CHAPTER II

PRELIMINARIES

In this chapter, we present some background in probability theory and Hilbert spaces, which will be used in this work.

2.1 Fundamental Results in Probability

In this section, we review some definitions and theorems in probability theory.

Definition 2.1. Let Ω be any set. A subset $\mathcal{F}$ of the power set of Ω is called a $\sigma-$algebra if

(i) $\varnothing \in \mathcal{F}$,

(ii) if $E \in \mathcal{F}$, then $E^c \in \mathcal{F}$,

(iii) if $E_1, E_2, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$.

Definition 2.2. Let Ω be a nonempty set and $\mathcal{F}$ be a $\sigma-$algebra. A *probability measure* P on $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$ satisfying

(i) $P(\varnothing) = 0$,

(ii) if $A_1, A_2, \ldots$ are disjoint members of $\mathcal{F}$, so that $A_i \cap A_j = \varnothing$ for all pairs i, j satisfying $i \neq j$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

The triple $(\Omega, \mathcal{F}, P)$, comprising a set Ω, a $\sigma-$algebra $\mathcal{F}$ of subsets of Ω and a probability measure P on $(\Omega, \mathcal{F})$, is called a *probability space.* The set Ω is the *sample space* and elements of $\mathcal{F}$ are called *events.*

Definition 2.3. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A *(real-valued) random variable* is a function $X : \Omega \to \mathbb{R}$ with the property that for every Borel set $\mathcal{B}$ in $\mathbb{R}$,

$$X^{-1}(\mathcal{B}) = \{\omega | X(\omega) \in \mathcal{B}\} \in \mathcal{F}.$$

Note that the $P(\{\omega | X(\omega) \in A\})$ is always denoted by $P(X \in A)$.

Theorem 2.4. *If X is a random variable and f is a Borel measurable function, then $f(X)$ is a random variable.*

Definition 2.5. Let X be a random variable on $(\Omega, \mathcal{F}, P)$ and $F_X : \mathbb{R} \to [0, 1]$ be defined by $F_X(x) = P(X \leq x)$ for $x \in \mathbb{R}$. Then F_X is called the *distribution function* of X.

Definition 2.6. The collection of random variables $\{X_i, 1 \leq i \leq n\}$ is called *independent* if for all $x_i \in \mathbb{R}, i = 1, 2, \ldots, n$

$$P\left(\bigcap_{i=1}^{n} \{X_i \leq x_i\}\right) = \prod_{i=1}^{n} P(X_i \leq x_i).$$

An infinite sequence $\{X_n, n \geq 1\}$ is called *independent* if for each positive integer n, the collection of random variables $\{X_i, 1 \leq i \leq n\}$ is independent.

Definition 2.7. Let X be a random variable and $g : \mathbb{R} \to \mathbb{R}$ be a measurable function. The *expectation* of the random variable $g(X)$ is defined as

$$E(g(X)) = \int_{\Omega} g(X)\, dP.$$

Definition 2.8. Let k be a positive integer and c be a constant. If $E(X - c)^k$ exists, we call it the *moment of order k about the point c*. If we take $c = E(X)$, we call $E(X - E(X))^k$ the *central moment of order k*.

Theorem 2.9. *Let X and Y be random variables and $a, b \in \mathbb{R}$. Then the following are true.*

(i) If $E(X)$ and $E(Y)$ are finite, then $E(aX + bY) = aE(X) + bE(Y)$.

(ii) If $X \leq Y$*, then* $E(X) \leq E(Y)$.

(iii) $|E(X)| \leq E(|X|)$.

(iv) If X *and* Y *are independent random variables, then* $E(XY) = E(X)E(Y)$.

Theorem 2.10. *For any random variable* X*, if* $E(X^m)$ *exists, so does* $E(X^n)$ *for* $n \leq m$.

Definition 2.11. If $E(X^2)$ exists, we call $E(X - E(X))^2$ the *variance of* X, and denoted by $Var(X)$. The quantity $\sqrt{Var(X)}$ is called the *standard deviation* of X.

Definition 2.12. Let X and Y be random variables. Then we define the *covariance* of X and Y, denoted as $Cov(X, Y)$, by

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Theorem 2.13. *Let* X, Y *and* Z *be random variables. Then*

(i) $Var(X) \leq E(X^2)$.

(ii) $Cov(X, Y) = E(XY) - E(X)E(Y)$.

(iii) $Cov(X, Y) = Cov(Y, X)$.

(iv) $Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$ *for all* $a, b \in \mathbb{R}$.

(v) $Cov(X, X) = Var(X)$.

(vi) If X *and* Y *are independent random variables, then* $Cov(X, Y) = 0$.

Definition 2.14. Let $\{X_n, n \geq 1\}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$.

(i) We say that X_n *converges in probability* to a random variable X if for every $\epsilon > 0$,

$$\lim_{n \to \infty} P(|X_n - X| > \epsilon) = 0.$$

(ii) We say that X_n *converges almost surely (a.s.)* to a random variable X if

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

(iii) We say that X_n *converges in distribution* to a random variable X if

$$\lim_{n\to\infty} P\left(X_n \leq x\right) = P\left(X \leq x\right).$$

Definition 2.15. ([9]) Let $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ be any sequences of real numbers. A sequence of random variables $\{X_n, n \geq 1\}$ is said to *converge completely* to a constant θ if $\sum_{n=1}^{\infty} a_n P\left(|X_n - \theta| > \epsilon b_n\right) < \infty$ for all $\epsilon > 0$.

Definition 2.16. ([4]) Let $\{X_n, n \geq 1\}$ be a sequence of random variables and $a_n > 0, b_n > 0$ and $q > 0$. If

$$\sum_{n=1}^{\infty} a_n E\left[b_n^{-1}|X_n| - \epsilon\right]_+^q < \infty \text{ for all } \epsilon > 0.$$

The above result is called the *complete moment convergence*, where $x_+^q = (x_+)^q$ and $x_+ = \max\{x, 0\}$.

Lemma 2.17. *(Borel-Cantelli Lemma) Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{A_n, n \geq 1\}$ be a sequence of events in $\mathcal{F}$.*

(i) If $\sum_{n=1}^{\infty} P\left(A_n\right) < \infty$, then

$$P\left(\bigcap_{k=1}^{\infty}\left(\bigcup_{n=k}^{\infty} A_n\right)\right) = 0.$$

(ii) If $\sum_{n=1}^{\infty} P\left(A_n\right) = \infty$ and $\{A_n, n \geq 1\}$ is independent, then

$$P\left(\bigcap_{k=1}^{\infty}\left(\bigcup_{n=k}^{\infty} A_n\right)\right) = 1.$$

Theorem 2.18. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables such that $\{X_n, n \geq 1\}$ converge completely to θ. Then $\{X_n, n \geq 1\}$ converges almost surely to θ.*

Proof. For $\epsilon > 0$, let $A_n = \{|X_n - \theta| > \epsilon\}$. Then

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(|X_n - \theta| > \epsilon) < \infty.$$

By Borel-Cantelli Lemma, $P\left(\bigcap_{k=1}^{\infty}\left(\bigcup_{n=k}^{\infty} A_n\right)\right) = 0$ and so $\{X_n, n \geq 1\}$ converges almost surely to θ. □

Theorem 2.19. *(Markov's Inequality). Let X be a random variable. Then, for any $a > 0$ and non-negative increasing function f such that $f(a) \neq 0$,*

$$P(X \geq a) \leq \frac{E(f(X))}{f(a)}.$$

Theorem 2.20. *(Hölder's Inequality). Let X and Y be random variables such that $E(|X|^p) < \infty$ and $E(|Y|^q) < \infty$ for any pair p, q such that $0 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$E(|XY|) \leq (E|X|^p)^{\frac{1}{p}} (E|Y|^q)^{\frac{1}{q}}.$$

2.2 Fundamental Results in Hilbert Spaces

In this section, we review some knowledge in Hilbert spaces.

Definition 2.21. A *metric (or distance)* on a nonempty set M is a map $d : M \times M \to \mathbb{R}$ satisfying

(i) $d(x, y) \geq 0$ for all $x, y \in M$,

(ii) $d(x, y) = d(y, x)$ for all $x, y \in M$,

(iii) $d(x, y) = 0$ if and only if $x = y$ for all $x, y \in M$,

(iv) $d(x,y) \leq d(x,z) + d(z,y)$ for all $x, y, z \in M$.

The pair (M, d) is called a *metric space.*

Definition 2.22. A metric space (M, d) is said to be *complete* if every Cauchy sequence in M converges.

Definition 2.23. Let (M, d) be a metric space. A subset $D \subseteq M$ is called *dense* if for all $x \in M$, $\{y \in M \mid d(x,y) < \epsilon\} \cap D \neq \varnothing$ for all $\epsilon > 0$.

Definition 2.24. A metric space (M, d) is called *separable* if it contains a countable dense subset.

Definition 2.25. Let M be a vector space over $\mathbb{R}$. A function $\|\cdot\| : M \to [0, \infty)$ is said to be a *norm* on M if

(i) $\|x\| = 0$ if and only if $x = 0$,

(ii) $\|cx\| = |c| \, \|x\|$ for any $x \in M$ and $c \in \mathbb{R}$,

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for any $x, y \in M$.

A vector space equipped with a norm is called a *normed linear space.*

Definition 2.26. Let M be a vector space over $\mathbb{R}$. An *inner product* on M is a map $\langle \cdot, \cdot \rangle : M \times M \to \mathbb{R}$ satisfying

(i) $\langle x, x \rangle \geq 0$ for any $x \in M$,

(ii) $\langle x, y \rangle = \langle y, x \rangle$ for each $x, y \in M$,

(iii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ for each $x, y \in M$ and $\alpha \in \mathbb{R}$,

(iv) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ for each $x, y, z \in M$,

(v) if $\langle x, x \rangle = 0$, then $x = 0$.

A real vector space equipped with an inner product is called a *(real) inner product space.*

Definition 2.27. A *Hilbert space* is a complete inner product space.

Definition 2.28. Let S be a subset of a vector space M. Then the smallest subspace containing S is called the *subspace of M generated by S* or the *subspace of M spanned by S*, denoted by $\langle S \rangle$. If $\langle S \rangle = M$, we say that M is spanned by S or S spans M.

Definition 2.29. Let M be a vector space over $\mathbb{R}$ and S a subset of M. We say that S is *linearly dependent* if there exist distinct elements $v_1, v_2, \ldots, v_n \in S$ and scalars $k_1, k_2, \ldots, k_n \in \mathbb{R}$, not all zero, such that $k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$.

We say that S is *linearly independent* if S is not linearly dependent. In other words, S is linearly independent if and only if for any distinct elements $v_1, v_2, \ldots, v_n \in S$ and any $k_1, k_2, \ldots, k_n \in \mathbb{R}$, if $k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$, then $k_i = 0$ for all $i = 1, 2, \ldots, n$.

Definition 2.30. A subset S of a vector space M is called a *basis* for M if

(i) S spans M, and

(ii) S is linearly independent.

Definition 2.31. Let M be an inner product space. We say that $u, v \in M$ are *orthogonal* if $\langle u, v \rangle = 0$.

Definition 2.32. A nonempty collection $\mathcal{O} = \{\mathbf{e}_\alpha \mid \alpha \in \Lambda\}$ of elements in an inner product space is said to be an *orthogonal set* if $\langle \mathbf{e}_\alpha, \mathbf{e}_\beta \rangle = 0$ for all $\alpha \neq \beta$ in Λ. If, in addition, $\mathbf{e}_\alpha$ has norm one for all $\alpha \in \Lambda$, then we say that the set $\mathcal{O}$ is an *orthonormal set*.

Theorem 2.33. *A Hilbert space is separable if and only if it has a countable orthonormal basis.*

Theorem 2.34. *Let $B = \{\epsilon_j\}_{j=1}^{\infty}$ be an orthonormal set in a real Hilbert space H. Then the following statements are equivalent:*

(i) B is an orthonormal basis for H,

(ii) $x = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j$ *for each* $x \in H$,

(iii) $\langle x, y \rangle = \sum_{j=1}^{\infty} \langle x, e_j \rangle \cdot \langle y, e_j \rangle$ *for each* $x, y \in H$,

(iv) $\|x\|^2 = \sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2$ *for every* $x \in H$.

2.3 Probability in Hilbert Spaces

In this section, let $(\Omega, \mathcal{F}, P)$ be a probability space, let H be a real Hilbert space with the norm $\|\cdot\|$ generated by inner product $\langle \cdot, \cdot \rangle$ and let $\{e_j, j \geq 1\}$ be an orthonormal basis in H. Let $\mathcal{B}$ be the $\sigma-$algebra generated by the class of all open subsets of H.

Definition 2.35. ([16]) A mapping $\mathbf{X} : \Omega \to H$ is called an *$H-$valued random vector* if $\mathbf{X}$ is $\mathcal{B}-$measurable, that is, for every $E \in \mathcal{B}$

$$\mathbf{X}^{-1}(E) = \left\{\omega \left| \mathbf{X}(\omega) \in E \right.\right\} \in \mathcal{F}.$$

Proposition 2.36. *([16]) Let H be a separable Hilbert space and let $\mathbf{X}$ be an $H-$valued random vector. Then $\|\mathbf{X}\|$ is a (real-valued) random vector.*

Definition 2.37. ([19]) Let $\mathbf{X}$ be an $H-$valued random vector. We say that $\mathbf{X}$ has *mean* $m \in H$ if $E(\langle \mathbf{X}, h \rangle) = \langle m, h \rangle$ for all $h \in H$.

Remark 2.38. Let $\mathbf{X}$ be an $H-$valued random vector. Then $\mathbf{X}$ has mean zero if and only if $E\left(X^{(j)}\right) = 0$ where $X^{(j)}$ denote the inner product $\langle \mathbf{X}, e_j \rangle$.

Definition 2.39. Let $\{\mathbf{X}_n, n \geq 1\}$ be a sequence of $H-$valued random vectors.

(i) The sequence $\{\mathbf{X}_n, n \geq 1\}$ is said to *converges in probability* to an $H-$valued random vector $\mathbf{X}$ if, for every $\epsilon > 0$,

$$\lim_{n \to \infty} P(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) = 0.$$

(ii) The sequence $\{\mathbf{X}_n, n \geq 1\}$ is said to *converges almost surely (a.s.)* to an $H-$valued random vector $\mathbf{X}$ if

$$P\left(\lim_{n\to\infty} \|\mathbf{X}_n - \mathbf{X}\| = 0\right) = 1.$$

2.4 Dependence Structure

In this section, we discuss the relationship of some dependence structures of random variables.

The independence assumption used in many results on probability limit theorems is too restricted in some applications. Therefore, some researchers have introduced new concepts relexing the independence structure. For example, Alam and Sexena [1] introduced the definition of negatively associated random variables as the following.

Definition 2.40. ([1]) A finite sequence of random variables $\{X_i, 1 \leq i \leq n\}$ is said to be *negatively associated (NA)*, if for every disjoint subsets A and B of $\{1, 2, \ldots, n\}$, we have

$$Cov\left(f\left(X_i, i \in A\right), g\left(X_j, j \in B\right)\right) \leq 0,$$

whenever f on $\mathbb{R}^{|A|}$ and g on $\mathbb{R}^{|B|}$ are coordinatewise nondecreasing functions and covariance exists. An infinite sequence of random variables is NA if every finite subfamily is NA.

The following is an example of a sequence of NA random variables.

Example 2.41. Let $\{X_i, 1 \leq i \leq n\}$ be a sequence of random variables taking values in $\{0, 1\}$ such that $\sum_{i=1}^{n} X_i = 1$. Then a sequence of random variables $\{X_i, 1 \leq i \leq n\}$ is NA. More details of the proof can be seen in [5].

In 1981, Ebrahimi and Ghosh [7] introduced other types of negative dependence. The definition is given as follows.

Definition 2.42. ([7]) A finite sequence of random variables $\{X_i, 1 \leq i \leq n\}$ is said to be *negatively upper orthant dependent (NUOD)* if for all $x_i \in \mathbb{R}, i = 1, 2, \ldots, n$,

$$P\left(\bigcap_{i=1}^{n}\{X_i \leq x_i\}\right) \leq \prod_{i=1}^{n} P(X_i \leq x_i).$$

It is called *negatively lower orthant dependent (NLOD)* if for all $x_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$,

$$P\left(\bigcap_{i=1}^{n}\{X_i > x_i\}\right) \leq \prod_{i=1}^{n} P(X_i > x_i).$$

If it is both NUOD and NLOD, it is called *negatively orthant dependent (NOD)*. A sequence of random variables $\{X_n, n \geq 1\}$ is said to be *NUOD, NLOD* or *NOD* if for each positive integer n, the sequence of random variables $\{X_i, 1 \leq i \leq n\}$ is NUOD, NLOD or NOD, respectively.

Joag-Dev and Proschan [12] also pointed out that NA random variables must be NOD but NOD random variables are not necessarily NA.

In 2009, Liu [17] extended the concept of NOD to a more general dependence structure, which is called extended negatively dependence. The definition is given as follows.

Definition 2.43. ([17]) A finite sequence of random variables $\{X_i, 1 \leq i \leq n\}$ is called *lower extended negatively dependent (LEND)* if there is some $M > 0$ such that, for all $x_i \in \mathbb{R}, i = 1, 2, \ldots, n$,

$$P\left(\bigcap_{i=1}^{n}\{X_i \leq x_i\}\right) \leq M\prod_{i=1}^{n} P(X_i \leq x_i).$$

It is called *upper extended negatively dependent (UEND)* if there is some $M > 0$ such that, for all $x_i \in \mathbb{R}, i = 1, 2, \ldots, n$,

$$P\left(\bigcap_{i=1}^{n}\{X_i > x_i\}\right) \leq M\prod_{i=1}^{n} P(X_i > x_i).$$

If it is both LEND and UEND, it is called *extended negatively dependent (END)*.

An infinite sequence of random variables $\{X_n, n \geq 1\}$ is call $LEND, UEND$ or END if for each positive integer n, the sequence random variables $\{X_i, 1 \leq i \leq n\}$ is LEND, UEND or END, respectively.

When $M = 1$, a sequence of END random variables is induced to NOD.

In 2013, Wang et al. [20] introduced a more general dependence structure called widely orthant dependence. The definition is given as follows.

Definition 2.44. ([20]) For a sequence of random variables $\{X_n, n \geq 1\}$:

(i) if there exists a sequence of real numbers $\{g_u(n), n \geq 1\}$ such that for each $n \geq 1$ and for all $x_i \in \mathbb{R}$ and $1 \leq i \leq n$,

$$P\left(\bigcap_{i=1}^{n}\{X_i > x_i\}\right) \leq g_u(n)\prod_{i=1}^{n} P\left(X_i > x_i\right),$$

then we say that the sequence $\{X_n, n \geq 1\}$ is *widely upper orthant dependent (WUOD)* with dominating coefficients $g_u(n), n \geq 1$;

(ii) if there exists a sequence of real numbers $\{g_l(n), n \geq 1\}$ such that for each $n \geq 1$ and for all $x_i \in \mathbb{R}$ and $1 \leq i \leq n$,

$$P\left(\bigcap_{i=1}^{n}\{X_i \leq x_i\}\right) \leq g_l(n)\prod_{i=1}^{n} P\left(X_i \leq x_i\right),$$

then we say that the sequence $\{X_n, n \geq 1\}$ is *widely lower orthant dependent (WLOD)* with dominating coefficients $g_l(n), n \geq 1$;

(iii) if $\{X_n, n \geq 1\}$ is both WUOD and WLOD, then we say that the sequence $\{X_n, n \geq 1\}$ is *widely orthant dependent (WOD)* with dominating coefficients $g_u(n)$ and $g_l(n), n \geq 1$.

When $g_u(n) = g_l(n) = M$ for some $M > 0$, a sequence of WOD random variables induced to END random variables.

From all above, we can see that WOD is more general than END, NOD, NA and independence.

The following properties of WOD random variables will be used in this work. Denote $g(n) = \max\{g_u(n), g_l(n)\}$ and $\log x = \ln(\max\{x, e\})$.

Lemma 2.45. *([6]) Let $p \geq 1$ and $\{Y_n, n \geq 1\}$ be a sequence of WOD random variables with $E(Y_n) = 0$ and $E|Y_n|^p < \infty$ for each $n \geq 1$. Then there exist positive constants $C_1(p)$ and $C_2(p)$ depending only on p such that*

$$E\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^{k} Y_i\right|\right)^p \leq [C_1(p) + C_2(p)g(n)] \log^p n \sum_{i=1}^{n} E|Y_i|^p$$

for $1 \leq p \leq 2$, and

$$E\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^{k} Y_i\right|\right)^p \leq C_1(p) \log^p n \sum_{i=1}^{n} E|Y_i|^p + C_2(p)g(n) \log^p n \left(\sum_{i=1}^{n} E|Y_i|^2\right)^{\frac{p}{2}}$$

for $p > 2$.

Proposition 2.46. *([21]) Let $\{Y_n, n \geq 1\}$ be a sequence of WOD random variables. If $f_n(\cdot)$ are all nondecreasing (or all nonincreasing) for $n \geq 1$, then $\{f_n(Y_n), n \geq 1\}$ is still WOD.*