Master of Science Dissertation

Laws of Small Numbers in Riesz Spaces

Nigel Musara
Supervisors: Prof. Bruce Watson and Dr. Wen-Chi Kuo
School of Mathematics
University of the Witwatersrand
704109@students.wits.ac.za

September 27, 2021

Abstract

In [23], Wen-Chi Kuo, Jessica Vardy and Bruce A. Watson extended the concept of the binomial and poisson distributions to random processes in Riesz spaces. In the current project, we extend the work of Kuo, Vardy and Watson to random processes with small hitting probability and yields laws of small numbers.

Acknowledgments

Firstly, I would like to thank the Lord for giving me strength through all the adversity to be able to finish my Master of Science.

I would also like to thank my supervisors, Dr. Wen-Chi Kuo and Prof. Bruce A. Watson, for their patience, support and advice for which I will be eternally grateful.

Declaration

I, Nigel Musara, declare this dissertation titled ”Laws of Small Numbers in Riesz Spaces“ my own work and that any work done by other researchers has been referenced accordingly. This is a submission at the University of the Witwatersrand in Johannesburg for the degree of Master of Science. There has been no prior submission for any degree to any other university.

Nigel Musara

September, 2021

Contents

Introduction

Translations of random processes to Riesz spaces were initially focused mainly on martingale theory as shown in [18, 25, 29, 13, 40, 41]. In various settings, there has been extensive research on the underlying structures of conditional expectation operators as shown in [16, 24, 32, 38, 46].

Kuo, Vardy and Watson formulated Markov processes in Riesz spaces making use of the Ando-Douglas-Nikodym theorem, a Riesz space version of the Radon-Nikodym theorem proved by Watson in [46].

Bernoulli processes and Poisson processes in Riesz spaces are researched in [23]. We will extend the work of Kuo, Vardy and Watson on Bernoulli processes and Poisson processes in Riesz spaces. In this dissertation, we gain insight into the underlying structure of stochastic processes particularly focusing on the laws of small numbers in Riesz spaces.

In the first chapter, we look at the classical probability and measure theoretic concepts upon which the translation from a probability setting to Riesz spaces is based. We also introduce the theory of laws of small numbers and the bounds of the approximations.

The theory of Riesz Spaces relevant to our work is introduced in Chapter 2. Aliprantis [1] and Zaanen [47, 48] provide most of the foundational Riesz Space theory used in this section. We also consider the earlier translations of stochastic processes to Riesz Spaces in this chapter.

In Chapter 3, we look at the Bernoulli processes as proven by Kuo, Vardy and Watson in [23] in Riesz Spaces. This will then lead us into our own translations of the laws of small numbers into Riesz Spaces.

We then conclude with our findings and thoughts on further research of laws of small numbers in Riesz spaces in the Conclusion section.

1 Preliminaries

1.1 Classical probability theory and measure theory

In this chapter, we present definitions and basic properties necessary to understand and conduct our work on translating classical results on laws of small numbers to the Riesz space setting. Other definitions and properties are presented in subsequent chapters where they are more relevant.

We begin by providing a review of probability theory since we aim to generalize results from $\mathcal{L}^1$ probability spaces to the Riesz space setting.

Let Ω denote a nonempty space. As noted in [6][Page 20], a class $\mathcal{F}$ of subsets of Ω is called a σ- algebra if:

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$

(iii) $A, B \in \mathcal{F}$ implies $A \cup B \in \mathcal{F}$

(iv) $A_1, A_2, A_3... \in \mathcal{F}$ implies $A_1 \cup A_2 \cup A_3... \in \mathcal{F}$

We note the below from [6][Page 24].

A set function is a real-valued function defined on some class of subsets of Ω. A set function $\mathbf{P}$ on $\mathcal{F}$ is a probability measure if it satisfies the below conditions:

(i) for $A \in \mathcal{F}$, $0 \leq \mathbf{P}(A) \leq 1$,

(ii) For empty set, $\mathbf{P}() = 0$, $\mathbf{P}(\Omega) = 1$,

(iii) if $A_1, A_2, ...$ is a disjoint sequence of subsets of $\mathcal{F}$ and if $\cup_{k=1}^{\infty} A_k \in \mathcal{F}$, then

$$\mathbf{P}(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k).$$

Condition (iii) is referred to as countable additivity.

Let $\mathcal{F}$ denote a σ- algebra and $\mathbf{P}$ denote a probability measure. Then, $(\Omega, \mathcal{F}, \mathbf{P})$ denotes a *probability space*.

1.2 Conditioning

The elementary results presented in this subsection can be found in [8, 42].

We note the below taken from [8][Definition 1.12].

Suppose that we have a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ consisting of a space Ω, a σ-algebra $\mathcal{F}$ of subsets of Ω and a probability measure $\mathbf{P}$ on the σ- algebra $\mathcal{F}$.

(i) Two events $A, B \in \mathcal{F}$ are said to be *independent* if

$$P(A \cap B) = P(A)P(B)$$

(ii) In general, n events $A_1, A_2, ..., A_n \in \mathcal{F}$ are said to be *independent*

$$P(A_{i1} \cap A_{i2} \cap ... \cap A_{ij}) = P(A_{i1})P(A_{i2})...P(A_{ij})$$

for any indices $1 \leq i_1 \leq i_2 \leq ... \leq i_j \leq n, n \in \mathbb{N}$.

We note the below taken from [8][Definition 1.11].

For $A, B \in \mathcal{F}$, the definition of conditional probability is given as

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

We note the below definition from [42][Definition 1.5].

A map $f : \Omega \to \mathbb{R}$ such that for every Borel set $B \subset \mathbb{R}$,

$$f^{-1}(B) \in \mathcal{F}$$

is called a *random variable*.

A special case of random variables is that of an **indicator function** where

$$I_B(x) = \begin{cases} 1, & \text{if } x \in B \\ 0, & \text{if } x \notin B \end{cases} \tag{1}$$

We note the below taken from [8][Definition 1.9].

A random variable $f : \Omega \to \mathbb{R}$ is said to be *integrable* if

$$\int_\Omega f d\mathbf{P} < \infty.$$

$\mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$ is the family of integrable random variables. For each $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$, the *expectation* of f is defined as,

$$\mathbf{E}(f) = \int_\Omega f d\mathbf{P}.$$

We note the below taken from [8][Definition 2.1].

Given an event $A \in \mathcal{F}$ such that $P(A) \neq 0$, the definition of the *conditional expectation* of an integrable random variable f is

$$\mathbf{E}[f|A] = \frac{\int_A f d\mathbf{P}}{P(A)}.$$

The below result is from [8][Exercise 2.2].

If f is the indicator function, that is, $f = \chi_B$ for $B \in \mathcal{F}$, then

$$\mathbf{E}(\chi_B|A) = \int_A \chi_B d\mathbf{P}.$$

Since

$$\begin{aligned}\mathbf{E}(\chi_B|A) &= \mathbf{E}(\chi_B|B, A)P(B|A) + \mathbf{E}(\chi_B|B^c, A)P(B^c|A)\\ &= 1.P(B|A) + 0.P(B|A)\\ &= P(B|A),\end{aligned}$$

the above equation is the same as the probability of B given A.

We will now provide a general definition of conditional expectation as provided in [8][Definition 2.4].

Let f be an integrable random variable on probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and let Σ be a σ- algebra with $\Sigma \subset \mathcal{F}$. Then, $\mathbf{E}[f|\Sigma]$ is defined to be a random variable such that

(i) $\mathbf{E}[f|\Sigma]$ is Σ- measurable,

(ii) For any $A \in \Sigma$,
$$\int_A \mathbf{E}[f|\Sigma] d\mathbf{P} = \int_A f d\mathbf{P}.$$

We now present the general properties of conditional expectation as provided in [42][Theorem 4.5].

Theorem 1.1 (Properties of the Conditional Expectation)**.** *Let Σ be a sub-σ-algebra of $\mathcal{F}$ and assume $f, g \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$. The conditional expectation as a mapping $f \to g$ has the following properties:*

(i) If $g = \mathbf{E}[f|\Sigma]$, then $\mathbf{E}[g] = \mathbf{E}[f]$. We have that
$$\mathbf{E}[\mathbf{1}|\Sigma] = \mathbf{1} \qquad \textit{(Idempotent)}$$

(ii) If f is non-negative then
$$g = \mathbf{E}[f|\Sigma] \qquad \textit{(Non-negative)}$$
is non-negative.

(iii) The map is linear. If a_1 and a_2 are constants, then
$$\mathbf{E}[a_1 f_1 + a_2 f_2|\Sigma] = a_1\mathbf{E}[f_1|\Sigma] + a_2\mathbf{E}[f_2|\Sigma] \qquad \textit{(Linearity)}$$

(iv) If h is a bounded Σ- measurable function, then

$$\mathbf{E}[fh|\Sigma] = h\mathbf{E}[f|\Sigma] \qquad \textit{(Averaging)}$$

(v) If $\Sigma_2 \subset \Sigma_1 \subset \mathcal{F}$, then

$$\mathbf{E}[f|\Sigma_2] = \mathbf{E}[\mathbf{E}[f|\Sigma_1]|\Sigma_2].$$

We now introduce the concept of a convex function as provided in [8][Page 31].

Definition 1.2.
*A function $f : \mathbb{R} \to \mathbb{R}$ is called **convex** if for any $x, y \in \mathbb{R}$ and any $\lambda \in [0, 1]$*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

We note the below taken from [8][Theorem 1.2].

Theorem 1.3. *(Jensen's Inequality)*
Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function and let θ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $f(\theta)$ is also integrable. Then

$$f(\mathbf{E}[\theta|\mathcal{G}]) \leq \mathbf{E}[f(\theta)|\mathcal{G}]$$

for any $\Sigma-$ algebra $\mathcal{G}$ on Ω contained in $\mathcal{F}$.

As shown in [31], the additive and multiplicative properties of expectations lead us to the following inequality.

Theorem 1.4. *(Tchebichev Inequality) [31][Page 11]*
If X is a simple random variable, then for every $\epsilon > 0$,

$$P[|X| \geq \epsilon] \leq \frac{1}{\epsilon^2}\mathbf{E}[X^2].$$

Proof.

$$\begin{aligned}
\mathbf{E}[X^2] &= \mathbf{E}[X^2 I_{|X| \geq \epsilon}] + \mathbf{E}[X^2 I_{|X| \leq \epsilon}] \\
&\geq \mathbf{E}[X^2 I_{|X| \geq \epsilon}] \\
&\geq \epsilon^2 \mathbf{E}[I_{|X| > \epsilon}] \\
&= \epsilon^2 P[|X| \geq \epsilon].
\end{aligned}$$

□

As shown in [31] on [Page 12], for an independent simple random variable X, $\mathbf{E}[(X - \mathbf{E}[X])^2]$ is called the variance of X. It is denoted by $Var(X)$.

Theorem 1.5. *(Bienayme equality) [31][Page 12]*
Let X_i, $i = 1, 2, ..., n$ be independent simple random variables then

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i). \tag{2}$$

Proof.
By definition, for $X_i, i = 1, 2, , ,$ that are independent,

$$\begin{aligned} Var(\sum_{i=1}^{n} X_i) &= \mathbf{E}[(\sum_{i=1}^{n} X_i - \mathbf{E}[\sum_{i=1}^{n} X_i])^2] \\ &= \mathbf{E}[(\sum_{i=1}^{n} X_i)^2 - 2(\sum_{i=1}^{n} X_i)\mathbf{E}[\sum_{i=1}^{n} X_i] + \mathbf{E}[\sum_{i=1}^{n} X_i]^2] \\ &= \mathbf{E}[(\sum_{i=1}^{n} X_i)^2] - 2\mathbf{E}[\sum_{i=1}^{n} X_i]\mathbf{E}[\sum_{i=1}^{n} X_i] + \mathbf{E}[\sum_{i=1}^{n} X_i]^2 \\ &= \mathbf{E}[(\sum_{i=1}^{n} X_i)^2] - \mathbf{E}[\sum_{i=1}^{n} X_i]^2 \end{aligned}$$

Now,

$$\begin{aligned} \mathbf{E}[(\sum_{i=1}^{n} X_i)^2] &= \mathbf{E}[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{E}[X_i X_j] \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbf{E}[\sum_{i=1}^{n} X_i]^2 &= (\sum_{i=1}^{n} \mathbf{E}[X_i])^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{E}[X_i]\mathbf{E}[X_j] \end{aligned}$$

Thus,

$$\begin{aligned} Var(\sum_{i=1}^{n} X_i) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{E}[X_i X_j] - \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{E}[X_i]\mathbf{E}[X_j] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}(\mathbf{E}[X_i X_j] - \mathbf{E}[X_i]\mathbf{E}[X_j]) \\ &= \sum_{i=1}^{n} Var(X_i), \end{aligned}$$

since X_i are independent. □

Theorem 1.6. *(Bernoulli Law of Large Numbers) [31][Page 11]*
Suppose that we have independent events $A_i, i = 1, 2, ...,$ whose probabilities, $P[A_i]$ are the same value p. Let $S_n = \sum_{i=1}^{n} I_{A_i}$. Then, for every $\epsilon > 0$,

$$P[|\frac{S_n}{n} - p| \geq \epsilon] \to 0$$

as $n \to \infty$.

Proof.
Applying the Tchebichev Inequality yields,

$$\begin{aligned} P[|\frac{S_n}{n} - p| \geq \epsilon] &= P[|S_n - \mathbf{E}[S_n]| \geq n\epsilon] \\ &\leq \frac{1}{\epsilon^2 n^2} Var(S_n). \end{aligned}$$

Now, by Bienayme equality, $Var(S_n) = Var(\sum_{i=1}^{n} I_{A_i}) = \sum_{i=1}^{n} Var(I_{A_i}) = p - p^2$.
Hence, it follows that as $n \to \infty$,

$$\begin{aligned} P[|\frac{S_n}{n} - p| \geq \epsilon] &\leq \frac{1}{\epsilon^2 n^2} Var(S_n) \\ &= \frac{p(1-p)}{\epsilon^2 n} \\ &\to 0 \end{aligned}$$

□

We present the notion of a signed measure and the related definitions and theorems as provided in [6][Page 446].

A *signed measure* on a measurable space $(\Omega, \mathcal{F})$ is a countably additive set function $\lambda(\cdot)$ defined for $A \in \mathcal{F}$.

A set $A \in \mathcal{F}$ is *totally positive* for a set function, λ, if for every $B \in \mathcal{F}$ with $B \subset A$, we have $\lambda(B) \geq 0$. Two measures, μ_1 and μ_2, are said to be *mutually singular* if there exist sets A and B, such that $A \cup B = \Omega$ and $A \cap B = \emptyset$ with $\mu_1(A) = 0$ and $\mu_2(B) = 0$. In this case, μ_1 is said to be concentrated on B and μ_2 is said to be concentrated on A.

Theorem 1.7. *(Hahn-Jordan decomposition theorem)[42][Theorem 4.3]*
Given a countably additive signed measure λ on $(\Omega, \mathcal{F})$, it can be written as $\lambda = \mu^+ - \mu^-$, the difference of two non-negative measures. Moreover, μ^+ and μ^- may be chosen to be mutually singular, that is, there are disjoint sets $\Omega_+, \Omega_- \in \mathcal{F}$ such that $\mu^+(\Omega_-) = \mu^-(\Omega_+) = 0$. In fact, Ω_+ and Ω_- can be taken to be subsets of Ω that are respectively totally positive and totally negative for λ. $\mu_\pm$ then become just the restrictions of λ to $\Omega_\pm$.

We note the below definition from [42][Definition 4.2].

A signed measure λ is said to be *absolutely continuous* with respect to a non-negative measure μ , $\lambda \ll \mu$ in symbols, if whenever $\mu(A)$ is zero for a set $A \in \mathcal{F}$, it is also true that $\lambda(A)$ is zero. When $\lambda \ll \mu$, μ is said to be a dominating measure for λ.

For establishing the existence of a conditional expectation of a random variable on $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$, the following theorem is needed.

Theorem 1.8. *(Radon-Nikodym theorem) [42][Theorem 4.4]*
If λ is absolutely continuous with respect to a non-negative measure μ, then there is $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$ such that

$$\lambda(A) = \int_A f d\mu$$

for all $A \in \mathcal{F}$. The function f is uniquely determined almost everywhere and is called the Radon-Nikodym derivative of λ with respect to μ. The function is denoted by

$$f = \frac{d\lambda}{d\mu}.$$

1.3 Laws of Small Numbers

The following review is a summary of the work done in one of the sections in "Laws of small numbers: Extremes and rare events" (see [17]). This section contains the foundational theory for the translation of laws of small numbers to Riesz Spaces. Only definitions and results relevant to this research will be given.

In [45], Ladislaus von Bortkiewicz showed the significance of the Poisson approximation of binomial distributions with practical examples. He made use of the number of cavalrymen killed by horse kicks as one of the various examples. Suppose we have an event wherein an individual is killed by a horse kick within one year. Assume R is a binary random variable that models this event, that is, $R \in [0,1]$ with $R = 1$ representing the occurrence of an accidental death and $R = 0$ representing the non-occurrence of an accidental death. Then, assuming the independence of individual lifetimes, the total number $K(n)$ of victims in a regiment of size n is binomial distributed. So, with $P(\cdot)$ denoting probability and $p = P(R = 1)$ being the mortality rate, we have

$$\begin{aligned} P(K(n) = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= B(n, p, k), \end{aligned}$$

for $k = 0, 1, ..., n$ with $B(n, p, k)$ denoting the binomial distribution with size n, probability of occurence p and in this example total number of victims k.

If the mortality rate, p, is small, then as $n \to \infty$ the occurrence of accidental death is a rare event and the binomial distribution can be approximated by the Poisson distribution

$$Po(\lambda, k) = e^{-\lambda} \frac{\lambda^k}{k!}, \tag{3}$$

where $k = 0, 1, ...$ with $\lambda = np$.

The approximation of the binomial distribution by the Poisson distribution is called the law of small numbers, written as $B(n, p) \sim P_{np}$.

Let $\mathcal{L}(K(n))$ denote the distribution of $K(n)$.

If $K(n)$ is a binomial $B(n, p)$-distributed random variable and $\tau(n)$ is a Poisson P_{np} -distributed random variable, then with $\sim_D$ representing the approximate equality of the distributions $\mathcal{L}(K(n)) = B(n, p)$ and $\mathcal{L}(\tau(n)) = P_{np}$ of $K(n)$ and $\tau(n)$, respectively, the above relation can be rewritten as $K(n) \sim_D \tau(n)$.

Let S be a sample space and $\mathcal{B}$ denote a σ- algebra . If $X_1, X_2, ...,$ is a sequence of independent copies of a random element X with values in S equipped with $\mathcal{B}$ and also assuming that the random variable $X_1, X_2, ...,$ are also independent of $K(n)$ and $\tau(n)$, then the above approximation suggests the following extension $(X_1.....X_{K(n)}) \sim_D (X_1.....X_{\tau(n)})$. This approximation is called the functional law of small numbers.

Since $(X_1.....X_{K(n)})$ and $(X_1.....X_{\tau(n)})$ are random vectors which might have different random length, point process are used to overcome this dimensionality problem. For $\mathcal{B}$ denoting a σ- algebra, any point $x \in S$ is identified with the related indicator function, that is,

$$\epsilon_x(B) = \begin{cases} 1, & \text{if } x \in B, B \in \mathcal{B} \\ 0, & \text{if } x \notin B \end{cases} \tag{4}$$

So, the random element X is identified with the random measure ϵ_X. The random vectors $(X_1.....X_{K(n)})$ and $(X_1.....X_{\tau(n)})$ turn into random finite measures: $N_n(B) = \sum_{i=1}^{K(n)} \epsilon_{X_i}(B)$ and $N_n^*(B) = \sum_{i=1}^{\tau(n)} \epsilon_{X_i}(B)$, $B \in \mathcal{B}$.

The set of finite point measures on $(S, \mathcal{B})$ is denoted by $\mathbb{M}$. If $\mathbb{M}$ is equipped with the smallest σ -algebra $\mathcal{M}$ such that for any $B \in \mathcal{B}$ the projections $\mu \to \mu(B)$, $\mu \in \mathbb{M}$ are measurable, then $N_n(\cdot)$ and $N_n^*(\cdot)$ are measurable random elements in $(\mathbb{M}, \mathcal{M})$. They are called point processes. N_n is called a binomial process and N_n^* is called a Poisson process. This is because the counting variable $K(n)$ is a binomial random variable and $\tau(n)$ is a Poisson random variable.

1.4 Bounds for the functional laws of small number

We quote the below from [17][Page 7].

The error of the approximations in the preceding section is measured with respect to:

(i) Hellinger distance

(ii) Variational distance

Definition 1.9. *[17][Page 7]*
*The **Hellinger distance** between distributions of random elements X and Y with values in some measurable space $(S, \mathcal{B})$ is defined by*

$$H(X, Y) = (\int (f^{\frac{1}{2}} - g^{\frac{1}{2}})^2 d\mu)^{\frac{1}{2}}, \tag{5}$$

where f, g are the densities of X and Y with respect to some dominating measure μ.

Definition 1.10. *[17][Page 7]*
*The **Variational distance** between the distributions of random elements X and Y is defined by*

$$d(X, Y) = sup_{B \in \mathcal{B}} |P(X \in B) - P(Y \in B)| \tag{6}$$

Definition 1.11. *[36][Page 41]*
*The **Kullback - Leibler distance** between distributions of random elements X and Y is defined as*

$$K(X,Y) = \int (-log\frac{f}{g})dX, \tag{7}$$

where f, g are the densities of X and Y with respect to some dominating measure μ.

Remark 1.12.
The Hellinger distance is bounded by the Kullback-Leibler distance. For more on this you may see [36][Page 41].

We note the below taken from [17][Page 8].

1.5 The Stein-Chen method

The Stein-Chen method of measuring the laws of small numbers in variational distance offers a great method of providing an upper bound. We detail the method below as provided in pages 12-13 of [17].

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and $X_1, \ldots, X_n$ be independent $0-1$ valued random variables with $\mathbf{E}[X_i = 1] = p_i$, i.e. $X_i = \chi_{A_i}$ for some $A_i \in \mathcal{A}$. Let $\lambda = \sum\limits_{i=1}^{n} p_i$.

We denote $Po(\cdot; \lambda)$ the Poisson distribution with mean λ. We recall that a random variable X is said to have Poisson distribution with parameter $\lambda > 0$ if

$$\mu(\{\omega \in \Omega | X(\omega) = k\}) = \frac{\lambda^k e^{-\lambda}}{k!} =: Po(k; \lambda).$$

Here, $\omega \in \Omega$, where Ω is the nonempty space, $\lambda = \mathbf{E}[X] = \int_\Omega X \, d\mu$ and further $\lambda = Var(X) := \mathbf{E}[X^2] - \mathbf{E}[X]^2$.

Set $W = \sum_{i=1}^n X_i$. The Stein-Chen method supplies an upper bound on how close the distribution of W is to $Po(\cdot; \lambda)$.

For $A \subset \mathbb{N}_0$, let $g(j, \lambda, A)$ for $\lambda > 0$ and $j \in \mathbb{N}_0$ defined recursively as, $g(0, \lambda, A) = 0$ and

$$\lambda g(j+1, \lambda, A) = jg(j, \lambda, A) + \chi_A(j) - Po(A, \lambda), \tag{8}$$

where

$$Po(A, \lambda) = \sum_{k \in A} Po(k, \lambda).$$

As W is an $\mathbb{N}_0$ valued function, $g(\cdot, \lambda, A)$ can be applied to W to give

$$\lambda g(W(\omega)+1, \lambda, A) = W(\omega)g(W(\omega), \lambda, A) + \chi_A(W(\omega)) - Po(A, \lambda).$$

If we set $W_j = \sum_{i \neq j} X_i$ then $W = W_i + X_i$ and, as X_i takes on only the values 0 and 1, so if $X_i(\omega) = 0$ then

$$X_i(\omega) g(W(\omega), \lambda, A) = 0 = X_i(\omega) g(W_i(\omega) + 1, \lambda, A)$$

while if $X_i(\omega) = 1$ then

$$X_i(\omega) g(W(\omega), \lambda, A) = g(W_i(\omega) + 1, \lambda, A) = X_i(\omega) g(W_i(\omega) + 1, \lambda, A).$$

Hence,

$$X_i g(W, \lambda, A) = X_i g(W_i + 1, \lambda, A).$$

Taking expectations we thus obtain

$$\mathbf{E}[X_i g(W, \lambda, A)] = \mathbf{E}[X_i g(W_i + 1, \lambda, A)]$$

which combined with the expectation of (8) gives

$$\begin{aligned} \lambda \mathbf{E}[g(W+1, \lambda, A)] &= \sum_{i=1}^{n} X_i g(W, \lambda, A) + \mu\{W \in A\} - Po(A, \lambda) \\ &= \sum_{i=1}^{n} \mathbf{E}[X_i g(W_i + 1, \lambda, A)] + \mu\{W \in A\} - Po(A, \lambda). \end{aligned}$$

Since the sequence $X_1, \ldots, X_n$ is independent, it follows that X_i and W_i are independent and hence X_i and $g(W_i + 1, \lambda, A)$ are independent giving that

$$\mathbf{E}[X_i g(W_i + 1, \lambda, A)] = \mathbf{E}[X_i] \cdot \mathbf{E}[g(W_i + 1, \lambda, A)] = p_i \mathbf{E}[g(W_i + 1, \lambda, A)].$$

Hence,

$$\sum_{i=1}^{n} p_i \mathbf{E}[g(W+1, \lambda, A)] = \sum_{i=1}^{n} p_i \mathbf{E}[g(W_i + 1, \lambda, A)] + \mu\{W \in A\} - Po(A, \lambda),$$

which can be re-arranged to give

$$\begin{aligned} \mu\{W \in A\} - Po(A, \lambda) &= \sum_{i=1}^{n} p_i \mathbf{E}[g(W+1, \lambda, A) - g(W_i + 1, \lambda, A)] \\ &= \sum_{i=1}^{n} p_i \mathbf{E}[g(W_i + X_i + 1, \lambda, A) - g(W_i + 1, \lambda, A)] \\ &= \sum_{i=1}^{n} p_i \mathbf{E}[X_i (g(W_i + 2, \lambda, A) - g(W_i + 1, \lambda, A))]. \end{aligned}$$

Again using the independence of X_i and W_i we have

$$\mu\{W \in A\} - Po(A, \lambda) = \sum_{i=1}^{n} p_i^2 \mathbf{E}[g(W_i + 2, \lambda, A) - g(W_i + 1, \lambda, A)].$$

Thus,

$$|\mu\{W \in A\} - Po(A,\lambda)| \;\; \leq \;\; \sum_{i=1}^{n} p_i^2 \alpha(\lambda),$$

where

$$\alpha(\lambda) = \sup_{A \in \mathcal{A}, j \in \mathbb{N}_0} |g(j+1,\lambda,A) - g(j,\lambda,A)|.$$

It can be shown (see [5]) that

$$\alpha(\lambda) \leq E(\lambda) := \frac{1-e^{-\lambda}}{\lambda}.$$

Now

$$\sum_{i=1}^{n} \frac{p_i^2}{\lambda} \leq \sum_{i=1}^{n} \frac{p_i}{\lambda} \max_{j=1,\ldots,n} p_j = \max_{j=1,\ldots,n} p_j$$

giving

$$|\mu\{W \in A\} - Po(A,\lambda)| \;\; \leq \;\; (1-e^{-\lambda}) \max_{j=1,\ldots,n} p_j.$$

So for rare events (that is, the p_i's are small and $\lambda < 1$) the error between the actual probability and the Poisson distribution is bounded by $\lambda \max_{j=1,\ldots,n} p_j$.

2 Riesz Spaces

Riesz spaces were first introduced by Frigyes Riesz in his 1928 paper "Sur la decomposition des operations fonctionelles lineaires". There is general agreement that even though mathematicians were scattered across multiple countries (mainly Japan, Russia and the United States) and made use of different terminology, the introduction of Riesz Spaces happened around 1928-1935. Around 1935, L. V. Kantorovitch began an extensive investigation of the algebraic and convergence properties of Riesz spaces with applications to linear operator theory. From this period, other mathematicians started making important contributions to the general theory of Riesz spaces. In 1936, a 'spectral theorem' for Riesz spaces was proved by H. Freudenthal. A.G. Pinsker, A.I. Judin, B.Z. Vulikh and several other mathematicians soon joined Kantorovitch in his research of Riesz spaces. In Japan between 1940 and 1944, H. Nakano, T. Ogasawara and K. Yosida, contributed further to the theory of Riesz spaces. Around the same time period, S. Kakutani and H.F. Bohnenblust in the United States also provided concrete representations of abstract L-spaces and M-spaces.

In this section, definitions and results which form the building blocks of Riesz space theory will be presented. The results presented here and any details of omitted proofs can be found in [49, 1, 33, 47, 48].

2.1 Ordered vector spaces

We provide the following definition taken from [47][Definition 1.1]

Definition 2.1.
A relation $\leq$ in X is called a partial ordering if;

(i) $x \leq x$ *for every* $x \in X$ *(reflexive)*

(ii) $x \leq y$ *and* $y \leq z$ *implies* $x \leq z$ *for every* $x, y, z \in X$ *(transitive)*

(iii) $x \leq y$ *and* $y \leq x$ *implies* $x = y$ *(anti-symmetric).*

Definition 2.2.
The non-empty set X is referred to as a partially ordered set if it is equipped with a partial ordering.

We quote the below definition from [47][Definition 4.1].

Definition 2.3. *(Linear Partial Ordering)*
A real vector space E (with elements x, y) is called an ordered vector space if it is partially ordered and:

(i) $x \leq y$ *implies* $x + z \leq y + z$ *for all* $z \in E$,

(ii) $x \geq 0$ *implies* $ax \geq 0$ *for every* $a \geq 0 \in \mathbb{R}$.

The following definitions of infimums and supremums and the adopted notation are from the explanations in [47] on pages 2 and 4.

Suppose X is a partially ordered set and Y is a non-empty subset of X. If we have $y \leq x$ for $x \in X$ and all $y \in Y$, we say is x an upper bound of Y. The subset Y is said to be bounded above.

The supremum of a subset, (x, y) of a partially ordered set X is denoted as $sup(x, y) = x \vee y$. An element, s, is said to be the supremum of the pair $x, y \in X$ if:

(i) the element, s, is an upper bound of the set $\{x, y\}$, that is , $x \leq s$ and $y \leq s$,

(ii) the element, s, is the least upper bound, that is, if $x, y \leq t$ then $s \leq t$.

Suppose X is a partially ordered set and Y is a non-empty subset of X. If we have $y \geq x$ for $x \in X$ and all $y \in Y$, we say is x a lower bound of Y. The subset Y is said to be bounded below. The infimum of a subset, (x, y) of a partially ordered set X is denoted as $inf(x, y) = x \wedge y$. An element, s, is said to be the infimum of the pair $x, y \in X$ if:

(i) the element, s, is a lower bound of the set $\{x, y\}$, that is, $s \leq x$ and $s \leq y$,

(ii) the element, s, is the greatest lower bound, that is, if $t \leq x, y$ then $s \leq t$.

We say the point $x_0 \in X$ is a maximal element of X if there does not exist an element that it is strictly larger than x_0. That is, $x_0 \leq x$ implies that $x_0 = x$ for $x \in X$. If there exists an element $x_0 \in X$ such that $x_0 \geq x$ for all $x \in X$, then x_0 is referred to as the largest element of X. The minimal element and the smallest element are defined in an analogous manner.

We now introduce the definition of a lattice as quoted from [47].

Definition 2.4. *[47][Definition 1.3(iv)]*
A partially ordered set X is referred to as a lattice if every subset of X consisting of two points has a supremum and an infimum.

We now quote the properties of a lattice as provided in [47][Page 4-5].

Every finite subset of a lattice X has a supremum and infimum. If the elements in the finite subset of a lattice X are $x_1, x_2, ..., x_n$, then the supremum and infimum of the set are denoted by $\vee_{i=1}^{n} x_i$ and $\wedge_{i=1}^{n} x_i$.

Definition 2.5. *[47][Definition 2.1]*
The lattice X is called distributive if for all $x, y, z \in X$,

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

We quote the below theorem which shows that the infimum operation is distributive over the supremum operation.

Theorem 2.6. *[47][Theorem 2.2]*
The lattice X is called distributive if and only if for all $x, y, z \in X$,

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

In Chapter 2 of [47][Page 13], the below definitions are introduced.

Definition 2.7.
If for every pair of $x, y \in E$, an ordered vector space, there exists $x \vee y$ and $x \wedge y$, then E is a Riesz space.

The collection of positive elements of E is defined by

$$E^+ = \{x : x \geq 0\}.$$

This collection of all $x \in E$ satisfying $x \geq 0$ is called the *positive cone* of E.

Theorem 2.8. *(Properties of the positive cone) [47][Page 16]*
Let E be an ordered vector space and suppose that E^+ is its positive cone. Then, for $x, y \in E^+$ and $a \in \mathbb{R}, a \geq 0$,

(i) $x, y \in E^+ \Rightarrow x + y \in E^+$,

(ii) $x \geq 0 \Rightarrow ax \geq 0$,

(iii) $x, -x \in E^+ \Rightarrow x = 0$,

Proof.

(i) Since $x, y \in E^+$, $x \geq 0, y \geq 0$. Hence, $x + y \geq x \geq 0$ so by transitivity of a partial order, we get
$$x + y \geq 0$$
and so by definition of a positive cone, $x + y \in E^+$.

(ii) Since $x \geq 0$ and $a \geq 0$, it follows from the positive homogeneity of a partial order that $ax \geq 0$.

(iii) Since $E^+ = \{x : x \geq 0\}$ and $x, -x \in E^+$,
$$\begin{aligned} -x \in E^+ &\Rightarrow -x \geq 0 \\ &\Rightarrow x \leq 0 \; and \; x \geq 0. \end{aligned}$$
Therefore, $x = 0$ as a partial ordering is anti-symmetric.

□

The below theorem is taken from [47][Page 16].

Theorem 2.9.
Let E be a Riesz space and suppose that E^+ is its positive cone. Then, for $x, y, z \in E$ and $a \in \mathbb{R}$,

(i) $x \geq y \Leftrightarrow x - y \in E^+$,

(ii) $x \geq y \Leftrightarrow x = x \vee y$ *and* $y = x \wedge y$,

(iii) $x \geq y \Leftrightarrow \begin{cases} ax \geq ay, & \textit{for } a > 0, \\ ax \leq ay, & \textit{for } a < 0, \end{cases}$

(iv) $(x+z) \vee (y+z) = (x \vee y) + z$ *. Likewise,* $(x+z) \wedge (y+z) = (x \wedge y) + z$,

Proof.

(i) Suppose $x \geq y$. Then since $x, y \in E^+$, it follows that $x \geq 0$ and $y \geq 0$. By subtracting y on both sides of the inequality $x \geq y$, we get $x - y \geq 0$. Now, suppose $x - y \in E^+$, then by definition of a positive cone,

$$x - y \geq 0.$$

By adding y on both sides of the inequality $x - y \geq 0$, we get $x \geq y$.

(ii) Since E is a Riesz Space, it has the lattice properties and partial ordering. Hence, $x \geq y$ implies that x is an upper bound of $y \in E$ and y is a lower bound of x in E. Thus, $x = x \vee y$ and $y = x \wedge y$.

(iii) Suppose $a > 0, a \in \mathbb{R}$. Then, by Linear Partial Ordering (ii), we get

$$x \geq y \Rightarrow ax \geq ay.$$

For $a < 0, a \in \mathbb{R}$, we have $-a > 0$. Let $z = -x - y$, then by Linear Partial Ordering (i), we have

$$\begin{aligned} x \geq y &\Rightarrow x + (-x - y) \geq y + (-x - y) \\ &\Rightarrow -y \geq -x. \end{aligned}$$

Since $(-a) > 0$, multiplying by $(-a)$ on both sides, we get

$$(-a)(-y) \geq (-a)(-x).$$

Hence, we have $ay \geq ax$.

(iv) Since $x \vee y \geq x$ and $x \vee y \geq y$, we have $(x \vee y) + z \geq x + z$ and $(x \vee y) + z \geq y + z$, since $x, y, z \in E^+$. Thus, $(x \vee y) + z \geq (x+z) \vee (y+z)$. Now, $(x+z) \vee (y+z) \geq y + z$ and $(x+z) \vee (y+z) \geq x + z$.
So, $(x+z) \vee (y+z) - z \geq x \vee y \Rightarrow (x+z) \vee (y+z) \geq (x \vee y) + z$.

Further, $(x \vee y) + z \geq x + z$ and $(x \vee y) + z \geq y + z$ so $(x \vee y) + z \geq (x+z) \vee (y+z)$.

$\square$

Definition 2.10. *[47][Page 17]*
Let E be Riesz space. Then, for any $x \in E$

(i) $x^+ = x \vee 0$,

(ii) $x^- = -(x \wedge 0) = (-x) \vee 0$,

(iii) $(x \vee -x) = |x|$.

Theorem 2.11. *[47][Theorem 5.1]*
Let E be a Riesz space with $x \in E$ such that $x \vee 0$ exists. Then,

(i) $x^+, x^- \in E^+$,

(ii) $x = x^+ - x^-$ *and* $x^+ \wedge x^- = 0$,

(iii) $|x| = x^+ + x^-$,

(iv) if $x, y \in E$ *and* x^+, y^+ *exist in* E^+, *then* $x \leq y \Leftrightarrow x^+ \leq y^+$ *and* $x^- \geq y^-$.

Proof.

(i) By definition,
$$x^+ = x \vee 0 \geq 0.$$
Thus, $x^+ \in E^+$.
By definition,
$$x^- = (-x) \vee 0 \geq 0.$$
Thus, $x^- \in E^+$.

(ii) By definition of x^+, x^- in Definition 2.10 (i) and Definition 2.10 (ii),
$$x^+ - x = (x \vee 0) - x = (-x) \vee 0 = x^-$$
and
$$x^+ \wedge x^- = (x + x^-) \wedge x^- = (x \wedge 0) + x^- = -x^- + x^- = 0.$$

(iii) By Definition 2.10 (iii), we have,
$$|x| = x \vee (-x) = (2x \vee 0) - x = 2(x \vee 0) - x = 2x^+ - (x^+ - x^-) = x^+ + x^-.$$

(iv) Suppose $x \leq y$. Since by part (ii) $x = x^+ - x^-$, we have
$$x \vee 0 \leq y \vee 0 \Rightarrow x^+ \leq y^+$$
and
$$(-x) \vee 0 \geq (-y) \vee 0 \Rightarrow x^- \geq y^-.$$

Conversely, suppose that $x^+ \leq y^+$ and $x^- \geq y^-$. Then,

$$y^+ + x^- \geq x^+ + y^-$$
$$\Rightarrow x^+ - x^- \leq y^+ - y^-.$$

Thus, by part (ii),

$$x \leq y.$$

□

Theorem 2.12 (Properties of the supremum and infimum)**.** *[47][Theorem 5.2] Let $x, y \in E$, a Riesz space. Then,*

(i) $x \vee y = (y - x)^+ + x = (x - y)^+ + y$*,*

(ii) $x \wedge y = y - (y - x)^+ = x - (x - y)^+$*,*

(iii) $x \vee y + x \wedge y = x + y$*,*

(iv) $x \vee y - x \wedge y = |x - y|$*,*

Proof.

(i)

$$\begin{aligned} x \vee y &= ((y - x) \vee 0) + x \\ &= (y - x)^+ + x. \end{aligned}$$

(ii)

$$x \vee y = ((x - y) \wedge 0) + y = -((y - x) \vee 0) + y = y - (y - x)^+.$$

(iii) Applying part (i) and (ii) of this theorem,

$$\begin{aligned} (x \vee y) + (x \wedge y) &= (x - y)^+ + y + x - (x - y)^+ \\ &= x + y. \end{aligned}$$

(iv)

$$\begin{aligned} (x \vee y) - (x \wedge y) &= (x - y)^+ + y - y + (y - x)^+ \\ &= (x - y)^+ + (y - x)^+ \\ &= |x - y|. \end{aligned}$$

□

The standard decomposition of $x \in E$, a Riesz space, is given by $x = x^+ - x^-$. This is the minimal decomposition of x as the difference of two non-negative vectors, that is, $x^+ \wedge x^- = 0$.

2.2 Riesz Space Inequalities and Distributive Laws

We denote E as a Riesz space. We will now present the inequalities and distributive laws of Riesz spaces.

We quote from [33][Theorem 12.1].

Theorem 2.13.
Suppose E is a Riesz Space and let $x, y \in E$. Then,

(i) $(x+y)^+ \leq x^+ + y^+$,

(ii) $(x+y)^- \leq x^- + y^-$,

(iii) $||x| - |y|| \leq |x+y| \leq |x| + |y|$.

Proof.

(i) By definition, $x + y \leq x^+ + y^+$ and $0 \leq x^+ + y^+$. Hence,

$$\begin{aligned} (x+y) \vee 0 &\leq x^+ + y^+ \\ &\Rightarrow (x+y)^+ \leq x^+ + y^+. \end{aligned}$$

(ii) By definition,

$$\begin{aligned} (x+y)^- &= (-(x+y)) \vee 0 \\ &= (-x-y)^+ \\ &\leq (-x)^+ + (-y)^+ \\ &= x^- + y^-. \end{aligned}$$

(iii) We will first show that $|x+y| \leq |x| + |y|$. Since by (i) $(x+y)^+ \leq x^+ + y^+$, breaking this down yields $x + y \leq x^+ + y^+$ and $0 \leq x^+ + y^+$. Also by (ii),

$$(x+y)^- \leq x^- + y^-.$$

Hence, $|x+y| \leq |x| + |y|$.
To show that $||x| - |y|| \leq |x+y|$, we substitute $x - y$ for x in the inequality $|x+y| \leq |x| + |y|$. Hence we get $|x| - |y| \leq |x-y|$ and $|y| = |(y-x) + x| \leq |y-x| + |x|$. So, $|y| - |x| \leq |y-x|$. Substituting $(-y)$ for y in $||x| - |y|| \leq |x-y|$ yields $||x| - |y|| \leq |x+y|$.

□

Theorem 2.14. *[33][Theorem 12.2]*
Suppose that D is a subset of a Riesz space E such that $x_0 = \vee_{x \in D} x$ exists in E, then for all $y \in E$,

$$x_0 \wedge y = \vee_{x \in D}(x \wedge y).$$

Proof.

Suppose that there exists $x_0 = \vee_{x\in D} x$ and let $y \in E$. Since $x_o \geq x \; \forall \; x \in D$, it follows that

$$x_0 \wedge y \geq x \wedge y \, \forall \, x \in D.$$

Therefore, $x_0 \wedge y$ is an upper bound of the set of elements $\{x \wedge y : x \in D\}$.

Let k be the upper bound of this set, then

$$k \geq x \wedge y = x + y - (x \vee y).$$

So,

$$k - y + (x \vee y) \geq x$$

holds $\forall x \in D$. It follows that

$$k - y + (x_0 \vee y) \geq x$$

holds $\forall x \in D$. Hence,

$$k - y + (x_0 \vee y) \geq \vee_{x\in D} x = x_0;$$

that is

$$k \geq y - (x_0 \vee y) + x_0,$$

showing that any upper bound of the set $\{x \wedge y : x \in D\}$ is greater or equal to $x_0 \wedge y$. □

Corollary 2.15. *[33][Corollary 12.3]*
Let E be a Riesz space. Suppose that x, y, z are elements of E. Then the following holds: $(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z)$ and $(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z)$.

So, a Riesz space is a distributive lattice with respect to its partial ordering.

For elements in a Riesz space, there is a decomposition property. We quote from [47][Theorem 6.4]

Theorem 2.16. *(Riesz decomposition property)*
Let E be a Riesz space. Suppose that $u, z_1, z_2 \in E^+$ satisfies $u \leq z_1 + z_2$. Then there exist u_1, u_2 in E^+ such that $u_1 \leq z_1$, $u_2 \leq z_2$ and $u = u_1 + u_2$.

Proof.

Let $u_1 = u \wedge z_1$ and $u_2 = u - u_1$. Hence, $0 \leq u_1 \leq z_1$. Since $u_1 \leq u$, it follows that $u - u_1 \in E^+$. From above $u = u_1 + u_2$, hence

$$\begin{aligned} u_2 &= u - u_1 \\ &= u - (u \wedge z_1) \\ &= (u - z_1)^+ \; by \; Theorem \; 2.12 \; (ii) \\ &\leq z_2 \end{aligned}$$

$\square$

Theorem 2.17. *[47][Theorem 6.5]*
Let $x, y \in E$ and let u, v, w be elements of the positive cone E^+. Then ,

(i) $(u + v) \wedge w \leq (u \wedge w) + (v \wedge w)$,

(ii) $(x + y) \vee w \leq (x \vee w) + (y \vee w)$,

(iii) $v \wedge w = 0$ *implies* $(u + v) \wedge w = u \wedge w$,

(iv) *if* $u \wedge v = 0$*, then* $(u + v) \wedge w = (u \wedge w) + (v \wedge w)$.

2.3 Order convergence and Uniform convergence

We will now generalize the results from sequences to directed sets and provide some fundamental results relating to the convergence of elements of a Riesz space.

We quote the below definition from [47].

Definition 2.18. *[47][Page 46]*
Let E be a Riesz space. A sequence $x_1 \leq x_2 \leq x_3...$ in E is said to be increasing and is denoted as $x_n \uparrow, n \in \mathbb{N}$, if $x = \vee_{n\in\mathbb{N}} x_n$ exists, then we write $x_n \uparrow x$. Similarly, a sequence $x_1 \geq x_2 \geq x_3...$ in E is said to be decreasing and is denoted as $x_n \downarrow, n \in \mathbb{N}$, if $x = \wedge_{n\in\mathbb{N}} x_n$ exists, then this is denoted as $x_n \downarrow x$. If $x_n \uparrow x$ or $x_n \downarrow x$, then x_n is said to converge monotonely.

Theorem 2.19. *(Properties of monotone convergence) [33][Theorem 15.2]*
Let x_n, y_n, $n = 1, 2, ...$ denote a sequence in a Riesz Space E.

(i) If $x_n \uparrow x$, then $x_{n_k} \uparrow x$ for every subsequence $(x_{n_k} : k = 1, 2, ...)$ such that $n_1 < n_2 < ...$,

(ii) For all $a \geq 0 \in \mathbb{R}$, if $x_n \uparrow x$, then $ax_n \uparrow ax$,

(iii) If $x_n \uparrow x$ and $y_n \uparrow y$ then $x_n + y_n \uparrow x + y$,

(iv) If $x_n \uparrow x, y_n \uparrow$ and $x_n + y_n \uparrow x + y$, then $y_n \uparrow y$.

We now define order convergence.

Definition 2.20.
A sequence x_n, $n = 1, 2, ...$ in E is said to converge in order to x if there exists a sequence $y_n \downarrow 0$ such that $|x_n - x| \leq y_n$ holds for all n. Note $x_n \to x$ denotes order convergence.

We quote the below from [47][Page 49].

Let E be a Riesz space. Suppose $0 < e \in E$. The sequence x_n, $n = 1, 2, ...$ in E is said to converge e-uniformly to x if for any number $\epsilon > 0$ there exists an index $n(\epsilon)$ such that $|x - x_n| \leq \epsilon e$ for all $n \geq n(\epsilon)$.

We note the below general properties of convergence as provided in [47][Theorem 10.2].

Theorem 2.21. *(General properties of convergence) Let E be a Riesz space and suppose we have $x_n, y_n, n = 1, 2, ...$ in E.*

(i) If $x_n \to x, y_n \to y$ and $a, b \in \mathbb{R}$, then $ax_n + by_n \to ax + by$. Also, $x_n \vee y_n \to x \vee y$ and $x_n \wedge y_n \to x \wedge y$.

(ii) If $x_n \to x$ and $x_n \geq y_n$ for all n then $x \geq y$.

(iii) If $x_n \to x$, then $(x_{n_k}) \to x$ where (x_{n_k}) is a subsequence of (x_n).

Monotone sequences are a special case of a directed set which is defined below as taken from [47][Definition 10.7].

Definition 2.22.
Let Λ be a non-empty set and denote the elements of Λ by α. Suppose E is a Riesz space and also assume that for each element $\alpha \in \Lambda$ there is a mapping $\alpha \mapsto x_\alpha$ which maps from Λ to E. Λ is said to be the index set for the family $(x_\alpha)_{\alpha\in\Lambda}$. The family $(x_\alpha)_{\alpha\in\Lambda}$ is said to be upwards directed (denoted $x_\alpha \uparrow_{\alpha\in\Lambda}$) if for any pair $\alpha_1, \alpha_2 \in \Lambda$ there exists $\alpha_3 \in \Lambda$ such that $x_{\alpha_3} \geq x_{\alpha_1} \vee x_{\alpha_2}$. If $x_\alpha \uparrow_{\alpha\in\Lambda}$ and $x = \vee_{\alpha\in\Lambda}(x_\alpha)$ we write $x_\alpha \uparrow_{\alpha\in\Lambda} x$. In this case, $(x_\alpha)_{\alpha\in\Lambda}$ is said to be upwards directed with supremum x. The family $(x_\alpha)_{\alpha\in\Lambda}$ is said to be downwards directed (denoted $x_\alpha \downarrow_{\alpha\in\Lambda}$) if for any pair $\alpha_1, \alpha_2 \in \Lambda$, there exists $\alpha_3 \in \Lambda$ such that $x_{\alpha_3} \leq x_{\alpha_1} \wedge x_{\alpha_2}$. If $x_\alpha \downarrow_{\alpha\in\Lambda}$ and $x = \wedge_{\alpha\in\Lambda}(x_\alpha)$ we write $x_\alpha \downarrow_{\alpha\in\Lambda} x$. In this case, $(x_\alpha)_{\alpha\in\Lambda}$ is said to be downwards directed with infimum x.

The results for directed sets are comparable to the results of sequences as shown in the lemma below.

Lemma 2.23. *[47][Lemma 2.16]*
Let Λ be a non-empty set.

(i) If $x_\alpha \uparrow_{\alpha\in\Lambda} x$, then $(x_\alpha : x_\alpha \geq x_{\alpha_0}) \uparrow_{\alpha\in\Lambda} x$ for any α_0 .

(ii) If $x_\alpha \uparrow_{\alpha\in\Lambda} x$ and $0 \leq a \in \mathbb{R}$, then $ax_\alpha \uparrow_{\alpha\in\Lambda} ax$.

Upward and downward directedness can also be defined for subsets D of Riesz space E that are not indexed as shown below.

Definition 2.24. *[47][Definition 10.9]*
The non-empty subset D of the Riesz space E is said to be upwards directed, if for any two elements x and y in D, there exists an element $z \in D$ such that $z \geq x \vee y$. A downwards directed set is similarly defined.

A subset, say A, of a Riesz space is said to be order bounded from above if there is a vector u such that $a \leq u$ for each $a \in A$. Subsets that are order bounded from below are defined similarly. If a subset of a Riesz space is order bounded from above and also order bounded from below, then the subset is called order bounded. If the elements x, y of the partially ordered set X satisfy $x \leq y$, then the non-empty set $[x, y] = \{z : z \in X, x \leq z \leq y\}$ is called an order interval in X. So, a set is order bounded if and only if it is contained in an order interval.

2.4 Riesz subspaces, ideals and bands

The below definitions are taken from [47][Definition 7.1] and [33][Definition 17.5].

Subsets of a Riesz space, E, are assumed to inherit the ordering from E.

(i) A subset V of E is called a Riesz subspace of E if;

  (a) V is a linear subspace of E,

  (b) $x, y \in V$, then $x \vee y, x \wedge y \in V$.

(ii) A subset S of E is said to be solid if it follows from $x \in S$ and $|y| \leq |x|$ that $y \in S$.

(iii) A subset S of E is called an ideal if;

  (a) S is a linear subspace of E,

  (b) if $x \in S$ and $|y| \leq |x|$, then $y \in S$.

  that is, S is a solid linear subspace of E. If S is an ideal generated by a singleton, then S is called a principal ideal.

(iv) An ideal B in E is called a band if whenever a subset of B has a supremum in E, then this supremum is already in B. The band generated by a singleton is called a principal band.

Theorem 2.25. *[47][Theorem 7.2]*
Let E be a Riesz space.

(i) Every band in E is an ideal and every ideal in E is a Riesz subspace. The trivial spaces-$\{0\}$ and the space E itself are bands,

(ii) Every Riesz subspace (respectively, ideal, band) of a Riesz subspace (respectively, ideal, band) of E is itself a Riesz subspace (respectively, ideal, band) of E.

Theorem 2.26. *[47][Theorem 7.4]*
Let E be a Riesz space.

(i) Any intersection of Riesz subspaces of E (respectively, ideal, band) is again a Riesz subspace (respectively, ideal, band).

(ii) Suppose that B is an ideal in E. If for each $D \subset B$ with D containing finite suprema, we have that $supD \in B$ if $supD$ exists, then B is a band in E.

For any non-empty subset D of E, the Riesz space generated by D is defined as the intersection of all Riesz subspaces of E containing D. The ideal and band generated by D can be similarly defined. The ideal generated by D is denoted by

A_D. If D is a singleton, then $A_D = A_x, x \in D$ and A_x is called a principal ideal. A band generated by a singleton is called a principal band.

2.5 Disjointness, algebraic and direct sums and disjoint complements

We now give the fundamental results of *disjointness* that are relevant to this dissertation.

We quote the definitions and following text from [47][Page 34].

Definition 2.27. *[47][Page 34]*
The elements x and y in a Riesz space, E, are said to be disjoint (denoted $\perp$) if $|x| \wedge |y| = 0$.

Definition 2.28. *[47][Page 34]*
If D is a non-empty subset of E, then the set

$$D^d = \{x \in E : x \perp y \; for \; all \; y \in D\} \tag{9}$$

is called the disjoint complement of D.

A disjoint complement of D^d is called the second disjoint complement of D and is denoted $D^{dd} = (D^d)^d$.

If D_1 and D_2 are non-empty subsets in E such that for every $d_1 \in D_1$ and $d_2 \in D_2$, we have $d_1 \perp d_2$ then D_1 and D_2 are said to be disjoint, denoted $D_1 \perp D_2$.

If A and B are subsets of the vector space V, the *algebraic sum* of A and B is defined by

$$A + B = \{x_1 + x_2 : x_1 \in A, x_2 \in B\}. \tag{10}$$

If A and B are linear subspaces of V, then $A + B$ is also a linear subspace. If $A \cap B = \{0\}$, then the representation $x = x_1 + x_2$ ($x_1 \in A, x_2 \in B$) is unique. In this case, $A + B$ is now called the *direct sum* of A and B denoted by $A \bigoplus B$.

Theorem 2.29. *[47][Theorem 8.1]*
If E is a Riesz space, then for $x, y, z \in E$:

(i) if $x \perp y$ and $|z| \leq |x|$, then $z \perp y$,

(ii) if $x \perp y$ and $a \in \mathbb{R}$, then $ax \perp y$,

(iii) if $x_1 \perp y$ and $x_2 \perp y$, then $(x_1 + x_2) \perp y$,

(iv) if D is a subset of E such that $x_0 = supD$ exists and $y \perp x$ holds for all $x \in D$, then $y \perp x_0$,

Proof.

(i) Suppose $x \perp y$ and $|z| \leq |x|$. Then,

$$0 \leq |z| \wedge |y| \leq |x| \wedge |y| = 0.$$

Hence, $|z| \wedge |y| = 0$

(ii) Suppose $x \perp y$ and $a \in \mathbb{R}$. If $|a| \leq 1$, then $|ax| \leq |x|$, Thus, by (i), $ax \perp y$. Now, suppose $|a| > 1$, then since

$$|ax| \wedge |ay| = |a|(|x| \wedge |y|) = 0,$$

it follows that $ax \perp ay$. Thus, $ax \perp y$ since $|y| \leq |ay|$.

(iii) Suppose $x_1 \perp y$ and $x_2 \perp y$. From Theorem 2.17, we have

$$0 \leq |x_1 + x_2| \wedge |y| \leq (|x_1| + |x_2|) \wedge |y| \leq |x_1| \wedge |y| + |x_2| \wedge |y| = 0.$$

Thus, $(x_1 + x_2) \perp y$.

(iv) Suppose D is a non-empty subset of E such that $x_0 = supD$ exists. Then, since $x_0 = sup(x : x \in D)$, we have

$$x_0 \vee z = sup(x \vee z : x \in D)$$

for any $z \in E$, in particular for $z = 0$. Hence,

$$x_0^+ = sup(x^+ : x \in D).$$

We also have

$$x_0 \wedge z = sup(x \wedge z : x \in D)$$

for any $z \in E$. Thus, for $z = 0$, we have $-x_0^- = inf(-x^- : x \in D)$ so $x_0^- = inf(x^- : x \in D)$.

Suppose now that $y \perp D$, then $y \perp x^+$ and $y \perp x^-$ for every $x \in D$. So,

$$x_0^+ \wedge |y| = sup(x^+ \wedge |y| : x \in D) = sup0 = 0$$

$$x_0^- \wedge |y| \leq x^- \wedge |y| = 0$$

for every $x \in D$. Thus, $x_0^- \wedge |y| = 0$. Hence, $x_0^+ \perp y$ and $x_0^- \perp y$. It follows that $x_0 \perp y$.

□

We quote the below from [33][Theorem 17.6].

Theorem 2.30. *[33] [Theorem 17.6]*
Let A and B be ideals in a Riesz space E. Then,

(i) $A + B$ is an ideal in E,

(ii) $A \perp B$ *(that is, for every* $x \in A$ *and every* $y \in B$*,* $x \perp y$*)* $\Leftrightarrow A \cap B = \{0\}$*,*

(iii) *if* $x, y \in A \bigoplus B$ *and* $x = x_1 + x_2$*,* $y = y_1 + y_2$ *(*$x_1, y_1 \in A, x_2, y_2 \in B$*), then* $x \leq y \Rightarrow x_1 \leq y_1$ *and* $x_2 \leq y_2$*.*

Proof.

(i) Since A and B are ideals in E, $A + B$ is a linear subspace of E. Suppose $x \in A + B, y \in E$ such that $|y| \leq |x|$. For $x_1 \in A, x_2 \in B$, we have $x = x_1 + x_2$. Hence,

$$y^+ \leq |y| \leq |x| \leq |x_1| + |x_2|.$$

It follows from Theorem 2.16 that there exists $u_1, u_2 \in E^+$ such that $y^+ = u_1 + u_2$, $0 \leq u_1 \leq |x_1|$, $0 \leq u_2 \leq |x_2|$. Since A and B are ideals, consequently $u_1 \in A, u_2 \in B$ so $y^+ \in A + B$. Similarly, it follows that $y^- \in A + B$. Hence, $A + B$ is an ideal.

(ii) Suppose $A \perp B$ and let $x \in A \cap B$. Then,

$$\begin{aligned} |x| \wedge |x| &= 0 \\ &\Rightarrow |x| = 0 \\ &\Rightarrow x = 0. \end{aligned}$$

Hence, $A \cap B = \{0\}$.

Now, suppose $A \cap B = \{0\}$ and $x_1 \in A, x_2 \in B$. Then,

$$|x_1| \wedge |x_2| \leq |x_1| \Rightarrow |x_1| \wedge |x_2| \in A$$

and

$$|x_1| \wedge |x_2| \leq |x_2| \Rightarrow |x_1| \wedge |x_2| \in B.$$

Hence, $|x_1| \wedge |x_2| = 0$ since the above shows that $|x_1| \wedge |x_2| \in A \cap B$.

(iii) Suppose that $u \geq 0$ with the decomposition $u = u_1 + u_2, u_1 \in A, u_2 \in B$. Separating u_1 and u_2 into their positive and negative parts gives $u = (u_1^+ + u_2^-) - (u_1^- + u_2^-)$.
Now,

$$u_1 \in A \Rightarrow |u_1| \in A.$$

Hence, $u_1^+ \in A, u_1^- \in A$.

□

The below remark and example were taken from [12][Page 13].

Remark 2.31.
For ideals $A, B \in E$, a Riesz Space, $A+B$ is also an ideal. However, for bands $D, F \in E$, a Riesz Space, $F+D$ is not necessarily a band as illustrated by the following example.

Example 2.32.
Consider $C([0,1])$, the space of continuous functions on $[0,1]$. We define $x \leq y$ and suppose $f(x) \leq f(y)$ holds for $x, y \in [0,1]$, $f \in C([0,1]$. Then, $C([0,1])$ is a lattice and an ordered vector space. Hence, $C([0,1])$ is a Riesz space.

Let A and B be bands in $C([0,1])$ defined by,

$$A = \{f : f(x) = 0, 0 \leq x \leq \frac{1}{2}\},$$

$$B = \{f : f(x) = 0, \frac{1}{2} \leq x \leq 1\},$$

Then, $A + B = \{f : f(\frac{1}{2}) = 0\}$ which is not a band.

Theorem 2.33. *[33][Theorem 19.2]*
Let D be a non-empty set of E. Then,

(i) D^d *is a band in* E,

(ii) $D \subset D^{dd}$ *and* $D^d = D^{ddd}$,

(iii) $D^d \cap D^{dd} = \{0\}$ *so* $D^d + D^{dd} = D^d \bigoplus D^{dd}$.

2.6 Projections

We quote the below theorem from [47][Theorem 11.1].

Theorem 2.34. *[47][Theorem 11.1]*
Let E be a Riesz space. Let A and B be disjoint ideals in E. If $A \bigoplus B = E$, then $B = A^d$ and $A = B^d$. Hence, $A = A^{dd}$, $B = B^{dd}$ and A and B are bands.

Proof.
Suppose A, B are disjoint ideals in E with $A \bigoplus B = E$, then $B \subseteq A^d$. Now, suppose $0 \leq u \in A^d$, then since A is disjoint from B, $u = u_1 + u_2$ with $0 \leq u_1 \in A$ and $0 \leq u_2 \in B$. Since $0 \leq u_1 \leq u \in A^d$ holds now, it follows that $u_1 \in A^d$. Also, $u \in A$. Hence, $u_1 = 0$. Thus, $u = u_2 \in B \Rightarrow A^d \subseteq B$. Therefore, $A^d = B$. □

We paraphrase the below from [47][Page 56-57].

A mapping between vector spaces, $T : V \to W$ is called a *linear operator* if for all $a, b \in \mathbb{R}$ and $x, y \in V$,

$$T(ax + by) = aT(x) + bT(y)$$

We define $T_1T_2x = T_1(T_2x)$ for any two operators, T_1, T_2. If $T^2 = TT = T$ then the operator $T : V \to V$ called idempotent.

If B is a band in a Riesz space E such that $B \bigoplus B^d = E$, then B is called a *projection band*. P_B is called the *band projection* on the projection band B.

Theorem 2.35. *[47][Theorem 11.4]*
Suppose that E is a Riesz space. The band $B \in E$ is a projection band if for any $u \in E^+$, we have

$$u_1 = \vee_{v \in B; 0 \leq v \leq u} v$$

exists. We call u_1 the component of u in B. if

$$u_2 = \vee_{w \in B^d; 0 \leq w \leq u} v$$

exists then u_2 is the component of u in B^d and we have that $u = u_1 + u_2$.

Theorem 2.36. *[47][Theorem 11.4]*
Let E be a Riesz space. For any $x \in E^+$, we denote the component of x in the projection band B by P_Bx. P_B is a mapping from E into itself and has the following properties;

(i) P_B is a projection,

(ii) for all $u \in E^+$, $0 \leq P_Bu \leq u$

P_B is called the band projection onto the projection band B.

To give more properties of band projections, we require the concept of *Archimedean* space. These spaces provide uniqueness of order limits in Riesz spaces.

We quote the below immediate definition and theorem from [47][Page 39].

Definition 2.37. *[47][Page 39]*
A Riesz space E is said to be Archimedean if

$$inf(n^{-1}u : n = 1, 2, ...) = 0$$

holds for every $u \in E^+$.

Theorem 2.38. *[47][Theorem 32.1]*
If B_1, B_2 are projection bands in the Archimedean Riesz space E then $B_3 = B_1 \cap B_2$ is a projection band and the corresponding band projections satisfy $P_1P_2 = P_3 = P_2P_1$.

This theorem leads to the following theorem.

Theorem 2.39. *[47][Theorem 32.2]*
If E is an Archimedean Riesz space with projection bands B_1, B_2 and the corresponding band projections P_1, P_2, then the following are equivalent:

(i) $B_1 \subset B_2$.

(ii) $P_1P_2 = P_1 = P_2P_1$.

(iii) $P_1 \leq P_2$.

2.7 Dedekind Completeness

We provide the below definitions and theorems taken from [47].

Definition 2.40. *[47][Theorem 12.1]*
A Riesz space E is called Dedekind complete if every non-empty subset of E^+ that is bounded above has a supremum.

Definition 2.41. *[47][Theorem 12.1]*
A Riesz space E is called Dedekind σ-complete if every non-empty countable subset of E^+ that is bounded above has a supremum.

Theorem 2.42. *[47][Theorem 12.3]*
Let E be a Riesz space. If E is Dedekind complete, then E is Dedekind σ-complete which implies E is Archimedean.

Theorem 2.43. *[33][Theorem 23.8]*
Let E be Dedekind complete. Then,

(i) if A and B are bands such that $A \cap B = \{0\}$, then $A \bigoplus B$ is also a band.

(ii) for any band B, $B \bigoplus B^d = E$.

Definition 2.44. *[47][Definition 23.5]*
Let E be a Riesz space and $e \in E$, $e \geq 0$. Then e is called a weak order unit of E if the band in E generated by e is E.

As shown in [33], every band is a principal band in a Dedekind complete space with weak order units. Since A_e is an ideal and $a \in \mathbb{R}$, every element $y \in E$ such that $|y| \leq |ae|$ is an element of A_e. Conversely, the set of all y satisfying the aforementioned inequality is an ideal in E that contains A_e. Thus, A_e can be explicitly written as follows,

$$A_e = \{y \in E||y| \leq |ae|, a \in \mathbb{R}\}.$$

This formula can be generalized in the case where D is a finite subset of E with elements $e_1, e_2, ..., e_n, n \in \mathbb{N}$ to,

$$A_D = \{y \in E||y| \leq \sum_{i=1}^{n} |a_i e_i|, a_i \in \mathbb{R}\}.$$

3 Linear properties, Conditional expectations and Independence in Riesz Spaces

3.1 Linear Operators

We now review linear operators and their properties in Riesz spaces. The results presented in this chapter and any details of omitted proofs can be found in [49, 33, 47, 48, 46, 24, 43]. Linear properties are very important in the translation of random processes to Riesz spaces.

From Chapter 2, we showed that a mapping between ordered vector spaces, $T : V \to W$ is called a *linear operator* if for all $a, b \in \mathbb{R}$ and $x, y \in V$,

$$T(ax + by) = aT(x) + bT(y).$$

Suppose that T is an operator. The space of all operators from V into W is denoted $\mathbb{L}(\mathrm{V}, \mathrm{W})$ and it is a vector space.

We note the below definition and subsequent paragraphs from [47][Page 122].

Definition 3.1.
Let V, W be ordered vector spaces and $T \in \mathbb{L}(\mathrm{V}, \mathrm{W})$.

(i) T is a positive operator, denoted $T \geq 0$, if T maps the positive cone of V into the positive cone of W.

(ii) T is called a regular operator if $T = T_1 - T_2$ for some positive operators T_1, T_2. The set of all regular operators between V and W is denoted by $\mathbb{L}_{\mathrm{r}}(\mathrm{V}, \mathrm{W})$.

(iii) The order interval $[g, f]$ is a subset of V of the form $\{h \in V | g \leq h \leq f\}$. T is called an order bounded operator if T maps order intervals of V into order intervals of W. The set of all order bounded operators from V into W is denoted by $\mathbb{L}_{\mathrm{b}}(\mathrm{V}, \mathrm{W})$.

Theorem 3.2. *[47][Theorem 18.3]*
Let V, W be ordered vector spaces and operator $T \in \mathbb{L}(\mathrm{V}, \mathrm{W})$. If T is a regular operator then T is order bounded.

Definition 3.3. *[47][Definition 19.1]*
The linear operator T between the Riesz spaces E and F, is said to be a Riesz homomorphism if for all $f, g \in E$,

$$T(f \vee g) = Tf \vee Tg.$$

Thus, every Riesz homomorphism is a positive operator.

We now show that this is a characterization of a Riesz homomorphism and hence every band projection is a Riesz homomorphism.
Consider the band projection P onto projection band B in Riesz space E. Since $0 \leq Pf \leq f$ for every $f \geq 0$ in E, it follows that $f \wedge g = 0$ implies that $Pf \wedge Pg = 0$.

We note the below definition from [47][Definition 21.1].

Definition 3.4.
Let E and F be Riesz spaces and $T : E \to F$. Then T is said to be an order continuous operator if and only if for any directed set $D \subset E$ with $D \downarrow 0$ we have that $\wedge_{f \in D} |Tf| = 0$.

Theorem 3.5. *[47][Theorem 21.2]*
Let E and F be Riesz spaces and consider operators T and S mapping from E to F.

(i) If T is order continuous, then for all scalars a, aT is order continuous.

(ii) If $T \geq 0$, then T is order continuous if and only if $D \downarrow 0$ in E implies $Tf \downarrow_{f \in D} 0$ in F.

(iii) If $T \geq 0$ is order continuous and $0 \leq S \leq T$, then S is order continuous.

3.2 Conditional expectations

We now have the tools that enable us to define the concept of a *conditional expectation* in a Riesz space. It is shown in [35] that defining stochastic processes through probability measure is equivalent to working through conditional expectations.

We now provide the definition and properties of a conditional expectation in the Riesz space setting as quoted and referenced below.

Theorem 3.6. *[24][Theorem 2.2]*
Let E be a Riesz space with weak order unit and T be a positive order continuous projection on E. There is a weak order unit e of E with $Te = e$ if and only if Tw is a weak order unit of E for each weak order unit w in E.

We note the below definition of a conditional expectation as provided in [24].

Definition 3.7. *[25][Definition 2.3]*
Let E be a Riesz space with weak order unit e. A positive order continuous projection T on E with range, $\mathcal{R}(T)$, a Dedekind complete Riesz subspace of E, is called a conditional expectation if Te is a weak order unit of E for each e in E.

In order to mirror the classical setting, we require that $\mathcal{R}(T)$ is Dedekind complete.

We say that a Riesz space, E, is *universally complete* if E is Dedekind complete and every subset of E which consists of mutually disjoint elements has a supremum in E. A Riesz space that is universally complete and contains E as an order dense Riesz subspace is called the *universal completion* of E. The universal completion is denoted by E^u.

We quote the below definition from [23][Page 3].

Let E be a Dedekind complete Riesz space with weak order unit and T a strictly positive conditional expectation on E. Then, the space E is said to be T-universally complete if for each increasing net (f_α) in E^+ with (Tf_α) order bounded in the universal completion E^u, we have that (f_α) is order convergent.

3.2.1 Commutation properties

Lemma 3.8. *[24][Lemma 3.1]*
Let E be a Riesz space with a weak order unit and T be a conditional expectation on E. Let B be the band in E generated by $0 \le g \in \mathcal{R}(T)$ and P be the band projection onto B. Then,

(i) $Tf \in B$ *for each* $f \in B$*,*

(ii) $Pf, (I-P)f \in \mathcal{R}(T)$ *for each* $f \in \mathcal{R}(T)$*, where I denotes the identity map,*

(iii) $Tf \in B^d$ *for each* $f \in B^d$*.*

Theorem 3.9. *[24][Theorem 3.2]*
Let E be a Dedekind complete Riesz space with weak order unit, T a conditional expectation on E and B be the band in E generated by $0 \le g \in \mathcal{R}(T)$, with associated band projection P. Then, $TP = PT$.

Proof.
Suppose $x \in B$ and $y \in B^d$ then Lemma 3.8 gives

$$\begin{aligned} TP(x+y) &= TPx \\ &= Tx \in B \end{aligned}$$

since $x \in B$. So, $PTx = Tx$.
Further by Lemma 3.8(iii), $Ty \in B^d$ so $PTy = 0$. Thus,

$$Tx = PTx = PTx + PTy = PT(x+y).$$

Hence, $TP = PT$ since $E = B \bigoplus B^d$. □

The above theorem shows that each conditional expectation on a Riesz space commutes with band projections associated with bands generated by elements from the range of the conditional expectation and underlies the averaging properties of conditional expectations.

We quote the below definition of an f-algebra from [24], [page 5].

Let E be a Dedekind complete Riesz space with weak order unit e. Then, E is said to be an f-algebra if the below properties hold:

(i) $xy \geq 0$ for all $x, y \in E^+$,

(ii) Multiplication is associative, that is, for all $x, y, z \in E$,

$$(xy)z = x(yz).$$

(iii) Multiplication is distributive, that is, for $a, b \in \mathbb{R}$ and $x, y \in E$,

$$z(ax + by) = azx + bzy$$

(iv) e is the multiplicative unit, that is, for $x \in E$,

$$ex = x.$$

(v) For $x, y \in E$, $x \wedge y \to (xz) \wedge y = (zx) \wedge y = 0$ for all $z \in E$.

Theorem 3.10. *[24] [Theorem 4.2]*
Let E be a Dedekind complete Riesz space with a weak order unit e and T be a conditional expectation on E with $Te = e$. For each $g \in \mathcal{R}(T)_+$, there exists a sequence (s_n) such that $s_n \uparrow g$ in order. Here, each s_n is of the form $\sum_{i=1}^{k} a_i P_i e$, where $a_i \in \mathbb{R}_+$ and P_i is a band projection which commutes with T.

Proof.
Suppose that B_j^n is the band generated by $(f - j2^{-n}e)^+$ and let $\tilde{P}_j^n$ denote the associated band projection for $j = 1, ..., n2^n$. Let $\tilde{P}_0^n = I$ and $\tilde{P}_{n2^n+1}^n = 0$. Denote $P_j^n = \tilde{P}_{j-1}^n - \tilde{P}_j^n$ and $\alpha_j^n = (j-1)2^{-n}$ for $j = 1, ..., 1 + n2^n$. Then,

$$s_n = \sum_{j=1}^{1+n2^n} \alpha_j^n P_j^n e \uparrow nf,$$

where this convergence is with respect to order. Since $(f - j2^{-n}e)^+ \in \mathcal{R}(T)$, Theorem 3.9 applies, giving that P_j^n commutes with T.

$\square$

Corollary 3.11. *[24][Theorem 4.3]*
Let E be a Dedekind complete Riesz space with a weak order unit e and T be a conditional expectation on E with $Te = e$. If E is an f- algebra and e is the multiplicative unit of the f- algebra, then $T(gf) = gTf$ for each $g \in \mathcal{R}(T)$ and $f \in E$.

Proof.
Suppose $f \in E^+$ and $g \in E^+$. By Theorem 3.10, there exist real numbers α_j^n and band projections P_j^n which commute with T such that $g_n \uparrow g$ e-uniformly, where

$$g_n = \sum_{j=1}^{k_n} \alpha_j^n P_j^n e \in \mathcal{R}(T)^+.$$

Multiplying the above by f gives

$$\begin{aligned} g_n f &= (\sum_{j=1}^{k_n} \alpha_j^n P_j^n e) f \\ &= \sum_{j=1}^{k_n} \alpha_j^n P_j^n f \end{aligned}$$

since e is the multiplicative unit and multiplication commutes with band projections because E is an f-algebra. Further, applying T to the above gives us

$$\begin{aligned} T(g_n f) &= T(\sum_{j=1}^{k_n} \alpha_j^n P_j^n f) \\ &= \sum_{j=1}^{k_n} \alpha_j^n T P_j^n f \\ &= \sum_{j=1}^{k_n} \alpha_j^n P_j^n T f \\ &= g_n(Tf). \end{aligned}$$

Also for $|f| \leq Me$, M a constant

$$\begin{aligned} |pf - qf| &= |p - q||f| \\ &\leq M|p - q| \end{aligned}$$

for $p, q \in E$. Thus, multiplication is e-uniformly continuous. It follows that $g_n f \uparrow gf$ and $g_n(Tf) \uparrow g(Tf)$ which along with order continuity of T gives us $T(g_n f) \uparrow T(gf)$. Combining these results completes the proof. □

3.3 Independence in Riesz spaces

Just as the probability measure and the multiplicative properties of the positive real numbers are essential in the study of independence in probability theory, the conditional expectation operator and composition are essential to the study of independence in Riesz spaces.

In [46], B.A Watson proves the Radon-Nikodym theorem in Riesz Spaces which like the Radon-Nikodym theorem in probability and measure theory, has enabled the proving of many results regarding independence in Riesz Spaces. Watson also proved the Hahn Decomposition theorem in Riesz Spaces in the same paper.

Let $\mathcal{B}(F)$ be the class of all band projections on a Dedekind complete Riesz Space E with $Pe \in F$, where P denotes a band projection and e, a weak order unit. F denotes a Dedekind complete Riesz subspace of a Riesz space E which contains $\mathcal{R}(T)$.

Let $f \in E$. Then, $J \in \mathcal{B}(F)$ is said to be positive with respect to (T, f) if $TPf \geq 0$ for all $P \in \mathcal{B}(F)$ such that $P \leq J$. T in this is a strictly positive conditional expectation operator acting on the Dedekind complete Riesz Space E with weak order unit $e = Te$. Similarly, $J \in \mathcal{B}(F)$ is said to be negative with respect (T, f) if $TPf \leq 0$ for all $P \in \mathcal{B}(F)$ such that $P \leq J$. J is said to be strictly positive if J is positive with respect to (T, f) and $TJf \neq 0$.

We will now state the Hahn-Jordan decomposition theorem in Riesz Spaces as proven by Watson in [46]. This theorem provides a decomposition of $\mathcal{B}(F) \to E$ with $P \to TPf$.

Theorem 3.12. *Hahn-Jordan Decomposition[46][Theorem 3.5]*
Let E be a Dedekind complete Riesz space with strictly positive conditional expectation operator, T, and weak order unit, $e = Te$. Let F be a Dedekind complete Riesz subspace of E with $\mathcal{R}(T) \subset F$ and let $f \in E$. There exists a band projection $Q \in \mathcal{B}(F)$ which is positive with respect to (T, f) and has $I - Q$ negative with respect to (T, f).

We note the below theorem from [46]. Please note that by F being an order closed Riesz subspace of a Riesz space E we mean that F is a Riesz subspace of E and if (x_α) is a net in F convergent to f in E then $f \in F$ and (x_α) converges to f in F.

Theorem 3.13. *(Radon-Nikodym)[46][Theorem 4.1]*
Let E be a T-universally complete Riesz space with weak order unit, $e = Te$, where T is a strictly positive conditional expectation operator on E. Let F be a closed Riesz subspace of E with $\mathcal{R}(T) \subset F$, there exists a unique $g \in F^+$ such that

$$TPf = TPg \tag{11}$$

for all $P \in \mathcal{R}(T)$.

Lemma 3.14. *[46][Lemma 5.6]*
Let T be a strictly positive conditional expectation operator on the T-universally complete Riesz space, E, with weak order unit e = Te. Let F be a closed Riesz subspace of E with $\mathcal{B}(F) \subset F$. The map T_F is additive on E^+ and has

$$TPTF(f) = TPf \tag{12}$$

for all $P \in \mathcal{B}(F), f \in E^+$.

Also note that the results can be extended to E as each $f \in E$ can be decomposed as $f = f^+ - f^-$, $f^+, f^- \in E$.

Corollary 3.15. *(Douglas-Ando)[46][Corollary 5.9]*
Let T be a strictly positive conditional expectation operator on the T-universally complete Riesz space, E, with weak order unit, $e = Te$. The subset F of E is a closed Riesz subspace of E with $\mathcal{R}(T) \subset F$ if and only if there is a conditional expectation T_F on E with $\mathcal{R}(T_F) = F$ and $TT_F = T = T_F T$. In this case $T_F f$ for $f \in E^+$ is uniquely determined by the property that

$$TPf = TPT_F f \tag{13}$$

for all band projections on E with $Pe \in F$.

We note the below definition of independence with respect to a conditional expectation operator T taken from [26] [Definition 4.1].

Definition 3.16.
Let E be a Dedekind complete Riesz space with conditional expectation T and weak order unit $e = Te$. Let P and Q be band projections on E, we say that P and Q are independent with respect to T and e if

$$TPTQe = TPQe = TQTPe$$

P and Q are said to be T-conditionally independent if they are defined as above.

Suppose our Riesz Space is the set of integrable functions, that is, $E = L^1(\omega, \mathcal{A}, \mu)$, where μ is a probability measure and T is the expectation operator

$$Tf = \int_\omega f d\mu = E[F]1. \tag{14}$$

We have that the band projections on E are maps of the form $pf = f\chi_A$ and $Qf = f\chi_B$, where $A, B \in \mathcal{A}$.

The above shows that, with the aforementioned conditions, the independence of P and Q corresponds to the classical independence of A and B.

As shown by Vardy and Watson in [43], the weak order unit $e = Te$ selected does not change the above definition.

Theorem 3.17. *[43][Theorem 3.2]*
Suppose E is a Dedekind complete Riesz space with conditional expectation, T, and let e be a weak order unit which is invariant under T. The band projections P and Q in E are T-conditionally independent if and only if $TPTQw = TPQw = TQTPw$ for all $w \in \mathcal{R}(T)$.

Corollary 3.18. *[43][Corollary 3.3]*
Let E be a Dedekind complete Riesz space with conditional expectation, T, and let e be a weak order unit which is invariant under T. Let P_i ,$i = 1, 2$, be band projections on E. The band projections, P_i ,$i = 1, 2$, are T-conditionally independent if and only if the closed Riesz subspaces $E_i = (P_i e, \mathcal{R}(T))$, $i = 1, 2$, are T-conditionally independent.

The above corollary shows that T-conditional independence of the closed Riesz subspaces $(hPe, R(T)i)$ and $(hQe, R(T)i)$ generated by Pe and $\mathcal{R}(T)$ and by Qe and $\mathcal{R}(T)$ respectively equates T-conditional independence of the band projections P and Q.

Theorem 3.19. *[43][Theorem 3.4]*
Let E_1 and E_2 be two closed Riesz subspaces of the T-universally complete Riesz space E with strictly positive conditional expectation operator T and weak order unit $e = Te$. Let S be a conditional expectation on E with $ST = T$. If $\mathcal{R}(T) \subset E_1 \cap E_2$ and $T_{<\mathcal{R}(S),E_i>}$ is the conditional expectation having as its range the closed Riesz subspace of E generated by $\mathcal{R}(S)$ and E_i , then the spaces E_1 and E_2 are S-conditionally independent, if and only if

$$T_i T_{<\mathcal{R}(S),E_{3-i}>} i = T_i S T_{<\mathcal{R}(S),E_{3-i}>}$$

$i = 1, 2$, where T_i is the conditional expectation commuting with T and having range E_i.

Corollary 3.20. *[43][Corollary 3.5]*
Let E_1 and E_2 be two closed Riesz subspaces of the T-universally complete Riesz space E with strictly positive conditional expectation operator T and weak order unit $e = Te$. If $\mathcal{R}(T) \subset E_1 \cap E_2$, then the spaces E_1 and E_2 are T- conditionally independent, if and only if

$$T_1 T_2 = T = T_2 T_1, \tag{15}$$

where T_i is the conditional expectation commuting with T and having range E_i.

Definition 3.21. *[43][Definition 3.8]*
Let E be a Dedekind complete Riesz space with conditional expectation T and

weak order unit $e = Te$. Let $(E_\lambda)_{\lambda\in\Lambda}$ be a family of closed Dedekind complete Riesz subspaces of E having $\mathcal{R}(T) \subset E_\lambda$ and let the Riesz space generated by all E_λ be denoted by $E_\Lambda =< \cup_{\lambda\in\Lambda} E_\lambda >$. The family is said to be T-conditionally independent if, for each pair of disjoint sets $\Lambda_1, \Lambda_2 \subset \Lambda$, we have that E_{Λ_1} and E_{Λ_2} are T-conditionally independent.

The below definition of T-conditional independence for sequences follows as provided in [43].

Definition 3.22. *[43][Definition 3.9]*
Let E be a Dedekind complete Riesz space with conditional expectation T and weak order unit $e = Te$. We say that the sequence $(s_n) \in E$ is T-conditionally independent if the family $< \{s_n\} \cup \mathcal{R}(T) >, n \in \mathbb{N}$ of Dedekind complete Riesz spaces is T-conditionally independent.

4 Bernoulli processes in Riesz spaces

In this chapter, we will introduce bernoulli processes in Riesz Spaces as proven by Kuo, Watson and Vardy in [23].

The multiplicative structure of an f-algebra is used to access the averaging properties of conditional expectation operators. F-algebras provide a multiplicative structure that is compatible with the order and additive structures on the space. For a Dedekind complete Riesz space E and a weak order unit e of E, the ideal of E generated by e, E^e, has a natural f-algebra since associativity, distributivity and positivity hold in E^e satisfying the definition of an f-algebra as defined on page 37. By setting $(Pe) \cdot (Qe) = PQe = (Qe) \cdot (Pe)$ for band projections P and Q and extending to E^e by use of Freudenthal's Theorem, a natural f-algebra structure can be constructed. The multiplicative structure is extended to the universal completion by this process and in E^+, under this multiplication; associativity, distributivity and positivity hold. In this setting, e is the multiplicative unit.

We recall the below definition from [23][Page 3].

Definition 4.1.
Let E be a Dedekind complete Riesz space with weak order unit and T a strictly positive conditional expectation on E. Then, the space E is said to be T-universally complete if for each increasing net (f_α) in E^+ with (Tf_α) order bounded in the universal completion E^u, we have that (f_α) is order convergent.

Wen-Chi Kuo, C.C.A Labuschagne and B.A Watson show in [24] that the space and the conditional expectation operator can be extended to make the extended space T-universally complete with respect to the extended T, if the space E is not originally complete. From [16, 22], the extended space is also known as the natural domain of T, denoted $dom(T)$.

We note the below results as provided in [23].

Theorem 4.2. *[23][Page 4]*
Suppose that E is T-universally complete with $\mathcal{L}^1(T) = dom(T) = E$. Since the universal completion of E (E^u) is an f-algebra, we can define a vector space

$$\mathcal{L}^2(T) := \{x \in \mathcal{L}^1(T) | x^2 \in \mathcal{L}^1(T)\}. \tag{16}$$

Proof.

For $f, g \in \mathcal{L}^2(T)$

$$\begin{aligned} 0 &\leq (f \pm g)^2 \\ &= f^2(\pm)2fg + g^2. \end{aligned}$$

Re-arranging gives us that

$$(\pm)2fg \leq f^2 + g^2$$

and

$$2|fg| \leq f^2 + g^2 \in \mathcal{L}^1(T) = E.$$

Thus, $fg \in \mathcal{L}^1(T) = E$. It also follows from the above that $\mathcal{L}^2(T)$ is a vector space. $\square$

We note the below lemma taken from [23][Lemma 3.1].

Lemma 4.3. *[23]*
Let E be a Dedekind complete Riesz space with weak order unit, e, and suppose that T is a conditional expectation operator on E with $Te = e$. If $f, g, fg \in E$ with $g \in \mathcal{R}(T)$ then $g \cdot Tf \in E$ and $T(fg) = g \cdot Tf$.

Theorem 4.4. *[23][Theorem 3.2]*
Let E be a T-universally complete Riesz space with weak order unit, e, where T is a strictly positive conditional expectation operator with $Te = e$ and let S be a conditional expectation operator on E with $TS = T = ST$. If $f \in \mathcal{L}^2(T)$, then $Sf \in \mathcal{L}^2(T)$.

Corollary 4.5. *[23][Corollary 3.3]*
Let E be a T-universally complete Riesz space with weak order unit, e, where T is a strictly positive conditional expectation operator with $Te = e$. Let S, J be conditional expectation operators on E with $TS = T = ST$, $TJ = T = JT$ and $JS = J = SJ$. If $f, g \in \mathcal{L}^2(T)$ then $S(f \cdot Jg) = Jg \cdot S(f)$.

Proof.
Since $g \in \mathcal{L}^2(T)$ it follows from Theorem 4.4 that $Jg \in \mathcal{L}^2(T)$. Now $f, Jg \in \mathcal{L}^2(T)$ so $f, Jg, f \cdot Jg \in E$ so from Lemma 4.3 it follows that $Jg \cdot Sf \in E$ and $S(f \cdot Jg) = Jg \cdot Sf$. $\square$

Theorem 4.6. *(Tchebichev's Inequality) [23] [Theorem 3.4]*
Let E be a Dedekind complete Riesz space with conditional expectation T and weak order unit $e = Te$. Let $f \in \mathcal{L}^2(T), f \geq 0$ and $\epsilon \in \mathbb{R}, \epsilon > 0$, then

$$TP_{(f-\epsilon e)^+}e \leq \frac{1}{\epsilon^2}T(f^2).$$

We note the below lemma taken from [23].

Lemma 4.7. *[23][Lemma 4.1]*
Let E be a T-universally complete Riesz space with weak order unit, $e = Te$, where T is a strictly positive conditional expectation operator on E. Let $f, g \in \mathcal{L}^2(T)$. If f, g are T-conditionally independent then

$$Tfg = Tf \cdot Tg = Tg \cdot Tf. \tag{17}$$

From Theorem 4.4, $(f - Tf) \in \mathcal{L}^2(T)$ since $f \in \mathcal{L}^2(T)$ implies $Tf \in \mathcal{L}^2(T)$. Thus, $(f - Tf)^2 \in \mathcal{L}^1(T)$. Hence, $\forall\, f \in \mathcal{L}^2(T)$, $T(f - Tf)^2$ exists. As shown in [23], the variance of f is now defined by

$$var(f) = T(f - Tf)^2 = Tf^2 - (Tf)^2.$$

We note the below result and following text taken from [23].

Theorem 4.8. *(Bienayme Equality) [23][Theorem 4.2]*
Let E be a T- universally complete Riesz space with weak order unit, $e = Te$, where T is a strictly positive conditional expectation operator on E. If $(f_k)_{k\in\mathbb{N}}$ is a T-conditionally independent sequence in $\mathcal{L}^2(T)$, then for each $n \in \mathbb{N}$,

$$var(\sum_{k=1}^{n} f_k) = \sum_{k=1}^{n} var(f_k). \tag{18}$$

Let $(P_k)_{k\in\mathbb{N}}$ be a sequence of T-conditionally independent band projections and E be a Dedekind Riesz space with weak order unit e and conditional expectation operator T with $Te = e$. $(P_k)_{k\in\mathbb{N}}$ is called a Bernoulli process if $TP_k e = f$ for all $k \in \mathbb{N}$ for some fixed $f \in E$. Hence, the pay-off at time n is $S_n = \sum_{j=1}^{n} P_j e$.

The following theorem is from the Bernoulli processes in Riesz Spaces paper by Kuo, Watson and Vardy [23].

Theorem 4.9 (Poisson Theorem)**.** *[23][Theorem 5.6]*
Let E be a T-universally complete Riesz space with weak order unit, $e = T_\lambda e$, where T is a strictly positive conditional expectation operator on E and $P_{n,k}$, $k = 1, ..., n$, $n \in \mathbb{N}$, be T-conditionally independent band projections with $TP_{n,k}e = \lambda_n$ for all $k = 1, ..., n$, $n \in \mathbb{N}$. If $S_n = \sum_{k=1}^{n} P_{n,k}e$ are T-conditionally independent with $TS_n = \lambda$, $n \in \mathbb{N}$, then for each $j = 0, 1, ...,$

$$TP_{S_n=je}e \to \frac{\lambda^j}{j!}e^{-\lambda}$$

e-uniformly as $n \to \infty$. Here, $P_{S_n=je}$ is the band projection on the band where $S_n = je$.

5 Stein-Chen Method in Riesz Spaces

In this section, we will now translate the Stein-Chen method of approximating the bounds of the laws of small numbers to Riesz Spaces.

Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $u = Tu$ (we avoid the use of e to denote the weak order unit as e is reserved for the Euler constant). Let $Q_1, \dots, Q_n$ be T-conditionally independent band projections on E, then $q_1 := Q_1 u, \dots, q_n := Q_n u$ are T-conditionally independent components of u.

Set $h_j := Tq_j$ and let $H := \sum_{i=1}^{n} h_j$, then $0 \le H \le nu$ and thus $0 \le H/n \le u$.

The next challenge is the lifting of $Po(k;\lambda)$ to the Riesz space setting. We recall that

$$Po(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

We note here that $Po(k;0) = \begin{cases} 1, & k = 0 \\ 0, & k > 0 \end{cases}.$

As $R(T)$ is a ring, in fact a universally complete f-algebra, and $H \in R(T)$ we have that H^k exists and is in $R(T)$ for all $k \in \mathbb{N}_0$, here $H^0 := u$. Thus, the first part of the lifting of Po to Riesz spaces can be achieved directly by setting

$$Po(k;H) = \frac{H^k e^{-H}}{k!}.$$

Let $P_H = I - e^{-H}$. In line with the observation made of the scalar case, we note that

$$Po(0,H) = ue^{-H} = (I - P_H)u.$$

Further to this,

$$e^{-H} = u + \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} H^i$$

and $(I - P_H)e^{-H} = (I - P_H)u$.

The inequality $1-\beta \le e^{-\beta}$ for all $0 \le \beta \le 1$ gives $u - v \le e^{-v}$ for all $0 \le v \le u$ after applying $u = \frac{v}{\beta}$. Applying this to $v = H/n$ we have $0 \le u - \frac{H}{n} \le e^{-H/n}$. Now iteratively multiplying in the f-algebra E^u we get

$$0 \le \left(u - \frac{H}{n}\right)^n \le e^{-H}.$$

We note that $Po(k,H) \in R(T)$ for all $k \in \mathbb{N}_0$.

We now say that a Riesz space random variable, f, composed of sums of components of u has T-conditional Poisson distribution with parameter $H \in R(T)$ if

$$T(I - P_{|f-ku|})u = Po(k;H). \tag{19}$$

[In this case $Tf = H$ and $Var_T(f) = T(f^2) - T(f)^2$, if $f^2 \in E$ since although multiplication of elements of E is defined it does not necessarily result in elements contained in E.]

In the above f is of the form $f = \sum_{i=0}^{m} ir_i$, where $r_i = R_i u$ and R_i is a band projection on E, for each $i = 0, \ldots, m$. Further, one can without loss of generality assume $\sum_{i=0}^{m} R_i = I$. In this case (19) becomes

$$Tr_i = Po(i; Tf) = \frac{(Tf)^i e^{-Tf}}{i!}, \tag{20}$$

for each $i = 1, \ldots, m$.

Let

$$w = \sum_{i=1}^{n} q_i.$$

Recall $g(j, \lambda, A)$ as defined in Equation 8. The next challenge is the lifting of $g(j, \lambda, A)$ to the Riesz space setting. It should be noted that many of these problems exist even extending this to a conditional version.

We begin by noting that the band, B_H, contains the band, B_w, generated by w, that is, $P_H \geq P_w$ as if this were not the case then $(I - P_H)w \neq 0$ but $TP_H = P_H T$ since $H \in R(T)$, see [24]. And as we have assumed T strictly positive so

$$0 \neq T(I - P_H)w = (I - P_H)Tw = (I - P_H)H = 0.$$

Therefore, a contradiction and thus $(I - P_H)w = 0$ giving the required result.

It now makes sense to define

$$(I - P_H)g(w, H, A) = 0.$$

It remains for us to define $P_H g(w, H, A)$. Recall that if E is T-universally complete, then $R(T)$ is a universally complete Riesz subspace of E - see [28] - and u is the algebraic unit for this f-algebra. In the band generated by H in the universally complete space $R(T)$, H is a weak order unit and hence a multiplicative unit. Thus, there is $J \in R(T)$ such that $P_H(H \cdot J) = P_H u = P_H(J \cdot H)$. Here J is a partial inverse of H and for efficiency we choose J so that $(I - P_H)J = 0$, if this were not the case we could replace J by $P_H J$.

We are now in a position to proceed with the definition of $g(j, H, A)$ in the Riesz space setting. We begin by setting $g(0, H, A) = 0$. For $j \geq 1$ we set

$$(I - P_H)g(j, H, A) = \frac{1}{j}(\chi_A(0) - \chi_A(j))(I - P_H)u \tag{21}$$

and for $j \geq 0$

$$P_H g(j + 1, H, A) = jJ \cdot g(j, H, A) + \chi_A(j)J - J \cdot Po(A, H).$$

Setting $g(j,H,A) = (I-P_H)g(j,H,A) + P_H g(j,H,A)$ defines $g(j,H,A)$ and we observe that

$$Hg(j+1,H,A) = j\cdot P_H g(j,H,A) + \chi_A(j)P_H u - P_H Po(A,H)$$

and

$$0 = j(I-P_H)g(j,H,A) + \chi_A(j)(I-P_H)u - (I-P_H)Po(A,H),$$

where we have used that $(I-P_H)Po(A,H) = \chi_A(0)(I-P_H)u$. Summing now gives

$$Hg(j+1,H,A) = j\cdot g(j,H,A) + \chi_A(j)u - Po(A,H). \tag{22}$$

We note here that $g(j,H,A)\in R(T)$ for all j.

Lemma 5.1.

$$Hg(j+1,H,A) = j!J^j e^H(Po(A\cap N(j),H) - Po(A,H)\,Po(N(j),H)), \tag{23}$$

where $N(j) = \{0,\dots,j\}$.

Proof.
Multiplying (22) by H^j and dividing by $j!$ gives

$$\frac{H^{j+1}}{j!}g(j+1,H,A) = \frac{H^j}{(j-1)!}g(j,H,A) + \frac{H^j}{j!}\chi_A(j)u - Po(A,H)\frac{H^j}{j!}. \tag{24}$$

Setting

$$G(j,H,A) = \frac{H^j}{(j-1)!}g(j,H,A),$$

(24) becomes

$$G(j+1,H,A) - G(j,H,A) = \frac{H^j}{j!}\chi_A(j)u - Po(A,H)\frac{H^j}{j!}. \tag{25}$$

Summing over j and using that $G(0,H,A)=0$, we have

$$G(j+1,H,A) = \sum_{0\le k\le j, k\in A}\frac{H^k}{k!} - Po(A,H)\sum_{k=0}^{j}\frac{H^k}{k!}. \tag{26}$$

We have from (26),

$$G(j+1,H,A) = e^H(Po(A\cap N(j),H) - Po(A,H)\,Po(N(j),H)), \tag{27}$$

and hence

$$\frac{H^{j+1}}{j!}g(j+1,H,A) = e^H(Po(A\cap N(j),H) - Po(A,H)\,Po(N(j),H)). \tag{28}$$

Multiplying by $J!J^j$ now gives (23).

□

Corollary 5.2.
$A \mapsto g(j, H, A)$ is an additive set function from $\mathcal{P}(\mathbb{N}_0) \to R(T)$ with $g(j, H, \emptyset) = 0 = g(j, H, \mathbb{N}_0)$, where $\mathcal{P}(\mathbb{N}_0)$ denotes the power set.

Proof.
From (21) we have

$$A \mapsto (I - P_H)g(j, H, A) = \frac{1}{j}(\chi_A(0) - \chi_A(j))(I - P_H)u$$

is an additive set function as $A \mapsto \chi_A(0) - \chi_A(j)$ is an additive set function on $\mathcal{P}(\mathbb{N}_0)$ for each j and $(I - P_H)u \in R(T)$. Further, for $A = \emptyset$ or $A = \mathbb{N} + 0$, $A \mapsto \chi_A(0) - \chi_A(j) = 0$ giving that $(I - P_H)g(j, H, A) = 0$.

From Lemma 5.1,

$$A \mapsto Hg(j + 1, H, A) = j!J^j e^H (Po(A \cap N(j), H) - Po(A, H)\, Po(N(j), H))$$

is an additive set function as $A \mapsto Po(A \cap N(j), H)$ and $A \mapsto Po(A, H)$ are additive set functions. Further $J, e^H, Po(A, H) \in R(T)$ giving that $Hg(j + 1, H, A) \in R(T)$. Now $G(0, H, A) = 0$ and since $Po(\emptyset, H) = 0$ it follows from the Lemma 5.1 that $Hg(j + 1, H, \emptyset) = 0$. Finally

$$Hg(j + 1, H, \mathbb{N}_0) = j!J^j e^H (Po(N(j), H) - Po(\mathbb{N}_0, H)\, Po(N(j), H)) = 0,$$

as $Po(\mathbb{N}_0, H) = u$.

□

Lemma 5.3.
For fixed i we have for $j < i$, $Hg(j + 1, H, \{i\})$ is negative and decreasing in j and

$$Hg(j + 1, H, \{j + 1\}) = -\frac{e^{-H}}{j + 1} \sum_{k=0}^{j} \frac{H^{k+1}}{k!}.$$

Proof.
From (23), for $j < i$, we have that

$$\begin{aligned} Hg(j + 1, H, \{i\}) &= -j!J^j e^H Po(\{i\}, H)\, Po(N(j), H) \\ &= -Po(\{i\}, H) \sum_{k=0}^{j} \frac{j!J^{j-k}}{k!} \\ &= -Po(\{i\}, H) \sum_{t=0}^{j} \frac{j!J^t}{(j - t)!} \end{aligned}$$

where $t = j - k$. Here

$$\frac{j!}{(j - t)!} = \begin{cases} j(j - 1)(j - 2) \ldots (j - t + 1), & t \in \{1, \ldots, j\} \\ 1, & t = 0 \end{cases}$$

which is increasing in j. Thus, $Hg(j+1,H,\{i\})$ is negative and decreasing on $\{0,\dots,i-1\}$. Finally

$$Hg(j+1,H,\{j+1\}) = -Po(\{j+1\},H)\sum_{k=0}^{j}\frac{j!J^{j-k}}{k!} = -e^{-H}\sum_{k=0}^{j}\frac{H^{k+1}}{k!(j+1)}$$

which simplifies to the stated formula.

□

Lemma 5.4.
For fixed i, we have for $j \geq i$, $Hg(j+1,H,\{i\})$ is non-negative and decreasing in j and

$$Hg(j+1,H,\{j\}) = Po(\{j+1,j+2,\dots\},H).$$

Proof.
From (23), for $j \geq i$, we have that

$$\begin{aligned}
Hg(j+1,H,\{i\}) &= j!J^j e^H(Po(\{i\},H) - Po(\{i\},H)\,Po(N(j),H)) \\
&= j!J^j e^H Po(\{i\},H)(u - Po(N(j),H)) \\
&= j!J^j Po(\{i\},H)\left(e^H - \sum_{k=0}^{j}\frac{H^k}{k!}\right) \\
&= Po(\{i\},H)\sum_{k=j+1}^{\infty}\frac{j!H^{k-j}}{k!} \\
&= Po(\{i\},H)\sum_{s=1}^{\infty}\frac{j!H^s}{(j+s)!}
\end{aligned}$$

where $s = k-j$. Here, for $s \in \mathbb{N}$,

$$\frac{j!}{(j+s)!} = \frac{1}{(j+1)\dots(j+s)}$$

which is decreasing in j. Thus, $Hg(j+1,H,\{i\})$ is non-negative and decreasing on for $j \geq i$.

Finally

$$Hg(j+1,H,\{j\}) = Po(\{j\},H)\sum_{k=j+1}^{\infty}\frac{j!H^{k-j}}{k!} = e^{-H}\sum_{k=j+1}^{\infty}\frac{H^k}{k!}$$

which simplifies to the stated formula.

□

Combining the above two lemmas we now obtain bounds

$$\Delta(j+1,H,A) := H(g(j+2,H,A) - g(j+1,H,A))$$

for $A \subset \mathbb{N}_0$.

Lemma 5.5.
Let $A \subset \mathbb{N}_0$ then

$$-(u - e^{-H}) \leq \Delta(j+1,H,A) \leq u - e^{-H}$$

Proof.
From Lemmas 5.3 and 5.4, $\Delta(j+1,H,\{i\}) \leq 0$ for all $i \neq j+1$ and for $i = j+1$ we have

$$\Delta(j+1,H,\{j+1\}) = e^{-H} \sum_{k=j+2}^{\infty} \frac{H^k}{k!} + \frac{e^{-H}}{j+1} \sum_{k=0}^{j} \frac{H^{k+1}}{k!}.$$

Now

$$\frac{1}{k!(j+1)} = \frac{k+1}{j+1}\frac{1}{(k+1)!} \leq \frac{1}{(k+1)!}$$

for $k = 0, \ldots, j$. Thus,

$$\Delta(j+1,H,\{j+1\}) \leq e^{-H} \sum_{k=j+2}^{\infty} \frac{H^k}{k!} + e^{-H} \sum_{k=1}^{j+1} \frac{H^k}{k!} = e^{-h}(e^h - u).$$

Now since $\Delta(j+1,H,\{i\}) \leq 0$ for all $i \neq j+1$, we have from additivity that

$$\Delta(j+1,H,A) \leq e^{-h}(e^h - u).$$

Further from Corollary 5.2

$$\Delta(j+1,H,A) \geq \Delta(j+1,H,\mathbb{N}_0 \backslash \{j+1\}) = 0 - \Delta(j+1,H,\{j+1\}) \geq -e^{-h}(e^h - u)$$

from the above computation.

$\square$

We recall that w is a sum of band projections and can thus be expressed as

$$w = \sum_{j=0}^{n} j r_j,$$

where $r_j = R_j u$ and $R_j, j = 0, \ldots, n$, are band projections with $R_i R_j = 0$ for all $i \neq j$ and $\sum_{j=0}^{n} R_j = I$. We thus set

$$g(w,H,A) = \sum_{j=0}^{n} R_j g(j,H,A) = \sum_{j=0}^{n} R_j u \cdot g(j,H,A).$$

It should be noted that, while $g(j, H, A) \in R(T)$, $g(w, H, A)$ is not generally in $R(T)$. With this definition we see that

$$g(w + ku, H, A) = \sum_{j=0}^{n} R_j g(j + k, H, A) = \sum_{j=0}^{n} R_j u \cdot g(j + k, H, A)$$

for $k \in \mathbb{N}_0$.

Applying R_j to (22) we obtain

$$H \cdot R_j g(j + 1, H, A) = jR_j u \cdot g(j, H, A) + \chi_A(j) R_j u - R_j Po(A, H). \quad (29)$$

Here we have used that

$$R_j(H \cdot g(j+1, H, A)) = R_j u \cdot (H \cdot R_j g(j+1, H, A)) = H \cdot (R_j u \cdot g(j+1, H, A)$$

as E is an $R(T)$ module and the multiplication is commutative. Further to this we note that

$$R_k w = \sum_{j=0}^{n} jR_j u = kR_k u$$

giving

$$jR_j u \cdot g(j, H, A) = R_j w \cdot g(j, H, A) = R_j u \cdot w \cdot g(j, H, A) = w \cdot R_j g(j, H, A).$$

Thus (22) can be rewritten as

$$H \cdot R_j g(j + 1, H, A) = w \cdot R_j g(j, H, A) + \chi_A(j) R_j u - R_j Po(A, H) \quad (30)$$

which when summed over j gives

$$H \cdot g(w + u, H, A) = w \cdot g(w, H, A) + \sum_{j \in A} R_j u - Po(A, H). \quad (31)$$

Taking the, T, conditional expectation of (31) we have

$$H \cdot Tg(w + u, H, A) = T(w \cdot g(w, H, A)) + \mathbb{P}_T(w \in A) - Po(A, H), \quad (32)$$

where $\mathbb{P}_T(w \in A) := \sum_{j \in A} TR_j u$ is the $R(T)$ valued conditional probability of $w \in A$. Here we have used that $Po(A, H) \in R(T)$ to give $TPo(A, H) = Po(A, H)$ and that T is an averaging operator (see [24, 28]) to give $T(H \cdot g(w + u, H, A)) = H \cdot Tg(w + u, H, A)$ since $H \in R(T)$.

Denoting $w_i := \sum_{j \neq i} q_j$, we have $w = w_i + q_i$ and as

$$(I - Q_i)w = \sum_{j} (I - Q_i) Q_j u = \sum_{j \neq i} (I - Q_i) Q_j u = (I - Q_i) w_i,$$

so

$$q_i \cdot w = w - (u - q_i) \cdot w_i = w - w_i + q_i \cdot w_i = q_i \cdot (u + w_i).$$

Thus,

$$Q_i w = Q_i(w_i + u)$$

giving

$$q_i \cdot g(w, A, H) = Q_i g(w, A, H) = Q_i g(w_i + u, A, H) = q_i \cdot g(w_i + u, A, H).$$

Here

$$w_i = \sum_{j=0}^{n} j r_j^i,$$

where $r_j^i = R_j^i u$ and $R_j^i, j = 0, \ldots, n$ are band projections with $R_k^i R_j^i = 0$ for all $k \neq j$ and $\sum_{j=0}^{n} R_j^i = I$ and

$$w_i + u = \sum_{j=0}^{n} (j+1) r_j^i$$

giving

$$g(w_i + u, A, H) = \sum_{j=0}^{n} R_j^i g(j+1, A, H).$$

In the above R_j^i are sums of products of $Q_k, k \neq i$ and thus $g(w_i + u, A, H)$ is in the Riesz subspace, E_i of E generated by $R(T), q_1 \ldots, q_{i-1}, q_{i+1}, \ldots, q_n$ which is T-conditionally independent of q_i, thus

$$T(q_i \cdot g(w_i + u, A, H)) = h_i \cdot Tg(w_i + u, A, H).$$

Combining the above gives

$$T(q_i \cdot g(w, A, H)) = T(q_i \cdot g(w_i + u, A, H)) = h_i \cdot Tg(w_i + u, A, H). \quad (33)$$

Combining (32) and (33) we have

$$\sum_i h_i \cdot T(g(w + u, H, A) - g(w_i + u, H, A)) = \mathbb{P}_T(w \in A) - Po(A, H). \quad (34)$$

Since $Q_i(w_i + q_i + u) = Q_i(w_i + 2u, H, A)$ and $(I - Q_i)(w_i + q_i + u) = (I - Q_i)(w_i + u)$ we have

$$g(w + u, H, A) - g(w_i + u, H, A) = Q_i(g(w_i + 2u, H, A) - g(w_i + u, H, A)).$$

Since q_i and $g(w_i + 2u, H, A) - g(w_i + u, H, A)$ are T-conditionally independent we have

$$T(q_i(g(w_i+2u, H, A) - g(w_i+u, H, A))) = h_i \cdot T(g(w_i+2u, H, A) - g(w_i+u, H, A))$$

and thus

$$T(g(w+u, H, A) - g(w_i+u, H, A)) = h_i \cdot T(g(w_i+2u, H, A) - g(w_i+u, H, A))$$

which when applied to (34) gives

$$\sum_i h_i^2 \cdot T(g(w_i+2u,H,A)-g(w_i+u,H,A)) = \mathbb{P}_T(w \in A)-Po(A,H). \quad (35)$$

As T is a positive operator (35) gives the inequality

$$|\mathbb{P}_T(w \in A)-Po(A,H)| \leq \sum_i h_i^2 \cdot T|g(w_i+2u,H,A)-g(w_i+u,H,A)|. \quad (36)$$

Applying the earlier results we have

$$\begin{aligned} |g(w_i+2u,H,A)-g(w_i+u,H,A)| &= \sum_{j=0}^{n_i} R_j^i |g(j+2,H,A)-g(j+1,H,A)| \\ &= \sum_{j=0}^{n_i} R_j^i J(u-e^{-H}) \end{aligned}$$

which combine to give

$$|\mathbb{P}_T(w \in A)-Po(A,H)| \leq \sum_i h_i^2 \sum_{j=0}^{n_i} T R_j^i J(u-e^{-H}) = \sum_i h_i^2 J(u-e^{-H}).$$

Note that

$$\sum_i h_i^2 J \leq \sup_{i=0,\ldots,n} h_i \cdot J \sum_{k=1}^{n} h_k = \sup_{i=0,\ldots,n} h_i \cdot JH = P_H \sup_{i=0,\ldots,n} h_i = \sup_{i=0,\ldots,n} h_i.$$

Thus,

$$|\mathbb{P}_T(w \in A)-Po(A,H)| \leq (u-e^{-H}) \sup_{i=0,\ldots,n} h_i.$$

6 Conclusion

In this dissertation, we have translated to the Riesz space setting the Stein-Chein method of measuring the error of approximation of the laws of small numbers.

The translation of the result to the Riesz space was reliant mainly on the work that Kuo, Watson, Vardy and others have done in translating random processes to Riesz spaces and the use of functional calculus.

References

[1] C.D. Aliprantis, K.C. Border, *Infinite Dimensional Analysis: A Hitchhiker's guide, Third edition*, Springer Verlag, 2006.

[2] Y. Azouzi, M. Trabelsi, L^p Spaces with Respect to Conditional Expectation on Riesz Spaces, *J. Math. Anal. Appl.*, **447**(2017), 798-816.

[3] A.D. Barbour, L.H.Y. Chen, An introduction to Stein's method, *Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, Vol 4, Singapore University Press and Word Scientific*, (2005), preface pages xi-xii.

[4] A.D. Barbour, P. Hall, On the rate of Poisson convergence, *Math. Proc. Camb. Phil. Soc.*, **95**(1984), 473-408.

[5] A.D. Barbour , G.K Eagleson, Poisson Approximation for Some Statistics Based on Exchangeable Trials, *Advances in Applied Probability*, **Vol. 15**(1983), 585-600.

[6] P. Billingsley, Probability and Measure, Anniversary Edition, Wiley, New Jersey, 2012.

[7] K. Boulabiar, G. Buskes, A. Triki, Results in f-algebras, *Positivity, Trends in Mathematics*, (2007), 73-96.

[8] Z. Brzezniak, *Basic Stochastic Processes*, Springer, 1998.

[9] G. Buskes, A. van Rooij, Almost f-algebras: Commutativity and the Cauchy-Schwartz inequality, *Positivity*, **4**(2000), 227-231.

[10] L.H.Y. Chen, L. Goldstein, Q.M. Shao, Normal approximation by Stein's method, *Springer verlag*, Berlin, 2011.

[11] L.H.Y. Chen, L. Goldstein, A. Röllin, Stein's method via induction, *Elec. J. Probab.*, **25**(2020), 1-49.

[12] E. De Jonge, A.C.M. Van Rooij, *Introduction to Riesz Spaces*, Mathematisch Centrum, Amsterdam, 1977.

[13] R. DeMarr, A martingale convergence theorem in vector lattices, *Canadian J.Math.*, **18**(1966), 424-432.

[14] P. Dey, Stein-Chen method for Poisson approximation, *ST414 course notes, University of Warwick*, 19**

[15] P. Dey, *Stein-Chen Method For Poisson Approximation*,University Of Warwick, New York, 2014.

[16] P.G. Dodds, C.B. Huijsmans, B. De Pagter, Characterizations of conditional expectation-type operators, *Pacific J. Math.*, **141**(1990), 55-57.

[17] M. Falk, J. Husler, R.D. Reiss, *Laws of small numbers: extremes and rare events, Third edition*, Springer, 2011.

[18] J.J. Grobler, Continuous stochastic processes in Riesz spaces: the Doob-meyer decomposition, *Positivity*, **14**(2010), 731-751.

[19] J.J. Grobler, Doob's optional sampling theorem in Riesz Spaces Positivity, 15 (2011), 617 - 637.

[20] J.J. Grobler, Jensen's and martingale inequalities in Riesz Spaces, *Indag. Math. (N.S.)*, **25**(2014), 275-295.

[21] J.J. Grobler, On the functional calculus in Archimedean Riesz spaces with applications to approximation theorems, *Quaestiones Math*, **11**(1988), 307-321.

[22] J.J. Grobler, B. De Pagter, Operators represantable as multiplication-conditional expectation operators, *J. Operator Theory*, **48**(2002), 15-40.

[23] W.C. Kuo, J.J. Vardy, B.A. Watson, Bernoulli processes in Riesz spaces, 2014.

[24] W.C. Kuo, C.C.A. Labuschagne, B.A. Watson, Conditional expectations on Riesz space, *J. Math. Anal. Appl.*, **303**(2005), 509-521.

[25] W.C. Kuo, C.C.A. Labuschagne, B.A. Watson, Discrete time stochastic processes on Riesz spaces, *Indag. Math. N.S*, **15**(2004), 435-451.

[26] W.C. Kuo, C.C.A. Labuschagne, B.A. Watson, Ergodic theory and the Strong Law of Large Numbers on Riesz spaces, *J. Math. Anal. Appl.*, **325**(2007), 422-437.

[27] W.C. Kuo, D.F. Rodda, B.A. Watson, Hájek-Rényi-Chow Maximal Inequality and a Strong Law of Large Numbers in Riesz Spaces, *J. Math. Anal. Appl.*, **481**(2020), 123462.

[28] W.C. Kuo, M.J. Rogans, B.A. Watson, Mixing inequalities in Riesz spaces, *J. Math. Anal. Appl.*, **456**(2017), 992-1004.

[29] C.C.A. Labuschagne, B.A. Watson, Discrete stochastic integration in Riesz spaces, *Positivity*, **14**(2010), 559-579.

[30] Z. Lin, Z. Bai, *Probability Inequalities*, Science Press Beijing, 2010.

[31] M. Loeve, *Probability Theory I*, 4th Edition, Springer-Verlag, New York, 1977.

[32] W.A.J. Luxemburg, B. De Pagter, Representations of positive projections II, *Positivity*, **9**(2004), 569-605.

[33] W.A.J. Luxemburg, A.C. Zaanen, Riesz Spaces I, North Holland, 1971.

[34] P. Meyer-Nieberg, Banach Lattices, *Springer-Verlag, Berlin*, 1991.

[35] M.M. Rao, Two Characterizations of Conditional Probability, Proc. Amer. Math. Soc., 19 (1976), 75-80.

[36] R.D Reiss, *A Course on Point Processes*, Springer- Verlag, New York, 1993.

[37] N. Ross, Fundamentals of Stein's method, *Probability Surveys*, **8**(2011), 210-293.

[38] H.H. Schaefer, *Banach lattices and positive operators*, Springer Verlag, Berlin, 1974.

[39] C. Stein, Approximate computation of expectations, *Lecture note - Monograph Series, Institute of Mathematical Statistics, Stanford University*, 1986.

[40] G. Stoica, Martingales in vector lattices, *Bull. Math. Soc. Sci. Math. Roumanie. (N.S.)*, **34(82)**(1990), 357-362.

[41] V. Troitsky, Martingales in Banach lattices. *Positivity*, **9**(2005), 437-456.

[42] S. R. S. Varadhan, *Probability Theory*, American Mathematical Soc., 2001.

[43] J.J. Vardy, B.A. Watson, Markov processes in Riesz spaces, *Positivity*, **16**(2012), 373-391 and erratum on page 393.

[44] L.M. Venter, P. van Eldik, Universally Complete Riesz Spaces and f-algebras, *South African J. Sci.*, **84**(1987), 343-346.

[45] L. von Bortkiewicz, *Das Gesetz der kleinen Zahlen*, Teubner, Leipzig.

[46] B.A. Watson, An Ando-Douglas type theorem in Riesz spaces, *Positivity*, **13**(2009), 543-558.

[47] A.C. Zaanen, *Introduction to Operator Theory in Riesz Spaces*, Springer Verlag, Berlin, 1997.

[48] A.C. Zaanen, *Lectures on Riesz spaces*, European Atomic Energy Community, Brussels, 1966.

[49] A.C. Zaanen, W.A.J Luxemburg, *Riesz Spaces*, North Holland Publishing, Amsterdam, 1971.