Chapter 1: **Stochastic Processes**

What are Stochastic Processes, and how do they fit in?

**Stats 210**: laid the foundations of both Statistics and Probability: the tools for understanding randomness.

**Stats 310**: develops the theory for understanding randomness in pattern: tools for estimating parameters (maximum likelihood), testing hypotheses, modelling patterns in data (regression models).

**Stats 325**: develops the theory for understanding randomness in process. A process is a sequence of events where each step follows from the last after a random choice.

What sort of problems will we cover in Stats 325?

Here are some examples of the sorts of problems that we study in this course.

**Gambler’s Ruin**

You start with $30 and toss a fair coin repeatedly. Every time you throw a Head, you win $5. Every time you throw a Tail, you lose $5. You will stop when you reach $100 or when you lose everything. What is the probability that you lose everything? **Answer:** 70%.
$X_1, X_2, X_3, X_4, \ldots$ (IID pattern)

1st order process

Markov Chain
Winning at tennis

What is your probability of winning a game of tennis, starting from the even score Deuce (40-40), if your probability of winning each point is 0.3 and your opponent’s is 0.7?

Answer: 15%.

Winning a lottery

A million people have bought tickets for the weekly lottery draw. Each person has a probability of one-in-a-million of selecting the winning numbers. If more than one person selects the winning numbers, the winner will be chosen at random from all those with matching numbers.

You watch the lottery draw on TV and your numbers match the winning numbers!!! Only a one-in-a-million chance, and there were only a million players, so surely you will win the prize?

Not quite... What is the probability you will win? Answer: only 63%.

Drunkard’s walk

A very drunk person staggers to left and right as he walks along. With each step he takes, he staggers one pace to the left with probability 0.5, and one pace to the right with probability 0.5. What is the expected number of paces he must take before he ends up one pace to the left of his starting point?

Answer: the expectation is infinite!
Pyramid selling schemes

Have you received a chain letter like this one? Just send $10 to the person whose name comes at the top of the list, and add your own name to the bottom of the list. Send the letter to as many people as you can. Within a few months, the letter promises, you will have received $77,000 in $10 notes! Will you?

**Answer:** It depends upon the response rate. However, with a fairly realistic assumption about response rate, we can calculate an expected return of $76 with a 64% chance of getting nothing!

**Note:** Pyramid selling schemes like this are prohibited under the Fair Trading Act, and it is illegal to participate in them.

Spread of SARS

The figure to the right shows the spread of the disease SARS (Severe Acute Respiratory Syndrome) through Singapore in 2003. With this pattern of infections, what is the probability that the disease eventually dies out of its own accord?

**Answer:** 0.997.
Markov’s Marvellous Mystery Tours

Mr Markov’s Marvellous Mystery Tours promises an All-Stochastic Tourist Experience for the town of Rotorua. Mr Markov has eight tourist attractions, to which he will take his clients completely at random with the probabilities shown below. He promises at least three exciting attractions per tour, ending at either the Lady Knox Geyser or the Tarawera Volcano. (Unfortunately he makes no mention of how the hapless tourist might get home from these places.)

What is the expected number of activities for a tour starting from the museum?

Answer: 4.2.

Structure of the course

- **Probability.** Probability and random variables, with special focus on conditional probability. Finding hitting probabilities for stochastic processes.

- **Expectation.** Expectation and variance. Introduction to conditional expectation, and its application in finding expected reaching times in stochastic processes.

- **Generating functions.** Introduction to probability generating functions, and their applications to stochastic processes, especially the Random Walk.

- **Branching process.** This process is a simple model for reproduction. Examples are the pyramid selling scheme and the spread of SARS above.
• **Markov chains.** Almost all the examples we look at throughout the course can be formulated as Markov chains. By developing a single unifying theory, we can easily tackle complex problems with many states and transitions like Markov’s Marvellous Mystery Tours above.

The rest of this chapter covers:

- quick revision of sample spaces and random variables;
- formal definition of stochastic processes.

# 1.1 Revision: Sample spaces and random variables

**Definition:** A **random experiment** is a physical situation whose outcome cannot be predicted until it is observed.

**Definition:** A **sample space**, $\Omega$, is a set of possible outcomes of a random experiment.

**Example:**

Random experiment: Toss a coin once.
Sample space: $\Omega = \{\text{head}, \text{tail}\}$

**Definition:** A **random variable**, $X$, is defined as a function from the sample space to the real numbers: $X : \Omega \to \mathbb{R}$.

That is, a random variable assigns a real number to every possible outcome of a random experiment.

**Example:**

Random experiment: Toss a coin once.
Sample space: $\Omega = \{\text{head, tail}\}$.
An example of a random variable: $X : \Omega \to \mathbb{R}$ maps “head” $\to 1$, “tail” $\to 0$.

Essential point: A random variable is a way of producing random real numbers.
1.2 Stochastic Processes

**Definition:** A **stochastic process** is a family of random variables, \( \{ X(t) : t \in T \} \), where \( t \) usually denotes time. That is, at every time \( t \) in the set \( T \), a random number \( X(t) \) is observed.

**Definition:** \( \{ X(t) : t \in T \} \) is a **discrete-time process** if the set \( T \) is finite or countable.

In practice, this generally means \( T = \{ 0, 1, 2, 3, \ldots \} \).

Thus a discrete-time process is \( \{ X(0), X(1), X(2), \ldots \} \): a random number associated with every time \( t = 0, 1, 2, \ldots \).

**Definition:** \( \{ X(t) : t \in T \} \) is a **continuous-time process** if \( T \) is not finite or countable.

In practice, this generally means \( T = [0, \infty) \) or \( T = [0, K] \) for some \( K \).

Thus a continuous-time process \( \{ X(t) : t \in T \} \) has a random number \( X(t) \) associated with **every instant** in time.

(Note that \( X(t) \) need not change at every instant in time, but it is allowed to change at any time; i.e. not just at \( t = 0, 1, 2, \ldots \), like a discrete-time process.)

**Definition:** The **state space**, \( S \), is the set of real values that any \( X(t) \) can take.

Every \( X(t) \) takes a value in \( \mathbb{R} \), but \( S \) will often be a smaller set: \( S \subseteq \mathbb{R} \). For example, if \( X(t) \) is the outcome of a coin tossed at time \( t \), then the state space is \( S = \{ 0, 1 \} \).

**Definition:** The state space \( S \) is **discrete** if it is finite or countable. Otherwise it is **continuous**.

The state space \( S \) is the set of states that the stochastic process can be in.
For Reference: Discrete Random Variables

1. **Binomial distribution**

   **Notation:** \( X \sim \text{Binomial}(n, p) \).

   **Description:** number of successes in \( n \) independent trials, each with probability \( p \) of success.

   **Probability function:**
   \[
   f_X(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad \text{for} \quad x = 0, 1, \ldots, n.
   \]

   **Mean:** \( E(X) = np \).

   **Variance:** \( \text{Var}(X) = np(1 - p) = npq \), where \( q = 1 - p \).

   **Sum:** If \( X \sim \text{Binomial}(n, p) \), \( Y \sim \text{Binomial}(m, p) \), and \( X \) and \( Y \) are independent, then
   \[ X + Y \sim \text{Bin}(n + m, p). \]

2. **Poisson distribution**

   **Notation:** \( X \sim \text{Poisson}(\lambda) \).

   **Description:** arises out of the Poisson process as the number of events in a fixed time or space, when events occur at a constant average rate. Also used in many other situations.

   **Probability function:**
   \[
   f_X(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad \text{for} \quad x = 0, 1, 2, \ldots
   \]

   **Mean:** \( E(X) = \lambda \).

   **Variance:** \( \text{Var}(X) = \lambda \).

   **Sum:** If \( X \sim \text{Poisson}(\lambda) \), \( Y \sim \text{Poisson}(\mu) \), and \( X \) and \( Y \) are independent, then
   \[ X + Y \sim \text{Poisson}(\lambda + \mu). \]
3. Geometric distribution

**Notation:** \( X \sim \text{Geometric}(p) \).

**Description:** number of failures before the first success in a sequence of independent trials, each with \( \mathbb{P}(\text{success}) = p \).

**Probability function:** \( f_X(x) = \mathbb{P}(X = x) = (1 - p)^x p \) for \( x = 0, 1, 2, \ldots \)

**Mean:** \( \mathbb{E}(X) = \frac{1 - p}{p} = \frac{q}{p} \), where \( q = 1 - p \).

**Variance:** \( \text{Var}(X) = \frac{1 - p}{p^2} = \frac{q}{p^2} \), where \( q = 1 - p \).

**Sum:** if \( X_1, \ldots, X_k \) are independent, and each \( X_i \sim \text{Geometric}(p) \), then \( X_1 + \ldots + X_k \sim \text{Negative Binomial}(k, p) \).

4. Negative Binomial distribution

**Notation:** \( X \sim \text{NegBin}(k, p) \).

**Description:** number of failures before the kth success in a sequence of independent trials, each with \( \mathbb{P}(\text{success}) = p \).

**Probability function:**
\[
 f_X(x) = \mathbb{P}(X = x) = \binom{k + x - 1}{x} p^k (1 - p)^x \text{ for } x = 0, 1, 2, \ldots
\]

**Mean:** \( \mathbb{E}(X) = \frac{k(1 - p)}{p} = \frac{kq}{p} \), where \( q = 1 - p \).

**Variance:** \( \text{Var}(X) = \frac{k(1 - p)}{p^2} = \frac{kq}{p^2} \), where \( q = 1 - p \).

**Sum:** If \( X \sim \text{NegBin}(k, p) \), \( Y \sim \text{NegBin}(m, p) \), and \( X \) and \( Y \) are independent, then \( X + Y \sim \text{NegBin}(k + m, p) \).
5. Hypergeometric distribution

**Notation:** \( X \sim \text{Hypergeometric}(N, M, n) \).

**Description:** Sampling without replacement from a finite population. Given \( N \) objects, of which \( M \) are ‘special’. Draw \( n \) objects without replacement. \( X \) is the number of the \( n \) objects that are ‘special’.

**Probability function:**

\[
f_X(x) = \mathbb{P}(X = x) = \binom{M}{x} \binom{N-M}{n-x} \binom{N}{n}^{-1} \quad \text{for} \quad \left\{ \begin{array}{l} x = \max(0, n + M - N) \\ \text{to} \quad x = \min(n, M). \end{array} \right.
\]

**Mean:** \( \mathbb{E}(X) = np \), where \( p = \frac{M}{N} \).

**Variance:** \( \text{Var}(X) = np(1 - p) \left( \frac{N - n}{N - 1} \right) \), where \( p = \frac{M}{N} \).

---

6. Multinomial distribution

**Notation:** \( X = (X_1, \ldots, X_k) \sim \text{Multinomial}(n; \ p_1, p_2, \ldots, p_k) \).

**Description:** there are \( n \) independent trials, each with \( k \) possible outcomes. Let \( p_i = \mathbb{P}(\text{outcome } i) \) for \( i = 1, \ldots, k \). Then \( X = (X_1, \ldots, X_k) \), where \( X_i \) is the number of trials with outcome \( i \), for \( i = 1, \ldots, k \).

**Probability function:**

\[
f_X(x) = \mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} \quad \text{for} \quad x_i \in \{0, \ldots, n\} \ \forall i \ \text{with} \ \sum_{i=1}^{k} x_i = n, \ \text{and where} \ \sum_{i=1}^{k} p_i = 1.
\]

**Marginal distributions:** \( X_i \sim \text{Binomial}(n, p_i) \) for \( i = 1, \ldots, k \).

**Mean:** \( \mathbb{E}(X_i) = np_i \) for \( i = 1, \ldots, k \).

**Variance:** \( \text{Var}(X_i) = np_i(1 - p_i) \), for \( i = 1, \ldots, k \).

**Covariance:** \( \text{cov}(X_i, X_j) = -np_i p_j \), for all \( i \neq j \).
Continuous Random Variables

1. Uniform distribution

**Notation:** \( X \sim \text{Uniform}(a, b) \).

**Probability density function (pdf):** \( f_X(x) = \frac{1}{b-a} \) for \( a < x < b \).

**Cumulative distribution function:**
\[
F_X(x) = P(X \leq x) = \frac{x-a}{b-a} \quad \text{for} \quad a < x < b.
\]
\[
F_X(x) = 0 \quad \text{for} \quad x \leq a, \quad \text{and} \quad F_X(x) = 1 \quad \text{for} \quad x \geq b.
\]

**Mean:** \( \mathbb{E}(X) = \frac{a+b}{2} \).

**Variance:** \( \operatorname{Var}(X) = \frac{(b-a)^2}{12} \).

2. Exponential distribution

**Notation:** \( X \sim \text{Exponential}(\lambda) \).

**Probability density function (pdf):** \( f_X(x) = \lambda e^{-\lambda x} \) for \( 0 < x < \infty \).

**Cumulative distribution function:**
\[
F_X(x) = P(X \leq x) = 1 - e^{-\lambda x} \quad \text{for} \quad 0 < x < \infty.
\]
\[
F_X(x) = 0 \quad \text{for} \quad x \leq 0.
\]

**Mean:** \( \mathbb{E}(X) = \frac{1}{\lambda} \).

**Variance:** \( \operatorname{Var}(X) = \frac{1}{\lambda^2} \).

**Sum:** if \( X_1, \ldots, X_k \) are independent, and each \( X_i \sim \text{Exponential}(\lambda) \), then \( X_1 + \ldots + X_k \sim \text{Gamma}(k, \lambda) \).
3. Gamma distribution

**Notation:** $X \sim \text{Gamma}(k, \lambda)$.

**Probability density function (pdf):**

$$f_X(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} \text{ for } 0 < x < \infty,$$

where $\Gamma(k) = \int_0^\infty y^{k-1} e^{-y} \, dy$ (the Gamma function).

**Cumulative distribution function:** no closed form.

**Mean:** $E(X) = \frac{k}{\lambda}$.

**Variance:** $\text{Var}(X) = \frac{k}{\lambda^2}$.

**Sum:** if $X_1, \ldots, X_n$ are independent, and $X_i \sim \text{Gamma}(k_i, \lambda)$, then

$$X_1 + \ldots + X_n \sim \text{Gamma}(k_1 + \ldots + k_n, \lambda).$$

4. Normal distribution

**Notation:** $X \sim \text{Normal}(\mu, \sigma^2)$.

**Probability density function (pdf):**

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \text{ for } -\infty < x < \infty.$$

**Cumulative distribution function:** no closed form.

**Mean:** $E(X) = \mu$.

**Variance:** $\text{Var}(X) = \sigma^2$.

**Sum:** if $X_1, \ldots, X_n$ are independent, and $X_i \sim \text{Normal}(\mu_i, \sigma_i^2)$, then

$$X_1 + \ldots + X_n \sim \text{Normal}(\mu_1 + \ldots + \mu_n, \sigma_1^2 + \ldots + \sigma_n^2).$$
Probability Density Functions

**Uniform**($a, b$)

\[
f_X(x) = \begin{cases} 1 & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}
\]

**Exponential**($\lambda$)

\[f(x) = \lambda e^{-\lambda x} \quad \text{for } x \geq 0, \lambda > 0\]

**Gamma**($k, \lambda$)

\[f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} \quad \text{for } x \geq 0, k, \lambda > 0\]

**Normal**($\mu, \sigma^2$)

\[f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for } x \in \mathbb{R}\]