Contributions to the theory of statistical estimation and other topics.

L. R. Sheaton.

D. Sc., 1959.
CONTENTS

A 1 On the Efficiency of the Method of Moments and Neyman's Type A Distribution.
2 Maximum Likelihood and the Efficiency of the Method of Moments.
3 Efficiency of the Method of Moments and the Gram-Charlier Type A Distribution.
4 Moment Estimators and Maximum Likelihood.
5 The Distribution of Moment Estimators.

B 1-5 A Determinantal Expansion for a Class of Definite Integral: Parts 1 to 5
6 Generalised Algebraic Continued Fractions related to Definite Integrals.

C 1 Inequalities for the Normal Integral, including a new continued fraction.
2 A semi-infinite Random Walk with Discrete steps.
3 Distributions associated with Random Walk and Recurrent Events. (J.G. Skellam and L.R. Shenton).
4 The Efficiency of Automatic Winding Machines with constant Patrolling time. (A.J. Howie and L.R. Shenton).
Thesis submitted by L.R. Shenton, B.Sc. (Manchester), Ph.D. (Edinburgh) to the University of Edinburgh for the degree of Doctor of Science.

October, 1959.
On the efficiency of the method of moments and Neyman’s type A distribution

BY L. R. SHENTON

In a paper by Neyman (1939) several types of compound Poisson distributions are derived and their application to empirical data mentioned. Neyman remarks that the method of fitting requires investigation. It is the object of this note to consider the efficiency of the method of moments and Neyman’s type A distribution of two parameters.

The efficiency of the method of moments, with particular reference to parameters of scale and location, has been discussed by R. A. Fisher (1921), in connexion with the Pearson system of frequency curves. More recently Fisher (1941) has used the covariance and information matrices to find an expression for the efficiency of the method of moments applied to the negative binomial distribution. The chief difficulty appears to be the evaluation of the information determinant, and the process given here may have applications in other cases.

1. NEYMAN’S TYPE A DISTRIBUTION

The probability function is given by
\[
\exp\left\{-m_1[1-e^{-m_2(x-1)}]\right\} = \sum_{x=0}^{\infty} P_x e^{x},
\]

with
\[
P_x = \frac{e^{-m_1}m_2^x}{x!} \left\{0^x + \lambda_1^x \frac{\lambda_2^{2x}}{2!} + \lambda_3^x \frac{\lambda_4^{3x}}{3!} + \cdots \right\},
\]

where
\[
\lambda_1 = m_1 e^{-m_2}, \quad 0^x = 1, \quad x = 0,
\]
\[
= 0, \quad x \neq 0.
\]

The cumulants \( K_r \) are given by
\[
\sum_{r=1}^{\infty} K_r e^{x} = \sum_{s=1}^{\infty} m_1 m_2^{s}(e^{x} - 1)^{s}/s!.
\]

The first few are recorded for later use
\[
K_1 = m_1 m_2, \quad K_2 = m_1 m_2(1 + m_2),
K_3 = m_1 m_2(1 + 3m_2 + 5m_4),
K_4 = m_1 m_2(1 + 7m_2 + 6m_2^2 + m_4),
K_5 = m_1 m_2(1 + 15m_2 + 25m_2^2 + 10m_2^3 + m_4),
K_6 = m_1 m_2(1 + 31m_2 + 90m_2^2 + 65m_2^3 + 15m_2^4 + m_4),
\]

The relations
\[
K_{r+1} = m_1 \left( K_r + \frac{r(r-1)}{2} K_{r-2} + \cdots + r K_1 + m_1 \right)
\]

and
\[
K_{r+1} = m_1 \left( K_r + \frac{\partial K_r}{\partial m_2} \right),
\]

are easily proved from (3), the second expression being useful if high-order cumulants are required.

By the method of moments \( m_1 \) and \( m_2 \) are estimated from
\[
\bar{\lambda}_1 = m_1 m_2, \quad \bar{\lambda}_2 = m_1 m_2(1 + m_2),
\]

where \( \bar{\lambda}_1 \) and \( \bar{\lambda}_2 \) are the first two sample moments about the mean. For large samples of \( n \), we find
\[
\begin{align*}
\text{var} m_1 &= m_1 [2 + m_2^2 + 2m_1(1 + m_2)^3]/m_2, \\
\text{var} m_2 &= [2 + m_2^2 + 2m_1(1 + m_2)^3]/m_1, \\
\text{cov}(m_1, m_2) &= -2[1 + m_1(1 + m_2)^3]/m_2,
\end{align*}
\]

so that the covariance matrix
\[
\begin{bmatrix}
\text{var} m_1 & \text{cov}(m_1, m_2) \\
\text{cov}(m_1, m_2) & \text{var} m_2
\end{bmatrix}
\]

has a determinant of value
\[
(1 + (1 + m_2)^3 + 2m_1(1 + m_2)^3)/n^2m_2.
\]

To find the efficiency of fitting the first two moments, we require the determinant of the information matrix.
2. Likelihood equations and information matrix

By differentiating the g.f. (1) we find

$$\frac{\partial P_x}{\partial m_2} = (xP_x - (x+1)P_{x+1})/m_2 \quad \text{and} \quad \frac{\partial P_x}{\partial m_1} = -P_x + \frac{(x+1)}{m_1m_2}P_{x+1}.$$  

(7)

The likelihood of the sample \((n_0, n_1, n_2, \ldots, n_n)\) is

$$L = \prod_{x=0}^{n} (P_x)^{n_x},$$

and so for optimum statistics \(\hat{m}_1, \hat{m}_2\)

$$\sum n_x(x+1)\frac{P_{x+1}}{P_x} = n\hat{m}_1\hat{m}_2 = n\bar{x},$$

with \(\sum\) indicating summation over the sample. \(m_1, m_2\) is therefore efficiently estimated by the mean, and to complete the solution we must approximate to

$$\sum n_x(x+1)\frac{P_{x+1}}{P_x} = n\bar{x}.$$  

(8)

If \(m_1, m_2\) are estimates by moments, they can be improved by (8). For \(m_1, m_2 = \hat{m}_1, \hat{m}_2 = \bar{x},\) and writing \(P_x(m_2)\) to indicate \(P_x\) with \(m_1 = \bar{x}/m_2,\) we have

$$\sum n_x(x+1)\frac{\partial}{\partial m_2} P_{x+1}(m_2) = \frac{1}{m_2}\sum n_x(x+1)\frac{P_{x+1}}{P_x} - \frac{(m_1 + 1)}{m_2} \sum \frac{(x+1)(x+2)}{P_x} \frac{P_{x+2} - (x+1)^2 P_{x+3}}{P_x^2}.$$  

(9)

Thus if

$$F(m_2) = \sum n_x(x+1)\frac{P_{x+1}}{P_x} - n\bar{x}^2 \quad \text{and} \quad \hat{m}_2 = m_2 + \delta m_2$$

then

$$\delta m_2 = -\frac{F(m_2)}{F'(m_2)},$$

where (9) gives \(F'(m_2).\) The improved solution is therefore obtained by the use of frequencies calculated in the moment solution.

We now proceed to consider the efficiency. Using equations (7) and (4) we have

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

$$E\left(-\frac{\partial^2}{\partial m_1 \partial m_2} \log P_x\right) = E\left(\frac{1}{P_x} \frac{\partial^2 P_x}{\partial m_1 \partial m_2}\right) = 1 + m_1 - \phi/m_1m_2^2 = \delta_{m_1}\delta_{m_2},$$

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

where

$$\phi = E((x+1)^2 P_{x+1}/P_x^3),$$

and for the information determinant

$$\begin{vmatrix} m_1^2m_1 & m_1^2m_2 \\ m_1m_1 & m_1m_2 \end{vmatrix} = n^2[(1 + m_2) \phi - m_1m_2^2(m_1 + m_1m_2 + m_2)/m_1m_2^2].$$  

(10)

3. The evaluation of \(\phi = E((x+1)^2 P_{x+1}/P_x^3)\)

Since

$$P_x = e^{-m_2} \sum_{x=1}^{\infty} \left(\frac{\lambda^x}{1!} + \frac{\lambda^2x}{2!} + \frac{\lambda^3x^2}{3!} + \ldots\right),$$

$$e^{-m_2} \sum_{x=1}^{\infty} \left(A_4 + A_1x + A_3x(x-1) + A_2x(x-1)(x-2) + \ldots\right),$$

where the \(A_i\)'s may be determined by the Gregory-Newton formula of interpolation, it is evident that we may set up orthogonal polynomials with respect to \(P_x,\) defined by

$$\sum_{x=0}^{\infty} \theta_r(x) \theta_s(x) P_x = 0, \quad r \neq s,$$

$$\pm 0, \quad r = s,$$

$$\theta_r(x) = x^r e^{-x} \sum_{x=1}^{\infty} \left(A_4 + A_1x + A_3x(x-1) + A_2x(x-1)(x-2) + \ldots\right).$$
with \( \theta_0 = 1 \), say. With these polynomials we may then find an expression for \((x + 1) P_{x+1}\) in the form
\[
(B_0 + B_1 \theta_1(x) + B_2 \theta_2(x) + \ldots) P_x,
\] where the B's are functions of \( m_1 \) and \( m_2 \). In fact we have
\[
\Sigma(x + 1) P_{x+1} \theta(x) = B_x \Sigma \theta(x) P_x.
\] Squaring (11) and summing, we have
\[
\phi = \sum_0^\infty (x + 1)^2 \frac{P_{x+1}^2}{P_x} = B_x^2 \Sigma P_x + B_x^2 \Sigma \theta(x)^2 P_x + \ldots + B_x^2 \Sigma \theta(x) P_x + B_x^2 \Sigma \theta(x) P_x + \ldots
\] (12)
This is a series of positive terms, the first two of which amount to \( m_1 m_2^2 [m_1(1 + m_2) + m_2]/(1 + m_2) \), and thus our expression (10) for the determinant of the information matrix is a series of positive terms. The first term of significance turns out to be the fourth in (12). The value of (12) can therefore be found in two stages; (a) the determination of the orthogonal polynomials, (b) the evaluation of the B's.

(a) The evaluation of the \( \theta \)'s will be illustrated by finding \( \theta_3(x) \). It appears most convenient to assume
\[
\theta_3(x) = (x - \lambda_1)^3 - \mu_3 + A[(x - \lambda_1)^2 - \mu_2] + B[(x - \lambda_1)]
\]
where \( \lambda_1 \) is \( m_1 m_2 \) and the \( \mu \)'s refer to the moments of \( P_x \) about the mean. The orthogonality conditions lead to
\[
\begin{align*}
\mu_4 &+ A\mu_3 + B\mu_2 = 0, \\
\mu_5 - \mu_2 \mu_3 + A[\mu_4 - \mu_2^2] &+ B\mu_3 = 0, \\
\mu_6 - \mu_2 \mu_4 + A[\mu_5 - \mu_3 \mu_2] &+ B\mu_4 = 0.
\end{align*}
\] (13)
The moments in (13) may be replaced by cumulants, using well-known formulae
\[
K_2 = \mu_2, \quad K_3 = \mu_3, \quad K_4 + 3K_2^2 = \mu_4,
\]
\[
K_5 + 10K_3 K_2 = \mu_5, \quad K_6 + 15K_4 K_2 + 10K_3^2 + 15K_2^3 = \mu_6
\]
e tc., and these in turn expressed in terms of \( m_1 \) and \( m_2 \) by means of (4).

(b) For \( B_3 \) we have
\[
B_3 \sum_0^\infty \theta(x)^2 P_x = \sum_0^\infty (x + 1) P_{x+1} \theta(x),
\]
s o that we require the value of
\[
\sum_0^\infty (x - \lambda_1 + \lambda_1) \theta_3(x - 1) P_x.
\]
Expanding and simplifying, this becomes
\[
\mu_4 - 3\mu_3 + 3(1 - \lambda_1) \mu_2 - \lambda_1 + A[\mu_3 - 2\mu_2 + \lambda_1] + B[\mu_2 - \lambda_1].
\]
Hence from (13) we have
\[
B_3^2 \sum_0^\infty \theta(x)^2 P_x = \begin{vmatrix}
\mu_4 & \mu_3 & \mu_2 \\
\mu_5 - \mu_2 \mu_3 & \mu_4 & \mu_3 \\
\mu_6 - \mu_2 \mu_4 & \mu_5 - \mu_3 \mu_2 & \mu_4
\end{vmatrix}
\]
\[
= \begin{vmatrix}
m_1 m_2^2 [1 + m_1(1 + m_2)(4 + m_2)]^2 \\
[2 + 2m_2 + m_2^2 + 2m_1(1 + m_2)]^2 \\
\end{vmatrix},
\]
where
\[
\alpha = 12 + 12m_2 + 6m_2^2 + 2m_2^3 + m_2^4 (48 + 144m_2 + 144m_2^2 + 80m_2^3 + 25m_2^4 + 2m_2^5)
\]
\[
+ 24m_2^3 (1 + m_2)^5 (2 + 2m_2 + m_2^2) + 12m_2^3 (1 + m_2)^6.
\] (14)
For the earlier polynomials it may be verified that
\[
\begin{align*}
\theta_0(x) &= 1, \quad \Sigma \theta(x)^2 P_x = 1, \quad B_0 = \Sigma \theta(x) P_x = 1, \\
\theta_1(x) &= x - m_1 m_2, \quad B_1 \Sigma \theta(x) P_x = m_1 m_2^2, \\
\theta_2(x) &= (x - m_1 m_2)^2 - \frac{1 + 3m_2 + m_2^3}{1 + m_2} (x - m_1 m_2) - m_1 m_2 (1 + m_2), \\
B_2 \Sigma \theta(x) P_x &= \frac{m_1 m_2^2}{1 + m_2} (1 + m_2)^2 + 2m_1 (1 + m_2)^3.
\end{align*}
\]
Inserting these values in (10) and multiplying by the covariance determinant (6), we have for the efficiency
\[ \frac{1}{E} > 1 + \frac{4m_2(1 + m_2)(1 + m_1(1 + m_2)(4 + m_2))^2}{\alpha}, \]
where \( \alpha \) is given by (14).

Table 1 gives an upper bound to the value of percentage efficiency, \( E \), for various values of \( m_1 \)

<table>
<thead>
<tr>
<th>( m_1 )</th>
<th>0-1</th>
<th>0-5</th>
<th>1-0</th>
<th>3-0</th>
<th>4-0</th>
<th>6-0</th>
<th>10-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1</td>
<td>96</td>
<td>93</td>
<td>93</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td>97</td>
</tr>
<tr>
<td>0-2</td>
<td>92</td>
<td>88</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>91</td>
<td>93</td>
</tr>
<tr>
<td>0-5</td>
<td>82</td>
<td>76</td>
<td>77</td>
<td>80</td>
<td>83</td>
<td>85</td>
<td>89</td>
</tr>
<tr>
<td>1-0</td>
<td>73</td>
<td>67</td>
<td>69</td>
<td>76</td>
<td>80</td>
<td>83</td>
<td>87</td>
</tr>
<tr>
<td>2-0</td>
<td>65</td>
<td>61</td>
<td>67</td>
<td>75</td>
<td>81</td>
<td>84</td>
<td>88</td>
</tr>
<tr>
<td>3-0</td>
<td>62</td>
<td>60</td>
<td>67</td>
<td>77</td>
<td>82</td>
<td>85</td>
<td>89</td>
</tr>
<tr>
<td>5-0</td>
<td>59</td>
<td>59</td>
<td>68</td>
<td>79</td>
<td>84</td>
<td>87</td>
<td>91</td>
</tr>
</tbody>
</table>

and \( m_2 \). It is clear from Table 1 that if \( m_2 \) is small, say less than 0-2, then \( E \) is generally in the region of 90% or above. For 0-2 < \( m_2 \) < 1-0, \( E \) may be as low as 70%, but otherwise it lies between 75 and 90%, and for these values it is not easy to decide whether the ‘improved’ fit is worth the additional computation involved. If \( m_2 \geq 1-0 \) and \( m_1 \leq 3-0 \), \( E \) may be less than 70% and in this case the maximum likelihood approximation (9) can be applied.

4. A numerical illustration

That Neyman’s type A distribution of two parameters and the negative binomial are related, in that they may arise as the result of heterogeneity in the population, has been pointed out by W. Feller (1943). The similarity of the two may be appreciated by comparing the first three moments:

\[
\begin{align*}
\text{Neyman's type A} & \\
\text{parameters } m_1 \text{ and } m_2 & \\
\mu_1 &= m_1 m_2 \\
\mu_2 &= m_1 m_2 (1 + m_2) \\
\mu_3 &= m_1 m_2 (1 + 3 m_2 + m_2^2) \\
\text{Negative binomial} & \\
\text{parameters } m_1 \text{ and } m_2 & \\
\mu_1 &= m_1 m_2 \\
\mu_2 &= m_1 m_2 (1 + m_2) \\
\mu_3 &= m_1 m_2 (1 + 3 m_2 + 2 m_2^2)
\end{align*}
\]

It is thus reasonable to suppose that data satisfactorily described by the negative binomial may be suitable for our present purpose provided the parameters satisfy the conditions \( m_2 \geq 1-0 \), \( m_1 \leq 3-0 \). Such a distribution is given by Ove Lundberg (1940) concerning insurance claims and incapacity caused by sickness or accident.

The moments of the distribution turn out to be
\[ \bar{x}_1 = 2-805,871, \quad \bar{x}_2 = 6-404,549, \]
from which, using the method of moments, we obtain the estimates
\[ m_1 = 2-187,723, \quad m_2 = 1-282,553. \]

 Frequencies for these values are obtained by using the expression due to Beall (Neyman, 1939), namely,
\[ (x + 1) P_{x+1} = m_1 m_2 e^{-m_1} \left\{ P_x + \frac{m_2}{1!} P_{x-1} + \frac{m_2^2}{2!} P_{x-2} + \ldots \right\}, \]
with
\[ m_1 m_2 e^{-m_1} = 0-778,1466, \quad P_0 = e^{-a}, \quad \text{where} \quad a = m_1 (1 - e^{-m_1}), \]
so that \( P_0 = 0-205,7682. \)
The fit by moments is given in column 3 of Table 2. For an improved fit we set up a table of values consisting of (a) \( (x+1) P_{x+1}/P_x \) (b) \( (x+1)^2 P_{x+1}^2/P_x^2 \), which is easily found from (a), (c) \( (x+2) P_{x+2}/P_x \) in each of which \( x \) takes the values 0, 1, 2, ..., to 15. In the present case we have

\[
\begin{align*}
\sum n_x (x+1) P_{x+1}/P_x &= 2983.569,210, \\
\sum n_x (x+1)^2 P_{x+1}^2/P_x^2 &= 10550.201,104, \\
\sum n_x (x+2) P_{x+2}/P_x &= 12126.916,881.
\end{align*}
\]

Using these in (9) we find \( \delta n_2 = -0.14863 \), and the improved estimates \( \hat{\lambda}_1 = 2.474,481, \hat{\lambda}_2 = 1.133,923 \). Since with these values \( \sum n_x (x+1) P_{x+1}/P_x = 2965.031,424 \), it is clear we are much nearer a maximum likelihood solution for which

\[
\sum n_x (x+1) P_{x+1}/P_x = n\hat{\lambda}_1 = 2963.
\]

The fit with the improved estimates is shown in column 4 of the table. The improvement is probably exaggerated by the value of \( \chi^2 \), where we have grouped the last five frequencies. A more reliable idea is given by \( \log_{10} L \), shown at the bottom of the table, and it is evident that an improvement has been effected.

REFERENCES


MAXIMUM LIKELIHOOD AND THE EFFICIENCY OF THE
METHOD OF MOMENTS

BY L. R. SHENTON

1. INTRODUCTION

One of the practical difficulties in estimation by maximum likelihood arises from the fact that, except in special cases, the resulting equations are complicated and not easily solved. The present paper approaches this problem by deriving an expansion based on moment approximations to the likelihood equations. These are also not easy to solve explicitly, but they furnish a method of deriving the efficiency of the method of moments in terms of an expansion of the determinant of the information matrix. The efficiency of moment fitting as applied to Pearson's distributions has been discussed by Fisher (1921), who more recently treated the case of the negative binomial (Fisher, 1941).

2. THE LIKELIHOOD EQUATIONS AND ORTHOGONAL POLYNOMIALS

Let \( P(x; \theta_1, \theta_2, \ldots, \theta_s) \) be the probability of the variate \( x \), depending on \( s \) parameters \( \theta_j \) \((j = 1, 2, \ldots, s)\). We shall assume that \( P(x; \theta_j) \) possesses first derivatives with respect to the \( \theta_j \), and that its moments \( \mu_j \) and their first derivatives exist and are finite. In this case (Kacmarz & Steinhaus, 1935) it is possible to construct a system of polynomials

\[
q_r(x) = \sum_{i=0}^{r} A_{rij} x^i \quad (A_{rij} \neq 0, r = 0, 1, 2, \ldots),
\]

which are orthogonal with respect to \( P(x; \theta) \), satisfying

\[
\int q_r q_s P(x; \theta) \, dx = \begin{cases} 
0 & (r \neq s), \\
\pm 0 & (r = s),
\end{cases}
\]

or

\[
\sum q_r q_s P(x; \theta) = \begin{cases} 
0 & (r \neq s), \\
\pm 0 & (r = s),
\end{cases}
\]

where the integral includes finite summation for a discrete variate. These may be replaced by

\[
\int q_r \phi_s(x) P(x; \theta) \, dx = \begin{cases} 
0 & (r > s), \\
\pm 0 & (r = s),
\end{cases}
\]

with a similar expression for (3), where \( \phi_j(x) \) is an arbitrary polynomial of degree \( j \).

It is convenient to assume for \( q_r \) the form

\[
q_r(x) = \sum_{i=1}^{r} B_{rij} (x^i - \mu_i^r) \quad (r = 1, 2, \ldots; \mu_0^0 = 1),
\]

where \( \mu_i^r \) is the \( s \)th crude moment of \( P(x; \theta) \).

Choosing \( \phi_j(x) = x^j \), the orthogonality conditions lead to

\[
\sum_{j=1}^{r} B_{rij} (\mu_{j+k}^r - \mu_j^r \mu_k^r) = 0 \quad (k = 1, 2, \ldots, r - 1),
\]

\[
-\int q_r^2 P(x; \theta) \, dx + \sum_{j=1}^{r} B_{rij} (\mu_{j+r}^r - \mu_j^r \mu_r^r) = 0.
\]
Maximum likelihood and the efficiency of the method of moments

Writing \( m'_r \) for sample moments, we have

\[
N \text{ cov} (m'_r, m'_s) = \mu'_r \mu'_s = (r, s) \quad \text{say},
\]

where \( N \) is the number in the sample, so that from (4) and (5)

\[
q_r(x) = \begin{vmatrix}
  x^r - \mu'_r & x^{r-1} - \mu'_r & \cdots & x - \mu'_1 \\
  (r, 1) & (r-1, 1) & \cdots & (1, 1) \\
  (r, 2) & (r-1, 2) & \cdots & (1, 2) \\
  \vdots & \vdots & \ddots & \vdots \\
  (r, r-1) & (r-1, r-1) & \cdots & (1, r-1)
\end{vmatrix},
\]

and from (5) and (6)

\[
\int q_r^2 P(x; \theta) \, dx = (-)^r \begin{vmatrix}
  (r, 1) & (r-1, 1) & \cdots & (1, 1) \\
  (r, 2) & (r-1, 2) & \cdots & (1, 2) \\
  \vdots & \vdots & \ddots & \vdots \\
  (r, r) & (r-1, r) & \cdots & (1, r)
\end{vmatrix} = \phi_r \quad \text{say}.
\]

Suppose further that

\[
\frac{\partial P_x}{\partial \theta_j} = \sum_{k=0}^{r} C_{jk} q_k P_x \quad (j = 1, 2, \ldots, s),
\]

where \( P_x = P(x; \theta) \). Then

\[
\int \frac{\partial P_x}{\partial \theta_j} \, dx = \phi_r C_{jr} \quad (j = 1, 2, \ldots, s; r = 0, 1, \ldots)
\]

since \( \frac{\partial}{\partial \theta_j} \int P_x \, dx = 0 \) provided the range of values of \( x \) is independent of the \( \theta_j \), and where we have assumed it is permissible to differentiate under the integration sign. Clearly \( C_{00} = 0 \).

Now the maximum likelihood equations for the estimation of the parameters \( \theta_j \) are

\[
\Sigma' \frac{\partial}{\partial \theta_j} \log P_x = 0 \quad \text{or} \quad \Sigma' \frac{1}{P_x} \frac{\partial P_x}{\partial \theta_j} = 0 \quad (j = 1, 2, \ldots, s),
\]

the summation being over the sample.

From (9) we have the following formal expansion to determine optimum statistics:

\[
C_{j1}(m'_1 - \mu'_1) + C_{j2} + C_{j3} + \cdots = 0
\]

\[
(j = 1, 2, \ldots, s),
\]
where \( C_r \) is given by (12), and \((k, l) \equiv N \text{cov}(m'_k, m'_l)\). The equations (13) will in general be very difficult to solve, and from a practical point of view of doubtful value. To use them at all we must curtail the series at a certain stage and assume that the remainders are negligible. Further, since the sampling variance of \( m'_s \) is large for \( s > 4 \) or 5, the inclusion of additional terms to improve mathematical accuracy will partly defeat its own object. The most suitable application would appear to be to large samples when the number of parameters is one or two, even in this case the usual method of using an inefficient estimator and successively approximating from the likelihood equation may be simpler. We turn to the main object of the paper which is an expansion for the efficiency of the method of moments.

3. Efficiency of the Method of Moments

Using the first \( s \) sample moments to estimate the \( s \) parameters, we have (Fisher, 1941) for the efficiency \( E_f \)

\[
\frac{1}{E_f} = \left| E \left[ \frac{1}{P_x^2} \frac{\partial P_x}{\partial \theta_j} \frac{\partial P_x}{\partial \theta_k} \right] \right| \begin{vmatrix} \text{var} \bar{\theta}_1 & \text{cov} (\bar{\theta}_1, \bar{\theta}_2) & \cdots & \text{cov} (\bar{\theta}_1, \bar{\theta}_s) \\ \text{cov} (\bar{\theta}_2, \bar{\theta}_1) & \text{var} \bar{\theta}_2 & \cdots & \text{cov} (\bar{\theta}_2, \bar{\theta}_s) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov} (\bar{\theta}_s, \bar{\theta}_1) & \text{cov} (\bar{\theta}_s, \bar{\theta}_2) & \cdots & \text{var} \bar{\theta}_s \end{vmatrix}
\]

\[
= \left| E \left[ \frac{1}{P_x^2} \frac{\partial P_x}{\partial \theta_j} \frac{\partial P_x}{\partial \theta_k} \right] \right| \left[ \text{cov} (m'_1 m'_s) \right] \left( \begin{smallmatrix} \text{det} \left( \bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_s \right) \\ \text{det} (m'_1, m'_2, \ldots, m'_s) \end{smallmatrix} \right)^2 \tag{14}
\]

where \( \bar{\theta}_j \) are consistent estimates of \( \theta_j \) obtained from the sample moments, \( [v_{jk}] \) is a matrix with \( j, k = 1, 2, \ldots, s \) and \( \frac{\partial (\theta_1, \theta_2, \ldots, \theta_s)}{\partial (\mu_1, \mu_2, \ldots, \mu_s)} \) is a Jacobian.

But from (9) we have

\[
E \left( \frac{1}{P_x^2} \frac{\partial P_x}{\partial \theta_j} \frac{\partial P_x}{\partial \theta_k} \right) = \int \left( \sum_{m=0}^{\infty} C_{jm} q_m \right) \left( \sum_{n=0}^{\infty} C_{kn} q_n \right) P_x \, dx
\]

\[
= \sum_{\lambda=1}^{\infty} C_{jm} C_{kl} \phi_{\lambda} \tag{15}
\]

Hence from (14) and (15)

\[
\frac{1}{E_f} = \phi_1 \phi_2 \cdots \phi_s \left( \frac{\partial \theta}{\partial \mu'} \right)^2 \left[ \sum_{\lambda=1}^{\infty} C_{jm} C_{kl} \phi_{\lambda} \right]
\]

\[
= \phi_1 \phi_2 \cdots \phi_s \left( \frac{\partial \theta}{\partial \mu'} \right)^2 \left( \begin{vmatrix} C_{11} & C_{21} & \cdots & C_{s1} \\ C_{12} & C_{22} & \cdots & C_{s2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1s} & C_{2s} & \cdots & C_{ss} \end{vmatrix} \phi_1 \phi_2 \cdots \phi_s + \begin{vmatrix} C_{11} & C_{21} & \cdots & C_{s1} \\ C_{12} & C_{22} & \cdots & C_{s2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1s} & C_{2s} & \cdots & C_{ss} \end{vmatrix} \phi_1 \phi_2 \cdots \phi_{s-1} \phi_s+1 \right) \tag{16}
\]

where we have used \( \frac{\partial \theta}{\partial \mu'} \) for the Jacobian appearing in (14), and where the determinants are of order \( s \times s \) whose terms consist of all distinct combinations of \( s \) from 1, 2, 3, \ldots, etc., used in the determinant

\[
\begin{vmatrix} C_{p1} & C_{q1} & \cdots & C_{sr} \\ C_{p2} & C_{q2} & \cdots & C_{wr} \\ \vdots & \vdots & \ddots & \vdots \\ C_{ps} & C_{qs} & \cdots & C_{ws} \end{vmatrix} \phi_p \phi_q \phi_r \cdots \phi_w
\]

with \( p, q, r, \) etc., as positive integers.
The first term of (16) after elementary operations on rows and columns of the determinant reduces to

\[ \phi_1 \phi_2 \ldots \phi_s \left( \frac{\partial}{\partial \mu} \right)^2 \left( \frac{1}{\phi_1 \phi_2 \ldots \phi_s \frac{\partial}{\partial \theta}} \right) \phi_1 \phi_2 \ldots \phi_s = 1. \]

Since \( \phi_r = \int q_r P_x dx \) and \( P_x \) is essentially positive for the range of values of \( x \) for which it is defined, it is clear that \( \phi_r \) is positive. Hence (16) consists of a series of positive terms, the first few of which will provide a lower bound to \( 1/E_f \) or upper bound to \( E_f \). In general it is quite sufficient to use the first two or three terms.

4. Application

As an example of the foregoing we take a 'hypergeometric' distribution recently discussed by J. G. Skellam (1948). From it, as Skellam shows, the negative binomial distribution may be found as a limiting case, for which the efficiency of the method of moments has been given. Taking the binomial distribution

\[ \phi(x) = \binom{n}{x} p^x q^{n-x}, \]

where the probability varies according to the law

\[ \phi(y) = k p^{a-1} (1-p)^{b-1} \quad (0 \leq p \leq 1), \]

with \( k \) a constant and \( a \) and \( b \) any real positive values, Skellam arrives at the probability function

\[ P(x; \alpha, \beta) = \binom{n}{x} (\alpha + x - 1)^{\alpha} (\beta + n - x - 1)^{\beta} \]

\[ (\alpha + \beta + n + 1)^{\alpha + \beta + n + 1} \]

with

\[ \binom{n}{x} = \frac{n(n-1) \ldots (n-x+1)}{x!} \quad \text{and} \quad N(x) = N(N-1) \ldots (N-x+1). \]

It is assumed that \( n \) is a known positive integer, so that there are two parameters \( \alpha \) and \( \beta \) for estimation. If the value of \( n \) is also to be estimated the derivation of \( E_f \) will be slightly more complicated. The distribution (17) may be regarded as a negative hypergeometric, in the sense that in the usual form of the hypergeometric probability distribution

\[ \psi(x) = \binom{n}{x} (Np)^{\alpha} (Nq)^{x-\alpha} N^x \]

we need only put \( Np = -\alpha, Nq = -\beta \). The negative binomial follows by putting \( \beta = n/c \) and letting \( n \to \infty \). Using known properties of the hypergeometric distribution and its associated orthogonal polynomials (Aitken & Gonin, 1935), we have

\[ \sum_{x=0}^{\infty} P_x x^r = F(-n, \alpha; \alpha + \beta; 1 - t), \]

\[ q_r(x) = F(n-r+1, -\alpha - r + 1; -\alpha - 2r + 2; -\Delta_x), \]

\[ \phi_r = r! \binom{n+r-1}{\alpha+r-1} (\alpha+r-1)^{\alpha+r-1} (x+\beta+r-n-1)^{\beta+r-1} \]

\[ (x+\beta+2r-1)^{\beta} (x+\beta+2r-2)^{\beta} \]

\[ (r > 0), \]

where

\[ F(a, b; c; d) = 1 + \frac{ab}{1!c} + \frac{a(a+1)b(b+1)d}{2!c(c+1)} + \ldots \]

and

\[ \Delta_x u_x = u_{x+1} - u_x. \]
Putting
\[ \frac{\partial P}{\partial x} = \sum_r A_r q_r P_x, \quad \frac{\partial P}{\partial \beta} = \sum_r B_r q_r P_\beta \]
it may be seen that
\[ A_0 = B_0 = 0, \]
\[ A_r \phi_r = \frac{(-r-1) n^{(r+1)} (\beta + r - 1)^{\alpha} (r-1)!}{(\alpha + \beta + 2r - 2)^{\alpha} (\alpha + \beta + r - 1)^\alpha}, \quad B_r \phi_r = \frac{-n^{(r+1)} (\alpha + r - 1)^{\alpha} (r-1)!}{(\alpha + \beta + 2r - 2)^{\alpha} (\alpha + \beta + r - 1)^\alpha}. \] (21)

Using (18) for the factorial moments we find
\[
\frac{1}{E_f} = 1 + \frac{A_r A_s}{B_r B_s} \left[ \phi_1 \phi_2 \phi_3 + \frac{A_1 A_2 A_3}{B_1 B_2 B_3} \frac{\phi_2 \phi_3 + A_1 A_4}{B_1 B_3} \phi_4 \cdots \right]
\]
\[ = 1 + \frac{4(n-2)(\alpha + \beta + 1)(\alpha + \beta + 3)(\alpha + \beta + 5)(\alpha - \beta)^2}{3(\alpha + 2)(\beta + 2)(\alpha + \beta + 4)^2(n + \alpha + \beta + 2)}
\]
\[ + \frac{2(n-1)(\alpha + 1)(\beta + 1)(\alpha + \beta)(\alpha + \beta + 4)^2(n + \alpha + \beta + 5)}{3(\alpha + 2)(\beta + 2)(\alpha + \beta + 4)^2(n + \alpha + \beta + 2)^2}
\]
\[ + \frac{3(n-2)(\alpha + \beta + 1)(\alpha + \beta + 7)(\alpha + 3)(\beta + 3)^2}{(\alpha + 3)(\beta + 3)(\alpha + \beta + 2)^2(n + \alpha + \beta + 3)^2}
\]
\[ + \frac{2(n-1)(\alpha + 1)(\beta + 1)(\alpha + \beta)(\alpha + \beta + 4)^2(n + \alpha + \beta + 5)}{3(\alpha + 2)(\beta + 2)(\alpha + \beta + 4)^2(n + \alpha + \beta + 2)^2}
\]
\[ + \cdots. \] (22)

From (22) it can be seen that:

(a) The general term associated with \[ A_r A_s \] \( (r + s) \) is of the form \( K n^{(r+1)} n^{(s+1)} (n - 1)^2 \), where \( K \) does not contain \( n \), so that for all admissible suffixes except \( r, s = 1, 2 \), we see that \( (n - 2) \) is a factor. Hence when \( n = 2 \), \( E_f = 1 \), so that the moment solution is efficient, a fact which is otherwise obvious.

(b) If \( \alpha \) and \( \beta \) are both small then
\[ E_f^{-1} = 1 + \frac{7}{18} \frac{(n-2)(n-3)}{(n+2)(n+3)} + \cdots, \]
which for practical values of \( n \), say \( \leq 20 \), suggests that \( E \) may not be below 80%.

(c) If \( \alpha \) is small, \( \beta \) not small, say \( \geq 1 \), then
\[ E_f^{-1} = 1 + \frac{2(n-2)(\beta^2(\beta + 1)(\beta + 3)(\beta + 5)}{3(n + \beta + 2)(\beta + 2)^2} + \cdots, \]
which very rarely reduces \( E_f \) below 70%.

(d) If \( \alpha \) and \( \beta \) are both large, the relative degrees in \( \alpha \) and \( \beta \) of the numerators and denominators suggest high efficiency.

Thus the discrete 'hypergeometric' type distribution (17), depending on two parameters \( \alpha, \beta \), with range \( n \) (assumed known), is simple to fit by the first two factorial moments, and this method of estimating the parameters rarely has low efficiency. (It is suggested that the first four terms of (22) should be used in any practical case.) We may mention that Mood (1943) gives this distribution in connexion with sampling inspection plans, and remarks on some of its properties.

Skellam's example on the secondary association of chromosomes in *Brassica* is:

<table>
<thead>
<tr>
<th>No. of associations</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>32</td>
<td>103</td>
<td>122</td>
<td>80</td>
</tr>
</tbody>
</table>
for which Skellam finds:

(a) \( \alpha = 6.168, \beta = 4.455 \) by moments,
(b) 'improved' maximum likelihood estimates, \( \hat{\alpha} = 6.12, \hat{\beta} = 4.42 \),
(c) \( \chi^2 = 0.76 \) for both methods of fitting.

It is clear that the improvement introduced by maximum likelihood methods is negligible, indicating high efficiency for the method of moments. Putting \( \alpha = 6, \beta = 4.5, n = 3 \) in (22), the first four terms give \( E_T = 1/1.001 \) in good agreement.

It is of interest to notice the form of (22) when we take \( \beta = n/c \) in (17) and let \( n \to \infty \), so that the distribution is now the negative binomial. We find

\[
\frac{1}{E_T} = 1 + \frac{4c}{3(1+c)(x+2)} + \frac{3c^2}{(1+c)^2(x+3)^2} + \ldots,
\]

as given by Fisher (1941) by a different method. Moreover, the likelihood equations (13) take the form

\[
\Sigma \frac{\partial}{\partial c} \log P_x = \Sigma' q_1(x)/c(1+c) = 0, \\
\Sigma \frac{\partial}{\partial x} \log P_x = \Sigma' \left( \frac{q_1(x)}{x(1+c)} - \frac{q_2(x)}{2(x+1)^2(1+c)^2} - \frac{q_3(x)}{3(x+2)^3(1+c)^3} - \ldots \right) = 0, \tag{23}
\]

where

\[
q_r(x) = x^{r-1}c(x+r-1)x^{r-2} = \frac{r(r-1)}{2!} c(x+r-1)^2 x^{r-2} - \ldots.
\]

The first two terms of (23) together with \( \Sigma'(x-\alpha c) = 0 \) correspond to the moment solution; we may use the first three terms to 'improve' this solution, neglecting the remainder which in any case consists of terms whose expectation is zero. Taking an example given by Haldane (1941):

\[
x = 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad \Sigma
\]

Frequency 167 267 271 185 111 61 27 8 3 1 1096

for which the first three factorial moments are

\[ m_{11} = 2.156,034, \quad m_{21} = 5.100,365, \quad m_{31} = 12.651,460, \]

we find from the first three terms of (23)

\[ 1.344,003x^2 - 7.805,548x^2 - 40.625,167x - 140.187,797 = 0, \]

a solution of which is \( x = 9.918 \); Haldane gives \( x = 9.900 \) as the maximum likelihood solution and \( \alpha = 10.44 \) as the moment solution, so that the use of (23) does in this case lead to an improvement. It would appear desirable to know under what conditions series such as (23) converge rapidly, and whether it is not just those cases for which the efficiency of moment fitting is low which give slow convergence.

In conclusion I wish to express my thanks to Prof. M. S. Bartlett for helpful criticisms.

REFERENCES*


* The method of the present paper is applied to the evaluation of a class of definite integral in a paper to appear in a forthcoming issue of the Proceedings of the Royal Society of Edinburgh.
EFFICIENCY OF THE METHOD OF MOMENTS AND THE
GRAM-CHARLIER TYPE A DISTRIBUTION

BY L. R. SHENTON, College of Technology, Manchester

1. INTRODUCTION

We propose to consider the efficiency of the method of moments applied to fitting the Gram-Charlier series of Type A. In practice, since the sampling variances of moments higher than the fourth are large, only the first four moments of the observed frequency data are in general used, so that we shall confine our attention to the Type A series taken as far as the term involving the fourth moment.

It has been pointed out by Fisher (1921, pp. 355–6) that in the region of the normal point the Pearson system of curves closely resembles the system for which the method of moments is efficient; for 80 % efficiency this region is restricted to \(2.62 < \beta_2 < 3.42, \beta_1 < 0.1\). The question arises as to how the departures from normality expressed by the Gram-Charlier type distribution affect the efficiency of moment estimation and how this compares with the situation for the Pearson system.

To evaluate the efficiency of moment estimation for the Gram-Charlier series we shall use a slight modification of the method given in a previous paper on maximum likelihood and moments (Shenton, 1950). It consists in developing a determinantal expansion, which, since its terms are positive, provides a lower bound to the efficiency; in particular cases when there is one parameter for estimation it is possible to set lower and upper bounds to the efficiency.

2. THE EXPANSION FOR THE INFORMATION DETERMINANT

The expansion for the efficiency \(E_f\) given in M.L. expression (16) is capable of a simpler form in the case when the frequency function has a polynomial factor, as for example is the case with Type A. For the information determinant is of the form

\[
\Delta = \left| \begin{array}{c}
\int_a^b A_j(x) A_k(x) w(x) \frac{B(x)}{B'(x)} dx \\
\end{array} \right| (j, k = 1, 2, \ldots, n),
\]

and it may be shown under fairly general conditions that

\[
\Delta = (-)^n \lim_{\gamma \to \infty} \left| \begin{array}{c}
[a]_{ns} [\beta]_{ss} \quad + [\beta]_{ss},
\end{array} \right|
\]

where (i)

\[
[a]_{ns} = \left[ \begin{array}{cccc}
\alpha_0 & \alpha_1 & \cdots & \alpha_s \\
\cdots & \cdots & \cdots & \cdots \\
\alpha_0 & \alpha_1 & \cdots & \alpha_s \\
\end{array} \right],
\]

and \([a]_{ns}^t\) is its transpose.

(ii) \(\alpha_j = \int_a^b A_j(x) w(x) dx \quad (j = 1, 2, \ldots, n; k = 0, 1, \ldots, s).\)

\(\beta_{jk} = \beta_{kj} = \int_a^b p_j(x) p_k(x) B(x) w(x) dx \quad (j, k = 0, 1, \ldots),\)

* I am indebted to Prof. M. G. Kendall for this suggestion.

† To be referred to as M.L.
and \( p_j(x) \), \( j = 0, 1, \ldots \), is a polynomial in \( x \) of precise degree \( j \). The expansion (2), apart from a factor depending on the covariance term, is a compact form of (16) in M.L. and can be shown to converge in certain cases by an appeal to Parseval’s theorem. In particular it converges to \( \Delta \) when \( a = -\infty \), \( b = \infty \), \( A_1(x) \) is a polynomial in \( x \), \( w(x) = e^{-ix^2}/\sqrt{2\pi} \) and \( B(x) \) is a polynomial always positive for real \( x \). Since (2) is equivalent to M.L. (16), the expansion will give an increasing sequence as an approximation to \( \Delta \). It is important to remember that the polynomial \( B(x) \) must be positive for \( x \) in \((a, b)\). We must therefore consider the conditions under which this is true for the case of Type A.

3. The Admissible Parameter Values for Type A

We consider the frequency function

\[
P_x dx = \left\{ 1 + \frac{a_3}{3!} H_3(x) + \frac{a_4}{4!} H_4(x) \right\} g(x) dx,
\]

(5)

where

\[
x = (X - m)/\sigma, \quad g(x) = e^{-ix^2}/\sqrt{2\pi}
\]

and \( H_r(x) = e^{-iy^2}y^r \) is the \( r \)th Hermite polynomial with the property

\[
\int_{-\infty}^{\infty} H_r(x) H_s(x) g(x) dx = r! \quad (r = s),
\]

and

\[
0 \quad (r \neq s).
\]

(6)

Using \( \frac{d}{dx} H_r(x) = rH_{r-1}(x) \) along with \( x^r = e^{iy^2} H_r(x) \) the moments of (5) may be seen to be

\[
\mu_{2r} = \frac{(2r)!}{2^r r!} \left( 1 + r(r-1)\frac{a_4}{3!} \right) \quad (r = 1, 2, \ldots),
\]

(7)

\[
\mu_{2r-1} = \frac{(2r-1)!}{2^{r-1} (r-1)!} \frac{a_3}{3!} \quad (r = 2, 3, \ldots),
\]

so that for the measures of skewness and kurtosis we have

\[\beta_1 = \mu_3^2/\mu_2^3 = a_3^3, \quad \beta_2 = \mu_4/\mu_2^2 = a_4 + 3\]

\[\gamma_1 = a_3, \quad \gamma_2 = a_4.\]

Now a term in the information determinant such as \( E \left( \frac{\partial}{\partial m} \log P_x \right)^2 \) involves the quartic \( B(x) = 24 + 4a_3 H_3(x) + a_4 H_4(x) \) in the denominator so that to apply (2) we must ensure that this is positive for \(-\infty < x < \infty\). Writing

\[
H = -a_3^2 - a_4^2, \quad J = -24a_4^2 - 2a_3^2 - 24a_3^2 - 6a_3^2 a_4,
\]

\[
I' = 24a_4 + 6a_3^2 + 12a_3, \quad \nabla = I'^3 - 27J^2,
\]

the quartic will have four imaginary roots if \( \nabla > 0 \) and either \( H > 0 \) or \( 2HI' - 3a_4 J > 0 \). (See, for example, Burnside & Panton, 1881, pp. 116 and 187.) Hence \( B(x) \) is positive for real \( x \) provided

\[
(8a_4 + 2a_3^2 + 4a_3^3)^3 > (24a_4^2 + 2a_3^2 + 24a_3^2 + 6a_3^2 a_4)^2,
\]

(8)

and

\[
4a_3^2 - a_4^2 + 4a_3 a_4 > 3a_3^2 a_4^2 + 4a_4^3.
\]

Table 1

<table>
<thead>
<tr>
<th>( a_4 )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
<th>3.0</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_4 )</td>
<td>0.026</td>
<td>0.070</td>
<td>0.123</td>
<td>0.242</td>
<td>0.306</td>
<td>0.370</td>
<td>0.435</td>
<td>0.499</td>
<td>0.562</td>
<td>0.645</td>
<td>1.039</td>
<td>1.100</td>
<td>0.990</td>
<td>0.658</td>
<td></td>
</tr>
</tbody>
</table>
The second condition implies that $0 < a_4 < 4$. Table 1 shows the restriction imposed by (8) on $a_4$ for given $a_4$ in $(0, 4)$. For later use we compare the scope of Type A under these restrictions with the Pearson system of curves, using a $(\gamma_1, \gamma_2)$ diagram. The region of validity of Type A is that enclosed by the curved boundary and the $\gamma_2$ axis in Fig. 1. Superimposed on this is the corresponding division for the Pearson system (Pearson, 1945, diagram XXXV, p. 66). It will be seen that the Pearson curves corresponding to the admissible Type A curves are Type IV, Type VII (including the heterotypic cases of these) and Types V and VI, putting them in approximate order of importance. Type IV and the symmetrical case of this, Type VII, occupy most of the Type A region. We notice that Types I and III do not share the region.

Fig. 1. Showing the region of validity of Type A (enclosed by curve and $\gamma_2$ axis) and the corresponding Pearson distribution.

4. EFFICIENCY OF THE FOUR PARAMETER CASE OF TYPE A

The parameters $m, \sigma, a_3, a_4$ in (5) may be estimated by moments using

$$\bar{\theta}_1 = \Sigma'X/N, \quad \bar{\theta}_j = \Sigma'(X - \bar{\theta}_1)/N, \quad (j = 2, 3, 4).$$

The efficiency $E_f$ is given by

$$E_f^{-1} = [I \times \{\text{cov}(\bar{\theta}_j, \bar{\theta}_k)\}], \quad (9)$$

where the information matrix is

$$I = \begin{bmatrix} E\frac{N\partial P_\mu}{\partial \theta_j} \frac{\partial P_\mu}{\partial \theta_k} \end{bmatrix}.$$ 

Since to order $N^{-1}$

$$\text{cov}(m_j, m_k) = \{\mu_j - \mu_k\mu_k + jk\mu_2\mu_j\mu_k - j\mu_j - k\mu_{j+1}\mu_{k-1} - k\mu_{j+1}\mu_{k-1}\}/N,$$

using (7) it may be verified that

$$N^4[\text{cov}(\bar{\theta}_j, \bar{\theta}_k)] = \sigma^{20} \begin{bmatrix} 2 - a_3^2 + a_4 & 6a_3 - a_5a_4 & -8 + 4a_4 \\ 6a_3 - a_5a_4 & 6 - a_3^2 + 9a_4 - a_4^2 & 12a_3 - a_5a_4 \\ -8 + 4a_4 & 12a_3 - a_5a_4 & 56 + 24a_4 - a_4^2 \end{bmatrix} = \sigma^{20}V \text{ say.} \quad (10)$$

For the information matrix $I$ we take in (1)-(4)

$$A_j(x)g(x) = \frac{\partial P_\mu}{\partial \theta_j} \quad (j = 1, 2, 3, 4);$$

$$B(x)g(x) = P_\mu;$$

$$w(x) = g(x);$$

$$P_k(x) = H_k(x), \quad (k = 0, 1, \ldots);$$

$$a = -\infty, \quad b = \infty;$$
so that
\[ \beta_{jk} = \int_{-\infty}^{\infty} H_j(x) H_k(x) P_{\sigma} dx, \quad (j, k = 0, 1, \ldots). \] (12)

The integrals in \((12)\) may be evaluated by using the recurrence relation
\[ H_j(x) H_k(x) = H_{j+k}(x) + j k H_{j+k-1}(x) + \frac{j(j-1)k(k-1)}{2!} H_{j+k-2}(x) + \ldots \]

Moreover,
\[ \frac{\partial P_{x}}{\partial \phi_1} = \left[ H_1(x) + \frac{a_3}{3!} H_3(x) + \frac{a_4}{4!} H_4(x) \right] \frac{g(x)}{\sigma}, \]
\[ \frac{\partial P_{x}}{\partial \phi_2} = \left[ H_2(x) - \frac{H_4(x)}{2} + \frac{a_3}{3!} H_3(x) + \frac{a_4}{4!} H_4(x) \right] \frac{g(x)}{2\sigma^2}, \]
\[ \frac{\partial P_{x}}{\partial \phi_3} = \frac{H_3(x) g(x)}{3! \sigma^3}, \]
\[ \frac{\partial P_{x}}{\partial \phi_4} = \frac{H_4(x) g(x)}{4! \sigma^4}. \]

Hence from \((11)\) and \((12)\) we find
\[ 1 \alpha_1 = \frac{1}{\sigma}, \quad 1 \alpha_4 = \frac{4a_3}{\sigma}, \quad 1 \alpha_5 = \frac{5a_4}{\sigma}, \quad 1 \alpha_j = 0, \quad (j \neq 1, 4, 5); \]
\[ 2 \alpha_2 = \frac{1}{\sigma^2}, \quad 2 \alpha_4 = -\frac{3!}{\sigma^2}, \quad 2 \alpha_5 = \frac{10a_3}{\sigma^2}, \quad 2 \alpha_6 = \frac{15a_4}{\sigma^2}, \quad 2 \alpha_j = 0, \quad (j \neq 2, 4, 5, 6); \]
\[ 3 \alpha_3 = \frac{1}{\sigma^3}, \quad 3 \alpha_j = 0, \quad (j \neq 3); \]
\[ 4 \alpha_4 = \frac{1}{\sigma^4}, \quad 4 \alpha_j = 0, \quad (j \neq 4); \]

and
\[ \beta_{j, 1} j! = 1 + j(j-1) \frac{a_1}{4!} j! \equiv b_j, \]
\[ \beta_{j, j+1} = j(j+1)! \frac{a_3}{2} \equiv c_j, \]
\[ \beta_{j, j+2} = j(j+2)! \frac{a_4}{3!} \equiv d_j, \]
\[ \beta_{j, j+3} = (j+3)! \frac{a_5}{3!} \equiv f_j, \]
\[ \beta_{j, j+4} = (j+4)! \frac{a_6}{4!} \equiv g_j, \quad (j = 0, 1, \ldots); \]
\[ \beta_{j, j+k} = 0, \quad (k = 5, 6, \ldots). \]

Substituting in \((2)\) and eliminating as many of the border elements as possible we are led to
\[ \frac{\sigma^{20}}{N^4} \left| I \right| = \lim_{s \to \infty} \frac{K_0(B_0, C_0, d_5, f_5, g_5)}{K_0(B_0, c_0, d_0, f_0, g_0)}, \] (15)
Efficiency of the method of moments

with

\[ B_s = b_5 - 20\alpha_3a_4 - 10\alpha_3a_4g_1 + 25\alpha_3^2b_1 + 100\alpha_3^2g_1 + 100\alpha_3a_4c_1 \]
\[ = 120 + 600\alpha_3 - 200\alpha_3^2 + 200\alpha_3^2a_4 - 25a_3^2, \]
\[ C_s = c_5 - 10\alpha_3g_2 - 15\alpha_4g_2 + 75\alpha_2^2c_1 + 150\alpha_2a_4b_2 \]
\[ = 1800\alpha_3 + 225\alpha_3^2a_4 - 300\alpha_3a_4, \]
\[ B_s = b_6 - 300\alpha_4g_2 + 225\alpha_4^2b_2 \]
\[ = 720 + 5400\alpha_4 - 450\alpha_4^2 + 225a_4^2, \]
\[ B_s = b_s, \quad s > 6, \quad C_s = c_s, \quad s > 5, \]

where \( K_s(p_k, q_l, r_m, \ldots) \) is a symmetric determinant of order \( s \) with elements

\[ p_j, \quad (j = k, k + 1, k + 2, \ldots), \]
\[ q_j, \quad (j = l, l + 1, l + 2, \ldots), \]

in the diagonal through \((1, 1)\),

in the diagonal through \((2, 1)\),

and so on,

all unspecified elements being zero.* Since we may write \( V \) given by (10) as \( K_5(b_0, c_0, d_0, f_0, g_0) \) we have for the joint efficiency of moment estimates for the four parameters of Type A

large sampling

\[ E_s \leq \frac{K_4(b_0, c_0, d_0, f_0, g_0)}{K_5(b_0, c_0, d_0, f_0, g_0)} \]

(17)

the ratio forming a decreasing sequence and the equality sign being applicable in the limit as \( s \to \infty \). Thus for a particular value of \( s \), (17) gives an upper bound for \( E_s \). For computational purposes we may use \( s = 6 \) or \( s = 7 \) in (17), and the determinants \( K_j(b_0, c_0, d_0, f_0, g_0) \), \( j = 5, 6, 7 \) may be simplified to become the first \( j - 2 \) rows and columns of the array (18) below:

\[
\begin{bmatrix}
2 + a_4 - a_3^2 & 6a_3 - a_3a_4 & -8 + 4a_4 & -10a_3 & 30a_4 \\
6a_3 - a_3a_4 & 6 + 9a_4 - a_3^2 - a_3^2 & 12a_3 - a_3a_4 & -30 + 15a_4 + 5a_3^2 & 120a_3 \\
-8 + 4a_4 & 12a_3 - a_3a_4 & 56 + 24a_4 - a_3^2 & 100a_3 + 5a_3a_4 & 360a_4 \\
-10a_3 & -30 + 15a_4 + 5a_3^2 & 100a_3 + 5a_3a_4 & 270 + 225a_4 - 25a_3^2 & 1200a_3 \\
30a_4 & 120a_3 & 360a_4 & 1200a_3 & 720 + 5400a_4 \\
\end{bmatrix}
\]

Table 2. Values of \( E_s \) in the joint estimation of four parameters for Type A

<table>
<thead>
<tr>
<th>( a_4 )</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.99</td>
<td>0.97</td>
<td>0.92</td>
<td>0.88</td>
<td>0.84</td>
<td>0.81</td>
<td>0.78</td>
<td>0.75</td>
<td>0.73</td>
<td>0.71</td>
<td>0.69</td>
<td>0.63</td>
<td>0.59</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 2 we give values of (17) for \( s = 6 \) for admissible values of the parameters \( a_4^2 \) and \( a_4 \)

(or \( a_4^2 \) and \( a_2 \)) so that \( E_s \) is certainly less than the figure shown. From Table 2 it is seen that

in the vicinity of the normal point \( E_s \) approaches unity; this is also suggested from (17), since

* \( K_j(b_0, c_0, d_0, f_0, g_0) \) is a symmetric determinant with elements in nine diagonals only; it may therefore be regarded as a generalized continuant and the notation is suggested by Muir's (1882, pp. 149–60) for a continuant determinant.
from (13) \(c, d, f, g\) tend to zero with \(a_3\) and \(a_4\). Type A is therefore similar to the Pearson system in this respect. Further, it appears that \(E_f < 0.80\) % for \(0.6 \leq a_4 < 4\) and \(a_3^2 > 0.1\) and low efficiency is to be expected in the vicinity of the Type A boundary shown in Fig. 1. Although the efficiency seems to increase for increasing \(a_4\) (with \(a_3^2 = 0.1\) say) this may be accounted for by the degree of approximation involved. To calculate further terms of (17) becomes laborious, but the following additional values assist in fixing the approximate position of the bound for the 80 % efficiency contour: \(a_4 = 0.1, a_3^2 = 0.01, E_f < 0.8950\) with \(s = 6, E_f < 0.8763\) with \(s = 7: a_4 = 0.2, a_3^2 = 0.05, E_f < 0.7029\) with \(s = 6\). However in the case when \(a_3^2 = 0\) the computation may be carried further, since in this case the determinants in (17) now factorize (using Laplacian expansion) and we find

\[
E_f \leq \frac{\overline{K}_3(b_0, d_0, g_0) \overline{K}_4(b_1, d_1, g_1)}{\overline{K}_5(b_0, d_0, g_0) \overline{K}_6(b_1, d_1, g_1) \overline{K}_7(b_2, d_2, g_2) \overline{K}_8(b_3, d_3, g_3) \overline{K}_9(b_4, d_4, g_4)}
\]

(19) or

\[
E_f \leq \frac{\overline{K}_3(b_0, d_0, g_0) \overline{K}_4(b_1, d_1, g_1) \overline{K}_5(b_2, d_2, g_2) \overline{K}_6(b_3, d_3, g_3) \overline{K}_7(b_4, d_4, g_4) \overline{K}_8(b_5, d_5, g_5)}{\overline{K}_9(b_0, d_0, g_0) \overline{K}_6(b_1, d_1, g_1) \overline{K}_7(b_2, d_2, g_2) \overline{K}_8(b_3, d_3, g_3) \overline{K}_9(b_4, d_4, g_4)}
\]

(20)

where

\[B_5 = b_5, \quad s > 6\]

\[B_6 = 120 + 600a_4 - 25a_3^2, \quad B_6 = 720 + 5400a_4 - 450a_3^2\]

and \(b, d, g\) are given in (14). The two expressions arise according as \(s\) is even or odd in (17). We use the notation \(\overline{K}_s(p, q, \ldots)\) to indicate a symmetric determinant of order \(s\) with elements

\[p_j, \quad (j = k, k + 2, k + 4, \ldots), \text{ in the diagonal through (1, 1)},\]

\[q_j, \quad (j = l, l + 2, l + 4, \ldots), \text{ in the diagonal through (2, 1)}, \text{ and so on,}\]

all unspecified elements being zero. Thus, for example,

\[
\overline{K}_3(b_0, d_0, g_0) = \begin{vmatrix}
    b_0 & d_0 & g_0 \\
    d_0 & b_2 & d_2 \\
    g_0 & d_2 & b_4
\end{vmatrix}, \quad \overline{K}_2(B_5, d_5, g_5) = \begin{vmatrix}
    B_5 & d_5 \\
    d_5 & b_7
\end{vmatrix}.
\]

(19) and (20) give upper bounds for \(E_f\) and decreasing sequences and taking \(s = 3, 4, 5\) we have the values in Table 3.

**Table 3. Successive values for \(E_f\) in the joint estimation of four parameters for Type A in the case \(a_3 = 0\)**

<table>
<thead>
<tr>
<th>(a_4)</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st approx.</td>
<td>0.99</td>
<td>0.97</td>
<td>0.92</td>
<td>0.88</td>
<td>0.84</td>
<td>0.81</td>
<td>0.78</td>
<td>0.75</td>
<td>0.73</td>
<td>0.71</td>
<td>0.69</td>
<td>0.63</td>
<td>0.59</td>
</tr>
<tr>
<td>2nd approx.</td>
<td>0.97</td>
<td>0.92</td>
<td>0.81</td>
<td>0.72</td>
<td>0.65</td>
<td>0.59</td>
<td>0.55</td>
<td>0.51</td>
<td>0.48</td>
<td>0.45</td>
<td>0.43</td>
<td>0.34</td>
<td>0.28</td>
</tr>
<tr>
<td>3rd approx.</td>
<td>0.97</td>
<td>0.92</td>
<td>0.81</td>
<td>0.72</td>
<td>0.65</td>
<td>0.59</td>
<td>0.54</td>
<td>0.50</td>
<td>0.46</td>
<td>0.43</td>
<td>0.41</td>
<td>0.31</td>
<td>0.25</td>
</tr>
<tr>
<td>4th approx.</td>
<td>0.97</td>
<td>0.92</td>
<td>0.81</td>
<td>0.71</td>
<td>0.63</td>
<td>0.57</td>
<td>0.52</td>
<td>0.47</td>
<td>0.43</td>
<td>0.40</td>
<td>0.37</td>
<td>0.26</td>
<td>0.20</td>
</tr>
<tr>
<td>5th approx.</td>
<td>0.97</td>
<td>0.91</td>
<td>0.80</td>
<td>0.71</td>
<td>0.63</td>
<td>0.57</td>
<td>0.52</td>
<td>0.47</td>
<td>0.43</td>
<td>0.40</td>
<td>0.37</td>
<td>0.26</td>
<td>0.19</td>
</tr>
<tr>
<td>6th approx.</td>
<td>0.97</td>
<td>0.91</td>
<td>0.80</td>
<td>0.71</td>
<td>0.63</td>
<td>0.57</td>
<td>0.52</td>
<td>0.47</td>
<td>0.43</td>
<td>0.40</td>
<td>0.37</td>
<td>0.25</td>
<td>0.17</td>
</tr>
</tbody>
</table>

The values in Table 3 are given correct to two figures; the first column of values correct to four figures reads: 0.9908, 0.9703, 0.9695, 0.9688, 0.9685 and 0.9680. The capricious behaviour of the successive approximations is noteworthy, so that although two successive terms may be nearly equal this provides little basis for supposing the remainder negligible. From Tables 2 and 3 it can now be seen that \(E_f < 80\) % for \(0.2 < a_4 < 4\cdot0, a_3^2 > 0.1\), and since
Efficiency of the method of moments

If \( E_f < 0.7029 \) for \( a_4 = 0.2, a_3^2 = 0.05 \), it would appear that the 80 \% efficiency contour is not in the region \( a_2 \geq 0.05, 0.2 < a_4 < 4.0 \). This may be compared with the case for Pearson’s Type IV distribution, except that the general case is one of some complexity. Taking Type IV in the form

\[
f(X) dX = y_0 \{1 + x^3(-16r+31) e^{-\nu \tan^{-1} x}\} dx,
\]

where

\[
x = \frac{X - m}{c}, \quad y_0 = \frac{1}{cF(r, \nu)}, \quad F(r, \nu) = e^{-\nu \tan^{-1} x} \int_0^\pi \sin \theta e^{\nu \theta} d\theta,
\]

the moments following the well-known recurrence relation

\[
\mu_s = -\frac{2 \nu c (s-1)}{r(r+1)} \mu_{s-1} + \frac{(s-1) c^2 (\nu^2 + r^2)}{r^2(r+1)} \mu_{s-2},
\]

(22)

with \( \mu_0 = 1, \mu_1 = 0, s \geq 2 \); we have for the other moments up to \( \mu_8 \):

\[
\begin{align*}
\mu_2 &= c^2 R^2 / r(r^2), \\
\mu_3 &= -4 c^2 \nu R^2 / r^2 r^3, \\
\mu_4 &= 3 c^2 R^2 ((r-2) R^2 + 8 \nu^2) / r^3 r^4, \\
\mu_5 &= -8 c^4 R^2 ((5r-12) R^2 + 24 \nu^2) / r^4 r^5, \\
\mu_6 &= 5 c^6 R^2 (3(r-2) R^4 + 8(13r-36) R^2 \nu^2 + 384 \nu^4) / r^5 r^6, \\
\mu_7 &= -12 c^6 \nu R^2 ((35r^2 - 238r + 360) R^4 + 8(77r^2 - 240) R^2 \nu^2 + 1920 \nu^4) / r^6 r^7, \\
\mu_8 &= 7 c^8 R^2 (15(r-2) R^4 + 8(170r^2 - 1259r + 2160) R^2 \nu^2 + 576(29r^2 - 240) R^2 \nu^2 + 1920 \nu^4) / r^7 r^8,
\end{align*}
\]

(23)

where \( \nu = \nu(r-1) \ldots (r-8+1) \) and \( R^2 = \nu^2 + \nu^2 \).

The information matrix, given by Fisher (1921) is

\[
I = \begin{bmatrix}
(r+1)(r+2)(r+4) A & -(r+1)(r+2) \nu A & (r+1)(r+2) B & -(r+1) \nu B \\
-(r+1)(r+2) \nu A & (r+1)(2r+8+\nu^2) A & -(r+1) \nu B & (r+2+\nu^2) B \\
(r+1)(r+2) B & -(r+1) \nu B & \frac{\partial^2}{\partial \nu^2} \log F & \frac{\partial^2}{\partial \nu \partial \nu} \log F \\
-(r+1) \nu B & (r+2+\nu^2) B & \frac{\partial^2}{\partial \nu \partial \nu} \log F & \frac{\partial^2}{\partial \nu^2} \log F
\end{bmatrix},
\]

(24)

where

\[
A^{-1} = c^2 \{(r+4)^2 + \nu^2\}, \quad B^{-1} = c\{(r+2)^2 + \nu^2\}
\]

It is evident from the form of the higher moments and the derivatives in \( I \) that \( E_f \) for joint estimation would be rather complicated. However, a comparison with Type A may be made in the special case \( \nu = 0 \). This situation may arise when we set out to fit a Type IV curve to a sample from a population in which \( \nu = 0 \), and where this fact is not known before the sample is drawn. The case should be distinguished from the one of simultaneous estimation of three parameters in a Type VII distribution which we consider in §5.

Using (23) we find

\[
N^4 \left[ \text{cov}(m_c, m_r) \right] = \frac{288 c^2 (r-2)}{(r-1)^5 (r-3)^5 (r-5)^3 (r-7)} (25)
\]

\[
\left. \frac{\partial^2 (\mu_1, \mu_2, \mu_3, \mu_4)}{\partial (m, c, \nu, \nu)} \right|_{\nu=0} = \frac{48 c^8}{r(r-1)^4 (r-2) (r-3)^2}. (26)
\]

From (24) we have

\[
| I | = \frac{PQ}{4(r+2)^4 (r+4)^2 c^4}, (27)
\]
where

\[ P = (r+2)^3 \frac{\phi(\frac{1}{2}r)}{2(r+1)(r+4)}, \]
\[ Q = (r+1)^2(r+2)^2 \{ \phi[\frac{1}{2}(r-1)] - \phi(\frac{1}{2}r) \} - 2(r+1)(r+4), \]
\[ \phi(x) = \frac{d^2}{dx^2} \log x! = \psi'(x+1) \quad \text{in the notation of the trigamma function.} \]

Hence using (25)–(27) we have for the efficiency of the method of moments in the joint estimation of the four parameters of Type IV

\[ E_f = \frac{32(r-3)(r-5)(r-7)(r+2)^4(r+4)^2}{r^5(r-1)^3(r-2)^3 PQ} \quad \text{when } \nu = 0. \]  

(28)

Table 4. Comparison of \( E_f \) for Type A and Type IV in the special case \( \gamma_2 = 0 \)

<table>
<thead>
<tr>
<th>( \gamma_2 )</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type IV</td>
<td>0.99</td>
<td>0.98</td>
<td>0.92</td>
<td>0.84</td>
<td>0.74</td>
<td>0.63</td>
<td>0.53</td>
<td>0.43</td>
<td>0.33</td>
<td>0.25</td>
<td>0.18</td>
<td>0.25</td>
</tr>
<tr>
<td>Type A</td>
<td>0.97</td>
<td>0.91</td>
<td>0.80</td>
<td>0.71</td>
<td>0.63</td>
<td>0.57</td>
<td>0.52</td>
<td>0.47</td>
<td>0.43</td>
<td>0.40</td>
<td>0.37</td>
<td>0.25</td>
</tr>
</tbody>
</table>

For comparison with the corresponding case for Type A we must have \( r = 3 + 6/a_4 \) and the values of \( E_f \) are shown in Table 4. Since the entry for Type A is an upper bound we may conclude that higher efficiency occurs whenever \( 0 < \gamma_2 < 0.6 \) with moment estimation for Type IV than for Type A, and for \( \gamma_2 > 0.6 \) there may be cases where the inefficiency for Type A is not as high as for Type IV. Type IV has \( E_f < 80\% \) for \( 0.35 \leq \gamma_2 \leq 1.5 \) while, for Type A, \( E_f \) is certainly below 80\% for \( 0.2 < \gamma_2 < 4.0 \).

5. THE SYMMETRICAL TYPE A AND COMPARISON WITH PEARSON’S TYPE VII

We now consider the case of Type A when \( a_3 = 0 \). Writing

\[ P_x dx = \left( 1 + \frac{a_4}{4!} H_4(x) \right) g(x) dx, \quad x = \frac{X - m}{\sigma}, \]  

(29)

we have

\[ \frac{\partial P_x}{\partial m} = \left( H_1(x) + \frac{a_4}{4!} H_5(x) \right) \frac{g(x)}{\sigma}, \]
\[ \frac{\partial P_x}{\partial \sigma} = \left( H_3(x) + \frac{a_4}{4!} [4H_4(x) + H_6(x)] \right) \frac{g(x)}{\sigma^2}, \]
\[ \frac{\partial P_x}{\partial a_4} = \frac{H_4(x) g(x)}{4! \sigma}. \]  

(30)

It is evident from the fact that the first of the equations in (30) is an odd function of \( x \) that for maximum likelihood estimators \( \hat{m} \) is uncorrelated with \( \hat{\sigma} \) and \( \hat{a}_4 \). The efficiency of the method of moments in this case then can be considered as follows:

**Case A. Efficiency of the mean \( \bar{m} \) (or \( m_1 \)).**

**Case B.** (i) Efficiency of \( \bar{\sigma} \) given \( a_4 \), where \( \bar{\sigma}^2 = m_2 \). (ii) Efficiency of \( \bar{a}_4 \) given \( \sigma \).

**Case C. Efficiency of \( \bar{\sigma} \) and \( \bar{a}_4 \), where \( \bar{\sigma}^2 = m_2 \) and \( a_4 + 3 = \frac{m_4}{m_2^2} \).

* The \( \phi \) notation is due to Pairman, *Tracts for Computers*, No. 1; the psi function notation is used by H. T. Davis, *Tables of the Higher Mathematical Functions*, Vol. II, Principia Press, Inc.
Here we have used \( m_r \) for the \( r \)th moment of the sample of \( N \) about its mean \( m \). Cases A and C are the most important in practice. Case B (i) may have interest for it corresponds to the efficiency of the estimation of \( \hat{\sigma} \) when \( r \) is known for the Type VII curve \((\nu = 0 \text{ in } (21))\). Jeffreys (1939b, p. 708) has indicated that with data of observational errors, when the sample size \( N \) is less than 500, the estimate of the index \((r + 2)/2\) is unreliable and a value 4 may be used in such cases. The parallel case with the symmetrical Type A would be to take \( a_4 = 2 \). Case B (ii) is not likely to occur in practice.

**Case A. Efficiency of the mean**

We have

\[
\frac{1}{N \text{ var } \hat{m}} = E \left( \frac{1}{P_x \text{ d}m} \right)^2 \quad \text{and} \quad N \text{ var } \hat{m} = \sigma^2,
\]

so that in (2) we take \( A_1(x) g(x) = \partial P_x/\partial m \) and following our previous method find

\[
E_f(\hat{m}) \leq \frac{\bar{K}_s(b_1, d_1, g_1)}{\bar{K}_s-1(b_3, D_3, g_3)}, \quad s \geq 1,
\]

(31)

with

\[
B_3 = 120 + 600a_4 - 25a_4^2, \quad B_s = b_s, \quad s = 5;
\]

\[
D_3 = 60a_4 - 5a_4^2, \quad D_s = d_s, \quad s = 3;
\]

using the notation of (14), (19) and (20). But we also have

\[
\frac{1}{E_f(\hat{m})} = \int_{-\infty}^{\infty} \frac{\{H_1(x) + (a_4/4!)H_3(x)\}^2 g(x) dx}{1 + (a_4/4!)H_4(x)}
\]

\[
= 1 + \frac{2a_4}{3} - \frac{a_4(4-a_4)}{6} \int_{-\infty}^{\infty} \frac{H_4^2(x) g(x) dx}{1 + (a_4/4!)H_4(x)},
\]

which leads to

\[
E_f(\hat{m}) \geq \left(1 + \frac{2a_4}{3} - \frac{a_4(4-a_4)}{6} \frac{\bar{K}_s^{-1}(b_3, D_3, g_3)}{\bar{K}_s(b_1, d_1, g_1)} \right)^{-1},
\]

(32)

in the same notation as (31). From (31) and (32) upper and lower bounds may be found for \( E_f(\hat{m}) \). If we wish to make a comparison with Pearson’s Type VII we find from (21), (23) and (24) with \( \nu = 0 \), \( E_f(\hat{m}) = (r - 1)(r + 4)/[(r + 1)(r + 2)] \). With \( s = 5 \) in (31) and (32) we thus find the values in Table 5. Interpolating in the table it is seen that \( E_f(\hat{m}) \geq 80 \% \) for \( 0 \leq \gamma_2 \leq 1.5 \), and \( E_f(\hat{m}) \leq 50 \% \) for \( 3.0 \leq \gamma_2 < 4.0 \). For Type VII \( E_f(\hat{m}) > 80 \% \) for \( 0 \leq \gamma_2 < 6.0 \) and is certainly higher than is the case for Type A. It may be noted that for Type A, since \( \text{var (median)} = 1/(4N^2 \bar{a}) \), where \( \bar{a} \) is the median ordinate, \( \{\text{var (median)}/\text{var } \hat{m}\} = 32\pi/2S = \pi(8 + a_4)^2 \) which is less than unity for \( a_4 > 2 \); thus for \( a_4 = 3, 0.55 < E_f \) (median) < 0.66. Hence for \( a_4 > 2 \) the median is a slightly better estimator than \( \hat{m} \).

<table>
<thead>
<tr>
<th>( \gamma_2 )</th>
<th>0.05</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i)</td>
<td>0.9996</td>
<td>0.9994</td>
<td>0.9979</td>
<td>0.9580</td>
<td>0.9299</td>
<td>0.8965</td>
<td>0.7951</td>
<td>0.6795</td>
<td>0.4527</td>
</tr>
<tr>
<td>(ii)</td>
<td>0.9996</td>
<td>0.9994</td>
<td>0.9801</td>
<td>0.9589</td>
<td>0.9324</td>
<td>0.9018</td>
<td>0.8141</td>
<td>0.7209</td>
<td>0.5493</td>
</tr>
<tr>
<td><strong>Type VII</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.9996</td>
<td>0.9950</td>
<td>0.9842</td>
<td>0.9714</td>
<td>0.9583</td>
<td>0.9455</td>
<td>0.9167</td>
<td>0.8929</td>
<td>0.8571</td>
</tr>
</tbody>
</table>

* (i) and (ii) refer to lower and upper bounds respectively.
Case B (i). Efficiency of $\bar{c}$ given $a_4$

From (30) and (2) we have

$$\frac{\sigma^2}{N \var{\bar{c}}} \geq \frac{4K_{s-2}(B_4, D_4, g_4)}{K_s(b_0, d_0, g_0)},$$

where

$$B_4 = b_4 - 4a_4d_2 + 4a_4^2b_2 - g_2^2 = 24 + 72a_4 - 25a_4^2 + 4a_4^3,$$
$$D_4 = d_4 - 2a_4g_2 - 15a_4d_3 + 30a_4^2b_2 = 480a_4 - 1204a_4^2 + 304a_4^3,$$
$$B_s = b_s, \quad s > 6, \quad D_s = d_s, \quad s > 4;$$

and the other symbols are those given in (14) and (20), so that since $4N \var{\bar{c}} = (a_4 + 2)\sigma^2$, we have

$$E_j(\bar{c} | a_4) \leq \frac{K_s(b_0, d_0, g_0)}{K_s(b_0, d_0, g_0) K_{s-2}(B_4, D_4, g_4)}. \tag{33}$$

A lower bound may also be found, for from

$$\int_{-\infty}^{\infty} \frac{(H_2(x) + (a_4/3!)H_4(x) + (a_4/4!)H_6(x))^2}{1 + (a_4/4!)H_4(x)} g(x) dx$$

there follows

$$E_j(\bar{c} | a_4) \geq \frac{4}{2 + a_4} \left( 2 + a_4 - \frac{a_4(4 - a_4)}{6} \right)^{-1} \frac{K_{s-1}(B_2, D_2, g_2)}{K_s(b_0, d_0, g_0)} \tag{34}$$

where

$$B_2 = b_2 + 4b_0 = 6 + a_4, \quad B_s = b_s, \quad s > 2, \quad D_2 = d_2 - 2g_0 = 6a_4, \quad D_s = d_s, \quad s > 2,$$

and the other symbols are those used in (33).

Table 6. $E_j(\bar{c} | a_4)$ for Type A and the corresponding case for Type VII

<table>
<thead>
<tr>
<th>$\gamma_2$</th>
<th>0.05</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A (i)*</td>
<td>0.9963</td>
<td>0.9920</td>
<td>0.9747</td>
<td>0.9500</td>
<td>0.9173</td>
<td>0.8759</td>
<td>0.7369</td>
<td>0.5684</td>
<td>0.2747</td>
</tr>
<tr>
<td>(ii)</td>
<td>0.9993</td>
<td>0.9920</td>
<td>0.9752</td>
<td>0.9525</td>
<td>0.9243</td>
<td>0.8908</td>
<td>0.7880</td>
<td>0.6704</td>
<td>0.4514</td>
</tr>
<tr>
<td>Type VII</td>
<td>0.9992</td>
<td>0.9893</td>
<td>0.9649</td>
<td>0.9341</td>
<td>0.9006</td>
<td>0.8667</td>
<td>0.7857</td>
<td>0.7143</td>
<td>0.6000</td>
</tr>
</tbody>
</table>

* (i) and (ii) refer to lower and upper bounds respectively.

For the corresponding problem for Type VII when $r$ is given, it follows since $\mu_2 = c^2/(r - 1)$ that $N \var{\bar{c}} = c^2r/[2(r - 3)]$, and from (24) $N \var{\bar{c}} = c^2(r + 4)/[2(r + 1)]$. Hence for the efficiency of $\bar{c}$ given $r$ we have $E_j(\bar{c} | r) = (r - 3)(r + 4)/(r(r + 1))$. The comparison with Type A is given in Table 6, taking $s = 6$ in (33) and (34). It will be seen that $E_j(\bar{c} | a_4) > 80\%$ for $0 \leq \gamma_2 \leq 1.0$ and $< 60\%$ for $2.0 < \gamma_2 < 4.0$ approximately. For $0 < \gamma_2 < 1.5$, $E_j$ is higher for Type A than for Type VII, but the difference is not marked. It may be noted that in the range $2.0 < \gamma_2 < 4.0$ the efficiency is in the region of $70\%$ or below in both cases and is unsatisfactory; for Type VII, however, it has been shown by Sichel (1949) that the method of frequency-moments has a highly satisfactory efficiency.
Case B (ii). Efficiency of \( \bar{a}_4 \) given \( \sigma \)

For the maximum likelihood estimator \( \hat{\alpha}_4 \) we have

\[
\frac{1}{N \text{var} \hat{\alpha}_4} = \frac{1}{\alpha_4^2} \left\{ \int_{-\infty}^{\infty} \frac{g(x) \, dx}{1 + (\alpha_4/4!) H_4(x)} - 1 \right\},
\]

and also

\[
\frac{1}{N \text{var} \hat{\alpha}_4} = \frac{1}{6\alpha_4(4-\alpha_4)} \left\{ 6 - \int_{-\infty}^{\infty} \frac{[H_4(x) - \frac{1}{2}]^2 \, dx}{1 + (\alpha_4/4!) H_4(x)} g(x) \, dx \right\},
\]

where

\[
B'_2 = b_2 + b_0 = 3 + \alpha_4, \quad B'_s = b_s, \quad s > 2;
\]

\[
D'_2 = d_2 + g_0 = 9\alpha_4, \quad D'_s = d_s, \quad s > 2;
\]

and \( b, d, g \) are otherwise given in (14).

For the estimation of \( \alpha_4 \) by moments we may use

(a) \( \bar{a}_4 + 3 = m_4/\sigma^4 \),  
(b) \( \bar{a}_4' = (m_4 - 6m_2 \sigma^2 + 3\sigma^4)/\sigma^4 \),  
(c) \( \bar{a}_4'' + 3 = m_4/m_2^2 \),

in which \( m_r = \Sigma(X - \bar{X})^r/N, \ r = 2, 4 \), and in large sampling we have approximately

\[
N \text{var} \bar{a}_4 = 9\beta + 20\alpha_4 - \alpha_4^2;
\]

\[
N \text{var} \bar{a}_4' = 24 + 72\alpha_4 - \alpha_4^2;
\]

\[
N \text{var} \bar{a}_4'' = 24 + 72\alpha_4 - 25\alpha_4^2 + 4\alpha_4^3.
\]

It is evident then that \( \bar{a}_4 \) is a poor estimator and indeed \( E_f(\bar{a}_4 \mid \sigma) \rightarrow \frac{1}{4} \) for small \( \alpha_4 \). The efficiency for the other two is given in Table 7, using \( s = 4 \) in (36) and (37).

**Table 7.** \( E_f(\bar{a}_4' \mid \sigma) \) and \( E_f(\bar{a}_4'' \mid \sigma) \) for Type A

<table>
<thead>
<tr>
<th>( \alpha_4 )</th>
<th>0-05</th>
<th>0-2</th>
<th>0-4</th>
<th>0-6</th>
<th>0-8</th>
<th>1-0</th>
<th>1-5</th>
<th>2-0</th>
<th>3-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_f(\bar{a}_4' \mid \sigma) ) (i)*</td>
<td>0-9753</td>
<td>0-8358</td>
<td>0-6722</td>
<td>0-5511</td>
<td>0-4602</td>
<td>0-3895</td>
<td>0-2660</td>
<td>0-1843</td>
<td>0-0771</td>
</tr>
<tr>
<td>(ii)</td>
<td>0-9759</td>
<td>0-8456</td>
<td>0-6880</td>
<td>0-5905</td>
<td>0-5101</td>
<td>0-4475</td>
<td>0-3373</td>
<td>0-2835</td>
<td>0-1660</td>
</tr>
<tr>
<td>( E_f(\bar{a}_4'' \mid \sigma) ) (i)</td>
<td>0-9755</td>
<td>0-8565</td>
<td>0-7213</td>
<td>0-6227</td>
<td>0-5508</td>
<td>0-4934</td>
<td>0-3867</td>
<td>0-3023</td>
<td>0-1448</td>
</tr>
<tr>
<td>(ii)</td>
<td>0-9761</td>
<td>0-8665</td>
<td>0-7490</td>
<td>0-6882</td>
<td>0-6105</td>
<td>0-5686</td>
<td>0-4904</td>
<td>0-4321</td>
<td>0-3118</td>
</tr>
</tbody>
</table>

* (i) and (ii) refer to lower and upper bounds respectively.

Thus in either case the efficiency is below 80% for \( \alpha_4 > 0-3 \) approximately, and although \( \bar{a}_4'' \) is a better estimator than \( \bar{a}_4' \) it cannot be regarded as satisfactory for a departure from normality exceeding \( \alpha_4 \) or \( \gamma_2 = 0-4 \). In other words, in large sampling from a Type A population with significant leptokurtosis, for which the parameters of scale and location are known (or the scale parameter alone is known) the method of moments should in general only be used as a first approximation.

Case C. The efficiency of the joint estimators \( \bar{\sigma} \) and \( \bar{a}_4 \)

Without going into details it is found using the previous methods that

\[
E_f(\bar{\sigma}, \bar{a}_4) \leq \frac{K_{s+2}(b_0, d_0, g_0)}{K_s(b_0, d_0, g_0) K_s(B_0, d_s, g_s)},
\]

with

\[
B_s = b_s, \ s > 6,
\]
the other symbols being given in (14) and (20). For the corresponding problem with Type VII, using (23) and (24),

\[ E_f(\bar{c}, \bar{r}) = \frac{6(r + 4)(r + 2)^3(r + 1)(r - 5)^3(r - 7)}{r^2(r - 1)^3(r - 2)(r - 3)Q}, \]

(39)

where \( Q \) is given in (27), and \( r = 3 + 6/\alpha_4 \). Taking \( s = 3 \) in (38) we find the values given in Table 8. For Type A then \( E_f(\bar{\sigma}, \bar{a}_4) < 80 \% \) for \( 0 < \alpha_4 < 4.0 \), and less than \( 50 \% \) for \( 1.5 < \alpha_4 < 4.0 \). On the other hand, the efficiency for Type VII is higher than for Type A for \( 0 < \alpha_4 < 0.5 \) and perhaps a little outside this range, since the Type A entries are upper bounds. Thus for the range \( 0 < \gamma_2 < 0.5 \) Type VII is nearer Fisher’s system of maximum moment efficiency (the exponential of a quartic) than is Type A.

Table 8. \( E_f(\bar{\sigma}, \bar{a}_4) \) for Type A and \( E_f(\bar{c}, \bar{r}) \) for Type VII

<table>
<thead>
<tr>
<th>( \alpha_4 )</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A</td>
<td>0.978</td>
<td>0.942</td>
<td>0.870</td>
<td>0.809</td>
<td>0.758</td>
<td>0.715</td>
<td>0.677</td>
<td>0.643</td>
<td>0.613</td>
<td>0.586</td>
<td>0.558</td>
<td>0.441</td>
<td>0.337</td>
<td>0.163</td>
</tr>
<tr>
<td>Type VII</td>
<td>0.996</td>
<td>0.986</td>
<td>0.947</td>
<td>0.888</td>
<td>0.815</td>
<td>0.731</td>
<td>0.642</td>
<td>0.550</td>
<td>0.460</td>
<td>0.372</td>
<td>0.291</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It may be worth noting that an upper bound may also be found for the efficiency, in joint estimation, of either \( \bar{\sigma} \) or \( \bar{a}_4 \). In the case of \( \bar{a}_4 \) we have

\[ N \text{ var } \bar{a}_4 = E\left(\frac{1}{P \partial \sigma}\right)^2 \Delta, \]

where

\[ \Delta = \begin{vmatrix} E\left(\frac{1}{P \partial \sigma}\right)^2 & E\left(\frac{1}{P^2 \partial \sigma} \partial \alpha_4\right) \\ E\left(\frac{1}{P^2 \partial a_4 \partial \sigma}\right) & E\left(\frac{1}{P \partial a_4}\right) \end{vmatrix}, \]

so that we find

\[ E\left(\frac{1}{P \partial \sigma}\right)^2 = \lim_{s \to \infty} \frac{4\overline{K}_{s-2}(B_4, D_4, g_4)}{\sigma^2 \overline{K}_s(b_0, d_0, g_0)}, \]

(40)

and

\[ \Delta = \lim_{s \to \infty} \frac{4\overline{K}_{s-2}(B_4, d_4, g_4)}{\sigma^2 \overline{K}_s(b_0, d_0, g_0)}, \]

(41)

in the notation of (14) and (33). Moreover, since \( \bar{a}_4 + 3 = m_4 / m_2 \) so that

\[ N \text{ var } \bar{a}_4 = 24 + 72\alpha_4 - 25a_4^2 + 4a_4^3 - B_4 \]

we find

\[ E_f(\bar{a}_4) = \lim_{s \to \infty} \left\{ \overline{K}_{s-2}(B_4, D_4, g_4) \right\} \lim_{s \to \infty} \left\{ \overline{K}_s(b_0, d_0, g_0) \right\} \]

(42)

\[ \frac{4\overline{K}_{s-2}(B_4, D_4, g_4)}{\sigma^2 \overline{K}_s(b_0, d_0, g_0)}, \]

(43)

since neither of the limits in (42) is zero. Further (43) may be written

\[
\begin{bmatrix}
B_4 & D_4 & g_4 & \ldots \\
D_4 & B_6 & d_6 & \ldots \\
g_4 & d_6 & b_5 & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 0 & \ldots \\
1 & B_4 & D_4 & g_4 & \ldots \\
0 & D_4 & B_6 & d_6 & \ldots \\
0 & g_4 & d_6 & b_5 & \ldots \\
\vdots & \vdots & \vdots & \ddots & \ddots
\end{bmatrix}_{s+1}
\]
and so is a Schwein's determinantal ratio inverted; since the $K_s(B_4, D_4, g_4)$ are essentially positive (43) therefore gives a decreasing sequence. As an illustration of this feature the values of $E_r(\bar{a}_4)$ for $a_4 = 1$ and $s = 1, 2, 3$ and 4 are 1-0000, 0-6560, 0-6165 and 0-6158. When $s = 4$ the values of $E_r(\bar{a}_4)$ are given in Table 9. The last row gives the corresponding value for Type VII found from (23) and (24) as

$$E_r(\bar{a}_4) = \frac{6(r + 2)^2(r + 1)^2(r - 5)(r - 7)}{r(r - 1)^2(r - 3)(r^2 - r + 18)Q}.\]

For Type A it is seen that $E_r(\bar{a}_4) < 80\%$ for $0.4 < \gamma_2 < 4.0$ and less than $50\%$ for $2.0 < \gamma_2 < 4.0$ approximately. $E_r(\bar{a}_4)$ for Type VII is higher in $0 < \gamma_2 < 0.4$.

<table>
<thead>
<tr>
<th>$a_4$</th>
<th>0.05</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A</td>
<td>0.9782</td>
<td>0.8714</td>
<td>0.7665</td>
<td>0.6676</td>
<td>0.6503</td>
<td>0.6158</td>
<td>0.5556</td>
<td>0.5026</td>
<td>0.3625</td>
</tr>
<tr>
<td>Type VII</td>
<td>0.9964</td>
<td>0.9474</td>
<td>0.8181</td>
<td>0.6519</td>
<td>0.4777</td>
<td>0.3130</td>
<td>0.0000</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

6. THE MAXIMUM LIKELIHOOD EQUATIONS

Considering the symmetrical Type A distribution for which

$$P_x dx = \left\{1 + \frac{a_4}{4!} H_4(x)\right\} g(x) dx, \quad x = \frac{X - m}{\sigma},$$

and writing for the likelihood function of a sample $(X_1, X_2, \ldots, X_N)$

$$L = \prod_{j=1}^{N} P_x,$$

we have, after simplification,

$$\frac{\partial \log L}{\partial m} = \Sigma \left\{H_4(x) - \frac{a_4 H_4(x)}{3! \Omega(x)}\right\},$$

$$\frac{\partial \log L}{\partial \sigma} = \frac{1}{\sigma} \Sigma \left\{H_2(x) - 4 + \frac{8 - a_4 H_4(x)}{2 \Omega(x)}\right\},$$

$$\frac{\partial \log L}{\partial a_4} = \frac{1}{a_4} \Sigma \left\{1 - \frac{1}{\Omega(x)}\right\},$$

in which $\Omega(x) = 1 + \frac{a_4}{4!} H_4(x)$ and $\Sigma$ denotes summation over the sample. For maximum likelihood estimates we require solutions of these expressions equated to zero. Using $\bar{m}, \bar{\sigma}, \bar{a}_4,$ the moment estimates as first approximations, we may find improved estimates $\hat{m}, \hat{\sigma}, \hat{a}_4$ from

$$\frac{\partial \log L}{\partial \hat{m}} + (\hat{m} - \bar{m}) \frac{\partial^2 \log L}{\partial \hat{m}^2} = 0,$$

$$\frac{\partial \log L}{\partial \hat{\sigma}} + (\hat{\sigma} - \bar{\sigma}) \frac{\partial^2 \log L}{\partial \hat{\sigma}^2} + (\hat{a}_4 - \bar{a}_4) \frac{\partial^2 \log L}{\partial \hat{\sigma} \partial \hat{a}_4} = 0,$$

$$\frac{\partial \log L}{\partial \hat{a}_4} + (\hat{\sigma} - \bar{\sigma}) \frac{\partial^2 \log L}{\partial \hat{a}_4 \partial \hat{\sigma}} + (\hat{a}_4 - \bar{a}_4) \frac{\partial^2 \log L}{\partial \hat{a}_4^2} = 0.$$
where
\[
\frac{\partial^2 \log L}{\partial \bar{m}^2} = \frac{1}{\sigma} \sum \left\{ \frac{1 + \bar{a}_4(4 + H_5(x)) - \bar{a}_4(4 - \bar{a}_4)(1 + H_5(x))}{3! \Omega(x)} \right\},
\]
\[
\frac{\partial^2 \log L}{\partial \bar{a}_4 \partial \sigma} = \frac{1}{\sigma} \sum \left\{ \frac{2 - 3H_5(x) + (24 - 10\bar{a}_4 - \bar{a}_4^2 H_5(x)) - (4 - \bar{a}_4)(8 - \bar{a}_4 - 2\bar{a}_4 H_5(x))}{2\Omega(x)} \right\},
\]
\[
\frac{\partial^2 \log L}{\partial \bar{a}_4^2} = \frac{1}{\bar{a}_4^2} \sum \left\{ \frac{1}{\Omega(x)} - \frac{(8 - \bar{a}_4 H_5(x))}{\Omega^2(x)} \right\},
\]
and 
\[
\Omega(x) = \Omega(x, \bar{m}, \sigma),
\]
and in which cross-product terms involving \(m\) such as \(\frac{\partial^2 \log L}{\partial \bar{m} \partial \sigma}\) are ignored since their expectations are zero. In the case when \(a_4 \leq 1.5\) approximately we see from Table 5 that \(\bar{m}\) is an estimate of \(m\) with 80\% or more efficiency in large sampling. Hence when \(a_4 \leq 1.5\) improved estimates of \(\sigma\) and \(a_4\) may be found from \(\bar{m}\) and \(\bar{a}_4\) using (48) and (49) only. The case when \(a_4 > 1.5\) is likely to be more involved. For this we could use (47) to improve \(\bar{m}\) (or the median estimate of \(m\)) and (48)–(49) to improve \(\bar{\sigma}\) and \(\bar{a}_4\). The process could be repeated if necessary.

The parallel problem with Pearson's Type VII has been discussed by Jeffreys (1939b; 1948, pp. 184–7), who uses a different form from (21) with \(v = 0\), involving orthogonal parameters \(\theta_j, j = 1, 2, 3\), in the sense that the expectation of such terms as \(\frac{\partial^2 \log L}{\partial \theta_j \partial \theta_3}\) is zero. A different approach to the problem has been described by Sichel (1947, 1949), whose method of frequency-moments is fairly efficient in the range where the method of moments is very inefficient. It would seem that if we are to find efficient estimates of the parameters involved, then the use of Type VII may be less laborious than Type A. It is clear that the solution of the likelihood equations for an asymmetrical Type A distribution would be no light undertaking.

7. Summary

The results found in this paper may be conveniently summarized as follows:

(a) For the Type A distribution of four parameters (given by (5)) the joint efficiency of the method of moments is less than 80\% for \(0.2 < \gamma_2 < 4.0, \gamma_4^2 \geq 0.05\); and falls rapidly with increasing skewness and leptokurtosis, and also in the vicinity of critical values of the parameters \(a_3, a_4\) which arise from the condition that the frequency must be positive. For the corresponding problem for Pearson's Type IV we have shown that the efficiency \(E_f\) is below 80\% for \(0.35 < \gamma_2 < 1.5, \gamma_4^2 = 0\) and above 80\% for \(0 < \gamma_2 < 0.35, \gamma_4^2 = 0\).

(b) For the symmetrical Type A
\[
P_x = \left[ 1 + \frac{a_4}{4!} H_4(x) \right] \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad x = \frac{X - m}{\sigma},
\]
and Pearson's Type VII
\[
f(X) = \frac{(\frac{1}{2}r)!}{(\frac{1}{2}r - 1)!} \frac{e^r}{\{c^2 + (X - m)^2\}^{\frac{1}{2}r + 2}}.
\]
we have:

<table>
<thead>
<tr>
<th>Parameters estimated</th>
<th>$\bar{m}$</th>
<th>$\sigma$ given $a_4$</th>
<th>$\bar{a}_4$ given $\sigma$</th>
<th>$\sigma$ and $\bar{a}_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_1 &lt; 80$ % Type A</td>
<td>$1.5 &lt; \gamma_2 &lt; 4.0$</td>
<td>$1.35 &lt; \gamma_2 &lt; 4.0$</td>
<td>$0.3 &lt; \gamma_2 &lt; 4.0$</td>
<td>$0.3 &lt; \gamma_2 &lt; 4.0$</td>
</tr>
<tr>
<td>Type VII</td>
<td>$\gamma_2 &gt; 6.0$</td>
<td>$\gamma_2 &gt; 1.4$</td>
<td>$-$</td>
<td>$-$</td>
</tr>
<tr>
<td>$E_1 &gt; 80$ % Type A</td>
<td>$0 \leq \gamma_2 \leq 1.5$</td>
<td>$0 \leq \gamma_2 \leq 1.35$</td>
<td>$0 \leq \gamma_2 &lt; 0.3$</td>
<td>$-$</td>
</tr>
<tr>
<td>Type VII</td>
<td>$0 \leq \gamma_2 \leq 6.0$</td>
<td>$0 \leq \gamma_2 &lt; 1.4$</td>
<td>$-$</td>
<td>$0 \leq \gamma_2 &lt; 0.42$</td>
</tr>
</tbody>
</table>

(c) In the vicinity of the normal point ($0 < \gamma_2 < 0.5$ approx.) the efficiency of the method of moments in joint estimation is higher for Type VII than for the symmetrical Type A; this remark also applies to Type IV and Type A with four parameters when $\gamma_2^2 = 0$. In both cases it appears that near the normal point efficiency is high. (Note. It has been shown by Fisher (1921) that the system of curves for which the method of moments gives maximum likelihood estimates is of the form $y = e^{ax}$, where $A(x) = bx^2 + cx^3 + dx^4$ and if $c/b$ and $d/b$ are small, satisfies a differential equation similar to Pearson's. From this point of view it might be anticipated that near the normal point the method of moments would have high efficiency for the Pearson system. But, as has been noted by Jeffreys (1939a), the differential equation of Fisher's curves cannot approximate to a Type IV unless $d$ is positive, in which case $y$ diverges ultimately. Furthermore, it appears that the Type A curves are not mathematical approximations near the normal point to the Fisher system and yet we have indicated that these have high efficiency. It may be that Type A and Type IV have approximately the same moments as the Fisher system near the normal point, after identification of the first four moments, which would explain the high efficiency.)

8. Conclusion

Except in the cases of slight departure from normality such as occur when $0 < \gamma_2 < 0.3$ and $0 < \gamma_2^2 < 0.05$, it appears that the method of moments is inefficient when applied to Type A in the form considered here. A departure from normality of this size would require a large sample to detect with any degree of certainty. Thus in the case of the symmetrical Type A with $\sigma$ known, using (37) we have $N \text{ var} \hat{a}_4 > 3(4 - a_4)(2 + a_4)$, so that if $a_4$ is to exceed $3/(\text{var} \hat{a}_4$) then $N > 5700$ approximately when $a_4 = 0.2$. In general, then, maximum likelihood estimates should be found whenever there is reason to believe the observed distribution arises from a Type A population.

It is possible that these remarks also apply to Edgeworth's form of Type A,

$$P_x = \left\{ 1 + \frac{a_3^2}{3!} H_3(x) + \frac{a_4}{4!} H_4(x) + \frac{10a_5^2}{6!} H_5(x) \right\} g(x).$$

It would not, however, be an easy matter to find the admissible parameter values for $P_x$ to be positive in this case. But the method described here would lead, in the general case, to an upper bound for the efficiency of estimation by moments. This would not provide grounds for deciding whether Edgeworth's form was better from the point of view of efficiency, for we should be comparing upper bounds.

The exact bearing of our results on the practical problem of estimating parameters from a sample from a leptokurtic population is not easy to assess. For the measure of efficiency
calculated is based on a large sample limit and may be lower for a finite sample. Again for a Pearson distribution in the so-called heterotypic region the efficiency of estimation by moments may be zero. Thus in sampling from a Type VII population \( y = y_0(c^2 + x^2)^{-\frac{1}{2}} \), we have for large samples \( \text{var} m_2 = 2r/[N(r - 3)(r - 1)^2] \) which is infinite for \( r = 3 \), so that the efficiency of an estimate depending on the second moment would be zero. But in finite sampling in this case the variance of the second moment would not be infinite so that the efficiency would not be as low as the large sample figure suggests. A similar remark applies to estimation by moments in the case of Type A when the parameters lie in the vicinity of the critical region (a border strip on the curvilinear boundary of Fig. 1). For in this region we have shown that the efficiency is low or zero, but in practice it may be that the method of moments is not so useless.

The comparison between the Pearson Type VII and symmetrical Type A distributions having the same first four moments suggests that the method of moments is more efficient with the former, at least in the vicinity of the normal point. Further, in a situation in which we consider that an observed distribution can be described either by a Pearson or Type A curve, the shape of the distribution may be important. Now the Pearson system is unimodal but this is not the case in general with Type A. The form of Type A considered here may have three modes, and from a practical point of view statistical data showing such a characteristic might be regarded as heterogeneous. Thus if we are thinking in terms of unimodal distributions and do not allow negative frequencies the admissible parameter values for Type A will not be as extensive as is implied in Fig. 1.

The method we have described in this paper could be applied to other forms of Type A, such as those based on the logarithmically transformed normal curve or Pearson’s Type III, provided the questions of convergence involved in (2) were settled. Discrete distributions such as Charlier’s Type B could also be treated.

REFERENCES

MOMENT ESTIMATORS AND MAXIMUM LIKELIHOOD

BY L. R. SHENTON

College of Science and Technology, Manchester

1. Let $P(x; \theta_1, \theta_2, \ldots, \theta_h)$ be the probability of the variate $x$, depending on the $h$ parameters $\theta_j$ ($j = 1, 2, \ldots, h$). It is assumed that $P(x; \theta)$ possesses first derivatives with respect to the $\theta_j$, and that the moments $\mu'_r$ and their first derivatives exist and are finite. Let $\{q_r(x)\}$ be the orthogonal system of polynomials associated with $P(x; \theta)$, and write

$$q_r(x) = \sum_{j=0}^{\infty} a_{rj} x^j \quad (a_{r0} = 0, \ r = 0, 1, \ldots),$$

where

$$\int q_r^2(x) P(x; \theta) \, dx = \phi_r,$$

$$\int q_r(x) q_s(x) P(x; \theta) \, dx = 0 \quad (r \neq s).$$

To avoid undue complication at this stage we assume $P(x; \theta)$ is continuous throughout its range. We reconsider the restrictions on $P$ in a subsequent section.

Now let the formal 'Fourier' expansions of the derivatives of $P$ be

$$\frac{\partial P(x; \theta)}{\partial \theta_j} = P(x; \theta) \left\{ A_{j0} q_0(x) + A_{j1} q_1(x) + \ldots \right\} \quad (j = 1, 2, \ldots, h).$$

For the partial sum of the series in (3) we write

$$\mathcal{A}_r^{(j)} = \sum_{k=0}^{r} A_{jk} q_k(x),$$

where

$$A_{jk} = \int q_k(x) \frac{\partial P(x; \theta)}{\partial \theta_j} \, dx$$

$$= - \int \frac{\partial q_k(x)}{\partial \theta_j} P(x; \theta) \, dx$$

(4)$^b$

if the range is independent of $\theta$ (which will be assumed throughout the subsequent development). We consider $\mathcal{A}_r^{(j)}(x)$ as an economized (or Tchebycheffian) polynomial approximation to $\frac{\partial \log P}{\partial \theta_j}$, for a given value of $r$. This being the case, we may estimate the parameters $\theta_j$ from the system of equations*.

$$\mathcal{A}_r^{(j)} = S(x) \mathcal{A}_r^{(j)} / N = 0 \quad (j = 1, 2, \ldots, h),$$

(5)$^b$

where the summation is over a random sample of $N$. If $\frac{\partial \log P}{\partial \theta_j}$ is a polynomial in $x$, then for a determined $r$, the corresponding equation in (5) will be the associated likelihood equation. Again under certain conditions (5), when $r \to \infty$, will become equivalent to the likelihood equations.

Our main objective here is to determine cov $(\hat{\theta}_j, \hat{\theta}_k)$ for large samples for a solution $(\hat{\theta}_{1r}, \hat{\theta}_{2r}, \ldots, \hat{\theta}_{hr})$ of (5). Since $\partial q_0 = 0$, and from (4)$^b$, $A_{j0} = 0$, we may say that the estimators are consistent (a modified definition of a consistent estimator or statistics has recently been given by Fisher (1956, p. 144). * See expression (13) of my 1950 paper (to be referred to as M.L.).
2.1. Now consider small variations in \( \theta_{jr} \) which from (5) satisfy
\[
\sum_{k=1}^{h} \delta \theta_{kr} \frac{\partial \mathcal{L}^{(r)}}{\partial \theta_{kr}} + \delta m_{i} \mathcal{A}^{(r)} = 0,
\]
where the second member refers to fluctuations in the sample moments and
\[
\delta m_{i} \mathcal{A}^{(r)} = \sum_{k=0}^{r} A_{jk} \sum_{i=0}^{k} \delta b_{ki} \delta m_{i}^{r},
\]
where \( m_{i}^{r} = S_{i}^{r}/N. \) But using (4b)
\[
\mathcal{E} \left( \frac{\partial \mathcal{L}^{(r)}}{\partial \theta_{kr}} \right) = \mathcal{E} \sum_{\lambda=0}^{r} A_{j\lambda} \frac{\partial \varphi_{\lambda}}{\partial \theta_{kr}}
= - \sum_{\lambda=0}^{r} A_{j\lambda} A_{k\lambda} \varphi_{\lambda}
= -(j, k)_{\varphi} = -(k, j)_{\varphi}.
\]
We now write (6) in matrix form
\[
((j, k)_{\varphi}) \{ \delta \theta \} = \{ \delta m_{i} \mathcal{A}^{(r)} \},
\]
where \( \{ u \} \) is a column vector and in particular \( \{ \delta \theta \} = \{ \delta \theta_{1}, \delta \theta_{2}, \ldots, \delta \theta_{h} \}. \) It now follows that
\[
\mathcal{E} \{ \delta \theta \} \{ \delta \theta \}' = (\{ (j, k)_{\varphi} \} \{ \delta m_{i} \mathcal{A}^{(r)} \} \{ \delta m_{i} \mathcal{A}^{(r)} \})^{-1},
\]
where \( \{ u \}' \) is the transposed of \( \{ u \}. \) Taking expected values we have
\[
\mathcal{E} \{ \delta \theta \} \{ \delta \theta \}' = ((j, k)_{\varphi})^{-1}/N,
\]
provided it can be proved that
\[
\mathcal{E} \{ \delta m_{i} \mathcal{A}^{(r)} \} (\delta m_{i} \mathcal{A}^{(r)})' = (j, k)_{\varphi}/N.
\]
First of all
\[
\mathcal{E} (\delta m_{i} \mathcal{A}^{(r)} \delta m_{i}') = \mathcal{E} \sum_{k=0}^{r} A_{jk} \sum_{\lambda=0}^{k} a_{\lambda k} \delta m_{i}^{r} \delta m_{i}'^{r}
= \sum_{k=0}^{r} A_{jk} \sum_{\lambda=0}^{k} a_{\lambda k} (\mu_{i}^{(r+1)} - \mu_{i}^{(r)} \mu_{i}')/N,
\]
after using the well-known expression for \( \text{cov}(m_{i}^{r}, m_{i}') \) and \( \mu_{i}' = \mathcal{E} m_{i}'. \) Hence
\[
\mathcal{E} (\delta m_{i} \mathcal{A}^{(r)} \delta m_{i}') = N^{-1} \sum_{k=0}^{r} A_{jk} \mathcal{E} [g_{k}(x) \{ x^2 - \mu_{i}^{(r)} \}]
= N^{-1} \sum_{k=0}^{r} A_{jk} \mathcal{E} [g_{k}(x), \text{ since } A_{j0} = 0.
\]
But \( \delta g_{k}(x)/N \) is a linear function of the sample moments, so
\[
\mathcal{E} (\delta m_{i} \mathcal{A}^{(r)} \delta m_{i}') = N^{-1} \mathcal{E} \left( \sum_{\lambda=0}^{r} A_{j\lambda} q_{\lambda}(x) \sum_{x=0}^{k} A_{k\nu} g_{\nu}(x) \right)
\]
from which (10) follows.

2.2. Returning to (9b) we find expressions for the asymptotic covariances as follows:
\[
N \var \hat{\theta}_{1r} = \Delta_{1r}^{(r)},
N \text{ cov } (\hat{\theta}_{1r}, \hat{\theta}_{2r}) = \Delta_{12}^{(r)},
\]
and in general
\[
N \text{ cov } (\hat{\theta}_{jr}, \hat{\theta}_{kr}) = \Delta_{jk}^{(r)},
\]
where \( \Delta_{jk}^{(r)} \) is the cofactor of \( (j, k)_{\varphi} \) in the matrix \( ((j, k)_{\varphi}) \) whose determinant is \( \Delta^{(r)}. \)
It is now of some interest to notice that if \( r \to \infty \) in (11), and we assume that the various expansions involved remain valid, then (11) gives the usual result for the covariance of maximum likelihood estimators in terms of weighted cofactors of

\[
\left( \int \left( \frac{\partial \log P}{\partial \theta_j} \frac{\partial \log P}{\partial \theta_k} \right) P \, dx \right).
\]

It may be noted that there is a parallel line of thought in Geary's (1942) proof, that the generalized variance (asymptotically) is a minimum for maximum likelihood estimators of the parameters of a population which can be described in terms of frequency compartments. Again it is instructive to compare the basic ideas with those which arise in curvilinear regression when the independent variable is discrete. In the latter case the residual variance decreases (in general) with each parameter introduced and finally (after using a finite number of parameters) becomes zero. Reference may be made to Aitken (1935, 1945). On the other hand, with a moment estimator, the variance decreases and reaches a limiting value only after including an infinite number of terms in the likelihood equations.

2-3. To put the matter in relief, suppose \( P \) depends on one parameter \( \theta \) only, then with

\[
A_s \phi_s = - \int P(x; \theta) \frac{\partial g_s(x)}{\partial \theta} \, dx,
\]

(12a)

\( \hat{\theta}_r \) is a solution of

\[
S \sum_{s=0}^r A_s g_s(x)/N = 0
\]

(12b)

with large sample variance given by

\[
(N \, \text{var } \hat{\theta}_r)^{-1} = A_1^2 \phi_1 + A_2^2 \phi_2 + \ldots + A_r^2 \phi_r
\]

(12c)

and, assuming the validity of the expansion, asymptotic efficiency given by

\[
\text{Eff. } \hat{\theta}_r = \frac{A_1^2 \phi_1 + A_2^2 \phi_2 + \ldots + A_r^2 \phi_r}{A_1^2 \phi_1 + A_2^2 \phi_2 + \ldots + \infty}.
\]

(12d)

It is clear from (12c) that we also have (for samples of \( N \))

\[
\text{var } \hat{\theta}_1 \geq \text{var } \hat{\theta}_2 \geq \text{var } \hat{\theta}_3 \geq \ldots
\]

(12e)

3. The property indicated in (12e) for a single parameter can be generalized. For it is clear from the expansion for the generalized variance determinant given in (16) of M.L. that \( \Delta^{(r)} \) is non-decreasing (considered as a function of \( r \)). One would expect the same sort of property to hold for \( \text{var } \hat{\theta}_r \). For simplicity consider

\[
(N \, \text{var } \hat{\theta}_r)^{-1} = \Delta^{(0)} + \Delta^{(1)}_{1,1}.
\]

Now it may be proved that

\[
\Delta^{(0)} \Delta^{(r+1)}_{1,1} - \Delta^{(0)} \Delta^{(r+1)}_{1,1} = \phi_{r+1}
\]

(13)

This result is achieved by writing

\[
(j, k)_{r+1} = (j, k)_{r} + A_{j, r+1} A_{k, r+1} \phi_{r+1},
\]
414 Moment estimators and maximum likelihood

and then introducing the bordered forms of the determinants involving these terms, finally appealing to a pivotal condensation identity (see, for example, Aitken (1946), p. 49). Hence

\[ \text{var} \hat{\vartheta}_1 \geq \text{var} \hat{\vartheta}_2 \geq \text{var} \hat{\vartheta}_3 \geq \ldots \ (j = 1, 2, \ldots, h), \]  

and this holds for large samples of \( N \).

4. As elementary examples of moment estimators we mention that for a Poisson, Binomial (assuming the index is known) and Normal distribution the form of \( \partial \log P_x / \partial \theta \) is a polynomial, so that \( \mathcal{A}_j^{(p)} = 0 \) merely gives the usual maximum likelihood estimators. As further illustrations we mention: (a) the Poisson distribution

\[ P_x = e^{-m}m^x(1 - e^{-m})^{-1}/x! \quad (x = 1, 2, \ldots), \]  

with

\[ \frac{\partial}{\partial m} \log P_x = \frac{x}{m} - \frac{1}{1 - e^{-m}}, \]  

so that

\[ A_0 = 0 \quad (s = 0, 2, 3, \ldots), \]
\[ A_1 \neq 0, \]

and

\[ \mathcal{S} \frac{\partial \log P_x}{\partial m} \equiv A_1 \left( m'_1 - m \right), \]  

which leads to the likelihood equation for \( \hat{m} \),

(b) the distribution with probability density

\[ P_x = e^{-a}(2 - a + (a - 1)x) \quad (0 < x < \infty, 1 < a < 2), \]

with

\[ q_0(x) = 1, \quad \varphi_0 = 1, \quad A_0 \varphi_0 = 0; \]
\[ q_1(x) = x - a, \quad \varphi_1 = -a^2 + 4a - 2, \quad A_1 \varphi_1 = 1; \]
\[ q_2(x) = (-a^2 + 4a - 2)x^2 + (4a^2 - 20a + 12)x + 2a^2 + 4a - 4, \]
\[ \varphi_2 = 4(-a^2 + 4a - 2)(-4a^3 + 15a^2 - 12a + 2), \]
\[ A_2 \varphi_2 = 4(1 - a), \quad \text{and so on.} \]

For the ‘linear’ moment estimator \( a_1 \),

\[ \mathcal{A}_1^{(1)} = A_1 (m'_1 - a_1) = 0 \]

and

\[ (N \text{var} a_1)^{-1} = A_1^2 \varphi_1 = (-a^2 + 4a - 2)^{-1}. \]

For the ‘quadratic’ moment estimator \( a_2 \),

\[ \mathcal{A}_1^{(2)} = 0 \equiv 4a_2^3 + a_2(m'_2 - 8m'_1 - 1) - m'_2 + 7m'_1 - 2 = 0 \]

and

\[ (N \text{var} a_2)^{-1} = \frac{4a - 3}{-4a^3 + 15a^2 - 12a + 2}, \]

Similarly, the ‘cubic’ moment estimator is given by

\[ \mathcal{A}_1^{(3)} = 15a_3^2 - a_3(m'_3 - 15m'_2 + 57m'_1 + 2) + a_3(2m'_3 - 29m'_2 + 100m'_1 - 28) - m'_3 + 14m'_2 - 44m'_1 + 16 = 0, \]

and for the variance

\[ (N \text{var} a_3)^{-1} = \frac{15a^2 - 20a + 6}{-15a^4 + 50a^3 - 48a^2 + 8}. \]
The variances of the successive estimators converge fairly rapidly to a limiting value, and for large $r$
\[
(N \text{ var } \alpha_{r+1})^{-1} - (N \text{ var } \alpha_r)^{-1} \sim \frac{4\pi}{(a-1)^3} \exp \left[ \frac{2-a}{a-1} - 4 \sqrt{\frac{(2-a)r}{a-1}} \right].
\]

We intend to make further use of this illustrative example on another occasion.

5.1. Application to the negative binomial distribution. Various methods of estimating the parameters have been considered by Anscombe (1950), and for three of them he gave some contours of large sample efficiency (see also Evans (1956) and Haldane (1941)). We consider the probability function
\[
P_x = \frac{\alpha(x+1) \ldots (x+x-1)}{x!} \frac{(\lambda \alpha)^x}{(\lambda + \alpha)^{x+1}} (x, \lambda > 0, x = 0, 1, \ldots).
\]

The orthogonal polynomials (Aitken & Gonin, 1935) are given by
\[
g_r(x) = (1 - \lambda \Delta_r/x)^{x+r-1} x!,
\]
where \(\Delta_r f(x) = f(x+1) - f(x)\), and there are the additional properties
\[
\begin{align*}
E g_r^2(x) &= \lambda r (\lambda + \alpha)^r (x+r-1)!/x! \\
&= \phi_r, \\
E x^{(s)} &= \lambda^{s-1} \prod_{r=1}^{s-1} \left(1 + \frac{r}{\alpha}\right) = \mu_{(s)}.
\end{align*}
\]
Moreover if
\[
\begin{align*}
\frac{\partial P_x}{\partial \alpha} &= P_x \sum A_s g_s(x), \\
\frac{\partial P_x}{\partial \lambda} &= P_x \sum B_s g_s(x),
\end{align*}
\]
then
\[
\begin{align*}
A_s \phi_s &= -\sum_{x=0}^{\infty} P_x \frac{\partial g_s(x)}{\partial \alpha}, \\
B_s \phi_s &= -\sum_{x=0}^{\infty} P_x \frac{\partial g_s(x)}{\partial \lambda}.
\end{align*}
\]
But
\[
\frac{\partial g_s(x)}{\partial \alpha} = \lambda s (s+\alpha-1) g_{s-1}(x)/\alpha^2 + \{\ln (1 - \lambda \Delta_r/x)\} g_s(x)
\]
\[
= \lambda s (s-1) g_{s-1}(x) + \lambda^2 s (s-1) g_{s-2}(x)/2x^2 - \ldots,
\]
and
\[
\begin{align*}
A_s \phi_s &= (-1)^{s+1} \lambda^s (s-1)!/\alpha^s (s > 1), \\
&= 0 \quad (s = 0 \text{ or } 1).
\end{align*}
\]
Similarly, using
\[
\begin{align*}
B_s \phi_s &= \sum_{x=0}^{\infty} \left(1 + (s-1)/\alpha\right) g_{s-1}(x), \\
B_s \phi_s &= 1 \quad (s = 1), \quad B_s = 0 \quad (s > 1).
\end{align*}
\]
Clearly from (21) and (22)
\[
\delta \left( \frac{\partial \log P_x}{\partial \alpha} \right) = 0,
\]
so that \(\hat{\alpha}\) and \(\hat{\lambda}\) are asymptotically uncorrelated (this was pointed out by Anscombe (1950); we mention in this connexion that it is also possible to choose asymptotically uncorrelated estimates for a Neyman Type A distribution with two parameters).
5.2. The truncated likelihood equations are (cf. (23) of M.L.)
\[ \frac{\alpha^2 q_2(x)}{2(\lambda + \alpha)^2(\alpha + 1)^2} - \frac{\alpha^2 q_3(x)}{3(\lambda + \alpha)^3(\alpha + 2)^2} + \frac{\alpha^4 q_4(x)}{4(\lambda + \alpha)^4(\alpha + 3)^2} - \ldots \]
\[ N^{-1}S_{g1}(x) = 0, \quad (24a) \]
\[ \frac{\alpha^2 q_2(x)}{2(\lambda + \alpha)^2(\alpha + 1)^2} - \frac{\alpha^2 q_3(x)}{3(\lambda + \alpha)^3(\alpha + 2)^2} + \frac{\alpha^4 q_4(x)}{4(\lambda + \alpha)^4(\alpha + 3)^2} - \ldots \]
\[ + \frac{(-\alpha)^r q_r(x)}{r(\lambda + \alpha)^r(\alpha + r - 1)^r} = 0 \quad (24b) \]
with 'solution'
\[ \lambda = S_x/N = m_1, \]
\[ \alpha = \alpha_{2r-3}. \quad (25) \]
Considered as an equation in \( \alpha \) it may be shown that (24b) after simplification is of degree \( 2r - 3 \), and in particular for \( r = 2 \) it corresponds to the 'moment' estimator. We have been unable to prove anything about the existence of solutions in the general case.

\[ \text{Fig. 1. 98} \% \text{ efficiency contours for three estimators of } \alpha \text{ for the negative binomial.} \]

For the variance of \( \alpha_{2r-3} \) we find (cf. M.L. (23), and Fisher (1941))
\[ (N \text{ var } \alpha_{2r-3})^{-1} = k(u_2 + u_3 + \ldots + u_r), \quad (26) \]
where
\[ k = \lambda^2/(2(\lambda + \alpha)^2(\alpha + 1)^2), \]
\[ u_2 = 2\lambda^2(s-1)/s(\lambda + \alpha)^2 \]
and the asymptotic efficiency of \( \alpha_{2r-3} \) is
\[ \text{Eff. } (\alpha_{2r-3}) = \frac{u_2 + u_3 + \ldots + u_r}{u_2 + u_3 + \ldots + u_r}. \quad (27) \]

An indication of the efficiency for the 'linear', 'cubic' and 'quintic' estimators of \( \alpha \) is given in Figs. 1 and 2, showing respectively the 98 and 90 \% contours. There is obviously a considerable gain if we use the 'cubic' as against the 'linear' estimator; the gain for the 'quintic' as against the 'cubic' is not so marked, as one might expect. Of course the efficiency improves for any of the estimators as \( \alpha \) increases (\( \lambda \) fixed) for this corresponds to an approach to the Poisson distribution with mean \( \lambda \).

5.3. From a practical point of view it is tedious to compute a root of (24b) if the degree is high, although one may be provided with a fairly good start (using an iterative process) from
the previous equation. To give some idea of the 'weight' of computation involved we remark that (using $N^{-1}s_{q(\alpha)} = m_{(\alpha)}$)

$$
\begin{align*}
N^{-1}s_{q_2}(x) &= m_{(\alpha)} - (1 + \alpha^{-1})m_{(1)}^2 \\
N^{-1}s_{q_3}(x) &= m_{(\alpha)} - 3(1 + 2\alpha^{-1})m_{(1)}m_{(2)} + 2(1 + 2\alpha^{-1})(1 + \alpha^{-1})m_{(3)}^3 \\
N^{-1}s_{q_4}(x) &= m_{(\alpha)} - 4(1 + 3\alpha^{-1})m_{(1)}m_{(3)} + 6(1 + 3\alpha^{-1})(1 + 2\alpha^{-1})m_{(1)}^2m_{(3)} \\
&
\end{align*}
$$

(28)

so that the truncated likelihood equations (24b) are:

**Linear**

$$
A(x_1) = \sum_{s=0}^{1} A_s x_1^s = 0,
\begin{align*}
A_1 &= m_{(\alpha)} - m_{(1)}^2, \\
A_2 &= -m_{(1)}^2, \\
(N \text{ var } x_1)^{-1} &= k u_2.
\end{align*}
$$

(29a)

**Cubic**

$$
B(x_2) = \sum_{s=0}^{3} B_s x_2^s = 0,
\begin{align*}
B_3 &= m_{(\alpha)} - m_{(1)}^2, \\
3B_2 &= 9m_{(1)}m_{(\alpha)} - 2m_{(\alpha)} - 7m_{(1)}^3 + 6m_{(\alpha)} - 9m_{(1)}^2, \\
B_1 &= 6m_{(1)}m_{(\alpha)} - 7m_{(1)}^3 - 2m_{(1)}^3, \\
3B_0 &= -14m_{(1)}^3; \\
(N \text{ var } x_2)^{-1} &= k(u_2 + u_3).
\end{align*}
$$

(29b)

**Quintic**

$$
C(x_3) = \sum_{s=0}^{5} C_s x_3^s = 0,
\begin{align*}
C_5 &= m_{(\alpha)} - m_{(1)}^2, \\
3C_4 &= -2m_{(\alpha)} + 12m_{(1)}m_{(\alpha)} - 10m_{(3)}^3 + 15m_{(\alpha)} - 18m_{(1)}^2, \\
6C_3 &= 3m_{(\alpha)} - 16m_{(1)}m_{(\alpha)} + 36m_{(1)}^2m_{(\alpha)} - 23m_{(1)}^3 - 12m_{(\alpha)} + 120m_{(3)}m_{(\alpha)} \\
&
\quad - 120m_{(1)}^3 + 36m_{(\alpha)} - 66m_{(1)}^2, \\
3C_2 &= 90m_{(1)}^2m_{(\alpha)} - 24m_{(1)}m_{(\alpha)} - 69m_{(3)}^3 + 72m_{(1)}m_{(\alpha)} - 110m_{(3)}^3 - 18m_{(1)}^2, \\
6C_1 &= 216m_{(1)}^2m_{(\alpha)} - 253m_{(1)}^2m_{(\alpha)} - 120m_{(1)}^3, \\
C_0 &= -23m_{(1)}; \\
(N \text{ var } x_3)^{-1} &= k(u_2 + u_3 + u_4).
\end{align*}
$$

(29c)

Fig. 2. 90 % efficiency contours for three estimators of $x$ for the negative binomial.
If we substitute in (29) the population value of $\alpha$ and neglect terms involving $N^{-1}$ and higher powers, it may be verified that

$$\delta A(\alpha) = \delta B(\alpha) = \delta C(\alpha) = 0.$$  \hspace{1cm} (30)

5.4. We now consider briefly two examples:

(i) Fisher (1941) in discussing the negative binomial distribution gives two examples based on tick data (p. 186). For his second example some results are summarized in the form:

$$N = 82$$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$\alpha_{2r-3}$</th>
<th>$\lambda = 0.560,9756$</th>
<th>$\sigma(\alpha_{2r-3})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.549</td>
<td>3.48/$\sqrt{N}$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.524</td>
<td>3.04/$\sqrt{N}$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.592</td>
<td>2.91/$\sqrt{N}$</td>
<td></td>
</tr>
<tr>
<td>$\infty$</td>
<td>1.778</td>
<td>2.80/$\sqrt{N}$</td>
<td></td>
</tr>
</tbody>
</table>

(ii) Haldane (1941) and quoted in M.L. p. 116:

$$N = 1096$$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$\alpha_{2r-3}$</th>
<th>$\lambda = 2.156,9343$</th>
<th>$\sigma(\alpha_{2r-3})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10.385</td>
<td>83.6/$\sqrt{N}$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9.918</td>
<td>82.8/$\sqrt{N}$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9.906</td>
<td>82.8/$\sqrt{N}$</td>
<td></td>
</tr>
<tr>
<td>$\infty$</td>
<td>9.900</td>
<td>82.8/$\sqrt{N}$</td>
<td></td>
</tr>
</tbody>
</table>

* Calculated for the values $\lambda = 2.157$, $\alpha = 10$.

The efficiency for $\alpha_1$ in (i) is only 64% while for (ii) it is 98%.


6.1. We now consider under what conditions it is possible to assert that $\Delta_1^1 + \Delta(\rightarrow N \text{var } \theta_1$ as $r \to \infty$, where $\theta_1$ is the maximum likelihood estimator (and similarly for the other covariances). Basically this reduces to the validity of

$$\mathbb{E} \left[ \frac{\partial \log P(x; \theta)}{\partial \theta_j} \frac{\partial \log P(x; \theta)}{\partial \theta_k} \right] = \sum_{\lambda=0}^{\infty} A_{jk} A_{\lambda \lambda} \phi_{\lambda},$$

which in turn depends on that of

$$\mathbb{E} \left[ \frac{\partial \log P(x; \theta)}{\partial \theta_j} \right]^2 = \sum_{\lambda=0}^{\infty} A_{jj}^{2} \phi_{\lambda}.$$

Thus we have to consider the conditions under which Parseval's formula holds for $(1/P) \left( \partial P/\partial \theta \right)$ with respect to the weight function $P(x; \theta)$. Clearly we must assume the existence of the moment sequence $\{\mu_s\}$ so that there is a tie-up with the problem of moments and a property of the solutions due to M. Riesz (see, for example, Shohat & Tamarkin (1943, pp. 61–6 and Theorem 2-20)). Riesz has shown that if there is a unique $\psi(x)$ (or what is called an extremal solution) such that

$$\int_{-\infty}^{\infty} x^s d\psi(x) = \mu_s^s \hspace{1cm} (s = 0, 1, \ldots),$$

then Parseval's formula,

$$\int_{-\infty}^{\infty} \{f(x)\}^2 d\psi(x) = \sum_{s=0}^{\infty} f_s^2,$$

where

$$f_s = \int_{-\infty}^{\infty} f(x) \psi_s(x) d\psi(x)$$

(33)
holds for every $f(x)$ of the class $L^2$ (this implies the measurability of $f(x)$ with respect to $\psi(x)$ and the convergence of $\int_{-\infty}^{\infty} (f(x))^2 \, d\psi(x)$). Thus in our application it is certainly necessary for the Stieltjes integral appearing as the first member of (32) to converge.

6-2. We now mention briefly various criteria for deciding whether the moment problem is determined (see Shohat & Tamarkin (1943), pp. vii–viii, 19–22; also Kendall (1943, pp. 106–10)). For the Stieltjes moment problem (variate range $0$ to $\infty$; $\psi(x)$ constant elsewhere) there is a unique solution, if $\int_{-\infty}^{\infty} f(x)^2 \, d\psi(x)$ is such that $a_s > 0$ and $\sum a_s$ diverges. A difficulty here is that it may not be possible to find a comparatively simple expression for $a_s$, as is the case, for example, with Neyman’s type A with two parameters. But for Poisson moments (mean $m$) it may be shown that $\lambda^{1/2} = m^{3/2}$ and $a_{2s} = (s-1)!/m^s$ so that $\sum a_s = \exp m + m^{-1} \sum s!/m^s$ which clearly diverges. Similarly, for the negative binomial (18) it appears that

$$
\sum_{s=0}^{\infty} \frac{P_s}{x+s} = \frac{1}{z+1} \frac{\lambda}{z+1} \frac{(1+\lambda z)}{1} \frac{(1+\lambda z)^s}{1+\lambda z} \frac{2!}{1} \frac{(1+2\lambda)^s}{1} \ldots
$$

is such that $a_s > 0$ and $\sum a_s$ diverges. A difficulty here is that it may not be possible to find a comparatively simple expression for $a_s$, as is the case, for example, with Neyman’s type A with two parameters. But for Poisson moments (mean $m$) it may be shown that $\lambda^{1/2} = m^{3/2}$ and $a_{2s} = (s-1)!/m^s$ so that $\sum a_s = \exp m + m^{-1} \sum s!/m^s$ which clearly diverges. Similarly, for the negative binomial (18) it appears that

$$
\sum_{s=0}^{\infty} \frac{P_s}{x+s} = \frac{1}{z+1} \frac{\lambda}{z+1} \frac{(1+\lambda z)}{1} \frac{(1+\lambda z)^s}{1+\lambda z} \frac{2!}{1} \frac{(1+2\lambda)^s}{1} \ldots
$$

so that after an equivalence transformation we find

$$
\sum_{s=0}^{\infty} \frac{P_s}{x+s} = \frac{1}{z+1} \frac{\lambda}{z+1} \frac{(1+\lambda z)}{1} \frac{(1+\lambda z)^s}{1+\lambda z} \frac{2!}{1} \frac{(1+2\lambda)^s}{1} \ldots
$$

which diverges since a term of the first infinite series is $O(1/\lambda)$. Hence if we assume the stochastic convergence of $x_{r-3}$ to the maximum likelihood estimator then (26) remains valid as $r \to \infty$.

We also note here that the normal distribution moments uniquely determine a distribution, as may be verified by an appeal to Carleman’s series test (Kendall, 1943, p. 109).

6-3. It is also possible to say something about Gram-Charlier distributions consisting of a finite number of terms. Thus for a Type A series based on a normal or Type III probability density, Parseval’s theorem holds provided the frequency is always positive (see, for example, Shenton (1954, pp. 80–2)). Thus the expansions given by me in the efficiency of the method of moments and the Gram-Charlier Type A distribution (1951) converge with certain parameter restrictions.

For example, for the maximum likelihood estimator $\hat{a}_4$ of $a_4$ in

$$
P(x; a_4) = C(x) g(x),
$$

where

$$
C(x) = 1 + H_4(x) a_4/4!, \quad g(x) = \exp \left( -\frac{1}{2} x^2 \right) /\sqrt{(2\pi)},
$$

$$
x = (X - m)/\sigma, \quad H_4(x) = \text{Hermite polynomial of degree 4},
$$

it turns out that if $\sigma$ is known,

$$
a_4^2/\text{var} \hat{a}_4 + 1 = \int_{-\infty}^{\infty} \frac{g(x)}{C(x)} \, dx.
$$
Parseval’s expansion holds and is equivalent to the convergence (as \( r \to \infty \)) to zero of

\[
\int_{-\infty}^{\infty} g(x) C(x) \left[ \frac{1}{C(x)} - \sum_{k=0}^{r} a_k g_k(x) \right]^2 \, dx
\]

(38)

with \( 0 < a_4 < 4 \), where \( \{g_k(x)\} \) is the orthogonal system with respect to \( g(x) C(x) \) (see the 1951 paper, (35) and (36)).

Similarly, if \( \psi(x) \) is a determined solution of a Stieltjes moment problem, then it may be shown that the various Parseval expansions which arise in connexion with the estimation of the parameters of a Gram-Charlier Type B development (consisting of a finite number of terms) expressible in the form

\[
\hat{\psi}(x) = \int_{0}^{x} \pi(t) \, d\hat{\psi}(t),
\]

(39)

where \( \pi(t) \) is a non-negative polynomial, converge (Shenton, 1957, pp. 153–6). Thus, in view of the remarks in §6.2, convergence questions arising in connexion with the estimation of parameters in Gram-Charlier Type B based on the Poisson and Negative Binomial (or geometric) distributions are readily settled.

7. Concluding remarks. The large sample moment estimators we have introduced here have the unusual property that although they involve higher moments this does not imply larger sampling variances; on the contrary the sampling variances decrease as higher moments are introduced. As far as we are aware examples of this sort of behaviour are rarely met in the literature. It may turn out that there is a similar property for the covariances. It must be mentioned, however, that the property of the variances might have been anticipated when it is recalled that the estimators (under certain conditions) ultimately converge to maximum likelihood estimators.

Our treatment here has been mainly formal and general, and covers discrete and continuous distributions. We reserve for another occasion a discussion of the formula for the covariance matrix (and its relation to the derivations given by earlier writers) and remarks on the distributions of the moment estimators.

I have to thank Mr A. Fletcher for drawing Figs. 1 and 2 and for assisting in some of the computations.

REFERENCES

1. In a previous paper (Shenton, 1958) we introduced a class of moment estimators depending on the sample moments. We now consider the asymptotic distribution of a moment estimator and give expressions for the first four cumulants, these suggesting that the distribution is asymptotically normal. As a special case, when the moment estimators depend on an infinity of moments, the cumulants are those of the maximum likelihood estimator (assuming one exists) and our expressions agree with those given by Haldane and Smith (1956).

Our main purpose here is to treat the problem for the case of a single parameter in general and attempt an approach which will lend itself to the development of the multi-parameter estimation problem. Thus although we give an illustrative example we do this to indicate what complexity is to be expected in the method, and not as a practical example, for most of these involve at least two parameters. The extension to several parameters involving simultaneous estimation is deferred, for there are several difficulties of simplification to overcome (see Haldane (1958)).

(In the press, Biometrika 46, 3 and 4, December 1959)

Three line drawings included.
In a previous paper (Shenton, 1958) we introduced a class of moment estimators depending on the sample moments. We now consider the sampling distribution of a moment estimator and give expressions for the first four cumulants, these suggesting that the distribution is asymptotically normal. As a special case, when the moment estimators depend on an infinity of moments, the cumulants are those of the maximum likelihood estimator (assuming one exists) and our expressions agree with those given by Haldane and Smith (1956).

Our main purpose here is to treat the problem for the case of a single parameter in general and attempt an approach which will lend itself to the development of the multi-parameter estimation problem. Thus although we give an illustrative example we do this to indicate what complexity is to be expected in the method, and not as a practical example, for most of these involve at least two parameters. The extension to several parameters involving simultaneous estimation is deferred, for there are several difficulties of complication to overcome (see Haldane (1953)).

2.1 Let $P(x, \theta)$ be the probability of the variate $x$ depending on the parameter $\theta$. For the sample and population moments we write

\[
\bar{m}_s = \frac{\sum_{j=1}^{n} x_j}{n}, \quad m_\theta^* = \frac{\sum_{j=1}^{n} (x_j - \bar{m}_s)^i}{n}\]
The Distribution of Moment Estimators

\[ \mu_s = \sum m_s, \quad \mu_s^* = \sum m_s^*. \]  
(1b)

It is assumed that the moments exist and that they are differentiable in some \( \theta \) - interval. The \( q \) th. moment estimator \( \theta \) (when it exists) satisfies the determinantal equation

\[ h(\phi) = 0, \quad \phi = \theta_q \]  
(2)

where

\[ h(\theta) = \begin{vmatrix}
0 & m_0 & m_1 & \cdots & m_r \\
0 & \mu_0 & \mu_1 & \cdots & \mu_r \\
\frac{\partial \mu_0}{\partial \theta} & \mu_1 & \mu_2 & \cdots & \mu_{r+1} \\
\frac{\partial \mu_1}{\partial \theta} & \mu_2 & \mu_3 & \cdots & \mu_{r+2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial \mu_r}{\partial \theta} & \mu_{r+1} & \mu_{r+2} & \cdots & \mu_{2r}
\end{vmatrix}. \]  
(3)

This is an alternative form of the expression (5) p.411 given by Shenton (1958) for the case of a single parameter. It may be derived by applying Schweins' theorem in determinants (see for example Aitken (1946) or Muir (1906)) to the series expansion (truncated) for a moment estimator appearing in (13) p.112 of my 1950 paper.

It will be noted that with the exception of the first row...
in (3), $h(\theta)$ consists of elements which are function of $\theta$.

It is convenient to expand $h(\theta)$ by its first row in terms of cofactors as follows:

$$h(\theta) = \sum_{s=0}^{q} m_s M_s(\theta)$$

$$= \langle m M \rangle.$$  (4a)

For brevity $M_s$ is written for $M_s(\theta)$, and in (4b) we have introduced a notation for the vector inner product of the vectors $m = [m_0, m_1, \ldots m_q]$ and $M = [M_0, M_1, \ldots M_q]$. It will be noted that $m$ has $q$ random components, for $m_0 = 1$.

2.3 From (2) we can write the stochastic $q$-dimensional Taylor expansion for $\theta_q$ about the point $\mu$, and in fact

$$\theta_q = \left\{ \exp\left(\sum_{m=1}^{\infty} \frac{1}{m!} M M = m - \mu, A = M \right) \right\} \theta_q$$

After the evaluation of the inner-product operator in (5) we have

$$\theta_q = \sum_{s=0}^{\infty} s! \frac{\delta(s, \theta)}{s}$$

where $s! \frac{\delta(s, \theta)}{s} \equiv \langle M M = m - \mu \rangle \theta_q$.  (6a)

(6b)
and in (6b) the partial derivatives of $\theta_q$ are to be evaluated at $m=\mu$, this being implied in the barred derivative symbol. For example,

$$
\Phi_1 = M_1 \frac{\partial \theta_1}{\partial m_1} + M_2 \frac{\partial \theta_2}{\partial m_2} + \ldots + M_q \frac{\partial \theta_q}{\partial m_q}
$$

where $\bar{\frac{\partial \theta_q}{\partial m_i}} = \frac{\partial \theta_q}{\partial m_i} \bigg|_{m=\mu}$ etc.

2.4 Expressions for $\Phi_s$. Differentiating (2) partially with respect to $m_s$ we have

$$
\frac{\partial \phi}{\partial m_s} \frac{\partial h(\phi)}{\partial \phi} + M_s = 0, \quad s=1,2,\ldots q; \quad M_s = \mathcal{M}_s(\phi),
$$

(7)

from which

$$
\mathcal{M}_s + \bar{h}_i \frac{\partial \theta_i}{\partial m_s} = 0
$$

(8)

where from (3)

$$
\bar{h}_i = \frac{\partial h(\phi)}{\partial \phi} = -\left< c \mathcal{M} \frac{\partial \theta}{\partial \theta} \right>
$$

(9)

We thus have

$$
\Phi_1 = -\left< c \mathcal{M} \frac{M}{\bar{h}_i} \right>.
$$

(10)
To find an expression for $\Phi_2$, we require terms such as $\frac{\partial^2 \Phi}{\partial m_1 \partial m_2}$ which could be found from (7) by further differentiation. This piecemeal procedure can be avoided, and indeed from (7) for arbitrary $A$

$$\frac{\partial h(\phi)}{\partial \phi} \left< \frac{A}{\partial m} \right. = - \left< \frac{A}{\partial m} \right> m_0$$

so that

$$\left< \frac{A}{\partial m} \right> \left\{ \frac{\partial h(\phi)}{\partial \phi} \left< \frac{A}{\partial \phi} \right> \right\} = - \left< \frac{A}{\partial m} \right> \left< \frac{A}{\partial \phi} \right> m_0$$

The expressions (10) and (14) give the first two terms of the Taylor expansion for $\Phi_2$, and they may be found in a similar way. Thus for $\Phi_3$ and higher terms may be found in this manner. Thus for $\Phi_3$ we operate on (13) at the following expressions.

$$\frac{\partial h(\phi)}{\partial \phi} \left< \frac{A}{\partial m} \right> \phi + \left< \frac{A}{\partial m} \right> \left\{ \frac{\partial h(\phi)}{\partial \phi} \left< \frac{A}{\partial \phi} \right> \right\} = - \left< \frac{A}{\partial m} \right> \left< \frac{A}{\partial \phi} \right> m_0$$

But

$$\left< \frac{A}{\partial m} \right> \frac{\partial h(\phi)}{\partial \phi} = \left< \frac{A}{\partial m} \right> \frac{\partial^2 h(\phi)}{\partial \phi^2} + \left< \frac{A}{\partial m} \right> m_0$$

so that substituting in (12) leads to

$$\frac{\partial h(\phi)}{\partial \phi} \left< \frac{A}{\partial m} \right> \phi = - \frac{\partial h(\phi)}{\partial \phi} \left< \frac{A}{\partial m} \right> \phi^2 - 2 \left< \frac{A}{\partial m} \right> \left< \frac{A}{\partial \phi} \right> m_0$$

(13)
and on putting $A = M$ and $m = \mu$, and using (8) we have,

$$
2! \Phi_2 = -\frac{\bar{h}_2}{\bar{h}_1^3} \left< M \delta M_0 \right>^2 + \frac{2}{\bar{h}_1^2} \left< M \delta M_0 \right> \left< M \delta M_0 \right>
$$

(14)

in which, after differentiating (3) twice,

$$
\bar{h}_2 = \frac{\bar{h}(\phi)}{\bar{\phi}^2} = -2 \left< \frac{\partial M_0}{\partial \theta} \frac{\partial m}{\partial \theta} \right> - \left< M \delta \delta m \right>.
$$

(15)

The expressions (10) and (14) give the first two terms of the Taylor expansion for $\theta_q$, and they involve only terms in the sample moment deviations $(m_i - \mu_i, \text{etc.})$ and the population parameter $\theta$. Expressions for $\Phi_3$ and higher terms may be found in a similar way. Thus for $\Phi_3$ we operate on (13) with $\left< A \delta \theta \right>$, replace $A$ by $M$ and $m$ by $\mu$. We thus arrive at the following expressions:

$$
3! \Phi_3 = \frac{\bar{h}_1 \bar{h}_3 - 3 \bar{h}_2^2}{\bar{h}_1^5} \left< M \delta M_0 \right>^3 + \frac{1}{\bar{h}_1^4} \left< M \delta M_0 \right> \left< M \delta M_0 \right> \left< M \delta M_0 \right>
$$

(16a)

(16b) and (16c) follow immediately from the identity $\left< \delta M_0 \right> = 0$.
The Distribution of Moment Estimators

\[ 4! \Phi_4 = \frac{13 \bar{h}_3}{\bar{h}_1^7} + \frac{4}{\bar{h}_1^4} < M_4 > + \frac{9}{\bar{h}_1^6} < M_3 > \frac{< M_3 >}{< M_2 >} \]

\[ + \frac{4}{\bar{h}_1^4} < M_3 > < M \frac{3}{\theta^3} > - \frac{117}{\bar{h}_1^5} < M_5 > < M \frac{2}{\theta^2} >^2 \]

\[ - \frac{33}{\bar{h}_1^5} < M_5 > < M \frac{2}{\theta^2} >^3 + \frac{42}{\bar{h}_1^4} < M_3 > < M \frac{2}{\theta^2} >^2 \]

\[ + \frac{45}{\bar{h}_1^4} < M_2 > < M \frac{2}{\theta^2} > < M \frac{2}{\theta^2} > \]

\[ (16b) \]

\[ \bar{h}_3 = -3 \frac{\frac{d}{d\theta} \frac{d^3}{d\theta^2} \alpha}{\frac{d^3}{d\theta^2} \alpha} - 3 \frac{\frac{d}{d\theta} \frac{d^2}{d\theta^2} \alpha}{\frac{d^3}{d\theta^2} \alpha} - \left( \frac{\frac{d^3}{d\theta^2} \alpha}{\frac{d^3}{d\theta^2} \alpha} \right) \]

\[ = \left( \frac{\frac{d^3}{d\theta^2} \alpha}{\frac{d^3}{d\theta^2} \alpha} \right) \]

\[ (16c) \]

\[ \bar{h}_4 = -4 \frac{\frac{d}{d\theta} \frac{d^4}{d\theta^4} \alpha}{\frac{d^4}{d\theta^4} \alpha} - 6 \frac{\frac{d}{d\theta} \frac{d^3}{d\theta^3} \alpha}{\frac{d^4}{d\theta^4} \alpha} - 4 \frac{\frac{d^2}{d\theta} \frac{d^3}{d\theta^3} \alpha}{\frac{d^4}{d\theta^4} \alpha} - \left( \frac{\frac{d^4}{d\theta^4} \alpha}{\frac{d^4}{d\theta^4} \alpha} \right) \]

\[ = \left( \frac{\frac{d^4}{d\theta^4} \alpha}{\frac{d^4}{d\theta^4} \alpha} \right) \]

\[ (16d) \]

Moreover, it is readily seen that

\[ \text{It is of interest to note that the expressions in (9), (15), (16c) and (16d) follow immediately from the identity } \left( \frac{\alpha}{\alpha} \right) = 0. \]

3. Expected Values of combinations of $\Phi'$.

3.1 We require the expected value of expressions such as $< A M > < B M > < C M > ^t \ldots$. Now for arbitrary $A$.
The Distribution of Moment Estimators

\[ E \left\{ \exp \left( y \left( \text{MA} \right) \right) \right\} = \left\{ 1 + \frac{y^2}{2! n^2} \psi_2 + \frac{y^3}{3! n^3} \psi_3 + \ldots \right\}^n \]  

(17)

where

\[ \psi_s = E \left\{ X A \right\}^s, \quad X_r = x - \mu_r. \]

Thus

\[ E \left\{ \text{MA} \right\}^2 = \frac{\psi_2}{n} \]
\[ E \left\{ \text{MA} \right\}^3 = \frac{\psi_2}{n^2} \]
\[ E \left\{ \text{MA} \right\}^4 = \frac{\psi_4}{n^3} + 3(n-1) \frac{\psi_2^2}{n^3} \]
\[ E \left\{ \text{MA} \right\}^5 = \frac{\psi_5}{n^4} + 10(n-1) \frac{\psi_2 \psi_3}{n^4} \]
\[ E \left\{ \text{MA} \right\}^6 = \frac{\psi_6}{n^5} + (n-1) \left( \frac{10 \psi_2^2}{n^5} + \frac{15 \psi_2 \psi_4}{n^5} + 15(n-1)(n-2) \frac{\psi_2^2}{n^5} \right) \]

Moreover it is readily seen that

\[ \psi_s = E \left\{ A_1(x-\mu_1) + A_2(x^2-\mu_2) + \ldots + A_l(x^q-\mu_q) \right\}^s \]
\[ = \left( A \mu \right)^s - \frac{s}{1} \left( A \mu \right)^{s-1} \left( A \mu \right)^2 + \left( \frac{s}{2} \right) \left( A \mu \right)^{s-2} \left( A \mu \right)^2 \]
\[ - \ldots + (-1)^{s-2} \left( \frac{s}{s-2} \right) \left( A \mu \right)^{s-2} \left( A \mu \right)^2 + (-1)^{(s-1)}(s-1) \left( A \mu \right)^s \]  

(19)

where \[ A = \frac{\Delta}{\tau} \]

in which \( A \mu \) is a symbolic multinomial expression to be
interpreted by replacing a term such as $\mu_1 \mu_2$ by $\mu_1 \mu_2$. For example, when $q=2$

$$\langle A \mu \rangle^2 = (A_0 \mu_2 + A_1 \mu_1 + A_2 \mu_0)^2$$

$$= A_0^2 + A_1^2 \mu + A_2^2 \mu + 2A_1 A_2 \mu + 2A_0 A_2 \mu + 2A_0 A_1 \mu_1 .$$

3.2 Derived Expressions. To evaluate an expression such as

$$\langle MA \rangle \langle MB \rangle \langle MC \rangle$$

we merely express this as

$$\mathcal{E} \frac{\partial^2}{\partial A \partial A} \langle B \rangle \langle MA \rangle^3$$

$A, B, C$ being arbitrary vectors independent of $m$ and $\theta$. Thus from

$$\mathcal{E} \langle MA \rangle^2 = \left( \langle A \mu \rangle^2 - \langle A \mu \rangle^2 \right) / n ,$$

$$n \mathcal{E} \langle MA \rangle \langle MB \rangle = \langle A \mu \rangle \langle B \mu \rangle - \langle A \mu \rangle \langle B \mu \rangle .$$

For example

$$n \mathcal{E} \langle M \rangle \langle BM \rangle = \langle M \rangle \langle B \mu \rangle ,$$

$$n \mathcal{E} \langle M \rangle \langle M \rangle = \langle M \rangle \langle M \rangle - \langle M \rangle \langle M \rangle = - \Delta \langle \frac{\partial M}{\partial \theta} \rangle ,$$

$$n \mathcal{E} \langle M \rangle^2 = - \Delta \langle \frac{\partial M}{\partial \theta} \rangle = \Delta \mu_1 ,$$

where $\Delta = |\mu_{10}, \mu_{11}, ..., \mu_{2q}|$.

Similarly for third order terms we find
The Distribution of Moment Estimators

11.

\[ n^2 \mathbb{E} \left\{ \frac{\partial}{\partial \gamma} \lambda \right\} = \mathbb{E} \left\{ \lambda \frac{\partial}{\partial \gamma} \right\} - \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} \\
+ 2 \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} + 2 \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} , \quad (22) \]

\[ n^2 \mathbb{E} \left\{ \frac{\partial}{\partial \gamma} \lambda \right\}^2 = \mathbb{E} \left\{ \lambda \frac{\partial}{\partial \gamma} \right\}^2 - \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} ^2 , \quad (23a) \]

\[ n^2 \mathbb{E} \left\{ \frac{\partial}{\partial \gamma} \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} = \mathbb{E} \left\{ \lambda \frac{\partial}{\partial \gamma} \right\} - 2 \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} \mathbb{E} \left\{ \lambda \right\} , \quad (23b) \]

Again without going into detail it will be found that

\[ n^3 \mathbb{E} \left\{ \frac{\partial}{\partial \gamma} \lambda \right\} = 3(n-1) \left( \Delta \bar{P}_1 \right)^2 + \mathbb{E} \left\{ \lambda \right\} ^2 , \quad (24a) \]

\[ n^3 \mathbb{E} \left\{ \frac{\partial}{\partial \gamma} \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} ^2 = -3(n-1) \Delta \bar{P}_1 \left( \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} \right) + \mathbb{E} \left\{ \lambda \right\} ^3 \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} \mathbb{E} \left\{ \lambda \right\} - \mathbb{E} \left\{ \lambda \right\} ^3 \left( \mathbb{E} \left\{ \lambda \right\} \frac{\partial}{\partial \gamma} \mathbb{E} \left\{ \lambda \right\} \right) . \quad (24b) \]

4. Moments of \( \theta \)

4.1 To work out the third and fourth standardized cumulants \( \gamma_1 = K_3 / K_2^{3/2} \), \( \gamma_2 = K_4 / K_2^2 \) to order \( n^{-1} \) and \( n^{-1} \) respectively, we require the expectations of the following terms:

**Linear** \( \Phi_2 \);

**Quadratic** \( \Phi_1, \Phi_1 \Phi_2, \Phi_2^2, \Phi_1 \Phi_3 \);

**Cubic** \( \Phi_3, \Phi_2 \Phi_2 \);

**Quartic** \( \Phi_4, \Phi_3 \Phi_2, \Phi_2 \Phi_2^2, \Phi_1 \Phi_3 \).
Omitting the algebra we quote some of the results in Table 1.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Expected Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi_1^2$</td>
<td>$\Delta \frac{\bar{h}}{n^2 \bar{h}^2} + O(n^3)$</td>
</tr>
<tr>
<td>$\Phi_2^2$</td>
<td>$\frac{\Delta \bar{h}^4}{n^2 \bar{h}^4} + \frac{3 \Delta^2}{4 n^2 \bar{h}^4} \langle \frac{\partial^2 \mu}{\partial \theta^2} \rangle + O(n^3)$</td>
</tr>
<tr>
<td>$\Phi_2^2$</td>
<td>$\frac{3 (n-1) \Delta^2}{n^2 \bar{h}^2} + \frac{\langle \frac{\partial M^4}{\partial \theta^2} \rangle}{n^3 \bar{h}^4}$</td>
</tr>
<tr>
<td>$\Phi_1^2 \Phi_2$</td>
<td>$\frac{-3 \Delta^2 \bar{h}}{2 n^2 \bar{h}^4} - \frac{3 \Delta^2}{n^2 \bar{h}^4} \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle + O(n^3)$</td>
</tr>
<tr>
<td>$\Phi_1^2 \Phi_2$</td>
<td>$\frac{-\langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle^3 \Delta}{n^3 \bar{h}^5} \left{ 5 \langle \frac{\partial M \partial \mu}{\partial \theta^2} \rangle + 6 \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle \right}$</td>
</tr>
<tr>
<td>$\Phi_1^2 \Phi_2$</td>
<td>$\frac{-6 \Delta}{n^3 \bar{h}^4} \left{ \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle^2 - \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle \right} + O(n^4)$</td>
</tr>
<tr>
<td>$\Phi_1^2 \Phi_2$</td>
<td>$\frac{15 \langle \frac{\partial M \partial \mu}{\partial \theta^2} \rangle^2 + 3 \frac{\partial^2 \bar{h}^2}{\Delta}}{n^2 \bar{h}^5} + O(n^4)$</td>
</tr>
</tbody>
</table>

(In this table $\bar{h}^2 = \frac{\bar{h}^2}{n} \left\{ \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle - \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle \right\} - \Delta \langle \frac{\partial M \partial \mu}{\partial \theta \theta} \rangle ^2 \).
4.2 The moments

\[ \mu^*_2(\theta_a) = \frac{\Delta}{n^3 h_1} + \frac{1}{n^2 h_1} \left\{ 3 \Delta^2 \bar{T} + \frac{1}{2} \Delta^2 \left( \frac{\partial^2 \bar{M}_2}{\partial \theta^2} \right) + 3 \Delta^2 \left( \frac{\partial \bar{M}_2}{\partial \theta} \right) \left( \frac{\partial \bar{M}_2}{\partial \theta} \right) \right\} + 2 \Delta \bar{h}_1^3 + O(n^{-3}) \]  

\[ \mu^*_3(\theta_a) = -\left( \frac{\partial \bar{M}_3}{\partial \theta} \right) + \frac{3 \Delta^2}{n^3 h_1^3} \left( \frac{\partial \bar{M}_3}{\partial \theta} \right) + O(n^{-3}) \]  

\[ \mu^*_4(\theta_a) = \frac{3(n-1) \Delta^4}{n^3 h_1^4} + \frac{\partial \bar{M}_4}{\partial \theta} + \frac{45 \Delta^3}{n^3 h_1^3 \bar{h}_1^3} \left( \frac{\partial^3 \bar{M}_2}{\partial \theta^3} \right) + \frac{30 \Delta^3}{n^3 h_1^3 \bar{h}_1^3} \left( \frac{\partial^2 \bar{M}_2}{\partial \theta^2} \right) \left( \frac{\partial \bar{M}_2}{\partial \theta} \right) + \frac{30 \Delta^2}{n^3 h_1^3 \bar{h}_1^3} \bar{T} \]  

\[ - \frac{2 \Delta^2}{n^3 h_1^3} \left\{ \left( \frac{\partial \bar{M}_3}{\partial \theta} \right)^2 + \frac{\partial \bar{M}_3}{\partial \theta} \right\} \]  

\[ + \frac{10 \Delta^2}{n^3 h_1^3} \left\{ 3 \left( \frac{\partial \bar{M}_2}{\partial \theta} \right)^2 \left( \frac{\partial \bar{M}_2}{\partial \theta} \right) + \left( \frac{\partial \bar{M}_2}{\partial \theta} \right)^3 \right\} + O(n^{-4}) \]  

4.3 We now find as first approximations to the standardized cumulants
The Distribution of Moment Estimators

\[ \gamma_1(\theta_q) \sim \frac{3 \Delta^2 \left< M \frac{\partial^2}{\partial \theta^2} \right> - \left< M M \frac{\partial^2}{\partial \theta^2} \right>}{\sqrt{(n \Delta^3 h_1^3)}}, \quad (26a) \]

\[ n \gamma_2(\theta_q) \sim \frac{\left< M M \frac{\partial}{\partial \theta} \right>^4}{\Delta^2 h_1^2} - \frac{12 \left< M M \frac{\partial^2}{\partial \theta^2} \right> \left\{ \left< M \frac{\partial^2}{\partial \theta^2} \right> + \left< M M \frac{\partial^2}{\partial \theta^2} \right> \right\}}{\Delta^2 h_1^2} + 12 \Delta \frac{\Delta^3 h_1^3}{h_1^3} \]

\[ + \frac{2 \Delta \left< M M \frac{\partial}{\partial \theta} \right>^2}{\Delta^2 h_1^2} - \frac{12 \left< M M \frac{\partial}{\partial \theta} \right> \left< M M \frac{\partial}{\partial \theta} \right>}{\Delta^2 h_1^2} + \frac{12 \left< M M \frac{\partial}{\partial \theta} \right> \left< M M \frac{\partial}{\partial \theta} \right>}{\Delta^2 h_1^2} - \frac{12 \left< M M \frac{\partial}{\partial \theta} \right> \left< M M \frac{\partial}{\partial \theta} \right>}{\Delta^2 h_1^2} + q, \quad (26b) \]

with a clear indication of asymptotic normality.

If \( P(x, \theta) \) is linear in \( \theta \) (a case which would arise for a Gram-Charlier distribution), the parameter governing the corrective term there is a considerable simplification in the moments, and indeed

\[ \mu_1 = \theta + O(n^{-2}), \quad (27a) \]

\[ \mu_2 \sim \frac{\Delta}{n h_1} + \frac{1}{n^2 h_1} \left\{ 3 \Delta \Psi - 2 \left< M \frac{\partial}{\partial \theta} \right> \left< M M \frac{\partial}{\partial \theta} \right> + 2 \Delta^3 h_1^3 \right\}, \quad (27b) \]

\[ \gamma_1(\theta_q) \sim \frac{\left< M M \frac{\partial}{\partial \theta} \right>^3}{\sqrt{(n \Delta^3 h_1^3)}}, \quad (27c) \]

\[ n \gamma_2(\theta_q) \sim \frac{\left< M M \frac{\partial}{\partial \theta} \right>^4}{\Delta^2 h_1^2} - \frac{12 \left< M M \frac{\partial}{\partial \theta} \right> \left< M M \frac{\partial}{\partial \theta} \right>}{\Delta^2 h_1^2} + \frac{12 \Delta \Psi}{\Delta^2 h_1^2} \]

\[ - \frac{12 \left< M M \frac{\partial}{\partial \theta} \right> \left< M M \frac{\partial}{\partial \theta} \right>}{\Delta^2 h_1^2} + q. \quad (27d) \]
It is of some interest to observe that the first order terms in $E(\theta_q)$ vanishes when the probability is linear in $\theta$; according to Haldane and Smith (1956) a similar property holds for maximum likelihood estimators in the case of linearity. We should expect this to be the case since (27a) holds for moment estimators of all orders (i.e. $q=1,2,...$). Haldane has suggested that estimators with this property should be described as being almost unbiased.

5. Maximum Likelihood Estimators.

If there is a moment estimator which gives the maximum likelihood solution (this may be the case for finite $q$, and in any event for $q \to \infty$ under certain conditions), then the expressions in (25) - (27) can be given in terms of the expected value of derivatives of $\log P(x, \theta)$. Thus (using $\overset{\cdot}{\cdot}$ to indicate correspondence)

$$- \left< \frac{d M_0}{d \theta} \right> / \Delta \overset{\cdot}{\cdot} E \frac{1}{p^2} \left( \frac{\partial P}{\partial \theta} \right)^2$$

(28a)

$$- \left< \frac{d^2 M_0}{d \theta^2} \right> / \Delta \overset{\cdot}{\cdot} E \frac{1}{p^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2}$$

(28b)

$$\left< \frac{\partial^2 M}{\partial \theta^2} \right> \overset{\cdot}{\cdot} E \left( \frac{\partial^2 \log P}{\partial \theta^2} - \frac{\partial \log P}{\partial \theta} \left( \frac{\partial^2 P}{\partial \theta^2} \right) + \frac{\partial \log P}{\partial \theta} \left( \frac{\partial^2 P}{\partial \theta^2} \right)^2 \right)$$

(29)
The derivation of most of these is straightforward, but for (28) we note that

$$\frac{\partial \log P}{\partial \theta} = - \frac{\partial^2}{\partial \theta} \left( \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)$$

where $\langle M \Theta \rangle = M_0 + M_1 \Theta + M_2 \Theta^2 + ...$

and (29) is derived by differentiating (27a), and (30) from (27b). For the sampling moments of the maximum likelihood estimate $\hat{\theta}$ we therefore find (after some non-trivial algebra)

$$\mathcal{E}(\hat{\theta}) = \hat{\theta} - \frac{\mathcal{E} \left( \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)}{2n \mathcal{I}^2}$$

$$\sigma^2(\hat{\theta}) = \frac{n-1}{n^2 \mathcal{I}} + \frac{\mathcal{E} \left( \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)^2 - \frac{1}{P^2} \frac{\partial^2 P}{\partial \theta^2} \frac{\partial^2 P}{\partial \theta^2} - \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^3 P}{\partial \theta^3} \frac{\partial^2 P}{\partial \theta^2}}{n^2 \mathcal{I}^3}$$

$$+ \frac{1}{2} \left( \frac{\mathcal{E} \left( \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)^2 - \left( \frac{\mathcal{E} \left( \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)^2 \right.}{\mathcal{I}^4} \right) + O(n^{-3}),$$

where

$$\mathcal{E} = \mathcal{E} \left( \frac{1}{P^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)$$
The Distribution of Moment Estimators

\[ \gamma_1(\hat{\theta}) = \frac{\kappa (\log P)^3}{\sqrt{n I^3}} - 3 \kappa \left( \frac{1}{p^2} \frac{\partial^2 P}{\partial \theta^2} \right) + O(n^{-3/2}), \quad (33c) \]

\[ n \gamma_2(\hat{\theta}) = -3 + \frac{\kappa (\log P)^4}{I^2} + \left\{ 24 \left( \frac{1}{p^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right)^2 - 12 \kappa \left( \frac{\log P}{p^2} \right) \frac{1}{p^2} \frac{\partial^3 P}{\partial \theta^3} \right\} / I^3 + O(n^{-1}), \quad (33d) \]

where \( I = \kappa (\log P)^2 \).

These expressions agree exactly (allowing for notational differences) with those given in (6) p. 100 of Haldane and Smith (1956). It is to be remarked that the forms given are not necessarily the simplest, and that there are alternatives for such terms as \( \kappa \left( \frac{1}{p^2} \frac{\partial P}{\partial \theta} \frac{\partial^2 P}{\partial \theta^2} \right) \) in terms of the expected value of powers of \( \partial \log P \) or their derivatives (in this connection reference may be made to Bartlett (1953) who gives an ingenious method of deriving such identities).

6. Illustrative Example.

In Shenton (1958) the asymptotic variances are given for the first three moment estimators of \( \theta \) in

\[ P_x = e^{-x} \left\{ \frac{x}{2} - \theta + (\theta - 1)x \right\}, \quad (0 < x < 0, 1 < \theta < 2). \]
Using the results in (27) and (33) we find for \( \gamma_1 \) and \( \gamma_2 \) for the first three moment estimators and the maximum likelihood estimator the following:

\[
\begin{align*}
\langle n \gamma_1(\theta) \rangle & \sim \frac{2\theta^3 - 12\theta^2 + 24\theta - 12}{(-\theta + 4\theta - 2)^{3/2}}, \\
\langle n \gamma_2(\theta) \rangle & \sim \frac{-6\theta + 48\theta^3 - 144\theta^2 + 192\theta - 84}{(-\theta + 4\theta - 2)^2}; \\
\langle n \gamma_3(\theta) \rangle & \sim \frac{128\theta^6 - 864\theta^5 + 2184\theta^4 - 2606\theta^3 + 1476\theta^2 - 300\theta - 16}{(\alpha^2)^{3/2}}, \\
\langle n \gamma_4(\theta) \rangle & \sim \frac{-6144\theta^9 + 59904\theta^8 - 243456\theta^7 + 568896\theta^6 - 901944\theta^5 + 1064706\theta^4 - 934512\theta^3 + 512\theta^2}{(\alpha^2)^2}.
\end{align*}
\]

where \( \alpha = 4\theta - 3, \beta = -4\theta + 15\theta^2 - 12\theta + 2; \)

\[
\begin{align*}
\langle n \gamma_1(\theta) \rangle & \sim p/(qr), \\
\langle n \gamma_2(\theta) \rangle & \sim p'/(q^3r^2),
\end{align*}
\]

where

\[
\begin{align*}
p &= 6750\theta^9 - 5130\theta^8 + 5040\theta^7 - 411280\theta^6 + 689472\theta^5 - 854400\theta^4 + 738688\theta^3 - 411280\theta^2 + 512\theta + 180, \quad q = 156\theta - 206 + 6, \\
r &= -156\theta^3 + 48\theta^2 + 8, \\
p' &= -4556250\theta^9 + 522450000\theta^8 - 283181000\theta^7 + 14732323200\theta^6 \\
& \quad - 7531774380\theta^5 + 28692814224\theta^4 - 75171819960\theta^3 \\
& \quad + 137243452800\theta^2 - 178312654848\theta + 166526475264\theta^5 \\
& \quad - 111281505024\theta^4 + 52015042560\theta^3 - 16167313152\theta^2 \\
& \quad + 30027417600 - 252062208.
\end{align*}
\]
\begin{align}
\sqrt{n} \gamma(\hat{\theta}) &\sim (\alpha+1) \left\{ (\alpha+4) I - (\alpha+1)^2 / \alpha \right\}^{3/2} / I, \\
\sim &\sim -1 / \left\{ \alpha \log \frac{1}{\alpha} \right\}^{3/2}, \quad (\theta \text{ nearly } 2) \\
\gamma^2(\hat{\theta}) &\sim -3 + (\alpha+1)^2 \left\{ (\alpha+3)(\alpha+7) I - (\alpha+1)^2 (\alpha^2 + 6 \alpha - 1) / \alpha^2 \right\} / (2 I^2), \\
\sim &\sim 1 / \left\{ 2 \alpha^2 \log \frac{1}{\alpha} \right\}^{2}, \quad (\theta \text{ nearly } 2),
\end{align}

where \( \alpha = \frac{2-\theta}{\theta-1} \), \( I = - (\alpha+1)^2 - (\alpha+1)^3 \, \text{Ei}(-\alpha), \text{Ei}(-x) = \int_{-x}^{\infty} \frac{e^{-u}}{u} \, du \).

It will be fairly evident from (34) that an investigation of the sampling distributions of moment estimators such as \( \Theta \) (or higher estimators) would be a considerable undertaking. An impression of the situation for this example is given in Figures 1-3. The gain in using higher estimators is evident from a glance at the variances in Figure 1. As for \( \gamma_1 \) and \( \gamma_2 \), the approach of the second and third estimators to the maximum likelihood estimator is adumbrated. The critical value of the parameter is \( \Theta = 2 \), for in this case the integrals appearing in the moments of the maximum likelihood estimator diverge.

* I am indebted to Mr. A. Fletcher for assisting in the construction of these diagrams.
Asymptotic Variance of Estimators of \( e \) in \( P = QA^2 - A + (e - 1) \cdot L \)

Maximum Likelihood

Linear Estimator

Cubic Estimator

Quadratic Estimator

Asymptotic Variance of Estimators of \( e \) in \( P = e^2 - \theta^2 + (\theta - 1) \cdot x \)
Several Estimators

Maximum Likelihood

\[ \sqrt{n} \gamma(\theta) \]
Several Estimators

\[ P = \frac{e^{1.5}}{4} \]

1.5
1.2
10
.8
.7
2-

\[ e + b \rightarrow x \]

Maximum Likelihood

MEP

OOP

2.0
A Determinantal Expansion for a Class of Definite Integral$^1$. Part 1

By L. R. Shenton

(Received 22nd October, 1949. Read 4th November, 1949. Revised M.S. received 29th September, 1952.)

1. Let $w(x)$ be a non-negative weight function for the finite interval $(a, b)$ such that $\int_a^b w(x)\, dx$ exists and is positive, and let $T_r(x)$, $r = 0, 1, 2, \ldots$ be the corresponding orthonormal system of polynomials. Then if $F(x)$ is continuous on $(a, b)$ and has "Fourier" coefficients

$$c_r = \int_a^b F(x) T_r(x) w(x)\, dx, \quad r = 0, 1, 2, \ldots,$$

Parseval's formula$^2$ gives

$$\int_a^b w(x) [F(x)]^2\, dx = \sum_{r=0}^{\infty} c_r^2. \quad (1)$$

We shall show that for the weight function $w(x) C(x)$ and $F(x) = A(x)/C(x)$, both satisfying the conditions above, $\sum_{r=0}^{\infty} c_r^2, \quad s = 0, 1, 2, \ldots,$ of Parseval's formula takes the form of a ratio of two determinants. The successive values of this determinantal ratio will be shown to provide a sequence of convergent approximants to the value of the integral. Moreover in the case when $C(x)$ is a polynomial, an expansion of the integral is given in terms of the roots of $C(x)$. The particular case when $C(x)$ is linear indicates the relation of the present method to the expression of integrals of the type

$$\int \frac{w(x)}{x+z}\, dx$$
as continued fractions. The case when the range of integration is infinite is to be treated in Part 2.

2. Let $\theta_r(x) = \sum_{s=0}^{r} a_{rs} x^s$, $a_{rr} \neq 0$, be a polynomial of degree $r$ in $x$, and

$$\int_a^b \theta_r(x) \theta_s(x) w(x) C(x)\, dx = \gamma_{rs} = \gamma_{sr}, \quad (2)$$

$^1$ Applications have been given in Biometrika, 37 (1950), 111 and 38 (1951), 58.

$^2$ See for example D. Jackson, Fourier Series and Orthogonal Polynomials (Carus Math. Mon., 1941), Ch. II, and p. 228, or G. Szegö, Orthogonal Polynomials (New York, 1939), Ch. III.
DETERMINANTAL EXPANSION FOR A CLASS OF DEFINITE INTEGRAL 45

where it is assumed that \( w(x) C(x) \) is non-negative and \( \int_a^b w(x) C(x) \, dx \) exists and is positive.

Further let \( p_s(x) = \sum_{s=0}^r p_s \theta_s(x) \) be an orthonormal system associated with the weight function \( w(x) C(x) \) on \((a, b)\), so that

\[
p_r(x) = \begin{vmatrix}
\theta_0(x) & \theta_1(x) & \ldots & \theta_r(x) \\
\gamma_{0,0} & \gamma_{0,1} & \ldots & \gamma_{0,r} \\
\gamma_{1,0} & \gamma_{1,1} & \ldots & \gamma_{1,r} \\
\vdots & \vdots & \ddots & \vdots \\
\gamma_{r-1,0} & \gamma_{r-1,1} & \ldots & \gamma_{r-1,r}
\end{vmatrix} \div \sqrt{(\Delta_{r-1} \Delta_r)},
\]

where

\[
\Delta_r = |\gamma_{0,0}, \gamma_{1,1}, \ldots, \gamma_{r,r}|.
\]

Hence the Fourier coefficients for \( A(x)/C(x) \), assumed continuous on \((a, b)\), are given by

\[
A_s = \int_a^b \frac{A(x) w(x) C(x)}{C(x)} p_s(x) \, dx \quad (s = 0, 1, 2, \ldots)
\]

\[
= \int_a^b A(x) w(x) \frac{\theta_s(x)}{\sqrt{(\Delta_{s-1} \Delta_s)}} \, dx
\]

\[
= |a_0, \gamma_{0,1}, \gamma_{1,2}, \ldots, \gamma_{s-1,s}| \div \sqrt{(\Delta_{s-1} \Delta_s)},
\]

where

\[
a_r = \int_a^b \theta_r(x) w(x) A(x) \, dx \quad (r = 0, 1, 2, \ldots).
\]

Hence, using Parseval's theorem, we find

\[
\int_a^b \frac{[A(x)]^2 w(x)}{C(x)} \, dx = \sum_{s=0}^\infty \frac{|a_0, \gamma_{0,1}, \gamma_{1,2}, \ldots, \gamma_{s-1,s}|^2}{\Delta_{s-1} \Delta_s}
\]

which, by Schweins' theorem on determinants \(^2\), may be written as

\[
\lim_{s \to \infty} \begin{vmatrix}
0 & a_0 & a_1 & \ldots & a_s \\
a_0 & \gamma_{0,0} & \gamma_{0,1} & \ldots & \gamma_{0,s} \\
a_1 & \gamma_{1,0} & \gamma_{1,1} & \ldots & \gamma_{1,s} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_s & \gamma_{s,0} & \gamma_{s,1} & \ldots & \gamma_{ss}
\end{vmatrix} \div |\gamma_{0,0}, \gamma_{1,1}, \ldots, \gamma_{ss}|.
\]

---

\(^1\) See Szegö, loc. cit., Ch. II.

In general the partial sums of the series (6) and the corresponding part of the determinantal ratio (7) form non-decreasing sequences. A more general result is found by applying Parseval’s formula to the functions

\[ \{A(x) \pm B(x)\}/C(x) \]

assumed continuous on \((a, b)\), whence

\[
\int_a^b \frac{A(x)B(x)w(x)}{C(x)} \, dx = \sum_{s=0}^{\infty} \Delta_{s-1} \Delta_s \begin{vmatrix} d_0 & \gamma_{01} & \gamma_{12} & \cdots & \gamma_{s-1, s} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_0 & \gamma_{00} & \gamma_{01} & \cdots & \gamma_{0s} \\ \beta_1 & \gamma_{10} & \gamma_{11} & \cdots & \gamma_{1s} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \beta_s & \gamma_{s0} & \gamma_{s1} & \cdots & \gamma_{ss} \end{vmatrix} - \lim_{s \to \infty} \frac{0}{\beta_s} \frac{\gamma_{ss}}{\gamma_s},
\]

in which \(\beta_s\) is given by (5) with \(B(x)\) replacing \(A(x)\).

It is convenient to use a matrix notation in (9) and write

\[
\int_a^b \frac{A(x)B(x)w(x)}{C(x)} \, dx = -\lim_{s \to \infty} \frac{\begin{vmatrix} 0 & \alpha_0 & \alpha_1 & \cdots & \alpha_s \\ \beta_0 & \gamma_{00} & \gamma_{01} & \cdots & \gamma_{0s} \\ \beta_1 & \gamma_{10} & \gamma_{11} & \cdots & \gamma_{1s} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \beta_s & \gamma_{s0} & \gamma_{s1} & \cdots & \gamma_{ss} \end{vmatrix}}{\beta_s \gamma_s}.
\]

The special case \(A(x) = B(x) = 1\), \(C(x) = x+z\) is known in the theory of continued fractions\(^1\) in which case, with \(\beta_s(x)\) suitably restricted, the determinants are of “continuant” type. The formula (10) is thus seen as an extension of this type of continued fraction. Since the convergents of a continued fraction satisfy a second order difference equation, a generalised continued fraction might be one for which the “convergents” satisfied a difference equation of the \(n\)-th order, this suggestion being given by Fürstenau in “Über Kettenbrüche höherer Ordnung”\(^2\). The determinants in (10) do not in general appear to satisfy any simple difference equation. It is of interest to note that Rogers\(^3\), in representing certain definite integrals as continued fractions, suggested that some form of algebraic fraction might exist for cases, such as \(\int_0^\infty \frac{t^n e^{-xt}}{e^t-1} \, dt\), which were intractable by his method.

\(^1\) See for example H. S. Wall, Continued Fractions (New York, 1948), Ch. XIII onwards, or O. Perron, Die Lehre von den Kettenbrüchen (Leipzig, 1913), Ch. 9.


3. We now turn to a further use of formula (10). Consider the determinant whose elements are definite integrals

\[
\Delta = \left| \int_a^b \frac{A_j(x) A_k(x) w(x) \, dx}{C(x)} \right|_n \quad (j, k = 1, 2, \ldots, n),
\]

(11)
in which \(A_j(x), j = 1, 2, 3, \ldots, n\), are continuous functions over \((a, b)\) and \(w(x) C(x)\) satisfies the same conditions as for (10). We may approximate \(\Delta\) by replacing each integral by the corresponding ratio (10) with the same \(s\), and using the notation

\[
\begin{bmatrix} \alpha_k \end{bmatrix} = \begin{bmatrix} \alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_s \end{bmatrix}
\]

to obtain

\[
\Delta = (-1)^n \lim_{s \to \infty} \left| \begin{array}{c} 0 \\ \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \vdots \\ \alpha_s \end{bmatrix} \\ \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \vdots \\ \gamma_s \end{bmatrix} \end{array} \right|^n \quad (j, k = 1, 2, \ldots, n),
\]

which, by an "extensional" identity in determinants, leads to

\[
\Delta = \left| \int_a^b \frac{A_j(x) A_k(x) w(x) \, dx}{C(x)} \right|_n
\]

(12)

\[
= (-1)^n \lim_{s \to \infty} \left| \begin{array}{cccc} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 \alpha_0 & 2 \alpha_0 & \cdots & n \alpha_0 \\ 1 \alpha_1 & 2 \alpha_1 & \cdots & n \alpha_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 \alpha_s & 2 \alpha_s & \cdots & n \alpha_s \end{array} \right| \left| \begin{array}{cccc} \gamma_0 & \gamma_0 & \cdots & \gamma_0 \\ \gamma_1 & \gamma_0 & \cdots & \gamma_0 \\ \gamma_1 & \gamma_1 & \cdots & \gamma_0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_s & \gamma_s & \cdots & \gamma_0 \end{array} \right|
\]

(13)

If the \(A_j(x), j = 1, 2, \ldots, n\), are linearly independent, then \(\Delta\) is positive. This follows from the fact that the quadratic form

\[
\sum_{j=1}^n \sum_{k=1}^n u_j u_k \int_a^b \frac{A_j(x) A_k(x) w(x) \, dx}{C(x)} = \int_a^b \left( \sum_{j=1}^n u_j A_j(x) \right)^2 \frac{w(x) \, dx}{C(x)}
\]

is positive definite. That a determinant with definite integral elements similar to \(\Delta\) is positive appears to be due to Kowalewski\(^2\). We thus see

\(^{1}\) Aitken, loc. cit.

\(^{2}\) G. Kowalewski, *Einführung in die Determinantentheorie* (Leipzig, 1925), 224.
that under the conditions attached to $A_j(x)$ and $w(x) C(x)$, the bordered determinant in (13) has the same sign as $(-1)^n$. It may be remarked that in certain cases the numerator of (13) reduces to a multiple of
\[ |y_{00}, y_{11}, \ldots, y_{nn}| \]
with a certain number of rows and corresponding columns deleted.

4. There is an alternative form for the expansion given in (8) when $C(x)$ is a polynomial. Let $C(x) = k \prod_{\lambda=1}^{n} (x-x_\lambda)$. Then by Christoffel's theorem (Szegö, loc. cit., 2.5), if $p_r(x)$ are orthonormal polynomials with respect to the weight function $w(x)$ on $(a, b)$, the orthogonal set with respect to $C(x) w(x)$ is $q_r(x)$ where
\[ q_r(x) = \begin{vmatrix} p_r(x) & p_{r+1}(x) & \cdots & p_{r+n}(x) \\ p_r(x_1) & p_{r+1}(x_1) & \cdots & p_{r+n}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_r(x_n) & p_{r+1}(x_n) & \cdots & p_{r+n}(x_n) \end{vmatrix} / C(x). \quad (14) \]

If $C(x)$ has a root of multiplicity $m$ at $x_k$ then the corresponding rows in (14) are to be replaced by the $0, 1, 2, \ldots, m-1$-th derivatives of $p(x)$ at $x_k$.

We further require a theorem of Darboux (Szegö, loc. cit., 3.2):
\[ \frac{p_r(x) p_{r+1}(x)}{p_r(y) p_{r+1}(y)} \bigg| \frac{d}{dy}(x-y) = -\frac{k_{r+1}}{k_r} \sum_{s=0}^{r} p_s(x) p_s(y), \quad (15) \]
where the recurrence relation for the polynomials $p_r(x)$ is
\[ p_r(x) = (xA_r + B_r) p_{r-1}(x) - C_r p_{r-2}(x) \quad (16) \]
with $A_r = k_r / k_{r-1}$, $C_r = A_r / A_{r-1}$, and where $k_r$ is the highest coefficient in $p_r(x)$. An extension of this is found by using the recurrence relation in (14), namely
\[ q_r(x) = (-1)^n \frac{k_{r+n}}{k_r} \sum_{s=0}^{r} |p_s(x_1), p_{r+1}(x_1), \ldots, p_{r+n}(x_1)| p_s(x) \quad (17) \]
with derivatives appearing in the rows of the determinants when $C(x) = 0$ has multiple roots. Using (14) and (17), we have
\[ \int_{a}^{b} q_r^2(x) C(x) w(x) dx = \int_{a}^{b} \frac{|p_r(x), p_{r+1}(x_1), \ldots, p_{r+n}(x_n)|}{C(x)} (-1)^n \frac{k_{r+n}}{k_r} \]
\[ \times \sum_{s=0}^{r} |p_s(x_1), p_{r+1}(x_1), \ldots, p_{r+n}(x_1)| C(x) p_s(x) w(x) dx \]
\[ = (-1)^n \frac{k_{r+n}}{k_r} |p_r(x_1), p_{r+1}(x_1), \ldots, p_{r+n}(x_1)| \cdot |p_{r+1}(x_1), p_{r+2}(x_1), \ldots, p_{r+n}(x_1)| \]
\[ = \phi_r \text{ say.} \quad (18) \]
If now $A(x)$ and $B(x)$ are polynomials of degree $L$ and $M$ respectively, we may write

$$A(x) = \sum_{\lambda=0}^{L} a_{\lambda} p_{\lambda}(x), \quad B(x) = \sum_{\lambda=0}^{M} b_{\lambda} p_{\lambda}(x).$$

Hence the Fourier coefficients of $A(x)/C(x)$ with respect to the orthogonal set $q_r(x)$ and the weight function $C(x)w(x)$ are given by

$$a_r, r = \int_{a}^{b} A(x)q_r(x)w(x)dx$$

$$= (-1)^n \sum_{\lambda=0}^{r} a_{\lambda} p_{\lambda}(x_1), p_{r+1}(x_2), p_{r+2}(x_3), \ldots, p_{r+n-1}(x_n)$$

with a similar expression for $B(x)$. In the expression $\sum_{\lambda=0}^{r} a_{\lambda} p_{\lambda}(x_1)$ it is to be understood that $a_\lambda = 0, \lambda > L$, and similarly, in $\sum_{\lambda=0}^{r} b_{\lambda} p_{\lambda}(x_1), b_\lambda = 0, \lambda > M$.

Hence if $A(x)/C(x), B(x)/C(x), w(x)/C(x)$ satisfy the conditions of Parseval's theorem, we have from (18) and (19)

$$\int_{a}^{b} A(x)B(x)w(x)dx \quad C(x)$$

$$= (-1)^n \sum_{\lambda=0}^{r} b_{\lambda} p_{\lambda}(x_1), p_{r+1}(x_2), p_{r+2}(x_3), \ldots, p_{r+n-1}(x_n)$$

$$= (-1)^n \sum_{\lambda=0}^{r} k_{r+n} p_{r}(x_1), p_{r+1}(x_2), \ldots, p_{r+n-1}(x_n) \quad (20)$$

If (20) is compared with (8) and (9) it will be observed that (a) the determinants in (20) are all of order $n$ whereas in (8) the order increases with the term, (b) whereas the partial sums of (8) may be expressed as (9) by Schweins' theorem, this does not appear to be the case with the partial sums of (20). It is however clear that there will be determinantal identities between the denominators $\Delta_{r-1}$ of (8) and $|p_\lambda(x_1), p_{r+1}(x_2), \ldots, p_{r+n-1}(x_n)|$ of (20). The special case $A(x) = B(x) = 1, C(x) = x + z (z \text{ real})$ gives the expansion

$$\int_{a}^{b} w(x)dx = \int_{a}^{b} x + z = (-1) \sum_{\lambda=0}^{\infty} k_{r+1} p_{\lambda}(-z) p_{r+1}(-z).$$

5. The relation of (21) to the corresponding continued fraction expansion appears from (7). For (21) is an example of (7) with $A(x) = 1, C(x) = x + z$ and $\theta_r(x) = p_r(x) \left\{ -\frac{k_{r+1}}{k_r p_r(-z) p_{r+1}(-z)} \right\}^{1/2}, p_r(x)$ being the orthonormal set with respect to $w(x)$. If however we take $\theta_r(x) = p_r(x)$ and use the recur-
rence relation (16), then \( x_0 = 1 \) and \( x_r = 0 \), \( r \neq 0 \). Moreover, writing \( x + z = k_1 \gamma + \delta + z \), where \( k_1 \neq 0 \), we find

\[
\gamma_{r,s} = (z-B_{r+1} A_{r+1}^{-1}) \delta_{r,s} + A_{r+1}^{-1} \delta_{r+1,s} + A_{r+1}^{-1} \delta_{r-1,s}, \quad (r, s = 1, 2, \ldots)
\]

and \( \gamma_{0,0} = z-B_1 A_1^{-1}, A_1 > 0 \).

From (7),

\[
\int_a^b w(x) \frac{dx}{x+z} = -\lim_{s \to \infty} N_s(z),
\]

where

\[
N_s(z) = \begin{vmatrix}
0 & 1 & 0 & 0 & \cdots \\
1 & z-B_1 A_1^{-1} & A_1^{-1} & 0 & \cdots \\
0 & A_1^{-1} & z-B_2 A_2^{-1} & A_2^{-1} & \cdots \\
0 & 0 & A_2^{-1} & z-B_3 A_3^{-1} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots \\
& & & & & z+1
\end{vmatrix}
\]

and \( D_s(z) \) is \( N_s(z) \) with the first row and column deleted. In other words we have the continued fraction expansion

\[
\int_a^b w(x) \frac{dx}{x+z} = \frac{1}{z-B_1 A_1^{-1} - z-B_2 A_2^{-1} - \cdots}.
\]

It is of some interest to notice another form for the expansion. Take \( \theta_r(x) = (x+z)^r \) so that \( x_r = \int_a^b (x+z)^r w(x) \frac{dx}{x+z} = m_r \), say, and is an Appell polynomial of degree \( r \) in \( z \). (These are treated by J. Geronimus, *Journal London Math. Soc.*, 6 (1931), 55.) Similarly for \( \gamma_{r,s} = m_{r+s+1} \). From (7), with the usual notation for persymmetric determinants,

\[
\int_a^b w(x) \frac{dx}{x+z} = -\lim_{s \to \infty} \frac{P(0, m_0, m_1, \ldots, m_{2s-1})}{P(m_1, m_2, \ldots, m_{2s-1})}.
\]

6. The expressions (21)-(24) indicate that there are relations between the various forms of the approximants to the definite integral. Consider

\[
\phi_r = \int_a^b (x+z) w(x) q_r^2(x) \frac{dx}{x+z},
\]

where \( \{q_r(x)\} \) is an orthogonal system with respect to \( (x+z) w(x) \), the coefficient of \( x^r \) in \( q_r(x) \) being unity. Then

\[
q_r(x) = (-)^r \theta_r^{-1} \theta_0(x), q_{10}, q_{11}, \ldots, q_{r-1,r}, \ldots, q_{r-1,1} \theta_r \frac{q_{r-1} \theta_0}{q_{r-1} \theta_0}, \ldots, q_{r-1,r} \theta_0 \frac{q_{r-1} \theta_0}{q_{r-1} \theta_0},
\]

where \( \theta_r(x) \) is an arbitrary polynomial of precise degree \( s \) with highest coefficient \( \theta_s \), and \( q_{s,s} = \int_a^b (x+z) q_s(x) \theta_s(x) w(x) \frac{dx}{x+z} \). But \( \phi_r \) is invariant with respect to the choice of \( \theta(x) \). Hence taking \( \theta_1(x) = (x+z)^r \), we find \( \phi_r = P(m_1, m_2, \ldots, m_{2s+1})/P(m_1, m_2, \ldots, m_{2s-1}) \). Again take \( \{\theta_s(x)\} \) to be
DETERMINANTAL EXPANSION FOR A CLASS OF DEFINITE INTEGRAL 51

the orthonormal set with respect to \( w(x) \), and use (14) and (15), so that
\[
\phi_r = - p_{r+1}(-z)/\{k_r k_{r+1} p_r(-z)\}.
\]
But from the recurrence relation (16) and the continued fraction expansion (23),
\[ p_s(-z) = (-)^s k_s D_s(z). \]
Hence
\[
P(m_1, m_2, ..., m_{2r+1}) = \frac{D_{r+1}(z)}{(k_0 k_1 ... k_r)^2}. \tag{25a}
\]
It therefore follows from (22) and (24) that
\[
P(0, m_0, m_1, ..., m_{2r+1}) = \frac{N_{r+1}(z)}{(k_0 k_1 ... k_r)^2}. \tag{25b}
\]
The relations (25) have been derived by Geronimus (loc. cit.) by another method.

7. As an illustration consider the hypergeometric function
\[
F(1, a; b; t) = \sum_{s=0}^{\infty} \frac{a(a+1) ... (a+s-1)}{b(b+1) ... (b+s-1)} \frac{\Gamma(b)}{\Gamma(a) \Gamma(b-a)} \int_0^1 \frac{w(x)}{1-xt} dx,
\]
where \( w(x) = x^{a-1}(1-x)^{b-a-1}, \quad 0 \leq x \leq 1, \quad a > 0, \quad b-a > 0, \)
and \( p_s(x) = \sqrt{r_s} F(s+b-1, -s; a; x), \)
\[
r_s = \frac{(2s+b-1) \Gamma(s+a) \Gamma(s+b-1)}{\Gamma(s+1) \Gamma(s+b-a) \Gamma(a)^2}, \quad k_s = (-)^s \frac{\Gamma(2s+b-1) \Gamma(a)}{\Gamma(s+a) \Gamma(s+b-1)}. \]
It follows from (21) with \( z = -t^{-1}, \quad |t| < 1, \) that
\[
F(1, a; b; t) = \sum_{s=1}^{\infty} \frac{\Gamma(b) \Gamma(s+a-1) \Gamma(s+b-2) \Gamma(s+b-a-1) \Gamma(s) t^{s-2}}{\Gamma(a) \Gamma(b-a) \Gamma(2s+b-2) \Gamma(2s+b-3) F(-s, 1-a-s; 2-b-2s; t) \times F(1-s, 2-a-s; 4-b-2s; t)}.
\tag{26}
\]
Moreover, using the recurrence relation for \( p_s(x) \) (see Szegö, loc. cit.; a slight change of notation is required in Szegö, 4.5.1) it will be found from (23) that \( F(1, a; b; t) \) is the even part of the continued fraction
\[
\frac{1}{1 - \frac{b_1 t}{1 - \frac{b_2 t}{1 - \frac{b_3 t}{1 - \cdots}}}}, \tag{27}
\]
where
\[
b_{2s+1} = \frac{(s+a)(s+b-1)}{(2s+b-1)(2s+b)}, \quad b_{2s+2} = \frac{(s+1)(s+b-a)}{(2s+b)(2s+b+1)}. \tag{27a}
\]

1 H. Bateman, Partial Differential Equations (New York, 1944), 392.
In a similar way the odd part of the continued fraction (27) arises from a consideration of the relation

\[ F(1, a; b; t) = 1 + \frac{t \Gamma(b)}{\Gamma(a) \Gamma(b - a)} \int_0^1 x w(x) dx, \]

and approximating to the integral by using \( x w(x) \) for \( w(x) \) in (23). There is also a corresponding series expansion similar to (26). If we call the \( s \)-th convergent of (27) \( n_s/d_s \), then the series expansion (26) may be derived from the identity

\[ \frac{n_s}{d_s} = b_1 b_2 \cdots b_{s-2} t^{s-2} \frac{d_{s-2} d_s}{d_{s-2} d_s}, \quad s = 2, 4, 6, \ldots \]

8. Finally, consider the relations between the persymmetric determinants and continued fraction convergents given by (25), in connection with the hypergeometric function. We find

\[ (k_0 k_1 \cdots k_r)^2 = \prod_{s=0}^{r} \frac{\Gamma(r+b+s)}{\Gamma(s+1) \Gamma(s+a) \Gamma(s+b-a)}, \]

\[ m_s = z^s \frac{\Gamma(a) \Gamma(b-a)}{\Gamma(b)} F(-s, a; b; -z^{-1}). \]

Inserting these in (25a) and (25b), we find that

\[ z^r P(F_1, F_2, \ldots, F_{2r-1}) = F(r+b-1, -r; a; -z) \prod_{s=0}^{r-1} \frac{s! (b-a)_s a_{s+1}}{b_{r+s}}, \]

\[ z^{r-1} P(0, F_0, F_1, \ldots, F_{2r-1}) = -\tilde{N}_r(z) \prod_{s=0}^{r-1} \frac{s! (b-a)_s a_s}{b_{r+s-1}}, \quad r = 1, 2, 3, \ldots, \]

where \( F_r = F(-a, a; b; -z^{-1}), \lambda_r = \lambda(\lambda+1) \cdots (\lambda+s-1) \) under the continued product sign, and \( \tilde{N}_r(z) \) is the numerator of the \( r \)-th convergent of

\[ \frac{1}{z+b_1 - z+b_2+b_3 - z+b_4+b_5 - \ldots}, \]

the \( b \)'s being given in (27a). A similar pair of relations would also be found by considering the weight function \( x w(x) \) in place of \( w(x) \). Burchnall\(^1\) has recently given similar expressions for \( P(F_0, F_1, \ldots, F_{2r-1}) \) and its minors.


\[ \text{College of Technology,}\]
\[ \text{Manchester, 1.} \]
A Determinantal Expansion for a Class of Definite Integral: Part 2.

Extracted from the
Edinburgh Mathematical Proceedings
Series 2—Vol. 10—Part II—December 1954
A Determinantal Expansion for a Class of Definite Integral
Part 2.

By L. R. Shenton.

(Received 24th April 1950. Read 3rd June 1950.
Revised M.S. received 18th August 1953.)

A. Introduction.

In a previous paper (Shenton, 1953) we have given an expansion for integrals of the form \( \int \frac{A(x)B(x)}{C(x)} w(x) \, dx \). This expansion may be expressed as a determinantal quotient or Schweinsian series. In the present paper we state more general terms under which the expansion holds and consider the case when the limits of integration are infinite and the weight function of the form \( A(x) e^{-x} \) or \( A(x) e^{-ix} \).

In particular we give expansions for \( \int_{0}^{\infty} e^{-ax} x^{n-1} \, dx \), the \( \Psi \) function, and \( \int_{-\infty}^{\infty} \frac{e^{-ix}}{C(x)} \, dx \), where \( C(x) \) is a positive polynomial.

We take this opportunity to remark that the method in this and the previous paper is closely related to the expansion of certain definite integrals as continued fractions. Indeed Tchebycheff (1859) uses an interpolation formula to give an expansion of a function in terms of orthogonal functions, these functions appearing as the denominators of the convergents of a continued fraction. As examples he gives

\[
\frac{1}{\pi} \int_{-1}^{1} \frac{du}{(x-u) \sqrt{1-u^2}} = \frac{1}{x} \frac{1}{2x} \frac{1}{3x} \cdots
\]

\[
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-ku} \, du}{x-u} = \frac{1}{\lambda x} \frac{1}{\lambda x} \frac{2}{\lambda x} \cdots, \lambda = \sqrt{(2k)},
\]

\[
\int_{0}^{\infty} \frac{e^{-ku} \, du}{x-u} = \frac{1}{kx-1} \frac{1^2}{kx-3} \frac{2^2}{kx-5} \cdots.
\]

The general method of expressing a definite integral as a continued fraction (i.e. a determinantal quotient of continuants) has been to
convert the integral into an infinite series, convergent or divergent, and to express this series as a continued fraction. With this procedure orthogonal polynomials appear in certain cases (Wall, 1945, pp. 192-202).

Romanovsky (1927) has treated Tchebycheff's method of interpolation and suggested that the interpolatory function might be used for points outside the range and for the case when the function is defined at an infinite number of points. The method we use is an extension of this and leads to a generalised type of continued fraction. Questions of convergence can be settled by an appeal to Parseval's theorem in the theory of orthogonal functions.

B. Parseval's Theorem.

We shall consider the formal expansion

\[ \int_a^b \frac{A(x)B(x)w(x)}{C(x)} \, dx = \sum_{s=1}^{\infty} \frac{a_0, \gamma_{01}, \gamma_{12}, \ldots, \gamma_{s-1, s}, \Delta_{s-1} \Delta_s}{\Delta_s} + \frac{\beta_0, \gamma_{01}, \gamma_{12}, \ldots, \gamma_{s-1, s}}{\Delta_{s-1} \Delta_s} \]

(1)

\[ = - \lim_{s \to \infty} \left[ \begin{array}{ccc} 0 & a_0 & a_1 & \cdots & a_s \\ \beta_0 & 0 & \beta_1 & \cdots & \beta_s \\ \gamma_{00} & \gamma_{01} & \gamma_{11} & \cdots & \gamma_{ss} \\ \gamma_{10} & \gamma_{11} & \gamma_{1s} & \cdots \end{array} \right] \]

(2)

where

\[ \begin{align*}
\alpha_s &= \int_a^b \theta_s(x)w(x)\frac{A(x)}{B(x)} \, dx, \\
\beta_s &= \int_a^b \theta_s(x)w(x) \, dx,
\end{align*} \]

(3)

\[ \gamma_{r,s} = \gamma_{s,r} = \int_a^b \theta_s(x) \theta_r(x)C(x)w(x) \, dx, \]

(4)

\[ \Delta_s \equiv \begin{vmatrix} \gamma_{00}, \gamma_{11}, \ldots, \gamma_{ss} \end{vmatrix}, \]

and

\[ p_s(x) = \begin{vmatrix} \theta_0(x) & \theta_1(x) & \cdots & \theta_s(x) \\ \gamma_{00} & \gamma_{01} & \cdots & \gamma_{0s} \\ \gamma_{10} & \gamma_{11} & \cdots & \gamma_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{s-1, 0} & \gamma_{s-1, 1} & \cdots & \gamma_{s-1, s} \end{vmatrix} \]

(1)

(2)

(3)

(4)

\[ \text{with } \theta_s(x) \text{ an arbitrary polynomial of precise degree } s. \text{ The set of polynomials } \{p_s(x)\} \text{ is an orthonormal system with respect to the} \]


weight function $C(x)w(x)$. In (2) we have introduced the notation

$$
\begin{vmatrix}
  a_0 & a_1 & a_2 \\
  b_0 & b_1 & b_2 \\
  c_0 & c_1 & c_2
\end{vmatrix} = \frac{|a_0, b_1, c_2|}{|b_1, c_2|},
$$

and so on for other orders. If $A(x), B(x)$ and $C(x)$ are polynomials of degrees $l, m$ and $n$ respectively, then there is the formal expansion

$$
\int_a^b \frac{A(x)B(x)w(x)}{C(x)} dx = \sum_{r=0}^{k_r+n} \sum_{\lambda=0}^r \frac{a_{\lambda} p_\lambda(x_1), p_{r+1}(x_2), \cdots p_{r+n-1}(x_n)}{p_{r+1}(x_1), p_{r+2}(x_2), \cdots p_{r+n}(x_n)},
$$

where $\{p_s(x)\}$ is an orthonormal set with respect to $w(x)$ on $(a, b)$, $k_s$ being the highest coefficient in $p_s(x)$, $C(x)$ has the roots $x_j, j=1, 2, \ldots, n$ (assumed distinct), and $A(x) = \sum_{\lambda=0}^r a_{\lambda} p_\lambda(x)$, $B(x) = \sum_{\lambda=0}^m b_{\lambda} p_\lambda(x)$.

In the expression $\sum_{\lambda=0}^r a_{\lambda} p_\lambda(x)$ it is to be understood that $a_{\lambda}=0$ if $\lambda > r$, and similarly in $\sum_{\lambda=0}^r b_{\lambda} p_\lambda(x), b_{\lambda}=0$ if $\lambda > m$.

We now consider the expansions (1), (2) and (6) in relation to Parseval's theorem, which may be stated as follows:

P. 1. Finite Range. Let

(i) $w(x)$ be a non-negative and measurable weight function such that $\int_a^b w(x) dx > 0$ and $\int_a^b x^nw(x) dx$ exists for $n = 0, 1, \ldots$,

(ii) $f(x) \sqrt{w(x)}$ be of the class $L^2(a, b)$,

(iii) $\{p_s(x)\} \sqrt{w(x)}$ be an orthonormal system with $p_s(x)$ a polynomial in $x$ of precise degree $s$.

Then

$$
\int_a^b |f(x)|^2 w(x) dx = \sum_{r=0}^\infty |f_r|^2,
$$

where

$$
\int_a^b f(x)p_r(x)w(x) dx = f_r.
$$

Similarly

$$
\int_a^b f(x)g(x)w(x) dx = \sum_{r=0}^\infty f_r g_r
$$

provided $g(x) \sqrt{w(x)}$ also belongs to $L^2(a, b)$.

---

A Determinantal Expansion for a Class of Definite Integral

P. 2. Range 0 to $\infty$. The theorem holds if the conditions of P. 1 are satisfied with the weight function $w(x) = e^{-x} \tilde{w}(x), \alpha > -1$, where $(\alpha) \tilde{w}(x)$ is a non-negative bounded measurable function, or (b) $w(x)$ is a non-negative polynomial of given degree.

P. 3. Range $-\infty$ to $\infty$. The theorem holds if the conditions of P. 1 are satisfied with $w(x) \equiv e^{-x^2} \tilde{w}(x)$, where $\tilde{w}(x)$ satisfies (a) or (b) of P. 2.

The statements in P. 2 and P. 3 when $w(x) = 1$ have been given by Szegö (loc. cit., pp. 104-106) who extended a method due to J. v. Neumann for a weight function of the form $e^{-x}$ (see Courant, R. and Hilbert, D., Methoden der Mathematischen Physik, Vol. 1 (Berlin 1931), pp. 81-2). Following v. Neumann and Szegö, we can deduce P. 2 from P. 1 provided it can be shown that if $m$ is a non-negative integer there exists for every $\epsilon > 0$ a polynomial $p_{n-1}(x)$ such that

$$S^2 = \int_0^\infty e^{-x} x^a \tilde{w}(x) \left( e^{-mx} - p_{n-1}(x) \right)^2 \, dx < \epsilon.$$  \hfill (7)

P. 2 (a) follows with $\tilde{w}(x) = 1$, and P. 3 (a) may be deduced from this. P. 2 (b) may be proved by an extension of the Neumann-Szegö method. We require the following properties of the Laguerre polynomials:

$$n! L_n^\alpha(x) = e^{x} x^{-a} \left( \frac{d}{dx} \right)^n e^{-x} x^{n+a}, \quad n = 0, 1, \ldots,$$  \hfill (8)

$$\int_0^\infty e^{-x} x^a L_n^a(x) L_m^a(x) \, dx = \frac{(n+a)!}{n!} \delta_{n,m}, \quad n, m = 0, 1, \ldots,$$  \hfill (9)

$$(1-\omega)^{a+1} \sum_{r=0}^\infty \omega^r L_r^a(x) = \exp \left\{ -\omega x/(1-\omega) \right\}, \quad |\omega| < 1,$$  \hfill (10)

$$L_n^\alpha(x) = L_{n+r}^\alpha(x) - \binom{r}{1} L_{n-1}^\alpha(x) + \binom{r}{2} L_{n-2}^\alpha(x) - \ldots.$$  \hfill (11)

Suppose now that

$$\tilde{w}(x) = a + bx + cx^2, \quad a \neq 0$$

$$\leq (|a| + |b| + |c|) (1 + x^2)$$

$$= k (1 + x^2).$$

Then

$$S^2 \leq k \int_0^\infty e^{-x} x^a + x^{n+a+2} \left( e^{-mx} - p_{n-1}(x) \right)^2 \, dx.$$  \hfill (12)

Take $p_{n-1}(x) = (1-\omega)^{a+1} \sum_{s=0}^{n-1} \omega^s L_s^a(x), \quad \omega = m/(m+1)$.
so that
\[ S^2 \leq k \int_0^\infty e^{-x} x^a \left[ (1 - \omega)^a + \sum \omega^s L_s^a (x) \right]^2 \, dx \]
\[ + k \int_0^\infty e^{-x} x^{a+2} \left[ (1 - \omega)^{a+1} \omega^s L_s^{a+2} (x) - (1 - \omega)^{a+1} \omega^s (2 - \omega) L_s^{a+2} (x) \right] \, dx \]

after using (10) and (11), the rearrangement of terms being justified since (10) is absolutely convergent. Hence
\[ S^2 \leq k \int_0^\infty \left( 1 - \omega \right)^{2a + 2} \omega^2 \frac{(n + a + 2)!}{(n - 2)!} \frac{F(1, n + a + 1; n + 1; \omega^2)}{n (n - 1) (n + a + 2) (n + a + 1)} + \frac{1}{n (n - 1) (n + a + 2)} \]

with the usual notation for the hypergeometric series. Term-by-term integration is justified since \( |L_n^a (x)| < e^x (n + a)! / n! \), so that \( \sum r^s L_r^a (x) L_s^a (x) \) converges uniformly for \( x \) in \( (0, A) \), \( A > 0 \) fixed, and by Schwarz's inequality
\[ \sum \int_0^\infty \omega^r s^{r+s} e^{-x} x^a |L_r^a (x)| |L_s^a (x)| \, dx \leq \sum \int_0^\infty \omega^r s^{r+s} \left[ \frac{(r + a)! (s + a)!}{r! s!} \right] \]
converges. Since \( \omega < 1 \) it is seen that, for \( \epsilon > 0 \), \( n \geq n (\epsilon, a) \) exists so that \( S^2 < \epsilon \). A similar proof applies to \( w (x) \) of any given degree.

P. 3 (b) follows from this (see Szegö, loc. cit., p. 105 (3)).

C. Illustrative Examples.

C. 1. Let \( I (p, q) = \int_{-1}^1 \frac{dx}{(x^2 + 2px + q) \sqrt{(1 - x^2)}} \),

where \( x^2 + 2px + q > 0 \) for \(-1 \leq x \leq 1\).

The conditions of P. 1 are satisfied with
\[ w (x) = (x^2 + 2px + q) / \sqrt{(1 - x^2)} \text{ and } f(x) = 1 / (x^2 + 2px + q). \]

Hence using (2) with
\[ A(x) = B(x) = 1, \quad C(x) = x^2 + 2px + q, \quad w (x) = 1 / \sqrt{(1 - x^2)}, \]

we have with
\[ \theta_s (x) = \sqrt{(2/\pi)} \cos s\phi, \quad \cos \phi = x, \quad s = 1, 2, \ldots \]
\[ \theta_0 (x) = \sqrt{(1/\pi)} \]

---

3 See Szegö, loc. cit., pp. 30-32.
A DETERMINANTAL EXPANSION FOR A CLASS OF DEFINITE INTEGRAL 83

the values
\[ \gamma_{s,s} = q + \frac{1}{4}, \quad s = 0, 2, 3, \ldots \]
\[ \gamma_{s,s+1} = p, \quad s = 1, 2, \ldots \]
\[ \gamma_{s,s+2} = \frac{1}{4}, \quad s = 1, 2, \ldots \]
\[ \gamma_{s,y} = 0, \quad r > s + 2 \]
\[ a_s = \beta_s = 0, \quad s \neq 0 \]
\[ a_0 = \beta_0 = 2\sqrt{2/\pi} \]

so that after slight simplification

\[
\frac{2}{\pi} I(p,q) = \begin{vmatrix}
0 & 1 & 0 & 0 & 0 & \ldots \\
1 & \frac{1}{2}q + \frac{1}{4} & p & \frac{1}{4} & 0 & \ldots \\
0 & p & q + \frac{3}{4} & p & \frac{1}{4} & \ldots \\
0 & \frac{1}{4} & p & q + \frac{1}{2} & p & \ldots \\
0 & 0 & \frac{1}{4} & p & q + \frac{1}{2} & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{vmatrix}, \quad (14)
\]

the expansion providing an increasing sequence.

Similarly, if \( q - p^2 \neq 0 \), then from

\[
\int_{-1}^{1} \frac{dx}{\sqrt{(1 - x^2)}} = (q - p^2) I + \int_{-1}^{1} \frac{(x + p)^2 dx}{(x^2 + 2px + q) \sqrt{(1 - x^2)}}
\]

we have

\[
\frac{2}{\pi} (q - p^2) I(p,q) = 2 \begin{vmatrix}
0 & 1 & 0 & 0 & \ldots \\
p & \frac{1}{2}q + \frac{1}{4} & p & \frac{1}{4} & 0 & \ldots \\
1 & p & q + \frac{3}{4} & p & \frac{1}{4} & \ldots \\
0 & \frac{1}{4} & p & q + \frac{1}{2} & p & \ldots \\
0 & 0 & \frac{1}{4} & p & q + \frac{1}{2} & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{vmatrix}, \quad (15)
\]

and this gives a decreasing sequence for \( I(p,q) \) if \( q > p^2 \).

An alternative expansion follows from (6) with
\[ x^2 + 2px + q = (x - \cos \theta_1)(x - \cos \theta_2), \quad \text{where } \theta_1 \text{ and } \theta_2 \text{ are complex, } \theta_1 \neq \theta_2: \]

\[
I(p,q) = \sum_{s=1}^{\infty} \frac{2\pi (\cos s \theta_1 - \cos s \theta_2)^2}{|\cos(s-1) \theta_1, \cos s \theta_2| \cdot |\cos s \theta_1, \cos (s+1) \theta_2|} \quad (16)
\]
The expansions (14)-(16) represent simple generalisations of the continued fraction development
\[
\frac{1}{\pi} \int_{-1}^{1} \frac{dx}{(z-x) \sqrt{1-x^2}} = \frac{1}{z} + \frac{1}{2} + \frac{1}{2} \frac{1}{z} + \cdots \quad |z| > 1
\]

and
\[
\sum_{s=1}^{\infty} \sec (s-1) \theta \sec s \theta \quad \text{with} \quad \cos \theta = z.
\]

Similar results hold for
\[
\int_{-1}^{1} \frac{\sqrt{(1-x^2)} \, dx}{x^2 + 2px + q} \int_{-1}^{1} \frac{1}{x^2 + 2px + q} \sqrt{1+x} \, dx
\]

and for \( C(x) \) a polynomial of higher degree than the second.

C. 2. We next consider
\[
\eta(b, a) = \int_0^\infty \frac{e^{-ax} x^{b-1} \, dx}{1 - e^{-x}} = \Gamma(b) \sum_{s=0}^{\infty} (s+a)^{-b}, \quad a > 0, \quad b > 1, \quad (17)
\]
and
\[
\theta(b, a) = \int_0^\infty \frac{e^{-ax} x^{b-1} \, dx}{1 + e^{-x}} = \Gamma(b) \sum_{s=0}^{\infty} (-)^s(s+a)^{-b}, \quad a > 0, \quad b > 0. \quad (18)
\]

With \( w(x) = e^{-x}/x \), which is non-negative, measurable and bounded \((\leq 1), w(x) = e^{-ax} x^{b-3} (1 - e^{-x}), \) and \( f(x) = x/ (1 + e^{-x}) \) so that \( f(x) \sqrt{w(x)} \) belongs to \( L^2(0, \infty) \), the conditions of P. 2. (a) are satisfied. In (2) we take \( w(x) = e^{-ax} x^{b-2}, \) \( C(x) = (1 - e^{-x})/x, A(x) = B(x) = 1, \beta_r(x) = x^r, \) so that \( \alpha_r = \beta_r = \Gamma(r+b-1) a^{1-b} - r, \gamma_r = - \Gamma(r+s+b-2) \Delta a^2 - r - b \), where \( \Delta a^n = (a + 1)^n - a^n. \)

Thus
\[
\eta(b, a) = \begin{vmatrix}
\Gamma(b) a^{1-b} & \Gamma(b+1) a^{-b-1} \\
\Gamma(b+1) a^{1-b} & \Gamma(b+2) \Delta a^2 - b \\
\Gamma(b) a^{-b} & \Gamma(b+1) \Delta a^{1-b} \\
\Gamma(b+1) a^{-b-1} & \Gamma(b+2) \Delta a^{1-b-1} \\
\end{vmatrix}
\]

In the special case \( b = 2, \Gamma(b+2) \Delta a^2 - b \) must be replaced by \(- \log (1 + 1/a).\)

Similarly
\[
\theta(b, a) = - \begin{vmatrix}
\Gamma(b) a^{-b} & \Gamma(b+1) a^{-b-1} \\
\Gamma(b+1) a^{-b-1} & \Gamma(b+2) \nabla a^{-b-2} \\
\end{vmatrix}
\]
where \( \nabla a^r = a^r + (a + 1)^r. \)

Again, using \( \theta(b, a) = \Gamma(b) a^{-b} - \theta(b, a + 1), \) we find that
A Determinantal Expansion for a Class of Definite Integral

\[ \theta(b,a) = \begin{vmatrix} \Gamma(b) a^{-b} & \Gamma(b) (a + 1)^{-b} & \Gamma(b+1) (a+1)^{-b-1} \\ \Gamma(b) (a + 1)^{-b} & \Gamma(b+1) \nabla (a + 1)^{-b} & \Gamma(b+1) \nabla (a+1)^{-b-1} \\ \Gamma(b+1) (a + 1)^{-b-1} & \Gamma(b+1) \nabla (a+1)^{-b-1} & \Gamma(b+2) \nabla (a+1)^{-b-2} \end{vmatrix}, \quad a > 0 \]

\[ b > 0. \]  

(21)

The expansions (19) and (20) are positive non-decreasing sequences while (21) is a positive non-increasing sequence. As a numerical illustration we take \( b = 2, \ a = 1 \) in (20) and (21) for which

\[ 2 \theta(2, 1) = \Sigma n^{-2}. \]  

For the first three approximations we have

\[
\begin{align*}
(20) & \\
152/93 = 1.634 & \\
33168/20187 = 1.64304 & \\
\end{align*}
\]

Thus \( 1.64304 < \Sigma n^{-2} < 1.644950 \), the correct value being \( 1.644934 \). Similarly from (19) we find for \( \Sigma n^{-2} \)

\[
\begin{align*}
\frac{1}{w} & = 1.44 & (w = \ln 2) \\
(4w - 1)/(3w - 1) & = 1.6421 \\
(104w - 42)/(74w - 33) & = 1.64475 \\
(16272w - 7790)/(11178w - 5627) & = 1.644928, \\
\end{align*}
\]

so that the fourth approximation is in error by \( 0.000,006 \). We note in passing that continued fractions for \( \sum_{s=0}^{\infty} (a+s)^{-b} \) in the particular cases \( b = 2 \) and \( b = 3 \) have been given by Stieltjes (1890) and rediscovered, although by a different method, by Rogers (1905). For example,

\[
\sum_{s=0}^{\infty} (a+s)^{-2} = \frac{1}{a} + \frac{a_1}{a - \frac{1}{2}} + \frac{a_2}{a - \frac{1}{2} + \frac{1}{2}} + \cdots, \quad a_p = \frac{p^4}{4(p^2 - 1)}. 
\]

With \( a = 1 \) in this, the eleventh and twelfth convergents to \( \Sigma n^{-2} \) are \( 1.65245 \) and \( 1.63856 \), indicating a slower rate of convergence than (19)-(21).

C. 3. The Psi function and related integrals.

We have

\[
\Psi(t) = \ln t - \int_0^t e^{-x} \left( \frac{e^x - 1}{x} \right) dx, \quad t > 0.
\]
Take \( w(x) = e^{-tx} \overline{w}(x) \), where \( \overline{w}(x) = (1 - e^{-x}) (x - 1 + e^{-x}) / x^3 \) is a non-negative bounded measurable function (its value for \( x = 0 \) being taken as \( \frac{1}{4} \)). Put \( f(x) = x / (1 - e^{-x}) \), so that \( f(x) \sqrt{w(x)} \) belongs to \( L^2(0, \infty) \). Then the conditions of P.2 (a) are satisfied. In (2), with \( A(x) = B(x) = 1 \),

\[
C(x) = (1 - e^{-x}) / x, w(x) = (x - 1 + e^{-x}) e^{-tx} / x^2,
\]

we have, taking \( \theta_r = x^r \),

\[
a_r = \beta_r = \int_0^\infty e^{-tx} x^{r-2} (x - 1 + e^{-x}) \, dx,
\]

\[
\gamma_{rs} = \gamma_{t, r} = \gamma_{u} = \int_0^\infty e^{-tx} x^{r-3} (x - 1 + e^{-x}) (1 - e^{-x}) \, dx, \quad u = r + s
\]

Thus

\[
\begin{align*}
a_0 &= -1 + (1 + t) \ln (1 + t^{-1}) \\
a_1 &= t - 1 - \ln (1 + t^{-1}) \\
a_r &= \left( -\frac{d}{dt} \right)^r a_0.
\end{align*}
\]

Similarly

\[
\begin{align*}
\gamma_{00} &= \gamma_0 = -\frac{1}{2} + \frac{3}{2} t (2 + t) \ln t - (1 + t) (2 + t) \ln (1 + t) + \frac{1}{2} (2 + t)^2 \ln(2 + t) \\
\gamma_{01} &= \gamma_{10} = \gamma_1 = -(1 + t) \ln t + (3 + 2t) \ln (1 + t) - (2 + t) \ln(2 + t) \\
\gamma_{02} &= \gamma_{11} = \gamma_{20} = \gamma_2 = \ln t - 2 \ln (1 + t) + \ln (2 + t) + 1/t (1 + t) \\
\gamma_r &= \left( -\frac{d}{dt} \right)^r \gamma_0.
\end{align*}
\]

Hence \( \Psi(t) = \frac{d}{dt} \ln \Gamma(t) \)

\[
= \begin{array}{l}
\ln t \\
a_0 \\
a_0' \\
a_0'' \\
\vdots
\end{array}
\begin{array}{cccc}
a_0 & a_0' & a_0'' & \cdots
\end{array}
\begin{array}{cccc}
\gamma_0 & \gamma_0' & \gamma_0'' & \cdots
\end{array}
\begin{array}{cccc}
\gamma_0 & \gamma_0' & \gamma_0'' & \cdots
\end{array}
\begin{array}{cccc}
\gamma_0 & \gamma_0' & \gamma_0'' & \cdots
\end{array}
\begin{array}{cccc}
\vdots & \vdots & \vdots & \vdots
\end{array}, \quad t > 0,
\]

in which superscripts denote derivatives. This is a non-increasing sequence. A non-decreasing sequence is found from

\[
\Psi'(t) = \int_0^\infty \left( \frac{e^{-tx} - e^{-tx}}{x - 1 - e^{-x}} \right) \, dx, \quad t > 0.
\]
A Determinantal Expansion for a Class of Definite Integral 87

In (2) we take \( A(x) = B(x) = 1, \quad C(x) = (1 - e^{-x})/x; \)
\( w(x) = (e^{-x} - 2e^{-2x} - xe^{-tx})/x^2, \quad t \geq 1.5, \) the restriction on \( t \) being necessary to ensure \( w(x) \geq 0. \) With \( \theta_r = x^r \) we find

\[
a_r = \beta_r = \int_0^\infty (e^{-x} - e^{-2x} - xe^{-tx}) x^{r-2} dx,
\]
and in particular

\[
a_0 = 1 - 2 \ln 2 + \ln t, \quad a_1 = \ln 2 - t^{-1}.
\]

Similarly

\[
\gamma_{rs} = \gamma_{sr} = \gamma_u = \int_0^\infty (e^{-x} - 2e^{-2x} + e^{-3x} - xe^{-x'} + xe^{-x(t+1)}) x^{-3} dx
\]
where \( u = r + s = 0, 1, \ldots, \)
and in particular

\[
\gamma_{00} = \gamma_0 = \frac{1}{2} + 4 \ln 2 - \frac{9}{2} \ln 3 + (1 + t) \ln(1 + t) - t \ln t
\]
\[
\gamma_{10} = \gamma_{01} = \gamma_1 = -4 \ln 2 + 3 \ln 3 + \ln t - \ln(1 + t)
\]
\[
\gamma_{20} = \gamma_{11} = \gamma_{02} = 2 \ln 2 - \ln 3 - 1/t(1 + t).
\]

As a numerical example put \( t = 1 \) in (22), so that

\[
\Psi(1) = \begin{bmatrix}
0 & 0.386294 & 0.306853 & 0.50000 \\
0.386294 & 0.284872 & 0.169899 & 0.212318 \\
0.306853 & 0.169899 & 0.212318 & 0.416667 \\
0.500000 & 0.212318 & 0.416667 & 1.138889 \\
\end{bmatrix}
\]

from which we have the first three approximations to Euler's constant \( C = 0.577216, \) namely, \( 0.52383, \quad 0.57651, \quad 0.57718. \) Similarly, from the expansion corresponding to (23) we have

\[
\Psi(2) = \begin{bmatrix}
0 & 0.306853 & 0.193147 & 0.250000 \\
0.306853 & 0.238376 & 0.117783 & 0.121015 \\
0.193147 & 0.117783 & 0.121015 & 0.194444 \\
0.250000 & 0.121015 & 0.194444 & 0.435185 \\
\end{bmatrix}
\]
so that using the recurrence relation \( \Psi'(1 + t) = \Psi'(t) + t^{-1} \) we have 
\( C < 0.60500, 0.57754, 0.57723 \). Hence \( 0.57718 < C < 0.57723 \).

A similar type of integral appears for

\[
J(t) = \ln \Gamma(t) - (t - \frac{1}{2}) \ln t - \frac{1}{2} \ln 2 \pi \\
= \int_0^\infty \left( \frac{1}{x} - \frac{1}{e^x - 1} \right) \frac{e^{-ix}}{x} \, dx, \quad t > 0,
\]

for which Stieltjes (1894) has given a continued fraction. With

\( w(x) = e^{-tx} \bar{w}(x) \), where \( \bar{w}(x) = \{ x - 2 + (x + 2) \, e^{-x} \} \, (1 - e^{-x})/x^4 \),

and \( f(x) = x/(1 - e^{-x}) \), so that \( w(x) \) is non-negative, bounded and measurable, and \( f(x) \sqrt{w(x)} \) belongs to \( L^2(0, \infty) \), the conditions of P. 2 (a) are satisfied. It may be verified that

\[
2J(t) = \begin{bmatrix}
0 & u^0 & u^1 & u^2 & \ldots \\
u^0 & v^0 & v^1 & v^2 & \ldots \\
u^1 & v^1 & v^2 & v^3 & \ldots \\
u^2 & v^2 & v^3 & v^4 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots 
\end{bmatrix}, \quad t > 0, \quad (24)
\]

where \( u^t \equiv \left( -\frac{d}{dt} \right)^t \left\{ t + \frac{1}{2} - t \,(1 + t) \ln (1 + t^{-1}) \right\}, \)

\[
v^t \equiv \left( -\frac{d}{dt} \right)^t \left\{ -\frac{(3t^2 + 2t^3)}{6} \ln t + \frac{2 \,(1 + t)}{3} \ln (1 + t) \right. \\
- \frac{(2 + t)^2 \,(1 + 2t)}{6} \ln (2 + t) + \frac{2 \,(1 + t)}{3} \right\}
\]

C. 4. Integrals of the form

\[
\int_{-\infty}^{\infty} \left[ \frac{A(x)}{C(x)} \right]^2 e^{-ix} \, dx, \quad C(x) > 0.
\]

As an illustration we consider in particular

\[
G(a, b) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} e^{-ix} \, dx \quad \text{b > a}^2.
\]

The conditions of P. 3 (b) are satisfied with \( w(x) = \frac{1}{\sqrt{(2\pi)}} \, e^{-ix} \bar{w}(x) \)

where \( \bar{w}(x) = x^2 + 2ax + b \), \( f(x) = g(x) = 1/\bar{w} \), so that \( f(x) \sqrt{w(x)} \) belongs to \( L^2(-\infty, \infty) \). In (2), take \( A(x) = B(x) = 1, C(x) = \bar{w}(x), \)

\( w(x) = \frac{1}{\sqrt{(2\pi)}} \, e^{-ix} \) and \( \theta_r(x) = H_r(x) = e^{ix} \left(-\frac{d}{dx}\right)^r e^{-ix} \). Then
A Determinantal Expansion for a Class of Definite Integral

\[ a_r = \beta_r = 1, \quad r = 0 \]
\[ = 0, \quad r = 1, 2, \ldots \]
\[ \gamma_{r,s} = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} (H_2(x) + 2aH_1(x) + b + 1) H_r(x) H_s(x) e^{-x^2} dx \]
\[ = (b + 2r + 1) r!, \quad r = s \]
\[ = 2ar!, \quad r = s + 1 \]
\[ = r!, \quad r = s + 2 \]
\[ = 0, \quad r > s + 2 \]

and so

\[ G(a, b) = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 \\
1 & b + 1 & (2a)1! & 2! & 0 \\
0 & 2a1! & (b + 1)1! & (2a)2! & 3! \\
0 & 2! & (2a)2! & (b + 5)2! & (2a)3! \\
0 & 0 & 3! & (2a)3! & (b + 7)3! \\
\end{pmatrix}, \quad (26) \]

which gives a non-decreasing sequence. Similarly, using

\[ (b - a^2) G(a, b) = 1 - \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} (x + a)^2 e^{-x^2} dx \]

we derive the non-increasing sequence

\[ (b - a^2) G(a, b) = \begin{pmatrix}
1 & a & 1 & 0 & 0 \\
0 & b + 1 & (2a)1! & 2! & 0 \\
1 & (2a)1! & (b + 3)1! & (2a)2! & 3! \\
0 & 2! & (2a)2! & (b + 5)2! & (2a)3! \\
0 & 0 & 3! & (2a)3! & (b + 7)3! \\
\end{pmatrix}, \quad (28) \]

The series expansions derived from (6) corresponding to (26) and (28) are respectively

\[ G(a, b) = \sum_{s=0}^{\infty} \frac{s!}{H_s(a), H_s+1(\beta)} \cdot \frac{H_{s+1}(\beta)^2}{|H_{s+1}(a), H_{s+2}(\beta)|} \]

and

\[ \begin{pmatrix}
G(a, b) = \frac{1}{b - a^2} - \frac{1}{b - a^2} \sum_{s=0}^{\infty} \frac{s!}{F_s F_{s+1}} \\
a + H_1(a) H_{s+1}(\beta)^2 \\
|H_s(a), H_{s+1}(\beta)| \cdot |H_{s+1}(a), H_{s+2}(\beta)| \\
\end{pmatrix} \]

\[ = \frac{1}{b - a^2} - \frac{1}{b - a^2} \sum_{s=0}^{\infty} \frac{s!G_s^2}{F_s F_{s+1}}, \quad (30) \]
where \( a, \beta \) are the roots of \( x^2 + 2ax + b = 0 \). The values of \( E_s, F_s \) and \( G_s \) are readily calculated from the recurrence relation
\[
H_{s+1}(x) = xH_s(x) - sH_{s-1}(x).
\]
Since
\[
G_s^2 + (b-a^2)E_s^2 = -4(b-a^2)H_{s+1}(a)H_{s+1}(\beta),
\]
it will be seen that the difference between the \((s+1)^{th}\) approximations arising from (29) and (30) is \( s!F_0/(b-a^2)F_s \). This may be used to assess the rate of convergence and also as a computational check.

It is interesting to observe that, when \( a = 0 \), (26) and (28) reduce to simple continuant quotients and give the even and odd part of the continued fraction
\[
G(0, b) = \frac{1}{b + \frac{1}{1 + \frac{2}{b + \frac{3}{1 + \frac{4}{b + \ldots}}}}}, \quad b > 0.
\]

By an equivalence transformation we have the Laplace (1805) continued fraction for the incomplete normal integral, namely
\[
G(0, t^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} dx = t^{-1} e^{\frac{t^2}{2}} \int_{t}^{\infty} e^{-x^2} dx
\]
\[
= \frac{t^{-1}}{t + \frac{1}{t + \frac{2}{t + \frac{3}{t + \ldots}}}}, \quad t > 0.
\]

<table>
<thead>
<tr>
<th>( s )</th>
<th>TERM OF (29)</th>
<th>( \Sigma )</th>
<th>TERM OF (30)</th>
<th>( \Sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.500000</td>
<td>.5000</td>
<td>1.166667</td>
<td>1.1667</td>
</tr>
<tr>
<td>1</td>
<td>.071429</td>
<td>.5714</td>
<td>.214286</td>
<td>.9524</td>
</tr>
<tr>
<td>2</td>
<td>.075630</td>
<td>.6471</td>
<td>.070028</td>
<td>.8824</td>
</tr>
<tr>
<td>3</td>
<td>.042596</td>
<td>.6897</td>
<td>.035062</td>
<td>.8473</td>
</tr>
<tr>
<td>4</td>
<td>.015594</td>
<td>.7052</td>
<td>.034370</td>
<td>.8129</td>
</tr>
<tr>
<td>5</td>
<td>.023337</td>
<td>.7286</td>
<td>.007269</td>
<td>.8057</td>
</tr>
<tr>
<td>6</td>
<td>.003307</td>
<td>.7319</td>
<td>.017955</td>
<td>.7877</td>
</tr>
<tr>
<td>7</td>
<td>.012806</td>
<td>.7447</td>
<td>.001422</td>
<td>.7863</td>
</tr>
<tr>
<td>8</td>
<td>.005688</td>
<td>.7453</td>
<td>.009729</td>
<td>.7765</td>
</tr>
<tr>
<td>9</td>
<td>.007125</td>
<td>.7524</td>
<td>.000181</td>
<td>.7764</td>
</tr>
<tr>
<td>10</td>
<td>.000380</td>
<td>.7524</td>
<td>.005401</td>
<td>.7710</td>
</tr>
<tr>
<td>11</td>
<td>.004019</td>
<td>.7564</td>
<td>.000000</td>
<td>.7710</td>
</tr>
<tr>
<td>12</td>
<td>.000010</td>
<td>.7565</td>
<td>.003051</td>
<td>.7679</td>
</tr>
<tr>
<td>13</td>
<td>.002293</td>
<td>.7588</td>
<td>.000336</td>
<td>.7679</td>
</tr>
<tr>
<td>14</td>
<td>.000063</td>
<td>.7588</td>
<td>.001743</td>
<td>.7661</td>
</tr>
<tr>
<td>15</td>
<td>.001318</td>
<td>.7601</td>
<td>.000873</td>
<td>.7660</td>
</tr>
<tr>
<td>16</td>
<td>.000104</td>
<td>.7602</td>
<td>.001004</td>
<td>.7650</td>
</tr>
<tr>
<td>17</td>
<td>.000762</td>
<td>.7610</td>
<td>.000115</td>
<td>.7649</td>
</tr>
<tr>
<td>18</td>
<td>.000120</td>
<td>.7611</td>
<td>.000580</td>
<td>.7643</td>
</tr>
<tr>
<td>19</td>
<td>.000441</td>
<td>.7616</td>
<td>.000122</td>
<td>.7642</td>
</tr>
<tr>
<td>20</td>
<td>.000119</td>
<td>.7617</td>
<td>.000335</td>
<td>.7639</td>
</tr>
</tbody>
</table>
A DETERMINANTAL EXPANSION FOR A CLASS OF DEFINITE INTEGRAL

As a numerical illustration we take $2a = b = 1$, and by comparison with the incomplete normal integral continued fraction development (Burgess, 1895) we expect rather slow convergence. Evaluating $s! F_s/(b - a^2) F_s$ for $s = 20$ we find that it is approximately 0.0027, so that we have only two-figure accuracy. In the table we give the terms and partial sums for the series (29) and (30), the identity (31) being used as a check.

We conclude then that $-7617 < G(0.5, 1.0) < -7639$, the correct value being -7628, 2634. The oscillatory nature of the terms is noteworthy, and this would be an awkward feature if we could not construct an enveloping sequence.

We intend to discuss later various forms for the numerators and denominators of the expansions considered here, including recurrence relations, noting the relation to the theory of continued fractions.

I am greatly indebted to the referees for a number of useful comments and criticisms, and to Dr W. Ledermann for some criticisms of Part 1.

REFERENCES.

Stieltjes, T. J. (1890). Sur quelques intégrales définies et leur développement en fractions continues.

COLLEGE OF TECHNOLOGY,
MANCHESTER, 1.
A Determinantal Expansion for a Class of Definite Integral; Part 3.

Extracted from the
Edinburgh Mathematical Proceedings
Series 2—Vol. 10—Part III—January 1966
A Determinantal Expansion for a Class of Definite Integral

Part 3. Generalised Continued Fractions

By L. R. Shenton

(Received 2nd March 1953.)

1. We have shown in [1] that under certain conditions the definite integral \( \int_a^b \frac{A^2(x) w(x)}{C(x)} \, dx \) may be approximated by a determinantal ratio. It is our object now to develop the theory when \( C(x) \) is a polynomial, showing the relation to the continued fraction form for \( \int_a^b \frac{w(x) \, dx}{z - x} \). In particular we shall give various forms for the approximants, and an integral form for the numerator.

2. From [1] we have the expansion

\[
F(z_1, z_2, \ldots, z_n) = \int_a^b w(x) \, dx = \sum_{s=0}^{\infty} \frac{\prod_{k=0}^{s} (z_k - x)}{\Delta_s - 1} \frac{\prod_{k=1}^{s} (z_k - x)}{\Delta_s} \tag{1}
\]

\[
= -\lim_{s \to \infty} \frac{\Delta_s}{\Delta_s} \begin{vmatrix}
0 & a_0 & a_1 & \cdots & a_s \\
&a_0 & \gamma_{01} & \gamma_{02} & \cdots & \gamma_{0s} \\
&\vdots & & \ddots & & \vdots \\
&\vdots & & & \ddots & \vdots \\
&\vdots & & & & \gamma_{ss} \\
\end{vmatrix} 
\tag{2}
\]

\[
= \lim_{s \to \infty} \frac{N_{s+1}(z_1, z_2, \ldots, z_n)}{D_{s+1}(z_1, z_2, \ldots, z_n)} \tag{3}
\]

where \(^1\)

\[
\alpha_s = \int_a^b \theta_s(z) w(x) \, dx, \quad \gamma_{s, z} = \int_a^b \theta_s(x) \theta_s(x) w(x) \prod_{k=1}^{n} (z_k - x) \, dx,
\]

\[
\Delta_s = D_{s+1}(z) = \begin{vmatrix}
\gamma_{00} & \gamma_{01} & \cdots & \gamma_{0s} \\
\gamma_{10} & \gamma_{11} & \cdots & \gamma_{1s} \\
\vdots & \vdots & & \vdots \\
\gamma_{s0} & \gamma_{s1} & \cdots & \gamma_{ss} \\
\end{vmatrix},
\]

\( w(x) \geq 0 \) for \( a \leq x \leq b \) \((a, b \text{ finite})\),

\[
\int_a^b w(x) \, dx \text{ exists and is positive},
\]

\(^1\) \( N_s(z_1, z_2, \ldots, z_n) \) and \( D_s(z_1, z_2, \ldots, z_n) \) etc. will be abbreviated to \( N_s(z) \) and \( D_s(z) \) when ambiguity is unlikely.
\( \Pi(z_1 - x) \equiv (z_1 - x)(z_2 - x) \ldots (z_n - x) > 0 \) and the \( z \)'s are distinct, \( \theta_s(x) \) is an arbitrary polynomial of precise degree \( s \) with highest coefficient \( \lambda_s \).

We shall refer to (1) and (2) as continued fractions (C.F.'s) of the \( n \)th order, and \( R_s(z_1, z_2, \ldots, z_n) = N_s/D_s \) as the \( s \)th approximant or convergent.

The expansions (1) and (2) arise from a consideration of the minimum value of

\[ S^2 = \int_a^b w(x) \Pi(z - x) \left\{ \frac{1}{\Pi(z - x)} - \sum_{s=0}^{r-1} A_s q_s(x) \right\}^2 \, dx \tag{4} \]

where \( q_s(x) = | \theta_0(x), \gamma_{01}, \gamma_{12}, \ldots, \gamma_{s-1,s} | / \{( - )^s \lambda_s \Delta_2 - 1 \} \),

and \( \{ q_s(x) \} \) is an orthogonal system with respect to the weight function \( w(x)\Pi(z - x) \), the highest coefficient in \( q_s(x) \) being unity. Indeed if we write

\[ \phi_r = \int_a^b w(x) \Pi(z_j - x) q_r^2(x) \, dx \tag{6} \]

then \( A_s \phi_r = \int_a^b q_r(x)w(x) \, dx \tag{7} \)

and \( S_{\text{Min}}^2 = F(z) - \sum_{s=0}^{r-1} A_s^2 \phi_s. \tag{8} \)

It may be remarked in passing that a consideration of the minimum value of \( \int_a^b (z - x)w(x) \{ (z - x)^{-1} - \sum A_s q_s(x) \}^2 dx \)

and of \( \int_a^b x(z - x)w(x) \{ (z - x)^{-1} - \sum A_s^1 q_s^1(x) \}^2 dx \)

leads to continued fractions for \( \int_a^b (z - x)^{-1} w(x) \, dx \) related to the ‘even’ and ‘odd’ parts of a Stieltjes type of continued fraction. The present approach shows immediately the central part played by orthogonal polynomials, and although in essence both these expressions were considered by Stieltjes [2], it is only at a later stage that the orthogonality property emerges.

3. We shall now consider various forms for \( N_s(z) \) and \( D_s(z) \). These arise by taking (a) \( \theta_s(x) = (v - x)^s \); (b) \( \theta_s(x) = p_s(x) \),

where \( \{ p_s(x) \} \) is an orthonormal system with respect to \( w(x) \), and \( p_s(x) \) has highest coefficient \( k_s \).
\( \theta_s(x) = q_s(x) = \left\lvert \begin{array}{c} p_s(x), p_s + 1(z_1), p_s + 2(z_2), \ldots, p_s + n(z_n) \\ k_{s+n}(z - x) \end{array} \right\rvert \)

and the system \( \{q_s(x)\} \) is orthogonal with respect to \( w(x)\Pi(z - x) \).

(a) Here \( a_s = \int_a^b (v - x)^r w(x) \, dx = m_s \) say,

\[ \gamma_{r,s} = \int_a^b (v - x)^r w(x) \Pi(z - x) \, dx = M_{r+s+n} \] say.

For particular choices of \( w(x) \), \( m_s \) is an Appell polynomial.

Further, if \( z_1 = z_2 = \ldots = z_n = v \) then \( \gamma_{r,s} = M_{r+s+n} \).

From (6) we have, in the notation of per-symmetric determinants,

\[ \phi_r = \frac{P_{r+1}(M_{s+1}, M_{s+2}, \ldots, M_{s+n})}{P_r(M_{s+1}, M_{s+2}, \ldots, M_{s+n})} \]

and

\[ \sum_{s=0}^{r-1} A^2_s \phi_s \]

\[ = \left| \begin{array}{cccc} 0 & m_0 & m_1 & \ldots & m_{r-1} \\ m_0 & M_0 & M_1 & \ldots & M_{r-1} \\ m_1 & M_0 & M_1 & \ldots & M_{r-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{r-1} & M_{r-2} & M_{r-1} & \ldots & M_{2r-2} \end{array} \right| \]

If the roots \( z_j \) are equal and \( z_j = v \), \( M_s \) is to be replaced by \( m_s \).

(b) The polynomials \( p_s(x) \) follow a recurrence relation

\[ p_s(x) = (A_x + B_s) p_{s-1}(x) - C_s p_{s-2}(x), \quad s = 2, 3, \ldots \]

\[ p_1(x) = (A_x + B_1) p_0(x), \quad p_0 = k_0 \]

\[ A_r = k_r/k_{r-1} > 0, \quad C_r = A_r/A_{r-1} > 0, \]

which may be written

\[ (x - z_1) p_{s-1}(x) = A_s^{-1} p_s(x) - (z_1 + B_s A_s^{-1}) p_{s-1}(x) + A_{s-1} p_{s-2}(x). \]

We require the following generalisation of (15):

\[ \Pi(x - z_j) p_{s-1}(x) = e_s - 1, n p_{s+n-1} + e_s - 1, n - 1 p_{s+n-2} + \ldots + e_0 p_s - 1 + e_{s-2,1} p_{s-2} + e_{s-3,2} p_{s-3} + \ldots + e_{s+n-1,0} p_{s-n-1} \]

where \( e_{s+n} = k_{s+n}/k_{s+n-1} \), and \( e_{s,r} \) is to be taken as zero if \( s < 0 \).

The notation is justified by the identity

\[ \int_a^b [p_{s-1}(x) \Pi(x - z)] p_r(x) w(x) \, dx = \int_a^b [p_r(x) \Pi(x - z)] p_{s-1}(x) w(x) \, dx. \]

For example,

\[ (x - z_1) (x - z_2) p_{s-1}(x) = f_{s-1} p_{s+1} + g_{s-1} p_s + h_{s-1} p_{s-1} + g_{s-2} p_{s-2} + f_{s-3} p_{s-3} \quad (s = 1, 2, \ldots) \]
where
\[ f_{s-1} = A_{s-1}^{-1} A_{s+1}^{-1}, \quad s = 1, 2, \ldots, \]
\[ - A_s B_{s+1} = B_{s+1} A_{s+1}^{-1} + z_1 + z_2 + B_s A_s^{-1}, \quad s = 1, 2, \ldots, \]
\[ h_{s-1} = A_s z_2 + B_s A_s^{-1} + A_s - 1, \quad s = 2, 3, \ldots \]
\[ h_0 = A_s z_2 + B_s A_s^{-1}. \]

From (6) and (7) we find, after using (16),
\[ \phi_r = \frac{(-)^n}{k_r^2} K_{r+1}(e_{00}, e_{01}, e_{02}, \ldots, e_{0n}), \quad r = 0, 1, \ldots \quad (18) \]
and
\[ \sum_{s=0}^{r-1} A_s^2 \phi_s = \frac{(-)^n}{k_r^2} K_{r-1}(e_{10}, e_{11}, \ldots, e_{1n}), \quad r = 1, 2, \ldots, K_0 = 1, \quad (19) \]

where we have introduced the notation \( K_r (e_{00}, e_{01}, \ldots, e_{0n}) \) for a generalised continuant determinant of order \( r \), symmetric, with elements \( e_{00}, e_{10}, e_{20}, \ldots \) in the diagonal through \((1,1)\), \( e_{01}, e_{11}, e_{21}, \ldots \) in the diagonal through \((1,2)\), and so on. We shall refer to these as continuants of the \( n \)th kind. Thus the ratio of continuants of the 1st kind is related to C.F.'s of the first order, the determinants concerned consisting of elements in three diagonals only. Similarly C.F.'s of the second order are associated with the ratio of continuants of the 2nd kind which in turn have elements in five diagonals only.

(c) Writing for simplicity
\[ \mid p_s(x), p_{s+1}(z_1), \ldots, p_{s+n}(z_n) \mid = A_s(x, z_1, z_2, \ldots, z_n) \]
we have
\[ q_s(x) = \frac{A_s(x, z_1, \ldots, z_n)}{k_{s+n} \Pi(z_j - x) A_s(z_1, z_2, \ldots, z_n)} \]
\[ = \frac{1}{k_s A_s(z_1, \ldots, z_n)} \sum_{r=0}^{s} \mid p_r(z_1), p_{s+1}(z_2), p_{s+2}(z_3), \ldots, p_{s+n-1}(z_n) \mid p_r(x) \quad (20) \]
by a generalisation of a theorem of Darboux quoted in [1], (17).

Hence from (6) and (7) we find
\[ \phi_r = \frac{1}{k_r k_{r+n}} A_{r+1}(z_1, z_2, \ldots, z_n) \quad (21) \]
and
\[ \sum_{s=0}^{r-1} A_s^2 \phi_s = \sum_{s=0}^{r-1} \frac{k_{s+n} \mid z_1, p_s+1(z_2), p_s+2(z_3), \ldots, p_{s+n-1}(z_n) \mid^2}{k_s A_s(z_1, z_2, \ldots, z_n) A_{s+1}(z_1, z_2, \ldots, z_n)} \quad (22). \]
4. If we now consider the value of $\Pi_{r}^{-1}$, we have from (18) and (21)

$$D_r(z_1, z_2, \ldots, z_n) = (-)^{n}K_r(e_0, e_1, \ldots, e_n)$$

$$= \frac{A_r(z_1, z_2, \ldots, z_n)}{|z_0^0, z_1^1, \ldots, z_n^{n-1}| \prod k_j^{j}}$$

(23)

since $A_r(z_1, z_2, \ldots, z_n) = \prod_{j=0}^{n-1} k_j |z_0^0, z_1^1, \ldots, z_n^{n-1}|$, and from (12) and (21)

$$D_r(z_1, z_2, \ldots, z_n) = \prod_{j=0}^{r-1} k_j^{j} P_r(M_n, M_{n+1}, \ldots, M_{n+2r-2})$$

(24)

and as a consequence

$$N_r(z_1, z_2, \ldots, z_n) = (-)^{n} k_0^{-2}K_r-1(e_{10}, e_{11}, \ldots, e_{1n})$$

$$= -\prod_{j=0}^{r-1} k_j^{j} \begin{vmatrix} 0 & m_0 & m_1 & \ldots & m_{r-1} \\ m_0 & M_n & M_{n+1} & \ldots & M_{n+r-1} \\ m_1 & M_{n+1} & \vdots & \vdots & \vdots \\ m_{r-1} & M_{n+r-1} & M_{n+2r-2} \end{vmatrix}$$

(25)

$$= \frac{A_r(z_1, \ldots, z_n)}{\prod k_j^{j} |z_0^0, z_1^1, \ldots, z_n^{n-1}|} \times \sum_{s=0}^{r-1} k_s A_s(z_1, \ldots, z_n) A_{s+1}(z_1, \ldots, z_n)$$

(25a)

When the roots $z_j$ are equal, the only change required in (23) is to replace

$$A_r(z_1, z_2, \ldots, z_n)$$

by $p_r(z_1), p_{r+1}(z_1), \ldots, p_{r+n-1}(z_1)$, where superscripts refer to derivatives and

$$(n-1)!! = (n-1)!(n-2)!! \ldots 1!0! .$$

A similar modification is required in (25a).

As an illustration we take $w(x) = 1/\sqrt{1 - x^2}$, $a = -1$, $b = 1$,
DETERMINANTAL EXPANSION FOR A CLASS OF DEFINITE INTEGRAL

with

\[ p_s(x) = \sqrt{\frac{2}{\pi}} \cos s\theta, \quad \cos \theta = x, \quad s = 1, 2, \ldots \]

\[ p_0(x) = \sqrt{\frac{1}{\pi}} k_0; \quad k_s = 2^{s-1} \sqrt{\frac{2}{\pi}}, \quad s = 1, 2, \ldots \]

and

\[ m_s = \frac{1}{1} \int_{-1}^{1} \frac{(z - x)^s}{\sqrt{(1 - x^2)}} \, dx = \Pi (z - 1)^{\alpha} P_{\alpha}(\frac{z}{\sqrt{(z^2 - 1)}}) \]

where \( \alpha \) is Legendre's polynomial. With \( t = z/\sqrt{(z^2 - 1)} \) we find from the modified form of (23), and (24),

\[
\begin{array}{cccccc}
& b_r & b_{r+1} & \ldots & b_{r+n-1} \\
& r c_r & (r+1) c_{r+1} & \ldots & (r+n-1) c_{r+n-1} \\
& r^2 b_r & (r+1)^2 b_{r+1} & \ldots & (r+n-1)^2 b_{r+n-1} \\
& r^3 c_r & (r+1)^3 c_{r+1} & \ldots & (r+n-1)^3 c_{r+n-1} \\
& \vdots & \vdots & \ddots & \vdots \\
\end{array}
\]

\[
\frac{1}{n!} \Gamma(n+1) \Gamma((r-1)^2) \Gamma(r) \Gamma((r-1) \Gamma(r+1)) \Gamma(r+n-1) \Gamma(r+n+1) \\
\]

\[
\begin{cases}
b_r = (t+1)^r + (t-1)^r \\
c_r = (t+1)^r - (t-1)^r \\
2 \chi = (n+r) (n+r-1) + (r-1) (r-2)
\end{cases}
\]

and \( \alpha \) stands for \( \alpha_t \). The result for \( n = 1 \) has been given by Geronimus [3].

5. A formula for the numerators. Consider the identity

\[
| x^0, p(z_1), p_{r+1}(z_2), \ldots, p_{r+n-1}(z_n) | = \Pi (z_j - x) \sum_{s=0}^{n-1} B_s \phi_s(x) \tag{27}
\]

where

\[
B_s \phi_s = \frac{1}{k_s} A_s(z_1, \ldots, z_n)
\]

Then

\[
\int_{a}^{b} \frac{| x^0, p(z_1), \ldots, p_{r+n-1}(z_n) |}{\Pi (z_j - x) A_s(z_1, \ldots, z_n)} w(x) \, dx = \sum_{s=0}^{r-1} k_s A_s(z_1, \ldots, z_n) A_{s+1}(z_1, \ldots, z_n)
\]

Hence from (22) it follows that

\[
N_r(z_1, z_2, \ldots, z_n) = \int_{a}^{b} \frac{| x^0, p(z_1), p_{r+1}(z_2), \ldots, p_{r+n-1}(z_n) |}{\Pi k_{r+1} A_s(z_1, \ldots, z_n)} \, w(x) \, dx \tag{28}
\]
since \( \prod_{1}^{n} (z_j - x) \mid z_1^0, z_2^1, \ldots, z_n^{n-1} \mid = \mid x^0, z_1^1, z_2^2, \ldots, z_n^n \mid \).

When \( z_1 = z_2 = \ldots = z_n = z \), the determinant in the numerator of (28) is to be replaced by \( \mid x^0, p_r(z), p^{(1)}_r(z), \ldots, p^{(n-1)}_{r+n-1}(z) \mid \), and that in the denominator by \( (z - x)^n(n - 1)! \).

In the particular case \( n = 1 \) we have the well-known formula

\[
N_r(z) = k_r^{-1} \int_a^z \frac{p_r(z) - p_r(x)}{z - x} w(x) \, dx,
\]

and from (23) \( D_r(z) = k_r^{-1} p_r(z) \), the ratio of these being the \( r \)th convergent of the C.F.

\[
\frac{A_r k_0^{-2} + C_2}{A_r z + B_1 - A_2 z + B_2 - A_3 z + B_3 - \ldots}.
\]

We shall consider the recurrence relations for the numerators and denominators of generalised C.F.'s, and some special properties of second order C.F.'s in Part 4.

**REFERENCES.**


**COLLEGE OF TECHNOLOGY,**

**MANCHESTER, 1.**

Extracted from the
*Edinburgh Mathematical Proceedings*
A Determinantal Expansion for a Class of Definite Integral

Part 4.

By L. R. SHENTON

(Received 15th October, 1953.)

1. We shall show in this part the relation of generalised C.F.'s to ordinary C.F.'s, in the main confining our attention to Stieltjes type fractions. Moreover we shall bring out the part played by Parseval's theorem in our development of the subject, and a property of extremal solutions of the Stieltjes moment problem given by M. Riesz.\(^1\)

2. The Stieltjes moment problem \(^2\) (S.M.P.) concerns itself with finding a bounded non-decreasing function \(\psi(x)\) in \((0, \infty)\) such that

\[
\int_0^\infty x^n d\psi(x) = \mu_n, \quad n = 0, 1, 2, \ldots,
\]

where the \(\mu\)'s are real. The solution offered by Stieltjes depends upon the characteristics of the C.F. associated with the formal expansion

\[
F(z, \psi) = \int_0^\infty x^n d\psi(x) \sim \frac{\mu_0}{z} - \frac{\mu}{z^2} + \frac{\mu_2}{z^3} - \ldots,
\]

namely

\[
F(z, \psi) \sim \frac{1}{az_1} + \frac{1}{a_z} + \frac{1}{a_4} + \frac{1}{a_5} + \ldots
\]

and the corresponding C.F. (obtained by contraction)

\[
F(z, \psi) \sim \frac{\lambda_1}{z + C_1} - \frac{\lambda_2}{z + C_2} - \frac{\lambda_3}{z + C_3} - \ldots
\]

A necessary and sufficient condition for the existence of a solution of the S.M.P. is that \(a_j > 0, \quad j = 1, 2, 3, \ldots\); the solution is unique if \(\sum_{j=1}^\infty a_j\) diverges. If \(\sum_{j=1}^\infty a_j\) converges there may be an

---

\(^1\) M. Riesz, "Sur le problème des moments," Arkiv för matematik, astronomi och fysik, 16 (12), 1-21; 16 (19), 1-21; 17 (16) 1-52.

\(^2\) See, for example, J. A. Shohat and J. D. Tamarkin, The Problem of Moments (American Mathematical Society Surveys No. 1, 1943).
infinity of solutions; of these there are two called extremal solutions, which have the property that \{\omega_{2s}(-z)\} and \{\omega_{2s+1}(-z)\} are orthogonal systems with respect to them, where \( \bar{\chi}_s(z)/\omega_s(z) \) is the \( s \)th convergent of (2).

The moments \( \mu_j \) and the elements \( a_j \) are related as follows:

\[
a_{2j} = \frac{\Delta_j^2}{\Delta_{j-1} \Delta_j}, \quad a_{2j+1} = \frac{\Delta_j^2}{\Delta_j \Delta_{j+1}}
\]

where

\[
\begin{align*}
\Delta_0 &= 1, \quad \Delta_1 = \mu_0, \quad \Delta'_0 = 1, \quad \Delta'_1 = \mu_1.
\end{align*}
\]

Stieltjes showed that when there is a solution of the S.M.P. then

\[
F(z, \psi) - \bar{\chi}_{2s}(z)/\omega_{2s}(z) = \min \frac{\int_0^\infty p_s^2(x) d\psi(x)}{z + x}
\]

\[
z\{\bar{\chi}_{2s+1}(z)/\omega_{2s+1}(z) - F(z, \psi)\} = \min \frac{\int_0^\infty x p_s^2(x) x d\psi(x)}{z + x}
\]

where \( p_s(-z) = 1, \ z > 0 \), and the minimum is taken over all polynomials of degree less than or equal to \( s \). For our present discussion it is important to write these in the form

\[
F(z, \psi) - \bar{\chi}_{2s+2}(z)/\omega_{2s+2}(z) = \min \left( \frac{1}{z + x} - \Pi_s(x) \right)^2
\]

\[
z\{\bar{\chi}_{2s+3}(z)/\omega_{2s+3}(z) - F(z, \psi)\} = \min \left( \frac{1}{z + x} - \Pi_s(x) \right)^2
\]

the forms (5) and (6) showing the obvious relation to Parseval's theorem. Indeed from a formal point of view all we have to do to find the even part of (2) is to write down the Parseval expansion for

\[
\int_0^\infty f(x)^2 d\bar{\psi}(x) \quad \text{where} \quad f(x) = (z + x)^{-1}, \quad \bar{\psi}(x) = \int_0^x (z + t) d\psi(t)
\]

and the Tchebicheffian orthogonal polynomials with respect to \( d\psi(x) \) are related to those for \( d\bar{\psi}(x) \) by the theorem of Christoffel (Part 1, 14). Similarly the odd part of (2) follows from \( \int_0^\infty f(x)^2 d\psi(x) \) where \( \psi(x) = \int_0^x f(\bar{z} + t) d\psi(t) \).
On the other hand it has been proved by Riesz (loc. cit.) that if \( \psi(x) \) is an extremal solution of the S.M.P. (or the unique solution) then Parseval's theorem applies to all \( f(x) \in L^2_\psi \). An extension of this which we require is the following: If \( \psi(x) \) is an extremal solution of the S.M.P., then Parseval's theorem applies to \( f(x) \in L^2_\psi \) where \( \psi(x) = \int_0^x \Pi(t)d\psi(t) \) and \( \Pi(t) \) is a non-negative polynomial of fixed degree which is not identically zero. Clearly \( \psi(x) \) is bounded and non-decreasing and the moments are given by

\[
\mu'_s = \int_0^\infty x^s d\psi(x) = \int_0^\infty x^s \Pi(x) d\psi(x)
\]

\[= \sum_{(r)} \gamma_r \mu_{r+s} \text{ where } \Pi(x) = \sum_{(r)} \gamma_r x^r.
\]

Thus if \( \psi(x) \) is the unique solution for a given sequence \( \{\mu_s\} \) (assuming such a solution exists) then \( \int_0^x \Pi(t)d\psi(t) \) is the unique solution for the sequence \( \{\Pi(\mu)\} \), provided \( \Pi(x) \) is a non-negative polynomial. It is of interest to recall that the S.M.P. is determined in the particular cases

\[
\mu_s = \int_0^\infty x^s x^{b-1} \exp(-kx^b) \, dx, \quad b > 0, \quad k > 0, \quad a \geq \frac{1}{2}
\]

(7)

\[
\mu_s = \int_0^\infty x^s f(x) x^{b-1} \exp(-kx^a) \, dx
\]

(8)

where \( f(x) \) is a positive bounded function on \((a, \infty)\), \( a > 0 \). Moreover quite apart from the theory of continued fractions, Hardy \(^2\) proved that the S.M.P. is determined for \( \psi(t) = \int_0^t \phi(x) \, dx, \phi(x) \geq 0 \), provided

\[
\int_0^\infty [\phi(t)]^q e^{\delta t} \, dt < \infty \text{ for } q \geq 1, \delta > 0.
\]

It will be noticed that the uniqueness of the moment problem \( \mu_n = \int_0^\infty x^n \phi(x) \, dx \) depends upon the order of magnitude of \( \phi(x) \) for large positive \( x \). Thus \( \phi(x) = \exp -x^a \) does not approach zero rapidly enough for \( x \to \infty \), and the Stieltjes C.F. corresponding to

\[
\int_0^\infty \exp -x^a \, dx = F(z) \text{ diverges by oscillation. But making the substitu-}
\]

\(^1\) T. J. Stieltjes, Oeuvres Complètes, Vol. 2, pp. 505-506, 518-520.

tion \( x = t^2 \) we have \( F(z) = \int_0^\infty \frac{e^{-y/t}dt}{z + t^2} \) and we shall show in the sequel that the second order C.F. for \( F(z) \) converges. In general it may be remarked that if \( \psi(x) \) is the solution of a determined S.M.P. then the C.F. (2), with elements given by (4) in terms of the moments, converges for \( z > 0 \); but the Stieltjes C.F. corresponding to \( \int_0^\infty \frac{d\psi(x)}{z + x^2} = F(z) \) may not converge, \( s \) being a positive integer greater than unity. However, the \( s^{\text{th}} \) order C.F. corresponding to \( F(z) \) does converge in this case.

2.0 We now state some properties of the convergents of Stieltjes C.F.'s which we require. We consider the expansions

\[
F(z) = \int_0^\infty \frac{d\psi(x)}{x + z} = \frac{b_1}{z + 1} + \frac{b_2}{z + 1 + 1} + \cdots \quad (= \chi_s(z) \text{ as } s \to \infty)
\]

\[
= \frac{b_1}{z + b_2 - z + b_3 - z + b_4 - z + b_5 + b_6 - \cdots} \quad (= \chi_2s(z) \text{ as } s \to \infty)
\]

assumed to be convergent for \( z > 0 \). Then we have the recurrence relations

\[
\begin{align*}
\omega_{2s}(z) &= (z + b_{2s-1} + b_{2s})\omega_{2s-2}(z) - b_{2s-1}b_{2s-2}\omega_{2s-4}(z) \\
\omega_{2s+1}(z) &= (z + b_{2s} + b_{2s+1})\omega_{2s-1}(z) - b_{2s-1}b_{2s} - b_{2s-1}\omega_{2s-3}(z) \quad s = 2, 3, \ldots
\end{align*}
\]

and similarly for \( \chi_s(z) \) with \( \chi_0 = 0, \chi_1 = b_1; \omega_0 = 1, \omega_1 = z \), from which we derive the determinantal relations

\[1\]

The following abbreviated notation for alternant types of determinants will be used throughout:

\[
\begin{vmatrix}
A_r(z_1), & B_s(z_2), & C_t(z_3)
\end{vmatrix}
\]

\[
\begin{pmatrix}
A_r(z_1) & B_r(z_2) & C_r(z_3) \\
A_s(z_1) & B_s(z_2) & C_s(z_3) \\
A_t(z_1) & B_t(z_2) & C_t(z_3)
\end{pmatrix}
\]

where any functional symbol cannot be separated from its argument. Thus

\[
\begin{vmatrix}
\chi_2r(z_1), & \omega_{2r+2}(z_2)
\end{vmatrix}
\]

\[
\begin{pmatrix}
\chi_{2r}(z_1) & \omega_{2r}(z_2) \\
\chi_{2r+2}(z_1) & \omega_{2r+2}(z_2)
\end{pmatrix}
\]

but \( \omega_r(z_1), \omega_{2r+2}(z_2) \) is unambiguous. Similarly when the symbol of functionality is tied to its suffix we shall write

\[
\begin{vmatrix}
A_r(z_1), B_s(z_2), C_t(z_3)
\end{vmatrix}
\]

\[
\begin{pmatrix}
A_r(z_1) & B_s(z_2) & C_t(z_3) \\
A_s(z_1) & B_s(z_2) & C_t(z_3) \\
A_t(z_1) & B_t(z_2) & C_t(z_3)
\end{pmatrix}
\]

Thus

\[
\begin{vmatrix}
f_r(x), p_r(z)
\end{vmatrix}
\]

\[
\begin{pmatrix}
f_r(x) & p_r(x) \\
f_j(z) & p_r(z)
\end{pmatrix}
\]
The relation between the even and odd convergents is given by

\[
\begin{align*}
| \omega_{2s}(0), \chi_{2s+2}(z) \rangle &= \omega_{2s}(0) \chi_{2s+1}(z), \quad \omega_{2s}(0) = \prod_{r=1}^{s} b_{2r}, \\
| \omega_{2s}(0), \omega_{2s+2}(z) \rangle &= \omega_{2s}(0) \omega_{2s+1}(z), \quad \chi_{2s+1}(0) = \prod_{r=0}^{s} b_{2r+1}.
\end{align*}
\]

(13)

From (12) it is easily proved that

\[
\begin{align*}
| \chi_{2r}(z_1), \omega_{2r+2}(z_1), \omega_{2r+4}(z_2) \rangle &= (z_1 - z_2) \omega_{2r+2}(z_2) \prod_{s=1}^{2r+1} b_s \\
| \chi_{2r+1}(z_1), \omega_{2r+3}(z_1), \omega_{2r+5}(z_2) \rangle &= z_1(z_2 - z_1) \omega_{2r+3}(z_2) \prod_{s=1}^{2r+2} b_s.
\end{align*}
\]

(14)

2.1 The orthonormal polynomials. We introduce the system \( \{ p_r(x) \} \) where

\[
\int_{0}^{\infty} p_r(x) p_s(x) d\psi(x) = \delta_{rs},
\]

(15)

with recurrence relation

\[
\begin{align*}
p_r(x) &= (A_r x - B_r) p_{r-1}(x) - C_r p_{r-2}(x), \quad p_{-1} = 0, \quad r = 1, 2, \ldots, \\
p_0(x) &= k_0, \quad p_1(x) = (A_1 x - B_1) k_0
\end{align*}
\]

(16)

and \( A_r = k_r/k_{r-1} > 0, \quad C_r = A_r/A_{r-1} > 0, \)

where \( k_r > 0 \) is the highest coefficient in \( p_r(x) \). Clearly \( B_r > 0, \) for

\[
B_r = A_r \int_{0}^{\infty} x p_{r-1}^2(x) d\psi(x).
\]

Moreover

\[
\begin{align*}
B_r/A_r = b_{2r} + b_{2r-1}, \quad &B_1/A_1 = b_2 \\
A_r^{-2} = b_{2r+1} b_{2r}, \quad &A_1^{-2} = b_3 b_2 \\
k_r^{-2} = \prod_{s=1}^{2r+1} b_s
\end{align*}
\]

(17)

by comparison with (11). We also require
\[ \begin{cases} k_r \omega_{2r}(z) = (-)^r p_r(-z) \\ k_r X_{2r}(z) = (-)^r \int_0^x \frac{p_r(-z) - p_r(x)}{z + x} \, d\psi(x) \end{cases} \]  
\[ r = 0, 1, 2, \ldots \]

3. A fundamental identity. We shall now prove an identity which relates the generalised convergents of
\[ F(z_1, z_2, \ldots z_n) = \int_0^\infty \frac{d\psi(x)}{f_n(x)}, \]
where \( f_n(x) = \prod_{\lambda=1}^n (x + z_\lambda) \), to those of \( F(z) \).

We consider
\[ \Delta_j(x) = \left| \int f_j(x) \prod_{r=1}^{j-1} \frac{p_r(-z_1), p_{r+1}(-z_2), \ldots p_{r+n-1}(-z_n)}{p_r(x)} \, d\psi(x) \right| \]
\[ = f_n(x) \sum_{s=0}^{r-1} A_{sj}(r) g_s(x) \]
\[ j = l, m; \quad r = 1, 2, \ldots, \]
where
\[ f_l(x) = \prod_{\lambda=1}^l (x + x_\lambda), \quad f_m(x) = \prod_{\lambda=1}^m (x + y_\lambda), \quad l < n, m < n, \]
\( f_l(x), f_m(x), f_n(x) \) are polynomials in \( x \) with real coefficients, \( f_n(x) > 0 \)
for \( x > 0 \) with distinct roots, and
\[ \int_0^x q_r(x) g_s(x) f_n(x) \, d\psi(x) = 0 \]
\[ = \phi_r, \quad r = s. \]

But from Part 1, paragraph 4, we have
\[ q_r(x) = (-)^n k_r + \sum_{s=0} \left| p_s(-z_1), p_{s+1}(-z_2), \ldots p_{s+n-1}(-z_n) \right| p_c(x) \]
\[ \phi_r = (-)^n k_r + \left| p_r(-z_1), \ldots p_{r+n-1}(-z_n) \right| p_{r+1}(-z_1), \ldots p_{r+n}(-z_n) \].

If now \( \psi(x) \) is a solution of a determined S.M.P., and the integrals
\[ \int_0^\infty \frac{(f_l(x))^2}{f_n(x)} \, d\psi(x) \]
\[ \int_0^\infty \frac{(f_m(x))^2}{f_n(x)} \, d\psi(x) \]

converge, then Parseval's theorem applies to the functions \( f_l(x)/f_n(x) \) and \( f_m(x)/f_n(x) \) giving, with respect to the distribution function \( \int_0^x f_n(t) \, d\psi(t) \),
\[ F(z_1, z_2, \ldots z_n) = \int_0^\infty \frac{f_l(x)f_m(x)}{f_n(x)} \, d\psi(x) = \sum_{j=0}^\infty A_{lj} A_{mj} \phi_j \]
\[ j = 0, 1, 2, \ldots \]
We also write
\[ \sum_{j=0}^{r-1} A_{ij} A_{mj} \phi_j = \chi_r(z_1, z_2, \ldots, z_n) = \chi_r(z) \]
and (24) gives the \( r \)th convergent of the \( n \)th order C.F. for \( F(z_1, z_2, \ldots, z_n) \).

But from (19)
\[ \int_0^\infty \frac{\Delta_l(x) \Delta_m(x)}{f_n(x)} \, d\psi(x) = \sum_{j=0}^{r-1} A_{ij} \bar{A}_{mj} \phi_j, \quad r = 1, 2, \ldots \]
where \( \bar{A}_{ij}(r) = | p_r(z_1) \cdots p_{r+n-1}(z_n) | A_{ij} \quad j = 0, 1, \ldots, r-1 \).

Hence from (24)
\[ \frac{\chi_r(z)}{\omega_r(z)} = | p_r(z_1), \ldots, p_{r+n-1}(z_n) | \int_0^\infty \frac{\Delta_m(x)f_i(x)}{f_n(x)} \, d\psi(x). \]

But since \( l < n \) and the roots of \( f_n(x) \) are distinct, we have
\[ f_i(x)f_n(x) = (-1)^{n-1} \begin{vmatrix} z_1^n & z_1^{n-1} & \cdots & 1 \\ z_2^n & z_2^{n-1} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ z_n^n & z_n^{n-1} & \cdots & 1 \end{vmatrix} \]

Using (27) and (18) in (26) we have
\[ \frac{\chi_r(z)}{\omega_r(z)} = \begin{vmatrix} g & f_m(z_1) & f_m(z_2) & \cdots & f_m(z_n) \\ g_1 & \omega_1(z_1) & \omega_1(z_2) & \cdots & \omega_1(z_n) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{r+n-1} & \omega_{r+n-2}(z_1) & \omega_{r+n-2}(z_2) & \cdots & \omega_{r+n-2}(z_n) \end{vmatrix} \]

where
\[ g = (-1)^n \begin{vmatrix} z_0^n & z_0^{n-2} \cdots z_0 & F_m(z_0)f_r(z_0) \end{vmatrix} / \begin{vmatrix} z_1^n & z_1^{n-2} \cdots z_1 \\ \vdots & \vdots & \ddots & \vdots \\ z_n^n & z_n^{n-2} \cdots z_n \end{vmatrix} \]
\[ g_s = (-1)^n \begin{vmatrix} z_s^n & z_s^{n-2} \cdots z_s & X_m(z_s)f_r(z_s) \end{vmatrix} / \begin{vmatrix} z_1^n & z_1^{n-2} \cdots z_1 \\ \vdots & \vdots & \ddots & \vdots \\ z_n^n & z_n^{n-2} \cdots z_n \end{vmatrix} \]
\[ F_m(z_s) = \int_0^\infty \frac{f_m(z_s)}{z_s+x} \, d\psi(x), \quad s = r, r+1, \ldots, r+n-1 \]

and (28) consists of the ratio of the determinant of order \( n+1 \) and the determinant obtained from this by deleting its first row and column. An alternative to (28) appears by using the partial fraction...
form of the elements in the first column, and we find

\[
\frac{X_r(z_\lambda)}{\omega_r(z_\lambda)} = - \sum_{s=1}^{n} \frac{F_s(z_1)f_s(-z_s)}{\prod_{r=1}^{n} (z_r-z_s)} f_m(-z_1) f_m(-z_2) \ldots f_m(-z_n)
\]

\[
\sum_{s=1}^{n} \chi_{2s}(z_2)f_s(-z_s)
\]

\[
\sum_{s=1}^{n} \chi_{2s+2n-2}(z_1)f_s(-z_s)
\]

\[
\sum_{s=1}^{n} \chi_{2s+2n-2}(z_2)f_s(-z_s)
\]

\[
\sum_{s=1}^{n} \chi_{2s+2n-2}(z_\lambda)f_s(-z_s)
\]

(29)

If we take \( \omega_r(z_\lambda) = | \omega_{2s}(z_1), \ldots, \omega_{2r+2n-2}(z_n) | / | z_1^0, z_1^1, \ldots, z_n^{n-1} | \) then (28) and (29) give expressions for \( X_r(z_\lambda) \) in terms of \( \chi_{2s}(z_\lambda) \) and \( \omega_{2s}(z_\lambda) \), \( s = r \) to \( r + n - 1 \), \( \lambda = 1 \) to \( n \). We note that if in addition to \( l < n \), we have \( m < n \), then \( l \) and \( m \) may be interchanged in (28) and (29) yielding an identity between two forms of the numerator \( \chi_{2s}(z_\lambda) \).

The confluent case of (28) - (29) in its general form is complicated, but particular cases may be obtained from first principles. Thus if \( z_1 = z_2 = \ldots = z_j, j \leq n \), then the limiting form appears by letting \( z_2 \rightarrow z_1, z_3 \rightarrow z_1, \ldots, z_j \rightarrow z_1 \) in succession, and subtracting appropriate columns. The complication arises from the fact that \( f_l(x) \) and \( f_m(x) \) may be functions of \( z_\lambda, \lambda = 1 \) to \( n \).

4.0 We shall now consider generalised continued fraction expansions for \( F(z_1, z_2, \ldots z_n) \) of four kinds, namely

(i) convergent increasing sequences,

(ii) convergent decreasing sequences,

(iii) convergent sequences,

(iv) convergent sequences involving an arbitrary parameter.

In the main we shall confine our attention to second order C.F.'s for \( F(z_1, z_2) \). We assume that \( \psi(x) \) is the unique solution of a S.M.P.
4.1 Increasing sequences. In (29) take \( f_i = f_m = 1 \), 
\( f_n(x) = (x + z_1)(x + z_2) > 0, z_1 \neq z_2 (z_2 = \bar{z}_1 \text{ if } z_1 \text{ is complex}) \) so that \( F_m(z_2) = 0 \), and 
\[
F(z_1, z_2) = \int_0^\infty \frac{dy(x)}{(x + z_1)(x + z_2)}
\]
\[
= \lim_{r \to \infty} \frac{1}{z_1 - z_2} \begin{vmatrix}
1 & 1 & 1 \\
\chi_2(z_1) & \chi_2(z_2) & \omega_2(z_1) \\
\chi_{2r+2}(z_1) & \chi_{2r+2}(z_2) & \omega_{2r+2}(z_1) \\
\omega_2(z_1) & \omega_2(z_2) & \omega_{2r+2}(z_1) \\
\omega_{2r+2}(z_1) & \omega_{2r+2}(z_2) & \omega_{2r+2}(z_1)
\end{vmatrix}
\]
where the expansion, in view of (22), is an increasing sequence. By (12) this may be written
\[
F(z_1, z_2) = \lim_{r \to \infty} \frac{1}{z_1 - z_2} \begin{vmatrix}
\chi_2(z_1) & \omega_2(z_1) & \omega_2(z_1) \\
\chi_{2r+2}(z_1) & \chi_{2r+2}(z_2) & \omega_{2r+2}(z_1) \\
\omega_2(z_1) & \omega_{2r+2}(z_1) & \omega_{2r+2}(z_1) \\
\omega_{2r+2}(z_1) & \omega_{2r+2}(z_2) & \omega_{2r+2}(z_1)
\end{vmatrix}
\]
where we use the abbreviation l.i.s. for limit of the increasing sequence.¹

If in particular \( z_1 = z_2 = z \) then by letting \( z_r \to z_1 \) in (30) we have
\[
F(z) = \int_0^\infty \frac{dy(x)}{(z + x)^{n+1}} = \text{l.i.s.} \begin{vmatrix}
\chi^{(n)}(z) & \omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z) \\
\chi^{(n)}(z) & \omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z) \\
\omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z) \\
\omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z)
\end{vmatrix} > 0.
\]

The general formula of this type is found similarly from (28) and gives
\[
\int_0^\infty \frac{dy(x)}{(z + x)^{n+1}} = \text{l.i.s.} \begin{vmatrix}
\chi^{(n)}(z) & \omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z) \\
\chi^{(n)}(z) & \omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z) \\
\omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z) \\
\omega^{(1)}(z) & \omega^{(2)}(z) & \ldots \omega^{(n)}(z)
\end{vmatrix}
\]
\[z > 0.
\]

¹ As a particular example suppose that by using (4) and an equivalence transformation we find the convergent expansion
\[
\frac{1}{2} \int_0^\infty x e^{-\sqrt{x} \frac{z}{x + z}} dx = \lim_{r \to \infty} \chi_r(z) = b_1 b_2 b_3 b_4 + \ldots \\
\sim 3! 5! 7! 9! \\
\frac{z}{z^2} + \frac{z^3}{z^4} + \ldots
\]
Then by (30) with \( z = \frac{i}{t} \), \( t > 0 \), we have a convergent expansion for
\[
F(i t, - i t) = \frac{1}{2} \int_0^\infty x e^{-\sqrt{x} \frac{z}{x + z}} dx = \frac{3!}{t^2 7!} + \frac{11!}{l^6} - \frac{15!}{l^8} + \ldots
\]
But the Stieltjes C.F for \( F(i t, - i t) = \frac{1}{2} \int_0^\infty \exp -\frac{x}{x + l^2} dx \) diverges by oscillation.
There are of course other forms of increasing sequences for $F(z_1, z_2)$; for example we could use $f_x = x$, $f_m = x$, $f_n = x^2 (x + z_1)(x + z_2)$ in (28). But (30) seems to be the simplest of this type.

4. 2 Decreasing sequences. When the roots $z_1, z_2$ are distinct and $(x + z_1)(x + z_2) = x^2 + 2px + q$, $q - p^2 > 0$, then we use the relation

$$ (q - p^2) F(z_1, z_2) = b_1 - \int_{0}^{\infty} \frac{(x + p)^2 d\psi(x)}{(x + z_1)(x + z_2)}. \quad (33) $$

Taking in (29) $f_x = f_m = x + p$, $f_n = (x + z_1)(x + z_2)$, $F_t(z) = -b_1$, we have

$$ (q - p^2) F(z_1, z_2) = b_1 - \text{l.i.s.} \begin{vmatrix} b_1 & \frac{1}{2}(z_2 - z_1) & \frac{1}{2}(z_1 - z_2) \\ \frac{1}{2}(z_2 + z_1) & \omega_{2r}(z_1) & \omega_{2r}(z_2) \\ \frac{1}{2}(z_2 + z_1) & \omega_{2r+2}(z_1) & \omega_{2r+2}(z_2) \end{vmatrix} \left(\frac{z_2 - z_1}{z_1 - z_2}\right). \quad (34) $$

and after using (12) this leads to

$$ F(z_1, z_2) = \text{l.d.s.} \frac{\left| \frac{\chi_{2r}(z_1)}{z_1 - z_2} + \frac{\chi_{2r+2}(z_2)}{z_1 - z_2} \right| + \left| \frac{\chi_{2r}(z_2)}{z_1 - z_2} + \frac{\chi_{2r+2}(z_1)}{z_1 - z_2} \right| - 2 \sum_{s=1}^{r+1} b_s}{(z_1 - z_2) \left| \omega_{2r}(z_2), \omega_{2r+2}(z_1) \right|}. \quad (35) $$

When $q - p^2 > 0$ it will be seen that the difference between corresponding convergents of (30) and (34) is

$$ \frac{-4 \sum_{s=1}^{r+1} b_s}{(z_1 - z_2) \left| \omega_{2r}(z_2), \omega_{2r+2}(z_1) \right|}. \quad (35) $$

and this exceeds the absolute error in either of them.

Again, taking $f_x = f_m = (x + p)^2, f_n = (x + z_1)(x + z_2)(x + p)^2, q - p^2 > 0$ in (28) (taking the limiting form with $z_3 = z_4 = p$), we find

$$ F(z_1, z_2) = \text{l.d.s.} \frac{\chi_{2r}(z_1) - \chi_{2r}(z_2)}{z_1 - z_2} \frac{\chi_{2r+2}(z_1) - \chi_{2r+2}(z_2)}{z_1 - z_2} \frac{\chi_{2r+4}(z_1) - \chi_{2r+4}(z_2)}{z_1 - z_2} \frac{\chi_{2r+6}(z_1) - \chi_{2r+6}(z_2)}{z_1 - z_2} \begin{vmatrix} 0 & 1 & 1 & 0 & 0 \\ \omega_{2r}(z_1) & \omega_{2r}(z_2) & \omega_{2r}(p) & \omega_{2r}(p) \\ \omega_{2r+2}(z_1) & \omega_{2r+2}(z_2) & \omega_{2r+2}(p) & \omega_{2r+2}(p) \\ \omega_{2r+4}(z_1) & \omega_{2r+4}(z_2) & \omega_{2r+4}(p) & \omega_{2r+4}(p) \\ \omega_{2r+6}(z_1) & \omega_{2r+6}(z_2) & \omega_{2r+6}(p) & \omega_{2r+6}(p) \end{vmatrix}. \quad (36) $$

1. l.d.s. means limit of the decreasing sequence.
In particular if $p = 0, q > 0$, then after using (12) and (13) we find
\[
\int_0^\infty \frac{d\psi(x)}{x^2 + q} \left| \chi_{2r+1}(z_1), \omega_{2r+3}(z_2), \omega'_{2r+5}(0) \right|^2 + \left| \chi_{2r+1}(z_2), \omega_{2r+3}(z_1), \omega'_{2r+5}(0) \right|^2 - 2q \omega_{2r+3}(2p) \prod_{s=1}^{2r+2} b_s
\]
\[
= \text{l.d.s.} \frac{(z_2 - z_1) \left| \omega_{2r+1}(z_1), \omega_{2r+3}(z_2), \omega'_{2r+5}(0) \right|^2}{(z_2 - z_1) \left| \omega_{2r+1}(z_1), \omega_{2r+3}(z_2), \omega'_{2r+5}(0) \right|^2}.
\]

Another interesting possibility is to use the expression
\[
q F(z_1, z_2) = b_1 - \int_0^\infty \frac{x(x + 2p) d\psi(x)}{(x + z_1)(x + z_2)}, \quad q > 0,
\]
where $p > 0, (x + z_1)(x + z_2) = x^2 + 2px + q$, taking $f_i = f_m = x(x + 2p)$,
\[
f_n = x(x + 2p)(x^2 + 2px + q) \quad \text{in (29)}
\]
with $z_1, z_2, z_3 = 2p, z_4 = 0$ and
\[
P_i(z_i) = (z_i - 2p)b_i - a_i
\]
\[
\sum_{s=1}^{4} F_i(z_i) f_m(-z_i) = -b_1, \quad \sum_{s=1}^{4} \chi(z_s) f_m(-z_s) = \chi(z_1) - \chi(z_2).
\]

Using (12) and (13) we find after some simplification
\[
\int_0^\infty \frac{d\psi(x)}{(x + z)^2} = \text{l.d.s.} - \left| \chi'_{2r+1}(z), \omega'_{2r+3}(z), \omega'_{2r+5}(2z) \right|^2, \quad z > 0.
\]

4.3 Convergent sequences. The approximations considered in 4.1 and 4.2 provide lower and upper bounds, but there are other approximations which merely converge. We shall briefly consider four simple types, derived from (28) – (29).

(i) $f_i = x, f_m = 1, f_n = x(x + z_1)(x + z_2), \quad z_1 + z_2, \quad z_1, z_2 > 0$.

Then $F(z_1, z_2) =
\[
\lim_{r \to \infty} \left| \chi_{2r+1}(z_1), \omega_{2r+3}(z_2) \right|^2 + \left| \chi_{2r+1}(z_2), \omega_{2r+3}(z_1) \right|^2 - (z_1 + z_2) \prod_{s=1}^{2r+2} b_s - (z_1 - z_2)(\omega_{2r+2}(z_1) - \omega_{2r+2}(z_2)) \prod_{s=1}^{2r+2} b_{s+1}
\]
\[
= \text{l.d.s.} \frac{(z_1 - z_2) \left| \omega_{2r+1}(z_1), \omega_{2r+3}(z_1) \right|^2}{(z_1 - z_2) \left| \omega_{2r+1}(z_1), \omega_{2r+3}(z_1) \right|^2}.
\]
In particular if \( z_1 = z_2 \), using (12) we find
\[
F(z, z) = \lim_{r \to \infty} \frac{\left| \chi'_{2r+1}(z), \omega_{2r+3}(z) \right| + \omega'_{2r+3}(z) \prod_{0}^{r} b_{2r+1}}{\left| \omega'_{2r+1}(z), \omega_{2r+3}(z) \right|}.
\]

(ii) \( q - p^2 > 0 \). In (33) take \( f_n = (x + p)^2, f_l = 1, f_n = x^2 + 2px + q \) and we find an expression which is exactly the same as (30). This brings to light an interesting identity, for we have the two expansions
\[
\int_{0}^{x} (x + p)^2 d\psi(x) = b_1 - \frac{1}{4}(z_1 - z_2) \left| X_{2r}(z_1), \omega_{2r+2}(z_2) \right| + \left| \chi_{2r}(z_2), \omega_{2r+2}(z_1) \right| - 2 \prod_{1}^{2r+1} b_1
\]

\[
= \lim_{r \to \infty} \left\{ b_1 - \frac{1}{4}(z_1 - z_2) \left| X_{2r}(z_1), \omega_{2r+2}(z_2) \right| + \left| \chi_{2r}(z_2), \omega_{2r+2}(z_1) \right| + 2 \prod_{1}^{2r+1} b_1 \right\}
\]

\[ q - p^2 > 0 \text{ or } z_1, z_2 > 0, (z_1 - z_2) \]

The difference between corresponding convergents of the two expansions in (43) is therefore \( (z_1 - z_2) \prod_{1}^{2r+1} b_s / \left| \omega_s(z_1), \omega_{2r+2}(z_2) \right| \).

In terms of the persymmetric determinants\(^1\) mentioned in Part 3, 3 (a) this comes to
\[
\begin{pmatrix}
0 & \mu_0^{(1)} & \mu_1^{(1)} & \cdots & \mu_r^{(1)} \\
\mu_0^{(1)} & -\mu_0 & -\mu_1 & \cdots & -\mu_r \\
\mu_1^{(1)} & -\mu_1 & -\mu_2 & \cdots & -\mu_{r+1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mu_r^{(1)} & -\mu_r & -\mu_{r+1} & \cdots & -\mu_{2r} \\
\end{pmatrix}
= r + 1 k_s^{-2}
\]

where
\[
\mu_s^{(0)} = \int_{0}^{x} x^s d\psi(x), \quad \mu_s^{(1)} = \int_{0}^{x} (x + p)x^s d\psi(x),
\]
\[
\mu_s^{(2)} = \int_{0}^{x} (x + p)^2x^s d\psi(x), \quad \bar{\mu}_s = \int_{0}^{x} (x^2 + 2px + q)x^s d\psi(x)
\]
\[
k_s^{-2} = \prod_{1}^{r+1} b_r
\]

and \( x^2 + 2px + q \) is non-negative for \( 0 \leq x < \infty \).

\(^1\) There is a similar identity for the diagonal determinants given in Part 3, 3 (b).
Determmantal Expansion for a Class of Definite Integral

(iii) If we use (38) with \( f_i = x, f_m = x + 2p, f_n = (x + z_1)(x + z_2) \) there follows the expansion

\[
F(z_1, z_2) = \lim_{r \to \infty} \frac{\left| \chi_{2r}(z_1), \omega_{2r+2}(z_2) \right|^+ + \left| \chi_{2r+1}(z_2), \omega_{2r+2}(z_1) \right|^+ + (z_1 z_2)^{-1}(z_1^2 + z_2^2)^{2r+1} \Pi b_s}{\left( z_1 - z_2 \right) \left| \omega_{2r}(z_2), \omega_{2r+2}(z_1) \right|}
\]

(45)

\[ z_1, z_2 > 0, \quad z_1 = z_2 \quad \text{or} \quad q - p^2 > 0, \]

and in particular

\[
F(z, z) = \lim_{r \to \infty} \frac{\left| \omega'_{2r}(z), \chi'_{2r+2}(z) \right|^+ + z^{-2} \Pi b_s}{\left| \omega_{2r+1}(z), \omega_{2r+3}(z) \right|}.
\]

(46)

(iv) Using (38) with \( f_i = x, f_m = x(x + 2p), f_n = x(x^2 + 2px + q) \) we find

\[
F(z_1, z_2) = \lim_{r \to \infty} \frac{\left| \chi_{2r+1}(z_1), \omega_{2r+3}(z_2) \right|^+ + \left| \chi_{2r+2}(z_2), \omega_{2r+3}(z_1) \right|^+ - (z_1 + z_2)^{2r+2} \Pi b_s}{\left( z_1 - z_2 \right) \left| \omega_{2r+1}(z_2), \omega_{2r+2}(z_1) \right|}
\]

(47)

\[ z^2 + 2px + q > 0 \quad \text{for} \quad x \geq 0 \]

and in particular

\[
F(z, z) = \lim_{r \to \infty} \frac{\left| \omega'_{2r+1}(z), \omega'_{2r+3}(z) \right|^+}{\left| \omega_{2r+1}(z), \omega_{2r+3}(z) \right|}, \quad z > 0. \]

(48)

4.4 Expansions with arbitrary parameters. An unusual type of expansion appears if in (29) we take \( f_i = f_m = x + p, f_n = (x + z)(x + p)^2 \) where we assume \( z > 0 \), and \( p \) real. Then

\[
F(z) = \int_0^\infty \frac{d\psi(x)}{z + x} = \text{l.i.s.} \lim_{r \to \infty} \frac{\left| \chi_{3r}(z) - \chi_{3r}(p), \omega_{2r+3}(p), \omega'_{2r+4}(p) \right|^+ + k^{-2}(\omega_{2r+3}(p) - \omega_{2r+3}(z)) \Pi b_s}{\left| \omega_{2r}(z), \omega_{2r+1}(p), \omega'_{2r+4}(p) \right|}
\]

(49)

for all real \( p \neq z \).

If \( p = z \) then

\[
F(z) = \text{l.i.s.} \lim_{r \to \infty} \frac{\left| \omega_{2r}(z), \omega'_{2r+3}(z), \chi'_{2r+4}(z) \right|^+ - k^{-2} \omega''_{2r+3}(z), \quad z > 0. \]

(50)

Similarly from \( f_i = f_m = x(x + p), f_n = x(x + z)(x + p)^2 \) we find

\[
F(z) = \text{l.d.s.} \lim_{r \to \infty} \frac{\left| \chi_{2r+1}(z) - \chi_{2r}(p), \omega_{2r+3}(p), \omega'_{2r+5}(p) \right|^+ - p z^{-1}(\omega_{2r+3}(p) - p \omega_{2r+3}(z)) \Pi b_s}{\left| \omega_{2r+1}(z), \omega_{2r+3}(p), \omega'_{2r+5}(p) \right|}
\]

(51)

\[ z > 0, \quad p \neq 0, \quad p + z. \]
In particular if \( p = 0 \) then for \( z > 0 \)

\[
F(z) = \lim_{r \to \infty} \frac{|x_{2r+1}(z) - x_{2r+1}(0), \omega'_{2r+3}(0), \omega''_{2r+5}(0) |}{| \omega_{2r+1}(z), \omega'_{2r+3}(0), \omega''_{2r+5}(0) |} 2r+2 \Pi b_s
\]

and if \( p = z > 0 \) then

\[
F(z) = \lim_{r \to \infty} \frac{| \omega_{2r+1}(z), \omega'_{2r+3}(z), \omega''_{2r+5}(z) | + 2(z^{-1} \omega_{2r+3}(z) - \omega'_{2r+3}(z)) \Pi b_s}{| \omega_{2r+1}(z), \omega'_{2r+3}(z), \omega''_{2r+5}(z) |}.
\]

Calling the \( r \)th convergents of (49) and (51) \( g_r(z, p) \) and \( \hat{g}_r(z, p) \) respectively, we may consider the sequences \{max \( g_r(z, p) \)\} and \{\( \min \hat{g}(z, p) \)\} as approximations to \( F(z) \), assuming that stationary values exist. That such values do exist is seen from the following asymptotic expansions:

\[
g_r(z, p) = \frac{x_{2r}(z)}{\omega_{2r}(z)} + o(p^{-1}) \text{ as } |p| \to \infty
\]

\[
\hat{g}_r(z, p) = \frac{x_{2r+1}(z)}{\omega_{2r+1}(z)} + \frac{2z}{p} + o(p^{-1}) \text{ as } |p| \to \infty.
\]

It is evident from (54) that for \( p \) large and negative \( g_r(z, p) \) is a closer approximation to \( F(z) \) than \( x_{2r}(z)/\omega_{2r}(z) \); similarly for \( \hat{g}_r(z, p) \). We shall return to a consideration of these approximations in a later part.

I am indebted to a referee for several useful comments on Parts 3 and 4.

Department of Mathematics,
College of Technology,
Manchester, 1.
A Determinantal Expansion for a Class of Definite Integral:
Part 5.

Extracted from the
*Edinburgh Mathematical Proceedings*
A Determinantal Expansion for a Class of Definite Integral

Part 5. Recurrence Relations

By L. R. Shenton

(Received 17th December, 1954).

1. We develop here the recurrence relations for the generalised C.F.'s introduced in Part 3 (Shenton 1 1956). In the main the discussion will be limited to second order C.F.'s, but results for higher orders will be given when these are not complicated.

We shall give three forms of recurrence relation, one involving recurrent determinants, and another corresponding to the even and odd parts of a Stieltjes C.F. In addition we shall show how to write down directly the recurrence relations for a second order C.F. being given the first order C.F. Several numerical examples are given in illustration.

2.0. We consider the C.F. "corresponding" to a determined Stieltjes moment problem, and write

\[ F(z) = \int dx \frac{d\mu(x)}{x + z} = \frac{b_1}{z} + \frac{b_2}{1 + z} + \frac{b_3}{1 + z + 1} + \ldots, \quad z > 0, \quad \ldots \quad (1) \]

and for the "contracted" form

\[ F(z) = \frac{a_0}{z} + \frac{a_1}{z + c_1} - \frac{a_2}{z + c_2} - \ldots \quad (2) \]

where

\[ a_s = b_{2s}, \quad b_{2s+1}, \quad s > 0; \quad a_0 = b_1, \]
\[ c_s = b_{2s-1} + b_{2s}, \quad s > 1; \quad b^*_s = b_s, \quad s > 1, \quad b^*_1 = 0. \]

---

1 We shall refer to the previous four papers on this subject as \( S_1, S_2, S_3, \) and \( S_4. \)

2 Stieltjes preferred to write the "corresponding" C.F. in the form

\[ F(z) = \frac{1}{a_1 z^2 + a_2 + a_3 z^3 + a_4 + \ldots} \]

in which case the Stieltjes moment problem is determined if \( \sum a_s \) diverges, the \( a \)'s being positive.
The \( s^\text{th} \) convergent of (1) will be written \( \chi_s(z)/w_s(z) \), and that of the even part \( \chi_{2s}(z)/w_{2s}(z) \), where \( \chi_0(z) = 0, \chi_1(z) = b_1, w_0(z) = 1, w_1(z) = z \). In the notation of \$3\$ the expression (2) becomes

\[
F(z) = \lim_{s \to \infty} \frac{a_0 K_{s-1} (\beta_1, \alpha_1)}{K_s (\beta_0, \alpha_0)}
\]  

(3)

where \( \beta_s = z + c_{s+1}, \alpha_s = \sqrt{\sqrt{a_{s-1}}} \) and \( K_s (\beta_0, \alpha_0) \) is a determinant of order \( s \) with elements \( \beta_0, \beta_1, \ldots \) along the diagonal through \( (1, 1) \) and elements \( \alpha_0, \alpha_1, \ldots \) along the diagonals through \( (2, 1) \) and \( (1, 2) \).

The second order C.F. can be written

\[
F(z_1, z_2) = \int_0^z \frac{d\psi(x)}{(x+z_1)(x+z_2)} = \lim_{s \to \infty} \frac{a_0 K_{s-1} (\gamma_1, \beta_1, \alpha_1)}{K_s (\gamma_0, \beta_0, \alpha_0)}
\]

(4a)

where

\[
a_s = \sqrt{(a_{s+1} a_{s+2})}, \\
\beta_s = (p + c_{s+1} + c_{s+2}) \sqrt{a_{s+1}}, \\
\gamma_s = q + pc_{s+1} + c_{s+2} + a_s + a_{s+1}, \\
a^*_s = a_s, s > 0, \quad a^*_0 = 0,
\]

\[
(x+z_1)(x+z_2) = x^2 + px + q > 0, x > 0.
\]

Similarly the third order C.F. may be expressed as

\[
F(z_1, z_2, z_3) = \int_0^z \frac{d\psi(x)}{(x+z_1)(x+z_2)(x+z_3)} = \lim_{s \to \infty} \frac{a_0 K_{s-1} (\delta_1, \gamma_1, \beta_1, \alpha_1)}{K_s (\delta_0, \gamma_0, \beta_0, \alpha_0)}
\]

(5a)

where

\[
a_s = \sqrt{(a_{s+1} a_{s+2} a_{s+3})}, \\
\beta_s = (p + c_{s+1} + c_{s+2} + c_{s+3}) \sqrt{(a_{s+1} a_{s+2})}, \\
\gamma_s = (q + p (c_{s+1} + c_{s+2}) + c_{s+2}^2 + c_{s+1} + c_{s+2} + a_{s+2} + a_{s+1} + a^*_s) \sqrt{a_{s+1}}, \\
\delta_s = r + gc_{s+1} + pc_{s+2} + c_{s+1}^2 + p (a_{s+1} + a^*_s) + a_{s+1} + a_{s+2} + 2a_{s+1} + a_{s+2} + 2a_{s+1} a_{s+2} + a^*_s c_{s+1} + a^*_s c_{s+2} + a^*_s c_{s+3} + a^*_s c_{s+4},
\]

\[
a^*_s = a_s, s > 0, \quad a^*_0 = 0,
\]

\[
(x+z_1)(x+z_2)(x+z_3) = x^3 + px^2 + qx + r > 0, x > 0.
\]

\[1\] \( K_s (\gamma_0, \beta_0, \alpha_0) \) is a determinant of order \( s \) with elements \( \gamma_0, \gamma_1, \ldots \) along the diagonal through \( (1, 1) \), \( \beta_0, \beta_1 \ldots \) along the diagonals through \( (2, 1) \) and \( (1, 2) \), and \( \alpha_0, \alpha_1, \ldots \) along the diagonals through \( (3, 1) \) and \( (1, 3) \). The determinant \( K_s (\gamma_0, \beta_0, \alpha_0) \) is symmetric with elements in five diagonals only, and may be regarded as a form of generalised continuant. The extension of the notation is obvious.
2.1 To develop the numerators and denominators of (4a) and (5a) in terms of recurrent determinants we require the following lemma.

**Lemma.** If \( K_s(h_1, g_1, f_1) \) is a determinant of order \( s \) with elements \( f_1, f_2, \ldots \) along the diagonal through \((1, 3)\), \( g_1, g_2, \ldots \) along the diagonal through \((1, 2)\), \( h_1, h_2, \ldots \) along the diagonal through \((1, 1)\) and so on, then

\[
K_{s-2}(g_1, h_2, g_2, f_2) K_s(g_0, f_0, h_1, g_1, f_1) = \frac{K_{s-1}(g_1, h_1, g_1, f_1)}{K_{s-1}(h_2, g_2, f_2)} \frac{K_{s-1}(h_1, g_1, f_1)}{K_{s-1}(g_0, h_1, g_1, f_1)} \prod_{\lambda = 0}^{s-2} f_\lambda.
\]

The proof is straightforward. For consider \( K_s(g_0, h_1, g_1, f_1) \).

Delete the first and last rows and columns, and use the remaining array as a pivot.

Applying (6) to the numerator and denominator of (4a) we find

\[
F(z_1, z_2) = \lim_{s \to \infty} a_0 \left| \begin{array}{cc} V_s & W_{s+1} \\ U_s & V_{s+1} \end{array} \right|.
\]

1 A determinant of the form \( K_s(\ldots, f_0, g_0, f_0, \ldots) \) of order \( s \), with elements \( f_0, f_1, \ldots \) along the first superdiagonal, \( g_0, g_1, \ldots \) along the principal diagonal, \( h_1, h_2, \ldots \) along the first subdiagonal, and so on, will be referred to as a **recurrent determinant**, or simply a **recurrent**.

By expanding a **recurrent** of this form by its last row, it will be seen that it follows a fourth order recurrence relation. Similarly a **recurrent** with \( n \) subdiagonals may be shown to follow a recurrence relation of order \( n + 1 \).

where \( U_s = K_s \left( \beta^*_s, a^*_s, \beta_1, a_0 \right) \)
\[
V_s = K_{s-1} \left( \beta_0, a_0, \beta_1, a_1 \right)
\]
\[
W_s = K_{s-2} \left( \beta_1, a_1, \beta_2, a_2 \right)
\]

\( \beta^*_s = \beta_s, a^*_s = a_s, s \geq 0; \quad \beta^*_{-1} = 0, \quad a^*_{-1} = 1, \)

and the recurrences \( U_s, V_s, W_s \) follow the relation
\[
y_s = \beta_s y_{s-1} - \gamma_s - 2\alpha^*_s y_{s-2} + \beta_s - 3\alpha^*_s y_{s-3} y_{s-1} - \beta_s - 3\alpha^*_s y_{s-4} a_s - 4 a_s - 5 y_{s-3}, \quad (8)
\]

with \( U_0 = 1, \quad U_1 = 0, \quad U_s = 0, \quad s < 0, \)
\( V_1 = 1, \quad V_2 = \beta_0, \quad V_s = 0, \quad s < 1, \)
\( W_2 = 1, \quad W_3 = \beta_1, \quad W_s = 0, \quad s < 2, \)

the values of \( \alpha_s, \beta_s, \gamma_s \) being given in (4b).

The \( s \)-th approximant to \( F(z_1, z_2) \) depends upon the six terms \( U_1, U_{s+1}, V_s, V_{s+1}, W_s, W_{s+1} \), each of which follows a recurrence relation of order four. Hence to advance the approximation process one stage, it is necessary to evaluate a value of each of \( U_s, V_s, W_s \), and this will involve twelve calculations. We shall show in a later section that \( | V_s, W_{s+1} | \) and \( | U_s, V_{s+1} | \) (or equivalent expressions) follow recurrence relations of order five, so that there is perhaps an economy to be gained by this method.

A decreasing sequence of upper bounds may be derived from the expression.
\[
F(z_1, z_2) = (q - \frac{1}{2}p^2)^{-1} a_0 - (q - \frac{1}{2}p^2)^{-1} \int_0^\infty \frac{(x + \frac{3}{4}p)^2 \, d\phi(x)}{2x^2 + px + q}, \quad (9)
\]

and it is not difficult to show that the difference between the \( s \)-th approximations that arise from (9) and (4a) is
\[
\frac{(q - \frac{1}{2}p^2)^{-1} \prod_{\lambda=0}^s a_\lambda}{K_s(\gamma_0, \beta_0, a_0)}, \quad (10)
\]

it being assumed that \( q - \frac{1}{2}p^2 > 0. \)

2.2 For third order C.F’s we require the following extension of the lemma:
A Determinantal Expansion for a Class of Definite Integral 171

\[ Q_{x}^{(0)} K_x \left( i_1, h_1, \varphi_1, f_1 \right) = \begin{vmatrix} T_{s+1}^{(1)} & T_{s+1}^{(0)} \\ T_{s-1}^{(1)} & T_{s-1}^{(0)} \end{vmatrix}, \]  

(11)

where \( Q_{x}^{(1)} \equiv K_x \left( g_1, h_1, \varphi_1, f_1 \right), \)

\[ T_{x}^{(1)} \equiv K_x \left( h_1, g_1, h_1, \varphi_1, f_1 \right). \]

The proof depends on pivotal condensation methods and closely follows that of the lemma. Each determinant in the compound determinant in (11) is replaced by a compound determinant, using identities similar to (6). For example

\[ \prod_{\lambda=0}^{s-1} f_\lambda \cdot T_{s}^{(1)} = \begin{vmatrix} Q_{x}^{(0)} & Q_{x}^{(1)} \\ Q_{x+1}^{(0)} & Q_{x+1}^{(1)} \end{vmatrix}. \]

We find in this way the relation

\[ \prod_{\lambda=-1}^{s-2} f_\lambda \cdot \prod_{\lambda=0}^{s-1} f_\lambda \cdot K_x \left( i_1, h_1, \varphi_1, f_1 \right) = \begin{vmatrix} Q_{s-2}^{(1)} & Q_{s-1}^{(0)} & Q_{s-1}^{(-1)} \\ Q_{s-1}^{(1)} & Q_{s}^{(0)} & Q_{s}^{(-1)} \\ Q_{s}^{(1)} & Q_{s+1}^{(0)} & Q_{s+1}^{(-1)} \end{vmatrix}, \]  

(12)

it being assumed that \( f_\lambda \neq 0, \lambda = -1, 0, 1, \ldots, s - 1. \)

Returning now to (5a) we derive from (12) an expansion for a third order C.F. in terms of recurrent determinants, namely

\[ F(z_1, z_2, z_3) = \int_0^\infty \frac{d\psi(x)}{(x+z_1)(x+z_2)(x+z_3)} \]

\[ = 1 \text{ i.e. s. } a_0 \begin{vmatrix} U_{s-1}, V_s, W_{s-1} \\ X_{s-1}, V_s, W_{s-1} \end{vmatrix} \]  

(13)

where

(i) \( U_s = K_{s-2} \left( \beta_1, \alpha_1, \gamma_2, \delta_3, \gamma_2, \beta_3, \alpha_3 \right), \)

\( V_s = K_{s-1} \left( \beta_0, \alpha_0, \gamma_1, \delta_2, \gamma_2, \beta_2, \alpha_2 \right), \)

\( W_s = K_s \left( \beta^{*}_{-1}, \alpha^{*}_{-1}, \gamma_0, \delta_1, \gamma_1, \beta_1, \alpha_1 \right), \)

\( X_s = K_{s+1} \left( \beta^{*}_{-2}, \alpha^{*}_{-2}, \gamma_0, \delta_0, \gamma_0, \beta_0, \alpha_0 \right); \)
\[ \begin{align*}
(ii) \quad & a_s^* = a_s, \quad \beta_s^* = \beta_s, \quad \gamma_s^* = \gamma_s, \quad s \geq 0; \\
& a_{s-1}^* = a_{s-2}^* = 1, \quad \beta_{s-1}^* = \beta_{s-2}^* = 0, \quad \gamma_{s-1}^* = 0, \\
& a_s, \beta_s, \gamma_s, \delta_s \text{ being given in (5b)};
\end{align*} \]

(iii) the recurrants \( U_s, V_s, W_s, X_s \) follow
\[
\begin{align*}
y_s &= \beta_{s-2}^* y_{s-1} - \gamma_{s-2}^* a_{s-3} y_{s-2} + \delta_{s-2} a_{s-3} a_{s-4} y_{s-3} - a_{s-3} a_{s-4} a_{s-5} y_{s-4} \\
&\quad + \beta_{s-4} a_{s-3} a_{s-5} a_{s-6} y_{s-5} - a_{s-3} a_{s-4}^2 a_{s-6} a_{s-7} y_{s-6},
\end{align*}
\]
with \( U_2 = 1, \quad U_3 = \beta_1, \quad U_s = 0, \quad s < 2; \)
\( V_1 = 1, \quad V_2 = \beta_0, \quad V_s = 0, \quad s < 1; \)
\( W_0 = 1, \quad W_1 = 0, \quad W_s = 0, \quad s < 0; \)
\( X_{-1} = 1, \quad X_0 = 0, \quad X_s = 0, \quad s < -1. \)

It will be seen that each of the elements \( U_s, V_s, W_s, X_s \), occurring in the \( s \)th convergent of a third order C.F. follows a sixth order recurrence relation, so that in setting up approximations to \( F(z_1, z_2, z_3) \) we have to perform in general twenty-four calculations to obtain each new approximation. Similarly for a C.F. of order \( n \) associated with the function \( F(z_1, z_2, \ldots, z_n) \), each approximation consists of the ratio of two \( n \)th order determinants, an element of either determinant consisting of a recurrent which satisfies a recurrence relation of order \( 2n \). In general then each new approximation to \( F(z_1, z_2, \ldots, z_n) \) will involve \( 2n(n+1) \) calculations, followed by the evaluation of two \( n \)th order determinants.\(^1\) We shall consider these more general C.F.'s and the associated recurrence relations in a forthcoming paper.

3. A Fifth Order Recurrence Relation.

3.1 We now establish a recurrence relation for the symmetric determinant \( K_s(h_1, g_1, f_1) \). Expand \( K_s \) by its last row and column.

---

\(^1\) A referee has indicated to me that the recurrence relation followed by these two \( n \)th order determinants will be of order \( \binom{2n}{n} \) in general, or a little less owing to the symmetry involved. Thus for a third order C.F. the numerator and denominator of the \( s \)th convergent will very likely satisfy a recurrence relation of order nineteen. Even if this recurrence could be found it might well be too complicated to be of much value, and the method of compound recurrent determinants seems to have a distinct advantage for C.F.'s of order three or more.
A Determinantal Expansion for a Class of Definite Integral 173

Then

\[
K_s = h_s K_{s-1} - g_{s-1}^* K_{s-2} + 2g_{s-1} f_{s-2} K_{s-2}^* - f_{s-2}^2 h_{s-1} K_{s-3}^* + f_{s-2}^2 f_{s-3}^2 K_{s-4},
\]

\[s = 4, 5, \ldots,\]  

\[(14)\]

where

\[
K_s^* = \begin{pmatrix}
K_{s-1} & f_{s-2}^* \\
& g_{s-1}
\end{pmatrix},
\]

and \(K_{s-1}\) is the matrix consisting of the elements of \(K_{s-1} (h_1, g_1, f_1)\).

For example,

\[
K_2^* = \begin{pmatrix}
h_1 & g_1 \\
f_1 & g_2
\end{pmatrix}, \quad K_3^* = \begin{pmatrix}
h_1 & g_1 & f_1 \\
g_1 & h_2 & g_2
\end{pmatrix}.
\]

But expanding \(K_s^*\) by its last row, we have

\[
K_s^* = g_s K_{s-1} - f_{s-1} K_{s-1}^*, \quad s = 3, 4, \ldots.
\]

\[(15)\]

Eliminating \(K_s^*\) from (14) and (15), we find

\[
g_{s-2} K_s = (h_s g_{s-2} - f_{s-2} g_{s-1}) K_{s-1}
\]

\[- (g_{s-1} g_{s-2} - h_{s-1} f_{s-2}) (g_{s-1} K_{s-2} - g_{s-2} f_{s-2} K_{s-3})
\]

\[- f_{s-2}^2 (h_{s-2} g_{s-1} - f_{s-1} g_{s-2}) K_{s-4} + f_{s-2} f_{s-3}^2 f_{s-4}^2 g_{s-1} K_{s-5},
\]

\[s = 3, 4, \ldots\], \quad K_{-2} = K_{-1} = 0, K_0 = 1.

\[(16)\]

The recurrence relation (16) satisfied by \(K_s (h_1, g_1, f_1)\) is of order five.

By a slight modification of the method employed here it may be shown that the recurrence relation for the asymmetric determinant

\[
K_s (h_1, g_1^1, f_1^1)\]

is of order six, but we do not require it in the present context.
There are three interesting special cases:

(a) \( f_j = 0 \), when (16) reduces to
\[
K_s = h_s K_{s-1} - g^2_{s-1} K_{s-2},
\]
as we should expect since \( K_s(h_1, g_1, 0) \) is now a "continuant" type of determinant.

(b) \( g_j = 0 \), when (16) becomes
\[
K_s = h_s K_{s-1} - h_{s-1}^2 f^2_{s-2} K_{s-3} + f^2_{s-2} f^2_{s-3} K_{s-4},
\]
which is the recurrence relation for the product of two "continuants," and indeed
\[
K_{2s}(h_1, 0, f_1) = K_s(h^*, f^*) K_s(h^{**}, f^{**}),
\]
\[
K_{2s+1}(h_1, 0, f_1) = K_{s-1}(h^*, f^*) K_s(h^{**}, f^{**}),
\]
where
\[
h^* = h_{2s-1}, \quad f^* = f_{2s-1},
\]
\[
h^{**} = h_{2s}, \quad f^{**} = f_{2s}.
\]

We shall treat the third example of reducibility in § 4, for it turns out to have several applications.

3.2 Now applying (16) to (4a) we may write the second order C.F. as
\[
F(z_1, z_2) = \lim_{s \to \infty} a_0 \frac{u_s}{v_s},
\]
where
\[
K_s(\gamma_0, \beta_0, a_0) = v_s \prod_{s=1}^2 a_s, \quad K_{s-1}(\gamma_1, \beta_1, a_1) = u_s \prod_{s=1}^2 a_s,
\]
and \( u_s, v_s \) follow the recurrence
\[
a_s \beta^1_{s-2} y_s = (\gamma_{s-1} \beta^1_{s-3} - a_{s-1} \beta^1_{s-2}) y_{s-1} - (\beta^1_{s-2} \beta^1_{s-3} - \gamma_{s-2}) (\beta^1_{s-3} y_{s-2} - \beta^1_{s-1} y_{s-3})
\]
\[
- (\gamma_{s-3} \beta^1_{s-2} - a_{s-2}) y_{s-4} + a_s \beta^1_{s-3} y_{s-5},
\]
where \( \beta^1_s \sqrt{a_{s-1}} = \beta_s \), and \( a_s, \beta_s, \gamma_s, a_s \) are given in (4b), the initial values being
\[
u_0 = 0, \quad a_1 v_1 = 1, \quad a_2 v_2 = \gamma_1, \quad u_s = 0, \quad s < 0;
\]
\[
v_0 = 1, \quad a_1 v_1 = \gamma_0, \quad a_2 v_2 = \gamma_0 \gamma_1 - \beta^2_0, \quad v_s = 0, \quad s < 0.
\]

---

1 The result has been noted by T. Muir, Proc. Edinburgh Math. Soc. ii (1884), 16-18.
3.3 As an illustration consider the "J" fraction expansion, convergent for \( z > 0 \).

\[
\int_0^1 \frac{dx}{x + z} = \frac{1}{z + \frac{1}{2} - z + \frac{1}{2} - z + \frac{1}{2} - \cdots}
\]

(19)

where \( a_s = s^2 / (16s^2 - 4) \), from which we deduce the second order C.F. expansion

\[
\int_0^1 \frac{dx}{x^2 + z^2} = z^{-1} \arctan z^{-1} = \text{l.i.s.} \frac{u_s}{v_s}, z = 0,
\]

(20)

where \( u_s \) and \( v_s \) follow

\[
a_s y_s = \left( z^2 + a_s + \frac{1}{4} \right) y_{s-1} + \left( z^2 + a_{s-1} + a_{s-2} - \frac{3}{4} \right) (y_{s-2} - y_{s-3})
\]

\[
- \left( z^2 + a_{s-3} + \frac{3}{4} \right) y_{s-4} + a_{s-3} y_{s-5}, \quad s = 3, 4, \ldots
\]

(21)

and the initial values are

\[
u_0 = 0, \quad u_1 = 12, \quad u_2 = 3 (60z^2 + 24); \]

\[
v_0 = 1, \quad v_1 = 4 (3z^2 + 1), \quad v_2 = 3 (60z^2 + 44z^4 + 3).
\]

(21a)

For example, using (21) it will be found that

\[
\begin{align*}
u_3 &= 16 (525z^4 + 410z^2 + 45), \\
v_3 &= 16 (525z^4 + 585z^2 + 135z^2 + 3), \\
u_4 &= 132,300z^8 + 153,300z^4 + 41,300z^2 + 1,800, \\
u_4 &= 132,300z^8 + 197,400z^6 + 80,640z^4 + 8,100z^2 + 75.
\end{align*}
\]

The expansion indicated in (20)-(21) is not the same (apart from the approximations for \( s = 0, 1 \)) as the even part of the hypergeometric C.F.

\[
z^{-1} \arctan z^{-1} = \frac{1}{z^2 + 1 + z^2 + 1 + z^2 + \cdots},
\]

where \( b_s = s^2 / (4s^2 - 1) \).

4. A Reducible Case of the Fifth Order Recurrence Relation.

4.1 The recurrence relation (16) reduces to a fourth order one when the C.F. in (2) takes on the special form

\[
F(z) = \frac{a_0}{z} - \frac{a_1}{z} - \frac{a_2}{z} - \frac{a_3}{z} - \cdots
\]

(22)

Proceeding formally at first and writing \( K_{s-1} (\gamma_1, \beta_1, a_1) = T^*_s \), \( K_s (\gamma_0, \beta_0, a_0) = T_s \), we find that the recurrence for \( T^*_s \) and \( T_s \) becomes
\[ y_s = (q + a_s) y_{s-1} + a_{s-1} (q - p^2 + a_{s-1} + a_{s-2}) y_{s-2} \]
\[ - a_{s-1} a_{s-2} (q - p^2 + a_{s-1} + a_{s-2}) y_{s-3} - a_{s-1} a_{s-2} a_{s-3} (q + a_{s-3}) y_{s-4} \]
\[ + a_{s-1} a_{s-2} a_{s-3}^2 y_{s-5}, \quad s = 3, 4, \ldots, \]

where \( q = z_1 z_2, \ p = z_1 + z_2, \ T_s = 0, \ s < 0, \ T_s^* = 0, \ s \lesssim 0. \)

Now (23) may be written
\[ \Phi_s (y) = \alpha_{s-1} \Phi_{s-1} (y) = 0, \]
where \( \Phi_s (y) \equiv y_s - (q + a_s - a_{s-1}) y_{s-1} - a_{s-1} (2a_{s-1} - p^2 + 2q) y_{s-2} \]
\[ - a_{s-1} a_{s-2} (q + a_{s-2} - a_{s-1}) y_{s-3} + a_{s-1} a_{s-2} a_{s-3} y_{s-4}. \]

But it is easily verified that \( \Phi_2 (T) = 0. \) Hence from (25)
\[ T_s = (q + a_s - a_{s-1}) T_{s-1} + a_{s-1} (2a_{s-1} - p^2 + 2q) T_{s-2} \]
\[ + a_{s-1} a_{s-2} (q + a_{s-2} - a_{s-1}) T_{s-3} - a_{s-1} a_{s-2} a_{s-3} T_{s-4}, \]
with \( T_0 = 1, \ T_1 = q + a_1, \ T_s = 0, \ s < 0. \)

Similarly it will be found that
\[ \Psi_s (T^*) = \alpha_{s-1} \Psi_{s-1} (T^*) = 0, \]
where \( \Psi_s (T^*) \equiv \Phi_s (T^*) - 2 \prod_{\lambda=0}^{s-1} \alpha_{\lambda}. \)

But since \( \Psi_2 (T^*) = 0, \) it follows that the recurrence for \( T_s^* \) is
\[ T_s^* = (q + a_s - a_{s-1}) T_{s-1}^* + a_{s-1} (2a_{s-1} - p^2 + 2q) T_{s-2}^* \]
\[ + a_{s-1} a_{s-2} (q + a_{s-2} - a_{s-1}) T_{s-3}^* - a_{s-1} a_{s-2} a_{s-3} T_{s-4}^* + 2 \prod_{\lambda=0}^{s-1} \alpha_{\lambda}, \]
with \( T_0^* = a_0, \ T_s^* = 0, \ s \lesssim 0. \)

4.2 Returning to §2 we observe that in (1) \( b_s > 0, \ s = 1, 2, \ldots. \)
Hence \( c_s \) cannot be zero for a Stieltjes C.F. We may ask the question then as to how the value of \( a_s \) in (22) must be restricted so that for certain values of \( z_1 \) and \( z_2 \) it will be true to assert that
\[ \frac{F(z_1) - F(z_2)}{z_2 - z_1} = \lim_{s \to \infty} \frac{T_s^*}{T_s}. \]
A Determinantal Expansion for a Class of Definite Integral 177

We can give at this stage only a partial answer to this question. First we refer to the theory of the Hamburger moment problem. Let the expansion in descending powers of \( z \) of \( F(z) \) be

\[
F(z) \sim \frac{\mu_0}{z} + \frac{\mu_2}{z^3} + \frac{\mu_4}{z^5} + \ldots
\]

and assume that

\[
\begin{cases}
  a_s \geq 0, & s = 0, 1, 2, \ldots, \\
  \sum_{n=1}^{\infty} \mu_{2n}^{-1/2n} = \infty.
\end{cases}
\]

Then there exists a unique bounded non-decreasing function \( \psi(x) \) in the interval \((-\infty, \infty)\) such that

\[
\int_{-\infty}^{\infty} x^{2n} \psi(x) \, dx = \mu_{2n}, \quad \int_{-\infty}^{\infty} x^{2n+1} \psi(x) \, dx = 0.
\]

All we have to do now is to justify the Parseval expansion for

\[
\int_{-\infty}^{\infty} \{f(x)\}^2 \psi(x) \, dx \quad \text{where} \quad f(x) = \{(x + z_1) (x + z_2)\}^{-1},
\]

\[
\bar{\psi}(x) = \int_{-\infty}^{x} (z_1 + t) (z_2 + t) \, d\psi(t),
\]

the argument being similar to that used in §2 of S4. It turns out then, using the theorem of M. Riesz (Shohat and Tamarkin, loc. cit., p. 62) regarding the solution of a determined Hamburger moment problem, that under the conditions in (30) we have

\[
F(z_1, z_2) = \int_{-\infty}^{\infty} \frac{\psi(x)}{(x + z_1) (x + z_2)} \, dx = 1, \ \text{i.e.} \ \sum_{s=1}^{\infty} \frac{\psi^s}{(x + z_1) (x + z_2)}, \quad \text{with} \quad z_1 = x_1 + iy_1, \quad z_2 = x_2 - iy_1,
\]

provided in addition \((x + z_1) (x + z_2)>0\) for all real \( x \). Again using (34) and (35) of S4 we can set up a decreasing sequence of upper bounds, the difference between corresponding \( s \)th approximants being

\[
\Pi \frac{a_s}{y_1^2 T_{1s}}, \quad \text{where} \quad y_1 = \text{Im} z.
\]

Secondly it may be possible to justify (31) if we are given a definite integral for the C.F. in (22) of the form \( \int_{a}^{b} \frac{\psi(x)}{x + z} \, dx \) and can justify the application of Parseval's theorem as in §B of S2. It would

---

be of some interest to know what are the weakest restrictions on the
\(a's\) in (22) to justify a statement similar to (31).

4.2. We now give several examples in illustration.
Example 1. Let

\[
F(z) = \tan \left( \frac{1}{2}z^{-1} \right),
\]

so that

\[
\mu_{2n} = 2B_{n+1} \frac{(2^{2n+1} - 1)}{(2n + 2)!},
\]

where \(B_n\) is a Bernoulli number and \(B_1 = 1/6, \ B_2 = 1/30\) etc. Using
Lambert’s C.F.\(^1\) for \(F(z)\) we have \(a_0 = \frac{1}{3}, \ a_s = (16s^2 - 4)^{-1}, s = 1, 2, \ldots\)
Moreover it is easily verified that \(\mu_{2n} \sim 1/\pi\), so that (30) is satisfied.
Hence with the appropriate value of \(a_3\) in (26) and (28) we have

\[
\sinh y_1 \equiv y_1 \left( \cosh y_1 + \cos x_1 \right) = 1, \ \text{i.e.} \ \frac{T^*}{T_s}, \ y_1 \neq 0,
\]

where \(x_1 = x_1, y_1 = (2^2 + 2) - y_1\). In particular if \(x_1 = 3, y_1 = 4\)
then \(\lambda T^* = 398, 499, 385, 800, \ \lambda T_4 = 19, 896, 118, 681, 110\) where \(\lambda^{-1} = a_4, a_3, a_2, a_1, a_0,\) giving the lower bound 0.020, 029, 001, 243, 260, and
using the corresponding upper bound it turns out that the error in
this cannot exceed \(1.6 \times 10^{-15}\).

In a similar way, taking \(F(z)\) to be \(2z - \cot \left( \frac{1}{2}z^{-1} \right)\), so that \(a_0 = 1/6, \ a_s = \left(4s^2 + 2\right) - 1, s = 1, 2, \ldots,\) it may be shown that

\[
\sinh y_1 \equiv y_1 \left( \cosh y_1 - \cos x_1 \right) = 2, \ \text{i.e.} \ \frac{T^*}{T_s}, \ y_1 \neq 0.
\]

In each case the recurrence relation for \(T_z\) is (26) and that for \(T_s^*\) is
(28) with the appropriate value of \(a_s, s = 0, 1, 2, \ldots\).

Example 2. It has been indicated by Stieltjes\(^2\) that

\[
\frac{1}{2} \int_{-\infty}^{\infty} \frac{\text{sech}(\frac{1}{2} \pi x) dx}{x + z} = \frac{1}{z} \left( 1^2 - 2^2 + 3^2 - \ldots \right), \ \text{Im} \ (z) \neq 0, \ \text{(34a)}
\]

\[
\frac{1}{2} \int_{-\infty}^{\infty} \frac{x \cosech(\frac{1}{2} \pi x) dx}{x + z} = \frac{1}{z} \left( 1.2 - 2.3 + 3.4 - \ldots \right), \ \text{Im} \ (z) \neq 0. \ \text{(34b)}
\]

It may be shown that the Hamburger moment problem is determined
in each case\(^3\), and it follows that for \(z_1 = x_1 + iy_1, \ z_2 = x_1 - iy_1, y_1 \neq 0\)

\[
\mu_n = \int_{-\infty}^{\infty} x^n e^{-by} dx, \ y = |x|^a, \ a > 1, \ b > 0, \ n = 0, 1, 2, \ldots \text{ is determined.}
\]

\(^1\) See for example Perron, O., *Die Lehre von den Kettenbrüchen*, p. 354, (Berlin, 1913).

\(^2\) Correspondance d'Hermite et de Stieltjes, p. 360 (Paris, 1905).

\(^3\) Compare also Wall, H. S., *Continued Fractions*, p. 366, Example 2 (New York, 1948). We may also recall that the Hamburger moment problem
the second order C.F.'s indicated in (31) converge to the corresponding
value of \( F(z_1, z_2) \).

**Example 3.** Let \( F(z) = \int_0^\infty \frac{g(x)dx}{x + z} \),
where \( g(x) = e^{-ix^2}/\sqrt{(2\pi)} \), with \( \text{Im} (z) > 0 \), so that
\[
F(z) = \frac{1}{z} - \frac{1}{z - z} - \frac{2}{z - z} - \ldots.
\] (35)

From P. 3, §B of 52 we can conclude that the second order C.F. converges for \( y_1 = 0 \), and indeed
\[
\int_0^\infty \frac{g(x)dx}{(x + x_1)^2 + y_1^2} = 1, \ \text{i.e.} \ \frac{T_s^*}{s-\infty T_s^*},
\]
where
\[
T_s^* = (x_1^2 + y_1^2 + 1) T_{s-1}^* + 2 (s - 1) (s - 1 + y_1^2 - x_1^2) T_{s-2}^*
+ (s - 1) (s - 2) (s - 2) w_{s+2}^* - 1 (s - 1) (s - 2) (s - 3) T_{s-4}^* + 2(s-1)!, \quad s = 2, 3, \ldots.
\] (36)

and the recurrence for \( T_s \) is exactly the same except that the factorial term is omitted. The initial values are
\[
T_s^* = 0, \quad s < 1, \quad T_1^* = 1
\]
\[
T_s = 0, \quad s < 0, \quad T_0 = 1, \quad T_1 = x_1^2 + y_1^2 + 1.
\]

A numerical example will be found in §C 4 of 52.

5. **A Recurrence Relation with Even and Odd Parts.**

5.1 For the C.F. given in (1) there are recurrence relations corresponding to the **even** and **odd** parts, namely
\[
w_{2s}(z) = w_{2s-1}(z) + b_{2s} w_{2s-2}(z),
\]
\[
w_{2s+1}(z) = z w_{2s}(z) + b_{2s+1} w_{2s-1}(z).
\] (37)
The question naturally arises as to whether there is a similar structure for higher order C.F.'s. We give here the result for a second order C.F. Starting with the form of the denominator of (4a) given in (30) of 54, we have, assuming for the moment \( z_1 + z_2 \),
\[
K_s(y_0, \beta_0, a_0) = \mid w_{2s}(z_1), w_{2s+2}(z_2) \mid \div (z_2 - z_1),
\] (38)
say. Using (37) we readily find that
\[ k_{2s} = y_{2s} + b_{2s+1} b_{2s} k_{2s-2}, \]  
(39)
where
\[ y_s = w_s(z_1)w_s(z_2). \]

Similarly if
\[ k_{2s-1} = \begin{vmatrix} w_{2s-1}(z_1), w_{2s+1}(z_2) \end{vmatrix} \]  
(40)
then
\[ k_{2s-1} = y_{2s-1} + b_{2s} b_{2s-1} k_{2s-3}. \]  
(41)
But
\[ y_{2s} = y_{2s-1} + b_{2s}^2 y_{2s-2} + b_{2s} \Theta_{2s-1}, \]  
(42)
where
\[ \Theta_{2s-1} = w_{2s-1}(z_1) w_{2s-2}(z_2) + w_{2s-1}(z_2) w_{2s-2}(z_1). \]  
Hence
\[ \Theta_{2s-1} = (z_1 + z_2) y_{2s-2} + b_{2s-1} [w_{2s-2}(z_1) w_{2s-3}(z_2) + w_{2s-2}(z_2) w_{2s-3}(z_1)]. \]

Hence
\[ \Theta_{2s-1} = (z_1 + z_2) y_{2s-2} + b_{2s-1} [w_{2s-2}(z_1) w_{2s-3}(z_2) + w_{2s-2}(z_2) w_{2s-3}(z_1)]. \]
and so from (42) we have
\[ y_{2s} - y_{2s-1} - b_{2s} y_{2s-2} = b_{2s}(z_1 + z_2) y_{2s-2} + 2b_{2s-1} y_{2s-3} + b_{2s-1} b_{2s-2} \Theta_{2s-3}, \]
from which, using (39) and (41), we deduce
\[ k_{2s} = k_{2s-1} + b_{2s}(z_1 + z_2 + b_{2s+1} + b_{2s} + b_{2s-1}) k_{2s-2} \]
\[ - b_{2s} b_{2s-1} b_{2s-2}(z_1 + z_2 + b_{2s} + b_{2s-1} + b_{2s-2}) k_{2s-4} \]
\[ - b_{2s} b_{2s-1} b_{2s-2} b_{2s-3} k_{2s-5} + b_{2s} b_{2s-1} b_{2s-2} b_{2s-3} b_{2s-4} k_{2s-6}. \]  
(43)

5.2 Similarly from (37) it follows that
\[ y_{2s+1} = z_1 z_2 y_{2s} + b_{2s+1} y_{2s-1} + b_{2s+1} \Phi_{2s}, \]  
(44)
where
\[ \Phi_{2s} = z_1 w_{2s}(z_1)w_{2s-1}(z_2) + z_2 w_{2s}(z_2) w_{2s-1}(z_1) \]
\[ = (z_1 + z_2) y_{2s-1} + b_{2s} [z_1 w_{2s-2}(z_1) w_{2s-1}(z_2) + z_2 w_{2s-3}(z_2) w_{2s-1}(z_1)]. \]

or
\[ \Phi_{2s} = (z_1 + z_2) y_{2s-1} + b_{2s} y_{2s-1} - 2z_1 z_2 b_{2s-1} y_{2s-2} + b_{2s} b_{2s-1} \Phi_{2s-2}. \]  
(45)

Thus
\[ y_{2s+1} = z_1 z_2 y_{2s} + b_{2s+1} y_{2s-1} + b_{2s+1}(z_1 + z_2) y_{2s-1} + 2z_1 z_2 b_{2s+1} b_{2s} y_{2s-2} \]
\[ + b_{2s+1} b_{2s}(y_{2s-1} - z_1 z_2 y_{2s-2} - b_{2s-1} y_{2s-3}), \]  
(46)
and so by (39) and (41) it appears that
\[ k_{2s+1} = z_1 z_2 k_{2s} + b_{2s+1}(z_1 + z_2 + b_{2s+2} + b_{2s+1} + b_{2s}) k_{2s-1} \]
\[ - b_{2s+1} b_{2s} b_{2s-1}(z_1 + z_2 + b_{2s+1} + b_{2s} + b_{2s-1}) k_{2s-3} \]
\[ - b_{2s+1} b_{2s} b_{2s-1} b_{2s-2} z_1 k_{2s-4} + b_{2s+1} b_{2s} b_{2s-1} b_{2s-2} b_{2s-3} k_{2s-5}. \]  
(47)

In the derivation of the recurrence formulæ (43) and (47) it has been assumed that s is large enough to avoid initial value idiosyncrasies. Allowance being made for these we finally have the theorem:
A Determinantal Expansion for a Class of Definite Integral

If the Stieltjes moment problem is determined and

\[ F(z) = \int_0^\infty \frac{d\phi(x)}{x+z} = z + \frac{1}{1+z} + \frac{1}{1+z} + \ldots \]  

then

\[ \lim_{s \to \infty} \frac{k^*_s}{k^*_{2s}} = \frac{F(z_2) - F(z_1)}{z_1 - z_2}, \quad (48) \]

where

(i) \( k^*_s \) and \( k_s \) follow, for \( s = 2, 3, \ldots \),

\[ w_{2s-1} = z_1 z_2 w_{2s-2} + a_{2s-2} w_{2s-3} - z_1 z_2 \gamma_{2s-1} w_{2s-6} + \delta_{2s-1} w_{2s-7} \]

\[ w_{2s} = z_1 z_2 + a_{2s} w_{2s-2} - \beta_{2s} w_{2s-4} - \gamma_{2s} w_{2s-5} + \delta_{2s} w_{2s-6}, \]

(ii) \( k^*_0 = 0, k^*_1 = k^*_2 = b_1, \quad k_s^* = 0, s < 0, \)

\[ k_0 = 1, \quad k_1 = z_1 z_2, \quad k_2 = z_1 z_2 + b_2 (z_1 + z_2 + b_3 + b_2), \quad k_s = 0, s < 0, \]

(iii) \( a_s = b_{s+1} (z_1 + z_2 + b_{s+1} + b_s + b_{s+1}), \quad \beta_s = b_s (z_1 - a_{s-1}), \)

\[ \gamma_s = b_s (z_2 b_{s+2} + b_{s+2} - b_{s+1} b_s - b_{s+1} b_{s+2}), \quad \delta_s = b_s (z_1 - b_{s+1} b_{s+2} - b_{s+1} b_{s+2}), \]

(iv) \( (x + z_1) (x + z_2) > 0 \) for \( x \geq 0 \).

It may be established that the 'odd' part \( k^*_s / k_{2s-1} \) arises from the second order C.F. associated with the integral in the expression

\[ z_1 z_2 \left( \frac{F(z_2) - F(z_1)}{z_1 - z_2} \right) = \int_0^\infty \frac{x(x + z_1 + z_2) d\phi(x)}{(x + z_1)(x + z_2)} \]

where \( x d\phi(x) \) is taken as the weight function. The 'odd' part of the sequence, unlike that of a Stieltjes C.F., does not in general provide a set of decreasing upper bounds, but there is a remarkable property which we now consider.

5.2 We shall prove the following identities:

\[ t_{2s+1} - t_2 = \frac{B_{2s+1} U_{2s+1}^{(1)} U_{2s+2}^{(1)}}{k_{2s+1} k_{2s+2}}, \quad s = 0, 1, 2, \ldots \]  

\[ (49a) \]

\[ t_{2s} - t_{2s-1} = -\frac{B_{2s} U_{2s}^{(1)} U_{2s+1}^{(1)}}{k_{2s-1} k_{2s}}, \quad s = 1, 2, \ldots \]  

\[ (49b) \]

where

\[ U_s^{(1)} = \frac{w_s(z_2) - w_s(z_1)}{z_2 - z_1}, \quad B_s = \prod_{t=1}^{s} b_t, \quad t_s = k^*_s / k_s. \]

\[ \text{It has been assumed throughout that } z_1 = z_2, \text{ but it is easily shown that the theorem still holds if } z_1 = z_2 \text{ and } (x + z_1) > 0 \text{ for } x \geq 0. \]
Introducing the expressions appearing in (30) and (47) of S4 for \( t_s \), we have \((z_1 + z_2)\)

\[
b_{2s+2}(z_1 - z_2)k_2k_{2s+1}(t_{2s+1} - t_2) = \left| \begin{array}{c} w_{2s+1}(z_1), \ w_{2s+2}(z_2) \end{array} \right| \frac{X_s}{w_{2s+1}(z_1)} - z_2 w_{2s+1}(z_1) w_{2s+2}(z_2) \]

(50)

where

\[
X_s = \frac{z_1 \xi_{2s+1}(z_1, z_2) + z_2 \xi_{2s+1}(z_2, z_1)}{z_1 + z_2} - \eta_{2s+1}(z_1, z_2) - \eta_{2s+1}(z_2, z_1),
\]

\[
Y_s = 2B_{2s+2} - \eta_{2s+1}(z_1, z_2) - \eta_{2s+1}(z_2, z_1),
\]

\[
\xi_{2s+1}(z_1, z_2) = \chi_{2s+1}(z_2) w_{2s+2}(z_1) - \chi_{2s+1}(z_1) w_{2s+2}(z_2),
\]

\[
\eta_{2s+1}(z_1, z_2) = \chi_{2s+1}(z_1) w_{2s+2}(z_2) - \chi_{2s+1}(z_2) w_{2s+2}(z_1).
\]

Now since \( \eta_{2s+1}(z, z) = B_{2s+2} \), the right member of (50) becomes

\[
(z_2 - z_1)\left(w_{2s+1}(z_2) w_{2s+2}(z_2) \mu_{2s+1}(z_1, z_2) + w_{2s+1}(z_2) w_{2s+2}(z_1) \mu_{2s+1}(z_2, z_1)\right),
\]

where \( \mu_{2s+1}(z_1, z_2) = \chi_{2s+1}(z_1) w_{2s+1}(z_2) - \chi_{2s+1}(z_2) w_{2s+1}(z_1) + B_{2s+2} \), and from which (49a) follows after simplification.\(^1\) The proof of (49b) is similar. We now deduce the expansion, valid under the conditions of the theorem in (48),

\[
F(z_1, z_2) = \sum_{s=1}^{\infty} \frac{(-1)^s B_s U_s^{(1)} U_{s+1}^{(1)}}{k_s k_{s+1}},
\]

(51)

This is the third type of expansion for \( F(z_1, z_2) \), the other two, given in earlier parts, being

\[
F(z_1, z_2) = \sum_{s=1}^{\infty} \frac{B_{2s+1}(U_s^{(1)} U_{s+1}^{(1)})^2}{k_s k_{2s+2}},
\]

(52)

\[
= \sum_{s=0}^{\infty} \frac{B_{2s} U_s^{(0)} U_{s+1}^{(0)}}{k_s k_{2s+1}},
\]

(53)

where

\[
U_s^{(0)} = \frac{z_1}{w_s(z_1)} \frac{z_2}{w_s(z_2)} \left( z_2 - z_1 \right).
\]

The expansions in (51)-(53) bear a striking resemblance to those for a first order C.F., namely

\[
F(z) = \sum_{s=1}^{\infty} \frac{(-1)^s B_s}{w_{s-1} w_s},
\]

(54)

\[
= \sum_{s=0}^{\infty} \frac{B_{2s+1}}{w_{2s+1} w_{2s+2}},
\]

(55)

\[
= \sum_{s=1}^{\infty} \frac{B_{2s}}{w_{2s-1} w_{2s+1}}
\]

(56)

where \( w_s \equiv w_s(z) \).

\(^1\) The results in (49) still hold if \( z_1 = z_2 \), and we merely introduce the confluent forms of \( k_s \) and \( U_s^{(1)} \).
A Determinantal Expansion for a Class of Definite Integral 183

Under the conditions set down in §2, (55) gives an increasing sequence (56) a decreasing sequence and (54) an enveloping sequence, provided \( z \) is real and positive. Correspondingly (52) gives an increasing sequence and (53) a converging sequence when \( z_1 \) and \( z_2 \) are complex conjugates with \( \text{Im } z_1 \neq 0 \). We can go a little further than this. Suppose the interval \((0, \infty)\) is "reducible," i.e., it may reduce to a sub-interval \((a, b)\) if \( \psi(x) = \psi(a) \) for \( x < a, a \geq 0 \), and \( \psi(x) = \psi(b) \) for \( x > b, b > a \). Now using the fact that the zeros of \( w_2(x) \) are distinct and lie entirely within \((-b, -a)\) and recalling that the degrees of the highest terms in \( w_{2s}(x) \) and \( w_{2s+1}(x) \) are \( s \) and \( s + 1 \) respectively, we easily see that

\[
\begin{align*}
U_{4s}^{(1)} &> 0, \quad U_{4s+2}^{(1)} > 0, \text{ for } z_1 \leq z_2 < -b, \\
U_{4s}^{(1)} &< 0, \quad U_{4s+2}^{(1)} < 0,
\end{align*}
\]

(57a)

\[ U_{4s}^{(1)} > 0, \text{ for } z_2 \geq z_1 > -b. \]

(57b)

Hence for \( z_1 \leq z_2 < -b \) it follows from (49) that \( \{t_s\} \) is an increasing sequence, whereas this is not necessarily so for \( z_2 \geq z_1 > -a \). Again suppose \( z_2 > 0, z_1 < -b \) but \( z_2 > \mid z_1 \mid \). Then \( U_s^{(1)} \) is clearly positive. But it is known (see 51, (18)) that

\[
\int_a^b (x + z_1)(x + z_2) q_s(x) \, d\psi(x) = B_{2s+1}(z_2 - z_1)^s k_s k_{2s+2},
\]

(58)

where \( \{q_s(x)\} \) is an orthogonal system with respect to the weight function \( (x + z_1)(x + z_2) \, d\psi(x) \). We deduce from (58), since \( k_0 = 1 \), the inequalities for the even denominators

\[
\begin{align*}
k_{4s} &> 0, & k_{4s+2} &< 0, & z_2 &> 0, \\
k_{4s+1} &< 0, & k_{4s+3} &> 0, & z_1 &< -b.
\end{align*}
\]

(59)

The result given for the odd denominators may be proved in a similar way. Hence \( \{t_s\} \) is a decreasing sequence for \( z_2 > 0, z_1 < -b, z_2 > \mid z_1 \mid \).

Lastly suppose \( z_1 \) and \( z_2 > 0 \). Then \( U_s^{(1)} > 0, U_s^{(0)} > 0 \), and so from (52) and (53) \( \{t_s\} \) is an enveloping sequence. A summary of the various possibilities appears in Table 1.

---

1 i.e. \( s_r < F < s_{r+1} \) where \( s_r \) is the sum of the first \( r \) terms of the series.

2 Exceptionally, \( w_{2s+1}(x) \) always has a zero \( x = 0 \).

3 It is assumed now that \( z_1, z_2 \) are entirely real.
TABLE 1.
NATURE OF SERIES FOR SECOND ORDER C.F.

<table>
<thead>
<tr>
<th>Arguments</th>
<th>Series</th>
<th>(51)</th>
<th>(52)</th>
<th>(53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1 = \bar{z}_2$, $\text{Im}(z_1) = 0.$</td>
<td>C</td>
<td>I</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>$z_1 &lt; -b$, $z_2 &lt; -b.$</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>$z_2 &gt; 0$, $z_1 &lt; -b$, $z_2 &gt;</td>
<td>z_1</td>
<td>.$</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>$z_1 &gt; 0$, $z_2 &gt; 0$.</td>
<td>E</td>
<td>I</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

C = Converges, D = Decreases, I = Increases, E = Envelopes.


Example 1. $F(z) = \frac{1}{z + 1 + \frac{1}{z + 1 + \ldots}} = \frac{1}{2\pi} \int_0^1 \frac{\sqrt{4x^{-1} - 1}}{x+z} dx,$
for $z > 0$.

The second order C.F. for $F(z, z) = \left\{z\sqrt{z^2 + 4z}\right\}^{-1}$, $z > 0$,
is given by $\lim_{s \to \infty} t_s$ where

$$k_1^s = 1, k_2^s = 1, k_1 = z^2, k_2 = z^2 + 2z + 2,$

and $k_s^s, k_s$ follow

$$w_{2s-1} = z^2w_{2s-2} + (2z + 3)w_{2s-3} - (2z + 3)w_{2s-5} - z^2w_{2s-6} + w_{2s-7},$$
$$w_{2s} = w_{2s-1} + (2z + 3)w_{2s-2} - (2z + 3)w_{2s-4} - w_{2s-5} + w_{2s-6}, \quad s = 2, 3, \ldots$$
$$k_s^s = 0, \quad s \leq 0, \quad k_s = 0, \quad s < 0.$$

From Table 1 the sequence $\{t_s\}$ is enveloping. In particular with $z=1$ the limit of the convergents is $1/\sqrt{5}$ and the first twenty are shown in Table 2.

TABLE 2.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$k_s^s$</th>
<th>$k_s$</th>
<th>$t_s$</th>
<th>$s$</th>
<th>$k_s^s$</th>
<th>$k_s$</th>
<th>$t_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
<td>11</td>
<td>10166</td>
<td>24276</td>
<td>0.44760</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5</td>
<td>0.2</td>
<td>12</td>
<td>23416</td>
<td>63555</td>
<td>0.44700</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>10</td>
<td>0.6</td>
<td>13</td>
<td>74431</td>
<td>166405</td>
<td>0.447288</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>30</td>
<td>0.37</td>
<td>14</td>
<td>194821</td>
<td>435665</td>
<td>0.447181</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>74</td>
<td>0.49</td>
<td>15</td>
<td>510096</td>
<td>1140574</td>
<td>0.447227</td>
</tr>
<tr>
<td>6</td>
<td>85</td>
<td>199</td>
<td>0.43</td>
<td>16</td>
<td>1335395</td>
<td>2986074</td>
<td>0.447208</td>
</tr>
<tr>
<td>7</td>
<td>235</td>
<td>515</td>
<td>0.456</td>
<td>17</td>
<td>3496170</td>
<td>7817630</td>
<td>0.447216</td>
</tr>
<tr>
<td>8</td>
<td>600</td>
<td>1355</td>
<td>0.443</td>
<td>18</td>
<td>9153025</td>
<td>26466835</td>
<td>0.4472125</td>
</tr>
<tr>
<td>9</td>
<td>1500</td>
<td>3540</td>
<td>0.449</td>
<td>19</td>
<td>23963005</td>
<td>53582855</td>
<td>0.4472140</td>
</tr>
<tr>
<td>10</td>
<td>4140</td>
<td>9276</td>
<td>0.4463</td>
<td>20</td>
<td>62735880</td>
<td>140281751</td>
<td>0.4472134</td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4472136</td>
</tr>
</tbody>
</table>
Example 2. \( F(z) = \frac{1}{z} \frac{\lambda}{1} \frac{\lambda}{z + 1 + z + 1 + \ldots} \)
\[
= \frac{1}{2\pi i} \int_{0}^{\infty} \frac{\sqrt{x[4\lambda - x]}}{x(x + z)} \, dx, \quad \lambda > 0
\]
\( z > 0. \)

With \( \lambda = \frac{1}{4} \) we consider \( F(\sqrt{2}, -\sqrt{2}) = \lim_{z \to \infty} \),

where \( k^s_r = X_s/4^s, k_s = Y_s/4^s \), and the recurrence for \( X_s \) and \( Y_s \) is
\[
w_{2s-1} = -8w_{2s-2} + 3w_{2s-3} - 3w_{2s-5} + 8w_{2s-6} + w_{2s-7},
\]
\[
w_{2s} = 4w_{2s-1} + 3w_{2s-2} - 3w_{2s-4} - 4w_{2s-5} + w_{2s-6}, \quad s = 2, 3, \ldots,
\]
with
\[
X_1 = 4, \quad X_2 = 16, \quad Y_0 = 1, \quad Y_1 = -8, \quad Y_2 = -30,
\]
\( X_s = 0, \quad s \leq 0, \quad Y_s = 0, \quad s < 0. \)

The sequence \( \{-t_s\} \) steadily increases to \( \frac{1}{2} \{\sqrt{(2 + \sqrt{2})} - \sqrt{(2 - \sqrt{2})}\} \),

and a few values are stated in Table 3.

<table>
<thead>
<tr>
<th>( s )</th>
<th>(-t_s)</th>
<th>( s )</th>
<th>(-t_s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5</td>
<td>8</td>
<td>0.541186</td>
</tr>
<tr>
<td>2</td>
<td>0.53</td>
<td>9</td>
<td>0.541937</td>
</tr>
<tr>
<td>3</td>
<td>0.537</td>
<td>10</td>
<td>0.541953</td>
</tr>
<tr>
<td>4</td>
<td>0.5396</td>
<td>11</td>
<td>0.541958</td>
</tr>
<tr>
<td>5</td>
<td>0.5408</td>
<td>12</td>
<td>0.5419602</td>
</tr>
<tr>
<td>6</td>
<td>0.5411</td>
<td>13</td>
<td>0.5419608</td>
</tr>
<tr>
<td>7</td>
<td>0.54116</td>
<td>( \infty )</td>
<td>0.5419610</td>
</tr>
</tbody>
</table>

Example 3. From
\[ F(z) = \frac{1}{z} \frac{a}{1} \frac{a + 1}{z + 1} \frac{a + 2}{z + 1} \frac{a + 3}{z + 1} \ldots \]
\[
= \frac{1}{\Gamma(a)} \int_{0}^{\infty} \frac{e^{-z x^{a-1}}}{x + z} \, dx, \quad a > 0, \quad z > 0,
\]

we derive the C.F. for \( F(iz, -iz) \) which we write
\[
\Phi(a, z) = \frac{1}{\Gamma(a)} \int_{0}^{\infty} \frac{e^{-z x^{a-1}}}{x^2 + z^2} \, dx = \text{l.i.s.} \left( t_{z^2} \right), \quad z \to 0,
\]

where \( k^\infty_r \) and \( k_s \) follow
$$w_{2s-1} = z^2 w_{2s-2} + (s - 1) (2a + 3s - 4) w_{2s-3}$$
$$- (s - 1) (s - 2) (a + s - 2) (a + 3s - 5) w_{2s-5}$$
$$- (s - 1) (s - 2) (a + s - 2) (a + s - 3) z^2 w_{2s-6}$$
$$+ (s - 1) (s - 2)^2 (s - 3) (a + s - 2) (a + s - 3) w_{2s-7},$$

$$w_{2s} = w_{2s-1} + (a + s - 1) (a + 3s - 2) w_{2s-2}$$
$$- (s - 1) (a + s - 1) (a + s - 2) (2a + 3s - 4) w_{2s-4}$$
$$- (s - 1) (s - 2) (a + s - 1) (a + s - 2) w_{2s-5}$$
$$+ (s - 1) (s - 2) (a + s - 1) (a + s - 2)^2 (a + s - 3) w_{2s-6},$$

with

$$k_0^* = 0, k_1^* = k_2^* = 1, k_0 = 1, k_1 = z^2, k_2 = z^2 + a(a + 1);$$

$$k_3^* = 0, k_4 = 0$$ for $$s < 0.$$

The sequence $$\{t_{2s}\}$$ is increasing and $$\{t_{2s+1}\}$$ is convergent. In particular the coefficients in the recurrence relations for $$\Phi(1, 1)$$ are set out in Table 4, and they are to be read off from the penultimate row upwards. Thus, suppose we have found the values of $$k_s$$ for $$s = 1, 2, 3$$ and 4; then from the column for $$s = 5$$ we see that $$k_5 = 1.k_1 + 14k_3 + 0.k_2 - 20.k_1 - 4k_0 + 0.k_1,$$ and similarly for $$k_s^*.$$

**TABLE 4.**

<table>
<thead>
<tr>
<th>Recurrence coefficients for $$\Phi(1, 1).$$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s)</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>(0)</td>
</tr>
<tr>
<td>(2)</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>(s)</td>
</tr>
</tbody>
</table>

A sequence of decreasing upper bounds is made available from

$$z^2 \Phi(1, z) = 1 - 2\Phi(3, z)$$

using the appropriate second order C.F. arising from

$$\Phi(3, z) = \frac{1}{z} + \frac{3}{z} + z + 1 + \frac{5}{z} + \frac{1}{z} + z + 1 + \frac{7}{z} + 1 + \ldots.$$

The corresponding multipliers in the recurrence formulae are, when $$z_1 = i, z_2 = -i,$$ those in Table 5.
A Determinantal Expansion for a Class of Definite Integral

TABLE 5.
Recurrence coefficients for $\Phi(3, 1)$.

<table>
<thead>
<tr>
<th>$s$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>8</td>
<td>28</td>
<td>22</td>
<td>50</td>
<td>42</td>
<td>78</td>
<td>68</td>
<td>112</td>
<td>100</td>
<td>152</td>
<td>138</td>
<td>198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Proceeding in this way we are led to the approximations given in Table 6, which also includes similar ones for $\Phi(1, 2)$.

TABLE 6.
Upper and lower bounds for $\Phi(1, 1)$ and $\Phi(1, 2)$.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$\Phi(1, 1)$ (1)</th>
<th>$\Phi(1, 1)$ (2)</th>
<th>$\Phi(1, 2)$ (1)</th>
<th>$\Phi(1, 2)$ (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.52</td>
<td>0.74</td>
<td>0.196</td>
<td>0.205</td>
</tr>
<tr>
<td>8</td>
<td>0.612</td>
<td>0.653</td>
<td>0.1992</td>
<td>0.1998</td>
</tr>
<tr>
<td>12</td>
<td>0.6199</td>
<td>0.6284</td>
<td>0.19931</td>
<td>0.19954</td>
</tr>
<tr>
<td>16</td>
<td>0.6200</td>
<td>0.6227</td>
<td>0.19946</td>
<td>0.19953</td>
</tr>
<tr>
<td>20</td>
<td>0.6204</td>
<td>0.6217</td>
<td>0.19950</td>
<td>0.19952</td>
</tr>
<tr>
<td>24</td>
<td>0.6209</td>
<td>0.6216</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(1) and (2) refer to lower and upper bounds respectively.

It will be noticed that the rate of convergence for $\Phi(1, 1)$ is rather slow, and that after 24 terms we can only assert that $0.6209 < \Phi(1, 1) < 0.6216$. For $\Phi(1, 2)$ the situation is better and twenty terms give accuracy in the fourth decimal place. Of course we could determine another set of upper bounds using (10), merely adding $s! s! / (2k_s)$ to $t_s$; however, there seems to be little improvement introduced in this way. According to Ser (1938) the values of the integrals are $\Phi(1, 1) = 0.62145$, $\Phi(1, 2) = 0.199510$.

7. Conclusion.

We intend to develop on another occasion the expansion of generalised C.F.'s using the compound determinants of § 2, each element of these determinants being a recurrent. We shall give two
types of recurrence formulae, and show that the evaluation of the convergents of a generalised C.F. of any order can be made a practical proposition.

I would like to put on record my appreciation of some stimulating remarks, and criticisms, of a referee.

REFERENCES.


College of Technology, Manchester.
Generalised Algebraic Continued Fractions related to Definite Integrals

By L. R. Shenton

(Received 11th May, 1955.)

1. Introduction.

The present paper is a continuation of the work initiated in [1]-[5]. In [5] I gave an expansion of the form

$$a_0|V_s, W_{s+1}| \div |U_s, V_{s+1}|$$

for the second order C.F. associated with

$$F(z_1, z_2) = \int_0^\infty \frac{d\psi(x)}{(x+z_1)(x+z_2)},$$

where $U_s, V_s, W_s$ satisfy a fourth-order recurrence relation, there being a similar expansion for third order C.F.'s. I shall now give simple expressions for $U_s, V_s, W_s$ (or related forms) in terms of $\chi_{2s}(z_1), \chi_{2s}(z_2), \omega_{2s}(z), \omega_{2s}(z_2)$, where

$$\frac{\chi_{2s}(x)}{\omega_{2s}(x)} = \frac{a_0}{x+c_1} - \frac{a_1}{x+c_2} - \ldots - \frac{a_{s-1}}{x+c_s},$$

and show that there is a remarkable relation between the recurrence formula for the first order C.F. and that satisfied by $U_s, V_s, W_s$. The generalised form of these results will be stated and proved.

Some remarks on other forms of generalisations of C.F.'s are given in conclusion, reference being made to the correspondence between Stieltjes and Hermite on this subject.

2. Expansions involving symmetric functions.

The basis of the discussion is the function defined by

$$F(z) = \frac{1}{k_1 z} + \frac{1}{k_2 + k_3 z} + \frac{1}{k_4 + \ldots} \quad (1)$$

1 By an algebraic continued fraction we mean one whose partial denominators (and perhaps numerators) are functions of a single variable $x$, contrasting with a C.F. with numerical elements.
in which it is assumed that the $k$'s are positive and $\Sigma k_s$ diverges. This implies the existence of a unique bounded non-decreasing function $\psi(x)$ in the interval $(0, \infty)$ such that

$$F(z) = \int_0^z \frac{d\psi(x)}{x+z}. \quad (2)$$

For our present purpose a form equivalent to (1) is preferred, namely

$$F(z) = \frac{b_1}{z} + \frac{b_2}{1+z} + \frac{b_3}{1+z+1} + \ldots \quad (\psi(x) = \chi_s(x) \text{ as } s \to \infty) \quad (3)$$

and its "contracted" form \(^1\)

$$F(z) = \frac{a_0}{z+c_1} - \frac{a_1}{z+c_2} - \frac{a_2}{z+c_3} - \ldots \quad \left( = \frac{P_s(z)}{Q_s(z)} \text{ as } s \to \infty \right). \quad (4)$$

Let $A(x)$, $B(x)$ and $C(x)$ be polynomials with real coefficients, $A(x)$ of degree $n$ or less, $B(x)$ of degree less than $n$, and suppose

$$C(x) = \prod_{\lambda=1}^{n} (x+z_\lambda) > 0 \quad \text{for} \quad x \geq 0, \quad z_\lambda \text{ distinct.}$$

From (28) of §3 [4], we have an $n$th order C.F. given by

$$F(z) = \left[ \int_0^\infty \frac{A(x)B(x)}{C(x)} \, d\psi(x) \right]^{\psi(z)} = \lim_{s \to \infty} \frac{P_s(A, B)}{Q_s(z)} \quad (5)$$

where

\begin{align}
\text{(a)} & \quad Q_s(z) = |Q_s(z_1), Q_s(z_2), \ldots, \Delta|, \\
\text{(b)} & \quad P_s(A, B) = \frac{(-1)^n}{\Delta} \begin{vmatrix}
-(A, B) & W_s(B) & W_{s+1}(B) & \ldots & W_{s+n-1}(B) \\
A(-z_1) & Q_s(z_1) & Q_{s+1}(z_1) & \ldots & Q_{s+n-1}(z_1) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A(-z_n) & Q_s(z_n) & Q_{s+1}(z_n) & \ldots & Q_{s+n-1}(z_n)
\end{vmatrix}
\end{align}

\(^1\) For convenience in later sections we shall assume that $a_0 = 1.$
with (i)
\[
(A, B) = \begin{vmatrix}
1 & 1 & \cdots & 1 \\
z_1 & z_2 & \cdots & z_n \\
\vdots & \vdots & \ddots & \vdots \\
z_1^{n-2} & z_2^{n-2} & \cdots & z_n^{n-2}
\end{vmatrix} \mathcal{\Delta}^n,
\] (6c)

\[
\beta_s = B(-z_s) \int_0^\infty \frac{A(x) - A(-z_s)}{x + z_s} \, d\phi(x), \quad s = 1, 2, \ldots, n,
\]

and (ii) \(W_s(B)\) similar to \((A, B)\) except that the last row of the numerator determinant is replaced by \(\gamma_1, \gamma_2, \ldots, \gamma_n\), where \(\gamma_{rs} = P_s(z_r) B(-z_r)\).

The elements of the determinants in (6a) and (6b) may be irrational or rational, depending upon whether the roots of the polynomial are complex or real. We shall now show how they may be transformed into rational forms, a step making them more amenable for computational purposes.

3. Rational Forms.

Let \(\Delta U_s^{(\lambda)}\) be the product of \(\Delta\) and the determinant formed from the array
\[
\begin{bmatrix}
z_1^0 & z_2^0 & z_3^0 & \cdots & z_n^0 \\
z_1^1 & z_2^1 & z_3^1 & \cdots & z_n^1 \\
\vdots & \vdots & \ddots & \vdots \\
z_1^{n-1} & z_2^{n-1} & z_3^{n-1} & \cdots & z_n^{n-1} \\
Q_s(z_1) & Q_s(z_2) & Q_s(z_3) & \cdots & Q_s(z_n)
\end{bmatrix}
\]
after deleting the row with index \(\lambda, \lambda = 0, 1, \ldots, n-1\), and let
\[
\Delta U_s^{(\lambda)} = \sum_{r=1}^n A_r^{(\lambda)} Q_s(z_r).
\] (8)

Moreover let \(A\) be an \(n \times n\) matrix whose \((i, j)\)th element is \(A_{ij}^{(n-i)}\); similarly let the elements of the \(n \times n\) matrices \(U_{s,n}\) and \(Q_{s,n}\) be \(U_{s+j-1}^{(n-1)}\) and \(Q_{s+j-1}(z_i)\) respectively. Then evidently
\[
AQ_{s,n} = \Delta U_{s,n}.
\] (9)

But by postmultiplying \(A\) by the \(n \times n\) alternant matrix \([z_i^{n-1}]\) it is found that
\[
|A| = \Delta^{n-1},
\] (10)
so that
\[ Q_s(z_i)^n = \left| U_x^{(n-1)} U_x^{(n-2)} \ldots U_x^{(0)} \right|. \] (11)

Similarly, by premultiplying the bordered matrix of \( Q_{s,n} \) [the determinant of which appears in (6b)] by \(^1\)

\[
\begin{bmatrix}
1 & O^T \\
O' & A
\end{bmatrix}
\]

where \( O = [0, 0, \ldots, 0] \),

it will be found that

\[ P_s(A, B) = (-1)^n \begin{vmatrix}
-(A, B) & W_s(B) & W_{s+1}(B) & \ldots & W_{s+n-1}(B) \\
U^{(n-1)}(A) & U_x^{(n-1)} & U_{x+1}^{(n-1)} & \ldots & U_{x+n-1}^{(n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
U^{(0)}(A) & U_x^{(0)} & U_{x+1}^{(0)} & \ldots & U_{x+n-1}^{(0)}
\end{vmatrix}, \] (12)

where \( U^{(k)}(A) \) is defined by the scheme in (7) except that the last row in the array is replaced by \( A(-z_1), A(-z_2), \ldots, A(-z_n) \). The expressions (12) and (11) for the numerator and denominator of the generalised C.F. associated with \( F(z)_{1}^{n} \) are now in terms of rational elements, each element being expressible in terms of the symmetric functions of the roots of \( C(x) = 0 \).

It is to be noted that we may write

\[ W_s(B) = \sum_{r=1}^{n} (-1)^{n-r} b_r V_s^{(n-r)} \], \] (13)

where \( B(x) = \sum_{r=1}^{n} b_r x^{n-r} \) and \(^2\) \( V_s^{(\lambda)}, \lambda = 0, 1, \ldots, (n-1) \), is similar to \( U_x^{(n-1)} \) except that the last row of the numerator determinant is

\( z_1^{\lambda} P_s(z_1), z_2^{\lambda} P_s(z_2), \ldots, z_n^{\lambda} P_s(z_n) \).

If the degree of \( A(x) \) is less than \( n \), the roles of \( A(x) \) and \( B(x) \) may be interchanged, yielding the identity

\[ P_s(A, B) = P_s(B, A). \] (14)

Again, if \( A(x) B(x) = A^*(x) B^*(x) \), where \( A^*(x) \) and \( B^*(x) \) are poly-

---

\(^1\) In this matrix \( O' \) means a column matrix of \( n \) components.

\(^2\) It will be observed that \( V_s^{(\lambda)} \) is the \( n \)th divided difference of \( z^\lambda P_s(z) \) with respect to the arguments \( z_1, z_2, \ldots, z_n \). Similarly \( U_x^{(n-1)} \) is the \( n \)th divided difference of \( Q_s(z) \).
nomials, the degree of $B^\star(x)$ being less than $n$, then
\[
F(z_\alpha)^1 = \lim_{s \to x} \frac{P_s(A^\star, B^\star)}{Q_s(z_\alpha)^1}
\]
\[
= \text{l.i.s.} \frac{P_s(B^\star, B^\star)}{Q_s(z_\alpha)^1} \text{ if } A^\star = B^\star.
\]

The only modification required in (11) and (12) when the roots $z_\alpha$ are not distinct is to replace each alternant by its confluent form; this presents no difficulties since each element is the ratio of alternants in the roots $z_\alpha$.

4. Recurrence relation satisfied by $V_s^{(j)}$, $U_s^{(j)}$.

It will be seen that $V_s^{(j)}$ and $U_s^{(j)}$ are linear functions of $P_s(z_j)$ and $Q_s(z_j)$, $j = 1, 2, ..., n$, respectively, the coefficients being independent of $s$. We shall now prove that they satisfy the same recurrence relation, and that this recurrence relation can be directly derived from that satisfied by $P_s(x)$ or $Q_s(x)$. We require the following theorem.

**Lemma.** If the polynomial $T_s(x)$ is defined by
\[
T_{s+1}(x) = (x + c_{s+1}) T_s(x) - \tilde{a}_s T_{s-1}(x), \quad s = 0, 1, ..., \\
T_0(x) = 1, \quad \tilde{a}_s = a_s \text{ for } s > 0, \quad \tilde{a}_s = 0 \text{ for } s \leq 0,
\]
then
\[
\prod_{\lambda=1}^{s} (x - z_\lambda) T_s(x) = \prod_{\lambda=1}^{s} (\xi_s - z_\lambda) T_s(x), \quad s = 0, 1, ..., 
\]
where $\xi_s$ is the finite difference operator $E - c_{s+1} + \tilde{a}_s E^{-1}$ and $ET_s(x) = T_{s+1}(x)$.

**Proof.** This proceeds on inductive lines. For $n = 1$
\[
(x - z_1) T_s(x) = T_{s+1}(x) - (z_1 + c_{s+1}) T_s(x) + \tilde{a}_s T_{s-1}(x),
\]
and for $n = 2$
\[
(x - z_1)(x - z_2) T_s(x)
\]
\[
= (x - z_2) T_{s+1}(x) - (x - z_2)(z_1 + c_{s+1}) T_s(x) + \tilde{a}_s(x - z_2) T_{s-1}(x), \quad s = 0, 1, ..., \\
= (\xi_{s+1} - z_2) T_{s+1}(x) - (z_1 + c_{s+1})(\xi_s - z_2) T_s(x) + \tilde{a}_s(\xi_{s-1} - z_2) T_{s-1}(x), \\
\]
\[
= \{E(\xi_s - z_2) - (z_1 + c_{s+1})(\xi_s - z_2) + \tilde{a}_s E^{-1}(\xi_s - z_2)\} T_s(x)
\]
\[
= (\xi_s - z_1)(\xi_s - z_2) T_s(x) \text{ for } s > 0.
\]

\footnote{1 In evaluating this expression a particular value of $s$ can only be inserted after the operational symbols have been absorbed.}
It may also be shown to hold for \( s = 0 \), when the fact that \( \tilde{a}_0 = \tilde{a}_{-1} = 0 \) conserves the identity. Now assume the truth of the statement for polynomials \( \Pi(x - z_a) \) of degree less than or equal to \( n - 1 \). Then

\[
\prod_{\lambda=1}^{n} (x-z_\lambda) \, T_\lambda'(x) = \prod_{\lambda=1}^{n-1} (x-z_\lambda) \{ T_{s+1}(x) - (z_n + c_{s+1}) \, T_s(x) + \tilde{a}_s \, T_{s-1}(x) \}, \quad s \geq 0,
\]

\[
= \prod_{\lambda=1}^{n-1} (\xi_{s+1} - z_\lambda) \, T_{s+1}(x) - (z_n + c_{s+1}) \, \prod_{\lambda=1}^{n-1} (\xi_s - z_\lambda) \, T_s(x)
\]

\[+ \tilde{a}_s \prod_{\lambda=1}^{n-1} (\xi_{s-1} - z_\lambda) \, T_{s-1}(x) \text{ for } s \geq 1 \quad (16)\]

Now, if \( s = 0 \),

\[
\prod_{\lambda=1}^{n} (x-z_\lambda) \, T_0(x) = \prod_{\lambda=1}^{n-1} (x-z_\lambda) \{ T_1(x) - (z_n + c_1) \, T_0(x) \}
\]

\[= \prod_{\lambda=1}^{n-1} (x-z_\lambda) \{ T_1(x) - (z_n + c_1) \, T_0(x) + \tilde{a}_0 \, T_{-1}(x) \}.
\]

Thus (16) still holds for \( s = 0 \). Hence

\[
\prod_{\lambda=1}^{n} (x-z_\lambda) \, T_s(x) = \left( E - (z_n + c_{s+1}) + \tilde{a}_s \, E^{-1} \right) \prod_{\lambda=1}^{n-1} (\xi_s - z_\lambda) \, T_s(x), \quad s \geq 0,
\]

from which the truth of the lemma follows. It is now evident that

\[
\prod_{\lambda=1}^{n} (\xi_s - z_\lambda) \, T_s(x) = 0 \text{ for } \begin{cases} x = z_1, z_2, \ldots, z_n \\ s \geq 0 \end{cases}
\]

and indeed if

\[
\Omega_s = \sum_{r=1}^{n} q_r \, T_s(z_r)
\]

then

\[
\prod_{\lambda=1}^{n} (\xi_s - z_\lambda) \, \Omega_s = 0, \quad s \geq 0,
\]

which is the recurrence relation for \( \Omega_s \). (It is to be understood that \( q_r \) is independent of \( s \).)

We now apply the lemma and (18) to the polynomials appearing in the numerators and denominators of C.F. (4); i.e. to \( P_s(z) \) and \( Q_s(z) \), where

\[
\frac{P_s(z)}{Q_s(z)} = \frac{a_0}{z + c_1} - \frac{a_1}{z + c_2} - \frac{a_2}{z + c_3} - \ldots - \frac{a_{s-1}}{z + c_s} \quad \begin{cases} s = 1, 2, \ldots \\ a_0 = 1 \end{cases}
\]

with

\[
\begin{cases} Q_0 = 1, \quad Q_1 = z + c_1, \\ P_0 = 0, \quad P_1 = 1. \end{cases}
\]
It is evident that we have proved that the recurrence relation followed by \( U_\lambda^{(s)} \) defined in (8) is

\[
\begin{align*}
C(-\xi_s) U_\lambda^{(s)} & = 0 \\
C(-E + c_{s+1} - \tilde{d}_s E^{-1}) U_\lambda^{(s)} & = 0,
\end{align*}
\]

where \( \tilde{d}_s = a_s \) for \( s > 0 \), \( \tilde{d}_s = 0 \) for \( s \leq 0 \), the \( a \)'s and \( c \)'s being given in (19). Similarly after a slight modification, the recurrence followed by \( V_\lambda^{(s)} \) is found to be

\[
C(-E + c_{s+1} - \hat{d}_s E^{-1}) V_\lambda^{(s)} = 0, \quad s = 1, 2, \ldots
\]

where \( \hat{d}_s = a_s \) for \( s > 1 \), \( \hat{d}_s = 0 \) for \( s \leq 1 \).

5. A fundamental system of solutions.

The recurrence relations (20a) and (20b), each of order \( 2n \), are exactly the same for \( s \geq 1 \). We shall now prove that \( U_\lambda^{(s)}, V_\lambda^{(s)}, \lambda = 1, 2, \ldots, n-1 \), is a fundamental system of solutions with respect to this common recurrence equation. For consider

\[
\Phi_{s,n} = \begin{bmatrix} U_{s,n} & U_{s+n,n} \\ JV_{s,n} & JV_{s+n,n} \end{bmatrix},
\]

where \( V_{s,n} \) is the same matrix function of the \( V \)'s as \( U_{s,n} \) is of the \( U \)'s, and \( J \) is an \( n \times n \) matrix with zeros everywhere except in the secondary diagonal, where there are units. It will be seen that

\[
\begin{bmatrix} A & O \\ O' & A^* \end{bmatrix} \begin{bmatrix} Q_{s,n} & Q_{s+n,n} \\ P_{s,n} & P_{s+n,n} \end{bmatrix} = \Delta \Phi_{s,n},
\]

where \( P_{s,n} \) is a matrix similar to \( Q_{s,n} \), with \( P(z) \) replacing \( Q(z) \), and

\[
A^* = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_1 & \cdots & z_n \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{n-1} & z_1^{n-1} & \cdots & z_n^{n-1} \end{bmatrix} \begin{bmatrix} A_1^{(n-1)} & 0 & \cdots & 0 \\ 0 & A_2^{(n-1)} & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n^{(n-1)} \end{bmatrix}.
\]

Hence, taking determinants in (22), we have

\[
\Delta^{2n} |\Phi_{s,n}| = |A| \cdot |A^*| (-1)^{n(n-1)} \Theta_s (z_1, z_2, \ldots, z_n),
\]

where

\[
\Theta_s (z_1, z_2, \ldots, z_n) = \prod_{i=1}^{n} (z - z_i).
\]
where

\[
\Theta_s(z_1, z_2, \ldots, z_n) = \begin{vmatrix}
Q_s(z_1) & Q_{s+1}(z_1) & \cdots & Q_{s+2n-1}(z_1) \\
P_s(z_1) & P_{s+1}(z_1) & \cdots & P_{s+2n-1}(z_1) \\
\vdots & \vdots & \ddots & \vdots \\
Q_s(z_n) & Q_{s+1}(z_n) & \cdots & Q_{s+2n-1}(z_n) \\
P_s(z_n) & P_{s+1}(z_n) & \cdots & P_{s+2n-1}(z_n)
\end{vmatrix},
\]

(24)

the general expression of this type being \(\Theta_{s+r}(z_{r+1}, z_{r+2}, \ldots, z_n)\), in which \(s\) indicates the suffix in the first column of the determinant, which is of order \(2(n-r)\), the arguments \(z_{r+1}, z_{r+2}, z_{r+2}, \ldots\) occurring in successive rows.

Now \(P_s(x)\) and \(Q_s(x)\) follow

\[y_s = (x+c_s)y_{s-1} - a_{s-1}y_{s-2}.\]

Using this on the last column of (24) and eliminating \(Q_{s+2n-1}(z_1)\) and \(P_{s+2n-1}(z_1)\), and so on until all the elements in the first two rows except the first four are removed, we find

\[\Theta_s(z_1, z_2, \ldots, z_n) = \alpha_s \Theta_{s+1}(z_2, z_3, \ldots, z_n) \prod_{\lambda=2}^{n} (z_{\lambda} - z_1)^2,\]

(25)

where \(\alpha_s = \prod_{\lambda=0}^{s} \alpha_{\lambda} = |Q_s(x), P_{s+1}(x)|\).

Continuing the condensation process in evidence in (25), it will be found that

\[\Theta_s(z_1, z_2, \ldots, z_n) = \alpha_s \prod_{2}^{n} (z_{\lambda} - z_1)^2 \cdot \alpha_{s+1} \prod_{3}^{n} (z_{\lambda} - z_2)^2 \cdots \alpha_{s+n-2} (z_{n-1} - z_n)^2 \cdot \alpha_{s+n-1}
\]

\[= \Delta^2 \prod_{s=1}^{n-1} \prod_{\lambda=0}^{s} a_{\lambda},\]

(26)

But \(|A^*| = (-1)^{n(n-1)} \Delta^{n-1}\) and from (10) \(|A| = \Delta^{n-1}\), so that from (23)

\[|\Phi_{s,n}| = \prod_{v-s}^{v} \prod_{\lambda=0}^{s} a_{\lambda}.\]

(27)
We have thus proved that $U^{(0)}$, $V^{(0)}$, $\lambda = 0, 1, 2, \ldots, n-1$ form a fundamental system of solutions of the recurrence relation (20). It is evident that the result still holds when $C(x) = 0$ has multiple roots.

6. Illustration.

We consider $n = 3$ and $C(x) = (x+\alpha)(x+\beta)(x+\gamma) > 0$ for $t \geq 0$, and the integral

$$I_m = \int_0^\infty \frac{x^m \psi(x)}{(x+\alpha)(x+\beta)(x+\gamma)}, \quad m = 0, 1, 2,$$

where it is assumed $\psi(x)$ satisfies the conditions of §2. The denominators of the associated generalised 3rd order C.F. all have the form

$$Q_s(\alpha, \beta, \gamma) = |U_s^{(2)}, U_{s+1}^{(1)}, U_{s+2}^{(0)}|,$$

where

$$U_s^{(2)} = \begin{vmatrix}
1 & 1 & 1 \\
\alpha & \beta & \gamma \\
Q_s(\alpha) & Q_s(\beta) & Q_s(\gamma)
\end{vmatrix} \div |\alpha^0, \beta^1, \gamma^2|,$$

$$U_s^{(1)} = \begin{vmatrix}
1 & 1 & 1 \\
\alpha^2 & \beta^2 & \gamma^2 \\
Q_s(\alpha) & Q_s(\beta) & Q_s(\gamma)
\end{vmatrix} \div |\alpha^0, \beta^1, \gamma^2|,$$

$$U_s^{(0)} = \begin{vmatrix}
\alpha^2 & \beta^2 & \gamma^2 \\
Q_s(\alpha) & Q_s(\beta) & Q_s(\gamma)
\end{vmatrix} \div |\alpha^0, \beta^1, \gamma^2|.$$

The confluent forms are easily constructed; for example if

(i) $\beta = \gamma$

$$U_s^{(2)} = \begin{vmatrix}
1 & 1 & 0 \\
\alpha & \beta & 1 \\
Q_s(\alpha) & Q_s(\beta) & Q_s'(\beta)
\end{vmatrix} \div |\alpha^0, \alpha^1, 1|,$$

(ii) $\alpha = \beta = \gamma$

$$U_s^{(2)} = \begin{vmatrix}
1 & 0 & 0 \\
\alpha & 1 & 0 \\
Q_s(\alpha) & Q_s(\alpha) & Q_s'(\alpha)
\end{vmatrix} \div |\alpha^0, \beta^1, \gamma^2|.$$

1 See for example H. W. Turnbull, Theory of Equations, p. 48 (Edinburgh, 1946).
2 Primes indicate derivatives.
For the numerators there are several possibilities, which we briefly indicate. Let \( A_s = x^s, s = 0, 1, 2 \); then it is easily seen that \( U^{(r)}(A_s) = \delta_{r,s}, r = 0, 1, 2 \) where \( \delta \) is the Kronecker delta symbol. Moreover,

\[
W_s(A_s) = (-1)^s \begin{vmatrix} 1 & 1 & 1 \\ x & \beta & \gamma \end{vmatrix} = (-1)^s V_s^{(s)},
\]

and from (6c)

\[(A_r, A_s) = 0, \ r+s \leq 2, \ r, s = 0, 1, 2.\]

We now have

\[I_0 = \lim_{s \to \infty} \frac{P_s(A_0, A_0)}{Q_s}, \tag{30}\]

\[I_1 = \lim_{s \to \infty} \frac{P_s(A_1, A_0)}{Q_s} = \lim_{s \to \infty} \frac{P_s(A_0, A_1)}{Q_s} \tag{31},\]

\[I_2 = \lim_{s \to \infty} \frac{P_s(A_2, A_0)}{Q_s} = \lim_{s \to \infty} \frac{P_s(A_0, A_2)}{Q_s} = \lim_{s \to \infty} \frac{P_s(A_1, A_1)}{Q_s}, \tag{32}\]

where \( Q_s \equiv Q_s(\alpha, \beta, \gamma) \) and

\[
P_s(A_0, A_0) = \begin{vmatrix} V_s^{(0)} & V_s^{(0)} & V_s^{(0)} \\ U_s^{(2)} & U_s^{(2)} & U_s^{(2)} \\ U_s^{(1)} & U_s^{(1)} & U_s^{(1)} \end{vmatrix},
\]

\[
P_s(A_1, A_0) = \begin{vmatrix} V_s^{(0)} & V_s^{(0)} & V_s^{(0)} \\ U_s^{(2)} & U_s^{(2)} & U_s^{(2)} \\ U_s^{(1)} & U_s^{(1)} & U_s^{(1)} \end{vmatrix},
\]

\[
P_s(A_2, A_0) = \begin{vmatrix} V_s^{(0)} & V_s^{(0)} & V_s^{(0)} \\ U_s^{(2)} & U_s^{(2)} & U_s^{(2)} \\ U_s^{(1)} & U_s^{(1)} & U_s^{(1)} \end{vmatrix},
\]

\[
P_s(A_1, A_1) = \begin{vmatrix} V_s^{(0)} & V_s^{(1)} & V_s^{(1)} \\ U_s^{(2)} & U_s^{(2)} & U_s^{(2)} \\ U_s^{(1)} & U_s^{(1)} & U_s^{(1)} \end{vmatrix}.
\]
Recurrence relation for \( V_s^{(1)} \) and \( U_s^{(1)} \). From (20a) and (20b) this is
\[
(E - c_{s+1} + \tilde{a}_s E^{-1} - \alpha)(E - c_{s+1} + \tilde{a}_s E^{-1} - \beta)(E - c_{s+1} + \tilde{a}_s E^{-1} - \gamma) U_s^{(1)} = 0,
\]
\[s \geq 0,
\]
with a similar expression for \( V_s^{(1)} \). On expansion this becomes
\[
U_{s+3}^{(1)} - \lambda_s U_{s+2}^{(1)} + \mu_s U_{s+1}^{(1)} - \nu_s U_s^{(1)} + \tilde{a}_s \mu_{s-1} U_{s+1}^{(1)} - \tilde{a}_s \tilde{a}_{s-1} \lambda_{s-2} U_{s-2}^{(1)}
\]
\[+ \tilde{a}_s \tilde{a}_{s-1} \tilde{a}_{s-2} U_{s-3}^{(1)} = 0, \quad \lambda = 0, 1, 2; \quad s \geq 0,
\]
(33)
where
\[
\begin{align*}
\lambda_s &= \alpha + \beta + \gamma + c_{s+1} + c_{s+2} + c_{s+3} \\
\mu_s &= \beta \gamma + \gamma \alpha + \alpha \beta + (\alpha + \beta + \gamma)(c_{s+1} + c_{s+2}) + c_{s+1}^2 + c_{s+1} c_{s+2} + c_{s+2}^2 \\
\nu_s &= (\alpha + c_{s+1})(\beta + c_{s+1})(\gamma + c_{s+1}) + (\alpha + \beta + \gamma)(\tilde{a}_s + \tilde{a}_{s+1})
\]
\[+ \tilde{a}_s \tilde{a}_{s+1} + \tilde{a}_{s+2},
\]
and
\[
\begin{align*}
U_0^{(0)} &= 1 & U_1^{(0)} &= c_1 & U_2^{(0)} &= c_1 c_2 - a_1 \\
U_0^{(1)} &= 0 & U_1^{(1)} &= -1 & U_2^{(1)} &= -c_1 - c_2 \\
U_0^{(2)} &= 0 & U_1^{(2)} &= 0 & U_2^{(2)} &= 1,
\end{align*}
\]
with \( \tilde{a}_s = a_s, s > 0; \tilde{a}_s = 0, s \leq 0; \) and \( a_s, c_s \) are given in the C.F. in (4).
Similarly \( V_s^{(1)} \) follows (33) for \( s \geq 1 \) provided \( \tilde{a}_s \) is replaced throughout by \( \hat{a}_s \) where \( \hat{a}_s = a_s, s > 1; \hat{a}_s = 0, s \leq 1; \) and the initial values are
\[
\begin{align*}
V_0^{(0)} &= 0 & V_1^{(0)} &= 0 & V_2^{(0)} &= 0 & V_3^{(0)} &= 1 \\
V_0^{(1)} &= 0 & V_1^{(1)} &= 0 & V_2^{(1)} &= 1 & V_3^{(1)} &= \alpha + \beta + \gamma + c_2 + c_3 \\
V_0^{(2)} &= 0 & V_1^{(2)} &= 1 & V_2^{(2)} &= \alpha + \beta + \gamma + c_2 & V_3^{(2)} &= \alpha^2 + \beta^2 + \gamma^2 + \beta \gamma + \gamma \alpha + \alpha \beta
\]
\[+ (c_2 + c_3)(\alpha + \beta + \gamma)
\]
\[+ c_2 c_3 - a_2.
\]
The recurrence relation (33) is the same (apart from minor changes due to the notation used) as that given for \( U_s, V_s, W_s, X_s \) under expression (13) in [5]. With the present notation and approach, however, there is no difficulty in writing down the recurrence relation for \( V_s^{(1)} \) and \( U_s^{(1)} \) together with their initial values.
Remarks on generalised C.F.'s.

Previous attempts at generalising C.F.'s have to a certain extent consisted in extending the recurrence relation satisfied by the numerator and denominator, contributions having been made by Jacobi [6], Hermite [7], Perron [8], Bateman [9], and Paley and Ursell [10]. In addition it seems that Stieltjes was very close to the theory developed here, for in the correspondence between Stieltjes and Hermite [11] Stieltjes indicated (see letter 167 written in 1889) that he had perfected a method of approximating the integral

\[ P = \int_a^b \frac{f(x)}{\phi(x)} \, dx, \]

using least squares. In fact he minimises the expression

\[ R_n = \int_a^b \frac{f(x)}{\phi(x)} \left(1 + x_1 \phi + x_2 \phi^2 + \ldots + x_n \phi^n\right)^2 \, dx \]  \hspace{1cm} (34)

with respect to \( x_1, x_2, \ldots, x_n \) and obtains the approximation

\[ P = \begin{vmatrix} 0 & c_0 & c_1 & \cdots & c_{n-1} \\ c_0 & c_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_n & c_{n+1} & \cdots & c_{2n-1} \end{vmatrix} \frac{1}{|c_1, c_3, \ldots, c_{2n-1}|}, \]

where \( c_k = \int_a^b \phi^k f(x) \, dx \).

Stieltjes goes on to show that there is an alternative form

\[ P = \sum_{\ell=1}^{n} \frac{A_{\ell}^2}{B_{\ell-1}B_{\ell}} \]

where\[ \begin{align*} A_{\ell} & = |c_0, c_2, \ldots, c_{2\ell-2}| \\ B_{\ell} & = |c_1, c_3, \ldots, c_{2\ell-1}|, \end{align*} \]

and then disposes of the remainder when the limits of integration are finite\(^1\). As an application he derives the even part of the C.F. for

\[ \int_a^b \frac{f(x)}{z-x} \, dx. \]

\(^1\) As far as I can gather Stieltjes does not state the restrictions on \( \phi \) and \( f \), but his method of proving that the remainder vanishes in the limit is noteworthy. Apparently, another letter by Stieltjes on this topic has been lost according to a footnote by the editors.
Hermite (letter 168 in the correspondence) remarks "votre nouveau point de vue pour obtenir l'approximation de l'intégrale
\[ \int_a^b \frac{f(x)}{x-z} \, dx, \]
constitue un très heureux et très grand progrès ...". However, as far as I can trace the only later references to this development [see Stieltjes (1889), "Sur un développement en fraction continue", Comptes Rendus cviii, p. 1297; also Stieltjes (1918) Oeuvres Complètes, "Recherches sur les fractions continues", Ch. viii, pp. 500-502. Groningen] suggest that the broader possibilities were not envisaged. It is perhaps worth recording here that the generalisation we have developed, springs from the minimisation of the expression
\[ \int_a^b C(x) \left( \frac{A(x)}{C(x)} - \pi_n(x) \right)^2 \, d\psi(x), \]
with respect to the coefficients in the polynomial \( \pi_n \), \( C(x) \) being in general positive in \( (a, b) \), convergence questions being settled by invoking Parseval's theorem. The difficulties in generalising (34) by taking \( \phi \) to be a quadratic or polynomial of higher degree are evident.

Lastly we may remark that Hermite evidently\(^1\) gave some thought to the question of generalising C.F's, but his contribution to the subject in 1893 [7] appears to have no relation to (33) or (34).

REFERENCES.

5. ———, ibid. (2), 10 (1957), 167–188.
8. O. Perron, Math. Annalen, 64 (1907), 1–76.

MANCHESTER COLLEGE OF SCIENCE AND TECHNOLOGY,
MANCHESTER, 1.

---

\(^1\) There are about a dozen references to the subject in letters written by Hermite to Stieltjes; see [11], Tome 2 and, for example, letters 234–236, 241, 242, 245, 265, 278, 388, 389, 391 and 392.
INEQUALITIES FOR THE NORMAL INTEGRAL INCLUDING A NEW CONTINUED FRACTION

BY L. R. SHENTON
College of Technology, Manchester

1. INTRODUCTION

In this note we consider the related normal integral ratios

\[ R(t) = \int_t^\infty g(x) \, dx / g(t), \]

sometimes called Mills's ratio, and

\[ \tilde{R}(t) = \int^t_0 g(x) \, dx / g(t), \]

where \( g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2} \), and \( R(t) + \tilde{R}(t) = 1/[2g(t)] \). We give a new continued fraction† for \( \tilde{R}(t) \) which is rapidly convergent for small values of \( t \) and which incidentally provides a new set of inequalities. The rapidity of convergence is compared with a series for \( \tilde{R}(t) \) and with the Laplace c.f. for \( R(t) \). This assessment is similar to recent work by Teichroew (1952) on the comparative rapidity of convergence of series and c.f.'s for the elementary function \( e^t \), \( \ln(1+x) \) and \( \arctan x \). Lane (1944) has also considered the same sort of thing for interpolation, comparing Newton's series and Thiele's c.f. in the case of the function \( 2^x \).

We also prove and generalize a conjecture of Birnbaum‡ (1950) that for \( t \geq 0 \),

\[ R(t) < 4/[3t + \sqrt{(t^2 + 8)}], \]

it being shown that there are two sequences of similar inequalities, increasing and decreasing to the limiting value \( R(t) \). In the Appendix we set out a brief summary of results relating to the normal integral.

2. THE CONTINUED FRACTION FOR \( \tilde{R}(t) \)

The new c.f. is given by

\[ \tilde{R}(t) = \frac{t}{1 - \frac{t^2}{3 - \frac{2t^2}{5 - \frac{3t^2}{7 - \frac{4t^2}{9 - \frac{5t^2}{11 + \cdots}}}}}}, \]

and is convergent for \( t > 0 \). This is derived from the c.f. representation of the series

\[ \tilde{R}(t) = \frac{t}{1 + \frac{t^3}{1.3 + \frac{t^5}{1.3.5 + \cdots}}}, \]

which may be expressed in terms of the confluent hypergeometric function

\[ _1F_1(a; b; x) = 1 + \frac{a x}{b} + \frac{a(a+1)x^2}{b(b+1)2!} + \cdots. \]

We then use the c.f. of Gauss (1812, pp. 134–5) in its confluent form

\[ _1F_1(1; b; x) = 1 + \frac{x}{1 - b + \frac{bx}{1 - b + \frac{2x}{b(b+3) - \frac{(b+1)x}{b+4} + \cdots}}}. \]

(See, for example, Perron, 1913, pp. 311–14; Wall, 1948, pp. 349–55.)

† Continued fraction will be abbreviated to c.f. throughout.
‡ Birnbaum actually conjectured \((t^2-1) R^2 - 3t R + 2 > 0\).
The expression (1) now follows since \( \bar{R}(t) = t_1 F_1(1; \frac{3}{2}; t^2) \). Denoting the \( s \)th convergent of (1) by \( r_s = p_s/q_s \), it follows that \( p_s \) and \( q_s \) satisfy the recurrence relations

\[
\begin{align*}
y_{2s+1} &= (4s + 1)y_{2s} + 2st^2y_{2s-1} \\
y_{2s} &= (4s - 1)y_{2s-1} - (2s - 1)t^2y_{2s-2}
\end{align*}
\]

(5)

with \( p_0 = 0, p_1 = t, q_0 = 1, q_1 = 1 \).

It may be shown that for \( t > 0 \), \( q_{4s} > 0, q_{4s+1} > 0 \), so that from (5) we have

\[
\begin{align*}
r_0 < r_1 < r_4 < r_5 < r_8 < \ldots < \bar{R} \quad (t > 0).
\end{align*}
\]

(6)

The situation for the remaining convergents is not so simple, but it transpires that

\[
\begin{align*}
r_{4s+2} > r_{4s+3} > r_{4s+6} > r_{4s+7} > \ldots > \bar{R} \\
\end{align*}
\]

(7)

for some positive integer \( s \). We therefore have the following simple inequalities

\[
\begin{align*}
r_1 &= t < r_4 &= \frac{105t + 5t^3}{105 - 30t^2 + 15t^4} < r_5 &= \frac{945t + 105t^3 + 81t^5}{945 - 210t^2 + 15t^4} < \bar{R} \quad (t > 0),
\end{align*}
\]

(8)

\[
\begin{align*}
r_2 &= \frac{3t}{3 - t^2} > r_3 &= \frac{15t + 2t^3}{15 - 3t^2} > r_6 &= \frac{10395t + 630t^3 + 63t^5}{10395 - 2835t^2 + 315t^4 - 15t^6} > \bar{R},
\end{align*}
\]

(9)

the ranges of \( t \) for (9) being \( 0 < t < \sqrt{3}, 0 < t < \sqrt{5}, 0 < t < \sqrt{8} - 283 \) respectively.

We now write the c.f. as

\[
\bar{R}(t) = t \left\{ \frac{t^2}{1} - \frac{2t^2}{3} + \frac{3t^2}{5 - 7} - \frac{4t^2}{9 - 11} + \frac{5t^2}{13 - 15} + \frac{6t^2}{17 - 19} + \frac{8t^2}{21} - \frac{10t^2}{23} - \ldots \right\}
\]

(10)

the convergents corresponding to the asterisks forming a decreasing sequence exceeding \( \bar{R}(t) \), the remaining convergents forming an increasing sequence less than \( \bar{R}(t) \), this state of affairs holding sooner or later for all \( t > 0 \). It certainly always holds for \( 0 < t < \sqrt{3} \). Otherwise, for example, if \( t = 2, 2.5, 3.0 \) it holds from the 3rd, 4th and 7th convergents, respectively, onwards. Further properties of the c.f. are listed in the Appendix.

**Numerical illustration**

We give in Table 1 the convergents for the examples \( t = 0.1 \) and \( t = 3.0 \). It will be seen that for \( t = 0.1 \) eight convergents give accuracy in the 22nd place of decimals, a rapid rate of convergence of about 6 decimal places for every two convergents. For \( t = 3.0 \) the convergence is slow at first, being due to the negative value of some of the early denominators. However, convergence soon accelerates, there being a gain of four decimal places between the 16th and 20th convergents. We may mention that the increase in the number of digits in the convergents of the c.f. creates a computational difficulty. Thus with \( t = 2.0 \) the 18th convergent involves 21 digits; with \( t = 2.5 \) employing an equivalence transformation to clear decimals, the 22nd convergent involves 35 digits. This situation can be avoided if we know the number of terms required, or by curtailing at the expense of some loss of accuracy. Three methods of using c.f.'s on high-speed computing machines are discussed by Teichroew (1952).

In Table 2 we give the values of the absolute differences \( |\Delta r_s| = |r_s - r_{s-1}| \), and the corresponding differences \( |\Delta T_s| = |T_s - T_{s-1}| \) for the series (referred to as series C)

\[
\bar{R}(t) = e^{4t^2} \left\{ \frac{t^3}{3} - \frac{t^5}{3 \cdot 2 \cdot 1!} + \frac{t^7}{5 \cdot 2^2 \cdot 2!} - \ldots \right\} = \lim_{s \to \infty} T_s,
\]
Table 1. Convergents of $R(t)$

<table>
<thead>
<tr>
<th>$t = 0.1$</th>
<th>$s$</th>
<th>$r_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 0.1$</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.100, 334, 4*</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.100, 334, 001, 3*</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.100, 334, 000, 953, 2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.100, 334, 000, 953, 438, 993</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.100, 334, 000, 953, 440, 116, 2*</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.100, 334, 000, 953, 440, 116, 180, 63*</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.100, 334, 000, 953, 440, 116, 180, 61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$t = 3.0$</th>
<th>$s$</th>
<th>$r_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1.5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-8.2</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>164.0</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>98.7*</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>112.515, 2*</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>112.515, 1</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>112.515, 152, 98*</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>112.515, 152, 96</td>
</tr>
</tbody>
</table>

Note. Convergents marked with an asterisk exceed the value of $R(t)$.

Table 2. Differences of successive convergents of $R(t)$ and of corresponding partial sums of the series $C$

| $s$ | $|\nabla r_s|$ | $|\nabla T_s|$ | $s$ | $|\nabla r_s|$ | $|\nabla T_s|$ |
|-----|--------------|--------------|-----|--------------|--------------|
| $t = 0.1$ |
| 2   | 3.33         | 3.17         |
| 4   | 9.39         | 9.30         |
| 6   | 15.12        | 15.24        |
| 8   | 22.18        | 21.10        |
| $t = 0.5$ |
| 2   | 1.45         | 1.24         |
| 4   | 4.34         | 4.26         |
| 6   | 8.68         | 7.13         |
| 8   | 12.64        | 11.36        |
| 10  | 16.36        | 15.61        |
| $t = 1.0$ |
| 2   | 0.50         | 0.27         |
| 4   | 2.63         | 2.49         |
| 6   | 4.21         | 4.39         |
| 8   | 7.31         | 6.17         |
| 10  | 10.27        | 9.47         |
| 12  | 13.15        | 12.77        |
| $t = 1.5$ |
| 2   | 1.45         | 1.17         |
| 4   | 0.24         | 0.16         |
| 6   | 2.35         | 2.63         |
| 8   | 4.26         | 3.14         |
| 10  | 6.11         | 5.19         |

Digits in bold type before the decimal place † imply that the decimal is to be multiplied by this power of ten; thus 1·45 means 0·045 and 1·45 means 4·5.

† The notation is due to A. C. Aitken (1925, footnote to p. 293).
for several values of \( t \). These are given for \( s \) even, since in this case \( \bar{R}(t) \) always lies between \( r_s \) and \( r_{s-1} \) for \( 0 < t < \sqrt{3} \) (and sooner or later for all \( t > 0 \)). Similar remarks apply to \( T_s \) and \( T_{s-1} \). Thus the error in the approximations \( r_s \) and \( T_s \) is less in value than \( |\nabla r_s| \) and \( |\nabla T_s| \) respectively. It will be observed that the C.F. converges ultimately more rapidly than the series, as far as the present analysis is concerned. Whether this state of affairs holds for larger values of \( s \) than those considered can only be conjectured.

Thus the error in the approximations \( r_s \) and \( T_s \) is less in value than \( Vr_s \) and \( VT_s \) respectively.

It will be observed that the C.F. converges ultimately more rapidly than the series, as far as the present analysis is concerned. Whether this state of affairs holds for larger values of \( s \) than those considered can only be conjectured.

For the early convergents the series has a slight advantage, especially for small \( t \); but for \( t > 1 \) the C.F. soon becomes more accurate, although for larger values of \( t \) there is the disadvantage of the negative approximations in the early stages. However, this disadvantage may be overcome by using the generalization of (1), namely,

\[
\bar{R}(t) = \sum_{r=0}^{s-1} \frac{t^{2r+1}}{1.3.5 \ldots (2r+1)} + \frac{t^{2s}}{1.3 \ldots (2s+1)} \left[ \frac{t}{2s+1} - \frac{(2s+1) t^2}{2s+3} + \frac{2t^2}{2s+5} - \frac{(2s+3) t^2}{2s+7} + \frac{4t^3}{2s+9} - \ldots \right].
\] (11)

If \( s \) is taken large enough in this expression (which converges for \( t > 0 \)) negative approximations are avoided. Numerical examples support the view that the rate of convergence of (11) is similar to that of (1), but from an analytical point of view there are difficulties.

The series (2) in its rate of convergence is not a serious rival to the expansions considered in Table 2. Thus the C.F. appears to be on the whole as good an approximation to \( \bar{R}(t) \) as any other.

### 3. THE LAPLACE CONTINUED FRACTION FOR \( R(t) \)

It can be shown (see, for example, Bromwich, 1926, p. 388) that

\[
R(t) = t \int_{-\infty}^{\infty} g(x) \, dx = \frac{t}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} x^s \, dx,
\] (12)

so that

\[
R(t) = \frac{1}{t} - \frac{1}{t^3} + \frac{1.3}{t^5} - \ldots + (-1)^{s+1} \frac{1.3 \ldots (2s-1)}{t^{2s+1}} + R_s(t),
\] (13)

where

\[
| R_s(t) | = t^{-(2s+1)} \int_{-\infty}^{\infty} \frac{x^{2s+2} g(x) \, dx}{x^2 + t^2} < \frac{1.3 \ldots (2s)}{t^{2s+3}},
\]

which states the familiar fact that the error involved in using any number of terms of (13) is always less in value than the first term omitted.

In terms of the hypergeometric function we have

\[
R(t) = \frac{1}{t} - \frac{1}{t^3} + \frac{1.3}{t^5} - \ldots + (-1)^{s+1} \frac{1.3 \ldots (2s-1)}{t^{2s+1}} + (-1)^{s+1} \frac{1.3 \ldots (2s+1)}{t^{2s+3}} \lim_{c \to \infty} F \left( \frac{1}{2}, \frac{s+3}{2} ; c ; \frac{2c}{t^2} \right),
\] (14)

so that using the C.F. of Gauss for this case we find

\[
R(t) = \frac{1}{t} - \frac{1}{t^3} + \frac{1.3}{t^5} - \ldots + (-1)^{s} \frac{1.3 \ldots (2s-1)}{t^{2s+1}} + (-1)^{s+1} \frac{1.3 \ldots (2s+1)}{t^{2s+3}} \left[ \frac{1}{t} + \frac{2s}{t} + \frac{2}{t} + \frac{2s+7}{t} + \frac{4}{t} + \frac{2s+7}{t} + \ldots \right],
\] (15)
the C.F. part of this converging for \( t > 0 \), the odd convergents forming a decreasing sequence, and the even convergents an increasing sequence. If we take \( s + 1 = 0 \) in (15) we derive the Laplace (1805) C.F.

\[
R(t) = \frac{1}{t + t + t + t + t + \ldots} \quad (16)
\]

The following results are required in the sequel:

(i) Denoting the \( st \)th convergent of (16) by \( \phi_s = \chi_s/\omega_s \), we see that \( \chi_s \) and \( \omega_s \) follow the recurrence relation

\[
\begin{align*}
\chi_s &= ty_{s-1} + (s-1)y_{s-2} \quad (s = 2, 3, \ldots), \\
\chi_0 &= 0, \quad \chi_1 = 1, \quad \omega_0 = 1, \quad \omega_1 = t.
\end{align*} \quad (17)
\]

From (17) it is easily shown that

\[
\begin{align*}
\phi_s - \phi_{s-1} &= (-)^s(1-s-1)!/(\omega_{s-1}\omega_s), \\
\phi_s - \phi_{s-2} &= (-)^s(s-2)!/(\omega_{s-2}\omega_s),
\end{align*} \quad (18)
\]

so that we have \( \phi_0 < \phi_2 < \phi_4 \ldots < R(t) < \ldots < \phi_2 < \phi_0 \) \quad (19)

(ii) The integral for \( R(t) \) may be written

\[
R(t) = \int_0^\infty \exp \left( -\frac{1}{2}x^2 - xt \right) dx,
\]

which after differentiation and the use of (17) yields

\[
R\omega_s - \chi_s = (-)^s\int_t^\infty (x-t)^s g(x) dx/g(t). \quad (20)
\]

(iii) It is well known that \( \omega_s \) is a Hermite polynomial and satisfies the differential equation

\[
d^2\omega_s/dt^2 + t \frac{d\omega_s}{dt} - s\omega_s = 0. \quad (22)
\]

Or if we set \( \omega_s = e^{-ut} w_s \), then

\[
d^2w_s/dt^2 = \left( \frac{1}{4}t^2 + s + \frac{1}{2} \right) w_s. \quad (23)
\]

We require an asymptotic formula for \( w_s \) for large \( s \). Following a method outlined in Jeffreys & Jeffreys (1946, § 17.122) we put \( w_s = \exp \left( \int_0^t \eta(\theta) d\theta \right) \), so that \( \eta^2 + \frac{d\eta}{dt} = \frac{1}{4}t^2 + s + \frac{1}{2} \). This leads to the asymptotic expressions (in \( s \))

\[
\begin{aligned}
\omega_{2s} &= e^{-1t} \frac{1}{(1 + t^2)(8s + 2)} \cosh \{\lambda_{2s}(t)\}, \\
\omega_{2s+1} &= e^{-1t} \frac{1}{(1 + t^2)(8s + 6)} \sinh \{\lambda_{2s+1}(t)\},
\end{aligned} \quad (24)
\]

where

\[
\lambda_s(t) = \frac{1}{2}t \sqrt{\left( \frac{3}{4}t^2 + s + \frac{1}{2} \right) + (s + \frac{1}{2}) \ln \left( \sqrt{(1 + t^2)(4s + 2)} + t/\sqrt{(4s + 2)} \right)}.
\]

Using an approximation to the factorial we find the following approximations to the difference between two convergents:

\[
\begin{align*}
\phi_{2s+1} - \phi_{2s} &= \frac{2}{(2\pi)^{1/2}} \left( \sqrt{(1 + t^2)(8s + 2)} \left( 1 + t^2(8s + 6) \right) \right)^{1/2} \exp \left( \frac{1}{2}t^2 - \frac{1}{2}t \sqrt{8s + 2} - \frac{1}{2}t \sqrt{8s + 6} \right), \\
\phi_{2s-1} - \phi_{2s} &= \frac{2}{(2\pi)^{1/2}} \left( \sqrt{(1 + t^2)(8s - 2)} \left( 1 + t^2(8s + 2) \right) \right)^{1/2} \exp \left( \frac{1}{2}t^2 - \frac{1}{2}t \sqrt{8s - 2} - \frac{1}{2}t \sqrt{8s + 2} \right),
\end{align*} \quad (25)
\]

The more complicated formulae for these differences found by using (24) in (18) is perhaps more reliable when \( t \) has a moderate value such as greater than three.
4. Numerical Comparison of Three Expansions for $R(t)$

We shall consider the expansions

$$R(t) = \frac{1}{1} \frac{1}{t} \frac{2}{t^2} \frac{3}{t^3} + \cdots$$

(A)

$$\bar{R}(t) = \frac{t}{1} \frac{t^2}{3} \frac{2t^2}{5} \frac{3t^2}{7} + \cdots$$

(B)

$$\tilde{R}(t) = e^{it} \left[ \frac{t}{1} \frac{t^3}{3 \cdot 2 !} + \frac{t^5}{5 \cdot 2^2 \cdot 2 !} - \frac{t^7}{7 \cdot 2^3 \cdot 3 !} + \cdots \right]$$

(C)

each of which can be used to approximate to $R(t)$ through the relation $R(t) + \bar{R}(t) = 1/(2\gamma)$, from the point of view of rapidity of convergence. We consider how many terms are required so that the error is $10^{-(n+1)} \times 2.5$. More precisely, if $C_{s-1}$ and $C_s$ are successive convergents of a C.F. (or partial sums of a series) providing lower and upper bounds, then we evaluate the least value of $s$ (an integer) for which $|C_{s-1} - C_s| \leq 10^{-(n+1)} \times 2.5$ is true approximately. This is about equivalent to ensuring accuracy in the $n$th place of decimals. For the Laplace C.F. we use the approximation

$$2\sqrt{(2\pi)} \exp \left\{ \frac{1}{2} t^2 - \frac{1}{2} t \sqrt{(4s + 2) - \frac{t}{2} \sqrt{(4s - 2)}} \right\} = 10^{-(n+1)} \times 2.5,$$

the solution of which is given by

$$s = \psi(t) + 1/[16\psi(t)], \quad \psi(t) = \left[ \left( \frac{1}{4} t + (1.4992 + 1.1513n) \right) \right]^2,$$

where the second term is unimportant for small $t$. As a simple crude formula for (A) we have $|\phi_{n-1} - \phi_n| = 2\sqrt{2\pi} \exp \left( \frac{1}{2} t^2 - 2t \sqrt{s} \right)$. Thus, approximately, the necessary number of terms varies directly as the square of the accuracy, and inversely as the square of $t$. We have already seen from §2 that the rates of convergence of (B) and (C) are very similar. Indeed, for small values of $t$ we have $r_{4s+2} - r_{4s+1} = \pi(T_{4s+2} - T_{4s+1}) / 2^{4s-1}$, so that (B) has an advantage in view of the factor $2^{4s-1}$. A similar relation for $t$ not small appears to be complicated.

A comparison of the three expansions is given in Table 3.

Table 3. Number of terms to achieve a given accuracy for three expansions

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>2.5 x 10^{-7}</th>
<th>2.5 x 10^{-13}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expansion</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>7,072</td>
<td>4</td>
</tr>
<tr>
<td>0.25</td>
<td>1,135</td>
<td>4</td>
</tr>
<tr>
<td>0.5</td>
<td>287</td>
<td>6</td>
</tr>
<tr>
<td>0.75</td>
<td>130</td>
<td>6</td>
</tr>
<tr>
<td>1.0</td>
<td>75</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td>2.0</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>2.5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>3.0</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>4.0</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>5.0</td>
<td>7</td>
<td>—</td>
</tr>
<tr>
<td>10.0</td>
<td>4</td>
<td>—</td>
</tr>
</tbody>
</table>
Remarks on Table 3

First of all it must be remarked that precise accuracy in a table of this kind is unnecessary. Most of the numbers in the region of 25 or less were found by direct calculation, but the rest are approximate. Again there is a slight difference between finding the number of terms \( s \) of an expansion to ensure an error less than \( \epsilon \) when the true value is known, as compared to relying on the expansion alone to give this information. From this point of view since the terms in (B) are in pairs greater than and less than the true value, the numbers given often give greater accuracy than is quoted. Thus four terms of (B) give accuracy in the ninth decimal place.

![Graph showing number of terms for accuracy](image)

**Fig. 1.** Number of terms \( s \) for accuracy to be of the order \( 10^{-(n+1) \times 2.5} \) for the Laplace c.f. and for series C.

It will be seen that the number of terms for a required accuracy increases with decreasing \( t \) for the Laplace c.f., the increase being rapid as \( t \) approaches zero, while for (B) and (C) there is a steady increase with increasing \( t \). Again for (A) the ratio of the number of terms for 12 places of decimals as compared to 6 is approximately three, whereas the corresponding factor for (B) and (C) is less and decreases as \( t \) increases. Expressed otherwise, we may say that increasing accuracy for the Laplace c.f. for a given value of \( t \) is only achieved at the expense of an increasing additional number of terms. To take an extreme case, we remark that for \( t = 0.1 \), the additional numbers of terms required for successive jumps of six places
of decimals, starting from zero, are: 7072, 16390, 25934, 35477, 45020, 54564 approximately. For the series (C) the corresponding numbers for the case \( t = 10 \) are: 187, 10, 10, 9, 9.

A graphical comparison of the expansions (A) and (C) showing the logarithm of the number of terms against the accuracy is given in Fig. 1.† It is evident from this that the usefulness of the series (C) falls off slowly with increasing \( t \), and from the flatness of the curves when \( n \) is large (for a fixed \( t \) it is clear that although convergence may be slow at first, sooner or later it becomes rapid. By contrast the Laplace c.f. falls off rapidly in its usefulness as \( t \) decreases, and there is little tendency to flatness for increasing \( n \) when \( t \) is fixed.

To discuss the behaviour of (B) for moderate or large values of \( t \) requires an asymptotic expression for the denominators of the convergents, and this is not at present available. From the numerical evidence of Table 3 it would appear that (B) may be quicker in converging than (C).

5. SCHWARZIAN INEQUALITIES AND BIRNBAUM’S CONJECTURE

Birnbaum (1942) gave the inequality \( R(t) > [-t + \sqrt{(t^2 + 4)}]/2 \) \( (t > 0) \), and later (1950) conjectured what amounts to \( R(t) < 4/[3t + \sqrt{(t^2 + 8)}] \) for \( t > 0 \). We prove generalizations of these, using the inequality of Schwarz, which in its simplest form states that if \( f_1(x) \) and \( f_2(x) \) are linearly independent functions, then

\[
\begin{vmatrix}
\int_a^b f_1^2(x) \, dx & \int_a^b f_1(x) f_2(x) \, dx \\
\int_a^b f_2(x) f_1(x) \, dx & \int_a^b f_2^2(x) \, dx
\end{vmatrix} > 0.
\]

Hence taking \( f_1(x) = (x-t)^s e^{-x^2} \), \( f_2(x) = (x-t)^{s+2} e^{-x^2} \) in (26) using (21) we find

\[
F_s(R) = \left| \frac{R \omega_s - \chi_s - \chi_{s+1} - R \omega_{s+1}}{\chi_{s+1} - R \omega_{s+1} - \omega_{s+1} \chi_s} \right| > 0 \quad (s = 0, 1, 2, \ldots ; t \geq 0).
\]

This leads to the two sets of inequalities‡:

\[
\begin{align*}
R &> [v_{2s+1} + (2s)\sqrt{(t^2 + 8s + 4)}]/(2\mu_{2s}) = R_{2s}, \\
R &< 2\pi_{2s+1} [v_{2s+1} + (2s + 1)\sqrt{(t^2 + 8s + 8)}] = R_{2s+1},
\end{align*}
\]

where \( \mu_s = \omega_s \omega_{s+1} - \omega_{s+1}^2 \), \( v_s = \omega_s \chi_s + \omega_{s+1} \chi_{s+1} - 2 \omega_{s+1} \chi_s \), \( \pi_s = \chi_s \chi_{s+1} - \chi_{s+1}^2 \), and \( \chi_s/\omega_s \) is the \( s \)th convergent of the Laplace c.f. The sign of the radical in (28) is easily determined provided we can show that for \( t > 0 \), (i) \( \mu_{2s} > 0 \) and (ii) \( \pi_{2s+1} > 0 \). For if this is the case we see that since

\[
F_s'(\phi_s) < 0, \quad F_s'(\phi_{s+2}) < 0, \quad F_s'(\phi_{s+1}) > 0,
\]

and also from (19) that \( R \) lies between \( \phi_{s+1} \) and \( \phi_{s+2} \), then

\[
\phi_{2s+2} < R_{2s} < R < \phi_{2s+1}.
\]

† I am indebted to Dr. J. A. Storrow for drawing this diagram.
‡ When \( t = 0 \), \( R_{2s}(0) = \frac{2^{2s-1} t(s-1)!}{(2s-1)! \sqrt{(2s+1)}} \), so that in the limit \( R(0) = \sqrt{\frac{1}{4} \pi} \) (from the formula of Wallis), providing a new demonstration of the value of the total area under the normal curve. A similar result follows from \( R_{2s+1}(0) \).
where \( R_{2s} \) is the larger root of \( F_s(R) = 0 \). Similarly, if (ii) holds, then by considering the sign of \( \hat{R}^2 F_{s+1}(\hat{R}^{-1}) \) when \( \hat{R}^{-1} = \phi_{2s+1} \), \( \phi_{2s+2} \), \( \phi_{2s+3} \), we see that

\[
\phi_{2s+2} < R < \phi_{2s+1} < \phi_{2s+3},
\]

where \( R_{2s+1} \) is the smaller root of \( F_{s+1}(R) = 0 \).

To prove (i) we have the integral representation \( \omega_s = \int_{-\infty}^{\infty} g(x) (t-x)^s dx \), so that the result follows immediately from the inequality of Schwarz. For (ii) there is the recurrence relation

\[
\pi_s = (s^2 - 1) \pi_{s-2} + \chi_{s-1}(\chi_{s+1} - s + 1 \chi_{s-1}),
\]

and if \( \sigma_s = \chi_s - s \chi_{s-2} \) then \( \sigma_{s+2} = (s-1) \sigma_s + t^2 \chi_s \). But \( \sigma_s = t, \sigma_4 = t^3 + t \), and so \( \sigma_{2s} \geq 0 \) since \( \chi_{2s} \geq 0 \) for \( t \geq 0 \). Thus \( \pi_{2s+1} > 0 \) since \( \pi_1 = 2 \).

From (19), (30) and (31) it is evident that \( \lim_{s \to \infty} R_s = R(t) \) (\( t \geq 0 \)). Murty (1952) has given a partial proof of some of the inequalities in (28). Sampford (1953) has shown that \( R < R_1 \) holds for the extended range \( t > -1 \). Similar extensions could be given for other cases also; for example, since the roots of \( F_s(R) = 0 \) are of opposite sign for \( 0 > t > -2 \), we have \( R > (t^2 - 2t + \sqrt{(t^2 + 12)(t^3 + 3)}) / (t^3 + 3) \) for \( t > -2 \).

It is obvious from (30) that \( R_{2s} \) increases from time to time with \( s \); but it is not obvious that \( \{R_{2s}\} \) is a monotonic increasing sequence. This can be proved by considering the recurrence relations

\[
F_s(y) = -s F_{s-1}(y) + (y \omega_s - \chi_s)^2 \quad (32)
\]

and

\[
F_s(y) = -F_{s-1}(y) + s(s-1) F_{s-2}(y) + t(y \omega_s - \chi_s) \times (y \omega_{s-1} - \chi_{s-1}) \quad (33)
\]

For from (32) it is evident that \( F_s(R_{s-1}) > 0 \), and using this in (33) we see that \( F_s(R_{s-2}) < 0 \). It follows that altogether we have

\[
F_{2s}(\phi_{2s}) < 0, \quad F_{2s}(R_{2s}) = 0, \quad F_{2s}(\phi_{2s+1}) > 0, \quad F_{2s}(R_{2s-2}) < 0,
\]

which in conjunction with (30) proves that \( R_{2s} > R_{2s-2} \). Similarly, we can prove \( R_{2s+2} < R_{2s+1} \).

Hence we have the result, to be compared with (19),

\[
R_0 < R_2 < R_4 < \ldots < R < \ldots < R_2 < R_4 < R_1 \quad (t \geq 0),
\]

in which \( R_0 \) is a better approximation than the Laplace \( \phi_{2s+2} \). As an indication of the closeness of approximation of these Schwarzian sequences, we give the value of the error

Table 4. Comparison of Schwarzian approximation to \( R(t) \)

<table>
<thead>
<tr>
<th>( t )</th>
<th>( R )</th>
<th>( 0-0 )</th>
<th>1-25</th>
<th>( 0-1 )</th>
<th>1-16</th>
<th>( 0-5 )</th>
<th>0-88</th>
<th>( 1-0 )</th>
<th>0-68</th>
<th>( 2-0 )</th>
<th>0-42</th>
<th>( 3-0 )</th>
<th>0-30</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_0 )</td>
<td>0-25</td>
<td>0-21 (1-11)</td>
<td>1-96 (0-48)</td>
<td>1-38 (0-15)</td>
<td>2-72 (1-21)</td>
<td>2-18 (2-46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_1 )</td>
<td>0-16</td>
<td>0-12 (1-55)</td>
<td>1-38 (0-50)</td>
<td>1-11 (0-10)</td>
<td>2-13 (2-74)</td>
<td>3-22 (2-10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_2 )</td>
<td>1-99</td>
<td>1-70 (1-10)</td>
<td>1-19 (0-30)</td>
<td>2-42 (1-56)</td>
<td>2-32 (2-28)</td>
<td>4-36 (3-24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_s )</td>
<td>( 1-s )</td>
<td>1-53 (1-42)</td>
<td>2-20 (1-37)</td>
<td>4-95 (2-12)</td>
<td>5-74 (4-69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_4 )</td>
<td>1-61</td>
<td>1-39 (0-95)</td>
<td>2-73 (0-20)</td>
<td>2-10 (1-24)</td>
<td>4-53 (3-53)</td>
<td>5-18 (4-22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_5 )</td>
<td>1-53</td>
<td>1-32 (1-34)</td>
<td>3-57 (1-17)</td>
<td>4-13 (3-25)</td>
<td>6-49 (5-74)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P.W.</td>
<td>0-00</td>
<td>3-76</td>
<td>3-93</td>
<td>2-73</td>
<td>1-51</td>
<td>0-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figures in the body of the table represent |Approximation – True value| with the same convention as in Table 2. Figures in parentheses in the same line as \( R_s \) (\( s = 0, 1, \ldots, 5 \)), refer to \( \phi_{s+2} \). P.W. refers to the Pólya-Williams inequality.
of $R_s, s = 0$ to 5, for six values of $t$ in Table 4. For comparison we also give the error for the Laplace convergent $\phi_{s+2}$, and that for the Pólya (1945) and Williams (1946) inequality

$$\int_0^t g(x) \, dx < \frac{1}{2}[1 - \exp (-2t^2/\pi)]^{|t|} \quad (t > 0).$$

Strictly speaking the latter is a different type of inequality, and although Pólya gives a generalization of it, convergency properties have not been considered.

Remarks on Table 4

The second row gives the value of $R(t)$ to two decimal places merely to indicate the importance of the errors in the approximations. It will be observed that the Schwarzian values are always better than the corresponding Laplace convergents, and this is particularly so for small values of $t$; thus for $t = 0.1$ the error in $R_s$ is within about 4%, whereas the error in $\phi_5$ is in the region of 100%. It is also interesting to notice that the P.W. values are better than the other approximations for $t < 1.0$, but deteriorate for $t > 1.0$. Of course for $t < 1.0$ higher Schwarzian approximations would ultimately become better than the P.W. inequality.

6. Summary and conclusion

We have given a new continued fraction (C.F.) for the normal integral ratio

$$R(t) = e^{it^2} \int_0^t e^{-i2x} \, dx,$$

which turns out to be rapidly convergent for small values of $t$ (say less than $\sqrt{3}$); for moderate values convergence is slow at first but becomes quite rapid in due course; at least this is supported by numerical evidence. The rate of convergence of this C.F. is compared with a series and also with the Laplace C.F. for the Mills's ratio $R(t) = e^{it^2} \int_t^\infty e^{-i2x} \, dx$. An interesting point about the Laplace C.F. is that the difference between the $s$th and $(s-1)$th convergents has the approximate magnitude $2\sqrt{(2\pi)} \exp \left(\frac{1}{2}t^2 - 2t\sqrt{s}/s\right)$, when $t$ is not large. This, together with numerical evidence, shows that fair accuracy is attainable with a few terms when $t$ is moderate to large, but the rate of convergence deteriorates rapidly as $t$ becomes small. In addition, for a fixed value of $t$ (not large) the necessary number of terms for a given accuracy varies directly as the square of the accuracy. The behaviour of the new C.F. for $R(t)$ is completely different, for it converges rapidly for small $t$ and deteriorates slowly as $t$ increases, whereas for a fixed value of $t$, increasing accuracy seems to demand much less in terms than the sequence of the accuracy.

We also give two sets of inequalities for $R(t)$, consisting of irrational fractions, these being generalizations of results by Birnbaum. We prove that one set increases, and the other decreases, monotonically, to the limit $R(t)$. A particular case includes the formula for the total area under the normal curve.

In an appendix we give some properties of the C.F.'s for $R(t)$ and $R(t)$ which are not readily available in the literature, and a brief summary of expansions and inequalities.
APPENDIX

(1) Properties of the Laplace continued fraction for \( R(t) \)

(a) \( R(t) = \sum_{s=0}^{\infty} \left(-t^s\right)/\left(\omega_0\omega_{s+1}\right) \).

(ii) \( R(t) = t \sum_{s=0}^{\infty} \left(2s\right)\left(\omega_{2s}\omega_{2s+1}\right) \).

(iii) \( R(t) = 1/\omega_1 - t \sum_{s=0}^{\infty} \left(2s+1\right)\left(\omega_{2s+1}\omega_{2s+3}\right) \).

These follow from the general theory of C.F.'s.

(b) (i) \( \chi_{s+1} = \omega_s + \chi_s \).

(ii) \( \chi_{s+1} = \omega_s + \omega_{s-1} + \omega_{s-2} + \ldots \).

Inductional methods may be used for (i), from which (ii) is deduced. A general theorem, of which (ii) is a particular case, using more elaborate methods is given by Sherman (1933).

(c) \( \frac{R\omega_s - \chi_s}{\chi_s - R\omega_{s+1}} = \frac{s + 1}{t + t + t + \ldots} \quad (t > 0) \).

The \( r \)th convergent of the c.f. leads to the approximation \( \phi_{r+s} \) to \( R(t) \). This result may be used as a computational check.

(d) Remainder formulæe:

\[
R\omega_s - \chi_s = \frac{2^{s-3}s!(1.3)\ldots(2s-3)}{\sqrt{\pi}} \int_0^\infty e^{-x^2} \frac{e^{-x^2}}{x^2} \ dx \quad (s = 0, 1, 2, \ldots),
\]

\[
\chi_{s-1} - R\omega_{s-1} = \frac{2^{s-2}(s-1)!}{\sqrt{\pi}} \int_0^\infty e^{-x^2} \frac{e^{-x^2}}{x^2} \ dx \quad (s = 1, 2, \ldots; t > 0).
\]

These follow from the general theory of C.F.'s.

(e) Systematic summation of asymptotic series for \( R(t) \):

\[
\frac{1}{t} - \frac{1}{t^3} + \frac{1.3}{t^5} - \ldots + (-1)^s \frac{1.3\ldots(2s-3)}{t^{2s-1}} = \frac{t}{t^2 + t^4 + t^6 + \ldots} = \frac{t}{t^2 - 1 + t^2 - 3 + \ldots + t^2 -(2s-1)}.
\]

(2) Properties of the new continued fraction for \( \bar{R}(t) \)

\[
\bar{R}(t) = \frac{t}{1 - 3 + \frac{2t^2}{5} - \frac{3t^2}{7} + \ldots} = \lim_{s \to \infty} r_s = \lim_{s \to \infty} p_s/q_s \quad (t > 0).
\]

(a) Differences of convergence:

(i) \( r_{2s+1} - r_{2s} = (-2)^s s(1.3\ldots(2s-1)) \left(t^{2s+1}/(q_{2s}q_{2s+1})\right) \).

(ii) \( r_{2s} - r_{2s-1} = (-2)^{(s-1)}(1.3\ldots(2s-1)) \left(t^{2s-1}/(q_{2s-1}q_{2s})\right) \).

(iii) \( r_{2s} - r_{2s-2} = (-)^{(s-1)}(4s-1)(2s-2) \left(t^{2s-1}/(q_{2s-2}q_{2s})\right) \).

(iv) \( r_{2s+1} - r_{2s-1} = (-)^{(s-1)}(4s+1)(2s-1) \left(t^{2s-1}/(q_{2s-1}q_{2s+1})\right) \).

(b) Formulae for denominators:

(i) \( q_{2s} = \int_{-\infty}^{\infty} g(x) x^{2s-2} e^{-t^2} \ dx \).

(ii) \( q_{2s+1} = \int_{-\infty}^{\infty} g(x) x^{2s+2} e^{-t^2} \ dx \).
Inequalities for the normal integral

(c) Remainder formulae:

(i) \( R_{q_2s} - p_{2s} = \int_0^t x^{2s}(x^2 - t^2)^p \exp \frac{1}{2}(t^2 - x^2) \, dx \).

(ii) \( R_{q_2s+1} - p_{2s+1} = \int_0^t x^{2s+1}(x^2 - t^2)^p \exp \frac{1}{2}(t^2 - x^2) \, dx \).

(d) Formulae for computational checks:

(i) \( \frac{R_{q_2s+1} - p_{2s+1}}{R_{q_2s} - p_{2s}} = \frac{(2s + 1) t^2}{4s + 3} - \frac{4s + 5}{4s + 7} - \ldots \)

n terms leading to the approximation \( r_{2s+n+1} \) for \( R(t) \).

(ii) \( \frac{R_{q_2s} - p_{2s}}{p_{2s-1} - R_{q_2s-1}} = \frac{2st^2}{4s + 1} - \frac{4s + 3}{4s + 5} - \ldots \)

n terms leading to the approximation \( r_{2s+n} \) for \( R(t) \).

(e) Partial sums of series for \( R(t) \):

\[ \sum_{r=0}^{s-1} \frac{q^{2r+1}}{1.3 \ldots (2 + 1)} = \frac{t}{1 - t^2 + 3 - t^2 + 5 - \ldots - t^5 + 2s - 1}. \]

(3) Miscellaneous results

(a) Factorial type of series for \( R(t) \) due to Schlömilch (1885)

\[ tR(t) = 1 + \frac{1}{\sqrt{\pi}} \sum_{r=1}^{\infty} (-2)^r \int_0^{\infty} x^{2r} e^{-x} x^r \, dx / (u_1 u_2 \ldots u_r), \]

where \( u_r = t^2 + 2r \), \( x^{2r} = x(x-1) \ldots (x-r+1) \) \( (t>0). \)

(This result is mentioned by Wishart (1927).)

In particular

\[ tR(t) = 1 - \frac{1}{u_1} + \frac{1}{u_1 u_2} - \frac{5}{u_1 u_2 u_3} + \frac{9}{u_1 u_2 u_3 u_4} - \frac{129}{u_1 u_2 u_3 u_4 u_5} + \ldots \]

(b) \( \int_0^t g(x) \, dx < a_n^{-1} (2\pi)^{-1} \left[ n \int_0^t \omega_n x n^{-1} e^{-x^2} \, dx \right]^{1/n} \)

where \( a_n = 2\pi^{-1/2} \left[ \Gamma \left( \frac{n}{2} + 1 \right) \right]^{1/n} \) \( (n = 2, 3, \ldots). \)

Due to Pólya (1945), the particular case \( n = 2 \) being given by Williams (1946). (See also Tate (1953).)

(c) Minimum property of the convergents of the Laplace c.f.:

\[ \min_{\pi_n} \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-x^2} x^{2s} \left( \frac{1}{x + \frac{1}{2}t^2} - \pi_n(x) \right) \, dx = R(t) - X_{2s+1}/\omega_{2s+1} \]

\[ \min_{\pi_n} \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-x^2} x^{2s+1} \left( \frac{1}{x + \frac{1}{2}t^2} - \pi_n(x) \right) \, dx = t[X_{2s+1}/\omega_{2s+1} - R(t)], \]

the minimum being taken over all polynomials of precise degree \( s \). These are implied in a general theorem due to Stieltjes. (See, for example, Shohat & Tamarkin (1943), p. 75.) In a similar way we find the following inequalities:

(i) \( (t^2 + 1) R(t) - t > 2t/(t^4 + 6t^2 + 15) \),

(ii) \( (t^2 + 1) R(t) - t > 2(t^2 + 14t^2 + 75)/(t^8 + 20t^6 + 160t^4 + 420t^2 + 525) \),

(iii) \( (t^3 + 3t) R(t) < t^3 - 2 - 3/(t^4 + 10t^2 + 35) \),

(iv) \( (t^2 + 3t) R(t) < t^2 - 3/(t^4 + 18t^2 + 119)/(t^6 + 28t^4 + 294t^2 + 1260t^2 + 2205) \),

(v) \( (t^4 + 6t^2 + 3) R(t) < t^4 - 5t > 41t/(t^6 + 15t^4 + 105t^2 + 315) \),

(vi) \( (t^6 + 10t^4 + 15t) R(t) < t^4 + 9t^2 + 8 - 5t/(t^4 + 21t^2 + 189t^2 + 693) \),

where \( t > 0 \) throughout.
The weight function in probit analysis.

If

$$\psi(t) = e^{-t^2} \int_{-\infty}^{t} e^{-x^2} \, dx \int_{t}^{\infty} e^{-x^2} \, dx,$$

then $\psi(t)$ is a decreasing function of $t^2$ (Hammersley, 1950; Tate, 1953).

(e) Rough formula for the normal distribution function:

$$-10 \log \int_{t}^{\infty} g(x) \, dx = 2t^2 + 4 + 10 \log t \quad \text{with an error less than 1 if} \quad 2 \leq t \leq 14$$

(Good, 1950).

REFERENCES

A SEMI-INFINITE RANDOM WALK WITH DISCRETE STEPS

BY L. R. SHENTON

Communicated by M. S. BARTLETT

Received 12 August 1954

1. A particle executes a random walk over the possible positions \( x = 0, 1, 2, \ldots \), its initial position being \( x = d \geq 0 \). At the \( n \)th step it occupies the position \( x \) with probability \( p_n(x \mid d) \) and is in the state \((n, x)\). The transition from \((n, x)\) to \((n+1, y)\) has the probability \( p_{x,y} \) given by

\[
\begin{align*}
p_{x,x-1} &= \frac{1}{2}K + 0, \\
p_{x,x} &= k, \\
p_{x,x+1} &= \frac{1}{2}K + 0 \quad (x > 0),
\end{align*}
\]

where

\[
\begin{align*}
p_{0,x} &= p_x \\ (0 \leq x \leq m), \\
p_{0,x} &= 0 \quad (x > m).
\end{align*}
\]

If the particle arrives at \( x = 0 \), the next step may lead to absorption with probability \( p_a \), where \( p_a = 1 - \sum_{x=0}^{m} p_s \), i.e. the particle is annihilated in this case. If \( p_a = 1 \), we say that \( x = 0 \) is an absorbing barrier. If \( p_a = 1 \), we say that \( x = 0 \) is a reflecting barrier. If \( p_a = 1 \), we say that \( x = 0 \) is a retaining barrier.

It will be seen that when \( p_a = 1 \), \( p_n(0 \mid d) \) represents the probability of annihilation (or absorption) in the course of \( n-1 \) steps.

We shall give an expression for \( p_n(x \mid d) \) in the form of a contour integral. Illustrations include the gambler's ruin problem, random walk with drift as recently discussed by Kac(2), and two problems given by Lauwerier(3).

2. We introduce the g.f.

\[
\Phi(x \mid d; t) = \sum_{n=0}^{\infty} t^n p_n(x \mid d) \quad (|t| < 1),
\]

which will be abbreviated to \( \Phi_x \).

A consideration of the transition probabilities leads to

\[
F_x(\Phi) = \delta_{x,d},
\]

where

\[
\begin{align*}
F_0 &= (1 - p_a t) \Phi_0 - \frac{1}{2}K't\Phi_1, \\
F_1 &= -p_a t \Phi_0 + (1 - kt) \Phi_1 - \frac{1}{2}K't\Phi_2, \\
F_x &= -p_a t \Phi_0 - \frac{1}{2}K't\Phi_{x-1} + (1 - kt) \Phi_x - \frac{1}{2}K't\Phi_{x+1} \quad (x = 2, 3, \ldots),
\end{align*}
\]

\[
\delta_{x,d} = 1 \quad \text{if } x = d, \\
= 0 \quad \text{if } x = d.
\]

It will be shown that the solution of (3) is

\[
\Phi(x \mid d; t) = \left( \frac{K}{K'} \right)^{(x-d)} \frac{1}{2\pi it \sqrt{|KK'|}} \int_{c} \frac{\Psi(x, d; z) 2\pi dz}{zH(z)(e^{\theta} - z)(z - e^{-\theta})},
\]

where \( H(z) \) is Hurwitz's zeta function.
A semi-infinite random walk with discrete steps

where

\[
\begin{align*}
H(z) &= \sum_{s=0}^{m} (\lambda z)^{s} p_{s} - k - \frac{1}{2} \sqrt{(K'K')} (z + z^{-1}) \quad (\lambda = \sqrt{(K'/K)}), \\
H_{x}(z) &= \sum_{s=0}^{x} (\lambda z)^{s} p_{s} - k - \frac{1}{2} \sqrt{(K'K')} (z + z^{-1}) \quad (x > 0), \\
&= - \frac{1}{2} \sqrt{(K'K')} z^{-1} \quad (x = 0),
\end{align*}
\]

(i) \[\Psi(x, d; z) = H_{x}(z) z^{d-x} - H_{x}(z^{-1}) z^{d+x},\]

(ii) \[t^{-1} = k + \sqrt{(K'K')} \cosh \theta,\]

(iii) \(C\) is a contour enclosing the origin within which \(\Psi'/zH(z)\) has no poles except possibly \(z = 0\). This is equivalent to requiring \(H(z)\) to be free from zeros within and on \(C\).

For if \(x \geq 2\),

\[
F_{x}(\Phi) = -\sum_{x=0}^{m} \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \frac{1}{2\pi i} \int_{C} \frac{\Psi(x, d; z) \cosh (\theta - \frac{1}{2} (z + z^{-1}))}{zH(z) \cosh (\theta - \frac{1}{2} (z + z^{-1}))} dz + \frac{1}{2} \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \frac{1}{2\pi i} \int_{C} \frac{(z^{d+1} - z^{d-1})}{zH(z)(\cosh \theta - \frac{1}{2} (z + z^{-1}))},
\]

i.e.

\[
F_{x}(\Phi) = \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \frac{1}{2\pi i} \int_{C} \frac{\Psi(x, d; z) z}{zH(z)} dz.
\] (5)

Similarly, the result may be shown to hold for \(0 < x < 2\). It remains to show that \(F_{x}(\Phi) = \delta_{x,d}\). Consider the three cases (a) \(x > m\), (b) \(0 < x \leq m\), (c) \(x = 0\).

(a) \(x > m\):

\[
F_{x}(\Phi) = \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \frac{1}{2\pi i} \int_{C} \left( z^{d-x-1} \frac{H(z)}{H(z)} \right) dz
\]

since the residue of \(z^{m}H(z^{-1}) z^{d-x-m}\) at the origin is zero.

(b) \(0 < x \leq m\):

\[
F_{x}(\Phi) = \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \frac{1}{2\pi i} \int_{C} \left( H_{x}(z) z^{d-x} - H_{x}(z^{-1}) z^{d+x} \right) \frac{dz}{zH(z)}.
\]

But the residue of \(\frac{z^{x}H_{x}(z)}{zH(z)} z^{d} \) at \(z = 0\) is zero, and

\[
\frac{z^{d-x}H_{x}(z)}{zH(z)} = z^{d-x-1} \sum_{s=-x+1}^{m} p_{s}(\lambda z)^{s}(zH(z)) \quad (0 < x < m),
\]

and so

\[
F_{x}(\Phi) = \delta_{x,d}.
\]

(c) \(x = 0\). Here

\[
F_{0}(\Phi) = \left( \frac{K}{K'} \right)^{\frac{1}{4}d} \frac{\sqrt{(KK')}}{4\pi i} \int_{C} \frac{(z^{d+1} - z^{d-1})}{zH(z)} dz
\]

= \delta_{0,d}.
In general (iv) of (4) can be taken to mean that the contour $C$ must enclose $z = 0$ and exclude any zeros of $H(z)$. In special cases this may be relaxed since
\[ H_x(z) z^{d-x} - H_x(z^{-1}) z^{d+x} \]
may have a factor in common with $H(z)$. Hence we have
\[ p_n(x \mid d) = \frac{1}{2\pi i} \int_C J_n(x, d; z) dz, \]
where
\[ J_n = \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \frac{H_x(z) z^{d-x} - H_x(z^{-1}) z^{d+x}}{zH(z)} k(z)^n, \]
and
\[ k(z) = k + \frac{1}{2} \sqrt{(KK')} (z + z^{-1}). \]

3. $C$ as the unit circle. The expression (6) still holds when $C$ is taken to be $|z| = 1$, provided $J_n(x, d; z)$ has no poles in or on the unit circle except $z = 0$. If $J_n(x, d; z)$ has poles at $z_j$ ($j = 1, 2, \ldots, s; z_j \neq 0$) inside the unit circle, then (6) is replaced by
\[ p_n(x \mid d) = \frac{1}{2\pi i} \int_{|z| = 1} J_n(x, d; z) dz - \sum_{j=1}^{s} \left[ \text{residue of } J_n(z) \right] z = z_j, \]
the summation excluding $z = 0$. This may be written
\[ p_n(x \mid d) = \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \int_{-\pi}^{\pi} S_x^d(\phi) S_d(\phi) (k + \sqrt{(KK') \cos \phi})^n d\phi - \sum_{j=1}^{s} \left[ \text{res } J_n(z) \right] z = z_j, \]
where
\[ S_x^d(\phi) = 3[\{H(z^{-1}) z^d \} \sin \phi + p_0 \cos \phi + \ldots + p_m \cos \phi (m < 0), \]
\[ S_d(\phi) = 3[\{H_x(z^{-1}) z^d \} \sin \phi + p_0 \cos \phi + \ldots + p_m \cos \phi (m < 0), \]
\[ \rho(\cos \phi) = |H(e^{i\phi})|^{2}, \]
\[ z = e^{i\phi}. \]

It will be observed that $S_x^d(\phi)/\sin \phi$ is a trigonometrical polynomial of degree $x$. Considered as function, of $\cos \phi$, it is of interest to note these polynomials are related to those considered by Szegö (4).† If $J_n(x, d; z)$ has a pole on the unit circle, say at $z = e^{iz}$, then the integral in (8) is to be taken as the principal value; and similarly for several such poles.

An expression for the g.f. follows from (4). In the case when $[H(z)]^{-1}$ has simple poles at $z_j$ inside the unit circle, we find
\[ \Phi(x \mid d; t) = \left( \frac{K}{K'} \right)^{\frac{1}{2}(x-d)} \int \frac{t \sqrt{(KK') \sin \theta H(e^{i\theta})}}{t \sqrt{(KK') \sin \theta H(e^{-i\theta})}} - \sum_{j=1}^{s} \frac{\Psi(x, d; z_j)}{z_j H'(z_j) [1 - k(z_j)]} \].

† Szegö considers, for example, the polynomials on the interval $[-1, +1]$ orthogonal with respect to the weight function $\rho(\cos \phi)$, where $\rho(X)$ is of precise degree $e$ and positive in $[-1, +1]$. It may be shown that there is a unique normalized representation of $\rho(\cos \phi)$, namely $|h(e^{i\phi})|^2$, such that $h(0) > 0$ and $h(z) > 0$ in $|z| < 1$. 

\[ If [H(z)]^{-1} has simple poles on the unit circle, then the only modification in (9) is to add the contributions resulting from the necessary indentations in the unit circle.
A semi-infinite random walk with discrete steps

4. Duration of the walk and probability of return

4.1. From (7) we have

$$\sum_{n=0}^{\infty} p_n(x \mid d) = \frac{1}{2 \pi i} \int_{|z|=1} 2zJ_0(x, d; z) \, dz - \sum_{j=1}^{s} \text{res} \left[ \frac{2zJ_0(x, d; z)}{(z \sqrt{K'} - \sqrt{K}) (\sqrt{K'} - z \sqrt{K})} \right]_{z=z_j}.$$  \hspace{1cm} (10)

Hence if \( K' > K \), noting that \( H(\sqrt{K'/K}) = -p_a \) we see that if \( p_a > 0 \) then \( z_j + \sqrt{(K/K')} \), and so for \( 0 \leq x \leq d \)

$$\sum_{n=0}^{\infty} p_n(x \mid d) = \frac{p_a^{-1}}{\frac{1}{2} K' - \frac{1}{2} K} \left( \frac{K}{K'} \right)^x H_x(\sqrt{K'/K}) - H_x(\sqrt{K'/K}).$$  \hspace{1cm} (11)

Hence in this case the probability that the particle will revisit the point \( x \) is

$$1 - \frac{(\frac{1}{2} K' - \frac{1}{2} K)p_a}{\left( \frac{K}{K'} \right)^x H_x(\sqrt{K'/K}) - \sum_{s=0}^{K} p_s + 1}.$$  \hspace{1cm} (12)

For if it visits a given point, that point becomes the starting point of the subsequent walk, irrespective of its previous history. If \( x > d \), taking account of the pole at \( z = 0 \), we have

$$\sum_{n=0}^{\infty} p_n(x \mid d) = \frac{p_a^{-1}}{\frac{1}{2} K' - \frac{1}{2} K} \left( \frac{K}{K'} \right)^x H_x(\sqrt{K'/K}) - \sum_{s=0}^{K} p_s + p_a \left( \frac{K}{K'} \right)^x.$$  \hspace{1cm} (13)

For completeness the following further results are noted:

$$\sum_{n=0}^{\infty} p_n(x \mid d) = \frac{(K'/K)^d}{\frac{1}{2} K - \frac{1}{2} K'} \left( \frac{K}{K'} \right)^x H_x(\sqrt{K'/K}) - \frac{K}{K'}.$$  \hspace{1cm} (14)

\((0 \leq x \leq d, K > K', p_0 = 1)\),

$$\frac{(K'/K)^d}{\frac{1}{2} K - \frac{1}{2} K'} \left( \frac{K}{K'} \right)^x H_x(\sqrt{K'/K}) - \frac{K}{K'}.$$  \hspace{1cm} (15)

\((x > d, K > K', p_0 = 1)\).

4.2. The expected number of steps to annihilation is, for \( K' > K \),

$$D(d) = 1 + p_a \sum_{s=0}^{\infty} s p_s(0 \mid d)$$

$$= 1 + p_a \left( \frac{K'/K}{K} \right)^d \frac{\pi i}{\sqrt{(K'/K)^d}} \int_0^{(z^2 - 1)^{-1/2} k(z)} \frac{dz}{zH(z)} \left[ \left( 1 - z \sqrt{K'/K} \right) - \left( 1 - z \sqrt{K/K} \right) \right]$$

$$- \sum_{j=1}^{s} \text{res} \left[ \frac{J_0(z) k(z)}{\left( 1 - k(z) \right)^2} \right]_{z=z_j}.$$  \hspace{1cm} (16)

$$= \frac{1}{p_a} + \frac{d}{K' - K} + \sum_{s=0}^{m} s p_s$$

$$= \frac{d}{p_a} + \frac{1}{\frac{1}{2} K' - \frac{1}{2} K} + \sum_{s=0}^{m} s p_s.$$  \hspace{1cm} (17)
The formula (17) gives the expected duration of the game in gambling against an infinitely rich adversary, the gambler winning at a trial with probability \( \frac{1}{2}K \), tying with probability \( k \), and losing with probability \( \frac{1}{2}K \), his initial capital being \( d \), and the stakes a unit of capital per trial. If, however, the gambler is reduced to penury, at the next trial he may be ruined with probability \( p_o \), win the right to spin again with probability \( p_0 \), or start again with capital of \( x \) units with probability \( p_x \).

5. Applications

5.1. Chance of ruin in gambling against infinitely rich adversary. Suppose the gambler’s chance of winning at any trial is \( p = \frac{1}{2}K \), and \( q = 1 - p = \frac{1}{2}K' \) is his chance of losing. Then the chance of ruin in the course of \( n \) games is given by (6) with \( p_o = 1 \),

\[
H(z) = 1 - \sqrt{(pq)(z + z^{-1})} = z^{-1} \sqrt{(pq)} \left( \sqrt{\frac{q}{p}} - z \right) \left( z - \sqrt{\frac{p}{q}} \right)
\]

and

\[
p_n(0 \mid d) = \frac{(pq)^{n+1}}{2\pi i} \left( \frac{q}{p} \right)^{1d} \int_C \frac{(z - z^{-1}) z^{d-1} (z + z^{-1})^n dz}{1 - (z + z^{-1}) \sqrt{(pq)}}
\]

where \( C \) is a contour surrounding the origin but not \( z = \sqrt{(pq)} \) or \( z = \sqrt{(q/p)} \). Hence if \( q > p \), taking \( C \) as the unit circle, and thus containing \( z = \sqrt{(pq)} \) we have (subtracting the residue at \( z = \sqrt{(pq)} \))

\[
p_n(0 \mid d) = 1 - \frac{\left( \frac{2}{\pi} \sqrt{(pq)} \right)^{n+1}}{\pi} \left( \frac{q}{p} \right)^{1d} \int_0^{\pi} \sin \phi \sin d \phi \cos^n \phi d\phi
\]

(18)

If, however, \( p > q \) we subtract the residue at \( z = \sqrt{(q/p)} \), and this leads to (18) again except that the first term is now \( (q/p)^d \). (See for example, Uspensky (5), p. 159.)

If the games are equitable, then the probability of ruin in the course of \( n \) games is

\[
\frac{1}{2\pi i} \int_C \frac{1 + z (z + z^{-1})^n}{1 - z (z + z^{-1})^n} z^d dz
\]

(19)

and using \( |z| = 1 \) indented at \( z = 1 \) as \( C \) we find

\[
p_n(0 \mid d) = 1 - \frac{1}{\pi} \int_0^{\pi} \sin d \phi \cot \left( \frac{\pi}{2} \phi \right) \cos^n \phi d\phi
\]

\[
= 1 - \frac{2}{\pi} \int_0^{\pi} \sin d \phi \cos^{n+1} \phi d\phi
\]

(20)

if \( n \) and \( d \) have the same parity. (See (5), p. 159.)

5.2. Particle with drift in the presence of a reflecting barrier. The boundary conditions in this case are \( p_0 = 0, p_1 = 1 \), so that \( H(z) = \frac{1}{2}z^{-1}(K'z^2 - K) \sqrt{(K'/K)} \) with \( k = 0 \). In (8) we have

\[
S_d(\phi) = \frac{1}{2} \sqrt{\frac{K'}{K}} \{ K' \sin (d - 1) \phi - K \sin (d + 1) \phi \},
\]

\[
S_x^2(\phi) = S_x(\phi) \quad (x > 0),
\]

\[
= -\frac{1}{2} \sqrt{(K'K)} \sin \phi \quad (x = 0),
\]
A semi-infinite random walk with discrete steps

and the residues of $J(z)$ occur at $z = \pm \sqrt{(K/K')}$. For $K' > K$. Writing $K' = 2p$, $K = 2q$, we find

$$p_n(x\mid d) = \frac{q^*}{2pq}(p - q)(q/p)^x(1 + (-1)^{x+d+n})$$

$$+ \frac{2}{\pi} q^* \left(\frac{q}{p}\right)^{1(x-d)} \int_0^{\pi} f_x(\phi)f_d(\phi) \tan^2 \phi (2 \sqrt{(pq) \cos \phi})^n d\phi,$$

(21)

where

(i) $f_x(\phi) = \cos x\phi - (p - q) \sin x\phi \cos \phi/\sin \phi$,

(ii) $q^* = q$ if $x = 0$,

$\quad = 1$ if $x > 0$,

(iii) $p > q$.

If $p < q$ the first term in (21) is zero for in this case $z = \pm \sqrt{(q/p)}$ is outside $|z| = 1$.

5.3. (i) Lauwerier (3) discusses the random walk where a particle starts at $z = 0$ in the presence of an elastic barrier at $z = m$. In our notation the boundary conditions are $p_0 = 0$, $p_1 = 1$ and $K' = K = 1$. Lauwerier does not give an expression for $p_n(x\mid d)$.

We have $H(z) = \frac{1}{2}(q-p)(m-1)$, and the zeros $z = \pm \frac{1}{\sqrt{(q-p)}}$ are outside $|z| = 1$ for $0 < q < 1$. Hence from (8) we have

$$p_n(x\mid d) = \left(\frac{1}{2q}\right)^* \frac{2}{\pi} \int_0^{\pi} \frac{(p \sin x\theta \cos \theta + q \cos x\theta \sin \theta)(p \sin d\theta \cos \theta + q \cos d\theta \sin \theta) \cos^n \theta d\theta}{p^2 \cos^2 \theta + q^2 \sin^2 \theta},$$

(22)

where

$$\left(\frac{1}{2q}\right)^* = 1 \quad (x \neq 0),$$

$$= \frac{1}{2q} \quad (x = 0).$$

The expression (22) corresponds to Lauwerier's $h_n(z)$ (see (3), p. 298), where $m - z = x$, the boundary being at $z = m$.

(ii) Lauwerier (3, p. 296) also discusses the random walk with start at $z = 0$ and a barrier at $z = m$, such that the particle arriving at the barrier may be absorbed with probability $p$, or move to $m - 1$ or $m + 1$ with probabilities $\frac{1}{q} = \frac{1}{2}(1 - p)$. In our present notation the start is at $x = d$, $K = K' = 1$, and at the barrier $p_0 = p$, $p_1 = \frac{q}{2}$. The expression for $p_n(x\mid d)$ can easily be found by using Kelvin's method of images. Thus the source at $x = d$ is equivalent to sources $\frac{1}{2}$ and $\frac{1}{2}$ at $+d$ and $-d$, and a source $\frac{1}{2}$ at $x = d$ and sink $-\frac{1}{2}$ at $x = -d$. The solution to the former is exactly the same as for a source $\frac{1}{2}$ at $x = d$ in the presence of a barrier with properties $p_0 = p$, $p_1 = q = 1 - p$, provided we double the value of the probability at the origin. The latter is equivalent to an absorbing barrier (as defined by Feller (1)) at $x = 0$ with a source $\frac{1}{2}$ at $x = d$. Hence

$$p_n(x\mid d) = \frac{1}{\pi} \int_0^{\pi} \sin x\theta \sin d\theta \cos^n \theta d\theta$$

$$+ \left(\frac{1}{q}\right)^* \frac{1}{\pi} \int_0^{\pi} \frac{(p \sin x\mid \theta \cos \theta + q \cos x\theta \sin \theta)(p \sin d\theta \cos \theta + q \cos d\theta \sin \theta) \cos^n \theta d\theta}{p^2 \cos^2 \theta + q^2 \sin^2 \theta},$$

† The expression (21) differs from that given by Kae (2) for this problem, the $\cos \phi$ factor not appearing in $f_x(\phi)$. In correspondence Mr Kae informs me that this was omitted from his article by a misprint.
where
\[
\left(\frac{1}{q}\right)^* = 1 \quad (x \neq 0),
\]
\[
= \frac{1}{q} \quad (x = 0).
\]

5.4 Some further special cases are given below:

(i) \( p_0 = 1; K = K' = 1: \)
\[
p_n(x \mid d) = \frac{2}{\pi} \int_0^\pi \sin x \phi \sin d \phi \cos^n \phi d\phi \quad (x > 0),
\]
\[
= 1 - \frac{1}{\pi} \int_0^\pi \sin d \phi \cot \left(\frac{1}{2} \phi\right) \cos^n \phi d\phi \quad (x = 0).
\] (23)

(ii) \( p_1 = 1; K = K' = 1: \)
\[
p_n(x \mid d) = \frac{2}{\pi} \int_0^\pi \cos x \phi \cos d \phi \cos^n \phi d\phi.
\] (24)

(iii) In (21) put \( x - d = x' \), and let \( x \) and \( d \to \infty \), and using well-known properties of oscillatory integrals we have
\[
p_n(x' \mid 0) = \frac{2^n}{\pi} q^{i(n-\pi)} p^{i(n-\pi)} \int_0^\pi \cos x' \phi \cos^n \phi d\phi.
\] (25)

(iv) \( p_0 = 1 - p, p_1 = p; \frac{1}{2} K = \frac{p}{2} k = 1 - 2p, \frac{1}{2} K' = p \quad (0 < p \leq \frac{1}{2}): \)
\[
p_n(x \mid d) = \frac{2}{\pi} \int_0^\pi \cos (x + \frac{1}{2} \phi) \cos (d + \frac{1}{2}) \phi (1 - 4p \sin^2 (\frac{1}{2} \phi))^n d\phi.
\] (26)

(v) \( p_0 = 1 - p, p_1 = p; \frac{1}{2} K = \frac{1}{2} K' = \frac{1}{2} p, k = 1 - p: \)
\[
p_n(x \mid d) = \frac{2^n}{\pi} \int_0^\pi \cos x \phi \cos d \phi (q + p \cos \phi)^n d\phi.
\] (27)

I am indebted to Dr J. A. Storrow of the Chemical Engineering Department, Manchester College of Technology, for drawing my attention to a problem similar to the one considered here.

REFERENCES

Distributions Associated with Random Walk and Recurrent Events

BY

J. G. SKELLAM and L. R. SHENTON

REPRINTED FROM THE JOURNAL OF THE ROYAL STATISTICAL SOCIETY
SERIES B (Methodological)
Vol. 19, No. 1, 1957
(pp. 64—118)
Distributions Associated with Random Walk and Recurrent Events

By J. G. Skellam,
The Nature Conservancy, London,

and

L. R. Shenton,
College of Science and Technology, University of Manchester

[Read before the Research Section of the Royal Statistical Society, January 23rd, 1957, Dr. H. E. Daniels in the Chair]

Introduction

1. Our concern here is with the relation between the distributions of the intervals of time (discrete or continuous) between successive events, and the number of events in a given time interval. Much work has been done on this subject for the case of continuous time, in particular with regard to industrial accidents, vehicular traffic, and the generation times of micro-organisms in the laboratory. For example, it is well known in the case of accident data and data referring to freely flowing traffic, that, at any rate approximately, the intervals between successive events follow a negative exponential distribution, while the number of events in a given interval is distributed as a Poisson variate. In discrete time there are a number of illustrations such as the distribution relating to runs of success in coin spinning, or those relating to the accumulated score in games such as billiards, table tennis, cricket and so on.

Our studies, starting in 1950, developed on rather different lines from those of Feller (1949). The majority of our results are new, and though there is some overlap, this has been included both for consistency and because we believe our formulation and methods to be of interest in themselves.

2. A stochastic process is imagined where a particle undergoes a sequence of non-negative displacements or jumps (see Fig. 1). The lengths of the jumps are regarded as independent values of the same random variable whose distribution (j) is considered as given, and is referred to as the primary distribution.

3. We seek:

(a) The expected number of landings (or events) \( \bar{n}, ndt \) at any point in the discrete case or infinitesimal interval in the continuous one.
(b) The distribution \((k)\) of the length of the interval \((O, N)\).
(c) The distribution \((k^2)\) of the length of the interval \((L, O)\).
(d) The distribution \((s)\) of the length of the interval \((L, N)\).
(e) The distribution \((u)\) of the number \((n)\) of landings in \((O, I)\) when \(B\) and \(O\) coincide.
(f) The distribution \((v)\) of the number \((n)\) of landings \((in O, I)\) when \(B\) precedes \(O\).

\(u\) and \(v\) are referred to as census distributions.

Interest is taken in the inter-relations between the theoretical properties of the various distributions and in particular the radius of convergence of their generating functions, the moments and limiting forms.

**PART 1. Discrete Distributions**

4. 1. Preliminaries on p.g.f.'s and notation

(a) Probability generating function.—For a p.g.f. we use

\[
P(0) = p(0) + p(1) + \theta^2 p(2) + \ldots,
\]

where the coefficient of \(\theta^t\) in \(P(0)\) is \(p(t)\), or the probability of a variate value of \(t\). We shall invariably use the parameters \(\theta\) and \(t\) in this respect; thus \(j(t)\) is associated with the p.g.f. \(J(\theta)\).

(b) Cumulative probability.—For the probability of a variate value greater than \(t\) \((t > 0)\) we use \(p^*(t)\), and for the probability of a variate value less than \(t\) we use \(*p(t)\). Thus

\[
p^*(t) = \sum_{s=t}^\infty p(s), \quad *p(t) = \sum_{s=0}^{t-1} p(s), \quad t = 0, 1, 2, \ldots
\]

and

\[
*p(t) + p^*(t) = 1.
\]

The g.f. \(\sum_{t=0}^\infty \theta^t p^*(t)\) is denoted by \(P^*(\theta)\).

(c) Truncated generating function.—For the positive tail of a p.g.f. we use \(P_t(\theta)\) defined by

\[
P_t(\theta) = \sum_{s=t}^\infty \theta^s p(s), \quad t = 0, 1, 2, \ldots
\]

Similarly for the complementary tail we use

\[
tP(\theta) = \sum_{s=0}^{t-1} \theta^s p(s), \quad t = 0, 1, 2, \ldots
\]

and clearly

\[
tP(\theta) + P_t(\theta) = P(\theta).
\]

It will be seen that in particular

\[
P_t(1) = p^*(t), \quad tP(1) = *p(t).
\]

(d) Reducible generating functions.—If \(P(\theta) = \sum_{t=0}^\infty \theta^t p(t)\) can be expressed as a power series in \(\varphi = \theta^l\) for any integer \(l > 1\) we say that \(P(\theta)\) is reducible. Otherwise \(P(\theta)\) is irreducible.

(e) Regular p.g.f. If the radius of convergence of a p.g.f. \(P(\theta)\) exceeds unity, and the p.g.f. is irreducible, then we shall say that \(P(\theta)\) is a regular p.g.f.

5. We now give a few elementary theorems relating to g.f.'s.

Theorem 1. The radius of convergence of the power series for \(P(\theta)\) is at least unity.

For

\[
|P(\theta)| < p(0) + p(1) + p(2) + \ldots = 1, \quad |\theta| < 1.
\]

so the series converges at all points within the unit circle.

Theorem 2. If the radius of convergence of \(P(\theta)\) is \(1 + r\) where \(r > 0\) then the moments of \(P(\theta)\) exist and are finite.

This follows from the fact that in the \(\theta\)-plane \(P(\theta)\) is analytic within \(|\theta| = 1 + r\). Hence
it is possible to define a circle \( |\theta - 1| < \varepsilon \) for some \( \varepsilon > 0 \) in which \( P(0) \) is also analytic so that in this circle the Taylor series converges, i.e.

\[
P(0) = 1 + \mu_1(0 - 1) + \mu_2(0 - 1)^2/2! + \ldots
\]

converges when \( |\theta - 1| < \varepsilon \), so that the factorial moment \( \mu_n \) of any stated order \( s \) exists and is finite. Hence the crude and central moments also exist and are finite. 

**Theorem 3.** The equation \( P(0) = 1 \) cannot have a root within the unit circle. If the radius of convergence of \( P(0) \) exceeds unity, then \( P(0) - 1 \) has a finite number of zeros in any circle lying entirely within its circle of convergence. If \( P(0) \) is a regular p.g.f. then the only zero satisfying \( |\theta| < 1 \) is \( \theta = 1 \).

The first part of the theorem follows immediately from \( |P(0)| < 1 \) (\( \rho(0) \neq 0 \)) for \( |\theta| < 1 \). For the next part we need only remark that an infinite number of zeros in any bounded region lying within the circle of convergence would imply the existence of a limit point of zeros (or singularity) and this is impossible for \( P(0) \) is analytic in its circle of convergence.

With regard to the last part of the theorem we observe first of all that \( \theta = 1 \) is clearly a simple zero of \( P(0) - 1 \). (A multiple zero at this point would imply \( \mu = 0 \), and \( P(0) \) would degenerate to the trivial distribution \( p(0) = 1 \), \( p(t) = 0 \) if \( t \neq 0 \).) Suppose in addition there is a zero at \( \theta = e^{i\psi} \), \( 0 < \psi < 2\pi \). Then, from the equations \( P(1) = P(e^{i\psi}) = 1 \), we must have

\[
p(t)(1 - \cos \psi t) = 0, \quad t = 1, 2, 3, \ldots
\]

It is not possible for \( p(t) = 0 \) unless \( \psi t \) is zero or a multiple of \( 2\pi \). Since \( \psi = 0 \), then excluding the case \( t = 0 \), it follows that \( \psi t \) is a multiple of \( 2\pi \), or that \( t\psi(2\pi) \) is an integer. Since \( t \) is itself an integer, \( \psi/(2\pi) \) must be a rational fraction, let us say \( q/l \) when expressed in its lowest terms, where \( q < l \), \( l > 1 \). Hence for \( p(t) = 0 \) \( (t = 0 \), \( t\psi/l \) is an integer and \( l \) therefore divisible by \( l \).

The whole set of non-zero frequencies can therefore be expressed in terms of \( p(s) \), \( \tau = 0, 1, 2, \ldots \) where \( \tau = t/l \), so that \( P(0) \) is therefore reducible, contrary to hypothesis. Hence a regular p.g.f. has \( \theta = 1 \) as the only root satisfying \( P(0) = 1 \), \( |\theta| < 1 \). We note a converse.

**Theorem 4.** If \( P(0) - 1 \) has a simple zero at \( \theta = 1 \) and no others on the unit circle, then \( P(0) \) is irreducible.

Assume the contrary to be true, namely that \( P(0) \) is reducible. Then

\[
P(0) = \sum_{n=0}^{\infty} \rho(n), \quad l > 1 \ (l \text{ an integer}).
\]

Hence \( P(0) = 1 \) when \( \theta = 1 \) so that \( P(0) - 1 \) has more than one zero on the unit circle, contrary to the assumption.

**Theorem 5.** The series for \( P(0) \) and \( P^*(0) \) have the same radii of convergence.

In the first place there is the relation between a g.f. \( P(0) \) and the cumulative g.f. \( P^*(0) \);

\[
(1 - 0) P^*(0) + 0 P(0) = 1, \quad \text{or}
\]

\[
P^*(0) - 1 = (1 - P(0)) 0/(1 - 0)
\]

valid for the common region of convergence of both \( P(0) \) and \( P^*(0) \).

\( P(0) \) and \( P^*(0) \) are either terminating polynomials or infinite series. In the former case the theorem is evident. In the latter case suppose \( P^*(0) \) converges for \( |\theta| < \rho \). Then

\[
(1 - (1 - 0) P^*(0))/(1 - 0)
\]

also converges for \( |\theta| < \rho \), \( \theta = 0 \) being an ordinary point of this function, and so \( P(0) \) con-

* It is of some interest to note that if in any power series \( a_n > 0 \), and \( \sum_{n=0}^{\infty} a_n x^n \) has radius of convergence equal to unity, then \( x = 1 \) is a singular point (Titchmarsh (1939), p. 214). Hence, if \( P(0) \) has unit radius of convergence, the factorial moment g.f. does not exist, and \( \theta = 1 \) is either a branch point or perhaps limit point of zeros approaching the unit circle from the outside or point where \( P(0) \) is otherwise not differentiable.
verges for \( | \theta | < \rho \). Similarly, if \( P(0) \) converges for \( | \theta | < \rho \), then, if \( \rho > 1 \), the radius of convergence of \( \{1 - \theta P(0)\}/(1 - 0) \) is also \( \rho \) since \( \theta = 1 \) is an ordinary point of this function; i.e. the radius of convergence of \( P^*(0) \) is also \( \rho \) where \( \rho > 1 \). But if \( \rho = 1 \) (it cannot be less than unity by Theorem 1) the radius of convergence of \( \{1 - \theta P(0)\}/(1 - 0) \) is also unity. Hence, if \( \rho = 1 \) or \( \rho > 1 \), the radius of convergence of \( P^*(0) \) is also \( \rho \).

It should however be noted that, although \( P(0) \) and \( P^*(0) \) have the same radius of convergence, yet \( | P(0) | < 1 \) everywhere on the unit circle, whereas a similar boundedness cannot be stated in general with regard to \( P^*(0) \).

6. The distributions* of \((\text{ON})\), \((\text{LO})\) and \((\text{LN})\)

Let \( J(\theta) \) be the p.g.f. of the “jump” distribution so that

\[
J(\theta) = j(0) + \theta j(1) + \theta^2 j(2) + \ldots,
\]

and write for the crude, factorial and central moments

\[
\mu_s = \mathcal{E}(t^s),
\]

\[
\mu_{(s)} = \mathcal{E}(t(t - 1) \ldots (t - s + 1)),
\]

\[
\mu_s = \mathcal{E}((t - \mu')^s),
\]

but in particular denote the mean by \( \mu \).

We consider a particle starting a random walk at \( B \) \((t' = 0)\), making its last “landing” in the interval \((0, H - 1)\) at \( L \) \((t' = x)\) and its next “landing” outside the interval at \( N \) \((t' = H + t)\) (see Fig. 2). (We assume \( H \geq 1 \).

![FIG. 2.](image)

Before deriving the distributions of the distances

(a) \((\text{ON})\), \hspace{1cm} (b) \((\text{LO}')\), \hspace{1cm} (c) \((\text{LN})\),

it is convenient to make some remarks on accumulated probability.

7a. Accumulated Probability

(a) The probability that the \( r \)th landing occurs at a distance \( t \) from the start is the coefficient of \( \theta^t \) in \( [J(\theta)]^r \). The expected number of landings is the accumulated probability

\[
\pi(t) = \text{coefficient of } \theta^t \text{ in } \sum_{r=1}^{\infty} [J(\theta)]^r,
\]

and

\[
\sum_{t=0}^{\infty} \pi(t) \theta^t = J(\theta)/(1 - J(\theta)).
\]

Clearly \( \pi(t) = 0 \) for \( t < 0 \) because all jumps are non-negative.

(b) It is interesting to ask what primary distribution (if any) bestows a constant value of \( \pi \)

* In the subsequent discussion we exclude the trivial P.D. for which \( j(0) = 1 \) and \( j(t) = 0 \) when \( t \neq 0 \) so that the “mean value” of a “jump” is definitely positive.
to all points \((t = 0, 1, 2, \ldots)\). The condition is clearly \(\pi(1 - 0) = J(0)/[1 - J(0)]\), giving the geometric distribution,
\[
J(0) = (1 - r)/(1 - r0),
\]
where
\[
r = 1/(1 + \pi),
\]
and \(1/\pi\) is the mean of the distribution.

(c) It is sometimes convenient to regard the source as the \(0^\text{th}\) landing with g.f. unity. With this convention the g.f. of \(\pi(t)\) is simply
\[
\Pi(\theta) = \sum_{t=0}^{\infty} \pi(t) \theta^t = 1/[1 - J(0)].
\]

(d) It is useful in the solution of problems to employ the function \(\pi(t \mid x)\) which denotes the expected number of landings at \(t\) subject to the condition that the particle has already landed at \(x\). Clearly \(\pi(t \mid x) = \pi(t - x)\), being zero if \(t < x\), whilst if \(t = x\) we adopt the convention of including the landing already at \(x\). Hence
\[
\Pi(\theta \mid x) = \sum_{t=x}^{\infty} \pi(t \mid x) \theta^t = \theta^x \sum_{t=x}^{\infty} \pi(t - x) \theta^{t-x} = \theta^x \Pi(\theta).
\]

(e) Since \(J(0)\) is analytic with modulus < 1 in \(|0| < 1\), then \(\Pi(0) = 1/[1 - J(0)]\) is analytic for \(|0| < 1\). There is always a singularity at \(0 = 1\).

(f) A special case of Cesàro's theorem (Bromwich, 1926, p. 150) on the comparison of divergent series is that if \(f(t)\) tends to a limit \(f\) as \(t \to \infty\) then \(f = \lim_{\theta \to 1} \{\sum_{t=1}^{\infty} f(t) \theta^t\}\). An application of the theorem to \(\Pi(0)\) shows that, if \(\pi(t)\) tends to a unique limit \(\pi\) as \(t \to \infty\), then
\[
\pi = \lim_{\theta \to 1} \{(1 - 0)/(1 - J(0))\}.
\]

If the distribution \(j\) has a mean \(\mu\), \(\pi = 1/\mu\), whilst if \(\mu\) is infinite \(\pi = 0\).

That \(\pi(t)\) does in fact tend to a limit when the P.D. is discrete and irreducible is implied in results due to Kolmogorov (1936) and proved by Erdös, Feller and Pollard (1949) and by Chung and Wolfowitz (1952).

7b. The Distribution of the Number of Landings at a Point

Let \(\omega(n : t)\) be the probability that a particle starting from the origin lands \(n\) times at the point \(t\), let \(\Omega(z : t) = \sum_{n=0}^{\infty} \omega(n : t) z^n\) be the corresponding p.g.f., and let \(\omega^*(n : t)\) be the probability of making \(n\) or more landings at \(t\).

A succession of landings at \(t\) may be regarded as a “success run” where the probability of continuing the run is \(j_0\) and of breaking it \((1 - j_0)\).

Clearly for \(n \geq 1\)
\[
\omega^*(n + 1 : t) = j_0 \omega^*(n : t)
\]
and
\[
\omega(n : t) = (1 - j_0) \omega^*(n : t).
\]

Hence
\[
\omega^*(n : t) = j_0^{n-1} \omega^*(1 : t)
\]
and
\[
\omega(n : t) = (1 - j_0) j_0^{n-1} \omega^*(1 : t),
\]
giving
\[
\Omega(z : t) = \omega(0 : t) + \omega^*(1 : t) (1 - j_0) z/(1 - j_0 z).
\]
The mean is then
\[
\pi(t) = \omega^*(1 : t)/(1 - j_0),
\]
allowing us to recast the p.g.f. in the form:
\[
\Omega(z : t) = 1 - \pi(t)(1 - j_0)(1 - z)/(1 - j_0 z).
\]
The conditions under which this distribution tends to a limit as $t \to \infty$ are clearly the same as for $\pi(t)$.

8. Distribution of ON

Let $k_H(t)$ be the probability that the length of ON = $t \geq 0$, and $K_H(0)$ the corresponding p.g.f.

The expected number of landings at a point $x(\geq 0)$ arising from the source at $B(t = -H)$ is $\pi(x \mid -H)$. But this is also

$$\sum_{i=0}^{\infty} k_H(t) \pi(x \mid t).$$

Hence

$$\pi(x \mid -H) = \pi(x + H) = \sum_{t=0}^{\infty} k_H(t) \pi(x - t), \quad x > 0. \quad (12)$$

The convolution on the right has g.f. $K_H(0) \Pi(0)$ [ $| 0 | < 1$], whilst the expression on the left has g.f. $\theta^{-H} \Pi(0)$ from (11). For the coefficients of $\theta^{x}$ ($x = 0, 1, 2 \ldots$ ) to be the same for both,

$$K_H(0) \Pi(0) = \theta^{-H} \Pi(0)$$

less the principal part† of its Laurent expansion. Hence

$$K_H(0) \Pi(0) = \frac{1}{2\pi i} \int_C \frac{\theta^{-H} \Pi(\xi)}{\xi - 0} d\xi \quad . \quad \quad \quad (13)$$

where $C$ is a closed contour lying inside the unit circle and enclosing the origin and the point $\xi = 0$. It follows that

$$\frac{K_H(0)}{1 - J(0)} = \frac{1}{2\pi i} \int_C \frac{\theta^{-H} d\xi}{[1 - J(\xi)][\xi - 0]} \quad . \quad \quad \quad (14)$$

An alternative form for (14) appears by writing

$$K_H(0) = \frac{1}{2\pi i} \int_C \frac{d\xi}{\xi H(\xi - 0)} - \frac{1}{2\pi i} \int_C \frac{[J(\xi) - J(0)] d\xi}{\xi H(\xi - 0)[1 - J(\xi)]} \quad . \quad \quad \quad (15)$$

in which the first integral is manifestly zero for $H > 1$, and so

$$K_H(0) = \frac{1}{2\pi i} \int_C \frac{\theta^{-H} [J(\xi) - J(0)] d\xi}{[1 - J(\xi)][\xi - 0]} \quad . \quad \quad \quad (15)$$

where $C$ is a simple closed contour within $| \xi | = 1$. (Since $J(\xi)$ is analytic within $| \xi | = 1$ the point $\xi = 0$ is now an ordinary point of the integrand, whereas in (14) it was a simple pole.)

9. Let $J(\xi)$ be a meromorphic function, and let the zeros of $J(\xi) - 1$ (in increasing order of magnitude) be $\xi_0 = 1$, $\xi_1$, $\xi_2 \ldots$ with multiplicities $\lambda_0 = 1$, $\lambda_1$, $\lambda_2 \ldots$ respectively. Let $C(R)$ be the circle $| \xi | = R > 1$, and suppose $R + | \xi_s |, s = 1, 2, \ldots$.

Now replace the contour in (14) by $C(R)$ and denote $\xi H[1 - J(\xi)][\xi - 0]$ by $1/G(\xi)$ so that

$$W(R, H) = \left| \frac{1}{2\pi i} \int_{C(R)} G(\xi) d\xi \right| < \frac{M(R)}{\theta^{-H}(R - | 0 |)}$$

where $M(R)$ is the maximum value of $| 1 - J(\xi) |^{-1}$ on $C(R)$.

† A Laurent expansion in an annulus containing $z$ and surrounding $z = a$ is of the form

$$f(z) = a_0 + a_1(z - a) + a_2(z - a)^2 \ldots + b_1(z - a)^{-1} + b_2(z - a)^{-2} \ldots$$

where $b_1(z - a)^{-1} + b_2(z - a)^{-2} + \ldots$ is called the principal part.
If a sequence of circles \( C(R_n) \), where \( R_n \to \infty \) with \( n \), can be drawn such that \( M(R_n) \) is \( O(R^{\alpha}) \), then \( W(R_n, H) \to O \) as \( n \to \infty \) for \( H > \rho \), and so in the limit

\[
\frac{1}{2\pi i} \oint G(\xi) \frac{d\xi}{\xi} + \Sigma (\text{Residues of } G(\xi) \text{ at its poles outside } C) = 0.
\]

Thus

\[
K_H(0) = [J(0) - 1] \sum_{s=0}^{\infty} \frac{a(\xi_s, H)}{(\xi_s - 1)!}.
\]  

(16)

As

\[
d(\xi_s, H) = D_{\xi_s}^{-1} \left( \frac{\xi - H(\xi - \xi_s)\xi_s}{(\xi - \theta)(1 - J(\xi))} \right) \quad \xi \to \xi_s
\]

In any event the first \( s \) terms of (16) provide an asymptotic expansion for \( H \) large, as is readily seen by considering \( W(R_n, H) \) when \( |\xi_{s-1}| < R_n < |\xi_s| \).

For the probability function of the first landing outside the interval (ON) we have

\[
k_H(t) = \sum_{s=0}^{\infty} \frac{b(\xi_s, t)}{(\xi_s - 1)!}.
\]

(17)

where

\[
b(\xi_s, t) = D_{\xi_s}^{-1} \left( \frac{(\xi - \xi_s)\xi_s J_{s-1}(\xi) \xi^{-H-1}}{J(\xi) - 1} \right) \quad \xi \to \xi_s.
\]

For example when the P.D. is the geometric \((1 - r)/(1 - r\theta)\), then from (16),

\[
K_H(0) = (1 - r)/(1 - r\theta),
\]

so that the first “landing” outside the interval is distributed in exactly the same way as the “jump” length. The geometric distribution seems to be unique among discrete distributions in this respect.

10. Limiting Form of the Distribution of the First “Landing” Outside the Interval

(a) We now consider the limiting form of \( K_H(0) \) when the length of interval \( H \) tends to infinity. It will be seen from (16) that a term in \( K_H(0) \) involves a factor \( \xi_s^{-H} \), where \( |\xi_s| > 1 \). Hence if \( J(\theta) \) is a terminating polynomial and is irreducible, then, from Theorem 3,

\[
|\xi_s| > 1, \quad s = 1, 2, \ldots
\]

and so

\[
\lim_{H \to \infty} K_H(0) = \frac{J(0) - 1}{\mu(0 - 1)}.
\]

(18a)

where \( \mu \) is finite (Theorem 2) and positive.

Similarly from (17), or from (18a) and (6),

\[
\lim_{H \to \infty} k_H(t) = J_{k+1}(1)/\mu = j_k(t + 1)/\mu,
\]

(18b)

or \( \mu k(x) \) is the probability of a value of the variate exceeding \( x \). The expressions (18a) and (18b) do in fact hold under much wider conditions. Indeed from (15) it is readily seen that

\[
K_H(0) = \text{coeff. of } \xi^H \text{ in } \frac{\xi}{1 - J(\xi)} \frac{J(\xi) - J(0)}{\xi - 0},
\]

and so by Cesàro’s theorem (§7f), provided that the limit exists,

\[
\lim_{H \to \infty} K_H(0) = \lim_{\xi \to 1} \frac{\xi(1 - \xi)}{1 - J(\xi)} \frac{J(\xi) - J(0)}{\xi - 0}
\]

\[
= \frac{J(0) - 1}{\mu(0 - 1)} = \frac{J_k(0) - 1}{\mu \theta} \quad \text{by } (6).
\]

(18z)
provided $\mu < \infty$, and then
\[
\lim_{H \to \infty} k_H(t) = j^*(t + 1)/\mu. \tag{18\%}
\]
Thus the existence of a definite form for $K_H(0)$ as $H \to \infty$ requires the finiteness of the mean of the P.D.

If, in addition to $\mu < \infty$, the variance of the P.D. is finite, then the mean of the distribution is
\[
\frac{1}{2} [j^*(t)/\mu].
\]
An idea of the form of the distribution of the first landing appears from (18%) from which it is easily seen that $k_\infty(t)$ is non-increasing for $t \geq 0$.

(b) The asymptotic form of $K_H(0)$ for the general class of P.D.'s for which all moments are infinite presents difficulties. For example, little is known of the general form of $J(0)$ for this class, especially with regard to the location of singularities. However it does seem that the behaviour of $J(0)$ in the region of $\theta = 1$ is fundamentally important.

As an illustration consider the case,
\[
J(0) = 1 - (1 - 0)^x, \quad 0 < x < 1.
\]
Then, if
\[
(1 - 0)^x = \sum_{s=0}^{\infty} a_s 0^x, \quad |0| < 1,
\]
it can be shown from (15) that
\[
K_H(0) = (1 - (1 - 0)^x) \sum_{s=0}^{H-1} a_s 0^s \left(\begin{array}{c} H-1 \\ s \end{array}\right) \tag{18c}
\]
\[
= (1 - 0)^x \sum_{s=0}^{\infty} a_{H+s} 0^s. \tag{18d}
\]
Evidently the mean and higher moments of $K_H(0)$ are infinite, like those of the P.D., and as $H \to \infty$ the value of $K_H(t)$ for fixed $t$ tends to zero.

(c) It is not difficult to show that for finite $H$ the mean of the $k_H$ distribution is finite if $\mu$ is finite. For from (15) and (45) the mean of $K_H$ is given by
\[
\tau(H) = \mu [1 + \nu(H)] - H. \tag{18e}
\]
where $\nu(H)$ is the mean of the C.D. for the interval $(0, H - 1)$ when the jump distribution is $J(0)$.

(We prove in a later section that the moments of the C.D. are finite for a finite interval.) Hence from (18e), if the mean $\mu$ is finite, then $\tau(H)$ as $H \to \infty$ behaves as $\mu \nu(H) - H$. Thus, for example, if $J(0) = 0 + (1 - 0)^{\gamma/\mu}, 1 < \gamma < 2$, with mean $\mu = 1$, it may be shown that
\[
\nu(H) \sim H + \frac{H^{2-\gamma}}{\Gamma(3 - \gamma)} + \frac{H^{2-2\gamma}}{\Gamma(4 - 2\gamma)} + \cdots, \tag{18f}
\]
whence $\tau(H)$ tends to infinity at least as rapidly as $H^{2-\gamma}/\Gamma(3 - \gamma)$.

The condition that $\kappa_\infty$ is a distribution with finite mean is easily proved from (18e) to imply that the P.D. has both first and second moments finite. The value of the mean is in fact $\frac{1}{2} j^*(t)/\mu$.

If the mean and variance of the P.D. are finite, then the variance of the $k_H$ distribution may be shown to be
\[
\tau_2 = \left[2H - \mu \nu(H)\right] \mu \nu(H) + (\mu_2 - \mu^2) \nu(H) + \mu_2 - 2\mu \sum_{s=1}^H \nu(s). \tag{18g}
\]
It is interesting to note that, if we seek a P.D. for which $K_\infty(0)$ is a repetition of itself, then such a distribution must satisfy
\[
J(0) = [J(0) - 1]/[\mu(0 - 1)],
\]
and thereby leads to a geometric P.D. with mean $\mu$.

Examples and illustrations of $K_H(0)$ are given in Table 5 and Fig. 3.
$J(\theta) \text{ and } K_H(\theta)$

<table>
<thead>
<tr>
<th>KEY</th>
<th>$J(\theta)$</th>
<th>$K(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H = 10$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$J(\theta)$ Rectangular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H = \infty$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$J(\theta)$ Triangular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H = \infty$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$J(\theta)$ Poisson</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3.
11. The Distribution of O'L

Let the distance O'L be \( y \) where BL is \( t' \). We refer to O'L as the part of LN complementary to ON and write \( k_H(y) \) for the probability that O'L has length \( y \).

Now the complement to the last "jump" in the interval will be \( y \) on those occasions when there is a "landing" at \( t' = H - 1 - y \), followed by a "jump" of \( y + 1 \) or more. Hence

\[
k_H(y) = \pi(H - y - 1) j^*(y + 1).
\]  

(19)

For the g.f. we have

\[
K_H(0) = \sum_{y=0}^{\infty} k_H(y) 0^y = \sum_{y=0}^{\infty} \pi(H - 1 - y)[j^*(y + 1) 0^y].
\]

(19a)

Since the g.f.

\[
\sum_{y=0}^{\infty} j^*(y + 1) x^y = [1 - J(x)]/[1 - x] \text{ by (6), } \quad |x| < 1,
\]

\[
\sum_{y=0}^{\infty} [j^*(y + 1) 0^y] z^y = [1 - J(0z)]/[1 - 0z], \quad |oz| < 1.
\]

The convolution on the right of (19a) is then the coefficient of \( z^H - 1 \) in

\[
\Pi(z)[1 - J(0z)]/[1 - 0z],
\]

giving

\[
K_H(0) = \text{coeff. of } z^H \text{ in } z[1 - J(0z)]/[1 - J(z)][1 - 0z].
\]

(20)

For example, if \( J(0) = (1 - r)/(1 - r0) \) or \( (1 - r)/\theta/(1 - r0) \), \( 0 < r < 1 \), then from (20)

\[
K_H(0) = \{1 - r + r(1 - \theta)(r0)^{H-1}\}/(1 - r0),
\]

and

\[
k_H(t) = (1 - r) r^t, \quad 0 < t < H - 2,
\]

\[
= r^t, \quad t = H - 1,
\]

\[
= 0, \quad t > H,
\]

(22)

which is a truncated exponential distribution.

Again if \( J(0) \) is the rectangular distribution

\[
j(t) = 1/H, \quad 0 < t < H - 1,
\]

(23)

then it may be shown from (20) that

\[
k_H(t) = \left(\frac{H}{H - 1}\right)^{H-t} \left(\frac{1}{H - t + 1}\right)\frac{1}{H^2}, \quad 0 < t < H - 1.
\]

(24)

12. Limiting Form of the Distribution of O'L

For \( H \to \infty \), we may proceed as in \( \S 10 \), and, under the assumption that the mean (\( \mu \)) of the P.D. is finite, derive

\[
K_{\infty}^*(0) = \lim_{H \to \infty} K_H(0) = \frac{J(0) - 1}{\mu(0 - 1)}.
\]

(25)

Thus the limiting forms of \( K_H(0) \) and \( K_H^*(0) \) are the same, but for finite \( H \) the two distributions are radically different, the complementary distribution having a finite range even though that of the P.D. is infinite.

13. Distribution of LN

Let \( s_H(t) \) be the probability that the "spanning jump" \( (LN) = t \). (See Fig. 2.) The "spanning jump" will be of length \( t \) on those occasions when there is a "landing" at

\[
t' = H - t, \quad H - t + 1, \ldots, H - 1,
\]


each followed by a single jump of $t$. Hence
\[ s_H(t) = j(t) \sum_{s=H-t}^{H-1} \pi(s). \] (26)

Now it is readily verified from (10) that
\[ \sum_{s=H-t}^{H-1} \pi(s) = \text{coeff. of } \theta^H \text{ in } \frac{\theta(1 - \theta)}{[1 - \theta][1 - J(\theta)]}. \]
We may express $s_H(t)$ in terms of the means $\nu(I)$ of two census distributions (see §18, iii (45)), for $\nu(I) + 1$ is the coefficient of $\theta^I$ in $\theta/[1 - (1 - \theta)(1 - J(\theta))]$.

Hence
\[ s_H(t) = \{\nu(H) - \nu(H - t)\} j(t), \quad t < H, \]
\[ = \{\nu(H) + 1\} j(t), \quad t > H. \] (27)

For the g.f. of the "spanning jump" we have
\[ S_H(0) = \text{coeff. of } z^H \text{ in } \frac{z^H}{1 - z} \left( J(0) - J(0)z \right). \] (28)

Example
If
\[ J(0) = \theta(1 - r)/(1 - r\theta), \quad 0 < r < 1, \quad s = 0 \text{ or } 1, \]
we have
\[ S_H(0) = (1 - r) \theta[1 - r + r(1 - r)(r\theta)^{H-1}]/(1 - r\theta)^2, \] (29)
leading to
\[ s_H(t) = (1 - r)^2 tr^{t-1}, \quad 0 \leq t \leq H - 1, \]
\[ = (1 - r)(H(1 - r) + r) r^{t-1}, \quad t > H, \] (30)
and the mean "spanning jump" is $(1 + r + rH)/(1 - r)$.

14. Limiting Form of the Distribution of LN

If we assume that the mean $\mu$ of the P.D. is finite, then again proceeding as in §10 we find
\[ S_\infty(0) = \lim_{H \to \infty} S_H(0) = \frac{0}{\mu} \frac{d}{d\theta} J(\theta). \] (31)

If the P.D. in addition has a finite second moment $\mu_2$, the mean "spanning jump" when $H \to \infty$ is then $\mu_2/\mu$, and this always exceeds the mean "jump" $\mu$ by $\mu_2/\mu$.

As examples we have,
\[ J(0) = e^{m(\theta-1)} (p\theta + q)^m \]
\[ S_\infty(0) = 0 e^{m(\theta-1)} (p\theta + q)^m \]
Some illustrations are given in Fig. 4.

15. Census Distribution Arising from a Fixed Origin

We now consider the distribution of the number of "landings" (the census distribution, or C.D.) in the interval BO' when the particle starts its walk at B and is to be regarded as out of the interval if it lands at I or beyond. (See Fig. 5 where there are 6 "landings".)

(It is to be noted that the starting point in the interval at B is not counted as a "landing", as for example would not be the case in counting the number of balls to reach a given score in cricket, or the number of animals required to be caught to yield a pre-assigned number of marked individuals. The modification that this difference in definition requires in the theorems we devise is slight.)
\[ J(\theta) \text{ AND } S_\infty(\theta) \]

\[ J(\theta) = 1/(2 - \theta) \]
\[ S_\infty(\theta) = \theta/(2 - \theta)^2 \]

**KEY**
- \[ J(\theta) \] (Rectangular)
- \[ S_\infty(\theta) \] (Triangular)

**FIG. 4.**

**FIG. 5.**
Let \( u(n, I) \) represent the probability of \( n \) "landings" in the interval, and \( U(z : I) \) its p.g.f. It is easily seen that
\[
 u(n + 1, I) = \sum_{t=0}^{I-1} j(t) u(n, I - t). \tag{32}
\]

For example
\[
 u(n, 1) = j_0^n j_1^*, \\
 u(n, 2) = t_0^{n-1} (j_0 j_1^* + nj_1 j_3^*), \\
 u(n, 3) = j_0^{n-2} \left( j_0^2 j_2^* + \binom{n}{1} j_0 j_1 j_3^* \right) + \binom{n}{2} j_1^2 j_1^*. \tag{33}
\]
and so on. Proceeding in this way, it will be seen that in general
\[
 u(n, I) = j_0^{n+1-I} \sum_{s=0}^{I-1} \binom{n}{s} A_{I,s}, \tag{34}
\]
where \( A_{I,s} \) is a non-negative term independent of \( n \). In particular, for large \( n \), with \( j_1 \neq 0 \),
\[
 u(n + 1, I) \sim j_0^{n+1-I} \binom{n}{I-1} j_1^{I-1} j_1^*. \tag{35}
\]

16. The g.f. for the Census Distribution

Introducing the generating function
\[
 H(n : 0) = \sum_{t=1}^{\infty} u(n, I) \theta^t, \quad n = 0, 1, 2, \ldots, \tag{36}
\]
it follows from (32) that
\[
 H(n + 1 : 0) = J(0)^* H(n : 0) \text{ for } |\theta| < 1.
\]
Hence
\[
 H(n : 0) = [J(0)]^n H(0 : 0), \quad |\theta| < 1,
\]
where
\[
 H(0 : 0) = \sum_{t=1}^{\infty} j_t^* \theta^t = J^*(0) - 1 = \theta [1 - J(0)] / (1 - \theta) \tag{37}
\]
by (6).

Thus we have for the probability of \( n \) "landings"
\[
 u(n, I) = \text{coeff. of } \theta^t \text{ in } \theta [J(0)]^n [1 - J(0)] / (1 - \theta), \tag{38}
\]
and for the p.g.f. of the C.D., with \( |zJ(0)| < 1 \),
\[
 U(z : I) = \text{coeff. of } \theta^t \text{ in } \theta \left( \frac{1}{1 - zJ(0)} \right) \frac{1 - J(0)}{1 - \theta}. \tag{39}
\]
For the "bivariate" generating function we have
\[
 U[z, 0] = \sum_{t=1}^{\infty} U(z : I) \theta^t \\
 = \frac{\theta}{1 - zJ(0)} \frac{1 - J(0)}{1 - \theta}, \quad |\theta| < 1, \quad |zJ(0)| < 1. \tag{39a}
\]

(A similar result, allowing for slight notational differences, is given by Feller (1949), (p. 110, 6.4).

The g.f. (39) may be written
\[
 U[z, 0] = \frac{\theta}{z(1 - \theta)} + \frac{z - 1}{z} \cdot \frac{\theta}{1 - \theta} \cdot \frac{1}{1 - zJ(0)}, \tag{39a}
\]
and writing \( J(0) = j_0 + j_1 \varphi(0) \), where \( | \varphi(0) | < 1 \) for \( | 0 | < 1 \),

\[
U[z, 0] = \frac{1}{z(1 - 0)} + \frac{(z - 1)}{z} \cdot \frac{0}{1 - 0} \cdot \frac{1}{1 - zj_0} \sum_{s=0}^{\infty} \left( \frac{z \varphi}{1 - zj_0} \right)^s.
\]

Hence

\[
U(z : I) = \frac{1 - j_0}{1 - zj_0} + (z - 1) \sum_{s=1}^{I-1} \frac{z^{s+1} A_s}{(1 - zj_0)^{s+1}},
\]

where the \( A_s \)'s are independent of \( z \), showing that the g.f. of the C.D. may always be expressed as the sum of \( I \) generating functions of Pascal type. The p.g.f. for the C.D. may also be expressed as a contour integral, for, from Theorem 1, \( J(0) \) is regular within \( | 0 | = 1 \), and hence

\[
U(z : I) = \frac{1}{2\pi i} \int_{C} \frac{0^{-I}}{1 - z J(0)} \frac{1 - J(0)}{1 - 0} d\theta,
\]

where \( C \) is a closed contour within the unit circle surrounding the origin, and where \( | z | < 1 \). We show later (§18 (iii)) that (40) remains valid for \( | z | = 1/j_0 \).

18. Properties of \( u(n, I) \) and \( U(z : I) \)

(i) Cumulative distribution.—From (37) we have, for the probability of \( n \) or more “landings” in the period,

\[
u^n(n, I) = \text{coeff. of } \theta^n \text{ in } 0[J(0)]^n/(1 - 0).
\]

Alternatively there is the contour integral

\[
u^n(n, I) = \frac{1}{2\pi i} \int_{C} \frac{0^{-I} [J(0)]^n}{1 - \theta} d\theta.
\]

(ii) Radius of Convergence (\( r \) of \( U(z : I) \))

Since \( U(z : I) \) is a p.g.f., its power series certainly converges for \( | z | < 1 \). Thus its radius of convergence is unity at least, and there is interest in deciding whether there are any singularities on the unit circle. (If there are not, all the moments of \( U(z : I) \) exist.) It is easily shown that in fact \( r = 1/j_0 > 1 \) (the case \( j_0 = 1 \) being excluded from our discussion as trivial). The result is immediately obvious for \( j_0 = 0 \), when the number of “landings” in the interval cannot be infinite, so that \( U(z : I) \) is a terminating polynomial. To prove \( r = 1/j_0 \) in general, we see that \( U[z, 0] \) is convergent in \( z \) and \( \theta \) for any \( | 0 | < 1 \) and the restriction on \( z \) implied in \( | z J(0) | < 1 \). But the convergence of \( U(z : I) \) as a function of \( z \) is independent of \( \theta \), and so it converges for \( | z/j_0 | < 1 \) and \( r = 1/j_0 \). We could also have arrived at this result from the expansion (39b).

(iii) Moments

Using \( \psi(n) \) for the \( s \)th factorial moment of the C.D. which exist from (ii) and Theorem 2, we have from (40)

\[
\psi(n) = \frac{s!}{2\pi i} \int_{C} \frac{0^{-I}}{1 - \theta} \left( \frac{J(0)}{1 - J(0)} \right)^s d\theta.
\]

In particular, with \( \psi \) for the mean,

\[
1 + \psi = \frac{1}{2\pi i} \int_{C} \frac{0^{-I} d\theta}{(1 - 0)[1 - J(0)]} = \sum_{n=0}^{I-1} \pi(n) \text{ [by 10]},
\]

and

\[
\psi'_2 = \frac{1}{2\pi i} \int_{C} \frac{0^{-I} J(0)[1 + J(0)] d\theta}{(1 - 0)[1 - J(0)]^2}.
\]
We shall show later (in §24) that, if \( J(0) \) is a regular P.D., then there are the approximate formulae:

\[
\nu = \frac{I}{\mu} + \frac{(\mu_2 - 2\mu^2)}{2\mu^2} + 0(1 + \varepsilon)^{-1}, \quad \ldots \quad \ldots \quad \ldots \quad (47)
\]

\[
\nu' = \frac{I(I + 1)}{\mu} + \frac{(2\mu_2 - 3\mu^2)}{\mu^3} + \frac{\left(\frac{\mu^4 - 3\mu_2^2 \mu_3}{2} - \frac{2}{3} \mu^2 \mu_3 \mu_3 + \frac{3}{2} \mu_2^2 \right)}{\mu^4} + 0(1 + \varepsilon)^{-1}, \quad (48)
\]

where the value of \( \varepsilon > 0 \) depends upon the magnitude of the first zero(s) of \( J(0) - 1 \) outside the unit circle. Feller (1949) has given expressions similar to these.

(iv) It is easily shown from (45) that for all P.D. \( \nu_1 = \nu_1(I) \) is an increasing function of \( I \).

For

\[
\nu_1(I) - \nu_1(I - 1) = \pi(I - 1) \to 0.
\]

(v) Effect of form of \( j(t) \) on \( u(n, I) \).

(a) As a crude description of the C.D., we remark that if \( j(0) \) is small the right-hand tail of the C.D. is relatively unimportant, whereas if \( j(0) \) is near to unity this tail of the C.D. becomes the dominant feature.

(b) If the range of \( j(t) \) is unlimited in the positive direction, then the probability of no "landings" is not zero.

(c) If \( j(t) = 0, t < k, k > 0, j(k) = 0 \), then the maximum number of landings possible in the interval cannot exceed \( (I - 1)/k \).

(d) If \( J(0) \) is not in reduced form, the C.D. can always be found from the reduced form of the P.D. For example, the C.D. \( U_{1s}(z : I) \) corresponding to \( J_{s}(0) = (1 + 0^s + 0^s)/3 \) is related to \( U_{3s}(z : I) \) corresponding to \( J_{3s}(0) = (1 + 0 + 0^3)/3 \) by \( U_{3s}(z : I^*) = U_{3s}(z : I) \) where

\[
I^* = 3I, 3I - 1, 3I - 2.
\]

(vi) Determinativeness of the C.D.

For a given P.D., the C.D. is uniquely determined by (38). But in general \( U(z : I) \) does not uniquely determine \( J(0) \). For example the C.D. corresponding to

\[
J(0) = \sum_{t=0}^{I-1} j(t) \theta^{t} + \theta^{I} z(0)
\]

is independent of the form of the polynomial \( z(0) \).

The common practice of studying natural processes by observing the census distribution with fixed observational interval therefore involves a loss of "information" in so far as it throws no light on the nature of the P.D. for values of the variate greater than the interval length.

Nevertheless the P.D. is uniquely determined by a complete set of values of \( u(n, I) \) for any particular fixed \( n \), as can be seen from (33, 37). For example, taking \( n = 0 \), \( u(0, I) = j^*(I) \), so that the cumulative function of the P.D. is immediately determined. It will also be apparent that sets of values of \( u(n, I) \) cannot be chosen arbitrarily; for example, in the case when \( n = 0 \), \( u(0, I) \) must be a non-increasing function of \( I \) in order that the P.D. be admissible.

19. Illustrations of C.D.'s

A simple example is the geometric type distribution

\[
J(0) = (1 - r) \theta^{s}/(1 - r^{s}), \quad s = 0, 1, 2 \ldots ; 0 < r < 1.
\]

Here \( U(z : I) \) is the coefficient of \( \theta^{t} \) in

\[
\frac{0}{1 - \theta}, \quad 1 - \theta^{I} = (1 - r^{s}) \theta^{s}
\]
(a) Rectangular

(b) Rectangular

(c) Bernoullian

(d) Poisson

(e) Triangular

FIG. 6.
When \( s = 0 \) we find \( U(z : I) = [r/(1 - (1 - r) z)]^r \) which is a Pascal distribution with mean \( (1 - r)/r \) and variance \( (1 - r)/r^2 \). Similarly, when \( s = 1 \), \( U(z, I) = [r + (1 - r) z]^{r-1}, \) a Bernoullian distribution with mean \( (1 - I)(1 - r) \) and variance \( (I - 1) r(1 - r) \). The fundamental change in the form of the two C.D.'s is embedded in the values of \( j_0 \) (see 18 (v) (c)).

A number of additional examples is given in Table 4.

An impression of the relative shapes of the P.D. and the corresponding C.D. appears in Fig. 6. In (a) and (c) the P.D.'s are simple polynomials and the C.D.'s are incipiently normal. By contrast the C.D.'s for (b) and (e) are markedly skew, even though \( I \) is not small. In fact (b) and (e) in Fig. 6 are exceptional in not being asymptotically normal.

20. Recurrence Formula for Setting up \( u(n, I) \)

Expression (32) may be used with advantage to evaluate (numerically) terms of the C.D. for moderate values of \( I \), especially when \( J(0) \) is a polynomial. First we set up the C.D. for \( I = 1 \) using

\[
u(n, 1) = [1 - j(0)](j(0))^n, \quad n = 0, 1 \ldots \ldots \ldots \ldots \ldots \ldots (50)
\]

We pass on to \( I = 2 \) using

\[
u(n + 1, 2) = j(0) u(n, 2) + j(1) u(n, 1),
\]

with

\[
u(0, 2) = 1 - j(0) - j(1) = j^*(2). \ldots \ldots \ldots \ldots \ldots \ldots (51)
\]

Next use

\[
u(n + 1, 3) = j(0) u(n, 3) + j(1) u(n, 2) + j(2) u(n, 1)
\]

with

\[
u(0, 3) = j^*(3), \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (52)
\]

and so on.

At first sight, from a practical point of view, the scheme appears to be adequate for finding the C.D. for values of \( I \) in the region of 10-20. However the number of digits held increases at each step, necessitating rounding-off, which in turn introduces an accumulated round-off error. From a very limited experience of this type of process it appears that the round-off error may be quite serious, even at an early stage.

21. Our first illustration is concerned with the number of balls required to reach a given score (or to score a given number of runs off a bowler) in first class cricket. The data (Table I) was extracted from detailed records lent to us by Mr. A. Wrigley, and covers four Test Match series for England v. (i) W. Indies, 1950; (ii) India, 1952; (iii) Australia, 1953; (iv) Australia, 1956.

Table 1

<table>
<thead>
<tr>
<th>Teams</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
<th>Mean Score per Ball</th>
<th>Wickets</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5,393</td>
<td>0.2890</td>
<td>80</td>
</tr>
<tr>
<td>E.</td>
<td>5,899</td>
<td>595</td>
<td>149</td>
<td>60</td>
<td>230</td>
<td>1</td>
<td>1</td>
<td>6,935</td>
<td>0.4214</td>
<td>56</td>
</tr>
<tr>
<td>W.I.</td>
<td>4,192</td>
<td>616</td>
<td>176</td>
<td>55</td>
<td>269</td>
<td>0</td>
<td>5</td>
<td>5,313</td>
<td>0.3623</td>
<td>40</td>
</tr>
<tr>
<td>(ii)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5,407</td>
<td>0.3283</td>
<td>70</td>
</tr>
<tr>
<td>E.</td>
<td>3,723</td>
<td>521</td>
<td>146</td>
<td>36</td>
<td>176</td>
<td>3</td>
<td>5</td>
<td>4,610</td>
<td>0.4032</td>
<td>70</td>
</tr>
<tr>
<td>I.</td>
<td>2,440</td>
<td>299</td>
<td>133</td>
<td>26</td>
<td>143</td>
<td>0</td>
<td>2</td>
<td>3,043</td>
<td>0.4387</td>
<td>70</td>
</tr>
<tr>
<td>(iii)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5,407</td>
<td>0.3283</td>
<td>70</td>
</tr>
<tr>
<td>E.</td>
<td>4,923</td>
<td>632</td>
<td>151</td>
<td>40</td>
<td>222</td>
<td>0</td>
<td>3</td>
<td>5,971</td>
<td>0.3902</td>
<td>74</td>
</tr>
<tr>
<td>A.</td>
<td>3,920</td>
<td>589</td>
<td>187</td>
<td>52</td>
<td>255</td>
<td>0</td>
<td>10</td>
<td>5,013</td>
<td>0.4387</td>
<td>72</td>
</tr>
<tr>
<td>(iv)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5,407</td>
<td>0.3283</td>
<td>70</td>
</tr>
<tr>
<td>E.</td>
<td>4,059</td>
<td>556</td>
<td>189</td>
<td>57</td>
<td>195</td>
<td>2</td>
<td>3</td>
<td>5,618</td>
<td>0.3902</td>
<td>74</td>
</tr>
<tr>
<td>A.</td>
<td>4,627</td>
<td>428</td>
<td>135</td>
<td>47</td>
<td>159</td>
<td>1</td>
<td>10</td>
<td>5,407</td>
<td>0.2979</td>
<td>77</td>
</tr>
<tr>
<td>Total</td>
<td>33,783</td>
<td>4,236</td>
<td>1,266</td>
<td>373</td>
<td>1,649</td>
<td>7</td>
<td>39</td>
<td>4,1353</td>
<td>0.3567</td>
<td>559</td>
</tr>
</tbody>
</table>


Ignoring the risk of being out from a ball, the P.D. is \( 0.816,942 + 0.102,4350 + 0.030,6140^3 + 0.009,0200^3 + 0.039,8706^4 + 0.000,1690^5 + 0.000,9430^6 \) \ldots \ldots \ldots \ldots \ldots \ldots . . . (53)
The form of the C.D. for \( I = 5 \) (based on the statistics for the series (i)-(iii)) is shown in Fig. 7.

Using (47), (48) and (53a), the first two moments of the C.D. are approximately

\[
\begin{align*}
v &= 2.8032I + 1.4578, \\
v^2 &= 18.8343I + 11.0852,
\end{align*}
\]

and a few values are given in Table 2A.

<table>
<thead>
<tr>
<th>( I )</th>
<th>10</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v_1 )</td>
<td>30</td>
<td>143</td>
<td>283</td>
</tr>
<tr>
<td>( v_2 )</td>
<td>199</td>
<td>953</td>
<td>1,895</td>
</tr>
<tr>
<td>( \sqrt{v_2} )</td>
<td>14.1</td>
<td>30.9</td>
<td>43.5</td>
</tr>
</tbody>
</table>

\( v^* \) is one added to \( v \).

If we assume the C.D. is normal for say \( I = 100 \), then it would be fairly safe to say that the number of balls required to score a hundred runs lies between 152 and 414 \((v^* \pm 3\sqrt{v_2})\); or if we allow 18 overs to an hour, then 100 runs would be expected to take between 1 hour 24 mins. and 3 hours 50 mins.

To see how the values in Table 2A turn out in practice, we extracted from the data of Table 1 the details of

\( a \) all those innings in which a player scored 50 or more,

\( b \) all those occasions on which a bowler had 50 or more runs scored off him in an innings (irrespective of whether this occurred in one particular spell of bowling or not).

| \( v^* \) | 127 | 137 |
| \( \sqrt{v_2} \) | 41 | 39 |
| Number of players | 98 | 122 |

We denote the number of balls for a batsman to score the first 50 runs (or more) by \( B_a \); more precisely, at the \((B_a - 1)\)th ball a player’s score is 49 or less, while at the \( B_a \)th ball it is 50 or more.

![Fig. 7.](image-url)
Similarly we denote the number of balls bowled by a bowler for the first 50 (or more) runs to be scored from him by \( B_n \). In Table 2a are given the means and standard deviation of \( B_n \) and \( B_w \) calculated from the data, comprising 98 and 122 observations respectively. Comparing with the theoretical values (i) in column four, derived from the P.D. (53), it will be seen that the observed s.d.'s exceed expectation, as might be anticipated on the basis that the data is heterogeneous (for example, variations introduced by teams, grounds, weather conditions, and so on). With regard to the means, that for bowlers is in good agreement with the theoretical (i), while there is a considerable discrepancy in the case of batsmen.

In this connection it may be remarked that one would expect bowling performances in Tests to be fairly consistent, for an “expensive” bowler is usually quickly taken off. Moreover, only those players bowl who are specialists. On the other hand, all players are in general called upon to bat, and it is likely that the mean score per ball for those in the lower order of batting is less than for the accredited batsmen. This would have the effect of giving a higher theoretical mean value of \( B_n \), when derived from a P.D. based on the performance of all players, than would be the case if based on the P.D. for the batsmen in form. Using a P.D. calculated from the completed innings of those players who scored 50 or more, it turns out that the mean number of balls to score 50 is 128, and the observed value 127 is now in good agreement.

22. Our second illustration concerns a P.D. which has an infinite number of terms; even so it turns out that there is a simple recurrence scheme for setting up the C.D. Suppose a particle executes a random walk in the presence of a reflecting barrier. The walker, being at \( x = y, y = 0 \), moves to \( y - 1 \) or \( y + 1 \) with equal probability. If it reaches \( x = 0 \), the next move is to \( x = 1 \) with certainty. The walk starts at \( x = 0 \). It is required to find the distribution of the number of returns to the origin in \( 2I - 1 \) moves. A related problem, that of the return to equilibrium in a coin tossing scheme, has been considered by Lévy (1939, p. 301) and by Feller (1949).

There is no difficulty in showing that the probability the walker returns to the origin for the first time in \( 2n \) moves is the coefficient of \( z^n \) in \( 1 - \sqrt{1 - z^2} \). Thus the C.D. is the distribution of the number of returns to the origin, and is given by

\[
\sum_{I=1}^{\infty} U(z : I) 0^I \equiv U[z, 0] = \frac{0\sqrt{1 - 0^2}}{(1 - 0)(1 - z + z\sqrt{1 - 0^2})}. \quad (54)
\]

This may be written in the form

\[
\frac{0(1 + 0)(1 - z - z\sqrt{1 - 0^2})}{\sqrt{1 - 0^2}} = (1 - 2z + z^20^2) \sum_{I=1}^{\infty} U(z : I) 0^I, \quad (54a)
\]

so that

\[
(1 - 2z) U(z : 2I) + z^2U(z : 2I - 2) = (1 - z) b_{I-1}, \quad I > 1, \quad (54b)
\]

where

\[
(1 - 0)^{-1} = \Sigma b_0 0^0, \quad U(z, 2) = 1.
\]

After equating coefficients of powers of \( z \), we are led to:

\[
\begin{align*}
&u(n, 2I) = 2u(n - 1, 2I) - u(n - 2, 2I - 2), \quad I > 1, \quad n = 2, 3, \ldots \\
&u(0, 2I) = u(1, 2I) = b_{I-1}, \\
&u(0, 2) = 1.
\end{align*}
\]

(55)

The expressions in (55) may be used to set up \( u(n, 2I) \), noting that \( b_0 = (2s - 1) b_{s-1}/s, b_s = 1 \). In this way the values in Table 3 were obtained.
Random Walk and Recurrent Events

Table 3

Values of \( u(n, 2I) \) for the P.D. \( J(0) = 1 - \sqrt{1 - \theta^2} \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( 2I = 6 )</th>
<th>( 2I = 12 )</th>
<th>( 2I = 18 )</th>
<th>( 2I = 24 )</th>
<th>( 2I = 30 )</th>
<th>( 2I = 36 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.375</td>
<td>0.2461</td>
<td>0.1964</td>
<td>0.1682</td>
<td>0.1494</td>
<td>0.13583</td>
</tr>
<tr>
<td>1</td>
<td>0.375</td>
<td>0.375</td>
<td>0.1964</td>
<td>0.1682</td>
<td>0.1494</td>
<td>0.13583</td>
</tr>
<tr>
<td>2</td>
<td>0.250</td>
<td>0.2187</td>
<td>0.1833</td>
<td>0.1602</td>
<td>0.1439</td>
<td>0.13172</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0.1641</td>
<td>0.1571</td>
<td>0.1442</td>
<td>0.1328</td>
<td>0.12349</td>
</tr>
<tr>
<td>4</td>
<td>0.0938</td>
<td>0.1208</td>
<td>0.1214</td>
<td>0.1169</td>
<td>0.11154</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0</td>
<td>0.0312</td>
<td>0.0806</td>
<td>0.0944</td>
<td>0.0974</td>
<td>0.09666</td>
</tr>
<tr>
<td>6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0439</td>
<td>0.0667</td>
<td>0.0762</td>
<td>0.08000</td>
</tr>
<tr>
<td>7</td>
<td>0.0176</td>
<td>0.0417</td>
<td>0.0554</td>
<td>0.06286</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.0039</td>
<td>0.0222</td>
<td>0.0370</td>
<td>0.04656</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.0</td>
<td>0.0095</td>
<td>0.0222</td>
<td>0.03222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.0029</td>
<td>0.0117</td>
<td>0.0263</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.0005</td>
<td>0.0052</td>
<td>0.01203</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.0</td>
<td>0.0018</td>
<td>0.00628</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.0</td>
<td>0.0005</td>
<td>0.00285</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.0</td>
<td>0.0001</td>
<td>0.00109</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00033</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are a number of interesting features about this example. Clearly from the table the C.D. shows no tendency to normality, even when \( 2I = 36 \). Again the radius of convergence of \( J(0) \) is unity (as is also the case for the reduced form) so the asymptotic expansions for the moments in (47, 48) do not hold (obviously the moments of \( J(0) \) are infinite). The moments however may be derived from first principles and are:

\[
v_1 = \frac{\chi - 1}{2} \sqrt{\frac{I}{\pi}},
\]

where \( \chi = \frac{2I(I + \frac{1}{2})}{I^2 I_{\sqrt{\pi}}} \)

and the interval is \( 2I \),

\[
v_2 = 2I - \chi - \chi^2 \sim 2I \left(1 - \frac{2}{\pi}\right) - 2 \sqrt{\frac{I}{\pi}},
\]

\[
v_3 = 2\chi^3 + 3\chi^2 - (2I - 3)\chi - 6I
\]

\[
\sim (16 - 4\pi) \left(\frac{I}{\pi}\right)^{\frac{3}{2}} + (12 - 6\pi) \frac{I}{\pi} + 6 \sqrt{\frac{I}{\pi}},
\]

\[
v_4 = 12I^2 + 26I - 13\chi - (4I + 12)\chi^2 - 6\chi^3 - 3\chi^4
\]

\[
\sim (12\pi^2 - 16\pi - 48) \frac{I^2}{\pi} - 48 \left(\frac{I}{\pi}\right)^{\frac{3}{2}} + (26\pi - 48) \frac{I}{\pi} - 26 \sqrt{\frac{I}{\pi}}.
\]  \( (56) \)

The measures of skewness and kurtosis of the C.D. have the asymptotic values,

\[
\beta_1 = \frac{2(4 - \pi^2)}{(\pi - 2)^3}, \quad \beta_2 = \frac{3\pi^2 - 4\pi - 12}{(\pi - 2)^3}.
\]  \( (57) \)

These are exactly the values for a normal distribution, with standard deviation \( \sqrt{2I} \), truncated at its mean. It can be shown that in fact the C.D. is asymptotically seminormal (see the next section). We may add that the approach to seminormality is not rapid, as is suggested for example by the term \( 2 \sqrt{\frac{I}{\pi}} \) in \( v_2 \).
23. Approach to Normality of the C.D.

We now indicate the form the C.D. takes when \( I \) becomes large, starting with (40) in the form

\[
U(e^{-a}, I ; C) = \frac{e^a}{2\pi i} \int_C \frac{0 - I}{e^a - J(0)} \frac{1 - J(0)}{1 - 6} d\theta, \quad \ldots \quad (58)
\]

where \( C \) is a closed contour within the unit circle containing \( \theta = 0 \) and where \( a > 0 \).

Our aim is to approximate to (58) by replacing the contour \( C \) by a new one containing the unit circle (assuming this to be allowable) and adjusting the resulting expression for any poles of \([e^a - J(0)]^{-1}\) thus introduced. The zeros of \( J(0) - e^a \) (for small \( a \)) and in particular those of \( J(0) - 1 \) are thus fundamental to the discussion. Some typical examples of the location of the zeros of the latter are shown in Fig. 8. Those for the Bernoullian, Poisson and Pascal distributions have an interesting feature in common; the zeros all lie on circles touching the unit circle at \( \theta = 1 \), of radii \( \frac{1}{p}, \infty, \lambda \) respectively.

24. Let \( J(0) \), independent of \( I \), be a regular p.g.f. with radius of convergence

\[
| 0 | = 1 + r \quad (r > 0).
\]

It can then be proved easily, by making use of continuity considerations (see e.g., Weber (1912) I, p. 148), that a circle \( | 0 | = 1 + r, \eta > 0, \eta = \eta(\theta_0) \), can always be drawn within the circle of convergence such that for sufficiently small \( \theta_0 \) no zero whatsoever of \( e^a - J(0) \) lies on it and only that zero \( \theta_0(x) \) which equals 1 when \( a = 0 \) lies within it. Taking this circle as a new contour \( C' \) we have

\[
U(e^{-a}, I ; C) = U(e^{-a}, I ; C') - \text{Residue at } \theta_0(x).
\]

Moreover on \( C' \), \( |e^a - J(0)| \) is bounded, and so

\[
U(e^{-a}, I ; C') = 0[1 + r(\theta_0)]^{-1}.
\]

Hence we have

**Theorem 6.**—If \( J(0) \) is a regular p.g.f. (independent of \( I \)), then*

\[
U(e^{-a}, I ; C') \sim e^a \frac{1 - \xi(\xi)}{1 - \xi} \cdot \xi^{-1}
\]

for \( a > 0 \) sufficiently small, and \( \xi \) the root of \( J(0) = e^a \) which \( \to 1 \) as \( a \to + 0 \).

We now require an expansion for \( \theta_0(x) \) which is a root of

\[
0 = 1 + (e^a - 1) \Phi(0),
\]

where

\[
\Phi(0) = (1 - 0)[1 - J(0)]. \ldots \ldots \ldots \ldots (59)
\]

Lagrange's expansion for the root of an equation (see e.g. Whittaker and Watson (1920), p. 133) is now applicable, the expansion converging for \( |e^a - 1| < |1 - J(0)| \) when \( \theta \) is on the circle \( | \theta - 1 | = \eta', 0 < \eta' < \eta(\theta_0) \). Hence if

\[
\min_{| \theta - 1 | = \eta'} | 1 - J(0) | = m(\theta_0),
\]

then

\[
\theta_0(x) = 1 + \sum_{n=1}^{\infty} \frac{(e^a - 1)^n}{n!} \frac{d^{n-1}}{d\theta^{n-1}} \left( \frac{1 - \theta}{1 - J(0)} \right)^n \bigg|_{\theta = 0} \ldots \ldots \ldots (60)
\]

* The prime in the expression denotes a derivative.
Fig. 8.—Zeros of $1 - J(0)$ for (a) Bernoullian $(q + p\lambda)^N$, (b) Poisson $e^{\mu(\theta-1)}$, (c) Pascal $[\lambda/(\lambda + 1 - 0)]^\lambda$, (d) Polynomial (53).
for \( |e^a - 1| < m(z_0) \), or \( |a| \) sufficiently small. In terms of the factorial moments of \( J(0) \) (which exist and are finite from Theorem 2), with \( \mu \neq 0 \) we have

\[
\theta_0(x) = 1 + \frac{e^a - 1}{\mu} - \frac{(e^a - 1)^2}{2\mu^3} \frac{\mu(a)}{2} - \frac{(e^a - 1)^3}{6\mu^5} \frac{\mu(a)}{3} + \ldots
\]

\[
= 1 + \frac{a}{\mu} - \frac{z^2(\mu(a) - \mu^2)}{2\mu^3} + o(z^2). \tag{61}
\]

Using (61) we derive (62a), and similarly (62b) and (62c):

\[
- \log \theta_0(x) = - \frac{a}{\mu} + \frac{\mu a z^2}{2\mu^3} + o(z^2), \tag{62a}
\]

\[
J'[\theta_0(x)] = \mu + \frac{\mu(\mu)}{\mu} + \frac{\mu(a)}{6\mu^3} - \frac{(\mu(a) - \mu)}{4\mu^2} \frac{a^2}{\mu} + o(a^2), \tag{62b}
\]

\[
\frac{1 - J[\theta_0(x)]}{1 - \theta_0(x)} = \mu + \frac{\mu(\mu)}{2\mu^3} + \frac{\mu(a)}{6\mu^3} - \frac{(\mu(a) - \mu)}{4\mu^2} \frac{a^2}{\mu} + o(a^2). \tag{62c}
\]

Hence finally we have the asymptotic expression for the Laplace transform of the C.D.,

\[
U(e^{-a}, I) \sim (1 + b_1 x + b_2 x^2) \exp(a_1 x + a_2 x^2), \tag{63}
\]

where

\[
a_1 = - \frac{l}{\mu} + 1, \quad a_2 = \frac{\mu I}{2\mu^3};
\]

\[
b_1 = - \frac{\mu(\mu)}{2\mu^3}, \quad b_2 = \frac{3(\mu(\mu)^2 - \mu(a))}{4\mu - \mu(a) - 3\mu^3}.
\]

so that, for a regular P.D. (independent of \( I \)) with radius of convergence exceeding unity, the C.D. is asymptotically normal with mean \( l/\mu \) and variance \( \mu I/\mu^3 \). The expressions for the moments given in (47, 48) arise from (63) as coefficients of powers of \((-a)\). A sharper approximation to \( U(z : I) \) could be found by including in (63) the dominant part of \( U(z : I ; C') \), extending the contour \( C' \) to the circle \( C'' \) which just includes \( \theta_0(z) \) the next root; similarly \( C'' \) could be extended, assuming of course that each new contour could be confined to the circle of convergence of \( J(0) \).

In general one would expect the approximation at each stage to lead to an improvement.

25. It is now of interest to examine the examples in §19 for asymptotic normality. For the geometric type P.D. \( J(0) = (1 - r) \theta/((1 - r)^2), \), with \( 0 < r < 1 \), which is reduced, the circle of convergence has radius \( 1/r > 1 \), so normality is assured. Similarly for examples i, vii, viii, ix and x of Table 4, the radius of convergence in each case exceeds unity, and normality again is ultimately reached. For the rest of the examples in Table 4 the P.D. depends upon \( I \), and as \( I \) increases, zeros of \( J(0) \) accumulate on the unit circle, which ultimately becomes a line of singularities. The method of this section cannot be applied in these cases, and in fact it is easy to see that the C.D.'s are not asymptotically normal. Indeed when \( J(0) = (1 - 0^T)/[(1 - 0)] \) the C.D. has the limiting form \( U(z : I) \sim z^{-1} + e^2(1 - z^{-1}) \) with moments \( \nu_1 \sim e - 1, \nu_2 \sim e(3 - e) \), not increasing with \( I \); similarly, when

\[
J(0) = \frac{2(1 + 2b + 3b^2 + \ldots + b^{l-1})}{I(I + 1)},
\]

\[
U(z : I) \sim z^{-1} + (1 - z^{-1}) \cosh(2z).
\]

Again, when \( J(0) = 1 - \sqrt{(1 - 0)} \), the reduced form of the example discussed in §22, it may be shown that

\[
U(e^{-a} : I) \sim e^{-a \cdot b^2} \left[ 1 + \frac{1/2}{\pi} \int_0^{\frac{b}{\pi}} e^{-x^2} dx \right], \quad b = (1 - e^{-a}) \sqrt{I},
\]

which corresponds to a normal distribution with variance \( 2I \), truncated at its mean.
26. Census Distribution Relating to an Arbitrary Start

We now consider the distribution of the number of "landings" in the interval \((0, I - 1)\) when the first "landing" in the interval has arisen from a source \(B\) which is \(H\) steps on the negative side of the origin \(t = 0\). (See Fig. 9.)

Let \(J(0)\) be the P.D. and \(v_H(n, I)\) the probability of \(n\) landings in the interval \((0, I - 1)\), with g.f. \(V_H(z : I)\). Let

\[
E_H(n : 0) = \sum_{I=1}^{\infty} v_H(n, I) \theta^I, \quad |\theta| < 1, \quad n = 0, 1, 2 \ldots
\]

Then since

\[
v_H(n + 1, I) = \sum_{I=0}^{I-1} k_H(t) u(n, I - t), \quad n = 0, 1, 2 \ldots \tag{64}
\]

and noting that

\[u(n, I - t) = 0, \quad t > I,\]

we have using (36),

\[E_H(n + 1 : 0) = K_H(0) H(n : 0). \tag{65}\]

Moreover

\[E_H(0 : 0) = \sum_{I=1}^{\infty} k_H^*(I) \theta^I = K_H^*(0) - 1. \tag{66}\]

Hence, from (37) and (65), and from (66) and (6),

\[
\begin{align*}
E_H(n : 0) &= 0K_H(0)[J(0)]^{n-1} [1 - J(0)][1 - \theta], \quad n = 1, 2, \ldots, \\
E_H(0 : 0) &= 0[1 - K_H(0)][1 - \theta],
\end{align*}
\]

for \(|\theta| < 1\).

Introducing the bivariate generating function of \(v_H(n, I)\), we have

\[
V_H(z, 0) = \sum_{I=1}^{\infty} \sum_{n=0}^{\infty} v_H(n, I) z^n \theta^I = \sum_{I=1}^{\infty} V_H(z : I) \theta^I
\]

valid for \(|\theta| < 1, |zJ(0)| < 1\).

As an alternative to the recurrence relation (64) it can be seen that

\[
v_H(n + 1, I) = \sum_{I=0}^{I-1} j(t) v_H(n, I - t), \quad n = 1, 2 \ldots \tag{69}
\]

so that

\[E_H(n + 1 : 0) = J(0) E_H(n : 0), \quad n = 1, 2, \ldots \tag{70}\]

Hence from (65) and (70)

\[K_H(0) H(n : 0) = J(0) E_H(n : 0), \quad n = 1, 2, \ldots\]
27. **Limiting Form of** $V_H(z : I)$

If the mean $\mu$ of the P.D. is finite, then, letting $H \to \infty$ in (68) and using $K_{\infty}$ from (18a), we have

$$V_{\infty}(z : I) = \text{coeff. of } 0^I \text{ in } \frac{\theta \mu(1 - \mu)[1 - z J(0)] + (z - 1) [1 - J(0)]}{\mu (1 - \mu)^2 [1 - z J(0)]}.$$  

(72)
In some ways a rather more convenient form is

\[ V_\omega(z : I) = 1 + I(1 - z^{-1})/\mu + W(z : I), \quad \ldots \quad (73) \]

where

\[ W(z : I) = \text{coeff. of } \theta^I \text{ in } \sum_{l=0}^{I} \frac{1 - z - \theta}{(1 - \theta)} \left[ 1 - zJ(0) \right]. \]

It is of interest to note that from (39a)

\[ V_\omega(z : I) = 1 + \frac{z - 1}{\mu} \sum_{s=1}^{I} U(z : s). \quad \ldots \quad (74) \]

Hence for the mean \( \nu^0 \) and variance \( \nu_2^0 \) of the \( V_\omega \) distribution we have

\[ \nu^0 = I/\mu, \quad \ldots \quad (75) \]

\[ \nu_2^0 = \frac{I}{\mu^2} - \frac{I^2}{\mu^3} + \frac{2}{\mu} \sum_{s=1}^{I} \nu(s), \quad \ldots \quad (76) \]

where \( \nu(s) \) is the mean of the census distribution \( U(z : s) \).

The property of the mean of \( V_\omega \) given in (75) is remarkable. It is worth emphasizing that it holds exactly for all P.D.'s with a finite mean, excepting those for which \( J(6) \) depends upon the interval length. That this should be so is apparent from the consideration that the expected number of landings in the interval is \( \frac{1}{\mu} \) and that \( \pi(t) = 1/\mu \), where the P.D. has a finite mean, and the source is at infinity.

Moreover all the moments of \( V_\omega \) are finite in this case since they depend upon \( \mu \) and the moments of \( U \); and the radius of convergence of \( V_\omega(z : I) \) as a function of \( z \) is \( 1/J(0) \). From (76) there is the inequality

\[ \sum_{s=1}^{I} \nu(s) > \frac{I}{2} \left( \frac{1}{\mu} - 1 \right), \]

so that since \( \nu(s) \) is non-decreasing (see §18, iv)

\[ \nu(I) > \frac{1}{2} \left( \frac{1}{\mu} - 1 \right). \]

28. A more general type of census distribution arises when there is a chance that the "jumper" may be absorbed or annihilated at any point in the period. For example, in the game of cricket, the batsman may score 0, 1, \ldots \infty from any one ball (allowing for the unlikely circumstance of scoring more than 6 from a ball as might arise with a benevolent or frustrated fielding side), or he may be out (annihilated, in a sense). The p.g.f. for the P.D. will now be of the form

\[ J^{(A)}(0) = p(j(0) + j(1) + \theta^2 j(2) + \ldots \} + q \theta^b \]

\[ = pJ(0) + q \theta^b; \quad \text{say,} \]

where \( b \) is arbitrary subject to \( b \geq I \) (though it is convenient to regard it as indefinitely great, particularly in considering the limiting form as \( I \to \infty \)), and where \( q \) is the probability of absorption, and \( p + q = 1 \). It is to be noted that the "jumper" may make a jump of \( t \geq 0 \), or it may be annihilated. The C.D. in this case is given by the coefficient of \( \theta^I \) in

\[ \frac{0}{I - zJ^{(A)}(0)} \cdot \frac{1 - J^{(A)}(0)}{1 - 0}, \]

or after rearrangement

\[ U^{(A)}(z : I) = \frac{p(1 - z)}{1 - pz} U(pz : I) + \frac{q}{1 - pz}, \quad \ldots \quad (77a) \]

where \( U(pz : I) \) is found from (38) or (40).
When, as in the present case, the P.D. includes a finite probability of absorption, the asymptotic form for large $I$ differs fundamentally from the cases considered in §§24 and 25, and is in fact geometric, that is,

$$U^{(A)}(z : I) \sim q/(1 - pq), \quad q \neq 0.$$

For evidently from (77a) it will be sufficient to prove that

$$p^n \sum_{s=0}^{\infty} u(s, I) \to 0$$

as $I \to \infty$ with $n$ fixed, for the vanishing of this expression when $n$ is large is obvious. But from (42), writing $[J(0)]^n = a_0 + a_1t + a_2t^2 + \ldots$, we have

$$\sum_{s=0}^{n-1} u(s, I) = \sum_{r=1}^{\infty} a_r.$$

Hence as $I$ becomes large, it follows that all terms in the expansion of $U(z : I)/(1 - pq)$ become vanishingly small, and the asymptotic form of $U^{(A)}(z : I)$ as stated above is proved.

The result is in exact agreement with the commonsense view that for infinitely large $I$ the number of landings is a "success run".

It may be noted that the probability associated with the value of $t$ at the moment of extinction is given by the coefficient of $t^r$ in

$$G(0) = (1 - p)/(1 - pJ(0)), \quad \ldots\ldots\ldots\ldots\ldots\ldots(77b)$$

a result given immediately by applying to a success run the well known relationship on multiplicative stochastic processes due to Watson (1889).

This type of process is well illustrated by cricket, where $t$ would be the total score attained by a batsman at the completion of his innings. The view that the distribution of $t$ is approximately geometric was put forward by Elderton (1945) and discussed by Wood (1945). The distribution on the present model, though dominated by the geometric, should have considerably greater variance, as is indeed the case. The geometric distribution is fitted in Fig. 9A to the actual results obtained from the four series of test matches referred to earlier.
**TABLE 5**

**Part 1.—Examples**

<table>
<thead>
<tr>
<th>No.</th>
<th>$J(0)$</th>
<th>$K_H(0)$ or $k_H(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>$(1 - r) \theta$</td>
<td>$0$</td>
</tr>
<tr>
<td>(ii)</td>
<td>$\frac{(1 - r) \theta^2}{1 - r \theta}$</td>
<td>$(1 - r) \left{ \frac{1 + (1 - r) \theta + (1 - r)(1 - \theta)\theta^{-1}}{(2 - r)(1 - r \theta)} \right}$</td>
</tr>
<tr>
<td>(iii)</td>
<td>$1 - \frac{\theta H}{H(1 - \theta)}$</td>
<td>$H = 2, 3, \ldots$</td>
</tr>
<tr>
<td>(iv)</td>
<td>$\frac{\theta H}{H(1 - \theta)}$, $H = 2, 3, \ldots$</td>
<td>$k_H(t) = (hH^{-1} - h^2)(H - 1)$, $t = 0, 1, \ldots, H - 1$</td>
</tr>
<tr>
<td>(v)</td>
<td>$\left{ \frac{r}{1 - (1 - r) \theta} \right}^2$, $0 &lt; r &lt; 1$, $r \left{ \frac{1 + (1 - r) \theta - (1 - r)(1 - \theta)\theta^{-1}}{2(1 - (1 - r) \theta)^2} \right}$</td>
<td></td>
</tr>
<tr>
<td>(vi)</td>
<td>$(8 + \theta + \theta^2)/10$</td>
<td>$\frac{2}{3} + \frac{1}{3} \theta + \frac{1}{3} (1 - \theta) \left( -\frac{1}{2} \right)^H$</td>
</tr>
<tr>
<td>(vii)</td>
<td>$1 - \frac{(1 - \theta)(a + b \theta)}{(\lambda_1 - \theta)(\lambda_2 - \theta)}$</td>
<td>$\frac{1}{(a + b)(\lambda_1 - \theta)(\lambda_2 - \theta)}$</td>
</tr>
<tr>
<td></td>
<td>$\times \left{ (\lambda_1 - 1)(\lambda_2 - 1)(a + b \theta) + \left( \frac{\theta - 1}{a} \right)(a + b \lambda_1)(a + b \lambda_2) - \left( -\frac{b}{a} \right)^H \right}$</td>
<td></td>
</tr>
<tr>
<td>(viii)</td>
<td>$0 + \frac{1}{2} (1 - \theta)^{3/2}$</td>
<td>$K_{\lambda_1}(\theta) = 1 - \frac{1}{2} (1 - \theta)^{3/2}$</td>
</tr>
<tr>
<td>(ix)</td>
<td>$0 + \frac{1}{\lambda} (1 - \theta)\lambda$, $1 &lt; \lambda &lt; 2$</td>
<td>$K_{\lambda_2}(\theta) = 1 - \frac{1}{\lambda} (1 - \theta)^{\lambda-1}$</td>
</tr>
</tbody>
</table>

* For admissible values of $a, b, \lambda_1, \lambda_2$.

**PART 2**

**The Continuous Case**

29. The setting of the process in continuous space or time involves only few changes in notation. Corresponding to the discrete probability masses $j(t), k(t)$ etc., are the probability elements $j(t) dt$, $k(t) dt$, etc., where $j(t)$ and $k(t)$ are now employed as density functions. The part played by the p.g.f.'s $J(\theta), K(\theta)$ etc., in the discrete case is taken over by the Laplace transforms ($LT$) $J(\lambda), X(\lambda)$, etc. A close analogy exists throughout between the discrete and continuous cases, the similarity being evident in the mathematical form of the results. Closed contours about the origin in the $\theta$-plane are replaced by infinite line integrals in the $\lambda$-plane.

The following abbreviated notation is sometimes employed

$$J(\lambda) = \int_0^\infty e^{-\lambda x} j(x) dx, \quad iJ(\lambda) = \int_0^t e^{-\lambda x} j(x) dx,$$

$$J(\lambda) = \int_0^\infty e^{-\lambda x} j(x) dx = iJ(\lambda) + J_i(\lambda).$$
30. In so far as a distribution function can be regarded as the limit of a sequence of approximating step functions with equally spaced discontinuity points when the density of those points is allowed to increase without limit, it is interesting to consider whether the results already obtained in the discrete case can be used to conjecture the corresponding results in the general case. The following line of argument, though in no sense a rigorous one, strongly suggests that this should be so.

Consider a general distribution to be represented approximately by discrete probability masses \( j_0, j_1, j_2, \ldots \) situated close together at minute intervals \( \delta = 1/c \). Then \( dF(t) = j_{st} \) \((j_s \) being 0 if \( s \) in not an integer). The substitution \( \theta = e^{-k\delta} \) \((|\lambda| > 0) \) in \( J(\theta) = \Sigma j_s \theta^s \) gives

\[
J(\theta) = \Sigma e^{-k\lambda st} j_{st} = \int_0^\infty e^{-k\lambda t} dF(t) = J(\lambda) \text{ the L.T.}
\]

If an interval comprises \( I \) steps its length is \( I' = I\delta \). The substitution \( \theta = e^{-k\delta} \) converts expressions of the type

\[
\frac{1}{2\pi i} \oint_C \Phi(J(\theta)) \frac{\theta^0}{1 - \theta} \quad \cdots \quad (78)
\]

(where \( C \) is a closed contour \(|\theta| = e^{-\gamma\delta} < 1\)) into

\[
\frac{\delta}{2\pi i} \oint_C e^{\lambda t'} \Phi(J(\lambda)) e^{-k\lambda} d\lambda/(e^{-k\delta} - 1)
\]

(where \( C' \) passes downwards in the right half plane parallel to the imaginary axis between end points \( \gamma = \pm 2\pi i/\delta \)). By choosing \( \delta \) arbitrarily small, the expression passes formally into

\[
\frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\lambda t'} \Phi(J(\lambda)) d\lambda/\lambda.
\]

(79)

This result may be regarded as the analogue of (78). It is of course, by reason of its form, the inverse of a L.T.

31. **Theorem 7**

If \( p^*(t) = \int p(x) dx \), its Laplace Transform is

\[
\mathcal{P}^*(\lambda) = [1 - \mathcal{P}(\lambda)]/\lambda.
\]

For

\[
\mathcal{P}^*(\lambda) = \int_0^\infty e^{-\lambda t} \left( \int p(x) dx \right) dt = \int_0^\infty p(x) \left( \int_0^x e^{-\lambda t} dt \right) dx
\]

\[
= \frac{1}{\lambda} \int_0^\infty (1 - e^{-\lambda x}) p(x) dx = [1 - \mathcal{P}(\lambda)]/\lambda.
\]

(80)

**Theorem 8**

If \( F_1(x) \) and \( F_2(x) \) are distribution functions, and \( F_n(x) \) and \( F_\sigma(x) \) the corresponding distribution functions associated with the sum of \( n \) independent random values, then

\[
F_n(x) \gg F_\sigma(x) \text{ if } F_1(x) \gg F_1(x) \text{ for all } x.
\]
Proof.—As is well known [Cramér (1945), p. 190] the distribution function of the sum of two independent random variables with distribution functions \( F_A \) and \( F_B \) is given by

\[
F_{A+B}(x) = \int_{-\infty}^{\infty} F_A(x-t) dF_B(t) = \int_{-\infty}^{\infty} F_B(x-t) dF_A(t).
\]

Now set

\[
F_{n-1}(x) - \mathcal{F}_{n-1}(x) = G_{n-1}(x) \geq 0.
\]

We then have

\[
G_n(x) = \int_{-\infty}^{\infty} F_{n-1}(x-t) dF_1(t) - \int_{-\infty}^{\infty} \mathcal{F}_1(x-t) d\mathcal{F}_{n-1}(t)
\]

\[
= \int_{-\infty}^{\infty} \left[ G_{n-1}(x-t) + \mathcal{F}_{n-1}(x-t) \right] dF_1(t)
\]

\[
- \int_{-\infty}^{\infty} \left[ F_1(x-t) - G_1(x-t) \right] d\mathcal{F}_{n-1}(t)
\]

\[
= \int_{-\infty}^{\infty} G_{n-1}(x-t) dF_1(t) + \int_{-\infty}^{\infty} G_1(x-t) d\mathcal{F}_{n-1}(t).
\]

This being clearly \( \geq 0 \) establishes the theorem.

Theorem 9

If \( P(t) \) is the cumulative function of a probability distribution for which \( 0 < t < \infty \) and

\[
\mathcal{P}(\lambda) = \int_{0}^{\infty} e^{-\lambda t} dP(t),
\]

then for \( \lambda \) real and non-negative \( 0 < \mathcal{P}(\lambda) < 1 \).

The theorem is obvious since \( e^{-\lambda t} < 1 \).

Theorem 10

If \( \mathcal{P}_H(\lambda) \) tends to a limit \( \mathcal{P}(\lambda) \) for fixed \( \lambda > 0 \) as \( H \to \infty \), and if \( \mathcal{P}[\lambda, \xi] \) is the bivariate transform

\[
\int_{0}^{\infty} e^{-\xi H} \mathcal{P}_H(\lambda) \, dH, \quad (0 < \xi < \infty),
\]

then

\[
\mathcal{P}(\lambda) = \lim_{\xi \to 0} \{ \xi \mathcal{P}[\lambda, \xi] \}.
\]

For given any \( \varepsilon > 0 \) and fixed \( \lambda \) there exists an \( X = X(\varepsilon, \lambda) \) such that

\[
\mathcal{P} < \mathcal{P}_H < \mathcal{P} + \varepsilon,
\]

for all \( H > X \).

Hence for \( \xi > 0 \)

\[
(\mathcal{P} - \varepsilon) \int_{X}^{\infty} e^{-\xi H} \, dH < \int_{X}^{\infty} e^{-\xi H} \mathcal{P}_H \, dH < (\mathcal{P} + \varepsilon) \int_{X}^{\infty} e^{-\xi H} \, dH,
\]
whence
\[ (\varphi - \varepsilon) e^{-\xi X} < \xi \int_{-\infty}^{\infty} e^{-\xi H} \varphi_H dH < (\varphi + \varepsilon) e^{-\xi X}. \]

From theorem 9,
\[ \xi \int_{-\infty}^{\infty} e^{-\xi H} \varphi_H dH < X\xi, \]

hence, on passing to the limit,
\[ \varphi - \varepsilon < \lim_{\xi \to 0^+} \{\xi \varphi[\lambda, \xi]\} < \varphi + \varepsilon, \]

and, since \( \varepsilon \) is arbitrarily small, this establishes the theorem. The result is the analogue of the form of Cesàro's theorem given in §7f.

**Theorem 11**

If \( P(t) \) is the cumulative function of the probability distribution of a non-negative random variable (excluding the trivial case of unit point mass at \( t = 0 \)), and if \( \varphi(\lambda) \) is the L.T. of the distribution, then \( |\varphi(\lambda)| \) decreases to \( P(0) \) as \( R(\lambda) \) increases indefinitely.

The theorem, which appears intuitively obvious, is readily proved from the consideration that all distribution functions are continuous to the right (Cramér (1946), pp. 57, 166) or by decomposing \( P(t) \) into a discontinuous and continuous part.

**Theorem 12**

If \( \varphi(\lambda) \) is the L.T. of the probability distribution of a non-negative random variable, then \([1 - \varphi(\lambda)]^{-1} \) for \( R(\lambda) \geq \gamma > 0 \) is regular.

This result follows from the previous theorem and the consideration that \( \varphi(\lambda) \) is a regular function in the right half plane [Widder, 1941, p. 57].

**Theorem 13**

If
\[ \mu = \lim_{T \to \infty} \int_{-\infty}^{\infty} t p(t) dt \]
is finite, and
\[ \mu^0 = \lim_{\lambda \to 0} \{[1 - P(\lambda)]/\lambda\}, \]

then \( \mu = \mu^0 \); whilst if \( \mu \) is infinite so is \( \mu^0 \).

(i) \( \mu \) finite. By theorem 7,
\[ [1 - P(\lambda)]/\lambda = \int_{-\infty}^{\infty} e^{-\lambda t} p^*(t) dt \quad (\lambda > 0). \]

By considerations of uniform convergence and continuity (see Titchmarsh (1939), p. 26),
\[ \lim_{\lambda \to 0} \int_{-\infty}^{\infty} e^{-\lambda t} p^*(t) dt = \int_{-\infty}^{\infty} p^*(t) dt \]

if the latter integral converges. That the value of this integral is \( \mu \) is readily seen by reversing the order of integration in
\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) dx. \]
Hence $\mu_0 = \mu$.

(ii) $\mu$ infinite. Suppose that $\mu^o$ were finite. Then given $\mu^0$ there exists a $T$ such that

$$\int_0^T t \, p(t) \, dt > \mu^0 + 1.$$ 

But

$$\int_0^T \frac{1 - e^{-\lambda t}}{\lambda} \, p(t) \, dt < \frac{1 - P(\lambda)}{\lambda} \quad (\lambda > 0).$$

Hence

$$1 + \mu^o - \frac{1 - P(\lambda)}{\lambda} \leq \int_0^T \frac{\lambda t - 1 + e^{-\lambda t}}{\lambda} \, p(t) \, dt \leq \left(\frac{\lambda T - 1 + e^{-\lambda T}}{\lambda T}\right) \int_0^T p(t) \, dt,$$

and this tends to 0 as $\lambda \to 0$ leading to a contradiction. Hence $\mu^o$ must also be infinite.

Accumulated Probability

32. The expected number of landings in an infinitesimal interval $t$ units from the source is $\pi(t) \, dt$, and, as in Part 1, the source is included as a landing if the element $dt$ contains it. Clearly $\pi(t) = 0$ for $t < 0$ and $\pi(0)$ is equivalent to the Dirac $\delta$-function.

Then

$$\pi(t) \, dt = \sum_{r=0}^{\infty} \text{Prob. \, \text{r}th \, landing \, falls \, in \, dt}.$$ 

Since the Laplace transform of the $r^{th}$ convolution of the P.D. is the $r^{th}$ power of the transform of the P.D., and the taking of the transform is a linear operation,

$$\Pi(\xi) = \int_0^\infty e^{-\xi t} \, \pi(t) \, dt = \sum_{r=0}^{\infty} \left[ \mathcal{F}(\xi)\right]^r = 1/[1 - \mathcal{F}(\xi)], \quad \mathbf{R}(< \xi > 0).$$

(81)

33. The density function $\pi(t \mid x)$ refers to the expected density of landings at $t$ subject to there having been a previous landing at $x$.

Clearly

$$\pi(t \mid x) = \pi(t - x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\xi(t - x)} \Pi(\xi) \, d\xi$$

by the inversion theorem.

The Laplace transform of $\pi(t \mid x)$ is then

$$\Pi(\lambda \mid x) = \int_0^\infty e^{-\lambda t} \pi(t \mid x) \, dt = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{-\xi x} \Pi(\xi) \left[ \int_0^\infty e^{-(\lambda - \xi)t} \, dt \right] d\xi$$

$$= \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{-\xi x} \Pi(\xi) \frac{d\xi}{(\lambda - \xi)}, \quad \mathbf{R}(\lambda) > \gamma,$$

(82)

the change in the order of integration being justified by uniform convergence.
It follows that the “bivariate” Laplace transform of $\pi(t \mid - H)$ ($H > 0$) defined as

$$\Pi(\lambda, \xi) = \int_0^\infty e^{-\lambda H} \Pi(\lambda \mid - H) \, dH$$

is

$$\Pi(\xi)(\lambda - \xi).$$

(83)

34. The P.D. for which $\pi(t)$ is a constant $c$ for all $t > 0$ is determined by

$$\Pi(\lambda) - 1 = \int_0^\infty e^{-\lambda t} c \, dt = c/\lambda.$$

It follows from (81) that $\mathcal{F}(\lambda) = c/(c + \lambda)$, the Laplace transform of $j(t) = e^{-\lambda t}c$, the exponential distribution.

35. We now enquire under what conditions $\pi(t \mid - H)$ tends to a limit as $H \to \infty$.

By Th. 10 (appropriately modified) the limiting form (if it exists) of the transform $\Pi(\lambda \mid - H)$ as $H \to \infty$ is given by

$$\lim_{\xi \to 0} \{\xi \Pi(\lambda, \xi)\},$$

and this by (83) is

$$\lim_{\xi \to 0} \{\xi \Pi(\xi)(\lambda - \xi)\}.$$

By (81) we then have

$$\Pi(\lambda \mid - \infty) = \frac{1}{\lambda} \lim_{\xi \to 0} \{\xi/[1 - \mathcal{F}(\xi)]\}.$$

Hence by theorem 13

$$\Pi(\lambda \mid - \infty) = \frac{1}{\lambda} \mu \quad (\mu \text{ finite}),$$

$$= 0 \quad (\mu \text{ infinite}).$$

(84)

Inversion of the transform gives

$$\pi(t \mid - \infty) = 1/\mu.$$

(85)

That $\pi(t)$ has a limit $\pi$ as $t \to \infty$ [in the sense that, for every fixed $\Delta > 0$, the mean density in the interval $(t, t + \Delta)$ tends to $\pi$] has been proved by Blackwell (1948) to be true for any P.D. provided that the variate values are not all integral multiples of a constant, that is, when the problem does not reduce to the simple discrete case. Other studies on the limiting properties of $\pi(t)$ under weaker conditions include Täcklind (1945), Feller (1941), Cox and Smith (1954), and Smith (1954).

Distribution of $ON$

36. We consider here the continuous case of the problem of §8. Observation starts at $t = 0$ and the probability density function $k_H(t)$ refers to the first observed occurrence (landing) for a process starting in the past at $t = - H$. Particular interest attaches later to the existence and nature of limiting forms as $H \to \infty$.

The accumulated probability density at the point $t = x$ is, in the notation of §33, $\pi(x \mid - H)$. But this must also be equal to

$$\int_0^x k_H(t) \pi(x \mid t) \, dt.$$

Noting that $\pi(x \mid t) = 0$ for $t > x$, we have

$$\pi(x \mid - H) = \int_0^\infty k_H(t) \pi(x - t) \, dt.$$

(86)

We are concerned only with solutions of this integral equation for $H > 0$. 
The L.T. with respect to \( x \) of the right-hand member of (86), by the convolution theorem, is the product of the transforms of \( k_H(t) \) and \( \pi(t) \). That of the left-hand member is given by (82). Hence

\[
\mathcal{X}_H(\lambda) \Pi(\lambda) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\xi H} \Pi(\xi) \frac{d\xi}{(\lambda - \xi)}
\]

(87)

where

\[
\Pi(\lambda) = [I - \mathcal{F}(\lambda)]^{-1}
\]

\[
\text{and } R(\lambda) > \gamma > 0.
\]

As in §8 for the discrete case, (87) may be written

\[
\mathcal{X}_H(\lambda) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\xi H} \frac{\mathcal{F}(\xi) - \mathcal{F}(\lambda)}{(\lambda - \xi)} d\xi,
\]

(88)

in which form \( \xi = \lambda \) is no longer a pole of the integrand and the only restriction on \( \lambda \) is \( R(\lambda) > 0 \).

Examples of \( k_H(t) \) are illustrated in Figs. 10 and 11, and in Table 6.

37. Limiting Form of the Distribution of ON as B moves to Minus Infinity.

The "bivariate" L.T. of \( k_H(t) \) defined as

\[
\mathcal{X}[\lambda, \xi] = \int_0^\infty e^{-\xi H} \mathcal{X}_H(\lambda) dH
\]

is given by (87) as

\[
\Pi(\xi)/[(\lambda - \xi) \Pi(\lambda)].
\]

By theorem 10 and result (81) or as in §35,

\[
\mathcal{X}_H(\lambda) = [\lambda \mu \Pi(\lambda)]^{-1} = [1 - \mathcal{F}(\lambda)]/(\lambda \mu),
\]

(88b)

where \( \mu \) the mean of the P.D. must be finite for \( k_{\infty}(t) \) to have definite meaning.

By theorem 7 it follows that

\[
k_{\infty}(t) = j^*(t)/\mu.
\]

(89)
\[ k(\tau) \quad \text{for } H = 2.5 \quad \text{and} \quad \begin{cases} j(t) = e^{-(t-\tau)}, & t \geq 1 \\ j(t) = 0, & t < 1 \end{cases} \]

\[(i) \quad k(\tau) = e^{-\frac{(\tau+1.5)}{5}} + (\tau+0.5) e^{-\frac{(\tau+0.5)}{5}} + \frac{1}{2}(\tau - 0.5)^2 e^{-\frac{(\tau-0.5)}{5}}

(ii) \quad = 1.34 e^{-\tau} \]

Fig. 11.
ZEROS OF $\hat{f}(\lambda) - 1$

$$\hat{f}(\lambda) = (\lambda + 1)^{-\gamma}, \ c = \gamma/8$$

$$j(t) = e^{-\epsilon} t e^{-\epsilon t}/\Gamma(\epsilon) ; \ t \geq 0, \ c > 0$$

$$= 0 \quad \text{otherwise}$$

\[ |\lambda + 1| = 1 \]

\[ \lambda = 0 \]

\[ x_{\lambda} = -1 \]

(i) \[ \hat{f}(\lambda) = (1 - e^{\lambda})/\lambda \]

\[
\begin{cases} 
    j(t) = 1, & 0 \leq t \leq 1 \\
    = 0 & \text{otherwise}
\end{cases}
\]

(ii) \[ \hat{f}(\lambda) = e^{\lambda}/(1 + \lambda) \]

\[
\begin{cases} 
    j(t) = 0, & t < 1 \\
    = -e^{-\epsilon}, & t \geq 1
\end{cases}
\]

FIG. 12.
38. A Series for $\mathcal{H}_H(\lambda)$ and the Limit as $H \to \infty$.

From (87) we have

$$\mathcal{H}_H(\lambda) = [\mathcal{F}(\lambda) - 1] \frac{1}{2\pi i} \int_L A(\xi, \lambda) \, d\xi, \quad R(\lambda) > \gamma > 0, \quad \ldots \quad (90)$$

where $L$ is the line $\gamma - i\infty, \gamma + i\infty$ and

$$A(\xi, \lambda) = e^{iH\xi}[(\xi - \lambda)(1 - \mathcal{F}(\xi))].$$

Let $C(R)$ be part of the circumference of a circle radius $R$, centre the origin, and lying for the most part in the negative half-plane, and which with the line $(\gamma - iR, \gamma + iR)$ forms a simple closed contour $D(R)$.

Let $\mathcal{F}(\xi)$ be a meromorphic function throughout the plane, not reducing to the transform of a unit point mass at the origin (for which the mean of the distribution would be zero). The mean $\mu$ of the P.D. is then $> 0$.

Now provided we can show that the integral in (90) over the contour $C(R)$ tends to zero as $R \to \infty$, then we have

$$\mathcal{H}_H(\lambda) = [\mathcal{F}(\lambda) - 1] \sum \{\text{Residues in } D(R)\},$$

where the residues of $A(\xi, \lambda)$ occur only at the zeros of $\mathcal{F}(\xi) - 1$, since $\xi = \lambda$ is outside $D(R)$.

Some examples of the zeros of $\mathcal{F}(\xi) - 1$ are illustrated in Fig. 12.

Denoting the zeros of $\mathcal{F}(\xi) - 1$ by $\xi = \xi_n, s = 0, 1, 2 \ldots$ with multiplicities $\lambda_n$, where in particular $\xi_0 = 0$, and $\lambda_0 = 1$, then by considering the integral form of $\mathcal{F}(\xi)$ it is easy to see that $R(\xi_n) < 0, s = 1, 2 \ldots$. Thus we obtain

$$\mathcal{H}_H(\lambda) = [\mathcal{F}(\lambda) - 1] \sum_{s=0}^{\infty} \frac{a(H, \lambda, \xi_n)}{(\lambda_n - 1)!}, \quad \ldots \quad (91)$$

where

$$a(H, \lambda, \xi_n) = D_\xi^{s-1} \{e^{H\xi}[(\xi - \lambda)^{3/4}][\xi - \lambda)(1 - \mathcal{F}(\xi))]) \bigg|_{\xi = \xi_n}.$$
39. If the distribution \( k_{\infty}(t) \) has a finite mean \( \tau \) this is given, on differentiating (88b), by

\[
\frac{d}{d\lambda} \mathcal{K}_{\infty}(\lambda) \bigg|_{\lambda=0} = \frac{1}{2\mu} \frac{d^2}{d\lambda^2} \mathcal{F}(\lambda) \bigg|_{\lambda=0}.
\]

If the P.D. has in addition to the mean a finite second moment \( \mu^2 \) then \( \tau = \frac{1}{2} \mu^2 / \mu \). Otherwise \( \tau \) is infinite.

40. Distribution of the Length of \( OL \).—As in the discrete case the point \( O \) is not included in the interval \( BO \), so that the last jump ends on or beyond \( O \). We require the distribution \( k_{Hc}(t) \) of the length of that part of the half open interval \( BO \) belonging to the last jump.

Clearly

\[
k_{Hc}(t) = 0, \quad t > H,
\]

\[
= \pi(H - t) j^*(t), \quad t < H, \quad \ldots \quad \ldots \quad (95)
\]

whilst at \( t = H \) there is a discrete probability mass \( j^*(H) \).

By the convention of \( \S 32 \), all these cases are covered by the single expression

\[
k_{Hc}(t) = \pi(H - t) j^*(t).
\]

Hence

\[
\mathcal{K}_H^c(\lambda) = \int_0^\infty e^{-\lambda t} k_{Hc}(t) \, dt = \int_0^\infty \pi(H - t) \left[ j^*(t) e^{-\lambda t} \right] dt,
\]

where \( \pi(0) \) is the Dirac \( \delta \)-function.

The L.T. with respect to \( H \) of the convolution on the right is \( \Pi(\xi) \varphi(\xi) \) where

\[
\varphi(\xi) = \int_0^\infty e^{-\xi t} j^*(t) e^{-\lambda t} \, dt = \mathcal{F}^*(\lambda + \xi). \quad \ldots \quad \ldots \quad (96)
\]

Hence

\[
\mathcal{K}_H^c(\lambda) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \Pi(\xi) \mathcal{F}^*(\lambda + \xi) e^{H\xi} \, d\xi. \quad \ldots \quad \ldots \quad (96b)
\]

Substituting for \( \Pi \) by (81) and \( \mathcal{F}^* \) by (80), we then have

\[
\mathcal{K}_H^c(\lambda) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{H\xi} \frac{[1 - \mathcal{F}(\lambda + \xi)] d\xi}{[1 - \mathcal{F}(\xi)][\lambda + \xi]} \quad \ldots \quad \ldots \quad (97)
\]

41. The Limiting Form of \( k_{Hc}(t) \) as \( H \to \infty \)

By (96b) the "bivariate" L.T.

\[
\mathcal{K}^c[\lambda, \xi] = \int_0^\infty e^{-tH} \mathcal{K}_H^c(\lambda) \, dH = \Pi(\xi) \mathcal{F}^*(\lambda + \xi). \quad \ldots \quad \ldots \quad (98)
\]

By theorem 10 and \( \S 35 \),

\[
\mathcal{K}_{\infty}^c(\lambda) = \lim_{\xi \to 0} \{ \xi \Pi(\xi) \mathcal{F}^*(\lambda + \xi) \} = \frac{1}{\mu} \mathcal{F}^*(\lambda), \quad \ldots \quad \ldots \quad (99)
\]

the existence of a definite limiting form being conditional on the mean \( \mu \) of the P.D. being finite.

By (89) it will be seen that the limiting forms of \( k_{Hc}(t) \) and \( k_{Hc}^c(t) \) as \( H \to \infty \) are alike, though for finite \( H \) they can be very different particularly with respect to the limited range of \( k_{Hc}^c(t) \) and the presence of a discrete probability mass at \( t = H \).

The conclusion from (99) that \( k_{\infty}(t) = j^*(t)/\mu \) may also be derived from (95) by noting that, under the conditions considered in \( \S 35 \), \( \pi(H - t) \to 1/\mu \) when \( t \) is finite and \( H \to \infty \).
42. The Distribution of LN.—LN is the jump by which the point O is “spanned”. By our
convention L lies to the left of O, though it is permissible for N to coincide with O.
Let \( s_H(t) \) be the probability density associated with the length \( (t) \) of jumps of this class.
Then
\[
s_H(t) = j(t) \int_{H-t}^H \pi(x) \, dx, \quad -\infty < t < \infty, \quad \ldots \quad \ldots \quad (100)
\]
by reason of \( j(t) \) being 0 for \( t < 0 \) and \( \pi(x) \) being the Dirac \( \delta \)-function for \( x = 0 \) and zero for \( x < 0 \).
Then
\[
s_H(t) = j(t) \frac{\gamma^+}{2\pi i} \int_{-\infty}^{\gamma^+} \frac{e^{i\theta} - e^{(H-\theta)}/0}{1 - \mathcal{F}(0)} \, d\theta, \quad \gamma > 0. \quad \ldots \quad \ldots \quad (101)
\]
Hence
\[
\mathcal{F}_H(\lambda) = \frac{1}{2\pi i} \int_{-\infty}^{\gamma^+} \frac{e^{i\theta}[\mathcal{F}(\lambda) - \mathcal{F}(\lambda + \theta)] \, d\theta}{0[1 - \mathcal{F}(0)]}. \quad \ldots \quad \ldots \quad (102)
\]
It should be noted that \( s_H(t) \) is subject to a discontinuity at \( t = H \), for if \( t > H \), a term in \( \pi(0) \) enters (100).

43A. The Census Distribution for an Interval Beginning at the Source.—The probability of
exactly \( n \) landings (or returns to the initial state) in an interval \( OI \) (which does not include \( I \) as
end point) when \( B = O \) is denoted by \( u(n, I) \).
Clearly
\[
\begin{align*}
  u(n + 1, I) &= \int_0^I j(t) u(n, I - t) \, dt, \quad n = 0, 1, 2, \ldots \\
  u(0, I) &= \int_I^\infty j(t) \, dt = j^*(I). 
\end{align*}
\]
Since \( u(n, I - t) = 0 \) for \( t > I \) the upper limit in the first integral may be replaced by \( \infty \).
Denoting the transform
\[
\int_0^\infty e^{-tI} u(n, I) \, dI
\]
by \( \mathcal{H}(n : \xi) \) and applying the convolution property of L.T.'s we have
\[
\mathcal{H}(n + 1 : \xi) = \mathcal{H}(n : \xi) \mathcal{F}(\xi), \quad \mathcal{H}(0 : \xi) = \mathcal{F}^*(\xi). \quad \ldots \quad \ldots \quad (103)
\]
Hence
\[
\mathcal{H}(n : \xi) = [\mathcal{F}(\xi)]^n[\mathcal{F}^*(\xi)]. \quad \ldots \quad \ldots \quad (104)
\]
Substituting for \( \mathcal{F}^* \) by (80) and inverting we have
\[
u(n, I) = \frac{1}{2\pi i} \int_{-\infty}^{\gamma^+} e^{i\xi} [\mathcal{F}(\xi)]^n[1 - \mathcal{F}(\xi)] \, d\xi/\xi. \quad \ldots \quad \ldots \quad (105)
\]
From (104) we have immediately the generating function,
\[
U[z, \xi] = \sum_{n=0}^\infty \mathcal{H}(n : \xi) \, z^n = \mathcal{F}^*(\xi)[1 - z\mathcal{F}(\xi)], \quad |z\mathcal{F}(\xi)| < 1,
\]
with the result that
\[ u(z : I) = \sum_{n=0}^{\infty} u(n, I) z^n = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{iz}[1 - f(\xi)] d\xi \quad [1 - z f(\xi)]^\frac{1}{\xi}. \] (106)

Examples of \( u(n, I) \) are given in Fig. 13 and Table 7.

43b. An Application to Chemical Theory.—The degradation of polymer chains is a subject for which many mechanisms have been advanced (see for example, Jellenek (1955)). In certain cases it is believed that, if one end of a molecule becomes activated under appropriate treatment, links of the chain are split off in extremely rapid succession until a side reaction steps in to stop the process. The molecule then remains stable unless reactivated at an end. It appears likely that the amount of monomer split off following each activation has a negative exponential distribution. The situation however is complicated by the fact that the original polymer chains are finite and differ in length. The most important case arises when the chain length has, so it is believed, a negative exponential distribution, but sometimes, as a result of radical recombination, such chains appear to be joined in pairs at random, the distribution of the combined lengths then being given by a simple convolution.

A typical question which can be asked of this picture would be “What is the distribution of the number of activations required to result in the complete disintegration of one molecular chain?”

If \( j(t) = ke^{-kt} \) represents the distribution of the quantity of monomer released at each “jump” [so that \( f(\lambda) = k/(k + \lambda) \)] and \( n \) the number of activations (which is one more than the number of “landings”), then the p.g.f. of the random variable \( n \) for molecules of fixed length \( I \) is

\[ G(z : I) = zU(z : I) = \frac{z}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{i\lambda} \frac{d\lambda}{\lambda + k(1 - z)} \]

by substituting for \( f(\lambda) \) in (106) and simplifying.

Hence

\[ G(z : I) = ze^{I/(e-1)}. \]

Taking

\[ dF(I) = e^{-I} e^s I^{s-1} dl, \quad s = 1, 2, \]

and integrating \( G(z : I) dF(I) \) over \( I \) we then have the required p.g.f.

\[ G(z) = z \int_0^\infty e^{-I(e+kz)} I^{s-1} dl = z \left( \frac{c}{e + k - k\lambda} \right)^s, \]

the result thereby being a Pascal distribution with index 1 or 2, displaced one unit to the right.
A more detailed application of the ideas of the present paper to the field of polymer chemistry are being developed at length elsewhere. We are indebted to Dr. Manfred Gordon for bringing this application to our notice.

44. The Radius of Convergence of $U(z : I).$—(a) If the P.D. has a discrete probability mass $j_0$ at $t = 0$ we show in this section that the radius of convergence ($p$) of the p.g.f. of the census distribution $(u)$ is $1/j_0$. One of the important consequences of this result is (by theorem 2) that all moments of the C.D. exist and are finite.

(b) The simplest case is that of a bounded continuous frequency function, $j_0$ being necessarily 0.

By (103)

$$u(n + 1, I) = \int_0^1 u(n, t) f(I - t) dt \leq b \int_0^1 u(n, t) dt,$$

where $b$ is an upper bound to $j(t)$ in $(0, \infty)$.

<table>
<thead>
<tr>
<th>j(t)</th>
<th>$kH(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) $\beta e^{-\beta t}, \beta &gt; 0$</td>
<td>$\beta e^{-\beta t}$</td>
</tr>
<tr>
<td>(ii) $te^{-t}$</td>
<td>$\frac{1}{2} e^{-t} {1 + t - (1 - t) e^{-2\beta H}}$</td>
</tr>
<tr>
<td>(iii) $(b^2 - a^2) e^{-bt} \sinh at/a,$</td>
<td>$(b^2 - a^2) e^{-bt} {b \sinh at \cosh bh + a \cosh at \sinh bh}$</td>
</tr>
<tr>
<td>$b &gt; a \geq 0,$</td>
<td></td>
</tr>
<tr>
<td>(iv) $\begin{cases} pe^{-t}, \ \text{Prob. } (t = 0) = 1 - p = q, \ 0 &lt; p \leq 1, \end{cases}$</td>
<td>$e^{-t}$</td>
</tr>
<tr>
<td>(v) $e^{-t}(1 + at)/(1 + a), \ a \geq 0,$</td>
<td>$[1 + a + at - a^2(1 + t) \exp {-(1 + 2a) \beta H/(1 + a)}] \times$</td>
</tr>
<tr>
<td>$0 &lt; t &lt; \infty.$</td>
<td>$e^{-t}/(1 + 2a).$</td>
</tr>
<tr>
<td>(vi) $\begin{cases} \frac{1}{c}, \ 0 &lt; t &lt; c, \ 0, \ t &gt; c, \end{cases}$</td>
<td>$\begin{cases} \sum_{s=0}^{r} A(H, s) - \sum_{s=1}^{r+1} A(H + t - c, s - 1), \ 0 &lt; t &lt; (r + 1) c - H; \ \sum_{s=0}^{r} A(H, s) - \sum_{s=1}^{r+1} A(H + t - c, s - 1), \ (r + 1) c - H &lt; t &lt; c; \ 0, \end{cases}$</td>
</tr>
<tr>
<td>(vii) $\begin{cases} 0, \ 0 &lt; t &lt; a, \ e^{a-t}, \ t \geq a, \end{cases}$</td>
<td>$\begin{cases} \sum_{s=0}^{r+1} B(H + t - a, s - 1), \ 0 &lt; t &lt; (r + 1) a - H; \ \sum_{s=0}^{r+1} B(H + t - a, s - 1), \ (r + 1) a - H &lt; t &lt; a; \ e^{a-t} \sum_{s=0}^{r+1} B(H, s), \ t &gt; a. \end{cases}$</td>
</tr>
</tbody>
</table>

$r$ is the least integer exceeding $H/c - 1.$

$A(H, s) = (sc - H)^s \exp \{((H - sc)/c)/[s! c^{s+1}]\}.$

$B(H, s) = (H - sa)^s e^{sa - H}/s!.$

The range of $j(t)$ and $kH(t)$ where not stated is $0 < t < \infty.$
TABLE 7.—Census Distributions Corresponding to Various P.D.'s
\(j(t)\) or \(U(z : I)\)

\[
\begin{align*}
(i) & \quad \frac{e^{-t} t^{-\frac{1}{2}}}{\Gamma(c)}, \quad c > 0, \\
& \quad \int_0^1 e^{-t} \left( \frac{t^{nc}}{\Gamma(nc)} - \frac{t^{nc+1}}{\Gamma(nc+1)} \right) dt, \quad n > 1; \\
& \quad 1 - \int_0^1 e^{-t} t^{-\frac{1}{2}}dt, \quad n = 0. \\
(b) & \quad P(2I, 2nc + 2c) - P(2I, 2nc), \quad n > 1; \\
& \quad P(2I, 2c), \quad n = 0.
\end{align*}
\]

where
\[
P(x, y) = 2^{-\frac{1}{2}y} \Gamma(\frac{1}{2} y - 1) \int_0^\infty e^{-tx} x^{y-1} dx.
\]

\[
\sum_{s=0}^{\infty} e^{-s} \left( \frac{\Gamma(nc + s + 1)}{\Gamma(nc + s + c + 1)} - \frac{\Gamma(nc + s)}{\Gamma(nc + s + c)} \right). \\
\sum_{x=nc}^{nc+c-1} e^{-x} x! , \quad c \text{ an integer.}
\]

\[e = p/q, \text{ a rational fraction in its lowest terms.}\]

\[
U(z : I) = (z - 1) z^{a/q} e^{-\frac{1}{2}x^2} \left( \sum_{s=0}^{p-1} \exp \left\{ \frac{I(z/s)}{qz} \right\} e^{2zqz/n(p)} \right) \\
+ (z - 1) e^{-1} \frac{\sin \pi x}{\pi} \int_0^\infty \frac{e^{-tx} x^y dx}{(1 + x)(x^2 - 2zx \cos \pi + z^2)}.
\]

(ii) \(pe^{-t}\),
\[\text{Prob. } (t = 0) = q,\]
\[p + q = 1, \quad 0 < p < 1,\]

\[
(n + 1) u(n + 1 : I) = [(2n + 1) q + pl] u(n : I), \\
u(0 : I) = pe^{-t}.
\]

\[
\frac{ae^{-a^2/4t}}{2\sqrt{\pi} t^2}, \quad a > 0,
\]

\[
\int_{a}^{(n+1)a} e^{-x^2} 2 \sqrt{2\pi} dx, \quad \chi = a/\sqrt{24}.
\]

(iv) \(\sqrt{\frac{a}{\pi t}} e^{-at}, \quad a > 0,\)
\[\int_0^{a\sqrt{24}} e^{-x^2} \left( \frac{x^{n+1}}{24} - \frac{x^{n+1}}{12} \right) dx, \quad n > 1; \\
1 - \int_0^{a\sqrt{24}} e^{-x^2} \left( \frac{x^{n-1}}{24} - \frac{x^{n-1}}{12} \right) dx, \quad n = 0.
\]

(v) \(0, \quad 0 \leq t < a;\)
\[ea^{-t}, \quad t \geq a,\]

\[
\begin{align*}
& 1 - \sum_{s=0}^{n-1} A(n, s), \quad n = [I/a]; \\
& \sum_{s=0}^{n-1} A(n + 1, s) - \sum_{s=0}^{n-1} A(n, s), \quad 0 < n < [I/a]; \\
& e^{a-I}, \quad n = 0.
\end{align*}
\]

\[A(n, s) = (I - na)^s e^{-a-I}/s!; \quad [I/a] \text{ means integer part of } I/a.\]

(vi) \((b^2 - a^2) e^{-bt} \frac{\sinh at}{a}, \quad b > a > 0,\)
\[e^{-bt} \left( \cosh (lw) + b \sinh (lw) \right)/w, \quad w^2 = a^2 + z(b^2 - a^2).\]
But

\[ u(0, I) = j^*(I) < 1, \]

so that

\[ u(1, I) \leq \int_0^I j(I - t) \, dt < bI, \]

\[ u(2, I) \leq \int_0^I (bI) j(I - t) \, dt < b^2 I^2 / 2!, \]

and so on, giving

\[ u(n, I) \leq (bI)^n / n!. \]

\[ \sum u(n, I) z^n \] then converges wherever \( \exp \{ bI \mid z \mid \} \) converges, and is therefore analytic in the whole \( z \)-plane.

(c) Similarly if \( j(t) \) is continuous but not bounded in the neighbourhood of \( t = 0 \) and if the order of contact is \( j(t) = O(t^{\alpha}) \) where \( 0 < \alpha < 1 \), then it may be shown that \( p = \infty \) by using

\[ u(n + 1, I) = \int_0^I \frac{u(n, I - t)}{t^\alpha} \{ t \, j(t) \} \, dt \]

for a positive \( k \), and continuing as in (b).

---

**Table 8A**

<table>
<thead>
<tr>
<th>Discrete Variable</th>
<th>Continuous Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J(0) ) converges within unit circle.</td>
<td>( \mathcal{J}(\lambda) ) converges in positive half-plane.</td>
</tr>
<tr>
<td>( J(0) = 1 ) has one zero at ( 0 = 1 ) and (</td>
<td>0</td>
</tr>
<tr>
<td>( K_{H}(0) = (1 - J(0)) \frac{1}{2\pi i} \int_{</td>
<td>\xi</td>
</tr>
<tr>
<td>( K_{\infty}(0) = \frac{J(0) - 1}{\mu(0 - 1)} ), ( \mu &lt; \infty )</td>
<td>( K_{\infty}(0) = \frac{\mathcal{J}(0) - 1}{\mu(0 - 1)} ), ( \mu &lt; \infty )</td>
</tr>
<tr>
<td>( k_{H}(t) = \sum_{s=0}^{\infty} D_{s}^{H-1} \left{ \frac{\xi^{H+1}(\xi - \xi_{s})}{(\lambda_{s} - 1)!} \right} \frac{J_{s+1}(\xi_{s})}{(J_{s+1}(\xi_{s}) - 1)} \right} \frac{\xi - \xi_{s}}{J_{s+1}(\xi_{s})} )</td>
<td>( k_{H}(t) = \sum_{s=0}^{\infty} D_{s}^{\lambda_{s}-1} \left{ \frac{e^{(\lambda_{s} + 1) H}(\xi - \xi_{s})\lambda_{s}}{(\lambda_{s} - 1)!} \right} \frac{\mathcal{J}(\xi) - \xi_{s}}{1 - \mathcal{J}(\xi)} \right} \xi - \xi_{s} )</td>
</tr>
<tr>
<td>where the zeros of ( J(0) - 1 ) are ( \xi_{s} ) with multiplicity ( \lambda_{s} ), ( s = 0, 1, \ldots ), ( \lambda_{0} = 1, \xi_{0} = 1 )</td>
<td>where the zeros of ( \mathcal{J}(0) - 1 ) are ( \xi_{s} ) with multiplicity ( \lambda_{s} ), ( s = 0, 1, \ldots ), ( \lambda_{0} = 1, \xi_{0} = 0 )</td>
</tr>
</tbody>
</table>
TABLE 8B

<table>
<thead>
<tr>
<th>Discrete Variable</th>
<th>Continuous Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{j}(0) = \frac{1}{2\pi i} \int_{</td>
<td>\varphi</td>
</tr>
<tr>
<td>$S_{\infty}(0) = \mu \frac{dJ(0)}{d\mu}$, $\mu &lt; \infty$</td>
<td>$\gamma &gt; 0$</td>
</tr>
<tr>
<td>$K_{j}(0) = \frac{1}{2\pi i} \int_{</td>
<td>\varphi</td>
</tr>
<tr>
<td>$k_{\infty}(t) = \sum_{s=t+1}^{\infty} j(s) / \mu$, $\mu &lt; \infty$</td>
<td>$R(\lambda) &gt; 0$</td>
</tr>
<tr>
<td>$U(z : I) = \frac{1}{2\pi i} \int_{</td>
<td>\varphi</td>
</tr>
<tr>
<td>$V_{H}(z : I) = \frac{1}{2\pi i} \int_{</td>
<td>\varphi</td>
</tr>
<tr>
<td>$V_{\infty}(z : I) = 1 + \frac{I}{\mu} (1 - z^{-1}) + \frac{z^{-1}(1 - z)^{2}}{2\pi i \mu} \int_{</td>
<td>\theta</td>
</tr>
</tbody>
</table>

$U(z : I)$ and $V_{H}(z : I)$ have radii of convergence $1/j_{0}$.

Where the contour of integration is denoted by $|\theta| < 1$, this is to be interpreted as a simple contour within the unit circle.

(d) From the result in the discrete case (§18 ii) and (b) and (c) above, we are led to believe that, for any admissible $j(t)$ continuous or otherwise, $\varphi$ is always $1/j_{0}$. As an example, consider the P.D. defined by

$$\text{Prob} \{0 < t < T\} = (1 - e^{-T})(1 - j_{0}),$$

$$\text{Prob} \{t = 0\} = j_{0}.$$

It is easily shown that

$$U(z : I) = \frac{1 - j_{0}}{1 - z j_{0}} \exp \left\{ - \frac{I(1 - z)}{1 - z j_{0}} \right\},$$

which has $\varphi = 1/j_{0}$ (including the case $j_{0} \to 0$).
Consider now the most general case where the distribution function of the P.D. denoted by \( F(t) \) may have an infinity of discontinuity points anywhere, \( F(0) \) being \( j_0 \), the probability mass at the origin. It is known (Cramér (1945) §6.6) that \( F(t) \) is always everywhere continuous to the right, and in particular therefore at \( F(0) \).

Hence, given \( \varepsilon > 0 \), there always exists a \( \delta = \delta(\varepsilon) > 0 \) such that \( F(t) < F(0) + \varepsilon \) for \( 0 < t < \delta \). Hence, if the step-function \( \mathcal{F}(t) \) is defined as \( j_0 + \varepsilon \) for \( 0 < t < \delta \) and 1 for \( \delta < t < \infty \), it follows that \( \mathcal{F}(t) > F(t) \) for \( 0 < t < \infty \).

By theorem 8, the respective \( n \)-fold convolutions satisfy \( \mathcal{F}_n(t) \geq F_n(t) \).

Since

\[
u^*(n, I) = F_n(I -)
\]

it follows that

\[
u^*(n, I) \leq \mathcal{F}_n(I -) = \mathcal{F}_n(n, I),
\]

the function which would correspond to \( u^* \) if the P.D. were \( F \).

The radius of convergence of \( U^*(z : I) \) is therefore not less than that of \( \mathcal{F}^*(z : I) \), and since (by theorem 5) \( U^* \) and \( U \) have the same radius of convergence, it follows that \( \rho \) is not less than the radius of convergence of \( \mathcal{F} \). Hence, \( \rho \) is not less than \( 1/j_0 \). Since \( \varepsilon \) is arbitrarily small it follows that \( \rho \) is not less than \( 1/j_0 \).

Now

\[
u(n + 1, I) = \int_0^I u(n, I - t) dF(t) = u(n, I) j_0 + \int_0^I u(n, I - t) dF(t) \geq u(n, I) j_0.
\]

It follows that

\[
u(n, I) u(n + 1, I) \leq 1/j_0,
\]

which is therefore an upper bound to \( \rho \). The theorem that \( \rho = 1/j_0 \) is thus completed.

45. The Census Distribution for an Interval beginning at an arbitrary point.—The observational interval \( OI \) (which does not include \( I \) as end point) is of length \( I \), and the period \( BO \) preceding the start of observation is of length \( H \).

The probability of exactly \( n \) landings in the observational interval is denoted by \( v_H(n, I) \).

Clearly

\[
v_H(n + 1, I) = \int_0^I k_H(t) u(n, I - t) dt, \quad n = 0, 1, 2, \ldots
\]

and

\[
v_H(0, I) = \int_I^\infty k_H(t) \ dt = k_H^*(I).
\]

(107)

Because \( u(n, I - t) = 0 \) for \( t > I \), the upper limit of the first integral may be replaced by \( \infty \), and the integral equation solved by taking L.T.'s with respect to \( I \).

If we put

\[
\varepsilon_H(n : \lambda) = \int_0^\infty e^{-\lambda I} v_H(n, I) \ dI,
\]

then

\[
\varepsilon_H(0 : \lambda) = \mathcal{H}_H^*(\lambda) = [1 - \mathcal{H}_H(\lambda)]/\lambda,
\]

\[
\varepsilon_H(n + 1 : \lambda) = \mathcal{H}_H^*(\lambda) \mathcal{H}(n : \lambda),
\]

where \( \mathcal{H}(n : \lambda) \) is defined in §43 and given by (104).
Denoting
\[ \sum_{n=0}^{\infty} \mathcal{P}_H(n : \lambda) z^n \] by \( V_H[z, \lambda] \),
it follows that
\[ V_H[z, \lambda] = \mathcal{F}_H^*(\lambda) + z\mathcal{F}_H(\lambda) \mathcal{B}[z, \lambda] \]
\[ = \frac{(z-1)\mathcal{F}_H(\lambda) + 1 - z\mathcal{B}(\lambda)}{\lambda[1 - z\mathcal{B}(\lambda)]}, \] (108)
whence
\[ V_H(z : I) = \frac{1}{2\pi i} \int_{y-i\infty}^{y+i\infty} \frac{e^{\lambda} (z-1)\mathcal{F}_H(\lambda) + 1 - z\mathcal{B}(\lambda)}{\lambda[1 - z\mathcal{B}(\lambda)]} d\lambda. \] (109)

46. The Limiting form of \( v_H(n, I) \) as \( H \to \infty \).—It is clear from (107) or (109) that \( v_H(n, I) \) will tend to a limit if and only if \( k_H \) does, and this requires the finiteness of the mean \( \mu \) of the P.D.

From (88b) and (109) it follows that
\[ V_H(z : I) = \frac{1}{2\pi i} \int_{y-i\infty}^{y+i\infty} \frac{e^{\lambda} (z-1)\mathcal{F}_H(\lambda) + 1 - z\mathcal{B}(\lambda)}{\lambda[1 - z\mathcal{B}(\lambda)]} d\lambda, \] (110)
which may be arranged as
\[ V_H(z : I) = 1 + \frac{1}{\mu z} \int_{y-i\infty}^{y+i\infty} \left[ \frac{(z-1)[1 - \mathcal{B}(\lambda)] + 1}{\lambda z[1 - z\mathcal{B}(\lambda)]} \right] e^{\lambda} d\lambda. \] (111)

The substitution \( z = 1 + \kappa \) converts (111) into a factorial moment g.f. It is immediately apparent that the mean \( \langle v^0 \rangle \) is \( I/\mu \), a result which also follows from first principles since
\[ \langle v^0 \rangle = \int_0^I \pi(t | - \infty) dt, \text{ and } \pi(t | - \infty) = 1/\mu. \]

47. Radius of convergence of \( V_H(z : I) \).—By comparing (109) with (106) it will be apparent that the radius of convergence of \( V_H(z : I) \) is the same as that of \( U(z : I) \) namely \( 1/\rho \). The result also holds as \( H \to \infty \) if \( v_H \) tends to a definite limiting form.

48. Distributions arising from an Alien first jump.—Where in a stochastic process the recurrent event under consideration is not the initial state, the distribution of the time of first occurrence \( a(t) \) will in general differ from \( j(t) \). Such for example occurs in the theory of counters discussed by Feller (1948). The modifications to our results necessitated by such a consideration are easily dealt with. In particular the first census distribution \( [\nu(n)] \) assumes the mathematical form of the second \( [v(n)] \) with \( a(t) \) substituted for \( k(t) \) throughout.

Summary

The main results in both the discrete and continuous cases are summarized side by side in Tables 8A and 8B.

Acknowledgments

We are greatly indebted to Mr. A. Wrigley for the loan of his detailed records of cricket, for without accurate descriptions of the game ball by ball the analysis in paragraph 21 would not have been possible. We gratefully acknowledge also the assistance given by Miss J. R. Proctor in preparing diagrams and handling numerical work.
Guide to Symbols and Notation Used

<table>
<thead>
<tr>
<th>Discrete</th>
<th>Continuous</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P.D.</td>
<td>P.D.</td>
<td>Primary Distribution.</td>
</tr>
<tr>
<td>C.D.</td>
<td>C.D.</td>
<td>Census Distribution.</td>
</tr>
<tr>
<td>$v_H(n, I)$</td>
<td>$f_H(\lambda)$</td>
<td>Truncated g.f.: “negative tail”.</td>
</tr>
<tr>
<td>$p_H(0)$</td>
<td>$f_H(\lambda)$</td>
<td>Truncated g.f.: “positive tail”.</td>
</tr>
<tr>
<td>$p_H(t)$</td>
<td>$f_H(\lambda)$</td>
<td>Distribution function for “positive tail”.</td>
</tr>
<tr>
<td>$q_H(t)$</td>
<td>$f_H(\lambda)$</td>
<td>Distribution function for “negative tail”.</td>
</tr>
<tr>
<td>$j_H(t)$</td>
<td>$J_H(\lambda)$</td>
<td>Jump probability.</td>
</tr>
<tr>
<td>$\pi(t)$</td>
<td>$\pi(t)$</td>
<td>Mean, factorial moment, crude moment, central moment of P.D.’s.</td>
</tr>
<tr>
<td>$\Pi(\theta)$</td>
<td>$\Pi(\lambda)$</td>
<td>Mean, factorial moment, crude moment, central moment of C.D.’s.</td>
</tr>
<tr>
<td>$\pi(t</td>
<td>x)$</td>
<td>$\pi(t</td>
</tr>
<tr>
<td>$\Pi(\theta</td>
<td>x)$</td>
<td>$\Pi(\lambda</td>
</tr>
<tr>
<td>$o(n : I)$</td>
<td>$\Omega(z : I)$</td>
<td>Distribution of landings at a point.</td>
</tr>
<tr>
<td>$k_H(t)$</td>
<td>$k_H(t)$</td>
<td>Distribution of first landing outside an interval.</td>
</tr>
<tr>
<td>$k_H(0)$</td>
<td>$k_H(0)$</td>
<td>Distribution of the part of the last jump included in the interval.</td>
</tr>
<tr>
<td>$\tau_z$</td>
<td>$\tau_z$</td>
<td>Mean and variance for $K_H(z : I)$.</td>
</tr>
<tr>
<td>$k_H^{(n)}(t)$</td>
<td>$k_H^{(n)}(t)$</td>
<td>Distribution of the “spanning” jump.</td>
</tr>
<tr>
<td>$k_H^{(n)}(0)$</td>
<td>$k_H^{(n)}(0)$</td>
<td>Distribution of number of landings in an interval.</td>
</tr>
<tr>
<td>$\sigma_k(t)$</td>
<td>$\sigma_k(t)$</td>
<td>Distribution of number of landings in an interval from a distant source.</td>
</tr>
<tr>
<td>$S_{k_H}(0)$</td>
<td>$S_{k_H}(0)$</td>
<td>Distribution of number of landings at a point.</td>
</tr>
<tr>
<td>$\nu(n, I)$</td>
<td>$\nu(n, I)$</td>
<td>Bivariate g.f. over intervals and number of landings.</td>
</tr>
<tr>
<td>$U(z : I)$</td>
<td>$U(z : I)$</td>
<td>Mean, factorial moment, crude moment, central moment of C.D.’s.</td>
</tr>
<tr>
<td>$p$</td>
<td>$p$</td>
<td>Radius of convergence of g.f. of C.D.</td>
</tr>
<tr>
<td>$V_H(n, I)$</td>
<td>$V_H(n, I)$</td>
<td>Distribution of number of landings in an interval from a distant source.</td>
</tr>
<tr>
<td>$V_H(z : I)$</td>
<td>$V_H(z : I)$</td>
<td>Distribution of number of landings from a distant source.</td>
</tr>
<tr>
<td>$V_H(z, \theta)$</td>
<td>$V_H(z, \theta)$</td>
<td>Bivariate g.f. over intervals and number of landings from a distant source.</td>
</tr>
<tr>
<td>$\nu_0^0$, $\nu_2^0$</td>
<td>$\nu_0^0$, $\nu_2^0$</td>
<td>Mean and variance for $V_H(z : I)$.</td>
</tr>
</tbody>
</table>

References


DISCUSSION ON MR. SKELLAM’S AND DR. SHENTON’S PAPER

Mr. R. L. Plackett: Recent work in renewal theory has been mainly concerned with the convergence of certain functions of time, such as the renewal density, towards their limiting values, and I think that the chief contribution of the authors in their reconsideration of some standard renewal problems lies in their emphasis on exact results for a finite interval of observation at a finite time from the source. The practical applications seem to be severely limited, but when illustrations are given, every student of probability texts will be grateful for the imagination of the authors in conducting him to the fresh breezes of the cricket field rather than the stale air of the gambling saloon.

The following remarks are confined to Part I. I begin by rejecting the authors’ convention in equation (10) whereby $\Pi(0) = 1/[1 - J(0)]$, because it implies that the expected number of landings at the point $t$ is $\pi(t)$ for $t > 0$ and $\pi(0) - 1$ for $t = 0$; and this entails trivial but tiresome distinctions between $H > 0$ and $H = 0$. Instead, I shall take $\Pi(0) = J(0)/[1 - J(0)]$, as in equation (9), which is exactly analogous to the usual formula relating Laplace transforms in continuous time. The expected number of landings is now $\pi(t)$ for $t > 0$.

An obvious feature of the paper is the large number of contour integrals, but the feeling of apprehension, which they induce in a reader not accustomed to seeing them marshalled in such strength, gradually fades with the realization that in many instances they are merely an elaborate way of specifying a coefficient in a series expansion, and that the extent to which they are introduced has not altogether been justified by the results obtained from them, although the ingenious arguments of Section 24, for example, are interesting in themselves. Indeed, contour integrals are capable of obscuring results which are essentially quite simple. Take, for example, the distribution of ON, which can be described in the parlance of industrial renewal as the residual life-time of the article in use at time $t = 0$. The authors express its probability generating function as a contour integral and the probabilities themselves as an infinite series of the type used by Alfred J. Lotka in his method for solving the integral equation of renewal theory. In fact, $k_H(t)$ can be obtained directly by subtracting from the renewals at $t' = H + t$ all those which have arrived from the interval $t = (H, H + t)$, so that

$$k_H(t) = \pi(H + t) - \sum_{y=0}^{\infty} \pi(H + y) (t - y).$$

The usual manipulations now show that the probability generating function satisfies

$$\theta^H K_H(0) = J(0) - \{1 - J(0)\} [\pi(0) + \pi(1) \theta + \cdots + \pi(H - 1) \theta^{H - 1}],$$

from which the factorial moments arise by successive differentiations at $\theta = 1$; or alternatively

$$K_H(0) = \{1 - J(0)\} [\pi(0) + \pi(H + 1) \theta + \pi(H + 2) \theta^2 + \cdots],$$

from which the limiting form of $k_H(0)$ is readily seen to be $\{1 - J(0)\}/\mu(1 - 0)$. If $K_H(0) = J(0)$ then

$$\pi(s) = \pi(s + H) = \pi(s + 2H) = \cdots$$

for $0 < s < H - 1$.

Since $\pi(t)$ tends to a limit, it must be constant, which proves the conjecture at the end of Section 9.

Another example where we come to a sharp halt with a contour integral occurs at equation (46), and it seems a pity to leave the exact expressions for the moments of the census distribution at this stage, because their analytical form is not uninteresting. Differentiating (39) $r$ times with respect to $z$ at $z = 1$,

$$\sum_{i=1}^{\infty} \frac{1}{i!} (i) 0^i = r! \left[ \frac{\theta}{1 - \theta} \right]^{i} \Pi'(0).$$
Discussion on Mr. Skellam’s and Dr. Shenton’s Paper

where \( v_0(I) \) denotes the 0th factorial moment of the census distribution for the interval \((0, I)\). In words, \( v_0(I) \) is \( I! \) times the sum of the first \( I \) terms of the \( r\)th convolution of the sequence \( \{\pi(t)\} \) with itself. We thus obtain the recurrence formula

\[
v_0(I) = r \sum_{i=0}^{I-1} \pi(i) v_{r-1}(I - i),
\]

in which \( v_0(I) = 1 \).

The limiting form of \( v_0(I) \) as \( I \to \infty \) can now be derived by finding the principal part of \( r! \pi(0)/(1 - 0) \) at \( \theta = 1 \). A simple way of doing this is to note that

\[
\mu(1 - \theta) \Pi(0) = 1 + c_1 \left( \frac{1 - \theta}{\mu} \right) + c_2 \left( \frac{1 - \theta}{\mu} \right)^2 + \ldots
\]

The coefficients required for the first four factorial moments are

\[
c_1 = \left( \mu(2) - 2\mu^2 \right)/2,
\]

\[
c_2 = \frac{3\mu(3)^2 - 2\mu(4)}{12},
\]

\[
c_3 = \frac{3\mu(4)^3 - 4\mu(4)\mu(5) + \mu^2(6)}{24},
\]

and

\[
c_4 = \frac{45\mu(5)^4 - 90\mu(5)^2\mu(6) + 20\mu^2(5)^2 + 30\mu^2\mu(5)\mu(4) - 6\mu^2(6)}{720}.
\]

For example, take \( r = 2 \). Then

\[
\sum_{i=1}^{\infty} v_2(i) \theta^l = 2 \left( \frac{0}{1 - 0} \right) \frac{1}{\mu(1 - 0)^2} \left[ 1 + 2c_1 \left( \frac{1 - \theta}{\mu} \right) + (2c_2 + c_1^2) \left( \frac{1 - \theta}{\mu} \right)^2 + \ldots \right],
\]

whence

\[
v_2(I) = \frac{I(I + 1)}{\mu^2} + \frac{4c_1I}{\mu^3} + \frac{4c_2 + 2c_1^2}{\mu^4},
\]

which agrees with (48). Similar, but more tedious, calculations show that the third cumulant of the census distribution has a limit of the form \( aI + b \) with \( a \neq 0 \).

With a source which precedes the origin by \( H \) units, denote the factorial moments of the census distribution by \( v_0(I; H) \). The number of landings in the interval of observation is

\[
n = n(0) + n(1) + \ldots + n(I - 1),
\]

where \( n(t) \) is the number of landings at time \( t \). The \( r\)th factorial moment of \( n(t) \) is \( r! \pi^{r-1}(0) \pi(t) \), and when \( t < u < v < \ldots \),

\[
\mathcal{E}\{n^r(t;n^u(t))n^r(v)\ldots\} = \mathcal{E}n^r(t)\mathcal{E}n^r(u - i)\mathcal{E}n^r(v - u)\ldots
\]

because the non-zero contributions to the expectation arise solely when there are landings at \( t \) and \( u \) and \( v \ldots \), and the process is regenerated at each landing point. These results show that shifting the source affects only a single argument of \( \pi(t) \) in \( v_0(I) \), and the analogy with \( H = 0 \) then indicates that

\[
v_0(I; H) = r \sum_{i=0}^{I-1} \pi(H + i) v_{r-1}(I - t; 0).
\]

That this is consistent with (68) appears on reorganizing that equation in the form

\[
V_H[z, \theta] = \frac{\theta}{(1 - \theta)} \left( 1 - \frac{1}{\Pi_H(0)/\Pi(0) + \Pi_H(0)/\Pi(0) \cdot 1 - (z - 1)/(\Pi(0))} \right),
\]

where

\[
\Pi_H(0) = \pi(H) + \pi(H + 1) 0 + \pi(H + 2) 0^2 + \ldots
\]

Equations (75) and (76) now follow from the general expression

\[
\lim_{H \to \infty} v_0(I; H) = \frac{r}{\mu} \sum_{i=1} v_{r-1}(s; 0).
\]

I have very much enjoyed working over the problems which the authors have brought before us this evening and I propose that we accord them a hearty vote of thanks.
Dr. D. R. Cox: It is a privilege to have the chance of congratulating Mr. Skellam and Dr. Shenton on a substantial and interesting contribution to an important problem in probability theory. My main criticism of their paper is of the relative failure to explain the relation between their work and numerous earlier investigations of the problem. For instance, the quantities \( \text{OL} \) and \( \text{ON} \) are backward and forward recurrence times and have been studied by von Smoluchowski, Doob, Bartlett and others, and a critical comparison of the present authors' new results with the earlier work would, I am sure, have been most instructive.

The authors' methods are satisfyingly self-contained, but there is one point in particular at which a different approach would have had some advantages. In Section 24, the asymptotic normality of the census distribution is studied by an interesting direct argument. However, Feller, in his book, pointed out that the probability that fewer than \( n \) landings occur in the interval \( \text{OL} \) is equal to the probability that the sum of \( n \) independent random variables, each with the distribution of the individual jumps, exceeds the length \( \text{OL} \). The central limit theorem can be used directly to approximate to this latter distribution, and, provided only that the basic distribution belongs to the domain of attraction of the normal law, asymptotic normality follows. (The authors' proof requires that the radius of convergence of the p.g.f. of the jump distribution exceeds unity, so that, in particular, all moments must exist.) If closer approximations to the census distribution are required, they can, of course, be obtained under weak conditions by an Edgeworth asymptotic expansion for the probability just mentioned.

Much of the paper is concerned with the manipulation of probability generating functions and Laplace transforms and it is natural to consider whether such generating functions can be given a physical significance enabling equations for them to be put down directly. One method of doing this for p.g.f.'s is by van Dantzig's method of collective marks. Consider a single jump in the discrete case, with probability distribution \( j_n \). (Let us call the event concerned a renewal.) Suppose also that at each point 1, 2, ... , \( n \) there is a probability \( 1 - z \) that a "catastrophe" will occur, the probabilities being mutually independent. Then the probability that the renewal occurs at \( n \) and precedes the first catastrophe is \( j_n z^n \). Hence the probability that the renewal precedes the catastrophe is \( \Sigma j_n z^n \), the p.g.f.

I do not know whether van Dantzig has extended this to deal with Laplace transforms, but the method of doing this is obvious. Consider a single renewal (\( R \)) occurring at a time from the origin with p.d.f., \( f(t) \), (absolute continuity is not necessary). Consider also a catastrophe (\( C \)) occurring in a Poisson process with parameter \( \lambda \). The probability that \( R \) occurs in \( (t, t + dt) \) not preceded by \( C \) is \( e^{-\lambda t} f(t) \) \( dt \). Hence the total probability that \( R \) precedes \( C \) is

\[
\Pr (R < C) = \int_0^\infty e^{-\lambda t} f(t) \, dt
\]

the Laplace transform of \( f(t) \).

Further if \( J(t) \) is the distribution function corresponding to \( f(t) \), the probability \( (1) \) can be obtained by arguing that the probability that \( C \) occurs in \( (t, t + dt) \) is \( \lambda e^{-\lambda t} dt \) and that it does so preceded by \( R \) is \( J(t) \lambda e^{-\lambda t} dt \). Hence

\[
\Pr (R < C) = \int_0^\infty \lambda J(t) e^{-\lambda t} \, dt;
\]

and the equality of \( (1) \) and \( (2) \) gives a probabilistic proof of the standard relation for the Laplace transform of an integral.

We need the Laplace transforms not only of p.d.f.'s but also of the renewal intensity (accumulated probability). In the continuous case, the authors denote this by \( \pi(t) \); it is such that \( \pi(t) \) \( dt \) is the expected number of renewals, of whatever serial number, in \( (t, t + dt) \), or alternatively is the derivative of the expected number of renewals in \( (0, t) \). The chance that \( C \) occurs in

\[
(t, t + dt) \]

and the expected number of \( R \)'s before this is \( \int_0^t \pi(x) \, dx \).

Hence the overall expected number of \( R \)'s preceding the first \( C \) is

\[
\int_0^\infty dt \int_0^t dx \pi(x) \lambda e^{-\lambda t} = \int_0^\infty \pi(t) e^{-\lambda t} \, dt = \Pi(\lambda),
\]
the Laplace transform of $\pi(t)$. Now the mean $\Pi(\lambda)$ has zero contribution if $C$ occurs before the first $R$ (chance, $1 - f(\lambda)$) and hence is

$$f(\lambda)(1 + \text{expected number of further } R\text{'s before a } C\text{, given that one } R \text{ has just occurred and no } C).$$

The regenerative nature of the processes leads us to

$$\Pi(\lambda) = f(\lambda)(1 + \Pi(\lambda)). \quad \quad \quad \quad \quad \quad (3)$$

Equation (3) is the transformed version of the integral equation of renewal theory; the slight difference from the authors' (81) arises because they have included the $R$ occurring at the origin in the definition of $\pi(t)$. It is likely that many relations connecting Laplace transforms and g.f.'s can be derived by this sort of argument.

Now while this method seems to me an interesting way of looking at these problems, it certainly does not appear from the above example that there is any very solid gain achieved by using it. But in some complicated problems, van Dantzig's method for p.g.f.'s does seem to simplify appreciably the writing down of recurrence relations. A very interesting example is a recent investigation by H. Kesten and J. Th. Runnenburg, of the Mathematical Centre at Amsterdam, of priority queueing problems.

I have great pleasure in seconding the vote of thanks to the authors for their very enjoyable paper.

The vote of thanks was put to the meeting and carried unanimously.

Dr. W. L. Smith: Dr. Cox has forestalled me on a number of points. There is much that I would like to say about this paper, but the Section's very sensible time restriction on contributions to the discussion forces me to limit my remarks to one or two aspects of it. The crucial theorem in recurrent-events theory is the convergence theorem:

$$\pi(t) \rightarrow \mu^{-1} \text{ as } t \rightarrow \infty.$$  

The difficult part of this theorem is the verification that $\pi(t)$ does indeed tend to a limit. The evaluation of this limit can be quite simple. Skellam and Shenton correctly credit Kolmogorov and Erdős, Feller and Pollard with the proof that this limit exists, but omit mentioning that the latter gentlemen also evaluate it. I am wondering why it has been thought worth while to devote so much space to this limit's re-evaluation, especially as we are not provided with a proof of its existence and as the power series argument employed has so much in common with the arguments in the first half of the paper by Erdős, Feller and Pollard. Similarly, in their discussion of the continuous process, Skellam and Shenton credit earlier authors with the proof that the relevant limit exists, but, curiously, omit to mention that these authors had also evaluated it. We are provided instead with a somewhat lengthy re-evaluation of this limit. Surely the simplest way to discover the value of the limit, if one is prepared to accept its existence, is by the kind of argument used by Doob in Trans. Amer. Math. Soc. (1948), 63, pp. 422-438 and by Blackwell in Pacific J. Math. (1953), 3, pp. 315-332, based on the law of large numbers. Incidentally, the approach taken in the present paper to the continuous case calls for, not Blackwell's renewal theorem, but the renewal density theorem which was first proved by Feller in Ann. Math. Statist. (1941), 12, pp. 243-267 and which is given in its most general form to date in my paper in Proc. Camb. Phil. Soc. (1955), 51, pp. 629-638.

To deal very briefly with some further points: The quantity LO is the same as the renewal age, in renewal theory, and as such is subjected to an extensive and deep analysis in the paper by Doob, to which I have already referred. The present paper ignores this work of Doob, and also the general proof of the convergence of the age-distribution given in Proc. Roy. Soc. Edinb. A (1954), 64, pp. 9-48. The quantities LO and ON are discussed under the names "backward and forward delay" by Cox and Smith in Biometrika (1954), 41, pp. 91-99, where it has already been announced that these quantities have the same ultimate distributions. The fact that all the moments of the census distribution are finite is well known, and follows at once from a result in sequential analysis due to Charles Stein, Ann. Math. Statist. (1946), 17, pp. 498-499, which is used by Blackwell in the proof of his renewal theorem. The asymptotic normality for the census distribution has already been proved in an extremely brief and elegant way by Feller, in Trans. Amer. Math. Soc. (1949), 67, pp. 98-119 under considerably weaker hypotheses than the ones used here. Lastly, I would like to draw attention to the prolific writings of the Hungarian, Lajos Takács, much of which has a direct bearing on the present subject.

Some of Mr. Plackett's remarks prompt me to mention a small investigation which I have just completed concerning the cumulants of a renewal process. I was most interested to learn...
Mr. Plackett's methods, but it does seem to me that they suffer from the mathematical disadvantage (and I am prepared to be corrected on this point if I am mistaken) that they require all the moments \( \mu_n \) of the renewal "lifetimes" distribution to be finite. To prove, say, that the variance of \( N_t \) (the number of renewals in \((0, t)\)) has a certain asymptotic form involving \( \mu_1, \mu_2 \) and \( \mu_3 \), I feel one should not have to assume the finiteness of \( \mu_n \) for all \( n > 3 \). Let us write \( \kappa_n(t) \) for the \( r \)th cumulant of \( N_t \) and \( P(x) \) for the distribution function of the sum of the first \( j \) renewal lifetimes. Suppose, for some \( j \), \( P(x) \) has an absolutely continuous component. Then I have proved that if \( \mu_{n+r+1} < \infty \) (\( n > 0 \)),

\[
\kappa_n(t) = \alpha_n t + \beta_n + \Omega_n(t),
\]

where both \( \Omega_n(t) \) and \( r^n \Omega_n(t) \) are functions of bounded variation which approach zero at infinity. The coefficients \( \alpha_n \) and \( \beta_n \) are functions of the first \((n+1)\) moments \( \mu_n \). I hope to publish this work soon, together with a table of the \( \alpha_n \) and \( \beta_n \) up to a reasonably high value of \( n \).

DR. D. E. BARTON: I should like to comment on a minor question raised in this paper, namely the semi-normal limit to a census distribution. This has been given in lectures in the Statistics Department at University College in the last few years as a direct corollary to the work of Feller. Feller's own proof (1949) is rather indirect. Since the authors of the paper just read have not given their proof perhaps it would be of value if I give our Fellerian proof, more especially since it leads to interesting generalization.

Feller's result may be re-written

\[
\Pi(z) = \left[ \frac{1}{n!} \frac{\partial^n}{\partial s^n} \left\{ (1 - s)[z + (1 - z) \sum_{j=0}^{\infty} u_j s^j]^{-1} \right\} \right]_{s=0},
\]

where

(i) \( \Pi(z) \) is the probability generating function of the number, \( r \), of events \( \varepsilon \) occurring in the set of time points \( 1, \ldots, n \);

(ii) the recurrent events \( \{\varepsilon\} \) are periodic (and the times of possible occurrence are measured in the scale of the period);

(iii) \( u_j = P(\varepsilon \text{ occurs at the } j\text{th time-point}) \), \( j > 1 \),

\[ u_1 = 1, \quad u_0 = 0. \]

When \( \varepsilon \) is the event that the number of heads equals the number of tails in a sequence of independent throws of an unbiased coin

\[ u_j = 2^j c_j \left( \frac{1}{4} \right)^j, \]

\[
\Pi(z) = \left[ \frac{1}{n!} \frac{\partial^n}{\partial s^n} \left\{ \frac{1}{\sqrt{1 - s}} \right\} \left\{ 1 - z(1 - \sqrt{1 - s}) \right\} \right]_{s=0},
\]

whence

\[
\delta(r^{(m)}) = \frac{m!}{n!} \left[ \frac{\Delta^n}{\Gamma(1 + x/2)} \right]_{x=0} = \frac{m!}{\Gamma(1 + m/2)} (1 + 0(n^{-1})).
\]

It follows that the \( m^{\text{th}} \) moment of \( r/\sqrt{2n} \) converges to that of a semi-normal variable. Thus \( x = r/\sqrt{2n} \) tends to be so distributed in the limit as \( n \to \infty \), by the Second Limit Theorem.

In essentially the same way, if

(i) the coin has respective probabilities \( a/(a+b) \), \( b/(a+b) \) of turning up heads and tails (where \( a \), \( b \) are co-prime integers);

(ii) \( \varepsilon \) is the event that in a sequence of \( i(a+b) \) throws there are \( ia \) heads (for some integer \( i \));

(iii) \( r \) is the number of occurrences of \( \varepsilon \) in a sequence \( i = 1, \ldots, n \);

then \( x = r/\sqrt{ab/(a+b)}n \) tends to a semi-normal limit as \( n \to \infty \).

This convergence might, on general grounds, be expected to be even slower than in the symmetrical case \( a = 1 = b \). But this does not seem to be the case. For instance, when \( a = 1 \), \( b = 2 \), \( n = 10 \) the mean and variance of \( x \) are 0.610, 0.2776 compared with the corresponding values 0.604, 0.2302 \((a = 1 = b, n = 10)\) and 0.798, 0.3635 \((n = \infty)\).

The generalized distribution is of great interest in examining the classical dice-throwing experiments. Thus Kerrich found that in 498 independent observations of an event, ""R, R"" whose theoretical probability was 1/6 there were four occasions when exactly one-sixth of the events were ""R, R"". This corresponds to \( n = 83, a = 1, b = 5, x = 0.40 \) and for a semi-normal variable \( P(x < 0.40) = 0.31 \) so that this result is by no means unreasonable.
Dr. C. L. Mallows: On page 77, equation (39b), the authors give a decomposition of the census distribution in the discrete case into the sum of a number of Pascal distributions. It seems of interest to consider a corresponding decomposition in the continuous case, this time into Poisson components; thus we may consider the situation where it is possible to write
\[ u(n, I) = \text{Prob. of } n \text{ events in interval } I \text{ (start at origin)} \]
\[ = \int \frac{m^n e^{-m}}{n!} dP_I(m), \]
where \( P_I(m) \) is the distribution function of the Poisson parameter \( m \). If this is possible, the underlying process may be regarded as being a heterogeneous mixture of purely random components.

Writing the L.T. of \( P_I(m) \) as
\[ P(m : \xi) = \int e^{-\xi t} P_I(m) dI. \]
We find from the equation next after (105),
\[ P(m : \xi) = \frac{1}{\xi} \left( 1 - \exp \left\{ -\frac{1 - f(\xi)}{f(\xi)} m \right\} \right), \]
where \( f(\xi) \) is the L.T. of the primary distribution. Thus for example if the P.D. is the negative exponential, with mean \( \mu \), we have
\[ f(\xi) = \frac{1}{1 + \mu \xi}, \]
whence the \( P_I(m) \) distribution is concentrated at \( m = I/\mu \).

Similarly for the case when the start is at \( -\infty \), we find from (110) with \( Q \) replacing \( P \),
\[ Q(m : \xi) = \frac{1}{\xi} - \frac{1}{\mu \xi} \frac{1 - f(\xi)}{f(\xi)} \exp \left\{ -\frac{1 - f(\xi)}{f(\xi)} m \right\} \] (2)

Now the interesting question is: when can \( P_I(m) \) and \( Q_I(m) \) be distribution functions? That is, when are the weights attached to the different \( m \)'s all \( \geq 0 \)? I remark that the decomposition for the discrete case given in the paper introduces weights \( A_k \) which may have either sign. This is a difficult one, and I can only indicate a way of deciding when the distributions certainly cannot exist. To do this, observe that a necessary (but not sufficient) condition that \( P_I(m) \) is a distribution function for all \( I > 0 \) is that \( \xi P(m : \xi) \) is a d.f. for all \( \xi > 0 \). This is always true in case (1), since \( (1 - f(\xi))/f(\xi) \) must be positive. However for case (2) \( \xi Q(m : \xi) \) is a d.f. only if
\[ 0 < \frac{1}{\mu \xi} \frac{1 - f(\xi)}{f(\xi)} \leq 1, \]
which requires
\[ f(\xi) > 1/1 + \mu \xi \text{ for all } \xi > 0. \]
An example when this condition is not satisfied is given by the strictly periodic process, each step being of length exactly \( \mu \). Then
\[ f(\xi) = e^{-\mu \xi}, \]
and the condition is not satisfied.

Professor Bartlett: A number of papers additional to the ones contained in the bibliography have been mentioned and all I wanted was to draw attention to a paper* by A. R. G. Owen which I mentioned in my 1949 Stochastic Processes paper. Owen was considering the number of points of interchange of the chromosome; the theory in his paper contains such distributions as the census distribution, not only for an interval starting from the origin but also for an interval starting away from the origin (in the terminology of the present paper).

Dr. I. J. Good: Suppose we put a zero at all points jumped over and a 1 at all points landed on. Then we get a random sequence if and only if the primary distribution is geometric for positive steps.

It then becomes obvious that the distribution of ON is the same as the distribution of the distance between 1's. This result was proved as a special case of the general theory at the end of Section 9. The result is analogous to the fact, proved I believe in Whitworth's *Choice and Chance*, that if \( k \) points are distributed uniformly on a circle that is suspended by a hook, then the distribution of the distance from the hook to the nearest point on its right is the same as that between any two consecutive points. This is clear, because, for all the circle knows, the hook is merely a \((k + 1)\)th point.

The random walk in Section 22, with reflection at the origin, is more strongly related to coin-tossing than the authors imply. The distribution of the number of returns to the origin must be the same for the doubly infinite random walk as for the one with reflection, because the distribution cannot be affected by the direction of the steps taken at the origin. The probability distribution of the time taken to return to the origin depends only on the distance from the origin but is independent of the sign of this distance. (This would be true for any symmetric random walk.) Incidentally the number of random walks with reflection at the origin, with \( n \) returns to the origin and finishing at the origin, is equal to the number of ordered trees in which the number of descendants is \( n \). This was pointed out and explained by T. E. Harris (*Trans. Amer. Math. Soc.* 73 (1952), 471-483).

The authors subsequently replied in writing as follows:

We greatly appreciate Dr. Plackett's remarks and value his constructive suggestions. The arguments and positive results which he has provided, make it amply clear in this field as in most others, that a variety of methods of approach are to be welcomed, no single one enjoying exclusive priority, and that devices and conventions introduced for simplicity in one approach may be troublesome in another. It is perhaps inevitable that different workers will prefer different tools, and that what appears simple to one may seem complicated or sophisticated to another, according to the familiarity or facility with which the tools are employed. We are grateful to Dr. Plackett for indicating that we have made use of only the most elementary aspects of complex variable theory. We excuse the expression of our results as contour integrals on the ground that they are the natural outcome of the unified approach we had intended.

The recurrence relation which Dr. Plackett gives is undoubtedly an interesting and fruitful one and we certainly agree that the expressions for the higher cumulants of the census distributions are complicated.

Dr. Cox draws attention to Feller's work on the asymptotic approach to normality of the census distribution, and Dr. Smith also mentions the same point. Feller's argument would certainly lend itself to the refinements necessary to turn it into a rigorous proof, but even so we would agree that his line of argument is the simplest way of demonstrating asymptotic normality under wide conditions. But without wishing to disparage early efforts it must be borne in mind that such results are often in practice of rather limited value unless something is known of the rapidity or slowness of the approach to normality, or unless further terms can be obtained. The contour integration methods we have used, go some way towards meeting the difficulty, and it would certainly be interesting to see how the Gram-Charlier development envisaged by Dr. Cox compares with the somewhat similar expressions arising by the method of §23. It is also important in this connection to explore the form of distribution through which, with increasing \( l \), normality is ultimately attained, and it may be of interest to remark that for

\[
J(\theta) = \frac{(8 + 6 + \theta^2)}{10}
\]

the transitional distribution is almost exactly Pearson Type III, the values of \( 6 + 3\beta_1 - 2\beta_2 \) for \( l = 5, 10, 20, 50, 100 \) being respectively 0.07, 0.04, 0.04, 0.02, 0.00. Little appears to be known on this important practical aspect of the subject, but the remarkable and surprising phenomenon of semi-normality (§22) should be sufficient to evoke a note of caution.

Dr. Cox has provided an extremely interesting physical interpretation of various expressions in renewal theory in which Laplace transforms are involved, and we are sure that all interested in this field must feel grateful to him. The model he constructs is by no means as artificial or divorced from real problems as might at first sight be surmised, and indeed, in many problems of polymer disintegration where the molecular chain lengths are exponentially distributed, we have already encountered a closely parallel situation.

It is unfortunate that our wording in theorems 7A (f) and 35 is capable of misinterpretation rendering it possible for anyone so minded to voice the insinuation which Dr. Smith has drawn, but there are sufficient references for the reader who wishes to delve into this difficult question.
to be left in no doubt as to where the credit for discovery lies. It is unfortunate however that Dr. Smith should regard this convergence theorem as "the" crucial theorem in recurrent event theory, for it plays no part in the finite problems in which we have been primarily interested. It is one thing to know that $\pi(t) \to 1/\mu$ as $t \to \infty$, and another to know how large $t$ must be before the theorem can be employed as an adequate approximation and the exact result dispensed with.

Dr. Smith regards it as well known that the moments of the census distribution are finite, but he furnishes us with only a single slanting reference. He also draws our attention to an announcement that the distributions of "backward and forward delay" are ultimately the same. Curiously enough, he omits to say that no proof is given, the result being "intuitively obvious". One might equally well argue that the so-called fundamental theorem of renewal theory, with which Dr. Smith seems preoccupied, is also intuitively obvious, but we feel sure that he would agree in this rather tricky field that intuition is no substitute for logical argument, nor oblique sources of cryptic information adequate alternatives to clearly enunciated theorems.

We were interested to learn of Dr. Barton's lucid proof of the asymptotic semi-normality of the distribution of the number of returns to equilibrium in a simple coin tossing scheme, and of his observations on the more general case. Dr. Barton judges the goodness of fit by comparison using the first two moments, but it is also of interest to fit a semi-normal curve to the discrete distribution and to consider the most appropriate position for, the origin, whether at $-\frac{1}{2}$, 0, or $+\frac{1}{2}$.

The intriguing problem which Dr. Mallows has tackled is one which we had not looked into. This problem to us seems worthy of further consideration.

We are grateful to Professor Bartlett for reminding us of A. R. G. Owen's paper on genetical and cytological cross-over, the importance of which we fully acknowledge.
THE EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

by

A.J. Howie and L.R. Shenton.

(23 pages with four line drawings)

(In the press, J.R.S.S., series B, 1959)
THE EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

by A.J. Howie and L.R. Shenton.

1. Introduction.

1.1 In certain types of winding machine used in cotton yarn preparation there is a fairly large number (usually between 100 and 300) of spindles each of which is used to wind yarn from a relatively small supply bobbin on to a larger bobbin. An automatic head patrols the machine in a fixed time and spends a fixed time servicing each spindle, i.e. replacing an exhausted supply bobbin by a full one and starting it, or knotting a yarn which has broken in re-starting it. All the unwinding is done at the same speed so we may speak of the 'length of yarn' on the bobbin and the 'unwinding time' of the bobbin, (i.e. the time required to unwind it completely in the absence of breaks) as equivalent. The supply bobbins on different spindles do not necessarily contain the same length of yarn but all the bobbins unsound successively on the same spindle have the same unwinding time. The patrolling time, i.e. the time between consecutive arrivals of the automatic head at a particular spindle can be varied within certain limits, and is set at the value which give maximum efficiency, i.e. the highest possible ratio of time in actual unwinding to total running time of the machine. In earlier models of the machine the arrival of the automatic head at a spindle which was running stopped the machine, so that the maximum efficiency was attained by making the patrolling time exceed by as little as possible the unwinding time of the largest supply bobbin on the machine.
since the amount of time lost as a result of breaks is a minimum when
the patrolling time has its smallest value. In later models, however,
the machine continues to run when the automatic head arrives at a
running spindle, so that patrolling times smaller than unwinding times
of some or all of the supply bobbins on the machine can be used. The
'servicing time', i.e. the time spent by the head at each spindle,
is the same whether the spindle is stopped or running, so the patrolling
time is strictly constant. With such machines the problem of finding
the patrolling time which will give maximum efficiency is more compli-
cated. It is convenient to consider separately two cases: (i) the
case in which there is only one spindle on the machine or there are
more than one but all supply bobbins have the same unwinding time,
random breaks in the yarn being taken into account; (ii) the case in
which supply bobbins with different unwinding times are being
simultaneously unwound on different spindles of the machine; here the
problem arises whether random breaks are ignored or taken into account.
The present paper is concerned with case (i); a later one, using the
results established here considers case (ii).

The essential features of the problem are the regular
stoppages of the spindle due to exhaustion of the supply bobbin, the
constant patrolling time, and the superimposed random stoppages due to
breaks. In the first and second of these features the situation
differs from that examined by Mack et al (1957): they do not
contemplate regular stoppages, and their patrolling time is a random
variable, being dependent on the number of machines (spindles) found
stopped during the patrol.

1.2 Fundamental Assumptions. It is assumed that breaks occur independently and with uniform probability throughout the length of the bobbin, i.e. the probability of a break in a small interval of time $\delta t$ is $\gamma \delta t + o(\delta t)$, $\gamma$ being constant: thus the number of breaks per bobbin follows a Poisson distribution with mean proportional to the length of yarn on the bobbin, and the interval between successive breaks has a negative exponential distribution. In some cases at least, this assumption corresponds fairly closely with the facts.

In practice the servicing time is of the order of 7 seconds and the patrolling time from 1.5 to 10 minutes (seldom less than 2 minutes). The servicing time does not always represent a loss of unwinding time; a spindle which is running when the automatic head arrives at it continues to unwind yarn while the head is present. In this first treatment of the problem, therefore, the servicing time will be ignored. The results obtained should be a good approximation to those attainable in practice.

1.3 Remarks on Notation. The length of yarn on the bobbin, or time required to unwind it in the absence of breaks, will be denoted by $a$, the patrolling time by $P$.

The probability that the interval between successive breaks
THE EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

exceeds \( x \) is taken to be \( e^{\mu x} \), \( \mu \) being a positive constant:

thus the number of breaks in a bobbin of unwinding time \( a \) follows Poisson's distribution with mean \( \mu a \).

Quantities \( i (=0, 1, 2, \ldots) \) and \( F (0 \leq F < P) \) are defined by the relation \( a = (i+1)P - F \); thus in the absence of breaks \( i+1 \) patrols are required to unwind the bobbin, and the efficiency \( E(i) \) is given by \( E(i) = a/(iP + P) \). As a result of breaks extra patrols may be required to unwind the bobbin; the probability that \( s \) extra patrols are required to unwind a bobbin of length \( a \) is denoted by \( P_s(i) \). If \( N \) such bobbins are unsound in succession on a spindle of the machine the total time spent in actual unwinding will be \( Na \) and the time for which the machine has been running will be, on the average, \( NP[i+1+\mu'(i)] \), where \( \mu'(i) \) is the mean value of \( s \); thus the average efficiency \( \bar{E}(i) \) will be given by \( \bar{E}(i) = a/P[i+1+\mu'(i)] \). The evaluation of \( \mu'(i) \) is thus one of the main purposes of the investigation.

1.4 Results Established. In \( \S \) 2 a fundamental recurrence relation satisfied by \( P_s(i) \) is derived: this leads to the determination of a generating function for \( P_s(i) \) and hence to the mean and variance. In \( \S \) 3 explicit expressions for the \( P_s(i) \) are obtained and are found to be expressible in terms of Laguerre polynomials. A recurrence relation convenient for numerical calculation is also derived and used in an example. Asymptotic normality of the distribution of \( s \) as \( i \) increases is established in \( \S \) 4. In \( \S \) 5 the properties of the average
THE EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME

2. The Distribution $P_s(1)$.

2.1 The Special Case $P > a$ (i.e., $P$ Poisson). In the absence of breaks a single patrol suffices to unwind the bobbin, and the efficiency is $a/P$. Each break necessitates an extra patrol so that the distribution $P_s(1)$ is Poisson with mean $\mu a$ and $E(1) = a/ (\mu a+1)P$.

2.2 Recurrence Relation. $P_s(1)$ is really a function of $F$ and we write it as $P_s(1, F)$ where the bobbin length is given by $a = (i+1)P - F = iP + F$, $0 < F < P$. If $s > 0$ there is at least one break, i.e., there is a first break. If $i > 0$ the first break may occur within the first patrol, or later. The effect of the first break within the first patrol depends on the position of the break within the patrol. If it occurs at $x$, where $0 < x < F$, the amount of yarn still to be unwound after the first patrol is complete is $a - x = iP + F - x = (i+1)P - (F + x)$, which requires $(i+1)$ patrols in the absence of breaks; but if $F < x < P$, the amount still to be unwound after the first patrol is complete is $iP - (x - F)$, which requires $i$ patrols in the absence of breaks. It follows that

\[
P_s(i, F) = e^{\mu F}P_s(i-1, F) + \int_0^F e^{-\mu x}P_s(i, F + x)dx + \mu e^{-\mu x}P_s(i-1, x - F)dx,
\]

\[
= e^{-\mu F}P_s(i-1, F) + \mu e^{\mu F} \int_0^F e^{-\mu x}P_s(i, x)dx + \mu e^{-\mu x} \int_0^{F-x} e^{-\mu x}P_s(i-1, x)dx.
\]

\[i, s = 1, 2, \ldots, \]
If \( s = 0 \) and \( i > 0 \), there may or may not be a first break, and the argument just given shows that if a first break occurs at \( x \) in the first patrol, where \( 0 < x < F \), the number of extra patrols caused by breaks must be at least one. Thus

\[
P_0(i,F) = e^{-\mu F} P_0(i-1,F) + \mu e^{-\mu F} \int_0^F e^{-\mu x} P_0(i-1,x) \, dx , \quad (1b)
\]

\[i = 1, 2, \ldots .\]

If \( i = 0 \) (the case of \( \S 2.1 \)) when \( s > 0 \) the above argument shows that

\[
P_s(0,F) = \mu e^{\mu F} \int_0^F e^{-\mu x} P_{s-1}(0,x) \, dx , \quad (1c)
\]

and when \( s = 0 \) there must be no breaks, so that

\[
P_0(0,F) = e^{-\mu F} . \quad (1d)
\]

The relations (1a) to (1d) are the fundamental recurrence relations satisfied by \( P_s(i,F) \).

2.3 Solution of the Integro-difference Equations. Let

\[
H_t(i,F) = \sum_{s=0}^{\infty} t^s P_s(i,F) , \quad |t| < 1 , \quad (2)
\]

so that from (1) we have
THE EFFICIENCY OF AUTOMATIC MACHINES WITH CONSTANT PATROLLING TIME.

\[ H_t(i, F) = e^{-\mu P} H_t(i-1, F) + t e^{\mu F} \int_F^P \mu e^{-\mu x} H_t(i, x) \, dx \]
\[ + e^{-\mu F} \int_0^F \mu e^{-\mu x} H_t(i-1, x) \, dx, \quad (3a) \]
\[ i = 1, 2, \ldots, \]

\[ H_t(0, F) = e^{-\mu F} + t e^{\mu F} \int_F^P \mu e^{-\mu x} H_t(0, x) \, dx \quad . \quad (3b) \]

Now introduce the bivariate g.f.

\[ W_{t,k}(F) = \sum_{i=0}^{\infty} k^i H_t(i, F), \quad |k| < 1, \quad (4) \]

which from (3) satisfies the integral equation

\[ (1 - k e^{-\mu P}) W(F) = e^{-\mu F} + t e^{\mu F} \int_F^P \mu e^{-\mu x} W(x) \, dx + k e^{-\mu F} \int_0^F \mu e^{-\mu x} W(x) \, dx \]

where \( W(F) \) is written for \( W_{t,k}(F) \).

Differentiating (5) with respect to \( F \) for constant \( P \) gives

\[ (1 - k e^{-\mu P}) \frac{dW}{dF} = \mu (1 - t) W, \quad (6) \]

so that

\[ W(F) = g \exp \frac{\mu (1 - t) F}{1 - k e^{-\mu P}} \]

where \( g \) is independent of \( F \). Substitution in (5) reveals \( g \), leading to
EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

\[ W_{t,k}(F) = \frac{t - ke^{-\mu P}}{1 - ke^{-\mu P}} \frac{\exp[\mu(1-t)F/G]}{-k + t \cdot \exp[\mu(1-t)P/G]} \]

where \( G = (1 - ke^{-\mu P}). \)

2.4 Mean and Variance of the Number of Extra Patrols. The mean of the patrol distribution is found from (7) by evaluating the coefficient of \( k^1 \) in

\[ \frac{dW}{dt} \bigg|_{t=1} \]

and we have

\[ \sum_{i=0}^{\infty} k^i \mu'_i(1) = \frac{b-k(l-K)}{(1-kK)(1-k)^2} - \frac{c}{(1-kK)(1-k)} \]

where for brevity \( \mu^P = c, \mu^P = b, \quad e^{-\mu^P} = K. \) Hence

\[ \mu'_1(1) = -1 - l + \frac{1-c+b(i+2)}{1-K} - \frac{b}{1-K}^2 + k^1 \left[ \frac{b-c+l - (2b-c+l)}{1-K} + \frac{b}{(1-K)^2} \right] \]

For example

\[ \mu'_1(0) = b-c \]
\[ \mu'_1(1) = 2b-c-l + K \quad (b-c+1) \]
\[ \mu'_1(2) = 3b-c-2 + K(2b-c+l) + K^2(b-c+l) \]

and further values can be found immediately from (9), or by using the recurrence relation (deducible from (8))
EFFICIENCY OF AUTOMATIC MACHINES WITH CONSTANT PATROLLING TIME.

\[ \mu'_i(i) = (2+K) \mu'_i(i-1) + (1+2K) \mu'_i(i-2) - K \mu'_i(i-3) \]

\[ = b-c, \quad i=0; \]

\[ = c-1+K, \quad i=1; \]

\[ = 0, \quad i \geq 2, \]

with \( \mu'_i(i) = 0 \) for \( i < 0 \).

It is not quite obvious from (9) that \( \mu'_i(i) \) is positive, as it must be by its nature. The situation becomes clearer if we write (in the notation of advancing differences)

\[ \Delta_i \mu'_i(i) = \left( \frac{b}{1-K} - 1 \right)(1 - K^{i+1}) + (b-c)K^{i+1} \]

\[ \geq 0, \quad i=0,1,..., \]

since \( b-1+e^{-b} > 0 \) for \( b > 0 \). But \( \mu'_i(0) > 0 \) for \( b > c \) and so \( \mu'_i(i) \) is a monotonic increasing function of \( i \) for \( b > c > 0, \quad b \neq 0 \).

It is readily seen that since \( K < 1 \), then for large \( i \)

\[ \mu'_i(i) \sim 1 \left( \frac{b}{1-K} - 1 \right) \]

2.4 The variance of the number of extra patrols can be found from the mean and second factorial moment, and it appears that

\[ \mu'_2(i) = (i+1)\alpha_0 + \sum_{r=1}^{4} q^r \alpha_r + (i+1)K^1 \sum_{r=0}^{3} q^r b_r + \]

\[ + K^1 \sum_{r=0}^{4} q^r c_r - K^2 \left( \sum_{r=0}^{2} q^r d_r \right)^2 \]

where \( q=1/(1-K) \); \( \alpha_0 = -qb + 2b(b+1)q^2 - 2b^2q^3, \quad \alpha_1 = c-b-1; \)

\( a_2 = 3b^2 - 2bc + 7b - 2c + 1, \quad a_3 = -5b^2 + 3bc - 6b, \quad a_4 = 5b^2; \)

\( b_0 = (b-c)^2 + 4(b-c) + 2, \quad b_1 = -6b^2 + 6bc - c^2 - 10b + 4c - 2; \)

\( b_2 = 9b^2 + 6b - 4bc, \quad b_3 = -4b^2; \)
EFFICIENCY OF AUTOMATIC MINDING MACHINES WITH CONSTANT PATROLLING TIME.

\[ c_0 = a_1, \quad c_1 = 2b^2 - c^2 + 5b - c + 1, \quad c_2 = 3b^2 + c^2 - 9b, \]
\[ c_3 = 10b^2 + 4b, \quad c_4 = b_3; \quad d_0 = -a_1, \quad d_1 = 2b^3 - c - 1, \quad d_2 = b. \]

This formula is somewhat complicated but for large \( i \) there is the asymptotic formula

\[ \mu_2(i) \sim (1 + 1)b \left( 1 - 2bK - K^2 \right) / (1 - K)^3 \quad (12a) \]

Moreover, in §3.3 we show that there is a recurrence relation for the second crude moment \( \mu_2'(i) \) for the number of extra patrols, and this can be used for computing purposes.

3. Expression for \( P_g(i,F) \).

We now derive a recurrence relation for \( P_g \) and two closed formulae in terms of Laguerre polynomials.

3.1 We write (7) in the form

\[ N(t,k) = (1-kK)^{W_{t,k}(F)} / (t-kK), \quad (13) \]

so that

\[ \frac{d \log N}{dt} = \frac{-c}{l - kK} - \frac{M}{tM - k} \left[ 1 - \frac{bt}{l - kK} \right], \quad (14a) \]
\[ \frac{d \log N}{dk} = \frac{c(l-t)K^2}{(l-kK)^2} + \frac{1}{tM - k} \left[ 1 - \frac{t(l-t)bK}{(l-kK)^2} \right], \quad (14b) \]

where \( M = \exp \left[ b(l-t)/(1-kK) \right] \).

Eliminating \( M \) and using

\[ \frac{d \log N}{dt} = \frac{d \log N}{dt} - \frac{1}{t-kK}, \]
\[ \frac{d \log N}{dk} = \frac{d \log N}{dk} - \frac{K}{1-kK} + \frac{K}{t-kK}, \]
we find a partial differential equation for \( w \), namely

\[
\left(1-k^2\right) - bk^2(l-t) t \frac{dw}{dt} + (1-k)(1-k-b) k \frac{dw}{dk} = \left[(1-k)(bt+k) + c(k+t)\right] w \tag{15}
\]

Extracting the coefficient of \( k^i \), we have the bivariate recurrence relation

\[
(s+1)P_s(1) - (b+2s+c-1)KP_s(1-1) + (s+1)(s+2)K^2P_s(1-2) - (b+c+1)P_s-1(1) + b(s+2+c)KP_s-1(1-1) = 0 \tag{16}
\]

\( i=0,1, \ldots \), \( s=1,2, \ldots \); \( P_s(1) = 0 \) for \( i < 0 \), with \( \Lambda \) from (7)

\[
\sum_{i=0}^{\infty} k^i P_i(1) = K \exp\left\{c/(1-kK)\right\} / (1-kK) \tag{16^*}
\]

But from the g.f. for Laguerre polynomials (see for example Bateman, p.189, (1953))

\[
\sum_{n=0}^{\infty} z^n L_n(x) = (1-z)^{-m-1} \exp\frac{zx}{z-1} \tag{16a}
\]

it follows that

\[
P_0(1) = e^{c-b} L_1(-c), \quad i=0,1, \ldots \tag{16a}
\]

As examples,

\[
\begin{cases}
  sP_s(0) = (b-c)P_{s-1}(0), \quad s=1,2, \ldots \tag{17a} \\
  P_0(0) = \exp(c-b)
\end{cases}
\]

\[
\begin{cases}
  (s+1)P_s(1) - (2b-c)P_{s-1}(1) = (bs+2s+c+1)KP_s(0) - bsKP_{s-1}(0), \tag{17b} \\
  P_0(1) = (1+c)\exp(c^2) 
\end{cases}
\]
Thus it is obvious from \((17a)\) that \(P_s(0)\) is a Poisson distribution with mean \(b-c\), and using \((16) - (16a)\) values of \(P_s(i)\) can be constructed recursively. Some examples of the distribution are given in Table 1 and Figure 1.

Table 1.

Probability Distribution of the Number of Extra Patrols.

<table>
<thead>
<tr>
<th></th>
<th>(b=1)</th>
<th>(c=0)</th>
<th>(i=0)</th>
<th>(i=2)</th>
<th>(i=4)</th>
<th>(b=1)</th>
<th>(c=0.5)</th>
<th>(i=0)</th>
<th>(i=2)</th>
<th>(i=4)</th>
<th>(b=1)</th>
<th>(c=1)</th>
<th>(i=0)</th>
<th>(i=2)</th>
<th>(i=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.3679</td>
<td>.0498</td>
<td>.0067</td>
<td></td>
<td></td>
<td>0</td>
<td>.6065</td>
<td>.1744</td>
<td>.0426</td>
<td></td>
<td>1.0000</td>
<td>.4737</td>
<td>.1595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.3679</td>
<td>.2833</td>
<td>.0801</td>
<td></td>
<td></td>
<td>1</td>
<td>.3033</td>
<td>.3472</td>
<td>.1760</td>
<td></td>
<td>.2932</td>
<td>.2596</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.1839</td>
<td>.5049</td>
<td>.1960</td>
<td></td>
<td></td>
<td>2</td>
<td>.0758</td>
<td>.2681</td>
<td>.2588</td>
<td></td>
<td>.1523</td>
<td>.2551</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.0613</td>
<td>.2845</td>
<td>.2466</td>
<td></td>
<td></td>
<td>3</td>
<td>.0126</td>
<td>.1343</td>
<td>.2314</td>
<td></td>
<td>.0575</td>
<td>.1726</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.0153</td>
<td>.1167</td>
<td>.0909</td>
<td></td>
<td></td>
<td>4</td>
<td>.0016</td>
<td>.0524</td>
<td>.1524</td>
<td></td>
<td>.0175</td>
<td>.0909</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.0031</td>
<td>.0431</td>
<td>.1356</td>
<td></td>
<td></td>
<td>5</td>
<td>.0002</td>
<td>.0172</td>
<td>.0808</td>
<td></td>
<td>.0045</td>
<td>.0398</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.0005</td>
<td>.0167</td>
<td>.0723</td>
<td></td>
<td></td>
<td>6</td>
<td>.0049</td>
<td>.0363</td>
<td>.0143</td>
<td></td>
<td>.0010</td>
<td>.0152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.0001</td>
<td>.0051</td>
<td>.0331</td>
<td></td>
<td></td>
<td>7</td>
<td>.0013</td>
<td>.0050</td>
<td>.0016</td>
<td></td>
<td>.0002</td>
<td>.0051</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.0014</td>
<td>.0134</td>
<td>.0154</td>
<td></td>
<td></td>
<td>8</td>
<td>.0003</td>
<td>.0049</td>
<td>.0016</td>
<td></td>
<td>.0016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.0004</td>
<td>.0049</td>
<td>.0016</td>
<td></td>
<td></td>
<td>9</td>
<td>.0001</td>
<td>.0050</td>
<td>.0001</td>
<td></td>
<td>.0004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.0001</td>
<td>.0016</td>
<td>.0005</td>
<td></td>
<td></td>
<td>10</td>
<td>.0005</td>
<td>.0016</td>
<td>.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>.0001</td>
<td>.0005</td>
<td>.001</td>
<td></td>
<td></td>
<td>11</td>
<td>.0001</td>
<td>.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 2 we give the values of \(\mu_1(i)\), \(\mu_2(i)\) and the standardised first two cumulants for a number of cases of \(P_s(i)\).

3.2 Comment on Tables 1 and 2. The distributions are initially \((i=0)\) Poissonian, and as \(i\) increases the skewness and kurtosis decrease steadily. Only unimodal distributions have been encountered in the cases considered. The value of \(c\) has a pronounced effect on the distribution for small values of \(i\), but this is 'damped' out as \(i\) increases.
EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

It is also to be noticed that the ratio of the mean to the variance is approximately constant, and in fact is not very different from the asymptotic value

$$\mu_1'(1)/\mu_2(1) \sim (b+1)(1-K)/b(1-2bK-K^2).$$  \hspace{1cm} (18)$$

Lastly there is an indication in Table 2 that $\gamma_1$ and $\gamma_2$ tend to zero as $t \to \infty$, suggesting that the distribution is asymptotically normal.

Table 2.

Skewness and Kurtosis of the Distribution of Extra Patrols.

<table>
<thead>
<tr>
<th></th>
<th>$\mu_1'(1)$</th>
<th>$\mu_2(1)$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$i=0$</td>
<td>$i=1$</td>
<td>$i=2$</td>
<td>$i=3$</td>
</tr>
<tr>
<td>a</td>
<td>1.00</td>
<td>1.74</td>
<td>2.37</td>
<td>2.98</td>
</tr>
<tr>
<td>b</td>
<td>.50</td>
<td>1.05</td>
<td>1.62</td>
<td>2.20</td>
</tr>
<tr>
<td>c</td>
<td>.37</td>
<td>.87</td>
<td>1.42</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>$i=0$</td>
<td>$i=1$</td>
<td>$i=2$</td>
<td>$i=3$</td>
</tr>
<tr>
<td>a</td>
<td>1.00</td>
<td>1.46</td>
<td>1.89</td>
<td>2.34</td>
</tr>
<tr>
<td>b</td>
<td>.50</td>
<td>1.01</td>
<td>1.52</td>
<td>2.04</td>
</tr>
<tr>
<td>c</td>
<td>.37</td>
<td>.87</td>
<td>1.42</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>$i=0$</td>
<td>$i=1$</td>
<td>$i=2$</td>
<td>$i=3$</td>
</tr>
<tr>
<td>a</td>
<td>1.00</td>
<td>.75</td>
<td>.71</td>
<td>.68</td>
</tr>
<tr>
<td>b</td>
<td>1.41</td>
<td>1.02</td>
<td>.86</td>
<td>.76</td>
</tr>
<tr>
<td>c</td>
<td>2.20</td>
<td>1.38</td>
<td>.96</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td>$i=0$</td>
<td>$i=1$</td>
<td>$i=2$</td>
<td>$i=3$</td>
</tr>
<tr>
<td>a</td>
<td>1.00</td>
<td>.76</td>
<td>.73</td>
<td>.65</td>
</tr>
<tr>
<td>b</td>
<td>2.00</td>
<td>1.29</td>
<td>.97</td>
<td>.77</td>
</tr>
<tr>
<td>c</td>
<td>5.42</td>
<td>1.80</td>
<td>.97</td>
<td>.68</td>
</tr>
</tbody>
</table>

a refers to $b=1$, $c=0.0$  \hspace{1cm} $\gamma_1 = \mu_3/\mu_2^{3/2}$  

b refers to $b=1$, $c=0.5$  \hspace{1cm} $\gamma_2 = \mu_4/\mu_2^2 - 3$  

c refers to $b=1$, $c=1.0$
Efficiency of Automatic Winding Machines With Constant Patroling Time.

3.3 Recurrence Relation for Higher Moments. The relation (16) may be used for evaluating the crude or factorial moments of the distribution. Thus by summing over \( s \) we find for the mean number of patrols \( p(1) = \mu'_1(1) + 1 + 1 \),

\[
p(1) - 2Kp(1-1) + K^2 p(1-2) = \left\{ (b-1) + 1 \right\} (i-K) + b \quad (19)
\]

with \( p(0) \) and \( p(1) \) given in \((9a)\). Whereas \((9b)\) is an homogeneous recurrence relation of order three, \((19)\) is non-homogeneous and of order two.

For the second crude moment there is the recurrence

\[
\mu'_2(i) - 2K\mu'_2(i-1) + K^2 \mu'_2(i-2) = \left\{ (b-1) + b-c \right\} \mu'_1(1) \\
- \left\{ (b-2) + b-c+1 \right\} \mu'_1(1) - (i-1)K^2 \mu'_1(1-2) \\
+ b(i+1-1K) - c \quad (20)
\]

and

\[
\mu'_2(1) - 2K\mu'_2(0) = (2b-c-1)\mu'_1(1) - (2b-c-1)K\mu'_1(0) + (2-K)b-c \\
\mu'_2(0) = (b-c)(b-c+1)
\]

In a similar way recurrence relations may be found for higher moments.

3.4 Closed Expressions for \( P_3(1) \). There are two formulae of a closed type for the probability of a extra patrols. We find
ICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

\[ P_s(i) = \left\{ X_s(i) - X_{s-1}(i+1) \right\} e^{-(i+1)b} \]
\[ s=1, 2, \ldots; \ i=0, 1, \ldots, \]

\[ P_o(i) = X_o(i) e^{-(i+1)b}, \]  \hspace{1cm} (21b)

where \( X_s(i) = \sum_{r=0}^{s} (-1)^r \frac{r!}{r!} \left[ c+(s-r)b \right]^r L_{1+s-r}^{(r)}(-c-(s-r)b), \)  \hspace{1cm} (22)

\( L_n^{(m)}(x) \) being a Laguerre polynomial of degree \( n \) in the argument \( x \), defined by

\[ e^{-x} x^m L_n^{(m)}(x) = \frac{1}{n!} (\frac{d}{dx})^n e^{-x} x^{m+n} \]  \hspace{1cm} (23)

The formula (21) is an immediate consequence of expanding the g.f. (7) in ascending powers of \( t \), and in writing down the coefficient of \( k^1 \) making use of (16*). From (21) we are able to deduce the relation

\[ \sum_{n=0}^{s} K^{-n} P_s(n+i) = X_s(i) e^{-(i+1)b}. \]  \hspace{1cm} (24)

A formula similar to (21), but involving Laguerre polynomials with negative upper index, has also been found. It concerns the cumulative distribution of extra patrols, in the form

\[ \sum_{s=0}^{r} P_s(i) = e^{-(i+1)b} \sum_{s=0}^{r} (-1)^s \binom{i+s}{s} L_{i+r}^{(-c)}(-c-(r-s)b). \]  \hspace{1cm} (25)

Since the member on the one side of (25) clearly approaches unity as \( r \to \infty \), it follows that the summatory part involving Laguerre polynomials

\[ X_{\infty}(i) = \sum_{s=0}^{\infty} \binom{s+i}{s} L_{s+i}^{(-c)}(-c-(s-r)b). \]
EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

approaches \( \exp\left\{ (i+1)b - c \right\} \) as \( r \to \infty \); as far as we are aware this property has not been noticed before.

Again it is to be observed that if we take \( i=0 \) in (21) and (25), then we are confronted with a somewhat novel form for the individual term and cumulative sum of a Poisson distribution. We have proved directly the equivalence of the expressions in the case of (25).


4.1 We now show that the distribution of extra patrols is asymptotically normal with mean \( 1(bq-1) \) and variance \( ib\left( q+2Kq^2-2Kbq^3 \right) \) where \( q=1/(1-K) \). The method of proof follows closely that given by Skellam and Shenton (1957) in their § 23 and § 24. Briefly, we have from (7)

\[
\Pi_t(i,F) = \frac{1}{2\sqrt{\pi j}} \int_{-\infty}^{\infty} k^{-(i+1)} \left( \frac{w}{t} \right)^{k/2} (F) dk , \quad (j=\sqrt{-1}), \quad (26)
\]

where the integral is taken round a closed simple contour including only the pole \( k=0 \) of the integrand.

Making the transformation \( kK=(w+t)/(w+1) \) we have

\[
\Pi_t(i,F) = \frac{(1-t)e^{-(i+1)b}}{2\sqrt{\pi j}} \int_{C} \frac{w(w+1)^{1/2} e^{c(w+1)}}{(w+t)^{1+1/2} w^{t-t(w+1)e^{bw}}} dw , \quad (27)
\]

where the contour \( C \) encloses \( w=-t \) and no other pole of the
EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

The integrand. The contour C is now chosen so that (i) the point \( w = -t \) is an internal point, (ii) the integrand tends to zero on it as \( i \to \infty \) (iii) it contains one and only one zero of \( w + t = (w+1)e^{bt} \). With this in view let

\[
f(w) = \frac{w(t^{-1}-1)}{e^{bw}-1} - w - 1,
\]

and consider the behaviour of \( f(w) \) in and on the rectangle

Clearly \( f(w) \) is regular in and on the rectangle, where in addition

on the boundary

\[
|w+1| > \left| \frac{w(t^{-1}-1)}{e^{bw}-1} \right|
\]

provided

\[
\frac{1}{h-1} > \frac{(t^{-1}-1)}{(1-e^{-bh})},
\]

which can be arranged for \( t \) nearly unity, that is \( t=1-\epsilon, \epsilon \) small and positive. Hence by Rouche's theorem \* \( f(w) \) has the same number of zeros in the rectangle as \( w+1 \), i.e. precisely one, no matter how large \( H \) and \( f \) are; thus \( f(w) \) has one and only one zero for which \( \Re w < -h \).

Now take \( C \) to be the circle \( |k| = \sqrt{(1+p)} \) where \( p > 0 \), on which \( |(w+1)/(w+1)| = \sqrt{(1+p)} > 1 \). The centre of the circle (in the \( w \)-plane) is \( \left(-1-(1-t)/p, 0\right) \) and its radius is \((1-t)\sqrt{(p+1)/p}\) so that on it \( \Re w < 0 \). Hence the integrand on this circle tends to zero as \( i \to \infty \) and

\[
H_t(i, F) \sim -\text{Residue of Integrand at the pole of } 1/f(w) \text{ for which } \Re w < 0.
\]

\* See for example Valiron (1942), p.392.
4.3 Putting \( w(\alpha) \) for the zero of \( f(w) \) for which \( \frac{2q}{w} < 0 \) and \( k_\alpha \) for the corresponding value of \( k \), we have

\[
k_\alpha = t \exp \left\{ \frac{2q}{\alpha} \frac{b(1-t)}{(1-Kk_\alpha)} \right\},
\]

where \( k = \sum_{n=0}^{\infty} A_n \alpha^n \), \( A_0 = 1 \), \( t = \exp(-\alpha) \).

By substitution it appears that

\[
A_1 = bc-1,
\]
\[
A_2 = \frac{1}{2} - 3bc/2 + \frac{1}{2}(b^2 - 2Kb)q^2 + Kb^2q^3.
\]

Omitting details, the following expansions will be found for the various factors of the integrand appearing in (28):

\[
k_\alpha^{-1}(1+1) \sim \exp \left[ -(1+1)A_1 \alpha + \frac{1}{2}(1+1)(A_1^2 - 2A_2) \alpha^2 + \ldots \right], \quad (30a)
\]
\[
(\exp^{-1} - KK_\alpha)/(1 - KK_\alpha) \sim 1 - \alpha q + \frac{1}{2} \alpha^2 \left[ 1 - (1 + 2A_1)K \right] q^2 + \ldots \quad (30b)
\]
\[
\left[ 1 - (1 - c)K_\alpha K/(1 - KK_\alpha) \right] \sim 1 + bKq^2 \alpha + bKq^3 \alpha^2 \left[ -\frac{1}{2} + A_1 + 2KqA_1 + bKq^2 \right] + \ldots
\]

Collecting the terms together and simplifying we have finally

\[
H_\alpha^t (f,F) = (1 + C_1 \alpha + C_2 \alpha^2 + \ldots) \exp (b_1 \alpha + \frac{1}{2} b_2 \alpha^2 + \ldots),
\]

where

\[
C_1 = (K - 1 + bK)q^2,
\]
\[
C_2 = (1 + K)q^2 - (7 + K)bKq^3 + 2(1 + 2K)b^2Kq^4,
\]
\[
b_1 = i + l - b(1 + l - c/b)q,
\]
\[
b_2 = (1 + l - c/b) \left( bq + 2bKq^2 - 2b^2Kq^3 \right).
\]
the $c$'s being independent of $i$ whereas each $b$ is a linear function of $i$. This shows that the distribution of the number of extra patrols is asymptotically normal. Moreover, if we evaluate the mean and variance from (31), exact agreement (neglecting powers of $K$ such as $K^4$ and higher) with the results stated in (9) and (12) will be found.

5. Efficiency of the Unwinding Process: Relation between $\bar{E}(i)$ and $P$.

The average efficiency, as mentioned in §1.2, is given by

$$\bar{E}(i) = \frac{\mu a}{b} (i + 1 + \mu_1(i)) \tag{32}$$

In practice, bobbins with fixed unwinding time $a$ have to be unwound and $P$ has to be varied, within the limits available on the machine, to give maximum efficiency. If $c$ is eliminated from (32) by means of the relation $(i+1)b = c = \mu a$, we get

$$\frac{1}{\bar{E}(1)} = \frac{b}{\mu a} \left[ (1 + \mu a + b)q - bq^2 + K^1(1 + \mu a - 1 b - (1 + \mu a - 1 b)q + bq^2) \right] \tag{33a}$$

$$= b \sum_{n=0}^{\infty} K^n (1 + \mu a - nb) \tag{33b}$$

The expressions in (33) give the efficiency as a function of $b$ for given $\mu a$; they are valid for the range $\mu a / 1 > b > \mu a / (1+1)$.

The extreme values, corresponding to $b = \mu a / 1$ and $b = \mu a / (1+1)$ will be denoted by $\bar{E}_1(i)$ and $\bar{E}_2(i)$ respectively. We find, with $\mu a = z$,
EFFICIENCY OF AUTOMATIC WINDING MACHINES WITH CONSTANT PATROLLING TIME.

\[ \frac{1}{E_1(i)} = \frac{1 + z - \exp(-z/2) - (1 - \exp(-z))(z/1)/(\exp(z/1) - 1)}{1(1 - \exp(-z/1))} \]  \hspace{1cm} (34a)

\[ \frac{1}{E_2(i)} = \frac{1 + z - \exp(-z) - (1 - \exp(-z))(z/(1+1))/(\exp(z/(1+1) - 1))}{(1+1)(1 - \exp(-z/(1+1))} \]  \hspace{1cm} (34b)

Consideration of the meaning of \( E(i) \) suggests that it increases steadily as \( b \) decreases from \( \mu a/i \) to \( \mu a/(i+1) \), a suggestion borne out by all the values calculated. In addition, \( E(i) \) has a saltus at \( b = \mu a/(i+1) \), for it follows easily from (34) that

\[ \frac{1}{E_1(i+1)} - \frac{1}{E_2(i)} = \frac{\exp(-\mu a)}{i+1} > 0 \]

*Mr. F. Bowman has succeeded in proving this property when \( 0 < \mu a < 4/3 \).

As an example of the order of efficiencies obtainable, Table 3 gives values of \( E_1(i) \) and \( E_2(i) \), and certain intermediate values of \( E(i) \), for \( i=0,1,\ldots,10 \) and \( \mu a=1 \). Higher values of these quantities are not likely to occur in practice. The graph of \( E(i) \) against \( b \) is shown in Figure 2, and extracts from it in Figures 3 and 4. The value of \( b \) (i.e., in effect \( P \)) for maximum efficiency, for given \( \mu a \), may be determined from this graph. A more detailed table or a set of
such graphs, for a range of values of $\mu a$, will be useful in cases where bobbins with different unwinding times are being simultaneously unwound on different spindles of the same machine. This fuller information and further aspects of the problem are under consideration and these are reserved for another occasion.

Table 3.

Values of $\bar{E}(1)$ as a function of $b$ ($\mu a=1$)

<table>
<thead>
<tr>
<th>$b$</th>
<th>$E(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\infty$</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td>1/2</td>
<td>43</td>
</tr>
<tr>
<td>1/3</td>
<td>56</td>
</tr>
<tr>
<td>1/5</td>
<td>69</td>
</tr>
<tr>
<td>1/7</td>
<td>61</td>
</tr>
<tr>
<td>1/9</td>
<td>69</td>
</tr>
<tr>
<td>1/11</td>
<td>77</td>
</tr>
<tr>
<td>1/13</td>
<td>71</td>
</tr>
<tr>
<td>1/15</td>
<td>76</td>
</tr>
<tr>
<td>1/17</td>
<td>82</td>
</tr>
<tr>
<td>1/19</td>
<td>76</td>
</tr>
<tr>
<td>1/21</td>
<td>81</td>
</tr>
<tr>
<td>1/23</td>
<td>85</td>
</tr>
</tbody>
</table>

$E(1)$ is expressed as a percentage.

Acknowledgements. The problem was brought to the attention of one of us by Mr. D. Brunnenschweiler of the Textile Industries Department in the Manchester College of Science and Technology. We are grateful to him for many helpful discussions and advice on practical issues.

We are also indebted to Mr. A. Fletcher for constructing the diagrams.
Glossary of Terms.

\[ a = (i+1)P - F = iP + \bar{F} \]  
Length of yarn on bobbin (1.3).

\[ b = \mu P \]  
(1.3).

\[ c = \mu F \]  
(1.3).

\[ F(i) = \text{Efficiency}; \quad E(i) = \text{Average Efficiency} \]  
(1.3).

\[ \bar{F} = \text{Fractional part of length of yarn}; \quad F + F = P \]  
(1.3).

\[ H_t(i,F) = \text{generating function for the number of extra patrols} \]  
(2.3).

\[ i = \text{'Integral' part of length of yarn} \]  
(1.3).

\[ K = \exp(-b) \]  
(1.3).

\[ L_n(x) \text{ is a Laguerre Polynomial of degree } n \text{ in the argument } x \]  
(3.1).

\[ P_s(i,F) = \text{probability of } s \text{ extra patrols to unwind a bobbin of} \]  

\[ q = 1/(1-K) \]  

\[ W_{t,k}(i,F) \text{ is a bivariate generating function} \]  
(2.3).

\[ \mu \] is a positive constant relating to the probability of a break.

\[ \mu'(i) = \text{the mean number of extra patrols for a bobbin of length } a. \]

\[ \mu_2(i) = \text{the variance of the number of extra patrols.} \]

\[ \binom{r}{s} = \frac{r!}{s!(r-s)!} \]

REFERENCES.


Legends to Line Drawings.

Figure 1  Probability Distribution for the Number of Extra Patrols.

Figure 2  Variation of Efficiency with Patrolling Time.

Figure 3  Variation of Efficiency with Patrolling Time:
           Extract from Fig. 2 on larger scale.

Figure 4  Variation of Efficiency with Patrolling Time;
           Further extract from Fig. 2 on larger scale.
Probability Distribution of Number of Extra Patrols

\[ b = 1, \quad c = 0 \]

\[ b = 1, \quad c = \frac{1}{2} \]

\[ b = 1, \quad c = 1 \]
Variation of Efficiency with Patrolling Time

Fig. 2