

Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Babylon
College of Education
Department of Mathematics



On the Measure Theory for Some Continuous Probability Distributions Density Functions

A thesis

*Submitted to the Council of the College of Education /
Babylon University in Partial Fulfillment of the requirements
for the
Degree of Master of Science in Mathematics .*

By

Khalid Fannukh Abu Khuder Al- Jubouri

Supervised by

Assist .Prof .Hussein Ahmed Ali

2009

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

٧

وَلَسَوْفَ يَعْطِيكَ رَبُّكَ فَتَرْضَى

صدق الله العلي العظيم

سورة الضحى: آية (5)



جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة بابل / كلية التربية

حول نظرية القياس لبعض توزيعات كثافة الاحتمال المستمرة

رسالة مقدمة إلى

قسم الرياضيات - كلية التربية - جامعة بابل

وكجزء من متطلبات نيل درجة الماجستير في علوم الرياضيات

من قبل

خالد فنوخ أبو خضر الجبوري

بإشراف

أ.م حسين احمد علي

2009م

1430هـ

Introduction

Probabilities are the most important applied topics in mathematics. They have great importance in knowing the probability distribution and its properties. Setting out from the principle of the joining among the knowledge, therefore, the researcher has turned to study the mixture between the probability and the measure theory which has the active role in simplifying the mathematic processes which are uncounted throughout the probability theory , therefore it has become the basic aim of present study .

The beauty of this topic starts from the researcher attempts to apply the theories and definitions on some cases which are impossible or difficult to be described from view of the probability theory's.

The present study is divided into four chapters.

Chapter one contains List of symbols, introduction ,literature reviews

Chapter two introduces many basic concepts ,definitions and theorems which are used in the later chapter like σ -field, measure, measurable function, simple function ,Lebesgue integration ,etc.

Chapter three gives some interesting connections and theorems about probability measure theory such as probability measure, random variable, convergence, convergence theorems ,etc.

Chapter four contains the results of calculate the Measure theory of some continuous function ,like normal , exponential , Cauchy, also of two functions for some continuous functions ,

Chapter five contains some conclusions and the recommendations.

Literature Reviews

There are some researchers who have worked in this field in the end of the twentieth century and beginning of twenty one century:

In 1909, Borel and Lebesgue introduced the basis for the measure-theoretical approach to portability. Important early work was done in 1909 by Borel - most notable Borel's form of the strong law of the large numbers. And the state of his Normal number Theorem is as : almost all numbers are strongly normal to all bases simultaneously (this all ten digits- 0,1, 2, ...,9 appear with equal asymptotic frequency in their decimal expansion, and similarly for the two digits 0, 1 in their dyadic expansion; likewise for duodecimal expansions, etc.). The expansion set are (Lebesgue) measure zero is clearly needed here, to cover such obvious exceptions as the rationales.

In 1933, Kolmogorov had published the basis for mathematical probability in his (German) 1933 monograph that made mathematical probability rigorous by being chary measure theory, hindered the acceptance of his own ideas by being chary in his use of measure theory terminology. For example, instead of defining a real random variable as a real measurable function ,he avoided the use of the adjective, going back to the defining property that the inverse image of an interval under the function that should be a set in his specified Borel field , he stated that his random variables were described elsewhere as measurable functions .He did not refer to the customary measure theory terminology in discussing almost everywhere convergence or convergence in measure, and went to the extreme of proving that the limit of a convergent sequence of random variables is a random variable. He obviously fell, and he was right, that the jump from Lebesgue to general measures was psychologically difficult.

In 1943, Cramer described the concepts of Lebesgue -measure and the Lebesgue integral. He then turned to the more general non-negative additive set functions on the Borel sets and to the Lebesgue Stieltjes integral. In particular he considered bounded, nonnegative, countable additive set functions on the Borel set of \mathbb{R} . Any such set function P corresponds to a unique (up to additive constant), nondecreasing point function $F(x)$ with the property that for any interval $(a, b]$.

$$F(b) - F(a) = P[(a, b]]$$

The function F is right continuous, and its discontinuity points form at most an enumerable set.

After giving these results, He narrowed his focus to those set functions for which $P(\mathbb{R}) = 1$. In this case he specified $F(x) = P[(-\infty, x]]$ and noted that

$$0 \leq F(x) \leq 1, \lim_{x \rightarrow -\infty} F(x) = 0 \text{ and } \lim_{x \rightarrow +\infty} F(x) = 1.$$

Either P or F could be used to define what Cramer called a distribution, P referred to as the probability function of the distribution, while F was called the distribution function. If F has a derivative, f , it is called the probability density or the frequency function of the distribution. (the probability derail)

In 1966, Lof, P. M, used acceptable definition of random sequences by using constructive measure theory, and our definition implies a sequence is random if a computer is incapable of compressing it.

In 1987, Dette, H.W, gave some investigation and results involving moments of probability measures on intervals of the real line and on the circle and defined the canonical moments of a probability measure, a basic familiarity with the moment spaces $\mu(a,b)$ and $\mu_n(a,b)$.

In 2001, Fets,T., used the construction of sets of jointed probability measures for the case that marginal sets of probability measures are generated by weighted focal sets. Different conditions on the choice of the weights of the joint focal sets and on the probability measures on these sets lead to different types of independence such as strong independence ,random set independence .

In 2002 ,Pap used many applications and formulas ,for example ,he began with lower or upper probabilities and the set of probability measures .

In 2004 ,Prokhorov gave the theory of weak convergence of probability measures on complete separable metric spaces ,and his results were examined on $C[0,1]$ and $D[0,1]$.Also gave notion of tightness play an important role both in the theory of weak convergence and in its applications .

Reference

- [1] AL–Hisnawi,W.A.J.,“Application of Central Limit Theorem on Probability Distribution by Using Differential Geometry “,Thesis of Master of Science in Mathematics, College of Education University of Babylon , 2006.
- [2] Ash,R.,”Real Analysis and Probability “, Academic Press,New York,1972.
- [3] Athreya, K.B. and Soumendera ,N.L .,“ Measure Theory and Probability“, 1998 .
- [4] Baelos , R.,” Lecture Notes Measure Theory and probability “ , Department of Mathematics , Purdue University , June 20 ,2003 .
- <http://www.math.purdue.edu/~bauelos/probability .pdf>
- [5] Bauer,H.”Measure Probability and integration”,Translated from the German by Robert ,Burckel .de Gruyter Studies in Mathematics ,26,Berlin: de Gruyter,x,v, 230 p,2001.
- [6] Bauer,H.”Probability”,Translated from the German by Robert ,Burckel.de Gruyter Studies in Mathematics ,23,Berlin :Walter de Gruyter, Gruyter,x,v,523p,1996.
- [7] Bremaud ,P. “An Introduction To Probabilistic Modeling “ ,Undergraduate Texts in Mathematics ,New York etc,: Springer – Verlag .xvi,207 p,1988.
- [8] Burrill , C.W., and Knudsen , J .R ., “ Real Variables “ ,Holt ,Rinehart and Winston .Inc ., New York ,1969 .
- [9] Chandalia.,“A Genyle Introduction to Measure Theory “ ,Department of computer science and Engineering , Suny- University at Buffalo , Buffalo, March 12, 2007
- [10] Davis, T.,“ Topics In Probability “ , August 14 ,2004 .

Reference

- [11] Dembo ,A.,“ Probability Theory :Stat 310 “, Department of Mathematics,Stanford University , September 20 , 2007 .
- [12] Esquerdo,P.J.R.”Convergence of Random Variables .Class Notes for Math 6601”,Department of Mathematics, University of Puerto Rico,Rio Piedras, Puerto Rico.
- [13] Fahady, K.S. and Shamoon , P. J., “ Probability “,Ministry of Higher Education and Scientific Research ,1990 .
- [14] Geiss,C.and Geiss, S.“ An Introduction To Probability Theory”, Department of Mathematics and Statistics ,University of Jyväskylä, March 7 ,2006.
<http://cse.ucdavis.edu/~chaos/courses/ncaso/Readings/Geiss.LectureNotesIntroprobTh.pdf>
- [15] Gohari,P.“Chapter 3: Random Variables and Probability Distribution “, Fall 2002.
<http://www.ece.concordia.ca/~gohari/engr371/3.pdf>
- [16] Gray,R.M.,“ Probability , Random Processes and Ergodic Properties “ , Standford University,August 20, 2008.
- [17] Gray,R.M. and Davisson, L.D.,“ Introduction to Statistical Signal Processing“ ,Cambridge University Press ,2004 .
- [18] Hewitt ,E. and Stromberg,K.”Real and Abstract Analysis .Amodern treatment of the theory of function of a real variable” ,3rd printing, Graduate Texts in Mathematics ,25,New York –Heidelberg Berlin:Springer –Verlag, x,476 p,1975 .
- [19] Jensen ,P. A .and Bard ,J .F., “ Probability Models .S3 Continuous Random Variables “ , 2001.

Reference

- <http://www.mw.uteXas.edu/~jensen/ORMM/supplements/models/probmode1/continuous.pdf>
- [20] “Chapter 4. Continuous Random Variables and Probability Distribution“, 2001 .
- http://www.am.qub.ac.uk/users/j.condon/SOR101/sor101_lectures_2up_issue_chapter4.pdf
- [21] Krogstad, H.E. ., “ MA8101 Measure and Integration Theory :The Basis of probability Theory “, September 8, 2003 .
- [22] Kurtz, T.G. ., “ Lecture on Stochastic Analysis “ , University of Wisconsin –Madison ,September 7, 2001 .
- [23] Kupferman ,R.,”Lecture notes in probability “, Institute of Mathematics ,The Hebrew University ,February 17, 2006 .
- [24] Lifton, J.H., “Measure Theory and Lebesgue Integration“ , 13 March , 1999.
- [25] Mood, A.M., Graybill ,F.A. and Boes, D.C.,”Introduction to the Theory of Statistics “ , 3d ed, McGraw –Hill, Inc. New York, 1974 .
- [26] Norris, J.R., “Advanced Probability “, University of Cambridge , 2003.
- [27] Papadimitrakis , M ., “Notes on Measure Theory “ Department of Mathematics , University of Crete , 2004 .
- [28] Pitman, J., “ Stat205 Probability Theory Topic: Probability Measures“, 2006 .
- [29] Prag, R., “Elementary Measure Theory and Statistics“, August 12 , 2006 .
- <http://math.nyu.edu/faculty/varadhan/limittheorems.html>
- [30] Rana ,I.K. ., “ Measure and Integration :Concepts , Examples and Exercises “, Department of Mathematics Bombay, India , May 2004 .

Reference

- [31] Royden ,H.L.,“ Real Analysis “, 3d ed ,Collier Macmillan Canada ,Inc. , USA ,1988.
- [32] Sattinger,D.H. ,“ Lecture Notes , Math 320 / 520 measure Theory and Integration“ , Department of Mathematics ,Yale University , 2004 .
- [33] Shiryaev,A,N.” Probability “Translated from the Russian by B,R.P,2 nd ed,Graduate Texts in Mathematics,95 , New York ,NY: Springer – Verlag. xiv,609 p,1995.
- [34] Suhov,Y., and Kelbert.,” Probability Statistics by Examples and”, 1d ed, Cambridge University ,New Yourk ,2005.
- [35] Suksompong,P.“ Measure Theoretic Probability Theory “, April 10, 2007.
- [36] Veldhuizen , T.L.,“ECE750-TXB Lecture 11: Probability“ Electrical & Computer Engineering,University of Waterloo,Canada,February 28,2007.
- [37] Waliji, M .“ Probability Theory “, August 11 , 2006 .
- [38] Walrand, J.“ Lecture Notes on Probability Theory and Random Processes “, Department of Electrical Engineering and Computer Sciences , University of California ,August 25 ,2004.
- [39] Whitt.,“Stochastic Models in FE Class Lecture Notes Random Variable, Expectation and Transformations” ,Thursday ,July 12, summer ,2007 .
- [40] Wikipedia,The Free Encyclopedia,“Convergence of Random Variables”,2005.
http://en.wikipedia.org/wiki/convergence_of_random_variables
- [41] Wikipedia,The Free Encyclopedia,“Law of Large Numbers”,2005.
http://en.wikipedia.org/wiki/law_of_large_numbers

Reference

[42] Wikipedia, The Free Encyclopedia, “Central of Limit Theorem”, 2005.

http://en.wikipedia.org/wiki/central_of_limit_theorem

[43] Wilde, I.F., “Measure Integration and Probability”, Mathematics

Department, King’s College London, 2007.

<http://www.math.kcl.ac.uk/~iwilde/notes/mip/mip.pdf>

Acknowledgments

First of all I offer my Praise and thanks to Almighty Allah .Who has granted me success throughout my M .A Study and enabled me to complete my present thesis . is to Allah for helping me to complete this thesis.

Then I would like to express my deep gratitude and appreciation , which go beyond words ,to my supervisor Assist. Prof .Hussein Ahmed Ali for suggesting the topic of this thesis ,and for his fatherly attention ,sincere help and encouragement throughout the preparation and writing of this thesis.

I'm Also greatly indebted to Dr. Lu'ay A.Al –Swidi,dean of the college of education ,for his scholarly guidance and sincere help , which he has always been ready to offer when needed ,and to all the staff of the department of Mathematics ,especially the head of the department of Mathematics, for their cooperation .

Words of thanks are also due to also Dr.Haider .K.Ammash for his support and encouragement.

Finally my love and gratitude are due to my family for their patience , help and encouragement especially my elder brother Dr .Chassib .

Abstract

In the last moment, a new method had appeared to mix among the different branches of mathematics with Statistics and Probability to solve many problems.

Depending on this principle. According to literature review of probability theory ,it is possible to apply on id (independent distribution). We apply this theory on some cases of one function (normal distribution ,exponential distribution and Cauchy distribution) and sum of two functions(sum of two normal distributions ,sum of normal distribution and exponential distribution and sum of normal distribution and Cauchy distribution) ,therefore it has become the basic aim of present study .

This thesis includes four chapters as follows :

The first chapter deals with some notations, terminology, and theorems from measure theory which we need in later chapter such as σ -field ,Borel σ -field, measure, Lebesgue measure , Lebesgue Stieltjes measure , measurable function ,simple function etc.

In the second chapter we give some definitions , concepts and theorems which are related to between probability theory and measure theory called (probability measure theory) such as probability measure ,random variable ,distribution function , independent ,expectation ,conditional ,convergence, convergence theorems ,etc.

In chapter three, the classical probability theory and the measure theory have been used to state the equality of probability and measure in one functions, and also when there are two functions including continuous

independent variables having the same distribution and continuous
independent variables having different distribution

then trying to find a function including the standard normal distribution and
standard Cauchy distribution at the same time.

We can put the following conclusions like :

1-We know that $\int_{-\infty}^{\infty} f(x)dx = 1$ in elementary probability theory . We also
find the same value in measure theory when we use as the simple function .

2-The sum two functions with i.i.d satisfy the
relation $M (f + g) = M (f) + M (g)$, when we apply it for the standard
normal distribution and standard Cauchy distribution .

3- The area under the curve of the function $\xi(w)$ doesn't equal one , which
means that this function isn't probability function on the two parts of interval
which are $(0, 1.851482)$ and $(1.851482, 30)$.

4- We find that the function $\xi_1(w) = \frac{\alpha \cos(\frac{w}{a})}{\beta\pi + w^\gamma}$ by using Gragger programmer
after multi experiment attempted . This function doesn't satisfy all conditions
probability function.

5- When $\xi_2(w) = \frac{f(w) + h(w)}{2}$,we find that it's a probability function satisfying
all the probability conditions on interval $(-30, 30)$.We show it is simple and

measurable function and , $\int_{-30}^{30} \xi_2 dM = 1$.

6- We can use the distribution function to show that the measure and probability measure give the same value in probability theory for the distributions we used .The same for the sum of two functions .

To future study ,the following are recommended :

- 1- Using another method can be solved This problem by like ; L^p space , fuzzy probability measure ,etc.
- 2- This method can be used to solve the problem when there is more than two distribution .
- 3- Try to use this method or any method to calculate the measure of sum of two discrete distributions which doesn't approximate the normal distribution .
- 4-If two functions have not any intersection points , try to use the following relation $M(f) = M(g)$ whenever $M(w:f(w) \neq g(w)) = 0$.
- 5-One can to use the relation $M(f) \leq M(g)$ if $f \leq g$ for two functions f and g .

<i>Subject</i>	<i>Page</i>
List of symbols	1
Introduction	2

Chapter One: Some Definitions and Concepts From Measure Theory **7-31**

1.1 Some Basic Concepts.....	7
1.2 σ - field.....	9
1.3 Borel σ -field.....	11
1.4 Lebesgue measure	12
1.5 Measurable set	13
1.6 Measure	13
1.7 Measurable function.....	18
1.8 Simple Function.....	24
1.9 Lebeque Integral	26

Chapter TWO : Probability Measure Theory
33-78

2.1 Some Basic Definitions of Probability	33
2.2 Probability Measure.....	34
2.3 Random Variables.....	36
2.4 Law of A random Variable	39
2.5 Some Special Continuous Distribution	44
2.6 Independence.....	45

2.7 Joint Distributed of r.vs.....	47
2.8 The moment generated function.....	51
2.9 Characteristic Function.....	51
2.10 Moments.....	53
2.11 Some Moment Inequalities.....	54
2.12 Convergence.....	65
2.13 Convergence Theorems.....	62
2.14 Limit Theorems.....	64
2.15 Relationship between of Convergence.....	68
2.16 The Law of Large Numbers.....	76
2.17 The Central Limit Theorem.....	77

*Chapter Three : Application of Measure Theory on
Some Probability Distributions* **80 - 102**

3.1 Case of one function.....	80
3.2 Case of two functions.....	86
3.3 Using The Measure and distribution function	99

Chapter Four : Conclusions and Recommendations

104-106

4.1. Conclusions.....	104
4.2 Recommendations.....	106
References.....	107

List of Symbols

Symbol	Meaning	Symbol	Meaning	Symbol	Meaning
A	Set or events	$\mu = E[X]$	The expected value	X^+	Max(X,o)
Ω	Sample space	(Ω, F, M)	Measure space	X^-	-Min(X,o)
Λ	Index set	\downarrow	Decreasing converge	\uparrow	Increasing converge
$\bigcup_{\lambda \in \Lambda} A_\lambda$	Union of a collection of s Ω sub set	$M_X(t)$	The moment generated function	$\varphi(w)$	Simple function
$\bigcap_{\lambda \in \Lambda} A_\lambda$	Intersection of a collection of s Ω sub set	$\Phi_X(t)$	Characteristic function	P	Probability measure
ϕ	Empty set	$f(x, y)$	The joint probability density function	(Ω, F, P)	Probability measure space
A_n	Sequence of sets A	$F(x, y)$	The joint distribution function	$\Phi_P(t)$	Characteristic function of P
\subset	Contained	$F_X(x)$	Marginal distribution function	$\min(X, Y)$	Minimum of X and Y
$1_A(\cdot)$	Indicator function of A	$f_X(x)$	Marginal density function	$\max(X, Y)$	Maximum of X and Y
\bar{A}	closure of set A	F	-field σ	$f(x)$	probability density function
\mathbb{R}	The real numbers	(Ω, F)	measurable space	M_*	Inner measure
N	The natural numbers	$B(P)$	Borel σ - field	$F(x)$	The cumulative distribution function
$i.o.$	Infinity often	M	measure		
X	Random variable or measurable function	\prod	Product	M^*	Outer measure

Measure theory provides a way to extend our notions of length, area, volume, etc. In order to be able to make use of measure and integrals, we need to know that the class of measurable sets is closed under certain types of operations. If we can assign a sensible notion of measure to a set, we are able to assign a sensible notion to its complement.

The major idea is to create a device by which we can measure sets in a fairly general context. Measure theory will allow us to do so. This is necessary for the following probability theory because we will build our probability distribution by using measure and the integrating random variables.

1.1 Some Basic Concepts

Definition (*Indicator function*) (1.1.1) [25]

The *indicator function* of A , denoted by $1_A(\cdot)$, is the function $1_A : \Omega \rightarrow \{0,1\}$, defined by

$$1_A(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{if } w \notin A \end{cases}$$

$1_A(\cdot)$ clearly "indicates" the set A in Ω .

Definition (1.1.2) [31]

An infinite set is said to be *countable* if its elements can be put in one to one correspondence with the natural number $1, 2, \dots$, in this case it is called a *countable infinite*. If the set has as many points as there is some interval of real numbers $[a,b]$, is called a *uncountable infinite set*.

Definition(1.1.3)] [15]

A sequence of sets $\{A_n\}$ is said to be *increasing* (or *nondecreasing*) if $A_n \subset A_{n+1}$ for each n . If $A_{n+1} \subset A_n$ for each $n \in N$,then the sequence $\{A_n\}$ is said to be *decreasing* (or *non increasing*).A *monotone* sequence is one which is either increasing or decreasing.

Definition (1.1.4) [6]

The function g of real number is called *bounded* if there is a positive constant K such $|g(w)| \leq K$, for all w .

Definition: (1.1.5) [18,31]

A real valued function g ,defined as a set Ω is said to be *continuous* at $w_2 \in \Omega$ if given $\varepsilon > 0$, there is a $\delta > 0$ such that for all w_1 in Ω with $|w_1 - w_2| < \delta$, we have $|g(w_1) - g(w_2)| < \varepsilon$.

Theorem(1.1.6)[18]

The function $g: \Omega \rightarrow \square$ be continuous if and only if for all open set U in \square , then $g^{-1}(U)$ is open set in Ω .

Definition (1.1.7) [18]

A real valued function g , defined on a set Ω is said to be *uniformly continuous* (on Ω) if given $\varepsilon > 0$, there is a $\delta(\varepsilon) > 0$ such that for all w_1 and w_2 in Ω with $|w_1 - w_2| < \delta$,we have $|g(w_1) - g(w_2)| < \varepsilon$

1.2 σ - field [3,37]

Suppose that Ω is a nonempty set. A collection F of subsets of Ω is called a σ -field (σ -algebra) on Ω if it contains Ω and is closed under the operations of taking complements and countable unions. That means,

I. $\emptyset, \Omega \in F$

II. If a set $A \in F$, then $A^c \in F$ (where $\Omega \setminus A = A^c$)

III. If $\{A_n\}$ be a sequence in F for $n=1,2,\dots$, then $\bigcup_{n=1}^{\infty} A_n \in F$.

The pair (Ω, F) , where Ω is a nonempty set, F is a σ -field of subsets of Ω , is called *measurable space*. A subset A of Ω is called *measurable set* if $A \in F$ and we write the $\sigma(A)$ as *the smallest σ -field containing A* it called σ -field generated by A .

Some Basic Properties (1.2.1) [6,43]

I. For any $A_1, A_2, \dots, A_n \in F$, put $A_{n+1} = A_{n+2} = \dots = \emptyset$. Then we see that

$$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^{\infty} A_i \in F, \text{ by (III) above.}$$

II. Let $A_1, A_2, \dots \in F$. Then since $\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c$, we see that $\bigcap_{n=1}^{\infty} A_n \in F$.

If we take $A_{n+1} = A_{n+2} = \dots = \Omega$, then we get $A_1 \cap A_2 \dots \cap A_n \in F$.

III. Let $A, B \in F$. Then $A \setminus B = A \cap B^c$ and so $A \setminus B \in F$.

Example (1.2.2)[27]

Let $\Omega = \square$ then the collection $\{A \subseteq \square, A \text{ is countable or } A^c \text{ is countable}\}$ is σ -field ?

Solution

By assumption $\Omega = \square$ and $A \subseteq \square$, then we see that $\Omega \in F$ and also A is countable in F , then $A^c \in F$.

Suppose that a countable A_1, A_2, \dots are all in F , if each of the A_n are countable,

then $\bigcup_n A_n$ is countable and so in F . If $A_{n_0}^c$ is countable for some n_0 , then

$(\bigcup_n A_n)^c = \bigcap_n A_n^c \subseteq A_{n_0}^c$ is countable and again $\bigcup_n A_n$ is in F . Since

$\bigcap_n A_n = (\bigcup_n A_n^c)^c$, then the countable intersection of sets in F is again in F .

Example(1.2.3)

Let $\Omega = N$, and let $B = \{i \in N : i = 2n \text{ for some } n \in N\}$. Set

$$A = \{A \subseteq N : \text{either } B \subseteq A \text{ or } A \cap B = \emptyset\}.$$

Show that A is a σ -field of subsets of N .

Solution

To check A is a σ -field :

1. $\emptyset \cap B = \emptyset$, so $\emptyset \in A$, $\Omega = N$ and $A \subseteq N$, then we see $\Omega \in A$.
2. If $A \in A$, then either $B \subseteq A$ – in which case $A^c \cap B = \emptyset$ or $A \cap B = \emptyset$ – in which case $B \subseteq A^c$. In either case, A^c is in F .
3. If A_1, A_2, \dots are all in A , then either B is a subset of one of the A_i – in which it is also a subset of $\bigcup A_i$, or $B \cap A_i = \emptyset$ for all i – in which case $B \cap (\bigcup A_i) = \emptyset$. In either case, $\bigcup A_i$ is in A .

1.3 Borel σ -field [7]

Let $\Omega = \mathbb{R}$, then the *Borel σ -field* is the smallest σ -field containing all open sets and closed set, usually written $B(\mathbb{R})$. In other words, it is the σ -field

$$B = \bigcap \{F : F \text{ is a } \sigma\text{-field, } u \text{ open set, then } u \in F \}. \quad \dots\dots\dots (1.1)$$

i.e.

It is the intersection of the class of all σ -fields that contain all the open sets and the elements of $B(\mathbb{R})$ are called *Borel sets*.

Theorem (1.3.1) [6,43]

There are many sub sets of \mathbb{R} belong to $B(\mathbb{R})$, like

- I. (a, b) for any $a < b$;
- II. $(-\infty, a)$ for any $a \in \mathbb{R}$;
- III. (a, ∞) for any $a \in \mathbb{R}$;
- IV. $[a, b]$ for any $a < b$;
- V. $(-\infty, a]$ for any $a \in \mathbb{R}$;
- VI. $[a, \infty)$ for any $a \in \mathbb{R}$;
- VII. $(a, b]$ for any $a < b$;
- VIII. $[a, b)$ for any $a < b$;
- IX. closed sub set of \mathbb{R} .

1.4 Lebesgue Measure [4,27]

The Lebesgue outer measure of a set $A \subseteq \mathbb{R}$ is defined as:

$$M^*(A) = \inf \{M(I_n) : M \text{ is a measure on interval } I_n \text{ and } A \subseteq \bigcup_n I_n \} \dots\dots\dots (1.2)$$

The outer as defined above, is again of generalized form . Consider an interval I with end points a and b such that $a < b$ instead of the set A , then the outer measure for the interval is defined as $b-a$.

The Lebesgue inner measure is defined as:

$$M_*(A) = \sup\{M^*(B) : B \subset A, B \text{ compact}\} \text{ [see 6]} \dots\dots\dots(1.3)$$

or

$$M_*(A) = M(\Omega) - M^*(\Omega \setminus A)$$

If the following is true $M^*(A) = M_*(A)$, then $M(A)$ (which denotes the values of $M^*(A)$ or $M_*(A)$) is said to be Lebesgue measure on the set A .

Remark (1.4.1)[2,3]

1-The interval $[a,b]$ such that $[a,b] = \{x \in \mathbb{P} ; a \leq x \leq b\}$ has measure $b-a$ and also that $(a,b), (a,b], [a,b)$ for any constants a, b .

2-The rational number has measure zero.

Example (1.4.2) [24]

Let $\{a_1, a_2, a_3, \dots, a_n\} \subset I$ be a finite set of points and $I = [a, b]$. Let $a < a_n < a_m < b$ for all $n < m$. Consider the set $A = I \setminus \{a_1, a_2, \dots, a_n\}$, the interval I without the points a_1, \dots, a_n . Let our measure be such that the measure of any interval with end points $a < b$ is $b-a$. Then ,

$$\begin{aligned} M(A) &= M([a, a_1) \cup (a_1, a_2) \cup \dots \cup (a_{n-1}, a_n) \cup (a_n, b]) \\ &= M([a, a_1)) + M((a_1, a_2)) + \dots + M((a_{n-1}, a_n)) + M((a_n, b]) \\ &= (a_1 - a) + (a_2 - a_1) + \dots + (a_n - a_{n-1}) + (b - a_n) = b - a = M(I) . \end{aligned}$$

1.5 Measurable set [31]

A set E is said to be measurable if for each set A we have

$$M^*(A) = M^*(A \cap E) + M^*(A \cap E^c).$$

Lemma (1.5.1)[2]

If $M^*(A) = 0$, then A is measurable.

Theorem(1.5.2)[31]

Every Borel set is measurable. In particular each open set and closed set is measurable.

1.6 Measure[4,6]

A *measure* M is any countable additive nonnegative set function of the measurable space (Ω, \mathcal{F}) . That is, $M : \mathcal{F} \rightarrow [0, \infty]$, with the properties :

I. $M(A) \geq 0$, for all $A \in \mathcal{F}$.

III. $M\left(\bigcup_n A_n\right) = \sum_n M(A_n)$, for any countable collection

of disjoint sets $A_n \in \mathcal{F}$.

Property (III) of a measure is called σ -*additivity* and sometimes the measure is also called σ -*additive measure*

A set A is said to be *finite measure* if $A \in \mathcal{F}$ and $M(A) < \infty$. Also it is said to be σ -*finite measure* if A is the union of a countable collection of measurable sets of finite measure.

A triple (Ω, \mathcal{F}, M) , where Ω is a nonempty set, \mathcal{F} is a σ -field of subsets of Ω , and M measure on Ω , is called a *measure space*.

Theorem (1.6.1)[30,43]

Let (Ω, \mathcal{F}, M) be a measure space. Then the following hold ;

I. If $A, B \in \mathcal{F}$ with $A \subseteq B$, then $M(A) \leq M(B)$. (*monotonicity*)

II. If the sequence $A_n \in \mathcal{F}$ and a set $A \subseteq \bigcup_{n=1}^{\infty} A_n$, then $M(A) \leq \sum_{n=1}^{\infty} M(A_n)$.
(*sub additivity*)

III. If $A_i \in \mathcal{F}$, $A_1 \subset A_2 \subset A_3 \subset \dots$, and $A = \bigcup_{i=1}^{\infty} A_i$, then we have
 $M(A) = \lim_{n \rightarrow \infty} M(A_n)$.

IV. If $A_i \in \mathcal{F}$, $A_1 \supset A_2 \supset A_3 \supset \dots$, $M(A_1) < \infty$, and $A = \bigcap_{i=1}^{\infty} A_n$, then we have
 $M(A) = \lim_{n \rightarrow \infty} M(A_n)$.

Proof

I. Write $B = A \cup (B \setminus A)$. Hence, by using (III) of definition (2.3)

$$M(B) = M(A) + M(B \setminus A) \geq M(A).$$

II. We can do many sets infinitely.

$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (A_1^c \cap \dots \cap A_{n-1}^c \cap A_n)$ where the sets in the last union are disjoint. Therefore,

$$M \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} M \left(A_n \cap A_1^c \cap \dots \cap A_{n-1}^c \right) \leq \sum_{n=1}^{\infty} M \left(A_n \right).$$

III. Put $B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus A_2$ and so on. Then $B_n \in F$ and

$(B_i \cap B_j = \phi)$ are for $i \neq j$. since $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A_i$, then

$$M \left(A \right) = M \left(\bigcup_{i=1}^{\infty} A_i \right) = M \left(\bigcup_{i=1}^{\infty} B_i \right) = \sum_{i=1}^{\infty} M \left(B_i \right) =$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n M \left(B_i \right) = \lim_{n \rightarrow \infty} M \left(\bigcup_{i=1}^n B_i \right) = \lim_{n \rightarrow \infty} M \left(\bigcup_{i=1}^n A_i \right) = \lim_{n \rightarrow \infty} M \left(A_n \right).$$

IV. Set $B_n = A_1 \setminus A_n$. Then $B_n \in F$ and $B_1 \subset B_2 \subset \dots$. Let

$$A = \bigcap_{n=1}^{\infty} A_n, \text{ then } A_1 \setminus A = \bigcup_n B_n \text{ giving } M \left(A_1 \setminus A \right) = \lim_n M \left(B_n \right), \text{ by (III).}$$

But $M \left(B_n \right) = M \left(A_1 \setminus A_n \right) = M \left(A_1 \right) - M \left(A_n \right)$ and so

$$M \left(A_1 \right) - M \left(A \right) = \lim_{n \rightarrow \infty} \left(M \left(A_1 \right) - M \left(A_n \right) \right), \text{ giving } M \left(A \right) = \lim_n M \left(A_n \right). \text{ The}$$

sequence $(M \left(A_n \right))$ is decreasing since we have $A_{n+1} \subseteq A_n$. \square

Theorem (1.6.2)[4]

Let F' be a semi field and let M be defined on F' . Suppose $M(\phi) = 0$ with the additional properties :

I. If $A \in F', A = \bigcup_{i=1}^n A_i, A_i \in F'$ disjoint, then $M \left(A \right) = \sum_{i=1}^n M \left(A_i \right)$ and

II. If $A \in F', A = \bigcup_{i=1}^{\infty} A_i, A_i \in F'$ disjoint, then $M \left(A \right) \leq \sum_{i=1}^{\infty} M \left(A_i \right)$.

Definition (Lebesgue - Stieltjes Measure) (1.6.3) [4,43]

Let $F : \mathbb{R} \rightarrow \mathbb{R}$ satisfy i. nondecreasing ii. right continuous. Let C be the class of sets of the form $(a, b]$ or $(b, \infty), -\infty \leq a < b < \infty$. Then there is a

measure $M_{\mathbb{F}}$ defined on $B(\mathbb{P})$, the smallest σ -field generated by C . such that

$$M_{\mathbb{F}}(a, b] = F(b) - F(a), \quad -\infty \leq a < b < \infty.$$

Theorem (1.6.4) [43]

If F is a distribution function corresponding to M , then

$$\text{I. } M((-\infty, x)) = F(x^-).$$

$$\text{II. } M((x, y]) = F(y) - F(x^+), x < y.$$

$$\text{III. } M((x, y)) = F(y^-) - F(x^+), x < y.$$

$$\text{IV. } M([x, y)) = F(y^-) - F(x^-), x < y.$$

$$\text{V. } M([x, y]) = F(y) - F(x^-), x < y.$$

$$\text{VI. } M(\mathbb{R}) = F(\infty) - F(-\infty).$$

Theorem (1.6.5) [4,22]

Let M be a Lebesgue-Stieltjes measure on \mathbb{R} . Define the distribution function $F: \mathbb{R} \rightarrow \mathbb{R}$, up to additive constants by $F(b) - F(a) = M(a, b]$, then F is an extended distribution.

Proof

Let $a < b$. Then $F(b) - F(a) = M(a, b] \geq 0$. Also, if $\{x_n\}$ is such that $x_1 > x_2 > \dots \rightarrow x$, then $M(x_1, x_n] \rightarrow 0$, by theorem (2.3.2), part (IV), since

$$\bigcap_{n=1}^{\infty} (x_1, x_n] = \emptyset \text{ and } (x_1, x_n] \downarrow \emptyset. \text{ Thus } F(x_n) - F(x) \rightarrow 0 \text{ implying that } F \text{ is}$$

right continuous. \square

Theorem (1.6.6) [4]

Suppose F is a distribution function on \mathbb{R} . There is a unique measure M on $B(\mathbb{R})$ such that $M(a, b] = F(b) - F(a)$(1.4)

Proof

Let $I = \{(a, b) : -\infty \leq a < b < \infty\}$. Set $F(\infty) = \lim_{x \uparrow \infty} F(x)$ and $F(-\infty) = \lim_{x \downarrow -\infty} F(x)$. These quantities exist since F is increasing. Define for any $M(a, b] = F(b) - F(a)$, for any $-\infty \leq a < b \leq \infty$, where

$F(\infty) > -\infty, F(-\infty) < \infty$. Suppose $(a, b] = \bigcup_{i=1}^n (a_i, b_i]$, where the union is

disjointed. By relabeling

We may assume that $a_1 = a, b_n = b, a_i = b_{i-1}$. Then $M(a_i, b_i] = F(b_i) - F(a_i)$ and

$$\sum_{i=1}^n M(a_i, b_i] = \sum_{i=1}^n (F(b_i) - F(a_i)) = F(b) - F(a) = M(a, b]$$
 which prove that conditions (I) of (2.3.5) holds.

for (II) of (2.3.5), Let $-\infty < a < b < \infty$ and $(a, b] \subset \bigcup_{i=1}^{\infty} (a_i, b_i]$, where the union is disjoint. By right continuity of F , given $\varepsilon > 0$ there is a $\delta > 0$ such that

$$F(a + \delta) - F(a) < \varepsilon,$$

or equivalently

$$F(a + \delta) < F(a) + \varepsilon$$

Similarly, there is a $\eta_i > 0$ such that

$$F(b_i + \eta_i) < F(b_i) + (\varepsilon 2^{-i}),$$

for all i . Now, $\{(a_i, b_i + \eta_i)\}$ forms an open cover for $[a + \delta, b]$. By compactness [see 31], there is a finite sub cover. Thus,

$$[a + \delta, b] \subset \bigcup_{i=1}^n (a_i, b_i + \eta_i) \quad \text{and} \quad (a + \delta, b] \subset \bigcup_{i=1}^n (a_i, b_i + \eta_i].$$

Therefore,

$$\begin{aligned} F(b) - F(a + \delta) &= M(a + \delta, b] \leq \sum_{i=1}^n M(a_i, b_i + \eta_i) = \sum_{i=1}^n F(b_i + \eta_i) - F(a_i) \\ &= \sum_{i=1}^n \{F(b_i + \eta_i) - F(b_i) + F(b_i) - F(a_i)\} \leq \sum_{i=1}^n \varepsilon 2^{-i} + \sum_{i=1}^n (F(b_i) - F(a_i)) \\ &\leq \varepsilon + \sum_{i=1}^{\infty} F(b_i) - F(a_i). \end{aligned}$$

Therefore

$$M(a, b] = F(b) - F(a) \leq 2\varepsilon + \sum_{i=1}^{\infty} F(b_i) - F(a_i) = 2\varepsilon + \sum_{i=1}^{\infty} M(a_i, b_i], \text{ for all}$$

$\varepsilon > 0$, then $M(a, b] \leq \sum_{i=1}^{\infty} M(a_i, b_i]$, proving (II) provided $-\infty < a < b < \infty$. \square

1.7 Measurable function [33,8]

Let $X : \Omega \rightarrow \mathbb{R}$ a function. Then X is said to be measurable (F-measurable) on Ω if $\{w \in \Omega \mid X(w) > a\}$ (1.5) is measurable (as a set) for every real number a .

And also we say that X is a measurable function; Suppose that (Ω, \mathcal{F}) and $(\Omega^*, \mathcal{F}^*)$ are measurable spaces. Then $X : \Omega \rightarrow \Omega^*$ is a measurable function to \mathcal{F} and \mathcal{F}^* if $X^{-1}(A) \in \mathcal{F}$, for all $A \in \mathcal{F}^*$ such that

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\} = \{X \in A\} \dots\dots\dots(1.6)$$

Remark(1.7.1)[2,6]

1. A continuous function on a measurable set is a measurable function.
2. Any interval the inverse of interval is a measurable set .
3. Every monotone function is a measurable function.

Note

The measurable function equal to Borel measurable function If we replace measurable set by Borel set .

Theorem(1.7.2)[4]

If (Ω, \mathcal{F}) be measurable space and $X : \Omega \rightarrow \mathbb{R}$ is measurable function, then the following conditions are equivalent .

- I. $\{\omega \in \Omega \mid X(\omega) < a\}$ for all $a \in \mathbb{R}$;
- II. $\{\omega \in \Omega \mid X(\omega) > a\}$ for all $a \in \mathbb{R}$;
- III. $\{\omega \in \Omega \mid X(\omega) \leq a\}$ for all $a \in \mathbb{R}$;
- IV. $\{\omega \in \Omega \mid X(\omega) \geq a\}$ for all $a \in \mathbb{R}$.

Theorem (1.7.3) [1]

If X and Y are measurable functions, so are $X+Y, XY, aX, \max(X, Y)$ and $\min(X, Y)$.

Proof

Step (1)

If $X(\omega) + Y(\omega) < a$, then $X(\omega) < a - Y(\omega)$, and there exists a rational r such the $X(\omega) < r < a - Y(\omega)$. So

$$\{w: X(w)+Y(w) < r\} = \bigcup_{r \text{ rational}} (\{w: X(w) < r\} \cap \{w: Y(w) < a-r\}) .$$

Step (2)

X^2 is measurable since $\{x: X^2 > a\} = \{x: X > \sqrt{a}\} \cup \{x: X < -\sqrt{a}\}$. The measurability of XY following since $XY = \frac{1}{2}[(X+Y)^2 - X^2 - Y^2]$. \square

In the same way we can prove that $aX, XY, \max(X, Y)$ and $\min(X, Y)$ are all these measurable functions .

Theorem (1.7.4) [3]

Let $\{X_n\}$ be a sequence of a measurable functions on (Ω, F) , then $\sup_n X_n, \inf_n X_n, \limsup_n X_n, \liminf_n X_n$, are measurable functions.

Lemma (1.7.5) [3,31]

If each function in sequence $\{X_n\}$ is measurable on a set Ω and if X is the point wise limit function of $\{X_n\}$, then X is measurable on Ω .

Proof

Let $w \in \Omega$ and $a \in \mathbb{R}$ such that $X(w) > a$. Let k be a natural number such that $X(w) > a + \frac{1}{k}$, Then by definition of limit, there exists a natural number m , such that $X_n(w) > a + \frac{1}{k}$ for all $n > m$. Thus

$$X(w) = \lim_{n \rightarrow \infty} X_n(w) > a + \frac{1}{k} > a.$$

This implies that $\{w: X(w) > a\} = \bigcup_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m+1}^{\infty} \{w: X_n(w) > a + \frac{1}{k}\}$.

Since this set is measurable and a was arbitrary, it follows that X is measurable. \square

Theorem (1.7.6) [4,31]

Suppose X is a measurable function, then all the following are measurable functions.

I. $X^n, n \geq 1$

II. $|X|^n, n > 0$

III. $X^+ = \max(X, 0)$

IV. $X^- = -\min(X, 0)$.

Example (1.7.7)

Let $X, Y : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be two measurable functions, where (Ω, \mathcal{F}) is a measurable space. Let X^*, Y^* , be defined by:

$$X^*(\omega) = X(\omega) \text{ if } X(\omega) \notin \{-\infty, \infty\} \text{ and } X^*(\omega) = 1 \text{ if } X(\omega) \in \{-\infty, \infty\}.$$

$Y^*(\omega)$ being defined in a similar way. Consider the partitions of Ω ,

$$\Omega = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \text{ and } \Omega = B_1 \cup B_2 \cup B_3 \cup B_4 \cup B_5,$$

where $A_1 = \{X \in (0, \infty)\}, A_2 = \{X \in (-\infty, 0)\}, A_3 = \{X = 0\},$

$A_4 = \{X = -\infty\}, A_5 = \{X = \infty\}$ and B_1, B_2, B_3, B_4, B_5 being defined in a

similar way with Y .

Show that

1. X^* and Y^* are measurable with respect to \mathcal{F} and $\mathcal{B}(\mathbb{R})$.
2. All A_i 's and B_j 's are elements of \mathcal{F} .
3. For all $B \in \mathcal{B}(\mathbb{R})$:

$$\{XY \in B\} = \bigcup_{i,j=1}^5 (A_i \cap B_j \cap \{XY \in B\})$$

4. $A_i \cap B_j \cap \{XY \in B\} = A_i \cap B_j \cap \{X^*Y^* \in B\}$, in the case when $1 \leq i \leq 3$ and $1 \leq j \leq 3$.

5. $A_i \cap B_j \cap \{XY \in B\}$ is either equal to ϕ or $A_i \cap B_j$, in the case when $i \geq 4$ or $j \geq 4$.

Solution

1. For all $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{*-1}(B)$ can be written as:

$$X^{*-1}(B) = (X^{-1}(B) \cap X^{-1}(\mathbb{R})) \cup (A \cap (\{X = +\infty\} \cup \{X = -\infty\})),$$
 where $A = \Omega$ if $1 \in B$, and $A = \phi$ otherwise. It follows that $X^{*-1}(B) \in \mathcal{F}$, and X^* is measurable. Similarly, Y^* is measurable.

2. All A_i 's and B_j 's are inverse images of Borel sets in \mathbb{R} , by measurable functions. They are therefore elements of \mathcal{F} .

3. Since $\Omega = \bigcup_{i,j} A_i \cap B_j$, for all $B \in \mathcal{B}(\mathbb{R})$, we have:

$$\{XY \in B\} = \bigcup_{i,j=1}^5 (A_i \cap B_j \cap \{XY \in B\})$$

4. For all $1 \leq i, j \leq 3$ and $w \in A_i \cap B_j$, $X(w) \in \mathbb{R}$ and $Y(w) \in \mathbb{R}$. In particular, $X(w) = X^*(w)$, and $Y(w) = Y^*(w)$. Hence, we conclude that $A_i \cap B_j \cap \{XY \in B\} = A_i \cap B_j \cap \{X^*Y^* \in B\}$.

5. Suppose $i \geq 4$ or $j \geq 4$. Then, for all $w \in A_i \cap B_j$, $X(w)Y(w)$ is either $-\infty$, 0 or $+\infty$. More specifically, $X(w)Y(w) = a$, with:

$a = -\infty$ if $(i, j) \in \{(1, 4), (2, 5), (4, 5), (5, 4), (5, 2), (4, 1)\}$, $a = 0$ if $(i, j) \in \{(3, 4), (3, 5), (4, 3), (5, 3)\}$, $a = \infty$ if $(i, j) \in \{(1, 5), (2, 4), (4, 4), (5, 5), (5, 1), (4, 2)\}$

Hence, given $B \in \mathcal{B}(\mathbb{R})$, $A_i \cap B_j \cap \{XY \in B\} = \emptyset$ if $a \notin B$, and

$$A_i \cap B_j \cap \{XY \in B\} = A_i \cap B_j \quad \text{if } a \in B.$$

Example (1.7.8)

Let $\Omega = \mathbb{N}$, and let $B = \{i \in \mathbb{N} : i = 2n \text{ for some } n \in \mathbb{N}\}$. Set

$$A = \{A \subseteq \mathbb{N} : \text{either } B \subseteq A \text{ or } A \cap B = \emptyset\}.$$

Show that $g : \mathbb{N} \rightarrow \mathbb{R}$ is \mathcal{F} -measurable if and only if g is constant on the set B .

Solution

Suppose that $g(i) = a$ for all i even, and consider $\{w : g(w) \leq a\}$, for any real a . Clearly this set either contains all of B or has an empty intersection with B ; i.e., it is in A . Therefore g is \mathcal{F} -measurable.

Suppose, on other hand, that $g(i_1) \neq g(i_2)$ for some pair of even integers i_1, i_2 . Without loss of generality, there is some a such that $g(i_1) < a < g(i_2)$. Then, letting $h_a = \{w : g(w) \leq a\}$, we have $h_a \cap B \neq \emptyset$ and $B \not\subseteq h_a$, and so h_a is not in A . Hence g is not \mathcal{F} -measurable.

1.8 Simple function [4,43]

Let (Ω, \mathcal{F}) be a measurable space. We say that the function $\varphi : \Omega \rightarrow \mathbb{R}$ is a *simple function* on (Ω, \mathcal{F}) , if and only if φ is of the form :

$$\varphi(w) = \sum_{i=1}^n a_i 1_{A_i} \quad \dots\dots\dots (1.7)$$

where A_i are disjointed measurable subsets which from Ω , $(\cup A_i = \Omega)$, $n \geq 1$, and $a_i \in \mathbb{R}$ for all $i = 1, \dots, n$.

Remark (1.8.1) [31]

1. Any simple function is a measurable function.
2. Any constant function is a simple function .

Theorem (1.8.2) [4,31]

If X and Y are simple functions , so are $\varphi + \psi$, $a\varphi$, $\varphi\psi$, $\max(\varphi, \psi)$ and $\min(\varphi, \psi)$.

Theorem (1.8.3) [4,43]

Let $X : \Omega \rightarrow [0, \infty]$ be a measurable function . There exists a sequence of simple functions $\{\varphi_n\}$ on Ω such that :

- I. $0 \leq \varphi_1(w) \leq \varphi_2(w) \leq \dots \leq X(w)$
- II. $\varphi_n(w) \rightarrow X(w)$ for each $w \in \Omega$.

Proof

Fix $n \geq 1$ and for $i = 1, 2, 3, \dots, n2^n$, define the measurable sets

$A_{n_i} = X^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}))$. Set $B_n = X^{-1}([n, \infty])$ and define the simple functions

$$\varphi_n(\omega) = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} 1_{A_{n_i}}(\omega) + n 1_{B_n}(\omega).$$

Clearly φ_n is a simple function and it satisfies $\varphi_n(\omega) \leq \varphi_{n+1}(\omega)$ and $\varphi_n \leq X$ for all ω . Fix $\varepsilon > 0$.

Let $\omega \in \Omega$. If $X(\omega) < \infty$, then pick n so large that $2^{-n} < \varepsilon$ and $X(\omega) < n$.

Then $X(\omega) \in [\frac{i-1}{2^n}, \frac{i}{2^n})$ for some $i = 1, 2, \dots, n2^n$. Thus $\varphi_n(\omega) = \frac{i-1}{2^n}$

and so, $X(\omega) - \varphi_n(\omega) < 2^{-n}$. \square

Example (1.8.4)

Let $X : (\Omega, \mathcal{F}) \rightarrow [0, +\infty]$ be a nonnegative and measurable function. For $n \geq 1$, put $A_n = \{X > 1/n\}$, and $\varphi_n = (1/n) 1_{A_n}$. Show that

1. φ_n is a simple function on (Ω, \mathcal{F}) with $\varphi_n \leq X$.
2. $A_n \uparrow \{X > 0\}$.

Solution

X being measurable, $A_n = \{X > 1/n\}$ is an element of the σ -field \mathcal{F} . Since $1/n \in \mathbb{R}$, from definition (2.6) it follows that $\varphi_n = (1/n) 1_{A_n}$ is a simple function on (Ω, \mathcal{F}) . Suppose that $\omega \in \Omega$. If $\omega \notin A_n$, then $\varphi_n(\omega) = 0 \leq X(\omega)$.

If $\omega \in A_n$, then $\varphi_n(\omega) = 1/n < X(\omega)$. In any case, $\varphi_n(\omega) \leq X(\omega)$. It follows that $\varphi_n \leq X$.

Let $n \geq 1$, if $\omega \in A_n$, then $X(\omega) > 1/n$ and in particular $X(\omega) > 1/(n+1)$.

So $\omega \in A_{n+1}$ and we see that $A_n \subseteq A_{n+1}$. For all $n \geq 1$, $A_n \subseteq \{X > 0\}$. It follows

that $\bigcup_{n=1}^{\infty} A_n \subseteq \{X > 0\}$. Conversely, if $X(\omega) > 0$, then there exists $n \geq 1$ such that

$X(\omega) > 1/n$. So $\{X > 0\} \subseteq \bigcup_{n=1}^{\infty} A_n$. We have proved that $A_n \subseteq A_{n+1}$ with

$$\bigcup_{n=1}^{\infty} A_n = \{X > 0\}, \text{ i.e. } A_n \uparrow \{X > 0\}. \quad \square$$

1.9 Lebesgue integral [18,43]

If $\varphi(\omega) = \sum_{i=1}^n a_i 1_{A_i}$ is a nonnegative measurable simple function and $E \in \mathcal{F}$ is measurable, define the Lebesgue integral of $\varphi(\omega)$ over the set E to be

$$\int_E \varphi(\omega) dM = \sum_{i=1}^n a_i M(A_i \cap E) \quad \dots\dots\dots (1.8)$$

Example (1.9.1)

Let $h: [2,5] \rightarrow \mathbb{R}$ be a function defined on $h(\omega) = 2$ if $\omega \in Q$, and $h(\omega) = 3$ if $\omega \in Q'$. Show that h is Lebesgue integral, where Q a rational number and Q' irrational number.

Solution

$$\int_2^3 h dM = \sum_{i=1,2} a_i M(A_i \cap E) = 2M([2,5] \cap Q) + 3M([2,5] \cap Q') =$$

$$2M(\phi) + 3M([2,5]) = 2(0) + 3(5-2) = 9. \quad (M(\phi) = 0)$$

Remarks (1.9.2) [33]

1. If $X \geq 0$ is a measurable function, we define the integral of X over E to be

$$\int_E X dM = \sup \left\{ \int_E \phi(w) dM : 0 \leq \phi(w) \leq X, \phi(w) \text{ simple} \right\} \dots\dots\dots(1.10)$$

2. If X is measurable and at least one of the integrals

$$\int_E X^+ dM, \int_E X^- dM \text{ is finite, where } X^+ = \max(X(w), 0)$$

and $X^- = -\min(X(w), 0)$, define the integral of X over E to be

$$\int_E X dM = \int_E X^+ dM - \int_E X^- dM \dots\dots\dots(1.11)$$

Theorem (1.9.3) [2,6]

Suppose that X, Y are measurable functions, let $E \in F$ and a, b are real constants, then:

I. If $0 \leq X \leq Y$, then $\int_E X dM \leq \int_E Y dM$.

II. If $A \subset B, A, B \in F$, and $X \geq 0$, then $\int_A X dM \leq \int_B X dM$.

III. If $X \geq 0$, then $\int_E (aX + bY) dM = a \int_E X dM + b \int_E Y dM$.

IV. If $X(w) = 0$ for all $w \in E$, then $\int_E X dM = 0$.

V. If $M(E) = 0$, then $\int_E X dM = 0$ for any $X \geq 0$.

VI. If $X \geq 0$, then $\int_E X dM = \int_\Omega \mathbf{1}_E X dM$.

VII. If X is integrable and bounded, then $\left| \int_\Omega X \right| \leq \int_\Omega |X| \leq \sup_w |X(w)|$.

Proof

I. If $\phi(w)$ is simple and $0 \leq \phi(w) \leq X$, then $0 \leq \phi(w) \leq X \leq Y$ and so, by definition of $\int_E Y dM$, we have $\int_E \phi dM \leq \int_E Y dM$.

Taking the supremum over all φ with $0 \leq \varphi \leq X$ gives the

$$\int_E X dM \leq \int_E Y dM$$

II. If $0 \leq \varphi(w) \leq X$, then $\int_A \varphi dM \leq \int_B \varphi dM$ (since $A \subset B$)

Taking the supremum over φ with $0 \leq \varphi \leq X$, gives the inequality

$$\int_A X dM \leq \int_B Y dM .$$

and on the same way we can prove III, IV, V, VI, VII. \square

Theorem (1.9.4)[11]

The integral $M(X)$ assigns a unique value to each X measurable and integrable function. Further,

I. $M(X) = M(Y)$ whenever $M(\{w: X(w) \neq Y(w)\}) = 0$.

II. M is linear, that is for any X, Y are measurable and integrable functions,

$$M(X + Y) = M(X) + M(Y) .$$

Note

$$M(X) = \int_{\Omega} X dM$$

Theorem (1.9.5) [12][18,43]

Let (Ω, F, M) be measure space. Suppose φ and ψ are simple functions with $\varphi \geq 0$ and $\psi \geq 0$.

I. For $E \in F$ define $M^*(E) = \int_E \varphi dM$, then M^* is measure on F .

$$\text{II. } \int_{\Omega} (\varphi + \psi) dM = \int_{\Omega} \varphi dM + \int_{\Omega} \psi dM \text{ , and also}$$

$$\int_{\Omega} (a\varphi + b\psi) dM = a \int_{\Omega} \varphi dM + b \int_{\Omega} \psi dM \text{ , for real constants } a \text{ and } b .$$

$$\text{III. } |\varphi| \text{ is a simple function and } \left| \int_{\Omega} \varphi(w) \right| \leq \int_{\Omega} |\varphi| dM \leq \sup_w |\varphi(w)| \epsilon .$$

Proof

I. $E_j \in F, E = \cup E_j$, then

$$\begin{aligned} M^*(E) &= \int_E \varphi dM = \sum_{i=1}^n a_i M(A_i \cap E) = \sum_{i=1}^n a_i \sum_{j=1}^{\infty} M(A_i \cap E_j) \\ &= \sum_{j=1}^{\infty} \sum_{i=1}^n a_i M(A_i \cap E_j) = \sum_{j=1}^{\infty} M^*(E_j) . \end{aligned}$$

II. Let us write $\varphi = \sum_{i=0}^n a_i 1_{A_i}$ and $\psi = \sum_{j=0}^m b_j 1_{B_j}$ where $a_0 = 0 = b_0$, the a_i s are

distinct and the b_j s are distinct so the $\{A_0, \dots, A_n\}$ is a partition of Ω and similarly, $\{B_0, \dots, B_m\}$ is a partition of Ω . The sum $\varphi + \psi$ is a simple

function, say, $\varphi + \psi = \sum_{l=0}^k d_l 1_{C_l}$, where $d_0 = 0$, the d_l s are distinct and

$\{C_0, \dots, C_k\}$ is a partition of Ω . Each d_l has the form $d_l = a_i + b_j$ for suitable

i, j and so $C_l = \bigcup_{I_l} A_i \cap B_j$ where $I_l = \{(i, j) : a_i + b_j = d_l\}$. Then

$$\int_{\Omega} (\varphi + \psi) dM = \sum_{l=0}^k d_l M(C_l)$$

and

$$\int_{\Omega} \varphi dM + \int_{\Omega} \psi dM = \sum_{i=0}^n a_i M(A_i) + \sum_{j=0}^m b_j M(B_j) .$$

The collection $\{A_i \cap B_j : 0 \leq i \leq n, 0 \leq j \leq m\}$ forms a partition of Ω

we have

$\int_{\Omega} (\varphi + \psi) dM = \sum_{l=0}^k d_l M(C_l) = \sum_{l=0}^k d_l M\left(\bigcup_{I_l} A_i \cap B_j\right) = \sum_{l=0}^k d_l \sum_{I_l} M(A_i \cap B_j)$, since distinct $A_i \cap B_j$ s are disjoint,

$$\begin{aligned} &= \sum_{l=0}^k \sum_{I_l} (a_i + b_j) M(A_i \cap B_j) \\ &= \sum_{i,j} (a_i + b_j) M(A_i \cap B_j) = \sum_{i,j} a_i M(A_i \cap B_j) + \sum_{i,j} b_j M(A_i \cap B_j) \\ &= \sum_i a_i M(A_i) + \sum_j b_j M(B_j) = \int_{\Omega} \varphi dM + \int_{\Omega} \psi dM \quad \square \end{aligned}$$

In the same way we can prove $\int_{\Omega} (a\varphi + b\psi) dM = a \int_{\Omega} \varphi dM + b \int_{\Omega} \psi dM$ and (III).

Remark (1.9.6) [24]

We see that the integral of a simple function is got by adding up its “elementary bits” and it does not matter how we write the simple function as such a sum. For example, consider $\varphi = a1_A$ so that $\int_{\Omega} \varphi dM = aM(A)$. If $A = B_1 \cup B_2$ with $B_1 \cap B_2 = \emptyset$, then we could also write φ as

$\varphi = a1_{B_1} + a1_{B_2}$. We have

$$\int_{\Omega} a1_A dM = \int_{\Omega} a1_{B_1} dM + \int_{\Omega} a1_{B_2} dM = aM(B_1) + aM(B_2) = aM(A) \dots (2.14)$$

Definition (L^p Spaces) (1.9.7)[4]

For $1 \leq p < \infty$, define the L^p norm of X by $\|X\|_p = \left(\int_{\Omega} |X|^p dM\right)^{\frac{1}{p}}$.

For $p = \infty$, define the L^∞ norm of X by

$$\|X\|_{\infty} = \inf\{K \in \mathbb{R} : M(\{w : |X(w)| \geq K\}) = 0\}, \text{ for } 1 \leq p \leq \infty,$$

the space L^p is the set $\{X : \|X\|_p < \infty\}$.

In this chapter we give some definitions, theorems which are related to probability measure such as random variables, distribution function, independent, expectation, etc.

There are many reasons why should such a student learn about probability measure theory.

The first reason is that the extension of probability to measure theory is fairly natural and intuitive since probabilities already measure sets .

Another reason is that using measure theory nicely generalizes probability theory. This allows continuous probabilities to be treated within the same framework (whereas without measure theory many concepts and theorems need to be defined and proved for continuous probabilities).

A third and final reason is that using measure theory allows one to use the analytic tools of measure theory .

2.1 Some Basic Definitions of Probability

Definition (*A random Experiment*) (2.1.1)[36]

A random experiment is the experiment whose results are not predictable and can be determined only after performing it and then observing the outcome.

Definition (*sample space*) (2.1.2)

Is a space Ω of all possible outcomes of a random experiment and each individual outcome is represented as a point w in that space Ω . Subsets of Ω are called event and each of them corresponds to a collection of outcomes .

2.2 Probability measure [9,14]

Let (Ω, \mathcal{F}) be a measurable space. A function $\Pi : \mathcal{F} \rightarrow [0,1]$ is called *probability measure* if it satisfies the following conditions:

I. $0 \leq P \leq 1$ for all $A \in \mathcal{F}$.

II. $P(\Omega) = 1$

III. If $A_1, A_2, \dots \in \mathcal{F}$ is a countable sequence of disjoint sets

$$A_n \cap A_k = \phi \text{ for } n \neq k, \text{ then } P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

A triple (Ω, \mathcal{F}, P) , where Ω is a nonempty set, \mathcal{F} is a σ -field of subsets of Ω , and P is a probability measure on (Ω, \mathcal{F}) is called

Probability measure space and if Ω equal any interval $(0,1), [0,1), \dots$ and $\mathcal{F} = \mathcal{B}$, then P_L is called *the Lebesgue probability measure*.

Some Basic Properties (2.2.1) [11,35]

We try to explain some properties of probability measure which are similar to those as in probability elementary theory:

I. If $A, B \in \mathcal{F}$, $A \subset B$, then $P(A) \leq P(B)$ (*Monotonicity*)

II. If $A, B \in \mathcal{F}$, $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$ (*Additivity*)

III. $P(A) + P(B) = P(A \cup B) + P(A \cap B) = P(A \setminus B) + 2P(A \cap B) + P(B \setminus A)$,

then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$\text{IV. } P\left(\bigcup_{k=1}^n A_k\right) = \sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) + \dots +$$

$$(-1)^{n+1} P(A_1 \cap \dots \cap A_n) \quad (\text{Inclusion-exclusion formula})$$

V. If A_1, \dots, A_n are measurable sets, not necessarily disjointed, then

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i) \quad (\text{Sub additivity})$$

VI. $\lim_{n \rightarrow \infty} P(A_n) = P(\lim_{n \rightarrow \infty} (A_n))$

VII. $P(\limsup_{n \rightarrow \infty} A_n) \geq \limsup_{n \rightarrow \infty} P(A_n)$

VIII. $P(\liminf_{n \rightarrow \infty} A_n) \leq \liminf_{n \rightarrow \infty} P(A_n)$

IX. If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\limsup_{n \rightarrow \infty} A_n) = 0$

IX. If A_1, A_2, \dots are assumed to be independent and

$$\sum_{n=1}^{\infty} P(A_n) = \infty, \text{ then } P(\limsup_{n \rightarrow \infty} A_n) = 1.$$

Theorem (2.2.2) [9,35]

Let (Ω, \mathcal{F}, P) be a probability measure space. Then the following assertions are true:

I. If $B \in \mathcal{F}$, then $P(B^c) = 1 - P(B)$

II. If $A, B \in \mathcal{F}$, then $P(A \setminus B) = P(A) - P(A \cap B)$

III. If $A_1, \dots, A_n \in \mathcal{F}$ such that $A_i \cap A_j = \phi$, for $i \neq j$, then

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \quad (\text{finite additivity})$$

IV. If $A_1, A_2, \dots \in \mathcal{F}$ such that $A_1 \subseteq A_2 \subseteq \dots$, then

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right) \quad (\text{continuity from below})$$

V. If $A_1, A_2, \dots \in \mathcal{F}$ such that $A_1 \supseteq A_2 \supseteq \dots$, then

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) \quad (\text{continuity from above})$$

VI. If $A_1, A_2, \dots \in \mathcal{F}$ and $A \subseteq \bigcup_{n=1}^{\infty} A_n$, then $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$.

Theorem (2.2.3) [4]

Let (Ω, \mathcal{F}, P) be a probability measure space, there exists complete probability measure space $(\Omega, \mathcal{F}^*, P^*)$ such that

I. $\mathcal{F} \subset \mathcal{F}^*$

II. If $A \in \mathcal{F}^*$, then $P(A) = P^*(A)$

III. If $A \in \mathcal{F}^*$, then $A = A_1 \cup A_2$ such that $A_2 \in \mathcal{F}$, $A_1 \subset B$,
 $B \in \mathcal{F}$, $P(B) = 0$

2.3 Random Variables [30,38]

A random variable X is a measurable function, $X: \Omega \rightarrow \mathbb{R}$.i.e. A real valued function $X(\omega)$ on Ω such that for every Borel set $B \in \mathcal{B}$, then $X^{-1}(B) = \{\omega: X(\omega) \in B\}$ is a random variable subset of Ω , $X^{-1}(B) \in \mathcal{F}$.

There are two types of random variables, *discrete* and *continuous*.

A *continuous random variable* is one which takes an infinite number of possible values. Continuous random variables are usually measurement. Examples include height, weight, etc.

The random variable has some properties such that for all X and Y are two random variables $\omega \in \Omega$ then :

I. $(aX + bY)(\omega) = aX(\omega) + bY(\omega)$ is a random variable $a, b \in \mathbb{R}$.

II. $(XY)(\omega) = X(\omega)Y(\omega)$ is a random variable .

III. If $Y(\omega) \neq 0$, then $(\frac{X}{Y})(\omega) = \frac{X(\omega)}{Y(\omega)}$ is a random variable .

IV. $|X(\omega)|$ is a random variable .

Example (2.3.1)

Let $\Omega = [0,1)$. Let $\mathcal{F} = \mathcal{B}$, the collection of all Borel subsets of $[0,1)$. Let $P = P_L$, the Lebesgue probability measure on $([0,1), \mathcal{B})$. For every positive integer n , let $X_n : [0,1) \rightarrow \mathbb{R}$ be defined by

$$X_n(\omega) = \begin{cases} 0, & \text{if } \omega < \frac{1}{2^n} \\ \frac{1}{2^n}, & \text{if } \frac{1}{2^n} \leq \omega < \frac{2}{2^n} \\ \frac{2}{2^n}, & \text{if } \frac{2}{2^n} \leq \omega < \frac{3}{2^n} \\ \vdots & \vdots \\ \frac{2^n-1}{2^n}, & \text{if } \frac{2^n-1}{2^n} \leq \omega < 1. \end{cases}$$

Show that each X_n is a random variable on $([0,1), \mathcal{B}, P_L)$. Describe precisely the σ -field $\mathcal{A}_n = \sigma(X_n)$. Let $\mathcal{A} = \bigcup_n \mathcal{A}_n$. prove that $\sigma(\mathcal{A}) = \mathcal{B}$.

Solution

For a given $n, k \in \{0, 1, \dots, n-1\}$ and $k/n \leq x < (k+1)/n$,
 $\{\omega : X_n(\omega) \leq x\} = [0, (k+1)/n) \in \mathcal{B}$,
 since \mathcal{B} contains all intervals. So X_n is a random variable. The “building blocks” of \mathcal{A}_n are the disjointed intervals $[k/n, (k+1)/n)$,
 $k = 0, \dots, n-1$. Then \mathcal{A}_n consists of all possible unions of these building blocks.

Clearly, each $\mathcal{A}_n \subseteq \mathcal{B}$, so also $\mathcal{A} \subseteq \mathcal{B}$. Then $\sigma(\mathcal{A}) \subseteq \mathcal{B}$, as $\sigma(\mathcal{A})$ is the

smallest σ -field containing A .

Since every real number can be approximated arbitrarily well by rationales,

given any $a \in [0, 1)$ we can write $[0, a) = \bigcup_m [0, k_m / n_m)$

for suitable n_m and $0 \leq k_m \leq n_m$. (Here $q_m = k_m / n_m$ is a sequence of rationales converging to a as $m \rightarrow \infty$.)

Hence also $[b, a) = [0, a) \setminus [0, b) \in \sigma(A)$ for $b < a$. Hence $B \subseteq \sigma(A)$.

Theorem (2.3.2) [35]

Let (Ω, \mathcal{F}, P) be a probability measure space. A function $X : \Omega \rightarrow \mathbb{R}$ satisfying $P(\{\omega \in \Omega; X(\omega) \in \{-\infty, \infty\}\}) = 0$ is a random variable

if and only if for every open interval (a, b) , $X^{-1}(a, b) \in \mathcal{F}$.

Proof

Since every open interval is a Borel set, if X is a random variable, then $X^{-1}((a, b)) \in \mathcal{F}$.

Now suppose that $X^{-1}(a, b) \in \mathcal{F}$ for every open interval (a, b) .

Consider $\mathcal{F}' = \{B \in \mathcal{F} : X^{-1}(B) \in \mathcal{F}\}$. We claim that \mathcal{F}' is a σ -field.

Note that $X^{-1}(\emptyset) = \emptyset \in \mathcal{F}$, so $\emptyset \in \mathcal{F}'$, $X^{-1}(\mathbb{R}) = \Omega \setminus X^{-1}(\{-\infty, \infty\})$

since

$P(X^{-1}(\{-\infty, \infty\}))$ is zero, $X^{-1}(\{-\infty, \infty\}) \in \mathcal{F}$;

Since \mathcal{F} is a σ -field, it is closed under complements, so $X^{-1}(\mathbb{R}) \in \mathcal{F}$, and $\mathbb{R} \in \mathcal{F}'$.

If $B \in \mathcal{F}'$, $X^{-1}(B) \in \mathcal{F}$, so $X^{-1}(\mathbb{R} \setminus B) = X^{-1}(\mathbb{R}) \setminus X^{-1}(B) \in \mathcal{F}$

because $X^{-1}(\square) \in \mathcal{F}$ and $X^{-1}(B) \in \mathcal{F}$.

Therefore, $P \setminus B \in \mathcal{F}'$, so \mathcal{F}' is closed under complements.

Now

Suppose $B_n \in \mathcal{F}'$, $n \in \mathbb{N}$. Then $X^{-1}(B_n) \in \mathcal{F}$ for all n

So $X^{-1}\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \bigcup_{n \in \mathbb{N}} X^{-1}(B_n) \in \mathcal{F}$ and hence \mathcal{F}' is closed under countable union.

Therefore, \mathcal{F}' is a σ -field, since every open set is a countable union of open intervals, \mathcal{F}' contains every open set;

Since the Borel σ -field is the smallest σ -field containing the open sets in \square , \mathcal{F}' contains every Borel set, so X is measurable and hence a random variable. \square

2.4 Law of a Random Variable [30]

Let (Ω, \mathcal{F}, P) be a probability measure space and $X: \Omega \rightarrow \square$ be a random variable, then

$P_X(B) = P(\omega \in \Omega : X(\omega) \in B)$ is called *the law or image probability measure* of the random variable X .

Definition (2.4.1) [28]

Let (Ω, \mathcal{F}) be measurable space, P and Q are two Probability measures defined on that space. The probability measure Q is said to be *absolutely continuous* with respect to probability measure P if for any $A \in \mathcal{F}$ so that $P(A)=0$ one has $Q(A) = 0$.

Definition (*Probability Density Function*) (2.4.2) [1]

The probability density function (p.d.f), is $f(x)$, of continuous random variable X , that is mean $f: \mathbb{R} \rightarrow [0, \infty)$

The p.d.f must satisfy the following two conditions:

I. $f(x) \geq 0$, for all x and

II. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Definition (Distribution Function) (2.4.3) [35,43]

Given a random variable $X: \Omega \rightarrow \mathbb{R}$, The (Cumulative) distribution function of X is $F_X(x): \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$F_X(x) = P_X(-\infty, x] = P(\{\omega: X(\omega) \leq x\})$$

The distribution P_X can be obtained from the distribution function by setting $P_X(-\infty, x] = F_X(x)$; that is uniquely determines P_X .

If X is a random variable, its cumulative distribution function $F_X(x)$ satisfies the following properties.

I. $0 \leq F_X(x) \leq 1$.

II. $F_X(x)$ is nondecreasing

III. $F_X(x)$ is right continuous :

that is

$$F_X(x) = \lim_{h \downarrow 0} F_X(x+h) \text{ for all } x \in \mathbb{R}$$

IV. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ and $\lim_{x \rightarrow \infty} F_X(x) = 1$

V. $f(x) = \frac{d}{dx} F(x)$

IV. $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) = \int_a^b f(x)dx$

for $a < b$

VI. A function F is the distribution function of some probability measure if and only if one has (II) (III) (IV).

VII. If F satisfies (II) (III) (IV), then there exists a unique probability measure P that has

$$P(a, b] = F(b) - F(a) \text{ for all } a, b \in \mathbb{R}.$$

VIII. $F_X(x)$ is continuous if and only if $P(X = x) = 0$.

IX. $F_X(x)$ is continuous if and only if P_X is continuous.

IX. $F_X(x)$ has at most countable many points of discontinuity.

Example (2.4.4)

Let (Ω, \mathcal{F}, P) be a probability measure space and X a random variable on (Ω, \mathcal{F}, P) .

1. Show that $F_X(x) = \lim_{\varepsilon \downarrow 0} F_X(x + \varepsilon)$.

2. Show that $\lim_{\varepsilon \downarrow 0} F_X(x - \varepsilon) \leq F_X(x)$.

3. Let $m = \sup\{x: F_X(x) < 1/2\}$. Show that $\lim_{\varepsilon \downarrow 0} F_X(m - \varepsilon) \leq 1/2 \leq F_X(m)$.

Solution

1. Define $A_n = \{w: X(w) \leq x + 1/n\}$ $n \in \mathbb{N}$, so the sets A_n form a decreasing nested sequence. Also, $\bigcap_n A_n = \{x: X(w) \leq x\}$. By A_n^c which form an

increasing nested sequence, we get $F_X(x) = \lim_{n \rightarrow \infty} F_X(x + 1/n)$. Since F_X is

nondecreasing, we deduce the result, where $\varepsilon = \frac{1}{n}$.

2. Define $A_n = \{w : X(w) \leq x - 1/n\}$, so the sets A_n form an increasing nested sequence. Also, $A = \bigcup_n A_n = \{w : X(w) < x\}$. We get $F_X(x) \geq P(A_n) = \lim_{n \rightarrow \infty} F_X(x - 1/n)$. Since F_X is nondecreasing, we deduce the result.

3. Let $m = \sup\{x : F_X(x) < 1/2\}$. Then $F_X(m) < 1/2$. Also, $F_X(m) \geq 1/2$; otherwise, if $F_X(m) < 1/2$, then by right continuity, $F_X(m') < 1/2$ for some $m' > m$, which is a contradiction.

Theorem (2.4.5) [4,34]

Suppose F is a distribution function. Then there is a probability measure space $(\Omega, \mathcal{F}, \mathbf{P})$ and a random variable X defined on this space such that $F = F_X$

Proof

We take $(\Omega, \mathcal{F}, \mathbf{P})$ with $\Omega = (0,1)$, $\mathcal{F} =$ Borel sets and \mathbf{P} Lebesgue probability measure. For each $w \in \Omega$. Define

$$X(w) = \sup\{y : F(y) < w\}.$$

We claim this is the desired random variable. Suppose we can show that for each $x \in \mathbb{R}$.

$$\{w \in \Omega : X(w) \leq x\} = \{w \in \Omega : w \leq F(x)\}. \quad \dots\dots\dots(2.1)$$

Clearly then X is measurable and also $\mathbf{P}\{X(w) \leq x\} = F(x)$,

proving that $F = F_X$. To prove (3.1)

$$\text{let } w_0 \in \{w \in \Omega : w \leq F(x)\}$$

That is, $w_0 \leq F(x)$. Then $x \notin \{y : F(y) < w_0\}$ and therefore

$$X(w_0) \leq x. \text{ Thus } \{w \in \Omega : w \leq F(x)\} \subset \{w \in \Omega : X(w) \leq x\}.$$

On the other hand, suppose $w_0 > F(x)$. Since F is right continuous, there exists $\varepsilon > 0$ such that $F(x + \varepsilon) < w_0$. Hence $X(w) \geq x + \varepsilon > x$. This shows that $w_0 \notin \{w \in \Omega : X(w) \leq x\}$ and concludes proof. \square

Theorem (2.4.6) [2,4]

For any given distribution function F satisfying the properties of F there exists probability measure Π on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ satisfying

$F(x) = \Pi((-\infty, x])$ for all $x \in \mathbb{R}$. Moreover, such Π is unique.

Theorem (2.4.7) [43]

If F is a distribution function corresponding to Π , then

- I. $\Pi((-\infty, x)) = F(x^-)$.
- II. $\Pi((x, y]) = F(y) - F(x^+)$, $x < y$.
- III. $\Pi((x, y)) = F(y^-) - F(x^+)$, $x < y$.
- IV. $\Pi([x, y)) = F(y^-) - F(x^-)$, $x < y$.
- V. $\Pi([x, y]) = F(y) - F(x^-)$, $x < y$.

Theorem (2.4.8) [30]

Assume that Π and Π^* are probability measures on $\mathcal{B}(\mathbb{R})$ and F and F^* are the corresponding distribution functions. Then the following assertions are equivalent:

- I. $\Pi = \Pi^*$.
- II. $F(x) = \Pi((-\infty, x]) = \Pi^*((-\infty, x]) = F^*(x)$, for all $x \in \mathbb{R}$.

2.5 Some Special Continuous Distributions [19, 20]

There are many special continuous distributions like, Uniform distribution, Normal distribution, Exponential distribution, Gamma distribution, Cauchy distribution, Beta distribution, Double Exponential (Laplace) distribution, Pareto distribution, Wald distribution, Weibull distribution, Students distribution. This thesis will explain some of the following distributions which we need in chapter four.

a. Normal Distribution

Normal distribution has the following formula

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0 \quad \dots\dots\dots (2.2)$$

with mean μ and variance σ^2 , μ and σ are parameters. The distribution

function is $F(x) = \frac{1}{2} + \frac{1}{\pi} \left[\frac{x}{3\sqrt{2}} + \sum_{n=1}^{\infty} \frac{e^{-\frac{n^2}{9}}}{n} \sin \frac{nx\sqrt{2}}{3} \right]$ where μ, σ are real

constants.

b. Exponential Distribution

Exponential distribution has the following formula

$$f(x) = \lambda e^{-\lambda x} \quad x \geq 0 \quad \lambda > 0 \quad \dots\dots\dots (2.3)$$

with means $\frac{1}{\lambda}$ and variance $\frac{1}{\lambda^2}$, λ be parameter. The distribution function is

$$F(x) = 1 - e^{-\lambda x} \quad x > 0 \quad \dots\dots\dots (2.4)$$

where λ is a real constant.

c. Cauchy Distribution

Cauchy distribution has the following formula

$$f(x) = \frac{1}{\pi\beta[1+(\frac{x-\alpha}{\beta})^2]} \quad -\infty < x < \infty, -\infty < \alpha < \infty, \beta > 0 \quad \dots\dots\dots (2.5)$$

the mean and the variance of this distribution do not exist, α and β are parameters .But

when $\alpha = 0$ and $\beta = 1$, the standard Cauchy has formula $\frac{1}{\pi(1+x^2)}$. The

distribution function is $F(x) = \frac{1}{\pi} \tan^{-1}(\frac{x-\alpha}{\beta}) + \frac{1}{2} \quad \dots\dots\dots (2.6)$

where α and β are real constants .

2.6 Independence [4,37]

I. The collection $\{F_i, 1 \leq i \leq n\}$ of σ - fields is said to be *independent* if

whenever $A_1 \in F_1, A_2 \in F_2, \dots\dots, A_n \in F_n$, then $P(\bigcap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i)$

II. A collection $\{X_i : 1 \leq i \leq n\}$ of random variables is said to be (*totally*)

independent if for any $\{B_i; 1 \leq i \leq n\}$ of Borel set in \mathbb{R}^n ,

$$P\{X_1 \in B_1, X_2 \in B_2, \dots\dots, X_n \in B_n\} = P\{\bigcap_{i=1}^n (X_i \in B_i)\} = \prod_{i=1}^n P\{X_i \in B_i\} .$$

III. The collection of measurable subsets $\{A_i, 1 \leq i \leq n\}$ in a σ - filed F is

independent if for any subset $I \subset \{1, 2, \dots\dots, n\}$ we have

$$P\{\bigcap_{i \in I} A_i\} = \prod_{i \in I} P\{A_i\} .$$

Lemma (2.6.1)[37]

Suppose X, Y are r.vs on probability measure space (Ω, \mathcal{F}, P) , with induced distribution α, β on \square respectively. Then X and Y are independent if and only if the distribution induced \square^2 by (X, Y) is $\alpha \times \beta$.

Proof

Let $A, B \in \mathcal{F}$, the distribution of X is $P_X(A) = P(X \in A) = \alpha$ and the distribution of Y is $P_Y(A) = P(Y \in A) = \beta$ then the distribution of X and Y on \square^2 is $P_{X,Y}(A, B) = P(x \in A, y \in B)$. Since r.vs X and Y are independent, then $P(x \in A) \cdot P(y \in B) = \alpha \times \beta$. By the same way we can prove the converse. \square

Definition (2.6.2) [4]

Two random variables $X: \Omega \rightarrow \square$ and $Y: \Omega \rightarrow \square$. Define $X, Y: \Omega \rightarrow \square^2$ the joint distribution by

$$P^2(A) = P[(X, Y) \in A], \quad A \in \mathcal{B}(\square^2)$$

P^2 is then Borle probability measure on (\square^2, \mathcal{B}) .

Mathematical Expected (2.6.3) [17]

Given probability measure space (Ω, \mathcal{F}, P) and a random variable $X: \Omega \rightarrow \square$, we define the expected by

$$E(X) = \int_{\Omega} X dP = \int_{\Omega} X(\omega) dP(\omega)$$

2.7 Joint Distributed of r.vs [25]

Given a random experiment with sample space Ω . Two random variables X and Y are said to be Jointly distributed if they are defined on the same probability sample space.

The sample points consists of ordered pairs (x,y) . The range of each r.v is a set of real numbers ,thus .

$$(X,Y): \Omega \rightarrow \square_1 \times \square_2 \text{ such that ; } (X,Y)(\omega) = (x,y)$$

Joint probability density function (j.p.d.f) [2.7.1] [13,25]

Let X and Y are two continuous random variables defined the same sample space Ω .The joint p.d.f of X and Y (or j.pdf) is denoted $f(x,y)$.

The j.pdf must satisfy :

$$\text{I. } f(x,y) \geq 0, \text{ for all } x \text{ and } y$$

$$\text{II. } \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$$

The marginal density function of X and Y are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy \quad \dots\dots\dots (2.7)$$

where $f_X(x)$ is the marginal density function of the r.v X .Similarly, the marginal density function of Y , $f_Y(y)$ is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y) dx \quad \dots\dots\dots (2.8)$$

Joint Distribution Function (2.7.2) [34]

The joint cumulative distribution function (J.c.d.f), or Joint simply the distribution function F of the two random variables X and Y is the function form \mathbb{R}^2 to $[0,1]$ and is defined for all x and $y \in \mathbb{R}^2$, by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dy dx \dots\dots\dots(2.6)$$

In other word, $F(x, y)$ denotes the probability that the two random variable X and Y takes on a value that is less than or equal.

We can show that the Joint continuous distribution function have the following prosperities

I. $0 \leq F(x, y) \leq 1$, $-\infty < x, y < \infty$

II. $F(-\infty, y) = \lim_{x \rightarrow -\infty} F(x, y) = 0$ for all y

I. $F(x, -\infty) = \lim_{y \rightarrow -\infty} F(x, y) = 0$

II. $F(\infty, \infty) = \lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F(x, y) = 1$

III. $f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$

IV. If $x_1 < x_2$ and $y_1 < y_2$, then $P(x_1 < X \leq x_2, y_1 < Y \leq y_2)$
 $= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$

V. $F(x, y)$ is right continuous in each argument; that is,

$$\lim_{0 < h \rightarrow 0} F(x + h, y) = \lim_{0 < h \rightarrow 0} F(x, y + h) = F(x, y)$$

The marginal distribution function of X and Y are given by

$$\begin{aligned}
 F_X(x) &= \lim_{y \rightarrow \infty} P(X \leq x, Y \leq y) = P(X \leq x) = F(x, \infty) \\
 &= \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dy dx \quad \dots\dots\dots(2.10)
 \end{aligned}$$

$$\begin{aligned}
 F_Y(y) &= \lim_{x \rightarrow \infty} P(X \leq x, Y \leq y) = P(Y \leq y) = F(\infty, y) \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^y f(x, y) dy dx \quad \dots\dots\dots(2.11)
 \end{aligned}$$

where F_X and F_Y are the marginal distribution of X and Y , respectively

Definition (2.7.3) [1,13]

(The Joint Probability density function for independent r.vs)

We say that two random variables (X, Y) are independent if $P(X \in A \text{ and } Y \in B) = P(X \in A)P(Y \in B)$.

The joint distribution r.vs X and Y are said to be independent if and only if their Joint distribution function $F(x, y)$ can be written as the product of the marginal function $F_X(x)$ and $F_Y(y)$

i.e

$$F(x, y) = F_X(x) \cdot F_Y(y) \quad \dots\dots\dots(2.12)$$

If X and Y are continuous random variables, with joint p.d.f $f(x, y)$, then X and Y are independent if and only if $f(x, y) = f_X(x) f_Y(y)$

Corollary (2.7.4)[11]

Suppose that r.v. X with a Borel measurable probability density function f and r.v. Y with a Borel measurable probability density function g are independent . Then ,the random variable $W=X+Y$ has the probability density function

$$f (w) = \int_{\mathbf{R}} f (w - y) f (y) dy$$

Proof

Fixing $w \in \mathbf{R}$, $h(x, y) = 1_{(x, y \leq w)}$, to get that

$$F(w) = P(X + Y \leq w) = Eh(X, Y) = \int_{\mathbf{R}} [\int_{\mathbf{R}} h(X, Y) dP(x)] dP(y).$$

Considering the inner integral for a fixed value of y , we have that

$$\int_{\mathbf{R}} h(X, Y) dP(x) = \int_{\mathbf{R}} 1_{(-\infty, w-y)}(x) dP(x) = P((-\infty, w-y))$$

$$= \int_{-\infty}^{w-y} f(x) dx,$$

Where the right most equality is by the existence of a density $f(x)$ for X .

Clearly , $\int_{-\infty}^{w-y} f(x) dx = \int_{-\infty}^w f(x-y) dx$. Thus we see that

$$F(w) = \int_{\mathbf{R}} [\int_{-\infty}^w f(x-y) dP(x)] dP(y) = \int_{-\infty}^w [\int_{\mathbf{R}} f(x-y) dP(y)] dP(x)$$

Since this applies for any $w \in \mathbf{R}$, it follows by definition that W has the probability density

$$f (w) = \int_{\mathbf{R}} f (w - y) f (y) dP(y) = E[f (w - y)].$$

With Y having density g .

2.8 The moment generated function (m.g.f) [16]

The moment generated function (or simply m.g.f) of the r.v, is denoted by $M_X(t)$ and if it exists ,is defined as

$$\Phi_P(t) = E[e^{tx}] = \int_{\Omega} e^{tx} dP \quad \dots\dots\dots (2.13)$$

where $t \in \mathbb{R}$.

Remark (2.8.1) [13]

There is a one to one correspondence between distribution functions and moment generating functions when the m.g.f exists.

If the m.g.f of the distribution of a r.v approaches that of the distribution of another r.v , then the distribution of the first r.v approaches that of the second r.v under the same limiting conditions.

2.9 Characteristic function [4,5,37]

Let P be a probability measure on Ω .Then ,the characteristic function of P is

$$\Phi_P(t) = E[e^{itx}] = \int_{\Omega} e^{itx} dP \quad \dots\dots\dots (2.14)$$

where $i = \sqrt{-1}$

If X is a random variable ,the characteristic function of the distribution \mathbb{P} induced by X will sometimes be denoted to $\Phi_X = \int e^{itx} P_X(dx)$. So if X and Y have the same law , they have the same characteristic function .Also , if the law of X has a density , that is , $P_X(dx) = f_X(x)dx$,then $\Phi_X(t) = \int e^{itx} f_X(x)dx$.

Theorems (2.9.1)[5]

Suppose P and Q are two probability measures on \mathbb{R}^n with characteristic functions Φ_1 and Φ_2 respectively. Suppose further that for each $t \in \mathbb{R}^n$, $\Phi_1(t) = \Phi_2(t)$. Then, $P = Q$

Proof

For $t \in \mathbb{R}^n$, we have $\Phi_P(t) = \Phi_Q(t)$, then

$\int_{\mathbb{R}^n} e^{itx} P_X(dx) = \int_{\mathbb{R}^n} e^{itx} Q_X(dx)$, we conclude that

$P_X(dx) = Q_X(dx)$, therefore $P = Q$. \square

Theorem (2.9.2)[33]

Let P_n, P be a probability measures on \mathbb{R}^n with distribution function F_n, F and characteristic functions Φ_n and Φ . Then, the following are equivalent.

I. $P_n \Rightarrow P$

II. For any bounded continuous function $g: \mathbb{R}^n \rightarrow \mathbb{R}$

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^n} g(x) dP_n = \int_{\mathbb{R}^n} g(x) dP$$

III. For any $t \in \mathbb{R}^n$, $\lim_{n \rightarrow \infty} \Phi_n(t) = \Phi(t)$

Proof

Suppose that (II) to prove (I).

For any bounded continuous function $g(x)$ then

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^n} g(x) dP_n = \int_{\mathbb{R}^n} g(x) dP, \text{ then}$$

we conclude that $dP_n = dP$, then $P_n \Rightarrow P$.

Suppose that (I) to prove (III), we have

$$\lim_{n \rightarrow \infty} \Phi_{P_n}(t) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} e^{itx} dP_n \quad \text{by (I)}$$

$$\text{then } \int_{\mathbb{R}} e^{itx} dP_n = \Phi(t) \quad \square$$

by the same way we can prove (II) .

Theorem (2.9.3)[26]

Suppose P_n is a sequence of probability measures on \mathbb{R} with characteristic functions Φ_n . Suppose that for each $t \in \mathbb{R}$, $\lim_{n \rightarrow \infty} \Phi_n(t) = \Phi(t)$ exists and Φ is continuous at 0 . Then , there is a probability distribution P such that Φ is the characteristic function of P . Further more , $P_n \Rightarrow P$.

Theorem(2.9.4)[5]

If P_n converges weakly to P , then Φ_n converges to Φ uniformly on every finite interval.

2.10 Moments [16]

Suppose X is a random variable .Then ,*the k th moment of X* is EX^k .The *k th absolute moment of X* is $E|X|^k$.

Theorem (2.10.1)[16]

Let X be a random variable. Suppose that the k th moment of X exists. Then, the characteristic function Φ of X is k times continuously differentiable ,

$$\text{and } \Phi^{(k)}(0) = i^k EX^k \quad .$$

2.11 Some Moment Inequalities

a. Markov's Inequality [23]

If X is a random variable that takes only nonnegative values , then for any constant $a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a} \dots\dots\dots (2.15)$$

b. Chebyshev's Inequality [13]

Let X be any random variable with mean μ and variance σ^2 for any $\varepsilon > 0$, we have

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \dots\dots\dots(2.16)$$

or equivalently $P(|X - \mu| < \varepsilon) > 1 - \frac{\sigma^2}{\varepsilon^2} \dots\dots\dots(2.17)$

Remarks [2.11.1][13]

1.The Chebyshev's Inequality can be obtained if we let

$g(x) = (X - \mu)^2$ and $a = \varepsilon^2$ in prove (a) , we get

$$P\{(X - \mu)^2 \geq \varepsilon^2\} \leq \frac{E(X - \mu)^2}{\varepsilon^2} = \frac{\text{var}(X)}{\varepsilon^2} \text{ hence}$$

$$P(|X - \mu| \geq \varepsilon) < \frac{\sigma^2}{\varepsilon^2}.$$

2.Since $P\{|X - \mu| \geq \varepsilon\} + P\{|X - \mu| < \varepsilon\} = 1$, then by applying the equation (2.16) we get the equation (2.17) .

3. The Chebyshev's Inequality can also be written as $P\{|X - \mu| \geq a\sigma\} \leq \frac{1}{a^2}$

c. Jensen's Inequality [34]

Let X be a r.v with finite mean $E[X]$, and let f be a convex function, then

$$f(E[x]) \leq E[f(x)] \quad \dots\dots\dots (2.18)$$

The function $f: \square \rightarrow \square$ is called convex if for every x, y and $0 < \lambda < 1$,
 $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.

Examples[2.11.2][4]

1- Let $g(w) = e^w$.

Then

$$\exp \int_{\Omega} w dP \leq \int_{\Omega} e^w dP$$

2- If $\Omega = \{1, 2, 3, \dots, n\}$ with the probability measure P defined $P\{i\} = \frac{1}{n}$ and the function $g(i) = w_i$, we obtain

$$\exp \left\{ \frac{1}{n} (w_1 + w_2 + \dots + w_n) \right\} \leq \frac{1}{n} \{ e^{w_1} + e^{w_2} + \dots + e^{w_n} \}.$$

Setting $y_i = e^{w_i}$. We obtain the Geometric mean inequality.

That is,

$$(y_1, y_2, \dots, y_n)^{\frac{1}{n}} \leq \frac{1}{n} \{ y_1 + y_2 + \dots + y_n \}$$

D. Kolmogorov's Inequality [23]

Let X_1, \dots, X_n be independent random variables such that $E[X_k] = 0$ and $\text{var}(X_k) = \sigma_k^2 < \infty$. Then for all $a > 0$,

$$P\left\{ \max_{1 \leq k \leq n} |X_1 + \dots + X_k| \geq a \right\} \leq \frac{1}{a^2} \sum_{i=1}^n \sigma_i^2 \quad \dots\dots\dots (2.19)$$

2.12 Convergence [2]

We will introduce some notions of the Convergence of sequences . Notice that we will often not explicitly state the independence of a function $X(w)$ on w . Hence, sets of the form $\{w: X(w) > 0\}$ will often be abbreviated $\{X > 0\}$. Let X_n be a sequence of random variables .There exist several kinds of converge of the sequence of random variables at any design r.vs which are.

Definition(*Convergence Almost Surely*) (a.s) (2.12.1)[35]

Let (Ω, \mathcal{F}, P) be a probability measure space .A sequence $\{X_n\}$ of random variables is said to converge to a random variable X *almost surely* (or *almost every where*), if $P(\lim_{n \rightarrow \infty} X_n(w) = X(w)) = 1$,for every $w \in \Omega$.

There are many equivalent statements (notations) for a sequence of random variables $\{X_n\}$ to converge almost surely to random variables X .

[see 35]

Some Properties of Convergence a.s (2.12.2) [33,35]

I.The following statements are all equivalent conditions (notations) for $X_n \xrightarrow{a.s} X$

1. $P[X_n \rightarrow X] < 1$
2. There exist $\varepsilon > 0, P\left(\liminf_{n \rightarrow \infty} [|X_n - X| \leq \varepsilon]\right) < 1$
3. There exist $\varepsilon > 0, P\left(\limsup_{n \rightarrow \infty} [|X_n - X| > \varepsilon]\right) > 0$

II. Uniqueness : If $X_n \xrightarrow{a.s} X$ and $X_n \xrightarrow{a.s} Y$, then $X = Y$ a.s

III. Suppose $X_n \xrightarrow{a.s} X$, $Y_n \xrightarrow{a.s} Y$, $a_n \rightarrow a$ and a_n real sequence, then

1. $X_n + Y_n \xrightarrow{a.s} X + Y$
2. $a_n X_n \xrightarrow{a.s} aX$
3. $X_n Y_n \xrightarrow{a.s} XY$
4. $\frac{X_n}{Y_n} \xrightarrow{a.s} \frac{X}{Y}$, $Y_n \neq 0$ and $Y \neq 0$

IV. If g be a continuous function, $X_n \xrightarrow{a.s} X$ then $g(X_n) \xrightarrow{a.s} g(X)$.

V. Suppose that for all $\varepsilon > 0$, $\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty$, then $X_n \xrightarrow{a.s} X$

VI. Let A_1, A_2, \dots , be independent. Then, $1_{A_n} \xrightarrow{a.s} 0$ if and only if $\sum_n P(A_n) < \infty$

Definition (*Convergence in Probability (Measure)*) (2.12.3) [32]

Let (Ω, \mathcal{F}, P) be a probability measure space. A sequence $\{X_n\}$ of random variables is said to converge to a random variable X in probability if

$$\lim_{n \rightarrow \infty} P\{\omega: |X_n(\omega) - X(\omega)| \geq \varepsilon\} = 0, \text{ for all } \varepsilon > 0$$

There many equivalent statements (notations) for a sequence of random variables $\{X_n\}$, to converge in probability to random variable X . [see 32]

Some Properties of Convergence in Probability(2.12.4)[21]

I. Uniqueness : If $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y$, then $X = Y$

II. Suppose $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$, and $a_n \xrightarrow{P} a$, then

$$1. X_n + Y_n \xrightarrow{P} X + Y$$

$$2. a_n X_n \xrightarrow{P} aX$$

$$3. X_n Y_n \xrightarrow{P} XY$$

$$4. \frac{X_n}{Y_n} \xrightarrow{P} \frac{X}{Y} \quad Y_n \neq 0 \text{ and } Y \neq 0$$

III. Let g be a continuous function, $X_n \xrightarrow{P} X$ then $g(X_n) \xrightarrow{P} g(X)$

IV. Let g be a continuous function at a , $X_n \xrightarrow{P} a$ then $g(X_n) \xrightarrow{P} g(a)$

V. **Fatou's Lemma:** If $0 \leq X_n \xrightarrow{P} X$, then $\liminf_{n \rightarrow \infty} E[X_n] \geq E[X]$.

VI. Suppose $X_n \xrightarrow{P} X$ and $|X_n| \leq Y$ with $E[Y] < \infty$, then $E[X_n] \rightarrow E[X]$

VII. Let A_1, A_2, A_3, \dots , be independent. Then $1_{A_n} \xrightarrow{P} 0$ if and only if $P(A_n) \rightarrow 0$.

Example (2.12.5)

Let (Ω, \mathcal{F}, P) be a probability measure space. It is shown that $X_n \rightarrow 0$ in probability if and only if $E\left[\frac{|X_n|}{1+|X_n|}\right] \rightarrow 0$, as $n \rightarrow \infty$.

Solution

Suppose $E\left[\frac{|X_n|}{1+|X_n|}\right] \rightarrow 0$.

Now, suppose $g(u) = u/(1+u)$ is increasing in $u \geq 0$ and bounded by 1, then

$$P(|X_n| > \varepsilon) = P\left(\frac{|X_n|}{1+|X_n|} > \frac{\varepsilon}{1+\varepsilon}\right) \leq \frac{\varepsilon}{1+\varepsilon} E\left[\frac{|X_n|}{1+|X_n|}\right] \rightarrow 0.$$

Conversely, since $g(u) \leq 1$ for all u and is increasing in u ,

$$E\left[\frac{|X_n|}{1+|X_n|}\right] \leq \frac{\varepsilon}{1+\varepsilon} P(|X_n| \leq \varepsilon) + P(|X_n| > \varepsilon) \rightarrow \frac{\varepsilon}{1+\varepsilon}.$$

Since $\varepsilon > 0$ is arbitrary.

Definition (Convergence in Distribution (in Law)) (2.12.6) [29]

Let (Ω, \mathcal{F}, P) be a probability measure space. A sequence $\{X_n\}$ of random variables is said to converge to a random variable X "in distribution" (usually abbreviated as D) if and only if

$$E[g(X_n)] \rightarrow E[g(X)] \quad \text{as } n \rightarrow \infty \quad \text{for all bounded and continuous function } g: \mathbb{R} \rightarrow \mathbb{R}$$

There are many equivalent statements (notations) for a sequence of random variables $\{X_n\}$ to converge in distribution to random variable X . [see 29]

Some properties of Converges in Distribution (1.12.7)[9,35]

I. Uniqueness : If $F_n \rightarrow F$ and $F_n \rightarrow F^*$, then $F = F^*$

II. Suppose $X_n \rightarrow X$, therefore

1. If $P[X \text{ is a discontinuous point of } g] = 0$, then $g(X_n) \rightarrow g(X)$

2. If $E[g(X_n)] \rightarrow E[g(X)]$ for every bounded real valued functions g such that $P[X \in D_g] = 0$ where D_g is the set of points of discontinuous of g , then $g(X_n) \rightarrow g(X)$ for g continuous.

3. **Fatou's lemma :**

If g is a nonnegative continuous, then $E[g(X)] \leq \liminf_{n \rightarrow \infty} E[g(X_n)]$.

III. If $X_n \rightarrow a$ and g is a continuous at a , then $g(X_n) \rightarrow g(a)$

Definition (Convergence in L^p) (2.12.8)[4]

Let (Ω, \mathcal{F}, P) be a probability measure space. A sequence $\{X_n\}$ of random variables is said to converge to a random variable X in " L^p " if and only if

$$E[|X_n - X|^p] \rightarrow 0 \quad n \rightarrow \infty, \quad 0 < p < \infty$$

There are many equivalent statements (notations) for a sequence of random variables $\{X_n\}$ to converge in L^p to random variable X . [see 29]

Definition (Convergence in Almost Uniformly (a.u.)) (2.12.9)[4]

Let (Ω, \mathcal{F}, P) be a probability measure space. A sequence $\{X_n\}$ of a random variables is said to converge to random variable X *almost uniformly* if give $\varepsilon > 0$, there is a set $A \in \mathcal{F}$ with $P(A) < \varepsilon$ such that $X_n \rightarrow X$ uniformly on A^c .

If X_n and Y_n be a sequence of random variables convergences to X and Y respectively for any constant a then :

$$\text{I. } X_n + Y_n \xrightarrow{a.u.} X + Y$$

$$\text{II. } a_n X_n \xrightarrow{a.u.} aX$$

$$\text{III. } X_n Y_n \xrightarrow{a.u.} XY$$

$$\text{IV. } \frac{X_n}{Y_n} \xrightarrow{a.u.} \frac{X}{Y} \quad Y_n \neq 0 \text{ and } Y \neq 0$$

Definition (Weak Converges) (2.12.10) [37]

A sequence of probability measure P_n on Ω converges weakly to P whenever $P(a) = P(b) = 0$, for $a < b \in \Omega$, we have

$$\lim_{n \rightarrow \infty} P_n[a, b] = P[a, b].$$

A sequence of r.v $\{X_n\}$ converges weakly to X if the induced probability measure P_n converges weakly to P . That is denoted by $P_n \Rightarrow P$ or $X_n \Rightarrow X$.

Theorem (2.12.11) [26,36]

Let P_n and P be probability measures on Ω . Denoted by F_n and F the corresponding distribution functions. The following are equivalent.

I. $P_n \Rightarrow P$ as $n \rightarrow \infty$

II. $F_n(x) \rightarrow F(x)$ for all $x \in \Omega$ such $F(x^-) = F(x)$ as $n \rightarrow \infty$

III. On some probability measure space (Ω, \mathcal{F}, P) , there exists random variables X and X_n with laws P , and P_n , respectively such that $X_n \rightarrow X$ a.s.

Proof

Suppose (II) to prove (I)

First, note that x is a continuity point of F if and only if $P(x) = 0$.

Let $a < b$ be continuity point of F .

Suppose $F_n(x) \rightarrow F(x)$ for each continuity point x of F . Then,

$$\lim_{n \rightarrow \infty} P_n[a, b] = \lim_{n \rightarrow \infty} F_n(b) - F_n(a) = F(b) - F(a) = P[a, b].$$

Suppose (I) to prove (II)

Suppose $P_n \Rightarrow P$. Then,

$$\lim_{n \rightarrow \infty} F_n(b) - F_n(a) = \lim_{n \rightarrow \infty} P_n[a, b] = P[a, b], \text{ then } F_n(x) \rightarrow F(x). \quad \square$$

In the same way we can prove (III)

Theorem (2.12.12) [5]

X_n converges weakly to X if and only if $E[g(X_n)] \rightarrow E[g(X)]$, for all g bounded and continuous function.

2.13 Convergence Theorems

The purpose of these theorems is to allow us to go from a sequence of functions to a limit function, and ultimately to be able to be a little bit more free when integrating with respect to various probability measures.

Theorem (*Bounded Convergence Theorem*)(2.13.1) [10,33]

If the sequence $\{X_n\}$ of random variables is uniform bounded and if $X_n \rightarrow X$ in measure as $n \rightarrow \infty$, then

$$\lim_{n \rightarrow \infty} \int_{\Omega} X_n dP = \int_{\Omega} X dP$$

Proof

Since $\left| \int_{\Omega} X_n dP - \int_{\Omega} X dP \right| = \left| \int_{\Omega} (X_n - X) dP \right| \leq \int_{\Omega} |X_n - X| dP$.

We need only prove that $X_n \rightarrow 0$ in measure and

$|X_n| \leq K$, then $\int_{\Omega} |X_n| dP \rightarrow 0$.

To see this

$$\int_{\Omega} |X_n| dP = \int_{|X_n| \leq \varepsilon} |X_n| dP + \int_{|X_n| > \varepsilon} |X_n| dP \leq \varepsilon + K P(\{w: |X_n(w)| > \varepsilon\})$$

and taking limits, $\limsup_{n \rightarrow \infty} \int_{\Omega} |X_n| dP \leq \varepsilon$.

and since $\varepsilon > 0$ arbitrary. \square

Theorem (*Fatou's Lemma*) (2.13.2)[28]

If for each $n \geq 1, X_n \geq 0$ is a random variables and $X_n \rightarrow X$ in measure as $n \rightarrow \infty$, then

$$\int_{\Omega} X dP \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n dP.$$

Proof

Suppose Y is bounded and it satisfies $0 \leq Y \leq X$. Then the sequence $Z_n = X_n \wedge Y = \min(X_n, Y)$ is uniformly bounded and

$$Z_n \rightarrow Z = X \wedge Y = Y.$$

Therefore, by the bounded convergence theorms,

$$\int_{\Omega} Y dP = \lim_{n \rightarrow \infty} \int_{\Omega} Z_n dP.$$

Since $\int_{\Omega} Z_n dP \leq \int_{\Omega} X_n dP$ for every n it follows that

$$\int_{\Omega} Y dP \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n dP.$$

As Y satisfying $0 \leq Y \leq X$ is arbitrary. \square

Theorem (*Monotone Convergence Theorem(MCT)*) (2.13.3)[38]

If for sequence $\{X_n\}$ of nonnegative functions, we have $X_n \rightarrow X$ monotonically, then

$$\int_{\Omega} X_n dP \rightarrow \int_{\Omega} X dP \text{ as } n \rightarrow \infty$$

Example (2.13.4)

Let (Ω, \mathcal{F}, P) be the Lebesgue probability measure space $([0,1), \mathcal{B}, P_L)$. Let $X : [0,1) \rightarrow \mathbb{R}$ be defined by $X(\omega) = \omega$. For every positive integer n , let

$X_n : [0,1) \rightarrow \mathbb{R}$ be defined by

$$X_n(\omega) = \begin{cases} 0, & \text{if } \omega < \frac{1}{2^n} \\ \frac{1}{2^n}, & \text{if } \frac{1}{2^n} \leq \omega < \frac{2}{2^n} \\ \frac{2}{2^n}, & \text{if } \frac{2}{2^n} \leq \omega < \frac{3}{2^n} \\ \vdots & \vdots \\ \frac{2^n-1}{2^n}, & \text{if } \frac{2^n-1}{2^n} \leq \omega < 1. \end{cases}$$

Use the variables X_n to prove carefully that $EX = \frac{1}{2}$

Solution

If $k2^{-n} \leq w < (k+1)2^{-n}$, then $(2k)2^{-n-1} \leq w < (2k+1)2^{-n-1}$ or

$(2k+1)2^{-n-1} \leq w < (2k+2)2^{-n-1}$. Then

$$X_n(w) = k2^{-n} = (2k)2^{-n-1} \leq X_{n+1}(w),$$

so X_n is increasing. Also, for every w , $|X_n(w) - X(w)| \leq 2^{-n}$, so X_n

converges to X pointwise. By the Monotone Convergence Theorem,

$EX_n \rightarrow EX$. But

$$EX_n = \sum_{k=0}^{2^n-1} k 2^{-2n} = 2^{n-1}(2^n-1)2^{-2n} \rightarrow \frac{1}{2} \text{ as } n \rightarrow \infty.$$

2.14 Limit Theorems

Theorem (Lebesgue's Monotone Convergence Theorem) (2.14.1)[4]

Suppose $\{X_n\}$ is a sequence of random variables satisfying.

I. $0 \leq X_1(w) \leq X_2(w) \leq \dots$, for every $w \in \Omega$ and

II. $X_n(w) \uparrow X(w)$, for every $w \in \Omega$. Then $\int_{\Omega} X_n dP \uparrow \int_{\Omega} X dP$.

Proof

Set $\alpha_n = \int_{\Omega} X_n dP$. Then α_n is a nondecreasing and it converges to $\alpha \in [0, \infty]$. Since

$$\int_{\Omega} X_n dP \leq \int_{\Omega} X dP, \text{ for all } n$$

we see that if $\alpha = \infty$, then $\int_{\Omega} X dP = \infty$. Assume $\int_{\Omega} X dP < \infty$. Since $\alpha \leq \int_{\Omega} X dP$.

We need to prove the opposite inequality. Let $0 \leq \varphi \leq X$ be simple and

$$\text{let } 0 < a < 1. \text{ Set } A_n = \{w: X_n(w) \geq a \varphi(w)\}.$$

Clearly $A_1 \subset A_2 \dots$. In addition,

suppose $w \in \Omega$. If $X(w) = 0$, then $\varphi(w) = 0$ and $w \in A_n$ then $a \varphi(w) < X(w)$ and since $X_n(w) \uparrow X(w)$, we have that $w \in A_n$ some n . Hence $\cup A_n = \Omega$.

Our notation proposition (1.6.1) $A_n \uparrow \Omega$.

Hence

$$\int_{\Omega} X_n dP \geq \int_{A_n} X_n dP \geq a \int_{A_n} \varphi(w) dP = a P_1(A_n).$$

Let $n \uparrow \infty$. By proposition (1.9.5) and proposition (1.6.1)

$$\alpha \geq \lim_{n \rightarrow \infty} \int_{\Omega} X_n dP \geq a P_1(\Omega) = a \int_{\Omega} \varphi dP \text{ and, therefore}$$

$$\int_{\Omega} \varphi dP \leq \alpha, \text{ for all simple } \varphi \leq X \text{ and } \sup_{\varphi \leq X} \int_{\Omega} \varphi dP \leq \alpha. \quad \square$$

Example (2.14.2)[4]

Let $\Omega = [0, \infty)$ and $X_n(w) = -\frac{1}{n}$ for all w . Then $\int_{\Omega} X_n dP = -\infty$, but

$X_n \rightarrow X$ where $X = 0$ and $\int_{\Omega} X dP = 0$, then X_n are not nonnegative.

Corollary (*First Borel – Cantelli Lemma*) (2.14.3) [33]

Let $\{A_n\}$ be a sequence of measurable sets . Suppose $\sum_{n=1}^{\infty} P(A_n) < \infty$.

Then $P(A_n \text{ i.o.}) = 0$.

Corollary (2.14.4) [29]

Let X be nonnegative random variable .Define $Q(A) = \int_{\Omega} X dP$.

Then Q is probability measure and $\int_{\Omega} Y dQ = \int_{\Omega} Y X dP$ for all nonnegative random variable Y .

Theorem (*Fatou's Lemma*) (2.14.5) [17]

Let $\{X_n\}$ be a sequence of nonnegative random variable.

Then $\int_{\Omega} \liminf_{n \rightarrow \infty} X_n dP \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n dP$.

Theorem (*The Dominated Convergence Theorem*) (2.14.6) [24,27]

Suppose X_n are random variables and $X_n(w) \rightarrow X(w)$.

Suppose there exists an integrable function Y such that

$|X_n(w)| \leq Y(w)$ for all w . Then $\int_{\Omega} X_n dP \rightarrow \int_{\Omega} X dP$

Proof

Since $X_n + Y \geq 0$, by Fatou's Lemma(2.13.2),

$\int_{\Omega} (X + Y) dP \leq \liminf_{n \rightarrow \infty} \int_{\Omega} (X_n + Y) dP$ since Y is integrable .

$\int_{\Omega} X dP \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n dP$.

Similarly , $Y - X_n \geq 0$, so

$$\int_{\Omega} (Y - X) d\mathbb{P} \leq \liminf_{n \rightarrow \infty} \int_{\Omega} (-X_n) d\mathbb{P} = -\limsup_{n \rightarrow \infty} \int_{\Omega} X_n d\mathbb{P}.$$

Therefore $\int_{\Omega} X d\mathbb{P} \geq \limsup_{n \rightarrow \infty} \int_{\Omega} X_n d\mathbb{P}$. \square

Example(2.14.7)

Let $\{A_n\}$ be a sequence of sets in \mathcal{F} , and set $A = \bigcup_n A_n$. Let X be a random variable such that $E[X]$ exists and $E|X| < \infty$. Show carefully that, as $n \rightarrow \infty$,

$$\int_{\bigcup_{k=1}^n A_k} X d\mathbb{P} \rightarrow \int_A X d\mathbb{P}.$$

Solution

For each n let $B_n = \bigcup_{k=1}^n A_k$. Then $B_n \subseteq B_{n+1}$ for every n .

Now we can always decompose $X = X^+ - X^-$, where $X^+ = \max\{X, 0\}$ and $X^- = -\min\{X, 0\}$. Then it is easy to see that $X^+ 1_{B_n}$ increases pointwise to $X^+ 1_A$, and $X^- 1_{B_n}$ increases pointwise to $X^- 1_A$.

Now

$$E[X 1_{B_n}] = \int_{\bigcup_{k=1}^n A_k} X d\mathbb{P} = \int_{\bigcup_{k=1}^n A_k} X^+ d\mathbb{P} - \int_{\bigcup_{k=1}^n A_k} X^- d\mathbb{P} = E[X^+ 1_{B_n}] -$$

$E[X^- 1_{B_n}]$. By the MCT, we have $E[X^+ 1_{B_n}] \rightarrow E[X^+ 1_A]$

and $E[X^- 1_{B_n}] \rightarrow E[X^- 1_A]$, so $E[X 1_{B_n}] \rightarrow E[X 1_A]$.

2.15 Some Relationship between Convergence [4,5]

There some theorem which represent the relationship between convergence which are .

a. Theorem [12,23]

Let $\{X_n\}$ be a sequence of a random variable on the probability measure space (Ω, \mathcal{F}, P) , if $X_n \xrightarrow{a.s} X$, then $X_n \xrightarrow{P} X$.

Remark (2.15.1) [23]

If X_n converges in probability it is not true that converges to X

Example(2.15.2) [23]

Let $\Omega = [0,1]$ with lebesgue probability measure .Consider the sequences of set $I_1 = [0, \frac{1}{2}]$, $I_2 = [\frac{1}{2}, 1]$, $I_3 = [0, \frac{1}{3}]$, $I_4 = [\frac{1}{3}, \frac{2}{3}]$, $I_5 = [\frac{2}{3}, 1]$, and so on .
If $X_n(\omega) = 1_{I_n(\omega)}$, then X_n tend to 0 in probability but not almost surely .

Theorem(2.15.3) [12]

The sequence of random variable $\{X_n\}$ on (Ω, \mathcal{F}, P) converges to X in probability measure if and only if every subsequence $\{X_{n_k}\}$ contains a further subsequence converging a.s to X

Proof

Let ε_k be sequence converging down to 0 .Then $P\{|X_n - X| > \varepsilon_k\} \rightarrow 0$, as $n \rightarrow \infty$ for each k .We therefore have a subsequence X_{n_k} satisfying

$$P\{|X_{nk} - X| > \varepsilon_k\} < \frac{1}{2^k}$$

Hence

$\sum_{k=1}^{\infty} P\{|X_{nk} - X| > \varepsilon_k\} < \infty$ and therefore by (the first Borel cantelli Lemma),

$P\{|X_{nk} - X| > \varepsilon_k \text{ i.o}\} = 0$ Thus $|X_{nk} - X| > \varepsilon_k$ eventually a.s .Thus

$X_{nk} \rightarrow X$ a.s .For the converse .

Let $\varepsilon > 0$ and put $y_n = P\{|X_n - X| > \varepsilon\}$ and consider the subsequence y_{nk} .If every X_{nk} subsequence has a subsequence X_{nkj} , such that $X_{nkj} \rightarrow X$ a.s . Then $\{y_{nk}\}$ has a subsequence $y_{nkj} \rightarrow 0$.Therefore $\{y_n\}$ converges to 0 and has that is $X_n \rightarrow 0$ in probability measure .

b. Theorem [40,16]

Let $\{X_n\}$ be a sequence of a random variable on the probability measure space (Ω, F, P) , if $X_n \xrightarrow{m.s} X$, then $X_n \xrightarrow{P} X$.

c. Theorem [22]

Let $\{X_n\}$ be a sequence of a random variable on the probability measure space (Ω, F, P) , if $X_n \xrightarrow{L^p} X$, then $X_n \xrightarrow{P} X$.

Proof

Let $\varepsilon > 0$,by using Chebyshev's inequality

$$P(\{\omega: |(X_n(\omega) - X(\omega))| > \varepsilon\}) \leq P(\{\omega: |(X_n(\omega) - X(\omega))|^p > \varepsilon^p\}) \leq$$

$$\frac{1}{\varepsilon^p} \int_{\Omega} |(X_n(\omega) - X(\omega))|^p dP$$

Remark (2.15.4) [22]

If X_n converges in probability it is not true that converges to X in L^p

Example (2.15.5) [5]

Let $\Omega = [0,1]$, \mathcal{F} the Borel σ -field and P Lebesgue probability measure

.Let $X_n = e^{n \cdot \mathbb{1}_{(0, \frac{1}{n})}}$, it is clearly that $X_n \mathbb{1}_{A_n}$ converges to X almost surely and in

probability, but $E[X_n^p] = \frac{e^{np}}{n} \rightarrow \infty$ not converges in L^p for any p .

d. Theorem [40,16]

Let $\{X_n\}$ be a sequence of random variable on the probability measure

space (Ω, \mathcal{F}, P) , if $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{D} X$.

Proof

Let $a \in \mathbb{R}$ be given, and set $\varepsilon > 0$. on the one hand

$$\begin{aligned} F_{X_n}(a) &= P(X_n \leq a, X \leq a + \varepsilon) + P(X_n \leq a, X > a + \varepsilon) \\ &= P(X_n \leq a \setminus X \leq a + \varepsilon)P(X_n \leq a + \varepsilon) + P(X \leq a, X > a + \varepsilon) \leq \\ &P(X_n \leq a + \varepsilon) + P(X < a - \varepsilon) \leq F_X(a + \varepsilon) + P(|X_n - X| > \varepsilon), \end{aligned}$$

Where we have used the fact that if A and B , then $P(A) \leq P(B)$.

By a similar argument

$$\begin{aligned} F_X(a - \varepsilon) &= P(X \leq a - \varepsilon, X_n \leq a) + P(X \leq a - \varepsilon, X_n > a) \\ &= P(X \leq a - \varepsilon \setminus X_n \leq a)P(X_n \leq a) + P(X \leq a - \varepsilon, X_n > a) \leq \\ &P(X_n \leq a) + P(X < X_n - \varepsilon) \leq F_{X_n}(a) + P(|X_n - X| > \varepsilon). \end{aligned}$$

Thus, we have obtained that

$$F_X(a - \varepsilon) - P(|X_n - X| > \varepsilon) \leq F_{X_n}(a) \leq F_X(a + \varepsilon) + P(|X_n - X| > \varepsilon).$$

Taking now $n \rightarrow \infty$, we have

$$F_X(a - \varepsilon) \leq \liminf_{n \rightarrow \infty} F_{X_n}(a) \leq \limsup_{n \rightarrow \infty} F_{X_n}(a) \leq F_X(a + \varepsilon).$$

Finally, since this inequality

holds for any $\varepsilon > 0$, we conclude that

$$\lim_{n \rightarrow \infty} F_{X_n}(a) = F_X.$$

e. Theorem [4,5]

Suppose $X_n \rightarrow X$ almost uniformly. then $X_n \rightarrow X$ in probability measure and almost surely.

Proof

Since $X_n \rightarrow X$ almost uniformly. Given $\varepsilon > 0$ there is a measurable set E such that $P(E) < \varepsilon$ and $X_n \rightarrow X$ uniformly on E^c . Let $\delta > 0$ be a given.

There is a $N = N(\delta)$ such that $|X_n - X| < \delta$ for all $n > N$ and for all

$w \in E^c$. That is $P\{|X_n(w) - X(w)| \geq \delta\} \subseteq E$ for all $n \geq N$. Hence, for all $n \geq N$

$$P\{|X_n(w) - X(w)| \geq \delta\} < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary we see that for all $\delta > 0$,

$$\lim_{n \rightarrow \infty} P\{|X_n(w) - X(w)| \geq \delta\} = 0, \text{ proving that } X_n \rightarrow X \text{ in probability.}$$

Next, for each n take $A_n \in \mathcal{F}$ with $P(A_n) < \frac{1}{n}$ and $X_n \rightarrow X$ uniformly on A_n^c .

If $E = \bigcup_{n=1}^{\infty} A_n$, then $X_n \rightarrow X$ on E^c and $P(E) = P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq P(A_n)$ for all n .

Thus $P(E) = 0$ and we have the almost sure convergence.

Example (2.15.6) [4]

Let $\Omega = [0, \infty)$ with lebesgue probability measure and define $X_n(\omega) = \mathbf{1}_{A_n}(\omega)$

where $A_n = \{\omega \in \Omega : n \leq \omega \leq n + \frac{1}{n}\}$, that $X_n \rightarrow 0$ a.s, in probability measure

and in L^p but that $X_n \not\rightarrow 0$ almost uniformly.

Theorem(2.15.7)

Suppose $X_n \rightarrow X$ in probability measure . Then there is a subsequence $\{X_{n_k}\}$ which converges almost uniformly to X .

Theorem (Egreff's theorem) (2.15.8) [4]

Suppose (Ω, \mathcal{F}, P) is finite probability measure space and that $X_n \rightarrow X$ a.s. Then $X_n \rightarrow X$ almost uniformly.

Proof

Let $\varepsilon > 0$ be given ,for k there is a $n(k)$ such that if

$$A_k = \bigcup_{n=n(k)}^{\infty} \{\omega \in \Omega : |X_n - X| \geq \frac{1}{k}\}, \text{ Then } P(A_k) < \frac{\varepsilon}{2k}. \text{ Thus if } A = \bigcup_{k=1}^{\infty} A_k,$$

then $P(A) \leq \sum_{k=1}^{\infty} P(A_k) < \varepsilon$. Now ,if $\delta > 0$ take k so large that $\frac{1}{k} < \delta$ and then

for any $n > n(k)$ and $\omega \notin A$, $|X_n(\omega) - X(\omega)| < \frac{1}{k} < \delta$, thus

$X_n \rightarrow X$ uniformly on A^c .

Note[5]

A scheme below explains the relationship between convergence like converge almost surely, converge in probability, converge in distribution , converge

In L^p and converge almost uniformly . Then new square are added to the scheme by the researcher due to his finding of the scheme in the various references , which are in the following :

From above scheme, we can see that for example :-

1. For continuous probability measure space , $X_n \xrightarrow{a.s} X$ if and only if $X_n \xrightarrow{P} X$
2. Suppose $X_n \xrightarrow{P} X$, for all n $|X_n| \leq Y$, $Y \in L^p$. Then, $x \in L^p$ and $X_n \xrightarrow{L^p} X$.
3. Suppose $X_n \xrightarrow{a.s} X$, for all n $|X_n| \leq Y$, $Y \in L^p$, Then, $x \in L^p$ and $X_n \xrightarrow{L^p} X$.
4. If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{D} X$
5. If $X_n \xrightarrow{D} X$, and if there exists $a \in \mathbb{R}$, $X = a$ a.s, then $X_n \xrightarrow{P} X$. Hence , when $X = 0$ a.s , $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{D} X$ are equivalent.

Theorem (2.15.9) [3,4]

Let $\{X_n\}$ be a sequence of independent random variables ,then :

- I. If X_n converges to X in probability, then it converges weakly.
- II. X_n converges weakly to a constant, it converges in probability .
- III. **Slutsky's Theorem** : If X_n converges weakly to X and Y_n converges weakly to a constant a , then $X_n + Y_n$ converges weakly to $X+a$ and $X_n Y_n$ converges weakly to $a X$.

Remark (2.15. 10) [4]

From (I) the converse is not true in general . However , if X is a degenerate distribution (takes a single value with probability one) . The converse is true .

Theorem (2.15.11) [37]

Suppose $X_n \rightarrow X$, and X is a degenerate distribution such that $P\{X = a\} = 1$. Then $X_n \xrightarrow{P} X$.

2.16 The Law of Large Numbers [41]

In probability theory, several Laws of large numbers say that the average of a sequence of random variables with a common distribution converges to their common expectation in the limit as the size of the sequence goes to infinity. The phrase “law of large numbers” is also sometimes used to refer to the principle that the probability of any possible event (even an unlikely one) occurring at least once in a series increase with the number of events in the series.

For example, the odds that we will win the lottery are every low, however, the odds that someone will win the lottery are quite good, provided that a large enough number of people purchased lottery tickets. There are two varieties of this law:

The weak law of large numbers and the strong law of large numbers. The weak law states that the average converges in probability to $E[\bar{X}_n]$. The strong law states that the average converges almost surely to $E[\bar{X}_n]$.

However, the strong law is significantly harder to prove, and requires a bit of additional machinery.

Theorem (*The Weak Law of Large Numbers*) (2.16.1) [23,41]

Let $\{X_n\}$ be a sequence of independent identically distributed (i.i.d) random variables, that is $F_{X_n} = F_X$ for all n with finite mean μ and finite variance ($\sigma^2 < \infty$),

let also

$$\bar{X}_n = \frac{S_n}{n} = \frac{(X_1 + X_2 + \dots + X_n)}{n}$$

be the sample mean of n random variables, then

$$\bar{X}_n \xrightarrow{P} \mu \text{ as } n \rightarrow \infty$$

i.e.

for every $\varepsilon > 0$

$$P(|\bar{X} - \mu| \geq \varepsilon) \rightarrow 0, \text{ as } n \rightarrow \infty \quad \dots\dots\dots(2.20)$$

Theorem (*The Strong Law of Large Numbers*) (2.16.2) [41]

Let $\{X_n\}$ be independent sequence and identically distribution (i.i.d) random variables with a finite mean $\mu = E(X_i)$ and finite fourth central moment $\mu_4 = E(X_i - \mu)^4$ for $i = 1, 2, \dots$, then

$$\bar{X}_n \xrightarrow{a.s} \mu.$$

$$\text{That is } P(\lim_{n \rightarrow \infty} \bar{X}_n = \mu) = 1. \quad \dots\dots\dots(2.21)$$

2.17 The Central Limit Theorem [23]

The central limit theorem is one of the most remarkable results of the theory of probability. In its simplest form, the theorem states the sum of a large number of independent observation from the same distribution has, under certain general conditions, an approximate normal distribution. Moreover, the approximation steadily improves as the number of observations increases. The theorem is considered as the basis of probability theory, although a better name would be a normal convergence theorem.

If X_n is sum of n real independent random variables

Theorem (2.17.1) [42]

Let $\{X_n\}$ be a sequence of (i.i.d) random variables with $E(X_n) = 0$ and $\text{var}(X_n) = 1$. Then, The sequence of random variables is

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

converges in distribution to a random variable $X \sim N(0,1)$

that is

$$\lim_{n \rightarrow \infty} P(\bar{X}_n \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{y^2}{2}} dy.$$

If $E(X_i) = \mu$ and $\text{var}(X_i) = \sigma^2$, then the same applies for

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma}.$$

The central limit theorem (CLT) is about running average rescaled by factor of \sqrt{n} . If we denote by \bar{Y}_n the running average

$$\bar{Y}_n = \frac{X_1 + X_2 + \dots + X_n}{n}, \text{ then the states that } P(\bar{Y}_n \leq \frac{a}{\sqrt{n}}) \rightarrow F(a)$$

where $F(a)$ is the cumulative distribution function of $N(0,1)$.

4.1 Conclusions

- 1- The functions of (standard normal distribution, exponential distribution, Cauchy distribution) are all measurable simple functions.
- 2- We know that $\int_{-\infty}^{\infty} f(x) dx = 1$ in elementary probability theory . We have find the same value in measure theory when we use the simple function .
- 3- The sum of two functions with i.d satisfies the relation $M(f + g) = M(f) + M(g)$, when we apply it for the standard normal distribution and standard Cauchy distribution .
- 4- The area under the curve of the function $\xi(w)$ doesn't equal one , that means that this function isn't a probability function on the two parts of interval $(0, 1.851482)$ and $(1.851482, 30)$.
- 5- When we change the number of division of X - axes (n), we find when, $n = 2.5$, that the sum of the two area is very close to one at the intersection point for all cases . But when we change the variance of two distributions , we find that $\sigma^2 = 5$.The sum of area is very near to the one at the intersection point 3.875
- 6- We find that the function $\xi_1(w) = \frac{\alpha \cos(\frac{w}{a})}{\beta \pi + w^\gamma}$ by using Gragher programmer after multi experimental attempts . This function doesn't satisfy all conditions of the probability function.
- 7- When we try to change the value of α, a and (α, a) together ,we find that all cases don't satisfy all of conditions of the probability function.
- 8- When $\xi_2(w) = \frac{f(w) + h(w)}{2}$,we find that it is a probability function which satisfying all the probability conditions on interval $(-30, 30)$.We show that it is a simple and measurable function and , $\int_{-30}^{30} \xi_2 dM = 1$.

- 9- We can use the distribution function to show that the measure and probability measure give the same value in probability theory for the distributions we used .The same use will be done for the sum of two functions.

4.2 Recommendations

- 1- This problem can be solved by using another method such as ; L^p space , fuzzy probability measure ,etc.
- 2- We can This solve the problem having method more than two distributions .
- 3- Using this method or any method to calculate the measure of sum of two discrete distributions which doesn't approximate to normal distribution .
- 4- If two functions have not any intersection points , then try to use the following relation $M(f) = M(g)$ whenever $M(w:f(w) \neq g(w)) = 0$
- 5- One can use the relation $M(f) \leq M(g)$ if $f \leq g$ for two functions f and g .

الخلاصة

ظهرَ في الآونة الأخيرة اتجاه للدمج بين فروع الرياضيات المختلفة مع الاحتمالية لحل المسائل الرياضية وانطلاقاً من هذا المبدأ ارتأى الباحث استخدام نظرية الاحتمالات ونظرية القياس لمحاولة لإيجاد توزيع لمتغير جديد والذي يساوي مجموع متغيرين لهم توزيعين مختلفين وذلك بتطبيق بعض النظريات والخواص من نظرية القياس والذي أصبح هدف أساسياً لهذا البحث .

وقد تضمن هذا البحث أربع فصول على النحو الآتي :-

يتناول الفصل الاول بعض الأفكار العامة والمصطلحات والنظريات من القياس والتي نحتاجها في نهاية البحث مثل حقل سكما , حقل سكما بورل , القياس , قياس ليبيك , قياس ليبيك استالجس , الدالة القابلة للقياس , الدالة البسيطة الخ تناول الفصل الثاني بعض التعاريف , المفاهيم والنظريات المتعلقة بقياس الاحتمالية قياس الاحتمالية , المتغير العشوائي , دالة التوزيع , الاستقلالية , التقارب ونظريات التقارب . أما في الفصل الثالث فقد تم استخدام نظرية الاحتمالية الكلاسيكية وكذلك نظرية القياس لبيان تساوي قيمة الاحتمال والقياس في حالة دالة واحدة وكذلك في حالة دالتين التي شملت متغيرات مستقلة مستمرة لها نفس التوزيع ومتغيرات مستقلة مستمرة لها توزيعات مختلفة , ومحاولة إيجاد دالة تشمل التوزيع الطبيعي القياسي وتوزيع كوشي القياسي . وقد تم التوصل الى استنتاجات في الفصل الرابع ومنها :

1- المعروف من الاحتمالية البسيطة أن $\int_{-\infty}^{\infty} f(x)dx = 1$ وكذلك تم إيجاد نفس القيمة في

نظرية القياس عندما

استخدمت كدالة بسيطة.

2-مجموع دالتي متغيرين (i.d) تحقق العلاقة $M(f + g) = M(f) + M(g)$ عند تطبيقها

للتوزيع الطبيعي القياسي والتوزيع الكوشي القياسي .

3-المساحة تحت منحنى الدالة (w) لا يساوي واحد , وهذا يعني أنها دالة غير احتمالية على

جزئي الفترة واللذان هما $(0,1.851482)$ و $(1.851482,30)$.

4- باستخدام برنامج الرسم تم إيجاد الدالة $\xi_1(w) = \frac{\alpha \cos(\frac{w}{a})}{\beta \pi + w^\gamma}$ بعد عدة محاولات تجريبية وهذه

الدالة لا تحقق جميع شروط الدالة الاحتمالية.

5- عند استخدام الدالة $\xi_2(w) = \frac{f(w) + h(w)}{2}$, تم إيجاد أنها دالة احتمالية يمكن أن تحقق

جميع الشروط الاحتمالية على الفترة (-30,30). وبيننا أنها دالة بسيطة وقابلة للقياس وتكاملها

$$\text{هو } \int_{-30}^{30} \xi_2 dM = 1 .$$

6- يمكن استخدام دالة التوزيع للبيان أن القياس واحتمالية القياس تعطي نفس القيمة في نظرية الاحتمال للتوزيعات التي تم استخدامها وكذلك نفس الشيء بالنسبة إلى مجموع للداليتين .

يوصي الباحث للدراسات المستقبلية في الفصل الرابع وكالاتي :

1- استخدام طرق أخرى لحل هذه المشكلة مثل L^p , وقياس احتمالية الاعداد الضبابية والخ .

فضاء

2- وهذه الطريقة يمكن استخدامها لحل المشاكل عندما نملك أكثر من توزيعين .

3- محاولة أستخدم هذه الطريقة أو أي طريقة لحساب القياس لمجموع توزيعين متقطعين والتي لا يقترب إلى التوزيع الطبيعي .

4- عند عدم امتلاك تمتلك داليتين أي نقاط تقاطع , يمكن إن نستخدم العلاقة التالية:

$$M(f) = M(g) \text{ whenever } M(w: f(w) \neq g(w)) = 0$$

5- يمكن أن نستخدم العلاقة التالية :

$$M(f) \leq M(g)$$

إذا كانت $f \leq g$ لأي داليتين f و g .