Probability Theory

Forrest Corcoran

Fall, Winter 2022

Contents

Chapter 1 MTH 664

1.1 Probability and Measure Spaces

1.1.1 Probability Spaces

Definition 1.1 Probability Space

A **probability space** is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ where

- Ω is the **outcome space**
- $\mathcal{F} \subset 2^\Omega$ is the σ-algebra of **events**
- $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a **probability function** satisifying the following properties:
  1. $\mathbb{P}(\varnothing) = 0$
  2. $\mathbb{P}(\Omega) = 1$
  3. $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ where $\{A_n : A_n \in \mathcal{F}\}_{n=1}^{\infty}$ are all pairwise disjoint

Example 1.1

Suppose $\Omega = \{1, 2, \dots, N\}$, $\mathcal{F} = \{A \subset \Omega\} := 2^\Omega$, and $\mathbb{P}(A) = \dfrac{|A|}{|\Omega|}$

Is $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space?

Proof.

$$\mathbb{P}(\varnothing) = \frac{|\varnothing|}{|\Omega|} = 0$$

$$\mathbb{P}(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$$

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \frac{|\bigcup_{n=1}^{\infty} A_n|}{|\Omega|} = \frac{|\sum_{n=1}^{\infty} A_n|}{|\Omega|}$$

Yes, $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space ■

Definition 1.2 σ-algebra

Let $\mathcal{F}$ be a collection of subsets in Ω. $\mathcal{F}$ is a σ-algebra on Ω if and only if

- $A_n \in \mathcal{F} \Rightarrow A_n^c \in \mathcal{F}$
- $A_n \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$
- $\varnothing \in \mathcal{F}$

☞ Remark

$$\forall A \in \mathcal{F},\ \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

Proof. First note that, by the definition of σ-algebra, $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$. Now observe that

$$A \cap A^c = \varnothing,\ A \cup A^c = \Omega$$
$$\Rightarrow \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) = 1$$
$$\Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

■

☞ Remark

$$\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(B \cup A)$$

Definition 1.3 Countable Subadditivity

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and A be a collection of events in $\mathcal{F}$ which are not necessarily disjoint. Then

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

Theorem 1.1 Continuity from Below

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and A be a collection of increasing events in $\mathcal{F}$, that is $\forall n \geq 1$

$$A_n \subseteq A_{n+1}$$

We define $\lim_n A_n = \bigcup_n A_n$. So, by countable subadditivity,

$$\mathbb{P}(\lim_n A_n) = \lim_n \mathbb{P}(A_n)$$

Proof. First note that since $A_n \subseteq A_{n+1}, \mathbb{P}(A_n) \leq \mathbb{P}(A_{n+1})$. Additionally, since $A_n \subseteq \Omega$ and $\mathbb{P}(\Omega) = 1, \mathbb{P}(A_n) \leq 1$. Then by Monotone Convergence Theorem, $\lim_{n\to\infty} \mathbb{P}(A_n)$ exists.

Now, set $A = \lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$ and define the sequence of events $\{B_n\}_{n=1}^{\infty}$ such that for each $n \in \mathbb{N}$,

$$B_1 = A_1$$
$$B_2 = A_2 \setminus A_1$$
$$B_n = A_n \setminus A_{n-1}$$

So, by this construction, the B_ns are disjoint and $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Additionally, $\mathbb{P}(B_n) = \mathbb{P}(A_n) - \mathbb{P}(A_{n-1})$. Then we have

$$\mathbb{P}(A) = \mathbb{P}(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathbb{P}(B_n)$$

$$= \Big[\sum_{n=1}^{\infty} \mathbb{P}(A_n) - \mathbb{P}(A_{n-1}) \Big] + \mathbb{P}(A_1)$$

$$= \lim_{n\to\infty} \Big[\sum_{n=1}^{\infty} (\mathbb{P}(A_n) - \mathbb{P}(A_{n-1})) + \ldots + (\mathbb{P}(A_2) - \mathbb{P}(A_1)) \Big] + \mathbb{P}(A_1)$$

$$= \lim_{n\to\infty} \mathbb{P}(A_n)$$

■

☞ **Remark**

We can also show that Continuity from Above, i.e.

$$A_n \supseteq A_{n+1} \Rightarrow \mathbb{P}(\lim_n A_n) = \lim_n \mathbb{P}(A_n)$$

holds. To show this, set $B_n^c = A_n$. By DeMorgan's Law

$$\Big(\bigcap_{n=1}^{\infty} B_n \Big)^c = \bigcup_{n=1}^{\infty} B_n^c = \bigcup_{n=1}^{\infty} A_n$$

and the proof follows as above.

1.1.2 Random Variables

Definition 1.4 Random Variables

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A **random variable** is a function $\mathrm{X} : \Omega \to \mathbb{R}$ such that $\mathrm{X}^{-1}([a, b]) \in \mathcal{F}$.

Definition 1.5 Distribution Function

A **distribution function** is a function

$$F_{\mathrm{X}}(x) = \mathbb{P}\big(\mathrm{X}^{-1}((-\infty, x])\big) = \mathbb{P}(\mathrm{X} \le x)$$

We say a distribution function is **absolutely continuous** if

$$F_{\mathrm{X}}(x) = \int_{-\infty}^{x} g(u)du \tag{1.1}$$

for some $g : \mathbb{R} \to \mathbb{R}$

☞ **Remark**

In general, $\dfrac{d}{d\mathrm{X}} F_{\mathrm{X}} = g(x)$ (except for some non-differentiable points).

Convergence

Definition 1.6 Convergence in Probability

Consider a sequence on random variables $\{X\}_{n=1}^{\infty}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say $\{X\}_{n=1}^{\infty}$ **converges in probability** to X if, $\forall \epsilon > 0$

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

Definition 1.7 Almost Sure Convergence

Now consider a sequence on random variables $\{X\}_{n=1}^{\infty}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say $\{X\}_{n=1}^{\infty}$ converges **almost surely** to X if $\forall \omega \in \Omega$,

$$\mathbb{P}(\{\omega \in \Omega : X_n(\omega) \not\to X(\omega)\}) = 0$$

Theorem 1.2

A sequence of random variables $(X_n)_{n=1}^{\infty}$ converges to X in probability if and only if every subsequence has a further subsequence that converges almost surely to X.

1.2 Expectation

Definition 1.8 Simple Function

A random variable X is called a **simple** or **discrete** random variable if it can be written as

$$X(\omega) = \sum_{j=1}^{n} a_j \mathbb{1}_{A_j}$$

where $a_j \in \mathbb{R}$ and $A_j \cap A_i = \varnothing$ for all $1 \leq i, j \leq n$. Note the function $\mathbb{1}_A$ is defined as

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \text{otherwise} \end{cases}$$

Definition 1.9 Expectation of Simple Function

The **expectation** of a discrete random variable X is given by

$$\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P} = \sum_{j=1}^{n} a_j \mathbb{P}(A_j)$$

Theorem 1.3

If X is a non-negative random variable, we can write

$$\mathbb{E}[X] = \sup\{\mathbb{E}[Y] : 0 \leq Y \leq X,\ Y \text{ simple}\}$$

Theorem 1.4

Given that any function may be approximated arbitrarily closely by a sequence of non-decreasing *simple functions* $(X_n)_{n=1}^{\infty}$*, we may define the expectation of any function* $X = \lim_{n\to\infty} X_n$ *by*

$$\mathbb{E}[X] = \lim_{n\to\infty}\left\{\sum_{j=0}^{n2^n-1}\frac{j}{2^n}\mathbb{P}(j2^{-n} \leq X < (j+1)^{-n}) + n\mathbb{P}(X \geq n)\right\}$$

Theorem 1.5

More generally, we write

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P} = \mathbb{E}[X^+] - \mathbb{E}[X^-]$$

Theorem 1.6

The following properties of the expect value are true:

1. $z \leq x \Rightarrow \mathbb{E}[z] \leq \mathbb{E}[x]$
2. $A \subset B \Rightarrow \mathbb{E}[X\mathbb{1}_A] \leq \mathbb{E}[X\mathbb{1}_B]$
3. $A = \bigcup_{n=1}^{\infty} A_n,\ A_n \subseteq A_{n+1} \Rightarrow \lim_{n\to\infty}\mathbb{E}[X\mathbb{1}_{A_n}] = \mathbb{E}[X\mathbb{1}_A]$

Theorem 1.7 Change of Variable

Let $X : \Omega \to \mathbb{R}$ *be a random variable and* $h : \mathbb{R}^+ \to \mathbb{R}^+$ *be a continuous function. Then*

$$\mathbb{E}[h \circ X] = \int_\Omega (h \circ X)\mathbb{P}(d\omega) = \int_{\mathbb{R}} h(X)F_X(dX)$$

Definition 1.10 Moments

Given a random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a distribution function F_X, we can define the **p-th order moment** of X to be

$$\mathbb{E}[X^p] = \int_\Omega X^p(\omega)\mathbb{P}(d\omega) = \int_{\mathbb{R}} x^p F_X(dx)$$

Furthermore, moments of absolute values ($|X|$) are refered to as **absolute moments** and are given by

$$\mathbb{E}[|X|^p] = p\int_0^\infty x^{p-1}\mathbb{P}(|X| > x)dx$$

1.2.1 Convexity

Theorem 1.8

The spaces $L^p(\Omega, \mathcal{F}, \mathbb{P})$ *with norms given by*

$$\|X\|_p = \left[\int_\Omega |X|^p d\mathbb{P}\right]^{1/p} = \left(\mathbb{E}[|X|]^p\right)^{1/p}$$

are Banach Spaces.

Moreover, the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert Space with an inner product given by

$$\langle X, Y \rangle = \mathbb{E}[XY] = \int_\Omega XY d\mathbb{P}$$

So we have $\|X\|_2 = \langle X, X \rangle^{1/2}$

Definition 1.11

A function ϕ is **convex** on an interval J if for all $a, b \in J$ and $0 \leq t \leq 1$,

$$\phi(ta + (1-t)b) \leq t\phi(a) + (1-t)\phi(b)$$

Theorem 1.9 Line of Support

Suppose ϕ is a convex function on an interval J. Then the following are true:

I. *If J is open:*

  i. *The left-hand and right-hand derivatives of ϕ (ϕ^- and ϕ^+ respectively) exists and are finite as well as non-decreasing on J with $\phi^- \leq \phi^+$*

  ii. *For each $x_0 \in J$, there exists a constant m such that $\phi(x) \geq \phi(x_0) + m(x - x_0)$, $\forall x \in J$.*

II. *If J if half open and the derivative of the open side is finite, the properties of I apply to the closed side with endpoint x_0.*

Theorem 1.10

Let X, Y be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Then the following inequalities hold:

a. *(**Jensen's Inequality**) If ϕ is a convex function on the interval J and $\mathbb{P}(X \in J) = 1$, then*

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$$

Moreover, if ϕ is strictly convex, the above inequality holds iff X is almost surely constant.

b. *(**Lyapounov Inequality**) If $0 < r < s$, then*

$$(\mathbb{E}[|X|^r])^{1/r} \leq (\mathbb{E}[|X|^s])^{1/s}$$

c. *(**Holder's Inequality**) Let $p \geq 1$. If $X \in L^p$, $Y \in L^q$, $\frac{1}{p} + \frac{1}{q} = 1$, then $XY \in L^1$ and*

$$\mathbb{E}[|XY|] \leq (\mathbb{E}[|X|^p])^{1/p}(\mathbb{E}[|Y|^q])^{1/q}$$

d. *(**Cauchy-Schwartz Inequality**) If $X, Y \in L^2$, then $XY \in L^1$ so we have*

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$$

e. *(**Minkowski's Inequality**) Let $p \geq 1$. if $x, Y \in L^p$ then*

$$\|X - Y\|_p \leq \|X\|_p + \|Y\|_p$$

f. (**Markov/Chebyshev Inequalities**) *Let* $p \geq 1$. *If* $X \in L^p$ *then for* $\lambda > 0$

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{\mathbb{E}[\,|X|^p\,]\mathbb{1}_{[|X|\geq\lambda]}}{\lambda^p} \leq \frac{\mathbb{E}[\,|X|^p\,]}{\lambda^p}$$

More generally, if h *is a non-negative increasing function on an interval containing the range of* X, *then*

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}[\,h(X)\mathbb{1}_{[X\geq\lambda]}\,]}{h(\lambda)}$$

1.2.2 L^p Spaces

Definition 1.12 L^p Probability Space

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $p \geq 1$. Then we define $L^p(\Omega, \mathcal{F}, \mathbb{P})$ to be

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) = \{X : \Omega \to \mathbb{R} : \mathbb{E}[\,|X|^p\,] < \infty\}$$

☞ Remark

For random variables $X, Y \in L^p$,

$$X = Y \iff \mathbb{E}[\,|X - Y|^p\,] = 0 \iff \mathbb{E}[\,|X|^p\,] = \mathbb{E}[\,|Y|^p\,]$$

Theorem 1.11

A sequence of random variables X_n *converges to* X *in* L^p *if:*

$$\lim_{n\to\infty} \mathbb{E}[X_n - X|^p] = 0$$

Theorem 1.12

$L^p(\Omega, \mathcal{F}, \mathbb{P})$ *is complete*

Definition 1.13 Uniform Integrability

A sequence of random variables X_n is said to be uniformly integrable if

$$\lim_{\lambda\to\infty} \sup_n \mathbb{E}[\,\{|X_n|\mathbb{1}_{[|X_n|\geq\lambda]}\}\,] = 0$$

Theorem 1.13 Fatou's Lemma

Let $X_n : \Omega \to [0, \infty]$ *be a sequence of non-negative random variables. Then*

$$\mathbb{E}[\lim_{n\to\infty} \inf X_n(\omega)] \leq \lim_{n\to\infty} \inf \mathbb{E}[X_n(\omega)]$$

We can also show that the reverse Fatou's Lemma holds. If $\exists Y : \Omega \to [0, \infty]$ *such that* $X_n \leq Y$ *for all* n *and* $\mathbb{E}[\,Y\,] < \infty$, *then*

$$\lim_{n\to\infty} \sup \mathbb{E}[X_n(\omega)] \leq \mathbb{E}[\lim_{n\to\infty} \sup X_n(\omega)]$$

Theorem 1.14

Consider $\{X_n\}_{n=1}^{\infty}$. *Then*

$$\mathbb{E}[|X_n - X|] \to 0 \iff \begin{cases} X_n \xrightarrow{\mathbb{P}} X \\ \{X_n\}_{n=1}^{\infty} \text{ is uniformly integrable} \end{cases}$$

1.2.3 Generating σ-algebras

Definition 1.14 Generating σ-algebras

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X : \Omega \to \mathcal{S}$, where $\mathcal{S}$ is a σ-algebra on S. Then $\sigma(X)$ is the **generating σ-algebra** generated by X which is the smallest σ-algebra such that X is a measurable map

☞ **Remark**

Recall: $X : (X, \mathcal{F}) \to (S, \mathcal{S})$ is measurable if $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{S}$.

Definition 1.15 Product σ-algebra

Let $(S_i, \mathcal{S}_i)$ be a finite collection of measurable spaces. $\mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \ldots \otimes \mathcal{S}_n$ is the **product σ-algebra** which is defined as the smallest σ-algebra on $S_1 \times S_2 \times \ldots \times S_n$ such that all projection maps,

$$T_k : S_1 \times S_2 \times \ldots \times S_n \to S_k$$

are measurable.

Definition 1.16 Product Measure

Let $(S_i, \mathcal{S}_i, \mu_i)$ be a measure space. The **product measure** $\mu_1 \times \mu_2 \times \ldots \times \mu_n$ on $\mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \ldots \otimes \mathcal{S}_n$ is defined as

$$\mu_1 \times \mu_2 \times \ldots \times \mu_n(B_1 \times B_2 \times \ldots \times B_n) = \mu_1(B_1)\mu_2(B_2)\ldots\mu_n(B_n)$$

Definition 1.17 Absolutely Continuous

Let $(X, \mathcal{S}, \mu)$ be a measure space. Consider the measure ν. ν is **absolutely continuous** with respect to μ if

$$\mu(A) = 0 \to \nu(A) = 0$$

$\forall A \in \mathcal{S}$. We denote this relationship by $\nu \ll \mu$.

Definition 1.18 Singular Measure

Let $(X, \mathcal{S}, \mu)$ be a measure space. Consider the measure ν. ν is **singular** with respect to μ if $\exists A \in \mathcal{S}$ with $\mu(A) = 0$ and $\nu(A^c) = 0$. We denote this relationship $\nu \perp \mu$.

Theorem 1.15 Lebesgue Decomposition

Let $(X, \mathcal{S}, \mu)$ *be a measure space. Given* $\nu : \mathcal{S} \to [0, \infty)$, *where* ν *is* σ *finite, there exist two unique measures,* $\nu_{ac}, \nu_s : \mathcal{S} \to [0, \infty)$ *such that,*

$$\nu = \nu_{ac} + \nu_s$$

and $\nu_{ac} \ll \mu$ and $\nu_s \perp \mu$.

Theorem 1.16 Radon-Nikodym

Let $(X, \mathcal{S}, \mu)$ be a measure space. Given $\nu : \mathcal{S} \to [0, \infty)$, there exists $h : X \to [0, \infty)$ such that

$$\nu_{ac}(A) = \int_A h d\mu$$

*$\forall A \in \mathcal{S}$. h is called a **density function**.*

1.3 Independence

Definition 1.19 Independence

A collection of random variables $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{S}_i)$ are **independent** if the distribution function, Q, defined as

$$Q(B) = \mathbb{P} \circ (X_1, X_2, ..., X_n)^{-1}(B)$$

where $B \in S_1 \times ... \times S_n$, equals the product measure $Q_1 \times Q_2 \times ... \times Q_n$, where

$$Q_i(B_i) = \mathbb{P} \circ X_i^{-1}(B_i)$$

Theorem 1.17

If $X_1, X_2, ..., X_n$ are independent random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}[|X_j|] < \infty$ for all $1 \leq j \leq n$, then

$$\mathbb{E}[X_1 X_2 ... X_n] = \mathbb{E}[X_1]\mathbb{E}[X_2]...\mathbb{E}[X_n]$$

Theorem 1.18

If X_1, X_2 are independent random variables with distributions Q_1, Q_2, respectively, then the distribution of $X_1 + X_2$ is given by the convolution

$$Q_1 * Q_2(B) = \int_{\mathbb{R}} Q_1(B - y)Q_2(dy)$$

Where B is an event and $B - y = \{b - y : b \in B\}$

Definition 1.20 i.i.d.

A sequence of independent random variables $X_1, X_2, \dots$ is **independent and identically distributed (i.i.d.)** if the distribution of X_n does not depend on n. That is, the distrubition is the same for all n.

1.3.1 Covariance & Variance

Definition 1.21 Covariance

Given two random variables X and Y in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, the **covariance** of X and Y is given by

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \end{aligned}$$

X and Y are said to be **uncorrelated** if $\mathrm{Cov}(X, Y) = 0$.

☞ Remark
Independent random variables are uncorrelated, however, uncorrelated random variables are not necessarily independent.

Definition 1.22 Variance
The **variance** of a random variable X is given by

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{Cov}(X, X)\\ &= \mathbb{E}[X - \mathbb{E}[X]]^2\\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

☞ Remark
The covariance term arises naturally in the variance of a sum of random variables:

$$\mathrm{Var}\Big(\sum_{j=1}^{n} X_j\Big) = \sum_{j=1}^{n} \mathrm{Var}(X_j) + 2\sum_{1\le i<j\le n} \mathrm{Cov}(X_i, X_j)$$

Theorem 1.19
If $X_1, X_2, \ldots, X_n$ are independent random variables in $L^2((\Omega, \mathcal{F}, \mathbb{P}))$, then

$$Var(X_1 + X_2 + \ldots + X_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n)$$

Theorem 1.20 Borel-Cantelli
Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of independent events. If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = 1$, then $\mathbb{P}(A_n \text{ i.o.}) = 1$, where $\mathbb{P}(A_n \text{ i.o.})$ is the probability that A_n occurs "infintely often".
Moreover, $\sum_{n=1}^{\infty} \mathbb{P}(_A n) < \infty$ then $\mathbb{P}(A_n \text{ i.o.}) = 0$.

1.3.2 Independent Random Maps

Definition 1.23
A family of random maps $\{X_t : t \in \Lambda\}$ is independent if and only if $\forall$ disjoint pairs of finite subsets Λ_1, Λ_2, any $V_1 \in L^2(\sigma(\{X_t : t \in \Lambda_1\}), V_2 \in L^2(\sigma(\{X_t : t \in \Lambda_2\})$ are uncorrelated.

Theorem 1.21
Let $Y_1, Y_2, \ldots, Y_n$ be random variables of $(\Omega, \mathcal{F}, \mathbb{P})$ and $Z : \Omega \to \mathbb{R}$. Z is $\sigma(Y_1, Y_2, \ldots, Y_n)$ measurable if and only if $\exists g : \mathbb{R}^n \to \mathbb{R}$ such that $Z = g(Y_1, Y_2, \ldots, Y_n)$.

Corollary 1.1

Suppose X_1, X_2 are independent random maps with values $(S_1, \mathcal{S}_1)$ and $(S_2, \mathcal{S}_2)$. Then for Borel measurable $g_i : \mathcal{S}_i \to \mathbb{R}$, $Z_1 = g_1(X_1)$ and $Z_2 = g_2(X_2)$ are independent.

Definition 1.24 Independent Events

A collection, $\mathcal{C}$, of events $A \in \mathcal{F}$ are **independent events** if the collection of indicator functions

$$\{\mathbb{1}_A : A \in \mathcal{C}\}$$

is a family of independent random maps.

☞ **Remark**

We denote an event A_n which occurs **eventually for all** n by

$$[A_n^c \text{ i.o.}]^c$$

i.e. A_n occurs for all but finitely many n.

1.4 Conditional Expectation

Definition 1.25 Conditional Expectation (L^2)

Let $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then the **conditional expectation of X given $\mathcal{G}$**, denoted $\mathbb{E}(X|\mathcal{G})$ is the $\mathcal{G}$-measurable orthogonal projection of X onto $L^2(\mathcal{G})$.

Definition 1.26 Conditional Expectation (L^1)

Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then a random variable Z is the **conditional expectation of X given $\mathcal{G}$**, $Z = \mathbb{E}(X|\mathcal{G})$ if

$$Z = \int_G X d\mathbb{P} = \int_G \mathbb{E}(X|\mathcal{G}) d\mathbb{P}$$

$\forall G \in \mathcal{G}$. Or, equivalently,

$$\mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(X|G)Z)$$

$\forall A \in \Gamma$, where $G = \{\mathbb{1}_G : G \in \mathcal{G}\}$

Theorem 1.22 Properties of Conditional Expectation

Let $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}, \mathcal{D}$ sub-σ-algebras of $\mathcal{F}$. Then the following hold (a.s.)

1. $\mathbb{E}(X|\{\Omega, \varnothing\}) = \mathbb{E}(X)$
2. $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$
3. $\mathbb{E}(cX + dY|\mathcal{G}) = c\mathbb{E}(X|\mathcal{G}) + d\mathbb{E}(Y|\mathcal{G})$ *where* $c, d \in \mathbb{R}$
4. $X \leq Y \Rightarrow \mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G})$
5. $\mathcal{D} \subset \mathcal{G} \Rightarrow \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{D}) = \mathbb{E}(X|\mathcal{D})$
6. $XY \in L^1$ *and* X *is* $\mathcal{G}$*-measurable, then* $\mathbb{E}(XY|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$
7. $\sigma(X)$ *independent of* $\mathcal{G}$*, then* $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$

8. Let ϕ be convex on a non-open interval J with finite left or right hand derivative at an end point of J. If $\mathbb{P}(X \in J) = 1$ and $\phi(X) \in L^1$, then

$$\phi(\mathbb{E}(X|\mathcal{G})) \leq \mathbb{E}(\phi(X)|\mathcal{G})$$

9. $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$, $p \geq 1$, then $\|\mathbb{E}(X|\mathcal{G})\|_p \leq \|X\|_p$
10. a. $X_n \xrightarrow{L^p} X \Rightarrow \mathbb{E}(X_n|\mathcal{G}) \xrightarrow{L^p} \mathbb{E}(X|\mathcal{G})$
    b. $0 \leq X_n \uparrow X$ a.s. $X_n, X \in L^1$, then $\mathbb{E}(X_n|\mathcal{G}) \uparrow \mathbb{E}(X|\mathcal{G})$ and $\mathbb{E}(X_n|\mathcal{G}) \xrightarrow{L^1} \mathbb{E}(X|\mathcal{G})$
    c. If $X_n \to X$ a.s. and $|X_n| \leq Y \in L^1$, then $\mathbb{E}(X_n|\mathcal{G}) \to \mathbb{E}(X|\mathcal{G})$ a.s.
11. Let $U, V : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_1, \mathcal{S}_1), (S_2, \mathcal{S}_2)$ respectively. Let $\phi : (S_1 \times S_2, \mathcal{S}_1 \otimes \mathcal{S}_2) \to \mathbb{R}$ be measurable. If U is $\mathcal{G}$-measurable, $\sigma(V)$ and $\mathcal{G}$ are independent, and $\mathbb{E}(|\phi(U, V)|) < \infty$, then

$$\mathbb{E}(\phi(U, V)|\mathcal{G}) = h(U)$$

    where $h(U) = \mathbb{E}(\phi(u, V))$
12. $\mathbb{E}(X|\sigma(Y, Z)) = \mathbb{E}(X|\sigma(Y))$ if (X, Y) and Z are independent.

1.4.1 Conditional Probability

Definition 1.27

Given $A \in \mathcal{F}$, the **conditional probability of A given $\mathcal{G}$** is

$$\mathbb{P}(A|\mathcal{G}) = \mathbb{E}(\mathbb{1}_A|\mathcal{G})$$

So, by orthogonality,

$$\mathbb{P}(A \cap G) = \int_G \mathbb{P}(A|\mathcal{G})\mathbb{P}(d\omega)$$

$\forall G \in \mathcal{G}$.

Moreover, $0 \leq \mathbb{P}(A|\mathcal{G}) \leq 1$, $\mathbb{P}(\emptyset|\mathcal{G}) = 0$, $\mathbb{P}(\Omega|\mathcal{G}) = 1$, and, given countable $\{A_n\}_{n=1}^{\infty}$,

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n|\mathcal{G}) = \sum_{n=1}^{\infty} \mathbb{P}(A_n|\mathcal{G})$$

Definition 1.28

Let $Y : (\Omega, \mathcal{F}, \mathbb{P}) \to (S, \mathcal{S})$ be a random map and $\mathcal{G}$ be a sub σ-algebra on $\mathcal{F}$. The **regular conditional distribution of Y given $\mathcal{G}$** is a function

$$(\omega, C) \mapsto \mathcal{Q}^{\mathcal{G}}$$

where $Q^{\mathcal{G}}(\omega, C) = \mathbb{P}^{\mathcal{G}}([Y \in C])(\omega)$ on $\Omega \times S$ such that

1. $\forall C \in \mathcal{S}$, $\mathcal{Q}^{\mathcal{G}}(\cdot, C) = \mathbb{P}([Y \in C]|\mathcal{G})$ a.s.
2. $\forall \omega \in \Omega$, $C \mapsto \mathcal{Q}^{\mathcal{G}}$ is a probability measure on $\Omega \times S$.

Definition 1.29

A topological space whose topology is induced by a metric is called **metrizable**. If a metrizable space is complete and seperable, it is called a **Polish Space**.

Theorem 1.23 Disintegration Formula

Given $f : (S, \mathcal{S}) \to \mathbb{R}$ *with* $f \in L^1$,

$$\mathbb{E}[f(Y)|\mathcal{G}] = \int_\Omega \int f(y)\mathcal{Q}^{\mathcal{G}}(\omega, dy)\mathbb{P}(d\omega)$$

Definition 1.30

Given $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$, B_n is a **partition** of $\mathcal{F}$ if B_n is disjoint, countable, and

$$\bigcup_{n\in\mathbb{N}} B_n = \Omega$$

Theorem 1.24

Let $\{B_n\}_{n\in\mathbb{N}}$ *be a paritition of* $\mathcal{F}$ *such that* $\mathbb{P}(B_n) > 0$ *for all* $n = 1, 2, \ldots$ *Let* $\mathcal{G} = \sigma(\{B_n\}_{n\in\mathbb{N}})$. *Then* $\forall A \in \mathcal{F}$,

$$\mathbb{P}(A|\mathcal{G})(\omega) = \frac{\mathbb{P}(A \cap B_n)}{\mathbb{P}(B_n)}$$

if $\omega \in B_n$

Example 1.2 Canonical Probability Space

Let $\Omega = S_1 \times S_2$, $\mathcal{F} = \mathcal{S}_1 \otimes \mathcal{S}_2$ and $\mathbb{P}$ be absolutely continuous with respect to $\mu = \mu_1 \times \mu_2$ and density f. We can view $\mathbb{P}$ as a joint coordinate distribution (X, Y) where $X(\omega) = x$,$Y(\omega) = y$ (i.e. $\omega = (x, y) \in S_1 \times S_2$). If we take the σ-algebra generated by the first coordinate, that is

$$\mathcal{G} = \{B \times S_2 : B \in \mathcal{S}_1\}$$

Then the regular conditional distribution of Y, given $\sigma(X)$ and $C \in \mathcal{S}_2$, is

$$\mathbb{P}([Y \in C]|\mathcal{G})(\omega) = \frac{\int_C f(x, y)\mu_2(dy)}{\int_{S_2} f(x, y)\mu_2(dy')}$$

where $A = S_1 \times C$.

Definition 1.31

The **conditional pdf of** Y **given** $X = x$, denoted $f(y|x)$ is the joint density section $y \mapsto f(x, y)$ normalized to a probability density function by dividing by the marginal pdf $f_X(x) = \int_{S_2} f(x, y)\mu_2(dy)$. This is given in general form by

$$f(y|x) = \frac{f(x, y)}{\int_{S_2} f(x, y)\mu_2(dy)}$$

1.4.2 Random Walks

Definition 1.32 Random Walks

Let $Z_1, Z_2, \dots, Z_n$ be a sequence of i.i.d. random variables. Then we can define a **random walk from** X by

$$S_{k,X} = X + \sum_{i=1}^{n} Z_i$$

where $X \in \mathbb{R}$ and $S_{0,X} = X$.

Theorem 1.25 Markov Property

Given i.i.d. random variables $Z_1, Z_2, \dots, Z_n$ and random walk $S_{k,X}$,

$$\mathbb{E}[S_{n,X} | \sigma(S_{n-1,X}, S_{n-2,X}, \dots, S_{0,X})] = \mathbb{E}[S_{n,X} | \sigma(S_{n-1,X})]$$

Furthermore, note that

$$\mathbb{E}[S_{n,X} | \sigma(S_{n-1,X})] = S_{n,X} + \mathbb{E}[S_{n-1,X}]$$

Definition 1.33 Stochastic Processes

A family of random maps $\{X_t : t \in \Lambda\}$ such that for each $t \in \Lambda$, $X_t : \Omega \to S_t$ is known as a **stochastic process**.

If the index set Λ is $1, 2, 3, \dots$, then $\{X_t : t \in \Lambda\}$ is refered to as a **discrete-parameter stochastic process**. If $\Lambda = [0, \infty)$, then $\{X_t : t \in \Lambda\}$ is known as a **continuous-parameter stochastic process**.

1.5 Martingales

Definition 1.34 Filtration

Let $\{\mathcal{F}_t : t \in T\}$ be a collection of σ-algebras such that T is a linearly ordered set and

$$\mathcal{F}_s \subseteq \mathcal{F}_t$$

for $s \leq t$. Then $\{\mathcal{F}_t : t \in T\}$ is called a **filtration**.

Furthermore, we say a collection of random variables $\{X_t\}$ is **adapted to** $\{\mathcal{F}_t : t \in T\}$ if X_t is $\mathcal{F}_t$ measurable for each $t \in T$

Definition 1.35 (Super/Sub) Martingales

Let $\{X_t\}$ be adapted to $\{\mathcal{F}_t : t \in T\}$. Then $\{X_t\}$ is a:

- **martingale** $\iff X_s = \mathbb{E}[X_t | \mathcal{F}_s]$ where ($s \leq t$)
- **supermartingale** $\iff X_s \geq \mathbb{E}[X_t | \mathcal{F}_s]$ where ($s \leq t$)
- **submartingale** $\iff X_s \leq \mathbb{E}[X_t | \mathcal{F}_s]$ where ($s \leq t$)

Exercise 1.1

Let $\{X_t\}_{t\in\mathbb{N}}$ be a sequence of iid random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Additionally, let $\{S_t\}_{t\in\mathbb{N}}$ be a sequence of random variables defined by

$$S_0 = X_1$$
$$S_t = S_{t+1} + X_t$$

Take the filtration $\mathcal{F}_t = \sigma(X_1, X_2, \ldots, X_t)$ and note that S_t is $\mathcal{F}_t$ adapted.
Is S_t a martingale, supermartingale, or submartingale? What is $\mathbb{E}[S_t + 1|\mathcal{F}_t]$?

Solution

By properties of conditional expectation we can get that

$$\mathbb{E}[S_{t+1}|\mathcal{F}_t] = S_t + \mathbb{E}[X_{t+1}]$$

So we can say that:

$$\text{if } \mathbb{E}[X_t + 1] = 0 \Rightarrow S_t = \mathbb{E}[S_t + 1|\mathcal{F}_t] \ \textbf{(martingale)}$$
$$\text{if } \mathbb{E}[X_t + 1] > 0 \Rightarrow S_t < \mathbb{E}[S_t + 1|\mathcal{F}_t] \ \textbf{(submartingale)}$$
$$\text{if } \mathbb{E}[X_t + 1] < 0 \Rightarrow S_t > \mathbb{E}[S_t + 1|\mathcal{F}_t] \ \textbf{(supermartingale)}$$

Definition 1.36

Let $\{X_n : n \geq 1\}$ be a martingale. Its **martingale difference sequence** is given by $\{Z_n\}$, where

$$Z_1 = X_1$$
$$Z_{k+1} = X_{k+1} - X_k$$

☞ Remark

If $X_n \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, $\forall n \geq 1$, $\{Z_n\}$ are uncorrelated. Moreover, if $X_n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and f is a bounded, $\mathcal{F}_n$ measurable function, then

$$\begin{aligned}\mathbb{E}[Z_{n+1}f(X_1, X_2, \ldots, X_n)] &= \mathbb{E}[\mathbb{E}[Z_{n+1}f(X_1, X_2, \ldots, X_n)|\mathcal{F}_n]] \\ &= \mathbb{E}[f(X_1, X_2, \ldots, X_n)\mathbb{E}(Z_{n+1}|\mathcal{F}_n)] = 0\end{aligned}$$

Theorem 1.26

(a) *If $\{X_n : n \geq 1\}$ is a martingale and $\phi(X_n)$ is a convex, integrable function, then $\{\phi(X_n) : n \geq 1\}$ is a submartingale.*

(b) *If $\{X_n\}$ is a submartingale and $\phi(X_n)$ is convex, non-decreasing, and integrable, then $\{\phi(X_n) : n \geq 1\}$ is a submartingale.*

Theorem 1.27 Doob's Maximal Inequality

Let $p \geq 1$ and $\{X_n\}$ be an $\{\mathcal{F}_n : 1 \leq k \leq n\}$ adapted martingale (or a non-negative submartingale) such that $\mathbb{E}[|X_n|^p] < \infty$. Then for $\lambda > 0$ and $M_n := \max\{|X_n|\}$ satisfies the following:

$$\mathbb{P}(M_n \geq \lambda) \leq \frac{1}{\lambda^p} \int_{[M_n > \lambda]} |X_n|^p d\mathbb{P} \leq \frac{1}{\lambda^p} \mathbb{E}[|X_n|^p]$$

Corollary 1.2 Kolmogorov's Inequality

Let $\{X_n\}$ be a martingale with $\mathbb{E}[X_k] = 0$ and $Var(X_k) < \infty$ for all $k = 1, 2, \dots, n$. Then for $M_n = \max\{X_n\}, \lambda > 0$,

$$\mathbb{P}(M_n > \lambda) \leq \frac{1}{\lambda^2} Var(S_n) = \frac{1}{\lambda^2} \sum_{k=1}^{n} Var(X_k) = \frac{1}{\lambda^2} \sum_{k=1}^{n} \mathbb{E}[|X_n|^2]$$

Theorem 1.28 Doobs Maximal Ineq. for Moments

(a) $\mathbb{E}[M_n] \leq \frac{e}{e-1}(1 + \mathbb{E}[|X_n|] \log(|X_n|))$

(b) *For $p > 1$, $\mathbb{E}[M_n^p] \leq q^p \mathbb{E}[|X_n|^p]$ where $\frac{1}{p} + \frac{1}{q} = 1$*

1.5.1 Stopping Times

Definition 1.37

Let $\{F_t\}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$. A random variable $\tau : \Omega \to T \cup \{\infty\}$ is a **stopping time** if

$$\{\tau \leq t\} = \{\omega \in \Omega : \tau(\omega) \leq t\}$$

. Furthermore, τ is an **optional stopping time** if

$$\{\tau < t\} = \{\omega \in \Omega : \tau(\omega) < t\} \in \mathcal{F}_t$$

$\forall t \in T$.

Exercise 1.2

Let $T = \mathbb{N}$ and $(S_i, \mathcal{S}_i, \mathbb{P}_i)$ be a series of Bernoulli probability spaces, where $S_i = \{\pm 1\}, \mathcal{S}_i = 2^{S_i}$, and $\mathbb{P}_i(\omega = +1) = 1 - \mathbb{P}(\omega = -1) = p$. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined as follows:

$$\Omega = \{\omega = (\omega_i)_{i \in \mathbb{N}} : \omega_i = \pm 1\} = S_1 \times S_2 \times \dots = \prod_{i=1}^{\infty} S_i$$

$$\mathcal{F} = \bigotimes_{i=1}^{\infty} \mathcal{S}_i$$

$$\mathbb{P} = \prod_{i=1}^{\infty} \mathbb{P}_i$$

Finally, define a set of random variables $X_i : (\omega) \to \omega_i$ (i.e. X_i is the value of ω (± 1) at iteration i. **What are $|\sigma(x_1)|$? $|\sigma(X_1, X_2)|$? $|\mathcal{F}|$?**

Solution

$$\sigma(X_1) = \{X_1^{-1}(+1), X_1^{-1}(-1), X_1^{-1}(\emptyset), X_1^{-1}(\pm 1)\} \Rightarrow |\sigma(X_1)| = 4$$
$$|\sigma(X_1, X_2)| = 16$$
$$|\mathcal{F}| = \infty$$

Definition 1.38 Pre-τ σ-algebra

Suppose $\{\mathcal{F}_t : t \in T\}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ and τ is an $\mathcal{F}_t$ stopping time. The **Pre-τ σ-algebra** $\mathcal{F}_\tau$ is defined as

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap [\tau \leq t] \in \mathcal{F}_t, \quad \forall t \in T\}$$

☞ **Remark**
If $\tau_1 \leq \tau_2$ are stopping times, then $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$

Definition 1.39 stopped process

The stochastic process $\{X_\tau \wedge t : t \geq 0\}$ is referred to as the **stopped process**, where

$$a \wedge b = \min\{a, b\}$$

Definition 1.40 progressively measurable

Let $T = [0, \infty)$ or $T = [0, t_0]$. A stochastic process $\{X_t : t \in T\}$ on $(S, \mathcal{S})$ is **progressively measurable** with respect to $\{\mathcal{F}_t\}$ is $(s, \omega) \mapsto x_s(\omega)$ from $[0, t_0] \times \Omega$ to S is measurable with respect to $\mathbb{B}[0, t_0] \otimes \mathcal{F}_t$ and $\mathcal{S}$ $\forall t \in T$.

Theorem 1.29

Suppose $\{X_t : t \in T\}$ is progressively measurable and τ is a stopping time. Then X_τ is $\mathcal{F}_\tau$ measurable. That is,

$$([X_\tau \in B] \cap [\tau \leq t]) \in \mathcal{F}_\tau \quad \forall B \in \mathcal{S}, \ \forall t \in T$$

Theorem 1.30

Let Y_n be an $\mathcal{F}_n$ martingale such that $Y_n \to Y \in L^1$. Then $Y_n = \mathbb{E}(Y|\mathcal{F}_n)$

Theorem 1.31

Let $T = \mathbb{N}$. Assume $\tau_1 \leq \tau_2$. Then

$$\mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_1}) = X_{\tau_1}$$

Corollary 1.3

Let $T = \mathbb{N}$, fix $n \in \mathbb{N}$ and assume X_t is an $\mathcal{F}_\tau$ measurable martingale. For any stopping time $\tau_1 \leq \tau_2$,

$$\mathbb{E}(X_{\tau_2 \wedge n}|\mathcal{F}_{\tau \wedge n}) = X_{\tau_1 \wedge n}$$

Theorem 1.32 Optional Stopping Time Theorem

Let X_t be an $\mathcal{F}_t$ martingale with $t \in T = \mathbb{N}$. Suppose $\tau_1 \leq \tau_2$ are stopping times such that:

1. $\mathbb{P}(\tau_2 < \infty) = 1$
2. $X_{\tau_2 \wedge n}$ *is uniformly integrable*

Then $\mathbb{E}(X_{\tau_2} | \mathcal{F}_{\tau_1}) = X_{\tau_1}$

Definition 1.41 Upcrossings

Let Z_n be $\mathcal{F}_n$ submartingale with $n \in \mathbb{N}$,

$$\mathbb{E}(Z_n | \mathcal{F}_m) \geq Z_m$$

Fix $a \leq b$ and define $\zeta_1 = 1$ and

$$\begin{cases} \zeta_{2k} = \inf\{n \geq 2k-1 | Z_n \geq b\} \\ \zeta_{2k+1} = \inf\{n \geq 2k | Z_n \leq a\} \end{cases}, \qquad k = 1, 2, \ldots$$

Now define $X_n = \max\{Z_n - a, 0\}$ and note that $\max$ is a convex function, so X_n is a submartingale by Jensen's Inequality. Also,

$$\begin{cases} X_{\zeta_{2k}} = \max\{Z_{\zeta_{2k}} - a, 0\} \geq b - a \\ X_{\zeta_{2k+1}} = \max\{Z_{\zeta_{2k+1}} - a, 0\} = 0 \end{cases}$$

The number of **Upcrossings** by time N is then given by $U_N = \sup\{k | \zeta_{2k} \leq N\}$.

Theorem 1.33 Doob's Upcrossing Inequality

$$\mathbb{E}(U_N) \leq \frac{\mathbb{E}(X_N) - \mathbb{E}(X_1)}{(b-a)} \leq \frac{\mathbb{E}(X_N)}{(b-a)}$$

1.6 Central Limit Theorem

Definition 1.42 Weak Convergence

A sequence of probabilities $\{Q_n\}_{n=1}^{\infty}$ **converges weakly** or, equivalently, **converges in probability** to a probability Q as $n \to \infty$ if

$$\lim_{n \to \infty} \int_{\mathbb{R}^k} g(x) Q_n(dx) = \int_{\mathbb{R}^k} g(x) Q(dx)$$

for all bounded functions $g : \mathbb{R}^k \to \mathbb{R}$. We denote converges in probability as $Q_n \Rightarrow Q$. Moreover, a sequence of random variables X_n with distributions Q_n converges to X with a distribution Q if $Q_n \Rightarrow Q$.

Theorem 1.34 Finite Dimensional Weak Convergence

Let $\{Q_n\}, Q$ be a sequence of probabilities. The following are equivalent:

1. $Q_n \Rightarrow Q$

2. $\int_{\mathbb{R}^k} f dQ_n \to \int_{\mathbb{R}^k} f dQ$ for all bounded, continuous f vanishing outside a compact set.
3. $\int_{\mathbb{R}^k} f dQ_n \to \int_{\mathbb{R}^k} f dQ$ for all infintely differentiable f vanishing outside a compact set.
4. For $F_n(x) = Q_n((-\infty, x_1] \times \ldots \times (-\infty, x_n])$ and $F(x) = Q((-\infty, x_1] \times \ldots \times (-\infty, x_n])$, $F_n(x) \to F(x)$ as $n \to \infty$

Theorem 1.35 Lindeberg Central Limit Theorem

For each n, let $X_{n,1}, \ldots, X_{n,k_n}$ be independent arrays of random variables such that $\mathbb{E}(X_{n,j}) = 0$, $\sigma_{n,j} = \left(\mathbb{E}(X_{n,j}^2)\right)^{1/2} < \infty$, $\sum_{j=1}^{k_n} \sigma_{n,j}^2 = 1$ and, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \mathbb{E}(X_{n,j}^2 \mathbb{1}_{[|X_{n,j}|>\epsilon]}) = 0 \qquad \text{(Lindeberg Condition)}$$

Then $\sum_{j=1}^{k_n} X_{n,j}$ converges in distribution to a standard normal distribution, $\mathcal{N}(0, 1)$.

Corollary 1.4 Classical Central Limit Theorem

Let $\{X_j\}$ be a sequence of random variables with $\mathbb{E}(X_j) = \mu$, $0 < \sigma^2 = VarX_j < \infty$. Then $\frac{\sum_{j=1}^{n}(X_j - \mu)}{(\sigma\sqrt{n})} \Rightarrow \mathcal{N}(0, 1)$. Equivalently, $\sum_{j=1}^{k_n} X_{n,j} \Rightarrow \mathcal{N}(n\mu, \sigma\sqrt{n})$.

Corollary 1.5 Lyapounov Central Limit Theorem

$\forall n$, let $X_{1,n}, X_{2,n}, \ldots, X_{n,k_n}$ be k_n independent random variables such that $\sum_{j=1}^{k_n} \mathbb{E}X_{n,j} = \mu$, $\sum_{j=1}^{k_n} VarX_{n,j} = \sigma^2 > 0$ and

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \mathbb{E}|X_{n,j} - \mathbb{E}X_{n,j}|^{2+\delta} = 0 \qquad \text{(Lyapounov Condition)}$$

for some $\delta > 0$. Then $\sum_{j=1}^{k_n} X_{n,j} \Rightarrow \mathcal{N}(\mu, \sigma^2)$.