Overview of Lecture

• Basic probability review
• Important distributions
• Poison Process
• Markov Chains
• Queuing Systems
Basic Probability concepts
Basic Probability concepts and terms

• Basic concepts by example: dice rolling
  – (Intuitively) coin tossing and dice rolling are *Stochastic Processes*
  – The outcome of a dice rolling is an *Event*
  – The values of the outcomes can be expressed as a *Random Variable*, which assumes a new value with each event.
  – The likelihood of an event (or instance) of a random variable is expressed as a real number in [0,1].
    • Dice: Random variable
    • instances or values: 1,2,3,4,5,6
    • Events: Dice=1, Dice=2, Dice=3, Dice=4, Dice=5, Dice=6
    • \( \text{Pr}\{\text{Dice}=1\} = 0.1666 \)
    • \( \text{Pr}\{\text{Dice}=2\} = 0.1666 \)
  – The outcome of the dice is independent of any previous rolls (given that the constitution of the dice is not altered)
Basic Probability Properties

• The sum of probabilities of all possible values of a random variable is exactly 1.
  \(- \Pr\{\text{Coin=HEADS}\} + \Pr\{\text{Coin=TAILS}\} = 1\)

• The probability of the union of two events of the same variable is the sum of the separate probabilities of the events
  \(- \Pr\{\text{Dice=1} \cup \text{Dice=2}\} = \Pr\{\text{Dice=1}\} + \Pr\{\text{Dice=2}\} = \frac{1}{3}\)
Properties of two or more random variables

• The tossing of two or more coins (such that each does not touch any of the other(s) ) simultaneously is called *(statistically) independent processes* (so are the related variables and events).

• The probability of the intersection of two or more independent events is the product of the separate probabilities:
  
  \[ \Pr\{ \text{Coin1=HEADS} \cap \text{Coin2=HEADS} \} = \Pr\{ \text{Coin1=HEADS} \} \cdot \Pr\{ \text{Coin2=HEADS} \} \]
Moments and expected values

• $m$th moment:

$$E[ X^m ] = \sum_{i} x_i^m \Pr \{ X = x_i \}$$

• Expected value is the 1st moment: $\mu_X = E[X]$

• $m$th central moment:

$$E[ (X - \mu_X)^m ] = \sum_{i} (x_i - \mu_X)^m \Pr \{ X = x_i \}$$

• 2nd central moment is called variance ($\text{Var}[X], \sigma_X$)
Conditional Probabilities

- The likelihood of an event can change if the knowledge and occurrence of another event exists.

- Notation: \( \text{Pr} \{ A \mid B \} = \frac{\text{Pr} \{ A \cap B \}}{\text{Pr} \{ B \}} \), \( \text{Pr} \{ B \} > 0 \)

- Usually we use conditional probabilities like this:

\[
\text{Pr} \{ A \} = \sum_i \text{Pr} \{ A \mid B_i \} \text{Pr} \{ B_i \}
\]
Conditional Probability Example

• Problem statement (Dice picking):
  – There are 3 identical sacks with colored dice. Sack A has 1 red and 4 green dice, sack B has 2 red and 3 green dice and sack C has 3 red and 2 green dice. We choose 1 sack randomly and pull a dice while blindfolded. What is the probability that we chose a red dice?

• Thinking:
  – If we pick sack A, there is 0.2 probability that we get a red dice
  – If we pick sack B, there is 0.4 probability that we get a red dice
  – If we pick sack C, there is 0.6 probability that we get a red dice
  – For each sack, there is 0.333 probability that we choose that sack

• Result ( \( R \) stands for “picking red”):
  \[
  \Pr\{R\} = \Pr\{R \mid A\} \Pr\{A\} + \Pr\{R \mid B\} \Pr\{B\} + \Pr\{R \mid C\} \Pr\{C\} = \\
  = 0.2 \cdot 0.333 + 0.4 \cdot 0.333 + 0.6 \cdot 0.333 = 0.4
  \]
Probability Distributions
Geometric Distribution

• Expresses the probability of number of trials needed to obtain the event $A$.
  - Example: what is the probability that we need $k$ dice rolls in order to obtain a 6?

• Formula: (for the dice example, $p=\frac{1}{6}$)

$$\Pr\{Y = k\} = p(1 - p)^{k-1}$$
Geometric distribution example

• A wireless network protocol uses a stop-and-wait transmission policy. Each packet has a probability $p_E$ of being corrupted or lost.

• What is the probability that the protocol will need 3 transmissions to send a packet successfully?
  – Solution: 3 transmissions is 2 failures and 1 success, therefore:
    \[
    \Pr \{Y = 3\} = (1 - p_E) p_E^2
    \]

• What is the average number of transmissions needed per packet?
  \[
  E[Y] = \sum_{i=1}^{\infty} i \Pr \{Y = i\} = \sum_{i=1}^{\infty} i (1 - p_E) p_E^{i-1} = (1 - p_E) \sum_{i=1}^{\infty} i p_E^{i-1} = (1 - p_E) \frac{1}{(1 - p_E)^2} = \frac{1}{1 - p_E}
  \]
Binomial Distribution

• Expresses the probability of some events occurring out of a larger set of events
  - Example: we roll the dice \( n \) times. What is the probability that we get \( k \) 6’s?

• Formula: (for the dice example, \( p = \frac{1}{6} \))

\[
\Pr\{Y = k\} = \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k}
\]
Binomial Distribution Example

• Every packet has $n$ bits. There is a probability $p_B$ that a bit gets corrupted.

• What is the probability that a packet has exactly 1 corrupted bit?
  \[ \Pr\{Y = 1\} = \frac{n!}{1!(n-1)!} p_B^1 (1 - p_B)^{n-1} = \frac{n!}{(n-1)!} p_B (1 - p_B)^{n-1} \]

• What is the probability that a packet is not corrupted?
  \[ \Pr\{Y = 0\} = \frac{n!}{0!(n)!} p_B^0 (1 - p_B)^n = (1 - p_B)^n \]

• What is the probability that a packet is corrupted?
  \[ p_E = 1 - \Pr\{Y = 0\} = 1 - (1 - p_B)^n \]
Solve this Problem

• We roll a regular dice and we toss a coin as many times as the roll indicates.
  – What is the probability that we get no tails?
  – What is the probability that we get 3 heads?

• Using the same process, we are asked to get 6 tails in total. If we don’t get 6 tails with the first roll, we roll again and repeat as needed. What is the probability that we need more than 1 rolls to get 6 tails?
Continuous Distributions

- Continuous Distributions have:
  - Probability Density function
    
    \[ f_X(x) : \mathbb{R} \to \mathbb{R}_+ \text{, such that } \int_{-\infty}^{\infty} f_X(x) \, dx = 1 \]

  - Distribution function:
    
    \[ \Pr \{ X \leq x \} = F_X(x) = \int_{-\infty}^{x} f(u) \, du \]
Continuous Distributions

• Normal Distribution: \( \varphi(x; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \)

• Uniform Distribution: \( f_U(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{elsewhere} \end{cases} \)

• Exponential Distribution: \( f_E(t; \lambda) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases} \)
Poisson Process

• It expresses the number of events in a period of time given the rate of occurrence.
• Formula

\[ P_m(t) = \frac{(\lambda t)^m e^{-\lambda t}}{m!} \]

• A Poisson process with parameter \( \lambda \) has expected value \( \lambda \) and variance \( \lambda \).
• The time intervals between two events of a Poisson process have an exponential distribution (with mean \( 1/\lambda \)).
Poisson dist. example problem

• An interrupt service unit takes $t_o$ sec to service an interrupt before it can accept a new one. Interrupt arrivals follow a Poisson process with an average of $\lambda$ interrupts/sec.

• What is the probability that an interrupt is lost?
  – An interrupt is lost if 2 or more interrupts arrive within a period of $t_o$ sec.

\[
\Pr \{ Y \geq 2 \}=1-\Pr \{ Y < 2 \}=1-P_0(t_o)-P_1(t_o)=
\]
\[
1- \frac{(\lambda t_o)^0 e^{-\lambda t_o}}{0!} - \frac{(\lambda t_o)^1 e^{-\lambda t_o}}{1!} = 1-(1+\lambda t_o)e^{-\lambda t_o}
\]
Markov Chains
The Markov Process

- Discrete time Markov Chain: is a process that changes states at discrete times (steps). Each change in the state is an event.
- The probability of the next state depends solely on the current state, and not on any of the past states:
  \[
  \Pr\{X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\} = \Pr\{X_{n+1} = j \mid X_n = i_n\}
  \]
Stationary Probabilities

• The transition probabilities after 2 steps are given by $P \times P = P^2$; transition probabilities after 3 steps by $P^3$, etc.

• Usually, $\lim_{n \to \infty} P^n$ exists, and shows the probabilities of being in a state after a really long period of time.

• Stationary probabilities are independent of the initial state
### Numeric Example

\[
\begin{pmatrix}
0 & 0 & 0.5 & 0.5 \\
0.5 & 0 & 0.5 & 0.5 \\
0.5 & 0.5 & 0 & 0 \\
0 & 0.5 & 0.5 & 0
\end{pmatrix},
\quad
\begin{pmatrix}
0 & 0.5 & 0 & 0.5 \\
2 & 0.5 & 0.5 & 0 \\
2 & 0.25 & 0.25 & 0.5 \\
2 & 0.375 & 0.3125 & 0.3125
\end{pmatrix}
\]

\[
P_1 = \begin{pmatrix} 1 & 0.5 & 0.5 \end{pmatrix},
\quad
P_2 = \begin{pmatrix} 1 & 0.25 & 0.5 & 0.25 \end{pmatrix},
\quad
P_4 = \begin{pmatrix} 1 & 0.3125 & 0.375 & 0.3125 \end{pmatrix}
\]

\[
\lim_{n \to \infty} P^n = \begin{pmatrix} 1 & 0.3333 & 0.3333 & 0.3333 \end{pmatrix}
\]

- No matter where you start, after a long period of time, the probability of being in a state will not depend on your initial state.
Example Problem

• A variable bit rate (VBR) audio stream follows a bursty traffic model described by the following two-state Markov chain:

  ![Markov Chain Diagram]

  - In the LOW state it has a bit rate of 50Kbps and in the HIGH it has a bit rate of 150Kbps.

• What is the average bit-rate of the stream?
Example Problem (cont’d)

• We find the stationary probabilities

\[
P = \begin{pmatrix} 0.99 & 0.01 \\ 0.1 & 0.9 \end{pmatrix}, \quad \lim_{n \to \infty} P^n \approx \begin{pmatrix} 0.91 & 0.09 \\ 0.91 & 0.09 \end{pmatrix}
\]

• Using the stationary probabilities we find the average bitrate:

\[
50 \cdot 0.91 + 150 \cdot 0.09 = 59 \text{ Kbps}
\]
Queuing Systems
Little’s Law:
Average number of tasks in system =

Average arrival rate \times Average response time
• What is the system performance?
  – average number of tasks in server
  – average number of tasks in queue
  – average turnaround time in whole system
  – average waiting time in the queue
Different Queuing Models

- Arrival process / service process / number of servers
- M/M/1
  - Markov / markov / 1 (poisson / poisson / 1 server)
- M/M/m
  - Markov / markov / 1 (poisson / poisson / m servers)
- M/G/1
  - Markov / general / 1
- G/M/1
- G/G/1
Results for M/M/1 Queue

- Definition: $\rho = \frac{\lambda}{\mu}$
- Average number in system = $\frac{\rho}{1-\rho}$
- Average number in queue = $\frac{\rho^2}{1-\rho}$
- Average waiting time in system = $\frac{1}{(\mu-\lambda)}$
- Average waiting time in queue = $\frac{\rho}{(\mu-\lambda)}$
Additional Slides
Transition Matrix

• Putting all the transition probabilities together, we get a transition probability matrix:
  
  - The sum of probabilities across each row has to be 1.

\[
\Pr\{X_{n+1} = j \mid X_n = i\} = P_{i,j}^{n,n+1} \rightarrow P_{n,n+1}^{n,n+1} = \begin{pmatrix}
0 & P_{00}^{n,n+1} & P_{01}^{n,n+1} & \cdots \\
1 & P_{10}^{n,n+1} & P_{11}^{n,n+1} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]
Finding the stationary probabilities

\[
\begin{bmatrix}
x_1 & x_2 & \cdots & x_n \\
x_1 & x_2 & \cdots & x_n \\
\vdots & \vdots & \ddots & \vdots \\
x_1 & x_2 & \cdots & x_n
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
= 
\begin{bmatrix}
x_1 & x_2 & \cdots & x_n \\
x_1 & x_2 & \cdots & x_n \\
\vdots & \vdots & \ddots & \vdots \\
x_1 & x_2 & \cdots & x_n
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} \cdot P \Rightarrow
\]

\[
\begin{bmatrix}
x_1 & x_2 & \cdots & x_n \\
x_1 & x_2 & \cdots & x_n \\
\vdots & \vdots & \ddots & \vdots \\
x_1 & x_2 & \cdots & x_n
\end{bmatrix}
\begin{bmatrix}
x_1 p_{1,1} + x_2 p_{2,1} + \cdots + x_n p_{n,1} \\
x_1 p_{1,1} + x_2 p_{2,1} + \cdots + x_n p_{n,1} \\
\vdots & \vdots & \ddots & \vdots \\
x_1 p_{1,1} + x_2 p_{2,1} + \cdots + x_n p_{n,1}
\end{bmatrix}
= 
\begin{bmatrix}
x_1 & x_2 & \cdots & x_n \\
x_1 & x_2 & \cdots & x_n \\
\vdots & \vdots & \ddots & \vdots \\
x_1 & x_2 & \cdots & x_n
\end{bmatrix}
\begin{bmatrix}
x_1 p_{1,1} + x_2 p_{2,1} + \cdots + x_n p_{n,1} \\
x_1 p_{1,1} + x_2 p_{2,1} + \cdots + x_n p_{n,1} \\
\vdots & \vdots & \ddots & \vdots \\
x_1 p_{1,1} + x_2 p_{2,1} + \cdots + x_n p_{n,1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
x_1 - x_1 p_{1,1} - x_2 p_{2,1} - \cdots - x_n p_{n,1} \\
x_2 - x_1 p_{1,2} - x_2 p_{2,2} - \cdots - x_n p_{n,2} \\
\vdots \\
x_n - x_1 p_{1,n} - x_2 p_{2,n} - \cdots - x_n p_{n,n}
\end{bmatrix}
= 0
\]
Markov Random Walk

- **Hitting probability (gambler’s ruin)** \( u_i = \text{Pr}\{X_T=0|X_0=i\} \)
  
  If \( \rho_k = \frac{q_1 q_2 \cdots q_k}{p_1 p_2 \cdots p_k} \) then gambler's ruin is

  \[
  u_i = \frac{\rho_i + \rho_{i+1} + \cdots + \rho_{N-1}}{1 + \rho_1 + \rho_2 + \cdots + \rho_{N-1}}
  \]

- **Mean hitting time (soujourn time)** \( v_i = \mathbb{E}[T|X_0=k] \)
  
  If \( \rho_k = \frac{q_1 q_2 \cdots q_k}{p_1 p_2 \cdots p_k} \) and \( \Phi_i = \rho_i \left( \frac{1}{q_1} + \frac{1}{q_2 \rho_1} + \cdots + \frac{1}{q_i \rho_{i-1}} \right) \)

  \[
  v_i = \frac{\Phi_1 + \Phi_2 + \cdots + \Phi_{N-1}}{1 + \rho_1 + \rho_2 + \cdots + \rho_{N-1}} \left( 1 + \rho_1 + \rho_2 + \cdots + \rho_{i-1} \right) - \left( \Phi_1 + \Phi_2 + \cdots + \Phi_{i-1} \right)
  \]
Example Problem

• A mouse walks equally likely in the following maze

```
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>D</td>
</tr>
</tbody>
</table>
```

• Once it reaches room D, it finds food and stays there indefinitely.
• What is the probability that it reaches room D within 5 steps, starting in A?
• Is there a probability it will never reach room D?
### Example Problem (cont’d)

\[ P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad P^2 = \begin{bmatrix} 0 & 0.75 & 0 & 0.25 \\ 0 & 0.75 & 0 & 0.25 \\ 0.25 & 0 & 0.25 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad P^3 = \begin{bmatrix} 0 & 0.375 & 0 & 0.375 \\ 0 & 0.375 & 0 & 0.375 \\ 0 & 0.375 & 0 & 0.625 \\ 0 & 0 & 0 & 1 \end{bmatrix} \]

\[ P^4 = \begin{bmatrix} 0 & 0.375 & 0 & 0.25 \\ 0 & 0.375 & 0 & 0.25 \\ 0 & 0.5625 & 0 & 0.4375 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad P^5 = \begin{bmatrix} 0 & 0.28125 & 0 & 0.4375 \\ 0 & 0.28125 & 0 & 0.4375 \\ 0.4375 & 0 & 0.71875 & 0.71875 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \lim_{n \to \infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \]
Poisson Distribution

• Poisson Distribution is the limit of Binomial when $n \rightarrow \infty$ and $p \rightarrow 0$, while the product $np$ is constant. In such a case, assessing or using the probability of a specific event makes little sense, yet it makes sense to use the “rate” of occurrence ($\lambda = np$).
  
  – Example: if the average number of falling stars observed every night is $\lambda$ then what is the probability that we observe $k$ stars fall in a night?

• Formula:

$$\Pr\{Y = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$$
Poisson Distribution Example

- In a bank branch, a rechargeable Bluetooth device sends a “ping” packet to a computer every time a customer enters the door.
- Customers arrive with Poisson distribution of \( \lambda \) customers per day.
- The Bluetooth device has a battery capacity of \( m \) Joules. Every packet takes \( n \) Joules, therefore the device can send \( u = \frac{m}{n} \) packets before it runs out of battery.
- Assuming that the device starts fully charged in the morning, what is the probability that it runs out of energy by the end of the day?

\[
\Pr \{ Y \geq u \} = 1 - \Pr \{ Y < u \} = 1 - \sum_{k=0}^{u-1} \frac{\lambda^k e^{-\lambda}}{k!}
\]
Markov Random Walk

• A Random Walk Is a subcase of Markov chains.

The transition probability matrix looks like this:

$$
p = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & \cdots & 0 \\
1 & p_1 & r_1 & q_1 & 0 & \cdots & 0 \\
2 & 0 & p_2 & r_2 & q_2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
N-1 & 0 & \cdots & 0 & p_{N-1} & r_{N-1} & q_{N-1} \\
N & 0 & \cdots & 0 & 0 & 0 & 1
\end{pmatrix}
$$