AN ABSTRACT OF THE THESIS OF

James Teng Wong (Name) for the Doctor of Philosophy (Degree)

in Mathematics (Major) presented on May 8, 1969 (Date)

Title: On the Generalization of the Distribution of the Significant Digits Under Computation

Redacted for Privacy

Abstract approved: _____ (Signature)

In this paper we use the set of all positive integers as a sample space whose probability density function is unknown. Then a generalization of the probability distribution of the most significant digits of the set of all physical constants is obtained on the strength of (i) a very general assumption imposed on the density function of the sample space, and (ii) a generalized invariance principle. The assumption is quite weak in the sense that it merely states that the occurrence of an event containing infinitely many elementary events is not impossible. The invariance principle, as is shown, is equivalent to another principle and to two functional equations. A function is constructed and, on the basis of the two foregoing stipulations that characterize the generalization, it is shown that this function is a unique solution, within a multiplicative positive constant, to another functional equation. The function so

constructed serves as a stepping stone in reaching our goal.

Having the generalization at our disposal, we deduce from it some of the consequences that are of interest. As it turns out, the deduction gives, on one hand, a proof to two empirical formulas published previously and, on the other, a fairly good agreement with the probabilities of three continuous density functions established in the literatures concerning the distribution of the leading digits under algebraic computation. In concluding the paper, a justification is made as to why a special case of the consequences of our result coincides with the probability of one of the three continuous density functions, even though our function is discrete.

On the Generalization of the Distribution of the Significant Digits Under Computation

by

James Teng Wong

A THESIS
submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1969

APPROVED:

Redacted for Privacy

Professor of Mathematics

In Charge of Major

Redacted for Privacy

Acting Chairman of Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 8, 1969

Typed by Charlene Laski for James Teng Wong

ACKNOWLEDGEMENT

I am ever grateful and indebted to my Major Professor, William M. Stone for the suggestion of the problem and for the guidance given to me throughout the period of research and writing of this thesis.

TABLE OF CONTENTS

ON THE GENERALIZATION OF THE DISTRIBUTION OF THE SIGNIFICANT DIGITS UNDER COMPUTATION

I. Introduction

It has been observed that in a well-used table of logarithms the pages containing the lower first or most significant digits, say 1, 2 and 3, are invariably better used, more ragged, than those beginning with the higher digits, say 8 and 9. This phenomenon was observed by Benford [1938]. No one could be expected to be interested in the actual condition of such a table but we may recall that the table is a base upon which some of our scientific studies are built. Consequently, an explanation was then sought by that observer in an attempt to obtain some meaningful measure of explanation for this peculiarity. For if we stipulate that the higher degree of decrepitude indicates more frequent usage we may then come to the conclusion that the numbers having the lower significant digits occur more frequently than those having the higher ones. As a result of his observation Benford remarked

> "A compilation of some 20,000 first digits taken from widely divergent sources shows that there is a logarithmic distribution of the first digits when the numbers are composed of four or more digits. An analysis shows that the numbers taken from unrelated subjects, such as a group of newspaper items, show a much better agreement with a logarithmic distribution than do numbers from mathematical tabulation or other formal data."

In particular, it is found that 30.6 percent of the observed physical constants have 1 as their first digit. This is in agreement with the common (base 10) logarithm of 2, 0.301 $\cdots$. Furthermore, the frequency of occurrence of the first, most significant, digits can be closely approximated by the logarithmic density function defined by the formula

$$(1.1) \qquad f(a) = \log_{10} \frac{a+1}{a}, \quad a = 1, 2, \cdots, 9,$$

as observed by Benford. For example, it may be observed that from Table I of Benford's paper 8.0 percent of the collected physical constants have 5 as their first digit and 4.7 percent of the observed constants have 9 as their first digit. These figures agree with $\log_{10} \frac{6}{5} = 0.079\cdots$ and $\log_{10} \frac{10}{9} = 0.046\cdots$. On the base of these consistencies Benford conjectured that the probability density function of the most significant digits of the set of all the physical constants is given by the $f(a)$ defined above, and he deduced that the probability of a digit in the q^{th} position is given by

$$(1.2) \qquad f(a_q) = \frac{\log_{10} \frac{a_1 a_2 \cdots a_q + 1}{a_1 a_2 \cdots a_q}}{\log_{10} \frac{a_1 a_2 \cdots a_{q-1} + 1}{a_1 a_2 \cdots a_{q-1}}}$$

where $a_1 a_2 \cdots a_q$ is a q-digit positive integer written in customary meaning of position and order in our decimal

system, and

$$(1.3) \qquad a_1 \varepsilon \{1,2,.....,9\}, \quad a_i \varepsilon \{0,1,....,9\}, \quad i=2,3,...,q.$$

At this point another result may be cited to substantiate the idea that nature seems to be in favor of odd digits over the even ones. In fact it has been found, Brown [1951], that in the production of random digits, including zero, the occurrence of odd digits is more favorable than that of even digits.

After the publication of Benford's paper, various successful papers have been published in connection with the derivation of the distribution function of the first significant digits of the set of all physical constants. Goudsmit and Furry [1944] has shown that the probability density of the first significant digits is independent of the probability density of the set of all physical constants, and the density obtained is the same as that of Benford's first empirical formula cited above. Pinkham [1961], on the other hand, observed that the collection of all known physical constants changes daily. For example, the population of a large city has a daily variation. However, he assumed that the probability distribution of the first significant digits is invariant under any scale change; that is, if all the physical constants were multiplied by a fixed real, positive number the distribution of the significant digits "would be

the same as before". On the strength of this invariance principle and the continuity condition on the distribution function of the underlying space (the collection of all the physical constants) Pinkham has shown that the probability distribution of the most significant digits is given by

$$D(n) = \log_{10} n, \quad n=2,3,\ldots,10. \tag{1.4}$$

Here $D(n)$ gives the probability that the first digit is $n-1$ or less.

Finally, in 1966 Flehinger [1966] observed that "the smallest population which contains the set of significant figures of all physical constants, past, present and future, must be the set of positive integers". It is this space the present investigation is concerned with and we shall obtain a generalization of the first significant digit concepts and a verification of Benford's formulas cited above.

CHAPTER II

Preliminaries

We shall state some of the well-known notions from the mathematical literature upon which the present investigation is based.

A probability P is a normed measure over a measurable space $(\Omega,\mathfrak{F})$; that is, P is a real-valued function on the sigma field $\mathfrak{F}$, which assigns to every set $A \in \mathfrak{F}$ a real number $P(A)$ such that

(a) $P(A) \geq 0$ for all $A \in \mathfrak{F}$,

(b) $P(\Omega) = 1$, and

(c) if $\{A_n\}$ is any denumerable sequence of disjoint events of $\mathfrak{F}$, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \tag{2.1}$$

If $A \in \mathfrak{F}$ and $B \in \mathfrak{F}$ and if $P(B) > 0$, then the conditional probability of A given B, denoted by $P(A|B)$, is defined by $P(A|B) = P(AB)/P(B)$, where $AB \equiv A \cap B$.

A random variable X is a real-valued function on Ω and X is $\mathfrak{F}$-measurable.

For convenience, whenever the meaning is clear from the context we shall write

(2.2) $$[X \leq x] \equiv \{\omega \in \Omega | X(\omega) \leq x\},$$

where x is a real number.

If X is a random variable its distribution function F_X is defined to be

(2.3) $$F_X(x) = P[X \leq x], \quad \text{for all} \quad x \in (-\infty, +\infty).$$

X is a discrete random variable if it is a random variable such that it assumes only a finite or countable number of real values and, for $A \in \mathfrak{B}$,

(2.4) $$P(A) = \sum_{\omega \in A} f(X(\omega)).$$

Then f is called a probability density function or distribution of X.

The notation of Theorem A.1 refers to the first theorem in the appendix.

CHAPTER III

Generalization

In this chapter we shall obtain a generalization of the distribution of the first significant digits. To begin with we shall define a few ideas, some of which characterize the results and the others simplify the typographical work.

Formulation of the Problem. Let Z_p be the set of all positive integers, and let P be the probability measure over the measurable space $(Z_p, \mathcal{G})$. The sigma field $\mathcal{G}$ is taken to be the set of all subsets of Z_p. Also, let F and f be the probability distribution and probability density function respectively of the random variable ξ on Z_p, where ξ is defined by the formula $\xi(i) = i$, for all $i \in Z_p$. So, we have

$$\sum_{i \in Z_p} f(i) = 1. \tag{3.1}$$

$$F(x) = \sum_{\substack{i \in Z_p \\ i \leq x}} f(i), \quad 1 \leq x < \infty, \tag{3.2}$$

$$F(\infty) = 1, \qquad F(x) = 0, \quad -\infty < x < 1. \tag{3.3}$$

Now we may consider the following question: what is the

probability of observing that a positive integer has its leftmost first $d(n)$ digits less than or equal to n and greater than or equal to $10^{d(n)-1}$, where $n \in Z_p$ and $d(n)$, as we shall see, denotes the number of digits of n. In particular, what is the probability of observing that a positive integer has the first four significant digits as specified, such as 3019. If the probability density function f were known the question could be answered as

$$\sum_{\substack{i \in Z_p \\ i_1 i_2 i_3 i_4 = 3019}} f(i) \quad .$$

Here P, and hence f, is unknown as part of the conditions imposed on the problem. However, from Benford's empirical formula cited above, we would expect that the probability of the event could be approximated by the value of a logarithmic function of a positive real number.

We observe that ξ so defined is a real valued function on Z_p such that

$$(3.4) \qquad \{i \in Z_p \mid \xi(i) \leq x\} = \{i \in Z_p \mid \xi(i) \leq [x]\}$$

$$= \begin{cases} \Phi, & x < 1, \\ \bigcup_{i=1}^{[x]} \{i\}, & x \geq 1. \end{cases}$$

where [ ] is the greatest integer function and Φ denotes the null set. Consequently, $\{i \in Z_p | \xi(i) \leq x\} \in \mathcal{G}$, for all real numbers x. Hence, by definition, ξ is measurable with respect to $\mathcal{G}$.

Relation. We define a relation $\overset{n}{\sim}$ on a subset of Z_p as follows: Let $S \subseteq Z_p$. Then, for all $x,y \in S$, we say that x is related to y of order n, denoted by $x \overset{n}{\sim} y$, if and only if

$$x_1x_2\ldots.x_n = y_1y_2\ldots.y_n, \tag{3.5}$$

where $x_1x_2\ldots.x_n$ and $y_1y_2\ldots.y_n$ are the first n digits of the positive integers x and y, respectively. According to the relation so defined the integers 17 and 1701 are related of order 1 and 2, but not of order 3 or higher. A convention is in order before we proceed. For all $n \in Z_p$, d(n) is defined to be the number of digits of n; that is, d is a real-valued function on Z_p. For example, d(1234) = d(8017) = d(9999) = 4. Having these notions we may now define

$$S_{d(n)} = \bigcup_{i=10^{d(n)-1}}^{\infty} \{i\} \tag{3.6}$$

and

$$A_n = \{x \in S_{d(n)} | x \overset{d(n)}{\sim} n\}. \tag{3.7}$$

For example, if $n = 1,2,\ldots,9$, then $d(n) = 1$ and $S_{d(n)} = \bigcup_{i=1}^{\infty}\{i\} = Z_p$; the A_n are $A_1, A_2, \ldots, A_9$ and we observe that, in this case, $\bigcup_{i=1}^{9} A_i = Z_p$. For $n = 100,101, \ldots,999$, $S_{d(n)} = \bigcup_{i=100}^{\infty}\{i\} = Z_p - \bigcup_{i=1}^{99}\{i\}$. Here the A_n are $A_{100}, A_{101}, \ldots, A_{999}$, and, of course, the union of the A_n is S_3.

The relation $\overset{d(n)}{\sim}$ is an equivalence relation on $S_{d(n)}$. Consequently, it partitions $S_{d(n)}$ into equivalence classes $\{A_n\}$, $n = 10^{d(n)-1}, 10^{d(n)-1}+1, \ldots, 10^{d(n)}-1$, such that the A_n are a non-empty, disjoint subset of Z_p.

Assumptions. Owing to the fact that the density function f of ξ is unknown it is impossible to compute the probability of every event in $\mathcal{A}$, except for Z_p itself and the null set Φ. However, in the course of obtaining a generalization of the distribution of the first significant digits a fairly general assumption is needed on f. It is stated explicitly as follows:

(i) For every countably infinite subset A of Z_p, $\sum_{i\in A} f(i) > 0$.

The restriction imposed on f is quite weak, in that it merely states that the occurrence of the event A is not impossible. Next, a generalized invariance principle may be stated as follows:

(ii) The probability distribution of the leftmost $d(n)$ digits is invariant under the linear transformation $\eta = c(\xi + 1)$, c being a positive, real constant. Assumption (ii) reduces to that of Pinkham if $d(n)$ is taken to be 1. The principle so stated in (ii) is equivalent, as is shown later, to the following statement, which we shall call the generalized invariance principle of the density distribution.

(iii) The probability density distribution of the leftmost $d(n)$ digits is invariant under the linear transformation $\eta = c(\xi + 1)$.

This last principle is not indispensable. However, later development reveals that the latter is more simple in implementation. At the expenditure of these two conditions a generalization is obtained and the continuity condition on the probability distribution function F can be relaxed, in contrast to Pinkham's assumption.

Main Results. In order to attain our goal it is necessary to have the following facts at our disposal. The idea used in the proof of Lemma 3.6 is similar to that employed by Shannon [1948], and the proof of Lemma 3.1 is trivial and so will be omitted.

Lemma 3.1. Let $n \in Z_p$. Then

$$S_{d(n)} = \bigcup_{i=10^{d(n)-1}}^{10^{d(n)}-1} A_i \,, \tag{3.8}$$

and the A_i are non-empty disjoint subsets of Z_p.

Lemma 3.2. For every positive integer n,

$$A_n = \bigcup_{m=0}^{\infty} \bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\} \,. \tag{3.9}$$

Proof. Let $x \in \bigcup_{m=0}^{\infty} \bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\}$. Then for some m_0, $m_0 = 0,1,2,\cdots$, $x \in \bigcup_{i=n10^{m_0}}^{(n+1)10^{m_0}-1} \{i\}$ and x is a positive integer. This implies that $n10^{m_0} \leq x \leq (n+1)10^{m_0}-1$, but $(n+1)10^{m_0}-1 = n10^{m_0} + (1-\frac{1}{10^{m_0}})10^{m_0}$ and $0 \leq 1 - \frac{1}{10^{m_0}} < 1$, for every m_0. So $x = (n+\theta)10^{m_0}$, $0 \leq \theta \leq 1 - \frac{1}{10^{m_0}}$.

It follows that

$$x_1 x_2 \cdots x_{d(n)} = n, \tag{3.10}$$

and, for every $n \in Z_p$, $10^{d(n)-1} \leq n$, but $x \geq n$, so $x \in S_{d(n)}$. Consequently, $x \in A_n$ and

$$(3.11) \qquad \bigcup_{m=0}^{\infty} \bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\} \subseteq A_n.$$

On the other hand, suppose that $x \in A_n$. Then, by definitions, $x \in S_{d(n)}$ and (3.10) holds. Also, x can be written as

$$(3.12) \qquad x = x_1x_2\cdots x_{d(n)}x_{d(n)+1}\cdots x_{d(x)}, \quad d(x) \geq d(n),$$

or

$$(3.13) \qquad x = x_1x_2\cdots x_{d(n)}\cdot x_{d(n)+1}\cdots x_{d(x)}10^{d(x)-d(n)}$$

$$= n10^{d(x)-d(n)} + \cdot x_{d(n)+1}\cdots x_{d(x)}10^{d(x)-d(n)},$$

which implies that

$$(3.14) \qquad n10^{d(x)-d(n)} \leq x \leq n10^{d(x)-d(n)} + 10^{d(x)-d(n)}-1,$$

since the inequality

$$(3.15) \qquad 0 \leq \cdot x_{d(n)+1}\cdots x_{d(x)}10^{d(x)-d(n)} \leq 10^{d(x)-d(n)}-1,$$

holds for every $x_i \in \{0,1,2,\cdots,9\}$, $i=d(n)+1$, $d(n)+2,\cdots,d(x)$. Furthermore, $d(x)-d(n)$ is a non-negative integer. It follows that

$$x \in \bigcup_{m=0}^{\infty} \bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\}, \tag{3.16}$$

and the assertion of (3.9) follows.

Lemma 3.3. The probability of the set of all positive integers having the leftmost $d(n)$ digits equal to $n, n \in Z_p$, is given by the series

$$\sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(n10^m-1)\}, \tag{3.17}$$

where F is the distribution function of ξ.

Proof. The set of all positive integers whose leftmost $d(n)$ digits equal to n is A_n. Since P is the probability measure on the sigma field $\mathcal{C}$, the power set of Z_p, the probability of A_n, denoted by $P(A_n)$, is

$$P(A_n) = P\left(\bigcup_{m=0}^{\infty} \bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\}\right) = \sum_{m=0}^{\infty} P\left(\bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\}\right). \tag{3.18}$$

The last equality holds, for P is a countably additive measure on $\mathcal{G}$, and for each m the sets are disjoint.

Consider now the equality

$$(3.19)\qquad \bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\} \cup \left(\bigcup_{i=1}^{n10^m-1} \{i\}\right) = \bigcup_{i=1}^{(n+1)10^m-1} \{i\}.$$

Applying the measure P to this last equation and using the definition of the distribution function F, we obtain

$$(3.20)\qquad P\left(\bigcup_{i=n10^m}^{(n+1)10^m-1} \{i\}\right) = P[\xi \leq (n+1)10^m-1] - P[\xi \leq n10^m-1]$$

$$= F[(n+1)10^m-1] - F(n10^m-1).$$

We recall that $[\xi \leq x] \equiv \{i \in Z_p \mid \xi(i) \leq x\}$. Hence, the lemma is proved.

Lemma 3.4. Let F be the distribution of ξ. Then the series

$$(3.21)\qquad \sum_{m=0}^{\infty} \{F(x10^m-1) - F(y10^m-1)\}$$

converges absolutely for all positive real numbers x and y.

Proof. First we show that the series

$$(3.22) \qquad \sum_{m=0}^{\infty} \{F(n10^m-1)-F(10^m-1)\}$$

converges absolutely for each $n \in Z_p$ by induction on n. For $n = 1$ the series vanishes indentically. Now suppose it is true for $n = k$. Consider the identity

$$(3.23) \quad \sum_{m=0}^{\infty} \{F[(k+1)10^m-1]-F(10^m-1)\} = \sum_{m=0}^{\infty}\{F[(k+1)10^m-1]-F(k10^m-1)\}$$

$$+ \sum_{m=0}^{\infty} \{F(k10^m-1)-F(10^m-1)\}.$$

The first series on the right hand side is precisely $P(A_k)$, by Lemma 3.3. Also, $0 < P(A_k) < 1$. The remaining series converges by the induction hypothesis. Therefore, the last equality holds from theorem A.1. Consequently, the series converges for all $n \in Z_p$. We note also that for each $n \in Z_p$ and every non-negative integer m, $10^m-1 \leq n10^m-1$. Therefore $0 \leq F(10^m-1) \leq F(n10^m-1)$. We have the absolute convergence of the series.

Now we are in the position to prove Lemma 3.4. Without loss of generality let us assume that $y \leq x$. For

$x > 0$ there exists a positive integer n such that $x \leq n$. It follows that $x10^m-1 \leq n10^m-1$ and

$$F(x10^m-1) \leq F(n10^m-1). \tag{3.24}$$

On the other hand, $y > 0$ implies either $y \geq 1$ or $0 < y < 1$.

Case 1. $1 \leq y$. Then $10^m-1 \leq y10^m-1$ and $F(10^m-1) \leq F(y10^m-1)$. So $-F(y10^m-1) \leq -F(10^m-1)$. By adding this last inequality to that of (3.24) we obtain, summing over m,

$$0 \leq \sum_{m=0}^{\infty} \{F(x10^m-1)\}-F(y10^m-1) \tag{3.25}$$

$$\leq \sum_{m=0}^{\infty} \{F(n10^m-1)-F(10^m-1)\} < \infty.$$

Case 2. $0 < y < 1$. Then for each arbitrary but fixed y in the open unit interval there exists a positive integer k such that $0 < \frac{1}{10^k} \leq y$. It follows that $\frac{10^m}{10^k} - 1 \leq y10^m-1$ and $-F(y10^m-1) \leq -F(\frac{10^m}{10^k} - 1)$. Again, using (3.24), we have

$$0 \leq \sum_{m=0}^{\infty} \{F(x10^m-1)-F(y10^m-1)\} \leq \sum_{m=0}^{\infty} \{F(n10^m-1)-F(\frac{10^m}{10^k}-1)\}$$

$$(3.26) \qquad = \sum_{m=0}^{k} F(n10^m-1) + \sum_{m=k+1}^{\infty} \{F(n10^m-1)-F(\frac{10^m}{10^k}-1)\}$$

$$= \sum_{m=0}^{k} F(n10^m-1) + \sum_{j=1}^{\infty} \{F(n10^{k+j}-1)-F(10^j-1)\} < \infty .$$

The series converges absolutely, and the assertion is proved.

Lemma 3.5. If $k \in Z_p$ and $k > 1$, then for every $r \in Z_p$ there exists a non-negative integer s such that

$$(3.27) \qquad k^s \leq 2^r < k^{s+1} .$$

Proof. We observe that for all real numbers $a > 0$,

$$(3.28) \qquad [a] \leq a < [a] + 1,$$

where $[\ ]$ denotes the greatest integer function. It follows that, for $k \in Z_p$ and $k > 1$,

$$(3.29) \qquad k^{[a]} \leq k^a < k^{[a]+1} .$$

Now $k > 1$ implies that for all $r \in Z_p$ $\log_k 2^r > 0$. Setting $a = \log_k 2^r$, we get

$$k^{[\log_k 2^r]} \leq k^{\log_k 2^r} = 2^r < k^{[\log_k 2^r]+1}, \tag{3.30}$$

and the proof is complete.

Lemma 3.6. Let L be a strictly increasing function on Z_p such that, for all m, $n \in Z_p$,

$$L(mn) = L(m) + L(n). \tag{3.31}$$

Then, for $n \in Z_p$,

$$L(n) = c \log_b n, \quad c > 0, \quad b > 1. \tag{3.32}$$

Proof. By induction it is easy to show, for $m \in Z_p$,

$$L(m^k) = k\, L(m), \tag{3.33}$$

for all non-negative integers k, since $L(1) = L(1) + L(1)$ implies $L(1) = 0$.

Now we show that $L(n) = c \log_b n$ for all $n \in Z_p$. For $n = 1$, $L(1) = 0$, and $\log_b 1 = 0$, if $b > 1$. Hence, the conclusion holds for $n = 1$. Now let n be a fixed but arbitrary integer greater than 1. Then by Lemma 3.5, for each $r \in Z_p$ there is a non-negative integer s such that

$$n^s \leq 2^r < n^{s+1}. \tag{3.34}$$

Using the strict monotonicity of L, we obtain

$$s\,L(n) \leq r\,L(2) < (s+1)\,L(n), \tag{3.35}$$

or, equivalently,

$$\frac{s}{r} \leq \frac{L(2)}{L(n)} < \frac{s+1}{r}. \tag{3.36}$$

On the other hand, using (3.34) again we obtain

$$s \log_b n \leq r \log_b 2 < (s+1)\log_b n, \tag{3.37}$$

or

$$\frac{s}{r} \leq \frac{\log_b 2}{\log_b n} < \frac{s+1}{r}. \tag{3.38}$$

It follows that

$$0 \leq \left|\frac{L(2)}{L(n)} - \frac{\log_b 2}{\log_b n}\right| < \frac{1}{r}, \tag{3.39}$$

and

$$L(n) = \frac{L(2)}{\log_b 2} \log_b n, \tag{3.40}$$

if r is allowed to increase without limit. The assertion follows.

Now we are in the position to prove the main results,

but first we must discuss the following theorem:

Theorem 3.1. The following statements are equivalent:

(a) The invariance principle, defined by (ii).

(b) $$\sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(10^{d(n)-1}10^m-1)\}$$

$$= \sum_{m=0}^{\infty} \{F[\frac{(n+1)10^m}{c} - 1]-F(\frac{10^{d(n)-1}10^m}{c} - 1)\},$$

where $c > 0$, $n = 10^{d(n)-1}, 10^{d(n)-1}+1, \cdots, 10^{d(n)}-1.$

(c) $$\sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(n10^m-1)\}$$

$$= \sum_{m=0}^{\infty} \{F[\frac{(n+1)10^m}{c} - 1]-F(\frac{n10^m}{c} - 1)\},$$

where c and n are defined as in (b).

(d) The invariance principle of the density distribution, given by (iii).

Proof. First of all, we observe that the infinite series, by Lemma 3.4, are well-defined. Part 1. (a) implies (b). Let $n \in Z_p$. Then

$$S_{d(n)} = \bigcup_{i=10^{d(n)-1}}^{10^{d(n)}-1} A_i \tag{3.41}$$

and the A_i are disjoint, by Lemma 3.1. Lemmas 3.2 and 3.3 give

$$P(A_i) = \sum_{m=0}^{\infty} \{F[(i+1)10^m-1]-F(i10^m-1)\}. \tag{3.42}$$

Using the fact that P has the countable additivity property and applying Theorem A.2 we obtain

$$P\left(\bigcup_{i=10^{d(n)-1}}^{n} A_i\right) = \sum_{i=10^{d(n)-1}}^{n} \sum_{m=0}^{\infty} \{F[(i+1)10^m-1]-F(i10^m-1)\}$$

$$= \sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(10^{d(n)-1}10^m-1)\}, \tag{3.43}$$

which is the probability of the set of all positive integers having the leftmost $d(n)$ digits less than or equal to n and greater than or equal to $10^{d(n)-1}$, for all n, $10^{d(n)-1} \leq n \leq 10^{d(n)}-1$. Hence, the last infinite series defines the distribution function G of the leftmost $d(n)$ digits such that

$$(3.44)\quad G(10^{d(n)}-1) = P\Big(\bigcup_{i=10^{d(n)-1}}^{10^{d(n)}-1} A_i\Big) = 1 - \sum_{i\in Z_p - S_{d(n)}} f(i),$$

where $Z_p - S_{d(n)}$ denotes the complement of the set $S_{d(n)}$ with respect to the set Z_p. By the definition of the distribution function of a random variable and Theorem A.3

$$P\Big(\bigcup_{i=10^{d(n)-1}}^{n} A_i\Big) = \sum_{m=0}^{\infty} \{P[\xi \le (n+1)10^m - 1] - P[\xi \le 10^{d(n)-1}10^m - 1]\}$$

$$(3.45)\qquad = \sum_{m=0}^{\infty} \{P[\xi+1 \le (n+1)10^m] - P[\xi+1 \le 10^{d(n)-1}10^m]\}$$

$$= \sum_{m=0}^{\infty} \{F_{\xi+1}[(n+1)10^m] - F_{\xi+1}(10^{d(n)-1}10^m)\};$$

by hypothesis, namely that the probability distribution of the leftmost $d(n)$ digits is invariant under the linear mapping $\eta = c(\xi+1)$, $c > 0$,

$$P\Big(\bigcup_{i=10^{d(n)-1}}^{n} A_i\Big) = \sum_{m=0}^{\infty} \{F_{c(\xi+1)}[(n+1)10^m] - F_{c(\xi+1)}(10^{d(n)-1}10^m)\}$$

$$(3.46)\qquad = \sum_{m=0}^{\infty} \{P[c(\xi+1) \le (n+1)10^m] - P[c(\xi+1) \le 10^{d(n)-1}10^m]\}$$

(3.46) continued

$$= \sum_{m=0}^{\infty} \{F[\frac{(n+1)10^m}{c} - 1] - F(\frac{10^{d(n)-1}10^m}{c} - 1)\}.$$

Part 2. (b) implies (c). For $n = 10^{d(n)-1}$ (b) gives

$$\sum_{m=0}^{\infty} \{F[(10^{d(n)-1}+1)10^m - 1] - F(10^{d(n)-1}10^m - 1)\}$$

(3.47)

$$= \sum_{m=0}^{\infty} \{F[\frac{(10^{d(n)-1}+1)10^m}{c} - 1] - F(\frac{10^{d(n)-1}10^m}{c} - 1)\}.$$

Here we have used the fact that $d(n) = d(10^{d(n)-1})$. Therefore (c) holds for $n = 10^{d(n)-1}$. It remains to be shown for $n = 10^{d(n)-1}+1$, $10^{d(n)-1}+2, \cdots, 10^{d(n)}-1$. For these values of n, $d(n-1) = d(n)$ and

$$\sum_{m=0}^{\infty} \{F(n10^m - 1) - F(10^{d(n)-1}10^m - 1)\}$$

(3.48)

$$= \sum_{m=0}^{\infty} F(\frac{n10^m}{c} - 1) - F(\frac{10^{d(n)-1}10^m}{c} - 1).$$

By Theorem A.1 subtracting the last equation from that of (b) is permissible and the algebraic operation yields the desired result.

Part 3. (c) implies (d). From the hypothesis and Theorem A.3,

$$\sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(n10^m-1)\} = P(A_n)$$

(3.49)

$$= \sum_{m=0}^{\infty} \{F_{\xi+1}[(n+1)10^m]-F_{\xi+1}(n10^m)\},$$

for $n = 10^{d(n)-1},\cdots,10^{d(n)}-1$. So the left side series in (c) gives the probability of the first $d(n)$ digits; it defines the probability density function of the leftmost $d(n)$ digits. On the other hand, the remaining series can be written as

(3.50) $$\sum_{m=0}^{\infty} \{F_{c(\xi+1)}[(n+1)10^m]-F_{c(\xi+1)}(n10^m)\}.$$

By hypothesis,

$$\sum_{m=0}^{\infty} \{F_{\xi+1}[(n+1)10^m]-F_{\xi+1}(n10^m)\}$$

(3.51) $$= \sum_{m=0}^{\infty} \{F_{c(\xi+1)}[(n+1)10^m]-F_{c(\xi+1)}(n10^m)\}.$$

Therefore, (c) asserts that the probability density function

of $d(n)$ digits, whose values are $P(A_n)$, $10^{d(n)-1} \leq n \leq 10^{d(n)}-1$, remains unchanged under linear transformation $\eta = c(\xi + 1)$.

Part 4. (d) implies (a). The proof is immediate by observing that the addition of $n - 10^{d(n)-1}+1$ equations, at the end of Part 3, for the values $10^{d(n)-1},\cdots,n$ yields

$$\sum_{m=0}^{\infty} \{F_{\xi+1}[(n+1)10^m]-F_{\xi+1}(10^{d(n)-1}10^m)\}$$

(3.52)

$$= \sum_{m=0}^{\infty} \{F_{c(\xi+1)}[(n+1)10^m]-F_{c(\xi+1)}(10^{d(n)-1}10^m)\}.$$

We recall that the series on the left side of the last equality is precisely

$$P\Big(\bigcup_{i=10^{d(n)-1}}^{n} A_i\Big), \tag{3.53}$$

and the proof is complete.

The series

$$\sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(10^{d(n)-1}10^m-1)\}, \tag{3.54}$$

as we have seen, defines a real-valued function G on Z_p. If n takes the values $10^{d(n)-1},\cdots,10^{d(n)}-1$, then $G(n)$

is the probability of the set of all positive integers having the leftmost $d(n)$ digits less than or equal to n and greater than or equal to $10^{d(n)-1}$. However, the function G so defined is neither vanishing for $n = 1$ nor monotone on Z_p. To show the non-monotonicity we first observe that, for each $n \in Z_p$,

$$d(n+1) = \begin{cases} d(n), & \text{if } n \neq 10^{d(n)}-1, \\ d(n)+1, & \text{if } n = 10^{d(n)}-1. \end{cases} \tag{3.55}$$

By Theorem A.1 and Lemma 3.4 the values of G at $n+1$ can be written as

$$G(n+1) = P(A_{n+1}) + G(n) - \sum_{m=0}^{\infty} \{F(10^{d(n+1)-1}10^m-1) - F(10^{d(n)-1}10^m-1)\}. \tag{3.56}$$

For $n \neq 10^{d(n)}-1$, $G(n+1) = P(A_{n+1}) + G(n) > G(n)$. On the other hand, for $n = 10^{d(n)}- 1$, say $n = 9$,

$$G(10) = P(A_{10}) + G(9) - P(\bigcup_{i=1}^{9} A_i)$$

$$= P(A_{10}) + G(9) - 1 < G(9),$$

for $0 < P(A_n) < 1$, all n. On the contrary, the real-valued function Q on Z_p,

$$(3.57) \qquad Q(n) = \sum_{m=0}^{\infty} \{F(n10^m-1)-F(10^m-1)\},$$

has the desired properties.

Theorem 3.2. The function Q defined above is a strictly increasing function on Z_p such that

$$(3.58) \qquad Q(kh) = Q(k) + Q(h),$$

for all $k, h \in Z_p$.

Proof. Lemma 3.4 asserts that Q is well-defined for all $n \in Z_p$. To show the monotonicity consider

$$(3.59) \qquad Q(n+1) = P(A_n) + Q(n),$$

and

$$(3.60) \qquad P(A_n) = \sum_{i \in A_n} f(i) > 0.$$

The positivity follows from stipulation (i) above, for obviously A_n is an infinite set. Therefore $Q(n+1) > Q(n)$ for all $n \in Z_p$.

Finally consider

$$Q(k) = \sum_{m=0}^{\infty} \{F(k10^m-1)-F(10^m-1)\}$$

(3.61)

$$= \sum_{m=0}^{\infty} \left\{ \sum_{j=2}^{k} [F(j10^m-1)-F((j-1)10^m-1)] \right\}$$

$$= \sum_{j=2}^{k} \left\{ \sum_{m=0}^{\infty} [F(j10^m-1)-F((j-1)10^m-1)] \right\},$$

which follows from Theorem A.2. Now assumption (ii) and Theorem 3.1 yield, taking $c = 1/h$, $h \in Z_p$,

$$Q(k) = \sum_{j=2}^{k} \sum_{m=0}^{\infty} [F(hj10^m-1)-F(h(j-1)10^m-1)]$$

$$= \sum_{m=0}^{\infty} \{F(hk10^m-1)-F(h10^m-1)\}$$

(3.62)

$$= \sum_{m=0}^{\infty} \{F(hk10^m-1)-F(10^m-1)\}$$

$$- \sum_{m=0}^{\infty} \{F(h10^m-1)-F(10^m-1)\}$$

$$= Q(hk) - Q(h).$$

This completes the proof.

Theorem 3.3. Let $n \in Z_p$. Then the probability of the set of all positive integers having the leftmost $d(n)$ digits less than or equal to n and greater than or equal to $10^{d(n)-1}$ is given by

$$\log_b \frac{n+1}{10^{d(n)-1}}, \quad 10^{d(n)-1} \leq n \leq 10^{d(n)}-1, \tag{3.63}$$

where $b = 10$ if $d(n) = 1$ and, for $d(n) \geq 2$, b must satisfy the equation

$$\log_b 10 + \sum_{i=1}^{10^{d(n)-1}-1} f(i) = 1. \tag{3.64}$$

Proof. Let $n \in Z_p$. Then $10^{d(n)-1} \leq n \leq 10^{d(n)}-1$. From the previous discussion it suffices to show that the function G gives the desired results.

Consider

$$\begin{aligned} G(n) &= \sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(10^{d(n)-1}10^m-1)\} \\ &= \sum_{m=0}^{\infty} \{F[(n+1)10^m-1]-F(10^m-1)\} \end{aligned} \tag{3.65}$$

(3.65) continued

$$- \sum_{m=0}^{\infty} \{F(10^{d(n)-1}10^m-1)-F(10^m-1)\}$$

$$= c_1 \log_{b_1}(n+1) - c_2 \log_{b_2}(10^{d(n)-1}).$$

Here we have used Lemmas 3.4 and 3.6 and Theorems A.1 and 3.2. The parameters, $c_1, c_2 > 0$ and $b_1, b_2 > 1$, are at our disposal. For no particular reason other than convenience we eliminate some of the parameters by choosing

$$(3.66) \qquad c_2 = c_1 \log_{b_1} b_2, \qquad c_1 = \frac{1}{\log_{b_1} b}, \quad b > 1.$$

Some manipulation reveals that

$$(3.67) \qquad G(n) = \log_b \frac{n+1}{10^{d(n)-1}},$$

for all $n \in Z_p$.

If $d(n) = 1$ then $1 \leq n \leq 9$ and

$$(3.68) \qquad P(\bigcup_{i=1}^{n} A_i) = G(n) = \log_b(n+1).$$

But

$$(3.69) \qquad \bigcup_{i=1}^{9} A_i = Z_p$$

and $P(Z_p) = 1 = \log_b 10$, therefore $b = 10$.

If $d(n) \geq 2$ then $10^{d(n)-1} \leq n \leq 10^{d(n)}-1$ and

$$(3.70) \qquad S_{d(n)} = \bigcup_{i=10^{d(n)-1}}^{10^{d(n)}-1} A_i = Z_p - \bigcup_{i=1}^{10^{d(n)-1}-1} \{i\}.$$

Therefore

$$(3.71) \qquad P(S_{d(n)}) = \log_b \frac{(10^{d(n)}-1)+1}{10^{d(n)-1}} = 1 - \sum_{i=1}^{10^{d(n)-1}-1} f(i),$$

since $P(\{i\}) = f(i)$ by definition. The theorem is proved.

Before proceeding let us observe that Theorem 3.3 yields

$$(3.72) \qquad G(n) = \log_{10}(n+1), \quad n = 1,2,\cdots,9,$$

which is the distribution of the first digits as obtained by Pinkham and the others.

Verifications. In concluding the chapter we verify Benford's empirical formulas cited in the introduction. For doing so the following corollary is essential:

Corollary 3.1. For any $n \in Z_p$

$$(3.73) \qquad P(A_n) = \log_b \frac{n+1}{n}, \quad 10^{d(n)-1} \leq n \leq 10^{d(n)}-1,$$

where b is defined as that of Theorem 3.3.

Proof. The proof is immediate if we observe

$$(3.74)\quad A_n \cup \Big(\bigcup_{i=10^{d(n)-1}}^{n-1} A_i\Big) = \bigcup_{i=10^{d(n)-1}}^{n} A_i,$$

$$10^{d(n)-1} \leq n \leq 10^{d(n)}-1.$$

By the countable additivity property of the probability measure P,

$$(3.75)\qquad P(A_n) = P\Big(\bigcup_{i=10^{d(n)-1}}^{n} A_i\Big) - P\Big(\bigcup_{i=10^{d(n)-1}}^{n-1} A_i\Big).$$

It follows that Theorem 3.3 asserts

$$(3.76)\qquad P(A_n) = \log_b \frac{n+1}{10^{d(n)-1}} - \log_b \frac{n}{10^{d(n)-1}}.$$

Now it is a trivial matter to establish Benford's first formula, for if in the Corollary we let n take on the values $1,2,\cdots,9$ then $b = 10$ by Theorem 3.3. Consequently, the Corollary establishes the claim.

To verify the remaining formula let $n \in Z_p$ such that $d(n) \geq 2$. Then n can be written in digital form as follows:

$$n = n_1 n_2 \cdots n_{d(n)-1}\, n_{d(n)} \,, \tag{3.77}$$

where

$$n_1 \in \{1,2,\cdots,9\},$$

$$n_i \in \{0,1,\cdots,9\}, \quad i = 2,3,\cdots,d(n). \tag{3.78}$$

Observe that

$$A_n = A_n \cap (\bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i}). \tag{3.79}$$

Applying Theorem A.4 we obtain

$$P(A_n) = P(\bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i}) \; P(A_n | \bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i}) \tag{3.80}$$

having

$$P(A_n | \bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i}) \tag{3.81}$$

as the conditional probability of A_n given that the compound event has occurred. That is, the conditional probability gives the probability of the event that the leftmost digits of the set of all the positive integers, $d(n)$ in

number, are $n_1 n_2 \cdots n_{d(n)}$, while it is known that the events $A_{n_1 n_2 \cdots n_{d(n)-1}0}$, $A_{n_1 n_2 \cdots n_{d(n)-1}1}$, $\cdots$, $A_{n_1 n_2 \cdots n_{d(n)-1}9}$, have occurred. A moment of consideration reveals that the probability of the set of all positive integers having their $d(n)^{th}$ digits equal to $n_{d(n)}$, given the first $d(n)-1$ digits being $n_1 n_2 \cdots n_{d(n)-1}$, is given by the conditional probability

$$(3.82) \quad P(A_n | \bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i}) = \frac{P(A_n)}{P(\bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i})}$$

$$= \frac{\log_b \frac{n+1}{n}}{\sum_{i=0}^{9} \log_b \frac{n_1 n_2 \cdots n_{d(n)-1} i+1}{n_1 n_2 \cdots n_{d(n)-1} i}}$$

$$= \frac{\log_b \frac{n+1}{n}}{\log_b \frac{n_1 n_2 \cdots n_{d(n)-1}+1}{n_1 n_1 \cdots n_{d(n)-1}}}$$

which is precisely Benford's second formula. By observing the fact that

$$\log_a b \log_b x = \log_a x, \quad x > 0, \; a, b > 1,$$

we restate the result as

Corollary 3.2. Let $n \in Z_p$ be $n = n_1 n_2 \cdots n_{d(n)}$ such that $d(n) \geq 2$. Then

$$P(A_n \mid \bigcup_{i=0}^{9} A_{n_1 n_2 \cdots n_{d(n)-1} i}) \tag{3.83}$$

$$= \frac{\log_{10} \frac{n_1 n_2 \cdots n_{d(n)}+1}{n_1 n_2 \cdots n_{d(n)}}}{\log_{10} \frac{n_1 n_2 \cdots n_{d(n)-1}+1}{n_1 n_2 \cdots n_{d(n)} - 1}} .$$

CHAPTER IV

Distribution of the Leading Digits under Computation

The study of errors arising in digital computation has been a center for investigation. This is not surprising, for if we recall that the advancement in the development of the high speed digital automatic computers has made it possible to carry out a long sequence of algebraic operations which previously was not possible. However, as is well-known, the digital automatic computers provide only a given precision, although multiple precision is obtainable through the use of special sub-routines. In any event, for a given precision, standard or multiple, the computed numbers may easily contain more digits than the given precision allows. Consequently, an error is involved in the computed answer due to rounding off so as to reduce the computed number back to the permissible range of the precision. An extensive collection of references of published results concerning errors in digital computation can be found in Rall [1964, 1965]. In particular, a profound investigation concerning rounding errors has been carried out by Wilkinson [1963], whose works have influenced the frontier research along this direction. To facilitate the discussion we cite the following example from Wilkinson. A rigorous bound of the cumulative effect of rounding errors is obtained on the extended product P_n of n numbers, each of which has a t

digit mantissa and the operation is done in a standard precision t. The product P_n is defined by

$$P_n = fl(x_1x_2x_3\cdots x_n). \tag{4.1}$$

A caution is in order. Here the x_i are real numbers, not digits. The notation signifies the floating point computation on n numbers. The algebraic operations in automatic digital computations always proceed from left to right. The quantities P_r are defined by the recursion formula:

$$\begin{aligned} P_1 &= x_1, \\ P_r &= fl(P_{r-1}x_r) \equiv P_{r-1}x_r(1 + \varepsilon_r), \end{aligned} \tag{4.2}$$

and

$$|\varepsilon_r| \leq \frac{1}{2}\, 10^{1-t}, \quad r = 2,3,\cdots,n. \tag{4.3}$$

The ε_r are rounding errors. It follows that

$$P_n \equiv x_1x_2\cdots x_n(1+\varepsilon_2)(1+\varepsilon_3)\cdots(1+\varepsilon_n). \tag{4.4}$$

Therefore

$$fl(x_1x_2\cdots x_n) \equiv x_1x_2\cdots x_n(1+E) \tag{4.5}$$

and

$$(4.6) \qquad (1 - \tfrac{1}{2}\, 10^{1-t})^{n-1} \leq 1 + E \leq (1 + \tfrac{1}{2}\, 10^{1-t})^{n-1}.$$

This is the most rigorous bound for the rounding procedure, namely adding $\frac{1}{2}\, 10^{-t}$ to each of the normalized computed numbers, $P_{r-1}x_r$. It is clear that the computed result attains its maximum error only where the numbers $x_1, x_2, \cdots x_n$ must be very special quantities. So the maximum error is not likely to occur in a given sequence of multiplication. For this reason we shift our attention to the theory of the most significant digit.

Hamming [1962] pointed out that in order to understand the effect of the cumulative error in the product or quotient it is necessary to investigate the distribution of the most significant digits. For if x_1 and x_2 are the two numbers with errors ε_1 and ε_2 respectively then

$$(4.7) \qquad (x_1+\varepsilon_1)(x+\varepsilon_2) = x_1x_2 + x_1\varepsilon_2 + x_2\varepsilon_1 + \varepsilon_1\varepsilon_2.$$

Hamming observed that the "leading digits in x_1 and x_2 tend to control the roundoff propagation and, by a similar argument, through division". However, we observe that not only do the leading digits of x_1 and x_2 influence the cumulative error but also the leading digits in the products $x_1\varepsilon_2$ and $x_2\varepsilon_1$ play an important role in the building up

of the error. Therefore, it is desirable to know the distribution of the leading digits under algebraic operations, multiplication and division. The distribution of the leading digits in the product and in the quotient of two numbers can be found in Hamming, obtained through a "private communication" from R.C. Prim, III. We cite them as follows: starting with the initial uniform density function of x, defined to be

$$f_i(x) = \frac{1}{9}, \quad x \in [1,10], \tag{4.8}$$

the density distribution f_p of the product, in floating-point form, of two numbers, selected independently from the initial distribution f_i, is given by

$$f_p(x) = \frac{10 \ln 10 - 9 \ln x}{81}, \quad x \in [1,10], \tag{4.9}$$

and the density distribution of the quotient, in floating-point form, of two numbers, selected independently from the initial distribution f_i, is given by

$$f_q(x) = \frac{1}{18}\left(1 + \frac{10}{x^2}\right), \quad x \in [1,10]. \tag{4.10}$$

Also, an assertion is made that the density distribution of "a long sequence of independent multiplications and/or divisions of numbers from any (reasonable) initial distribution"

has the form

$$(4.11) \qquad f_s(x) = \frac{1}{(\ln 10)x}, \quad x \in [1,10].$$

For the purpose of comparing the foregoing three continuous functions with that of Corollary 3.1, namely $P(A_n) = \log_{10}(1 + \frac{1}{n})$, $n = 1,2,\cdots,9$, we plot these four density functions in two different graphs. One of them shows the relative values of these functions at $n = 1,2,\cdots,9$ and the other gives the comparison of the probabilities attained by these functions at the same points. The numerical values of each function have been carried out to four decimal places and then rounded off to three places by adding $\frac{1}{2}\,10^{-3}$ to the computed values in four decimal places. The figures are placed in the appendix.

The first figure reveals that the graphs of the density functions at the n distinct points agree closely, except for $n = 1$. However, the comparison between these three continuous functions and a discrete function would not show the main features possessed by the functions. On the other hand, the comparison between two probabilities of an event associated with the density functions provides a relative measure of the difference in the occurrence of the leading digits under the algebraic computation. Also, the second figure shows that, in all four cases, the numbers having

smaller leading digits occur more frequently than those possessing greater leading digits. Furthermore, the probabilities of the leading digits of $f_s(x)$ coincide with those of $P(A_n)$ at $n = 1,2,\cdots,9$. The fact that these two density functions yield the same probability is not surprising. Let us first recall that our second assumption stipulates that the probability distribution of the leftmost $d(n)$ digits is invariant under the linear transformation $\eta = c(\xi + 1)$, where c is any positive real number. Then Theorem 3.1 asserts that this principle is equivalent to the invariance principle of the density distribution. In particular, c, being any real number, could be of the form

$$c = \frac{xy\cdots z}{tu\cdots v}, \tag{4.12}$$

where each symbol in the equation is a real number. Thus, the transformation $\eta = c(\xi + 1)$ has the effect of mapping our original space Z_p into cZ_p, which is defined as follows:

$$cZ_p = \{cx \mid x \in Z_p \text{ and } c \text{ is any fixed real number}\}. \tag{4.13}$$

For example, for $c = \frac{9}{2}$, $\frac{9}{2}Z_p = \left\{\frac{9}{2}x \mid x \in Z_p\right\} = \left\{\frac{9}{2}, \frac{18}{2}, \cdots, \frac{9n}{2}\cdots\right\}$

Secondly, it is now apparent that "a long sequence of independent multiplications and/or divisions of numbers" is precisely an element of the set cZ_p.

BIBLIOGRAPHY

Apostol, Tom M., 1957. Mathematical analysis. Reading, Massachusetts, Addison-Wesley. 559 p.

Benford, Frank, 1938. The law of anomalous numbers. Proceedings of the American Philosophical Society, 78: 551-572.

Brown, G.W., 1951. History of RAND'S random digits - summary. In: Monte Carlo method: Proceedings of a symposium, Los Angeles, 1949. Washington, D.C. p. 31-32. (U.S. National Bureau of Standards. Applied Mathematics Series no. 12).

Flehinger, B.J., 1966. On the probability that a random integer has initial digit A. American Mathematical Monthly 73: 1056-1061.

Goudsmit, S.A. and W.H. Furry. 1944. Significant figures of numbers in statistical tables. Nature 154: 800-801.

Hamming, R.W. 1962. Numerical methods for scientists and engineers. New York, McGraw-Hill. 411 p.

Parzen, E. 1960. Modern probability theory and its applications. New York, 464 p.

Pinkham, R.S. 1961. On the distribution of first significant digits. Annals of Mathematical Statistics 32: 1223-1230.

Rall, Louis B. 1964. Error in digital computation. Vol. 1. 324 p. (U.S. Army. Mathematics Research Center. Publication no. 14).

Rall, Louis B. 1965. Error in digital computation. Vol. 2. New York, Wiley. 288 p. (U.S. Army. Mathematics Research Center. Publication no. 15).

Shannon, C.E. 1948. The mathematical theory of communication. Bell System Technical Journal 27: 379-423, 623-656.

Tucker, H.G. 1967. A graduate course in probability. New York, Academic. 273 p.

Wilkinson, J.H. 1963. Rounding errors in algebraic proceses. New Jersey, Prentice-Hall. 161 p.

APPENDIX

APPENDIX

The following facts are well-established in the literature; for convenience we cite from Apostol [1957], Tucker [1967] and Parzen [1960].

Theorem A.1. Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be convergent series. Then for every pair of constants α and β, the series

$$\sum_{n=1}^{\infty} (a\alpha_n + \beta b_n) \tag{A.1}$$

converges and

$$\sum_{n=1}^{\infty} (a\alpha_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n. \tag{A.2}$$

Theorem A.2. Suppose $f(m,n) \geq 0$ for all m,n. Assume that

$$\sum_{n=1}^{\infty} f(m,n) \tag{A.3}$$

converges for each fixed $m = 1,2,\cdots,$ and that

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f(m,n) \tag{A.4}$$

converges. Then

(A.5a) $\sum_{m=1}^{\infty} f(m,n)$ converges for each $n = 1,2,\cdots,$

(A.5b)

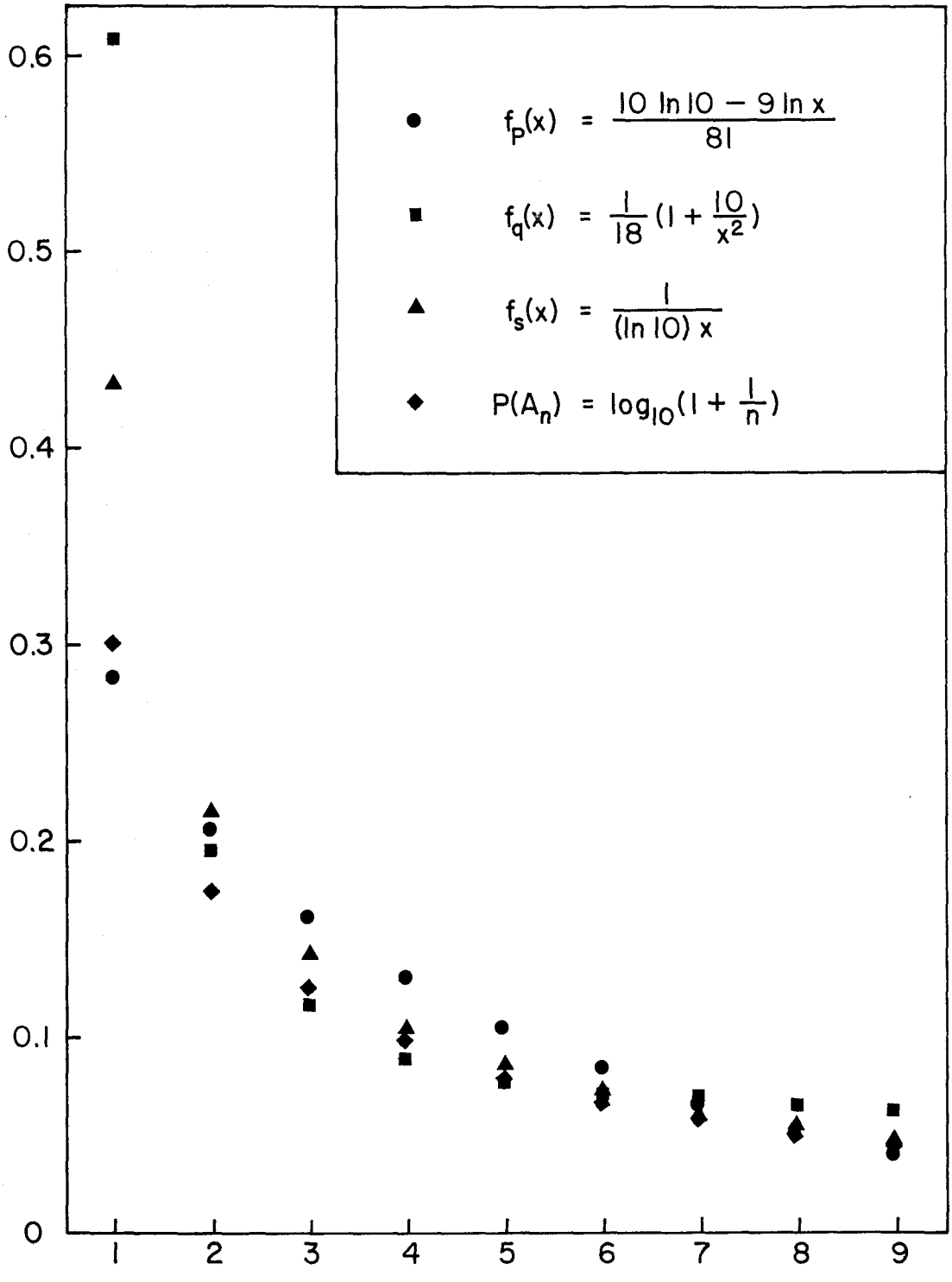
$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(m,n)$ converges and is equal to $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f(m,n)$.

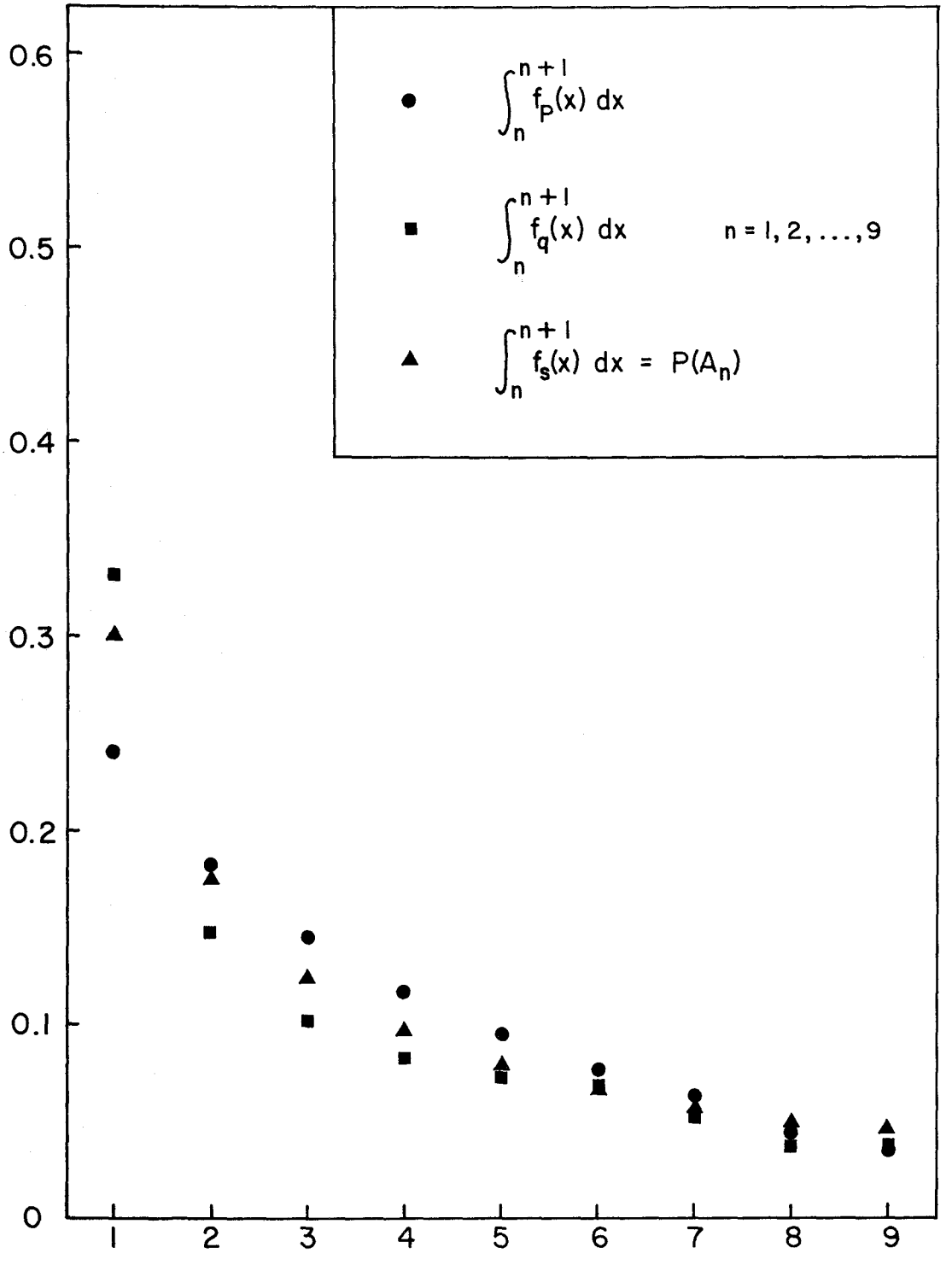
Theorem A.3. Let X be a random variable and let a,b be any real numbers such that $a > 0$. Then the distribution function of the random variable $Y = aX + b$ is given by

$$(A.6) \qquad F_{aX+b}(y) = P[aX+b \leq y] = P[X \leq \frac{y-b}{a}] = F_X(\frac{y-b}{a}), \quad -\infty < y < \infty.$$

Theorem A.4. For every $n + 1$ events $A_0, A_1, \cdots, A_n$ for which $P(A_0A_1\cdots A_n) > 0$ we have

$$(A.7) \qquad P(A_0A_1\cdots A_n) = P(A_0)P(A_1|A_0)\cdots P(A_n|A_0A_1\cdots A_{n-1}).$$


0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
$f_P(x) = \frac{10 \ln 10 - 9 \ln x}{81}$
$f_q(x) = \frac{1}{18}(1 + \frac{10}{x^2})$
$f_s(x) = \frac{1}{(\ln 10)\, x}$
$P(A_n) = \log_{10}(1 + \frac{1}{n})$


0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
$\int_n^{n+1} f_P(x)\,dx$
$\int_n^{n+1} f_q(x)\,dx$
$n = 1, 2, \ldots, 9$
$\int_n^{n+1} f_s(x)\,dx = P(A_n)$