

**INTRODUCTION TO  
MATHEMATICAL PROBABILITY**

INTRODUCTION  
TO  
MATHEMATICAL PROBABILITY

BY  
J. V. USPENSKY  
*Professor of Mathematics, Stanford University*

FIRST EDITION  
SEVENTH IMPRESSION

McGRAW-HILL BOOK COMPANY, Inc.  
NEW YORK AND LONDON  
1937

COPYRIGHT, 1937, BY THE  
MCGRAW-HILL BOOK COMPANY, INC.

---

PRINTED IN THE UNITED STATES OF AMERICA

*All rights reserved. This book, or  
parts thereof, may not be reproduced  
in any form without permission of  
the publishers.*

## PREFACE

This book is an outgrowth of lectures on the theory of probability which the author has given at Stanford University for a number of years. At first a short mimeographed text covering only the elementary parts of the subject was used for the guidance of students. As time went on and the scope of the course was gradually enlarged, the necessity arose of putting into the hands of students a more elaborate exposition of the most important parts of the theory of probability. Accordingly a rather large manuscript was prepared for this purpose. The author did not plan at first to publish it, but students and other persons who had opportunity to peruse the manuscript were so persuasive that publication was finally arranged.

The book is arranged in such a way that the first part of it, consisting of Chapters I to XII inclusive, is accessible to a person without advanced mathematical knowledge. Chapters VII and VIII are, perhaps, exceptions. The analysis in Chapter VII is rather involved and a better way to arrive at the same results would be very desirable. At any rate, a reader who does not have time or inclination to go through all the intricacies of this analysis may skip it and retain only the final results, found in Section 11. Chapter VIII, though dealing with interesting and historically important problems, is not important in itself and may without loss be omitted by readers. Chapters XIII to XVI incorporate the results of modern investigations. Naturally they are more complex and require more mature mathematical preparation.

Three appendices are added to the book. Of these the second is by far the most important. It gives an outline of the famous Tshebysheff-Markoff method of moments applied to the proof of the fundamental theorem previously established by another method in Chapter XIV.

No one will dispute Newton's assertion: "In scientiis addiscendis exempla magis prosunt quam praecepta." But especially is it so in the theory of probability. Accordingly, not only are a large number of illustrative problems discussed in the text, but at the end of each chapter a selection of problems is added for the benefit of students. Some of them are mere examples. Others are more difficult problems, or even important theorems which did not find a place in the main text. In all such cases sufficiently explicit indications of solution (or proofs) are given.

The book does not go into applications of probability to other sciences. To present these applications adequately another volume of perhaps larger size would be required.

No one is more aware than the author of the many imperfections in the plan of this book and its execution. To present an entirely satisfactory book on probability is, indeed, a difficult task. But even with all these imperfections we hope that the book will prove useful, especially since it contains much material not to be found in other books on the same subject in the English language.

J. V. USPENSKY.

STANFORD UNIVERSITY,  
September, 1937.

# CONTENTS

	PAGE
PREFACE . . . . .	v
INTRODUCTION . . . . .	1
Scope of Probability—Necessity and Chance—Genesis of Idea of Probability—Classical Definition of Mathematical Probability—Its Limitations—Short Sketch of Historical Development of Theory of Probability.	
CHAPTER I	
COMPUTATION OF PROBABILITY BY DIRECT ENUMERATION OF CASES . . . . .	14
Simple Problems Solved by Direct Application of Definition—Permutations, Arrangements and Combinations—Some Illustrative Problems—Problems for Solution.	
CHAPTER II	
THEOREMS OF TOTAL AND COMPOUND PROBABILITY . . . . .	27
Mutually Exclusive or Incompatible Events—Theorem of Total Probability in Restricted and Generalized Form—Conditional Probability—Theorem of Compound Probability—Independent Events—Compound Probability of Independent Events—Problems Illustrating the Combined Use of Theorem of Total and Compound Probabilities—Problems for Solution.	
CHAPTER III	
REPEATED TRIALS . . . . .	44
Independent and Dependent Trials—Bernoullian Trials—Most Probable Number of Successes—Approximate Formulas in Case of a Large Number of Trials—Markoff's Method of Continued Fractions—Problems for Solution.	
CHAPTER IV	
PROBABILITIES OF HYPOTHESES AND BAYES' THEOREM . . . . .	60
Bayes' Formula for Probability a Posteriori—Probability of Future Events—Estimation of Probabilities from Observed Frequencies—Criticism of Underlying Assumptions—Problems for Solution.	
CHAPTER V	
USE OF DIFFERENCE EQUATIONS IN SOLVING PROBLEMS OF PROBABILITY . . . . .	74
Simple Chains of Trials—Problem of Runs—Examples Requiring Use of Partial Difference Equations—Problems for Solution.	
CHAPTER VI	
BERNOULLI'S THEOREM . . . . .	96
Statement and Proof of Bernoulli's Theorem—Theorem of Cantelli—Significance of Bernoulli's Theorem—Application to Games of Chance—	

	PAGE
Experimental Verification of Bernoulli's Theorem—Appendix: Buffon's Needle Problem—Problems for Solution.	
CHAPTER VII	
APPROXIMATE EVALUATION OF PROBABILITIES IN BERNOULLIAN CASE . . . . .	119
Statement of Problem—Analytical Representation of Required Probability—Analysis Leading to Final Approximate Formula—Numerical Examples—Problems for Solution.	
CHAPTER VIII	
FURTHER CONSIDERATIONS ON GAMES OF CHANCE. . . . .	139
Problems Concerning Ruin of Gamblers When Number of Games is Unlimited—Stakes Equal—Case of Unequal Stakes—Probability of Ruin in Case of a Limited Number of Games—Problems for Solution.	
CHAPTER IX	
MATHEMATICAL EXPECTATION. . . . .	161
Stochastic Variables—Definition of Mathematical Expectation—Mathematical Expectation of Sums—Examples—Mathematical Expectation of Products—Dispersion and Standard Deviation—Dispersion of Sums—Examples—Problems for Solution.	
CHAPTER X	
THE LAW OF LARGE NUMBERS . . . . .	182
Tshebysheff's Lemma—Law of Large Numbers—Sufficient Conditions for Its Validity—Some Experimental Results—Case of Infinite Dispersion: Theorem of Markoff—Law of Large Numbers for Identical Variables: Theorem of Khintchine—Problems for Solution.	
CHAPTER XI	
APPLICATION OF THE LAW OF LARGE NUMBERS . . . . .	208
Poisson's Theorem—Divergence Coefficient—Lexis' and Poisson's Cases—Tschuprow's Theorem—Examples of Dependent Trials—Experimental Results for Simple Chains—Problems for Solution.	
CHAPTER XII	
PROBABILITIES IN CONTINUUM . . . . .	235
Notion of Probability for Continuous Variables—Case of One Variable—Density of Probability—Distribution Function, Moments and Characteristic Function—Examples—Case of Two or More Variables—Geometrical Problems—Problems for Solution.	
CHAPTER XIII	
THE GENERAL CONCEPT OF DISTRIBUTION . . . . .	260
Increasing Functions and Stieltjes' Integrals—Representation of Distribution Function by Stieltjes' Integral—Moments and Characteristic Function—Inequalities for Moments—Composition of Distribution Functions—Examples—Determination of Distribution When Its Characteristic Function Is Known—Composition of Characteristic Functions—Examples—Problems for Solution.	

CONTENTS

ix

CHAPTER XIV

FUNDAMENTAL LIMIT THEOREMS . . . . .	283
--------------------------------------	-----

Introductory Remarks—Laplace-Liapounoff's Theorem—Fundamental Lemma—Proof of Laplace-Liapounoff's Theorem—Particular Cases—Estimation of Error Term—Hypothesis of Elementary Errors—Examples Concerning Dependent Variables—Problems for Solution.

CHAPTER XV

NORMAL DISTRIBUTION IN TWO DIMENSIONS. LIMIT THEOREM FOR SUMS OF INDEPENDENT VECTORS. ORIGIN OF NORMAL CORRELATION . . . . .	308
--	-----

Properties of Two-dimensional Normal Distribution—Limit Theorem for Sums of Independent Vectors—Application to Trials with Three or More Alternatives—Pearson's " $\chi^2$ -Test"—Hypothetical Explanation of Normal Correlation—Problems for Solution.

CHAPTER XVI

DISTRIBUTION OF CERTAIN FUNCTIONS OF NORMALLY DISTRIBUTED VARIABLES . . . . .	331
---	-----

Distribution of Sums of Squares—Distribution of Sums of Squares of Deviations from the Mean—"Student's" Distribution—Distribution of the Correlation Coefficient of a Sample—Problems for Solution.

APPENDIX I

EULER'S SUMMATION FORMULA—STIRLING'S FORMULA—SOME DEFINITE INTEGRALS . . . . .	347
--	-----

APPENDIX II

METHOD OF MOMENTS AND ITS APPLICATIONS. . . . .	356
---	-----

Introductory Remarks—Continued Fractions Associated with Power Series—Continued Fractions Associated with  $\int_a^b \frac{d\varphi(x)}{z-x}$ . Properties of Denominators of Their Convergents—Equivalent Point Distributions—Tshebysheff's Inequalities—Application to Normal Distribution—Tshebysheff-Markoff's Fundamental Theorem—Application to Sums of Independent Variables.

APPENDIX III

ON A GAUSSIAN PROBLEM . . . . .	396
---------------------------------	-----

Statement of the Problem—Analytical Expression for  $P_n(x)$  and Its Derivative—More General Recurrence Relation—Main Inequalities—Solution of the Gaussian Problem.

TABLE OF PROBABILITY INTEGRAL. . . . .	407
--	-----

INDEX. . . . .	409
----------------	-----



## APPENDIX I

**1. Euler's Summation Formula.** Let  $f(x)$  be a function with a continuous derivative  $f'(x)$  in an interval  $(a, b)$  where  $a$  and  $b > a$  are arbitrary real numbers. The notation

$$\sum_{\substack{n \leq b \\ n > a}} f(n)$$

will be used to designate the sum extended over all integers  $n$  which are  $> a$  and  $\leq b$ . It is an important problem to devise means for the approximate evaluation of the above sum when it contains a considerable number of terms.

Let  $[x]$ , as usual, denote the largest integer contained in a real number  $x$ , so that

$$x = [x] + \theta$$

where  $\theta$ , so-called "fractional part" of  $x$ , satisfies the inequalities

$$0 \leq \theta < 1.$$

Considered as functions of a continuous variable  $x$ , both  $[x]$  and  $\theta$  have discontinuities for integral values of  $x$ . The function

$$\rho(x) = \frac{1}{2} - \theta = [x] - x + \frac{1}{2}$$

is likewise discontinuous for integral values of  $x$ . Besides, it is a periodic function of  $x$  with the period 1; that is, we have

$$\rho(x + 1) = \rho(x)$$

for any real  $x$ . With this notation adopted we have the following important formula:

$$(1) \quad \sum_{\substack{n \leq b \\ n > a}} f(n) = \int_a^b f(x) dx + \rho(b)f(b) - \rho(a)f(a) - \int_a^b \rho(x)f'(x) dx$$

which is known as "Euler's summation formula."

**Proof.** Let  $k$  be the least integer  $> a$  and  $l$  the greatest integer  $\leq b$ . The sum in the left member of (1) is, by definition,

$$f(k) + f(k+1) + \cdots + f(l)$$

and we must show that this is equal to the right member. To this end we write first

$$\int_a^b \rho(x)f'(x)dx = \int_a^k \rho(x)f'(x)dx + \int_l^b \rho(x)f'(x)dx + \sum_{j=k}^{j=l-1} \int_j^{j+1} \rho(x)f'(x)dx.$$

Next, since  $j$  is an integer,

$$\int_j^{j+1} \rho(x)f'(x)dx = \int_j^{j+1} \left( j - x + \frac{1}{2} \right) f'(x)dx = -\frac{f(j) + f(j+1)}{2} + \int_j^{j+1} f(x)dx$$

and

$$\sum_{j=k}^{j=l-1} \int_j^{j+1} \rho(x)f'(x)dx = -\frac{f(k) + f(l)}{2} - \sum_{n=k+1}^{n=l-1} f(n) + \int_k^l f(x)dx.$$

On the other hand,

$$\begin{aligned} \int_a^k \rho(x)f'(x)dx &= \int_a^k \left( k - 1 - x + \frac{1}{2} \right) f'(x)dx = -\frac{f(k)}{2} - \rho(a)f(a) + \int_a^k f(x)dx \\ \int_l^b \rho(x)f'(x)dx &= \int_l^b \left( l - x + \frac{1}{2} \right) f'(x)dx = -\frac{f(l)}{2} + \rho(b)f(b) + \int_l^b f(x)dx, \end{aligned}$$

so that finally

$$\begin{aligned} \int_a^b \rho(x)f'(x)dx &= -f(k) - f(k+1) - \cdots - f(l) + \\ &\quad + \rho(b)f(b) - \rho(a)f(a) + \int_a^b f(x)dx; \end{aligned}$$

whence

$$\sum_{\substack{n \leq b \\ n > a}} f(n) = \int_a^b f(x)dx + \rho(b)f(b) - \rho(a)f(a) - \int_a^b \rho(x)f'(x)dx,$$

which completes the proof of Euler's formula.

**Corollary 1.** The integral

$$\int_0^x \rho(z)dz = \sigma(x)$$

represents a continuous and periodic function of  $x$  with the period 1. For

$$\sigma(x+1) - \sigma(x) = \int_x^{x+1} \rho(z)dz = \int_0^1 \rho(z)dz = \int_0^1 \left( \frac{1}{2} - z \right) dz = 0.$$

If  $0 \leq x \leq 1$ ,

$$\sigma(x) = \int_0^x \left( \frac{1}{2} - z \right) dz = \frac{x(1-x)}{2}$$

and in general

$$\sigma(x) = \frac{\theta(1-\theta)}{2}$$

where  $\theta$  is a fractional part of  $x$ . Hence, for every real  $x$

$$0 \leq \sigma(x) \leq \frac{1}{8}.$$

Supposing that  $f''(x)$  exists and is continuous in  $(a, b)$  and integrating by parts, we get

$$\int_a^b \rho(x)f'(x)dx = \sigma(b)f'(b) - \sigma(a)f'(a) - \int_a^b \sigma(x)f''(x)dx,$$

which leads to another form of Euler's formula:

$$\sum_{\substack{n \leq b \\ n > a}} f(n) = \int_a^b f(x)dx + \rho(b)f(b) - \rho(a)f(a) - \sigma(b)f'(b) + \sigma(a)f'(a) + \int_a^b \sigma(x)f''(x)dx.$$

**Corollary 2.** If  $f(x)$  is defined for all  $x \geq a$  and possesses a continuous derivative throughout the interval  $(a, +\infty)$ ; if, besides, the integral

$$\int_a^{\infty} \rho(x)f'(x)dx$$

exists, then for a variable limit  $b$  we have

$$(2) \quad \sum_{\substack{n \leq b \\ n > a}} f(n) = C + \int f(b)db + \rho(b)f(b) + \int_b^{\infty} \rho(x)f'(x)dx$$

where  $C$  is a constant with respect to  $b$ .

It suffices to substitute for

$$\int_a^b \rho(x)f'(x)dx$$

the difference

$$\int_a^{\infty} \rho(x)f'(x)dx - \int_b^{\infty} \rho(x)f'(x)dx$$

and separate the terms depending upon  $b$  from those involving  $a$ .

**2. Stirling's Formula.** Factorials increase with extreme rapidity and their exact computation soon becomes practically impossible. The question then naturally arises of finding a convenient approximate

expression for large factorials, which question is answered by a celebrated formula usually known as "Stirling's formula," although, in the main, it was established by de Moivre in connection with problems on probability. De Moivre did not establish the relation to the number

$$\pi = 3.14159 \dots$$

of the constant involved in his formula; it was done by Stirling.

In formula (2) it suffices to take  $a = \frac{1}{2}$ ,  $f(x) = \log x$ , and replace  $b$  by an arbitrary integer  $n$  to arrive at the remarkable expression

$$\log(1 \cdot 2 \cdot 3 \cdots n) = C + \left(n + \frac{1}{2}\right) \log n - n + \int_n^\infty \frac{\rho(x) dx}{x}$$

where  $C$  is a constant. For the sake of brevity we shall set

$$\omega(n) = \int_n^\infty \frac{\rho(x) dx}{x}.$$

Now

$$\int_n^\infty \frac{\rho(x) dx}{x} = \int_n^{n+1} \frac{\rho(x) dx}{x} + \int_{n+1}^{n+2} \frac{\rho(x) dx}{x} + \dots$$

and

$$\begin{aligned} \int_k^{k+1} \frac{\rho(x) dx}{x} &= \int_0^1 \frac{\rho(u) du}{u+k} = \int_0^{\frac{1}{2}} \frac{\rho(u) du}{u+k} + \int_{\frac{1}{2}}^1 \frac{\rho(u) du}{u+k} = \\ &= \int_0^{\frac{1}{2}} \frac{(\frac{1}{2}-u) du}{u+k} + \int_{\frac{1}{2}}^1 \frac{(\frac{1}{2}-u) du}{u+k} = \frac{1}{2} \int_0^{\frac{1}{2}} \frac{(1-2u)^2 du}{(k+u)(k+1-u)}. \end{aligned}$$

Hence

$$\omega(n) = \frac{1}{2} \int_0^{\frac{1}{2}} (1-2u)^2 F_n(u) du$$

where

$$F_n(u) = \sum_{k=n}^{\infty} \frac{1}{(k+u)(k+1-u)}.$$

Since

$$(k+u)(k+1-u) = k(k+1) + u - u^2,$$

it follows that for  $0 < u < \frac{1}{2}$

$$\begin{aligned} (k+u)(k+1-u) &> k(k+1) \\ (k+u)(k+1-u) &< (k+\frac{1}{2})^2 < (k+\frac{1}{2})(k+\frac{3}{2}). \end{aligned}$$

Thus for  $0 < u < \frac{1}{2}$

$$F_n(u) < \sum_{k=n}^{\infty} \frac{1}{k(k+1)} = \frac{1}{n}$$

$$F_n(u) > \sum_{k=n}^{\infty} \frac{1}{(k+\frac{1}{2})(k+\frac{3}{2})} = \frac{1}{n+\frac{1}{2}}$$

Making use of these limits, we find that

$$\omega(n) < \frac{1}{2n} \int_0^1 (1-2u)^2 du = \frac{1}{12n}$$

$$\omega(n) > \frac{1}{2n+1} \int_0^1 (1-2u)^2 du = \frac{1}{12(n+\frac{1}{2})}$$

and consequently can set

$$\omega(n) = \frac{1}{12(n+\theta)}$$

where

$$0 < \theta < \frac{1}{2}$$

Accordingly

$$\log(1 \cdot 2 \cdot 3 \cdots n) = C + \left(n + \frac{1}{2}\right) \log n - n + \frac{1}{12(n+\theta)}$$

The constant  $C$  depends in a remarkable way on the number  $\pi$ . To show this we start from the well-known expression for  $\pi$  due to Wallis:

$$\frac{\pi}{2} = \lim_{n \rightarrow \infty} \left( \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdots \frac{2n}{2n-1} \cdot \frac{2n}{2n+1} \right),$$

which follows from the infinite product

$$\sin x = x \left(1 - \frac{x^2}{\pi^2}\right) \left(1 - \frac{x^2}{4\pi^2}\right) \left(1 - \frac{x^2}{9\pi^2}\right) \cdots$$

by taking  $x = \pi/2$ . Since

$$\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdots \frac{2n}{2n-1} \cdot \frac{2n}{2n+1} = \left[ \frac{2 \cdot 4 \cdot 6 \cdots 2n}{1 \cdot 3 \cdot 5 \cdots (2n-1)} \right]^2 \frac{1}{2n+1}$$

we get from Wallis' formula

$$\sqrt{\pi} = \lim_{n \rightarrow \infty} \left[ \frac{2 \cdot 4 \cdot 6 \cdots 2n}{1 \cdot 3 \cdot 5 \cdots (2n-1)} \frac{1}{\sqrt{n}} \right],$$

On the other hand,

$$\begin{aligned} 2 \cdot 4 \cdot 6 \cdots 2n &= 2^n \cdot 1 \cdot 2 \cdot 3 \cdots n \\ 1 \cdot 3 \cdot 5 \cdots (2n-1) &= \frac{1 \cdot 2 \cdot 3 \cdots 2n}{2^n \cdot 1 \cdot 2 \cdot 3 \cdots n} \end{aligned}$$

so that

$$\sqrt{\pi} = \lim \left\{ \frac{2^{2n}(1 \cdot 2 \cdot 3 \cdots n)^2}{1 \cdot 2 \cdot 3 \cdots 2n} \cdot \frac{1}{\sqrt{n}} \right\}, \quad n \rightarrow \infty$$

or, taking logarithms

$$\log \sqrt{\pi} = \lim [2n \log 2 + 2 \log (1 \cdot 2 \cdot 3 \cdots n) - \log (1 \cdot 2 \cdot 3 \cdots 2n) - \frac{1}{2} \log n]$$

But, neglecting infinitesimals,

$$\begin{aligned} \log (1 \cdot 2 \cdot 3 \cdots n) &= C + (n + \frac{1}{2}) \log n - n \\ \log (1 \cdot 2 \cdot 3 \cdots 2n) &= C + (2n + \frac{1}{2}) \log 2n - 2n \end{aligned}$$

whence

$$\lim [2n \log 2 + 2 \log (1 \cdot 2 \cdot 3 \cdots n) - \log (1 \cdot 2 \cdot 3 \cdots 2n) - \frac{1}{2} \log n] = C - \frac{1}{2} \log 2.$$

Thus

$$\log \sqrt{\pi} = C - \frac{1}{2} \log 2, \quad C = \log \sqrt{2\pi}$$

and finally

$$(3) \quad \log (1 \cdot 2 \cdot 3 \cdots n) = \log \sqrt{2\pi} + \left(n + \frac{1}{2}\right) \log n - n + \frac{1}{12(n + \theta)}; \quad 0 < \theta < \frac{1}{2}$$

This is equivalent to two inequalities

$$e^{\frac{1}{12n+6}} < \frac{1 \cdot 2 \cdot 3 \cdots n}{\sqrt{2\pi n} n^n e^{-n}} < e^{\frac{1}{12n}}$$

which show that for indefinitely increasing  $n$

$$\lim \frac{1 \cdot 2 \cdot 3 \cdots n}{\sqrt{2\pi n} n^n e^{-n}} = 1.$$

This result is commonly known as Stirling's formula.

For a finite  $n$  we have

$$1 \cdot 2 \cdot 3 \cdots n = \sqrt{2\pi n} n^n e^{-n} \cdot e^{\omega(n)}$$

where

$$\frac{1}{12(n + \frac{1}{2})} < \omega(n) < \frac{1}{12n}$$

The expression

$$\sqrt{2\pi n} n^n e^{-n}$$

is thus an approximate value of the factorial  $1 \cdot 2 \cdot 3 \cdots n$  for large  $n$  in the sense that the ratio of both is near to 1; that is, the relative error is small. On the contrary, the absolute error will be arbitrarily large for large  $n$ , but this is irrelevant when Stirling's approximation is applied to quotients of factorials.

In this connection it is useful to derive two further inequalities.

Let  $m < n$ ; we have, then,

$$F_m(u) - F_n(u) = \sum_{k=m}^{k=n-1} \frac{1}{(k+u)(k+1-u)};$$

and further, supposing  $0 < u < \frac{1}{2}$ ,

$$F_m(u) - F_n(u) < \sum_{k=m}^{k=n-1} \frac{1}{k(k+1)} = \frac{1}{m} - \frac{1}{n}$$

$$F_m(u) - F_n(u) > \sum_{k=m}^{k=n-1} \frac{1}{(k+\frac{1}{2})(k+\frac{3}{2})} = \frac{1}{m+\frac{1}{2}} - \frac{1}{n+\frac{1}{2}}$$

Hence,

$$\omega(m) - \omega(n) < \frac{1}{12m} - \frac{1}{12n}, \quad \omega(m) - \omega(n) > \frac{1}{12(m+\frac{1}{2})} - \frac{1}{12(n+\frac{1}{2})}$$

and, if  $l$  is a third arbitrary positive integer,

$$\omega(m) + \omega(l) - \omega(n) < \frac{1}{12m} + \frac{1}{12l} - \frac{1}{12n}$$

$$\omega(m) + \omega(l) - \omega(n) > \frac{1}{12(m+\frac{1}{2})} + \frac{1}{12(l+\frac{1}{2})} - \frac{1}{12(n+\frac{1}{2})}.$$

**3. Some Definite Integrals.** The value of the important definite integral

$$\int_0^{\infty} e^{-t^n} dt$$

can be found in various ways. One of the simplest is the following: Let

$$J_n = \int_0^{\infty} e^{-t^n} dt$$

in general where  $n$  is an arbitrary integer  $\geq 0$ . Integrating by parts one can easily establish the recurrence relation

$$J_n = \frac{n-1}{2} J_{n-2};$$

whence

$$J_{2m} = \frac{1 \cdot 3 \cdot 5 \cdots (2m-1)}{2^m} J_0$$

$$J_{2m+1} = \frac{1 \cdot 2 \cdot 3 \cdots m}{2}$$

On the other hand,

$$J_{n+1} + 2\lambda J_n + \lambda^2 J_{n-1} = \int_0^\infty e^{-t} t^{n-1} (t + \lambda)^2 dt,$$

which shows that

$$J_{n+1} + 2\lambda J_n + \lambda^2 J_{n-1} > 0$$

for all real  $\lambda$ . Hence, the roots of the polynomial in the left member are imaginary, and this implies

$$J_n^2 < J_{n+1} J_{n-1}.$$

Taking  $n = 2m$  and  $n = 2m + 1$  and using the preceding expression for  $J_{2m}$  and  $J_{2m+1}$ , we find

$$\frac{2 \cdot 4 \cdot 6 \cdots 2m}{1 \cdot 3 \cdot 5 \cdots (2m-1)} \frac{1}{\sqrt{4m+2}} < J_0 < \frac{2 \cdot 4 \cdot 6 \cdots 2m}{1 \cdot 3 \cdot 5 \cdots (2m-1)} \frac{1}{\sqrt{4m}}$$

But

$$\lim_{m \rightarrow \infty} \frac{2 \cdot 4 \cdot 6 \cdots 2m}{1 \cdot 3 \cdot 5 \cdots (2m-1)} \frac{1}{\sqrt{m}} = \sqrt{\pi};$$

hence

$$J_0 = \int_0^\infty e^{-t} dt = \frac{1}{2} \sqrt{\pi}.$$

Here substituting  $t = \sqrt{a}u$ , where  $a$  is a positive parameter, we get

$$\int_0^\infty e^{-au^2} du = \frac{1}{2} \sqrt{\frac{\pi}{a}}.$$

As a generalization of the last integral we may consider the following one:

$$V = \int_0^\infty e^{-au^2} \cos bu \, du.$$

The simplest way to find the value of this integral is to take the derivative

$$\frac{dV}{db} = - \int_0^\infty e^{-au^2} \sin bu \cdot u \, du$$

and transform the right member by partial integration. The result is



$$\frac{dV}{db} = -\frac{b}{2a}V$$

or

$$d(Ve^{\frac{b^2}{4a}}) = 0,$$

whence

$$V = Ce^{-\frac{b^2}{4a}}.$$

To determine the constant  $C$ , take  $b = 0$ ; then

$$C = (V)_{b=0} = \int_0^{\infty} e^{-au^2} du = \frac{1}{2}\sqrt{\frac{\pi}{a}},$$

so that finally

$$\int_0^{\infty} e^{-au^2} \cos budu = \frac{1}{2}\sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}}.$$

The equivalent form of this integral is as follows:

$$\int_{-\infty}^{\infty} e^{-au^2} \cos budu = \int_{-\infty}^{\infty} e^{-au^2 + iau} du = \sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}}.$$

## APPENDIX II

### METHOD OF MOMENTS AND ITS APPLICATIONS

**1. Introductory Remarks.** To prove the fundamental limit theorem Tshebysheff devised an ingenious method, known as the "method of moments," which later was completed and simplified by one of the most prominent among Tshebysheff's disciples, the late Markoff. The simplicity and elegance inherent in this method of moments make it advisable to present in this Appendix a brief exposition of it.

The distribution of a mass spread over a given interval  $(a, b)$  may be characterized by a never decreasing function  $\varphi(x)$ , defined in  $(a, b)$  and varying from  $\varphi(a) = 0$  to  $\varphi(b) = m_0$ , where  $m_0$  is the total mass contained in  $(a, b)$ . Since  $\varphi(x)$  is never decreasing, for any particular point  $x_0$ , both the limits

$$\begin{aligned}\lim \varphi(x_0 - \epsilon) &= \varphi(x_0 - 0) \\ \lim \varphi(x_0 + \epsilon) &= \varphi(x_0 + 0)\end{aligned}$$

exist when a positive number  $\epsilon$  tends to 0. Evidently

$$\varphi(x_0 - 0) \leq \varphi(x_0) \leq \varphi(x_0 + 0).$$

If

$$\varphi(x_0 - 0) = \varphi(x_0 + 0) = \varphi(x_0),$$

then  $x_0$  is a "point of continuity" of  $\varphi(x)$ . In case

$$\varphi(x_0 + 0) > \varphi(x_0 - 0),$$

$x_0$  is a point of discontinuity of  $\varphi(x)$ , and the positive difference

$$\varphi(x_0 + 0) - \varphi(x_0 - 0)$$

may be considered as a mass concentrated at the point  $x_0$ . In all cases  $\varphi(x_0 - 0)$  is the total mass on the segment  $(a, x_0)$  excluding the end point  $x_0$ , whereas  $\varphi(x_0 + 0)$  is the mass spread over the same segment including the point  $x_0$ .

The points of discontinuity, if there are any, form an enumerable set, whence it follows that in any part of the interval  $(a, b)$  there are points of continuity.

If for any sufficiently small positive  $\epsilon$

$$\varphi(x_0 + \epsilon) > \varphi(x_0 - \epsilon),$$

$x_0$  is called a "point of increase" of  $\varphi(x)$ . There is at least one point of increase and there might be infinitely many. For instance, if

$$\begin{aligned}\varphi(x) &= 0 & \text{for } a \leq x \leq c \\ \varphi(x) &= m_0 & \text{for } c < x \leq b,\end{aligned}$$

then  $c$  is the only point of increase. On the other hand, for

$$\varphi(x) = m_0 \frac{x-a}{b-a}$$

every point of the interval  $(a, b)$  is a point of increase. In case of a finite number of points of increase the whole mass is concentrated in these points and the distribution function  $\varphi(x)$  is a step function with a finite number of steps.

Stieltjes' integrals

$$\int_a^b d\varphi(x) = m_0, \quad \int_a^b x d\varphi(x) = m_1, \quad \dots \quad \int_a^b x^i d\varphi(x) = m_i$$

represent respectively the whole mass  $m_0$  and its moments about the origin of the order 1, 2, . . .  $i$ . When the distribution function  $\varphi(x)$  is given, moments  $m_0, m_1, m_2, \dots, m_i$  (provided they exist) are determined. If, however, these moments are given and are known to originate in a certain distribution of a mass over  $(a, b)$ , the question may be raised with what error the mass spread over an interval  $(a, x)$  can be determined by these data? In other words, given  $m_0, m_1, m_2, \dots, m_i$ , what are the precise upper and lower bounds of a mass spread over an interval  $(a, x)$ ? Such is the question raised by Tshebysheff in a short but important article "Sur les valeurs limites des intégrales" (1874).<sup>1</sup> The results contained in this article, including very remarkable inequalities which indeed are of fundamental importance, are given without proof. The first proof of these results and the complete solution of the question raised by Tshebysheff was given by Markoff in his eminent thesis "On some applications of algebraic continued fractions" (St. Petersburg, 1884), written in Russian and therefore comparatively little known.

Suppose that  $\rho_i$  is the limit of the error with which we can evaluate the mass belonging to the interval  $(a, x)$  or, which is almost the same, the value of  $\varphi(x)$ , when moments  $m_0, m_1, m_2, \dots, m_i$  are given. If, with  $i$  tending to infinity,  $\rho_i$  tends to 0 for any given  $x$ , then the distribution function  $\varphi(x)$  will be completely determined by giving all the moments

$$m_0, m_1, m_2, \dots$$

One case of this kind, that in which

$$m_0 = 1, \quad m_{2k} = \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2k-1)}{2^k}, \quad m_{2k+1} = 0$$

<sup>1</sup> *Jour. Liouville*, Ser. 2, T. XIX, 1874.

was considered by Tshebysheff in a later paper, "Sur deux théorèmes relatifs aux probabilités" (1887)<sup>1</sup> devoted to the application of his method to the proof of the limit theorem under certain rather general conditions. The success of this proof is due to the fact that moments, as given above, uniquely determine the normal distribution

$$\varphi(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-u^2} du$$

of the mass 1 over the infinite interval  $(-\infty, +\infty)$ .

After these preliminary remarks and before proceeding to an orderly exposition of the method of moments, it is advisable to devote a few pages to continued fractions associated with power series, for continued fractions are the natural tools in questions of the kind we shall consider.

## 2. Continued Fractions Associated with Power Series. Let

$$\phi(z) = \frac{A_1}{z^{\alpha_1}} + \frac{A_2}{z^{\alpha_2}} + \frac{A_3}{z^{\alpha_3}} + \dots; \quad (A_1 \neq 0)$$

be a power series arranged according to decreasing powers of  $z$  where the smallest exponent  $\alpha_1$  is positive. We consider this power series from a purely formal point of view merely as a means to form a sequence of rational fractions

$$\frac{A_1}{z^{\alpha_1}}, \quad \frac{A_1}{z^{\alpha_1}} + \frac{A_2}{z^{\alpha_2}}, \quad \frac{A_1}{z^{\alpha_1}} + \frac{A_2}{z^{\alpha_2}} + \frac{A_3}{z^{\alpha_3}}, \dots$$

and we need not be concerned about its convergence.

Evidently  $1/\phi(z)$  can again be expanded into power series, arranged according to decreasing powers of  $z$ . Let its integral part, containing non-negative powers of  $z$ , be denoted by  $q_1(z)$ , and let the fractional part

$$\frac{B_1}{z^{\beta_1}} + \frac{B_2}{z^{\beta_2}} + \frac{B_3}{z^{\beta_3}} + \dots$$

containing negative powers of  $z$ , be denoted by  $-\phi_1(z)$ , so that

$$\frac{1}{\phi(z)} = q_1(z) - \phi_1(z).$$

In the same way

$$\frac{1}{\phi_1(z)}$$

can be represented thus:

$$\frac{1}{\phi_1(z)} = q_2(z) - \phi_2(z)$$

<sup>1</sup> Oeuvres complètes de P. L. Tshebysheff, Tome 2, p. 482.

where  $q_2(z)$  is a polynomial and

$$\phi_2(z) = \frac{C_1}{z^{\gamma_1}} + \frac{C_2}{z^{\gamma_2}} + \frac{C_3}{z^{\gamma_3}} + \dots,$$

the power series containing only negative powers of  $z$ . Further, we shall have

$$\frac{1}{\phi_2(z)} = q_3(z) - \phi_3(z)$$

with a certain polynomial  $q_3(z)$  and a power series

$$\phi_3(z) = \frac{D_1}{z^{\delta_1}} + \frac{D_2}{z^{\delta_2}} + \frac{D_3}{z^{\delta_3}} + \dots$$

containing negative powers of  $z$ , and so on. Thus we are led to consider a continued fraction (finite or infinite)

$$(1) \quad \frac{1}{q_1} - \frac{1}{q_2} - \frac{1}{q_3} - \dots$$

associated with  $\phi(z)$  in the sense that the formal expansion of

$$\frac{1}{q_1} - \frac{1}{q_2} - \dots - \frac{1}{q_i} - \phi_i(z)$$

into a power series will reproduce exactly  $\phi(z)$ . The continued fraction (1) is again considered from a purely formal standpoint as a mere abbreviation of the sequence of its convergents

$$\frac{P_1}{Q_1} = \frac{1}{q_1}; \quad \frac{P_2}{Q_2} = \frac{1}{q_1} - \frac{1}{q_2}; \quad \frac{P_3}{Q_3} = \frac{1}{q_1} - \frac{1}{q_2} - \frac{1}{q_3}; \dots$$

The polynomials

$$P_1, P_2, P_3, \dots \\ Q_1, Q_2, Q_3, \dots$$

can be found step by step by the recurrence relations

$$(2) \quad \left. \begin{aligned} P_i &= q_i P_{i-1} - P_{i-2} \\ Q_i &= q_i Q_{i-1} - Q_{i-2} \end{aligned} \right\} i = 2, 3, 4, \dots \\ P_1 = 1, \quad P_0 = 0 \\ Q_1 = q_1, \quad Q_0 = 1$$

from which the following identical relation follows:

$$(3) \quad P_i(z)Q_{i-1}(z) - Q_i(z)P_{i-1}(z) = 1,$$

showing that all fractions

$$\frac{P_i(z)}{Q_i(z)}$$

are irreducible. Evidently degrees of consecutive denominators of convergents form an increasing sequence and the degree of  $Q_i(z)$  is at least  $i$ . Since

$$\begin{aligned} \frac{1}{q_1} - \frac{1}{q_2} - \dots - \frac{1}{q_{i+1} - \phi_{i+1}(z)} &= \frac{P_i(q_{i+1} - \phi_{i+1}(z)) - P_{i-1}}{Q_i(q_{i+1} - \phi_{i+1}(z)) - Q_{i-1}} = \\ &= \frac{P_{i+1} - P_i\phi_{i+1}(z)}{Q_{i+1} - Q_i\phi_{i+1}(z)} \end{aligned}$$

we can write

$$\phi(z) = \frac{P_{i+1} - P_i\phi_{i+1}(z)}{Q_{i+1} - Q_i\phi_{i+1}(z)}$$

in the sense that the formal development of the right-hand member is identical with  $\phi(z)$ . By virtue of relation (3)

$$\phi(z) - \frac{P_i}{Q_i} = \frac{1}{Q_i(Q_{i+1} - Q_i\phi_{i+1})}$$

The degree of  $Q_i$  being  $\lambda_i$  and that of  $Q_{i+1}$  being  $\lambda_{i+1}$ , the expansion of

$$Q_i(Q_{i+1} - Q_i\phi_{i+1})$$

in a series of descending powers of  $z$  begins with the power  $z^{\lambda_i + \lambda_{i+1}}$ . Hence,

$$\phi(z) - \frac{P_i}{Q_i} = \frac{M}{z^{\lambda_i + \lambda_{i+1}}} + \dots$$

and, since  $\lambda_{i+1} \geq \lambda_i + 1$ , the expansion of

$$\phi(z) - \frac{P_i}{Q_i}$$

begins with a term of the order  $2\lambda_i + 1$  in  $1/z$  at least. This property characterizes the convergents  $P_i/Q_i$  completely. For let  $P/Q$  be a rational fraction whose denominator is of the  $n$ th degree and such that in the expansion of

$$\phi(z) - \frac{P}{Q}$$

the lowest term is of the order  $2n + 1$  in  $1/z$  at least. Then  $P/Q$  coincides with one of the convergents to the continued fraction (1). Let  $i$  be determined by the condition

$$\lambda_i \leq n < \lambda_{i+1}.$$

Then

$$\begin{aligned} \phi(z) - \frac{P_i}{Q_i} &= \frac{M}{z^{\lambda_i + \lambda_{i+1}}} + \dots \\ \phi(z) - \frac{P}{Q} &= \frac{N}{z^{2n+1}} + \dots \end{aligned}$$

whence in the expansion of

$$\frac{P}{Q} - \frac{P_i}{Q_i}$$

the lowest term will be of degree  $2n + 1$  or  $\lambda_i + \lambda_{i+1}$  in  $1/z$ . Hence, the degree of

$$PQ_i - P_iQ$$

in  $z$  is not greater than both the numbers

$$\lambda_i - n - 1 \quad \text{and} \quad n - \lambda_{i+1}$$

which are both negative while

$$PQ_i - P_iQ$$

is a polynomial. Hence, identically,

$$PQ_i - P_iQ = 0$$

or

$$\frac{P}{Q} = \frac{P_i}{Q_i}$$

which proves the statement.

**3. Continued Fraction Associated with  $\int_a^b \frac{d\varphi(x)}{z-x}$ .** Let  $\varphi(x)$  be a never decreasing function characterizing the distribution of a mass over an interval  $(a, b)$ . The moments of this distribution up to the moment of the order  $2n$  are represented by integrals

$$\begin{aligned} m_0 &= \int_a^b d\varphi(x), & m_1 &= \int_a^b x d\varphi(x), \\ & & m_2 &= \int_a^b x^2 d\varphi(x), \dots & m_{2n} &= \int_a^b x^{2n} d\varphi(x). \end{aligned}$$

Let

$$\Delta_0 = m_0; \quad \Delta_1 = \begin{vmatrix} m_0 m_1 \\ m_1 m_2 \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} m_0 m_1 m_2 \\ m_1 m_2 m_3 \\ m_2 m_3 m_4 \end{vmatrix}; \quad \dots \quad \Delta_n = \begin{vmatrix} m_0 m_1 \dots m_n \\ m_1 m_2 \dots m_{n+1} \\ \dots \dots \dots \\ m_n m_{n+1} \dots m_{2n} \end{vmatrix}.$$

If  $\varphi(x)$  has not less than  $n + 1$  points of increase, we must have

$$\Delta_0 > 0, \quad \Delta_1 > 0, \quad \dots \quad \Delta_n > 0,$$

and conversely, if these inequalities are satisfied,  $\varphi(x)$  has at least  $n + 1$  points of increase. To prove this, consider the quadratic form

$$\phi = \int_a^b (t_0 + t_1 x + \dots + t_n x^n)^2 d\varphi(x)$$

in  $n + 1$  variables  $t_0, t_1, \dots, t_n$ . Evidently

$$\phi = \sum m_{ij} t_i t_j \quad (i, j = 0, 1, 2, \dots, n)$$

so that  $\Delta_n$  is the determinant of  $\phi$  and  $\Delta_0, \Delta_1, \dots, \Delta_{n-1}$  its principal minors. The form  $\phi$  cannot vanish unless  $t_0 = t_1 = \dots = t_n = 0$ . For if  $x = \xi$  is a point of increase and  $\phi = 0$ , we must have also

$$\int_{\xi-\epsilon}^{\xi+\epsilon} (t_0 + t_1 x + \dots + t_n x^n)^2 d\varphi(x) = 0$$

for an arbitrary positive  $\epsilon$ , whence by the mean value theorem

$$(t_0 + t_1 \eta + \dots + t_n \eta^n)^2 \int_{\xi-\epsilon}^{\xi+\epsilon} d\varphi(x) = 0 \quad (\xi - \epsilon < \eta < \xi + \epsilon)$$

or

$$t_0 + t_1 \eta + \dots + t_n \eta^n = 0$$

because

$$\int_{\xi-\epsilon}^{\xi+\epsilon} d\varphi(x) > 0.$$

Letting  $\epsilon$  converge to 0, we conclude

$$t_0 + t_1 \xi + \dots + t_n \xi^n = 0$$

at any point of increase. Since there are at least  $n + 1$  points of increase the equation

$$t_0 + t_1 x + \dots + t_n x^n = 0$$

would have at least  $n + 1$  roots and that necessitates

$$t_0 = t_1 = \dots = t_n = 0.$$



Hence, the quadratic form  $\phi$ , which is never negative, can vanish only if all its variables vanish; that is,  $\phi$  is a definite positive form. Its determinant  $\Delta_n$  and all its principal minors  $\Delta_{n-1}, \Delta_{n-2}, \dots, \Delta_0$  must be positive, which proves the first statement.

Suppose the conditions

$$\Delta_0 > 0, \quad \Delta_1 > 0, \quad \dots \quad \Delta_n > 0$$

satisfied and let  $\varphi(x)$  have  $s < n + 1$  points of increase. Then the integral representing  $\phi$  reduces to a finite sum

$$\phi = p_1(t_0 + t_1\xi_1 + \dots + t_n\xi_1^n)^2 + p_2(t_0 + t_1\xi_2 + \dots + t_n\xi_2^n)^2 + \dots + p_s(t_0 + t_1\xi_s + \dots + t_n\xi_s^n)^2$$

denoting by  $p_1, p_2, \dots, p_s$  masses concentrated in the  $s$  points of increase  $\xi_1, \xi_2, \dots, \xi_s$ . Now, since  $s \leq n$  constants  $t_0, t_1, \dots, t_n$ , not all zero, can be determined by the system of equations

$$\begin{aligned} t_0 + t_1\xi_1 + \dots + t_n\xi_1^n &= 0 \\ t_0 + t_1\xi_2 + \dots + t_n\xi_2^n &= 0 \\ &\dots \dots \dots \\ t_0 + t_1\xi_s + \dots + t_n\xi_s^n &= 0. \end{aligned}$$

Thus  $\phi$  vanishes when not all variables vanish; hence, its determinant  $\Delta_n = 0$ , contrary to hypothesis.

From now on we shall assume that  $\varphi(x)$  has at least  $n + 1$  points of increase. The integral

$$\int_a^b \frac{d\varphi(x)}{z - x}$$

can be expanded into a formal power series of  $1/z$ , thus

$$\int_a^b \frac{d\varphi(x)}{z - x} = \frac{m_0}{z} + \frac{m_1}{z^2} + \frac{m_2}{z^3} + \dots + \frac{m_{2n}}{z^{2n+1}} + \dots$$

and this power series can be converted into a continued fraction as explained in Sec. 2. Let

$$\frac{P_1}{Q_1}, \frac{P_2}{Q_2}, \dots, \frac{P_n}{Q_n}, \frac{P_{n+1}}{Q_{n+1}}$$

be the first  $n + 1$  convergents to that continued fraction. I say that the degrees of their denominators are, respectively, 1, 2, 3,  $\dots, n + 1$ . Since these degrees form an increasing sequence, it suffices to show that there exists a convergent with the denominator of a given degree

$$s \leq n + 1.$$

This convergent  $P/Q$  is completely determined by the condition that in a formal expansion of the difference

$$\int_a^b \frac{d\varphi(x)}{z-x} - \frac{P}{Q}$$

into a power series of  $1/z$ , terms involving  $1/z, 1/z^2, \dots, 1/z^{2s}$  are absent. This is the same as to say that in the expansion of

$$Q(z) \int_a^b \frac{d\varphi(x)}{z-x} - P(z)$$

there are no terms involving  $1/z, 1/z^2, \dots, 1/z^s$ . The preceding expression can be written thus:

$$\int_a^b \frac{Q(x)d\varphi(x)}{z-x} + \int_a^b \frac{Q(z) - Q(x)}{z-x} d\varphi(x) - P(z) = \frac{A}{z^{s+1}} + \dots$$

Since

$$\int_a^b \frac{Q(z) - Q(x)}{z-x} d\varphi(x) - P(z)$$

is a polynomial in  $z$ , it must vanish identically. That gives

$$(4) \quad P(z) = \int_a^b \frac{Q(z) - Q(x)}{z-x} d\varphi(x).$$

To determine  $Q(z)$  we must express the conditions that in the expansion of

$$\int_a^b \frac{Q(x)d\varphi(x)}{z-x}$$

terms in  $1/z, 1/z^2, \dots, 1/z^s$  vanish. These conditions are equivalent to  $s$  relations

$$(5) \quad \int_a^b Q(x)d\varphi(x) = 0, \quad \int_a^b xQ(x)d\varphi(x) = 0, \quad \dots \quad \int_a^b x^{s-1}Q(x)d\varphi(x) = 0,$$

which in turn amount to the single requirement that

$$(6) \quad \int_a^b \theta(x)Q(x)d\varphi(x) = 0$$

for an arbitrary polynomial  $\theta(x)$  of degree  $\leq s-1$ .

Conversely, if there exists a polynomial  $Q(z)$  of degree  $s$  satisfying conditions (5), and  $P(z)$  is determined by equation (4), then  $P(z)/Q(z)$  is a convergent whose denominator is of degree  $s$ . For then the expansion of

$$\int_a^b \frac{d\varphi(x)}{z-x} - \frac{P(z)}{Q(z)}$$

lacks the terms in  $1/z, 1/z^2, \dots, 1/z^{2s}$ .

Let

$$Q(z) = l_0 + l_1z + l_2z^2 + \dots + l_{s-1}z^{s-1} + z^s.$$

Then equations (5) become

$$\begin{aligned} m_0l_0 + m_1l_1 + m_2l_2 + \dots + m_{s-1}l_{s-1} + m_s &= 0 \\ m_1l_0 + m_2l_1 + m_3l_2 + \dots + m_sl_{s-1} + m_{s+1} &= 0 \\ &\dots \dots \dots \\ m_{s-1}l_0 + m_sl_1 + m_{s+1}l_2 + \dots + m_{2s-2}l_{s-1} + m_{2s-1} &= 0. \end{aligned}$$

This system of linear equations determines completely the coefficients  $l_0, l_1, \dots, l_{s-1}$  since its determinant  $\Delta_{s-1} > 0$ .

The existence of a convergent with the denominator of degree

$$s \leq n + 1$$

being established, it follows that the denominator of the  $s$ th convergent  $P_s/Q_s$  is exactly of degree  $s$ . The denominator  $Q_s$  is determined, except for a constant factor, and can be presented in the form:

$$Q_s = \frac{c}{\Delta_{s-1}} \begin{vmatrix} 1 & z & z^2 & \dots & z^s \\ m_0 & m_1 & m_2 & \dots & m_s \\ m_1 & m_2 & m_3 & \dots & m_{s+1} \\ & & & \dots & \dots \\ m_{s-1} & m_s & m_{s+1} & \dots & m_{2s-1} \end{vmatrix}$$

A remarkable result follows from equation (6) by taking  $Q = Q_s$  and  $\theta = Q_{s'}$ ; namely,

$$(7) \quad \int_a^b Q_s Q_{s'} d\varphi(x) = 0 \quad \text{if} \quad s \neq s'$$

while

$$\int_a^b Q_s^2 d\varphi(x) > 0 \quad (s \leq n).$$

In the general relation

$$Q_s = q_s Q_{s-1} - Q_{s-2}$$

the polynomial  $q_s$  must be of the first degree

$$q_s = \alpha_s z + \beta_s,$$

which shows that the continued fraction associated with

$$\int_a^b \frac{d\varphi(x)}{z - x}$$

has the form

$$\frac{1}{\alpha_1 z + \beta_1} - \frac{1}{\alpha_2 z + \beta_2} - \frac{1}{\alpha_3 z + \beta_3} - \dots$$

The next question is, how to determine the constants  $\alpha_s$  and  $\beta_s$ . Multiplying both members of the equation

$$Q_s = (\alpha_s z + \beta_s) Q_{s-1} - Q_{s-2} \quad (s \geq 2)$$

by  $Q_{s-2} d\varphi(z)$ , integrating between limits  $a$  and  $b$ , and taking into account (7), we get

$$0 = \alpha_s \int_a^b z Q_{s-1} Q_{s-2} d\varphi(z) - \int_a^b Q_{s-2}^2 d\varphi(z).$$

On the other hand, the highest terms in  $Q_{s-1}$  and  $Q_{s-2}$  are

$$\alpha_1 \alpha_2 \dots \alpha_{s-1} z^{s-1}, \quad \alpha_1 \alpha_2 \dots \alpha_{s-2} z^{s-2}.$$

Hence,

$$z Q_{s-2} = \frac{1}{\alpha_{s-1}} Q_{s-1} + \psi$$

where  $\psi$  is a polynomial of degree  $\leq s - 2$ . Referring to equation (6), we have

$$\int_a^b z Q_{s-2} Q_{s-1} d\varphi(z) = \frac{1}{\alpha_{s-1}} \int_a^b Q_{s-1}^2 d\varphi(z)$$

and consequently

$$(8) \quad \frac{\alpha_s}{\alpha_{s-1}} = \frac{\int_a^b Q_{s-2}^2 d\varphi(z)}{\int_a^b Q_{s-1}^2 d\varphi(z)}.$$

Suppose that the following moments are given:  $m_0, m_1, \dots, m_{2n}$ ; how many of the coefficients  $\alpha_s$  can be found? Evidently  $\alpha_1 = 1/m_0$ . Furthermore,  $Q_0 = 1$  and  $Q_1$  is completely determined given  $m_0$  and  $m_1$ . Relation (8) determines  $\alpha_2$ , and  $Q_2$  will be completely determined given  $m_0, m_1, m_2, m_3$ . The same relation again determines  $\alpha_3$ , and  $Q_3$  will be determined given  $m_0, m_1, \dots, m_5$ . Proceeding in the same way, we conclude that, given  $m_0, m_1, m_2, \dots, m_{2n}$ , all the polynomials

$$Q_0, Q_1, Q_2, \dots, Q_n$$

as well as constants

$$\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_{n+1}$$

can be determined. It is important to note that all these constants are positive.

Proceeding in a similar manner, the following expression can be found

$$\beta_s = -\alpha_s \frac{\int_a^b z Q_{s-1}^2 d\varphi(z)}{\int_a^b Q_{s-1}^2 d\varphi(z)}.$$

It follows that constants

$$\beta_1, \beta_2, \dots, \beta_n$$

are determined by our data, but not  $\beta_{n+1}$ . For if  $s = n + 1$ , the integral

$$\int_a^b z Q_n^2 d\varphi(z)$$

can be expressed as a linear function of  $m_0, m_1, \dots, m_{2n+1}$  with known coefficients. But  $m_{2n+1}$  is not included among our data; hence,  $\beta_{n+1}$  cannot be determined.

**4. Properties of Polynomials  $Q_s$ . Theorem.** *Roots of the equation*

$$Q_s(z) = 0 \quad (s \leq n)$$

are real, simple, and contained within the interval  $(a, b)$ .

**Proof.** Let  $Q_s(z)$  change its sign  $r < s$  times when  $z$  passes through points  $z_1, z_2, \dots, z_r$ , contained strictly within  $(a, b)$ . Setting

$$\theta(z) = (z - z_1)(z - z_2) \cdots (z - z_r)$$

the product

$$\theta(z)Q_s(z)$$

does not change its sign when  $z$  increases from  $a$  to  $b$ . However,

$$\int_a^b \theta(z)Q_s(z)d\varphi(z) = 0,$$

and this necessitates that

$$\theta(z)Q_s(z)$$

or  $Q_s(z)$  vanishes in all points of increase of  $\varphi(z)$ . But this is impossible, since by hypothesis there are at least  $n + 1$  points of increase, whereas the degree  $s$  of  $Q_s$  does not exceed  $n$ . Consequently,  $Q_s(z)$  changes its sign in the interval  $(a, b)$  exactly  $s$  times and has all its roots real, simple, and located within  $(a, b)$ .

It follows from this theorem that the convergent

$$\frac{P_n}{Q_n}$$

can be resolved into a sum of simple fractions as follows:

$$(9) \quad \frac{P_n(z)}{Q_n(z)} = \frac{A_1}{z - z_1} + \frac{A_2}{z - z_2} + \cdots + \frac{A_n}{z - z_n}$$

where  $z_1, z_2, \dots, z_n$  are roots of the equation  $Q_n(z) = 0$  and in general

$$A_k = \frac{P_n(z_k)}{Q'_n(z_k)}.$$

The right member of (9) can be expanded into power series of  $1/z$ , the coefficient of  $1/z^k$  being

$$\sum_{\alpha=1}^n A_\alpha z_\alpha^{k-1}.$$

By the property of convergents we must have the following equations:

$$\begin{aligned} \sum_{\alpha=1}^n A_\alpha &= m_0 \\ \sum_{\alpha=1}^n A_\alpha z_\alpha &= m_1 \\ &\dots \dots \dots \\ \sum_{\alpha=1}^n A_\alpha z_\alpha^{2n-1} &= m_{2n-1}. \end{aligned}$$

These equations can be condensed into one,

$$(10) \quad \sum_{\alpha=1}^n A_\alpha T(z_\alpha) = \int_a^b T(z) d\varphi(z)$$

which should hold for any polynomial  $T(z)$  of degree  $\leq 2n - 1$ .

Let us take for  $T(z)$  a polynomial of degree  $2n - 2$ :

$$T(z) = \left[ \frac{Q_n(z)}{(z - z_\alpha) Q'_n(z_\alpha)} \right]^2.$$

Then

$$T(z_\alpha) = 1, \quad T(z_\beta) = 0 \quad \text{if} \quad \beta \neq \alpha$$

and consequently, by virtue of equation (10),

$$A_\alpha = \int_a^b \left[ \frac{Q_n(z)}{(z - z_\alpha) Q'_n(z_\alpha)} \right]^2 d\varphi(z) > 0.$$

Thus constants  $A_1, A_2, \dots, A_n$  are all positive, which shows that  $P_n(z_k)$

has the same sign as  $Q'_n(z_k)$ . Now in the sequence

$$Q'_n(z_1), Q'_n(z_2), \dots, Q'_n(z_n)$$

any two consecutive terms are of opposite signs. The same being true of the sequence

$$P_n(z_1), P_n(z_2), \dots, P_n(z_n),$$

it follows that the roots of  $P_n(z)$  are all simple, real, and located in the intervals

$$(z_1, z_2); (z_2, z_3); \dots (z_{n-1}, z_n).$$

Finally, we shall prove the following theorem:

**Theorem.** For any real  $x$

$$Q'_n(x)Q_{n-1}(x) - Q'_{n-1}(x)Q_n(x)$$

is a positive number.

**Proof.** From the relations

$$\begin{aligned} Q_s(z) &= (\alpha_s z + \beta_s)Q_{s-1}(z) - Q_{s-2}(z) \\ Q_s(x) &= (\alpha_s x + \beta_s)Q_{s-1}(x) - Q_{s-2}(x) \end{aligned}$$

it follows that

$$\begin{aligned} \frac{Q_s(z)Q_{s-1}(x) - Q_s(x)Q_{s-1}(z)}{z-x} &= \alpha_s Q_{s-1}(z)Q_{s-1}(x) + \\ &+ \frac{Q_{s-1}(z)Q_{s-2}(x) - Q_{s-1}(x)Q_{s-2}(z)}{z-x} \end{aligned}$$

whence, taking  $s = 1, 2, 3, \dots, n$  and adding results,

$$\frac{Q_n(z)Q_{n-1}(x) - Q_n(x)Q_{n-1}(z)}{z-x} = \sum_{s=1}^n \alpha_s Q_{s-1}(x)Q_{s-1}(z).$$

It suffices now to take  $z = x$  to arrive at the identity

$$Q'_n(x)Q_{n-1}(x) - Q'_{n-1}(x)Q_n(x) = \sum_{s=1}^n \alpha_s Q_{s-1}(x)^2.$$

Since  $Q_0 = 1$  and  $\alpha_s > 0$ , it is evident that

$$Q'_n(x)Q_{n-1}(x) - Q'_{n-1}(x)Q_n(x) > 0$$

for every real  $x$ .

**5. Equivalent Point Distributions.** If the whole mass can be concentrated in a finite number of points so as to produce the same  $l$  first moments as a given distribution, we have an "equivalent point distribu-

tion" in respect to the  $l$  first moments. In what follows we shall suppose that the whole mass is spread over an infinite interval  $-\infty, \infty$  and that the given moments, originating in a distribution with at least  $n + 1$  points of increase, are

$$m_0, m_1, m_2, \dots, m_{2n}.$$

The question is: Is it possible to find an equivalent point distribution where the whole mass is concentrated in  $n + 1$  points? Let the unknown points be

$$\xi_1, \xi_2, \dots, \xi_{n+1}$$

and the masses concentrated in them

$$A_1, A_2, \dots, A_{n+1}.$$

Evidently the question will be answered in the affirmative if the system of  $2n + 1$  equations

$$(A) \quad \begin{aligned} \sum_{\alpha=1}^{n+1} A_{\alpha} &= m_0 \\ \sum_{\alpha=1}^{n+1} A_{\alpha} \xi_{\alpha} &= m_1 \\ \sum_{\alpha=1}^{n+1} A_{\alpha} \xi_{\alpha}^2 &= m_2 \\ &\dots \dots \dots \\ \sum_{\alpha=1}^{n+1} A_{\alpha} \xi_{\alpha}^{2n} &= m_{2n} \end{aligned}$$

can be satisfied by real numbers  $\xi_1, \xi_2, \dots, \xi_{n+1}$ ;  $A_1, A_2, \dots, A_{n+1}$ , the last  $n + 1$  numbers being *positive*. The number of unknowns being greater by one unit than the number of equations, we can introduce the additional requirement that one of the numbers  $\xi_1, \xi_2, \dots, \xi_{n+1}$  should be equal to a given real number  $v$ . The system (A) may be replaced by the single requirement that the equation

$$(11) \quad \sum_{\alpha=1}^{n+1} A_{\alpha} T(\xi_{\alpha}) = \int_{-\infty}^{\infty} T(x) d\varphi(x)$$

shall hold for any polynomial  $T(x)$  of degree  $\leq 2n$ . Let  $Q(x)$  be the polynomial of degree  $n + 1$  having roots  $\xi_1, \xi_2, \dots, \xi_{n+1}$  and let  $\theta(x)$  be an arbitrary polynomial of degree  $\leq n - 1$ . Then we can apply equation (11) to

$$T(x) = \theta(x)Q(x).$$



Since  $Q(\xi_\alpha) = 0$ , we shall have

$$(12) \quad \int_{-\infty}^{\infty} \theta(x)Q(x)d\varphi(x) = 0$$

for an arbitrary polynomial  $\theta(x)$  of degree  $\leq n - 1$ . Presently we shall see that requirement (12) together with  $Q(v) = 0$  determines  $Q(x)$ , save for a constant factor if

$$Q_n(v) \neq 0.$$

Dividing  $Q(x)$  by  $Q_n(x)$ , we have identically

$$Q(x) = (\lambda x + \mu)Q_n(x) + R_{n-1}(x)$$

where  $R_{n-1}(x)$  is a polynomial of degree  $\leq n - 1$ . If  $\theta(x)$  is an arbitrary polynomial of degree  $\leq n - 2$ ,

$$(\lambda x + \mu)\theta(x)$$

will be of degree  $\leq n - 1$ . Hence

$$\int_a^b (\lambda x + \mu)\theta(x)Q_n(x)d\varphi(x) = 0$$

by (6), and (12) shows that

$$\int_a^b \theta(x)R_{n-1}(x)d\varphi(x) = 0$$

for an arbitrary polynomial  $\theta(x)$  of degree  $\leq n - 2$ . The last requirement shows that  $R_{n-1}(x)$  differs from  $Q_{n-1}(x)$  by a constant factor. Since the highest coefficient in  $Q(x)$  is arbitrary, we can set

$$R_{n-1}(x) = -Q_{n-1}(x).$$

In the equation

$$Q(x) = (\lambda x + \mu)Q_n(x) - Q_{n-1}(x)$$

it remains to determine constants  $\lambda$  and  $\mu$ . Multiplying both members by  $Q_{n-1}(x)d\varphi(x)$  and integrating between  $-\infty$  and  $\infty$ , we get

$$\lambda \int_{-\infty}^{\infty} xQ_{n-1}Q_n d\varphi(x) = \int_{-\infty}^{\infty} Q_n^2 d\varphi(x)$$

or

$$\frac{\lambda}{\alpha_n} \int_{-\infty}^{\infty} Q_n^2 d\varphi(x) = \int_{-\infty}^{\infty} Q_{n-1}^2 d\varphi(x).$$

But

$$\frac{\int_{-\infty}^{\infty} Q_{n-1}^2 d\varphi(x)}{\int_{-\infty}^{\infty} Q_n^2 d\varphi(x)} = \frac{\alpha_{n+1}}{\alpha_n}$$

whence

$$\lambda = \alpha_{n+1}.$$

The equation

$$0 = Q(v) = (\alpha_{n+1}v + \mu)Q_n(v) - Q_{n-1}(v)$$

serves to determine  $\mu$  if  $Q_n(v) \neq 0$ . The final expression of  $Q(x)$  will be

$$Q(x) = \left( \alpha_{n+1}(x - v) + \frac{Q_{n-1}(v)}{Q_n(v)} \right) Q_n(x) - Q_{n-1}(x).$$

Owing to recurrence relations

$$Q_2 = (\alpha_2x + \beta_2)Q_1 - Q_0; \quad Q_3 = (\alpha_3x + \beta_3)Q_2 - Q_1; \quad \dots$$

$$Q_n = (\alpha_nx + \beta_n)Q_{n-1} - Q_{n-2},$$

it is evident that

$$Q, Q_n, Q_{n-1}, \dots, Q_1, Q_0 = 1$$

in a Sturm series. For  $x = -\infty$ , it contains  $n + 1$  variations and for  $x = \infty$  only permanences. It follows that the equation

$$Q(x) = 0$$

has exactly  $n + 1$  distinct real roots and among them  $v$ . Thus, if the problem is solvable, the numbers  $\xi_1, \xi_2, \dots, \xi_{n+1}$  are determined as roots of

$$Q(x) = 0.$$

Furthermore, all unknowns  $A_\alpha$  will be positive. In fact, from equation (11) it follows that

$$A_\alpha = \int_{-\infty}^{\infty} \left[ \frac{Q(x)}{(x - \xi_\alpha)Q'(\xi_\alpha)} \right]^2 d\varphi(x) > 0.$$

Now we must show that constants  $A_\alpha$  can actually be determined so as to satisfy equations (A). To this end let

$$P(x) = \int_{-\infty}^{\infty} \frac{Q(x) - Q(z)}{x - z} d\varphi(z) = \left[ \alpha_{n+1}(x - v) + \frac{Q_{n-1}(v)}{Q_n(v)} \right] P_n(x) - P_{n-1}(x).$$

Then

$$Q(x) \int_{-\infty}^{\infty} \frac{d\varphi(z)}{x - z} - P(x) = \int_{-\infty}^{\infty} \frac{Q(z)d\varphi(z)}{x - z}$$

and, on account of (12), the expansion of the right member into power series of  $1/x$  lacks the terms in  $1/x, 1/x^2, \dots, 1/x^n$ . Hence, the expansion of

$$\int_{-\infty}^{\infty} \frac{d\varphi(z)}{x - z} - \frac{P(x)}{Q(x)}$$

lacks the terms in  $1/x, 1/x^2, \dots, 1/x^{2n+1}$ ; that is,

$$\frac{P(x)}{Q(x)} = \frac{m_0}{x} + \frac{m_1}{x^2} + \dots + \frac{m_{2n}}{x^{2n+1}} + \dots$$

On the other hand, resolving in simple fractions,

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - \xi_1} + \frac{A_2}{x - \xi_2} + \dots + \frac{A_{n+1}}{x - \xi_{n+1}}.$$

Expanding the right member into power series of  $1/x$  and comparing with the preceding expansion, we obtain the system (A). By the previous remark all constants  $A_\alpha$  are positive. Thus, there exists a point distribution in which masses concentrated in  $n + 1$  points produce moments  $m_0, m_1, \dots, m_{2n}$ . One of these points  $v$  may be taken arbitrarily, with the condition

$$Q_n(v) \neq 0$$

being observed, however.

**6. Tshebysheff's Inequalities.** In a note referred to in the introduction Tshebysheff made known certain inequalities of the utmost importance for the theory we are concerned with. The first very ingenious proof of them was given by Markoff in 1884 and, by a remarkable coincidence, the same proof was rediscovered almost at the same time by Stieltjes. A few years later, Stieltjes found another totally different proof; and it is this second proof that we shall follow.

Let  $\varphi(x)$  be a distribution function of a mass spread over the interval  $-\infty, \infty$ . Supposing that a moment of the order  $i$ ,

$$\int_{-\infty}^{\infty} x^i d\varphi(x) = m_i,$$

exists, we shall show first that

$$\begin{aligned} \lim l^i(m_0 - \varphi(l)) &= 0 \\ \lim l^i\varphi(-l) &= 0 \end{aligned}$$

when  $l$  tends to  $+\infty$ . For

$$\int_l^{\infty} x^i d\varphi(x) \geq l^i \int_l^{\infty} d\varphi(x) = l^i[\varphi(+\infty) - \varphi(l)]$$

or

$$l^i(m_0 - \varphi(l)) \leq \int_l^{\infty} x^i d\varphi(x).$$

Similarly

$$\left| \int_{-\infty}^{-l} x^i d\varphi(x) \right| \geq l^i \int_{-\infty}^{-l} d\varphi(x) = l^i\varphi(-l)$$

or

$$l^i \varphi(-l) \leq \left| \int_{-\infty}^{-l} x^i d\varphi(x) \right|.$$

Now both integrals

$$\int_l^{\infty} x^i d\varphi(x) \quad \text{and} \quad \int_{-\infty}^{-l} x^i d\varphi(x)$$

converge to 0 as  $l$  tends to  $+\infty$ ; whence both statements follow immediately. Integrating by parts, we have

$$\begin{aligned} \int_0^l x^i d\varphi(x) &= l^i [\varphi(l) - m_0] - i \int_0^l [\varphi(x) - m_0] x^{i-1} dx \\ \int_{-l}^0 x^i d\varphi(x) &= (-1)^{i-1} l^i \varphi(-l) - i \int_{-l}^0 x^{i-1} \varphi(x) dx, \end{aligned}$$

whence, letting  $l$  converge to  $+\infty$ ,

$$m_i = \int_{-\infty}^{\infty} x^i d\varphi(x) = -i \int_0^{\infty} [\varphi(x) - m_0] x^{i-1} dx - i \int_{-\infty}^0 x^{i-1} \varphi(x) dx.$$

If the same mass  $m_0$ , with the same moment  $m_i$ , is spread according to the law characterized by the function  $\psi(x)$ , we shall have

$$m_i = \int_{-\infty}^{\infty} x^i d\psi(x) = -i \int_0^{\infty} [\psi(x) - m_0] x^{i-1} dx - i \int_{-\infty}^0 x^{i-1} \psi(x) dx,$$

whence

$$(13) \quad \int_{-\infty}^{\infty} x^{i-1} [\varphi(x) - \psi(x)] dx = 0.$$

Suppose the moments

$$m_0, m_1, m_2, \dots, m_{2n}$$

of the distribution characterized by  $\varphi(x)$  are known. Provided  $\varphi(x)$  has at least  $n + 1$  points of increase, there exists an equivalent point distribution, defined in Sec. 5 and characterized by the step function  $\psi(x)$  which can be defined as follows:

$$\begin{aligned} \psi(x) &= 0 & \text{for} & \quad -\infty < x < \xi_1 \\ \psi(x) &= A_1 & \text{for} & \quad \xi_1 \leq x < \xi_2 \\ \psi(x) &= A_1 + A_2 & \text{for} & \quad \xi_2 \leq x < \xi_3 \\ & \dots & & \dots \\ \psi(x) &= A_1 + A_2 + \dots + A_n & \text{for} & \quad \xi_n \leq x < \xi_{n+1} \\ \psi(x) &= A_1 + A_2 + \dots + A_{n+1} & \text{for} & \quad \xi_{n+1} \leq x < +\infty, \end{aligned}$$

provided roots  $\xi_1, \xi_2, \dots, \xi_{n+1}$  of the equation  $Q(x) = 0$  are arranged in an increasing order of magnitude.

Equation (13) will hold for  $i = 1, 2, 3, \dots, 2n$  or, which is the same, the equation

$$(14) \quad \int_{-\infty}^{\infty} \theta(x)[\varphi(x) - \psi(x)]dx = 0$$

will hold for an arbitrary polynomial  $\theta(x)$  of degree  $\leq 2n - 1$ . The function

$$h(x) = \varphi(x) - \psi(x)$$

in general has ordinary discontinuities. We can prove now that  $h(x)$ , if not identically equal to 0 at all points of continuity, changes its sign at least  $2n$  times.<sup>1</sup> Suppose, on the contrary, that it changes sign  $r < 2n$  times; namely, at the points

$$a_1, a_2, \dots, a_r.$$

Taking

$$\theta(x) = (x - a_1)(x - a_2) \cdots (x - a_r),$$

equation (14) will be satisfied, while the integrand

$$\theta(x)h(x),$$

if not 0, will be of the same sign, for example, positive. Let  $\xi$  be any point of continuity of  $h(x)$ . If  $\xi = a_i$  ( $i = 1, 2, \dots, r$ ) then  $h(a_i) = 0$  since  $h(x)$  changes sign at  $a_i$ . If  $\xi$  does not coincide with any one of the numbers  $a_1, a_2, \dots, a_r$ , then for an arbitrarily small positive  $\epsilon$  we must have

$$\int_{\xi-\epsilon}^{\xi+\epsilon} \theta(x)h(x)dx = 0.$$

But by continuity

$$\theta(x)h(x)$$

remains in the interval  $(\xi - \epsilon, \xi + \epsilon)$  for sufficiently small  $\epsilon$  above a certain positive number unless  $h(\xi) = 0$ . Thus, if  $h(x)$  does not vanish at all points of continuity (in which case  $\varphi(x)$  and  $\psi(x)$  do not differ essentially), it must change sign at least  $2n$  times. Let us see now where the change of sign can occur. In the intervals

$$-\infty, \xi_1 \text{ and } \xi_{r+1}, +\infty$$

<sup>1</sup>A function  $f(x)$  is said to change sign once in  $(a, b)$  if in this interval there exists a point or points  $c$  such that, for instance,  $f(x) \geq 0$  in  $(a, c)$  and  $f(x) \leq 0$  in  $(c, b)$ , equality signs not holding throughout the respective intervals. The change of sign occurs  $n$  times if  $(a, b)$  can be divided in  $n$  intervals in which  $f(x)$  changes sign once.

$\varphi(x) - \psi(x)$  evidently cannot change sign. Within each of the intervals

$$\xi_{i-1}, \xi_i$$

there can be at most one change of sign, since  $\psi(x)$  remains constant there, and  $\varphi(x)$  can only increase. The sign may change also at the points of discontinuity of  $\psi(x)$ ; that is, at the points  $\xi_1, \xi_2, \dots, \xi_{n+1}$ . Altogether,  $\varphi(x) - \psi(x)$  cannot change sign more than  $2n + 1$  times and not less than  $2n$  times.

Since  $\psi(x) = 0$  so far as  $x < \xi_1$  and  $\varphi(\xi_1 - \epsilon)$  is not negative for positive  $\epsilon$ , we must have

$$\varphi(\xi_1 - \epsilon) - \psi(\xi_1 - \epsilon) \geq 0.$$

Also  $\psi(x) = m_0$  for  $x > \xi_{n+1}$  and  $\varphi(x) \leq m_0$ , so that

$$\varphi(\xi_{n+1} + \epsilon) - \psi(\xi_{n+1} + \epsilon) \leq 0.$$

At first let us suppose

$$\varphi(\xi_1 - \epsilon) - \psi(\xi_1 - \epsilon) > 0, \quad \varphi(\xi_{n+1} + \epsilon) - \psi(\xi_{n+1} + \epsilon) < 0.$$

In this case  $\varphi(x) - \psi(x)$  must change sign an *odd* number of times; that is, not less than  $2n + 1$  times. Since this cannot happen more than  $2n + 1$  times, the number of times  $\varphi(x) - \psi(x)$  changes its sign must be exactly  $2n + 1$ . These changes occur once within each interval

$$\xi_{i-1}, \xi_i$$

and in each of the points  $\xi_1, \xi_2, \dots, \xi_{n+1}$ . When the change of sign occurs in the interval  $(\xi_{i-1}, \xi_i)$  where  $\psi(x)$  remains constant, because  $\varphi(x)$  never decreases, we must have for sufficiently small  $\epsilon$

$$(15) \quad \varphi(\xi_i - \epsilon) - \psi(\xi_i - \epsilon) > 0.$$

But the sign changes in passing the point  $\xi_i$ ; therefore,

$$(16) \quad \varphi(\xi_i + \epsilon) - \psi(\xi_i + \epsilon) < 0.$$

The equalities

$$\varphi(\xi_1 - \epsilon) - \psi(\xi_1 - \epsilon) = 0, \quad \varphi(\xi_{n+1} + \epsilon) - \psi(\xi_{n+1} + \epsilon) = 0$$

cannot both hold for all sufficiently small  $\epsilon$ . For then there would not be a change of sign at  $\xi_1$  and  $\xi_{n+1}$ , so that the number of changes would not be greater than  $2n - 1$  which is impossible. Therefore, let

$$\varphi(\xi_1 - \epsilon) - \psi(\xi_1 - \epsilon) = 0 \quad \text{and} \quad \varphi(\xi_{n+1} + \epsilon) - \psi(\xi_{n+1} + \epsilon) < 0.$$

Then there will be exactly  $2n$  changes of sign: one in each of the intervals

$$\xi_{i-1}, \xi_i$$

and in each of the points  $\xi_2, \xi_3, \dots, \xi_{n+1}$ . The inequalities (15) and (16) would hold for  $i \geq 2$ , but

$$\varphi(\xi_1 - \epsilon) - \psi(\xi_1 - \epsilon) = 0, \quad \varphi(\xi_1 + \epsilon) - \psi(\xi_1 + \epsilon) < 0$$

for all sufficiently small  $\epsilon$ .

Now let

$$\varphi(\xi_{n+1} + \epsilon) - \psi(\xi_{n+1} + \epsilon) = 0 \quad \text{and} \quad \varphi(\xi_1 - \epsilon) - \psi(\xi_1 - \epsilon) > 0$$

for all sufficiently small positive  $\epsilon$ . Then there will be exactly  $2n$  changes of sign: In each of the points  $\xi_1, \xi_2, \dots, \xi_n$  and in each of the  $n$  intervals

$$\xi_{i-1}, \xi_i.$$

The inequalities (15) and (16) will again hold for  $i \leq n$ , but

$$\varphi(\xi_{n+1} - \epsilon) - \psi(\xi_{n+1} - \epsilon) > 0 \quad \text{and} \quad \varphi(\xi_{n+1} + \epsilon) - \psi(\xi_{n+1} + \epsilon) = 0$$

for all sufficiently small  $\epsilon$ . Letting  $\epsilon$  converge to 0, we shall have

$$\begin{aligned} \varphi(\xi_i - 0) &\geq \psi(\xi_i - 0) \\ \varphi(\xi_i + 0) &\leq \psi(\xi_i + 0) \end{aligned}$$

for  $i = 1, 2, 3, \dots, n + 1$  in all cases. Then, since

$$\varphi(\xi_i) \geq \varphi(\xi_i - 0); \quad \varphi(\xi_i) \leq \varphi(\xi_i + 0),$$

we shall have also

$$\begin{aligned} \varphi(\xi_i) &\geq \psi(\xi_i - 0) \\ \varphi(\xi_i) &\leq \psi(\xi_i + 0) \end{aligned}$$

or, taking into consideration the definition of the function  $\psi(x)$

$$\begin{aligned} \varphi(\xi_i) &\geq \sum_{l=1}^{i-1} \frac{P(\xi_l)}{Q'(\xi_l)} \\ \varphi(\xi_i) &\leq \sum_{l=1}^i \frac{P(\xi_l)}{Q'(\xi_l)}. \end{aligned}$$

These are the inequalities to which Tshebysheff's name is justly attached. For a particular root  $\xi_i = v$  they can be written thus:

$$(17) \quad \begin{aligned} \varphi(v) &\geq \sum_{\xi_l < v} \frac{P(\xi_l)}{Q'(\xi_l)} \\ \varphi(v) &\leq \sum_{\xi_l \leq v} \frac{P(\xi_l)}{Q'(\xi_l)} \end{aligned}$$

with the evident meaning of the extent of summations. Another, less explicit, form of the same inequalities is

$$(18) \quad \begin{aligned} \varphi(v) &\geq \psi(v-0) \\ \varphi(v) &\leq \psi(v+0). \end{aligned}$$

As to  $P(x)$  and  $Q(x)$ , they can be taken in the form:

$$\begin{aligned} P(x) &= [\alpha_{n+1}(x-v)Q_n(v) + Q_{n-1}(v)]P_n(x) - Q_n(v)P_{n-1}(x) \\ Q(x) &= [\alpha_{n+1}(x-v)Q_n(v) + Q_{n-1}(v)]Q_n(x) - Q_n(v)Q_{n-1}(x). \end{aligned}$$

Thus far we have assumed that  $v$  was different from any root of the equation

$$Q_n(x) = 0,$$

but all the results hold, even if

$$Q_n(v) = 0.$$

To prove this, we note first that when a variable  $v$  approaches a root  $\xi$  of  $Q_n(x)$ , one root of  $Q(x)$  (either  $\xi_1$  or  $\xi_{n+1}$ ) tends to  $-\infty$  or  $+\infty$ , while the remaining  $n$  roots approach the  $n$  roots  $x_1, x_2, \dots, x_n$  of the equation

$$Q_n(x) = 0.$$

If  $\xi_1$  tends to negative infinity, it is easy to see that

$$\frac{P(\xi_1)}{Q'(\xi_1)}$$

tends to 0. In this case the other quotients

$$\frac{P(\xi_l)}{Q'(\xi_l)}$$

tend respectively to

$$\frac{P_n(x_1)}{Q'_n(x_1)}, \frac{P_n(x_2)}{Q'_n(x_2)}, \dots$$

If  $\xi_{n+1}$  tends to positive infinity the quotients

$$\frac{P(\xi_l)}{Q'(\xi_l)}, l = 1, 2, \dots, n$$

approach respectively

$$\frac{P_n(x_l)}{Q'_n(x_l)}, l = 1, 2, 3, \dots, n,$$

while

$$\frac{P(\xi_{n+1})}{Q'(\xi_{n+1})}$$



tends to 0. Now take  $v = \xi - \epsilon$  and  $v = \xi + \epsilon$  in (17) and let the positive number  $\epsilon$  converge to 0. Taking into account the preceding remarks, we find in the limit

$$\begin{aligned} \varphi(\xi - 0) &\geq \sum_{x_1 < \xi} \frac{P_n(x_1)}{Q'_n(x_1)} \\ \varphi(\xi + 0) &\leq \sum_{x_1 \leq \xi} \frac{P_n(x_1)}{Q'_n(x_1)}, \end{aligned}$$

whence again

$$\begin{aligned} \varphi(\xi) &\geq \sum_{x_1 < \xi} \frac{P_n(x_1)}{Q'_n(x_1)} \\ \varphi(\xi) &\leq \sum_{x_1 \leq \xi} \frac{P_n(x_1)}{Q'_n(x_1)}. \end{aligned}$$

But these inequalities follow directly from (17) by taking  $v = \xi$ .

Since

$$\psi(v + 0) - \psi(v - 0) = \frac{P(v)}{Q'(v)}$$

it follows from inequalities (18) that

$$0 \leq \varphi(v) - \psi(v - 0) \leq \frac{P(v)}{Q'(v)}.$$

On the other hand, one easily finds that

$$\frac{P(v)}{Q'(v)} = \frac{1}{\alpha_{n+1}Q_n(v)^2 + Q'_n(v)Q_{n-1}(v) - Q'_{n-1}(v)Q_n(v)}.$$

But referring to the end of Sec. 4,

$$Q'_n(v)Q_{n-1}(v) - Q'_{n-1}(v)Q_n(v) = \sum_{s=1}^n \alpha_s Q_{s-1}(v)^2,$$

whence

$$\alpha_{n+1}Q_n(v)^2 + Q'_n(v)Q_{n-1}(v) - Q'_{n-1}(v)Q_n(v) = Q'_{n+1}(v)Q_n(v) - Q'_n(v)Q_{n+1}(v).$$

Finally,

$$0 \leq \varphi(v) - \psi(v - 0) \leq \frac{1}{Q'_{n+1}(v)Q_n(v) - Q'_n(v)Q_{n+1}(v)}.$$

If  $\varphi_1(v)$  is another distribution function with the same moments

$$m_0, m_1, m_2, \dots, m_{2n},$$

we shall have also

$$0 \leq \varphi_1(v) - \psi(v - 0) \leq \frac{1}{Q'_{n+1}(v)Q_n(v) - Q'_n(v)Q_{n+1}(v)},$$

and as a consequence,

$$(19) \quad |\varphi_1(v) - \varphi(v)| \leq \chi_n(v)$$

—a very important inequality. Here for brevity we use the notation

$$\chi_n(v) = \frac{1}{Q'_{n+1}(v)Q_n(v) - Q'_n(v)Q_{n+1}(v)}.$$

**7. Application to Normal Distribution.** An important particular case is that of a normal distribution characterized by the function

$$\varphi(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-u^2} du.$$

In this case it is easy to give an explicit expression of the polynomials  $Q_n(x)$ . Let

$$H_n(x) = e^{x^2} \frac{d^n e^{-x^2}}{dx^n}.$$

Integrating by parts, one can prove that for  $l \leq n - 1$

$$\int_{-\infty}^{\infty} e^{-x^2} x^l H_n(x) dx = 0.$$

Hence, one may conclude that  $Q_n(x)$  differs from  $H_n(x)$  by a constant factor. Let

$$Q_n(x) = c_n H_n(x).$$

To determine  $c_n$ , we may use the relation

$$H_n(x) = -2xH_{n-1}(x) - 2(n-1)H_{n-2}(x)$$

which can readily be established. Introducing polynomials  $Q_n$ , this relation becomes

$$Q_n(x) = -2x \frac{c_n}{c_{n-1}} Q_{n-1}(x) - 2(n-1) \frac{c_n}{c_{n-2}} Q_{n-2}(x).$$

Hence,

$$\frac{c_n}{c_{n-2}} = \frac{1}{2n-2}, \quad \alpha_n = -2 \frac{c_n}{c_{n-1}}, \quad \beta_n = 0.$$

Since  $H_0(x) = Q_0(x) = 1$ , we have  $c_0 = 1$ ; also

$$\alpha_1 = \frac{1}{m_0} = 1 = -2 \frac{c_1}{c_0}$$

whence  $c_1 = -\frac{1}{2}$ . The knowledge of  $c_0$  and  $c_1$  together with the relation

$$c_n = \frac{c_{n-2}}{2n-2}$$

allows determination of all members of the sequence  $c_2, c_3, c_4, \dots$ . The final expressions are as follows:

$$c_{2m} = \frac{1}{2^m \cdot 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2m-1)}$$

$$c_{2m+1} = \frac{-1}{2^{m+1} \cdot 2 \cdot 4 \cdot 6 \cdot \dots \cdot 2m}$$

From the above relation between  $H_n(x), H_{n-1}(x), H_{n-2}(x)$  and owing to the fact that  $H_n(x)$  is an even or odd polynomial, according as  $n$  is even or odd, one finds

$$H_{2m}(0) = (-2)^m \cdot 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2m-1),$$

while another relation

$$H'_n(x) = -2nH_{n-1}(x),$$

following from the definition of  $H_n(x)$ , gives

$$H'_{2m-1}(0) = (-2)^m \cdot 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2m-1).$$

These preliminaries being established, we shall prove now that

$$\chi_n(v) = \frac{1}{c_n c_{n+1} (H'_{n+1}(v)H_n(v) - H'_n(v)H_{n+1}(v))}$$

attains its maximum for  $v = 0$ . Let

$$\Omega(v) = H'_{n+1}(v)H_n(v) - H'_n(v)H_{n+1}(v).$$

Then, taking into account the differential equation for polynomials  $H_n(v)$ :

$$H''_n(v) = 2vH'_n(v) - 2nH_n(v)$$

we find that

$$\frac{d\Omega}{dv} = 2v\Omega - 2H_n(v)H_{n+1}(v).$$

On the other hand,

$$\Omega = -H_{n+1}(v) \frac{d}{dv} \frac{H_n(v)}{H_{n+1}(v)},$$

and denoting roots of the polynomial  $H_{n+1}(v)$  in general by  $\xi$ .

$$\frac{d}{dv} \frac{H_n(v)}{H_{n+1}(v)} = - \sum \frac{H_n(\xi)}{H'_{n+1}(\xi)} \frac{1}{(v-\xi)^2}$$

Consequently

$$\Omega = H_{n+1}(v)^2 \sum \frac{H_n(\xi)}{H'_{n+1}(\xi)} \frac{1}{(v - \xi)^2}.$$

Again

$$H_n(v)H_{n+1}(v) = H_{n+1}(v)^2 \sum \frac{H_n(\xi)}{H'_{n+1}(\xi)} \frac{v - \xi}{(v - \xi)^2},$$

and so

$$\frac{d\Omega}{dv} = 2H_{n+1}(v)^2 \sum \frac{H_n(\xi)}{H'_{n+1}(\xi)} \frac{\xi}{(v - \xi)^2} = \frac{-H_{n+1}(v)^2}{n+1} \sum \frac{\xi}{(v - \xi)^2}.$$

Roots of the polynomial  $H_{n+1}(x)$  being symmetrically located with respect to 0, we have:

$$\sum \frac{\xi}{(v - \xi)^2} = - \sum \frac{\xi}{(v + \xi)^2} = 2v \sum \frac{\xi^2}{(v^2 - \xi^2)^2},$$

and finally

$$\frac{d\Omega}{dv} = -2v \frac{H_{n+1}(v)^2}{n+1} \sum \frac{\xi^2}{(v^2 - \xi^2)^2}.$$

Hence

$$\frac{d\Omega}{dv} > 0 \quad \text{if} \quad v < 0; \quad \frac{d\Omega}{dv} < 0 \quad \text{if} \quad v > 0$$

that is,  $\Omega(v)$  attains its maximum for  $v = 0$  and  $\chi_n(v)$  attains its maximum for  $v = 0$ . Referring to the above expressions of  $c_{2m}$ ,  $c_{2m+1}$ ;  $H_{2m}(0)$ ,  $H'_{2m+1}(0)$ , we find that

$$\begin{aligned} \chi_{2m}(0) &= \frac{2 \cdot 4 \cdot 6 \cdots 2m}{3 \cdot 5 \cdot 7 \cdots (2m+1)} \\ \chi_{2m+1}(0) &= \frac{2 \cdot 4 \cdot 6 \cdots 2m}{3 \cdot 5 \cdot 7 \cdots (2m+1)}. \end{aligned}$$

In Appendix I, page 354, we find the inequality

$$\frac{2 \cdot 4 \cdot 6 \cdots 2m}{1 \cdot 3 \cdot 5 \cdots (2m-1)} \frac{1}{\sqrt{4m+2}} < \frac{\sqrt{\pi}}{2}$$

whence

$$\frac{2 \cdot 4 \cdot 6 \cdots 2m}{3 \cdot 5 \cdot 7 \cdots (2m+1)} < \sqrt{\frac{\pi}{4m+2}}$$

Thus, in all cases

$$\chi_n(v) \leq \chi_n(0) < \sqrt{\frac{\pi}{2n}},$$

whence, by virtue of inequality (19),

$$|\varphi_1(v) - \varphi(v)| < \sqrt{\frac{\pi}{2n}}$$

Thus any distribution function  $\varphi_1(v)$  with the moments

$$m_0 = 1, \quad m_{2k-1} = 0, \quad m_{2k} = \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2^k} \quad (k \leq n)$$

corresponding to

$$\varphi(v) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^v e^{-u^2} du$$

differs from  $\varphi_1(v)$  by less than

$$\sqrt{\frac{\pi}{2n}}$$

Since this quantity tends to 0 when  $n$  increases indefinitely, we have the following theorem proved for the first time by Tshebysheff:

*The system of infinitely many equations*

$$\int_{-\infty}^{\infty} d\varphi(x) = 1; \quad \int_{-\infty}^{\infty} x^{2k-1} d\varphi(x) = 0; \quad \int_{-\infty}^{\infty} x^{2k} d\varphi(x) = \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2^k}$$

$$k = 1, 2, 3, \dots$$

uniquely determines a never decreasing function  $\varphi(x)$  such that  $\varphi(-\infty) = 0$ ; namely,

$$\varphi(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-u^2} du.$$

**8. Tshebysheff-Markoff's Fundamental Theorem.** When a mass = 1 is distributed according to the law characterized by a function  $F(x, \lambda)$  depending upon a parameter  $\lambda$ , we say that the distribution is variable. Notwithstanding the variability of distribution, it may happen that its moments remain constant. If they are equal to moments of normal distribution with density

$$\frac{e^{-u^2}}{\sqrt{\pi}}$$

then by the preceding theorem we have rigorously

$$F(x, \lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-u^2} du$$

no matter what  $\lambda$  is.

Generally moments of a variable distribution are themselves variable. Suppose that each one of them, when  $\lambda$  tends to a certain limit (for instance  $\infty$ ), tends to the corresponding moment of normal distribution. One can foresee that under such circumstances  $F(x, \lambda)$  will tend to

$$\varphi(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-u^2} du.$$

In fact, the following fundamental theorem holds:

**Fundamental Theorem.** *If, for a variable distribution characterized by the function  $F(x, \lambda)$ ,*

$$\lim \int_{-\infty}^{\infty} x^k dF(x, \lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} x^k dx; \quad \lambda \rightarrow \infty$$

for any fixed  $k = 0, 1, 2, 3, \dots$ , then

$$\lim F(v, \lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^v e^{-x^2} dx; \quad \lambda \rightarrow \infty$$

uniformly in  $v$ .

**Proof.** Let

$$m_0, m_1, m_2, \dots, m_{2n}$$

be  $2n + 1$  moments corresponding to a normal distribution. They allow formation of the polynomials

$$Q_0(x), Q_1(x), \dots, Q_n(x) \text{ and } Q(x)$$

and the function designated in Sec. 6 by  $\psi(x)$ . Similar entities corresponding to the variable distribution will be specified by an asterisk. Since

$$m_k^* \rightarrow m_k \quad \text{as} \quad \lambda \rightarrow \infty$$

and since  $\Delta_n > 0$ , we shall have

$$\Delta_n^* > 0$$

for sufficiently large  $\lambda$ . Then  $F(x, \lambda)$  will have not less than  $n + 1$  points of increase and the whole theory can be applied to variable distribution. In particular, we shall have

$$0 \leq \varphi(v) - \psi(v - 0) \leq \chi_n(v)$$

(20)

$$0 \leq F(v, \lambda) - \psi^*(v - 0) \leq \chi_n^*(v).$$

Now  $Q_s^*(x)$  ( $s = 0, 1, 2, \dots, n$ ) and  $Q^*(x)$  depend rationally upon

$m_k^*(k = 0, 1, 2, \dots, 2n)$ ; hence, without any difficulty one can see that

$$\begin{aligned} Q_s^*(x) &\rightarrow Q_s(x); & s = 0, 1, 2, \dots, n \\ Q^*(x) &\rightarrow Q(x) \end{aligned}$$

as  $\lambda \rightarrow \infty$ ; whence,

$$\chi_n^*(v) \rightarrow \chi_n(v).$$

Again

$$\psi^*(v - 0) \rightarrow \psi(v - 0)$$

as  $\lambda \rightarrow \infty$ . A few explanations are necessary to prove this. At first let  $Q_n(v) \neq 0$ . Then the polynomial  $Q(x)$  will have  $n + 1$  roots

$$\xi_1 < \xi_2 < \xi_3 < \dots < \xi_{n+1}.$$

Since the roots of an algebraic equation vary continuously with its coefficients, it is evident that for sufficiently large  $\lambda$  the equation

$$Q^*(x) = 0$$

will have  $n + 1$  roots:

$$\xi_1^* < \xi_2^* < \xi_3^* < \dots < \xi_{n+1}^*$$

and  $\xi_k^*$  will tend to  $\xi_k$  as  $\lambda \rightarrow \infty$ . In this case, it is evident that  $\psi^*(v - 0)$  will tend to  $\psi(v - 0)$ . If  $Q_n(v) = 0$ , it may happen that  $\xi_1^*$  or  $\xi_{n+1}^*$  tends respectively to  $-\infty$  or  $+\infty$  as  $\lambda \rightarrow \infty$ , while the other roots tend to the roots

$$x_1, x_2, \dots, x_n$$

of the equation

$$Q_n(x) = 0.$$

But the terms in  $\psi^*(v - 0)$  corresponding to infinitely increasing roots tend to 0, and again

$$\psi^*(v - 0) \rightarrow \psi(v - 0).$$

Now

$$\chi_n(v) < \sqrt{\frac{v}{2n}}$$

Consequently, given an arbitrary positive number  $\epsilon$ , we can select  $n$  so large as to have

$$\chi_n(v) < \sqrt{\frac{v}{2n}} < \epsilon.$$

Having selected  $n$  in this manner, we shall keep it fixed. Then by the preceding remarks a number  $L$  can be found so that

$$\begin{aligned} \chi^*(v) &< \sqrt{\frac{\pi}{2n}} < \epsilon \\ |\psi(v-0) - \psi^*(v-0)| &< \epsilon \end{aligned}$$

for  $\lambda > L$ . Combining this with inequalities (20), we find

$$|F(v, \lambda) - \varphi(v)| < 3\epsilon$$

for  $\lambda > L$ . And this proves the convergence of  $F(v, \lambda)$  to  $\varphi(v)$  for a fixed arbitrary  $v$ . To show that the equation

$$\lim F(v, \lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} dx$$

holds uniformly for a variable  $v$  we can follow a very simple reasoning due to Pólya. Since  $\varphi(-\infty) = 0$ ,  $\varphi(+\infty) = 1$  and  $\varphi(x)$  is an increasing function, one can determine two numbers  $a_0$  and  $a_n$  so that

$$\begin{aligned} \varphi(x) \leq \varphi(a_0) < \frac{\epsilon}{2} & \quad \text{for } x \leq a_0 \\ 1 - \varphi(x) \leq 1 - \varphi(a_n) < \frac{\epsilon}{2} & \quad \text{for } x \geq a_n. \end{aligned}$$

Next, because  $\varphi(x)$  is a continuous function, the interval  $(a_0, a_n)$  can be subdivided into partial intervals by inserting between  $a_0$  and  $a_n$  points  $a_1 < a_2 < \dots < a_{n-1}$  so that

$$0 < \varphi(a_{k+1}) - \varphi(a_k) < \frac{\epsilon}{2}$$

for  $k = 0, 1, 2, \dots, n-1$ . By the preceding result, for all sufficiently large  $\lambda$

$$F(a_0, \lambda) < \frac{\epsilon}{2}; \quad 1 - F(a_n, \lambda) < \frac{\epsilon}{2}$$

and

$$|F(a_k, \lambda) - \varphi(a_k)| < \frac{\epsilon}{2}; \quad k = 1, 2, \dots, n-1.$$

Now consider the interval  $(-\infty, a_0)$ . Here for  $v \leq a_0$

$$0 \leq F(v, \lambda) < \frac{\epsilon}{2}; \quad 0 < \varphi(v) < \frac{\epsilon}{2}$$

and

$$|F(v, \lambda) - \varphi(v)| < \epsilon.$$



For  $v$  belonging to the interval  $(a_n, +\infty)$

$$0 \leq 1 - F(v, \lambda) < \frac{\epsilon}{2}, \quad 0 < 1 - \varphi(v) < \frac{\epsilon}{2}$$

whence again

$$|F(v, \lambda) - \varphi(v)| < \epsilon.$$

Finally, let

$$a_k \leq v < a_{k+1} \quad (k = 0, 1, 2, \dots, n-1).$$

Then

$$\begin{aligned} F(v, \lambda) - \varphi(v) &\geq F(a_k, \lambda) - \varphi(a_{k+1}) = \\ &= [F(a_k, \lambda) - \varphi(a_k)] + [\varphi(a_k) - \varphi(a_{k+1})] \\ F(v, \lambda) - \varphi(v) &\leq F(a_{k+1}, \lambda) - \varphi(a_k) = \\ &= [F(a_{k+1}, \lambda) - \varphi(a_{k+1})] + [\varphi(a_{k+1}) - \varphi(a_k)]. \end{aligned}$$

But

$$\begin{aligned} F(a_k, \lambda) - \varphi(a_k) &> -\frac{\epsilon}{2}; \quad \varphi(a_k) - \varphi(a_{k+1}) > -\frac{\epsilon}{2} \\ F(a_{k+1}, \lambda) - \varphi(a_{k+1}) &< \frac{\epsilon}{2}; \quad \varphi(a_{k+1}) - \varphi(a_k) < \frac{\epsilon}{2} \end{aligned}$$

whence

$$-\epsilon < F(v, \lambda) - \varphi(v) < \epsilon.$$

Thus, given  $\epsilon$ , there exists a number  $L(\epsilon)$  depending upon  $\epsilon$  alone and such that

$$|F(v, \lambda) - \varphi(v)| < \epsilon$$

for  $\lambda > L(\epsilon)$  no matter what value is attributed to  $v$ .

The fundamental theorem with reference to probability can be stated as follows:

Let  $s_n$  be a stochastic variable depending upon a variable positive integer  $n$ . If the mathematical expectation  $E(s_n^k)$  for any fixed  $k = 1, 2, 3, \dots$  tends, as  $n$  increases indefinitely, to the corresponding expectation

$$E(x^k) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2} dx$$

of a normally distributed variable, then the probability of the inequality

$$s_n < v$$

tends to the limit

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^v e^{-x^2} dx$$

and that uniformly in  $v$ .

In very many cases it is much easier to make sure that the conditions of this theorem are fulfilled and then, in one stroke, to pass to the limit theorem for probability, than to attack the problem directly.

## APPLICATION TO SUMS OF INDEPENDENT VARIABLES

9. Let  $z_1, z_2, z_3, \dots$  be independent variables whose number can be increased indefinitely. Without losing anything in generality, we may suppose from the beginning

$$E(z_k) = 0; \quad k = 1, 2, 3, \dots$$

We assume the existence of

$$E(z_k^2) = b_k$$

for all  $k = 1, 2, 3, \dots$ . Also, we assume for *some* positive  $\delta$  the existence of absolute moments

$$E|z_k|^{2+\delta} = \mu_k^{(2+\delta)}; \quad k = 1, 2, 3, \dots$$

Liapounoff's theorem, with which we dealt at length in Chap. XIV, states that the probability of the inequality

$$\frac{z_1 + z_2 + \dots + z_n}{\sqrt{2B_n}} < t,$$

where

$$B_n = b_1 + b_2 + \dots + b_n$$

tends uniformly to the limit

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^t e^{-x^2} dx$$

as  $n \rightarrow \infty$ , provided

$$\frac{\mu_1^{(2+\delta)} + \mu_2^{(2+\delta)} + \dots + \mu_n^{(2+\delta)}}{B_n^{1+\frac{\delta}{2}}} \rightarrow 0.$$

Liapounoff's result in regard to generality of conditions surpassed by far what had been established before by Tshebysheff and Markoff, whose proofs were based on the fundamental result derived in the preceding section. Since Liapounoff's conditions do not require the existence of moments in an infinite number, it seemed that the method of moments was not powerful enough to establish the limit theorem in such a general form. Nevertheless, by resorting to an ingenious artifice, of which we made use in Chap. X, Sec. 8, Markoff finally succeeded in proving the limit theorem by the method of moments to the same degree of generality as did Liapounoff.

Markoff's artifice consists in associating with the variable  $z_k$  two new variables  $x_k$  and  $y_k$  defined as follows:

Let  $N$  be a positive number which in the course of proof will be selected so as to tend to infinity together with  $n$ . Then

$$\begin{aligned} x_k &= z_k, & y_k &= 0 & \text{if } |z_k| \leq N \\ x_k &= 0, & y_k &= z_k & \text{if } |z_k| > N. \end{aligned}$$

Evidently  $z_k, x_k, y_k$  are connected by the relation

$$z_k = x_k + y_k$$

whence

$$(21) \quad E(x_k) + E(y_k) = 0.$$

Moreover

$$(22) \quad E(x_k^2) + E(y_k^2) = E(z_k^2) = \mu_k$$

$$E|x_k|^{2+\delta} + E|y_k|^{2+\delta} = E|z_k|^{2+\delta} = \mu_k^{(2+\delta)},$$

as one can see immediately from the definition of  $x_k$  and  $y_k$ .

Since  $x_k$  is bounded, mathematical expectations

$$E(x_k^l)$$

exist for all integer exponents  $l = 1, 2, 3, \dots$  and for  $k = 1, 2, 3, \dots$ .

In the following we shall use the notations

$$\begin{aligned} |E(x_k^l)| &= c_k^{(l)}; & l &= 1, 2, 3, \dots \\ c_1^{(2)} + c_2^{(2)} + \dots + c_n^{(2)} &= B'_n \\ \mu_1^{(2+\delta)} + \mu_2^{(2+\delta)} + \dots + \mu_n^{(2+\delta)} &= C_n. \end{aligned}$$

Not to obscure the essential steps of the reasoning we shall first establish a few preliminary results.

**Lemma 1.** Let  $q_k$  represent the probability that  $y_k \neq 0$ ; then

$$q_1 + q_2 + \dots + q_n \leq \frac{C_n}{N^{2+\delta}}$$

**Proof.** Let  $\varphi_k(x)$  be the distribution function of  $z_k$ . Since  $y_k \neq 0$  only if  $|z_k| > N$ , the probability  $q_k$  is not greater than

$$\int_{-N}^{-\infty} d\varphi_k(x) + \int_N^{\infty} d\varphi_k(x).$$

On the other hand,

$$\int_{-N}^{-\infty} |x|^{2+\delta} d\varphi_k(x) + \int_N^{\infty} |x|^{2+\delta} d\varphi_k(x) \leq \mu_k^{(2+\delta)}.$$

But

$$\int_{-\infty}^{-N} |x|^{2+\delta} d\varphi_k(x) + \int_N^{\infty} |x|^{2+\delta} d\varphi_k(x) \geq N^{2+\delta} \left\{ \int_{-\infty}^{-N} d\varphi_k(x) + \int_N^{\infty} d\varphi_k(x) \right\},$$

whence

$$q_k \leq \int_{-\infty}^{-N} d\varphi_k(x) + \int_N^{\infty} d\varphi_k(x) \leq \frac{\mu_k^{(2+\delta)}}{N^{2+\delta}}.$$

The inequality to be proved follows immediately.

**Lemma 2.** *The following inequality holds:*

$$1 \geq \frac{B'_n}{B_n} \geq 1 - \frac{C_n}{B_n N^\delta}.$$

**Proof.** From

$$E|y_k|^{2+\delta} \leq \mu_k^{(2+\delta)}$$

which is a consequence of the second equation (22) it follows that

$$E(y_k^2) \leq \frac{\mu_k^{(2+\delta)}}{N^\delta}.$$

The first equation (22)

$$c_k^{(2)} + E(y_k^2) = b_k$$

gives

$$b_k \geq c_k^{(2)} \geq b_k - \frac{\mu_k^{(2+\delta)}}{N^\delta}.$$

Taking the sum for  $k = 1, 2, 3, \dots, n$ , we get

$$B_n \geq B'_n \geq B_n - \frac{C_n}{N^\delta},$$

whence

$$1 \geq \frac{B'_n}{B_n} \geq 1 - \frac{C_n}{B_n N^\delta}.$$

**Lemma 3.** *For  $e \geq 3$ ,*

$$\frac{c_1^{(e)} + c_2^{(e)} + \dots + c_n^{(e)}}{B_n^{\frac{e}{2}}} \leq \left( \frac{N^2}{B_n} \right)^{\frac{e-2}{2}}.$$

**Proof.** This inequality follows immediately from the evident inequalities

$$c_k^{(e)} \leq E|x_k|^e \leq N^{e-2} E(x_k^2) \leq N^{e-2} b_k.$$

**Lemma 4.** *The following inequality holds*

$$\frac{c_1^{(1)} + c_2^{(1)} + \dots + c_n^{(1)}}{B_n^{\frac{1}{2}}} \leq \left( \frac{C_n}{N^{2+\delta}} \right)^{\frac{1}{2}}$$

**Proof.** Since

$$E(x_k) + E(y_k) = 0,$$

we have

$$c_k^{(1)} = |E(x_k)| = |E(y_k)| \leq E|y_k|.$$

On the other hand, by virtue of Schwarz's inequality

$$\begin{aligned} [E|y_1| + E|y_2| + \dots + E|y_n|]^2 &\leq \\ &\leq (q_1 + q_2 + \dots + q_n) \sum_{k=1}^n E(y_k^2) \leq B_n \frac{C_n}{N^{2+\delta}}, \end{aligned}$$

whence the statement follows immediately.

If the variable integer  $N$  should be subject to the requirements that both the ratios

$$\frac{C_n}{N^{2+\delta}} \quad \text{and} \quad \frac{N^2}{B_n}$$

should tend to 0 when  $n$  increases indefinitely, then the preceding lemmæ would give three important corollaries. But before stating these corollaries we must ascertain the possibility of selecting  $N$  as required. It suffices to take

$$N = (B_n C_n)^{\frac{1}{4+\delta}}.$$

Then

$$\frac{N^2}{B_n} = \frac{C_n}{N^{2+\delta}} = \left( \frac{C_n}{B_n^{1+\frac{\delta}{2}}} \right)^{\frac{2}{4+\delta}} \rightarrow 0$$

by virtue of Liapounoff's condition.

Also

$$\frac{C_n}{B_n N^{\delta}} = \left( \frac{C_n}{N^{2+\delta}} \right)^{\frac{\delta}{2+\delta}} \cdot \left( \frac{C_n}{B_n^{1+\frac{\delta}{2}}} \right)^{\frac{2}{2+\delta}}$$

will tend to 0. By selecting  $N$  in this manner we can state the following corollaries:

**Corollary 1.** *The sum*

$$q_1 + q_2 + \dots + q_n$$

tends to 0 as  $n \rightarrow \infty$ .

Corollary 2. The ratio

$$\frac{B'_n}{B_n}$$

tends to 1.

Corollary 3. The ratio

$$\frac{c_1^{(e)} + c_2^{(e)} + \cdots + c_n^{(e)}}{B_n^{\frac{e}{2}}}$$

tends to 0 for all positive integer exponents  $e$  except  $e = 2$ .

10. Let  $F_n(t)$  and  $\phi_n(t)$  represent, respectively, the probabilities of the inequalities

$$\frac{z_1 + z_2 + \cdots + z_n}{\sqrt{2B_n}} < t$$

$$\frac{x_1 + x_2 + \cdots + x_n}{\sqrt{2B_n}} < t.$$

By repeating the reasoning developed in Chap. X, Sec. 8, we find that

$$|F_n(t) - \phi_n(t)| \leq q_1 + q_2 + \cdots + q_n.$$

Hence,

$$\lim (F_n(t) - \phi_n(t)) = 0 \quad \text{as} \quad n \rightarrow \infty$$

by Corollary 1. It suffices therefore to show

$$\phi_n(t) \rightarrow \frac{1}{\sqrt{\pi}} \int_{-\infty}^t e^{-x^2} dx \quad \text{as} \quad n \rightarrow \infty,$$

and that can be done by the method of moments. By the polynomial theorem

$$\left( \frac{x_1 + x_2 + \cdots + x_n}{\sqrt{2B_n}} \right)^m = \sum \frac{m!}{\alpha! \beta! \cdots \lambda!} \frac{S_{\alpha, \beta, \dots, \lambda}}{2^{\frac{m}{2}} B_n^{\frac{m}{2}}}$$

where the summation extends over all systems of positive integers  $\alpha \geq \beta \geq \cdots \geq \lambda$  satisfying the condition

$$\alpha + \beta + \cdots + \lambda = m$$

and  $S_{\alpha, \beta, \dots, \lambda}$  denotes a symmetrical function of letters  $x_1, x_2, \dots, x_n$  determined by one of its terms

$$x_1^\alpha x_2^\beta \cdots x_l^\lambda$$

if  $l$  represents the number of integers  $\alpha, \beta, \dots, \lambda$ . Since variables  $x_1, x_2, \dots, x_n$  are independent, we have

$$E\left(\frac{x_1 + x_2 + \dots + x_n}{\sqrt{2B_n}}\right)^m = \sum \frac{m!}{\alpha! \beta! \dots \lambda!} \frac{G_{\alpha, \beta, \dots, \lambda}}{2^{\frac{m}{2}} B_n^{\frac{m}{2}}}$$

where  $G_{\alpha, \beta, \dots, \lambda}$  is obtained by replacing powers of variables by mathematical expectations of these powers. It is almost evident that

$$\frac{|G_{\alpha, \beta, \dots, \lambda}|}{B_n^{\frac{m}{2}}} \leq \frac{c_1^{(\alpha)} + c_2^{(\alpha)} + \dots + c_n^{(\alpha)}}{B_n^{\frac{\alpha}{2}}} \cdot \frac{c_1^{(\beta)} + c_2^{(\beta)} + \dots + c_n^{(\beta)}}{B_n^{\frac{\beta}{2}}} \dots \frac{c_1^{(\lambda)} + c_2^{(\lambda)} + \dots + c_n^{(\lambda)}}{B_n^{\frac{\lambda}{2}}}$$

Now if not all the exponents  $\alpha, \beta, \dots, \lambda$  are  $= 2$  (which is possible only when  $m$  is even), by virtue of Corollary 3 the right member as well as

$$\frac{G_{\alpha, \beta, \dots, \lambda}}{B_n^{\frac{m}{2}}}$$

tends to 0. Hence

$$E\left(\frac{x_1 + x_2 + \dots + x_n}{\sqrt{2B_n}}\right)^m \rightarrow 0$$

if  $m$  is odd.

But for even  $m$  we have

$$(23) \quad E\left(\frac{x_1 + x_2 + \dots + x_n}{\sqrt{2B_n}}\right)^m = \frac{m! G_{2, 2, \dots, 2}}{2^{\frac{m}{2}} B_n^{\frac{m}{2}}} \rightarrow 0.$$

Let us consider now ( $m$  being even)

$$\left(\frac{B_n'}{B_n}\right)^{\frac{m}{2}} = \left(\frac{c_1^{(2)} + c_2^{(2)} + \dots + c_n^{(2)}}{B_n}\right)^{\frac{m}{2}} = \sum \frac{\frac{m!}{2}}{\lambda! \mu! \dots \omega!} \frac{H_{\lambda, \mu, \dots, \omega}}{B_n^{\frac{m}{2}}}$$

where summation extends over all systems of positive integers

$$\lambda \geq \mu \geq \dots \geq \omega$$

satisfying the condition

$$\lambda + \mu + \dots + \omega = \frac{m}{2}$$

and  $H_{\lambda, \mu, \dots, \omega}$  is a symmetric function of  $c_1^{(2)}, c_2^{(2)}, \dots, c_n^{(2)}$  determined by its term

$$(c_1^{(2)})^\lambda (c_2^{(2)})^\mu \dots (c_n^{(2)})^\omega,$$

$l$  being the number of subscripts  $\lambda, \mu, \dots, \omega$ . Apparently

$$\frac{H_{\lambda, \mu, \dots, \omega}}{B_n^{\frac{m}{2}}} \leq \frac{(c_1^{(2)})^\lambda + (c_2^{(2)})^\lambda + \dots + (c_n^{(2)})^\lambda}{B_n^\lambda} \dots \frac{(c_1^{(2)})^\omega + (c_2^{(2)})^\omega + \dots + (c_n^{(2)})^\omega}{B_n^\omega}.$$

Besides

$$c_k^{(2)} \leq N^2, \quad (c_k^{(2)})^e \leq N^{2e-2} c_k^{(2)} \leq N^{2e-2} b_k$$

and

$$\frac{(c_1^{(2)})^e + (c_2^{(2)})^e + \dots + (c_n^{(2)})^e}{B_n^e} \leq \left(\frac{N^2}{B_n}\right)^{e-1} \rightarrow 0$$

if  $e > 1$ . Thus

$$\frac{H_{\lambda, \mu, \dots, \omega}}{B_n^{\frac{m}{2}}} \rightarrow 0$$

if not all subscripts  $\lambda, \mu, \dots, \omega$  are equal to 1. It follows that

$$\left(\frac{B'_n}{B_n}\right)^{\frac{m}{2}} - \left(\frac{m}{2}\right)! \frac{H_{1,1, \dots, 1}}{B_n^{\frac{m}{2}}} \rightarrow 0.$$

But by Corollary 2

$$\frac{B'_n}{B_n} \rightarrow 1$$

and evidently  $H_{1,1, \dots, 1} = G_{2,2, \dots, 2}$ . Hence

$$\left(\frac{m}{2}\right)! \frac{G_{2,2, \dots, 2}}{B_n^{\frac{m}{2}}} \rightarrow 1$$

and this in connection with (23) shows that for an even  $m$

$$E\left(\frac{x_1 + x_2 + \dots + x_n}{\sqrt{2B_n}}\right)^m \rightarrow \frac{m!}{2^m \left(\frac{m}{2}\right)!}.$$

Finally, no matter whether the exponent  $m$  is odd or even, we have

$$\lim E\left(\frac{x_1 + x_2 + \dots + x_n}{\sqrt{2B_n}}\right)^m = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^m e^{-x^2} dx.$$



Tshebysheff-Markoff's fundamental theorem can be applied directly and leads to the result:

$$\lim \phi_n(t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^t e^{-x^2} dx$$

uniformly in  $t$ . On the other hand, as has been established before,

$$\lim [F_n(t) - \phi_n(t)] = 0$$

uniformly in  $t$ . Hence, finally

$$\lim F_n(t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^t e^{-x^2} dx$$

uniformly in  $t$ .

And this is the fundamental limit theorem with Liapounoff's conditions now proved by the method of moments. This proof, due to Markoff, is simple enough and of high elegance. However, preliminary considerations which underlie the proof of the fundamental theorem, though simple and elegant also, are rather long. Nevertheless, we must bear in mind that they are not only useful in connection with the theory of probability, but they have great importance in other fields of analysis.

## APPENDIX III

### ON A GAUSSIAN PROBLEM

1. In a letter to Laplace dated January 30, 1812,<sup>1</sup> Gauss mentions a difficult problem in probability for which he could not find a perfectly satisfactory solution. We quote from his letter:

Je me rappelle pourtant d'un problème curieux duquel je me suis occupé il y a 12 ans, mais lequel je n'ai pas réussi alors à résoudre à ma satisfaction. Peut-être daignerez-vous en occuper quelques moments: dans ce cas je suis sur que vous trouverez une solution plus complète. La voici: Soit  $M$  une quantité inconnue entre les limites 0 et 1 pour laquelle toutes les valeurs sont ou également probables ou plus ou moins selon une loi donnée: qu'on la suppose convertie en une fraction continue

$$M = \frac{1}{a} + \frac{1}{a'} + \dots$$

Quelle est la probabilité qu'en s'arrêtant dans le développement à un terme fini  $a^{(n)}$  la fraction suivante

$$\frac{1}{a^{(n+1)}} + \frac{1}{a^{(n+2)}} + \dots$$

soit entre les limites 0 et  $x$ ? Je la désigne par  $P(n, x)$  et j'ai en supposant toutes les valeurs également probables

$$P(0, x) = x.$$

$P(1, x)$  est une fonction transcendante dépendant de la fonction

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{x}$$

que Euler nomme inexplicable et sur laquelle je viens de donner plusieurs recherches dans un mémoire présenté à notre Société des Sciences qui sera bientôt imprimé. Mais pour le cas où  $n$  est plus grand, la valeur exacte de  $P(n, x)$  semble intraitable. Cependant j'ai trouvé par des raisonnements très simples que pour  $n$  infinie

$$P(n, x) = \frac{\log(1+x)}{\log 2}$$

<sup>1</sup> Gauss' Werke, X, 1, p. 371.

Mais les efforts que j'ai fait lors de mes recherches pour assigner

$$P(n, x) = \frac{\log(1+x)}{\log 2}$$

pour une valeur très grande de  $n$ , mais pas infinie, ont été infructueux.

The problem itself and the main difficulty in its solution are clearly indicated in this passage. The problem is difficult indeed, and no satisfactory solution was offered before 1928, when Professor R. O. Kuzmin succeeded in solving it in a very remarkable and elegant way.

**2. Analytical Expression for  $P_n(x)$ .** We shall use the notation  $P_n(x)$  for the probability which Gauss designated by  $P(n, x)$ . The first question that presents itself is how to express  $P_n(x)$  in a proper analytical form. Let  $\delta(v_1, v_2, \dots, v_n, x)$  be an interval whose end points are represented by two continued fractions:

$$\frac{1}{v_1} + \frac{1}{v_2} + \dots + \frac{1}{v_n + x} \quad \text{and} \quad \frac{1}{v'_1} + \frac{1}{v'_2} + \dots + \frac{1}{v'_n}$$

with positive integer incomplete quotients  $v_1, v_2, \dots, v_n$ , while  $x$  is a positive number  $\leq 1$ . Two such intervals corresponding to two *different* systems of integers  $v_1, v_2, \dots, v_n$  and  $v'_1, v'_2, \dots, v'_n$  do not overlap; that is, do not have common inner points. For, if they had a common inner point represented by an irrational number  $N$  (which we can always suppose), we should have for some positive  $x' < 1$  and  $x'' < 1$

$$N = \frac{1}{v_1} + \frac{1}{v_2} + \dots + \frac{1}{v_n + x'} = \frac{1}{v'_1} + \frac{1}{v'_2} + \dots + \frac{1}{v'_n + x''}$$

But that is impossible unless  $v'_1 = v_1, v'_2 = v_2, \dots, v'_n = v_n$ .

A number  $M$  being selected at random between 0 and 1 and converted into a continued fraction

$$M = \frac{1}{v_1} + \frac{1}{v_2} + \dots + \frac{1}{v_n + \xi}$$

if the quantity  $\xi$  turns out to be contained between 0 and  $x < 1$ ,  $M$  must belong to one (and only one) of the intervals  $\delta(v_1, v_2, \dots, v_n, x)$  corresponding to one of all the possible systems of  $n$  positive integers  $v_1, v_2, \dots, v_n$ . Since  $M$  has a uniform distribution of probability and

since the length of the interval  $\delta(v_1, v_2, \dots, v_n, x)$  is

$$(-1)^n \left[ \frac{1}{v_1 + \frac{1}{v_2 + \dots + \frac{1}{v_n + x}}} - \frac{1}{v_1 + \frac{1}{v_2 + \dots + \frac{1}{v_n}}} \right]$$

the required probability  $P_n(x)$  will be expressed by the sum

$$P_n(x) = \sum_{v_1, v_2, \dots, v_n} (-1)^n \left[ \frac{1}{v_1 + \frac{1}{v_2 + \dots + \frac{1}{v_n + x}}} - \frac{1}{v_1 + \frac{1}{v_2 + \dots + \frac{1}{v_n}}} \right]$$

extended over all systems of positive integers  $v_1, v_2, \dots, v_n$ . In general let

$$\frac{P_i}{Q_i} = \frac{1}{v_1 + \frac{1}{v_2 + \dots + \frac{1}{v_i}}} \quad (i = 1, 2, \dots, n)$$

be a convergent to the continued fraction

$$\frac{1}{v_1 + \frac{1}{v_2 + \dots + \frac{1}{v_n}}}$$

Then the above expression for  $P_n(x)$  can be exhibited in a more convenient form:

$$(1) \quad P_n(x) = \sum_{v_1, v_2, \dots, v_n} (-1)^n \left[ \frac{P_n + xP_{n-1}}{Q_n + xQ_{n-1}} - \frac{P_n}{Q_n} \right].$$

By the very definition of  $P_n(x)$  we must have  $P_n(1) = 1$ ; hence the important relation

$$(2) \quad \sum \frac{1}{Q_n(Q_n + Q_{n-1})} = 1.$$

This result can also be established directly by resorting to the original expression of  $P_n(1)$  and performing summation first with respect to  $v_1$ , then with respect to  $v_2$ , etc.

Relation (2) can be interpreted as follows: Let  $\delta$  in general be the length of an interval  $\delta(v_1, v_2, \dots, v_n, 1)$ . Then

$$\Sigma \delta = 1$$

summation being extended over the (enumerable) set of intervals  $\delta$ .

**3. The Derivative of  $P_n(x)$ .** In attempting to show that  $P_n(x)$  tends uniformly to a limit function as  $n \rightarrow \infty$  it is easier to begin with its derivative  $p_n(x)$ . Series

$$\sum \frac{1}{(Q_n + xQ_{n-1})^2}$$

obtained by formal derivation of (1) is uniformly convergent in the interval (0, 1). For

$$Q_n > \frac{Q_n + Q_{n-1}}{2}$$

whence

$$\frac{1}{(Q_n + xQ_{n-1})^2} < \frac{2}{Q_n(Q_n + Q_{n-1})}$$

and the series

$$\sum \frac{2}{Q_n(Q_n + Q_{n-1})} = 2$$

is convergent. Hence

$$\frac{dP_n(x)}{dx} = p_n(x) = \sum \frac{1}{(Q_n + xQ_{n-1})^2}.$$

Since

$$Q_n = v_n Q_{n-1} + Q_{n-2}$$

we have

$$p_n(x) = \sum_{v_1, v_2, \dots, v_n} \frac{1}{\left(Q_{n-1} + \frac{1}{v_n + x} Q_{n-2}\right)^2} \cdot \frac{1}{(v_n + x)^2}$$

and, performing summation with respect to  $v_1, v_2, \dots, v_{n-1}$  for constant  $v_n$

$$\sum_{v_1, v_2, \dots, v_{n-1}} \frac{1}{\left(Q_{n-1} + \frac{1}{v_n + x} Q_{n-2}\right)^2} = p_{n-1}\left(\frac{1}{v_n + x}\right),$$

whence

$$p_n(x) = \sum_{v_n=1}^{\infty} p_{n-1}\left(\frac{1}{v_n + x}\right) \frac{1}{(v_n + x)^2}$$

or else

$$(3) \quad p_n(x) = \sum_{v=1}^n p_{n-1} \left( \frac{1}{v+x} \right) \frac{1}{(v+x)^2}$$

—an important recurrence relation which permits determining completely the sequence of functions

$$p_1(x), p_2(x), \dots$$

starting with  $p_0(x) = 1$ .

**4. Discussion of a More General Recurrence Relation.** In discussing relation (3) the fact that  $p_0(x) = 1$  is of no consequence. We may start with any function  $f_0(x)$  subject to some natural limitations, and form a sequence

$$f_1(x), f_2(x), f_3(x), \dots$$

by means of the recurrence relation

$$(4) \quad f_n(x) = \sum_{v=1}^n f_{n-1} \left( \frac{1}{v+x} \right) \frac{1}{(v+x)^2}$$

The following properties of  $f_n(x)$  follow easily from this relation:

a. If

$$f_0(x) = \frac{a}{1+x}$$

then

$$f_n(x) = \frac{a}{1+x}; \quad n = 1, 2, 3, \dots$$

For

$$f_1(x) = a \sum_{v=1}^{\infty} \left( \frac{1}{v+x} - \frac{1}{v+x+1} \right) = \frac{a}{1+x}$$

whence the general statement follows immediately.

b. If

$$\frac{m}{1+x} \leq f_0(x) \leq \frac{M}{1+x}$$

then

$$\frac{m}{1+x} \leq f_n(x) \leq \frac{M}{1+x}$$

Follows from (a) and equation (4) itself.

As a corollary we have: Let  $M_n$  and  $m_n$  be the precise upper and lower bounds of

$$(1+x)f_n(x) \quad (n = 0, 1, 2, \dots)$$

in the interval  $0 \leq x \leq 1$ . Then

$$\begin{aligned} M_0 &\geq M_1 \geq M_2 \geq \dots \\ m_0 &\leq m_1 \leq m_2 \leq \dots \end{aligned}$$

c. We have

$$\begin{aligned} \int_0^1 f_n(x) dx &= \sum_{v=1}^n \int_0^1 f_{n-1} \left( \frac{1}{v+x} \right) \frac{dx}{(v+x)^2} = \\ &= \int_1^\infty f_{n-1} \left( \frac{1}{u} \right) \frac{du}{u^2} = \int_0^1 f_{n-1}(x) dx = \int_0^1 f_0(x) dx. \end{aligned}$$

d. The following relations can easily be established by mathematical induction:

$$\begin{aligned} f_n(x) &= \sum f_0 \left( \frac{P_n + xP_{n-1}}{Q_n + xQ_{n-1}} \right) \frac{1}{(Q_n + xQ_{n-1})^2} \\ f_{2n}(x) &= \sum f_n \left( \frac{P_n + xP_{n-1}}{Q_n + xQ_{n-1}} \right) \frac{1}{(Q_n + xQ_{n-1})^2} \\ f_{3n}(x) &= \sum f_{2n} \left( \frac{P_n + xP_{n-1}}{Q_n + xQ_{n-1}} \right) \frac{1}{(Q_n + xQ_{n-1})^2} \\ &\dots \end{aligned}$$

Let us suppose now that the function  $f_0(x)$  defined in the interval

$$0 \leq x \leq 1$$

possesses a derivative everywhere in this interval and let  $\mu_0$  be an upper bound of  $|f'_0(x)|$  while  $M$  is an upper bound of  $|(1+x)f_0(x)|$ . Then by property (b)

$$|f_n(x)| \leq M; \quad |f_{2n}(x)| \leq M; \quad |f_{3n}(x)| \leq M, \dots$$

The function  $f_n(x)$  represented by the series

$$f_n(x) = \sum f_0(u) \frac{1}{(Q_n + xQ_{n-1})^2}$$

where  $u$  stands for

$$\frac{P_n + xP_{n-1}}{Q_n + xQ_{n-1}}$$

has a derivative; for the series obtained by a formal differentiation

$$f'_n(x) = \sum f'_0(u) \frac{(-1)^n}{(Q_n + xQ_{n-1})^4} - 2 \sum f_0(u) \frac{Q_{n-1}}{(Q_n + xQ_{n-1})^3}$$

is uniformly convergent and represents  $f'_n(x)$ . Now

$$\frac{Q_{n-1}}{(Q_n + xQ_{n-1})^3} < \frac{1}{Q_n^2}$$

and

$$Q_n^2 > \frac{Q_n(Q_n + Q_{n-1})}{2}.$$

Hence

$$\left| 2 \sum f_0(u) \frac{Q_{n-1}}{(Q_n + xQ_{n-1})^3} \right| < 4M \sum \frac{1}{Q_n(Q_n + Q_{n-1})} = 4M$$

by virtue of (2). On the other hand, the inequality

$$Q_n(Q_n + Q_{n-1}) = (v_n Q_{n-1} + Q_{n-2})[(v_n + 1)Q_{n-1} + Q_{n-2}] > 2Q_{n-1}(Q_{n-1} + Q_{n-2})$$

holding for  $n \geq 2$  together with an evident inequality

$$Q_1(Q_1 + Q_0) \geq 2$$

shows that

$$Q_n(Q_n + Q_{n-1}) > 2^n \quad (n \geq 2).$$

Thus

$$(Q_n + xQ_{n-1})^4 > Q_n^2 \cdot Q_n^2 > \frac{Q_n(Q_n + Q_{n-1})}{2} \cdot \frac{Q_n(Q_n + Q_{n-1})}{2} > 2^{n-2} Q_n(Q_n + Q_{n-1})$$

and consequently

$$\left| \sum f'_0(u) \frac{(-1)^n}{(Q_n + xQ_{n-1})^4} \right| < \frac{\mu_0}{2^{n-2}}.$$

Hence, we may conclude that

$$\mu_1 = \frac{\mu_0}{2^{n-2}} + 4M$$

is an upper bound of  $|f'_n(x)|$ . Similarly, starting with the second equation in (d), we find that

$$\mu_2 = \frac{\mu_1}{2^{n-2}} + 4M$$



is an upper bound of  $|f'_{2n}(x)|$ , and so forth. In general, the recurrence relation

$$\mu_k = \frac{\mu_{k-1}}{2^{n-2}} + 4M \quad (k = 1, 2, 3, \dots)$$

determines upper bounds of

$$|f'_n(x)|, |f'_{2n}(x)|, |f'_{4n}(x)|, \dots$$

It is easy to see that in general

$$\mu_k < \frac{\mu_0}{2^{k(n-2)}} + \frac{4M}{1 - 2^{-(n-2)}}$$

so that for sufficiently large  $n$

$$\mu_k < 5M.$$

### 5. Main Inequalities. Let

$$\varphi_0(x) = f_0(x) - \frac{m_0}{1+x}.$$

Then

$$\begin{aligned} f_n(x) - \frac{m_0}{1+x} &= \varphi_n(x) = \\ &= \sum \varphi_0(u) \frac{1}{(Q_n + xQ_{n-1})^2} > \frac{1}{2} \sum \varphi_0(u) \frac{1}{Q_n(Q_n + Q_{n-1})}. \end{aligned}$$

Since the intervals  $\delta$  defined at the end of Sec. 2 do not overlap and cover completely the whole interval  $(0, 1)$ , we may write:

$$l = \frac{1}{2} \int_0^1 \varphi_0(x) dx = \frac{1}{2} \sum \int_{(\delta)} \varphi_0(x) dx = \frac{1}{2} \sum \varphi_0(u_1) \frac{1}{Q_n(Q_n + Q_{n-1})},$$

the latter part following from the mean value theorem and  $u_1$  being a number contained within the interval  $\delta$ . By subtraction we find

$$f_n(x) - \frac{m_0}{1+x} - l > \frac{1}{2} \sum [\varphi_0(u) - \varphi_0(u_1)] \frac{1}{Q_n(Q_n + Q_{n-1})}$$

and, since both  $u$  and  $u_1$  belong to the same interval  $\delta$ ,

$$\varphi_0(u) - \varphi_0(u_1) > -\frac{\mu_0 + m_0}{Q_n(Q_n + Q_{n-1})} > -\frac{\mu_0 + m_0}{2^n}.$$

Consequently,

$$f_n(x) - \frac{m_0}{1+x} - l > -\frac{\mu_0 + m_0}{2^{n+1}},$$

and a fortiori

$$f_n(x) > \frac{m_0 + l - 2^{-n}(\mu_0 + m_0)}{1 + x}.$$

It follows that

$$(5) \quad m_1 \geq m_0 + l - 2^{-n}(\mu_0 + m_0).^*$$

In a similar way, considering the function

$$\psi_0(x) = \frac{M_0}{1 + x} - f_0(x)$$

and setting

$$l_1 = \frac{1}{2} \int_0^1 \psi_0(x) dx,$$

we shall have

$$f_n(x) < \frac{M_0 - l_1 + 2^{-n}(\mu_0 + M_0)}{1 + x},$$

whence

$$(6) \quad M_1 \leq M_0 - l_1 + 2^{-n}(\mu_0 + M_0).$$

Further, from (5) and (6)

$$M_1 - m_1 \leq M_0 - m_0 + 2^{-n+1}(\mu_0 + M_0) - l - l_1.$$

But

$$l + l_1 = \frac{1}{2} \log 2 \cdot (M_0 - m_0) = (1 - k)(M_0 - m_0); \quad k < 0.66,$$

so that finally

$$M_1 - m_1 < k(M_0 - m_0) + 2^{-n+1}(\mu_0 + M_0).$$

Starting with  $f_n(x)$ ,  $f_{2n}(x)$ , . . . instead of  $f_0(x)$ , in a similar way we find

$$M_2 - m_2 < k(M_1 - m_1) + 2^{-n+1}(\mu_1 + M_1)$$

$$M_3 - m_3 < k(M_2 - m_2) + 2^{-n+1}(\mu_2 + M_2)$$

$$M_n - m_n < k(M_{n-1} - m_{n-1}) + 2^{-n+1}(\mu_{n-1} + M_{n-1}).$$

From these inequalities it follows that

$$M_n - m_n < (M_0 - m_0)k^n + 2^{-n+1} [\mu_0 k^{n-1} + \mu_1 k^{n-2} + \cdots + \mu_{n-1} + M_0 k^{n-1} + M_1 k^{n-2} + \cdots + M_{n-1}].$$

Without losing anything in generality, we may suppose that  $f_0(x)$  is a positive function. Then

\*  $M_i$ ,  $m_i$  are used here with the same meaning as  $M_{ni}$ ,  $m_{ni}$  in Sec. 4.

$$M_k \leq M_0, \quad \mu_k < 5M_0 \quad (k = 1, 2, 3, \dots)$$

at least for sufficiently large  $n$ . Owing to these inequalities we shall have

$$(7) \quad M_n - m_n < (M_0 - m_0)k^n + \mu_0 \left(\frac{k}{2}\right)^{n-1} + \frac{6M_0}{(1-k)2^{n-1}}.$$

This inequality shows that sequences

$$\begin{aligned} M_0 &\geq M_1 \geq M_2 \geq \dots \\ m_0 &\leq m_1 \leq m_2 \leq \dots \end{aligned}$$

approach a common limit  $a$ . The following method can be used to find the value of this limit. Let  $N$  be an arbitrary sufficiently large integer and  $n$  the integer defined by

$$n^2 \leq N < (n+1)^2.$$

Then

$$\frac{m_n}{1+x} \leq f_{nn}(x) \leq \frac{M_n}{1+x},$$

and therefore

$$\frac{m_n}{1+x} \leq f_N(x) \leq \frac{M_n}{1+x}.$$

The last inequality permits presenting  $f_N(x)$  thus:

$$(8) \quad f_N(x) = \frac{a}{1+x} + \theta(M_n - m_n); \quad |\theta| < 1,$$

whence

$$\int_0^1 f_N(x) dx = \int_0^1 f_0(x) dx = a \log 2 + \theta'(M_n - m_n), \quad |\theta'| < 1,$$

and, because  $M_n - m_n$  ultimately becomes as small as we please in absolute value,

$$a \log 2 = \int_0^1 f_0(x) dx.$$

Equation (8) shows clearly that the sequence of functions

$$f_0(x), f_1(x), f_2(x), \dots$$

defined by the recurrence relation (4) approaches uniformly the limit function

$$\frac{a}{1+x}$$

where

$$a = \frac{1}{\log 2} \int_0^1 f_0(x) dx.$$

**6. Solution of the Gaussian Problem.** It suffices to apply the preceding considerations to the case  $f_0(x) = p_0(x) = 1$ . In this case  $M_0 = 2$ ,  $m_0 = 1$ ,  $\mu_0 = 0$  and

$$a = \frac{1}{\log 2}.$$

Consequently,

$$p_N(x) = \frac{1}{(1+x) \log 2} + \theta \left( k^n + \frac{3}{(1-k) \cdot 2^{n-3}} \right); \quad |\theta| < 1$$

where  $n = [\sqrt{N}]$ . It suffices to integrate this expression between limits 0 and  $t < 1$  to find

$$P_N(t) = \frac{\log(1+t)}{\log 2} + \lambda \left( k^n + \frac{3}{(1-k)2^{n-3}} \right); \quad |\lambda| < t.$$

As  $N \rightarrow \infty$

$$P_N(t) \rightarrow \frac{\log(1+t)}{\log 2}$$

as stated by Gauss. Moreover,

$$\left| P_N(t) - \frac{\log(1+t)}{\log 2} \right| < t \left( k^n + \frac{3}{(1-k)2^{n-3}} \right)$$

for sufficiently large, but finite  $N$ .

TABLE OF THE PROBABILITY INTEGRAL

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-t^2} dt$$

$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
0.00	0.0000	0.65	0.2422	1.30	0.4032	1.95	0.4744
0.01	0.0040	0.66	0.2454	1.31	0.4049	1.96	0.4750
0.02	0.0080	0.67	0.2486	1.32	0.4066	1.97	0.4756
0.03	0.0120	0.68	0.2517	1.33	0.4082	1.98	0.4761
0.04	0.0160	0.69	0.2549	1.34	0.4099	1.99	0.4767
0.05	0.0199	0.70	0.2580	1.35	0.4115	2.00	0.4772
0.06	0.0239	0.71	0.2611	1.36	0.4131	2.02	0.4783
0.07	0.0279	0.72	0.2642	1.37	0.4147	2.04	0.4793
0.08	0.0319	0.73	0.2673	1.38	0.4162	2.06	0.4803
0.09	0.0359	0.74	0.2703	1.39	0.4177	2.08	0.4812
0.10	0.0398	0.75	0.2734	1.40	0.4192	2.10	0.4821
0.11	0.0438	0.76	0.2764	1.41	0.4207	2.12	0.4830
0.12	0.0478	0.77	0.2794	1.42	0.4222	2.14	0.4838
0.13	0.0517	0.78	0.2823	1.43	0.4236	2.16	0.4846
0.14	0.0557	0.79	0.2852	1.44	0.4251	2.18	0.4854
0.15	0.0596	0.80	0.2881	1.45	0.4265	2.20	0.4861
0.16	0.0636	0.81	0.2910	1.46	0.4279	2.22	0.4868
0.17	0.0675	0.82	0.2939	1.47	0.4292	2.24	0.4875
0.18	0.0714	0.83	0.2967	1.48	0.4306	2.26	0.4881
0.19	0.0753	0.84	0.2995	1.49	0.4319	2.28	0.4887
0.20	0.0793	0.85	0.3023	1.50	0.4332	2.30	0.4893
0.21	0.0832	0.86	0.3051	1.51	0.4345	2.32	0.4898
0.22	0.0871	0.87	0.3078	1.52	0.4357	2.34	0.4904
0.23	0.0910	0.88	0.3106	1.53	0.4370	2.36	0.4909
0.24	0.0948	0.89	0.3133	1.54	0.4382	2.38	0.4913
0.25	0.0987	0.90	0.3159	1.55	0.4394	2.40	0.4918
0.26	0.1026	0.91	0.3186	1.56	0.4406	2.42	0.4922
0.27	0.1064	0.92	0.3212	1.57	0.4418	2.44	0.4927
0.28	0.1103	0.93	0.3238	1.58	0.4429	2.46	0.4931
0.29	0.1141	0.94	0.3264	1.59	0.4441	2.48	0.4934
0.30	0.1179	0.95	0.3289	1.60	0.4452	2.50	0.4938
0.31	0.1217	0.96	0.3315	1.61	0.4463	2.52	0.4941
0.32	0.1255	0.97	0.3340	1.62	0.4474	2.54	0.4945
0.33	0.1293	0.98	0.3365	1.63	0.4484	2.56	0.4948
0.34	0.1331	0.99	0.3389	1.64	0.4495	2.58	0.4951
0.35	0.1368	1.00	0.3413	1.65	0.4505	2.60	0.4953
0.36	0.1406	1.01	0.3438	1.66	0.4515	2.62	0.4956
0.37	0.1443	1.02	0.3461	1.67	0.4525	2.64	0.4959
0.38	0.1480	1.03	0.3485	1.68	0.4535	2.66	0.4961
0.39	0.1517	1.04	0.3508	1.69	0.4545	2.68	0.4963
0.40	0.1554	1.05	0.3531	1.70	0.4554	2.70	0.4965
0.41	0.1591	1.06	0.3554	1.71	0.4564	2.72	0.4967
0.42	0.1628	1.07	0.3577	1.72	0.4573	2.74	0.4969
0.43	0.1664	1.08	0.3599	1.73	0.4582	2.76	0.4971
0.44	0.1700	1.09	0.3621	1.74	0.4591	2.78	0.4973
0.45	0.1736	1.10	0.3643	1.75	0.4599	2.80	0.4974
0.46	0.1772	1.11	0.3665	1.76	0.4608	2.82	0.4976
0.47	0.1808	1.12	0.3686	1.77	0.4616	2.84	0.4977
0.48	0.1844	1.13	0.3708	1.78	0.4625	2.86	0.4979
0.49	0.1879	1.14	0.3729	1.79	0.4633	2.88	0.4980
0.50	0.1915	1.15	0.3749	1.80	0.4641	2.90	0.4981
0.51	0.1950	1.16	0.3770	1.81	0.4649	2.92	0.4982
0.52	0.1985	1.17	0.3790	1.82	0.4656	2.94	0.4984
0.53	0.2019	1.18	0.3810	1.83	0.4664	2.96	0.4985
0.54	0.2054	1.19	0.3830	1.84	0.4671	2.98	0.4986
0.55	0.2088	1.20	0.3849	1.85	0.4678	3.00	0.49865
0.56	0.2123	1.21	0.3869	1.86	0.4686	3.20	0.49831
0.57	0.2157	1.22	0.3888	1.87	0.4693	3.40	0.49806
0.58	0.2190	1.23	0.3907	1.88	0.4699	3.60	0.49841
0.59	0.2224	1.24	0.3925	1.89	0.4706	3.80	0.498928
0.60	0.2257	1.25	0.3944	1.90	0.4713	4.00	0.499668
0.61	0.2291	1.26	0.3962	1.91	0.4719	4.50	0.499967
0.62	0.2324	1.27	0.3980	1.92	0.4726	5.00	0.499997
0.63	0.2357	1.28	0.3997	1.93	0.4732		
0.64	0.2389	1.29	0.4015	1.94	0.4738		

## INDEX

- A**
- Arrangements, 18
- B**
- Bayes' formula (theorem), 61  
 Bernoulli criterion, 5  
 Bernoulli theorem, 96  
 Bernoulli trials, 45  
 Bernstein, S., inequality, 205  
 Bertrand's paradox, 251  
 Buffon's needle problem, 113, 251  
     Barbier's solution of, 253
- C**
- Cantelli's theorem, 101  
 Cauchy's distribution, 243, 275  
 Characteristic function, composition of,  
     275  
     of distribution, 240, 264  
 Coefficient, correlation, 339  
     divergence, 212, 214, 216  
 Combinations, 18  
 Compound probability, theorem of, 31  
 Continued fractions, 358, 361, 396  
     Markoff's method of, 52  
 Continuous variables, 235  
 Correlation, normal (*see* Normal correlation)  
 Correlation coefficient, distribution of,  
     339
- D**
- Difference equations, ordinary, 75, 78  
     partial, 84  
 Dispersion, definition, 172  
     of sums, 173  
 Distribution, Cauchy's, 243, 275  
     characteristic function of, 264  
     of correlation coefficient, 339  
     Distribution, determination of, 271  
         equivalent point, 369  
         general concept of, 263  
         normal (Gaussian), 243  
         Poisson's, 279  
         "Student's," 339  
     Distribution function of probability,  
         239, 263  
     Divergence coefficient, empirical, 212  
         Lexis' case, 214  
         Poisson's case, 214  
         theoretical, 212  
         Tschuprow's theorem, 216
- E**
- Elementary errors, hypothesis of, 296  
 Ellipses of equal probability, 311, 328  
 Estimation of error term, 295  
 Euler's summation formula, 177, 201,  
     303, 347  
 Events, compound, 29  
     contingent, 3  
     dependent, 33  
     equally likely, 4, 5, 7  
     exhaustive, 6  
     future, 65  
     incompatible, 37  
     independent, 32, 33  
     mutually exclusive, 6, 27  
     opposite, 29  
 Expectation, mathematical, 161  
     of a product, 171  
     of a sum, 165
- F**
- Factorials, 349  
 Fourier theorem, 241  
 French lottery, 19, 108  
 Frequency, 96  
 Fundamental lemma (*see* Limit theorem)  
 Fundamental theorem (*see* Tshchysheff-  
     Markoff theorem)

- G
- Gaussian distribution, 243  
 Gaussian problem, 396  
 Generating function of probabilities,  
 47, 78, 85, 89, 93, 94
- H
- Hermite polynomials, 72  
 Hypothesis of elementary errors, 296
- I
- Independence, definition of, 32, 33
- K
- Khintchine (*see* Law of large numbers)  
 Kolmogoroff (*see* Law of large numbers;  
 Strong law of large numbers)
- L
- Lagrange's series, 84, 150  
 Laplace-Liapounoff (*see* Limit theorem)  
 Laplace's problem, 255  
 Laurent's series, 87, 148  
 Law of large numbers, generalization  
 by Markoff, 191  
 for identical variables (Khintchine),  
 195  
 Kolmogoroff's lemma, 201  
 theorem, 185  
 Tshebysheff's lemma, 182  
 Law of repeated logarithm, 204  
 Law of succession, 69  
 Lexis' case, 214  
 Liapounoff condition (*see* Limit theorem)  
 Liapounoff inequality, 265  
 Limit theorem, Bernoullian case, 131  
 for sums of independent vectors, 318,  
 323, 325, 326  
 fundamental lemma, 284  
 Laplace-Liapounoff, 284  
 Line of regression, 314  
 Lottery, French (*see* French lottery)
- M
- Marbe's problem, 231  
 Markoff's theorem, infinite dispersion,  
 191
- Markoff's theorem, for simple chains, 301  
 Markoff-Tshebysheff theorem (*see*  
 Tshebysheff-Markoff theorem)  
 Mathematical expectation, definition of,  
 161  
 of a product, 171  
 of a sum, 165  
 Mathematical probability, definition of,  
 6  
 Moments, absolute, 240, 264  
 inequalities for, 264  
 method of (Markoff's), 356*ff.*
- N
- Normal correlation, 313  
 origin of, 327  
 Normal distribution, Gaussian, 243  
 two-dimensional, 308
- P
- Pearson's " $\chi^2$ -test," 327  
 Permutations, 18  
 Point, of continuity, 261, 356  
 of increase, 262, 356  
 Poisson series, 182, 293  
 Poisson's case, 214  
 Poisson's distribution, 279  
 Poisson's formula, 137  
 Poisson's theorem, 208, 294  
 Polynomials, Hermite (*see* Hermite)  
 Probability, approximate evaluation of,  
 by Markoff's method, 52  
 compound, 29, 31  
 conditional, 33  
 definition (classical) of, 6  
 total, 27, 28  
 Probability integral, 128  
 table of, 407
- R
- Relative frequency, 96  
 Runs, problem of, 77
- S
- Simple chains, 74, 223, 297  
 Markoff's theorem for, 301  
 Standard deviation, 173

Stieltjes' integrals, 261  
Stirling's formula, 349  
Stochastic variables, 161  
Strong law of large numbers (Kolmogoroff), 202  
"Student's" distribution, 339

## T

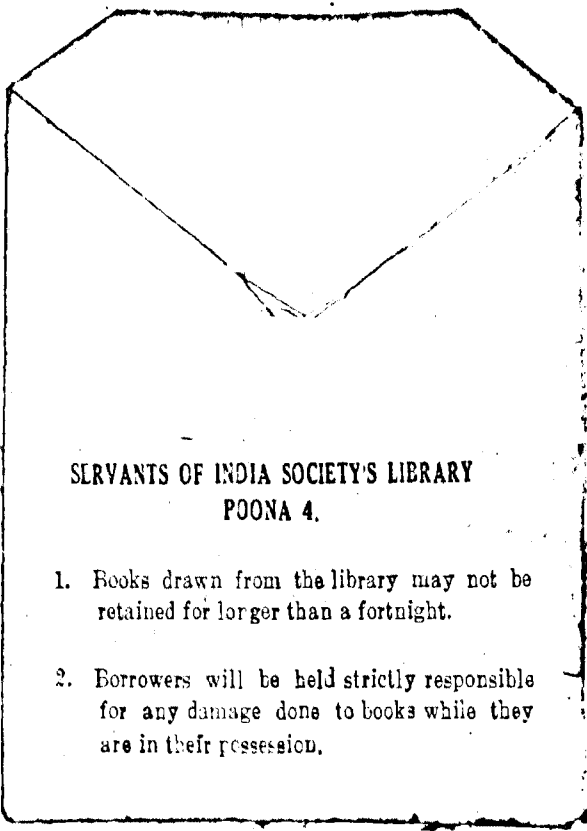
Table of probability integral, 407  
Tests of significance, 331  
Total probability, theorem of, 27, 28  
Trials, dependent, independent, repeated, 44, 45

Tschuprow (*see* Divergence coefficient)  
Tshebysheff-Markoff theorem, fundamental, 304, 384  
    application, 388  
Tshebysheff's inequalities, 373  
Tshebysheff's inequality, 204  
Tshebysheff's lemma, 182  
Tshebysheff's problem, 199

## V

Variables, continuous, 235  
    independent, 171  
    stochastic, 161  
Vectors (*see* Limit theorem)





**SERVANTS OF INDIA SOCIETY'S LIBRARY  
POONA 4.**

1. Books drawn from the library may not be retained for longer than a fortnight.
2. Borrowers will be held strictly responsible for any damage done to books while they are in their possession.