Chapters 1-3. Background

STAT 462/662 Introduction to Stochastic Processes

Fall 2021

Introduction to Probability Theory

- Probabilities defined on events
- Conditional probability
- Independent events
- Bayes' formula

Probabilities Defined on Events

Consider a sample space S.

For each event E of S, we assume that $P(E)$ is defined and satisfies the following three conditions:

(1) $0 \leq P(E) \leq 1$

(2) $P(S) = 1$

(3) For any sequence of event $E_1, E_2, \ldots$ that are mutually exclusive, that is, events for which $E_n \cap E_m = \varnothing$ when $n \neq m$, then

$$P\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} P(E_n)$$

Probabilities Defined on Events (cont'd)

Useful formula:

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- When A and B are mutually exclusive (that is, when $A \cap B = \varnothing$), then

$$P(A \cup B) = P(A) + P(B)$$

In the same way, we can write

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C) \end{aligned}$$

Example 1: Probabilities defined on events

The experiment consists of the flipping of a coin.
If we assume that a head is equally likely to appear as a tail, then

$$P(H) = P(T) = \frac{1}{2}$$

Now, suppose that we toss two coins. Then,

$$S = \{(H, H), (H, T), (T, H), (T, T)\}$$

Each outcome has probability $\frac{1}{4}$.
Find the probability that either the first or the second coin falls heads.

Conditional Probability

Suppose that we toss two dice and that each of the 36 possible outcomes is equally likely to occur and hence has probability $\frac{1}{36}$.

Suppose that we observe that the first die is a four.
Given this information, what is the probability that the sum of the two dice equals six?

Possible outcomes: $(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)$
Hence, the desired probability will be $\frac{1}{6}$.

General formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{when } P(B) > 0$$

Example 2: Conditional Probability

Suppose that cards numbered one through ten are placed in a hat, mixed up, and then one of the cards is drawn.

If we are told that the number on the drawn card is at least five, then what is the conditional probability that it is ten?

Independent Events

Two events A and B are said to be **independent** if

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

Also, A and B are **independent** if

$$P(A|B) = P(A)$$

Bayes' Formula

Let A and B be events. We may express A as

$$A = (A \cap B) \cup (A \cap B^c)$$

Since $(A \cap B)$ and $(A \cap B^c)$ are mutually exclusive, we have that

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(A|B)P(B) + P(A|B^c)P(B^c) \end{aligned}$$

This implies that $P(A)$ is a weighted average of the conditional probability of A given that B has occurred and the conditional probability of A given that B has not occurred.
Using

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i),$$

we have

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} \qquad : \text{Bayes' formula}$$

Random Variables

- Discrete random variables
- Continuous random variables
- Expectation of a random variable
- Jointly distributed random variables
- Limit theorems
- Stochastic processes

Discrete Random Variables

A random variable that can take on at most a countable number of possible values is said to be **discrete**.

For a discrete random variable X, we define the **probability mass function** of X by $P(X = x)$.

- Bernoulli random variable
- Binomial random variable
- Geometric random variable
- Poisson random variable

Bernoulli Random Variable

Suppose that a trial, or an experiment, whose outcome can be classified as either "success" or as a "failure" is performed. We let

$$X = \begin{cases} 1 & \text{if outcome is a success} \\ 0 & \text{if outcome is a failure} \end{cases}$$

The probability mass function of X is given by

$$P(X=0) = 1 - p = q \quad \text{and} \quad P(X=1) = p,$$

where p is the probability that the trial is a "success".
Simply,

$$P(X=x) = p^x q^{1-x} \quad \text{for} \quad x = 0, 1$$

- $X \sim Bernoulli(p)$
- $E(X) = p$
- $Var(X) = pq$

Binomial Random Variable

Suppose that n independent trials, each of which results in a "success" with probability p and in a "failure" with probability $1-p$, are to be performed.

If X represents the number of successes that occur in the n trials, then X is said to be a **binomial** rv with parameters (n, p).

- $X \sim B(n,p)$
- $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \ldots, n$
- $E(X) = np$ and $Var(X) = npq$
- By the binomial theorem, the probabilities sum to one.

$$\sum_{x=0}^{n} P(X=x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1$$

- Binomial theorem: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

Example 3: Binomial Random Variable

It is known that any item produced by a certain machine will be defective with probability 0.1, independently of any other item.

What is the probability that in a sample of three items, at most one will be defective?

$X =$ number of defective items

$P(X \text{ is at most one}) = P(X = 0) + P(X = 1) = 0.972$

Geometric Random Variable

Suppose that independent trials, each having probability p of being a success, are performed until a success occurs.
If we let X be the number of trials required until the first success, then X is said to be a **geometric** rv with parameter p.

- $X \sim Geo(p)$
- $P(X = x) = (1 - p)^{x-1} p$ for $x = 1, 2, \ldots$
- $E(X) = \frac{1}{p}$ and $Var(X) = \frac{q}{p^2}$

Poisson Random Variable

A random variable X, taking on one of the values $0, 1, 2, \ldots$, is said to be a **Poisson** random variable with parameter λ, if for some $\lambda > 0$,

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Since

$$\sum_{x=0}^{\infty} P(X = x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!} = 1,$$

we have

$$e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^{x}}{x!}$$

- $X \sim Poi(\lambda)$
- $E(X) = \lambda$ and $Var(X) = \lambda$

Poisson Random Variable (cont'd)

An important property of the Poisson random variable is that it may be used to approximate a binomial random variable when n is large and p is small.

Suppose that $X \sim Bin(n, p)$ and let $\lambda = np$. Then,

$$
\begin{aligned}
P(X = x) &= \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \\
&= \frac{n!}{(n-x)!x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{n^x} \frac{\lambda^x}{x!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^x} \\
&\approx \frac{e^{-\lambda}\lambda^x}{x!} \quad \text{for } n \text{ large and } p \text{ small}
\end{aligned}
$$

Example 4: Poisson Random Variable

1. Suppose that the number of typographical errors on a single page of the book has a Poisson distribution with parameter $\lambda = 1$.

   Find the probability that there is at least one error on one page.

   $P(X \geq 1) \approx 0.633$

2. If the number of accidents occurring on a highway each day is a Poisson random variable with parameter $\lambda = 3$, what is the probability that no accidents occur today?

   $P(X = 0) \approx 0.05$

Continuous Random Variables

X is a continuous random variable if there exists a nonnegative function $f(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set B of real numbers

$$P(X \in B) = \int_B f(x)dx$$

The cumulative distribution function is

$$F(a) = P(X \in (-\infty, a]) = \int_{-\infty}^{a} f(x)dx$$

- Uniform random variable
- Exponential random variable
- Gamma random variable
- Normal random variable

Uniform Random Variable

A random variable is said to be **uniformly distributed** over the interval (a, b) if its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

- $X \sim Unif(a, b)$
- $E(X) = \frac{1}{2}(a + b)$ and $Var(X) = \frac{1}{12}(b - a)^2$

Exponential Random Variable

A continuous random variable whose probability density function is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

is said to be an exponential random variable with parameter λ.

- $X \sim Exp(\lambda)$
- $E(X) = \frac{1}{\lambda}$ and $Var(X) = \frac{1}{\lambda^2}$

Gamma Random Variable

A continuous random variable whose density is given by

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

for some $\alpha > 0$ and $\beta > 0$, is said to be a gamma random variable with parameters α and β.

- $X \sim Gamma(\alpha, \beta)$
- $\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx$: Gamma function
- For a positive integer n, $\Gamma(n) = (n-1)!$
- $E(X) = \frac{\alpha}{\beta}$ and $Var(X) = \frac{\alpha}{\beta^2}$

Gamma Random Variable (cont'd)

- A gamma distribution with shape parameter $\alpha = 1$ and scale parameter β is an exponential distribution with parameter β.
- The sum of gamma (n_i,β) random variables has a gamma ($\sum n_i$, β) distribution.
- The sum of n exponential (β) random variables is a gamma (n,β) random variable.

Normal Random Variable

X is a normal random variable with parameters μ and σ^2 if the density of X is given by

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

- $X \sim N(\mu, \sigma^2)$
- $E(X) = \mu$ and $Var(X) = \sigma^2$

Expectation of a Random Variable

- Discrete case

$$E(X) = \sum_x xP(X = x)$$

- Continuous case

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

- For any real-valued function g,

$$\begin{aligned} E[g(X)] &= \sum_x g(x)P(X = x) \quad \text{for a discrete r.v. } X \\ &= \int_{-\infty}^{\infty} g(x)f(x)dx \quad \text{for a continuous r.v. } X \end{aligned}$$

- If a and b are constants, then $E[aX + b] = aE[X] + b$.
- $Var(X) = E\left[(X - E(X))^2\right] = E(X^2) - (E[X])^2$

Jointly Distributed Random Variables

- Joint distribution functions
- Independent random variables
- Covariance and variance of sums of random variables

Joint distribution functions

For any two random variables X and Y, the joint cumulative probability function of X and Y is given by

$$F(x, y) = P(X \leq x, Y \leq y) \quad \text{for} \quad -\infty < x, y < \infty$$

If X and Y are random variables and g is a function of two variables, then

$$\begin{aligned} E[g(X, Y)] &= \sum_y \sum_x g(x, y) P(X = x, Y = y) \quad &: \text{discrete case} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \quad &: \text{continuous case} \end{aligned}$$

For any constants a and b,

$$E(aX + bY) = aE(X) + bE(Y)$$

Independent random variables

The random variables X and Y are said to be independent if, for all x and y,

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

Similarly, we have

$$\begin{aligned} P(X = x, Y = y) &= P(X = x)P(Y = y) \quad &: \text{discrete case} \\ f(x, y) &= f_X(x)f_Y(y) \quad &: \text{continuous case} \end{aligned}$$

If X and Y are independent, then for any functions h and g

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

Covariance and variance of sums of random variables

The covariance of any two random variables X and Y is defined by

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

If X and Y are independent, then $Cov(X,Y)=0$.

For any random variables X, Y, Z and constant c,

- $Cov(X,X)=Var(X)$
- $Cov(X,Y)=Cov(Y,X)$
- $Cov(cX,Y)=cCov(X,Y)$
- $Cov(X,Y+Z)=Cov(X,Y)+Cov(X,Z)$

Example 5: Sums of Independent Poisson Random Variables

Let X and Y be independent Poisson random variables with respective means λ_1 and λ_2.
Calculate the distribution of $X + Y$.

The event $\{X + Y = n\} = \{X = k, Y = n - k\}$ for $0 \leq k \leq n$

$$\begin{aligned}
P(X + Y = n) &= \sum_{k=0}^{n} P(X = k, Y = n - k) \\
&= \sum_{k=0}^{n} P(X = k) P(Y = n - k) \\
&= \sum_{k=0}^{n} \frac{e^{-\lambda_1} \lambda_1^k}{k!} \frac{e^{-\lambda_2} \lambda_2^{n-k}}{(n-k)!} \\
&= \frac{e^{-(\lambda_1 + \lambda_2)} (\lambda_1 + \lambda_2)^n}{n!}
\end{aligned}$$

Hence, $X + Y$ has a Poisson distribution with mean $\lambda_1 + \lambda_2$.

Example 6: Order Statistics

Let $X_1, \ldots, X_n$ be iid continuous random variables with probability distribution F and density function f.

If we let $X_{(i)}$ denote the ith smallest of these random variables, then $X_{(1)}, \ldots, X_{(n)}$ are called the **order statistics**.

For example, if $n = 3$ and $X_1 = 4$, $X_2 = 5$, $X_3 = 1$, then $X_{(1)} = 1$, $X_{(2)} = 4$, $X_{(3)} = 5$.

The density function of $X_{(i)}$ is

$$f_{X_{(i)}}(x) = \frac{n!}{(n-i)!(i-1)!} f(x)(F(x))^{i-1}(1-F(x))^{n-i}$$

: quite intuitive

Limit Theorems

Strong Law of Large Numbers

Let $X_1, X_2, \ldots$ be a sequence of independent random variables having a common distribution, and let $E(X_i) = \mu$. Then, with probability 1,

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \mu \quad \text{as } n \to \infty$$

Central Limit Theorem (holds for any distribution of X)

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables, each with mean μ and variance σ^2. Then the distribution of $\frac{X_1+X_2+\cdots+X_n-n\mu}{\sigma\sqrt{n}}$ tends to the standard normal as $n \to \infty$. That is,

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx$$

Stochastic Processes

A **stochastic process** $\{X(t); t \in T\}$ is a collection of random variables. That is, for each $t \in T$, $X(t)$ is a random variable.

- The index t is often interpreted as **time**
- $X(t)$ is the **state** of the process at time t. For example,
  - the total number of customers that have entered a supermarket by time t
  - the number of customers in the supermarket at time t
- The set T is called the **index** set of the process.
  - When T is a countable set, a **discrete-time** process.
  - If T is an interval of the real line, a **continuous-time** process.
- The **state space** of a stochastic process is defined as the set of all possible values that $X(t)$ can assume.
- Thus, a stochastic process is a family of random variables that describes the evolution through time of some (physical) process.

Conditional Probability and Conditional Expectation

- Discrete Case
- Continuous Case
- Computing Expectations by Conditioning
- Computing Variances by Conditioning
- Computing Probabilities by Conditioning

Discrete Case

If X and Y are discrete random variables, then the conditional probability mass function of X given that $Y = y$ is defined by

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

for all values of y such that $P(Y = y) > 0$.

The conditional expectation of X given that $Y = y$ is defined by

$$E(X|Y = y) = \sum_x x P(X = x | Y = y)$$

If X is independent of Y, then

$$P(X = x | Y = y) = P(X = x)$$

Example 7

X and Y are independent Poisson random variables with respective means λ_1 and λ_2. Calculate the conditional expected value of X given that $X+Y=n$.

First calculate the conditional pmf of X given that $X+Y=n$.

$$\begin{aligned} P(X=k|X+Y=n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} \\ &= \frac{P(X=k, Y=n-k)}{P(X+Y=n)} \\ &= \frac{P(X=k)P(Y=n-k)}{P(X+Y=n)} \\ &= \binom{n}{k}\Big(\frac{\lambda_1}{\lambda_1+\lambda_2}\Big)^k\Big(\frac{\lambda_2}{\lambda_1+\lambda_2}\Big)^{n-k} \end{aligned}$$

That is, $X|X+Y=n \sim B(n, \frac{\lambda_1}{\lambda_1+\lambda_2})$.
Thus, $E(X|X+Y=n) = n\frac{\lambda_1}{\lambda_1+\lambda_2}$

Continuous Case

If X and Y have a joint pdf $f(x, y)$, then the conditional pdf of X, given that $Y = y$, is defined for all values of y such that $f_Y(y) > 0$, by

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

The conditional expectation of X, given that $Y = y$, is defined for all values of y such that $f_Y(y) > 0$, by

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

Computing Expectations and Variances by Conditioning

$E(X|Y)$ is a function of the random variable Y. At $Y = y$, the value is $E(X|Y = y)$. Note that $E(X|Y)$ is itself a random variable.

For all random variables X and Y,

$$E(X) = E[E(X|Y)]$$

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$

Example 8

Suppose that the expected number of accidents per week at an industrial plant is four. Suppose also that the numbers of workers injured in each accident are independent random variables with a common mean of 2. Assume also that the number of workers injured in each accident is independent of the number of accidents that occur. What is the expected number of injuries during a week?

Let
$N =$ the number of accidents
$X_i =$ the number of workers injured in the ith accident
$(i = 1, 2, \cdots)$
Then, $\sum_{i=1}^{N} X_i =$ the total number of injuries

Find $E(\sum_{i=1}^{N} X_i)$.

Example 8 (cont'd)

$$\begin{aligned}
E\left(\sum_{i=1}^{N} X_i\right) &= E\left[E\left(\sum_{i=1}^{N} X_i \middle| N\right)\right] \\
&= E(NE(X)) \\
&= E(N)E(X) \\
&= 8 \\
\text{since} \quad E\left(\sum_{i=1}^{N} X_i \middle| N = n\right) &= E\left(\sum_{i=1}^{n} X_i \middle| N = n\right) \\
&= E\left(\sum_{i=1}^{n} X_i\right) \quad (\text{b/c } N \perp X_i) \\
&= \sum_{i=1}^{n} E(X_i) = nE(X)
\end{aligned}$$

$\sum_{i=1}^{N} X_i$ is called a **compound random variable**.

Example 9

In Example 8, let us assume that N is a Poisson random variable with mean λ and X is a iid random variable with mean μ and variance σ^2. Find $Var(\sum_{i=1}^{N} X_i)$.

$$\begin{aligned} Var\Big(\sum_{i=1}^{N} X_i\Big) &= E\Big[Var\Big(\sum_{i=1}^{N} X_i \Big| N\Big)\Big] + Var\Big[E\Big(\sum_{i=1}^{N} X_i \Big| N\Big)\Big] \\ &= E[NVar(X)] + Var[NE(X)] \\ &= E(N)Var(X) + [E(X)]^2 Var(N) \\ &= \lambda(\sigma^2 + \mu^2) \end{aligned}$$

① $$Var\Big(\sum_{i=1}^{N} X_i \Big| N = n\Big) = Var\Big(\sum_{i=1}^{n} X_i \Big| N = n\Big) = nVar(X)$$

② $$E\Big(\sum_{i=1}^{N} X_i \Big| N = n\Big) = E\Big(\sum_{i=1}^{n} X_i \Big| N = n\Big) = nE(X)$$

Computing Probabilities by Conditioning

$$P(X = x) = \sum_y P(X = x|Y = y)P(Y = y) \quad : \text{discrete case}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dy \quad : \text{continuous case}$$

Example 10

Suppose that the number of people who visit a yoga studio each day is a Poisson random variable with mean λ. Suppose further that each person who visits is, independently, female with probability p or male with probability $1 - p$. Find the joint probability that exactly n women and m men visit the academy today.

Define
$N =$ the number of visitors
$N_1 =$ the number of female visitors
$N_2 =$ the number of male visitors

That is, we want to find $P(N_1 = n, N_2 = m)$

Example 10 (cont'd)

Since $P(X = x) = \sum_y P(X = x|Y = y)P(Y = y)$,

$$\begin{aligned} P(N_1 = n, N_2 = m) &= \sum_{i=0}^{\infty} P(N_1 = n, N_2 = m|N = i)P(N = i) \\ &= P(N_1 = n, N_2 = m|N = n + m)P(N = n + m) \\ &= e^{-\lambda p}\frac{(\lambda p)^n}{n!} e^{-\lambda(1-p)}\frac{(\lambda(1-p))^m}{m!} \end{aligned}$$

① $P(N_1 = n, N_2 = m|N = n + m) = P(N_1 = n|N = n + m)$
$\Rightarrow$ Binomial$(n + m, p)$

② $P(N = n + m) \Rightarrow$ Poisson(λ)

Example 10 (cont'd)

Interestingly,

$$\begin{aligned} P(N_1 = n) &= \sum_{m=0}^{\infty} P(N_1 = n, N_2 = m) = e^{-\lambda p} \frac{(\lambda p)^n}{n!} \\ &\Rightarrow \text{Poisson}(\lambda p) \\ P(N_2 = m) &= \sum_{n=0}^{\infty} P(N_1 = n, N_2 = m) = e^{-\lambda(1-p)} \frac{(\lambda(1-p))^m}{m!} \\ &\Rightarrow \text{Poisson}(\lambda(1-p)) \end{aligned}$$

We will discuss this in Chapter 2 (Poisson process) later.

Arithmetic sequence

The n-th term of an arithmetic sequence with initial value a_1 and common difference of successive numbers d is given by $a_n = a_1 + (n-1)d$. Such a sequence also follows the recursive relation $a_n = a_{n-1} + d$ for every integer $n \geq 1$.
An arithmetic series is the sum of the numbers in a finite arithmetic sequence.

$$S_n = \sum_{i=1}^{n} a_i = \frac{1}{2} n\{2a_1 + (n-1)d\}.$$

Example: Find the 20th term for $1, 4, 7, 10, 13, \cdots$.
$a_1 = 1$, $d = 3$, thus the nth term $a_n = 3n - 2$.
The 20th term is 58.

Geometric sequence

The n-th term of a geometric sequence with initial value a_1 and common ratio r is given by $a_n = a_1 r^{n-1}$. Such a geometric sequence also follows the recursive relation $a_n = r a_{n-1}$ for every integer $n \geq 1$.

A geometric series is the sum of the numbers in a geometric sequence.

$$S_n = \sum_{i=1}^{n} a_i = \begin{cases} \frac{a_1(1-r^n)}{1-r} & \text{if } |r| \neq 1 \\ a_1 n & \text{if } |r| = 1 \end{cases}$$

An infinite geometric series is

$$\sum_{i=1}^{\infty} a_i = \begin{cases} \frac{a_1}{1-r} & \text{if } |r| < 1 \\ \infty & \text{if } |r| \geq 1 \end{cases}$$

Example: Sum up all the terms for the following sequence:
$2, 1, 0.5, 0.25, 0.125, 0.0625, \cdots$
$a_1 = 2$, $r = 0.5$, thus the sum is 4

Difference sequence

A difference sequence $\{b_n\}$ is the difference between the successive terms of another sequence $\{a_n\}$. That is, $b_n = a_{n+1} - a_n$. Now, $\{b_n\}$ can be an arithmetic sequence or geometric sequence.

For example, we have a sequence $\{1, 4, 9, 16, \ldots\}$. Its difference is $\{3, 5, 7, \ldots\}$. That is, $a_1 = 1, a_2 = 4, a_3 = 9, a_4 = 16, \ldots$. The difference sequence is $b_1 = a_2 - a_1 = 3$, $b_2 = a_3 - a_2 = 5$, $b_3 = a_4 - a_3 = 7$. That is, $\{b_n\}$ is an arithmetic sequence with initial value 3 and common difference 2.

The original sequence is

$$a_n = a_1 + \sum_{i=1}^{n-1} b_i$$