

Theory of Chances: Novel Stochastic Chains: Linear Algebraic Approach

Garimella Rama Murthy*

International Institute of Information Technology, Gachibowli, Hyderabad-500032, Telangana, India

Abstract

In this research paper, the author proposes the idea of allowing the chance (like probability) to assume negative and complex values. Novel stochastic chains based on real/complex valued L^{2p} chances are formally discussed. Also, the theory of L^1 chances is discussed. Chance number theory is briefly discussed.

Keywords: Mathematics; Bayes; Length; Probability Theory; Chance Concept

1. Introduction

Mathematics as a branch of human endeavour had early beginnings with geometry. The efforts of mathematicians led to the fields such as trigonometry, algebra etc. Many branches of mathematics had their beginnings in attempting to solve problems associated with natural and/or artificial phenomena. One such branch was classical probability theory. Efforts of mathematicians such as Pascal, Bayes, Bernoulli etc. led to various basic concepts of classical probability theory.

In trying to abstract the notion of length of a subset of real line, Lebesgue formalized the “theory of measure” [20]. Probability theory, was considered as the theory associated with measure space of finite (one) measure. Thus, probability theory was provided a rigorous axiomatic footing. Researchers realized that some uncertainties associated with human reasoning cannot be modeled using probability theory. Thus, fuzzy logic was proposed as an alternative to model certain types of uncertainty.

After understanding the concept of fuzzy set, the author reasoned the need for negative membership function in [19]. This innovative idea motivated the author to see if a consistent theory can be developed in which the probabilities (called “chances”) are allowed to assume negative values. More generally, the author proposed to allow “chances” to assume even complex values (complex numbers). Also, the author proposed to formalize a consistent theory of stochastic chains based on real, complex valued chances. The associated effort resulted in this research paper.

• Motivation for the theory of chances

In traditional as well as modern probability theory, on spaces with finite, countable number of outcomes it is assumed that the probability of every outcome is a nonnegative number lying in the interval $[0,1]$. After careful contemplation, the author reasoned that there are many experiments in which some of the outcomes are “favourable” to the events (defined on the space of outcomes), while the other outcomes are totally “un-favourable” (to the events considered). More explicitly, in many real world situations, outcomes have diametrically opposite possibility of occurring in some events.

Specifically, let us consider an experiment in which we observe temperatures in a certain region of our planet. Let the temperatures (outcomes in the experiment) assume positive as well as negative values. We consider the events: “HOT” and “COLD” day. A temperature value of - 25 degrees Centigrade has “negative” CHANCE in the event “HOT” days (and not “zero” or small positive value in the “HOT” Day event). Similarly, a temperature of +50 degrees Centigrade has “NEGATIVE

CHANCE” in the event COLD day.

It is possible to give several real world experiments in which “chances” of outcomes necessarily assume negative values (For instance “summer season” related events SUNNY, CLOUDY, RAINY and the associated temperature values). Thus the author was naturally motivated to develop a detailed “THEORY OF CHANCES” on spaces with “uncountable” as well as “countable” outcomes.

• Early attempts

In theoretical physics literature, physicists were naturally led to negative probabilities in modeling problems. For some reason, such models were abandoned. Also probabilists encountered characteristic functions, whose inverse Laplace transforms are not probability densities. For a variety of reasons, such characteristic functions were also not utilized in stochastic modeling.

This research paper is organized as follows. In section 2, we discuss the theory of L^{2p} real valued chances and the associated stochastic chains. In section 3, we allow the chances to assume complex numbered values. The associated complex chance based stochastic chains are discussed in this section. In Section 4, probabilistic/chance number theory is discussed. In Section 5, theory of L^1 chances is discussed and also the associated stochastic chains are briefly discussed. In Section 6, theory of random sets and graded sets is briefly discussed. The research paper concludes in Section 7.

2. Theory of L^{2p} Real Valued Chances: Novel Stochastic Chains

Probability theory was successful in modeling a certain type of uncertainty. Researchers proposed the need for other types of uncertainty such as fuzzy uncertainty. After a careful understanding of various types of uncertainty, the author proposes a novel theory of uncertainty, called the theory of chances. This theory constitutes a variation of probability theory [9].

*Corresponding author: Garimella Rama Murthy, Professor, International Institute of Information Technology, Gachibowli, Hyderabad-500032, Telangan, India, Tel: 040 6653 1000; E-mail: rammurthy@iiit.ac.in

Received August 11, 2014; Accepted October 28, 2014; Published November 08, 2014

Citation: Murthy GR (2014) Theory of Chances: Novel Stochastic Chains: Linear Algebraic Approach. J Appl Computat Math 3: 191. doi:10.4172/2168-9679.1000191

Copyright: © 2014 Murthy GR. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

For the sake of simplicity and concreteness, we first consider space of chances with finite/countable number of outcomes.

Two outcome case

• Let us first consider a “chance space” with two outcomes i.e. $\{w_1, w_2\}$ we have

$$Q(w_1) \geq 0 \text{ or } Q(w_2) < 0$$

$$Q(w_1) \geq 0 \text{ or } Q(w_1) < 0$$

$$Q(w_1) \geq 0 \text{ or } Q(w_2) \geq 0$$

$$Q(w_1) \leq 0 \text{ or } Q(w_2) \leq 0$$

Also, we must necessarily have that

$$[Q(w_1)]^2 + [Q(w_2)]^2 \leq 1$$

Thus, in this simple case, the chances of outcomes lie on a circle of unit radius centered at the origin. More generally, we consider a “Chance space of p^{th} kind” in which case

$$[Q(w_1)]^{2p} + [Q(w_2)]^{2p} = 1$$

“p” an integer. When $p=1$, we have a basic chance space. Also, generalized chances (parametrized by “p”) are constrained to lie on an arbitrary hypersurface or manifold. There are uncountably many such chance spaces.

Remark

Even in the case of a probability space, with two outcomes, the probabilities of outcomes lie on a straight line. Thus, probabilities, like chances are also constrained. More generally, we can consider general constraint sets on which the “generalized chances” lie.

It should be noted that we are interested in the “normalized chance” of events. Thus, the “normalized chance” of an event is the sum of squares of chances of constituent outcomes.

Finite outcome case

Now, let us consider a “chance space” (unlike probability space) with finitely many outcomes (say N outcomes i.e. w_1, w_2, \dots, w_N). In such a space, chances of some/all outcomes assume negative values satisfying the condition

$$[Q(w_1)]^2 + [Q(w_2)]^2 + \dots + [Q(w_N)]^2 = 1$$

i.e. the set of all possible chances (of outcomes) constitute the vectors on the unit hypersphere in N dimensional Euclidean space. More generally, we can consider a “chance space of p^{th} kind” in which,

$$[Q(w_1)]^{2p} + [Q(w_2)]^{2p} + \dots + [Q(w_N)]^{2p} = 1$$

When p equals one, we have a “basic chance space”. Thus, there are uncountably many chance spaces (corresponding to vectors on the unit hyper sphere).

Definition

It should be kept in mind that the “normalized chance” of an event is the sum of squares of chances of the constituent outcomes.

Remark

One of the main reasons for “normalization” is the following: Suppose, we consider a chance space with countably many outcomes.

Let $Q(\bar{w}_1) = d$ and $Q(\bar{w}_2) = -d$

Consider the event $C = \{w_1, w_2\}$ then

$$Q(c) = 0$$

Which is sometimes not reasonable while defining the chance of the event C.? But the normalized chance of event C is $\bar{Q}(c) = 2d^2$

Remark

Since each chance space corresponds to a point on the unit hypersphere, one can introduce the notions such as “linearly independent chance spaces”, “orthogonal chance spaces” etc. Detailed results are derived using the well-known results of linear algebra.

Countable outcome case

We now consider the “chance space” with countable number of outcomes. The chances of outcomes assume positive or negative values such that

$$\sum_{i=1}^{\infty} [Q(w_i)]^2 = 1$$

On treating the set of chance values as vectors, the set of all possible chance vectors (of outcomes) constitute the points on the infinite dimensional unit hyper sphere (in the infinite dimensional Euclidean space). Thus, there are uncountable many possible chance spaces. But it is well known that the set of “square summable” sequences constitute the “discrete Hilbert space”. Thus the set of all possible chance vectors in the infinite dimensions is a subset of the discrete Hilbert space. More generally, we have

$$\sum_{i=1}^{\infty} [Q(w_i)]^{2p} = 1$$

Let us return to the basic chance space, when the number of outcomes is countable. Treating the chance vectors (in the countable outcome case) as infinite dimensional vectors, results of infinite-dimensional linear algebra can be utilized to relate the chance spaces.

Remark

We have another interesting case in which the chances of outcomes can be negative and we have the normalization condition that

$$\sum_{i=1}^N |Q(\bar{w}_i)| = 1 \text{ or } \sum_{i=1}^{\infty} |Q(\bar{w}_i)| = 1$$

Remark

In the finite/countable chance spaces (previously considered), concepts like “conditional chance” (like conditional probability), “independent events” etc are easily defined and utilized. Also, using the concept of “normalized chance”, “modified distribution function” is defined and studied (in the case of finite/countable chance spaces). In summary, the concepts/Lemmas/Theorems of probability theory have their parallels in the theory of L^{2p} chances.

Novel Stochastic Chains: Real Valued Chance Case:

In the following, we consider an interesting class of chance/random processes defined on chance spaces [Rama3].

- Consider a sequence of random variables $\{X_i\}_{i=0}^{\infty}$ defined on a chance space (Ω, β, Q) Let the sequence satisfy the Markovian property i.e. Let $m < n$. We have

$$Q\{X_n = j | X_0, X_1, \dots, X_m\} = Q\{X_n = j | X_m\}$$

Thus, let the one-step state transition matrix be denoted by R, where

$$R_{ij} = Q\{X_{n+1} = j | X_n = i\}$$

Also, let the sequence of chance vectors (be row vectors) be denoted by $\{v_n, n \geq 0\}$ where jth component of v_n be $Q\{X_n = j\}$. it is easy to see that given v_0 is a chance vector, the sequence of vectors $\{v_n, n \geq 1\}$ constitute chance vectors if R is an orthogonal matrix. Thus, as in the case of Markov chains, we have the following recurrence equation involving chance vectors and the state transition matrix.

$$v_{n+1} = v_n R \text{ for } n \geq 0.$$

Remark

This investigation is motivated by the Hopfield associative memory [Hop]. The main observation is that the state of Hopfield network (at all-time instants) lies on the hypercube [2]. This observation motivated the author to conceive of dynamical systems (linear/non-linear) whose state lies on a constraint set such as unit hypersphere, bounded lattice, complex hypercube etc [3,8]. The author expects that such dynamical systems naturally arise in modeling of neural and other types of systems.

Now we investigate the convergence behavior of the above linear dynamical system.

- Linear System with State Dynamics on the Unit Hypersphere:

Consider a discrete time dynamical system with the following state dynamics $v(n+1) = v(n)R$ for $n \geq 0$. where

R is an “orthogonal” matrix; V (n) is a row vector for all n. It is easy to see that, if V (0) is constrained to lie on the unit hypersphere, then V (n) lies on unit hypersphere for all “n”. The following lemma on the spectrum of R is well known

Lemma 2.1: All the eigenvalues of an orthogonal matrix lie on the unit circle.

Using the above lemma, we make the following conclusions on the stability of such linear dynamical system.

- Suppose V (0) is the left eigenvector corresponding to any one of the eigenvalues on the unit circle. Let the eigenvalue be

λ_{\max} Then, we have

$$v(n) = \lambda_{\max}^n v(0)$$

Thus, we necessarily have that

$$\lim_{n \rightarrow \infty} |v(n)| = |v(0)|$$

Hence, in this case, we have convergence to an equilibrium chance vector (which happens to be the initial vector).

- Suppose V(0) is an arbitrary row vector that lies on the unit hyper sphere i.e.

Since R is an orthogonal matrix, we have that $V^T(0) = RD$ for some column vector D.

Thus, we have

$$R D D^T R^T = 1$$

Hence, we have the following equation

$$V(n+1) = D^T R^n$$

Suppose the orthogonal matrix R is diagonalizable i.e.

$$R = \sum_{i=1}^M \lambda_i F_i$$

Then, we necessarily have that

$$R^n = \sum_{i=1}^M \lambda_i^n F_i$$

Hence, we necessarily have that

$$V(n+1) = D^T R^n = \sum_{i=1}^M \lambda_i^n D^T F_i$$

Using Cauchy-Schwarz inequality and the fact that all the eigenvalues lie on the unit circle, we have

$$|V(n+1)| \leq \sum_{i=1}^M |D^T F_i| \text{ For all } n.$$

From the above equation, the limiting behavior can be upper bounded.

- Now, we consider the case where “R” is “orthogonal” as well as “symmetric” i.e.

$$R^T = R \text{ and } R^T = R^{-1}$$

Imply that $R^2=I$ The following claim follows easily and the proof is avoided for brevity

- Claim: The eigenvalues of a symmetric, orthogonal matrix are either +1 or -1.

Thus, we necessarily have

$$R = F_1 - F_2$$

Where F_1 is the sum of residue matrices corresponding to the eigenvalue “one” and F_2 is the sum of residue matrices corresponding to the eigenvalue -1 (minus one). With the above Lemma, we are able to prove the following Convergence Theorem regarding the stability of linear dynamical system.

$$V(n+1) = V(n)R \text{ For } n \geq 0$$

Since $R^2=I$ we immediately conclude that the above linear dynamical system is periodic.

Theorem 2.1: Consider the linear dynamical system

$$V(n+1) = V(n)R \text{ For } n \geq 0$$

With the row vector V (0) lying on the unit hyper sphere. The limiting behavior of the above dynamical system is characterized by the following two cases:

- cycle of length 2 ($V(0), V(0)R, V(0)V(0)R, \dots$)
- convergence to a single vector (when $V(0)$ or $V(0)F_2 = 0$)

The proof follows and is avoided for brevity.

Remark

In view of the above Theorem, novel stochastic chains characterized by symmetric orthogonal matrix converge to an equilibrium chance vector when $V(0)$ or $V(0)F_2 = 0$ in many practically interesting cases; such a choice of initial chance vector can always be made. The practical utility of such chains need to be addressed.

- Suppose ‘V (0)’ is infinite dimensional and $V(0) V^T(0) = 1$. Then,

by definition $V(O)$ belongs to the “discrete Hilbert space”. Also, all the vectors $V(n)$ for all n lie on the infinite dimensional unit hyper sphere. The limiting behavior in the infinite dimensional case is similar to that in finite dimensional case.

Remark

It should be noted that Markov chains exist in discrete time as well as continuous time. But it is possible to reason that the continuous time counterparts of the novel stochastic chains (discussed above) do not exist. Specifically, the sojourn time (in each state) being Poisson and the underlying discrete time stochastic chain being “ L^2 -norm based chance process” is simultaneously impossible.

Uncountable outcome case

We now consider the case where the set of outcomes is uncountable. For the sake of simplicity, we consider the real valued chance case.

- Let $f(x)$ be the “chance density function” which can assume positive as well as negative values. Let B be an event defined on the chance space

$$\int_B f(x)dx = Q(B)$$

- As in the countable outcome case, the “normalized chance density” is defined as

$$g(x) = f^2(x) \text{ with } \int g(x) dx = 1$$

Thus, the chance density function $f(x)$ belongs to the Hilbert space i.e. L^2 space.

Remark

As in the countable outcome case, concepts in traditional probability theory (in the case where probability space has uncountable many outcomes) have parallel counterparts in the case where the chance space has uncountable many outcomes. Detailed explanation can be found in [8].

Note

In the following section, we consider a generalization of the case considered here. For the purposes of clarity and completeness, parallels of the concepts considered in this section are explicitly included.

3. Theory of L^{2p} Complex Valued Chances: Novel Stochastic Chains

In the previous section, we allowed chances (unlike probabilities) to assume negative values. A natural question arises as to whether the chances can be allowed to assume “complex (numbered) values”. More interestingly, we propose and formalize a consistent “theory of complex valued chances”. As in the previous section, based on complex valued chances, we formalize the “theory of complex novel stochastic chains”.

Consider a set Ω with finitely many outcomes $\{w_1, w_2, \dots, w_N\}$. Let the chance $c(\bar{w}_i)$ be a complex number for $i=1,2,\dots, N$. The real and imaginary parts of $c(\bar{w}_i)$ can assume positive or negative values.

Furthermore, let $0 \leq |c(\bar{w}_i)| \leq 1$ with the condition $\sum_{i=1}^N |c(\bar{w}_i)|^2 = 1$

More generally, we can consider a “complex chance space of p th kind” in which we have $\sum_{i=1}^N |c(\bar{w}_i)|^{2p} = 1$

When $p=1$, we have a “Basic complex chance space” (as in the real

valued case). For the purposes of brevity, in the following discussion, we only consider the basic chance space. Thus, in this case, the chance vector (with the components being the chance values of the outcomes) lies on the “complex N dimensional hyper sphere”.

Definition

The ‘normalized chance’ of an event is equal to the sum of squares of magnitudes of constituent outcomes (implicitly, the normalized chance of an outcome is equal to the square of magnitude of the complex valued chance).

Note

In this theory of complex valued chances, we have concepts that are parallel of those in the traditional probability theory. We utilize normalized chances to formalize the definitions. For the purposes of illustration, some definitions are provided below.

- **Chance Mass Function:** It is a function which assigns complex valued (with positive as well as negative real parts) chances to each outcome in the chance space (The normalization condition is satisfied by the complex valued chances).

The normalized chance mass function assigns to each outcome, the squared magnitude of the associated complex valued chance.

- **Random Variable:** It is a function from the space of outcomes (in the chance space) to the real/complex number field.

- **Cumulative Chance Distribution Function:** It is a function based on the events associated with the random variable. Given a random variable X , the cumulative chance distribution $F_X(x)$ is given by the following expression:

$$F_X(x) = \bar{c}\{w : X(w) \leq x\}$$

Where $\bar{c}\{.\}$ is the normalized chance of the event. Thus, in terms of chances, we have the following expression for $F_X(x)$

$$F_X(x) = \sum_{\{w_i : X(w_i) \leq x\}} |c(w_i)|^2$$

Where $|\cdot|$ denotes the absolute value of the complex number.

- **Random/Stochastic Process:** Chance based random process is a collection (countable or uncountable) of random variables defined on the chance space (Ω, β, c) . Study of such random processes proceeds along the same lines as stochastic processes defined on the traditional probability space.

Complex valued chance based novel stochastic chains

Now, we consider a countable state space random process (defined on a chance space) $\{Y_i\}_{i=1}^\infty$ which satisfies Markovian-Type property i.e. letting $m < n$, we have $c\{Y_n = j | Y_0, Y_1, Y_2 \dots Y_m\} = c\{Y_n = j | Y_m\}$

Thus, such a stochastic process (complex valued chance based) is completely characterized by the one step state transition matrix denoted by S , where

$$S_{ij} = c\{Y_{n+1} = j | Y_n = i\}$$

Now, let

$$c\{Y_n = i\} = \mu_i(n)$$

For the sake of simplicity, let us first consider the case where the state space associated with all the random variables is a fixed finite number. Thus, at each time instant, we have the chance vector (with

complex valued components) $\bar{\mu}(n)$ Specifically, let the chance vector at time “zero” ($n = 0$) lie on the complex unit hyper sphere i.e.

$\bar{\mu}(0)\bar{\mu}^*(0) = 1$ Where $\bar{\mu}^*(.)$ denotes the complex conjugate transpose of vector (matrix). Thus, by the Markovian type property, we have

$$\bar{\mu}(n+1) = \bar{\mu}(n)S \quad (**)$$

Consider the case where “S” is a unitary matrix. In such a case, it is easy to verify the following claim.

Claim: Given that $\bar{\mu}(0)$ is a chance vector (with complex valued components) lying on the complex unit hyper sphere, the sequence of vectors $\{\bar{\mu}_n; n \geq 1\}$ also constitute chance vectors lying on the complex unit hyper sphere.

Now, we would like to investigate the equilibrium behavior of the linear dynamical system specified by (**). The following Lemma from linear algebra is very helpful.

Lemma: All the eigenvalues of a unitary matrix lie on the unit circle.

Note: As in the case of real valued chance based random process, results related to equilibrium behavior are easily derived. Details are avoided for brevity.

We now consider the case where S is a “unitary” as well as “Hermitian” matrix. i.e.

$$S^* = S \text{ and } S = S^{-1}$$

Thus, using basic results in linear algebra, we necessarily have that “S” is diagonalizable and

$$S^2 = I$$

Consequently, we have the following Lemma.

Lemma: The eigenvalues of a Hermitian, Unitary matrix are either +1 or -1. Hence, we have that

$$S = G_1 - G_2$$

Where G_1 is the sum of residue matrices corresponding to the eigenvalue “one” and G_2 is the sum of residue matrices corresponding to the eigenvalue -1 (minus one).

The above inferences naturally lead to the following Theorem (in the same spirit of Theorem (2.1) in Section 2).

Theorem 3.1: Consider the linear dynamical system

$$\bar{\mu}(n+1) = \bar{\mu}(n)S \quad \text{for } n \geq 0$$

The row vector $\bar{\mu}(0)$ lying on the complex unit hyper sphere. The steady state (equilibrium) behavior of the above dynamical system is characterized by the following two cases:

- cycle of length 2 ($\bar{\mu}(0), \bar{\mu}(0)S, \bar{\mu}(0), \bar{\mu}(0)S, \dots$)
- convergence to a single vector when $\bar{\mu}(0)G_1 = 0$ or $\bar{\mu}(0)G_2 = 0$

The proof follows based on the above facts and is avoided for brevity.

- Summary of the Results:

In the same spirit of Markov chains, we have a sequence of random variables $\{Y_i\}_{i=1}^{\infty}$ associated with the chance vectors (with complex valued components) $\{\bar{\mu}_n\}_{n=0}^{\infty}$ we have the following linear dynamical system.

$$\bar{\mu}(n+1) = \bar{\mu}(n)S$$

Where S is a unitary and Hermitian matrix. The novel stochastic chain converges to an equilibrium chance vector when

$$\bar{\mu}(0)G_1 = 0 \text{ or } \bar{\mu}(0)G_2 = 0$$

It is expected that in many practical interesting cases, such a choice of initial chance vector can always be made. We expect that such novel stochastic chains (with associated complex valued chance vectors) are of utility in science, engineering, economics and other fields.

Remark

The generalization to countable state space stochastic chains follows in a straightforward manner

Remark: We also propose another definition of “normalized chance”. The normalized chance of an outcome (in finite or countable chance spaces) is the magnitude of the chance (which can be complex valued). Thus, the normalized chance of an event is the sum of normalized chances of the constituent outcomes. We consider the case where the sum of magnitudes of chances is one.

Generalizations

- It is well known that discrete time Markov chains have multidimensional extensions, called Markov Random Fields. So, along these lines, we are currently investigating multidimensional extensions of “chance based stochastic chains”.

- It is also known that discrete time Markov chains have “k-dependent (number of past steps) Markov chains”. So a natural extension is to study “k-dependent chance stochastic chains”.

- Furthermore, it is also known that “quaternions”, “Clifford/Grossmann algebra” are considered as generalizations of complex numbers. Thus chances are allowed to assume such generalized structures.

4. Waring Problem: Probabilistic/Chance Number Theory

In view of the waring problem the following sums are naturally considered. In elementary mathematics, these sums are naturally studied.

$$1 + 2 + 3 + \dots + m = \sum_{i=1}^m i = \frac{m(m+1)}{2} = \frac{1}{2}m^2 + \frac{1}{2}m$$

$$1^2 + 2^2 + \dots + m^2 = \sum_{i=1}^m i^2 = \frac{m(m+1)(2m+1)}{6} = \frac{1}{3}m^3 + \frac{1}{2}m^2 + \frac{1}{6}m$$

$$1^3 + 2^3 + \dots + m^3 = \sum_{i=1}^m i^3 = \frac{m^2(m+1)^2}{4} = \frac{1}{4}m^4 + \frac{1}{2}m^3 + \frac{1}{4}m^2$$

In general, in the spirit of above sums, let us define the following most general sum

$$h^{(s)}(m) = 1^s + 2^s + \dots + m^s$$

For general ‘s’, Faulhaber’s formula provides the expression for the associated polynomial. It is well known that these polynomials can be expressed in terms of Bernoulli polynomials.

On observing the coefficients of polynomials $h^{(1)}(.), h^{(2)}(.), h^{(3)}(.)$ it is clear that they sum to one. From the Faulhaber polynomials, we realize that not all the coefficients are non-negative. Thus, the following definitions are motivated.

Definition: A probability polynomial is a polynomial all of whose

coefficients are non-negative and sum up to one.

Definition: A probability-type polynomial is a polynomial the sum of whose coefficients is one. The coefficients need not all be positive numbers.

We have the following Theorem on the above sums.

Theorem 4.1: Let

$$h^{(s)}(m) = 1^s + 2^s + \dots + m^s = a_{s+1}m^{s+1} + a_s m^s + \dots + a_1 s$$

We claim that $h^{(s)}(\cdot)$ is a “probability” polynomial or “probability-type” polynomial. We can also call ‘probability type’ polynomials as “chance” polynomials [9].

Proof: Follows from the argument in Wikipedia article **Q.E.D.**

Remark

The implications of the above Theorem are fully being investigated in [8]. It should be noted that with Faulhaber formula stated as an identity, the above result can be obtained by substituting n equal to 1.

Also, Faulhaber realized that the ‘sum’, $h^{(s)}(m)$ for odd values of ‘ s ’ can be expressed as a polynomial in ‘ a ’, where

$$h^{(2p+1)}(m) = \sum_{k=1}^m k^{(2p+1)} = c_1 a^2 + c_2 a^3 + \dots + c_p a^{p+1} \text{ and } a = \frac{m(m+1)}{2}$$

Lemma 4.1: In the above equation/expression, $\sum_{j=1}^p c_j = 1$ i.e. the above polynomial is a “probability-type” polynomial or chance polynomial as a function of ‘ a ’.

Proof: Since $a = \frac{m(m+1)}{2}$ and $h^{(2p+1)}(m) = c_1 a^2 + c_2 a^3 + \dots + c_p a^{p+1}$

it is clear that

$$h^{(2p+1)}(1) = c_1 + c_2 + \dots + c_p$$

Thus, by the above Lemma,

$$c_1 + c_2 + \dots + c_p = 1 \quad \text{Q.E.D.}$$

Remark: We realize that for various values of “ s ”, the coefficients of polynomials $\{h^{(s)}(\cdot)\}$ constitute “interesting” “chance” mass functions corresponding to discrete chance/random variables (assuming integer values) [9]. We expect these “chance” mass functions to naturally arise in applications such as physics, chemistry etc (in the spirit of Fermi-Dirac statistics, Bose-Einstein statistics etc.)

For instance, we can compute the ‘entropy’ of the above ‘interesting’ probability/chance mass functions when the coefficients of Faulhaber polynomials are all non-negative.

Consider the value of “ s ” to be 1, 2, 3. It is easy to see that the entropy values are

$$H(1)(X_1) = \frac{1}{2} \log_2 2 + \frac{1}{2} \log_2 2 = 1 \text{bit}$$

$$H(2)(X_2) = \frac{1}{3} \log_2 3 + \frac{1}{2} \log_2 2 + \frac{1}{6} \log_2 6 = \frac{4}{6} + \frac{1}{2} \log_2 3$$

$$H(3)(3) = \frac{1}{4} \log_2 4 + \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 = 2 \text{bits.}$$

Remark

It should be noted that the Shannon entropy cannot be defined in the case where some of the chances assume negative values. But other entropy definitions such as Tsallis entropy could be defined. To be more exact Tsallis proposed an entropy measure which in the case of a

discrete random variable is given by

$$s_q(p) = \frac{1}{q-1} (1 - \sum_x (p(x))^q).$$

Thus, it is clear that when $q=2$, Tsallis entropy could be defined for chance mass functions of a discrete chance variable. More generally for all even integer values of ‘ q ’, Tsallis entropy of chance mass functions can be defined. It will be interesting to see if the Tsallis entropy of chance mass functions associated with the Faulhaber polynomials $\{h^{(s)}(\cdot) \text{ for } s \geq 4\}$ is strictly increasing (and thus can be interpreted in the spirit of second law of Thermodynamics).

In view of the fact that probability/chance polynomials are associated with Shannon/Tsallis entropy (based on the coefficients), we are naturally led to the following discussion.

- Every polynomial with non-negative coefficients, can be divided by the sum of coefficients (non-zero sum) to arrive at a probability polynomial (without affecting the zeroes of the polynomial)

- Every polynomial with arbitrary (negative as well as positive) coefficients, can be divided by the sum of coefficients (assumed to be non-zero) to arrive at a chance/probability polynomial (without affecting the zeroes of the polynomial).

- Consider a finite set of non-negative elements i.e. $A = \{x_1, x_2, \dots, x_N\}$ Divide each of the elements by the sum of all elements in A . i.e. define

$$p_i = \frac{x_i}{\sum_{j=1}^N x_j} \text{ for } 1 \leq i \leq N.$$

It is evident that p_i 's are probabilities Thus Shannon entropy of set A can be defined as

$$H(A) = -\sum_{i=1}^N p_i \log_2 p_i$$

Such an entropy measure quantifies, how different are the elements of the finite set from the constant value (i.e. different from a uniform probability mass function).

Remark: The generalization of above definition for a countable set with non-negative elements can easily be made when the sum of all elements converges to a non-negative value

- If the elements of set A are allowed to assume negative values, then the normalized absolute value of the elements i.e.

$$p_i = \frac{|x_i|}{\sum_{j=1}^N |x_j|} \text{ for } 1 \leq i \leq N$$

can be utilized to compute the Shannon entropy (of the associated probability mass function). We call such an entropy as. L^1 -Shannon entropy

- More generally, L^p - Shannon entropy of a finite set (with possibly negative elements) is based on the following normalization:

$$p_i = \frac{(|x_i|)^p}{\sum_{j=1}^N (|x_j|)^p} \text{ for } 1 \leq i \leq N$$

The above approach can be generalized to countable sets in a straightforward manner.

Remark: Other types of entropy such as Tsallis/Renyii entropy of finite/countable sets are defined based on suitable normalization.

Definition: Every polynomial whose coefficients are all positive and are all equal to the same constant can be normalized to arrive at a “maximum Shannon entropy” polynomial.

• For instance, every cyclotomic polynomial constitute maximum entropy polynomial.

• Lemma 4.2: Faulhaber polynomials $\{h^{(s)}(n)\}$ are probability polynomials only for $s = 1, 2, 3$ and they are all probability-type polynomials for $s \geq 4$. Thus, there are infinitely many Faulhaber polynomials that are chance polynomials (probability-type polynomials).

Proof: It is well known that the following expression for Faulhaber polynomials holds true.

$$h^{(p)}(n) \sum_{k=1}^n k^p = \sum_{j=0}^p \binom{p}{j} \frac{B_{p-j}}{j+1} n^{j+1}$$

Where B_j 's are Bernoulli numbers.

For $p=1, 2, 3$, we have the following expressions:

$$h^{(1)}(n) = \frac{n(n+1)}{2} = \frac{1}{2}n^2 + \frac{1}{2}n$$

$$h^{(2)}(n) = \frac{n(n+1)(2n+1)}{6} = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

$$h^{(3)}(n) = \frac{n^2(n+1)}{4} = \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}n^2$$

Thus, the above three polynomials are pure “probability” polynomials. For $p=4$. We have the following expression.

$$h^{(3)}(n) = \sum_{j=0}^4 \binom{4}{j} \frac{B_{4-j}}{j+1} n^{j+1} = B_4 n + 2B_3 n^2 + 2B_2 n^3 + B_1 n^4 + \frac{B_0}{5} n^5$$

Substituting for the Bernoulli numbers, we have the following

$$h^{(4)}(n) = -\frac{1}{30}n + \frac{1}{3}n^3 + \frac{1}{2}n^4 + \frac{1}{5}n^5$$

Thus, since the Bernoulli number B_4 is a negative real number, $h^{(4)}(n)$ is a probability type (chance) polynomial and not a probability polynomial. Now let us consider $p \geq 5$

Since, $\binom{p}{j} \frac{1}{j+1}$ is a positive real number for all integer values of $\{p, j\}$ and Bernoulli numbers $\{B_4, B_8 \dots\}$ etc are all negative real numbers, for all $p \geq 4$ the Faulhaber polynomials $\{h^{(p)}(n)\}$ constitute probability type (chance) polynomials (and not probability polynomials).

• Thus, there are infinitely many Faulhaber polynomials that are chance polynomials **Q.E.D.**

Note

It can be easily shown that there are infinitely (countable) many Bernoulli numbers that are negative real numbers. Also, there are countably many positive Bernoulli numbers. Thus, positive and negative valued Bernoulli numbers can be put in one-to-one correspondence with the set of all natural numbers (positive and negative valued).

• From the Formula for Faulhaber polynomials, the following facts can readily be derived

• For all values of ‘p’, coefficient of $n^{p+1} = \binom{p}{p} \frac{B_0}{p+1} = \frac{1}{p+1}$

• For all values of ‘p’, coefficient of $n^p = \binom{p}{p-1} \frac{B_1}{p} = \frac{1}{2}$

• For all values of ‘p’, coefficient of $n = B_p$. Thus, if ‘p’ is odd, the coefficient of ‘n’ is zero. More generally,

$$\binom{p}{j} \frac{B_{p-j}}{j+1} n^{j+1} = 0 \text{ if and only if } B_{p-j} = 0$$

Also,

$$\binom{p}{j} \frac{B_{p-j}}{j+1} < 0 \text{ if and only if } B_{p-j} < 0$$

• For all values of ‘p’, coefficient of $n^2 = \frac{p}{2} B_{p-1}$. Thus, if ‘p’ is an even number greater than or equal to 4, then coefficient of n^2 is zero.

• From the Faulhaber formula, we also have that

$$\sum_{j=0}^p (-1)^j \binom{p+1}{j} B_j = (p+1) \text{ with } B_1 = -\frac{1}{2}$$

• For all integer values of ‘p’, we have that

$$\sum_{q=0}^p \binom{2p+2}{2q} (2 - 2^{2q}) B_{2q} [9^{p+1-q} - 1] = 2^{2p+2} (2p+2)$$

• **Lemma 4.3:** Let the Bernoulli polynomial for integer ‘p’ be denoted by $B_p(x)$ and let b_p denote the Bernoulli number. The following identity is satisfied for all values of ‘p’.

$$B_{p+1}(2) = (p+1) + b_{p+1}$$

Also, for even values of ‘p’ we have that

$$B_{p+1}(2) = (p+1)$$

Proof:

The following expression relating Faulhaber polynomials and Bernoulli polynomials is well known:

$$h^{(p)}(x) = \sum_{k=1}^p k^p = \frac{B_{p+1}(x+1) - b_{p+1}}{p+1}$$

Rearranging the equation, we have that

$$B_{p+1}(x+1) = (p+1) h^{(p)}(x) + b_{p+1}$$

From Theorem 4.1, we note that Faulhaber polynomial is either a probability polynomial or probability-type polynomial i.e. $h^{(p)}(1) = 1$ hence, we readily have that for all values of ‘p’

$$B_{p+1}(2) = (p+1) + b_{p+1}$$

Also, it is well known that the all odd Bernoulli numbers are zero. Hence, we readily have that for even values of ‘p’,

$$B_{p+1}(2) = (p+1)$$

• Lemma 4.4: The product of any two chance/probability polynomials corresponding to the above sums ($\{h^{(i)}(m)\}, \{h^{(j)}(n)\}$) is another chance/probability polynomial [Rama3].

Proof: Let $k(m, n) = (\{h^{(i)}(m)\} \{h^{(j)}(n)\})$

It is well known that coefficients of the resulting polynomial (i.e. the product polynomial) are the convolution of the coefficients corresponding to the two chances/probability polynomials

$$k(1,1) = h^{(1)}(1)h^{(1)}(1) = (1)(1) = 1$$

Also the same result holds true if $m=n$.

Corollary: If a probability/chance polynomial (in unknown 'n') can be expressed as the product of two polynomials and one of them is a probability/chance polynomial, then the other is also a probability/chance polynomial.

Remark: The above Lemma shows that the product of finitely many chance/probability polynomials is another chance/probability polynomial. Thus this class of polynomials is closed under multiplication.

Remark: In view of Lemma 4.4, we realize that higher powers of first three Faulhaber polynomials i.e. $\{h^{(j)}(m): 1 \leq j \leq 3\}$ are also probability polynomials. Specifically, we first consider higher powers of first Faulhaber polynomial i.e. $h^{(1)}(m) = \frac{m(m+1)}{2}$. The following result readily follows:

Lemma 4.5: The "k"th power of $h^{(1)}(m)$ (for $k \geq 2$) constitutes a probability polynomial, whose coefficients constitute a Binomial probability mass function i.e. binomial $(k, \frac{1}{2})$. Thus, it also follows that "k"th power of $h^{(3)}(m)$ (for $k \geq 2$) constitutes a probability polynomial whose coefficients constitute Binomial $(2k, \frac{1}{2})$.

Proof:

$$[h^{(1)}(m)]^k = \left[\frac{m(m+1)}{2}\right]^k = \frac{1}{2^k} \sum_{j=0}^k \binom{k}{j} m^{k+j} = \sum_{j=0}^k \binom{k}{j} \left(\frac{1}{2}\right)^j \left(\frac{1}{2^{k-j}}\right) m^{k+j}$$

Thus $[h^{(1)}(m)]^k$ ('k' fold product of sum of first 'm' integers) is a probability polynomial, whose coefficients constitute a Binomial probability mass function i.e. Binomial $(k, \frac{1}{2})$.

Since $h^{(3)}(m) = [h^{(1)}(m)]^2$ it follows that "k"th power of $h^{(3)}(m)$ constitutes a probability polynomial whose coefficients constitute Binomial $(2k, \frac{1}{2})$.

In view of the above lemma, it will be interesting to interpret the higher powers of probability polynomials $h^{(2)}(m)$. Also it is interesting to understand whether the Shannon entropy of such probability polynomials is increasing/decreasing with "k".

- It is easy to see that

$$[h^{(r)}(m)][h^{(s)}(m)] = \left[\sum_{j=1}^m j^r\right] \left[\sum_{k=1}^m k^s\right] = \sum_{j=1}^m \sum_{k=1}^m j^r k^s$$

The above expression is reminiscent of similar expression in Eisenstein series and its generalization by Ramanujan. Specifically, Eisenstein series are special cases of Ramanujan's more general function [1,6].

$$\phi_{r,s}(q) = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} k^r n^s q^{kn}$$

where 'r' and 's' are non-negative integers. Trivially $\phi_{r,s}(q) = \phi_{r,s}(q)$. We are naturally led to the study of finite version of the above function.

$$\phi_{r,s}(q) = \sum_{k=1}^M \sum_{n=1}^M k^r n^s q^{kn} = \phi(r,s,q,M)$$

Suppose $\{r,s\}$ are considered to be constants. Then $\phi(r,s,q,M)$ a bi-variate polynomial (function) in $\{q,M\}$. It is easy to see that

$\phi(r,s,q,M)|_{q=1} = \sum_{k=1}^M \sum_{n=1}^M k^r n^s = [h^{(r)}(M)][h^{(s)}(M)]$. Thus, as a bi-variate polynomial, it is a Probability Chance Polynomial i.e. coefficients are probabilities/chances.

Probability/Chance theoretic interpretation of probability/Chance polynomials: random/chance fields

Consider a collection of discrete random/chance variables, assuming integer values be denoted by $\{X_m\}_{m=1}^{\infty}$. As in the case of Faulhaber polynomials, we have a collection of marginal probability/chance mass functions associated with this collection of independent random/chance variables. We associate the following transformed collection of random/chance variables:

$$Y_m(n) = n^{X_m} \text{ for } 1 \leq m \leq \infty \text{ where } n \text{ is a deterministic variable.}$$

We consider the expected value of the new random/chance variables i.e. $E(Y_m(n))$. Specifically in the case of Faulhaber sums, it is easy to see that

$$E(Y_m(n)) = h^{(m)}(n) = \sum_{k=1}^n k^m$$

Thus, in the case of Faulhaber sums, we have the expectation of an infinite collection of random/chance variables (i.e. we have the associated random/chance process). Also, since X_m, s are independent random/chance variables, it is easy to see that

$$E(Y_r(n) Y_s(n)) = E(Y_r(n)) E(Y_s(n))$$

Note

In this computation of expectation, we use the fact that the probability/chance mass function of sum of two independent random/chance variables is the convolution of their individual probability/chance mass functions.

- Also, if the value of 'm' is fixed and the variable (like the discrete time parameter) 'n' is increased, we have a countably infinite collection of random/chance variables that are independent (i.e. we have a random/chance process). For instance in the case of 'm=1', the random variable X_m assumes the values $\{m^2, m\}$ with the associated probabilities $\{\frac{1}{2}, \frac{1}{2}\}$. As 'n' is increased this probability mass function is constant.

- Thus, as 'm', 'n' are varied, we have a bi-variate random/chance process or also denoted as a 'random/chance field'. Specifically, the collection of Faulhaber polynomials correspond to the expected values of the random/chance variables belonging to the random field. Other moments can easily be computed. It is expected that this probability/chance theoretic interpretation can have associated explanation from physics/chemistry view point.

- The polynomial $\{h^{(s)}(m)\}$ can also be interpreted as the expectation of a discrete chance/random variable assuming higher powers of "m" with chances/probabilities corresponding to the coefficients (of the probability polynomial). Thus various results from probability/chance theory can be invoked in association with the above polynomials.

In view of the above discussion, we are naturally motivated to arrive at the following generalization:

Definition: A power series is defined to be a "probabilistic power series" if all the coefficients are non-negative and their sum converges to one.

Definition: A power series is defined to be a “chance power series” if some of the coefficients are negative and their sum converges to one.

Consider a discrete random variable, W which assumes non-negative integer values with the associated probability/chance mass function. Define the associated random variable

$$Z = x^w \text{ where } x \text{ is a deterministic variable.}$$

It is easy to see that the expectation of the random variable Z is a probability/Chance power series.

Algebra of Faulhaber Polynomials

It is well known that for ‘odd’ values of ‘p’, the Faulhaber polynomial has factors n^2 and $(n+1)^2$. Also, for even values of ‘p’, Faulhaber polynomial has factors ‘n’, $(n+1)$ and $(n + \frac{1}{2})$. Thus the following hold true:

- When any (even or odd) Faulhaber polynomial is divided by $h^{(1)}(n) = \frac{n(n+1)}{2}$ then the remainder is zero and the quotient is a probability/chance polynomial.

- When the Faulhaber polynomial for even values of ‘p’ is divided by $n(n+1)(n + \frac{1}{2})$ then the remainder is a zero and the quotient is a polynomial whose coefficients add upto $\frac{1}{3}$

Sum of Pairwise, Triplewise Products of Consecutive Integers: Polynomial Expressions

It is easy to see that the following formula can be derived.

$$\sum_{i=1}^n \sum_{j=1}^n (i)(j) = (1+2+\dots+n)^2 - (1^2+2^2+\dots+n^2) = \frac{3n^4 + 2n^3 - 3n^2 - 2n}{12} = [h^{(1)}(n)]^2 - h^{(2)}(n)$$

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (k)(i)(j) = (1+2+\dots+n)^3 - (1^3+2^3+\dots+n^3) = \frac{n^6 + 3n^5 + 3n^3 - 2n^2}{4} = [h^{(1)}(n)]^3 - h^{(3)}(n)$$

$$i \neq j \neq k$$

- In the spirit of the above sums, we now arrive at the most general case. It is not clear why such expressions should be polynomials. The following reasoning shows that they are infact polynomials.

$$\sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_p=1}^n (i_1)(i_2)\dots(i_p) = (1+2+\dots+n)^p - (1^p+2^p+\dots+n^p) = [h^{(1)}(n)]^p - h^{(p)}(n)$$

$$i_1 \neq i_2 \neq \dots i_p$$

It is easy to reason that the above expression represents a polynomial (for fixed ‘p’ and variable ‘n’), the sum of whose coefficients is zero. Let us name label such polynomials as “Zero Polynomials”. In view of such polynomials, the following Lemma is interesting.

Lemma 4.6: “Zero” polynomials with the coefficients from the field of real numbers form a “ring” (as an algebraic structure)

Proof: Follows from the verification of axioms of a ring. Q.E.D

- Using a similar reasoning, it can be shown that the sum of pairwise, triplewise etc products of consecutive integers (for fixed ‘p’) constitute “zero” polynomials in the variable ‘n’.

- Using the above and similar formulae, congruence properties of

products of first ‘n’ integers etc can be derived.

Goal: To study the properties of above chance/probability polynomials and other polynomials naturally derived from them.

- It is easy to derive the following identities

$$\sum_{n=1}^M h^{(1)}(n) = \frac{M(M+1)(M+2)}{6} = \frac{M^3 + 3M^2 + 2M}{6}$$

Also, it is easy to see the above function is a probability polynomial.

$$\sum_{n=1}^M h^{(2)}(n) = \frac{M(M+1)(M^2 + 3M + 2)}{12}$$

Once again the resulting function is a probability polynomial.

- From the expressions for $\{h^{(s)}(n)\}$ for various values of ‘s’, it is easy to see that $\{0, -1\}$ are the common zeros of all the polynomials. Thus, we are motivated to define the following function:

$$g(s, n) = \sum_{i=1}^n i^s$$

where ‘n’ assumes integer values only but ‘s’ could be arbitrary real/complex number. It is easy to see that $g(s, n+1) = g(s, n) + (n+1)^s$

Thus, by mathematical induction on ‘n’, it is easy to reason that $\{0, -1\}$ are zeros of the above function.

- One interesting goal is to determine whether the zeros correspond to a “soluble group” and thus understand the geometry of zeros.

- In the case of the function $g(s,n)$, allowing $\{n, s\}$ to be arbitrary complex numbers leads to the study of a function of two complex numbers.

- Probability/Chance Polynomials: Electrical Network Theory: It should be noted that the first three Faulhaber polynomials are probability polynomials and all others are probability type polynomials (chance polynomials). First three probability polynomials encountered above (with the sums considered) have zeroes in the left half s-plane only with a zero at location “zero” in s-plane. Such polynomials arise naturally in the case of electrical networks (and also stable linear time invariant systems). In fact “positive real functions” are well studied in network theory.

Remark

Every polynomial with non-negative coefficients can be normalized to arrive at probability polynomial without affecting the zeroes (i.e. both polynomials have same set of zeroes). Also every probability polynomial with strictly positive coefficients cannot have real zeroes on the positive real axis. If some coefficients of a probability polynomial are zeroes, then the polynomial is unstable in the sense of Routh’s stability test.

- The class of all polynomials can be divided into the following classes:

(A) Polynomials which are such that the sum of all coefficients is zero. We call them “Zero polynomials”

(B) Polynomials which are such that the sum of all coefficients is a non-zero quantity.

Remark

Polynomials in the above class (B) can be normalized to arrive at a chance polynomial (realcomplex valued chances) without affecting the zeroes.

• In the following discussion (in this subsection), we consider polynomials that donot have a zero at the origin (i.e. n=0). If a zero exists at origin it is factored out.

• **Lemma 4.7:** For $p \geq 4$ Faulhaber polynomials $h^{(p)}(n)$ have zeroes in the right half s-plane and are thus unstable in the sense of Routh's stability test

• **Proof:** It is well known that a necessary but not sufficient condition for a polynomial to have zeroes in the left half s-plane is that all the coefficients of the polynomial be present and all of them have positive sign (in the presence of atleast one positive coefficient). Thus, using Lemma 4.2, it is clear that the Faulhaber polynomials $\{h^{(p)}(n)\}$ for $p \geq 4$ contain zeroes at the origin as well as right half s-plane. In the sense of Routh's stability test they unstable polynomials. Applying Routh's stability test on various Faulhaber polynomials, the number of zeroes in the right half plane could be determined. Q.E.D.

Note: All strictly chance polynomials are unstable in the sense of Routh's stability test.

• **Multi-Variate Chance/Probability Polynomials**

Based on the above discussion, we are naturally led to the following definition.

Definition: A homogeneous/non-homogeneous, seperable/nonseperable multi-variate probability polynomial is one where all the coefficients are non-negative and the sum of all the coefficients is one. It is a "chance" polynomial when some coefficients are negative, but the sum of them is one.

Example: We now provide a seperable, non-homogeneous bivariate probability polynomial which naturally arises in connection with the above sums.

$$[h^{(r)}(m)][h^{(s)}(n)] = \left[\sum_{j=1}^m j^r \right] \left[\sum_{k=1}^n k^s \right] = \sum_{j=1}^m \sum_{k=1}^n j^r k^s$$

• **Goal:** To study properties of Uni/Multi-Variate, seperable/nonseperable, homogeneous/non-homogeneous probability/chance polynomials. Of particular interest are the probability/chance quadratic forms.

• In view of the definition of probability/chance polynomials, we are naturally led to the following definitions. We only consider chance polynomials keeping in mind that similar definitions hold true for probability polynomials:

Definition: A polynomial is called "Sub-Chance Polynomial" when the sum of all of its coefficients is smaller than one.

Definition: A polynomial is called "Super-Chance Polynomial" when the sum of all of its coefficients is greater than one.

It is easy to see that the product of finitely many sub-chancesuper chance polynomials is a sub-chancesuper chance polynomial.

• **Novel Power Series based on Faulhaber—Type Sums**

In the above discussion, we studied some properties of Faulhaber polynomials. Now, we propose a power series in variable 'x', whose coefficients constitute Faulhaber polynomials in M i.e. $h^{(p)}(M)$ integer values of 'p'.

Consider the following power series:

$$g(x) = \frac{1}{1-x} + \frac{1}{1-2x} + \dots + \frac{1}{1-Mx}$$

with 'M' being an integer. In the Region of Convergence (ROC) i.e. $\{x : |x| > 1\}$, the corresponding coefficient sequence constitutes a right-sided/causal sequence. It is easy to see that

$$g(x) = 1 + \sum_{p=1}^{\infty} h^{(p)}(M) x^p \text{ where}$$

$$h^{(p)}(M) = \sum_{n=1}^M n^p \text{ i.e. Faulhaber sum.}$$

• The above power series naturally motivates the study of power series in variable 'x' whose coefficients are interesting structured polynomials. For instance, the coefficients constitute arbitrary probability/chance polynomials, [4].

• As a natural generalization, we consider the following power series in 'x':

$$\frac{1}{1-e_1x} + \frac{1}{1-e_2x} + \dots + \frac{1}{1-e_Mx} \text{ where}$$

$$\{e_i\} \text{ are consecutive integers i.e. } e_{i+1} = e_i + 1$$

Once again the coefficient sequence can be expressed in terms of Faulhaber sums i.e.

$$\sum_{j=1}^M e_j^p = h^{(p)}(e_M) - h^{(p)}(e_1 - 1)$$

Thus, the coefficients are the difference of two Faulhaber polynomials. Hence the sum of coefficients of such polynomials is zero.

• In the spirit of the above discussion, it is only natural to consider a power series in 'x' whose coefficients are ratio of two structured polynomials (say probability/chance polynomials) as in the case of hyper-geometric series.

• Now let us consider another power series:

$$\tilde{g}(x) = \frac{1}{1-x} + \frac{1}{1-\frac{x}{2}} + \dots + \frac{1}{1-\frac{x}{M}}$$

with the associated Region of Convergence (ROC) $\{x : |x| > M\}$ In the ROC, the coefficient sequence constitutes a right sided sequence. It is easy to see that

$$\tilde{g}(x) = 1 + \sum_{p=1}^{\infty} \phi^{(p)}(M) x^p \text{ where}$$

$$\phi^{(p)}(M) = \sum_{n=1}^M \frac{1}{n^p}$$

i.e truncated Euler Sum as in Zeta function

• Now consider a power series of the following form:

$$k(x) = \frac{1}{(1-x)(1-2x)\dots(1-Mx)} \text{ with ROC } \{x : |x| > 1\}$$

By partial fraction expansion method, K(x) can be written as

$$k(x) = \frac{c_1}{1-x} + \frac{c_2}{1-2x} + \dots + \frac{c_M}{1-Mx}$$

The constants i.e. $\{c_i\}$ scan easily be evaluated (as the poles are simple and distinct). Also the coefficient sequence of K(x) are easily related to the Faulhaber-type sums.

• Alternatively, by a suitable choice of constants i.e. $\{d_i\}$ s in the following power series $\frac{d_1}{1-x} + \frac{d_2}{1-2x} + \dots + \frac{d_M}{1-Mx}$

various interesting Faulhaber-type sums can be obtained as the coefficient sequence (of the associated power series).

5. Theory of L¹ Chances: Stochastic Chains

In the technical report [9] and the research paper [10], the author formalized the theory of chances. In those research works, L^{2p}-E" chances were formalized in detail. Also stochastic chains based on such chances were formalized.

In the following, we specifically discuss L¹-E" chances and the stochastic chains based on them.

- For the sake of concreteness, we consider the case where the cardinality of space of outcomes is finite. Let it be denoted by α . Also, let the collection of subsets of α be denoted by β . Also, let γ be the "chance measure". These entities satisfy the following: $\alpha = A \cup B$, where $A \cap B = \emptyset$ and

$$\begin{aligned} \gamma(w) &\geq 0 \text{ for } w \in A, \\ \gamma(w) &\leq 0 \text{ for } w \in B, \end{aligned}$$

Also, for all $w, \gamma(w) \in [-1, 1]$.

Furthermore

$$\sum_{w \in \alpha} \gamma(w) = 1$$

The chance measure of any set, C in β is computed in the following manner i.e. $\gamma(c) = \sum_{w \in c} \gamma(w)$

- The above condition (6.1) leads to the following classification of L¹- chances

Type-A L¹- chances: For all $w \in \alpha, \gamma(w) \in [-1, 1]$ For instance, it can be shown that the coefficients of Faulhaber polynomials $h^{(p)}(n) = \sum_{k=1}^n k^p$ for $p > 3$.

Type-B L¹- chances: For some $w \in \alpha, |\gamma(w)| > 1$ Specifically the L¹- chances value is negative for some 'w' and greater than -1. But $\sum_{w \in \alpha} \gamma(w) = 1$

For instance, in section (4), we realized that $h^{(s)}(m)$ for odd values of "s" can be expressed as a polynomial in 'a',

Where

$$h^{(2p+1)}(m) = \sum_{k=1}^m k(2p+1) = c_1 a^2 + c_2 a^3 + \dots + c_p a^{p+1} \text{ and } a = \frac{m(m+1)}{2}$$

It can be seen that as a polynomial in 'a', the coefficients of $h^{(2p+1)}(m)$ constitute "Type-B" L¹- chances For instance, we have that

$$h^{(7)}(m) = \sum_{k=1}^m k^{(7)} = 2a^4 - \frac{4}{3}a^3 + \frac{1}{3}a^2$$

- The concepts such as conditional chance, independence etc are consistently defined. Thus an effort is made to develop a consistent theory of chances where the space of outcome has finite/countable/uncountable number of outcomes.

- Also joint/marginal chance variables are defined like random variables. In the case of discrete chance variables, there is an associated chance mass function. For instance, let us consider few examples that are parallels of probability mass functions (of certain classical random variables):

(A) Bernoulli--Type Chance Variable: Consider an experiment in which there are only two possible outcomes i.e. $\alpha = \{w_1, w_2\}$

$$\text{Let } \gamma(w_1) = q \text{ and } \gamma(w_2) = 1 - q \text{ with } q < 0$$

It should be noted that we can have Type (A) or Type (B) chances. The above provides the chance mass function of Bernoulli chance variable.

(B) Binomial Chance Variable, Y: Consider the following chance mass function (with $C(Y = w_i) = \gamma(w_i)$ denoting the chance of the outcome w_i). Let there be M possible outcomes

$$\gamma(w_j) = \binom{M}{j} (q)^j (1-q)^{M-j} \text{ with } q < 0$$

Once again we can have Type (A) or Type (B) chances.

(C) Geometric Chance Variable: Consider an experiment with countably many outcomes. The following chance mass function corresponds to a geometric chance variable:

$$\gamma(w_j) = (q)^j (1-q) \text{ with } q < 0 \text{ for } 0 \leq j < \infty$$

In this case, we are restricted to Type (A) chances.

(D) Poisson Chance Variable: Once again consider an experiment with countably many outcomes. The Poisson chance mass function is given by

$$\gamma(w_j) = \frac{e^{-\mu} \mu^j}{j!} \text{ with } \mu < 0$$

Also in the case of continuous chance variables there is an associated chance density function.

Remark: Any finite/infinite sequence of real numbers (positive as well as negative valued) which can be normalized to sum (converge) to one, constitutes a Chance Mass Function. As in the random variable case, certain chance mass functions find many applications in science and technology.

Stable Continuous Time Linear Time Invariant Systems: Chance Density Functions

- Suppose we consider a chance space which consists of uncountable number of outcomes. In such a case we are naturally led to consider the chance density function. Once again, it is clear that such a function can assume negative values, but must be integrable. In this connection, we realize that the impulse response of a Continuous Time (CT), Linear Time Invariant (LTI) system must be integrable. If we consider the Laplace transform of such a function (i.e the transfer function of a CT LTI system), we are naturally led to consider the special case of systems with rational transfer functions. It is well known that the impulse response of stable CT LTI system (whose Laplace transform is a rational function) is integrable and hence all the poles of the transfer function lie in the left half s-plane. Thus such impulse response functions constitute Chance Density Functions. Hence many interesting examples of chance density functions can be specified by considering stable LTI systems.

- It is interesting to note that the rational transfer functions all of whose poles lie on the negative real axis, lead to interesting examples of Chance Density Functions (i.e. linear combination of decaying exponential signals). This class of chance density functions are defined in the same spirit of the family of Erlang distributions.

- Any real valued function (not necessarily non-negative) whose integral is finite can be normalized to arrive at a chance density.

- As in the case of stochastic processes in probability theory, we define a “chance process” to be a “countable/uncountable” collection of “chance variables” (like random variables) indexed by the “time” parameter. Similarly “chance” fields are defined.

- Stochastic chains in the spirit of Markov chains (based on L^{2p} chances) are discussed in [9]. In the following, we briefly discuss stochastic chains based on L^1 chances. Detailed discussion is based on the ideas documented in [9].

Definition: A state transition chance matrix, C is defined as one where the elements can be negative/zero/positive assuming values in the interval $[-1, 1]$ and the row sums are all equal to one. The matrix can be finite dimensional or infinite dimensional. For instance, the following matrix is a state transition chance matrix of a stochastic chain satisfying Markovian type property.

$$C = \begin{pmatrix} 0.6 & 0.8 & -0.4 \\ 0.3 & -0.2 & 0.9 \\ -0.5 & 0.8 & 0.7 \end{pmatrix}$$

Now we consider a finite dimensional “chance matrix”, C leading to the definition of a “chance process”. In this case the chance process is a stochastic chain in the spirit of a Discrete Time Markov chain.

(1) Let $\mathcal{G}_n(j)$ denote the chance that the Discrete Time stochastic chain is in the state ‘j’ at time ‘n’. Also let \mathcal{G}_n denote the row vector of chances that the process is in one of the states. Since the stochastic chain satisfies the Markovian type property, we have that

$$\bar{v}_{n+1} = \bar{v}_n C \text{ for } n \geq 0$$

Let \bar{e} denote a column vector of ‘ones’. By definition of C, we have $C\bar{e} = \bar{e}$

(2) Thus, one of the eigenvalues of state transition chance matrix, C is “one”. Also, it is easy to see that the product of any two chance matrices is another chance matrix. Also higher powers of a chance matrix is another chance matrix.

Lemma 5.1:

Suppose the state transition chance matrix C is diagonalizable and has spectral radius ‘one’. Then the limiting equilibrium chance mass function (equilibrium chance distribution) exists and is given by the left eigenvector corresponding to eigenvalue ‘one’ of the matrix C.

Proof: Since C is diagonalizable, we have that $C = \sum_{i=1}^M \mu_i \hat{f}_i \hat{g}_i$ where $\hat{f}_i \hat{g}_i$ are the left, right eigenvectors of C corresponding to the eigen value μ_i respectively.

Since spectral radius is one (i.e. all the eigenvalues other than ‘one’ are strictly inside the unit circle), we have that

$$\lim_{n \rightarrow \infty} \bar{\mathcal{G}}_{n+1} = \overline{\mathcal{G}(0)} \hat{f}_1 \hat{e}$$

Where \hat{f}_1, \hat{e} are the left, right eigenvector of C corresponding to the eigen value one respectively.

Thus, we have that $\lim_{n \rightarrow \infty} \bar{\mathcal{G}}_{n+1} = \hat{f}_1$

Thus, the equilibrium chance vector is the left eigenvector of C corresponding to the eigenvalue one of C.

Note: The proof could be generalized for non-diagonalizable L^1 -chance matrices with spectral radius one.

Complex Valued L^1 . Chances

In [9], we have discussed the idea that $L_{2,p}$ chances can be complex valued. In fact novel stochastic chains based on such complex valued chances are discussed. Now we consider the case of L^1 - chance complex valued chances. In other words chances associated with outcomes are complex valued, lie within the unit circle and the sum of all chances is equal to one. Naturally, we are motivated to consider the state transition chance matrix with L^1 - chances along the rows i.e. components of the matrix are complex valued, lie in the unit circle and along all the rows they add upto one. As in the real valued chance case, under reasonable assumptions existence and uniqueness of equilibrium chance mass function (chance distribution) is established.

Mathematicians developed structures such as quaternions, octanions generalizing the idea of field of complex numbers. Logically, it should be possible to allow the chances to be quaternions, octanions.

Motivation for Alternative Theories of Uncertainty

Consider a space of outcomes \mathcal{A} , which contains finitely many outcomes. In classical/modern probability theory, the “probability” of an outcome lies in the interval $[0,1]$ and the sum of probabilities of all outcomes is one. In the above section, we considered an alternative theory of uncertainty i.e. “theory of chances”.

Suppose we now consider the most general scenario. The chances in the space of outcomes are such that the sum of chances of all outcomes assumes two possible values:

- (i) zero value and
- (ii) Non-zero value not necessarily equal to one.

Theory of Balanced Chances [9]

In the above discussion and [9], we allowed the chances (L^{2p} chances, L^1 chances) to assume negative values as well as complex (numbered) values unlike probabilities. Also, in the case of finite/countable number of outcomes, the associated chances are assumed to sum to one (as in the case of probabilities). We now consider the case where the space of outcomes is finite.

But in this case, the chances of all possible outcomes (lie within the interval $[-1, 1]$ in the real case and in the unit circle in the complex numbered case) add upto “zero”. Thus, the chance of outcomes assume positive as well as negative values. We call such chances “Balanced Chances”. In summary, real valued (positive or negative valued) or complex valued chances of all possible outcomes add upto zero. $f \in \{f_{\min}, f_{\min} + 1, f_{\min} + 2, \dots, f_{\max}\}$

Theory of Symmetric Balanced Chances

In this theory we have a special case where for every outcome with “positive” chance, there is a corresponding outcome with chance that is equal to it in magnitude but opposite in polarity (i.e. outcome with negative chance).

- In the case where the number of outcomes is countable (and not finite) the sum of chances of all outcomes (possibly negative or complex valued) converges to a zero value.

Dichotomy of Absolutely Summable Sequences

Suppose a function of single variable is integrable. The value of

integral is either zero or non-zero. It is easy to see that if the value of integral is non-zero, then the function can be normalized to arrive at a chance density function (by dividing the function with the non-zero value of the integral).

Note: We considered certain bounded infinite sequences/functions whose associated infinite series sum/integral is infinity (i.e. they diverge/are non integrable). In [13], it is shown that such infinite sequences/functions can be “modified” to arrive at the associated infinite series/integral that are convergent/integrable. Thus, we can associate such sequences/functions with “chance spaces”/“probability spaces” (associated chance/probability mass functions or “chance/probability density functions”).

When the space of outcomes contains uncountably many outcomes, the “balanced” chance density is integrable and the integral is equal to zero. Naturally, we consider the state transition balanced chance matrix with L^1 – balanced chances along the rows i.e. components of the matrix are real/complex valued (each of them lie in the interval $[-1, 1]$ in the real case and within the unit circle in the complex case) and along the rows they add upto one. As in the case of real/ complex valued ordinary L^1 chances, with L^1 – balanced chances under reasonable assumptions, existence and uniqueness of equilibrium chance mass function is established.

An Application of the Theory of Chances: Graph Theory

In [11], the author defined a probability mass function on the vertices of a graph. The definition involves normalization of the vertex degree distribution. It is given by:

$$p_i = \frac{\text{degree of } i^{\text{th}} \text{ vertex}}{\text{sum of degrees of all vertices of graph}} \quad \text{for } 1 \leq i \leq M$$

(Vertex set V)

Using this probability mass function, the author associated Shannon entropy with the graph i.e. $H(p_1, p_2, \dots, p_M)$.

Motivated by the above idea, in the following, we consider a directed, weighted graph. Degree of i^{th} vertex = In Degree of Vertex $V_{in}^{(i)}$ - Out Degree of Vertex $V_{out}^{(i)} = \widehat{V}_i$

Suppose, we consider graphs for which

$$D = \sum_{i \in V} (V_{in}^{(i)} - V_{out}^{(i)}) \neq 0$$

Now, we normalize \widehat{V}_i by D for $1 \leq i \leq M$ i.e.

$$q_i = \frac{\widehat{V}_i}{D} \quad \text{for } 1 \leq i \leq M$$

It is easy to see that $\{q_1, q_2, \dots, q_m\}$ is a “chance mass function” and not necessarily a probability mass function (since the out degree of a vertex can be larger than the in degree of a matrix). The case where D equals zero is considered in [12]. As discussed above, let us consider a weighted, directed graph. As in the case of [11], the weights are normalized i.e.

$$\widehat{W}_{ij} = \frac{W_{ij}}{\sum_{j=1}^M W_{ij}} = \frac{W_{ij}}{E_i}$$

It should be noted that in the above equation, \widehat{W}_{ij} can assume negative values (more generally values can be complex numbers). Under the assumption that $E_i \neq 0$ for all i the matrix of normalized weights i.e. $\{\widehat{W}_{ij}\}$ is a state transition chance matrix. Thus, a stochastic

(chance) chain is naturally defined based on such a matrix (as in the case of a Discrete Time Markov chain associated with the graph when all the weights are non-negative). Such a stochastic chain (in the spirit of Markov Chain) is naturally associated with a weighted, directed graph in which the weights can assume negative values.

- Detailed development of “theory of chances” is based on the work documented in [Rama3].
- **Note:** It should be noted that chance/probability spaces with finite/countable number of outcomes are naturally associated with polynomials/power series. Also they are naturally associated with finite/infinite dimensional vector spaces.

6. Theory of Random Sets (based on random intervals) and Theory of Granular Sets: Novel Models of Uncertainty

Based on our results dealing with theory of chances, we are naturally led to the following question:

Q: Can the measure of an interval or more generally a subset of real line be a random variable?

Based on very practical considerations (in sensor fusion), we are led to the answer that such a thing can happen. Let us consider a collection of ‘M’ sensors measuring temperature (or some other variable based on some physical phenomena). Let the crisp measurements measured by the sensors be $\{x_1, x_2, \dots, x_M\}$. Since the sensors are not absolutely sure of the recorded value, a “tolerance” is added on the left side as well as right side. In the simplest case, the tolerance value is deterministic, say “T”. It is also the same on the left side and the right side of the crisp measurement. Thus the interval valued outputs from the sensors are given by $[x_1 - T, x_1 + T], [x_2 - T, x_2 + T], \dots, [x_M - T, x_M + T]$

But in the case of many practical applications, the tolerance is not a constant value. It is characterized by probabilistic uncertainty and constitutes a random variable. It is reasonable to assume that the random tolerance value is constant on the left as well as right side of tolerance. Thus, the outputs of the sensors are given by $[x_1 - D_1, x_1 + D_1], [x_2 - D_2, x_2 + D_2], \dots, [x_M - D_M, x_M + D_M]$

Hence, in this case, the outputs of sensors are RANDOM INTERVALS. It is reasonable to assume that the random variables D_1, D_2, \dots, D_M are independent and identically distributed.

The “length” of interval from the first sensor is $2D_1$. Thus the length is a random variable.

As in the case of Lebesgue’ theory of measure, the “measure” of an arbitrary subset of real line is characterized by a random variable. The author is currently developing a formal theory of random sets in which the “measure” of a set is a random variable. The theory could be related to stochastic calculus.

Theory of Graded Sets

In the theory of rough sets, a set A is specified by a tuple of sets (A_{lower}, A_{upper}) where A_{lower} is a subset of A_{upper}

Detailed theory of rough sets has been developed and applied in various fields of human endeavour. The author conceived the following generalization of a rough set called graded set.

$A = (A_1, A_2, \dots, A_L)$ where the AC sets satisfy the condition that they are a finite collection of nested sets i.e

A_i is subset of A_{i+1} for $1 \leq i \leq (L-1)$

Such a set potentially arises in many applications. One such application is sensor fusion of interval valued measurements. Particularly we consider the “F” fusion function proposed by Schmidt et al. We now explain the approach to arrive at the fused interval based on finitely many measurement intervals.

Let $[a_1, b_1], [a_2, b_2], \dots, [a_L, b_L]$ be the measurement intervals and let there be “F” faulty intervals (we do not know which ones are faulty). The fused interval estimate $[c, d]$ is obtained in the following manner:

c (i.e. left end point of fused estimate) is determined in the following manner. Arrange the left end points i.e. a_i 's from smallest to largest value. Counting down from the highest value i.e. a_L consider the $f + 1^{th}$ measurement. Consider it as ‘c’. d (i.e. right end point of fused estimate) is determined in the following manner. Arrange the right end points i.e. b_i 's from smallest to largest value. Counting up from the smallest value i.e. b_1 consider the $f + 1^{th}$ measurement.

Consider it as ‘d’.

The fused interval estimate is $[c, d]$. It is clear that the end points c, d depend on the number of faulty intervals (from say sensors). We are interested in understanding how graded sets naturally arise when the number of faulty intervals (from sensors) is varied from a smallest to largest value.

Specifically, let $f \in \{f_{min}, f_{min} + 1, f_{min} + 2, \dots, f_{max}\}$ From the above specified method of computing the left and right end points of fused estimate, it is easy to see that when f is increased from f_{min} to f_{max} the collection of fused intervals (labeled as) A_i 's lead to a graded set i.e.

A_i is a subset of A_{i+1} for $f_{min} \leq i \leq f_{max}$

i.e. they are nested intervals constituting a “graded” set. Thus, in summary, with the number of faulty intervals/sensors as a parameter, the output of fused function is a graded set.

• Random graded set

Consider the case where the number of faulty sensors is a discrete random variable. As a consequence, the output of “F”- fusion function assumes various intervals with associated probabilities. Hence when the number of faulty sensors is varied from the lowest to the highest value, we arrive at what we call as a RANDOM GRADED SET.

The above concept motivates the idea where the outcomes in a probability space are mapped to sets (unlike the concept of a random variable where the outcomes are mapped to real/complex numbers) i.e. the domain of mapping is outcomes in a finite/countable set and the range is sets. We now propose another interesting model of uncertainty

Fuzzy rough set:

A Rough set, B is specified as

$$B = [B_{lower} B_{upper}] \text{ with } B_{lower} \text{ a subset of } B_{upper}$$

Also, $B_{lower} B_{upper}$ are crisp sets. we now define a Fuzzy Rough set as the one where $B_{lower} B_{upper}$ are fuzzy sets with B_{lower} a subset of B_{upper}

Thus, since B_{lower} a subset of B_{upper} associated membership functions satisfy the following:

$$\mu_{B_{lower}}(x) \leq \mu_{B_{upper}}(x) \text{ for every } x \in X \text{ the universe of discourse}$$

Hence, such a “model of uncertainty” applies when the associated membership function of Fuzzy Rough Set, B i.e. $\mu_B(\cdot)$ has the “lower”

and “upper” limits.

Goal: To develop the detailed theory of FUZZY—ROUGH SETS: We are first interested in simple operations such as union and intersection: (I) Intersection of Fuzzy Rough Sets B, C:

$B = [B_{lower} B_{upper}], C = [c_{lower} c_{upper}]$ and $D = B \cap C$ the operation is fuzzy intersection.

The membership function of Fuzzy Rough set D has lower and upper limits:

$$\mu_D(x) = [\min\{\mu_{B_{lower}}(x), \mu_{C_{lower}}(x)\}, \min\{\mu_{B_{upper}}(x), \mu_{C_{upper}}(x)\}]$$

Union of Fuzzy Rough Sets B, C:

$$B = [B_{lower} B_{upper}], C = [c_{lower} c_{upper}] \text{ and } E = B \cup C$$

the operation is fuzzy union. The membership function of Fuzzy Rough set E has lower and upper limits:

$$\mu_E(x) = [\max\{\mu_{B_{lower}}(x), \mu_{C_{lower}}(x)\}, \max\{\mu_{B_{upper}}(x), \mu_{C_{upper}}(x)\}]$$

Connectionist structures: models of uncertainty

In applied mathematics graphs are proposed as the simplest connection structures. They found many applications in computer science and other research areas. Introducing probabilistic uncertainty, random graphs are proposed. Such random graphs are obtained by starting with certain number of isolated vertices and randomly adding certain edges between the vertices. Detailed theory of random graphs with several interesting results is developed.

Generalization of random graphs

We now propose a “connectionist structure” in which there are certain number of isolated vertices and edges of varying “connection strength” (positive or negative) are randomly added between them. Thus in our formulation, the connection strength of an edge between two vertices is a discrete/continuous random variable (Note: In the case of random graph, the connection strength is a Bernoulli random variable). The connection strength could be normalized to assume finite/countably many values lying in $[0,1]$. The questions of interest in such connectionist structures deal with probability distributions on graphs of various connections strength profiles. We expect many applications for such a model in science and engineering.

**In the case of such non-deterministic connection structures, same issues questions/theorems/concepts as in the case of conventional random graphs are addressed and derived. For instance, one possible goal of such a study is to determine at what stage a particular property of the graph is likely to arise. We now provide some formal results

• Let their N vertices and the connection strength on the associated edges be represented by independent random variables i.e. $\{X_i\}_{i=1}^N$

Definition: Let the associated connection structure (generalized random graph) as being “connected” if the connection strength on all edges is greater than or equal to a certain lower limit, say ‘k’.

Let the associated random variable Z be

$$Z = \text{Min}\{X_1, X_2, \dots, X_N\}$$

$$\text{Prob}\{\text{Generalized Random Graph is connected}\} = \text{Prob}\{Z \geq k\} \\ = \text{prob}\{X_1 \geq k, X_2 \geq k, \dots, X_N \geq k\}$$

Since the random variables $\{X_i\}_{i=1}^N$ are independent, we have that

Prob {Generalized Random Graph is connected} = $\prod_{i=1}^N p\{X_i \geq K\}$

Suppose the random variables $\{X_i\}_{i=1}^N$ are identically distributed. Then, we have

Prob {Generalized Random Graph is connected} = $(p\{X_i \geq k\})^N$

Using Chebyshev inequality, this quantity can be easily upper bounded. Equivalently, the cumulative distribution function (CDF) of the random variable Z is given by

Prob $\{Z \leq k\} = F_z(k) = 1 - (1 - F_{X_i}(k))^N$

If we now fix 'k' and let N tend to infinity (i.e. the number of vertices in the generalized random graph approaches infinity)

i.e. asymptotically the generalized random graph becomes connected.

Let the random variables $\{X_i\}_{i=1}^N$ corresponding to edge connection strengths be non-negative, independent and identically distributed with mean μ Define a new random variable Y in the following manner:

$$Y = \frac{1}{N} \left(\sum_{i=1}^N X_i \right)$$

It is easy to see that the expected value of random variable Y is given by

$$E[Y] = \mu$$

This gives the average connection strength on "N" edges of a generalized random graph. As in the discussion above, we would like to compute the probability that the "connection strength" on all "N" edges of a random graph is atleast L_μ (where "L" is an integer).

Prob $\{X_1 \geq L_\mu, X_2 \geq L_\mu, \dots, X_N \geq L_\mu\}$

Now, we invoke Markoff inequality for upper bounding the probability

Prob $\{X_i \geq L_\mu\}$
 Prob $\{X_i \geq L_\mu\} \leq \frac{\mu}{L_\mu}$ i.e. We have that prob $\{X_i \geq L_\mu\} \leq \frac{1}{L}$

Hence, we readily have the following bound on the probability that the connection strength on all N edges is atleast L_μ

Prob $\{X_1 \geq L_\mu, X_2 \geq L_\mu, \dots, X_N \geq L_\mu\} \leq \left(\frac{1}{L}\right)^N$

• Thus, we could also be interested in determining the probability distribution on graphs/connectionist structures of certain minimum connection strength on all edges (For example cycles of certain minimum strength, open paths of certain minimum strength etc).

• We can consider such a connectionist structure (generalized random graph) as being an "ordinary random graph if the connection strength on an edge exceeds a lower limit.

• Fuzzy Uncertainty based Connectionist Structures

In this model of uncertainty on a connectionist structure (such as graph) with fixed number of vertices, the connection strength is associated with a Fuzzy set (of finite/countable/uncountable support). The membership values are naturally associated with connection strength of edges/links. We now introduce the following concept as a special case

Definition: A fuzzy graph is a connectionist structure in which the "connection strength" of an edge assumes only two values ("link exists" or "link does not exist") with associated membership values.

we are interested in issues such as the

(i) Membership value that a fuzzy graph is a tree, a spanning tree etc

(ii) Membership value that a fuzzy graph is connected

(iii) Membership value associated with closed paths (cycles), open paths etc

7. Conclusions

In this research paper, the innovative idea of allowing "chances of outcomes" to assume negative as well as complex (numbered) values is proposed. Based on such an idea, the theory of real/complex valued L^{2p} chances is discussed. Also, novel class of real/complex valued chance based stochastic chains are discussed. The theory of L^1 chances is also discussed. The research area of chance number theory is briefly discussed.

References

- Berndt BC (2006) Number Theory in the Spirit of Ramanujan American Mathematical Society.
- Hopfield JJ (1982) Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proceedings of National Academy of Sciences USA 79: 2554-2558.
- Jagadeesh G, Praveen D, Rama Murthy G (2007) Heteroassociative Memories on the Complex Hypercube. Proceedings of 20th IJCAI Workshop on Complex Valued Neural Networks and Neuro-Computing: Novel Methods, Applications and Implementations.
- Knopp K (1990) Theory and Application of Infinite Series. Dover Publications Inc, New York.
- Papoulis A (2014) Probability, Random Variables and Stochastic Processes. McGrawHill Inc, New York.
- Ramanujan R (1916) On Certain Arithmetical Functions. Transactions of the Cambridge Philosophical Society. XXII: 159-184.
- Rama Murthy G (2008) Innovative Arithmetical Functions: Infinite Products. International Journal of Algorithms, Computing and Mathematics.
- Rama Murthy G (2014) On a generalized approach to the Waring Problem: Algorithmic Number Theory " Manuscript in preparation
- Rama Murthy G (2009) Theory of Chances: Novel Stochastic Chains: Quantity Theory and Unification. IIIT Technical report.
- Rama Murthy G (2010) Stochastic Chains: Matrix Power Series Equations: Algebraic Geometry: Quantity Theory. International Congress of Mathematicians.
- Rama Murthy G (2013) Graphs: Associated Markov Chains. Proceedings of International Conference on Applied Mathematics and Theoretical Computer Science, Kanyakumari.
- Rama Murthy G (2014) Applications of Theory of Chances in Graph Theory. Manuscript in Preparation.
- Rama Murthy G (2011) Bounded Infinite Sequences/Functions: Orders of Infinity. International Journal of Algorithms, Computing and Mathematics.
- Rama Murthy G (1989) Transient and Equilibrium Analysis of Computer Networks: Finite Memory and Matrix Geometric Recursions. Ph.D thesis, Purdue University, West Lafayette.
- Rama Murthy G, Multi-Dimensional Neural Networks: Unified Theory, New Age International publishers, New Delhi
- Rama Murthy G (2008) Hybrid Neural Networks. Proceedings of International Conference on Power System Analysis, Control and Optimization (PSACO-2008).
- Rama Murthy G (2008) A Novel Class of Generative Neural Networks. Proceedings of 4th International Conference on Natural Computation, ICNC'08, IEEE.

18. Rama Murthy G (2014) Mathematical Theory of Chances, Quantity Theory and Unification. IIIT Technical Report in Preparation.
19. Rama Murthy G, Vasanth Iyer (2007) Distributed Wireless Sensor Network Architecture: Fuzzy Logic based Sensor Fusion. Proceedings of EUSL FAT 2007 (European Society for Fuzzy Logic and Technology), Ostrava, Czech Republic.
20. Royden HL (2011) Real Analysis Prentice Hall of India.