MATH 4205 / MATH 7620
Topics in Probability Theory and Stochastic Processes

Instructor: Dr. Wu Tian Talia

Department of Mathematics

Hong Kong Baptist University

Lecture 1

Review of Probability Theory

- Review concepts of: conditional probability; independent events; Bayes' formula; discrete random variables; continuous random variables; expectation of a random variable; jointly distributed random variables
- Introduce stochastic processes

1.1 Set Notations

- A set: $X = \{1, 2, 3\}$ (3 elements), $Y = \{a, b\}$ (2 elements), $\Omega = \{(1, 2), (3, 4), (4, 5)\}$ (3 elements)
- Empty set: $X = \emptyset$ (0 elements)
- A set of sets: $\mathcal{F} = \{\{1, 2, 3\}, \{1, 3, 5, 7\}\}$ (2 elements, each is a set)
- Union: $X = \{1, 2\}$, $Y = \{2, 3\}$, $X \cup Y = \{1, 2, 3\}$
- Intersection: $X = \{1, 2\}$, $Y = \{2, 3\}$, $X \cap Y = \{2\}$
- Complement: If $X = \{1, 2, 3, 4, 5\}$ is the universal set and $Y = \{1, 2\}$, then $Y^c = \{3, 4, 5\}$.
- Subset: If $X = \{1\}$, $Y = \{1, 2, 3\}$, then $X \subset Y$.
- Superset: If $X = \{1\}$, $Y = \{1, 2, 3\}$, then $Y \supset X$.
- Element: If $X = \{a, b, c\}$, then $a \in X$, $b \in X$, and $c \in X$.

1.2 Conditional Probabilities

- Given an event F has occurred, the probability that an event E also occurs is

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

Example 1.1 Suppose a card is drawn from a deck of ten cards numbered 1 through 10. If we are told that the number of the drawn card is 5 or more, then what is the conditional probability that it is 10?

Solution: Let $E = \{10\}$ and $F = \{5, 6, 7, 8, 9, 10\}$. Then,

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(\{10\})}{P(\{5, 6, 7, 8, 9, 10\})} = \frac{\frac{1}{10}}{\frac{6}{10}} = \frac{1}{6}.$$

$\square$

1.3 Independent Events

- Two events E and F are *independent* if

$$P(E|F)\cdot P(F) = P(E \cap F) = P(E)P(F).$$

- Equivalently,

$$P(E|F) = P(E).$$

That is, knowledge of occurrence of F does not affect the probability that E occurs.

Example 1.2 Two fair dice are tossed. Let E be the event that the sum is 6. Let F be the event that the first die equals 4. Then

$$P(E \cap F) = P(\{(4,2)\}) = \frac{1}{36}.$$

But

$$P(E)P(F) = \frac{5}{36} \cdot \frac{1}{6} = \frac{5}{216}.$$

Hence, E and F are dependent. What if the sum is 7?

$$P(E) = \frac{6}{6\times 6} = \frac{1}{6}$$

$$\begin{cases} 1+6 \\ 2+5 \\ 3+4 \end{cases} \quad P(F) = \frac{1}{6}$$

$$P(E \cap F) = P\{(4,3)\} = \frac{1}{36}$$

1.4 Bayes' Formula

$F_i \cap F_j = \phi \quad i,j=1,\cdots n,\ i \neq j$

- If $F_1, F_2, \ldots, F_n$ are mutually exclusive and $F_1 \cup F_2 \cup \ldots \cup F_n = \Omega$, that is, $\{F_1, \ldots, F_n\}$ is a *partition* of Ω, then

$$P(E) = \sum_{i=1}^{n} P(E \cap F_i) = \sum_{i=1}^{n} P(E|F_i)P(F_i).$$

This is called the *total probability formula.*

- Therefore,

updated prob from prior probability

posterior probability

$$P(F_j|E) = \frac{P(E \cap F_j)}{P(E)} = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}.$$

This is known as *Bayes' formula.*

$P(\text{Fire}) = 1\%$ $\quad P(\text{Smoke}) = 10\%$ $\quad P(\text{Smoke}|\text{Fire}) = 90\%$ $\quad P(\text{Fire}|\text{Smoke}) = \frac{P(F \cap S)}{P(S)} = \frac{P(S|F)\cdot P(F)}{P(S)}$

$= \frac{0.9 \times 0.01}{0.1} = 0.09 = 9\%$

Example 1.3 There are two urns. The first contains 2 white balls and 7 black balls. The second contains 5 white and 6 black balls. A fair coin is flipped. If it is a head (tail), then a ball from the first (second) urn is drawn. What is the conditional probability that the outcome of the toss was heads given that a white ball was selected?

Solution: Let W be the event that a white is drawn. Let H be the event that the coin comes up head. Note that $\Omega = H \cup H^c$.

$$P(H|W) = \frac{P(H \cap W)}{P(W)} = \frac{\frac{\cancel{2}}{9} \times \frac{1}{\cancel{2}}}{\frac{67}{198}} = \frac{\cancel{198}\,22}{\cancel{9} \times 67} = \frac{22}{67}$$

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

$$P(W|H) = \frac{2}{9}$$

$$P(W|T) = \frac{5}{11}$$

$$P(W) = P(W|H) \times P(H) + P(W|T)P(T)$$

$$= \frac{2}{9} \times \frac{1}{2} + \frac{5}{11} \times \frac{1}{2} = \frac{1}{9} + \frac{5}{22} = \frac{22+45}{198} = \frac{67}{198}$$

1.5 Random Variables

- A *random variable* X is a function on Ω, i.e. $X : \Omega \to \mathbb{R}$.
- E.g. $\{X = a\}$ denotes the set of outcomes such that $X = a$. Thus, it is an event. Some other events: $\{a \leq X \leq b\}$, $\{X \geq a\}$, etc.
- For simplicity, we write $P(\{X = a\})$ as $P\{X = a\}$, etc.

Example 1.4 A biased coin has a probability p of coming up heads. We toss the coin repeatedly until the first head appears. Let N be the number of flips required.

$$\begin{aligned}
P\{N=1\} &= P\{H\} = p \\
P\{N=2\} &= P\{(T,H)\} = (1-p)p \\
P\{N=3\} &= P\{(T,T,H)\} = (1-p)^2 p \\
&\vdots \\
P\{N=n\} &= P\{(\underbrace{T,T,\ldots,T}_{n-1},H)\} = (1-p)^{n-1}p.
\end{aligned}$$

As expected,

$$\sum_{n=1}^{\infty} P\{N=n\} = p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\cdot\frac{1}{1-(1-p)} = 1.$$

Here, we used the formula $1+x+x^2+\cdots=\frac{1}{1-x}$. $\square$

geometric series

1.6 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 (pmf)* $p(a)$ of X by

$$p(a) = P\{X = a\}.$$

- If $x_1, x_2, \ldots$ are the possible values of X, then

$$\sum_{i=1}^{\infty} p(x_i) = 1.$$

- The *cumulative distribution function* (cdf) F is given by

$$F(a) = \sum_{i:\, x_i \le a} p(x_i).$$

Example 1.5 Suppose X has a probability mass function given by

$$p(1) = \frac{1}{2}, \quad p(2) = \frac{1}{3}, \quad p(3) = \frac{1}{6}.$$

Then, the cdf F of X is given by

$$F(a) = \begin{cases} 0, & a < 1 \\ \frac{1}{2}, & 1 \leq a < 2 \\ \frac{5}{6}, & 2 \leq a < 3 \\ 1, & 3 \leq a. \end{cases}$$

distribution of a R.V is the most infomative way to describe its features because it contains all possible information about the random/prbablistic behavior.

1.6.1 Bernoulli Random Variable

- Suppose that the outcome of a trial (experiment) is either "success" or "failure".
- Suppose that the probability of a success is p.
- If we let $X = 1$ for a success and $X = 0$ for a failure, then the probability mass function of X is

$$\begin{aligned} p(0) &= P\{X = 0\} = 1 - p \\ p(1) &= P\{X = 1\} = p. \end{aligned}$$

- Such a random variable is called a *Bernoulli* random variable.

1.6.2 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.
- Let X be the number of successes in the n trials.
- Then, X is called a *binomial* random variable with parameters (n, p).
- Notation: $X \sim \text{Binomial}(n, p)$
- The probability mass function of X is

$$p(i) = P\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i}, \qquad i = 0, 1, \ldots, n$$

where

$$\binom{n}{i} = \frac{n!}{(n-i)!i!}.$$

Example 1.6 If an item produced by a machine will be defective with probability 0.1, independent of any other item. What is the probability that in a sample of three items, at most one will be defective?

Solution: Let X be the number of defective items in the sample. Then,

$$\begin{aligned}
& P\{X=0\}+P\{X=1\} \\
= & \binom{3}{0}(0.1)^0(0.9)^3+\binom{3}{1}(0.1)^1(0.9)^2 \\
= & 0.972.
\end{aligned}$$

□

1.6.3 Geometric Random Variable

- Suppose that independent trials, each having probability p of being a success, are run until a success occurs.
- Let X be the number of trials required until the first success. Then, X is called a *geometric* random variable with parameter p.
- The probability mass function of X is

$$p(n) = P\{X = n\} = (1-p)^{n-1}p, \qquad n = 1, 2, \ldots$$

1.6.4 Poisson Random Variable

- A random variable X is called a *Poisson* random variable with parameter λ if

$$p(i) = P\{X = i\} = \boxed{e^{-\lambda} \cdot \frac{\lambda^i}{i!},} \qquad i = 0, 1, 2, \ldots$$

- It is used to model the number of customers arriving by a unit of time. The λ is the *arrival rate* (mean number of customers per unit of time).

Example 1.7 If the number of accidents on a highway each day is Poisson with $\lambda = 3$, what is the probability that no accidents occur today?

Solution:

$$P\{X = 0\} = e^{-3} \approx 0.0498.$$

$\square$

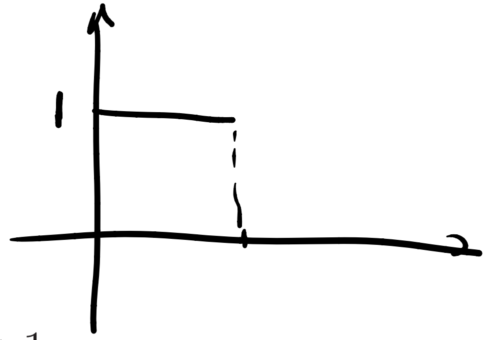
1.7 Continuous Variables

- Let X be a random variable whose set of possible values is uncountable.
- We assume that there is a nonnegative function $f(x)$ such that

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx.$$

  The function f is called the *probability density function* (pdf) of X.
- In particular, $P\{X = a\} = \int_a^a f(x)dx = 0$.
- *Cumulative distribution function* (cdf): $F(a) = \int_{-\infty}^a f(x)dx$.
- $\frac{dF(a)}{da} = f(a)$.
- $P\{a \leq X \leq b\} = F(b) - F(a)$.

1.7.1 Uniform Random Variable



- X is said to be *uniformly distributed* over $(0, 1)$ if

$$f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

- $P\{a \leq X \leq b\} = \int_a^b f(x)dx = b - a$ for $0 < a < b < 1$.
- More generally, X is uniformly distributed over (α, β) if $f(x) = \frac{1}{\beta - \alpha}$ on (α, β) and $f(x) = 0$ otherwise.

Example 1.8 Let X be uniformly distributed over $(0, 10)$. Calculate $P\{1 \leq X < 6\}$.

Solution:

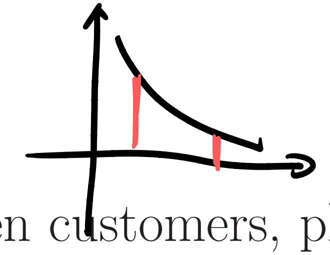
$$P\{1 < X < 6\} = \int_1^6 \frac{1}{10 - 0}\, dx = \frac{1}{2}.$$

□

1.7.2 Exponential Random Variable

- The pdf of an *exponential* random variable with parameter λ is

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



- It is used to model the inter-arrival time between customers, phone calls, storms, etc.

1.7.3 Normal Random Variable

- X is *normally distributed* with mean μ and variance σ^2 if

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty.$$

$X \sim N(\mu, \sigma^2)$

- If X has mean μ and variance σ^2, then $Z = \frac{X-\mu}{\sigma}$ is normally distributed with mean 0 and variance 1. Such a Z is said to have the *standard normal distribution.*

1.8 Expectation of a Random Variable

1.8.1 The Discrete Case

- Denote by $x_1, x_2, \ldots$ the possible values of a discrete random variable X, i.e. $p(x_i) > 0$ where p is the probability mass function of X. The *expected value* of X is defined by

$$E[X] = \sum_i x_i \cdot p(x_i).$$

weighted sum

- Sometimes, we write $E[X] = \sum_x x \cdot p(x)$ for simplicity.

Example 1.9 The expectation of a Bernoulli random variable with parameter p is

$$E[X] = 0 \cdot p(0) + 1 \cdot p(1) = 0 \cdot (1-p) + 1 \cdot p = p.$$

□

Example 1.10 A Poisson random variable with parameter λ has possible values $0, 1, 2, \ldots$. Its expectation is

$$\begin{aligned} E[X] &= \sum_{i=0}^{\infty} i \cdot \frac{e^{-\lambda}\lambda^i}{i!} = \sum_{i=1}^{\infty} i \cdot \frac{e^{-\lambda}\lambda^i}{i!} = \sum_{i=1}^{\infty} \frac{e^{-\lambda}\lambda^i}{(i-1)!} \\ &= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

Taylor series of a fun of f(x) at x=a.

□

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2!}$$

1.8.2 The Continuous Case

- We assume that a continuous random variable X has a pdf $f(x)$. Then, its *expectation* is defined by

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Example 1.11 The expectation of a random variable uniformly distributed over (α, β) is

$$E[X] = \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx = \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} = \frac{\alpha + \beta}{2}.$$

$\square$

Example 1.12 The expectation of a random variable normally distributed with mean μ and standard deviation σ is

$$\begin{aligned}
E[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y+\mu) \cdot e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y \cdot e^{-\frac{y^2}{2\sigma^2}} dy + \frac{\mu}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy \\
&= 0 + \mu \cdot \int_{-\infty}^{\infty} f(x) dx \\
&= \mu.
\end{aligned}$$

The integral of $y \cdot e^{-\frac{y^2}{2\sigma^2}}$ is 0 because it is an odd function. □

1.8.3 Expectation of a Function of a Random Variable

- Let X be a random variable with pdf f_X (or pmf in discrete case). Let $Y = g(X)$ for some function g. We'd like to determine $E[Y]$.
- Since Y is a random variable, we can first determine its pdf f_Y and then obtain $E[Y]$ as $\int y f_Y(y)dy$.
- Luckily, we have the following result which greatly simplifies the computation. It allows us to reuse the pdf of X.

Proposition 1.13

$$E[g(X)] = \begin{cases} \sum_i g(x_i)p(x_i), & \textit{if } X \textit{ is discrete} \\ \int_{-\infty}^{\infty} g(x)f(x)dx, & \textit{if } X \textit{ is continuous.} \end{cases}$$

Example 1.14 Let X be uniformly distributed over $(0, 1)$. Then

$$E[X^3] = \int_0^1 x^3 \cdot 1dx = \left[\frac{x^4}{4}\right]_0^1 = \frac{1}{4}. \qquad \square$$

1.9 Jointly Distributed Random Variables

① marginal distribution
② covariance
③ cond. distribution

1.9.1 Joint Distribution Functions

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

$$F(a, b) = P\{X \le a, Y \le b\}.$$

- If X and Y are both discrete, the *joint probability mass function* of X and Y is

$$p(x, y) = P\{X = x, Y = y\}.$$

- The pmf of X and the pmf of Y can be recovered from $p(x, y)$:

$$p_X(x) = \sum_j p(x, y_j) \qquad p_Y(y) = \sum_i p(x_i, y).$$

- If both X and Y are continuous and if there is a function $f(x,y)$ so that

$$P\{a \le X \le b, c \le Y \le d\} = \int_c^d \int_a^b f(x,y)\,dx\,dy,$$

then $f(x,y)$ is called the *joint probability density function* of X and Y. The pdf of X and Y can be recovered from $f(x,y)$ as

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx.$$

- In the discrete case, the expectation of $g(X,Y)$ is

$$E[g(X,Y)] = \sum_j \sum_i g(x_i,y_j)p(x_i,y_j).$$

- In the continuous case, the expectation of $g(X,Y)$ is

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)\,dx\,dy.$$

$E[X] = \sum_{i=1}^{6} x_i P(x_i) = (1+\cdots+6)\cdot\frac{1}{6} = \frac{1}{6}\cdot\frac{7\times 6}{2} = \frac{7}{2}$

Example 1.15 The expected sum when three fair dice are rolled is:

$$E[X] = E[X_1 + X_2 + X_3] = E[X_1] + E[X_2] + E[X_3] = \frac{7}{2} + \frac{7}{2} + \frac{7}{2} = \frac{21}{2}.$$

Here, X_i is the value of the ith die and X is the total. □

Example 1.16 Let $g(X,Y) = X + Y$. Then, in continuous case,

Fubini's theorem.

$$\begin{aligned}
E_{X,Y}[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f(x,y)\, dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\, dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\, dxdy \\
&= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y)\, dy\right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x,y)\, dx\right] dy \\
&= \int_{-\infty}^{\infty} x f_X(x) dx + \int_{-\infty}^{\infty} y f_Y(y) dy \\
&= E_X[X] + E_Y[Y]. \quad \square
\end{aligned}$$

$E[\sum x_i] = \sum E[x_i]$ "linearity"

objective: to use "rolling two dice" to show

how to use $E[g(X,Y)] = \sum_j \sum_i g(x_i, y_j) p(x_i, y_j)$.

to find $E[X+Y] = E[X] + E[Y]$

$p(X=x, Y=y) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

$E(X+Y) = \sum_{j=1}^{6} \sum_{x=1}^{6} (x+y) \cdot p(X=x, Y=y)$

$= \frac{1}{36} \sum_{y=1}^{6} \sum_{x=1}^{6} (x+y)$

$= \frac{1}{36} \cdot \sum_{y=1}^{6} \cdot \left(6y + \frac{7 \times 6}{2}\right) = \frac{1}{36} \cdot \sum_{y=1}^{6} (6y+21) = 7.$

$E[Y] = E[X] = \sum_{x=1}^{6} x_i \, p(x_i) = \frac{1}{6} \times (1 + \cdots + 6) = \frac{21}{6} = \frac{7}{2}$

so $E[X+Y] = E[X] + E[Y]$

1.9.2 Independent Random Variables

- Two random variables X and Y are said to be *independent* if for all a, b,

$$P\{X \leq a, Y \leq b\} = P\{X \leq a\}P\{Y \leq b\}.$$

That is the events $\{X \leq a\}$ and $\{Y \leq b\}$ are independent. The above can also be written as $F(a, b) = F_X(a)F_Y(b)$.

- Independence can also be expressed as

$$\begin{aligned} p(x, y) &= p_X(x)p_Y(y) \qquad \text{(discrete case)} \\ f(x, y) &= f_X(x)f_Y(y) \qquad \text{(continuous case)} \end{aligned}$$

Proposition 1.17

If X and Y are independent, then

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

In particular, $E[XY] = E[X]E[Y]$.

independent $\Longrightarrow$ Cov $= 0$ (uncorrelated)

$\not\Longleftarrow$ e.g. $X \sim N(0, \sigma^2)$, $Y = X^2$

$\text{Cov}(X, Y) = \text{Cov}(X, X^2)$

$= E(X^3) - E[X] \cdot E[X^2]$

$= 0.$

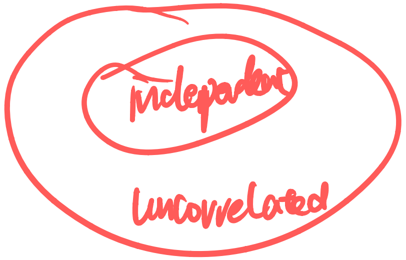
1.9.3 Covariance and Variance

- *Covariance* of X and Y:

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

$\text{Cov}(X, Y) = 0$, X, Y are uncorrelated

- If X and Y are independent, then $\text{Cov}(X,Y) = 0$.



- $\text{Cov}(X,X) = \text{Var}(X)$

- Covariance of sums:

$$\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \text{Cov}(X_i, Y_j).$$

- Covariance measures the joint variability of two R.V. indicating how they tend to change together.
- Sign - show the tendency.

1.10 Stochastic Processes

- A *stochastic process* $\{X_t |\ t \in T\}$ is simply a collection of random variables.
- The index t is often time.
- The X_t is the *state* at time t.
- T can be *discrete-time* $T = \{0, 1, 2, 3, \ldots\}$ or *continuous-time* $T = [0, \infty)$.
- E.g. X_n can be the closing price of a stock on day n for $n = 0, 1, 2, \ldots$
- E.g. X_t can be the coordinates of a car at time t for $t \geq 0$.

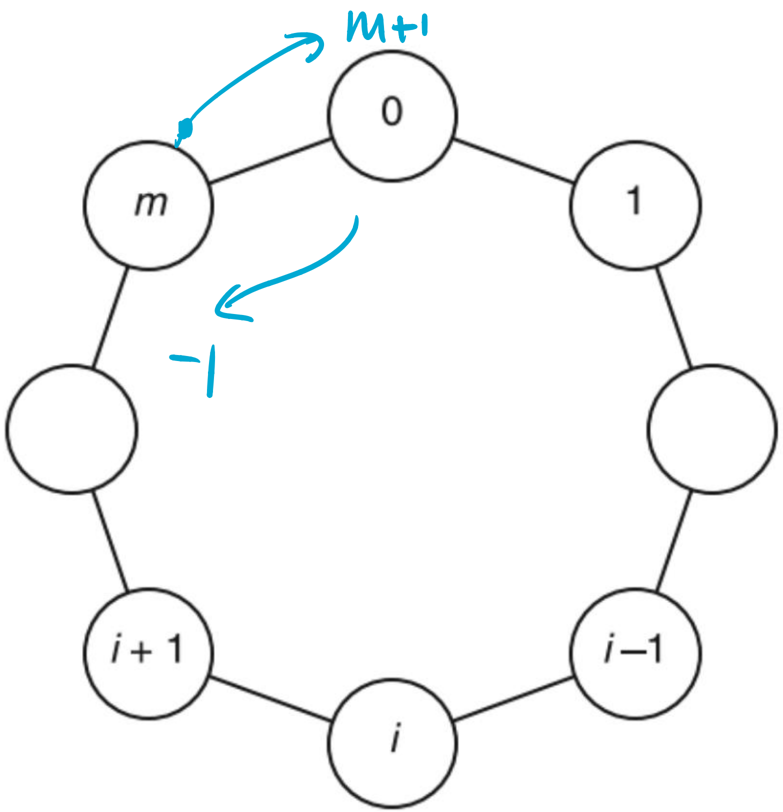
Example 1.18 Consider a particle that moves along a set of $m+1$ nodes, labeled $0, 1, \ldots, m$, that are arranged around a circle. At each step the particle moves one position in either clockwise or counterclockwise. Let X_n be the position of the particle after its nth step. Suppose that

$$P\{X_{n+1} = i+1 | X_n = i\} = P\{X_{n+1} = i-1 | X_n = i\} = \frac{1}{2}.$$

Here, $m+1$ is considered 0 and -1 is considered m. Suppose the particle starts at position 0, i.e. $X_0 = 0$. The process ends when all nodes have been visited at least once. What is the probability that node i, $i = 1, 2, \ldots, m$, is the last one visited? That is, all nodes except i have been visited at least once before i is visited.

Note that node 0 is always the first node visited.

Particle moving around a circle



Solution:

$P(E) = \sum_{i=1}^{m} P(E|F_i) P(F_i)$

$$\begin{aligned} P\{i \text{ is the last}\} &= P\{i \text{ is the last} \mid \text{visits } i-1 \text{ before } i+1\} \cdot P\{\text{visits } i-1 \text{ before } i+1\} \\ &\quad + P\{i \text{ is the last} \mid \text{visits } i+1 \text{ before } i-1\} \cdot P\{\text{visits } i+1 \text{ before } i-1\} \\ &= P\{i \text{ is the last} \mid \text{visits } i-1 \text{ before } i+1\} \cdot p_i \\ &\quad + P\{i \text{ is the last} \mid \text{visits } i+1 \text{ before } i-1\} \cdot (1-p_i). \end{aligned}$$

Suppose that $i-1$ is visited before $i+1$. Consider the first time that the particle is at $i-1$. Then,

$$P\{i \text{ is the last} \mid \text{visits } i-1 \text{ before } i+1\} = P\{m-1 \text{ steps backward before 1 step forward}\}.$$

Likewise, we have

$$\begin{aligned} P\{i \text{ is the last} \mid \text{visits } i+1 \text{ before } i-1\} &= P\{m-1 \text{ steps forward before 1 step backward}\} \\ &= P\{m-1 \text{ steps backward before 1 step forward}\}. \end{aligned}$$

Therefore, we have

$$P\{i \text{ is the last}\} = P\{m-1 \text{ steps backward before 1 step forward}\}.$$

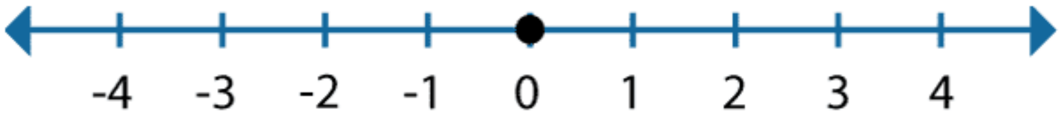
Since $P\{m-1 \text{ steps backward before 1 step forward}\}$ is independent of i, it must be equal to $1/m$. □

This example shows that $P\{m-1 \text{ up before 1 down}\} = \frac{1}{m}$.

$P\{k \text{ up before } m \text{ down}\} = \frac{m}{k+m}$

Random walk

- Random walk is a stochastic process that describes a path formed by a sequence of independent random steps in a mathematical space, e.g., on a number line.
- A simple symmetric random walk in one dimension takes steps of +1 (forward) or -1 (backward) with equal probability $\frac{1}{2}$.



- It's a type of **Markov process**, meaning the next position depends only on the **current** position, not on the path taken to get there.

In the particle moving along a circle example

- The objective is to find the probability that the walk, starting at 0, moves $m-1$ steps forward before moving 1 step backward.
- Let u_k represent the probability of reaching $m-1$ before -1 starting from state k, the desired probability is u_0.
- Boundary conditions:
  - $k = m - 1, u_{m-1} = 1$
  - $k = -1$, the walk has reached -1, so $u_{-1} = 0$

$P($m-1 forward before 1 backward starting from 0$) = \mu_0$

In the particle moving along a circle example

- Recurrence relation:

For any state k, where $-1 < k < m - 1$, the walk moves to $k+1$ or $k-1$ with the same probability $\frac{1}{2}$.

$k=0 \quad u_1 - 2u_0 + u_{-1} = 0$

$k=1 \quad u_2 - 2u_1 + u_0 = 0$

$u_1 = 2u_0$

$u_2 = 2u_1 - u_0 = 3u_0$

- The probability u_k satisfies

$$u_k = \frac{1}{2}u_{k+1} + \frac{1}{2}u_{k-1}$$

Rewrite this equation

$$u_{k+1} - 2u_k + u_{k-1} = 0$$

$$u_0 = \frac{1}{m}$$

Random walk

- A **symmetric** random walk is a martingale, so the same conclusion can be obtained using the martingale stopping theorem.

- A Brownian motion can be regarded as a continuous version of a symmetric random walk.

Example 1.19 Suppose that a gambler is equally likely to either win or lose \$1 on each gamble. What is the probability of winning \$$k$ before losing \$$n$? You may use the fact that $P\{\text{down } n \text{ before up } 1\} = \frac{1}{n+1}$.

Solution: First, note that

$$P\{\text{up 1 before down } n\} = 1 - \frac{1}{n+1} = \frac{n}{n+1}.$$

By conditioning,

$$\begin{aligned} P\{\text{up 2 before down } n\} &= P\{\text{up 2 before down } n|\text{up 1 before down } n\} \cdot \frac{n}{n+1} \\ &= P\{\text{up 1 before down } n+1\} \cdot \frac{n}{n+1} = \frac{n+1}{n+2} \cdot \frac{n}{n+1}. \end{aligned}$$

Repeating this argument, we obtain

$$\begin{aligned} &P\{\text{win } k \text{ before lose } n\} \\ &= P\{\text{win } k-1 \text{ before lose } n+1\} \cdot \frac{n}{n+1} = \cdots \\ &= P\{\text{win 1 before lose } n+k-1\} \cdot \frac{n+k-2}{n+k-1} \cdot \cdots \cdot \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{n}{n+k}. \end{aligned}$$

□

Example 1.19 Suppose that a gambler is equally likely to either win or lose \$1 on each gamble. What is the probability of winning \$$k$ before losing \$$n$? You may use the fact that $P\{\text{down } n \text{ before up } 1\} = \frac{1}{n+1}$.

Solution: First, note that

$$P\{\text{up 1 before down } n\} = 1 - \frac{1}{n+1} = \frac{n}{n+1}.$$

By conditioning,

$$\begin{aligned} P\{\text{up 2 before down } n\} &= P\{\text{up 2 before down } n | \text{up 1 before down } n\} \cdot \frac{n}{n+1} \\ &= P\{\text{up 1 before down } n+1\} \cdot \frac{n}{n+1} = \frac{n+1}{n+2} \cdot \frac{n}{n+1}. \end{aligned}$$

Repeating this argument, we obtain

$$\begin{aligned} & P\{\text{win } k \text{ before lose } n\} \\ = \; & P\{\text{win } k-1 \text{ before lose } n+1\} \cdot \frac{n}{n+1} = \cdots \\ = \; & P\{\text{win 1 before lose } n+k-1\} \cdot \frac{n+k-2}{n+k-1} \cdot \cdots \cdot \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{n}{n+k}. \end{aligned}$$

□

Example 1.19 Suppose that a gambler is equally likely to either win or lose \$1 on each gamble. What is the probability of winning $\$k$ before losing $\$n$? You may use the fact that $P\{\text{down } n \text{ before up } 1\} = \frac{1}{n+1}$.

Solution: First, note that

$$P\{\text{up 1 before down } n\} = 1 - \frac{1}{n+1} = \frac{n}{n+1}.$$

By conditioning,

$$\begin{aligned} P\{\text{up 2 before down } n\} &= P\{\text{up 2 before down } n | \text{up 1 before down } n\} \cdot \frac{n}{n+1} \\ &= P\{\text{up 1 before down } n+1\} \cdot \frac{n}{n+1} = \frac{n+1}{n+2} \cdot \frac{n}{n+1}. \end{aligned}$$

Repeating this argument, we obtain

$$\begin{aligned} &P\{\text{win } k \text{ before lose } n\} \\ = \;& P\{\text{win } k-1 \text{ before lose } n+1\} \cdot \frac{n}{n+1} = \cdots \\ = \;& P\{\text{win 1 before lose } n+k-1\} \cdot \frac{n+k-2}{n+k-1} \cdot \cdots \cdot \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{n}{n+k}. \end{aligned}$$

$\square$