View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Governors State University

Governors State University
OPUS Open Portal to University Scholarship

All Student Theses | Student Theses

Summer 2016

A Little Aspect of Real Analysis, Topology and Probability

Asmaa A. Abdulhameed
Governors State University

Follow this and additional works at: http://opus.govst.edu/theses

Part of the Mathematics Commons

Recommended Citation

Abdulhameed, Asmaa A., "A Little Aspect of Real Analysis, Topology and Probability" (2016). *All Student Theses.* Paper 81.

For more information about the academic degree, extended learning, and certificate programs of Governors State University, go to http://www.govst.edu/Academics/Degree_Programs_and_Certifications/

Visit the Governors State Mathematics Department
This Thesis is brought to you for free and open access by the Student Theses at OPUS Open Portal to University Scholarship. It has been accepted for inclusion in All Student Theses by an authorized administrator of OPUS Open Portal to University Scholarship. For more information, please contact opus@govst.edu.

A LITTLE ASPECT OF REAL ANALYSIS, TOPOLOGY AND PROBABILITY

By

Asmaa A. Abdulhameed

B.A., Trinity Christian College, 2012

Thesis

Submitted in partial fulfillment of the requirements

For the Degree of Master of Science,
With a Major in Mathematics

Governors State University
University Park, IL 60484

2016

Table of Contents

List of Figures

List of Tables

ABSTRACT

The body of the paper is divided into three parts:

Part one: include definitions and examples of the metric, norm and topology along with some important terms such as metrics $\ell_1, \ell_2, \ell_\infty$ and vector norm $|x|_p$ which also known as the L^p space. In case of the topology space this concept adjusted to be a unit ball that is distance one from a point unit circle.

Part two: demonstrate the measure and probability which is one of the main topics in this work. This section serves as an introduction for the remaining part. It explains the construction of the σ-algebra and the Borel set, and the usefulness of the Borel set in the probability theory.

Part Three: to understand the measure, one has to understand finitely additive measure and countable additivity of subsets, besides knowing the definition of ring, σ-ring and their measure extension. One also has to differentiate between the terms premeasure, outer measure and measure from their domains and additivity conditions, which is clarified in the form of table. The advantage of measurable function and the induced measure is explained, as both share to define the probability as a measure, the probability that has values in the Borel set. As a summary of this section, a Borel σ-algebra shows a special role in real-life probability because numerical data, real numbers, is gathered whenever a random experiment is performed.

In this work a simple technique that is supported by pictorial presentation mostly is used, and for easiness most proofs are replaced by examples. I have freely borrowed a lot of material from various sources, and collected them in the manner that makes this thesis equipped with a little aspect of the real analysis, topology and probability theory.

INTRODUCTION

Probability space has its own vocabulary, which is inherited from probability theory. In modern real analysis instead of the word set, the word space used, the word space is used to title a set that has been endowed with a special structure. We therefore begin by presenting a brief list of real analysis, probability and topological space dialect.

Analysts' Term	*Probabilists' Term*	*Topological Term*
Normed Measure space $(X, \mathcal{M}, \mu)$	Probability space $(\Omega, \mathcal{F}, \mathbb{P})$ or called field of probability	Topological Space $(X, \mathcal{T})$
σ-algebras $\sigma(\mathcal{A})$ and Borel set $\mathcal{A}$: subset	σ-field $\sigma(E)$ and Borel set E: Event	$\sigma(\mathcal{T})$ and Borel σ-algebras The Borel subsets of $\mathbb{R}^n$ is the σ-algebra generated by the open sets of $\mathbb{R}^n$ (with a compact metric space)
Measurable function f	Measurable function (Random variable)	Measurable space
Points in space	Outcome, ω	Points in space
Convergence in measure	Convergence in probability	Convergence in ε-ball
the induced Measure μ_F (Lebesgue-Stieltjes measure)	the induced Measure μ_F (Distribution)	Support Measure Theory (Lebesgue-Stieltjes measure)
Continuous function is measurable	Measure induced by continuous random variable	Continuous function is measurable

Table 1 Interpretation of measure theory and probability theory concepts (adopted from: Gerald B. Folland, 1999).

PART ONE : MERTIC, NORMED VECTOR SPACE AND TOPOLOGICAL SPACES

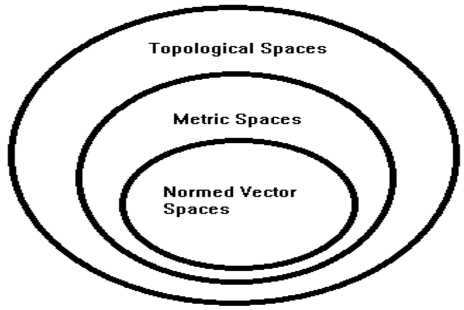


Figure 1 Metric space → topological space, but not all topological spaces can be metric

1. Metric Space (Distance Function) $(\boldsymbol{X}, \boldsymbol{d})$

A metric space as a pair (X, d), where X is a set and d is a function that assign a real number $d(x, y)$ to every pair $x, y \in X$ written as; $d: X \times X \longrightarrow \mathbb{R}$, and the Cartesian product $X \times X = \{ (x, y) | \, x \in X \quad and \;\; y \in X \}$.

The distance function is $d(x, y)$ and d map the x and y to positive real number [0,∞) such that:

1. $d(x, y) \geq 0$ for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$,
2. $d(x, y) = d(y, x)$ for all $x, y \in X$;
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Condition (3) is called the triangle inequality because (when applied to $\mathbb{R}^2$ with the usual metric) it says that the sum of two sides of the triangle is at least as big as the third side$(a + b \geq c)$.

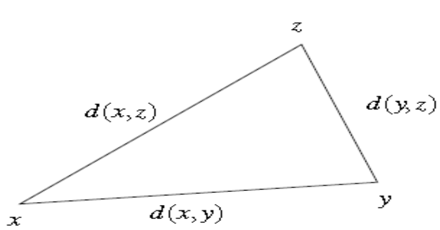


Figure 2 Diagram of triangle equality in metric space (source: Jiři Lebl, 2011)

The plane $\mathbb{R}^2$ with usual metric d_1 is obtained from the measure of the horizontal and vertical distance, and add the two together. $d_1\big((x_1, y_1), (x_2, y_2)\big) = |x_1 - x_2| + |y_1 - y_2|$.

The metric d_2 obtained from Pythagoras's theorem (called the Euclidean metric). For $n = 2$, it is defined by $d_2(x, y) = \sqrt{((x_1 - y_1)^2 + (x_2 - y_2)^2)}$, where $x = (x_1, x_2)$, $y = (y_1, y_2)$. For $n > 2$, $d_2(x, y) = \sqrt{((x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots (x_n - y_n)^2)}$, where $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_2)$.

The plane $\mathbb{R}^2$ with the supremum metric d_∞. $d_\infty\big((x_1, y_1), (x_2, y_2)\big) = max\,\{|x_1 - x_2|, |y_1 - y_2|\}$. We can represent it as $d_\infty: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$.

Note. Some reference use the notation $\ell_1, \ell_2, \ell_\infty$ instead of d_1 , d_2 and d_∞.

2. Norms and Normed Vector Spaces (V, ||·||)

The norm is best thought of as a length, defined for functions in space, where it must satisfy some properties. In linear algebra, we define a vector $\vec{x}$ as an ordered tuple of numbers $\vec{x}$ = $(x_1, x_2, \dots, x_n)$. If X is a vector space over $\mathbb{R}$, a norm is a function $\|\cdot\|: X \to [0, \infty)$ having the following properties:

N1 $\|x\| > 0$ when $x \neq 0$ and $\|x\| = 0$ if and only if $x = 0$.

N2 $\|\alpha x\| = |\alpha| \cdot \|x\|$ for all $x \in X$ and $\alpha \in \mathbb{R}$, α a scalar.

N3 $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

Note. A general vector norm $|x|$, sometimes written with a double bar $\|x\|$ as above, is a nonnegative norm.

The distance between two numbers x_1 and x_2 in $\mathbb{R}$ is the absolute value of their difference is $d(x_1, x_2) = |x_2 - x_1|$, in other words, a normed vector space is automatically a metric space.

But a metric space may have no algebraic (vector) structure (i.e., it may not be a vector space) so the concept of a metric space is a generalization of the concept of a normed vector space.

The triangle inequality can be shown as absolute value: (Satish Shirali and Harkrishan L. Vasudeva, 2011):

$$| x_1 | = \begin{cases} x_1 & x_1 \geq 0 \\ -x_1 & x_1 < 0 \end{cases}$$

$$|x_1 | = \sqrt{{x_1}^2}$$

$$| - x_1 | = | x_1 |$$

$$|x_1 + x_2| \leq |x_1| + |x_2|$$

or alternatively $$|x_1 - x_2| \geq |x_1| - |x_2|$$

$$| x_2 - x_1 | \leq | x_1 | + | x_2 |$$

2.1 Different Kinds of Norms

Types of vector norms are summarized in the following table, together with the value of the norm for example vector $v = (1,2,3)$:

Distance	Norm	Symbol	Value	Approx.
d_1	$L^1 - norm$	$\|x\|_1$	1+2+3=6	6.000
d_2	$L^2 - norm$	$\|x\|_2$	$(1^2 + 2^2 + 3^2)^{1/2} = \sqrt{14}$	3.742
d_3	$L^3 - norm$	$\|x\|_3$	$(1^3 + 2^3 + 3^3)^{1/3} = 6^{2/3}$	3.302
d_4	$L^4 - norm$	$\|x\|_4$	$(1^4 + 2^4 + 3^4)^{1/4} = 2^{1/4}.\sqrt{7}$	3.146
d_∞	$L^\infty - norm$	$\|x\|_\infty$	3	3.000

Table 2. Different kind of norms, (adopted from: Eric Weisstein, 2015).

2.2 Ball in ℓ_p

Let $X = \mathbb{R}^n$, the distance function between x and y

$$d_p(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|\right)^p$$

where $x = (x_1, x_2, x_3, \dots, x_n)$, $y = (y_1, y_2, y_3, \dots, x_n)$ and the sequence $x - y = (x_1 - y_1, x_2 - y_2, x_3 - y_3, \dots, x_n - y_n)$ are in $\mathbb{R}^n$ with $0 < p < 1$, is the metric space $(\mathbb{R}^n, d_p)$ denoted by ℓ_n^p.

The ℓ_n^p norm is a locally convex topological vector space. All the points lying at a given distance from a fixed point that is $\{p \in \mathbb{R}^2 | d(0,p) < 1\}$ in ℓ_n^p norm are convex.

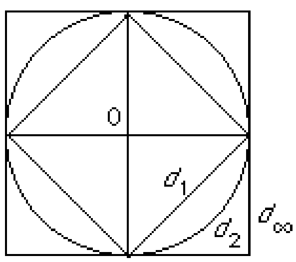


Figure 3 Ball in $\mathbb{R}^2$ with the $\ell_1, \ell_2, \ell_\infty$ norms. The unit ball (and so all balls) in ℓ_p norm are convex.

2.3 L^p Space (Function Spaces)

Definition. Let $(X, \mathcal{M}, \mu)$ be a three tuple space, the set $L_\mu^p(X)$ consists of equivalence classes of a functions $f: X \to \mathbb{R}$, where L^pnorm of $f \in L_\mu^p(X)$ is given by:

$$L_\mu^p(X) = \left\{ f: \int_X |f|^p d\mu < \infty \right\}$$

A function $f: X \to \mathbb{R}$ is measurable depends on the measure μ with integral is taken over a set X and $1 \le p < \infty$. The set $L_\mu^p(X) = L^p(X, d\mu)$ is a vector space with a norm.

Let X be the closed unit interval, so that the continuous functions on $[0,1]$ define the norm of X to be.

$$|f|_p = \left(\int_X |f|^p \, dx \right)^{1/p} = \|f\| = \int_0^1 |f(x)| dx$$

2.4 Distance Between Functions, L^p-distance

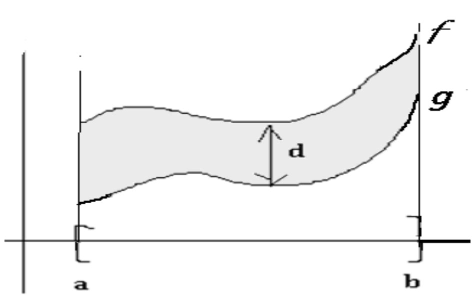
If $f, g \in L^p$, the L^p-distance between function space elements f and g is essentially the area between them, $\|f - g\|_p = (\int |f - g|^p)^{1/p}$. (adopted from: William Johnston, 2015).

The set of all continuous functions on $[a, b]$ can also be equipped with metric. For example, let X the collection of all continuous function from a to b, $f: [a, b] \to \mathbb{R}$. In the space of all continuous functions on $[a, b]$, the distance between f and g is defined to be;

$$d(f, g) = \int_a^b |f(x) - g(x)|$$

There are different ways to represent the distance between two functions in X such as:

$$d_1(f, g) = \ sup_{a \le x \le b} |f(x) - g(x)|$$



Distance between two function, L^p-distance.

3. Topological space

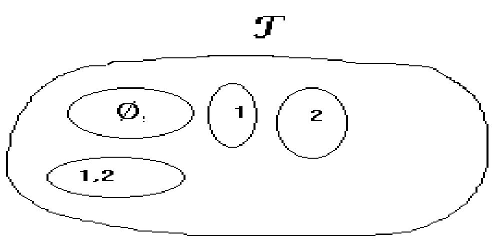
A Topology on a set X is a set of subsets, called open subset, including X and the empty set, such that any union of open sets is open, and any finite intersection of open sets is open.

In other words, A collection of subsets $\mathcal{T}$ of X is called a topology on X, if:

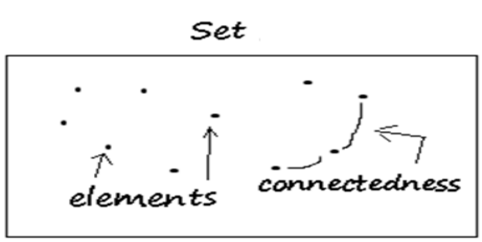
1. $X \in \mathcal{T}, \emptyset \in \mathcal{T}$,
2. $A \in \mathcal{T}$ and $B \in \mathcal{T}$, then their intersection $\in \mathcal{T}$,
3. The union of sets in any sub-collection of $\mathcal{T}$ is an element of $\mathcal{T}$ (i.e $S \subset \mathcal{T}$ is any collection of sets in $\mathcal{T}$ then $\bigcup_{A \in S} A \in \mathcal{T}$).

In a topological space, every open set open in $\mathcal{T}$ can be written as the union of countable open sets.

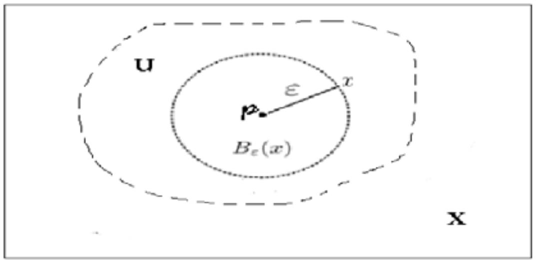
Example. The ordered pair $(X, \mathcal{T})$ is a topological space, with $X = \{1,2\}$ and $\mathcal{T} = \{\emptyset, \{1\},\{2\},\{1,2\}\}$.



A topological space X is called path connected if every two points in X are connected by a path, that is, for all $x, y \in X$ there is a continuous function $f: [0,1] \to X$ such that $f(0) = x$, and $f(1) = y$. (John B. Etnyre, 2014).



If p is an element of a metric space (X, d), given $x \in X$ and $\varepsilon > 0$, the collection of points x in X within distance ε from p is the ε-ball, define as $B_\varepsilon(x) = \{x \in X | d(x, p) < \varepsilon\}$ or $B_\varepsilon(x) = \{x \in X | |p - x| < \varepsilon\}$.



Open Ball, ε-ball

The open balls of metric space form a basis for a topological space, called the topology induced by the metric d.

A collection of subsets $\{U_a\}_{a \in J}$ of X is a cover of X if $X = \bigcup_{a \in J} U_a$ (here J is an indexing set, so for example if $J = \{1, 2, \dots n\}$ then $\{U_a\}_{a \in J} = \{U_1, U_2, \dots U_n\}$). (John B. Etnyre, 2014),

A topology space X is compact if every cover of X by open sets has a finite subcover, in other words if $\{U_a\}_{a \in J}$ is a collection of open sets in X that cover X then there is a subset $I \subset J$ such that I is finite and $X = \bigcup_{a \in I} U_a$, which means that every open cover has a finite refinement.

3.1 Continuous Function in Metric and Topology

An element x of a topological space is said to be the limit of a sequence x_n of other points when, for any given integer N, there is a neighborhood of x which contains every x_n for $n \geq N$. A sequence is said to be convergent if it has a limit. More clearly we say that the function $f: \mathbb{R}^n \to \mathbb{R}^m$ is continuous at the point $x_1 \in \mathbb{R}^n$ if $\forall \varepsilon > 0, \exists\, \delta > 0$ such that, $\|x - x_1\| < \delta \Longrightarrow |f(x) - f(x_1)| < \varepsilon$.

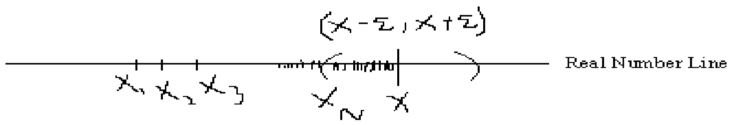


Figure 4 The convergence of the sequence (source Ben Garside, 2014)

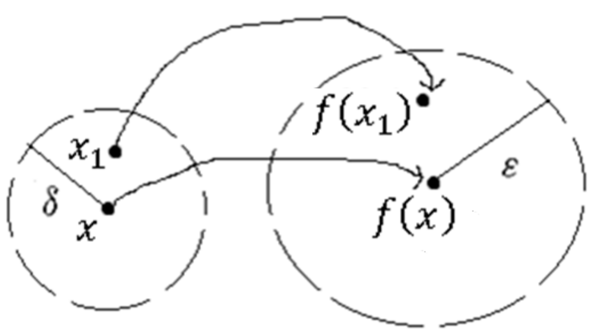
To generalize to metric space we replace the difference vector norms $\|x - x_1\|$ with distance metric $d(x, x_1)$ with letting $f(x_1)$ to be an arbitrarily small open ball around $f(x)$ and by keeping x_1 to be a sufficiently small open ball around x.(adopted from: Sadeep Jayasumana, 2012).

Definition: The (X, d_X) and $\left(X_1, d_{X_1}\right)$ are arbitrary metric spaces. A function $f\colon X \to X_1$ is continuous at a point $x \in X$ if, for any $\varepsilon > 0$, there is an $\delta > 0$ such that for each $x_1 \in X_1$: $d_X(x, x_1) < \delta$ implies $d_{X_1}\big(f(x), f(x_1)\big) < \varepsilon$.

In other words, f is continuous at x if, for any $\varepsilon > 0$, there exist a $\delta > 0$ such that

$$f(B_X(x|\delta)) \subset B_{X_1}(f(x)|\varepsilon)$$

If f is continuous at every point of X, we simply say that f in continuous on X.



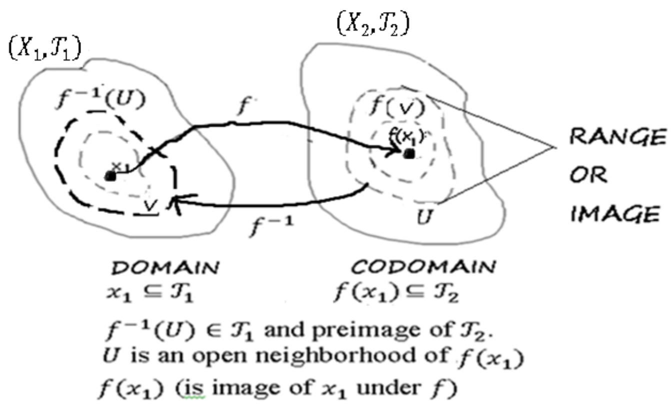
The planes $X = \mathbb{R}^2$ and $X_1 = \mathbb{R}^2$ (adopted from: Jianfei Shen,2012),

Using topological terms we define: a function $f\colon (X_1, \mathcal{T}_1) \to (X_2, \mathcal{T}_2)$ between two topological spaces is said to be continuous if whenever $U \in \mathcal{T}_2$, then $f^{-1}(U) \in \mathcal{T}_1$ is the preimage of $\mathcal{T}_2$.

More easily, we say that a function between two topological spaces is continuous if preimages of open sets are open, that is for any open neighborhood U of $f(x_1)$, there exists an open neighborhood V of x_1 such that $f(V) \subset U$, as in the following theorem.

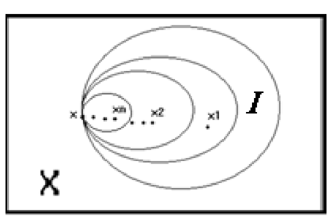
Theorem. given $f: X_1 \to X_2$, the following are equivalent : (Jianfei Shen,2012)

- f is continuous on X_1 by the ε- δ definition
- For every $x_1 \in X_1$, if $x_n \to x_1$ in X, then $f(x_n) \to f(x_1)$ in X_2
- If U is closed in X_2, then $f^{-1}(U)$ is closed in X_1
- If U is open in X_2, then $f^{-1}(U)$ is open in X_1



3.2 Limit Point

A point x of a topological space is said to be a limit point of a subset I if every neighborhood of x intersects I in some point other than x.



That is, in a metric space (X, d), if $I = \{x_1, x_2 \dots x_n \dots\}$, and $\forall \varepsilon > 0$, $x_i \in B_\varepsilon(x)$ for some $x_i \in I$ then x is a limit point of I.

PART TWO : OUTLINES OF MEASURE AND PROBABILITY SPACES

1. Measure and Measure Space

A measure on a set X, is a systematic way to assign a number to each suitable subset of that set, intuitively interpreted as its size. In this sense, it generalizes the concepts of length, area, and volume.

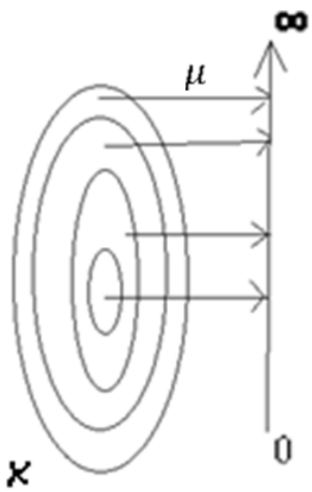
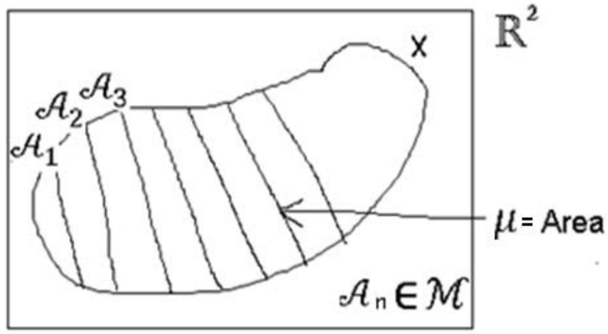


Figure 5 Measure on a set, S (source: Rauli Susmel, 2015)

The measure of $\mathcal{A}$ denoted as $\mu(\mathcal{A})$ is a measure μ on set X associates to the subset $\mathcal{A}$ of X. Measurable set is set that have a well-defined measure, $(X, \mathcal{M})$ called a measurable space and the members of $\mathcal{M}$ are the measurable sets.

A measure space $(X, \mathcal{M}, \mu)$ consisting of a set X, class of subsets $\mathcal{A}\ of\ X$ represented by $\sigma(\mathcal{A})$ which is equal to $\mathcal{M}$, and μ attaching a nonnegative number to each set $\mathcal{A}$ in $\mathcal{M}$. The measure μ is required to have properties that facilitate calculations involving limits along sequences.



1.1 Construction of Sigma Algebra $\sigma(\mathcal{A})$

Construction of sigma algebra for a set X is done by identifying a subsets $\mathcal{A}'s$ of the elements we wish to measure, and then defining the associating sigma algebra by generating the sigma algebra on $\mathcal{A}$, $\sigma(\mathcal{A})$.

There are three requirements in generating $\sigma(\mathcal{A}) = \mathcal{M}$:

1. The set X is an element in $\mathcal{M}$,

2. The complement of any element in $\mathcal{M}$ will also be in $\mathcal{M}$,

3. For any two elements in $\mathcal{M}$ their union also in $\mathcal{M}$.

For example, let set X = {A,B,C,D,E,F,G,H}, and let subset of X be $\mathcal{A}_1 = \{B, C, D\}$ and $\mathcal{A}_2 = \{F, G, H\}$. Then the algebra $\mathcal{M}$ that generated by $\mathcal{A}$ denoted by $\sigma(\{\mathcal{A}_1, \mathcal{A}_2\})$ or simply by $\sigma(\mathcal{A})$.

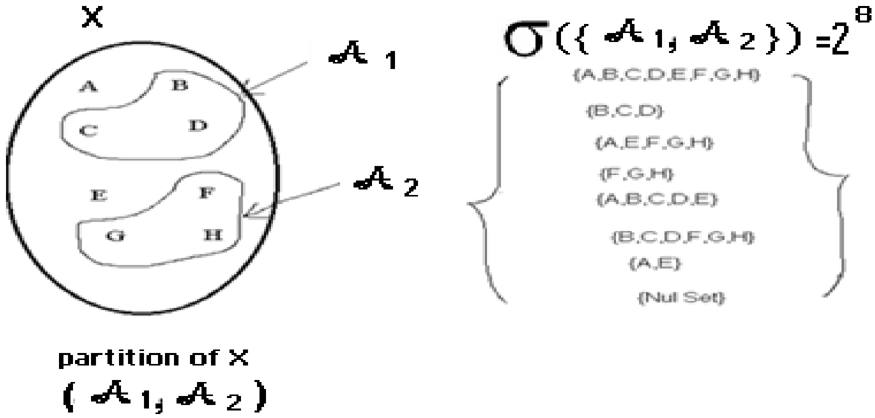


Figure 6 Constructing sigma algebra, (adopted from: Viji Diane Kannan, 2015).

In other words, measurable set contain the subset $\mathcal{A}_1$ and $\mathcal{A}_2$, make a class of measurable sets $\mathcal{M}$ is;

$\sigma(\mathcal{A}) =$

$\{(A, B, C, D, E, F, G, H), (B, C, D), (A, E, F, G, H), (F, G, H), (A, B, C, D, E), (B, C, D, F, G, H), (A, E), (\emptyset)\}$

Typical notation	Measure Terminology
X	Collection of objects - Set
$A, B, C, \dots$	a member of X
$\mathcal{A}_1, \mathcal{A}_2, \dots \mathcal{A}_n$	subset of X
σ-algebra of $\mathcal{A}$	measurable sets $\mathcal{M}$ or Borel sets $\mathcal{M}$

Table 3. Terminology of measure theory

1.2 The Inverse Function is Measurable Function

A function between measurable spaces is said to be measurable if the preimage of each measurable set is measurable, and this is equivalent of the inverse and measurable function with the continuous function is shown also between topological spaces.

Suppose $f: X \longrightarrow Y$ and let $a \subseteq X$, and $A \subseteq Y$. The image "a" and the inverse image of "A" are:

$$f(a) = \{y | y = f(x) \text{ for some } x \in a\}$$

$$f^{-1}(A) = \{x | f(x) \in A\}$$

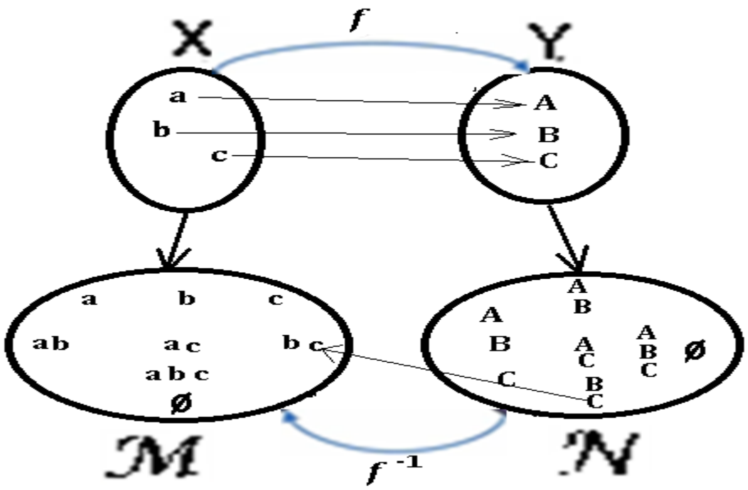


Figure 7 mapping $f: X \longrightarrow Y$ (adopted from: Viji Diane Kannan, 2015)

Definition. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be two measurable spaces. The function $f: X \longrightarrow Y$ is said to be $(\mathcal{M}, \mathcal{N})$ measurable if $f^{-1}(\mathcal{N}) \subseteq \mathcal{M}$. Say that $f: (X, \mathcal{M}) \longrightarrow (Y, \mathcal{N})$ is measurable if $f: X \longrightarrow Y$ is an $(\mathcal{M}, \mathcal{N})$ measurable function.

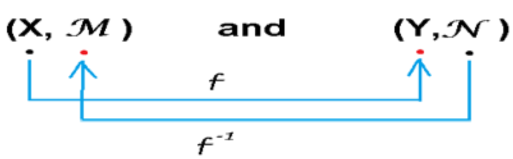


Figure 8 f is a measurable function because f^{-1} exist, (source: Viji Diane Kannan, 2015).

If $(a_i)_{i \in I}$ is a collection of subsets of X and $(A_i)_{i \in I}$ is a collection of subsets of Y then

$$f\left(\bigcup_{i \in I} a_i\right) = \bigcup_{i \in I} f(a_i)$$

$$f^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f^{-1}(A_i)$$

2. Probability Space

Probability is mapping in a measurement system where the domain of Ω is a field of outcomes and the domain of $\mathcal{F}$ is the unit interval [0,1] (see figure 9). The measurement triple is $(\Omega, \mathcal{F}, \mathbb{P})$ called the probability field or space. A probability space is needed for each experiment or collection of experiments that we wish to describe mathematically.

The ingredients of a probability space are:

- Ω is a set (sometimes called a sample space in elementary probability). Elements ω of Ω are sometimes called outcomes. When Ω is a finite set, we have the probability $(\omega) = \frac{1}{|\Omega|}$, where $|\Omega|$ is the number of points Ω.
- $\mathcal{F}$ is a σ-algebra (or σ -field, we will use these terms synonymously) of subsets of Ω.
- $\mathbb{P}$, probability measure, is a function from $\mathcal{F}$ to [0,1] with total probability $\mathbb{P}(\Omega) = 1$ and such that if $E_1, E_2, \ldots \in \mathcal{F}$ are pairwise disjoint events, meaning that $E_i \cap E_j = \emptyset$ whenever $i \neq j$, then

$$\mathbb{P}\left[\bigcup_{j=1}^{\infty} E_j\right] = \sum_{j=1}^{\infty} \mathbb{P}[E_j]$$

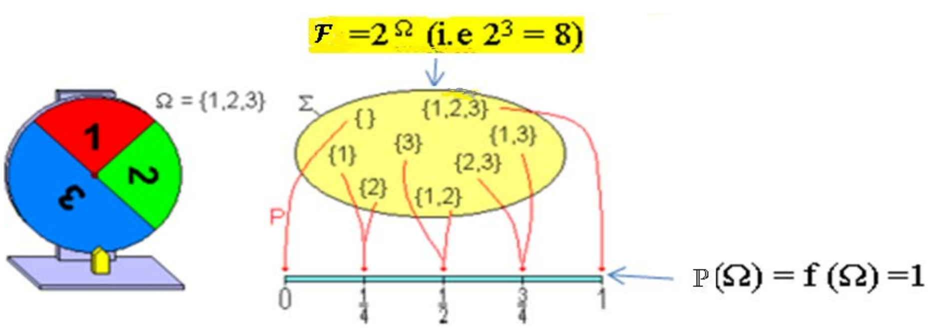


Figure 9 A probability measure mapping the probability space for 3 events to the unit interval (source: Probability measure, Wikipedia, cited 2014).

Typical notation	Set Terminology	Probability jargon
Ω	Collection of objects	Sample Space
ω	a member of Ω	Outcome
E	subset of Ω	Event (that some outcome in E occurs)
$\mathcal{F}$	σ-algebra of E	Measurable sets or Borel sets
$\mathbb{P}$		Probability Measure

Table 4. Terminology of set theory and probability theory (adopted from: Manuel Cabral Morais, 2009-2010)

2.1 Sample Space Ω

The set of all possible outcomes is called the sample space, If we flip a coin a single time and are interested just in which side turns up heads or tails, then we can take the sample space to be, symbolically, the set {H, T}. Ω represents the sample space and ω represent some generic outcome; hence, Ω = {H, T} for this experiment, and we would write ω = H to indicate that the coin came up heads.(T W Epps, 2013).

2.2 The σ-field and Class of Events $\mathcal{F}$

The field of events is the collection of events or the subset of Ω whose occurrence is of interest to us and about which we have the information. σ-field is a nonempty class of subsets of Ω closed under the formation of complements and countable unions.(RJS, 2012).

Here are some examples:

Event, E_1 = { the outcome of the die is even } = {2,4,6}

Event, E_2 = { exactly two tosses come out tails} = {(0,1,1),(1,0,1),(1,1,0)}

Event, E_3 = { the second toss comes out tails and the fifth heads}

$$=\{\omega = (x_i)_{i=1}^{\infty}\} \in \{0,1\}^{\mathbb{N}} \mid x_2 = 1 \; and \; x_5 = 0\}$$

In discrete sample space the class $\mathcal{F}$ of events contains all subsets of the sample space, in complicated sample spaces the case is different, since in general theory $\mathcal{F}$ is a σ-field in sample space. $\mathcal{F}$ follow the following axioms:

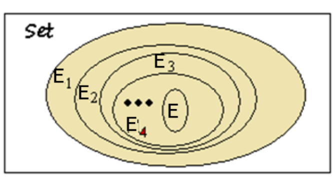
- $\Omega \in \mathcal{F}$ and $\emptyset \in \mathcal{F}$ (the sample space Ω and the empty set $\emptyset$ are events)
- if $E \in \mathcal{F}$ then $E^c \in \mathcal{F}$. (E^c) is the complement of E, this is the set of points of Ω that do not belong to E,
- if $E_1, E_2, E_3 \dots \in \mathcal{F}$ then also $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$, in other words the union of sequence of events is also an event.(Benedek Valkó, 2015)

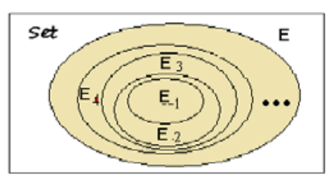
2.3 Sequences of Events in Probability

We deal here with sequences of events and various types of limits of such sequences. The limits are also events. We start with two simple definitions (Kyle Siegrist, 2015).

Suppose that $(E_1, E_2, \dots)$ is a sequence of events.

1. The sequence is increasing $\{E_n\} \uparrow$ if $E_n \subseteq E_{n+1}$ for every $n \in \mathbb{N}_+$, and we can define the limit, $\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n$.
2. The sequence is decreasing $\{E_n\} \downarrow$ if $E_{n+1} \subseteq E_n$ for every $n \in \mathbb{N}_+$, and we can define the limit, $\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n$.





A sequence of decreasing events and their intersection A sequence of increasing events and their union

(source: Kyle Siegrist,2015).

2.4 Continuity Theorems of Probability

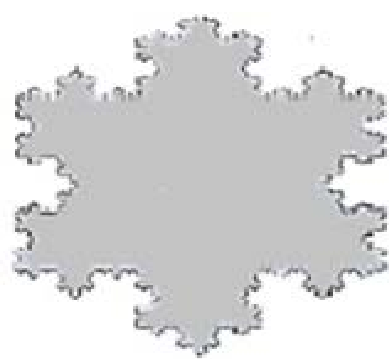
A function is continuous if it preserves limits. Thus, the following results are the continuity theorems of probability, suppose that $(E_1, E_2, \ldots)$ is a sequence of events:

- Continuity theorem for increasing events: If the sequence is increasing then $\lim_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(\lim_{n\to\infty} E_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} E_n\right)$ and,
- Continuity theorem for decreasing events: If the sequence is decreasing then $\lim_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(\lim_{n\to\infty} E_n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} E_n\right)$

3. General Idea of Borel Set

A Borel Set in a topological space is a set which can be formed from open sets through operations of countable union, countable intersection and complement. The set of Borel sets which are intersections of countable collections of open sets is called G_δ and the set of countable unions of closed sets is called F_σ. We can consider set of type $F_{\sigma\delta}$, which are the intersections of countable collections of sets each which is an F_σ. Similarly we construct classes $G_{\delta\sigma}$, $F_{\sigma\delta\sigma}$, etc. (H.L. Royden,1988).

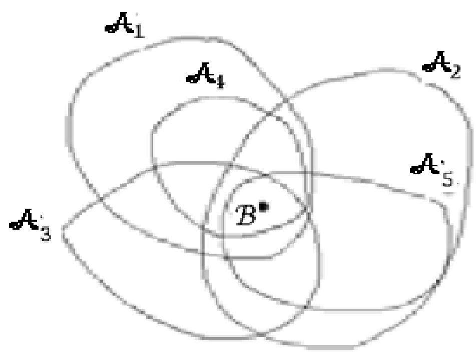
Then we can take countable unions and countable intersection of those new sets, and get more new sets, all of which will be Borel, too. This process will never stop. (Dr. Nikolai Chernov, 2012-1014)



Note. F_σ : F refer to closed sets and the subscript σ is addition. G_δ: G refer to open sets and the subscript is δ intersection.

The following properties of Borel sets are clear.

1. If $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5, \ldots \subset \mathcal{B}$ are open, then $\bigcap_{i=1}^{\infty} \mathcal{A}_i$ is a Borel set;
2. All closed sets are Borel.



In the above graph $\mathcal{B}$ is contained in a finite collection of $\mathcal{A}_i$.(adopted Tai-Danae Bradley, 2015)

If a is a real number, $\{a\}$ can be written as the countable intersection $\{a\} = \bigcap_{n=1}^{\infty}[a, a+\frac{1}{n})$, so it's a Borel set. Further monotonicity tells us that $\mu(\{a\}) = \lim_{n\to\infty} \mu\left([a, a+\frac{1}{n})\right) = \lim_{n\to\infty} \frac{1}{n} = 0$ and so the singleton {a} has measure zero.

Example. σ-algberas are closed under countable union, it follows that the interval $(a, b) \in \mathcal{B}$, where $(a, b) \in \mathbb{R}$.

$$(a, b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n}\right]$$

The Borel σ-algebra on $\mathbb{R}$ called $\mathcal{B}(\mathbb{R})$ is generated by any of the following collection of intervals

$$\{(-\infty, b) : b \in \mathbb{R}\}, \quad \{(-\infty, b] : b \in \mathbb{R}\}, \quad \{(a, \infty) : a \in \mathbb{R}\}, \quad \{[a, \infty) : a \in \mathbb{R}\}$$

Proposition1: Let $X = \mathbb{R}$ and let us consider the following collections

$\mathcal{A}_1 = \{(a,b): a \leq b\}$, $\mathcal{A}_2 = \{(a,b]: a \leq b\}$, $\mathcal{A}_3 = \{[a,b): a \leq b\}$, $\mathcal{A}_4 = \{[a,b]: a \leq b\}$

Then $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{A}_1)$ = $\{(a,b): a \leq b\}$ and $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{A}_2)$ = $\{(a,b]: a \leq b\}$ and so on.

Proposition2: Let $n \in \mathbb{N}$ and $X = \mathbb{R}^n$ and consider $\mathcal{A} = \{(a_1,b_1] \times (a_2,b_2] \times \dots (a_n,b_n]: a_i \leq b_i, i = 1,\dots,n\}$ then $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{A})$.

The Borel sets of $\mathbb{R}^n$ is defined by replacing intervals with rectangles as

$$[x = (x_1, \dots x_n): a_i \leq x_i \leq b_i, i = 1, \dots k]$$

and sometimes called hyperrectangles in the definition of $\sigma(\mathcal{A})$. (RJS 2012).

3.1 Complicated Borel Sets (Cantor set)

Definition. If we take the interval [0,1] as a Borel set, and we remove the open interval $\left(\frac{1}{3}, \frac{2}{3}\right)$, then we have $\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$ and continue this way by removing middle third and taking the intersection of countably many Borel sets, we end with a Borel set that called Cantor set (H. L. Royden and P.M. Fitzpatrick, 2010)

To construct the Cantor set of the interval Ω = [0,1], the subset we remove is the disjoint union

$$E_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$$

$$E_2 = \left(\frac{1}{9}, \frac{2}{9}\right) \bigcup \left(\frac{7}{9}, \frac{8}{9}\right)$$

$$E_3 = \left(\frac{1}{27}, \frac{2}{27}\right) \bigcup \left(\frac{7}{27}, \frac{8}{27}\right) \bigcup \left(\frac{19}{27}, \frac{20}{27}\right) \bigcup \left(\frac{25}{27}, \frac{26}{27}\right)$$

$$E_4 = \left(\frac{1}{81}, \frac{2}{81}\right) \bigcup \left(\frac{7}{81}, \frac{8}{81}\right) \bigcup \left(\frac{19}{81}, \frac{20}{81}\right) \bigcup \left(\frac{25}{81}, \frac{26}{81}\right) \bigcup \left(\frac{55}{81}, \frac{56}{81}\right) \bigcup \left(\frac{61}{81}, \frac{62}{81}\right) \bigcup \left(\frac{73}{81}, \frac{74}{81}\right) \bigcup \left(\frac{79}{81}, \frac{80}{81}\right)$$

$$\vdots$$

$$\lim_{n \to \infty} \mu \left(\bigcup_{i=1}^{n} E_n \right) = \frac{1}{3} = \left(\frac{1}{9} + \frac{1}{9}\right) + \left(\frac{1}{27} + \frac{1}{27} + \frac{1}{27} + \frac{1}{27}\right) \dots = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} \dots = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = 1$$

The total length removed is geometric progression

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = 1 \ or \ \sum_{n=0}^{\infty} \frac{2^n}{3^{n+1}} = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} \ldots = \frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) = 1$$

The subset we remove from [0, 1] which is of measure 1, is the disjoint union $\bigcup_1^{\infty} E_n$.

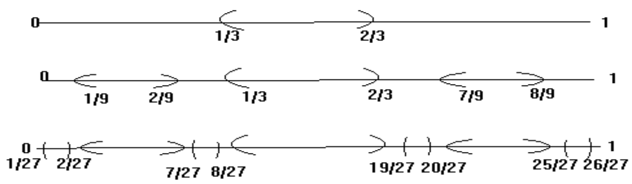


Figure 10 Borel set (adopted from H. L. Royden and P.M. Fitzpatrick,2010)

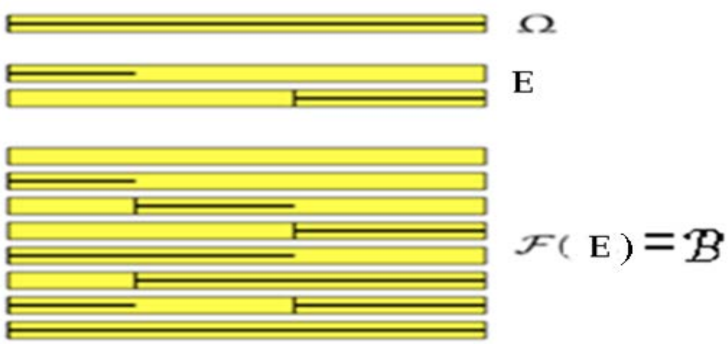


Figure 11 Borel sets, E is a collection of subsets of Ω . The minimal field containing E, denoted $\mathcal{F}(E)$, is the smallest field containing E. (source Albert Tarantola, 2007).

3.2 Borel σ-algebra in Topology

Defintion: Let $(X, \mathcal{T})$ be a topological space. The σ-algebra $\mathcal{B}(X) = \mathcal{B}(X, \mathcal{T}) = \sigma(\mathcal{T})$ that is generated by the open sets is called the Borel σ-algebra on X. The element $\mathcal{A} \in \mathcal{B}(X, \mathcal{T})$ are called Borel measurable sets. (Paolo Dai Pra, cited 2016).

Borel Function. Let $X_1, \mathcal{T}_1$ and $X_2, \mathcal{T}_2$ be two topological spaces, a function $f: X_1 \to X_2$ is said to be Borel function iff for any open set $U \subset \mathcal{T}$ its preimage $f^{-1}(U) \subset X$ is a Borel set.

Borel Measure. If μ measure is defined on Borel σ-algebra of all open sets in $\mathbb{R}^n$, then μ is called Borel measure.

3.3 Borel Set of Probability Space

Consider the following probability space $([0,1], \mathcal{B}([0,1]), \mu)$, the sample space is the real interval [0,1]. A measure is a function $\mathbb{P}: \mathcal{F} \to [0,1]$ which assigns a nonnegative extended real number $\mathbb{P}(E)$ to every subset E in $\mathcal{F}$, which means we extended the probability measure $\mathbb{P}$ from E to the entire σ-field $\mathcal{F}$, that is $\sigma(E) = \mathcal{B}$

$\mathcal{F}$ consist of the empty set and all sets that are finite unions of intervals of the form (a,b]. In more detail, if a set $E \in \mathcal{F}$ is nonempty, it is of the form

$$E = (a_1, b_1] \cup \dots \cup (a_n, b_n],$$

$$where \quad 0 \leq a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_n < b_n \leq 1, n \in \mathbb{N}$$

Then

$$\mathbb{P}\left\{\bigcup_{i=1}^{n} |a_i, b_i|\right\}$$

Example. Consider the Borel σ-algebra on [0,1], denoted by $\mathcal{B}$ [0,1]. This is the minimal σ-algebra generated by the elementary event $\{[0,b), 0 \leq b \leq 1\}$. This collection contain things like:

$$\left[\frac{1}{2}, \frac{2}{3}\right], \left[0, \frac{1}{2}\right] \bigcup \left(\frac{2}{3}, 1\right], \left\{\frac{2}{3}, 1\right\}, \left[\left\{\frac{1}{2}\right\}, \left\{\frac{2}{3}\right\}\right]$$

4. Measurable Space in Probability

Given two measurable spaces (Ω_1,$\mathcal{F}_1$) and (Ω_2,$\mathcal{F}_2$), a mapping or a transformation from $\mathbb{P}: \Omega_1 \to \Omega_2$ i.e. a function $\omega_2 = \mathbb{P}(\omega_1)$ that assigns for each point $\omega_1 \in \Omega_1$ a point $\omega_2 = \mathbb{P}(\omega_1) \in \Omega_2$ is said to be measurable if for every measurable set $E \in \mathcal{F}_2$, there is an inverse image

$$\mathbb{P}^{-1}(E) = \{\omega_1 : \mathbb{P}(\omega_1) \in E\} \in \mathcal{F}_1.$$

The function $f: \Omega \to \mathbb{R}$ defined on $(\Omega, \mathcal{F}, \mathbb{R})$ is called $\mathcal{F}$-measurable if $f^{-1}(E) = \{\omega: f(\omega_1) \in E\} \in \mathcal{F}$ for all $E \in \mathcal{B}(\mathbb{R})$. i.e., the inverse f^{-1} maps all of the Borel sets $E \subset \mathbb{R}$ to $\mathcal{F}$. The function in this context is a random variable.

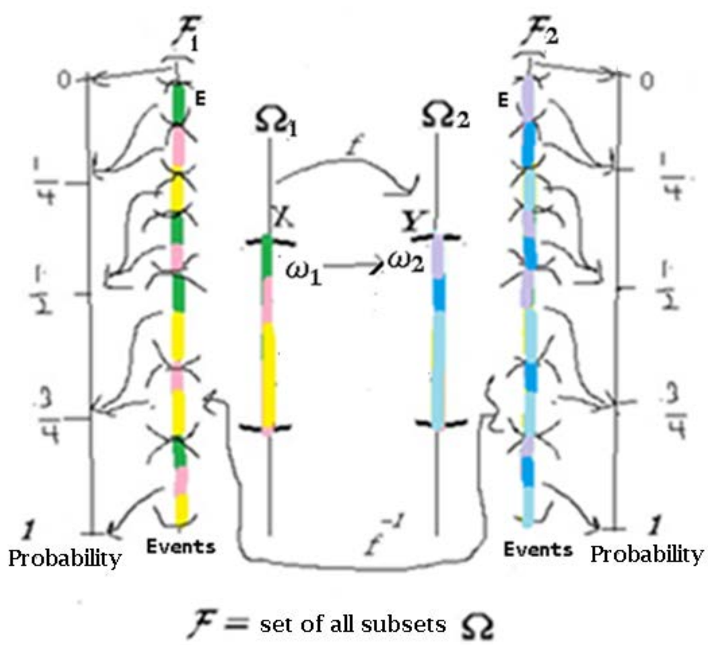


Figure 12 Two measurable spaces

5. Random Variable (R.V)

Definition. A random variable is numerical function of the outcome of random experiment, can be represented as $R: \omega \ \to \mathrm{R}(\omega)$.

Knowing that the random variables by definition are measurable functions and can be limited to a subset of measurable functions through the measurable mapping $[R: \omega \ \to \mathrm{R}(\omega)] \in \mathcal{F}$.

The random variable is a simple random variable if it has finite range (assumes only finitely many values) and if $[R: \omega \ \to \mathrm{R}(\omega) = \mathbb{R}] \in \mathcal{F}$ for each R. (adopted from Patrick Billingsley, 1995).

The function or the transformation R, has the sample space (domain) of the random experiment and the range is the real number line, where $\omega_1, \omega_2, \omega_3$ subset Ω and the function R assigns a real number $\mathrm{R}(\omega_1)$, $\mathrm{R}(\omega_2)$ and $\mathrm{R}(\omega_3)$.

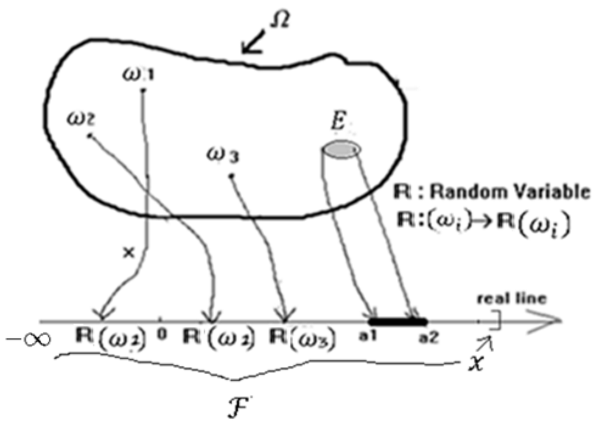


Figure 13 The transformation of some sample points $\omega_1, \omega_2, \omega_3$ and real line segment $[a_1, a_2]$ of event E on the real number line. (adopted from:Prateek Omer, 2013).

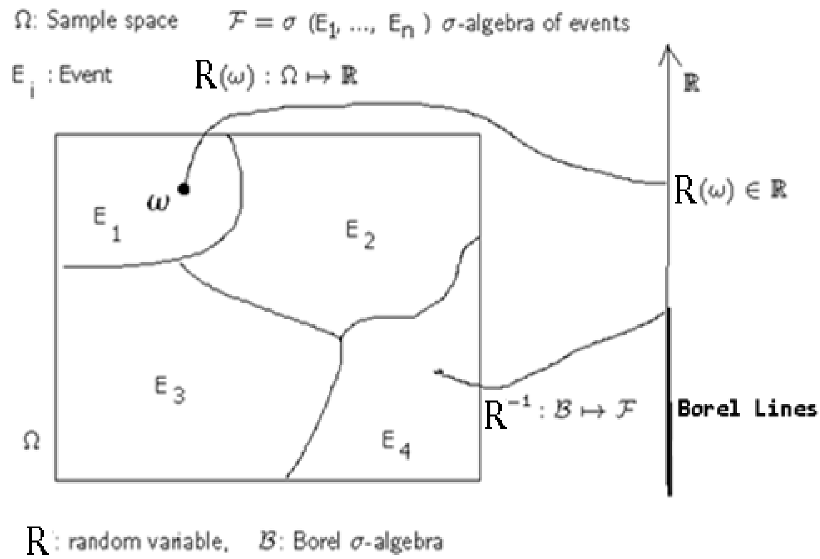
If the Outcomes are grouped into sets of outcomes called event, then E is a σ-algebra on Ω generated by R.

5.1 Measurable Function is a Random Variable

Let R be a function from a measurable space to the Borel line, R is measurable if the inverse image of every measurable space set in the Borel line is a measurable set in the measurable space.

$$\text{Random Variable (Measurable Function)}$$
$$\equiv \mathrm{R}^{-1}\text{that map(set in the Borel line) into the measurable space } \mathcal{F}$$
$$\equiv \mathrm{R}^{-1}(E): \mathcal{B} \to \mathcal{F}$$

In the probability the measurable functions are random variables. They are functions whose variable represent outcomes ω of an experiment that can happen to be random.(Sunny Auyang, 1995)



Measurable functions. (source**:** Florian Herzong 2013).

PART THREE: MEASURE , PROBABILITY THEORY AND DISTRIBUTION FUNCTION

1. Lebesgue's Idea of Integral (subdivide the range)

Lebesgue began with what he stated a problem of integration, which is considered the modern method of integration that has a relation with interval, the basic measure idea.

Lebesgue's idea is to assign to each bounded function f defined on a bounded interval $\mathbb{I} = [a, b]$ a number, called it integral and denoted by $\int_a^b f(x)dx = \int_a^b f = \int_{\mathbb{I}} f$.(Charles Swartz. 1994)

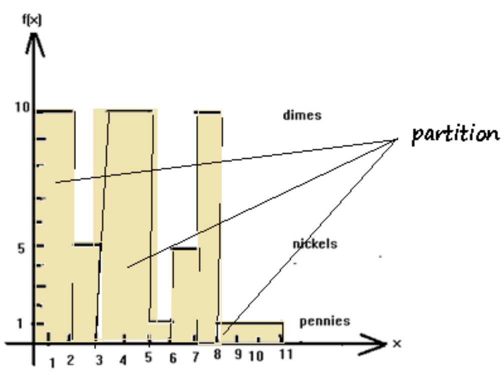
Example. (adopted from Bruckner & Thomson, 2008). Suppose that we have coins: 5 dimes 2 nickels and 4 pennies, where a dime is worth 10 cents, a nickel is worth 5 cents, a penny is worth 1 cent.

First: we take an amount have computation (1) as:

10+10+10+10+10+5+5+1+1+1+1=64

The sum of the partition between the interval [0,11] gives the area by Riemann integration as:

$\int_0^{11} f(x)dx = 64$



Second: we take an amount have computation (2) as:

5(10)+2(5)+4(1)=64

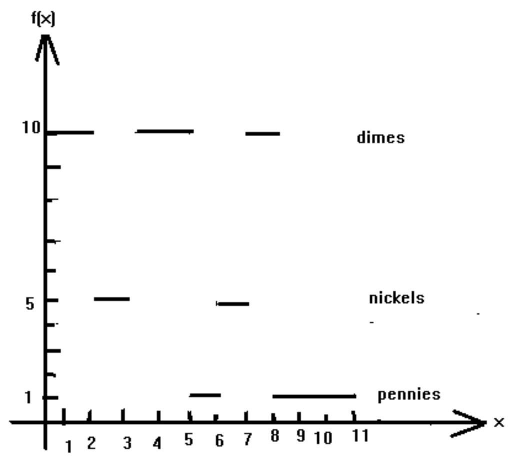


Figure 14. Example about Lebesgue integral, (source: Bruckner & Thomson, 2008)

Looking to the computation (2) (the Lebesgue idea); let $|\mathbb{I}|$ be the length of interval , if $\mathcal{A}$ is a finite union of pairwise-disjoint intervals, $\mathcal{A} = |\mathbb{I}_1| \cup \ldots |\mathbb{I}_n|$, and the measure of $\mathcal{A}$ is $\mu(\mathcal{A}) = |\mathbb{I}_1| + \cdots + |\mathbb{I}_n|$.

Going back to the last example; let $\mathcal{A}_1 = \{x : f(x) = 1\}$, $\mathcal{A}_5 = \{x : f(x) = 5\}$, and $\mathcal{A}_{10} = \{x : f(x) = 10\}$, then $\mu(\mathcal{A}_1) = 4, \mu(\mathcal{A}_5) = 2$ and $\mu(\mathcal{A}_{10}) = 5$

The sum $(1)\mu(\mathcal{A}_1) + (5)\mu(\mathcal{A}_5) + (10)\mu(\mathcal{A}_{10})$, given the number 1,5, and 10, is the value of the function f and $\mu(\mathcal{A}_i)$ is how often the value i is taken on.

Assume that the function $f(x)$ has range:

$$m \le f(x) < M \; for \; all \; x \in [a, b].$$

Instead of partition the interval $[a, b]$ we partition $[m, M]$ and have

$$m = y_0 < y_1 < \cdots < y_n = M.$$

For $i = 1, 2, \ldots, n$,

$\mathcal{A}_i = \{x : y_{i-1} \le f(x) < y_i\}$, thus the partition of the range induces a partition of the interval$[a, b]$:

$$[a, b] = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \ldots \cup \mathcal{A}_n$$

Where the sets $\{\mathcal{A}_i\}$ are clearly pairwise disjoint, and we can form the sums $\sum y_i \, \mu(\mathcal{A}_i)$ above and $\sum y_{i-1} \, \mu(\mathcal{A}_i)$ below. These two sums approximate the integration, the approximation of the integration reaches a limit as the norm of the partition reaches zero (squeeze theorem). The idea of partitioning the range is a Lebesgue's contribution.

2. Additive Set Function and Countable Additivity

Set function

A set function $\mu: \mathcal{M} \to [0, \infty]$, is a function whose domain is a class of sets $\mathcal{M}$. An extended real value set function is a measure μ if its domain is a semi ring.

Usually the input of the set function is a set and the output is a number. Often the input is a set of real numbers, a set of points in Euclidean space, or a set of points in some measure space.

The Lebesgue measure is a set function that assigns a non-negative real number to each set of real numbers. Let μ be a set function, we have the following:

1. μ is **increasing** if, for all $A, B \in \mathcal{A}$ with $A \subseteq B$,

$$\mu(A) \leq \mu(B)$$

2. μ is **additive** if, for all disjoint sets $A, B \in \mathcal{A}$ with $A \cup B \in \mathcal{A}$,

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

3. μ is (**σ- additive) countable additive** if, $A \in \mathcal{A}$ and $A = \bigcup_{n=1}^{N} A_n$ where $A_1, \dots A_n \in \mathcal{A}$ are pairwise disjoint sets (sequence of disjoint sets), then $\mu(\bigcup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mu(A_n)$.

4. μ is **countable subadditive** if, for all sequences $(A_n : n \in \mathbb{N})$ in $\mathcal{A}$ then $\mu(\bigcup_n A_n) \leq \sum_n \mu(A_n)$. (J.R.Norris, cited 2016).

Example. The length is a σ-additive measure on the family of all bounded intervals in$\mathbb{R}$.

$$|\mathbb{I}| = \sum_{k=1}^{\infty} |\mathbb{I}_k|$$

Example. Probability is an additive function. If the event E is a disjoint union of a finite sequence of event $E_1, \dots, E_n$, then

$$\mathbb{P}(E) = \sum_{k=1}^{\infty} \mathbb{P}(E_k)$$

3. Construction of Measure on $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$

Definition. **An interval** $\mathbb{I} \subset \mathbb{R}$. Its length is $|\mathbb{I}| = b - a$

The useful property of the length is additivity. If $\mathbb{I}$ is a disjoint union of a finite family $\{\mathbb{I}_n\}_{k=1}^{n}$ of intervals, then $|\mathbb{I}| = \sum_{k=1}^{n} |\mathbb{I}_n|$. If there are no gap between $\mathbb{I}_k$ then, each intervals $\mathbb{I}_k$ has necessarily the end points a_0, a_N, we conclude that $|\mathbb{I}| = a_N - a_0 = \sum_{i=0}^{N-1}(a_{i+1} - a_i) = \sum_{k=1}^{n} |\mathbb{I}_n|$

Definition. **A rectangle** $A \subset \mathbb{R}^2$ is a set of the form $A = \mathbb{I}_1 \times \mathbb{I}_2 = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{I}_1, y \in \mathbb{I}_2\}$ where $\mathbb{I}_1, \mathbb{I}_2$ are intervals. The area of a rectangle is $\text{Area}(A) = |\mathbb{I}_1| \times |\mathbb{I}_2|$. (Alexander Grigoran, 2007).

We claim that the area is also additive: if the rectangle is a disjoint union of a finite family of rectangles, that is $A = \bigcup_k A_k$, then $area(A) = \sum_{k=1}^{n} \text{area}(A_k)$.

Consider a particular case, when the rectangles $A_1, \dots A_k$ form a regular tiling of A, that is $A = \mathbb{I}_1 \times \mathbb{I}_2$ when $\mathbb{I}_1 = \bigcup_i \mathbb{I}_1$ and $\mathbb{I}_2 = \bigcup_j \mathbb{I}_2$ and assume that all rectangles A_k have the form

$\mathbb{I}_{1_i} \times \mathbb{I}_{2_j}$. Then $Area(A) = \sum_i \left|\mathbb{I}_{1_i}\right| \sum_j \left|\mathbb{I}_{2_j}\right| = \sum_{ij} \left|\mathbb{I}_{1_i}\right| \left|\mathbb{I}_{2_j}\right| = \sum_k Area(A_k)$. (Alexander Grigoran, 2007).

Definition. **Volume of domain in** $\mathbb{R}^3$. The construction is similar to the area.

Consider all boxes in $\mathbb{R}^3$, that is, the domain of the form $A = \mathbb{I}_1 \times \mathbb{I}_2 \times \mathbb{I}_3$ where $\mathbb{I}_1, \mathbb{I}_2, \mathbb{I}_3$ are intervals in $\mathbb{R}$, and set

$$vol(A) = |\mathbb{I}_1|.\,|\mathbb{I}_2|.\,|\mathbb{I}_3|$$

Then the volume is also an additive function.

1D	2D		General
Interval I	Rectangle R	$\Longrightarrow$	Space X
Length	Area	$\Longrightarrow$	Measure μ
Integral $\int_a^b f\,dx$	Double Integral $\iint_R f(x,y)\,dx\,dy$	$\Longrightarrow$	Lebesgue Integral $\int_X f\,d\mu$

(source: Dr. Nikolai Chernov,1994)

4. Hyper Rectangles and Hyperplanes $\{\boldsymbol{H_i}\}$

An interval of the form $[a, b] \subset \mathbb{R}^3$ has length $b - a$. This is consistent with the fact that a point has length zero, because the point $\{a\}$ can be written as $[a, a]$. The additivity of length implies that the intervals of the form $[a, b], (a, b)$ and $[a, b)$ all have the same length $b - a$. (Christopher Genovese, 2006).

To see why, notice that the interval $[a, b]$ can be written as a union pairwise disjoint intervals in three ways:

$$[a, b] = [a, a] \cup (a, b) \cup [b, b] = [a, a] \cup (a, b] = [a, b) \cup [b, b]$$

Additivity tells us that for any such disjoint partition of $[a, b]$, the sum of the lengths of the pieces must be $b - a$.

Since the boundaries $[a, a]$ and $[b, b]$ have length 0 (they are zero dimensional sets), all four types of intervals have the same length. This still holds if a or b are infinite; (a, ∞), $(-\infty, b)$, and $(-\infty, \infty)$ all have length ∞.

Similarly, the area of rectangle $[a_1, b_1] \times [a_2, b_2]$ is $(b_1 - a_1)(b_2 - a_2)$ and the volume of a rectangular box $[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ is $(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)$. These hold for rectangles and rectangular box with any part of the boundary missing because the boundary is a lower dimensional set. (Christopher Genovese, 2006).

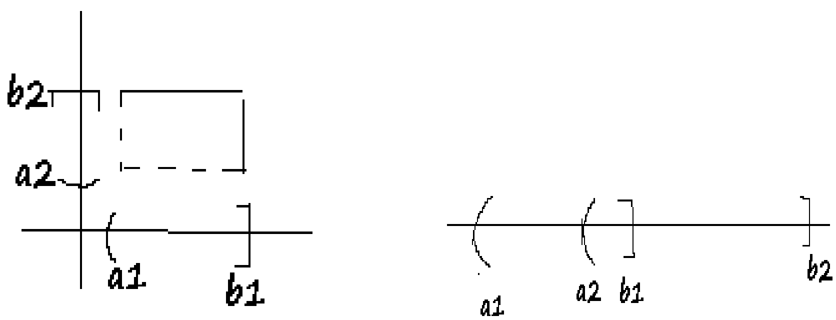


Figure 15 Two dimensional rectangle of half open intervals $(a_1, b_1] \times (a_2, b_2]$, and 1 dimension $(a_1, b_1] \cap (a_2, b_2] = (a_1, b_2]$

In Euclidean spaces $\mathbb{R}^n$, the set of the points $(x_1, x_2, \dots, x_n)$ of $\mathbb{R}^n$ that fulfill a linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, where $b, a_1, a_2, \dots, a_n$ are real numbers and, of these the coeffiecents $a_1, a_2, \dots, a_n$ do not all vanish together is called hyperplane $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$.

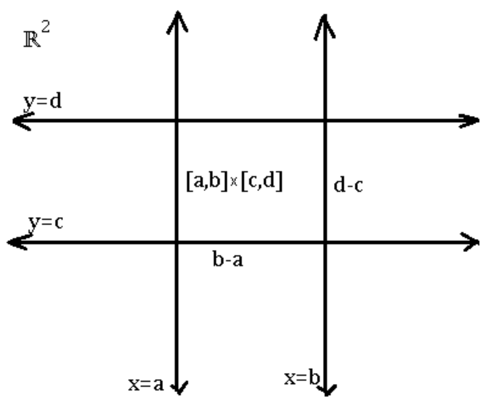
The space $\mathbb{R}$ is a straight line and the space $\mathbb{R}^2$ in a plane. Accordingly, the sets in $\mathbb{R}$ will often be called linear, and those in $\mathbb{R}^2$ plane sets. In $\mathbb{R}$ hyperplanes coincide with points, in $\mathbb{R}^2$ and $\mathbb{R}^3$

they are straight lines and planes respectively. If $b_1 - a_1 = b_2 - b_2 = \cdots = b_n - a_n > 0$, the interval $[a_1, b_1; a_2, b_2; \dots; a_n, b_n]$ is termed cube (square in $\mathbb{R}^2$). (Stanisław Saks, cited 2016)

5. Lebesgue Measures on $\mathbb{R}$ and $\mathbb{R}^n$

Theorem. There exists a unique Borel measure μ on $\mathbb{R}$ such that, for all $a, b \in \mathbb{R}$ with $a < b$, we have $\mu\big((a, b]\big) = b - a$. The measure μ is called Lebesgue measure on $\mathbb{R}$.

For Lebesgue measure on $\mathbb{R}^n$, we define the Cartesian product of two closed intervals $[a, b]$ and $[c, d]$ in $\mathbb{R}$ as the rectangle in $\mathbb{R}^2$ bounded by the lines $x = a, x = b, y = c$ and $y = d$.



(adopted from: Prof. Quibb,2013)

Its volume is the product of the lengths of the respective intervals; in this case, it is $(b - a)(d - c)$, in equation form, with the Cartesian product, $\mu([a, b] \times [c, d]) = (b - a)(d - c)$.

This approach is extended to three dimensions by considering the Cartesian product of three closed intervals, which takes the form of a rectangular prism, the volume of this set being the length times the width times the height. (Prof. Quibb,2013).

Then Lebesgue measure in $\mathbb{R}^n$ is defined as

$$\mu\big((a, b]\big) = (b_1 - a_1).(b_2 - a_2) \dots (b_n - a_n),$$

Which is the product of the sides

$$\mu((a,b]) = \prod_{k=1}^{n} \mu((a_k, b_k]),$$

Lebesgue measure in $\mathbb{R}^n$ can also be written as

$$\mu(|\mathbb{I}|) = \prod_{i=1}^{n} (b_i - a_i)$$

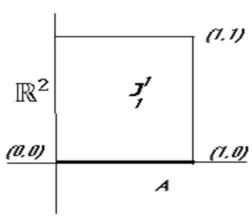
Where $\mathbb{I} = [a_1,b_1) \times \ldots \times [a_n,b_n)$.

Example. (Prof. Quibb,2013)

Let the single square J_1^1 in 2-dimensions $[0,1] \times [0,1]$ contain set A, having side length 1, the volume is 1.

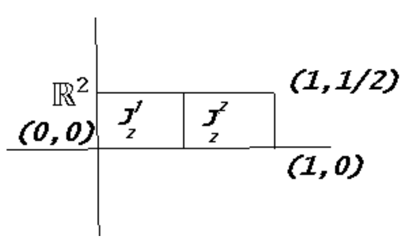
(The interval, is an uncountable set, but it can be shown by a succession of approximations by squares in $\mathbb{R}^2$ that this set has Lebesgue Measure 0)

$$\mu(J_1) = \mu\left(\bigcup_{i=1}^{1} J_1^i\right) = 1$$



The next case is two squares J_2^1 and J_2^2, each side length $\frac{1}{2}$, the total volume $\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$

$$\mu(J_2) = \mu\left(\bigcup_{i=1}^{2} J_2^i\right) = \frac{1}{2}$$



6. Basic Definitions and Properties

a) An algebra

Theorem. A class of sets $E \subset \mathcal{P}(\Omega)$ is an algebra if and only if the following three properties hold:

i. $\Omega \in E$

ii. E is closed under complements.

iii. E is closed under intersections.

b) A ring $\mathfrak{R}$

A nonempty collection of sets $\mathfrak{R}$ is a ring if

i. Empty set $\emptyset \in \mathfrak{R}$,

ii. if $A, B \in \mathfrak{R} \Rightarrow A \cup B$, and $A \backslash B \in \mathfrak{R}$,

iii. it follows $A \cap B \in \mathfrak{R}$ because $A \cap B = B \backslash (B \backslash A)$ is also in $\mathfrak{R}$.

The ring is closed under the formation of symmetric difference and intersection.

c) A semiring $\mathfrak{S}$

A semiring is a non empty class $\mathfrak{S}$ of sets such that

i. If $A \in \mathfrak{S}$ and $B \in \mathfrak{S}$, then $A \cap B \in \mathfrak{S}$ and

ii. If $A, B \in \mathfrak{S} \Rightarrow A \backslash B \in \mathfrak{S}$

For set $X = \mathbb{R}$ the set of $\mathfrak{S} = \{(a, b]: \infty < a \leq b < \infty\}$

d) σ-ring (sigma ring)

The σ-ring is a nonempty class S of sets such that,

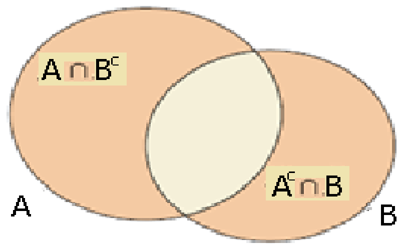
i. If $A, B \in S$ then $A - B \in S$.

ii. If $\mathcal{A}_i \in S$, $i = 1,2, \dots$, then $\bigcup_i^\infty \mathcal{A}_i \in S$. (Ross, Stamatis and Vladas, 2014)

Remarks:

Symmetric difference of two sets A and B is denoted by $A\Delta B$ and defined by

$$A\Delta B = (A - B) \cup (B - A) = (A \cap B^c) \cup (A^c \cap B)$$



Symmetric set difference. (source: Steve Cheng ,2008)

6.1 Measure Extension from Ring To σ-ring

The construction procedure of obtaining the σ-ring $\sum(\mathfrak{R})$from ring $\mathfrak{R}$ is as follows: denote by $\mathfrak{R}_\sigma$ the family of all countable unions of elements from $\mathfrak{R}$. Clearly, $\mathfrak{R} \subset \mathfrak{R}_\sigma \subset \sum(\mathfrak{R})$. Then denote by $\mathfrak{R}_{\delta\sigma}$ the family of all countable intersections from $\mathfrak{R}_\sigma$ so that

$$\mathfrak{R} \subset \mathfrak{R}_\sigma \subset \mathfrak{R}_{\delta\sigma} \subset \sum(\mathfrak{R})$$

Define similarly $\mathfrak{R}_{\sigma\delta\sigma}$, $\mathfrak{R}_{\sigma\delta\sigma\delta}$, etc. we obtain an increasing sequence of families of sets all being subfamilies of $\sum(\mathfrak{R})$. In order to exhaust $\sum(\mathfrak{R})$ by this procedure, one has to apply it uncountable many times. (adopted from: Alexander Grigoryan, 2007)

7. The Outer Measure

Definition. Given a measure space $(X, \mathcal{M}, \mu)$ for any set $\mathcal{A}$ subset of X, the outer measure μ^* is a σ-subadditive function on $\mathcal{P}(X)$.

The purpose of constructing an outer measure on all subsets of X is to pick out a class of subsets (to be called measurable) in such a way as to satisfy the countable additivity property.

a) The idea

To extend the length function into bigger subsets of power set of the real line and that what we call the outer measure, represented as a function $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$, which basically describe the size.

b) The Technique

We know that $\mu\colon |\mathcal{A}_i| \rightarrow [0, \infty]$ describes the size of the interval of a subset of real numbers. Therefore we take a set of real numbers and call it $\mathcal{M}$ and cover the set $\mathcal{M}$ with intervals of subset $\mathcal{A}$ (we could make two intervals, of $\mathcal{A}$) and find an arbitrary set containing at most countably infinite numbers of intervals such that $\mathcal{M}$ is contained within the union of the intervals, $\mathcal{M} \subseteq \bigcup_{i=1}^{\infty} \mathcal{A}_i$, see the figure below.

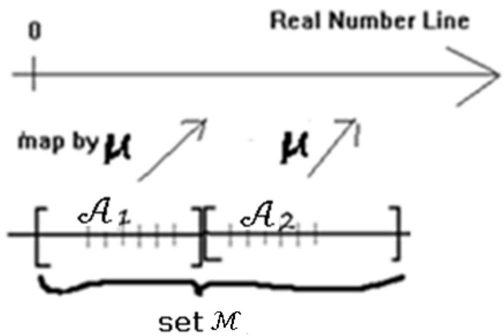


Figure 16 Lebesgue outer measure. (adopted from: Ben Garside cited, 2015)

As the term outer measure is defined on the finite (infinite) sequence of the intervals, we can compute the sum $\sum_{i=1}^{\infty} \mu(\mathcal{A}_i)$, in our case is $\mu(\mathcal{A}_1) + \mu(\mathcal{A}_2)$, that have the length. The μ map these intervals to the extended real number by this we get the size of set $\mathcal{M}$ that is less than or equal the sum:

$$\mu^*(\mathcal{M}) \leq \sum_{i=1}^{\infty} \mu(\mathcal{A}_i)$$

The μ^* is an outer measure and this is why we should examine the biggest set $\mathcal{M}$ of the union of the subsets $\mathcal{A}$ on the real number line

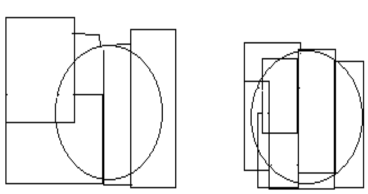
$$\mu^*(\mathcal{M}) = \inf\left\{\sum_{i=1}^{\infty} \mu(\mathcal{M}_i) | \ \forall \text{ coverings } \mathcal{A}_i \in \mathcal{A} \text{ such that } \mathcal{M} \subseteq \bigcup_{i=1}^{\infty} \mathcal{A}_i\right\}$$

But the sum of the lengths of the intervals is larger than the measure of $\mathcal{M}$. Therefore, the outer measure of $\mathcal{M}$ should be smallest (infimum) of the set of all such lengths of the intervals that cover $\mathcal{M}$. (adopted from Paul Loya, 2013)

Thus, to justify that, we say that any set on real numbers is bounded below, and the set always has an infimum and hence, we can work this out with for all $\mathcal{M}$ as a set, subset in the real line.

Figure 17 Set $\mathcal{M} \subseteq \mathbb{R}$, left half open intervals covering $\mathcal{M}$.

In $\mathbb{R}^2$, the cover of $\mathcal{M}$ consists of rectangles. For example, a disk in $\mathbb{R}^2$ and the cover represented by the union of the rectangles gives the lower approximation to $\mu^*(\mathcal{M})$ which is better than the upper approximation.



(source: Paul Loya, 2013)

7.1 Premeasure

Definition. Let $\mathcal{M} \subset \mathcal{P}(X)$ be a Boolean algebra, a premeasure on $\mathcal{M}$ is a function $\mu: \mathcal{M} \rightarrow [0, +\infty]$ on a Boolean algebra $\mathcal{M}$ that can be extended to a measure $\mu: \mathcal{B} \rightarrow [0, +\infty]$ on a σ-algebra $\mathcal{B} \subset \mathcal{M}$ and satisfying:

- $\mu(\emptyset) = 0$, if $\{\mathcal{A}_j\}_{j=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{M}$, and
- If $\bigcup_1^{\infty} \mathcal{A}_j \in \mathcal{M}$, then $\mu\left(\bigcup_{j=1}^{\infty} \mathcal{A}_j\right) = \sum_{j=1}^{\infty} \mu\left(\mathcal{A}_j\right)$.

Therefore the collection $\mathcal{M} \subset \mathcal{P}(X)$ of finite unions of half open intervals is Boolean algebra, and if $\mathcal{A} \subset \mathcal{M}$ is a disjoint union, $\mathcal{A} = \bigcup_{j=1}^{\infty}(a_j, b_j]$ then the function $\mu(\mathcal{A}) = \sum_{j=1}^{n} b_j - a_j$ is a premeasure on $\mathcal{M}$. Premeasure has a right additive property.

Finiteness and σ-finiteness are defined same for premeasure in the same way as for measures.

	Domain	Additivity condition
Premeasure	Boolean algebra $\mathcal{M}$	countably additive, when possible
Outer measure	all of $\mathcal{P}(X)$	Monotone countably subadditive
Measure	σ-algebra containing $\mathcal{A}$	countably additive

Table 5 The relationship between premeasure, outer measure and measure.(source: U.F., MAA6616 Course Notes Fall 2012).

7.2 Extending the Measure by Carathedory's Theorem

Definition. On the measure space $(X, \mathcal{M}, \mu)$, let $\mathcal{A}$ be a subsets of a sets X, and let $\mathcal{M} = \sigma(\mathcal{A})$. Suppose that μ is a mapping from $\mathcal{A}$ to $[0, \infty]$ that satisfies a countable additivity on $\mathcal{A}$.

Then, μ can be extended uniquely to a measure on $(\mathrm{X}, \mathcal{M})$. That is, there exists a unique measure μ^* on $(\mathrm{X}, \mathcal{M})$ such that $\mu^*(\mathbb{I}) = \mu(\mathbb{I})$ for all $\mathbb{I} \in \mathcal{A}$.(adopted from.MIT - 2008)

a) The Idea

Taking $(X, \mathcal{M}, \mu)$, first we start with defined a set function μ on a small paving $\mathcal{M}$, then we use μ to define a set function on all subset of X. Finally, we apply Carathedory's theorem in order to obtain measure. (J. Hoffman-Jorgensen,1994).

b) The Technique

Let X be a set, and let μ^* be an extended real valued function defined on $\mathcal{P}(X)$ such that;

1. $\mu^*(\emptyset) = 0$
2. If $\mathcal{A}, \mathcal{M} \in \mathcal{P}(X)$, $\mathcal{A} \subset \mathcal{M}$, then $\mu^*(\mathcal{A}) \leq \mu^*(\mathcal{M})$, this property called monotonicity.
3. μ^* is countably subadditive, i.e whenever $\mathcal{M} \in \mathcal{P}(X)$ and $(\mathcal{A}_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{P}(X)$ with $\mathcal{M} \subset \bigcup_{n=1}^{\infty} \mathcal{A}_n$, it follows that $\mu^*(\mathcal{M}) \leq \sum_{n=1}^{\infty} \mu^*(\mathcal{A}_n)$.

Remarks. μ^* is automatically subadditive in the sense that $\mathcal{M}, \mathcal{A}_1, \dots, \mathcal{A}_n \in \mathcal{P}(X)$ such that $\mathcal{M} \subset \mathcal{A}_1, \cup \dots \cup \mathcal{A}_n$, it follow that $\mu^*(\mathcal{M}) \leq \mu^*(\mathcal{A}_1) + \cdots + \mu^*(\mathcal{A}_n)$.

c) Carathedory's approach

Proposition. Let X be a nonempty set, and let $\mathfrak{S}$ be a semiring on X, and let $\mu: \mathfrak{S} \to [0, \infty]$ be a measure on $\mathfrak{S}$. Consider the collection:

$$\mathcal{P}_{\sigma}^{\mathfrak{S}}(X) = \left\{ \mathcal{M} \subset X: \text{there exist } (\mathcal{A}_n)_{n=1}^{\infty} \subset \mathfrak{S}, \text{with } \mathcal{M} \subset \bigcup_{n=1}^{\infty} \mathcal{A}_n \right\}.$$

Define the map $\mu: \mathcal{P}_{\sigma}^{\mathfrak{S}}(X) \to [0, \infty]$ by

$$\mu(\mathcal{M}) = inf \left\{ \sum_{n=1}^{\infty} \mu(\mathcal{A}_n): (\mathcal{A}_n)_{n=1}^{\infty} \subset \mathfrak{S}, \mathcal{M} \subset \bigcup_{n=1}^{\infty} \mathcal{A}_n \right\}, \forall \mathcal{M} \in \mathcal{P}_{\sigma}^{\mathfrak{S}}(X)$$

Then the map $\mu^*: \mathcal{P}(X) \to [0, \infty]$, defined by $\mu^*(\mathcal{M}) = \begin{cases} \mu(\mathcal{M}) & if\ \mathcal{M} \in \mathcal{P}_{\sigma}^{\mathfrak{S}}(X) \\ \infty & if\ \mathcal{M} \notin \mathcal{P}_{\sigma}^{\mathfrak{S}}(X) \end{cases}$

is an outer measure on X. (adopted from. Gabriel Nagy, 2011).

Caratheodory's definition is equivalent to Lebesgue's definition, but the advantage of Caratheodory's definition is that it works for unbounded sets too. (Paul Loya, 2013).

Remark. the set M is said to be bounded if $M \subset B(x, r)$ for some $x \in X$ and $r > 0$. This means if there an open ball with radius r then the entire set M in contain within the open ball.

8. Lebesgue, Uniform Measure on $[\mathbf{0}, \mathbf{1}]$ and Borel Line

Real intervals are the simplest sets whose measures or lengths are easy to define. A Borel line is a measurable set on the real line $\mathbb{R}$, therefore a Borel line is a measurable space $(\mathbb{R}, \mathcal{B})$. The idea of measure can be extended to more complicated sets of real numbers, leading to the Borel measure and eventually to the Lebesgue measure.

In order to discuss the Lebesgue measure in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we can replace the probability measure $\mathbb{P}$ with the measure μ, such hat the measure space is $(\Omega, \mathcal{F}, \mu)$. The Lebesgue measure that signed to any subset [0,1] is known as the uniform measure over the interval of length one and uniquely extend it to the Borel sets of [0,1], hence by applying the extension theorem we can conclude that there exist a probability measure $\mathbb{P}$, called Lebesgue or uniform measure defined on entire Borel σ algebra. The μ here as a set function is not the same as a calculus function, because its domain consists of sets [0,1] instead of elements, where ω_i are members of event,E that generate σ-field$= \sigma(E) = \mathcal{F}$.

The real line with Borel σ-algebra plays a special role in real-life probability because numerical data, real numbers, are recorded whenever a random experiment is performed. The number assigned to outcome of an experiment is the random variable.

Example. (Michael Kozdron, 2008). Suppose that $\mathbb{P}$ is the uniform probability measure on $\left([0,1], 2^{[0,1]}\right)$. This means that if $a < b$ and, $\mathbb{p}\{[a, b]\} = \mathbb{p}\{(a, b)\} = b - a$. The probability that falls in the interval $\left[0, \frac{1}{4}\right]$ is $\frac{1}{4}$, the probability that falls in the interval $\left[\frac{1}{3}, \frac{1}{2}\right]$ is $\frac{1}{6}$ and the probability that falls in either the interval $\left[0, \frac{1}{4}\right]$ or $\left[\frac{1}{3}, \frac{1}{2}\right]$ is:

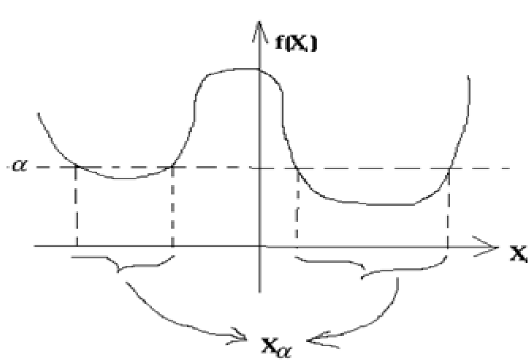
$$\mathbb{p}\left\{\bigcup_{i=1}^{n} |a_i, b_i| = \sum_{i=1}^{\infty} \mathbb{p}\{|a_i, b_i|\} = \sum_{i=1}^{\infty} (b_i - a_i)\right\}$$

$\mathbb{P}\left\{\left[0, \frac{1}{4}\right] \cup \left[\frac{1}{3}, \frac{1}{2}\right]\right\} = \mathbb{P}\left\{\left[0, \frac{1}{4}\right]\right\} + \mathbb{p}\left\{\left[\frac{1}{3}, \frac{1}{2}\right]\right\} = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}.$

9. The Advantage of Measurable Function

As we have shown before every continuous function is measurable. That is, continuous functions $f: \mathbb{R} \to \mathbb{R}$ are measurable. The interval $(-\infty, a)$ is an open set in the range, and so $f^{-1}\big((-\infty, a)\big)$ is an open set in the domain, and every open set is measurable so this gives us that f is measurable. For example, let $f(x)$ be a real function defined on space X. This functoion is called measurable if and only if, for every real number α, the set: (Frank Porter cited 2016)

$$X_\alpha = \{x | f(x) < \alpha, where \;\; x \in X\}$$



(source. Frank Porter cited 2016)

Measurable functions ensure the compatibility of the concepts of magnitudes and qualities. The measurable functions are important because we can use their inverse to project the measurable nature of the Borel line into an arbitrary set and turn it into a measurable space.

10. Induced Measure

Recall that R(ω) is a function and the random variable R: $\Omega \to \mathbb{R}$.

I. With $(\Omega, \mathcal{F}, \mu)$ as a probability space, consider the following:

The σ-algebra (of event) on Ω generate by R, denoted by $\sigma(R) = \sigma(E_1, E_2, \dots) = \mathcal{F}$ is:

$$\sigma(R) = \{R^{-1}(E):\ E \text{ Borel in } \mathbb{R}\}$$

The event that R takes values in the set E is

$$\{R \in E\} = \{\omega \in \Omega | R(\omega) \in E\} = R^{-1}(E)$$

Thus

$$\sigma(R) = \big\{\{R \in E\} \text{ Borel in } \mathbb{R}\big\}$$

Where $R^{-1}(E)$ is a measurable function and $\sigma(R)$ is the smallest σ-algebra on Ω that makes R measurable.

$$\mu(R \in E) = \mu(R^{-1}(E))$$

is the probability that the values of R are in the set E and measurable R is a random variable.

II. With $(\Omega, \mathcal{F}, \mathbb{P})$ as a measure space with total measure one, consider the following:

For all Borel sets $E \subseteq \mathbb{R}$, we define the probability $\mathbb{P}(E)$ of E by

$$\mathbb{P}_f(E) \equiv \mathbb{P}(R \in E) = \mathbb{P}(R^{-1}(E))$$

Here, we introduced f as a function that take R from the measurable space to $\mathbb{R}$ and the $\mathbb{P}_f$ is a probability measure induced on $\mathbb{R}$, hence $\mathbb{P}_f$ is the distribution of R.

This means that the domain where R is define not important as value and specific nature of Ω space not important as the probability of the value $\omega \in \Omega$. Thus if R is a random variable, then R induces a probability measure $\mathbb{P}$ on $\mathbb{R}$ defined by taking the primage for every Borel set of E.

$$\mathbb{P}_f(E) = \mathbb{P}\{\omega \in \Omega: R(\omega) \in E\} = \mathbb{P}\{R \in E\} \ = \mathbb{P}\{R^{-1}(E)\}$$

Where the probability measure $\mathbb{P}_f$ is called the distribution of R or $\sigma(R)$ as in (I):

$$R = \{\mathbb{P}(R \in E), E \text{ Borel in } \mathbb{R}\}$$

and this distribution lists the probabilities of each Borel (output) set.

Briefly, the random variable transforms the probability spaces: (Michael Kozdron,2008).

$$(\Omega, \mathcal{F}, \mathbb{P}) \overset{f}{\rightarrow} \left(\mathbb{R}, \mathcal{B}, \mathbb{P}_f\right)$$

and if we know the probability $\mathbb{P}$ of the value of every R, then we know the distribution. (Roman Vershynin, 2008).

Example. let f be a function that take R from the measurable space to the Borel line, $f: R \rightarrow \mathbb{R}$. Let f^{-1} be the inverse of the measurable function that map the event, E from the Borel line into an arbitrary set Ω. The probability that this event occurs given by

$$\mathbb{P}\{f \in \mathcal{B}\} = \mathbb{P}(f^{-1}(\mathcal{B})) = \text{the probability that the value of } f \text{ are in the set } \mathcal{B}$$

If $E_1, E_2, \ldots$ are pairwise disjoint Borel sets, then

$$\mathbb{P}_f\left(\bigcup_{n=1}^{\infty} E_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} f^{-1}(E_n)\right) = \sum_{n=1}^{\infty} \mathbb{P}\,(f^{-1}(E_n)) = \sum_{n=1}^{\infty} \mathbb{P}_f(E_n)$$

The measure $\mathbb{P}_f: \mathcal{B} \rightarrow [0,1]$ has the probabilistic behavior of the data that the function represents. (adopted from.Paul Loya, 2013).

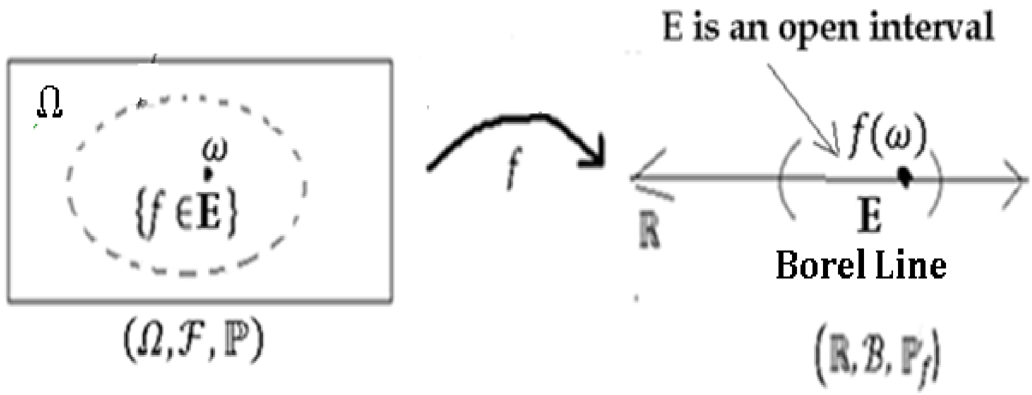


Figure 18 $\mathbb{P}_f$ is a probability measure on $(\mathbb{R}, \mathcal{B})$ such that for all Borel sets $E \in \mathbb{R}$, the $\mathbb{P}_f(E) =$ probability that f lies in the event set E.(adopted from Paul Loya, 2013).

notation	Measure Terminology	notation	Equivalent Probability Terminology
μ	Lebesgure measure	$\mathbb{P}$	Probability measure
X	sets of elements of real numbers, in general has also been referred to as $\mathbb{R}$	Ω	sets of elements of real numbers, in general has also been referred to as $\mathbb{R}$
$\mathcal{M}$	Borel algebra of subset of X, in general has also been referred to as $\mathcal{B}$	$\mathcal{F}$	Borel algebra of subset of Ω, in general has also been referred to as $\mathcal{B}$

Table.6 Terminology for both measure space and the equivalent probability space.

11. Probability Space and the Distribution Function

Definition. If $R: \Omega \to \mathbb{R}$ is a random variable, the distribution function of R written $F: \mathbb{R} \to [0,1]$ is defined by: (Michael Kozdron, 2008).

$$F(x) = \mathbb{P}_f\{(-\infty, x])\} = \mathbb{P}\{\omega | R(\omega) \leq \omega\} = \mathbb{P}\{R \leq x\}$$

For every $x \in \mathbb{R}$.

$$R(\text{random variable}) \leftrightsquigarrow \mathbb{P}_f(\text{law of } R) \leftrightsquigarrow F(\text{distribution function of } R)$$

Theorem: If F is a nondecreasing, right-continuous function satisfying $\lim_{x\to\infty+} F(x) = 1$ and $\lim_{x\to-\infty-} F(x) = 0$, then there exist on some probability space a random variable R for which $F(x) = \mathbb{P}[R \le x]$. (Patrick Billingsley, 1995)

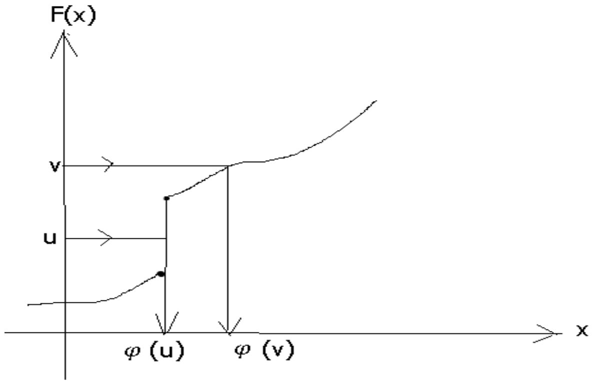
Proof: (by Patrick Billingsley, 1995)

For the probability space, take $(\Omega, \mathcal{F}, \mathbb{P}) = (\mathbb{R}, \mathcal{B}, \mu)$, the open unit interval Ω is $(0,1)$, $\mathcal{F}$ consists of the Borel subsets of $(0,1)$, and $\mathbb{P}(E) = \mu(E)$ is Lebesgue measure of E. To understand the method, suppose at first that F is continuous and strictly increasing. Then F is a one-to-one mapping of ℝ onto (0,1); let $\varphi: (0,1) \to \mathbb{R}$, be the inverse mapping. For $0 < \omega < 1$, let $R(\omega) = \varphi(\omega)$. Since φ is increasing, certainly R is measurable $\mathcal{F}$. If $0 < u < 1$, then $\varphi(u) \le x$ if and only if $u \le F(x)$. Since $\mathbb{P}$ is a Lebesgue measure, $\mathbb{P}[R \le x] = \mathbb{P}[\omega \in (0,1) | \varphi(\omega) \le x] = \mathbb{P}[\omega \in (0,1) | \omega \le F(x)] = F(x)$ as required.

If F has discontinuities or is not strictly increasing, it is defined as follows:

$$\varphi(u) = inf[x | u \le F(x)]$$

For $0 < u < 1$. Since F is nondecreasing, $[x | u \le F(x)]$ is an interval stretching to ∞; since F is right continuous, this interval is closed on the left. For $0 < u < 1$, therefore, $[x | u \le f(x)] = [\varphi(u), \infty]$ and so $\varphi(u) \le x$ if and only if $u \le F(x)$. If $R(\omega) = \varphi(\omega)$ for $0 < \omega < 1$, then by the same reasoning as before, R is a random variable and $\mathbb{P}[R \le x] = F(x)$.



(adpoted from: Patrick Billingsley,1995)

This provides a simple proof for the probability distribution F.(Patrick Billingsley, 1995).

The constructed distribution $F(x)$ that defined by $\mu(E) = \mathbb{P}[R \in E]$ for $E \in \mathcal{B}$ satisfies $\mu(-\infty,x] = F(x)$ and hence $\mu(a,b] = F(b) - F(a)$.

12. Lebesgue-Stieltjes Measure

Definition. Assume μ is a finite Borel measure on $\mathbb{R}$ and let $F(x) = \mu\big((-\infty, x]\big) = \mu(E)$, in which $\mu(E)$ is the length of the interval $(a,b]$ and therefore μ = Lebesgue measure on $\mathcal{B}$, then F is the distribution function of μ. The measure μ_F is called the Lebesgue-Stieltjes measure associated with F. (adopted from Kyle Siegrist, 2015)

The Lebesgue Stieltjes measure μ_F induced by a distribution function F is the measure defined $\mu_F(a,b] = F(b) - F(a)$ for interval $(a,b]$ and for other Borel sets on the real line by the method of measure extension. (A. K. Basu, 2012).

Definition. $\mathcal{P}(X)$ denote the set of all subset of X , and for $\mathcal{A} \subseteq \mathbb{R}$,

$$\mu(\mathcal{A}) = \inf\left\{\sum_{i=1}^{\infty}\big(F(b_i -) - F(a_i +)\big)\middle|\, \mathcal{A} \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i)\right\}$$

Which means, we look at all covering of $\mathcal{A}$ with open intervals and for each of these open covering, we add the length of the individual open intervals and take the infimum of all such numbers obtained. (Kuttler, 2014). The Lebesgue-Stieltjes measure is considered as a generalization of one dimensional Lebesgue measure.

12.1 Lebesgue-Stieltjes Measure as a Model for Distributions

The measure μ that associates with the distribution function F on $\mathbb{R}$ is Lebesgue-Stieltjes measure and has the following properties:

i. increasing $[a < b \ implies \ F(a) \leq F(b)]$ and

ii. right continuous $\left[\lim_{x \to x_0^+} F(x) = F(x_0)\right]$.

The Lebesgue-Stieltjes measure is a Borel measure on $\mathbb{R}$ which assume finite values on compact set (Rudi Weikard, 2016), and if the function F is to be defined as a Borel measure we need

$$\underset{i \to \infty}{lim}\, \mu_F \left(\left(a, a + \frac{1}{i}\right]\right) = \lim_{i \to \infty}\left[F\left(a + \frac{1}{i}\right) - F(a)\right] = \lim_{x \to a^+} F(x) - F(a) = 0$$

where $a_+ = \lim_{\Delta \to 0}(a + \Delta)$, that is $F(x)$ is continuous to the right.

The formula $\mu_F(a,b] = F(b) - F(a)$ sets up a one to one correspondence between Lebesgue-Stieltjes measures and distribution functions, where two distribution functions that differ by a constant are identified.

13. Properties of the Distribution Function

Theorem.The distribution function of $F(x)$ of a random variable R has the following properties: (Roman Vershynin, 2008)

i. F is nondecreasing and $0 \leq F \leq 1$

$F(x) \to 1$ as $x \to \infty$, i.e., $\lim_{x\to\infty+} F(x) = 1$

$F(x) \to 0$ as $x \to -\infty$, i.e., $\lim_{x\to\infty-} F(x) = 0$

ii. $F(x-) = \mathbb{P}(R < x), x \in \mathbb{R}$ and F has limit from the left

iii. $F(x+) = F(x), x \in \mathbb{R}$ and F is continuous from the right

F has left hand limits at all points (these are jumps on the distribution function, and these jumps mean the probability that R takes this specific value).

Suppose F is the distribution function of a real-valued random variable, if $a, b \in \mathbb{R}$ and $a < b$ then: (adopted from Kyle Siegrist, 2015).

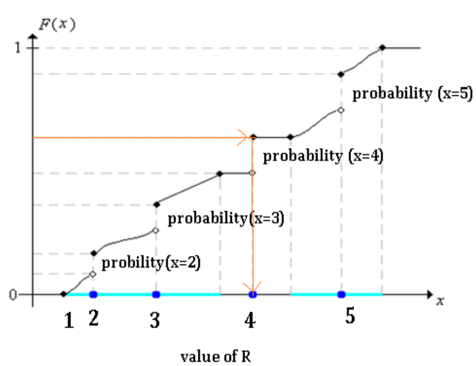
$$\mu\{a\} = \mathbb{P}(R = a) = F(a) - F(a-)$$

$$\mu(a, b] = \mathbb{P}(a < R \leq b) = F(b) - F(a)$$

$$\mu(a, b) = \mathbb{P}(a < R < b) = F(b-) - F(a)$$

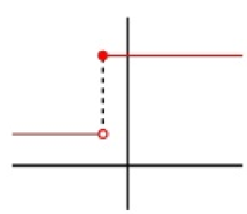
$$\mu[a. b] = \mathbb{P}(a \leq R \leq b) = F(b) - F(a-)$$

$$\mu[a, b) = \mathbb{P}(a \leq R < b) = F(b-) - F(a-)$$



Distribution function

(adopted from Kyle Siegrist, 2015)



$F(x)$ has Left hand limit and right continuous

(adopted from Ali Hussien Muqaibel)

14. Relate the Measure μ to the Distribution Function, F

The distribution of R is determined by its value of the form $\mathbb{P}(R \in E), E = (-\infty, x], x \in \mathbb{R}$ because such intervals form a generating collection of $\mathcal{F}$ and measure is uniquely determined by its value on a collection of generators, $R = \sigma\big((-\infty, x], x \in \mathbb{R}\big)$.

Consider the probability space defined as $(\Omega, \mathcal{F}, \mu)$ where Ω is a set, $\mathcal{F}$ is a Borel algebra of subset of Ω, and $\mu: \mathcal{F} \to \mathbb{R}$ a probability measure. The law of a random variable is always a probability measure on $(\mathbb{R}, \mathcal{B})$, thus the distribution function induced by μ is the function $F: \mathbb{R} \to [0,1]$ defined by:

$$F(x) = \mu\{(-\infty, x)\} \; \forall \; x \in \mathbb{R}$$

and, since

$$(-\infty, x] = \bigcap_{n=1}^{\infty} \left(-\infty, x + \frac{1}{n}\right)$$

is the countable intersection of open sets, we see that $(-\infty, x] \in \mathcal{B}$ for every $x \in \mathbb{R}$. Thus $\mu\{(-\infty, x]\}$ make sense. (Michael Kozdron, 2008).

15. Distribution Function on $\mathbb{R}^n$

Proposition. If $R_1, \dots, R_n$ are random variables, the $(R_1, \dots, R_n)$ is a random vector. (Roman Vershynin, 2008).

We can generalize the construction to many random variables $R_1, \dots, R_n$, once we know $F(x_1, \dots x_n) = \mu_F(R_1 \leq x_1, \dots, R_n \leq x_n)$. In this case $\Omega = \mathbb{R}^n$, $\mathcal{F} =$ Borel algebra generated by open sets of $\mathbb{R}^n$. (Joseph G. Conlon, 2010).

Further generalized to countably infinite set of variables $R_1, R_2, \dots$, provided the distribution of the variables satisfy an obvious consistency condition. Thus for

$$1 \le j_1 < j_2 < \cdots < N_j$$

Let $F_{j_1,j_2,\ldots,j_N}(x_1, x_2, x_N)$ be the distribution of $R_{j1}, R_{j2}, \ldots, R_{jN}$. Consistency then just means:

$$\lim_{x_k \to \infty} F_{j1,j2,\ldots,jN}(x_1, x_2, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_N)$$

$$= F_{j_1,j_2,\ldots,j_{k-1},j_{k+1},\ldots,j_N}\,(x_1, x_2, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_N)$$

for all possible choice of the $j_i \ge 1$, $x_i \in \mathbb{R}$. Here $\Omega = \mathbb{R}^\infty$ and $\mathcal{F}$ is the σ-field generated by finite dimensional rectangles.

The notion of a distribution function on $\mathbb{R}^n$, $n \ge 2$, is more complicated than the one dimensional case, and we may work with compact space $\overline{\mathbb{R}}$, that includes the $(-\infty, \infty)$.

Let $n = 3$, and let μ be finite measure on $\mathcal{B}(\mathbb{R}^3)$, then

$$F(x_1, x_2, x_3) = \mu\{\omega \in \mathbb{R}^3\colon\ \omega_1 \le x_1,\ \omega_2 \le x_2,\ \omega_3 \le x_3\}, (x_1, x_2, x_3) \in \mathbb{R}^3$$

By analogy with the one-dimensional case, we expect that F is a distribution function corresponding to μ. The formula $\mu_F(a, b] = F(b) - F(a)$ will be replaced by:

$$\mu_F(a, b] = \Delta_{b_1 a_1}\, \Delta_{b_2 a_2} \Delta_{b_3 a_3} F(x_1, x_2, x_3)$$

where Δ is the difference operator,

Our operation is introduced as follows:

If $G\colon \mathbb{R}^n \to \mathbb{R}, \Delta_{b_i a_i}\, G(x_1, \ldots x_n)$ is defined as

$$G(x_1, \ldots, x_{i-1}, b_i, x_{i+1}, \ldots, x_n) - G(x_1, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_n)$$

Lemma. If $a \le b$, that is, $a_i \le b_i, i = 1{,}2{,}3$ then

$$\mu_F(a, b] = \Delta_{b_1 a_1}\, \Delta_{b_2 a_2} \Delta_{b_3 a_3} F(x_1, x_2, x_3)$$

Where

$$\Delta_{b_1 a_1}\,\Delta_{b_2 a_2}\Delta_{b_3 a_3}F(x_1, x_2, x_3)$$

$$= F(b_1, b_2, b_3) - F(a_1, b_2, b_3) - F(b_1, a_2, b_3) - F(b_1, b_2, a_3) + F(a_1, a_2, b_3)$$

$$+ F(a_1, b_2, a_3) + F(b_1, a_2, a_3) - F(a_1, a_2, a_3) \geq 0$$

Thus $\mu_F(a, b]$ is not simply $F(b) - F(a)$

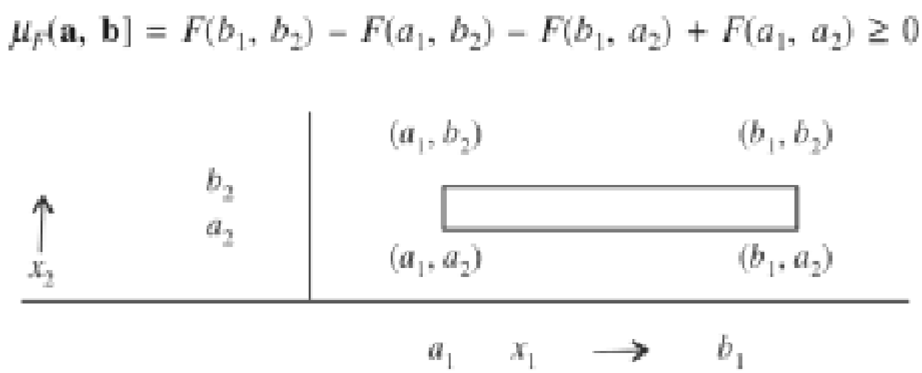


Figure 19 Cartesian product of $\mu\big((a_1, b_1] \times (a_2, b_2]\big)$.(source: A. K. BASU, 2012)

Non Example. (adopted from Rick Durrett, 2013)

Given $F(x, y)$ as below:

$$F(x,y) = \begin{cases} 1 & if\ x, y \geq 1 \\ 2/3 & if\ x \geq 1\ and\ 0 \leq y < 1 \\ 2/3 & if\ y \geq 1\ and\ 0 \leq x < 1 \\ 0 & otherwise \end{cases}$$

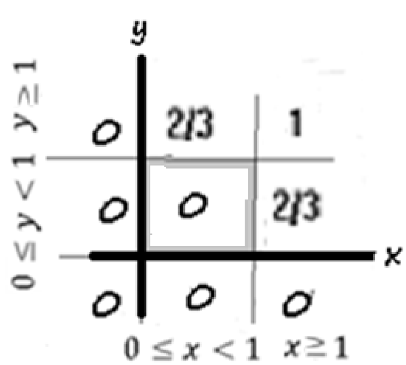


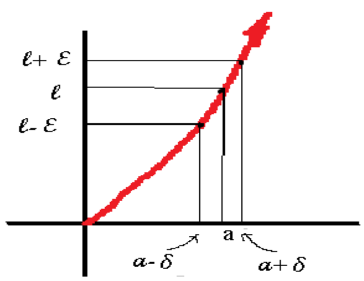
Figure 20 Non Example (adopted from Rick Durrett, 2013)

Suppose F gives rise to a measure μ on $\mathbb{R}^2$, $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ is right continuous, i.e.

$$F(a, b) = \mu((-\infty, a] . (-\infty, b])$$

F is increasing: if $x \leq y$ then $F(x) \leq F(y)$

- $F(x+) = F(x) \quad for\ x \in \mathbb{R}$ and F is continuous from the right
- $F(x-) = \mathbb{P}(X < x) \quad for\ x \in \mathbb{R}$ and F has limits from the left
- $F(-\infty) = 0$
- $F(\infty) = 1$



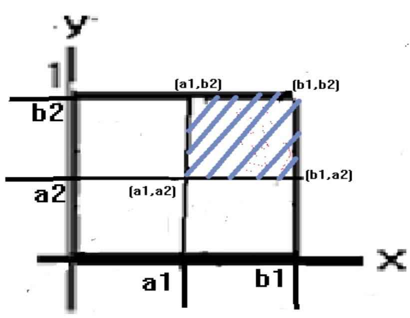


Figure 21 Solution for Non Example (Dr. Seshendra Pallekonda, 2015)

From the Cartesian product of $\mu\big((a_1, b_1] \times (a_2, b_2]\big)$ we have.

$$\mu_F(a, b] = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \geq 0$$

Then

$$\mu\big((a_1, b_1] \times (a_2, b_2]\big)$$

$$= \mu\big((-\infty, b_1] \times (-\infty, b_2]\big) - \mu\big((-\infty, a_1]$$

$$\times \big(-\infty, b_2]\big) - \mu\big((-\infty, b_1] \times (-\infty, a_2]\big) + \mu\big((-\infty, a_1] \times (-\infty, a_2]\big)$$

$$= F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2)$$

Let $a_1 = a_2 = 1 - \varepsilon$ and $b_1 = b_2 = 1$ and letting $\varepsilon \to 0$

$$\lim_{\varepsilon\to 0} \mu((1-\varepsilon, 1] \times (1-\varepsilon, 1])$$

$$= \lim_{\varepsilon\to 0} F(1,1) - \lim_{\varepsilon\to 0} F(1-\varepsilon, 1) - \lim_{\varepsilon\to 0} F(1, 1-\varepsilon) + \lim_{\varepsilon\to 0} F(1-\varepsilon, 1-\varepsilon)$$

Using $a_1 = a_2 = 1 - \varepsilon$ and $b_1 = b_2 = 1$ and letting $\varepsilon \to 0$

$$\mu(\{1,1\}) = 1 - \frac{2}{3} - \frac{2}{3} + 0 = -\frac{1}{3} \; which \; is < 0$$

$$\mu(\{1,0\}) = \mu(\{0,1\}) = \frac{2}{3}$$

$\mu(\{1,1\}) = -\frac{1}{3} < 0$

CONCLUSION IN GENERAL

Measure theory is foundation of probability and probabilities are special sorts of measures. The probability uses unlike terminology than that of measure but one can notice the similarity in the definition of measurable functions, and the Borel measure, in both theory or spaces.

In fact learning the main concept of measure theory as probability theory is easier than understanding this concept through real analysis. The reason is that probability is rich in applied examples beside the description of the probability jargon noteworthy pictured the idea rather than in real analysis or measure theory do. Thus most of the books about the measure theory consider the probability thought with the measure analysis.

The study includes a list of the measure, probability and topology components (elements) and examines the structure of their triple spaces in order to relate each one to the other. The measure can be written in different ways and the idea of the triple space is the same thus it can be applied to variety of fields of science, and it is not the matter if each one has driven from the other, it is the matter of how to use and understand the component of the structure.

To complete the conclusion I should thank all the writers, who make their science resources available online, that I used on forming this topic and hoping that this material has been assembled with minor mistake.

REFERENCES

Auyang, S. Y. (1995). *How is Quantum Field Theory Possible?* Oxford university press.

Basu, A. K. (2012). *Measure Theory and Probability*. Delhi: Asoke K. Ghosh.

Billingsley, P. (1995). *Probability and Measure, 3rd Edition, University of Chicago.* John Wiley & Sons.

Bradley, T.-D. (2015). *The Borel-Cantelli Lemma*. Retrieved from Mathema: http://www.math3ma.com/mathema/2015/9/21/the-borel-cantelli-lemma

Cheng, S. (April 27, 2008). *A crash Course on the Lebesgue Integral and Measure Theory*. Retrieved 2016, from http://www.gold-saucer.org/math/lebesgue/lebesgue.pdf

Chernov, N. (2012-2014). *Teaching by Nikolai Chernov at University of Alabama at Birmingham.* Retrieved from Real Analysis, MA 646-2F : http://people.cas.uab.edu/~mosya/teaching/teach.html

Durrett, R. (April 21, 2013). *Probability: Theory and Examples, Edition 4.1*. Cambridge University Press in 2010.

Epps, T. W. (2014). *Probability and Statistical Theory for Appiled Researchers, university of Virgina, USA.* World Scientific.

Etnyre, J. B. (2014). *Recollections from point set topology and overview of quotient spaces.* Retrieved 2016, from Georgia Institute of Technology: http://people.math.gatech.edu/~etnyre/class/6452Fall14/TopologyReview.pdf

Folland, G. B. (1999). *Real Analysis, Modern Techiques and Their Application, 2nd Edition.* John Wiley & Sons Inc.

G.Conlon, J. (2010, 12 17). *Math 625, Probability and Random Processes*. Retrieved 2016, from University of Michigan: http://www.math.lsa.umich.edu/~conlon/math625/chapter1.pdf

Garside, B. (2014, Oct). *functional analysis*. Retrieved Sep 2015, from https://www.youtube.com/playlist?list=PLAvgI3H-gclZa-DVTMyUIAxM-X8NSikwu

Garside, S. B. (2014, Sep 5). *Introduction to Metric Spaces*. Retrieved from https://www.youtube.com/watch?v=etP21xln8iQ

Genovese, C. (2006, spring). *Intermediate Probability*. Retrieved from Carnegie Mellon University: http://www.stat.cmu.edu/~genovese/class/iprob-S06/

Grigoryan, A. (2007-2008). *Measure Theory and Probability.* Retrieved from Lecture Notes, University of Bielefeld: https://www.math.uni-bielefeld.de/~grigor/mwlect.pdf

Herzong, F. (2013). *Probability Theory.* Retrieved from https://www.ethz.ch/content/dam/ethz/special-interest/mavt/dynamic-systems-n-control/idsc-dam/Lectures/Stochastic-Systems/Probability.pdf

Hoffman-Jorgensen, J. (1994). *Probability With a View Towards Statistics.* Chapman & Hall.

Jayasumana, S. (2012, June 11). *Sadeep's Tech Blog.* Retrieved from Blogger: http://sadeepj.blogspot.com/2012/06/understanding-riemannian-manifolds-part.html

Kannan, V. D. (n.d.). *Measure Theory for Applied Research*. Retrieved Oct 31, 2015, from https://www.youtube.com/channel/UCFwohpuBUlR3gdkJq3bn6_Q

Kannan, V. D. (n.d.). *Measure Theory for Applied Research*. Retrieved 2015, from Measurable Functions: https://www.youtube.com/watch?v=JoL0b6Az2CY

Kozdron, M. (winter 2008). *University of British Columbia.* Retrieved cited 2016, from Lecture Notes: http://stat.math.uregina.ca/~kozdron/

Kuttler. (2014, 3 21). *Real and Abstract Analysis.* Retrieved 2016, from Lecture Notes: https://www.math.byu.edu/~klkuttle/541book.pdf

Leadbetter, R., Cambanis , S., & Pipiras, V. (2014). *A Basic Course in Measure and Probability: Theory for Applications, 1st Ed.* Cambridge university press.

Lebl, J. (2011, Fall). *Basic Analysis: Introduction to Real Analysis.* Retrieved 2015, from University of Pittsburgh supplements: http://www.math.pitt.edu/~frank/pittanal2121.pdf

Loya, P. (2013). *Lebesgue's Remarkable Theory of Measure and Integration with Probability.* Retrieved cited 2016, from Binghamton university: http://www.math.binghamton.edu/loya/479-S13/Loyabook.pdf

MIT. (Fall 2008). *Massachusetts Institute of Technology.* Retrieved 2016, from Fundamentals of Probability, 6.436J / 15.085J, Lecture2: http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-436j-fundamentals-of-probability-fall-2008/lecture-notes/

Morais, M. C. (2009-2010). Lecture Notes - Probability Theory. In I. S. Tecnico.

Muqaibel, A. H. (2016, Mar 6). *SlideShare.* Retrieved from Random Variables Dr. Ali Hussein Muqaibel Ver. 2.3 1: http://www.slideshare.net/MuhannadSaleh/2-random-variables-notes-2p3

Nagy, G. (FALL 2001). *Notes in One File.* Retrieved from REAL ANALYSIS, Kansas State University: https://www.math.ksu.edu/~nagy/real-an/real-an-old/notes.pdf

Norris, J. (2010). *Probability and Measure.* Retrieved from http://www.statslab.cam.ac.uk/~james/Lectures/pmall.pdf

Omer, P. (2012, Dec 3). *Probability and Random Variables.* Retrieved from http://www.slideshare.net/pratomer/ch2-probability-and-random-variables-pg-81

Pallekonda, S. (2015, Dec). Senior Research Analyst . *Idaho Department Of Labor*.

Porter, F. (n.d.). *The Jordan Measure.* Retrieved cited 2016, from Physics 129a, Measure Theory: http://www.hep.caltech.edu/~fcp/math/measureTheory/measureTheory.pdf

Pra, P. D. (cited 2016). *Basic Measure Theory.* Retrieved from University of Padua: http://www.math.unipd.it/~daipra/didattica/galileiana-15/Parte-I.pdf

Quibb. (2013). *Lebesgue Measure.* Retrieved cited 2016, from Professor Quibb: http://quibb.blogspot.com/2013/02/lebesgue-measure.html

RJS. (2012, 1 28). *University of Texas at Dallas.* Retrieved from Some Basic Classes of Subsets of a Space X: http://www.utdallas.edu/~serfling/handouts/SomeBasicClassesOfSets.pdf

Royden, H. L. (1988). *Real Anaylsis, 3rd Edition.* Prentive-Hall, Inc.

Royden, H., & Fitzpatrick, P. M. (2010). *Real Analysis, 4th Edition.* Prentice Hall.

Saks, S. (2016). *Home Page of Stefan Banach.* Retrieved from Theory of the Integral, with additional notes by Stefan Banach: http://kielich.amu.edu.pl/Stefan_Banach/e-saks.html

Satish, S., & Harkrishan, V. L. (2011). *Multivariable Analysis.* Springer-Verlag London Limited.

Shen, J. (2012). *School of Economics, The University of New South Wales, Australia.* Retrieved from Mathematics for Economics and Finance, Lecture Notes (14/02/2012): http://jianfeishen.weebly.com/uploads/4/7/2/6/4726705/math2012.pdf

Siegrist, K. (1997-2015). *Probability Convergence.* Retrieved from University Alabama: http://www.math.uah.edu/stat/prob/Convergence.html#con

Susmel, R. (8/2015). *Bauer College of Business*. Retrieved from http://www.bauer.uh.edu/rsusmel/phd/sR-0.pdf.

Swartz, C. (1994). *Measure, Integration and function Space.* World scientific publishing .

Tarantola, A. (2007). *Mapping of Probabilities Theory for the Interpretation of Uncertain Physical Measurements.* Cambridge University Press.

Tarantola, A. (2007). Mapping of Probabilities Theory for the Interpretation of Uncertain Physical Measurements .

Thomson, B. S., Bruchner, J. B., & Bruchner, A. M. (2008). *Elementary Real Analysis.* Prentice Hall.

UF. (Fall 2012). *University of Florida College of Liberal Arts and Sciences,.* Retrieved from MAA6616 Course Notes : http://people.clas.ufl.edu/mjury/files/6616notesshort.pdf

Valkó, B. (2015). *Introduction to the Theory of Probability.* Retrieved from Math/Stat 431: http://www.math.wisc.edu/~valko/courses/431/431.html

Vershynin, R. (2007-2008). *MAT 235A / 235B: Probability, UM Personal World Wide Web Server* . Retrieved 2016, from Lecture Notes in Probability Theory, Function of Random Variables: http://www-personal.umich.edu/~romanv/teaching/2007-08/235B/lecture-notes.pdf

Weikard, R. (2016, Feb 29). *Real Analysis.* Retrieved from Lecture notes for MA 645/646: http://people.cas.uab.edu/~weikard/teaching/reals.pdf

Weisstein, E. (2015). *Vector Norm, (creator of MathWorld).* Retrieved 2015, from Created, developed, and nurtured by Eric Weisstein at Wolfram Research: http://mathworld.wolfram.com/Norm.html

Wikipedia. (n.d.). *Probability Measure*. Retrieved 2014, from https://en.wikipedia.org/wiki/Probability_measure

William, J. (2015). *The Lebesgue Integral for Undergraduates.* The Mathematical Association of America.