$.

Then from Lemma 3.2.3,

(3.2.15) $\phi_o^2(p + m + 1) - \phi_o^2(p + m) = \varepsilon_m - \varepsilon_{m+1} \geq 0$ for any p and every m. Also from the definition of f_o^2 and ϕ_o^2, recalling that $q = p + \frac{1}{2}$,

$$\phi_o^2(p+m+1) = p+m+1 - \left[\frac{\lceil (q+m+1) \rceil}{\sqrt{(p+m+1)}}\right]^2 = p+m+1 - \frac{(q+m)^2}{(p+m)^2} f_o^2(p+m)$$

The right member above clears up to give

$$\phi_o^2(p + m + 1) = \frac{1}{4} - \left\{\varepsilon_m \frac{[2(p + m) + 1]^2}{4(p + m)^2} - \frac{1}{16(p + m)^2}\right\}.$$

3.2

From the definition of ε_{m+1} this gives

$$(3.2.16)\qquad \varepsilon_{m+1} = \varepsilon_m \frac{[2(p+m)+1]^2}{4(p+m)^2} - \frac{1}{16(p+m)^2} .$$

If m is replaced by $m+1$ in the above and the resulting ε_{m+1} is replaced by its value as given by (3.2.16), the result is

$$\varepsilon_{m+2} = \varepsilon_m \frac{[2(p+m)+1]^2[2(p+m+1)+1]^2}{4(p+m)^2\, 4(p+m+1)^2} - \frac{[2(p+m+1)+1]^2}{4(p+m+1)^2\, 16(p+m)^2} - \frac{1}{16(p+m+1)^2} .$$

Continuing this process gives for $n = 1, 2, 3, - - - - -$

$$(3.2.17)\qquad \varepsilon_{m+n} = \varepsilon_m \prod_{i=0}^{n-1} \frac{[2(p+m+i)+1]^2}{4(p+m+i)^2} - \sum_{i=o}^{n-1} \frac{1}{16(p+m+i)^2} \prod_{j=o}^{n-2-i} \frac{[2(p+m+1+i+j)+1]^2}{4(p+m+1+i+j)^2}$$

in which the last product is defined as 1 when $i = n-1$.

Suppose that for some p, say p_o, and for some m, the corresponding ε_m, say ε'_m, the inequality $\varepsilon'_m < 0$ is true. Then, if $\varepsilon'_m = -\delta_o$, we have $\delta_o > 0$.

If ε'_{m+n} is the corresponding function of $p_o + m + n$, then from (3.2.15) and the above assumption

$$\varepsilon'_{m+n} \le \varepsilon'_m < 0. \qquad \text{If } -\delta_n = \varepsilon'_{m+n}, \text{ then } \delta_n > 0. \qquad \text{From}$$

(3.2.17), if $p_1 = p_o + m$,

3.2

$$(3.2.18)\quad \delta_n = \delta_o \prod_{i=o}^{n-1} \frac{[2(p_1+i)+1]^2}{4(p_1+i)^2} + \sum_{i=o}^{n-1} \frac{1}{16(p_1+i)^2} \prod_{j=o}^{n-2-i} \frac{[2(p_1+i+j+1)+1]^2}{4(p_1+i+j+1)^2} .$$

Since $\delta_o > 0$, every factor of both terms of the right member above is positive. Consider now the coefficient of δ_o and its behavior with increasing n.

$$\prod_{i=o}^{n-1} \frac{[2(p_1+i)+1]^2}{4(p_1+i)^2} = \prod_{i=o}^{n-1} \left\{ 1 + \frac{4(p_1+i)+1}{4(p_1+i)^2} \right\}$$ which is of the

form $1 + u_i$ in which u_i is of the order -1 in $(p_1 + i)$ and, consequently, $\sum_{i=o}^{\infty} u_i$ is divergent while $\sum_{i=o}^{\infty} u_i^2$ is convergent. Thus from a well-known theorem (Ref. 18) the infinite product $\prod_{i=o}^{\infty} (1 + u_i)$ is divergent and as a consequence

$$\lim_{n \to \infty} \delta_n = \infty .$$

Thus if any ε_m is negative for any p or any m, it will lead to an unbounded sequence of ε_m's and will result in an unbounded value of ϕ_o^2 for some p. This contradicts Lemma 3.2.5. Therefore, $\varepsilon_m > 0$ for any p and all m and as a consequence

$$\phi_o^2 < \frac{1}{4} \quad \text{for all } p > 0.$$

To establish the second part of the lemma, from (3.2.15) and (3.2.16)

$$\varepsilon_m \frac{[2(p+m)+1]^2}{4(p+m)^2} - \frac{1}{16(p+m)^2} < \varepsilon_m .$$ Which gives

$$\varepsilon_m < \frac{1}{4[4(p+m)+1]} \quad \text{for any } p \text{ and every } m.$$

Since $\varepsilon_m > 0$ for any p and every m, then obviously

$\lim_{(p+m) \to \infty} \varepsilon_m = 0$ and it follows from the definition of ε_m

3.2-3

$$\lim_{p \to \infty} \phi_o^2 = \frac{1}{4} .$$

Lemma 3.2.7. $\lim_{p \to \infty} f_o(4\phi_o^2 - 1) = 0$.

Proof: From Lemma 3.2.4, $4\phi_o^2 > \frac{4p}{4p+1}$ for all $p > 0$. Since $f_o > 0$ it can be used as a multiplier, giving

$\frac{4f_o p}{4p+1} - f_o < 4f_o\phi_o^2 - f_o$. From Lemma 3.2.6, $4\phi_o^2 < 1$, or $4f_o\phi_o^2 - f_o < 0$. Combining these two inequalities gives

$$(3.2.19) \qquad -\frac{1}{4}\left[\frac{f_o}{p+1/4}\right] < 4f_o\phi_o^2 - f_o < 0 .$$

From (3.2.10)

$$f_o = p^{1/2}\left[\left(1+\frac{1}{2p}\right)^{2p} e^{-1}\right]^{1/2} e^{\frac{p\theta_1 - q\theta_2}{12pq}} \quad 0 < \theta_1 < 1,\ 0 < \theta_2 < 1$$

and thus

$$\frac{f_o}{p} = \left\{\left[\left(1+\frac{1}{2p}\right)^{2p} e^{-1}\right]^{1/2} e^{\frac{p\theta_1 - q\theta_2}{12pq}}\right\} p^{-1/2} .$$

It follows from this by the use of (3.2.14) that

$$\lim_{p \to \infty} \frac{f_o}{p} = 0.$$

Thus from (3.2.19), $4f_o\phi_o^2 - f_o$ for all values of p lies between zero and a quantity whose limit is zero as p becomes infinite, and it follows,

$$\lim_{p \to \infty} (4f_o\phi_o^2 - f_o) = 0.$$

3.3 The existence of unique moment estimates of the parameters.

In section 3.1 it was shown that the moment estimates for p will satisfy the equation $2\alpha_3\phi_o^3 + 4f_o\phi_o^2 - f_o = 0$. From Lemma 3.2.3, ϕ_o^2 is a single-valued function of p. Thus showing that the above equation has a unique root for ϕ_o^2 will be equivalent to showing that there exists

3.3

a unique moment estimate for p. Further, the equations for the determination of the other parameters, once p is known, yield a unique value in each case. Thus the existence of a unique estimate for p assures the existence of a unique estimate of each of the other parameters. Consequently, the establishment of the following theorem will be equivalent to proving the existence of a unique set of moment estimates of the parameters.

Theorem 3.3.1. The moment equation (3.1.7), i.e. $2\alpha_3\phi_o^3 + 4f_o\phi_o^2 - f_o = 0$, always has one and only one admissible root for ϕ_o for any finite α_3.

Proof: An admissible root for ϕ_o is defined as a value of ϕ_o such that the corresponding value of ϕ_o^2 will lie within the range of values as established in the previous Lemmas. That is

$$\frac{p}{4p+1} < \phi_o^2 < \frac{1}{4},$$ with upper and lower limits for ϕ_o^2 being 1/4 and 0, respectively.

Writing

$$F(\alpha_3, \phi_o, f_o) = 2\alpha_3\phi_o^3 + 4f_o\phi_o^2 - f_o \quad \text{gives}$$

$$F(\alpha_3, \phi_o, f_o) \equiv F(-\alpha_3, -\phi_o, f_o).$$

Thus a change in the sign for α_3 will result in a change in the sign for the root for ϕ_o, but will not change the absolute value of the root. Since p is determined from ϕ_o^2, the value of p will be invariant under a change in the sign of α_3. Thus it will be necessary to consider only $\alpha_3 > 0$. As previously mentioned, a discussion of the modification necessary to adjust for a negative skewness will be found in Appendix II.

From Lemma 3.2.3, $2\alpha_3\phi_o^3$ is a monotonic increasing function of p. Further, from Lemma 3.2.6, if the positive root for ϕ_o is used

3.3

by taking p sufficiently large $2\alpha_3\phi_o^3$ can be made arbitrarily close to $\frac{\alpha_3}{4}$.

From Lemma 3.2.7, $4f_o\phi_o^2 - f_o$ can be made arbitrarily small by taking p sufficiently large. Consequently, for any $\alpha_3 > 0$, there will exist some $p > 0$, say p_o, such that

(3.3.1) $$F[\alpha_3, \phi_o(p_o), f_o(p_o)] > 0 .$$

Since $f_o > 0$ for all $p > 0$, $F(\alpha_3, \phi_o, f_o)$ and $\frac{F(\alpha_3, \phi_o, f_o)}{f_o}$ as functions of p will be positive, or negative, or vanish together.

$$\frac{F(\alpha_3, \phi_o, f_o)}{f_o} = \frac{2\alpha_3\phi_o^3}{f_o} + 4\phi_o^2 - 1.$$

From the definition of ϕ_o^2,

$$\frac{\phi_o^3}{f_o} = \phi_o\left[\frac{p}{f_o} - f_o\right] = \phi_o\left[\frac{\sqrt{(p+1)}}{\sqrt{p}} - f_o\right] .$$

From Lemma 3.2.4,

$$\lim_{p\to o}[\phi_o] = 0.$$

From the definition of f_o, it is obvious

$$\lim_{p\to o} f_o = 0. \quad \text{Also} \quad \lim_{p\to o}\frac{\sqrt{(p+1)}}{\sqrt{q}} = \frac{1}{\sqrt{\pi}} .$$

Combining the above gives

$$\lim_{p\to o}\frac{\phi_o^3}{f_o} = 0 .$$

Thus for any finite $\alpha_3 > 0$, by the selection of sufficiently small value of p, $\frac{2\alpha_3\phi_o^3}{f_o} + 4\phi_o^2$ can be made less than any assigned quantity, and specifically it can be made less than one. Thus there will exist some p, say p_1, such that

3.3

$$\frac{F[\alpha_3, \phi_o(p_1), f_o(p_1)]}{f_o} < 0, \quad \text{or, since } f_o > 0 \text{ for all } p > 0,$$

$$F[\alpha_3, \phi_o(p_1), f_o(p_1)] < 0 . \tag{3.3.2}$$

From Lemma 3.2.1, ϕ_o^2 and f_o are continuous functions of p for all $p > 0$ and, consequently, $F(\alpha_3, \phi_o, f_o)$ is a continous function of p over the same range. Then in view of (3.3.1) and (3.3.2) there will be some value of p, say p_m, such that $p_1 < p_m < p_o$, and $F[\alpha_3, \phi_o(p_m), f_o(p_m)] = 0$. Further, each p_1 and p_o is greater than zero and p_m being between them is obviously greater than zero. That is, p_m is an admissible value of p. Therefore, for any finite value of α_3, the moment equation,

$2\alpha_3\phi_o^3 + 4f_o\phi_o^2 - f_o = 0$, will have at least one admissible root for p.

To establish the uniqueness of the root for $F(\alpha_3, \phi_o, f_o) = 0$, ϕ_o will be regarded as the independent variable, and f_o will be regarded as an independent parameter, subject only to the condition $f_o > 0$. Under this interpretation $F(\alpha_3, \phi_o, f_o)$ becomes a two parameter family of curves. Since f_o and ϕ_o are mutually dependent through p, the true curve for $F(\alpha_3, \phi_o, f_o)$, for any fixed α_3 will be a single member of this family of curves. Consequently, this interpretation cannot reduce the number of zeros for $F(\alpha_3, \phi_o, f_o)$. As previously noted, it will be necessary to consider only the case $\alpha_3 > 0$. Since ϕ_o is now being regarded as the independent variable, $F(\alpha_3, \phi_o, f_o)$ will be designated by $F(\phi_o)$, and $F'(\phi_o)$ and $F''(\phi_o)$ will designate its first and second partial derivative, respectively, with respect to ϕ_o. Then

$$F(\phi_o) = 2\alpha_3\phi_o^3 + 4f_o\phi_o^2 - f_o$$

$$F'(\phi_o) = 6\alpha_3\phi_o^2 + 8f_o\phi_o$$

3.3

$$F''(\phi_o) = 12\alpha_3\phi_o + 8f_o$$

Since $f_o > 0$ and $\alpha_3 > 0$, $F(\phi_o)$ will have a relative minimum at $\phi_o = 0$ and a relative maximum at $\phi_o = -\frac{4f_o}{3\alpha_3}$. Then

$$F(0) = -f_o < 0 \quad \text{and}$$

$$F(-\frac{4f_o}{3\alpha_3}) = \frac{f_o(64f_o^2 - 27\alpha_3^2)}{27\alpha_3^2}$$

The graph for $F(\phi_o)$ will have three possible forms depending on the value of $64f_o^2 - 27\alpha_3^2$, since each f_o and α_3 is positive.

Case 1. $64f_o^2 - 27\alpha_3^2 < 0$. The relative maximum point on the curve is below the horizontal axis and there is only one zero for $F(\phi_o)$. Consequently, $F(\alpha_3, \phi_o, f_o) = 0$ has only one root and the theorem is established since this root has been shown to be an admissible one in the first part of the proof. A typical curve is shown in Figure 3a., page 34.

Case 2. $64f_o^2 - 27\alpha_3^2 = 0$. Under this condition $F(\phi_o)$ will have one positive zero and a double negative zero. A typical curve is shown in Figure 3b, page 34.

Case 3. $64f_o^2 - 27\alpha_3^2 > 0$. Under this condition $F(\phi_o)$ will have a single positive zero and two distinct negative zeros. Figure 3c on page 34 shows a typical curve.

It has been established in the first part of the proof of the theorem that the unique positive zero in each of the above cases is an admissible value of ϕ_o and thus in each case there will be a unique positive, admissible root for the moment equation, $F(\alpha_3, \phi_o, f_o) = 0$. However, it will be necessary to investigate the negative zeros for $F(\phi_o)$ to see whether or not they lead to admissible roots for the moment equation, $F(\alpha_3, \phi_o, f_o) = 0$. Case 1 has already been disposed of.

3.3

Case 2. $64f_o^2 - 27\alpha_3^2 = 0$, or $\alpha_3 = \frac{8f_o}{3\sqrt{3}}$

Under this condition

$$F(\phi_o) = f_o\left[\frac{16\phi_o^3}{3\sqrt{3}} + 4\phi_o^2 - 1\right] \quad \text{and}$$

$$F'(\phi_o) = f_o\phi_o\left[\frac{48\phi_o}{3\sqrt{3}} + 8\right].$$

Any double root of $F(\phi_o) = 0$ must also be a root of $F'(\phi_o) = 0$. Thus the negative root of $F(\phi_o) = 0$ is $\phi_o = -\frac{\sqrt{3}}{2}$ which gives as the root $\phi_o^2 = \frac{3}{4}$. From Lemma 3.2.6, $\phi_o^2 < \frac{1}{4}$ for all $p > 0$ and, consequently, the negative root in this case is not an admissible one. Thus there is a unique root, or moment estimate, for p in this case.

Case 3. $64f_o^2 - 27\alpha_3^2 > 0$, or $f_o > \frac{3\sqrt{3}\alpha_3}{8}$.

The two roots of $F(\phi_o) = 0$ will have to lie on opposite sides of the abscissa of the maximum point, i.e. $\phi_o = -\frac{4f_o}{3}$. Let ϕ_o' and ϕ_o'' represent the zeros lying to the left and right, respectively, of the maximum point for the curve $F(\phi_o)$. Then

$$\phi_o' < \frac{-4f_o}{3\alpha_3} \quad \text{or} \quad \phi_o'^2 > \frac{16f_o^2}{9\alpha_3^2} > \frac{3}{4} > \frac{1}{4}$$ and thus by Lemma 3.2.6, ϕ_o' is not an admissible value of ϕ_o.

The admissibility of ϕ_o'', which lies between 0 and $-\frac{4f_o}{3\alpha_3}$ will be investigated now. Since

$$f_o > \frac{3\sqrt{3}\alpha_3}{8}, \quad -\frac{4f_o}{3\alpha_3} < -\frac{1}{2}$$

and the maximum point for the curve $F(\phi_o)$ is to the left of $\phi_o = -\frac{1}{2}$. Since $F(-\frac{1}{2}) = -\frac{\alpha_3}{4} < 0$, the zero for $F(\phi_o)$ must lie to the left of $\phi_o = -\frac{1}{2}$. That is, if $F(\phi_o'') = 0$, $\phi_o'' < -\frac{1}{2}$ and, consequently, $\phi_o''^2 > \frac{1}{4}$

3.3-4

and thus from Lemma 3.2.6, ϕ_o'' is not an admissible value of ϕ_o.

Since neither of the negative zeros of $F(\phi_o)$ leads to an admissible value of ϕ_o, $F(\alpha_3, \phi_o f_o) = 0$ has a unique admissible root and, consequently, there exists a unique moment estimate for p in this case.

Combining the conclusions of the three cases above establishes the theorem.

(Figures 3a, b, and c will be found on page 34.)

3.4 Formulas for the variance and covariance of the moment estimates of the parameters. The basic relationships between the moments and the parameters, as given in section 3.1 are:

$$a^{-1/2} f_o = \nu_1; \quad a^{-1} p = \nu_2; \quad a^{-3/2}(p + 1/2) f_o = \nu_3;$$

in which ν_s is the s^{th} moment for z about the unknown point -B.

In an applied problem m'_s, the s^{th} sample moment about -B, would be used and the basic functional relationships from which the moment estimates of the parameters are obtained are:

(3.4.1) $$a^{-1/2} f_o = m'_1$$

(3.4.2) $$a^{-1} p = m'_2$$

(3.4.3) $$a^{-3/2}(p + 1/2) f_o = m'_3$$

The large sample variance and covariance of sample moments about a fixed point are: (Ref. 11)

$$\sigma^2_{m'_s} = \frac{\nu_{2s} - \nu_s^2}{n}; \quad \mathrm{Cov}(m'_s, m'_t) = \frac{\nu_{s+t} - \nu_s \nu_t}{n}$$

in which n is the number of observations in the sample. These give:

(3.4.4) $$\sigma^2_{m'_1} = \frac{a^{-1} \phi_o^2}{n}$$

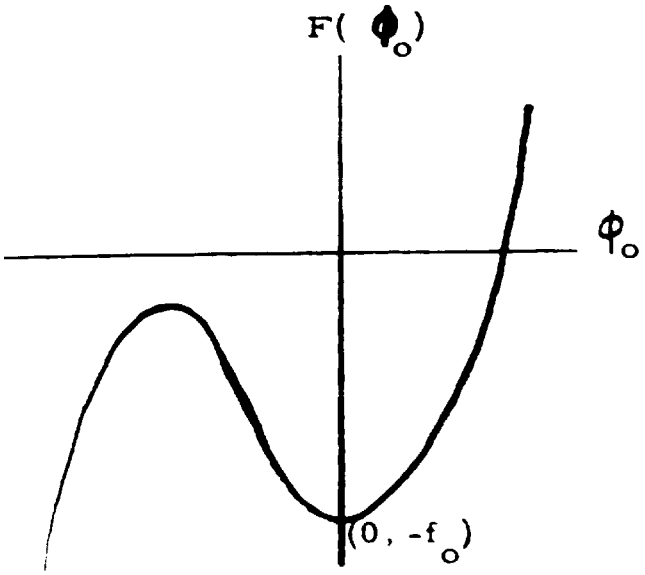


Figure 3a

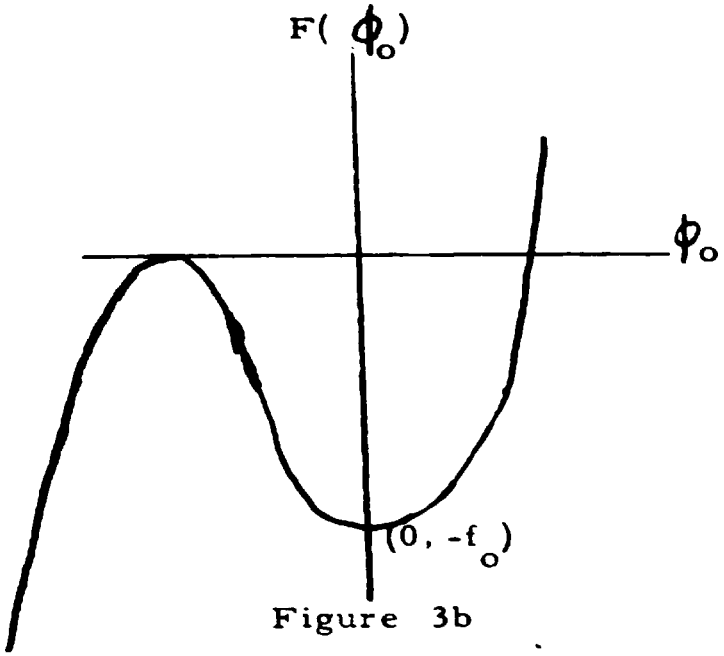


Figure 3b

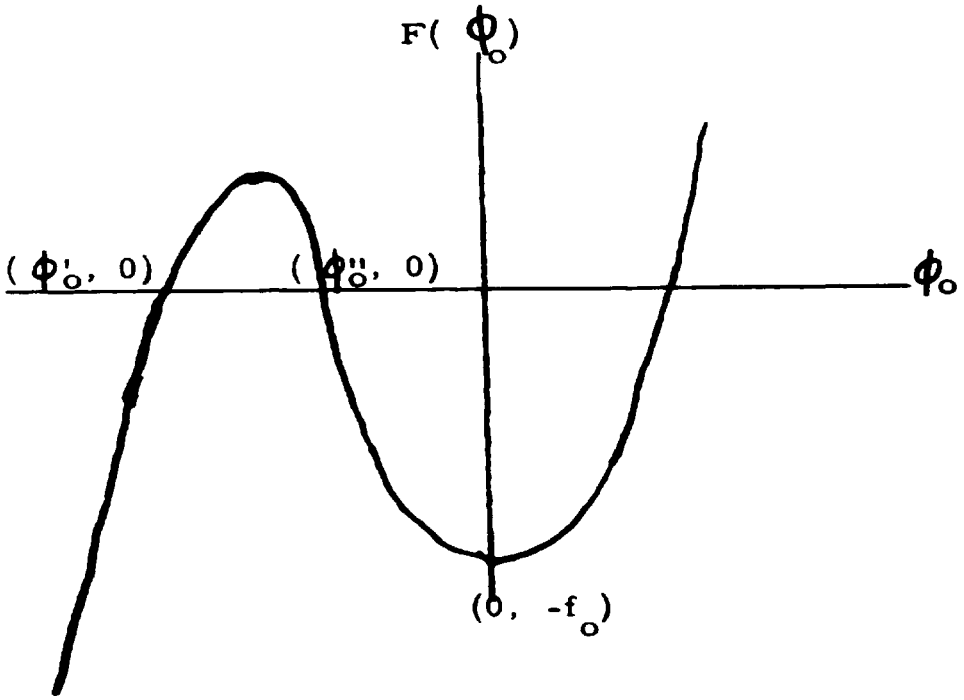


Figure 3c

Typical curves for $F(\alpha_3, \phi_o, f_o)$ in which α_3 and f_o are regarded as independent parameters subject only to the conditions $\alpha_3 > 0$ and $f_o > 0$.

3.4

$$(3.4.5) \qquad \sigma^2_{m'_2} = \frac{p}{a^2 n}$$

$$(3.4.6) \qquad \sigma^2_{m'_3} = \frac{\phi_o^2 (p + 1/2)^2 + 2p^2 + \frac{7p}{4}}{a^3 n}$$

$$(3.4.7) \qquad \mathrm{Cov}(m'_1 m'_2) = (f_o)(2a^{3/2} n)^{-1}$$

$$(3.4.8) \qquad \mathrm{Cov}(m'_1, m'_3) = (2p\,\phi_o^2 + 2\,\phi_o^2 + f_o)(2a^2 n)^{-1}$$

$$(3.4.9) \qquad \mathrm{Cov}(m'_2, m'_3) = 3f_o(p + 1/2)(2a^{5/2} n)^{-1}$$

From (3.4.1), (3.4.2), and (3.4.3) is obtained the relationship

$$\frac{1}{2p} = \frac{m'_3}{m'_1 m'_2} - 1 .$$

Taking the differential of both sides gives

$$-\frac{\delta p}{2p^2} = \frac{m'_1 m'_2 \delta m'_3 - m'_1 m'_3 \delta m'_2 - m'_2 m'_3 \delta m'_1}{(m'_1 m'_2)^2} .$$

Squaring both sides and taking expected values in accordance with the usual procedure and then substituting the relationships from (3.4.4) through (3.4.9) gives

$$(3.4.10) \qquad \sigma^2_p = \frac{p[4\,\phi_o^2 p(p + 1) + p^2 - f_o^2]}{n f_o^2} .$$

Similarly from (3.4.1), (3.4.2), and (3.4.3) are obtained

$$2a = \frac{m'_1}{m'_3 - m'_1 m'_2}$$

and

$$\delta a = \frac{m'_3 \delta m'_1 - m'_1 \delta m'_3 + m'^2_1 \delta m'_2}{2(m'_3 - m'_1 m'_2)^2} .$$

Squaring both sides and following the same procedure as above yields as a result,

3.4

$$(3.4.11)\qquad \sigma_a^2 = \frac{a^2}{nf_o^2}\,[4\,\phi_o^2(p+1) + p] = \frac{a^2}{p^2}\left(\sigma_p^2 + \frac{p}{n}\right)$$

From (3.4.2), $\frac{p}{a} = m'_2$

and

$$a\,\delta p - p\,\delta a = a^2\,\delta m'_2 \ ,$$

which yields by the same procedure as before,

$$(3.4.12)\qquad \mathrm{Cov}(a,p) = \frac{a}{nf_o^2}\,[4p\,\phi_o^2(p+1) + p^2 - f_o^2] = \frac{a}{p}\,\sigma_p^2 \ .$$

As previsouly noted k is the first moment about $-B$. That is, $k = V_1$. However, this does not justify writing $\sigma_k^2 = \sigma_{V_1}^2 = \frac{\phi_o^2}{an}$ since k is beingtreated as an independent parameter and its value is determined from the moment equations rather than computed directly from the sample data. In fact, k could not be computed until B is known. Thus the variance for k will be gotten from the moment equations. From (3.4.1)

$$k = \frac{f_o}{a^{1/2}}$$

and

$$\delta k = \frac{2a\,\delta f_o - f_o\,\delta a}{2a^{3/2}}$$

From the definitions of f_o, $\log f_o = \log \sqrt{(p + \frac{1}{2})} - \log \sqrt{p}$ and

$$\delta f_o = f_o[\psi(p + 1/2) - \psi(p)] \qquad \text{in which } \psi(p) = \frac{d \log \sqrt{p}}{dp} \ .$$

If $\Omega \equiv 2[\psi(p + 1/2) - \psi(p)]$, then

$$(3.4.13)\qquad \delta k = \frac{f_o(\Omega a\,\delta p - \delta a)}{2a^{3/2}} \ .$$

Employing the same procedure as in the previous cases give

$$(3.4.14)\qquad \sigma_k^2 = \frac{f_o^2}{4ap^2}\,[(p\Omega - 1)^2\,\sigma_p^2 + pn^{-1}] \ .$$

3.4

If B is known, the value of a and p would be completely determined by (3.4.1) and (3.4.2) and under this condition it can be shown that (3.4.14) becomes $\frac{\phi_o^2}{an}$ which agrees with $\sigma_{\gamma_1}^2$. However, if B were known, k could be computed directly from the sample data, i.e. the first moment about - B.

Multiplying both sides of (3.4.13) by δ_a gives

$$\delta_a \delta_k = \frac{k[a\Omega\,\delta a\,\delta p - (\delta a)^2]}{2a}, \quad \text{since } f_o = a^{1/2}k, \text{ or}$$

(3.4.15) $$\text{Cov}(a,k) = \frac{k(a^2\Omega\,\sigma_p^2 - p\,\sigma_a^2)}{2ap}.$$ Further

(3.4.16) $$\text{Cov}(p.k) = \frac{k(p\Omega - 1)\,\sigma_p^2}{2p}.$$

The above formulas give the variances and covariances of the three independent parameters. The dependent parameters are determined from a, p, and k by the following relationships: $A^2 = p$, $B = k - \mu_1'$, $b = 2aB$, $4ac = b^2$, $Aa' = a$, $Ab' = b$, and finally $A(A + c') = c$. By the use of these equations and employing the same procedure as was used in the derivation of the preceding formulas, the formulas for the variances of the dependent parameters are found to be as follows:

(3.4.17) $$\sigma_A^2 = \sigma_p^2(4p)^{-1}$$

(3.4.18) $$\sigma_B^2 = \sigma_k^2 + \frac{\phi_o^2}{an}$$

(3.4.19) $$\sigma_b^2 = 4[a^2 p\,\sigma_B^2 + a^2 k\Omega^2\,\sigma_p^2 + p\,\sigma_a^2(B^2 - k\Omega)]p^{-1}$$

(3.4.20) $$\sigma_c^2 = B^2[\sigma_b^2 - \frac{bka}{p}\Omega\sigma_p^2 + B(2k - 3B)\,\sigma_a^2]$$

(3.4.21) $$\sigma_{a'}^2 = \frac{a^2}{4p^3}(\sigma_p^2 + \frac{4p}{n})$$

(3.4.22) $$\sigma_{b'}^2 = [4p^2\sigma_b^2 - b\,\sigma_p^2(3b + pk\Omega + 4ak)](4p^3)^{-1}$$

(3.4.23) $$\sigma_{c'}^2 = [4p^2\sigma_c^2 + (p+c)\,\sigma_p^2\{p + c - 2bk(p\Omega - 1) - 2bB\}](4p^3)^{-1}$$

3.5

3.5. Comparison of the variances of the moment estimates of the independent parameters as obtained from the Z-function and as obtained from the Gamma type function. In many cases the parameters can be estimated from a given set of observations by use of either the Z-function or the Gamma type function. When the graduation is carried out, the χ^2 for the test of goodness of fit in both cases may be small. However, almost without exception in numerical problems worked by the author, the χ^2 resulting from graduation by the Z-function was the smaller. This is illustrated by the second example of section 6.1. In these cases the criterion for the determination of which function to use would be the relative magnitude of the sampling error of the estimates as determined by the two methods. Consequently, an investigation into the relative magnitude of the variances of these two sets of estimates is in order. The quantities to be compared will be the ratio of the variance of the parameter to the square of the estimate of the parameter.

The Gamma type function can be written

$$f(x)dx = \frac{1}{a_3 \lceil p_3} \left[\frac{x - B_3}{a_3}\right]^{p_3 - 1} e^{-\frac{x-B_3}{a_3}} dx, \qquad B_3 \leq x < \infty .$$

If $\nu_{s:x}$ represents the s^{th} moment for x about B_3, then

$$\nu_{s:x} = a_3^s \, p_3(p_3 + 1)(p_3 + 2) - - - - - (p_3 + s - 1) . \tag{3.5.1}$$

By method of moments p_3 is determined from the relation

$$p_3 = \frac{4}{\alpha_3^2}. \tag{3.5.2}$$

If p_1 denotes the parameter as determined by the method of moments from the Z-function, then from (3.1.8) the relationship for the determination of p_1 is

$$\alpha_3 = \frac{f_o(1 - 4\phi_o^2)}{2\phi_o^3} , \tag{3.5.3}$$

3.5

in which f_o and ϕ_o are functions of p_1. Then the relationship between p_1 and p_3 as determined from any given set of observations is

$$(3.5.4) \qquad p_3 = \frac{16\phi_o^6}{f_o^2(1 - 4\phi_o^2)^2} .$$

From Lemmas 3.2.4 and 3.2.6

$$(3.5.5) \qquad \frac{1}{4} > \phi_o^2 > \frac{p}{4p + 1} \quad \text{or} \quad -1 < -4\phi_o^2 < -\frac{4p}{4p + 1}$$

or

$$0 < 1 - 4\phi_o^2 < \frac{1}{4p + 1},$$

for any $p > 0$, and specifically for p_1. The above inequalities give

$$(3.5.6) \qquad f_o^2(1 - 4\phi_o^2)^2 < \frac{f_o^2}{(4p_1 + 1)^2} .$$

Putting (3.5.5) and (3.5.6) in (3.5.4) gives

$$(3.5.7) \qquad p_3 > \frac{16p_1^3}{f_o^2(4p_1 + 1)} .$$

From the definition of ϕ_o^2, it follows that $f_o^2 < \frac{4p_1^2}{4p_1 + 1}$ and consequently (3.5.7) becomes

$$(3.5.8) \qquad p_3 > 4p_1 .$$

That is, the estimate of p obtained by using the Gamma type function exceeds four times the estimate obtained by use of the Z-function.

Let

$$r_{p_3} \equiv \frac{\sigma_{p_3}^2}{p_3^2} \quad \text{and} \quad r_{p_1} \equiv \frac{\sigma_{p_1}^2}{p_1^2} .$$

Then from (3.4.10)

$$\sigma_{p_1}^2 = \frac{p_1}{nf_o^2}\left[4\phi_o^2 p_1(p_1 + 1) + p_1^2 - f_o^2\right].$$

3.5

The inequalities $\phi_o^2 < \frac{1}{4}$ and $f_0^2 > p - \frac{1}{4}$ give

$$\sigma_{p_1}^2 < \frac{p_1(8p_1^2 + 1)}{4nf_o^2}$$

and consequently

$$(3.5.9) \qquad r_{p_1} < \frac{8p_1^2 + 1}{4nf_o^2 p_1} .$$

From the variance of p_3 as given in (Ref. 17) and the definition of r_{p_3} is obtained

$$(3.5.10) \qquad r_{p_3} = \frac{6p_3(p_3 + 1)(p_3 + 5)}{p_3 n} .$$

Combining (3.5.9) and (3.5.10) gives, if the ratio $[r_{p_3}][r_{p_1}]^{-1}$ is represented by $R_{p_3:p_1}$

$$R_{p_3:p_1} > \frac{24(p_3^2 + 6p_3 + 5)f_o^2 p_1}{p_3(8p_1^2 + 1)} > \frac{24(p_3 + 6)p_1 f_o^2}{8p_1 + 1} ,$$

since $p_3 > 0$. From (3.5.8) $p_3 > 4p_1$ and consequently

$$R_{p_3:p_1} > \frac{24(4p_1 + 6)p_1 f_o^2}{8p_1^2 + 1} > 12 f_o^2 \left[1 + \frac{12p_1 - 1}{8p_1^2 + 1}\right] .$$

Thus for $p_1 > \frac{1}{12}$ it follows immediately

$$R_{p_3:p_1} > 12 f_o^2 .$$

The table of Appendix 1 shows $12 f_o^2 > 1$ for $p \geq .22$. Also f_o^2 is a monotonic increasing function of p. Thus for any $p > .22$, $r_{p_3} > r_{p_1}$, and their ratio is an increasing function of p_1 or p_3. Further f_o^2 is of the same order of magnitude as p_1 and thus $R_{p_3:p_1}$ is of the order of magnitude of $12 p_1$.

3.5

Let a_3 and a_1 designate the parameters as estimated from the Gamma type function and the Z-function respectively. Then a_3 and a_1 are not identically corresponding parameters, but they play similar roles in the two functions.

For the joint estimation of p_3, a_3, and B_3 the variance of a_3 is found (Ref. 11) as follows: From (3.5.1)

$$a_3 = \frac{\nu_{2:x} - \nu^2_{1:x}}{\nu_{1:x}} .$$

Taking the differential of both sides, squaring, and taking the expected value gives, after replacing $\nu_{s:x}$ by its value in terms of a_3 and p_3 as given by (3.5.1)

$$\sigma^2_{a_3} = \frac{a_3^2(2p_3 + 3)}{np_3} .$$

If $r_{a_3} \equiv \frac{\sigma^2_{a_3}}{a_3^2}$ then

(3.5.11) $$r_{a_3} = \frac{2p_3 + 3}{np_3} .$$

If $r_{a_1} \equiv \sigma^2_{a_1}/a_1^2$ from (3.4.11)

$$r_{a_1} = \frac{4\phi_o^2(p_1 + 1) + p_1}{nf_o^2}$$

and since $\phi_o^2 < \frac{1}{4}$,

(3.5.12) $$r_{a_1} < \frac{2p_1 + 1}{nf_o^2} .$$

If $R_{a_3:a_1} \equiv [r_{a_3}][r_{a_1}]^{-1}$ then (3.5.11) and (3.5.12) give

3.5

$$R_{a_3:a_1} > \left[2 + \frac{3}{p_3}\right]\left[\frac{f_o^2}{2p_1 + 1}\right] .$$

From Lemma 3.2.6, $f_o^2 > p - \frac{1}{4}$ and consequently,

$$R_{a_3:a_1} > \frac{1}{4}\left[2 + \frac{3}{p_3}\right]\left[\frac{4p_1 + 1}{2p_1 + 1}\right] .$$

This can be written

$$(3.5.13) \qquad R_{a_3:a_1} > 1 + \frac{3(4p_1 - 2p_3 - 1)}{4p_3(2p_1 + 1)} .$$

It cannot be stated on the basis of the above that r_{a_3} is greater than r_{a_1}. However, (3.5.13) does show that the ratio of these two quantities is greater than a quantity which does not differ appreciably from one for large values of either p_1 or p_3. Further, for small values of p_1, ϕ_o^2 is considerably less than .25 and as a consequence the inequality (3.5.13) is not a very close one, $R_{a_3:a_1}$ being considerably greater than the right member of the inequality.

Let k_1 designate the parameter k as determined from the Z-function. In order to get a parameter of the Gamma type function which corresponds more closely to k_1 than does B_3, consider $k_3 = \mu_1' - B_3$. Then k_3 is the first moment about the lower limit of the range, or it is the lower limit of the range if the variates are measured about their mean. Since -B is the lower limit of the range for the variate which obeys the z distribution function, and $k_1 = \mu_1' + B$, k_1 will play the same role for the z variate as k_3 does for the x variate, that is, each is the lower limit of the range if the variates are measured from their mean.

To find the variance of k_3, from (3.5.1), $k_3 = a_3 p_3$. The usual procedure yields

$$\sigma_{k_3}^2 = E\left[a_3^2(\delta p_3)^2 + 2a_3 p_3 \delta p_3 \delta a_3 + p_3^2(\delta a_3)^2\right] .$$

3.5

From the relationship $a_3^2 p_3 = \mu_2$ is obtained

$$\delta a_3 \delta p_3 = \frac{[\delta \mu_2]^2 - 4a_3^2 p_3^2 [\delta a_3]^2 - a_3^4 [\delta p_3]^2}{4a_3^3 p_3} .$$

Using this in the above gives

$$\sigma_{k_3}^2 = \frac{\sigma_{\mu_2}^2 + a_3^4 \sigma_{p_3}^2 - 2a_3^2 p_3^2 \sigma_{a_3}^2}{2a_3^2} .$$

From reference 11,

$$\sigma_{\mu_2}^2 = \frac{\mu_4 - \mu_2^2}{n} .$$

If μ_4 and μ_2 are replaced in this by their values in terms of a_3 and p_3 as obtained from (3.5.1) the result is

$$\sigma_{\mu_2}^2 = \frac{2a_3^4 p_3(p_3 + 3)}{n} .$$

Using the expressions for $\sigma_{p_3}^2$ and $\sigma_{a_3}^2$ which have been obtained earlier in this section gives

$$\sigma_{k_3}^2 = \frac{a_3^2 p_3(3p_3^2 + 17p_3 + 15)}{n} .$$

From (3.5.1) $k_3 = a_3 p_3$, and if $r_{k_3} = \sigma_{k_3}^2 / k_3^2$, then it follows

$$(3.5.14) \qquad r_{k_3} = \frac{3p_3^2 + 17p_3 + 15}{np_3} .$$

Replacing $\sigma_{p_1}^2$ in (3.4.14) by its value as given by (3.4.10) gives,

$$\sigma_{k_1}^2 = \frac{1}{4a_1 p_1} \left\{ \frac{(p_1 \Omega - 1)^2 [4\phi_o^2 p_1(p_1 + 1) + p_1^2 - f_o^2] + f_o^2}{n} \right\} .$$

From (3.1.1) $k_1 = a_1^{-\frac{1}{2}} f_o$ and thus if $r_{k_1} = \sigma_{k_1}^2 / k_1^2$

3.5

$$(3.5.15) \qquad r_{k_1} = \frac{(p_1\Omega - 1)^2[4\phi_o^2 p_1(p_1+1) + p_1^2 - f_o^2] + f_o^2}{4np_1 f_o^2} .$$

From definition $\Omega = 2[\psi(p + \frac{1}{2}) - \psi(p)]$ and thus from (3.2.2)

$$(3.5.16) \qquad 2f_o \frac{df_o}{dp_1} = f_o^2 \Omega$$

and from Lemma 3.2.3, $2f_o \frac{df_o}{dp} < 1$. Consequently $f_o^2\Omega < 1$.

From (3.2.3) $\frac{df_o}{dp} > 0$, and since $f_o > 0$, (3.5.16) gives $\Omega > 0$. Since $p_1 > 0$, combining the above gives

$$0 < p_1\Omega < \frac{p_1}{f_o^2}$$

or

$$-1 < p_1\Omega - 1 < \frac{\phi_o^2}{f_o^2} < \frac{1}{4f_o^2},$$

since $\phi_o^2 < \frac{1}{4}$. From the table 1 of Appendix I, $4f_o^2 > 1$ for $p_1 = .44$. Also from (3.2.3) $\frac{df_o}{dp} > 0$, and consequently f_o^2 is a monotonic increasing function of p. Thus for any $p_1 > .44$, $1/4f_o^2 < 1$. Consequently for any $p_1 \geq .44$, we have $(p_1\Omega - 1)^2 < 1$. Putting this in (3.5.15) gives

$$r_{k_1} < \frac{4\phi_o^2(p_1+1) + p_1}{4nf_o^2} .$$

Again using the inequality $\phi_o^2 < \frac{1}{4}$ gives

$$(3.5.17) \qquad r_{k_1} < \frac{2p_1 + 1}{4nf_o^2} .$$

If $R_{k_3:k_1} = [r_{k_3}][r_{k_1}]^{-1}$ (3.5.14) and (3.5.17) gives

$$R_{k_3:k_1} > 4f_o^2\left[3p_3 + 17 + \frac{15}{p_3}\right]\left[\frac{1}{2p_1+1}\right] > \frac{4(3p_3+17)f_o^2}{2p_1+1} .$$

3.5

From (3.5.8) $p_3 > 4p_1$. Putting this in the above gives

$$R_{k_3:k_1} > \frac{4(12p_1 + 17)f_o^2}{2p_1 + 1} > 24f_o^2 .$$

From table 1 of Appendix I, for $p \geq .44$, $24f_o^2 > 1$. Actually this inequality is true for a much smaller value of p but this smaller value cannot be justified here as the condition $p \geq .44$ was imposed in the derivation of the inequality (3.5.17). Further f_o^2 is an increasing function of p and of the same order. Thus $R_{k_3:k_1}$ increases with p and is of the same order of magnitude as $24p_1$.

PART IV

Estimation by Method of Maximum Likelihood

4.1 Equations for the determination of the maximum likelihood estimates of the parameters and outline of the procedure for their solution.

From (2.3.2) the probability density function for z is

$$f(z)dz = \frac{2a^p}{\sqrt{p}} (z + B)^{2p-1} e^{-a(z+B)^2} dz, \quad -B \leq z < \infty .$$

From definition, $B = k - \mu_1^1$. If w represents the z variate when measured about the mean, i.e. $w = z - \mu_1^1$, the probability density function for w becomes,

$$(4.1.1) \qquad f(w)dw = \frac{2a^p}{\sqrt{p}} (w + k)^{2p-1} e^{-a(w+k)^2} dw, \quad -k \leq w < \infty.$$

The independent parameters are now the same as in the case by method of moments, a, k, and p. It can be shown that the maximum likelihood equations derived from f(z)dz and regarding a, p, and B as the independent parameters lead to the same results as those derived from (4.1.1) and regarding a, p, and k as the independent parameters. Thus there is no loss in generality in the change in origin along with the corresponding change in independent parameters.

If $L = \Sigma \log f(w)dw$ in which the summation is taken over the sample values of w and n is the number of observations in the sample, then

$$(4.1.2) \qquad L = n \log 2 + np \log a - n \log \sqrt{p} + (2p - 1) \sum_{i=1}^{n} \log(w_i + k) - a \sum_{i=1}^{n} (w_i + k)^2 .$$

If m_2 is the second central sample moment and $\psi(p) \equiv \frac{d \log \sqrt{p}}{dp}$

4.1

the equations* for the determination of the estimates of the parameters by the method of maximum likelihood are:

$$(4.1.3) \qquad p - a(m_2 + k^2) = 0$$

$$(4.1.4) \qquad (2p - 1) R - 2nak = 0 \qquad R \equiv \sum_{i=1}^{n} \frac{1}{w_i + k}$$

$$(4.1.5) \qquad n \log a - n\psi(p) + 2 \sum_{i=1}^{n} \log (w_i + k) = 0 .$$

M. L. equations in general, and the ones occurring here in particular, do not lend themselves to an exact algebraic solution. Two methods of approximating the roots for M. L. equations are given by Fisher and Koshal (Ref. 13, 19 and 20). These methods are based on the expansion of either the partial derivatives of L, or of L itself, about a point, usually the moment estimates of the parameters. Neither of these methods is applicable to the above system of equations since $\partial^i L/\partial a^i$ is of order a^{-i} and thus if $a < 1$, which it usually is, $\partial^i L/\partial a^i$ will increase in magnitude with i. Thus the remainder term will be of increasing magnitude unless the first approximation is extremely close to the true roots of the system. Since there is no way of measuring the accuracy of this first approximation, it will be necessary to determine the roots of the above M. L. equations by some other method.

The elimination of first a and then p between (4. 1. 3) and (4. 1. 4) gives,

$$(4.1.6) \qquad p = \frac{R(m_2 + k^2)}{2\,[R(m_2 + k^2) - nk]}$$

and

* This system of equations will be referred to as the M. L. (Maximum Likelihood) equations and any set of parameters which satisfy the M. L. equations will be referred to as the M. L. estimates.

4.1

$$(4.1.7) \qquad a = \frac{R}{2[R(m_2 + k^2) - nk]} .$$

Eliminating a between (4.1.3) and (4.1.5) gives

$$(4.1.8) \qquad n \log p - n\psi(p) + 2 \sum_{i=1}^{n} \log (w_i + k) - n \log (m_2 + k^2) = 0 .$$

For simplicity of notation the following functions of k will be defined:

$$K = (m_2 + k^2) \qquad p_k = \frac{RK}{2(RK - nk)}$$

$$P(k) = \prod_{i=1}^{n} (w_i + k)^{\frac{2}{n}} \qquad U(k) = \log p_k - \psi(p_k) + \log \frac{P(k)}{K} .$$

The procedure for solving the M. L. equations now becomes: For any value of k, R can be calculated from the sample observations and the corresponding value of p_k computed from (4.1.6). Likewise the value of P(k) corresponding to any k can be calculated from the sample observations. Thus for any k, U(k) can be evaluated. Two values of k, say k_1 and k_2, can be found such that $U(k_1) < 0$ and $U(k_2) > 0$. By successive approximations the interval k_1, k_2 can be shortened as much as desired. Since U(k) is a continuous function of k, there will be some k, say k_o, between k_1 and k_2 such that $U(k_o) = 0$.

The moment estimate of k can be used as a starting value. If a calculating machine is available the above process is neither difficult nor complicated and U(k) can be evaluated in a short time for any k. Having found the zero for U(k), the M. L. estimates for p and a are obtained from (4.1.6) and (4.1.7) respectively. The M. L. estimates for the dependent parameters can then be found from the appropriate relationships as given in the preceding sections.

The analysis and proof of the existence of the roots of the M. L. equations will be discussed in a later section. Before this can be done it will be necessary to establish some of the properties of the functions

4.1-2

R(k) and P(k).

4.2 Study of the functions $R(k) \equiv \sum_{i=1}^{n} \frac{1}{w_i + k}$ and $P(k) \equiv \prod_{i=1}^{n} (w_i + k)^{2/n}$.

Before investigating the roots for the M. L. equations it will be necessary to establish some of the properties of the functions which are defined above and occur in these equations. The study of these properties will be organized into a series of lemmas. Where there is no danger of ambiguity the argument, k, will be omitted.

It is assumed that the set of observations, $\{w_i\}$, $i = 1, 2, \cdots n$ satisfy the following conditions:

a) w_i is finite for every i.

b) n is finite but greater than 2.

c) $\sum_{i=1}^{n} w_i = 0$, i.e. the sample observations are measured from the mean of the sample.

d) $w_i \neq w_j$ for some i and some j, i.e. $m_2 > 0$.

e) $w_i \leq w_j$ for $i < j$, for every i and j.

The following lemmas are valid for any set of observations which satisfy the above conditions.

Lemma 4.2.1. $w_i + k \geq 0$ for all i.

Proof: The statement is obvious from the range of values for w, i.e. $-k \leq w < \infty$.

Corollary: $k > 0$. This is an immediate consequence of the Lemma and properties c) and d) above.

Lemma 4.2.2 If m_r is the r^{th} central sample moment for z, then

$$R = \sum_{r=0}^{\infty} (-1)^r \frac{n m_r}{k^{r+1}},$$

for any $k \geq \max |w_i|$.

4.2

Proof: By actual division

$$\frac{1}{k + w_i} = \sum_{r=o}^{\infty} (-1)^r \frac{w_i^r}{k^{r+1}} .$$

If $k > \max |w_i|$ this series will converge absolutely for every i. Then from definition of R,

$$R = \sum_{i=1}^{n} \sum_{r=o}^{\infty} (-1)^r \frac{w_i^r}{k^{r+1}} .$$

Since the summation over r converges absolutely for every i, the order of summation can be reversed. Thus

$$R = \sum_{r=o}^{\infty} (-1)^r \frac{1}{k^{r+1}} \sum_{i=1}^{n} w_i^r = \sum_{r=o}^{\infty} (-1)^r \frac{nm_r}{k^{r+1}} .$$

Lemma 4.2.3. $R = \frac{d[\log g(k)]}{dk}$, in which $g(k) \equiv \prod_{i=1}^{n} (w_i + k)$.

Proof: By definition $R = \sum_{i=1}^{n} \frac{1}{w_i + k}$. If $g'(k) = \frac{d[g(k)]}{dk}$,

then $R = \frac{g'(k)}{g(k)}$, (Ref. 2), and it follows as a consequence that $R = \frac{d[\log g(k)]}{dk}$.

Lemma 4.2.4 For values of $k \geq \max |w_i|$,

$$\frac{nk}{R} = k^2 - m_2 + \frac{m_3}{k} + \frac{m_2^2 - m_4}{k^2} + \text{terms of order } k^{-3} .$$

Proof: From Lemma 4.2.2 for $k \geq \max |w_i|$,

(4.2.1) $$\frac{n}{R} = \frac{k}{\sum_{r=o}^{\infty} (-1)^r \frac{m_r}{k^r}} .$$

In order to expand the right member in a power series in k, write

4.2

$$(4.2.2)\qquad \frac{k}{\sum_{r=0}^{\infty}(-1)^r \frac{m_r}{k^r}} \equiv \lambda_{-1}k + \lambda_o + \lambda_1 k^{-1} + \lambda_2 k^{-2} + \lambda_3 k^{-3} \cdots\cdots$$

or

$$k \equiv (\lambda_{-1}k + \lambda_o + \lambda_1 k^{-1} + \lambda_2 k^{-2} + \lambda_3 k^{-3} \cdots)(m_o - m_1 k^{-1} + m_2 k^{-2} \cdots) .$$

Since (4.2.2) is an identity in k, the λ's can be found by equating coefficients of like powers of k. This gives: $\lambda_{-1} = 1$; $\lambda_o = 0$; $\lambda_1 = -m_2$; $\lambda_2 = m_3$; $\lambda_3 = m_2^2 - m_4$; $\lambda_4 = m_5 - 2m_2m_3$; and $\lambda_5 = -m_6 + 2m_2m_4 + m_3^2 - m_2^3$. Putting these values in (4.2.2) and multiplying both sides of (4.2.1) by k gives

$$\frac{nk}{R} = k^2 - m_2 + m_3 k^{-1} + (m_2^2 - m_4)k^{-2} + \text{terms } 0(k^{-3}) .$$

Corollary: $\lim_{k\to\infty} \frac{RK}{nk} = 1.$

Proof: From the previous Lemma $\frac{nk}{R} = k^2 - m_2 + D_1(k)$, in which $D_1(k)$ is a function of degree -1 in k and is finite for all $k \geq \max |w_i|$. Then by division

$$\frac{K}{k^2 - m_2 + D_1(k)} = 1 + \frac{2m_2 - D_1(k)}{k^2 - m_2 + D_1(k)} ,$$

$D_1(k)$ is finite for all $k > \max |w_i|$ and m_2 is finite and thus the limit of the second term of the right member above is zero as k becomes infinite. Consequently,

$$\lim_{k\to\infty} \frac{RK}{nk} = \lim_{k\to\infty} \frac{k^2 + m_2}{k^2 - m_2 + D_1(k)} = 1.$$

Lemma 4.2.5 $\lim_{k\to -w_1} \frac{RK}{nk} = \infty .$

Proof: From the definition of R, it is obvious $\lim_{k\to -w_1} R = \infty .$

4.2

$K = k^2 + m_2 > 0$ for all k. Thus it follows $\lim_{k \to -w_1} \frac{RK}{nk} = \infty$.

Lemma 4.2.6 $R > \frac{n}{k}$ for any k.

Proof: (By mathematical induction.) Let $u_i \equiv w_i + k$, then from Lemma 4.2.1, $u_i \geq 0$ for every i. Also,

$$\frac{u_i}{u_j} + \frac{u_j}{u_i} = \frac{2u_iu_j + u_i^2 - 2u_iu_j + u_j^2}{u_iu_j} = 2 + \frac{(u_i - u_j)^2}{u_iu_j} \geq 2$$

for all i and all j. Then for n = 2,

$$(u_1 + u_2)(\frac{1}{u_1} + \frac{1}{u_2}) = 2 + (\frac{u_1}{u_2} + \frac{u_2}{u_1}) \geq 4 = n^2 .$$

For n = 3,

$$(u_1 + u_2 + u_3)(\frac{1}{u_1} + \frac{1}{u_2} + \frac{1}{u_3}) = 3 + \frac{u_1}{u_2} + \frac{u_2}{u_1} + \frac{u_1}{u_3} + \frac{u_3}{u_1} + \frac{u_2}{u_3} + \frac{u_3}{u_2} \geq 9 = n^2 .$$

Thus for n = 2 and n = 3,

$$\sum_{i=1}^{n} u_i \sum_{i=1}^{n} \frac{1}{u_i} \geq n^2 .$$

The equality can hold only if $u_i = u_j$ for every i and every j. However, this is a contradiction of condition d) and thus the inequality cannot hold.

Now assume

$$(4.2.3) \qquad \sum_{i=1}^{n} u_i \sum_{i=1}^{n} \frac{1}{u_i} > n^2 .$$

Then

4.2

$$\sum_{i=1}^{n+1} u_i \sum_{i=1}^{n+1} \frac{1}{u_i} = \sum_{i=1}^{n} u_i \sum_{i=1}^{n} \frac{1}{u_i} + u_{n+1} \sum_{i=1}^{n} \frac{1}{u_i} + \frac{1}{u_{n+1}} \sum_{i=1}^{n} u_i + 1$$

$$= \sum_{i=1}^{n} u_i \sum_{i=1}^{n} \frac{1}{u_i} + \sum_{i=1}^{n} \left[\frac{u_{n+1}}{u_i} + \frac{u_i}{u_{n+1}} \right] + 1.$$

$\sum_{i=1}^{n} \left[\frac{u_{n+1}}{u_i} + \frac{u_i}{u_{n+1}} \right]$ contains n terms of the type $\left[\frac{u_i}{u_j} + \frac{u_j}{u_i} \right]$ each of which is greater than two. Thus from (4.2.3)

$$\sum_{i=1}^{n+1} u_i \sum_{i=1}^{n+1} \frac{1}{u_i} > n^2 + 2n + 1 = (n+1)^2$$

which is the same form as (4.2.3) with n + 1 replacing n. Replacing u_i by $w_i + k$ gives

$$\sum_{i=1}^{n} (w_i + k) \sum_{i=1}^{n} \frac{1}{w_i + k} > n^2 .$$

From condition c) this becomes $nkR > n^2$ or $R > \frac{n}{k}$ since from Corollary of Lemma 4.2.1, $k > 0$.

Corollary. $R > \frac{nk}{K}$ for any k.

Proof: Since $m_2 > 0$, $K = k^2 + m_2 > k^2$ and thus $\frac{n}{k} > \frac{nk}{K}$ and consequently

$$R > \frac{n}{k} > \frac{nk}{K} \quad \text{for any } k.$$

Lemma 4.2.7. $\sum_{i=1}^{n} \frac{w_i}{(w_i + k)^2} < 0$.

Proof: From conditions c) and d) there will be at least one i for which $w_i < 0$ and at least one other i for which $w_i > 0$. Also from condition e) $w_j \geq w_i$ for $j > i$. Further let the set of observations be ordered as follows:

4.2

$w_i \leq 0$ for $i \leq n_1$; $w_i > 0$ for $i > n_1$ for some $n_1 < n$.

Also let $w_{n_1} \geq w_i$ for $i \leq n_1$ and $w_{n_1+1} \leq w_i$ for $i > n_1$ and finally $w_n \geq w_i$ for all $i \leq n$.

From condition c)

$$w_n = -\sum_{i=1}^{n-1} w_i$$

and

$$\sum_{i=1}^{n} \frac{w_i}{(w_i + k)^2} = \sum_{i=1}^{n-1} \frac{w_i}{(w_i + k)^2} - \sum_{i=1}^{n-1} \frac{w_i}{(w_n + k)^2}$$

or

$$(4.2.4) \quad \sum_{i=1}^{n} \frac{w_i}{(w_i + k)^2} = \sum_{i=1}^{n} \frac{w_i [(w_n + k)^2 - (w_i + k)^2]}{(w_i + k)^2 (w_n + k)^2} = \sum_{i=1}^{n} w_i \overline{W}_i$$

in which

$$\overline{W}_i \equiv \frac{(w_n + k)^2 - (w_i + k)^2}{(w_n + k)^2 (w_i + k)^2} .$$

Then since $w_n \geq w_i$ for any $i < n$, $\overline{W}_i \geq 0$ for every i. Also since $w_{n_1} \geq w_i$ for $i \leq n_1$, $(w_{n_1} + k)^2 \geq (w_i + k)^2$ for $i \leq n_1$. Consequently,

$$(4.2.5) \quad \overline{W}_{n_1} \leq \overline{W}_i \text{ for } i \leq n_1 .$$

In like fashion,

$$(4.2.6) \quad \overline{W}_{n_1+1} \geq \overline{W}_i \text{ for any } i > n_1 .$$

Since $w_{n_1} \leq 0$ and $w_{n_1+1} > 0$, and from Lemma 4.2.1 Corollary, $k > 0$, $(w_{n_1} + k)^2 < (w_{n_1+1} + k)^2$ and consequently,

$$(4.2.7) \quad \overline{W}_{n_1} > \overline{W}_{n_1+1} .$$

Then from (4.2.4)

4.2

$$(4.2.8)\qquad \sum_{i=1}^{n} \frac{w_i}{(w_i + k)^2} = \sum_{i=1}^{n_1} w_i \overline{W}_i + \sum_{i=n_1+1}^{n} w_i \overline{W}_i .$$

From condition c),

$$\sum_{i=1}^{n_1} w_i = - \sum_{i=n_1+1}^{n} w_i = -S ,$$

in which S is some positive quantity. Then from (4.2.5)

$$\left| \sum_{i=1}^{n_1} w_i \overline{W}_i \right| \geq \overline{W}_{n_1} \left| \sum_{i=1}^{n_1} w_i \right| = S\overline{W}_{n_1}$$

and from (4.2.6)

$$\left| \sum_{i=n_1+1}^{n} w_i \overline{W}_i \right| \leq W_{n_1+1} \left| \sum_{i=n_1+1}^{n} w_i \right| = S\overline{W}_{n_1+1} .$$

Combining these two inequalities with (4.2.7) gives

$$\left| \sum_{i=1}^{n_1} w_i \overline{W}_i \right| > \left| \sum_{i=n_1+1}^{n} w_i \overline{W}_i \right| .$$

Since $w_i \leq 0$ for all $i \leq n_1$ and $w_i > 0$ for $i > n_1$; and $\overline{W}_i \geq 0$ for all i, it follows

$$\sum_{i=1}^{n_1} w_i \overline{W} < 0 \quad \text{and} \quad \sum_{i=n_1+1}^{n} w_i \overline{W}_i > 0.$$

Thus the first term of the right member of (4.2.8) is negative and the second term is positive. However, the negative term has the larger absolute value, and it follows as a consequence from (4.2.8)

$$\sum_{i=1}^{n} \frac{w_i}{(w_i + k)^2} < 0.$$

Corollary. $$\sum_{i=1}^{n} \frac{1}{w_i + k} < k \sum_{i=1}^{n} \frac{1}{(w_i + k)^2} .$$

4.2

Proof: $$\sum_{i=1}^{n} \frac{1}{w_i + k} \equiv \sum_{i=1}^{n} \frac{w_i}{(w_i + k)^2} + k \sum_{i=1}^{n} \frac{1}{(w_i + k)^2} .$$

From Lemma 4.2.1 $w_i + k \geq 0$ for all k and thus

$$\sum_{i=1}^{n} \frac{1}{w_i + k} > 0$$

for all k, and the Corollary follows immediately from the above identity.

Lemma 4.2.8. $\frac{nk}{RK}$ is a monotonic increasing function of k

Proof: Writing $M = \frac{nk}{RK}$, and $R' = \frac{dR}{dk} = -\sum_{i=1}^{n} \frac{1}{(w_i + k)^2}$ gives

$$\frac{dM}{dk} = n \left\{ \frac{RK - k(R'K + 2kR)}{R^2K^2} \right\} .$$

Writing

$$R_2 = \sum_{i=1}^{n} \frac{1}{(w_i + k)^2} ,$$

gives $R' = -R_2$ and

$$\frac{dM}{dk} = n \left\{ \frac{R(m_2 - k^2) + kR_2(m_2 + k^2)}{R^2 K^2} \right\} .$$

From the Corollary of Lemma 4.2.7 $kR_2 > R$ and thus

$$\frac{dM}{dk} > \frac{2nm_2}{RK^2} > 0 ,$$

since each R and m_2 is obviously positive.

Thus $\frac{dM}{dk} > 0$ for all k, and M is a monotonic increasing function of k.

Lemma 4.2.9. $P(k) = \prod_{i=1}^{n} (w_i + k)^{2/n} \leq k^2$ for all k.

4.2

Proof: It is a well known inequality (Ref. 14) that for any $u_i \geq 0$,

$$\frac{1}{n} \sum_{i=1}^{n} u_i \geq \left[\prod_{i=1}^{n} u_i \right]^{1/n} .$$

From Lemma 4.2.9, $w_i + k \geq 0$ for every i, and thus

$$\left[\frac{1}{n} \sum_{i=1}^{n} (w_i + k) \right]^2 \geq \prod_{i=1}^{n} (w_i + k)^{2/n} . \tag{4.2.9}$$

From condition c) $\sum_{i=1}^{n} (w_i + k) = nk$ and thus (4.2.9) becomes

$$k^2 > \prod_{i=1}^{n} (w_i + k)^{2/n} .$$

Corollary.

$$\frac{\prod_{i=1}^{n} (w_i + k)^{2/n}}{K} < 1 \quad \text{for every } k.$$

Proof: Since $m_2 > 0$, it follows that $K > k^2$ and thus from the previous Lemma,

$$\frac{\left[\prod_{i=1}^{n} (w_i + k) \right]^{2/n}}{K} \leq \frac{k^2}{K} < 1.$$

Lemma 4.2.10. For sufficiently large values of k,

$$\prod_{i=1}^{n} (w_i + k)^{2/n} = k^2 - m_2 + \frac{2}{3} m_3 k^{-1} + \frac{1}{2} (m_2^2 - m_4) k^{-2} + \text{terms } 0(k^{-3}).$$

Proof: Define q_r as the sum of the products of the w's taken r at a time. Then, since $q_1 = \sum_{i=1}^{n} w_i = 0$

$$\prod_{i=1}^{n} (w_i + k) = k^n + q_2 k^{n-2} + q_3 k^{n-3} + \cdots q_n \quad \text{and}$$

$$\left[\prod_{i=1}^{n} (w_i + k) \right]^{2/n} = k^2 (1 + q_2 k^{-2} + q_3 k^{-3} + \cdots q_n k^{-n})^{2/n} . \tag{4.2.10}$$

4.2

Since q_r is finite for every $r \leq n$ and the expression in the parentheses above contains a finite number of terms, it is obvious that there will exist some k, say k', such that for $k \geq k'$, $q_2 k^{-2} + q_3 k^{-3} + q_4 k^{-4} + \cdots + q_n k^{-n} < 1$. Thus the right member of (4.2.10) can be expanded by the binomial theorem with the expansion being valid and giving an absolutely convergent series for any $k \geq k'$ (Ref. 24). This expansion gives,

$$(4.2.11)\quad \left[\prod_{i=1}^{n} (w_i + k)\right]^{2/n} = k^2 + \frac{2}{n} q_2 + \frac{2}{n} q_3 k^{-1} + \left[\frac{2}{n} q_4 + \frac{2-n}{n^2} q_2^2\right] k^{-2} + \left[\frac{2}{n} q_5 + \frac{2(2-n)}{n^2} q_2 q_3\right] k^{-3} + \left[\frac{2}{n} q_6 + \frac{2-n}{n^2}(q_3^2 + 2q_2 q_4) + \frac{(2-n)(2-2n)}{3n^3} q_2^3\right] k^{-4} + \text{terms } 0(k^{-5}) .$$

If $S_r = \sum_{i=1}^{n} w_i^r$, the following relationships are obtained from Newton's formula for symmetric functions (Ref. 2).

$$q_1 = S_1 = 0;\ q_2 = -\frac{1}{2} S_2;\ q_3 = \frac{1}{3} S_3;\ q_4 = \frac{1}{8}(S_2^2 - 2S_4) .$$

By definition $S_r = nm_r$. Putting these values in (4.2.11) gives

$$\prod_{i=1}^{n} (w_i + k)^{2/n} = k^2 - m_2 + \frac{2}{3} m_3 k^{-1} + \frac{1}{2}(m_2^2 - m_4) k^{-2} + \text{terms } 0(k^{-3}) .$$

Corollary.

$$\lim_{k \to \infty} \frac{\prod_{i=1}^{n} (w_i + k)^{2/n}}{K} = 1.$$

Proof: From the previous Lemma

$$\prod_{i=1}^{n} (w_i + k)^{2/n} = k^2 - m_2 + D_2(k)$$

in which $D_2(k)$ is finite for large k and of degree -1 in k. Then

$$\frac{\prod_{i=1}^{n} (w_i + k)^{2/n}}{K} = 1 - \frac{2m_2 - D_2(k)}{k^2 + m_2}$$

4.2

Since $D_2(k)$ is finite for large k, and m_2 is finite, the second term of the right member approaches zero as k becomes infinite. Thus

$$\lim_{k\to\infty} \frac{\prod_{i=1}^{n} (w_i + k)^{2/n}}{K} = 1.$$

Lemma 4.2.11.

$$\frac{\prod_{i=1}^{n} (w_i + k)^{2/n}}{K} \text{ is a monotonic increasing function of k.}$$

Proof: In the notation introduced earlier, the function in this Lemma becomes $\frac{P(k)}{K}$. If this function is designated by H, and H' represents its derivative with respect to k, then $\log H = \log P(k) - \log K$, and from Lemma 4.2.3,

$$\frac{H'}{H} = \frac{2}{n} R - \frac{2k}{K} \quad \text{or} \quad H' = \frac{2Hk}{K}\left[\frac{RK}{nk} - 1\right].$$

From Lemma 4.2.1 $w_i + k \geq 0$ for all admissible k. Consequently $P(k) \geq 0$ and since $K > 0$, it follows that $H > 0$. From Lemma 4.2.6 $\frac{RK}{nk} - 1 > 0$. It follows immediately from these inequalities that $H' \geq 0$ for all admissible k, and $H' > 0$ for all $k > -w_1$. Thus H is a monotonic increasing function of k for all admissible values of k.

Lemma 4.2.12. If

$$R_2 = \sum_{i=1}^{n} \frac{1}{(w_i + k)^2}, \quad \text{and} \quad g = \prod_{i=1}^{n} (w_i + k),$$

then

$$R_2 = \frac{nk^{2n-2} + 2(n-3)q_2k^{2n-4} + 2(n-6)q_3k^{2n-5} + V(k)}{g^2}$$

in which V(k) is a polynomial of degree 2n - 6 in k, consisting of a finite number of terms each with a finite coefficient* .

* Such a polynomial will be referred to as a finite polynomial of a specified degree.

4.2

Proof: In the proof the following notation will be used:

$g_r = \frac{g}{w_r + k}$ for any r such that $1 \leq r \leq n$ and $q_{r:j} = q_r$ less those terms of q_r which contain w_j, i.e. for $n = 4$ $q_{3:1} = w_2 w_3 w_4$. If R_2 is cleared of fractions, in this notation it becomes

$$(4.2.12) \qquad R_2 = \frac{\sum_{r=1}^{n} g_r^2}{g^2} .$$

From the definition of g_r

$$(4.2.13) \qquad g_1 = \sum_{j=o}^{n-1} q_{j:1}\, k^{n-1-j} , \text{ with } q_{o:1} \text{ defined as unity.}$$

From the relationship $q_{1:i} = -w_i$ it is a straightforward operation to obtain the general relationship for any $r \leq n$ in which $q_o = 1$ and $q_1 = 0$,

$$(4.2.14) \qquad q_{r:i} = \sum_{j=o}^{r} (-1)^j q_{r-j}\, w_i^j .$$

Replacing the subscript 1 by i in (4.2.13)

$$g_i^2 = \left[\sum_{j=o}^{n-1} q_{j:i}\, k^{n-1-j}\right]^2 = k^{2n-2} + 2q_{1:i}\, k^{2n-3} + (q_{1:i}^2 + 2q_{2:i})k^{2n-4} + 2(q_{1:i} q_{2:i} + q_{3:i})\, k^{2n-5} + V_1(k)$$

in which $V_1(k)$ is a finite polynomial of degree $2n - 6$ in k. Putting the relationship (4.2.14) in this gives

$$g_i^2 = k^{2n-2} - 2w_i k^{2n-3} + 2(q_2 + 3w_i^2)k^{2n-4} + 2(q_3 - 2w_i q_2 - 2w_i^3)k^{2n-5} + V_1(k) .$$

Since $\sum_{i=1}^{n} w_i = 0$ and $\sum_{i=1}^{n} w_i^r = S_r$, summing the above over i gives

$$\sum_{i=1}^{n} g_i^2 = nk^{2n-2} + (2nq_2 + 3S_2)k^{2n-4} + (2nq_3 - 4S_3)k^{2n-5} + V_2(k)$$

in which $V_2(k)$ is a finite polynomial of degree $2n - 6$ in k. Using the relationships connecting S_r and q_r from Lemma 4.2.10 the above becomes

4.2

$$(4.2.15)\qquad \sum_{i=1}^{n} g_i^2 = nk^{2n-2} + 2(n-3)q_2k^{2n-4} + 2(n-6)q_3k^{2n-5} + V(k)$$

in which $V(k)$ is a finite polynomial of degree $2n - 6$ in k. In fact $V(k)$ is the same function as that which was previously designated by $V_2(k)$.

From the definition of g, the common denominator for R_2 is obviously g^2. Since $q_1 = 0$, g^2 can be written,

$$(4.2.16)\qquad g^2 = \prod_{i=1}^{n} (k + w_i)^2 = (k^n + q_2k^{n-2} + q_3k^{n-3} + \cdots + q_n)^2 .$$

Putting (4.2.15) in (4.2.12) gives

$$(4.2.17)\qquad R_2 = \frac{nk^{2n-2} + 2(n-3)q_2k^{2n-4} + 2(n-6)q_3k^{2n-5} + V(k)}{g^2}$$

in which, as previously stated, $V(k)$ is a finite polynomial of degree $2n - 6$ in k.

Before stating the next lemma, the below designated functions of k will be defined as follows:

$$H(k) \equiv \frac{P(k)}{K} = \frac{\prod_{i=1}^{n} (w_i + k)^{2/n}}{k^2 + m_2} \qquad \text{(see Lemma 4.2.11)}$$

$$M(k) \equiv \frac{nk}{RK} \quad \text{(see Lemma 4.2.8)}$$

$$G(k) \equiv \frac{1}{\delta + \log H(k)} \quad \text{in which } \delta \text{ is an arbitrary non-zero parameter, independent of } k.$$

$$T(k) \equiv [M(k)]^{-1}$$

$G'(k)$ and $T'(k)$ will represent the derivatives of these functions with respect to k.

As a further matter of notation, the argument k will be omitted in writing the above functions. Exception will be made to this in case a special value of the argument is to be used, or if such omission gives

4.2

rise to ambiguity.

Lemma 4.2.13. $\lim_{k\to\infty} G' = 0$; $\lim_{k\to\infty} T' = 0$; and $\lim_{k\to\infty} \frac{G'}{T'} = \frac{1}{\delta^2}$.

Proof: From the definition of T,

$$T' = - \frac{k^2(-kR' - R) + m_2(-kR' + R)}{nk^2}$$

in which $R' = \frac{dR}{dk} = - R_2$ as defined in Lemma 4.2.8. Thus,

$$T' = - \frac{k^2(kR_2 - R) + (kR_2 + R)m_2}{nk^2} \quad . \tag{4.2.18}$$

From Lemma 4.2.3

$$R = \frac{d[\log g(k)]}{dk} = \frac{g'}{g}$$

and from (4.2.16)

$$g = k^n + q_2 k^{n-2} + q_3 k^{n-3} + q_4 k^{n-4} \cdots q_n .$$

Consequently,

$$R = \frac{nk^{n-1} + (n-2)q_2 k^{n-3} + (n-3)q_3 k^{n-4} \cdots q_{n-1}}{g} \quad . \tag{4.2.19}$$

Then using R_2 as given in Lemma 4.2.12 gives

$$kR_2 - R = \frac{-4q_2 k^{2n-3} - 9q_3 k^{2n-4} + X(k)}{g^2} \tag{4.2.20}$$

in which $X(k)$ is a finite polynomial of degree $2n - 5$ in k. Similarly,

$$kR_2 + R = \frac{2nk^{2n-1} + 4(n-2)q_2 k^{2n-3} + (4n-15)q_3 k^{2n-4} + X_1(k)}{g^2} \tag{4.2.21}$$

in which $X_1(k)$ is a finite polynomial of degree $2n - 5$ in k.

Putting (4.2.20) and (4.2.21) in (4.2.18) gives

4.2

$$T' = \frac{4q_2k^{2n-3} + 9q_3k^{2n-4} - 2nm_2k^{2n-5} + X_2(k)}{ng^2} .$$

Using the known relationships between q_r, S_r, and m_r in the above gives

$$(4.2.22) \qquad T' = - \frac{4m_2k^{2n-3} - 3m_3k^{2n-4} + X_3(k)}{g^2}$$

in which $X_3(k)$ is a finite polynomial of degree $2n - 5$ in k.

It is obvious from (4.2.16) that the coefficient of k^{2n} in g^2 cannot vanish. Thus for any k and any set of observations the denominator of T' as given by (4.2.22) is of degree $2n$ in k. Similarly, $m_2 \neq 0$ and consequently the numerator of T' as given by (4.2.22) is of degree $2n - 3$ in k for any k and any set of observations. Thus T' is of degree -3 in k, and since $X_3(k)$ is a finite polynomial in k it follows that $\lim_{k\to\infty} T' = 0$.

From the definitions of the functions concerned, G can be written

$$G = \frac{1}{\sigma + \log P(k) - \log K}$$

and consequently

$$(4.2.23) \qquad G' = - \frac{2}{n} \left[R - \frac{nk}{K}\right] [G]^2 .$$

Using R from (4.2.19) gives

$$R - \frac{nk}{K} = \frac{(nm_2 - 2q_2)k^{n-1} - 3q_3k^{n-2} + Y_3(k)}{gK}$$

in which $Y_3(k)$ is a finite polynomial of degree $n - 3$ in k. Using the relationships $-2q_2 = S_2 = nm_2$ and $3q_3 = S_3 = nm_3$ in the above and putting the resulting expression for $R - nk(K)^{-1}$ in (4.2.23) gives

$$(4.2.24) \qquad G' = - \left[\frac{4m_2k^{n-1} - 2m_3k^{n-2} + Y_1(k)}{gK}\right] [G]^2 .$$

4.2

in which $Y_1(k)$ is a finite polynomial of degree $n - 3$ in k.

From the Corollary of Lemma 4.2.10, $\lim_{k\to\infty} H = 1$ and thus $\lim_{k\to\infty} \log H = 0$. It follows then from the definition of G that $\lim_{k\to\infty} G = \frac{1}{\delta}$.

The first factor of the right member of (4.2.24) is obviously of order k^{-3} since the coefficient of the leading term of neither the numerator nor the denominator can vanish. Thus the quantity approaches zero as k becomes infinite. By hypothesis δ is independent of k and non-vanishing. Consequently the second factor of the right member of (4.2.24) remains finite for all k. Therefore it follows that

$$\lim_{k\to\infty} G' = 0 . \tag{4.2.25}$$

From (4.2.22) and (4.2.24)

$$\frac{G'}{T'} = \frac{[4m_2k^{n-1} - 2m_3k^{n-2} + Y_1(k)]\,[k^n + q_2k^{n-2} \cdots + q_n]\,[G]^2}{[k^2 + m_2]\,[4m_2k^{2n-3} - 3m_3k^{2n-4} + X_3(k)]}$$

or

$$\frac{G'}{T'} = \frac{[4m_2k^{2n-1} + Y_2(k)]\,[G]^2}{4m_2k^{2n-1} + X_4(k)}$$

in which each $Y_2(k)$ and $X_4(k)$ is a finite polynomial of degree $2n - 2$ in k. Since $m_2 \neq 0$ it is obvious from the above that

$$\lim_{k\to\infty} \frac{G'}{T'} = \lim_{k\to\infty} [G]^2 = \frac{1}{\delta^2} .$$

Lemma 4.2.14. If n_1 is the frequency of w_1 in the set of observations $\{w_i\}$, $i = 1, 2, 3, \cdots n$, w_1 being the smallest observation, then $\lim_{k\to -w_1} \frac{R_2}{R^2} = \frac{1}{n_1}$.

Proof: From the definition of R_2,

4.2

$$R_2 = \frac{n_1}{(w_1 + k)^2} + \frac{1}{(w_{n_1+1} + k)^2} + \cdots + \frac{1}{(w_n + k)^2} .$$

Clearing of fractions gives, in the notation of Lemma 4.2.12

$$R_2 = \frac{n_1 \prod_{i=n_1+1}^{n} (w_i + k)^2 + \sum_{i=n_1+1}^{n} g_i^2}{\prod_{i=1}^{n} (w_i + k)^2} .$$

From definition of R,

$$R^2 = \frac{\left[n_1 \prod_{i=n_1+1}^{n} (w_i + k) + \sum_{i=n_1+1}^{n} g_i \right]^2}{\prod_{i=1}^{n} (w_i + k)^2} .$$

Then for $k \neq -w_1$

$$\frac{R_2}{R^2} = \frac{n_1 \prod_{i=n_1+1}^{n} (w_i + k)^2 + \sum_{i=n_1+1}^{n} g_i^2}{\left[n_1 \prod_{i=n_1+1}^{n} (w_i + k) + \sum_{i=n_1+1}^{n} g_i \right]^2} .$$

$\sum_{i=n_1+1}^{n} g_i$ consists of a finite number of terms, each having a finite number of factors. Further, $(w_1 + k)$ is a factor of each term. Thus by taking k sufficiently close to $-w_1$, g_i can be made arbitrarily small for every $i \geq n_1 + 1$. Consequently

$$\lim_{k \to -w_1} \sum_{i=n_1+1}^{n} g_i = 0$$

and thus

4.2-3

$$\lim_{k \to -w_1} \frac{R_2}{R^2} = \frac{n_1 \prod_{i=n_1+1}^{n} (w_i + k)^2}{\left[n_1 \prod_{i=n_1+1}^{n} (w_i + k)\right]^2} = \frac{1}{n_1} .$$

Lemma 4.2.15. $\lim_{k \to \infty} \frac{R_2}{R^2} = \frac{1}{n}$.

Proof: Writing R_2 as given by (4.2.17) of Lemma 4.2.12 and R as given by (4.2.19) of Lemma 4.2.13 gives,

$$\frac{R_2}{R^2} = \frac{nk^{2n-2} + 2(n-3)q_2 k^{2n-4} + 2(n-6)q_3 k^{2n-5} + V(k)}{\left[nk^{n-1} + (n-2)q_2 k^{n-3} + (n-3)q_3 k^{n-4} + \cdots q_{n-1}\right]^2} .$$

Since each of the numerator and denominator is a polynomial of degree $2n-2$ in k, repeated use of l'Hopital's rule gives $\lim_{k \to \infty} \frac{R_2}{R^2} = \frac{1}{n}$.

4.3 Roots of the maximum likelihood equations and the existence of an admissible solution. Admissible values of the independent parameters are as follows: $0 < p < \infty$; $0 < a < \infty$; and $-w_1 \leq k < \infty$, in which $w_1 \leq w_i$ for every $i > 1$. Any set of values of a, p, and k which satisfies the maximum likelihood equations and consists of admissible values for each of the parameters will be termed an admissible solution.

Theorem 4.3.1. The M. L. equations will have no admissible solution in which $p \leq \frac{1}{2}$.

Proof: One of the M. L. equations, (4.1.4), is

(4.3.1) $\quad (2p - 1)R = 2nak.$

From condition 1' of section 1.2, $a > 0$ and from the Corollary of Lemma 4.2.1, $k > 0$. Thus the right member of (4.3.1) is positive for all admissible values of the parameters. From the definition of R and the fact that $w_i + k \geq 0$ for all k, it follows that $R > 0$. Thus if $2p - 1 < 0$ the left member of (4.3.1) is negative. Consequently no set of admissible

4.3

values for a, p, and k in which $p < \frac{1}{2}$ can satisfy (4.3.1). If $p = \frac{1}{2}$, it would be necessary for either a or k to be zero to satisfy (4.3.1). From condition 1, of section 1.2 and Corollary of Lemma 4.2.1 neither a nor k can be zero. Thus for $p = \frac{1}{2}$ no admissible values of a and k can satisfy (4.3.1).

Thus no admissible solution for (4.3.1) can exist in which $p \leq \frac{1}{2}$ and since this equation is a member of the M.L. system of equations the theorem is established.

Theorem 4.3.2. The range of values for p as determined by the M.L. equations is $\frac{1}{2} < p < \infty$.

Proof: The relationship, (4.1.6), for the determination of the M.L. estimate of p for any k is

$$p = \frac{RK}{2(RK - nk)} = \frac{1}{2[1 - M(k)]} \qquad (4.3.2)$$

in the notation of Lemma 4.2.8. From the Corollary of Lemma 4.2.6, $R > \frac{nk}{K}$ and thus $M(k) < 1$ for all k. Further from Lemma 4.2.5

$$\lim_{k \to -w_1} [M(k)]^{-1} = \infty, \quad \text{or} \quad \lim_{k \to -w_1} M(k) = 0 .$$

Using these in limits in (4.3.2) gives $p > \frac{1}{2}$ for all k and $\lim_{k \to -w_1} p = \frac{1}{2}$.

From the Corollary of Lemma 4.2.4 $\lim_{k \to \infty} M(k) = 1$ and thus $\lim_{k \to \infty} p = \infty$. From Lemma 4.2.8, M(k) is a monotonic increasing function of k. Consequently p, as determined from (4.3.2) will be monotonically increasing with k, and $\frac{1}{2} < p < \infty$.

Theorem 4.3.3. The value of a as determined by the M.L. equations is positive.

Proof: From (4.1.6) the relationship for the determination of a

* In this section the frequent occurrence of special values for k and the later introduction of another variable, ξ, make it inadvisable to continue omitting the argument in writing the functions M, H, G, and T.

4.3

for any k is $a = 2K\{1 - M(k)\}^{-1}$. From the Corollary of Lemma 4.2.6 $M(k) < 1$ for all k and consequently $1 - M(k) > 0$. Since $K = k^2 + m_2 > 0$, it is obvious that $a > 0$ for all k.

Theorem 4.3.4. There will exist an admissible solution for the M.L. equations determined by any set of observations $\{w_i\}$ meeting the conditions a) through e) of section 4.2 and for which the frequency curve is not monotonic decreasing.

Proof: From Theorems 4.3.2 and 4.3.3 the values of a and p as obtained from the M.L. equations through (4.1.6) and (4.1.7) respectively for any admissible value of k are admissible values. Further, the condition that the frequency curve is not monotonic decreasing is equivalent to the condition $p > \frac{1}{2}$. Thus to establish the theorem it is sufficient to show that some admissible value of k is a zero of the left member of (4.1.8) in which p is regarded as a dependent parameter to be determined by (4.1.6). That is, the theorem will be established by showing that some admissible k is a zero for

$$U(k) = \log p_k - \psi(p_k) + \log H(k), \tag{4.3.3}$$

in which p_k is defined by the right member of (4.1.6) and H(k) is defined in Lemma 4.2.13.

From Theorem 4.3.2 $\lim_{k \to -w_1} p_k = \frac{1}{2}$ and consequently $\log p_k$ and $\psi(p_k)$ each remains finite as k approaches $-w_1$. Since w_i is finite for every i and n is finite, $\prod_{i=1}^{n} (w_i + k)$ can be made arbitrarily small by taking k sufficiently close to $-w_1$. $k^2 + m_2$ remains finite as k approaches $-w_1$. Thus it follows,

$$\lim_{k \to -w_1} [\log H(k)] = -\infty, \quad \text{and} \quad \lim_{k \to -w_1} U(k) = -\infty .$$

Consequently there will be some k, say k_1, such that $k_1 > -w_1$ and

$$U(k_1) < 0. \tag{4.3.4}$$

4.3

We next show that there exists some admissible k for which $U(k) > 0$. Let $\mathcal{E}$ be any real and positive quantity and define the function,

$$(4.3.5) \qquad \mathcal{E}(\mathcal{E}, k) = \mathcal{E} - \frac{R(k^2 + m_2)}{nk} = \mathcal{E} - T(k)$$

in the notation of Lemma 4.2.13.

If the $\mathcal{E}$-axis is taken as the vertical axis, the graph for $\mathcal{E}(\mathcal{E}, k) = 0$ will have the following properties:

1) $\mathcal{E} > 1$ for every k, from the Corollary of Lemma 4.2.6.

2) $\lim_{k \to \infty} \mathcal{E} = 1$, from the Corollary of Lemma 4.2.4.

3) $\mathcal{E}$ is a monotonic decreasing function of k, from Lemma 4.2.8.

4) $\lim_{k \to -w_1} \mathcal{E} = \infty$, from Lemma 4.2.5.

5) $\mathcal{E}(\mathcal{E}, k)$ is a continuous function of each $\mathcal{E}$ and k over the range $1 < \mathcal{E} < \infty$ and $-w_1 < k < \infty$, from the definitions of the functions concerned.

Thus the graph for $\mathcal{E}(\mathcal{E}, k) = 0$ will be a continuous curve over the interval $k > -w_1$. It will approach from the right as a vertical asymptote the line $k = -w_1$, and will approach from above as a horizontal asymptote the line $\mathcal{E} = 1$.

The graph for $\mathcal{E}(\mathcal{E}, k) = 0$ is shown by the broken line in figure 4, page 72.

Let δ be any arbitrarily assigned positive parameter and define the function $\mathcal{G}(\mathcal{E}, k)$ as follows,

$$(4.3.6) \qquad \mathcal{G}(\mathcal{E}, k) = e^{-\left[\frac{\delta \mathcal{E} - 1}{\mathcal{E}}\right]} - H(k),$$

with H(k) being defined as in Lemma 4.2.13. The graph for $\mathcal{G}(\mathcal{E}, k) = 0$ will be a one parameter family of curves. Solving $\mathcal{G}(\mathcal{E}, k) = 0$ for $\mathcal{E}$ gives

$$(4.3.7) \qquad \mathcal{E} = \frac{1}{\delta + \log H(k)} = G(k)$$

in the notation of Lemma 4.2.13.

4.3

If the $\mathcal{E}$- axis is taken as the vertical axis, the graph for $\mathcal{G}(\mathcal{E}, k) = 0$ as drawn from (4.3.7) will have the following properties for any $\delta > 0$.

1) $\mathcal{E} > \frac{1}{\delta}$ for all k, from the Corollary of Lemma 4.2.9.

2) $\lim_{k \to -w_1} \mathcal{E} = \frac{1}{\delta}$, from the Corollary of Lemma 4.2.10.

3) From the Corollary of Lemma 4.2.9, $H(k) < 1$ for all admissible values of k and from Lemma 4.2.11, H(k) is a monotonic increasing function of k. It is obvious then that $\mathcal{E}$ as defined by (4.3.7) is a monotonic decreasing function of k for all $\delta > 0$.

4) H(k) is a continuous function of k, increasing monotonically from zero to arbitrarily close to unity as k increases from $-w_1$ without limit. Thus log H(k) will vary continuously from an unbounded negative quantity to arbitrarily close to zero as k varies over its range of values. It is obvious that there will exist, for any finite $\delta > 0$, some k, say k_2, such that: $k_2 > -w_1$ and $\log H(k_2) = -\delta$. From this it follows $\lim_{k \to k_2} \mathcal{E} = \infty$ for any $\delta > 0$.

5) $\mathcal{G}(\mathcal{E}, k)$ is a continuous function of each $\mathcal{E}$ and k over the range $\frac{1}{\delta} < \mathcal{E} < \infty$ and $k_2 < k < \infty$, from the definitions of the functions concerned.

Thus the graph for $\mathcal{G}(\mathcal{E}, k) = 0$ will be a continuous curve for $k > k_2$ for any $\delta > 0$. It will approach from the right the line $k = k_2$ as a vertical asymptote and will approach from above as a horizontal asymptote the line $\mathcal{E} = \frac{1}{\delta}$, as $\mathcal{E}$ and k respectively become infinite.

For any $\delta > 0$, the vertical asymptote for $\mathcal{G}(\mathcal{E}, k) = 0$ is to the right of the vertical asymptote for $\mathcal{E}(\mathcal{E}, k) = 0$ since $k_2 > -w_1$. Further, the curve for $\mathcal{E}(\mathcal{E}, k) = 0$ will have a finite ordinate at $k = k_2$ for any $\delta > 0$.

For any $\delta > 1$ the horizontal asymptote for the curve $\mathcal{E}(\mathcal{E}, k) = 0$ is above the horizontal asymptote for the curve $\mathcal{G}(\mathcal{E}, k) = 0$. Further,

4.3

the graph for $\mathcal{Y}(\mathcal{E}, k) = 0$ will have a finite abscissa for $\mathcal{E} = 1$ for any $\delta > 1$.

It follows from the above that for any $\delta > 1$, the curve for $\mathcal{Y}(\mathcal{E}, k) = 0$ will intersect the curve for $\mathcal{E}(\mathcal{E}, k) = 0$ for some finite $k > k_2 > -w_1$. The point of intersection for these two curves is obviously a function of δ. For this reason k_δ will represent the abscissa of the point and $\mathcal{E}_\delta$ will represent its ordinate.

A typical curve for $\mathcal{Y}(\mathcal{E}, k) = 0$ for some $\delta > 1$ is shown by the solid curve of figure 4, page 72.

We now investigate the behavior of $\mathcal{E}_\delta$ and k_δ as $\delta \to 1$ and specifically the existence of $\mathcal{E}_\delta$ and k_δ if $\delta = 1$.

That is, as $\delta \to 1$, the horizontal asymptote for the curve for $\mathcal{Y}(\mathcal{E}, k) = 0$ approaches the horizontal asymptote for the curve for $\mathcal{E}(\mathcal{E}, k) = 0$ and for $\delta = 1$ they have a common horizontal asymptote, $\mathcal{E} = 1$. The question to be answered is: For $\delta = 1$ will the two curves have a common point so that $\mathcal{E}_\delta$ and k_δ will still exist when $\delta = 1$?

There will be two possibilities:

1) $\lim_{\delta \to 1} k_\delta$ is finite. That is, the points of intersection for the two curves will always be to the left of some fixed point, and consequently $\mathcal{E}_\delta$ would be greater than some fixed quantity which is greater than unity for all $\delta \geq 1$. Under this condition the solution of the problem is obvious as will be pointed out later.

2) $\lim_{\delta \to 1} k = \infty$ and $\lim_{\delta \to 1} \mathcal{E}_\delta = 1$. Under this condition the point of intersection of the two curves moves to the right continuously without approaching a finite limit point as $\delta \to 1$.

The following is based on the assumption that the second of the above possibilities is the one which prevails.

Let $\mathcal{E}_1$ be the ordinate for the curve $\mathcal{E}(\mathcal{E}, k) = 0$ at any point

4.3

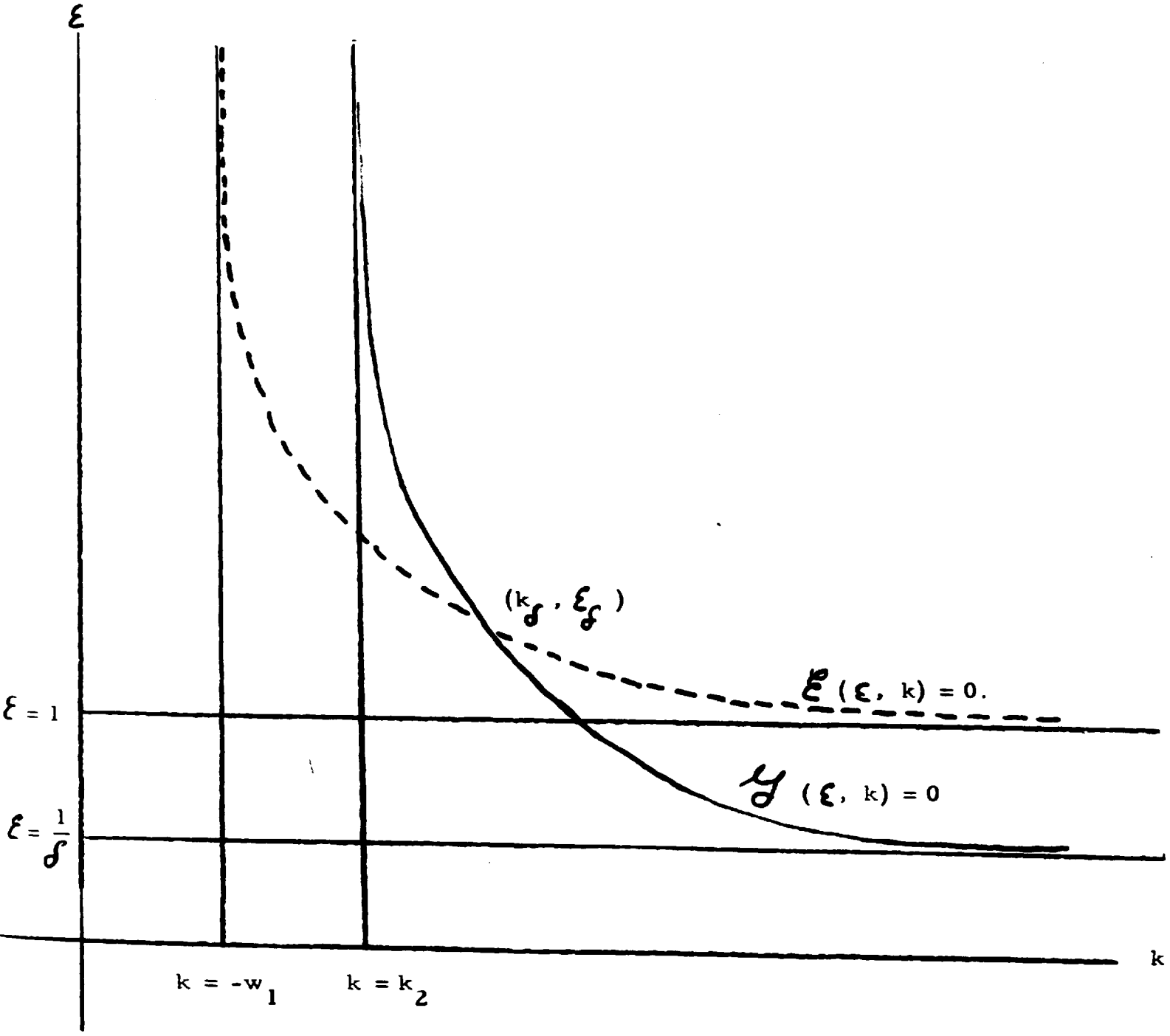


Fig. 4.

Graph showing the intersection of the curve for $\Xi(\varepsilon, k) = 0$ and a typical member of the system of curves for $\Upsilon(\varepsilon, k) = 0$ for some $\delta > 1$.

4.3

on the curve and let ε_2 be the ordinate for the curve $\mathcal{Y}(\varepsilon, k) = 0$ for any point on the curve. Then ε_1 and ε_2 will possess the following properties:

a) As δ decreases the vertical asymptote for $\mathcal{Y}(\varepsilon, k) = 0$ will move to the left. However, so long as δ remains finite, this asymptote will remain at a finite distance to the right of the line $k = -w_1$. For $\delta = 1$ the vertical asymptote for $\mathcal{Y}(\varepsilon, k) = 0$ is that value of k for which $H(k) = e^{-1}$ and for this value of k, ε_1 is finite. Consequently for any $\delta \geq 1$ there will exist some k, say k_4, such that, if $k = k_2$ is the vertical asymptote for $\mathcal{Y}(\varepsilon, k) = 0$, then

$$(4.3.8) \qquad \varepsilon_2(k_4) > \varepsilon_1(k_4) \quad \text{and} \quad k_4 > k_2 > -w_1 .$$

b) As δ decreases continuously the horizontal asymptote of $\mathcal{Y}(\varepsilon, k) = 0$ will move upward continuously and as $\delta \to 1$ the asymptote will approach the line $\varepsilon_2 = 1$ from below.

c) The definitions of the functions involved give

$$(4.3.9) \qquad \frac{\varepsilon_1}{\varepsilon_2} = \delta\,[T(k)] + [T(k)][\log H(k)].$$

Then from the Corollary of Lemma 4.2.4 and the Corollary of Lemma 4.2.10

$$(4.3.10) \qquad \lim_{k \to \infty} \frac{\varepsilon_1}{\varepsilon_2} = \delta.$$

Thus for sufficiently large k and for any $\delta > 1$, the ordinate for $\mathcal{Y}(\varepsilon, k) = 0$ is greater than the ordinate for $\mathcal{Y}(\varepsilon, k) = 0$. Further, as $\delta \to 1$ the ordinates of the two curves approach equality. From (4.3.9) alone it cannot be stated whether the approach to equality is from $\frac{\varepsilon_1}{\varepsilon_2} > 1$ or $\frac{\varepsilon_1}{\varepsilon_2} < 1$. However, this conclusion can be drawn from the following: For any $\delta > 1$ the asymptote for ε_2 is below the asymptote for ε_1 and consequently for any $\delta > 1$ there will exist some k, say

4.3

k_5, such that $\xi_1(k) > \xi_2(k)$ for any $k > k_5$. Thus it can be concluded that the ratio of (4.3.10) approaches unity from above. That is,

$$\lim_{\delta \to 1^+} \left[\frac{\xi_1}{\xi_2}\right] = 1^+ .$$

d) ξ_2 as defined here is identical with G(k) as defined in Lemma 4.2.13 and from (4.2.23)

$$\frac{d\,\xi_2}{dk} = G' = -K_1\left[\frac{1}{\delta + \log H(k)}\right]^2 ,$$

in which K_1 is positive and independent of δ. Thus for any fixed k, the rate at which ξ_2 approaches its horizontal asymptote increases monotonically with decreasing δ.

e) From Lemma 4.2.13, since T(k) and G(k) as defined there are identical with ξ_1 and ξ_2 respectively,

$$\lim_{k \to \infty} \frac{\xi_2'}{\xi_1'} = \frac{1}{\delta^2}$$

in which ξ_1' and ξ_2' represent the derivatives with respect to k. Thus the relative rate at which ξ_2 and ξ_1 approach their respective horizontal asymptotes increases with decreasing δ and for $\delta = 1$ the ratio of their slopes becomes unity or the "tails" of these two curves of figure 4 become "parallel".

f) From Lemma 4.2.6 $\frac{R}{nk} > \frac{1}{k^2}$ and thus from definition

$$\xi_1 = \frac{R}{nk}(k^2 + m_2) > 1 + \frac{m_2}{k^2} > 1 .$$

That is, ξ_1 for all k is greater than a quantity which is greater than 1.

In view of b), c), d), e), and f) above, it can be stated: For $\delta = 1$ even though the horizontal asymptotes for the curves $\xi(\xi, k) = 0$ and $\eta(\xi, k) = 0$ coincide at $\xi_1 = \xi_2 = 1$. there will exist some k, say k_6, such that

4.3

(4.3.11) $\xi_2(k_6) \leq \xi_1(k_6)$ and $k_6 > k_2 > -w_1$.

Since each $\xi(\varepsilon, k)$ and $\eta(\varepsilon, k)$ is a continuous function of k for $k > k_2$, it follows as a consequence of (4.3.8) and (4.3.11) that the graphs for $\xi(\varepsilon, k) = 0$ and $\eta(\varepsilon, k) = 0$ must intersect at some finite point for any $\delta \geq 1$.

Algebraically the above means: For any $\delta \geq 1$ there will exist a k_δ and an ε_δ such that

$$(4.3.12) \qquad \varepsilon_\delta = \frac{[R(k_\delta)]\,[k_\delta^2 + m_2]}{nk_\delta},$$

and

$$(4.3.13) \qquad \frac{\prod_{i=1}^{n} (k_\delta + w_i)^{2/n}}{k_\delta^2 + m_2} = e^{-\frac{(\delta\varepsilon_\delta - 1)}{\varepsilon_\delta}}$$

in which

$$R(k_\delta) = \sum_{i=1}^{n} \frac{1}{w_i + k_\delta}.$$

Davis (Ref. 7) gives the following asymptotic expansion for $\psi(p)$

$$\psi(p) = \log p - \frac{1}{2p} + \sum_{r=1}^{\infty} \frac{(-1)^r B_r}{2r\, p^{2r}}$$

in which B_r is the r^{th} Bernoulli number. Then

$$\psi(p) < \log p - \frac{1}{2p} - \frac{1}{12p^2} + \frac{1}{120p^4}$$

since the first term omitted is negative and the expansion is asymptotic. The above is equivalent to,

$$\log p - \psi(p) > \frac{1}{2p} + \frac{1}{3}\left[\frac{1}{2p}\right]^2 - \frac{2}{15}\left[\frac{1}{2p}\right]^4.$$

From Theorem 4.3.2., $p_k > \frac{1}{2}$ for any k. Thus for any k,

4.3

$$\left[\frac{1}{2p_k}\right]^4 < \left[\frac{1}{2p_k}\right]^2$$

and as a result

$$\frac{1}{3}\left[\frac{1}{2p_k}\right]^2 - \frac{2}{15}\left[\frac{1}{2p_k}\right]^4 > \frac{1}{5}\left[\frac{1}{2p_k}\right]^2 .$$

Consequently $\log p_k - \psi(p_k) > \frac{1}{2p_k} + \frac{1}{5}\left[\frac{1}{2p_k}\right]^2$.

Putting this inequality in (4.3.3) gives

$$(4.2.14) \qquad U(k) > \frac{1}{2p_k} + \frac{1}{5}\left[\frac{1}{2p_k}\right]^2 + \log H(k).$$

In the expression for p_k as defined by (4.1.6) replacing $T(k_\delta)$ by its equivalent as given by (4.3.12) gives

$$(4.3.15) \qquad \frac{1}{2p_k} = 1 - \frac{1}{\xi_\delta} .$$

Then from (4.2.13) and the definition of H(k) as given in Lemma 4.2.13,

$$(4.3.16) \qquad \log H(k_\delta) = -\frac{\delta\xi_\delta - 1}{\xi_\delta} .$$

Letting $k = k_\delta$ and using the values as given by (4.3.15) and (4.3.16) in (4.3.14) gives

$$(4.3.17) \qquad U(k_\delta) > \frac{1}{5}\left[1 - \frac{1}{\xi_\delta}\right]^2 - (\delta - 1) .$$

If the first possibility referred to earlier in the proof is the prevailing condition, $\frac{1}{5}[1 - \frac{1}{\xi_\delta}]^2$ will have a finite lower limit and δ can be taken close enough to unity to make $(\delta - 1)$ less than this finite lower limit. Then for the corresponding value of k_δ, say k'_3, it follows that

$$(4.3.18) \qquad U(k'_3) > 0 \quad \text{and} \quad k'_3 > k_2 > -w_1 .$$

If the second possibility is the prevailing condition, then from (4.3.12) and (4.3.13) k_δ and ξ_δ will exist for $\delta = 1$. If k''_3 and ξ''_3 represent the values of k_δ and ξ_δ respectively which correspond to $\delta = 1$, then it follows from (4.3.17) that

4.3 - 4

$$(4.3.19) \qquad U(k_3'') > \frac{1}{5}\left[1 - \frac{1}{\xi_3''}\right]^2 > 0 \quad \text{and} \quad k_3'' > k_2 > -w_1 .$$

Since each k_3' and k_3'' is greater than $-w_1$ each is admissible. Thus it can be stated that regardless of which of the two possibilities prevail, there will exist some admissible k, say k_3, such that

$$(4.3.20) \qquad U(k_3) > 0 \quad \text{and} \quad k_3 > k_2 > -w_1 .$$

$P(k)$, $R(k)$, $\psi(p_k)$ and $\log p_k$ are all continuous functions of k; the latter two because p_k is a continuous function of k. Thus $U(k)$ as defined by (4.3.3) is a continuous function of k. From (4.3.4) $U(k_1) < 0$ and from (4.3.20) $U(k_3) > 0$. Consequently, there must exist some k, say $\hat{k}$, such that $k_3 > \hat{k} > k_1$ and $U(\hat{k}) = 0$.

Since each k_1 and k_3 is an admissible value of k, and $\hat{k}$ lying between them is obviously an admissible value of k and the theorem is established.

4.4. Existence of a sufficient set of estimates. From (4.1.1) $f(w)$ can be written

$$(4.4.1) \qquad f(w) = e^{[-\log\sqrt{p} + \log 2 + p \log a + (2p-1)\log(w+k) - a(w+k)^2]} .$$

If the following notation is adopted; $A_1 = 2p - 1$, $A_2 = a$, $X_1 = \log(w + k)$; $X_2 = (w + k)^2$; and $B_1 = \log 2 + p \log a - \log\sqrt{p}$; then (4.4.1) can be written,

$$(4.4.2) \qquad f(w) = e^{B_1 + A_1X_1 + A_2X_2} .$$

If $(w + k)$ is regarded as the variate, i.e. let $w + k = x$, then $f(w, a, k, p)$ becomes $f(x, a, p)$ which is a continuous and non-zero function over the admissible range for the parameters. Further the partial derivative of the function will exist for $p \geq 1$.

Thus, if k is regarded as known, according to the familiar theorem

4.4 - 5

due to Koopman (Ref. 17) there will exist a set of jointly sufficient estimators for a and p for $p \geq 1$. However, there does not exist a set of jointly sufficient estimators for a, p, and k.

4.5. Analysis of the roots of the M. L. equations relative to their maximizing the likelihood function*. Since the roots of the M. L. equations are zeros for the partial derivatives of the logarithm of the likelihood function with respect to the parameters, they may represent either a maximum point, a point of inflection, or even a minimum point for this function. The purpose of this section is the establishment of the type of point which the M. L. estimates represent for the likelihood function.

Theorem 4.5.1. If $(\hat{a}, \hat{p}, \hat{k})$ is a solution for the M. L. equations as given in section 4.1, then:

a) $(\hat{a}, \hat{p}, \hat{k})$ satisfy all the necessary conditions for maximizing the likelihood function, $L(w_i, a, p, k)$, as given by (4.1.2) and

b) $(\hat{a}, \hat{p}, \hat{k})$ cannot minimize the function $L(w_i, a, p, k)$.

Proof: In the proof the following notation will be employed:

$$L_{\hat{a}:\hat{k}} = \frac{\partial^2 L}{\partial a \, \partial k}$$

evaluated at $(\hat{a}, \hat{p}, \hat{k})$, with corresponding notation for the second partial derivatives with respect to the other parameters.

$$\hat{R} = \sum_{i=1}^{n} \frac{1}{w_i + \hat{k}} ; \quad \hat{R}_2 = \sum_{i=1}^{n} \frac{1}{(w_i + \hat{k})^2} ; \quad \psi'(\hat{p}) = \frac{d^2[\log \Gamma(\hat{p})]}{dp^2}$$

evaluated at $p = \hat{p}$. $|H|$ will represent the Hessian for $L(w_i, a, p, k)$ with each partial derivative being evaluated at $(\hat{a}, \hat{p}, \hat{k})$. The usual notation relative to minors of various order of $|H|$ will be employed. In all

---

* The function $L(w_i, a, p, k)$ is the logarithm of the joint frequency function of the sample. It will be referred to as the likelihood function.

4.5

cases the elements of $|H|$ and its minors are evaluated at $(\hat{a}, \hat{p}, \hat{k})$.

Since $(\hat{a}, \hat{p}, \hat{k})$ satisfy the M. L. equations: (4.1.3), (4.1.4), and (4.1.5); the following relationships will be satisfied:

$$(4.5.1) \qquad \hat{p} = (m_2 + \hat{k}^2)\,\hat{a}$$

$$(4.5.2) \qquad (2\hat{p} - 1)\hat{R} = 2n\hat{a}\hat{k}$$

$$(4.5.3) \qquad n \log \hat{a} - n\psi(\hat{p}) + 2 \sum_{i=1}^{n} \log (w_i + \hat{k}) = 0 .$$

The Hessian for $L(w_i, a, p, k)$, with the parameters being taken in the order a, p, k becomes,

$$|H| = \frac{n^3}{\hat{a}^2} \begin{vmatrix} -\hat{p} & 1 & -2\hat{a}\hat{k} \\ 1 & -\psi'(\hat{p}) & 2\hat{R}_1 \\ -2\hat{a}\hat{k} & 2\hat{R}_1 & -[(2\hat{p} - 1)\hat{R}_{2:1} + 2\hat{a}] \end{vmatrix} .$$

in which $n\hat{R}_1 = \hat{R}$ and $n\hat{R}_{2:1} = \hat{R}_2$.

Expanding the above determinant by method of minors and using the relation (4.5.2) gives

$$|H| = \frac{n^3}{\hat{a}^2}(2\hat{p} - 1)\left\{\frac{1}{\hat{k}}(\hat{k}\hat{R}_{2:1} + \hat{R}_1) - 2\hat{R}_1 - \psi'(\hat{p})\left[\frac{\hat{p}}{\hat{k}}(\hat{R}_{2:1}\hat{k} + \hat{R}_1) - (2\hat{p} - 1)\hat{R}_1^2\right]\right\} + 2\hat{R}_1^2 \cdot \frac{n^3}{\hat{a}^2} .$$

By writing $\psi'(\hat{p}) = \frac{1}{\hat{p}} + \Delta(\hat{p})$ and using the relationships (4.5.1), (4.5.2), and (4.5.3) the above clears up to

$$(4.5.4) \qquad |H| = \frac{n^3}{\hat{a}^2}\left\{[-2\hat{p}\,\Delta(\hat{p})]\left[\frac{\hat{R}_{2:1}\hat{k}(m_2 + \hat{k}^2) + \hat{R}_1(m_2 - \hat{k}^2)}{\hat{R}_1(m_2 + \hat{k}^2)^2}\right] + \frac{\hat{R}_1}{\hat{p}}\right\}.$$

Differentiating (4.5.1) and (4.5.3) gives,

$$\psi'(\hat{p}) = \frac{1}{\hat{p}} + \frac{2[\hat{R}_1(\hat{k}^2 + m_2) - \hat{k}]}{\hat{k}^2 + m_2} \cdot \left.\frac{d[k]}{dp}\right|_{\substack{k=\hat{k} \\ p=\hat{p}}}$$

4.5

From (4.1.6) it follows

$$\left.\frac{dp}{dk}\right|_{\substack{k=\hat{k}\\ p=\hat{p}}} = \frac{\hat{R}_1^2(m_2+\hat{k}^2)^2}{2\hat{p}^2[\hat{R}_{2:1}\hat{k}(m_2+\hat{k}^2)+\hat{R}_1(m_2-\hat{k}^2)]}$$

From the Corollary of Lemma 4.2.7, $kR_2 > R$ for any k and from definition $nR_2 = R_{2:1}$ and $nR = R_1$. Thus $\hat{R}_{2:1}\hat{k} > \hat{R}_1$, and as a consequence the denominator above cannot vanish. Substituting this expression for its equivalent in $\psi'(\hat{p})$ gives

$$\psi'(\hat{p}) = \frac{1}{\hat{p}} + \frac{\hat{R}_1^2[\hat{k}^2+m_2][\hat{R}_1(\hat{k}^2+m_2)-\hat{k}]}{\hat{p}^2[\hat{R}_{2:1}\hat{k}(m_2+\hat{k}^2)+\hat{R}_1(m_2-\hat{k}^2)]} \quad . \tag{4.5.5}$$

From the definition of $\Delta(\hat{p})$ the second term of the right member of (4.5.5) becomes $\Delta(\hat{p})$. Putting this value in (4.5.4) gives

$$\frac{\hat{a}^2}{n^3}|H| = -\frac{2\hat{R}_1[\hat{R}_1(m_2+\hat{k}^2)-\hat{k}]}{m_2+\hat{k}^2} + \frac{\hat{R}_1^2}{\hat{p}} \quad ,$$

or

$$\frac{\hat{a}^2}{n^3}|H| = -\frac{(2\hat{p}-1)}{\hat{p}}\hat{R}_1^2 + \frac{2\hat{R}_1\hat{k}}{m_2+\hat{k}^2} \quad .$$

From (4.5.1) and (4.5.2)

$$\frac{2\hat{p}-1}{\hat{p}}\hat{R}_1 = \frac{2\hat{k}}{m_2+\hat{k}^2}$$

and as a consequence, since $a \neq 0$,

$$|H| = 0. \tag{4.5.6}$$

If $|H_{11}|$ represents the minor of $L_{\hat{a}\hat{a}}$, then from definition,

$$|H_{11}| = n^2 \begin{vmatrix} -\psi'(\hat{p}) & 2R_1 \\ 2\hat{R}_1 & -[(2\hat{p}-1)\hat{R}_{2:1}+2\hat{a}] \end{vmatrix}$$

4.5

Expanding the determinant and using (4.5.2) to eliminate a gives

$$|H_{11}| = n^2\left\{\frac{[\psi'(\hat{p})][2\hat{p} - 1][\hat{R}_{2:1}\hat{k} + \hat{R}_1]}{\hat{k}} - 4\hat{R}_1^2\right\} .$$

Writing $\psi'(\hat{p})$ as given by (4.5.5) and putting

$\hat{R}_3 = \hat{R}_{2:1}\hat{k} + \hat{R}_1$ gives

$$\frac{1}{n^2}|H_{11}| = \frac{(2\hat{p} - 1)\hat{R}_3}{\hat{p}\hat{k}} + \frac{2\hat{R}_1\hat{R}_3[\hat{R}_1(m_2 + \hat{k}^2) - \hat{k}]}{\hat{p}[\hat{R}_3(m_2 + \hat{k}^2) - 2\hat{k}^2\hat{R}_1]} - 4\hat{R}_1^2 .$$

Clearing this of fractions and using the relations (4.5.1) and (4.5.2) gives

$$\frac{1}{n^2}|H_{11}| = \frac{2[\hat{R}_{2:1}\hat{k} + \hat{R}_1 - 2\hat{R}_1^2\hat{k}]^2}{\hat{R}_1[\hat{R}_{2:1}\hat{k}(m_2+\hat{k}^2) + \hat{R}_1(m_2 - \hat{k}^2)]}$$

As previously explained from Lemma 4.2.7, the denominator above is positive for all k. Obviously the numerator is positive for all k. Therefore,

(4.5.7) $$|H_{11}| > 0 .$$

If $|H_{22}|$ represents the minor of $L_{\hat{p}\hat{p}}$, then from definition,

$$|H_{22}| = n^2\begin{vmatrix} \frac{\hat{p}}{\hat{a}^2} & 2k \\ 2\hat{k} & (2\hat{p} - 1)\hat{R}_{2:1} + 2\hat{a} \end{vmatrix}$$

Expanding and using the relations (4.5.1) and (4.5.2) gives

$$\frac{1}{n^2}|H_{22}| = \frac{2}{\hat{R}_1}[\hat{R}_{2:1}\hat{k}(m_2 + \hat{k}^2) + \hat{R}_1(m_2 - \hat{k}^2)] .$$

From Lemma 4.2.7, $\hat{R}_{2:1}\hat{k} > \hat{R}_1$ and, consequently,

(4.5.8) $$|H_{22}| > 0.$$

If $|H_{33}|$ represents the minor of $L_{\hat{k}\hat{k}}$, then from definition

4.5 - 6

$$|H_{33}| = \frac{n^2}{\hat{a}^2}\begin{vmatrix} -\hat{p} & 1 \\ 1 & -\psi'(\hat{p}) \end{vmatrix} = \frac{n^2}{\hat{a}^2}[\hat{p}\,\psi'(\hat{p}) - 1].$$

In $\psi'(p)$ as given by (4.5.5), from the Corollary of Lemma 4.2.6, $R_1(k^2 + m_2) > k$ for any k. Also the denominator of the second term of the right member of (4.5.5) has been shown to be positive. Thus it follows that $\hat{p}\,\psi'(\hat{p}) - 1 > 0$, since $\hat{p} > 0$. It can be shown, independently of the relationship (4.5.5), from any of the several expansions for $\psi'(p)$, (Ref. 7), that for any $p > 0$, $p\,\psi'(p) - 1 > 1$. Thus it follows,

(4.5.9) $$|H_{33}| > 0.$$

The following principal minors of order 1 are gotten from definition

(4.5.10) $$|H_{11} \cdot H_{22}| = L_{\hat{k}\hat{k}} = -n[(2\hat{p} - 1)\hat{R}_{2:1} + 2\hat{a}] < 0.$$

since from Theorem 4.3.2 $\hat{p} > \frac{1}{2}$.

(4.5.11) $$|H_{11} \cdot H_{33}| = L_{\hat{p}\hat{p}} = -n\,\psi'(\hat{p}) < 0$$ for all $\hat{p} > 0$ from the definition of $\psi'(p)$, (Ref. 7).

(4.5.12) $$|H_{22} \cdot H_{33}| = L_{\hat{a}\hat{a}} = -\frac{n\hat{p}}{\hat{a}^2} < 0.$$

Thus from the relationships (4.5.6) through (4.5.12) it is evident that the Hessian of the likelihood function vanishes and that every principal minor of it has the sign $(-1)^r$, in which r is the order of the minor, at the point $(\hat{a}, \hat{p}, \hat{k})$. It follows from the theorems relative to maxima and minima of functions of several variables (Ref. 4) that $(\hat{a}, \hat{p}, \hat{k})$ meets all the necessary conditions of a maximum point of the likelihood function. It further follows that $(\hat{a}, \hat{p}, \hat{k})$ cannot be a minimum point for the likelihood function. The theorem is thus established.

4.6. Formulas for the variance and covariance of M. L. estimates of the parameters. The conventional formulas for the variance of

4.6

maximum likelihood estimates as well as the efficiency of such estimates require either that the range of the variates be independent of the parameters, or that the distribution function and its derivatives with respect to the parameters vanish at the extremity of any range which is dependent on the parameters. (Ref. 17).

The range for w, whose probability density function is given by (4.1.1), is obviously a function of the parameters. Consequently some investigation into the behavior of the distribution function and its derivatives at the lower end of the range will be necessary.

In this section for simplicity the parameters a, p, and k will be designated by θ_1, θ_2, and θ_3 respectively. Further, unless there is a possibility of ambiguity, the arguments in $f(w, \theta_1, \theta_2, \theta_3)$ will be omitted and the quantity written f.

The reason for the conditions of the first paragraph is the necessity of the two relationships:

$$\text{a)}\quad \frac{\partial}{\partial \theta_i}\left[\int_a^b f\, dw\right] = \int_a^b \frac{\partial f}{\partial \theta_i}\, dw$$

and

$$\text{b)}\quad \int_a^b \left[\frac{\partial (\log f)}{\partial \theta_i}\right]\left[\frac{\partial (\log f)}{\partial \theta_j}\right] f\, dw = \int_a^b \frac{\partial^2 f}{\partial \theta_i \partial \theta_j}\, dw\,.$$

The usual formulas for the variance of the M. L. estimates and the efficiency of these estimates will be valid if a) and b) are true even though the range is dependent on the parameters. Thus it becomes necessary to establish the conditions under which a) and b) are true.

In the notation of this section, $f(w, a, p, k)dw$ becomes

$$(4.6.1)\qquad f(w, \theta_1, \theta_2, \theta_3)dw = \frac{2\theta_1^{\theta_2}}{\Gamma\theta_2}\,(w+\theta_3)^{2\theta_2-1}\, e^{-\theta_1(w+\theta_3)^2}\, dw\,.$$

f is a continuous function of w, θ_1, θ_2, and θ_3 over the range

4.6

$w_1 \geq -\theta_3$; $\theta_2 \geq \frac{1}{2}$ and $\theta_1 > 0$. Also $\frac{\partial f}{\partial \theta_i}$ is continuous over the same range for $i = 1, 2$. If $i = 3$, then $\theta_2 \geq 1$ is necessary for continuity at $w = -\theta_3$.

It is obvious that for $\theta_2 \geq 1/2$ both f and its derivatives, except for derivatives with respect to θ_3, will vanish at $w = -\theta_3$. The first derivative with respect to θ_3 will vanish at $w = -\theta_3$, for $\theta_2 \geq 1$.

Lemma 4.6.1. $\frac{\partial}{\partial \theta_i}\left[\int_{-\theta_3}^{\infty} f \cdot dw\right] = \int_{-\theta_3}^{\infty} \frac{\partial f}{\partial \theta_i} \cdot dw$, $i = 1, 2, 3$ for any $\theta_1 > 0$, $\theta_2 \geq 1$ and any θ_3 for:

a) all θ_i in which θ_3 is independent of θ_1 and θ_2, or

b) θ_1 and θ_2 are functionally dependent on θ_3 through either the M. L. equations or the moment equations if $n \neq 2n_1$, n_1 being the frequency of occurrence of w_1 in the set of observations $\{w_i\}$.

Proof: Since both f and its derivatives with respect to the parameters are continuous over the range of the parameters,

$$(4.6.2) \qquad \frac{\partial}{\partial \theta_i}\left[\int_{-\theta_3}^{\infty} f(w, \theta_i)\, dw\right] = \int_{-\theta_3}^{\infty} \frac{\partial f}{\partial \theta_i}\, dw + f(-\theta_3, \theta_i)\, \frac{d\theta_3}{d\theta_i},$$

in which $f(w, \theta_i) \equiv f(w, \theta_1, \theta_2, \theta_3)$. Since $f(-\theta_3, \theta_i) = 0$ for $\theta_1 > 0$, $\theta_2 \geq 1/2$, and all θ_3, the second term of the right member above will vanish if $\frac{d\theta_3}{d\theta_i}$ is finite.

For $i = 3$, $\frac{d\theta_3}{d\theta_i} = 1$ and the lemma is obviously true.

If θ_1 and θ_2 are functionally independent of θ_3, it is equally obvious that the lemma is true.

Suppose now that θ_1 and θ_2 are functionally dependent on θ_3 through the M. L. equations. Then, since $\theta_1 = a$ and $\theta_3 = k$, it follows from (4.1.7) that

4.6

$$(4.6.3) \qquad \frac{d\theta_3}{d\theta_1} = \frac{R^2}{2\theta_1^2(nR + n\theta_3 R_2 - 2\theta_3 R^2)} \quad .$$

To investigate the denominator for zeros, refer to section 4.5. The Hessian as defined there, when written in terms of θ_i and expanded becomes, for any set of $\{\theta_i\}$, $i = 1, 2, 3$.

$$|H| = \frac{n^3}{\theta_1^2}\left\{[\theta_3(R_{2:1} - 2R_1^2) + R_1][1 - \theta_2\psi'(\theta_2)] + R_1^2\theta_3[\frac{2}{2\theta_2 - 1} - \psi'(\theta_2)]\right\}$$

$$\left\{\frac{2\theta_2 - 1}{\theta_3}\right\} \quad .$$

From Reference 7

$\psi'(\theta_2) < \frac{1}{\theta_2} + \frac{1}{2\theta_2^2} + \frac{1}{6\theta_2^3}$. Further for

$2\theta_2 - 1 > 0$, $\frac{2}{2\theta_2 - 1} > \frac{1}{\theta_2} + \frac{1}{2\theta_2^2} + \frac{1}{4\theta_2^3}$. Consequently,

$\frac{2}{2\theta_2 - 1} - \psi'(\theta_2) > \frac{1}{12\theta_2^3} > 0$ and since $\theta_3 > 0$, it follows

$\theta_3 R_1^2 \left[\frac{2}{2\theta_2 - 1} - \psi'(\theta_2)\right] > 0$.

From section 4.5, $1 - \theta_2\psi'(\theta_2) < 0$ for any θ_2. Thus it is obvious that, if for any θ_3, $\theta_3(R_{2:1} - 2R_1^2) + R_1 = 0$, it will follow that $|H| > 0$. But from (4.5.6), $|H| = 0$ for any set $\{\theta_i\}$ $i = 1, 2, 3$, which satisfy the M. L. equations. Therefore it follows that for any θ_1 and θ_2 which are functionally dependent on θ_3 through the M. L. equations

$$\theta_3(R_{2:1} - 2R_1^2) + R_1 \neq 0.$$

From the definition of $R_{2:1}$ and R_1

$\theta_3(R_{2:1} - 2R_1^2) + R_1 = \frac{1}{n^2}(nR + n\theta_3 R_2 - 2\theta_3 R^2) \neq 0$. Thus

4.6

$nR + n\theta_3 R_2 - 2\theta_3 R^2$ can have no zeros among the set $\{\theta_i\}$, $i = 1, 2, 3$ which are functionally related through the M. L. equations. It follows then that the right member of (4.6.3) cannot become infinite through the vanishing of the denominator.

To investigate the behavior of $\frac{d\theta_3}{d\theta_1}$ as R becomes infinite (4.6.3) will be written

$$\frac{d\theta_3}{d\theta_1} = \frac{1}{2\theta_1^2\left(\frac{n}{R} + n\theta_3 \frac{R_2}{R^2} - 2\theta_3\right)} .$$

Since $\lim_{\theta_3 \to -w_1} R = \infty$, it follows from (4.1.7) that

$$\lim_{\theta_3 \to -w_1} \theta_1 = \frac{1}{2(m_2 + w_1^2)} .$$

From Lemma 4.2.14,

$$\lim_{\theta_3 \to -w_1} \left[\frac{n}{R} + n\theta_3 \frac{R_2}{R^2} - 2\theta_3\right] = -w_1\left(\frac{n}{n_1} - 2\right) .$$

Combining these two limits gives

$$\lim_{\theta_3 \to -w_1} \frac{d\theta_3}{d\theta_1} = \frac{2(m_2 + w_1^2)^2}{-w_1\left(\frac{n}{n_1} - 2\right)} .$$

Thus for $n \neq 2n_1$, $\frac{d\theta_3}{d\theta_1}$ will remain finite as R becomes infinite.

From Lemma 4.2.15, since $\lim_{\theta_3 \to \infty} R = 0$, it is obvious that

$$\lim_{\theta_3 \to \infty} \frac{d\theta_3}{d\theta_1} = 0 .$$

Thus for any θ_1, θ_2, and θ_3 which are functionally related through the M. L. equations, $\frac{d\theta_3}{d\theta_1}$ is finite and consequently the second term of the right member of (4.6.2) will vanish for $i = 1$.

If $i = 2$ and θ_1, θ_2, and θ_3 are functionally related through the M. L. equations, then from (4.1.3)

4.6

$$\theta_2 = \theta_1(m_2 + \theta_3^2) \qquad \text{and}$$

$$\frac{d\theta_3}{d\theta_2} = \frac{1}{(m_2 + \theta_3^2)\frac{d\theta_1}{d\theta_3} + 2\theta_1\theta_3} .$$

From the argument relative to $\frac{d\theta_3}{d\theta_1}$ it is obvious that, if $\theta_3(R_{2:1} - 2R_1^2) + R_1 < 0$, then $|H| > 0$, which is a contradiction of (4.5.6) for any set of $\{\theta_i\}$ $i = 1, 2, 3$, which are functionally related through the M. L. equations. Thus it follows, since it has already been shown that $\theta_3(R_{2:1} - 2R_1^2) + R_1 \neq 0$, that

$$\theta_3(R_{2:1} - 2R_1^2) + R_1 = \frac{1}{n^2}(nR + n\theta_3 R_2 - 2\theta_3 R^2) > 0 .$$

Consequently, $\frac{d\theta_1}{d\theta_3}$ as given by (4.6.3) cannot be negative. Thus, since each θ_1 and θ_3 is positive, the denominator of $\frac{d\theta_3}{d\theta_2}$ as given above will have no zero. Therefore, for all θ_1, θ_2, and θ_3, which are functionally related through the M. L. equations $\frac{d\theta_3}{d\theta_2}$ will remain finite and the second term of the right member of (4.6.2) will vanish in this case. Thus the lemma is established for all θ_1, θ_2, and θ_3 which are functionally related through the M. L. equations.

Now suppose the parameters are related functionally through the moment equations. Replacing population moments by sample moments in (3.1.5) gives,

$$\frac{d\theta_3}{d\theta_1} = \frac{2(m_3 + 2m_2\theta_3)}{1 - 4m_2\theta_1} . \tag{4.6.4}$$

From (3.1.4)

$$\phi_o^2 = \theta_1 m_2$$

and from Lemma 3.2.6

$$\phi_o^2 < \frac{1}{4} .$$

Thus, if θ_1, θ_2, and θ_3 are functionally related through the moment

4.6

equations the denominator of the right member of (4.6.4) will have no zero.

From Theorem 3.3.1 the moment estimates are finite for any set of observations. Thus for any m_2 and m_3 which are finite, the numerator of the right member of (4.6.4) is finite.

It follows from the above that for all θ_1, θ_2, and θ_3 which are functionally related through the moment equations, the right member of (4.6.4) remains finite and, consequently, in the case $i = 1$ the second terms of the right member of (4.6.2) will vanish.

For the case $i = 2$, again replacing population moments by sample moments in (3.1.2) gives,

$$\frac{d\theta_2}{d\theta_3} = (m_2 + \theta_3^2)\frac{d\theta_1}{d\theta_3} + 2\theta_1\theta_3 . \tag{4.6.5}$$

From (3.1.5), with population moments replaced by sample moments, since each θ_1 and θ_3 is positive,

$$m_3 + 2m_2\theta_3 = \frac{\theta_3}{2\theta_1} > 0 .$$

Since $1 - 4m_2\theta_1$ has already been shown to be positive, it follows that $\frac{d\theta_3}{d\theta_1}$ as given by (4.6.4) is positive. Consequently, each term of the right member of (4.6.5) is positive and $\frac{d\theta_2}{d\theta_3} > 0$ for all θ_1, θ_2, and θ_3 which are functionally related through the moment equations. Thus $\frac{d\theta_3}{d\theta_2}$ is finite over the same set of the parameters, and in this case the second term or the right member of (4.6.2) will vanish.

From the preceding it can be concluded that for $i = 1, 2, 3$; $f(-\theta_3, \theta_1, \theta_2, \theta_3)\frac{d\theta_3}{d\theta_i} = 0$ for all θ_1, θ_2, and θ_3 which are: 1) independent, 2) functionally dependent through the M.L. equations, or 3) functionally dependent through the moment equations. It follows then that under any of these conditions the second term of the right member of (4.6.2) will vanish and thus the lemma is established.

4.6

Lemma 4.6.2. If $f \equiv f(w, \theta_1, \theta_2, \theta_3)$ then

$$\int_{-\theta_3}^{\infty} \left[\frac{\partial(\log f)}{\partial \theta_i}\right] \left[\frac{\partial(\log f)}{\partial \theta_j}\right] f\,dw = -\int_{-\theta_3}^{\infty} \frac{\partial^2(\log f)}{\partial \theta_i \partial \theta_j} f\,dw$$

for $i, j = 1, 2, 3$; and for $\theta_1 > 0$, $\theta_2 \geq 1$, and any θ_3 if

a) θ_3 is independent of θ_2 and θ_1 or

b) θ_1, θ_2, and θ_3 are functionally related through either the M. L. equations or the moment equations unless the frequency of the smallest observation of the set $\{w_i\}$ is half the total number of observations, i.e. $n \neq 2n_1$.

Proof: From Lemma 4.6.1 and from the fact that $\int_{-\theta_3}^{\infty} f \cdot dw = 1$,

$$\frac{\partial}{\partial \theta_i}\left[\int_{-\theta_3}^{\infty} f \cdot dw\right] = \int_{-\theta_3}^{\infty} \frac{\partial f}{\partial \theta_i}\,dw = 0 \quad \text{for} \quad i = 1, 2, 3, \text{ or}$$

$$\int_{-\theta_3}^{\infty} \left[\frac{\partial(\log f)}{\partial \theta_i}\right] f\,dw = 0. \tag{4.6.6}$$

If $\theta_2 \geq 1$ and $\theta_1 > 0$,

$\frac{\partial(\log f)}{\partial \theta_i} f = \frac{\partial f}{\partial \theta_i}$ is a continuous function of all the parameters over their respective specified ranges and thus can be differentiated with respect to θ_j, $j = 1, 2, 3$. Thus, differentiating (4.6.6) with respect to θ_j, gives as a result,

$$\int_{-\theta_3}^{\infty} \left[\frac{\partial^2(\log f)}{\partial \theta_i \partial \theta_j}\right] f\,dw = -\int_{-\theta_3}^{\infty} \left[\frac{\partial(\log f)}{\partial \theta_i}\right]\left[\frac{\partial(\log f)}{\partial \theta_j}\right] f\,dw - \frac{\partial f}{\partial \theta_i}\bigg|_{w=-\theta_3} \left[\frac{d\theta_3}{d\theta_j}\right]. \tag{4.6.7}$$

For $\theta_2 \geq 1$, $\frac{\partial f}{\partial \theta_i}\Big|_{w=-\theta_3} = 0$, for $i = 1, 2, 3$. Also it can be shown as in Lemma 4.6.1 by the substitution of j for i that $\frac{d\theta_3}{d\theta_j}$ is

4.6

is finite for $j = 1, 2, 3$ under the conditions set forth in the lemma. Thus under these conditions the second term of the right member of (4.6.7) will vanish for $i, j = 1, 2, 3$. It is sufficient that $\theta_2 \geq 1/2$ except for the case $i = j = 3$ but for that case $\theta_2 \geq 1$ is necessary to ensure the vanishing of the second term.

With the vanishing of the second term of the right member of (4.6.7) it becomes

$$\int_{-\theta_3}^{\infty} \left[\frac{\partial(\log f)}{\partial \theta_i}\right]\left[\frac{\partial(\log f)}{\partial \theta_j}\right] f\,dw = -\int_{-\theta_3}^{\infty} \left[\frac{\partial^2(\log f)}{\partial \theta_i \partial \theta_j}\right] f\,dw$$

and the lemma is established.

In view of Lemmas 4.6.1 and 4.6.2 the usual formulas and theory concerning the variance of maximum likelihood estimators, and their minimum variance properties (Ref. 12, 6, 17) are valid even though the range is a function of one of the parameters.

The Hessian used here is not to be confused with that used in section 4.5. Here the elements are expected values of the second partial derivatives. In the previous section the elements are the second partial derivatives evaluated at the point represented by the M.L. estimator for the parameters.

Returning now to designating the parameters by a, p, and k, (4.1.1) gives,

$$\log f = \log 2 + p \log a - \log \sqrt{p} + (2p - 1)\log(w + k) - a(w + k)^2 .$$

Taking the expected values* of the proper second partial derivatives and defining $f_o = \frac{\sqrt{(p + 1/2)}}{\sqrt{p}}$, as in Part III, gives

*To ensure the existence of the integral in evaluating $E\left[\frac{\partial^2 \log f}{\partial k^2}\right]$ it is necessary that $p > 1$.

4.6

$$\Delta = \begin{vmatrix} \frac{p}{a^2} & \frac{2f_o}{a^{1/2}} & -\frac{1}{a} \\ \frac{2f_o}{a^{1/2}} & \frac{(4p-3)a}{p-1} & -\frac{2a^{1/2}f_o}{p-1/2} \\ -\frac{1}{a} & \frac{-2a^{1/2}f_o}{p-1/2} & \psi'(p) \end{vmatrix}$$

If $\hat{a}$, $\hat{p}$, and $\hat{k}$ represent the M. L. estimates of the parameters, then

$$\text{Var.}(\hat{a}) = \frac{\Delta_{11}}{n\Delta}; \qquad \text{Var.}(\hat{k}) = \frac{\Delta_{22}}{n\Delta}; \qquad \text{Var.}(\hat{p}) = \frac{\Delta_{33}}{n\Delta};$$

$$\text{Cov.}(\hat{a},\hat{k}) = \frac{\Delta_{12}}{n\Delta}; \qquad \text{Cov.}(\hat{a},\hat{p}) = \frac{\Delta_{13}}{n\Delta}; \qquad \text{Cov.}(\hat{p},\hat{k}) = \frac{\Delta_{23}}{n\Delta}.$$

The variances and covariances of the dependent parameters b, B, c, A, a', b', and c' will be gotten from the formulas of section 3.4, as these dependent parameters are related to the independent parameters through the same functional relationships as in the previous section. However, whenever Var.(a), Var.(p) or Var. (k) appears in one of these formulas, it is to be replaced by the appropriate function of a, p, and/or k as given by (3.4.10), (3.4.11) or (3.4.14).

PART V

Test Criterion for the Z-function

5.1 **Criterion for testing the hypothesis that a given random sample comes from a population whose probability density function is f(z)dz.** From section 3.1 the moments for z about -B are given by the relationship

$$(5.1.1) \qquad a^{s/2}\,\nu_s = \frac{\Gamma(p + \frac{s}{2})}{\Gamma(p)} .$$

For s = 2 and s = 4, (5.1.1) becomes

$$(5.1.2) \qquad a\,\nu_2 = p$$

and

$$(5.1.3) \qquad a^2\,\nu_4 = p(p + 1).$$

From the two relationships, (5.1.2) and (5.1.3), it is obvious that for any p and any a a necessary moment criterion for a population which is distributed according to the Z-function is

$$(5.1.4) \qquad p\,\nu_4 - (p + 1)\,\nu_2^2 = 0 .$$

This identical moment criterion can be obtained by the following procedure. If $\mu'_{s:x}$ represents the s^{th} moment about the origin for the Gamma type distribution, the relationship $\mu'_{s:x} = a^s\,\nu_{2s}$ is easily derived. Making this transformation on the familiar Gamma type moment criterion (Ref. 5) leads identically to (5.1.4) above.

The elimination of p from (5.1.4) by use of the relationship

$$p = \frac{\nu_1\,\nu_2}{2[\,\nu_3 - \nu_1\,\nu_2]} ,$$

obtained from ν_1, ν_2, and ν_3 as given in the early part of section 3.1 and then converting to central moments gives

5.1

$$\frac{2\mu_3 k + k(\mu_4 - 5\mu_2^2) - 2\mu_2\mu_3}{\mu_3 + 2\mu_2 k} = 0, \text{ in which } k = \mu_1' + B.$$

From (3.1.5) $k = 2a(\mu_3 + 2\mu_2 k)$. Thus if $\mu_3 + 2\mu_2 k = 0$, it would follow that $k = 0$, since $a \neq 0$. This in turn requires that $\mu_3 = 0$. Therefore for $\mu_3 \neq 0$, the denominator of the preceding expression cannot vanish, and the relationship which the central moments for z and k must obey becomes,

$$(5.1.5) \qquad 2\mu_3 k + k(\mu_4 - 5\mu_2^2) - 2\mu_3\mu_2 = 0.$$

This is not a satisfactory test as it involves k, which is the first moment for z about the unknown point -B and thus could not be calculated from the sample data.

From the moment equations of section 3.1 is obtained the relationship $k = \mu_2^{1/2} f_o \phi_o^{-1}$. Substituting this in (5.1.5) gives

$$(5.1.6) \qquad 2\alpha_3 f_o^2 + f_o\phi_o(\alpha_4 - 5) - 2\alpha_3\phi_o^2 = 0$$

as the relationship which the standard moments for z and f_o and ϕ_o must satisfy. However from (3.1.7) the standard moments for z and f_o and ϕ_o must also satisfy

$$(5.1.7) \qquad 2\alpha_3\phi_o^3 + 4f_o\phi_o^2 - f_o = 0.$$

Since both f_o and ϕ_o are functions of p, (5.1.6) and (5.1.7) constitute two equations in one variable. If the variates obey the Z-function these two equations will be consistent. It is not algebraically expedient to eliminate p between these two equations and thereby find the functional relationship between α_3 and α_4. Thus another method of testing the consistency of (5.1.6) and (5.1.7) will have to be devised.

Eliminating ϕ_o from (5.1.6) and (5.1.7) by use of the identity $\phi_o^2 \equiv p - f_o^2$ and solving the resulting equations for p and f_o, noting that from Lemma 3.2.6 $4\phi_o^2 - 1 < 0$ and thus ϕ_o and α_3 must have the

5.1-2

same signs, gives

$$(5.1.8) \qquad p = \frac{4\beta^{1/2}}{(\alpha_4 - 5 + \beta^{1/2})(\alpha_4 + 3 + \beta^{1/2})}$$

and

$$(5.1.9) \qquad f_o^2 = \frac{32\,\alpha_3^2}{(\alpha_4 - 5 + \beta^{1/2})^2(\alpha_4 + 3 + \beta^{1/2})}$$

in which $\beta = (\alpha_4 - 5)^2 + 16\alpha_3^2$.

Let p^* and f_o^* be the numerical values of p and f_o respectively obtained by substituting the sample standard moments in (5.1.8) and (5.1.9). Further, let $f_o^{**} = f_o(p^*)$ as found in the tables of Appendix I. If $f_o^{**} = f_o^*$, the equations (5.1.6) and (5.1.7) will be consistent.

In order to set up the test for the hypothesis that a given set of observations comes from a population whose probability distribution function is the Z-function consider the quantities

$$T^* = \frac{p^*}{(f_o^*)^2} \quad \text{and} \quad T^{**} = \frac{p^*}{(f_o^{**})^2}.$$

As previously stated, the hypothesis that the distribution has the Z-function as its probability distribution function is equivalent to the hypothesis that the equations (5.1.6) and (5.1.7) are consistent. This in turn is equivalent to the hypothesis that $T^* = T^{**}$, with T^{**} being the known constant $\frac{p^*}{f_o^2(p^*)}$.

Thus the test criterion will be:

$$T_c = T^{**} - T^* = 0.$$

5.2 Variance of the testing criterion. In order to investigate the sampling error or the testing criterion developed in the preceding section consider the quantity

5.2

(5.2.1) $$T^* = \frac{p^*}{f_o^{*2}} = \frac{\beta^{1/2}(\lambda + \beta^{1/2})}{8\beta_1}$$

in which $\lambda = \alpha_4 - 5$, and $\beta_1 = \alpha_3^2$. Taking the differential of both sides of (5.2.1) gives,

$$\delta_{T^*} = \frac{\beta_1[\beta^{1/2}\delta\lambda + (\frac{1}{2}\beta^{-1/2}\lambda + 1)\delta\beta] - (\beta^{1/2}\lambda + \beta)\delta\beta_1}{8\beta_1^2}$$

Since $\beta = \lambda^2 + 16\beta_1$, $\delta\beta = 2\lambda\delta\lambda + 16\delta\beta_1$. Putting this in the above gives,

$$\delta_{T^*} = \frac{\beta_1(\beta^{1/2} + \lambda)^2\delta\lambda - \lambda[\lambda(\lambda + \beta^{1/2}) + 8\beta_1]\delta\beta_1}{8\beta^{1/2}\beta_1^2}.$$

Writing $\lambda_1 = \lambda + \beta^{1/2}$ in this, squaring, and taking expected values gives

(5.2.2) $$\sigma^2_{T^*} = \frac{\beta_1^2\lambda_1^4\sigma_\lambda^2 - 2\beta_1\lambda_1^2\lambda[\lambda\lambda_1 + 8\beta_1]\operatorname{cov}(\lambda\beta_1) + \lambda^2[\lambda\lambda_1 + 8\beta_1]^2\sigma^2_{\beta_1}}{64\beta\beta_1^4}.$$

Converting each λ and β_1 from functions of standard moments to functions of central moments and taking differentials, gives

$$\delta\lambda\delta\beta_1 = \frac{2\mu_3\delta\mu_3\delta\mu_4}{\mu_2^5} - \frac{4\mu_3\mu_4\delta\mu_2\delta\mu_3}{\mu_2^6} - \frac{3\mu_3^2\delta\mu_2\delta\mu_4}{\mu_2^6} + \frac{6\mu_3^2\mu_4\delta^2\mu_2}{\mu_2^7}.$$

Taking expected value of each side of this and using the familiar formulas (Ref. 11) for the variance and covariance of sample moments about the mean gives,

(5.2.3) $$\operatorname{Cov.}(\lambda, \beta_1) = \frac{\alpha_3}{n}\,2[\alpha_7 - \alpha_5(2\alpha_4 + 3)] + 3\alpha_3[\alpha_4 + 2\alpha_4^2 + 8 - \alpha_6 + 4\alpha_3^2].$$

The following familiar formulas for σ_λ^2 and $\sigma^2_{\beta_1}$ are given in

5.2

reference 3,

$$(5.2.4) \qquad \sigma^2_{\beta_1} = \frac{\alpha_3^2}{n}\left[4(\alpha_6 - 6\alpha_4 + 9) + \alpha_3(35\alpha_3 + 9\alpha_3\alpha_4 - 12\alpha_5)\right].$$

$$(5.2.5) \qquad \sigma^2_{\lambda} = \frac{1}{n}\left[\alpha_8 + \alpha_4(4\alpha_4^2 - \alpha_4 - 4\alpha_6) + 8\alpha_3(2\alpha_3 + 2\alpha_3\alpha_4 - \alpha_5)\right].$$

Thus $\sigma^2_{T^*}$ can easily be computed from (5.2.2) with the aid of (5.2.3), (5.2.4), and (5.2.5) as the right member in each case is in terms of the standard moments.

In the investigation of a large sample relative to its distribution being of the Z-function type, one would compare the value of T_c as calculated by the previous section with the value of σ_{T^*} as obtained from (5.2.2).

PART VI

Numerical Illustrations

6.1 By method of moments.

Example 1. The following data relative to lost articles which were found and turned in came from reference 25. I am indebted to the Bell System Technical Journal for permission to use it.

Table 6.1.1. Lost Articles

Lost Articles	0	1	2	3	4	5	6	7
Frequency	169	134	74	32	11	2	0	1

If the "Lost Articles" are represented by z, and m_s and $a_{s:m}$ represent the s^{th} central and standard sample moment respectively, then: the sample number, n = 423 and

$$\bar{z} = 1.037825 \quad \text{and } m_2 = 1.270437; \quad m_2^{1/2} = 1.127137$$
$$a_{3:m} = 1.223336, \quad a_{4:m} = 4.888157, \quad a_{5:m} = 16.612754$$
$$a_{6:m} = 71.781417, \quad a_{7:m} = 331.3167973, \quad a_{8:m} = 1625.223052.$$

First the test criterion of section 5.1 will be applied. From formulas (5.1.8) and (5.1.9) respectively, $p^* = .32024$, and $(f^*)^2 = .16377$, which gives $T^* = 1.95532$. Further from the table of Appendix I, $(f^{**})^2 = .166998$ and $T^{**} = 1.91762$. Thus $T^{**} - T^* = .0377$.

Applying the formula (5.2.2) gives $\sigma_{T^*} = .48894$. Thus $\frac{T^{**} - T^*}{\sigma_{T^*}} = .077106$ and the hypothesis that the sample comes from a population with the Z-function for its probability distribution function is strongly supported.

The equation (3.1.8) and the procedure outlined in section 3.1 yields

6.1

for $p = .363$, $\alpha_3 = 1.223223$. Thus .363 will be taken as the moment estimate for p. Then (3.1.4) and (3.1.2) give for $p = .363$, $a = .127149$ and $k = 1.258763$ respectively. From definition $B = k - \bar{z} = .220938$.

The use of the relationships from sections 1.2 and 3.4 yield as values for the moment estimates of the parameters and their respective standard deviations

Table 6.1.2. Parameters for the Z-function.

Parameter	Value	Standard deviation
p	.363	.032632
a	.127149	.009254
k	1.258763	.016696
A	.602495	.027081
B	.220938	.057290
b	.056184	.018711
c	.006207	.003355
a'	.211037	.019494
b'	.093252	.020450
c'	.592194	.025172

The lower limit of the range for z is thus -.220938. If the observed values are regarded as the midpoints of the "class" this lower limit is consistent with the observed data. The transformation (1.1.1) becomes

(6.1.1) $$t = .211037\, z^2 + .093252\, z - .592194.$$

In the graduation of this data it is necessary to use a table of incomplete Gamma functions (Ref. 22). The argument u of that table corresponds to xA^{-1} in the notation of this paper. Thus from (1.2.4) the

6.1

transformation becomes

$$(6.1.2) \qquad u = \frac{x}{\sqrt{p}} = .211037\,(z + .220938)^2 .$$

Regarding each observed value of z as the midpoint of the class, graduated frequencies as shown in table 6.1.3 are obtained.

Table 6.1.3. Graduated frequencies by Z-function.

z	Graduated Frequency	Observed Frequency
0	173.297	169
1	129.898	134
2	70.968	74
3	32.481	32
4	11.925	11
5	3.469	2
6	.796	0
7	.166	1

Grouping the last three classes, the resulting χ^2_2 from the table above is found to be .9065. Since the degrees of freedom are determined by 6 - 4 = 2, it is found $P(\chi^2_2 > .9065)$ is between .50 and .70, with linear interpolation giving .64.

For the purpose of comparison, the results of graduating the data of table 6.1.1 by the Gamma type function are given here. The lower end of the range is found to be -.804902 which is not consistent with the observed data. Regarding the observed values as the midpoint of the respective classes, the graduated frequencies are shown in table 6.1.4.

6.1

Table 6.1.4. Graduated frequencies by the Gamma type function.

z	Graduated Frequency	Observed Frequency
-1	8.984	0
0	148.917	169
1	146.362	134
2	74.184	74
3	29.463	32
4	10.307	11
5	3.334	2
6	1.046	0
7	.403	1

Grouping together the last three classes, the resulting χ^2_3 is found to be 13.6667. For three degrees of freedom it is found $P(\chi^2_3 > 13.6667) < .01$.

Example 2. For the data of this example the "fit", as indicated by the test criterion as well as the graduated frequencies, for both the Z-function and the Gamma type function is so good as to lead one to suspect the randomness of the sample. It is presented here to bring out two points: The improvement in the relative magnitude of the standard deviations of the parameters when estimated by the Z-function; and the improvement in the already small χ^2 of the Gamma type function which is brought about by use of the Z-function.

The distribution of 1000 scores made by students on a placement test is as follows:

6.1

Table 6.1.5 Distribution of scores.

Score	19	20	21	22	23	24	25	26	27	28	29
Frequency	1	11	53	140	228	244	180	94	36	10	3

Representing the scores by z, then in the notation of example 1, $n = 1000$, $\bar{z} = 23.815$, $m_2 = 2.546775$, $m_2^{1/2} = 1.595862$, $a_{3:m} = .179848$, $a_{4:m} = 2.968457$, $a_{5:m} = 1.670967$, $a_{6:m} = 14.416708$, $a_{7:m} = 15.058569$, $a_{8:m} = 93.818537$.

Applying the test criterion of section 5.1 gives from formulas 5.1.8 and 5.1.9 respectively: $p^* = 8.5849$ and $(f^*)^2 = 8.3387$. From these according to (5.2.1), $T^* = 1.029524$. From a table of Gamma functions, $(f^{**})^2 = 8.3388$, which gives $T^{**} = 1.029513$. From formula (5.2.2) $\sigma_{T^*} = .217910$. Thus $\frac{T^{**} - T^*}{\sigma_{T^*}} = .00005$ and as mentioned before the "fit" is almost too good.

To approximate the root for (3.1.7), if F(p) designates its left member, then $F(8.31) = -.000011$ and $F(8.32) = .000055$. Thus 8.311, obtained by interpolation, will be taken as the moment estimate for p. The use of the relationships as indicated in example 1 yields values of the moment estimates for the parameters and their standard deviations as given in the table 6.1.6 below.

It will be noticed that the lower limit of the range, -B, is lower than that indicated by the observations. However, no graduated frequencies will be found in the classes below the observed lower limit in the graduation of a set of 1000 observations as the probability of such an occurrence is less than .00001.

6.1

Table 6.1.6. Parameters for the Z-function.

Parameter	Value	Standard Deviation
p	8.311	.37601
a	.0966	.00450
k	9.1350	.05049
B	-14.6800	.07138
b	-2.8375	.13265
c	20.8275	.99171
A	2.8829	.06521
a'	.0335	.000843
b'	-.9843	.06734
c'	4.3417	.15477

The transformation (1.1.1) becomes $t = .0335\, z^2 - .9843z + 4.3417$. The skewness of the Gamma type function into which the Z-function is transformed is $\frac{2}{A} = .6937$.

Regarding each observed value of z as the midpoint of its class, graduated frequencies as shown in table 6.1.7 are obtained.

The χ^2 calculated from the following table is .5902. The degrees of freedom are given by 11 - 4 = 7 and for seven degrees of freedom it is found that $P(\chi^2_7 > .5902) > .99$ again indicating a "fit" almost too good.

6.1

Table 6.1.7. Graduated frequencies by Z-function.

z	Graduated Frequency	Observed Frequency
19	1.47	1
20	12.34	11
21	54.80	53
22	140.11	140
23	224.89	228
24	240.63	244
25	179.03	180
26	95.51	94
27	37.38	36
28	10.94	10
29	2.90	3

For the purpose of comparison, the results of graduating the data of table 6.1.5 by the Gamma type function are given.

The criterion for the Gamma function (Ref. 3 and 5) gives $\delta = -.01599$ with the corresponding $b_2 = -.0082587$ for which $\sigma_{b_2} = .02155$. The notations used are that of the two references cited. Since the deviation in b_2 is less than four-tenths of its standard deviation the indicated "fit" is highly significant.

For the Gamma type function as written in section 3.5 moment estimates for the parameters are as follows: $p_3 = 123.7178$, $a_3 = .1435$ and $B_3 = 6.0645$.

Regarding each observed value as the midpoint of its class, graduated frequencies as shown in table 6.1.8 are obtained. The corresponding χ^2 is .6668. The degrees of freedom being seven, it is found that

6.1

$P(\chi^2_7 > .6668) > .99.$

It would appear from the above that the Gamma type function is as well adapted to this example as the Z-function. However, an important difference becomes apparent through consideration of the first point mentioned in the presentation of the example.

In the notation of section 3.5, the following values are obtained for the parameter estimates by the Gamma type function: $r_{p_3} = .77855$, $r_{a_3} = .002024$, and $r_{k_3} = .38827$.

The corresponding quantities for the Z-function are: $r_{p_1} = .00195$, $r_{a_1} = .00217$, and $r_{k_3} = .0003056$. Thus the ratios become: $R_{p_3:p_1} = 401.314$, $R_{a_3:a_1} = .93255$, and finally $R_{k_3:k_1} = 1270.517$.

Table 6.1.8. Graduated frequencies by the Gamma type function.

z	Graduated Frequency	Observed Frequency
19	1.71	1
20	12.47	11
21	54.08	53
22	139.23	140
23	225.61	228
24	241.88	244
25	179.08	180
26	94.86	94
27	37.03	36
28	10.98	10
29	3.07	3

6.2

6.2 By method of maximum likelihood. Consider again the distribution of scores given in Table 6.1.5.*

In the notation of section 4,1, taking values of k as listed below gives the following table of values.

Table 6.2.1 Values of U(k)

k	p_k	U(k)
8.801	7.720546	.000000043
8.8016	7.721596	.000000004
8.802	7.722295	-.000000024

Designating the maximum likelihood estimates by $\hat{a}$, $\hat{k}$, etc. the above gives as correct to four decimal places, $\hat{k} = 8.8016$, $\hat{p} = 7.7216$ and from (4.1.7) $\hat{a} = .0965$. Then $\hat{B} = \hat{k} - \bar{z} = -15.0134$.

The standard deviations for $\hat{a}$, $\hat{p}$, and $\hat{k}$ are found from the formulas of section 4.6. The values of the dependent parameters are found from the formulas of section 1.2. The standard deviations for these dependent parameters are found from formulas (3.4.17) through (3.4.23) in which σ_p^2, σ_a^2, σ_k^2, are replaced by the functional relationships (3.4.10), (3.4.11), and (3.4.14), respectively.

Table 6.2.2 gives a summary of the values of the parameters and their respective standard deviations as obtained from the formulas enumerated above.

---

* As previously noted this data "fits" almost too well. However, it is selected to illustrate the tecnique of the M.L. method due to its also showing that small though the χ^2 obtained previously by each the Z-function and the Gamma type function may be, it is even smaller here.

6.2

Table 6.2.2. Parameters for the Z-function

Parameter	Value	Standard Deviation
$\hat{p}$	7.7216	2.7727
$\hat{k}$	8.8016	1.6440
$\hat{a}$	.0965	.0044
$\hat{B}$	-15.0134	.0735
$\hat{b}$	-2.8976	.1363
$\hat{c}$	21.7517	1.0221
$\hat{A}$	2.7788	.0630
$\hat{a}'$	.0347	.00088
$\hat{b}'$	-1.0428	.0409
$\hat{c}'$	5.0490	.1597

It will be noticed that the standard deviations for $\hat{k}$ and $\hat{p}$ are larger than in the case of the moment estimates by the Z-function. However, if $\hat{k}$ and $\hat{p}$ are regarded as independently estimated, the respective standard deviations become:

$\sigma_{\hat{k}} = .04997$, $\sigma_{\hat{a}} = .00109$ and $\sigma_{\hat{p}} = .08504$ from the method of reference 12.

The transformation (1.1.1) becomes

$t = .0347\, z^2 - 1.0428\, z + 5.0490$. The skewness of the Gamma type function into which the Z-function is transformed is $\frac{2}{\hat{A}} = .7197$.

Regarding the observed values of z as the midpoints of their respective classes, graduated frequencies as shown in Table 6.2.3 are obtained.

6.2

Table 6.2.3. Graduated Frequencies by the Z-function

z	Graduated Frequencies	Observed Frequencies
19	1.439	1
20	12.142	11
21	54.770	53
22	140.517	140
23	225.170	228
24	240.421	244
25	178.684	180
26	95.411	94
27	37.469	36
28	11.023	10
29	2.465	3
30	.489	0

Grouping together the last two classes, i.e. $z = 29$ and $z = 30$, the resulting χ^2 is found to be .5727. The degrees of freedom are given by $11 - 4 = 7$ for which it is found $P(\chi^2_7 > .5727) > .99$. As in the case of this problem when graduated by the method of moments the "fit" is almost good enough to lead one to suspect the randomness of the sample.

APPENDIX I

Tables of Values of f_o and ϕ_o^2.

The procedure employed in the construction of the table is as follows:

From definition, $f_o(p) = \frac{\Gamma(p + 1/2)}{\Gamma p}$. Thus for any value of p such that $1 < p < 2$, $f_o(p)$ can be evaluated by division of the two Gamma functions involved. These Gamma functions can be found in any of the several tables available, e.g. (Ref. 7). Knowing the value of $f_o(p)$ for any given value of p, values of $f_o(p \pm 1)$ are readily obtained by use of the formula

$$f_o(p + 1) = \frac{p + 1/2}{p} f_o(p) .$$

Values of $\phi_o^2(p)$ are obtained from the definition of the function, i.e.

$$\phi_o^2 = p - f_o^2 .$$

Table 1. $f_o(p)$, $f_o^2(p)$, and $\phi_o^2(p)$ for values of p from zero through one at intervals of two-hundredths.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
.00	.00000 00000	.00000 00000	.00000 00000	.00
.02	.03450 17710	.00119 03722	.01881 96278	.02
.04	.06724 07699	.00452 13211	.03547 86789	.04
.06	.09839 43736	.00968 14528	.05031 85472	.06
.08	.12811 41800	.01641 32431	.06358 67569	.08
.10	.15653 45082	.02450 30523	.07549 69477	.10
.12	.18377 10726	.03377 18071	.08622 81929	.12
.14	.20992 58559	.04406 88650	.09593 11350	.14
.16	.23508 87834	.05526 67361	.10473 32639	.16
.18	.25933 94856	.06725 69688	.11274 30312	.18
.20	.28274 87525	.07994 68570	.12005 31430	.20
.22	.30537 97494	.09325 67913	.12674 32087	.22
.24	.32728 90358	.10711 81130	.13288 18870	.24
.26	.34852 74246	.12147 13657	.13852 86343	.26
.28	.36914 06032	.13626 47849	.14373 52151	.28
.30	.38916 93716	.15145 27998	.14854 72002	.30
.32	.40865 36117	.16699 77744	.15300 22256	.32
.34	.42762 49092	.18286 30630	.15713 69370	.34
.36	.44611 51521	.19901 87289	.16098 12711	.36
.38	.46415 27488	.21543 77742	.16456 22258	.38
.40	.48176 36789	.23209 62423	.16790 37577	.40
.42	.49897 17871	.24897 28443	.17102 71557	.42
.44	.51579 89965	.26604 86048	.17395 13952	.44
.46	.53226 55016	.28330 65642	.17669 34358	.46
.48	.54838 99444	.30073 15311	.17926 84689	.48

Table 1. Continued.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
.50	.56418 95836	.31830 98862	.18169 01138	.50
.52	.57968 03875	.33602 93516	.18397 06484	.52
.54	.59487 73272	.35387 90344	.18612 09656	.54
.56	.60979 37490	.37184 84163	.18815 15837	.56
.58	.62444 15401	.38992 72370	.19007 27630	.58
.60	.63883 67725	.40811 24219	.19188 75781	.60
.62	.65298 63380	.42639 11576	.19360 88424	.62
.64	.66690 21318	.44475 84534	.19524 15466	.64
.66	.68059 39345	.46320 81037	.19679 18963	.66
.68	.69407 09381	.48173 44670	.19826 55330	.68
.70	.70734 17589	.50033 23639	.19966 76361	.70
.72	.72041 45026	.51899 70556	.20100 29444	.72
.74	.73329 67921	.53772 41853	.20227 58147	.74
.76	.74599 58167	.55650 97585	.20349 02415	.76
.78	.75851 83539	.57535 00932	.20464 99068	.78
.80	.77087 08047	.59424 17975	.20575 82025	.80
.82	.78305 92196	.61318 17414	.20681 82586	.82
.84	.79508 93235	.63216 70323	.20783 29677	.84
.86	.80696 65383	.65119 49939	.20880 50061	.86
.88	.81869 60045	.67026 31478	.20973 68522	.88
.90	.83028 26003	.68936 91964	.21063 08036	.90
.92	.84173 09571	.70851 10041	.21148 89959	.92
.94	.85304 54785	.72768 65884	.21231 34116	.94
.96	.86423 03531	.74689 41032	.21310 58968	.96
.98	.87528 95742	.76613 18387	.21386 81613	.98
1.00	.88622 69254	.78539 81633	.21470 18367	1.00

Table 2. $f_o(p)$, $f_o^2(p)$, and $\phi_o^2(p)$ for values of p from zero through ten at intervals of two-tenths.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
.0	.0000 00000	.0000 00000	.0000 00000	.0
.2	.2827 48753	.0799 46857	.1200 53143	.2
.4	.4817 63679	.2320 96242	.1679 03758	.4
.6	.6388 36772	.4081 12421	.1918 87579	.6
.8	.7708 70805	.5942 41798	.2057 58203	.8
1.0	.8862 26926	.7853 98163	.2147 01837	1.0
1.2	.9896 20634	.9793 48999	.2206 51001	1.2
1.4	1.0839 68277	1.1749 87226	.2250 12774	1.4
1.6	1.1712 00749	1.3717 11194	.2282 88806	1.6
1.8	1.2526 65058	1.5691 69748	.2308 30252	1.8
2.0	1.3293 40387	1.7671 45865	.2328 54135	2.0
2.2	1.4019 62565	1.9654 99034	.2345 00966	2.2
2.4	1.4710 99805	2.1641 34636	.2358 65364	2.4
2.6	1.5372 00983	2.3629 86862	.2370 13138	2.6
2.8	1.6006 27574	2.5620 08631	.2379 91369	2.8
3.0	1.6616 75484	2.7611 65414	.2388 34586	3.0
3.2	1.7205 90421	2.9604 31397	.2395 68603	3.2
3.4	1.7775 78931	3.1597 86856	.2402 13144	3.4
3.6	1.8328 16557	3.3592 16532	.2407 83468	3.6
3.8	1.8864 53927	3.5587 08419	.2412 91581	3.8
4.0	1.9386 21398	3.7582 52925	.2417 47075	4.0
4.2	1.9894 32674	3.9578 42364	.2421 57636	4.2
4.4	2.0389 87597	4.1574 70421	.2425 29579	4.4
4.6	2.0873 74412	4.3571 31936	.2428 68064	4.6
4.8	2.1346 71549	4.5568 22622	.2431 77378	4.8

Table 2. Continued.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
5.0	2.1809 49073	4.7565 38859	.2434 61141	5.0
5.2	2.2262 69897	4.9562 77654	.2437 22346	5.2
5.4	2.2706 90733	5.1560 36405	.2439 63595	5.4
5.6	2.3142 62935	5.3558 12932	.2441 87068	5.6
5.8	2.3570 33169	5.5556 05360	.2443 94640	5.8
6.0	2.3990 43980	5.7554 12018	.2445 87982	6.0
6.2	2.4403 34310	5.9552 31545	.2447 68455	6.2
6.4	2.4809 39875	6.1550 62663	.2449 37337	6.4
6.6	2.5208 93554	6.3549 04311	.2450 95689	6.6
6.8	2.5602 25684	6.5547 55553	.2452 44447	6.8
7.0	2.5989 64312	6.7546 15495	.2453 84505	7.0
7.2	2.6371 35464	6.9544 83455	.2455 16545	7.2
7.4	2.6747 63303	7.1543 58727	.2456 41273	7.4
7.6	2.7118 70338	7.3542 40730	.2457 59270	7.6
7.8	2.7484 77573	7.5541 28969	.2458 71031	7.8
8.0	2.7846 04620	7.7540 22890	.2459 77110	8.0
8.2	2.8202 69871	7.9539 22145	.2460 77855	8.2
8.4	2.8554 90553	8.1538 26298	.2461 73702	8.4
8.6	2.8902 82860	8.3537 35011	.2462 64989	8.6
8.8	2.9246 62033	8.5536 48007	.2463 51993	8.8
9.0	2.9586 42409	8.7535 64904	.2464 35096	9.0
9.2	2.9922 37546	8.9534 85532	.2465 14468	9.2
9.4	3.0254 60229	9.1534 09597	.2465 90403	9.4
9.6	3.0583 22561	9.3533 36887	.2466 63113	9.6
9.8	3.0908 36012	9.5532 67253	.2467 32747	9.8
10.0	3.1230 11432	9.7532 00404	.2467 99596	10.0

Table 3. $f_o(p)$, $f_o^2(p)$, and $\phi_o^2(p)$ for values of p from zero through fifty at intervals of one-half.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
.0	.0000 00000	.000 00000	.0000 00000	.0
.5	.5641 89584	.318 30989	.1816 90114	.5
1.0	.8862 26926	.785 39816	.2147 01837	1.0
1.5	1.1283 79168	1.273 23955	.2267 60453	1.5
2.0	1.3293 40387	1.767 14587	.2328 54135	2.0
2.5	1.5045 05557	2.263 53697	.2364 63029	2.5
3.0	1.6616 75484	2.761 16541	.2388 34586	3.0
3.5	1.8054 06668	3.259 49324	.2405 06763	3.5
4.0	1.9386 21398	3.758 25293	.2417 47075	4.0
4.5	2.0633 21906	4.257 29729	.2427 02712	4.5
5.0	2.1809 49073	4.756 53886	.2434 61141	5.0
5.5	2.2925 79895	5.255 92258	.2440 77420	5.5
6.0	2.3990 43980	5.755 41202	.2445 87982	6.0
6.5	2.5009 96249	6.254 98224	.2450 17757	6.5
7.0	2.5989 64312	6.754 61550	.2453 84505	7.0
7.5	2.6933 80576	7.254 29893	.2457 01067	7.5
8.0	2.7846 04620	7.754 02289	.2459 77110	8.0
8.5	2.8729 39281	8.253 78011	.2462 19882	8.5
9.0	2.9586 42409	8.753 56490	.2464 35096	9.0
9.5	3.04 19 35709	9.253 37286	.2466 27136	9.5
10.0	3.1230 11432	9.753 20040	.2467 99596	10.0
10.5	3.2020 37588	10.253 04472	.2469 55279	10.5
11.0	3.2791 62004	10.752 90345	.2470 96552	11.0
11.5	3.3545 15568	11.252 77470	.2472 25297	11.5
12.0	3.4282 14822	11.752 65687	.2473 43134	12.0

Table 3. Continued.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
12.5	3.5003 64072	12.252 54863	.2474 51370	12.5
13.0	3.5710 57106	12.752 44885	.2475 51146	13.0
13.5	3.6403 78635	13.252 35661	.2476 43394	13.5
14.0	3.7084 05456	13.752 27103	.2477 28974	14.0
14.5	3.7752 07473	14.252 19146	.2478 08536	14.5
15.0	3.8408 48508	14.752 11726	.2478 82739	15.0
15.5	3.9053 87041	15.252 04794	.2479 52060	15.5
16.0	3.9688 76792	15.751 98299	.2480 17010	16.0
16.5	4.0313 67268	16.251 92205	.2480 77950	16.5
17.0	4.0929 04192	16.751 86473	.2481 35275	17.0
17.5	4.1535 29912	17.251 81073	.2481 89270	17.5
18.0	4.2132 83727	17.751 75976	.2482 40236	18.0
18.5	4.2722 02195	18.251 71160	.2482 88405	18.5
19.0	4.3303 19386	18.751 66599	.2483 34015	19.0
19.5	4.3876 67119	19.251 62275	.2483 77253	19.5
20.0	4.4442 75159	19.751 58169	.2484 18311	20.0
20.5	4.5001 71404	20.251 54267	.2484 57335	20.5
21.0	4.5553 82038	20.751 50551	.2484 94488	21.0
21.5	4.6099 31682	21.251 47011	.2485 29887	21.5
22.0	4.6638 43515	21.751 43633	.2485 63668	22.0
22.5	4.7171 39396	22.251 40408	.2485 95919	22.5
23.0	4.7698 39959	22.751 37323	.2486 26766	23.0
23.5	4.8219 64716	23.251 34372	.2486 56278	23.5
24.0	4.8735 32132	23.751 31544	.2486 84558	24.0
24.5	4.9245 59710	24.251 28834	.2487 11663	24.5
25.0	4.9750 64051	24.751 26231	.2487 37688	25.0

Table 3. Continued.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
25.5	5.0250 60929	25.251 23734	.2487 62660	25.5
26.0	5.0745 65332	25.751 21331	.2487 86691	26.0
26.5	5.1235 91535	26.251 19022	.2488 09782	26.5
27.0	5.1721 53127	26.751 16797	.2488 32031	27.0
27.5	5.2202 63073	27.251 14655	.2488 53449	27.5
28.0	5.2679 33740	27.751 12589	.2488 74111	28.0
28.5	5.3151 76947	28.251 10598	.2488 94022	28.5
29.0	5.3620 03985	28.751 08674	.2489 13265	29.0
29.5	5.4084 25665	29.251 06817	.2489 31826	29.5
30.0	5.4544 52330	29.751 05022	.2489 49780	30.0
30.5	5.5000 93897	30.251 03288	.2489 67124	30.5
31.0	5.5453 59869	30.751 01608	.2489 83923	31.0
31.5	5.5902 59371	31.250 99984	.2490 00165	31.5
32.0	5.6348 01157	31.750 98408	.2490 15921	32.0
32.5	5.6789 93647	32.250 96884	.2490 31157	32.5
33.0	5.7228 44925	32.750 95404	.2490 45964	33.0
33.5	5.7663 62780	33.250 93971	.2490 60289	33.5
34.0	5.8095 54697	33.750 92578	.2490 74223	34.0
34.5	5.8524 27896	34.250 91228	.2490 87722	34.5
35.0	5.8949 89325	34.750 89914	.2491 00858	35.0
35.5	5.9372 45692	35.250 88641	.2491 13593	35.5
36.0	5.9792 03458	35.750 87399	.2491 26008	36.0
36.5	6.0208 68871	36.250 86196	.2491 38038	36.5
37.0	6.0622 47950	36.750 85021	.2491 49793	37.0
37.5	6.1033 46527	37.250 83883	.2491 61171	37.5
38.0	6.1441 70220	37.750 82769	.2491 72308	38.0

Table 3. Continued.

p	$f_o(p)$	$f_o^2(p)$	$\phi_o^2(p)$	p
38.5	6.1847 24481	38.250 81691	.2491 83094	38.5
39.0	6.2250 14565	38.750 80633	.2491 93666	39.0
39.5	6.2650 45578	39.250 79609	.2492 03906	39.5
40.0	6.3048 22444	39.750 78605	.2492 13940	40.0
40.5	6.3443 49952	40.250 77631	.2492 23687	40.5
41.0	6.3836 32724	40.750 76675	.2492 33251	41.0
41.5	6.4226 75260	41.250 75750	.2492 42515	41.5
42.0	6.4614 81903	41.750 74838	.2492 51617	42.0
42.5	6.5000 56889	42.250 73956	.2492 60440	42.5
43.0	6.5384 04306	42.750 73087	.2492 69131	43.0
43.5	6.5765 28146	43.250 72246	.2492 77545	43.5
44.0	6.6144 32263	43.750 71416	.2492 85838	44.0
44.5	6.6521 20423	44.250 70612	.2492 93878	44.5
45.0	6.6895 96265	44.750 69819	.2493 01811	45.0
45.5	6.7268 63349	45.250 69052	.2493 09484	45.5
46.0	6.7639 25112	45.750 68292	.2493 17079	46.0
46.5	6.8007 84924	46.250 67558	.2493 24417	46.5
47.0	6.8374 46037	46.750 66831	.2493 31691	47.0
47.5	6.8739 11643	47.250 66128	.2493 38724	47.5
48.0	6.9101 84824	47.750 65430	.2493 45698	48.0
48.5	6.9462 68607	48.250 64756	.2493 52439	48.5
49.0	6.9821 65915	48.750 64087	.2493 59135	49.0
49.5	7.0178 79623	49.250 63440	.2493 65597	49.5
50.0	7.0534 12505	49.750 62797	.2493 72034	50.0

APPENDIX II

Technique to be employed in a graduation problem by method of moments in case of negative skewness.

At the close of section 3. 1 attention was called to the fact that a slight modification in technique will be necessary for a negative skewness. Details of this modification are made evident by analysis of the moment equations and their solution. As pointed out in section 3. 3, the only difference in finding the roots for the moment equations in the cases $\alpha_3 > 0$ and $\alpha_3 < 0$ is in the sign of ϕ_o. ϕ_o must have the same sign as α_3.

Since each p and a is a function of ϕ_o^2, their values will be independent of the sign of ϕ_o and α_3. The difference comes in the determination of k and B.

There are two lines of procedure which may be employed in the case of $\alpha_3 < 0$. Let a_1, k_1, etc. designate the parameters corresponding to $\alpha_3 < 0$, while a, k, etc. represent the equivalent parameters corresponding to an α_3 of the same absolute value but of opposite sign.

The first line of procedure is: If $\alpha_3 < 0$ and $\phi_o < 0$, then $a_1 = \frac{(-\phi_o)^2}{\mu_2}$ and $(a_1)^{1/2} = -\frac{\phi_o}{\mu_2^{1/2}}$. Then $k_1 = \frac{f_o}{(a_1)^{1/2}}$ would be negative, and $-B_1 = -k_1 + \mu_1'$ would be the upper end of the range for z.

Further, the following relationships will hold. $a_1 = a$, $k_1 = -k$, $B_1 = k_1 - \mu_1' = -k - \mu_1'$, or $B_1 = B - 2k$. Then $b_1 = 2a_1B_1 = 2a(B - 2k) = b - 4ak$. $c_1 = \frac{b^2}{4a_1} = \frac{b^2 - 8abk + 16k^2}{4a}$ $p_1 = p$ and $A_1 = A$. The dependent parameters would then have the following functional relationship. $a_1' = a'$, $b_1' = b' - 4a'k$, and, finally,

$c'_1 = c' - 2a'k + 4b'k^2$.

The transformation (1. 1. 1a) becomes, if t_1 represents the value of t for $\alpha_3 < 0$,

$$t_1 = a'(z - 2k)^2 + b'(z - 2k) + c'.$$

Thus, according to this line of procedure, the values of the parameters are found for α_3 having the same absolute value but with a positive sign. In finding the value of t corresponding to any given z, the z variate will be decreased by 2k. Further, since k is the upper limit of the range if measured from the mean, the area corresponding to any t_1 will represent the percentage of area above the corresponding value of z.

The second line of procedure is: Starting from the relationship $B_1 = -(k + \mu'_1)$ from the first line of procedure, the other relationships may be written: $b_1 = -2a(k + \mu'_1)$ and $c_1 = a(k + \mu'_1)^2$ Which gives for the dependent parameters, $a'_1 = a'$, $b'_1 = -\frac{2a(k + \mu'_1)}{A}$, and $c'_1 = \frac{a(k + \mu'_1)^2 - A^2}{A}$ and the transformation (1. 1. 1a) now becomes

$$t_1 = a'z^2 - \frac{2a(k + \mu'_1)z}{A} + \frac{a(k + \mu'_1)^2 - A^2}{A}.$$

However, for $\alpha_3 > 0$ the transformation (1. 1. 1a) can be written $t = a'z^2 + \frac{2a(k - \mu'_1)z}{A} + \frac{a(k - \mu'_1)^2 - A^2}{A}$. Thus it is obvious that $t_1(-z, k + \mu'_1) \equiv t(z, k - \mu'_1)$. That is, to find the value of t_1 corresponding to any z, in case $\alpha_3 < 0$, the parameters will be found just as for $\alpha_3 > 0$, then use $k + \mu'_1$ as the parameter of location in place of $k - \mu'_1$, and change the sign of z. Again the area found in the table corresponding to any t_1 will be the percentage of area lying above that value of z which corresponds to the t_1.

LIST OF REFERENCES

1. Baker, G. A., Transformation of Non-normal Frequency Distributions into Normal Distributions. Annals of Mathematical Statistics, Vol. 5 (1934), p. 113.

2. Bôcher, M., Introduction to Higher Algebra, MacMillan Co., New York, 1915.

3. Carver, H. C., Frequency Curves. Handbook of Mathematical Statistics, H. L. Rietz, Editor-inChief, Houghton Mifflin Company, New York, 1924.

4. Chaundy, T., The Differential Calculus, Oxford University Press, London, 1935.

5. Craig, C. C., A New Exposition and Chart for the Pearson System of Frequency Curves. Annals of Mathematical Statistics, Vol. 7, No. 1, March 1936, p. 16.

6. Cramér, H., Mathematical Methods of Statistics, Princeton University Press, 1946.

7. Davis, E. T., Tables of Higher Mathematical Functions, Vol. 1. Principia Press, Inc., Bloomington, Ind., 1933.

8. Edgeworth, F. Y., Representation of Statistics by Mathematical Formula. Journal of Royal Statistical Society, Vol. 61, (1898), p. 170; Vol. 62, (1899), p. 125, and Vol. 63, (1900), p. 72.

9. Edgeworth, F. Y., Use of Analytical Geometry to Represent Certain Kinds of Statistics. Journal of Royal Statistical Society, Vol. 77, (1913-14), p. 300; p. 415; p. 653; and p. 838.

10. Edgeworth, F. Y., On Mathematical Representation of Statistical Data. Journal of Royal Statistical Society, Vol. 79, (1916), p. 455; Vol. 80, (1917), p. 65; p. 266; and p. 411.

11. An Editorial, The Probable Errors of Frequency Constants, Biometrica, Vol. 2, (1902-03), p. 273.

12. Fisher, R. A., On the Mathematical Foundations of Theoretical Statistics. Philosophical Transactions of the Royal Society of London, Series A, Vol. 222 (1921), p. 309.

13. Fisher, R. A., Theory of Statistical Estimations. Proceedings of the Cambridge Philosophical Society, Vol. 22, (1925), p. 700.

14. Hardy, G. H., Littlewood, J. E., and Polya, G., Inequalities. Cambridge University Press, London, 1934.

15. Kapteyn, J. C., Skew Curves in Biology and Statistics. Published by the Astronomical Laboratory at Groningen. P. Noordhoff, Publisher, Groningen, 1903.

16. Kapteyn, J. C. and Van Uven, M. J., Skew Curves in Biology and Statistics, 2nd. paper. Published by the Astronomical Laboratory at Groningen, Hoitsema Brothers, Groningen, 1916.

17. Kendall, M. G., The Advanced Theory of Statistics, Vol. 2. Chas. Griffin & Co. Ltd., London, 1948.

18. Knopp, K., Theory and Application of Infinite Series. Blackie & Sons, Ltd., London, 1949.

19. Koshal, R. S., Maximal Likelihood and Minimal χ^2 in Relation to Frequency Curves, Annals of Eugenics, Vol. 9 (1939), p. 209.

20. Koshal, R. S., Application of the Method of Maxima Likelihood to the Improvement of Curves Fitted by Method of Moments, Journal of Royal Statistical Society, Vol. 96 (1933), p. 303.

21. Olshen, C. A., Transformation of the Pearson Type III Distribution, Annals of Mathematical Statistics, Vol. 9 (1938).

22. Pearson, E. S., Tables of Incomplete Gamma Functions. Cambridge University Press, London, 1946.

23. Rietz, H. L. Frequency Distributions Obtained by Certain Transformation of Normally Distributed Variates, Annals of Mathematics, Vol. 23 (1922), p. 291.

24. Smail, L. L., Elements of Theory of Infinite Processes. McGraw-Hill, New York, 1923.

25. Thorndike, Frances, Applications of Poisson's Probability Summation, The Bell System Technical Journal, Vol. 5, No. 4. October 1926.

26. Wicksell, S. D., On the Genetic Theory of Frequency, Arkiv för Matematik, Astronomi, och Fysik, Vol. 12 (1917), No. 20, p. 1.

27. Yuan, P. T., Logarithmic Frequency Distributions and Semi-logarithmic Correlation Surfaces, Annals of Mathematical Statistics, Vol. 4, (1933), p. 30.