PROBABILITY THEORY APPLICATIONS ON TIME SCALES

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Mathematics

by
Sevcan KAHRAMAN

July 2008
İZMİR

We approve the thesis of **Sevcan KAHRAMAN**

Assoc. Prof. Dr. Ünal UFUKTEPE
Supervisor

Prof. Dr. Ali ÇALIŞKAN
Committee Member

Assist. Prof. Dr. Serdar ÖZEN
Committee Member

11 July 2008
Date

Prof. Dr. Oğuz YILMAZ
Head of the Mathematics Department

Prof. Dr. Hasan BÖKE
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my adviser Assoc. Prof. Dr. Ünal UFUKTEPE for his continual presence, invaluable guidance and endless patience throughout this work, and most thanks for providing me his deep friendship. Without his help this thesis would not be possible.

I would like to thank Prof. Dr. Ali ÇALIŞKAN and Assist. Prof. Dr. Serdar ÖZEN for being in my thesis committee, it was a great honor.

My special thanks go to my dear friend Roberto LO PREJATO for encouraging me during my thesis. Also I'm particularly grateful Hatice KUŞAK for answering all my questions about time scale. Also I want to thank to my wonderful friend İbrahim GÜR who set me on my path.

Finally, I am deeply indebted to my mother for always being with me and encouraging me in my life in many ways.

ABSTRACT

PROBABILITY THEORY APPLICATION ON TIME SCALES

In this thesis we introduce probability theory adapted on time scales. In this study, we have modified some concepts of probability theory on time scales. $\Delta-$Probability Function has been constructed on time scales. After providing some basic concepts of Probability Theory, the review of the fundamental concepts of time scale has been provided. Time scale unifies continuous and discrete analysis. With the help of the definition of $\Delta-$Measure and the $\Delta-$Probability Function, random variable has been constructed on time scale. Mathematical expectation and some probabilistic inequalities are provided to Time scale.

ÖZET

ZAMAN SKALALARINDA OLASILIK KURAMI UYGULAMALARI

Bu tezde, zaman scalasına uyarlanmış olasılık teorisini çalıştık. Bu çalışmada olasılık kuramının bazı kavramlarını zaman skalasına uyarladık. Delta olasılık fonksiyonunu zaman skalasında tanımladık. Olasılık kuramıyla ilgili bazı temel kavramlardan söz ettikten sonra, zaman skalasının kısa bir tekrarını verdik. Zaman skalası sürekli ve ayrık analizi birleştirir. Delta-ölçüm ve delta-olasılık fonksiyonları yardımıyla zaman skalasındaki rassal değişkenleri tanımladık. Matematiksel beklenen değer ve bazı olasılıksal eşitsizlikleri zaman skalasına uyarladık.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Time scale unifies continuous and discrete analysis. Probability theory on time scale has been constructed on time scale. This thesis is organized as follows: In chapter one, same basic concepts about Probability Theory are given. In chapter two, we gave the review of the fundamental concepts of Time Scale. In chapter three, Measure on time scale was discussed. In chapter four, $\Delta-$Probability Function and random variable have been constructed on time scale. We also have derived some properties of expectations of random variables on time scales, they need not be discrete or be continuous only. In addition, expectation, variance and some important probabilistic inequalities are modified to Time scales.

1.1. Basic concepts of Probability Theory

We express our behavior of believing in change by the use of words such as ″random″ or ″probability″. The mathematical theory of probability incorporates these concepts of chance. Such a theory formalizes these concepts as a collection of axioms. Any experiment involving randomness can be modelled by a probability space. Such a space comprises of the set Ω of all possible outcomes of by the experiment, the set $\mathcal{F}$ of events, and the probability measure P.
This chapter contains the essential ingredients about the probability theory:

1.1.1. Events and sets

In everyday life we come across many phenomena which can not be predicted in advance or many experiments whose outcomes may not be known precisely. We may know that the outcome has to be one of the several possibilities. The weight of a newborn baby can be known before the birth. However, when a coin is tossed, we know that the outcome has to be a head or a tail. In this case, possible outcomes should be finite, but also the observation of the phenomenon such as the weight of a new born baby or the weather condition of certain region,the number of possible outcomes is infinite. The concept of random experiment is very important for studying probability theory.

1. Sample Space: The set of all possible outcomes of an experiment is called the sample space and is denoted by Ω.

2. Outcomes: The results of an experiment are called outcomes. w is also called a sample point.

3. Events: Any subset of a sample space Ω is called an event.

1.1.2. Fields and $\sigma-$Fields

Let $\mathcal{F}$ = {all events associated with the experiment E on Ω}. If A is an event A^c is also an event which means that the event A is not occurred.

Definition 1.1 *Let $\mathcal{F}$ be a class of the sets of all subsets of Ω. If these conditions are satisfied;*

a) If $A, B \in \mathcal{F}$ then $A \cup B \in \mathcal{F}$ and $A \cap B \in \mathcal{F}$,

b) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$,

c) The empty set belongs to $\mathcal{F}$,

then any collection $\mathcal{F}$ of subsets of Ω is called a field.

Lemma 1.1 *Every field contains the empty set $\varnothing$ and the whole space Ω.*

The class containing only $\varnothing$ and Ω is a field which is called the smallest field (trivial field).

The power set consisting of every subset of Ω is also a field and is the largest field.

Theorem 1.1 *The intersection of arbitrary number of fields is a field.*

Proof (Royden 1988). □

Since events are any subsets of a sample space it can be considered that a collection of these events is a field:

Example 1.1 *Suppose a coin is tossed three times. The set of possible outcomes is finite and is denoted by*

$$\Omega = \{HHH, TTT, TTH, HTT, THT, THH, HTH, HHT\}$$

$$= \{w_1, w_2, w_3, w_4, ..., w_8\}$$

Let A be the event that at least one head turns up, then

$$A = \{w_1, w_3, w_4, w_5, w_6, w_7, w_8\}$$

then $A^c = \{TTT\} = w_2$ represents the event that no head turns up. This is fine when Ω is a finite set, but we will deal with the common situation when Ω is infinite. For example, a coin is tossed repeatedly until the first head turns up. We are concerned with the number of tosses before this happens. The set of all possible outcomes is the set

$$\Omega = \{w_1, w_2, w_3, ...\}$$

In this case the experiment E has an infinite number of outcomes. Let A be an event that the first head occurs after an even number of tosses, then

$$A = \{w_2, w_4, w_6, ...\}$$

This is an infinite countable union of members of Ω and we require that such a set belong to $\mathcal{F}$.

Definition 1.2 *A collection $\mathcal{F}$ of events of Ω is called a $\sigma-$field if it satisfies the following conditions:*

a) $\emptyset$ and $\Omega \in \mathcal{F}$,

b) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$,

c) If $A_1, A_2, ... \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Remark 1.1 *A $\sigma-$field is always a field, but a field need not be a $\sigma-$field.*

Example 1.2 *Let $\Omega = \{1, 2, 3, 4, 5, 6, ...\}$ and $\mathcal{D} = \{A : A \subset \Omega$, either A contains a finite number of points or A^c contains a finite number of points$\}$ Then $\mathcal{D}$ is a field , but not a $\sigma-$field. Because, the set $A = \{A_n\}_{n=1}^{\infty}$ and its complement is not belong to the set $\mathcal{D}$.*

Here are some examples of $\sigma-$fields:

Example 1.3 *The smallest $\sigma-$field associated with Ω is the collection $\mathcal{F} = \{\emptyset, \Omega\}$.*

Example 1.4 *If A be any subset of Ω then $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ is a $\sigma-$field.*

Example 1.5 *A power set of Ω, which is written $\mathcal{P}$ and contains all subsets of Ω, is obviously a $\sigma-$field.*

Theorem 1.2 *The intersection of an arbitrary number of $\sigma-$fields is a $\sigma-$field.*

Example 1.6 *Let us consider $\mathcal{D} = \{[a, b) \in \Omega, a, b \in \mathbb{R}, a < b\}$ is not a field, since this class is not closed under complementations (nor under countable intersections). Let us define $\sigma(\mathcal{D}) = \mathcal{B}$ be the minimal $\sigma-$field containing $\mathcal{D}$, since*

$$
\begin{aligned}
[a, b] &= \cap_{n=1}^{\infty}[a, b + \frac{1}{n}) \text{ (by countable intersection)} \\
(-\infty, a) \cup (b, \infty) &= ([a, b])^c \text{ (by complementation)} \\
(a, b) &= (-\infty, b) \cap (a, \infty) \ (a < b)
\end{aligned}
$$

$(a, b], (-\infty, a]$, *etc. also belong to the set $\mathcal{D}$ for a,b ϵ $\mathbb{R}$.*

$\mathcal{B}$ is called the 'Borel Field' of subsets of the real line. The sets of $\mathcal{B}$ are called 'Borel sets'.
From now on with each experiment E, we denote Ω as the space of all outcomes of the random experiment and $\mathcal{F}$ as the $\sigma-$field of events. This pair $(\Omega, \mathcal{F})$ is called a measurable space.

1.1.3. Probability Measure

Definition 1.3 *A Probability measure P on $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$, satisfying*
a) $P(\emptyset) = 0$ and $P(\Omega) = 1$
b) If $A_1, A_2 \ldots$ is a collection of disjoint members of $\mathcal{F}$, in that case $A_i \cap A_j = \emptyset$ for all pairs i, j satisfying $i \neq j$ then

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

The triplet $(\Omega, \mathcal{F}, P)$ is called a probability space.

A probability measure is a special example of what is called a measure on a pair $(\Omega, \mathcal{F})$. A measure is a function $\mu : \mathcal{F} \to [0, \infty)$ satisfying $\mu(\emptyset) = 0$ together with the condition b) as mentioned above. A measure μ is a probability measure if $\mu(\Omega) = 1$ (Grimmett and Stirzaker 2006).

Proposition 1.1 *Let $A_1, A_2, \ldots$ be subsets of* Ω.

a) If $A_1 \subseteq A_2 \subseteq \ldots$, then $A_n \to A = \bigcup_{n=1}^{\infty} A_n$. (We denote this by $A_n \uparrow A$.)

b) If $A_1 \supseteq A_2 \supseteq \ldots$, then $A_n \to A = \bigcap_{n=1}^{\infty} A_n$. (This is written as $A_n \downarrow A$.)

Proof (Royden 1988). □

Elementary Properties of Probability Measure:

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Then,

a) $P(\emptyset) = 0$.

b) P is finitely additive: if $A_1, \ldots, A_n$ are (pairwise) disjoint, then

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

c) For each A, $P(A^c) = 1 - P(A)$.

d) If $A \subseteq B$, then $P(B \backslash A) = P(B) - P(A)$, where $B \backslash A$ is the set such that $B \backslash A = B \cap A^c$.

e) P is *monotone*: if $A \subseteq B$ then $P(A) \leq P(B)$.

d) For all A and B (disjoint or not),

$$P(A \cup B) + P(A \cap B) = P(A) + P(B).$$

f) P is *(finitely) subadditive*: for all A and B, disjoint or not,

$$P(A \cup B) \leq P(A) + P(B).$$

1.1.4. Random Variables

A random variable is a quantity that is measured in connection with a random experiment. If Ω is a sample space, and the outcome of the experiment is w, a measuring process is carried out to obtain a number $X(w)$. Then a random variable is a real valued function on a sample space. For example, the experiment is to throw a coin 15 times, and X be the number of heads. We take $\Omega = \{all\ sequencesof\ length\ 15\ with\ components\ H(head),\ T(tail)\}$ For this Ω we have 2^{15} points all together. A typical sample point is $w = \{HTHTHTHTTTTTTTT\}$, so for this point $X(w) = 4$.

Aa another example, consider picking a person at random from a certain city and measure his/her height and weight, this is the set of all pairs (x, y), where x

denotes the height and y denotes the weight of this person. Let X be the ratio of height to weight; that is $X(w) = \frac{x}{y}$.
If we are interested in a random variable X defined on a given sample space we generally want to know the probability of events involving X.
If X is a random variable on the probability space $(\Omega, \mathcal{F}, P)$, we are generally interested in calculating probabilities of events involving X, that is we generally want to know $P(w : X(w) \in \mathcal{B})$ for various Borel sets $\mathcal{B}$.

Definition 1.4 *A random variable is a function $X : \Omega \to \mathbb{R}$ with the property that* inverse image under X *of* $B \in \mathcal{B}(\mathbb{R})$ *is the subset of* Ω *given by*

$$X^{-1}(B) = \{\omega : X(\omega) \in B\},$$

Definition 1.5 *A random variables is a function $X : \Omega \to \mathbb{R}$ such that $X^{-1}(B) \in \mathcal{F}$ for every $B \in \mathcal{B}(\mathbb{R})$.*

Definition 1.6 ***(A σ-algebra Generated by a Random Variable:)*** *The family of events that are inverse images of Borel sets under a random variable is a σ-algebra on Ω.*

The way in which probabilities are calculated will depend on the particular nature of X. The random variable X is said to be "discrete" if and only if the set of possible values of X is finite. In this case, if $\{x_1, x_2, ..., x_n\}$ are the values of X that belong to $\mathcal{B}$ then

$$\begin{aligned} P(X \in \mathcal{B}) &= P(X = x_1, X = x_2, ...) \\ &= \sum_{x \in B} P(X = x_i) = \sum_{x \in B} p_X(x_i) \end{aligned}$$

where $p_X(x)$ is the probability mass function of X, defined by $p_X(x) = P(X = x)$.

1.1.5. Distribution Function

Definition 1.7 *The distribution function of a random variable X is the function $F : \mathbb{R} \to [0, 1]$ given by $F(x) = P(X \le x)$.*
If the random variable is discrete then:

$$F(x) = P(X \le x) = \sum_{x_i \le x} p_X(x_i)$$

where $p_X(x_i)$ is a probability mass function. If the random variable is continuous then:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

where $f(x)$ is the probability density function. Distribution Function satisfies the following conditions:

1) If $a \leq b$ then $F(a) \leq F(b)$

2) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1$

Proof 1) $A = \{x : X < a\}$ and $B = \{x : X < b\}$ since

$$\begin{aligned} A \subseteq B \text{ then } P(A) &\leq P(B) \text{ which is} \\ P(X \leq a) &\leq P(X \leq b) \text{ then} \\ F(a) &\leq F(b). \end{aligned}$$

2) $\lim_{x \to -\infty} P(X \leq x) = P(\emptyset) = 0$

$\lim_{x \to \infty} P(X \leq x) = P(\Omega) = 1.$ □

1.1.6. Expected Value

Definition 1.8 *Expected value is an averaging process for random variables. If random variable X is discrete then expected value of X is given by*

$$E[X] = \sum_{x} x p_X(x_i).$$

Similarly, if X is continuous random variable with probability density function f, then the expected value of X is defined as follows:

$$E[X] = \int_{-\infty}^{\infty} x f(x)dx.$$

Expected value has the following properties:

1. Constants preserved: For any constant c, then $E[cX] = cE[X]$.

2. Monotonicity: If $X \leq Y$, then $E[X] \leq E[Y]$.

3. Linearity: For $a, b \in \mathbb{R}$, $E[aX + bY] = aE[X] + bE[Y]$.

4. Continuity: If $X_n \to X$, then $E[X_n] \to E[X]$.

5. Separation: For any independent two random variables X and Y,

$$E[XY] = E[X]E[Y].$$

1.1.7. Variance

Definition 1.9 *Let X be a random variable on* $(\Omega, \mathcal{F}, P)$. *If* $k > 0$, *the number* $E[X^k]$ *is called the k'th. moment of X. If* $k = 1$ *then* $E[X]$ *is called the mean of X. The variance of X is defined as follows:*

$$Var[X] = E[(X - E[X])^2]$$

Variance has the following properties:

1.$Var[X] = E[X^2] - (E[X])^2$.

2.$Var[cX] = c^2 Var[X]$.

CHAPTER 2

TIME SCALE

A time scale is an arbitrary nonempty closed subset of the real numbers. In 1988 the calculus of time scales was initially introduced by Stefan Hilger in his Ph.D thesis. Stefan Hilger and his supervisor Bernd Auldbach have constructed time scale in order to create a theory that can unify discrete and continuous analysis. In this section we are going to introduce the theory of time scale.

Definition 2.1 *A time scale is a closed subset of the real numbers. Thus,* $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{N}$ *and union of any closed intervals such as* $[5, 7] \cup [8, 11]$, *or any closed intervals plus some single points for instance* $[4, 6] \cup [9, 14] \cup \{16, 19\}$, *are some examples of time scales, whereas* $\mathbb{Q}$ *and* $\mathbb{C}$ *are not time scales. Because they are any subsets of real numbers that are not closed or not union of closed intervals or single points. Time scale is denoted by the symbol* $\mathbb{T}$.

Definition 2.2 *Let* $\mathbb{T}$ *be a time scale. We define the forward jump operator* $\sigma : \mathbb{T} \to \mathbb{T}$ *by*

$$\sigma(t) := \inf\{s \ \in \ \mathbb{T} : s > t\}$$

and the backward jump operator $\rho : \mathbb{T} \to \mathbb{T}$ *by*

$$\rho(t) := \sup\{s \ \in \ \mathbb{T} : s < t\}.$$

If $\sigma(t) > t$, *t is called right-scattered, if* $\rho(t) < t$, *t is called left-scattered, if the points that are both right-scattered and left-scattered are called isolated, if* $\sigma(t) = t$, *then t is called right-dense, if* $\rho(t) = t$, *then t is called left-dense, if the points that are both right-dense and left-dense are called dense points.*
For a special case if $t = \max \mathbb{T}$, $\sigma(t) = t$ *and if* $t = \min \mathbb{T}$, $\rho(t) = t$.

Definition 2.3 *The function* $\mu : \mathbb{T} \to [0, \infty)$ *is defined by*

$$\mu(t) := \sigma(t) - t.$$

is called graininess function.

2.1. Derivative on Time Scale

Now we consider a function $f : \mathbb{T} \to \mathbb{R}$ and define so-called delta or Hilger derivative of f at a point $t \in \mathbb{T}^k$. $\mathbb{T}^k$ is a new term derived from $\mathbb{T}$ if $\mathbb{T}$ has a left scattered maximum m, then $\mathbb{T}^k = \mathbb{T} - \{m\}$. Otherwise, $\mathbb{T}^k = \mathbb{T}$.

Definition 2.4 *Assume $f : \mathbb{T} \to \mathbb{R}$ is a function and let $t \in \mathbb{T}^k$. Then we define delta derivative $f^{\Delta}(t)$ to be the number provided it exists with the property that given any $\epsilon > 0$, there is a neighborhood U of t (i.e, $U = (t - \delta, t + \delta) \cap \mathbb{T}$ for some $\delta > 0$) such that*

$$|(f(\sigma(t)) - f(s)) - f^{\Delta}(t)(\sigma(t) - s)| \leq \epsilon|\sigma(t) - s| \quad for \;\; all \;\; s \in U.$$

We call $f^{\Delta}(t)$ the delta derivative of f at t. Moreover, we say that f is Δ-differentiable on $\mathbb{T}^k$ provided $f^{\Delta}(t)$ exists for all $t \in \mathbb{T}^k$. The function $f^{\Delta} : \mathbb{T}^k \to \mathbb{R}$ is then called the delta derivative of f on $\mathbb{T}^k$.

Theorem 2.1 *Assume $f : \mathbb{T} \to \mathbb{R}$ is a function and let $t \in \mathbb{T}^k$. Then we have the following:*

(i) *If f is differentiable at t, then f is continuous at t.*

(ii) *If f is continuous at t and t is right-scattered, then f is differentiable at t with*

$$f^{\Delta}(t) = \frac{f(\sigma(t)) - f(t)}{\mu(t)}$$

(iii) *If t is right-dense, then f is differentiable at t if and only if the limit*

$$\lim_{s \to t} \frac{f(t) - f(s)}{t - s}$$

exist as a finite number, in this case

$$f^{\Delta}(t) = \lim_{s \to t} \frac{f(t) - f(s)}{t - s}$$

.

(iv) *If f is differentiable at t, then $f(\sigma(t)) = f(t) + \mu(t)f^{\Delta}(t)$.*

2.2. Integration on Time Scale

Definition 2.5 *A function is called a rd − continuous if it is continuous at all right dense points and its left sided limits exists at all left dense points.* C_{rd} *denotes the class of rd − continuous functions.*

Definition 2.6 *A function* $F : \mathbb{T} \to \mathbb{R}$ *is called an* antiderivative *of* $f : \mathbb{T} \to \mathbb{R}$ *provided* $F^{\Delta}(t) = f(t)$ *holds for every* $t \in \mathbb{T}^{\kappa}$*. The* indefinite integral *is defined by*

$$\int f(t)\Delta t = F(t) + c,$$

where c is arbitrary constant. We define the Cauchy integral by

$$\int_a^b f(t)\Delta t = F(b) - F(a), \quad for\ all\ a, b \in \mathbb{T}.$$

Theorem 2.2 *Let* $F : \mathbb{T} \to \mathbb{R}$ *be an rd-continuous function, for* $t_0 \in \mathbb{T}$ *and* $t \in \mathbb{T}^{\kappa}$

$$F(t) := \int_{t_0}^{t} f(\tau)\Delta\tau$$

is antiderivative of f.

Proposition 2.1 *If* $F : \mathbb{T} \to \mathbb{R}$ *is rd-continuous function and* $t \in \mathbb{T}^{\kappa}$*, then*

$$\int_t^{\sigma(t)} f(s)\Delta s = \mu(t)f(t).$$

Proof (Bohner and Peterson 2001). □

Theorem 2.3 *Let* $a, b \in \mathbb{T}$ *and* $f \in C_{rd}$*.*

(i) If $\mathbb{T} = \mathbb{R}$*, then*

$$\int_a^b f(t)\Delta t = \int_a^b f(t)dt,$$

where the integral on the right hand side is the usual Riemann integral from calculus.

(ii) If $\mathbb{T} = \mathbb{Z}$*, then*

$$\int_a^b f(t)\Delta t = \begin{cases} \sum_{t=a}^{b-1} f(t) & , \quad a < b\,; \\ 0 & , \quad a = b; \\ -\sum_{t=a}^{b-1} f(t)\ , & \quad a > b. \end{cases} \tag{2.1}$$

(iii) If $[a,b]$ *consists only isolated points, then*

$$\int_a^b f(t)\Delta t = \begin{cases} \sum_{t\epsilon[a,b)} f(t)\mu(t) & , \quad a < b\,; \\ 0 & , \quad a = b; \\ -\sum_{t\epsilon[b,a)} f(t)\mu(t) & , \quad a > b. \end{cases} \tag{2.2}$$

Example 2.1 *Let us compute* $\int_0^2 2^t \Delta t$ *on the time scales* $\mathbb{T} = \mathbb{R}$, $\mathbb{T} = \mathbb{Z}$, *and* $\mathbb{T} = \frac{1}{2}\mathbb{Z}$.
If $\mathbb{T} = \mathbb{R}$, *then*

$$\int_0^2 2^t \Delta t = \frac{2^2 - 2^0}{\ln 2} = \frac{3}{\ln 2}$$

If $\mathbb{T} = \mathbb{Z}$, *then*

$$\int_{t_0}^2 2^t \Delta t = \sum_{t=0}^{1} 2^t = 2^0 + 2^1 = 3$$

If $\mathbb{T} = \frac{1}{2}\mathbb{Z}$, *then*

$$\int_0^2 2^t \Delta t = \sum_{t=0}^{\frac{3}{2}} 2^t \frac{1}{2} = \frac{1}{2}(2^0 + 2^{\frac{1}{2}} + 2^1 + 2^{\frac{3}{2}}) = \frac{3 + \sqrt{2} + \sqrt{8}}{2}.$$

2.2.1. Double Integration on Time Scale

Integration on time scale was studied by G.Guseinov in 2003 (Guseinov 2003). Multiple Lebesgue integration on time scale was introduced by M. Bohner and G. Guseinov in 2006 (Bohner and Guseinov 2006). The relationship between Riemann and Lebesgue multiple integrals is investigated in that paper.

Definition 2.7 *Let* $\mathbb{T}_1$ *and* $\mathbb{T}_2$ *be two time scales. For* $i = 1, 2$ *let* σ_i, ρ_i *and* Δ_i *denote the forward jump operator, backward jump operator, and the delta differentiation operator, respectively on* $\mathbb{T}_i$. *Suppose* $a < b$ *are points in* $\mathbb{T}_1$, $c < d$ *are points in* $\mathbb{T}_2$. $[a,b)$ *be a half closed bounded interval in* $\mathbb{T}_1$, *and* $[c,d)$ *be a half closed bounded interval in* $\mathbb{T}_2$. *Let us introduce a rectangle in* $\mathbb{T}_1 \times \mathbb{T}_2$ *by*

$$R = [a,b) \times [c,d) = \{(t,s) : t\,\epsilon[a,b), s\,\epsilon[c,d)\}.$$

First we define Riemann integrals over rectangles R. A partition of $[a,b)$ *is any ordered subset* $P = \{t_0, t_1, ...t_n\} \subset [a,b]$ *where* $a = t_0 < t_1 < ... < t_n = b$. *Similarly,* $S = \{s_0, s_1, ...s_k\} \subset [c,d]$ *where* $c = s_0 < s_1 < ... < s_k = d$. *The number n and k are arbitrary positive integers. The collection of the intervals*

$$P_i = \{[t_{i-1}, t_i); 1 \leq i \leq n\}$$

$$S_j = \{[s_{j-1}, s_j); 1 \le j \le k\}$$

is called a $\Delta-$*partition of* $[a,b)$ *and* $[c,d)$ *respectively. Let us set*

$$R_{ij} = [t_{i-1}, t_i) \times [s_{j-1}, s_j), \quad \textit{where} \quad 1 \le i \le n,\ 1 \le j \le k.$$

We call the collection

$$P = \{R_{ij} : 1 \le i \le n,\ 1 \le j \le k\}$$

a $\Delta-$*partition of* R*. The set of all* $\Delta-$*partitions of* R *is denoted by* $P(R)$*. Let* f *be a bounded function we set*

$$M = \sup\{f(t,s) : (t,s) \in R\} \quad \textit{and} \quad m = \inf\{f(t,s) : (t,s) \in R\}$$

and for $1 \le i \le n,\ 1 \le j \le k.$

$$M_{ij} = \sup\{f(t,s) : (t,s) \in R_{ij}\} \quad \textit{and} \quad m_{ij} = \inf\{f(t,s) : (t,s) \in R_{ij}\}.$$

The upper Darboux Δ*-sum* $U(f,P)$ *and the lower Darboux* Δ*-sum* $L(f,P)$ *of* f *with respect to* P *are defined by*

$$U(f,P) = \sum_{n=1}^{n} \sum_{j=1}^{k} M_{ij}(t_i - t_{i-1})(s_j - s_{j-1}), \quad L(f,P) = \sum_{n=1}^{n} \sum_{j=1}^{k} m_{ij}(t_i - t_{i-1})(s_j - s_{j-1})$$

The upper Darboux integral $U(f)$ *of* f *from* a *to* b *defined by*

$$U(f) = \inf\{U(f,P)\}$$

and the lower Darboux integral $L(f)$ *of* f *from* a *to* b *defined by*

$$L(f) = \sup\{U(f,P)\}.$$

Definition 2.8 f *is said to be* Δ*-integrable over* R *if* $L(f) = U(f)$*. In this case, we write* $\int\int_R f(t)\Delta_1 t \Delta_2 s$ *for this common value and call this integral the Riemann* Δ*-integral.*

Definition 2.9 *A bounded function* f *on* R *is* Δ*-integrable if and only if for each* $\epsilon > 0$*, there exists* $\delta > 0$ *such that*

$$U(f,P) - L(f,P) < \epsilon.$$

for all partitions $P \in P_\delta(R)$*.*

Definition 2.10 *Let* $E \subset \mathbb{T}_1 \times \mathbb{T}_2$ *be a bounded set and* f *be a bounded function defined on the set* E*. Let* $R = [a, b) \times [c, d) \subset \mathbb{T}_1 \times \mathbb{T}_2$ *be a rectangle containing* E *(obviously such a rectangle exist) and define* F *on* R *as follows:*

$$F(t,s) = \begin{cases} f(t,s), & \text{if } (t,s) \in E \\ 0, & \text{if } (t,s) \in R \setminus E \end{cases}$$

Then f *is said to be Riemann* $\Delta-$*integrable over* E *if* F *is Riemann* $\Delta-$*integrable over* R*, we write*

$$\int_E \int f(t,s)\Delta_1 t \Delta_2 s = \int_R \int F(t,s)\Delta_1 t \Delta_2 s.$$

(Bohner and Guseinov 2006).

CHAPTER 3

MEASURE ON TIME SCALE

In this chapter we will introduce the concept of $\Delta-$Measure. $\Delta-$measure and $\nabla-$measure were first defined by Guseinov in 2003 (Guseinov 2003). Then in further study, the relationship between Lebesgue $\Delta-$integral and Riemann $\Delta-$integral were introduced in detail by Guseinov and Bohner (Bohner and Guseinov 2003). Delta measurability of sets was studied by Rzezuchousky (Rzezuchousky 2005). And after that, Cabada and Vivero (Cabada and Vivero 2004) were developed this concept.

3.1. Definition of $\Delta-$Measure

Let $\mathbb{T}$ be a time scale, σ and ρ be the forward and the backward jump functions defined on $\mathbb{T}$ respectively as given in Definition 2.2. We shall denote by $\mathcal{F}_1$ the class of all bounded left closed and right open intervals of the form

$$\mathcal{F}_1 = \{[a, b) \cap \mathbb{T} : a, b \in \mathbb{T}, a \leq b\}.$$

It is easy to see that if $a = b$, then interval is equal to the empty set.
We define the function $m_1 : \mathcal{F}_1 \to [0, +\infty)$ that assigns each interval $[a, b)$ to its length such that,

$$m_1([a, b)) = b - a. \tag{3.1}$$

where $\mathcal{F}_1$ is the set of all semi-closed intervals. The set function m_1 is countably additive measure on $\mathcal{F}_1$. Indeed,

1) $m_1([a, b)) = b - a \geq 0$ since $b \geq a$.

2) $m_1(\bigcup_{i=1}^{\infty}[a_i, b_i)) = \sum_{i=1}^{\infty} m_1([a_i, b_i))$ for pairwise disjoint intervals $[a_i, b_i)$.

3) $m_1(\emptyset) = m_1([a, a)) = a - a = 0$ holds for all $a \in \mathbb{T}$.

For describing the Caratheodory extension μ_Δ of m_1, first we will discuss the outer measure on all subsets of $\mathbb{T}$ by using m_1.

Definition 3.1 *Let* E *be any subset of* $\mathbb{T}$*. If there exists at least one finite or countable system of intervals* $I_j \in \mathcal{F}_1$ $(j = 1, 2, ...)$ *such that* $E \subset \bigcup_j I_j$*, then*

$$m_1^*(E) = \inf_{E \subset \bigcup_j I_j} \sum_j m_1(I_j)$$

is called the outer measure of the set E, where the infimum is taken over all coverings of E by a finite or countable system of intervals $I_j \in \mathcal{F}_1$*.*

The outer measure is always nonnegative and also could be infinite so that in general we have $0 \leq m_1^*(E) \leq \infty$.

Definition 3.2 *A set* $A \subset \mathbb{T}$ *is called* m_1^**-measurable or* Δ*-measurable if the following equality,*

$$m_1^*(E) = m_1^*(E \cap A) + m_1^*(E \cap A^c)$$

holds for all subsets E of $\mathbb{T}$*.*

Now defining the family

$$M(m_1^*) = \{ A \subset \mathbb{T} : A \text{ is } \Delta - \text{measurable}\},$$

the Lebesgue $\Delta-$measure, denoted by μ_Δ, is the restriction of m_1^* to $M(m_1^*)$.

Proposition 3.1 *The family* $\mathrm{M}(m_1^*)$ *of all* m_1^**-measurable* $(\Delta - \text{measurable})$ *subsets of* $\mathbb{T}$ *is a* σ*-algebra.*

Proposition 3.2 *If* E_n *is an increasing sequence of sets in a time scale* $\mathbb{T}$*, then*

$$\mu_\Delta(\cup_{n=1}^{\infty} E_n) = \lim_{n\to\infty} \mu_\Delta(E_n).$$

Similarly, if $\{E_n\}$ *is a decreasing sequence of sets in a time scale* $\mathbb{T}$*, then*

$$\mu_\Delta(\cap_{n=1}^{\infty} E_n) = \lim_{n\to\infty} \mu_\Delta(E_n).$$

In other words, Δ*-measure on* $\mathbb{T}$ *is continuous.*

Theorem 3.1 *For each* $t_0 \epsilon \mathbb{T} - \{max\mathbb{T}\}$ *the single point set* $\{t_0\}$ *is* $\Delta-$*measurable and its* $\Delta-$*measure is given by* $\mu_\Delta(\{t_0\}) = \sigma(t_0) - t_0$*.*

Proof **Case1.** Let $\{t_0\}$ be right scattered. Then $\{t_0\} = [t_0, \sigma(t_0)) \epsilon \mathcal{F}_1$. So $\{t_0\}$ is $\Delta-$measurable and $\mu_\Delta(\{t_0\}) = \sigma(t_0) - t_0$.

Case2. Let $\{t_0\}$ be right dense. Then there exists a decreasing sequence $\{t_k\}$ of points of $\mathbb{T}$ such that $t_0 \leq t_k$ and $t_k \downarrow t_0$. Since $\{t_0\} = \bigcap_{k=1}^{\infty}[t_0, t_k) \epsilon \mathcal{F}_1$. Therefore $\{t_0\}$ is $\Delta-$measurable. By *Preposition* 3.2

$$\mu_\Delta(\cap_{k=1}^{\infty}[t_0, t_k)) = \lim_{k\to\infty} \mu_\Delta([t_0, t_k)) = \lim_{k\to\infty} t_k - t_0 = 0.$$

which is the desired result. □

Every kind of interval can be obtained from an interval of the form $[a, b)$ by adding and subtracting the end points a and b. Then each interval of $\mathbb{T}$ is $\Delta-$measurable.

Theorem 3.2 *If $a, b \in \mathbb{T}^k$ and $a \leq b$, then*

a) $\mu_\Delta([a, b)) = b - a.$

b) $\mu_\Delta((a, b)) = b - \sigma(a).$

If $a, b \in \mathbb{T}$ and $a \leq b$, then

c) $\mu_\Delta((a, b]) = \sigma(b) - \sigma(a).$

d) $\mu_\Delta([a, b]) = \sigma(b) - a.$

Proof (Guseinov 2003). □

Definition 3.3 *We say that f is $\Delta-$measurable if for ever $\alpha \in \mathbb{R}$ the set*

$$f^{-1}([-\infty, \alpha)) = \{t \in \mathbb{T} : f(t) < \alpha\}$$

is $\Delta-$measurable.

Theorem 3.3 ***Monotone Convergence Theorem Adapted to Time Scale:*** *Let $\{f_n\}_{n\in N}$ be an increasing sequence of nonnegative $\Delta-$measurable functions defined on a $\Delta-$measurable set E and let $f(t) = \lim f_n(t)$ $\Delta-$a.e., then*

$$\int_E \lim f(s)\Delta s = \lim \int_E f_n(s)\Delta s$$

Proof As in (Aliprantis and Burkinshaw 1998). □

CHAPTER 4

PROBABILITY THEORY ON TIME SCALE

Probability Theory involves discrete and continuous random variables. In this chapter we introduce $\Delta-$Probability Measure which unifies both continuous and discrete cases. Also some distribution functions are adapted to time scale. Uniform and Normal random variables are constructed on time scale. Some *Mathematica* applications of Normal random variable on time scale are given. Very important inequalities on probability theory are also adapted to time scale.

4.1. Random Variable on Time Scale

Let $X_{\mathbb{T}}$ be a real valued function from $\Omega_{\mathbb{T}}$ to $\mathbb{T}$ such that $X_{\mathbb{T}} : \Omega_{\mathbb{T}} \to \mathbb{T}$. The inverse image under $X_{\mathbb{T}}$ of $B_{\mathbb{T}} \in \mathcal{B}(\mathbb{T})$ is the subset of $\Omega_{\mathbb{T}}$ given by

$$X_{\mathbb{T}}^{-1}(B_{\mathbb{T}}) = \{w : X_{\mathbb{T}}(w)\, \epsilon B_{\mathbb{T}},\ w\, \epsilon\, \Omega_{\mathbb{T}}\}$$

For example, let $\Omega_{\mathbb{T}} = [1,2] \cup \{3\} \cup [5,7]$ be a sample space of an experiment, and $X(w) = 2w + 1$ be a random variable on $\Omega_{\mathbb{T}}$, so,

$$X(\Omega_{\mathbb{T}}) = [3,5] \cup \{7\} \cup [11,15].$$

Definition 4.1 *A random variable is a function $X_{\mathbb{T}} : \Omega_{\mathbb{T}} \to \mathbb{T}$ such that $X_{\mathbb{T}}^{-1}(B_{\mathbb{T}}) \in \mathcal{F}_1$ for every $B_{\mathbb{T}} \in \mathcal{B}(\mathbb{R})$.*

4.2. $\Delta-$Probability

$\Delta-$Probability is a real valued set function P that assigns to each event A in the sample space $\Omega_{\mathbb{T}}$ a number $P_{\Delta}(A)$, called the probability of the event A. We define $\Delta-$Probability function on a Time Scale as follows:

$$P_{\Delta}(A) = \begin{cases} \frac{\mu_{\Delta}(A)}{\mu_{\Delta}(\Omega_{\mathbb{T}})}, & \text{if } A \subset \Omega_{\mathbb{T}} \\ 0, & \text{otherwise.} \end{cases}$$

Since this function satisfies the following conditions:

i) $P_{\Delta}(A) = \frac{\mu_{\Delta}(A)}{\mu_{\Delta}(\Omega_{\mathbb{T}})} \geq 0$ by the definition of Δ-measure.

ii) $P_\Delta(\Omega_\mathbb{T}) = 1$ since, $P_\Delta(\Omega_\mathbb{T}) = \frac{\mu_\Delta(\Omega_\mathbb{T})}{\mu_\Delta(\Omega_\mathbb{T})} = 1$.

iii) Let A_i be pairwise disjoint subsets of Ω_T;

$$\begin{aligned} P_\Delta(\bigcup_{i=1}^{\infty} A_i) &= \frac{\mu_\Delta(\bigcup_{i=1}^{\infty} A_i)}{\mu_\Delta(\Omega)} \\ &= \frac{\mu_\Delta(A_1)}{\mu_\Delta(\Omega_\mathbb{T})} + \frac{\mu_\Delta(A_2)}{\mu_\Delta(\Omega_\mathrm{T})} + ... + \frac{\mu_\Delta(A_n)}{\mu_\Delta(\Omega_\mathrm{T})} + ... \\ &= P_\Delta(A_1) + P_\Delta(A_2) + ... + P_\Delta(A_n) + ... \\ &= \sum_{i=1}^{\infty} P_\Delta(A_i) \end{aligned}$$

Then $P_\Delta(A)$ is a probability function on a time scale.

Definition 4.2 *A random variable $X_\mathbb{T} = a$ on time scale $\mathbb{T}$ is called almost surely $(a.s)$ if*

$$P_\Delta(X_\mathbb{T} \neq a) = 0.$$

4.2.1. Properties of the $\Delta-$Probability

Theorem 4.1 *Let $(\Omega_\mathbb{T}, \mathcal{F}_1, P)$ be a probability space. Then,*

1) $P_\Delta(\varnothing) = 0$.

2) P_Δ is finitely additive: if $A_1, \ldots, A_n$ are (pairwise) disjoint, then

$$P_\Delta(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P_\Delta(A_i).$$

3) For each A, $P_\Delta(A^c) = 1 - P_\Delta(A)$.

4) If $A \subseteq B$, then

$$P_\Delta(B \backslash A) = P_\Delta(B) - P_\Delta(A).$$

5) P_Δ is monotone*: if $A \subseteq B$ then $P_\Delta(A) \leq P_\Delta(B)$.*

6) For all A and B (disjoint or not),

$$P_\Delta(A \cup B) + P_\Delta(A \cap B) = P_\Delta(A) + P_\Delta(B).$$

7) P_Δ is (finitely) subadditive*: for all A and B, disjoint or not,*

$$P_\Delta(A \cup B) \leq P_\Delta(A) + P_\Delta(B).$$

Proof 1) If we use the definition of $\Delta-$Probability
$P_\Delta(\varnothing) = P_\Delta([a, a)) = \frac{\mu_\Delta([a,a))}{\mu_\Delta(\Omega_\mathbb{T})} = \frac{a-a}{\mu_\Delta(\Omega_\mathbb{T})} = 0$.

2) This property can be showed easily by using Axiom3 of being $\Delta-$Probability.

3) We know that $A \cup A^c = \Omega$,

$1 = P_\Delta(\Omega) = P_\Delta(A \cup A^c) = P_\Delta(A) + P_\Delta(A^c)$.

4) Since A and $B \cap A^c$ are disjoint sets,we can write $B = A \bigcup (B \bigcap A^c)$,

Take the $\mu_\Delta$measure of both sides and divide by $\mu_\Delta(\Omega_\mathbb{T})$ then we get

$$\begin{aligned} \frac{\mu_\Delta(B)}{\mu_\Delta(\Omega_\mathbb{T})} &= \frac{\mu_\Delta(A)}{\mu_\Delta(\Omega_\mathbb{T})} + \frac{\mu_\Delta(B \cap A^c)}{\mu_\Delta(\Omega_\mathbb{T})} \\ P_\Delta(B) &= P_\Delta(A) + P_\Delta(B \backslash A) \end{aligned}$$

5) We know that if $A \subseteq B$ then $\mu_\Delta(A) \leq \mu_\Delta(B)$.

and divide both sides by $\mu_\Delta(\Omega_\mathrm{T})$ then we get

$$\begin{aligned} \frac{\mu_\Delta(A)}{\mu_\Delta(\Omega_\mathbb{T})} &\leq \frac{\mu_\Delta(B)}{\mu_\Delta(\Omega_\mathbb{T})} \\ P_\Delta(A) &\leq P_\Delta(B). \end{aligned}$$

6) Since

$$(A \cap B) \cup (A \backslash B) = A \tag{4.1}$$

$$(A \cap B) \cup (B \backslash A) = B \tag{4.2}$$

take the measure of both sides of (4.1) and (4.2) and divide by $\mu_\Delta(\Omega_\mathbb{T})$ then we get;

$$\begin{aligned} \frac{\mu_\Delta(A \cap B)}{\mu_\Delta(\Omega_\mathbb{T})} + \frac{\mu_\Delta(A \setminus B)}{\mu_\Delta(\Omega_\mathbb{T})} &= \frac{\mu_\Delta(A)}{\mu_\Delta(\Omega_\mathbb{T})} \\ \frac{\mu_\Delta(A \cap B)}{\mu_\Delta(\Omega_\mathbb{T})} + \frac{\mu_\Delta(B \setminus A)}{\mu_\Delta(\Omega_\mathbb{T})} &= \frac{\mu_\Delta(B)}{\mu_\Delta(\Omega_\mathbb{T})} \end{aligned}$$

then we have

$$P_\Delta(A \cap B) + P_\Delta(A \backslash B) = P_\Delta(A) \tag{4.3}$$

$$P_\Delta(A \cap B) + P_\Delta(B \backslash A) = P_\Delta(B) \tag{4.4}$$

By adding these equations we get

$$P_\Delta(A \cap B) + P_\Delta(A \backslash B) + P_\Delta(A \cap B) + P_\Delta(B \backslash A) = P_\Delta(A) + P_\Delta(B)$$

which is

$$P_\Delta(A \cap B) + P_\Delta(A \cup B) = P_\Delta(A) + P_\Delta(B)$$

7) This property can be showed easily by using Property 5. □

Example 4.1 *An experiment is performed. During a day, the first observation is taken at any time from midnight to one p.m., second observation is taken at any time from 2 to 3 p.m. and so on. The sample space is*
$\Omega_{\mathbb{T}} = [0,1] \cup [2,3] \cup ... \cup [22,23]$. *Let A be the event which observations have been either during* $[2,3]$ *or* $[8,9]$ *that is* $A = [2,3] \cup [8,9]$ *and these two intervals are disjoint, by using the Theorem 3.2 we can compute*

$$\mu_\Delta(\Omega_{\mathbb{T}}) = \mu_\Delta([0,1] \cup [2,3] \cup ... \cup [22,23]) = 23$$

$$P_\Delta(A) = P_\Delta([2,3]) + P_\Delta([8,9])$$
$$P_\Delta(A) = \frac{\mu_\Delta([2,3])}{\mu_\Delta(\Omega_{\mathbb{T}})} + \frac{\mu_\Delta([8,9])}{\mu_\Delta(\Omega_{\mathbb{T}})} = \frac{4}{23}$$

Example 4.2 *A bus stops at the station every day at some time between* $[0,1]$,$[8,9]$, $[12,13]$, $[16,17]$, $[18,19]$, $[20,21]$, $[22,23]$. *Evaluate the probability of the bus arrive the station during* $[1,9]$ *or* $[1,9)$ *or* $(1,9]$ *or* $(1,9)$.
Solution:

$$\Omega_{\mathbb{T}} = [0,1] \cup [8,9] \cup [12,13] \cup [16,17] \cup [18,19] \cup [20,21] \cup [22,23]$$

$$\mu_\Delta(\Omega_{\mathbb{T}}) = \mu_\Delta([0,1] \cup [8,9] \cup [12,13] \cup [16,17] \cup [18,19] \cup [20,21] \cup [22,23]) = 23$$

$$\begin{aligned} P_\Delta([1,9]) &= \frac{11}{23} \\ P_\Delta([1,9)) &= \frac{8}{23} \\ P_\Delta((1,9]) &= \frac{4}{23} \\ P_\Delta((1,9)) &= \frac{1}{23}. \end{aligned}$$

4.3. $\Delta-$Probability Density Function on Time Scale

$\Delta-$Probability Density Function on time scale satisfies the following conditions:

1) $f_\Delta(x) \geq 0$ for all $x \in \mathbb{T}$.
2) $\int_{-\infty}^{\infty} f_\Delta(x)\Delta x = 1$.

4.4. $\Delta-$Distribution Function

For the density function f_Δ on time scale $\Delta-$Distribution Function is defined as follows:

$$F_\Delta(x) = P_\Delta(X \le x) = \int_{-\infty}^{x} f_\Delta(t)\Delta t$$

$\Delta-$Distribution Function has the following conditions:

1)F is increasing function.

2)$\lim_{x\to-\infty} F_\Delta(x) = 0$ and $\lim_{x\to+\infty} F_\Delta(x) = 1$

where $F_\Delta(a) = P_\Delta(X \le a) = \int_{-\infty}^{a} f_\Delta(t)\Delta t$

Proof 1) $A = \{x : X < a\}$ and $B = \{x : X < b\}$ where $a < b$. Since $A \subset B$

$$\begin{aligned} P_\Delta(X \le a) &\le P_\Delta(X \le b) \ \ then \\ F_\Delta(a) &\le F_\Delta(b). \end{aligned}$$

2)$\lim_{x\to-\infty} P_\Delta(X \le x) = P_\Delta(\varnothing) = 0$

$\lim_{x\to+\infty} P_\Delta(X \le x) = P_\Delta(\Omega) = 1.$ □

4.5. Some $\Delta-$Distribution Functions

In this section we are going to adapt some well known distribution functions on time scale. Uniform and Normal Distributions are studied on time scale. In addition, some *Mathematica* applications for Normal Distribution on time scale are given.

4.5.1. Uniform Random Variable on Time Scale

Let $\mathrm{T} = [t_0, t_1] \cup [t_2, t_3] ... \cup [t_{2n}, t_{2n+1}]$ where $t_0 = a$ and $t_{2n+1} = b$ and $S_R = \{t_1, t_3, ..., t_{2n+1}\}$ be the set of all right scattered points of T. Uniform $\Delta-$probability function on T can be defined as follows:

$$f_\Delta(t) = \begin{cases} \frac{1}{\mu_\Delta([a,b))}, & \text{if } t \epsilon \mathrm{T} \\ 0, & \text{otherwise.} \end{cases}$$

$\mu_\Delta(\bigcup_{i=1}^{\infty}(A_i) = \sum_{i=1}^{\infty}\mu_\Delta(A_i)$ for pairwise disjoint sets (A_i)

$$\begin{aligned}\mu_\Delta(\bigcup_{i=1}^{\infty}A_i) &= \sum_{i=1}^{\infty}\mu_\Delta(A_i)\\ &= \sum_{i=1}^{\infty}\mu_\Delta([t_{i-1},t_i))\\ &= \mu_\Delta([a,b)).\end{aligned}$$

and since this function satisfies the following condition:

$$\begin{aligned}\int_{-\infty}^{\infty}f_\Delta(t)\Delta t &= \int_{-\infty}^{a}f_\Delta(t)\Delta t+\int_{a}^{b}f_\Delta(t)\Delta t+\int_{b}^{\infty}f_\Delta(t)\Delta t\\ &= \int_{a}^{t_1}f_\Delta(t)\Delta t+\int_{t_1}^{t_2}f_\Delta(t)\Delta t+\int_{t_2}^{t_3}f_\Delta(t)\Delta t+...+\\ &\quad\int_{t_{2n-1}}^{t_{2n}}f_\Delta(t)\Delta t+\int_{t_{2n}}^{t_{2n+1}}f_\Delta(t)\Delta t+\int_{t_{2n+1}}^{\sigma(t_{2n+1})}f_\Delta(t)\Delta t\end{aligned}$$

since $\int_t^{\sigma(t)} f(s)\Delta s = \mu(t)f(t)$, for scattered points t_i, where $i = 1,3,5,...,2n+1$.

$$\begin{aligned}&= \frac{1}{\mu_\Delta([a,b))}\Big\{\sum_{i=0}^{n}\mu_\Delta([t_{2i+1},\sigma(t_{2i+1}))) + \sum_{i=0}^{n}\mu_\Delta([t_{2i},t_{2i+1}))\Big\}\\ &= \frac{1}{\mu_\Delta([a,b))}[t_2-t_1+t_1-t_0+t_4-t_3+t_3-t_2+...+\sigma(t_{2n})-t_{2n-1}+t_{2n+1}-t_{2n}+b-b]\\ &= \frac{t_{2n+1}-t_0}{\mu_\Delta([a,b))} = \frac{b-a}{b-a} = 1.\end{aligned}$$

then $f_\Delta(t)$ is a probability density function on $\Omega_{\mathbb{T}}$. Also Uniform $\Delta-$Distribution function on time scale is defined as follows:

$$F_\Delta(t) = \begin{cases} 0, & t<a \text{ or } t\epsilon(t_i,\sigma(t_i))\ where\ t_i\ is\ right\ scattered\\ \frac{\mu_\Delta([a,t)\cap\Omega)}{\mu_\Delta([a,b))}, & \text{if } t\epsilon\text{T}\\ 1, & t\geq b\end{cases}$$

This function satisfies the all distribution function properties:

$$\begin{aligned}\lim_{t\to+\infty}F_\Delta(t) &= \lim_{t\to+\infty}\frac{\mu_\Delta([a,t)\cap\Omega)}{b-a}\\ &= \lim_{t\to+\infty}\frac{\mu_\Delta[a,t]_{\mathbb{T}}}{b-a}\\ &= \lim_{t\to+\infty}\frac{\sigma(t)-a}{b-a}\\ &= \frac{b-a}{b-a} = 1\end{aligned}$$

since a is right dense $\sigma(a) = a$ and $max(\mathbb{T}) = b = \sigma(b)$

$$\begin{aligned}
\lim_{t\to-\infty} F_\Delta(t) &= \lim_{t\to-\infty} \frac{\mu_\Delta([a,t)\cap\Omega)}{b-a} \\
&= \lim_{t\to-\infty} \frac{\mu_\Delta(\{a\})_{\mathbb{T}}}{b-a} \\
&= \frac{\sigma(a)-a}{b-a} \\
&= \frac{a-a}{b-a} = 0.
\end{aligned}$$

4.5.2. Normal Random Variable on Time Scale

Simon (Simon et al 2005) studied on Gaussian Bell for time scales. He introduced a function $f(x)$ such that

$$f(x) = c \prod_{t\epsilon[0,x)} (1+\mu(t))^{-t}.$$

We take this function as probability density function on time scale T_+ every $x \in T_+$ such that $T_+ = \{x \in T | x \geq 0\}$.

$f_\Delta(x) = c\prod_{t\in[0,x)}(1+\mu(t))^{-t}$ every $x \in T_+$

If we take $\mathbb{T} = hZ^+$ and $h > 0$, then this function satisfies the following condition:

$$\sum_{n=1}^{\infty} f_\Delta(hn) = 1$$

since

$f_\Delta(hn) = c\prod_{k=0}^{n-1}(1+h)^{-kh} = c(1+h)^{-h\sum_{k=0}^{n-1}k} = c(1+h)^{hn(1-n)/2}$ every $n \in N_0$.

By substituting $x = hn$ then we get

$$f_\Delta(x) = c[(1+h)^{1/h}]^{-x(x-h)/2} \tag{4.5}$$

and every $x \in T_+$ and for all functions of $f_\Delta(x)$ we can find the values of c's which satisfy the following condition

$$\sum_{n=1}^{\infty} f_\Delta(x) = \sum_{n=1}^{\infty} c[(1+h)^{1/h}]^{-x(x-h)/2} = 1$$

If we take $h = 1, 1/2, 1/4, 1/8, 1/32$ then by using *Mathematica* we can find the values of c respectively which are $c = 0.60915,\ c = 0.350264,\ c = 0.187143,$ $c = 0.0966356,\ c = 0.0248695$. Then $f_\Delta(x)$ is a probability density function on time

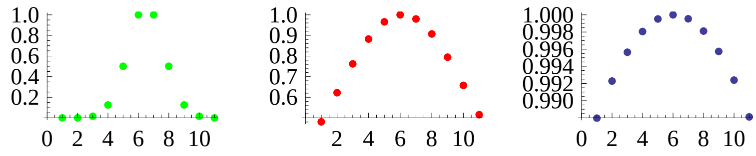

Figure 4.1. Graph of $f_\Delta(1n)$, $f_\Delta(\frac{1}{4}n)$, and $f_\Delta(\frac{1}{32}n)$

scale $\mathbb{T}$.

The graph of $f_\Delta(x)$ for $h = 1, 1/4, 1/32$ are respectively given as follows:

By observing Figure 4.1, it is easy to see that if $h \to 0$ then the function $f_\Delta(x)$ will coincide with the Gaussian Bell density function. Now, let us see this result by analytically.

Lemma 4.1 $\lim_{h\to 0} f_\Delta(x) = ce^{-x^2/2}$

Proof Let take $f_\Delta(x)$ in Equation (4.5)

$f_\Delta(x) = c[(1+h)^{1/h}]^{-x(x-h)/2}$

By taking the natural logarithm of both sides,

$\ln f_\Delta(x) = \ln c - \frac{x(x-h)}{2} \ln[(1+h)^{1/h}]$

by taking the limit of both sides when $h \to 0$

$$\begin{aligned}
\lim_{h\to 0} \ln f_\Delta(x) &= \lim_{h\to 0} \ln c - \frac{x(x-h)}{2} \ln[(1+h)^{1/h}] \\
&= \ln c - \lim_{h\to 0} \frac{x(x-h)}{2} \lim_{h\to 0} \ln[(1+h)^{1/h}] \\
&= \ln c - \frac{x^2}{2} \ln(e) \\
&= \ln c - \frac{x^2}{2} \\
\ln(\lim_{h\to 0} \ln f_\Delta(x)) &= e^{\ln c - x^2/2} \\
&= ce^{-x^2/2}.
\end{aligned}$$

□

We get Gaussian Probability density function.

4.6. $\Delta-$Expected Value on Time Scale

Expected value is an averaging process for random variables x_k's on time scales. The expectation gives an idea of the central tendency which is used as a parameter of location of the probability distribution of X.

Since the definition of $\Delta-$integral on time scale involves discrete and continuous case, and besides expected value on time scale involves discrete and continuous case, then the Expected value on time scale is given by

$$E_{\Delta}[X] = \int_{-\infty}^{\infty} x f_{\Delta}(x)\Delta x.$$

where f_{Δ} is a $\Delta-$probability density function on time scale.

Definition 4.3 (Simple random variable:) *The function I_{A_k} defined by*

$$I_{A_k} = \begin{cases} 1, & \text{if } w \in A_k \\ 0, & \text{otherwise}. \end{cases}$$

is called characteristic function of $\Omega_{\mathbb{T}}$ and a linear combination of the characteristic functions as

$$X = \sum_{k=1}^{n} x_k I_{A_k}$$

is called a simple random variable on time scale $\Omega_{\mathbb{T}}$ where A_k are pairwise disjoint sets with

$$A_k = \{w : X(w) = x_k\} \;\; i = 1, 2, ...n.$$

and $\bigcup_{i=1}^{n} A_k = \Omega_{\mathbb{T}}$.

The indefinite $\Delta -$ *integral* or indefinite $\Delta-$expectation of X over $A \in \mathcal{F}_1$ on hN is defined as follows:

$E_{\Delta}[XI_A] = \int_A X\Delta P = E_{\Delta}(\sum_k x_k I_{AA_k}) = \sum_k x_k P_{\Delta} AA_k.$

4.6.1. Properties of $\Delta-$Expected Value on Time Scale

Theorem 4.2

1) *If X and Y are jointly distributed random variables, then;*

$$E_{\Delta}[X + Y] = E_{\Delta}[X] + E_{\Delta}[Y]$$

where $a, b \in \mathbb{R}$.

2) *$E_\Delta[cX] = cE_\Delta[X]$.*

3) *If $X \geq 0$ then $E_\Delta[X] \geq 0$.*

4) *For any two independent jointly distributed random variables X and Y,*

$$E_\Delta[XY] = E_\Delta[X]E_\Delta[Y].$$

5) *If $X_n \to X$, then $E_\Delta[X_n] \to E_\Delta[X]$.*

Proof

1)

$$\begin{aligned} E_\Delta[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_\Delta(x,y)\Delta x\Delta y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_\Delta(x,y)\Delta x\Delta y + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_\Delta(x,y)\Delta x\Delta y \end{aligned}$$

chance of order

$$\begin{aligned} &= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_\Delta(x,y)\Delta y\Delta x + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_\Delta(x,y)\Delta x\Delta y \\ &= \int_{-\infty}^{\infty} x f_X(x)\Delta x + \int_{-\infty}^{\infty} y f_Y(y)\Delta y \\ &= E_\Delta[X] + E_\Delta[Y]. \end{aligned}$$

2) $E_\Delta[cX] = \int_{-\infty}^{\infty} cxf_\Delta(x)\Delta x = c\int_{-\infty}^{\infty} xf_\Delta(x)\Delta x = cE_\Delta[X]$.

3) Since $f_\Delta(x) \geq 0$ for all $x \,\varepsilon\, \mathbb{T}$. then $E_\Delta[X] = \int_{-\infty}^{\infty} xf_\Delta(x)\Delta x \geq 0$.

4) Let X and Y be any two independent jointly distributed random variables, then

$$E_\Delta[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\Delta x\Delta y$$

change of order

$$\begin{aligned} E_\Delta[XY] &= \int_{-\infty}^{\infty} xf_X(x)\int_{-\infty}^{\infty} yf_Y(y)\Delta y\Delta x \\ &= E_\Delta[Y]\int_{-\infty}^{\infty} xf_X(x)\Delta x \\ &= E_\Delta[X]E_\Delta[Y]. \end{aligned}$$

5) There exist a monotone increasing sequence of simple functions X_n converging to X. $E_\Delta[X_n]$ is a monotone increasing sequence which converges to a limit value related with X_n sequences. By using the Monotone Convergence Theorem on time scale,

$$\begin{aligned}
E_\Delta[X_n] &= \int_{-\infty}^{\infty} x_n f_{X_n}(x)\Delta x \\
&= \lim_{n\to\infty} \int_{-\infty}^{\infty} x_n f_{X_n}(x)\Delta x \\
&= \int_{-\infty}^{\infty} \lim_{n\to\infty} x_n f_{X_n}(x)\Delta x \\
&= \int_{-\infty}^{\infty} x f_X(x)\Delta x \\
&= E_\Delta[X].
\end{aligned}$$

□

Corollary 4.1 *If $X \leq Y$ a.s., then* $E_\Delta[X] \leq E_\Delta[Y]$.

Proof Since $Y > X$ a.s., $Y - X > 0$ a.s. and since 3.th property of the expected value
$E_\Delta[Y - X] \geq 0$. By linearity property
$E_\Delta[Y - X] = E_\Delta[Y] - E_\Delta[X] \geq 0$.
This implies $E_\Delta[X] \leq E_\Delta[Y]$. □

4.7. $\Delta-$Variance on Time Scale

Let X be a random variable on $(\Omega_{\mathbb{T}}, \mathcal{F}_1, P)$. If $k > 0$, the number $E_\Delta[X^k]$ is called the k.th moment of X. If k=1 then $E_\Delta[X]$ is called the $\Delta-$mean of X.
$Var_\Delta[X] = E[(X - E_\Delta[X])^2]$ is called the $\Delta-$Variance of X. $\Delta-$Variance on time scale has the following properties:
1. $Var_\Delta[X] = E_\Delta[X^2] - (E_\Delta[X])^2$.
2. $Var_\Delta[cX] = c^2 Var_\Delta[X]$.

4.7.1. Moments

$E_\Delta[X - a]^k$ $(k = 1, 2, 3...)$ is the k'th $\Delta-$moment of X about a. $E_\Delta[X]^k$ is called the k'th $\Delta-$moment about the origin.

$E_\Delta[|X-a|]^k$ $(r > 0)$ is called the r′th $\Delta-$absolute moment of X about a. $E_\Delta[|X|]^k$ is called the k.th $\Delta-$absolute moment.

$E_\Delta[(X - E_\Delta[X])^2]$ the second $\Delta-$moment about the $\Delta-$expected value is called the $\Delta-$variance of X on a time scale $\mathbb{T}$ and as we discussed above denoted by $Var_\Delta[X]$.

$$E_\Delta[(X - E_\Delta[X])^2] = E_\Delta[X^2] - (E_\Delta[X])^2 \geq 0.$$

$\Delta-$Moment Generating Function

If $\Delta-$moments of all orders exist and are finite then

$$M(\theta) = 1 + (E_\Delta[X])\theta + (E_\Delta[X^2])\frac{\theta^2}{2} + ... + (E_\Delta[X^n])\frac{\theta^n}{n}...$$

is called the $\Delta-$moment generating function of X and also

$$E_\Delta[X^n] = \frac{d^n M(\theta)}{d\theta^n}\Big|_{\theta=0}.$$

4.7.2. Probabilistic Inequalities on Time Scale

M. Bohner and R.Agarval (Agarwal and Bohner 2001) studied about inequalities on time scales. In this section we modify some important Probabilistic Inequalities on time scales.

Lemma 4.2 (C_r-Inequality on time scale)

$$E_\Delta|X+Y|^r \leq C_r E_\Delta|X|^r + C_r E_\Delta|Y|^r$$

$$C_r = \begin{cases} 1, & if\, r \leq 1 \\ 2^{r-1}, & if\, r > 1 \end{cases}$$

Proof If $a > 0$ and $b > 0$ then

$$\begin{aligned} \left(\frac{a}{a+b}\right)^r + \left(\frac{b}{a+b}\right)^r &\geq 1 \;\; for\, all \;\; r \leq 1 \\ a^r + b^r &> (a+b)^r \end{aligned}$$

Hence for all w, $a = |X(w)|$, $b = |Y(w)|$ and $r \leq 1$ we get

$$\begin{aligned} |X(w)|^r + |Y(w)|^r &\geq (|X(w)| + |Y(w)|)^r \\ &\geq |X(w) + Y(w)|^r. \end{aligned}$$

now let us consider the function

$$\Psi(p) = p^r + (1-p)^r \quad (r \geq 1) \;\; (0 < p < r)$$

This function has a minimum value when $p = \frac{1}{2}$. Thus $a > 0$ and $b > 0$,

$$\begin{aligned} \Big(\frac{a}{a+b}\Big)^r + \Big(\frac{b}{a+b}\Big)^r &\geq 2^{-(r-1)} \\ 2^{r-1}(a^r + b^r) &\geq (a+b)^r. \end{aligned}$$

So for all w and all $r > 1$

$$\begin{aligned} 2^{r-1}(|X(w)|^r + |Y(w)|^r) &\geq (|X(w)| + |Y(w)|)^r \\ &\geq |X(w) + Y(w)|^r. \end{aligned}$$

The proof is completed. □

Lemma 4.3 (Hölder's Inequality on time scale)

$$E_\Delta|XY| \leq \sqrt[r]{E_\Delta|X|^r}.\sqrt[s]{E_\Delta|Y|^s}$$

where $r > 1$ *and* $\frac{1}{r} + \frac{1}{s} = 1$.

Proof For nonnegative real numbers a and b, the basic inequality

$$a^{\frac{1}{r}}.b^{\frac{1}{s}} \leq \frac{a}{r} + \frac{b}{s} \tag{4.6}$$

holds. Now suppose without lost of generality, that

$$(E_\Delta|X|^r)^{1/r}.(E_\Delta|Y|^s)^{1/s} \neq 0.$$

Apply (4.6) to $a = |X(w)|^r/E_\Delta|X|^r$ and $b = |Y(w)|^s/E_\Delta|Y|^r$ and taking expectations of both sides then we get

$$\begin{aligned} E_\Delta\Big[\frac{|X(w)|}{(E_\Delta|X|^r)^{\frac{1}{r}}}.\frac{|Y(w)|}{(E_\Delta|Y|^s)^{\frac{1}{s}}}\Big] &\leq E_\Delta\Big[\frac{1}{r}\frac{|X(w)|^r}{E_\Delta|X|^r} + \frac{1}{s}\frac{|Y(w)|^s}{E_\Delta|Y|^s}\Big] \\ &= \frac{1}{r}\frac{E_\Delta|X(w)|^r}{E_\Delta|X|^r} + \frac{1}{s}\frac{E_\Delta|Y(w)|^s}{E_\Delta|Y|^s} \\ &= \frac{1}{r} + \frac{1}{s} = 1. \end{aligned}$$

This directly yields Hölder's Inequality adapted to time scales.

$$E_\Delta|X(w)Y(w)| \leq (E_\Delta|X|^r)^{\frac{1}{r}}.(E_\Delta|Y|^s)^{\frac{1}{s}}.$$

□

Corollary 4.2 (Cauchy-Schwartz Inequality on time scale) *If we put $r = s = 2$ then*

$$E_\Delta|XY| \leq \sqrt{E_\Delta|X|^2.E_\Delta|Y|^2}$$

is called Cauchy Schwartz Inequality on time scale.

Proof We can show the proof easily by using Lemma 4.3 above. □

Lemma 4.4 (Minkowski's Inequality on time scale)

$$\sqrt[r]{E_\Delta|X+Y|^r} \leq \sqrt[r]{E_\Delta|X|^r} + \sqrt[r]{E_\Delta|Y|^r}.$$

where $r \geq 1$.

Proof If we separate the expectation and applying Hölder's inequality

$$\begin{aligned}
E_\Delta|X+Y|^r &= E_\Delta(|X+Y|.|X+Y|^{r-1}) \\
&\leq E_\Delta(|X|.|X+Y|^{r-1}) + E_\Delta(|Y|.|X+Y|^{r-1}) \\
&\leq (E_\Delta|X|^r)^{\frac{1}{r}}.(E_\Delta|X+Y|^{(r-1)s})^{\frac{1}{s}} + (E_\Delta|Y|^r)^{\frac{1}{r}}.(E_\Delta|X+Y|^{(r-1)s})^{\frac{1}{s}} \\
&\leq (E_\Delta|X+Y|^{(r-1)s})^{\frac{1}{s}}\{(E_\Delta|X|^r)^{\frac{1}{r}} + (E_\Delta|Y|^r)^{\frac{1}{r}}\} \\
(E_\Delta|X+Y|^r)^{(1-\frac{1}{s})} &\leq (E_\Delta|X|^r)^{\frac{1}{r}} + (E_\Delta|Y|^r)^{\frac{1}{r}}
\end{aligned}$$

with $1 - \frac{1}{s} = \frac{1}{r}$ we get the desired result.

$$(E_\Delta|X+Y|^r)^{\frac{1}{r}} \leq (E_\Delta|X|^r)^{\frac{1}{r}} + (E_\Delta|Y|^r)^{\frac{1}{r}}.$$

□

Convex Function: Let f be a real valued Borel function defined on an open interval I which is finite or infinite subset of real numbers, is said to be convex if for every pair of points x_1, x_2 of I,

$$f\Big(\frac{x_1+x_2}{2}\Big) \leq \frac{1}{2}f(x_1) + \frac{1}{2}f(x_2).$$

An alternative definition of a convex function is that, for every $x_0 \epsilon I$, there exist a number $\lambda(x_0)$ such that for all $x \epsilon I$,

$$\lambda(x_0)(x - x_0) \leq f(x) - f(x_0). \tag{4.7}$$

Lemma 4.5 (Jensen's Inequality on time scale) *If f is convex and $E_\Delta[X]$ is finite, then*

$$f(E_\Delta[X]) \le E_\Delta[f(X)].$$

Proof Let X be a random variable whose values lie in I, replacing x_0 by $E_\Delta[X]$ and x by X in Inequality (4.7), then we have

$$\lambda(E_\Delta[X])(X(w) - E_\Delta[X]) \le f(X(w)) - f(E_\Delta[X]).$$

Taking expectations and left hand side vanished, so we get the desired result. □

Lemma 4.6 (Markov's Inequality on time scale) *If $X \ge 0$ and $a > 0$*

$$P_\Delta(X \ge a) \le \frac{E_\Delta[X]}{a}.$$

Proof If we separate the $\Delta-$integral

$$\begin{aligned} E_\Delta[X] = \int_0^\infty x f_\Delta(x)\Delta x &= \int_0^a x f_\Delta(x)\Delta x + \int_a^\infty x f_\Delta(x)\Delta x \\ &\ge \int_a^\infty x f_\Delta(x)\Delta x \end{aligned}$$

Since $x f_\Delta(x) \ge a f_\Delta(x)$ when $x \ge a$,

$$\begin{aligned} &\ge \int_a^\infty a f_\Delta(x)\Delta x \\ &= a\int_a^\infty f_\Delta(x)\Delta x \\ &= aP_\Delta(X \ge a) \\ E_\Delta[X] &\ge aP_\Delta(X \ge a) \\ P_\Delta(X \ge a) &\le \frac{E_\Delta[X]}{a}. \end{aligned}$$

□

Lemma 4.7 (Chebyshev's Inequality on time scale) *Let $X \ge 0$ be a random variable with mean $E_\Delta[X]$, and $\Delta-$variance $Var_\Delta[X] = \sigma_\Delta^2$ and $a > 0$, then*

$$P_\Delta(|X - E_\Delta[X]| \ge k\sigma_\Delta) \le \frac{1}{k^2}$$

Proof Since $\frac{(X-E_\Delta[X])^2}{\sigma_\Delta^2} \ge 0$ (*nonnegative random variable*)

$$E_\Delta\Big[\frac{(X - E_\Delta[X])^2}{\sigma_\Delta^2}\Big] = \frac{E_\Delta[(X - E_\Delta[X])^2]}{\sigma_\Delta^2} = 1$$

by using Markov's inequality

$$P_\Delta\Big(\frac{(X - E_\Delta[X])^2}{\sigma_\Delta^2} \geq k^2\Big) \leq \frac{1}{k^2}$$

which means that

$$P_\Delta(|X - E_\Delta[X]| \geq k\sigma_\Delta) \leq \frac{1}{k^2}.$$

$\square$

REFERENCES

Agarwal R. and M. Bohner. 2001. Inequalities on Time Scale: A survey. *Math. Ineq. Apps* 4(4): 535-557.

Aliprantis, C.D. and Burkinshaw O. 1998. Principles of Real Analysis. Academic Press, San Diego.

Bohner, M. and Peterson A. 2001. *Dynamic Equations on time scales: An introduction with Applications* Birkhauser: Boston.

Bohner, M. and A.Peterson. 2003. *Advanced in Dynamic Equations on time scales* Birkhauser: Boston. Inc. MA.

Bohner, M. and Guseinov G., 2003. Riemann and Lebesgue Integration. Advances in Dynamic Equations on Time Scales. 117-163.

Bohner, M. and G.Guseinov. 2006. Multiple Lebesgue Integration on time scales. *J. Math. Anal. Appl.* 285: 107-127.

Cabada, A., and R.P.Vivero. 2004. Expression of the Lebesque $\Delta-$Integral on time scales as a usual Lebesgue Integral; Application to the Calculus of $\Delta-$Antiderivatives. *Elseiver* 4: 291-310.

Erge, L. Peterson A. and Simon M. 2005. Square Integrability of Gaussian Bells on Time Scale. *Computers and Mathematics with Apps*. Elsevier. 49: 871-883.

Grimmett, G. and Stirzaker D. 2006.*Probability and random processes* Oxford University Press.

Guseinov, G. 2003. Integration on time scales. *J. Math. Anal. Appl.* 285: 107-127.

Hilger, S. 1997. Differential and difference calculus unified. *Nonlinear Analysis, Theory, Methods and Applications* 30: 2683-2694.

Royden, H.L. 1988. *Real Analysis* Macmillan: Newyork.

Rzezuchousky 2005. A note on measures on time scales *Demonstratio Mathematica,* 38(1): 79-84.

Ufuktepe, Ü. and A. Yantır 2005. Measure on time scales with Mathematica. *Lecture notes in computer science*. 3482: 529-537.