SOME PROPERTIES OF THE DISTRIBUTION FUNCTIONS OF THE INDEX NUMBERS OF SIMPLE CONTINUED FRACTIONS

A THESIS SUBMITTED

by

[illegible]

to

THE DEPARTMENT OF MATHEMATICS

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in the subject of

MATHEMATICAL STATISTICS

University of Ottawa

Ottawa, Ontario

October, 1963

ABSTRACT

A brief review of the probability theory of simple continued fractions is given, as developed by Borel, Kuzmin and others. Some hitherto unpublished Properties of the distribution functions $M_n(x)$ as defined in Kuzmin's work have been derived.

Various aspects of the dependence of index numbers are discussed. Finally a statistical analysis is made to study empirically the distribution functions of a_n which are directly related to $M_{n-1}(x)$.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

I am deeply indebted to Dr. C.J. van der Maas who, in his capacity as thesis adviser, provided the guidance and inspiration so necessary to the completion of this work.

Chapter I

INTRODUCTION

The application of Probability Theory to continued fractions could be said to have started when Gauss in 1812 stated the following problem:

If a_n is the n^{th} incomplete quotient (or the n^{th} index number) of a real number in its simple continued fraction expansion, then

$$\lim_{n \to \infty} P\left\{a_n \geq k\right\} = \frac{1}{\log 2} \log\left(1 + \frac{1}{k}\right).$$

where $P\left\{a_n \geq k\right\}$ denotes the Probability that a_n is greater than or equal to k. However a proof for the above theorem by Gauss is not known.

Even before the Gaussian theorem was proved, Borel [1] 1909, Bernstein [2] 1912 and Khintchine [3], [4] obtained some results in this theory. Borel gave the upper and lower limits for $P\left\{a_{n+1} = k\right\}$ and $P\left\{a_{n+1} > k\right\}$ as follows:

$$\frac{2}{3k(k+1)} < P\left\{a_{n+1} = k\right\} < \frac{3}{(k+1)(k+2)}$$

$$\frac{2}{3(k+1)} < P\left\{a_{n+1} > k\right\} < \frac{3}{k+2}$$

[Note: The interval for the first inequality can be made smaller than Borel gave it. We can namely write

$$\frac{1}{k(k+2)} \leq P\left\{a_{n+1} = k\right\} \leq \frac{2}{(k+1)^2}$$

This leads to a form for the second inequality which is also different from the one given above. By summation we get

$$\frac{1}{2}\left[\frac{2k+3}{(k+1)(k+2)}\right] \leq P\left\{a_{n+1} > k\right\} \leq 2\sum_{r=1}^{\infty} \frac{1}{(k+r+1)^2} = 2\psi'(k+1)$$

where $\psi'(x) = \sum_{r=1}^{\infty} \frac{1}{(r+x)^2}$.

The results of Borel and Bernstein amount to the following theorem:

Let $\phi(n)$ be a positive increasing function of n . Then

A) If $\sum \frac{1}{\phi(n)}$ converges, $a_n(\alpha) \leq \phi(n)$ for almost all α and for all large n . $[n > N(\alpha)]$;

B) If $\sum \frac{1}{\phi(n)}$ diverges $a_n(\alpha) > \phi(n)$ for almost all α and for an infinity of n .

Taking $\phi(n) = n$, one obtains that the $a_n(\alpha)$'s are not bounded for almost all α . The above theorem was improved by Dyson [5] and Khintchine [3] .

In 1928 the first proof of the Gaussian theorem was given by Kuzmin [6], [7] . He stated the theorem in the following form:

Let α be a real number in the closed interval $(0, 1)$, whose simple continued fraction expansion is given by (1.1)

$$\alpha = \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots + \frac{1|}{|a_n} + \cdots \qquad (1.1)$$

Let

$$z_n(a) = \frac{1}{|a_{n+1}} + \frac{1}{|a_{n+2}} + \frac{1}{|a_{n+3}} + \cdots \tag{1.2}$$

and

$$M_n(x) = \text{Prob}\left\{z_n(a) < x\right\} \tag{1.3}$$

Then the theorem is

$$\lim_{n \to \infty} M_n(x) = \frac{\log (1+x)}{\log 2}$$

From (1.2) one can observe that

$$z_n(a) = \frac{1}{a_{n+1}(a) + z_{n+1}(a)}$$

and $a_{n+1}(a) = \left[\frac{1}{z_n(a)}\right]$ (integral part of $\frac{1}{z_n(a)}$)

so one obtains that

$$z_{n+1}(a) = \frac{1}{z_n(a)} - \left[\frac{1}{z_n(a)}\right]$$ from which it follows that

$$M_{n+1}(x) = \sum_{k=1}^{\infty} M_n\left(\frac{1}{k}\right) - M_n\left(\frac{1}{k+x}\right) \tag{1.4}$$

It can be verified easily that the function $M_n(x) = c \log (1+x)$ (c is a constant independent of n) satisfies the equation (1.4). This suggested to Gauss the asymptotic behaviour of $M_n(x)$.

If it is further assumed that the distribution function of a (ie $M_0(x)$) has a continuous bounded derivative it can be proved from (1.4) that $M'_n(x)$ satisfies the following relation,

$$M'_{n+1}(x) = \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} M'_n\left(\frac{1}{k+x}\right) \qquad (1.5)$$

Using the equation (1.5), Kuzmin proved that

$$\lim_{n \to \infty} M'_n(x) = \frac{1}{(1+x)\log 2}$$

Further he gave the following inequality

$$\left| M'_n(x) - \frac{1}{1+x} \cdot \frac{1}{\log 2} \right| < A e^{-\lambda\sqrt{n}} \qquad (1.6)$$

Where A and λ are positive constants. This inequality is very useful in obtaining further important inequalities [7] e.g.

$$\left| \frac{\text{Prob}\,(a_{n_1} = r_1;\ a_{n_2} = r_2;\ \ldots\ a_{n_t} = r_t;\ a_{n_{t+1}} = r)}{\text{Prob}\,(a_{n_1} = r_1;\ a_{n_2} = r_2;\ \ldots\ a_{n_t} = r_t)} - \frac{\log\left\{1 + \frac{1}{r(r+2)}\right\}}{\log 2} \right|$$

$$< \frac{A}{r(r+1)} e^{-\lambda\sqrt{n_{t+1} - n_t}}.$$

The second proof of the Gaussian theorem was given by Levy [8] in 1929. He also proved that $\frac{q_{n-1}}{q_n}$ (q_n is the denominator of the n^{th} convergent of the simple continued fraction) has the same asymptotic distribution function as a_n. Further he [9] studied the behaviour of the geometric mean of the a_n's, the n^{th} root of q_n and the divergence of the series $\Sigma \frac{1}{s_n}$ where $s_n = a_1 + a_2 + \cdots + a_n$. Khintchine [10], [11] proved all the results mentioned above using the following theorem and its generalised equivalent for many variables.

If $f(r) = O(r^{1-\delta})$, ($f(r)$ is a non-negative function of r and $\delta > 0$), then almost everywhere

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(a_k) = \sum_{r=1}^{\infty} f(r) \frac{\log\left\{1+\frac{1}{r(r+2)}\right\}}{\log 2} \qquad (1.7)$$

Since the function $f(r) = r$ does not satisfy the conditions of the theorem, one cannot obtain any result about the arithmetic mean of the a_n's from (1.7). However Khintchine proved the following result for the aritimetic mean;

For each $\epsilon > 0$

$$\text{Prob}\left\{\left|\frac{\frac{a_1+a_2+\dots+a_n}{n}}{\frac{\log n}{\log 2}} - 1\right| > \epsilon\right\} \to 0 \text{ as } n\to\infty$$

or in other words $\dfrac{\frac{a_1+a_2+\dots+a_n}{n}}{\frac{\log n}{\log 2}}$ converges in probability to 1 .

Theorem (1.7) is proved by A. Kroschew [12] where $f(r)$ satisfies less stringent conditions.

The third proof for the Gaussian theorem was given by Denjoy [13], [14]. He generalised some results of Borel. Further he obtained the following results.

1) $M_n(x, t)$ tends uniformly to $\dfrac{t.\log(1+x)}{\log 2}$ as $n\to\infty$

where $M_n(x, t) = \text{Prob}(z_n(u) < x)$ for $0 < u < t$.

$$2) \qquad \lim_{n \to \infty} \text{Prob}\left\{a_n \geq k,\ a_{n+1} \geq l\right\} = \frac{\log\left(1 + \frac{1}{kl}\right)}{\log 2} .$$

[In this connection XXXXXXXXXXXXXX notes exchanged between Lévy and Denjoy [15], [16] are of some interest].

The fourth proof was given by Doeblin [17].

The aim of this work is to study some properties of the distribution functions $M_n(u)$ defined by (1.3) in Kuzmin's work. $Z_n(u)$ defined in (1.2) can also be defined in the following way.

$$Z_0(u) = \frac{1}{u}$$

$$Z_1(u) = \frac{1}{u} - \left[\frac{1}{u}\right]$$

$$Z_2(u) = \frac{1}{Z_1(u)} - \left[\frac{1}{Z_1(u)}\right]$$

$$\vdots$$

$$Z_{n+1}(u) = \frac{1}{Z_n(u)} - \left[\frac{1}{Z_n(u)}\right] .$$

If one writes $T(u) = \frac{1}{u} - \left[\frac{1}{u}\right]$,

$$Z_n(u) = T[Z_{n-1}(u)] = T[T\,Z_{n-2}(u)] \cdots T^{n-1}\,Z_1(u) = T^n(u)$$

or $Z_n(u)$ is obtained by the iteration of the function $T(u)$. Since Z is a composition of the T's the function $T(u)$ is studied and its Laplace transform is obtained in the second chapter.

In the third chapter the properties of $M_0(x)$, $M_1(x)$ and $M_2(x)$ are studied. If one defines the distribution function $F_n(k)$ as $\text{Prob}(a_n \leq k)$, then

$$
\begin{aligned}
F_n(k) &= \text{Prob}\,(a_n \le k) \\
&= \text{Prob}\,\left(\left[\frac{1}{Z_{n-1}}\right] \le k\right) \\
&= \text{Prob}\,\left(\frac{1}{Z_{n-1}} < k+1\right) \\
&= 1 - M_{n-1}\left(\frac{1}{k+1}\right) \ldots \qquad (1.8)
\end{aligned}
$$

$\therefore$ From $M_n(x)$ one can calculate $F_n(k)$ very easily. $F_1(k)$, $F_2(k)$, $F_3(k)$ and $F_\infty(k)$ are calculated and their graphs are drawn.

In the fourth chapter all the properties obtained for $M_2(x)$ are generalised for $M_n(x)$. Functions $R_n(t)$ are defined by the following integral relation

$$R_{n+1}(t) = \int_0^\infty \frac{R_n(u)\, u\, J_1(2\sqrt{ut})}{(e^u - 1)\sqrt{ut}}\, du \quad \text{and} \quad R_0(t) = 1 \quad \text{where}$$

$J_1(x)$ is the Bessel function.

$$\text{Then} \quad M_n(x) = \int_0^\infty \frac{R_n(t)\,(1 - e^{-xt})}{e^t - 1}\, dt .$$

In the fifth chapter an expression for the joint probability $(a_{n+p} = \ell,\ a_n = k)$ is obtained and the statistical dependence (defined in chapter 5) of a_n and a_{n+p} is tabulated for $n=1$ and $p=1$ and 2.

In the last chapter the functions $F_n(k)$ are studied empirically by taking a sample of 1000 random numbers. The empirical distribution functions $F_1^*(k)$, $F_2^*(k)$ ... $F_{15}^*(k)$ are tabulated and their graphs are drawn. χ^2-test on the approximation of the empirical distribution functions $F_n^*(k)$ by $F_\infty(k)$ is carried out.

Chapter II

LAPLACE TRANSFORM OF T(x)

As defined in the previous chapter

$$T(x) = \frac{1}{x} - \left[\frac{1}{x}\right] \quad \text{for } 0 < x \leq 1$$

$$= 0 \qquad \text{otherwise}$$

$T(x)$ is bounded and continuous except at the points $x = \frac{1}{k}$ where k is an integer. In each of the intervals $\left(\frac{1}{k} \leq x < \frac{1}{k+1}\right)$, $T(x)$ is continuous and behaves there like $\frac{1}{x}$ since $\left[\frac{1}{x}\right]$ is constant in each interval.

By definition the Laplace transform of $T(x)$ is $\int_0^1 e^{-sx}\, T(x)dx$. Denote it by $H(s)$

$$H(s) = \int_0^1 e^{-sx}\, T(x)dx \tag{2.1}$$

$$= \int_0^1 e^{-sx} \left\{\frac{1}{x} - \left[\frac{1}{x}\right]\right\} dx ,$$

$$= \lim_{n\to\infty} \int_{\frac{1}{n}}^1 e^{-sx} \left\{\frac{1}{x} - \left[\frac{1}{x}\right]\right\} dx, \text{ where } n \text{ is a}$$

positive integer.

Dividing the interval $(\frac{1}{n}, 1)$ into $(n-1)$ intervals $(\frac{1}{n}, \frac{1}{n-1})$, $(\frac{1}{n-1}, \frac{1}{n-2})$ ... $(\frac{1}{2}, 1)$, one can write

$$\theta(s) = \lim_{n\to\infty} \sum_{k=1}^{n-1} \int_{\frac{1}{k+1}}^{\frac{1}{k}} e^{-sx} \left\{\frac{1}{x} - [\frac{1}{x}]\right\} dx$$

or $\theta(s) = \lim_{n\to\infty} S_n$ where

$$S_n = U_1 + U_2 + \dots + U_{n-1}$$

and $U_k = \int_{\frac{1}{k+1}}^{\frac{1}{k}} e^{-sx} \left\{\frac{1}{x} - [\frac{1}{x}]\right\} dx$

Now $U_k = \int_{\frac{1}{k+1}}^{\frac{1}{k}} e^{-sx} \left\{\frac{1}{x} - k\right\} dx$ since $[\frac{1}{x}] = k$ in the interval

$(\frac{1}{k+1}, \frac{1}{k})$,

$$= \int_{\frac{1}{k+1}}^{\frac{1}{k}} (\frac{1}{x} - k)\left(1 - \frac{sx}{1!} + \frac{s^2x^2}{2!} + \dots + \frac{(-1)^r s^r x^r}{r!} + \dots\right) dx$$

$$= \sum_{r=0}^{\infty} \int_{\frac{1}{k+1}}^{\frac{1}{k}} (\frac{1}{x} - k) \frac{(-1)^r s^r x^r}{r!} dx$$

$$= [\log x - kx]_{\frac{1}{k+1}}^{\frac{1}{k}} + \sum_{r=1}^{\infty} \left[\frac{(-1)^r s^r x^r}{r.r!} - \frac{(-1)^r s^r x^{r+1}}{(r+1)!}\right]_{\frac{1}{k+1}}^{\frac{1}{k}}$$

$$= [\log \frac{k+1}{k} - \frac{1}{k+1}] + \sum_{r=1}^{\infty} \frac{(-1)^r s^r}{(r+1)!r} \left\{\frac{1}{k^r} - \frac{1}{(k+1)^r} - \frac{r}{(k+1)^{r+1}}\right\}$$

$$\therefore S_n = \sum_{k=1}^{n-1} \left[\log(\frac{k+1}{k}) - \frac{1}{k+1} + \sum_{r=1}^{\infty} [\frac{(-1)^r s^r}{(r+1)!r} \left\{\frac{1}{k^r} - \frac{1}{(k+1)^r} - \frac{r}{(k+1)^{r+1}}\right\}]\right]$$

$$= \left(\log n - 1 - \frac{1}{2} - \frac{1}{3} \cdots \frac{1}{n}\right) + \sum_{r=1}^{\infty} \frac{(-1)^r s^r}{(r+1)!\,r} \left\{1 - \frac{1}{n^r} - \frac{r}{2^{r+1}} - \frac{r}{3^{r+1}} \cdots - \frac{r}{n^{r+1}}\right\}$$

$$\lim_{n \to \infty} S_n = 1 - C + \sum_{r=1}^{\infty} \frac{(-1)^r s^r}{r(r+1)!} \left\{1 + r - r\,\zeta(r+1)\right\}$$

where C is Euler's constant and $\zeta(r+1) = \sum_{k=1} \frac{1}{k^{r+1}}$ where ζ is Riemann's zeta function.

Finally

$$T(s) = (1 - C) + \sum_{r=1}^{\infty} \frac{(-1)^r s^r}{r!} \left(\frac{1}{r} - \frac{\zeta(r+1)}{r+1}\right) \qquad (2.2)$$

From (2.2) one obtains that

$$\int_0^1 x^r\, T(x)dx = \left(\frac{1}{r} - \frac{(r+1)}{r+1}\right) = M_r$$

Let $\varphi(s)$ be the Steiltjes transform of $T(x)$.

$$\varphi(s) = \int_0^1 \frac{T(x)}{s+x}\, dx$$

$$= \frac{1}{s} \int_0^1 T(x)\left(1 - \frac{x}{s} + \frac{x^2}{s^2} \cdots\right) dx$$

$$= \frac{1}{s}\left[(1 - C) - \frac{M_1}{s} + \frac{M_2}{s^2} + \cdots + \frac{(-1)^n M_n}{s^n} + \cdots\right]$$

substituting the values for M_r one obtains

$$\varphi(s) = \frac{1 - C}{s} - \frac{1}{s^2}\left(1 - \frac{\zeta(2)}{2}\right) + \frac{1}{s^3}\left(\frac{1}{2} - \frac{\zeta(3)}{3}\right) + \cdots$$

$$= \frac{1 - C}{s} - \frac{1}{s}\log\left(1 + \frac{1}{s}\right) + \frac{\zeta(2)}{2s^2} - \frac{\zeta(3)}{3s^3} + \cdots + \frac{(-1)^r \zeta(r)}{r \cdot s^r} + \cdots$$

Substituting the formula

$$\zeta(n) = \int_0^\infty \frac{x^{n-1}}{(n-1)!}\frac{dx}{e^x-1}$$

$$Q(s) = \frac{1-C}{s} - \frac{1}{s}\log\left(1+\frac{1}{s}\right) + \int_0^\infty \frac{e^{-\frac{x}{s}} - 1 + \frac{1}{s}(1-e^{-x})}{x(e^x-1)}\,dx$$

Substituting

$$-C = \int_0^\infty \frac{1-e^{-x}-x}{x(e^x-1)}\,dx$$

$$Q(s) = \frac{1}{s} - \frac{1}{s}\log\left(1+\frac{1}{s}\right) + \log\Gamma\left(1+\frac{1}{s}\right) \qquad (2.3)$$

where Γ is the gamma function.

Chapter III

PROPERTIES OF $M_o(x)$, $M_1(x)$ AND $M_2(x)$

$M_o(x)$:-

If a has uniform distribution then $\text{Prob}(a \leq x) = x$ {in the interval (0.1)} .

i.e. $M_o(x) = x$ and therefore

$F_1(k) = 1 - \frac{1}{k+1}$ by (1.6)

$$f_1(k) = \text{Prob}(a_1 = k) = F_1(k) - F_1(k-1) = \frac{1}{k(k+1)} \ldots \quad (3.1)$$

Since $F_1(1) = \frac{1}{2}$, 1 is the median of the distribution function of a_1 .

$f_1(1) > f_1(k)$ for all $k > 1$. Since $\frac{1}{k(k+1)}$ decreases as k increases.

∴ 1 is the mode for the distribution function of a_1 .

$$E(k^r) = \sum_{k=1}^{\infty} \frac{k^r}{k(k+1)} = \sum_{k=1}^{\infty} \frac{k^{r-1}}{k+1} = \infty \quad \text{for } r \geq 1 .$$

∴ All the moments of $F_1(k)$ are infinite. However the moments for negative powers of k do all exist. For example

$$E(\tfrac{1}{k}) = \sum_{k=1}^{\infty} \frac{1}{k^2(k+1)} = \frac{\pi^2}{6} - 1 = .6493 .$$

$M_1(x)$:-

$$M_1(x) = \sum_{k=1}^{\infty} M_o(\tfrac{1}{k}) - M_o(\tfrac{1}{k+x})$$

$$= \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+x}\right) = C + \psi(x) \ldots \qquad (3.2)$$

where C is Euler's constant and

$$\psi(x) = \frac{d \log x!}{dx}$$ ($x!$ stands for $\Gamma(1+x)$ where Γ is the gamma function).

And $F_2(k) = 1 - C - \psi\left(\frac{1}{k+1}\right)$ by (1.8)

From this

$$f_2(k) = \text{Prob}\,(a_2 = k) = \psi\left(\frac{1}{k}\right) - \psi\left(\frac{1}{k+1}\right).$$

Since $\psi\left(\frac{1}{k}\right) - \psi\left(\frac{1}{k+1}\right)$ is a decreasing function of k, 1 is the mode for this distribution function also. But it is not the median since $F_2(1) = .386 \ldots$. From (3.2) one obtains that

$$[illegible]'(x) = \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} = \psi'(x)$$ which shows that $\psi_1(x)$

is decreasing.

Again all the moments are infinite since

$$E(k^r) = \sum_{k=1}^{\infty} k^r \left[\psi\left(\frac{1}{k}\right) - \psi\left(\frac{1}{k+1}\right)\right]$$

$$= \sum_{k=1}^{\infty} \frac{k^r}{k(k+1)} \psi'\left(\frac{1}{k+\theta_k}\right) \quad \text{where } 0 < \theta_k < 1$$

$$> \sum_{k=1}^{\infty} \frac{k^{r-1}}{k+1} \psi'(1) \quad \text{since } \psi(x) \text{ is decreasing}$$

$$= \infty \quad \text{for } r \geq 1$$

However it can be verified easily that all the moments for negative powers of k exist.

The bounds of $M_1(x)$

Consider

$$M_1(x) = \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+x}\right) = \sum_{k=1}^{\infty} \frac{x}{k(k+x)}$$

Since $k \leq k+x \leq k+1$

$$\sum_{k=1}^{\infty} \frac{x}{k^2} \geq M_1(x) \geq \sum_{k=1}^{\infty} \frac{x}{k(k+1)}$$

i.e. $$x\, M_1'(0) \geq M_1(x) \geq M_0(x) \quad \ldots \quad (3.3)$$

From (3.3) one obtains that

$$F_1(k) \geq F_2(k) \quad \text{for all } k$$

But $$f_1(1) > f_2(1)$$

and $$f_1(k) < f_2(k) \quad \text{for all } k > 1 .$$

Some properties of $M_1(x)$

Now some important properties of $M_1(x)$ which will be used later are given below. First we have

$$M_1(x) = \int_0^{\infty} \frac{1 - e^{-tx}}{e^t - 1}\, dt \quad [18] \quad \ldots . \quad (3.4)$$

which can be proved easily.

This formula can be expressed as a finite integral as follows

$$M_1(x) = \int_0^1 \frac{1 - v^x}{1 - v}\, dv \ldots \qquad (3.4^1)$$

From (3.4^1), one observes that $M_1(x)$ for rational values of x can be expressed as the sum of a finite number of terms one has namely:

$$M_1\left(\frac{h}{k}\right) = \log\left(\frac{1}{k}\right) - \frac{\pi}{2}\cot\frac{\pi}{k} + \sum_{r=1}^{< \frac{k}{2}} \cos\frac{2\pi h r}{k} \log\left[2 - 2\cos\frac{2\pi r}{k}\right]$$ [19]

<u>Maclaurin expansion of $M_1(x)$</u>

This is given by

$$M_1(x) = x\,\zeta(2) - x^2\,\zeta(3) + \ldots + (-1)^{n-1} x^n\,\zeta(n+1) + \ldots \qquad (3.4^2)$$

where ζ is again representing Riemann's zeta function.

By differentiating the formula (3.4), one obtains

$$M_1'(x) = \int_0^\infty \frac{t\, e^{-tx}}{e^t - 1}\, dt \ldots \qquad (3.5)$$

i.e. $\quad M_1'(x)$ is the Laplace transform of $\dfrac{t}{e^t - 1}$.

<u>$M_2(x)$</u>

By definition

$$S_2(x) = \sum_{k=1}^{\infty} \psi\left(\frac{1}{k}\right) - \psi\left(\frac{1}{k+x}\right)$$

$$= \sum_{k=1}^{\infty} \left[\zeta(2) \left(\frac{1}{k} - \frac{1}{k+x}\right) - \zeta(3) \left(\frac{1}{k^2} - \frac{1}{(k+x)^2}\right) + \cdots \right]$$

$$= \zeta(2) \left(\psi(x) - \psi(0)\right) + \zeta(3) \left(\psi'(x) - \psi'(0)\right) + \cdots$$

$$= \sum_{r=0}^{\infty} \frac{\zeta(r+2)}{r!} \left[\psi^{(r)}(x) - \psi^{(r)}(0) \right]$$

where $\psi^{(r)}(x)$ is the r^{th} derivative of $\psi(x)$. $S_2(x)$ can be expressed by a more rapidly convergent series in the following manner. One has namely

$$S_2(x) = \sum_{k=1}^{\infty} \psi\left(\frac{1}{k}\right) - \psi\left(\frac{1}{k+x}\right)$$

$$= \sum_{k=1}^{\infty} \left\{ \sum_{r=1}^{\infty} \frac{1}{r + \frac{1}{k+x}} - \frac{1}{r + \frac{1}{k}} \right\}$$

$$= \sum_{k=1}^{\infty} \left\{ \sum_{r=1}^{\infty} \frac{x}{[r(k+x)+1][r\,k+1]} \right\}$$

Interchanging the order of summation one obtains

$$S_2(x) = \sum_{r=1}^{\infty} \frac{1}{r^2} \left\{ \sum_{k=1}^{\infty} \frac{x}{(k+x+\frac{1}{r})(k+\frac{1}{r})} \right\} \tag{3.6}$$

$$= \sum_{r=1}^{\infty} \frac{1}{r^2} \left[\sum_{k=1} \left\{ \frac{1}{k+\frac{1}{r}} - \frac{1}{k+x+\frac{1}{r}} \right\} \right]$$

$$M_2(x) = \sum_{r=1}^{\infty} \frac{1}{r^2} \left\{ \psi(x + \tfrac{1}{r}) - \psi(\tfrac{1}{r}) \right\} \qquad (3.7)$$

The bounds of $M_2(x)$

From the formula (3.6) and the inequality

$k + \frac{1}{r} \leq k + x + \frac{1}{r} \leq k + 1 + \frac{1}{r}$ it follows

$$\sum_{r=1}^{\infty} \frac{1}{r^2} \sum_{k=1}^{\infty} \frac{x}{(k+\frac{1}{r})^2} \geq \sum_{r=1}^{\infty} \frac{1}{r^2} \sum_{k=1}^{\infty} \frac{x}{(k+x+\frac{1}{r})(k+\frac{1}{r})} \geq \sum_{r=1}^{\infty} \frac{1}{r^2} \sum_{k=1}^{\infty} \frac{x}{(k+1+\frac{1}{r})(k+\frac{1}{r})}$$

i.e. $\qquad x\,M_2'(0) \geq M_2(x) \geq M_0(x) \ldots \qquad (3.8)$

From this it follows that

$F_3(k) \leq F_1(k)$

However it can be verified that

$f_3(1) < f_1(1)$

but $\qquad f_3(k) > f_1(k)$ for all $k > 1$.

From (3.7)

$M_2'(x) = \sum_{r=1}^{\infty} \frac{1}{r^2} \psi'(x + \frac{1}{r})$ which shows that

$M_2'(x)$ is decreasing.

$M_2(x)$ as an infinite integral

Using the expression (3.4) in (3.7) one obtains that

$$R_2(x) = \sum_{k=1}^{\infty} \frac{1}{k^2} \left\{ \int_0^{\infty} \frac{e^{-\frac{t}{k}} - e^{-t(x+\frac{1}{k})}}{e^t - 1} \, dt \right\}$$

$$= \int_0^{\infty} \left\{ \sum_{k=1}^{\infty} \frac{1}{k^2} e^{-\frac{t}{k}} \left(\frac{1 - e^{-tx}}{e^t - 1} \right) dt \right\}$$

[Summation under the integral sign is justified since the series (3.7) is uniformly convergent].

$$R_2(x) = \int_0^{\infty} \frac{R(t)(1 - e^{-tx})}{e^t - 1} \, dt \quad \ldots \qquad (3.9)$$

where $$R(t) = \sum_{k=1}^{\infty} \frac{1}{k^2} e^{-\frac{t}{k}} \quad \ldots \qquad (3.10)$$

Since this function $R(t)$ will be used in obtaining some important formulae in the next section, some properties of this function are derived here.

Upper and lower bounds for $R(t)$

$R(t)$ is positive and decreasing for all positive values of t.

$$0 < R(t) \leq \zeta(2) \quad \text{for} \quad 0 < t < \infty .$$

But by applying a known formula [19] one obtains more precise bounds for $R(t)$.

$R(t)$ can be written as

$$R(t) = \sum_{k=1}^{\infty} u_k \quad \text{where} \quad u_k = \frac{1}{k^2} e^{-\frac{t}{k}} .$$

It can be verified easily that

$$U_{k+1} \le U_k \quad \text{and} \quad U_k \ge 0 \quad \text{for all} \quad k \ge 1$$

Then the theorem gives the result

$$U_1 + \int_1^\infty U_x dx \ge R(t) \ge \int_1^\infty U_x dx$$

i.e. $$e^{-t} + \int_1^\infty \frac{1}{x^2} e^{-\frac{t}{x}} dx \ge R(t) \ge \int_1^\infty \frac{1}{x^2} e^{-\frac{t}{x}} dx$$

$$e^{-t} + \frac{1}{t}(1-e^{-t}) \ge R(t) \ge \frac{1}{t}(1-e^{-t}) \ldots \qquad (3.11)$$

From this the upper and lower limits for $M_2(x)$ can be obtained as

$$\int_0^\infty \frac{1-e^{-tx}}{e^t-1}\left[\frac{(1-e^{-t})}{t} + e^{-t}\right] \ge M_2(x) \ge \int_0^\infty \frac{(1-e^{-tx})}{e^t-1}\frac{(1-e^{-t})}{t}\, dt$$

i.e. $$\log(1+x) + [\Psi(x+1) - \Psi(1)] \ge M_2(x) \ge \log(1+x) .$$

<u>Laplace transform of R(t)</u>

$$R(t) = \sum_{k=1}^\infty \frac{1}{k^2} e^{-\frac{t}{k}} \xrightarrow{\;s\;} \sum_{k=1}^\infty \frac{1}{k^2}\frac{1}{s+\frac{1}{k}}$$

$$= \sum_{k=1}^\infty \frac{1}{k(ks+1)}$$

$$= \sum_{k=1}^\infty \left(\frac{1}{k} - \frac{1}{k+\frac{1}{s}}\right)$$

$$= C + \Psi\left(\frac{1}{s}\right) = M_1\left(\frac{1}{s}\right)$$

$$H(t) \xrightarrow{o} H_1\left(\frac{1}{s}\right) \ldots \qquad (3.12)$$

($\xrightarrow{o}$ means that the right hand side is the Laplace transform of left hand side).

<u>H(t) as an infinite integral</u>

From (3.12) $H(t)$ has its Laplace transform as $H_1\left(\frac{1}{s}\right)$ or in otherwords $H_1\left(\frac{1}{s}\right)$ is the inverse transform of $H(t)$. Now the inverse transform of $H_1\left(\frac{1}{s}\right)$ is obtained as an infinite integral using the Bessel function $J_1(x)$ and thereby $H(t)$ is expressed as an infinite integral.

Consider

$$J_1(a) = \frac{a}{2} \sum_{k=0}^{\infty} \frac{(-1)^k a^{2k}}{k!\,(k+1)!\,2^{2k}}$$

and

$$\frac{J_1(2\sqrt{at})}{\sqrt{at}} = \sum_{k=0}^{\infty} \frac{(-1)^k a^k t^k}{k!\,(k+1)!}$$

$$\frac{J_1(2\sqrt{at})}{\sqrt{at}} \xrightarrow{o} \sum_{k=0}^{\infty} \frac{(-1)^k a^k}{(k+1)!\,s^{k+1}} = \frac{1}{a}\left(1 - e^{-\frac{a}{s}}\right)$$

$$\therefore \quad \frac{a\,J_1(2\sqrt{at})}{\sqrt{at}\,(e^a-1)} \xrightarrow{o} \frac{\left(1 - e^{-\frac{a}{s}}\right)}{(e^a - 1)}$$

$$\int_0^\infty \frac{u\, J_1(2\sqrt{ut})}{\sqrt{ut}\,(e^u-1)}\,du \;\xrightarrow{\quad}\; \int_0^\infty \frac{1-e^{-\frac{u}{s}}}{e^u-1}\,du = \Psi_1\left(\frac{1}{s}\right)$$

From the above formula and (3.12) one obtains

$$R(t) = \int_0^\infty \frac{u\, J_1(2\sqrt{ut})}{\sqrt{ut}\,(e^u-1)}\,du \ldots \qquad (3.13)$$

From this $\Psi_2(x)$ can be expressed as a double integral

$$\Psi_2(x) = \int_0^\infty \frac{1-e^{-tx}}{e^t-1} \int_0^\infty \frac{u\, J_1(2\sqrt{ut})}{\sqrt{ut}\,(e^u-1)}\,du\,dt .$$

<u>$R(t)$ in terms of the Laplace transform of $T(x)$</u>

From Chapter II

$$R(s) = \int_0^1 e^{-sx}\,T(x)dx = \sum_{n=0}^\infty \frac{(-1)^n s^n}{n!}\,\mu_n$$

where $\mu_n = \left[\frac{1}{n} - \frac{\zeta(n+1)}{n+1}\right]$

$$R(t) = \sum_{k=1}^\infty \frac{1}{k^2}\, e^{-\frac{t}{k}}$$

$$= \sum_{k=1}^\infty \frac{1}{k^2} \left\{ \sum_{r=0}^\infty \frac{(-1)^r t^r}{k^r\, r!} \right\}$$

$$= \sum_{r=0}^\infty \frac{(-1)^r t^r}{r!}\,\zeta(r+2)\ldots \qquad (3.14)$$

$$\int_0^1 x\,T(x)\,e^{-tx}\,dx = \sum_{n=0}^\infty \frac{(-1)^n}{n!}\,t^n\,\mu_{n+1}$$

and $$\int_0^1 x^2\,T(x)\,e^{-tx}\,dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\,t^n\,\mu_{n+2}$$

$$\int_0^1 (t\,x^2-2x)\,T(x)e^{-tx}\,dx = [-2+\zeta(2)] + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}t^n}{(n-1)!}\left\{\frac{2\mu_{n+1}}{n} + \mu_{n+1}\right\}$$

$$= [-2+\zeta(2)] + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!}\,t^n\,(n+2)\,\mu_{n+1}\,,$$

substituting the value of μ_n

$$= [-2+\zeta(2)] + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}\,t^n(n+2)}{(n+1)!} + \sum_{n=1}^{\infty} \frac{(-1)^n\,t^n}{n!}\,\zeta(n+2)$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^{n+1}\,t^n}{n!} + \sum_{n=0}^{\infty} \frac{(-1)^{n+1}\,t^n}{(n+1)!} + \sum_{n=0}^{\infty} \frac{(-1)^n\,t^n}{n!}\,\zeta(n+2)$$

$$= -e^{-t} + \frac{(e^{-t}-1)}{t} + R(t)$$

$$\therefore R(t) = \int_0^1 (t\,x^2-2x)\,T(x)\,e^{-tx}\,dx + e^{-t} + \frac{(1-e^{-t})}{t} \qquad (3.15)$$

Analogous to the treatment of $M_1'(x)$, one can easily derive from (3.9) for $M_2'(x)$:

$$M_2'(x) = \int_0^{\infty} \frac{t}{e^t-1}\,R(t)\,e^{-tx}\,dt \;\ldots \qquad (3.16)$$

which shows that $\frac{t}{e^t-1}\,R(t)$ is the inverse transform of $M_2'(x)$.

From (3.16) and (3.11) one obtains that

$$\int_0^\infty e^{-t} \cdot e^{-tx}\, dt - \int_0^\infty \frac{t\, e^{-t(1+x)}}{e^t - 1}\, dt \geq \phi_2(x) \geq \int_0^\infty e^{-t} \cdot e^{-tx}\, dt \;.$$

i.e.
$$\frac{1}{1+x} - \psi'(x+1) \geq \phi_2(x) \geq \frac{1}{1+x} \qquad (3.17)$$

Using Euler's formula to the infinite series (3. 7), the values of $\phi_2(x)$ for $x = \frac{1}{k}$ are calculated. The details of this calculation and the tabulated values of $\phi_2(x)$ are given in Appendix (A). In table I, $F_1(k)$, $F_2(k)$, $F_3(k)$ and $F_\infty(k)$ are tabulated and in table II, $f_1(k)$, $f_2(k)$, $f_3(k)$ and $f_\infty(k)$ are tabulated. From these two tables one can observe the following inequalities

$$F_2(k) < F_\infty(k) < F_3(k) < F_1(k) \quad \text{for all } k$$

where
$$F_\infty(k) = \frac{\log\left(1 + \frac{k}{k+2}\right)}{\log 2}$$

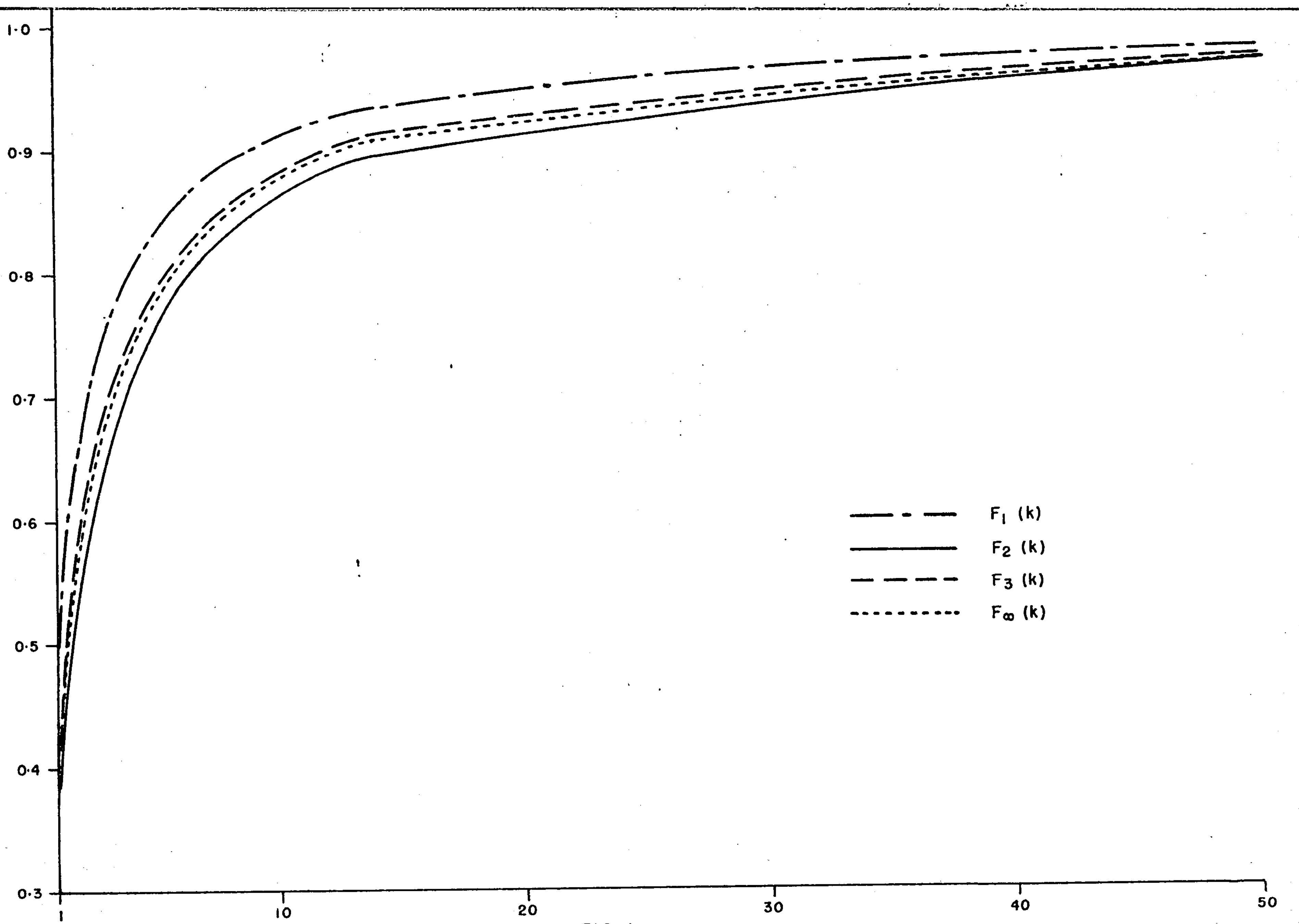
and
$$f_3(k) > f_2(k) \quad \text{for } k = 1 \text{ and } 2$$

$$f_3(k) < f_2(k) \quad \text{for } k \geq 3 \;.$$

Table II shows that the convergence of $f_n(k)$ to $f_\infty(k)$ as $k \to \infty$ is very rapid. The graphs of $F_1(k)$, $F_2(k)$, $F_3(k)$ and $F_\infty(k)$ are given in Figure 1. This figure shows the convergence of $F_n(k)$ to $F_\infty(k)$ as $n \to \infty$ is also rapid. In figure 2 the graphs of $(\phi_0(x) - \phi_\infty(x))$, $(\phi_1(x) - \phi_\infty(x))$, $(\phi_2(x) - \phi_\infty(x))$ are given. From figure 2 one can see that these graphs alternate in sign.

Fig 1

Distribution functions $F_1(k)$, $F_2(k)$, $F_3(k)$ and $F_\infty(k)$ (though these functions are discrete, it is more convenient to draw them as continuous when we are drawing the four curves in the same graph).



1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
1
10
20
30
40
50
F_1 (k)
F_2 (k)
F_3 (k)
F_∞ (k)

Fig 2

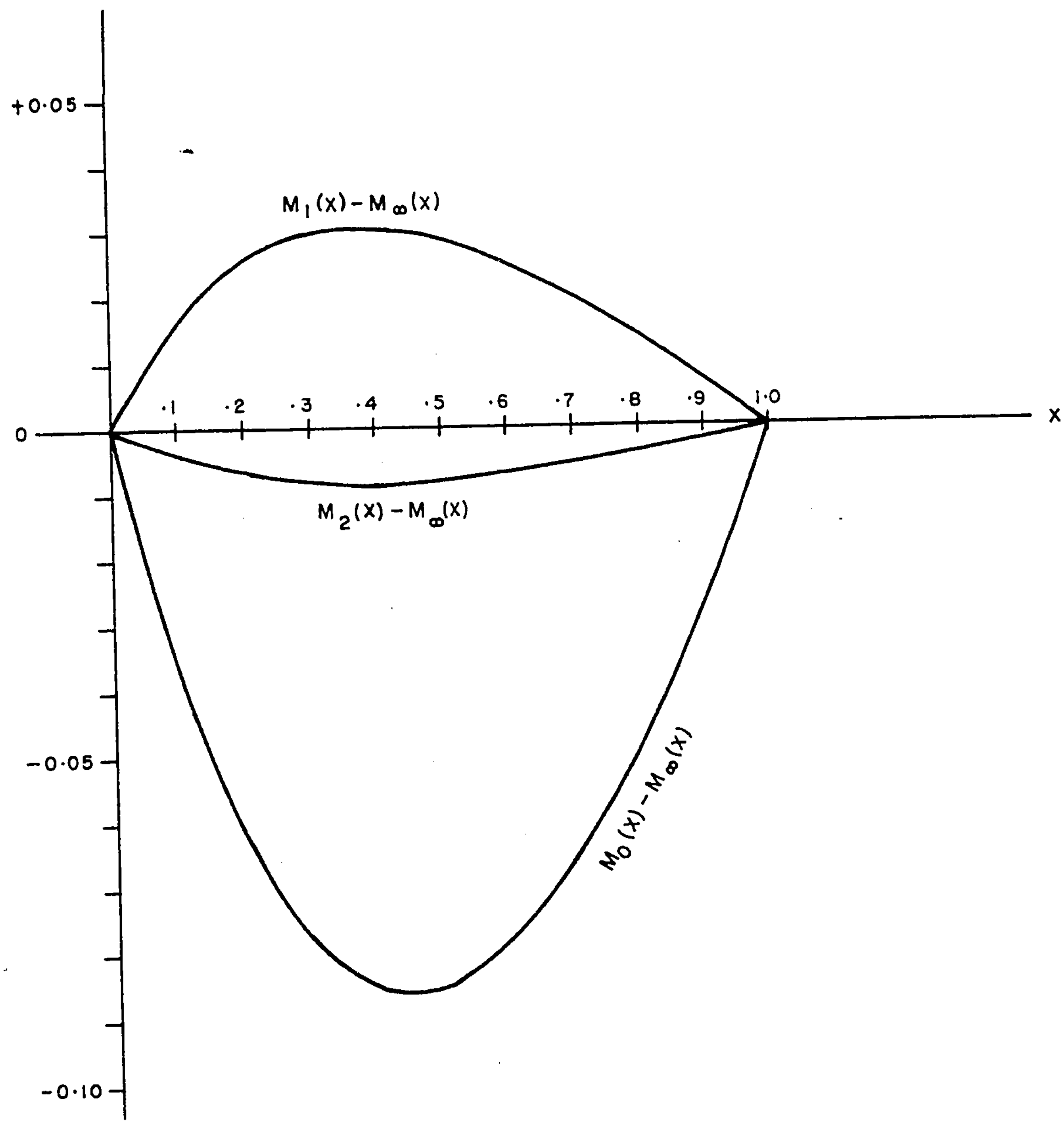


TABLE I

THEORETICAL DISTRIBUTION FUNCTIONS $F_1(k)$, $F_2(k)$, $F_3(k)$ AND $F_\infty(k)$

k	$F_1(k)$	$F_2(k)$	$F_3(k)$	$F_\infty(k)$
1	.5000	.3865	.4238	.4150
2	.6667	.5550	.5940	.5849
3	.7500	.6505	.6062	.6780
4	.8000	.7120	.7442	.7369
5	.8333	.7549	.7842	.7775
6	.8571	.7870	.8131	.8073
7	.8750	.8117	.8354	.83[illegible]
8	.8889	.8316	.8528	.8479
9	.9000	.8468	.8669	.8624
14	.9333	.8953	.9102	.9069
49	.9800	.9678	.9725	.9714
99	.9900	.9839	.9865	.9857
∞	1.0000	1.0000	1.0000	1.0000

TABLE 2

THEORETICAL VALUES OF FUNCTIONS $f_1(k)$, $f_2(k)$, $f_3(k)$ AND $f_\infty(k)$

k	$f_1(k)$	$f_2(k)$	$f_3(k)$	$f_\infty(k)$
1	.5000	.3865	.4238	.4150
2	.1667	.1685	.1702	.1699
3	.0833	.0955	.0922	.0931
4	.0500	.0615	.0580	.0589
5	.0333	.0429	.0400	.0406
6	.0238	.0321	.0289	.0298
7	.0179	.0247	.0223	.0227
8	.0139	.0193	.0174	.0179
9	.0111	.0158	.0141	.0145
$10 \leq k < 15$	.0333	.0485	.0432	.0445
$15 \leq k < 50$	.0467	.0725	.0624	.0645
$50 \leq k < 100$	.0100	.0161	.0140	.0143
$100 \leq k < \infty$	.0100	.0161	.0135	.0143

CHAPTER IV

GENERAL LAWS OF $M_n(x)$

All the properties derived in the previous chapter are generalised for $M_n(x)$ in this chapter. The inverse transform $Q_n(t)$ of $M_n(x)$ is obtained and a few properties of $Q_n(t)$ are discussed.

(A) <u>Difference equation of $M_n(x)$</u>

$M_n(x)$ for all $n \geq 1$ is defined by the equation

$$M_n(x) = \sum_{k=1}^{\infty} M_n(\tfrac{1}{k}) - M_n(\tfrac{1}{k+x}) \quad \text{for} \quad 0 \leq x \leq 1 .$$

But even for values $x > 1$, the right hand side of the equation is defined since $\frac{1}{k+x}$ is always less than 1. Therefore one can define $M_n(x)$ for all positive values of x by the above equation.

Now consider

$$M_{n+1}(x+1) - M_{n+1}(x) = \sum_{k=1}^{\infty} M_n(\tfrac{1}{k+x}) - M_n(\tfrac{1}{k+x+1})$$

$$= M_n(\tfrac{1}{1+x}) - \lim_{k \to \infty} M_n(\tfrac{1}{k+x+1})$$

$$= M_n(\tfrac{1}{1+x}) \quad \text{since} \quad M_n(0) = 0$$

$$M_{n+1}(x+1) - M_{n+1}(x) = M_n(\tfrac{1}{1+x}) \ldots \qquad (4.1)$$

If one denotes by $F(x) = \lim_{n \to \infty} M_n(x)$ it follows from (4.1) that

$$F(x+1) - F(x) = F\left(\frac{1}{1+x}\right)$$

or $$Q(x) = -Q\left(\frac{1}{x}\right) \quad \text{where} \quad Q(x) = F(1+x) .$$

As mentioned earlier in Chapter 1 the function $Q(x) = c \log x$ satisfies this equation. [Here $M_0(x)$ need not be x] .

Expected values of a_n

$$F_n(k) = 1 - M_{n-1}\left(\frac{1}{k+1}\right)$$

and therefore $f_n(k) = \text{Prob}(a_n = k) = M_{n-1}\left(\frac{1}{k}\right) - M_{n-1}\left(\frac{1}{k+1}\right)$

$$E(k^r) = \sum_{k=1}^{\infty} k^r \quad M_{n-1}\left(\frac{1}{k}\right) - M_{n-1}\left(\frac{1}{k+1}\right)$$

$$= \sum_{k=1}^{\infty} \frac{k^r}{k(k+1)} M'_{n-1}\left(\frac{1}{k+\theta_k}\right)$$

$$> \sum_{k=1}^{\infty} \frac{k^{r-1} m}{k+1} \quad \text{where } m \text{ is the lower bound of } M'_{n-1}(x)$$

$$= \infty \quad \text{for } r \geq 1 .$$

Similarly it can be verified that all the moments for negative powers of k exist.

([illegible]) Upper and lower bounds for $M_n(x)$

According to the lemma of Kuzmin [7] if

$$\frac{\lambda}{1+x} < M'_n(x) < \frac{\gamma}{1+x} \quad \text{then}$$

$$\frac{t}{1+x} < \Psi'_{n-1}(x) < \frac{T}{1+x}$$

or if

$$\frac{t}{1+x} < \Psi'_{0}(x) < \frac{T}{1+x}$$

then

$$\frac{t}{1+x} < \Psi'_{n}(x) < \frac{T}{1+x} \quad \text{for all } n .$$

In case $\Psi_0(x) = x$ $\Psi'_0(x) = 1$ and therefore

$$\frac{1}{1+x} \leq \Psi'_{0}(x) \leq \frac{2}{1+x}$$

$$\frac{1}{1+x} \leq \Psi'_{n}(x) \leq \frac{2}{1+x}$$

From this it follows that

$$\log(1+x) \leq \Psi_n(x) \leq 2\log(1+x) .$$

For $F_n(k)$, one considers values in $(0, \frac{1}{2})$. For that interval

$$\log(1+x) \leq \Psi_n(x) \leq \frac{3}{2}\log(1+x) \quad \text{for} \quad 0 \leq x \leq \frac{1}{2}$$

$$\underline{x\,\Psi'_n(0) \geq \Psi_n(x) \geq x}$$

This result is proved for $n = 1$ and 2 in Chapter 3.

First an expression for $\Psi'_n(x)$ is obtained and from that the above inequality is obtained.

$$\Psi'_0(x) = 1$$

$$\Psi'_1(x) = \sum_{k=1}^{\infty} \frac{1}{(k+x)^2}$$

$$m_2'(x) = \sum_{k_2=1}^{\infty} \sum_{k_1=1}^{\infty} \frac{1}{[k_2k_1+k_1x+1]^2} = \sum_{k_2=1}^{\infty}\sum_{k_1=1}^{\infty} \frac{1}{[f(k_2,k_1)+xf(k_1)]^2}$$

In general it can be proved by induction that

$$m_n'(x) = \sum_{k_n=1}^{\infty} \sum_{k_{n-1}=1}^{\infty} \cdots \sum_{k_1=1}^{\infty} \frac{1}{[f(k_n,k_{n-1}\cdots k_1)+x\, f(k_{n-1},k_{n-2}\cdots k_1)]^2}$$

where $f(k_1) = 1$

$$f(k_2, k_1) = k_2k_1 + 1$$

$$\vdots$$

$$f(k_n k_{n-1}\cdots k_1) = k_n f(k_{n-1},k_{n-2}\cdots k_1) + f(k_{n-2}\cdots k_1) .$$

These $f(k_n, k_{n-1}\cdots k_1)$ can be identified with the denominators q_n of the convergents of a simple continued fraction. We therefore write here q_n instead $f(k_n\cdots k_1)$.

$$m_n'(x) = \sum^{(n)} \frac{1}{[q_n+x\, q_{n-1}]^2} \cdots \qquad (4.2)$$

where $\sum^{(n)}$ means the summation over all possible values of q_n and q_{n-1}. From (4.2) one can observe easily that <u>$m_n'(x)$ is a decreasing function</u> of x for all n.

Integrating (4.2) from 0 to x one has

$$m_n(x) = \sum^{(n)} \frac{x}{q_n(q_n+x\, q_{n-1})} \qquad (4.3)$$

Since $m_n(1) = 1$

$$\sum^{(n)} \frac{1}{q_n(q_n+q_{n-1})} = 1$$

$$\psi_n(x) = \sum^{(n)} \frac{x}{q_n(q_n+x\,q_{n-1})} \geq \sum^{(n)} \frac{x}{q_n(q_n+q_{n-1})} = x$$

Similarly $\psi_n(x) \leq \sum^{(n)} \frac{x}{q_n^2} = x\,\psi_n'(0)$

$$\therefore \quad x\,\psi_n'(0) \geq \psi_n(x) \geq x$$

$$\psi_n'(0) = \sum^{(n)} \frac{1}{q_n^2} > \sum^{(n)} \frac{1}{q_n(q_n+q_{n-1})} = 1$$

and $$\psi_n'(1) = \sum^{(n)} \frac{1}{(q_n+q_{n-1})^2} < \sum^{(n)} \frac{1}{q_n(q_n+q_{n-1})} = 1$$

$$\therefore \quad \psi_n'(0) > 1$$

and $$\psi_n'(1) < 1$$

$$\psi_n'(\tfrac{1}{2}) = \sum^{(n)} \frac{1}{[q_n+\frac{1}{2}q_{n-1}]^2} < 1$$

Consider

$$[q_n+x\,q_{n-1}]^2 = q_n^2 + 2x\,q_nq_{n-1} + x^2q_{n-1}^2$$

Now $$(q_n^2 + 2x\,q_nq_{n-1} + x^2q_{n-1}^2) - q_n(q_n + q_{n-1})$$

$$= 2x\,q_nq_{n-1} + x^2q_{n-1}^2 - q_nq_{n-1}$$

$$= q_n(2x-1) + x^2q_{n-1} < q_n(2x - 1 + x^2) < 0$$

if $(x+1)^2 < 2$

or $x < (\sqrt{2}-1)$

$$M_n'(x) = \sum^{(n)} \frac{1}{(q_n + x\, q_{n-1})^2} > \sum^{(n)} \frac{1}{q_n(q_n + q_{n-1})} \quad \text{for} \quad 0 < x < (\sqrt{2}-1)$$

or $$M_n'(x) > 1 \quad \text{for} \quad 0 \leq x \leq (\sqrt{2}-1)$$

and $$M_n'(x) < 1 \quad \text{for } \frac{1}{2} \leq x \leq 1$$

Since $M_n'(x)$ is decreasing and continuous, there exists a constant c_n such that

$$M_n'(c_n) = 1 \quad \text{for} (\sqrt{2}-1 < c_n < \tfrac{1}{2})$$

$$\text{i.e.} \quad (.4142 < c_n < .5)$$

From this follows that

$$f_{n+1}(k) = M_n(\tfrac{1}{k}) - M_n(\tfrac{1}{k+1})$$

$$= \frac{1}{k(k+1)} \; M_n'(\tfrac{1}{k+\theta})$$

$$< \frac{1}{k(k+1)} \quad \text{for} \quad k = 1$$

and $f_{n+1}(k) > \dfrac{1}{k(k+1)}$ for $k \geq 3$.

By direct substitution, one can verify that

$$f_{n+1}(2) > \frac{1}{k(k+1)}$$

$$\therefore \; f_{n+1}(1) < f\,(1) \quad \text{for all} \quad n \geq 1$$

$$f_{n+1}(k) > f_1(k) \quad \text{for all} \quad k \geq 2 \quad \text{and for all} \quad n \geq 1 .$$

Since $M'_n(x)$ is decreasing

$$M'_n(1) \leq M'_n(x) \leq M'_n(0)$$

$$\sum_{k=1}^{\infty} \frac{1}{(k+x)^2} M'_n(1) \leq \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} M'_n\left(\frac{1}{k+x}\right) \leq \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} M'_n(0)$$

i.e. $$M'_n(1)\psi^1(x) \leq M'_{n+1}(x) \leq M'_n(0)\,\psi^1(x)$$

From this follows that

$$M'_n(1)\ \psi(x) \leq M_{n+1}(x) \leq M'_n(0)\ \psi(x)$$

(C) <u>Expressions for $M_n(x)$</u>

An expression for $M_n(x)$ similar to (3.7) will be derived here

$$M_{n+1}(x) = \sum^{(n+1)} \frac{x}{q_{n+1}(q_{n+1}+x\,q_n)}$$

$$= \sum^{(n+1)} \frac{1}{q_n} \left[\frac{1}{q_{n+1}} - \frac{1}{q_{n+1}+q_n x}\right]$$

$$= \sum^{(n+1)} \frac{1}{q_n} \left[\frac{1}{k\,q_n+q_{n-1}} - \frac{1}{k\,q_n+q_{n-1}+q_n\,x}\right]$$

$$= \sum^{(n)} \sum_{k=1} \frac{1}{q_n} \left[\frac{1}{k\,q_n+q_{n-1}} - \frac{1}{k\,q_n+q_{n-1}+q_n\,x}\right]$$

$$= \sum^{(n)} \frac{1}{q_n^2} \sum_{k=1} \left[\frac{1}{k+\frac{q_{n-1}}{q_n}} - \frac{1}{k+x+\frac{q_{n-1}}{q_n}}\right]$$

$$= \sum^{(n)} \frac{1}{q_n^2} \left\{ \psi\left(\frac{x+q_{n-1}}{q_n}\right) - \psi\left(\frac{q_{n-1}}{q_n}\right) \right\} \qquad (4.4)$$

Using the Maclaurin expansion of $M_n(x)$, $M_{n+1}(x)$ can be written in the following way

$$M_{n+1}(x) = \sum_{k=1}^{\infty} M_n\left(\frac{1}{k}\right) - M_n\left(\frac{1}{k+x}\right)$$

$$= \sum_{k=1}^{\infty} \left[\left(\frac{1}{k} - \frac{1}{k+x}\right) M_n'(0) + \left(\frac{1}{k^2} - \frac{1}{(k+x)^2}\right) \frac{M_n''(0)}{2!} + \cdots \right]$$

$$= \sum_{k=1}^{\infty} \frac{M_n^{(k)}(0)}{k!} \left[\psi^{(k-1)}(x) - \psi^{(k-1)}(0) \right] \frac{(-1)^{k-1}}{(k-1)!}$$

where $\psi^{(k)}(x)$ is the k^{th} derivative of $\psi(x)$.

(D) Inverse transform of $M_n(s)$

In Chapter III, it is known that the function $R(t)$ satisfies the property that

$$M_2(x) = \int_0^{\infty} \frac{1 - e^{-tx}}{e^t - 1} R(t)dt .$$

Here in general $R_n(t)$ is defined where $R_n(t)$ satisfies the following equation

$$M_{n+1}(x) = \int_0^{\infty} \frac{R_n(t)\,(1 - e^{-tx})}{e^t - 1} dt .$$

[Def]:- Let $R_0(t) = 1$

$$R_1(t) = \int_0^{\infty} \frac{U\,J_1(2\sqrt{Ut})}{\sqrt{Ut}\,e^U - 1} dU .$$

Define $R_{n+1}(t) = \int_0^\infty \frac{R_n(u) \cdot u \cdot J_1(2\sqrt{ut})}{\sqrt{ut}\,(e^u - 1)}\, du$.

$R_1(t)$ is the function $R(t)$ defined in Chapter III . $R_n(t)$ exists for all n .

Now
$$\Psi_{n+1}(x) = \int_0^\infty \frac{R_n(t)\,(1 - e^{-tx})}{e^t - 1}\, dt \quad \ldots \qquad (4.5)$$

Proof: For $n = 0$ and $n = 1$ (4.5) is true. Let it be true for n . Then

$$\Psi_n(x) = \int_0^\infty \frac{R_{n-1}(t)\,(1 - e^{-tx})}{e^t - 1}\, dt$$

Consider
$$\Psi_{n+1}(x) = \sum_{k=1}^\infty \int_0^\infty \frac{R_{n-1}(t)}{(e^t - 1)} \left(e^{-\frac{t}{k+x}} - e^{-\frac{t}{k}}\right) dt$$

$$= \int_0^\infty \sum_{k=1}^\infty \frac{R_{n-1}(t)}{e^t - 1} \left(e^{-\frac{t}{k+x}} - e^{-\frac{t}{k}}\right) dt \quad \ldots \qquad (4.6)$$

Now
$$\sum_{k=1}^\infty e^{-\frac{t}{k+x}} - e^{-\frac{t}{k}} = \sum_{k=1}^\infty \left\{ \left(\frac{t}{k} - \frac{t}{k+x}\right) - \frac{t^2}{2!}\left(\frac{1}{k^2} - \frac{1}{(k+x)^2}\right) + \cdots \right\}$$

$$= t[\psi(x) - \psi(0)] + \frac{t^2}{2!}[\psi'(x) - \psi'(0)] + \frac{t^3}{3!2!}[\psi''(x) - \psi''(0)]$$

$$+ \cdots + \frac{t^n}{n!(n-1)!}[\psi^{n-1}(x) - \psi^{n-1}(0)] + \cdots$$

But
$$\psi(x) - \psi(0) = \int_0^\infty \frac{1 - e^{-ux}}{e^u - 1}\, du$$

$$\psi'(x) - \psi'(0) = \int_0^\infty - \frac{u(1 - e^{-ux})}{e^u - 1}\, du$$

$$\vdots$$

$$\psi^n(x) - \psi^n(0) = \int_0^\infty \frac{(-1)^n u^n (1-e^{-ux})}{(e^u - 1)}\, du \; .$$

Using these formulae

$$\sum_{k=1}^\infty e^{-\frac{t}{k+x}} - e^{-\frac{t}{k}} = \int_0^\infty \frac{1 - e^{-ux}}{e^u - 1} \left[\frac{t}{1!0!} - \frac{t^2 u}{2!1!} + \frac{t^3 u^2}{3!2!} - \frac{t^4 u^3}{4!3!} \cdots + \right] du$$

$$= \int_0^\infty \frac{1 - e^{-ux}}{e^u - 1} \left\{ t \sum_{k=0}^\infty \frac{(-1)^k t^k u^k}{k!\,(k+1)!} \right\} du$$

$$= \int_0^\infty \frac{(1 - e^{-ux}) t}{(e^u - 1)} \frac{J_1(2\sqrt{ut})}{\sqrt{ut}}\, du$$

Substituting this in the formula (4.6)

$$R_{n+1}(x) = \int_0^\infty \frac{R_{n-1}(t)}{e^t - 1} \left\{ \int_0^\infty \frac{(1 - e^{-ux}) t}{(e^u - 1)} \frac{J_1(2\sqrt{ut})}{\sqrt{ut}}\, du \right\} dt$$

$$= \int_0^\infty \int_0^\infty \frac{R_{n-1}(t)\, t\, (1 - e^{-ux})\, J_1(2\sqrt{ut})}{(e^t - 1)\; (e^u - 1)\; \sqrt{ut}}\, du\, dt \; .$$

Interchanging the order of integration

$$= \int_0^\infty \int_0^\infty \frac{R_{n-1}(t)\, t\, (1 - e^{-ux})\, J_1(2\sqrt{ut})}{(e^t - 1)\; (e^u - 1)\; \sqrt{ut}}\, dt\, du$$

$$= \int_0^\infty \frac{1 - e^{-ux}}{(e^u - 1)} \left\{ \int_0^\infty \frac{R_{n-1}(t)\, t\, J_1(2\sqrt{ut})}{(e^t - 1)\ \sqrt{ut}}\, dt \right\} du$$

$$= \int_0^\infty \frac{1 - e^{-ux}}{e^u - 1}\, R_n(u)\, du \text{ . from the definition of } R_n(u)$$

$$M_n(x) = \int_0^\infty \frac{1 - e^{-ux}}{e^u - 1}\, R_{n-1}(u)\, du \quad \text{for all} \quad n \geq 1$$

Now from the difference equation (4.1) and the above formula

$$M_n(\tfrac{1}{1+x}) = \int_0^\infty \frac{R_n(u)\,(e^{-ux} - e^{-u(x+1)})}{e^u - 1}\, du$$

$$= \int_0^\infty R_n(u)\, e^{-u(x+1)}\, du \ .$$

i.e. $M_n(\frac{1}{x}) = \int_0^\infty R_n(u) e^{-ux}\, du$ which shows that $R_n(u)$ is the inverse transform of $M_n(\frac{1}{x})$

Again from (4.5) one obtains

$$M_n'(x) = \int_0^\infty \left(\frac{R_{n-1}(t)\, t}{(e^t - 1)}\right) e^{-tx}\, dt \ \ldots \qquad (4.6)$$

<u>$R_n(t)$ as an infinite sum</u>

In Chapter III $R_1(t)$ is expressed as the sum of an infinite series. Now from (4.6) $R_n(t)$ is expressed as an infinite sum.

$$M_n'(x) = \sum^{(n)} \frac{1}{[q_n + x\, q_{n-1}]^2} \quad \text{from } (4.2)$$

$M_n'(x)$ is the transform of the function

$$\sum^{(n)} \frac{t\, e^{-\frac{q_n}{q_{n-1}} t}}{q_{n-1}^2}$$

$$= \sum^{(n-1)} \sum_{k=1}^{\infty} \frac{t\, e^{-[k + \frac{q_{n-2}}{q_{n-1}}] t}}{q_{n-1}^2}$$

$$= \sum^{(n-1)} \frac{e^{-\frac{q_{n-2}}{q_{n-1}} t}}{q_{n-1}^2} \sum_{k=1}^{\infty} t\, e^{-tk}$$

$$= \frac{t}{(e^t - 1)} \sum^{(n-1)} \frac{e^{-\frac{q_{n-2}}{q_{n-1}} t}}{q_{n-1}^2}$$

$$\therefore R_{n-1}(t) = \sum^{(n-1)} \frac{e^{-\frac{q_{n-2}}{q_{n-1}} t}}{q_{n-1}^2}$$

<u>Upper and lower bounds for $R_n(t)$</u>

If $c_1 \frac{(1 - e^{-t})}{t} < R_n(t) < c_2 \frac{(1 - e^{-t})}{t}$ then

$$c_1 \frac{(1 - e^{-t})}{t} < R_{n+1}(t) < c_2 \frac{(1 - e^{-t})}{t} .$$

Consider

$$R_{n+1}(t) = \int_0^{\infty} \frac{R_n(u)\; u\, J_1(2\sqrt{ut})}{\sqrt{ut}\; (e^u - 1)}\, du$$

$$< \int_0^{\infty} \frac{c_2 (1 - e^{-u})\; u\, J_1(2\sqrt{ut})}{u \quad \sqrt{ut}\, (e^u - 1)}\, du$$

$$= C_2 \int_0^\infty \frac{J_1(2\sqrt{ut})}{\sqrt{ut}} e^{-u} \, du = \frac{C_2(1 - e^{-t})}{t} .$$

Similarly the other inequality also follows. Since

$$\frac{(1 - e^{-t})}{t} < R_1(t) < \frac{2(1 - e^{-t})}{t}$$

by induction $\dfrac{(1 - e^{-t})}{t} < R_n(t) < 2\,\dfrac{(1 - e^{-t})}{t}$.

CHAPTER V

STATISTICAL DEPENDENCE OF a_i's

Continued fractions and systematic fractions have certain properties in common. Both are used in approximating real numbers by rational numbers. However the continued fractions give better and simpler approximations than the systematic fractions. On the other hand systematic fractions are more convenient to use in arithmetical calculations (for example; sum of two systematic fractions can be easily calculated). From the Appendix B it can be seen that the distribution functions of the index numbers of a systematic fraction are much simpler than those of the simple continued fraction. In particular where α is uniformly distributed in $(0, 1)$, the $a_n(\alpha)$'s (index numbersof α) of a systematic fraction are statistically independent. This was proved by Borel in 1909 for a binary fraction.

i.e. If α is uniformly distributed and

$$\alpha = \frac{a_1(\alpha)}{2} + \frac{a_2(\alpha)}{2^2} + \dots + \frac{a_n(\alpha)}{2^n} + \dots$$

where $a_n(\alpha) = 0$ or 1 for all n, then

$$\text{Prob}\left\{a_n(\alpha) = k \,;\, a_m(\alpha) = \ell\right\} = \text{Prob}\left\{a_n(\alpha) = k\right\} \cdot \text{Prob}\left\{a_m(\alpha) = \ell\right\}$$

for all m and n.

In general for any systematic fraction with root 'g', the

$a_n(\omega)$'s are statistically independent if ω is uniformly distributed [20]

Now considering the problem of independence for continued fractions, it can be shown that a_n and a_{n+1} of a simple continued fraction are not independent. Before obtaining the general expression for $\text{Prob}(a_n = k, a_{n+1} = r)$, expressions for $\text{Prob}(a_1 = k, a_2 = r)$, $\text{Prob}(a_2 = k, a_3 = r)$ and $\text{Prob}(a_3 = r, a_1 = k)$ are obtained.

<u>$\text{Prob}(a_1 = k, a_2 = r)$</u>

$$\text{Prob}(a_1 = k, a_2 = r) = \frac{1}{k+\frac{1}{r+1}} - \frac{1}{k+\frac{1}{r}} \qquad (5.1)$$

[$a_1 = k$ implies that ω must be in $\left(\frac{1}{k+1}, \frac{1}{k}\right)$ and $a_2 = r$ determines the sub-interval $\left(\frac{1}{k+\frac{1}{r+1}}, \frac{1}{k+\frac{1}{r}}\right)$] .

Now $$\text{Prob}(a_1 = k) \cdot \text{Prob}(a_2 = r) = \frac{1}{k(k+1)}\left\{\psi\left(\frac{1}{r}\right) - \psi\left(\frac{1}{r+1}\right)\right\} \qquad (5.2)$$

from Chapter III.

If a_1 and a_2 are independent (5.1) and (5.2) must be equal for all pairs of values of k and r. For $k = 1, r = 1$

$$(5.1) = \frac{1}{6} = .1666\ldots$$

and $$(5.2) = \frac{1}{2} \cdot .386 = .193$$

Therefore a_1 and a_2 are not independent.

<u>$\text{Prob}(a_2 = k, a_3 = r)$</u>

$$\text{Prob}(a_1 = m,\ a_2 = k,\ a_3 = r) = M_0\left(\frac{1}{m+\frac{1}{k+\frac{1}{r}}}\right) - M_0\left(\frac{1}{m+\frac{1}{k+\frac{1}{r+1}}}\right).$$

From this

$$\text{Prob}(a_2 = k,\ a_3 = r) = \sum_{m=1}^{\infty} M_0\left(\frac{1}{m+\frac{1}{k+\frac{1}{r}}}\right) - M_0\left(\frac{1}{m+\frac{1}{k+\frac{1}{r+1}}}\right)$$

$$= M_1\left(\frac{1}{k+\frac{1}{r+1}}\right) - M_1\left(\frac{1}{k+\frac{1}{r}}\right) \quad \text{since}$$

$$\sum_{m=1}^{\infty} M_0\left(\frac{1}{m+x}\right) - M_0\left(\frac{1}{m+y}\right) = M_1(y) - M_1(x).$$

Therefore $\text{Prob}(a_2 = k,\ a_3 = r) = \psi\left(\frac{1}{k+\frac{1}{r+1}}\right) - \psi\left(\frac{1}{k+\frac{1}{r}}\right)$ (5.3)

$$\text{Prob}(a_2 = k)\cdot\text{Prob}(a_3 = r) = \left\{\psi\left(\tfrac{1}{k}\right) - \psi\left(\tfrac{1}{k+1}\right)\right\}\left\{M_2\left(\tfrac{1}{r}\right) - M_2\left(\tfrac{1}{r+1}\right)\right\} \quad (5.4)$$

Again for $k = 1 \quad r = 1$

(5.3) = .1448

and (5.4) = .16405

Therefore a_2 and a_3 are not independent.

<u>$\text{Prob}(a_1 = k,\ a_3 = r)$</u>

$$\text{Prob}(a_1 = k,\ a_2 = m,\ a_3 = r) = \frac{1}{k+\frac{1}{m+\frac{1}{r}}} - \frac{1}{k+\frac{1}{m+\frac{1}{r+1}}}$$

From this

$$\text{Prob}(a_1 = k,\ a_3 = r) = \sum_{m=1}^{\infty} \frac{1}{k+\cfrac{1}{m+\cfrac{1}{r}}} - \frac{1}{k+\cfrac{1}{m+\cfrac{1}{r+1}}}$$

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

$$= \sum_{m=1}^{\infty} \frac{\frac{1}{r(r+1)}}{[(m+\frac{1}{r})k+1][(m+\frac{1}{r+1})k+1]}$$

$$= \sum_{m=1}^{\infty} \frac{1}{k^2} \left[\frac{1}{m+\frac{1}{r+1}+\frac{1}{k}} - \frac{1}{m+\frac{1}{r}+\frac{1}{k}}\right]$$

$$= \frac{1}{k^2} \left\{ \psi\left(\frac{1}{r}+\frac{1}{k}\right) - \psi\left(\frac{1}{r+1}+\frac{1}{k}\right) \right\} \qquad (5.5)$$

$$\text{Prob}(a_1 = k) \cdot \text{Prob}(a_3 = r) = \frac{1}{k(k+1)} \left[M_2\left(\frac{1}{r}\right) - M_2\left(\frac{1}{r+1}\right)\right] \qquad (5.6)$$

For $k = 1$, $r = 1$

(5.5) = .2199

and (5.6) = .2145

Therefore a_1 and a_3 are not independent.

<u>$\text{Prob}(a_n = k,\ a_{n+1} = r)$</u>

Remembering the definition of Z_n

$$\text{Prob}(a_{n+1} = r,\ a_n = k) = \text{Prob}\left\{ \left[\frac{1}{Z_n}\right] = r,\ a_n = k \right\}$$

$$= \text{Prob}\left\{ r \le \frac{1}{Z_n} < r+1\ ;\ a_n = k \right\}$$

$$= \text{Prob}\left\{\frac{1}{r+1} < Z_n \leq \frac{1}{r} ;\ a_n = k\right\}$$

But $Z_n = \frac{1}{Z_{n-1}} - [\frac{1}{z_{n-1}}] = \frac{1}{z_{n-1}} - a_n$, so we have

$$\text{Prob}(a_n = k,\ a_{n+1} = r) = \text{Prob}\left\{\frac{1}{r+1} < \frac{1}{z_{n-1}} - a_n \leq \frac{1}{r} ,\ a_n = k\right\}$$

$$= \text{Prob}\left\{\frac{1}{r+1} < \frac{1}{z_{n-1}} - k \leq \frac{1}{r}\right\}$$

$$= \text{Prob}\left\{\frac{1}{k+\frac{1}{r}} \leq z_{n-1} < \frac{1}{k+\frac{1}{r+1}}\right\}$$

$$= M_{n-1}\left(\frac{1}{k+\frac{1}{r+1}}\right) - M_{n-1}\left(\frac{1}{k+\frac{1}{r}}\right) \qquad (5.7)$$

$$\text{Prob}(a_{n+1} = r) \cdot \text{Prob}(a_n = k) = [M_{n-1}(\tfrac{1}{k}) - M_{n-1}(\tfrac{1}{k+1})][M_n(\tfrac{1}{r}) - M_n(\tfrac{1}{r+1})] \qquad (5.8)$$

Now from Kusmin's theorem

$M_n(x) \rightarrow \frac{\log(1+x)}{\log 2}$ uniformly for $0 \leq x \leq 1$ as $n \rightarrow \infty$

Therefore for large values of n

$$\text{Prob}[a_{n+1} = r;\ a_n = k] = \frac{\log(1+\frac{1}{k+\frac{1}{r+1}}) - \log(1+\frac{1}{k+\frac{1}{r}})}{\log 2} \qquad (5.9)$$

and

$$\text{Prob}(a_{n+1} = r) \cdot \text{Prob}(a_n = k) = \frac{\log[1+\frac{1}{k(k+2)}][\log(1+\frac{1}{r(r+2)})]}{(\log 2)^2} \qquad (5.10)$$

For $k = 1 \quad r = 1$

$(5.9) = .15185$

and $(5.10) = .17189$

Therefore even for large values of n, a_n and a_{n+1} are not independent. By the same argument as in (5.7) it can be obtained that

$$\mathrm{Prob}(a_{n+1} \geq r \quad a_n \geq k) = \frac{\log(1+\frac{1}{k\,r})}{\log 2}$$

and
$$\mathrm{Prob}(a_{n+1} \geq r) \cdot \mathrm{Prob}(a_n \geq k) = \frac{\log(1+\frac{1}{k}) \cdot \log(1+\frac{1}{r})}{(\log 2)^2}$$

which are in general not equal.

By calculations to those given above one can show more generally that a_n and a_m are not independent for any values of n and m. Now to have an idea of the dependence between a_n and a_m, a measure of dependence can be defined in the following manner.

Dependence between the relations $(a_n = k)$ and $(a_m = r)$

$$D_{n,m}(k, r) = \left| \frac{\mathrm{Prob}(a_n = k,\ a_m = r)}{\mathrm{Prob}(a_n = k) \cdot \mathrm{Prob}(a_m = r)} - 1 \right|$$

If a_n and a_m are independent $D_{n,m}(k, r) = 0$ for every pair k and r, and conversely also. Therefore the magnitude $D_{n,m}(k, r)$ gives a measure for dependence. $D_{n,m}(k, r)$ for $(n=1, m=2)$, $(n=1, m=3)$ and $D_{n,n+1}(k, r)$ for large values of n are calculated and tabulated (tables 3, 4, 5 at the end of this chapter) using the formulae obtained.

<u>Expression for the joint Prob($a_{n+p} = r$, $a_n = k$)</u>

Joint Prob($a_1 = k$, $a_2 = k_2 \cdots a_n = k_n$)

$$= \left| K_o\left(\cfrac{1}{k_1 + \cfrac{1}{k_2 + \cdots + \cfrac{1}{k_n}}}\right) - K_o\left(\cfrac{1}{k_1 + \cfrac{1}{k_2 + \cdots + \cfrac{1}{k_n + 1}}}\right) \right| \qquad (5.11)$$

$$= \left| \frac{p_n}{q_n} - \frac{p_n + p_{n-1}}{q_n + q_{n-1}} \right| \quad \text{where}$$

$\frac{p_n}{q_n}$ is the n^{th} convergent of $\cfrac{1}{k_1 + \cfrac{1}{k_2 + \cdots + \cfrac{1}{k_n}}}$

$\therefore$ joint Prob($a_1 = k_1$, $a_2 = k_2$, $\cdots$ $a_n = k_n$)

$$= \frac{1}{q_n(q_n + q_{n-1})} \qquad (5.12)$$

conditional Prob($a_{n+1} = k \mid a_1 = k_1$, $a_2 = k_2 \cdots a_n = k_n$)

$$= \frac{\text{Prob}(a_1 = k_1,\ a_2 = k_2,\ \cdots a_n = k_n,\ a_{n+1} = k)}{\text{Prob}(a_1 = k_1,\ a_2 = k_2,\ \cdots a_n = k_n)}$$

$$= \frac{q_n(q_n + q_{n-1})}{(k\,q_n + q_{n-1})(k\,q_n + q_n + q_{n-1})} \qquad (5.13)$$

and $$\text{Prob}(a_{n+1} = k) = \overset{(n)}{\Sigma} \frac{1}{(q_{n-1}+k\,q_n)(q_{n-1}+k\,q_n+q_n)} \qquad (5.34)$$

Before obtaining the general expression for $\text{Prob}(a_{n+p} = r,\ a_n = k)$ joint $\text{Prob}(a_{n+2} = r,\ a_n = k)$ is derived

$$\text{Prob}(a_{n+2} = r,\ a_{n+1} = k_{n+1},\ a_n = k_n \ldots a_1 = k_1)$$

$$= \frac{1}{(q_n + r\,q_{n+1})(q_n + r\,q_{n+1} + q_{n+1})} \quad \text{from } (5.12)$$

substituting $q_{n+1} = k_{n+1}\,q_n + q_{n-1}$, one obtains

$$= \frac{1}{[(r\,k_{n+1} + 1)q_n + r\,q_{n-1}][(r\,k_{n+1} + k_{n+1} + 1)q_n + (r+1)q_{n-1}]}$$

$$= \frac{1}{q_n^2\,(r)(r+1)} \quad \frac{1}{[k_{n+1} + \frac{1}{r} + \frac{q_{n-1}}{q_n}][k_{n+1} + \frac{1}{r+1} + \frac{q_{n-1}}{q_n}]}$$

$$= \frac{1}{q_n^2}\left[\frac{1}{k_{n+1} + \frac{1}{r+1} + \frac{q_{n-1}}{q_n}} - \frac{1}{k_{n+1} + \frac{1}{r} + \frac{q_{n-1}}{q_n}}\right]$$

Now let k_{n+1} assume all values from 1 to ∞. Then

$$\text{Prob}(a_{n+2} = r,\ a_1 = k_1,\ a_2 = k_2 \ldots a_n = k_n)$$

$$= \sum_{k_{n+1}=1}^{\infty} \frac{1}{q_n^2}\left[\frac{1}{k_{n+1} + \frac{1}{r+1} + \frac{q_{n-1}}{q_n}} - \frac{1}{k_{n+1} + \frac{1}{r} + \frac{q_{n-1}}{q_n}}\right]$$

$$= \frac{1}{q_n^2}\left\{\psi\left(\frac{1}{r}+\frac{q_{n-1}}{q_n}\right) - \psi\left(\frac{1}{r+1}+\frac{q_{n-1}}{q_n}\right)\right\}$$

$$= \frac{1}{(k_n q_{n-1}+q_{n-2})^2}\left[\psi\left(\frac{1}{r}+\frac{q_{n-1}}{k_n q_{n-1}+q_{n-2}}\right) - \psi\left(\frac{1}{r+1}+\frac{q_{n-1}}{k_n q_{n-1}+q_{n-2}}\right)\right]$$

Write $k_n = k$ and let $k_1 k_2 \ldots k_{n-1}$ take all the values from 1 to ∞

$$\text{Prob}(a_{n+2} = r,\ a_n = k) = \sum^{(n-1)} \frac{\psi\left(\frac{1}{r}+\frac{q_{n-1}}{k\, q_{n-1}+q_{n-2}}\right) - \psi\left(\frac{1}{r+1}+\frac{q_{n-1}}{k\, q_{n-1}+q_{n-2}}\right)}{(k\, q_{n-1}+q_{n-2})^2} \tag{5.15}$$

From this formula it follows that

$$\text{Prob}(a_{n+2} = r,\ a_n = k) < \frac{M'_{n-1}\left(\frac{1}{k}\right)}{k^2}\ \text{Prob}(a_2 = r) \tag{5.16}$$

From the formula (5.16) it appears that the joint $\text{Prob}(a_{n+p} = r,\ a_n = k)$ is more directly related in a way to the product $\text{Prob}(a_n = k) \cdot \text{Prob}(a_p = r)$ than to the product $\text{Prob}(a_n = k) \cdot \text{Prob}(a_{n+p} = r)$.

Joint $\text{Prob}(a_{n+p} = r,\ a_n = k)$

$a_n = k$ implies that z_{n-1} is in $\left(\frac{1}{k+1}, \frac{1}{k}\right)$ and the distribution function of z_{n-1} is $M_{n-1}(x)$.

Therefore from (5.11) it follows that

Joint $\text{Prob}(a_{n+p} = r,\ a_{n+p-1} = k_{p-1},\ \dots\ a_{n+1} = k_1,\ a_n = k)$

$$= \left| M_{n-1}\left(\cfrac{1}{k+\cfrac{1}{k_1+\cfrac{1}{k_2 \ddots + \cfrac{1}{k_{p-1}+\cfrac{1}{r}}}}}\right) - M_{n-1}\left(\cfrac{1}{k+\cfrac{1}{k_1+\cfrac{1}{k_2 \ddots + \cfrac{1}{k_{p-1}+\cfrac{1}{r+1}}}}}\right) \right|$$

(Instead of a with distribution function $M_o(x)$, here z_{n-1} with $M_{n-1}(x)$ is considered)

Let $\cfrac{1}{k_1+\cfrac{1}{k_2+\dots+\cfrac{1}{k_{p-1}}}} = \dfrac{p'_{p-1}}{q'_{p-1}}$ and $\dfrac{p'_r}{q'_r}$ be the r^{th} convergent of this continued fraction. With this notation,

$$\text{Prob}(a_{n+p} = r,\ a_{n+p-1} = k_{p-1},\ \dots\ a_{n+1} = k_1,\ a_n = k)$$

$$= \left| M_{n-1}\left(\frac{1}{k+\dfrac{r\,p'_{p-1}+p'_{p-2}}{r\,q'_{p-1}+q'_{p-2}}}\right) - M_{n-1}\left(\frac{1}{k+\dfrac{(r+1)p'_{p-1}+p'_{p-2}}{(r+1)q'_{p-1}+q'_{p-2}}}\right) \right|$$

$$= \sum^{(n-1)} \frac{1}{q_{n-1}^2}\left[\frac{1}{k+\dfrac{q_{n-2}}{q_{n-1}}+\dfrac{r\,p'_{p-1}+p'_{p-2}}{r\,q'_{p-1}+q'_{p-2}}} - \frac{1}{k+\dfrac{q_{n-2}}{q_{n-1}}+\dfrac{(r+1)p'_{p-1}+p'_{p-2}}{(r+1)q'_{p-1}+q'_{p-2}}}\right]$$

Since $\psi_{n-1}(x) = \sum^{(n-1)} \frac{x}{q_{n-1}(q_{n-1} + x\, q_{n-2})}$

$$= \sum^{(n-1)} \frac{1}{q_{n-1}^2} \frac{\frac{1}{(r\, q'_{p-1} + q'_{p-2})(r\, q'_{p-1} + q'_{p-1} + q'_{p-2})}}{(k + \frac{q_{n-2}}{q_{n-1}} + \frac{r\, p'_{p-1} + p'_{p-2}}{r\, q'_{p-1} + q'_{p-2}})(k + \frac{q_{n-2}}{q_{n-1}} + \frac{(r+1)p'_{p-1} + p'_{p-2}}{(r+1)q'_{p-1} + q'_{p-2}})} \tag{5.17}$$

$$= \sum^{(n-1)} \frac{1}{(q_{n-1}k + q_{n-2})^2} \frac{\frac{1}{(r\, q'_{p-1} + q'_{p-2})(r\, q'_{p-1} + q'_{p-1} + q'_{p-2})}}{(1 + \frac{r\, p'_{p-1} + p'_{p-2}}{r\, q'_{p-1} + q'_{p-2}} \frac{1}{x_n})(1 + \frac{(r+1)p'_{p-1} + p'_{p-2}}{(r+1)q'_{p-1} + q'_{p-2}} \frac{1}{x_n})}$$

Where $x_n = k + \frac{q_{n-2}}{q_{n-1}}$

Now $\frac{q_{n-2}}{q_{n-1}}$ lies between 0 and 1

Therefore $k \leq x_n \leq k + 1$

$$\frac{1}{k+1} \leq \frac{1}{x_n} \leq \frac{1}{k}$$

and

$$0 < \frac{r\, p'_{p-1} + p'_{p-2}}{r\, q'_{p-1} + q'_{p-2}} < 1 \;;\; 0 < \frac{(r+1)p'_{p-1} + p'_{p-2}}{(r+1)q'_{p-1} + q'_{p-2}} < 1$$

Using these inequalities and the formula

$$\psi'_{n-1}\left(\frac{1}{k}\right) = \sum^{(n-1)} \frac{1}{(q_{n-1} + q_{n-2}\frac{1}{k})^2} = \sum^{(n-1)} \frac{k^2}{(k\, q_{n-1} + q_{n-2})^2}$$

one obtains that

Joint $\text{Prob}(a_{n+p} = r,\ a_{n+p-1} = k_{p-1},\ \ldots\ a_n = k)$

$$\geq \sum^{(n-1)} \frac{k^2}{(k\, q_{n-1} + q_{n-2})^2} \; \frac{\dfrac{1}{(r\, q'_{p-1} + q'_{p-2})(r\, q'_{p-1} + q'_{p-1} + q'_{p-2})}}{(k+1)^2}$$

$$= \frac{M'_{n-1}\left(\frac{1}{k}\right)}{(k+1)^2} \quad \frac{1}{(r\, q'_{p-1} + q'_{p-2})(r\, q'_{p-1} + q'_{p-1} + q'_{p-2})}$$

Now allowing $k_1, k_2 \ldots k_{p-1}$ assume all the values from 1 to ∞

$$\text{Prob}(a_{n+p} = r,\ a_n = k) \;\geq\; \frac{M'_{n-1}\left(\frac{1}{k}\right)}{(k+1)^2}\ \text{Prob}(a_p = r) \qquad (5.18)$$

Similarly using the other inequality

$$\text{Prob}(a_{n+p} = r\ \ a_n = k) \;<\; \frac{M'_{n-1}\left(\frac{1}{k}\right)}{k^2}\ \text{Prob}(a_p = r) \qquad (5.19)$$

(5.18) and (5.19) gave the inequalities in terms of $M'_{n-1}\left(\frac{1}{k}\right)$. Now to get the upper and lower bounds for (5.17) in terms of $\text{Prob}(a_n = k)$, one can write (5.17) in the following way

$$\text{Prob}(a_{n+p} = r,\ a_{n+p-1} = k_{p-1}\ \ldots\ a_{n+1} = k_1,\ a_n = k)$$

$$= \sum^{(n-1)} \frac{\dfrac{1}{(r\, q'_{p-1})(q'_{p-2})(r\, q'_{p-1} + q'_{p-1} + q'_{p-2})}}{(k\, q_{n-1}+q_{n-2})(k\, q_{n-1}+q_{n-1}+q_{n-2})\left(1+\dfrac{1}{x_n}\,\dfrac{r\, p'_{p-1}+p'_{p-2}}{r\, q'_{p-1}+q'_{p-2}}\right)\left(1+\dfrac{1}{x_{n+1}}\,\dfrac{rp'_{p-1}+p'_{p-1}+p'_{p-2}}{rq'_{p-1}+q'_{p-1}+q'_{p-2}}\,\dfrac{1}{x_n+1}\right)}$$

$$\geq \sum^{(n-1)} \frac{\dfrac{1}{(r\,q'_{p-1} + q'_{p-2})(r\,q'_{p-1} + q'_{p-1} + q'_{p-2})}}{(k\,q_{n-1} + q_{n-2})(k\,q_{n-1} + q_{n-2} + q_{n-1})(1 + \frac{1}{k})\ (1)}$$

$$= \text{Prob}(a_n = k)\ \frac{1}{(r\,q'_{p-1} + q'_{p-2})(r\,q'_{p-1} + q'_{p-1} + q'_{p-2})} \cdot (1 - \frac{1}{k+1})$$

Again summing over all $k_1\ k_2 \ldots k_{p-1}$, one obtains that

$$\text{Prob}(a_{n+p} = r,\ a_n = k) \geq \text{Prob}(a_n = k) \cdot \text{Prob}(a_p = r)(1 - \frac{1}{k+1}) \qquad (5.20)$$

Similarly using the inequalities

$$1 + \frac{1}{x_n}\ \frac{r\,p'_{p-1} + p'_{p-2}}{r\,q'_{p-1} + q'_{p-2}} > 1$$

and

$$1 + \frac{1}{x_n}\ \frac{r\,p'_{p-1} + p'_{p-1} + p'_{p-2}}{r\,q'_{p-1} + q'_{p-1} + q'_{p-2}} = \frac{1}{x_n + 1} > 1 - \frac{1}{k+1}$$

$$\text{Prob}(a_{n+p} = r,\ a_n = k) \leq \text{Prob}(a_n = k) \cdot \text{Prob}(a_p = r)(1 + \frac{1}{k}) \qquad (5.21)$$

From (5.20) and (5.21)

$$\left| \frac{\text{Prob}(a_{n+p} = r,\ a_n = k)}{\text{Prob}(a_p = r) \cdot \text{Prob}(a_n = k)} - 1 \right| < \frac{1}{k} \qquad (5.22)$$

i.e. For large values of k

$$\text{Prob}(a_p = r) \cdot \text{Prob}(a_n = k) = \text{Prob}(a_{n+p} = r,\ a_n = k)$$

Since $\lim\limits_{p \to \infty} \text{Prob}(a_p = r) = \lim\limits_{p \to \infty} \text{Prob}(a_{n+p} = r) = \dfrac{\log(1 + \frac{1}{r(r+2)})}{\log 2}$

one can say that for large values of k and for large p,

$$\text{Prob}(a_{n+p} = r,\ a_n = k) = \text{Prob}(a_{n+p} = r) \cdot \text{Prob}(a_n = k)$$

In this connection Khintchine [10] by generalising Kuzmin's result, established the result that

$$\left| \frac{\text{Prob}(a_{n+p} = r;\ a_n = k)}{\text{Prob}(a_{n+p} = r)\ \text{Prob}(a_n = k)} - 1 \right| < 6A\, e^{-\lambda\sqrt{p-1}} .$$

This inequality shows that as $p \to \infty$, a_{n+p} and a_n are independent or two sufficiently far apart indices of a continued fraction are independent.

TABLE 3

$$D_{12}(k, r) = \left| \frac{\text{Prob}(a_1 = k \; ; \; a_2 = r)}{\text{Prob}(a_1 = k)\ \text{Prob}(a_2 = r)} - 1 \right|$$

r \ k	1	2	3	4	5	6	7	8	9	10
1	.1365	.036	.110	.152	.175	.196	.209	.219	.227	.234
2	.0140	.014	.014	.0115	.009	.006	.004	.002	.001	.0001
3	.042	.008	.0385	.0575	.0699	.0792	.0857	.0910	.0954	.0986
4	.092	.006	.0542	.0819	.0993	.1116	.1206	.1276	.1332	.1176
5	.108	.024	.0821	.1141	.1344	.1487	.1588	.1667	.1738	.1778
6	.114	.0385	.1027	.1380	.1368	.1751	.1861	.1944	.2018	.2057
7	.111	.0589	.1274	.1640	.1663	.2027	.2141	.2227	.2297	.2350
8	.169	.0223	.0978	.1380	.1619	.1800	.1932	.2014	.2106	.2146
9	.136	.0603	.1356	.1763	.2009	.2177	.2298	.2390	.2461	.2518
10	.165	.0445	.1244	.1661	.1921	.2095	.2222	.2315	.2392	.2450

TABLE 4

$$D_{n,n+1}(k,\ r) = \left| \frac{\log\left(1 + \frac{1}{(r\ k+k+1)\ (r\ k+r+1)}\right) \cdot \log 2}{\log\left(1 + \frac{1}{k(k+2)}\right) \log\left(1 + \frac{1}{r(r+2)}\right)} - 1 \right| \quad \text{for large } n$$

r \ k	1	2	3	4	5	6	7	8	9	10
1	.1175	.002	.0513	.083	.104	.119	.131	.138	.144	.1515
2	.002	.008	.007	.002	.006	.005	.002	.002	.0013	.0056
3	.0513	.007	.019	.037	.060	.0535	.058	.065	.068	.078
4	.083	.002	.037	.0615	.0698	.0905	.1063	.1197	.1043	.1064
5	.104	.006	.060	.0698	.0742	.0935	.1004	.0884	.0967	.1018
6	.119	.005	.0535	.0905	.0935	.1376	.1649	.1290	.1526	.1625
7	.131	.002	.058	.1063	.1004	.1649	.1636	.1035	.1925	.1466
8	.138	.002	.065	.1197	.0884	.1290	.1035	.1742	.1690	.1507
9	.144	.0013	.068	.1043	.0967	.1526	.1925	.1690	.1290	.1394
10	.1515	.0056	.078	.1064	.1018	.1625	.1466	.1507	.1394	.1916

TABLE 5

$$D_{13}(k, r) = \left| \frac{\text{Prob}(a_1 = k \; ; \; a_3 = r)}{\text{Prob}(a_1 = k) \cdot \text{Prob}(a_3 = r)} - 1 \right|$$

r \ k	1	2	3	4	5	6	7	8	9	10
1	.0372	.0076	.0288	.0414	.0492	.0496	.0334	.0631	.0653	.0674
2	.0059	.0021	.0017	.0027	.0038	.0039	.0056	.0058	.0057	.0057
3	.0130	.0054	.0142	.0200	.0217	.0271	.0255	.0248	.0262	.0267
4	.0276	.0086	.0230	.0300	.0366	.0314	.0463	.0484	.0494	.0485
5	.0550	.0012	.0233	.0438	.0362	.0656	.0630	.0494	.0494	.0533
6	.0156	.0098	.0380	.0330	.0690	.0828	.0464	.0706	.0572	.0866
7	.0225	.0058	.0280	.0430	.0553	.0388	.0640	.0836	.0802	.0852
8	.0575	.0103	.0370	.0590	.0636	.0582	.0861	.0964	.0805	.0775
9	.0496	.0106	.0485	.0630	.0647	.0756	.0796	.0865	.0827	.0867
10	.0940	.0086	.0470	.0760	.0666	.0973	.0955	.1134	.1044	.0962

CHAPTER VI

STATISTICAL ANALYSIS OF THE DATA

As the expressions for $M_n(x)$ are theoretically very complicated and direct calculations are almost impossible, a statistical analysis of the distribution functions was also made. A sample of 1000 random numbers [21] was taken and for each number the index numbers $a_1, a_2, \dots a_{15}$ were calculated. From the data the distribution functions $F_n^*(k)$ and the frequency functions $f_n^*(k)$ were tabulated (Tables 6 and 7) for $n = 1$ to 15. Throughout this chapter $F_n^*(k)$, $f_n^*(k)$ denote respectively the distribution and frequency functions of the sample.

The analysis of the sample suggests that the distribution functions $F_n^*(k)$ behave like the asymptotic distribution function $F_\infty(k) = \dfrac{\log(1+\frac{k}{k+2})}{\log 2}$ even for small values of $n (n \geq 3)$. The theoretical frequency functions $f_1(k)$, $f_2(k)$, $f_3(k)$ and the inequality of Kuzmin

$$\left| f_n(k) - \frac{\log(1 + \frac{1}{k(k+2)})}{\log 2} \right| < \frac{A}{k(k+1)} \, e^{-\lambda\sqrt{n-1}}$$

suggest that the convergence of $f_n(k)$ to $\dfrac{\log(1 + \frac{1}{k(k+2)})}{\log 2}$ is rapid not only as $n \to \infty$ but also as $k \to \infty$. However in the frequency functions $f_n^*(k)$ of this sample this property was not very much pronounced. This may be due to the small size of the sample.

Test of goodness of fit for $f_n^*(k)$: To test the approximation of $f_n^*(k)$ to $f_\infty(k) = \dfrac{\log[1 + \frac{1}{k(k+2)}]}{\log 2}$, the χ^2 test is used. The procedure of this test is given below [22].

Let hypothesis H be that $f_n^*(k)$ can be approximated by $\dfrac{\log[1 + \frac{1}{k(k+2)}]}{\log 2}$. Let $\chi_o^2 = \sum_{i=1}^{r} \dfrac{(\gamma_i - N p_i)^2}{Np_i}$ where γ_i = the observed frequency of the i^{th} class

N = number of observations in the sample.

Np_i = expected frequency (under hypothesis H).

r = number of classes into which the sample is divided.

χ_o^2 is a measure of deviation between the two distributions $\{f_n^*(k), f_\infty(k)\}$. If χ_o^2 is equal to 0, the two distributions are identical. The larger the χ_o^2, the more the deviation between the two distributions.

And further for large values of $N (N > 50)$, χ_o^2 is distributed as a χ^2 with $(r-1)$ degrees of freedom. Let χ_p^2 be the value of χ^2 such that

$$\text{Prob}(\chi^2 > \chi_p^2) = \frac{p}{100} = p\,\%$$

(generally $p = 5$ and 1 are taken).

Now if the hypothesis H is true, then it is practically impossible, in one single sample, to encounter a value of χ_o^2 exceeding χ_p^2.

Therefore if one finds a value $\chi_o^2 > \chi_p^2$ in the sample, one accordingly says that the sample shows a significant deviation from the hypothesis H and the hypothesis H is rejected at least until further data are available. The probability that H is falsely rejected is p %. If, on the other hand, $\chi^2 \leq \chi_p^2$ this will be regarded as consistent with hypothesis H. However one isolated result of this kind cannot be considered as sufficient evidence of the truth of the hypothesis H. The test must be repeated to make a positive statement.

Since the sample consists of 1000 observations one can use the above test. For n = 2 to 15 the values of χ_o^2 are calculated and tabulated (Table 8) and they are compared with the values χ_p^2 for p = 5 and p = 1 from the χ^2 tables. Since all the χ_o^2's $< \chi_p^2$ it follows that all the frequency functions $f_n^*(k)$ for n = 2 to 15 can be approximated by $f_\infty(k)$.

TABLE 6

FREQUENCY FUNCTIONS OF a_n FOR $n = 1$ TO 15

k	$f^*_1(k)$	$f^*_2(k)$	$f^*_3(k)$	$f^*_4(k)$	$f^*_5(k)$	$f^*_6(k)$	$f^*_7(k)$	$f^*_8(k)$	$f^*_9(k)$	$f^*_{10}(k)$	$f^*_{11}(k)$	$f^*_{12}(k)$	$f^*_{13}(k)$	$f^*_{14}(k)$	$f^*_{15}(k)$
1	500	388	409	414	414	414	417	424	406	382	421	408	397	423	408
2	170	188	179	182	160	178	158	164	154	191	169	178	168	167	144
3	66	94	90	86	92	106	86	97	101	110	92	109	103	91	100
4	46	62	48	62	52	50	57	50	62	60	55	60	57	61	73
5	33	42	47	41	47	41	35	40	35	38	40	33	40	47	49
6	22	39	34	32	19	33	39	37	32	27	35	28	36	23	24
7	22	18	23	29	20	21	22	24	20	16	13	29	18	23	26
8	20	22	19	13	24	14	19	19	16	17	15	16	15	18	19
9	12	13	17	14	18	15	15	14	18	16	18	18	20	17	13
10 ≤ k 45	41	34	44	39	57	46	43	58	46	37	52	38	46	47	57
15 ≤ k < 50	48	69	54	70	62	56	70	48	75	67	65	54	61	53	58
50 ≤	20	31	36	18	27	26	39	25	35	39	25	29	39	30	29
Greatest Value	258	1984	18796	763	5020	25000	2777	9803	11467	1343	726	2223	4997	20408	333

TABLE 7

DISTRIBUTION FUNCTIONS OF a_n FOR $n = 1$ TO 15

k	F_1^*	F_2^*	F_3^*	F_4^*	F_5^*	F_6^*	F_7^*	F_8^*	F_9^*	F_{10}^*	F_{11}^*	F_{12}^*	F_{13}^*	F_{14}^*	F_{15}^*	$F_\infty(k)$
1	500	388	409	414	414	414	417	424	406	382	421	408	397	423	408	415
2	670	576	588	596	582	592	575	588	559	573	590	586	565	590	552	585
3	736	670	678	682	674	698	661	685	659	683	682	695	668	681	652	678
4	782	732	726	744	726	748	718	733	721	743	737	755	725	742	725	737
5	815	774	773	785	773	789	751	773	756	781	777	788	765	789	774	778
6	837	813	807	817	792	822	790	810	788	808	812	816	801	812	798	807
7	859	831	830	846	812	843	812	834	808	824	825	845	815	835	824	830
8	879	853	849	859	836	857	831	853	824	841	840	861	830	853	843	848
9	891	866	866	873	854	872	846	867	842	857	858	879	850	870	856	862
$10 \le k < 15$	932	900	910	912	911	918	889	925	908	894	910	917	896	917	913	907
$15 \le k < 50$	980	969	964	982	973	974	959	973	963	961	975	971	957	970	971	971
$50 \le k < \infty$	1000	1000	1000	1000	1000	1000	999	999	999	979	998	998	998	998	998	1000

TABLE 8

VALUES OF x_o^2 FOR EACH n = 1 TO 15

n	x_o^2
2	12.369
3	8.025
4	10.164
5	12.226
6	6.585
7	9.910
8	13.355
9	8.167
10	14.680
11	8.839
12	9.628
13	10.595
14	4.833
15	15.921

$x^2_{.05}$ = 19.675 for 11 degrees of freedom

$x^2_{.01}$ = 24.725

APPENDIX A

CALCULATION OF $M_2(x)$

Euler's summation formula: - Let $f(x)$ be a function having derivative of order $(2k+1)$ which is continuous. Then

$$f_1 + f_2 + \dots + f_n = \int_1^n f(x)dx + \frac{1}{2}(f_n + f_1)$$

$$+ \frac{B_2}{2!}(f'_n - f'_1) + \frac{B_4}{4!}(f'''_n - f'''_1) + \dots$$

$$\dots + \frac{B_{2k}}{2k!}(f_n^{2k-1} - f_1^{2k-1}) + R_k$$

where $$R_k = \int_1^n P_{2k+1}(x)\, f^{2k+1}(x)dx \; [23]$$

$B_2, B_4 \dots B_{2k} \dots$ are the Bernoulli numbers and $P_k(x)$ is the k^{th} Bernoulli polynomial. [$P_k(x) = \frac{1}{k!}(B+x)^k$ in symbolic notation $B^k = B_k$]

f_n denotes the value of f at n.

Further if

1) $\sum_1^\infty f_n$ converges

2) $\int_1^\infty f(x)dx$ converges

3) $f_n^k \to 0$ as $n \to \infty$ for $r = 1, 2 \dots (2k+1)$

the Euler's summation formula can be written as

$$\sum_{n=1}^{\infty} f_n = \int_1^{\infty} f(x)dx + \frac{1}{2} f_1 - \frac{B_2}{2!} f_1' - \frac{B_4}{4!} f_1''' \cdots - \frac{B_{2k}}{2k!} f_1^{2k-1} + R_k \qquad (A.1.1)$$

where $\quad R_k = \int_1^{\infty} p_{2k+1}(x)\, f^{2k+1}(x)dx$

since $\quad |p_{2k}(x)| < \frac{B_{2k}}{2k!}$

and $\quad R_k = -\int_1^{\infty} p_{2k}(x)\, f^{2k}(x)dx$

$$|R_k| < \frac{B_{2k}}{2k!} f_1^{2k-1}$$

In particular if f^{2k} and f^{2k-2} have the same sign and keep their sign, then the error

$$|R_k| < \left|\frac{B_{2k+2}}{2k+2!} f_1^{2k+1}\right| \qquad [24]$$

i.e. The error is numerically less than the first neglected term and has the same sign.

<u>The calculation of $F_2(y)$ with Euler's formula</u>

$$F_2(y) = \sum_{n=1}^{\infty} \frac{1}{n^2} \left\{\Psi(y+\tfrac{1}{n}) - \Psi(\tfrac{1}{n})\right\}$$

$$= \sum_{n=1}^{\infty} f_n$$

$$\sum_{n=1}^{\infty} f_n = \int_1^{\infty} \frac{\Psi(y+\frac{1}{x}) - \Psi(\frac{1}{x})}{x^2} dx + \frac{1}{2} [\ \Psi(y+1) - \Psi(1)]$$

$$- \frac{B_2}{2!}[f_1'] - \frac{B_4}{4!}[f_1'''] - \cdots - \frac{B_{2k}}{2k!}[f_1^{2k-1}] + R_k \ .$$

Since f_n satisfies all the properties for the formula (A.1.1), we have

$$H_2(y) = \log(1+y) + \frac{1}{2}[\psi(y+1) - \psi(1)] - \frac{B_2}{2!}f_1' \cdots + R_k \ .$$

$$\left[\text{since } \int_1^\infty \frac{\psi(y+\frac{1}{x}) - \psi(\frac{1}{x})}{x^2}\,dx = \log(1+y)\right]$$

$$f(n) = \frac{1}{n^2}[\psi(y+\tfrac{1}{n}) - \psi(\tfrac{1}{n})]$$

$$f'(n) = \frac{-1}{n^4}[\psi(y+\tfrac{1}{n}) - \psi(\tfrac{1}{n})] - \frac{2}{n^3}[\psi'(y+\tfrac{1}{n}) - \psi'(\tfrac{1}{n})]$$

$$f''(n) = \frac{1}{n^6}[\psi''(y+\tfrac{1}{n}) - \psi'(\tfrac{1}{n})] + \frac{6}{n^5}[\psi'(y+\tfrac{1}{n}) - \psi'(\tfrac{1}{n})] + \frac{6}{n^4}[\psi(y+\tfrac{1}{n}) - \psi(\tfrac{1}{n})]$$

$$f'''(n) = -[\frac{1}{n^8}\overline{\psi^3(y+\tfrac{1}{n}) - \psi^3(\tfrac{1}{n})} + \frac{12}{n^7}\overline{\psi^2(y+\tfrac{1}{n}) - \psi^2(\tfrac{1}{n})} + \frac{36}{n^6}\overline{\psi^1(y+\tfrac{1}{n}) - \psi^1(\tfrac{1}{n})}$$

$$+ \frac{26}{n^5}\overline{\psi(y+\tfrac{1}{n}) - \psi(\tfrac{1}{n})}]$$

$$f^4(n) = +[\ \cdots\cdots\cdots\cdots\cdots\ ]$$

$$f^5(n) = -[\ \cdots\cdots\cdots\cdots\cdots\ ]$$

Substituting the values f_1^1, f_1^3, f_1^5 in the formula

$$H_2(y) = \log(1+y) + \frac{1}{2}[\psi(y+1) - \psi(1)] + \frac{1}{12}[2\,\overline{\psi(y+1) - \psi(1)} + \overline{\psi'(y+1) - \psi'(1)}]$$

$$- \frac{1}{720}[\psi^3(y+1) - \psi^3(1) + 12\,\overline{\psi^2(x+1) - \psi^2(1)} + 36\,\overline{\psi^1(x+1) - \psi^1(1)} +$$

$$+ 24\,\overline{\psi(x+1) - \psi(1)}]$$

$$+ \frac{1}{30240}\left[\psi^5(y+1)-\psi^5(1)+30\,\overline{\psi^4(y+1)-\psi^4(1)}+300\,\overline{\psi^3(x+1)-\psi^3(1)}+1200\,\overline{\psi^2(x+1)-\psi^2(1)}\right.$$

$$\left.+\ 1800\,\overline{\psi'(x+1)-\psi'(1)}+720\,\overline{\psi(x+1)-\psi(1)}\right]+R_3$$

$|R_3| < .0026$ and $-$ve. [since f^4 and f^6 have same sign].

neglecting R_3 .

$$F_2(y) = \log(1+y) + a_0\,\overline{\psi(y+1)-\psi(1)} + a_1\,\overline{\psi^1(y+1)-\psi^1(1)}$$

$$+ a_2\,\overline{\psi''(y+1)-\psi''(1)} + a_3\,\overline{\psi^3(y+1)-\psi^3(1)} + a_4\,\overline{\psi^4(y+1)-\psi^4(1)}$$

$$+ a_5\,\overline{\psi^5(y+1)-\psi^5(1)} .$$

$a_0 = .6571$

$a_1 = .09286$

$a_2 = .02301$

$a_3 = .00853$

$a_4 = .000992$

$a_5 = \frac{1}{30240} = .00003303$

$\psi(y)$, $\psi^1(y)$, $\psi^2(y)$, $\psi^3(y)$, $\psi^4(y)$, are tabulated from [25]. $\psi^5(y)$ is not tabulated: But a_5 is very small. The maximum value of the term

$$a_5\,\overline{\psi^5(y+1)-\psi^5(1)} \text{ is } |a_5(\psi^5(2)-\psi^5(1))| < .00062$$

Therefore the values of $F_2(y)$ for different values of $y (0 < y < 1)$ are calculated and tabulated (Table 9). The error is $< .001$ for

$0 < y \leq \frac{1}{2}$.

Therefore the values are correct to 3 decimal places.

TABLE 9

$H_2(y)$

y	$\log(1+y)$	$\psi(y+1)-\psi(1)$	$-(\psi'(y+1)-\psi'(1))$	$\psi''(y+1)-\psi''(1)$	$-(\psi^3(y+1)-\psi^3(1))$	$\psi^4(y+1)-\psi^4(1)$	$H_2(y)$
1	.69315	.50000	.25000	.25000	.375000	.75000	1.0014
1/2	.40547	.28038	.15458	.16791	.27003	.57251	.57640
1/3	.28768	.19499	.11183	.12574	.20855	.45773	.40617
1/4	.22314	.14975	.08761	.10037	.16940	.37520	.31400
1/5	.18232	.12151	.07200	.08349	.14246	.31883	.25596
1/6	.15413	.10216	.06111	.07144	.12285	.27691	.21591
1/7	.13356	.08821	.05310	.06245	.10799	.24473	.18706
1/8	.11777	.07760	.04692	.05543	.09627	.21908	.16475
1/9	.10535	.06925	.04204	.04985	.08687	.193326	.14733
1/10	.09531	.06255	.03808	.04529	.07914	.18118	.13324
1/15	.06456	.04206	.02588	.03105	.05472	.12627	.08992
1/50	.01980	.01202	.00799	.00970	.01732	.04045	.02750
0	0	0	0	0	0	0	0

APPENDIX B

SIMILARITIES BETWEEN SYSTEMATIC AND SIMPLE CONTINUED FRACTIONS

As one can observe, there is some similarity between a simple continued fraction expansion and a systematic fraction expansion of a real number. For simple continued fractions the operator $T(\alpha) = \frac{1}{\alpha} - [\frac{1}{\alpha}]$ is iterated where as in systematic fractions the corresponding operator is $T(\alpha) = g\alpha - [g\alpha]$ where the integer g is the root of the number system.

Let

$$\alpha = \frac{a_1}{g} + \frac{a_2}{g^2} + \dots + \frac{a_n}{g^n} + \dots$$

where g is a positive integer ≥ 2 and $0 \leq a_n < g$ for each n.

Denote α by α_0 and define

$$\alpha_n = \frac{a_{n+1}}{g} + \frac{a_{n+2}}{g^2} + \dots + \frac{a_{n+r}}{g^r} + \dots \quad .$$

Then one can easily observe that

$[\alpha g] = a_1$ and in general

$[\alpha_n g] = a_{n+1}$

$$\alpha_{n+1} = g\alpha_n - [g\alpha_n] \qquad (B.1)$$

The distribution function $D_n(x)$ is defined as

$D_n(x) = \text{Prob}(a_n < x)$.

<u>Relation between $D_n(x)$ and $D_{n+1}(x)$</u>

$D_{n+1}(x) = \text{Prob}(a_{n+1} < x)$

$= \text{Prob}(a_n g - [a_n g] < x)$ by 8.1

$= \text{Prob}(\sum_{k=0}^{g-1} k \le a_n g < k+x)$ since

$[a_n g] = k$ implies $k \le a_n g < k+1$ and k can assume values from 0 to $(g-1)$.

$$D_{n+1}(x) = \sum_{k=0}^{g-1} \text{Prob}\left(\frac{k}{g} \le a_n < \frac{k+x}{g}\right)$$

$$= \sum_{k=0}^{g-1} D_n\left(\frac{x+k}{g}\right) - D_n\left(\frac{k}{g}\right) \qquad (8.2)$$

So if $D_o(x) = x$, then $D_n(x) = x$ for all n . And furthermore

$$G_n(k) = \text{Prob}(a_n \le k) = D_{n-1}\left(\frac{k+1}{g}\right) = \frac{k+1}{g}$$

Therefore $g_n(k) = \text{Prob}(a_n = k) = \frac{k+1}{g} - \frac{k}{g} = \frac{1}{g}$ which is independent of both n and k .

<u>$D_n(x)$ in terms of $D_o(x)$</u>

From the definition for a_n , it follows that

$a_n = g^n \cdot a - [g^n \cdot a]$

$$g^n \cdot \alpha = a_1 g^{n-1} + a_2 g^{n-2} + \cdots + a_n + \frac{a_{n+1}}{g} + \frac{a_{n+2}}{g^2} + \cdots + \frac{a_{n+r}}{g^r} + \cdots$$

and therefore $[g^n \cdot \alpha] = a_1 g^{n-1} + a_2 g^{n-2} + \cdots + a_n$

From this relation, it follows that

$$D_n(x) = \text{Prob}(g^n\alpha - [g^n \cdot \alpha] < x)$$

$$= \text{Prob}\Big(\sum_{k=0}^{g^n-1} k \leq g^n\alpha < k+x\Big)$$

$$= \sum_{k=0}^{g^n-1} D_o\Big(\frac{k+x}{g^n}\Big) - D_o\Big(\frac{k}{g^n}\Big) \qquad (B.3)$$

From the equation (B.3) if one assumes that $D_o(X)$ has a continuous bounded derivative, then

$$D_n'(x) = \sum_{k=0}^{g^n-1} \frac{1}{g^n} D_o'\Big(\frac{k+x}{g^n}\Big) \qquad (B.4)$$

From (B.4) one can observe that all properties of $D_n'(x)$ can be easily deduced directly from $D_o'(x)$.

<u>Inverse transform of $D_n'(x)$</u>

Let $$D_n'(x) = \int_0^\infty e^{-tx} \beta_n(t)dt$$

From (B.4)

$$D_n'(x) = \int_0^\infty \sum_{k=0}^{g^n-1} \frac{1}{g^n} \beta_o(t) e^{-\frac{t(k+x)}{g^n}} dt$$

$$= \int_0^\infty \frac{1-e^{-t}}{1-e^{-\frac{t}{g^n}}} \beta_o(t) e^{-\frac{tx}{g^n}} \frac{dt}{g^n}$$

$$= \int_0^\infty \frac{1 - e^{-g^n \cdot t}}{1 - e^{-t}} \cdot \beta_o(g^n \cdot t) \cdot e^{-tx}\, dt \; .$$

Therefore if $D_o^!(x)$ has an inverse transform $\beta_o(t)$ then

$$\beta_n(t) = \frac{1 - e^{-g^n \cdot t}}{1 - e^{-t}} \; \beta_o(g^n \cdot t) \; .$$

REFERENCES

[1] E. Borel — Rend. Circ. Mat. palermo 27 (1909) 247-271.

[2] F. Bernstein — Math. Ann. 71 (1912) 417-439.

[3] A. Khintchine — Math. Z. 18 (1923) 289-306.

[4] A. Khintchine — Math. Ann. 92 (1924) 115-125.

[5] F.J. Dyson — J. London Math. Soc. 18 (1943) 40-43.

[6] R.O. Kuzmin — Dokl. Akad. Nauk SSSR (1928) 375-380.

[7] A. Khintchine — Kettenbruche, Teubner, Leipzig (1956) (Mathematisch naturwissenschaftlich Bibliothek 3).

[8] P. Levy — Bull. Soc. Math. France 57 (1929) 178-194.

[9] P. Levy — Compositio Math. 3 (1936) 286- 303.

[10] A. Khintchine — Compositio Math. 1 (1935) 361-382.

[11] A. Khintchine — Compositio Math. 3 (1936) 276-285.

[12] A. Groschew — Rec. Math. Moscow 4 (1935) 509-518.

[13] A. Denjoy — C.R. Acad. Sci. Paris 202 (1936) 371-374.

[14] A. Denjoy — C.R. Acad. Sci. Paris 202 (1936) 537-540.

[15] A. Denjoy — C.R. Acad. Sci. Paris 202 (1936) 1140-1142.

[16] P. Levy — C.R. Acad. Sci. Paris 202 (1936) 812-813.

[17] W. Doeblin — Compositio Math. 7 (1940) 353-371.

[18] Whittaker and Watson — Modern Analysis; Cambridge University Press, London (1952).

[19] T.J.I. Bromwich — An introduction to the theory of infinite series, Macmillan and Company, London (1949).

[20] M. Kac — Statistical Independence in Probability, analysis and number theory. The carus Mathematical Monographs No. 12, New York (1959).

[21] The Rand Corporation — A Million random digits (with 1000 normal deviates). The Free Press Publishers, Glenco, Illinois.

[22] H. Cramer — Mathematical Methods of Statistics, Princeton University Press, Princeton (1954).

[23] K. Knopp — Theory and application of infinite series, Blackie and Son Limited, London (1949).

[24] J.F. Steffensen — Interpolation, Chelsea Publishing Company, New York (1950).

[25] H.T. Davis — Tables of Higher Mathematical functions, Vol. II, The Principia Press, Bloomington, Indiana (1935).

[26] P. Levy — Rend. circ. Mat. Palermo (2) 1 (1950) 170-208.