An Introduction to Stochastic Calculus

Haijun Li

lih@math.wsu.edu
Department of Mathematics
Washington State University

Week 1
Outline

1. Basic Concepts from Probability Theory
   - Notations
   - Random Vectors

2. Stochastic Processes
   - Basic Definition
   - Distributional Properties
   - Dependence Structure
Sample or outcome space \( \Omega := \{ \text{all possible outcomes } \omega \text{ of the underlying experiment} \} \).
Notations

- Sample or outcome space $\Omega := \{\text{all possible outcomes } \omega \text{ of the underlying experiment}\}$.
- $\sigma$-field or $\sigma$-algebra $\mathcal{F}$: A non-empty class of subsets (or observable events) of $\Omega$ closed under countable union, countable intersection and complements.
**Notations**

- **Sample or outcome space** $\Omega := \{\text{all possible outcomes } \omega \text{ of the underlying experiment}\}$.

- **$\sigma$-field or $\sigma$-algebra** $\mathcal{F}$: A non-empty class of subsets (or observable events) of $\Omega$ closed under countable union, countable intersection and complements.

- **Probability measure** $P(\cdot)$ on $\mathcal{F}$: $P(A)$ denotes the probability of event $A$. 

**Haijun Li**

*An Introduction to Stochastic Calculus*
Notations

- Sample or outcome space $\Omega := \{\text{all possible outcomes } \omega \text{ of the underlying experiment}\}$.
- $\sigma$-field or $\sigma$-algebra $\mathcal{F}$: A non-empty class of subsets (or observable events) of $\Omega$ closed under countable union, countable intersection and complements.
- Probability measure $P(\cdot)$ on $\mathcal{F}$: $P(A)$ denotes the probability of event $A$.
- Random variable $X : \Omega \mapsto \mathbb{R}$ is a real-valued measurable function defined on $\Omega$. That is, events $X^{-1}(a, b) \in \mathcal{F}$ are observable for all $a, b \in \mathbb{R}$. 

Induced probability measure $P_X(B) := P(X \in B) = P(\{\omega : X(\omega) \in B\})$, for any Borel set $B \subseteq \mathbb{R}$.

Distribution function $F_X(x) := P(X \leq x)$, $x \in \mathbb{R}$. 

Haijun Li  
An Introduction to Stochastic Calculus  
Week 1 3 / 16
Notations

- Sample or outcome space $\Omega := \{\text{all possible outcomes } \omega \text{ of the underlying experiment}\}$.

- $\sigma$-field or $\sigma$-algebra $\mathcal{F}$: A non-empty class of subsets (or observable events) of $\Omega$ closed under countable union, countable intersection and complements.

- Probability measure $P(\cdot)$ on $\mathcal{F}$: $P(A)$ denotes the probability of event $A$.

- Random variable $X : \Omega \mapsto \mathbb{R}$ is a real-valued measurable function defined on $\Omega$. That is, events $X^{-1}(a, b) \in \mathcal{F}$ are observable for all $a, b \in \mathbb{R}$.

- Induced probability measure $P_X(B) := P(X \in B) = P(\{\omega : X(\omega) \in B\})$, for any Borel set $B \subseteq \mathbb{R}$.
Sample or outcome space $\Omega := \{\text{all possible outcomes } \omega \text{ of the underlying experiment}\}$.

$\sigma$-field or $\sigma$-algebra $\mathcal{F}$: A non-empty class of subsets (or observable events) of $\Omega$ closed under countable union, countable intersection and complements.

Probability measure $P(\cdot)$ on $\mathcal{F}$: $P(A)$ denotes the probability of event $A$.

Random variable $X : \Omega \mapsto \mathbb{R}$ is a real-valued measurable function defined on $\Omega$. That is, events $X^{-1}(a, b) \in \mathcal{F}$ are observable for all $a, b \in \mathbb{R}$.

Induced probability measure $P_X(B) := P(X \in B) = P(\{\omega : X(\omega) \in B\})$, for any Borel set $B \subseteq \mathbb{R}$.

Distribution function $F_X(x) := P(X \leq x), x \in \mathbb{R}$. 

Haijun Li
An Introduction to Stochastic Calculus
Week 1
Random variable $X$ is said to be \textit{continuous} if the distribution function $F_X$ has no jumps, that is,

$$\lim_{h \to 0} F_X(x + h) = F_X(x), \forall x \in \mathbb{R}.$$
Continuous and Discrete Random Variables

- Random variable $X$ is said to be **continuous** if the distribution function $F_X$ has no jumps, that is,

$$\lim_{h \to 0} F_X(x + h) = F_X(x), \ \forall x \in \mathbb{R}.$$ 

Most continuous distributions of interest have a density $f_X \geq 0$:

$$F_X(x) = \int_{-\infty}^{x} f_X(y) dy, \ x \in \mathbb{R}$$

where $\int_{-\infty}^{\infty} f_X(y) dy = 1$. 

Continuous and Discrete Random Variables

- Random variable $X$ is said to be continuous if the distribution function $F_X$ has no jumps, that is,

$$\lim_{h \to 0} F_X(x + h) = F_X(x), \ \forall x \in \mathbb{R}.$$ 

Most continuous distributions of interest have a density $f_X \geq 0$:

$$F_X(x) = \int_{-\infty}^{x} f_X(y) dy, \ x \in \mathbb{R}$$

where $\int_{-\infty}^{\infty} f_X(y) dy = 1$.

- Random variable $X$ is said to be discrete if the distribution function $F_X$ is a pure jump function:

$$F_X(x) = \sum_{k : x_k \leq x} p_k, \ x \in \mathbb{R}$$

where the probability mass function $\{p_k\}$ satisfies that $1 \geq p_k \geq 0$ and $\sum_{k=1}^{\infty} p_k = 1$. 
Expectation, Variance and Moments

A General Formula
For a real-valued function $g$, the expectation of $g(X)$ is given by $Eg(X) = \int g(x)dF_X(x)$. The $k$-th moment of $X$ is given by $E(X^k) = \int x^k dF_X(x)$. The mean $\mu_X$ (or “center of gravity”) of $X$ is the first moment. The variance (or “spread out”) of $X$ is defined as $\sigma_X^2 = \text{var}(X) := E(X - \mu_X)^2$. Clearly $\sigma_X^2 = E(X^2) - \mu_X^2$. If the variance exists, then the Chebyshev inequality holds: $P(|X - \mu_X| > k\sigma_X) \leq \frac{1}{k^2}$, $k > 0$. That is, the probability of tail regions that are $k$ standard deviations away from the mean is bounded by $1/k^2$. 
A General Formula

For a real-valued function \( g \), the expectation of \( g(X) \) is given by

\[
Eg(X) = \int g(x) dF_X(x).
\]

- The \( k \)-th moment of \( X \) is given by
  \[
  E(X^k) = \int x^k dF_X(x).
  \]
  The mean \( \mu_X \) (or “center of gravity”) of \( X \) is the first moment.

The variance (or “spread out”) of \( X \) is defined as

\[
\sigma^2_X = \text{var}(X) := E(X - \mu_X)^2.
\]

Clearly

\[
\sigma^2_X = E(X^2) - \mu^2_X.
\]

If the variance exists, then the Chebyshev inequality holds:

\[
P(|X - \mu_X| > k \sigma_X) \leq \frac{1}{k^2}, \quad k > 0.
\]

That is, the probability of tail regions that are \( k \) standard deviations away from the mean is bounded by \( 1/k^2 \).
A General Formula

For a real-valued function $g$, the expectation of $g(X)$ is given by

$$Eg(X) = \int g(x) dF_X(x).$$

- The $k$-th moment of $X$ is given by $E(X^k) = \int x^k dF_X(x)$. The mean $\mu_X$ (or “center of gravity”) of $X$ is the first moment.
- The variance (or “spread out”) of $X$ is defined as $\sigma^2_X = \text{var}(X) := E(X - \mu_X)^2$. Clearly $\sigma^2_X = E(X^2) - \mu^2_X$. 
Expectation, Variance and Moments

A General Formula

For a real-valued function $g$, the expectation of $g(X)$ is given by
$$Eg(X) = \int g(x) dF_X(x).$$

- The $k$-th moment of $X$ is given by $E(X^k) = \int x^k dF_X(x)$. The mean $\mu_X$ (or “center of gravity”) of $X$ is the first moment.
- The variance (or “spread out”) of $X$ is defined as $\sigma^2_X = \text{var}(X) := E(X - \mu_X)^2$. Clearly $\sigma^2_X = E(X^2) - \mu^2_X$.
- If the variance exists, then the Chebyshev inequality holds:
  $$P(|X - \mu_X| > k\sigma_X) \leq k^{-2}, \quad k > 0.$$  
  That is, the probability of tail regions that are $k$ standard deviations away from the mean is bounded by $1/k^2$. 
Random Vectors

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

- $\mathbf{X} = (X_1, \ldots, X_d) : \Omega \mapsto \mathbb{R}^d$ denotes a $d$-dimensional random vector, where its components $X_1, \ldots, X_d$ are real-valued random variables.
Random Vectors

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

- $\mathbf{X} = (X_1, \ldots, X_d) : \Omega \mapsto \mathbb{R}^d$ denotes a $d$-dimensional random vector, where its components $X_1, \ldots, X_d$ are real-valued random variables.

- The induced probability measure: $P_{\mathbf{X}}(B) = P(\mathbf{X} \in B)$
  $:= P(\{\omega : \mathbf{X}(\omega) \in B\})$ for all Borel subsets $B$ of $\mathbb{R}^d$. 

Haijun Li  
An Introduction to Stochastic Calculus  
Week 1
Random Vectors

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

- $X = (X_1, \ldots, X_d) : \Omega \mapsto \mathbb{R}^d$ denotes a $d$-dimensional random vector, where its components $X_1, \ldots, X_d$ are real-valued random variables.

- The induced probability measure: $P_X(B) = P(X \in B) \ := P(\{\omega : X(\omega) \in B\})$ for all Borel subsets $B$ of $\mathbb{R}^d$.

- The distribution function $F_X(x) := P(X_1 \leq x_1, \ldots, X_d \leq x_d)$, $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. 

Random Vectors

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

- $\bm{X} = (X_1, \ldots, X_d) : \Omega \mapsto \mathbb{R}^d$ denotes a $d$-dimensional random vector, where its components $X_1, \ldots, X_d$ are real-valued random variables.

- The induced probability measure: $P_\bm{X}(B) = P(\bm{X} \in B) := P(\{\omega : \bm{X}(\omega) \in B\})$ for all Borel subsets $B$ of $\mathbb{R}^d$.

- The distribution function $F_\bm{X}(\bm{x}) := P(X_1 \leq x_1, \ldots, X_d \leq x_d)$, $\bm{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

- If $\bm{X}$ has a density $f_\bm{X} \geq 0$, then

$$F_\bm{X}(\bm{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f_\bm{X}(\bm{x}) \, d\bm{x}$$

with $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_\bm{X}(\bm{x}) \, d\bm{x} = 1$. 
Random Vectors

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

- $X = (X_1, \ldots, X_d) : \Omega \mapsto \mathbb{R}^d$ denotes a $d$-dimensional random vector, where its components $X_1, \ldots, X_d$ are real-valued random variables.

- The induced probability measure: $P_X(B) = P(X \in B) := P(\{\omega : X(\omega) \in B\})$ for all Borel subsets $B$ of $\mathbb{R}^d$.

- The distribution function $F_X(x) := P(X_1 \leq x_1, \ldots, X_d \leq x_d)$, $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

- If $X$ has a density $f_X \geq 0$, then
  \[
  F_X(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f_X(x) \, dx
  \]
  with $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x) \, dx = 1$.

- For any $J \subseteq \{1, \ldots, d\}$, let $X_J := (X_j; j \in J)$ be the $J$-margin of $X$. The marginal density of $X_J$ is given by
  \[
  f_{X_J}(x_J) = \int f_X(x) \, dx_{J^c}.
  \]
The expectation or mean value of $X$ is denoted by

$$\mu_X = EX := (E(X_1), \ldots, E(X_d)).$$
- The expectation or mean value of $X$ is denoted by 
  \[ \mu_X = E_X := (E(X_1), \ldots, E(X_d)) \].

- The covariance matrix of $X$ is defined as 
  \[ \Sigma_X := (\text{cov}(X_i, X_j); i, j = 1, \ldots, d) \]
  where the covariance of $X_i$ and $X_j$ is defined as 
  \[ \text{cov}(X_i, X_j) := E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})] = E(X_i X_j) - \mu_{X_i} \mu_{X_j}. \]
Expectation, Variance, and Covariance

- The expectation or mean value of $X$ is denoted by
  $$\mu_X = E\mathbf{X} := (E(X_1), \ldots, E(X_d)).$$

- The covariance matrix of $X$ is defined as
  $$\Sigma_X := (\text{cov}(X_i, X_j); i, j = 1, \ldots, d)$$
  where the covariance of $X_i$ and $X_j$ is defined as
  $$\text{cov}(X_i, X_j) := E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})] = E(X_iX_j) - \mu_{X_i}\mu_{X_j}.$$ 

- The correlation of $X_i$ and $X_j$ is denoted by
  $$\text{corr}(X_i, X_j) := \frac{\text{cov}(X_i, X_j)}{\sigma_{X_i}\sigma_{X_j}}.$$ 

It follows from the Cauchy-Schwarz inequality that
$$-1 \leq \text{corr}(X_i, X_j) \leq 1.$$
Independence and Dependence

- The events $A_1, \ldots, A_n$ are independent if for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P(\cap_{j=1}^k A_{i_j}) = \prod_{j=1}^k P(A_{i_j}).$$
Indepedence and Dependence

- The events $A_1, \ldots, A_n$ are independent if for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

  \[ P(\cap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}). \]

- The random variables $X_1, \ldots, X_n$ are independent if for any Borel sets $B_1, \ldots, B_n$, the events $\{ X_1 \in B_1 \}, \ldots, \{ X_n \in B_n \}$ are independent.
Independence and Dependence

The events $A_1, \ldots, A_n$ are independent if for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P(\cap_{j=1}^k A_{i_j}) = \prod_{j=1}^k P(A_{i_j}).$$

The random variables $X_1, \ldots, X_n$ are independent if for any Borel sets $B_1, \ldots, B_n$, the events $\{X_1 \in B_1\}, \ldots, \{X_n \in B_n\}$ are independent.

The random variables $X_1, \ldots, X_n$ are **independent** if and only if $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^n F_{X_i}(x_i)$, for all $(x_1,\ldots,x_n) \in \mathbb{R}^n$. 
Independence and Dependence

The events $A_1, \ldots, A_n$ are independent if for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P(\bigcap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}).$$

The random variables $X_1, \ldots, X_n$ are independent if for any Borel sets $B_1, \ldots, B_n$, the events $\{X_1 \in B_1\}, \ldots, \{X_n \in B_n\}$ are independent.

The random variables $X_1, \ldots, X_n$ are independent if and only if $F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$, for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

The random variables $X_1, \ldots, X_n$ are independent if and only if $E[\prod_{i=1}^{n} g_i(X_i)] = \prod_{i=1}^{n} E g_i(X_i)$ for any real-valued functions $g_1, \ldots, g_n$.
Independence and Dependence

The events $A_1, \ldots, A_n$ are independent if for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P(\bigcap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}).$$

The random variables $X_1, \ldots, X_n$ are independent if for any Borel sets $B_1, \ldots, B_n$, the events $\{X_1 \in B_1\}, \ldots, \{X_n \in B_n\}$ are independent.

The random variables $X_1, \ldots, X_n$ are independent if and only if

$$F_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i), \text{ for all } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

The random variables $X_1, \ldots, X_n$ are independent if and only if

$$E[\prod_{i=1}^{n} g_i(X_i)] = \prod_{i=1}^{n} Eg_i(X_i) \text{ for any real-valued functions } g_1, \ldots, g_n.$$

In the continuous case, the random variables $X_1, \ldots, X_n$ are independent if and only if

$$f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i), \text{ for all } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$
Two Examples

Let $\mathbf{X} = (X_1, \ldots, X_d)$ have a $d$-dimensional Gaussian distribution. The random variables $X_1, \ldots, X_d$ are independent if and only if $\text{corr}(X_i, X_j) = 0$ for $i \neq j$. 
Two Examples

Let \( X = (X_1, \ldots, X_d) \) have a \( d \)-dimensional Gaussian distribution. The random variables \( X_1, \ldots, X_d \) are independent if and only if \( \text{corr}(X_i, X_j) = 0 \) for \( i \neq j \).

For non-Gaussian random vectors, however, independence and uncorrelatedness are not equivalent. Let \( X \) be a standard normal random variable. Since both \( X \) and \( X^3 \) have expectation zero, \( X \) and \( X^2 \) are uncorrelated:

\[
\text{cov}(X, X^2) = E(X^3) - E(X)E(X^2) = 0.
\]

But \( X \) and \( X^2 \) are clearly dependent (co-monotone). Since \( \{X \in [-1, 1]\} = \{X^2 \in [0, 1]\} \), we obtain

\[
P(X \in [-1, 1], X^2 \in [0, 1]) = P(X \in [-1, 1])
\]

\[
> [P(X \in [-1, 1])]^2 = P(X \in [-1, 1])P(X^2 \in [0, 1]).
\]
Autocorrelations

For a time series $X_0, X_1, X_2, \ldots$ the autocorrelation at lag $h$ is defined by $\text{corr}(X_0, X_h)$, $h = 0, 1, \ldots$. 

Log-returns $X_t := \log S_t - S_{t-1}$, where $S_t$ is the price of a speculative asset (equities, indexes, exchange rates and commodity) at the end of the $t$-th period. If the relative returns are small, then $X_t \approx S_t - S_{t-1} / S_{t-1}$. Note that the log-returns are scale-free, additive, stationary, ...

Stylized Fact #1: Log-returns $X_t$ are not iid (independent and identically distributed) although they show little serial autocorrelation.

Stylized Fact #2: Series of absolute $|X_t|$ or squared $X_t^2$ returns show profound serial autocorrelation (long-range dependence).
Autocorrelations

- For a time series $X_0, X_1, X_2, \ldots$ the autocorrelation at lag $h$ is defined by $\text{corr}(X_0, X_h), h = 0, 1, \ldots$.

- Log-returns $X_t := \log \frac{S_t}{S_{t-1}}$, where $S_t$ is the price of a speculative asset (equities, indexes, exchange rates and commodity) at the end of the $t$-th period. If the relative returns are small, then $X_t \approx \frac{S_t - S_{t-1}}{S_{t-1}}$. Note that the log-returns are scale-free, additive, stationary, ....
Autocorrelations

- For a time series $X_0, X_1, X_2, \ldots$ the autocorrelation at lag $h$ is defined by $\text{corr}(X_0, X_h)$, $h = 0, 1, \ldots$.

- Log-returns $X_t := \log \frac{S_t}{S_{t-1}}$, where $S_t$ is the price of a speculative asset (equities, indexes, exchange rates and commodity) at the end of the $t$-th period. If the relative returns are small, then $X_t \approx \frac{S_t - S_{t-1}}{S_{t-1}}$. Note that the log-returns are scale-free, additive, stationary, ....

- **Stylized Fact #1:** Log-returns $X_t$ are not iid (independent and identically distributed) although they show little serial autocorrelation.
For a time series $X_0, X_1, X_2, \ldots$ the autocorrelation at lag $h$ is defined by $\text{corr}(X_0, X_h)$, $h = 0, 1, \ldots$.

Log-returns $X_t := \log \frac{S_t}{S_{t-1}}$, where $S_t$ is the price of a speculative asset (equities, indexes, exchange rates and commodity) at the end of the $t$-th period. If the relative returns are small, then $X_t \approx \frac{S_t - S_{t-1}}{S_{t-1}}$. Note that the log-returns are scale-free, additive, stationary, ....

**Stylized Fact #1:** Log-returns $X_t$ are not iid (independent and identically distributed) although they show little serial autocorrelation.

**Stylized Fact #2:** Series of absolute $|X_t|$ or squared $X_t^2$ returns show profound serial autocorrelation (long-range dependence).
A stochastic process $X := (X_t, t \in T)$ is a collection of random variables defined on some space $\Omega$, where $T \subseteq \mathbb{R}$.
A stochastic process $X := (X_t, t \in T)$ is a collection of random variables defined on some space $\Omega$, where $T \subseteq \mathbb{R}$.

If index set $T$ is a finite or countably infinite set, $X$ is said to be a discrete-time process. If $T$ is an interval, then $X$ is a continuous-time process.
A stochastic process $X := (X_t, t \in T)$ is a collection of random variables defined on some space $\Omega$, where $T \subseteq \mathbb{R}$.

If index set $T$ is a finite or countably infinite set, $X$ is said to be a discrete-time process. If $T$ is an interval, then $X$ is a continuous-time process.

A stochastic process $X$ is a (measurable) function of two variables: time $t$ and sample point $\omega$.

Fix time $t$, $X_t = X_t(\omega), \omega \in \Omega$, is a random variable.

Fix sample point $\omega$, $X_t = X_t(\omega), t \in T$, is a sample path.
A stochastic process \( X := (X_t, t \in T) \) is a collection of random variables defined on some space \( \Omega \), where \( T \subseteq \mathbb{R} \).

If index set \( T \) is a finite or countably infinite set, \( X \) is said to be a discrete-time process. If \( T \) is an interval, then \( X \) is a continuous-time process.

A stochastic process \( X \) is a (measurable) function of two variables: time \( t \) and sample point \( \omega \).

Fix time \( t \), \( X_t = X_t(\omega), \omega \in \Omega \), is a random variable.

Fix sample point \( \omega \), \( X_t = X_t(\omega), t \in T \), is a sample path.

**Example:** An autoregressive process of order 1 is given by

\[
X_t = \phi X_{t-1} + Z_t, \quad t \in \mathbb{Z},
\]

where \( \phi \) is a real parameter. Time series models can be understood as discretization of stochastic differential equations.
Finite-Dimensional Distributions

- All possible values of a stochastic process $X = (X_t, t \in T)$ constitute a function space of all sample paths $(X_t(\omega), t \in T), \forall \omega \in \Omega$. 

Example: A stochastic process is called Gaussian if all its finite-dimensional distributions are multivariate Gaussian. The distribution of this process is determined by the collection of the mean vectors and covariance matrices.
Finite-Dimensional Distributions

- All possible values of a stochastic process $X = (X_t, t \in T)$ constitute a function space of all sample paths $(X_t(\omega), t \in T), \forall \omega \in \Omega$.

- Specifying the distribution of $X$ on this function space is equivalent to specifying which information is available in terms of the observable events from the $\sigma$-field generated by $X$.

Example: A stochastic process is called Gaussian if all its finite-dimensional distributions are multivariate Gaussian. The distribution of this process is determined by the collection of the mean vectors and covariance matrices.
Finite-Dimensional Distributions

- All possible values of a stochastic process $X = (X_t, t \in T)$ constitute a function space of all sample paths $(X_t(\omega), t \in T), \forall \omega \in \Omega$.

- Specifying the distribution of $X$ on this function space is equivalent to specifying which information is available in terms of the observable events from the $\sigma$-field generated by $X$.

- The distribution of $X$ can be described by the distributions of the finite-dimensional vectors

  $$(X_{t_1}, \ldots, X_{t_n}), \text{ for all possible choices of times } t_1, \ldots, t_n \in T.$$
Finite-Dimensional Distributions

- All possible values of a stochastic process $X = (X_t, t \in T)$ constitute a function space of all sample paths $(X_t(\omega), t \in T), \forall \omega \in \Omega$.

- Specifying the distribution of $X$ on this function space is equivalent to specifying which information is available in terms of the observable events from the $\sigma$-field generated by $X$.

- The distribution of $X$ can be described by the distributions of the finite-dimensional vectors

  $$(X_{t_1}, \ldots, X_{t_n}), \text{ for all possible choices of times } t_1, \ldots, t_n \in T.$$  

- **Example:** A stochastic process is called Gaussian if all its finite-dimensional distributions are multivariate Gaussian. The distribution of this process is determined by the collection of the mean vectors and covariance matrices.
The expectation function of a process $X = (X_t, t \in T)$ is defined as

$$\mu_X(t) := \mu_{X_t} = EX_t, \; t \in T.$$
Expectation and Covariance Functions

The expectation function of a process $X = (X_t, t \in T)$ is defined as

$$\mu_X(t) := \mu_{X_t} = EX_t, \quad t \in T.$$ 

The covariance function of $X$ is given by

$$C_X(t, s) := \text{cov}(X_t, X_s) = E[(X_t - EX_t)(X_s - EX_s)], \quad t, s \in T.$$ 

In particular, the variance function of $X$ is given by

$$\sigma^2_X(t) = C_X(t, t) = \text{var}(X_t), \quad t \in T.$$
The expectation function of a process $X = (X_t, t \in T)$ is defined as
$$\mu_X(t) := \mu_{X_t} = \mathbb{E}X_t, \ t \in T.$$ 

The covariance function of $X$ is given by
$$C_X(t, s) := \text{cov}(X_t, X_s) = \mathbb{E}[(X_t - \mathbb{E}X_t)(X_s - \mathbb{E}X_s)], \ t, s \in T.$$ 

In particular, the variance function of $X$ is given by
$$\sigma_X^2(t) = C_X(t, t) = \text{var}(X_t), \ t \in T.$$

**Example:** A Gaussian white noise $X = (X_t, 0 \leq t \leq 1)$ consists of iid $N(0, 1)$ random variables. In this case its finite-dimensional distributions are given by, for any $0 \leq t_1 \leq \cdots \leq t_n \leq 1$,
$$P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n) = \prod_{i=1}^n P(X_{t_i} \leq x_i) = \prod_{i=1}^n \Phi(x_i), \ \forall x \in \mathbb{R}^n.$$

Its expectation and covariance functions are given by $\mu_X(t) = 0$,
$$C_X(t, s) = \begin{cases} 1 & \text{if } t = s \\ 0 & \text{if } t \neq s \end{cases}$$
A process $X = (X_t, t \in T)$ is said to be \textit{strictly stationary} if for any $t_1, \ldots, t_n \in T$

$$(X_{t_1}, \ldots, X_{t_n}) = d (X_{t_1+h}, \ldots, X_{t_n+h}).$$

That is, its finite-dimensional distribution functions are invariant under time shifts.
Dependence Structure

A process $X = (X_t, t \in T)$ is said to be strictly stationary if for any $t_1, \ldots, t_n \in T$

$$(X_{t_1}, \ldots, X_{t_n}) \overset{d}{=} (X_{t_1+h}, \ldots, X_{t_n+h}).$$

That is, its finite-dimensional distribution functions are invariant under time shifts.

A process $X = (X_t, t \in T)$ is said to have stationary increments if

$$X_t - X_s \overset{d}{=} X_{t+h} - X_{s+h}, \quad \forall t, s, t + h, s + h \in T.$$
A process $X = (X_t, t \in T)$ is said to be **strictly stationary** if for any $t_1, \ldots, t_n \in T$

$$(X_{t_1}, \ldots, X_{t_n}) =_d (X_{t_1+h}, \ldots, X_{t_n+h}).$$

That is, its finite-dimensional distribution functions are invariant under time shifts.

A process $X = (X_t, t \in T)$ is said to have **stationary increments** if

$$X_t - X_s =_d X_{t+h} - X_{s+h}, \ \forall t, s, t+h, s+h \in T.$$ 

A process $X = (X_t, t \in T)$ is said to have **independent increments** if for all $t_1 < \cdots < t_n$ in $T$,

$$X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$$

are independent.
Strictly Stationary vs Stationary

A process $X$ is said to be stationary (in the wide sense) if

$$\mu_X(t + h) = \mu_X(t), \text{ and } C_X(t, s) = C_X(t + h, s + h).$$
Strictly Stationary vs Stationary

A process $X$ is said to be **stationary (in the wide sense)** if

$$
\mu_X(t + h) = \mu_X(t), \text{ and } C_X(t, s) = C_X(t + h, s + h).
$$

If second moments exist, then the strictly stationarity implies the stationarity.
A process $X$ is said to be **stationary (in the wide sense)** if

$$
\mu_X(t + h) = \mu_X(t), \quad \text{and} \quad C_X(t, s) = C_X(t + h, s + h).
$$

If second moments exist, then the strictly stationarity implies the stationarity.

**Example:** Consider a strictly stationary Gaussian process $X$. The distribution of $X$ is determined by $\mu_X(0)$ and $C_X(t, s) = g_X(|t - s|)$ for some function $g_X$. In particular, for Gaussian white noise $X$, $g_X(0) = 1$ and $g_X(x) = 0$ for any $x \neq 0$. 
A stochastic process $X = (X_t, t \geq 0)$ is called an Poisson process with intensity rate $\lambda > 0$ if
Homogeneous Poisson Process

A stochastic process \( X = (X_t, t \geq 0) \) is called an **Poisson process** with intensity rate \( \lambda > 0 \) if

- \( X_0 = 0 \),

Simulation of Poisson Processes

Simulate iid exponential \( \text{Exp}(\lambda) \) random variables \( Y_1, Y_2, \ldots \), and set \( T_n := \sum_{i=1}^{n} Y_i \). The Poisson process can be constructed by \( X_t := \# \{ n : T_n \leq t \} , t \geq 0 \).

Example: Claims arriving in an insurance portfolio.
Homogeneous Poisson Process

A stochastic process $X = (X_t, t \geq 0)$ is called an **Poisson process** with intensity rate $\lambda > 0$ if

- $X_0 = 0$,

- it has stationary, independent increments, and
Homogeneous Poisson Process

A stochastic process $X = (X_t, t \geq 0)$ is called an **Poisson process** with intensity rate $\lambda > 0$ if

- $X_0 = 0$,
- it has stationary, independent increments, and
- for every $t > 0$, $X_t$ has a Poisson distribution $\text{Poi}(\lambda t)$.
A stochastic process $X = (X_t, t \geq 0)$ is called an Poisson process with intensity rate $\lambda > 0$ if

- $X_0 = 0$,
- it has stationary, independent increments, and
- for every $t > 0$, $X_t$ has a Poisson distribution $\text{Poi}(\lambda t)$.

### Simulation of Poisson Processes

Simulate iid exponential $\text{Exp}(\lambda)$ random variables $Y_1, Y_2, \ldots$, and set $T_n := \sum_{i=1}^n Y_i$. The Poisson process can be constructed by

$$X_t := \#\{n : T_n \leq t\}, \ t \geq 0.$$
Homogeneous Poisson Process

A stochastic process \( X = (X_t, t \geq 0) \) is called an **Poisson process** with intensity rate \( \lambda > 0 \) if

- \( X_0 = 0 \),
- it has stationary, independent increments, and
- for every \( t > 0 \), \( X_t \) has a Poisson distribution \( \text{Poi}(\lambda t) \).

### Simulation of Poisson Processes

Simulate iid exponential \( \text{Exp}(\lambda) \) random variables \( Y_1, Y_2, \ldots, \) and set \( T_n := \sum_{i=1}^{n} Y_i \). The Poisson process can be constructed by

\[
X_t := \#\{n : T_n \leq t\}, \quad t \geq 0.
\]

**Example:** Claims arriving in an insurance portfolio.