T-1691

INTEGRATION THEORY AND APPLICATIONS OF LEBESGUE-STIELTJES AND RIEMANN-STIELTJES INTEGRATION

By

Richard C. Walker

11025/64

A Thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Mathematics).

Signed: Richard C. Walker
Richard C. Walker

Golden, Colorado
Date: April 29, 1975

Approved: John O. Kork
Dr. John O. Kork
Thesis Advisor

Robert A. Walsh
Dr. Robert A. Walsh
Head, Department of Mathematics

Golden, Colorado
Date: April 29, 1975

ABSTRACT

A measure-theoretic development of the general integral is presented. Its specializations to the Lebesgue-Stieltjes, Riemann-Stieltjes, Lebesgue, and Riemann integrals are discussed. Particular emphasis is placed on the role of the Stieltjes measure function in determining the σ-field and Lebesgue-Stieltjes measure with respect to which the Lebesgue-Stieltjes integral may be defined.

Riemann-Stieltjes integration is applied to basic physical problems, including a gravitational attraction problem involving polar coordinates. The foundations of probability theory in Lebesgue-Stieltjes measure and integration are discussed and an application of Lebesgue-Stieltjes integration to the Poisson stochastic process is given.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

INDEX OF NOTATION

N	set of natural numbers
R	set of real numbers
R^k	Euclidian k-dimensional space
$\{x: Q(x)\}$	set of x such that $Q(x)$ is true
$\{x \mid Q(x)\}$	set of x given $Q(x)$
$x \in A$	x is an element of A
R^+	$\{x: 0 \leq x < +\infty\}$
R^*	$R \cup \{-\infty, +\infty\}$; extended real numbers
$A \times B$	Cartesian product of sets
$\langle x \rangle$	greatest integer in x
$\{x_i\}_{i=1}^{n}$	finite sequence
$\{x_i\}_{i=1}^{\infty}$	infinite sequence
$\bigtimes_{i=1}^{n} A_i$ or $\prod_{i=1}^{n} A_i$	Cartesian product
$\prod_{i=1}^{n} x_i$	product of numbers of a sequence
$\mathcal{P}$	semi-ring of half-open intervals
$\mathcal{P}^k$	$\{\bigtimes_{i=1}^{k} A_i : A_i \in \mathcal{P}\}$
$\mathcal{B}$	σ-field of Borel sets in R
$\mathcal{B}^k$	σ-field of Borel sets in R^k
$A - B$	complement of B with respect to A

ϕ	empty set
$\mu(A)$	set functional value
$g \circ f$ or $f(g)$	composition of functions
$X(\omega)$	value of the random variable X
$(x_1,x_2,\ldots,x_n)$	n-tuple
$(X,Y)(t)$	ordered functional value $(X(t),Y(t))$
χ_A	characteristic function of A
$N(t)$	number of events by time t
S_n	time of nth event
$\mathcal{L}^k$	k-dimensional Lebesgue-measurable sets
$\mathcal{L}_F^k$	k-dimensional Lebesgue-Stieltjes measurable sets

INTRODUCTION

The purpose of this dissertation is to provide the reader with a knowledge of the construction of the general integral, its specialization to integrals of the Lebesgue and Lebesgue-Stieltjes types, the relationships between Riemann-Stieltjes and Lebesgue-Stieltjes integrals, and examples of the application of Riemann-Stieltjes and Lebesgue-Stieltjes integration.

Special attention is given to the following concepts: (1) the construction by use of extension theorems of measure spaces over which Lebesgue-Stieltjes integration may be defined, (2) the one-to-one correspondence between a given Lebesgue-Stieltjes measure and its associated Stieltjes measure function, (3) the computationally efficient form of the Riemann-Stieltjes integral when expressed as a Riemann integral plus a summation, and (4) the notation of the Lebesgue-Stieltjes integral by which it explicitly but compactly displays information about the structure of the measure space used in the integral's construction.

The author gratefully acknowledges the assistance of Professor John O. Kork, thesis advisor, for patiently reading the manuscript and making many helpful suggestions.

Special thanks are due Professor Ardel J. Boes for stimulating the author's interest in measure theory and Professor William R. Astle for assisting the author in the selection of a thesis topic and stimulating the author's interest in probability theory.

MEASURE THEORETIC DEVELOPMENT OF THE INTEGRAL

The purpose of this chapter is to discuss the measure - theoretic derivation of the general integral. Theorems and methods of construction introduced in this chapter will provide the means by which the integral will be specialized to the Lebesgue and Lebesgue-Stieltjes types of integrals in subsequent chapters.

This chapter is divided into four sections: Classes of Subsets, Measure, Simple Functions and Measureable Functions, and Definition of the Integral. The first section discusses closure properties of some important classes of subsets. The generation of larger classes of subsets from smaller ones, motivated by the desired extension of set functions defined on the smaller classes, is the topic of the second section. The third section discusses types of point functions for which the integral may be defined. The integral is then defined in the fourth section.

Classes of Subsets

The construction of set functions which will be used in defining the integral relies on properties of relevant classes of subsets.

Definitions of Classes: In the definitions below, given by Kingman and Taylor (1966, p. 15-16), each subset is to be regarded as a subset of some fixed non-empty reference set X.

DEFINITION 2.1: A *semi-ring* is a class $\mathcal{S}$ of subsets such that

(i) $\phi \in \mathcal{S}$,

(ii) $A, B \in \mathcal{S}$ implies $A \cap B \in \mathcal{S}$,

(iii) $A, B \in \mathcal{S}$ implies $A - B = \bigcup_{i=1}^{n} E_i$ for some finite sequence $\{E_i\}_{i=1}^{n}$ of sets in $\mathcal{S}$.

DEFINITION 2.2: A *ring* is a non-empty class $\mathcal{R}$ of subsets such that

(i) $A, B \in \mathcal{R}$ implies $A \cap B \in \mathcal{R}$,

(ii) $A, B \in \mathcal{R}$ implies $(A - B) \cup (B - A) \in \mathcal{R}$.

DEFINITION 2.3: A *field* is a ring $\mathcal{F}$ of subsets which comtains X.

DEFINITION 2.4: A *σ-ring* is a ring $\mathcal{L}$ such that, for each sequence $\{A_i\}_{i=1}^{\infty}$ of sets in $\mathcal{L}$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{L}$.

DEFINITION 2.5: A *σ-field* is a σ-ring $\mathcal{F}$ which contains X.

DEFINITION 2.6: A <u>monotone</u> <u>class</u> is a class $\mathcal{M}$ of subsets such that for each monotone sequence $\{A_i\}_{i=1}^{\infty}$ of sets in $\mathcal{M}$, $\lim_{i\to\infty} A_i \ \varepsilon \ \mathcal{M}$.

<u>Properties</u> <u>of</u> <u>Classes</u> <u>of</u> <u>Subsets</u>: Notions of closure of classes of subsets with respect to the operations of union, intersection, and complementation are valuable in the development of measure theory and concepts of integration.

A ring is closed under finite union and finite symmetric difference. A σ-ring or σ-field is closed under countable union and countable intersection. A field or σ-field is closed under the operation of complementation (Kingman and Taylor, 1966, p. 15-16).

THEOREM 2.1 (Kingman and Taylor, 1966, p. 16): Given any class $\mathcal{C}$ of subsets, there is a unique ring $\mathcal{R}$, σ-ring $\mathcal{L}$, monotone class $\mathcal{M}$, field $\mathcal{F}$, or σ-field $\mathcal{B}$ such that if $\mathcal{R}'$, $\mathcal{L}'$, $\mathcal{M}'$, $\mathcal{F}'$, or $\mathcal{B}'$ are any other ring, σ-ring, monotone class, field, or σ-field containing $\mathcal{R}$, $\mathcal{L}$, $\mathcal{M}$, $\mathcal{F}$, or $\mathcal{B}$, respectively, then

$$\mathcal{R} \subset \mathcal{R}', \quad \mathcal{L} \subset \mathcal{L}', \quad \mathcal{M} \subset \mathcal{M}', \quad \mathcal{F} \subset \mathcal{F}', \quad \mathcal{B} \subset \mathcal{B}'.$$

One says that $\mathcal{R}$, $\mathcal{L}$, $\mathcal{M}$, $\mathcal{F}$, and $\mathcal{B}$ are the ring, σ-ring, monotone class, field, or σ-field, respectively, generated by $\mathcal{C}$.

THEOREM 2.2 (Kingman and Taylor, 1966, p. 18): The monotone class $\mathcal{M}$ generated by a ring $\mathcal{R}$ is the same as the σ-ring generated by $\mathcal{R}$.

Some examples of classes of subsets are now given.

EXAMPLE 2.1: Let $\mathcal{P} = \{(a,b] : a \le b,\ a, b \in R\}$. Then $\mathcal{P}$ is a semi-ring.

Proof:

(i) $(a,a] = \phi \in \mathcal{P}$.

(ii) $(a,b] \in \mathcal{P}$ and $(c,d] \in \mathcal{P}$ implies $(a,b] \cap (c,d] \in \mathcal{P}$ since

$$(a,b] \cap (c,d] = \begin{cases} \phi & ;\ a=b \text{ or } c=d \text{ or } b \le c \text{ or } d \le a, \\ (a,d] & ;\ c \le a \le d \le b, \\ (c,b] & ;\ a \le c \le b \le d, \\ (a,b] & ;\ c \le a \le b \le d, \\ (c,d] & ;\ a \le c \le d \le b. \end{cases}$$

Therefore, $(a,b] \cap (c,d] \in \mathcal{P}$ since the conditions specified for each set on the right side of the above equality exhaust all possibilities. These sets are contained in $\mathcal{P}$ by definition of $\mathcal{P}$ and by (i).

(iii) $(a,b] \in \mathcal{P}$ and $(c,d] \in \mathcal{P}$ implies $(a,b] - (c,d] \in \mathcal{P}$ since

$$\begin{aligned}(a,b] - (c,d] &= (a,b] \cap (R-(c,d]) = \\ &= (a,b] \cap ((-\infty,c] \cup (d,\infty)) = \\ &= ((a,b] \cap (-\infty,c]) \cup ((a,b] \cap (d,\infty)) = \\ &= ((a,b] \cap (e,c]) \cup ((a,b] \cap (d,f]) \text{ ; } e \le a, \ f \ge b.\end{aligned}$$

The last step is valid, since $a \varepsilon R$ and $b \varepsilon R$ imply the existence of $e \varepsilon R$ and $f \varepsilon R$ such that $e \le a$ and $f \ge b$. Thus

$$(a,b] \cap (-\infty,c] = (a,b] \cap (e,c] \quad \text{and}$$

$$(a,b] \cap (d,\infty) = (a,b] \cap (d,f].$$

Each of these sets is contained in $\mathcal{P}$ by (ii); therefore their union is a finite union of sets from $\mathcal{P}$ as required by Definition 2.1(iii).

Thus (i), (ii), and (iii) imply $\mathcal{P}$ is a semi-ring.

EXAMPLE 2.2: Let $\mathcal{R} = \{A : A$ is a finite union of disjoint sets of $\mathcal{P}\}$. It may be shown that $\mathcal{R}$ is the ring generated by $\mathcal{P}$. If the reference set is R, $\mathcal{R}$ is also a field.

Measure

DEFINITION 2.7 (Royden, 1968, p. 52): A set function is a function which associates with each set contained in some class of sets an extended real number. If f is a set function defined on a class $\mathcal{C}$ of subsets, one writes $f:\mathcal{C} \to R^*$.

Let us now proceed toward the definition of measure, a type of set function in terms of which the integral may be defined.

Definition of measure: For completeness of the text and for future reference the definitions of outer measure and measurable sets are included in this discussion.

DEFINITION 2.8 (Kingman and Taylor, 1966, p. 59): An outer measure on the set X is a set function μ^* defined on the collection $\mathcal{C}$ of all subsets of X such that

(i) $\mu^*(\phi) = 0$,

(ii) $B \in \mathcal{C}$ and $A \subset B$ imply $\mu^*(A) \leq \mu^*(B)$ (μ^* is monotone),

(iii) for each sequence $\{A_n\}_{n=1}^{\infty}$ from $\mathcal{C}$, $A \subset \bigcup_{n=1}^{\infty} A_n$ implies $\mu^*(A) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$ (μ^* is countably subadditive).

DEFINITION 2.9 (Kingman and Taylor, 1966, p. 74): Suppose μ^* is an outer measure on X. A set E is said to be μ^*-measurable if for each $A \subset X$

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A - E).$$

DEFINITION 2.10 (Kingman and Taylor, 1966, p. 54): A measure on a collection $\mathcal{C}$ of sets is a set function $\mu : \mathcal{C} \rightarrow R^*$ such that

(i) $\mu(\phi) = 0$,

(ii) for each sequence $\{A_n\}_{n=1}^{\infty}$ of disjoint sets from $\mathcal{C}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$, $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ (μ is σ-additive),

(iii) for each $A \in \mathcal{C}$, $\mu(A) \geq 0$ (μ is non-negative).

Note that conditions (i) and (ii) for outer measure imply non-negativity of μ^* and conditions (ii) and (iii) for measure imply monotonicity of μ.

Extension of Measures: The natural domain of definition for a measure is a σ-ring, since if $\{E_i\}_{i=1}^{\infty}$ is a sequence from a ring $\mathcal{C}$, then $\bigcup_{i=1}^{\infty} E_i = \bigcup_{n=1}^{\infty} A_n$ for some sequence $\{A_n\}_{n=1}^{\infty}$ of disjoint sets from the σ-ring $\mathcal{L}(\mathcal{C})$ generated by $\mathcal{C}$. Because each σ-ring is a monotone class, one might ask how convergence of monotone sequences from a ring $\mathcal{C}$ is related to σ-additivity of measures that might be defined on $\mathcal{C}$. The following theorem concerns not only measures, but also σ-additive set functions in general.

THEOREM 2.3 (Adapted from Kingman and Taylor, 1966, p. 56): Suppose $\mathcal{C}$ is a ring and $\mu:\mathcal{C} \to R^*$ is finitely additive.

(i) If μ is σ-additive, then for each $E \in \mathcal{C}$,

$$\lim_{n\to\infty} \mu(I_n) = \lim_{n\to\infty} \mu(D_n) = \mu(E)$$

for each increasing sequence $\{I_n\}_{n=1}^{\infty}$ from $\mathcal{C}$ converging to E and for each decreasing sequence $\{D_n\}_{n=1}^{\infty}$ from $\mathcal{C}$ converging to E such that $\mu(D_n) < \infty$ for some n.

(ii) If, for each $E \in \mathcal{C}$, $\lim_{n\to\infty} \mu(I_n) = \mu(E)$ for each increasing sequence $\{I_n\}_{n=1}^{\infty}$ converging to E, then μ is σ-additive.

(iii) If $\mu(E) < \infty$ for each $E \in \mathcal{C}$ and if

$\lim_{n\to\infty} \mu(E_n) = 0$ for each decreasing sequence $\{E_n\}_{n=1}^{\infty}$ from $\mathcal{C}$ converging to ϕ, then μ is σ-additive.

DEFINITION 2.11 (Kingman and Taylor, 1966, p. 58): Suppose for classes $\mathcal{C} \subset \mathcal{D}$ of subsets there are set functions $\mu:\mathcal{C} \to R^*$ and $\nu:\mathcal{D} \to R^*$. Then ν is an <u>extension</u> of μ (μ is the <u>restriction</u> of ν to $\mathcal{C}$) if for each $A \in \mathcal{C}$, $\nu(A) = \mu(A)$.

THEOREM 2.4 (Kingman and Taylor, 1966, p. 65): If $\mu:\mathcal{C} \to R^+$ is a non-negative finitely additive set function on a semi-ring $\mathcal{C}$, there is a unique extension ν of μ defined on the ring $\mathcal{R}(\mathcal{C})$ generated by $\mathcal{C}$. The set function ν is non-negative and finitely additive on $\mathcal{R}(\mathcal{C})$.

DEFINITION 2.12 (Kingman and Taylor, 1966, p. 59): A set function $\mu:\mathcal{C} \to R^*$ is <u>σ-finite</u> if for each set $E \in \mathcal{C}$ there is a sequence $\{A_n\}_{n=1}^{\infty}$ from $\mathcal{C}$ such that $E \subset \bigcup_{n=1}^{\infty} A_n$ and $\mu(A_n)$ is finite for each n.

THEOREM 2.5 (Extension Theorem) (Kingman and Taylor, 1966, p. 77): Suppose $\mathcal{C}$ is a ring of subsets of X such that $X = \bigcup_{n=1}^{\infty} A_n$ for some sequence $\{A_n\}_{n=1}^{\infty}$ from $\mathcal{C}$. If $\mu:\mathcal{C} \to R^+$ is a measure on $\mathcal{C}$, then there is an extension of μ to a measure ν on the σ-ring $\mathcal{L}(\mathcal{C})$ generated by $\mathcal{C}$. If μ is σ-finite on $\mathcal{C}$, the extension is unique and is σ-finite on $\mathcal{L}(\mathcal{C})$.

DEFINITION 2.12 (Kingman and Taylor, 1966, p. 81): A measure $\mu : \mathcal{C} \rightarrow \{x : 0 \leq x \leq \infty\}$ is a <u>complete measure</u> if for each $A \in \mathcal{C}$ such that $\mu(A) = 0$, $E \subset A$ implies $E \in \mathcal{C}$.

THEOREM 2.6 (Adapted from Kingman and Taylor, 1966, p. 81): If μ is a measure on a σ-ring $\mathcal{C}$ and $\mathcal{L} = \{A : A \in \mathcal{C}$ or A is a subset of some set $B \in \mathcal{C}$ such that $\mu(B) = 0\}$, there is a unique extension ν of μ to $\mathcal{L}$ and ν is a complete measure on $\mathcal{L}$. The σ-ring $\mathcal{L}$ is called the completion of $\mathcal{C}$ with respect to μ.

<u>Simple Functions and Measurable Functions</u>

The next consideration is the types of point functions for which the integral might be defined. A few definitions will provide notational simplicity.

DEFINITION 2.13: A <u>measure space</u> is a set X together with a measure μ defined on a specified σ-field $\mathcal{F}$ of subsets of X. Such a structure is denoted by the triple $(X, \mathcal{F}, \mu)$.

The following concepts will be defined relative to a measure space $(X, \mathcal{F}, \mu)$.

DEFINITION 2.14 (Royden, 1968, p. 68): The <u>characteristic function</u> χ_A of the set A is the function defined by

$$\chi_A(x) = \begin{cases} 1 & ; x \in A, \\ 0 & ; x \notin A. \end{cases}$$

DEFINITION 2.15 (Kingman and Taylor, 1966, p. 102): An $\mathcal{F}$-dissection of the set X is a finite sequence $\{E_i\}_{i=1}^{n}$ of sets from $\mathcal{F}$ such that

(i) The sets E_i $(i = 1, 2, \ldots, n)$ are disjoint,

(ii) $X = \bigcup_{i=1}^{n} E_i$.

DEFINITION 2.16 (Kingman and Taylor, 1966, p. 102): A function $f:X \to R$ is called $\mathcal{F}$-simple if there exists an $\mathcal{F}$-dissection $\{E_i\}_{i=1}^{n}$ of X and a sequence $\{c_i\}_{i=1}^{n}$ from R such that

$$f(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x) \quad ; x \in X.$$

If only one field is being considered, one speaks of functions as being simple rather than $\mathcal{F}$-simple.

DEFINITION 2.17 (Kingman and Taylor, 1966, p. 103): A set B is in the class $\mathcal{B}^*$ of Borel sets in R^* if it is the union of a Borel set in R with a subset of $\{-\infty, +\infty\}$.

DEFINITION 2.18 (Kingman and Taylor, 1966, p. 103): A function $f:X \to R^*$ is $\mathcal{F}$-measurable if $f^{-1}(B) \in \mathcal{F}$ for each set $B \in \mathcal{B}^*$.

If only one field is being considered, one speaks of functions as being measurable rather than $\mathcal{F}$-measurable.

To test for measurability of a function f, one needs only to check the inverse images of sets from a smaller class which generates $\mathcal{B}^*$, as indicated by the following theorem.

THEOREM 2.7 (Kingman and Taylor, 1966, p. 103): A function $f:X \to R^*$ is $\mathcal{F}$-measurable if and only if any one of the following conditions is satisfied:

(i) $\{x : f(x) \leq c\} \in \mathcal{F}$ for each $c \in R$,

(ii) $\{x : f(x) > c\} \in \mathcal{F}$ for each $c \in R$,

(iii) $\{x : f(x) \geq c\} \in \mathcal{F}$ for each $c \in R$,

(iv) $\{x : f(x) < c\} \in \mathcal{F}$ for each $c \in R$.

COROLLARY (Kingman and Taylor, 1966, p. 104): Any $\mathcal{F}$-simple function is $\mathcal{F}$-measurable.

THEOREM 2.8 (Kingman and Taylor, 1966, p. 104): Any non-negative $\mathcal{F}$-measurable function $f:X \to \{x : 0 \leq x \leq +\infty\}$ is the limit of a monotone increasing sequence of non-negative $\mathcal{F}$-simple functions.

Definition of the Integral

Since the most fundamental type of measurable function is a non-negative simple function, it is natural to first define the integral of such a function and then to extend this definition to measurable functions in general.

Integration of Non-negative Simple Functions: Suppose $f:X \to \{x : 0 \leq x \leq +\infty\}$ is an $\mathcal{F}$-simple function expressible as

$$f(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x)$$

where $\{E_i\}_{i=1}^{n}$ forms an $\mathcal{F}$-dissection of X and $\{c_i\}_{i=1}^{n}$ is a sequence from $\{x : 0 \leq x < \infty\}$. Define

$$\int f d\mu = \sum_{i=1}^{n} c_i \mu(E_i)$$

and call this sum the integral of f with respect to μ (Kingman and Taylor, 1966, p. 110). Under the convention that if $c_i = 0$ and $\mu(E_i) = \infty$ for some i then $c_i\mu(E_i) = 0$, the sum is always defined. Additivity of μ may be used to verify that $\int f d\mu$ is independent of which representation of the above form for f is used.

Integration of Non-negative Measurable Functions: Since any non-negative $\mathcal{F}$-measurable function is the limit of a monotone non-decreasing sequence of non-negative $\mathcal{F}$-simple functions by Theorem 2.8, one is led to make the following definition.

Suppose $f:X \to \{x : 0 \leq x \leq +\infty\}$ is $\mathcal{F}$-measurable. Then there exists a monotone non-decreasing sequence $\{f_n\}_{n=1}^{\infty}$ of non-negative simple functions converging to f. Define

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu \text{ (Kingman and Taylor, 1966, p. 111).}$$

This limit is always defined (may be $+\infty$) since monotonicity of $\{f_n\}_{n=1}^{\infty}$ implies monotonicity of $\{\int f_n d\mu\}_{n=1}^{\infty}$. Thus $\int f d\mu$ is well-defined if and only if $\lim_{n\to\infty}\int f_n d\mu$ does not depend on which of the monotone non-decreasing sequences converging to f is used. The main step in showing that the definition of $\int f d\mu$ is proper is to show that if g is any simple function such that $g \leq f$, then $\lim_{n\to\infty}\int f_n d\mu \geq \int g d\mu$

for an arbitrary monotone non-decreasing sequence $\{f_n\}_{n=1}^{\infty}$ converging to f.

DEFINITION 2.19 (Kingman and Taylor, 1966, p. 113): A non-negative measurable function f is integrable with respect to μ if $\int f d\mu$ is finite.

Integration of Measurable Functions: Given a measurable function $f:X \to R^*$, define a function $f_+:X \to \{x : 0 \le x \le +\infty\}$ by

$$f_+(x) = \begin{cases} f(x) & ; x \in \{x : f(x) \ge 0\}, \\ 0 & ; \text{elsewhere} \end{cases}$$

and define a function $f_-:X \to \{x : 0 \le x \le +\infty\}$ by

$$f_-(x) = \begin{cases} -f(x) & ; x \in \{x : f(x) \le 0\}, \\ 0 & ; \text{elsewhere.} \end{cases}$$

By Theorem 2.7 f_+ and f_- are both measurable. Since f_+ and f_- are both non-negative and since $f = f_+ - f_-$, define

$$\int f d\mu = \int f_+ d\mu - \int f_- d\mu$$

whenever f_+ and f_- are integrable.

DEFINITION 2.20 (Kingman and Taylor, 1966, p. 114): A measurable function $f:X \to R^*$ is integrable with respect to μ if f_+ and f_- are integrable.

Restriction of the Integral to Measurable Subsets: Suppose $A \in \mathcal{F}$. Define for a function $f:X \to R^*$

$$\int_A f d\mu = \int f \chi_A d\mu$$

whenever $\int f \chi_A d\mu$ is defined. Then $\int_A f d\mu$ will be defined if

(i) $f\chi_A$ is a non-negative measurable function, or if

(ii) $f\chi_A$ is a measurable integrable function (Kingman and Taylor, 1966, p. 114). Under condition (i) $\int_A f d\mu$ will be non-negative (possibly $+\infty$) and under condition (ii) $\int_A f d\mu$ will be finite.

Properties of the Integral

A few properties of the general integral which will be applied to Lebesgue-Stieltjes integrals in subsequent chapters are now given.

THEOREM 2.9 (Kingman and Taylor, 1966, p. 115): Suppose $(X, \mathcal{F}, \mu)$ is a measure space, A and B are disjoint sets in $\mathcal{F}$, and $f:X \to R^*$, $g:X \to R^*$ are two functions integrable (over X) with respect to μ. Then f is integrable over A, f + g and $|f|$ are integrable (over X), and

(i) $\int_{A\cup B} f d\mu = \int_A f d\mu + \int_B f d\mu$,

(ii) f is finite except on a set of measure zero,

(iii) $\int (f + g) d\mu = \int f d\mu + \int g d\mu$,

(iv) $|\int f d\mu| \leq \int |f| d\mu$,

(v) for any $c \in R$, cf is integrable and $\int cf d\mu = c\int f d\mu$,

(vi) $f \geq 0$ implies $\int f d\mu \geq 0$; $f \geq g$ implies $\int f d\mu \geq \int g d\mu$,

(vii) if $f \geq 0$ and $\int f d\mu = 0$, then $f = 0$ except possibly on a set of measure zero,

(viii) $f = g$ implies $\int f d\mu = \int g d\mu$,

(ix) if $h:X \to R^*$ is $\mathcal{F}$-measurable and $|h| \leq f$, then h is integrable.

THEOREM 2.10 (Monotone Convergence) (Kingman and Taylor, 1966, p. 119): Suppose $\{f_n\}_{n=1}^{\infty}$ is a monotone non-decreasing sequence of non-negative measurable functions $f_n:X \to R^*$ ($n \in N$) and $f_n(x) \to f(x)$ for all $x \in X$. Then

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$$

in the sense that if f is integrable $\int f_n d\mu \to \int f d\mu$; while if f is not integrable, either $\int f_n d\mu < \infty$ for all n but $\int f_n d\mu \to +\infty$, or there is an integer N such that $\int f_n d\mu = +\infty$ for each $n \geq N$.

THEOREM 2.11 (Bounded Convergence) (Kingman and Taylor, 1966, p. 121): If $g:X \to R^+$ is integrable, $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions $f_n:X \to R^*$ ($n \in N$) such that $|f_n| \leq g$ ($n \in N$) and $f_n \to f$ as $n \to \infty$, then f is integrable and $\int f_n d\mu \to \int f d\mu$ as $n \to \infty$.

DEVELOPMENT OF THE LEBESGUE INTEGRAL

The purpose of this chapter is to define the concept of Lebesgue measure and to establish the Lebesgue integral as a special case of the general integral.

Lebesgue Measure

One type of measure, called Lebesgue measure, is of particular interest, being literally an extension of the concepts of length, area, and volume of figures.

Consider a "length" function ℓ defined on the semi-ring $\mathcal{P} = \{(a,b] : a \in R, b \in R\}$ by

$$\ell((a,b]) = b - a \quad ; a \in R, b \in R.$$

By Theorem 2.4 there is a unique extension ϕ of ℓ defined on the ring $\mathcal{E} = \{A : A \text{ is a finite union of disjoint members of } \mathcal{P}\}$ generated by $\mathcal{P}$. It may be shown that ϕ is σ-additive on $\mathcal{E}$; thus ϕ is a measure. Furthermore, since $\phi(A) < \infty$ for each set $A \in \mathcal{P}$, ϕ is σ-finite on $\mathcal{E}$. Therefore, the Extension Theorem implies that ϕ may be extended uniquely to a measure μ defined on (and σ-finite on) the σ-ring $\mathcal{B}$ generated by $\mathcal{E}$. $\mathcal{B}$ is the σ-field of Borel sets in R. It is worth noticing that, by Theorem 2.6, there is a unique complete measure defined on the completion $\mathcal{L}$ of $\mathcal{B}$ with respect to μ. $\mathcal{L}$ is called the class of Lebesgue-measurable sets.

Construction of k-dimensional ($k \in N$) Lebesgue measure follows the same lines as that for one-dimensional Lebesgue measure. Denote by $\prod_{i=1}^{k} P_i$ sets of the form $\{(x_1, x_2, \ldots, x_k) : x_i \in P_i\}$ for sequences$\{P_i\}_{i=1}^{k}$ from $\mathcal{P}$ (Gemignani, 1972, p. 84). The "length" function ℓ_k for the k-dimensional case may then be defined on the semiring $\mathcal{P}^k = \{\prod_{i=1}^{k} P_i : P_i \in \mathcal{P}\ (i=1, 2, \ldots, k)\}$.by

$$\ell_k\left(\prod_{i-1}^{k} P_i\right) = \ell(P_1)\ell(P_2)\cdots\ell(P_k) \ ; \ \prod_{i=1}^{k} P_i \in \mathcal{P}^k$$

(For k=2 and k=3, ℓ_k may be thought of as an area function and a volume function, respectively).

Extension of ℓ_k leads to a measure μ_k defined on the σ-field $\mathcal{B}^k$ of Borel sets in R^k, and the completion of $\mathcal{B}^k$ with respect to μ_k is called the class of k-dimensional Lebesgue measurable sets((Kingman and Taylor, 1966, p. 79).

Lebesgue Integration

The concept of Lebesgue integration is the specialization of the general integral to the measure space $(R^k, \mathcal{L}^k, \mu_k^k)$, where μ_k represents Lebesgue measure on R^k and $\mathcal{L}^k$ is the σ-field of μ_k-measurable sets.

In the one-dimensional measure space $(R, \mathcal{L}, \mu)$, where μ is Lebesgue measure on R and $\mathcal{L}$ is the σ-field of Lebesgue measurable sets, it is usual to use the notation

$$\int_A f(x)dx \quad \text{for} \quad \int_A f d\mu$$

to denote the Lebesgue integral of an $\mathcal{L}$-measurable function $f:R \to R^*$ over a set $A \in \mathcal{L}$ (Kingman and Taylor, 1966, p. 124).

If A is an interval with endpoints a and b, one may write

$$\int_a^b f(x)dx \quad \text{for} \quad \int_A f(x)dx.$$

In the k-dimensional measure space $(R^k, \mathcal{L}^k, \mu_k)$, if $f:R^k \to R^*$ is an $\mathcal{L}^k$-measurable function and $A \in \mathcal{L}^k$, one writes

$$\int\int_A \cdots \int f(x_1, x_2, \ldots, x_k) d(x_1, x_2, \ldots, x_k)$$

for $\int_A f d\mu_k$.

DEVELOPMENT OF THE LEBESGUE-STIELTJES INTEGRAL

A type of measure which finds application in probability theory and in other areas is a Lebesgue-Stieltjes measure. The purpose of this chapter is to construct a general Lebesgue-Stieltjes measure and to discuss the type of integral which might be defined with respect to such a measure.

Lebesgue-Stieltjes Measure

DEFINITION 4.1 (Kingman and Taylor, 1966, p. 95): A k-dimensional Stieltjes measure function is a real-valued function F defined on R^k which is monotone non-decreasing and continuous on the right in each variable separately ($k \in N$).

The derivation of Lebesgue-Stieltjes measure for $k = 1$ will be presented in detail.

One-Dimensional Lebesgue-Stieltjes Measure: Suppose a function $F:R \to R$ is monotone non-decreasing and everywhere continuous on the right. Then F is a (one-dimensional) Stieltjes measure function by Definition 4.1. Define a set function ℓ_F on the semi-ring $\mathcal{P} = \{(a,b]: a \leq b,\ a \in R,\ b \in R\}$ by

$$\ell_F((a,b]) = F(b) - F(a) \quad ; (a,b] \in \mathcal{P}.$$

The set function ℓ_F is finitely additive on $\mathcal{P}$ because for each set $(a,b] \in \mathcal{P}$ and for each monotone increasing sequence $\{x_i\}_{i=1}^n$ from (a,b),

$$
\begin{aligned}
\ell_F((a,b]) &= F(b) - F(a) = \\
&= F(b) - F(x_n) + F(x_n) - F(x_{n-1}) + \ldots \\
&\qquad + F(x_1) - F(a) = \\
&= \sum_{i=1}^{n+1} \ell_F((x_i, x_{i+1}])
\end{aligned}
$$

where $x_{n+1} = b$. Monotonicity of F implies ℓ_F is non-negative. Thus Theorem 2.4 may be used to extend ℓ_F to a unique non-negative and finitely additive set function ϕ_F defined on the ring $\mathcal{E} = \{A\text{: A is a finite union of disjoint members of } \mathcal{P}\}$ generated by $\mathcal{P}$.

To see that ϕ_F is σ-additive and therefore is a measure on $\mathcal{E}$, suppose ϕ_F is not σ-additive. Since $\phi_F(A)<\infty$ for each $A \in \mathcal{E}$, Theorem 2.3,iii implies the existence of a monotone decreasing sequence $\{A_n\}_{n=1}^\infty$ from $\mathcal{E}$ converging to ϕ such that $\lim_{n\to\infty} \phi_F(A_n) \neq 0$. Since each A_n is a finite union of elements of $\mathcal{P}$, monotonicity of $\{A_n\}_{n=1}^\infty$ implies there is a monotone decreasing sequence $\{(a_n,b_n]\}_{n=1}^\infty$ from $\mathcal{P}$ converging to ϕ such that $\lim_{n\to\infty} \ell_F((a_n,b_n]) > 0$. Thus $\{b_n\}_{n=1}^\infty$ is a sequence converging to $\lim_{n\to\infty} a_n$ and there exists $\delta > 0$ for which $F(b_n) - F(\lim_{n\to\infty} a_n) > \delta$ for each n, contradicting the right-continuity of F. Therefore, ϕ_F is

σ-additive on $\mathcal{E}$.

The measure ϕ_F is σ-finite on $\mathcal{E}$ because each set in $\mathcal{E}$ is a finite union of members of $\mathcal{P}$ and $\phi_F(P) < \infty$ for each $P \in \mathcal{P}$. Thus, since R may be covered by a countable union of members of $\mathcal{E}$, the Extension Theorem may be used to extend ϕ_F to a unique measure μ_F defined on the σ-ring $\mathcal{B}$ generated by $\mathcal{E}$. The σ-ring $\mathcal{B}$ is the σ-field of Borel sets in R.

As in the discussion of Lebesgue measure, one may use Theorem 2.6 to extend μ_F to a complete measure c_F defined on the completion $\mathcal{L}_F$ of $\mathcal{B}$ with respect to μ_F. Note that $\mathcal{L}_F$ depends on F, as is illustrated by the following examples.

EXAMPLE 4.1: If $F(x) = x \quad ; x \in R$, then μ_F is Lebesgue measure in R and $\mathcal{L}_F = \mathcal{L}$ (Kingman and Taylor, 1966, p. 96).

EXAMPLE 4.2: If $F(x) = c \quad ; c$ a constant, $x \in R$, then $\mu_F(R) = 0$ and $\mathcal{L}_F$ is the σ-field of all subsets of R (Kingman and Taylor, 1966, p. 96).

<u>Multi-Dimensional Lebesgue-Stieltjes Measure</u>: Suppose F is a k-dimensional Stieltjes measure function ($k \in N$). Define a set function ℓ_F^k on the semi-ring $\mathcal{P}^k$ by

$$\ell_F^k(I) = \sum_{i=1}^{2k} y_i F(V_i) \quad ; I \in \mathcal{P}^k \tag{1}$$

where the V_i are the 2k vertices of the set I and $y_i = 1$ for the vertex in which each coordinate is largest and

$y_i = (-1)^r$ otherwise, r being the number of coordinates of the vertex V_i which are at their upper bounds (Kingman and Taylor, 1966, p. 96).

It may be shown that ℓ_F^k is non-negative and finitely additive; thus ℓ_F^k has a unique extension ϕ_F^k defined on the ring $\mathcal{E}^k$ generated by $\mathcal{P}^k$. As in the one-dimensional case, ϕ_F^k is a σ-finite measure on $\mathcal{E}^k$. Since R^k may be covered by a countable union of members of $\mathcal{E}^k$, the Extension Theorem guarantees a unique extension of ϕ_F^k to a measure μ_F^k defined on $\mathcal{B}^k$, the σ-field of Borel sets in R^k. There is, by Theorem 2.6, a complete measure c_F^k which is an extension of μ_F^k to the completion $\mathcal{L}_F^k$ of $\mathcal{B}^k$ with respect to μ_F^k. $\mathcal{L}_F^k$ is called the class of k-dimensional Lebesgue-Stieltjes measurable sets, and depends on F.

EXAMPLE 4.3: If $k = 2$ and $F(x,y) = xy$; $x, y \in R$, then

$$\ell_F(\{(x,y): a < x \leq b,\ c < y \leq d\}) = bd - ad - bc + ac =$$
$$= d(b - a) - c(b - a) = (d - c)(b - a)$$

corresponding to the area of the rectangular region. Accordingly, the μ_F derived from ℓ_F is 2-dimensional Lebesgue measure.

EXAMPLE 4.4: Let $\langle x\rangle$ denote the greatest integer in x throughout this paper. If $k = 2$ and

$$F(x,y) = \begin{cases} \left[\sum_{n=0}^{\langle x\rangle} \frac{e^{-\lambda}\lambda^n}{n!}\right]\left[\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{y} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt\right] & ;\ x \geq 1,\ y \in R \\ 0 & ;\ x < 1,\ y \in R \end{cases}$$

with fixed numbers $\lambda > 0$, $\sigma > 0$, $\mu \in R$, then $\mathcal{L}_F^2 = \mathcal{L}^2$. Since F has the form of a joint distribution function for two independent random variables X and Y with the Poisson distribution with parameter λ and the Normal distribution with mean μ and variance σ^2, respectively, μ_F^2 may be interpreted as representing a probability measure defined on the class $\mathcal{L}_F^2$ which is, in this case, the class of 2-dimensional Lebesgue measurable sets. Note that even though the domain of the first variable (x) is the whole real number line, the greatest integer notation in the expression for F(x,y) insures that the random variable X takes on non-integer values with probability zero and may thus be regarded as a discrete random variable.

Lebesgue-Stieltjes Integration

Lebesgue-Stieltjes integration is the specialization of the general integral to a space $(R^k, \mathcal{L}_F^k, \mu_F^k)$ $(k \in N)$, where μ_F^k is the k-dimensional Lebesgue-Stieltjes measure on R^k determined by a specified Stieltjes measure function F. If $g: R^k \to R^*$ is an $\mathcal{L}_F^k$-measurable function and $A \in \mathcal{L}_F^k$, one writes

$$\iint_A \cdots \int g(x_1, x_2, \ldots, x_k)\, dF(x_1, x_2, \ldots, x_k) \text{ for } \int_A g\, d\mu_F .$$

Note that one does not write the Lebesgue-Stieltjes integral of a measurable function $g: R \to R^*$ over an interval as

$$\int_a^b g(x)\, dF(x),$$

since

$$\int_{[a,b]} g(x)\, dF(x) = \int_{[a,b)} g(x)\, dF(x) + \int_{\{b\}} g(x)\, dF(x) =$$

$$= \int_{[a,b)} g(x)\, dF(x) + g(b)[F(b) - \lim_{x \to b^-} F(x)]$$

and the expression in brackets might not be zero.

DEVELOPMENT OF THE RIEMANN-STIELTJES INTEGRAL

The Riemann-Stieltjes integral is a special case of the Lebesgue-Stieltjes integral. However, its definition closely resembles that of the Riemann integral. Therefore, this chapter is divided into two sections: The Riemann Integral and Riemann-Stieltjes Integration. The first section presents the definition of the Riemann integral. The second section then uses the same pattern of argument in defining the Riemann-Stieltjes integral and compares it to the Lebesgue-Stieltjes integral.

The Riemann Integral

The derivation to follow is adapted from that of Royden (1968, p. 73). Consider a bounded real-valued function f defined on a closed interval [a,b], and denote an arbitrary subdivision of [a,b] by

$$a = \xi_0 < \xi_1 < \dots < \xi_n = b \quad ; n \epsilon N.$$

For each such subdivision define

$$S = \sum_{i=1}^{n} \left\{ \left[\sup_{\xi_{i-1}<x<\xi_i} f(x) \right] \left[\xi_i - \xi_{i-1} \right] \right\}$$

and define

$$s = \sum_{i=1}^{n} \left\{ \left[\inf_{\xi_{i-1}<x<\xi_i} f(x) \right] \left[\xi_i - \xi_{i-1} \right] \right\}.$$

Then define the upper Riemann integral of f over [a,b] by

$$\overline{\int_a^b} f(x)dx = \inf S$$

and define the lower Riemann integral of f over [a,b] by

$$\underline{\int_a^b} f(x)dx = \sup s$$

where the infinum and supremum are taken over all possible subdivisions of [a,b].

If the upper and lower Riemann integrals are equal, one says that f is Riemann integrable over [a,b] and calls the common value the Riemann integral of f over [a,b]. It will be denoted henceforth by

$$\mathcal{R}\int_a^b f(x)dx.$$

Note that if the Riemann integral of f from a to b is defined, it is finite, since f is bounded on [a,b].

The following theorem gives precise conditions under which a function is Riemann integrable.

THEOREM 5.1 (Kingman and Taylor, 1966, p. 129): Let μ represent Lebesgue measure on R. A bounded function f:[a,b] → R is Riemann integrable if and only if the set A of points in [a,b] at which f is discontinuous satisfies $\mu(A) = 0$. If f is Riemann integrable, then it is Lebesgue integrable and

$$\int_a^b f(x)dx = \mathcal{R}\int_a^b f(x)dx.$$

Riemann-Stieltjes Integration

A Riemann-Stieltjes integral is expressible as a Riemann integral plus a summation and is evaluated as such. In this section the Riemann-Stieltjes integral will be defined and a theorem concerning the evaluation of Riemann-Stieltjes integrals will be presented. Finally, the Riemann-Stieltjes and Lebesgue-Stieltjes integrals will be compared.

Definition of the Riemann-Stieltjes Integral: The definition which follows is adapted from that presented by Gaughan (1969, p. 138).

Let $g:[a,b] \to R$ be a bounded function and let $F:[a,b] \to R$ be a monotone non-decreasing function. For any subdivision S defined by

$$a = \xi_0 < \xi_1 < \dots < \xi_n = b \quad ; n \in N$$

define

$$U(S,g) = \sum_{i=1}^{n} \sup_{\xi_{i-1}<x<\xi_i} g(x)[F(\xi_i) - F(\xi_{i-1})] \text{ and}$$

$$L(S,g) = \sum_{i=1}^{n} \inf_{\xi_{i-1}<x<\xi_i} g(x)[F(\xi_i) - F(\xi_{i-1})].$$

If $\inf U(S,g) = \sup L(S,g)$, where the infinum and supremum are taken over all possible subdivisions of $[a,b]$, denote their common value by

$$\int_{[a,b]} g(x)dF(x)$$

and call it the Riemann-Stieltjes integral of g with respect to F over the interval [a,b].

If $\int_{[a,b]} g(x)dF(x)$ is defined and is finite, one says that g is Riemann-Stieltjes integrable with respect to F over [a,b].

Note that if $F(x) = x$ for all $x \varepsilon R$, the Riemann-Stieltjes integral above reduces to the Riemann integral

$$\mathcal{R}\int_a^b g(x)dx.$$

Theorem on the Evaluation of Riemann-Stieltjes Integrals: The following theorem and variations thereof are sometimes referred to as Duhamel's Theorem (Widder, 1961, p. 73).

THEOREM 5.2: Let F be a real-valued, bounded, non-decreasing function with domain a closed interval [a,b]. Suppose F is non-differentiable on at most a finite set $D = \{d_j\}_{j=1}^m$. Define a function $f:[a,b] \to [0,+\infty)$ by

$$f(x) = \begin{cases} 0 & ; x \varepsilon D, \\ \frac{dF(x)}{dx} & ; x \varepsilon [a,b] - D. \end{cases}$$

Suppose g is a real-valued function which is defined and bounded on [a,b] and is continuous on (a,b). Then g is Riemann-Stieltjes integrable with respect to F on [a,b] and

$$\int_{[a,b]} g(x)dF(x) =$$

$$= \mathcal{R}\int_a^b g(x)f(x)dx + \sum_{j=1}^{m} g(d_j)[\lim_{x\to d_j^+} F(x) - \lim_{x\to d_j^-} F(x)].$$

Proof: Let S be a subdivision of [a,b] defined by

$$a = \xi_0 < \xi_1 < \dots < \xi_n = b. \qquad (1)$$

One says that a subdivision S^* is a refinement of S if each element of S is an element of S^*, and one writes $S \subset S^*$ (Gaughan, 1968, p. 137).

For each subdivision defined by (1) and for each bounded function $h:[a,b] \to R$, define the following notation.

$$M(h,\xi_{i-1},\xi_i) = \sup_{\xi_{i-1}<x<\xi_i} h(x),$$

$$m(h,\xi_{i-1},\xi_i) = \inf_{\xi_{i-1}<x<\xi_i} h(x).$$

Then for each subdivision S of [a,b]

$$U(S,h) = \sum_{i=1}^{n} M(h,\xi_{i-1},\xi_i)[F(\xi_i) - F(\xi_{i-1})] \text{ and}$$

$$L(S,h) = \sum_{i=1}^{n} m(h,\xi_{i-1},\xi_i)[F(\xi_i) - F(\xi_{i-1})].$$

Let us now prove that

$$S \subset S^* \text{ implies } \begin{cases} U(S^*,g) \leq U(S,g) \text{ and} \\ L(S^*,g) \geq L(S,g). \end{cases} \qquad (2)$$

First, suppose S^* contains exactly one more element ξ^* than does S, and that $\xi_{i-1} < \xi^* < \xi_i$ where $\xi_{i-1}, \xi_i \in S$. Clearly

$$M(g,\xi_{i-1},\xi^*) \leq M(g,\xi_{i-1},\xi_i) \text{ and}$$
$$M(g,\xi^*,\xi_i) \leq M(g, \xi_{i-1},\xi_i).$$

Therefore,

$$M(g,\xi_{i-1},\xi_i)[F(\xi_i)-F(\xi_{i-1})]-\{M(g,\xi_{i-1},\xi^*)[F(\xi^*)-F(\xi_{i-1})] + $$
$$+ M(g,\xi^*,\xi_i)[F(\xi_i)-F(\xi^*)]\} =$$
$$= [M(g,\xi_{i-1},\xi_i)-M(g,\xi_{i-1},\xi^*)][F(\xi^*)-F(\xi_{i-1})] +$$
$$+ [M(g,\xi_{i-1},\xi_i)-M(g,\xi^*,\xi_i)][F(\xi_i)-F(\xi^*)] \geq 0,$$

the equality being obtained by adding and subtracting $M(g,\xi_{i-1},\xi_i)F(\xi^*)$ in the first expression. If S^* contains k more elements than S, application of this procedure k times and rearrangement of terms gives

$$U(S^*,g) \leq U(S,g).$$

The case for $L(S^*,g) \geq L(S,g)$ is similar. Since S was an arbitrary subdivision of [a,b] and S^* was extended to an arbitrary refinement of S, (2) is proved.

Let us now show that inf U(S,g) and sup L(S,g) are defined and are equal. For any subdivision S, let Σ_1 denote a summation over all intervals (ξ_{i-1},ξ_i) which contain no elements of D and let Σ_2 denote a summation over all intervals (ξ_{i-1},ξ_i) which contain elements of D. For each subdivision define $||S|| = \max_i (\xi_i - \xi_{i-1})$. If

S is any fixed subdivision of $[a,b]$ and $\{S_k\}_{k=1}^{\infty}$ is any sequence of subdivisions such that $S \subset S_1 \subset S_2 \subset \ldots$ for which $||S_k|| \to 0$ as $k \to \infty$, then (2) implies that $\{U(S_k,g)\}_{k=1}^{\infty}$ is a monotone non-increasing sequence of real numbers.

Denote for each $k \in N$

$$U(S_k,g) = \sum\nolimits_1 M_k(g,\xi_{i-1},\xi_i)[F(\xi_i) - F(\xi_{i-1})] + \\ + \sum\nolimits_2 M_k(g,\xi_{i-1},\xi_i)[F(\xi_i) - F(\xi_{i-1})], \tag{3}$$

where the ξ_i are understood to belong to their respective subdivisions S_k $(k \in N)$.

Since F is continuous on $[a,b] - D$, g is continuous on (a,b), and both F and g are bounded on $[a,b]$,

$$\sum\nolimits_2 M_k(g,\xi_{i-1},\xi_i)[F(\xi_i) - F(\xi_{i-1})] \to \\ \to \sum_{j=1}^{m} g(d_j)[\lim_{x \to d_j^+} F(x) - \lim_{x \to d_j^-} F(x)] \tag{4}$$

as $k \to \infty$.

Now for each $k \in N$ and for each interval (ξ_{i-1},ξ_i) included in Σ_2,

$$F(\xi_i) - F(\xi_{i-1}) = \mathcal{R}\int_{\xi_{i-1}}^{\xi_i} f(t)dt$$

by the Fundamental Theorem of Calculus. Therefore, the first term on the right side of (3) is

$$\sum_1 M_k(g,\xi_{i-1},\xi_i) \mathcal{R}\int_{\xi_{i-1}}^{\xi_i} f(t)dt.$$

But by definition of the Riemann integral, $\mathcal{R}\int_{\xi_{i-1}}^{\xi_i} f(t)dt$ may be expressed as

$$\mathcal{R}\int_a^b f(t)\chi_{(\xi_{i-1},\xi_i)}(t)dt =$$

$$= \lim_{k\to\infty} \sum_{i=1}^{k} M(f\chi_{(\xi_{i-1},\xi_i)},\omega_{i-1},\omega_i)(\omega_i - \omega_{i-1})$$

where the ω_i's represent the ξ_i's defining S_k, $k \varepsilon N$. Thus as $k \to \infty$, the first term on the right side of (3) reduces to

$$\lim_{k\to\infty}\sum_1 M_k(g,\xi_{i-1},\xi_i)\lim_{k\to\infty}\sum_{i=1}^{k} M(f\chi_{(\xi_{i-1},\xi_i)},\omega_{i-1},\omega_i)(\omega_i - \omega_{i-1}) =$$

$$= \lim_{k\to\infty}\sum_1 M_k(g,\xi_{i-1},\xi_i)M_k(f,\xi_{i-1},\xi_i)(\xi_i - \xi_{i-1}). \tag{5}$$

Continuity of g and f on [a,b] - D implies that, for any $c \varepsilon [a,b] - D$ and for any δ for which $0 < \delta < \min_j (|c - d_j|)$

$$M_k(g,c-\delta,c+\delta)M_k(f,c-\delta,c+\delta) \to g(f(c)) \text{ as } k \to \infty. \tag{6}$$

Therefore, since $f\circ g$ is continuous on [a,b] - D, Theorem 5.1 and (6) imply that (5) may be written as

$$\lim_{k\to\infty}\sum_1 M_k(f\ g,\xi_{i-1},\xi_i)(\xi_i - \xi_{i-1}) = \mathcal{R}\int_a^b g(t)f(t)dt. \tag{7}$$

Since (4) and (7) did not depend on which sequence of subdivisions was used,

$$\inf U(S,g) = \mathcal{R}\int_a^b g(t)f(t)dt + \sum_{j=1}^{m} g(d_j)[\lim_{x\to d_j^-} F(x) - \lim_{x\to d_j^+} F(x)]$$

as was expected. An identical result holds for inf $L(S,g)$ and the theorem is proved.

<u>Comparison with the Lebesgue-Stieltjes Integral</u>: To observe the relationship between Riemann-Stieltjes and Lebesgue-Stieltjes integration, define a function $G:R \to \{x:0 \leq x < \infty\}$ by

$$G(x) = \begin{cases} 0 & ; x < a, \\ F(x) & ; a \leq x \leq b, \\ F(b) & ; b < x. \end{cases}$$

Since G is continuous on $(-\infty,a)$ and on (b,∞) and is identical to F on $[a,b]$, G is monotone non-decreasing and everywhere continuous on the right. Thus G is a Stieltjes measure function and defines a Lebesgue-Stieltjes measure μ_G on the σ-ring $\mathcal{E}$ generated by the semi-ring $\mathcal{P}$ of half-open intervals of the form $(c,d]$. Hence the Riemann-Stieltjes integral

$$\int_{[a,b]} g(x)dF(x)$$

is exactly

$$\int_{[a,b]} g d\mu_G = \int_{[a,b]} g(x)dG(x).$$

A HIERARCHY OF TYPES OF INTEGRALS

One is now in a position to recognize a hierarchy of types of integrals. Such an ordering scheme is depicted in Figure 1, in which each integral listed is a special case of all those above it in the chart.

General integral

Lebesgue-Stieltjes integral

Lebesgue integral　　Riemann-Stieltjes integral

Riemann integral

Figure 1. A Hierarchy of Types of Integrals. Each integral listed is a special case of all those above it in the chart.

APPLICATIONS OF STIELTJES INTEGRATION

A basic application of a Stieltjes-type integral is its use in expressing the cumulative value of some physically measurable quantity such as length, area, volume, mass, or force. This chapter gives three examples of such applications; mass of a wire, moment of inertia, and gravitational attraction.

Mass of a Wire

Consider a wire of length L; denote distance from one end by x. Let point masses of mass $m_1, m_2, \ldots, m_n$ be at respective points $x_1 < x_2 < \cdots < x_n$ along the wire. Suppose the density of the wire at all other points is a function $\rho : [0,L] - \{x_i\}_{i=1}^n \to R^+$ which is uniformly continuous on each of the open intervals $(0,x_1)$, (x_1,x_2), ..., (x_n,L) (Fig. 1a). In case $x_1 = 0$ or $x_n = L$, then $(0,x_1)=\phi$ or $(x_n,L) = \phi$, respectively, and ρ may be said to be uniformly continuous on ϕ by convention.

Define a function $\rho^*: R \to R^+$ by

$$\rho^*(x) = \begin{cases} \rho(x) & ; x \varepsilon [0,L] - \{x_i\}_{i=1}^n, \\ 0 & ; \text{elsewhere}. \end{cases}$$

Let $x_0 = 0$. Then for fixed $x \varepsilon [0,\infty)$ define

$$k = \max \{i : x_i \leq x\}, \text{ and}$$

$$y(x) = \min (x,L).$$

Define a function $F:R \to R^{+}$ by

$$F(x) = \begin{cases} 0 & ; \; x < 0, \\ \sum_{i=1}^{k} \mathcal{R}\int_{x_{i-1}}^{x_i} \rho^{*}(t)dt + \mathcal{R}\int_{x_k}^{y(x)} \rho^{*}(t)dt + \sum_{i=1}^{k} m_i & ; \; x \geq 0, \end{cases}$$

where $\mathcal{R}\int(\cdot)dt$ denotes the Riemann integral. Note that if $x = x_i$ for some i $(i = 1, 2, \ldots, n)$, then $x_k = y(x)$ and $\mathcal{R}\int_{x_k}^{y(x)} \rho^{*}(t)dt = 0$. Therefore, in any case, $F(x)$ represents the total mass distributed along the interval $(-\infty, x]$ (Fig. 2b).

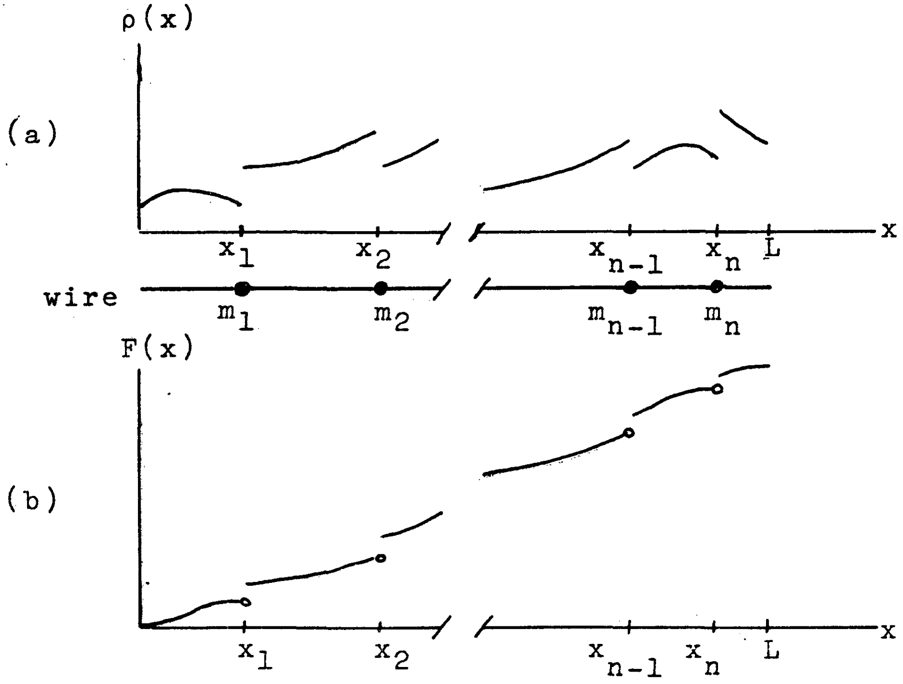


Figure 2. (a) Wire Density ρ as a Uniformly Continuous Function of Length x on Each of the Intervals $(0,x_1)$, (x_1,x_2),...,(x_n,L). (b) The Cumulative Mass F as a Function of Length x. The jumps at points x_1, x_2,...,x_n are of heights $m_1, m_2, \ldots, m_n$, respectively.

The function F is defined on R (instead of merely on $[0,L]$) and F is monotone non-decreasing and everywhere continuous on the right; thus F is a Stieltjes measure function.

Define a set function ℓ_F on the semi-ring $\mathcal{P} = \{(a,b]: a \leq b \text{ and } a \in R,\ b \in R\}$ by

$$\ell_F((a,b]) = F(b) - F(a) \quad ; a \leq b \text{ and } a \in R,\ b \in R.$$

An argument identical to that of page 21 may then be used to extend ℓ_F to a unique Lebesgue-Stieltjes measure μ_F defined on the σ-field of Borel sets in R. For any Borel set A, $\mu_F(A)$ represents the mass contained in that part of the wire whose x-coordinates are elements of A. Since μ_F is the Legesgue-Stieltjes measure determined by F, one writes $\mu_F(A) = \int_A d\mu_F$ as the Lebesgue-Stieltjes integral

$$\int_A dF(x).$$

The above discussion used a measure-theoretic approach to illustrate an application of Lebesgue-Stieltjes measure theory. Clearly, once F was defined, Theorem 5.2 (with $g(x) \equiv 1$) could have been used directly to write the mass distributed over any interval I with endpoints a and b as

$$\mathcal{R}\int_a^b \rho^*(t)dt + \sum_{m_i \in I} m_i$$

which has the Stieltjes form

$$\int_I dF(x).$$

Moment of Inertia

This section will make use of Theorem 5.2.

Consider a plane lamina on a region $a \le x \le b$, $0 \le y \le f(x)$ where f is a continuous, bounded function on $[a,b]$ and $a \ge 0$, $b \ge 0$. Suppose a finite number of point masses of mass $m_1, m_2, \ldots, m_k$ are distributed throughout the interior of the lamina at coordinates (x_1,y_1), (x_2,y_2), $\ldots$, (x_k,y_k), respectively. Furthermore, suppose that at all other points (x,y) in the region, density (mass/ unit area) is a continuous function $\rho(x)$ of x alone (Fig. 3). Let us find the moment of inertia of the lamina about the y-axis.

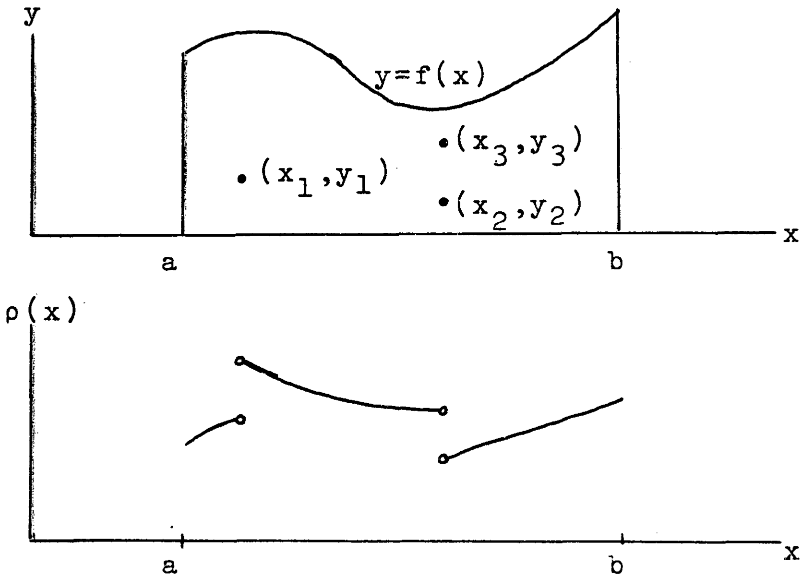


Figure 3. A Typical Distribution of Point Masses at Coordinates (x_i,y_i), and the Density $\rho(x)$ as a Continuous Function of x except at $x_1, x_2, \ldots, x_k$.

Denote an arbitrary subdivision of $[a,b]$ by

$$a = \xi_0 < \xi_1 < \dots < \xi_n = b \quad ; n \in N.$$

For each such subdivision, erect ordinates at each ξ_i of heights $\sup\limits_{\xi_{i-1}<x<\xi_i} f(x)$ and $\inf\limits_{\xi_{i-1}<x<\xi_i} f(x)$. For each ξ_i, consider the rectangular regions

$R_{1i} = \{(x,y) : \xi_{i-1} \le x \le \xi_i,\ 0 \le y \le \sup\limits_{\xi_{i-1}<x<\xi_i} f(x)\}$ and

$R_{2i} = \{(x,y) : \xi_{i-1} \le x \le \xi_i,\ 0 \le y \le \inf\limits_{\xi_{i-1}<x<\xi_i} f(x)\}$ (Fig. 4).

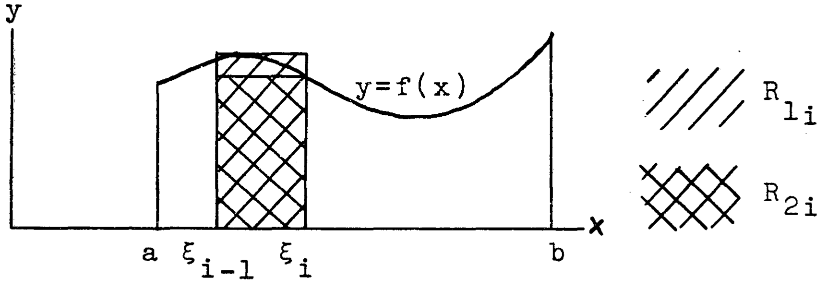


Figure 4. Construction of R_{1i} and R_{2i} on a Typical Interval $\xi_{i-1} \le x \le \xi_i$.

Define a function $M:[a,b] \to R$ by

$$M(x) = \mathcal{R}\int_a^x \rho(t)\,dt + \sum_{x_j \le x} m_j.$$

Then the mass of R_{1i} is

$$\sup_{\xi_{i-1}<x<\xi_i} f(x)[M(\xi_i) - M(\xi_{i-1})]$$

and the mass of R_{2i} is

$$\inf_{\xi_{i-1}<x<\xi_i} f(x)[M(\xi_i) - M(\xi_{i-1})].$$

If I_i is the moment of inertia with respect to the y-axis of the region $R_i = \{(x,y) : \xi_{i-1} \leq x \leq \xi_i,\ 0 \leq y \leq f(x)\}$, then for each subdivision

$$\sum_{i=1}^{n} (\xi_{i-1})^2 \inf_{\xi_{i-1} < x < \xi_i} f(x)\ [M(\xi_i) - M(\xi_{i-1})] \leq$$
$$\leq \sum_{i=1}^{n} I_i \leq \sum_{i=1}^{n} (\xi_i)^2 \sup_{\xi_{i-1} < x < \xi_i} f(x)\ [M(\xi_i) - M(\xi_{i-1})]. \quad (1)$$

Now taking the supremum of the right member and the infinum of the left member of (1) over all possible subdivisions, and using Theorem 5.2 one obtains the common value

$$\int_{[a,b]} x^2\ f(x)dM(x),$$

which is equal to $\sum_{i=1}^{n} I_i$ regardless of which set of subdivisions is considered. Since the moment of inertia of a body is the sum of the 2d moments of the disjoint parts comprizing it, the total moment of inertia I of the lamina with respect to the y-axis is

$$I = \int_{[a,b]} x^2\ f(x)dM(x).$$

Gravitational Attraction

A particle of mass m is placed at the pole of a polar coordinate system in the plane of a bounded lamina L which does not include the particle. Assume that $\{(r,\theta): (r,\theta) \in L\}$ is measurable with respect to the σ-field of 2-dimensional Borel sets in R^2. A finite number of point masses of mass $m_1, m_2, \ldots, m_k$ are distributed throughout the lamina at points (r_1,θ_1), (r_2,θ_2), ..., (r_k,θ_k), respectively (Fig. 5). At all other points of L the density $p(r,\theta)$ (mass/ unit area) is a function which is piecewise continuous in each variable separately. Let us find the component along the polar axis of the gravitational attractive force exerted on the particle by the lamina.

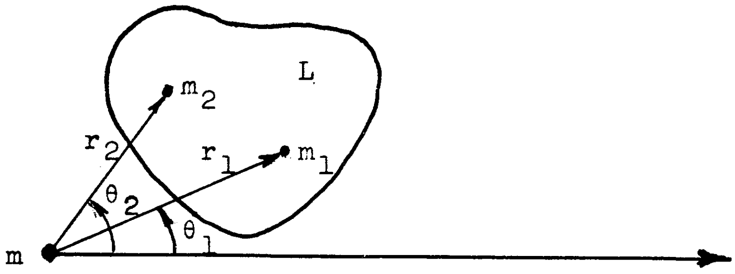


Figure 5. A Typical Lamina L with Point Masses m_1 and m_2 at Coordinates (r_1,θ_1) and (r_2,θ_2), Respectively.

Assume all points of L are designated by coordinate pairs (r,θ) such that $r > 0$, $0 \leq \theta < 2\pi$.

Define real numbers α and β by

$$\alpha = \inf_L r, \qquad \beta = \sup_L r.$$

For each $\ell \in N$, define sequences $\{s_{\ell i}\}_{i=0}^{2^{\ell-1}}$ and $\{t_{\ell j}\}_{j=0}^{2^{\ell-1}}$ by

$$s_{\ell i} = \alpha + i\left(\frac{\beta - \alpha}{2^{\ell-1}}\right) \quad ; \; i = 1, 2, \ldots, 2^{\ell-1}, \tag{1}$$

$$t_{\ell j} = j\left(\frac{2\pi}{2^{\ell-1}}\right) \quad ; \; j = 1, 2, \ldots, 2^{\ell-1}. \tag{2}$$

Then $\{s_{\ell i}\}_{i=0}^{2^{\ell-1}}$ and $\{t_{\ell j}\}_{j=0}^{2^{\ell-1}}$ partition $[\alpha,\beta]$ and $[0,2\pi]$, respectively, into $2^{\ell-1}$ equal intervals. Furthermore, each interval of any selected partition is successively halved, quartered, etc. as successive values of ℓ are chosen. Such partitions together determine successive partitions of the annular region $A = \{(r,\theta): \alpha \le r \le \beta,\ 0 \le \theta < 2\pi\}$ into $2^{2(\ell-1)}$ sectors. Let a typical element be represented by

$$A_{\ell ij} = \{(r,\theta): s_{\ell(i-1)} \le r \le s_{\ell i},\ t_{\ell(j-1)} \le \theta \le t_{\ell j}\};$$

$$\ell \in N \text{ and } i = 1,\ldots, 2^{\ell-1},\ j = 1,\ldots, 2^{\ell-1}.$$

Then the sequence of partitions $\{\{A_{\ell ij}\}\}_{\ell=1}^{\infty}$ is monotone in that for each $\ell \in N$, $A_{\ell ij} \subset A_{(\ell-1)pq}$ for some $p, q \in \{1, 2,\ldots, 2^{\ell-2}\}$.

Define a function $\rho : R^2 \to R^+$ by

$$\rho(r,\theta) = \begin{cases} p(r,\theta) & ; \ (r,\theta) \in L - \{(r_j,\theta_j)\}_{j=1}^{k}, \ 0 \le \theta < 2\pi, \\ 0 & ; \ \text{elsewhere}, \end{cases}$$

and define

$$M(r,\theta) = R\int_{-\infty}^{r} \left[R\int_{-\infty}^{\theta} \rho(u,v)\,u\,du \right] dv + \sum_{\substack{r_j \le r \text{ and} \\ \theta_j \le \theta}} m_j \ ; \ (r,\theta) \in R^2 .$$

Since ρ is non-negative and piecewise continuous in each variable separately, $M : R^2 \to R^+$ is monotone non-decreasing and piecewise continuous in each variable separately, thus is a 2-dimensional Stieltjes measure function. Note that the total mass of L is given by

$$\int_{R^2} dM(r,\theta),$$

since $\rho(s,t) = 0$ for all points

$$(s,t) \notin \{(r,\theta) : (r,\theta) \in L \text{ and } 0 \le \theta < 2\pi\}.$$

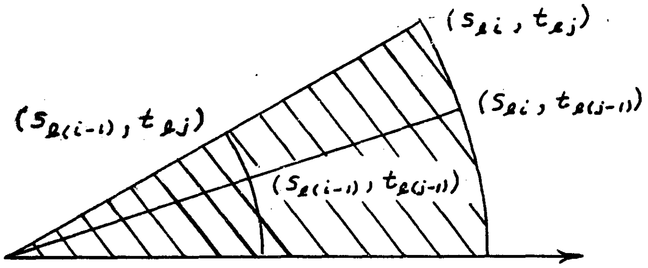
For any of the partitions defined by (1) and (2), the mass contained in the sector $A_{\ell i j}$ is given by (Fig. 6)

$$M(s_{\ell i}, t_{\ell j}) - M(s_{\ell(i-1)}, t_{\ell j}) -$$
$$- M(s_{\ell i}, t_{\ell(j-1)}) + M(s_{\ell(i-1)}, t_{\ell(j-1)})$$

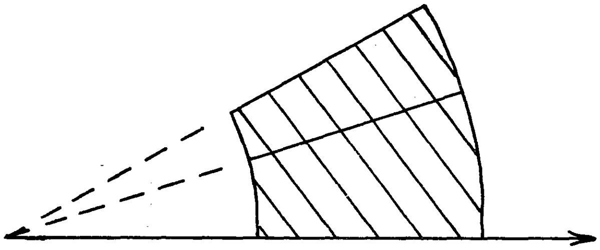
which is precisely the form of equation (1), p.23 . Thus the mass of $A_{\ell i j}$ is

$$\int_{A_{\ell i j}} dM(r,\theta).$$

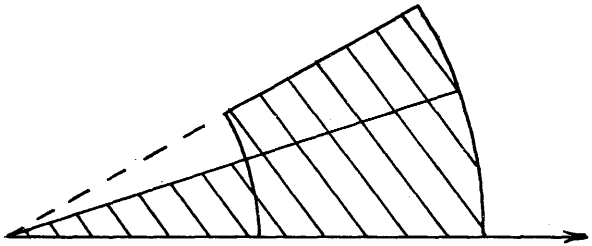
Figure 6. Pictorial Derivation of the Mass of $A_{\ell i_j}$.



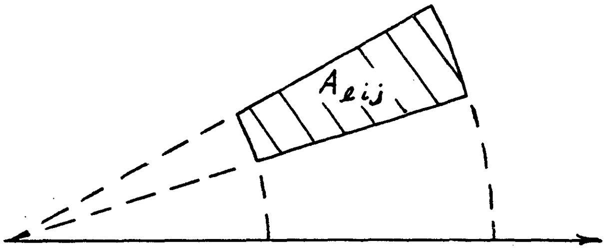
$M(s_{\ell i}, t_{\ell j})$



$M(s_{\ell i}, t_{\ell j}) - M(s_{\ell (i-1)}, t_{\ell j})$



$M(s_{\ell i}, t_{\ell j}) - M(s_{\ell (i-1)}, t_{\ell j}) + M(s_{\ell (i-1)}, t_{\ell (j-1)})$



$M(s_{\ell i}, t_{\ell j}) - M(s_{\ell (i-1)}, t_{\ell j}) - M(s_{\ell (i-1)}, t_{\ell j}) +$
$+ M(s_{\ell (i-1)}, t_{\ell (j-1)})$

Newton's law of gravitation states that the gravitational force F exerted by two particles of masses m_1 and m_2 separated by a distance r is given by

$$F = G\frac{m_1m_2}{r^2}$$

where G is the gravitational constant.

It is a well-known fact that the attractive force exerted by an object with non-zero volume on a point mass is the same as if the mass of the object were concentrated at its center of mass.

Let $\ell \in N$ be fixed but arbitrary. If $(r_{\ell i}^*, \theta_{\ell j}^*)$ represents the center of mass of $A_{\ell ij}$, then the attractive force $F_{\ell ij}$ exerted by $A_{\ell ij}$ on m is given by

$$F_{\ell ij} = G\frac{m(\text{mass of } A_{\ell ij})}{(r_{\ell i}^*)^2} = G\frac{m}{(r_{\ell i}^*)^2}\int_{A_{\ell ij}} dM(r,\theta) =$$

$$= Gm\int_{A_{\ell ij}} \frac{1}{(r_{\ell i}^*)^2}\, dM(r,\theta).$$

Therefore, the component of $F_{\ell ij}$ along the polar axis is

$$F_{\ell ij}\cos\theta_{\ell j}^* = Gm\int_{A_{\ell ij}} \frac{\cos\theta_{\ell j}^*}{(r_{\ell i}^*)^2}\, dM(r,\theta).$$

Since the total force exerted on a mass by a second body is the vector sum of the forces exerted on the mass by each disjoint portion of the second body, the total force F in the direction of the polar axis exerted on m is given by

$$F = \sum_{i=1}^{2^{\ell-1}} \sum_{j=1}^{2^{\ell-1}} F_{\ell ij} = \sum_{i=1}^{2^{\ell-1}} \sum_{j=1}^{2^{\ell-1}} Gm \int_{A_{\ell ij}} \frac{\cos\theta^*_{\ell j}}{(r^*_{\ell i})^2}\, dM(r,\theta) =$$

$$= Gm \int_L \left(\sum_{i=1}^{2^{\ell-1}} \sum_{j=1}^{2^{\ell-1}} \frac{\cos\theta^*_{\ell j}}{(r^*_{\ell i})^2}\, \chi_{A_{\ell ij}}(r,\theta) \right) dM(r,\theta), \qquad (3)$$

where $\chi_{A_{\ell ij}}$ is the characteristic function of $A_{\ell ij}$. Placing the double sum under the integral is justified by linearity of the integral in general, since for each ℓ, the $A_{\ell ij}$'s are disjoint. Note that (3) is true for each $\ell \in N$, since ℓ was arbitrary.

For each $\ell \in N$ define a function $f_\ell : R^2 \to R$ by

$$f(r,\theta) = \begin{cases} 0 & ; (r,\theta) \in R^2 - L, \\ \displaystyle\sum_{i=1}^{2^{\ell-1}} \sum_{j=1}^{2^{\ell-1}} \frac{\cos\theta^*_{\ell j}}{(r^*_{\ell i})^2}\, \chi_{A_{\ell ij}}(r,\theta) & ; (r,\theta) \in L. \end{cases}$$

Since f_ℓ is a linear combination of characteristic functions, it is $\mathcal{B}^2$-measurable.

Let us now determine whether the functions f_ℓ have a limit as $\ell \to \infty$. Let $(s,t) \in L$. Clearly, there exists a unique decreasing sequence $\{A_\ell\}_{\ell=1}^{\infty}$ of sets $A_{\ell ij}$ defined above such that $\lim_{\ell\to\infty} A_\ell = \{(s,t)\}$. Hence, the sequence of points $\{(r^*_\ell, \theta^*_\ell)\}_{\ell=1}^{\infty}$ converges to the limit point (s,t);

that is, $\lim_{\ell\to\infty} r_\ell^* = s$ and $\lim_{\ell\to\infty} \theta_\ell^* = t$. Therefore, since $s > 0$

$$\lim_{\ell\to\infty} \frac{\cos\theta_\ell^*}{(r_\ell^*)^2} = \frac{\lim_{\ell\to\infty} \cos\theta_\ell^*}{\lim_{\ell\to\infty} (r_\ell^*)^2} = \frac{\cos t}{s^2}. \qquad (4)$$

Define $f(R,\theta) = \dfrac{\cos\theta}{r^2}$; $(r,\theta) \in L$. Then since (s,t) was arbitrary in L, equation (4) implies that

$$f_\ell \to f \text{ pointwise as } \ell \to \infty. \qquad (5)$$

Finally, for each ℓ,

$$|f_\ell(r,\theta)| = \frac{|\cos\theta_\ell^*|}{(r_\ell^*)^2} \le \frac{1}{(r_\ell^*)^2} \le \frac{1}{\alpha^2} < \infty \qquad (6)$$

and

$$\int_L \frac{1}{\alpha^2}\, dM(r,\theta) = \frac{1}{\alpha^2}\int_L dM(r,\theta) = \frac{1}{\alpha^2}\,(\text{mass of } L) < \infty. \qquad (7)$$

Hence, (7) asserts that $\frac{1}{\alpha^2}$ is integrable; (3), (6), and (4) imply $\{f_\ell\}_{\ell=1}^{\infty}$ is a sequence of measurable functions bounded above by $1/\alpha^2$ and $f_\ell \to f$ pointwise as $\ell \to \infty$. Thus by the Bounded Convergence Theorem, f is integrable and

$$\int_L f_\ell(r,\theta)dM(r,\theta) \to \int_L f(r,\theta)dM(r,\theta) \text{ as } \ell \to \infty.$$

Multiplying by Gm, the total component along the polar axis of the gravitational force exerted on m by L is

$$F = Gm \int_L \frac{\cos\theta}{r^2}\, dM(r,\theta).$$

LEBESGUE-STIELTJES MEASURE IN PROBABILITY THEORY

This chapter investigates how, by means of a random variable, the probability of events in an abstract space $(\Omega, \mathcal{F}, P)$ may be placed in one-to-one correspondence with the measure of appropriate sets on a space $(R^k, \mathcal{B}^k, \mu_F^k)$, where μ_F^k is a k-dimensional Lebesgue-Stieltjes measure. The application of Lebesgue-Stieltjes measure and integration to concepts of conditional probability, expectation, and independence are then discussed. Finally, an example of Lebesgue-Stieltjes integration applied to the study of a stochastic process is presented.

Induced Probability Spaces

Let $\mathcal{F}$ be a σ-field of subsets of a specified set Ω.

DEFINITION 8.1 (Kingman and Taylor, 1966, p. 96): A probability measure on $\mathcal{F}$ is a measure $P:\mathcal{F} \to \{x: 0 \leq x \leq +\infty\}$ such that $P(\Omega) = 1$. The space $(\Omega, \mathcal{F}, P)$ is called a probability space and Ω alone is the sample space (Chung, 1968, p. 21). Elements ω belonging to Ω are called outcomes (Kingman and Taylor, 1966, p. 285).

DEFINITION 8.2 (Kingman and Taylor, 1966, p. 285): A random variable on a probability space $(\Omega, \mathcal{F}, P)$ is a measurable function from Ω into R^*.

THEOREM 8.1 (Chung, 1968, p. 33): Suppose $(\Omega, \mathcal{F}, P)$ is a probability space, $X:\Omega \to R^*$ is a random variable on $(\Omega, \mathcal{F}, P)$, and define a set function $\mu:\mathcal{B} \to R^*$ by

$$\mu(B) = P\{X^{-1}(B)\} = P\{\omega:X(\omega) \in B\} \quad ; B \in \mathcal{B}.$$

Then $(R, \mathcal{B}, \mu)$ is a probability space ($\mathcal{B}$ is the σ-field of Borel sets in R).

Proof: One must show that $\mu:\mathcal{B} \to R^*$ is a measure such that $\mu(R) = 1$.

Since X is a random variable, $X^{-1}(B) \in \mathcal{F}$ for each $B \in \mathcal{B}$. Since P is a measure, $P\{X^{-1}(B)\} \geq 0$ for each $B \in \mathcal{B}$. Choose an arbitrary sequence $\{B_n\}_{n=1}^{\infty}$ of disjoint sets from $\mathcal{B}$. Since X is a function, $\{X^{-1}(B_n)\}_{n=1}^{\infty}$ is a sequence of disjoint sets from $\mathcal{F}$. Thus

$$\mu \bigcup_{n=1}^{\infty} B_n = P\left(X^{-1} \bigcup_{n=1}^{\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} X^{-1}(B_n)\right) =$$

$$= \sum_{n=1}^{\infty} P(X^{-1}(B_n)) = \sum_{n=1}^{\infty} \mu(B_n).$$

Finally, $\mu(R) = P(X^{-1}(R)) = P(\Omega) = 1$. Thus μ is a probability measure on $\mathcal{B}$, and $(R, \mathcal{B}, \mu)$ is a probability space.

The space $(R, \mathcal{B}, \mu)$ is called the probability space <u>induced</u> by X on $(\Omega, \mathcal{F}, P)$.

The Distribution Function

DEFINITION 8.3 (Kingman and Taylor, 1966, p. 96): A distribution function is a function $F:R \to R$ such that

(i) F is monotone non-decreasing,

(ii) F is everywhere continuous on the right,

(iii) $F(x) \to 0$ as $x \to -\infty$, $F(x) \to 1$ as $x \to +\infty$.

Therefore, a distribution function is a Stieltjes measure function which satisfies property (iii) above. Let us now see how Lebesgue-Stieltjes measure is related to probability theory.

THEOREM 8.2 (Adapted from Chung, 1968, p. 23): If $(R, \mathcal{B}, \mu)$ is a probability space, then μ determines a distribution function F through the correspondence

$$\mu((-\infty,x]) = F(x) \quad ; x \in R.$$

Furthermore, μ is the unique Lebesgue-Stieltjes measure on $(R, \mathcal{B})$ determined by F.

Proof: Monotonicity of μ implies F is monotone non-decreasing. For arbitrary $x \in R$ let $\{x_n\}_{n=1}^{\infty}$ be any monotone decreasing sequence from R converging to x. Since μ is σ-additive, Theorem 2.3,i applies, and

$$\lim_{n\to\infty} \mu(-\infty,x_n] = \mu(-\infty,x].$$

Thus F is continuous on the right. Similarly,

$$\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} \mu(-\infty,x] = \mu(\phi) = 0.$$

Finally, since μ is a probability measure,

$$\lim_{x\to+\infty} F(x) = \lim_{x\to+\infty} \mu(-\infty,x] = \mu(R) = 1.$$

That μ is the unique Lebesgue-Stieltjes measure on $(R, \mathcal{B})$ determined by F is seen by defining a set function ϕ on the semi-ring $\mathcal{P}$ of intervals of the form (a,b] by

$$\phi((a,b]) = F(b) - F(a) \quad ; a, b \in R \text{ and } a \leq b$$

and then using Theorem 2.4 and the Extension Theorem to uniquely extend ϕ to the measure μ defined on $\mathcal{B}$, the σ-field generated by $\mathcal{P}$.

In summary, given a probability space $(\Omega, \mathcal{F}, P)$ and a random variable $X:\Omega \to R^*$, the probability space $(R, \mathcal{B}, \mu)$ induced by the correspondence

$$\mu(B) = P\{X^{-1}(B)\} \quad ; B \in \mathcal{B}$$

is completely described by the distribution function $F:R \to \{x:x \geq 0\}$ defined by

$$F(x) = \mu(-\infty,x] \quad ; x \in R.$$

Furthermore, it is observed that F is a Stieltjes measure function, and that μ is the unique Lebesgue-Stieltjes measure on $\mathcal{B}$ determined by F.

Functions of Random Variables and Multi-Dimensional Probability Spaces

The following theorems concerning random variables are usually proved for measurable functions in general. In probability theory, however, the spaces induced by random variables are as important in their own right as the behavior of the measurable functions themselves.

In this section, all random variables will be understood to be defined on a probability space $(\Omega, \mathcal{F}, P)$.

DEFINITION 8.4 (Chung, 1968, p. 33): A function $f:R \to R^*$ is said to be <u>Borel measurable</u> if f is $\mathcal{B}$-measurable, where $\mathcal{B}$ is the σ-field of Borel sets in R.

THEOREM 8.3 (Chung, 1968, p. 34): If X is a random variable and f is a Borel measurable function, then $f(X)$ is a random variable.

Proof: Let B be a Borel set. Then $f^{-1}(B) \varepsilon \mathcal{B}$, since f is Borel measurable. Also, $X^{-1}(f^{-1}(B)) \varepsilon \mathcal{F}$ since X is $\mathcal{F}$-measurable. Now

$$[f(X)]^{-1}(B) = (X \circ f)^{-1}(B) = (f^{-1}X^{-1})(B) = X^{-1}(f^{-1}(B)) \varepsilon \mathcal{F}.$$

Since B was arbitrary, $f(X)$ is $\mathcal{F}$-measurable; therefore, $f(X)$ is a random variable from the space $(\Omega, \mathcal{F}, P)$ to the space induced by f on the X-induced space.

DEFINITION 8.5 (Chung, 1968, p. 34): A function $f:R^n \to R^*$ of n variables is said to be <u>Borel measurable</u> if it is $\mathcal{B}^n$-measurable, where $\mathcal{B}^n$ is the σ-field of Borel sets in R^n.

If X and Y are random variables on $(\Omega, \mathcal{F}, P)$ denote by (X,Y) the function defined by

$$(X,Y)(\omega) = (X(\omega),Y(\omega)) \quad ; \omega \varepsilon \Omega,$$

where $(X(\omega),Y(\omega))$ is to be regarded as an ordered pair in R^2.

THEOREM 8.4 (Chung, 1968, p. 35): If X and Y are random variables and f is a Borel measurable function of two variables, then f(X,Y) is a random variable.

Proof: Let B be a Borel set in $\mathcal{B}$. Then $f^{-1}(B) \in \mathcal{B}^2$. Denote $f^{-1}(B)$ by $A_1 \times A_2$ where $A_1 \in \mathcal{B}$ and $A_2 \in \mathcal{B}$. Now

$$(X,Y)^{-1}(A_1 \times A_2) = \{a : X(a) \in A_1 \text{ and } Y(a) \in A_2\} =$$
$$= X^{-1}(A_1) \cap Y^{-1}(A_2) \in \mathcal{F}$$

since X and Y are $\mathcal{F}$-measurable. Since B was arbitrary, f(X,Y) is $\mathcal{F}$-measurable and therefore is a random variable.

In Theorem 8.4, (X,Y) was considered as a function from the space $(\Omega, \mathcal{F})$ into the space $(R^2, \mathcal{B}^2)$. By defining a set function $\nu : \mathcal{B}^2 \to R$ by

$$\nu(B) = P\{(X,Y)^{-1}(B)\} \quad ; B \in \mathcal{B}^2,$$

one has the induced probability space $(R^2, \mathcal{B}^2, \nu)$. The associated 2-dimensional distribution function has precisely the form of the 2-dimensional Stieltjes measure function of Definition 4.1. Accordingly, ν is a 2-dimensional Lebesgue-Stieltjes measure.

COROLLARY (Chung, 1968, p. 35): If X is a random variable and $f : R \to R$ is a continuous function, then f(X) is a random variable. In particular,

$$X^r \;;\; r \in N, \quad |X|^r \;;\; r \in R, \quad e^{-\lambda X} \;;\; \lambda \in R$$

are random variables. If X and Y are random variables, then

$$\max(X,Y), \quad \min(X,Y), \quad X + Y, \quad X - Y, \quad X \cdot Y, \quad X/Y$$

are random variables, the last provided $Y(\omega) \neq 0$ for each $\omega \in \Omega$.

Similar results hold for functions of a finite number of random variables; for a countably infinite number of random variables, one has the following theorem.

THEOREM 8.5 (Chung, 1968, p. 35): If $\{X_j\}_{j=1}^{\infty}$ is a sequence of random variables, then

$$\inf_j X_j, \quad \sup_j X_j, \quad \liminf_j X_j, \quad \limsup_j X_j$$

are random variables, not necessarily finite valued.

Conditional Probability

In this section, all concepts will be understood relative to a probability space $(\Omega, \mathcal{F}, P)$.

DEFINITION 8.6 (Chung, 1968, p. 21): The _trace of a subset $\Delta \subset \Omega$ on $\mathcal{F}$_ is the collection $\Delta \cap \mathcal{F}$ defined by

$$\Delta \cap \mathcal{F} = \{\Delta \cap A : A \in \mathcal{F}\}.$$

DEFINITION 8.7 (Chung, 1968, p. 21): For a specified subset $\Delta \subset \Omega$ such that $P(\Delta) > 0$, define a probability measure $P_\Delta : \Delta \cap \mathcal{F} \to \{x : 0 \leq x \leq 1\}$ by

$$P_\Delta(A) = \frac{P(A \cap \Delta)}{P(\Delta)} \quad ; A \in \Delta \cap \mathcal{F}.$$

The probability space $(\Delta \cap \Omega, \Delta \cap \mathcal{F}, P_\Delta)$ is called the _trace of $(\Omega, \mathcal{F}, P)$ on Δ_.

DEFINITION 8.8 (Kingman and Taylor, 1966, p. 272): Given subsets $A \in \mathcal{F}$, $B \subset \Omega$ such that $P(B) \neq 0$, the <u>conditional probability of A given B</u>, denoted by $P(A|B)$, is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Note that $P(A|B) = P_B(A \cap B)$; $A \in \mathcal{F}$.

<u>Expectation</u>

The concept of expectation is that of integration with respect to the probability measure.

DEFINITION 8.9 (Kingman and Taylor, 1966, p. 287): If X is a random variable on a probability space $(\Omega, \mathcal{F}, P)$, the expectation $E(X)$ of X is defined by

$$E(X) = \int_\Omega X dP,$$

provided X is integrable with respect to P.

THEOREM 8.6 (Kingman and Taylor, 1966, p. 291): If $(R, \mathcal{B}, \mu)$ is the space induced by a random variable X on the probability space $(\Omega, \mathcal{F}, P)$, $F: R \to \{x : 0 \leq x \leq 1\}$ is the distribution function determined by μ, and $g: R \to R^*$ is a Borel measurable function, then

$$E[g(X)] = \int_\Omega g(X) dP = \int_R g(t) dF(t)$$

provided that either integral exists.

Proof: By Theorem 8.3 , $g(X)$ is a random variable and thus is $\mathcal{F}$-measurable. Define functions $g_+:R \to R^+$ and $g_-:R \to R^+$ by

$$g_+(t) = \begin{cases} g(t) & ; \ g(t) > 0, \\ 0 & ; \ g(t) \leq 0, \end{cases}$$

$$g_-(t) = \begin{cases} -g(t) & ; \ g(t) < 0, \\ 0 & ; \ g(t) \geq 0. \end{cases}$$

Borel measurability of g_+ and g_- follows from Theorem 2.7 and $\mathcal{F}$-measurability of g_+ and g_- then follows from Theorem 8.3 . Therefore, let $\{g_{+n}\}_{n=1}^{\infty}$ be a monotone non-decreasing sequence of $\mathcal{B}$-simple functions converging to g_+. Referring to Definition 2.16, for each n, g_{+n} may be represented in terms of a $\mathcal{B}$-dissection $\{B_i\}_{i=1}^{j_n}$ of R as

$$g_{+n}(t) = \sum_{i=1}^{j_n} c_i \chi_{B_i}(t) \ ; \ t \in R.$$

By definition of the integral of a simple function,

$$\int_R g_{+n} d\mu = \sum_{i=1}^{j_n} c_i \mu(B_i).$$

But $\mu(B_i) = P(X^{-1}(B_i))$ $(i = 1, 2, \ldots, j_n)$ by Theorem 8.1 , and so

$$\int_R g_{+n} d\mu = \sum_{i=1}^{j_n} c_i P(X^{-1}(B_i)) = \sum_{i=1}^{j_n} c_i \int_{X^{-1}(B_i)} dP =$$
$$= \sum_{i=1}^{j_n} \int_{X^{-1}(B_i)} c_i dP = \sum_{i=1}^{j_n} \int_{X^{-1}(B_i)} g_{+n}(X) dP = \int_{\Omega} g_{+n}(X) dP, \tag{1}$$

the last equality following from Theorem 2.9,i, since X a function implies $X^{-1}(B_i) \cap X^{-1}(B_k) = \phi$ unless $i = k$.

Application of the Monotone Convergence Theorem then gives

$$\lim_{n\to\infty} \int_R g_{+n} d\mu = \int_R g_+ d\mu,$$

$$\lim_{n\to\infty} \int_\Omega g_{+n}(X) dP = \int_\Omega g_+(X) dP,$$

provided either side exists. Term-for-term equality implied by (1) then gives

$$\int_\Omega g_+(X) dP = \int_R g_+ d\mu.$$

An identical argument for g_- gives

$$\int_\Omega g_-(X) dP = \int_R g_- d\mu.$$

The right side is expressible as the Lebesgue-Stieltjes integral

$$\int_R g(x) dF(x)$$

by Theorem 8.2.

Independence and Product Measures

DEFINITION 9.10 (Chung, 1968, p. 47): Let $\{X_i\}_{i=1}^n$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$. The random variables $X_1, X_2, \ldots, X_n$ are independent if for each sequence $\{B_i\}_{i=1}^n$ of Borel sets in R,

$$P\left(\bigcap_{i=1}^n X_i^{-1}(B_i)\right) = \prod_{i=1}^n P(X_i^{-1}(B_i)).$$

Throughout the remainder of this section, the following notational correspondences are observed, relative to Definition 8.10 .

(i) $(R^n, \mathcal{B}^n, \mu^n)$: the probability space induced by the correspondence

$$\mu^n(B^n) = P\{(X_1, X_2, \ldots, X_n)^{-1}(B^n)\} \; ; \; B^n \in \mathcal{B}^n,$$

(ii) $(R, \mathcal{B}, \mu_i)$: the probability space induced by the correspondence

$$\mu_i(B) = P\{X_i^{-1}(B)\} \; ; \; B \in \mathcal{B},$$

(iii) $F^n : R^n \rightarrow \{t : 0 \leq t \leq 1\}$: the n-dimensional distribution function determined by μ^n,

(iv) $F_i : R \rightarrow \{t : 0 \leq t \leq 1\}$: the distribution function determined by μ_i.

THEOREM 8.7 (Chung, 1968, p. 47): The following statements are equivalent:

(i) $X_1, X_2, \ldots, X_n$ are independent random variables.

(ii) For each sequence $\{B_i\}_{i=1}^n$ from $\mathcal{B}$,

$$P\left(\bigcap_{i=1}^{n} X_i^{-1}(B_i)\right) = \prod_{i=1}^{n} P\{X_i^{-1}(B_i)\}.$$

(iii) For each sequence $\{B_i\}_{i=1}^n$ from $\mathcal{B}$,

$$\mu^n\left(\bigtimes_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} \mu_i(B_i).$$

(iv) For each sequence $\{t_i\}_{i=1}^n$ from R,

$$F^n(t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} F_i(t_i).$$

(v) For each sequence $\{t_i\}_{i=1}^n$ from R,

$$P\left(\bigcap_{i=1}^n X_i^{-1}((-\infty,t_i])\right) = \prod_{i=1}^n P\{X_i^{-1}((-\infty,t_i])\}.$$

Proof that (ii) implies (iii):

$$P\left(\bigcap_{i=1}^n X_i^{-1}(B)\right) = P\left(\bigcap_{i=1}^n \{\omega : X_i(\omega) \in B_i\}\right) =$$

$$= P\{\omega : (X_1(\omega), X_2(\omega), \ldots, X_n(\omega)) \in B_1 \times B_2 \times \cdots \times B_n\} =$$

$$= P\left[(X_1, X_2, \ldots, X_n)^{-1}\left(\bigtimes_{i=1}^n B_i\right)\right] =$$

$$= \mu^n\left(\bigtimes_{i=1}^n B_i\right) \quad ; B_i \in \mathcal{B},$$

and

$$\prod_{i=1}^n P\{X_i^{-1}(B_i)\} = \prod_{i=1}^n \mu_i(B_i) \quad ; B_i \in \mathcal{B}.$$

Thus

$$\mu^n\left(\bigtimes_{i=1}^n B_i\right) = \prod_{i=1}^n \mu(B_i).$$

Proof that (iii) implies (iv): In particular, for any sequence $\{t_i\}_{i=1}^n$.

$$\bigtimes_{i=1}^n (-\infty, t_i] \in \mathcal{B}^n \text{ and}$$

$$(-\infty, t_i] \in \mathcal{B}, \ (i = 1, 2, \ldots, n).$$

Therefore, if (iii) is true,

$$\mu^n\left(\bigtimes_{i=1}^n (-\infty, t_i]\right) = \prod_{i=1}^n \mu_i((-\infty, t_i]).$$

Writing each side in terms of the (uniquely determined) distribution functions,

$$F^n(t_1,t_2,\ldots,t_n) = \prod_{i=1}^{n} F_i(t_i).$$

The other implications follow from Definition 9.10 and the correspondence between $(\Omega, \mathcal{F}, P)$ and the induced spaces $\{(R, \mathcal{B}, \mu_i)\}_{i=1}^{n}$ and $(R^n, \mathcal{B}^n, \mu^n)$.

Since $\mu^n\left(\bigtimes_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} \mu_i(B_i)$,

$$\int_{R^n} d\mu^n \text{ is written } \iint_{R^n}\cdots\int d\mu_1 d\mu_2 \cdots d\mu_n.$$

THEOREM 8.8 (Chung, 1968, p. 48): If $\{X_i\}_{i=1}^{n}$ are independent random variables and $\{f_i : R \to R^*\}_{i=1}^{n}$ are Borel measurable functions, then $\{f_i(X_i)\}_{i=1}^{n}$ are independent random variables.

THEOREM 8.9 (Chung, 1968, p. 49): If X and Y are independent random variables having finite expectations, then

$$E(XY) = E(X)E(Y).$$

Proof: Denote elements $X(\omega)$ of the space induced by X on Ω by x and denote elements $Y(\omega)$ of the space induced by Y on Ω by y. Then

$$E(XY) = \int_\Omega XY dP = \iint_{R^2} xy\, d\mu_1 d\mu_2 = \int_R x\, d\mu_1 \int_R y\, d\mu_2 =$$

$$= \int_\Omega X dP \int_\Omega Y dP = E(X)E(Y).$$

Example of Lebesgue-Stieltjes Integration Applied to a Stochastic Process

> Suppose that electrical pulses having random impulses arrive at a counter in accordance with a Poisson process with rate λ. The amplitude of the pulses is assumed to decrease at an exponential rate. That is, we suppose that if the pulse has an amplitude of Y units upon arrival, then its amplitude at a time t units later will be $Ye^{-\alpha t}$. We further suppose that the initial amplitudes of the pulses are independent and have a common distribution F (Ross, 1970, p. 19).

Before proceeding with the analysis, let us first review a few facts about the Poisson process.

DEFINITION 8.11 (Ross, 1970, p. 13): Denote by $\{N(t):t \geq 0\}$ a counting process, where $N(t)$ represents the total number of events which have occurred up to and including time t. A process $\{N(t):t \geq 0\}$ is said to be a <u>Poisson process</u> if

(i) $N(0) = 0$,

(ii) for each monotone increasing finite sequence $\{t_i\}_{i=1}^{n}$ ($n \in N$), the random variables

$$N(t_1) - N(t_0),\ N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$$

are independent (i.e. $\{N(t):t \geq 0\}$ has independent increments),

(iii) for each $s, t \geq 0$, there exists $\lambda \in R$ such that

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad ; n \in N \cup \{0\}.$$

THEOREM 8.10 (Ross, 1970, p. 18): Denote by S_i the time of the i^{th} event. Then given that $N(t) = n$, the n arrival times $S_1, S_2, \ldots, S_n$ considered as unordered random variables are distributed independently and uniformly in the interval $[0,t]$.

Returning to the electronic counter process, let us first describe the probability space at hand.

For fixed $t \geq 0$, one is interested in the number of pulses $N(t)$ which have arrived by time t, together with the arrival amplitudes Y_i $(i = 1, 2, \ldots, N(t))$ which together with the respective arrival times determine the amplitudes $Y_i(t)$ at time t. Denote realizations of the random variables $N(t)$ and Y_i by n and y_i, respectively. Define

$A_0 = \{(0)\}$,

$A_n = \{(n, y_1, y_2, \ldots, y_n) : y_i \in R \ (i = 1,2,\ldots,n)\}$; $n \in N$.

Then the range Λ of the random vector $(N(t), Y_1, Y_2, \ldots, Y_{N(t)})$ is

$$\Lambda = \bigcup_{n=0}^{\infty} A_n .$$

For each $n \in N \cup \{0\}$, the collection of subsets of A_n of the form $\{n\} \times \mathcal{B}^n$, where $\mathcal{B}^n$ is the σ-field of Borel sets in R^n, is a σ-field $\mathcal{A}_n$ of subsets of A_n.

Let $\mathcal{A}$ be the σ-field generated by all sets of the form

$$\left\{ \prod_{n=0}^{\infty} C_n : C_n \in \mathcal{A}_n \ (n \in N) \right\}.$$

Then the pair $(\Lambda, \mathscr{A})$ is a measurable space.

Define a function $G:R \to [0,1]$ by

$$G(x) = P\{N(t) \leq x\} = \begin{cases} \sum_{n=0}^{\langle x\rangle} \frac{e^{-\lambda t}(\lambda t)^{\langle x\rangle}}{\langle x\rangle !} & ; x \geq 0, \\ 0 & ; x < 0. \end{cases}$$

Then G is a distribution function and defines a unique probability measure $\phi:\mathscr{B} \to [0,1]$ such that for each $B \in \mathscr{B}$,

$$\phi(B) = P\{N(t) \in B\} = P\{N(t) \in B \cap (N \cup \{0\})\} = \int_B dG(x).$$

Similarly, the common distribution F of the arrival amplitudes Y_i $(i \in N)$ determines unique identical measures $\psi_i:\mathscr{B} \to [0,1]$ such that for each $n \in N$,

$$\psi_i(B) = P\{Y_i \in B\} = \int_B dF(y) \quad ; B \in \mathscr{B}, \ (i \in N).$$

Since the Y_i's are independent, Definition 8.10 implies that for fixed $n \in N$ and for each sequence $\{B_i\}_{i=1}^n$ from $\mathscr{B}$,

$$P\{Y_i \in B_i \ (i = 1,2,\ldots,n) | N(t) = n\} = \prod_{i=1}^{n} P\{Y_i \in B_i\} =$$

$$= \iint\cdots\int_{\prod_{i=1}^{n} B_i} dF(y_1)dF(y_2)\cdots dF(y_n).$$

Let us now determine a Stieltjes integral representation of the distribution function of the total amplitude $Y(t)$ at any fixed time $t \geq 0$. Now

$$Y(t) = \sum_{i=1}^{N(t)} Y_i(t) = \sum_{i=1}^{N(t)} Y_i e^{-\alpha(t-S_i)}$$

where Y_i is the amplitude of the i^{th} pulse at the time S_i

of arrival at the counter.

Let $q \varepsilon R$. Suppose for the moment that $\alpha = 0$ so that $Y_i(t) = Y_i$ ($i = 1, 2, \ldots, N(t)$). One now uses the fact that

$$P\{Y(t) \leq q\} = \sum_{n=0}^{\infty} P\{Y(t) \leq q | N(t) = n\} P\{N(t) = n\} = \qquad (1)$$
$$= \int_R P\{Y(t) \leq q | N(t) = n\} dG(x).$$

Also

$$P\{Y(t) \leq q | N(t) = 0\} = P\{(0)\} = 0$$

since there is no amplitude to be measured,

$$P\{Y(t) \leq q | N(t) = 1\} = P\{(1, y_1) : y_1 \leq q\} = \int_{\Delta_1(q)} dF(y_1)$$

where $\Delta_1(q) = \{y_1 : y_1 \leq q\}$ as in Figure 7,

$$P\{Y(t) \leq q | N(t) = 2\} = P\{(2, y_1, y_2) : y_1 + y_2 \leq q\} =$$
$$= \iint_{\Delta_2(q)} dF(y_1) dF(y_2)$$

where $\Delta_2(q) = \{(y_1, y_2) : y_1 + y_2 \leq q\}$ as in Figure 8,

$$P\{Y(t) \leq q | N(t) = n\} = P\{(n, y_1, y_2, \ldots, y_n) : \sum_{i=1}^{n} y_i \leq q\} = \qquad (2)$$
$$\iint \cdots \int_{\Delta_n(q)} dF(y_1) dF(y_2) \cdots dF(y_n)$$

where $\Delta_n(q) = \{(y_1, y_2, \ldots, y_n) : \sum_{i=1}^{n} y_i \leq q\}$. The portion of the set $\Delta_3(q)$ for which $y_1 \geq 0$, $y_2 \geq 0$, $y_3 \geq 0$ is shown in Figure 9.

Figure 7. The Linear Region $\Delta_1(q)$.

Figure 8. The Planar Region $\Delta_2(q)$.

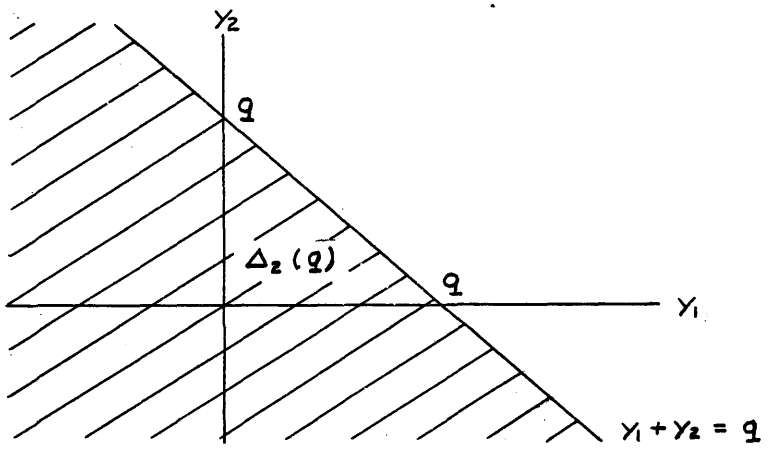
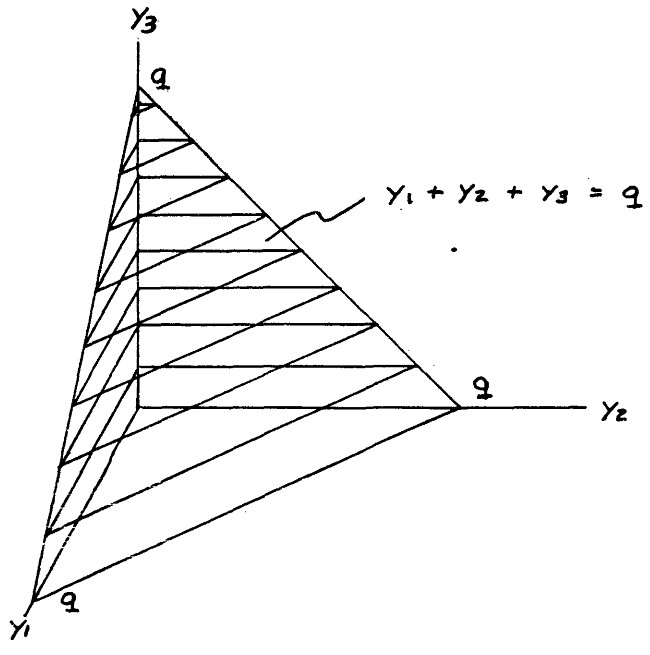


Figure 9. The Volumetric Region $\Delta_3(q)$ for $y_1 \geq 0$, $y_2 \geq 0$, $y_3 \geq 0$.



Combining (1) and (2) yeilds

$$P\{Y(t) \leq q\} = \int_R \left[\iint_{\Delta_n(q)} \cdots \int dF(y_1)dF(y_2)\cdots dF(y_n) \right] dG(x).$$

Now suppose $\alpha \neq 0$ so that $Y(t) = \sum_{i=1}^{N(t)} Y_i e^{-\alpha(t - S_i)}$.

Then for fixed $n \in N$ and for fixed $q \in R$,

$$P\{Y(t) \leq q | N(t) = n\} = P\{\sum_{i=1}^{N(t)} Y_i e^{-\alpha(t-S_i)} \leq q | N(t) = n\} =$$

$$= P\{\sum_{i=1}^{n} Y_i e^{-\alpha(t - S_i)} \leq q\}. \tag{3}$$

Now Theorem 8.8 implies that for each sequence $\{s_i\}_{i=1}^n$ from $[0,t]$,

$$P\{S_1 \leq s_1, S_2 \leq s_2, \ldots, S_n \leq s_n\} = \prod_{i=1}^n P\{S_i \leq s_i\} =$$

$$= \prod_{i=1}^n \left[\frac{1}{t} R \int_0^{s_i} ds \right] = \frac{1}{t^n} \left(\prod_{i=1}^n s_i \right). \tag{4}$$

Now differentiating (4) with respect to s_i, $i = 1, 2, \ldots, n$, one obtains

$$P\{S_1 = s_1, S_2 = s_2, \ldots, S_n = s_n\} = \frac{1}{t^n} ds_1 ds_2 \cdots ds_n.$$

Therefore, conditioning on the S_i, $i = 1, 2, \ldots, n$, in (3) gives

$$P\{\sum_{i=1}^{n} Y_i e^{-\alpha(t - S_i)} \leq q\} =$$

$$= \iint_{[0,t]^n}\cdots\int P\{\sum_{i=1}^{n} Y_i e^{-\alpha(t-S_i)} \leq q \,|\, (S_1,\ldots,S_n) = (s_1,\ldots,s_n)\} \frac{ds_1\cdots ds_n}{t^n} =$$

$$= \frac{1}{t^n}\iint_{[0,t]^n}\cdots\int P\{\sum_{i=1}^{n} Y_i e^{-\alpha(t - s_i)} \leq q\} ds_1 ds_2\cdots ds_n. \qquad (5)$$

Consider the integrand in (5). For each i = 1, 2,...,n,

$$P\{Y_i e^{-\alpha(t - s_i)} \leq q\} = P\{Y_i \leq qe^{\alpha(t - s_i)}\} =$$

$$= \int_{(-\infty, qe^{\alpha(t-s_i)}]} dF(y_i).$$

If N(t) = 1,

$$P\{Y_1 e^{-\alpha(t-s_1)} \leq q\} = P\{(1,y_1): y_1 \leq qe^{\alpha(t-s_1)}\} =$$

$$= \int_{\Delta_1(q,s_1)} dF(y_1)$$

where $\Delta_1(q,s_1) = \{y: y \leq qe^{\alpha(t - s_1)}\}$ as in Figure 10 .

If N(t) = 2,

$$P\{Y_1 e^{-\alpha(t-s_1)} + Y_2 e^{-\alpha(t-s_2)} \leq q\} =$$

$$= P\{(2,y_1,y_2): y_1 e^{-\alpha(t-s_1)} + y_2 e^{-\alpha(t-s_2)} \leq q\} =$$

$$= \int_{\Delta_2(q,s_1,s_2)} dF(y_1)dF(y_2)$$

where $\Delta_2(q,s_1,s_2) = \{(y_1,y_2):y_1e^{-\alpha(t-s_1)} + y_2e^{-\alpha(t-s_2)} \leq q\}$ as in Figure 11.

If $N(t) = n$,

$$P\{\sum_{i=1}^{n} Y_i e^{-\alpha(t-s_i)} \leq q\} = \iint\cdots\int_{\Delta_n(q,s_1,\ldots,s_n)} dF(y_1)\cdots dF(y_n)$$

where $\Delta_n(q,s_1,\ldots,s_n) = \{(y_1,\ldots,y_n):$

$$y_1e^{-\alpha(t-s_1)} + \ldots + y_ne^{-\alpha(t-s_n)} \leq q\}.$$

Note that linearity of the terms following the colon indicates that the $\Delta_n(q,s_1,\ldots,s_n)$ have the same shape as the $\Delta_n(q)$ $(n = 1, 2, \ldots)$ considered earlier (that is, they are bounded by a line, plane, etc.), but that the dimensions are scaled down by the exponential components of the expressions, depending on how long before time t the i^{th} event took place.

Therefore, combining (1), (3), (5), and (6) yields

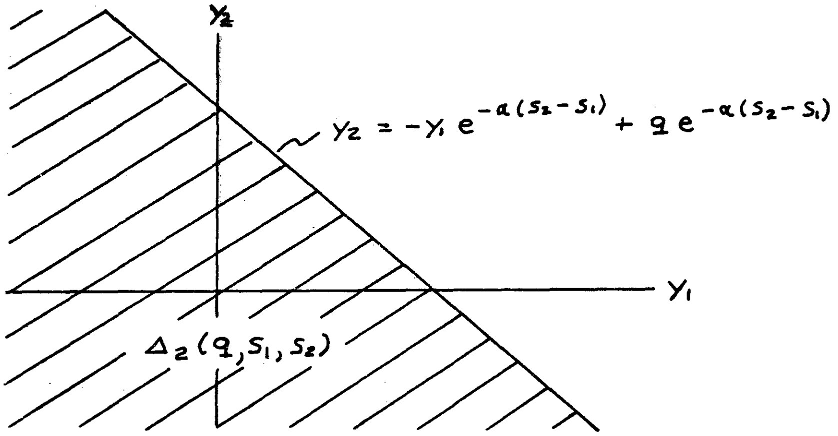
$P\{Y(t) \leq q\} =$

$$= \int_R \frac{1}{t^x}\left\{\iint\cdots\int_{[0,t]^{\langle x\rangle}}\left[\iint\cdots\int_{\Delta_{\langle x\rangle}(q,s\ ,\ldots,s_{\langle x\rangle})} dF(y_1)\cdots dF(y_{\langle x\rangle})\right] ds_1\cdots ds_{\langle x\rangle}\right\}dG(x).$$

This Lebesgue-Stieltjes representation is useful in that it contains in somewhat explicit fashion the information used in its derivation. However, if the distribution function of the amplitudes on arrival (F), the rate

Figure 10. The Linear Region $\Delta_1(q,s_1)$.

Figure 11. The Planar Region $\Delta_2(q,s_1,s_2)$.



of decay (α), and the rate of the process (λ) were specified, one would probably work with the characteristic function $E\, e^{iuY(t)}$ in hopeful anticipation that its form might be recognizable as the characteristic function of some commonly encountered distribution.

CONCLUSIONS

The type of integral which may be defined on a given space is completely determined by the structure of the σ-field under consideration and the measure defined on the σ-field. Lebesgue-Stieltjes integration is integration over a σ-field of subsets of R^k on which a Lebesgue-Stieltjes measure is defined (this σ-field contains the Borel sets in R^k). Any Lebesgue-Stieltjes measure is completely determined by its Stieltjes measure function.

The Lebesgue-Stieltjes integral, though computationally awkward as compared to the Riemann-Stieltjes integral, is useful in that its form displays explicitly the information about the structure of the space used in the construction of the integral. This type of integral is therefore widely used in the study of probability and stochastic processes, since the form of the integral for an expectation indicates the manner in which a space is partitioned when conditioning with respect to random variables, and since the integral representation for the probability of an event displays the role of the distribution function.

The Riemann-Stieltjes integral has the advantage that it may be expressed as a Riemann integral plus a summation, and thus is computationally efficient. This type of integral is especially useful in expressing cumulative values of physical quantities such as length, area, volume, mass, or force. When possible, Lebesgue-Stieltjes integrals are evaluated as Riemann-Stieltjes integrals.

LITERATURE CITED

Chung, Kai Lai, 1968, A course in probability theory: New York, Harcourt, Brace and World, Inc., p. 1 - 49.

Gaughan, Edward, 1968, Introduction to analysis: Belmont, Calif., Wadsworth Publishing Company, p. 135 - 180.

Gemignani, Michael C., 1972, Elementary topology, 2d ed.: Reading, Mass., Addison-Wesley Publishing Company, p. 84.

Kingman, J. F. C., and Taylor, S. J., 1966, Introduction to measure and probability: Cambridge, G. B., Cambridge University Press. 401 p.

Ross, Sheldon M., 1970, Applied probability models with optimization applications: San Francisco, Calif., Holden-Day, p. 13 - 19.

Royden, H. L., 1968, Real analysis, 2d ed.: New York, The Macmillan Company, p. 149 - 180.

Rudin, Walter, 1964, Principles of mathematical analysis: New York, McGraw-Hill, Inc., p. 106 - 107.

Widder, David V., 1961, Advanced calculus: Englewood Cliffs, N. J., Prentice Hall, Inc. p. 70 - 75.