1. State and prove Little’s theorem.

2. **Poisson Arrivals See Time Averages (PASTA)**
   
   Show that for a queuing system, if the arrival process is Poisson and the inter-arrival times and service times are independent,
   
   \[
   \lim_{t \to \infty} P(N(t) = n \mid \text{an arrival occurred just after time } t) = \lim_{t \to \infty} P(N(t) = n)
   \]
   
   Hint: \(a_a(t) = \lim_{\delta \to 0} P[N(t) = n | A(t, t+\delta)] = \lim_{\delta \to 0} \frac{P[A(t, t+\delta)|N(t)=n]P[N(t)=n]}{P[A(t, t+\delta)]}, \) where, \(A(t, t+\delta)\) is the event that an arrival occurs in the interval \((t, t+\delta)\).

3. **Reversibility of birth-death process:**
   
   An ergodic birth-death process at steady state is time reversible.
   
   To prove the above result it needs to be shown that the rate at which a birth-death process goes from state \(i\) to \(i+1\) in steady state \(\pi(i)q_{i+1,i},\) is equal to the rate at which it goes from \(i+1\) to \(i\) in steady state \(\pi(i+1)q_{i+1,i},\) Note that there exists only a unique path from state \(i\) to \(i+1\) and vice-versa. Hence as the number of such transitions goes to infinity as \(t \to \infty,\) it follows that the rate of transitions from \(i\) to \(i+1\) equals the rate from \(i+1\) to \(i.\)

4. All the following queueing systems are birth-death processes and hence reversible. Solve the following using results from the previous question and truncated reversible processes.

   a. For an \(M/M/1\) queue specify the birth-death coefficients, the condition for ergodicity and show that the stationary distribution is of the form \(\pi_k = \rho^k(1-\rho)\) for \(\rho = \frac{\lambda}{\mu}.\)

   i. Let \(W\) be the time spent by a typical new arrival in the \(M/M/1\) queue before he begins his service. Show that the distribution of \(W\) for the \(M/M/1\) queue is given by

   \[
   F(W \leq x) = 1 - \rho e^{-\mu x}, x > 0, \text{ and } P(W = 0) = 1 - \rho.
   \]

   Hint: Let \(\Phi_W(x)\) be the m.g.f. of \(W.\) Use

   \[
   \Phi_W(s) = \mathbb{E}[e^{sW}] = \mathbb{E}[\mathbb{E}[e^{sW} | Q]] = \mathbb{E}\left[\mathbb{E}\left[e^{s(\sum_{k=1}^{Q-1} S_k + S_Q)}\right]\right]
   \]

   where, \(Q \sim \rho^n(1-\rho).\) \(S_k \sim \text{exp}(\mu), 1 \leq k \leq Q - 1,\) are the service times for the \(k^{th}\) customer and \(S_Q\) is the residual service time of \(Q^{th}\) customer. Note that \(S_Q \sim \text{exp}(\mu)\) (why?).
(b) Consider a queuing system where arrivals tend to get discouraged when more people are present in the queue. One way to model such a system is to choose \( \lambda_k = \frac{\alpha}{k+1}, \forall k \geq 0 \) and \( \mu_k = \mu, \forall k \geq 1 \). Show that the ergodic condition is \( \alpha/\mu < \infty \) and the stationary distribution is given by \( \pi_k \sim \text{Poisson} \left( \frac{\alpha}{\mu} \right) \).

(c) Show that for an \( M/M/\infty \) queue, the stationary distribution \( \pi_k \sim \text{Poisson} \left( \frac{\lambda}{\mu} \right) \) and the ergodic condition is \( \lambda/\mu < \infty \).

(d) i. Consider a system with unlimited waiting room and constant external arrival rate \( \lambda \). The system provides for a maximum of \( m \)-servers. Such systems are usually modelled by an \( M/M/m \) queue, which has \( \lambda_k = \lambda \) and \( \mu_k = \min\{k\mu, m\mu\} \). Prove that for an \( M/M/m \) queue, the stationary distribution is given by

\[
\pi_k = \begin{cases} 
\pi_0 \left( \frac{\lambda}{m\mu} \right)^k \frac{1}{k!}, & \text{for } k \leq m \\
\pi_0 \left( \frac{\lambda}{m\mu} \right)^m \frac{m^m}{m!}, & \text{for } k \geq m
\end{cases}
\]

where,

\[
\pi_0 = \frac{1}{\sum_{k=0}^{m-1} \frac{(km)^k}{k!} + \frac{(m)^m}{m!} \left( \frac{1}{1-\rho} \right)}
\]

for \( \rho = \frac{\lambda}{m\mu} \).

ii. Erlang’s C formula: Show that the probability no server is available to serve a customer for an \( M/M/m \) queuing system is given by

\[
\frac{(m\mu)^m}{m!} \left( \frac{1}{1-\rho} \right) \left/ \left( \sum_{k=0}^{m-1} \frac{(km)^k}{k!} + \frac{(m)^m}{m!} \left( \frac{1}{1-\rho} \right) \right) \right.
\]

(e) In practice the waiting room or buffer of the system has finite capacity. An example of such a finite capacity system is an \( M/M/1/K \) queue. Show that for an \( M/M/1/K \) queue

\[
\pi_k = \begin{cases} 
\pi_0 \left( \frac{\lambda}{\mu} \right)^k \frac{1-\rho}{1-(\frac{\lambda}{\mu})^{k+1}}, & \text{for } 0 \leq k \leq K \\
0, & \text{otherwise}
\end{cases}
\]

(f) The \( M/M/m/m \) (\( m \)-server loss system) is another example for finite capacity system.

i. Show that for this queue

\[
\pi_k = \begin{cases} 
\pi_0 \left( \frac{\lambda}{\mu} \right)^k \frac{1}{k!}, & \text{for } k \leq m \\
0, & \text{for } k > m
\end{cases}
\]

where,

\[
\pi_0 = \frac{1}{1 + \sum_{k=1}^{m} \frac{1}{k!} \left( \frac{\lambda}{\mu} \right)^k}
\]

ii. Erlang’s B formula: Prove that the fraction of time that all \( m \) servers are busy is given by

\[
\frac{\frac{1}{m!} \left( \frac{\lambda}{\mu} \right)^m}{1 + \sum_{k=1}^{\infty} \frac{1}{k!} \left( \frac{\lambda}{\mu} \right)^k}
\]
5. **Tandem queues**

Two queues have one server each, and all service times are independent and exponentially distributed, with parameter $\mu_i$ for queue $i$. Customers arrive at the first queue at the instants of a Poisson process of rate $\lambda < \min\{\mu_1, \mu_2\}$, and on completing service immediately enter the second queue. The queues are in equilibrium. Show that

(a) the output of the first queue is a Poisson process with rate $\lambda$, and that the departures before time $t$ are independent of the length of the queue at time $t$.

(b) the waiting times of a given customer in the two queues are not independent.

**Hint:** Consider an arriving customer when the queues are in equilibrium, and let $W_i$ be his waiting time (before service) in the $i^{th}$ queue. With $T$ the time of arrival, argue that $P\{W_1 = 0, W_2 = 0\} > P\{\text{arrival finds entire system empty}\} = P\{Q_1(T) = 0, Q_2(T) = 0\} = P\{Q_1(T) = 0\}P\{Q_2(T) = 0\}$, what is the relationship between $P\{Q_1 = 0\}$ and $P\{Q_2 = 0\}$ for this system?

6. **Pollaczek-Khintchine (PK) formula**

**M/G/1 Queue:** Consider a single-channel queue with Poisson arrivals at rate $\lambda$ and a general independent service with common distribution $G$. Let $S$ denote the service time distribution with $\mathbb{E}(S) = 1/\mu$ and $\rho = \lambda/\mu$.

**Embedding of discrete time Markov chain:**

Let

- $X_n = $ number of customers in the system left behind by departure of $n^{th}$ customer $C_n$.
- $Y_n = $ number of arrivals that occur during the service time $S_n$ of $C_n$.

Derive the PK-formula from the following steps

(a) Show that $X_{n+1} = X_n - 1 + \delta_n + Y_{n+1}$, where $\delta_n = 1$ if $X_n = 0$, $\delta_n = 0$ otherwise; and argue that $\{X_n\}$ is a Markov chain.

**Hint:** Consider the two cases $X_n = 0$ and $X_n \geq 1$ separately.

(b) Argue that $(Y_n|S_n) \sim \text{Poisson}(\lambda S_n)$.

(c) Show that $\mathbb{E}(Y_n) = \rho$, $\text{Var}(Y_n) = \rho + \lambda^2 \text{Var}(S_n)$.

**Hint:** $\text{Var}(Y_n) = \mathbb{E}[\text{Var}(Y_n|S_n)] + \text{Var}[\mathbb{E}(Y_n|S_n)]$

(d) Argue that at stationarity, $\text{Cov}(X_n, \delta_n) = -\pi_0 \mathbb{E}[X_n] = -(1-\rho)\mathbb{E}[X_n]$.

**Hint:** $\mathbb{E}[X_n] = \mathbb{E}[\delta_n]$ and $\mathbb{E}[X_{n+1}] = \mathbb{E}[X_n] - 1 + \mathbb{E}[\delta_n] + \mathbb{E}[Y_{n+1}]$ and at stationarity, $\mathbb{E}[X_{n+1}] = \mathbb{E}[X_n]$.

(e) Show that $\mathbb{E}[X_n] = \rho + \frac{\lambda^2 \mathbb{E}[S_n^2]}{2(1-\rho)}$

**Hint:** $\text{Var}(X_{n+1}) = \text{Var}(X_n) + \delta_n + \text{Var}(Y_{n+1})$ and at stationarity $\text{Var}(X_{n+1}) = \text{Var}(X_n)$.

Let customer $C_n$ wait for time $W_n$ in the system and leave $X_n$ customers behind. Under FIFO, the customers left behind by a departure is precisely those who arrived while the departing customer was in the system. Thus $X_n$ is the number of Poisson arrivals during $W_n$ and $(X_n|W_n) \sim (Y_n|S_n)$.
(f) Prove using above $E[X_n] = \lambda E[W_n]$, and conclude the PK-formula

\[
\begin{align*}
\text{Average waiting time, } E[W] &= \frac{1}{\mu} + \frac{\lambda E[S^2]}{2(1 - \rho)} \\
\text{Average delay, } E[D] &= E[W] - \frac{1}{\mu} = \frac{\lambda E[S^2]}{2(1 - \rho)}
\end{align*}
\]

(g) Consider a communication channel in which every data transmission takes one unit of time and each transmission is corrupted independently with probability $p$. Data packets arrive at the transmitter queue according to a Poisson process with rate $\lambda$. The transmitter sends the Head of the Line (HOL) packet and waits one time unit for an ACK signal from the receiver. If the ACK signal is received within one time unit, the transmitter sends the next packet, otherwise it retransmits the original packet. Assuming data packets are not delayed, and there is no maximum number of retransmissions, find average packet waiting time in the queue and in the system (i.e., waiting time + $E[S]$) using PK formula.

*Hint: Note that if there are $k$ retransmissions, a given packet has to wait for $1 + 2k$ time units. The probability that there is no ACK is $1 - p^2 = q$. Thus $P\{W = 1 + 2k\} = q^k(1 - q)$, $k \geq 0$. Then the first two moments of the service time are given by $E[S] = \sum_{k=0}^{\infty}(1 + 2k)(1 - q)q^k$ and $E[S^2] = \sum_{k=0}^{\infty}(1 + kn)^2(1 - q)q^k$. Use $\sum_{k=0}^{\infty}k^2q^k = \frac{q + q^2}{(1 - q)^3}$.

7. Consider the closed queueing network in the figure below. There are three customers who cycle between node 1 and node 2. Both nodes use FCFS service and have exponentially distributed i.i.d service times. The service times at one node are also independent of those at the other node and are independent of the customer being served. The server at node $i$ has mean service time $1/\mu_i$, $i = 1, 2$. Assume to be specific that $\mu_2 < \mu_1$.

(a) Draw a birth-death process transition diagram for the number of customers at node 1 and find the stationary probability of each state.

(b) Find the time-average rate at which customers leave node 1.

(c) Find the time-average rate at which a given customer cycles through the system. *Hint: Each third departure from node 1 is a departure of the same customer, and this corresponds to each third downward transition in the queue.*
The failure rate or hazard rate function:
Consider a continuous random variable $X$ having distribution function $F$ and corresponding density function $f$. The failure rate or hazard rate function is defined by

$$H(t) = \frac{f(t)}{F(t)}$$

To interpret $H(t)$, think of $X$ as the lifetime of a process. Suppose we want to find the probability the process will not survive an additional time $dt$, given the knowledge it has survived for $t$ time units. This probability can be evaluated as follows

$$P\{X \in (t, t+dt)|X > t\} = \frac{P\{X \in (t, t+dt), X > t\}}{P\{X > t\}}$$

$$= \frac{P\{X \in (t, t+dt)\}}{P\{X > t\}}$$

$$\approx \frac{f(t)dt}{F(t)} = H(t)dt$$

Thus $H(t)$ represents the probability density that a $t$ time unit old process will fail.

8. $M/G/1$ processor sharing queue

Suppose that customers arrive in accordance with a Poisson process having rate $\lambda$ and each customer requires a random amount of work distributed according to $G$, with corresponding density function $g$. The server can process work at the rate of one unit of work per unit time, and divides his time equally among all the customers present in the system. Thus, whenever there are $n$ customers in the system, each will receive service work at a rate of $1/n$ per unit time.

Let $H(t) = g(t)/G(t)$ denote the hazard rate function of the service distribution, and suppose that $\lambda E[S] < 1$, where $E[S]$ is the mean of $G$. To analyse the system, let the state of the system at any time be the ordered vector of the amounts of work already performed on customers still in the system. That is, if there are $n$ customers present in the system, the state is $\pi = (x_1, x_2, \ldots, x_n)$, $x_1 \leq x_2 \leq \cdots \leq x_n$, where $x_1, x_2, \ldots, x_n$ is the amount of work performed on these $n$ customers.

Let $p(\pi) = \prod_{i=1}^{n} p_i(x_i)$ denote the limiting probability density of the state $\pi$ and the limiting probability that the system is empty.

Let $e_i(\pi) = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$

(a) Argue that the forward transitions $\pi \rightarrow e_i(\pi)$ occurs at rate $H(x_i)/n$, and the reverse transitions $e_i(\pi) \rightarrow \pi$ occurs at rate $\lambda g(x_i)$.

Hint: Note that the transition $e_i(\pi) \rightarrow \pi$ corresponds to a new arrival to the system requiring work $x_i$. Since work required is independent of arrival process, the result follows from Bernoulli splitting of Poisson process. The transition $\pi \rightarrow e_i(\pi)$ corresponds to a customer who has already completed $x_i$ units of work leaving the system.

(b) Show that the reverse process is a system of the same type, with customers arriving at a Poisson rate $\lambda$, having workloads distributed according to $G$ and with the state representing the ordered residual workloads of customers presently in the system.

Hint: You need to check the validity of the equations. $p(\pi)q_{\pi,e_i(\pi)} = p(e_i(\pi))q_{e_i(\pi),\pi}$

(c) Using above, show that

i. $P\{n \text{ customers in the system}\} = (\lambda E[S])^n (1 - \lambda E[S]), n \geq 0.$
ii. the conditional distribution of the ordered amounts of work already performed, given \( n \) in the system, 
\[
p(\bar{x} | n) = n! \prod_{i=1}^{n} \frac{G(x_i)}{G(M)}.
\]